



HAL
open science

Monitoring en temps réel de la vitesse de déplacement sur dispositif connecté : modélisation mathématique sur plateforme mobile interfacée avec une base de données d'entraînement et d'audit physiologique

Thomas Carbillet

► To cite this version:

Thomas Carbillet. Monitoring en temps réel de la vitesse de déplacement sur dispositif connecté : modélisation mathématique sur plateforme mobile interfacée avec une base de données d'entraînement et d'audit physiologique. Traitement du signal et de l'image [eess.SP]. Université Grenoble Alpes, 2019. Français. NNT : 2019GREAM013 . tel-02269138

HAL Id: tel-02269138

<https://theses.hal.science/tel-02269138v1>

Submitted on 22 Aug 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



THÈSE

Pour obtenir le grade de

DOCTEUR DE LA COMMUNAUTÉ UNIVERSITÉ GRENOBLE ALPES

Spécialité : Mathématiques Appliquées

Arrêté ministériel : 25 mai 2016

Présentée par

Thomas CARBILLET

Thèse dirigée par **Stéphane LABBE**, Professeur , Université Grenoble Alpes

préparée au sein du **Laboratoire Jean Kuntzmann** dans **l'École Doctorale Mathématiques, Sciences et technologies de l'information, Informatique**

Monitoring en temps réel de la vitesse de déplacement sur dispositif connecté : modélisation mathématique sur plateforme mobile interfacée avec une base de données d'entraînement et d'audit physiologique

Thèse soutenue publiquement le **12 avril 2019** devant le jury composé de :

Monsieur STEPHANE LABBE

PROFESSEUR, UNIVERSITE GRENOBLE ALPES, Directeur de thèse

Madame SOPHIE DONNET

CHARGE DE RECHERCHE, INRA-AGRO-PARISTECH, Rapporteur

Monsieur ARNAUD FERRY

PROFESSEUR, UNIVERSITE PARIS 5, Rapporteur

Madame VERONIQUE BILLAT

PROFESSEUR, UNIVERSITE D'EVRY-VAL D'ESSONNE, Examineur

Madame ADELIN LECLERCQ SAMSON

PROFESSEUR, UNIVERSITE GRENOBLE ALPES, Président

Monsieur FRANCK CORSET

MAITRE DE CONFERENCES, UNIVERSITE GRENOBLE ALPES, Examineur

Madame CHARLOTTE DION

MAITRE DE CONFERENCES, SORBONNE UNIVERSITES - PARIS, Examineur

Monsieur JEAN-RENAUD PYCKE

MAITRE DE CONFERENCES, UNIVERSITE D'EVRY-VAL D'ESSONNE, Examineur

Remerciements

Quatre belles années pour avoir un aperçu du monde de la recherche et puis s'en va. Je tire ma révérence à ce monde dans lequel je me suis toujours senti en marge malgré une fascination certaine. Nous formions, avec le Laboratoire Jean Kunzmann et l'entreprise BillaTraining, un trio éphémère avec ces réussites et ces déboires. Je suis heureux d'avoir pu vivre cette aventure qui a su me faire grandir professionnellement et humainement. Je tiens à remercier toutes les personnes avec qui j'ai pu échanger, même brièvement, durant cette thèse.

Plus particulièrement, je tiens à remercier Stéphane pour son encadrement, son soutien tout particulièrement pendant les derniers mois de thèse et d'avoir su comprendre certains de mes choix personnels. Je remercie également Véronique pour l'opportunité. Les moments partagés auront été une grande source d'inspiration pour moi malgré nos chemins qui se séparent. Je tiens à remercier Jean-Renaud pour ces discours mathématiques et philosophiques, un vrai livre ouvert. Je remercie également Nicolas qui a su me débloquer au milieu de cette thèse avec qui nous avons commencé de belles choses. J'aimerais aussi remercier Adeline pour sa collaboration et ses conseils, j'ai beaucoup aimé cette période lors de ma thèse. Je remercie Charlotte pour s'être déplacée à Grenoble et avoir pris le temps de me donner plusieurs pistes de recherche.

Je remercie chaque membre du Jury pour le temps consacré lors de la lecture du manuscrit et de la soutenance de thèse. Merci à Stéphane Labbé, Sophie Donnet, Arnaud Ferry, Véronique Billat, Adeline Leclercq Samson, Franck Corset, Charlotte Dion et Jean-Renaud Pycke.

Sur un autre registre, je remercie énormément Claire, Florent et Luc. Nous avons pu travailler dans une ambiance de travail remarquable. Les bières sont très bonnes. Jean Pierre, un grand monsieur à qui je dois un grand merci, toujours à la retombée du ballon. Je lui dois

la relecture pointilleuse de ma thèse parmi beaucoup d'autres choses dont la liste exhaustive prendrait des pages entières.

Je tiens à remercier spécialement mes parents, Christine et Jean-Luc, et ma sœur, Mélanie pour absolument tout. Ce remerciement sort largement du cadre de cette thèse, car ils savent être là au bon moment et au bon endroit en toute circonstance. Une belle définition de la famille qui me fait réaliser la chance que j'ai de les avoir tous les trois. Un autre remerciement spécial va à celle qui partage ma vie depuis plus de sept ans maintenant, Charline. Tu as su m'écouter, me résonner et me conseiller pendant cette aventure. Nous avons vécu tellement de belles choses pendant ces quatre années, un grand merci. Je remercie également Carole, Gilles, Léa, Clément et la petite dernière Ovsanna pour avoir su me sortir de mon quotidien. J'allais terminer les remerciements avec le paragraphe précédent, mais comme il me reste un peu de place... Je ne sais pas s'ils le méritent, mais je remercie quand même mes six frères : Colas, Brice, Kévin, Antoine, Rémy et Thomas ainsi que leur compagne. J'espère que vous finirez du bon côté de la vitre teintée (sauf Rémy).

Finalement, j'aimerais me remercier moi pour tout le travail fourni pendant la thèse. J'aimerais me remercier moi pour ne pas avoir abandonné. J'aimerais me remercier moi pour avoir été moi-même pendant ces quatre belles années.

Résumé

L'amélioration de la performance en course à pied, lors de courses sur du plat, raisonne avec l'actualité et l'objectif de voir un athlète descendre en dessous des deux heures sur Marathon. Cependant, il existe peu d'équipes travaillant de façon transverse sur les sujets de préparation et de stratégie de course à destination du grand public. Les programmes d'entraînement, fondés sur l'expérience des entraîneurs, ne sont pas, ou peu, personnalisés ce qui conduit à des risques de blessures et une perte de la motivation. Une limite de l'analyse actuelle des programmes d'entraînement semble avoir été atteinte et l'entreprise *BillaTraining*[®] a pour objectif de construire un pont entre la recherche scientifique et les coureurs amateurs.

L'objectif principal de ce travail de thèse est double. Premièrement, nous avons essayé d'apporter une contribution à la recherche dans le domaine de la course à pied. Après avoir accumulé et formaté des données d'entraînements et de courses provenant de différentes sources, nous avons cherché à décrire des phénomènes tels que l'accélération humaine ou encore les différentes stratégies de course employées sur Marathon pour des coureurs dont les performances de durée de course sont comprises entre 2h30 et 4 heures.

Deuxièmement, nous avons développé une application web intégrant les trois phases de la méthode *BillaTraining*[®]. La première étape est un audit énergétique qui n'est autre qu'un entraînement de trente minutes à la sensation, permettant de connaître les capacités physiques d'un coureur. La seconde étape, le radar énergétique, est le bilan de l'audit et agit comme point d'entrée de l'entraînement sur-mesure qui est la troisième et dernière étape.

Afin de répondre à ces deux objectifs, nous avons intégré des notions de physiologie, de mathématiques et d'informatique.

Les connaissances en physiologie sont basées sur l'état de l'art ainsi que les recherches passées et présentes du professeur Véronique Billat. Par extension, il s'agit du coeur de métier de l'entreprise *Billa Training*[®]. Les idées de recherche présentent dans cette thèse émanent de la physiologie.

Les mathématiques, quand à elles, nous ont permis de décrire certains phénomènes physiologiques grâce notamment aux statistiques. Nous avons eu l'occasion d'utiliser le modèle d'Ornstein-Uhlenbeck de retour à la moyenne ou encore la méthode PELT (Pruned Exact Linear Time) pour la détection d'instant de ruptures dans une série temporelle.

Finalement, l'informatique permet de faire communiquer les mathématiques et la physiologie à des fins de recherche scientifique et commerciales.

Mots clés

Course à pied ; Entraînement ; Détection de ruptures ; Application Web ; Temps réel

Abstract

The improvement running performance has become a major topic lately. We are getting closer to running a marathon in under 2 hours. However, there are not so many professionals working transversally regarding pre-race and in-race preparation concerning the general public. Training plans are based on trainers' experience and are often not custom-made. This exposes the runners to injury risk and motivation loss. It seems that the current analysis of training plans has reached a limit. The aim for *BillaTraining*[®] is to go beyond this limit by connecting the research with the general public of runners.

This PhD has two main goals. The first one is trying to contribute to the research about running. After gathering and formatting trainings and races data from different origins, we tried to isolate and describe how humans run marathons including 2.5 to 4-hour performances. We studied acceleration, speed and heart rate time series among other things, with the idea of understanding the different running strategies.

The second one is the development of a web application embracing the three steps of the *BillaTraining*[®] method. The first step is an energetic audit which is a 30-minute running session guided by the runner's sensations. The second step is the energetic radar which is the results of the audit. The last step is a tailor-made training plan built depending on the runner's objectives.

In order to come up with a solution, we had to bring together Physiology, Mathematics and Computer Science.

The knowledge we had in Physiology was based on professor Véronique Billat's past and current researches. These researches are now part of *BillaTraining*[®] and are central for the

growth of the company.

We used Mathematics to try to describe physiological phenomenons thanks to Statistics. By applying the Ornstein-Uhlenbeck model, we found that humans are able to run at an even acceleration. By using the PELT (Pruned Exact Linear Time) method we automated changepoints detection in time series.

Finally, Computer Science allowed a communication between Physiology and Mathematics for research, as well as marketing training tools at the forefront of innovation.

Keywords

Running ; Training plan ; Changepoints ; Web application ; Real-time

Table des matières

| | |
|--|-----------|
| Introduction | 13 |
| I Contexte physiologique et d'entreprise | 19 |
| 1 Problématiques et enjeux | 20 |
| 2 Etat de l'art et positionnement de la thèse | 21 |
| 2.1 Joseph B. Keller (1973) | 23 |
| 2.2 Frank Mathis (1989) | 26 |
| 2.3 William Woodside (1991) | 28 |
| 2.4 Horst Behncke (1993) | 30 |
| 2.5 Daniel Whitt (2009) | 32 |
| 2.5.1 Le modèle probabiliste | 32 |
| 2.5.2 Le processus stochastique | 34 |
| 2.6 Amandine Aftalion et Frédéric Bonnans (2013) | 35 |
| 2.7 Positionnement de la thèse | 39 |
| 3 L'entreprise <i>BillaTraining</i> [®] et ses méthodes | 41 |
| 3.1 L'entreprise en bref | 42 |
| 3.2 Le test de RabiT [®] | 44 |
| 3.3 Le Radar énergétique | 46 |
| 3.4 L'entraînement sur mesure | 48 |
| 3.5 La Signature de Vitesse [®] | 49 |
| 4 Etat des lieux de la concurrence | 50 |
| 5 Le matériel à disposition | 52 |

| | | |
|--|--|-----------|
| 5.1 | <i>Cosmed</i> [®] - K5 - Analyseur de gaz | 52 |
| 5.2 | <i>Garmin</i> [®] - 920XT - Montre de sport | 54 |
| 5.3 | <i>Physioflow</i> [®] - PF07 Enduro - Débit cardiaque | 55 |
| 5.4 | <i>Centaure Metrix</i> [®] - Locometrix - Accéléromètre triaxial | 56 |
| | Bilan | 56 |
| II Première contribution : La segmentation du test de RabiT[®] | | 58 |
| 1 | Problématiques et enjeux | 59 |
| 2 | La détection de ruptures | 61 |
| 2.1 | Le problème théorique de détection de ruptures | 62 |
| 2.2 | Le choix de la méthode de détection de ruptures | 64 |
| 3 | Application au test de RabiT [®] | 69 |
| 3.1 | Construction des solutions | 70 |
| 3.2 | Sélection et filtrage de la meilleure solution | 71 |
| 4 | Création d'un micro-service web | 74 |
| 4.1 | Architectures monolithique VS. micro service | 74 |
| 4.2 | Application Programming Interface (API) | 77 |
| 4.3 | Création d'un micro service web pour segmenter le test de RabiT [®] | 79 |
| | Bilan | 81 |
| III Deuxième contribution : Vers une nouvelle stratégie de course sur mara- | | |
| | thon | 85 |
| 1 | Problématique et enjeux | 86 |
| 2 | [Article] Detecting the marathon asymmetry with a statistical signature | 87 |
| 2.1 | Introduction | 88 |
| 2.2 | Methods | 90 |
| 2.2.1 | Methodology | 90 |
| 2.2.2 | Statistical study | 90 |
| 2.3 | Results | 93 |

| | | |
|-------|---|-----|
| 2.3.1 | Speed decreasing trend | 93 |
| 2.3.2 | Speed strong asymmetry behavior | 93 |
| 2.3.3 | Speed statistical signature | 97 |
| 2.4 | Discussion | 98 |
| 2.5 | Conclusion | 101 |
| 3 | [Article En Soumission] The 2018 marathon world record shows the positive asymmetry pacing strategy effectiveness | 101 |
| 3.1 | Introduction | 102 |
| 3.2 | Materials and Methods | 104 |
| 3.2.1 | Methodology | 104 |
| 3.2.2 | Statistical study | 104 |
| 3.2.3 | Critical Speed | 106 |
| 3.3 | Results | 107 |
| 3.3.1 | Statistical study of the whole marathon | 107 |
| 3.3.2 | Statistical study of the marathon split in two | 108 |
| 3.3.3 | Statistical study of the marathon split in four | 109 |
| 3.3.4 | Critical Speed | 111 |
| 3.4 | Discussion | 112 |
| 3.5 | Conclusion | 113 |
| | Bilan | 114 |

| | | |
|-----------|--|------------|
| IV | Troisième contribution : Caractérisation de l'accélération d'un coureur à pied | 116 |
| 1 | Problématique et enjeux | 117 |
| 2 | [Article] Humans are able to self-paced constant running acceleration until exhaustion | 118 |
| 2.1 | Abstract | 118 |
| 2.2 | Introduction | 119 |
| 2.3 | Methods | 120 |

| | | |
|----------|---|------------|
| 2.3.1 | Subjects | 120 |
| 2.3.2 | Experimental Design | 121 |
| 2.3.3 | Acceleration trials | 121 |
| 2.3.4 | Mathematical Model | 122 |
| 2.4 | Results | 124 |
| 2.4.1 | Comparison of the three different acceleration exercises | 124 |
| 2.4.2 | Characterizing the human faculty to maintain a constant acceleration | 126 |
| 2.5 | Discussion | 132 |
| 2.5.1 | The self-paced acceleration range values compared with current use of acceleration protocols in laboratory | 133 |
| 2.5.2 | The possible mechanisms allowing humans to control its body acceleration in accordance with prior results obtained on manual task | 134 |
| 2.6 | Conclusion | 134 |
| | Bilan | 136 |
| V | Quatrième contribution : Vers un produit fini | 137 |
| 1 | Problématiques et enjeux | 138 |
| 2 | Création d'une application web d'entraînement en course à pied | 139 |
| 2.1 | Architecture de l'application web | 140 |
| 2.2 | Les fonctionnalités de l'application web | 141 |
| 2.2.1 | La logique métier | 142 |
| 2.2.2 | Le web UI/UX | 146 |
| 2.2.3 | Les analyses scientifiques | 156 |
| 2.2.4 | Le paiement en ligne | 158 |
| 3 | Prototype d'une application mobile temps réel | 160 |
| 3.1 | Architecture de l'application mobile | 161 |
| 3.2 | Les fonctionnalités de l'application mobile | 163 |

| | | |
|-------|--|------------|
| 3.2.1 | Le temps réel | 163 |
| 3.2.2 | Le mobile UI/UX | 165 |
| 3.2.3 | Le web UI/UX 2 | 174 |
| | Bilan | 175 |
| | Conclusion | 177 |
| | A Comparison of the three acceleration intensity instructions | 182 |
| | B Measured data VS. Ornstein-Uhlenbeck model | 185 |
| | C Comparison of the three estimated mean accelerations | 188 |
| | Liste des tableaux | 191 |
| | Table des figures | 194 |
| | Bibliographie | 201 |

Introduction

Contexte social et économique

L'être humain est capable de courir depuis la nuit de temps. C'est en apprenant à chasser afin de se nourrir qu'il a développé cette compétence [39]. Souvent bien moins rapide au sprint que la proie qu'il convoite, l'être humain a adopté une allure modérée, obligeant le gibier à répéter son effort jusqu'à son épuisement. La différence majeure entre l'animal et l'être humain est la manière de réguler la température corporelle. Après chaque sprint, la proie doit marquer un instant d'arrêt pour éliminer la chaleur alors que l'être humain évacue le surplus de chaleur grâce à la transpiration par la peau.

De nos jours, courir pour chasser n'est plus nécessaire, cependant, la course à pied a d'autres bénéfices sur le corps humain. Une série d'études a montré les bienfaits sur l'homme d'une pratique régulière de la course à pied sur la santé mentale et la gestion des émotions[41, 73, 81], sur la réduction des maladies cardio-vasculaires[112, 70, 176], le diabète[175] et bien d'autres maladies. Certains Etats songent à adopter une stratégie préventive face à ces maladies en donnant la possibilité aux médecins de prescrire une pratique sportive sur ordonnance.

De par son histoire et ses bienfaits, la course à pied s'est imposée comme l'un des sports les plus pratiqués sur terre. L'idée de courir est simple et il y a peu de barrières à l'entrée ; une tenue et une paire de chaussure adéquates suffisent. La course à pied a même été qualifiée de sport le plus accessible au monde[157]. Il est cependant difficile d'obtenir des indicateurs récents et fiables sur le nombre de coureur à pied dans le monde. Rassemblant les données de

11 pays européens parmi les plus développés entre 1960 et 2013, le professeur Jeroen Scheerder note deux tournant dans l'adoption de la course à pied. Le premier tournant date des années 1960-1970 [151] où s'opère une révolution culturelle et une volonté de la population de se maintenir en forme. Le second tournant date des années 1990 et l'adoption de la course à pied par les femmes et les séniors. Un troisième tournant pourrait se dessiner avec l'apparition des courses à thèmes (courses en couleur, courses déguisées, ...). Le nombre de coureurs à pied dans l'Europe des 28 est estimé à 49,9 millions de personnes soit environ 10% de la population totale. Le chiffre des dépenses annuelles liées à la course à pied représente 9,6 milliards d'euros soit un panier moyen d'environ 190 euros par coureur et par an [151].

Profitant de l'augmentation du nombre de coureurs à pied, le marathon est devenu une épreuve phare. Le nombre de personnes franchissant la ligne d'arrivée d'un marathon a rapidement augmenté entre 2000 et 2013 passant de 400 000 à 1,6 millions dans le monde[151]. Il en va de même pour les athlètes de haut niveau qui convoitent le marathon et cherche à repousser les limites humaines. Descendre en dessous des 2 heures sur un marathon est devenu un phénomène sportif, scientifique et de marketing. Grâce a une équipe multi-disciplinaire, Eliud Kipchoge s'est rapproché de cette prouesse en battant le record du monde d'une minute et 18 secondes pour un temps final de 2 heures, 1 minute et 39 secondes.

Malgré leur enthousiasme, les coureurs amateurs préparant un marathon, bien trop souvent, se blessent et perdent en motivation à cause d'un programme d'entraînement inefficace[38, 46]. Les causes principales sont des séances d'entraînement trop longues, un nombre de séances hebdomadaires trop élevé et des séances d'entraînement trop peu variées. Plus particulièrement, nous décrierons dans le [chapitre I](#) une philosophie d'entraînement basé sur le plaisir, la performance et l'écoute de soi.

En aval de la phase d'entraînement se pose la question de la stratégie de course. C'est un point capital souvent négligé par les coureurs à pied. Dans l'imaginaire collectif, on retrouve des stratégies de courses basiques comme la vitesse constante ou le *negative split*. Nous savons pourtant qu'une allure de course variée permet d'augmenter ses performances[21, 17, 18, 20].

De même, le [chapitre III](#) propose des pistes de réflexion sur la manière d’aborder un marathon. Une des raisons ayant poussé le professeur Véronique Billat à fonder la société *BillaTraining*[®] est la construction d’un pont entre la recherche et le grand public. Le but est de pouvoir appliquer les dernières méthodes scientifiques d’entraînement et de stratégie de course.

Contexte technologique

Près de 50% des coureurs assidus (plus de deux courses par semaine) possèdent un objet connecté [99] enregistrant tout type de données telles que la vitesse, la fréquence cardiaque, le temps d’appui au sol, etc...

Dans ce marché en plein essor, on distingue deux groupes, chacun ciblant un segment de coureurs bien différent.

Le premier groupe, celui des téléphones portables et des applications mobiles, se concentre sur des coureurs débutants ou occasionnels. L’expérience utilisateur du coureur est soignée et donne un historique des séances d’entraînement passées. Certaines applications proposent des programmes d’entraînement, rarement personnalisés. La limite des téléphones portables est la précision des données enregistrées. Les applications présentes sur le marché utilisent uniquement le capteur GPS du téléphone. La fusion de capteurs permettrait de palier cette imprécision mais consommerait plus d’énergie en contrepartie. L’avantage est la possibilité de se connecter en temps réel avec le coureur afin de pouvoir le guider pendant son entraînement et le jour de sa course.

Le second groupe, celui des montres de sport, cible les coureurs confirmés en cherchant à développer des objets connectés permettant d’enregistrer des signaux de grandes qualités. Le coureur a accès à des données précises et extrêmement variées. La limite ici est l’impossibilité d’utiliser les techniques de temps réel puisque les dispositifs ne sont pas connectés à internet mais se synchronisent lorsqu’un réseau wifi pré-enregistré est détecté.

Cette révolution des objets connectés a permis l’accumulation d’un nombre incalculable de données. Nous pensons qu’il est temps de donner un sens à ces données en les analysant pour

proposer des entraînements sur-mesure et des stratégies de course personnalisées. Idéalement, nous souhaitons développer un service se situant à l'intersection des téléphones portables et des montres de sport. En d'autre terme, il est possible d'utiliser la fonctionnalité *temps réel* des téléphones portables avec la précision des montres de sport.

Objectifs

Ce travail de thèse s'inscrit dans le but de développer des liens directs entre la recherche et le grand public afin qu'il puisse bénéficier des dernières avancées scientifiques dans le domaine de la course à pied. Nous souhaitons proposer des alternatives aux méthodes d'entraînement et de stratégie de course existantes en proposant aux coureurs les clés qui leurs permettront de mieux se connaître. Nous croyons que la science est un outil permettant de décrire puis de prédire le monde dans lequel nous vivons. Nous avons toujours cherché à à faire parti d'équipes interdisciplinaires afin de répondre aux deux objectifs suivants :

- Automatiser totalement la méthode *BillaTraining*[®]. L'objectif ici est de construire les moyens par lesquels nous pourrions toucher le plus grand nombre de personnes. Nous souhaitons proposer un service abordable, fiable et opérationnel pour le coureur à pied afin qu'il découvre de nouvelles méthodes d'entraînement sur-mesure. Grâce aux mathématiques et à l'informatique nous avons construit une application web automatisée intégrant la méthode *BillaTraining*[®].
- Avancer dans notre réflexion sur la stratégie de course optimale sur marathon. En partant des données du terrain issues des objets connectés, nous avons décrit la façon dont les sportifs courent un marathon. Nous souhaitons également donner des perspectives de recherche afin de construire un modèle prédictif, personnalisé de stratégie de course.

Organisation du mémoire

Ce mémoire de thèse est constitué de cinq chapitres, organisés comme suit.

Le [chapitre I](#) présente deux parties. La première partie est consacrée à l'état de l'art des différentes stratégies de course à pied au moment de l'écriture de cette thèse. Des premières stratégies à vitesse constante des années 1970 aux stratégies à vitesses variables, nous détaillerons les travaux de nos prédécesseurs. La seconde partie, quand à elle, développe les quatre étapes de la méthode *BillaTraining*[®] (Audit, radar énergétique, entraînement personnalisé, stratégie de course). Ce chapitre permet de comprendre les concepts clés de physiologie que vous pourrez retrouver dans la suite de cette thèse. Aussi, à la fin de ce chapitre, nous avons détaillé précisément les objectifs de ces trois ans de recherche.

Le [chapitre II](#) se concentre sur la segmentation automatique d'un signal de vitesse en fonction du temps enregistré par le biais de montres de sport. L'objectif est de récupérer les paramètres intéressants du signal, ce qui constitue la première étape en vue de la création d'un entraînement sur-mesure et de l'établissement d'une stratégie de course. Cette partie mêle mathématiques et informatique dans le but de créer un micro service web que nous pourrions utiliser dans diverses applications.

Le [chapitre III](#) donne un regard sur l'état de notre réflexion concernant la stratégie de course. Premièrement, nous proposons de nouveaux outils permettant de décrire la façon de courir un marathon suivant différents segments (sexes, niveau de performance, profil de course). Deuxièmement, nous appliquons ces mêmes outils sur le récent record du monde établi à Berlin par Eliud Kipchoge afin de vérifier nos théories.

Le [chapitre IV](#), grâce à une étude sur l'accélération humaine, donne des perspectives sur les consignes que nous pourrions donner en temps réel à un coureur pendant un marathon afin qu'il change de vitesse ou de sensation pour battre ses précédents records.

Le [chapitre V](#) décrit les produits issus de ces trois ans de recherche ayant un but commercial. Nous présenterons deux outils. Le premier est une application web qui a vu le jour en

production en Juin 2018 et qui permet de souscrire et de faire vivre la méthode *BillaTraining*[®] de manière automatisée. Le second produit est un prototype d'application mobile permettant d'enregistrer les données de vitesse et de fréquence cardiaque d'un coureur à pied et d'interagir avec lui en temps réel grâce aux technologies web.

Chapitre I

Contexte physiologique et d'entreprise

Dans ce premier chapitre, nous allons positionner la thèse selon deux axes. Le premier décrit le contexte physiologique. Nous allons vous décrire l'idée à l'origine de ce travail puis son évolution au sein de l'entreprise *BillaTraining*[®]. Nous verrons comment la thèse se place en marge des travaux existants grâce à l'interdisciplinarité entre la physiologie, les mathématiques et l'informatique. Le deuxième axe donne le contexte d'entreprise. Nous décrierons succinctement la méthode *BillaTraining*[®] créée par le professeur Véronique Billat, de l'audit permettant d'estimer la forme d'un coureur au plan de course sur mesure, puis, nous donnerons le contexte concurrentiel. Enfin nous présenterons le matériel que nous avons eu l'occasion d'utiliser, surtout lors des premiers mois de cette thèse.

Sommaire

| | | |
|----------|--|-----------|
| 1 | Problématiques et enjeux | 20 |
| 2 | Etat de l'art et positionnement de la thèse | 21 |
| 2.1 | Joseph B. Keller (1973) | 23 |
| 2.2 | Frank Mathis (1989) | 26 |
| 2.3 | William Woodside (1991) | 28 |
| 2.4 | Horst Behncke (1993) | 30 |
| 2.5 | Daniel Whitt (2009) | 32 |
| 2.5.1 | Le modèle probabiliste | 32 |

| | | |
|----------|---|-----------|
| 2.5.2 | Le processus stochastique | 34 |
| 2.6 | Amandine Aftalion et Frédéric Bonnans (2013) | 35 |
| 2.7 | Positionnement de la thèse | 39 |
| 3 | L'entreprise <i>BillaTraining</i>[®] et ses méthodes | 41 |
| 3.1 | L'entreprise en bref | 42 |
| 3.2 | Le test de RabiT [®] | 44 |
| 3.3 | Le Radar énergétique | 46 |
| 3.4 | L'entraînement sur mesure | 48 |
| 3.5 | La Signature de Vitesse [®] | 49 |
| 4 | Etat des lieux de la concurrence | 50 |
| 5 | Le matériel à disposition | 52 |
| 5.1 | <i>Cosmed</i> [®] - K5 - Analyseur de gaz | 52 |
| 5.2 | <i>Garmin</i> [®] - 920XT - Montre de sport | 54 |
| 5.3 | <i>Physioflow</i> [®] - PF07 Enduro - Débit cardiaque | 55 |
| 5.4 | <i>Centaure Metria</i> [®] - Locometrix - Accéléromètre triaxial | 56 |
| | Bilan | 56 |

1 Problématiques et enjeux

Comment optimiser sa vitesse de déplacement en course à pied ? L'idée de cette thèse est née après 30 ans de recherche et d'expériences sur le terrain du Professeur Véronique Billat. Suite à un surentraînement et plusieurs blessures liées à ce phénomène, le Professeur Billat décide de travailler sur la compréhension des mécanismes de la course à pied. Elle en vient à la conclusion qu'un entraînement dense, long et rébarbatif est chronophage, augmente le risque de blessures physiques, affecte négativement la motivation et met le coureur dans un état de fatigue avant même le jour de la course. L'entreprise *BillaTraining*[®] (Cf. sous chapitre 3) découle du savoir faire du Professeur Billat en accompagnant les coureurs à pied dans leurs

objectifs grâce à un entraînement et une stratégie de course personnalisée. Un des enjeux de cette thèse est de proposer une nouvelle stratégie de course à pied pour des courses sur le plat comprises entre le 10 km et le Marathon. Par ailleurs, cette thèse a pour vocation d'amener les coureurs à pied à remettre en question le paradigme de la vitesse constante en courant avec des allures variées et en suivant leurs sensations du jour. Pour ce faire, nous proposons un nouveau modèle prédictif et personnalisé de stratégie de course à pied basé sur des variations de vitesse. Nous souhaitons que ce modèle soit simple à essayer en conditions réelles ; c'est-à-dire uniquement à l'aide d'une montre GPS ou d'un téléphone portable.

2 Etat de l'art et positionnement de la thèse

Nous allons, dans cette partie, aborder des différentes tentatives de modélisation de stratégies prédictives de courses à pied. Commençons par rendre hommage aux précurseurs ayant travaillé sur la problématique de prévision de la performance en proposant des modèles que nous pouvons décomposer en 2 grandes catégories :

- Les modèles empiriques : il s'agit des modèles de prévision de la performance basés sur les records mondiaux du 100 m au 1000 km et plus généralement de résultats d'expérimentations. Nous pouvons citer, sans être exhaustif, les travaux de Kennelly [104, 105], Meade [125, 126], Grosse-Lordemann et Müller [74], Tornvall [161], Keller [100], Craig [50] et Harman [80].
- Les modèles bioénergétiques : basés sur les caractéristiques physiologique des coureurs à pied (e.g. $\dot{V}O_2$, fréquence cardiaque, fréquence respiratoire, ...). Nous pouvons citer, sans être exhaustif, les travaux de Hill [88], Henry [85, 86, 87], Scherrer [152], Ettema [59], Margaria [120], Di Prampero [143], Morton [129] et Billat [27]

Toutefois, il s'avère que la frontière entre les modèles empiriques et bioénergétiques est faible, les uns s'inspirant des autres. Cette classification a pour seul but de distinguer les approches mathématiques (i.e. les modèles empiriques) des approches physiologiques (i.e. les modèles bioénergétiques).

Pour une approche plus complète détaillant chacun de ces précurseurs, nous vous invitons à vous procurer la thèse de Richard Hugh Morton datant de 1984 sur le sujet "A mathematical and computer simulation model of the running athlete" [127] ainsi que la thèse de Valéry Bocquet datant de 2002 sur le sujet "Modélisation de l'endurance humaine, de l'espace à l'individu" [37] notamment le chapitre 3 intitulé "Les modèles de prévision de la performance".

Les technologies progressent de jour en jour et nous permettent d'enregistrer des données mécaniques et physiologiques de plus en plus précises. Il devient possible d'analyser des courses passées et même de surveiller en temps réel les caractéristiques d'un coureur à pied. Un des objectifs de cette thèse est d'essayer de comprendre cet amas de données et de leur donner du sens. Plus précisément, nous voulons comprendre la façon dont sont réalisées les performances en course à pied du 10km au Marathon. Plusieurs auteurs ont abordé ce sujet au cours du XXème et XXIème siècle en essayant de modéliser la façon optimale de courir. Les mots clés "pacing + strategy + running" m'ont permis d'établir la liste des scientifiques traitant du sujet :

- Keller, "A theory of competitive running", 1973 [100] et "Optimal Velocity in a race", 1974 [101]
- Mathis, "The effect of fatigue on running strategies", 1989 [124]
- Woodside, "The optimal strategy for running a race (a mathematical model for world records from 50 m to 275 km)", 1991 [177]
- Behncke, "A mathematical model for the force and energetics in competitive running", 1993 [15]
- Whitt, "Optimal strategies for stochastic models of running performance", 2009 [173]
- Aftalion et Bonnans "Optimizing of running strategies based on anaerobic energy and variations of velocity", 2013 [2]

Détaillons un à un les modèles proposés par ces scientifiques par ordre chronologique.

2.1 Joseph B. Keller (1973)

Joseph Keller est le premier auteur à proposer une stratégie optimale en course à pied. Basé sur la seconde loi de Newton, l'objectif de Keller est de trouver le temps T le plus court pour parcourir la distance D . Keller donne quatre contraintes physiques et physiologiques à son problème :

1. La force maximale horizontale qu'un coureur peut exercer, dénommée F .
2. La force de résistance qui s'oppose au coureur, proportionnelle à la vitesse, $\frac{v(t)}{\tau}$, où τ est une constante de temps.
3. Le taux d'énergie fourni par l'approvisionnement en oxygène σ .
4. La quantité finie d'énergie présente dans le corps du coureur avant le départ E_0 .

Les quatre quantités ci-dessus peuvent être obtenues de deux manières différentes. La première par une estimation en effectuant une somme des moindres carrés entre le modèle de Keller et au moins 4 records du monde. La deuxième par une mesure physiologique de chacune des quatre grandeurs.

Le modèle de Keller s'applique aux distances comprises entre 50 mètres et 10 kilomètres. Pour les distances supérieures, l'auteur pense que d'autres contraintes physiologiques pourraient apparaître comme l'accumulation et le recyclage de l'acide lactique. Néanmoins, la théorie de Keller est segmentée en deux catégories par l'introduction de la distance critique (D_c) qu'il considère comme étant égale à 291 mètres. La première catégorie sont les courses inférieures à la distance critique qu'il qualifie de "sprint court". Le coureur part avec une vitesse nulle et accélère le plus vite possible à sa vitesse maximale qu'il maintient ensuite jusqu'à la fin de la course. La deuxième correspond aux courses supérieures ou égales à la distance critique. Les dites courses se décomposent alors en trois phases. Accélération maximale pendant 1 à 2 secondes, puis vitesse constante et enfin un léger lâché prise à la fin de la course. La figure [I.1](#) donne une représentation graphique d'une course supérieure ou égale à D_c . Le palier à vitesse constante de la figure [I.1](#), qui représente la majorité de la course par

la même occasion, est décrit par Keller comme un point de vue accepté par les physiologistes et expérimentales des années 80 comme le moyen d'effectuer le temps le plus court sur une distance donnée sans toutefois citer de sources.

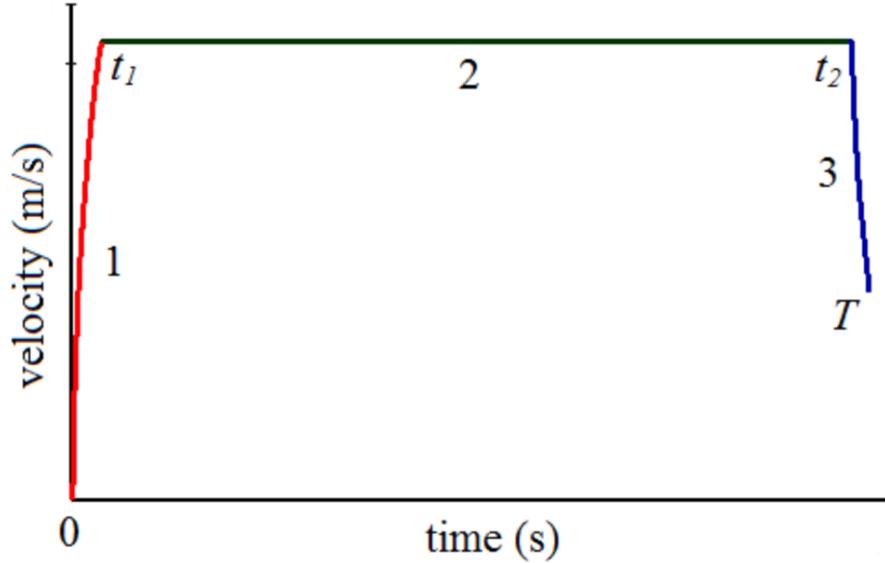


FIGURE I.1 – Tracé de la vitesse v en fonction du temps t pour une course supérieure à la distance critique. La phase 1 pour $0 \leq t \leq t_1$ représente l'accélération de 1 à 2 s, la phase 2 pour $t_1 \leq t \leq t_2$ représente la phase de vitesse constante et la phase 3 pour $t_2 \leq t \leq T$ est le lâché prise final.

Etant donné que nous n'essayons pas de modéliser le sprint dans cette thèse, nous allons nous concentrer sur la présentation du modèle pour les courses supérieurs à D_c . Il s'agit d'un problème de commande optimale où Keller cherche à minimiser le temps. Il part de l'équation I.1 en supposant que le temps de course T est lié à la vitesse à l'instant t , $v(t)$ et la distance totale de course D .

$$D = \int_0^T v(t)dt \quad (\text{I.1})$$

La vitesse $v(t)$ est donnée par l'équation différentielle I.2

$$\frac{dv}{dt} + \frac{v}{\tau} = f(t) \quad (\text{I.2})$$

où τ est supposé constant et $\frac{v}{\tau}$ est la force de résistance qui s'oppose au coureur par unité de masse et $f(t)$ est la force de propulsion du coureur par unité de masse.

L'équation différentielle I.3 donne l'équilibre en oxygène du coureur avec $E(t)$ l'oxygène disponible à l'instant t dans les muscles du coureur par unité de masse. La consommation d'oxygène pendant la course est supposée être utilisée au taux fv . Enfin, l'approvisionnement en oxygène est représenté par σ .

$$\frac{dE(t)}{dt} = \sigma + f(t)v(t) \quad (\text{I.3})$$

Pour finir avec la formulation du problème, Keller pose quatre contraintes à son problème d'optimisation. La première indique que la vitesse initiale est toujours nulle (équation I.4a). La deuxième que la force exercée à l'instant t par le coureur par unité de masse ne peut pas dépasser la force maximale F qu'il est capable d'exercer (équation I.4b). La troisième est que l'énergie dans le corps du coureur au départ de la course est finie et s'écrit E_0 (équation I.4c). La dernière impose le fait que l'énergie présente dans le corps du coureur par unité de masse ne peut pas être négative (équation I.4d).

$$v(0) = 0 \quad (\text{I.4a})$$

$$f(t) \leq F \quad (\text{I.4b})$$

$$E(0) = E_0 \quad (\text{I.4c})$$

$$E(t) \geq 0 \quad (\text{I.4d})$$

L'objectif étant de trouver les fonctions $v(t)$, $f(t)$ et $E(t)$ en minimisant le temps pour parcourir la course T . Il s'agit d'un problème de commande optimale.

Dans sa reformulation du problème, Keller exprime $f(t)$ et $E(t)$ en fonction de $v(t)$. L'équation I.5 représentent les trois fonctions vitesses solutions du problème d'optimisation pour respectivement $0 \leq t \leq t_1$, $t_1 \leq t \leq t_2$ et $t_2 \leq t \leq T$ pour des courses supérieures à 291

m et inférieures à 10 km (Cf. figure I.1 pour une représentation graphique).

$$v(t) = \begin{cases} F\tau \left(1 - e^{-\frac{t}{\tau}}\right), & 0 \leq t \leq t_1 \\ \left(\frac{2\tau(E_0 + \sigma T)}{2 + \tau}\right)^{1/2}, & t_1 \leq t \leq t_2 \\ \left(\sigma\tau + [v^2(t_2) - \sigma\tau]e^{2\frac{(t_2-t)}{\tau}}\right)^{1/2}, & t_2 \leq t \leq T \end{cases} \quad (\text{I.5})$$

C'est donc Keller qui à partir de 1973 est le premier à proposer une stratégie de course en trois phases. Keller explique dans sa conclusion que son modèle est encore trop simple et qu'une incorporation d'autres paramètres physiologiques est nécessaire afin de l'améliorer. Par ailleurs, deux perspectives de recherche présentes dans la conclusion de Keller nous semblent intéressantes. La première perspective est la prise en compte du dénivelé et donc du profil de la course. En effet, seules les courses sur le plat peuvent appliquer la théorie de Keller. La seconde perspective est la recherche de stratégie de course concernant d'autres sports comme la natation ou le cyclisme pour ne donner que deux exemples.

2.2 Frank Mathis (1989)

Frank Mathis est le premier à faire remarquer que la théorie de Keller ne comporte aucun facteur de fatigue. Dans son papier de 1989, il décide de partir de la théorie de Keller et d'ajouter un facteur fatigue afin de déterminer la stratégie de course optimale. Pour ce faire, Mathis suppose que la force propulsive du coureur $f(t)$ est bornée par une fonction qui décroît au fur et à mesure de la course. L'auteur pense que cette décroissance provient du déficit d'oxygène (le différentiel entre la quantité d'oxygène fournie par la respiration et celle consommée) du coureur. Il modélise la décroissance grâce à l'équation I.6.

$$d(t) = \max\left\{0, \int_0^t f v ds - \sigma t\right\} \quad (\text{I.6})$$

σ correspond à l'approvisionnement en oxygène, $v(t)$ est la vitesse du coureur à l'instant t .

Cette idée implique la modification de la contrainte sur la force propulsive de l'article de

Keller ($f(t) \leq F$) qui devient l'équation I.7.

$$f(t) \leq F - cd(t) \quad (\text{I.7})$$

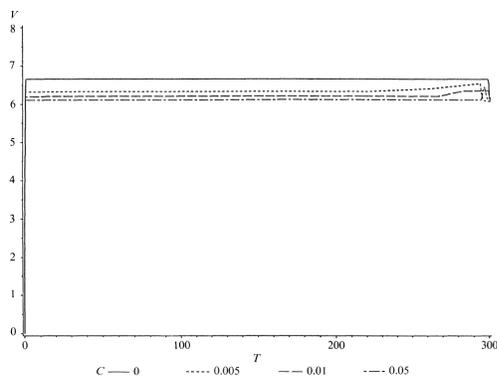
F correspond à la force maximale que le coureur peut exercer et c est une contrainte physiologique liée à la condition physique du coureur. L'auteur ne donne aucune signification concrète de c .

Mathis fait remarquer que pour $c = 0$, on retrouve la théorie de Keller. Pour $c \neq 0$, il discrétise le problème et cherche toutes les vitesses v_i pour chaque instant $i = 0, 1, 2, \dots, N$. Il pose alors le problème d'optimisation cherchant à maximiser la distance (\hat{D}) pour un temps donné (équation I.8)

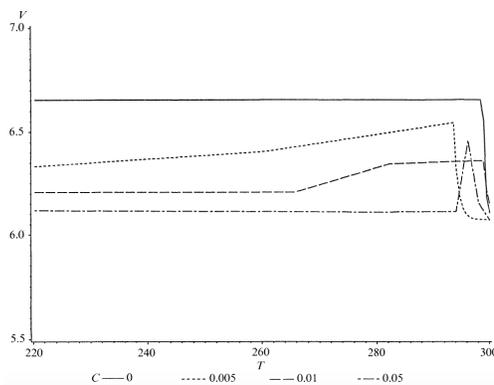
$$\begin{aligned} \text{fonction objectif,} \quad & \hat{D} = \sum_{i=1}^N \frac{1}{2}(v_{i-1} + v_i)\Delta t_i \\ \text{contraintes,} \quad & v_0 = 0, \\ & \left(1 + \frac{1}{2}\Delta t_i \tau\right)v_i - \left(1 - \frac{1}{2}\Delta t_i \tau\right)v_{i-1} \leq \Delta t_i \left[F - \frac{1}{2}c(d_i + d_{i-1})\right], \quad (\text{I.8}) \\ & E_i = E_0 + \sigma t_i - \frac{1}{2}v_i^2 - \tau \sum_{j=1}^i \frac{1}{2}(v_{j-1}^2 + v_j^2)\Delta t_j \geq 0, \\ & d_i = \max\{0, E_0 - E_i\}, \quad i = 1, 2, 3, \dots, N \end{aligned}$$

Les différents résultats s'affichent sur la figure I.2 pour différentes valeurs de c ($c \neq 0$) et avec la théorie de Keller comme moyen de comparaison ($c = 0$). La figure I.2a donne l'ensemble de la course (300 secondes dans l'article de Mathis) et la figure I.2b donne un détail de la fin de la course (80 secondes dans l'article de Mathis).

Pour résumer la stratégie de course de Mathis, le coureur accélère sur un très court intervalle de temps au début de la course. Il tient ensuite une vitesse constante sur la majorité de la course puis accélère à nouveau sur la fin de la course pour finalement décélérer dans les tous derniers instants.



(a) Course de 300 secondes avec différentes valeurs de c



(b) 80 dernières secondes de la course de 300 secondes avec différentes valeurs de c

FIGURE I.2 – Résultats de différentes stratégies de courses selon Mathis. Les courbes sont issus de l'article de 1989 de Mathis.

2.3 William Woodside (1991)

Deux ans plus tard, William Woodside décide de reprendre la théorie de Joseph Keller basée sur la vitesse constante (voir 2.1) et de l'étendre aux courses supérieures à 10 km. Après l'explication de la théorie de Keller, Woodside étend le modèle de son prédécesseur en y ajoutant un facteur fatigue qui permet de donner une prédiction de stratégie de course jusqu'à 275 km (distance maximale parcourue par un homme en 24h à l'époque; Ce record est aujourd'hui de 290 km sur route et 303 km sur piste). Pour rappel, Keller propose deux stratégies de courses distinctes. La première pour les courses inférieures à 219 m où il préconise la recherche d'une accélération maximale pendant toute la durée de la course. La seconde stratégie concerne les courses comprises entre 291 m et 10 km où il propose un modèle en trois phases. La première phase consiste à courir à accélération maximale pendant 1 à 2 secondes. La deuxième phase est le maintien d'une vitesse constante jusqu'à la quasi fin de course. La troisième et dernière phase est une décélération progressive lors des dernières secondes de la course. Keller indiquait dans son article que pour les courses supérieures à 10 km, un facteur fatigue devait être ajouté à son problème d'optimisation. Woodside introduit alors dans son papier la contrainte de fatigue. Le problème d'optimisation à désormais cinq contraintes physiques et physiologiques.

1. La force maximale qu'un coureur peut exercer, dénommée F (Keller, 1974)
2. La force de résistance qui s'oppose au coureur, proportionnelle à la vitesse $\frac{v(t)}{\tau}$ (Keller, 1974)
3. Le taux d'énergie fourni par l'approvisionnement en oxygène σ (Keller, 1974)
4. La quantité d'énergie présente dans le corps du coureur avant le départ, E_0 (Keller, 1974)
5. La fatigue représentée par un taux additionnel de perte d'énergie au temps t qui est proportionnel à l'énergie déjà dépensée depuis le début de la course au temps t , symbolisée par k (Woodside, 1991)

L'équation I.9 donne la fatigue du coureur à l'instant t selon Woodside.

$$k \int_0^t f(t)v(t)dt \quad (\text{I.9})$$

L'auteur modifie en conséquence l'équation différentielle donnant l'équilibre en oxygène du coureur de Keller. l'équation I.3 devient I.10

$$\frac{dE(t)}{dt} = \sigma + f(t)v(t) - k \int_0^t f(t)v(t)dt \quad (\text{I.10})$$

Pour donner finalement la même philosophie de course que Keller (globalement à vitesse constante) avec les 3 fonctions de vitesse données par l'équation I.11

$$v(t) = \begin{cases} F\tau(1 - e^{-\frac{t}{\tau}}), & 0 \leq t \leq t_1 \\ V, & t_1 \leq t \leq t_2 \\ \left\{ \left(V^2 + kV^2t_2 - \sigma\tau - \frac{1}{2}kV^2\tau \right) e^{-2\frac{(t-t_2)}{\tau}} + \sigma\tau + \frac{1}{2}kV^2(\tau - 2t) \right\}^{\frac{1}{2}}, & t_2 \leq t \leq T \end{cases} \quad (\text{I.11})$$

avec $V = \left[\frac{2\tau(E_0 + \sigma T)}{\tau + 2T + kT^2} \right]^{\frac{1}{2}}$ la vitesse constante sur la majorité de la course.

Pour estimer les cinq coefficients physiques et physiologiques F , τ , σ , E_0 et k , Woodside effectue une analyse des moindres carrés selon le même procédé que Keller en s'appuyant sur les records du monde passés. F et τ sont estimés à partir des records du monde sur courses courtes (inférieures à 291 m) tandis que E_0 , σ et k sont estimés à partir des courses longues. Il est à noter que Woodside considère que pour les courses inférieures à 2 ou 3 km, les effets de la fatigue sont négligeables.

En plus d'introduire le facteur fatigue, il est intéressant de noter la segmentation proposée par l'auteur entre Femmes et Hommes pour la détermination des cinq contraintes physiologiques. Nous trouvons cette idée intéressante et il pourrait être pertinent de pousser plus loin la segmentation en regroupant par plage de performance également. Ceci irait dans le sens de la personnalisation de la stratégie de course.

2.4 Horst Behncke (1993)

Horst Behncke est le premier à apporter une alternative au modèle de stratégie de course de Keller en détaillant les échanges d'énergie dans le corps humain. L'auteur commence par donner les trois sources d'énergie nécessaires à la course à pied.

1. ATP et créatine phosphate (métabolisme anaérobie alactique)
2. Glycolyse (métabolisme anaérobie lactique)
3. Oxydation des lipides et des glucides (métabolisme aérobie)

Behncke propose un modèle hydraulique basé sur les théories de Margaria [119] et Morton [128]. Le modèle est composé de trois compartiments correspondant aux trois sources d'énergie décrite ci-dessus : créatine phosphate (P), glycolyse (L) et oxydation des lipides et glucides (O). La figure I.3 donne une représentation graphique du modèle hydraulique en question. M_O et M_L représentent les débits maxima pour respectivement les apports en énergie de O

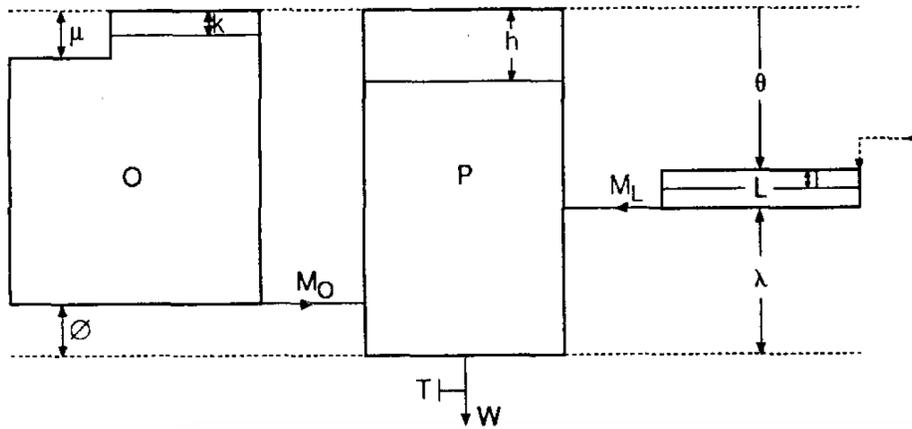


FIGURE I.3 – Figure issue des travaux de Behncke. La figure représente un modèle hydraulique à 3 composants représentant les 3 sources d'énergie : Créatine phosphate (P), Glycolyse (L) et Oxydation des lipides et glucides (O)

vers P et de L vers P . Les fonctions k , h et l représentent le niveau des réservoirs d'énergie en fonction du temps. W correspond à la puissance musculaire. θ , λ , ϕ et μ sont les paramètres géométriques du système.

Après avoir détaillé les courses de courtes distances (sprint), l'auteur propose d'étudier les courses longues. Pour ce faire, il propose de réunir les réservoirs L et P car il fait l'hypothèse que les effets liés à la créatine phosphate sont négligeables sur de telles distances (supérieures à 5000 m). Après résolution de son problème d'optimisation, Behncke tombe sur une solution optimale en trois phases. La première phase est une accélération maximale pendant un court instant (comme dans le modèle de Keller). La seconde phase est une vitesse augmentant légèrement en fonction du temps de course. Cette deuxième phase correspond à la majorité de la course. L'équation I.12 modélise la phase deux selon l'auteur.

$$v(t) = v(T_0)e^{-(t-T_0)/\tau} \quad (\text{I.12})$$

où $v(T_0)$, T_0 et τ sont trois constantes calculées sur les précédents records du monde. La troisième et dernière phase correspond à une augmentation de la vitesse sur quelques secondes.

Behncke, dans son article, introduit la composante bioénergétique au modèle de Keller. Il propose un modèle complexe dans lequel la seconde loi de Newton est complétée par un

modèle hydraulique comprenant trois sources d'énergies correspondant aux différentes filières énergétiques. Il conserve néanmoins sur le paradigme de vitesse constante comme meilleure stratégie de course pour les distances longues.

2.5 Daniel Whitt (2009)

Après avoir présenté la théorie déterministe de Keller et ses extensions par Mathis et Woodside, Daniel Whitt propose d'aborder le problème sous un angle stochastique de deux manières différentes. Premièrement grâce à un modèle probabiliste basique sur l'énergie du coureur. Deuxièmement en introduisant un processus stochastique sur la consommation d'oxygène. L'auteur conserve le même paradigme de course à vitesse constante que ses prédécesseurs avec une légère différence à la fin de la course. Au lieu de décélérer sur la fin de la course, Whitt propose une accélération. Il tient cette intuition de l'observation de différents records, notamment celui de Haile Gebresellasié sur 10.000m sur piste en 1998. Comme ses prédécesseurs, il explique que cette brève accélération de fin de course est due à la volonté de battre ses adversaires plus que battre le record du monde.

2.5.1 Le modèle probabiliste

Whitt commence par retenir tous les aspects de la théorie de Keller à l'exception de la quantité d'énergie présente dans le corps du coureur au départ de la course, E_0 , qu'il considère comme une variable normale aléatoire avec faible variance. Pour rappel, Keller considère cette quantité comme constante. Dans son premier modèle, Whitt introduit la variable normale aléatoire X ($X \sim \mathcal{N}(m, \sigma^2)$) pour remplacer E_0 dans le modèle de Keller. Cependant, il conserve E_0 pour déterminer l'instant où le coureur n'a plus d'énergie. X est utilisée pour construire la stratégie de course uniquement.

Premièrement, l'auteur observe les prédictions de performance d'un coureur pour différente valeur de E_0 définie par E_* . Il considère la distribution discrète suivante $E_0 - 0.5E_0 \leq E_* \leq E_0 + 0.5E_0$. Après une résolution numérique des équations de Keller pour différentes valeurs de E_* , il introduit la fonction g où pour chaque E_* , il existe un temps final T tel que

$g(E_*) = T$. Il trouve alors une fonction affine en deux morceaux avec un minimum en E_0 . La figure I.4 illustre ce résultat.

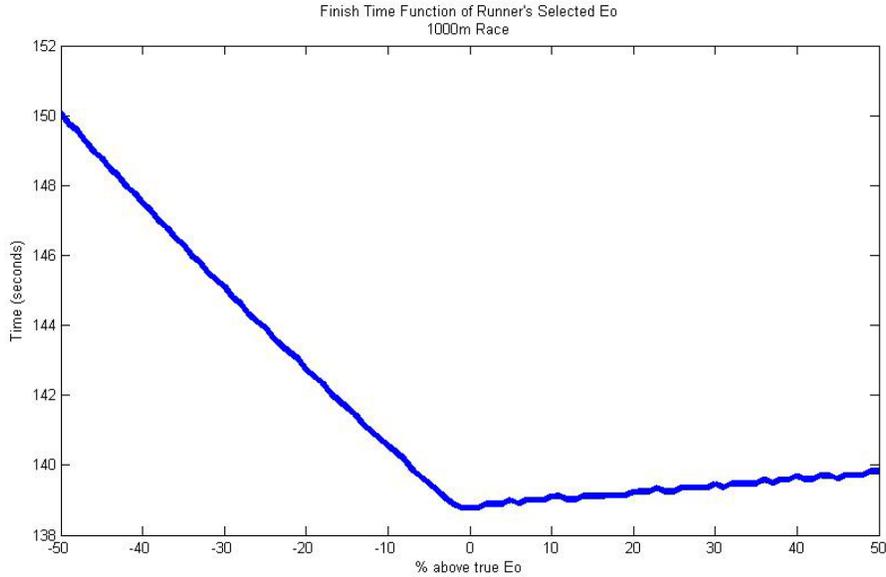


FIGURE I.4 – D’après Whitt, 2009. La figure représente la distribution des temps finaux par rapport à la valeur de E_* pour un coureur sur 1000m

Ce résultat est valable pour toutes les distances hors sprints selon la définition de Keller de distance critique (291m à 10km). La fonction g prend toujours la même forme. Seuls les coefficients directeur des deux morceaux varient avec la distance de course.

En s’appuyant sur les résultats de la figure I.4, il définit la partie de gauche avec la fonction $h(x) = c + dx$ et la partie de droite par la fonction $f(x) = a + bx$. a, b, c et d étant des constantes. L’objectif de l’auteur à ce point est de déterminer les paramètres a, b, c et d . Il propose d’évaluer numériquement les 3 équations de Keller (3.12-3.14 de l’article de Keller de 1974) afin de trouver les quatre paramètres décrits ci-dessus.

L’auteur cherche ensuite à trouver la valeur de m (moyenne de la variable normale aléatoire X) optimale afin que le coureur minimise son temps final T . Il pose alors les équations I.13.

$$g(x) = \begin{cases} f(x) = a + bx, & x > 0 \\ h(x) = c + dx, & x \leq 0 \end{cases} \quad (\text{I.13})$$

Il pose finalement une expression analytique de l'espérance de $g(x)$ (équation I.14) après une suite de calculs.

$$E[g(x)] = \left(1 - \Phi(-m/\sigma)\right) \left(a + b \left[m + \sigma \frac{\phi(-m/\sigma^2)}{1 - \Phi(-m/\sigma^2)}\right]\right) + \Phi(-m/\sigma) \left(c + d \left[m + \sigma \frac{-\phi(-m/\sigma^2)}{\Phi(-m/\sigma^2)}\right]\right) \quad (\text{I.14})$$

où Φ représente la fonction de répartition et ϕ la densité de probabilité de la variable normale aléatoire X standardisée. L'auteur en déduit alors un m optimal en minimisant l'équation I.14. Il trouve que le m est légèrement supérieur à la valeur E_0 de l'article de Keller préconisant au coureur de légèrement surestimer son E_0 au départ de la course.

