



# **Estimation de l'état fonctionnel de l'opérateur**

**Mémoire**

**Olivier Gagnon**

**Maîtrise en génie électrique**  
Maître ès sciences (M.Sc.)

Québec, Canada

© Olivier Gagnon, 2017

# **Estimation de l'état fonctionnel de l'opérateur**

**Mémoire**

**Olivier Gagnon**

Sous la direction de :

Marc Parizeau, directeur de recherche

# Résumé

L'estimation de l'état fonctionnel de l'opérateur (c'est-à-dire le patron multidimensionnel de conditions physiologiques et comportementales qui régule les performances) a un grand potentiel pour l'augmentation de la sécurité et de la fiabilité de systèmes critiques. L'apprentissage automatique, qui a connu des avancées importantes au cours des dernières années, est une avenue à explorer pour effectuer cette estimation.

Une problématique dans l'utilisation de ces techniques est la formalisation de l'OFS en une mesure objective permettant de fournir un signal d'entraînement à l'apprentissage automatique. Ce mémoire présente une mesure, la performance dynamique décontextualisée, permettant d'utiliser ces techniques pour estimer l'état fonctionnel de plusieurs participants, pour plusieurs tâches expérimentales différentes.

Cet ouvrage explore également les performances obtenues par plusieurs techniques d'apprentissage automatique dans divers contextes. Entre autres, la généralisation des modèles entraînés à de nouveaux participants ou de nouvelles tâches expérimentales et l'utilisation du contexte expérimental sont étudiées.

# Abstract

The assessment of an operator's functional state (i.e., the multidimensional pattern of human psycho-physiological conditions that mediates performance) has great potential for increasing safety and reliability of critical systems. Machine learning, which has had success in recent years, is a technique which should be investigated for this task.

An open question in the use of machine learning algorithms for the assessment of the operator's functional state is the formalization of the operator's state in an objective measure that can provide a training signal for the algorithms. This Master's thesis introduces the decontextualized dynamic performance, a measure which enables the use of machine learning for many experimental tasks and many participants simultaneously.

This work also explores the performances obtained by machine learning techniques in some contexts. The generalization of the trained models to new participants, or new tasks as well as the utilization of the training context is investigated.

# Table des matières

Résumé	iii
Abstract	iv
Table des matières	v
Liste des tableaux	vii
Liste des figures	viii
Acronymes	x
Remerciements	xii
Introduction	1
<b>1 L'état fonctionnel de l'opérateur</b>	<b>4</b>
1.1 Opérateur . . . . .	6
1.2 Environnement . . . . .	9
1.3 Tâche . . . . .	9
1.4 Estimation . . . . .	11
1.5 Objectifs de recherche . . . . .	17
<b>2 Apprentissage automatique</b>	<b>19</b>
2.1 Définition . . . . .	19
2.2 Algorithmes évalués . . . . .	22
2.3 Évaluation des algorithmes . . . . .	25
2.4 Entraînement . . . . .	29
<b>3 Méthode</b>	<b>34</b>
3.1 Signaux physiologiques et comportementaux . . . . .	34
3.2 Collecte de données . . . . .	38
3.3 Définition de la valeur cible . . . . .	44
3.4 Apprentissage automatique . . . . .	57
<b>4 Résultats</b>	<b>62</b>
4.1 Collecte de données . . . . .	62

4.2 Performances . . . . .	69
4.3 Apprentissage automatique . . . . .	74
<b>Conclusion</b>	<b>82</b>
Formalisation de l'OFS . . . . .	82
Évaluation de l'apprentissage automatique . . . . .	83
Généralisation . . . . .	83
Utilisation du contexte . . . . .	84
Conclusion et travaux futurs . . . . .	84
<b>A Code</b>	<b>86</b>
<b>B Données</b>	<b>87</b>
<b>C Résultats supplémentaires</b>	<b>89</b>
<b>Bibliographie</b>	<b>97</b>

# Liste des tableaux

3.1	Signaux du BioHarness . . . . .	36
3.2	Signaux du dispositif de suivi oculaire . . . . .	36
3.3	Valeurs calculées par RHRV . . . . .	37
3.4	Signaux calculés . . . . .	38
3.5	Hyperparamètres utilisés pour la recherche aléatoire . . . . .	60
4.1	Proportion des étiquettes d'un DDP fonctionnel pour chaque tâche et chaque participant. . . . .	73
4.2	Proportion des étiquettes de l'état fonctionnel dans le jeu d'entraînement et de test. . . . .	73
4.3	Proportion des étiquettes de l'état fonctionnel en fonction de chaque condition expérimentale. . . . .	74
4.4	Résultats des expériences d'apprentissage automatique . . . . .	76
B.1	Valeurs utilisés pour les expériences d'apprentissage automatique . . . . .	88
C.1	Résultats complet des expériences d'apprentissage automatique pour tous les participants, tâches, classifieurs et modélisations indépendant. . . . .	96

# Liste des figures

1.1	Modèle de l'état fonctionnel de l'opérateur. . . . .	5
1.2	Capteurs permettant d'estimer l'OFS. . . . .	11
1.3	Diagramme de la performance en fonction de la demande . . . . .	16
2.1	Frontières de décision du SVM . . . . .	23
2.2	Arbre de décision pour la classification du jeu de données iris. . . . .	24
2.3	Frontières de décisions de quatre classifieurs . . . . .	26
2.4	Matrice de confusion . . . . .	28
2.5	Recherche d'hyper paramètres . . . . .	31
2.6	Validation croisée à k-plis . . . . .	32
3.1	BioHarness 3.0 de Zephyr . . . . .	35
3.2	Le EyeTribe . . . . .	36
3.3	Représentation de la tâche de n-back . . . . .	40
3.4	Exemple de la tâche de rotation mentale . . . . .	41
3.5	Exemple de la tâche de rotation mentale . . . . .	41
3.6	Exemples de la tâche de recherche visuelle . . . . .	43
3.7	Temps de réaction d'une session expérimentale . . . . .	46
3.8	Filtrage à l'aide d'un filtre moyenneur à fenêtre glissante . . . . .	48
3.9	Filtrage à l'aide d'un filtre exponentiel à fenêtre glissante . . . . .	49
3.10	Filtrage à l'aide d'un filtre médian à fenêtre glissante . . . . .	50
3.11	Comparaison des types de filtres . . . . .	51
3.12	Comparaison de plusieurs techniques de seuillage . . . . .	55
3.13	Matrice de corrélation entre les combinaisons de paramètres du DDP . . . . .	56
3.14	Exemple de la DDP calculé sur une session. . . . .	57
4.1	Distribution de l'âge des participants . . . . .	63
4.2	Valeurs de différents signaux physiologiques et comportementaux en fonction du temps. . . . .	64
4.3	Distributions de quatre signaux en fonction des participants. . . . .	66
4.4	Distribution de quatre signaux en fonction de la tâche expérimentale. . . . .	67
4.5	Signaux biologiques en fonction des conditions expérimentales . . . . .	68
4.6	Distribution des signaux biologiques en fonction du DDP . . . . .	69
4.7	Distribution des signaux biologiques en fonction du DDP . . . . .	70
4.8	Proportion des réponses correctes . . . . .	71



4.9	Distribution des temps de réaction dans les tâches expérimentales . . . .	72
4.10	MCC calculé sur le jeu de test. . . . .	75
4.11	Distribution des MCC calculés pour chaque tâche et chaque participant.	77
4.12	MCC de la classification par les forêts aléatoires en fonction des participants.	78
4.13	MCC d'une classification en fonction des tâches expérimentales . . . . .	79
4.14	Distribution des MCC avec ou sans contexte. . . . .	81

# Acronymes

- CART** arbre de classification et de régression (classification and regression tree)
- DDP** performance dynamique décontextualisée (decontextualized dynamic performance)
- ECG** électrocardiogramme
- EEG** électroencéphalogramme
- HF** hautes fréquences (high frequencies)
- HRV** variabilité du rythme cardiaque (heart rate variability)
- IID** indépendant et identiquement distribués
- KNN** K-voisins les plus proches (K-nearest neighbors)
- LF** basses fréquences (low frequencies)
- LVSN** laboratoire de vision et systèmes numériques
- MCC** coefficients de corrélation de Matthews (Matthew's correlation coefficients)
- MRI** imagerie par résonance magnétique (magnetic resonance imaging)
- NaN** nombre invalide (Not a Number)
- OFS** état fonctionnel de l'opérateur (operator functional state)
- PSD** densité spectrale de puissance (power spectral density)
- SA** conscience de la situation (situational awareness)
- SVM** séparateur à vastes marges (support vector machines)
- ULF** fréquences ultras basses (ultra low frequencies)
- VLF** fréquences très basses (very low frequencies)

Nous ne sommes savants que de  
la science présente.

---

Michel Equyem de Montaigne

# Remerciements

L'écriture de ce mémoire n'aurait pas été possible sans le support de mes proches.

Plus spécialement, j'aimerais remercier mes parents pour leur support matériel et pour leurs encouragements dans la poursuite de mes études. J'ai été choyé de pouvoir compter sur mes parents et de sentir toute leur fierté lorsque je parlais de mes projets en cours.

J'aimerais aussi remercier mon grand-père Gaétan et mon père de m'avoir donné le goût aux sciences. C'est grâce à vous que j'ai entrepris mes études en génie.

Au cours de mon cheminement scolaire, j'ai eu des professeurs qui m'ont marqué et qui m'ont permis de développer une véritable culture et une éducation. Si je complète ce mémoire, c'est aussi parce que j'ai appris à réfléchir sur la vie en dehors des études. J'aimerais donc remercier mon ancien professeur de philosophie Alexandre Provencher-Gravel.

Un mémoire ne pourrait être ce qu'il est sans le support d'un directeur de recherche. J'aimerais donc remercier mon directeur de recherche Marc Parizeau. C'est lui qui m'a enseigné à être critique et à toujours être prêt à remettre en question les hypothèses suivies, qu'elles soient de moi ou d'un spécialiste.

Cette maîtrise a été réalisée avec le support d'une bourse BMP. Dans le cadre de celle-ci, j'ai donc eu la chance de recevoir un financement de Thalès Canada. J'aimerais également remercier mon superviseur en milieu pratique, Daniel Lafond ainsi que Jean-François Gagnon. Ceux-ci ont été d'une grande aide pendant mes travaux.

Finalement, il est impératif que je remercie Flore Morneau-Sévigny pour ses encouragements et son support. Les longues heures d'études ont été beaucoup plus légères à porter en sachant que tu m'accompagnais avec tes propres travaux.

# Introduction

Au cours des dernières années, la charge de travail demandée aux opérateurs, c'est-à-dire les personnes contrôlant divers procédés, est de plus en plus importante [20]. Cette surcharge de travail a poussé le développement de systèmes d'automatisation. Cependant, malgré les automatisations disponibles, les humains demeurent toujours essentiels dans plusieurs tâches. Il est donc important de considérer les opérateurs en tant qu'élément du processus lors du design de ces systèmes.

Pour comprendre les interactions entre les travailleurs et leurs tâches, il est nécessaire de modéliser à la fois les opérateurs, les tâches, l'environnement et leurs interactions. Pour certains auteurs, cette modélisation est donnée par le modèle de l'état fonctionnel de l'opérateur (operator functional state, OFS) [90]. L'analyse de l'état fonctionnel de l'opérateur permet donc de répondre aux phases de la taxinomie *sense-assess-augment*. La taxinomie *sense-assess-augment* est utilisée par l'armée de l'air américaine dans le développement de procédés d'augmentation des opérateurs [33]. Celle-ci est analogue au *sense-plan-act* utilisée en robotique [5]. L'estimation de l'OFS permet d'effectuer les étapes *sense* et *assess*, l'activation des méthodes d'automatisation permet d'effectuer l'étape *augment*.

L'estimation de l'OFS a plusieurs utilités. Elle permet de détecter l'augmentation du risque d'erreurs humaines, de contrôler les processus d'automatisation adaptatifs et d'évaluer l'efficacité des interfaces entre les opérateurs et les procédés contrôlés par ceux-ci.

D'autres avancées technologiques récentes rendent également l'estimation de l'OFS plus simple. Entre autres, les avancées récentes dans les capteurs physiologiques portatifs permettent d'obtenir des données physiologiques beaucoup plus facilement qu'il y a quelques années. Dans le domaine de l'apprentissage automatique, les avancées des dernières années permettent d'envisager des performances qui étaient difficiles à atteindre jusqu'à récemment.

Pour ces raisons, l'estimation de l'état fonctionnel de l'opérateur est un sujet qui doit être étudié plus en profondeur.

## Les capteurs portatifs

Les capteurs physiologiques portatifs ont connu une popularité sans précédent au cours des dernières années. Auparavant, ceux-ci étaient généralement dispendieux et leur utilisation, compliquée. Pour ces raisons, les études employant ces capteurs devaient se faire dans un contexte de laboratoire. De plus, le coût élevé de ces équipements limitait le type et le nombre de capteurs accessible. Récemment, les coûts des capteurs physiologiques ont grandement diminué. De plus, la facilité d'utilisation de ceux-ci les a rendus utilisables hors du laboratoire.

La disponibilité de ces capteurs est notée par plusieurs auteurs [64, 78]. D'autres auteurs ont également démontré des liens entre les signaux physiologiques et comportementaux et l'OFS [85, 87, 90]. Pour cette raison, l'utilisation des capteurs portables pour l'estimation de l'état fonctionnel de l'opérateur est un sujet de recherche qui mérite d'être approfondi.

## Apprentissage automatique

Les performances des algorithmes d'apprentissage automatique ont grandement augmenté au cours des dernières années. En effet, les avancées récentes en apprentissage automatique ont permis leur application à des tâches jugées jusqu'à présent comme très difficiles pour ces algorithmes. Parmi les récents succès de l'apprentissage automatique, on retrouve des avancées en robotique [51], en diagnostic médical [28] et en vision artificielle [48]. Ces avancées récentes posent la question quant à savoir si ces techniques permettent d'estimer l'OFS de façon acceptable.

## Estimation de l'état fonctionnel de l'opérateur

Ce document présente des recherches effectuées pour évaluer la possibilité et l'efficacité de l'usage de l'apprentissage automatique et des signaux physiologiques et comportementaux pour l'estimation de l'état fonctionnel de l'opérateur. Les recherches effectuées comprennent entre autres la création de tâches informatisées pour évaluer les opérateurs, une collecte de données avec des participants et l'essai de plusieurs techniques d'apprentissage automatique.

# Structure du document

Ce document est divisé en quatre chapitres.

Les deux premiers chapitres du document présentent le contexte théorique dans lequel le projet évolue. Le premier chapitre aborde le concept de l'OFS. C'est à ce moment que la définition de celui-ci est présentée. De plus, les objectifs de recherche de ce projet y sont également discutés. Le chapitre suivant traite de l'apprentissage automatique. Les principes de bases et les techniques utilisés dans le reste de cet ouvrage y sont décrits.

Les deux derniers chapitres sont consacrés à l'application de l'apprentissage automatique pour faire l'estimation de l'OFS. La méthode utilisée ainsi que les détails d'implémentation peuvent être trouvés au troisième chapitre. Finalement, cet ouvrage présente les résultats des expériences d'apprentissage automatique dans le dernier chapitre.

# Chapitre 1

## L'état fonctionnel de l'opérateur

L'objet de ce projet est l'estimation de l'état fonctionnel de l'opérateur. Cependant, bien que le sujet ait brièvement été abordé dans l'introduction, la définition formelle et détaillée n'a pas encore été faite. Dans [40], l'OFS est défini comme :

« The variable capacity of the operator for effective task performance in response to task and environmental demand, and under the constraints imposed by cognitive and physiological processes that control and energise behaviour. »

Dans [87], l'OFS est défini de la manière suivante :

« OFS is defined as the multidimensional pattern of human psychophysiological condition that mediates performance in relation to physiological and psychological costs. OFS results from the synthesis of operator characteristics, current operator condition, and the operator's interaction with operational requirements. »

Dans ces deux citations, on peut distinguer trois éléments qui sont nécessaires à la modélisation de l'OFS : l'opérateur, la tâche et le contexte. Le concept de l'OFS nécessite donc de comprendre les interactions entre ces trois éléments. La figure 1.1 présente le modèle de l'OFS qui est utilisé dans ce projet. Ce modèle compte les trois composantes principales de l'OFS, l'opérateur, la tâche et l'environnement ou le contexte. Dans ce modèle, un opérateur, aussi appelé participant, est en train d'effectuer une tâche, c'est-à-dire qu'il fait des actions pour accomplir certains objectifs. Celui-ci a un état physiologique et cognitif de base et l'environnement dans lequel se trouve l'opérateur a une influence sur cet état. Les actions qu'il pose permettent d'obtenir une certaine performance dans la tâche en cours. Cette tâche a également un effet sur l'état du



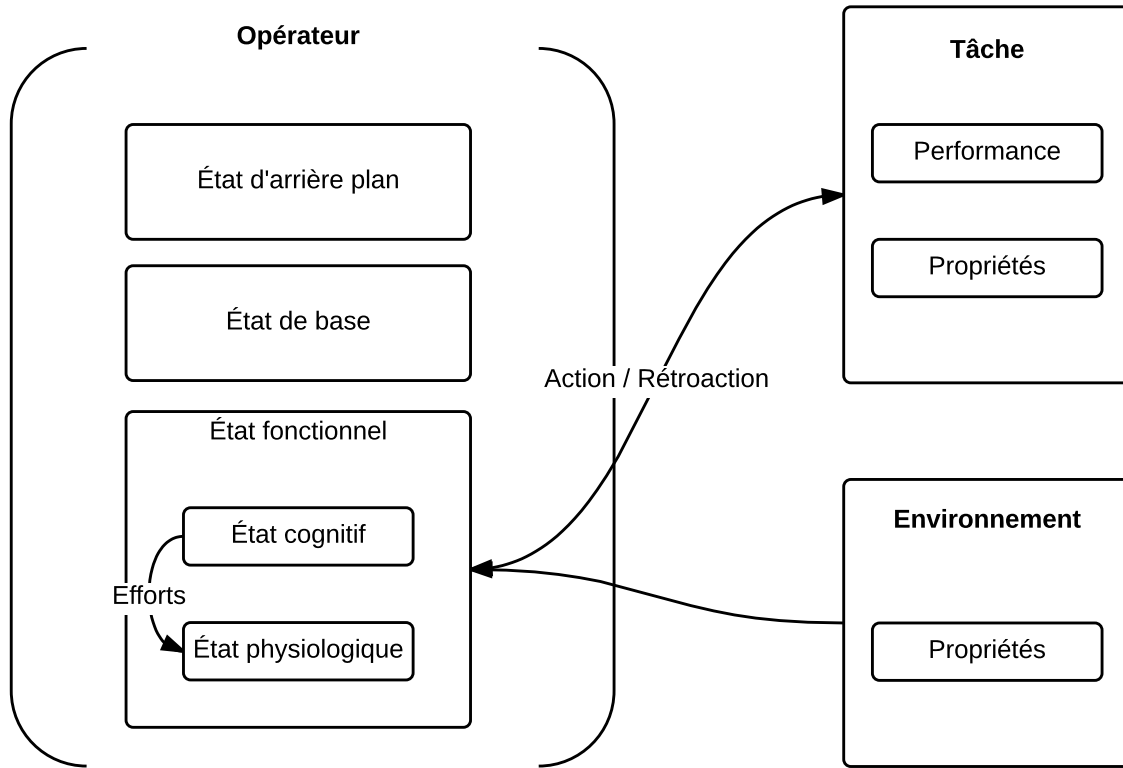


FIGURE 1.1 – Modèle de l'état fonctionnel de l'opérateur.

participant. Les sections 1.1, 1.2 et 1.3 présentent les composantes du modèle de l'OFS ainsi que leurs interactions.

Le concept de l'OFS rend possible la modélisation de la performance d'une personne dans le cas où celle-ci effectue une tâche dans un contexte particulier. Par exemple, on peut utiliser l'OFS pour mesurer l'état d'un pilote d'avion en train de faire un atterrissage la nuit. Dans cet exemple, l'opérateur est le pilote d'avion, la tâche effectuée est l'atterrissage et le contexte est donné par le fait que cette action est effectuée pendant la nuit.

L'étude de l'OFS est importante puisque celle-ci permet de limiter les risques dus à une performance dégradée de la part d'un opérateur dans une situation dangereuse. Par exemple, détecter l'état non fonctionnel d'un contrôleur aérien permet la délégation de ses responsabilités à un autre opérateur ou à un processus d'automatisation. De plus, les avancées récentes dans le domaine des capteurs de signaux physiologiques portatifs [46] et des dispositifs de suivi oculaire [81] à faible coût rendent l'utilisation des méthodes d'estimations de l'OFS hors du laboratoire envisageable.

## 1.1 Opérateur

Comme son nom l'indique, l'estimation de l'OFS consiste à estimer l'état de l'opérateur. L'opérateur est donc une composante nécessaire à l'estimation de l'OFS. Dans le contexte de l'estimation de l'OFS, l'opérateur est la personne qui dirige un procédé ou effectue une tâche. Les contrôleurs aériens, les opérateurs de machinerie lourde, les conducteurs sont tous des exemples d'opérateur dans le contexte de l'analyse de l'OFS. Ceux-ci peuvent interagir avec différents processus d'automatisation des tâches. L'estimation de l'OFS permet également d'activer ou d'éteindre ces processus d'automatisation en réaction à l'état de l'opérateur. Par exemple, dans le cas où une baisse de l'OFS est détectée, certaines automatisations de la tâche peuvent être activées pour permettre à l'opérateur de se concentrer sur les objectifs principaux [36].

Les personnes qui accomplissent une tâche tentent de réaliser un but. Lors de ce processus, ceux-ci produisent des signaux physiologiques, psychologiques et comportementaux. Ces signaux reflètent l'état de l'opérateur. Cet état peut être divisé en trois parties : l'état d'arrière-plan, l'état de base et l'état fonctionnel. L'état d'arrière-plan représente l'état moyen d'un opérateur au repos qui n'a pas de responsabilités ou de buts. Celui-ci peut être vu comme une signature représentant l'opérateur [87]. C'est l'état d'un opérateur tel qu'il pourrait être pris n'importe quel jour. L'état de base est l'état d'un opérateur juste avant de s'engager dans une tâche. Il correspond aux conditions dans lesquelles cette tâche se déroulera. Celui-ci est semblable à l'état d'arrière-plan, mais est plus représentatif de la situation actuelle. Finalement, l'état opérationnel représente l'OFS d'un opérateur activement engagé dans une tâche dans des conditions spécifiques. Il s'agit donc de l'interaction entre l'état d'arrière-plan avec les demandes de la tâche et les conditions actuelles.

L'état fonctionnel de l'opérateur a une influence sur plusieurs signaux physiologiques, psychologiques et comportementaux produits par l'opérateur. Ces signaux peuvent être mesurés et utilisés pour estimer l'OFS.

Le concept d'opérateur est en soi un modèle complexe. En effet, les êtres humains ne peuvent pas être réduits à un modèle simple. Il existe de grandes différences entre les participants dans l'OFS. De plus, les opérateurs ne sont pas que des participants qui accomplissent des expériences. Ceux-ci arrivent au laboratoire avec leurs propres qualifications, stratégies et motivations. Par exemple, une personne venant faire une expérience tentera de réaliser des buts implicites, comme retourner chez lui plus rapidement, lorsqu'il participera à cette expérience. Ces éléments peuvent avoir une certaine

influence sur les performances de celui-ci ainsi que sur l'OFS. Un modèle faisant fit de cette complexité et de cette variance risque donc de mal estimer l'OFS.

### **1.1.1 Dimensions de l'OFS**

En plus des signaux physiologiques, psychologiques et comportementaux, plusieurs autres facteurs doivent être considérés lors de l'analyse de l'état fonctionnel de l'opérateur. En effet, l'opérateur n'est pas pris indépendamment du contexte. D'une part, l'environnement extérieur à une influence sur l'OFS. D'autre part, l'opérateur a également des particularités qui affectent l'analyse de l'OFS. Par exemple, lorsqu'un participant effectue une tâche, il peut être fatigué ou non, stressé ou non, plus ou moins motivé, etc. Ces conditions pouvant affecter l'opérateur sont appelées les dimensions de l'OFS.

#### **Fatigue**

Bien que plusieurs auteurs aient démontré l'association entre la fatigue et la baisse de l'OFS [32], celle-ci est un concept qui, dans le contexte de l'OFS, n'est pas formellement défini. Il s'agit en fait de plusieurs concepts différents de fatigue qui ont tous leur explication. Dans le cadre de ce projet, deux concepts sont utilisés : la fatigue mentale [37] et le manque de sommeil [3, 44]. La fatigue mentale est définie comme un état qui se développe dans le temps, particulièrement lorsqu'une personne effectue un travail ou vie des événements stressants [87]. Le manque de sommeil consiste, comme son nom l'indique, en une quantité insuffisante de sommeil. Le rythme circadien [37] a également une influence sur l'état fonctionnel de l'opérateur et doit être considéré lors de son étude.

#### **Charge de travail**

La charge de travail est une condition pouvant avoir une grande influence sur l'OFS. Celle-ci comprend les trois composantes suivantes : la pression temporelle, les changements de contextes et le niveau de traitement de l'information [61]. La charge de travail est manipulée dans plusieurs études portant sur l'estimation de l'OFS [26, 41, 90].

#### **Autres**

En plus de la fatigue et de la charge de travail, plusieurs autres facteurs intrinsèques à chaque opérateur peuvent avoir une influence sur l'OFS. Par exemple, l'hydratation [76], la maladie [82] ou la motivation du participant [80] sont tous des facteurs dont l'influence sur l'OFS a été démontrée dans des études précédentes.

### 1.1.2 Signaux physiologiques, psychologiques et comportementaux

Les signaux physiologiques sont des mesures produites par le corps lors de son fonctionnement habituel. Le rythme cardiaque, le rythme respiratoire et la température corporelle sont des exemples de mesures physiologiques. Les signaux psychologiques quant à eux sont les variables produites par le cerveau lors de son fonctionnement. Ceux-ci peuvent être mesurés grâce à l'imagerie par résonance magnétique (magnetic resonance imaging, MRI) ou un électroencéphalogramme (EEG). Finalement, les signaux comportementaux sont des variables qui indiquent les comportements de l'opérateur en temps réel. Par exemple, les dispositifs de suivi oculaire peuvent permettre de détecter les fixations du regard ou les saccades.

Plusieurs auteurs ont abordé les liens entre des mesures physiologiques et l'OFS. Par exemple, [85] démontre un lien entre la variabilité du rythme cardiaque (heart rate variability, HRV) et l'OFS. Dans Wilson et Russell [91] et dans Wilson [87], les signaux de rythmes respiratoires sont utilisés avec d'autres variables pour faire une estimation de l'OFS.

Dans Hockey [38], une théorie permettant d'expliquer le lien entre les signaux physiologiques et l'OFS est présentée. Cette théorie explique ces variations dans les signaux physiologiques par une allocation des ressources cognitives à certains buts plutôt que d'autres. Dans ce modèle, lors d'un effort élevé, une mobilisation des ressources supplémentaires implique une plus grande activation du système sympathique. C'est cette activation du système sympathique par rapport au système parasympathique qui est détectée dans les signaux physiologiques.

Les mesures psychologiques sont utilisées dans plusieurs études de l'estimation de l'état fonctionnel de l'opérateur à l'aide d'apprentissage automatique [14, 43, 59, 88, 94]. Celles-ci sont souvent employées en combinaison avec des signaux physiologiques. Cette combinaison est appelée les signaux psychophysiologiques.

Les mesures comportementales sont également employées pour l'estimation de l'état fonctionnel de l'opérateur [23, 24, 60]. Celles-ci sont souvent utilisées en combinaison avec des signaux physiologiques. La combinaison des valeurs physiologiques et comportementales est dite physiocomportementale.

## 1.2 Environnement

Dans le cadre du projet, l'environnement est défini comme l'ensemble de facteurs qui peuvent influencer l'OFS et qui ne sont pas directement liés à la tâche en cours. Par exemple, le niveau de bruit actuel ou la température sont des variables reliées à l'environnement. Dans la figure 1.1, on peut voir que celui-ci influence l'état physiologique de l'opérateur.

Dans Wilson [87], plusieurs de ces facteurs sont présentés. On y retrouve notamment la pression atmosphérique [2], le manque d'oxygène [35], le bruit et les vibrations [38], l'accélération soutenue [15] et la chaleur [42]. Ces effets de l'environnement sur l'OFS nécessitent des conditions extrêmes pour avoir une influence remarquable sur l'OFS. En effet, ceux-ci sont des conditions qui sont observées, entre autres, chez des pilotes de chasse [15, 35].

Dans le cadre de ce projet, un seul facteur environnemental est étudié. Il s'agit du moment de la journée. Celui-ci n'est pas seulement dû au milieu. Il s'agit en fait d'une combinaison entre le rythme circadien de l'opérateur et l'environnement. Cette relation entre le moment de la journée et l'OFS est étudiée puisque celle-ci est une condition qui affecte les opérateurs dans presque tous les contextes. En effet, le manque d'oxygène n'affecte qu'un faible pourcentage des opérateurs alors que plusieurs personnes doivent travailler de nuit ou pendant une longue durée.

## 1.3 Tâche

La dernière composante du modèle présenté à la figure 1.1 consiste en la tâche. Dans le contexte actuel, la tâche correspond à l'activité que l'opérateur est en train d'effectuer. Celle-ci comporte un ou plusieurs buts. Le type de travail peut être très varié. Par exemple, il peut s'agir d'un travail cognitif, comme écrire un mémoire, ou d'un travail physique, comme planter un arbre. Dans ce projet, les tâches utilisées sont de nature cognitive.

### 1.3.1 Tâches et OFS

Plusieurs caractéristiques des tâches ont un effet sur l'OFS. Premièrement, la charge mentale a une influence sur l'état fonctionnel de l'opérateur du participant qui est en train d'effectuer celle-ci [61]. Deuxièmement, la charge physique d'une tâche a également une influence sur son état. Finalement, la conscience de la situation (situational awareness,

SA), c'est-à-dire la capacité à conceptualiser la situation actuelle, est aussi associée à l'OFS [25].

Plusieurs liens unissent les tâches et l'état de l'opérateur. Premièrement, les tâches imposent une certaine charge mentale ou physique à l'opérateur. Celui-ci est donc affecté par l'intensité de celle-ci. De plus, certaines tâches comprennent des rétroactions qui influencent l'opérateur et sa perception de la quantité de travail. À son tour, le participant utilise une stratégie pour répondre aux exigences et celle-ci module la performance de celui-ci.

### 1.3.2 Caractéristiques des tâches

Lors du choix des tâches pour le projet, un des aspects à considérer est leur validité. Certaines sont complexes et permettent d'étudier l'OFS dans un contexte opérationnel. Par exemple, un simulateur de vol permet d'analyser l'OFS dans une situation réaliste. Étant donné leur complexité, l'estimation de l'OFS est plus difficile dans ces tâches. Celles-ci permettent cependant d'étudier la combinaison de plusieurs capacités simultanément ainsi que leur intégration. Par contre, dans ces tâches, l'identification des états à risque est plus difficile. De plus, dans un contexte opérationnel réaliste, les stratégies utilisées par les opérateurs peuvent venir masquer les dégradations de l'OFS.

Pour pallier les faiblesses de l'utilisation de tâches complexes, il est également possible d'avoir recours à des tâches synthétiques. Celles-ci rendent possible l'analyse de certaines compétences individuellement. Par exemple, la tâche du n-back [12, 54] permet d'évaluer la mémoire de travail dans un environnement contrôlé. L'utilisation de ces tâches rend la manipulation des caractéristiques de celles-ci, comme la charge de travail, plus facile. Cependant, les tâches synthétiques ont le désavantage de n'étudier qu'une seule compétence et, par conséquent, sous-estiment l'effet des interactions entre les différentes compétences et l'OFS. Par exemple, certains comportements peuvent être associés à une dégradation de l'OFS dans une tâche, mais pas dans une autre. Une étude utilisant seulement une tâche synthétique est donc limitée dans la portée des conclusions qu'elle peut tirer.

Dans ce projet, l'étude de l'estimation de l'état fonctionnel de l'opérateur se fait à l'aide de tâches synthétiques. Cependant, pour limiter les désavantages liés à leur utilisation, une batterie de tâches est utilisée. Celle-ci permet d'évaluer plusieurs compétences cognitives fondamentales. De plus, l'usage de plusieurs tâches rend possible l'étude de la capacité de généralisation des algorithmes d'apprentissage automatique pour

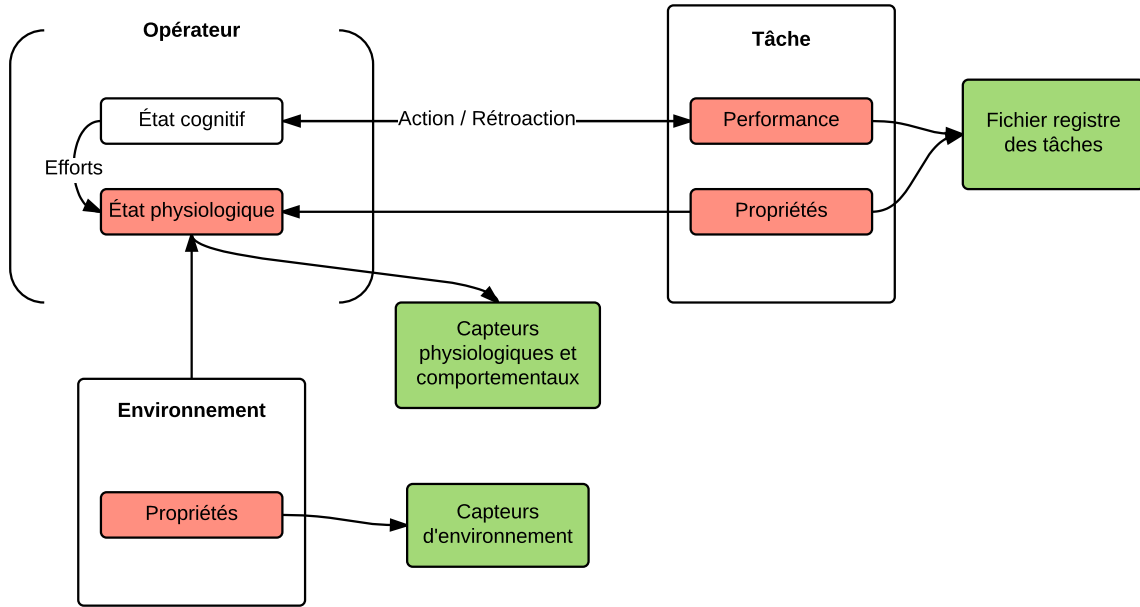


FIGURE 1.2 – Capteurs permettant d’estimer l’OFS.

l’estimation de l’OFS dans plusieurs contextes différents. Les trois tâches qui composent cette batterie sont le n-back [12, 54], une tâche de recherche visuelle [16] et une tâche de rotations mentales [58]. Celles-ci permettent d’étudier respectivement la mémoire de travail, la recherche visuelle et la perception spatiale. Une description de ces tâches est donnée au chapitre 3, section 3.2.2.

