



Contribution au développement d'un système portable automatique d'aide à la détermination de l'état d'urgence d'un blessé polytraumatisé. Apprentissage supervisé de classes de choc hémorragique.

Guillaume Jean-Paul Claude Becq

► To cite this version:

Guillaume Jean-Paul Claude Becq. Contribution au développement d'un système portable automatique d'aide à la détermination de l'état d'urgence d'un blessé polytraumatisé. Apprentissage supervisé de classes de choc hémorragique.. Autre. Université Joseph-Fourier - Grenoble I, 2004. Français. <NNT : >. <tel-00123992>

HAL Id: tel-00123992

<https://theses.hal.science/tel-00123992v1>

Submitted on 11 Jan 2007

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

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



HAL Authorization

UNIVERSITÉ JOSEPH FOURIER – GRENOBLE 1
ÉCOLE DOCTORALE INGÉNIERIE POUR LA SANTÉ, LA COGNITION ET
L'ENVIRONNEMENT (ÉDISCE)

Doctorat

Modèles et Instruments en Médecine et Biologie

Guillaume BECQ

CONTRIBUTION AU DÉVELOPPEMENT D'UN SYSTÈME
PORTABLE AUTOMATIQUE D'AIDE À LA DÉTERMINATION DE
L'ÉTAT D'URGENCE D'UN BLESSÉ POLYTRAUMATISÉ
APPRENTISSAGE SUPERVISÉ DE CLASSES DE CHOC HÉMORRAGIQUE

Thèse dirigée par : Pierre BACONNIER et Sylvie CHARBONNIER

Soutenue le 10 décembre 2004

Jury :

Président :	PR Christian Jutten (LIS, UJF Grenoble 1)
Rapporteur :	PR Bernard Dubuisson (Heudiasyc CNRS, UTC Compiègne)
Rapporteur :	PR Lotfi Senhadji (LTSI Inserm, Univ. Rennes 1)
Membre du jury :	PR Pierre Baconnier (TimC, UJF Grenoble 1)
Membre du jury :	MCF Sylvie Charbonnier (Lag, UJF Grenoble 1)
Membre du jury :	MC Lionel Bourdon (CRSSA, Grenoble)
Membre du jury :	CR Marie-Christine Chambrin (Inserm, Lille)

À François,

Remerciements

Le laboratoire seul apprend les difficultés réelles de la science à ceux qui le fréquentent, il leur montre que la science pure a toujours été la source de toutes les richesses que l'homme acquiert et de toutes les conquêtes réelles qu'il fait sur les phénomènes de la nature. C'est là en outre une excellente éducation pour la jeunesse, parce qu'elle lui fait comprendre que les applications actuelles si brillantes des sciences ne sont que l'épanouissement de travaux antérieurs, et que ceux qui, aujourd'hui, profitent de leurs bienfaits, doivent un tribut de reconnaissance à leurs devanciers qui ont péniblement cultivé l'arbre de la science sans le voir fructifier.

Claude Bernard, *Introduction à l'étude de la médecine expérimentale*, 1984 (1865), Champs Flammarion, p. 209.

Il faut avant tout remercier les personnes présentes, proches, celles qui nous influencent, nous guident, nous portent dans leur cœur et nous aident à grandir. Nous devons aussi remercier les personnes qui nous font le plaisir de soutenir ce travail. Voici une liste non exhaustive de celles ci en espérant que les personnes oubliées me pardonneront. Je tiens à remercier :

La Délégation Générale pour l'Armement (DGA) qui a financé ce projet et les correspondants qui ont bien voulu m'encadrer M. Sarron et M. Pérès. Travailler pour eux fut une réelle motivation.

Le président du jury, Christian Jutten qui m'honore de sa présence. Je tiens à le remercier particulièrement pour m'avoir fait découvrir les réseaux de neurones grâce à son ouvrage réalisé en collaboration avec Jeanny Hérault.

Les rapporteurs, Bernard Dubuisson et Lotfi Senhadji qui me font l'honneur de lire ces quelques lignes et de participer au jury.

Sylvie Charbonnier, pour son support, son tutorat. Je la remercie particulièrement pour avoir absorbé mes humeurs parfois ténébreuses, et mes visions souvent étranges de la réalité. Le chemin que nous avons parcouru nous a permis de franchir certains seuils. Son pragmatisme m'a permis de pouvoir me recadrer dans les moments de submersions totales.

Pierre Baconnier, pour son travail de formation aux modélisations appliquées à la médecine et à la biologie et pour la confiance qu'il m'a accordée durant mon DEA et ma thèse. Je le remercie encore pour ses relectures profondes et sa correction des erreurs créées par des infinitifs et participes passés du premier groupe abusifs.

Lionel Bourdon, pour son accueil chaleureux dans son laboratoire, la formation médicale qu'il m'a transmise. J'ai beaucoup appris à ses côtés : gestion d'une équipe, construc-

tion de projets, relations humaines, interdisciplinarité des sciences... Merci encore de m'avoir prêté une partie d'un bureau déjà bien rempli.

Marie-Christine Chambrin, pour avoir accepté de faire partie du jury. Je la remercie vivement pour les échanges de données, de connaissances, de tickets de trolley bus... et m'excuse encore pour les dépôts minutes devant la gare de la Part Dieu. Je tiens à remercier également l'équipe de Lille et les concepteurs de NaineForge.

M. Demongeot, pour nous faire partager sa passion, mais aussi son investissement dans tous nos travaux. Nous lui devons beaucoup.

M. Viret, pour m'avoir accueilli au sein de son établissement. Lors de notre première rencontre, juste avant ma formation militaire, M. Viret m'avait demandé *théorie* ou *pratique*. J'avais répondu *pratique*, ce qui m'avait permis de rencontrer Lionel Bourdon. J'espère que ce travail pratique permettra de mettre en valeur des notions plus fondamentales théoriques mais il me semble que les deux soient fortement liées dans ce travail. Le général Alain Buguet, pour son accueil dans son équipe et sa bonhomie communicative. Frédéric Canini pour son partage de sa connaissance sur les neurosciences et les différents chemins de vie, de la réalité à la fiction, de la science en général, et son attachement à la cuisine qui devrait être un point des plus importants dans la démarche scientifique pour l'application pratique évidente de l'observation des phénomènes physico-chimiques. Alain Queyroy pour ses nombreuses discussions critiques sur nos façons de faire et de penser et les garde barrières que nous devons nous imposer, sa grande connaissance des arts martiaux. Mon grand regret est de n'avoir pu le rencontrer sur une piste d'escrime, son ouverture d'esprit et ses aides précieuses sont utiles pour ceux qui l'écoutent. Philippe d'Aleo, pour toute son implication dans les expérimentations ainsi qu'Alain Roux et Nadine Fidier. Florian Chapotot, pour sa collaboration, sa motivation qu'il a su me faire partager, ses échanges scientifiques très fructueux. J'espère que tu ne te laisseras pas détourner de ta voie. Sarah Mercier et Céline Chevrier pour nos échanges de techniques, recettes, savoirs pratiques. La Biologie Appliquée et les zootechniciens en particulier l'Adj. chef Daniel Maître pour ses aides précieuses ainsi que Vévé et Maud, pour avoir su s'occuper des animaux avec respect, ainsi qu'une disponibilité complète lors des expérimentations et David Popieul pour ses blagues. Le service électronique Simon Pla et PAB pour leurs techniques sur mesures. Le service informatique, en particulier Philippe Loudenot et François Petit, Philippe, Stéphane, Patrick et les autres, débordés par l'inhomogénéité d'un parc informatique pour chercheurs par définition tous différents l'un de l'autre.

Mme Chérut, qui m'a longuement accueillie au sein de l'école qu'elle a su si bien diriger en collaboration avec René Feuillet et M. Baudois, mais aussi pour son constant regard vers les biotechnologies et le monde du vivant.

Loïc Biot, pour sa convivialité, son pragmatisme et son intérêt pratique porté vers la SpO_2 . L'équipe de Lille, Marie Christine Chambrin et Samir Sharshar. Pierre-Yves Carry et Jean-Pierre Perdrix pour leur accueil amical. Luc Quintin pour nous avoir fait partager ses techniques indochinoises.

Anne Sophie Silvent pour sa quiétude et ses conseils, ses orientations bibliographiques, le partage que nous avons réalisé et l'échange de connaissances durant la thèse. Benoît Froidurot parce qu'il a fait une thèse financée par la DGA, qu'il m'a aidé pour les démarches administratives et m'a bien soutenu moralement sur la fin. Ptit Lu et Gwendal Simon parce qu'ils ont aussi fait une thèse et qu'ils ont échangé de précieux conseils. Barbara pour les conseils pratiques L^AT_EX plus TP et café au labo du LIS. Mes Camarades de Pension, mes collocs, le père Du Plessix qui suivent d'autres voies intéressantes $\bar{A}$.

Ma famille, pour son soutien sans faille, sa branche forte et solide. Leur influence est incommensurable. Ils ont su me donner le goût à la vie, aux arts, sciences et techniques dans toutes leurs formes et me font partager leurs expériences. J'espère que la mienne leur est fructueuse, en tous cas, elle essaie de porter le plus d'eux en elle. Ma femme, Cécile, pour sa douceur, qu'elle me transmet dans sa façon d'être, ses visions, et ses travaux artistiques. La vie est plus agréable à ses côtés, particulièrement lorsque le pire arrive à l'improviste. Je pense qu'elle m'a fait pousser des ailes, passer de l'état chenille à papillon plutôt que de larve à mouche, mais rien n'est moins sûr. Anne, ma belle maman, pour sa douceur et sa générosité. Sébastien, Bénédicte, Juliette, Sophie (Fréd) et Marion, mes frères et sœurs chéris. Special thanks to Bénédicte et ses collaborateurs pour les rectifications anglaises !

Préface

Lors de mes premiers travaux, j'ai été confronté à des mesures de pression artérielle, d'activité motrice et de fréquence cardiaque traitées en régression par des réseaux de neurones constructifs et perceptrons multicouches. J'avais pu noter l'intérêt évident de l'association de ces petites unités et de la grande richesse des données obtenues sur le monde du vivant.

Ensuite, au cours de mon diplôme de fin d'étude, j'ai travaillé sur la mise en place d'un système capable de délivrer un médicament antihypertenseur. Cette étude m'avait permis de me familiariser avec des séries chronobiologiques appliquées au domaine médical. La séparation des individus sous forme de bons répondeurs, moyen répondeurs, mauvais répondeurs m'avait permis de mettre en évidence la différence de comportements d'un individu à l'autre, de l'effet d'adaptation de l'organisme (différence de comportement d'un individu sur lui-même au cours du temps), en laissant de nombreux doutes subsister sur les modèles utilisés.

Ensuite, j'ai réalisé un travail d'étude de la saturation pulsée en oxygène SpO_2 [1]. Durant cette étude, visant à mettre au point un système de monitoring intelligent de la SpO_2 , notre approche avait été poussée par une démarche de supervision issue des méthodes de contrôle de processus industriel et de diagnostic des défaillances [2, 3]. Une identification des paramètres de modèles linéaires n'avait conduit qu'à une incertitude sur la possibilité de réaliser celle-ci et nous n'avions pu conclure sur la plausibilité de notre modèle d'explication des variations de SpO_2 inscrit dans un graphe causal. Cependant cette étude nous avait permis d'observer plusieurs caractéristiques des enregistrements de données physiologiques sur systèmes soumis à des contraintes diverses : là encore, j'avais observé la variabilité des mesures d'un individu à l'autre, la variabilité d'un même signal sur un même individu et enfin les relations temporelles complexes d'une variable avec une autre. La possibilité de l'existence de régions dans un espace à plusieurs dimensions avait été abordée en annexe (avec représentation dans un espace de phase).

Lors de mon service militaire, j'ai pu travailler sur des données de sommeil et des enregistrements polysomnographiques [4], et j'ai pu me familiariser avec les différentes techniques de mesures, d'enregistrements, de prétraitement des signaux et de mise en place de base de données. C'est alors que j'ai pu effectuer mes premiers pas dans la recherche de classification en différents états de veille et de vigilance sur ces données polysomnographiques. J'avais été surpris par la possibilité de mettre en relation différents individus, différentes données grâce aux scorings réalisés par des experts. Une grande surprise venait aussi de l'utilisation abondante de réseaux de neurones dans ces problèmes de classification et de la très grande difficulté pour valider les résultats obtenus, mais aussi la grande richesse de leur utilisation.

Le sujet proposé par la DGA m'a tout de suite motivé et part de la discussion suivante :

Lionel : « L'armée veut mettre au point un prototype permettant de diagnostiquer l'état de santé d'un blessé polytraumatisé à partir de capteurs portables, tu crois qu'une machine peut le faire ? »

Guillaume : « Si toi tu penses que tu peux le faire avec les données recueillies par les capteurs, alors je crois que la machine pourra. »

... Mais ce n'est pas si simple et le travail est long. En ouvrant plusieurs voies, nous nous sommes rendus compte des différents domaines de la science imbriqués les uns dans les autres. Cette thèse présente une modélisation de ce prototype et les différentes démarches que nous avons entreprises pour indiquer des pistes à approfondir.

Elle s'inscrit dans la continuité de ma recherche de modèles dans les sciences du vivant.

Table des matières

1	Introduction	1
I	Contexte de l'étude	5
2	Présentation de l'étude	7
2.1	Objectifs de l'étude	7
2.2	Le Service de santé des armées	10
2.3	Le polytraumatisé	12
3	Bibliographie	15
3.1	Le choc hémorragique	15
3.2	Le diagnostic des systèmes	19
3.3	Modèles de choc hémorragique	27
3.4	Conclusion du chapitre	29
II	Étude réalisée	31
4	Matériel	33
4.1	Expérimentation Animale	33
4.2	Mesures réalisées	39
4.3	Extraction de caractéristiques	46
4.4	Conclusion du chapitre	62
5	Méthodes	65
5.1	Espace de représentation des données	65
5.2	Apprentissage artificiel, RdF	70
5.3	Techniques de validation	79
5.4	Conclusion du chapitre	86
III	Résultats	89
6	Présentation des données	91
6.1	Analyse de la population	91
6.2	Analyse monodimensionnelle	94
6.3	Représentation multidimensionnelle	100

7	Classification	105
7.1	Choix des échantillons	105
7.2	Recherche des caractéristiques pertinentes	105
7.3	Analyse des modèles	110
7.4	Rejet en ambiguïté et rejet en distance	116
7.5	Nettoyage des données	118
7.6	Conclusion du chapitre	122
8	Visualisation de l'évolution du choc	123
8.1	Maillage de l'espace	123
8.2	Espace à 3 dimensions	124
8.3	Espace à 2 dimensions	124
8.4	Prise en compte du rejet	125
8.5	Conclusion du chapitre	126
9	Prise en compte du temps dans la décision	129
9.1	Changement de granularité	129
9.2	Caractéristiques dynamiques	130
9.3	Conclusion du chapitre	134
IV	Comparaison avec les travaux de Glass et coll.	135
10	Article de Glass détaillé	137
10.1	Présentation de l'article et commentaires associés	137
10.2	Conclusion du chapitre	140
11	Influence de la validation sur les résultats	141
11.1	Validation par sous échantillonnage aléatoire	141
11.2	Validation animal par animal	145
11.3	Conclusion du chapitre	146
12	Classification par perte de sang	147
12.1	Validation Animal par Animal	147
12.2	Évaluation par Leave One (Animal) Out	149
12.3	Conclusion du chapitre	150
13	Discussion	153
V	Conclusion	155
14	Conclusion	157
A	Article Chapitre Sommeil	163
A.1	Introduction	163
A.2	Materials and Methods	164
A.3	Results	171
A.4	Discussion	173

A.5 Conclusion	174
B Indices de Gravités	177
C Techniques de validation	185
C.1 Matrice de confusion	185
D Correspondances avec le Dr Glass	189
D.1 Courrier du 18 mars 2004 adressé à Glass	189
D.2 Réponse du 19 mars 2004 adressée à Becq	190
D.3 Réponse du 19 mars adressée à Glass	191
E Liens	193

Table des figures

2.1	Triage : Récupération des blessés	9
3.1	Modèle à 3 niveaux après lésion traumatique	18
3.2	Élaboration d'un système de décision par reconnaissance de formes	25
3.3	Principes de la reconnaissance de formes	26
4.1	Stabulation de l'animal	34
4.2	Synoptique de l'expérimentation	35
4.3	Intubation de l'animal	36
4.4	Cathétérisme	37
4.5	Stabilisation	37
4.6	Choc	38
4.7	Un signal brut	41
4.8	Évolution des signaux au cours de l'expérimentation	45
4.9	Segmentation temporelle des Signaux	47
4.10	Évolution des Signaux (autre découpage temporel)	48
4.11	Électrocardiogramme : complexes	52
4.12	Corrélogramme d'un <i>ECG</i>	53
4.13	Détection de fréquence cardiaque	54
4.14	Corrélogrammes d'un <i>ECG</i> et d'une respiration	56
4.15	Détection de rythme respiratoire	57
5.1	Transformation vers une distribution normale	68
5.2	Représentation graphique de la base de données	69
5.3	Apprentissage de modèles	71
5.4	Erreurs Nature-Homme-Machine	72
5.5	Paramètre de pondération des kPPV flous	80
6.1	Évolution moyenne des caractéristiques (Début)	96
6.2	Évolution moyenne des caractéristiques (Fin)	97
6.3	Évolution moyenne des caractéristiques (Début)	98
6.4	Évolution moyenne des caractéristiques (Fin)	99
6.5	1D	101
6.6	2D	102
6.7	3D	103
6.8	4D	104
7.1	SFS $\mathcal{S}_*$	106

7.2	Comparaison $\mathcal{S}_*$	108
7.3	SFS $\mathcal{S}_P$	109
7.4	Comparaison $\mathcal{S}_P$	110
7.5	Arbre de décision : $\mathcal{S}_P$, non élagué	114
7.6	Arbre de décision : $\mathcal{S}_P$, non élagué	115
7.7	Arbre de décision : $\mathcal{S}_P$, élagué	116
7.8	Influence du coût de rejet en ambiguïté sur l'erreur	117
7.9	Influence du coût de rejet en distance sur l'erreur	118
8.1	Évolution dans un espace à 3 dimensions	125
8.2	Évolution dans un espace à 2 dimensions	126
8.3	Espace à 3 dimensions et rejet	127
9.1	Granularité temporelle	130
9.2	Transformation vers une distribution normale (<i>vqmFC</i>)	132
9.3	Performance des caractéristiques dynamiques	133
11.1	Influence de la taille des ensembles d'évaluation	142
11.2	Sélection de caractéristiques SFS 300-300	143
11.3	Comparaison des classifieurs	144
13.1	Classification selon perte de sang	154
A.1	Human hypnogram	164
A.2	Electrophysiological behaviour	166
A.3	Representation of the database	167
A.4	Effect of samples size	168
A.5	Parameters selection kNN, Parzen	171
A.6	Effect of feature tranformation	172

Liste des tableaux

4.1	Réglage du débit de spoliation	39
4.2	Mesures	44
4.3	Classification retenue pour le choc hémorragique	61
4.4	Caractéristiques invasives de l'étude	62
4.5	Caractéristiques non invasives de l'étude	63
5.1	Matrice de confusion (cas 3 classes)	81
5.2	Classifieurs	86
6.1	Valeurs de repos	92
6.2	Valeurs a 50%	93
6.3	Valeurs physiologiques au cours d'une hémorragie contrôlée (Becq et coll.)	94
7.1	Matrices de confusion pour le classifieur quadratique	111
7.2	Matrices de confusion pour le classifieur PMC	111
7.3	Modélisation Quadratique par classes	112
7.4	Matrice de confusion pour un rejet en ambiguïté	117
7.5	Nettoyage : ambiguïté experte, transitions courtes	119
7.6		120
7.7	Nettoyage 1PPV $\mathcal{S}_P$ classifieur quadratique	120
7.8	Nettoyage Condensation $\mathcal{S}_*$	121
7.9	Nettoyage Condensation $\mathcal{S}_P$	121
9.1	Transformations pour les caractéristiques dynamiques	132
10.1	Valeurs physiologiques au cours d'une hémorragie contrôlée (Glass et coll.)	139
11.1	Classification par animal. Problème à 4 Classes suivant le choc	145
12.1	Définition des classes suivant la perte de sang	147
12.2	Classification par animal. Problème à 8 Classes suivant la perte de sang . .	148
12.3	Table de confusion - Classification par perte de sang	149
12.4	Classification par animal. Problème à 8 Classes suivant la perte de sang . .	150
12.5	Table de confusion - Classification par perte de sang	151
12.6	Table de confusion - Classification par perte de sang 3 classes	151
A.1	Feature statistics	165
B.1	TS : Trauma Score	178
B.2	RTS : Revised Trauma Score	179

B.3	PHI : Prehospital Index	179
B.4	CRAMS : Circulation, Respiration, Abdomen, Motor, Speech Scale	180
B.5	AP : Anatomic Profile	180
B.6	ASCOT : A Severity Characterization of Trauma	182
B.7	Score de Lindsey	182
B.8	Poids pour les calculs de Probabilité de Survie	183
C.1	Matrice de confusion (cas 2 classes)	185

Résumé

Dans cette thèse, nous proposons et évaluons la mise en place de la chaîne complète de traitement de l'information pour la détermination de l'état de santé d'un blessé polytraumatisé. Nous faisons l'hypothèse que l'état de santé du blessé est observable dans un espace généré à partir de caractéristiques extraites sur des mesures provenant de capteurs portables. Au cours du polytraumatisme, son état se déplace dans des régions propres à chaque classe de choc hémorragique rencontrée, indépendante de l'individu. Le but de cette étude est de trouver ces régions par apprentissage supervisé pour un espace de dimension le plus faible possible (minimisation du nombre de capteurs). Pour cela, nous avons mis au point un modèle animal de choc hémorragique sur de jeunes porcs subissant une perte de sang continue à débit contrôlé. Dans un premier temps l'apprentissage a été réalisé en observant l'état instantané, l'évolution dynamique étant vue comme contextuelle. Dans un second temps nous avons évalué l'apport de l'information dynamique.

Mots clés – Analyse exploratoire multidimensionnelle, Apprentissage supervisé, Capteurs portables, Choc hémorragique, Classification, Classifieur, Diagnostic, Electrocardiogramme, Monitoring, Oxymétrie, Physiologie, Porcs, Reconnaissance de formes, Respiration, Triage.

Title

Contribution to the development of a portable device scoring the health state of a trauma patient.

Abstract

This dissertation describes and evaluates the complete information processing chain proposed to score the health state of a trauma patient. The hypothesis is that the state can be observed in a representative space of features obtained from wearable sensors. During the polytraumatism, the state evolves in various regions corresponding to classes of hemorrhagic shock. These regions are supposed to be independent between individuals. The aim of the study is to recognize these regions. For that, a supervised learning heuristic is proposed and try to find the minimum dimension by feature subset selection (for a minimization of sensors). An animal model of swine enduring controlled continuous blood loss was elaborated. Features were extracted and proposed to an expert for scoring. Supervised learning was realized in a first time in a state space and in a second time in a

phase space by evaluating the benefit of dynamical information.

Keywords – Classification, Classifier, Diagnosis, Electrocardiogram, Hemorrhagic shock, Monitoring, Multivariate analysis, Oximetry, Pattern recognition, Physiology, Respiratory, Supervised learning, Swine, Triage, Wearable Sensors.

Notations

Probabilité

X	une variable aléatoire
P	une probabilité
f	une densité de probabilité
F	une fonction de répartition $F(x) = P(X < x)$
$\hat{x}$	estimation de la variable x
$\mathbb{N}$	l'ensemble des entiers naturels
$\mathbb{R}^d$	l'espace euclidien de dimension d
$\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_d \end{pmatrix}$	un vecteur de $\mathbb{R}^d$
$\mathbf{x} = (x_1, \dots, x_d)^T$	un vecteur avec $.^T$, opérateur transposée
$\mathbf{M}$	une matrice
M_{ij}	élément de la ligne i et de la colonne j de la matrice $\mathbf{M}$
$\mathbf{M}^{-1}$	une matrice inverse d'une matrice carrée $\mathbf{M}$
$ \mathbf{M} $	le déterminant de la matrice $\mathbf{M}$
$d(\mathbf{x}, \mathbf{y})$	la distance entre $\mathbf{x}$ et $\mathbf{y}$
μ	moyenne d'un échantillon
σ	écart-type d'un échantillon
$\sharp$	cardinal d'un ensemble (nombre d'éléments)
$\mathbf{1}_{(m,n)}$	Matrice $m \times n$ remplie de 1
$\mathbf{I}_n$	Matrice identité (1 sur la diagonale, 0 ailleurs)
$\log$	fonction logarithme en base 10
$\ln$	fonction logarithme en base 2

Temporalité

F_e	Fréquence d'échantillonnage
T_e	Période d'échantillonnage
Δt	Durée d'une fenêtre temporelle

Apprentissage Artificiel

$\mathcal{X}$	l'espace de représentation des objets (des formes)
$\mathcal{U}$	l'espace de supervision (des sorties désirées)
$\mathcal{S}$	l'ensemble des exemples, instances, unités statistiques disponibles $\mathcal{S} = \{(\mathbf{x}_i, u_i), i = 1, \dots, n\}$
$\mathcal{A}$	l'ensemble des unités statistiques pour l'apprentissage
$\mathcal{T}$	l'ensemble des unités statistiques pour le test
$\mathbf{z}_i = (\mathbf{x}_i, \mathbf{u}_i)$	un exemple (élément d'un échantillon d'apprentissage)
$\mathbf{x}_i$	un vecteur de $\mathcal{X}$
x_{ij}	la valeur de la coordonnée j de la description de l'objet $\mathbf{x}_i$ dans $\mathbb{R}^d$
u_i	la supervision ou sortie désirée associée à $\mathbf{x}_i$
$f : \mathcal{X} \rightarrow \mathcal{U}$	la fonction cible (celle que l'on cherche à apprendre)
$\mathcal{H}$	l'espace des hypothèses d'apprentissage
$h \in \mathcal{H}$	une hypothèse produite par un apprenant (un algorithme d'apprentissage)
$\mathbf{y} = h(\mathbf{x})$	la prédiction faite par l'hypothèse h sur la description $\mathbf{x}$ d'un exemple.
$\mathbf{X}$	la matrice (n,d) des observations
$l(f, h)$	la perte (ou distance) entre la fonction cible et une hypothèse
$R_{réel}$	le risque réel
R_{emp}	le risque empirique
$err_{\mathcal{A}}$	erreur obtenue sur $\mathcal{A}$, estimation de R_{emp}
$err_{\mathcal{T}}$	erreur obtenue sur $\mathcal{T}$, estimation de $R_{réel}$
$\mathcal{C}$	l'ensemble des classes
C	le nombre de classes ($\sharp \{\mathcal{C}\}$)
ω_k	une classe de $\mathcal{C}$

Médical

Abd	Effort respiratoire (abdomen)
$animal$	Étiquette (nom) de l'animal
CO_2	Dioxyde de carbone
$classes$	Classes
Δ_{PA}	Différentielle de pression artérielle
ECG	Électrocardiogramme
EEG	Électroencéphalogramme
E_tCO_2	CO_2 de fin d'expiration (angl. <i>end tidal CO₂</i>)
FC	Fréquence cardiaque (angl. <i>HR, heart rate</i>)
$LFHF_{PoulsB}$	Basses fréquences sur hautes fréquences pouls Biopac
$LFHF_{PoulsK}$	Basses fréquences sur hautes fréquences pouls Kontron
$LFHF_{PVC}$	Basses fréquences sur hautes fréquences pour la pression veineuse centrale
PA	Pression Artérielle (angl. <i>blood pressure</i>)
PAD	Pression artérielle diastolique
PAM	Pression artérielle moyenne
PAS	Pression artérielle systolique
$PerteSang$	Pourcentage de volume sanguin prélevé
$\tilde{V}t_{Abd}$	Pseudo volume courant abdomen (angl. <i>tidal volume</i>)
$\tilde{V}t_{Tho}$	Pseudo volume courant thorax (angl. <i>tidal volume</i>)
$PoulsB$	Onde de pouls (oxymètre biopac)
$PoulsK$	Onde de pouls (oxymètre Kontron)
PVC	Pression veineuse centrale
RR_{Abd}	Rythme Respiratoire (ou fréquence respiratoire) (angl. <i>respiratory rate</i>) au niveau de l'abdomen
RR_{Tho}	Rythme respiratoire (ou fréquence respiratoire) (angl. <i>respiratory rate</i>) au niveau du thorax
SaO_2	Saturation en oxygène
SpO_2	Saturation pulsée en oxygène
σ_{PA}	Écart-type de la pression artérielle
σ_{ECG}	Écart-type ECG
σ_{PoulsK}	Écart-type pouls Kontron
σ_{PoulsB}	Écart-type pouls Biopac
T_{amb}	Température ambiante
T_{cut}	Température cutanée (scutularis interne)
T_{int}	Température interne (rectale)
$temps$	Instants de l'expérimentation
Tho	Effort respiratoire (thorax)

Chapitre 1

Introduction

La vie est courte, l'art est long, l'occasion fugitive, l'expérience trompeuse, le jugement difficile.

Aphorisme attribué à Hippocrate

Le soldat sur le théâtre d'opération est exposé à un risque de blessures important. Si il est blessé, sa survie dépend d'un grand nombre de paramètres : l'ampleur des traumatismes qu'il a subi, la perte de sang résultante, le temps de prise en charge par les services de soutien aux blessés, la mise en place d'un garrot par une personne proche... Toutes ces considérations semblent de bon sens, néanmoins, on peut se demander quel peut être le processus qui permet d'optimiser le processus de récupération et de soutien de ces blessés polytraumatisés.

La mise en place d'un soutien optimal passe par la maîtrise des différents facteurs qui entrent en jeu au cours de l'évolution et de la prise en charge du ou des traumatismes subis : d'un point de vue phénoménologique par la compréhension des mécanismes physiologiques opérants sur la victime, d'un point de vue conséquentialiste, par la mise en place d'une structure efficace de récupération du blessé.

Afin d'organiser au mieux la récupération des blessés, et soigner le plus grand nombre de blessés présents sur un théâtre d'opération, la Délégation Générale pour l'Armement (DGA) subventionne un nombre important de recherches avant de pouvoir faire appliquer celles ci par les différents corps de l'Armée, en routine, sur le terrain. Dans le cadre du suivi médical des forces engagées, le Service de Santé des Armées (SSA) est généralement subventionné par la DGA pour réaliser des recherches spécifiques. Par exemple, le travail développé dans cette thèse fait l'objet d'un article du projet d'étude amont (PEA) 98.08.21 ayant pour thème le « soutien médical précoce aux blessés polytraumatisés ». Cette problématique intéresse d'autres unités oxylogiques civiles¹ ou de médecine d'urgence (angl. *emergency medical services, EMS*), où la prise en charge des blessés est un souci permanent. Le but des études présentes dans ce PEA est de mieux définir les facteurs entrant en jeu dans la prise en charge des blessés et d'en apprécier leur part respective dans les différents domaines concernés. Dans un souci de récupération du plus grand nombre de personnes à secourir, ces études essaieront de définir les priorités et les moyens à mettre en place pour récupérer les personnes qui nécessitent une intervention rapide, au détriment

¹services d'aide médicale urgente (SAMU), service mobile d'urgence et de réanimation (SMUR)

d'autres blessés, qui peuvent attendre (les vrais patients ?).

En nous plaçant sur le théâtre d'opération et en considérant les techniques modernes existantes, nous pouvons approximativement imaginer le nombre d'informations qu'il est possible de faire transiter de la personne incapacitée au centre logistique responsable des décisions de récupération. Nous pouvons arriver à un nombre important de mesures à envisager et il est légitime de se demander lesquelles sont les plus pertinentes. Ainsi, l'information véhiculée ne doit pas noyer l'information pertinente et de nombreux choix doivent être réalisés pour apporter une aide réellement efficace et rapide, lors de la crise.

Notre travail se situe au niveau du polytraumatisé et essaie d'évaluer si un système portable peut être mis en place pour fournir au centre logistique un indice sur son état de santé. Nous nous intéressons donc à la chaîne complète de traitement de l'information à mettre en place pour pouvoir délivrer un tel indice à partir de mesures réalisées sur l'individu par des capteurs préalablement disposés sur le corps, caricaturalement « sous gilet pare-balles ».

Il est important de préciser que notre étude se situe au niveau du blessé polytraumatisé, et que l'observation de cet organisme traumatisé en est la partie centrale. La principale fonction de l'appareil envisagé consistera donc à observer la réaction de l'organisme face à des traumatismes multiples dont l'évolution est généralement liée à la perte de sang résultante. La création et la mise en place d'un tel appareil s'accompagne donc de l'étude des mécanismes physiologiques qui rentrent en jeu à la suite de perturbations brutales, et du diagnostic (ou de l'aide au diagnostic) que l'on peut en faire avec les moyens modernes à notre disposition.

Afin de modéliser les réactions de l'organisme soumis à de telles agressions, nous avons travaillé sur un modèle animal, le jeune porc anesthésié subissant une perte de sang à débit contrôlé. Des capteurs invasifs (capteurs nécessitant une chirurgie) et non invasifs (capteurs portables) ont été utilisés. Nous avons pu réaliser plusieurs enregistrements et ainsi constituer une base de données sur laquelle nous avons lancé nos recherches exploratoires afin d'effectuer une reconnaissance des phénomènes concernés et d'évaluer différents systèmes portables envisageables.

Un opérateur humain a réalisé une classification des données en leur attribuant un score reflet d'une classe de choc hémorragique correspondant à un état particulier de l'organisme de l'animal. Nous avons alors essayé de réaliser un transfert de cette connaissance experte, en essayant de faire apprendre à la machine le score manuel réalisé par l'expert. Cet apprentissage supervisé nous a servi de base pour mettre en évidence l'intérêt de certains capteurs par rapport à d'autres. Il nous a permis de valider ou réfuter certaines hypothèses mais aussi de nous interroger sur les possibilités liées à l'induction d'un système cognitif dans un contexte idiosyncrasique.

Ce rapport présente la méthode que nous avons suivie pour mettre en évidence les problèmes liés à la mise en place d'un tel système portable. Il soutient fondamentalement la thèse selon laquelle un système portable de diagnostic de classe de choc hémorragique peut être applicable sur n'importe quel individu, en utilisant un nombre minimal de capteurs et en utilisant une information dynamique minimale.

La première partie développera le contexte de notre étude, les objectifs du système envisagé et les hypothèses retenues pour y parvenir. Nous présenterons les différentes parties impliquées dans la récupération des blessés et en particulier le Service de santé des Armées dont l'action mixte se situe au niveau opérationnel et au niveau de la recherche médicale. Ceci nous permettra d'indiquer la position dans laquelle nous nous plaçons dans cette

étude. Nous dresserons un état de l'art des différents domaines scientifiques (informatique, traitement du signal, automatique, biologique, médical...), conduisant à la création du système de traitement de l'information envisagé. Cette partie permettra donc de préciser nos hypothèses et de situer notre étude.

La deuxième partie présentera l'étude réalisée d'un point de vue pragmatique. Le chapitre matériel permettra de présenter l'expérimentation animale et physique réalisée, les différents instruments impliqués pour acquérir les données et les outils informatiques retenus pour traiter les signaux et extraire les caractéristiques pertinentes de l'étude. Le chapitre méthodes présentera les techniques retenues pour élaborer notre système et valider ses performances. Nous nous intéresserons en particulier aux différents modes de représentation et leurs implications sur la modélisation des phénomènes en jeu. Nous développerons la fonction cognitive de notre travail qui a consisté à faire apprendre à une machine un savoir expert. Nous présenterons donc les techniques d'apprentissage supervisé utilisées et préciseront leurs liens avec les processus de reconnaissance de formes utilisés en automatique. Nous aborderons enfin les techniques de validation qui leur sont associées.

La troisième partie présentera les résultats de l'étude et comparera l'influence des différentes techniques employées. Nous présenterons dans un premier temps les résultats d'une analyse au cas par cas et observerons les caractéristiques retenues de l'étude, une à une. Nous présenterons les avantages d'une analyse par une approche multidimensionnelle sur ce type de données et mettront en pratique l'apprentissage supervisé et la reconnaissance de formes présentés dans la partie précédente. Nous présenterons la sélection de caractéristiques réalisée par notre démarche et comparerons les différents modèles obtenus. Nous présenterons les cartographies et cinématiques réalisables qui pourraient permettre un suivi plus pratique de l'évolution de l'état de santé du blessé et permettre ainsi une aide à la décision. Nous apprécierons l'influence de la prise en compte du temps dans cette étude et ses implications dans les choix technologiques envisagés.

Enfin, nous présenterons les résultats de l'équipe de Glass parus dans un article récent et qui développe un système très proche du nôtre, et évaluerons les effets de l'utilisation de techniques similaires de validation sur notre modèle pour mettre en évidence l'intérêt de notre démarche.

Il sera alors largement temps de tirer les conclusions de notre étude, d'en apprécier les implications, et d'en émettre des perspectives.

Première partie

Contexte de l'étude

Chapitre 2

Présentation de l'étude

Ce chapitre a pour but d'inclure notre sujet dans son contexte de soutien médical précoce au blessé dont la fonction majeure est réalisée par le Service de santé des Armées. Il nous permettra de préciser les limites de notre étude et de présenter les hypothèses de travail retenues. La section 2.1 présentera le problème de récupération des blessés sur un théâtre d'opération, les problèmes de triage effectués lors de situations de catastrophes et les principales hypothèses retenues pour la création et la mise en place d'un système portable capable de faciliter le travail des forces en présence pour faciliter et améliorer la récupération des blessés. La section 2.2 présentera le rôle du Service de santé des armées dans cette récupération, le soutien logistique apporté par le Groupe Médical ainsi qu'une partie historique sur le rôle de l'Armée dans le soutien médical aux blessés. Enfin, la section 2.3 portera notre attention sur la pathologie retenue pour notre étude correspondant au cas le plus fréquemment retrouvé sur le terrain et en présentera les principales caractéristiques. Ceci nous permettra de cibler notre étude sur le facteur le plus important influençant le temps de survie du blessé, le choc hémorragique et son étude par le biais d'une expérimentation animale.

2.1 Objectifs de l'étude

2.1.1 Contexte du théâtre d'opération

L'article récent de Howard Champion, l'un des plus brillants spécialistes de la prise en charge des blessés polytraumatisés civils ou militaires [5], nous dresse un tableau particulièrement complet des blessures de combats auxquelles sont confrontés les soldats sur les théâtres d'opération. L'auteur indique que les causes principales des pertes, pour des blessés potentiellement secourables, sont dues aux exsanguinations et aux lésions directes du système nerveux central. L'accent est mis sur le fait que la perte de sang au niveau des membres périphériques compte pour plus du double de morts potentiellement secourables, et suggère que l'attention soit portée, en priorité, sur le maintien d'une homéostasie, impliquant un soutien par des agents extérieurs pour réduire ou bloquer une hémorragie non accessible ou non contrôlée.

Selon ce même auteur, les statistiques de mortalité indiquent qu'en 150 ans, la proportion d'hommes tués sur le terrain (angl. *killed in action*) n'a pratiquement pas bougé et se maintient aux alentours de 20% : les personnes secourables auraient pu être sauvées par prise en charge de leur perte de sang dans un temps acceptable.

Des observations similaires ont été rapportées dans le document qui nous a servi de base pour notre recherche [6]. Ce document dresse un premier bilan des méthodes existantes et il répertorie les moyens à mettre en place pour un diagnostic précoce d'hypovolémie majeure chez l'homme jeune dans une situation d'urgence telle que le champ de bataille, correspondant aux personnes décrites par Champion.

2.1.2 Situation de catastrophes

Les différentes organisations qui se mettent en place lors d'une situation de catastrophe, dépendent de la nature des multiples scénarios possibles : catastrophes naturelles, guerres, accidents, épidémies, liés à des produits toxiques, dans des environnements extrêmes, dépendant de la pathologie des blessures, des catastrophes, de la présence des unités médicales mobilisés, etc. [7]. En dépit du nombre des unités déployées et de leur adéquation avec la situation, la situation de catastrophe se définit souvent par un nombre insuffisant d'unités pour secourir les blessés présents sur le site où a eu lieu une crise.

Les différentes étapes de prise en charge se décomposent en différents lieux et ont pour but de gérer le flux des blessés : récupération des victimes sur le site, triage sur un site primaire (attente ou prise en charge suivant la gravité des blessures), évacuation et répartition vers des centres secondaires (dès que la gravité diminue) [8]. La notion de triage (angl. *triage*) consiste à définir des consignes pour déterminer les priorités dans le suivi des patients et la décision de leur prise en charge¹. De cette façon nous pouvons reprendre la formule de Tissot « le triage est un acte médico-chirurgical adapté à une situation très particulière : l'afflux de blessés » [7].

Un processus de récupération des blessés peut être résumé par le schéma proposé Fig. 2.1 en notant, toutefois, que le soldat blessé, à l'heure actuelle, ne peut transmettre son état de santé sans l'aide d'une personne. Seul un coéquipier présent à ses côtés peut communiquer cette situation ou bien lui même si il est en mesure de réaliser son propre diagnostic et son repérage. Le schéma proposé Fig. 2.1 correspond donc à une vision futuriste proposée par la DGA pour un meilleur soutien du blessé. Cette vision nécessite le développement d'un système portable s'inscrivant dans le cadre d'une recherche de télémédecine pour l'avant. Le système doit être capable de communiquer un suivi en temps réel de l'état de santé d'un soldat. Il permet en outre de définir sa localisation et les différentes informations médicales standards du soldat. Ceci pourrait permettre une meilleure prise en charge pour la récupération des blessés lors d'un contexte où sont invalidés de nombreux individus.

Ainsi, pour un soutien médical précoce aux blessés polytraumatisés, la DGA souhaiterait savoir si la mise en place de capteurs disposés sur le soldat en opération pourrait permettre de fournir des informations favorables à sa récupération et à une optimisation de son triage, et ceci dans un souci de récupérer le plus grand nombre d'individus. La prise de mesures ambulatoires est de nos jours de plus en plus accessibles par le biais de capteurs intégrés à des gilets (angl. *lifeshirt*) [9], à un ensemble d'habillement [10] ou dans l'utilisation de matériaux textiles intelligent [11] (projet européen Wealthy²). Si ces solutions génèrent un ensemble de mesures plus ou moins fiables accessibles par monitoring, elles ne proposent néanmoins pas d'analyse multidimensionnelle de celles ci dans

¹<http://www.nato.int>, <http://www.sante.gouv.fr>

²Projet européen Wealthy : <http://www.wealthy-ist.com/>

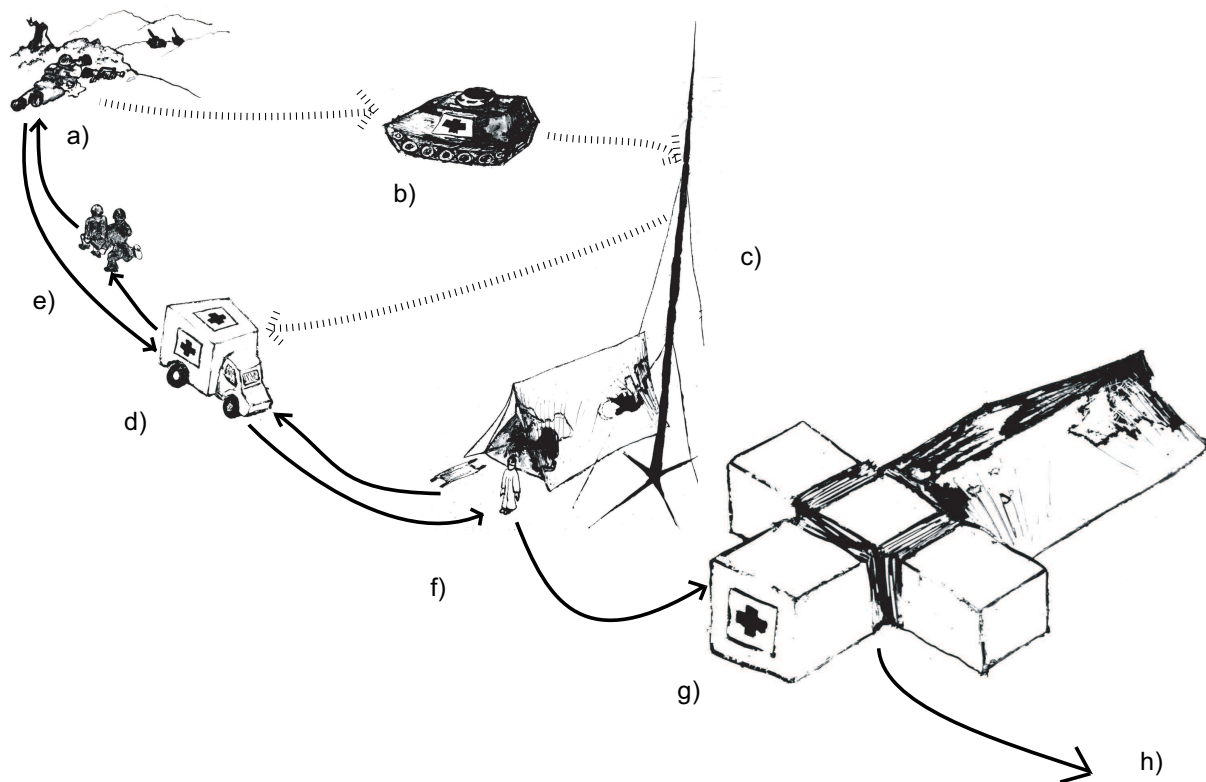


FIG. 2.1 – Triage : Récupération des blessés. Le blessé a) est immobilisé en position latérale de survie. Un véhicule d'avant blindé sanitaire b) repère les soldats et fait transiter les informations de faibles puissances sur des distances plus importantes vers le centre logistique c). Les officiers décident suivant l'urgence de la situation, la mobilisation des forces disponibles pour la récupération des blessés. Les véhicules disponibles d) sont envoyés au plus près du blessé. Les brancardiers secouristes e) récupèrent le blessé et apportent les premiers secours. Le malade est alors véhiculé vers le centre de tri f) qui décide ou non de la prise en charge au bloc opératoire ou en réanimation dans l'hôpital militaire modulaire de proximité g). Dès que l'état du patient est stabilisé, un transport vers un hôpital militaire ou civil plus important h) est effectué.

l'élaboration d'une aide au diagnostic sous la forme d'un indice de gravité ou d'un score facilitant la récupération des blessés ou le triage de ceux-ci.

Notre étude essaie de répondre à la demande du Service de santé des armées pour la récupération de ces blessés en proposant une chaîne complète de traitement de l'information sur les mesures obtenues par des capteurs portables (angl. *wearable sensors*) semblables à ceux insérés dans ces vêtements dits intelligents (angl. *smart clothes*) et dont la robustesse doit être éprouvée (au moins en laboratoire). Elle propose une analyse multidimensionnelle pour la délivrance d'un diagnostic par des techniques de reconnaissance de formes.

2.1.3 Hypothèse

Nous émettons l'hypothèse que l'état de santé du blessé est observable dans un espace représentatif. Cet état est déterminé par des caractéristiques calculées à partir de mesures réalisées par les différents capteurs portables installés sur le blessé polytraumatisé. Au cours du polytraumatisme, l'état se déplace dans des régions propres à chaque classe de choc hémorragique rencontré, indépendantes de l'individu. D'un point de vue

expérimental, nous pensons qu'il est possible de pouvoir tirer d'un ensemble d'individus, des données consistantes permettant de reconnaître (apprendre) ces différentes classes de choc hémorragique.

Moyens pour valider l'hypothèse

Les moyens seront mis en œuvre pour caractériser les régions propres à chaque classe de choc hémorragique par apprentissage supervisé pour un espace de dimension le plus faible possible (minimisation du nombre de capteurs), chez des cochons subissant une perte de sang en continu :

- Dans un premier temps en observant l'état instantané, l'évolution dynamique n'étant pas prise en compte. La prise de décision sera alors effectuée suivant une approche conditionnelle (ou contextuelle), et s'appuiera sur une étude de la distribution conjointe des variables.
- Dans un second temps en évaluant l'apport de l'information dynamique : d'une part sur la durée de la fenêtre de calcul des caractéristiques, d'autre part en incluant dans l'étude des caractéristiques dynamiques.

La suite de ce rapport présentera les différents étages de cette chaîne de traitement de l'information et les différents choix réalisés. Mais avant de présenter le système dans sa particularité, il est intéressant de revenir sur les différentes parties impliquées dans la prise en charge des blessés militaires. En France, le soutien médical aux blessés polytraumatisés sur le théâtre d'opération est réalisé par le Service de santé des armées. Il est assisté logistiquement par les régiments médicaux qui mettent à disposition leurs forces et matériels opérationnels. La partie recherche est dirigée par la Délégation générale pour l'armement (DGA), qui a financé ce travail et débloqué des crédits pour l'acquisition du matériel et la mise en place des expérimentations.

2.2 Le Service de santé des armées

Le Service de santé des armées (SSA) a de multiples fonctionnalités, puisqu'il doit assurer un suivi des personnels militaires aussi bien en temps de paix qu'en temps de guerre. La présentation disponible sur le portail électronique officiel du ministère de la défense nous indique ³ que le Service de santé des armées a pour vocation d'assister le personnel militaire selon différents axes :

- Pratique d'un suivi médical (visites périodiques, visites d'aptitude),
- Surveillance de l'hygiène générale et individuelle (salubrité des casernements, contrôle de l'eau et de l'alimentation, désinsectisation etc.);
- Contrôle de l'entraînement physique militaire,
- Vaccinations légales et réglementaires (B.C.G., antidiphtérique, antitétanique, antipoliomyélitique, antiméningococcique, antihépatitiques, fièvre jaune etc.),
- Éducation sanitaire sur les fléaux sociaux (alcoolisme, drogue, tabagisme, tuberculose, maladies sexuellement transmissibles, SIDA etc.),
- Études épidémiologiques, cliniques et fondamentales sur la prévention des risques en environnement hostile (pathologie infectieuse, bruit d'armes, coup de chaleur etc.).

³source : <http://www.defense.gouv.fr>

En opération, son rôle recouvre l'ensemble des actions concourant, sur les plans individuel et collectif, à la mise en condition et à la conservation du potentiel humain par une prise en charge intégrale et cohérente, dans un cadre interarmées, des combattants, des malades et des blessés des forces.

Le Service de santé a aussi un rôle important de recherche pour l'amélioration et le développement de moyens pour la prise en charge, la protection et la survie des combattants confrontés à des conditions extrêmes.

Le Centre de recherches du Service de santé des armées (CRSSA) qui nous a accueilli durant ces travaux a, en particulier, cette vocation.

2.2.1 Historique

Si, pour une armée moderne, la nécessité d'un service de santé organisé est une évidence, pour des raisons sociales, humaines et matérielles (un soldat moderne est un homme qui a une forte valeur ajoutée), il n'en a pas toujours été ainsi. Pendant longtemps, les soins sur le champ de bataille ont été réduits à des gestes élémentaires effectués par le blessé lui même, ou par ses camarades de combat. Parfois ils étaient pratiqués par des chirurgiens, ou barbiers, que le chef de guerre rémunérait pour son usage personnel et celui de ses hommes.

- Du Moyen-Âge au XVIIIe siècle : Le changement d'échelle des conflits, avec le développement des armes à feu et la constitution de grandes armées, a rendu nécessaire une organisation spécifique, afin de faire face à la fois au traitement des blessures multiples et au risque épidémique, tout aussi meurtrier. La France fut l'une des premières à se doter d'un tel service. La naissance du Service de santé était consacrée à la fin du règne de Louis XIV par l'édit du 17 janvier 1708.
- Les temps modernes : La deuxième grande date pour le Service de santé des armées, c'est la loi de 1882 qui pour la première fois donne son autonomie technique au service de santé. Cette loi de 1882 est complétée en 1889, par une loi qui donne au service de santé son autonomie administrative. Troisième grande étape, c'est l'apparition du service de santé sous sa forme interarmées avec en 1948 la création d'une direction unique, puis en 1968 le regroupement des services de santé de l'armée de terre, de la marine, de l'armée de l'air et des troupes coloniales en un seul service de santé, qui obéit sur le plan technique à un chef unique, le directeur central.
- Dernière étape, en juillet 1991, le plan ARMÉES 2000, renforce l'autonomie du SSA autour de 2 pôles majeurs :
 - la composante propre au service, regroupant tous les moyens du soutien interarmées, de caractère technique ou logistique,
 - la composante intégrée aux forces, adaptée aux besoins spécifiques de chaque armée.