Pour résumer le premier modèle proposé par Whitt, il commence par fixer la distance totale de course et la variance (σ^2) de la variable normale aléatoire X ($X \sim \mathcal{N}(m, \sigma^2)$) (qui représente l'énergie du coureur au départ de la course). Il calcule ensuite les paramètres a,b,c et d pour déterminer les coefficients directeurs des deux pentes des fonctions affines par morceaux de la distribution des temps finaux en fonction des différentes valeurs de X pour $m = E_0$. Il finit par déterminer le m optimal en minimisant l'équation I.14.

2.5.2 Le processus stochastique

Dans ce deuxième modèle, Whitt propose de s'attaquer au taux d'approvisionnement en oxygène de Keller ($\bar{\sigma}$). En effet, il note que Keller considère que tous les coureurs ont le même $\bar{\sigma}$. Il qualifie σ comme une distribution normale autour de la valeur constante de $\bar{\sigma}$ trouvée par Keller (équation I.15).

$$\sigma(t) = \bar{\sigma} + KB_t \quad (\text{I.15})$$

où K est une constante qui ajuste la variance du modèle et B_t est un mouvement Brownien standard. L'auteur finit par tomber sur un problème d'optimisation qu'il ne peut pas résoudre en l'état car l'énergie du coureur n'est pas contrainte à rester positive ou nulle. Par définition, l'énergie d'un coureur ne pouvant pas être négative, il n'existe aucune solution possible. Le

concept du problème bien posé [54, 76] n'est pas respecté.

In fine, Daniel Whitt ne propose pas de réelle stratégie de course mais à l'intérêt d'introduire le fait que tous les coureurs sont différents et qu'il semble difficile de tous les classer suivant une ou deux catégories (Femmes et Hommes). Ce travail de Whitt fait poser deux questions. La première est la possibilité d'une stratégie de course personnalisée et unique à chaque individu. La seconde est l'introduction d'une (ou plusieurs) variable de contrôle pour vérifier la bonne réalisation de cette stratégie en temps réelle.

2.6 Amandine Aftalion et Frédéric Bonnans (2013)

Quatre ans plus tard, Aftalion et Bonnans proposent une approche basée sur le modèle de Keller combinée avec l'approche de Morton et son analogie hydraulique. La nouveauté apportée dans cet article est l'introduction de variation de vitesse découlant des travaux indiquant qu'un coureur ne conserve pas une allure constante durant un exercice de course à pied [21, 79]. Les auteurs présentent un problème d'optimisation en commençant par rappeler l'équation de mouvement de l'article de Keller (équation I.2). En parallèle de l'énergie aérobie (oxydation des lipides et des glucides) décrite par Keller, les auteurs introduisent l'énergie anaérobie $e_{an}(t)$ (anaérobie lactique (Glycolyse) et anaérobie alactique (créatine phosphate)). L'équation I.16 donne les quatre contraintes du problème.

$$\begin{aligned} 0 &\leq f(t) \leq f_{max}, \\ v(0) &= 0, \\ e_{an}(0) &= e_{an}^0, \\ e_{an}(t) &\geq 0 \end{aligned} \tag{I.16}$$

Le but étant de minimiser le temps final T en sachant la distance $d = \int_0^T v(t)dt$.

Aftalion et Bonnans présentent alors les quatre étapes qui leurs ont permis d'arriver à une nouvelle stratégie de course à pied. Nous allons les détailler dans la suite de cette sous-section.

Etape 1 : Ajustement et justification du modèle de Keller

Après une explication du modèle hydraulique, les auteurs écrivent l'équation I.17.

$$\frac{de_{an}}{dt} = \bar{\sigma} - fv \quad (\text{I.17})$$

$\bar{\sigma}$ correspond à la consommation maximale d'énergie ($\dot{V}O_2max$). Ils retombent sur l'équation d'énergie proposée par Keller avec une différence. Ce que Keller décrit comme étant l'énergie aérobique, Aftalion et Bonnans le caractérisent par l'accumulation du déficit en oxygène.

Après optimisation, il est intéressant de remarquer, grâce à la figure I.5, que la solution optimale donne une course en trois phases, ce qui confirme la théorie de Keller sur ce point.

1. La première phase étant une forte accélération à force maximale (f_{max}) sur un court instant.
2. La seconde phase est le maintien d'une vitesse constante sur la majorité de la course.
3. La troisième et dernière phase correspond à une vitesse décroissante.

Etape 2 : Amélioration de la phase d'accélération afin d'atteindre $\dot{V}O_2 \max$

Partant des études [21, 79] montrant que la consommation d'oxygène (σ) n'est pas constante sur la course mais monte progressivement jusqu'à atteindre un plateau ($\dot{V}O_2 \max$), les auteurs posent l'équation I.18.

$$\frac{de_{an}}{dt} = \begin{cases} \sigma_r + \lambda(\bar{\sigma} - \sigma_r)(e_{an}^0 - e_{an}) - fv, & \lambda(e_{an}^0 - e_{an}) < 1 \\ \bar{\sigma} - fv, & \lambda(e_{an}^0 - e_{an}) \geq 1 \end{cases} \quad (\text{I.18})$$

où $1/\lambda = e_{an}^0(1 - \varphi)$ et φ et σ_r deux constantes.

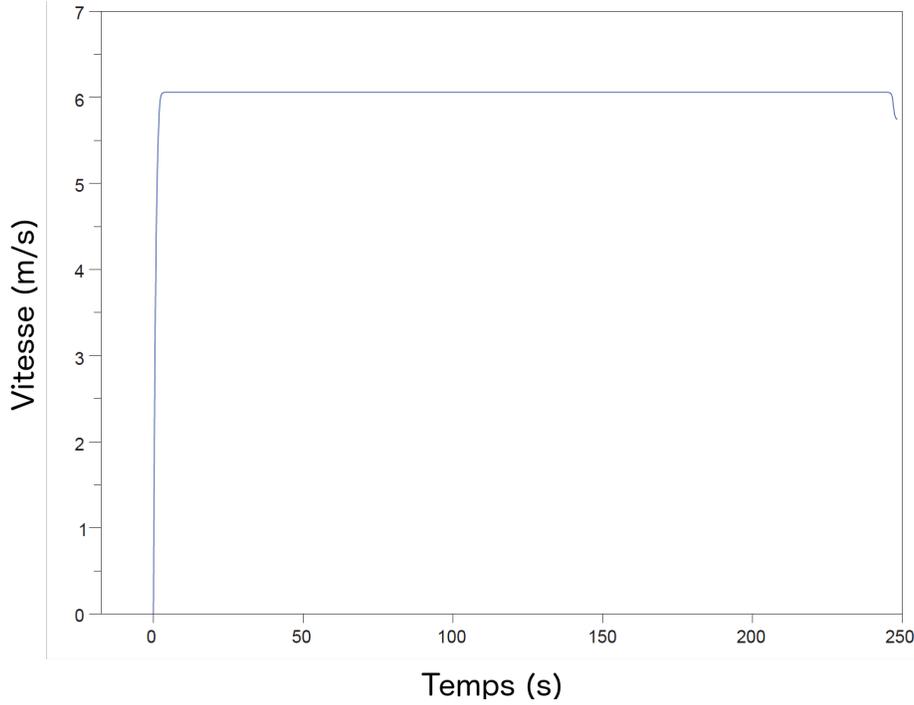


FIGURE I.5 – Figure issue des travaux d’Aftalion et Bonnans. La figure représente la vitesse en fonction du temps à l’étape 1 du modèle d’Aftalion et Bonnans

Étape 3 : Baisse de $\dot{V}O_2$ à la fin de la course

Reprenant les travaux montrant une légère baisse de la consommation d’oxygène [79] à la fin d’un exercice en course à pied, les auteurs introduisent une nouvelle constante appelée énergie critique (e_{crit}). L’équation I.19 donne la reformulation de l’équation d’énergie.

$$\frac{de_{an}}{dt} = \begin{cases} \bar{\sigma} \frac{e_{an}}{e_{an}^0 e_{crit}} + 0.75\bar{\sigma} \left(1 - \frac{e_{an}}{e_{an}^0 e_{crit}}\right) - fv, & \frac{e_{an}}{e_{an}^0} < e_{crit} \\ \bar{\sigma} - fv, & \frac{e_{an}}{e_{an}^0} \geq e_{crit} \text{ et } \lambda(e_{an}^0 - e_{an}) \geq 1 \\ \sigma_r + \lambda(\bar{\sigma} - \sigma_r)(e_{an}^0 - e_{an}) - fv, & \lambda(e_{an}^0 - e_{an}) < 1 \end{cases} \quad (\text{I.19})$$

les résultats sont présentés sur la figure I.6. Désormais, le modèle comprend les quatre phases suivantes :

1. Accélération à force maximale au démarrage de la course.
2. Légère décroissance de la vitesse jusqu’au milieu de la course.
3. Remontée légère de la vitesse.

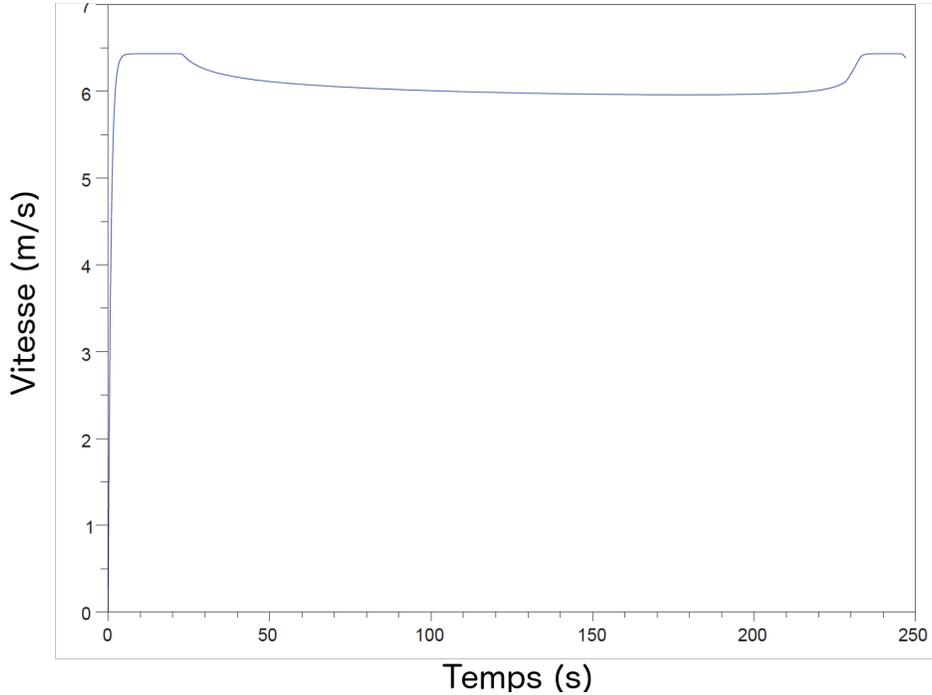


FIGURE I.6 – Figure issue des travaux d’Aftalion et Bonnans. La figure représente la vitesse en fonction du temps à l’étape 3 du modèle d’Aftalion et Bonnans

4. Accélération jusqu’à épuisement de l’énergie.

Etape 4 : Recréation d’énergie anaérobie pendant la course

Suite aux travaux de Morton et Billat [130] sur la recréation d’énergie pendant un exercice de course à pied, les auteurs décident d’ajouter un terme à l’équation d’énergie. Ils introduisent le facteur de recréation d’énergie $\nu(a)$ où a est l’accélération définie par $a = \frac{dv}{dt}$. Le facteur de recréation d’énergie est défini par l’équation I.20.

$$\nu(a) = \begin{cases} 0, & a \geq 0 \\ c|a|^2, & a \leq 0 \end{cases} \quad (\text{I.20})$$

c est une constant. En effet, Morton et Billat [130] ont trouvé que lors des phases de décélération, une partie de l’énergie anaérobie se recréait. La figure I.7 illustre le modèle final. On aperçoit sur la figure I.7 une vitesse périodique sur la majorité de la course.

L’article est intéressant par sa capacité à introduire la recréation d’énergie anaérobie

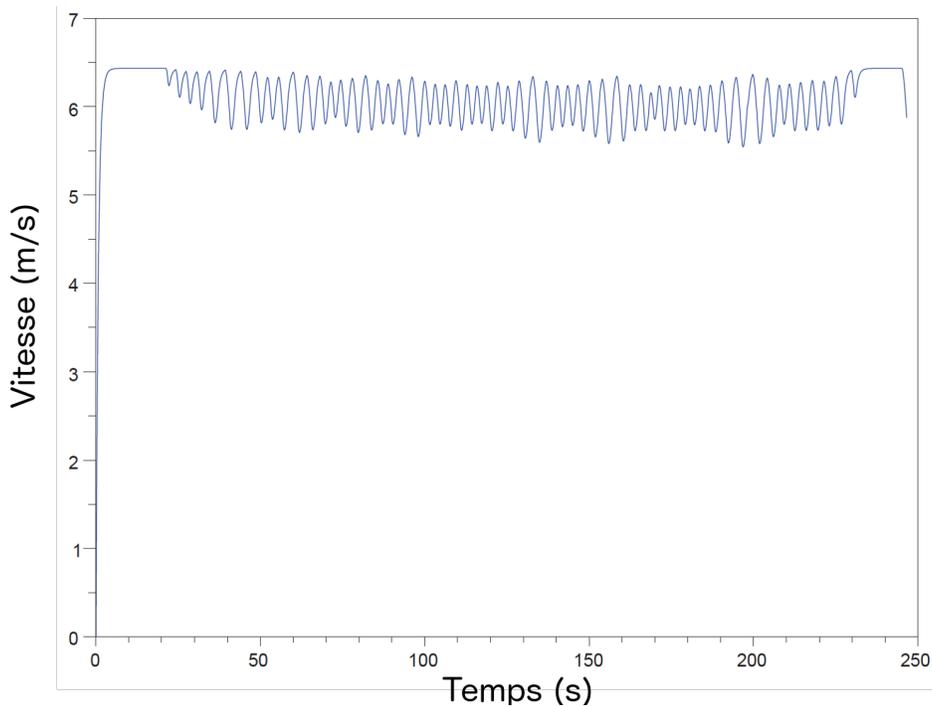


FIGURE I.7 – Figure issue des travaux d’Aftalion et Bonnans. La figure représente la vitesse en fonction du temps à l’étape 4 du modèle d’Aftalion et Bonnans

comme facteur permettant d’améliorer la performance en course à pied. Cependant, une telle stratégie de course semble difficile à réaliser en conditions réelles (sur le terrain) au vue de la périodicité très courte de la vitesse.

2.7 Positionnement de la thèse

Nous avons vu, dans ce sous chapitre consacré à l’état de l’art, deux grandes catégories afin de construire une stratégie de course optimale. La première basée sur l’équation de mouvement proposée par Keller propose une stratégie de course globalement à allure constante. Elle a le mérite de donner une bonne prédiction du temps final pour un athlète mais ne décrit pas la manière de courir de façon réaliste (Keller, Mathis, Woodside et Whitt). La seconde intègre, en plus de la théorie de Keller, un modèle hydraulique représentant les trois sources énergétiques, à savoir, créatine phosphate, glycolyse et oxydation des lipides et glucides (Behncke, Aftalion et Bonnans).

A la lecture de ces articles scientifiques, nous constatons qu’il est difficile de proposer une

stratégie de course réaliste. De plus, ces stratégies dites optimales (souvent basées sur les records du monde), représentent une façon de courir théorique et ne sont pas personnalisées. Posons donc le problème de la stratégie de course différemment. Nous savons que nous faisons face au dilemme suivant :

- Si nous voulons mesurer le maximum d'information pour un coureur, il nous faut lourdement l'équiper. Cette étape est très chronophage et ne permet pas d'avoir un échantillon de sportif conséquent. Par ailleurs, le jour de la course, le coureur n'aura accès qu'à sa vitesse et sa fréquence cardiaque.
- Si nous voulons les données d'un maximum de coureurs, nous n'auront pas accès aux données de consommation d'oxygène, volume d'éjection systolique et autres variables physiologiques.

L'idée est de combiner les deux approches en étudiant des phénomènes précis sur des petits groupes de personnes et en les validant sur le grand public. Dans tous les cas, nous pensons que nous devons ramener notre modèle de stratégie de course à la vitesse uniquement.

D'un côté, la physiologie doit nous aider à poser le problème de base afin de trouver une stratégie optimale personnalisée de course. Etant donnée qu'il est difficile de contrôler les variables physiologiques ($\dot{V}O_2$, FC, VES, ...) d'un coureur, la physiologie doit nous aider à associer une vitesse à un état physiologique. En effet, nous avons démontré que l'être humain est capable de contrôler son accélération. Nous le verrons dans le [chapitre IV](#). D'un autre côté, les mathématiques permettent d'apporter un cadre aux théories physiologiques en proposant une (des) solution(s) optimale(s) de stratégie de course à pied. A l'interface, l'informatique permet d'articuler la physiologie et les mathématiques.

Cette thèse se positionne à l'intersection de la physiologie, des mathématiques et de l'informatique afin de proposer une nouvelle stratégie de course personnalisée, réaliste et exploitable par tous les coureurs au vu des technologies disponibles actuellement du 10 km au Marathon. La suite du chapitre est dédiée au contexte d'entreprise.

3 L'entreprise *BillaTraining*[®] et ses méthodes

BillaTraining[®] est une entreprise qui offre un suivi de l'activité sportive et physique sous les 2 formes de transactions commerciales B2B ¹ et B2C ². Elle permet aux sportifs de se connecter avec leurs sensations en proposant un entraînement varié, sur mesure et composé de séances courtes au lieu de suivre un entraînement monotone, standardisé qui oublie l'aspect agréable de l'activité physique. Bien que faisant profiter d'une expérience ludique, la méthode proposée par *BillaTraining*[®] se base sur 30 ans de recherches menées par le professeur Véronique Billat. La méthode d'entraînement *BillaTraining*[®] peut se résumer en quatre étapes. La première est la réalisation d'un audit énergétique appelé test de RabiT[®] (*Running Advisor BillaTraining*) que nous aborderons dans la section 3.2. L'analyse du test de RabiT[®] donne lieu à un bilan de la forme du sportif représenté par le Radar Énergétique détaillé dans la section 3.3. Ensuite, un entraînement sur mesure d'une durée de trois mois découle de ces résultats, voir section 3.4. La dernière étape, donne une stratégie de course au sportif pour son objectif (course plate uniquement du 10km au Marathon). Cette étape porte le nom de Signature de Vitesse[®] et sera exposée en section 3.5. La figure I.8 met en image l'enchaînement de la méthode *BillaTraining*[®].



FIGURE I.8 – Les 4 étapes de la méthode *BillaTraining*[®] : Le RabiT[®], le Radar Énergétique, L'entraînement sur mesure, La Signature de Vitesse[®]

¹peut être écrit B to B ou Business to Business, cette annotation désigne l'ensemble des activités d'une entreprise visant une clientèle d'entreprises.

²peut être écrit B to C ou Business to Consumer, cette annotation désigne l'ensemble des activités d'une entreprise visant directement les consommateurs finaux.

3.1 L'entreprise en bref

BillaTraining[®] est une Société par Action Simplifiée ³ au capital sociale de 38.000 € fondée en juillet 2014 par le Pr. Véronique Billat qui siège désormais au conseil scientifique de l'entreprise. La présidence est assurée par Docteur Jean Pierre Koralzstein. Par ailleurs, *BillaTraining*[®] est composée de 2 salariés étudiants en doctorat à temps plein sous une convention CIFRE ⁴. Claire Molinari qui a débuté son doctorat en physiologie en 2016 sur le sujet "Energie à $\dot{V}O_{2max}$ facteur(s) limitant(s) à vitesse constante ou variable" et moi-même. La figure I.9 représente l'organigramme de la société au mois d'Août 2018.

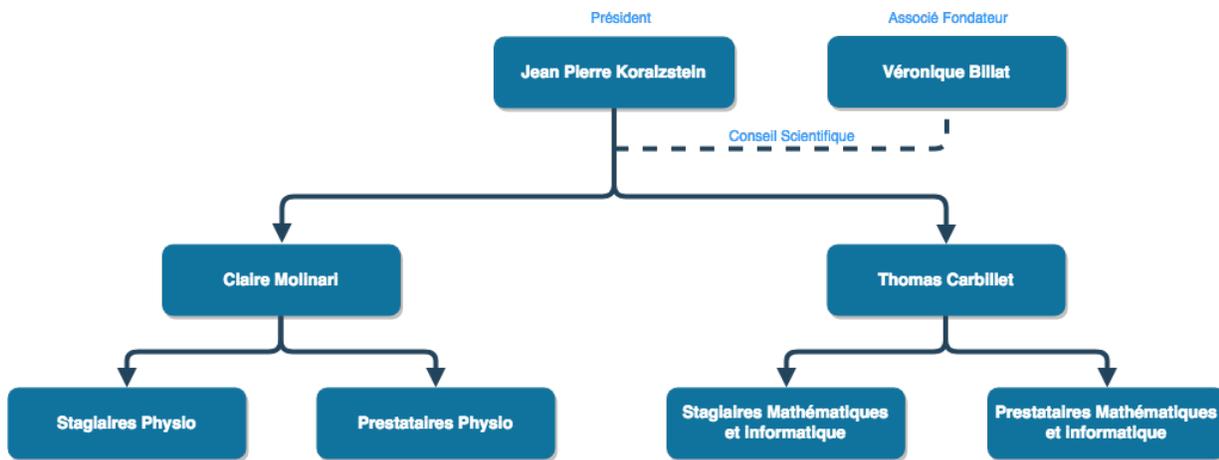


FIGURE I.9 – Organigramme de l'entreprise *BillaTraining*[®] au mois d'Août 2018

La figure I.10 présente le logo de l'entreprise *BillaTraining*[®]

Comme évoqué plus tôt dans le rapport, l'offre *BillaTraining*[®] se décompose en 2 grandes familles, le B2B et le B2C. Nous pouvons voir le détail de cette offre sur la figure I.11.

³société commerciale fixée par les articles L.227-1 à L.227-20 et L.244-1 à L.244-4 du code de commerce français

⁴la Convention Industrielle de Formation par la Recherche a été créée en 1981 par l'Association Nationale de la Recherche Technique (ANRT) et est financée par le Ministère de l'Enseignement Supérieur et de la Recherche. L'objectif de ces conventions est d'établir un contrat de recherche entre un laboratoire, un étudiant en première année de doctorat et une entreprise dans laquelle l'étudiant travaillera et mènera ses recherches.



FIGURE I.10 – Logo de l’entreprise *BillaTraining*[®]

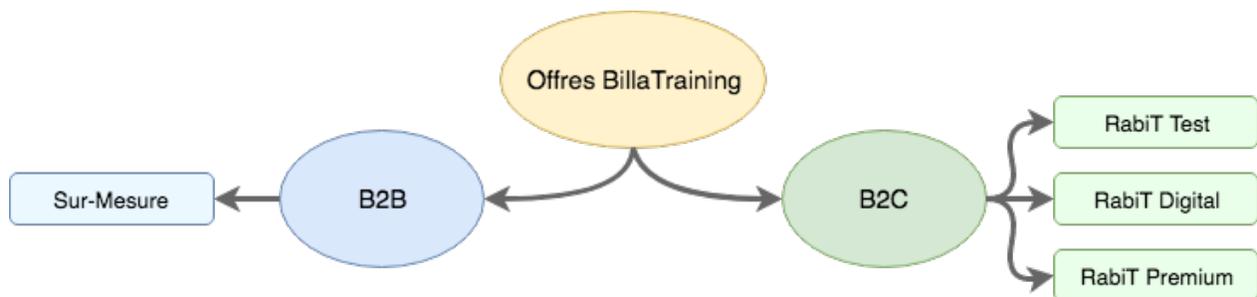


FIGURE I.11 – Les offres de l’entreprise *BillaTraining*[®]

Offre B2C

Le marché cible de l’entreprise pour le B2C sont les CSP+⁵ dans la tranche d’âge 30-70 ans, indifféremment Femmes et Hommes. Sur la partie droite de la figure I.11 on distingue 3 offres B2C. La première, le RabiT Test, est un audit physiologique. Les clients particuliers de l’entreprise *BillaTraining*[®] se déplacent à l’institut *BillaTraining*[®] et effectuent un test de RabiT[®] (voir section 3.2 pour le détail du test de RabiT[®]) dans les conditions réelles, à savoir, un chemin, une route ou encore une piste de stade d’athlétisme mais en aucun cas un tapis roulant. Les clients sont équipés de matériels scientifiques de pointe (voir section 5) permettant d’analyser leur effort et de leur livrer un bilan personnalisé grâce au radar énergétique (analyse des points forts et des points faibles - voir section 3.3), $\dot{V}O_{2max}$, masse grasse et autres données physiologiques. Cette offre permet d’avoir un état de la forme actuelle du coureur. La seconde offre est le RabiT Digital qui comprend l’offre RabiT test à laquelle s’ajoute un entraînement sur mesure d’une durée de trois mois (voir section 3.4). Cette offre

⁵signe désignant les Catégories Socio-Professionnelles favorisées. Cette catégorie comprend les chefs d’entreprise, les artisans et commerçants, les cadres, les professions intellectuelles supérieures et les professions intermédiaires.

a été conçue pour la préparation d'une course sur route (du 10 km au marathon) ou sur chemin (trails). Finalement, l'offre RabiT Premium reprend le RabiT Digital augmenté d'un suivi personnalisé par un expert en physiologie *BillaTraining*[®] et d'un plan de course pour le jour J appelé Signature de Vitesse[®] (voir section 3.5).

Offre B2B

Le segment B2B consiste en la collaboration avec des entreprises soucieuses de développer le bien être de leurs salariés, de valider un produit (vélos, chaussures, skis, ...) ou de créer une offre en marque blanche utilisant le savoir faire *BillaTraining*[®]. L'offre B2B ne fait pas directement l'objet de cette thèse car elle fait l'objet d'études spécifiques et/ou confidentielles. Nous la présentons ici à titre informatif.

Par ailleurs, l'entreprise *BillaTraining*[®] a reçu diverses distinctions et aides lors de ces 3 dernières années. *BillaTraining*[®] est Jeune Entreprise Innovante⁶ depuis 2015, a remporté le prix DIGITEO⁷ en 2015, a reçu une aide de BPI France⁸ de 35.000 € en 2016 pour la validation de la méthode *BillaTraining*[®], et est accréditée Crédit Impôt Innovation⁹ depuis 2018. Autant d'indicateurs d'innovation permettant à l'entreprise *BillaTraining*[®] de faciliter le transfert entre recherche et grand public.

La suite de cette partie est consacrée à l'explication des quatre étapes de la méthode *BillaTraining*[®] : le test de RabiT[®], le radar énergétique, l'entraînement sur mesure et la signature de vitesse[®]

3.2 Le test de RabiT[®]

La première étape de la méthode *BillaTraining*[®] est le test de RabiT[®] (acronyme de *Running Advisor BillaTraining*). Il s'agit d'un audit physiologique d'une trentaine de minutes permettant de dresser un bilan du coureur à pied. Ainsi, nous détectons les points forts et points faibles physiques et physiologiques du sportif. Ce test a été créé par le Professeur

⁶<https://www.service-public.fr/professionnels-entreprises/vosdroits/F31188>

⁷<http://www.digiteo.fr/>

⁸<https://www.bpifrance.fr/>

⁹<https://www.entreprises.gouv.fr/politique-et-enjeux/credit-impot-innovation>



FIGURE I.12 – La méthode *BillaTraining*[®] - Etape 1 - Le test de RabiT[®]

Véronique Billat en 2013 dans le but de regrouper les informations des tests en laboratoire et des tests de terrain (vameval, test de Cooper, test de Léger, ...). La figure I.13 présente le test de RabiT[®] dans son intégralité. Il est question de courir en suivant sa perception et

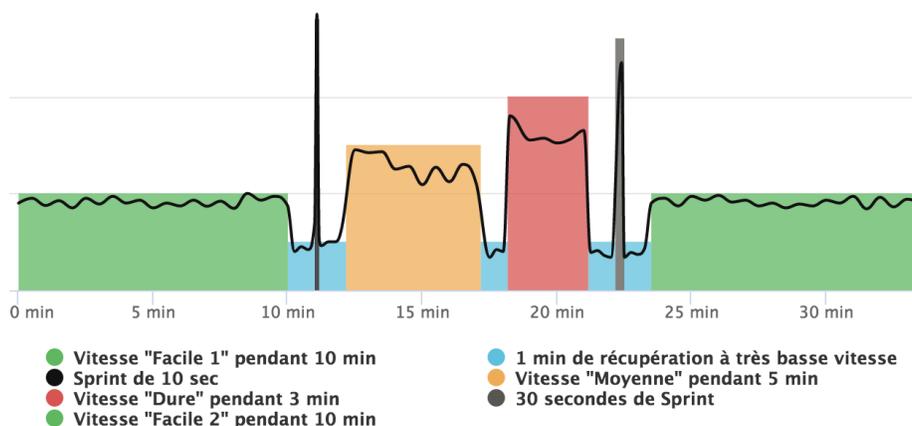


FIGURE I.13 – Test de RabiT[®]

d'éviter de s'imposer des allures constantes. Il se décompose en onze paliers comme suit :

1. 1er palier, courir à sensations "faciles" pendant 10 minutes.
2. 2ème palier, 1 minute de repos à très faible intensité ou marche.
3. 3ème palier, sprinter jusqu'à atteindre sa vitesse maximale pendant au moins 10 secondes.
4. 4ème palier, 1 minute de repos à très faible intensité ou marche.
5. 5ème palier, courir à sensations "moyennes" pendant 10 minutes.
6. 6ème palier, 1 minute de repos à très faible intensité ou marche.
7. 7ème palier, courir à sensations "dures" pendant 3 minutes.

8. 8ème palier, 1 minute de repos à très faible intensité ou marche.
9. 9ème palier, sprinter pendant 30 secondes.
10. 10ème palier, 1 minute de repos à très faible intensité ou marche.
11. 11ème palier, courir à sensations "faciles" pendant 10 minutes.

Finalement, la courbe noire présente sur la figure [I.13](#) donne un exemple de la vitesse en fonction du temps pour un test de RabiT[®] déjà réalisé. L'avantage de ce test est qu'il est adaptable à tous les sports d'endurance. La durée des paliers sera alors modifiée sans en impacter l'ordre.

La première contribution majeure de cette thèse est l'automatisation du traitement du test de RabiT[®]. En effet, le traitement encore manuel rend la diffusion de la méthode *BillaTraining*[®] peu applicable au grand nombre de coureurs. C'est la première étape en vue de la création d'une application web autonome (ne nécessitant aucun traitement manuel). Nous développerons largement l'automatisation du test de RabiT[®] dans le [chapitre II](#). Finalement, il est à noter que le test de RabiT[®] peut se faire de deux façons. Soit en présentiel avec un physiologiste qui équipe le coureur de multiples capteurs (voir section [5](#) pour la liste des principaux capteurs utilisés par l'entreprise *BillaTraining*[®]). Les paramètres physiologiques et les qualités physiques sont alors directement mesurés. Soit à distance avec une simple montre de sport ou un téléphone portable. Grâce à la base de données issue de 30 ans de mesures, *BillaTraining*[®] peut estimer les paramètres physiologiques et les qualités physiques uniquement à partir des séries temporelles de vitesse et de fréquence cardiaque.

3.3 Le Radar énergétique

A partir des résultats du test de RabiT[®], le Professeur Véronique Billat a développé le Radar Énergétique. Cet outil permet visuellement et instantanément d'évaluer les points forts et les points à développer pour un coureur. Le Radar énergétique se décompose en 4 facteurs d'évaluation de la performance :



FIGURE I.14 – La méthode *BillaTraining*[®] - Etape 2 - Le Radar Energétique

- La force musculaire est la force développée par les muscles à chaque appui au sol. C'est un facteur très important pour atteindre une vitesse maximale élevée mais également le facteur primordial pour améliorer, solliciter et maintenir $\dot{V}O_2$ max sur des durées importantes.
- L'efficacité cardiaque correspond à la capacité du coeur d'envoyer du sang aux muscles.
- La tolérance à l'acidose est la capacité à répéter les efforts intenses et à tolérer l'accumulation d'acide lactique.
- La perception regroupe les fonctions cognitives qui se rapportent aux prises de décisions. En d'autres termes, la capacité à connaître ses allures, d'accélérer et de ralentir en fonction des sensations ressenties. Ce facteur représente la capacité à gérer l'effort.

La figure I.15 est un exemple de Radar énergétique où chacun des quatre facteurs de la performance est noté sur 10. Un des avantages du Radar énergétique réside dans la possibilité

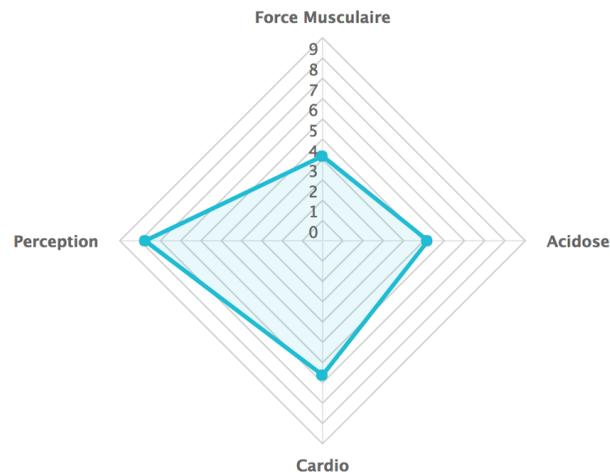


FIGURE I.15 – Exemple de Radar énergétique. Les 4 facteurs de la performance (force musculaire, efficacité cardiaque, tolérance à l'acidose et perception) sont notés sur 10

de discriminer les coureurs à $\dot{V}O_2$ max équivalent. Les résultats du test de RabiT[®] combinés à l'interprétation du Radar énergétique permettent la création d'un plan d'entraînement sur mesure.

3.4 L'entraînement sur mesure



FIGURE I.16 – La méthode *BillaTraining*[®] - Etape 3 - l'entraînement sur mesure

La troisième étape de la méthode *BillaTraining*[®] est un entraînement sur mesure. Créé à partir des résultats du test de RabiT[®] et de l'interprétation du Radar énergétique, le programme d'entraînement, d'une durée de trois mois, permet d'améliorer les facteurs de la performance du coureur à raison de deux séances d'entraînement spécifiques par semaine, toutes deux inférieures à 1 heure et une séance longue en fin de semaine d'une durée comprise entre 1 et 2 heures. Toutes les séances d'entraînement spécifiques sont classées suivant un des quatre facteurs de la performance (Force, Cardio, Acidose, Perception) et du niveau du sportif ($\dot{V}O_2$ max). La figure I.17 donne un exemple de séance d'entraînement spécifique en perception.

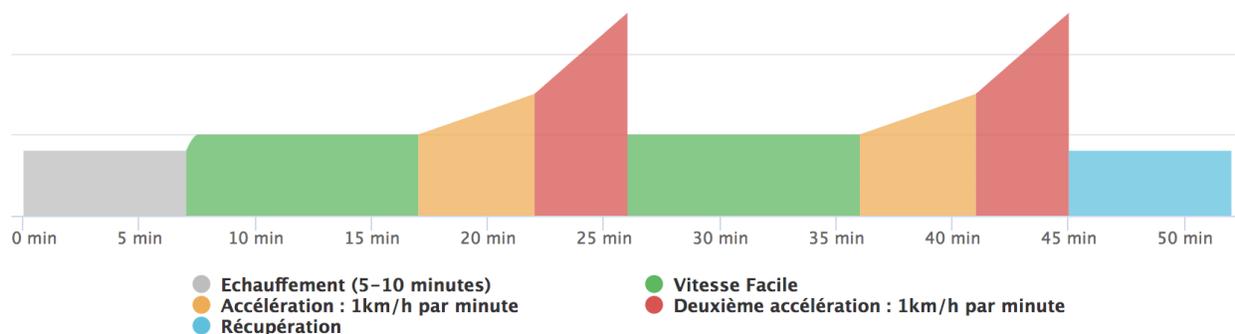


FIGURE I.17 – Exemple de séance d'entraînement spécifique *BillaTraining*[®] en Perception

La particularité de cet entraînement est la variation d'allure/d'intensité au sein d'une même séance. En effet, l'entraînement en endurance est connu pour induire une augmentation de la consommation maximale d'oxygène ($\dot{V}O_2 \text{ max}$), qui est une mesure globale de la capacité conjointe des systèmes pulmonaire, cardiovasculaire et musculaire à extraire, transporter et utiliser l'oxygène. Cette augmentation de $\dot{V}O_2 \text{ max}$ reflète, entre autres, d'une augmentation du nombre et de l'activité des mitochondries, le lieu de la respiration cellulaire [134, 97]. Les recherches du Professeur Billat ont conduit à déterminer que l'entraînement en endurance doit être basé sur un mélange d'exercices de faible intensité et de volume élevé et d'exercices intermittents de haute intensité (au-dessus de la vitesse en aérobie maximale) et/ou des exercices de force. Ces combinaisons permettent d'optimiser la performance en aérobie (c'est-à-dire $\dot{V}O_2 \text{ max}$) [179, 9]. Étant donné que les augmentations des paramètres mitochondriaux coïncident avec une augmentation de $\dot{V}O_2 \text{ max}$ dans l'ensemble de l'organisme [77], on peut supposer qu'une combinaison d'entraînements de faible intensité et de haute intensité augmente la biogenèse des mitochondries. L'étude de cas du cycliste centenaire Robert Marchand a démontré que les effets de l'entraînement *BillaTraining*[®] à base de variation d'allures permettait d'augmenter le $\dot{V}O_2 \text{ max}$ ainsi que les performances sportives [33].

3.5 La Signature de Vitesse[®]



FIGURE I.18 – La méthode *BillaTraining*[®] - Etape 4 - La Signature de Vitesse[®]

La dernière étape de la méthode *BillaTraining*[®] est la signature de vitesse[®]. Il est question de donner une stratégie de course personnalisée au coureur pour le jour de sa course grâce aux résultats de son entraînement. La première version de la signature de vitesse[®] date de 2015. Sa construction repose sur l'expérience terrain et les connaissances en physiologie

du Professeur Véronique Billat. La figure I.19 donne un exemple de signature de vitesse[®] pour un client de l'entreprise *BillaTraining*[®] sur marathon. L'idée est de découper la course

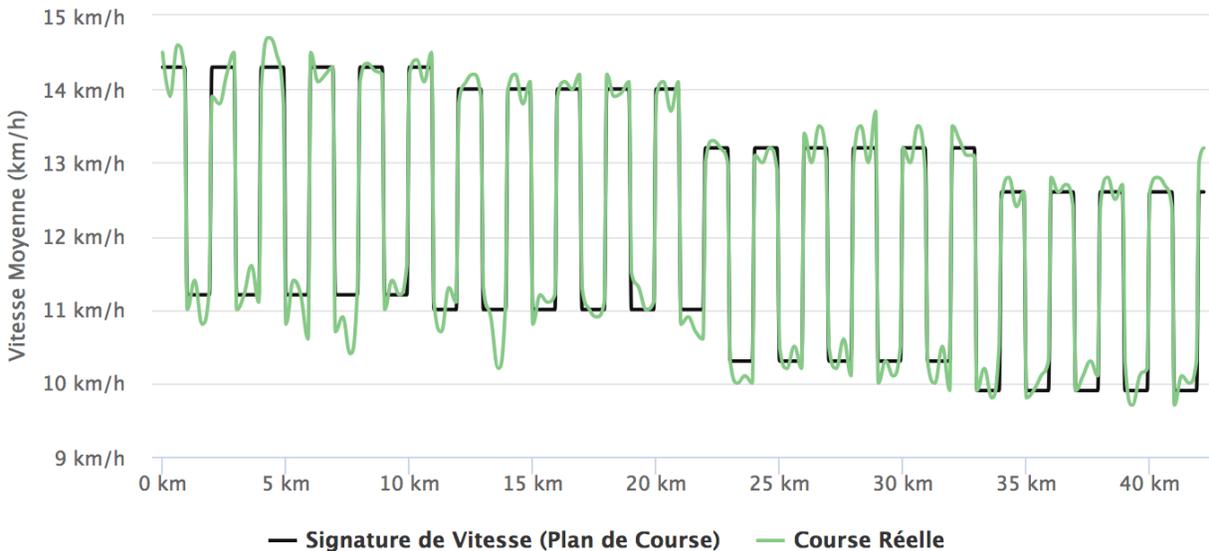


FIGURE I.19 – Exemple de signature de vitesse[®] pour une cliente de *BillaTraining*[®] sur marathon

en quatre partie grâce à quatre vitesses moyennes cibles constantes selon une fonction en escalier. Cependant, au lieu de courir à allure constante sur chacune des quatre parties, il est demandé au coureur de varier sa vitesse tous les kilomètres en fonction des résultats du test de RabiT[®], du Radar énergétique et des trois mois d'entraînement [26].

La deuxième contribution majeure de cette thèse est d'essayer d'apporter un cadre mathématique à la signature de vitesse[®] qui repose actuellement sur l'expérience du Professeur Billat. Le chapitre III propose un raisonnement mathématique descriptif posant les bases d'une nouvelle version de la signature de vitesse[®].

4 Etat des lieux de la concurrence

L'objet de cette partie est de positionner l'entreprise *BillaTraining*[®] par rapport à la concurrence afin d'en illustrer les différences. Les concurrents de l'entreprise *BillaTraining*[®] sont ceux qui, à partir des données collectées par des objets connectés propriétaires ou non, fournissent des recommandations/informations intégrées à ces mêmes objets connectés ou au sein

d'un service différents (type plateforme en ligne ou seulement API¹⁰). A ce jour, de nombreuses entreprises, développent des algorithmes d'interprétation de données issues d'objets connectés pour des usages très variés. Le tableau I.1 présente les principaux concurrents de l'entreprise *BillaTraining*[®], leurs offres, avantages et limites. Au sens de Porter [142], l'entre-

| | Objets connectés | Algorithmes d'interprétation de données |
|-------------|---|--|
| Entreprises | <ul style="list-style-type: none"> • <i>Garmin</i>[®] • <i>Polar</i>[®] • <i>Suunto</i>[®] | <ul style="list-style-type: none"> • Strava • Peak Performance • Runkeeper • Runtastic |
| Périmètre | International | International |
| Offres | <ul style="list-style-type: none"> • Cardio-fréquence mètre GPS • Restitution des données de vitesse, fréquence cardiaque, cadence, altitude, ... • Estimation de $\dot{V}O_2$max et indice de récupération | <ul style="list-style-type: none"> • Collecte et traitement des données issues d'objets connectés ou de coaches. • Hub de services • Super intégrateur indépendant du matériel |
| Avantages | <ul style="list-style-type: none"> • Qualité des données enregistrées. • Marketing et Communication | <ul style="list-style-type: none"> • Ergonomie des plateformes utilisateurs • Traitement de données multi-appareils et multi-sports • Véritable Réseau Social • Marketing et Communication |
| Limites | <ul style="list-style-type: none"> • Conseils généraux non personnalisés • Non utilisation des données stockées • Valeur d'usage qui peut être limitée | <ul style="list-style-type: none"> • Pas d'analyse scientifique • Programme d'entraînement grand public |

TABLE I.1 – Principaux concurrents de l'entreprise *BillaTraining*[®] : périmètre, offre, avantages et limites

¹⁰En informatique, une interface de programmation applicative (souvent désignée par le terme API pour application programming interface) est un ensemble normalisé de classes, de méthodes ou de fonctions qui sert de façade par laquelle un logiciel offre des services à d'autres logiciels. https://fr.wikipedia.org/wiki/Interface_de_programmation

prise *BillaTraining*[®] adopte la stratégie de différenciation pour ces services B2C en essayant de proposer un service innovant et en marge de la concurrence. Pour illustrer ce propos, voici les deux avantages compétitifs identifiés par l'entreprise *BillaTraining*[®] :

1. La légitimité scientifique s'appuie sur 30 ans de recherche et de validation opérationnelle par le terrain. Collecte d'une masse élevée de données sur le terrain, permettant le développement d'algorithmes robustes et fiables.
2. L'analyse personnalisée des données clients nous permet de proposer un programme d'entraînement sur mesure qui se corrige et se met à jour à chaque fois que le coureur termine une séance.

5 Le matériel à disposition

L'entreprise est munie d'un laboratoire mobile, équipé de dispositifs permettant d'évaluer la forme physique des sportifs et constitué de multiples capteurs permettant l'analyse des données biomécaniques et bioénergétiques des sportifs. La présente section vise à détailler le matériel présent dans le laboratoire de l'entreprise. En effet, la qualité de l'acquisition de données sur le terrain est cruciale car toutes nos recherches et découvertes découlent de celle-ci. L'enjeu a été de trouver du matériel portable afin d'effectuer toute nos mesures sur le terrain.

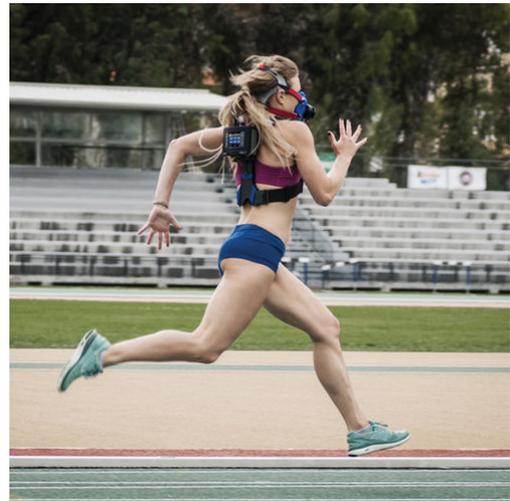
5.1 *Cosmed*[®] - K5 - Analyseur de gaz

L'entreprise *Cosmed*^{®11} est spécialisée dans la fabrication de matériel cardio-pulmonaire, métabolique et de composition corporelle depuis 1980. L'entreprise est basée en Italie avec son siège social à Rome et distribue ses produits à travers le monde. Le laboratoire *BillaTraining*[®] utilise le K5 de la gamme "cardio-pulmonaire". La figure I.20 montre deux photos du K5. Il s'agit d'un analyseur des échanges gazeux respiratoire qui permet de mesurer notamment le

¹¹<http://www.cosmed.com>



(a) Photo d'un K5 provenant du site de *Cosmed*[®]



(b) Photo d'un K5 porté par un coureur sur son terrain d'entraînement. La photo provient du site de *Cosmed*[®]

FIGURE I.20 – Photos de l'unité centrale d'un K5. (a) correspond à la photo d'un K5 provenant du site de *Cosmed*[®]. (b) correspond à la photo d'un K5 porté par un coureur sur son terrain d'entraînement. La photo provient du site de *Cosmed*[®]

$\dot{V}O_2$ ¹² et le $\dot{V}CO_2$ ¹³ d'un sportif en cycle à cycle ou grâce à une chambre de mélange. Le cycle à cycle renvoie une valeur de $\dot{V}O_2$ à chaque cycle respiratoire (un cycle correspond à une inspiration suivi d'une expiration) en calculant la différence entre la quantité d'oxygène inspirée et la quantité d'oxygène expirée. De même pour le $\dot{V}CO_2$ avec la quantité de CO_2 expirée. La chambre de mélange donne une valeur moyenne du $\dot{V}O_2$ et du $\dot{V}CO_2$ sur une période de trente secondes. L'avantage du cycle à cycle par rapport à la chambre de mélange est pour les efforts courts. En effet, si l'on prend l'exemple d'un sprinter courant le 100 m en 10 secondes, la valeur moyenne renvoyée par la chambre des mélanges sur trente secondes sera erronée. L'entreprise *BillaTraining*[®] se sert du K5 pour mesurer notamment le $\dot{V}O_2$ max et le $\dot{V}CO_2$ max des coureurs à pied. Il est à noter que le K5 est l'appareil de référence pour mesurer les échanges gazeux sur le terrain [117, 137, 12].

¹²Consommation d'oxygène. Il s'agit du volume d'oxygène qui peut être consommé par unité de temps lors d'un exercice physique [25]

¹³Production de dioxyde de carbone. Il s'agit du volume de dioxyde de carbone expiré après avoir transporté de l'oxygène à travers le corps [25]

5.2 *Garmin*[®] - 920XT - Montre de sport

L'entreprise *Garmin*^{®14} est spécialisée dans la technologie GPS pour l'automobile, l'aviation le nautisme et les activités sportives. Le siège social est basé à Olathe dans l'état du Kansas aux Etats-Unis. *Garmin*[®] distribue ses produits à travers le monde. Le laboratoire *BillaTraining*[®] utilise le modèle 920XT. La figure I.21 présente deux photos de la montre de sport. La montre 920XT permet d'enregistrer des données mécaniques (cadence de foulée,



(a) Photo de la montre 920XT provenant du site de *Garmin*[®]



(b) Photo de la montre 920XT portée par un coureur sur son terrain d'entraînement. La photo provient du site de *Garmin*[®]

FIGURE I.21 – Photos de la montre de sport 920XT. (a) correspond à la photo de la montre 920XT provenant du site de *Garmin*[®]. (b) correspond à la photo de la montre 920XT portée par un coureur sur son terrain d'entraînement. La photo provient du site de *Garmin*[®]

oscillation verticale du centre de masse, temps d'appui au sol, vitesse et accélération) et la fréquence cardiaque sous condition de porter une ceinture pectorale adaptée. Une estimation du $\dot{V}O_2$ max et du temps de récupération sont aussi fournies. L'entreprise *BillaTraining*[®] se sert de la montre 920XT pour les données mécaniques et la fréquence cardiaque. La montre 920XT est équipée des technologies GPS-GLONASS¹⁵, d'un baromètre et d'un accéléromètre. Plusieurs études scientifiques supportent le choix de cette montre de sport pour des mesures

¹⁴<https://www.garmin.com>

¹⁵Le GPS est un réseau de 31 satellites développé par les américains et la technologie GLONASS est un réseau de 24 satellites développé par les russes. <https://beebom.com/what-is-glonass-and-how-it-is-different-from-gps/>

hors du laboratoire [43, 109, 136].

5.3 *Physioflow*[®] - PF07 Enduro - Débit cardiaque

L'entreprise *Physioflow*^{®16} est spécialisée dans la mesure précise et continue du débit cardiaque. L'entreprise possède deux sièges sociaux. Le premier, historique, est basé à Petit Ebersviller en France. Le second est basé à Bristol au Etats-Unis. *Physioflow*[®] distribue ses produits à travers le monde. Le laboratoire *BillaTraining*[®] utilise le modèle PF07 Enduro. La figure I.22 présente la photo du dispositif de mesure du débit cardiaque PF07 de la marque *Physioflow*[®]. Le débit cardiaque est mesuré par impédancemétrie grâce à des électrodes pla-

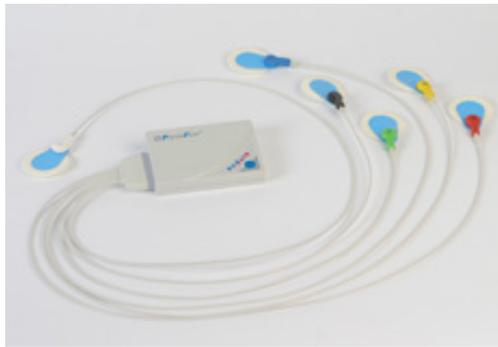


FIGURE I.22 – Photo de l'appareil *Physioflow*[®] PF07

cées sur le buste du sportif. Le placement des électrodes à une importance clé pour la précision des mesures. Le débit cardiaque (\dot{Q}_c) est la quantité de sang pompée par le coeur toutes les minutes. Il est défini par l'équation I.21

$$\dot{Q}_c = F_c \cdot VES \quad (\text{I.21})$$

où F_c est la fréquence cardiaque en battements par minute et VES le volume d'éjection systolique en millilitres. Plusieurs études scientifiques valident le choix de ce dispositif pour des mesures non intrusives sur des sujets en bonne santé [72, 160, 115].

¹⁶<https://www.physioflow.com/>

5.4 Centaure Metrix® - Locometrix - Accéléromètre triaxial

La société *Centaure Metrix*¹⁷ est spécialisée dans l'analyse de la marche et de la course à pied grâce à un accéléromètre triaxial avec une acquisition de 100Hz. Le siège social est basé à Evry en France. L'entreprise distribue ses produits à travers le monde. Le laboratoire *BillaTraining*[®] utilise le modèle Locometrix. La figure I.23 présente deux photos du dispositif. Cet appareil se fixe sur la région lombaire à l'aide d'une ceinture et enregistre les mouve-



(a) Photo du Locometrix provenant du site de *Centaure Metrix*[®]



(b) Photo du Locometrix portée par une sportive. La photo provient du site de OpenPR

FIGURE I.23 – Photos de l'appareil Locometrix. (a) correspond à la photo du Locometrix provenant du site de *Centaure Metrix*[®]. (b) correspond à la photo du Locometrix portée par une sportive. La photo provient du site de OpenPR

ments du centre de gravité du coureur. Les deux valeurs restituées par l'appareil dont se sert l'entreprise *BillaTraining*[®] sont l'accélération horizontale du coureur et l'amortissement des semelles (onde de choc de hautes fréquences). Plusieurs études scientifiques supportent le choix de ce dispositif pour des mesures hors laboratoire [11, 68].

Bilan

Nous avons pu identifier six travaux scientifiques, antérieurs à cette thèse et traitant de stratégie optimale en course à pied. Joseph Keller, le pionnier, propose un modèle en trois phases. La première est une accélération maximale pendant quelques secondes puis le maintien d'une vitesse constante pour finir avec une vitesse décroissante dans les derniers instants

¹⁷<http://www.centaure-metrix.com/index.html>

de la course. Viennent ensuite Mathis puis Woodside qui introduisent un facteur de fatigue non pris en compte par Keller. Behncke présente ensuite une théorie basée sur un modèle hydraulique représentant les trois sources d'énergie (créatine phosphate, acide lactique et oxydation des lipides et glucides). Whitt part ensuite des fondamentaux de la théorie de Keller en traitant la consommation d'oxygène comme un processus stochastique. Finalement, Aftalion et Bonnans prennent en compte la recréation d'énergie provenant des travaux de Morton et Billat [130]. Nous avons vu que cette thèse se place en marge de ces recherches par la volonté de proposer une stratégie de course personnalisée, basée sur les capacités physiques d'un coureur, réalisable et sur des distances longues (10 km au Marathon).

Dans un contexte concurrentiel basé sur les objets connectés et l'amas de données, cette thèse est au coeur de l'entreprise *BillaTraining*[®] qui souhaite démocratiser une méthode d'entraînement et un plan de course sur mesure là où se croisent physiologie, mathématiques et informatique.

Finalement, dans ce premier chapitre, nous avons souligné deux contributions majeures auxquelles nous en rajoutons deux nouvelles, à savoir :

- L'automatisation du traitement du test de RabiT[®] (voir [chapitre II](#)).
- La description d'une façon de courir optimale. Ceci ouvre la voie vers une nouvelle stratégie de course personnalisée (voir [chapitre III](#)).
- La validation de l'hypothèse selon laquelle l'être humain est capable de courir à accélération constante (voir [chapitre IV](#)).
- La création d'une plateforme web intégrant les trois premières étapes de la méthode *BillaTraining*[®] (voir [chapitre V](#)).

Chapitre II

Première contribution : La segmentation du test de RabiT[®]

Dans ce deuxième chapitre, je vais vous présenter toutes les étapes, de manière plus ou moins détaillée, qui m'ont permis de développer un micro service ayant pour but la segmentation du test de RabiT[®]. De l'idée à la réalisation, ce chapitre se décompose en quatre étapes. La première pose les problématiques et enjeux propres à la segmentation du test de RabiT[®]. La deuxième présente la théorie de détection de ruptures. La troisième étape est l'application de la théorie sur la détection de ruptures au test de RabiT[®]. La quatrième et dernière étape présente la mise en production de la solution retenue.

Sommaire

| | | |
|----------|---|-----------|
| 1 | Problématiques et enjeux | 59 |
| 2 | La détection de ruptures | 61 |
| 2.1 | Le problème théorique de détection de ruptures | 62 |
| 2.2 | Le choix de la méthode de détection de ruptures | 64 |
| 3 | Application au test de RabiT[®] | 69 |
| 3.1 | Construction des solutions | 70 |
| 3.2 | Sélection et filtrage de la meilleure solution | 71 |
| 4 | Création d'un micro-service web | 74 |

| | | |
|------------------------|--|-----------|
| 4.1 | Architectures monolithique VS. micro service | 74 |
| 4.2 | Application Programming Interface (API) | 77 |
| 4.3 | Création d'un micro service web pour segmenter le test de RabiT [®] | 79 |
| Bilan | | 81 |

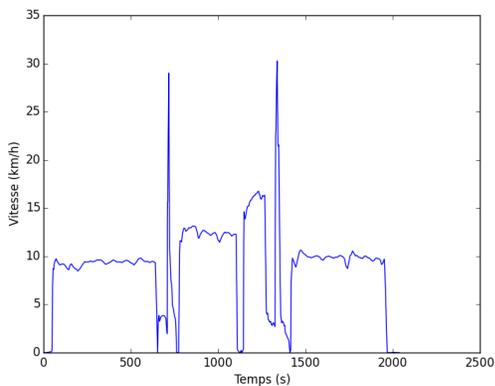
1 Problématiques et enjeux

Le premier verrou scientifique auquel j'ai dû faire face lors de ma thèse est l'automatisation du traitement du test de RabiT[®]. En effet, comme expliqué dans la section 3.2 du chapitre I, le test de RabiT[®] était encore traité à la main par les physiologistes BillaTraining. Cette étape d'analyse manuelle était chronophage, rébarbative et susceptible d'induire des erreurs par son manque d'objectivité. L'objectif à terme est d'automatiser toute la méthode BillaTraining et le traitement manuel du test de RabiT[®] présentait un goulot d'étranglement. Nous cherchons une méthode de détection des phases du test de RabiT[®] selon les quatre critères suivants :

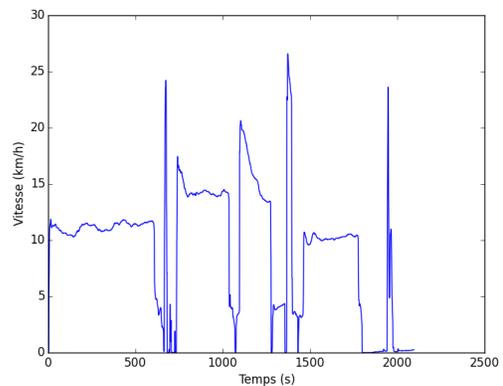
- **Rapide** - Le développement des services en ligne dont la qualité a considérablement augmentée ces dernières années impose le choix d'une solution quasi instantanée afin de ne pas faire attendre l'utilisateur final.
- **Précis** - Le test de RabiT[®] est à la base de la méthode BillaTraining. Nous cherchons une solution qui permet d'identifier tous les paliers du test de RabiT[®] dans leur intégralité.
- **Robuste** - Nous avons remarqué à quel point il pouvait être difficile de respecter tous les temps imposés par le test de RabiT[®] à la lettre. C'est pourquoi, nous cherchons une solution tolérant de légers écarts dans les temps des paliers du test.
- **Disponible** - L'application web que nous réaliserons devra permettre d'intégrer la solution retenue en production. Nous chercherons un moyen efficace de rendre la segmentation du test de RabiT[®] disponible en continu (24 heures sur 24 et 7 jours sur 7).