## 1.4 Estimation

Le but du projet est d’étudier l’état fonctionnel de l’opérateur dans le but de pouvoir faire l’estimation de celui-ci. Cette section présente les méthodes qui ont été employées dans la littérature pour faire cette estimation ainsi que les défis méthodologiques auxquels il faut faire face.

L’estimation de l’OFS se fait à l’aide de données provenant de trois types de capteurs. Ceux-ci peuvent mesurer des signaux psychophysiques et comportementaux, ainsi que des informations sur l’environnement ou sur les tâches en cours. Les capteurs de signaux physiologiques permettent de recueillir des informations sur l’état de l’opérateur. Les capteurs d’informations sur l’environnement et des tâches en cours offrent un aperçu du contexte. La figure 1.2 présente le modèle de l’OFS ainsi que les capteurs qui peuvent

être utilisés dans l'estimation de l'OFS.

Deux méthodes sont principalement utilisées pour étudier l'OFS : les méthodes statistiques et l'apprentissage automatique.

### 1.4.1 Méthodes statistiques

Plusieurs études ont été faites pour déterminer l'effet de certains facteurs sur l'OFS. Par exemple, certains auteurs étudient le lien entre les signaux physiologiques tels que la HRV et l'état fonctionnel de l'opérateur [85]. D'autres auteurs abordent la relation entre le cycle circadien et l'OFS [3]. Ces recherches sont faites en utilisant des tests statistiques. Celles-ci démontrent donc l'existence d'une relation entre l'OFS et différents signaux ou différentes conditions.

Cependant, les études permettant de faire une association entre une condition physiologique ou environnementale et l'OFS ne peuvent pas être utilisées pour faire une estimation de l'OFS [22]. En effet, bien que celles-ci soient importantes pour établir les liens entre des facteurs et l'OFS, les tests statistiques utilisés ne sont pas faits pour produire des inférences. De plus, ces études ne prennent pas nécessairement en compte la combinaison des signaux physiologiques. Cela a pour effet de limiter l'utilisation de méthodes statistiques pour faire l'estimation de l'OFS à partir d'une combinaison de ces signaux.

### 1.4.2 Apprentissage automatique

L'apprentissage automatique permet d'utiliser des observations qui ont déjà été effectuées pour prédire une valeur cible sur une nouvelle observation. L'usage de l'apprentissage automatique pour estimer l'état fonctionnel de l'opérateur est donc une technique appropriée. Plusieurs auteurs utilisent ces techniques pour estimer l'OFS [53, 90, 92].

Bien que plusieurs auteurs utilisent l'apprentissage automatique pour faire l'estimation de l'état fonctionnel de l'opérateur, plusieurs choix liés à l'OFS peuvent influencer les performances de ces estimations. Dans cette section, plusieurs choix méthodologiques faits par différents auteurs sont présentés pour permettre de juger des avantages et des inconvénients de ces choix.



## Choix des signaux

Bien que plus de détails sur l'apprentissage automatique soient donnés au chapitre 2, il est important de savoir que celui-ci permet d'associer une valeur en entrée à une valeur en sortie. Pour ce faire, ces algorithmes utilisent des exemples du passé qui leur ont été fournis sous forme d'un jeu de données d'entraînement. Ce jeu de données prend la forme d'un ensemble de paires de valeurs contenant des signaux d'entrée et des valeurs cibles qui leur sont associées. Le choix de ces valeurs est crucial. En effet, si les exemples ne contiennent pas d'information sur les valeurs en sortie, aucun algorithme d'apprentissage automatique ne pourra apprendre de liens valides.

Une des premières étapes qui doit être effectuée lors de l'apprentissage automatique consiste à choisir les signaux d'entrées. Plusieurs auteurs qui emploient ces méthodes utilisent une combinaison de signaux physiologiques et psychologiques [90, 95]. Certains ont également utilisé une combinaison de signaux physiologiques et comportementaux [23, 72].

Les paragraphes suivants présentent des signaux physiologiques, comportementaux et psychologiques qui sont utilisés dans la littérature pour faire l'estimation de l'OFS.

**Rythme cardiaque** Le rythme cardiaque est défini comme le nombre de contractions du cœur par minutes. Pour un adulte au repos, celui-ci se situe entre 50 et 90 battements par minute [63]. Il peut être obtenu à l'aide des signaux produits par un électrocardiogramme (ECG).

**Variabilité du rythme cardiaque** La variabilité du rythme cardiaque (heart rate variability, HRV) est la variation dans le temps, de la durée entre des battements cardiaques consécutifs [1]. La HRV reflète la capacité du cœur à s'adapter à des stimuli imprévisibles. L'analyse de celle-ci permet d'obtenir des informations sur le système nerveux autonome. Plusieurs techniques permettent d'en faire son analyse : l'analyse dans le domaine temporel, l'analyse dans le domaine fréquentiel et l'analyse non linéaire. Celles-ci sont effectuées à l'aide des signaux provenant d'un ECG.

**Rythme respiratoire** Le rythme respiratoire est défini par le nombre de respirations par minutes. Pour un adulte, le rythme respiratoire normal se situe sous 20 respirations par minutes [19].

**Posture** Dans ce projet, la posture correspond à une mesure de l'angle du torse par rapport au sol. Cette mesure est utilisée dans certains systèmes de mesures portatifs [46, 84].

**Température corporelle** La température corporelle est mesurée en degré Celsius. Celle-ci est utilisée dans plusieurs études sur l'OFS [49, 50, 53].

**Clignements** Les clignements se produisent quand la paupière couvre complètement l'œil [72]. Ceux-ci sont des indicateurs de la fatigue [7] et de la charge de travail [11]. Le clignement des yeux peut être détecté par un dispositif de suivi oculaire.

**Position du regard** La position du regard est définie par la direction du regard de l'opérateur. Celle-ci peut être donnée par la déviation en degré par rapport à un regard centré. Elle peut également être donnée par la position du regard par rapport à un objet fixe, comme un écran d'ordinateur. Cette position peut être obtenue grâce à un dispositif de suivi oculaire.

**Dilatation de la pupille** La dilatation de la pupille est un signal comportemental permettant d'estimer l'OFS [55, 72]. La valeur de la dilatation de la pupille est obtenue par un dispositif de suivi oculaire.

**Saccades et Fixations** Les fixations sont les moments où le regard de l'opérateur se concentre sur un point en particulier. Les saccades sont les moments où le regard de l'opérateur est en mouvement. Les saccades et les fixations donnent des informations comportementales sur l'opérateur [72]. L'état du regard peut être calculé à partir des données du dispositif de suivi oculaire.

**EEG** L'électroencéphalogramme permet de mesurer l'activité électrique du cerveau. Celui-ci fonctionne grâce à des électrodes placées sur le scalp du participant. L'utilisation d'un EEG permet d'obtenir des signaux psychologiques permettant d'estimer l'OFS [30, 59, 73].

Étant donné qu'un des buts secondaires du projet est de vérifier la faisabilité de l'utilisation de ces algorithmes d'estimation de l'OFS dans une plateforme mobile [30], l'utilisation des signaux psychologique comme l'EEG n'est pas recommandée. En effet, bien que certains auteurs aient obtenus de meilleures performances en utilisant un

EEG [43], ceux-ci comportent trop d'artéfacts dus aux mouvements pour être utilisés dans un contexte hors du laboratoire [73].

### Valeur cible

La valeur cible est la valeur que l'algorithme d'apprentissage automatique doit estimer en fonction d'une valeur en entrée. Son choix est crucial pour obtenir des conclusions valides. Comme il a été mentionné précédemment, l'OFS est un modèle permettant d'expliquer des phénomènes observés dans les performances des opérateurs. De cette façon, il est analogue au concept de l'intelligence. Cependant, comme l'intelligence, l'OFS ne se mesure pas directement. Pour permettre l'utilisation des algorithmes d'apprentissage automatique, il est donc nécessaire de définir une mesure intermédiaire permettant d'estimer l'OFS à un moment donné.

La façon intuitive d'obtenir cette valeur est d'utiliser la performance de l'opérateur dans les différentes tâches. Puisque le but de l'estimation de l'OFS est de prévenir les chutes de performances, son utilisation comme valeur cible peut sembler intéressante. Cependant, celle-ci n'est pas directement associée à l'OFS. En effet, lorsqu'un participant se retrouve dans un état sous-fonctionnel, celui-ci peut « protéger » la performance dans sa tâche principale au prix d'un effort supplémentaire [39]. Ce phénomène est représenté à la figure 1.3. Cette figure présente la performance d'un opérateur en fonction de la charge de travail. On peut y remarquer qu'il est possible pour un participant de conserver la même performance lorsque la charge de travail augmente en fournissant un effort supplémentaire. Éventuellement, la charge de travail peut être suffisamment élevée et il devient impossible de fournir un effort pour garder un niveau de performance acceptable. De plus, on peut remarquer que dans le cas où la charge de travail est faible, la performance peut également se dégrader. Ce phénomène se nomme le désengagement [13].

Deux méthodes sont utilisées dans la littérature pour obtenir une valeur cible représentative de l'OFS. La première consiste à employer les dimensions de l'OFS comme une estimation de l'état fonctionnel de l'opérateur. Ces dimensions sont des états pour lesquels un lien avec l'OFS est connu. La fatigue, le stress ou la charge de travail sont des exemples de dimensions de l'OFS. Par exemple, certains auteurs manipulent la charge de travail de la tâche effectuée et utilisent cette charge de travail comme valeur cible pour les algorithmes d'apprentissage automatique [24, 89, 90]. Cette technique repose sur le raisonnement suivant : puisqu'un lien a été démontré entre l'OFS et la charge de travail, l'estimation de la charge de travail à l'aide de l'apprentissage automatique permet d'obtenir une estimation de l'OFS.

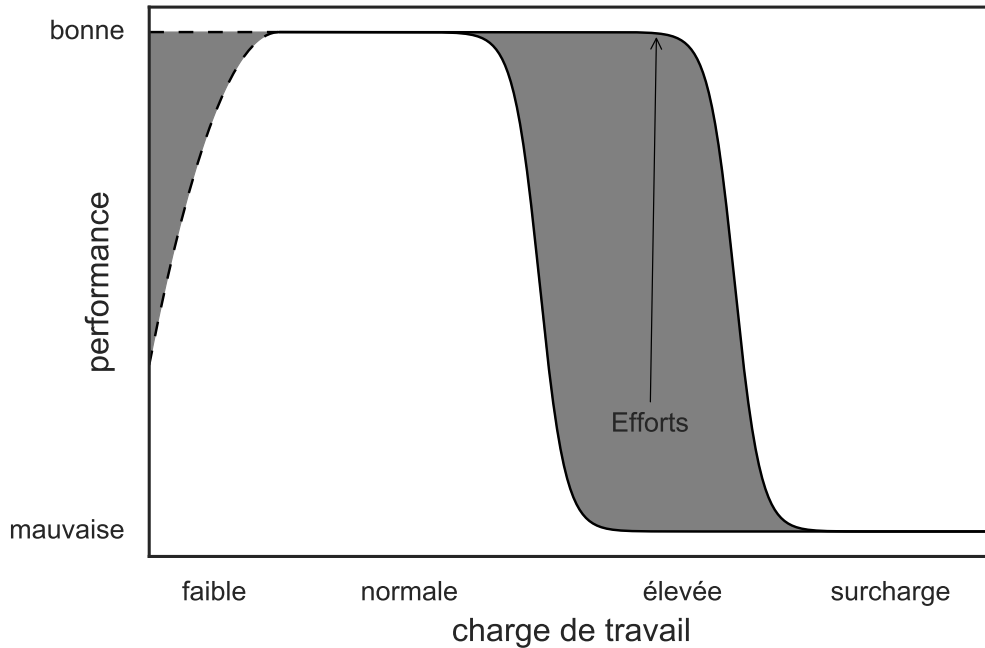


FIGURE 1.3 – Diagramme de la chute de performance en fonction des demandes de la tâche.

La deuxième méthode consiste à utiliser une mesure de performance dans un but secondaire à la tâche [86, 92]. En effet, bien que la performance dans la tâche principale puisse être protégée par un changement de stratégie ou un effort supplémentaire, une dégradation de l'OFS affecte la performance dans les buts secondaires [39].

Ces deux méthodes présentent chacune leurs avantages et leurs inconvénients. L'utilisation des dimensions de l'OFS, telles que la charge de travail et la fatigue, a l'avantage d'être objective et plus facile à contrôler. Cependant, ces dimensions sont difficiles à généraliser à plusieurs tâches ou plusieurs participants. En effet, il est difficile de comparer les niveaux de charge de travail dans des tâches différentes. Par exemple, comment comparer une charge de travail élevée pour un contrôleur aérien et un pilote d'avion ? De même, il est difficile de comparer le niveau de fatigue entre plusieurs opérateurs.

L'utilisation de la performance secondaire a le bénéfice d'être plus directement liée au concept de l'OFS. Cependant, comme dans le cas de l'usage des dimensions de l'OFS, la performance est difficile à comparer entre les différentes tâches ou les différents participants.

## 1.5 Objectifs de recherche

Dans Brouwer *et al.* [10], plusieurs conseils concernant l'utilisation de l'apprentissage automatique sont formulés. Ce mémoire poursuit quatre objectifs de recherche en lien avec ces conseils. Ceux-ci sont la formalisation de l'OFS dans un contexte d'apprentissage automatique, l'évaluation de la capacité des algorithmes d'apprentissage automatique à estimer l'OFS, l'étude de la capacité de ces algorithmes à généraliser à de nouveaux participants et à de nouvelles tâches et, finalement, l'évaluation de la possibilité d'utiliser le contexte pour permettre d'améliorer les performances de ceux-ci. Ces objectifs rendent possible l'évaluation d'une manière objective les performances pouvant être atteintes dans l'estimation de l'OFS à l'aide de l'apprentissage automatique dans plusieurs contextes.

### 1.5.1 Formalisation de l'OFS

L'usage de l'apprentissage automatique pour estimer l'OFS n'est pas une méthode simple. En effet, plusieurs choix dans le design d'une expérience d'apprentissage automatique peuvent avoir une influence sur la possibilité d'employer les résultats dans un contexte pratique. Bien que plusieurs auteurs utilisent l'apprentissage automatique pour faire l'estimation de l'OFS, il existe plusieurs techniques différentes pour appliquer celui-ci à ce problème. De plus, aucune méthodologie proposée ne permet de faire cet apprentissage dans un contexte comprenant plusieurs tâches et plusieurs participants à la fois.

Pour cette raison, le premier objectif de ce projet est de proposer une formalisation de l'OFS permettant d'appliquer des algorithmes d'apprentissage automatique. Concrètement, cet objectif comprendra le choix des signaux d'entrée ainsi que la valeur cible qui seront utilisés par les algorithmes.

### 1.5.2 Évaluation de l'apprentissage automatique

La performance obtenue lors de l'apprentissage automatique dépend de plusieurs choix dans la modélisation de l'OFS. Il est difficile de comparer les résultats de deux études qui n'utilisent pas les mêmes signaux en entrée, les mêmes tâches, ni les mêmes valeurs cibles. Ce projet de recherche vise donc l'évaluation des performances obtenues par l'apprentissage automatique et permettra de comparer plusieurs algorithmes entre eux pour déterminer si un de ceux-ci permet d'obtenir de meilleures performances.

### 1.5.3 Généralisation

Une autre question abordée par ce projet est de déterminer la capacité de généralisation et de synthèse des modèles d'apprentissage automatique pour l'estimation de l'OFS. Cet objectif de recherche vise à répondre aux deux interrogations suivantes : est-ce que ces algorithmes obtiennent de meilleures performances lorsqu'ils sont entraînés et évalués avec les données d'un seul participant ou une seule tâche à la fois ? Est-ce que les modèles d'apprentissage automatique entraînés sur les données de plusieurs participants peuvent être utilisés sur les données d'un nouveau participant ?

### 1.5.4 Utilisation du contexte

Enfin, ce projet a comme objectif de recherche de déterminer l'effet de l'utilisation du contexte sur les performances des algorithmes d'apprentissage automatique. Le contexte correspond aux conditions externes qui pourraient avoir une influence sur l'OFS. En effet, le modèle présenté à la figure 1.1 inclut des interactions avec l'environnement. Cependant, les informations provenant de l'environnement sont rarement utilisées lors d'expériences étudiant l'apprentissage automatique pour faire l'estimation de l'OFS. Dans certains cas, ces informations sont employées comme cible d'apprentissage alors que, dans d'autres cas, cette information n'est tout simplement pas utilisée. Une évaluation de l'apport de ces informations sur la capacité d'apprentissage et de généralisation fait donc partie des objectifs de ce mémoire.

# Chapitre 2

## Apprentissage automatique

Jusqu'à maintenant, le concept de l'OFS fut présenté ainsi que plusieurs méthodes d'estimation de celui-ci. Une technique utilisée par certains auteurs pour faire celle-ci est l'apprentissage automatique. Il est donc important de décrire plus en détail ce qu'est l'apprentissage automatique et les éléments nécessaires à son application dans la détermination de l'état fonctionnel de l'opérateur. Plus précisément, on décrit ici les différentes familles d'algorithmes, les modèles qui seront utilisés dans ce projet ainsi que les méthodes pour appliquer ces algorithmes dans l'estimation de l'OFS sont abordés dans ce chapitre.

### 2.1 Définition

L'apprentissage automatique permet de trouver une solution à une catégorie de problèmes pour lesquels les algorithmes à base de règle et de séquence d'instruction peinent à trouver une solution. Dans plusieurs cas, bien que cet algorithme permettant de trouver une solution soit inconnu, nous disposons de plusieurs exemples de la bonne solution. Par exemple, bien qu'il n'existe pas d'algorithme à base de règle pour reconnaître efficacement des caractères manuscrits, plusieurs exemples de lettres manuscrites ainsi que le caractère représenté sur chacune de celles-ci sont disponibles.

L'apprentissage automatique consiste à utiliser les expériences passées pour apprendre la solution à un problème [4]. Ces algorithmes utilisent en entrée des jeux de données qui contiennent des observations, c'est-à-dire un ensemble de valeurs. Celles-ci sont souvent organisées sous forme de vecteurs et sont notées  $\mathbf{x}$ . Ces techniques utilisent de grandes quantités de données qui sont organisées sous la forme de jeux de données.

Pour faire ses prédictions, l'apprentissage automatique a deux composantes principales. La première composante est l'algorithme ou le modèle employé. La deuxième composante est le jeu de données utilisé. En effet, la qualité des données est un critère fondamental permettant à un algorithme d'apprentissage automatique d'obtenir de bonnes performances.

Les algorithmes d'apprentissage ont une influence significative sur la capacité d'apprendre des informations à partir des données fournies en entrée. Les différents algorithmes ont plusieurs propriétés qui influencent les performances sur plusieurs critères. Par exemple, certains demandent plus de ressources computationnelles que d'autres. Les algorithmes ont également des propriétés différentes qui leur permettent de mieux fonctionner sur certains types de données.

### 2.1.1 Apprentissage supervisé et non supervisé

On peut séparer l'apprentissage automatique en deux grandes familles d'algorithmes : l'apprentissage supervisé et l'apprentissage non supervisé.

#### **Apprentissage non supervisé**

Les algorithmes d'apprentissage non supervisé ne disposent que de données en entrée. Ceux-ci utilisent alors des jeux de données ayant la forme présentée à l'équation 2.1. Ces modèles permettent de trouver des similitudes dans les données et d'extraire des caractéristiques intéressantes dans celles-ci. Par exemple, un algorithme pourrait recevoir la liste de tous les achats effectués par un client dans un commerce. La tâche de l'algorithme consiste à faire des regroupements de clients ou trouver des similitudes entre les comportements d'acheteur.

Dans le cas des algorithmes non supervisés, les jeux de données sont représentés à l'aide de la notation suivante :

$$\mathcal{D} = \{\mathbf{x}_i\}_{i=1}^N, \tag{2.1}$$

où  $\mathcal{D}$  représente le jeu de donnée,  $i$  représente l'indice dans le jeu de donnée,  $\mathbf{x}$  est une observation et  $N$  est la cardinalité du jeu de donnée.



## Apprentissage supervisé

Lors d'un apprentissage supervisé, nous disposons de données d'entrées ainsi que la valeur de sortie correspondante. Dans ce cas, les jeux de données ont la forme suivante :

$$\mathcal{D} = \{\mathbf{x}_i, y_i\}_{i=1}^N \quad (2.2)$$

où  $y_i$  représente une étiquette correspondant à la sortie associée à l'entrée  $\mathbf{x}$ .

En apprentissage automatique, on utilise l'équation suivante :

$$y = g(\mathbf{x}|\theta) \quad (2.3)$$

Dans cette équation,  $g$  correspond au modèle utilisé,  $\mathbf{x}$  est le vecteur en entrée,  $\theta$  désigne les paramètres du modèle et  $y$  représente la valeur prédite.

La prédiction peut être évaluée grâce à la fonction de perte. Cette fonction retourne une valeur correspondant à la justesse de la prédiction. Dans le cas où celle-ci est correcte, cette fonction retourne une valeur de 0. Si elle est incorrecte, la valeur retournée est 1. La fonction de perte est notée de la façon suivante :

$$L(y_{prediction}|y_{vrai}) \quad (2.4)$$

L'apprentissage automatique est un ensemble de méthodes permettant de déterminer les paramètres  $\theta$  permettant de réduire le plus possible les erreurs dans la prédiction de  $y$ . En d'autres termes, on vise à effectuer l'opération suivante :

$$\operatorname{argmin}_{\theta} \left( \sum_{i=0}^N L(g(\mathbf{x}_i|\theta) | y_i) \right) \quad (2.5)$$

**Classification et Régression** Deux types de données en sortie peuvent être utilisés : des valeurs catégorielles et des valeurs réelles. Une tâche pouvant être faite à l'aide des algorithmes supervisés est la classification. Dans cette tâche, une observation est associée à une étiquette. Le but de la tâche est de prédire l'étiquette associée à une nouvelle donnée. Par exemple, si nous désirions entraîner un modèle à différencier des photos de chats et de chiens [65], il faut un jeu de données dans lequel plusieurs photos de chiens et de chats sont disponibles, ainsi que la correspondance associant chaque photo à la bonne classe.

Les algorithmes supervisés peuvent aussi être utilisés pour faire de la régression. Dans celle-ci, la valeur en sortie correspond à un nombre réel et le modèle doit estimer cette

valeur à partir de l'observation. Par exemple, une tâche de régression peut consister à estimer le niveau de pollution par rapport à plusieurs variables obtenues par des capteurs [21].

## 2.2 Algorithmes évalués

Parmi les familles d'algorithmes présentés dans les sections précédentes, les familles les plus aptes à estimer l'état fonctionnel de l'opérateur sont les algorithmes d'apprentissage supervisé. En effet, les données provenant des différents capteurs physiologiques et comportementaux sont utilisées en entrée, alors que les données provenant des tâches effectuées sont employées pour produire une valeur cible.

Il existe plusieurs modèles pouvant être utilisés pour l'apprentissage supervisé. Il est donc important d'évaluer plusieurs algorithmes pour avoir une mesure de la performance sur les données effectives. En effet, les données utilisées peuvent avoir des propriétés inconnues ayant une influence importante sur la performance de certains algorithmes.

Dans le cadre du projet, trois modèles différents et leur application à l'estimation de l'OFS sont évalués : le K-voisins les plus proches (K-nearest neighbors, KNN), le séparateur à vastes marges (support vector machines, SVM) et les forêts d'arbres aléatoires.

### 2.2.1 KNN

L'algorithme du KNN est une méthode non paramétrique d'apprentissage automatique [18, 29]. L'algorithme d'entraînement consiste simplement à mémoriser tous les points donnés lors de l'entraînement et leur étiquette. Au moment de l'utilisation, les  $k$  points les plus près sont sélectionnés et leur étiquette est consultée. La classification se fait en prenant le mode des étiquettes de ces  $k$  points.

### 2.2.2 Séparateurs à Vastes Marges

Les SVM représentent une méthode d'apprentissage automatique non paramétrique [17]. Leur modèle consiste en la somme de l'influence d'un sous-ensemble des observations.

L'algorithme du SVM consiste à placer un hyperplan séparateur entre les instances des éléments des deux classes à l'étude. Les paramètres de cet hyperplan sont choisis de façon à maximiser la marge entre les deux classes. Pour classifier de nouveaux exemples,

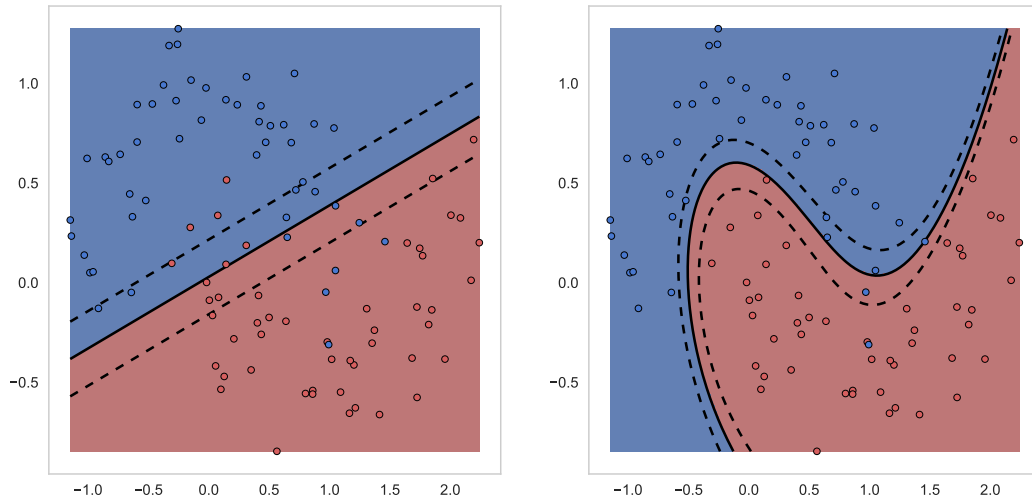


FIGURE 2.1 – Fonctionnement d'un SVM linéaire (gauche) et d'un SVM à noyau gaussien (droite)

on calcule l'influence d'un sous ensemble de points dans le jeu d'entraînement qui sont près de la marge. Les points de ce sous-ensemble sont appelés les vecteurs de support. Les autres exemples sont ignorés puisqu'ils n'influencent pas le placement de la marge.

Une variante du SVM très utilisée consiste à multiplier tous les points du vecteur d'entrée par un noyau. Le SVM maximise alors les marges dans l'espace du noyau. Cette technique est appelée le truc du noyau, ou *kernel trick* [77]. Un exemple de SVM linéaire et avec noyau gaussien est présenté à la figure 2.1. Dans cette figure, le SVM doit classifier deux groupes de points. La frontière de décision entre les deux classes est représentée par la ligne. Les lignes pointillées représentent les marges autour de la frontière de décision qui ont été conservées par l'algorithme. Cette figure met en évidence l'utilité du noyau gaussien, qui permet de séparer des données qui ne sont pas linéairement séparables. En utilisant le truc du noyau, on peut donc obtenir une frontière de décision arbitrairement complexe.

### 2.2.3 Arbres de décisions

Les arbres de décisions sont des classifieurs non paramétriques. Ils fonctionnent en appliquant une suite de décisions binaires organisées sous forme d'arbre [4]. Pour apprendre les règles composant ceux-ci, l'algorithme d'arbre de classification et de régression (classification and regression tree, CART) peut être utilisé [71]. Un exemple

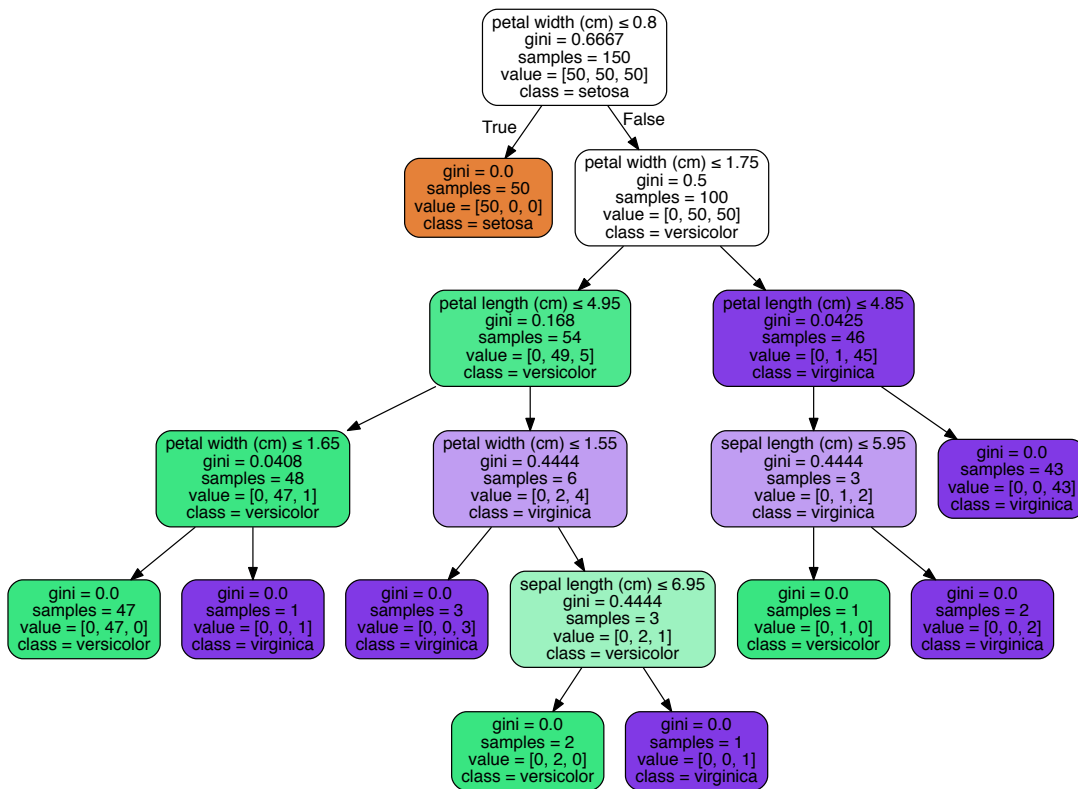


FIGURE 2.2 – Arbre de décision pour la classification du jeu de données iris.

d'arbre de décision est présenté à la figure 2.2.

Les arbres de décisions ont plusieurs propriétés qui les rendent populaires. Contrairement à la plupart des algorithmes, ils fonctionnent aussi bien avec des données qui ne sont pas normalisées ou prétraitées. De plus, le classifieur est un arbre de décision qui peut être interprété par des humains.

Cependant, les arbres de décisions peuvent être trop complexes, et ainsi, faire du sur apprentissage. De plus, l'apprentissage de ceux-ci est considéré comme instable, c'est-à-dire que de petites variations dans les données en entrée peuvent provoquer des changements importants dans les modèles obtenus.

## 2.2.4 Forêt d'arbres aléatoires

La classification par forêt d'arbre aléatoire est une technique permettant de limiter les faiblesses des arbres de décisions [9]. Ce modèle consiste à entraîner plusieurs arbres de décisions sur des sous-ensembles de données différentes. Lors de la classification, chaque arbre vote et la classe sélectionnée est la classe ayant reçu le plus de votes. Cette technique atténue la propriété d'instabilité des arbres de décisions.

## 2.3 Évaluation des algorithmes

La meilleure façon d'évaluer la performance d'un algorithme d'apprentissage automatique par rapport à un autre est de mesurer sa performance sur un jeu de données. Pour pouvoir se prononcer sur la capacité de ces algorithmes à faire des prédictions valables, il faut donc avoir une procédure d'évaluation des performances rigoureuse.

L'évaluation de la performance d'un classifieur est faite pour deux raisons. La première est de déterminer quel classifieur permet d'obtenir l'estimation de l'OFS la plus fiable. La deuxième est d'évaluer la performance de ce classifieur sur de nouvelles données. Pour faire cette évaluation, il est nécessaire d'avoir un jeu de données d'entraînement et un jeu de données de tests. Il faut également déterminer une mesure permettant de comparer les performances des algorithmes utilisés.

### 2.3.1 Choix du classifieur

Les classifieurs présentés dans les sections précédentes reposent sur des algorithmes différents. Pour cette raison, les décisions prises par chacun de ces algorithmes seront différentes. La figure 2.3 présente les frontières de décisions produites par ces modèles sur le jeu de données *iris*. On peut voir dans cette figure que les décisions prises par chaque classifieur sont différentes, bien que le taux de classifications réussi soit semblable.

Dans ces conditions, le lecteur est en droit de se demander quel classifieur est le meilleur. Malheureusement, il n'y a pas de façon de connaître le classifieur qui fonctionnera le mieux sur un jeu de données, sauf en l'essayant et en évaluant ses performances.