Depuis 1996, avec les débuts de la professionnalisation, le Service adapte ses moyens au nouveau format des armées et à leurs nouvelles conditions d'emploi, ce qui se traduit par des effectifs en personnels de carrière et sous contrat en augmentation, ainsi que par une adaptation aux nouvelles structures de commandement.

C'est dans ce contexte que chaque individu acquiert une valeur de plus en plus importante et que la prise en charge de celui ci nécessite un cadre privilégié.

2.2.2 Le groupe médical, le régiment médical

Pour la récupération des blessés, le personnel médical militaire du SSA est assisté par une force armée dotée d'un ensemble de moyens humains et logistiques, le groupe médical. Sa mission principale est le soutien santé du combattant qui doit être assuré d'être traité au plus tôt et de bénéficier des techniques médico-chirurgicales les plus modernes. Les événements politico-stratégiques récents ont mis en évidence la nécessité de disposer d'un régiment médical capable de remplir des missions tant dans le cadre d'intervention extérieures que dans celui d'actions à vocation humanitaire.

La fonction du groupe médical est donc la suivante :

- La mobilisation et l'exploitation des matériels disponibles (Camions super lourds, Véhicules de transport logistique avec remorque pour le transport des éléments techniques modulaires, Véhicules de l'avant blindés version sanitaire, Chariots élévateurs types 15 et Transmanut, Véhicules légers, Motos, Camions lourds tous terrains).
- Le déploiement des hôpitaux militaires et la récupération des blessés.
- Le brancardier secouriste doit être capable de prendre en charge une personne malade ou blessée et d'agir face à une éventuelle urgence simple. Lors de l'engagement de son unité sur un théâtre d'opération ou pendant les périodes d'exercice, il assure le ramassage des blessés. Il doit donc savoir observer une victime, pratiquer en urgence les gestes des premiers secours et être en mesure d'assurer le transport de la victime.

Une importante organisation logistique et médicale se met donc en place autour des individus blessés, victimes des combats. Les principales pathologies sont liées aux différents traumatismes occasionnés par les armes de guerre destructrices. Les victimes sont alors référencées sous la désignation générale de polytraumatisés.

2.3 Le polytraumatisé

« Le polytraumatisé est un blessé qui présente deux ou plusieurs lésions traumatiques graves, périphériques, viscérales ou complexes entraînant une répercussion respiratoire ou circulatoire ». Cette définition classique de A. Trillat et A. Patel peut être élargie, et l'on peut considérer aussi comme un polytraumatisé tout blessé qui présente des traumatismes périphériques multiples qui, par effet de sommation, vont mettre en jeu le pronostic vital [12].

La particularité de la prise en charge des blessés polytraumatisés réside dans la nécessité de réaliser un diagnostic sur l'ensemble de l'organisme. Il faut donc faire appel à un savoir multidisciplinaire. Les principaux ouvrages font une synthèse de ces particularités en développant d'une part les moyens pour *évaluer* l'état de santé du blessé polytraumatisé, d'autre part les méthodes pour *prendre en charge* les différentes pathologies associées aux différents organes, membres ou entités fonctionnelles touchées. La première étape de la prise en charge du blessé polytraumatisé consistera donc à surveiller les perturbations des principales fonctions vitales : d'une part les perturbations de la fonction circulatoire ou respiratoire et d'autre part les perturbations graves de la conscience, voire le coma [13, 14].

Le monitoring du blessé se fera donc avec pour optique la recherche ou le suivi d'une de ces perturbations. Les perturbations circulatoires rencontrées chez les polytraumatisés sont généralement le collapsus, le choc, l'arrêt circulatoire. Les perturbations majeures de la respiration sont les dyspnées graves, les hypoxémies, les asphyxies et l'apnée. L'état de

coma fait courir un risque vital à court terme en menaçant la liberté des voies aériennes et en perturbant la ventilation. Trois types de mécanismes permettent d'expliquer la perturbation d'une fonction vitale :

- Le traumatisme direct de la structure servant de support à cette fonction (crâne, thorax, rachis).
- Le choc traumatique⁴ dont la composante est l'*hypovolémie*.
- Une complication évolutive telle que l'embolie graisseuse, la défaillance mono ou polyviscérale, l'infection (choc septique).

Une attention particulière se fera donc sur le suivi de la ventilation, du remplissage vasculaire, de l'oxygénation tissulaire et de l'hypothermie [13].

Nous nous intéressons, dans notre étude à la recherche des symptômes liés à des polytraumatismes qui ne sont pas pris en charge et dont l'évolution conduit à un état de type choc hémorragique. Comme plusieurs règles peuvent être énoncées suivant la nature des traumatismes, nous excluons en particulier les traumatismes crâniens dont la résultante n'est pas liée à un choc hémorragique mais à des perturbations intracérébrales avec atteinte directe du système nerveux. L'hypovolémie responsable du choc traumatique est de loin le facteur le plus aggravant de la situation du traumatisé. L'état traumatique est alors fonction du choc hémorragique. Le pronostic dépend alors de la durée et de l'intensité du choc [16]. C'est sur ce phénomène que nous avons porté notre attention pour réaliser notre étude et évaluer la mise en place d'un système portable capable de diagnostiquer l'état de santé du blessé polytraumatisé.



⁴Un état de choc [15] se définit comme la diminution aiguë et durable de la perfusion tissulaire, compromettant les apports nutritionnels cellulaires. Cette situation conduit à une ischémie de l'ensemble des viscères, avec une hiérarchie qui tend à préserver le cerveau et le cœur, au détriment du tube digestif, des muscles, de la peau, puis des poumons, du foie et des reins.

Chapitre 3

Bibliographie

L'Éternel Dieu fit pousser du sol des arbres de toute espèce, agréables à voir et bons à manger, et l'arbre de la vie au milieu du jardin, et l'arbre de la connaissance du bien et du mal.

Bible, Genèse, 2.9

Dans ce chapitre, nous présentons les principales connaissances et sources dont nous nous sommes inspirés pour mettre en place notre système et comprendre les mécanismes rentrant en jeu dans l'organisme étudié. Nous présentons dans la section 3.1, les connaissances médicales liées au choc hémorragique actuellement disponibles. Dans la section 3.2, nous présentons les techniques actuelles de diagnostic, d'une part médicales, d'autre part des sciences dérivées de l'automatique et de celle liées au développement de l'informatique. Enfin, dans la section 3.3, nous dressons une analyse des différents systèmes disponibles actuellement qui permettent d'évaluer l'état de santé d'un blessé et les principales orientations pour en préciser les modes de fonctionnement et en prédire l'évolution.

3.1 Le choc hémorragique

Les principales références [17, 18, 19] exposent les différents mécanismes observés rentrant en jeu lors du choc hémorragique qui ont largement été étudiés sur le plan expérimental depuis de nombreuses années, principalement chez le chien par Wiggers [19]. Nous reprenons en partie, ici, le document du Dr Declety disponible sur le site du corpus de la faculté de médecine de Grenoble¹ augmenté des précisions apportées par Schadt [18], qui propose un article de revue sur le choc hémorragique chez les mammifères.

3.1.1 Définition

Le choc hémorragique est une insuffisance circulatoire aiguë par spoliation sanguine majeure et durable, rencontrée dans des circonstances aussi variées que la traumatologie, la pathologie digestive ou l'obstétrique. Le pronostic est fonction de la durée et de la gravité du choc, et par conséquent dépend de la rapidité du diagnostic et du délai de mise en route d'un traitement approprié.

¹<http://www-sante.ujf-grenoble.fr/SANTE/alpesmed/corpus.htm>

3.1.2 Physiopathologie

Le choc hémorragique est la principale cause de choc hypovolémique. Sa survenue est fonction de l'importance et de la vitesse de la spoliation sanguine, ainsi que de l'efficacité des mécanismes compensateurs. Ceux-ci sont réduits par une anémie chronique, une hypovolémie antérieure, une insuffisance cardiaque, l'anesthésie, les bêta-bloquants ou les vasodilatateurs. Chez un sujet normal, la baisse de la pression artérielle (PA) survient pour une spoliation sanguine d'environ 25 à 40% de la masse sanguine (valeur normale = 70 ml/kg).

Trois étapes successives peuvent être distinguées au cours du choc hémorragique :

- une phase de *choc compensé* où l'hypoperfusion tissulaire est contrebalancée par des mécanismes adaptatifs circulatoires (pas de baisse de PA),
- une phase de *choc décompensé* avec l'apparition d'un cercle vicieux d'aggravation progressive et d'évolution fatale en l'absence de traitement (baisse de PA),
- le *choc irréversible* défini par un point de non retour quelque soit la thérapeutique instituée.

Phase sympathoexcitatrice

Phase sympathoexcitatrice : phase d'excitation sympathique (angl. *sympathoexcitatory phase*). Observée jusqu'à une perte de 25 à 35% du volume sanguin total. Cette phase est aussi qualifiée d'hémorragie modérée (angl. *moderate hemorrhage*). Le débit cardiaque diminue, les réflexes cardiovasculaire engendrent une élévation de fréquence cardiaque et une augmentation de la résistance dans les muscles cutanés et la région splanchnique. La résistance vasculaire chute dans le cerveau et le comportement de la résistance vasculaire coronaire est ambigu. On constate une augmentation de la résistance vasculaire systémique inversement proportionnelle à la chute de débit cardiaque, ce qui compense normalement la chute de pression artérielle. Ceci permet d'alimenter les régions vitales (cœur, cerveau) en priorité.

- Mécanismes efférents : Vasoconstriction sympathique bien connue mais qui dépend : de la conscience de l'animal, de mécanismes afférents (barorécepteurs etc.), donc système d'autorégulation de la vasoconstriction capable de modérer la baisse de tension artérielle pas très bien connu. Il apparaît cependant que l'augmentation de la conduction sympathique vasoconstrictrice est responsable de la hausse de la résistance vasculaire systémique.
- Mécanismes afférents : les barorécepteurs artériels ont un rôle important. Si une barodénervation est effectuée, la fréquence cardiaque n'augmente pas, la résistance vasculaire systémique diminue, la PAM aussi. Les récepteurs cardiaques et pulmonaires pourraient rentrer aussi en compte dans l'atténuation de la réponse sympathoexcitatrice.

Phase sympathoinhibitrice

Phase sympathoinhibitrice : phase d'inhibition sympathique (angl. *sympathoinhibitory phase*). Les études réalisées durant la seconde guerre mondiale mettent en évidence une chute dramatique de la pression artérielle accompagnant une syncope lors de choc hémorragique important. La chute de pression est obtenue à la suite d'une diminution critique du volume sanguin central, sur des mammifères conscients (hommes, lapins, chiens...) ou

lors de l'occlusion de la veine cave. Chez le rat, on observe une chute de pression dès le début du choc. Dans tous les cas, cette chute de pression s'accompagne d'une chute de la fréquence cardiaque. Chez l'homme, la chute de pression est dramatique et semble provenir d'une médiation vagale. Cependant les 2 chutes sont indépendantes. La chute de pression est aussi indépendante du débit cardiaque.

La chute de pression lors de la diminution du volume sanguin central est due à une chute précipitée de la résistance périphérique totale. Le flux sanguin augmente dans les veines et les muscles squelettiques, alors qu'il diminue dans la peau, et dans les os. La résistance périphérique s'écroule.

- Mécanismes efférents : la vasodilatation des muscles squelettiques semble être modulée par l'activation des fibres vasodilatatrices. Cependant, l'inhibition centrale de l'activité sympathique semble être le facteur déterminant de la chute de pression sanguine. La libération d'autres substances (en particulier libération de catécholamines par la surrénale) n'y peut rien.
- Mécanismes afférents : l'action des barorécepteurs artériels a été discuté lors de la phase d'activation sympathique. Lors de la phase d'inhibition, les récepteurs cardiopulmonaires semblent jouer un rôle important en particulier par leurs afférences vagales. Le rôle des récepteurs sensoriels cardiaques lors d'hypovolémie profonde semble important. Cependant, il existe de nombreuses incertitudes sur la localisation et le mode d'action de ces afférences (action locale, générale ou suivant les espèces animales). Néanmoins, la libération de substances vasoconstrictrice est certainement modulée par les barorécepteurs artériels, un signal des récepteurs sensoriels cardiaques ou bien les deux.

Ces différentes observations sont encore obtenues avec les expérimentations suivantes : réduction de volume sanguin central par phlébotomie, par tilt-tests (angl. *foot-down tilting*) ou par des tests de *lower body negative pressure* (LBNP).

Conséquences de l'hypoxie tissulaire

L'hypoperfusion entraîne rapidement une hypoxie cellulaire. L'hypoxie se manifeste au niveau de tous les viscères de l'organisme provoquant une libération dans le sang de substances vasoactives (prostaglandines, kinines, histamine, leucotriènes, oxyde nitrique) ou cardioactives (*Myocardial Depressant Factor*), aggravant davantage la défaillance circulatoire. La transition d'un état de choc compensé à celui de choc décompensé serait marquée au niveau microcirculatoire par la perte du tonus vasoconstricteur des artérioles précapillaires. Tous ces facteurs aboutissent à une réduction irréversible du flux microcirculatoire (obstruction capillaire), d'où l'apparition d'un syndrome de défaillance multi-viscéral : insuffisance rénale et hépato-cellulaire, œdème pulmonaire, gastrite hémorragiques... De manière précoce, la circulation splanchnique se trouve sacrifiée, d'où l'apparition d'une ischémie intestinale et d'une augmentation de la perméabilité intestinale. Ceci pourrait favoriser la translocation bactérienne.

Action des anesthésiques

L'action des anesthésiques dépend des agents et des espèces² :

²pour une description des effets obtenus sur le modèle animal du porc, le lecteur est reporté à [20]

- Le pentobarbital sodium, l’halothane, la ketamine et le propofol atténuent ou même font disparaître la réponse vasoconstrictrice lors de la phase sympathoexcitatrice au cours de l’hémorragie (ou lors de la diminution du volume central au cours de test de LBNP ou lors d’une occlusion graduée de la veine cave). Ceci est due à l’action centrale de ces drogues.
- L’Halothane réduit aussi la contractilité cardiaque et est un vasodilatateur direct.
- Les opioïdes (agonistes μ), le fentanyl ou l’alfentanil, n’affectent pas la phase sympathoexcitatrice.

On notera que lorsque la phase sympathoexcitatoire est atténuée, il est difficile de mettre en évidence le début de la phase sympathoinhibitrice. L’alfentanil et les opioïdes (agonistes μ) suppriment la phase sympathoinhibitrice. Pour les hypothèses des mécanismes d’interactions des systèmes nerveux sympathiques et parasympathiques en particulier lors de syncope vaso-vagale (angl. *fainting*, *vaso-vagal syncope*) on peut encore consulter Dickinson [21] qui précise l’importance des barorécepteurs des gros troncs veineux et de l’oreillette droite.

3.1.3 Conclusion choc hémorragique

De nombreux facteurs interviennent au cours du polytraumatisme et du choc hypovolémique. Un schéma à 3 niveaux des différents médiateurs est proposé par Rose [22] et représenté Fig. 3.1.

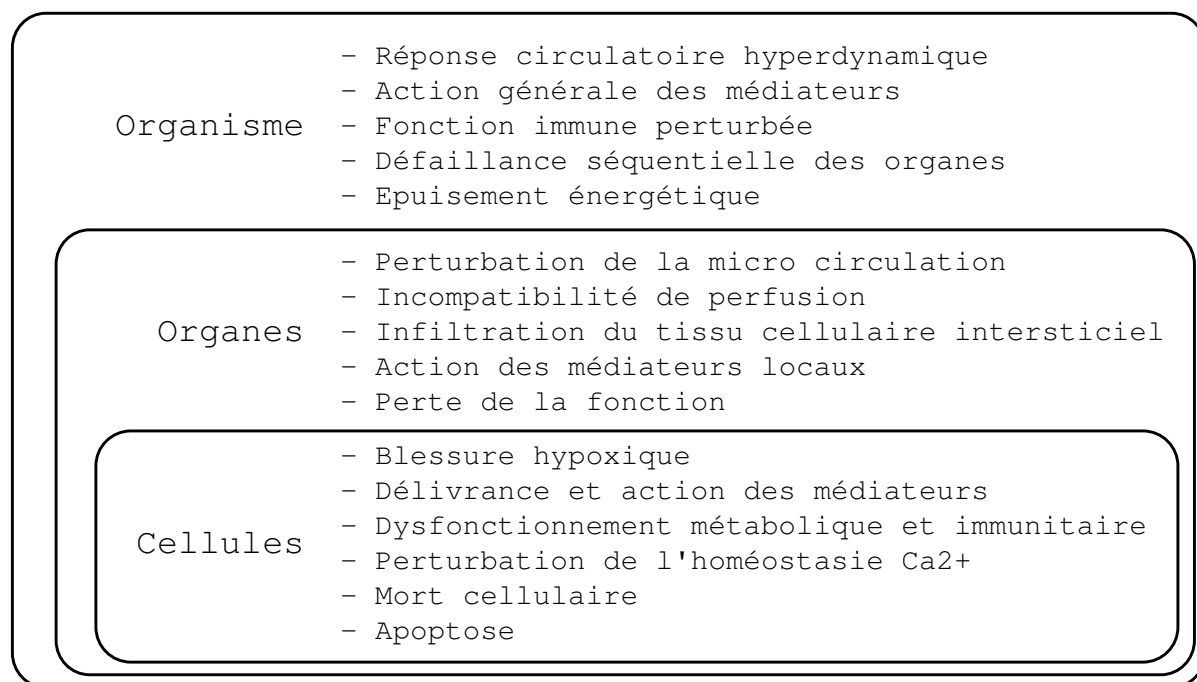


FIG. 3.1 – Modèle à 3 niveaux de l’action des médiateurs suite à une lésion traumatique d’après Rose [22]

Dans notre étude, nous plaçons notre observation d’un point de vue global, comme une résultante de toutes les actions des différentes régulations et contre-régulations de l’organisme pour se maintenir dans un état homéostatique. Nous observons des descripteurs généraux qui intègrent ces différentes informations (les variables observées sont les résultantes d’une sommation de l’organisme). Une attention particulière est à porter sur

l'utilisation d'agents anesthésiques, pour savoir quelle est notre interférence avec le mode de fonctionnement du système que nous étudions. Notre approche est systémique.

3.2 Le diagnostic des systèmes

3.2.1 Diagnostic médical

L'utilisation du terme de diagnostic (angl. *diagnosis*) a longtemps été réservé à son emploi médical. Son origine prend naissance dans les fondements de la médecine moderne initiée par Hippocrate dans le Corpus. Dans celui-ci, on parle peu de diagnostic mais souvent de pronostic comme l'indique Sournia dans son ouvrage « Histoire du diagnostic en médecine » [23]. Il précise encore cette démarche hippocratique : « le diagnostic est la conclusion portée sur un état pathologique à un instant donné, le pronostic annonce l'évolution dans le temps : les deux notions ne sont pas contradictoires, le pronostic dépend du diagnostic même si celui-ci n'est pas clairement exprimé. Pour Hippocrate, l'intervention du médecin n'est pas justifiée si le pronostic doit être fatal, mais pour cela le médecin doit *juger* de l'état du malade : la *crise* joue un grand rôle dans la vision hippocratique de l'évolution des maladies (gr. *krînô*, juger, apprécier, *krisis*, décision) et son traitement. »

Dans l'ouvrage d'Heggin [24], le diagnostic est présenté sous une version plus moderne : « chaque diagnostic posé par le médecin est en dernière analyse un diagnostic différentiel, c'est à dire une évaluation, une estimation, une différenciation des divers symptômes de la maladie. Pour faire une évaluation exacte, il est nécessaire de connaître la signification d'un *symptôme*. Un diagnostic différentiel se devra donc en premier lieu d'établir d'une manière générale dans quelles maladies on peut observer un symptôme donné. On doit faire état d'une part des *signes classiques* des diverses maladies, et d'autre part des *variantes* connues. »

Démarche diagnostique dans un monde multiparamétré

L'étude des signes, symptômes et syndromes associés à une maladie est généralement désignée sous le terme de sémiologie. De façon à réaliser son acte de diagnostic, parfois qualifié d'art, le médecin essaie donc de fonder son raisonnement sur le recoupement de ses observations avec ses propres connaissances : « Le médecin réalise souvent une juste pondération des différents symptômes. Il existe des symptômes tellement éloquents que leur seule présence permet de conclure avec certitude à l'existence d'une maladie déterminée. Dans certains cas, la présence simultanée de deux signes est nécessaire au diagnostic. Dans d'autres cas ce n'est qu'une triade de symptômes qui permettent le diagnostic » [24].

En d'autres termes, le médecin essaie de rechercher les véritables facteurs discriminants qui pourront lui servir à diagnostiquer une maladie. Les autres facteurs éloquents deviennent redondants, alors, dans ces conditions, on peut appliquer le principe du rasoir d'Occam³.

³Guillaume d'Occam ou d'Ockham : Philosophe anglais (Ockham, Surrey, v. 1285 - Munich v. 1349). Franciscain, excommunié, partisan du nominalisme dans la querelle des universaux et auteur d'une logique qui distingue les objets de pensée des catégories de la connaissance. Sa pensée a influencé la logique médiévale et préparé la doctrine de Luther en ébranlant les bases de la théologie médiévale [25]. Le principe du rasoir d'Occam est un argument rhétorique qui affirme qu'il ne sert à rien de multiplier les

On peut donc se demander si il existe une marche à suivre en pratique pour élaborer un diagnostic médical. Hegglin nous aide encore dans cette démarche en précisant que pour ce faire « le médecin doit se baser sur des considérations séméiologiques, celles qui correspondent le mieux à la situation se présentant au lit du malade. Il faut donc partir des symptômes ou syndromes (groupes de symptômes) et tenter de les subdiviser autant que le permet l'état actuel de la recherche, afin d'aboutir à l'entité nosologique. En pratique, lors de l'établissement du diagnostic différentiel, le médecin se trouve en présence d'un signe conducteur qui oriente sa réflexion et pose l'indication des examens complémentaires. Ce symptôme conducteur peut être révélé par l'anamnèse (par exemple une douleur abdominale haute), aussi bien que par l'examen clinique (par exemple une splénomégalie) ou par des résultats de laboratoire (par exemple la formule sanguine). » Les principales observations qui doivent être entreprises pour effectuer cette démarche conduisant au diagnostic de la maladie sont donc :

1. l'**anamnèse** et la **compréhension** de la personne du malade.
2. les **résultats de l'examen clinique**.
3. les **données du laboratoire**.

L'Anamnèse est une notion intéressante dans la démarche du médecin qui consiste à laisser parler librement le patient au sujet de sa maladie. Dans cette pratique, le médecin peut encore orienter subtilement vers une information pertinente. C'est aussi lors de cette phase que le médecin s'intéresse à l'historique du patient, ses affections antérieures, son environnement etc.

Dans cette situation, le médecin essaie d'éviter le plus de chemins possibles et essaye de se focaliser sur le cas particulier. C'est une stratégie qui est commode pour ne pas passer trop de temps à explorer toutes les pistes possibles mais à en déterminer quelques unes. En quelque sorte le médecin se met au diapason de façon à percevoir au mieux la réalité de son patient pour essayer de se focaliser sur ce qui l'intéresse et ne pas être dérouté par d'autres variables qui pourrait tromper l'observation.

3.2.2 Version Moderne

Approche intelligence artificielle

La présentation de Fieschi dans son ouvrage d'intelligence artificielle [27] propose un essai synthétique des différentes démarches diagnostiques : « dans le cadre de la décision médicale, qu'elle soit diagnostique, pronostique, thérapeutique ou préventive, les modélisations utilisées jusqu'à ces dernières années étaient basées principalement sur la théorie de la décision et les méthodes probabilistes ou statistiques. Les méthodes sont bien connues (celles qui se fondent sur la théorie de Bayes en particulier). Le médecin ne reconnaît pas son raisonnement et sa démarche diagnostique dans ce « calcul ». En particulier devant un malade particulier il n'utilise pas une démarche statistique. Si l'on se réfère à la classification établie par Murphy et rappelée par Williams il y aurait 4 types d'approches diagnostiques :

« entités » inutiles (lat. *Essentia non sunt multiplicanda praeter necessitatem*) ou qu'il est vain de faire avec plusieurs ce que l'on peut faire avec peu (lat. *Frustra fit per plura, quod fieri potest per pauciora*), en d'autres termes qu'une explication simple vaut mieux qu'une explication compliquée [26]...

l'approche exhaustive : non applicable au chevet d'un patient dont on ne connaît pas la maladie.

celle du « GESTALT » : référence à la théorie du GESTALT (mot allemand signifiant mot-à-mot : configuration, forme) développée par les psychologues allemands des années 30 pour proposer une nouvelle approche des phénomènes psychologiques de la perception. Elle nie la possibilité de décomposer tous les phénomènes en une somme de constituants élémentaires. C'est-à-dire qu'elle affirme que l'interprétation de certains éléments ne peut être dérivée de la somme de ses parties. Elle inclut des notions d'intuition et de données non verbales, comme par exemple la reconnaissance instantanée par un expert du faciès d'un patient présentant une péritonite. Cette approche est liée à l'identification du trait pertinent.

l'approche algorithmique : les règles de connaissance ne sont pas ambiguës et proposent de ce fait un guide pour agir de manière séquentielle. Ces algorithmes réalisent une véritable « architecture de base » permettant de repérer les principales situations qui peuvent se présenter. Il semble que ces deux dernières approches soient impliquées dans la phase de génération d'hypothèses du médecin. Toutefois, un organigramme ou le parcours d'une structure figée, telle une arborescence, ne permet pas de s'adapter par exemple à la contrainte d'un malade, et ne peut pas prendre en compte une maladie qui se présente sous des formes multiples, ne permet pas de fixer un contexte, c'est-à-dire un ensemble de faits qui doivent être pris en compte globalement pour décrire une situation dans laquelle une connaissance est applicable. En un mot, une telle approche manque de perspicacité.

hypothético-déductive : L'approche hypothético-déductive conditionne la saisie des données elle-même. En effet, l'hypothèse étant formulée, les observations complémentaires nécessaires sont réalisées pour confirmer ou infirmer cette hypothèse. Il arrive que cette approche ait pour trame un modèle physiopathologique. Toutefois, ce modèle est rarement complet et l'on retrouve ici le problème de la décision sous incertitude. Nous pensons que ces différentes approches ne sont pas mutuellement exclusives et que suivant le problème auquel il est confronté, le médecin utilise un subtil mélange de ces différentes approches. »

Mais dans cet ouvrage, on traite principalement des systèmes experts, pas d'apprentissage par la machine mais plutôt une mise à plat de la connaissance en vue de créer un système de diagnostic. La plupart des systèmes experts ainsi créés sont alors construits à partir d'une logique des prédicats du premier ordre ou par l'introduction d'une logique floue.

Approche qualité

De nos jours, le médecin a de moins en moins vocation à exercer son art dans la pratique du diagnostic. La version imposée par la société s'inscrit dans une démarche plus systématique. Les notions de qualité, de catégorisation, de classification sont de plus en plus rencontrées en médecine, d'une part pour des raisons de suivi et de traçabilité, d'autre part pour des raisons d'optimisation des pratiques médicales dans un domaine où le nombre de paramètres devient très important. En particulier, le paramètre économique prend une place de moins en moins négligeable et devrait reboucler sur une vision plus Hippocratique de la médecine. C'est cette orientation que la France a choisie en instituant son programme de médicalisation des systèmes d'information (PMSI). Dans celui ci, les

notions de diagnostic principal (DP) et de diagnostic associé (DA) sont introduites. Le diagnostic s'inclut alors dans un cadre défini le plus rigoureusement possible, et s'appuie sur une classification internationale des maladies (CIM). On constate encore un essai de hiérarchisation des connaissances, au même titre que n'importe quelle taxonomie.

PMSI [28] : ce programme, mis en place en France à partir de 1985 dans les établissements hospitaliers publics puis participant au service public est destiné d'une part à « connaître les différents types de malades et d'autre part à savoir attribuer à ces différents types de malades la part de dépenses hospitalières qu'ils ont consommée selon les différentes fonctions » (ministère de la Santé). Il sert donc à l'analyse médico-économique de l'activité hospitalière. – Le fonctionnement du PMSI peut être résumé ainsi : dans le service hospitalier ou plus précisément l'unité fonctionnelle (UF) est rédigé le compte rendu d'hospitalisation (CRH), d'où l'on extrait le résumé médical de séjour (RMS) comportant les mots-clés dérivés de la classification internationale des maladies (CIM). À partir de ce document est créé le résumé d'unité médicale (RUM) par la sélection d'un seul diagnostic principal (DP). Ce RUM est communiqué au département de l'information médicale (DIM) ; il constitue dans le cas du séjour dans un seul service le résumé de sortie standardisé (RSS), sinon ce dernier est composé à partir des différents RUM de chaque service. Le résumé de sortie anonyme (RSA) est un RSS modifié, communicable à l'administration. Le groupage des RSS aboutit à la classification en catégories majeures de diagnostic (CMD, p. ex. affections du système circulatoire), eux mêmes divisés en groupes homogènes de malades (GHM, p. ex. hypertension artérielle). En fonction des points d'indice synthétique d'activité (ISA) attribués à chaque GHM, les établissements hospitaliers MCO (médecine-chirurgie-obstétrique) se positionnent sur l'échelle des coûts.

Diagnostic principal : DP [28] Terme de gestion hospitalière. Ce DP est considéré « comme ayant mobilisé l'essentiel de l'effort médical et soignant » (ministère de la Santé). Sa mention dans le résumé d'unité fonctionnelle (RUM) permet de classer le séjour considéré dans la catégorie majeure de diagnostic (CMD) adéquate et d'établir correctement le PMSI.

Diagnostic associé : DA [28], Terme de gestion hospitalière. « Dans le cas d'un séjour multi-unité, il correspond aux diagnostics principaux des résumés d'unités médicales (RUM) non retenus comme diagnostics principaux (DP) ou aux diagnostics associés enregistrés sur les différents RUM décrivant le séjour » (ministère de la Santé) V. DP 2°, RUM et PMSI.

3.2.3 Diagnostic des Systèmes

Nous vous présentons ici une introduction d'un ouvrage récent de diagnostic des systèmes écrit par un spécialiste du domaine, Bernard Dubuisson [29], qui essaie de situer le diagnostic dans un contexte d'automatique par une approche systémique.

« Les étapes du processus de diagnostic sont définies à partir de la notion de mode de fonctionnement. Un mode de fonctionnement peut être défini comme l'ensemble des états du système qui le placent dans une situation de production et de sécurité donnée.

Le diagnostic est appliqué à un système, ce mot étant entendu dans le sens le plus général possible et pouvant tout aussi bien décrire une voiture qu'un procédé chimique, un être humain ou une rivière...

Le suivi du mode de fonctionnement d'un système est généralement décrit comme un processus en trois étapes, chacune méritant une attention particulière :

- la détection du mode sous lequel le système fonctionne,
- l'identification et la localisation de la cause d'un mauvais fonctionnement éventuel,
- l'action sur le système qui découle des deux étapes précédentes. Il peut s'agir du maintien dans le même mode de fonctionnement, de la correction du fonctionnement, ou même de son arrêt si nécessaire. »

Ainsi, le diagnostic vise à détecter dans quel mode de fonctionnement est l'état d'un système, c.-à-d., différencier et choisir les modes de fonctionnement qu'ils soient bons ou mauvais. Zwingelstein [2] adopte une approche similaire dans son ouvrage de diagnostic des défaillances pour les systèmes industriels et en précise les termes adoptés par la communauté internationale. « La définition donnée par les instances nationales et internationales ⁴ est la suivante : « Le diagnostic est l'identification de la cause probable de la (ou des) défaillances à l'aide d'un raisonnement logique fondé sur un ensemble d'informations provenant d'une inspection, d'un contrôle ou d'un test ». La défaillance est définie par « l'altération ou la cessation de l'aptitude d'un ensemble à accomplir sa ou ses fonctions requises avec les performances définies dans les spécifications techniques ».

On peut encore intégrer le diagnostic dans un cadre plus large de supervision et de surveillance des systèmes [30] et le rapprocher du monitoring (angl. *monitoring*). Les fonctions réalisées sont alors les suivantes :

- Détecter la présence d'une défaillance et avertir au moyen d'une alarme.
- Localiser l'endroit du procédé où la défaillance est apparue et si possible en préciser la nature.
- Identifier une valeur associée à la défaillance de façon à préciser l'importance de celle-ci.

Les procédures de localisation et d'identification correspondent au diagnostic et nécessitent une connaissance des causes de défaillances et des types de défauts associés au système. Pour éclaircir ces notions de défaillances, de leur cause et de leurs effets, Zwingelstein insiste sur le principe de décomposition fonctionnelle qu'il présente comme indispensable. Il propose alors un résumé pour sélectionner la méthode diagnostic la plus appropriée à un système industriel donné en proposant un recensement des besoins et des connaissances disponibles. Les éléments indispensables à étudier étant les suivants :

- nature des causes de défaillance à localiser,
- connaissance des symptômes associés aux défaillances induites par les causes,
- maîtrise des moyens de mesure des symptômes,
- maîtrise de moyens de traitement des symptômes,
- connaissance des mécanismes physiques entre les causes et les effets,
- inventaire du retour d'expérience,
- recensement des expertises disponibles,
- définition du niveau de confiance dans le diagnostic,
- identification des utilisateurs finals du diagnostic. »

Mais avant de décrire les démarches pour l'élaboration d'un modèle de diagnostic il est important de revenir sur la classification actuelle qui semble émerger des différents groupes de travail sur les modèles de diagnostic.

⁴Association française de normalisation (AFNOR), International Electrotechnical Commission (IEC).

Les différents modèles de diagnostic

Les différents types de modèles de diagnostic en automatique peuvent être répertoriés ainsi selon l'approche de [2], ils sont repris sous la désignation d'*approche analytique* et d'*approche heuristique* par Evsukoff [30] :

méthodes internes de diagnostic Ces méthodes sont basées sur une connaissance physique du modèle (angl. *model-based approach*) : on part d'un ensemble de lois physiques qui fournissent un ensemble de règles analytiques entre les variables qui caractérisent le système étudié. L'observation du processus est confronté à ce modèle. La différence entre les variables observées et les variables prédites par la modélisation mathématique forme les résidus. Des algorithmes de détection sont appliquées sur ces résidus pour déclencher un processus de décision ou un diagnostic.

méthodes externes de diagnostic Ces méthodes sont utilisées lorsqu'un modèle formel n'est pas disponible mais lorsque seules les observations du système constituent la connaissance du système. On adopte alors une approche de type reconnaissance de formes en recherchant un modèle sur l'ensemble des observations disponibles constituant l'ensemble d'apprentissage et en appliquant ces modèles à de nouvelles observations. Les modèles générés par cette approche sont de différents types : statistiques, flous, réseau de neurones, systèmes experts. La création de ces modèles est réalisée en plusieurs étapes qui correspondent à : une recherche des observations les plus pertinentes pour la création du modèle, une recherche des caractéristiques les plus pertinentes qui constituent le vecteur forme de l'étude, une recherche des partitions de l'espace des formes permettant d'obtenir les régions de décisions correspondant aux modes de fonctionnement [31]. On peut représenter schématiquement les principes d'une telle approche par les Fig. 3.2 et 3.3.

Conclusion Dans tous les cas, on recherche un mode de fonctionnement, c.-à-d., un critère appliqué à un modèle correspondant le plus possible à la réalité. Suivant la connaissance disponible sur ce modèle, on diagnostiquera une défaillance ou un mode de fonctionnement par observation de l'état du système et le rapprochement ou l'éloignement de celui-ci à une région déterminée. La recherche de ces modèles résidera donc dans la détermination de ces régions de fonctionnement qu'il est pratique de représenter par des formes dans des espaces appropriés. Dans une approche expérimentale, la confrontation des données et l'identification de paramètres sera alors une étape essentielle pour la construction des modèles. Les ordinateurs et leur grande capacité de calculs permettent désormais de faire émerger automatiquement des formes dans un ensemble de données représentant l'état d'un système (c'est l'utilisation de l'informatique dans son sens étymologique).

Au final, le diagnostic du système reposera sur une véritable chaîne de traitement de l'information. La plus courante peut être schématisée par la succession étagée suivante [34] : Phénomène → Capteur → Signal → Traitement du signal → Signal → Extraction de caractéristiques → Données → Traitement de données → Données d'ordre sémantique élevé → Organe de décision → Décision.

Dans notre cas, on peut transformer la chaîne en : Phénomène → Capteur → Mesure → Prétraitement du signal → Signal → Extraction de caractéristiques → Caractéristiques (Données) → Transformation → Représentation → Classification → Classe (Décision).

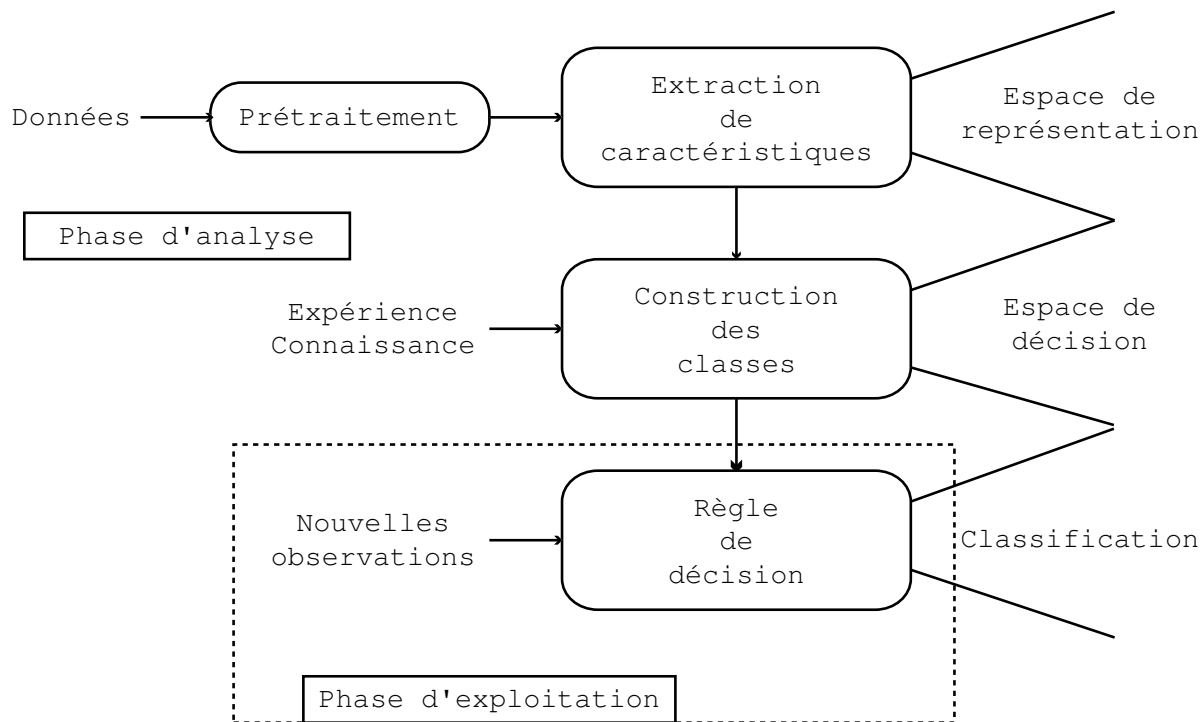


FIG. 3.2 – Élaboration d'un système de décision par reconnaissance de formes [32]

3.2.4 Découverte de connaissance à partir de données

L'utilisation de l'informatique (en particulier par l'intermédiaire de la théorie de l'information) facilite grandement la recherche de modèles basés sur une reconnaissance de formes (angl. *pattern recognition*). Celle-ci est généralement réalisée par un apprentissage artificiel (angl. *machine learning*) des données expérimentales disponibles. Une large branche de l'intelligence artificielle leur est consacrée. Depuis un certain nombre d'années, on constate l'émergence de nouveaux domaines tels que la découverte de connaissance (angl. *knowledge discovery*). La différence avec le diagnostic (étymologiquement, ce qui conduit au savoir) c'est que celui-ci fait référence à une maladie ou une défaillance, à la découverte de sa nature et de son instant d'apparition, alors que la découverte de connaissance fait référence à l'émergence de... n'importe quel type de connaissance. L'utilisation des techniques de fouille de données (angl. *data mining*), essaie de faire émerger cette connaissance comme une différenciation dans un ensemble vaste de données. Le but est essentiellement le même que le diagnostic et est lié au caractère cognitif de n'importe quel modèle d'explication d'un phénomène, d'une forme, d'un symptôme, en vue de comprendre, de réaliser un pronostic, ou encore de contre-réagir sur le phénomène observé. Aussi les différentes méthodes pour essayer de rendre intelligible (lat. *intelligere*, comprendre) les différents systèmes étudiés seront d'un grand intérêt. Elles sont l'objet des sciences cognitives.

Application en médecine

Les méthodes d'intelligence artificielle ont donné lieu à de nombreux travaux en médecine depuis une trentaine d'année, en essayant de tenter de capturer la connaissance experte des praticiens [35, 36]. La revue *Artificial Intelligence in Medicine* en témoigne.

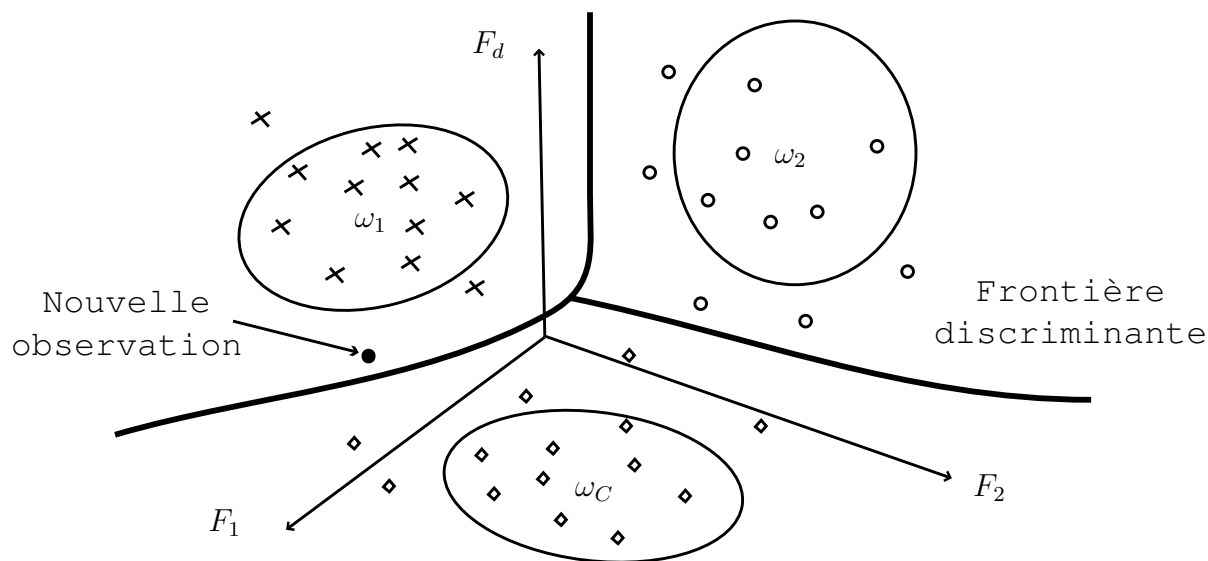


FIG. 3.3 – Principes de la reconnaissance de formes [33, 32] : F_i caractéristiques, ω_i : classe attribuée

Plusieurs études utilisent des méthodes d'apprentissage supervisé, entre autres, on peut retenir les études que nous avons explorées, qui sont confrontées à des problématiques ou des méthodes proches des nôtres comme la validation et la comparaison de classifieurs [37, 38], la sélection d'attributs [39, 40], la classification diagnostic en unités de soins intensifs [41, 42]...

3.2.5 Diagnostic du choc hémorragique, cas de notre étude

Dans notre étude, nous utiliserons un savoir expert concernant une problématique donnée : le polytraumatisé vu comme un blessé susceptible d'être à l'un des différents stades d'évolution d'un choc hémorragique. Pour cela, nous demanderons à un médecin de diagnostiquer les différentes classes correspondant à des états de choc rencontrés sur un animal subissant une perte de sang à débit contrôlé, par l'utilisation de ses propres connaissances et l'observation des paramètres retenus extraits des signaux recueillis au cours d'une expérimentation. Nous essaierons alors de faire apprendre plusieurs modèles informatiques de cette classification experte à une machine par apprentissage supervisé. Nous réaliserons ainsi une reconnaissance de formes. Nous nous sommes donc largement servi de l'ouvrage Diagnostic et Reconnaissance de formes de M. Dubuisson [33] dans un premier temps puis de l'ouvrage récent de Cornuéjols [26] sur l'apprentissage artificiel.

Peu d'articles évaluent les effets du choc hémorragique d'une façon multidimensionnelle ou en essayant de réaliser un modèle proche de celui que nous proposons. D'un point de vue historique, différents domaines médicaux ont mis en place un certain nombre de grilles d'évaluation, de classifications ou de scores, de façon à standardiser et optimiser leurs pratiques. Nous présenterons dans un premier temps les principaux scores utilisés en réanimation et ceux utilisés plus spécialement en traumatologie. Nous développerons ensuite les recherches de modélisation qui s'intéressent spécifiquement au choc hémorragique. En particulier, nous présenterons brièvement l'article de Glass paru en troisième année de thèse qui se rapproche de notre travail et sur lequel nous reviendrons en dernière partie de ce rapport.

3.3 Modèles de choc hémorragique

3.3.1 Évaluation en réanimation et traumatologie

En réanimation, plusieurs indices sont utilisés de façon à catégoriser les blessés et permettre ainsi une meilleure efficacité lors de la prise en charge des patients par application de principes généraux au cas individuel. Ils permettent en outre d'évaluer l'efficacité des soins apportés par une analyse a posteriori. Ces indices, fondés dans un premier temps sur l'observation pratique des médecins et la détermination des paramètres qui conditionnent au mieux l'évolution, sont désormais réalisés par des méthodes statistiques [43].

Les indices de gravité en réanimation permettent :

- d'une part, de quantifier un état de façon à obtenir une échelle graduée pour l'approche diagnostique.
- d'autre part, pour un certains nombre de ces scores, de réaliser un pronostic sur la probabilité de survie du malade.

En traumatologie, on distingue principalement deux types de scores et un mixte :

- les scores physiologiques, calculés à partir de paramètres cliniques (paramètres physiologiques, examens biologiques, atteinte neuro-physiologique...)
- les scores anatomiques qui permettent d'apprécier la gravité en fonction des lésions anatomiques observées. Ces derniers ne peuvent être donc calculés efficacement qu'à l'hôpital après un bilan lésionnel précis par un personnel compétent.
- Enfin, les scores qui associent les deux types de paramètres servent surtout à l'évaluation de la qualité des soins [44].

Une analyse des principaux scores est disponible dans [45, 46] et est proposée en annexe B. Ces scores sont consultables et mis à jour selon l'amélioration de leur précision, sur différents sites comme par exemple celui de la Société Française d'Anesthésie et de Réanimation : <http://www.sfar.org>. D'autres scores sont disponibles suivant les spécialités recherchées que nous ne développerons pas dans cette étude mais reportons le lecteur intéressé sur le site du projet *Medical Algorithms* pour exemple (<http://www.medal.org/>) qui en répertorie un grand nombre (et d'autres encore).

Le résultat d'une telle analyse montre que, quel que soit le score utilisé ou l'échelle mise en place pour évaluer l'état du patient et pronostiquer sa survie, les valeurs physiologiques restent d'une importance majeure dans la création du score qui va aider le médecin à faire le choix de la thérapeutique à adopter, ou pour évaluer son urgence. Ces scores sont en général réalisés à l'admission du malade au centre de soins intensifs. D'une part, le relevé des différents paramètres requiert un coût en temps de travail pour le personnel hospitalier, d'autre part, l'estimation de la survie par ces modèles se fait par rapport à un état, réalisé à l'instant t de la personne traumatisée. Pour l'aider à la prise en charge de ces individus, et surveiller leurs évolutions pour adapter la thérapie, un appareil de monitoring de plus en plus performant est mis en place.

Le système que nous proposons va essayer de prendre en compte un ensemble de mesures réalisées en continu. Il permettra donc de suivre l'évolution du blessé. L'information anatomique lésionnelle pour caractériser les différents traumatismes et les associer à une classification de type *abbreviated injury scale* (AIS), classification internationale des maladies (CIM) ou *anatomic profile* (AP), ne peut être facilement envisagée sur le terrain. De la même façon, un score permettant de diagnostiquer par un système autonome l'état de conscience de l'individu n'est pas évident à mettre en place (classification des états

de conscience par électroencéphalographie difficilement réalisable avec l'avancée des recherches actuelles, en tout cas, pas à notre connaissance). Aussi, le système que nous proposons ne se fonde que sur un ensemble de mesures physiologiques. Nous avons donc choisi de créer celui ci à partir d'une classification réalisée à partir des différents phases de choc hémorragique décrites par la littérature et définies en début de chapitre.

On fait de plus l'hypothèse que les soldats secourables font partie d'un ensemble lésionnel correspondant à une section d'un gros tronc veineux. Pour une blessure plus importante (score de type A pour le score AP p. ex.), le facteur pronostic est très faible pour l'individu et ne permet pas la mobilisation des moyens de récupération (nous sommes alors dans un cas d'urgence dépassée, les secours ne pouvant arriver sur les lieux avant le décès du blessé qui nécessite une prise en charge immédiate). Nous nous bornerons dans notre travail à l'étude des mécanismes associés à une perte de sang continue compatible avec des blessés secourables sur le terrain.

L'utilisation d'agents anesthésiques pour la mise en place d'un modèle de perte de sang à débit contrôlé sur l'animal implique que le facteur neurophysiologique de l'état de conscience soit considéré comme nul. Notre travail utilisera donc un modèle avec absence de conscience (anesthésie), ce qui aura pour effet de biaiser les résultats pour un état de choc peu avancé.

3.3.2 Modélisation mathématique du choc hémorragique

Cannon utilise un modèle mathématique [47] comprenant 3 paramètres pour la modélisation de l'évolution des variables au cours du choc hémorragique. Les variables retenues sont la *PAM*, la fonction neurologique (angl. *neurologic function*) et le déficit acido basique systémique (angl. *systemic acidosis/base deficit*). Un système d'équations différentielles construit sur des variations autour de différentes valeurs d'équilibre et limites, essaie de mettre en évidence les relations physiologiques complexes entre les variables du modèle. Le système proposé permet de pronostiquer le temps de survie par observation des courbes ainsi générées. Les trajectoires dépendent donc des états initiaux. Ils mettent en évidence l'existence d'une valeur limite pour la pression artérielle et compare celle ci aux modèles de Guyton et Cromwell [17].

3.3.3 Pronostic, hémorragie non contrôlée et réanimation

Roberts [48] essaie de prédire la mortalité d'animaux ayant subi des chocs hémorragiques importants (veineux ou artériels non-contrôlés) et ayant été pris en charge selon différents protocoles de réanimation par administration de fluides pour un retour à une pression artérielle moyenne de base. Pour cela il utilise des modèles mathématiques composés de réseaux de neurones polynomiaux de façon à prédire la survie à 20 *min* ou 60 *min*. Les variables de *PAS* et *PAM* sont utilisées comme entrées des réseaux. Les données proviennent d'expérimentation animale : 142 porcs ayant subi une lésion du foie ou une aortotomie. Une splénectomie a été systématiquement réalisée (animaux de 40 ± 5 kg) car on constate une constriction importante de la rate au cours du choc chez cet animal pour des individus de cet âge (ce qui peut induire une différence physiologique importante par rapport à la réponse observée chez l'homme). Leur conclusion indique que la *PAM* ne peut, à elle seule, discriminer les survivants des non survivants, que le nombre d'animaux est insuffisant pour leur modèle et que l'utilisation de méthodes non contrôlées pour l'hé-

morragie peut affecter la discrimination du modèle. La détermination d'un pronostic a 20-30 minutes est soulignée et indique un intérêt réel pour les problèmes de triage civil ou militaire puisqu'il correspond aux moments des prise de décision.