Pour un rappel des consignes du test de RabiT[®], veuillez vous référer à la section 3.2 du chapitre I.

Avant de présenter en détail la solution retenue pour la détection des phases du test de RabiT[®], je vais vous présenter deux cas typiques réalisés par deux clients de l'entreprise BillaTraining. Le premier est un coureur qui respecte au plus près le test de RabiT[®]. Le second est un coureur qui oublie de couper sa montre à la fin du test. La figure II.1 présente les données brutes de vitesse en fonction du temps pour respectivement le premier (figure II.1b) et le deuxième cas (figure II.1a). Les deux tests présents sur la figure II.1 me serviront de



(a) Test de RabiT[®] réalisé par un client de l'entreprise BillaTraining ayant respecté les consignes du test. Vitesse (km/h) en fonction du temps (s)



(b) Test de RabiT[®] réalisé par un client de l'entreprise BillaTraining ayant oublié de couper sa montre. Vitesse (km/h) en fonction du temps (s)

FIGURE II.1 – Deux tests de RabiT[®] réalisés par deux clients de l'entreprise BillaTraining. Vitesse (km/h) en fonction du temps (s)

fil rouge pour ce chapitre. Pour trouver et implémenter une solution de détection de ruptures dans le test de RabiT[®], j'ai eu accès à une base de données de cinquante fichiers contenant un signal de vitesse en fonction du temps. Quarante de ces fichiers sont des tests de RabiT[®] jugés comme "valide" (qui auraient été acceptés par un physiologiste BillaTraining) et dix comme "non-valide" (qui auraient été rejetés par un physiologiste BillaTraining). Globalement, nous devons traiter des séries temporelles discrètes de vitesse avec un échantillonnage pouvant varier de 1 et 3 secondes, ce qui correspond à un nombre de points compris entre 700 et 2500 environ.

Nous pouvons dès lors formuler le problème qui nous accompagnera dans tout ce chapitre.

L'objectif est de détecter les onze segments du test de RabiT[®] sachant la durée théorique pour chaque segment (facile 1 (600 s), repos (60 s), sprint (10 s), repos (60 s), moyen (300 s), repos (60 s), dur (180 s), repos (60 s), trente secondes (30 s), repos (60 s), facile 2 (600 s)) en essayant de reproduire le degré d'acceptation d'un physiologiste BillaTraining.

Nous sommes face à un problème de détection de ruptures rétrospectives car nous avons l'intégralité du signal temporel de vitesse. Dans les sections 2 et 3, je présenterai la méthode mathématique et algorithmique retenue pour la détections de ruptures des phases du test de RabiT[®].

2 La détection de ruptures

Etudiant les différentes possibilités de détection de ruptures, j'ai eu l'occasion de découvrir Rebecca Killick, chercheur à l'université de Lancaster au Royaume-Uni. Après plusieurs lectures, ses deux articles [108, 107] et son site internet [106] semblaient répondre à notre problématique de segmentation du test de RabiT[®]. De plus, dans un souci d'intégration à notre futur application web, l'auteur propose un package développé avec le langage de programmation R¹ testé et validé par la communauté mathématique et informatique.

De mes lectures, il apparaît que pour les applications les plus courantes, nous pouvons diviser le problème de détection de ruptures en deux grandes familles de méthode. La première est l'heuristique de segmentation binaire (Binary Segmentation(BS)). La deuxième est la programmation dynamique (Segment Neighbourhood (SN), Optimal Partitioning (OP) et Pruned Exact Linear Time (PELT), ...).

La suite de cette section est consacrée au choix de la méthode en vue de la segmentation du test de RabiT[®].

¹La documentation complète du package R se trouve à l'url : <https://cran.r-project.org/web/packages/changepoint/changepoint.pdf>

2.1 Le problème théorique de détection de ruptures

Commençons par poser le problème de détection de ruptures de manière globale, c'est-à-dire, indépendamment de la méthode utilisée. Il s'agit de relier des instants de ruptures candidats à un modèle mathématique. Les instants de ruptures sont les points dans une série de données où l'on observe un changement dans les propriétés statistiques. On étudie généralement les changements dans la distribution de la série de données suivant trois techniques différentes. La première se conduit par l'étude des changements dans la moyenne. La seconde par l'étude des changements dans la variance. La troisième et dernière par l'étude des changements dans la moyenne et la variance combinées. D'autres techniques existent comme la densité spectrale de puissance mais n'ont pas été considérées dans cette thèse.

Comme nous pouvons le voir sur la figure [II.1](#) représentant des données brutes de vitesse en fonction du temps de deux test de RabiT[®] en conditions réelles, il apparaît que les valeurs de vitesse oscillent autour de plusieurs valeurs moyennes correspondant aux paliers du test de RabiT[®]. La variance ne dépend pas des paliers, par conséquent, afin de trouver les instants de ruptures du test de RabiT[®], nous étudierons uniquement les changements dans la moyenne de la série temporelle de vitesse.

Définissons à présent, en termes mathématiques, la série temporelle de vitesse, v , que nous allons utiliser tout au long de cette section. Il est possible de retrouver cette définition dans l'article [\[108\]](#).

Soit $v_{1:n} = (v_1, \dots, v_n)$ un signal de vitesse discret. Notre modèle aura m instants de ruptures aux emplacements $\tau_{1:m} = (\tau_1, \dots, \tau_m)$ où chaque τ_i est un entier entre 1 et $n - 1$ inclus. On introduit d'ores et déjà deux instants de ruptures $\tau_0 = 0$ et $\tau_{m+1} = n$ qui correspondent respectivement au début et à la fin du signal. Les instants de ruptures sont ordonnés tels que $\tau_i < \tau_j$ si et seulement si $i < j$. Le i -ème segment contiendra les données $v_{(\tau_{i-1}+1):\tau_i} = (v_{(\tau_{i-1}+1)}, \dots, v_{\tau_i})$. Pour tenter de faire comprendre au lecteur la théorie de détection de ruptures, prenons l'exemple d'un signal ayant un seul instant de rupture que nous appelons τ_1 . Il convient alors d'effectuer le test du rapport de vraisemblance logarithmique qui sera notre fonction de coût.

Dans la littérature, la fonction de coût est communément une fonction de vraisemblance logarithmique négative que nous cherchons à minimiser [90, 44]. C'est aussi la fonction de coût retenue par Rebecca Killick [108, 107] et implémentée par la même occasion dans le package R "changeoint" que nous utiliserons dans la section 3. D'autres fonctions existent comme la perte quadratique ou les sommes cumulées. De telles fonctions n'ont pas été étudiées dans cette thèse.

Revenons à notre test statistique où nous allons tester l'hypothèse nulle (H_0) face à l'hypothèse alternative (H_1) suivantes :

- H_0 : Il n'existe pas d'instant de rupture. Le maximum de vraisemblance logarithmique est $\log p(v_{1:n}|\hat{\theta})$ où $p(\bullet)$ est la fonction de densité de probabilité associée à la distribution de vitesse v et $\hat{\theta}$ est l'estimateur du maximum de vraisemblance pour la moyenne du signal [71].
- H_1 : Il existe un instant de rupture. La maximum de vraisemblance logarithmique devient $ML(\tau_1) = \log p(v_{1:\tau_1}|\hat{\theta}_1) + \log p(v_{(\tau_1+1):n}|\hat{\theta}_2)$ où $\hat{\theta}_1$ et $\hat{\theta}_2$ sont les estimateurs du maximum de vraisemblance pour les moyennes des deux segments séparés par τ_1 .

L'équation II.1 donne le test statistique pour détecter un instant de rupture.

$$LR = 2 \left[\max_{\tau_1} ML(\tau_1) - \log p(v_{1:n}|\hat{\theta}) \right] \quad (\text{II.1})$$

Il existe dès lors deux catégories de problèmes d'optimisation de recherche d'instant de ruptures. La première regroupe les problèmes contraints. Dans ce cas, on connaît le nombre d'instant de ruptures à l'avance. Le risque de sur-segmentation ou de sous-segmentation est alors élevé. La deuxième regroupe les problèmes pénalisés. Dans ce cas, on ne connaît pas le nombre d'instant de ruptures exacts. Afin d'éviter la sur-segmentation ou la sous-segmentation, on introduit le concept de pénalité. Notons c la pénalité que nous allons comparer à LR . Si $LR > c$, on rejette H_0 ce qui veut dire qu'un instant de rupture a été détecté en τ_1 . Sinon, on ne peut pas rejeter H_0 et aucun instant de rupture n'est détecté. Nous avons opté pour la deuxième approche puisque les interactions de l'être humain avec le dispositif d'enregis-

trement (montre de sport et téléphone portable principalement) ont une forte influence sur le signal de vitesse. En effet, si le coureur lance l'enregistrement des données au début de sa phase d'échauffement ou oublie de le couper à la fin du test de RabiT[®], cela induit un changement dans le nombre d'instant de ruptures.

Généralisons maintenant le problème d'optimisation pénalisée à de multiples instants de ruptures. L'idée reste la même que pour un seul instant de rupture. Il convient d'identifier les maxima de $ML(\tau_{1:m})$ pour toutes les combinaisons de $\tau_{1:m}$. L'équation II.2 définit le problème de détection pour de multiples instants de ruptures.

$$\sum_{i=1}^{m+1} [\mathcal{C}(v_{(\tau_{i-1}+1):\tau_i})] + \beta f(m) \quad (\text{II.2})$$

On cherche à minimiser l'équation II.2 pour identifier les instants de ruptures optimaux. \mathcal{C} est une fonction de coût (vraisemblance logarithmique négative) sur un segment candidat et $\beta f(m)$ correspond à la pénalité, qui est l'équivalent de c pour un problème avec de multiples instants de ruptures.

Le choix de la pénalité est encore sujet à discussion dans la littérature [4, 154, 75, 139, 35]. Cependant, au vue de notre problème, il semble judicieux d'opter pour la méthode CROPS (Changepoints Detection for a Range of Penalties) [83, 84] car elle permet de choisir une plage de valeurs de pénalités contrairement à d'autres méthodes.

2.2 Le choix de la méthode de détection de ruptures

Nous avons introduit, au début de cette section, quatre méthodes de détection de ruptures (Binary Segmentation (BS), Segment Neighbourhood (SN) Optimal Partitioning (OP) et Pruned Exact Linear Time (PELT)). L'objectif ici est d'expliquer brièvement les trois méthodes BS, SN et OP encore largement utilisées pour résoudre des problèmes de détection de ruptures puis de détailler la méthode PELT qui est la solution que nous avons retenue pour segmenter le test de RabiT[®]. Nous aborderons les avantages et les inconvénients de chaque méthode et nous justifierons le choix de la méthode PELT.

Binary Segmentation (BS)

De par son caractère pionnier, la méthode BS est largement implantée dans la littérature et de nombreux programmes informatiques de détection de ruptures. Née dans les années 1970 [155, 156] cette méthode a su plaire grâce à sa simplicité et sa rapidité de calcul. Pour résumer son fonctionnement, la méthode BS repose sur la méthode de détection d'un instant de rupture unique (voir 2.1) que l'on applique à l'ensemble des données. Dans le cas où un instant de rupture est détecté, les données sont découpées en deux segments sur lesquels on utilise à nouveau la méthode de détection de rupture pour un instant unique et ainsi de suite. la méthode BS s'arrête dès lors que plus aucun instant de rupture n'est détecté.

Mathématiquement, la méthode BS cherche à minimiser l'équation II.3 (voir 2.1 pour les notations).

$$\sum_{i=1}^{m+1} \left[\mathcal{C}(v_{(\tau_{i-1}+1):\tau_i}) \right] + \beta m \quad (\text{II.3})$$

l'algorithme BS est très rapide grâce sa complexité quasi-linéaire en $\mathcal{O}(n \log(n))$. Cette qualité est la raison pour laquelle la méthode BS est encore largement utilisée pour résoudre des problèmes de détection de ruptures. Cependant, il n'est pas garanti de trouver le minimum global de l'équation II.3 par cette méthode. En d'autres termes, la méthode BS est une approche approximative au problème de recherche d'instant de ruptures.

Segment Neighbourhood (SN)

Les problèmes de détection de ruptures devenant de plus en plus complexes et demandant une augmentation perpétuelle de la précision impliquèrent la recherche d'alternatives à l'algorithme BS. C'est en 1989 qu'apparait un algorithme exact de recherche d'instant de ruptures [10] : la méthode SN.

Basée sur le concept de la programmation dynamique, la méthode SN parcourt plusieurs fois l'ensemble des données sous contrainte d'un nombre d'instant de ruptures défini à l'avance, M . Nous sommes donc dans le cas d'un problème d'optimisation sous contrainte, par opposition à un problème d'optimisation pénalisé.

La méthode SN cherche à trouver la fonction de coût la plus faible en parcourant toutes les segmentations possibles de 0 à M.

Le principale défaut de cette méthode réside dans sa complexité élevée $\mathcal{O}(Mn^2)$ rendant l'algorithme lent en comparaison de la méthode BS.

Optimal Partitioning (OP)

Le troisième algorithme que nous avons envisagé est une méthode plus récente, datant de 2005 [96]. Egalement basée sur la programmation dynamique, l'idée est de résoudre un problème d'optimisation pénalisé de manière exacte. Cette méthode, appelée Optimal Partitioning (OP), a pour objectif de minimiser l'équation II.4

$$\sum_{i=1}^{m+1} [\mathcal{C}(v_{(\tau_{i-1}+1):\tau_i}) + \beta] \quad (\text{II.4})$$

Ceci revient à minimiser l'équation II.2 pour $f(m) = m$.

L'algorithme OP est un processus récursif dont le but est d'optimiser à chaque pas de temps t en se basant sur la solution optimale des pas de temps précédents $1..(t-1)$. Plus précisément, on considère les instants compris entre le dernier instant de rupture connu et l'instant t comme candidats pour déterminer un potentiel nouvel instant de rupture. Le système II.5 pose les bases de l'algorithme OP.

$$\begin{cases} F(0) = -\beta \\ F(t) = \min_{0 \leq \tau < t} [F(\tau) + \mathcal{C}(v_{(\tau+1):t}) + \beta] \end{cases} \quad (\text{II.5})$$

Le système II.5 est résolu pour chaque pas de temps $t = 1..n$ où τ représente le dernier instant de rupture connu. $F(n)$ donne alors le coût optimal et en stockant les valeurs de τ , nous obtenons les instants de ruptures de notre problème. L'algorithme 1 présente l'implémentation de la méthode OP [108].

La méthode OP calcule l'équation II.5 pour $t = 1 : n$ et nécessite une minimisation pour $\tau = 0 : (t-1)$ ce qui donne une complexité en $\mathcal{O}(n^2)$. L'algorithme OP est donc plus rapide

Algorithme 1 : Optimal Partitioning

Entrée : Une série temporelle de la forme (v_1, v_2, \dots, v_n) où $v_i \in \mathbb{R}$.
Une fonction de coût \mathcal{C} .
Une pénalité β constante qui ne dépend pas du nombre et de l'emplacement des instants de ruptures.

Sortie : Les instants de ruptures enregistrés dans $cp(n)$

Initialisation : n = longueur de la série temporelle, $F(0) = -\beta$ et $cp(0) = NULL$

```
1 pour  $\tau^* = 1, \dots, n$  faire
2    $F(\tau^*) = \min_{0 \leq \tau \leq \tau^*} [F(\tau) + \mathcal{C}(v_{(\tau+1):\tau^*}) + \beta]$ 
3    $\tau' = \arg\{\min_{0 \leq \tau \leq \tau^*} [F(\tau) + \mathcal{C}(v_{(\tau+1):\tau^*}) + \beta]\}$ 
4    $cp(\tau^*) = (cp(\tau'), \tau')$ 
5 fin
```

que l'algorithme SN en restant une méthode exacte.

Une façon de rendre l'algorithme OP plus efficace est de procéder à ce que l'on appelle un élagage des candidats aux instants de ruptures. Le prochain paragraphe détaille une méthode d'élagage appelée PELT (**P**runed **E**xact **L**inear **T**ime).

Pruned Exact Linear Time (PELT)

Il faut attendre 2012 pour voir apparaître une méthode améliorant l'OP [108]. Appelée **P**runed **E**xact **L**inear **T**ime (PELT), cette méthode est basée sur la méthode OP en y ajoutant une étape d'élagage (*pruning* en anglais). Il s'agit d'une méthode exacte solutionnant le problème d'optimisation pénalisé.

Le principe de l'étape d'élagage est de supprimer les valeurs de τ qui ne pourront jamais être un instant de ruptures candidat, à chaque itération. Killick et al. définissent alors la condition d'élagage suivante :

Il existe une constante K telle que pour tout $t \leq s \leq T$,

$$\mathcal{C}(v_{(t+1):s}) + \mathcal{C}(v_{(s+1):T}) + K \leq \mathcal{C}(v_{(t+1):T}) \quad (\text{II.6})$$

Alors pour $t > s$, si

$$F(t) + \mathcal{C}(v_{(t+1):s}) + K \geq F(s) \quad (\text{II.7})$$

alors à un temps futur $T \geq s$, t ne peut pas être le dernier instant de ruptures précédent T . t peut alors être retiré de la liste des candidats et ne sera plus cherché dans les futures itérations. Le pseudo-code implémentant la méthode PELT est donné par l’algorithme 2. L’intérêt de la méthode PELT réside dans sa complexité, inférieure ou égale à celle de la

Algorithme 2 : Pruned Exact Linear Time

Entrée : Une série temporelle de la forme (v_1, v_2, \dots, v_n) où $v_i \in \mathbb{R}$.
 Une fonction de coût \mathcal{C} .
 Une pénalité β constante qui ne dépend pas du nombre et de l’emplacement des instants de ruptures.
 Une constante K qui satisfait l’équation II.6.

Sortie : Les instants de ruptures enregistrés dans $cp(n)$

Initialisation : n = longueur de la série temporelle, $F(0) = -\beta$, $cp(0) = NULL$ et $R_1 = \{0\}$

```

1 pour  $\tau^* = 1, \dots, n$  faire
2    $F(\tau^*) = \min_{\tau \in R_{\tau^*}} [F(\tau) + \mathcal{C}(v_{(\tau+1):\tau^*}) + \beta]$ 
3    $\tau^1 = \arg\{\min_{\tau \in R_{\tau^*}} [F(\tau) + \mathcal{C}(v_{(\tau+1):\tau^*}) + \beta]\}$ 
4    $cp(\tau^*) = (cp(\tau^1, \tau^1))$ 
5    $R_{\tau^*+1} = \{\tau \in R_{\tau^*} \cup \{\tau^*\} : F(\tau) + \mathcal{C}(v_{(\tau+1):\tau^*}) + K \leq F(\tau^*)\}$ 
6 fin

```

méthode OP, rendant la méthode PELT plus rapide en pratique. En effet, la complexité de la méthode PELT peut être en $\mathcal{O}(n)$, dans le cas le plus favorable et sous certaines conditions [108]. Dans le cas le plus défavorable, si aucun élagage n’a eu lieu, la complexité sera en $\mathcal{O}(n^2)$.

Afin d’avoir une vue globale sur les méthodes de segmentation présentées dans cette thèse, j’ai regroupé leurs caractéristiques dans le tableau II.1. J’ai choisi la méthode PELT qui per-

| Algorithme | Type d’optimisation | Complexité |
|-----------------------|---------------------|-------------------------------------|
| Binary Segmentation | Pénalisé | $\mathcal{O}(n \log(n))$ |
| Segment Neighbourhood | Contraint | $\mathcal{O}(Mn^2)$ |
| Optimal Partitioning | Pénalisé | $\mathcal{O}(n^2)$ |
| PELT | Pénalisé | $\mathcal{O}(n) - \mathcal{O}(n^2)$ |

TABLE II.1 – Récapitulatif des méthodes de segmentation envisagées pour segmenter le test de RabiT[®]

met de résoudre un problème d’optimisation pénalisé de façon exact et avec une complexité

acceptable. Cela correspond aux attentes pour la segmentation du test de RabiT[®]. Bien évidemment, d'autres méthodes de segmentation existent comme la méthode FPOP (Functional Pruning Optimal Partitioning) [118].

La méthode PELT a été retenue pour deux raisons. La première raison est que la rapidité de l'algorithme PELT me convient largement en donnant une solution quasi instantanée (voir section 3.1). En effet, l'entreprise BillaTraining souhaite passer d'un traitement manuel du test de RabiT[®] à un traitement automatique. Cependant, je n'écarte pas la possibilité d'implémenter la méthode FPOP dans le futur afin de gagner encore en rapidité. La deuxième raison est l'accès au package R développé par Killick et al.². Dans un souci de mettre rapidement sur le marché une application web autonome, l'utilisation d'un package testé et validé par la communauté scientifique balaye les coûts de développement liés au temps et aux erreurs.

Dans le prochain paragraphe, nous allons appliquer la théorie de détection d'instant de ruptures au test de RabiT[®] via la méthode PELT.

3 Application au test de RabiT[®]

Pour rappel, le but de ce chapitre est d'automatiser l'analyse du test de RabiT[®] qui est aujourd'hui le goulot d'étranglement de la méthode BillaTraining. La segmentation du test de RabiT[®] peut être vu comme un problème d'optimisation pénalisé avec changement dans la moyenne. On souhaite une méthode exacte et rapide, c'est à dire, quasi-instantanée si l'on se place du côté de l'utilisateur final. La méthode PELT semble répondre à ce besoin. Il est temps de l'appliquer. Cette sous section se décompose ainsi en deux parties.

Dans la première partie, nous allons nous concentrer sur la construction des solutions. Nous rappelons au lecteur que même si nous avons une idée approximative du nombre d'instant de ruptures dans le test de RabiT[®] (environ 12), il est impossible d'avoir la certitude que le coureur démarre sa montre en même temps que le début du premier palier "facile" et la coupe à la fin du deuxième palier "facile".

²<https://cran.r-project.org/web/packages/changepoint/changepoint.pdf>

La deuxième partie consiste à sélectionner la meilleure solution parmi les solutions calculées par la méthode PELT. De plus, nous présenterons la fonction de filtrage qui nous permet de statuer sur la conformité ou la non conformité du test de RabiT®.

3.1 Construction des solutions

Démarrons cette partie par le téléchargement du package intitulé "*changepoint*"³, développé par Rebecca Killick et al. Nous pouvons cloner le *repository* sur Github⁴. Le package "*changepoint*" est développé avec le langage de programmation **R**.

Faisons une parenthèse sur **R** qui est un langage de programmation libre principalement conçu pour les statistiques et l'analyse de données. La première version de **R** date de 1993 et la dernière version connue (3.5.1) est disponible depuis le 2 juillet 2018. Pour ma part, j'ai utilisé la version 3.3.3 de **R** pour toute la suite de ce chapitre. Pour notre application, **R** possède deux intérêts. Le premier est qu'il existe de nombreux *packages*, comme "*changepoint*", que nous pouvons adapter facilement. le second réside dans sa compatibilité avec le langage Python grâce au projet *rpy2*⁵. Ceci nous permettra par la suite de créer le micro service web pour segmenter automatiquement le test de RabiT®.

Revenons à présent à notre problématique, à savoir, la construction des solutions de segmentation du test de RabiT® grâce à l'algorithme PELT du package "*changepoint*". Nous avons à notre disposition cinquante tests de RabiT® décomposés comme suit :

- Quarante tests jugés comme "valide", c'est-à-dire qui auraient été acceptés par un physiologiste BillaTraining.
- Dix tests jugés comme "non-valide", c'est-à-dire qui auraient été rejetés par un physiologiste BillaTraining.

L'idée est de créer un programme s'approchant du comportement d'un physiologiste BillaTraining. Comme évoqué précédemment, nous nous concentrons sur les changements dans la

³<https://cran.r-project.org/web/packages/changepoint/changepoint.pdf>

⁴<https://github.com/rkillick/changepoint/>

⁵https://rpy2.readthedocs.io/en/version_2.8.x/

moyenne du signal de vitesse. La fonction *cpt.mean* du package "*changepoint*" a justement pour objectif le calcul des positions optimales des instants de ruptures pour un problème de changements dans la moyenne. Par conséquent, nous utiliserons la fonction *cpt.mean* dans la suite de ce chapitre. Voyons désormais la façon de l'utiliser.

$$cpt.mean(data, penalty, pen.value, method)$$

La fonction *cpt.mean* prend quatre arguments en entrée, en voici le détail :

- **data** : Vecteur contenant les données dans lesquelles nous souhaitons trouver les instants de ruptures.
- **penalty** : Nous choisissons la méthode "CROPS" qui permet de définir un vecteur de longueur 2 donnant une plage de pénalité donnée par l'argument *pen.value*.
- **pen.value** : vecteur de longueur 2 (min, max).
- **method** : PELT

Le tableau II.2 donne les résultats de la fonction *cpt.mean* pour les 50 tests de RabiT® à notre disposition. En regardant de plus près le tableau II.2, on voit un nombre de solutions moyen de 36,6. Le temps de calcul moyen est de 0,79 secondes, le temps médian est de 0,76 secondes et l'écart type est de 0,10 secondes. Comme attendu, les temps de calcul de l'algorithme PELT sont très rapide. A présent, il faut choisir la solution s'approchant de l'oeil du physiologiste en mettant en place des règles de sélection. Le but de la prochaine sous section est de déterminer cette solution optimale en pratique.

3.2 Sélection et filtrage de la meilleure solution

La décision ou non d'analyser un test de RabiT® se fait en trois étapes.

1. Réduction de l'ensemble des solutions.
2. Choix de la solution optimale.

3. Décision sur la conformité / non conformité du test de RabiT®.

Le but de cette sous section est de détailler ces trois étapes afin d'aider le lecteur à comprendre la méthodologie employée pour la segmentation du test de RabiT®.

Pour la première étape, reprenons le test de RabiT® théorique. La figure II.2 montre, pour rappel, le test et ses 11 phases. On peut apercevoir 12 instants de ruptures dans le test

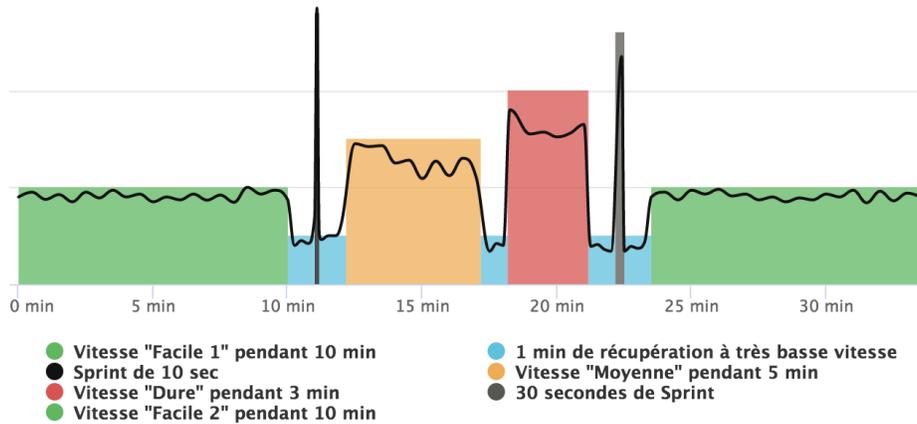


FIGURE II.2 – Les 11 phases du test de RabiT®

de RabiT®, à savoir :

- **1er instant** : Début de la phase "facile" à $t = 0$
- **2ème instant** : Fin de la phase "facile 1" / Début de la phase "repos 1" à $t = 600$
- **3ème instant** : Fin de la phase "repos 1" / Début de la phase "sprint" à $t = 660$
- **4ème instant** : Fin de la phase "sprint" / Début de la phase "repos 2" à $t = 670$
- **5ème instant** : Fin de la phase "repos 2" / Début de la phase "moyenne" à $t = 730$
- **6ème instant** : Fin de la phase "moyenne" / Début de la phase "repos 3" à $t = 1030$
- **7ème instant** : Fin de la phase "repos 3" / Début de la phase "dure" à $t = 1090$
- **8ème instant** : Fin de la phase "dure" / Début de la phase "repos 4" à $t = 1270$
- **9ème instant** : Fin de la phase "repos 4" / Début de la phase "trente secondes" à $t = 1330$

- **10ème instant** : Fin de la phase "trente secondes" / Début de la phase "repos 5" à $t = 1360$
- **11ème instant** : Fin de la phase "repos 5" / Début de la phase "facile 2" à $t = 1420$
- **12ème instant** : Fin de la phase "facile 2" à $t = 2020$

De cette information, nous pouvons en tirer la limite basse de notre filtrage. En effet, un test qui comporte moins de 12 instants de ruptures ne sera pas considéré comme un test de RabiT[®] valide et sera, par conséquent, rejeté par l'algorithme. Il est plus difficile de fixer une limite haute de manière universelle. En réalité, les moments choisis par le coureur pour lancer et couper la mesure sont assez aléatoires. L'expérience pratique montre qu'une limite haute de 18 instants de ruptures permet de se rapprocher de la tolérance accordée par un physiologiste BillaTraining. Ainsi, nous retenons uniquement les solutions ayant un nombre d'instant de ruptures compris entre 12 et 18 inclus.

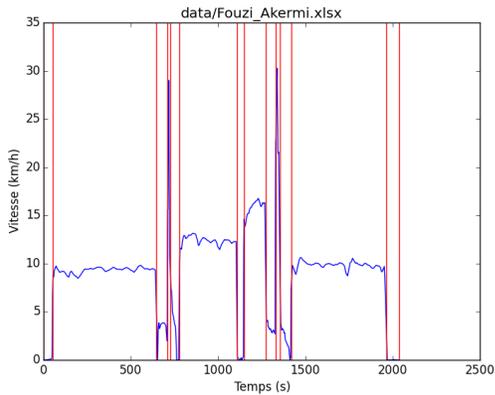
La deuxième étape, que nous avons baptisé le **choix de la solution optimale**, consiste à vérifier deux éléments dans le signal de vitesse. Le premier élément est l'enchaînement des phases du test de RabiT[®]. Prenons par exemple un test où la phase "dure" se retrouve avant la phase "moyenne". Il convient de le rejeter puisque le test de RabiT[®] est avant tout un audit (i.e. test de forme) ou l'ordre des phases à un sens d'un point de vue physiologique. Le second élément est la durée des phases du test de RabiT[®]. Prenons un autre exemple où cette fois-ci, le coureur termine la phase "trente secondes" en 10 secondes. Logiquement le test devra être rejeté.

Suite à mes discussions avec les physiologistes BillaTraining et le Professeur Billat, nous avons convenu de tolérances sur les temps de chaque phase du test de RabiT[®] afin que ce dernier garde sa qualité d'évaluation de la forme d'un coureur. Le tableau [II.3](#) rapporte ces tolérances.

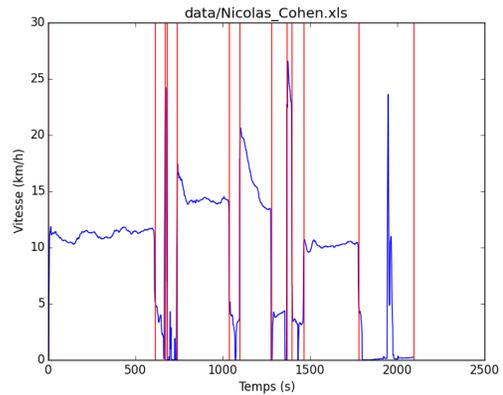
Finalement, la troisième étape, porte sur l'acceptation / le rejet du test de RabiT[®]. Cette étape regroupe l'étape 1 et l'étape 2. En partant de la liste contenant les solutions comprises entre 12 et 18 instants de ruptures, un algorithme vient chercher la solution se rapprochant

au plus près de la théorie du test de RabiT[®]. Une fois cette solution isolée, on statue sur la conformité / non conformité du test.

Pour finir cette sous section, je vous laisse avec la figure II.3 qui présente deux tests de RabiT[®]. Un conforme (figure II.3a) et un non conforme (figure II.3b).



(a) Test de RabiT[®] conforme et validé par l'algorithme



(b) Test de RabiT[®] non conforme et rejeté par l'algorithme

FIGURE II.3 – Tests de RabiT[®] en conditions réelles. (a) correspond à un test conforme et validé par l'algorithme. (b) correspond à un test non conforme et rejeté par l'algorithme

4 Création d'un micro-service web

La dernière marche à franchir pour lever ce premier verrou est la mise en production de notre solution. C'est exactement ce que nous allons voir dans cette section par le biais de la création d'un micro service web et d'une API grâce au *framework*⁶ Django. Commençons par expliquer ce qu'est un micro service, à quoi il sert et quels sont les avantages à l'utiliser.

4.1 Architectures monolithique VS. micro service

L'idée de micro service est relativement récente et date du début des années 2010. Elle se place comme alternative aux architectures monolithiques, encore largement utilisées aujourd'hui. Voyons les différences entre ces deux types d'architecture et essayons de justifier notre choix du micro service pour la segmentation du test de RabiT[®].

⁶framework est le mot anglais pour cadre de travail

L'architecture monolithique

Une architecture monolithique peut être vue comme une application autonome construite en un seul bloc. Toute la logique et l'accès à la base de données se trouve sur le même serveur. De plus, toutes les briques logiques sont interconnectées et une forte dépendance existe entre elles. Pour plus de clarté, la figure II.4 donne une vue schématique d'une architecture monolithique. On y distingue quatre fonctionnalités : **le paiement en ligne**, **les notifications**, **la gestion de l'entraînement** et **l'authentification des utilisateurs**. Ces quatre fonctionnalités sont interconnectées et se situent dans la même application. L'avantage principale de ce type d'architecture est la rapidité de développement qui permet de tester rapidement une idée sur le marché à moindre coût. Cependant, la maintenance est complexe car la moindre modification d'une seule fonctionnalité peut impacter le fonctionnement de l'ensemble de l'application.

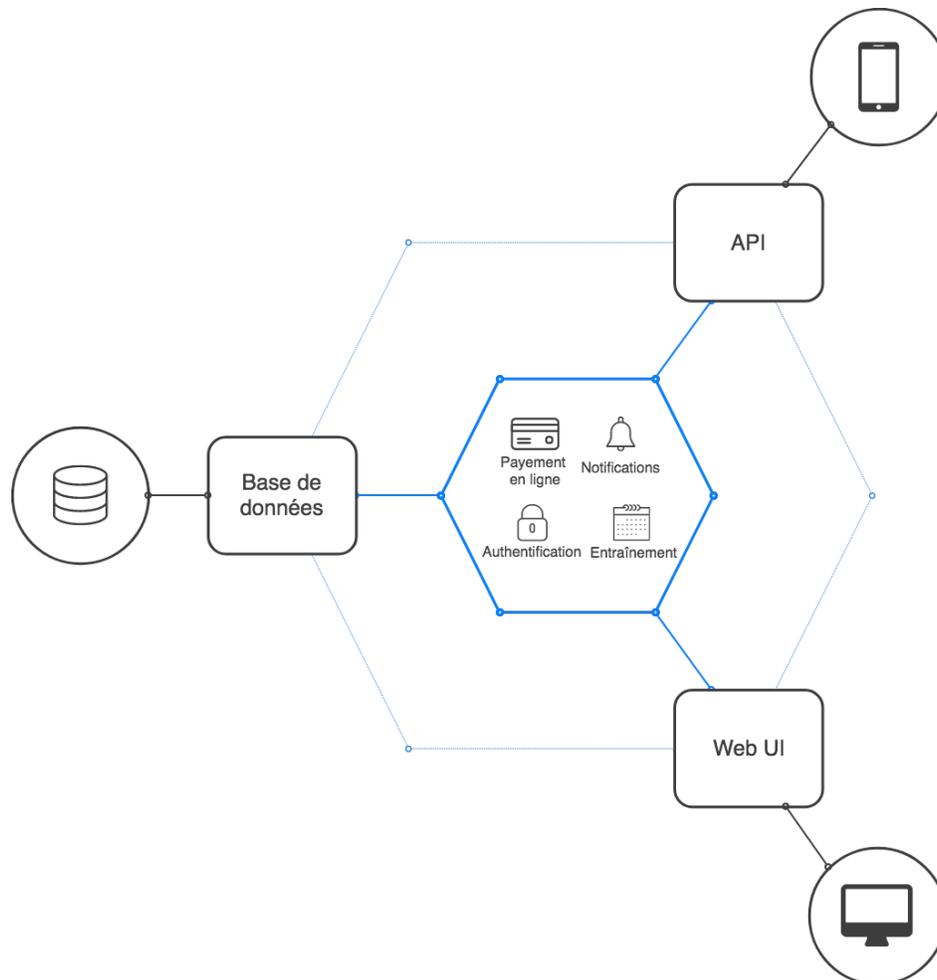


FIGURE II.4 – Schéma d'une application construite sur une architecture monolithique

Beaucoup d'entreprise choisissent ce type d'architecture à leurs débuts. Netflix ou encore Ebay utilisaient les architectures monolithiques. Les deux entreprises ont progressivement migrées vers une architecture en micro services. En effet, l'architecture en micro services offre de nombreux avantages notamment lorsque le nombre d'utilisateurs croît fortement.

L'architecture en micro services

L'architecture en micro services prend le contre-pied de l'architecture monolithique. Il convient de séparer chaque fonctionnalité en les rendant autonomes et indépendantes les unes des autres. La figure II.5 donne une vue schématique d'une architecture en micro services. Chacune des quatre fonctionnalité (**le paiement en ligne, les notifications, la gestion de l'entraînement** et **l'authentification des utilisateurs**) est transformée en service et possède sa base de données. Les avantages d'une telle architecture sont multiples. Premièrement,

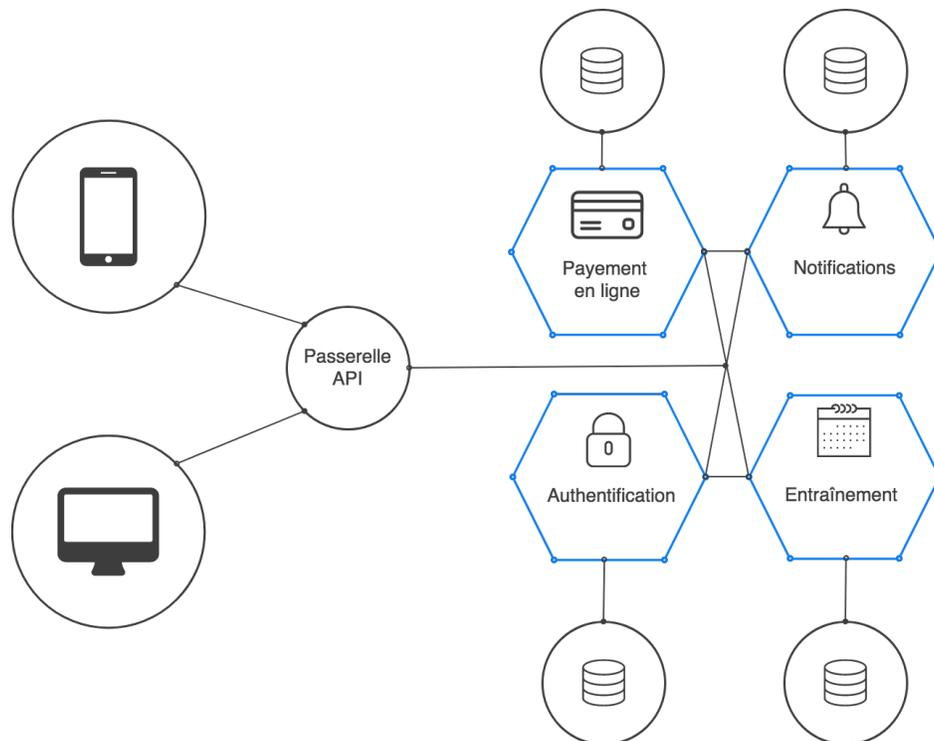


FIGURE II.5 – Schéma d'une application construite sur une architecture en micro services

la maintenance, les tests et le fonctionnement sont indépendants des autres services. Par exemple, si le service de notification est en panne, les utilisateurs pourront toujours gérer leur entraînement, s'authentifier et payer en ligne. Deuxièmement, les besoins de l'applica-

tion sont facilement ajustables. En effet, si je fais face à une forte demande de création de programme d'entraînement suite à une action de communication, je peux augmenter la taille des serveurs pour **le paiement en ligne et la gestion de l'entraînement** uniquement. Finalement, chaque service peut être développé dans un langage différent et adapté. Par exemple, je peux utiliser le langage Python pour un service nécessitant l'emploi des mathématiques et le langage Javascript pour créer un service en temps réel.

Le désavantage de l'architecture en micro services et le temps de développement, plus long que celui de l'architecture monolithique.

En prenant en compte les avantages et inconvénients des deux solutions, j'ai décidé de développer l'algorithme de segmentation du test de RabiT[®] comme un micro service. La première raison est l'utilisation du langage R combiné au langage Python qui ne permet pas une intégration direct dans les applications BillaTraining existantes en PHP. La seconde raison est que nous souhaitons rendre indépendante la segmentation du test de RabiT[®] puisque nous l'utiliserons dans plusieurs applications.

D'un point de vue pratique, je pense que les deux architectures sont valables. Même si les bonnes pratiques et l'effet de mode nous pousse à utiliser une architecture en micro services, une architecture monolithique ou hybride (monolithique et micro services) peut être la solution ; surtout pour une jeune entreprise avec peu de ressources humaines en développement qui souhaite tester rapidement l'intérêt du marché pour un produit ou service en ligne.

4.2 Application Programming Interface (API)

Sur les figures [II.4](#) et [II.5](#) apparaît la notion d'API. Le but de cette sous section est d'expliquer brièvement le concept d'API. En d'autres termes, nous allons répondre aux deux question suivantes : Qu'est ce qu'une API? A quoi sert une API?

API est l'acronyme anglais de "**A**pplication **P**rogramming **I**nterface" pouvant être traduit en français par "Interface de programmation". Une API est une partie d'un serveur destinée à

recevoir des requêtes (protocole HTTP ⁷) et envoyer des réponses (formats XML⁸ ou JSON⁹). Pour vulgariser, je dirai qu'une API est une boîte noire vers laquelle on envoie des données et qui nous envoie un résultat en retour.

La création d'une API permet de standardiser les requêtes et les réponses et de créer des routines indépendantes. Souvent, les API suivent des standards (REST, SOAP, ...). J'ai utilisé le standard REST pour sa notoriété (i.e. il est facile de trouver de la documentation sur internet) et sa polyvalence [60]. L'avantage d'utiliser une API REST est que l'on est proche des standards HTTP ce qui est très avantageux et qui évite les erreurs lors du développement d'applications web. Je conseille au lecteur le cours "Utilisez des API REST dans vos projets web" disponible sur la plateforme OpenClassrooms ¹⁰ s'il désire mieux comprendre le fonctionnement d'une API REST.

La figure II.6 donne une représentation schématique de l'API de segmentation du test de RabiT[®]. Sur la gauche de la figure II.6, nous pouvons voir trois entrées. Il s'agit de

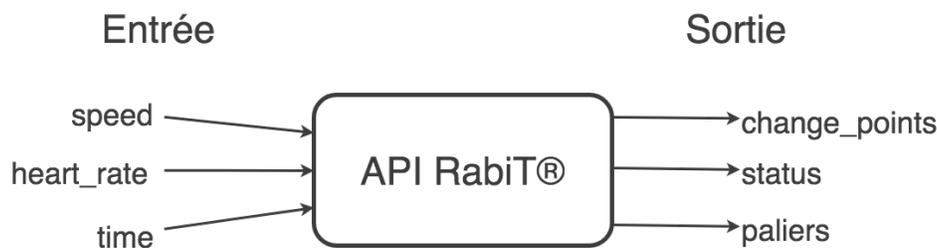


FIGURE II.6 – Schéma de l'API de segmentation du test de RabiT[®]. A gauche de la figure se trouvent les entrées (requête). A droite se trouvent les sorties (réponse)

trois tableaux de données de même dimension et qui correspondent aux données enregistrées durant le test de RabiT[®]. *speed* représente les données de vitesse en kilomètres par heure, *heart_rate* la fréquence cardiaque en battements par minute et *time* le temps en secondes.

Le centre de la figure II.6 représente la boîte noire qui analyse les données issues du test de RabiT[®] en essayant de le segmenter.

La droite de la figure II.6 donne les sorties (i.e. les résultats) de la segmentation. La première

⁷Hypertext Transfer Protocol

⁸Extensible Markup Language

⁹JavaScript Object Notation

¹⁰<https://openclassrooms.com/fr/courses/3449001-utilisez-des-api-rest-dans-vos-projets-web>

sortie, appelée *change_points* renvoie un tableau contenant tous les instants de ruptures trouvés par l'algorithme de détection de ruptures. La deuxième sortie, appelée *status* renvoie une chaîne de caractères permettant de statuer sur l'acceptation / le rejet du test de RabiT[®]. La troisième sortie, appelée *paliers*, est renvoyée uniquement si le test de RabiT[®] est rejeté. Il s'agit de l'explication du rejet du test.

La dernière partie de ce chapitre dévoile les étapes de mise en production de l'algorithme de segmentation du test de RabiT[®] grâce au *framework* Django.

4.3 Création d'un micro service web pour segmenter le test de RabiT[®]

Dans le monde du développement, il n'est pas rare d'entendre la phrase "ne pas réinventer la roue". Cela signifie qu'il est inutile de coder ce qui existe déjà. C'est avec cette philosophie qu'est né le concept de *framework*, ou cadre de travail en français. Construit par des développeurs expérimentés, l'objectif du *framework* est de faciliter le travail de l'ensemble des développeurs. En utilisant un *framework*, le développeur programmera plus vite, de manière plus propre et aura plus de facilité à maintenir son code.

Dans le cadre de la segmentation du test de RabiT[®], j'ai décidé d'utiliser le *framework* Django¹¹ écrit en Python. Il s'agit d'un ensemble de classes et de méthodes spécialement écrites afin de donner les composants de base à la création d'une application web reposant sur le langage Python. Le choix du *framework* n'est pas anodin, en effet, l'algorithme de détection de ruptures a été écrit en Python également (voir section 3.1).

Une fois l'API de segmentation du test de RabiT[®] terminée et testée localement sur mon ordinateur, il est temps de la passer en production, c'est-à-dire la rendre disponible en continu (i.e. 24h sur 24 et 7 jours sur 7). La finalité étant de construire une application web autonome (i.e. avec le moins d'interventions humaines possibles).

J'ai choisi un serveur dédié chez OVH pour héberger l'application web. Le rapport perfor-

¹¹<https://www.djangoproject.com/>

mance / coût me semble intéressant pour démarrer. Je suis conscient que cette solution est temporaire et que je devrai migrer l'application vers un hébergement *cloud*. L'avantage d'un hébergement *cloud* réside dans sa flexibilité permettant d'absorber un trafic (i.e. nombre d'utilisateur en simultané) plus important. L'inconvénient est le coût, souvent supérieur ou très imprévisible.

L'étape suivante est la mise en place du protocole HTTPS afin d'utiliser une communication sécurisée avec notre API fraîchement créée. Bien que pas (encore) obligatoire, le protocole HTTPS est fortement recommandé lorsqu'un transfert de données d'un point A à un point B a lieu. Ce transfert de données est justement le but d'une API. Pour sécuriser notre API, j'ai utilisé Certbot¹² et Let's Encrypt¹³.

La dernière étape de mise en production consiste à lier le serveur HTTP présent sur notre serveur OVH avec notre application Django. Pour ce faire, il est nécessaire d'utiliser une **Web Server Gateway Interface** (WSGI). Pour ce projet, j'ai utilisé la WSGI Gunicorn¹⁴. La figure II.7 positionne les étapes de mise en production de notre API les unes par rapport aux autres.

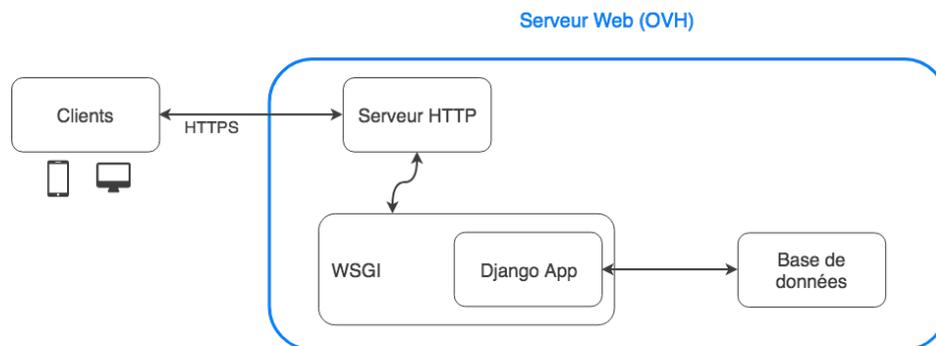


FIGURE II.7 – Schéma de l'architecture de mise en production de l'application de segmentation du test de RabiT®

Notre API est désormais fonctionnelle, c'est-à-dire que nous pouvons lui envoyer une requête par le protocole HTTPS et nous recevrons une réponse en retour. Les figures II.8 et II.9 présentent deux réponses obtenues via notre API. La figure II.8 correspond à une

¹²<https://certbot.eff.org/>

¹³<https://letsencrypt.org/>

¹⁴<https://gunicorn.org/>

réponse découlant sur l'acceptation du test de RabiT[®]. La figure II.9, quant à elle, entraîne le rejet du test de RabiT[®] et en donne les causes. On voit que le paliers "facile 2" porte le label "ERREUR".

```
1 {
2   "change_points": [0, 611, 663, 674, 728, 1030, 1087, 1267, 1325, 1356, 1417, 2012],
3   "status": "RABIT_CONFORME"
4 }
```

FIGURE II.8 – Réponse de l'API de segmentation du test de RabiT[®] au format JSON pour un test accepté.

```
1 {
2   "change_points": [0, 608, 669, 680, 740, 1035, 1158, 1213, 1339, 1428, 1459, 1549, 1882],
3   "status": "RABIT_NON_CONFORME_NON_TRAITABLE",
4   "paliers": {
5     "msg_dur": "OK",
6     "msg_facile1": "OK",
7     "msg_facile2": "ERREUR",
8     "msg_moyen": "OK",
9     "msg_repos1": "OK",
10    "msg_repos2": "OK",
11    "msg_repos3": "OK",
12    "msg_repos4": "OK",
13    "msg_repos5": "OK",
14    "msg_sprint": "OK",
15    "msg_trenteSeconde": "OK"
16  }
17 }
```

FIGURE II.9 – Réponse de l'API de segmentation du test de RabiT[®] au format JSON pour un test rejeté.

Bilan

Afin de trouver une solution à la segmentation du test de RabiT[®], j'ai commencé par me documenter sur les techniques de détection d'instantants de ruptures. L'algorithme PELT semblait répondre à notre besoin par sa précision et sa rapidité d'exécution.

En appliquant l'algorithme PELT au test de RabiT[®], je me suis aperçu qu'il remplissait effectivement son rôle avec des temps d'exécution inférieurs à la seconde. J'ai alors construit les filtres et les tests permettant de retenir la solution optimale de segmentation du test de RabiT[®] avec l'aide des physiologistes BillaTraining. Ceci m'a permis de statuer sur l'accep-

tation / le rejet du test de RabiT[®]. Une fois le test de RabiT[®] accepté, les résultats de la segmentation permettent de construire le Radar Énergétique.

Par la suite, j'ai eu l'occasion de mettre en production la solution de segmentation grâce au *framework* Django en construisant une API REST.

Deux perspectives me semblent intéressantes dans ce chapitre. La première est l'utilisation de nouveaux algorithmes de détection de ruptures afin d'améliorer la qualité de la segmentation du test de RabiT[®]. La seconde est le concept de rupture en temps réel qui permet de segmenter des données arrivant en continu par opposition au concept de rupture rétrospective qui traite un signal entier. Cette idée de temps réel me semble très intéressante par la suite car elle nous permettrait de corriger le coureur ou lui demander de refaire une phase du test non reconnue.

Ce deuxième chapitre a été l'occasion de répondre librement à une problématique d'entreprise. Partir d'une feuille blanche sans savoir quelle direction prendre pour finalement aboutir à une solution utile a été un voyage passionnant pour moi.

| Coureur | Nombre de Solutions | Nombre IR min | Nombre IR max | Complexité en temps (s) |
|---------|---------------------|---------------|---------------|-------------------------|
| 1 | 45 | 6 | 57 | 0.83 |
| 2 | 37 | 0 | 46 | 0.85 |
| 3 | 35 | 0 | 40 | 0.87 |
| 4 | 28 | 1 | 36 | 0.70 |
| 5 | 24 | 1 | 32 | 0.73 |
| 6 | 38 | 4 | 52 | 0.85 |
| 7 | 30 | 2 | 38 | 0.74 |
| 8 | 22 | 0 | 26 | 0.68 |
| 9 | 40 | 1 | 51 | 0.78 |
| 10 | 37 | 1 | 52 | 0.74 |
| 11 | 42 | 2 | 54 | 0.84 |
| 12 | 37 | 2 | 47 | 1.32 |
| 13 | 24 | 1 | 33 | 0.70 |
| 14 | 32 | 2 | 44 | 0.74 |
| 15 | 21 | 0 | 26 | 0.69 |
| 16 | 30 | 1 | 37 | 0.76 |
| 17 | 33 | 0 | 39 | 0.74 |
| 18 | 47 | 1 | 63 | 0.89 |
| 19 | 32 | 1 | 36 | 0.74 |
| 20 | 32 | 2 | 43 | 0.76 |
| 21 | 65 | 5 | 83 | 0.91 |
| 22 | 36 | 2 | 48 | 0.82 |
| 23 | 52 | 0 | 65 | 0.73 |
| 24 | 44 | 0 | 62 | 0.93 |
| 25 | 30 | 2 | 38 | 0.84 |
| 26 | 34 | 1 | 40 | 0.72 |
| 27 | 35 | 1 | 44 | 0.74 |
| 28 | 46 | 0 | 58 | 0.78 |
| 29 | 30 | 0 | 38 | 0.72 |
| 30 | 22 | 0 | 28 | 0.71 |
| 31 | 44 | 0 | 51 | 0.80 |
| 32 | 30 | 2 | 38 | 0.76 |
| 33 | 31 | 0 | 36 | 0.73 |
| 34 | 44 | 1 | 61 | 0.88 |
| 35 | 42 | 1 | 52 | 0.71 |
| 36 | 36 | 1 | 46 | 0.75 |
| 37 | 30 | 0 | 35 | 0.71 |
| 38 | 28 | 0 | 32 | 0.74 |
| 39 | 35 | 7 | 46 | 0.71 |
| 40 | 56 | 0 | 75 | 0.90 |
| 41 | 31 | 2 | 39 | 0.78 |
| 42 | 25 | 1 | 29 | 0.68 |
| 43 | 90 | 1 | 127 | 0.92 |
| 44 | 22 | 1 | 30 | 0.68 |
| 45 | 30 | 0 | 37 | 0.73 |
| 46 | 36 | 1 | 47 | 0.85 |
| 47 | 36 | 3 | 43 | 0.81 |
| 48 | 44 | 1 | 57 | 0.84 |
| 49 | 44 | 0 | 55 | 0.89 |
| 50 | 36 | 0 | 40 | 0.76 |

TABLE II.2 – Résultats de la fonction *cpt.mean* du package R *change point* appliquée au signal de vitesse du test de RabiT[®]. On peut lire le nombre de solutions, le nombre d’instantants de ruptures (IR) min, le nombre d’instantants de ruptures (IR) max et le temps de calcul de l’algorithme en secondes.

| Phase | temps théorique (s) | limite basse (s) | limite haute (s) |
|-----------------|---------------------|------------------|------------------|
| Facile | 600 | 500 | 700 |
| Repos | 60 | 40 | 90 |
| Sprint | 10 | 5 | 20 |
| Moyenne | 300 | 270 | 330 |
| Dure | 180 | 160 | 200 |
| Trente secondes | 30 | 22 | 38 |

TABLE II.3 – Tolérances sur les temps pour chaque phase du test de RabiT®.

Chapitre III

Deuxième contribution : Vers une nouvelle stratégie de course sur marathon

Dans ce troisième chapitre, nous allons voir les travaux que nous avons eu l'occasion sur la stratégie de course optimale. Bien que nous n'ayons pas eu le temps de créer un outil prédictif capable de générer des stratégies de course personnalisées, nous avons pu décrire la façon de courir un marathon en proposant une philosophie à adopter. Par le biais de deux articles scientifiques, nous avons mis en évidence un nouvel outil, appelé signature statistique, permettant de déterminer si le marathon a été couru de manière optimale indépendamment du genre, niveau de performance et profil de course (i.e. Berlin ou Paris).

Sommaire

| | | |
|----------|--|-----------|
| 1 | Problématique et enjeux | 86 |
| 2 | [Article] Detecting the marathon asymmetry with a statistical signature | 87 |
| 2.1 | Introduction | 88 |
| 2.2 | Methods | 90 |
| 2.2.1 | Methodology | 90 |

| | | |
|----------|--|------------|
| 2.2.2 | Statistical study | 90 |
| 2.3 | Results | 93 |
| 2.3.1 | Speed decreasing trend | 93 |
| 2.3.2 | Speed strong asymmetry behavior | 93 |
| 2.3.3 | Speed statistical signature | 97 |
| 2.4 | Discussion | 98 |
| 2.5 | Conclusion | 101 |
| 3 | [Article En Soumission] The 2018 marathon world record shows the positive asymmetry pacing strategy effectiveness . . . | 101 |
| 3.1 | Introduction | 102 |
| 3.2 | Materials and Methods | 104 |
| 3.2.1 | Methodology | 104 |
| 3.2.2 | Statistical study | 104 |
| 3.2.3 | Critical Speed | 106 |
| 3.3 | Results | 107 |
| 3.3.1 | Statistical study of the whole marathon | 107 |
| 3.3.2 | Statistical study of the marathon split in two | 108 |
| 3.3.3 | Statistical study of the marathon split in four | 109 |
| 3.3.4 | Critical Speed | 111 |
| 3.4 | Discussion | 112 |
| 3.5 | Conclusion | 113 |
| | Bilan | 114 |

1 Problématique et enjeux

Pour la seconde contribution scientifique, nous souhaitons comprendre la façon de courir des marathonien, en gardant à l'esprit l'idée finale de proposer une stratégie de course person-

nalisée. Avant de prédire de telles stratégies sur-mesures, nous pensons qu’il est nécessaire de passer par une phase de description en partant des données réelles, que nous avons pu extraire de différentes applications web. Nous retrouvons, sur ces applications, des données de vitesses moyennes par kilomètre. Cet échantillonnage nous convient car nous cherchons à optimiser la stratégie de course de manière globale.