### 2.3.2 Jeux de données

Pour évaluer la performance d'un algorithme d'apprentissage automatique, il faut tester celui-ci sur des données qui n'ont jamais été vues ou utilisées par celui-ci. En effet,

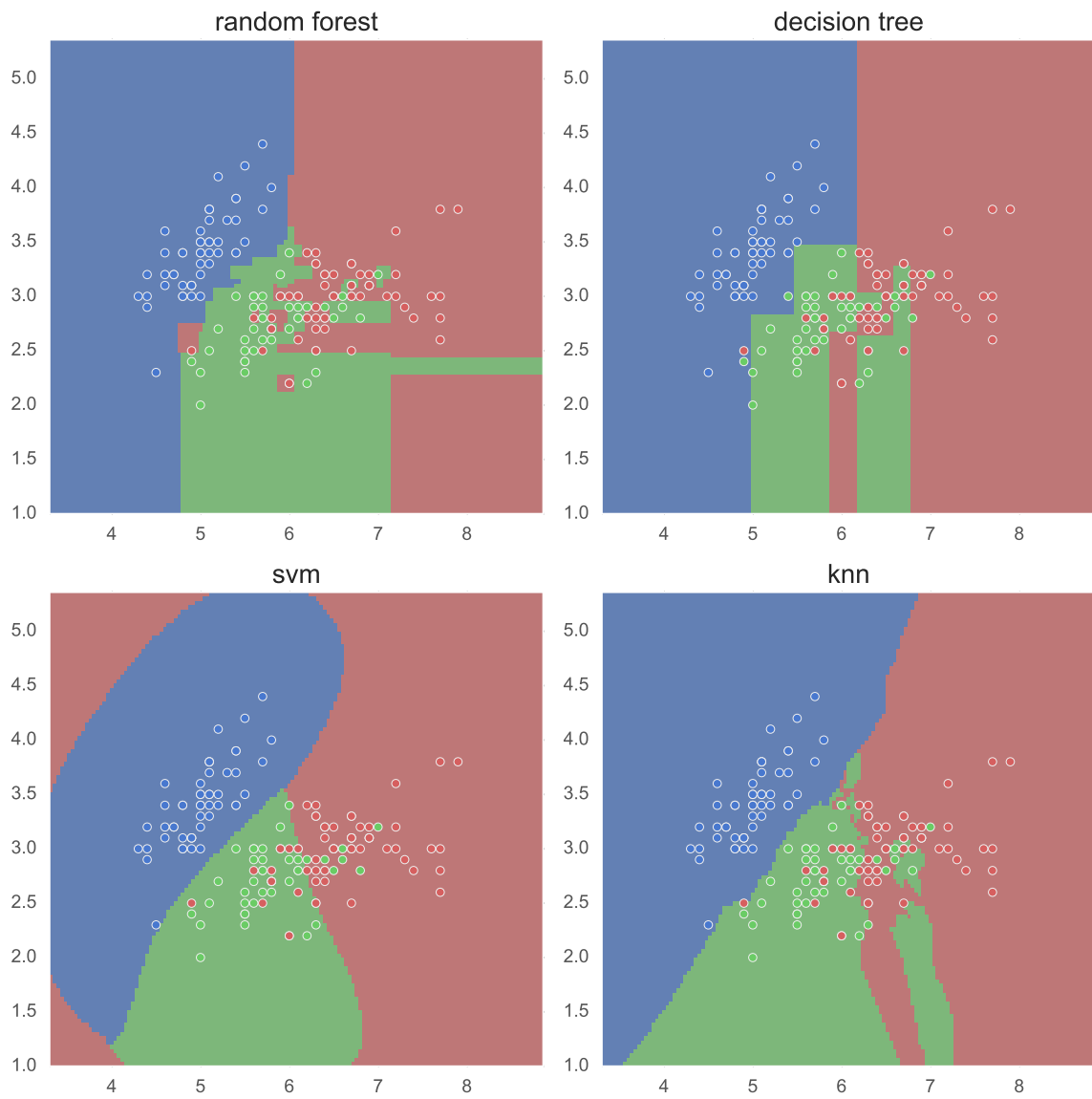


FIGURE 2.3 – Frontières de décisions obtenues par quatre classifieurs sur deux dimensions des données dans le jeu de données iris.

dans le cas où l'on utiliserait les mêmes données pour l'entraînement et le test, un dictionnaire mémorisant les données obtiendrait une performance parfaite. Pour cette raison, le jeu de donnée récolté sera divisé en deux jeux distincts, un jeu de test et un jeu d'entraînement. Le jeu d'entraînement servira à l'entraînement et le jeu de test servira à évaluer la performance finale de l'algorithme d'apprentissage automatique.

Il est très important de limiter l'utilisation du jeu de test. Comme le rapporte Jensen et Cohen [45], réutiliser le jeu de test pour faire plusieurs comparaisons diminue beaucoup la portée des conclusions pouvant être effectuées avec celui-ci. En effet, chaque utilisation du jeu de test augmente les chances que les conclusions tirées soient dues au hasard.

Le partitionnement du jeu de données en un jeu d'entraînement et un jeu de test peut grandement influencer les conclusions que l'on peut faire des expériences d'apprentissage automatique. L'évaluation que l'on fait des performances d'un modèle n'est valide que sur des jeux de données indépendant et identiquement distribués (IID). Dans le cas de l'estimation de l'OFS, ces conditions ne sont pas respectées. En effet, les données étant séquentielles par nature, celles-ci ne sont donc pas indépendantes. De plus, les différents participants produisent des données qui ont une distribution différente.

Il est donc important de prendre des précautions pour limiter les problèmes occasionnés par les dépendances temporelles et les changements de domaines. Pour pouvoir tirer les conclusions les plus valables de la performance attendue des algorithmes d'apprentissage automatique hors du laboratoire, il faut faire une séparation du jeu de données qui reflètera l'utilisation potentielle de ces algorithmes. Par exemple, si l'on désire estimer la performance de ces algorithmes sur les données d'un nouveau participant, il est nécessaire que les données d'entraînement et les données de test proviennent de participants différents. De façon générale, il est également préférable que le jeu d'entraînement et le jeu de test contiennent des données de chaque classe dans la même proportion.

### 2.3.3 Mesure de la performance

Lorsqu'un classifieur est utilisé sur un jeu de données, celui-ci produit une valeur d'étiquette prédite par observation. Cependant, il est nécessaire de traduire cette prédiction en une mesure de la performance. Cette mesure de la performance peut être notée de façon générale par la fonction de perte présentée à l'équation 2.4.

Pour mesurer la performance des algorithmes de classification, plusieurs mesures ont été proposées [4, 70]. Dans le cas à deux classes, quatre situations sont possibles. Un

		Prédiction		total
		p	n	
Valeur	p'	Vrai Positif	Faux Négatif	P'
	n'	Faux Positif	Vrai Négatif	Neg'
total		P	Neg	

FIGURE 2.4 – Matrice de confusion

échantillon avec une classe positive obtient une prédiction positive, il s'agit alors d'un vrai positif. Un échantillon négatif obtient une classe négative, il s'agit alors d'un vrai négatif. Un échantillon positif obtient une classe négative, ceux-ci sont appelés des faux négatifs. Finalement, lorsqu'un échantillon négatif obtient une classe positive, on parle d'un faux positif. La figure 2.4 présente la matrice de confusion qui regroupe ces quatre cas de figure. Dans cette section, les vrais positifs sont représentés par la variable  $tp$ , les vrais négatifs sont représentés par la variable  $tn$ , les faux positifs sont représentés par  $fp$  et les faux négatifs sont représentés par  $fn$ . Le nombre d'observations prédites est représenté par  $N$ .

Ces quatre cas sont la façon la plus complète de présenter la performance de la classification. Cependant, plusieurs situations nécessitent d'obtenir une seule mesure de la performance, et non quatre. Pour ces situations, d'autres mesures de performances peuvent être utilisées [4].

Une mesure fréquemment utilisée est l'exactitude. Cette mesure peut être calculée à l'aide de l'équation suivante :

$$\text{exactitude} = \frac{tp + tn}{N} \quad (2.6)$$

L'exactitude est une mesure qui permet d'évaluer la performance d'un classifieur. Cependant, cette mesure n'est pas valide dans le cas où la taille des classes n'est pas proche. En effet, l'importance relative des différentes classes accorde une importance plus importante à la classe la plus fréquente.

Pour pallier cette faiblesse, d'autres mesures de la performance en classifications sont disponibles [70]. Les coefficients de corrélation de Matthews (Matthew's correlation



coefficients, MCC) [56] prennent en compte à la fois les vrais positifs, les vrai négatifs, les faux positifs et les faux négatifs dans leur mesure de la performance. Cette mesure peut entre autres être utilisée lorsque les classes ne sont pas balancées.

Les MCC sont en fait la moyenne géométrique de la *Markedness* et de la *Informedness*. Les équations permettant de calculer ces trois mesures sont les suivantes :

$$\text{Markedness} = \frac{tp}{tp + fn} - \frac{fn}{fp + tn} \quad (2.7)$$

$$\text{Informedness} = \frac{tp}{tp + fp} - \frac{fn}{fn + tn} \quad (2.8)$$

$$\text{MCC} = \sqrt{\text{Markedness} \cdot \text{Informedness}} \quad (2.9)$$

D'un point de vue pratique, les MCC sont une mesure qui peut prendre une valeur de -1 à 1. Dans le cas d'une prédiction parfaite, les MCC ont une valeur de 1 et dans une prédiction contraire, les MCC ont une valeur de -1. Les prédictions aléatoires donnent des MCC de 0.

## 2.4 Entraînement

Les algorithmes présentés à la section 2.2 permettent de faire des prédictions à partir d'observations du passé. Ces prédictions peuvent être évaluées grâce aux techniques présentées dans la section 2.3. Le dernier aspect à déterminer pour pouvoir utiliser les algorithmes d'apprentissage automatique pour estimer état fonctionnel de l'opérateur est la procédure d'entraînement et de choix des modèles. Cette procédure comprend les prétraitements effectués sur les données, le choix des hyper paramètres ainsi que l'évaluation des modèles.

### 2.4.1 Prétraitement

À l'état brut, les données sur lesquelles l'apprentissage automatique est effectué ne permettent pas toujours d'obtenir les meilleures performances possible. De plus, les données peuvent être incomplètes. Pour pallier ces problèmes, il est donc généralement nécessaire d'effectuer un prétraitement sur les données reçues en entrée.

Le premier prétraitement devant être effectué consiste à traiter les données manquantes. En effet, lorsque les techniques d'apprentissage automatique sont utilisées sur des données réelles, celles-ci sont généralement imparfaites. Dans plusieurs cas, des données sont

absentes. Deux méthodes peuvent être employées dans le cas des données manquantes. La première méthode consiste simplement à ne pas utiliser les observations dans lesquelles certaines valeurs sont manquantes. Une autre technique est d'utiliser un estimateur tel que la moyenne ou la médiane pour calculer les valeurs qui ne sont pas disponibles.

Pour plusieurs algorithmes, les données utilisées doivent être standardisées pour chaque dimension. Par exemple, le SVM avec noyau gaussien nécessite des données avec une moyenne centrée autour de 0 et un écart-type de 1 pour fonctionner de façon optimale. Il est donc préférable de soustraire par la moyenne et de diviser par l'écart-type du jeu d'entraînement pour chaque dimension. Pour augmenter la stabilité du processus, il est également possible d'utiliser la médiane et l'étendue interquartile plutôt que la moyenne et l'écart-type.

Un élément important à garder en tête lors du choix et de l'implémentation des prétraitements est de ne jamais utiliser des données provenant du jeu de données de tests. Les valeurs employées pour calculer les estimations doivent uniquement provenir du jeu de données d'entraînement.

## 2.4.2 Choix des hyper paramètres

Les algorithmes présentés dans la section 2.2 ont chacun des paramètres pour lesquels des valeurs doivent être sélectionnées. Ceux-ci sont appelés des hyper paramètres. Par exemple, dans le cas du KNN, le nombre de voisins utilisés pour faire une prédiction doit être choisi. Ces paramètres ont une grande influence sur la performance que le classifieur peut obtenir. Cependant, il n'existe pas nécessairement de valeur fonction analytique permettant d'obtenir une valeur optimale pour ces hyper paramètres. Celle-ci dépend généralement des données utilisées.

Pour sélectionner les hyper paramètres, il faut procéder en deux étapes. Premièrement, il faut obtenir une liste d'hyper paramètres. Deuxièmement, il faut estimer la performance de ces hyper paramètres afin de choisir ceux qui permettront d'obtenir la meilleure performance en généralisation.

### Liste des hyper paramètres

Pour obtenir les hyper paramètres qui seront évalués, deux techniques sont fréquemment utilisées : la recherche en grille [4] et la recherche aléatoire [6]. La recherche en grille consiste à essayer les hyper paramètres de façon systématique. La procédure de recherche se termine une fois toutes les valeurs évaluées. Par exemple, dans le cas d'un classifieur

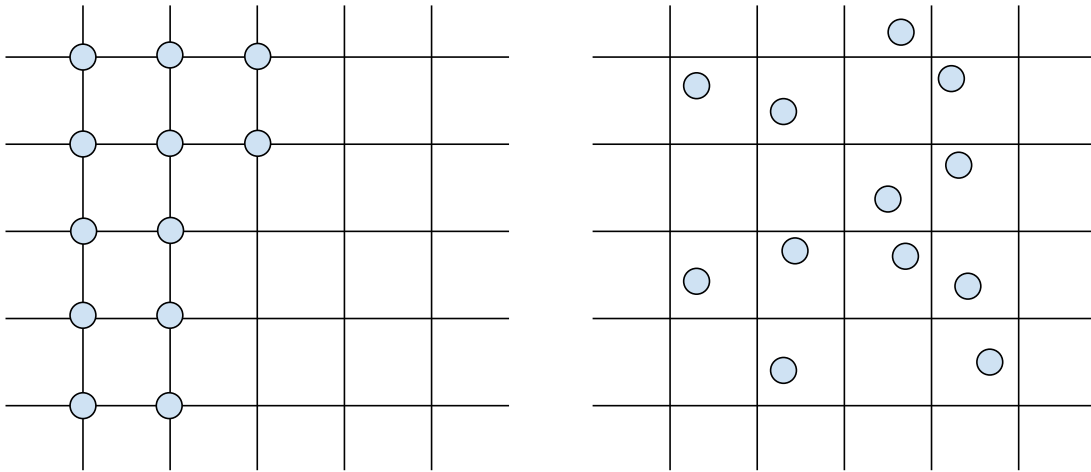


FIGURE 2.5 – Recherche d’hyper paramètres en grille (gauche) et recherche d’hyper paramètres aléatoire (droite)

du type KNN, un seul hyper paramètre doit être déterminé : la valeur de  $k$ . Pour une recherche en grille, les valeurs de 1, 5, 10, 15,  $\dots$ , 50 sont évaluées dans l’ordre.

La recherche aléatoire consiste à échantillonner les hyper paramètres aléatoirement à partir d’une certaine distribution de probabilité. La recherche se termine après un nombre d’itérations fixe ou bien lorsqu’une performance suffisamment élevée est obtenue. Par exemple, pour un classifieur du type SVM, les hyper paramètres  $C$ , le type de noyau, la taille du noyau gaussien et le degré du noyau polynomial doivent être sélectionnés. En utilisant une recherche aléatoire, on utilise une distribution logarithmique pour la taille du noyau gaussien et le paramètre  $C$  et une distribution uniforme pour le degré du noyau polynomial et le type de noyau. La recherche aléatoire offre un avantage par rapport à la recherche en grille, puisqu’elle permet de diminuer le temps nécessaire pour obtenir un modèle aussi performant qu’une recherche en grille [6]. La figure 2.5 illustre les deux procédés. Dans la figure de gauche, la recherche se fait systématiquement et dans l’ordre. Dans la figure de droite, la recherche se fait aléatoirement. Il est intéressant de remarquer que, dans le cas de la recherche aléatoire, les points évalués ne sont pas régulièrement espacés comme dans la recherche en grille.

### Validation croisée

Une fois les hyper paramètres obtenus, ceux-ci doivent être employés pour entraîner un classifieur et évaluer la performance de celui-ci. Puisque le jeu de données de test

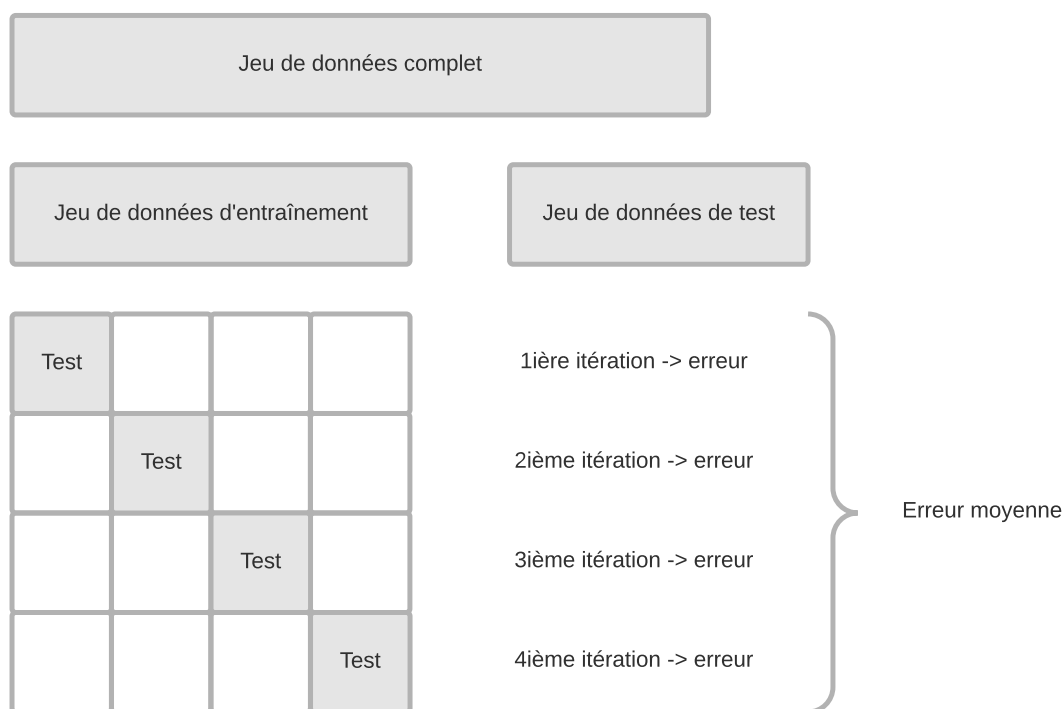


FIGURE 2.6 – Validation croisée à k-plis

ne peut être utilisé, la technique employée consiste à prendre une portion du jeu de données d'entraînement qui sera conservé pour faire l'estimation de la performance sur de nouvelles données. Ce jeu est appelé le jeu de validation. On peut alors estimer la performance sur le jeu de test en utilisant la performance obtenue sur ce jeu de validation comme estimateur.

Pour augmenter la validité de cette estimation, il est possible d'utiliser la validation croisée à k-plis. Cette méthode consiste à séparer le jeu de données d'entraînement en  $k$  plis de tailles égales. On utilise alors ces  $k$  plis pour effectuer  $k$  entraînements. Ces entraînements utilisent  $k - 1$  plis comme données d'entraînement et le pli restant comme jeu de validation. La performance sur le jeu de test peut donc être estimée en faisant la moyenne des performances sur les  $k$  plis. En utilisation la validation croisée, il est possible d'obtenir un estimé de la variance sur ces  $k$  plis. Le meilleur ensemble d'hyperparamètres est celui dont la moyenne des scores de validations des  $k$

plis est la plus élevée. Ce processus est illustré à la figure 2.6. Une fois la combinaison d'hyper paramètres optimale déterminée, on réentraîne le classifieur en utilisant toutes les données d'entraînement.

# Chapitre 3

## Méthode

Le chapitre 1 expose le concept de l'OFS et son influence connue sur les personnes. Le chapitre 2 quant à lui brosse un tableau des algorithmes d'apprentissage automatiques et de leur utilisation dans le cas général.

Dans ce chapitre, la méthodologie employée pour récolter les données et les traitements effectués pour prédire l'OFS sont présentés. Plus précisément, les capteurs de signaux physiologiques et comportementaux, les expériences psychométriques sur ordinateur, la collecte de données avec des participants et la méthodologie d'apprentissage automatique y sont abordés.

### 3.1 Signaux physiologiques et comportementaux

Au chapitre 1, des signaux physiologiques et comportementaux étant associés à l'OFS sont présentés. Ces signaux sont produits naturellement par divers systèmes du corps humain. Par exemple, le rythme cardiaque et le rythme respiratoire sont des signaux physiologiques. Les signaux comportementaux sont des variables qui changent en fonction de notre comportement. Le niveau de fixation du regard est donc un signal comportemental.

Les sections 3.1.1 et 3.1.2 présentent les deux capteurs utilisés pour enregistrer ces signaux lors de la collecte de données. Les caractéristiques d'intérêts pouvant être extraites à partir d'un prétraitement sur les données brutes sont présentées à la section 3.1.3.

#### 3.1.1 Signaux physiologiques

Les signaux physiologiques sont captés à l'aide du BioHarness 3.0 de la compagnie Zephyr [46]. Le BioHarness est une suite de capteurs physiologiques intégrés dans un



FIGURE 3.1 – BioHarness 3.0 de Zephyr

seul dispositif. Ce dispositif est porté par un participant à l'aide d'une bande élastique placée sur la poitrine. Une photo de ce capteur est présentée à la figure 3.1.

Le tableau 3.1 présente les signaux du BioHarness qui sont utilisés.

L'avantage d'utiliser le BioHarness plutôt que plusieurs capteurs séparés est que celui-ci retourne toutes les variables directement consolidées, synchronisées et horodatées dans un seul fichier. En plus des signaux consolidés présentés au tableau 3.1, le BioHarness enregistre le signal brut de l'ECG, des accéléromètres et du capteur de respiration. Cependant, ceux-ci sont échantillonnés à des fréquences différentes et sont donc enregistrés dans des fichiers séparés.

### 3.1.2 Signaux comportementaux

Les signaux oculaires peuvent être utilisés pour l'estimation de l'OFS [23]. Ces valeurs sont captées à l'aide du dispositif de suivi oculaire EyeTribe [81]. Celui-ci fonctionne avec un émetteur et un récepteur infrarouge qui sont dans un boîtier placé directement sous l'écran d'ordinateur. La figure 3.2 illustre l'usage du EyeTribe.

Les signaux captés à l'aide de celui-ci sont présentés dans le tableau 3.2.

TABLE 3.1 – Signaux du BioHarness

Paramètres	Unités	Commentaires
Rythme respiratoire	BPM	
Amplitude de la respiration	LSB	
Confiance sur Rythme respiratoire	%	
Rythme cardiaque	BPM	
Variabilité du rythme cardiaque	ms	
Amplitude ECG	V	
ECG	V	
Confiance ECG	%	
Température du BioHarness	°C	
Température estimée	°C	
Accélération	$m s^{-2}$	Dans 3 axes, min et max
Activité	$m s^{-2}$	Moyenne des accélérations sur une minute
Posture	degree	



FIGURE 3.2 – Représentation de l'utilisation du EyeTribe par un participant. Le EyeTribe est le dispositif placé sous l'écran.

TABLE 3.2 – Signaux du dispositif de suivi oculaire

Paramètres	Unités
Position du regard dans l'écran d'ordinateur en x	pixel
Position du regard dans l'écran d'ordinateur en y	pixel
Rayon de la pupille	pixel
Fixation du regard	booléen



TABLE 3.3 – Valeurs calculées par RHRV

	$f_{min}$	$f_{max}$
ULF	0 Hz	0.03 Hz
VLF	0.03 Hz	0.05 Hz
LF	0.05 Hz	0.15 Hz
HF	0.15 Hz	0.4 Hz

### 3.1.3 Prétraitement

En plus des données brutes récoltées par le BioHarness et le EyeTribe, plusieurs signaux peuvent être calculés à partir de ces données. Ceux-ci sont des variables qui ont été étudiées dans la littérature et dont la relation avec l’OFS a été démontrée.

#### Variabilité du rythme cardiaque

La variabilité du rythme cardiaque [8] est présentée dans la section 1.1.2. Celle-ci est calculée automatiquement par le BioHarness. En plus de la HRV, l’analyse du contenu spectral du signal cardiaque donné par l’ECG est effectuée. Il existe plusieurs façons de produire celle-ci. Dans le cadre du projet, l’analyse fréquentielle de la HRV est calculée à l’aide de la densité spectrale de puissance (power spectral density, PSD) [83].

Le calcul des différentes fréquences est effectué à l’aide du paquet RHRV [34], disponible dans le langage de programmation R. Plusieurs bandes de fréquences sont utilisées. Celles-ci sont présentées dans le tableau 3.3. Les fréquences ultras basses (ultra low frequencies, ULF), les fréquences très basses (very low frequencies, VLF), les basses fréquences (low frequencies, LF) et les hautes fréquences (high frequencies, HF) sont calculées. De plus, la valeur du quotient entre les LF et les HF est calculée. Une fenêtre de 300 s est utilisée. Le calcul du contenu fréquentiel de l’ECG produit donc un vecteur contenant une valeur pour chaque bande de fréquences (ULF à HF) 5 minutes à partir du moment où le participant porte le BioHarness. Cette valeur est calculée pour chaque seconde.

#### Signaux oculaires

Le dispositif de suivi oculaire permet d’obtenir la position du regard de l’utilisateur par rapport à l’écran. Cette caractéristique peut être utilisée pour estimer l’OFS. Cependant, elle ne peut pas être utilisée directement dans un contexte avec plusieurs tâches. En effet, cette caractéristique est fortement liée à la tâche en cours d’exécution et les informations

TABLE 3.4 – Signaux calculés

Paramètres	Unités
Proportion du temps en saccade (yeux)	
Détection du clignement des yeux	
Durée du clignement des yeux	s
Fréquence de clignement des yeux	Hz

apportées par celle-ci ne se généralisent pas à plusieurs tâches.

Pour limiter cette faiblesse des valeurs brutes des dispositifs de suivi oculaire, plusieurs caractéristiques d'intérêt pouvant être calculées à partir des signaux bruts [23, 72]. Ces caractéristiques sont présentées au tableau 3.4. Ils sont également décrits dans les paragraphes suivants.

**Proportion du temps en saccade** Le EyeTribe retourne un champ dans les données brutes indiquant si le regard est une fixation ou non. La proportion des valeurs en fixation pendant une seconde est calculée.

**Clignement des yeux** Les clignements des yeux sont calculés à partir des signaux du EyeTribe. Un clignement d'œil est défini comme un ensemble de données manquantes d'une durée de 10 ms à 1.5 s. Les données manquantes plus courtes et plus longues sont considérées comme étant du bruit dans le dispositif de suivi oculaire. La moyenne de temps en clignement est ensuite calculée.

### 3.1.4 Signaux utilisés

Le tableau B.1 présenté à l'annexe B présente la liste complète des signaux utilisés lors des expériences d'apprentissage automatique.

## 3.2 Collecte de données

Pour améliorer la capacité des modèles d'apprentissage automatique à généraliser sur de nouvelles tâches, trois tâches expérimentales ont été utilisées lors de la collecte de données. Celles-ci sont des tâches cognitives classiques. Elles sont implémentées en utilisant le cadriciel Psychopy [67, 68]. Ce cadriciel permet la présentation de stimuli visuels, sonores et le chronométrage des temps de réaction.

### 3.2.1 Conditions expérimentales

Pour faire un apprentissage automatique permettant d'estimer l'état fonctionnel de l'opérateur, il est nécessaire d'avoir des données correspondant à un état fonctionnel et non fonctionnel. Pour s'assurer d'avoir une variabilité dans les données, des manipulations ont été effectuées. Deux manipulations en lien avec les dimensions de l'OFS, qui sont présentées à la section 1.1.1, sont effectuées lors des tâches expérimentales : les expérimentations se déroulent à deux moments de la journée séparés, et la charge de travail de chaque tâche est manipulée pendant l'expérience.

#### Moment de la journée

Comme mentionné au chapitre 1, la fatigue est liée à l'état fonctionnel de l'opérateur [32]. Pour influencer l'OFS des participants ainsi que l'état physiologique des participants, la collecte de données s'est donc déroulée en deux sessions distinctes pour chaque participant. Une session est effectuée en matinée et une autre session est effectuée en fin d'après-midi. Les deux sessions sont effectuées la même journée.

Puisque la fatigue est mal définie dans la littérature, l'effet attendu de cette manipulation est simplement d'augmenter la variabilité dans les données physiologiques, comportementales et dans les performances des participants.

#### Charge de travail

La charge de travail est fortement associée à l'OFS [26, 41, 90]. Plusieurs auteurs utilisent directement la charge de travail comme valeur cible pour estimer l'OFS.

Pour manipuler la charge de travail, chaque tâche a été implémentée avec deux niveaux de difficulté. Étant donné que la façon de varier ce niveau est différente pour chaque tâche, une description de celui-ci sera présentée indépendamment pour chacune des tâches.

### 3.2.2 Tâches

Pour la collecte de données, une batterie comprenant trois tâches informatisées a été créée à partir de tâches expérimentales connues et le cadriciel Psychopy [67, 68]. Les trois tâches sont le n-back, une tâche de rotation mentale et une tâche de recherche visuelle.

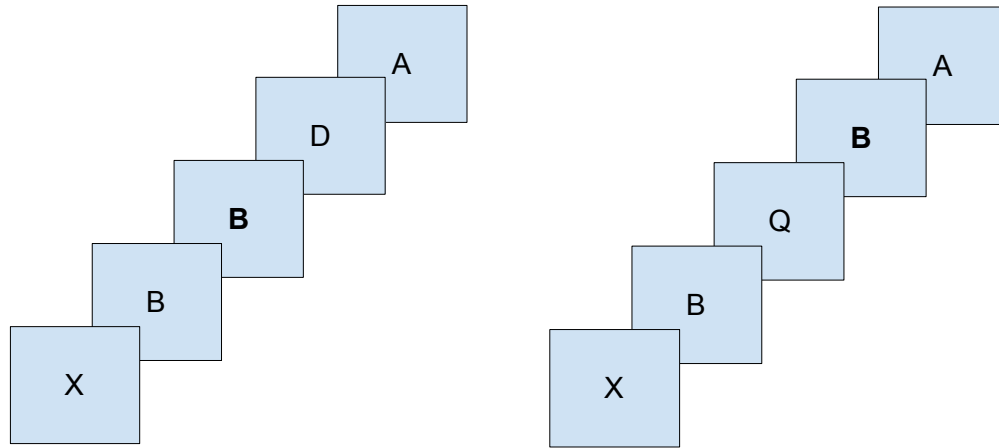


FIGURE 3.3 – Représentation de la tâche de n-back. Dans la figure de gauche, un 1-back est effectué. La lettre B est répétée. Dans la figure de droite, un 2-back est effectué. Le B est encore une fois répété. Les stimuli positifs sont en gras.

Dans celles-ci, des stimuli sont présentés aux participants. Le participant doit alors fournir une réponse à l'aide du clavier d'ordinateur ou de la souris. Le temps entre la présentation des stimuli et la réponse du participant est mesuré.

### N-back

La tâche du n-back [12, 54] est utilisée pour évaluer la mémoire de travail. Lors de cette tâche, une séquence de stimuli est présentée au participant, un stimulus à la fois. Le participant garde le stimulus en mémoire alors que de nouveaux stimuli lui sont présentés. La tâche du participant consiste à décider si le stimulus présenté est le même que celui  $n$  itérations précédentes.

Dans l'implémentation utilisée, les stimuli correspondent à des lettres et sont affichés sur un écran d'ordinateur. Chaque stimulus est présenté pour une durée de 2 s. Les stimuli sont séparés par une pause de 0.5 s où rien n'est affiché à l'écran. Le participant indique, à l'aide du clavier, si la lettre est la même que celle présentée  $n$  lettre auparavant ou une lettre différente. Les stimuli sont générés aléatoirement et sont différents pour chaque participant. La proportion de stimuli positifs, c'est à dire de lettres identiques à la lettre  $n$  itérations auparavant est de 50%. La figure 3.3 représente la tâche de n-back.

La charge de travail a été manipulée en changeant la valeur de  $n$ . La valeur de  $n$  est de 1 pour la charge de travail faible et de 2 pour la charge de travail élevée. Le but primaire

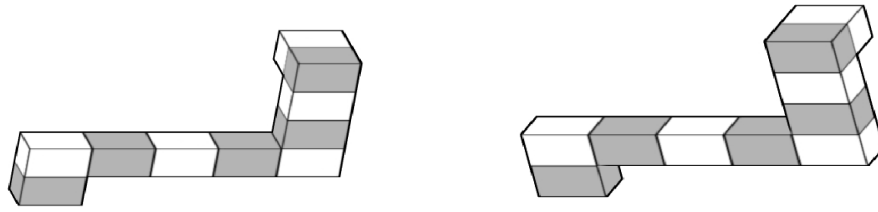


FIGURE 3.4 – Une paire de formes de la tâche de rotations mentales. Une rotation est effectuée.

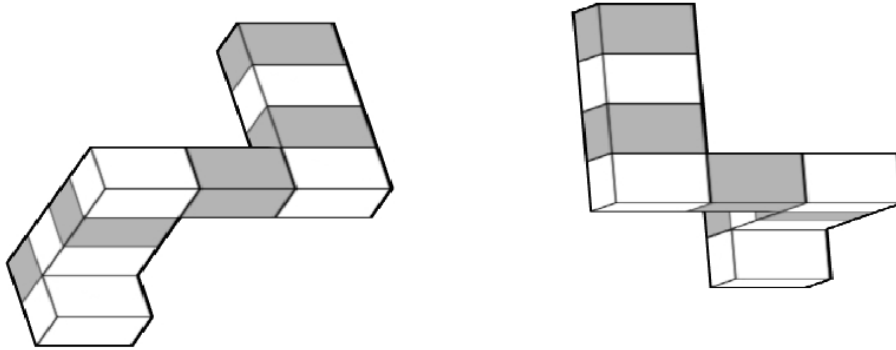


FIGURE 3.5 – Une paire de formes de la tâche de rotations mentales. Une rotation et une symétrie sont effectuées.

du participant dans cette tâche est l'exactitude de la réponse. Le but secondaire est de fournir celle-ci le plus rapidement possible. L'exactitude de la réponse et le temps mesuré entre la présentation des stimuli et la réponse sont enregistrés.