3.3.4 Classification par perte de sang, hémorragie contrôlée

Glass et coll. ont publié un article [49] en mars 2003 au cours de la troisième année de thèse concernant l'utilisation des techniques d'intelligence artificielle pour identifier le compromis cardiovasculaire dans un modèle de choc hémorragique. Leurs résultats indiquent qu'en utilisant uniquement l'ECG, et en calculant sur celui ci les caractéristiques d'intervalle RR, d'amplitude QRS et d'intervalle RS, il est possible d'obtenir un modèle capable de diagnostiquer une classe correspondant à un niveau de perte de sang. 19 animaux de 25 à 35 *kg* mâles et femelles ont été utilisés. L'hémorragie contrôlée a été réalisée par ponction, suivant des paliers (aliquots) de 10, 20, 30, 35, 40, 45, 50% du volume sanguin total.

Étant donné la forte similarité avec notre étude, nous avons repris une grande partie de cet article que nous avons commenté et présenté dans la partie IV. Des questions ont été posées à son équipe avec une première réponse relatant un intérêt certain pour ce type d'études et d'un futur éclaircissement sur les questions posées, malheureusement sans suite depuis (dernier envoi le 20 mars 2004, cf. Annexe D).

La mise en place d'une expérimentation par perte de sang suivant des paliers, a été envisagée lors de la mise en place du protocole mais nous ne l'avons pas retenue. En effet, il nous semble que la succession de ces paliers fait partir l'individu sur une base nouvelle à chaque étape. Il serait intéressant d'étudier ces différents paliers, sur différents individus, en partant de l'état stable de base uniquement. Nous avons retenu une perturbation de type rampe car elle correspond à une réalité physiologique, la phlébotomie ou la perte de sang continue non garrottée d'un individu isolé. Le paramètre de débit de la perte est constant (entrée de type échelon, perte de sang de type rampe).

3.4 Conclusion du chapitre

Le corpus médical a acquis une base de connaissance importante concernant l'étude du choc hémorragique et les différents signes permettent de diagnostiquer l'évolution du choc hypovolémique résultant en plusieurs classes. Néanmoins, de nombreux facteurs entrent en jeu dans la réactivité du système : au niveau cellulaire, au niveau des organes, au niveau de l'organisme pris dans son intégralité, l'ensemble relié par de nombreux agents médiateurs, au niveau de la réponse du système nerveux et en particulier au niveau des voies afférentes que constituent les barorécepteurs... La grande diversité des variables micro, méso et macroscopiques liés au système pris dans son intégralité peut expliquer qu'il existe encore de nombreuses incertitudes sur la compréhension des mécanismes et des résultantes que l'on peut observer sur les variables intégrées.

Nous avons vu la démarche du diagnostic médical qui se rapproche de plus en plus d'une approche systémique employée en automatique lors de la surveillance des systèmes. Dans la possibilité de mettre en place un système capable de déterminer la classe de choc hémorragique d'un blessé polytraumatisé, nous nous sommes rendu compte que les techniques de reconnaissances de formes utilisées par des méthodes externes de diagnostic

(aucune connaissance sur le modèle physique décrivant les relations de causes à effets n'est utilisée, mais on utilise la connaissance d'un expert que l'on essaie de cerner et mimer), nécessitait un apprentissage supervisé et par là-même une incursion dans une approche d'intelligence artificielle et de sciences cognitives.

Aussi, nous avons remarqué que les principaux scores créés par nos confrères, résultaient de plusieurs tentatives : dans un premier temps, un grand souci de classification, souvent résumé sous forme de tableaux et de scores a été présenté de façon à catégoriser les maladies non plus par particularités mais par grandes classes de patients de façon à faciliter la démarche intellectuelle du médecin, d'autre part, l'utilisation de statistiques de plus en plus recherchées et principalement de régression de type logistique, permet encore de prédire la mortalité des patients et donc d'en pronostiquer la survie et de tout mettre en place pour améliorer celle-ci.

L'utilisation des techniques d'analyse multidimensionnelle pour la compréhension des phénomènes observés au cours du choc hémorragique, semble être une voie prometteuse. L'utilisation de l'informatique rend possible cette approche et c'est celle-ci que nous avons essayée d'utiliser dans notre étude. Nous avons de plus souligné l'intérêt de travailler sur le modèle animal porcin, comme de nombreuses équipes l'ont constaté, et nous avons donc appliqué notre démarche à ce modèle.

La méthode que nous avons choisie ressemble beaucoup à celle utilisée par Glass et coll. Ceci conforte en partie les choix que nous avons réalisés et augmente la croyance en notre modèle. Les principales différences sont : la perte de sang à débit contrôlé ; une classification de l'état de choc par un expert correspondant à une constatation physiologique (choc compensé, décompensé, irréversible) ; des algorithmes de classification et des méthodes de validation permettant d'obtenir une estimation des performances en généralisation la plus réaliste possible et non dépendante des individus.

Deuxième partie

Étude réalisée

Chapitre 4

Matériel

Il faut donc nécessairement, après avoir disséqué sur le mort, disséquer sur le vif, pour mettre à découvert et voir fonctionner les parties intérieures ou cachées de l'organisme; c'est à ces sortes d'opérations qu'on donne le nom de vivisections, et sans ce mode d'investigation, il n'y a pas de physiologie ni de médecine scientifique possibles : pour apprendre comment l'homme et les animaux vivent, il est indispensable d'en voir mourir un grand nombre, parce que les mécanismes de la vie ne peuvent se dévoiler et se prouver que par la connaissance des mécanismes de la mort.

Claude Bernard, *Introduction à l'étude de la médecine expérimentale*, 1984 (1865), Flammarion, Paris, p. 140.

Ce chapitre présente la partie expérimentale de notre étude qui a consisté en la mise en place d'un protocole de choc hémorragique sur des animaux, et une acquisition d'un ensemble de mesures regroupées au sein d'enregistrements qui ont servi de base pour le traitement des données. Nous présentons dans un premier temps la prise en charge des animaux, le protocole utilisé et les réglages nécessaires au bon déroulement de ces expérimentations dans la section 4.1. Nous présentons dans la section 4.2, les mesures réalisées et les principales notions liées à l'acquisition des mesures plus particulièrement dans notre contexte médical. Enfin, nous présentons dans la section 4.3, les caractéristiques retenues en vue d'une analyse de données (apprentissage artificiel et reconnaissance de formes) et les méthodes que nous avons utilisées pour calculer celles-ci.

4.1 Expérimentation Animale

4.1.1 Modèle Animal

Protocole expérimental

Nous avons mis au point ce protocole expérimental à partir des articles disponibles et de la connaissance de divers confrères. Ce protocole, soumis au Comité d'Éthique (Comité Consultatif pour l'utilisation des animaux en recherche expérimentale) du CRSSA, permet une acquisition des données utilisable en routine et respecte les recommandations du décret 87-848 relatif aux expériences sur les animaux vertébrés.

24 animaux ont été utilisés lors des expérimentations qui se sont déroulées de mars 2002 à juillet 2003. 5 animaux ont servi pour acquérir l'expérience nécessaire et la mise en place d'un protocole utilisable en routine dans les conditions définies ci dessous.

Les animaux, porcs charcutiers (espèce *Sus scrofa*), jeunes adultes, femelles, ont été stabulés dans un box commun par 3 ou 4 pour les 6 premiers animaux puis dans des box individuels (Fig. 4.1) au département de biologie appliquée (BA) du CRSSA. L'équipe des zootechniciens de ce département a pris en charge le transport, la stabulation, la nutrition et le soin des animaux pendant une durée minimale de 8 jours d'accommodation des animaux.



FIG. 4.1 – Stabulation de l'animal : l'animal est stabulé en cage individuelle, d'une capacité supérieure à celle recommandée par la communauté européenne. Un accès à l'eau par biberon est proposé ad libitum, une nourriture composée de granules enrichis en vitamines et minéraux ainsi que des légumes frais coupés en morceaux est servie dans une mangeoire quotidiennement

Nous nous sommes largement inspirés des publications concernant l'utilisation du modèle animal porcin mais aussi des pratiques réalisées sur l'homme :

- Les planches anatomiques [50]
- L'anesthésie-analgésie [20]
- Le choc hémorragique [51]
- L'utilisation du matériel [52, 53]

Le choix de l'espèce a été justifié par défaut : il n'existe plus que 2 modèles animaux de choc hémorragique qui soit encore retenus par la communauté scientifique internationale : le rat et le porc (le modèle chien n'étant plus accepté pour des raisons sociales). Le modèle rat est presque exclusivement tourné vers la réponse immunitaire et cellulaire locale du choc hémorragique, tandis que le second est utilisé pour étudier les réponses cardiovasculaires, ventilatoires et neurologiques au choc. Le modèle porc s'imposait aussi de par la dimension ergonomique [54].

Expérimentation animale

Le synoptique de l'expérimentation est donné Fig. 4.2. Le temps total de l'expérimentation peut durer 8 h (de 8 h à 16 h).

1. Travail sur porc (*Sus scrofa*) juvénile femelle de 20 kg.

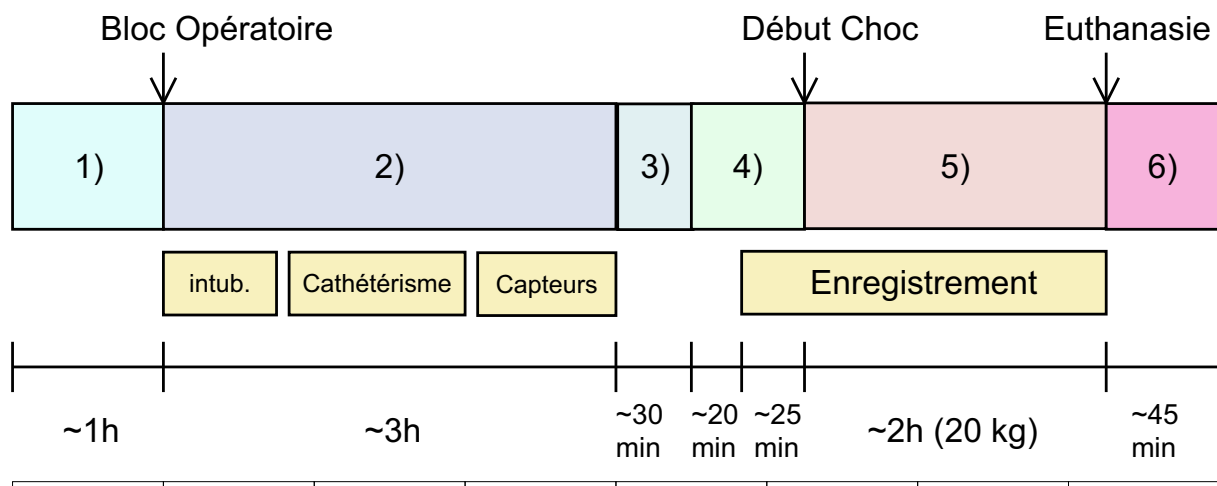


FIG. 4.2 – Synoptique de l'expérimentation : 1) Sédation, Préparation de l'animal 2) Chirurgie, Anesthésie Gazeuse, Instrumentation de l'animal 3) Respiration Contrôlée 4) Stabilisation 5) Spoliation Sanguine 6) Nettoyage du bloc

2. Prémédication par *Azapérone* (2 mg/kg , *Stresnil*®) injecté par voie intra musculaire (i. m.). La prémédication a pour but de tranquilliser l'animal et de diminuer la posologie des anesthésiques administrés, elle potentialise leurs effets. « L'*Azapérone* est un neuroleptique du groupe des butyrophénones spécialement mis au point pour les suidés. Administré 15 à 20 min avant l'anesthésie, l'animal restant dans son enclos au calme, il produit une indifférence vis-à-vis de son entourage et diminue son activité motrice. Il prévient bien le stress chez certains animaux nerveux et facilite leur capture. La durée d'action varie de 1 à 3 h suivant la dose employée qui peut s'élever jusqu'à 4 mg/kg i. m. » [20]
3. Induction anesthésique 30 min après par *Kétamine* (15 mg/kg , *Imalgène 1000*®) et *Xylazine* (2 mg/kg , *Rompun 2%*®). « La *Kétamine*, anesthésique dissociatif, est très largement utilisée car facile à administrer et dotée d'une large marge de sécurité. Elle procure une psychodyslepsie (le sujet semble dissocié de son environnement), une analgésie superficielle, une sédation, une immobilisation avec des mouvements anormaux et une catalepsie. Elle laisse le réflexe laryngé intact, procure une mauvaise myorelaxation musculaire et la salivation est abondante. C'est pourquoi on l'associe à la *Xylazine* qui assure une myorelaxation excellente et une analgésie importante. La dépression cardiorespiratoire est alors importante et à surveiller. » [20]
4. L'animal est alors préparé suivant la procédure vétérinaire d'usage dans une salle prévue à cet effet : pesage, nettoyage de l'animal, rasage des soies de l'animal au endroit d'abords chirurgicaux ou aux endroits de mise en place des électrodes. L'animal est ensuite conduit au bloc opératoire, où il est disposé sur la table d'opération.
5. Avant introduction de la sonde endotrachéale, injection de *Buprénorphine* (0.3 mg dans 5 ml de sérum physiologique, *Temgesic*®). Ce dérivé morphinique permet d'obtenir une analgésie plus importante appréciée lors de l'intubation et des premiers temps opératoires. Une dépression respiratoire est observée et demande donc une intubation rapide.
6. Intubation endotrachéale par voie naturelle (Fig. 4.3 et ventilation mécanique par un mélange air-oxygène ($F_i\text{O}_2$: 50%, V_E : 4 à 5 l/min) enrichi en *isoflurane* (entre

0.5 et 1%). Le passage à l'anesthésie gazeuse permet un confort appréciable lors des

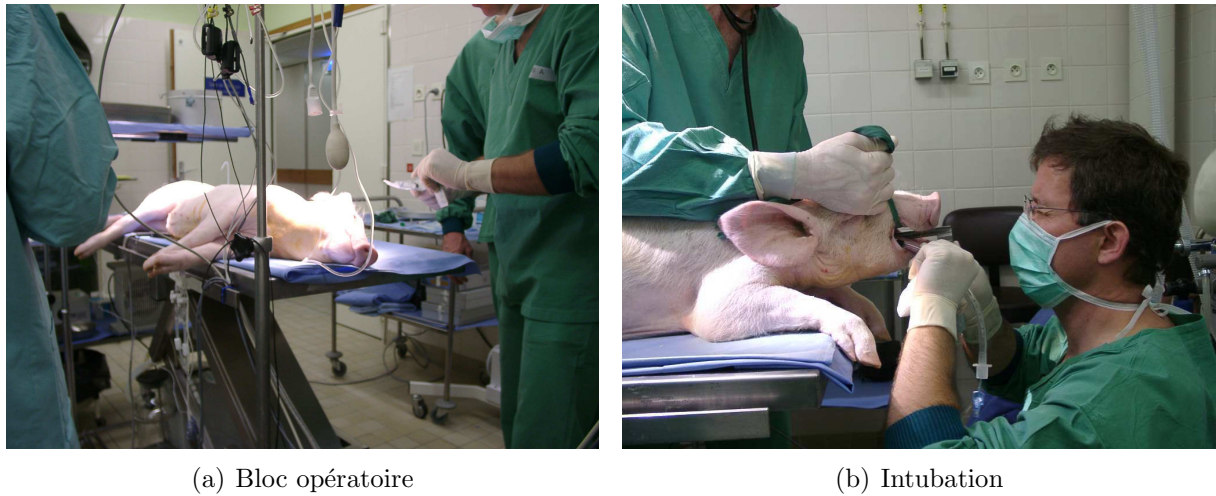


FIG. 4.3 – Intubation de l'animal : l'animal est intubé en position ventrale. Le voile du palais est légèrement relevé. L'anatomie est légèrement différente de celle de l'homme.

différents temps opératoires. Les dérivés halogénés ont un temps d'action rapide et la régulation de leur concentration gazeuse par monitoring permet un moyen d'action sûr pour le contrôle de la profondeur d'anesthésie, la myorelaxation et l'analgésie. Bien que diminuant le tonus sympathique, l'*isoflurane* induit une faible dépression myocardique et son métabolisme hépatique est quasi nul. Il ne sensibilise pas le myocarde aux catécholamines. La diminution de pression artérielle qu'il entraîne est due à la diminution de la résistance périphérique totale. Il induit une diminution de la résistance coronaire. On veillera donc à respecter un temps d'élimination et de récupération pour diminuer ces effets délétères sur le modèle de choc hémorragique retenu.

7. Cathétérisme carotidien et jugulaire externe après dénudation chirurgicale (Fig. 4.4) : un cathéter 2 voies (Gauge *G20*) est amené dans la crosse de l'aorte ; un cathéter 3 voies (*G22*) est amené dans l'oreillette droite à l'abouchement de la veine cave supérieure.
8. Pose des capteurs et réglages des paramètres de calibration. Vérification de la bonne stabilité des instruments. Les différents capteurs disposés sont présentés dans la section 4.2
9. A la fin de cette période de préparation, l'animal est déplacé de la table chirurgicale sur le hamac de suspension (position naturelle) ou bien placé sur la table d'opération en position decubitus latérale qui sont les positions les plus adaptées pour ce type d'expérimentation [51]. Les capteurs sont réajustés et une perfusion à la pompe de *propofol* (20 à 60 mg/h, *Diprivan 1% ®*) est mise en place par abord intra veineux (i. v.). L'utilisation de cet hypnotique légèrement analgésique permet d'obtenir une excellente myorelaxation. Son délai d'action rapide permet une utilisation par pousse-seringue comparable à une anesthésie gazeuse. Une administration de « bolus » est réalisée lors de la constatation d'un réveil de l'animal et le débit de la perfusion est adapté à la réaction selon les valeurs indiquées précédemment. Les modifications cardiovasculaires sont minimales d'après [20].

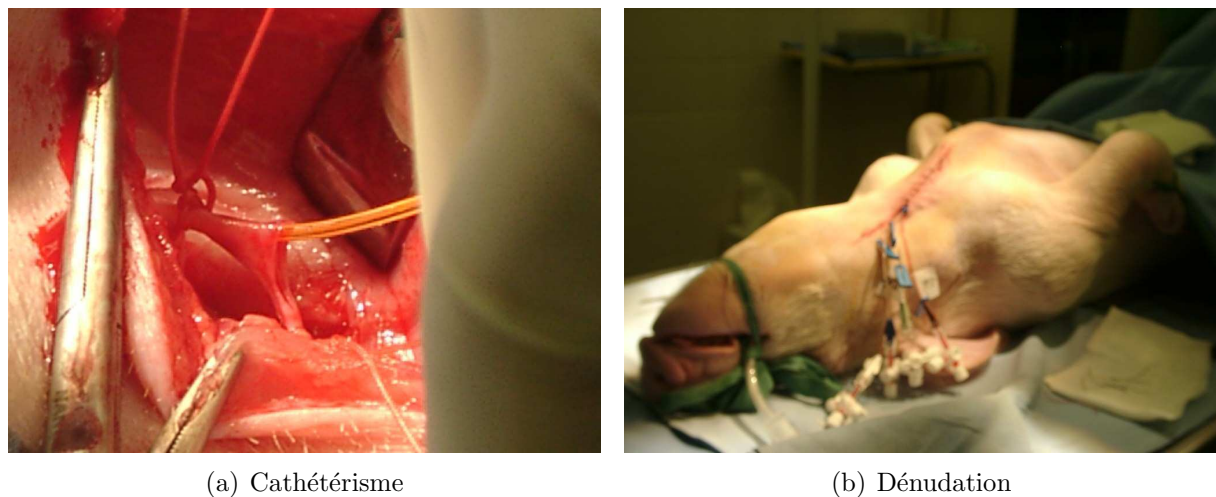


FIG. 4.4 – Cathétérisme : un cathéter carotidien et un cathéter veineux sont introduits pour la mesure, l'apport de drogues anesthésiques et les prélèvements sanguins.

10. La ventilation mécanique est alors arrêtée, ainsi que l'anesthésie gazeuse. A ce moment, l'animal est en ventilation spontanée à l'air et en position naturelle. L'animal est alors laissé dans cet état jusqu'à ce que la stabilisation des variables physiologiques (PA , FC , RR , E_tCO_2) soit constatée (Fig. 4.5).

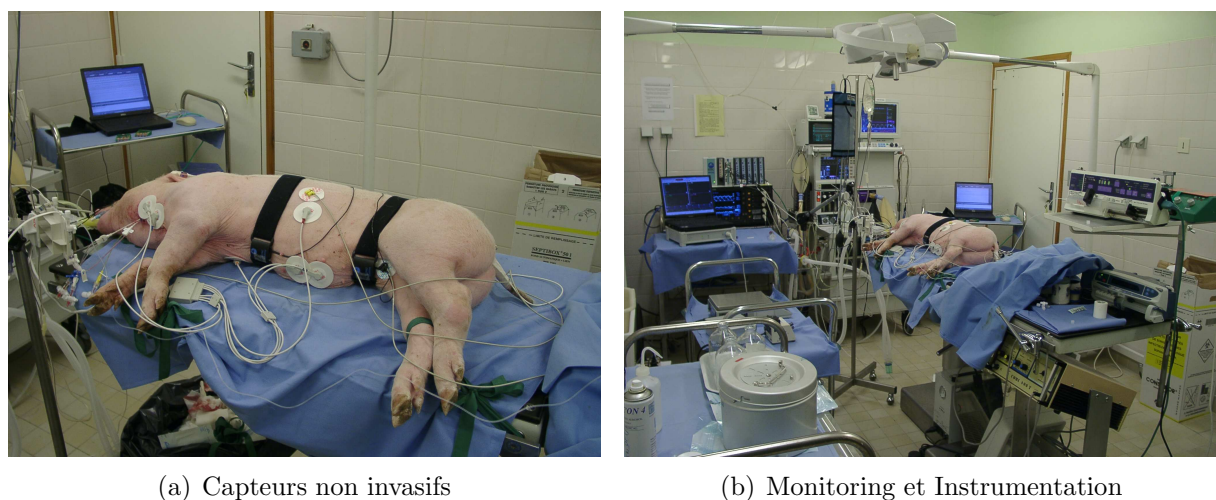
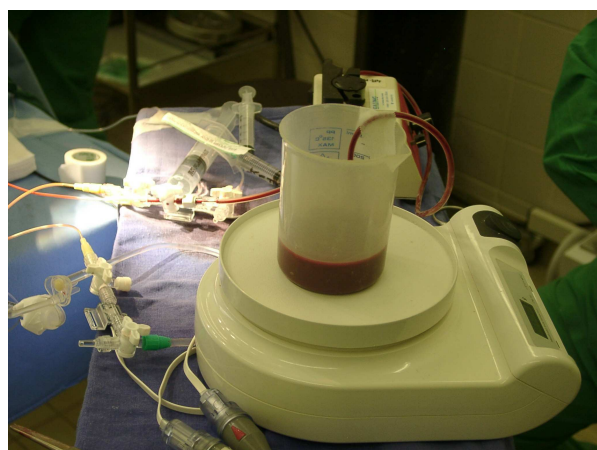
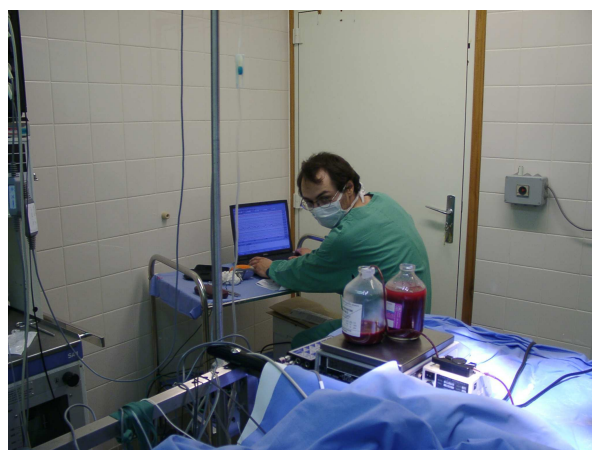


FIG. 4.5 – Stabilisation de l'animal : Après la mise en place des capteurs, l'animal récupère pour retrouver des constantes physiologiques dans un état stable.

11. Puis, la spoliation sanguine contrôlée (prélèvement de sang à la pompe par l'une des voies veineuses) est commencée à un débit de 5 à 7 ml/min de façon à provoquer une perte sanguine de 50% du volume sanguin estimé (sur la base de 69 ml/kg) en 90 min .
12. La spoliation sanguine est maintenue jusqu'au décès de l'animal qui survient en $1\text{ h }30\text{ min}$ à 3 h ($2\text{ h }08\text{ min} \pm 20\text{ min}$), à l'issue d'une phase plus ou moins prolongée de choc hémorragique, irréversible après une spoliation de 900 à 1300 ml , soit 70 à 90% du volume sanguin total.



(a) Balance



(b) Prise de mesures



(c) Contrôle

FIG. 4.6 – Choc hémorragique : a) mise en place de la balance et de la pompe à galet. b) et c) le débit est constamment contrôlé pour atteindre la cible de 50% du volume sanguin total en 1 h30 min. Les poids et instants sont relevés à intervalles réguliers.

13. Le décès est confirmé par injection de *pentobarbital* (*Doléthal*®) selon une procédure d'euthanasie vétérinaire usuelle par injection et surdose de barbiturique, dès que des troubles respiratoires importants sont observés.

Choc Hémorragique

De façon à obtenir une spoliation continue à débit contrôlé correspondant à une perte de 50% du volume sanguin total en 1 h 30 min, il est nécessaire d'ajuster le débit de la spoliation au poids des animaux qui diffère d'un individu à l'autre. Les réglages suivants ont été appliqués (Tab. 4.1) et correspondent à des calculs de débits volumiques rapportés à des débits massiques. La bonne pratique de ces consignes a été constamment observée au cours de l'expérimentation et le débit de la pompe péristaltique a été ajusté lorsque des dérives ont été constatées.

Nous avons considéré un volume sanguin pour l'animal de 69 ml/kg ($\approx 1/14.5$) par rapprochement avec les valeurs généralement acceptées chez l'homme et en nous référé-

rençant aux livres de biologie sur les animaux de laboratoires [55]¹. D'une façon plus rigoureuse, l'âge de l'animal devrait rentrer en compte pour ce calcul, en particulier chez le porc dont la masse grasseuse atteint une large proportion chez certains individus (ce que nous n'avons pas observé chez nos jeunes individus sains).

TAB. 4.1 – Réglage du débit de spoliation

Poids (kg)	ΔV_1 (ml)	ΔT_{25} (MM : SS)	Poids (kg)	ΔV_1 (ml)	ΔT_{25} (MM : SS)
15	5.75	04 :21	23	8.81	02 :50
16	6.13	04 :05	24	9.20	02 :43
17	6.51	03 :50	25	9.58	02 :37
18	6.90	03 :38	26	9.96	02 :31
19	7.28	03 :26	27	10.34	02 :25
20	7.66	03 :16	28	10.73	02 :20
21	8.05	03 :06	29	11.11	02 :15
22	8.43	02 :58	30	11.49	02 :11

ΔV_1 : Volume à prélever en une minute

ΔT_{25} : Durée pour prélever 25 g (≈ 25 ml) de sang

4.2 Mesures réalisées

4.2.1 Quelques définitions

Notions de Capteurs

Premier étage de notre chaîne de traitement de l'information, le capteur réalise une véritable perception physique de l'environnement et peut être considéré comme un organe sensoriel [56].

Pour chaque mesure un transducteur est utilisé qui traduit un phénomène physique difficilement observable en une autre grandeur physiquement plus appréhensible. Depuis plus d'une centaine d'année, l'électricité a joué un rôle fondamental dans ce domaine, et depuis l'apparition des ordinateurs et de leur fort pouvoir intégratif, la transduction en signal électrique est quasiment systématique. Cette grandeur physique est dite analogique car théoriquement analogue à la grandeur physique primaire. Elle véhicule un signal par variation d'une grandeur continue. L'enregistrement de celle-ci sur un ordinateur se fait après passage du signal dans un convertisseur analogique-numérique. Le signal subit alors une quantification en amplitude et un échantillonnage en fréquence (une discrétisation du signal à la fois en amplitude et en temps) afin que la machine puisse stocker les données sous forme binaire. On obtient un signal brut (numérique échantillonné, stockage en bits : 1 bit = 1 élément binaire, abbr. angl. *binary digit*). Les opérations d'échantillonnage et de quantification peuvent générer des bruits et des distortions du signal si elles sont mal réalisées. Le bon conditionnement de ces signaux est en partie l'objet du traitement du signal [57, 58, 2, 59].

L'expert médical a l'habitude de travailler à une échelle de temps donnée, son œil est alors habitué à la reconnaissance de certaines caractéristiques sur cette représentation

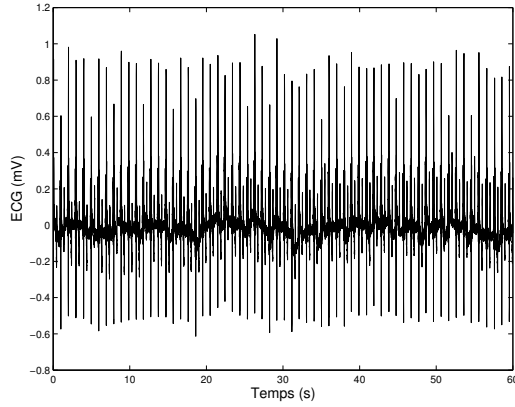
¹Les valeurs référencées par Glass et coll. semblent mieux adaptées au modèle du porc

graphique. L'expertise dépend alors de la fenêtre d'observation utilisée. L'utilisation de l'informatique permet de mettre en évidence l'influence de l'échelle de visualisation sur le contenu des données stockées et permet de préciser les limites du traitement du signal. Un exemple d'un signal électrocardiographique discrétisé est représenté Fig. 4.7 qui met en évidence les variations de la ligne de base (composantes basse fréquence) sur des fenêtres de longues durées (60 s), les particularités liées à chaque battement cardiaque pour des fenêtres plus courtes (2 s), ou les limites d'interprétation liées à l'échantillonnage des données (fenêtres de 100 ms pour une période d'échantillonnage de 2 ms).

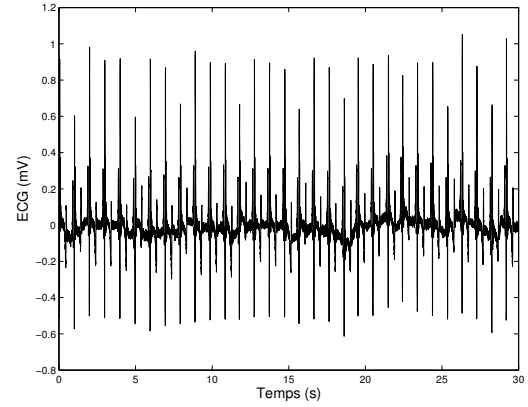
Étalonnage d'une chaîne de mesure

Pour établir la relation entre le signal réel et le signal quantifié on effectue l'étalonnage (angl. *calibration*) de la chaîne de mesure qui permet d'établir la correspondance entre la grandeur physique et la valeur traitée par la machine. Cette opération n'est pas négligeable et requiert parfois de longues heures (suivant la loi qui existe entre la grandeur physique et la grandeur artificielle), en particulier lors du suivi de la dérive des appareils électroniques pour la mesure de la température. L'étalonnage permet en outre de fixer les intervalles de confiance de mesure car elle donne la précision de la mesure.

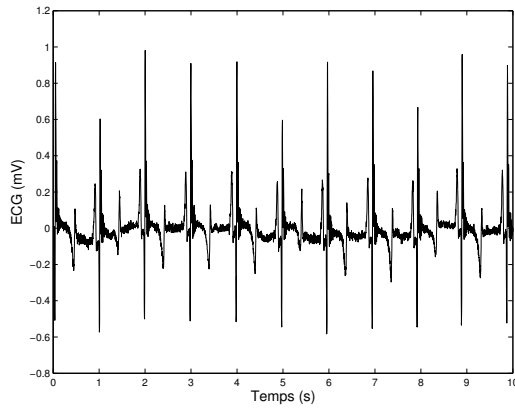
La calibration est un étalonnage simplifié qui consiste à passer un objet de référence de façon à obtenir l'unité de la grandeur désirée. Cette opération est généralement effectuée lors du passage de la sortie du transducteur, généralement en volt, vers l'appareil d'acquisition pour un stockage dans l'unité désirée.



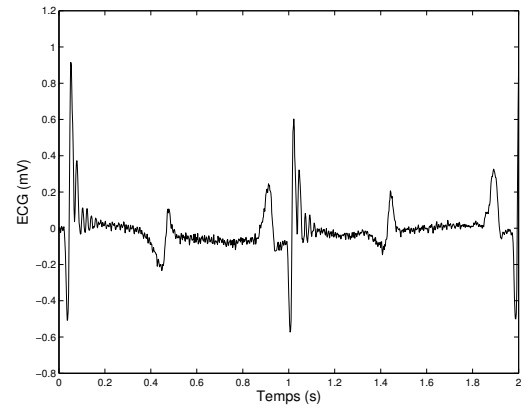
(a) 60 s



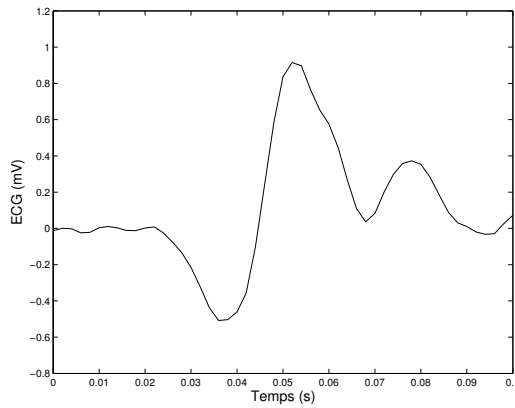
(b) 30 s



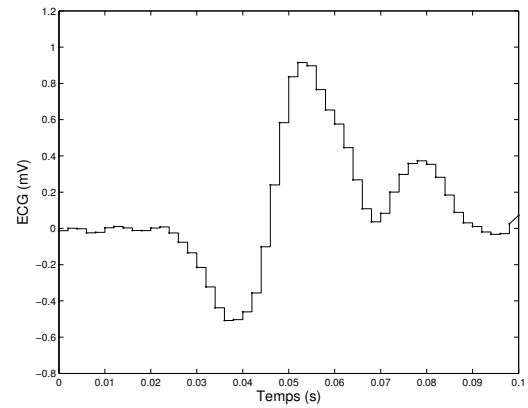
(c) 10 s



(d) 2 s



(e) 0.1 s



(f) 0.1 s

FIG. 4.7 – Un signal brut : un électrocardiogramme (ECG) échantillonné à une fréquence de 500 Hz (période d'échantillonnage de 2 ms) sur 16 bits ($2^{16} = 65536$ niveaux de quantifications, c'est à dire des valeurs codées entre -32768 et $+32767$). Une opération d'étalonnage est nécessaire pour retrouver l'analogie avec la grandeur physique. On notera l'influence de l'échelle de temps sur la représentation du signal. L'échantillonnage apparaît pour des fenêtres de faibles durées : e) le signal est tracé avec une interpolation linéaire, ce qui peut induire en erreur sur la réalité de l'information disponible ; f) les valeurs stockées sont représentés par des points, le signal est représenté graphiquement par ces valeurs bloquées.

Signaux, séries temporelles

Chaque grandeur mesurée est transformée en un signal numérisé en échantillon discret. Si X est une variable associée à la grandeur mesurée et $x(t)$ la réalisation à l'instant t de cette variable, la discrétisation du signal conduit à $x(k) = x(t_k)$, $t_k = t_0 + kT_e$ avec T_e période d'échantillonnage et t_0 instant initial sur la durée d'observation (instant de référence). Les valeurs discrétisées du signal continu sur la période d'observation, sont alors considérées comme une suite ordonnée de valeurs numériques.

On parlera de série temporelles pour qualifier un ensemble de données séquentielles issues de l'observation ou encore, pour désigner un ensemble de données relevées à des instants successifs.

Bruit

On désignera par bruit, tout ce qui gêne à la détermination de la grandeur observée et véhiculé par le signal. Ainsi, un signal pourra être bruité par un phénomène n'entrant pas dans la modélisation de la grandeur observée.

Artefacts Les bruits générés par l'influence de l'homme sur la nature par son action propre ou par le biais d'appareils réalisés par l'homme, sont généralement désignés sous le terme d'artefacts. Ils peuvent être produits par un mouvement (p. ex. le capteur de SpO_2 est très sensible à ce type de perturbations), un déplacement d'un capteur non détectable par l'instrument de mesure (p. ex. le décollement d'une électrode ECG) ou qui n'a pas été conçu pour minimiser ces déplacements (p. ex. la variation de la pression ambiante sur un appareil de mesure de pression artérielle), l'influence d'un appareil extérieur au capteur qui viendra perturber le signal mesuré (50 Hz de la distribution électrique en France capté par les électrodes)... Le résultat de ce type de bruits se traduit souvent par des sauts de la ligne de base du signal, l'introduction de composantes périodiques à des fréquences impropres au signal mesuré, dans le pire des cas à une saturation du signal par addition de bruits conduisant à un dépassement des valeurs du domaine de définition sur laquelle le signal est mesuré.

Outliers Les outliers sont des valeurs qui vont faire sortir la grandeur mesurée de son domaine de définition, elles sont hors normes.

Chronologie

Le temps prend une place importante dans l'étude des phénomènes biologiques. La chronobiologie² en étudie ces implications.

Périodicité, rythmes, fréquences En biologie, on aura tendance à parler de rythmes pour qualifier la répétition des phénomènes. En traitement du signal, on parlera plus couramment de fréquences et de périodes, ces deux notions étant inverses l'une de l'autre³.

²Société Francophone de Chronobiologie : <http://www.univ-st-etienne.fr:1042/sfc/>

³À une période T , on associera une fréquence F , son inverse : $F = 1/T$

Ordre de grandeurs de variations chronobiologiques La plupart des entités biologiques ont une tendance innée à se comporter selon un rythme spécifique dans des conditions déterminées. Les rythmes souvent observés, généralement liés aux conditions journalières de l'environnement ou liés à des processus métaboliques, sont les suivants :

- Rythmes nycthéméraux Rythme de $24\ h$
- Rythmes circadiens Rythmes d'environ $24\ h$
- Rythmes circasemiadiens Rythmes d'environ $12\ h$
- Rythmes ultradiens Rythmes $> 24\ h$
- Rythmes infradiens Rythmes $< 24\ h$

Les rythmes les plus fréquents sont observés sur des courtes périodes de temps : les rythmes cardiaques et respiratoires varient d'une espèce à l'autre (plus l'animal est petit, plus les rythmes sont rapides). Chez l'homme on observe un rythme qui varie de 40 à 300 battements par minute (bpm) et un rythme respiratoire qui varie de 0 à 60 cycles par minute (cpm). Une arythmie sinusale d'origine respiratoire est généralement observée sur des individus sains et met en évidence l'influence d'un phénomène rythmé sur un autre. Des rythmes plus courts δ ($0.5 - 4.5\ Hz$), θ ($4.5 - 8.5\ Hz$), α ($8.5 - 11.5\ Hz$), σ ($11.5 - 15.5\ Hz$), β ($11.5 - 22\ Hz$), γ ($22 - 45\ Hz$) sont encore observés sur les électroencéphalogrammes et sont caractéristiques de l'activité du cerveau.

Librairie de données

Quelques conseils pour la création d'enregistrements de données médicales sont disponible dans la littérature [60, 61] ou sur internet (<http://www.physionet.org/> et <http://ecg.mit.edu/>). Une attention particulière est portée sur la mise en place d'un format standard de fichiers pour chaque individu enregistré, de façon à pouvoir échanger plus facilement les données entre les équipes multidisciplinaires qui pourraient avoir à travailler ensemble. Par exemple, le format EDF (*European Data Format*) a été créé pour le projet Siesta de l'équipe de Bob Kemp [62], dans cette optique. D'un formalisme simple, il met en évidence les avantages et les inconvénients liés aux enregistrements et aux stockage des données brutes.

Les progrès réalisés en matière de stockage de données par les langages structurés tels que le XML semblent particulièrement adaptés pour la constitution de systèmes de gestion de base de données médicales (*Patient Data Management System*, PDMS). Néanmoins, nous n'avons pu utiliser ces techniques qui nécessitaient encore trop de développement.

4.2.2 Données brutes

Les données ont été enregistrés sur un ordinateur portable ($1.2\ GHz$, $528\ Mo$, $28\ Go$ de disque dur) par l'intermédiaire du logiciel AcqKnowledge®v.3.7.2, Biopac Inc. La fréquence de base a été fixée à $500\ Hz$. Les enregistrements ont été débutés et arrêtés manuellement. L'ensemble des 14 signaux enregistrés peut être consulté Tab. 4.2.

Six modules de transduction et d'amplification des signaux (1 module oxymétrique OXY100C, 2 modules de respiration RSP100C et 3 modules de températures SKT100C), ont été branchés à une unité d'acquisition MP150, pour la mesure des signaux de SpO_2 (capteur TSD123B), T_{int} (capteur TSD202F), T_{cut} (capteur TSD202B), T_{amb} (capteur TSD202A), Abd (capteur TSD201) et Tho (capteur TSD201). Les mesures de pression artérielle, pression veineuse centrale, et électrocardiographiques ainsi qu'une mesure de pouls périphérique

TAB. 4.2 – Mesures

Voie	Abr.	Nom	Position	F_e	Unité	Prétrait.
1	<i>ECG</i>	Électrocardiogramme	Axe cardiaque	500 Hz	mV	-
2	<i>PoulsK</i>	Onde de pouls (oxymètre Kontron)	Groin	500 Hz	V	-
3	<i>PA</i>	Pression artérielle	Cathéter carotidien	500 Hz	mmHg	-
4	<i>SpO₂</i>	Saturation pulsée en oxygène (oxymètre Biopac)	Oreille scutularis interne	15.625 Hz	%	-
5	<i>Tho</i>	Effort respiratoire (pneumogramme) (jauge de contrainte)	Circonférence thoracique	125 Hz	%	LP 0.5 – 1 Hz
6	<i>Abd</i>	Effort respiratoire (pneumogramme) (jauge de contrainte)	Circonférence abdominale	125 Hz	%	LP 0.5 – 1 Hz
7	<i>PVC</i>	Pression veineuse centrale	Cathéter jugulaire externe	500 Hz	mmHg	-
8	<i>PoulsB</i>	Onde de pouls (oxymètre Biopac)	Oreille jonction tête	500 Hz	%	-
9	<i>T_{cut}</i>	Température cutanée	Intérieur oreille	15.625 Hz	°C	-
10	<i>T_{int}</i>	Température interne	Rectum (5 à 10 cm)	15.625 Hz	°C	-
11	<i>CO₂</i>	Dioxyde de carbone	Capnigramme prélèvement sonde trachéale	125 Hz	mmHg	-
14	<i>T_{amb}</i>	Température ambiante	Oreille scutularis externe	15.625 Hz	°C	-

par un capteur oxymétrique ont été centralisées par le biais d'un moniteur Kontron Instruments Ltd. Un câble adapté réalisé par le service technique nous a permis de relier la sortie de cet appareil au système d'acquisition par l'intermédiaire de fiches Jack ($\phi 3.5\text{ mm}$ mono mâles) au module MP150.

Les différents temps opératoires ont été saisis manuellement par l'intermédiaire d'un journal synchronisé avec le logiciel d'acquisition. Le poids du sang prélevé sur les différents animaux au cours de la spoliation a été mesuré sur une balance électronique, approximativement tous les 25 g et les instants de mesure correspondant saisis informatiquement. Le débit de la perte de sang a été régulé de façon à obtenir la valeur cible déterminé précédemment.

L'analyse des données a été réalisée avec le logiciel Matlab v. 6.5 (Mathwork Inc.). Les boîtes d'outils *neural network* v. 4.0.2, *signal processing* v. 6.0 et *statistics toolbox* v. 4.0 ont été utilisées. Des algorithmes spécifiques ont été développés de façon à importer les données au format AcqKnowledge (*.acq) sous Matlab⁴. Les données ont été partagées avec l'équipe de Marie Christine Chambrin (Lille) et peuvent être récupérées à l'adresse électronique : <http://frodon.univ-lille2.fr/naineforge/frodon/public/> après une demande d'adhésion supervisée par Chambrin et Becq.

⁴Nous avons encadrés des étudiants pour réaliser un convertisseur afin de transformer les données au format European Data Format (*.edf), plus généralement connu par la communauté scientifique.

Les algorithmes pour la calibration, la synchronisation, le traitement des données, la classification, l'analyse des résultats et la visualisation ont été développés sous Matlab.

Un exemple de l'évolution des différents signaux enregistrés est présenté Fig. 4.8. Cet exemple met en évidence certaines particularités de ces signaux biologiques : changement de rythmes au cours de l'expérimentation, changement de formes des phénomènes cycliques, apparition et disparition de bruits et artefacts...

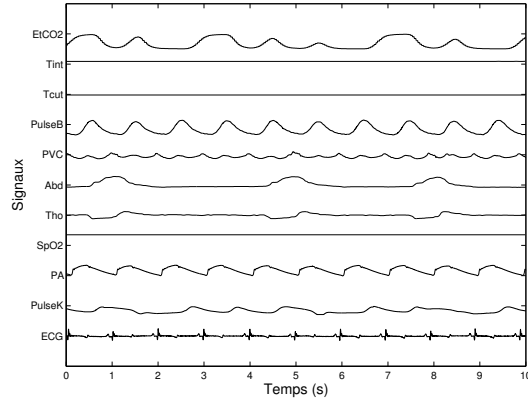
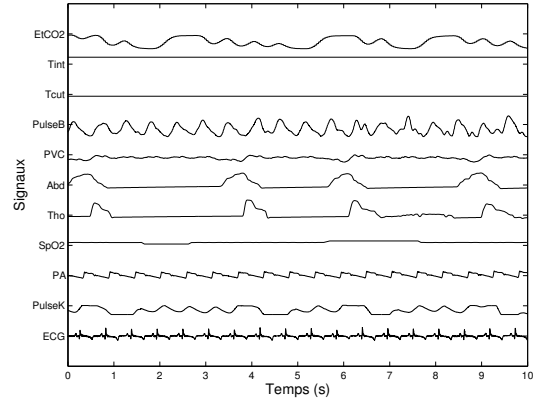
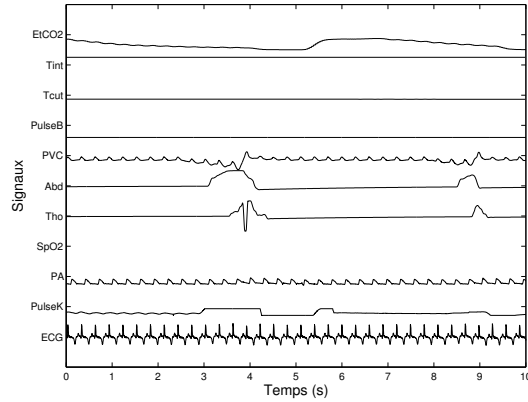
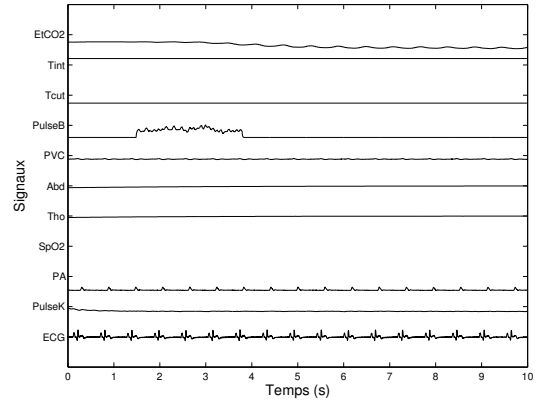
(a) $t = -25 \text{ min}$ (b) $t = 30 \text{ min}$ (c) $t = 1 \text{ h } 30 \text{ min}$ (d) $t = 2 \text{ h}$

FIG. 4.8 – Évolution des signaux au cours de l'expérimentation sur des fenêtres de 10 s par rapport au temps $t_{choc} = 0$ correspondant au début de la spoliatio : Les formes des signaux changent au cours de l'expérimentation, néanmoins sur des fenêtres temporelles plus réduites, les propriétés statistiques et fréquentielles semblent être respectées. L'influence d'une variable sur l'autre n'est pas évidente et les bruits sont plus ou moins marqués au cours de l'expérimentation.

4.3 Extraction de caractéristiques

On devient cuisinier, mais on naît rôtiisseur.

Brillat Savarin, Physiologie du Goût, champs flammarion, 1982, Paris.

Dans cette section, nous présentons les différentes techniques qui nous ont permis d'extraire des signaux bruts de l'étude, les différentes caractéristiques utilisées pour la reconnaissance de formes. Sans être exhaustif, nous définirons et indiquerons les traitements couramment effectués en automatique et traitement du signal ou dans des études similaires à la nôtre. Nous développerons en particulier, les techniques d'extraction utilisée en analyse polysomnographique (étude des tracés de sommeil) dont nous nous sommes largement inspirés pour la création de notre système. Enfin, nous présenterons les techniques utilisées sur les grandeurs biologiques couramment utilisées en anesthésie-réanimation pour évaluer l'état de santé d'un patient.

4.3.1 Quelques définitions

Caractéristiques

Les caractéristiques sont des variables extraites des signaux, supposées contenir une information pertinente pour décrire l'état du système. Etymologiquement, le terme de caractéristique fait référence à une variable associée aux signes gravés dans un signal. En apprentissage artificiel, on préfère l'emploi des termes d'attributs, de paramètres ou encore de descripteurs de l'espace de représentation plutôt que celui de caractères ou caractéristiques (angl. *feature*).

Découpages temporels

Afin de pouvoir extraire une caractéristique du signal observé, nous devons définir une fenêtre temporelle. L'hypothèse de stationnarité doit être vérifiée pour pouvoir effectuer une estimation fiable de celle-ci. On distingue différentes façons de procéder suivant le choix de la fenêtre temporelle (angl. *windowing*). On notera Δt la durée d'une fenêtre (sa largeur temporelle) (angl. *window length*). Plusieurs techniques peuvent être employées selon une complexité algorithmique croissante :

- Fenêtrages fixes : la largeur de la fenêtre Δt est constante et de durée fixe. Le signal total est découpé en plusieurs fenêtres de même durée.
- Fenêtrages avec ou sans recouvrement (angl. *overlapping*) : selon l'instant initial constituant le début de la fenêtre d'étude, les fenêtres successives qui découpent le signal peuvent se recouvrir partiellement selon un rapport α dit de recouvrement. L'utilisation de recouvrement permet de réaliser une statistique plus fine au prix d'un calcul plus long.
- Fenêtrage adaptatif : aussi connu sous le terme de segmentation du signal : le signal est découpé selon différents critères. En général, le début et la fin de la fenêtre temporelle⁵ sont déterminés dès qu'un changement de comportements est dé-

⁵On peut utiliser ici des techniques de reconnaissance de formes pour détecter le changement de modes de fonctionnement du processus qui génère le signal

tecté [63, 64, 65]. Les nombreuses méthodes de segmentation par des méthodes d'analyse non linéaire ou stochastique sur des signaux *EEG* sont référencées dans [66, 67].

Parmi ces techniques nous avons retenu celle utilisant un fenêtrage fixe car elle présente l'avantage d'être simple et ne nécessite pas de travaux supplémentaires. L'effet d'un tel découpage sur nos données est représenté Fig. 4.9 : chaque fenêtre est répertoriée par un indice k correspondant son instant initial $t_k = t_0 + k\Delta t$.