La recherche d’une stratégie de course optimale a assez peu évolué depuis 1973 et les travaux de Keller [100, 101]. Nous renvoyons le lecteur à la section 2 du chapitre I pour plus de détails. Plutôt que de repartir de la théorie de Keller, nous avons choisi d’aborder le problème par les données de vitesses issues de près de 300 marathons. De ces données, nous avons produit un premier article scientifique [24] en introduisant l’idée d’asymétrie. Nous avons identifié deux types d’asymétrie. La première, l’asymétrie positive, indique que le coureur a passé plus de temps (ou distance) en dessous de sa vitesse moyenne. Respectivement, la deuxième asymétrie, nommée négative, indique que le coureur a passé plus de temps (ou distance) au dessus de sa vitesse moyenne. En partant du concept d’asymétrie et du nouveau record du monde sur marathon établi par Eliud Kipchoge à Berlin en 2018, nous avons écrit un second article scientifique [145].

Notre objectif de recherche de stratégie de course optimale est double. Premièrement, nous souhaitons proposer une stratégie de course sur-mesure au grand public, basée sur la philosophie de course des meilleurs athlètes. Deuxièmement, nous souhaitons proposer une stratégie de course permettant de passer sous la barre symbolique des deux heures sur marathon.

2 [Article] Detecting the marathon asymmetry with a statistical signature

Abstract

Lately, the sub two-hour marathon attempt in Monza was still based on the belief that constant speed is the best way of running. This idea is relayed by marathon organizers who

offer pace-group leaders to help the runners to maintain a target race speed. The purposes of this study are to verify the hypotheses that 1. The mass runners try to maintain a constant speed without succeeding. 2. Marathoners run in an asymmetric way and this turns out to be visible in the speed time series. Those two points are independent of the gender, the level of performance (2h30-3h40) and the profile of the race (Paris vs Berlin). Before considering a predictive running strategy for optimizing personal marathon running performance, here we shed light on some significant statistical features by analyzing speed time series data recorded by 273 runners' GPS. We started with looking for a trend in the speed time series. By means of Kendall's non-parametric rank correlation coefficient we exhibited a decreasing trend in speed data, whichever the level of performance, gender (Male and Female) and race profile (Berlin and Paris marathons). Going deeper in the study we applied a systematic analysis of the asymmetry of speed via classical statistical measures of skewness. Among them the quantiles of the average speed, i.e. the proportion of the race run above or below the final average speed. The combination of the trend and the asymmetry lead to building up a statistical signature for the speed time series which is identical regardless the level of performance, gender and race profile.

Keywords

Running; Runners; Endurance; Kendall; Strava; Signature

2.1 Introduction

Nowadays the number of runners in the USA equals the whole French population, i.e. 65 millions with an increasing rate of 50% over the last ten years¹. In the same way, the number of marathon finishers increased, from 353,000 finishers in the early 2000s to 507,600 in 2016 [147]. In spite of an increasing volume of data split available on running community websites no study has systematically investigated a marathon running pacing strategy for the various levels of performance, from 2h30 to 3h40.

¹<https://www.statista.com/statistics/227423/number-of-joggers-and-runners-usa/>

Indeed variability in pacing has been studied in respect of short and middle-distance running (e.g. 3000 m to 10 km) [113, 8, 31, 48, 148, 65], these studies having especially focused on the influence of pacing on metabolic and performance measures. Ely et al. (2008) reported that elite runners completing a marathon had few changes in pace during a marathon suggesting low speed variability [57]. Even lately, the sub two-hour marathon attempt in Monza was still based on the belief that constant speed is the best way of running following the prior idea that optimal is the even pace according to the seminal model of Keller (1973) [100]. In the same way recreational marathon runners adopt the paradigm of constant speed by running with a pace-group leader from 3h00 to 4h30 every 15 min provided by the marathon organizers. Therefore the idealized paradigm of constant speed prevails in theory and perverts the practice. However, considering the physiological limitation (glycogen availability) that means that the runners have to choose the ideal speed allowing them to get the optimal performance without hitting the set of fatigue 10 km or less before the arrival (“the famous Marathon wall”) !

By applying methods from mathematical statistics we test the two following hypotheses: first, most marathoners run at a speed that is significantly increasing or decreasing ; second, they run in a significantly asymmetric way leading them, for instance, to run most of the time at a speed above their average speed including the final speed fall. In addition we test a third hypothesis that these features above are independent of 1. the gender, 2. the level of performance (2h30-3h40) and 3. the profile of the race (Paris vs Berlin) by analyzing speed time series data currently recorded by 273 runners’ GPS.

More precisely, the following three points have been studied:

- (1) The first one is that we notice a decreasing trend for the speed time series while running a marathon. We therefore characterize this trend statistically via Kendall’s τ [103, 3].
- (2) Beyond the decreasing trend we point out the presence of strong positive or negative skewness in speed time series. In other words, in the case of negative asymmetry, runners spend much more time and run a longer longer distance above the overall marathon average speed. The combination of the trend and the asymmetry lead to building up a statistical

signature for the speed time series which appears to be identical regardless the level of performance, gender and race profile.

(3) Finally we introduce a new statistical signature combining the trend and the asymmetry analyses to help runners and coaches to be able to go beyond the sole performance analysis through chronometers results. In addition before considering a predictive running strategy for optimizing personal marathon running performance, here we shed light on some significant hidden running patterns.

2.2 Methods

2.2.1 Methodology

To achieve this study we examined the time and average speed per kilometer run by 273 marathons finishers all during 2017. These data were collected from the public platform Strava².

We say that a marathon has a strong positive (resp. strong negative) asymmetry if the runner spent at least 54% of the time below (resp. above) marathon average speed. More formally in a sample of speeds measured at regular intervals, a strong positive (resp. negative) asymmetry corresponds to a percentile of the average speed greater than 54% (resp. less than 46%.)

We study the marathons through different layers introducing two categories and six groups of runners. In the first category, we take the whole dataset (273 marathons), in the second one, we only take the strongly asymmetric marathons (191 marathons).

Then we split each of the two categories into seven groups: All (A), Men (M), Women (W), Less than 3 hours (-3), Greater-than or equal 3 hours (+3), Paris (P) and Berlin (B). Table III.1 gives the frequencies of the categories and groups.

All races were run using a Global Positioning System (GPS) device.

2.2.2 Statistical study

²<https://www.strava.com/>

| | All | Men | Women | < 3 hours | ≥ 3 hours | Paris | Berlin |
|------------------|-----|-----|-------|-----------|----------------|-------|--------|
| Whole Dataset | 273 | 230 | 43 | 130 | 143 | 140 | 133 |
| Strong asymmetry | 191 | 162 | 29 | 91 | 100 | 97 | 94 |

Table III.1 – Number of marathons run divided into two categories: All Marathons (273 Marathons) and Strongly Asymmetric Speed (191 races) each subdivided in seven groups: All, Men, Women, Less than 3 hours marathons, Greater-than and equal to 3 hours marathons, Paris and Berlin.

2.2.2.1 Characterizing the trend Being aware that the accuracy of the GPS used by runners is 0.3% in open sky conditions and 1.6% in open forest or building proximity we consider the speed variable as 'continuous' even if it appears as discontinuous. Furthermore, to test the influence of the race profile, we chose two opposite marathons: Berlin having 25% of open sky conditions versus 75% of open forest/building proximity and Paris with 5% of channels, 25% of open sky and 70% of open forest/building proximity.

In order to detect the decreasing trend of speed during the marathons we are going to use the Kendall's τ [103, 3]. Kendall's τ is a non-parametric coefficient computed from a times series v_1, v_2, \dots, v_n . Here v_i is the i -th value of a speed, $i < j$ meaning that i indicates a period of time prior to j . For instance the speed v increases with time if $v_1 < v_2 < \dots < v_n$. Consider the symmetric kernel

$$K(v_j, v_i) = K(v_i, v_j) = \begin{cases} 1 & \text{if } i < j \text{ and } v_i < v_j, \\ 0 & \text{if } v_i = v_j, \\ -1 & \text{if } i < j \text{ and } v_i > v_j. \end{cases} \quad (\text{III.1})$$

$$\tau = \frac{2}{n(n-1)} \sum_{i < j} K(v_i, v_j) \quad (\text{III.2})$$

Kendall's τ is defined by Equation (III.2), the sum being preformed over the $n(n-1)/2$ distinct unordered couples of indices $\{i, j\}$, so that τ takes values in between $[-1, 1]$. The Kendall's τ equals 1 (resp. -1) only if the speed time series is increasing (resp. decreasing). Furthermore a constant speed time series leads to the value $\tau = 0$. We know from the field of mathematical statistics that Kendall's is suited to detect a decreasing or increasing linear

trend in a time series. It enjoys the qualities of being robust and non-parametric [102]. A p-value close to zero indicates an increasing (resp. decreasing) trend if $\tau > 0$ (resp. $\tau < 0$). Our data correspond to $n = 42$, v_i being the average speed during the i -th kilometer, $1 \leq i \leq 42$.

2.2.2.2 Characterizing the asymmetry The asymmetry of the time series $(v_i)_{1 \leq i \leq 42}$ will be specified in two ways:

- Time spent below vs. time spent above marathon average speed. We compute the percentage of time spent above marathon average speed for each runner.
- Comparison of the marathon median speed (v_{med}) and the marathon average speed (\bar{v}). We consider the sign of $\bar{v} - v_{med}$.

For the rest of this paper, we use the first coefficient (mathematically equivalent to the second one) and we consider that runners having a percentage of time spent above average speed of 54% (resp. 46%) or greater-than (resp. lesser-than) are strongly negatively (resp. positively) asymmetric. Our motivation is twofold. First this coefficient is easy to be used by every runner and coach. Second it is mathematically well-established and robust statistically.

2.2.2.3 Combining trend and asymmetry Next we studied trend and asymmetry jointly. The idea was to build up a statistical signature to describe marathons speed time series. The signature is composed of two signs (\pm, \pm) . Thus it takes four possible values, $(++)$, $(+-)$, $(-+)$ and $(--)$. The first sign is $+$ or $-$ according to whether the Kendall τ is positive or negative. The second sign is $+$ or $-$ according to whether the asymmetry is positive or negative. In this way each runner is easily categorized. In turn, the set of signatures of one marathon, of one gender, or any other subsample can be gathered in a 2×2 matrix whose coefficients are the frequencies or relative frequencies of the four signatures in the following ordering:

$$\begin{pmatrix} -- & -+ \\ +- & ++ \end{pmatrix}$$

2.3 Results

2.3.1 Speed decreasing trend

To characterize the trend of the speed time series we use the Kendall's τ described in section 2.2.2.1. Table III.2 shows the results of the Kendall's τ analysis for the speed time series.

Within the 273 marathons, 81.7% of runners present a decreasing trend in speed time series with an average Kendall's τ of -0.54 . The six groups (Men, Women, < 3 hours, ≥ 3 hours, Paris, Berlin) highlight the same idea with the percentage of runners having a decreasing trend in the range 72.9% - 90% and averages Kendall's τ close to -0.54 . Looking at the p-values all less than 0.05, we can reject the null hypothesis stating that the speed time series has a constant trend.

From Table III.2 we also see that 13.9% of the runners seemed to run the marathon with an increasing trend with a Kendall's τ equals to 0.30. The tendency goes into the same direction if we take a close look at the groups with the percentage of runners in the range 7.9% - 20.3% and averages Kendall's τ close to 0.30.

An another interesting fact is that we have 12 runners with a Kendall's τ in the range -0.05 - 0.05 . This suggests that a minority of runners can hold a constant speed during a marathon.

To conclude with the trend of the speed time series we clearly observe a majority of negative Kendall's tau values strengthening the idea that spontaneously runners decrease their speed during the race regardless of the gender, performance or race profile. These computations agree with the assessment from the literature [150, 22].

2.3.2 Speed strong asymmetry behavior

Let us now study asymmetry behavior of the speed distribution. In order to do so, we use the percentile of the average speed described in subsection 2.2.2.2 by computing the time spent above and below the runners' marathon average speed. Both the whole dataset and strong speed asymmetry categories have been taken into consideration in this section.

| Group | Trend | Nb. Runner | % Runner | Avg. Kendall Tau | Avg. p-value |
|-----------|-------|------------|----------|------------------|--------------|
| All | ↗ | 38 | 13.9 | 0.30 | 0.39 |
| | → | 12 | 4.4 | 0 | < 0.05 |
| | ↘ | 223 | 81.7 | -0.54 | < 0.05 |
| Men | ↗ | 34 | 14.8 | 0.30 | 0.38 |
| | → | 9 | 3.9 | 0 | < 0.05 |
| | ↘ | 187 | 81.3 | -0.55 | < 0.05 |
| Women | ↗ | 4 | 9.3 | 0.32 | 0.44 |
| | → | 3 | 7 | 0 | < 0.05 |
| | ↘ | 36 | 83.7 | -0.55 | < 0.05 |
| < 3 hours | ↗ | 21 | 16.2 | 0.29 | 0.38 |
| | → | 6 | 4.6 | 0 | < 0.05 |
| | ↘ | 103 | 79.2 | -0.54 | < 0.05 |
| ≥ 3 hours | ↗ | 17 | 11.9 | 0.32 | 0.40 |
| | → | 6 | 4.2 | 0 | < 0.05 |
| | ↘ | 120 | 79.2 | -0.55 | < 0.05 |
| Paris | ↗ | 11 | 7.9 | 0.25 | 0.32 |
| | → | 3 | 2.1 | 0 | < 0.05 |
| | ↘ | 126 | 90 | -0.57 | < 0.05 |
| Berlin | ↗ | 27 | 20.3 | 0.32 | 0.42 |
| | → | 9 | 6.8 | 0 | < 0.05 |
| | ↘ | 97 | 72.9 | -0.52 | < 0.05 |

Table III.2 – Kendall's τ and p-values for describing the trend in speed time series in a marathon (increasing: ↗, constant: → or decreasing: ↘) among the six groups (Men, Women, < 3 hours, ≥ 3 hours, Paris and Berlin)

| Group | Asymmetry | Nb. Runner | % Runner |
|----------------|-----------|------------|----------|
| All | strong | 191 | 70 |
| | weak | 82 | 30 |
| Men | strong | 162 | 70.4 |
| | weak | 68 | 29.6 |
| Women | strong | 29 | 67.4 |
| | weak | 14 | 32.6 |
| < 3 hours | strong | 91 | 70 |
| | weak | 39 | 30 |
| \geq 3 hours | strong | 100 | 69.9 |
| | weak | 43 | 30.1 |
| Paris | strong | 97 | 69.3 |
| | weak | 43 | 30.7 |
| Berlin | strong | 94 | 70.7 |
| | weak | 39 | 29.3 |

Table III.3 – Repartition of runners between strong and weak asymmetry in speed time series regardless of positive or negative asymmetry.

First Table III.3 presents the repartition of runners between strong and weak asymmetries in speed time series regardless of positivity or negativity. We can see that there is, regardless of gender, performance or race profile, about 70% of runners being strongly asymmetric and 30% being weakly asymmetric.

To go further and study the sign of the asymmetry, Table III.4 displays the results for the two categories and the seven groups (All, Men, Women, < 3 hours, \geq 3 hours, Paris, Berlin). In most cases, asymmetry is slightly more negative than positive. This means that runners tend to run more time above their marathon average speed. We can see that for the whole dataset category, 56.4% of the runners have a negative asymmetry. This percentage increase for the strong speed asymmetry category to 67.7% meaning that the weak asymmetry runners tend to have a positive asymmetry.

Only the less than 3 hours group for the whole dataset category has a majority of runners with a positive asymmetry. This could be linked to their performance and the fact that they know better how to run a marathon. This idea is accentuated by the percentage of runners of the greater than or equals to 3 hours group (66.4%) which is the highest for negative asymmetry.

| Group | Asymmetry | Whole Dataset | | Strong Speed Asymmetry | |
|----------------|-----------|---------------|-----------|------------------------|-----------|
| | | Nb. Runners | % Runners | Nb. Runners | % Runners |
| All | + | 119 | 43.6 | 75 | 39.3 |
| | - | 154 | 56.4 | 116 | 67.7 |
| Men | + | 101 | 43.9 | 62 | 38.3 |
| | - | 129 | 56.1 | 100 | 61.7 |
| Women | + | 18 | 41.9 | 13 | 44.8 |
| | - | 25 | 58.1 | 16 | 55.2 |
| < 3 hours | + | 71 | 54.6 | 43 | 47.3 |
| | - | 59 | 45.4 | 48 | 52.7 |
| \geq 3 hours | + | 48 | 33.6 | 32 | 32 |
| | - | 95 | 66.4 | 68 | 68 |
| Paris | + | 54 | 38.6 | 33 | 34 |
| | - | 86 | 61.4 | 64 | 66 |
| Berlin | + | 65 | 48.9 | 42 | 44.7 |
| | - | 68 | 51.1 | 52 | 55.3 |

Table III.4 – Repartition of runners by whole dataset vs. strong asymmetry and positive vs. negative asymmetry in seed time series.

It seems that the place where the marathons have been run as an importance as well. Indeed, Paris has more negatively asymmetric runners (Whole dataset: 61.4%; Strong speed asymmetry: 66%) than Berlin (Whole dataset: 51.1%; Strong speed asymmetry: 55.3%)

The gender does not seem significant enough to give a conclusion on the tendency of speed asymmetry.

To conclude with the asymmetry behavior of the speed time series, we highlighted the following three facts thanks to tables III.3 and III.4:

- 70% of the runners run in a strong asymmetric way regardless of the gender, level of performance and race profile.
- The asymmetry tends to be negative, meaning that runners are spending more time above their marathon average speed than below.
- The asymmetry sign and repartition clearly depends on the level of performance and the race profile but not markedly on the gender.

| Trend | Asymmetry | Men (%) | Women (%) | < 3h (%) | ≥ 3h (%) | Paris (%) | Berlin (%) | All (%) |
|-------|-----------|------------|-----------|-----------|-----------|-----------|------------|------------|
| - | - | 122 (53.8) | 25 (58.1) | 56 (43.1) | 91 (63.6) | 83 (59.3) | 64 (48.1) | 147 (53.8) |
| - | + | 68 (29.6) | 14 (32.6) | 51 (39.2) | 31 (21.7) | 44 (31.4) | 38 (28.6) | 82 (30) |
| + | + | 33 (14.3) | 4 (9.3) | 20 (15.4) | 17 (11.9) | 10 (7.1) | 27 (20.3) | 37 (13.6) |
| + | - | 7 (3) | 0 (0) | 3 (2.3) | 4 (2.8) | 3 (2.1) | 4 (3) | 7 (2.6) |

Table III.5 – Subdivided repartition of the runners in the signature for the whole dataset category by the following seven groups: Men, Women, Less than 3 hours, Greater than or equals to 3 hours, Paris, Berlin, All

| Trend | Asymmetry | Men (%) | Women (%) | < 3h (%) | ≥ 3h (%) | Paris (%) | Berlin (%) | All (%) |
|-------|-----------|-----------|-----------|-----------|----------|-----------|------------|------------|
| - | - | 97 (59.9) | 16 (55.2) | 47 (51.6) | 66 (66) | 64 (66) | 49 (52.1) | 113 (59.2) |
| - | + | 30 (18.5) | 10 (34.5) | 23 (25.3) | 17 (17) | 23 (23.7) | 17 (18.1) | 40 (20.9) |
| + | + | 32 (19.8) | 3 (10.3) | 20 (22) | 15 (15) | 10 (10.3) | 25 (26.6) | 35 (18.3) |
| + | - | 3 (1.9) | 0 (0) | 1 (1.1) | 2 (2) | 0 (0) | 3 (3.2) | 3 (1.6) |

Table III.6 – Subdivided repartition of the runners in the signature for the strong asymmetry category by the following seven groups: Men, Women, Less than 3 hours, Greater than or equals to 3 hours, Paris, Berlin, All

2.3.3 Speed statistical signature

The study of trend and asymmetry of speed time series taken jointly lead to building up the statistical signature described in subsection 2.2.2.3.

The first significant result appears from the following tables III.5 and III.6 giving the four signature frequencies for the seven groups: Men, Women, less than 3 hours, greater-than or equal to 3 hours, Paris, Berlin and All. Table III.5 displays the results for the whole dataset category while table III.6 shows the results for the strong asymmetry category.

It is clear that the repartition of the four signatures is stable. The frequency is significantly decreasing with respect to the ranking $(--)$, $(-+)$, $(++)$, $(+-)$ of the signatures.

From tables III.5 and III.6, the group "All" is the global statistical signature i.e. repartition of the runners by tendency and asymmetry. In both Tables we can see that the $(- -)$ cluster represents more than half of the runners (whole dataset category: 53.8% ; strong asymmetry category: 59.2%) followed by $(- +)$ and $(+ +)$ representing the rest (whole dataset category: 43.6% ; strong asymmetry category: 39.2%). The $(+ -)$ cluster represents a minority of runners. The main difference between the whole dataset and strong asymmetry categories is the interval between $(- +)$ and $(+ +)$. While $(- +)$ is twice larger than $(+ +)$ in the whole dataset category, they are nearly equals in the strong asymmetry category.

In table III.5 we see that the order of the statistical signature for the whole dataset category is the same regardless of gender, level of performance and race profile (- -) (- +) (+ +) (+ -). In table III.6 we see that the order of the statistical signature for the strong asymmetry category is the same regardless of gender and level of performance (- -) (- +) (+ +) (+ -) but changes for the race profile (- -) (+ +) (- +) (+ -). In both categories, the exact repartition of the runners varies.

2.4 Discussion

To the best of our knowledge, this is the first study describing the asymmetric behavior of runners during marathons, asymmetry which appears to be independent of any categorization of runners and races. Additionally we introduced a new statistical signature to analyze the trend and the asymmetry jointly. The main findings of this research are as follow:

1. 70% of the runners tend towards having a strong asymmetric behavior in the speed time series while running a marathon regardless of the gender, level of performance and race profile.
2. The asymmetry tends to be negative, meaning that runners are spending more time above their marathon average speed than below.
3. The statistical signature of a marathon keeps the same order (- -) (- +) (+ +) (+ -) whichever the gender, the level of performance and the race profile.

However, we are aware that in this paper, we compared men and women looking at their absolute performance and it would be interesting to study their relative performance compared to world record as well. Beside this reserve regarding the gender effect, thanks to the statistical signature tool we confirmed the fact that marathon runners cannot maintain a target even pace. They are going to reach a point of fatigue which has been described in literature as a matter of high fractional use of VO₂max (68 to 100% of VO₂max) decreasing during the race [121].

The extreme physical endurance demand has induced a high post-race rectal temperatures due to dehydration and metabolic rate [45, 49] introducing a precipitous rise in body temperature from 40.9 to 41.9°C [122]. The increase in body temperature appears due to the inability of the runner to maintain the well-known "glycogen shunt" necessary for sustaining the muscle rapidity of contraction at the instant of the runner stabilizes. Such rapid contraction requires that the energy is delivered fast so that the muscle has the power requirements of rapid energy expenditure that are ultimately met by a slower averaged consumption of carbon and oxygen from blood [158] as in the last 10 kilometers of a marathon race. Therefore it is the reason why the paradigm of even pace maintained until exhaustion [138, 170], leads to a much lower critical speed than using intermittent speed [130]. This impossible recovery of energy among muscle contraction induces an increase of energy cost of running [82] leading the marathoner to suffer a lack of energy while requirements per kilometer increase. In order to reach an optimal performance, runners have to make decisions about how and when to invest their energy. In other words marathoners have to deal with exercise intensity regulation during the race [159]. Although the physiological regulation of pacing strategy during exercise is still in debate for a marathon, runners tend to adopt a constant speed and seem to try keeping it for as long as possible.

Pacing strategy has been rather described as the efficient use of energetic resources during athletic competition so that all available energy stores are "used before finishing a race, but not so far from the end of a race that a meaningful slowdown can occur" [163]. Most of the time, marathoners are unsatisfied with their race. We distinguish two kind of behavior. First, runners considering that they did not give their best (increasing trend in speed time series). Second, runners being completely exhausted due to the famous "marathon wall" [53]. The present study showed that the non optimal employed pacing strategy in non elite marathon runners was independent of gender, level of performance and marathon profile. This goes into the same direction as a prior study done on the recent world record marathon runners using also a statistical analysis of high frequency (1 km) split data [6]. This last study may conclude that for realistic marathon conditions (i.e. no treadmill), an approximation to

an optimal pacing strategy should rest on two principles:

- Principle 1 (the even principle). Where on-course wind and/or gradient variations are insignificant, minimize variations from average power.
- Principle 2 (the parallel principle). Where on-course wind and gradient variations exist, apply small (0-10%) variations in power (speed) in parallel to environmental variations to maintain a constant running speed (i.e. exert more (less) power uphill (downhill) or into head-winds (tail winds)).

This assumption of the ideal race at constant power [78, 123, 177, 162, 144, 140] is based on the the 1973 Keller's model identifying a runner with a Newtonian particle. However, a prior study has interestingly reported that older runners have a still more even slow pace than the younger one. Here we did not focus on the age effect, but considering that running performance has a structural basis [172] and that muscle mass decreases with age, older runners cannot bear speed variation due to their narrow metabolic scope associated with their loss of strength that is amplified during a marathon which is often described as a "second (slow) death" by the runners at the arrival.

The future of marathon performed with the integrity of human homeostasis is now more and more considered to be associated with the concept of self-paced run rather than having to follow a constant target speed [169, 89, 95]. This strategy is adopted in elite runners especially because they have a sufficient metabolic scope [138], a high speed reserve between average marathon and 1000 meters paces [30]. As underlined above oscillations in energy metabolism is associated with the recovery of phosphocreatine pattern necessary for maintaining the muscle strength to take off the ground at each step [94]. Therefore, the future of performance while running a marathon will probably be associated with a new paradigm of dissipative system with dynamic instabilities. The visible part is the speed variation that can not be any more considered as a mistake but rather as an intelligent fit of metabolism adaptation in real time. That is why, here we propose a "qualitative" approach of the marathon race beyond the only recognized performance criteria: the time. The performance, beyond the final time, is also a way to succeed a marathon. A qualitative analysis could be a necessary

step to break the two-hour marathon which has been a real recurrent issue for almost 10 years [98, 164].

2.5 Conclusion

Running a marathon at a perfect even pace seems nearly impossible. The research on that matter could explore other pacing strategies and apply them in real conditions in order to have a proper validation. We think that speed variation is the optimal way of running but we still need to find a way to build custom-made strategies by clustering by gender, level of performance or race profile for instance. Meanwhile we can think of two research opportunities to go further regarding this paper. One would be to try to find universal speed patterns for marathons and see if these match the statistical signature. This could be the first step for clustering runners with the idea of predicting pacing strategies in future works. The other one would be to study heart rate trend and asymmetry and then combine the results with those from this paper. The cardiac cost could help us to link both speed and heart rate time series at the same time.

3 [Article | En Soumission] The 2018 marathon world record shows the positive asymmetry pacing strategy effectiveness

Abstract

Recently, Kenyan Eliud Kipchoge lowered the marathon world record by more than one minute in Berlin. The Olympic champion already made an impression 6 months earlier while trying to complete the first unofficial sub two-hour marathon in Monza. The difference between the two marathons is that in Monza, the instructions were to run at a constant pace whereas in Berlin the speed was significantly variable. The purpose of this study is to analyze the running pattern to achieve this new best performance beyond the well-known negative

split with a more accurate analysis of the speed time series. More precisely, instead of trying to keep a constant pace until fatigue like mass runners would do, Eliud Kipchoge ran in what we named a positive asymmetric way. One of the feature of this asymmetry in the speed time series is that he ran 66% of the distance at a speed less-than or equal to his marathon average speed. Indeed, he started and finished very fast. The 6 minutes, 8 seconds that Kipchoge ran from 40k to the finish line is the fastest known speed in any major marathon. The strategy of running with both fast start and fast finish is made possible thanks to a high speed reserve (i.e. difference between the critical speed and the average marathon speed).

Keywords

Marathon; Kipchoge; Strategy; Running

3.1 Introduction

Eliud Kipchoge lowered the world record in the 2018 Berlin marathon in 02h01min39s with a variable pace (see figure III.2). He established this new world record by running 1 minute and 18 seconds less than the previous one. The Kenyan made history by breaking the existing world record with the biggest margin in more than 50 years (Former record holder Derek Clayton broke it by 2 minutes and 24 seconds in 1967). Eliud Kipchoge started the race with three pacers working as shields in front of him. The first one dropped out after 14k and the second one after 16k. The last pacer brought him to 25k and then Eliud Kipchoge finished the race by his own. In reality this was not Eliud Kipchoge's fastest marathon. During the unofficial Nike Breaking2 marathon he ran in 02h00min25s in Monza, Italy.

One remarkable fact is that he ran from 40k to the finish line in 6 minutes, 8 seconds setting in the meantime the fastest known finish for a major marathon. The finish contributed to the negative split. The negative split way of running raises the unresolved question of the racing strategy again.

The paradigm stating that the best way of running is by adopting an even pace still pervades theory and practice. This even pace based on the belief that constant speed is the best way

of running was used for the sub two-hour marathon attempt in Monza. This same idea is relayed by marathon organizers who offer pace-group leaders to help the runners to maintain a target constant race speed. Most of the time, marathon analyses are not personalized but rely on the data of all the runners [78, 116, 5, 150, 58]. We wanted to take a deep look at the 2018 marathon record because we think that the combination of both personalized and overall analyses are necessary to improve our understanding of running.

In the 2018 MIT workshop [34] we presented a possible ideal pacing strategy for running a marathon. This way of running is positively asymmetric in the sense introduced in the previous paper [24]. One particular case is to run with a bathtub shape strategy by spending 2/3 of the race under the average speed (cf. figure III.2). In the case of negative asymmetry, runners spend much more time and run a longer distance above the overall marathon average speed. Therefore, the aim of this study is to test the hypothesis that this new marathon performance was run with a positive asymmetry and try to understand the reproduction of asymmetric patterns. In addition, we check that the positive asymmetry could be performed considering the high speed difference between his critical speed (computed with his personal bests from 1,500m to 10,000m, cf. table III.7) and his average marathon speed [28, 30].

The exercise intensity that defined the critical speed is the speed that could be sustained almost indefinitely (subcritical threshold) and above which the exercise duration (supra-critical threshold intensities cf. [28]) is restricted. Accordingly, the amount of work performed above the critical speed is referred to as the anaerobic distance capacity (ADC) and provides an estimate of the capacity to complete the anaerobic work [59].

Furthermore, very recently, the anaerobic speed reserve has been reconsidered as a condition to be able to run a fast finish and perform a final speed on distance from 1,500 to 10,000 meters [149]. This seems to be the case for modern marathon races also.

The purposes of this study are, first, to use the new positive asymmetry paradigm to the new marathon world record by analyzing the whole race, the race split in two half marathons segments and the race split in four 10 kilometers segments and, second, thanks to this positive asymmetry strategy, to examine the possible anaerobic distance capacity reserve of Eliud

Kipchoge for a future new record.

3.2 Materials and Methods

3.2.1 Methodology

To achieve this study, we examined the time per kilometer ran by Eliud Kipchoge. Unofficial splits data were collected from the SCC Events website³. At the time of this study, we could not find official data for the new world record. As SCC Events was the organizer of the 45th BMW Berlin-Marathon in 2018, those were the most accurate data that have been made public.

We computed the average speed per kilometer by dividing the distance per time.

3.2.2 Statistical study

3.2.2.1 Characterizing the trend In order to detect the trend (increasing or decreasing) of speed during the marathon we used the Kendall's τ [103, 3]. Kendall's τ is a non-parametric coefficient computed from a times series v_1, v_2, \dots, v_n . Here v_i is the i -th value of speed, $i < j$ means that i indicates a period of time prior to j . For instance, the speed v increases with time if $v_1 < v_2 < \dots < v_n$.

Consider the symmetric kernel

$$K(v_j, v_i) = K(v_i, v_j) = \begin{cases} 1 & \text{if } i < j \text{ and } v_i < v_j, \\ 0 & \text{if } v_i = v_j, \\ -1 & \text{if } i < j \text{ and } v_i > v_j. \end{cases} \quad (\text{III.3})$$

$$\tau = \frac{2}{n(n-1)} \sum_{i < j} K(v_i, v_j) \quad (\text{III.4})$$

Kendall's τ is defined by Equation III.4, the sum being performed over the $n(n-1)/2$ distinct unordered couples of indices $\{i, j\}$, so that τ takes values in between $[-1, 1]$. The Kendall's τ

³http://results.scc-events.com/2018/?pid=leader_public

equals 1 (respectively -1) only if the speed time series is continuously increasing (respectively decreasing). Furthermore a constant speed time series leads to the value $\tau = 0$. We know from the field of mathematical statistics that Kendall's τ is suited to detect a decreasing or increasing linear trend in a time series. It enjoys the qualities of being robust and non-parametric [102].

3.2.2.2 Characterizing the asymmetry The asymmetry of the time series $(v_i)_{1 \leq i \leq 42}$ will be specified by figuring out the distance spent below vs. distance spent above the marathon average speed. We computed the percentage of distance spent above the marathon average speed for Eliud Kipchoge. Recall from [24] that we consider that runners having a percentage of distance spent above (respectively under) average speed greater than 50% are negatively (respectively positively) asymmetric. The asymmetric coefficient defined above is easy to be used and is mathematically well-established and robust statistically.

3.2.2.3 Combining trend and asymmetry Finally we studied trend and asymmetry jointly. We used the statistical signature [24] to describe the 2018 Berlin marathon speed time series for Eliud Kipchoge. The signature is made of two signs (\pm, \pm) . Thus it takes four possible values, $(++)$, $(+-)$, $(-+)$ and $(--)$. The first sign is $+$ (respectively $-$) according to whether the Kendall τ is positive (respectively negative). The second sign is $+$ (respectively $-$) according to whether the asymmetry is positive (respectively negative).

The idea behind the use of the statistical signature is to try to describe the new marathon world record by studying the three following point:

- Trend, asymmetry and statistical signature of the entire race (only the first 42 kilometers have been taken into account for simplicity reasons).
- Trend, asymmetry and statistical signature of the race split in two 21 kilometers parts (only the first 42 kilometers have been taken into account for simplicity reasons).
- Trend, asymmetry and statistical signature of the race split in four 10 kilometers parts (only the first 40 kilometers have been taken into account for simplicity reasons).

| Date | Place | Distance (m) | Time (s) |
|--------------|------------------|--------------|----------|
| 31 May 2004 | Hengelo (NED) | 1,500 | 213.2 |
| 20 July 2004 | London (GBR) | 1,609 | 230.4 |
| 06 May 2011 | Doha (QAT) | 3,000 | 447.6 |
| 04 June 2005 | Eugene, OR (USA) | 3,200 | 487.6 |
| 02 July 2004 | Roma (ITA) | 5,000 | 766.5 |
| 26 May 2007 | Hengelo (NED) | 10,000 | 1,606.0 |

Table III.7 – Personal all time bests for Eliud Kipchoge.

3.2.3 Critical Speed

After the statistical study, we will compare the running strategy (i.e. the shape of the speed time series) of Eliud Kipchoge with his critical speed [42, 52]. In order to compute Eliud Kipchoge’s critical speed we used his personal bests for distances from 1,500 to 10,000 meters. Table III.7 displays these values.

The critical velocity was calculated by fitting the equation III.5 [59]:

$$D_{lim} = \alpha + \beta * t_{lim} \quad (III.5)$$

using an ordinary linear least squares method, where α is considered to be the anaerobic distance running capacity in meters, and the slope β stands for the critical speed [92]. Figure III.1 shows the results. The critical speed is equal to 6.10 m/s (i.e. 21.97 km/h) and the anaerobic capacity is equal to 235 m .

In the results, we will look at the difference between the theoretical distance above the critical speed and the theoretical anaerobic distance capacity. In other words, we will look at the maximal anaerobic capacity (d_{cs}) and the actual distance spent above the critical speed (d_{acs}) keeping the same time interval. The obtained value (Δ_d) is compared to the anaerobic capacity. A value less than the anaerobic capacity [52] indicates a potential way to improve the performance especially with a variable pace allowing recovery phase as reported for the critical speed model for intermittent exercise [130]. A negative value is nearly impossible.

The athlete has been able to spare his anaerobic reserve probably because of his ability to maintain a high $\dot{V}O_2$ steady state even when he ran close to his critical speed as previously

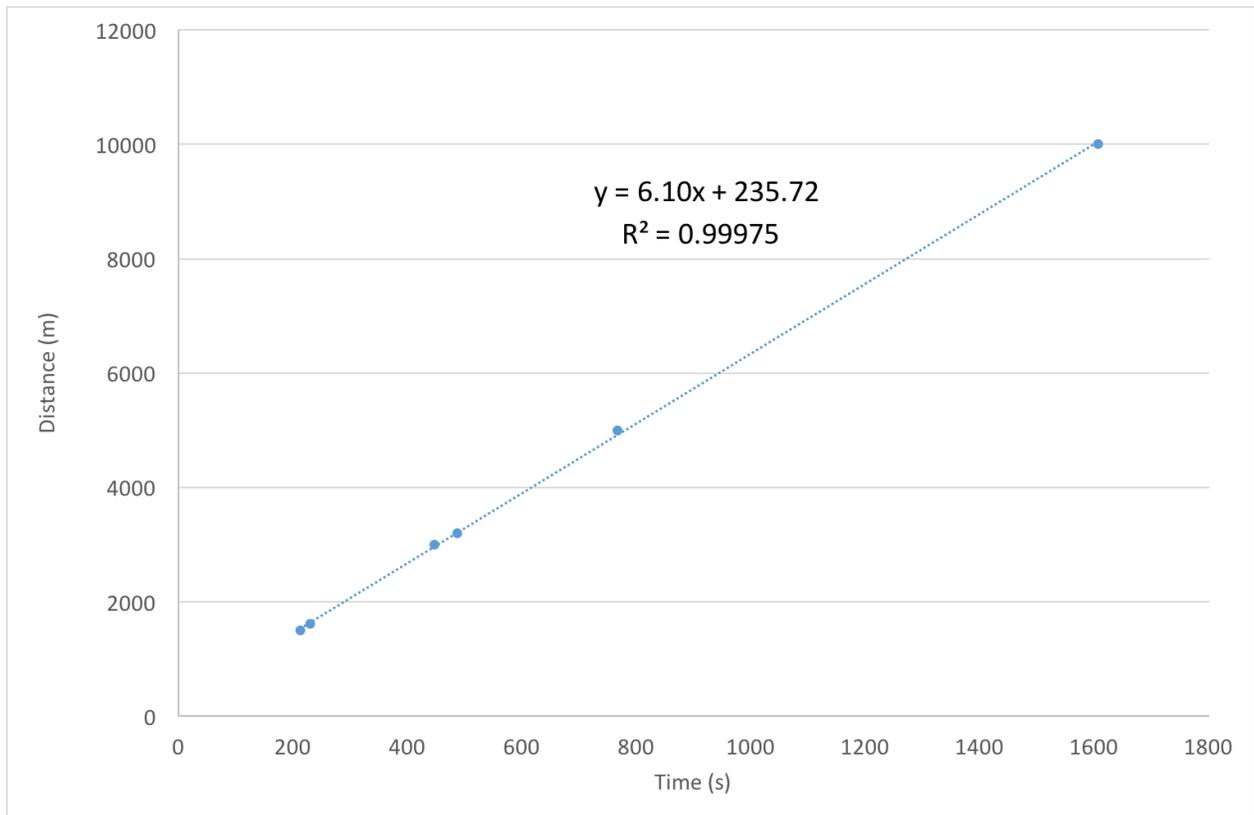


Figure III.1 – Critical speed and anaerobic capacity of Eliud Kipchoge. Linear regression computed from 1,500 m to 10,000 m.

reported in high level runners [29].

3.3 Results

3.3.1 Statistical study of the whole marathon

First we focus on the entire marathon minus the last 195 meters to get a time series given per kilometer.

By studying the trend, we see that the final Kendall's τ value is 0.28 meaning that there is an increasing trend. In other words, the speed of Eliud Kipchoge is globally increasing during the marathon.

The second step is to study the asymmetry. By computing the distance spent above and under the marathon average speed, we found out that Kenyan ran 54 % of the marathon under his average speed and 66 % at a speed less than or equal to his average speed. This is

the consequence of a positive asymmetry with 23 segments ran under, 14 segments ran above and 5 segments ran at the average speed.

Figure III.2 shows two series. The first one is a bar chart representing the speed time series. Light (respectively dark) bars are for the kilometers spent above (respectively under) the marathon average speed. The second is the average speed of the race.

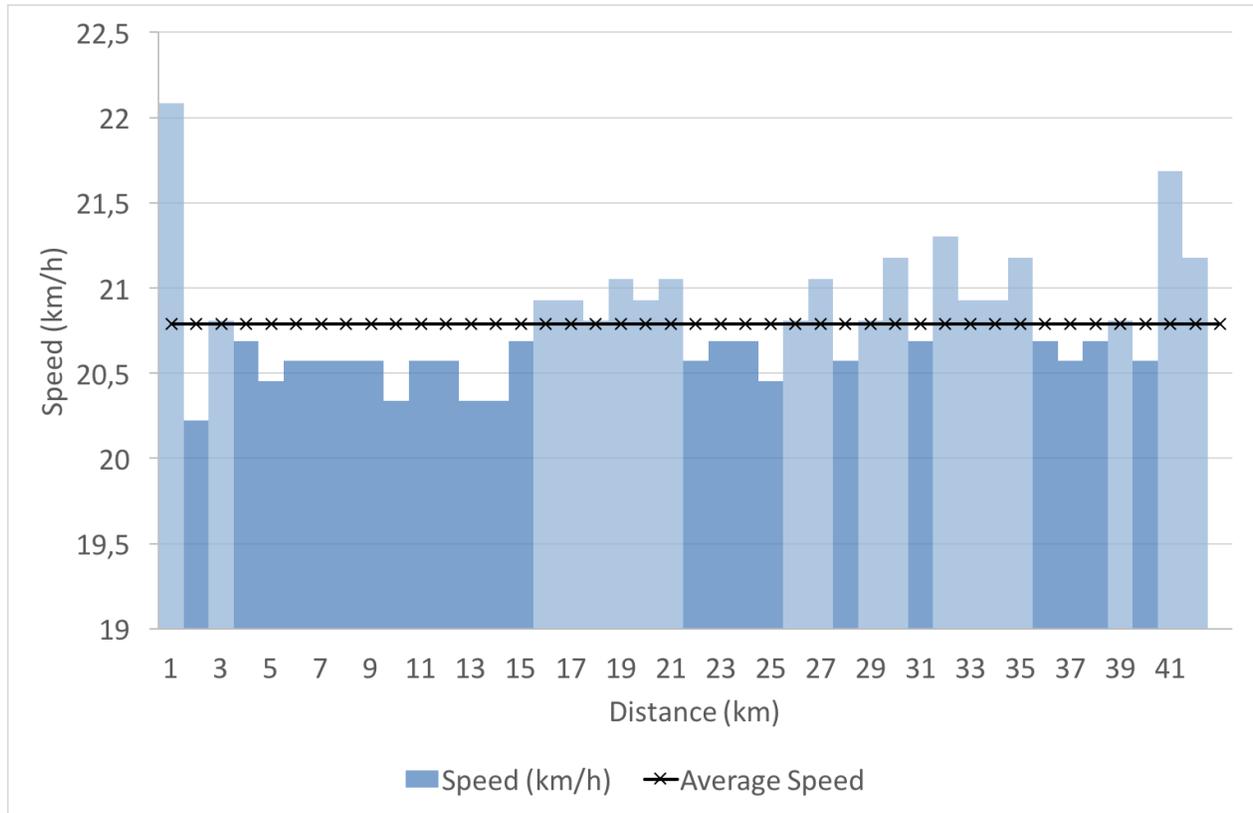


Figure III.2 – 2018 marathon world record. The bar chart represents the speed time series. The line chart stands for the average speed on the segment. Speed values are computed for each kilometer from pace values.

Finally, when joining the trend and asymmetry, Eliud Kipchoge has a (+ +) signature.

In order to go deeper in the analysis of the race, we decided to split the marathon into two equal parts of 21 kilometers each.

3.3.2 Statistical study of the marathon split in two

We now split the marathon into two equal parts of 21 kilometers. The first 21 kilometers will be identified as the first half-marathon and the next 21 kilometers as the second half-

| Part | Average speed | Kendall's τ | Number of km under/above average speed | Signature |
|------------------|---------------|------------------|--|-----------|
| From 1 to 21 km | 20.74 | +0.29 | 13/8 | (+ +) |
| From 22 to 42 km | 20.87 | +0.24 | 13/8 | (+ +) |

Table III.8 – Average Speed, Kendall's τ , number of kilometers under and above average speed and signature (see in text) for each half-marathon of the 2018 marathon world record.

marathon.

By computing the Kendall's τ , we found a value of 0.29 for the first half-marathon and a value of 0.24 for the second half-marathon. These results highlight the fact that in both cases the speed of the runner is globally increasing.

By computing the asymmetry, we notice that Eliud Kipchoge ran 13 of the 21 kilometers under his half-marathon average speed (i.e. 62 %) for both half-marathons. The average speed is 20.74 km.h^{-1} for the first half-marathon and 20.87 km.h^{-1} for the last one. The Kenyan ran his second half-marathon faster than the first one (i.e. negative split). Table III.8 presents the results of average speed, Kendall's τ , number of kilometers under and above average speed and signature for each half-marathon. Another interesting fact is that the runner has more variations in his speed in the second half-marathon than in the first one. Figure III.3 presents two pictures. The first half-marathon is on the left-hand side (III.3a) and the second half-marathon is on the right hand side (III.3b). On both plots, the first series is a bar chart representing the speed time series. Light (respectively dark) bars are for the kilometers spent above (respectively under) the marathon average speed. The second series is the average speed of the half-marathon.

Both half-marathons have a positive asymmetry and a positive trend giving a (+ +) signature.

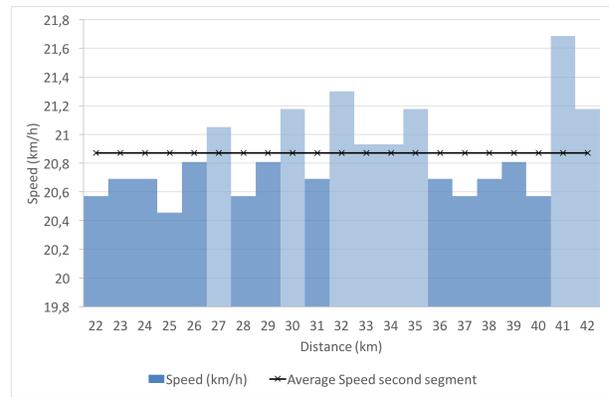
We are now going to take the analysis one step further by splitting the marathon into four equal parts.

3.3.3 Statistical study of the marathon split in four

In order to make the computation easier we focus on the first 40 kilometers of the race. The first part gathers the data from 1 to 10 kilometers, the second part gathers the data from



(a) First half of the 2018 marathon world record



(b) Second half of the 2018 marathon world record

Figure III.3 – 2018 marathon world record. On the left-hand side is the first half of the marathon. On the right-hand side is the second half. Bar charts represent the speed time series. Line charts stand for the average speed on the segment.

| Part | Average speed | Kendall's τ | Number of km under/above average speed | Signature |
|------------------|---------------|------------------|--|-----------|
| From 1 to 10 km | 20.69 | -0.36 | 8/2 | (- +) |
| From 11 to 20 km | 20.74 | +0.61 | 5/5 | (+ 0) |
| From 21 to 30 km | 20.82 | +0.26 | 7/3 | (+ +) |
| From 31 to 40 km | 20.85 | -0.42 | 6/4 | (- +) |

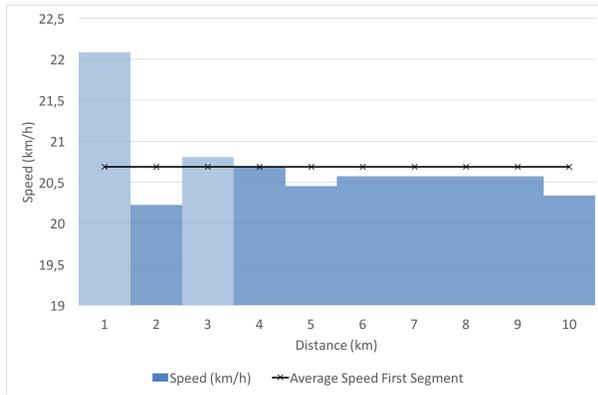
Table III.9 – Average Speed, Kendall's τ , number of kilometers under and above average speed and signature for the four segments (10km each) of the 2018 marathon world record.

11 to 20 kilometers, the third part gathers the data from 21 to 30 kilometers and the fourth part gathers the data from 31 to 40 kilometers.

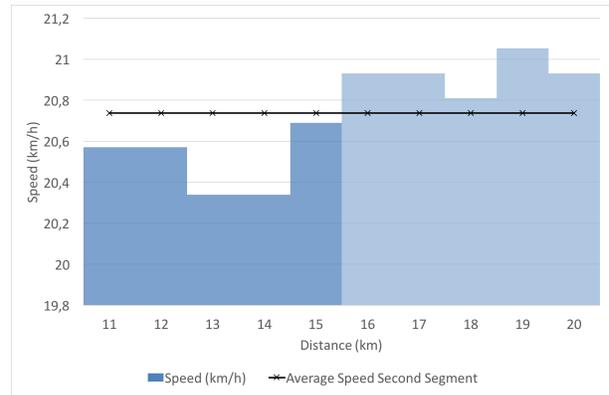
Table III.9 shows the results for the computation of the average speed, the Kendall's τ , the asymmetry and the signature for the four segments. These results highlight the three following facts. The first one is that the average speed of the runner increase continuously from the first to the last segments. The second one is the presence of speed variations during the race. From 1 to 10 km and from 31 to 40 km the runner has a decreasing speed. On the other hand, from 11 to 20 km and from 21 to 30 km, the runner has an increasing speed. The last one is an asymmetry always positive except for the second segment where the asymmetry is null. Turning the second symmetric segment to an asymmetric one could be a way of improvement.

Figure III.4 shows the four parts of the newly established marathon world record. Each plot displays two series. The first one is a bar chart of the speed and the second one is the

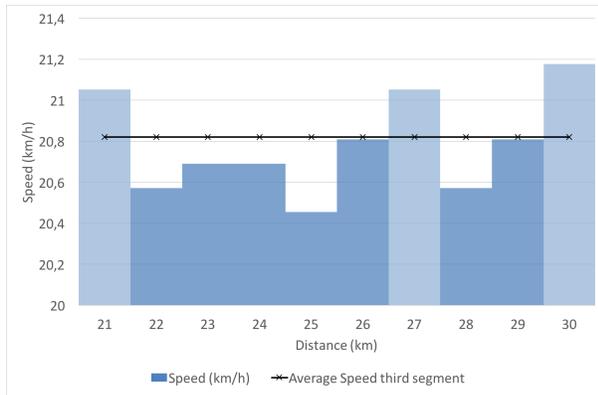
average speed of the segment in question.



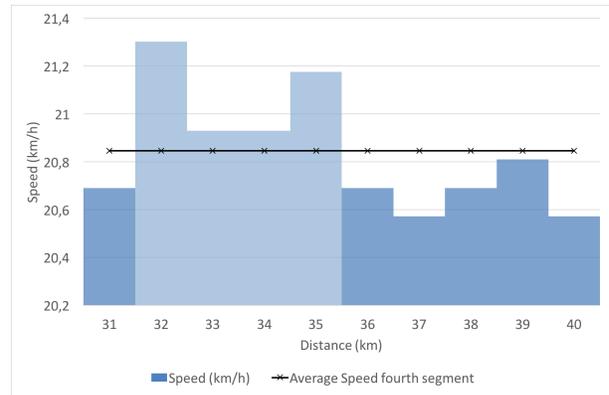
(a) First 10k of the 2018 marathon world record



(b) Second 10k of the 2018 marathon world record



(c) Third 10k of the 2018 marathon world record



(d) Fourth 10k of the 2018 marathon world record

Figure III.4 – 2018 marathon world record. Top left is the first 10k of the marathon. Top right is the second 10k. Bottom left is the third 10k. Bottom right is the fourth 10k. Bar charts represent the speed time series. Line charts stand for the average speed on the segment.

3.3.4 Critical Speed

The runner spent just the first kilometer above his critical speed. This implies that $d_{acs} = 1000$ m and that the runner spent 163 s above his critical speed. We now compute the theoretical distance if the runner ran at his critical speed for 163 s. We found that $d_{cs} = 994.7$. We deduce that $\Delta_d = 5.3$ m. By comparing Δ_d and the anaerobic capacity we clearly see a great difference between those two values. A way of improvement for the runner could be to spend more time above his critical speed during the marathon.

3.4 Discussion

This study analyses for the first time the newly established marathon world record under the prism of asymmetry. We could find lots of publications speaking about pacing variability in running endurance events. The authors generally focus on a group of runners instead of a single individual. For instance, we note a 1 to 5 % of variability in speed for 3,000 m to 10,000 m for competitive runners [32], 3 % on marathons [78], 5.4 % for the top ten finishers of a 100 km ultra-marathon [110]. The speed variation coefficient of Eliud Kipchoge is 2.2 % giving a relatively even pace strategy. Behind this speed stability we looked for the asymmetry of the speed time series of the runner. This is the first study showing that it may be possible to establish a world record by running below the average pace thanks to a very fast start and final.

From Maron and al. [121], we know that the marathon should not be a constant speed race. They showed that at 23.4 Miles, a rather good runner (2h36) reached 100 % of his $\dot{V}O_{2max}$. Also, they showed that the oxygen uptake was between 68 and 100 % of $\dot{V}O_{2max}$ (in the prior and last Miles) due to a final acceleration as in the 2018 marathon world record.

The difference of performance between high level marathon runners (2h11 - 2h16) and elite marathon runners (less-than 2h11) is known to be the ability to run fast on 1,000 m rather than $\dot{V}O_{2max}$. $\dot{V}O_{2max}$ was not discriminant for distinguishing these two levels of performance [30].

With the idea of optimizing the pacing strategy, we think that the comparison of the average speed with the critical speed is also interesting. We noticed that the average speed represents 94 % of the critical speed. The value is close to the personal best one-hour record meaning that the runner is able to double his time limit by only decreasing his speed by 5 %. Accordingly, the amount of work performed above critical speed is referred to as the anaerobic distance capacity (ADC) and provides an estimate of the capacity to complete anaerobic work. Specifically, running speed over 800- and 1,500-m races has been shown to be regulated by the prevailing anaerobic (oxygen independent) store (estimated by the ADC) at each instant of the race up until the all-out phase of the race over the last several

meters [21]. High-level marathoners have been shown to be able to run at $\dot{V}O_2max$ during some phase of the race [121] and to have high performance even on 1,000m [30]. Finally, another way of increasing the performance would be for the runner to use a bigger part of his anaerobic capacity. He used only 2 % of his anaerobic capacity during the 2018 marathon world record.

3.5 Conclusion

Among the simple analyses stating that Eliud Kipchoge ran negative splits, we wanted to describe this newly established marathon world record in depth. The final idea is to find an optimal tailor-made running strategy for every marathon runner. By studying the running strategy of world records, we think that this could lead us to a better understanding of what is a good way of running.

The main findings of this study are as follow:

- The analysis of the trend seems to show that the speed of Eliud Kipchoge increased along the race. The average speed of the second 21km is greater than the first one. Even when we split the race in four parts the average speed of the next 10km is always greater than the previous ones.
- Eliud Kipchoge spent 66 % of the total distance at a speed less-than or equal to the final average speed.
- Each of the two 21km shows the same pattern of asymmetry. The runner spent 62% of the distance at a speed less-than or equal to the average speed of the 21km.
- By splitting the race in four parts of 10k, we notice that three of them are asymmetric except the second 10km that is symmetric.

In addition with these four findings, we identified potentially two ways of improvement. The first one would be the reproduction of an asymmetric pattern. This raise the question of what could be the best asymmetric pattern for optimizing the performance and which is

the best period for the pattern to repeat. The second way would be the use the anaerobic capacity. By spending more time above his critical speed Eliud Kipchoge could improve his performance and perhaps break the 2-hour marathon barrier.

Acknowledgements

The authors would like to thank Dr. Jean Pierre Koralsztein for his advice and for rereading this article.

Disclosures

No conflict of interest, financial or otherwise are declared by the authors.

Bilan

Lors de ce chapitre, nous avons mis en place la première étape en vue de la création d'une stratégie de course personnalisée. Par l'analyse de données de vitesse en fonction du temps (ou de la distance), nous avons essayé de comprendre les différentes façons de courir un marathon. Notre étude s'est déroulée en deux temps. Premièrement, nous avons étudié les données de 273 marathons, issues des marathons de Berlin et Paris 2017. Deuxièmement, nous nous sommes concentrés sur le récent record du monde établi à Berlin par Eliud Kipchoge. Lors de ces deux travaux, nous avons introduits le concept de signature statistique. Il s'agit d'un couple de signe (\pm, \pm) où le premier signe correspond au signe du τ de Kendall et le second au signe de l'asymétrie. Le τ de Kendall nous permet de caractériser la tendance de la course. Un τ de Kendall positif indique que le coureur suit une vitesse globalement croissante. A l'inverse, un τ de Kendall négatif indique une vitesse décroissante. L'asymétrie, quand à elle, est définie comme positive si le coureur passe plus de temps (ou distance) en dessous de sa vitesse moyenne sur le marathon, négative sinon.

Les conclusions du premier article scientifique (c.f. [2](#)) [[24](#)] sont les suivantes :

- 70% des coureurs semblent avoir une forte asymétrie dans leur vitesse de course sur

marathon, indépendamment du genre, niveau de performance et profil de course (i.e. marathons de Berlin ou Paris).

- L'asymétrie est principalement à tendance négative. Cela signifie que les coureurs passent plus de temps au dessus de leur vitesse moyenne qu'en dessous.
- La signature statistique sur marathon conserve le même ordre $(--)$ $(-+)$ $(++)$ $(+-)$ indépendamment du genre, niveau de performance et profil de course.

A la vue des résultats, nous pensons que les signatures $(-+)$ et $(++)$, c'est-à-dire, celles qui donnent une asymétrie positive, sont les plus performantes. En effet, il est préférable de passer plus de temps en dessous de sa vitesse moyenne qu'au dessus. Le coureur se fatigue moins et peut ainsi prétendre à une meilleure performance finale. D'autre part, nous pensons qu'il est possible de diviser un marathon en plusieurs parties égales et de reproduire la signature statistique sur chaque partie afin de vérifier si la stratégie utilisée lors d'un marathon a été optimale. Le second article scientifique (c.f. 3) [145] étudiant le nouveau record du monde semble aller dans ce sens.