### Rotation mentale

La tâche de rotation mentale [16] est utilisée pour évaluer la perception spatiale. Dans cette tâche, deux formes en trois dimensions sont présentées simultanément au participant. Celui-ci doit indiquer s'il s'agit de la même forme à laquelle on a appliqué une rotation, ou s'il s'agit d'une opération de symétrie et de rotation. La figure 3.4 présente un exemple de symétrie et de rotation et la figure 3.5 présente un exemple de la même forme ayant subi seulement une rotation. Selon Cooper [16], le temps de réaction d'un participant à la tâche de rotation mentale est proportionnel au niveau de rotation entre les deux figures.

Les stimuli sont tirés de Peters et Battista [69]. Pour chaque paire de stimuli, le type de formes (avec ou sans symétrie) est choisi aléatoirement. Les stimuli sont présentés à l'écran jusqu'à ce que le participant entre une réponse à l'aide du clavier.

La charge de travail a été manipulée en changeant le niveau de rotation entre les figures. Dans la tâche avec un niveau de charge de travail faible, les figures sont en rotation de  $5^\circ$  à  $40^\circ$  les unes par rapport à l'autre. Dans la tâche avec un niveau de charge de travail élevé, les figures sont en rotation de  $125^\circ$  à  $170^\circ$  les unes par rapport à l'autre. Deux mesures de performances, l'exactitude du résultat et le temps entre l'affichage du stimulus et la réponse du participant sont enregistrées.

## Recherche visuelle

Comme son nom l'indique, la tâche de recherche visuelle est utilisée pour évaluer la capacité à rechercher un élément à l'aide de la vue. Elle est une adaptation de la tâche décrite dans Mocharnuk [58]. Dans celle-ci, le participant doit identifier une cible dans un écran rempli de distractions en utilisant la souris d'ordinateur. La cible et les distractions sont des lettres. Dans la condition de charge de travail faible, le participant doit trouver la lettre A. Dans la condition de charge de travail élevée, le participant doit trouver la voyelle qui n'a pas subi de rotation. Des exemples de la charge de travail faible et élevée sont illustrés respectivement dans le haut et le bas de la figure 3.6 .

Comme il a été mentionné plus tôt, le participant indique sa réponse en cliquant sur la cible à l'aide de la souris d'ordinateur. L'écran reste affiché tant que le participant ne clique pas sur la cible, s'il clique ailleurs, l'essai continue. Le temps de réaction est enregistré.

### 3.2.3 Déroulement

Le déroulement de la passation des participants s'est effectué en suivant la procédure présentée dans cette section. Chaque participant a été rencontré individuellement au cours de l'hiver 2015. Ceux-ci étaient rencontrés dans un local du laboratoire de vision et systèmes numériques (LVSN) de l'Université Laval le matin et en fin d'après-midi. Les participants ont ensuite fourni leur consentement à l'expérience en remplissant un formulaire explicatif. À l'exception de ce formulaire, les rencontres du matin et de l'après-midi sont identiques.

Une fois le formulaire de consentement rempli, le participant est équipé du BioHarness. La calibration du dispositif de suivi oculaire est ensuite effectuée.

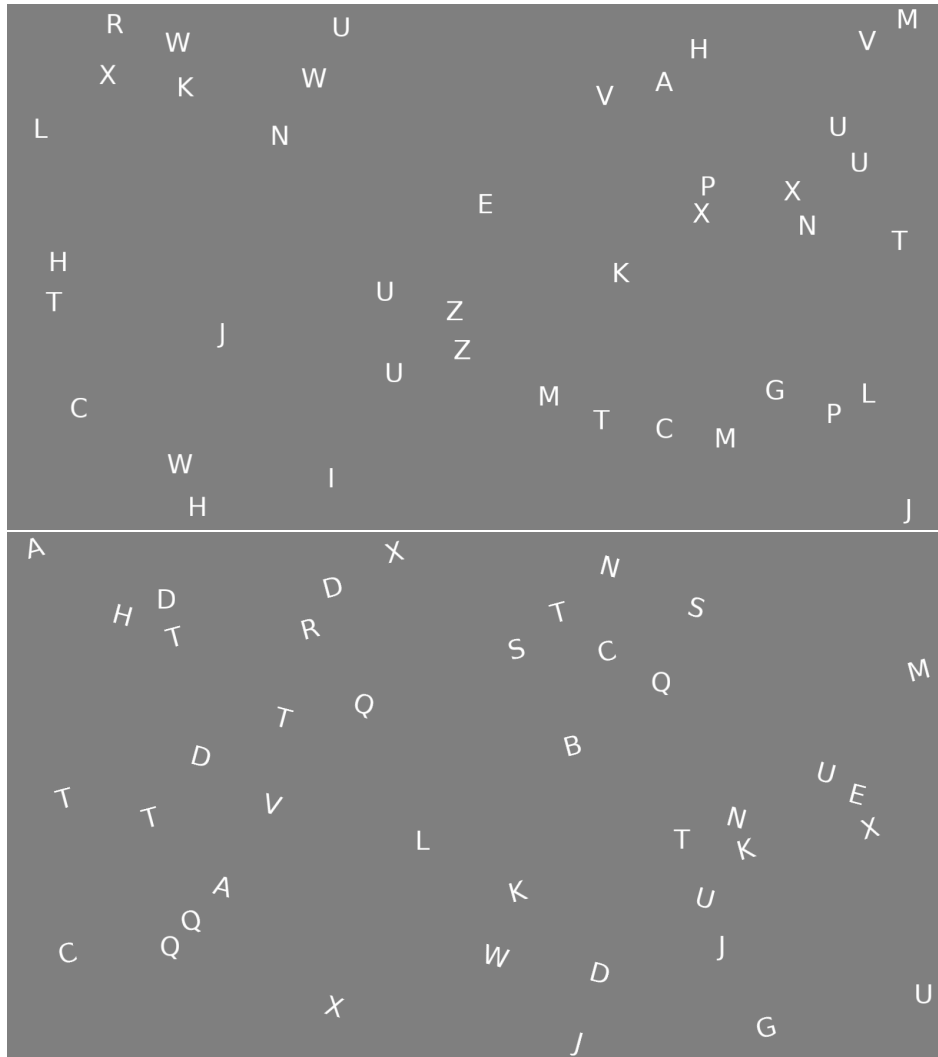


FIGURE 3.6 – Deux exemples de la tâche de recherche visuelle. La partie du haut représente la charge de travail faible et la partie du bas correspond à la charge de travail élevée. Dans la partie du haut, la lettre cible est la lettre "A" et se trouve à droite de la partie supérieure. Dans la partie du bas, la lettre cible est le "U" et se trouve à droite de la partie inférieure de l'écran.

Pour chaque tâche expérimentale, les instructions sont données par écrit aux participants. Une session de pratique d'environ deux minutes est effectuée pour chaque condition expérimentale. L'ordre des tâches effectuées est choisi aléatoirement. Quatre participants ont commencé par la condition de charge de travail faible. Les quatre autres ont commencé par la charge de travail élevée. La session en matinée et la session en fin d'après-midi sont toujours effectuées le même jour. L'ordre des tâches et des conditions expérimentales est le même pour la session du matin et du soir.

### 3.3 Définition de la valeur cible

Au chapitre 1, un des objectifs de recherche consistait à formaliser l'OFS dans un contexte d'apprentissage automatique. Au chapitre 2, les composantes nécessaires pour considérer l'estimation de l'OFS comme un problème d'apprentissage automatique sont présentées.

Le premier élément nécessaire pour faire un apprentissage automatique est le choix des valeurs fournies en entrée. Dans ce cas, ces valeurs correspondent aux signaux de la section 3.1.

Le deuxième élément nécessaire à l'apprentissage supervisé est le choix des valeurs de sorties. Dans le problème de l'estimation de l'état fonctionnel de l'opérateur, ces valeurs correspondent à l'estimation de cet état. Il s'agit de ce qu'on appelle la vérité terrain (*ground truth*).

Cependant, comme le montre le modèle présenté au chapitre 1, l'état fonctionnel de l'opérateur est un état qui ne peut pas être mesuré directement. Deux mesures sont employées dans la littérature pour fournir une vérité terrain lors de l'utilisation d'algorithmes d'apprentissage automatique pour l'estimation de l'état fonctionnel de l'opérateur. Certains auteurs prennent la charge de travail de la tâche effectuée comme mesure de l'état fonctionnel [24, 89, 90]. D'autres auteurs utilisent la mesure de la performance du participant dans la tâche courante [86, 92].

Pour obtenir un ensemble de valeurs permettant d'avoir une estimation de l'OFS, les différentes tâches présentées dans la section 3.2.2 comportent une mesure de la réalisation du but secondaire.



### 3.3.1 Performance Dynamique Décontextualisée

Pour utiliser ces mesures provenant de différentes tâches et plusieurs participants, une mesure de l'OFS créée par l'auteur de ce mémoire et pouvant être utilisée comme vérité terrain, la performance dynamique décontextualisée (decontextualized dynamique performance, DDP), est décrite dans ce chapitre [31]. La DDP représente une mesure de la performance qui est décontextualisée, c'est-à-dire qui permet de comparer la performance de sujets différents et de tâches différentes. Elle rend donc possible l'utilisation d'algorithmes d'apprentissage automatique pour plusieurs tâches simultanément.

Pour respecter le modèle de l'état fonctionnel de l'opérateur présenté au chapitre 1 ainsi que pour être utile comme signal d'apprentissage automatique, la DDP a certaines propriétés.

Premièrement, la métrique doit pouvoir être utilisée pour plusieurs participants et plusieurs tâches simultanément. Par exemple, certains sont naturellement meilleurs dans une tâche que d'autres. Il faut donc que la mesure puisse indiquer la performance d'un participant relativement à lui-même pour indiquer si celui-ci est dans un état fonctionnel ou non. De même, la performance dans les différentes tâches est fondamentalement différente. Pour étudier la généralisation des algorithmes d'apprentissage automatique à plusieurs tâches et à plusieurs participants, la vérité terrain doit donc pouvoir permettre d'obtenir une valeur indépendamment de la tâche ou du participant.

Deuxièmement, puisque l'état fonctionnel de l'opérateur est un concept qui ne varie pas rapidement, la mesure utilisée devrait également être stable dans le temps. En effet, bien que l'état fonctionnel de l'opérateur puisse varier dans le temps, cet état ne doit pas changer à chaque essai. De plus, même si les tâches ont été choisies pour leur stabilité dans la charge de travail, la variation des performances dans celles-ci n'est pas négligeable.

Troisièmement, la DDP est une mesure binaire. Celle-ci indique si l'opérateur est dans un état fonctionnel ou non.

### 3.3.2 Procédure

La DDP consiste à transformer des valeurs provenant de mesures de performances secondaires obtenues dans différentes tâches et de les transformer en une valeur binaire indiquant l'état fonctionnel de l'opérateur. Ces données de performances obtenues lors de la prise de données ont la forme d'une suite de tuple contenant le temps de réaction

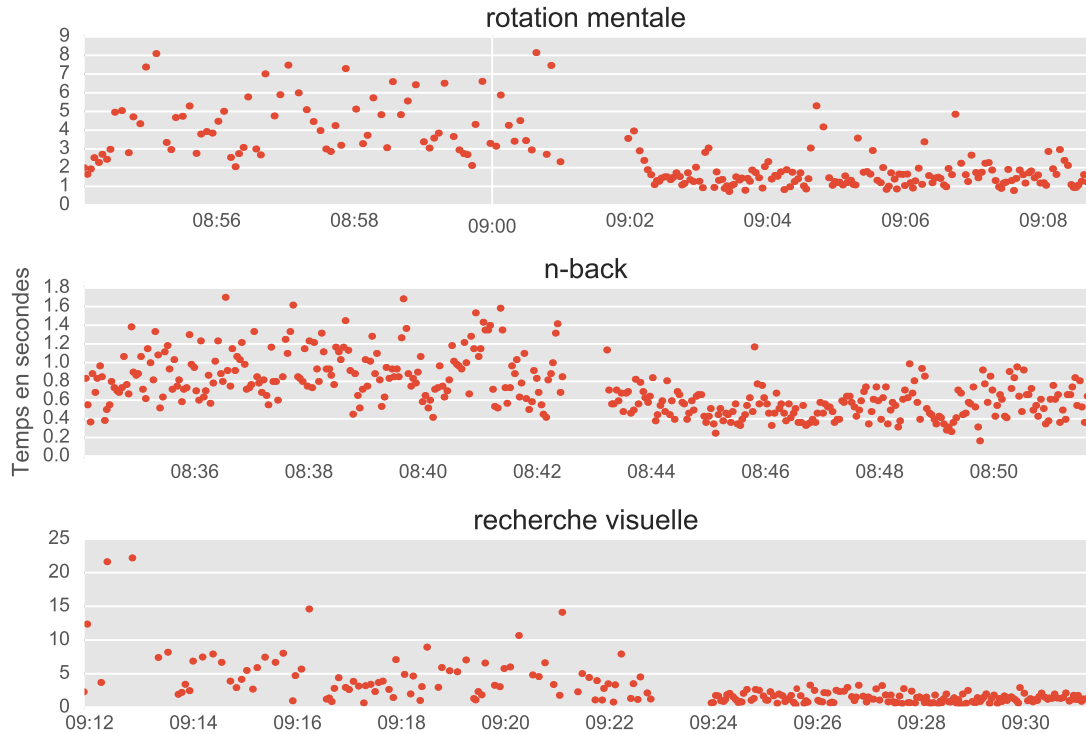


FIGURE 3.7 – Temps de réaction d’une session expérimentale. La charge de travail était élevée à gauche et faible à droite.

pour un essai de la tâche en cours et un horodatage indiquant le moment auquel l’essai s’est terminé. Une fois la procédure de la DDP appliquée, on obtient une suite de valeurs à fréquence fixe indiquant l’OFS estimé correspondant à ces valeurs de performances. Un exemple de temps de réaction est présenté à la figure 3.7. Cette figure présente les temps de réaction obtenus par un participant lors de la collecte de données. Il est intéressant de remarquer que les valeurs obtenues dans les différentes tâches n’ont pas la même échelle temporelle. En effet, les temps de réaction dans la tâche de rotation mentale se situent entre 0.5 s et 9 s alors que ceux du n-back se situent entre 0.25 s et 2 s.

Puisque le calcul de la DDP tient pour acquis que l’OFS ne change pas rapidement, la première étape du calcul de la DDP consiste à appliquer un filtre sur les données de performances. Différentes techniques de filtrage pouvant être utilisées pour le calcul de la DDP sont présentées dans la section suivante.

Une fois les données filtrées, celles-ci sont ensuite transformées en valeurs binaires à l’aide d’une méthode de seuillage. La valeur de ce seuil est déterminée indépendamment pour chaque tâche et chaque participant. Les techniques de seuillage pouvant être utilisées

sont présentées à la section 3.3.4.

### 3.3.3 Filtrage

Tel que discuté précédemment, la DDP représente un état qui change lentement dans le temps. Cependant, les données de performances qui sont recueillies lors de la collecte de données sont bruitées. Un exemple de ces données de performances est présenté à la figure 3.7. Afin d'éliminer les hautes fréquences contenues dans les temps de réaction, il est nécessaire de filtrer les temps de réaction reçus en entrée.

Tous les filtres présentés dans cette section sont à fenêtres glissantes, c'est-à-dire que ceux-ci utilisent une fenêtre temporelle avant le pas de temps actuel. Dans tous les cas, la procédure est la même, on commence par échantillonner les signaux à une fréquence fixe de 1 Hz. Les valeurs manquantes dans le vecteur sont laissées comme telles, c'est-à-dire qu'elles ont comme valeur la mention nombre invalide (Not a Number, NaN). Les fonctions de filtrage utilisées ne tiennent pas compte des valeurs manquantes et ajustent les équations en conséquence.

Le choix du filtre utilisé pour le filtrage des temps de réponse dépend des propriétés qui sont recherchées. En effet, dans le calcul de la DDP, le filtre a pour but d'atténuer les variations de performances ponctuelles.

La procédure de filtrage prend en entrée une séquence de temps de réaction  $\mathbf{x}$  et produit en sortie une séquence  $\mathbf{p}$  correspondant aux temps de réactions filtrées. La fréquence du signal de sortie est de 1 Hz.

#### Filtre moyenneur à fenêtre glissante

Le filtre moyenneur [57] consiste à appliquer la moyenne des  $M$  dernières valeurs avant la valeur courante. Il est calculé grâce à l'équation suivante :

$$p[n] = \frac{1}{N} \sum_{\ell=0}^{M-1} x[n - \ell], \quad (3.1)$$

où la constante  $N$  représente le nombre de points qui ne sont pas NaN et la constante  $M$  représente la taille en seconde du filtre. Les valeurs NaN sont ignorées.

La figure 3.8 présente la filtration du signal de la figure 3.7 en utilisant plusieurs tailles de filtre. On peut remarquer dans cette figure que la taille du filtre utilisé a une plus grande influence pour les valeurs correspondant à la charge de travail élevée que pour la charge de travail faible.

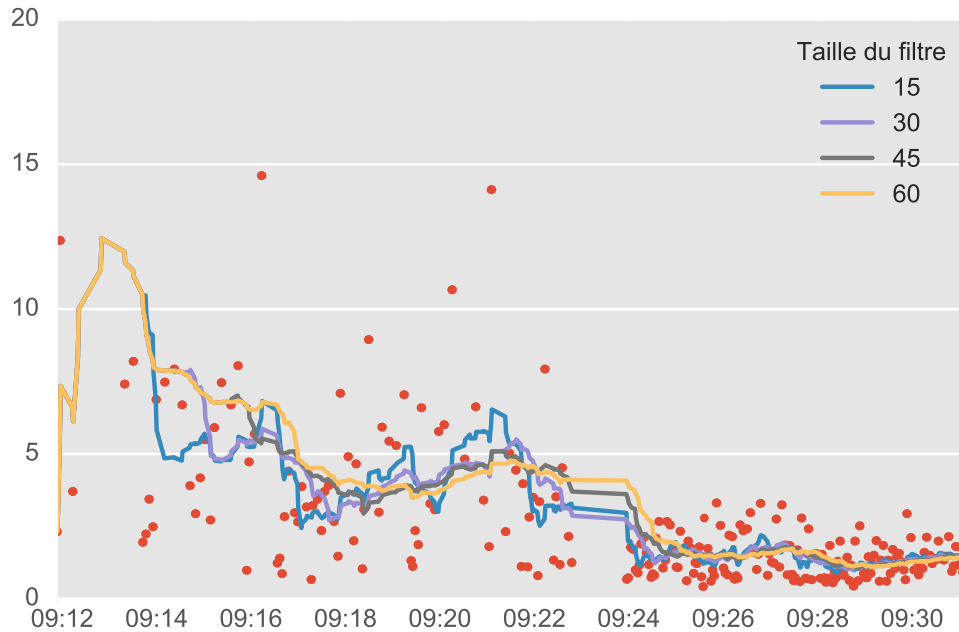


FIGURE 3.8 – Filtrage à l’aide d’un filtre moyenneur à fenêtre glissante sur les performances d’un participant dans la tâche de recherche visuelle. La charge de travail élevée est à gauche.

Le filtre moyenneur permet d’avoir un échantillonnage à la seconde des données. Par contre, les valeurs extrêmes peuvent avoir une grande influence sur les données en sorties.

### Filtre exponentiel à fenêtre glissante

Le filtre exponentiel à fenêtre glissante [57] est calculé à l’aide de l’équation suivante :

$$p[n] = \alpha p[n - 1] + x[n], 0 < \alpha < 1 \quad (3.2)$$

où  $\alpha$  correspond à une constante représentant l’influence des données précédentes.

Dans le cas des données manquantes, celles-ci sont tout simplement ignorées et traitées comme si elles étaient égales à 0.

La figure 3.9 présente la filtration du signal de la figure 3.7 en utilisant plusieurs tailles de filtre.

Dans le filtre exponentiel à fenêtre glissante, l’importance des données qui sont plus loin dans le temps est atténuée. Par contre, les valeurs extrêmes peuvent avoir une grande

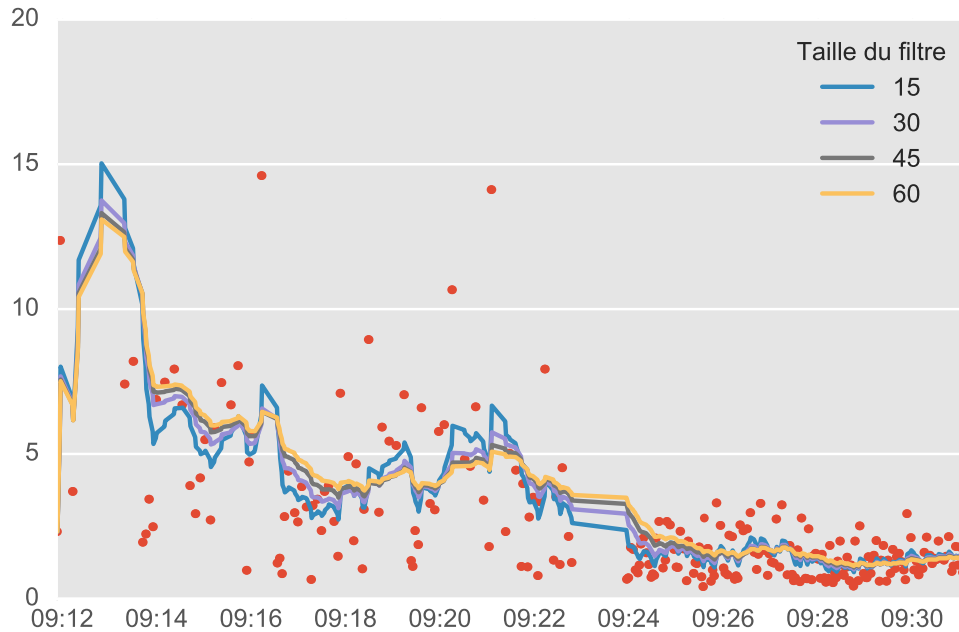


FIGURE 3.9 – Filtrage à l’aide d’un filtre exponentiel à fenêtre glissante sur les performances d’un participant dans la tâche de recherche visuelle. La charge de travail élevée est à gauche.

influence sur les données en sorties.

### Filtre médian à fenêtre glissante

Le filtre médian [57] consiste à calculer la médiane des valeurs valides dans les  $M$  dernières secondes. On peut le calculer grâce cette équation :

$$p[n] = \text{med} \{x[n - K], \dots, x[n - 1], x[n], x[n + 1], \dots, x[n + K]\} \quad (3.3)$$

La figure 3.10 présente la filtration du signal de la figure 3.7 en utilisant plusieurs tailles de filtre.

Le filtre exponentiel à fenêtre glissante permet d’avoir un échantillonnage à la seconde de nos données. De plus, l’importance des valeurs extrêmes ponctuelles est atténuée.

### Comparaison des filtres

La figure 3.11 présente une comparaison des différents types de filtres. On peut remarquer

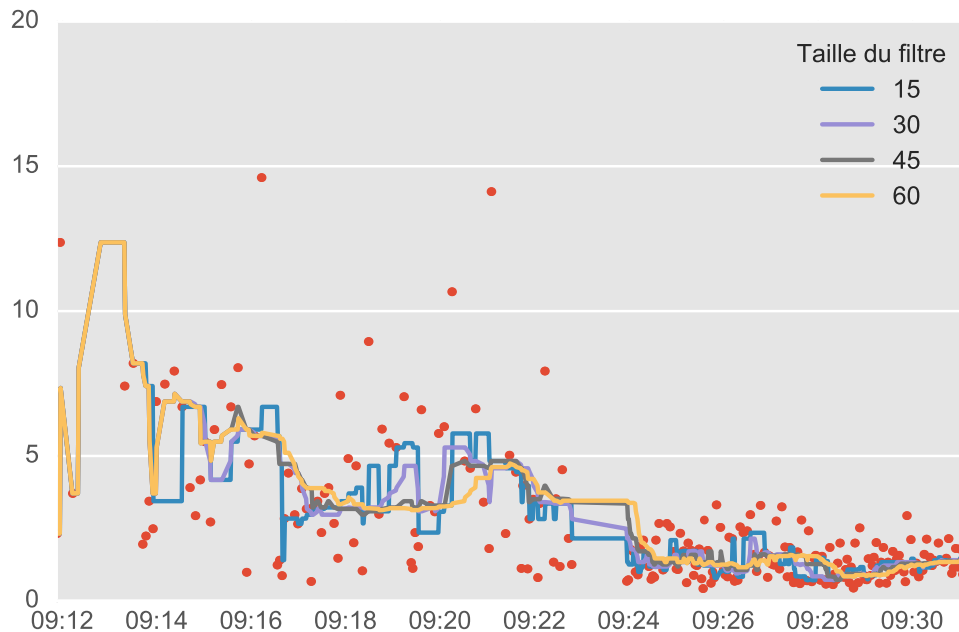


FIGURE 3.10 – Filtrage à l’aide d’un filtre médian à fenêtre glissante sur les performances d’un participant dans la tâche de recherche visuelle. La charge de travail élevée est à gauche.

que le type de filtre change la forme des valeurs obtenues en sorties. Cependant, étant donné que ces valeurs seront seuillées en utilisant les techniques présentées à la section suivante, la forme du filtre utilisé a une importance limitée.

### 3.3.4 Seuillage

L’étape du seuillage est effectuée directement après avoir filtré les données. Dans le contexte de la DDP, le seuillage consiste à choisir une valeur de performance en dessous de laquelle le participant doit être considéré comme non fonctionnel. La méthode de seuillage reçoit en entrée le vecteur  $\mathbf{p}$ , c’est-à-dire les valeurs de performances qui ont été filtrées à l’étape précédente. Celles-ci sont déjà filtrées à l’aide d’une des techniques de la section 3.3.3.

L’étape de seuillage permet de créer un modèle parcimonieux de l’état fonctionnel. En effet, en discretisant la cible en deux niveaux, la DDP pose l’hypothèse qu’il existe une variation d’au moins deux niveaux de l’état fonctionnel de l’opérateur. Le seuil trouvé correspond alors à un niveau de performance qui, pour un participant et une tâche

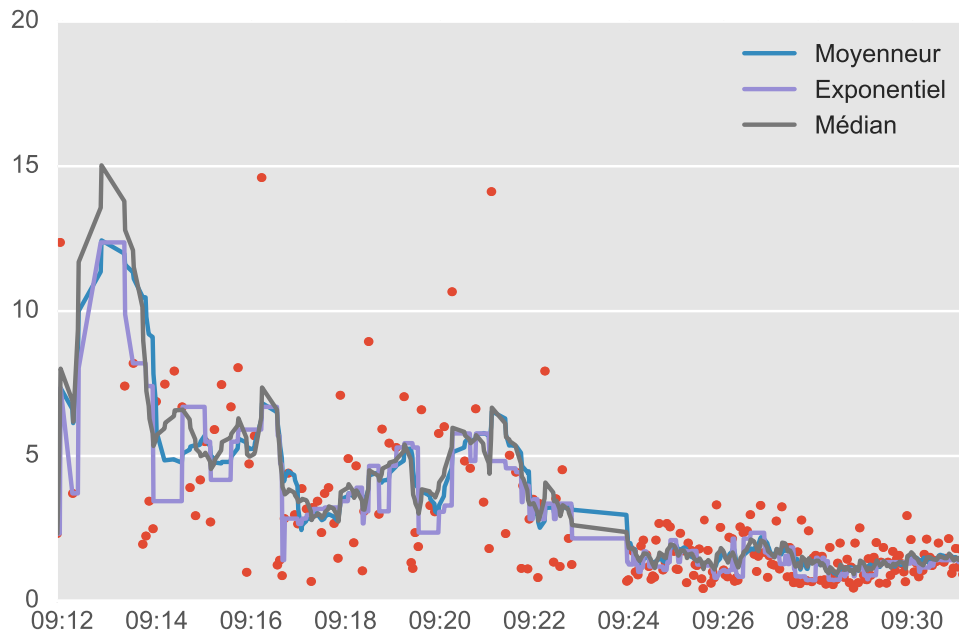


FIGURE 3.11 – Comparaison des types de filtres sur les performances d’un participant dans la tâche de recherche visuelle.

donnée, est indicateur d’un état sous fonctionnel.

Deux approches peuvent être utilisées pour choisir un seuil. Il est possible de choisir un seuil avec une valeur fixe. Il est également possible d’utiliser des techniques de seuillage adaptatif.

### Valeur fixe

Le seuillage à l’aide d’une valeur fixe consiste à utiliser une valeur prédéfinie de performance en dessous de laquelle l’état du participant est considéré comme non fonctionnel. Pour que cette technique puisse produire des comparaisons valables entre les tâches et les participants, ce seuil est déterminé indépendamment pour chaque tâche et chaque participant. En effet, les performances des différentes tâches ne sont pas dans le même ordre de grandeur et peuvent même être des quantités différentes. De plus, chaque participant peut avoir un niveau de performance de base qui rend le choix d’un seuil de performance absolue plus difficile.

Deux techniques sont présentées ici pour permettre d’utiliser un seuil fixe dans le choix

de l'état fonctionnel : l'utilisation d'un score- $\mathcal{Z}$  [47], et l'utilisation d'un percentile [75].

**Score- $\mathcal{Z}$**  Le score- $\mathcal{Z}$  où la variable centrée réduite permet d'obtenir une distribution des performances avec une moyenne nulle et un écart-type avec une valeur fixe. Pour chaque point  $i$  du vecteur  $\mathbf{p}$ , on peut calculer le score- $\mathcal{Z}$  équivalent à l'aide de la formule suivante :

$$\text{score-}\mathcal{Z}_i = \frac{p_i - \mu}{\sigma}, \quad (3.4)$$

où  $\mu$  représente la moyenne pour la tâche et le participant et  $\sigma$  représente l'écart-type pour la tâche et le participant.

En utilisant :

$$s = Z_s * \sigma + \mu, \quad (3.5)$$

on peut calculer la valeur de seuil  $s$  correspondant à un seuil  $Z_s$  d'une valeur fixe. Par exemple, dans la situation où une performance qui se situe à plus d'une fois et demie l'écart-type de la moyenne d'un participant dans une certaine tâche, on peut utiliser les équations précédentes pour obtenir si chacune des valeurs du vecteur  $\mathbf{p}$  correspond à un état fonctionnel.

**Percentile** L'utilisation du percentile pour trouver un seuil consiste à choisir comme valeur seuil  $S$ , celle qui correspond au  $p_s$ -ième percentile. C'est-à-dire que celle-ci est la valeur catégorisant  $(1 - p_s)$  exemple comme non fonctionnel.

L'utilisation du percentile se fait donc de la façon suivante : pour un participant et une tâche donnée, une valeur du vecteur  $\mathbf{p}$  correspond à un état fonctionnel si elle se trouve sous une valeur de percentile donnée.

### Seuillage adaptatif

Les méthodes présentées dans les sections précédentes ont l'avantage d'être simples à implémenter et à utiliser. Cependant, la détermination d'une valeur fixe correspondant au seuil entre l'état fonctionnel et sous-fonctionnel est problématique. En effet, le modèle de l'OFS ne détermine pas de valeur permettant de séparer les performances obtenues à partir d'un seuil. De plus, une valeur de seuil valide pour une tâche n'est pas nécessairement valide pour une autre.

Cependant, la problématique de seuillage a déjà été abordée dans d'autres domaines, notamment dans le traitement d'image. Dans Sezgin et Sankur [79], on rapporte plusieurs



techniques pour trouver un seuil adaptatif permettant séparer l'avant-plan de l'arrière-plan d'une image. En supposant que la catégorisation de l'OFS correspond à trouver l'avant-plan et l'arrière-plan d'un signal, il est donc possible d'utiliser les techniques présentés dans cet article. Trois méthodes permettant de séparer les valeurs du vecteur  $\mathbf{p}$  sont présentées dans cette section :

- la méthode d'Otsu [62] ;
- l'entropie minimale de Li [52] ;
- la méthode de Yen [93].

**Méthode d'Otsu** La méthode d'Otsu [62] est utilisée pour séparer un signal en deux classes. La méthode consiste à minimiser la somme pondérée de la variance intra-classe des classes d'avant-plan et d'arrière-plan.

Le seuil optimal de la méthode d'Otsu peut être obtenu grâce cette équation :

$$T_{opt} = \arg \max \left\{ \frac{P(T)[1 - P(T)][m_f(T) - m_b(T)]^2}{P(T)\sigma_f^2(T) + [1 - P(T)]\sigma_b^2(T)} \right\} \quad (3.6)$$

Dans celle-ci,  $T_{opt}$  correspond au seuil optimal,  $P(T)$  correspond à la quantité normalisée de valeur ayant l'intensité  $T$ ,  $m_f(T)$ ,  $m_b(T)$ ,  $\sigma_f^2(T)$  et  $\sigma_b^2(T)$  correspondent à la moyenne et la variance de la classe d'avant-plan et d'arrière-plan si la valeur de seuil est égale à  $T$ .