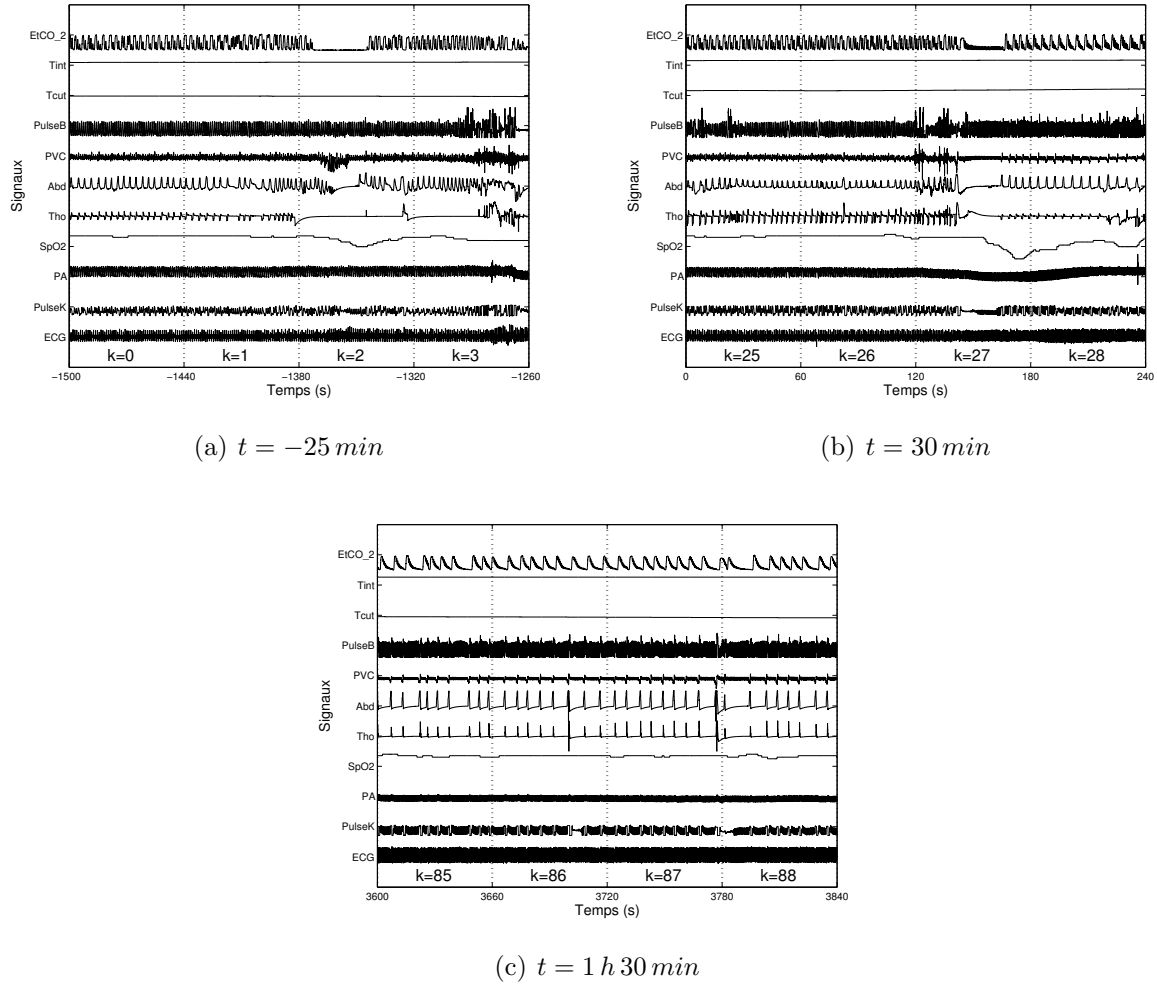


FIG. 4.9 – Segmentation des signaux au cours de l'expérimentation par des fenêtres de 60 s. Le calcul de caractéristiques propres à chaque signal se fera sur les fenêtres d'analyse repérées par l'index k et débutant aux instants $t_k = t_0 + k\Delta t$ en prenant $t_0 = t_{choc} - 25 \text{ min}$ avec $t_{choc} = 0 \text{ min}$ correspondant au début de la spoliation sanguine.

La difficulté d'une segmentation du signal par des fenêtres de durée fixe sur des signaux biologiques réside dans la différence des temps caractéristiques des signaux. Une recherche des temps caractéristiques adaptées à chaque grandeur biologique en unité de soins intensifs (*FC*, *RR*, *SpO₂*, etc.) à été réalisée par l'équipe de Marie-Christine Chambrin [68] et met en évidence la présence de différents temps caractéristiques pour chaque grandeur physiologique mesurée.

Un exemple de la difficulté du choix de cette fenêtre d'étude peut être apprécié par comparaison des Fig. 4.8 et 4.10. Par exemple, l'instabilité de la *SpO₂* sera mise en évidence sur une fenêtre de 60 s mais ne sera disponible qu'au terme de l'analyse de cette

fenêtre (c.-à-d. au bout de 60 s). Au contraire, sur une fenêtre plus courte, le résultat de l'analyse sur la fenêtre sera disponible plus rapidement (au terme de 10 s), mais l'épisode de désaturation ne sera plus interprété de la même façon : la différenciation entre un débranchement et une hypoxie réelle ne sera plus possible. Il existe alors un compromis à réaliser entre vitesse de détection et précision de l'interprétation.

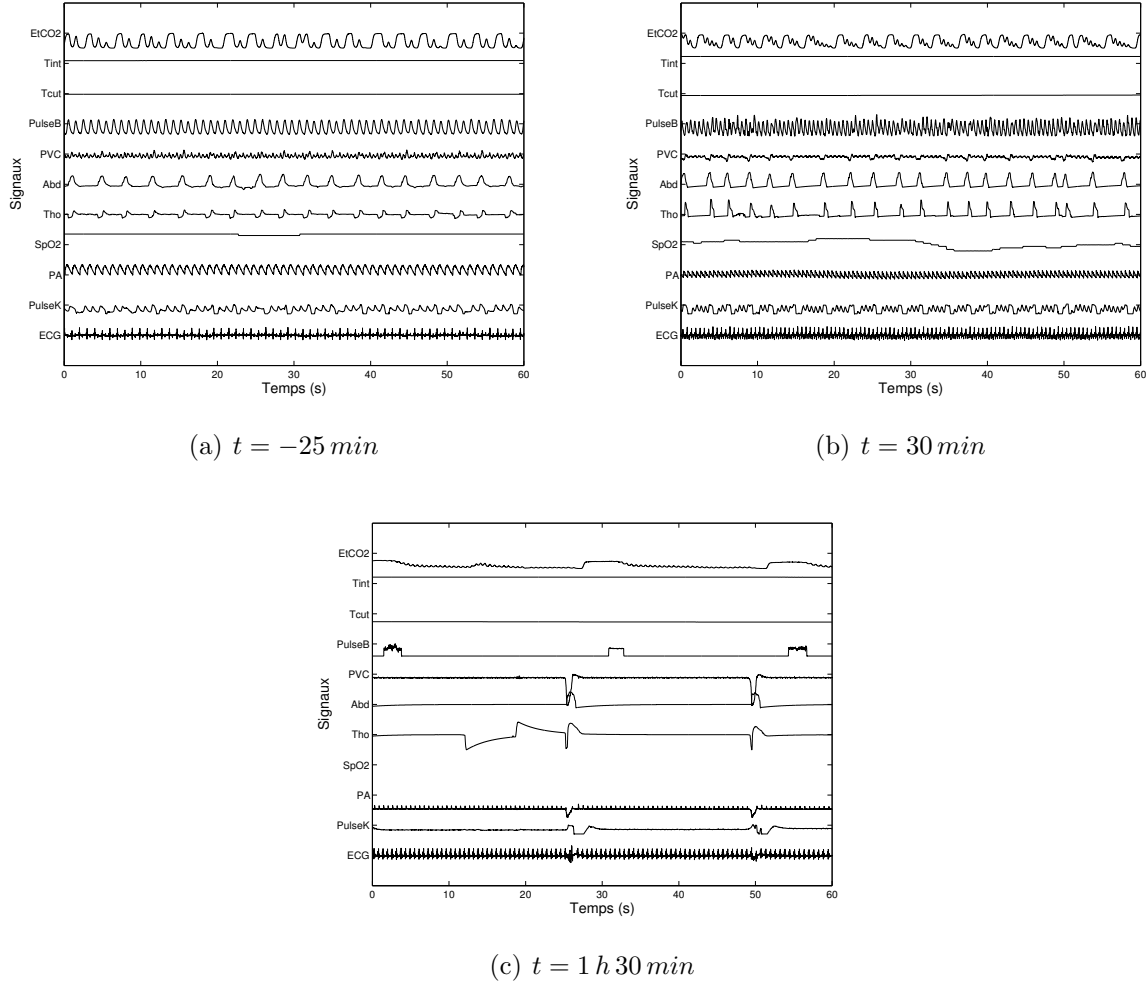


FIG. 4.10 – Évolution des signaux au cours de l'expérimentation sur des fenêtres de 60 s. Suivant la durée de la fenêtre utilisée, l'estimation des caractéristiques peut être plus ou moins précise. On peut comparer avec la Fig. 4.8

Prétraitement des données brutes

Pour éliminer les composantes qui constituent le bruit, suivant la connaissance sur la nature de celui-ci, sa périodicité, etc., différentes techniques sont généralement employées. Nous ne citerons pas toutes ces techniques et renvoyons le lecteur vers les ouvrages de traitement du signal [58, 57, 59].

La technique la plus couramment utilisée pour éliminer le bruit consiste à projeter le signal sur une base de filtres constituée de fonctions temporelles déterminées et paramétrées. Ces fonctions étant choisies de façon à être le plus orthogonal possible aux bruits présents sur le signal, et en ne retenant que les valeurs associées aux projections sur certaines bases, on réalise l'opération de filtrage qui fournit un signal filtré. Dans ces bases de filtres, les plus connues réalisent des filtrages fréquentiels en projetant le signal sur des bases de fonctions sinusoïdales (décomposition en séries de Fourier) ou en projetant sur des bases d'ondelettes selon un fenêtrage temporel variable (Δt variable, décomposition en ondelettes). C'est à dire en tenant compte des formes prises par la succession des différentes réalisations temporelles⁶ (en pratique, on passera d'une représentation temporelle à une représentation fréquentielle).

La technique la plus générale consiste à déterminer les caractéristiques statistiques (c'est-à-dire l'étude des distributions des différentes réalisations) des valeurs générées par les différentes réalisations sans tenir compte de l'aspect temporel sur la durée de la fenêtre d'étude (hypothèse d'ergodicité). Selon le nombre de réalisations obtenues sur le signal et suivant les propriétés statistiques de celui ci et des bruits parasites, on récupère une information correspondant aux caractéristiques statistiques du signal recherché. Elles reposent toutes sur une notion importante : la théorie de l'estimation, qui définit quelles sont les erreurs que nous pouvons commettre en effectuant ces calculs de paramètres en utilisant telle ou telle méthode.

Estimation de paramètres

L'estimation consiste à donner des valeurs approchées aux paramètres d'une population à l'aide d'un échantillon de n observations issues de cette population [69].

On estimera donc un paramètre en étudiant un ensemble de valeurs prises par une variable aléatoire. La valeur estimée pourra être biaisée par rapport à la valeur réelle de la caractéristique, le biais étant alors calculé comme la différence entre la valeur réelle et la valeur estimée. On essaiera de réaliser des estimations selon les modes propres recherchés et en particulier, des sondages stratifiés sur des populations inhomogènes pourront conduire à des résultats erronés. On cherchera donc à réaliser des estimations robustes, c'est à dire peu sensibles à la présence de points aberrants, artefacts ou autres outliers. Pour cela on pourra encore utiliser des techniques de rééchantillonnage, qui seront particulièrement utiles pour l'estimation des performances des modèles d'apprentissage que nous traiterons par la suite ou encore d'autres techniques d'estimations robustes (algorithme EM, angl. *Expectation Maximization algorithm*).

⁶On peut noter que si la forme de référence est le signal lui même, ou un autre signal de nature différente, on parle de corrélation et on utilise un corrélateur temporel

Extraction de caractéristiques dans le domaine du monitoring

Certaines méthodes consistent à faire une recherche systématique des paramètres statistiques sur les mesures réalisées. Nous citerons par exemple l'approche de Vyzas en psycho-physiologie [70, 71] très orientée statistique ou celle de Karazynski [72] qui utilise les différents moments statistiques des signaux, mais qui exclut systématiquement le contexte médical (les caractéristiques couramment utilisées en physiologie ne sont pas utilisées). On peut encore réaliser les recherches de moments statistiques fréquentiels sur les signaux électroencéphalographiques qui s'y prêtent bien [73, 74] en particulier pour la classification d'épilepsie [75].

Souvent, on cherchera à effectuer un calcul permettant de caractériser les tendances des signaux. Une revue des pratiques couramment employées pour les systèmes de monitoring est présentée par Avent [76] pour l'extraction de tendances et l'élimination des bruits (filtrage de Kalman, etc.).

Les méthodes sont nombreuses et auront pour but d'identifier les paramètres d'un modèle donné. Les principales caractéristiques recherchées seront donc de natures différentes : soit définies par une notion mathématique, soit adaptées au savoir-expert et à la recherche d'une information particulière souvent associée à la reconnaissance d'un motif temporel particulier. L'extraction de caractéristiques se fera donc souvent en association avec l'opération de filtrage définie précédemment (fréquences fondamentales, ondelettes, etc.).

De nouvelles méthodes pour la détection de motifs temporels particuliers font l'objet de recherches prometteuses et permettent en outre de segmenter le signal en temps réel. Néanmoins, une approximation par agrégation par morceaux (PAA, angl. *Piecewise Approximate Aggregation*), bien que simple de conception, semble être très compétitive par rapport à d'autres techniques plus complexes [77, 78, 79, 80]. Cette technique (PAA) ressemble beaucoup à celle que nous avons utilisée (calcul de caractéristiques sur un fenêtrage par morceaux) et c'est pour cela que nous la citons ici.

L'estimation de caractéristiques liées à des couplages entre différents signaux (caractéristiques multivariées) est encore possible : par exemple, le temps de transit du pouls correspond au décalage entre un battement cardiaque repéré sur un *ECG* et une onde pulsée repérée sur un oxymètre de pouls [81], le déphasage de la respiration thoraco-abdominale peut mettre en évidence certaines apnées du sommeil [82], enfin une fonction de cohérence entre signaux peut mettre en évidence un trouble épileptique [75]. Ces estimations, entre autres, complexifient encore les calculs.

Analyse Polysomnographique

Les recherches effectuées dans ce domaine depuis une trentaine d'années se sont orientées vers des méthodes d'extraction de caractéristiques et de reconnaissance de formes. Nous avons utilisé celles-ci dans une étude préalable sur des données polysomnographiques disponibles au laboratoire (cf. Annexe A). Cette étude nous a servi de base pour la mise en place du système utilisé dans la présente étude.

L'analyse polysomnographique s'inscrit dans un cadre plus général de l'analyse de l'électroencéphalogramme (*EEG*). Le terme *époque* (*epoch*, en anglais) est généralement utilisé pour désigner la fenêtre temporelle correspondant à la fenêtre d'étude ou de re-

cherches de paramètres⁷.

Sur ces *époques* de signal, le praticien recherche les signes et critères pour caractériser la phase de sommeil dans laquelle est plongé le sujet d'étude. Ainsi, dans un premier temps, l'expert étudie la forme, l'amplitude et la répartition de certains événements apparaissant ou disparaissant au cours du sommeil. Il synthétisera ensuite ces informations en attribuant à chaque époque un score. Le consensus établi par la conférence de praticiens spécialistes du sommeil et utilisé de nos jours en clinique et au laboratoire, disponible dans le manuel dit de Rechtschaffen and Kales (R&K) [83] pour diagnostiquer les pathologies associées aux troubles du sommeil, constitue en réalité, une base de filtres et prétraitements sur les signaux polysomnographiques pour l'étude du sommeil. Cependant, implémenter les règles expertes de celui-ci dans un automate, se révèle difficile en terme mathématique [67].

4.3.2 Mesures courantes et caractéristiques retenues

Les caractéristiques que nous avons retenues sont généralement disponibles en clinique dans les services d'anesthésie réanimation via les différents appareils de monitoring ou les appareils d'enregistrements [52]. Néanmoins, les algorithmes de calcul de ceux-ci restent la propriété du constructeur. L'utilisation de nombreuses techniques de filtrages ne permet pas de comprendre le comportement des machines et les mesures données sont souvent incohérentes avec l'observation du signal brut pour certains modes de fonctionnement atypiques. Les progrès actuels tendent à minimiser les fausses valeurs calculées par les machines souvent génératrices de fausses alarmes. Les études polysomnographiques utilisent aussi un grand nombre de ces mesures et développent des algorithmes de détection de signes pathognomiques sur différents matériels de mesures, en particulier pour le suivi des syndromes d'apnées du sommeil. Une description et une comparaison des capteurs utilisés dans ce domaine (qui sont pour la plupart proches de ceux utilisables pour notre domaine) et de leur apport clinique est donné dans le task force [82].

Électrocardiogramme

Les nombreux manuels médicaux d'électrocardiographie et de physiologie générale sont les premières sources pour une aide précieuse pour la pose des électrodes et la recherche des particularités et trouble de l'*ECG* [84, 85, 86]. Les caractéristiques les plus fréquemment détectées sont :

- Complexes de l'ECG : ondes P-QRS-T (Fig. 4.11). L'analyse de la distribution de ces complexes permet de mettre en évidence une pathologie associée au cœur, à une souffrance cardiaque. On associe souvent l'étude de ces complexes à l'étude de la variabilité cardiaque [87].
- Fréquence cardiaque : FC qui correspond au nombre de battements du cœur par seconde. C'est aussi l'inverse de la durée entre deux battements généralement repérés par les complexes R.
- Intervalle RR : Durée entre 2 complexes R. Il correspond à l'inverse de la fréquence cardiaque instantanée.

⁷Cette fenêtre dont la durée est habituellement fixée à 20 s, 30 s ou 60 s, constitue une véritable *époque* au sens temporel si l'*EEG* est enregistré avec une période d'échantillonnage usuelle de 2 ms ou 4 ms (500 Hz ou 250 Hz)

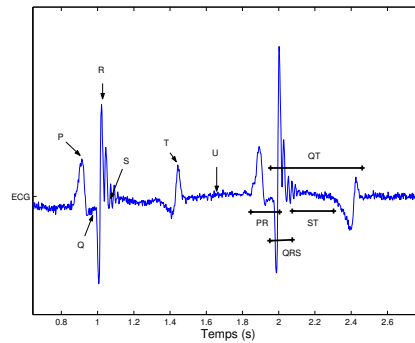


FIG. 4.11 – Les différents complexes (déflexions) d’un électrocardiogramme ainsi que les intervalles couramment utilisés [88].

- Variabilité cardiaque : l’étude des variations du rythme cardiaque permet de mettre en évidence certains signes d’insuffisances cardiaques [89] et semble être le reflet de la balance sympatho-vagale au niveau du système nerveux central [90, 91, 92].
- Énergie : amplitude du QRS [49].

Un filtrage passe bande permet d’éliminer les bruits parasites [93] de l’*ECG* :

- Interférences de la source d’alimentation électrique (p. ex. en France 50 Hz , aux États-Unis 60 Hz).
- Bruit de contact des électrodes.
- Artefacts de mouvements.
- Saut de la ligne de base et amplitude de l’*ECG* modulés par la respiration.
- Bruit d’instrumentation généré par les appareils électroniques utilisés pour le traitement du signal.
- Bruit électrochirurgicaux (bistouri électrique, défibrillateur etc.).

Caractéristiques retenues Le filtrage de l’électrocardiogramme s’obtient facilement en soustrayant la ligne de base [94] puis en utilisant un filtrage de type passe-bande [93] pour exclure les bruits parasites. Dans un premier temps on peut penser à utiliser une autocorrélation pour détecter la fréquence principale (ou par dualité le spectre du signal) [95], mais des problèmes surviennent au cours de l’expérimentation, d’une part par augmentation du rythme cardiaque, d’autre part par augmentation de l’amplitude des différents complexes, ce qui se traduit par une détection de rythmes secondaires associés aux complexes T, confondus avec les complexes R, et a pour effet de multiplier la fréquence réelle par deux. Le corrélogramme représenté Fig 4.12 permet de mettre en évidence l’évolution des différents rythmes au cours du choc hémorragique et la difficulté de détection de ceux-ci par cette méthode. Il correspond à une représentation sous forme d’image de la succession des différentes valeurs obtenues par les fonctions d’autocorrélation calculées aux instants $t_k = t_0 + k\Delta t$. La fonction d’autocorrélation étant définie par $C_{xx}(t_k, \tau) = \int_{t_k}^{t_k + \Delta t} x(t)x(t - \tau)dt$ avec τ paramètre de translation. La valeur de la fonction de corrélation traduit la similitude du signal avec lui-même pris aux instants $t - \tau$. Une valeur forte positive (représentée en rouge sur la figure) indique une corrélation du signal sur lui-même et ses valeurs traduites de τ . Une valeur négative (représentée en bleu sur la figure) indique une forte anticorrélation.

D’autres techniques consistent à utiliser des complexes QRS types pour identification

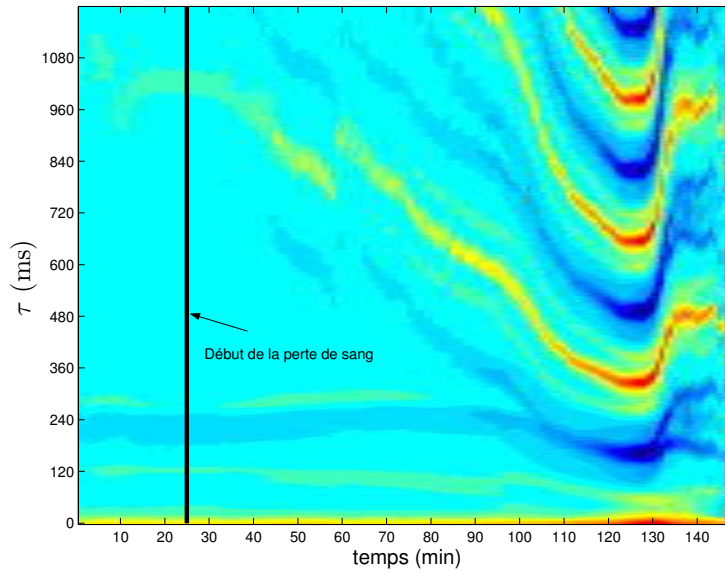


FIG. 4.12 – Corrélogramme de l'ECG de l'animal *Pig020409* : l'instant t_k définissant le début de la fenêtre du calcul de la fonction d'autocorrélation sur une largeur Δt de 60 s est en abscisse, et le paramètre de translation τ en ordonnée. La couleur correspond à la valeur de la fonction de corrélation : une couleur bleue foncée correspond à une valeur négative, une couleur bleue ciel correspond à une valeur nulle, une couleur verte et jaune à une valeur faiblement positive, enfin une couleur rouge correspond à une forte valeur positive. Les rythmes évoluent au cours de l'expérimentation. La forme jaune puis rouge est engendrée par une corrélation positive des complexes QRS de l'ecg aux paramètres de translation correspondant aux intervalles RR de l'ecg (l'inverse de l'intervalle RR correspond à la fréquence cardiaque FC). On constate une diminution de cet intervalle au cours de l'expérimentation jusqu'à l'obtention d'un minimum suivi d'une augmentation brusque et d'une arythmie finale. Le passage du jaune au rouge indique une augmentation de l'amplitude des complexes QRS ; la forme bleue foncée correspond à une anticorrélation entre le complexe QRS et le complexe T (ce complexe s'amplifie au cours de l'expérimentation). Enfin, la netteté des formes s'accroît et s'explique par une diminution de variabilité du rythme d'apparition des complexes QRS au cours de l'expérimentation. Ces tableaux pourraient être de bons indicateurs pour le suivi de l'évolution d'un blessé.

(nécessitant une reconnaissance et un apprentissage sur chaque nouvel individu, ou en utilisant un échantillon stocké dans une base déjà constituée) ou encore utiliser des ondelettes pour détecter les complexes QRS. De nombreux travaux montrent l'intérêt et les difficultés techniques de telles méthodes [96] en particulier au niveau du choix entre précision et filtrage (compromis temps-fréquence). Nous n'avons pas appliqué ces méthodes qui demandent une étude à temps plein pour comparer l'intérêt d'une technique par rapport à une autre (et partir de détection réalisée manuellement pour servir de référence), mais sommes partis d'un algorithme adaptatif simple (inspiré de [93] et [97]) qui nous fournissait les meilleurs résultats par essai erreur (comparaison des méthodes de corrélation [98], périodogrammes et seuillages adaptatifs.).

Calcul de la fréquence cardiaque sur une fenêtre de Δt s d'ECG

Filtrage passe bande (10 Hz – 50 Hz) à réponse impulsionnelle finie (100 coefficients pour notre étude, fenêtrage de Hamming)

Redressement : $y^* = \text{abs}(y) = \text{sign}(y) * y$

Calcul de l'écart type σ de y^*

Recherche des valeurs de y^* supérieures à $3.29\sigma^8$

Recherche des couples de valeurs successives telles que la différence de temps entre celles-ci soit supérieure à 200 ms (ce qui implique une FC inférieure à 300 bpm)⁹

Le nombre de pics R (n_R) correspond au nombre de valeurs retenues.

La fréquence cardiaque FC est alors : $FC = \frac{n_R * 60}{\Delta t}$

Un exemple d'application est présenté Fig. 4.13.

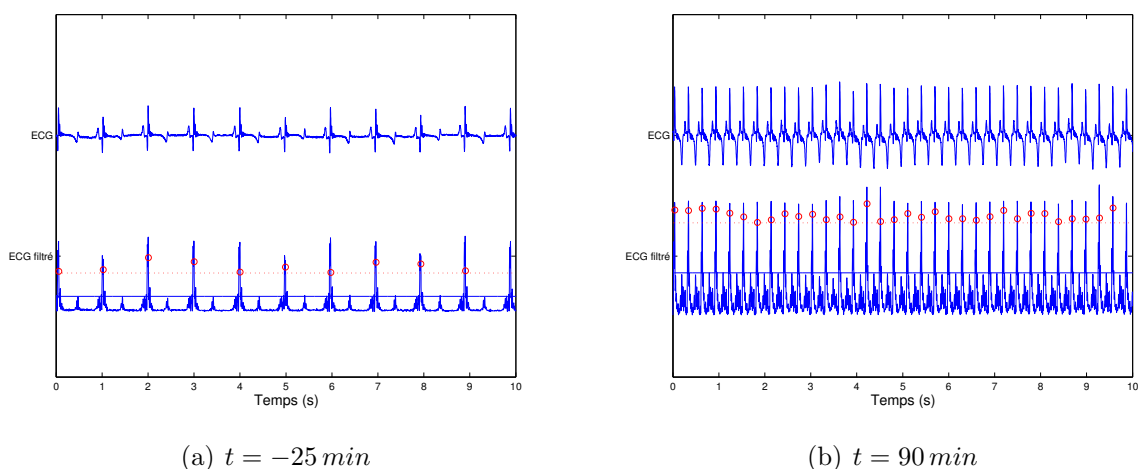


FIG. 4.13 – Détection de fréquence cardiaque : La ligne continue (bleue) correspond à l'écart-type du signal, la ligne en pointillé (rouge) correspond au seuil de détection (3.29 fois l'écart-type). Une période réfractaire de 200 ms est utilisée pour ne retenir qu'un seul point. Le signal est filtré, redressé et sa tendance est éliminée. a) Avant la perte de sang, le signal *ECG* a une fréquence cardiaque estimée de $66 \pm 6\text{ bpm}$. b) 90 min après le début de la spoliatio, l'amplitude de l'onde T a augmenté, ce qui implique l'adaptation du seuil pour une détection des pics R. la FC est maintenant de $204 \pm 6\text{ bpm}$. L'utilisation de l'écart-type relatif à la fenêtre considérée permet une adaptation du seuil pour la détection du pic R.

Ventilation

Les mesures de référence pour les signaux de ventilation sont obtenues par l'utilisation d'un pneumotachomètre qui permet de quantifier précisément la valeur des débits entrants et sortants dans le système respiratoire. Son utilisation nécessite l'utilisation d'un masque ou d'une sonde d'intubation. Il impose donc une contrainte de mesure. On peut lui associer des mesures de pression permettant d'obtenir un ensemble complet pour une étude précise de la physiologie respiratoire des systèmes vivants [99]. Lorsque l'on a accès à ces mesures contraignantes mais qui font références, un calcul de paramètres respiratoires peut être envisagé. Un exemple de calculs cycle à cycle est fourni en temps réel dans [100, 101].

⁸intervalle à 3.29σ , ce qui correspond à une probabilité de $1/1000$ de réalisation pour une loi de Laplace Gauss pour un *ECG* en absence de signal, i. e. constitué uniquement de bruit.

⁹l'utilisation d'une valeur dépendante de l'écart-type du signal pourrait être une recherche intéressante.

Les mesures non invasives par capteurs de pression nasale, par plethysmographie d'inductance respiratoire, par jauge de contrainte, capteurs piézo-électriques, impédance thoracique (souvent associée aux déviations de la ligne de base de l'*ECG*) ou encore par capteurs thermiques (souvent mesuré au niveau du flux nasal) permettent d'obtenir des informations sur la ventilation. Les deux premières mesures peuvent, après calibration, apporter une information quantitative sur l'amplitude respiratoire. Les dernières ne permettent d'obtenir qu'une information qualitative [82]. Toutes ces mesures permettent d'obtenir une estimation quantitative de la fréquence respiratoire.

Pour notre étude, nous avons utilisé des sangles thoraciques et abdominales aux propriétés élastiques (nylon) reliées à une jauge de contrainte.

Caractéristiques ventilatoires retenues Les sangles thoraciques et abdominales restent sensibles aux mouvements de l'animal (perte du signal après mouvements puis reprise du signal) et leur positionnement constitue une étape délicate. Les informations obtenues sur les deux capteurs utilisés sont généralement redondantes sauf lors des calculs de couplage ou lors de recherche de phase entre ces deux mesures mais nous n'avons pu évaluer ces paramètres.

La variation des rythmes du signal au cours de l'expérimentation peut être représentée par un corrélogramme, dont un exemple est fourni Fig. 4.14. Ce corrélogramme met en évidence les variations des rythmes de la respiration comme pour l'*ECG*. Son calcul est très simple et pourrait servir pour un suivi de l'évolution respiratoire du soldat.

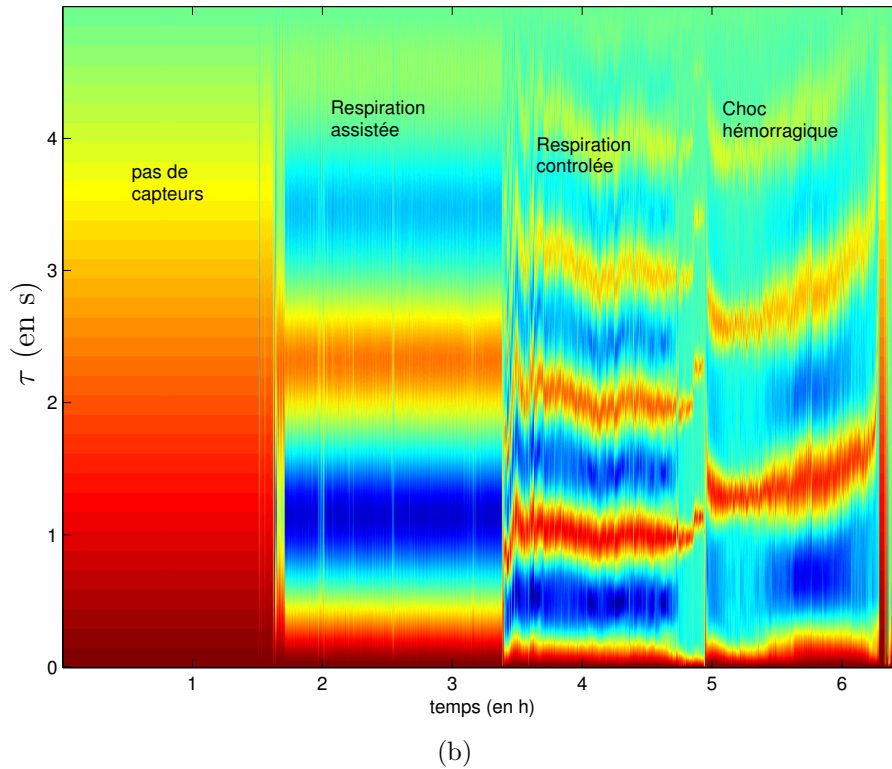
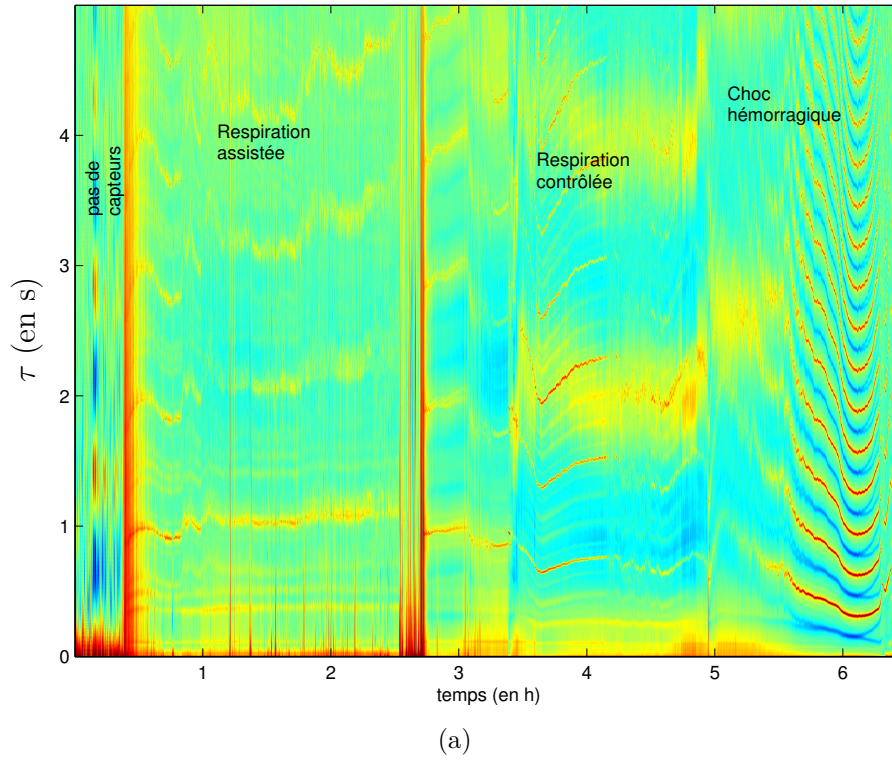


FIG. 4.14 – Corrélogrammes de a) l'électrocardiogramme et de b) la respiration ($\Delta t = 10s$) de l'animal *Pig020411* au cours de l'expérimentation depuis son entrée au bloc opératoire. Le rythme cardiaque peut être facilement évalué comme indiqué Fig. 4.12. On procède de la même façon pour évaluer les rythmes respiratoires : au cours du choc hémorragique on constate une augmentation de la durée des cycles respiratoires (la fonction de corrélation est maximale pour une translation de plus en plus grande). La modification de l'épaisseur du motif associé au rythme respiratoire indique qu'au cours de l'expérimentation l'animal passe d'une respiration synchrone (rétrécissement du motif), à une respiration de Cheyne Stokes (élargissement du motif), suivie de gaspes et d'un arrêt respiratoire final (disparition du motif).

Peu de références concernant le calcul de fréquence respiratoire sont disponibles, excepté la méthode de Eberhard et Baconnier qui fonctionne en temps réel sur des mesures pneumotachométriques dont nous nous sommes largement inspirés [101]. Nous avons donc développé l'algorithme ci dessous après raisonnement par essai erreur en essayant de garder une résolution acceptable sans passer trop de temps à développer de nouveaux algorithmes.

Calcul de la fréquence respiratoire sur une fenêtre de Δt s de respiration

Retrait de la ligne de base par retrait de la tendance correspondant à la droite de meilleur ajustement obtenue par régression linéaire

Filtrage passe bande ($0.1\text{ Hz} - 1\text{ Hz}$) à réponse impulsionnelle finie (100 coefficients pour notre étude, fenêtrage de Hamming)

Calcul de l'écart type σ du signal filtré y^*

Si on a $\sigma < 0.01$ on est en présence de bruit (seuil empirique suivant la calibration du capteur dont la valeur varie entre -1 et $+1$)

Sinon Rechercher les valeurs de y^* supérieures à σ

Rechercher les couples de valeurs successives telles que la différence de temps entre celles-ci soit supérieure à 100 ms

Le nombre de pics (n_R) correspond au nombre de valeurs retenues.

Le rythme respiratoire RR est alors : $RR = \frac{(n_R)*60}{\Delta t}$

Un exemple d'application est présenté Fig. 4.13.

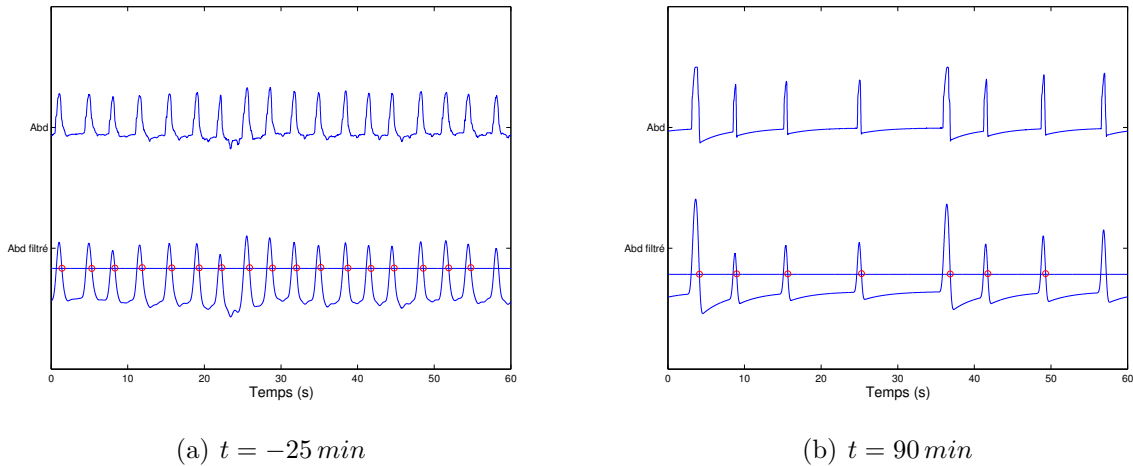


FIG. 4.15 – Détection de rythme respiratoire : La ligne continue correspond à l'écart-type du signal qui correspond au seuil de détection du pic. Une période réfractaire de 100 ms est utilisée pour ne retenir qu'un seul point. Le signal est filtré, redressé et la tendance est éliminée. a) Avant la perte de sang, le signal de respiration a un rythme respiratoire estimée de $17 \pm 1\text{ cpm}$. b) 90 min après le début de la spoliation, les respirations ont un rythme qui fluctue autour de $7 \pm 1\text{ cpm}$ (dyspnée, syndrome d'une respiration de type Cheyne-Stokes). Pour avoir une bonne estimation du rythme moyen, il faut utiliser un fenêtrage long adapté au rythme respiratoire (fenêtres de 60 s par exemple).

Le calcul d'une estimation du volume d'air ventilé dans les poumons a été réalisé par intégration de l'amplitude respiratoire. Nous avons désigné cette estimation par le terme

de pseudo-volume, puisque l'estimation de cette grandeur ne permet pas de déterminer le volume réel dans les poumons, les mesures n'étant pas étalonnées (et ne pouvant l'être). Les valeurs des caractéristiques $\tilde{V}_{t_{Tho}}$ et $\tilde{V}_{t_{Abd}}$ seront donc semi-quantitatives et les résultats obtenus sur ces grandeurs seront qualitatifs.

Calcul du pseudo-volume respiratoire sur une fenêtre de Δt s de respiration

On enlève la tendance (angl. *detrending*)

On repère les zones d'inspiration (positives(> 0.01)), et d'expiration (négatives(< -0.01)), les zones de pauses (comprises entre -0.01 et 0.01).

Les volumes inspirés V_{insp} et expirés V_{exp} correspondent à la somme des variations d'amplitude que multiplie la période d'échantillonnage. (A_{abd} et A_{tho} ont été fixés avec un gain de 10 et un filtre passe bande de 0.5 Hz – 10 Hz .)

Le pseudo-volume courant est obtenu en faisant la moyenne de ces volumes :

$$\tilde{V}t = \frac{1}{2}(V_{insp} + V_{exp}) \propto \frac{1}{2} \int_{insp,exp} \text{abs} \left(\frac{dA}{dt} \right) dt \quad (4.1)$$

Pression Sanglante

La pression artérielle de référence est obtenue par introduction d'un cathéter dans une voie artérielle. Ceci permet d'obtenir une mesure continue, après calibration du capteur de pression piézo électrique disposé sur le fluide interne au cathéter, vecteur de la pression mesurée à l'intérieur de l'organisme. Une mesure instantanée peut être obtenue par mesure auscultatoire (techniques des sons de Korotkoff) et nécessite l'utilisation d'un brassard gonflable adaptée à l'anatomie de l'individu observé. C'est la technique couramment utilisée par les médecins lors des consultations. Les techniques oscillométriques permettent un calcul automatique des pressions artérielles systoliques, diastoliques et moyennes par un appareil. Les mesures sont alors réalisées par un brassard au bras, au poignet, voire au doigt [10]. Si pour des valeurs de mesures normales, les données sont bien corrélées à une mesure de référence invasive, les valeurs obtenues pour des grandeurs pathologiques (hypotension et hypertension) sont erratiques pour les appareils actuels [102].

Pression Artérielle Nous n'avons pas réalisé de moyennes des différentes pressions systoliques, diastoliques et moyennes instantanées, après détection des cycles générés par battements du cœur, qui auraient été plus proches d'une définition médicale, mais nous avons choisi de réaliser ces estimations algorithmiquement plus simples :

$$PAS = \max_{t_k < t < t_k + \Delta t} PA(t) \quad (4.2)$$

$$PAM = \frac{1}{\Delta t} \int_{t_k}^{t_k + \Delta t} PA(t) dt \quad (4.3)$$

$$PAD = \min_{t_k < t < t_k + \Delta t} PA(t) \quad (4.4)$$

$$\Delta_{PA} = PAS - PAD \quad (4.5)$$

$$\sigma_{PA} = \sigma_{t_k < t < t_k + \Delta t} PA(t) \quad (4.6)$$

La *PAS* ainsi calculée correspond à la valeur maximale des différentes systoles présentes dans la fenêtre d'observation. De la même façon, la *PAD* correspond à la valeur minimale sur la fenêtre d'observation des différentes diastoles obtenues sur les différents battements. Nous avons observé qu'une arythmie sinusale d'origine respiratoire participait à la variabilité de ces caractéristiques, cependant, nous avons considéré que les caractéristiques ainsi estimées seraient suffisantes pour notre étude.

Pression Veineuse Centrale En accord avec l'expert médical, nous avons décidé d'inclure un rapport de puissances spectrales entre les basses fréquences ($0.01 \text{ Hz} \leq f < 0.7 \text{ Hz}$) sur les hautes fréquences ($0.7 \text{ Hz} \leq f < 10 \text{ Hz}$) pour le calcul de $LFHF_{PVC}$ obtenue sur la pression veineuse centrale. Nous pensions que ce rapport permettrait de mettre en évidence l'influence d'un couplage cardio-respiratoire par observation des fluctuations de la *PVC* au cours des expérimentations. En effet, au cours des expérimentations, nous avons constaté que des changements de rythmes étaient marqués sur la mesure de *PVC* et, de par la nature du site de mesure (cœur droit), correspondaient à la fois aux influences cardiaques et respiratoires dont les bandes de fréquence correspondent à celles choisies.

Oxymétrie de Pouls

L'oxymétrie de pouls a été mise au point par Aoyagi en 1974 par l'association de deux techniques : la spectrophotométrie d'absorption pour la mesure de l'oxymétrie, et la photopléthysmographie pour la détection de l'onde de pouls [103]. La mesure de la saturation en oxygène du sang, par des techniques de transillumination du sang, avait été décrite dans les années 1930. Elle avait été utilisée en 1941 dans un oxymètre d'oreille, mis au point pour expliquer les pertes de connaissance des aviateurs britanniques. La saturation en oxygène exprime le contenu en oxygène, sous la forme du rapport de la quantité d' O_2 fixée par l'Hémoglobine (*Hb*) en fonction de la capacité totale de fixation de l'*Hb*. Dans l'oxymétrie de pouls, les contenus relatifs des deux hémoglobines (*HbO₂* et *Hb* réduite) sont mesurées par spectrophotométrie d'absorption, principe reposant sur la loi de Lambert-Beer. Désormais, des capteurs fonctionnent par absorption ou par réflexion permettant une utilisation plus souple (p. ex. capteurs posés sur le front). Au cours de chaque cycle cardiaque, l'absorption lumineuse varie cycliquement, avec en diastole une absorption lumineuse due au sang veineux, aux tissus, à l'os et aux pigments, et en systole une augmentation de l'absorption par addition d'une partie ondulatoire liée à la partie pulsatile du sang artériel.

La valeur de référence pour la saturation du sang en oxygène (*SaO₂*) est obtenue par des techniques d'oxymétrie gazeuse qui nécessitent le dosage d'un échantillon sanguin et donc un prélèvement invasif. Les oxymètres de pouls qui réalisent une mesure non invasive, ne permettent d'obtenir qu'une mesure reflet de la *SaO₂*, qui est préférentiellement désignée sous le terme de saturation pulsée en oxygène (*SpO₂*).

Le site de mesure peut influencer le temps de réponse des capteurs oxymétriques. Une comparaison des différents matériels existants montre une grande incertitude entre la mesure de *SpO₂* et *SaO₂*. On note de plus une différence suivant les constructeurs et technologies embarquées au sein du capteur. La corrélation devient douteuse en dessous de 55% pour les capteurs actuels [103]. On constate une perte du signal lors du choc et une grande sensibilité aux bruits (artefacts). Néanmoins, les capteurs d'oxymétrie se

sont imposés par leur souplesse d'utilisation lors du suivi des patients en anesthésie et en réanimation lors du suivi d'hypoxémie. Des protocoles d'utilisation bien adaptées aux patients suivis permettent de réduire les nombreuses alarmes (faux positifs) associés à cette mesure [104, 105]. La perte du signal est observée lors d'une hypoxémie sévère par vasoconstriction périphérique, ou lors d'une hypothermie sévère. Nous porterons une attention particulière sur ce signal.

Mesures et caractéristique retenues Deux capteurs ont été utilisés. Le premier a été relié à l'appareil de monitoring Kontron (K), l'autre à notre système d'enregistrement Biopac (B).

σ_{PoulsB} et σ_{PoulsK} Une estimation de la variation de l'amplitude du signal nous est donnée par l'estimation de son écart type (plus robuste que l'estimation de l'amplitude du signal). Cette caractéristique nous renseignerait sur la dynamique de la circulation périphérique.

$LFHF_{PoulsB}$ et $LFHF_{PoulsK}$ Le rapport BFHF (ou LFHF en anglais), rapport des basses fréquences sur les hautes fréquences est supposé donner une indication sur le couplage cardiorespiratoire.

SpO_2 La saturation pulsée en oxygène nous apporte une information sur l'oxygénation du sang. Elle apporte encore une information sur la circulation périphérique et l'exclusion des territoires associés au site de mesure.

Températures

Il existe deux types de capteurs permettant de mesurer la température pour les systèmes biologiques. Les thermomètres à mercure étalonnés restent les références pour un étalonnage précis. L'utilisation de capteurs électroniques nécessite un étalonnage constant et une vérification fréquente de la dérive temporelle des appareils est nécessaire pour s'assurer de leur stabilité. La référence est alors obtenue par une sonde de platine. Les sondes utilisées couramment sur les animaux sont les thermocouples (jonction de deux métaux de propriétés thermoélectriques différentes) et les thermistances (variations d'impédance, p. ex. sonde de platine). Les points de mesure sur l'organisme sont nombreux et peuvent se situer au niveau de régions ciblées. On retiendra les principales liées à la température cutanée, buccale, interne (centrale par sonde œsophagienne, le plus couramment rectale).

Température interne, cutanée et ambiante Nous avons estimé la température moyenne sur la fenêtre d'observation pour T_{int} , T_{cut} et T_{amb} :

$$T_{xxx} = \frac{1}{\Delta t} \int_{t_k}^{t_k + \Delta t} T_{xxx}(t) dt \quad (4.7)$$

Capnimétrie

Le capnimètre permet de mesurer le CO_2 rejeté par l'organisme. Il permet d'apporter une indication quantitative sur les échanges gazeux de l'organisme. La valeur du pla-

teau obtenu en fin d'expiration (angl. *End Tidal CO₂*, E_tCO_2) est le paramètre le plus couramment enregistrée.

Choc hémorragique

Nous avons utilisé une extrapolation linéaire pour transformer les échantillons réalisés approximativement tous les 25 g en série temporelle et ainsi resynchroniser les données entre elles. Les valeurs ainsi obtenues ont été rapportées en terme de pourcentage du volume sanguin total par rapport au poids total de l'animal suivant le rapport empirique de 69 ml/kg développé en 4.1.1.

4.3.3 Ensembles de capteurs

Nous avons utilisé plusieurs types de capteurs au cours de l'expérimentation de façon à obtenir des informations de qualités et contenu informationnel différents. Nous avons regroupés les données provenant de ces capteurs suivant deux ensembles auxquels nous ferons référence par la suite :

- l'ensemble invasif (resp. non invasif) : regroupe les données provenant de capteurs qui nécessitent un placement par effraction du territoire cutané et nécessitent donc un acte chirurgical (resp. non effraction du territoire cutané sans acte chirurgical). L'emploi de l'adjectif effractif semblerait plus approprié car invasif n'est pas trouvable dans un dictionnaire médical [28] (angl. *invasive*).
- l'ensemble portable (resp. non portable) (angl. *wearable*, resp. *not wearable*) : regroupe les données provenant de capteurs qui peuvent être portés (resp. ne peuvent pas être portés) par le combattant sur le terrain. Pour notre étude, nous considérons que les capteurs non invasifs utilisés sur l'animal lors de l'expérimentation sont portables.

4.3.4 Classes

Quatre classes ont été retenues pour la réalisation de la classification par un expert. Elles correspondent aux différentes états rencontrées au cours du choc hémorragique et sont rappelées Tab. 4.3. La description médicales de ces différentes classes est précisée

TAB. 4.3 – Classification retenue pour le choc hémorragique

Classe	État
1	Rien À Signaler
2	Choc Compensé
3	Choc Décompensé
4	Choc Irréversible

en 3.1.

L'expert s'est constitué un ensemble de données qu'il jugeait pertinentes pour réaliser cette classification, provenant de capteurs invasifs et de capteurs portables ($\{animal, temps, PerteSang, PAM, PAD, PAS, FC, T_{int}, T_{cut}, T_{amb}, RR_{Abd}, RR_{Tho}, \tilde{V}t_{Abd}, \tilde{V}t_{Tho}\}$). Les caractéristiques ont été calculées sur des fenêtres temporelles de 60 s. Il

a alors utilisé un raisonnement causal en faisant glissé une fenêtre de sept minutes et a attribué une classe aux données entrantes sur des périodes d'une minute en fonction de l'état diagnostiqué. Ainsi, pour chaque période d'une minute, une classe est disponible. L'expert a ainsi traité les enregistrements des animaux disponibles, de façon à fournir pour chaque animal, pour chaque instant, un score correspondant à la classe de choc hémorragique rencontrée.

4.4 Conclusion du chapitre

Nous avons vu que la réalisation de l'expérimentation animale associée à notre modèle de choc hémorragique à débit contrôlé nécessitait des règles protocolaires précises : d'une part assurer une efficacité maximale des temps opératoires et une cohésion de l'équipe au sein du bloc opératoire ; d'autre part pour un respect de l'animal pour lui éviter toute souffrance inutile (analgésie) et prise de conscience de son état (anesthésie) si les contraintes fixées par le modèle animal retenu le permettent ; enfin pour la mise en place de la chaîne d'acquisition et le placement des capteurs invasifs et portables retenus pour permettre un enregistrement des signaux dans les meilleures conditions.