Nous avons créé un outil capable de déterminer la qualité de la stratégie de course à posteriori. Nos futures recherches vont désormais se concentrer sur un outil de prédiction qui donnerait un temps de passage au kilomètre par exemple. Encore mieux, nous pensons qu'il est possible de prévoir la stratégie de course en temps réel en partant de la forme du jour du coureur. Nous croyons qu'il est possible d'identifier, chez les coureurs, des états physiologiques (e.g. vitesse, fréquence cardiaque, fréquence respiratoire, ...). L'historique de ces états physiologiques depuis le début de la course, couplé aux informations détenues sur le coureur avant la course nous permettrait de demander aux coureurs d'accélérer ou de ralentir en temps réel.

Chapitre IV

Troisième contribution :

Caractérisation de l'accélération d'un coureur à pied

Le quatrième chapitre nous a permis de réfléchir à une possible variable de transition intervenant lorsque le coureur change de vitesse ou d'état physiologique. L'idée de la signature de vitesse est de donner des valeurs de vitesses cibles à chaque kilomètre et des transitions pour atteindre ces valeurs. Nous pensons que l'accélération est à même de réaliser ce rôle. Nous allons présenter nos travaux sur la capacité de l'être humain à distinguer trois différentes intensités d'accélération ("facile", "moyenne" et "dure") qui pourraient intervenir dans la transition entre chaque vitesse cible.

Sommaire

| | | |
|----------|---|------------|
| 1 | Problématique et enjeux | 117 |
| 2 | [Article] Humans are able to self-paced constant running acceleration until exhaustion | 118 |
| 2.1 | Abstract | 118 |
| 2.2 | Introduction | 119 |
| 2.3 | Methods | 120 |

| | | |
|-------|---|------------|
| 2.3.1 | Subjects | 120 |
| 2.3.2 | Experimental Design | 121 |
| 2.3.3 | Acceleration trials | 121 |
| 2.3.4 | Mathematical Model | 122 |
| 2.4 | Results | 124 |
| 2.4.1 | Comparison of the three different acceleration exercises | 124 |
| 2.4.2 | Characterizing the human faculty to maintain a constant acceleration | 126 |
| 2.5 | Discussion | 132 |
| 2.5.1 | The self-paced acceleration range values compared with current use of acceleration protocols in laboratory | 133 |
| 2.5.2 | The possible mechanisms allowing humans to control its body acceleration in accordance with prior results obtained on manual task | 134 |
| 2.6 | Conclusion | 134 |
| | Bilan | 136 |

1 Problématique et enjeux

Nous rappelons au lecteur que la stratégie de course, comme nous la concevons, ne doit pas être une vitesse constante que l'on essaye de maintenir sur toute la course mais un enchaînement de vitesses variées (c.f. figure I.19). Deux questions se posent alors, quand changer de vitesse et comment le faire ? Nous allons essayer de répondre au "comment" dans ce chapitre.

L'hypothèse est la suivante : nous pensons que l'être humain peut contrôler assez aisément son accélération, contrairement à sa vitesse. L'idée est de vérifier cette hypothèse en proposant aux coureurs d'accélérer progressivement jusqu'à épuisement et regarder si l'accélération est constante pendant tout l'effort. Les coureurs devront reproduire trois fois ce test selon trois

intensités d'accélération différentes, "facile", "moyenne" et "dure". L'article scientifique (c.f. [section 2](#)) [23] a été écrit dans le but de valider l'hypothèse selon laquelle, l'être humain est capable de distinguer au moins trois intensités d'accélération différentes dont les valeurs sont constantes.

2 [Article] Humans are able to self-paced constant running acceleration until exhaustion

2.1 Abstract

Although it has been experimentally reported that speed variations is the optimal way of optimizing his pace for achieving a given distance in a minimal time, we still do not know what the optimal speed variations (i.e accelerations) are. At first, we have to check the hypothesis that human is able to accurately self-pacing its acceleration and this even in a state of fatigue during exhaustive self-pacing ramp runs. For that purpose, 3 males and 2 females middle-aged, recreational runners ran, in random order, three exhaustive acceleration trials. We instructed the five runners to perform three self-paced acceleration trials based on three acceleration intensity levels: "soft", "medium" and "hard". We chose a descriptive modeling approach to analyse the behaviour of the runners. Once we knew that the runners were able to perceive three acceleration intensity levels, we proposed a mean-reverting process (Ornstein-Uhlenbeck) to describe those accelerations: $da_t = -\theta(a_t - a)dt + \sigma dW_t$ where a is the mean acceleration, a_t is the measured acceleration at each time interval t , θ the ability of the runner to correct the variations around a mean acceleration and σ the human induced variations. The goodness-of-fit of the Ornstein-Uhlenbeck process highlights the fact that humans are able to maintain a constant acceleration and are able to precisely regulate their acceleration (regardless of its intensity) in a run leading to exhaustion in the range from 1 minutes 36 seconds to 20 minutes.

Keywords

Acceleration; Exercise; Running; Self-paced; Performance; Ornstein-Uhlenbeck

2.2 Introduction

Endurance running is considered to have played a major role in human evolution. Humans have developed the ability to fine-tune their running speed in order to run for several days and still catch their fastest prey [39]. Indeed, it has been reported that speed variation is the ideal way of optimizing its pace for achieving a given distance in a minimal time [62, 63, 51, 31, 36, 148, 166, 163]. Hence, the rate at which the velocity of a body changes with time requires positive and negative accelerations and then strength variations according to Newton's second law of motion. However, there is a direct relationship between force impulse and running acceleration [93] and the minimum-jerk model [61] predicts that to save energy and optimize performance, running must be as smooth as possible and variations in acceleration must be close to $0 \text{ m}\cdot\text{s}^{-3}$. Furthermore, there is a clearly non-random speed fluctuation during self-paced exercise [165, 135, 1, 51]. It has been suggested that speed variability operates through a feed forward mechanism [7] and we know that most of the races are stochastic when the velocity is not imposed [22]. Despite the large published body of work on pacing strategy and speed control, this is the first study to have examined acceleration control during running. We know that the fastest runners on middle and long-distance run with light speed variations [31, 57, 78]. Now that sensors make the recording of acceleration data easy, we would like to check if humans are able to integrate instructions in acceleration. The purpose is to take part in a runner's training and race in future studies. Given the fact that we are looking for a stochastic process to describe the ability for humans to maintain a constant acceleration and the hypothesis that self-paced accelerations lightly fluctuate around a constant value, we thought about the Ornstein-Uhlenbeck process [167]. Such process is described as a stochastic process with a mean-reverting property. The use of Ornstein-Uhlenbeck process is present in fields such as mathematical finance [153] (volatilities

of asset prices and dynamics of interest rates) and biological processes [146]. Before going deeper in the effective mechanism of speed variation control during exhaustive exercise, we have to verify human's ability to maintain constant acceleration in a conscious way until exhaustion. The present study tests the hypothesis whereby humans are able to maintain a constant acceleration in a self-paced trial with no external cues regardless of the speed and the magnitude of acceleration and fatigue. For that purpose, we study more specifically the following two points:

(1) The ability for humans to maintain three different intensity levels of constant acceleration when so instructed. Those three intensities are described as "soft", "medium" and "hard" which are the instructions we gave to the runners for respectively a slow, a medium and a hard constant acceleration.

(2) The characterization of the acceleration data by introducing a new interpretation using a mean-reversing process (Ornstein-Uhlenbeck).

The paper is organized as follows. Section 2.3 describes the experimental protocol and presents the mathematical model. Section 2.4 compares the three acceleration intensities for each runner to see if we notice a genuine difference and gives an estimation of the parameters of the Ornstein-Uhlenbeck model for characterizing human accelerations. Finally, section 2.5 and 2.6 conclude the paper.

2.3 Methods

2.3.1 Subjects

The study population is composed of three males and two females recreational runners (aged 38 ± 3 yrs., total running distance per week: 36.1 ± 4.3 km; body weight: 66.9 ± 12.4 kg and height 171.1 ± 6.7 cm). All subjects were first informed of the risks and constraints associated with the protocol and gave their written, informed consent to participation. The present study conformed to the precepts of the Declaration of Helsinki and all procedures were approved by the local investigational review board (Saint Louis Hospital, Paris, France). On a physiological aspect, we cannot compare openly males and females. Given this fact plus

the sample size of the study, we choose to focus on a descriptive modeling approach to analyse the runner behaviours. No statistical tests will be applied.

2.3.2 Experimental Design

Subjects ran alone and performed three types of track sessions until exhaustion, in random order and with a one-hour interval between sessions: the first, second and third track tests were self-paced acceleration trials at respectively slow, medium and high accelerations. We characterize those accelerations as "soft", "medium" and "hard". Velocity and acceleration were measured with a GPS-enabled accelerometer of 10 and 100 Hz the Minimax from Catapult Sports (Pty Ltd, Victoria, Australia). The GPS and accelerometer signals were sampled at 5 and 50 Hz, respectively, and averaged per second. The difference between the real distance (track) and the recorded distance (GPS) was less than 1% and 0.92% over 800 m and 1500 m, respectively. This result agrees with previous GPS studies for maximal efforts run by humans or horses [14, 111].

2.3.3 Acceleration trials

In the self-paced acceleration trial protocol, the runners performed three freely paced acceleration sets in which they were asked to maintain constant acceleration by progressively increasing their speed until exhaustion. The trials were run at three constant acceleration values, based on ratings of perceived acceleration ("soft", "medium" and "hard"), in random order. In all of the the self-paced acceleration trials, the runners started at an initial velocity between 2.2 and 3.05 $m.s^{-1}$ and then increased their velocity at three different, constant accelerations ("soft", "medium" and "hard" intensities). There was a one-hour interval between acceleration trials. In the self-paced acceleration trials set, the runners were not provided with any external information other than the distance covered. All tests were performed between 3pm and 6pm on wind-free, spring days ($< 2 m.s^{-1}$ according to an anemometer, the Windwatch from Alba, Silva, Sweden) with a temperature of 20°C, as in a previous study of the energetics of middle-distance running [19].

2.3.4 Mathematical Model

We know that most of the races are stochastic when the velocity is not imposed [22]. Our hypothesis analyzing the measured data was that humans can maintain a constant acceleration with little variations around a mean value. Therefore, we are characterizing the human acceleration on self-paced exercises with three different instruction intensities ("soft", "medium" and "hard") with a Ornstein-Uhlenbeck process. From the model to the simulation, this subsection explains the method used for characterizing the runners self-paced acceleration when so instructed.

2.3.4.1 The Ornstein-Uhlenbeck process To assess the subjects' ability to maintain self-paced, constant acceleration, we check that the equation for constant-acceleration motion is satisfied. If the acceleration of the runner (a) is constant then the runner's speed at time t , denoted $v(t)$, is a linear function of time $v(t) = v(0) + at$, where $v(0) = v_0$ is the initial speed. Here, v_0 is within the range of 2.2 to 3.05 $m.s^{-1}$. As a consequence, the distance $d(t)$ covered at time t has a closed form obtained directly by integration as $d(t) = 0.5at^2 + v_0.t$ if we consider that $d(0) = 0$. Hence, the theoretical relationship between distance and time is a polynomial of second order. As mentioned in the introduction, in order to save energy and optimize performance, running must be as smooth as possible (variation in acceleration must be close to 0 $m.s^{-3}$) during self-paced exercise. In this study, we want to see if the following analogy is possible: self-paced acceleration is as smooth as possible during self-paced exercise. Let us denote $a(t)$ the acceleration at time t . The objective of the paper is to propose a reasonable stochastic model for $a(t)$. Assuming that each runner has an internal variability and an autocorrect ability, we define $a(t)$ as the solution of:

$$da(t) = -\theta[a(t) - a]dt + \sigma dW(t) \tag{IV.1}$$

The above equation is the definition of a mean-reverting process also called Ornstein-Uhlenbeck process. $a(t)$ ($m.s^{-2}$) is the acceleration of the runner at time t (s). a ($m.s^{-2}$) is the mean acceleration. We distinguish two human factors, θ depicting the autocorrect ability of

the runner (that is also the speed reversion parameter) and σ the human variability (that is also the short term standard deviation). The stochastic nature of the model is brought by the Brownian motion $W(t)$.

A solution in terms of integral of a mean-reverting process can be written as the following equation [69]:

$$a(t) = a(0)e^{-\theta t} + a(1 - e^{-\theta t}) + \sigma e^{-\theta t} \int_0^t e^{\theta s} dW(s) \quad (\text{IV.2})$$

The three parameters a , θ and σ are unknown. The objective is to estimate them from the data and to validate a posteriori the fit obtained with model (IV.1). From equation (IV.2), we would like to highlight the fact that unlike the beginning of the current subsection, $a(t)$ is not consistent anymore. In addition $a(t)$ and $v(t)$ are not linear in time. The next subsection focuses on the estimation of the parameters a , θ and σ from the data.

2.3.4.2 Parameter estimation Parameter estimation from discrete observation $a(i\Delta)$ for $i = 0, \dots, n$ can be performed by optimizing a least square criteria. Indeed, we have

$$a(t) = e^{-\theta\Delta t} a(t - \Delta t) + (1 - e^{-\theta\Delta t})a + \sigma \sqrt{\frac{1 - e^{-2\theta\Delta t}}{2\theta}} (W(t) - W(t - \Delta t)) \quad (\text{IV.3})$$

Let us introduce a new parametrization: $\alpha = (1 - e^{-\theta\Delta t})a$, $\beta = e^{-\theta\Delta t}$ and $\gamma = \sigma \sqrt{\frac{1 - e^{-2\theta\Delta t}}{2\theta}}$. The least square criteria is then defined as $(\hat{\alpha}, \hat{\beta}) = \underset{(\alpha, \beta)}{\operatorname{argmin}} \sum_{i=1}^n (a(i\Delta) - \beta a((i-1)\Delta) - \alpha)^2$ and $\hat{\gamma}^2 = \frac{1}{n} \sum_{i=1}^n (a(i\Delta) - \hat{\beta} a((i-1)\Delta) - \hat{\alpha})^2$. The estimators correspond to the maximum likelihood estimators. They are consistent, unbiased and asymptotically normal. The solution of $\hat{\alpha}$ and $\hat{\beta}$ are explicit. The Fisher information matrix can be computed explicitly, that corresponds to the asymptotic variance denoted $\hat{\Sigma}$.

To obtain the estimation of the original parameters, we use the following transformation: $\phi(\alpha, \beta, \gamma) = (\frac{\alpha}{1-\beta}, \frac{-\ln(\beta)}{\Delta}, \gamma \sqrt{\frac{-2\ln(\beta)}{\Delta(1-\beta^2)}})$. The estimator of (a, θ, σ) are then defined as $(\hat{a}, \hat{\theta}, \hat{\sigma}) = \phi(\hat{\alpha}, \hat{\beta}, \hat{\gamma})$. By the delta method, we can prove that their are asymptotically normal: $(\hat{a}, \hat{\theta}, \hat{\sigma}) \xrightarrow{\mathcal{L}} \mathcal{N}((a^*, \theta^*, \sigma^*), \mathcal{J}_\phi \hat{\Sigma} \mathcal{J}'_\phi)$ where $(a^*, \theta^*, \sigma^*)$ are the true (unknown) parameters and \mathcal{J}_ϕ is the Jacobian of ϕ . This allows to compute confidence interval for (a, θ, σ) .

2.4 Results

2.4.1 Comparison of the three different acceleration exercises

In order to know if humans are able to perceive different intensity levels of acceleration, we instructed the five runners to perform three self-paced constant acceleration trials (see subsection 2.3.3). The result of the three trials for the first male runner is shown in Figure IV.1.

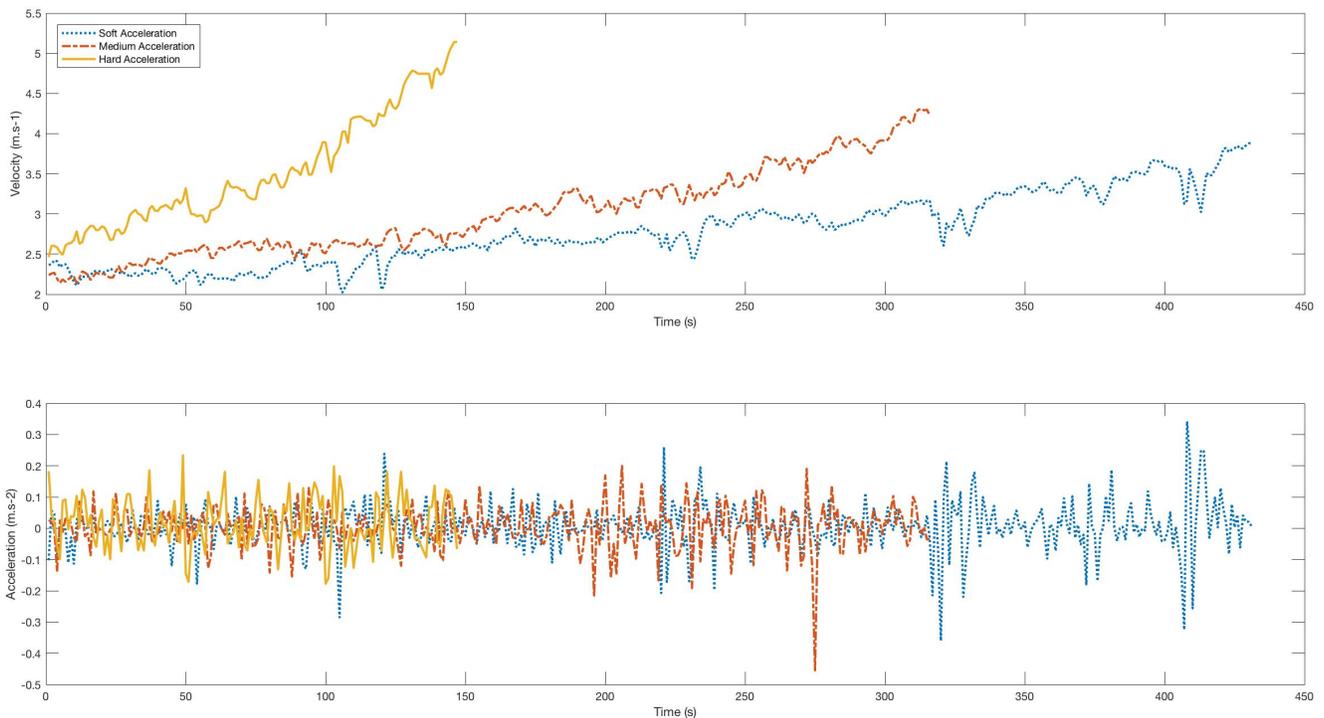


Figure IV.1 – Male Runner 1 - Top chart: Velocity (m/s) VS. Time (s); Bottom chart: Acceleration(m/s²) VS. Time (s) - Comparison of the three acceleration intensity instructions: "soft" (blue curve), "medium" (red curve) and "hard" (yellow curve)

On each plot, the velocity corresponding to "hard" intensity has the highest slope, the "medium" intensity has the medium slope and the "soft" intensity acceleration trial has the smallest one. Results of the four other runners can be found in the appendix A, Figures A.1, A.2, A.3, A.4. Table IV.1 presents the distance and time at exhaustion for each runner and each trial.

| Trial | | Final Distance (m) | Final Time (s) |
|------------------|----------|--------------------|----------------|
| Soft | Male 1 | 1120 | 409 |
| | Male 2 | 4403 | 1204 |
| | Male 3 | 1718 | 602 |
| | Female 1 | 1931 | 603 |
| | Female 2 | 1605 | 570 |
| Average "Soft" | | 2155.4 | 677.6 |
| Medium | Male 1 | 888 | 296 |
| | Male 2 | 2359 | 595 |
| | Male 3 | 923 | 282 |
| | Female 1 | 1139 | 333 |
| | Female 2 | 1029 | 337 |
| Average "Medium" | | 1267.6 | 368.6 |
| Hard | Male 1 | 491 | 143 |
| | Male 2 | 762 | 166 |
| | Male 3 | 436 | 96 |
| | Female 1 | 832 | 225 |
| | Female 2 | 756 | 219 |
| Average "Hard" | | 655.4 | 169.8 |

Table IV.1 – Distance and Time at exhaustion for each runner and for each acceleration trial intensity: "soft", "medium" and "hard"

| | Male 1 | Male 2 | Male 3 | Average Males | Female 1 | Female 2 | Average Females |
|-------------------------------------|--------|--------|--------|---------------|----------|----------|-----------------|
| $\frac{\text{soft}}{\text{medium}}$ | 1.26 | 1.87 | 1.86 | 1.66 | 1.70 | 1.56 | 1.63 |
| $\frac{\text{medium}}{\text{hard}}$ | 1.81 | 3.09 | 2.21 | 2.37 | 1.37 | 1.36 | 1.37 |

Table IV.2 – Ratios of distances to exhaustion for "soft" vs. "medium" and "medium" vs. "hard" acceleration intensities for each runner

The three different intensity levels are distinct for each runner. The times to exhaustion are in the range 96 s (1 minute and 36 seconds) to 1 204 s (approximately 20 minutes). The average time to exhaustion for "soft" intensities is 677.6 seconds (approximately 11 minutes and 6 seconds), for "medium" intensities is 368.6 seconds (6 minutes and 26 seconds) and for "hard" intensities is 169.8 seconds (2 minutes and 49 seconds).

Moreover the average ratios of the "Soft" to "Medium" instructions on "Distance" and "Time" are respectively 1.65 and 2.22 and the average ratios of the "Medium" to "Hard" instructions on "Distance" and "Time" are respectively 1.81 and 2.32. Tables IV.2 and IV.3 display those results. Note that we do not apply any statistical test to determine the significance of the three regimes because of the low sample size (3 males and 2 females). This highlights the

| | Male 1 | Male 2 | Male 3 | Average Males | Female 1 | Female 2 | Average Females |
|-------------------------------------|--------|--------|--------|---------------|----------|----------|-----------------|
| $\frac{\text{soft}}{\text{medium}}$ | 1.38 | 2.02 | 2.13 | 1.84 | 1.81 | 1.69 | 1.75 |
| $\frac{\text{medium}}{\text{hard}}$ | 2.07 | 3.58 | 2.94 | 2.86 | 1.48 | 1.54 | 1.51 |

Table IV.3 – Ratios of times to exhaustion for "soft" vs. "medium" and "medium" vs. "hard" acceleration intensities for each runner

fact that humans are able to feel three different intensities of acceleration and divide distance and time to exhaustion from one intensity to another.

The fact that the three regimes are clearly distinct on the plot strengthens the idea that humans are able to maintain at least three different intensities of acceleration.

2.4.2 Characterizing the human faculty to maintain a constant acceleration

As shown before, we can consider that humans are able to perceive different intensity levels of acceleration. The objective is how to characterize those accelerations. To do so, the Ornstein-Uhlenbeck model (IV.3) is applied.

The goal is to show that it is possible to characterize the human acceleration. We would like to give an estimation and interpretation of the parameters of the Ornstein-Uhlenbeck model $(\hat{a}, \hat{\theta}, \hat{\sigma})$ and their confidence interval. By applying the method described in subsection 2.3.4.2, we have estimated $\hat{a}, \hat{\theta}$ and $\hat{\sigma}$ and the confidence interval for \hat{a} . Results of the estimated parameters and confidence intervals for \hat{a} are shown in Table IV.4.

We observe the following two points from Table IV.4:

1. The mean autocorrect abilities $(\hat{\theta})$ of the runners are steady with a slight drop in the "Hard" trial. This means that the runners are less able to correct their mistakes during the "Hard" trials.
2. The human variability $(\hat{\sigma})$ is steady too with a slight rise in the "Hard" trial. This could be an answer to the drop of autocorrect ability. This means that the runner variability tends toward increasing during "Hard" trial.

The next step consists in checking if our model could explain the reality. By its very nature, an Ornstein-Uhlenbeck process is Gaussian. This property allows us to compute the

| Trial | | \hat{a} | $\hat{\theta}$ | $\hat{\sigma}$ | C.I |
|------------------|----------|-----------|----------------|----------------|------------------|
| Soft | Male 1 | 0.0036 | 1.9803 | 0.1188 | [0.0032; 0.0039] |
| | Male 2 | 0.0026 | 1.3763 | 0.0827 | [0.0025; 0.0027] |
| | Male 3 | 0.0046 | 1.4311 | 0.0868 | [0.0044; 0.0048] |
| | Female 1 | 0.0038 | 1.6404 | 0.0600 | [0.0037; 0.0039] |
| | Female 2 | 0.0033 | 1.4077 | 0.0713 | [0.0031; 0.0035] |
| Average "Soft" | | 0.0036 | 1.5672 | 0.0839 | - |
| Medium | Male 1 | 0.0049 | 1.5569 | 0.1008 | [0.0045; 0.0054] |
| | Male 2 | 0.0040 | 1.2028 | 0.0818 | [0.0037; 0.0042] |
| | Male 3 | 0.0103 | 1.9777 | 0.1372 | [0.0098; 0.0109] |
| | Female 1 | 0.0066 | 1.6293 | 0.0565 | [0.0063; 0.0068] |
| | Female 2 | 0.0047 | 1.7663 | 0.0669 | [0.0045; 0.0050] |
| Average "Medium" | | 0.0061 | 1.6266 | 0.0886 | - |
| Hard | Male 1 | 0.0172 | 1.0275 | 0.0982 | [0.0159; 0.0185] |
| | Male 2 | 0.0160 | 1.4493 | 0.1054 | [0.0151; 0.0169] |
| | Male 3 | 0.0253 | 0.7115 | 0.1400 | [0.0213; 0.0293] |
| | Female 1 | 0.0088 | 1.2892 | 0.0450 | [0.0085; 0.0091] |
| | Female 2 | 0.0065 | 1.3800 | 0.0907 | [0.0059; 0.0072] |
| Average "Hard" | | 0.0147 | 1.1715 | 0.0958 | - |

Table IV.4 – Mean acceleration (\hat{a}), autocorrect ability ($\hat{\theta}$) and human variability ($\hat{\sigma}$) estimated for each runner and for each acceleration trial intensity: "soft", "medium" and "hard" and confidence interval (C.I) for \hat{a}

| | Male 1 | Male 2 | Male 3 | Female 1 | Female 2 | Average |
|-----------------------|--------|--------|--------|----------|----------|---------|
| $\frac{soft}{medium}$ | 0.73 | 0.65 | 0.45 | 0.56 | 0.70 | 0.62 |
| $\frac{medium}{hard}$ | 0.28 | 0.25 | 0.41 | 0.75 | 0.72 | 0.48 |

Table IV.5 – Ratios of estimated acceleration for "soft" vs. "medium" and "medium" vs. "hard" acceleration intensities for each runner

quantiles exactly. Once such confidence boundaries are computed, we check the percentage of the measured acceleration data present inside.

Ratios of estimated acceleration for "soft" vs. "medium" and "medium" vs. "hard" acceleration intensities in Table IV.5 seem to confirm the hypothesis that humans are able to distinct at least 3 intensities of acceleration.

In order to evaluate the validity of the mean-reverting process applied to acceleration, the goodness-of-fit of the model has been tested in three steps:

1. Checking the ability of the Ornstein-Uhlenbeck model to distinguish three distinct mean accelerations.
2. Checking for non autocorrelation of residuals with the Durbin-Watson test.
3. Testing for stationarity with the Augmented Dickey-Fuller (ADF) Test.

2.4.2.1 Checking the ability of the Ornstein-Uhlenbeck model to distinguish

three distinct mean accelerations Let us now highlight the ability of our model to distinguish the three intensities of acceleration. In order to do so, we use the delta method to compute a confidence interval for each estimated acceleration \hat{a} . The Figure IV.2 displays the three different estimated mean accelerations for the male runner 1, symbolized by three dots. The 95 % confidence intervals for each estimated mean acceleration have been plotted as well symbolizing by error bars. It appears that the suggested Ornstein-Uhlenbeck model seems to distinguish three distinct mean accelerations: Figure IV.2 and Figures C.1, C.2, C.3, C.4 for the four other runners show that the confidence intervals do not overlap. However, because of the low sample size, this conclusion is purely descriptive. A statistical test will be applied on the future on a larger dataset.

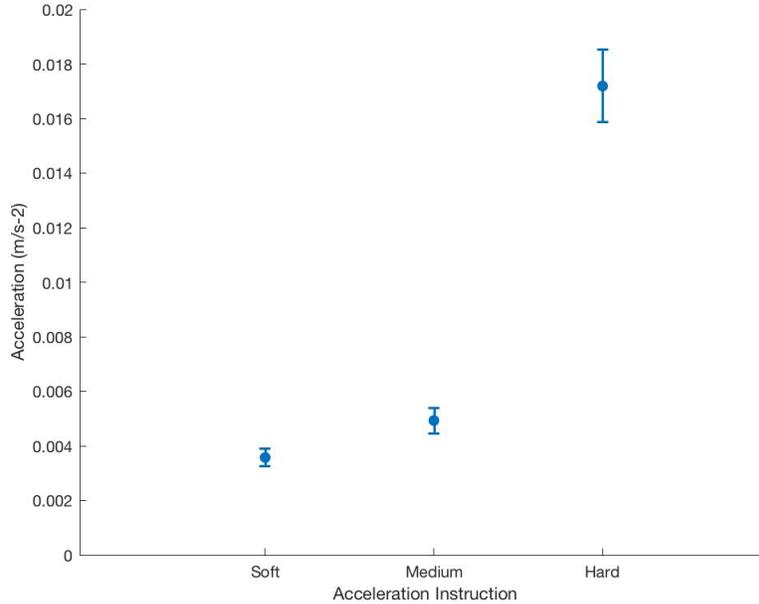


Figure IV.2 – Male Runner 1 - Comparison of the three estimated mean accelerations

The Figure IV.3 shows the result of one simulation for the first male runner. We can distinguish 3 information per chart with two curves and one filled region on each chart. The black curve (dotted) represents the measured data (i.e. measured accelerations), the red one (solid) represents one random simulation of the Ornstein-Uhlenbeck model (equation (IV.3) subsection 2.3.4) and the filled region stands for the 95% confidence interval.

We can quickly observe that the confidence interval is consistent with the reality of our data and that most of the recorded acceleration data seem to be inside the confidence interval. Results for the four other runners can be found in appendix B, Figures B.1 to B.4.

To confirm that a significant amount of measured acceleration are inside the confidence boundaries, Table IV.6 shows that the values of percentages of measured data found in the 95% confidence boundaries are above 92%. Thus the Ornstein-Uhlenbeck model described in subsection 2.3.4.2 seems to be consistent and could be used to describe the reality.

To go further, it would be interesting to perform self-paced acceleration following the same protocol with first, more runners in order to check if there is a trend in the ratios, second, splitting the runners into two categories (men and women) to check if there is a gender effect on those ratios.

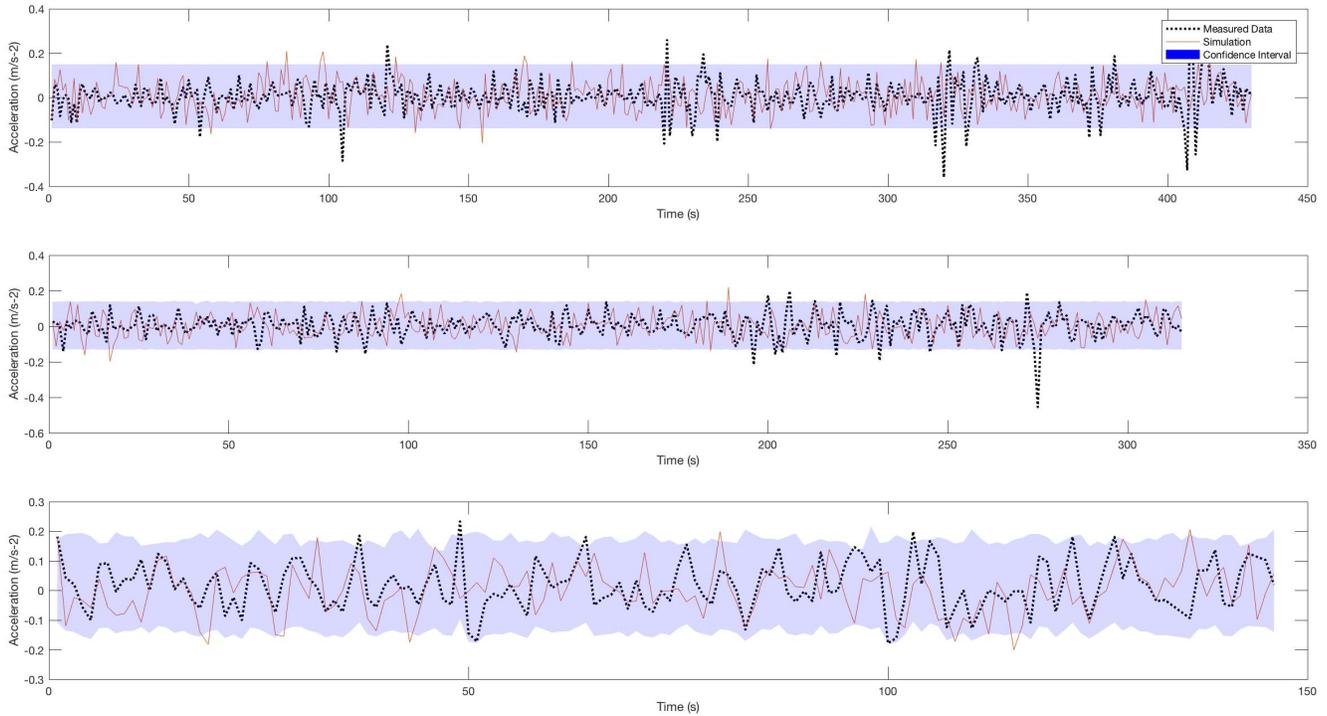


Figure IV.3 – Male Runner 1 - From top to bottom: Soft, Medium and Hard Acceleration Intensities - Measured data VS. Ornstein-Uhlenbeck model

| | | Male 1(%) | Male 2(%) | Male 3(%) | Female 1(%) | Female 2(%) |
|--------------|--------|-----------|-----------|-----------|-------------|-------------|
| Acceleration | Soft | 94.42 | 95.43 | 94.02 | 97.51 | 97.89 |
| | Medium | 95.24 | 97.14 | 94.33 | 94.29 | 95.63 |
| | Hard | 99.32 | 97.60 | 96.00 | 92.89 | 95.45 |

Table IV.6 – Results of percentages of measured data found in the 95% confidence interval for acceleration for each trial and each runner.

| | Male 1 | Male 2 | Male 3 | Female 1 | Female 2 |
|--------|--------|--------|--------|----------|----------|
| Soft | 0.94 | 0.33 | 0.89 | 0.95 | 0.21 |
| Medium | 0.86 | 0.33 | 0.84 | 0.82 | 0.96 |
| Hard | 0.52 | 0.63 | 0.008 | 0.61 | 0.29 |

Table IV.7 – p – values of Durbin-Watson statistic to check for autocorrelation among residuals for each runner and for each acceleration trial intensity ("soft", "medium" and "hard")

2.4.2.2 Testing for non autocorrelation of residuals In this subsection, we want to detect the presence of autocorrelation among residuals. In our case, the residuals is a vector defined by the difference between our measured data $a(t)$ and the estimated acceleration parameter from the Ornstein-Uhlenbeck model (\hat{a}). We perform the Durbin-Watson statistic on the following linear regression:

$$a(t) - a(t - \Delta_t) = \hat{\theta}(a(t) - \hat{a})\Delta_t + \eta(t), \quad (\text{IV.4})$$

where $\eta(t)$ is the vector of residuals. If the model is correct, the residuals are independent and their distribution follows a normal distribution described as follows: $\mathcal{N}(0, \sigma^2 \Delta t)$

In order to perform the Durbin-Watson statistic, we define the null hypothesis (H_0) as follows: the residuals of the linear regression (IV.4) are uncorrelated. The alternative hypothesis (H_1) is that there is autocorrelation among the residuals. In order to test H_0 against H_1 , we used the function `dwtest`¹ from Matlab ®. The results of the test are shown in Table IV.7.

The range of p – value is included between 0.08 and 0.98. Given those results, we cannot reject the null hypothesis stating that the residuals are uncorrelated.

2.4.2.3 Testing for stationarity In this subsection, we test for no unit root and whether the process of acceleration is stationary or not to insure that the estimated mean accelerations neither grow nor decrease over time. We define the null Hypothesis (H_0) as follows: the unit root is present in the measured acceleration time series. The alternative hypothesis (H_1) is defined as follows: the accelerations time series is trend-stationary. In order to test H_0

¹See the documentation on `dwtest`: <https://fr.mathworks.com/help/stats/dwtest.html>

| | Male 1 | Male 2 | Male 3 | Female 1 | Female 2 |
|--------|-----------|-----------|-----------|-----------|-----------|
| Soft | < 1.0e-03 |
| Medium | < 1.0e-03 |
| Hard | < 1.0e-03 |

Table IV.8 – p – value of Augmented Dickey-Fuller static to check for stationarity for each runner and for each acceleration trial intensity ("soft", "medium" and "hard")

against H_1 , we used the function *adftest*² from Matlab®.

The results of the test are shown in Table IV.8. Given those results, we reject the null hypothesis stating that the simulation of acceleration is non-stationary.

2.5 Discussion

Our results showed that recreational runners are able to control their acceleration until exhaustion at three significantly different accelerations values perceived to be "soft", "medium" and "hard". Indeed, our data showed that runners are able to apply distinct, subjective acceleration values when so instructed and maintain constant acceleration until exhaustion, regardless of the acceleration value. In addition, our results showed that self-paced acceleration fitted the following Ornstein-Uhlenbeck process $da_t = -\theta(a_t - a)dt + \sigma dW_t$ where a is the mean acceleration, θ is the human ability for autocorrection, σ the human variation and a_t the measured acceleration. Indeed, a fine-grain analysis of the data shows that accelerations are obtained by the application of brief and frequent corrections designed to reach a mean target value a , and that there is no drift in the accelerations. Moreover, the fact that there is no unit-root in the acceleration time series shows that possible strong deviance from the mean value a are rapidly corrected and vanishes by mean-reverting corrections. Further studies are now required for characterizing the deceleration of a recreational runner.

Given the fact that this is the first study having examine the human ability to maintain self-paced acceleration, we are going to discuss the two following points:

1. The self-paced acceleration range values compared with current use of acceleration protocols in laboratory.

²See the documentation on *adftest*: <https://fr.mathworks.com/help/econ/adftest.html>

2. The possible mechanisms allowing humans to control its body acceleration in accordance with prior results obtained on manual task.

2.5.1 The self-paced acceleration range values compared with current use of acceleration protocols in laboratory

The range of these perceived acceleration values is in accordance with observed values in middle-distance running [21]. Furthermore, these freely chosen accelerations also correspond to the imposed acceleration values frequently used in treadmill ramp protocols for determining VO₂max [132, 141]. In ramp protocols, the work rate is ramped up as a continuous increase and then a continuous acceleration. Given that the linearity of the oxygen uptake response is a major discriminating cardiovascular feature for assessing exercise intolerance, it is important to be sure that the work rate profile is linear and then the acceleration is constant. That is why the tests are currently performed on a treadmill. Even though manufacturers have developed a range of technologies for enabling ramp tests (e.g. controlled cycle ergometer), some subjects have difficulty walking and running on a treadmill and then reaching their maximum VO₂ and speed at VO₂max. It has been argued that the treadmill use induces a higher maximum metabolic rate and requires a mode of exercises that more closely approximates some activities of daily living [13, 64]. In a much easier protocol, the present study shows that male and female middle-aged, recreational runners were able to self-monitor acceleration and thus reproduced an outdoor ramp protocol. Indeed, the present results showed that it is possible to apply a self-paced, ramp-like running protocol on a real running track. Regardless of the acceleration intensity level ("soft", "medium" or "hard"), the protocol uses a continuous change in speed and brings the subjects to exhaustion in approximately 2 minutes and 30 seconds to 20 minutes; this meets the criteria for clinical exercise testing issues by the relevant international organizations (American College of Sports Medicine 2000, American Thoracic Society 2003, [133, 67, 174]).

2.5.2 The possible mechanisms allowing humans to control its body acceleration in accordance with prior results obtained on manual task

Here we showed that human are accurately able to accelerate smoothly and constantly, with acceleration vs. time values close to 0 m.s^{-2} (even near to exhaustion). However, the question of how acceleration is controlled remains to be determined. There is still no studies having worked on the ability of self-pace its acceleration in locomotion. Therefore we can only discuss versus prior studies on arm and hand movement in catching tasks. In that tasks, neurophysiologists have demonstrated that the human brain imposes (in a top-down fashion) its rules of interpretation on sensory data. It transforms the perceived world according to rules of symmetry and stability and kinematic laws derived from principles of maximum smoothness [16]. They hypothesized that the first possibility was that humans can perceive distance very accurately [131]. It has been proved that humans possess efficient mechanisms for building an abstract representation of distances, which can be transferred from one perceptual context to another [131].

This paper seemed to show that whatever the variation in velocity, from "soft" acceleration intensity level at 0.0026 m.s^{-2} to ten times that at 0.0253 m.s^{-2} in the "hard" acceleration intensity level in self-paced acceleration trial, the acceleration was well controlled and steady, with a human ability to autocorrect and a minimal jerk value. The concept of "jerk-cost" has been validated for running and walking stride smoothness [91].

2.6 Conclusion

To the best of our knowledge, a human's ability to maintain constant acceleration in a conscious way until exhaustion has not yet been investigated. We therefore decided to test the hypothesis whereby humans are able to maintain a constant acceleration, regardless of the velocity and the magnitude of acceleration. Indeed, despite the large, published body of work on pacing strategy and speed control (especially concerning feedback vs. feed forward power output control and central vs. peripheral mechanisms), this is the first study to have

examined acceleration control during running.

The main findings of this research are as follows:

1. The measured data seemed to show that runners are able to apply at least three distinct and subjective acceleration intensities when so instructed.
2. The measured data seemed to show that the runners tend towards reducing the time and distance to exhaustion between the "soft" and the "medium" instruction as well as between the "medium" and "soft" instructions.
3. The simulation using the Ornstein-Uhlenbeck process seemed to show that accelerations are obtained by the application of brief and frequent corrections designed to reach a mean target value a .
4. The simulation showed that the runners tend towards increasing the mean acceleration between the "soft" and the "medium" instruction as well as between the "medium" and "soft" instructions.

Our results show that runners are able to control their acceleration until exhaustion at three significantly different accelerations values perceived to be "soft", "medium" and "hard". According to Newton's Second Law, humans could control their force of motion as it is directly proportional to the acceleration opening the door on future study related to human physiology and human energetic. We can think of two research opportunities go further. One would be to study the deceleration and see if humans are also able to control it. Once we successfully describe both acceleration and deceleration with mathematical models, we could think about alternating acceleration and deceleration phases to see how runners would react and study the sequence. The other would be gathering data from many runners, both males and females, in order to establish a predictive model for human acceleration. This could be done using Stochastic Differential Mixed-Effects Equations.

Bilan

Nous avons pu, dans ce chapitre, montrer que l'être humain est capable de maintenir trois intensités d'accélération différentes dont la valeur est à chaque fois constante. Ces trois intensités perçues, "facile", "moyenne" et "dure", ont pu être décrites grâce à l'application d'un modèle d'Ornstein-Uhlenbeck. Cependant, ce modèle, descriptif, ne nous permet pas de prédire l'accélération humaine. C'est dans cette direction que nous souhaitons diriger nos recherches à présent. En effet, les transitions entre les vitesses (ou états physiologiques) du coureur sont un sujet clé afin de construire une stratégie de course.

Ce chapitre clos la partie recherche de cette thèse. Nous allons à présent passer aux étapes de développements informatiques qui ont permis l'automatisation de la méthode BillaTraining.

Chapitre V

Quatrième contribution : Vers un produit fini

Cet ultime chapitre est consacré à la digitalisation de *BillaTraining*[®]. Afin de gagner des parts de marché, diffuser son savoir au plus grand nombre et innover, l'entreprise cherche à automatiser la gestion de ses clients et construire des outils évolutifs. Deux produits ont été créés pendant cette thèse. Le premier est une application web permettant l'accompagnement des coureurs à pied dans leurs différents objectifs. L'application web propose des programmes d'entraînement personnalisés. Le second produit est une application mobile permettant d'interagir en temps réel avec les coureurs à pied le désirant.

Sommaire

| | | |
|----------|---|------------|
| 1 | Problématiques et enjeux | 138 |
| 2 | Création d'une application web d'entraînement en course à pied | 139 |
| 2.1 | Architecture de l'application web | 140 |
| 2.2 | Les fonctionnalités de l'application web | 141 |
| 2.2.1 | La logique métier | 142 |
| 2.2.2 | Le web UI/UX | 146 |
| 2.2.3 | Les analyses scientifiques | 156 |
| 2.2.4 | Le paiement en ligne | 158 |

| | | |
|--------------|--|------------|
| 3 | Prototype d'une application mobile temps réel | 160 |
| 3.1 | Architecture de l'application mobile | 161 |
| 3.2 | Les fonctionnalités de l'application mobile | 163 |
| 3.2.1 | Le temps réel | 163 |
| 3.2.2 | Le mobile UI/UX | 165 |
| 3.2.3 | Le web UI/UX 2 | 174 |
| Bilan | | 175 |

1 Problématiques et enjeux

En regardant le marché de l'entraînement sportif pour le grand public, on distingue deux types d'acteurs. Le premier type concerne les entraîneurs ou groupe d'entraîneurs délivrant un service individualisé à leurs clients en faisant évoluer le programme d'entraînement au fur et à mesure. Le second type propose des programmes d'entraînement génériques et non personnalisés basés sur la distance à parcourir le jour de la course (marathon, semi-marathon, 10km, etc...). On note aisément l'opposition entre la qualité du suivi et la quantité de client potentiellement absorbable.

Suite à ce constat, un troisième type d'acteur commence à se répandre sur le marché de l'entraînement en course à pied. Il s'agit d'applications web et / ou mobile permettant de générer automatiquement un programme d'entraînement sur-mesure. L'entreprise *BillaTraining*[®] se positionne comme étant une solution scientifique répondant à la problématique de l'entraînement personnalisé pour le grand public.

L'objectif de l'entreprise *BillaTraining*[®] est double. Premièrement, il s'agit de répandre les méthodes d'entraînement scientifiques sur le marché de la course à pied. L'intérêt est d'avoir un passage direct entre la recherche et l'utilisateur final. Deuxièmement, il est question de trouver la "vache à lait"¹ de *BillaTraining*[®] afin de pérenniser la société.

¹https://fr.wikipedia.org/wiki/Matrice_BCG

Nous avons décidé de mettre en place deux nouveaux produits comme fer de lance de la stratégie digitale. Le premier produit est une application web 100% automatisée, de la gestion des utilisateurs aux programmes d'entraînement en passant par le paiement en ligne, permettant aux coureurs d'avoir un entraînement sur-mesure et évolutif (voir [sous-section 2.1](#)). Le second produit est un prototype d'application mobile permettant de s'affranchir des montres connectées et d'améliorer l'expérience client grâce à un enregistrement en temps réel de la vitesse et de la fréquence cardiaque (voir [sous-section 3.1](#)).

2 Création d'une application web d'entraînement en course à pied

Nous aborderons dans cette section les étapes ayant permis la mise en place de l'application web *RabiT*[®] by *BillaTraining*[®]. Par application web, nous entendons une architecture client-serveur ne nécessitant aucune installation sur la machine de l'utilisateur et accessible via un navigateur web (e.g. Mozilla Firefox, Google Chrome, Internet Explorer, Safari).

Précédemment, dans la [sous-section 4.1](#) du [chapitre II](#), nous avons introduit les architectures monolithique et en micro service. Après réflexion, nous avons décidé de développer une application hybride (i.e. avec des briques monolithiques et des micro services). Nous souhaitons tester l'intérêt des utilisateurs finaux pour une application d'entraînement sur-mesure en course à pied dans un premier temps.

Nous définissons ici deux termes que nous allons utiliser dans toute cette section :

- *Front-end* : Définit tout ce qui se rapporte aux interactions entre l'utilisateur final et la machine. C'est la partie client de l'architecture client-serveur. Nous utilisons les langages HTML, CSS et JavaScript afin de construire la meilleure expérience utilisateur possible [55, 56].
- *Back-end* : Définit tout ce qui n'est pas visible par l'utilisateur final. C'est la partie serveur de l'architecture client-serveur. Nous utilisons des langages comme *Php*, *Ruby* ou *Python* afin de construire la logique de l'application [178, 171].

2.1 Architecture de l'application web

Afin de comprendre les étapes de création de l'application web d'entraînement en course à pied, la figure V.1 donne une vue globale des briques à développer ou à connecter.

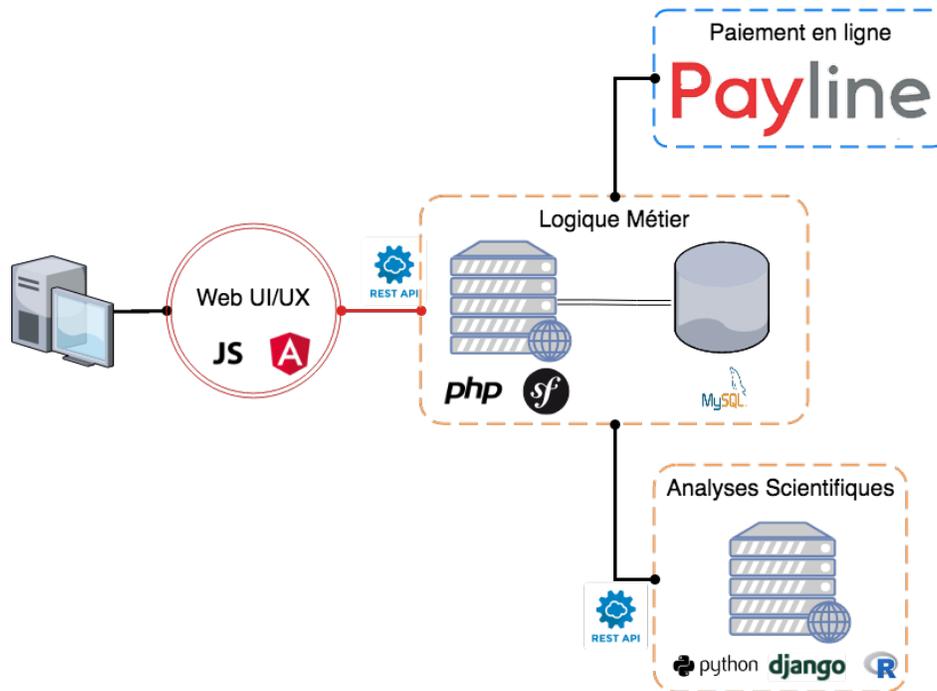


FIGURE V.1 – Schéma global de l'application web *RabiT*[®] by *BillaTraining*[®].

On distingue 4 briques distinctes que nous pouvons diviser en deux catégories. La première catégorie rassemble les briques à développer par la société *BillaTraining*[®], à savoir :

- Le Web UI/UX : "UI" signifie *User Interface* en anglais et peut être traduit par l'interface utilisateur. En d'autres termes, il est question de tous les éléments graphiques que verra le client en navigant sur l'application. "UX" signifie *User Experience* en anglais et peut être traduit par l'expérience utilisateur. Un "UX designer" essaiera de rendre la navigation du client la plus simple et la plus intuitive possible. Nous avons fait le choix d'utiliser *Angular*², un *framework* développé en JavaScript, afin de développer l' "UI/UX" qui attirera le plus de client. Tout le *front-end* de notre application sera contenu dans cette brique.

²<https://angular.io/>

- La logique métier : Cette brique regroupe la logique de l'application, l'authentification et la gestion des accès. Invisible par le client, elle est le cerveau de notre application. Cette brique et la partie centrale de notre *back-end* car elle permet l'interconnexion avec d'autres services et est la seule pouvant communiquer avec le *front-end*. Nous avons fait le choix d'utiliser *Symfony*³, un *framework* développé en *Php*.
- Les analyses scientifiques : Initialement conçu pour la segmentation du test de *RabiT*[®], ce micro service devra pouvoir accueillir le calcul de la signature de vitesse dans le futur. Il s'agit d'une autre partie de notre *back-end* développé avec le *framework Django*. Nous avons vu les étapes du développement de cette brique dans la [sous-section 4.3](#) du [chapitre II](#).

La seconde catégorie regroupe les briques développées par des tiers auxquelles nous devons venir nous connecter afin d'utiliser leurs services. Il s'agit ici de la brique de paiement en ligne. L'intérêt d'utiliser des services tiers est de pouvoir se concentrer sur son coeur de métier et d'accélérer ainsi le développement. Nous avons opté pour *Payline*⁴ comme solution de paiement en ligne pour notre application.

2.2 Les fonctionnalités de l'application web

Passons à présent aux fonctionnalités de notre application. Nous avons établi un cahier des charges dans le but de cadrer le développement de notre application web. C'est une phase indispensable au bon déroulement d'un projet web qui permet de définir la feuille de route et de fixer les choix technologiques.

Le projet a pour but de créer une application web ayant pour but la gestion des entraînements en course à pied et l'accompagnement jusqu'au jour de la course des clients de l'entreprise *BillaTraining*[®]. L'application devra intégrer l'intégralité de la méthode *BillaTraining*[®], soit, les quatre étapes : *RabiT*[®], Radar Énergétique, Entraînement sur-mesure et signature de vitesse. Par conséquent, la plateforme permettra aux coureurs d'enregistrer leurs données

³<https://symfony.com/>

⁴<https://www.payline.com/>

physiologiques et d'entraînement (e.g. vitesse, fréquence cardiaque, temps, distance, poids, taille, ...). Ces données seront analysées automatiquement et les coureurs pourront suivre leur progression. L'entreprise *BillaTraining*[®] souhaite que le parcours client, de l'inscription en passant par la facturation et jusqu'au suivi des progrès ne nécessite aucune intervention humaine.

2.2.1 La logique métier

Suite à la présentation du schéma global de l'application *RabiT*[®] by *BillaTraining*[®] (voir figure V.1), il est temps de rentrer dans le détail de la brique *logique métier*. La figure V.2 présente les dix fonctionnalités à développer. Nous les détaillerons une à une dans cette section.

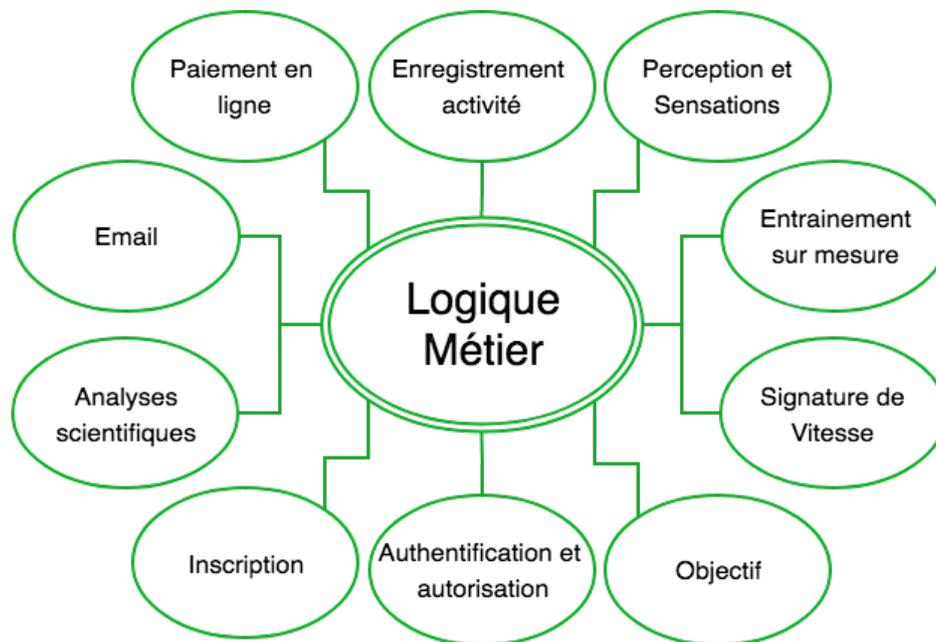


FIGURE V.2 – Schéma des fonctionnalités de la brique *logique métier* de l'application *RabiT*[®] by *BillaTraining*[®].

Inscription La fonctionnalité *inscription* permet, comme son nom le laisse entendre, l'inscription des clients à la plateforme *RabiT*[®] by *BillaTraining*[®]. Lors de la première connexion, l'utilisateur devra rentrer différentes informations ayant pour objectif de l'identifier tout au long de son expérience et d'apprendre à le connaître pour lui fournir un entraînement per-

sonnalis . Les informations demand es sont pr esent es dans le tableau V.1.

| Param tre | Description | Type | Exemple |
|-----------|---|---------|------------------------------------|
| gender | La personne est un homme ou une femme | bool en | 1 |
| name | Le pr enom de la personne | texte | Thomas |
| surname | Le nom de famille | texte | Carbillet |
| birthdate | La date de naissance au format YYYY-MM-DD | date | 1990-06-14 |
| height | La taille en cm | entier | 178 |
| weight | Le poids en kg | entier | 65 |
| email | L'adresse email | texte | thomas.carbillet@billatraining.com |
| password | Le mot de passe pour les futures connexions | texte | lorem |

TABLE V.1 – Liste exhaustive des param tres, de leur description, de leur type et exemple de la fonctionnalit  *inscription* de la brique *logique m tier* de l'application *RabiT*[®] by *BillaTraining*[®].

Authentification et autorisation Pour mieux comprendre la fonctionnalit  *authentification et autorisation*, il faut s parer authentification d'autorisation. L'authentification est la phase durant laquelle le coureur rentre son couple email / mot de passe permettant   l'application de v rifier son existence et de retrouver son historique. L'autorisation, quand   elle, attribue un r le au coureur et indique   l'application les pages et actions avec lesquelles il peut interagir. Pour l'instant, seul un r le a  t  d fini dans l'application, le r le de "coureur". Par la suite, nous souhaitons ajouter deux autres r les. Le premier, celui de "coach" qui permettrait   un entra neur de suivre tous les coureurs qu'il encadre sur la m me plateforme. Le second, celui de "scientist" qui donnerait un acc s et des outils analysant la base de donn es dans le but d' crire de nouveaux articles scientifiques et d'am liorer la plateforme.

Objectif La fonctionnalit  *objectif* doit permettre   l'utilisateur d'ajouter une ou plusieurs  ch ances auxquelles il souhaite participer. Cette fonctionnalit  est essentielle en vue de la cr ation d'un entra nement sur mesure. L'entra nement sera diff rent si le coureur pr pare un  v nement pr cis ou s'il souhaite s'entra ner dans le but de se maintenir en forme. Le tableau V.2 pr sente les param tres n cessaires   l'enregistrement d'un objectif.