**Entropie minimale de Li** Dans Li et Lee [52], une méthode de seuillage automatique basée sur l'entropie croisée est présentée. La technique consiste à minimiser l'entropie entre le signal binarisé et le signal d'origine. Le seuil optimal peut être obtenu grâce cette équation :

$$T_{opt} = \arg \min \left\{ \sum_{P(i) < T} P(i) \log \left( \frac{P(i)}{m_f(t)} \right) + \sum_{P(i) > T} P(i) \log \left( \frac{P(i)}{m_b(t)} \right) \right\} \quad (3.7)$$

**Méthode de Yen** La méthode de Yen [93] consiste à maximiser la somme de l'entropie du signal d'avant-plan et d'arrière-plan. Le seuil optimal peut être obtenu grâce à l'équation suivante :

$$T_{opt} = \arg \max \left\{ -\log \left\{ \sum_{g=0}^T \left[ \frac{p(g)}{P(T)} \right]^2 \right\} - \log \left\{ \sum_{g=T+1}^G \left[ \frac{p(g)}{1 - P(T)} \right]^2 \right\} \right\} \quad (3.8)$$

## Comparaison des méthodes de seuillage

Les techniques présentées dans cette section permettent de séparer les valeurs obtenues suite à la filtration en deux groupes : les valeurs correspondant à un état fonctionnel et les valeurs associées à un état sous-fonctionnel. Cette séparation en au moins deux classes est nécessaire pour utiliser les algorithmes d'apprentissage automatique de classification.

La figure 3.12 présente le résultat des techniques de seuillage présenté dans cette section. Celle-ci montre les histogrammes des temps de réaction d'un participant sur les trois tâches de l'expérience en fonction du niveau de charge de travail. Bien que la figure montre que le seuil obtenu est près du croisement de la densité des distributions des temps de réaction en fonction du niveau de charge de travail, il est important de garder en tête que le but de ce seuillage n'est pas d'identifier le niveau de charge de travail, mais les situations où la performance de l'opérateur est dégradé.

### 3.3.5 Choix du filtre et du seuil

Plusieurs paramètres ayant une influence dans le résultat de la DDP ont été présentés au cours de cette section. Cependant, peu de justifications théoriques permettant de faire le choix de ceux-ci sont présentées. En effet, puisque l'OFS est un modèle conceptuel de l'état d'un opérateur, il n'y a pas d'équivalence objective permettant d'évaluer la validité de la DDP. Cependant, une étude des effets des différents paramètres de la DDP peut permettre d'isoler des valeurs permettant d'obtenir des caractéristiques souhaitables pour l'apprentissage automatique.

Sur la figure 3.12, on peut voir que le seuil trouvé par la méthode d'Otsu sépare l'histogramme d'une façon débalancé. Tel qu'indiqué dans Sezgin et Sankur [79], la méthode d'Otsu donne des résultats satisfaisants lorsque le nombre de pixels dans chaque classe est approximativement égal. Dans le cas présent, cette hypothèse est potentiellement fautive. La méthode d'Otsu n'est donc pas une méthode adéquate pour le calcul de la DDP.

La figure 3.13 présente un dendrogramme ainsi que la matrice de corrélation des DDP calculées en utilisant différentes valeurs de type de filtre, de taille de fenêtre et de méthode de seuillage. Un dendrogramme est une figure permettant de visualiser un regroupement hiérarchique. Dans ce cas-ci, les regroupements sont effectués en fonction de la corrélation entre les DDP. L'ordre des corrélations dans la matrice de corrélation est en fonction du regroupement effectué par le dendrogramme. Cette figure permet de

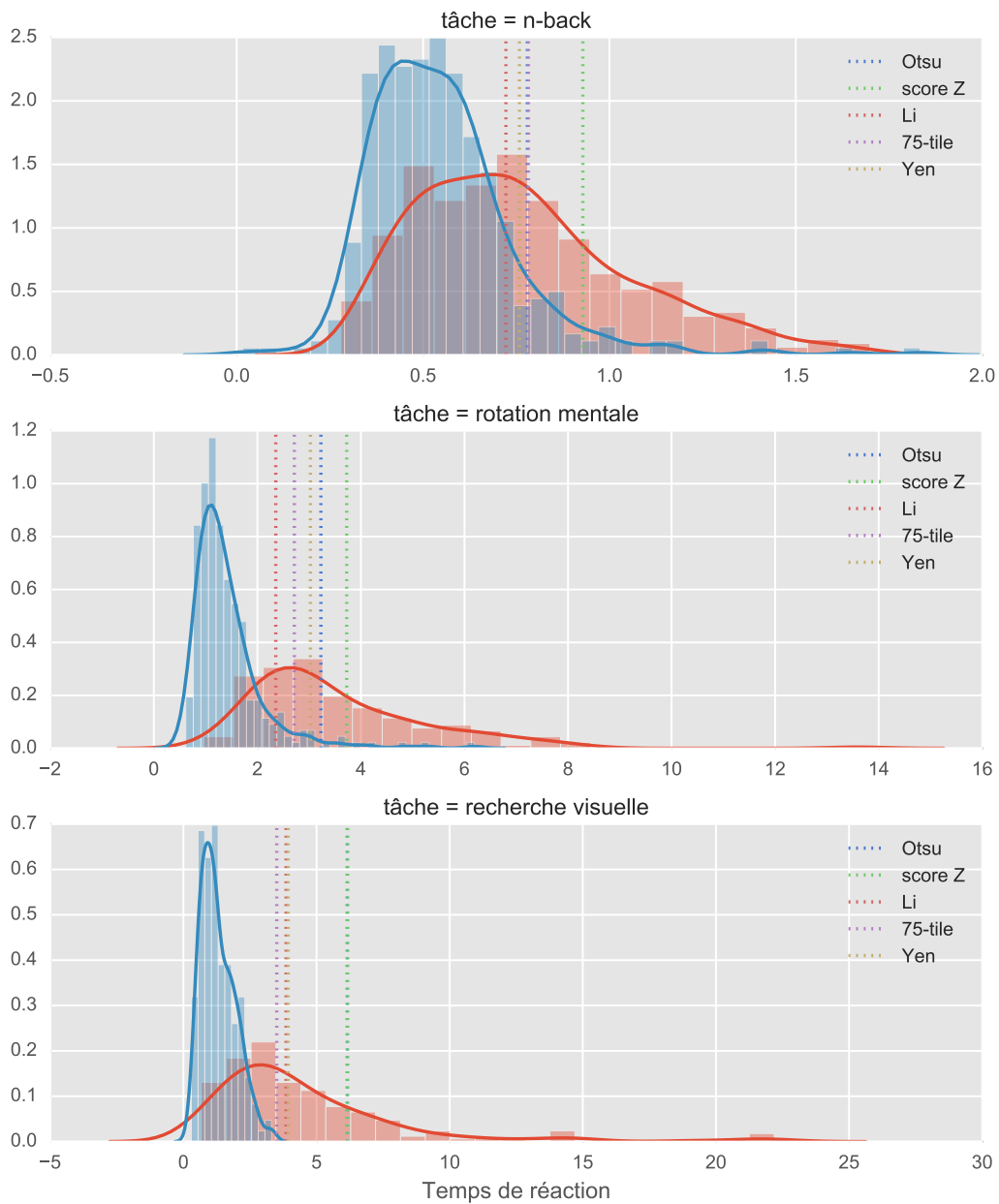


FIGURE 3.12 – Comparaison de plusieurs techniques de seuillage. L’histogramme bleu représente la condition de charge de travail faible et l’histogramme rouge correspond à la charge de travail élevé.  $score - Z$  représente un seuil fixé à un score- $Z$  de 1 et 75 p-tiles représentent un seuil fixé au 75e percentile.

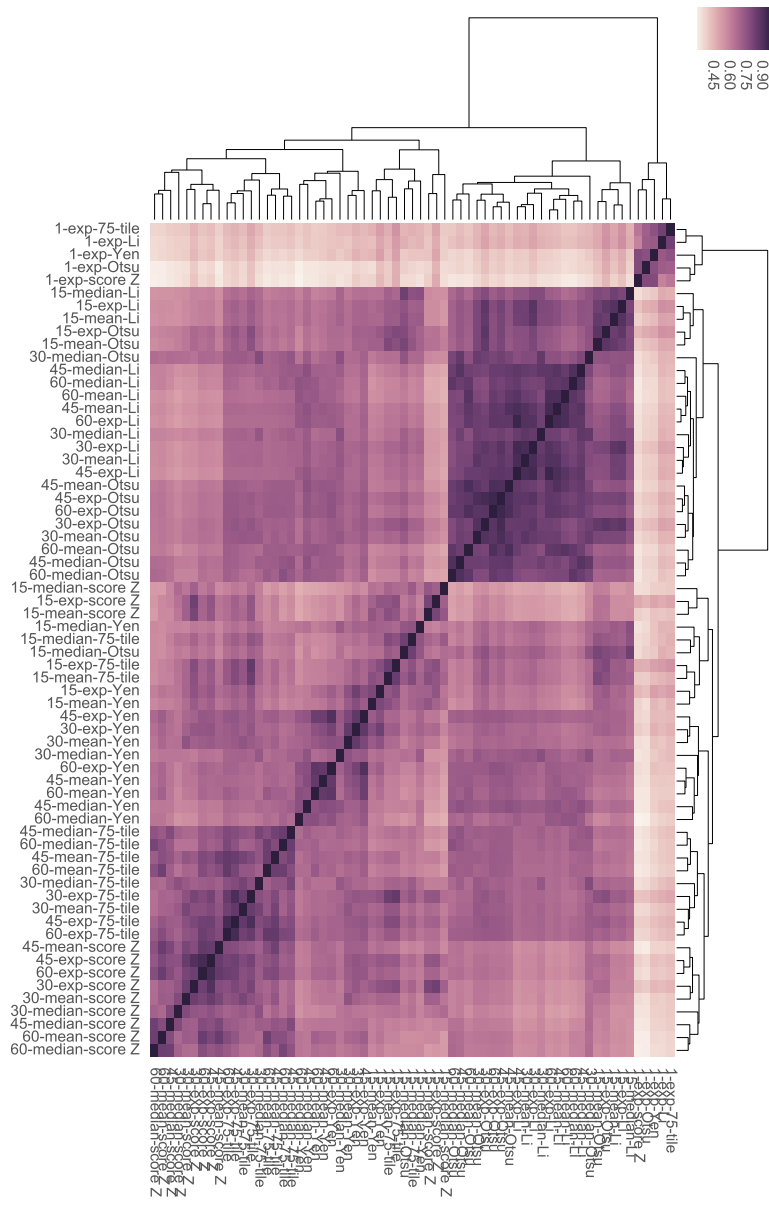


FIGURE 3.13 – Matrice de corrélation entre les différentes combinaisons de paramètres du DDP.

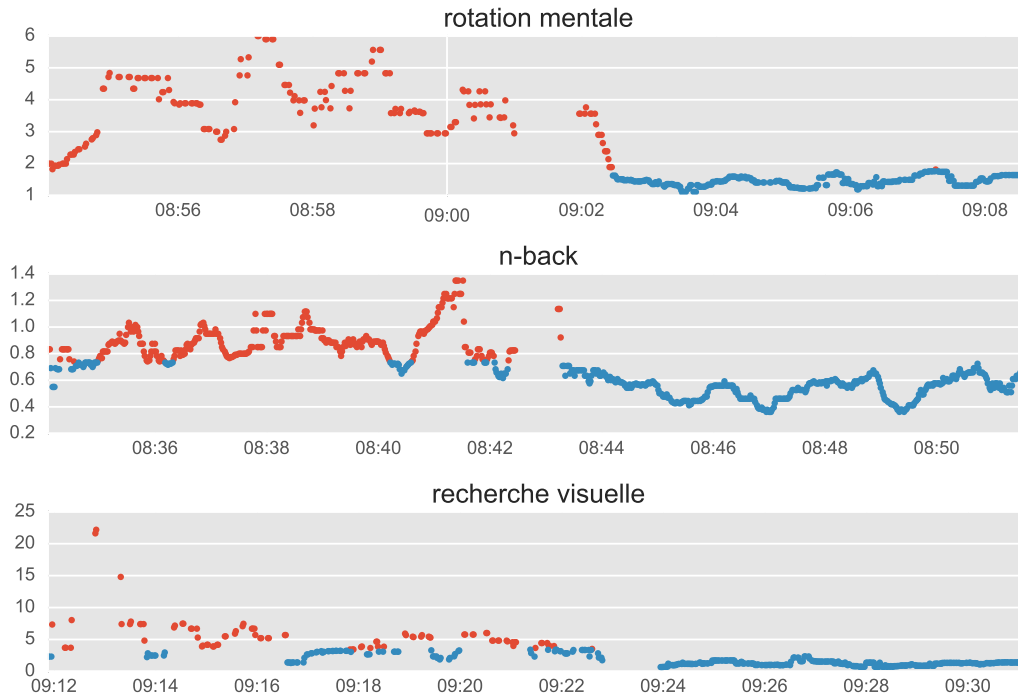


FIGURE 3.14 – Exemple de la DDP calculé sur une session.

faire le constat que trois principaux groupes de paramètres ont une influence.

- Les DDP calculées avec la méthode d’Otsu sont fortement corrélées entre elles, mais peu avec les autres DDP ;
- Les DDP calculées sans filtre sont fortement corrélées entre elles, mais peu avec les autres ;
- Les autres DDP sont corrélées entre elles.

La figure 3.14 présente le résultat du calcul de la DDP sur une session. Le calcul a été fait avec un filtre médian et une fenêtre de 30 secondes.

### 3.4 Apprentissage automatique

Tel que décrit au chapitre 2, les algorithmes d’apprentissage automatique sont des modèles permettant d’apprendre la relation entre un vecteur d’entrée  $\mathbf{x}$  et une valeur en sortie  $y$  à l’aide des expériences du passé. Dans le cas de l’utilisation de ces algorithmes pour l’estimation de l’état fonctionnel de l’opérateur, le vecteur  $\mathbf{x}$  utilisé correspond aux valeurs données par les capteurs physiologiques et comportementaux ainsi que les

caractéristiques d'intérêt calculées à partir de ces signaux, présentées à la section 3.1.3. La valeur de sortie  $y$  quant à elle est calculée à l'aide de la DDP, présentée à la section 3.3.

Le chapitre 2 décrit l'apprentissage automatique de manière générale. Il est toutefois important de rappeler que la méthodologie utilisée pour entraîner et tester la performance de l'apprentissage automatique a une importance primordiale. En effet, une mauvaise méthodologie dans l'application de l'apprentissage automatique peut causer des lacunes dans les généralisations pouvant être attendues des résultats. Cette section présente la méthodologie utilisée lors de ces expériences.

### 3.4.1 Procédure

Cette section présente la procédure utilisée lors des expériences d'apprentissage automatique.

#### Chargement des données

Cette étape consiste à charger les données du BioHarness, du EyeTribe ainsi que celles provenant des différentes tâches à partir du disque dur. Celles-ci sont synchronisées à la seconde près en utilisant la moyenne dans le cas des données ayant une fréquence plus élevée qu'un hertz et en utilisant une interpolation linéaire pour les données ayant une fréquence égale ou inférieure à un hertz.

#### Prétraitement des données

Le prétraitement des données consiste à calculer tous les prétraitements présentés à la section 3.1.3. Les prétraitements sur les données physiologiques, comportementales et le calcul de la DDP sont effectués à ce moment. Les données sont ensuite consolidées dans un seul fichier qui pourra être chargé directement du disque pour toutes les expériences.

#### Division des données

Tel que mentionné au chapitre 2, l'évaluation des algorithmes d'apprentissage automatique nécessite la division du jeu de données en un jeu d'entraînement et de test [4]. Pour éviter d'avoir une grande dépendance temporelle entre le jeu d'entraînement et de test, les données sont séparées en utilisant la méthode suivante. Le premier 20% des données de chaque bloc expérimental sont employées comme jeu de test. Les blocs expérimentaux sont définis en fonction du moment de la journée, de la tâche expérimentale, du niveau de charge et du participant. Par exemple, l'exécution du n-back, avec un niveau de

charge faible, lors de la session du matin du participant numéro 1 représente un bloc expérimental.

Cette méthode a plusieurs avantages par rapport à un échantillonnage aléatoire. Premièrement, l'utilisation des blocs expérimentaux permet de s'assurer que les distributions des participants, du niveau de charge de travail, de la session de travail et de la tâche expérimentale sont équivalents dans le jeu d'entraînement et le jeu de test. Deuxièmement, en utilisant les premiers 20% des données, la réplicabilité du processus est simplifiée. Finalement, la dépendance temporelle entre les données du jeu d'entraînement et du jeu de test est limitée.

### **Évaluation de la performance**

L'évaluation des performances des différents algorithmes s'effectue à l'aide des MCC [56]. Cette technique est utilisée puisqu'elle permet la comparaison de la performance entre des jeux de données ayant une distribution de classes différentes. L'équation permettant de calculer les MCC est présenté à la section 2.3.3.

Plusieurs types de séparations sont utilisées pour démontrer divers aspects de la généralisation des modèles entraînés. Ces modèles sont entraînés en utilisant la validation croisée et sont évalués sur le jeu de test selon les modalités suivantes :

**Modèle général** Le modèle général consiste à entraîner un seul modèle avec les données de tous les participants et de toutes les tâches.

**Modèle intra tâche** Le modèle intra tâche consiste à entraîner un modèle par tâche et à évaluer celui-ci sur cette même tâche.

**Modèle intra participant** Le modèle intra participant consiste à entraîner un modèle par participant et à évaluer celui-ci sur ce même participant.

**Modèle inter participant** Le modèle inter participant consiste à entraîner un modèle sur les données de plusieurs participants et à évaluer celui-ci sur un participant qui n'a pas été utilisé pour l'entraînement.

Pour rapporter la performance sur le jeu de test, les données sont séparées par tâche et par participant. La performance est donc évaluée pour chaque partition en utilisant le classifieur entraîné précédemment sur les données appropriées.

TABLE 3.5 – Hyperparamètres utilisés pour la recherche aléatoire

	Paramètre	Min.	Max.	Distribution
KNN	k	1	100	uniforme
SVM	C	0.001	100	logarithmique
	$\gamma$	0.001	100	logarithmique
Random Forest	nombre d'estimateurs	1	100	uniforme

### Sélection des hyper paramètres

Pour obtenir de bonnes performances, il est important de choisir les bonnes valeurs pour les hyper paramètres des algorithmes utilisés. Cependant, il est primordial de ne pas utiliser le jeu de test pour choisir ceux-ci. Dans le cadre de ce projet, la validation croisée et la recherche aléatoire sont utilisées.

**Validation croisée** La validation croisée est une méthode permettant d'estimer la performance en généralisation d'un algorithme sans utiliser le jeu de données de test.

Cette validation est effectuée en séparant le jeu de données d'entraînement en 4 plis. Cette validation consiste donc en 4 entraînements sur 3 des 4 plis et une évaluation sur le pli restant. La moyenne obtenue permet donc d'obtenir une estimation de la capacité de généralisation d'un algorithme ou d'un ensemble d'hyper paramètres.

**Recherche aléatoire** Selon Bergstra et Bengio [6], l'utilisation de la recherche aléatoire est plus efficace pour optimiser les hyper paramètres que la recherche en grille. De plus, cette méthode est implémentée dans la plupart des bibliothèques d'apprentissage automatique, ce qui facilite la réplique des résultats. La procédure d'entraînement consiste à entraîner chaque classifieur 50 fois à l'aide d'hyper paramètres choisis aléatoirement. Le classifieur est entraîné indépendamment pour chaque paire de jeux d'entraînement et de validation et la performance moyenne est calculée. Les hyper paramètres choisis sont ceux dont la performance moyenne est la plus élevée. On réentraîne ensuite le classifieur sur toutes les données en utilisant les hyper paramètres choisis.

Les hyper paramètres qui ont été manipulés sont présentés au tableau 3.5. Ce tableau présente les distributions utilisées dans la recherche aléatoire d'hyper paramètres.



## **Classifieurs utilisés**

Cette section présente les classifieurs utilisés ainsi que les détails de leur implémentation. Une description plus détaillée de leur fonctionnement est présentée dans la section 2.

Trois classifieurs sont utilisés pour démontrer la performance de base de l'apprentissage automatique. Ceux-ci sont tous implémentés dans le cadriciel Scikit-learn [66]. Les trois classifieurs utilisés sont :

- Une forêt d'arbre aléatoire ;
- Un SVM ;
- Un KNN.

## **Standardisation des données**

Pour tous les classifieurs utilisés, les données sont standardisées. Pour ce faire, la médiane et l'étendue interquartile du jeu d'entraînement sont calculées pour chaque dimension. Les valeurs sont ensuite standardisées en soustrayant la valeur de la médiane calculée et en divisant par l'étendue interquartile.

# Chapitre 4

## Résultats

Ce chapitre présente les résultats obtenus en utilisant la méthodologie décrite au chapitre précédent. Ces résultats sont présentés selon la même structure que le chapitre 3. Premièrement, les données récoltées lors de la collecte de données sont données en deux volets : les valeurs des signaux physiologiques et comportementaux ainsi que les performances obtenues lors des tâches expérimentales. Des observations sur la nature des données sont apportées. Par exemple, les distributions des différents signaux physiologiques et des différentes mesures de performances sont présentées. Finalement, les résultats des expériences d'apprentissage automatique sont présentés.

### 4.1 Collecte de données

La collecte de données s'est déroulée entre le mois de février 2015 et le mois de juillet 2015. Les capteurs qui ont été utilisés sont présentés à la section 3.1. Les tâches utilisées sont présentées à la section 3.2. Les trois tâches ont été évaluées pour chaque participant sur deux niveaux de difficulté.

Les participants recrutés sont des étudiants de l'Université Laval. Leur âge moyen est de 26,12 ans et l'écart type est de 3,40 ans. Le groupe était composé de 6 hommes et 2 femmes. La figure 4.1 montre la distribution de l'âge des participants.

Chaque participant a effectué deux sessions de collecte de données : une le matin et une en fin d'après-midi.

Bien que les données utilisées soient ordonnées dans le temps, les algorithmes étudiés dans les prochaines sections n'utilisent pas cet ordre et considèrent chaque pas de temps comme indépendant. Pour cette raison, les données provenant des différents capteurs

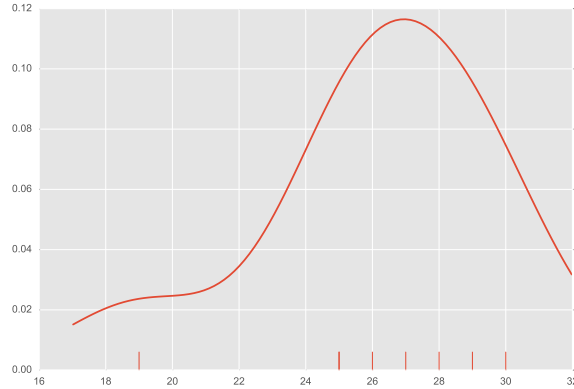


FIGURE 4.1 – Distribution de l'âge des participants

sont considérées indépendamment du temps.

Comme il a été mentionné au chapitre 2, plusieurs algorithmes d'apprentissage automatique posent l'hypothèse que les données sont indépendantes et identiquement distribuées (IID). Puisque celles-ci proviennent de participants et de tâches différentes, celles-ci ne sont donc pas nécessairement IID. Les distributions des signaux de la section 4.1.1 en fonction des conditions de la collecte de données sont présentées dans cette section.

### 4.1.1 Signaux physiologiques

Lors des sessions de collecte de données avec les participants, des signaux physiologiques et comportementaux ont été recueillis. Ceux-ci ont la forme d'une série temporelle de valeurs correspondant à la sortie de chaque capteur à une fréquence de 1 Hz.

La figure 4.2 présente la sortie de quatre capteurs dans le temps. Dans cette figure, on peut identifier six groupes de données. Il s'agit de trois tâches expérimentales avec deux niveaux de charge de travail pour chacune de ces tâches. Dans tous les cas, la charge de travail élevée suit la charge de travail faible. Un élément intéressant de cette figure est le fait que les données physiologiques sont bruitées. Cet élément peut avoir une influence sur les performances des algorithmes d'apprentissage automatique.

#### Effet du participant sur les signaux

Les données récoltées lors de la collecte de données l'ont été sur plusieurs participants. Cependant, les participants sont tous différents physiologiquement.

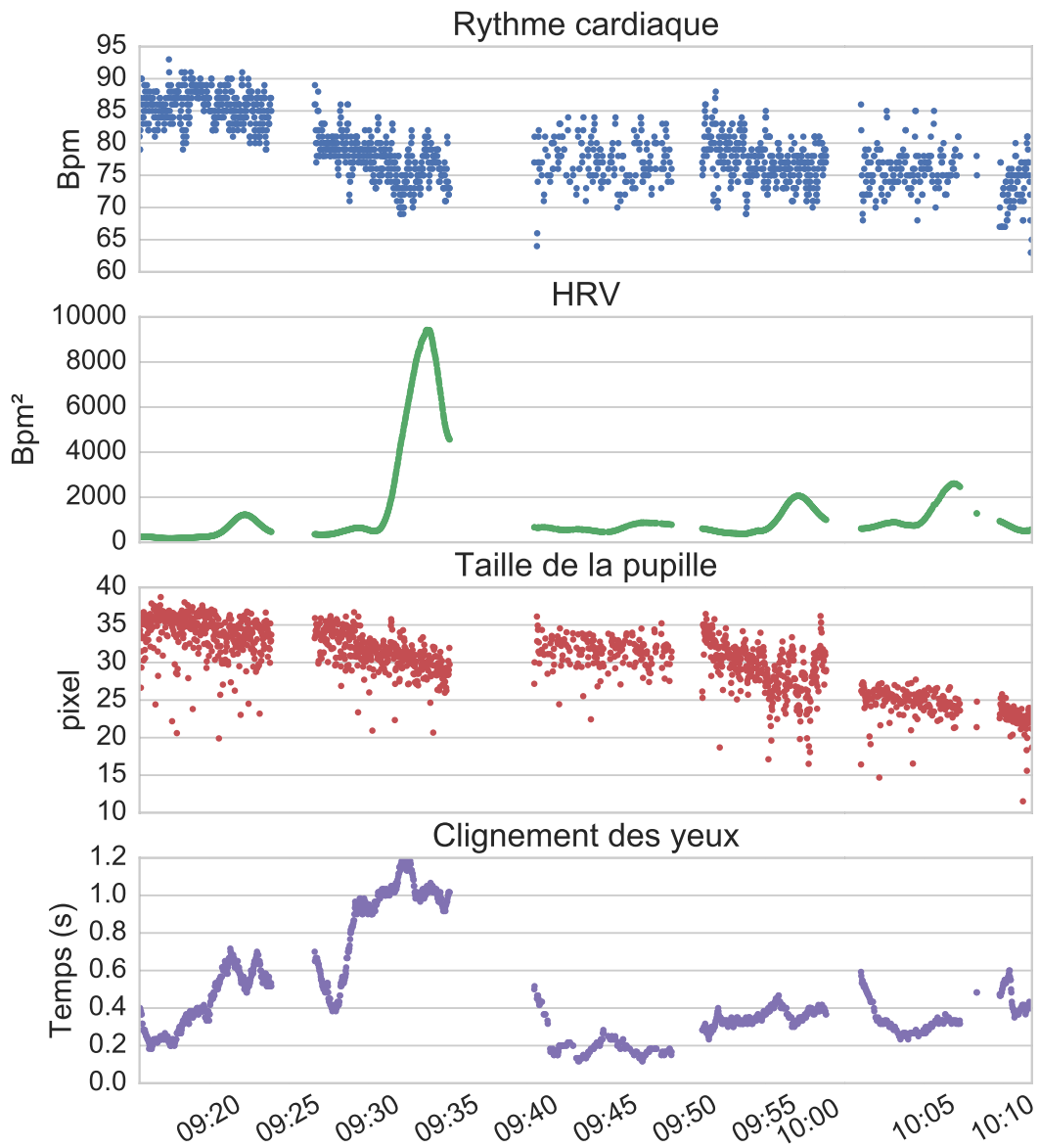


FIGURE 4.2 – Valeurs de différents signaux physiologiques et comportementaux en fonction du temps.

Dans la figure 4.3, les distributions de quatre signaux physiologiques et comportementaux sont présentées en fonction du participant. Ces distributions sont présentées sous forme de boîte à moustaches, ou *boxplot* en anglais. Cette figure montre les différences dans les distributions des signaux physiologiques et comportementaux provenant des différents participants. En effet, il semble que les données provenant de participants séparés ne sont pas identiquement distribuées. Les points qui se trouvent à l'extérieur des boîtes sont des points qui sont à plus de 1.5 écarts-types de la moyenne. Bien que ceux-ci semblent nombreux, il s'agit en fait d'une faible proportion des échantillons. Ils ne semblent nombreux qu'à cause du grand nombre d'échantillons disponibles.

### **Effet de la tâche sur les signaux**

Comme dans le cas des participants, les tâches effectuées ont également un effet sur les distributions des différents signaux physiologiques.

La figure 4.4 présente la distribution de quatre signaux provenant de quatre capteurs pour un seul participant, niveau de charge de travail et moment dans la journée. Cette figure est un diagramme de violon. Un diagramme de violon est similaire à un diagramme en boîte ou *boxplot*. Cependant, au lieu d'un rectangle, la distribution du signal est indiquée. Cette distribution est obtenue avec la méthode de Parzen-Rosenblatt [74]. Pour des fins esthétiques, une réflexion de cette courbe est effectuée autour de l'axe vertical.

Cette figure montre que les tâches effectuées ont une influence sur la distribution des signaux provenant des différents capteurs.

### **Effet des manipulations expérimentales sur les signaux**

Au chapitre précédent, il est mentionné que les manipulations de la charge de travail et du moment de la journée sont effectuées pour augmenter la variabilité dans les données. Il est donc intéressant d'observer l'effet de ces conditions sur les données physiologiques et comportementales.

La figure 4.5 montre les signaux physiologiques et comportementaux par rapport aux conditions expérimentales, c'est à dire, le moment de la journée et la charge de travail. Dans celle-ci, on peut remarquer que les signaux ne sont pas affectés uniformément par les différentes conditions. Certaines variables sont plus affectées par la charge de travail, d'autres sont plus affectées par le moment de la journée et, finalement, certaines demeurent stables, peu importe la condition.

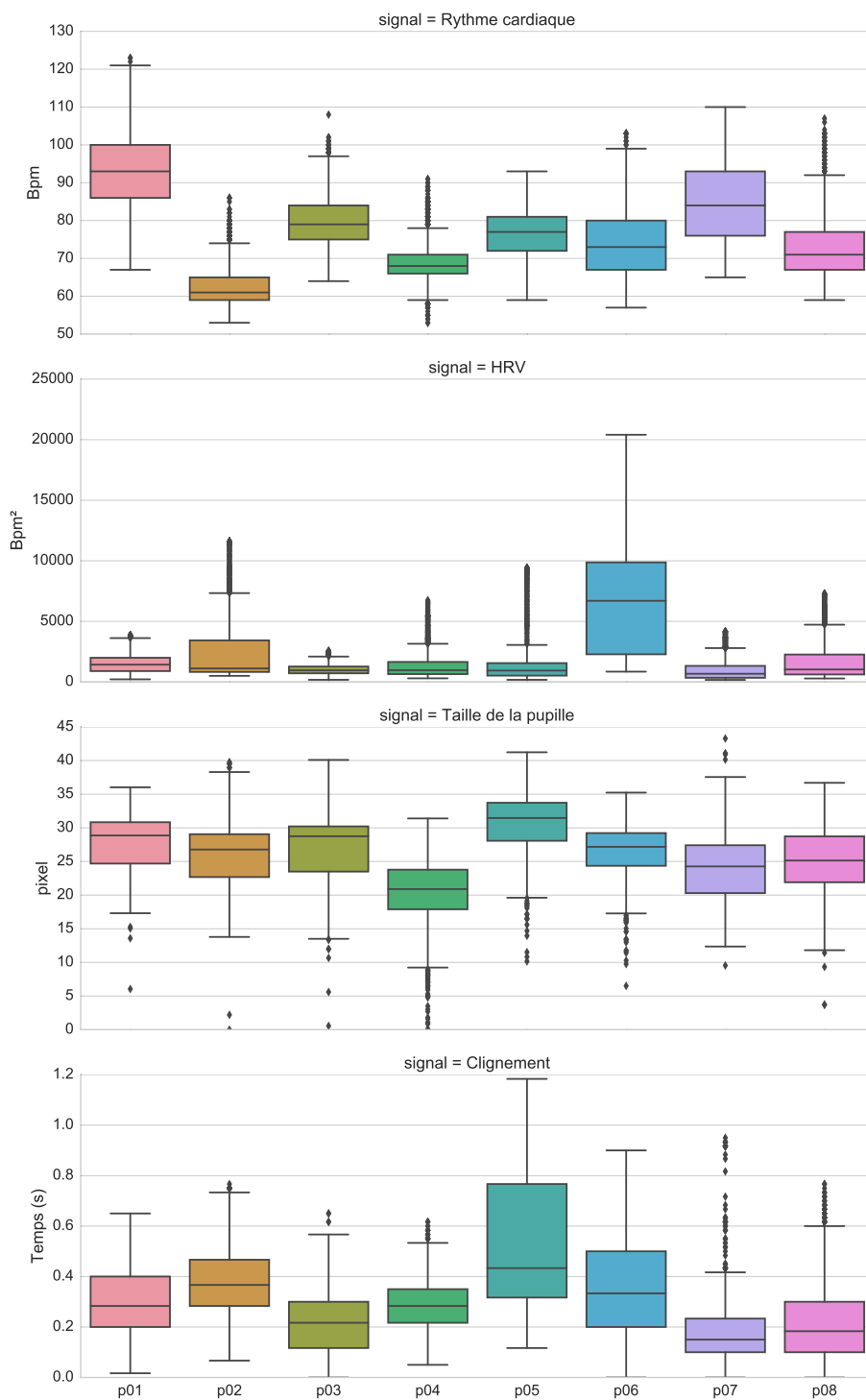


FIGURE 4.3 – Distributions de quatre signaux en fonction des participants.