Nous avons développé les techniques utilisées pour traiter les signaux et extraire les caractéristiques qui nous semblaient pertinentes sur des fenêtres temporelles de $\Delta t = 60\text{ s}$ en développant les particularités et difficultés associées à ces signaux biologiques.

Enfin, nous avons présenté le processus adopté par l'expert pour classer les données en quatre stades de choc hémorragique.

L'ensemble de ces caractéristiques ou attributs constitue le matériel qui nous a servi pour la suite de l'étude et l'apprentissage supervisé des classes de l'expert. Un résumé des caractéristiques est proposée Tab. 4.4 et 4.5.

TAB. 4.4 – Caractéristiques invasives de l'étude

F_i	Abréviation	Caractéristique	Unité
F_1	<i>temps</i>	Début de l'intervalle servant à l'estimation des caractéristiques	(s)
F_2	<i>animal</i>	Nom de l'animal (date de l'expérimentation)	(yyyyddmm)
Caractéristiques invasives			
F_3	<i>PerteSang</i>	Pourcentage de volume sanguin prélevé	(%)
F_4	<i>PAM</i>	Pression artérielle moyenne	(mmHg)
F_5	Δ_{PA}	Différentielle de pression	(mmHg)
F_6	<i>PAD</i>	Pression artérielle diastolique	(mmHg)
F_7	<i>PAS</i>	Pression artérielle systolique	(mmHg)
F_8	σ_{PA}	écart type de la pression artérielle	(mmHg)
F_9	<i>LFHF_{PVC}</i>	Basses fréquences sur hautes fréquences pour la pression veineuse centrale	sans dimension
F_{10}	T_{int}	Température interne	(°C)
F_{11}	E_tCO_2	End tidal CO_2	(mmHg)

TAB. 4.5 – Caractéristiques non invasives de l'étude

F_i	Abréviation	Caractéristique	Unité
Caractéristiques portables			
F_{12}	σ_{ECG}	Écart type <i>ECG</i>	(V)
F_{13}	FC	Fréquence cardiaque	(bpm)
F_{14}	RR_{Tho}	Rythme respiratoire (ou fréquence respiratoire) au niveau du thorax	(cpm)
F_{15}	$\tilde{V}t_{Tho}$	Pseudo volume courant thorax	%
F_{16}	RR_{Abd}	Rythme Respiratoire (ou fréquence respiratoire) au niveau de l'abdomen	(cpm)
F_{17}	$\tilde{V}t_{Abd}$	Pseudo volume courant abdomen	%
F_{18}	SpO_2	Saturation pulsée en oxygène	%
F_{19}	σ_{PoulsK}	Écart type pouls Kontron	V
F_{20}	$LFHF_{PoulsK}$	Basses fréquences sur hautes fréquences pouls Kontron	sans dimension
F_{21}	σ_{PoulsB}	Écart type pouls Biopac	%
F_{22}	$LFHF_{PoulsB}$	Basses fréquences sur hautes fréquences pouls Biopac	sans dimension
F_{23}	T_{cut}	Température cutanée	(°C)
F_{24}	T_{amb}	Température ambiante	(°C)
F_{25}	<i>classes</i>	classes	(h)

bpm : battements par minute ; cpm : cycles par minute.

Chapitre 5

Méthodes

L'« inflation expérimentale » n'est pas moins pernicieuse que l'inflation économique : on a des instruments, on les utilise massivement et on en tire une masse infinie de données desquelles, à la fin, on ne sait rien tirer. Les données emplissent des bibliothèques entières, dorment dans des archives poussiéreuses et personne ne s'en occupe plus.

René Thom, *Paraboles et Catastrophes*, 1980, champs Flammarion, Paris, p. 54.

Dans ce chapitre, nous exposons les méthodes que nous avons utilisées pour construire et valider les classifications réalisées sur les données (les caractéristiques extraites précédemment). Dans la section 5.1, nous présentons les transformations retenues de façon à obtenir des mesures comparables entre les données. La section 5.2 expose les techniques de reconnaissance de formes et les différents classifieurs utilisés pour l'apprentissage supervisé de la classification des états de choc hémorragique réalisé par l'expert. La section 5.3 expose les techniques de validation et les stratégies adoptées pour obtenir les classifieurs les plus performants.

5.1 Espace de représentation des données

Dans ce chapitre, les données feront référence aux caractéristiques extraites précédemment. La recherche d'un espace bien adapté à la nature des données est une étape importante pour amener à une représentation pour laquelle les outils mathématiques couramment utilisés sont efficaces. Cette recherche s'accompagne d'un certain nombre de transformations appliquées aux données. Ces transformations permettent d'obtenir un espace des formes (espace de représentation dans lequel la classification est réalisée) où la notion de distance à un sens. Ces techniques font l'objet de l'analyse de données qui répond à bon nombre de ces questions [106] et dont nous rappelons quelques définitions.

5.1.1 Quelques définitions

Les données ont été regroupées dans un tableau de façon à disposer la masse d'information sous forme rectangulaire dans lequel les lignes ($i = 1, \dots, n$) représentent les

observations, appelés plus généralement *unités statistiques*; les colonnes ($j = 1, \dots, d$) représentent les *attributs*, *caractères* ou *caractéristiques* observés [107].

La *matrice des observations* associées aux d caractéristiques retenues pour l'étude et aux n observations est généralement notée $\mathbf{X}$. Les observations sont issues de différents *individus*, ou éléments d'un ensemble déterminé appelé *population* [108].

Pour notre étude nous avons regroupé les différentes observations temporelles réalisées sur nos différents individus (nos animaux) ainsi que la supervision désirée, dans le tableau suivant :

$$\mathbf{M} = \begin{pmatrix} x_1(t_0(1)) & x_2(t_0(1)) & \cdots & x_d(t_0(1)) & u(t_0(1)) \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ x_1(t_f(1)) & x_2(t_f(1)) & \cdots & x_d(t_f(1)) & u(t_f(1)) \\ \vdots & & \vdots & & \vdots \\ x_1(t_k(i)) & x_2(t_k(i)) & \cdots & x_d(t_k(i)) & u(t_k(i)) \\ \vdots & & \vdots & & \vdots \\ x_1(t_1(N)) & x_2(t_1(N)) & \cdots & x_d(t_1(N)) & u(t_1(N)) \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ x_1(t_f(N)) & x_2(t_f(N)) & \cdots & x_d(t_f(N)) & u(t_f(N)) \end{pmatrix} \quad (5.1)$$

Avec $t_k(i) = t_0(i) + k\Delta t$ pour chaque individu i , x_j désignant la j ème caractéristique étudiée et $u(t_k(i))$ la classification réalisée par un expert sur l'observation réalisée à l'instant t_k sur l'individu i . Nous regroupons ainsi toutes les données dans un tableau $\mathbf{M}$ notre base de données. L'ensemble des exemples (observations associées à la classification de l'expert, unités statistiques) disponibles est alors noté $\mathcal{S} = \{(\mathbf{x}_i, u_i), i = 1, \dots, n\}$.

$\mathbf{M}$ regroupe donc deux tableaux : $\mathbf{X}$ contient n observations en ligne sur d caractéristiques en colonne; $\mathbf{U}$ contient les n classifications de l'expert et constitue la matrice de supervision associée à la matrice des observations. Le terme générique de la matrice $\mathbf{X}$ des observations sera noté x_{ij} par la suite, i représentant l'indice de la ligne pointant sur une observation et j l'indice de la caractéristique pointée. On note alors que chaque observation de $\mathbf{X}$ se situe dans un espace $\mathcal{X}$ de $\mathbb{R}^d$ appelé espace de représentation. L'espace de représentation, utilisé pour la classification, est encore désigné d'espace des formes et les différentes observations le constituant sont encore appelés des vecteurs formes. Les classes ou étiquettes associées se situent dans un espace $\mathcal{U}$ dit de supervision.

Le tableau $\mathbf{X}$ des observations peut contenir des données de natures différentes : qualitatives ou quantitatives. Dans le cadre de notre étude, nous considérons le cas de variables quantitatives mesurées dans des échelles différentes¹. Afin de pouvoir réaliser une analyse exploratoire multidimensionnelle sur cet ensemble de données dont les ordres de grandeurs sont différents, il est important d'effectuer une opération qui rende les données comparables entre elles et indépendantes de l'échelle choisie. Il s'agit alors de réaliser les changements de variables adéquats pour rendre les caractéristiques comparables entre elles².

¹Pour l'étude d'un tableau de données contenant des variables qualitatives et quantitatives, il est conseillé de tout décomposer en modalités qualitatives et passer à un tableau disjonctif complet pour obtenir un tableau de données de même type. Nous renvoyons le lecteur à l'ouvrage de Celeux [109]

²Les principales étapes préliminaires à une analyse de données sont décrites dans les ouvrages de références [106, 107].

5.1.2 Changement de représentation par changement de variables

On parle encore de codage par changement de variables [106].

Blanchiment, variables centrées réduites, zscore

Le blanchiment (angl. *whitening process, whitening transformation*) des données consiste à mettre les données à la même échelle. Pour cela, une transformation est appliquée telle que dans la nouvelle base de représentation, la matrice de covariance des données soit égale à la matrice identité. Il peut être réalisé sous forme multidimensionnelle [110] en considérant les distributions conjointes des données. Lorsque la transformation des données est réalisée sur chaque caractéristique individuelle (cas monodimensionnel), l'opération est généralement désignée sous le terme de *zscore* et a pour but de fournir des variables centrées réduites (donc indépendante d'une échelle de mesure).

Elle consiste alors à réaliser le changement de variable suivant :

$$z_{ij} = \frac{(x_{ij} - \mu_j)}{\sigma_j} \quad (5.2)$$

avec μ_j moyenne des observations disponibles sur la caractéristique x_j :

$$\mu_j = \frac{1}{n} \sum_{i=1}^n x_{ij} \quad (5.3)$$

et σ_j écart-type de la caractéristique x_j ,

$$\sigma_j = \frac{1}{(n-1)} \left(\sum_{i=1}^n (x_{ij} - \mu_j)^2 \right)^{1/2} \quad (5.4)$$

Au terme de cette étape, pour chaque caractéristique, les données sont centrées sur l'origine et ont une variance égale à un.

Transformations

L'opération de blanchiment amène les données dans une hypersphère de centre 0 et de rayon 1 mais la distribution des données n'est pas forcément une loi de Laplace Gauss et le calcul des distances dans cet espace sera faussé. De façon à se rapprocher le plus d'une distribution de Laplace Gauss centrée réduite, on peut utiliser les transformations couramment utilisées en statistiques comme les transformations de type $z = x^\alpha$ décrite dans [110] ou celles décrites par Gasser [111] sur des données électroencéphalographiques (utilisées pour rechercher les conséquences de ces transformations sur les distributions des données observées sur des fenêtres temporelles différentes). Celles ci sont les suivantes : $\log(x)$, $\log(1+x)$, $\sqrt{x}$, $\sqrt[3]{x}$, $\log((x)/(1-x))$, $1/(\sqrt{x})$, $\arcsin(\sqrt{x})$.

5.1.3 Représentation des données de notre étude

Base de données initiale

Les différentes caractéristiques extraites constituant la base de données sont résumées dans les tableaux Tab. 4.4 et 4.5.

Cette base de données initiale regroupent des caractéristiques dont les échelles et les distributions diffèrent d'une variable à l'autre. Afin d'utiliser un espace plus approprié pour une analyse exploratoire multidimensionnelle, nous avons effectué les transformations suivantes.

Changements de représentation retenues

Nous avons observé l'influence de chaque transformation décrite dans la section 5.1.2 sur la distribution de chaque caractéristique de l'ensemble des observations disponibles. Des transformations de type logarithmique $\log$ ont été retenue pour les caractéristiques de $LFHF_{PVC}$, $LFHF_{PoulsB}$ et $LFHF_{PoulsK}$, car elles permettaient d'obtenir des distributions se rapprochant d'une distribution de Laplace Gauss. Le résultat graphique d'une telle transformation est présenté Fig. 5.1.

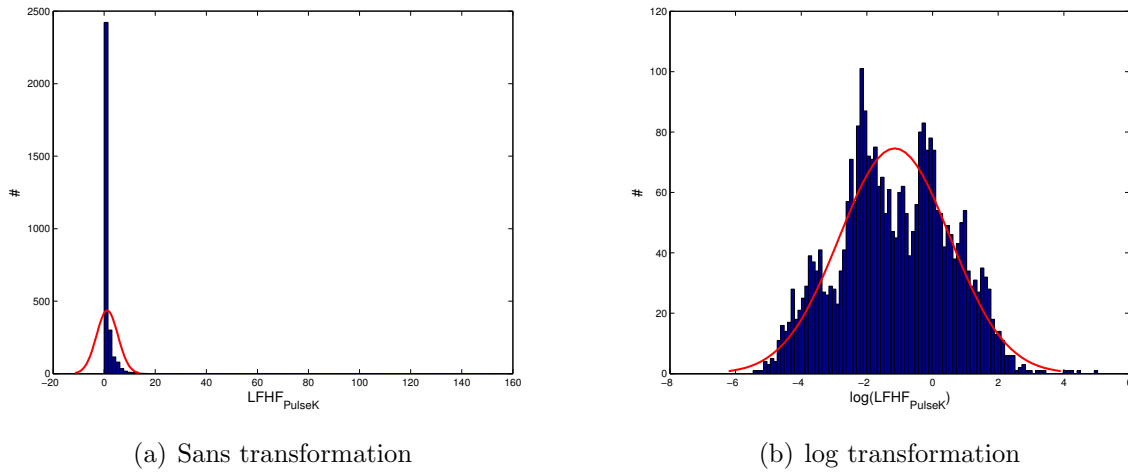


FIG. 5.1 – Effet de la transformation $\log$ vers une distribution normale : la caractéristique du rapport de puissance des basses fréquences sur les hautes fréquences de l'onde de pouls Kontron peut être représentée par son histogramme : a) la distribution de la caractéristique suit une loi de type exponentielle très différente d'une loi de Laplace Gauss (en rouge), b) une transformation logarithmique permet d'ajuster au mieux la distribution de cette variable sur une loi de Laplace-Gauss.

A la suite de ces transformations, un *zscore* a été systématiquement réalisé de façon à obtenir des variables centrées réduites. Nous avons retenu cette méthode car celle-ci ne fait aucune hypothèse sur la distribution conjointe des caractéristiques à cette étape du traitement de l'information. Ceci permet d'effectuer une sélection de variables par une méthode ascendante que nous développerons par la suite.

La base de données ainsi obtenue peut alors être représentée graphiquement Fig. 5.2.

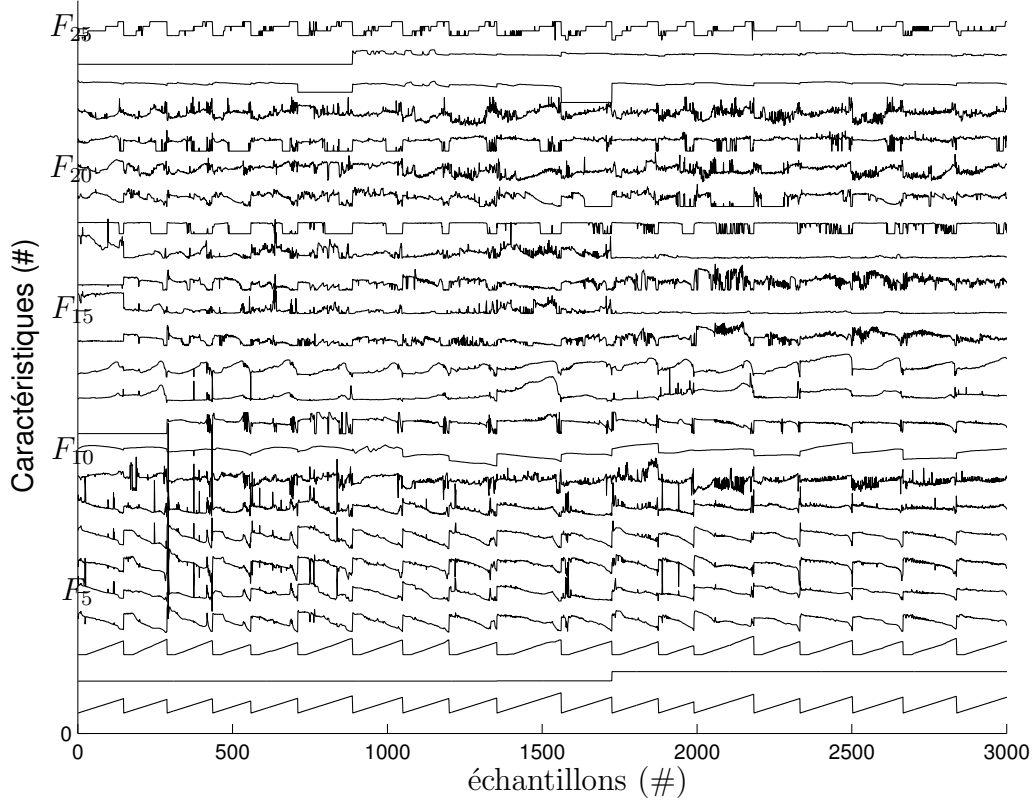


FIG. 5.2 – Représentation graphique de la base de données. Les échantillons des différents animaux mis bout à bout sont reportés en abscisse. Les différentes caractéristiques définies Tab. 4.4 et 4.5 sont en ordonnée. Les dents de scie observées correspondent au changement d'animaux. Ce graphique met en évidence les comportements communs des animaux au cours de l'expérimentation : la variation de fréquence cardiaque (F_{13}) au cours de l'expérimentation pour chaque animal en donne un exemple.

Dans cet espace où les données ont subi des transformations et ont été centrées réduites, nous pouvons alors utiliser une distance euclidienne définie par :

$$d(\mathbf{x}, \mathbf{y}) = \left(\sum_{j=1, \dots, d} (x_j - y_j)^2 \right)^{1/2} = (\mathbf{x} - \mathbf{y})^T \mathbf{I}_d (\mathbf{x} - \mathbf{y}) \quad (5.5)$$

L'étude réalisée en annexe A montre que le résultat de ces transformations permet de comparer différents classifieurs sans en avantager un par rapport à un autre et permet ainsi une comparaison non biaisée des résultats obtenus. Les changements de représentations spécifiques à chaque école de classifieurs (codage flou par exemple) ont été incluses dans les classifieurs spécifiques de façon à partir d'une base de données dont les échelles étaient comparables.

5.2 Apprentissage artificiel, reconnaissance de formes

Ce chapitre présente les méthodes utilisées pour l'apprentissage supervisé du score de choc hémorragique réalisé par un expert. Dans un premier temps nous présentons l'objectif visé par l'analyse exploratoire multidimensionnelle dans un espace de représentation et principalement les objectifs visés par l'apprentissage supervisé. Nous présenterons alors les algorithmes que nous avons retenus.

5.2.1 Objectifs de l'analyse exploratoire

« L'analyse exploratoire aura pour but non seulement de présenter mais encore d'analyser, de découvrir, parfois de vérifier ou de prouver, éventuellement mettre à l'épreuve certaines hypothèses » [107]. L'analyse exploratoire a donc pour but de faire émerger des *modèles* reliant les éléments d'une base de donnée. On parle de *reconnaissance de formes* (angl. *pattern recognition*) lorsqu'on effectue une prise de décision pour regrouper des données. On parle de *classification* ou de *discrimination* lorsqu'on attribue une classe à un ensemble de données. On parle de *régression* lorsqu'on cherche à expliquer une variable quantitative par une combinaison d'autres variables. Suivant la connaissance disponible sur la base de données, on parle de méthodes *supervisées* lorsque celles ci sont guidées par une connaissance extérieure et de méthodes *non supervisées* lorsqu'aucune connaissance a priori n'est disponible. L'emploi plus rigoureux de *classement* (angl. *classification*) est utilisé lorsqu'on effectue une classification par apprentissage supervisé et l'emploi de *classification automatique* (angl. *clustering*) est préféré pour une méthode non supervisée.

Depuis l'apparition de l'informatique et l'émergence des domaines d'*intelligence artificielle*, on parle d'*apprentissage artificiel* ou *apprentissage automatique* (angl. *machine learning*), pour regrouper les méthodes permettant de construire un modèle de la réalité à partir d'observations, soit en améliorant un modèle partiel, soit en créant complètement le modèle à partir d'algorithmes.

On parlera de *fouille de données* (*data mining*), ou *extraction de connaissances à partir de données* (ECD) (angl. *knowledge discovery in data*, KDD), lorsque le processus complet d'extraction sera pris en compte par la machine : stockage de la base de donnée, sélection des données, nettoyage des données, apprentissage, présentation des modèles à l'utilisateur pour validation [26].

Le problème fondamental de l'apprentissage réside dans la dualité entre *précision* du modèle et *généralisation*. Par rapport aux données présentes dans la base de données, il va falloir trouver un modèle qui puisse être généralisable à d'autres observations, et correspondre pour le mieux aux données présentes. Pour certaines bases et certaines méthodes on aura tendance à privilégier un apprentissage généraliste et l'on essaiera d'éviter un *surapprentissage* ou un *apprentissage par cœur*.

5.2.2 Définitions liées à l'apprentissage supervisé

Le schéma Fig.5.3 propose une représentation des acteurs impliqués dans l'apprentissage supervisé : à l'aide d'un savoir expert, une machine va chercher un modèle optimal d'interprétation de plusieurs observations de la nature. D'une façon plus générale, le problème de l'apprentissage supervisé nécessite la définition de plusieurs acteurs [26] :

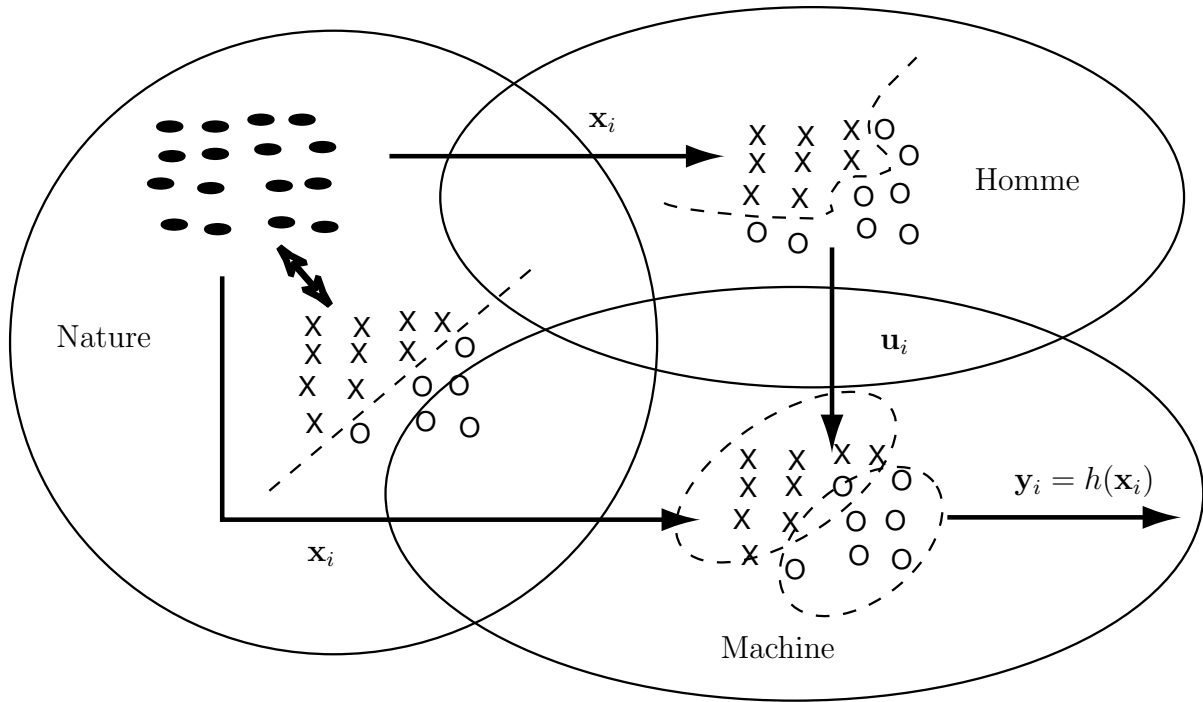


FIG. 5.3 – Synoptique de l'apprentissage d'un modèle. On fait l'hypothèse qu'un modèle existe dans la nature permettant d'associer à chaque observation $\mathbf{x}_i$ une classe sur le schéma. Un homme interprète les observations recueillies et construit son propre modèle qui associe à chaque $\mathbf{x}_i$ une réponse désirée $\mathbf{u}_i$. La machine construit son modèle dont la sortie est $\mathbf{y}_i = h(\mathbf{x}_i)$, à partir des observations de la nature $\mathbf{x}_i$ et du modèle proposé par l'homme : elle réalise un apprentissage. Chaque modèle (nature, homme, machine) peut être différent.

- la *nature* (l'*environnement*...) : dont les phénomènes fournissent les données, les observations et dont on suppose qu'il existe une structure interne représentable sous la forme de relations dont on essaie de tirer des modèles. Les observations ou vecteurs formes $\mathbf{x}_i$ sont dans un espace d'entrée $\mathcal{X}$.
- l'*oracle* (le *professeur*, le *superviseur*, l'*expert*...) qui propose un résultat. Ainsi, pour chaque observation de $\mathcal{X}$, l'oracle propose une *réponse désirée* (*étiquette*, *classe*...) $\mathbf{u}_i$ dans un espace de sortie $\mathcal{U}$.
- le *modèle* (l'*apprenant*, la *machine*, le *classifieur*...) : qui réalise une fonction entre les observations fournies par la nature et une réponse dans l'espace de sortie. Cette fonction h est choisie dans un espace de fonctions $\mathcal{H}$ de façon à obtenir $\mathbf{y}_i = h(\mathbf{x}_i)$ dans l'espace de sortie $\mathcal{Y}$ (h correspond à une hypothèse reliée au modèle cherché).
- la *tâche d'apprentissage* (*phase d'apprentissage*...) va consister à trouver un modèle dont la réponse est la plus proche de celle prévue par l'oracle. En d'autres termes, l'apprentissage consiste à optimiser un modèle de façon à minimiser l'erreur entre sa sortie et celle prévue par l'oracle. De façon à pondérer les erreurs réalisées, on introduit une fonction de *coût* ou *perte* (angl. *loss function*) $l(\mathbf{u}_i, h(\mathbf{x}_i))$. L'apprentissage consiste alors à minimiser l'espérance du coût ou de la perte, ou encore minimiser le *risque* défini par :

$$R_{\text{Réel}}(h) = \int_{\mathbf{z}=\mathbf{x} \times \mathbf{u}} l(\mathbf{u}_i, h(\mathbf{x}_i)) dF(\mathbf{x}, \mathbf{u}) \quad (5.6)$$

Le calcul de ce risque nécessite la connaissance totale de $\mathcal{X}$ et $\mathcal{U}$ et de leurs relations

exprimées par la densité de probabilité conjointe $dF(\mathbf{x}, \mathbf{u})$. Cela n'est généralement pas possible et seul un échantillon est disponible $\mathcal{S} = \{(\mathbf{x}_1, \mathbf{u}_1), (\mathbf{x}_2, \mathbf{u}_2), \dots, (\mathbf{x}_m, \mathbf{u}_m)\}$. L'estimation de ce risque permettra de se rendre compte de la performance de l'apprentissage. Le processus permettant de minimiser ce risque constituera la phase d'apprentissage. En d'autres termes, l'apprentissage dépendra des observations de la nature disponibles, des interprétations de l'oracle et du principe inductif permettant de réaliser l'estimation et la minimisation de ce risque réel.

Ayant défini ces différents acteurs, nous pouvons représenter les possibilités d'erreurs qui peuvent s'insérer parmi les différents acteurs de ce processus d'apprentissage. La figure Fig. 5.4 propose une représentation schématique des différentes erreurs qui peuvent être générées lors de l'utilisation d'une technique d'apprentissage supervisé.

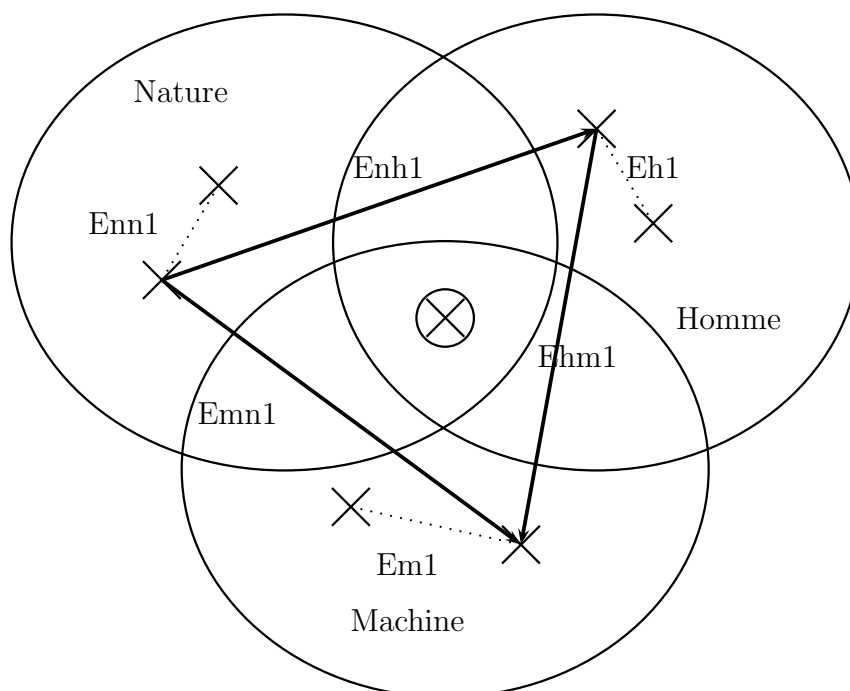


FIG. 5.4 – Erreurs de modélisation entre la nature, l'homme et la machine : Les croix représentent des modèles qui décrivent les données. Enn1) Existe-t-il un modèle qui régit la nature ? il existe toujours des erreurs de mesures qui viennent fausser la réalité. Enh1) Erreur de modélisation entre l'homme (oracle, superviseur, professeur) et la nature. Eh1) Erreur de modélisation entre hommes (Erreur interscorer, etc.). Ehm1) Erreur de modélisation entre l'homme et la machine (étudiant) après un apprentissage supervisé. Em1) Différences entre les différents modèles obtenus par la machine après apprentissage supervisé par l'homme. Les différentes combinaisons possibles des sources d'erreur sont possible. La machine peut être plus proche de la nature que la vision experte après apprentissage. Lorsqu'il n'y a pas d'erreurs entre les modèles, on se trouve au centre du schéma. C'est ce que l'on cherche à obtenir avec l'apprentissage supervisé.

On utilise le terme de classifieur pour désigner l'opérateur réalisant la fonction h lorsque la sortie désirée fait partie d'un ensemble fini $\mathcal{C}$ de C valeurs possibles appelées classes ω_i ou étiquettes.

5.2.3 Classifieurs retenus

L'apprentissage supervisé des classes de choc hémorragique proposées par l'expert a été testé sur différents types de classifieurs de façon à retenir le plus performant. Dans cette partie, nous présentons les modèles retenus et leurs spécificités.

Classifieurs basés sur la Règle de Bayes

De nombreux classifieurs utilisent la règle de Bayes (5.7) qui donne une performance de classification maximale lorsque toutes les propriétés statistiques du phénomène observé sont connues. Leur description est largement décrite dans [110, 33, 26]. L'attribution d'une classe au vecteur $\mathbf{x}$ se fait par comparaison des probabilités conditionnelles a posteriori $P(\omega_i|\mathbf{x})$ dont le calcul est défini par la règle de Bayes :

$$P(\omega_i|\mathbf{x}) = \frac{f(\mathbf{x}|\omega_i)P(\omega_i)}{f(\mathbf{x})} \quad (5.7)$$

avec

$$f(\mathbf{x}) = \sum_{k=1}^C f(\mathbf{x}|\omega_k)P(\omega_k) \quad (5.8)$$

$P(\omega_i|\mathbf{x})$ constitue la probabilité a posteriori d'appartenir à la classe ω_i pour l'observation $\mathbf{x}$. $P(\omega_i)$ correspond à la probabilité a priori d'appartenir à la classe ω_i sans connaissance de l'observation $\mathbf{x}$, $f(\mathbf{x}|\omega_i)$ correspond à la densité de probabilité de $\mathbf{x}$ sachant ω_i .

La règle de décision d'un classifieur utilisant la règle de Bayes sera obtenue en attribuant la classe qui maximise la probabilité a posteriori d'appartenance à une classe suivant l'observation :

$$h(\mathbf{x}) = \arg \max_{i \in \{1, \dots, C\}} P(\omega_i|\mathbf{x}) \quad (5.9)$$

Ce qui revient à trouver :

$$h(\mathbf{x}) = \arg \max_{i \in \{1, \dots, C\}} f(\mathbf{x}|\omega_i)P(\omega_i) \quad (5.10)$$

L'apprentissage supervisé consistera donc à évaluer les différentes densités de probabilités pour l'observation dans son espace de représentation représenté par le vecteur forme $\mathbf{x}$. Pour cela, on peut émettre plusieurs hypothèses sur la nature de la densité de probabilité $f(\mathbf{x}|\omega_i)$. On est ramené à un problème d'inférence statistique (comment d'un ensemble d'observations, trouver la propriété statistique de l'ensemble de la population...).

Modèles paramétriques L'estimation de la densité de probabilité $f(\mathbf{x}|\omega_i)$ repose sur un modèle gaussien multidimensionnel défini par l'équation :

$$p(\mathbf{x}|\omega_i) = \frac{1}{(\sqrt{2\pi})^d \sqrt{|\Sigma|}} \exp \left(-\frac{1}{2} (\mathbf{x} - \mu)^T \Sigma^{-1} (\mathbf{x} - \mu) \right) \quad (5.11)$$

avec pour paramètres, sa moyenne μ , et sa matrice de covariance Σ , définies par :

$$\mu = E[\mathbf{x}] = (\mu_1, \mu_2, \dots, \mu_d)^T \quad (5.12)$$

$$\Sigma = E[(\mathbf{x} - \mu)(\mathbf{x} - \mu)^T] \quad (5.13)$$

Suivant les hypothèses faites sur la matrice de covariance, et la manière de réaliser l'estimation de celle-ci, on obtient les classifieurs suivants :

Classifieur linéaire (angl. *linear*) La matrice de covariance $\Sigma = \Sigma_i$ est identique pour chaque classe. Cela revient à considérer une distance de Mahalanobis, entre le vecteur $\mathbf{x}$ et le vecteur des moyennes μ , défini par :

$$d_{Mahal}(\mathbf{x}, \mathbf{y}) = (\mathbf{x} - \mathbf{y})^T \Sigma^{-1} (\mathbf{x} - \mathbf{y}) \quad (5.14)$$

avec Σ identique pour chaque classe.

Les isoprobabilités sont des ellipses identiques centrées sur les moyennes de chaque classe. Le résultat de la classification d'un tel opérateur se traduit par une classification suivant des régions séparées par des frontières linéaires.

Classifieur quadratique (angl. *quadratic*) Les matrices de covariance Σ_i sont estimées pour chaque classe à partir des observations disponibles dans l'ensemble qui sert à l'apprentissage $\mathcal{A}$, selon l'algorithme suivant :

Estimation de la matrice de variance-covariance par classe

Début Pour chaque classe ω_k

Repérer les observations de cette classe $\mathcal{A}_{\omega_k} = \{\mathbf{z} \in \mathcal{A} | \mathbf{u} = \omega_k\}$

Soit $n_{\omega_k} = \sharp \{\mathcal{A}_{\omega_k}\}$, le nombre d'observations de cette classe.

Estimer le vecteur moyenne $\mu_{\omega_k} = (\mu_{\omega_k 1}, \dots, \mu_{\omega_k d})^T$ sur $\mathcal{A}_{\omega_k}$

$$\hat{\mu}_{\omega_k j} = \frac{1}{n_{\omega_k}} \sum_{i=1}^{n_{\omega_k}} x_{ij} \quad (5.15)$$

Dupliquer celui-ci pour le soustraire à la matrice des observations de cette classe de façon à centrer les données.

$$\mathbf{Xc}_{\omega_k} = \mathbf{X}_{\omega_k} - \mathbf{1}_{(n_{\omega_k}, 1)} \mu_j^T \quad (5.16)$$

Estimer la matrice de variance-covariance :

$$\hat{\Sigma}_{\omega_k} = \frac{1}{(n_{\omega_k} - 1)} (\mathbf{Xc}_{\omega_k}^T \mathbf{Xc}_{\omega_k}) \quad (5.17)$$

Fin pour Fin

Les isoprobabilités sont des ellipses indépendantes pour chaque classe, centrées sur chaque moyenne. Le résultat de la classification d'un tel opérateur se traduit par une classification suivant des régions séparées par des frontières quadratiques.

Classifieur naïf (angl. *naive*) Les variables sont considérées comme indépendantes l'une de l'autre, ce qui revient à écrire la densité de probabilité sous la forme :

$$f(\mathbf{x} | \omega_i) = \prod_{j=1}^d f(x_j | \omega_i) \quad (5.18)$$

Chaque variable subit une classification selon une loi de Gauss unidimensionnelle en estimant la moyenne et la variance pour chaque classe de la fonction :

$$f(x_j|\omega_i) = \frac{1}{\left(\sqrt{2\pi\sigma_j^2}\right)} \exp\left(-\frac{1}{2}\left(\frac{x_j - \mu_j}{\sigma_j}\right)^2\right) \quad (5.19)$$

Modèles non paramétriques La densité de probabilité est définie par l'équation 5.20 qui ne fait pas d'hypothèses sur l'utilisation d'un modèle statistique, ce qui laisse plus de liberté sur la densité possible.

$$f(\mathbf{x}|\omega_i) = \frac{k_{n_i}}{n_i V_{n_i}} \quad (5.20)$$

avec n_i nombre de représentants de la classe ω_i dans le volume V_{n_i}

Classifieur par k plus proches voisins La densité de probabilité est estimée en fixant le nombre de voisins k_n et en regardant le volume qu'ils occupent. Le choix du maximum a posteriori par la règle de Bayes est équivalent à une procédure de vote sur les k plus proches voisins du vecteur à classer. L'algorithme que nous avons retenu est un algorithme basique de classification par les k plus proches voisins (kPPV), qui peut se présenter ainsi :

kPPV

Début

Soit $\mathbf{x}$ de classe inconnue

Soit k le nombre de voisins recherchés ($1 \leq k \leq n$)

Calculer toutes les distances de $\mathbf{x}$ aux $n_{\mathcal{A}}$ données de l'apprentissage $\mathcal{A}$.

Créer un nouvel ensemble $\mathcal{A}_o$ comportant les données de $\mathcal{A}$ ordonnées suivant les distances croissantes calculées précédemment.

Choisir les k premières valeurs dans $\mathcal{A}_o$. Ce sont les k plus proches voisins de $\mathbf{x}$.

Si une classe est majoritaire parmi ses k plus proches voisins alors attribuer cette classe à $\mathbf{x}$.

Sinon plusieurs classes sont majoritaires alors faire une attribution aléatoire de ces classes majoritaires.

Fin si

Fin

Le résultat d'une telle classification se traduit par la détermination de frontières non linéaires (linéaires par morceaux) entre les classes ³

Classifieur par estimation de Parzen avec noyau Gaussien La densité de probabilité est estimée en évaluant l'influence des observations d'apprentissage présentes dans un volume fixé V_n . Pour cela, à chaque observation, un noyau est attribué qui correspond

³Pour un algorithme des 1PPV, les frontières correspondent à certains parcours d'un diagramme de Voronoï, dual d'une triangulation de Delaunay.

à une densité de fonction $K(\cdot)$. La somme de ces noyaux sur l'ensemble d'apprentissage réparti par classe fournit la densité de probabilité dans cette région suivant :

$$f(\mathbf{x}|\omega_i) = \frac{1}{n_i V_{n_i}} \sum_{j=1}^{n_i} K\left(\frac{\mathbf{x} - \mathbf{x}_{i,j}}{h_{n_i}}\right) \quad (5.21)$$

avec

$$V_{n_i} = h_{n_i}^d = n_i^{-\alpha} \quad (5.22)$$

Dans notre étude nous avons choisi le noyau Gaussien suivant :

$$K(u) = \frac{1}{(\sqrt{2\pi})^d} \exp\left(-\frac{1}{2}(u^2)\right) \quad (5.23)$$

Le résultat d'une telle classification se traduit par des régions séparées par des frontières non linéaires.

Modèle aléatoire L'attribution d'une classe se fait aléatoirement suivant une loi de distribution équiprobable :

$$P(\omega_i|\mathbf{x}) = \frac{1}{C} \quad (5.24)$$

Cette règle permet de mettre en évidence les performances obtenues par une classification aléatoire.

Classifieurs basés sur un calcul direct de frontières

Réseaux de neurones

Quelques définitions Nous ne décrivons que brièvement les réseaux de neurones qui font l'objet de travaux abondants et de plusieurs ouvrages très bien fournis [112, 113, 114, 115].

Le réseau de neurones artificiel (RN) (angl. *artificial neural network*) est un algorithme qui réalise une fonction entre plusieurs entrées et plusieurs sorties par la mise en association de plusieurs unités fonctionnelles appelées neurones artificiels. Chaque neurone artificiel a la particularité de recevoir plusieurs entrées et de ne disposer que d'une seule sortie. Il réalise une réponse à ces différentes entrées. La comparaison formelle à une cellule neuronale animale provient de la correspondance entre les synapses (les entrées du système $\mathbf{x}$), leur longueur ou propriétés conductrices (les poids $\mathbf{w}$ associées à chaque entrée), l'intégration réalisée au sein du soma (généralement une fonction somme $a = \mathbf{w}^t \mathbf{x}$), et le potentiel d'action déclenché au niveau de l'axone (sortie y du système) lorsque l'intégration au sein du soma des messages synaptiques dépasse un certain potentiel (par l'intermédiaire d'une fonction sigmoïdale $f(a)$ et d'un biais b).

On parle encore de perceptron multicouches (PMC) (angl. *multilayer perceptron* MLP) lorsque le réseau de neurones est composé de plusieurs couches (une couche faisant référence à un ensemble de neurones n'ayant pas de connexions entre eux) avec : une couche d'entrée (généralement reliée aux vecteurs forme du système), une couche de sortie reliée aux sorties du système (qui correspondent aux classes pour un problème de classification ou aux variables expliquées pour un problème de régression) et de couches cachées reliant

la couche d'entrée à celle de sortie. La comparaison de la sortie obtenue à la sortie désirée fournit une erreur, qui peut, lors de la phase d'apprentissage permettre d'ajuster les poids du réseau par rétropropagation du gradient de l'erreur. La minimisation de cette erreur (qui est une erreur empirique) permet de converger vers une architecture stable mais qui souvent correspond à une solution locale. Plusieurs techniques d'optimisation permettent de rechercher le réseau le mieux adapté à l'étude, en particulier l'utilisation de plusieurs initialisation aléatoire pour les poids du modèle permet de retenir le réseau le plus performant.

Architecture retenue Pour notre étude, nous avons utilisé un perceptron multicouches souvent utilisé en discrimination [114]. Son architecture est la suivante :

- une couche d'entrée constituée de d unités reliée à première couche cachée constituée de d unités avec fonction de transfert de type tangente hyperbolique $y = 2 / (1 + \exp(-2x)) - 1$
- une couche cachée constituée de d unités avec fonction de transfert linéaire $y = x$.
- une couche de sortie constituées de C unités avec fonction de transfert de type sigmoïdales $y = 1 / (1 + \exp(-x))$.
- L'initialisation des poids $\mathbf{W}$ du réseau a été déterminé aléatoirement et la modification de ceux-ci a été réalisée par rétropropagation du gradient selon un algorithme de Levenberg-Marquardt.
- Pour chaque apprentissage effectué, un ensemble de 5 réseaux (5 initialisations différentes) a été testé et celui obtenant les meilleures performances a été retenu.

Les perceptrons multicouches (PMC) sont des classifieurs dont le comportement dans l'espace des formes, correspond à une recherche directe des séparatrices de classification. Ils permettent d'obtenir des séparatrices non linéaires plus ou moins complexes. Nous n'avons pas réalisé d'étude pour déterminer la structure la mieux adaptée (nombre de neurones dans la couche cachée), mais nous avons fixé la structure présentée précédemment. La mise en compétition de plusieurs réseaux initialisés aléatoirement nous a permis de retenir le plus performant sans connaître sa complexité.

Induction d'Arbres de Décisions

Quelques définitions Cette méthode est aussi connue sous le terme de méthode de segmentation [107], car elle cherche à résoudre les problèmes de discrimination ou de régression en segmentant de façon progressive l'échantillon de façon à construire un arbre de décision binaire.

Les techniques d'induction d'arbres de décision essaient d'expliquer une variable y en recherchant quelle variable explicative F_j permet d'obtenir la meilleure performance lorsque l'on essaie d'optimiser un critère en recherchant un seuil qui permette de séparer les données d'apprentissage en 2 sous ensembles (2 segments pour la variable d'étude). La technique itérative continue dans les 2 sous ensembles créés. Le résultat de l'itération est généralement représenté sous la forme d'un arbre de décision qui en facilite la compréhension.

Algorithme retenu Pour notre étude, nous avons utilisé l'algorithme C4.5 [116]⁴, en utilisant le critère de GainRatio, fondé sur le calcul de plusieurs critères.

A chaque itération, le calcul est réalisé sur l'ensemble de données $\mathcal{S}$. Le GainRatio est défini par :

$$\text{GainRatio}(\mathcal{S}) = \text{Gain}(\mathcal{S}) / \text{SplitInfo}(\mathcal{S}) \quad (5.25)$$

Qui nécessite le calcul du Gain ($\sharp\{S\}$ correspond au cardinal de l'ensemble $\mathcal{S}$) :

$$\text{Gain}(\mathcal{S}) = \text{Info}(\mathcal{S}) - \sum_{k=1}^n \frac{\sharp\{\mathcal{S}_k\}}{\sharp\{\mathcal{S}\}} \text{Info}(\mathcal{S}_k) \quad (5.26)$$

Et du critère de séparabilité défini par :

$$\text{SplitInfo}(\mathcal{S}) = - \sum_{k=1}^n \frac{\sharp\{\mathcal{S}_k\}}{\sharp\{\mathcal{S}\}} \ln \left(\frac{\sharp\{\mathcal{S}_k\}}{\sharp\{\mathcal{S}\}} \right) \quad (5.27)$$

Avec $\mathcal{S}_k$ l'ensemble défini par la k ème segmentation de l'ensemble $\mathcal{S}$ sur la caractéristique F_j évaluée. Pour les attributs quantitatifs continus de notre étude, la segmentation en sous ensemble $\mathcal{S}_k$ a été réalisé suivant n_j valeurs seuils. Celles-ci ont été obtenues en prenant les valeurs intermédiaires des différentes réalisations ordonnées, pour chaque caractéristique présente dans l'ensemble d'apprentissage.

Ces différentes fonctions sont basées sur le calcul de la fonction Entropie ou Information (Info) :

$$\text{Info}(\mathcal{S}) = - \sum_{i=1}^C P(\omega_i) \ln (P(\omega_i)) \quad (5.28)$$

Avec $P(\omega_i)$ correspondant à la probabilité a priori d'appartenir à la classe ω_i sur l'ensemble $\mathcal{S}$ ⁵.

Pour chaque itération, la valeur seuil qui maximise le GainRatio de chaque caractéristique est retenue. La caractéristique obtenant le GainRatio maximal est retenue. Un critère d'arrêt correspondant à un minimum d'échantillons dans l'ensemble $\mathcal{S}$ permet de stopper la segmentation. La classe majoritaire sur ce minimum d'échantillons est attribuée à ce segment.

La classification ainsi réalisée correspond alors à une séparation des classes par des hyperplans.

Classifieurs Flous

De façon à obtenir une information graduelle sur l'appartenance des observations à chaque classe, nous avons utilisé le classifieur des k plus proches voisins flous (kPPV flous) correspondant à la version floue de celui utilisé dans un cas net⁶.

⁴L'implantation de l'algorithme proposé par Quinlan sous Matlab a fait l'objet des travaux de Lukas Zoubek [117] dont j'ai assuré l'encadrement.

⁵On peut voir cette mesure comme une mesure d'impureté sur l'ensemble $\mathcal{S}$: la fonction est minimale pour un ensemble contenant des échantillons d'une seule classe ($P(\omega_i) = 1$, $\text{Info}(\mathcal{S}) = 0$), le mélange est homogène; elle est maximale pour un ensemble contenant autant d'échantillons dans une classe et dans l'autre pour un problème à deux classes par exemple ($P(\omega_i) = 1/2$, $\text{Info}(\mathcal{S}) = 1$), le mélange est inhomogène.

⁶Ce travail a fait l'objet du travail de DEA de Grégory Brouillet [118], dont j'ai assuré l'encadrement.

Algorithme retenu On attribue une fonction d'appartenance floue à chacun des points de l'espace d'apprentissage [119] et on stocke celles-ci dans le classifieur. On calcule une nouvelle fonction d'appartenance pour chaque nouveau point à classer en pondérant suivant l'influence des points voisins stockés.

Pour chaque point constituant la base d'apprentissage, on utilise l'initialisation des fonctions d'appartenance par la technique des K_{init} PPV : la fonction d'appartenance à la classe ω_j pour chaque $\mathbf{x}$ de l'ensemble d'apprentissage $\mathcal{A}$, de classe ω_i , se fait suivant la répartition de ses K_{init} plus proches voisins dans $\mathcal{A}$ suivant l'équation :

$$u_{init,\omega_j}(\mathbf{x}) = \begin{cases} 0.51 + 0.49 \left(\frac{n_{\omega_j}}{K_{init}} \right), & \text{si } \omega_j = \omega_i \\ 0.49 \left(\frac{n_{\omega_j}}{K_{init}} \right), & \text{si } \omega_j \neq \omega_i. \end{cases} \quad (5.29)$$

n_{ω_j} correspond aux nombres de voisins trouvés qui appartiennent à la classe ω_j . u_{ω_j} correspond à la fonction d'appartenance définie pour les ensembles flous. La classification est alors réalisée en utilisant l'algorithme suivant, proche de celui des kPPV

kPPV flous

Début

Soit $\mathbf{x}$ de classe inconnue

Soit k le nombre de voisins recherchés ($1 \leq k \leq n$)

Trouver les kPPV de $\mathbf{x}$

Pour i variant de 1 à C

Calculer $u_{\omega_i}(\mathbf{x})$ suivant l'équation 5.30

Fin Pour

Fin

L'équation définissant la fonction d'appartenance de $\mathbf{x}$ à la classe ω_i est définie par l'équation :

$$u_{\omega_i}(\mathbf{x}) = \frac{\sum_{j=1}^k u_{init,\omega_i}(\mathbf{x}_j) \left(1/d(\mathbf{x}, \mathbf{x}_j)^{\frac{2}{(m-1)}} \right)}{\sum_{j=1}^k \left(1/d(\mathbf{x}, \mathbf{x}_j)^{\frac{2}{(m-1)}} \right)} \quad (5.30)$$

Où m représente un paramètre de pondération sur la proximité des points ($m > 1$). Pour m proche de 1, la pondération est grande pour les plus proches voisins et faible pour les voisins éloignés. Pour m grand, la pondération des voisins plus distants est prise en compte. L'influence du choix de m peut être évaluée Fig. 5.5.

L'attribution d'une classe se fait sur la fonction d'appartenance maximale. Le résultat d'une telle classification se rapproche d'un classifieur par estimateur de Parzen, de par la nature des noyaux générés.