Enregistrement activit  La fonctionnalit  *enregistrement activit * a pour but de transformer les donn es brutes issues des dispositifs connect s en un format l ger et pratique (i.e.

| Paramètre | Description | Type | Exemple |
|-----------|---|-------|-------------------|
| name | Nom de l'objectif | texte | Marathon de Paris |
| date | Date de l'objectif au format YYYY-MM-DD | date | 2019-04-14 |
| distance | Nombre de kilomètres à parcourir | réel | 42,195 |

TABLE V.2 – Liste exhaustive des paramètres, de leur description, de leur type et exemple de la fonctionnalité *objectif* de la brique *logique métier* de l'application *RabiT*[®] by *BillaTraining*[®].

le format JSON⁵). Par ailleurs, cette fonctionnalité permet aussi de classer l'activité selon son type (e.g. test de *RabiT*[®], séance de force musculaire, libre). Le tableau V.3 présente les paramètres utiles à l'utilisation de la fonctionnalité *enregistrement activité*.

| Paramètre | Description | Type | Exemple |
|----------------------|--|-------|--|
| <i>activity_type</i> | Type d'activité | texte | Perception |
| data | Données issues des dispositifs connectés (e.g. vitesse, fréquence cardiaque, accélération) | json | [<i>speed</i> : [3.4, 5.6], <i>fc</i> : [187, 189], <i>time</i> : [10, 20]] |
| date | Date de l'activité au format YYYY-MM-DD | date | 2018-12-01 |

TABLE V.3 – Liste exhaustive des paramètres, de leur description, de leur type et exemple de la fonctionnalité *enregistrement activité* de la brique *logique métier* de l'application *RabiT*[®] by *BillaTraining*[®].

Perception et sensations Lors de la création d'une application entièrement automatisée, il peut être difficile de faire du sur-mesure. Afin de palier à cette situation et afin de mieux comprendre les coureurs, nous avons décidé de leur demandé de nous faire part de leurs ressentis. La fonctionnalité *perception et sensations* permet de récolter ces informations sur un coureur. Le but final étant d'utiliser des outils d'analyse de langage afin de décrypter la personnalité d'un coureur et améliorer son entraînement sur-mesure. Le tableau V.4 présente les paramètres nécessaires à la fonctionnalité *perception et sensations*.

⁵https://www.w3schools.com/whatis/whatis_json.asp

| Paramètre | Description | Type | Exemple |
|------------|---|--------|---|
| feelings | Difficulté de l'exercice notée entre 0 et 5 | entier | 3 |
| perception | Ressenti du coureur | texte | Reprise après une semaine off. Décrassage avant marathon. |

TABLE V.4 – Liste exhaustive des paramètres, de leur description, de leur type et exemple de la fonctionnalité *perception et sensations* de la brique *logique métier* de l'application *RabiT*[®] by *BillaTraining*[®].

Analyses scientifiques La fonctionnalité *analyses scientifiques* permet d'interfacer les briques *logique métier* et *analyses scientifiques* par le biais d'une API⁶. Toutes les requêtes effectuées à la brique *analyses scientifiques* passent par cette fonctionnalité. Nous récupérons, grâce à cette fonctionnalité, le résultat de la segmentation d'un test de *RabiT*[®] et les notes du radar énergétique associées. La première et la deuxième étapes de la méthode *BillaTraining*[®] sont gérées par cette fonctionnalité.

Entraînement sur-mesure La fonctionnalité *entraînement sur-mesure* permet de générer un programme d'entraînement personnalisé à un coureur sous conditions qu'il ait effectué un test de *RabiT*[®] au préalable. L'entraînement, en plus d'être personnalisé, devra pouvoir se remettre à jour et évoluer en fonction de l'historique fournit par le coureur. Trois séances d'entraînement sont données par semaine, au fur et à mesure de la progression. L'algorithme de création d'entraînement est confidentiel et ne sera pas développé ici. La fonctionnalité *entraînement sur-mesure* gère la troisième étape de la méthode *BillaTraining*[®]. Il s'agit également de la seule fonctionnalité payante de l'application *RabiT*[®] by *BillaTraining*[®].

Signature de vitesse La fonctionnalité *signature de vitesse* est censée générer une stratégie de course personnalisée. Nous avons décidé de ne pas implémenter cette fonctionnalité dans la première version de notre application web étant donné que nous n'avons pas de modèle de stratégie de course assez précis. L'idée est de donner, à terme, un temps de passage par kilomètre par exemple. C'est cette fonctionnalité qui gèrera la quatrième étape de la méthode

⁶API est un acronyme venant de l'anglais *Application Programming Interface* et désigne un ensemble de tâches, de protocoles de communication et d'outils permettant le développement de la logique d'une application.

BillaTraining[®].

Paiement en ligne Sur le même principe que la fonctionnalité *analyses scientifiques*, la fonctionnalité *paiement en ligne* permet de réaliser l'interface entre la brique *logique métier* et *paiement en ligne* via une API lorsque le coureur souhaite souscrire à un entraînement sur-mesure d'une durée de trois mois.

Email Nous souhaitons centraliser l'envoi d'email à nos clients. Ceci nous permet de gérer l'envoi de newsletter, d'email de confirmation d'inscription, de réinitialisation de mot de passe et d'email au support technique. Nous avons opté pour la solution offerte par mailgun⁷.

2.2.2 Le web UI/UX

Continuons avec la brique *web UI/UX*, qui, rappelons-le, doit gérer toute l'expérience et l'interface client [40, 47, 66, 168]. Nous avons identifié 12 fonctionnalités qui correspondront au 12 vues que le coureur verra en navigant sur l'application *RabiT*[®] by *BillaTraining*[®]. La figure V.3 recense les 12 vues, nous allons les détailler une par une dans la suite de ce paragraphe.

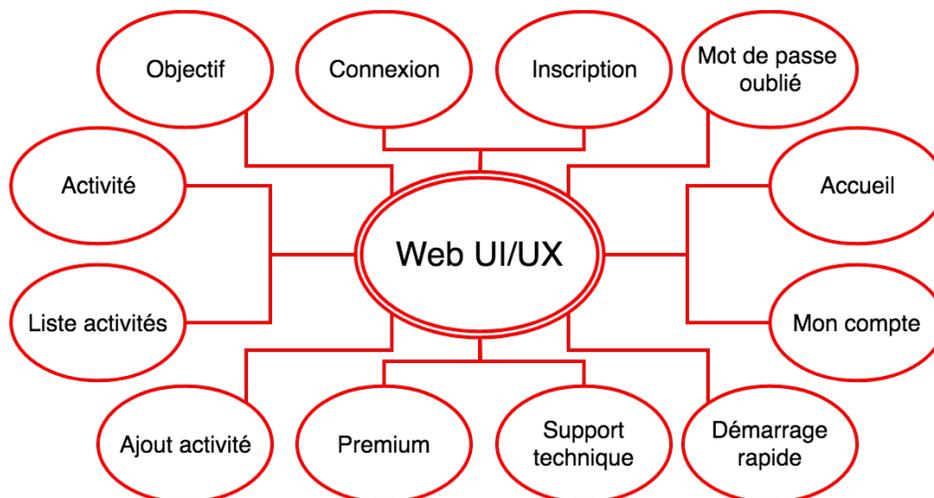


FIGURE V.3 – Schéma des fonctionnalités de la brique *web UI/UX* de l'application *RabiT*[®] by *BillaTraining*[®].

⁷<https://www.mailgun.com>

Inscription Cette étape est incontournable lors de la première connexion à la plateforme. Il est demandé à l'utilisateur de renseigner ses informations personnelles afin d'essayer la méthode *BillaTraining*[®]. Une fois inscrit, le coureur reçoit un email de confirmation et est redirigé vers la page d'accueil. La figure V.4 correspond à la vue *inscription*.

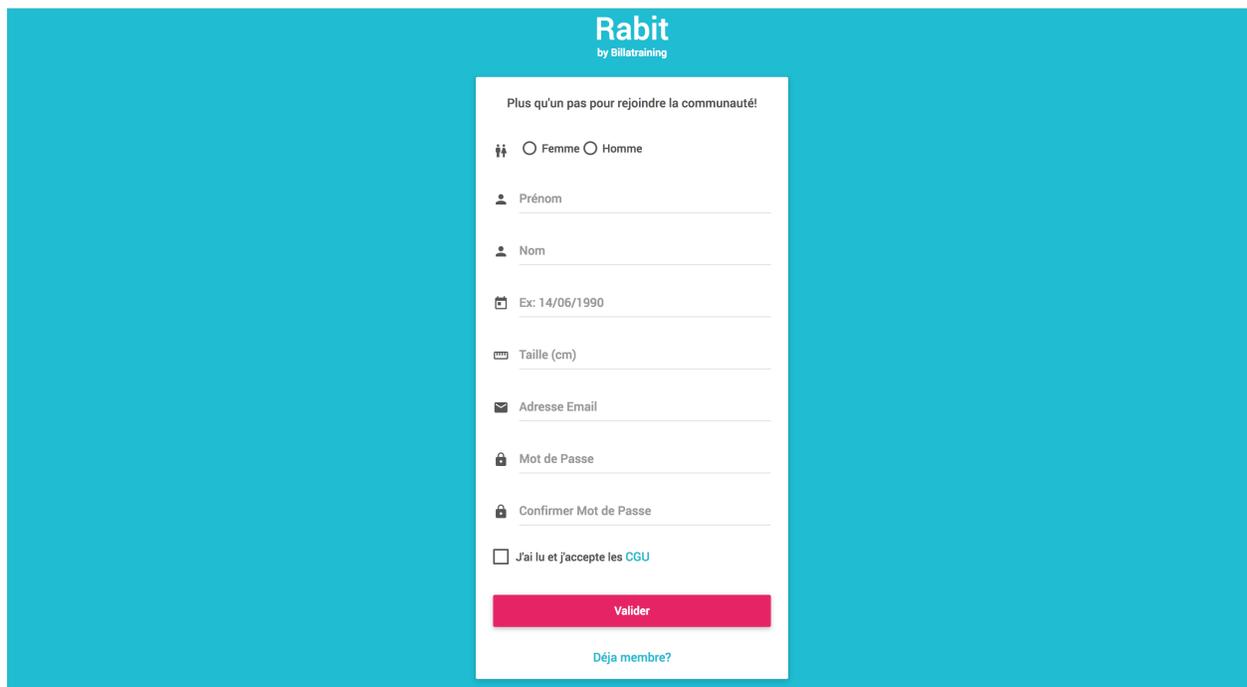
The image shows a registration form for 'Rabit by BillaTraining' on a teal background. At the top, the logo 'Rabit by BillaTraining' is displayed. Below it, a white box contains the text 'Plus qu'un pas pour rejoindre la communauté!'. The form includes a gender selection with radio buttons for 'Femme' and 'Homme'. There are input fields for 'Prénom', 'Nom', 'Date de naissance' (with an example 'Ex: 14/06/1990'), 'Taille (cm)', 'Adresse Email', 'Mot de Passe', and 'Confirmer Mot de Passe'. A checkbox for 'J'ai lu et j'accepte les CGU' is present. At the bottom of the form is a red 'Valider' button and a blue link 'Déjà membre?'.

FIGURE V.4 – Vue *inscription* de l'application *RabiT*[®] by *BillaTraining*[®].

Connexion Lors des connexions suivantes à l'espace *RabiT*[®] by *BillaTraining*[®], le coureur pourra s'identifier via la page de connexion. Il lui suffira d'entrer son email et son mot de passe afin d'être redirigé sur la page d'accueil. En cas d'échec, un message d'erreur apparaît. La figure V.5 présente la page *connexion*.

Mot de passe oublié Dans le cas où le coureur ne se rappelle plus de son mot de passe, nous avons prévu la page *mot de passe oublié*. L'utilisateur renseigne son identifiant, il reçoit ensuite un email afin de réinitialiser son mot de passe. Si l'identifiant n'existe pas (i.e. absent de la base de données), un message d'erreur apparaît à l'écran. La figure V.6 dévoile la page *mot de passe oublié*.

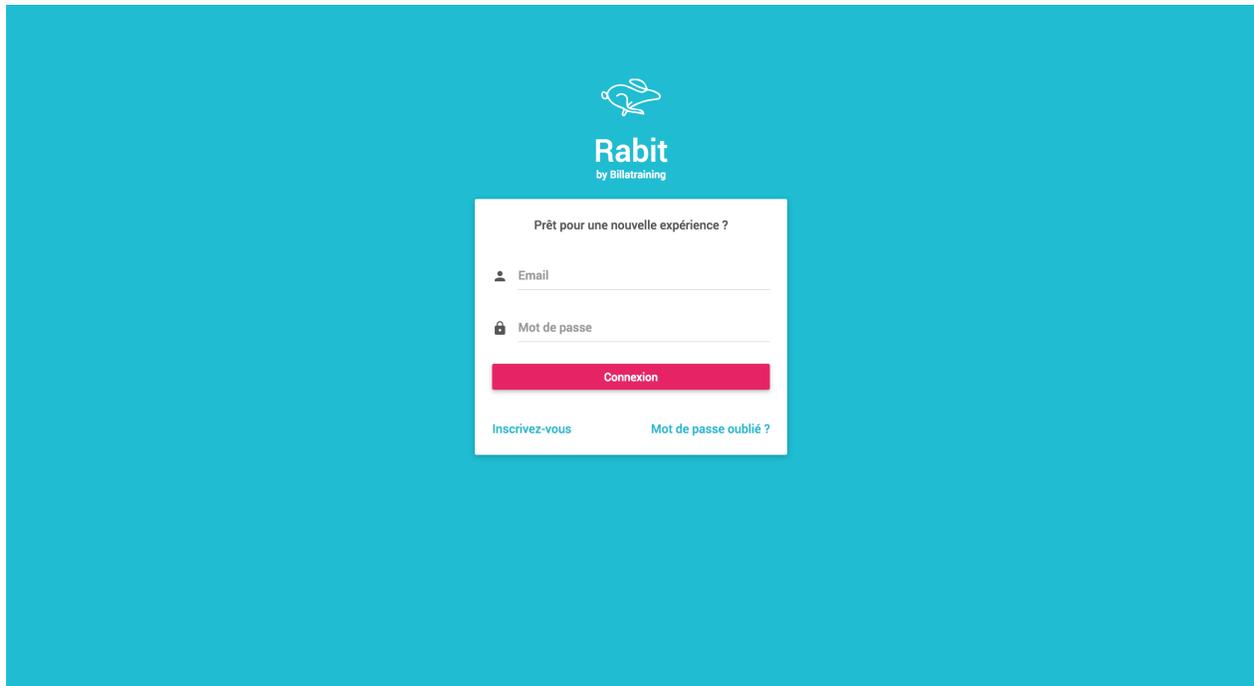


FIGURE V.5 – Vue *connexion* de l'application *RabiT*[®] by *BillaTraining*[®].

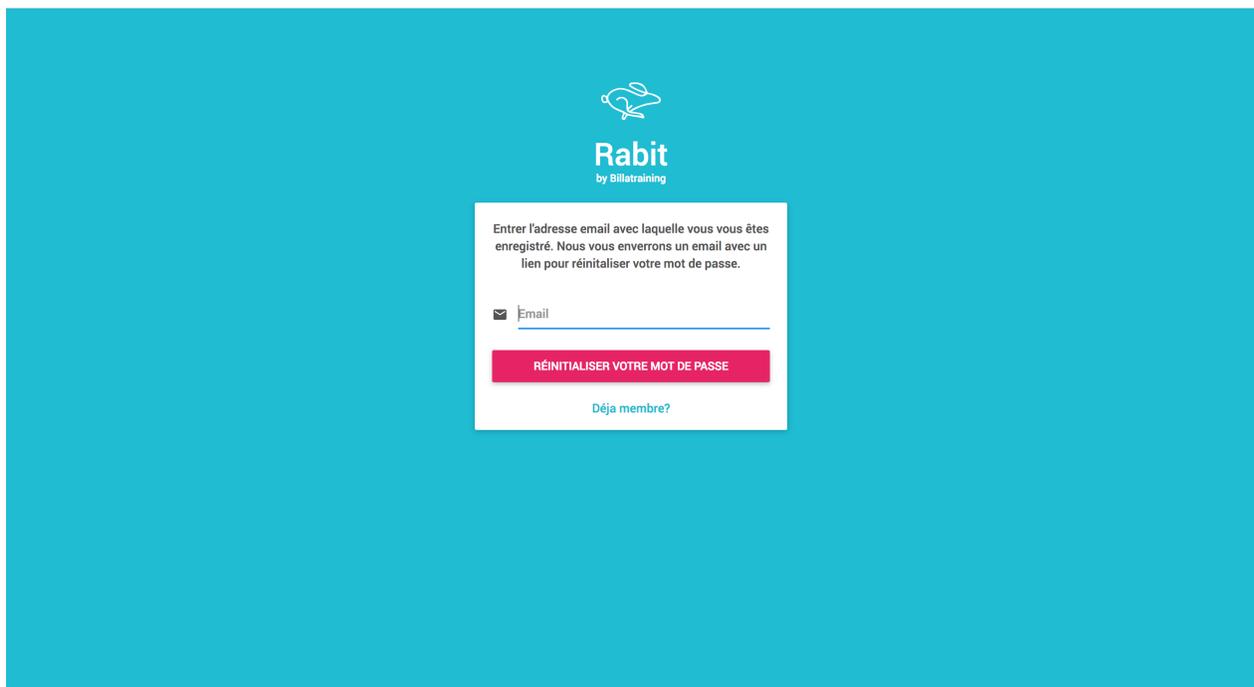


FIGURE V.6 – Vue *mot de passe oublié* de l'application *RabiT*[®] by *BillaTraining*[®].

Accueil Une fois identifié (inscription ou connexion), le coureur accède à la page *accueil*. Cette vue est un tableau de bord permettant de résumer la progression et les prochaines séances d'entraînement du coureur. On y retrouve cinq rubriques, appelées tuiles, à savoir :

- La forme de ma forme : c'est un graphique en forme de radar reprenant les 4 facteurs de la performance (Cardio, Perception, Acidose et Force). A chaque test de *RabiT*[®] réalisé par le coureur, un nouveau radar énergétique apparaît. Nous affichons uniquement les trois tests de *RabiT*[®] les plus récents. La figure V.7 donne une représentation du radar énergétique.
- Le calendrier : cette tuile sert de navigation. Il y est renseigné toutes les activités passées ainsi que les séances d'entraînement futures. La figure V.7 donne une représentation du calendrier.
- VO2max et Sensations : on y retrouve, d'une part, la valeur de VO2max en $ml.min^{-1}.kg^{-1}$ et, d'autre part, la variation de vitesse pour chacun des trois paliers de sensations, "facile", "moyenne" et "dure". La figure V.8 donne une représentation du calendrier.
- Dernière activité : affiche la dernière activité réalisée par le coureur. La figure V.8 donne une représentation de la dernière activité.
- Prochain Objectif : on peut y voir le prochain objectif du coureur. Le nombre de jour avant l'échéance, l'intitulé, la distance et la date y sont indiqués. La figure V.8 donne une représentation du prochain objectif.

Objectif Dans le but de préparer au mieux le coureur lors de son entraînement personnalisé, nous avons ajouté la possibilité de renseigner un objectif. Cette fonctionnalité est présente sur la page d'accueil et accessible via un *pop-up*. Trois informations sont alors nécessaires à la création d'un objectif, l'intitulé, la distance en kilomètres et la date. La figure V.9 présente la page *objectif*.

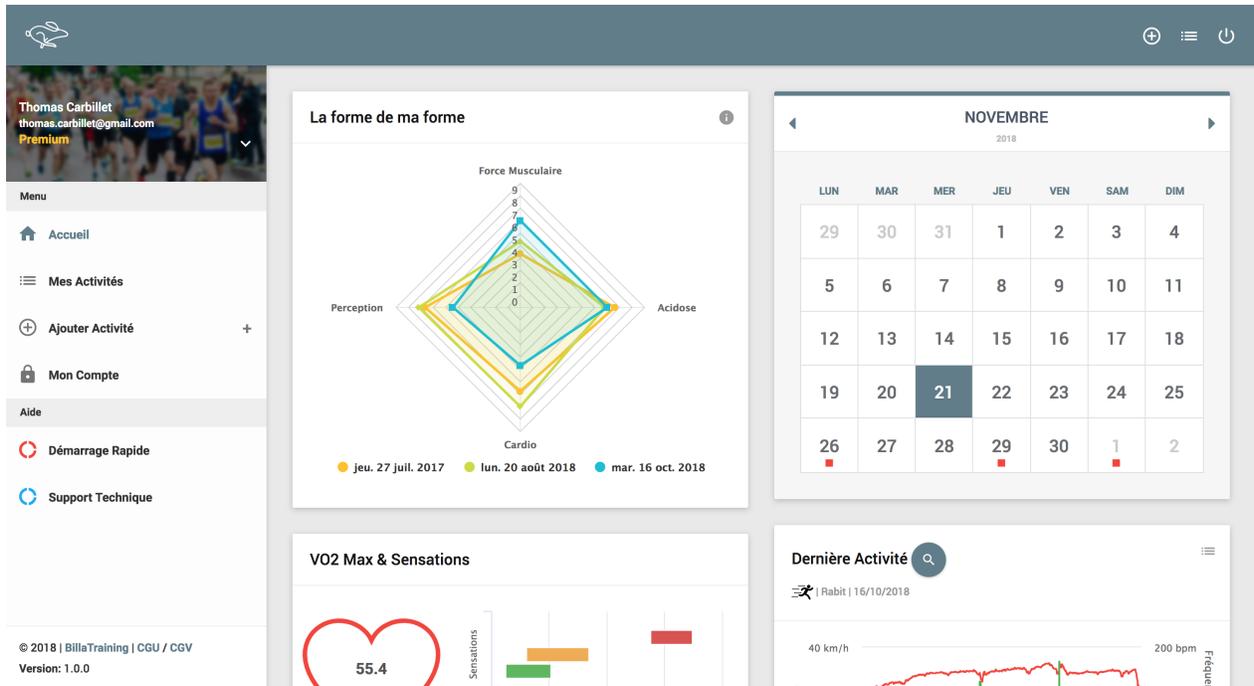


FIGURE V.7 – Première vue *accueil* de l'application *RabiT*[®] by *BillaTraining*[®].

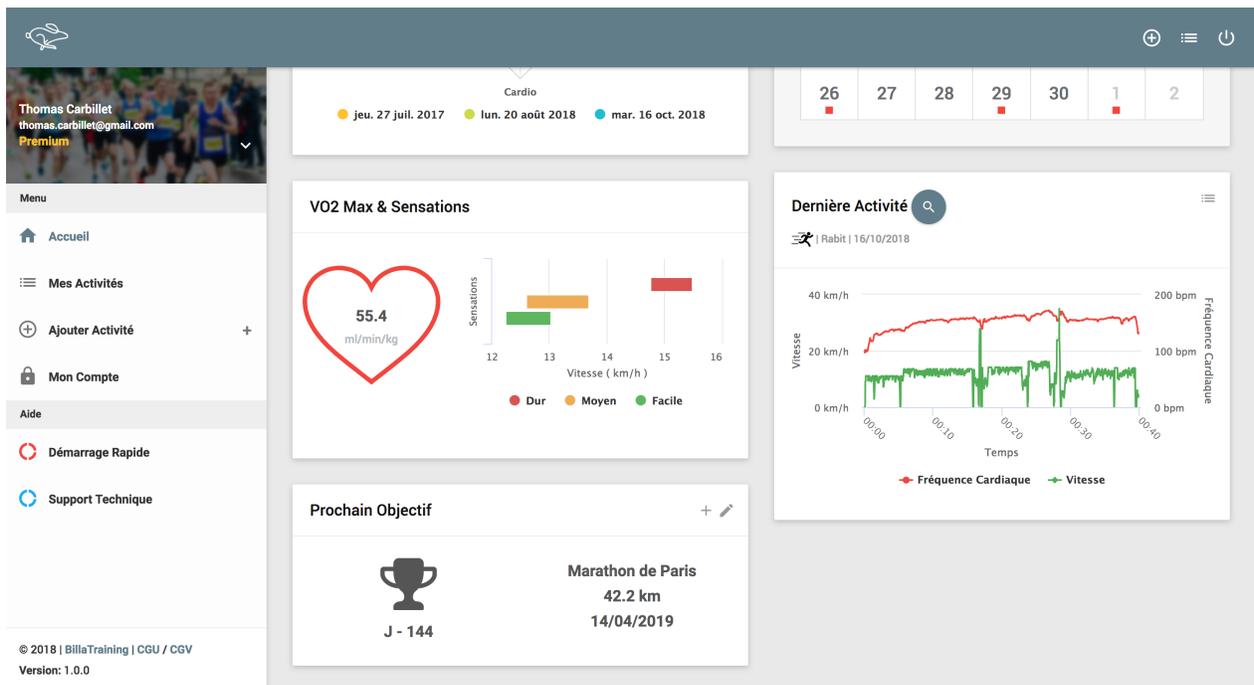


FIGURE V.8 – Seconde vue *accueil* de l'application *RabiT*[®] by *BillaTraining*[®].

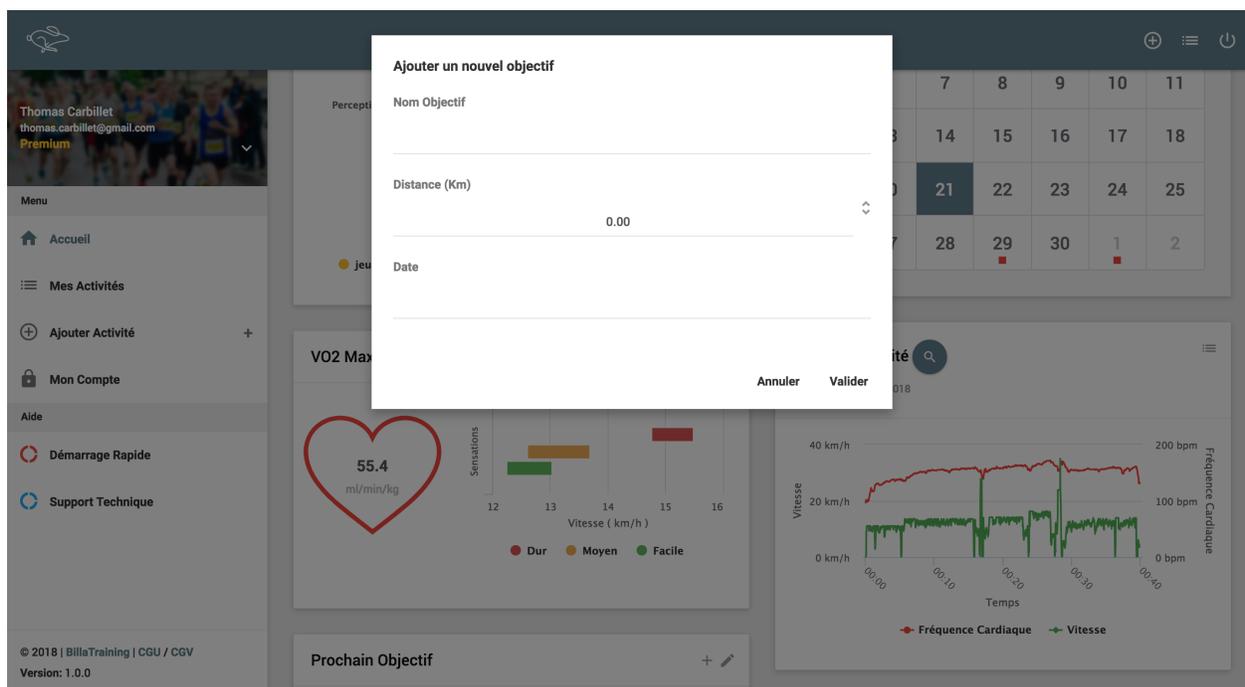


FIGURE V.9 – Vue *objectif* de l'application *RabiT*[®] by *BillaTraining*[®].

Mon compte Afin que le coureur puisse modifier ses informations personnelles, nous avons prévu la vue *mon compte*. Nous distinguons trois tuiles sur cette page.

- Informations Personnelles : on y retrouve le genre, le prénom, le nom de famille, la date de naissance, la taille et l'adresse email du coureur.
- Changement de Mot de Passe : l'utilisateur peut modifier son mot de passe
- Code Promo : ceci permet d'appliquer des réductions au programme premium (i.e. trois mois d'entraînement sur mesure). Nous utiliserons cette fonctionnalité pour nos partenaires ou encore pour des occasions spéciales telles que Noël ou anniversaire.

La vue *mon compte* est disponible sur la figure [V.10](#).

Démarrage rapide La page *démarrage rapide* donne les premières indications au coureur fraîchement inscrit sur le déroulé de la méthode *BillaTraining*[®]. A l'aide de schémas, de vidéos et de textes, l'objectif de cette vue est d'aider l'utilisateur à prendre en main la plateforme *RabiT*[®] by *BillaTraining*[®] le plus rapidement possible. Les trois figures [V.11](#), [V.12](#) et [V.13](#) donnent un aperçu de la page *démarrage rapide*.

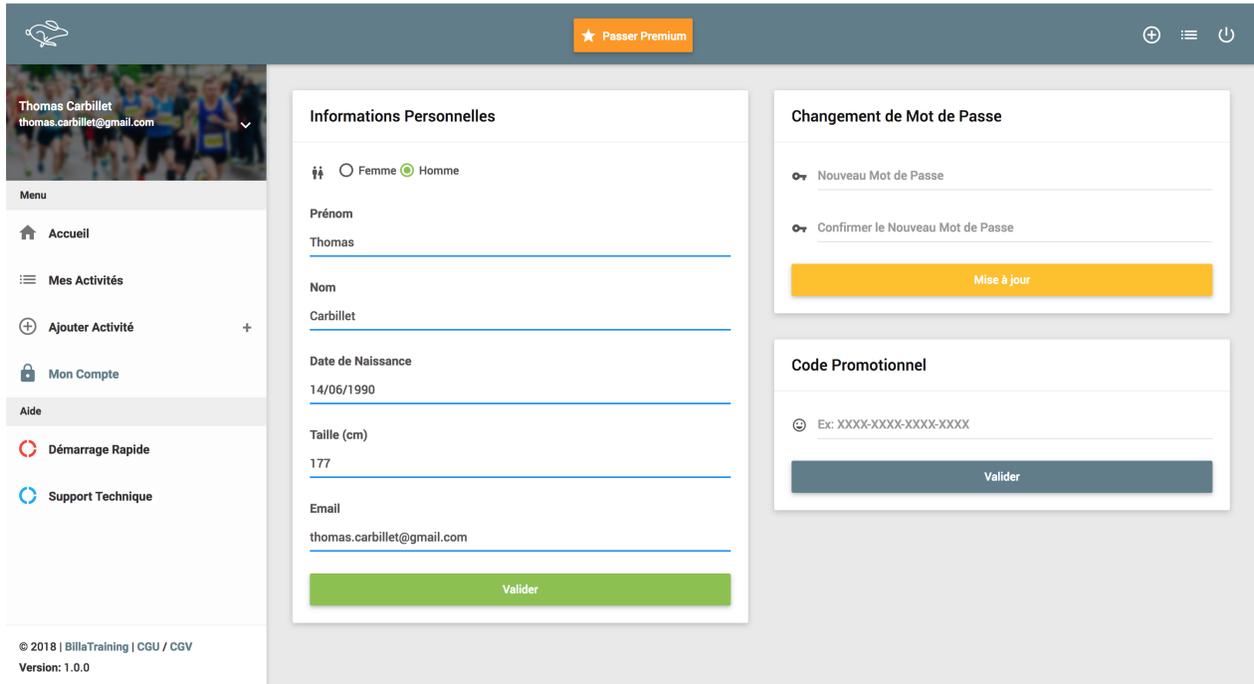


FIGURE V.10 – Vue *mon compte* de l'application *RabiT*[®] by *BillaTraining*[®].

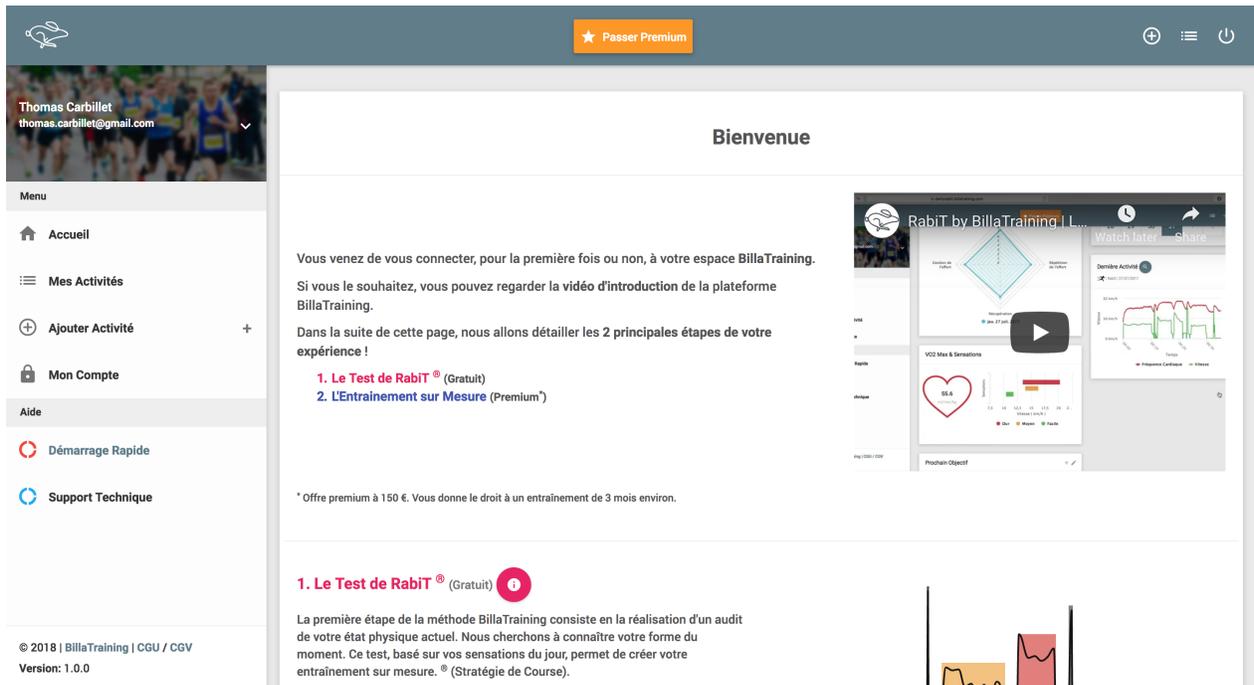


FIGURE V.11 – Première vue *démarrage rapide* de l'application *RabiT*[®] by *BillaTraining*[®].

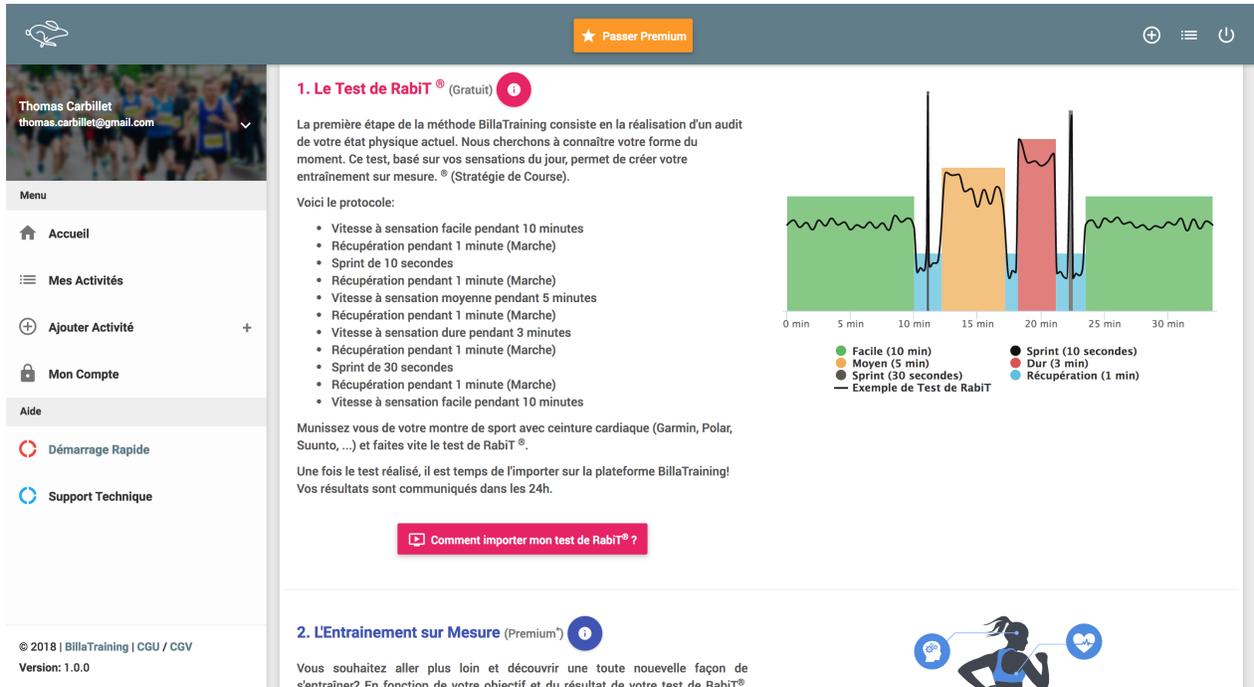


FIGURE V.12 – Seconde vue *démarrage rapide* de l'application *RabiT[®]* by *BillaTraining[®]*.

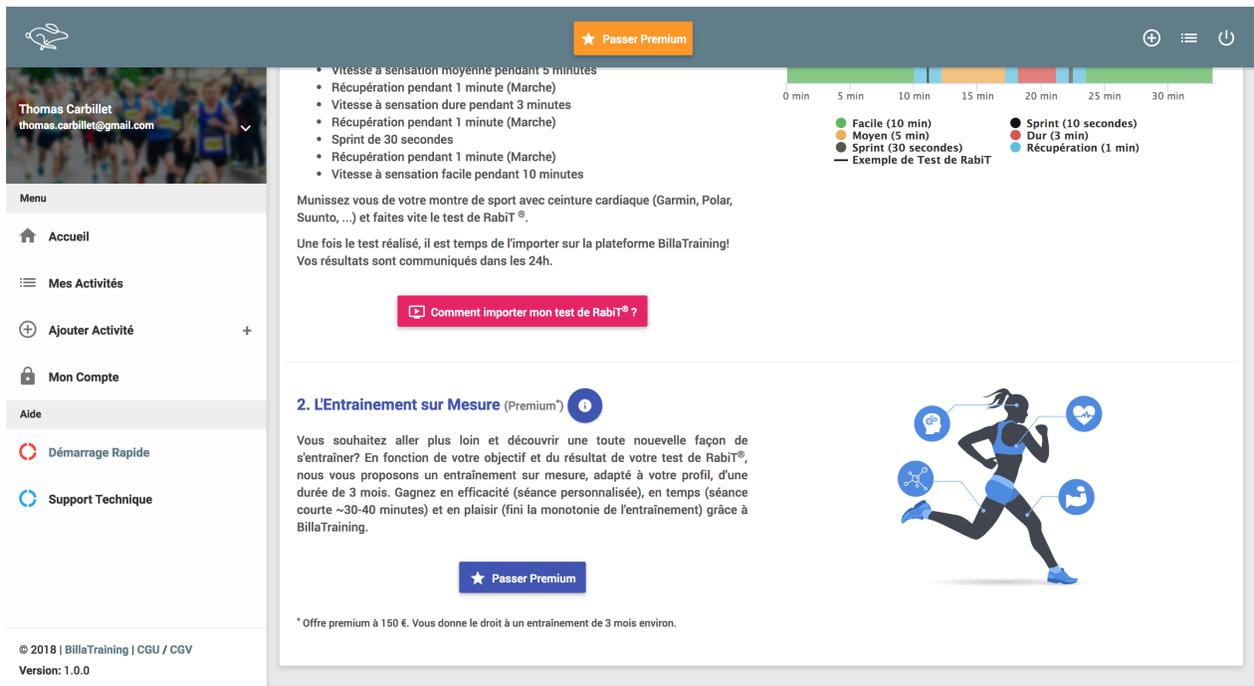


FIGURE V.13 – Troisième vue *démarrage rapide* de l'application *RabiT[®]* by *BillaTraining[®]*.

Ajout activité Pour générer puis mettre à jour le programme d'entraînement, nous demandons au coureur de charger les données issues de son dispositif connecté dans notre plateforme. Deux solutions s'offre au coureur dès lors. La première est le chargement manuel de fichier au format FIT⁸ ou TCX⁹. La seconde est la création d'un lien avec le compte *Strava*¹⁰ du coureur. Le coureur peut ensuite classer son activité parmi les catégories RabiT, Force Musculaire, Cardio, Perception, Acidose, Compétition ou Libre. Voir la figure V.14 pour une représentation graphique de la vue *ajout activité*.



FIGURE V.14 – Vue *ajout activité* de l'application *RabiT*[®] by *BillaTraining*[®].

Liste activités Sur cette page, on retrouve un tableau regroupant tout l'historique du coureur. Cela permet au coureur de visionner ou de modifier un activité passée. Dans ce tableau, on retrouve notamment les colonnes Catégorie, Date, Difficulté et Commentaire. La figure V.15 présente la vue *liste activités*.

⁸<https://en.wikipedia.org/wiki/FITS>

⁹https://en.wikipedia.org/wiki/Training_Center_XML

¹⁰*Strava* est une application web permettant de stocker les données d'entraînements et de courses d'un athlète. <https://www.strava.com/>

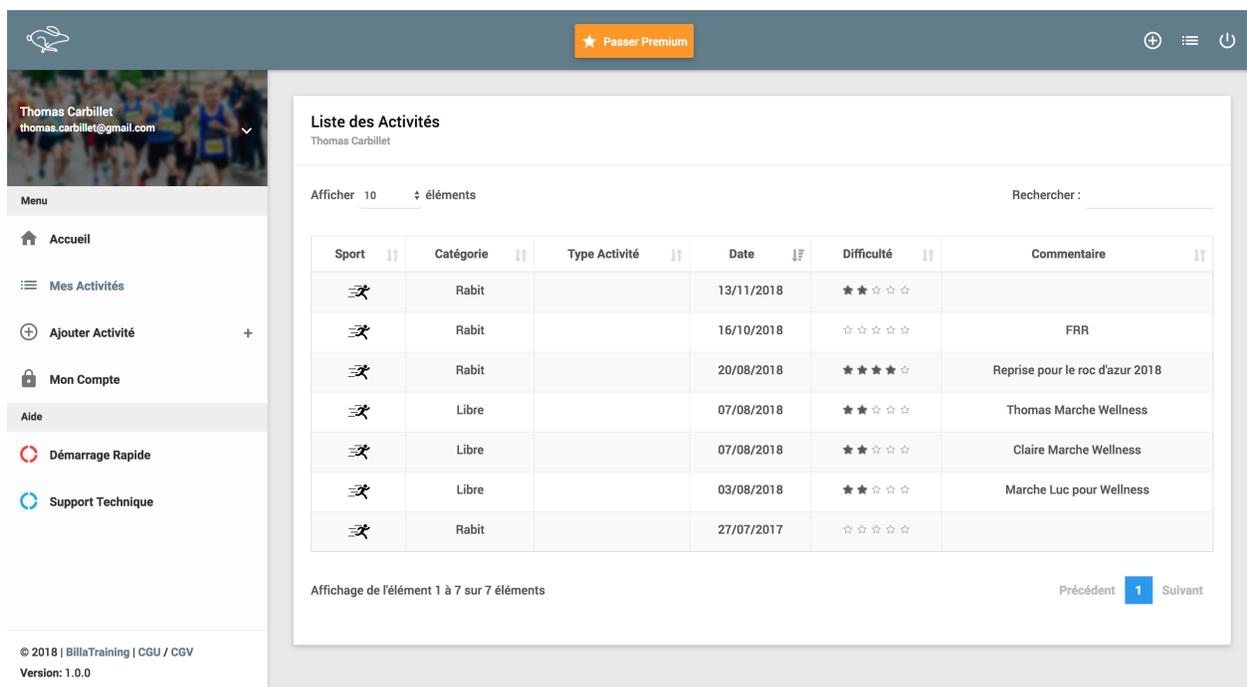


FIGURE V.15 – Vue *liste activités* de l'application *RabiT*[®] by *BillaTraining*[®].

Activité Cette vue permet au coureur de visualiser et commenter ses activités. On y retrouve les cinq tuiles suivantes :

- Indicateurs clés : au nombre de quatre, on y retrouve la distance parcourue, le temps total, la fréquence cardiaque maximal atteinte et la vitesse moyenne du coureur sur l'activité. Cela permet à l'utilisateur d'avoir un résumé rapide de sa séance.
- Trace *GPS* : le coureur peut retracer son activité sur un plan.
- Courbes : on y retrouve toutes les données issues du dispositif connecté que possède le coureur. Les courbes affichées peuvent varier en fonction du dispositif. Les données disponibles dans la majorité des cas sont la vitesse, la fréquence cardiaque et l'altitude.
- Difficulté de l'exercice : le coureur peut noter la difficulté de l'exercice de 1 à 5 (i.e. très facile à très dur).
- Décrivez vos sensations : le coureur peut laisser libre court aux sensations qu'il a éprouvé durant la séance. L'objectif final étant d'utiliser des techniques d'analyse de texte afin

de repérer la personnalité du coureur et encore améliorer la personnalisation de son entraînement.

Les figures V.16 et V.17 donnent la représentation graphique de la vue *activité*.

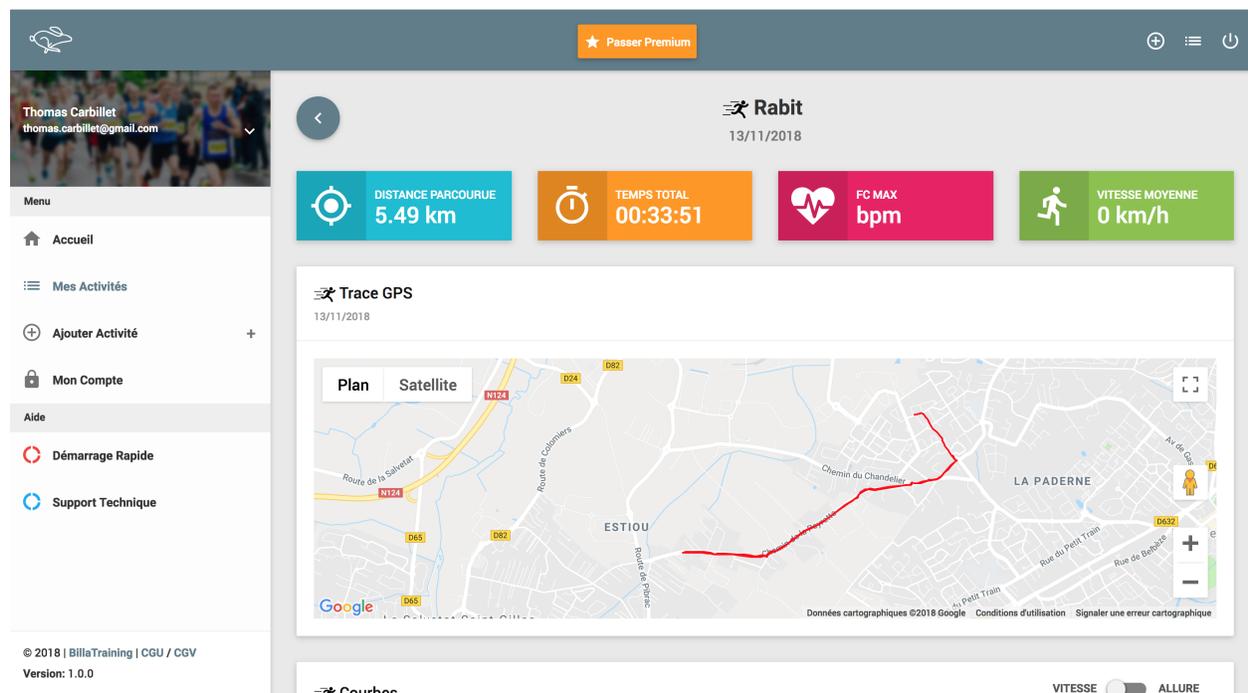


FIGURE V.16 – Première vue *activité* de l'application *RabiT*[®] by *BillaTraining*[®].

Premium Il s'agit d'un *pop-up* qui est transverse à toutes les pages du site. Le coureur peut alors payer pour bénéficier d'un entraînement personnalisé de trois mois. La figure V.18 montre le *pop-up premium* en action.

Support technique La vue *support technique* est également un *pop-up* transverse à toutes les pages de l'application. En cas de problème, le coureur peut ainsi envoyer un message directement à l'équipe *BillaTraining*[®]. La figure V.19 présente le *pop-up support technique* en action.

2.2.3 Les analyses scientifiques

La brique *analyses scientifiques*, comme le montre la figure V.20, possède deux fonctions. La première est la segmentation du test de *RabiT*[®] (c.f. chapitre II). La seconde est l'attribution

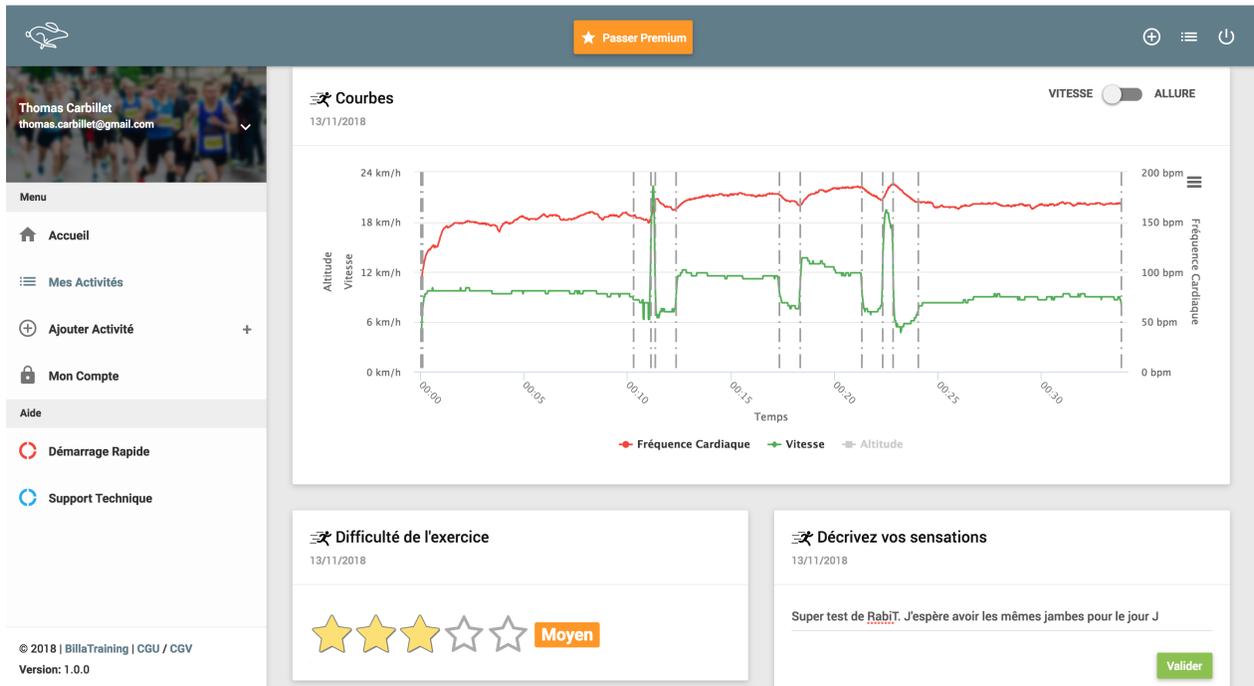


FIGURE V.17 – Deuxième vue *listeactivité* de l'application *RabiT*[®] by *BillaTraining*[®].

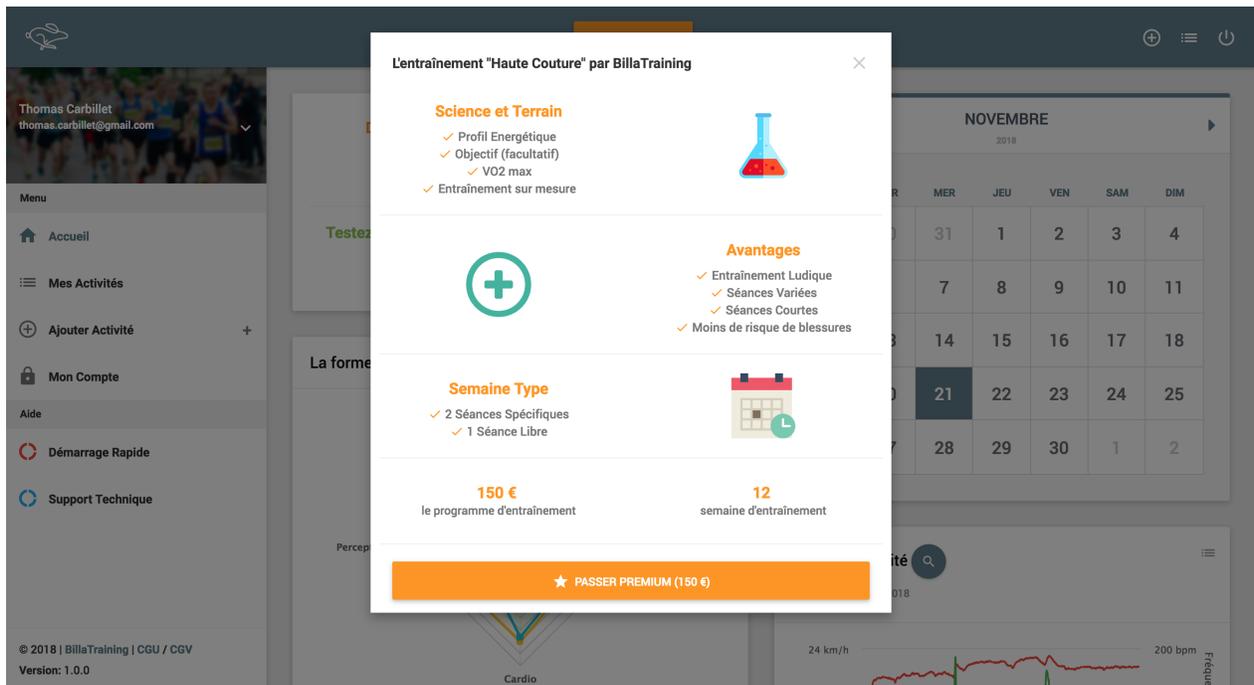


FIGURE V.18 – Vue *premium* de l'application *RabiT*[®] by *BillaTraining*[®].

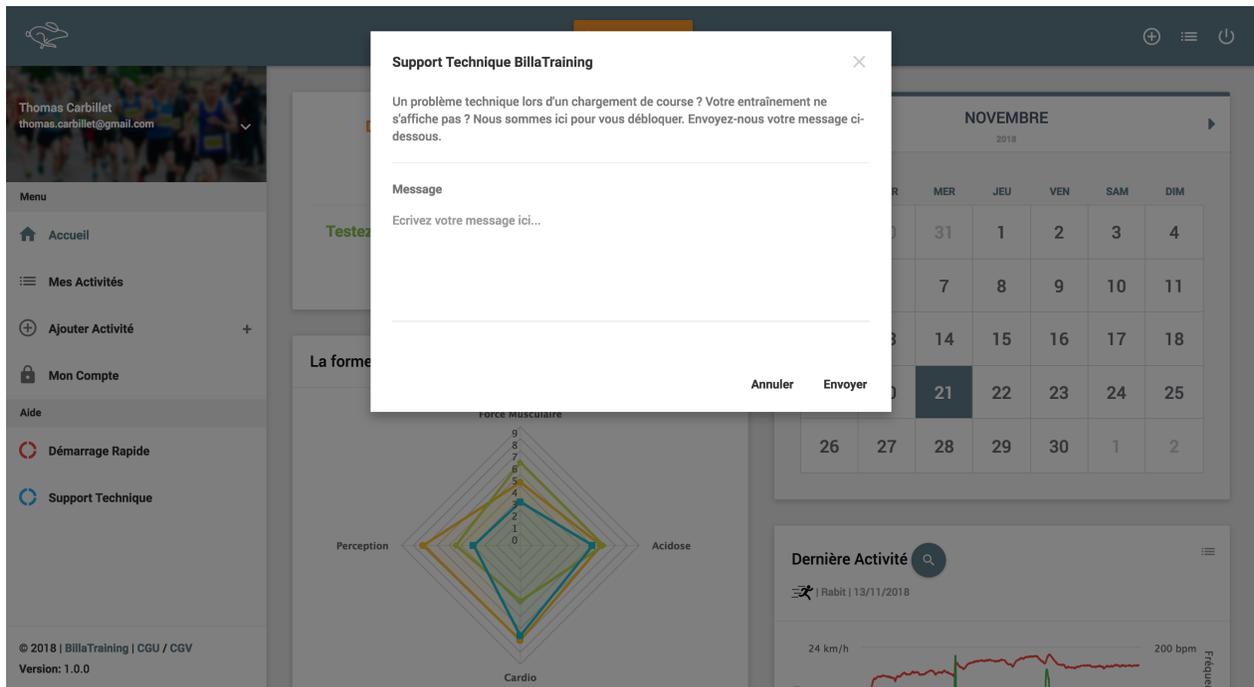


FIGURE V.19 – Vue *support technique* de l’application *RabiT*[®] by *BillaTraining*[®].

des notes du radar énergétique. L’algorithme étant confidentiel, il ne sera pas détaillé ici.

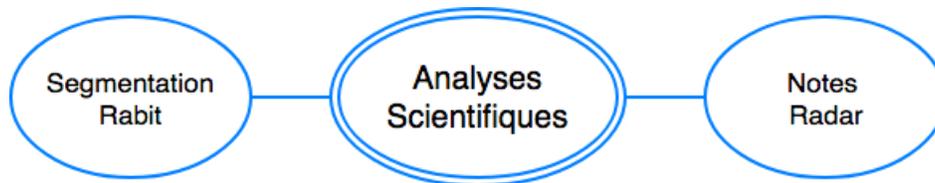


FIGURE V.20 – Schéma des fonctionnalités de la brique *analyses scientifiques* de l’application *RabiT*[®] by *BillaTraining*[®].

2.2.4 Le paiement en ligne

La dernière brique à laquelle nous devons venir connecter l’application *RabiT*[®] by *BillaTraining*[®] est le *paiement en ligne*. Comme nous cherchons à automatiser notre application de bout en bout, le choix d’une solution de paiement en ligne est donc indispensable. Dans le but de choisir la solution qui correspond le mieux à notre application, nous nous sommes fixés, à minima les trois critères suivants :

- Sécurisé : nous avons cherché une solution de paiement en ligne qui n’a jamais fait face à une intrusion. Par ailleurs, l’organisme devait être certifié PCI DSS (Payment Card

Industry Data Security Standard) avant d'éviter toute fraude de paiement.

- Fiable : afin de ne pas rater une commande, nous cherchons un système disponible en permanence. De plus, le système devra conserver un historique de toutes les commandes clients.
- Facile à utiliser : nous souhaitons un système facile à intégrer et flexible qui pourra grandir avec notre application (e.g. forte augmentation du nombre de commande).