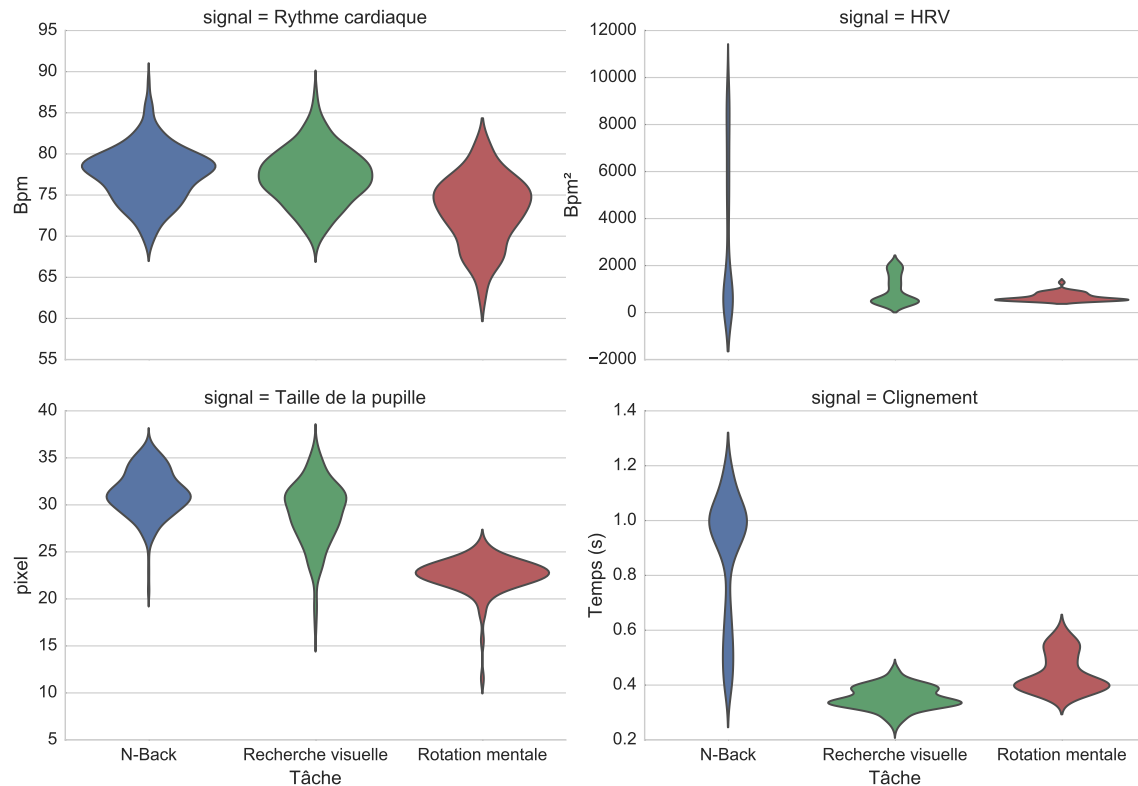


FIGURE 4.4 – Distribution de quatre signaux en fonction de la tâche expérimentale.

### Signaux en fonction de la DDP

La DDP présentée au chapitre 3 permet d’obtenir une mesure de l’OFS à partir de la performance d’un participant dans une tâche. Dans ce contexte, il est intéressant de voir les distributions des différents signaux physiologiques en fonction de cette mesure.

Les figures 4.6 et 4.7 présentent la distribution de quatre signaux physiologiques et comportementaux en fonction de la DDP. Dans la figure 4.6, les données proviennent d’un participant pour toutes les tâches, toutes les charges de travail pour le matin et l’après-midi. Dans la figure 4.7, les données proviennent de la collecte de données d’un participant pour une seule charge, avec un niveau de charge de travail élevée pour la session du matin.

Ces figures mettent en évidence le fait qu’un seul signal physiologique n’est pas suffisant pour distinguer si un état est fonctionnel ou non.

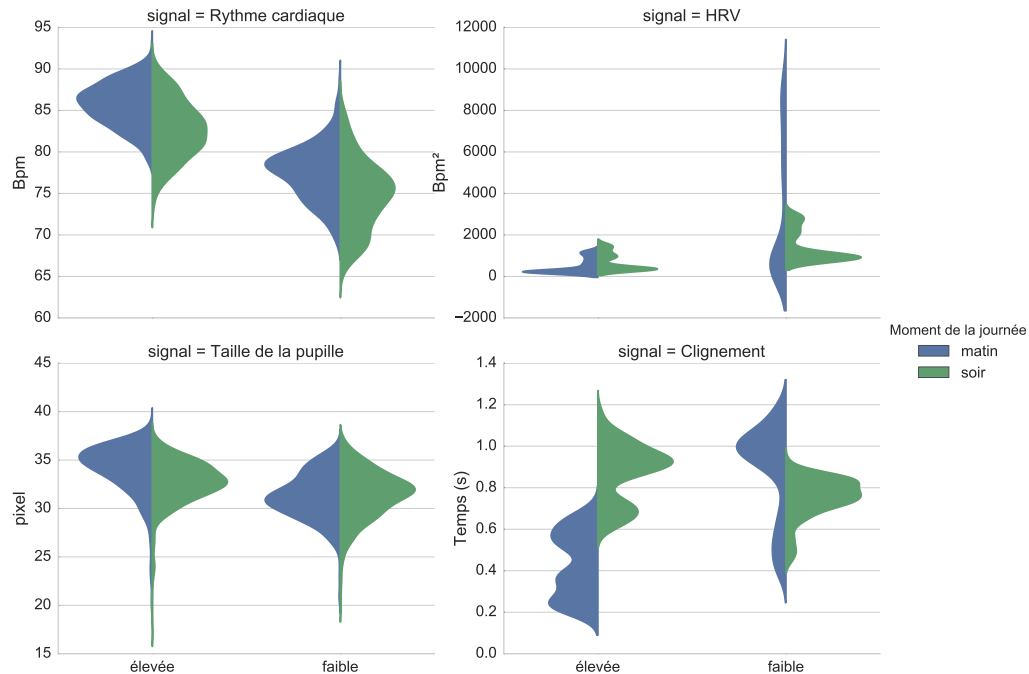


FIGURE 4.5 – Distributions de quatre signaux en fonction du moment de la journée (en bleu et en vert) et de la charge de travail (gauche et droite).

#### 4.1.2 Conclusions

Les figures présentées dans cette section permettent de faire certaines observations. Premièrement, les signaux physiologiques de chaque participant et de chaque tâche sont très différents. Cette caractéristique peut influencer la capacité des algorithmes d'apprentissage automatique à généraliser sur de nouveaux participants et de nouvelles tâches. Cela affectera donc l'objectif de recherche de l'étude de la capacité de généralisation, présentée à la section 1.5.3.

Deuxièmement, les informations contextuelles telles que le niveau de charge de travail, le moment de la journée et la tâche en cours d'exécution influencent les distributions de plusieurs signaux biologiques. Il est donc intéressant d'étudier l'influence de l'ajout de ces informations dans le processus d'apprentissage automatique pour vérifier si cet ajout permet d'obtenir de meilleures performances. Cette vérification permettra de remplir l'objectif de recherche présenté à la section 1.5.4, c'est-à-dire d'évaluer le gain en performance des algorithmes d'apprentissage automatique pouvant être obtenu par l'ajout d'information contextuelle.



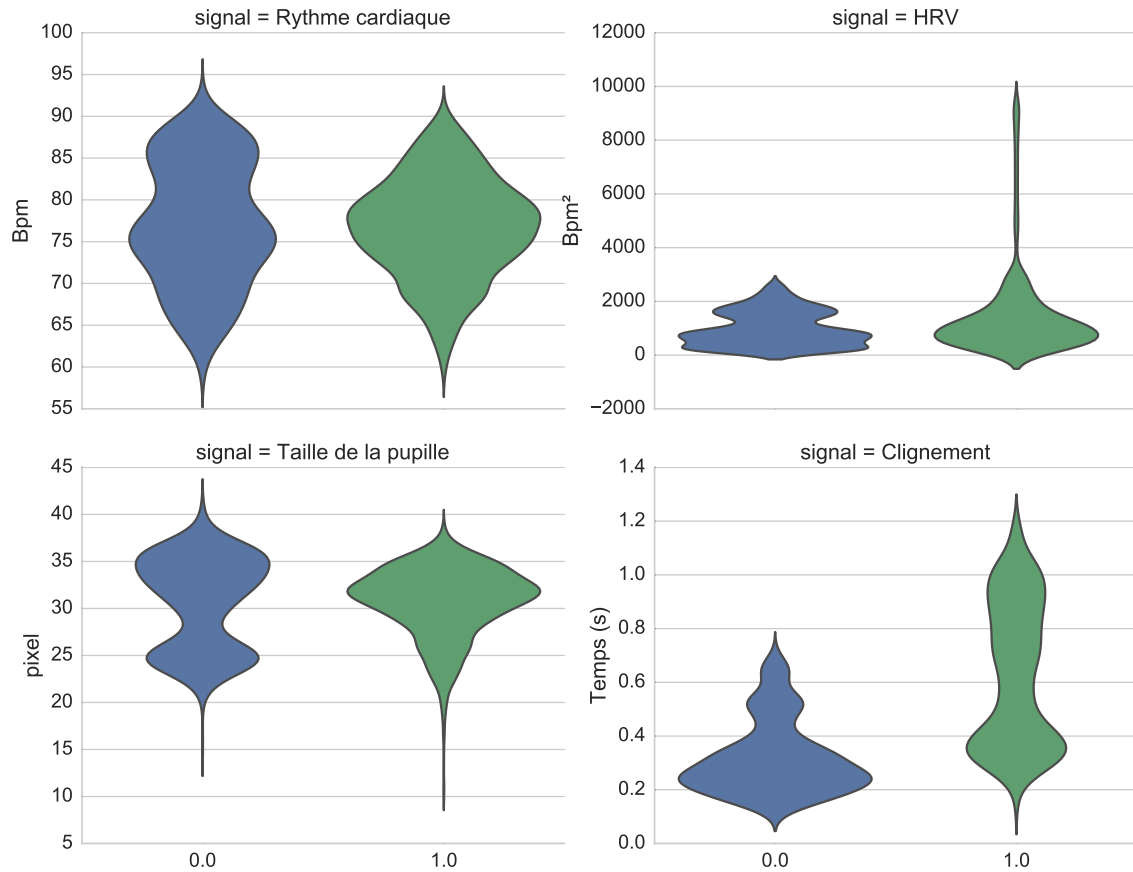


FIGURE 4.6 – Distribution des signaux biologiques d’un participant en fonction du DDP.

## 4.2 Performances

Les tâches utilisées lors de la collecte de données sont le n-back, la tâche de recherche visuelle et la tâche de rotation mentale. Elles permettent d’enregistrer des mesures de la performance instantanée du participant. Celles-ci sont présentées à la section 3.2.2. Les tâches de n-back, de recherche visuelle et de rotation mentale enregistrent le temps de réaction du participant. Le n-back et la rotation mentale enregistrent également si le participant a obtenu la bonne réponse ou non.

Cette section présente les performances obtenues par les participants dans ces tâches et les conclusions qui peuvent en être tirées. L’influence des manipulations expérimentales sur les performances est également abordée dans cette section.

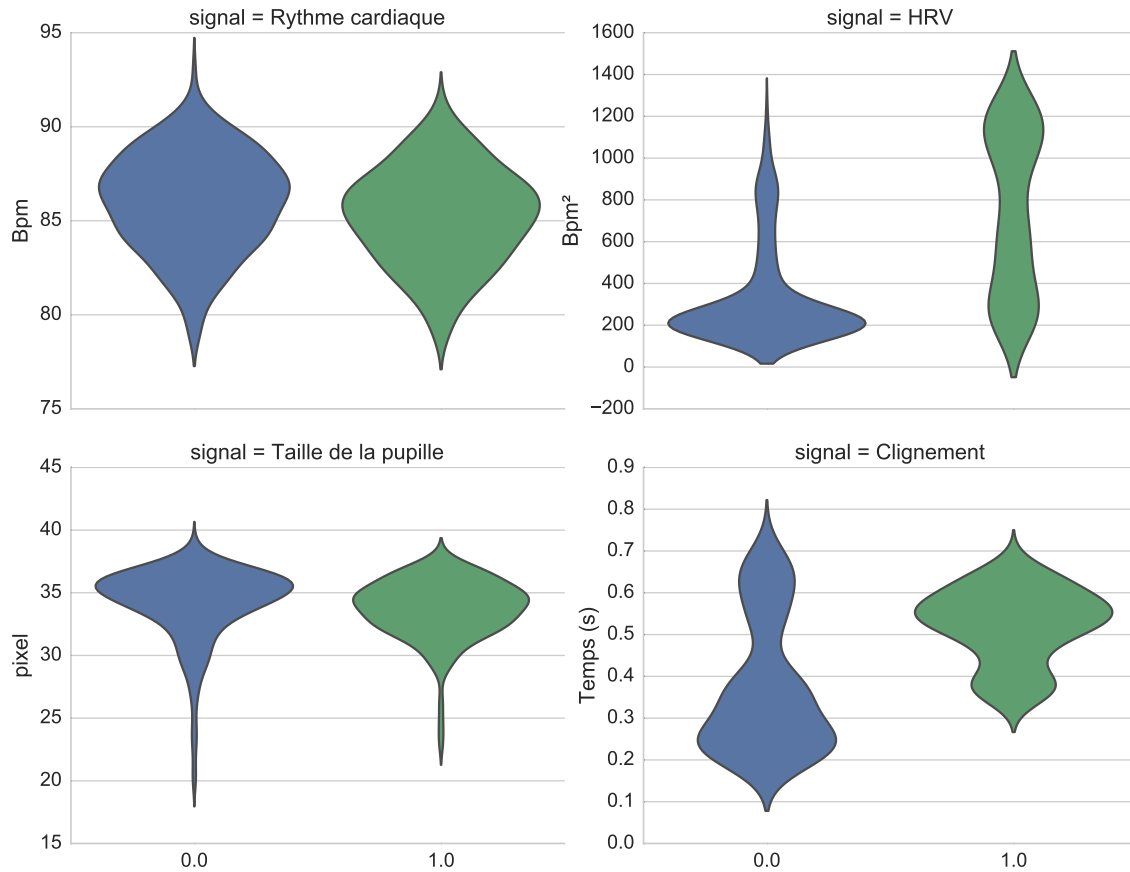


FIGURE 4.7 – Distribution des signaux biologiques d’un niveau de charge de travail, d’une session et d’un participant en fonction du DDP.

#### 4.2.1 Exactitude des réponses

L’exactitude des réponses consiste à savoir si le participant obtient la bonne réponse. Dans la tâche du n-back, le participant doit indiquer si la lettre présentée à l’écran est la lettre cible. Dans la tâche de rotations mentales, le participant doit indiquer si les formes à l’écran correspondent à la même forme en trois dimensions. L’exactitude des réponses est un but primaire.

La figure 4.8 présente la proportion moyenne d’essais réussis pour un participant en fonction de la tâche, du niveau de difficulté et du moment de la journée. Dans cette figure, il faut noter le fait que les performances sont très élevées pour plusieurs conditions. En effet, la performance se trouve sous la barre des 80 % seulement pour une seule session d’un participant.

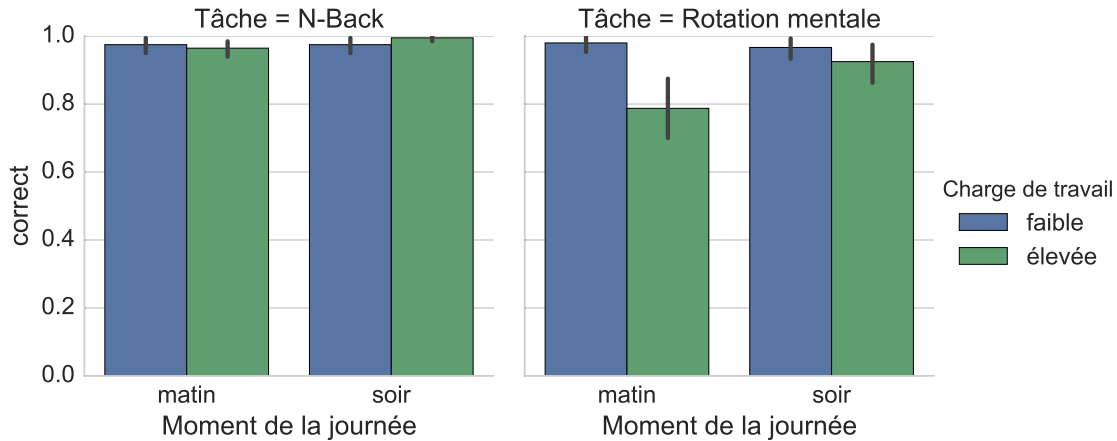


FIGURE 4.8 – Proportion des réponses correctes dans les tâches de n-back et de rotation mentale pour un participant.

La faible plage dynamique de la performance des participants pour l'objectif principal des différentes tâches est prédite par la théorie. Comme indiqué au chapitre 1, dans un contexte où l'OFS d'un participant est diminué, celui-ci peut généralement soutenir un niveau de performances acceptables dans le but primaire en fournissant un effort supplémentaire.

## 4.2.2 Temps de réaction

Les temps de réaction, c'est-à-dire le délai entre l'affichage d'un stimulus et l'enregistrement de la réponse du participant, sont enregistrés dans les trois tâches de la collecte de données. La figure 4.9 présente les distributions des temps de réaction d'un participant pour les trois tâches. Cette figure met en évidence l'effet qu'ont les conditions d'expérimentations, soit la charge de travail et le moment de la journée, sur ces distributions. Contrairement à l'objectif principal présenté à la section précédente, le temps de réaction constitue un objectif secondaire de ces tâches expérimentales. Il est donc attendu que celui-ci soit plus affecté par les manipulations expérimentales.

## 4.2.3 DDP

Comme il est indiqué au chapitre 3, la DDP est utilisée pour estimer l'OFS à partir de la performance dans les différentes tâches. Pour valider cette méthode, la proportion des étiquettes de l'état fonctionnel et sous-fonctionnel est calculée pour chaque participant et

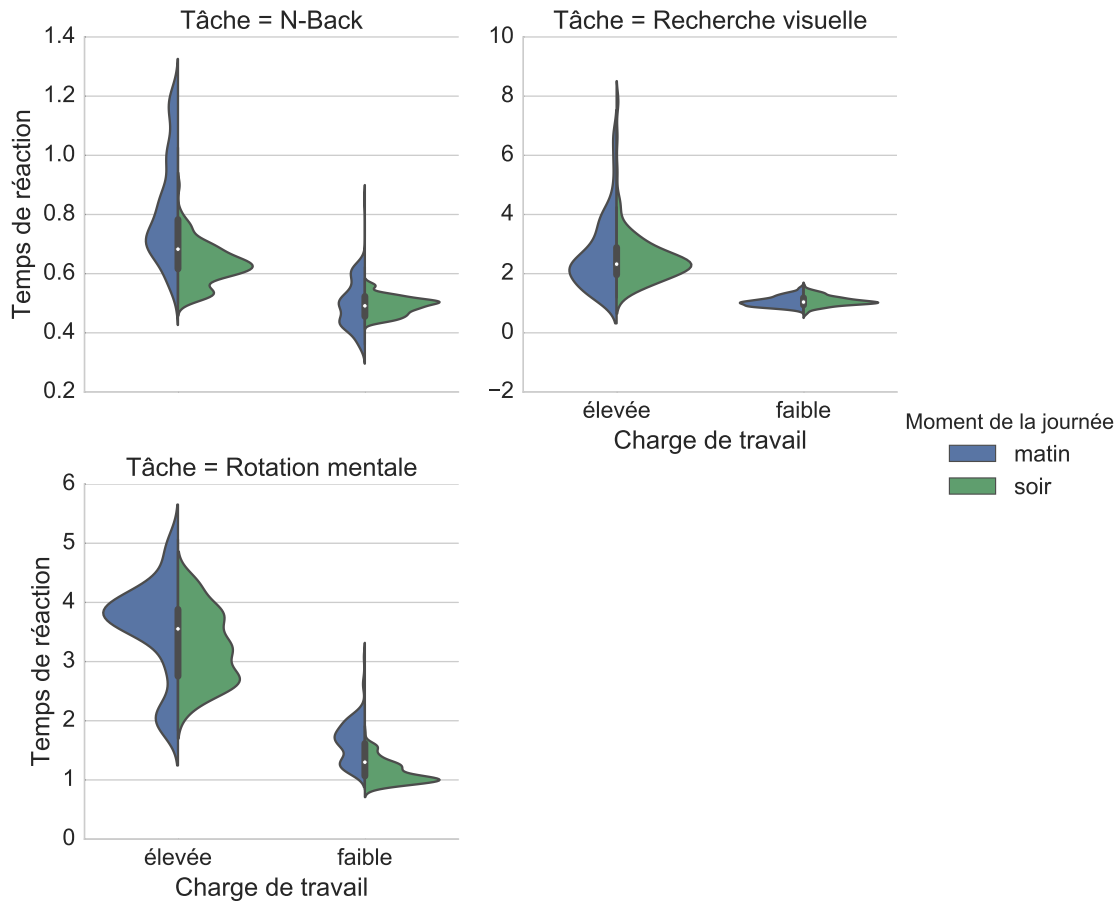


FIGURE 4.9 – Distribution des temps de réaction d’un participant dans les trois tâches expérimentales.

chaque tâche. Ces proportions sont présentées au tableau 4.1. Cette vérification permet de valider qu’aucune section de l’expérience ne présente un cas où le calcul de la DDP n’a pas été en mesure de détecter deux classes.

Comme il est mentionné au chapitre 3, le jeu de données est séparé en deux jeux : un jeu d’entraînement pour déterminer les paramètres des différents modèles, et un jeu de test pour évaluer la performance en généralisation de ces algorithmes. De plus, comme il est mentionné dans le chapitre 2, pour obtenir des performances acceptables de la part des algorithmes d’apprentissage automatique, il est important d’entraîner ceux-ci en utilisant des données tirées des mêmes distributions que pour le domaine d’application. Le tableau 4.2 présente la proportion d’états identifiés comme fonctionnels dans le jeu de données d’entraînement et de test pour chaque participant et chaque tâche.

TABLE 4.1 – Proportion des étiquettes d’un DDP fonctionnel pour chaque tâche et chaque participant.

Participant	Rotation Mentale	N-Back	Recherche Visuelle
1	0.51	0.61	0.63
2	0.56	0.52	0.67
3	0.57	0.65	0.64
4	0.70	0.55	0.64
5	0.31	0.62	0.65
6	0.67	0.64	0.53
7	0.66	0.70	0.46
8	0.60	0.28	0.60

TABLE 4.2 – Proportion des étiquettes de l’état fonctionnel dans le jeu d’entraînement et de test.

Participant	Tâche	Entraînement	Test	Différence
1	N-Back	0.61	0.58	0.036
	Rotation Mentale	0.52	0.44	0.078
	Recherche Visuelle	0.63	0.65	0.020
2	N-Back	0.53	0.49	0.046
	Rotation Mentale	0.56	0.58	0.021
	Recherche Visuelle	0.69	0.60	0.098
3	N-Back	0.68	0.57	0.110
	Rotation Mentale	0.58	0.55	0.026
	Recherche Visuelle	0.66	0.57	0.089
4	N-Back	0.57	0.45	0.120
	Rotation Mentale	0.71	0.67	0.044
	Recherche Visuelle	0.65	0.60	0.050
5	N-Back	0.64	0.55	0.089
	Rotation Mentale	0.13	0.59	0.460
	Recherche Visuelle	0.67	0.59	0.077
6	N-Back	0.69	0.44	0.250
	Rotation Mentale	0.72	0.51	0.200
	Recherche Visuelle	0.49	0.63	0.140
7	N-Back	0.70	0.69	0.008
	Rotation Mentale	0.65	0.71	0.055
	Recherche Visuelle	0.42	0.54	0.120
8	N-Back	0.25	0.37	0.110
	Rotation Mentale	0.66	0.34	0.320
	Recherche Visuelle	0.65	0.41	0.250

TABLE 4.3 – Proportion des étiquettes de l’état fonctionnel en fonction de chaque condition expérimentale.

charge de travail	matin	soir	toutes
faible	0.91	0.92	0.92
élevée	0.09	0.23	0.16
toutes	0.55	0.63	0.59

Le tableau 4.2 permet donc de vérifier que le jeu de test ne contient pas de déséquilibre qui pourrait perturber les mesures de la performance des algorithmes d’apprentissage automatique.

Puisque les résultats de performances de l’apprentissage automatique seront également faits pour chaque participant et chaque tâche séparément, il est important de vérifier que chacun de ces groupes contient des exemples de chaque classe.

Le tableau 4.3 présente la proportion d’échantillons classés comme fonctionnels dans le jeu de données en fonction des conditions expérimentales. Ces résultats sont attendus puisque les manipulations expérimentales correspondent à des conditions ayant une influence sur l’OFS.

## 4.3 Apprentissage automatique

Dans cette section, les résultats des expériences d’apprentissage automatique sont présentés. Ceux-ci sont importants puisqu’il s’agit de la technique employée pour faire l’estimation de l’OFS. Ces résultats permettront donc de répondre à l’objectif de l’évaluation de la performance des algorithmes d’apprentissage automatique présenté à la section 1.5.2. Ces mêmes résultats permettent également d’évaluer la performance en généralisation, qui est également un objectif de recherche.

### 4.3.1 Apprentissage supervisé

Les algorithmes de SVM, KNN et de forêts aléatoires sont présentés plus en détail au chapitre 2. Les algorithmes ont été utilisés avec les modélisations suivantes : modèle unique, par tâche, par participant, en généralisation de participant et en généralisation de tâches. Dans le cas d’un modèle unique, un seul modèle est entraîné avec des données de tous les participants. Dans le cas de la modélisation par *tâche*, trois modèles sont

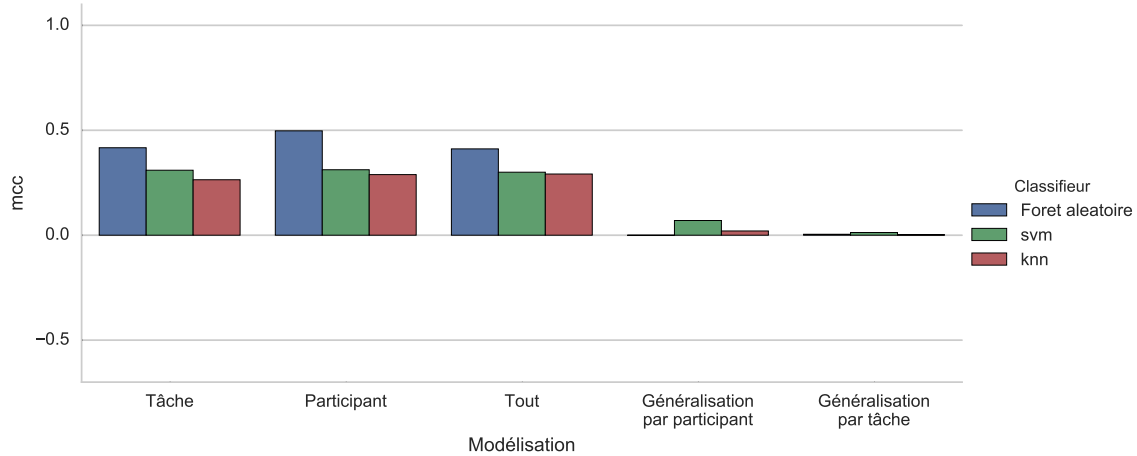


FIGURE 4.10 – MCC calculé sur le jeu de test.

entraînés, soit un par tâche. Dans le cas de la modélisation par *participant*, huit modèles sont entraînés, soit un par participant. Finalement, dans le cas des modélisations de généralisation, les modèles sont entraînés avec les données de sept participants ou deux tâches, et sont testés sur les valeurs du jeu de test correspondant au participant ou la tâche qui n'a pas été utilisée pour l'entraînement.

Les hyper paramètres des algorithmes sont trouvés à l'aide de 50 itérations de recherche aléatoire. Les paramètres utilisés dans la recherche aléatoire sont présentés à la table 3.5. La performance est évaluée en prenant le meilleur MCC moyen sur 4 plis.

La figure 4.10 et le tableau 4.4 présentent le MCC de la classification du jeu de test. À partir de cette figure, il est possible de faire remarquer certains éléments. Premièrement, les forêts aléatoires sont les meilleurs classifieurs pour la modélisation par tâche, par participant et pour toutes les données simultanément. Le fait que les forêts aléatoires performant mieux que les autres classifieurs est en accord avec la littérature. En effet, une évaluation systématique des classifieurs disponibles sur un grand ensemble de jeux de données révélait que les meilleurs classifieurs étaient les forêts aléatoires et les SVMs [27].

Ces résultats semblent montrer des performances semblables à celles obtenues par Hogervorst *et al.* [43]. Dans cet article, la classification du niveau de charge de travail d'un seul participant dans une tâche de n-back en utilisant des caractéristiques physiologiques et comportementales permet d'obtenir une précision entre 0.6 et 0.8. Il est toutefois difficile de faire une comparaison plus en profondeur étant donné que la méthodologie

TABLE 4.4 – Résultats des expériences d’apprentissage automatique

Modélisation	Classifieur	mcc
Tâche	Forêt aléatoire	0.42
	SVM	0.31
	KNN	0.26
Participant	Forêt aléatoire	0.50
	SVM	0.31
	KNN	0.29
Tout	Forêt aléatoire	0.41
	SVM	0.30
	KNN	0.29
Généralisation par participant	Forêt aléatoire	0.00
	SVM	0.07
	KNN	0.02
Généralisation par tâche	Forêt aléatoire	0.00
	SVM	0.01
	KNN	0.00

utilisée n’utilise qu’une seule tâche et ne permet pas de faire des comparaisons entre plusieurs tâches expérimentales ou plusieurs participants.

Deuxièmement, la performance en généralisation, c’est-à-dire sur de nouveaux participants ou sur de nouvelles tâches, est proche de 0, c’est-à-dire du hasard. La faible performance des classifieurs en généralisation n’est pas surprenante. En effet, comme il est indiqué au chapitre 1, l’OFS et les signaux physiologiques et comportementaux sont associés à un état de base qui est différent pour chaque individu. La faible performance en généralisation peut donc être expliquée par le fait que les distributions provenant de participants différents sont différentes. Comme il est mentionné dans le chapitre 2, une des hypothèses sur laquelle reposent les algorithmes d’apprentissage automatique est que les exemples sont IID.

Cependant, pour bien comprendre la performance obtenue, cette mesure n’est pas suffisante. Dans la figure 4.11, la distribution des MCC calculé indépendamment pour chaque tâche et chaque participant est présentée. Cette figure est un diagramme en violon où la distribution est tronquée après le dernier point à chaque extrémité de la distribution pour éviter d’avoir une longue queue pouvant s’étendre au-delà de la valeur maximale du MCC. Cette figure permet de faire remarquer la variabilité dans la performance des algorithmes d’apprentissage automatique. En effet, la performance de la classification obtenue pour un participant et une tâche donnée n’est pas stable. En effet,



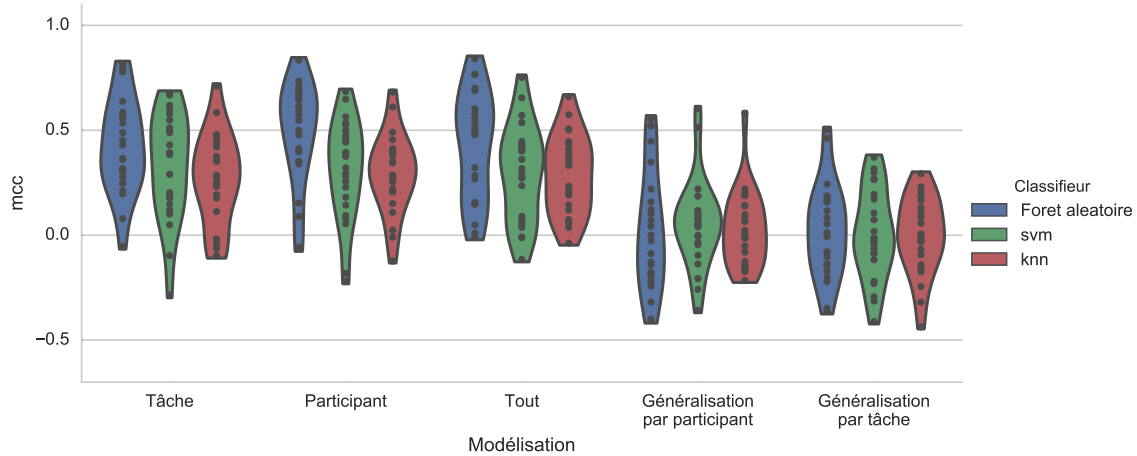


FIGURE 4.11 – Distribution des MCC calculés pour chaque tâche et chaque participant.

pour un même classifieur, la performance pour des participants et des tâches différentes peut être très variable. Les résultats complets peuvent être trouvés à l’annexe C au tableau C.1.

Les figures 4.12 et 4.13 permettent de mettre en relief cette variabilité. Dans la figure 4.12, la performance de la classification à l’aide des forêts aléatoires est présentée en fonction des participants. On peut remarquer dans cette figure que les participants ne sont pas tous aussi difficiles à classifier. Par exemple, le participant numéro huit est mieux classifié que le numéro cinq.

La figure 4.13 quant à elle présente les résultats de la classification à l’aide des forêts aléatoires en fonction des tâches expérimentales.

Cette figure montre que la tâche de rotations mentales est plus difficile à classifier que la tâche de recherche visuelle. La performance de la classification de la tâche de n-back est variable en fonction de la modélisation.

Le fait que les performances en test soient aussi variables en fonction des participants et en fonction des tâches permet en partie d’expliquer les faibles performances des algorithmes d’apprentissage automatique dans la généralisation à de nouveaux participants et à de nouvelles tâches.