5.3 Techniques de validation

Suivant la méthode utilisée pour générer l'ensemble d'apprentissage et l'ensemble de test et suivant le critère retenu pour évaluer les performances de la classification, on peut

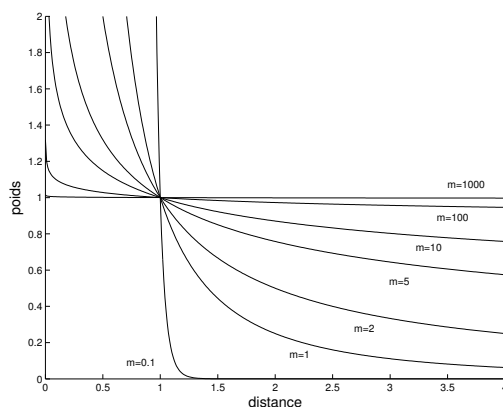


FIG. 5.5 – Paramètre de pondération des kPPV flous : influence de pondération suivant la valeur du paramètre m suivant la distance du voisin pour le calcul de la fonction d'appartenance. On définit une sorte de noyau paramétré par m .

obtenir des résultats différents, qui peuvent conduire à l'acceptation d'un modèle alors que celui-ci a été évalué dans des conditions impropres. Nous nous sommes donc penchés sur cette étape de validation qui nous semblait importante et critique (c'est en particulier, sur ce point que nous critiquons les résultats obtenus par Glass dans la partie IV).

Le résultat de notre recherche montre que de nombreuses techniques de validation peuvent être utilisées suivant la nature du problème posé et la pertinence des résultats recherchés [120, 121, 122, 109, 33, 107, 26]. Nous présentons ici les techniques que nous avons retenues. Nous en proposons d'autres, pour référence, en annexe C.

5.3.1 Quelques définitions

Tableaux de Contingence, Matrices de confusion

On peut comparer les classes obtenues après apprentissage, à celles désirées, en dressant un tableau de contingence (dit aussi de dépendance, ou tableau croisé) des différentes observations appariées, matrice $C \times C$, dont l'élément générique donne le nombre d'exemples de la classe i qui ont été classés dans la classe j . On parle plus souvent en apprentissage artificiel de matrice de confusion (angl. *confusion matrix*).

La matrice de concordance pour un cas à 3 classes est représentée Tab. 5.1 et permet une généralisation au cas d'un problème multiclassé. Dans cette matrice, on note n le nombre total d'observations appariées. Il est commode de diviser l'ensemble des éléments de la matrice par n de façon à obtenir une matrice de concordance exprimée en pourcentage de façon à comparer plus facilement deux matrices construites avec un nombre d'observations différentes. C'est en général cette méthode que nous utiliserons par la suite.

Erreur, risque empirique, risque réel

Pour un problème à plusieurs classes, les opérateurs de performances décrits pour un cas à deux classes décrits en annexe C sont difficilement applicables⁷. On préférera alors

⁷On peut encore utiliser des statistiques pour évaluer les performances de classement et comparer les classifications : coefficient Tau de concordance de Kendall, coefficient Kappa de Cohen, test de χ^2 (angl.

TAB. 5.1 – Matrice de confusion (cas 3 classes)

		Machine (prédiction du modèle)		
Expert (sortie réelle)		ω_1	ω_2	ω_3
	ω_1	M_{11}	M_{12}	M_{13}
	ω_2	M_{21}	M_{22}	M_{23}
	ω_3	M_{31}	M_{32}	M_{33}
		$\sum_{i=1}^3 M_{i1}$	$\sum_{i=1}^3 M_{i2}$	$\sum_{i=1}^3 M_{i3}$
		$\sum_{i=1}^3 \sum_{j=1}^3 M_{ij} = n$		

utiliser la concordance (angl. *accuracy*, Acc) entre l'expert et la machine ou encore l'erreur de classement Err (plus rigoureusement, taux d'erreur), définis par :

$$Acc = \frac{1}{n} \sum_{i=j} M_{ij} \quad (5.31)$$

$$Err = \frac{1}{n} \sum_{i \neq j} M_{ij} = 1 - Acc \quad (5.32)$$

n étant le nombre d'échantillons sur lequel est effectué la classification.

Le risque empirique est la perte moyenne mesurée sur l'échantillon d'apprentissage $\mathcal{A}$:

$$R_{emp}(h) = \frac{1}{m} \sum_{i=1}^m l(\mathbf{u}_i, h(\mathbf{x}_i)) \quad (5.33)$$

En considérant une fonction de coût $\{0, 1\}$ on obtient :

$$l(\mathbf{u}_i, h(\mathbf{x}_i)) = \begin{cases} 0 & \text{si } \mathbf{u}_i = h(\mathbf{x}_i) \\ 1 & \text{si } \mathbf{u}_i \neq h(\mathbf{x}_i) \end{cases} \quad (5.34)$$

Pour notre problème de classification, le risque empirique est alors donné, en reprenant les notations de la matrice de confusion, par :

$$\hat{R}_{emp}(h) = \frac{1}{m} \sum_{i \neq j} M_{ij} \quad (5.35)$$

m étant la taille de l'échantillon d'apprentissage. Cette estimation correspond à l'erreur obtenue sur l'ensemble d'apprentissage (plus rigoureusement, taux d'erreur apparent).

De la même façon, une estimation du risque réel est donnée par :

$$\hat{R}_{réel}(h) = \frac{1}{t} \sum_{i \neq j} M_{ij} \quad (5.36)$$

t étant la taille de l'échantillon de test. Cette estimation correspond à l'erreur obtenue sur l'ensemble de test (plus rigoureusement taux d'erreur de reclassement).

goodness of fit), test de McNemar [26, 123].

5.3.2 Cas de notre étude

Dans notre étude, nous avons été confronté à un problème comportant plusieurs classes ($C = 4$). Nous avons systématiquement divisé l'ensemble des exemples disponibles en un ensemble d'apprentissage et un ensemble de test de façon à obtenir une estimation de l'erreur empirique et une estimation de l'erreur réelle. L'ensemble d'apprentissage a systématiquement été stratifié de façon à obtenir une répartition équiprobable des classes ($P(\omega_i)=1/C$). Ceci a été réalisé de façon à ne pas privilégier une classe par rapport à une autre.

Choix des indicateurs de performance

Nous avons décidé de retenir comme indicateurs de performance, l'erreur apparente et l'erreur de reclassement, car ils correspondent à des estimations simples du risque empirique R_{emp} et du risque réel $R_{réel}$:

- $err_{\mathcal{A}}$: erreur pour une fonction de coût $\{0, 1\}$ obtenue entre le classement de l'apprenant sur l'ensemble d'apprentissage $\mathcal{A}$. Elle correspond au nombre d'échantillons mal classés rapporté au nombre total d'échantillons de $\mathcal{A}$.
- $err_{\mathcal{T}}$: obtenue de la même façon sur l'ensemble de test $\mathcal{T}$.

L'avantage de ces indicateurs réside dans leur simplicité de calcul et de compréhension. De plus, lorsqu'ils sont comparés l'un et l'autre, ils permettent de se rendre compte du comportement du classifieur évalué, de son potentiel à obtenir une généralisation ($err_{\mathcal{A}}$ et $err_{\mathcal{T}}$ confondus) ou son potentiel à réaliser un surapprentissage (apprentissage par cœur) ($err_{\mathcal{A}}$ et $err_{\mathcal{T}}$ distants).

Choix des ensembles d'apprentissage et de test

Nous avons alors porté notre attention sur la génération des ensembles d'apprentissage et des ensembles de test qui nous paraissaient un point critique de toutes induction et évaluation par apprentissage supervisé. Notre étude regroupant à la fois des données provenant d'individus différents (de différents animaux) et d'instantanés différents (données du même animal à différents instantanés de l'expérimentation) et l'objet de la thèse étant de trouver des modèles indépendants des individus (l'appareil portable appris sur un certain nombre d'individus doit être applicable sur n'importe quel individu nouveau), nous avons utilisé une technique de leave one (animal) out pour évaluer une performance indépendante des individus présents dans l'ensemble d'apprentissage. Une explication de la méthode de la technique de leave one out est fournie en annexe C dont on résume le principe appliqué à notre étude (les individus correspondant aux animaux de l'étude) : Un ensemble d'apprentissage est créé en utilisant les observations provenant de $N - 1$ individus en excluant les données de l'individu i . Un rééchantillonnage est effectué sur cet ensemble $\mathcal{A}_i$ de façon à obtenir un ensemble équiprobable pour chaque classe par tirage aléatoire dans cet ensemble. L'apprentissage est alors réalisé et l'erreur d'apprentissage $err_{\mathcal{A}_i}$ évaluée. L'ensemble de test $\mathcal{T}_i$ est constitué des observations de l'individu exclu et permet de calculer l'erreur de test $err_{\mathcal{T}_i}$. Le processus est répété sur les N individus, en changeant l'indice i de l'individu exclu. Un calcul de l'erreur moyenne en apprentissage

et de l'erreur moyenne en test est alors réalisé :

$$err_{\mathcal{A}} = \frac{1}{N} \sum_{i=1}^N err_{\mathcal{A}_i} \quad (5.37)$$

$$err_{\mathcal{T}} = \frac{1}{N} \sum_{i=1}^N err_{\mathcal{T}_i} \quad (5.38)$$

Par abus de notation, dans le cas du leave one (animal) out, nous noterons ces erreurs $err_{\mathcal{A}}$ et $err_{\mathcal{T}}$.

Cette technique de rééchantillonnage permet d'obtenir une estimation d'un risque réel associé à $err_{\mathcal{T}}$ très pessimiste, puisque l'estimation est réalisée sur des données provenant d'individus n'ayant pas servi pour la création du modèle. C'est une analyse dans le pire cas.

Sélection de sous ensembles de caractéristiques

Le nombre de sous ensembles que l'on peut obtenir à partir d'un ensemble constitué de d composantes est égal à 2^d . La recherche du sous espace constitué des caractéristiques les plus performantes pour notre problème de classification ne peut être effectué exhaustivement sans obtenir une explosion combinatoire lorsque d devient important. De façon à minimiser le nombre d'opérations nécessaires pour effectuer une recherche d'un sous espace minimal, plusieurs méthodes peuvent être envisagées.

Pour notre étude, nous avons effectué une sélection ascendante séquentielle (angl. *sequential forward selection*, SFS) des caractéristiques les plus pertinentes en utilisant comme critère la minimisation de l'erreur de test. La technique repose sur l'algorithme itératif suivant :

Sequential forward selection

Début Soit $I_r = \{\emptyset\}$, ensemble des indices des caractéristiques retenues.

Soit $I_d = \{1, \dots, d\}$, ensemble des indices des caractéristiques disponibles.

Tant que $I_d \neq \{\emptyset\}$

Pour chaque indice i disponible dans I_d ,

Faire

$I_c(i) = I_r \cup i$, ensemble courant des indices des caractéristiques à évaluer

Créer $S_{I_c(i)}$, l'ensemble des observations repérées par sélection des caractéristiques indicées par $I_c(i)$

Calculer le critère $\Xi(i)$ associé à l'ensemble $S_{I_c(i)}$

Fin Pour

Choisir l'indice i_r qui optimise le critère retenu

Mettre à jour les ensembles :

$I_r = I_r \cup i_r$

$I_d = I_d \setminus i_r$

Fin Tant que Fin

Au final, on récupère l'ensemble I_r ensemble ordonné des indices des caractéristiques retenues par SFS. L'évaluation se fait donc progressivement dans un espace à 1 dimension $\{i_{r1}\}$, 2 dimensions $\{i_{r1}, i_{r2}\}$, ..., d dimensions $\{i_{r1}, i_{r2}, \dots, i_{rd}\}$, en retenant les caractéristiques les plus performantes. Cette sélection ne nécessite que l'évaluation de $\frac{(d+1)d}{2}$ sous ensembles de caractéristiques.

Amélioration de l'apprentissage

Confronté à certaines données incohérentes dans les régions frontalières de classification, nous avons essayé plusieurs méthodes pour essayer d'améliorer l'apprentissage et ainsi obtenir de meilleures performances.

Rejet Dans un premier temps nous avons essayé de rejeter les données dont l'attribution à une classe n'était pas franche car celle-ci était partagée entre plusieurs classes (rejet en ambiguïté), ou dont l'attribution à une classe n'était pas fiable car trop éloigné des observations ayant servies pour l'apprentissage (rejet en distance). Ces notions de rejet sont expliqués dans [33, 110]. Nous avons appliqué ces rejets avec un classifieur quadratique, en définissant deux nouvelles classes : ω_0 correspond à la classe de rejet en ambiguïté et ω_{-1} correspondant à la classe de rejet en distance. Les nouvelles règles de classification pour un classifieur utilisant la règle de Bayes sont alors les suivantes :

$$\begin{cases} \mathbf{x} \text{ appartient à } \omega_i & \text{si } \max_{i=1, \dots, C} P(\omega_i | \mathbf{x}) \geq 1 - C_r \\ \text{sinon } \mathbf{x} \text{ est rejeté en ambiguïté} & (\mathbf{x} \text{ appartient à } \omega_0) \end{cases} \quad (5.39)$$

avec C_r coût de rejet en ambiguïté.

$$\begin{cases} \mathbf{x} \text{ appartient à } \omega_i & \text{si } f(\mathbf{x}) \geq C_d \\ \text{sinon } \mathbf{x} \text{ est rejeté en distance} & (\mathbf{x} \text{ appartient à } \omega_{-1}) \end{cases} \quad (5.40)$$

avec C_d coût de rejet en distance.

L'évaluation de la performance est alors effectuée sur les données qui n'ont pas été rejetées dans l'espoir d'améliorer les performances du classifieur.

Nettoyage des données Nous avons essayé d'éliminer les exemples qui nous paraissaient douteux selon différentes méthodes, de façon à obtenir des ensembles de données plus fiables. Nous avons procédé de deux façons : la première s'appuie sur une élimination des données instables du scoring expert (changement de décision de l'expert pour deux classifications successives), la seconde s'appuie sur une condensation des données servant à la création de l'ensemble d'apprentissage.

Élimination des transitions de l'expert Dans cette technique, on utilise une connaissance temporelle a priori pour éliminer les données qui nous paraissent peu robustes voire nuisibles pour la création de l'ensemble d'apprentissage : On enlève les éléments de la base de données correspondant à des classifications successives de l'expert qui ne sont pas identiques, c.-à-d. telles que $u(k) \neq u(k-1)$. Ces observations correspondent aux transitions entre classes.

Nettoyage par 1 PPV et condensation Nous avons utilisé une méthode de nettoyage constituée de deux étapes tirée de [26], afin d'éliminer le plus grand nombre d'exemples non stables par cette méthode. Le premier algorithme consiste à éliminer les exemples qui ne montrent pas de cohérence de classes avec leur voisinage. Pour cela, il utilise l'algorithme suivant :

Algorithme de Nettoyage par 1PPV

Début

Diviser aléatoirement l'ensemble d'apprentissage $\mathcal{A}$ en deux sous-ensemble $\mathcal{A}_1$ et $\mathcal{A}_2$

Tant que la stabilisation de $\mathcal{A}_1$ et $\mathcal{A}_2$ n'est pas réalisée **faire**

Classer tous les points de $\mathcal{A}_1$ sur $\mathcal{A}_2$ par la règle du 1PPV

Éliminer de $\mathcal{A}_1$ tous les points dont la classe n'est pas la même que celle de leur plus proche voisin dans $\mathcal{A}_2$

Classer tous les points de $\mathcal{A}_2$ sur le nouveau $\mathcal{A}_1$ par la règle du 1PPV

Éliminer de $\mathcal{A}_2$ tous les points dont la classe n'est pas la même que celle de leur plus proche voisin dans $\mathcal{A}_1$

Fin tant que

L'ensemble d'apprentissage nettoyé est composé de $\mathcal{A}_1 \cup \mathcal{A}_2$

Fin

Le second algorithme permet de ne garder que les exemples stables suivant la procédure définie ci-dessous et donc éliminer ceux qui n'apportent rien à la classification par 1PPV :

Algorithme de Condensation

Début

Ordonner les m exemples d'apprentissage de $\mathbf{x}_1$ à $\mathbf{x}_m$

Initialiser $\mathcal{S}$ par $\mathbf{x}_1$ et $\mathcal{G}$ par $\mathbf{x}_2$ à $\mathbf{x}_m$

Tant que $\mathcal{S}$ et $\mathcal{G}$ ne sont pas stabilisés **faire**

Pour chaque point g_i de $\mathcal{G}$ **faire**

Si le 1PPV de g_i dans $\mathcal{S}$ n'a pas la même classe que g_i **alors**

 Enlever g_i de $\mathcal{G}$ et le mettre dans $\mathcal{S}$

Fin Si

Fin Pour

Fin Tant que

L'ensemble d'apprentissage condensé est $\mathcal{S}$

Fin

5.3.3 Analyse statistique

Pour chaque analyse statistique, nous avons déterminé le modèle obtenant l'erreur en test moyenne $err_{\mathcal{T}}$ minimale. Celui-ci a servi de référence pour une comparaison avec les performances obtenues sur les autres résultats. Une comparaison de l'identité des distributions des erreurs $err_{\mathcal{T}}$ appariées, obtenues sur chaque individu (animal) a été réalisée.

TAB. 5.2 – Classifieurs utilisées dans l'étude et leurs abréviations

Nom du classifieur	Abréviation (paramètre)
Règle de Bayes	Bayes
paramétrique	
linéaire	Lin
quadratique	Quad
naïf	Naïf
non paramétrique	
k plus proches voisins	kPPV (k)
estimateur de Parzen	Parzen (α)
Aléatoire	Aléa
Réseau de neurone	RN
perceptron multi couche	PMC
Arbre de décision	
C4.5	C4.5
Flou	
k plus proches voisins flou	kPPVf (k, K_{init}, m)

Nous avons donc utilisé un test de Wilcoxon. Ce test non-paramétrique nous paraissait approprié par rapport au nombre peu élevé d'individus testés ($N = 17$). Les tests ont été effectués au seuil de signification $p = 0.05$. Les modèles ont été jugés significativement différents pour $p < 0.05$. Les modèles présentant une différence significative avec le modèle obtenant une erreur moyenne en test err_T la plus faible possible, ont été annotés du symbole † sur les graphiques.

5.4 Conclusion du chapitre

La classification des caractéristiques extraites sur les signaux bruts constitue la partie centrale de notre étude. Pour réaliser un apprentissage supervisé de la classe attribuée par l'expert et évaluer cet apprentissage dans des conditions réalistes, nous avons présenté les différentes notions liées à l'apprentissage artificiel. Après une description des choix réalisés pour la représentation des données, les transformations effectuées pour obtenir un espace des formes le mieux adapté, nous avons abordé les notions d'apprentissage supervisé nécessaire à la compréhension des modèles de classification que nous avons entraînés (les classifieurs utilisés reliés à leur espace de représentation et à leur données d'apprentissage). Nous avons alors présenté les classifieurs utilisés en spécifiant leurs propriétés. Un résumé des différents classifieurs utilisés est disponible Tab. 5.2.

Nous avons développé les choix réalisés sur l'évaluation des résultats et les problèmes d'estimations associées. Cette évaluation nous a servi de critère pour réaliser une sélection de modèles. Enfin nous avons présenté quelques techniques pour essayer d'optimiser les performances par rejet ou élimination de certaines données (élimination des classes de l'expert dans les régions de transitions, nettoyage des données par condensation).

Pour la plupart de ces techniques, nous avons donc essayé de respecter au mieux les points suivants :

- Tirages indépendants et identiquement distribués (iid).
- Indépendance temporelle.
- Indépendance des individus.
- Équiprobabilité des classes.

Troisième partie

Résultats

Chapitre 6

Présentation des données

Le physiologiste et le médecin doivent donc toujours considérer en même temps les organismes dans leur ensemble et dans leurs détails, sans jamais perdre de vue les conditions spéciales de tous les phénomènes particuliers dont la résultante constitue l'individu. Toutefois les faits particuliers ne sont jamais scientifiques : la généralisation seule peut constituer la science.

Claude Bernard, Introduction à l'étude de la médecine expérimentale, 1984, Flammarion, Paris, p. 140

Dans ce chapitre, nous présentons les données suivant différentes approches de façon à cerner au mieux celles ci et comprendre l'avantage des méthodes exposées précédemment que nous appliquerons dans le chapitre suivant. La section 6.1 présente les données prises individuellement et les particularités propres à chaque individu. La section 6.2 présente une analyse statistique monodimensionnelle fréquemment utilisée et ses avantages et désavantages. Enfin, la section 6.3 présente les données en les resituant dans leur contexte multidimensionnel et permet de mieux appréhender les effets des classifications envisagées et des différents modes de fonctionnement de l'organisme.

6.1 Analyse de la population

Un premier bilan de la population permet de regrouper les données propres à chaque individu et les particularités obtenues lors des expérimentations.

L'analyse des caractéristiques moyennes des animaux au repos et après une perte de sang de 50% du volume sanguin présentée Tab. 6.1 et Tab. 6.2 permet de mettre en évidence les individus qui ressortent de la population (indiqué en gras dans ces tableaux).

TAB. 6.1 – Valeurs de repos

Animal	Nom (Date)	Poids	PAM	PAD	PAS	ΔPA	T_{int}	E_tCO_2
#1	20020409	19	75 (12)	43 (21)	104 (10)	61 (18)	37.7 (0.2)	- (-)
#2	20020411	19	82 (2)	62 (4)	98 (1)	36 (4)	37.4 (0.0)	- (-)
#3	20020423	19.3	98 (12)	70 (52)	116 (14)	46 (62)	36.4 (0.0)	48 (2)
#4	20020425	17.1	80 (2)	61 (3)	92 (1)	31 (3)	37.2 (0.2)	57 (8)
#5	20020516	18.8	68 (2)	45 (12)	88 (13)	43 (25)	36.2 (0.2)	38 (2)
#6	20020521	17.6	89 (2)	65 (2)	117 (2)	52 (2)	36.7 (0.1)	51 (2)
#7	20020523	17.2	80 (2)	60 (2)	100 (3)	40 (2)	36.5 (0.3)	52 (1)
#8	20020528	14.3	70 (3)	46 (2)	94 (4)	48 (4)	35.2 (0.0)	46 (1)
#9	20020627	20.7	73 (5)	51 (14)	93 (14)	42 (22)	34 (0.1)	44 (2)
#10	20020702	21	74 (3)	51 (3)	94 (3)	43 (3)	35.9 (0.0)	47 (1)
#11	20020813	20.9	53 (20)	35 (30)	73 (15)	37 (28)	35.9 (0.1)	47 (1)
#12	20030526	29.6	85 (10)	57 (16)	108 (11)	51 (19)	37.4 (0.1)	62 (18)
#13	20030527	31	69 (7)	52 (21)	81 (3)	29 (21)	36.1 (0.0)	48 (4)
#14	20030617	20.8	89 (4)	67 (5)	110 (5)	44 (4)	36.7 (0.0)	53 (2)
#15	20030619	20.5	63 (0)	43 (0)	83 (1)	40 (1)	35.4 (0.0)	48 (0)
#16	20030625	18.8	74 (4)	55 (6)	97 (9)	42 (13)	36.8 (0.1)	41 (1)
#17	20030702	19.1	67 (2)	47 (2)	86 (1)	39 (2)	36.1 (0.1)	41 (1)
#18	20030708	18.6	73 (5)	56 (5)	87 (3)	31 (4)	34.5 (0.0)	42 (1)
#19	20030709	21.4	73 (4)	55 (4)	91 (3)	37 (2)	36.3 (0.1)	44 (1)
$\mu(\sigma)$:		20.3 (3.9)	76 (10)	54 (9.2)	95 (12)	42 (7.9)	36 (0.96)	48 (6.1)
Moyenne (Écart-type)								

Animal	FC	RR_{Abd}	RR_{Tho}	SpO_2	T_{cut}	T_{amb}
#1	48 (4)	12.1 (0.6)	12 (0.5)	98.9 (0.4)	36.6 (1.1)	- (-)
#2	87 (4)	29.3 (0.9)	29.4 (0.8)	96.3 (2.1)	35.2 (0)	- (-)
#3	60 (8)	28.8 (10.8)	27.4 (7.6)	96.1 (1.7)	34.4 (0.1)	- (-)
#4	59 (10)	22.2 (7.4)	24.1 (5.4)	81.7 (10.2)	35.6 (0.6)	- (-)
#5	65 (4)	12.6 (4.7)	15.9 (2.2)	95.5 (1.3)	34.7 (0.8)	- (-)
#6	36 (2)	10.6 (7.9)	18.2 (0.7)	98.4 (0.4)	- (-)	- (-)
#7	52 (1)	17.4 (3.1)	22.7 (2.8)	95.6 (0.7)	32.8 (0.4)	38.1 (2.7)
#8	46 (2)	7.32 (8.2)	28 (9.6)	93.8 (1.0)	33.3 (2.8)	37.7 (5.5)
#9	83 (21)	11.6 (7.7)	18.7 (3.8)	96.6 (0.4)	32 (0.1)	30.7 (0.4)
#10	98 (8)	15.3 (3.8)	15.1 (4.2)	88.6 (7.6)	33.3 (0.6)	30.5 (0.7)
#11	93 (9)	16.8 (4.3)	19.3 (1.4)	97.2 (0.5)	- (-)	34.6 (1.8)
#12	104 (19)	14 (9.6)	14.2 (8.8)	98.1 (1)	34.7 (0.2)	30.5 (0.5)
#13	91 (9)	21.4 (1.9)	18.7 (3)	97.5 (1)	31 (0.5)	25.4 (1.5)
#14	68 (6)	35.9 (20.6)	36.6 (18)	95.4 (5.0)	35.5 (0.1)	32.1 (0.3)
#15	46 (1)	21.8 (0.5)	19.5 (3.5)	95.9 (0.4)	35 (0.0)	29.9 (1.0)
#16	188 (5)	20.9 (2.6)	20.8 (2.8)	81.2 (19.5)	35.8 (0.1)	30.9 (2.2)
#17	98 (2)	36.4 (8.7)	28.3 (8.1)	95.2 (1.1)	35 (0.1)	29.1 (0.2)
#18	94 (7)	22.8 (10.5)	24.4 (13.8)	77.4 (27.6)	33.3 (0.2)	29.7 (1.1)
#19	103 (5)	18.8 (5.4)	21.7 (4.0)	90.7 (2.2)	35.1 (0.1)	27.2 (1.7)
$\mu(\sigma)$	80 (34)	20 (8.2)	22 (6.1)	93 (6.4)	34 (1.5)	31 (3.7)

TAB. 6.2 – Valeurs a 50%

Animal	<i>PAM</i>	<i>PAD</i>	<i>PAS</i>	ΔPA	T_{int}	E_tCO_2
#1	34 (0)	24 (1)	51 (1)	27 (0)	37.6 (0.0)	- (-)
#2	16 (0)	11 (0)	64 (63)	53 (63)	37.7 (0.0)	- (-)
#3	28 (0)	24 (0)	38 (0)	14 (0)	37 (0.0)	42 (1)
#4	3 (1)	-15 (17)	22 (19)	37 (36)	35.6 (0.1)	14 (16)
#5	49 (10)	26 (4)	84 (9)	58 (6)	37.4 (0.0)	49 (12)
#6	36 (1)	29 (2)	43 (4)	14 (5)	37.6 (0.1)	58 (1)
#7	30 (1)	23 (1)	39 (1)	16 (0)	37.7 (0.0)	54 (1)
#8	23 (1)	18 (1)	29 (1)	11 (2)	35.7 (0.0)	37 (4)
#9	18 (4)	5 (10)	27 (11)	22 (21)	33.1 (0.0)	30 (34)
#10	44 (1)	34 (1)	61 (2)	26 (1)	34.8 (0.0)	50 (0)
#11	35 (2)	32 (2)	39 (1)	7 (1)	35.8 (0.0)	42 (5)
#12	30 (1)	20 (1)	60 (11)	40 (10)	38.5 (0.0)	34 (2)
#13	- (-)	- (-)	- (-)	- (-)	- (-)	- (-)
#14	44 (0)	31 (0)	63 (0)	32 (0)	37.1 (0.0)	37 (2)
#15	35 (1)	28 (1)	54 (1)	26 (1)	35.6 (0.0)	45 (0)
#16	50 (0)	39 (0)	72 (1)	33 (1)	38.4 (0.0)	37 (2)
#17	34 (0)	26 (1)	51 (1)	26 (1)	37 (0.0)	38 (0)
#18	41 (0)	32 (1)	59 (2)	27 (2)	34.7 (0.0)	36 (2)
#19	51 (0)	38 (2)	74 (0)	36 (2)	37 (0.0)	36 (1)
$\mu(\sigma)$	33 (13)	24 (13)	52 (17)	28 (14)	37 (1.4)	40 (10)

Animal	<i>FC</i>	<i>RR_{Abd}</i>	<i>RR_{Tho}</i>	<i>SpO₂</i>	T_{cut}	T_{amb}
#1	161 (3)	12 (0)	12 (1)	99 (0.0)	35.9 (0.1)	- (-)
#2	129 (3)	20 (0)	20 (0)	0.0 (0.0)	33.9 (0.1)	- (-)
#3	171 (2)	14 (1)	14 (1)	0.0 (0.0)	32 (0.1)	- (-)
#4	62 (39)	1 (1)	1 (1)	0.0 (0.0)	32.4 (0.2)	- (-)
#5	161 (16)	2 (2)	1 (1)	17.4 (12.2)	33.9 (0.0)	- (-)
#6	99 (2)	6 (1)	6 (1)	0.0 (0.0)	18.4 (0)	- (-)
#7	128 (6)	16 (1)	15 (1)	0.0 (0.0)	32.7 (0.1)	31.1 (0.1)
#8	171 (3)	12 (2)	13 (3)	96.6 (0.5)	31.3 (0.2)	29.5 (0.2)
#9	127 (23)	0 (1)	1 (1)	24.9 (24.3)	29.1 (0.1)	28.3 (0.1)
#10	174 (1)	6 (1)	6 (1)	8.6 (14.8)	31 (0.0)	27.7 (0.1)
#11	184 (3)	17 (8)	11 (6)	55.7 (50.9)	0 (0)	33 (0.1)
#12	227 (2)	23 (15)	30 (2)	45.8 (37.5)	32.4 (0.1)	27.3 (0.2)
#13	- (-)	- (-)	- (-)	- (-)	- (-)	- (-)
#14	85 (3)	48 (1)	46 (3)	32.4 (45.8)	32.2 (0.0)	29.6 (0.0)
#15	150 (1)	11 (7)	26 (8)	94.7 (0.3)	34.7 (0.0)	30.4 (0.0)
#16	261 (0)	27 (2)	21 (12)	0.0 (0.0)	36.4 (0.0)	33.2 (0.4)
#17	203 (3)	30 (2)	8 (10)	0.0 (0.0)	33.8 (0.1)	27.6 (0.1)
#18	156 (3)	26 (3)	26 (4)	68.5 (28.4)	32.5 (0.0)	28.4 (0.1)
#19	190 (1)	21 (2)	8 (3)	90.8 (0.8)	35 (0.0)	26.9 (0.4)
$\mu(\sigma)$	160.02 (49)	16(12)	15(12)	35 (39)	30 (8.5)	29 (2.1)

Le poids moyen des animaux est de $20 \pm 4 \text{ kg}$, la fréquence cardiaque au repos est de $80 \pm 34 \text{ bpm}$ et la température interne moyenne au repos de $36 \pm 1 \text{ }^\circ\text{C}$. Ces quelques exemples présentent des écart-types importants sur la plupart des caractéristiques, et indiquent que la population est variée.

L’animal #13 est décédé pour une perte de sang correspondant à 45%, c’est pourquoi les valeurs des caractéristiques ne sont pas disponibles à 50% et ont été notées – Tab. 6.2.

Sur les 19 animaux ayant fourni des données exploitables, les individus #6 et #11 ont été exclus de l’analyse de données car la caractéristique T_{cut} n’était pas disponible (capteur hors service). Nous avons exclu la caractéristique E_tCO_2 , les données n’étant pas disponibles sur les individus #1 et #2. Nous avons exclu la caractéristique T_{amb} , les données n’étant pas disponibles sur les individus #1, #2, #3, #4, #5 et #6. Les résultats présentés pour la classification porteront donc sur 17 animaux.

6.2 Analyse monodimensionnelle

6.2.1 Cas moyen

La période de base correspond à la période débutant à $t_0 = -25 \text{ min}$ et finissant à $t_{choc} = 0 \text{ min}$, instant du début de la spoliation. Les valeurs physiologiques moyennes au cours de l’hémorragie contrôlée sont disponibles dans la dernière ligne des tableaux Tab. 6.1 et Tab. 6.2, et résumées Tab. 6.3. Les caractéristiques ont été calculées sur une

TAB. 6.3 – Valeurs physiologiques au cours d’une hémorragie contrôlée (Becq et coll.)

	Données (mean $\pm$ SD)	
	Ligne de base	Hémorragie de 50% (0.49–0.51)
Pression Artérielle Moyenne (mmHg)	75 ± 12	32 ± 14
Pression Artérielle Systolique (mmHg)	95 ± 14	50 ± 22
Pression Artérielle Diastolique (mmHg)	54 ± 19	22 ± 16
Température interne ($^\circ\text{C}$)	36.2 ± 0.9	36.4 ± 1.4
Fréquence Cardiaque (bpm)	80 ± 35	154 ± 51
RR_{Tho} (cpm)	20 ± 11	15 ± 12
RR_{Abd} (cpm)	22 ± 9	13 ± 11
SpO_2 (%)	93 ± 10	35 ± 42
Température cutanée ($^\circ\text{C}$)	34.3 ± 1.6	33.1 ± 1.8

fenêtre temporelle de $\Delta t = 60 \text{ s}$ et les calculs statistiques effectuées en ne retenant que les valeurs disponibles (exclusion des données manquantes) sur les 19 animaux de l’étude. Ces valeurs sont comparables à celles obtenues par d’autres équipes et en particulier par Glass et coll.(cf. Tab. 10.1 au chap. 10).

6.2.2 Évolution moyenne par variable

Dans cette section, nous analysons les évolutions des différentes caractéristiques traitées une à une. Nous pouvons réaliser une moyenne ou un calcul de médiane sur l’ensemble des animaux aux différents instants sur lesquels sont calculés ces caractéristiques pour évaluer l’évolution de ces variables. Comme la durée des enregistrements, fixée par la réactivité de l’organisme des animaux, n’est pas la même pour tous les animaux, nous devons tenir compte d’un alignement des échantillons :

Par rapport au début de l'expérimentation pour mettre en valeur la transition occasionnée par la perte de sang. Le résultat de cet alignement est représenté graphiquement Fig. 6.1 et 6.3. $t = 25 \text{ min}$ correspond au début de la spoliation.

Par rapport à la fin de l'expérimentation pour mettre en évidence les phénomènes observés lors des derniers combats de l'organisme. Le résultat de cet alignement est représenté graphiquement Fig. 6.2 et 6.4. La valeur de référence pour l'alignement, correspondant à la mort de l'animal a été fixé arbitrairement à $t = 240 \text{ min}$ sachant que la durée maximale pour un animal a été de 182 min (#10).

Pour chaque caractéristique étudiée, à chaque instant de l'expérimentation, on peut réaliser une estimation du comportement statistique de la population (ensemble des animaux), par un calcul de moyenne, de médiane ou écart-type. Les données manquantes ne sont pas prises en compte dans ces calculs. Ainsi, pour un alignement par rapport au début de l'expérimentation, le nombre d'animaux entrant dans le calcul va diminuer au cours de l'expérimentation en correspondance avec la perte de l'animal. Inversement, pour un alignement par rapport à la fin de l'expérimentation, le temps est compté à rebours, et plus on remonte avant la fin de l'expérimentation, moins on aura d'animaux. La visualisation de l'effondrement du nombre d'animaux rentrant dans le calcul des paramètres ainsi calculés est représentée Fig. 6.1 a) et Fig. 6.2 a).

Si l'on regarde la caractéristique de fréquence cardiaque, par exemple : En alignant les données par rapport au début de l'expérimentation (Fig. 6.1 l)), la montée de fréquence cardiaque (tachycardie) dès le début de la spoliation ($t_{choc} = 25 \text{ min}$) est mise en évidence ; En alignant les données par rapport à la fin de l'expérimentation (Fig. 6.2 l)), on met en évidence une diminution brutale de la fréquence cardiaque (bradycardie). Les résultats de l'observation sont donc différents selon la méthode d'alignement utilisée.

La forte variabilité provient en partie du nombre de sujets traités ($N = 19$). Elle peut encore être expliquée par le fait qu'on regarde la caractéristique en une seule dimension, c.-à-d., on considère chaque variable comme indépendante des autres variables. Nous verrons par la suite que ceci ne correspond pas à la réalité et que la projection ainsi réalisée, si elle permet d'appréhender les paramètres les plus pertinents, ne permet pas de conclure.

Pour les caractéristiques de pression artérielle (PAM , PAS , PAD , Δ_{PA} , σ_{PA}), on constate un comportement similaire sur les différentes caractéristiques : l'effondrement au cours de l'expérimentation. La lecture du chapitre 3.1, nous indique que nous devrions observer un plateau de pression correspondant à un réflexe sympathique. On peut se demander si l'anesthésie n'a pas été en moyenne trop poussée au cours des différentes expérimentations (effet du propofol sur le réflexe sympathique).

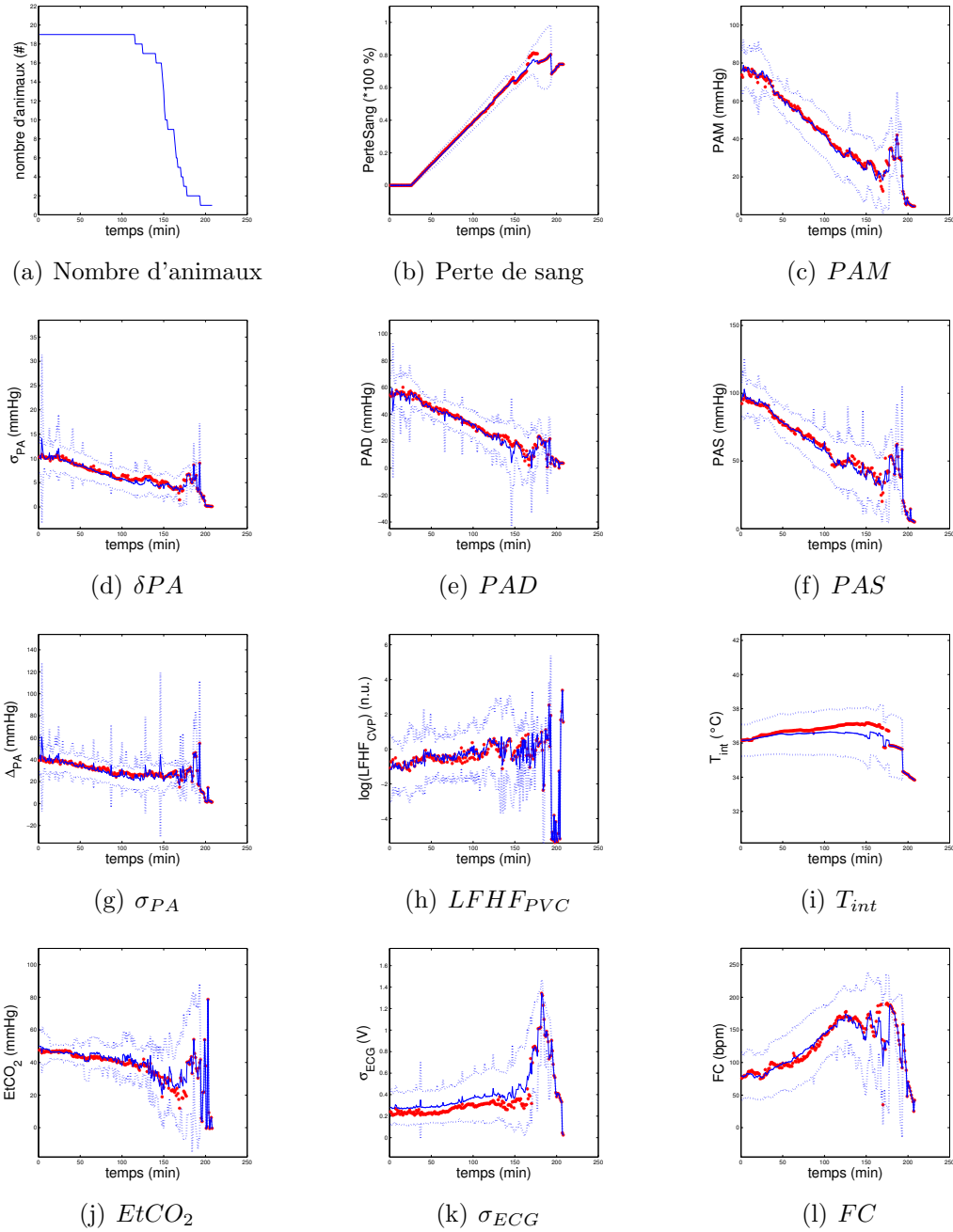


FIG. 6.1 – Évolution moyenne des caractéristiques (Début) : Les données sont alignées par rapport au début de l'expérimentation ($t = 0 \text{ min}$). Le nombre d'animaux entrant en jeu pour le calcul des paramètres statistiques est donné en a) et diminue au cours de l'expérimentation. Les points représentent les moyennes réalisées aux différents instants de l'expérimentation, les astérisques correspondent aux valeurs médianes. Les lignes en pointillés correspondent aux écart-types.

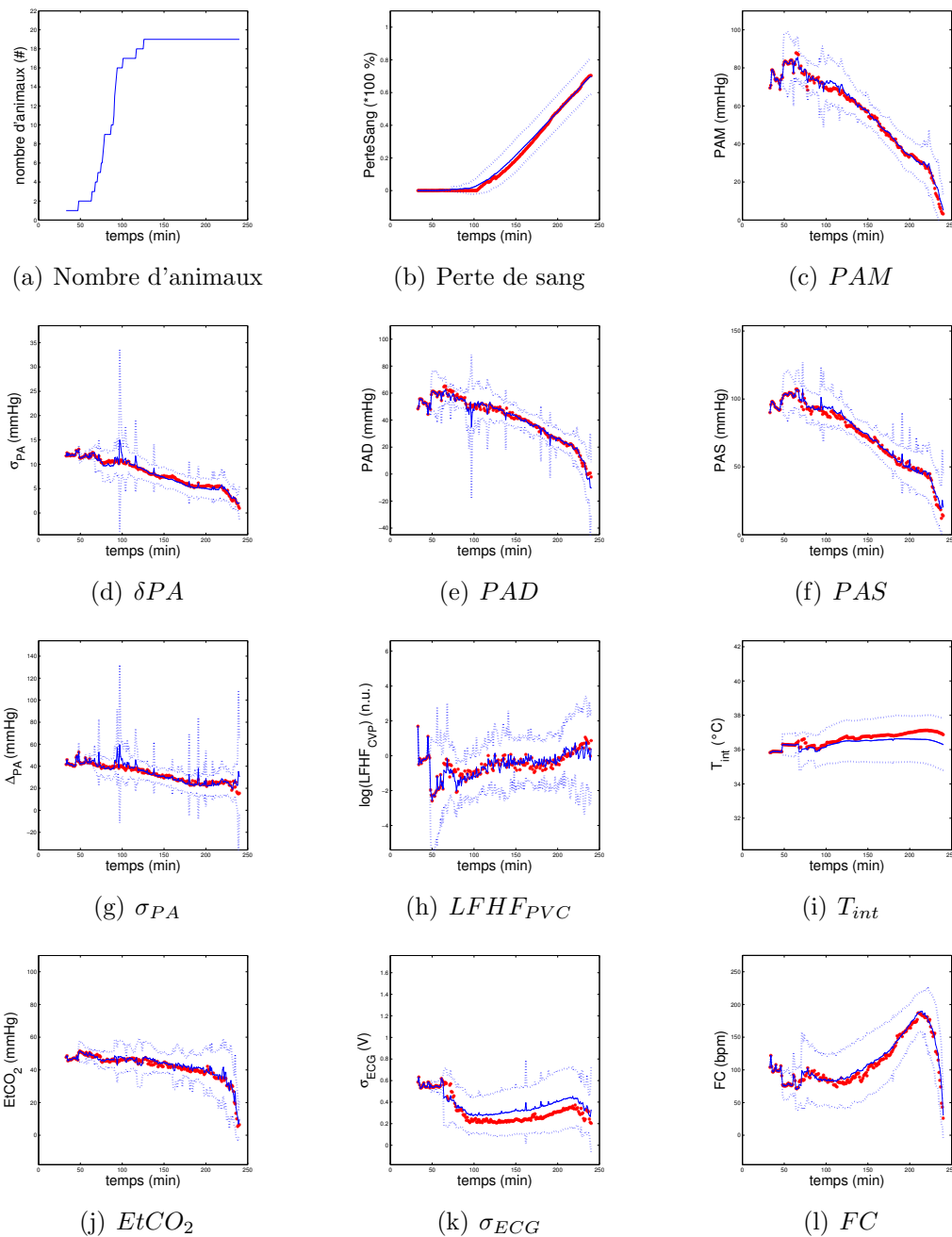


FIG. 6.2 – Évolution moyenne des caractéristiques (Fin) : les données sont alignées par rapport à la fin de l'expérimentation (fixé arbitrairement à $t = 240 \text{ min}$). Les points représentent les moyennes réalisées aux différents instants de l'expérimentation, les astérisques correspondent aux valeurs médianes. Les lignes en pointillés correspondent aux écart-types.

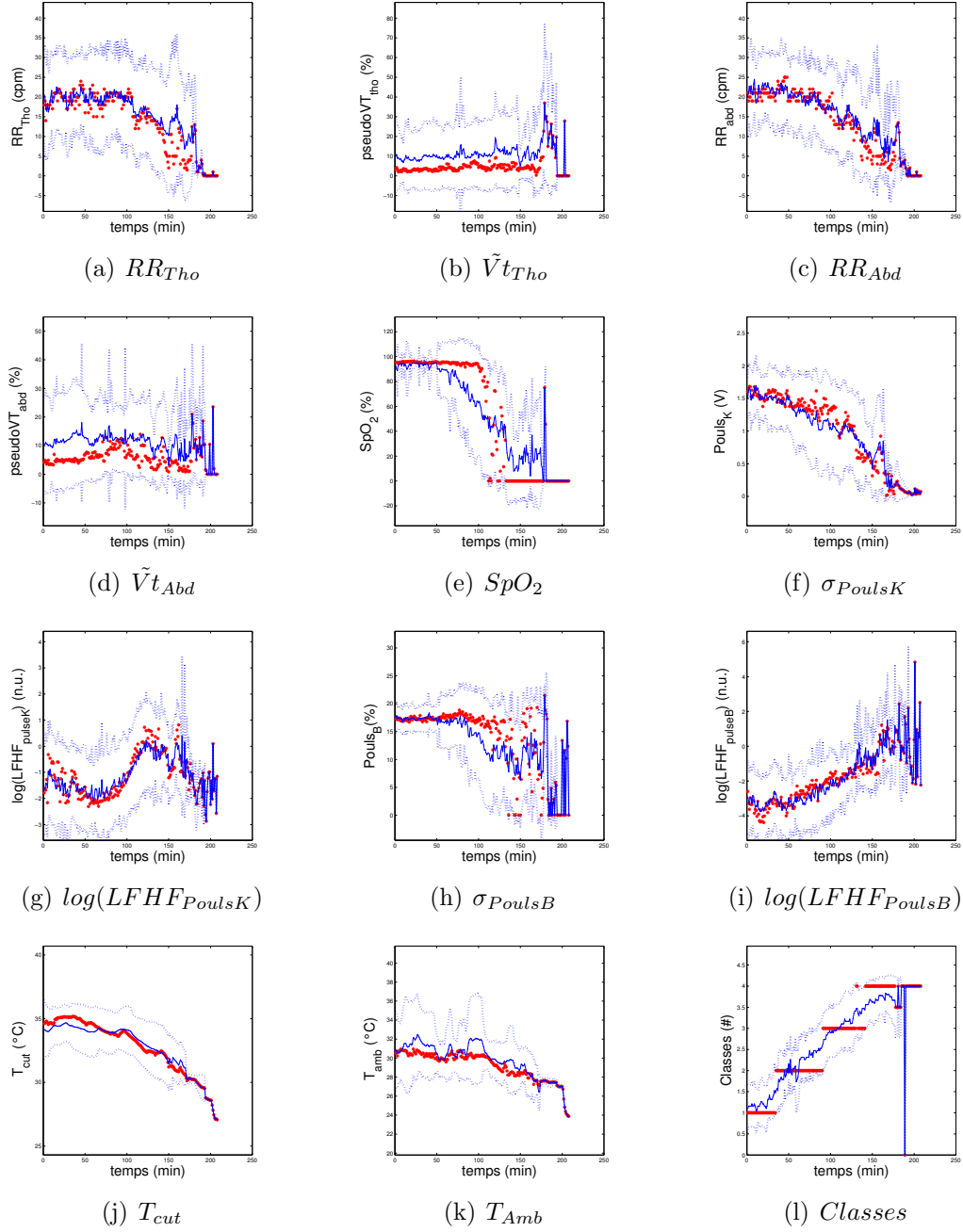


FIG. 6.3 – Évolution moyenne des caractéristiques (Début) : les points représentent les moyennes réalisées aux différents instants de l'expérimentation, les astérisques correspondent aux valeurs médianes. Les lignes en pointillés correspondent aux écart-types.

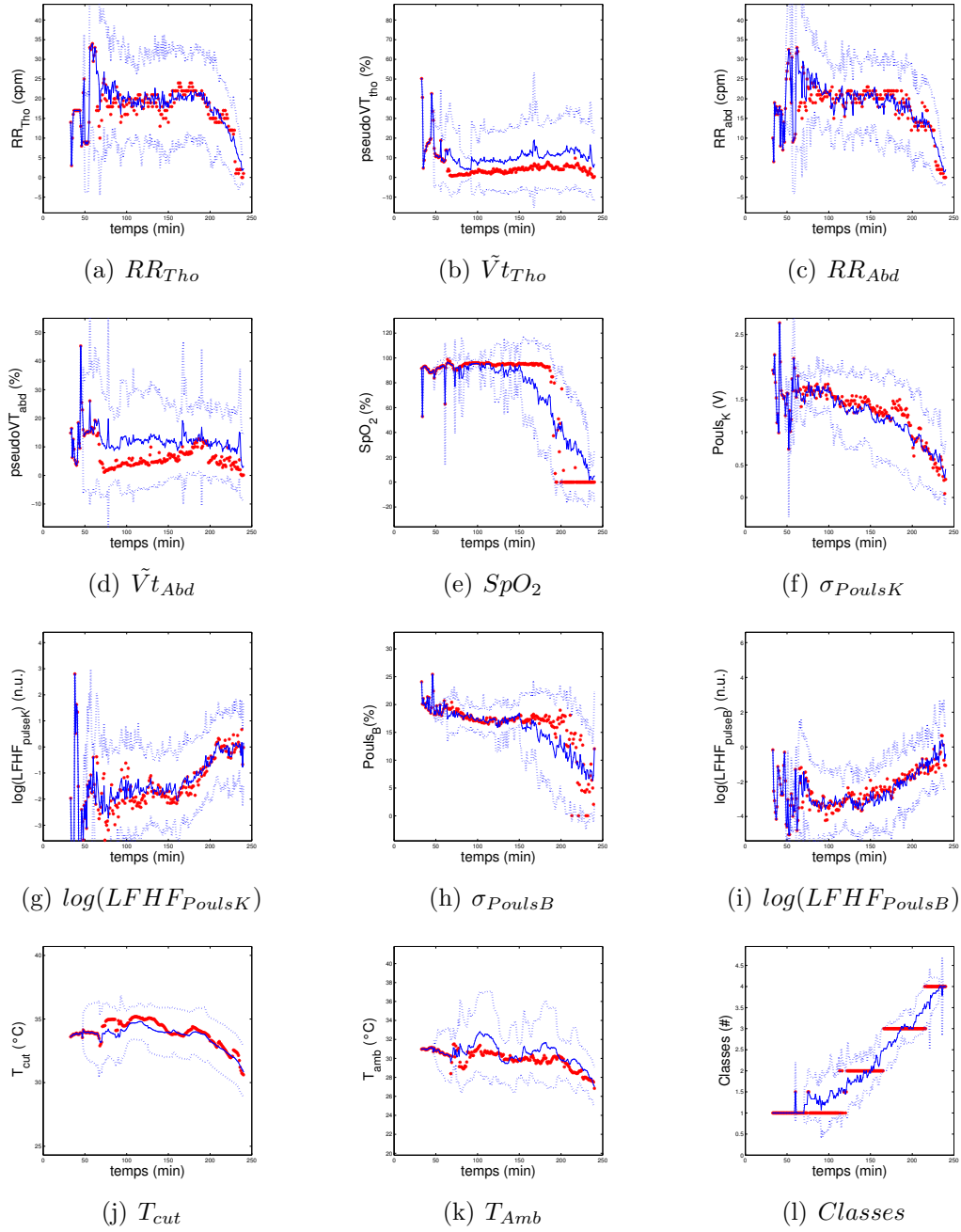


FIG. 6.4 – Évolution moyenne des caractéristiques (Fin) : les points représentent les moyennes réalisées aux différents instants de l'expérimentation, les astérisques correspondent aux valeurs médianes. Les lignes en pointillés correspondent aux écart-types.