Notre choix s'est porté sur la solution de paiement *Payline*. La figure V.21 permet de comprendre le fonctionnement de cette solution de paiement. Nous allons à présent détailler la

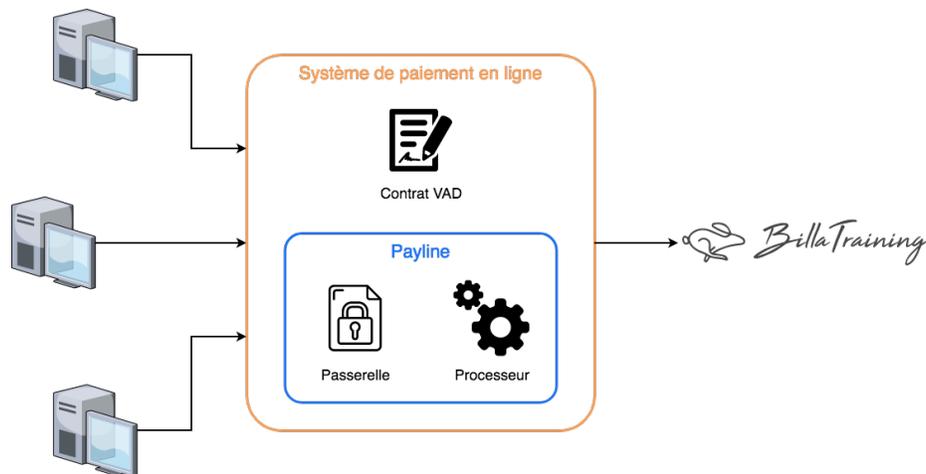


FIGURE V.21 – Schéma de fonctionnement du paiement en ligne de l'application *RabiT*[®] by *BillaTraining*[®].

figure V.21. Premièrement, sur la gauche de la figure, nous avons les clients. Ce sont ceux qui achètent l'entraînement de trois mois personnalisé. Munis de leur carte bancaire, nos clients rentrent leurs informations. Ensuite, les données bancaires transitent vers le système de paiement en ligne, c'est le centre de la figure V.21. D'une part nous avons un contrat VAD (Vente A Distance), délivré par notre banque, pour nous permettre d'encaisser les paiements en ligne. D'autre part, notre prestataire *Payline* chiffre les informations de paiement, grâce à la passerelle, et vérifie si le paiement est autorisé, grâce au processeur. Finalement, le paiement parvient à l'entreprise *BillaTraining*[®].

D'autres solutions de paiement ont été envisagées en amont. Nous pouvons en citer deux

et expliquer les raisons principales pour lesquelles nous les avons écartées. La première est *Stripe*¹¹. Même si son intégration est très simple et ne nécessite pas de contrat VAD avec la banque, nous l'avons écartée car les commissions sur chaque paiement sont élevées. La seconde est une solution propriétaire. Bien que moins chère que *Stripe* ou *Payline*, cette solution nécessite la certification PCI DSS. Il est par conséquent plus difficile de la mettre en place.

3 Prototype d'une application mobile temps réel

En parallèle du développement de l'application web d'entraînement *RabiT*[®] by *BillaTraining*[®], nous avons eu l'occasion de développer un prototype d'application mobile capable d'interagir en temps réel avec le coureur. Nous avons dû choisir entre deux approches, le mobile natif ou le mobile hybride. Les deux méthodes donnent un rendu assez similaire pour des applications simples. Nous allons étudier les avantages et les inconvénients de ces deux méthodes.

Le développement mobile natif consiste à développer une application mobile par plateforme (e.g. iOS, Android, ...). Ceci implique la connaissance des langages de programmation propre à chaque plateforme, par exemple, *Swift* et *Objective-C* pour les applications distribuées sur iOS et *Java* et *Xml* pour Android. Nous avons identifié les avantages suivants :

- Rapide, l'application charge vite.
- Réactive, les interactions avec l'utilisateur sont prises en compte sans délais.
- Très bonne interface utilisateur grâce à l'utilisation des composants natifs (e.g. boutons, couleurs, champs de formulaire, ...).
- Accès à tous les capteurs du téléphone (e.g. appareil photo, puce NFC, capteur de proximité, ...).

Cependant, la maîtrise nécessaire de plusieurs langages de programmation et le fait de devoir

¹¹<https://stripe.com/>

maintenir deux dépôts de code source rend l'utilisation de cette méthode chronophage et coûteuse.

Le développement mobile hybride permet l'écriture d'un seul code source commun aux plateformes iOS et Android. Le code est construit en Javascript, HTML et CSS. S'agissant de technologies que nous connaissons déjà, nous n'avons pas besoin de consacrer du temps à l'apprentissage de nouveaux langages. Avec cette approche, il est également possible d'interagir avec certains capteurs du téléphone (e.g. *GPS*, Accéléromètre, ...) grâce à l'utilisation d'un *wrapper* jouant le rôle de pont avec les fonctionnalités natives du téléphone. Voici la liste des avantages que nous avons identifié :

- Un seul code source pour toutes les plateformes.
- Facile à construire grâce aux langages HTML/CSS/Javascript.
- Accès à la majorité des capteurs du téléphone.
- Rapide à développer

Les limites de cette méthode sont au niveau des performances, plus limitées que le développement natif. Etant donné que nous cherchons à construire un prototype et que nous n'avons pas les ressources humaines nécessaires au développement natif, nous avons choisi d'utiliser les techniques du développement hybride.

Cette section présente l'architecture et les fonctionnalités du prototype.

3.1 Architecture de l'application mobile

Nous allons, dans un premier temps, présenter l'architecture de l'application mobile. La figure V.22 présente l'architecture. Nous avons ajouté trois briques supplémentaires à l'architecture de l'application web (c.f. figure V.1).

- Mobile UI/UX : Le développement de cette brique a été réalisé avec *Ionic*¹² qui intègre le *framework Angular*. Nous avons essayé de nous rapprocher le plus possible d'une

¹²<https://ionicframework.com/>

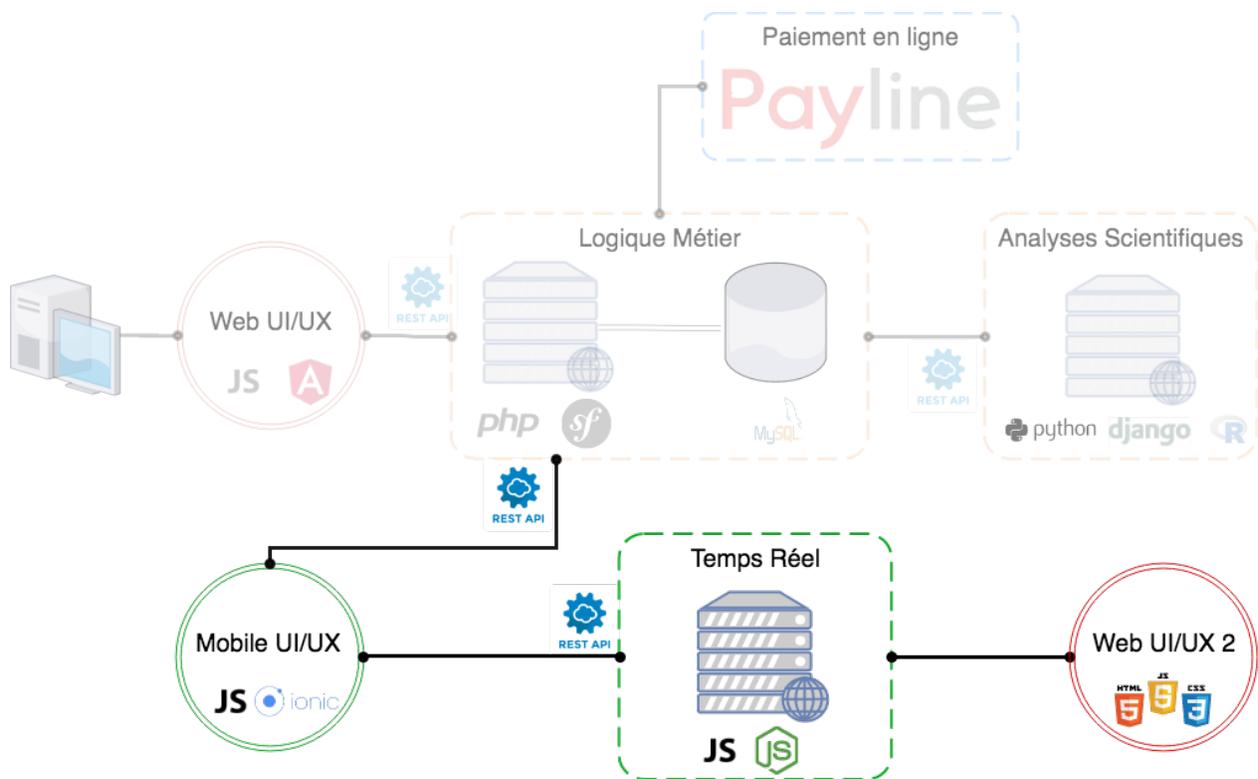


FIGURE V.22 – Schéma globale de l'application mobile *RabiT*[®] by *BillaTraining*[®]

expérience et d'une interface utilisateur native.

- Web UI/UX 2 : Développé avec le trio HTML/CSS/Javascript, il s'agit d'une vue très simple permettant d'interagir en temps réel avec les utilisateurs de l'application mobile.
- Temps réel : Afin de pouvoir interagir avec le coureur pendant son entraînement / sa course, nous avons du créer une nouvelle brique regroupant les fonctionnalités liées au temps réel. Nous avons choisi l'environnement serveur *Node.js*¹³ qui permet de développer en utilisant le langage Javascript côté serveur. L'intérêt de ce serveur est qu'il peut gérer les entrées et sorties de manière asynchrone et en temps réel.

Nous pouvons réutiliser la brique *logique métier*, et donc toutes ses fonctionnalités, que nous avons développées pour l'application web. C'est l'avantage de séparer le *front-end* et le *back-end*.

¹³<https://nodejs.org/en/>

3.2 Les fonctionnalités de l’application mobile

Nous allons présenter, dans cette sous-section, l’ensemble des fonctionnalités relatives à l’application mobile. Le projet a pour but de visionner, en temps réel, la vitesse et la fréquence cardiaque du coureur à pied et de pouvoir interagir avec ce dernier lors de son entraînement ou sa course.

3.2.1 Le temps réel

Interagir en temps réel avec le coureur est un des objectifs de cette thèse. D’une part, nous souhaitons récupérer les données de vitesse et fréquence cardiaque en temps réel, et, d’autre part, nous souhaitons pouvoir donner des consignes et poser des questions au coureur lors de son entraînement. Les *WebSockets* semblent la solution la plus appropriée pour résoudre notre problématique.

Afin de comprendre ce qu’est un *WebSocket*, commençons par détailler les trois couches du modèle *TCP/IP* [114]. La figure V.23 donne une représentation du modèle. En bas de la figure

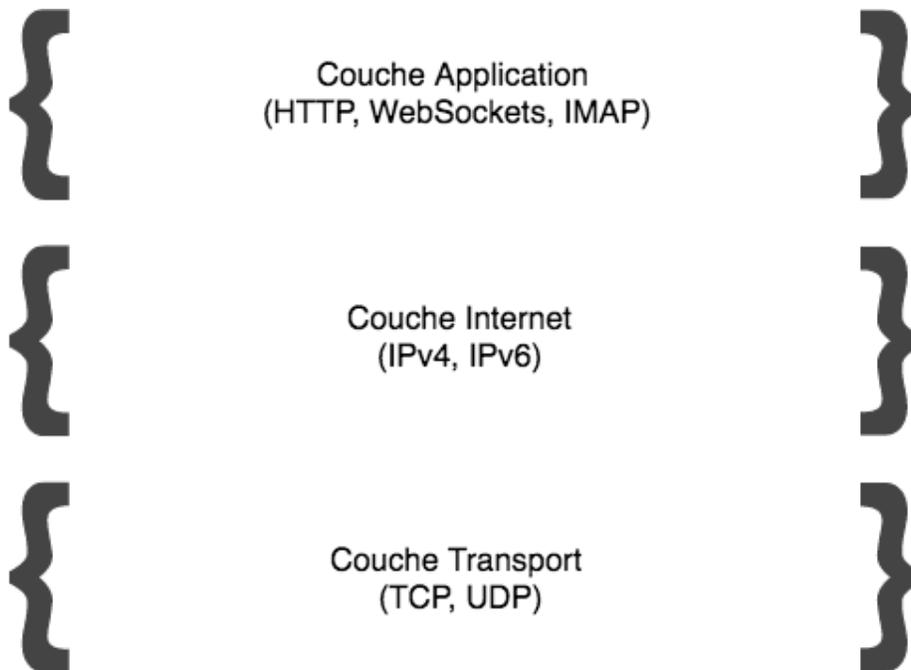


FIGURE V.23 – Les trois couches du modèle *TCP/IP*

V.23, on retrouve la couche de transport. C’est le domaine d’action des protocoles *TCP*

(Transmission Control Protocol) et *UDP* (User Datagram Protocol). Les deux protocoles permettent d'envoyer des bits de données (ou paquets) sur internet. En d'autres termes, ils servent de navettes permettant de transiter des données d'un point A à un point B. La différence principale entre les deux protocoles est que le *TCP* s'assure que les paquets qu'il envoie arrive à destination et dans le bon ordre alors que l'*UDP* envoie les paquets en continu sans s'assurer de leurs bonne réception. Le *TCP* est largement plus utilisé que l'*UDP*.

Au centre de la figure V.23, la couche internet avec l'attribution des fameuses adresses IP (Internet Protocol). Si l'on reprend notre exemple de navettes transitant des données, les points A et B ont alors chacune une adresse IP unique. Cela permet d'identifier qui envoie les données et qui les reçoit.

Finalement, en haut de la figure V.23, la couche application permet de créer, modifier ou encore supprimer les données à envoyer par paquet. On y retrouve des protocoles comme HTTP, IMAP (envoi / réception d'email) ou WebSockets. Dans notre exemple de navettes, cette couche permet de savoir à quel endroit du point A ou B livrer les données.

On sait désormais que le protocole WebSocket se trouve dans la couche application du modèle *TCP/IP* tout comme le protocole HTTP. Présentons tout d'abord le protocole HTTP. Le protocole HTTP envoie une requête au serveur depuis le client et le serveur répond au client. Le désavantage d'un tel mécanisme est que chaque fois que le client envoie des informations au serveur ou souhaite mettre à jour ses données, l'utilisateur est obligé de rafraîchir la page en formulant une nouvelle requête. C'est le principe de la barre d'url dans un navigateur.

Afin de ne pas être obligé de rafraîchir la page sans arrêt, il est possible, depuis 2005, de faire appel à la technologie AJAX¹⁴. Cela permet d'effectuer des requêtes HTTP sans recharger la page. Cependant, l'utilisateur doit demander la mise à jour, depuis le client, au serveur en cliquant sur un bouton par exemple.

Pour aller plus loin dans le temps réel, il est préférable que l'utilisateur n'est pas à interagir avec le client afin de demander les mises à jour côté serveur. C'est dans ce but bien précis que sont nés les WebSockets en 2011. Une seule communication *TCP* est établie lors de la

¹⁴[https://en.wikipedia.org/wiki/Ajax_\(programming\)](https://en.wikipedia.org/wiki/Ajax_(programming))

première connexion du client au serveur (requête). Les mises à jour se font ensuite automatiquement. Ceci fait du WebSocket un protocole de communication full-duplex (i.e. le client et le serveur peuvent s'échanger des informations en même temps). Nous pouvons donc voir les WebSockets comme une amélioration du protocole HTTP.

Nous avons utilisé le langage de programmation Javascript et le *framework* Node.js afin d'implémenter les fonctionnalités de temps réel de notre application mobile côté *back-end*.

3.2.2 Le mobile UI/UX

La brique *mobile UI/UX* est le *front-end* de l'application mobile. La figure V.24 rassemble les cinq fonctionnalités de cette brique. Nous allons les détailler une par une.

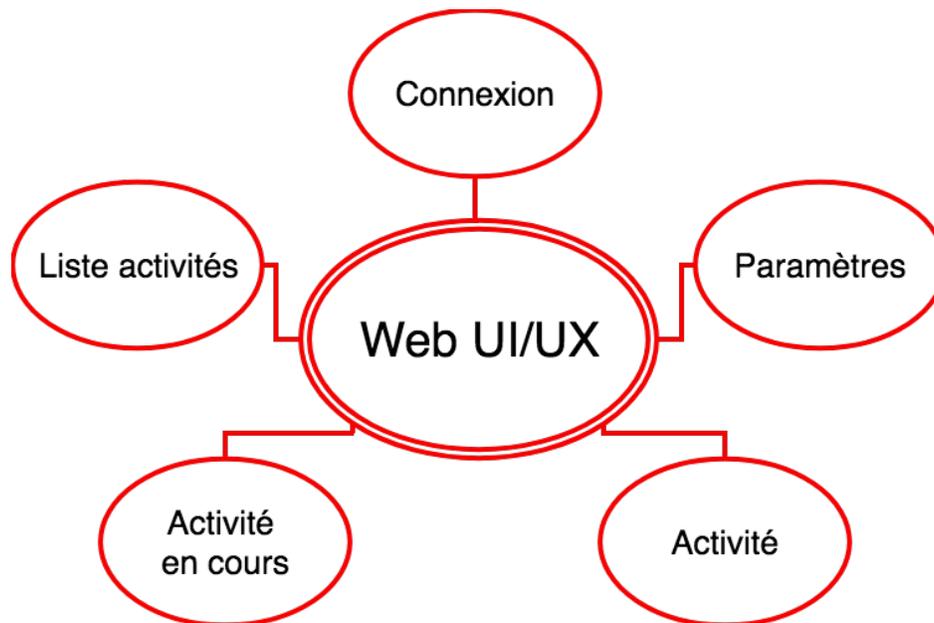


FIGURE V.24 – Schéma des fonctionnalités de la brique *mobile UI/UX* de l'application mobile *RabiT*[®] by *BillaTraining*[®]

Connexion Sur le même modèle que l'application web, l'utilisateur devra se connecter à l'application mobile en utilisant un couple email / mot de passe. La brique *logique métier* étant commune aux deux applications, l'utilisateur conserve le même couple d'identifiant. En cas d'échec, un message d'erreur apparaîtra. La figure V.25 présente la vue *connexion*.

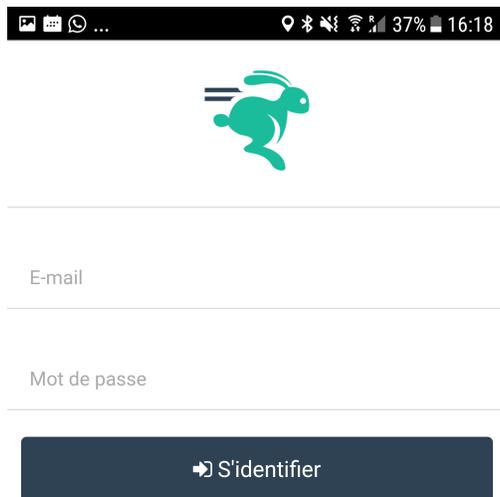


FIGURE V.25 – Vue *connexion* de l'application mobile.

Entraînement Grâce à la fonctionnalité *entraînement*, le coureur peut retrouver la liste de ses séances d'entraînement (c.f. figure V.26) et peut voir le détail de chaque séance à l'aide d'un bref descriptif et d'un schéma explicatif (c.f. figure V.27). Le type d'activité (cardio, force, perception ou acidose) est identifiable par le biais d'icônes.

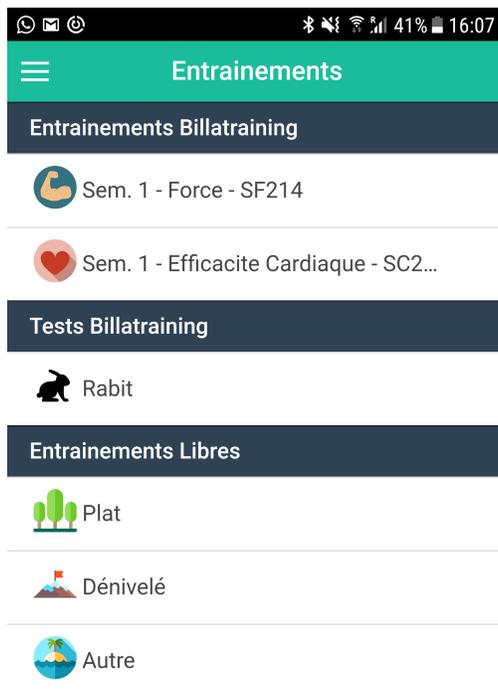


FIGURE V.26 – Première vue *entraînement* de l'application mobile.

Activité La fonctionnalité *activité* permet d'enregistrer les données de vitesse et de fréquence cardiaque du coureur. Cette fonctionnalité interagit avec les capteurs du téléphone portable afin de récupérer la vitesse instantanée et le nombre de battements cardiaques par minute. Pour la vitesse, lorsque le coureur clique sur le bouton "GPS" (c.f figure V.28), l'application, par le biais d'*Ionic*, déclenche la géolocalisation¹⁵.

Ainsi, il nous est possible de récupérer la longitude, la latitude et l'altitude en temps réel du coureur. Pour la fréquence cardiaque, lorsque le coureur appuie sur le bouton "Ceinture"

¹⁵<https://ionicframework.com/docs/native/geolocation/>



FIGURE V.27 – Deuxième vue *entraînement* de l'application mobile.

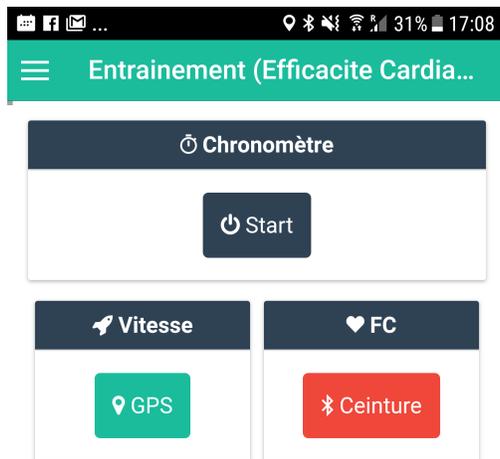


FIGURE V.28 – Première vue *activité* de l'application mobile.

(c.f figure V.28), nous pouvons la récupérer uniquement si le coureur porte une ceinture pectorale *bluetooth*. Dans ce cas, *Ionic* nous permet d'interagir avec le le *bluetooth* basse consommation¹⁶ (*Bluetooth Low Energy* en anglais). L'enregistrement de la fréquence cardiaque est optionnel alors que celui de la vitesse est obligatoire.

Une fois la vitesse et la fréquence cardiaque lancées, le coureur peut appuyer sur le bouton "Start" (c.f figure V.29).

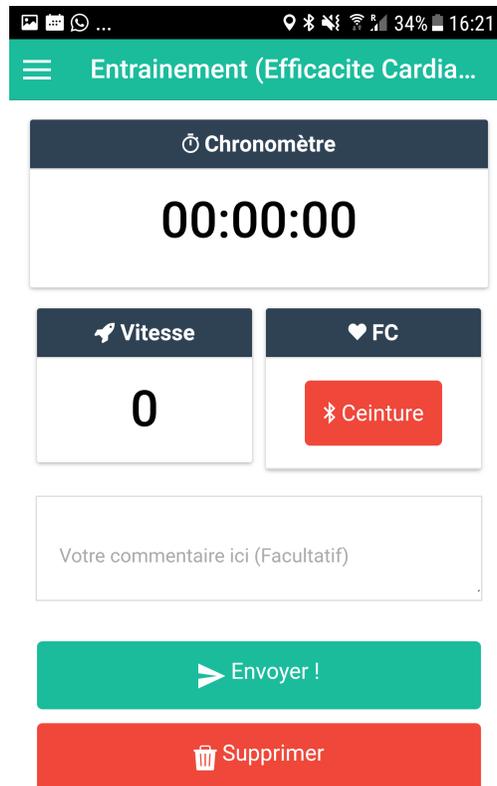


FIGURE V.29 – Deuxième vue *activité* de l'application mobile.

Le chronomètre démarre et des consignes vocales sont donnés au coureur pendant son entraînement. A la fin de son activité, le coureur peut laisser un commentaire puis appuyer sur "Envoyer" (c.f figure V.29).

Finalement, comme on peut le voir sur la figure V.30, le coureur peut visionner le tracé de ses courbes de vitesse et de fréquence cardiaque.

¹⁶<https://ionicframework.com/docs/native/ble/>



FIGURE V.30 – Troisième vue *activité* de l'application mobile.

Liste activités La fonctionnalité *liste activités* donne un accès à l'historique du coureur. On peut en voir un exemple sur la figure V.31.

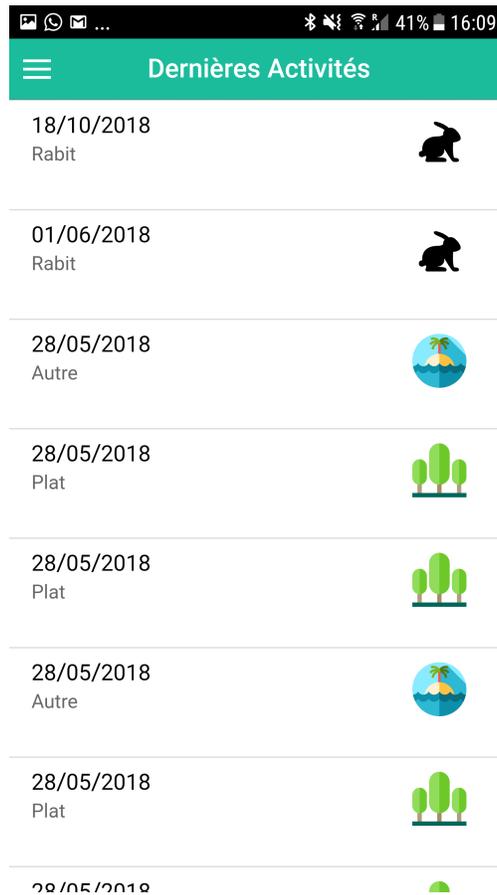


FIGURE V.31 – Vue *liste activités* de l'application mobile.

Paramètres Nous avons regroupé les deux options disponibles dans l'application dans la vue *paramètres* (c.f. figure V.32).

La première option permet de modifier la voix utilisées pour les consignes vocales. Le choix se porte entre une voix féminine ou une voix masculine. La deuxième option permet d'activer ou désactiver l'enregistrement en temps réel des données de vitesse et de fréquence cardiaque lorsque le coureur le désire. Par l'activation du temps réel, l'application mobile se connecte au serveur *temps réel* (voir sous-sous-section 3.2.1) via *WebSocket*. Dès lors que le coureur lance une activité, les données de vitesse et de fréquence cardiaque sont envoyées au serveur *temps réel* qui envoie lui même les données à la brique *web UI/UX 2*.

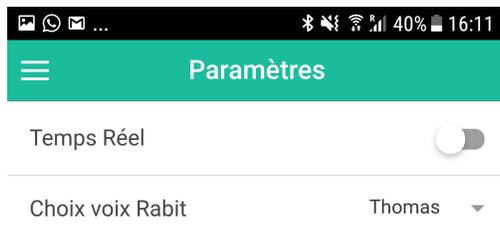


FIGURE V.32 – Vue *paramètres* de l'application mobile.

3.2.3 Le web UI/UX 2

La brique *web UI/UX 2* permet de visualiser les données de fréquence cardiaque et de vitesse du coureur en temps réel lorsque l'option temps réel est activée. Afin de permettre à un physiologiste d'interagir avec un coureur à pied, nous avons créé une interface web, simple, dotée de deux vues.

La première vue donne la liste de tous les coureurs ayant activés l'option temps réel en activité, la figure V.33 en donne un exemple.

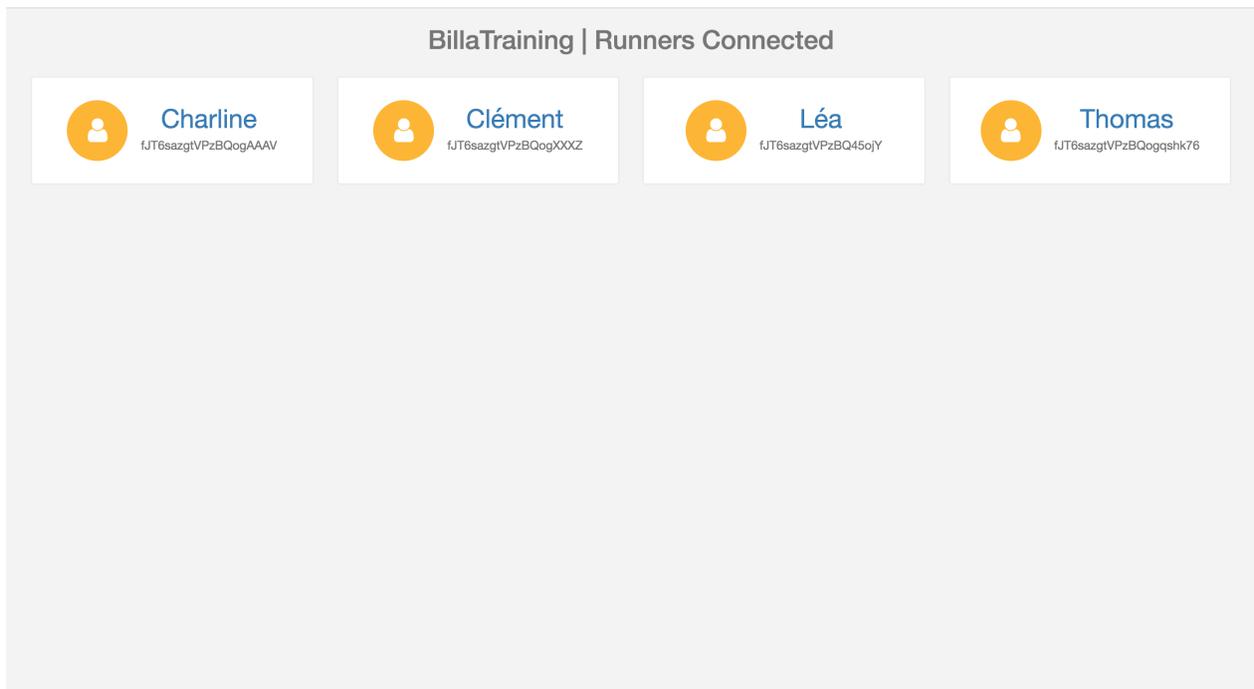


FIGURE V.33 – Première vue de la brique *web UI/UX 2*. On y distingue 4 coureurs à pied connectés en temps réel.

La seconde vue s'affiche lorsque le physiologiste clique sur un coureur afin d'avoir le détail de son activité (c.f. figure V.34).

Un graphique contenant les données de vitesse et de fréquence cardiaque apparaît. Ce graphique se met à jour dès qu'une nouvelle donnée est reçue sur le serveur *temps réel*. D'autre part, on retrouve une multitude de tuiles contenant diverses données afin d'aider le physiologiste à interpréter les courbes de vitesse et fréquence cardiaque. On y retrouve la fréquence cardiaque maximale, minimale et instantanée, la vitesse instantanée, le temps total depuis le début de l'activité et la distance totale. On peut également apercevoir une tuile (en bas de

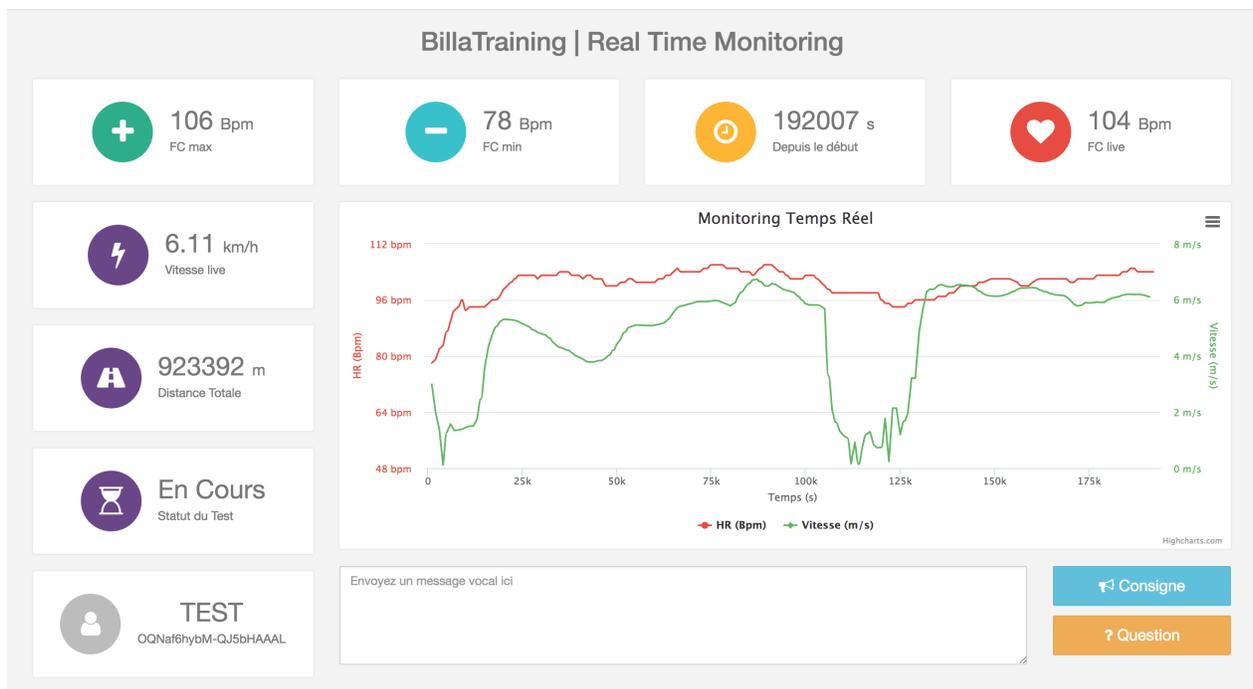


FIGURE V.34 – Deuxième vue de la brique *web UI/UX 2*.

la Figure V.34) permettant de donner une consigne ou poser une question au coureur.

A terme, il est prévu d'intégrer ce prototype web à l'application web *RabiT[®]* by *BillaTraining[®]*.

Bilan

Cet ultime chapitre nous a permis de détailler la stratégie digitale de l'entreprise *BillaTraining[®]*. L'application *RabiT[®]* by *BillaTraining[®]* a été déployée en production fin juin 2018. Elle est disponible à l'adresse <https://rabit.billatraining.com>. Les prochaines étapes liées sont liées au développement de l'application. Nous souhaitons itérer en ajoutant de nouvelles fonctionnalités et en supprimant celle inutiles. C'est en discutant avec nos utilisateurs que nous pourrions comprendre la façon d'améliorer l'application et de trouver le bon prix de vente de notre entraînement personnalisé. Un autre projet en cours est la traduction des contenus en anglais dans le but d'élargir notre audience.

L'application mobile, quand à elle, nécessite encore du travail avant d'être déployée sur les plateformes *Apple Store* et *Google Play*. En effet, nous avons remarqué que l'utilisation du capteur *GPS* seul ne donnait pas des données assez précises. La solution se trouve dans la

fusion de capteurs dans le but de créer un signal de vitesse de meilleure qualité. Ce projet est en cours au sein de l'entreprise *BillaTraining*[®].

Conclusion

En réponse aux besoins d'innovation de l'entreprise *BillaTraining*[®], ce travail de thèse propose une approche globale et pluridisciplinaire ayant deux objectifs. Le premier objectif est de poser les fondations permettant de mieux comprendre comment courir un marathon de façon optimale dans le but de proposer, à terme, un modèle prédictif de stratégie de course individualisée. Le second objectif est de créer une plateforme automatisée intégrant le savoir-faire de l'entreprise *BillaTraining*[®] dans le but d'assurer son développement commercial et de construire une passerelle entre la recherche en course à pied et le grand public.

Les deux sections suivantes résument les principaux résultats de ce travail de thèse et présentent certaines perspectives qui permettraient à la fois d'améliorer ces travaux et d'en exploiter les performances.

Résultats

Nous allons présenter cette section consacrée aux résultats en deux parties. La première partie sera dédiée à la stratégie de course sur marathon. La seconde partie traitera de la plateforme automatisée d'entraînement en course à pied.

Stratégie de course sur marathon

Durant ce travail de thèse, nous avons eu l'occasion d'étudier les données de vitesse issues de près de 300 marathons sous forme de temps de passage au kilomètre. Cette étude, purement descriptive, a donné lieu à la création d'un outil statistique permettant de qualifier la performance d'un coureur à pied. L'outil, que nous appelons signature statistique, est une

matrice deux lignes, deux colonnes dont le but est d'avoir un regard croisé sur la tendance et l'asymétrie du signal de vitesse lors d'un marathon indépendamment du genre, du parcours et de la performance. Il s'agit d'un indicateur permettant de voir si le coureur a réalisé une course optimale ou s'il aurait pu faire mieux.

Pour la tendance, si un coureur a une vitesse globalement décroissante sur le marathon, il sera noté (-). Au contraire, si la vitesse est globalement croissante, le coureur sera noté (+). Pour l'asymétrie, si un coureur passe plus de temps (ou distance) au dessus de sa vitesse moyenne qu'en dessous (i.e. asymétrie négative), il sera noté (-). S'il passe plus de temps (ou distance) en dessous de sa vitesse moyenne qu'au dessus (i.e. asymétrie positive), il sera noté (+).

La combinaison de la tendance couplée à l'asymétrie forme quatre catégories. Par exemple, la première catégorie (--) regroupe les coureurs dont la vitesse a été décroissante et qui ont passé plus de temps au dessus de leur vitesse moyenne qu'en dessous. La logique est la même pour les trois autres catégories (-+)(+-)(++).

Notre outil a permis de voir que la majorité des coureurs à pied sur marathon courent de manière asymétrique. Asymétrie négative dans la majorité des cas si l'on fait abstraction du genre, de la performance et du parcours. Il semblerait donc que la majorité des coureurs passent plus de temps au dessus de leur vitesse moyenne. D'autre part, il apparaît dans cette étude que les coureurs à pied les plus performant passent plus de temps en dessous de leur vitesse moyenne qu'au dessus (i.e. asymétrie positive) indépendamment de la tendance croissante ou décroissante. Les signatures (-+) et (++) seraient optimales, la tendance de la vitesse, croissante ou décroissante, n'ayant pas d'impact. L'asymétrie positive de la vitesse semble être la clé de l'optimisation de la performance indépendamment de la tendance.

En complément, nous avons étudié le cas particulier du récent recordman du monde sur marathon, Eliud Kipchoge. En appliquant notre outil à ses temps de passage au kilomètre, il en ressort une signature (++).

En parallèle, nous avons voulu rentrer plus en détail dans la course du Kenyan en découpant la course en deux semi-marathons. Chaque semi-marathon reproduit le même motif asymé-

trique, le coureur a passé 62% de la distance à une vitesse inférieure ou égale à sa vitesse moyenne. En découpant la course en quatre, nous avons remarqué que trois 10km sur quatre présentaient une asymétrie positive, le second 10km étant symétrique.

Cette étude nous permet de penser qu'il est possible d'avoir un aperçu de la performance relative au potentiel propre du coureur sur marathon. D'autre part, nous pensons qu'une stratégie de course optimale peut être obtenue en reproduisant le même motif d'asymétrie positive plusieurs fois. Le motif et le nombre de reproductions dépendraient du coureur et seraient individualisés.

Plateforme web d'entraînement

Le deuxième axe de ce travail de thèse portait sur la création d'un outil commercial de diffusion du savoir scientifique de l'entreprise *BillaTraining*[®]. Nous rappelons au lecteur les quatre étapes de la méthode *BillaTraining*[®] : Audit énergétique (Test de *RabiT*[®]), Radar énergétique, plan d'entraînement personnalisé et stratégie de course (signature de vitesse). En débutant le développement de l'application, nous nous sommes vite rendus compte que le traitement du test de *RabiT*[®] présentait un goulot d'étranglement. La solution que nous avons retenue afin de segmenter automatiquement la série temporelle de vitesse de l'audit énergétique est l'algorithme PELT. Cet algorithme, construit sur le paradigme de la programmation dynamique, permet de repérer en une fraction de seconde les instants de ruptures d'une série temporelle. Nous avons ensuite intégré cet algorithme à un micro-service web construit grâce au *framework Django*.

Nous avons construit en parallèle une application web dont le *front-end* s'appuie sur le *framework Angular* et le *back-end* sur le *framework Symfony*. Nous avons implémenté toutes les fonctionnalités permettant d'accéder aux trois premières étapes de la méthode *BillaTraining*[®] en automatisant une application web de bout en bout. La quatrième étape, la stratégie de course, nécessite une validation scientifique avant de proposer ce service aux coureurs amateurs.

Limites et perspectives

Pour l'ultime section de cette thèse, nous allons vous présenter les limites et les perspectives relatives à notre travail. Cette partie comporte deux axes. Le premier présentera les limites et perspectives des résultats propres à la stratégie de course. Le second exposera les limites et les perspectives de la plateforme web.

Stratégie de course sur marathon

L'objectif de la signature de vitesse est de proposer au coureur à pied une stratégie de course sur mesure afin qu'il puisse battre ses records personnels. Un exemple de restitution serait de donner au coureur ses temps de passage cibles au kilomètre avant le départ de la course. Cependant, cette méthode est contradictoire avec la philosophie de l'entreprise *BillaTraining*[®]. En effet, lors de son entraînement, l'athlète amateur est invité à courir à la sensation en portant une attention particulière aux réactions de son organisme. Or, la stratégie de course, dans sa première version, allait contre ce principe et obligeait le coureur à suivre des consignes sans prendre en compte ses sensations du jour. Dans l'idéal, nous souhaiterions laisser au coureur son libre arbitre et intervenir uniquement lorsqu'il ne semble plus s'écouter. C'est pourquoi nous avons imaginé d'itérer sur la première version de l'algorithme de la signature de vitesse. La deuxième version pourrait faire intervenir des changements d'états physiologiques au lieu des changements de vitesses. Par état physiologique nous entendons une matrice de données regroupant par exemple la vitesse, l'accélération, la fréquence cardiaque ou encore la fréquence respiratoire. L'entraînement du coureur nous permettrait de qualifier au moins trois états physiologiques : facile, moyen et dur.

Plateforme web d'entraînement

La plateforme d'entraînement *BillaTraining*[®] est, à ce jour, disponible uniquement en français. Un développement futur que nous avons envisagé serait de traduire l'application en anglais et en espagnol afin de toucher un public plus large qui est contraint par la barrière de

la langue. Ceci devrait s'accompagner d'une gestion des paiements dans différentes devises et d'une stratégie commerciale visant l'international.

Un autre développement que nous souhaitons intégrer est la signature de vitesse. En l'état, l'application intègre les trois premières étapes de la méthode *BillaTraining*[®] puisque nos recherches sur la stratégie de course n'ont pas encore abouti à un résultat pouvant être automatisé.

Une fois l'internationalisation et l'intégration de la signature de vitesse terminées, nous pensons à une possible collaboration avec des chercheurs spécialisés en analyse et traitement du langage. Le but étant de traiter les commentaires des coureurs à pied que nous avons accumulés afin d'améliorer notre service. Il serait alors possible d'intégrer les résultats d'une telle analyse à notre algorithme d'entraînement personnalisé.

Finalement, nous ne souhaitons pas nous arrêter à la course à pied et au marathon. D'autres sports comme le cyclisme, la natation ou encore le ski de fond, ont un fort potentiel d'optimisation. L'expertise scientifique permettrait, pour ces sports, de passer d'un entraînement quantitatif et non personnalisé à un entraînement qualitatif et sur mesure.

Annexe A

Comparison of the three acceleration intensity instructions

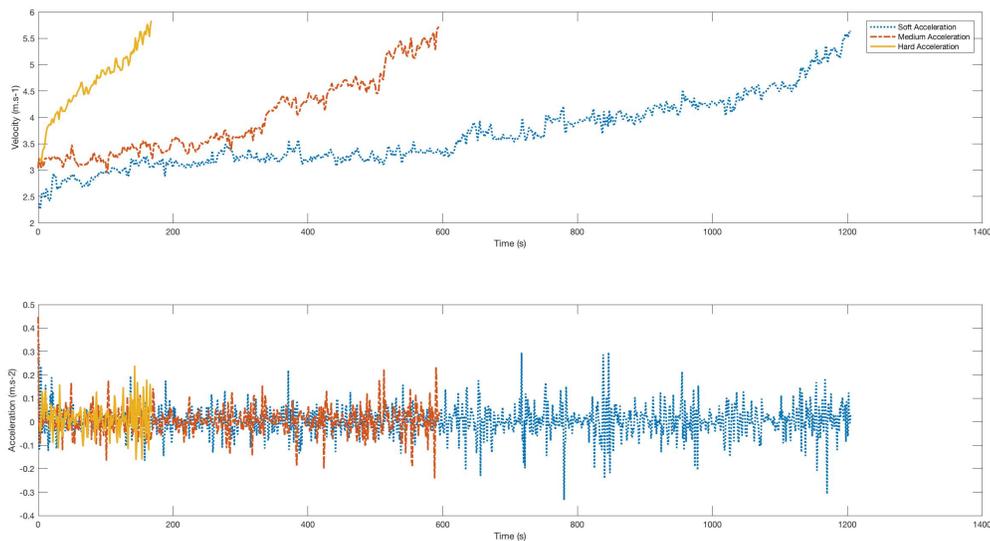


FIGURE A.1 – Male Runner 2 - Top chart : Velocity (m/s) VS. Time (s) ; Bottom chart : Acceleration(m/s²) VS. Time (s) - Comparison of the three acceleration intensity instructions : "soft" (blue curve), "medium" (red curve) and "hard" (yellow curve)

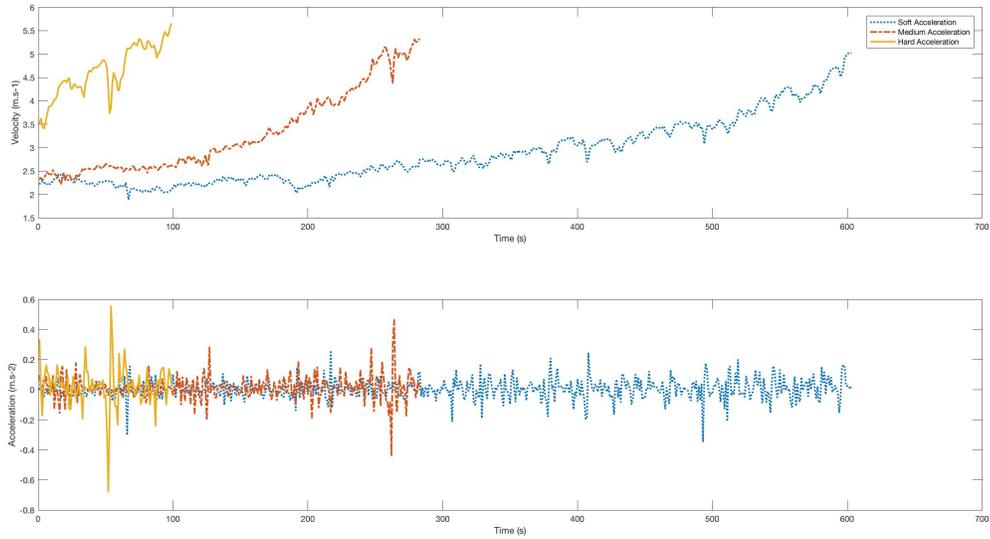


FIGURE A.2 – Male Runner 3 - Top chart : Velocity (m/s) VS. Time (s) ; Bottom chart : Acceleration(m/s²) VS. Time (s) - Comparison of the three acceleration intensity instructions : "soft" (blue curve), "medium" (red curve) and "hard" (yellow curve)

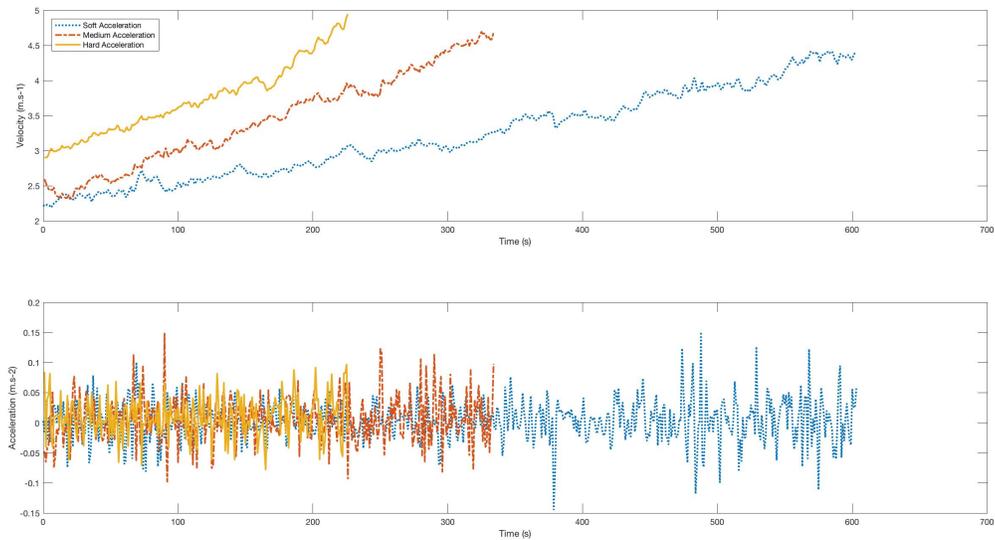


FIGURE A.3 – Female Runner 1 - Top chart : Velocity (m/s) VS. Time (s) ; Bottom chart : Acceleration(m/s²) VS. Time (s) - Comparison of the three acceleration intensity instructions : "soft" (blue curve), "medium" (red curve) and "hard" (yellow curve)

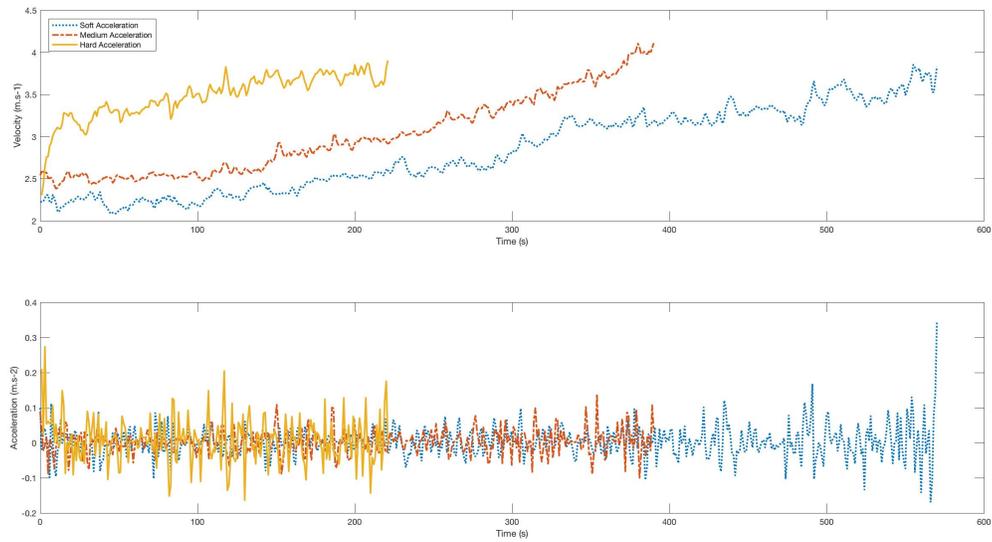


FIGURE A.4 – Female Runner 2 - Top chart : Velocity (m/s) VS. Time (s); Bottom chart : Acceleration(m/s²) VS. Time (s) - Comparison of the three acceleration intensity instructions : "soft" (blue curve), "medium" (red curve) and "hard" (yellow curve)

Annexe B

Measured data VS.