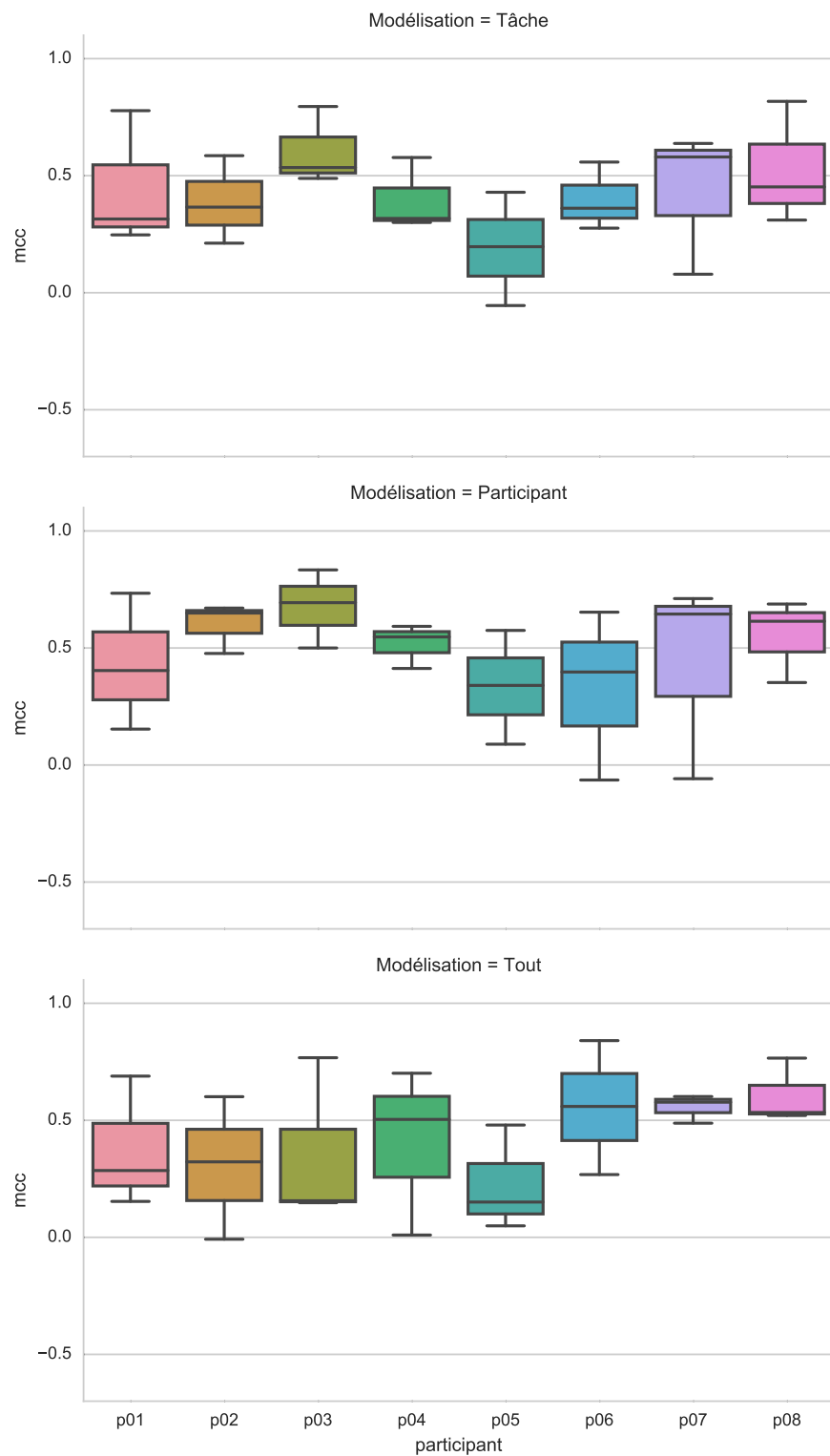


FIGURE 4.12 – MCC de la classification par les forêts aléatoires en fonction des participants.

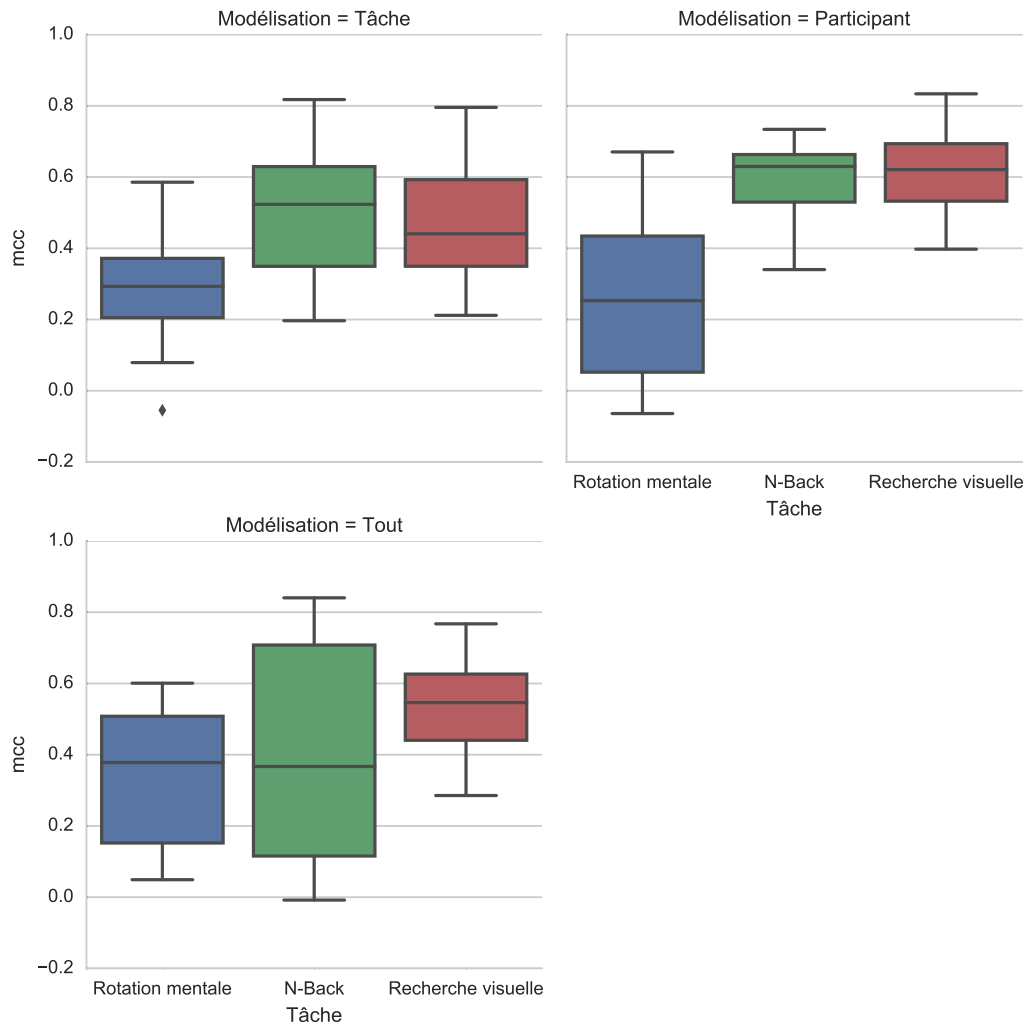


FIGURE 4.13 – MCC de la classification par les forêts aléatoires en fonction des tâches expérimentales.

### 4.3.2 Apprentissage avec contexte

Jusqu'à présent, les objectifs d'évaluation de l'utilisation de l'apprentissage automatique pour l'estimation de l'OFS ainsi que l'objectif d'évaluation de la performance de cette estimation en généralisation ont été remplis. Cependant, la performance en généralisation, c'est-à-dire lorsque cette estimation est effectuée sur une nouvelle tâche ou sur un nouveau participant, est très proche du hasard.

Pour augmenter les performances, il est intéressant d'utiliser toutes les informations disponibles. Par exemple, les conditions expérimentales comme la charge de travail et le moment de la journée sont enregistrés. La figure 4.14 présente la performance obtenue en utilisant cette information contextuelle en plus de l'information provenant des capteurs physiologiques et comportementaux. On remarque que la performance obtenue est supérieure lorsque l'on ajoute l'information contextuelle en plus de l'information des capteurs physiologiques et comportementaux. Parmi les résultats particulièrement prometteurs, les résultats en généralisation des forêts aléatoires sont presque équivalents aux résultats obtenus pour les autres modélisations. Ce résultat est expliqué par le fait que l'effet de la charge de travail et du moment de la journée est observé sur tous les participants.

Bien que l'augmentation de la performance semble une bonne chose, celle-ci ne peut être utile que dans une situation où cette information est disponible. En effet, même si cette information est relativement simple à obtenir lors de l'entraînement, elle est souvent impossible à obtenir dans les conditions de tests. Dans les applications réelles de l'estimation de l'OFS, il est généralement difficile d'obtenir une estimation fiable de la charge de travail.

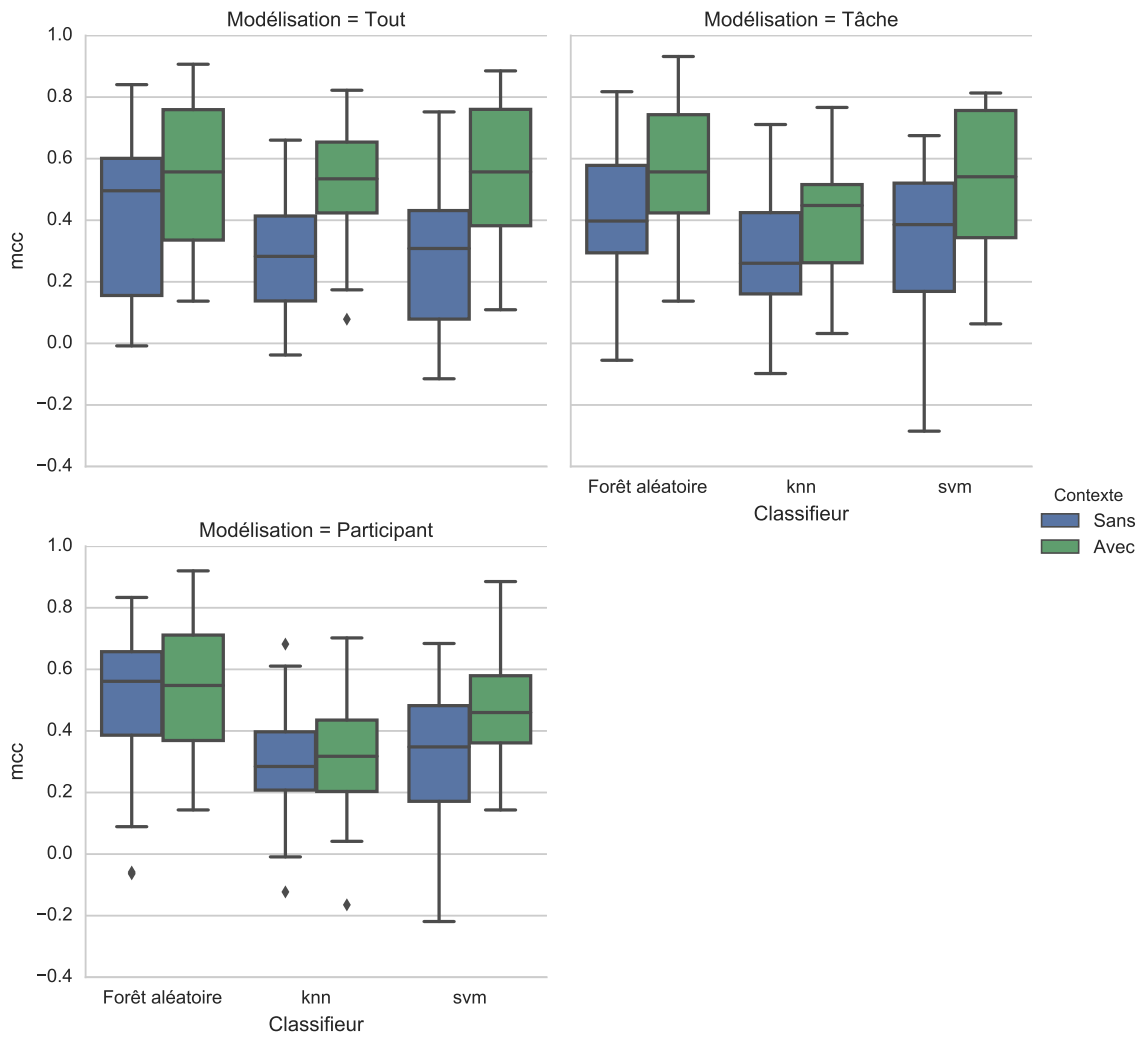


FIGURE 4.14 – Distribution des MCC avec ou sans contexte.

# Conclusion

Ce projet avait pour but d'évaluer la possibilité d'utiliser l'apprentissage automatique pour estimer l'état fonctionnel de l'opérateur. Pour ce faire, une définition de l'OFS ainsi que des objectifs de recherche plus précis sont énoncés au chapitre 1. Ensuite, au chapitre 2, les principes de l'apprentissage automatique sont exposés. Au chapitre 3, la façon dont celui-ci est utilisé pour estimer l'OFS est décrite. Finalement, au chapitre 4, les résultats et les performances obtenus lors des expériences d'estimation de l'OFS sont présentés.

## Formalisation de l'OFS

Le premier objectif de recherche, qui est présenté au chapitre 1, section 1.5.1, est de formuler le problème de l'estimation de l'état fonctionnel de l'opérateur d'une manière permettant la comparaison dans plusieurs contextes.

Dans ce projet, une méthode d'estimation de l'OFS à partir de la performance, la performance dynamique décontextualisée, est proposée. Cette technique utilise la performance d'un opérateur dans une tâche pour estimer l'état de celui-ci. La DDP est décontextualisée puisque cet état est indépendant de la tâche ou du participant. Pour cette raison, celle-ci peut être employée pour l'évaluation des performances des algorithmes d'apprentissage automatique.

Pour montrer que l'objectif de formalisation de l'OFS en une mesure pouvant être utilisée pour des expériences d'apprentissage automatique, nous devons mettre en lumière plusieurs éléments de ce mémoire. Premièrement, l'utilisation simultanée de données provenant de plusieurs tâches expérimentales séparées montre que la DDP permet de mesurer un concept général d'OFS qui ne se limite pas à une seule tâche ou à un seul participant.

De plus, comme le montrent les résultats du chapitre 4, particulièrement ceux de la

section 4.3.1, les performances des algorithmes d'apprentissage automatique supervisé qui utilisent la DDP comme signal d'apprentissage ont une performance semblable lorsqu'ils modélisent un seul participant, une seule tâche à la fois ou l'ensemble des participants et des tâches.

Ces éléments permettent donc de conclure que l'objectif de la formalisation de l'OFS est rempli par la présentation de la DDP.

## Évaluation de l'apprentissage automatique

Le deuxième objectif, qui est présenté au chapitre 1, section 1.5.2, est d'évaluer la performance des algorithmes d'apprentissage automatique pour estimer l'OFS en utilisant les données physiologiques et psychologiques ainsi que d'évaluer le meilleur algorithme pour ce faire.

Les résultats présentés dans le chapitre 4 montrent la performance obtenue lors de l'estimation de l'OFS par apprentissage automatique. Ces résultats mettent en lumière la possibilité d'utiliser que l'apprentissage automatique pour estimer l'OFS. De plus, ces résultats indiquent également que le meilleur classifieur permettant cette estimation est la forêt d'arbres aléatoires.

Cependant, bien que ces résultats suggèrent qu'il soit possible d'utiliser l'apprentissage automatique pour estimer l'état fonctionnel de l'opérateur, ceux-ci illustrent également que la précision de cette estimation est limitée. D'autres travaux devront être réalisés pour permettre d'améliorer la précision de cette estimation afin qu'elle soit suffisamment élevée pour une utilisation dans un contexte pratique.

## Généralisation

Tel que présenté au chapitre 1, section 1.5.3, le troisième objectif est d'évaluer la généralisation des algorithmes d'apprentissage automatique qui sont utilisés. Bien que les algorithmes utilisés permettent d'estimer l'état fonctionnel de l'opérateur, est-ce que ces mêmes algorithmes permettent d'obtenir une estimation de l'OFS sur de nouveaux participants ou sur des données provenant de nouvelles tâches ?

Comme le montrent les résultats du chapitre 4, la généralisation des algorithmes et des données utilisées est limitée. En effet, les performances obtenues lors des expériences sur la généralisation sont très près de celles obtenues en classifiant les données au

hasard. Cette performance en généralisation indique que les techniques utilisées ne permettent pas d'extraire des indicateurs généraux permettant d'estimer l'OFS sur plusieurs participants.

## Utilisation du contexte

Le quatrième objectif, présenté au chapitre 1 section 1.5.4, est d'évaluer l'effet du contexte d'une tâche en cours sur la performance de l'estimation de l'état fonctionnel de l'opérateur.

Les résultats du chapitre 4 montrent que l'utilisation du contexte, tel que la charge de travail imposé et le moment de la journée où la tâche est effectuée permettent de prédire la DDP beaucoup mieux que sans cette information. Bien que l'utilisation du contexte soit plus difficile à mettre en œuvre dans un contexte pratique, le fait qu'il soit possible d'augmenter les performances en généralisation à l'aide du contexte laisse croire qu'il serait possible de développer des techniques pour intégrer des informations semblables dans de nouvelles applications.

## Conclusion et travaux futurs

Les travaux effectués dans le cadre de ce projet mettent en lumière les conclusions suivantes. L'utilisation de l'apprentissage automatique pour l'estimation de l'état fonctionnel de l'opérateur est possible. Cependant, pour obtenir des performances permettant d'utiliser l'apprentissage automatique dans un contexte pratique, des travaux futurs devront être effectués.

Premièrement, pour avoir confiance envers la prédiction obtenue par les différents modèles, il est nécessaire d'augmenter la performance pouvant être obtenue par ceux-ci. Il est cependant important de considérer le fait qu'il est peut-être impossible d'obtenir une performance parfaite et que la performance maximale d'estimation de l'OFS est inconnue. Deuxièmement, la capacité de transférer un modèle à de nouvelles tâches et de nouveaux participants devra être développée.

Deux avenues de recherches peuvent être développées pour permettre ces avancées. La première est l'utilisation de l'information temporelle des données qui pourrait permettre d'augmenter les performances et la généralisation des modèles d'apprentissage automatique. Il serait également pertinent d'augmenter la quantité de données utilisées



en augmentant le nombre de participants et de sessions expérimentales. La deuxième est l'utilisation d'autres signaux physiologiques qui peuvent potentiellement fournir de nouvelles informations pouvant être exploitées par les modèles d'apprentissage automatique.

# Annexe A

## Code

Le code utilisé pour le calcul de la DDP est disponible à l'adresse suivante :

<https://github.com/leconteur/DDP>

Le code utilisé pour les expériences d'apprentissage automatique est disponible à l'adresse suivante :

[https://github.com/leconteur/experiences\\_memoire](https://github.com/leconteur/experiences_memoire)

Les tâches expérimentales sont disponibles à l'adresse suivante :

<https://github.com/leconteur/psychotasks>

# Annexe B

## Données

Le tableau B.1 présente les signaux qui sont utilisés lors des expériences d'apprentissage automatique. Ces données sont disponibles à l'adresse suivante :

[https://github.com/leconteur/experiences\\_memoire/tree/master/datasets\\_begin](https://github.com/leconteur/experiences_memoire/tree/master/datasets_begin)

TABLE B.1 – Valeurs utilisés pour les expériences d’apprentissage automatique

Signal	Description
HR	Rythme cardiaque
BR	Rythme respiratoire
Posture	Posture en degré
BRAmplitude	Amplitude de la respiration
ECGAmplitude	Amplitude de l’ECG
CoreTemp	Température corporelle estimé par le BioHarness
HRV	Variance du rythme cardiaque
ULF	Ultra-basses fréquences du rythme cardiaque
VLF	Très-basses fréquences du rythme cardiaque
LF	basses fréquences du rythme cardiaque
HF	hautes fréquences du rythme cardiaque
LFHF	Ratio des basses fréquences sur les hautes fréquences
avg_x	Position horizontale du regard dans l’écran
avg_y	Position verticale du regard dans l’écran
fix	Fixation du regard
lefteye_psize	Taille de la pupille gauche
righteye_psize	Taille de la pupille droite
eye_velocity	Velocité du regard en pixel
blink_mean_60_seconds	Moyenne du nombre de clignement sur une minute
blink_mean_10_seconds	Moyenne du nombre de clignement sur dix secondes
saccade_mean_60_seconds	Moyenne du temps en saccade sur une minute
saccade_mean_10_seconds	Moyenne du temps en saccade sur dix secondes
fixation_mean_60_seconds	Moyenne du temps en fixation sur une minute
fixation_mean_10_seconds	Moyenne du temps en fixation sur dix secondes
involuntary_fixation_mean_60_seconds	Moyenne du temps en fixation involontaire sur une minute
involuntary_fixation_mean_10_seconds	Moyenne du temps en fixation involontaire sur dix secondes

# Annexe C

## Résultats supplémentaires

Modélisation	participant	Tâche	Classifieur	mcc
Tâche	p01	Rotation mentale	Foret aleatoire	0.25
Tâche	p01	N-Back	Foret aleatoire	0.78
Tâche	p01	Recherche visuelle	Foret aleatoire	0.31
Tâche	p02	Rotation mentale	Foret aleatoire	0.59
Tâche	p02	N-Back	Foret aleatoire	0.37
Tâche	p02	Recherche visuelle	Foret aleatoire	0.21
Tâche	p03	Rotation mentale	Foret aleatoire	0.53
Tâche	p03	N-Back	Foret aleatoire	0.49
Tâche	p03	Recherche visuelle	Foret aleatoire	0.80
Tâche	p04	Rotation mentale	Foret aleatoire	0.32
Tâche	p04	N-Back	Foret aleatoire	0.30
Tâche	p04	Recherche visuelle	Foret aleatoire	0.58
Tâche	p05	Rotation mentale	Foret aleatoire	-0.05
Tâche	p05	N-Back	Foret aleatoire	0.20
Tâche	p05	Recherche visuelle	Foret aleatoire	0.43
Tâche	p06	Rotation mentale	Foret aleatoire	0.28
Tâche	p06	N-Back	Foret aleatoire	0.56
Tâche	p06	Recherche visuelle	Foret aleatoire	0.36
Tâche	p07	Rotation mentale	Foret aleatoire	0.08
Tâche	p07	N-Back	Foret aleatoire	0.58
Tâche	p07	Recherche visuelle	Foret aleatoire	0.64
Tâche	p08	Rotation mentale	Foret aleatoire	0.31
Tâche	p08	N-Back	Foret aleatoire	0.82
Tâche	p08	Recherche visuelle	Foret aleatoire	0.45
Tâche	p01	Rotation mentale	svm	0.19
Tâche	p01	N-Back	svm	0.67
Tâche	p01	Recherche visuelle	svm	0.60
Tâche	p02	Rotation mentale	svm	0.05
Tâche	p02	N-Back	svm	0.10
Tâche	p02	Recherche visuelle	svm	0.29
Tâche	p03	Rotation mentale	svm	0.38
Tâche	p03	N-Back	svm	0.51
Tâche	p03	Recherche visuelle	svm	0.67
Tâche	p04	Rotation mentale	svm	0.18
Tâche	p04	N-Back	svm	-0.29

Modélisation	participant	Tâche	Classifieur	mcc
Tâche	p04	Recherche visuelle	svm	0.58
Tâche	p05	Rotation mentale	svm	-0.10
Tâche	p05	N-Back	svm	0.12
Tâche	p05	Recherche visuelle	svm	0.15
Tâche	p06	Rotation mentale	svm	0.29
Tâche	p06	N-Back	svm	0.50
Tâche	p06	Recherche visuelle	svm	0.55
Tâche	p07	Rotation mentale	svm	0.20
Tâche	p07	N-Back	svm	0.39
Tâche	p07	Recherche visuelle	svm	0.39
Tâche	p08	Rotation mentale	svm	0.43
Tâche	p08	N-Back	svm	0.62
Tâche	p08	Recherche visuelle	svm	0.49
Tâche	p01	Rotation mentale	knn	0.11
Tâche	p01	N-Back	knn	0.44
Tâche	p01	Recherche visuelle	knn	0.36
Tâche	p02	Rotation mentale	knn	-0.05
Tâche	p02	N-Back	knn	0.18
Tâche	p02	Recherche visuelle	knn	0.19
Tâche	p03	Rotation mentale	knn	0.35
Tâche	p03	N-Back	knn	0.24
Tâche	p03	Recherche visuelle	knn	0.71
Tâche	p04	Rotation mentale	knn	0.23
Tâche	p04	N-Back	knn	-0.10
Tâche	p04	Recherche visuelle	knn	0.37
Tâche	p05	Rotation mentale	knn	-0.02
Tâche	p05	N-Back	knn	-0.04
Tâche	p05	Recherche visuelle	knn	0.25
Tâche	p06	Rotation mentale	knn	0.23
Tâche	p06	N-Back	knn	0.58
Tâche	p06	Recherche visuelle	knn	0.42
Tâche	p07	Rotation mentale	knn	-0.06
Tâche	p07	N-Back	knn	0.46
Tâche	p07	Recherche visuelle	knn	0.27
Tâche	p08	Rotation mentale	knn	0.48
Tâche	p08	N-Back	knn	0.46
Tâche	p08	Recherche visuelle	knn	0.29
Participant	p01	Rotation mentale	Foret aleatoire	0.15
Participant	p01	N-Back	Foret aleatoire	0.73
Participant	p01	Recherche visuelle	Foret aleatoire	0.40
Participant	p02	Rotation mentale	Foret aleatoire	0.67
Participant	p02	N-Back	Foret aleatoire	0.48
Participant	p02	Recherche visuelle	Foret aleatoire	0.65
Participant	p03	Rotation mentale	Foret aleatoire	0.50
Participant	p03	N-Back	Foret aleatoire	0.69
Participant	p03	Recherche visuelle	Foret aleatoire	0.83
Participant	p04	Rotation mentale	Foret aleatoire	0.41
Participant	p04	N-Back	Foret aleatoire	0.55
Participant	p04	Recherche visuelle	Foret aleatoire	0.59
Participant	p05	Rotation mentale	Foret aleatoire	0.09
Participant	p05	N-Back	Foret aleatoire	0.34
Participant	p05	Recherche visuelle	Foret aleatoire	0.58
Participant	p06	Rotation mentale	Foret aleatoire	-0.06

Modélisation	participant	Tâche	Classifieur	mcc
Participant	p06	N-Back	Foret aleatoire	0.65
Participant	p06	Recherche visuelle	Foret aleatoire	0.40
Participant	p07	Rotation mentale	Foret aleatoire	-0.06
Participant	p07	N-Back	Foret aleatoire	0.64
Participant	p07	Recherche visuelle	Foret aleatoire	0.71
Participant	p08	Rotation mentale	Foret aleatoire	0.35
Participant	p08	N-Back	Foret aleatoire	0.61
Participant	p08	Recherche visuelle	Foret aleatoire	0.69
Participant	p01	Rotation mentale	svm	0.08
Participant	p01	N-Back	svm	0.29
Participant	p01	Recherche visuelle	svm	0.32
Participant	p02	Rotation mentale	svm	0.14
Participant	p02	N-Back	svm	-0.18
Participant	p02	Recherche visuelle	svm	0.52
Participant	p03	Rotation mentale	svm	0.50
Participant	p03	N-Back	svm	0.48
Participant	p03	Recherche visuelle	svm	0.46
Participant	p04	Rotation mentale	svm	0.39
Participant	p04	N-Back	svm	0.23
Participant	p04	Recherche visuelle	svm	0.39
Participant	p05	Rotation mentale	svm	-0.22
Participant	p05	N-Back	svm	0.28
Participant	p05	Recherche visuelle	svm	0.38
Participant	p06	Rotation mentale	svm	0.06
Participant	p06	N-Back	svm	0.68
Participant	p06	Recherche visuelle	svm	0.44
Participant	p07	Rotation mentale	svm	0.26
Participant	p07	N-Back	svm	0.56
Participant	p07	Recherche visuelle	svm	0.09
Participant	p08	Rotation mentale	svm	0.53
Participant	p08	N-Back	svm	0.65
Participant	p08	Recherche visuelle	svm	0.18
Participant	p01	Rotation mentale	knn	0.35
Participant	p01	N-Back	knn	0.25
Participant	p01	Recherche visuelle	knn	0.02
Participant	p02	Rotation mentale	knn	0.11
Participant	p02	N-Back	knn	0.21
Participant	p02	Recherche visuelle	knn	0.49
Participant	p03	Rotation mentale	knn	0.38
Participant	p03	N-Back	knn	0.46
Participant	p03	Recherche visuelle	knn	0.40
Participant	p04	Rotation mentale	knn	0.15
Participant	p04	N-Back	knn	0.21
Participant	p04	Recherche visuelle	knn	0.41
Participant	p05	Rotation mentale	knn	-0.12
Participant	p05	N-Back	knn	0.27
Participant	p05	Recherche visuelle	knn	0.34
Participant	p06	Rotation mentale	knn	-0.01
Participant	p06	N-Back	knn	0.68
Participant	p06	Recherche visuelle	knn	0.27
Participant	p07	Rotation mentale	knn	0.29
Participant	p07	N-Back	knn	0.40
Participant	p07	Recherche visuelle	knn	0.22

Modélisation	participant	Tâche	Classifieur	mcc
Participant	p08	Rotation mentale	knn	0.38
Participant	p08	N-Back	knn	0.61
Participant	p08	Recherche visuelle	knn	0.28
Tout	p01	Rotation mentale	Foret aleatoire	0.15
Tout	p01	N-Back	Foret aleatoire	0.69
Tout	p01	Recherche visuelle	Foret aleatoire	0.29
Tout	p02	Rotation mentale	Foret aleatoire	0.60
Tout	p02	N-Back	Foret aleatoire	-0.01
Tout	p02	Recherche visuelle	Foret aleatoire	0.32
Tout	p03	Rotation mentale	Foret aleatoire	0.15
Tout	p03	N-Back	Foret aleatoire	0.16
Tout	p03	Recherche visuelle	Foret aleatoire	0.77
Tout	p04	Rotation mentale	Foret aleatoire	0.50
Tout	p04	N-Back	Foret aleatoire	0.01
Tout	p04	Recherche visuelle	Foret aleatoire	0.70
Tout	p05	Rotation mentale	Foret aleatoire	0.05
Tout	p05	N-Back	Foret aleatoire	0.15
Tout	p05	Recherche visuelle	Foret aleatoire	0.48
Tout	p06	Rotation mentale	Foret aleatoire	0.27
Tout	p06	N-Back	Foret aleatoire	0.84
Tout	p06	Recherche visuelle	Foret aleatoire	0.56
Tout	p07	Rotation mentale	Foret aleatoire	0.49
Tout	p07	N-Back	Foret aleatoire	0.58
Tout	p07	Recherche visuelle	Foret aleatoire	0.60
Tout	p08	Rotation mentale	Foret aleatoire	0.52
Tout	p08	N-Back	Foret aleatoire	0.77
Tout	p08	Recherche visuelle	Foret aleatoire	0.53
Tout	p01	Rotation mentale	svm	0.08
Tout	p01	N-Back	svm	0.65
Tout	p01	Recherche visuelle	svm	-0.01
Tout	p02	Rotation mentale	svm	0.09
Tout	p02	N-Back	svm	0.05
Tout	p02	Recherche visuelle	svm	0.41
Tout	p03	Rotation mentale	svm	0.44
Tout	p03	N-Back	svm	0.40
Tout	p03	Recherche visuelle	svm	0.54
Tout	p04	Rotation mentale	svm	0.24
Tout	p04	N-Back	svm	0.07
Tout	p04	Recherche visuelle	svm	0.45
Tout	p05	Rotation mentale	svm	-0.11
Tout	p05	N-Back	svm	0.04
Tout	p05	Recherche visuelle	svm	0.32
Tout	p06	Rotation mentale	svm	-0.01
Tout	p06	N-Back	svm	0.57
Tout	p06	Recherche visuelle	svm	0.27
Tout	p07	Rotation mentale	svm	0.36
Tout	p07	N-Back	svm	0.42
Tout	p07	Recherche visuelle	svm	0.28
Tout	p08	Rotation mentale	svm	0.43
Tout	p08	N-Back	svm	0.75
Tout	p08	Recherche visuelle	svm	0.30
Tout	p01	Rotation mentale	knn	0.14
Tout	p01	N-Back	knn	0.43