6.3 Représentation multidimensionnelle

Les résultats suivants ont été obtenus pour des caractéristiques calculées sur des fenêtres de 60 s (L'effet du fenêtrage sera discuté au chapitre 9). Les tracés des différents individus dans différents espaces formes possibles permettent de représenter les données par des nuages de points. A chaque point nous avons attribué la couleur correspondant à la classe attribuée par l'expert. Le code de couleur retenu est le suivant :

Vert ¹ Classe ω_1 – Rien A Signaler

Jaune ² Classe ω_2 – Choc compensé

Rouge orangé ³ Classe ω_3 – Choc décompensé

Rouge sang ⁴ Classe ω_4 – Choc irréversible

Les points ont été tracés aléatoirement de façon à éviter le recouvrement d'une classe par une autre. En effet, si l'on trace les points d'une façon ordonnée suivant la classe, en utilisant la même grosseur de point pour chaque classe, il y a risque de recouvrement de la classe tracée en premier par la classe tracée en dernier.

Dans une première représentation (Fig. 6.5), la donnée des points ne dépend que d'une seule caractéristique et du temps (représentation paramétrique). Les formes des nuages correspondant aux classes s'entrecroisent et la discrimination entre classes paraît difficile.

Dans une représentation à 2 caractéristiques (Fig. 6.6), le temps ayant été exclu de la représentation (représentation explicite), des formes apparaissent et on peut se demander si une discrimination de celles-ci est possible. On note le changement important des formes lors des représentations conjointes des variables, en particulier dans un espace $\{FC, T_{cut}\}$ (Fig. 6.6 b)), où il semble exister une influence mutuelle.

Dans un espace de représentation à 3 caractéristiques (Fig. 6.7), les nuages de points correspondant aux différentes classes de choc hémorragique semblent prendre forme, laissant néanmoins subsister quelques doutes et incertitudes. Une classification des données semble possible et c'est ce que nous avons évalué dans la partie suivante de l'étude.

Enfin, on peut encore essayer de représenter le comportement des variables en 4 dimensions. Pour cela, on choisit une caractéristique dont le domaine de définition sera découpé en intervalles de même longueur, créant ainsi une partition simple de l'espace. Les données de chaque intervalle sont alors sélectionnées et un affichage de celles ci dans un espace à 3 dimensions constitué des dernières caractéristiques est alors réalisé correspondant à chaque partition réalisée selon la première caractéristique retenue. Chaque point peut être représenté par sa classe. Le résultat d'une telle démarche peut être observé Fig. 6.8 et met en évidence la possibilité d'une classification dans un espace de cette dimension. La segmentation du domaine de définition de la variable de SpO_2 a été réalisée en considérant les valeurs seuils de 85% et 95% fréquemment utilisées dans le domaine du monitoring en réanimation.

¹jade, viridian léger, vert léger : RVB (128, 255, 128)

²jaune de chrome, jaune : RVB (255, 255, 0)

³corail, rouge orangé, orangé : RVB (255, 64, 0)

⁴sang, brun van dyck, marron : RVB (128, 0, 0)

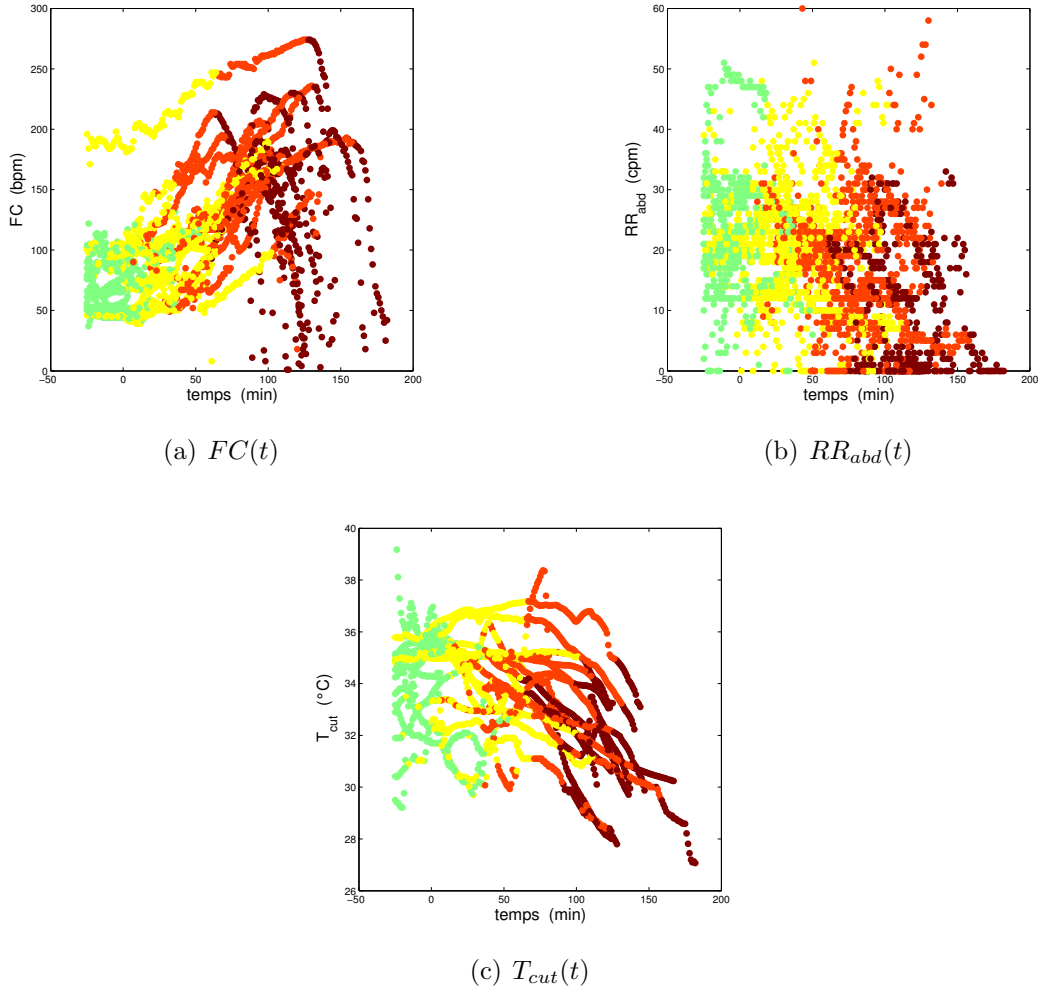


FIG. 6.5 – Espace de représentation à une dimension (1D), une caractéristique, en fonction du temps. À chaque point représentant une observation de la base de données est associé une couleur correspondant à la classe donnée par l'expert. Le code de couleur retenu est le suivant : vert – classe ω_1 (Rien A Signaler), jaune – classe ω_2 (choc compensé), rouge orangé – ω_3 (choc décompensé), rouge sang – ω_4 (choc irréversible).

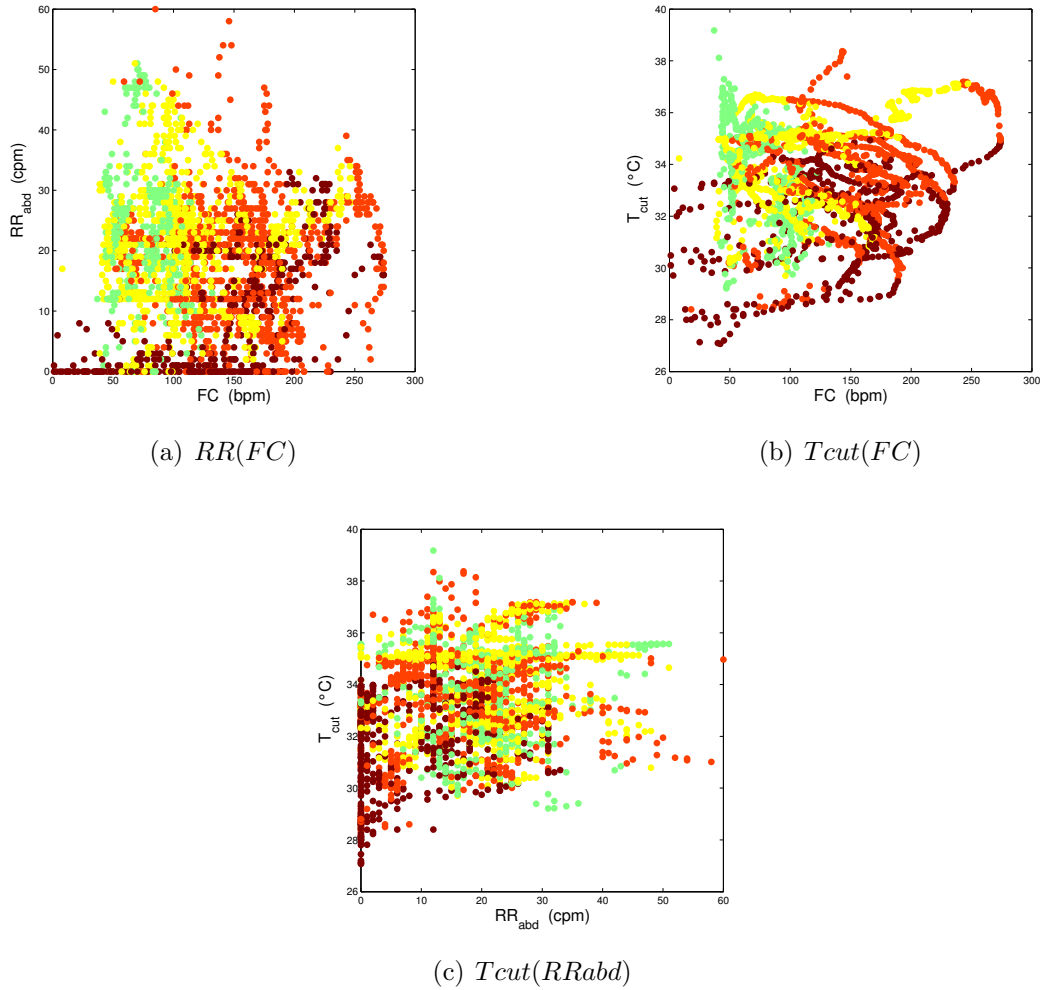


FIG. 6.6 – Espace de représentation à deux dimensions (2D), deux caractéristiques en fonction l'une de l'autre. À chaque point représentant une observation de la base de données est associé une couleur correspondant à la classe donnée par l'expert. Le code de couleur est identique à la figure Fig. 6.5. Les 3 espaces de représentation a) b) et c) indiquent que certains espaces sont plus favorables que d'autres pour une classification

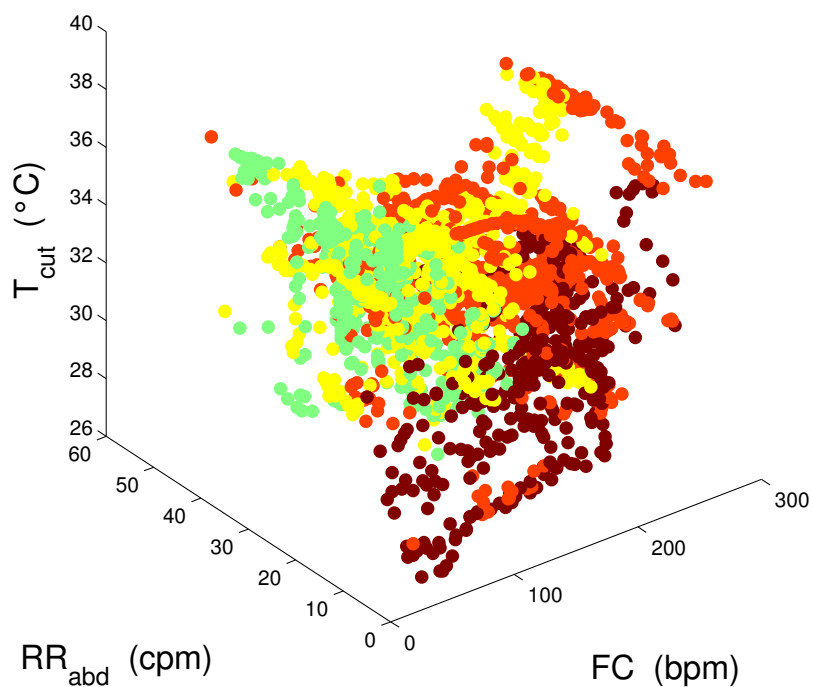


FIG. 6.7 – Espace de représentation à trois dimensions (3D), trois caractéristiques en fonction les unes des autres. À chaque point représentant une observation de la base de données est associé une couleur correspondant à la classe donnée par l'expert. Le code de couleur est identique à la figure Fig. 6.5. On remarque la distribution des points des différentes classes en plusieurs régions de l'espace.

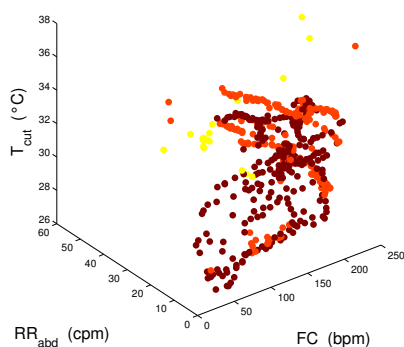
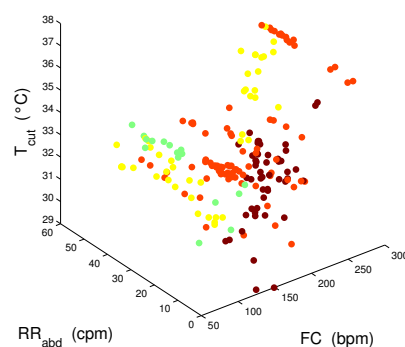
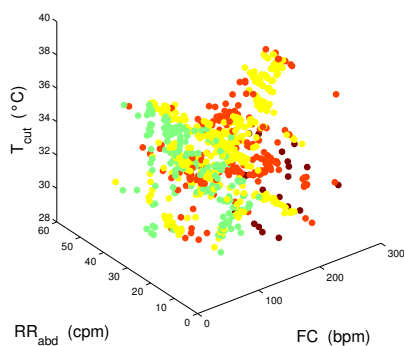
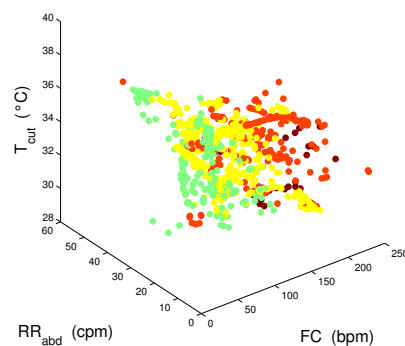
(a) $SpO_2 \leq 25\%$ (b) $25 < SpO_2 \leq 85\%$ (c) $85 < SpO_2 \leq 95\%$ (d) $95 < SpO_2 \leq 100\%$

FIG. 6.8 — Espace de représentation à quatre dimensions (4D), quatre caractéristiques en fonction les unes des autres. À chaque point représentant une observation de la base de données est associé une couleur correspondant à la classe donnée par l'expert. Le code de couleur est identique à la figure Fig. 6.5. La répartition des classes semble former différentes régions de l'espace. Les valeurs seuils de 85% et 95% fréquemment utilisées dans le domaine du monitoring en réanimation ont été retenues.

Chapitre 7

Classification

Lorsque l'on a affaire à une phénoménologie, rendre une situation intelligible signifie avant tout isoler dans cette phénoménologie, dans la morphologie donnée dans l'espace substrat, des éléments identifiables, reconnaissables, stables. J'estime que le but de toute théorie scientifique est l'explication d'une morphologie empirique (ou expérimentale).

René Thom, Paraboles et Catastrophes, p. 90, champs Flammarion, 1980, Paris

Dans ce chapitre, nous présentons les résultats de la classification sur la base des données en utilisant les méthodes exposées précédemment. C'est donc dans cette partie que nous évaluons le système envisagé suivant les différents critères qui peuvent être discutés : le choix des échantillons (7.1), la recherche des caractéristiques pertinentes (7.2), l'analyse des modèles appris (7.3) et l'influence du rejet de certaines données (7.4).

7.1 Choix des échantillons

Dans cette section, nous avons utilisé une évaluation des performances de type leave one (animal) out. Celle ci nous semblait bien appropriée pour obtenir une estimation non biaisée de l'estimation du risque réel, tout en permettant de mettre en évidence les effets des individus. Nous avons lancé une procédure de SFS en minimisant le critère de l'erreur réelle sur un ensemble contenant toutes les caractéristiques disponibles et un ensemble ne contenant que les caractéristiques portables. Chaque procédure a été appliquée en effectuant un tirage aléatoire de façon à obtenir un nombre égal de représentants sur chaque classe pour l'apprentissage de façon à obtenir une équiprobabilité des classes a priori. Le calcul des probabilité a posteriori selon la règle de Bayes s'est donc fait en utilisant cette équiprobabilité a priori lorsque nous avons utilisé cette règle.

7.2 Recherche des caractéristiques pertinentes

Un classifieur quadratique a été retenu pour effectuer une sélection de caractéristiques selon la procédure SFS, dans un premier temps en considérant l'ensemble des caractéristiques invasives et portables de l'étude (toutes les caractéristiques disponibles, ensemble

noté $\mathcal{S}_*$ par la suite), dans un second temps en considérant l'ensemble des caractéristiques non invasives (ensemble des caractéristiques portables, ensemble noté $\mathcal{S}_P$ par la suite). Le choix du classifieur quadratique a été déterminé par sa rapidité algorithmique pour l'apprentissage et pour l'évaluation. On montrera par la suite que ce classifieur est de plus très compétitif.

7.2.1 Toutes les caractéristiques

La progression de la sélection obtenue sur toutes les caractéristiques (invasives et portables) est représentée Fig. 7.1. L'erreur de test $err_{\mathcal{T}}$ atteint un minimum lorsque quatre caractéristiques sont retenues ($\{PAM, PAS, FC, T_{cut}\}$). On ne peut obtenir une performance inférieure à 20% d'erreurs entre l'expert et la machine lorsque celle-ci est évalué avec l'ensemble de test ($err_{\mathcal{T}}$, critère de sélection de la méthode). Ceci indique que la machine ne peut apprendre un modèle meilleur que celui obtenu avec ces quatre caractéristiques et ceci, même si elle a accès à l'ensemble des données dont l'expert a pu se servir pour réaliser son scoring manuel. Ces erreurs peuvent être expliquées par les ambiguïtés existantes aux frontières des différentes zones de classes de choc qui ont été trouvées par la machine.

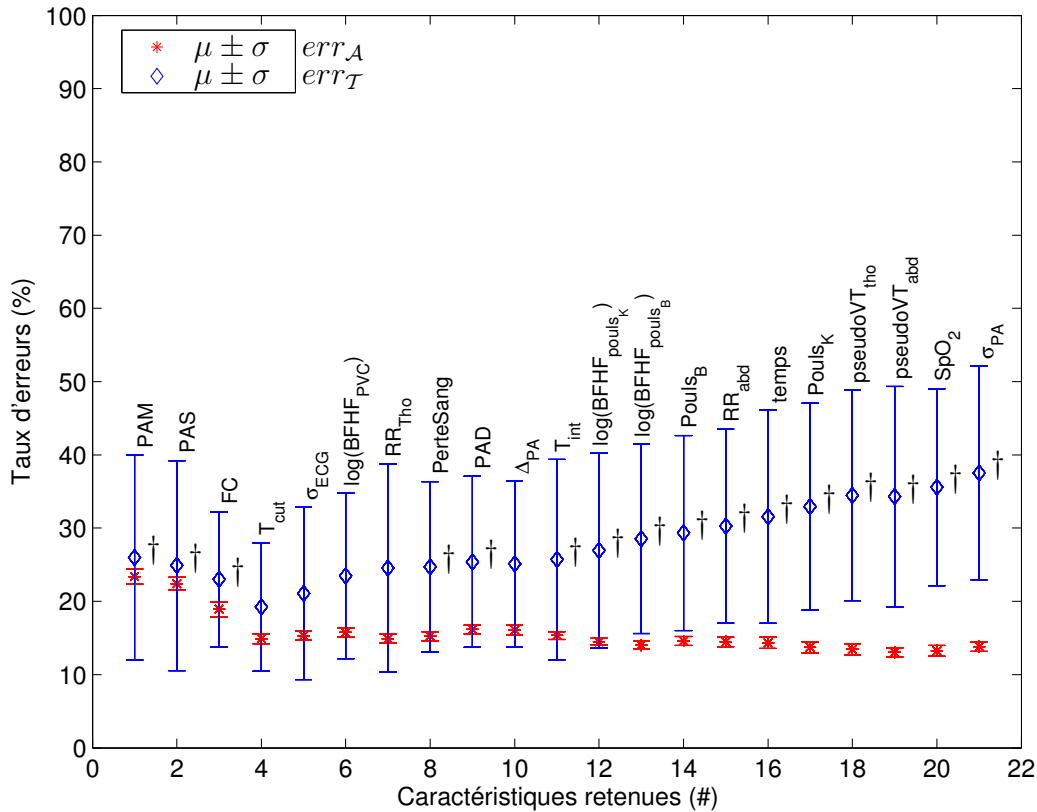


FIG. 7.1 – SFS $\mathcal{S}_*$: Sélection ascendante de caractéristiques par un classifieur quadratique sur l'ensemble des caractéristiques disponibles par leave one (animal) out. Les valeurs notées $\dagger$ indiquent une différence significative pour un test de Wilcoxon entre les échantillons de l'ensemble en question et ceux de l'ensemble de référence $\{PAM, PAS, FC, T_{cut}\}$, déterminé par obtention de l'erreur moyenne en test $err_{\mathcal{T}}$ la plus faible.

Ceci peut encore s'expliquer par des erreurs de classification réalisées par l'expert lors du scoring manuel. L'expert décide toujours d'assigner un score même si il existe une ambiguïté entre classes qu'il lève par connaissance de son score précédent ou en utilisant un autre démarche diagnostique propre à sa connaissance experte. Il est alors très difficile pour la machine, qui travaille à un pas de temps ou qui adopte une démarche calculatoire, de trouver une classification correspondant à une bascule d'un score à l'autre lors d'une succession d'états transitoires.

Comparaison des classifieurs

Une comparaison des différents classifieurs est obtenue Fig. 7.2 en utilisant les quatre meilleures caractéristiques retenues par le classifieur quadratique ($\{PAM, PAS, FC, T_{cut}\}$). Le meilleur classifieur est le perceptron multicouches avec une erreur moyenne obtenue sur l'ensemble d'apprentissage $err_A = 24.6\%$ et une erreur moyenne obtenue sur l'ensemble de test $err_T = 19.7\%$. Pour ce classifieur, l'erreur de test semble inférieure à l'erreur en apprentissage, mais la forte variance ne permet pas de conclure sur leur différence. La comparaison avec les autres modèles indique qu'il n'y a pas de différence significative (test de rang de Wilcoxon au seuil de 5% par comparaison au modèle obtenant l'erreur de test err_T la plus faible possible) avec un classifieur quadratique ou d'autres classifieurs en utilisant un ajustement de paramètres adéquat. Les classifieurs obtenant une faible différence entre err_A et err_T indiquent un bon comportement en généralisation, ce qui est à leur avantage. Nous privilégions donc ceux ci par rapport à ceux obtenant un écart important entre ces deux erreurs.

7.2.2 Caractéristiques portables

L'application d'une procédure SFS sur un ensemble de caractéristiques portables est représenté Fig. 7.3. Les quatre première caractéristiques retenues sont ($\{FC, RR_{Abd}, T_{cut}, PoulsK\}$). On peut noter qu'il n'y a pas d'amélioration significative après la sélection de T_{cut} .

Comparaison des classifieurs

Une comparaison de tous les classifieurs est obtenue Fig. 7.4 et aboutit aux mêmes conclusions que lors de l'utilisation de toutes les caractéristiques disponibles. Le meilleur classifieur est le perceptron multicouches avec une erreur empirique et une estimation de l'erreur réelle de 40.9% et 42.6%. La variance importante obtenue, et ceci quel que soit le classifieur, est due à la grande différence obtenue entre les différents animaux (certains obtiennent des scores très proches de l'expert tandis que d'autres obtiennent de nombreuses erreurs).

La procédure SFS appliquée à un classifieur PMC conduit à la sélection des ensembles $\mathcal{S}_* = \{PAM, PAS, FC, T_{cut}\}$ et $\mathcal{S}_P = \{FC, RR_{Abd}, T_{cut}, SpO_2\}$. La similarité de ces ensembles avec ceux retenus par le classifieur quadratique est en faveur du choix de ces caractéristiques. Nous avons retenus ces deux ensembles par la suite qui nous paraissaient très pertinents en accord avec l'expert.

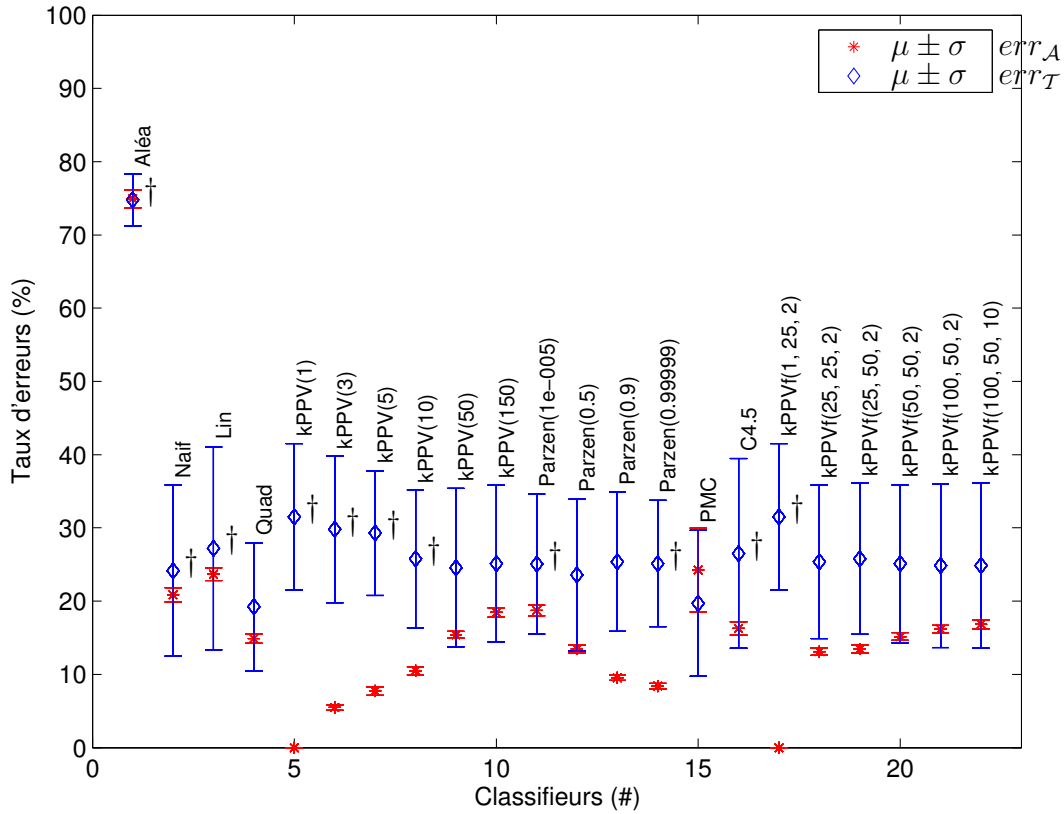


FIG. 7.2 – Comparaison $\mathcal{S}_*$: comparaison des classifieurs pour l'ensemble des caractéristiques disponibles retenues par leave one (animal) out. Les valeurs notées $\dagger$ indiquent une différence significative pour un test de Wilcoxon entre les échantillons de l'ensemble en question et ceux de l'ensemble de référence obtenu pour un classifieur quadratique, déterminé par obtention de l'erreur moyenne en test $err_{\mathcal{T}}$ la plus faible.

7.2.3 Matrices de confusion

L'interprétation de ces erreurs peut encore être expliquée en regardant la matrice de confusion représentée Tab. 7.1 pour un classifieur quadratique et Tab. 7.2 pour un classifieur PMC. La comparaison du score expert et du score de la machine indique qu'un grand nombre d'erreurs provient de l'incertitude entre les classes adjacentes, exprimant les incertitudes liées à une ambiguïté au niveau des frontières de décision.

Les matrices de confusion ont été obtenues en estimant la moyenne et l'écart-type par leave one (animal) out. Un écart-type important indique une différence importante entre les différents individus.

La comparaison des matrices de confusion pour des classification réalisées avec les ensembles $\mathcal{S}_*$ et $\mathcal{S}_P$ montre que l'augmentation des erreurs provient des classes intermédiaires ω_2 et ω_3 . Ainsi pour le classifieur quadratique, on passe d'une concordance de 15.2% à 8.2% pour la classe ω_2 , 19.2% à 8% pour la classe ω_3 en apprentissage et 17.7% à 8.7% pour la classe ω_2 , 22% à 7.8% pour la classe ω_3 en test. La classe ω_2 a tendance à être classée en classe ω_1 et la classe ω_3 en classe ω_2 pour ce classifieur. Les mêmes constatations sont réalisées pour le classifieur PMC à l'exception de l'erreur de classification sur la classe ω_3 qui est plutôt classée en ω_4 par la machine.

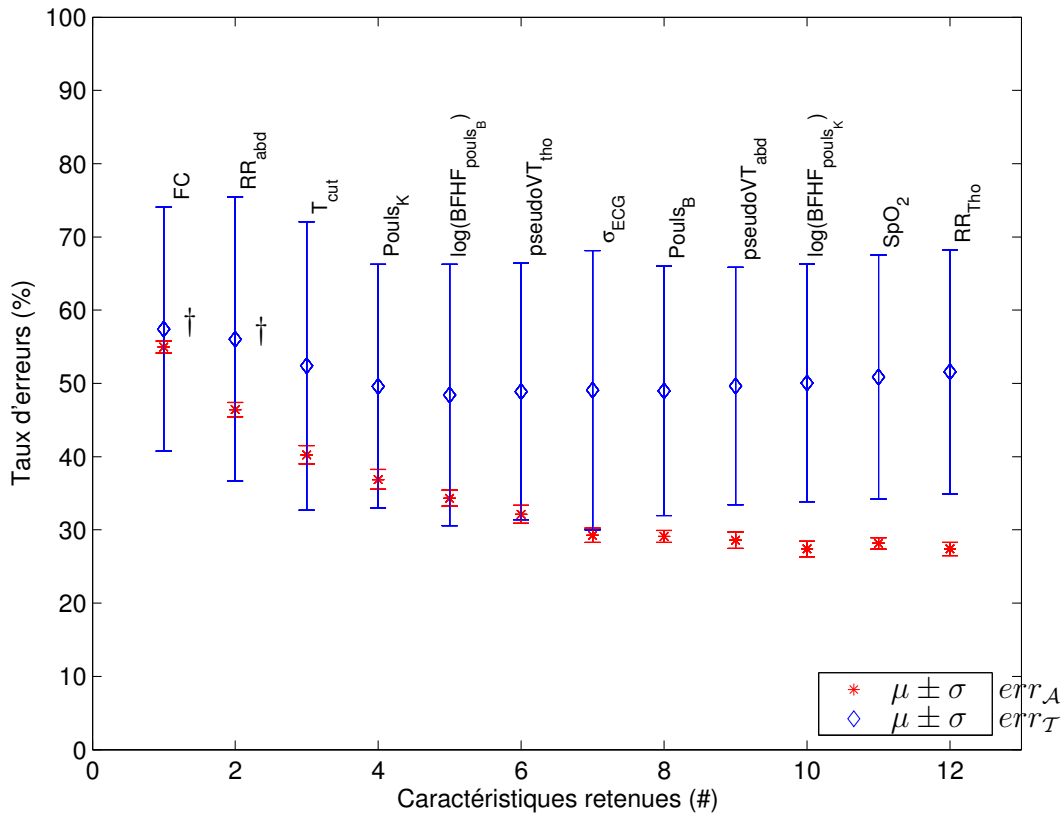


FIG. 7.3 – SFS $\mathcal{S}_P$: Sélection ascendante de caractéristiques par un classifieur quadratique sur l'ensemble des caractéristiques portables par leave one (animal) out. Les valeurs notées $\dagger$ indiquent une différence significative pour un test de Wilcoxon entre les échantillons de l'ensemble en question et ceux de l'ensemble obtenant une moyenne d'erreur en test μ_{err_T} minimale $\{FC, RR_{Abd}, T_{cut}, Poul_K, \log(LFHF_{Poul_B})\}$, déterminé par obtention de l'erreur moyenne en test err_T la plus faible.

Les gains obtenus par le classifieur PMC sur le classifieur quadratique, en particulier en test, sont obtenus grâce au meilleur taux de classification en classe ω_2 et ω_3 . En effet, on passe d'une concordance de 8.7% pour la classe ω_2 pour le classifieur quadratique à une concordance de 14.9% pour le classifieur PMC. La concordance pour la classe ω_3 passe de 7.8% à 16.2%. Ce gain est en partie réalisé sur une meilleure discrimination des classes ω_1, ω_2 et ω_2, ω_3 par le PMC. La classe ω_4 est généralement bien discriminée et ceci quel que soit le classifieur utilisé.

La comparaison des résultats obtenus sur les ensembles d'apprentissage pour ces deux classifieurs indique que les erreurs sont identiquement réparties, ce qui peut conduire à une remise en question des données d'apprentissage (la classification de l'expert). Avec les résultats obtenus en test, nous pouvons conclure que le PMC obtient une meilleure généralisation pour notre problème en discriminant mieux les classes ω_2 et ω_3 .

L'agrégation des classes $\{\omega_1, \omega_2\}$ et $\{\omega_3, \omega_4\}$ pour ces caractéristiques fait chuter l'erreur err_T à 19.5% par exemple, pour le classifieur quadratique et 15.6% pour le PMC. Ceci indique que le système utilisé pour diagnostiquer deux états, ω_1 : rien à signaler ou choc compensé, contre, ω_2 : choc décompensé ou choc irréversible, est intéressant avec sensibilité de 87% pour ω_1 et une spécificité de 73% pour ω_2 pour le classifieur quadratique

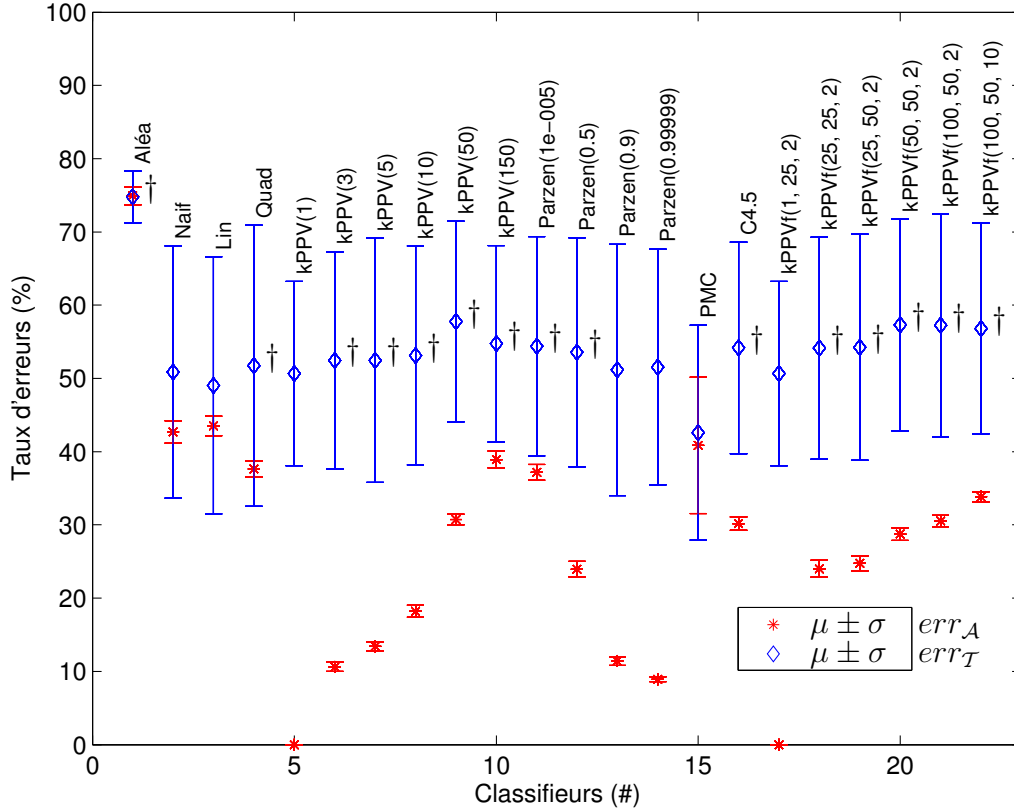


FIG. 7.4 – Comparaison $\mathcal{S}_P$: comparaison des classifieurs pour l'ensemble des caractéristiques portables retenues par leave one (animal) out. Les valeurs notées † indiquent une différence significative pour un test de Wilcoxon entre les échantillons de l'ensemble en question et ceux de l'ensemble de référence obtenu pour un classifieur perceptron multicouches, déterminé par obtention de l'erreur moyenne en test $err_{\mathcal{T}}$ la plus faible.

et une sensibilité de 89.2% pour ω_1 et une spécificité de 79.3% pour ω_2 pour le PMC.

7.3 Analyse des modèles

7.3.1 Classifieur quadratique

Un exemple d'estimation des paramètres du modèle Gaussien multidimensionnel obtenue lors de l'apprentissage d'un classifieur quadratique pris au hasard est disponibles Tab. 7.3. Sur cet exemple, nous avons effectué l'opération inverse du zscore de façon à retrouver les grandeurs d'origine associées à chaque caractéristique. Cet exemple met en évidence la différence des liens entre variables qui évolue au cours des différentes classes de choc hémorragique. La matrice de variance-covariance pouvant être décomposée selon :

$$\Sigma = \mathbf{\Gamma} \mathbf{R} \mathbf{\Gamma} \quad (7.1)$$

avec $\mathbf{\Gamma}$ matrice des écart-types et $\mathbf{R}$ matrice des corrélations.

L'analyse de la matrice de covariance montre l'évolution du lien des paramètres du modèle au cours de l'expérimentation. Ainsi, pour la classe ω_4 , on observe que l'ensemble

TAB. 7.1 – Matrices de confusion pour le classifieur quadratique

$S_* = \{PAM, SpO_2, FC, PAS\}$								
Expert	Apprentissage Machine				Test Machine			
	ω_1	ω_2	ω_3	ω_4	ω_1	ω_2	ω_3	ω_4
ω_1	22.5 ± 0.4	2.2 ± 0.4	0.3 ± 0.2	0.1 ± 0.0	16.4 ± 10.1	2.2 ± 4.4	0.1 ± 0.4	0.1 ± 0.2
ω_2	6.9 ± 0.8	15.2 ± 1	2.8 ± 0.3	0.14 ± 0.1	10.2 ± 11.9	17.7 ± 11.0	5.2 ± 7.3	0.2 ± 0.5
ω_3	0.4 ± 0.1	3.9 ± 0.5	19.2 ± 0.5	1.5 ± 0.2	0.8 ± 1.4	4.8 ± 6.2	22.0 ± 6.8	2.9 ± 3.7
ω_4	0.0 ± 0.0	0.1 ± 0.1	3.7 ± 0.5	21.2 ± 0.5	0.0 ± 0.0	0.0 ± 0.2	2.4 ± 2.0	15.0 ± 8.6
$err_A : 21.9 \pm 0.4$					$err_T : 28.8 \pm 6.1$			
$S_P = \{FC, T_{cut}, RR_{Abd}, SpO_2\}$								
	ω_1	ω_2	ω_3	ω_4	ω_1	ω_2	ω_3	ω_4
ω_1	24.0 ± 0.2	0.7 ± 0.2	0.2 ± 0.1	0.1 ± 0.0	17.2 ± 11.4	0.9 ± 1.5	0.6 ± 1.4	0.1 ± 0.3
ω_2	13.2 ± 0.9	8.2 ± 1.6	2.9 ± 0.8	0.7 ± 0.2	17.9 ± 14.8	8.7 ± 13.4	5.6 ± 13.2	1.2 ± 2.9
ω_3	3.5 ± 0.4	7.8 ± 1.0	8.0 ± 1.0	5.7 ± 0.4	3.9 ± 6.8	10.4 ± 11.7	7.8 ± 7.2	8.4 ± 8.4
ω_4	0.2 ± 0.0	0.5 ± 0.3	2.7 ± 0.5	21.6 ± 0.5	0.1 ± 0.5	0.9 ± 2.2	2.7 ± 6.5	13.7 ± 5.0
$err_A : 38.2 \pm 0.5$					$err_T : 52.7 \pm 6.2$			

TAB. 7.2 – Matrices de confusion pour le classifieur PMC

$\mathcal{S}_* = \{PAM, SpO_2, FC, PAS\}$								
Expert	Apprentissage Machine				Test Machine			
	ω_1	ω_2	ω_3	ω_4	ω_1	ω_2	ω_3	ω_4
ω_1	18.6 ± 4.5	5.9 ± 4.4	0.2 ± 0.0	0.4 ± 1.3	16.1 ± 11.8	2.6 ± 3.9	0.1 ± 0.4	0.0 ± 0.0
ω_2	3.9 ± 4.6	17.7 ± 5.1	3.4 ± 2.2	0.0 ± 0.1	3.3 ± 4.2	26.7 ± 16.0	3.3 ± 4.7	0.0 ± 0.0
ω_3	0.2 ± 0.5	3.4 ± 1.4	19.2 ± 2.5	2.2 ± 1.5	0.2 ± 0.4	3.3 ± 2.9	24.9 ± 8.4	2.2 ± 3.0
ω_4	0.1 ± 0.1	0.4 ± 0.3	3.4 ± 2.7	21.1 ± 2.6	0.0 ± 0.2	0.3 ± 0.5	2.5 ± 3.2	14.6 ± 9.0
$err_A : 23.4 \pm 2.5$					$err_T : 17.7 \pm 7.5$			
$\mathcal{S}_P = \{FC, T_{cut}, RR_{Abd}, SpO_2\}$								
	ω_1	ω_2	ω_3	ω_4	ω_1	ω_2	ω_3	ω_4
ω_1	19.8 ± 5.3	2.9 ± 3.0	1.2 ± 4.5	1.0 ± 1.9	15.6 ± 12.0	1.7 ± 2.4	1.1 ± 4.5	0.4 ± 0.8
ω_2	10.2 ± 3.7	7.9 ± 4.4	4.7 ± 5.1	2.3 ± 1.4	12.4 ± 12.6	14.9 ± 19.3	3.0 ± 4.0	3.0 ± 6.7
ω_3	3.2 ± 1.7	4.1 ± 2.9	10.3 ± 4.7	7.4 ± 3.8	2.3 ± 5.5	4.0 ± 5.3	16.2 ± 12.3	7.9 ± 10.7
ω_4	0.2 ± 0.3	1.0 ± 2.7	4.1 ± 3.8	19.7 ± 5.8	0.1 ± 0.5	0.3 ± 0.9	3.9 ± 5.3	13.1 ± 9.2
$err_A : 42.3 \pm 3.4$					$err_T : 40.1 \pm 8.8$			

des variables ont une corrélation positive non négligeable (p. ex. 0.39 entre FC et T_{cut} , 0.61 entre FC et RR_{Abd} , 0.5 entre RR_{Abd} et SpO_2). Pour cette classe correspondant au choc irréversible, toutes les variables diminuent ensemble dans le temps et ceci quel que soit l'individu concerné.

L'analyse des moyennes tend à prouver qu'on observe : une fréquence cardiaque importante (tachycardie) en choc compensé ($FC = 109 \text{ bpm}$) et décompensé ($FC = 155 \text{ bpm}$) suivie d'une diminution en choc irréversible (bradycardie) ($FC = 145 \text{ bpm}$), en considérant une évolution des valeurs de la classe ω_1 à ω_4 . La température cutanée moyenne est plus importante lors de la phase de compensation ($T_{cut} = 34.2^\circ C$) que dans les autres classes et particulièrement au repos ($T_{cut} = 33.8^\circ C$). Elle diminue lors de la phase décompensée ($T_{cut} = 33.6^\circ C$) et diminue encore lors du choc irréversible ($T_{cut} = 31.7^\circ C$). Le rythme respiratoire diminue : on passe de 23 cpm à 20 cpm puis 18 cpm pour finir à 10 cpm). La SpO_2 est de plus en plus basse, elle passe de 98% en classe ω_1 à 87.5% en classe ω_2 puis 55% en classe ω_3 pour finir à une moyenne de 19% en classe ω_4 .

TAB. 7.3 – Modélisation quadratique par classe : Estimations des vecteurs moyennes, des matrices de écart-types et des matrices de corrélations par classe.

Classe	$\hat{\mu}$			$\hat{\Gamma}$			$\hat{\mathbf{R}}$			
		FC	T_{cut}	RR_{Abd}	SpO_2		FC	T_{cut}	RR_{Abd}	SpO_2
ω_1	FC	75.14	22.1	0	0	0	1	-0.40	-0.16	-0.16
	T_{cut}	33.82	0	1.88	0	0	-0.40	1	0.23	0.04
	RR_{Abd}	23.42	0	0	10.4	0	-0.16	0.23	1	0.12
	SpO_2	93.30	0	0	0	9.02	-0.16	0.04	0.12	1
ω_2	FC	109.12	47	0	0	0	1	0.24	0.08	0.06
	T_{cut}	34.22	0	1.73	0	0	0.24	1	0.13	0.12
	RR_{Abd}	20.92	0	0	9.77	0	0.08	0.13	1	-0.00
	SpO_2	87.51	0	0	0	22.8	0.06	0.12	-0.00	1
ω_3	FC	155.12	51.6	0	0	0	1	0.23	-0.03	-0.39
	T_{cut}	33.58	0	1.97	0	0	0.23	1	0.20	0.24
	RR_{Abd}	18.77	0	0	10.7	0	-0.03	0.20	1	0.15
	SpO_2	55.31	0	0	0	44.6	-0.39	0.24	0.15	1
ω_4	FC	145.29	59.7	0	0	0	1	0.39	0.61	0.37
	T_{cut}	31.68	0	1.77	0	0	0.39	1	0.34	0.03
	RR_{Abd}	10.26	0	0	9.48	0	0.61	0.34	1	0.50
	SpO_2	18.57	0	0	0	34.5	0.37	0.03	0.50	1

Ces résultats sont en accord avec les descriptions obtenues dans la littérature [17, 18]. On observe une légère différence pour le rythme respiratoire qui devrait s'accélérer au cours du choc compensé, puis diminuer ensuite. Cette augmentation est en général prise en considération dans les tables de scoring (p. ex. Tab. B.1 et autres en Annexe B).

La diminution observée dans notre expérimentation, pourrait être due à un effet trop puissant de l'anesthésie qui aurait pour effet de minimiser la commande des centres respiratoires au niveau du système nerveux central. Cet effet sur le système nerveux est confirmé par la réactivité sympathique faible observée sur les signaux de pression artérielle. Il semble donc que le propofol (Diprivan®) a un effet délétère sur la réactivité sympathique (en particulier lors de la phase sympathoexcitatrice) normalement observée lors la phase de compensation du choc hémorragique.

7.3.2 Arbres de décisions

On peut remarquer que les arbres de décision ne nécessitent pas de normalisation des données et que le résultat visuel permet une appréhension rapide des modèles générés. Ces considérations ont fait des arbres de décision un outil pratique et assez bien accueilli en médecine. Cependant, dans cette discipline qui introduit un grand nombre de facteurs pour la mise en place d'un diagnostic, on arrive souvent à des arbres de décisions d'une grande complexité qui rend le modèle caduc.

Une exemple d'arbre généré par l'ensemble de caractéristiques portables en utilisant l'algorithme C4.5 est représenté Fig. 7.5. Cet arbre met en évidence que le premier paramètre retenu par l'algorithme est celui de *SpO₂* qui, suivant des considérations mathématiques, joue un rôle discriminant. L'arbre est alors séparé en 2 sous arbres : celui de droite correspond aux classes ω_1 , ω_2 et ω_3 ; celui de gauche correspond aux classes

ω_3 et ω_4 . Les classes associées aux feuilles de l'arbre et les nombres d'échantillons de la base d'apprentissage associés à cette branche indiquent que la classe ω_4 ($SpO_2 < 48\%$, $RR_{Abd} < 14\text{ cpm}$, $FC < 184\text{ bpm}$, $\forall T_{cut}$) est assez bien déterminée par cette branche. Cet exemple est en accord avec les observations médicales et pourrait s'énoncer ainsi : quelle que soit la température, si la saturation est très basse¹, la respiration lente, et qu'on observe une bradycardie, le patient est dans un état de choc irréversible.

Le parcours de la branche de droite conduit à des situations plus complexes qui mettent en évidence les indécisions du modèle en particulier pour la détermination de la classe ω_1 ou ω_2 suivant la T_{cut} , la FC et le rythme respiratoire. Les nombreuses bifurcations de l'arbre peuvent encore être expliquées par le fait que les hyperplans générés par les conditions associées aux fourches de l'arbre et définies par une seule variable, ne sont pas appropriés. Elles mettent en évidence un lien entre différentes variables (des séparatrices de décisions non univariées).

Lorsqu'on utilise un arbre de décision composé de nombreux niveaux (nombreuses branches et feuilles terminales), on risque de favoriser un apprentissage par cœur. Au contraire, l'utilisation d'un arbre de décision élagué (angl. *pruned tree*), c'est à dire dont le nombre de ramifications a été réduit par affectation de la décision à la classe majoritaire sur l'ensemble des branches descendantes, permet d'obtenir une meilleure généralisation. L'utilisation d'une méthode de leave one (animal) out met en évidence cette généralisation et indique qu'un manque de précision permet de mieux classer. Exemple : pour un classifieur créé à partir de 4 caractéristiques, avec un élagage, on obtient $err_{\mathcal{T}}$ 53.7% et sans élagage $err_{\mathcal{T}}$ 59.4% mais on ne peut rejeter l'hypothèse d'égalité de ces distributions ($p = 0.87$). De tels arbres sont représentés Fig. 7.6 et 7.7.