Ornstein-Uhlenbeck model

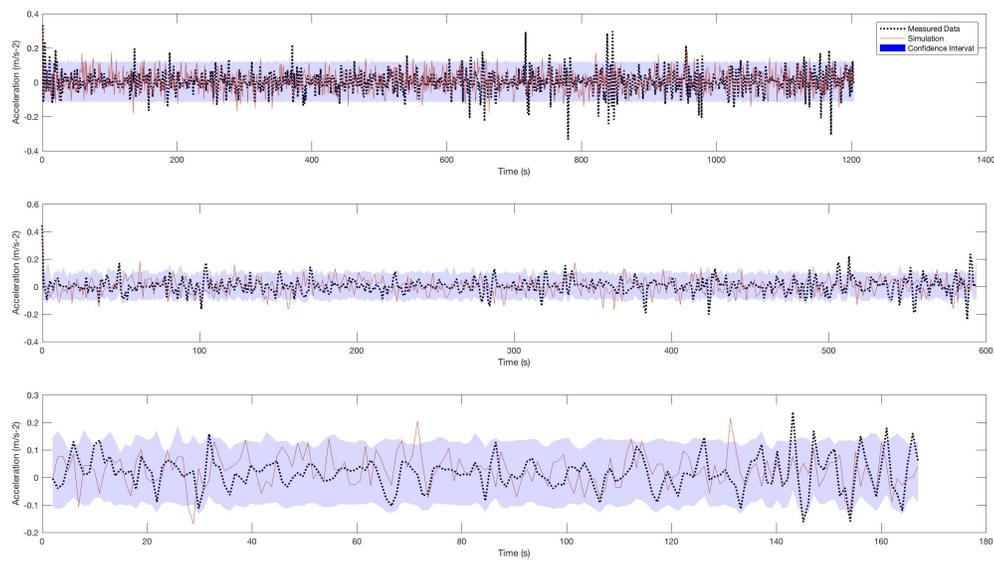


FIGURE B.1 – Male Runner 2 - From top to bottom : Soft, Medium and Hard Acceleration Intensities - Measured data VS. Ornstein-Uhlenbeck model

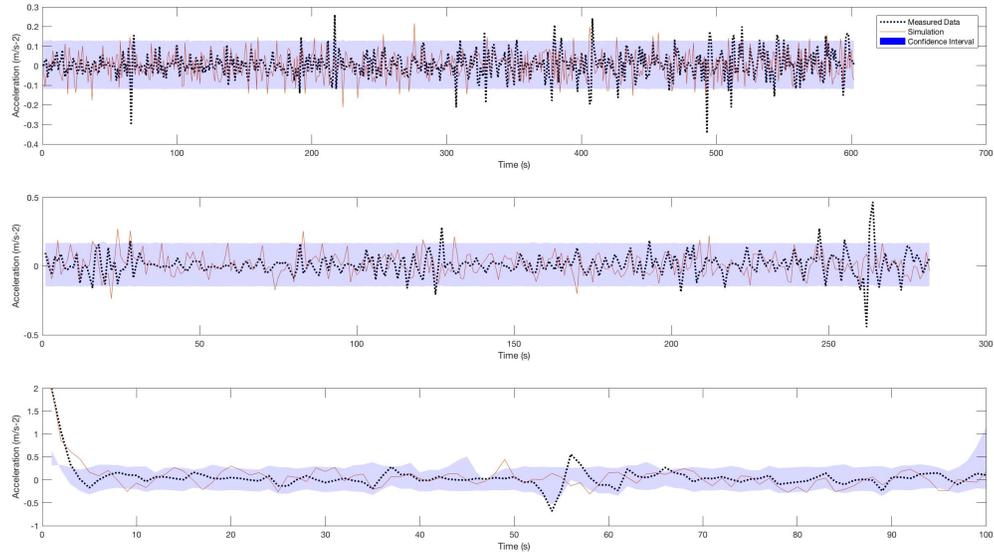


FIGURE B.2 – Male Runner 3 - From top to bottom : Soft, Medium and Hard Acceleration Intensities - Measured data VS. Ornstein-Uhlenbeck model

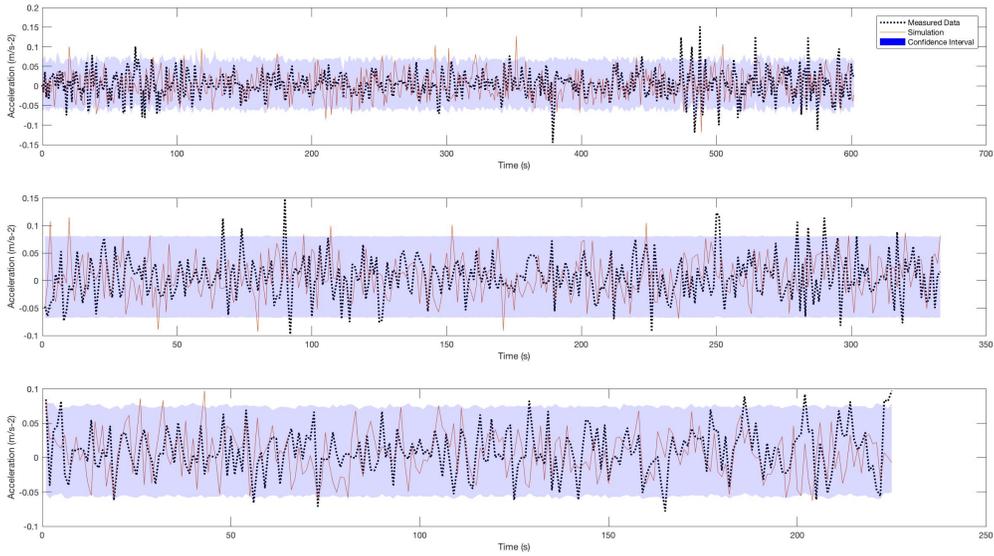


FIGURE B.3 – Female Runner 1 - From top to bottom : Soft, Medium and Hard Acceleration Intensities - Measured data VS. Ornstein-Uhlenbeck model

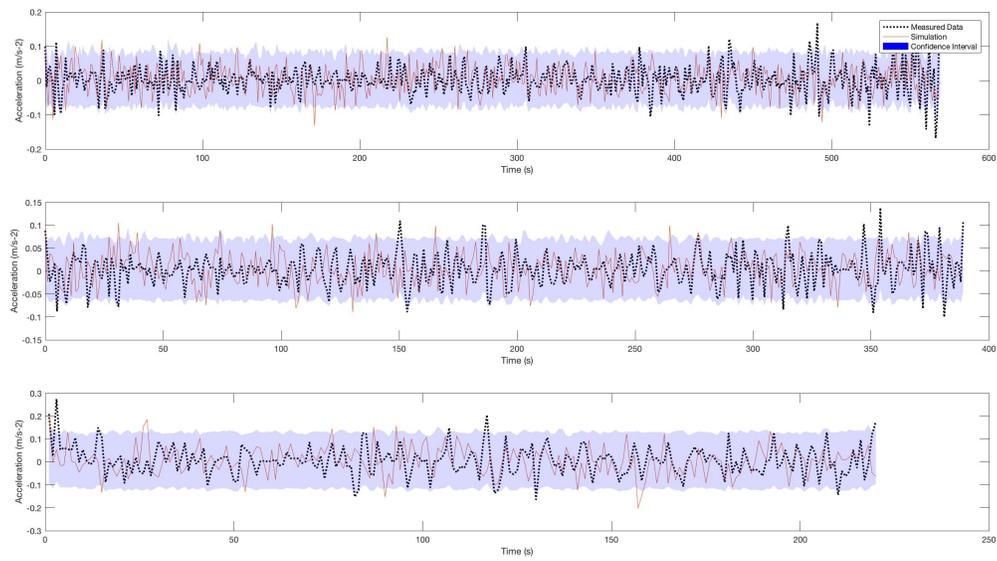


FIGURE B.4 – Female Runner 2 - From top to bottom : Soft, Medium and Hard Acceleration Intensities - Measured data VS. Ornstein-Uhlenbeck model

Annexe C

Comparison of the three estimated mean accelerations

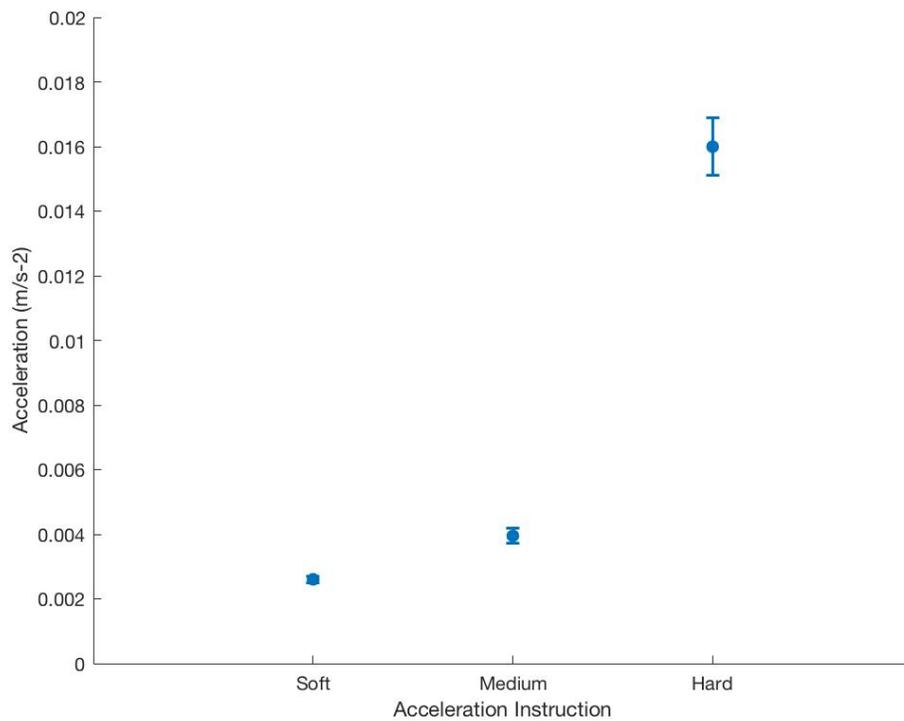


FIGURE C.1 – Male Runner 2 - Comparison of the three estimated mean accelerations

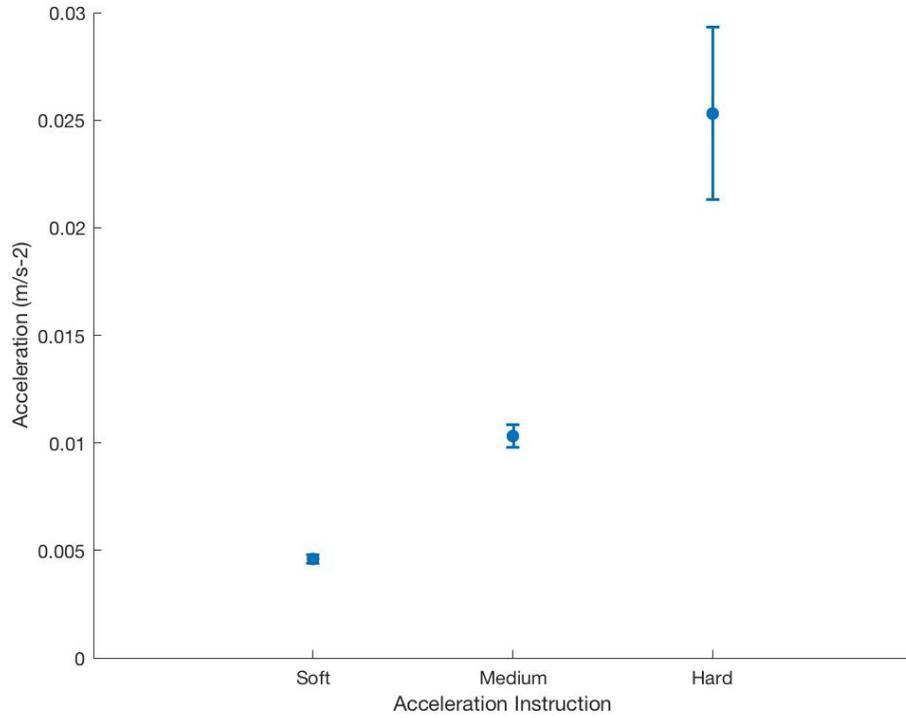


FIGURE C.2 – Male Runner 3 - Comparison of the three estimated mean accelerations

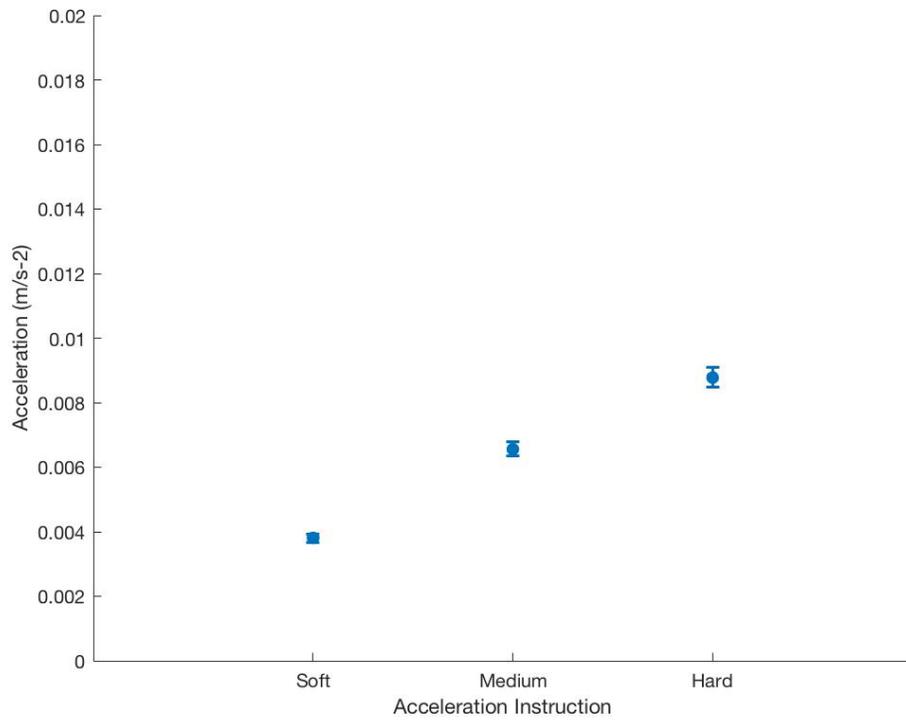


FIGURE C.3 – Female Runner 1 - Comparison of the three estimated mean accelerations

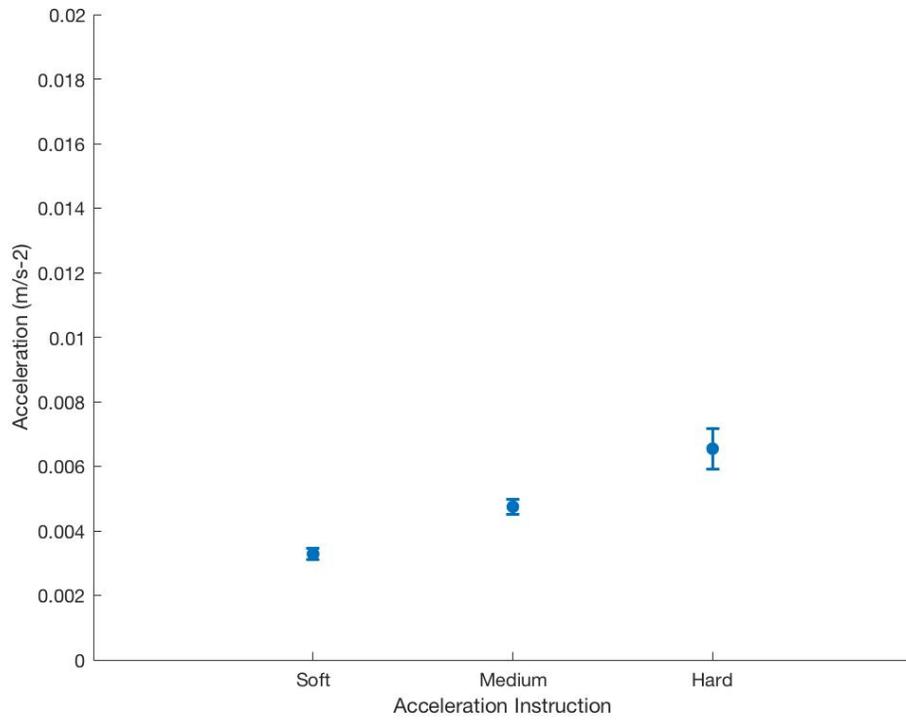


FIGURE C.4 – Female Runner 2 - Comparison of the three estimated mean accelerations

Liste des tableaux

| | | |
|-------|---|----|
| I.1 | Principaux concurrents de l'entreprise <i>BillaTraining</i> [®] : périmètre, offre, avantages et limites | 51 |
| II.1 | Récapitulatif des méthodes de segmentation envisagées pour segmenter le test de RabiT [®] | 68 |
| II.2 | Résultats de la fonction <i>cpt.mean</i> du package R <i>changeoint</i> appliquée au signal de vitesse du test de RabiT [®] . On peut lire le nombre de solutions, le nombre d'instantants de ruptures (IR) min, le nombre d'instantants de ruptures (IR) max et le temps de calcul de l'algorithme en secondes. | 83 |
| II.3 | Tolérances sur les temps pour chaque phase du test de RabiT [®] | 84 |
| III.1 | Number of marathons run divided into two categories: All Marathons (273 Marathons) and Strongly Asymmetric Speed (191 races) each subdivided in seven groups: All, Men, Women, Less than 3 hours marathons, Greater-than and equal to 3 hours marathons, Paris and Berlin. | 91 |
| III.2 | Kendall's τ and p-values for describing the trend in speed time series in a marathon (increasing: ↗, constant: → or decreasing: ↘) among the six groups (Men, Women, < 3 hours, ≥ 3 hours, Paris and Berlin) | 94 |
| III.3 | Repartition of runners between strong and weak asymmetry in speed time series regardless of positive or negative asymmetry. | 95 |
| III.4 | Repartition of runners by whole dataset vs. strong asymmetry and positive vs. negative asymmetry in seed time series. | 96 |

| | |
|--|-----|
| III.5 Subdivided repartition of the runners in the signature for the whole dataset category by the following seven groups: Men, Women, Less than 3 hours, Greater than or equals to 3 hours, Paris, Berlin, All | 97 |
| III.6 Subdivided repartition of the runners in the signature for the strong asymmetry category by the following seven groups: Men, Women, Less than 3 hours, Greater than or equals to 3 hours, Paris, Berlin, All | 97 |
| III.7 Personal all time bests for Eliud Kipchoge. | 106 |
| III.8 Average Speed, Kendall's τ , number of kilometers under and above average speed and signature (see in text) for each half-marathon of the 2018 marathon world record. | 109 |
| III.9 Average Speed, Kendall's τ , number of kilometers under and above average speed and signature for the four segments (10km each) of the 2018 marathon world record. | 110 |
| IV.1 Distance and Time at exhaustion for each runner and for each acceleration trial intensity: "soft", "medium" and "hard" | 125 |
| IV.2 Ratios of distances to exhaustion for "soft" vs. "medium" and "medium" vs. "hard" acceleration intensities for each runner | 125 |
| IV.3 Ratios of times to exhaustion for "soft" vs. "medium" and "medium" vs. "hard" acceleration intensities for each runner | 126 |
| IV.4 Mean acceleration (\hat{a}), autocorrelability ($\hat{\theta}$) and human variability ($\hat{\sigma}$) estimated for each runner and for each acceleration trial intensity: "soft", "medium" and "hard" and confidence interval (C.I) for \hat{a} | 127 |
| IV.5 Ratios of estimated acceleration for "soft" vs. "medium" and "medium" vs. "hard" acceleration intensities for each runner | 128 |
| IV.6 Results of percentages of measured data found in the 95% confidence interval for acceleration for each trial and each runner. | 130 |

| | | |
|------|---|-----|
| IV.7 | <i>p – values</i> of Durbin-Watson statistic to check for autocorrelation among residuals for each runner and for each acceleration trial intensity ("soft", "medium" and "hard") | 131 |
| IV.8 | <i>p – value</i> of Augmented Dickey-Fuller static to check for stationarity for each runner and for each acceleration trial intensity ("soft", "medium" and "hard") | 132 |
| V.1 | Liste exhaustive des paramètres, de leur description, de leur type et exemple de la fonctionnalité <i>inscription</i> de la brique <i>logique métier</i> de l'application <i>RabiT</i> [®] by <i>BillaTraining</i> [®] | 143 |
| V.2 | Liste exhaustive des paramètres, de leur description, de leur type et exemple de la fonctionnalité <i>objectif</i> de la brique <i>logique métier</i> de l'application <i>RabiT</i> [®] by <i>BillaTraining</i> [®] | 144 |
| V.3 | Liste exhaustive des paramètres, de leur description, de leur type et exemple de la fonctionnalité <i>enregistrement activité</i> de la brique <i>logique métier</i> de l'application <i>RabiT</i> [®] by <i>BillaTraining</i> [®] | 144 |
| V.4 | Liste exhaustive des paramètres, de leur description, de leur type et exemple de la fonctionnalité <i>perception et sensations</i> de la brique <i>logique métier</i> de l'application <i>RabiT</i> [®] by <i>BillaTraining</i> [®] | 145 |

Table des figures

| | | |
|------|---|----|
| I.1 | Tracé de la vitesse v en fonction du temps t pour une course supérieure à la distance critique. La phase 1 pour $0 \leq t \leq t_1$ représente l'accélération de 1 à 2 s, la phase 2 pour $t_1 \leq t \leq t_2$ représente la phase de vitesse constante et la phase 3 pour $t_2 \leq t \leq T$ est le lâché prise final. | 24 |
| I.2 | Résultats de différentes stratégies de courses selon Mathis. Les courbes sont issus de l'article de 1989 de Mathis. | 28 |
| I.3 | Figure issue des travaux de Behncke. La figure représente un modèle hydraulique à 3 composantes représentant les 3 sources d'énergie : Créatine phosphate (P), Glycolyse (L) et Oxydation des lipides et glucides (O) | 31 |
| I.4 | D'après Whitt, 2009. La figure représente la distribution des temps finaux par rapport à la valeur de E_* pour un coureur sur 1000m | 33 |
| I.5 | Figure issue des travaux d'Aftalion et Bonnans. La figure représente la vitesse en fonction du temps à l'étape 1 du modèle d'Aftalion et Bonnans | 37 |
| I.6 | Figure issue des travaux d'Aftalion et Bonnans. La figure représente la vitesse en fonction du temps à l'étape 3 du modèle d'Aftalion et Bonnans | 38 |
| I.7 | Figure issue des travaux d'Aftalion et Bonnans. La figure représente la vitesse en fonction du temps à l'étape 4 du modèle d'Aftalion et Bonnans | 39 |
| I.8 | Les 4 étapes de la méthode <i>BillaTraining</i> [®] : Le RabiT [®] , le Radar Energétique, L'entraînement sur mesure, La Signature de Vitesse [®] | 41 |
| I.9 | Organigramme de l'entreprise <i>BillaTraining</i> [®] au mois d'Août 2018 | 42 |
| I.10 | Logo de l'entreprise <i>BillaTraining</i> [®] | 43 |

| | | |
|------|---|----|
| I.11 | Les offres de l'entreprise <i>BillaTraining</i> [®] | 43 |
| I.12 | La méthode <i>BillaTraining</i> [®] - Etape 1 - Le test de RabiT [®] | 45 |
| I.13 | Test de RabiT [®] | 45 |
| I.14 | La méthode <i>BillaTraining</i> [®] - Etape 2 - Le Radar Energétique | 47 |
| I.15 | Exemple de Radar énergétique. Les 4 facteurs de la performance (force musculaire, efficacité cardiaque, tolérance à l'acidose et perception) sont notés sur 10 | 47 |
| I.16 | La méthode <i>BillaTraining</i> [®] - Etape 3 - l'entraînement sur mesure | 48 |
| I.17 | Exemple de séance d'entraînement spécifique <i>BillaTraining</i> [®] en Perception | 48 |
| I.18 | La méthode <i>BillaTraining</i> [®] - Etape 4 - La Signature de Vitesse [®] | 49 |
| I.19 | Exemple de signature de vitesse [®] pour une cliente de <i>BillaTraining</i> [®] sur marathon | 50 |
| I.20 | Photos de l'unité centrale d'un K5. (a) correspond à la photo d'un K5 provenant du site de <i>Cosmed</i> [®] . (b) correspond à la photo d'un K5 porté par un coureur sur son terrain d'entraînement. La photo provient du site de <i>Cosmed</i> [®] | 53 |
| I.21 | Photos de la montre de sport 920XT. (a) correspond à la photo de la montre 920XT provenant du site de <i>Garmin</i> [®] . (b) correspond à la photo de la montre 920XT portée par un coureur sur son terrain d'entraînement. La photo provient du site de <i>Garmin</i> [®] | 54 |
| I.22 | Photo de l'appareil <i>Physioflow</i> [®] PF07 | 55 |
| I.23 | Photos de l'appareil Locometrix. (a) correspond à la photo du Locometrix provenant du site de <i>Centaure Metrix</i> [®] . (b) correspond à la photo du Locometrix portée par une sportive. La photo provient du site de OpenPR | 56 |
| II.1 | Deux tests de RabiT [®] réalisés par deux clients de l'entreprise <i>BillaTraining</i> . Vitesse (km/h) en fonction du temps (s) | 60 |
| II.2 | Les 11 phases du test de RabiT [®] | 72 |

| | | |
|-------|---|-----|
| II.3 | Tests de RabiT [®] en conditions réelles. (a) correspond à un test conforme et validé par l’algorithme. (b) correspond à un test non conforme et rejeté par l’algorithme | 74 |
| II.4 | Schéma d’une application construite sur une architecture monolithique . . . | 75 |
| II.5 | Schéma d’une application construite sur une architecture en micro services . | 76 |
| II.6 | Schéma de l’API de segmentation du test de RabiT [®] . A gauche de la figure se trouvent les entrées (requête). A droite se trouve les sorties (réponse) . . . | 78 |
| II.7 | Schéma de l’architecture de mise en production de l’application de segmentation du test de RabiT [®] | 80 |
| II.8 | Réponse de l’API de segmentation du test de RabiT [®] au format JSON pour un test accepté. | 81 |
| II.9 | Réponse de l’API de segmentation du test de RabiT [®] au format JSON pour un test rejeté. | 81 |
| III.1 | Critical speed and anaerobic capacity of Eliud Kipchoge. Linear regression computed from 1,500 m to 10,000 m. | 107 |
| III.2 | 2018 marathon world record. The bar chart represents the speed time series. The line chart stands for the average speed on the segment. Speed values are computed for each kilometer from pace values. | 108 |
| III.3 | 2018 marathon world record. On the left-hand side is the first half of the marathon. On the right-hand side is the second half. Bar charts represent the speed time series. Line charts stand for the average speed on the segment. . | 110 |
| III.4 | 2018 marathon world record. Top left is the first 10k of the marathon. Top right is the second 10k. Bottom left is the third 10k. Bottom right is the fourth 10k. Bar charts represent the speed time series. Line charts stand for the average speed on the segment. | 111 |

| | |
|--|-----|
| IV.1 Male Runner 1 - Top chart: Velocity (m/s) VS. Time (s); Bottom chart: Acceleration(m/s ²) VS. Time (s) - Comparison of the three acceleration intensity instructions: "soft" (blue curve), "medium" (red curve) and "hard" (yellow curve) | 124 |
| IV.2 Male Runner 1 - Comparison of the three estimated mean accelerations . . . | 129 |
| IV.3 Male Runner 1 - From top to bottom: Soft, Medium and Hard Acceleration Intensities - Measured data VS. Ornstein-Uhlenbeck model | 130 |
| V.1 Schéma global de l'application web <i>RabiT</i> [®] by <i>BillaTraining</i> [®] | 140 |
| V.2 Schéma des fonctionnalités de la brique <i>logique métier</i> de l'application <i>RabiT</i> [®] by <i>BillaTraining</i> [®] | 142 |
| V.3 Schéma des fonctionnalités de la brique <i>web UI/UX</i> de l'application <i>RabiT</i> [®] by <i>BillaTraining</i> [®] | 146 |
| V.4 Vue <i>inscription</i> de l'application <i>RabiT</i> [®] by <i>BillaTraining</i> [®] | 147 |
| V.5 Vue <i>connexion</i> de l'application <i>RabiT</i> [®] by <i>BillaTraining</i> [®] | 148 |
| V.6 Vue <i>mot de passe oublié</i> de l'application <i>RabiT</i> [®] by <i>BillaTraining</i> [®] | 148 |
| V.7 Première vue <i>accueil</i> de l'application <i>RabiT</i> [®] by <i>BillaTraining</i> [®] | 150 |
| V.8 Seconde vue <i>accueil</i> de l'application <i>RabiT</i> [®] by <i>BillaTraining</i> [®] | 150 |
| V.9 Vue <i>objectif</i> de l'application <i>RabiT</i> [®] by <i>BillaTraining</i> [®] | 151 |
| V.10 Vue <i>mon compte</i> de l'application <i>RabiT</i> [®] by <i>BillaTraining</i> [®] | 152 |
| V.11 Première vue <i>démarrage rapide</i> de l'application <i>RabiT</i> [®] by <i>BillaTraining</i> [®] . . . | 152 |
| V.12 Seconde vue <i>démarrage rapide</i> de l'application <i>RabiT</i> [®] by <i>BillaTraining</i> [®] . . . | 153 |
| V.13 Troisième vue <i>démarrage rapide</i> de l'application <i>RabiT</i> [®] by <i>BillaTraining</i> [®] . . . | 153 |
| V.14 Vue <i>ajout activité</i> de l'application <i>RabiT</i> [®] by <i>BillaTraining</i> [®] | 154 |
| V.15 Vue <i>liste activités</i> de l'application <i>RabiT</i> [®] by <i>BillaTraining</i> [®] | 155 |
| V.16 Première vue <i>activité</i> de l'application <i>RabiT</i> [®] by <i>BillaTraining</i> [®] | 156 |
| V.17 Deuxième vue <i>listeactivité</i> de l'application <i>RabiT</i> [®] by <i>BillaTraining</i> [®] | 157 |
| V.18 Vue <i>premium</i> de l'application <i>RabiT</i> [®] by <i>BillaTraining</i> [®] | 157 |
| V.19 Vue <i>support technique</i> de l'application <i>RabiT</i> [®] by <i>BillaTraining</i> [®] | 158 |

| | |
|--|-----|
| V.20 Schéma des fonctionnalités de la brique <i>analyses scientifiques</i> de l'application <i>RabiT</i> [®] by <i>BillaTraining</i> [®] | 158 |
| V.21 Schéma de fonctionnement du paiement en ligne de l'application <i>RabiT</i> [®] by <i>BillaTraining</i> [®] | 159 |
| V.22 Schéma globale de l'application mobile <i>RabiT</i> [®] by <i>BillaTraining</i> [®] | 162 |
| V.23 Les trois couches du modèle <i>TCP/IP</i> | 163 |
| V.24 Schéma des fonctionnalités de la brique <i>mobile UI/UX</i> de l'application mobile <i>RabiT</i> [®] by <i>BillaTraining</i> [®] | 165 |
| V.25 Vue <i>connexion</i> de l'application mobile. | 166 |
| V.26 Première vue <i>entraînement</i> de l'application mobile. | 167 |
| V.27 Deuxième vue <i>entraînement</i> de l'application mobile. | 168 |
| V.28 Première vue <i>activité</i> de l'application mobile. | 169 |
| V.29 Deuxième vue <i>activité</i> de l'application mobile. | 170 |
| V.30 Troisième vue <i>activité</i> de l'application mobile. | 171 |
| V.31 Vue <i>liste activités</i> de l'application mobile. | 172 |
| V.32 Vue <i>paramètres</i> de l'application mobile. | 173 |
| V.33 Première vue de la brique <i>web UI/UX 2</i> . On y distingue 4 coureurs à pied connectés en temps réel. | 174 |
| V.34 Deuxième vue de la brique <i>web UI/UX 2</i> | 175 |
| A.1 Male Runner 2 - Top chart : Velocity (m/s) VS. Time (s); Bottom chart : Acceleration(m/s ²) VS. Time (s) - Comparison of the three acceleration intensity instructions : "soft" (blue curve), "medium" (red curve) and "hard" (yellow curve) | 182 |
| A.2 Male Runner 3 - Top chart : Velocity (m/s) VS. Time (s); Bottom chart : Acceleration(m/s ²) VS. Time (s) - Comparison of the three acceleration intensity instructions : "soft" (blue curve), "medium" (red curve) and "hard" (yellow curve) | 183 |

A.3 Female Runner 1 - Top chart : Velocity (m/s) VS. Time (s) ; Bottom chart : Acceleration(m/s²) VS. Time (s) - Comparison of the three acceleration intensity instructions : "soft" (blue curve), "medium" (red curve) and "hard" (yellow curve) 183

A.4 Female Runner 2 - Top chart : Velocity (m/s) VS. Time (s) ; Bottom chart : Acceleration(m/s²) VS. Time (s) - Comparison of the three acceleration intensity instructions : "soft" (blue curve), "medium" (red curve) and "hard" (yellow curve) 184

B.1 Male Runner 2 - From top to bottom : Soft, Medium and Hard Acceleration Intensities - Measured data VS. Ornstein-Uhlenbeck model 185

B.2 Male Runner 3 - From top to bottom : Soft, Medium and Hard Acceleration Intensities - Measured data VS. Ornstein-Uhlenbeck model 186

B.3 Female Runner 1 - From top to bottom : Soft, Medium and Hard Acceleration Intensities - Measured data VS. Ornstein-Uhlenbeck model 186

B.4 Female Runner 2 - From top to bottom : Soft, Medium and Hard Acceleration Intensities - Measured data VS. Ornstein-Uhlenbeck model 187

C.1 Male Runner 2 - Comparison of the three estimated mean accelerations . . . 188

C.2 Male Runner 3 - Comparison of the three estimated mean accelerations . . . 189

C.3 Female Runner 1 - Comparison of the three estimated mean accelerations . . 189

C.4 Female Runner 2 - Comparison of the three estimated mean accelerations . . 190

Liste des Algorithmes

- 1 Optimal Partitioning 67
- 2 Pruned Exact Linear Time 68

Bibliographie

- [1] C. R. Abbiss and P. B. Laursen. Describing and understanding pacing strategies during athletic competition. *Sports Med*, 38 :239–252, 2008.
- [2] A. Aftalion and J. F. Bonnans. Optimization of running strategies based on anaerobic energy and variations of velocity. *SIAM J. Appl. Math*, 74 :1615–1636, 2014.
- [3] A. Agresti. *Analysis of Ordinal Categorical Data*. John Wiley and Sons, 2nd edition, 2010.
- [4] H. Akaike. A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19 :716–723, 1974.
- [5] J. Alvarez-Ramirez and E. Rodriguez. Scaling properties of marathon races. *Physica A*, 365 :509–520, 2006.
- [6] S. D. Angus. Did recent world record marathon runners employ optimal pacing strategies? *J Sports Sci*, 32 :31–45, 2014.
- [7] L. Ansley, P. J. Robson, A. St-Clair-Gibson, and T. D. Noakes. Anticipatory pacing strategies during supra-maximal exercise lasting more than 30 s. *Med Sci Sports Exerc*, 36 :309–314, 2004.
- [8] M. Ariyoshi, K. Yamaji, and R. Shephard. Influence of running pace upon performance : Effects upon treadmill endurance time and oxygen cost. *Eur J Appl Physiol Occup Physiol*, 41 :83–91, 1979.

- [9] T. A. Astorino, R. P. Allen, D. W. Roberson, and M. Jurancich. Effect of high-intensity interval training on cardiovascular function, vo2max, and muscular force. *J Strength Cond Res*, 26 :138–145, 2012.
- [10] I. E. Auger and C. E. Lawrence. Algorithms for the optimal identification of segment neighborhoods. *Bull Math Biol*, 51 :39–54, 1989.
- [11] B. Auvinet, G. Berrut, C. Touzard, L. Moutel, N. Collet, D. Chaleil, and E. Barrey. Reference data for normal subjects obtained with an accelerometric device. *Gait posture*, 16 :124–134, 2002.
- [12] C. Baldari, M. Meucci, F. Bolletta, M. Gallotta, G. Emerenziani, and L. Guidetti. Accuracy and reliability of cosmed k5 portable metabolic device versus simulating system. In *Sport Sci Health*, 2015.
- [13] B. Balke and R. W. Ware. An experimental study of physical fitness of air force personnel. *U S Armed Forces Med J*, 10 :675–688, 1959.
- [14] J. C. Barbero-Alvarez, A. Coutts, J. Granda, V. Barbero-Alvarez, and C. Castagna. The validity and reliability of a global positioning satellite system device to assess speed and repeated sprint ability (rsa) in athletes. *J Sci Med Sport*, 13 :232–235, 2009.
- [15] H. Behncke. A mathematical model for the force and energetics in competitive running. *J Math Biol*, 31 :853–878, 1993.
- [16] A. Berthoz. *Neurobiology of "Umwelt"*. Springer, 2009.
- [17] V. Billat. *Physiologie et méthodologie de l'entraînement : de la théorie à la pratique*. Editions De Boeck, 3ème édition edition, 2012.
- [18] V. Billat, R. H. Morton, N. Blondel, S. Berthoin, V. Bocquet, J. P. Koralsztein, and T. Barstow. Oxygen kinetics and modelling of time to exhaustion whilst running at various velocities at maximal oxygen uptake. *J Appl Physiol*, 82 :178–187, 2000.

- [19] V. Billat, P. Lepretre, A. Heugas, and J. Koralsztein. Energetics of middle-distance running performances in male and female junior using track measurements. *Jpn J Physiol*, 54 :125–135, 2004.
- [20] V. Billat, E. Wesfreid, C. Kapfer, J. P. Koralsztein, and Y. Meyer. Nonlinear dynamics of heart rate and oxygen uptake in exhaustive 10,000 m runs : influence of constant vs. freely paced. *J Physiol Sci*, 56 :103–111, 2006.
- [21] V. Billat, L. Hamard, J. P. Koralsztein, and R. H. Morton. Differential modeling of anaerobic and aerobic metabolism in the 800-m and 1.500-m run. *J Appl Physiol*, 107 : 478–487, 2009.
- [22] V. Billat, L. Mille-Hamard, Y. Meyer, and E. Wesfreid. Detection of changes in the fractal scaling of heart rate and speed in a marathon race. *Physica A*, 388 :3798–3808, 2009.
- [23] V. Billat, N. J.-B. Brunel, T. Carbillet, S. Labbé, and A. Samson. Humans are able to self-paced constant running accelerations until exhaustion. *Physica A*, 506 :290–304, 2018.
- [24] V. Billat, T. Carbillet, M. Correa, and J. R. Pycke. Detecting the marathon asymmetry with a statistical signature. *Physica A*, 515 :240–247, 2019.
- [25] V. L. Billat. *VO2 max à l'épreuve du temps*. de boeck, 2013.
- [26] V. L. Billat. *Entraînement pratique et scientifique à la course à pied*. de boeck, 2015.
- [27] V. L. Billat and J. P. Koralsztein. Significance of the velocity at vo2max and time to exhaustion at this velocity. *Sports Med*, 22 :90–108, 1996.
- [28] V. L. Billat, J. C. Renoux, J. Pinoteau, B. Petit, and J. P. Koralsztein. Times to exhaustion at 90, 100 and 105(maximal aerobic speed) and critical speed in elite long distance runners. *Arch Physiol Biochem*, 103 :129–135, 1995.

- [29] V. L. Billat, V. Binsse, B. Petit, and J. P. Koralsztein. High level runners are able to maintain a vo2 steady-state below vo2max in an all-out run over their critical velocity. *Arch Physiol Biochem*, 106 :38–45, 1998.
- [30] V. L. Billat, A. Demarle, J. Slawinski, M. Paiva, and J. P. Koralsztein. Physical and training characteristics of top-class marathon runners. *Med Sci Sports Exerc*, 33 : 2089–2097, 2001.
- [31] V. L. Billat, J. Slawinski, M. Danel, and J. P. Koralsztein. Effect of free versus constant pace on performance and oxygen kinetics in running. *Med Sci Sports Exerc*, 33 :2082–2088, 2001.
- [32] V. L. Billat, A. Demarle, M. Paiva, and J. P. Koralsztein. Effect of training on the physiological factors of performance in elite marathon runners (males and females). *Int J Sports Med*, 23 :336–341, 2002.
- [33] V. L. Billat, G. Dhonneur, L. Mille-Hamard, L. L. Moyec, I. Momken, T. Launay, J. P. Koralsztein, and S. Besse. Case studies in physiology : Maximal oxygen consumption and performance in a centenarian cyclist. *J Appl Physiol*, 122 :430–434, 2017.
- [34] V. L. Billat, T. Carbillet, and J. R. Pycke. How to run the marathon in less than 2 hours : Old and modern approach building a bridge between physiology and mathematics. In *Running and Science meets Physics and Engineering, MIT workshop, Boston*, April 13th 2018.
- [35] L. Birge and P. Massart. Minimal penalties for gaussian model selection. *Probab Theory Rel*, 138 :33–73, 2007.
- [36] D. Bishop, D. Bonetti, and B. Dawson. The influence of pacing strategy on vo2 and supramaximal kayak performance. *Med Sci Sports Exerc*, 34 :1041–1047, 2002.
- [37] V. Bocquet. *Modélisation de l’endurance humaine, de l’espèce à l’individu*. PhD thesis, Université René Descartes Paris V, 2002.

- [38] A. M. P. Bovens, G. M. E. Janssen, H. G. W. Vermeer, J. H. Hoeberigs, M. P. E. Janssen, and F. T. J. Verstappen. Occurrence of running injuries in adults following a supervised training program. *Int J Sports Med*, 10 :186–190, 1989.
- [39] D. M. Bramble and D. E. Lieberman. Endurance running and the evolution of homo. *Nature*, 432 :345–352, 2004.
- [40] L. Buley. *The User Experience Team of One : A Research and Design Survival Guide*. Rosenfeld Media, 2013.
- [41] K. E. Callen. Mental and emotional aspects of long-distance running. *Psychosomatics*, 24 :133–134, 1983.
- [42] V. Carbone and S. Savaglio. Scaling laws and forecasting in athletic world records. *J Sports Sci*, 19 :477–484, 2001.
- [43] J. Cassirame, R. Vanhaesebrouck, S. Chevrolat, and L. Mourot. Accuracy of the garmin 920 xt hrm to perform hrv analysis. *Australas Phys Eng Sci Med*, 40 :831–839, 2017.
- [44] J. Chen and A. K. Gupta. *Parametric Statistical Change Point Analysis*. Birkhauser, 2012.
- [45] S. N. Cheuvront and E. M. Haymes. Thermoregulation and marathon running. *Sports Med*, 31 :743–762, 2001.
- [46] J. N. Chorley, J. C. Cianca, J. G. Divine, and T. D. Hew. Baseline injury risk factors for runners starting a marathon training program. *Clin J Sport Med*, 12 :18–23, 2002.
- [47] A. Cooper, R. Reimann, D. Cronin, and C. Noessel. *About Face : The Essentials of Interaction Design*. Wiley, 2014.
- [48] F. Cottin, Y. Papelier, F. Durbin, J. P. Koralsztein, and V. L. Billat. Effect of fatigue on spontaneous velocity variations in human middle-distance running : Use of short-term fourier transformation. *Eur J Appl Physiol*, 87 :17–27, 2002.

- [49] E. F. Coyle. Physiological regulation of marathon performance. *Sports Med*, 37 : 306–311, 2007.
- [50] A. B. Craig. Evaluation and predictions of world running and swimming records. *J Sports Med Phys Fitness*, 3 :14–21, 1963.
- [51] J. de Koning, C. Foster, A. Bakkum, S. Kloppenburg, C. Thiel, T. Joseph, J. Cohen, and J. P. Porcari. Regulation of pacing strategy during athletic competition. *PLoS One*, 6 :e15863, 2011.
- [52] J. Dekerle, X. Nesi, and H. Carter. The distance - time relationship over a century of running olympic performances : A limit on the critical speed concept. *J Sports Sci*, 24 : 1213–1221, 2006.
- [53] J. J. Diaz, E. J. Fernández-Ozcorta, and J. Santos-Concejero. The influence of pacing strategy on marathon world records. *Eur J Sport Sci*, 18 :781–786, 2018.
- [54] A. L. Dontchev and T. Zolezzi. *Well-Posed Optimization Problems*. Springer, 1993.
- [55] J. Duckett. *HTML and CSS : Design and Build Websites*. John Wiley and Sons, 2011.
- [56] J. Duckett. *JavaScript and JQuery : Interactive Front-End Web Development*. Wiley, 2014.
- [57] M. Ely, D. Martin, S. Cheuvront, and S. Montain. Effect of ambient temperature on marathon pacing is dependent on runner ability. *Med Sci Sports Exerc*, 40 :1675–1680, 2008.
- [58] W. S. Erdmann and P. Lipinska. Kinematics of marathon running tactics. *Hum Mov Sci*, 32 :1379–1392, 2013.
- [59] J. H. Ettema. Limits of human performance and energy production. *Arbeitsphysiologie*, 22 :45–54, 1966.
- [60] R. T. Fielding. *Architectural Styles and the Design of Network-based Software Architectures*. PhD thesis, University of California, Irvine, 2000.

- [61] T. Flash and N. Hogan. The coordination of arm movements : an experimentally confirmed mathematical model. *J Neurosci*, 5 :1688–1703, 1985.
- [62] C. Foster, A. Snyder, N. Thompson, M. Green, M. Foley, and M. Schrage. Effect of pacing strategy on cycle time trial performance. *Med Sci Sports Exerc*, 25 :383–388, 1993.
- [63] C. Foster, M. Schrage, A. Snyder, and N. Thompson. Pacing strategy and athletic performance. *Sports Med*, 17 :77–85, 1994.
- [64] V. Froelicher and J. Myers. *Exercise and the Heart*. Elsevier, 2000.
- [65] M. Garcin, M. Daniel, and V. L. Billat. Perceptual responses in free vs. constant pace exercise. *Int J Sports Med*, 29 :453–459, 2008.
- [66] J. J. Garrett. *The Elements of User Experience : User-Centered Design for the Web and Beyond*. New Riders, 2010.
- [67] R. Gibbons, G. Balady, J. Beasley, J. Bricker, W. Duvernoy, V. Froelicher, D. Mark, T. Marwick, B. McCallister, P. J. Thompson, W. Winters, F. Yanowitz, J. Ritchie, R. Gibbons, M. Cheitlin, K. Eagle, T. Gardner, A. J. Garson, R. Lewis, R. O’Rourke, and T. Ryan. Guidelines for exercise testing : a report of the american college of cardiology/ american heart association task force on practice guidelines. *J Am Coll Cardiol*, 30 :260–311, 1997.
- [68] S. Gillain, E. Warzee, F. Lekeu, V. Wojtasik, D. Maquet, J.-L. Croisier, E. Salmon, and J. Petermans. The value of instrumental gait analysis in elderly healthy, mci or alzheimer’s disease subjects and a comparison with other clinical tests used in single and dual-task conditions. *Ann Phys Rehabil Med*, 52 :453–474, 2009.
- [69] D. T. Gillespie. Exact numerical simulation of the ornstein-uhlenbeck process and its integral. *Phys. Rev. E*, 54 :2084–2091, 1996.

- [70] M. Gleeson, N. C. Bishop, D. J. Stensel, M. R. Lindley, S. S. Mastana, and M. A. Nimmo. The anti-inflammatory effects of exercise : Mechanisms and implications for the prevention and treatment of disease. *Nat Rev Immunol*, 11 :607–615, 2011.
- [71] V. G. T. Gomez, J. A. G. Venegas, and A. E. C. Franco. Likelihood ratio test for a change-point in the mean of independent normal observations. *Proceedings of the 2013 Industrial and Systems Engineering Research Conference*, 2013.
- [72] N. Gordon, C. R. Abbiss, A. J. Maiorana, K. J. Marston, and J. J. Peiffer. Intrarater reliability and agreement of the physioflow bioimpedance cardiography device during rest, moderate and high-intensity exercise. *Kinesiology*, 50 :140–149, 2018.
- [73] J. H. Greist, M. H. Klein, R. R. Eischens, J. F. an Alan S Gurman, and W. P. Morgan. Running as treatment for depression. *Compr Psychiatry*, 20 :41–54, 1979.
- [74] H. Grosse-Lodermann and E. A. Muller. Der einfluss der leistung und der arbeitsgeschwindigkeit auf das arbeitsmaximum und den wirkungsgrad beim radfahren. *Arbeitsphysiologie*, 9 :454–475, 2005.
- [75] X. Guyon and J. f Yao. On the underfitting and overfitting sets of models chosen by order selection criteria. *J Multivar Anal*, 70 :221–249, 1999.
- [76] J. Hadamard. Sur les problèmes aux dérivées partielles et leur signification physique. *Princeton University Bulletin*, 13 :49–52, 1902.
- [77] L. M. Hamard, C. Breuneval, A. S. Rousseau, P. Grimaldi, and V. L. Billat. Transcriptional modulation of mitochondria biogenesis pathway at and above critical speed in mice. *Mol Cell Biochem*, 405 :223–232, 2015.
- [78] T. Haney and J. Mercer. A description of variability of pacing in marathon distance running. *Int J Exerc Sci*, 4 :133–140, 2011.
- [79] C. Hanon, J. M. Leveque, C. Thomas, and L. Vivier. Pacing strategy and vo₂ kinetics during a 1500m race. *Int J Sports Med*, 29 :206–211, 2008.

- [80] E. A. . Harman, H. G. Knuttgen, P. N. Frykman, and J. F. Patton. Exercise endurance time as function of percent maximal power production. *Med Sci Sports Exerc*, 19 : 480–485, 1987.
- [81] D. V. Harris. *Comparative effectiveness of running therapy and psychotherapy*. Hemisphere, w. p. morgan and s. e. goldston edition, 1987.
- [82] C. Hausswirth and D. Lehenaff. Physiological demands of running during long distance runs and triathlons. *Sports Med*, 31 :679–689, 2001.
- [83] K. Haynes, I. A. Eckley, and P. Fearnhead. Efficient penalty search for multiple changepoint problems. *arXiv*, 2014.
- [84] K. Haynes, I. A. Eckley, and P. Fearnhead. Computationally efficient changepoint detection for a range of penalties. *J Comput Graph Stat*, 26 :134–143, 2015.
- [85] F. M. Henry. A note on physiological limits and the history of the mile run. *Res Q*, 25 :483–484, 1954.
- [86] F. M. Henry. Time-velocity equations and oxygen requirements of "all-out" and "steady-pace" running. *Res Q*, 25 :164–177, 1954.
- [87] F. M. Henry. Prediction of world records in running sixty yards to twenty six miles. *Res Q*, 26 :147–158, 1955.
- [88] A. V. Hill. *Muscular movement in man*. McGraw-Hil, 1927.
- [89] J. S. Hogg, J. G. Hopker, and A. R. Mauger. The self-paced vo2max test to assess maximal oxygen uptake in highly trained runners. *Int J Sports Physiol Perform*, 10 : 172–177, 2017.
- [90] L. Horvath. The maximum likelihood method of testing changes in the parameters of normal observations. *Ann Stat*, 21 :671–680, 1993.
- [91] A. Hreljac. Stride smoothness evaluation of runners and other athletes. *Gait Posture*, 11 :199–206, 2000.

- [92] R. L. Hughson, C. J. Orok, and L. E. Staudt. A high velocity treadmill running test to assess endurance running potential. *Int J Sports Med*, 5 :23–25, 1984.
- [93] J. P. Hunter, R. N. Marshall, and P. J. McNair. Relationships between ground reaction force impulse and kinematics of sprint-running acceleration. *J Appl Biomech*, 21 :31–43, 2005.
- [94] S. Iotti, M. Borsari, and D. Bendahan. Oscillations in energy metabolism. *Biochim Biophys Acta*, 1797 :1353–1361, 2010.
- [95] A. E. Iyoho, L. N. MacFadden, and L. J. Ng. Prediction of performance reduction in self-paced exercise as modulated by the rating of perceived exertion. *Eur J Appl Physiol*, 115 :675–690, 2015.
- [96] B. Jackson, J. Scargle, D. Barnes, S. Arabhi, A. Alt, P. Gioumousis, E. Gwin, P. Sangtrakulcharoen, L. Tan, and T. T. Tsai. An algorithm for optimal partitioning of data on an interval. *IEEE Signal Process Lett*, 12 :105–108, 2005.
- [97] R. A. Jacobs, D. Fluck, T. C. Bonne, S. Burgi, P. M. Christensen, M. Toigo, and C. Lundby. Improvements in exercise performance with high-intensity interval training coincide with an increase in skeletal muscle mitochondrial content and function. *J Appl Physiol*, 115 :785–793, 2013.
- [98] M. J. Joyner, J. R. Ruizand, and A. Lucia. The two-hour marathon : who and when? *J Appl Physiol*, 110 :275–277, 2011.
- [99] U. S. KantarSport. Communiqué de presse sur la pratique du running en france, Mars 2014.
- [100] J. B. Keller. A theory of competitive running. *Phys Today*, 26 :43, 1973.
- [101] J. B. Keller. Optimal velocity in a race. *Am Math Mon*, 81 :474–480, 1974.
- [102] M. Kendall and A. Stuart. *The Advance Theory of Statistics*. Griffin and Co, 1963.

- [103] M. G. Kendall. A new measure of rank correlation. *Biometrika*, 30 :81–93, 1938.
- [104] A. E. Kennelly. An approximate law of fatigue in the speeds of racing animals. *Proc Am Acad Arts Sci*, 42 :275–331, 1906.
- [105] A. E. Kennelly. Changes during the last twenty years in the world’s speed records on racing animals. *Proc Am Acad Arts Sci*, 61 :487–523, 1926.
- [106] R. Killick. Introduction to optimal changepoint detection algorithms.
- [107] R. Killick and I. A. Eckley. changepoint : An r package for changepoint analysis. *J Stat Softw*, 58, 2014.
- [108] R. Killick, P. Fearnhead, and I. A. Eckley. Optimal detection of changepoints with a linear computational cost. *J Am Stat Assoc*, 107 :1590–1598, 2012.
- [109] G. L. Kraft and R. A. Roberts. Validation of the garmin forerunner 920xt fitness watch vo2peak test. *Int J Inno Edu Res*, 5, 2017.
- [110] M. I. Lambert, J. P. Dugas, M. C. Kirkman, G. G. Mokone, and M. R. Waldeck. Changes in running speeds in a 100 km ultra-marathon race. *J Sports Sci Med*, 3 : 167–173, 2004.
- [111] P. Larsson and K. Henriksson-Larsen. The use of dgps and simultaneous metabolic measurements during orienteering. *Med Sci Sports Exerc*, 33 :1919–1924, 2001.
- [112] D. C. Lee, R. R. Pate, C. J. Lavie, X. Sui, T. S. Church, and S. N. Blair. Leisure-time running reduces all-cause and cardiovascular mortality risk. *J Am Coll Cardiol*, 64 : 472–481, 2014.
- [113] L. A. Léger and R. J. Ferguson. Effect of pacing on oxygen uptake and peak lactate for a mile run. *Eur J Appl Physiol Occup Physiol*, 32 :251–257, 1974.
- [114] C. Leiden and M. Wilensky. *TCP / IP For Dummies*. John Wiley and Sons, 6th edition edition, 2009.

- [115] P. M. Lepretre, J. P. Koralsztein, and V. L. Billat. Effect of exercise intensity on relationship between vo2max and cardiac output. *Med Sci Sports Exerc*, 36 :1357–1363, 2004.
- [116] Z. Lin and F. Meng. Empirical analysis on the runners’ velocity distribution in city marathons. *Physica A*, 490 :533–541, 2018.
- [117] D. J. Macfarlane. Open-circuit respirometry : a historical review of portable gas analysis systems. *Eur J Appl Physiol*, 117 :2369–2386, 2017.
- [118] R. Maidstone, T. Hocking, G. Rigai, and P. Fearnhead. On optimal multiple change-point algorithms for large data. *Stat Comput*, 27 :519–533, 2017.
- [119] R. Margaria. *Biomechanics and energetics of muscular exercises*. Oxford University Press, 1976.
- [120] R. Margaria, P. Aghemo, and F. Pinera-Linas. A simple relation between performance in running and maximal aerobic power. *J Appl Physiol*, 38 :351–352, 1975.
- [121] M. B. Maron, S. M. Horvath, J. E. Wilerson, and J. A. Gliner. Oxygen uptake measurements during competitive marathon running. *J Appl Physiol*, 40 :836–838, 1976.
- [122] M. B. Maron, J. A. Wagner, and S. M. Horvath. Thermoregulatory responses during competitive marathon running. *J Appl Physiol Respir Environ Exerc Physiol*, 42 : 909–914, 1977.
- [123] R. Maronski and K. Rogowski. Minimum-time running : A numerical approach. *Acta Bioeng Biomech*, 13 :83–86, 2011.
- [124] F. Mathis. The effect of fatigue on running strategies. *SIAM Review*, 31 :306–309, 1989.
- [125] G. P. Meade. An analytic study of athletic records. *Sci Mon*, 2 :509–600, 1916.
- [126] G. P. Meade. Consistent running records. *Science*, 124 :1025, 1956.

- [127] R. H. Morton. *A mathematical and computer simulation model of the running athlete*. PhD thesis, Massey University, 1984.
- [128] R. H. Morton. A three component model of human bioenergetics. *J Math Biol*, 24 : 451–466, 1986.
- [129] R. H. Morton. A 3-parameter critical power model. *Ergonomics*, 39 :611–619, 1996.
- [130] R. H. Morton and L. V. Billat. The critical power model for intermittent exercise. *Eur J Appl Physiol*, 91 :303–307, 2004.
- [131] M. Mossio, M. Vidal, and A. Berthoz. Traveled distances : new insights into the role of optic flow. *Vision Res*, 48 :289–303, 2008.
- [132] J. Myers and D. Bellin. Ramp exercise protocols for clinical and cardiopulmonary exercise testing. *Sports Med*, 30 :23–29, 2000.
- [133] J. Myers, N. Buchanan, D. Smith, J. Neutel, E. Bowes, D. Walsh, and V. F. Froelicher. Individualized ramp treadmill. observations on a new protocol. *Chest*, 101 :236S–241S, 1992.
- [134] G. A. Nader. Concurrent strength and endurance training : from molecules to man. *Med Sci Sports Exerc*, 38 :1965–1970, 2006.
- [135] T. D. Noakes, A. S. C. Gibson, and E. V. Lambert. From catastrophe to complexity : a novel model of integrative central neural regulation of effort and fatigue during exercise in humans. *Br J Sports Med*, 38 :511–514, 2004.
- [136] L. Norman and G. Lachapelle. Odometry performance of gnss portable and carrier phase-based receivers. *International Conference on Indoor Positioning and Indoor Navigation*, 2015.
- [137] B. S. Overstreet, D. R. Bassett, S. E. Crouter, B. C. Rider, and B. B. Parr. Portable open-circuit spirometry systems. *J Sports Med Phys Fitness*, 57 :227–237, 2017.

- [138] C. C. Peterson, K. A. Nagy, and J. Diamond. Sustained metabolic scope. *Proc Natl Acad Sci U.S.A.*, 87 :2324–2328, 1990.
- [139] F. Picard, S. R. M. Lavielle, C. Vaisse, and J. J. Daudin. A statistical approach for array cgh data analysis. *BMC Bioinformatics*, 2005.
- [140] A. B. Pitcher. Optimal strategies for a two-runner model of middle-distance running. *SIAM J Appl Math*, 70 :1032–1046, 2009.
- [141] J. Porszasz, R. Casaburi, A. Somfay, L. J. Woodhouse, and B. J. Whipp. A treadmill ramp protocol using simultaneous changes in speed and grade. *Med Sci Sports Exerc*, 35 :1596–1603, 2003.
- [142] M. Porter. *L’avantage concurrentiel*. DUNOD, 2003.
- [143] P. E. D. Prampero, G. Atchou, J. C. Bruckner, and C. Moi. The energetic of endurance running. *Eur J Appl Physiol Occup Physiol*, 55 :259–266, 1986.
- [144] W. G. Pritchard. Mathematical models of running. *Society for Industrial and Applied Mathematics*, 35 :359–379, 1993.
- [145] J.-R. Pycke, T. Carbillet, and V. L. Billat. The 2018 marathon world record shows the positive asymmetry pacing strategy effectiveness. Sent to Journal of Applied Physiology, 2018.
- [146] L. M. Ricciardi and L. Sacerdote. The ornstein-uhlenbeck process as a model for neuronal activity. *Biol Cybern*, 35 :1–9, 1979.
- [147] Running-USA. 2016 running usa annual marathon report.
- [148] L. E. Sandals, D. M. Wood, S. B. Draper, and D. V. James. Influence of pacing strategy on oxygen uptake during treadmill middle-distance running. *Int J Sports Med*, 27 :37–42, 2006.

- [149] G. N. Sandford, A. E. Kilding, A. Ross, and P. B. Laursen. Maximal sprint speed and the anaerobic speed reserve domain : the untapped tools that differentiate the world's best male 800 m runners. *Sports Med*, 2018.
- [150] A. Santos-Lozano, P. S. Collado, C. Foster, A. Lucia, and N. Garatachea. Influence of sex and level on marathon pacing strategy. insights from the new york city race. *Int J Sports Med*, 35 :933–938, 2014.
- [151] J. Scheerder, K. Breedveld, and J. Borgers. *Running across Europe : The Rise and Size of One of the Largest Sport Markets*. Palgrave Macmillan UK, 2015.
- [152] J. Scherrer, M. Samson, and A. Paleologue. Etude du travail musculaire de la fatigue. *J Appl Physiol*, 46 :887–916, 1954.
- [153] R. Schobel and J. Zhu. Stochastic volatility with an ornstein-uhlenbeck process : An extension. *Review of Finance*, 3 :23–46, 1999.
- [154] G. Schwarz. Estimating the dimension of a model. *Ann Stat*, 6 :461–464, 1978.
- [155] A. J. Scott and M. Knott. A cluster analysis method for grouping means in the analysis of variance. *Biometrics*, 30 :507–512, 1974.
- [156] A. Sen and M. S. Srivastava. On tests for detecting change in mean. *Ann Stat*, 3 : 98–108, 1975.
- [157] V. Shteinbakh and V. Gerlitsyn. *Soviet sport : the success story*. Raduga, 1987.
- [158] R. G. Shulman and D. L. Rothman. The "glycogen shunt" in exercising muscle : A role for glycogen in muscle energetics and fatigue. *Proc Natl Acad Sci U S A*, 98 :457–461, 2001.
- [159] B. L. M. Smits, G.-J. Pepping, and F. J. Hettinga. Pacing and decision making in sport and exercise : The roles of perception and action in the regulation of exercise intensity. *Sports Med*, 44 :763–775, 2014.

- [160] R. Suwannarat, K. Srihirun, J. Brito, and D. Suksom. Acute effects of normoxic and hypoxic high-intensity interval exercise on hemodynamics and cutaneous blood flow in youth soccer players. *J Exerc Physiol Online*, 20 :92–98, 2017.
- [161] G. Tornvall. *Assessment of physical capabilities*. Acta Physiologica Scandinavica, 1963.
- [162] M. A. Townsend. Road-racing strategies. *Med Sci Sports Exerc*, 14 :253–243, 1982.
- [163] R. Tucker and T. Noakes. The physiological regulation of pacing strategy during exercise : a critical review. *Br J Sports Med*, 43, 2009.
- [164] R. Tucker and J. Santos-Concejero. The unlikeliness of an imminent sub-2-hour marathon : Historical trends of the gender gap in running events. *Int J Sports Physiol Perform*, 12 :1017–1022, 2017.
- [165] R. Tucker, A. Bester, E. V. Lambert, T. D. Noakes, C. L. Vaughan, and A. S. C. Gibson. Non random fluctuations in power output during self paced exercise. *Br J Sports Med*, 40 :912–917, 2006.
- [166] R. Tucker, M. I. Lambert, and T. D. Noakes. An analysis of pacing strategies during men’s world-record performances in track athletics. *Int J Sports Physiol Perform*, 1 : 233–245, 2006.
- [167] G. Uhlenbeck and L. Ornstein. On the theory of the brownian motion. *Phys Rev*, 36 : 823, 1930.
- [168] R. Unger. *A Project Guide to UX Design : For user experience designers in the field or in the making*. New Riders, 2012.
- [169] D. Van-Biesen, F. Hettinga, K. McCulloch, and Y. C. Vanlandewijck. Pacing profiles in competitive track races : Regulation of exercise intensity is related to cognitive ability. *Front Physiol*, 7 :624, 2016.
- [170] A. Vanhatalo, A. M. Jones, and M. Burnley. Application of critical power in sport. *Int J Sports Physiol Perform*, 6 :128–136, 2011.

- [171] W. S. Vincent. *REST APIs with Django : Build powerful web APIs with Python and Django*. Independently published, 2018.
- [172] P. G. Weyand and J. A. Davis. Running performance has a structural basis. *J Exp Biol*, 208 :2625–2631, 2005.
- [173] D. Whitt. Optimal strategies for stochastic models of running performance. May 2009.
- [174] P. M. Will and J. D. Walter. Exercise testing : improving performance with a ramped bruce protocol. *Am Heart J*, 138 :1033–1037, 1999.
- [175] P. T. Williams. Reduced total and cause-specific mortality from walking and running in diabetes. *Med Sci Sports Exerc*, (46) :933–939, 2014.
- [176] K. R. Wilund. Is the anti-inflammatory effect of regular exercise responsible for reduced cardiovascular disease? *Clin Sci (Lond)*, 112 :543–555, 2007.
- [177] W. Woodside. The optimal strategy for running a race (a mathematical model for world records from 50 m to 275 km). *Math Comput Model*, 15 :1–12, 1991.
- [178] F. Zaninotto and F. Potencier. *The Definitive Guide to symfony*. Apress, 2007.
- [179] E. Ziemann, T. Grzywacz, M. Luszczuk, R. Laskowski, R. A. Olek, and A. L. Gibson. Aerobic and anaerobic changes with high-intensity interval training in active college-aged men. *J Strength Cond Res*, 25 :1104–1112, 2011.