Modélisation	participant	Tâche	Classifieur	mcc
Tout	p01	Recherche visuelle	knn	0.36
Tout	p02	Rotation mentale	knn	0.12
Tout	p02	N-Back	knn	0.19
Tout	p02	Recherche visuelle	knn	0.23
Tout	p03	Rotation mentale	knn	0.35
Tout	p03	N-Back	knn	0.21
Tout	p03	Recherche visuelle	knn	0.66
Tout	p04	Rotation mentale	knn	0.06
Tout	p04	N-Back	knn	0.23
Tout	p04	Recherche visuelle	knn	0.41
Tout	p05	Rotation mentale	knn	0.06
Tout	p05	N-Back	knn	0.04
Tout	p05	Recherche visuelle	knn	0.37
Tout	p06	Rotation mentale	knn	-0.04
Tout	p06	N-Back	knn	0.50
Tout	p06	Recherche visuelle	knn	0.19
Tout	p07	Rotation mentale	knn	0.38
Tout	p07	N-Back	knn	0.51
Tout	p07	Recherche visuelle	knn	0.07
Tout	p08	Rotation mentale	knn	0.45
Tout	p08	N-Back	knn	0.57
Tout	p08	Recherche visuelle	knn	0.33
Généralisation par participant	p01	Rotation mentale	Foret aleatoire	-0.18
Généralisation par participant	p01	N-Back	Foret aleatoire	0.12
Généralisation par participant	p01	Recherche visuelle	Foret aleatoire	-0.14
Généralisation par participant	p02	Rotation mentale	Foret aleatoire	0.52
Généralisation par participant	p02	N-Back	Foret aleatoire	-0.13
Généralisation par participant	p02	Recherche visuelle	Foret aleatoire	-0.32
Généralisation par participant	p03	Rotation mentale	Foret aleatoire	-0.22
Généralisation par participant	p03	N-Back	Foret aleatoire	-0.24
Généralisation par participant	p03	Recherche visuelle	Foret aleatoire	0.10
Généralisation par participant	p04	Rotation mentale	Foret aleatoire	-0.03
Généralisation par participant	p04	N-Back	Foret aleatoire	-0.40
Généralisation par participant	p04	Recherche visuelle	Foret aleatoire	0.22
Généralisation par participant	p05	Rotation mentale	Foret aleatoire	0.00
Généralisation par participant	p05	N-Back	Foret aleatoire	-0.40
Généralisation par participant	p05	Recherche visuelle	Foret aleatoire	-0.19
Généralisation par participant	p06	Rotation mentale	Foret aleatoire	0.07
Généralisation par participant	p06	N-Back	Foret aleatoire	0.04
Généralisation par participant	p06	Recherche visuelle	Foret aleatoire	0.35
Généralisation par participant	p07	Rotation mentale	Foret aleatoire	0.55
Généralisation par participant	p07	N-Back	Foret aleatoire	-0.09
Généralisation par participant	p07	Recherche visuelle	Foret aleatoire	0.45
Généralisation par participant	p08	Rotation mentale	Foret aleatoire	0.16
Généralisation par participant	p08	N-Back	Foret aleatoire	0.52
Généralisation par participant	p08	Recherche visuelle	Foret aleatoire	-0.19
Généralisation par participant	p01	Rotation mentale	svm	0.09
Généralisation par participant	p01	N-Back	svm	-0.21
Généralisation par participant	p01	Recherche visuelle	svm	-0.10
Généralisation par participant	p02	Rotation mentale	svm	-0.14
Généralisation par participant	p02	N-Back	svm	0.51
Généralisation par participant	p02	Recherche visuelle	svm	0.05
Généralisation par participant	p03	Rotation mentale	svm	-0.26

Modélisation	participant	Tâche	Classifieur	mcc
Généralisation par participant	p03	N-Back	svm	-0.01
Généralisation par participant	p03	Recherche visuelle	svm	-0.03
Généralisation par participant	p04	Rotation mentale	svm	0.12
Généralisation par participant	p04	N-Back	svm	-0.05
Généralisation par participant	p04	Recherche visuelle	svm	0.04
Généralisation par participant	p05	Rotation mentale	svm	-0.36
Généralisation par participant	p05	N-Back	svm	0.00
Généralisation par participant	p05	Recherche visuelle	svm	0.10
Généralisation par participant	p06	Rotation mentale	svm	0.22
Généralisation par participant	p06	N-Back	svm	0.09
Généralisation par participant	p06	Recherche visuelle	svm	0.19
Généralisation par participant	p07	Rotation mentale	svm	0.60
Généralisation par participant	p07	N-Back	svm	0.08
Généralisation par participant	p07	Recherche visuelle	svm	0.11
Généralisation par participant	p08	Rotation mentale	svm	0.06
Généralisation par participant	p08	N-Back	svm	0.19
Généralisation par participant	p08	Recherche visuelle	svm	0.07
Généralisation par participant	p01	Rotation mentale	knn	0.02
Généralisation par participant	p01	N-Back	knn	0.20
Généralisation par participant	p01	Recherche visuelle	knn	0.14
Généralisation par participant	p02	Rotation mentale	knn	-0.17
Généralisation par participant	p02	N-Back	knn	0.00
Généralisation par participant	p02	Recherche visuelle	knn	-0.13
Généralisation par participant	p03	Rotation mentale	knn	-0.16
Généralisation par participant	p03	N-Back	knn	-0.16
Généralisation par participant	p03	Recherche visuelle	knn	-0.08
Généralisation par participant	p04	Rotation mentale	knn	0.07
Généralisation par participant	p04	N-Back	knn	-0.01
Généralisation par participant	p04	Recherche visuelle	knn	-0.01
Généralisation par participant	p05	Rotation mentale	knn	-0.22
Généralisation par participant	p05	N-Back	knn	-0.16
Généralisation par participant	p05	Recherche visuelle	knn	-0.01
Généralisation par participant	p06	Rotation mentale	knn	0.10
Généralisation par participant	p06	N-Back	knn	-0.05
Généralisation par participant	p06	Recherche visuelle	knn	0.14
Généralisation par participant	p07	Rotation mentale	knn	0.58
Généralisation par participant	p07	N-Back	knn	-0.14
Généralisation par participant	p07	Recherche visuelle	knn	0.22
Généralisation par participant	p08	Rotation mentale	knn	0.08
Généralisation par participant	p08	N-Back	knn	0.19
Généralisation par participant	p08	Recherche visuelle	knn	0.19
Généralisation par tâche	p01	Rotation mentale	Foret aleatoire	0.18
Généralisation par tâche	p01	N-Back	Foret aleatoire	0.19
Généralisation par tâche	p01	Recherche visuelle	Foret aleatoire	0.12
Généralisation par tâche	p02	Rotation mentale	Foret aleatoire	0.50
Généralisation par tâche	p02	N-Back	Foret aleatoire	-0.00
Généralisation par tâche	p02	Recherche visuelle	Foret aleatoire	-0.14
Généralisation par tâche	p03	Rotation mentale	Foret aleatoire	0.16
Généralisation par tâche	p03	N-Back	Foret aleatoire	-0.36
Généralisation par tâche	p03	Recherche visuelle	Foret aleatoire	-0.10
Généralisation par tâche	p04	Rotation mentale	Foret aleatoire	0.46
Généralisation par tâche	p04	N-Back	Foret aleatoire	-0.14
Généralisation par tâche	p04	Recherche visuelle	Foret aleatoire	-0.08

Modélisation	participant	Tâche	Classifieur	mcc
Généralisation par tâche	p05	Rotation mentale	Foret aleatoire	-0.01
Généralisation par tâche	p05	N-Back	Foret aleatoire	-0.09
Généralisation par tâche	p05	Recherche visuelle	Foret aleatoire	-0.20
Généralisation par tâche	p06	Rotation mentale	Foret aleatoire	0.01
Généralisation par tâche	p06	N-Back	Foret aleatoire	0.16
Généralisation par tâche	p06	Recherche visuelle	Foret aleatoire	-0.17
Généralisation par tâche	p07	Rotation mentale	Foret aleatoire	0.24
Généralisation par tâche	p07	N-Back	Foret aleatoire	-0.22
Généralisation par tâche	p07	Recherche visuelle	Foret aleatoire	0.08
Généralisation par tâche	p08	Rotation mentale	Foret aleatoire	0.05
Généralisation par tâche	p08	N-Back	Foret aleatoire	0.17
Généralisation par tâche	p08	Recherche visuelle	Foret aleatoire	-0.35
Généralisation par tâche	p01	Rotation mentale	svm	0.19
Généralisation par tâche	p01	N-Back	svm	-0.09
Généralisation par tâche	p01	Recherche visuelle	svm	-0.29
Généralisation par tâche	p02	Rotation mentale	svm	0.32
Généralisation par tâche	p02	N-Back	svm	0.11
Généralisation par tâche	p02	Recherche visuelle	svm	-0.03
Généralisation par tâche	p03	Rotation mentale	svm	0.17
Généralisation par tâche	p03	N-Back	svm	-0.31
Généralisation par tâche	p03	Recherche visuelle	svm	-0.23
Généralisation par tâche	p04	Rotation mentale	svm	-0.03
Généralisation par tâche	p04	N-Back	svm	0.02
Généralisation par tâche	p04	Recherche visuelle	svm	-0.06
Généralisation par tâche	p05	Rotation mentale	svm	-0.01
Généralisation par tâche	p05	N-Back	svm	0.27
Généralisation par tâche	p05	Recherche visuelle	svm	0.07
Généralisation par tâche	p06	Rotation mentale	svm	0.37
Généralisation par tâche	p06	N-Back	svm	-0.01
Généralisation par tâche	p06	Recherche visuelle	svm	-0.22
Généralisation par tâche	p07	Rotation mentale	svm	0.30
Généralisation par tâche	p07	N-Back	svm	0.27
Généralisation par tâche	p07	Recherche visuelle	svm	-0.04
Généralisation par tâche	p08	Rotation mentale	svm	-0.41
Généralisation par tâche	p08	N-Back	svm	-0.12
Généralisation par tâche	p08	Recherche visuelle	svm	-0.08
Généralisation par tâche	p01	Rotation mentale	knn	0.16
Généralisation par tâche	p01	N-Back	knn	0.13
Généralisation par tâche	p01	Recherche visuelle	knn	0.01
Généralisation par tâche	p02	Rotation mentale	knn	0.09
Généralisation par tâche	p02	N-Back	knn	-0.32
Généralisation par tâche	p02	Recherche visuelle	knn	-0.07
Généralisation par tâche	p03	Rotation mentale	knn	0.06
Généralisation par tâche	p03	N-Back	knn	-0.44
Généralisation par tâche	p03	Recherche visuelle	knn	-0.03
Généralisation par tâche	p04	Rotation mentale	knn	-0.09
Généralisation par tâche	p04	N-Back	knn	-0.02
Généralisation par tâche	p04	Recherche visuelle	knn	-0.14
Généralisation par tâche	p05	Rotation mentale	knn	0.19
Généralisation par tâche	p05	N-Back	knn	0.29
Généralisation par tâche	p05	Recherche visuelle	knn	0.21
Généralisation par tâche	p06	Rotation mentale	knn	-0.02
Généralisation par tâche	p06	N-Back	knn	0.18

Modélisation	participant	Tâche	Classifieur	mcc
Généralisation par tâche	p06	Recherche visuelle	knn	-0.25
Généralisation par tâche	p07	Rotation mentale	knn	0.23
Généralisation par tâche	p07	N-Back	knn	-0.15
Généralisation par tâche	p07	Recherche visuelle	knn	0.11
Généralisation par tâche	p08	Rotation mentale	knn	-0.17
Généralisation par tâche	p08	N-Back	knn	-0.03
Généralisation par tâche	p08	Recherche visuelle	knn	-0.00

TABLE C.1 – Résultats complet des expériences d'apprentissage automatique pour tous les participants, tâches, classifieurs et modélisations indépendant.

# Bibliographie

- [1] U. Rajendra ACHARYA, K. Paul JOSEPH, N. KANNATHAL, Choo Min LIM et Jasjit S. SURI : Heart rate variability : A review, 2006. ISSN 01400118.
- [2] John ADOLFSON et Thomas BERGHAGE : *Perception and performance under water*. John Wiley & Sons, 1974.
- [3] Torbjorn AKERSTEDT : Work hours and operator functional status. *NATO Science Series Sub Series I Life and Behavioural Sciences*, 355 :65–80, 2003.
- [4] Ethem ALPAYDIN : *Introduction to Machine Learning Adaptive Computation and Machine Learning*. MIT Press, 2014. ISBN 0262012111.
- [5] Haruhiko ASADA et Jean-Jacques E. SLOTINE : *Robot analysis and control*. John Wiley & Sons, 1988. ISBN 0471830291.
- [6] James BERGSTRA et Yoshua BENGIO : Random Search for Hyper-Parameter Optimization. *Journal of Machine Learning Research*, 13 :281–305, 2012. ISSN 1532-4435.
- [7] Malcolm A BONNER et Glenn F WILSON : Heart Rate Measures of Flight Test and Evaluation. *International Jnl. of Aviation Psychology*, 12(1) :63–77, 2001.
- [8] Sansanee BOONNITHI et Sukanya PHONGSUPHAP : Comparison of Heart Rate Variability Measures for Mental Stress Detection. *In 2011 Computing in Cardiology*, pages 85–88. IEEE, 2011.
- [9] Leo BREIMAN : Random Forests. *Machine Learning*, 45(1) :5–32, 2001. ISSN 1573-0565. URL <http://dx.doi.org/10.1023/A:1010933404324>.
- [10] Anne-Marie BROUWER, Thorsten O. ZANDER, Jan B. F. van ERP, Johannes E. KORTELING et Adelbert W. BRONKHORST : Using neurophysiological signals that

- reflect cognitive or affective state : six recommendations to avoid common pitfalls. *Frontiers in Neuroscience*, 9(April) :1–11, 2015. ISSN 1662-453X.
- [11] Daniel BRUNEAU, M. Angela SASSE et John MCCARTHY : The Eyes Never Lie : The Use of Eye Tracking Data in HCI Research. *Proceedings of the CHI*, page 25, 2002.
- [12] Sarah E BULLARD, Mélina GRISS, Sonia GREENE et Anna GEKKER : Encyclopedia of Clinical Neuropsychology. *Archives of Clinical Neuropsychology*, 28(1), 2012. ISSN 0887-6177. URL <http://acn.oxfordjournals.org/cgi/doi/10.1093/arclin/acs103>.
- [13] Scott C. BUNCE, Kurtulus IZZETOGLU, Hasan AYAZ, Patricia SHEWOKIS, Meltem IZZETOGLU, Kambiz POURREZAEI et Banu ONARAL : *Implementation of fNIRS for Monitoring Levels of Expertise and Mental Workload*, pages 13–22. Springer Berlin Heidelberg, Berlin, Heidelberg, 2011. ISBN 978-3-642-21852-1. URL [http://dx.doi.org/10.1007/978-3-642-21852-1\\_2](http://dx.doi.org/10.1007/978-3-642-21852-1_2).
- [14] Jordan a. CANNON, Pavlo a. KROKHMAL, Russell V. LENTH et Robert MURPHEY : An algorithm for online detection of temporal changes in operator cognitive state using real-time psychophysiological data. *Biomedical Signal Processing and Control*, 5(3) :229–236, jul 2010. ISSN 17468094.
- [15] Randall M CHAMBERS : Effects of acceleration on pilot performance. Rapport technique, DTIC Document, 1963.
- [16] Lynn a COOPER : Mental rotation of random two-dimensional shapes. *Cognitive Psychology*, 7 :20–43, 1975. ISSN 00100285.
- [17] Corinna CORTES et Vladimir VAPNIK : Support-vector networks. *Machine Learning*, 20(3) :273–297, 1995. ISSN 08856125.
- [18] Thomas COVER et Peter HART : Nearest neighbor pattern classification. *IEEE Transactions on Information Theory*, 13(1) :21–27, 1967. ISSN 0018-9448.
- [19] Michelle A. CRETIKOS, Rinaldo BELLOMO, Ken HILLMAN, Jack CHEN, Simon FINFER et Arthas FLABOURIS : Respiratory rate : The neglected vital sign. *Medical Journal of Australia*, 188(11) :657–659, 2008. ISSN 0025729X.
- [20] Werner J. A. DAHM : United States Air Force Report on A Vision for Air Force Science & Technology. *Technology Horizons*, 1(May) :171, 2010.

- [21] S DE VITO, E MASSERA, M PIGA, L MARTINOTTO et G DI FRANCIA : On field calibration of an electronic nose for benzene estimation in an urban pollution monitoring scenario. *Sensors and Actuators, B : Chemical*, 129(2) :750–757, 2008. ISSN 09254005.
- [22] Francesco DI NOCERA : On reliability and stability of psychophysiological indicators for assessing operator functional states. *NATO Science Series Sub Series I Life and Behavioural Sciences*, 355 :162–176, 2003.
- [23] Gautier DURANTIN, Jean-François GAGNON, Sébastien TREMBLAY et Frédéric DEHAIS : Using near infrared spectroscopy and heart rate variability to detect mental overload. *Behavioural brain research*, 259 :16–23, 2014. ISSN 1872-7549. URL <http://www.ncbi.nlm.nih.gov/pubmed/24184083>.
- [24] Kevin DURKEE, Alexandra GEYER, Scott PAPPADA, Andres ORTIZ et Scott GALSTER : Real-time workload assessment as a foundation for human performance augmentation. *In International Conference on Augmented Cognition*, pages 279–288. Springer, 2013.
- [25] Mica R. ENDSLEY : Toward a Theory of Situation Awareness in Dynamic Systems. *Human Factors : The Journal of the Human Factors and Ergonomics Society*, 37 (1) :32–64, mar 1995. ISSN 00187208.
- [26] Stephen H. FAIRCLOUGH et Louise VENABLES : Prediction of subjective states from psychophysiology : A multivariate approach. *Biological Psychology*, 71(1) :100–110, 2006. ISSN 03010511.
- [27] Manuel FERNÁNDEZ-DELGADO, Eva CERNADAS, Senén BARRO et Dinani AMORIM : Do we Need Hundreds of Classifiers to Solve Real World Classification Problems? *Journal of Machine Learning Research*, 15 :3133–3181, 2014. ISSN 1532-4435. URL <http://jmlr.org/papers/v15/delgado14a.html>.
- [28] David FERRUCCI, Anthony LEVAS, Sugato BAGCHI, David GONDEK et Erik T MUELLER : Watson : beyond jeopardy! *Artificial Intelligence*, 199 :93–105, 2013.
- [29] Evelyn FIX et JL HODGES : Discriminatory analysis. Nonparametric discrimination : Consistency properties. *Statistical Review/Revue Internationale de Statistique*, 57 :238–247, 1989. ISSN 03067734. URL <http://www.jstor.org/stable/10.2307/1403797>.

- [30] Jean-François GAGNON, Daniel LAFOND et Martin RIVEST : Sensor-Hub : A Real-Time Data Integration and Processing Nexus for Adaptive C2 Systems. *In ADAPTIVE 2014, The Sixth International Conference on Adaptive and Self-Adaptive Systems and Applications*, pages 63–67, 2014. ISBN 9781612083414.
- [31] Olivier GAGNON, Jean-François GAGNON, Daniel LAFOND et Marc PARIZEAU : Comparing Methods for Assessing Operator Functional State. *IEEE International Multi-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support (CogSIMA)*, pages 88–92, mar 2016.
- [32] Anthony W K GAILLARD : Fatigue assessment and performance protection. *NATO Science Series Sub Series I Life and Behavioural Sciences*, 355 :24–35, 2003.
- [33] Scott GALSTER et Erica JOHNSON : Sense-assess-augment : A taxonomy for human effectiveness. Rapport technique, Air Force Research Laboratory, Wright-Patterson Air Force Base, OH, USA, 2013.
- [34] Constantino A GARCÍA, Abraham OTERO, Jesús PRESEDO, Xosé VILA et Paulo FÉLIX : A software toolkit for nonlinear Heart Rate Variability analysis. *Computing in Cardiology Conference*, pages 393–396, 2013. ISSN 2325-8861.
- [35] DAVID P GRADWELL : Hypoxia and hyperventilation. *Ernsting's aviation medicine. 4th ed. London : Hodder Arnold*, pages 41–56, 2006.
- [36] Andrea HAARMANN, Wolfram BOUCSEIN et Florian SCHAEFER : Combining electrodermal responses and cardiovascular measures for probing adaptive automation during simulated flight. *Applied ergonomics*, 40(6) :1026–40, nov 2009. ISSN 1872-9126. URL <http://www.ncbi.nlm.nih.gov/pubmed/19520358>.
- [37] G Robert J HOCKEY : Changes in operator efficiency as a function of environmental stress, fatigue, and circadian rhythms. *In Handbook of human perception and performance*, pages 1–49. John Wiley & Sons, 1986.
- [38] G Robert J HOCKEY : Compensatory control in the regulation of human performance under stress and high workload : A cognitive-energetical framework. *Biological Psychology*, 45(1-3) :73–93, 1997. ISSN 03010511.
- [39] G Robert J HOCKEY : *Operator functional state as a framework for the assessment of performance degradation*. IOS Press, 2003.



- [40] G. Robert J. HOCKEY : *Operator functional state : the assessment and prediction of human performance degradation in complex tasks*, volume 355. IOS Press, Amsterdam, NL, 2003.
- [41] G Robert J HOCKEY, Peter NICKEL, Adam C ROBERTS et Michael H ROBERTS : Sensitivity of candidate markers of psychophysiological strain to cyclical changes in manual control load during simulated process control. *Applied ergonomics*, 40 (6) :1011–8, dec 2009. ISSN 1872-9126. URL <http://www.ncbi.nlm.nih.gov/pubmed/19482260>.
- [42] Robert HOCKEY : *Stress and fatigue in human performance*, volume 3. John Wiley & Sons Inc, 1983.
- [43] Maarten Andreas HOGERVORST, Anne-marie BROUWER et Jan B F Van ERP : Combining and comparing EEG , peripheral physiology and eye-related measures for the assessment of mental workload. *Frontiers in neuroscience*, 8(322) :1–14, 2014. ISSN 1662-453X.
- [44] James HORNE : *Why we sleep : the functions of sleep in humans and other mammals*. Oxford University Press, 1988.
- [45] David D. JENSEN et Paul R. COHEN : Multiple comparisons in induction algorithms. *Machine Learning*, 38(3) :309–338, 2000. ISSN 08856125.
- [46] James a. JOHNSTONE, Paul a. FORD, Gerwyn HUGHES, Tim WATSON et Andrew T. GARRETT : Bioharness™ multivariable monitoring device. Part I : Validity. *Journal of Sports Science and Medicine*, 11(3) :400–408, 2012. ISSN 13032968.
- [47] Erwin KREYSZIG : *Advanced Engineering Mathematics*. John Wiley & Sons, 2010. ISBN 9780470458365. URL <https://books.google.ca/books?id=UnN8DpXI74EC>.
- [48] Alex KRIZHEVSKY, Ilya SUTSKEVER et Geoffrey E HINTON : Imagenet classification with deep convolutional neural networks. *In Advances in neural information processing systems*, pages 1097–1105, 2012. ISBN 9781627480031.
- [49] Mohit KUMAR, Matthias WEIPPERT, Reinhard VILBRANDT, Steffi KREUZFELD et Regina STOLL : Fuzzy Evaluation of Heart Rate Signals for Mental Stress Assessment. *IEEE Transactions on Fuzzy Systems*, 15(5) :791–808, oct 2007. ISSN 1063-6706. URL <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4343114>.

- [50] Saroj KL LAL et Ashley CRAIG : A critical review of the psychophysiology of driver fatigue. *Biological psychology*, 55(3) :173–94, feb 2001. ISSN 0301-0511. URL <http://www.ncbi.nlm.nih.gov/pubmed/11240213>.
- [51] Ian LENZ, Honglak LEE et Ashutosh SAXENA : Deep learning for detecting robotic grasps. *The International Journal of Robotics Research*, 34(4-5) :705–724, apr 2015. ISSN 0278-3649.
- [52] Chun Hung LI et C K LEE : Minimum cross entropy thresholding. *Pattern Recognition*, 26(4) :617–625, 1993.
- [53] Ekaterina M. LOBACHEVA, Yulia N. GALATENKO, Rozaliya F. GABIDULLINA, Vladimir V. GALATENKO, Eugene D. LIVSHITZ, Taras P. LUKASHENKO, Dmitriy P. VETROV, Alexey E. LEBEDEV, Vladimir E. PODOL'SKII, Vyacheslav V. LEBEDEV, Sergey a. ISAYCHEV, Alexandr M. CHERNORIZOV et Yuriy P. ZINCHENKO : Automated Real-time Classification of Functional States based on Physiological Parameters. In *Procedia - Social and Behavioral Sciences*, volume 86, pages 373–378, oct 2013.
- [54] MACKWORTH, JANE F : Paced memorizing in a continuous task. *Journal of experimental psychology*, 58 :206–211, 1959. ISSN 0022-1015.
- [55] Sandra P MARSHALL : Method and apparatus for eye tracking and monitoring pupil dilation to evaluate cognitive activity, 2000.
- [56] Brian W MATTHEWS : Comparison of the predicted and observed secondary structure of T4 phage lysozyme. *BBA - Protein Structure*, 405(2) :442–451, 1975. ISSN 00052795.
- [57] Sanjit Kumar MITRA et Yonghong KUO : *Digital Signal Processing : A Computer-based Approach*. McGraw-Hill international edition electrical engineering series. McGraw-Hill/Irwin, 2001. ISBN 9780071181754. URL <https://books.google.ca/books?id=16nwGwAACAAJ>.
- [58] John B MOCHARNUK : Visual target acquisition and ocular scanning performance. *Human factors*, 20(5) :611–631, 1978. ISSN 0018-7208.
- [59] Christian MÜHL, Camille JEUNET, Fabien LOTTE et Maarten Andreas HOGERVORST : EEG-based workload estimation across affective contexts. *Frontiers in Neuroscience*, 8(June) :1–15, 2014.

- [60] Alain MUZET et Joceline ROGE : Detecting operator drowsiness during a complex task through behavioural and physiological parameters. *NATO Science Series Sub Series I Life and Behavioural Sciences*, 355 :81–89, 2003.
- [61] Mark A NEERINCX, Henk van DOORNE et Mark RUIJSENDAAL : Attuning computer-supported work to human knowledge and processing capacities in Ship Control Centres. *Cognitive task analysis*, pages 341–362, 2000.
- [62] Nobuyuki OTSU : A Threshold Selection Method from Gray-Level Histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, 9(1) :62–66, 1979. ISSN 0018-9472.
- [63] Richard L. PAGE, José A. JOGLAR, Mary A. CALDWELL, Hugh CALKINS, Jamie B. CONTI, Barbara J. DEAL, N.A. Mark ESTES, Michael E. FIELD, Zachary D. GOLDBERGER, Stephen C. HAMMILL, Julia H. INDIK, Bruce D. LINDSAY, Brian OLSHANSKY, Andrea M. RUSSO, Win-Kuang SHEN, Cynthia M. TRACY et Sana M. AL-KHATIB : 2015 ACC/AHA/HRS Guideline for the Management of Adult Patients With Supraventricular Tachycardia. *Circulation*, 64 (21) :10.1161/CIR.0000000000000311, 2015. ISSN 0009-7322.
- [64] Alexandros PANTELOPOULOS et Nikolaos G BOURBAKIS : A Survey on Wearable Sensor-Based Systems for Health Monitoring and Prognosis. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 40(1) :1–12, jan 2010. ISSN 1094-6977. URL <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5306098>.
- [65] Omkar M PARKHI, Andrea VEDALDI, Andrew ZISSERMAN et CV JAWAHAR : Cats and Dogs. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2012.
- [66] F PEDREGOSA, G VAROQUAUX, A GRAMFORT, V MICHEL, B THIRION, O GRISEL, M BLONDEL, P PRETTENHOFER, R WEISS, V DUBOURG, J VANDERPLAS, A PASSOS, D COURNAPEAU, M BRUCHER, M PERROT et E DUCHESNAY : Scikit-learn : Machine Learning in Python. *Journal of Machine Learning Research*, 12 :2825–2830, 2011.
- [67] Jonathan W. PEIRCE : PsychoPy-Psychophysics software in Python. *Journal of Neuroscience Methods*, 162 :8–13, 2007. ISSN 01650270.

- [68] Jonathan W PEIRCE : Generating Stimuli for Neuroscience Using PsychoPy. *Frontiers in neuroinformatics*, 2(January) :10, 2008. ISSN 1662-5196.
- [69] Michael PETERS et Christian BATTISTA : Applications of mental rotation figures of the Shepard and Metzler type and description of a mental rotation stimulus library. *Brain and Cognition*, 66(3) :260–264, 2008. ISSN 02782626.
- [70] David Martin W. POWERS : Evaluation : from precision, recall and F-measure to ROC, informedness, markedness and correlation. *Journal of Machine Learning Technologies*, 2(1) :37–63, 2011.
- [71] J. Ross QUINLAN : C4. 5 : programs for machine learning. *Machine Learning*, 240 :302, 1993. ISSN 08856125.
- [72] Nicolas REGIS, Frédéric DEHAIS, Catherine TESSIER et Jean-François GAGNON : Ocular metrics for detecting attentional tunnelling. *Human Factors and Ergonomics Society–Chapter Europe, Toulouse, France*, pages 121–132, 2012. URL <http://www.hfes-europe.org/books/proceedings2012/registis.pdf>.
- [73] Kirsten REINECKE, Marjolijn CORDES, Christiane LERCH, Flora KOUTSANDRÉOU, Michael SCHUBERT, Michael WEISS et Jochen BAUMEISTER : From Lab to Field Conditions : A Pilot Study on EEG Methodology in Applied Sports Sciences. *Applied Psychophysiology and Biofeedback*, 36(4) :265–271, 2011. ISSN 1573-3270. URL <http://dx.doi.org/10.1007/s10484-011-9166-x>.
- [74] Murray ROSENBLATT : Remarks on Some Nonparametric Estimates of a Density Function, 1956. ISSN 0003-4851.
- [75] Prasanna K SAHOO, SAKC SOLTANI et Andrew KC WONG : A survey of thresholding techniques, 1988. ISSN 0734189X.
- [76] MICHAEL N SAWKA : Body fluid responses and hypohydration during exercise-heat stress. *Human performance physiology and environmental medicine at terrestrial extremes*, pages 227–266, 1988.
- [77] Bernhard SCHILKOPF : The kernel trick for distances. *In Advances in Neural Information Processing Systems 13 : Proceedings of the 2000 Conference*, volume 13, page 301. MIT Press, 2001.
- [78] Fernando SEOANE, Inmaculada MOHINO-HERRANZ, Javier FERREIRA, Lorena ALVAREZ, Ruben BUENDIA, David AYLLÓN, Cosme LLERENA et Roberto GIL-PITA

- : Wearable biomedical measurement systems for assessment of mental stress of combatants in real time. *Sensors (Basel, Switzerland)*, 14(4) :7120–41, 2014. ISSN 1424-8220.
- [79] Mehmet SEZGIN et Bülent SANKUR : Survey over image thresholding techniques and quantitative performance evaluation. *Journal of Electronic imaging*, 13(1) :146–168, 2004.
- [80] Herbert A SIMON : Motivational and emotional controls of cognition. *Psychological review*, 74 :29–39, 1967. ISSN 0033-295X.
- [81] Henrik SKOVGAARD, John Paulin HANSEN et Emilie MØLLENBACH : Gaze Tracking Through Smartphones. *Workshop on Gaze Interaction in the Post-WIMP World CHI*, 2013.
- [82] AP SMITH, DAJ TYRRELL, K COYLE et JS WILLMAN : Selective effects of minor illnesses on human performance. *British Journal of Psychology*, 78 :183–188, 1987. ISSN 00071269.
- [83] Hans Von STORCH et Francis W ZWIERS : Statistical Analysis in Climate Research. *Journal of the American Statistical Association*, 95 :1375, 1999. ISSN 01621459. URL <http://ebooks.cambridge.org/ref/id/CB09780511612336>.
- [84] P TREMOULET, P BARTON, P CRAVEN, C CORRADO, G MAYER et K STIBLER : DARPA Improving warfighter information intake under stress-augmented cognition phase 3 concept validation experiment (CVE) analysis report for the Lockheed Martin ATL Team. *DARPA/IPTO, Arlington*, 2005.
- [85] Giedrius VARONECKAS : Heart rate variability in the evaluation of functional status during training. *NATO Science Series Sub Series I Life and Behavioural Sciences*, 355 :90–106, 2003.
- [86] Raofen WANG, Jianhua ZHANG, Yu ZHANG et Xingyu WANG : Assessment of human operator functional state using a novel differential evolution optimization based adaptive fuzzy model, 2011.
- [87] Glen F. WILSON : Operator Functional State Assessment. Rapport technique October 2003, North Atlantic Treaty Organisation, 2004.
- [88] Glenn F WILSON et Frank FISHER : Cognitive task classification based upon topographic EEG data. *Biological psychology*, 40(1-2) :239–50, may 1995. ISSN 0301-0511. URL <http://www.ncbi.nlm.nih.gov/pubmed/7647185>.

- [89] Glenn F WILSON, Jared D LAMBERT et Chris A RUSSELL : Performance Enhancement with Real-Time Physiologically Controlled Adaptive Aiding. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 44(13) :61–64, jul 2000. ISSN 1071-1813. URL <http://pro.sagepub.com/lookup/doi/10.1177/154193120004401316>.
- [90] Glenn F WILSON et Chris A RUSSELL : Operator Functional State Classification Using Multiple Psychophysiological Features in an Air Traffic Control Task. *Human Factors : The Journal of the Human Factors and Ergonomics Society*, 45(3) :381–389, 2003. ISSN 1547-8181.
- [91] Glenn F. WILSON et Christopher A. RUSSELL : Real-Time Assessment of Mental Workload Using Psychophysiological Measures and Artificial Neural Networks. *HUMAN FACTORS*, 45(4) :635–643, 2003. ISSN 1547-8181.
- [92] Shaozeng YANG et Jianhua ZHANG : An adaptive human–machine control system based on multiple fuzzy predictive models of operator functional state. *Biomedical Signal Processing and Control*, 8(3) :302–310, may 2013. ISSN 17468094.
- [93] Jui Cheng YEN, Fu Juay CHANG et Shyang CHANG : New criterion for automatic multilevel thresholding. *IEEE Transactions on Image Processing*, 4(3) :370–378, 1995. ISSN 10577149.
- [94] Zhong YIN et Jianhua ZHANG : Identification of temporal variations in mental workload using locally-linear-embedding-based EEG feature reduction and support-vector-machine-based clustering and classification techniques. *Computer methods and programs in biomedicine*, 115(3) :119–34, jul 2014. ISSN 1872-7565. URL <http://www.ncbi.nlm.nih.gov/pubmed/24821400>.
- [95] Jian-Hua ZHANG, Pan-Pan QIN, Jörg RAISCH, Ru-Bin WANG, ZHANG PAN-PAN, Jörg RAISCH, Ru-Bin WANG et Qin JIAN-HUA : Predictive modeling of human operator cognitive state via sparse and robust support vector machines. *Cognitive Neurodynamics*, 7(5) :395–407, oct 2013. ISSN 1871-4080.