¹on retient en général $< 85\%$ car les appareils de SpO_2 sont assez mal corrélés avec la SaO_2 en dessous d'une certaine valeur - dans la littérature 55% [103]

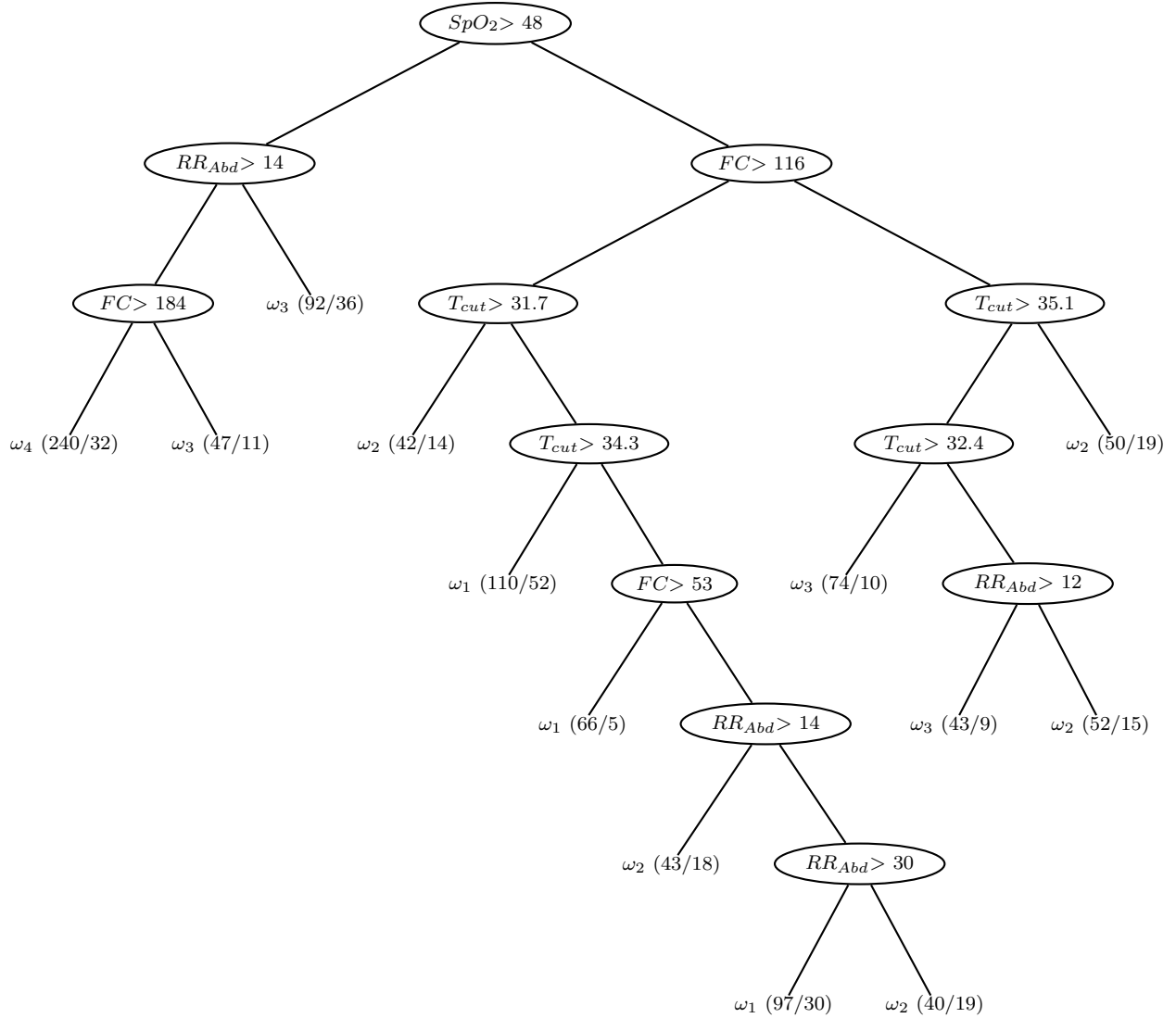
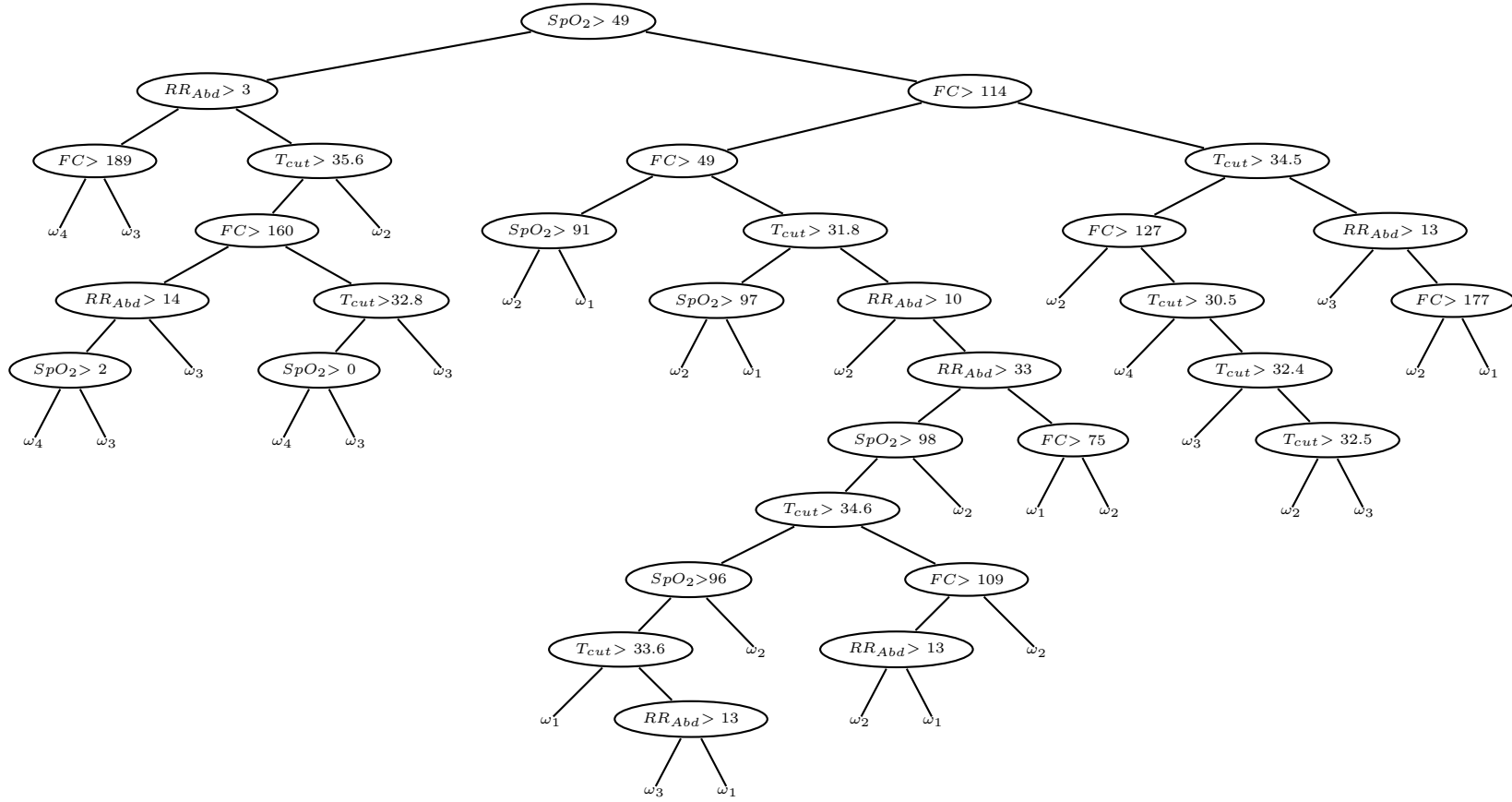
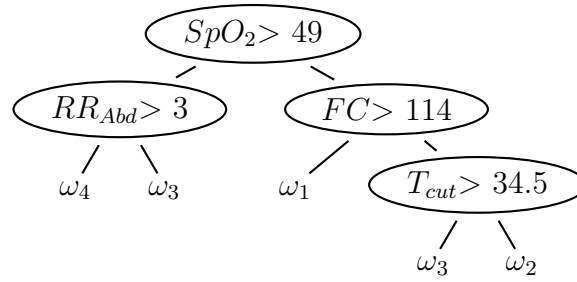


FIG. 7.5 – Arbre de décision C4.5 non élagué obtenu à partir des caractéristiques portables $\mathcal{S}_P$. La convention utilisée est la suivante pour cette représentation : lorsque la condition est réalisée au niveau d'une fourche d'un arbre, on emprunte le chemin de droite. Les feuilles terminales correspondent à la classe attribuée aux données réalisant les conditions associées aux branches parcourues. Ainsi, cet arbre de décision associera au vecteur $\mathbf{x}_i = (FC = 134, RR_{Abd} = 27, SpO_2 = 89, T_{cut} = 34)^T$ la classe $u_i = \omega_2$ en parcourant le chemin : $SpO_2 > 48, FC > 115, T_{cut} < 35.1, T_{cut} > 32.4, RR_{Abd} > 12$. On peut associer à chaque feuille le nombre d'exemples de la base d'apprentissage et leur situation après construction du modèle ($\sharp\{BienClassé\}/\sharp\{MalClassé\}$).

FIG. 7.6 – Arbre de décision non élagué : caractéristiques portables S_P

FIG. 7.7 – Arbre de décision élagué : caractéristiques portables $\mathcal{S}_P$

L'arbre élagué représenté Fig. 7.7 fournit un condensé qui peut être traduit dans la base de règles suivantes, correspondant au modèle animal de notre étude :

- Rien à signaler : Si SpO_2 présente et FC basse ($< 114 bpm$).
- Choc compensé : Si présence de SpO_2 , FC élevée ($> 114 bpm$) et T_{cut} haute ($> 34.5^\circ C$). Le facteur discriminant entre classe ω_2 et ω_3 est ici la température cutanée.
- Choc décompensé : deux cas :
 - Si présence de SpO_2 , avec FC élevée ($> 114 bpm$) et température cutanée basse ($< 34.5^\circ C$) (début de vasoconstriction périphérique).
 - Si absence de SpO_2 mais respiration.
- Choc irréversible : Si absence de SpO_2 et absence de respiration (délai entre chaque respiration $> 20 s$).

7.4 Rejet en ambiguïté et rejet en distance

Une étude de l'influence du coût de rejet, présenté section 5.3.2, a été réalisée sur l'ensemble des caractéristiques portables retenues $\mathcal{S}_P = \{FC, RR_{Abd}, SpO_2, T_{cut}\}$. Les résultats obtenus en faisant varier le coût de rejet en ambiguïté C_r , en utilisant une validation par leave one (animal) out et un calcul de l'erreur sur les données non rejetées, sont représentés Fig. 7.8. L'interprétation de ces courbes peut être réalisée de cette façon. Lorsque le taux de rejet augmente (le nombre d'observations classées diminue), si :

- Les données rejetées étaient mal classées, l'ensemble des données restantes bien classées ne varie pas et l'ensemble des données restantes mal classées diminue : le taux d'erreur calculé sur les données classées diminue.
- Les données rejetées étaient bien classées, les données restantes sont mal classées et le taux d'erreur augmente.
- Les données rejetées font partie d'un ensemble mixte de données bien et mal classées, le taux d'erreur ne varie pas.

Pour notre expérimentation, on s'aperçoit que lorsque 50% des données sont rejetées en ambiguïté, on obtient une erreur réelle $err_{\mathcal{T}}$ de $45 \pm 25\%$ ce qui constitue un gain faible. Il faut enlever un grand nombre de données (et tomber sur un seuil très bas) pour obtenir une performance meilleure en moyenne (un taux d'erreur plus bas). Il est important de noter que les classes affectées par le rejet en ambiguïté sont les classes ω_2 et ω_3 comme nous l'avons vu précédemment. Il semblerait que les données de ces classes rejetées participent équitablement à l'erreur ou à la concordance. L'analyse de la matrice de confusion pour un coût de rejet de 60% associé à un taux de rejet en ambiguïté montre que les données rejetées proviennent des classes ω_2 et ω_3 sur des données qui étaient bien classées.

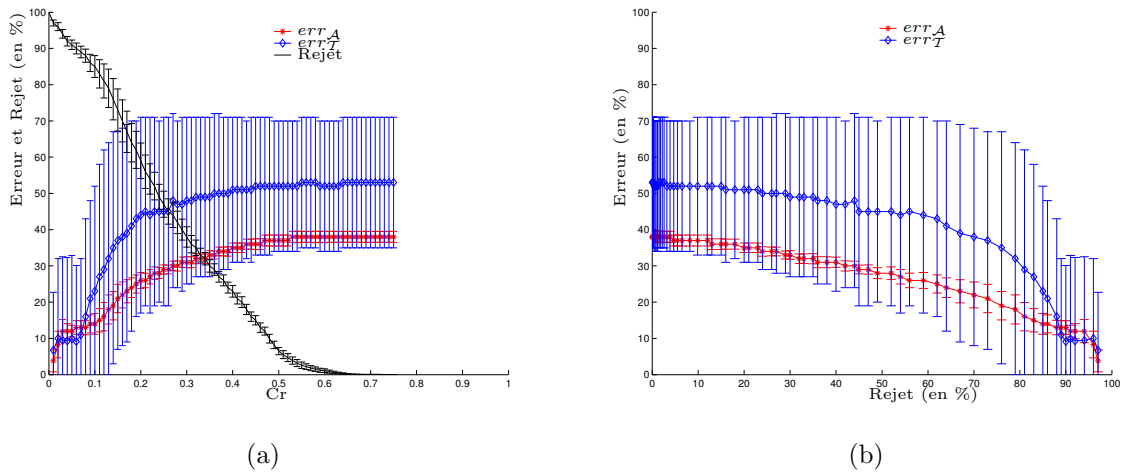


FIG. 7.8 – Influence du coût de rejet en ambiguïté sur l'erreur pour l'ensemble $\mathcal{S}_P$: a) représente le taux de données rejetées (en %) et le taux erreur (en %) en fonction du coût en rejet (Cr). Le taux d'erreurs a été calculé par rapport aux nombres de données non rejetées. Lorsque celui-ci augmente, le pourcentage de données rejetées diminuent. Pour un Cr de l'ordre de 0.1, 80% des données sont rejetées. b) représente le taux d'erreur en fonction du nombre de données rejetées. La nécessité de rejeter un grand nombre de données pour faire diminuer l'erreur indique une discrimination difficile entre classe.

TAB. 7.4 – Matrice de confusion pour un rejet en ambiguïté $Cr = 0.3$: Les principales ambiguïtés proviennent des classes ω_2 et ω_3 . Les données rejetées sont prises aussi bien parmi les données bien classées que parmi les données mal classées lorsque l'on compare cette matrice à celle présentée Tab. 7.1.

	ω_0	ω_1	ω_2	ω_3	ω_4
ω_1	4.34 ± 6.03	14.09 ± 10.30	0.27 ± 0.98	0.03 ± 0.14	0.03 ± 0.14
ω_2	14.96 ± 14.88	11.70 ± 11.90	2.15 ± 4.83	3.80 ± 10.59	0.68 ± 2.21
ω_3	16.73 ± 10.39	2.94 ± 6.01	4.52 ± 7.69	1.71 ± 2.32	4.62 ± 7.04
ω_4	5.32 ± 4.57	0.04 ± 0.17	0.08 ± 0.33	1.29 ± 3.87	10.68 ± 5.24

La variance de l'erreur indique que l'effet de ce rejet est différent selon les individus de l'étude. Ceci pourrait s'expliquer par la proximité ou l'éloignement entre les données de certains individus. Ceci permet de conclure que le mauvais taux de classification provient en partie d'une discrimination difficile entre classes (correspondant aux données rejetées), mais surtout d'une irrégularité de classification entre les individus. Enfin, on ne peut exclure une certaine ambiguïté dans la classification de l'expert, qui pourrait, pour certains individus dont les données sont proches, induire une mauvaise évaluation du système.

Les résultats obtenus en faisant varier le coût de rejet en distance C_d sont représentés Fig. 7.9. Le graphique met en évidence un nombre important de points rejetés sans changement de comportement sur la classification en conservant une erreur réelle err_T à la même valeur (40%).

Cependant, on note l'augmentation de la variation, ce qui indique là encore un changement de comportement sur l'effet de ce rejet selon les individus traités. On constate une augmentation de l'erreur lorsqu'on effectue un rejet important, ce qui indique que les données rejetées en distance n'interviennent pas dans le processus de bonne ou mauvaise classification en moyenne sur les individus de la base de données, sauf lorsqu'on touche aux données de plus en plus proches, la classification est alors moins performante et on

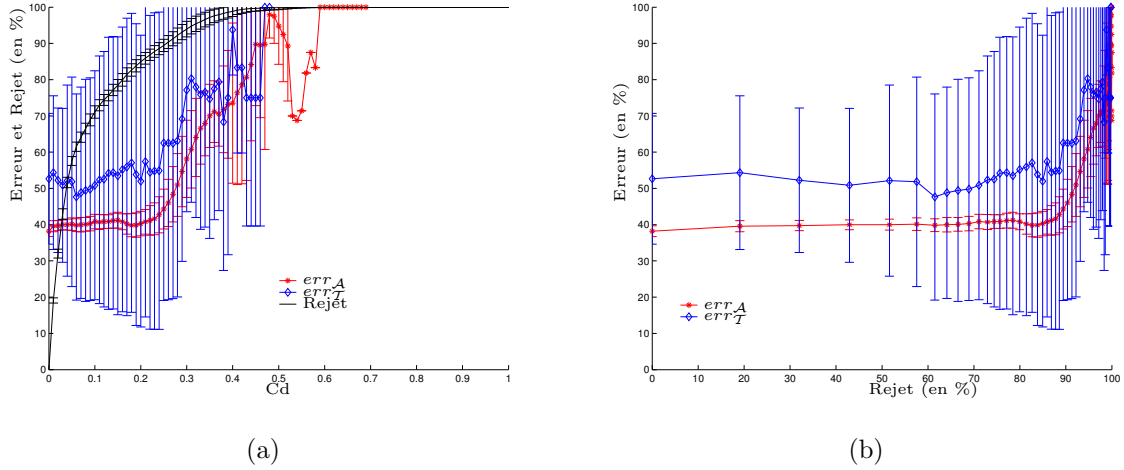


FIG. 7.9 – Influence du coût de rejet en distance sur l'erreur pour l'ensemble $\mathcal{S}_P$: a) représente le taux de données rejetées et le taux d'erreur en fonction du coût de rejet en distance (C_d). Lorsque celui-ci augmente, le pourcentage de données rejetées augmente. Pour un C_d de l'ordre de 0.2, 80% des données sont rejetées. b) représente le taux d'erreur en fonction du nombre de données rejetées. Plus on rejette de données, moins la classification est bonne, plus la variabilité augmente : notre base de données contient des données proches dont la classification est difficile ; il est nuisible d'enlever trop de données pour l'apprentissage de notre base de données.

constate l'influence des données ambiguës (ou des différences entre les individus les plus proches). L'application d'un rejet ne permet d'obtenir qu'une légère amélioration des performances en moyenne. De plus, son emploi ne s'effectue qu'a posteriori, uniquement lors de l'évaluation du système.

7.5 Nettoyage des données

Il est intéressant de rejeter les données incompatibles avec l'apprentissage dès la création de la base de données. Cette opération est généralement désignée sous le terme de *nettoyage des données*. Nous avons procédé de deux façons : la première s'appuie sur une élimination des données instables du scoring expert (changement de décision de l'expert pour deux classifications successives), la seconde s'appuie sur une condensation des données servant à la création de l'ensemble d'apprentissage, par deux techniques différentes.

7.5.1 Élimination des transitions de l'expert

L'élimination des instabilités de l'expert dont la technique a été présentée section 5.3.2 fait passer d'une base de données constituées de 3004 observations à une base de données constituées de 2369 observations ; on ne retient donc que 78.86% des exemples. Le tableau Tab. 7.5 présente les résultats obtenus sur chaque ensemble utilisé lors de la validation par leave one (animal) out pour un classifieur quadratique et un classifieur PMC.

La comparaison de ces valeurs avec celles obtenues sans nettoyage indique que l'élimination des transitions courtes supposées douteuses pour la classification de l'expert, n'apporte pas de gain de performance et ceci pour un classifieur de type quadratique ou perceptron multicouches. On peut donc expliquer la mauvaise classification, ou bien

TAB. 7.5 – Nettoyage : ambiguïté experte, transitions courtes, classifieur quadratique et PMC. On part de l'ensemble des données brutes $\mathcal{S}$ et on applique la procédure de nettoyage des transitions courtes sur les classifications réalisées par l'expert. On obtient un ensemble nettoyé $\mathcal{N}$. L'ensemble d'apprentissage $\mathcal{A}$ est obtenu après tirage aléatoire d'un ensemble équiprobable de représentants de chaque classe.

#	$\sharp\{\mathcal{S}\}$	$\sharp\{\mathcal{N}\}$	$\sharp\{\mathcal{A}\}$	classifieur quadratique				PMC			
				$\mathcal{S}_*$		$\mathcal{S}_P$		$\mathcal{S}_*$		$\mathcal{S}_P$	
				$err_{\mathcal{A}}$	$err_{\mathcal{T}}$	$err_{\mathcal{A}}$	$err_{\mathcal{T}}$	$err_{\mathcal{A}}$	$err_{\mathcal{T}}$	$err_{\mathcal{A}}$	$err_{\mathcal{T}}$
#1	2499	1995	1424	16.1	20.0	36.0	64.8	25.3	12.4	32.7	43.4
#2	2504	2013	1316	16.7	8.6	33.7	56.4	15.3	9.3	50.3	32.1
#3	2497	2005	1384	16.3	33.3	36.6	44.9	16.3	20.4	32.2	45.6
#4	2520	2026	1400	15.7	29.8	37.0	41.9	20.6	8.9	30.1	35.5
#5	2493	2018	1432	16.0	21.2	35.9	49.0	16.5	24.5	31.7	32.5
#7	2482	2002	1392	16.0	30.9	37.1	51.9	22.2	22.2	49.8	51.9
#8	2494	2022	1404	15.8	18.7	35.3	60.7	24.5	19.3	32.3	41.3
#9	2490	2014	1364	12.9	41.6	34.6	40.9	21.6	43.5	34.8	37.0
#10	2437	1976	1364	16.0	23.2	32.6	64.3	28.7	13.0	30.0	62.8
#12	2506	2008	1396	16.5	31.2	36.2	41.3	16.8	15.9	39.2	12.3
#13	2531	2038	1436	16.2	18.6	35.4	46.9	15.8	15.0	35.6	29.2
#14	2452	1969	1360	13.9	44.8	35.2	52.1	18.8	25.5	36.6	51.0
#15	2495	1988	1492	15.4	58.4	35.7	93.3	32.4	10.1	34.8	77.9
#16	2474	1967	1472	15.9	32.4	34.2	87.1	31.5	24.7	43.2	58.2
#17	2481	1978	1452	18.7	24.5	35.7	67.5	27.1	9.8	33.7	37.4
#18	2471	2032	1448	17.5	26.6	37.5	59.5	25.9	32.9	32.9	42.8
#19	2478	2000	1504	16.4	33.7	38.4	30.7	21.5	16.3	36.0	10.8
$\mu \pm \sigma$	2488	81%	57%	16 ± 1	29 ± 12	36 ± 1	56 ± 16	22 ± 6	19 ± 9	36 ± 6	41 ± 17

par une non adaptation des caractéristiques retenues (la classification n'est pas possible dans l'espace de représentation choisi), ou bien une incohérence de classement de l'expert, non pas lors du classement des données sur un individu, mais lors de la comparaison des données entre plusieurs individus.

7.5.2 Nettoyage par 1-ppv et condensation

Les résultats obtenus après nettoyage par 1PPV et condensation sont présentés Tab. 7.6, 7.7, 7.8 et 7.9.

On constate que la méthode du nettoyage par 1 ppv enlève peu de données (en moyenne on conserve 75% des données), ce qui indique que l'ensemble de données est relativement riche en données proches de classes différentes. Lorsqu'on condense les données, on tombe sur un ensemble contenant au contraire, très peu de données (en moyenne 14%). La comparaison des résultats à ceux obtenus sans nettoyage permet de conclure que le nettoyage des données réalisé de cette façon ne permet pas d'obtenir de meilleure performance. Ceci pourrait provenir de l'incohérence de classification de l'expert ou des différences entre individus.

TAB. 7.6 – Nettoyage 1PPV $\mathcal{S}_*$ classifieur quadratique

Animal	$\sharp\{B\}$	$\sharp\{N\}$	ratio (%)	err_A (%)	err_T (%)
#1	2499	2158	86.35	15.20	15.86
#2	2504	2173	86.78	17.06	8.57
#3	2497	2174	87.06	15.36	39.45
#4	2520	2182	86.58	14.34	13.70
#5	2493	2155	86.44	15.91	16.55
#7	2482	2145	86.42	16.17	29.62
#8	2494	2160	86.60	16.29	19.33
#9	2490	2165	86.94	13.08	42.20
#10	2437	2108	86.49	16.73	26.57
#12	2506	2164	86.35	15.92	30.43
#13	2531	2196	86.76	16.59	18.58
#14	2452	2126	86.70	12.97	42.70
#15	2495	2155	86.37	15.17	55.03
#16	2474	2134	86.25	17.04	24.70
#17	2481	2124	85.61	15.80	21.47
#18	2471	2150	87.00	17.29	29.47
#19	2478	2157	87.04	15.06	33.13
$\mu \pm \sigma$	2489 ± 23	2155 ± 22	86.6 ± 0.4	15.6 ± 1.3	27.5 ± 12.2

TAB. 7.7 – Nettoyage 1PPV $\mathcal{S}_P$ classifieur quadratique

Animal	$\sharp\{S\}$	$\sharp\{N\}$	err_A (%)	err_T (%)
#1	2499	1877 (75.11%)	35.01	64.13
#2	2504	1861 (74.32%)	33.36	59.28
#3	2497	1849 (74.04%)	34.35	48.29
#4	2520	1909 (75.75%)	37.03	54.03
#5	2493	1838 (73.72%)	34.93	48.34
#7	2482	1873 (75.46%)	35.62	51.85
#8	2494	1860 (74.57%)	34.67	60
#9	2490	1887 (75.78%)	34.67	40.25
#10	2437	1865 (76.52%)	31.39	67.14
#12	2506	1881 (75.05%)	38.74	42.02
#13	2531	1915 (75.66%)	34.51	45.13
#14	2452	1851 (75.48%)	32.81	54.16
#15	2495	1874 (75.11%)	34.39	93.28
#16	2474	1817 (73.44%)	32.62	84.70
#17	2481	1814 (73.11%)	35.45	68.71
#18	2471	1875 (75.88%)	36.46	56.64
#19	2478	1904 (76.83%)	37.93	30.72
$\mu \pm \sigma$	2489 ± 23	1867 ± 29 (75 ± 1)	34.9 ± 1.9	57 ± 15.7

TAB. 7.8 – Nettoyage Condensation $\mathcal{S}_*$

Animal	$\sharp\{S\}$	$\sharp\{N\}$	Quadratique		MLP	
			err_A (%)	err_T (%)	err_A (%)	err_T (%)
#1	2499	224 (8.96%)	27.77	22.06	22.97	15.17
#2	2504	253 (10.10%)	27.08	10.71	26.47	7.85
#3	2497	224 (8.97%)	29.16	29.93	29.41	24.48
#4	2520	230 (9.12%)	34.37	33.87	26.51	19.35
#5	2493	235 (9.42%)	32.43	21.19	30.71	19.86
#7	2482	251 (10.11%)	30.76	31.48	27.34	19.75
#8	2494	196 (7.85%)	25.78	30	30.14	16.66
#9	2490	220 (8.83%)	21.29	46.75	29.62	38.31
#10	2437	222 (9.10%)	34.84	25.12	34.55	13.04
#12	2506	213 (8.49%)	34.28	34.78	28.90	27.53
#13	2531	233 (9.20%)	31.81	24.77	25	15.92
#14	2452	230 (9.38%)	28.03	44.27	27.85	40.10
#15	2495	244 (9.77%)	31.09	58.38	19.28	20.13
#16	2474	226 (9.13%)	25	31.17	22.22	55.88
#17	2481	222 (8.94%)	35.13	28.83	27.85	11.04
#18	2471	246 (9.95%)	27.70	37.57	27.67	28.90
#19	2478	210 (8.47%)	30.64	52.40	23.33	28.91
$\mu \pm \sigma$	2489 ± 23	228 ± 15 (9.2 $\pm$ 0.6%)	29.8 ± 3.9	33.1 ± 12	27 ± 3.7	23.7 ± 12.7

TAB. 7.9 – Nettoyage Condensation $\mathcal{S}_P$

Animal	$\sharp\{B\}$	$\sharp\{N\}$	Quadratique		MLP	
			err_A (%)	err_T (%)	err_A (%)	err_T (%)
#1	2499	375 (15.00%)	36.57	69.65	39.28	20
#2	2504	337 (13.45%)	31.70	35.71	43.61	50.71
#3	2497	337 (13.49%)	33.33	37.41	33.82	40.81
#4	2520	363 (14.40%)	38.39	50.80	42.5	43.54
#5	2493	365 (14.64%)	35.83	41.05	34.82	29.80
#7	2482	371 (14.94%)	39.40	51.85	34.42	51.23
#8	2494	330 (13.23%)	35.86	48	33.72	39.33
#9	2490	338 (13.57%)	35.93	46.75	29.71	46.75
#10	2437	318 (13.04%)	39	57.97	31.01	57.48
#12	2506	343 (13.68%)	34.88	41.30	29.62	23.91
#13	2531	347 (13.70%)	42.41	50.44	42.04	26.54
#14	2452	342 (13.94%)	37	50.52	38.69	57.29
#15	2495	343 (13.74%)	35	92.61	53.27	42.28
#16	2474	363 (14.67%)	36.05	88.82	36.87	60
#17	2481	338 (13.62%)	38.63	63.80	37.22	30.06
#18	2471	349 (14.12%)	37.75	56.64	28.19	27.74
#19	2478	352 (14.20%)	38.09	40.36	33.17	25.90
$\mu \pm \sigma$	2489 ± 23	348 ± 15 (14 $\pm$ 0.6%)	36.8 ± 2.5	54.3 ± 16.4	36.6 ± 6.3	39.6 ± 13

7.6 Conclusion du chapitre

Nous avons vu qu'un apprentissage supervisé des classes de choc hémorragique conduisait à une performance de 40% d'erreur pour un ensemble de caractéristiques portables. L'erreur du système chute à 20% lorsqu'on a accès aux caractéristiques invasives. Cette erreur de 20% est encore obtenue lorsqu'on utilise un ensemble de caractéristiques portables mais sur un problème à deux classes en regroupant $\{\omega_1, \omega_2\}$ et $\{\omega_3, \omega_4\}$.

La sélection ascendante nous indique que : les caractéristiques *PAM*, *PAS*, *FC* et *T_{cut}* sont à choisir lorsque nous pouvons accéder aux données invasives et portables ; les caractéristiques *FC*, *RR_{Abd}*, *T_{cut}* et *SpO₂* sont à retenir lorsque nous avons accès uniquement aux données portables.

La comparaison des classifieurs nous indique que le perceptron multicouches obtient les meilleures performances. Cependant, les performances obtenues par les autres classifieurs s'en rapprochent et permettent de déterminer les choix à réaliser sur certains paramètres (paramètre *k* des plus proches voisins, coefficient α de Parzen etc.).

L'utilisation d'une analyse dans le pire cas en utilisant la technique du leave one (animal) out nous permet d'obtenir une erreur pessimiste, mais qui place le système dans un contexte d'utilisation réaliste, c.-à-d. lorsque le système est utilisé sur un individu inconnu. Cette analyse permet encore de mettre en évidence le comportement particulier de certains individus dans notre base de données et indique qu'il faudrait utiliser plus d'individus pour pouvoir obtenir une estimation moins biaisée.

Enfin, un rejet des données en ambiguïté et en distance ainsi qu'un nettoyage des données ne semble pas améliorer les performances du système. Ces techniques mettent en évidence la difficulté liée aux ambiguïtés de notre problème de classification.

Chapitre 8

Visualisation de l'évolution de l'état de choc hémorragique

*Lorsque les prémisses de cette nouvelle approche eurent été formulées, je découvris de façon inattendue qu'elles constituaient la négation des vieilles « lois de la pensée » et le fondement d'un système non-aristotélicien. J'ai nommé « Sémantique générale » son *modus operandi*. Les prémisses en sont très simples et peuvent s'énoncer au moyen d'une analogie :*

- Une carte n'est pas le territoire. (les mots ne sont pas les choses qu'ils représentent.)*
- Une carte ne couvre pas tout le territoire (Les mots ne peuvent pas couvrir tout ce qu'ils représentent.)*
- Une carte est auto-réflexive. (Dans le langage nous pouvons parler à propos du langage.)*

Alfred Korzybski, Une carte n'est pas le territoire, p. 64, éditions de l'éclat, 2001

Dans ce chapitre, nous présentons les résultats graphiques obtenus d'après les classificateurs appris sur les données expérimentales. Les différentes régions correspondant aux différentes classes de choc hémorragique permettent de mettre en évidence l'évolution de l'état de santé des animaux dans ces espaces multidimensionnels. Dans un premier temps, nous présentons une approche liée à une représentation tridimensionnelle permettant de visualiser l'ensemble des trajectoires. Dans un second temps, nous proposons une représentation bidimensionnelle par projection sur un hyperplan. La succession des réalisations des valeurs permet d'obtenir une séquence correspondant à l'évolution de l'état de santé (cinématique, dessin animé, film).

8.1 Maillage de l'espace

En effectuant un certain maillage de l'espace, on peut obtenir un ensemble de points espacés régulièrement les uns des autres. A chaque point, on peut associer le résultat de la classification obtenue par un des modèles induit par apprentissage. On peut faire l'hypothèse que chaque point représente alors le sommet d'un sous espace multidimensionnel et

que la classe qui lui est associée, est associée à l'ensemble des valeurs incluses dans cette région.

On réalise ainsi une cartographie de l'espace qui nous permet de visualiser le résultat de la classification sur l'espace entier. Une nouvelle donnée aura pour affectation la classe de la région dans laquelle elle se situe.

8.2 Espace à 3 dimensions

Le maillage de l'espace consiste en un ensemble de points à 3 dimensions. L'effet de ce maillage a pour effet de découper l'espace en parallélépipèdes rectangles.

Les logiciels informatiques (en particulier Matlab), offrent la possibilité de faire des figures en 3 dimensions. L'utilisation de la couche α permet de jouer sur la transparence des points et permet ainsi d'appréhender pour le mieux une classification dans l'espace.

Le résultat d'un classifieur quadratique sur les caractéristiques de FC , PAM , T_{int} est proposé Fig. 8.1. L'évolution des différents animaux est représentée par les diverses lignes brisées.

8.3 Espace à 2 dimensions

L'évolution de l'état de l'individu dans un espace à 2 dimensions correspond à une représentation plus habituelle. Pour construire cet espace, on peut réaliser des coupes dans un espace à trois dimensions tel que celui décrit précédemment, orthogonalement à l'une des trois variables. On peut alors projeter l'évolution de la trajectoire de l'animal au cours de l'expérimentation sur ces coupes ainsi définies. Le résultat de tels découpages et de telles projections est proposé Fig. 8.2 pour un animal subissant une perte de sang. Les coupes réalisées sont fonction de la troisième variable qui évolue au cours de l'expérimentation.

Dans cet exemple, on observe le changement de formes des régions de classification au cours de l'expérimentation. Ces changements de formes sont dus aux changements de coupes dans l'espace à 3 dimensions qui dépendent de l'évolution de la troisième variable (T_{cut} dans l'exemple Fig. 8.2). Cette variable semble *cachée* pour un observateur en deux dimensions, mais sa connaissance influence la formation des régions de classification. La prise en considération de cette dernière variable peut donc permettre de réaliser un diagnostic plus juste, un raisonnement moins naïf.

La visualisation des trajectoires à travers les régions de l'espace indique des comportements similaires entre individus : en classe ω_1 et ω_2 , la trajectoire tend à se maintenir dans les régions correspondant à ces classes ; en classe ω_3 , la trajectoire évolue par sauts plus importants, l'état se déplace vers la région de classe ω_4 ; lorsque l'état entre dans la région correspondant à la classe ω_4 , il y a systématiquement inversion de la trajectoire vers le point de fréquence respiratoire nulle et de fréquence cardiaque nulle. L'analyse du comportement de ces trajectoires, répétée d'un individu à l'autre, montre une certaine cohérence de la classification réalisée par la machine qui est moins évidente dans la classification proposée par l'expert.

Nous pensons que les informations représentées sous cette forme permettent de mieux appréhender l'évolution de l'état de santé du blessé. Nous ne pouvons néanmoins, pour le moment, tirer de conclusions sans avoir pu valider et tester ce type de monitoring sur le terrain ou en clinique.

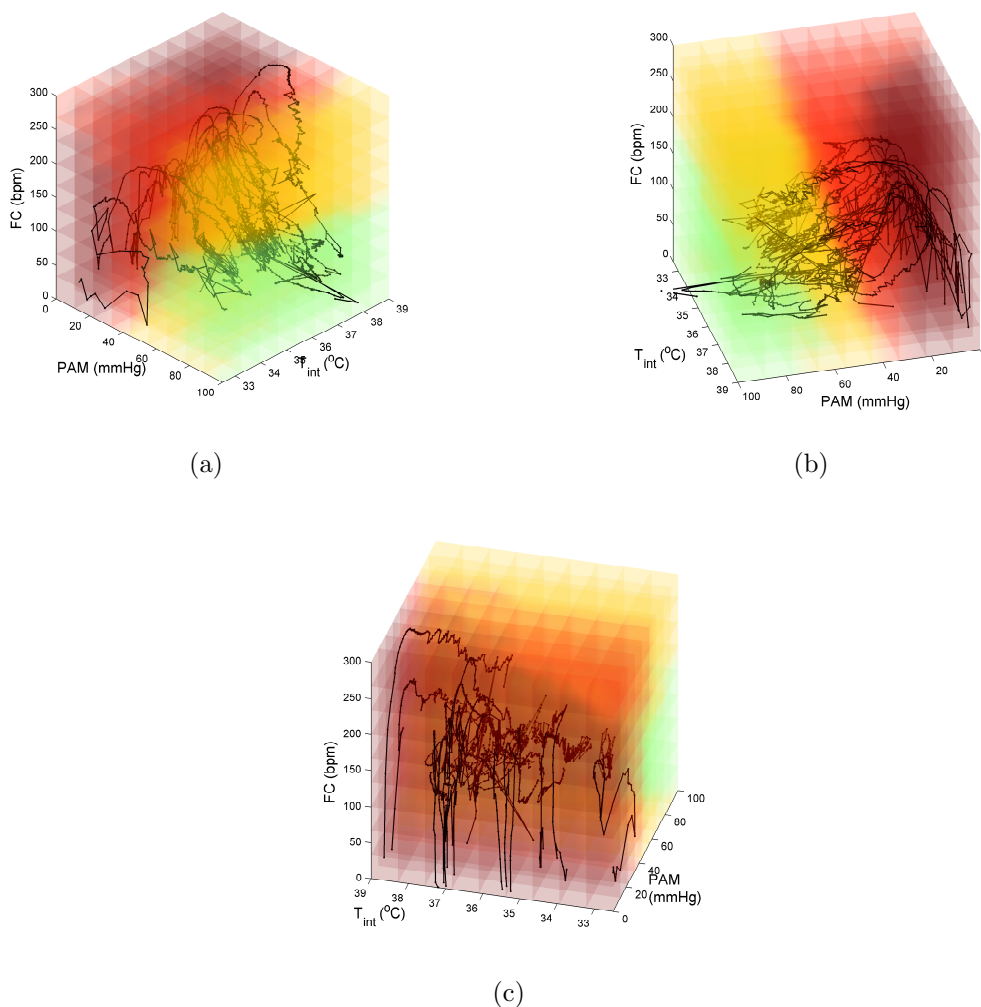


FIG. 8.1 – Évolution dans un espace à 3 dimensions selon trois angles de vue : l'évolution de l'état des animaux est représentée par une ligne brisée. Les animaux traversent les différentes régions associées aux états de choc hémorragique. Le code de couleur retenu est le suivant : vert – classe ω_1 (Rien A Signaler), jaune – classe ω_2 (choc compensé), rouge orangé – ω_3 (choc décompensé), rouge sang – ω_4 (choc irréversible)

8.4 Prise en compte du rejet

On peut mettre en évidence les zones de rejet en distance (transparent) ou les zones de rejet en ambiguïté (bleu clair). Un exemple d'une telle classification dans un espace à 3 dimensions peut être visualisé Fig. 8.3.

Les données de l'ensemble d'apprentissage ne permettent d'obtenir qu'une connaissance insuffisante sur l'ensemble de l'espace et laissent donc un certain vide dans lequel il est difficile de prendre une décision par manque de connaissance : on obtient ainsi une zone de *terra incognita* en ce qui concerne les régions correspondant au rejet en distance. Les régions associées à une ambiguïté dans la prise de décision correspondent aux régions frontalières entre classes, où il est difficile de prendre une décision.

On constate la quasi disparition des classes ω_2 et ω_3 ce qui met en évidence l'ambiguïté existant entre ces deux classes et la bonne discrimination existant entre les classes ω_1 et

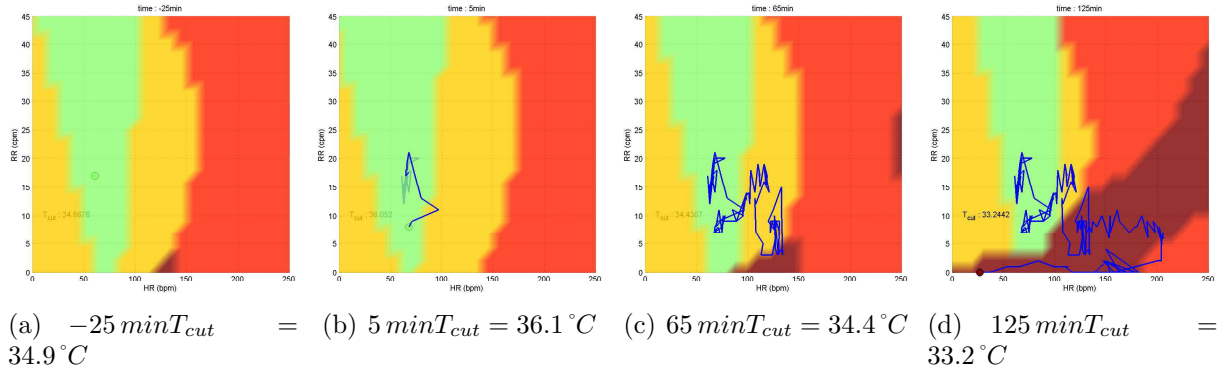


FIG. 8.2 – Évolution du vecteur d'état projeté dans un espace à 2 dimensions (FC en abscisse, RR en ordonnée) pour l'individu #5. Le vecteur d'état évolue dans un espace à 3 dimensions (Fréquence Cardiaque, Rythme Respiratoire, et Température Cutanée) dans lequel on réalise une coupe suivant un plan (Fréquence Cardiaque, Rythme Respiratoire) qui change en fonction de la température cutanée. On projette sur ce plan la trajectoire de l'individu depuis le début de l'expérimentation 25 minutes avant le début de la perte de sang correspondant au temps $t = 0 \min$. Au cours de l'expérimentation, les régions changent en fonction de la variable *cachée* T_{int} .

ω_4 . Cette visualisation confirme l'origine de ces ambiguïtés qui ont été constatées au chapitre 7.

8.5 Conclusion du chapitre

Nous avons vu que les régions de classification, trouvées par la machine par apprentissage supervisé de la classification experte, dans un espace de représentation bien choisi, semblaient bien adaptées à la problématique de l'étude.

Nous avons pu voir que les trajectoires des individus dans cet espace évoluaient en fonction de la région dans laquelle se trouvait le vecteur forme associé à l'état de l'individu. L'observation de comportements similaires entre individus dans ces différentes régions vient confirmer la validité des régions de classification proposée par la machine, pu

Enfin, la visualisation graphique nous a permis de mettre en évidence l'influence des coûts de rejet en ambiguïté, par la disparition des régions correspondant aux classes ω_2 et ω_3 , et de rejet en distance, par la disparition des régions non explorées lors de l'expérimentation (données non disponibles).

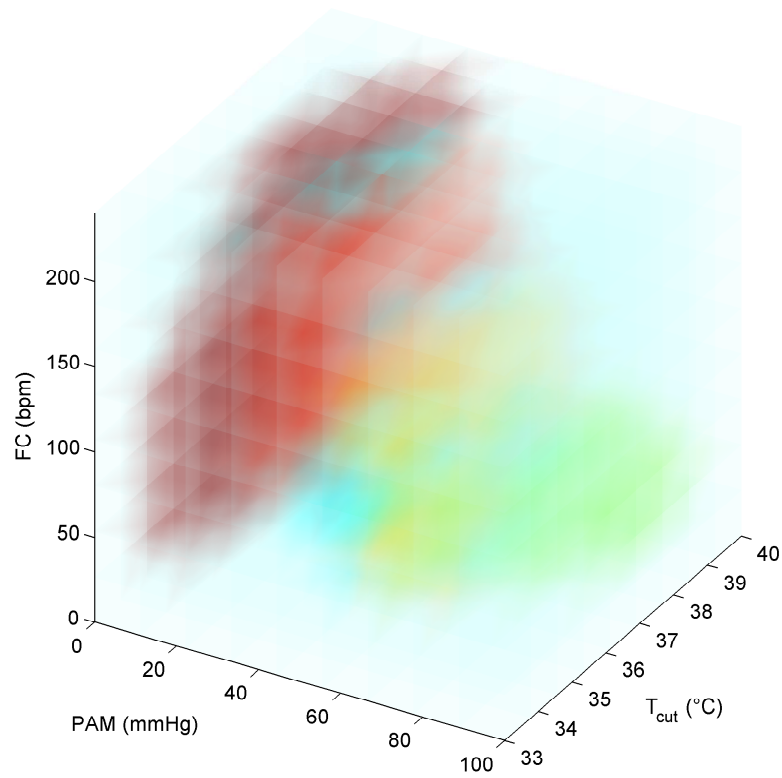


FIG. 8.3 – Espace à 3 dimensions et rejet en distance et en ambiguïté (régions en bleu clair transparentes) pour un seuil de rejet en distance $Cd = 0.01$ et un seuil de rejet en ambiguïté fixé à $Cr = 0.4$. Les régions de décision ainsi déterminées sont à comparer avec celles obtenues Fig. 8.1 a).

Chapitre 9

Prise en compte du temps dans la décision

Dans ce chapitre, nous évaluons l'influence de la prise en compte d'une information dynamique, c.-à-d., liée au temps. Dans un premier temps (9.1), nous évaluerons l'influence de la granularité temporelle des fenêtres de calcul des caractéristiques sur la classification. Dans un second temps (9.2), nous évaluerons la prise en compte de caractéristiques dynamiques, calculées suivant plusieurs méthodes.

9.1 Changement de granularité

Les résultats présentés précédemment ont été obtenus avec des caractéristiques calculées sur des fenêtres temporelles de $\Delta t = 60\text{ s}$. On peut se demander si l'utilisation d'un fenêtrage différent induit un résultat différent. Nous avons donc comparé les résultats obtenus en utilisant les quatre meilleures caractéristiques retenues ($\mathcal{S}_* : \{PAM, PAS, FC, T_{cut}\}$, $\mathcal{S}_P : \{FC, T_{cut}, SpO_2, RR_{Abd}\}$), un classifieur quadratique et des fenêtres temporelles de 10 s , 20 s , 30 s et 60 s . Les sorties des classifieurs ont été comparées aux classifications de l'expert (réalisées sur 60 s) segmentées suivant la granularité temporelle Δt de l'étude (p. ex. pour une fenêtre d'étude de 10 s , chaque classification réalisée sur 60 s par l'expert a été découpée en 6 classes successives identiques).

Les résultats obtenus sont présentés Fig. 9.1. Un test de comparaison de rang signé (test de Wilcoxon) a été réalisé et jugé significatif au seuil $\alpha = 5\%$ ($\dagger$ indique une significativité statistique sur la figure). Les résultats indiquent que l'utilisation d'une fenêtre avec $\Delta t = 60\text{ s}$ conduit aux meilleures performances du système, et ceci pour un ensemble de caractéristiques invasives ou portables. Ce résultat nous permet de confirmer l'utilisation de cette granularité lors de la première partie de l'étude et conforte une meilleure estimation des paramètres du modèles (en particulier respiratoire). En utilisant cette granularité, la base d'apprentissage et de test contient moins d'échantillons, ce qui est plus pratique pour les calculs réalisés, et conduit à des résultats plus performants. Néanmoins, on augmente le temps de réponse du système en utilisant cette fenêtre de longue durée (une réponse toutes les minutes) et l'on induit ainsi un retard à la détection (une fenêtre glissante pourrait permettre de lisser cet effet de retard).

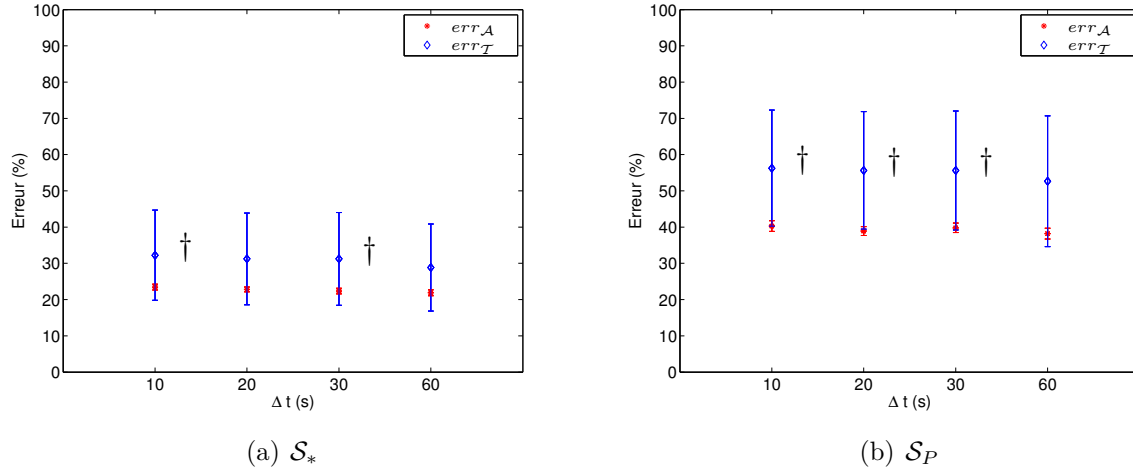


FIG. 9.1 – Granularité temporelle : Effet de la modification de la largeur temporelle des fenêtres pour le calcul des caractéristiques sur la classification réalisée par un classifieur quadratique. a) sur un ensemble $\mathcal{S}_*$: $\{PAM, PAS, FC, T_{cut}\}$, b) $\mathcal{S}_P$: $\{FC, T_{cut}, SpO_2, RR_{Abd}\}$. Une différence significative au seuil $p = 0.05$ entre les données traitées et celles obtenant l'erreur en test err_T la plus faible, est représenté par le symbole $\dagger$.

9.2 Caractéristiques dynamiques

L'ajout de caractéristiques dynamiques complexifie redoutablement l'étude du système. Il existe une complication inhérente à la structure des modèles dynamiques et à leur identification. En particulier, l'estimation de paramètres dynamiques devient difficile : la granularité de nos échantillons temporels et la possibilité réduite d'utiliser la notion de limite ne permettent pas de calculer la vitesse d'évolution d'une caractéristique à proprement parler mais elle va dépendre du grain de la fenêtre temporelle sur laquelle on observe celle-ci et d'une dimension (dite de plongement) dans laquelle on fait plonger notre modèle initial déterminé à un instant (statique). Dans le cas de séries temporelles, l'étude d'une vitesse moyenne sur une durée $D = K\Delta t$ revient à étudier la phase d'une caractéristique à l'instant $t_k - K\Delta t$. On peut en faire la démonstration suivante. Pour chaque caractéristique, on a une vitesse instantanée :

$$v_j(t) = \lim_{h \rightarrow 0} \frac{1}{h} d(x_j(t), x_j(t+h)) = \dot{x}_j(t) \quad (9.1)$$

Dans le cas d'une série temporelle, on peut réaliser l'estimation suivante :

$$\hat{v}_j(t_k, \Delta t) = \frac{1}{\Delta t} (x_j(t_k + \Delta t) - x_j(t_k)) \quad (9.2)$$

Un calcul moyen sur K mesures tel que $D = K\Delta t$ donne :

$$\bar{v}_j(t_k, D) = \frac{1}{K} \sum_{i=1}^K \hat{v}_j(t_k + (i-1)\Delta t, \Delta t) \quad (9.3)$$

$$\bar{v}_j(t_k, D) = \frac{1}{K} \frac{1}{\Delta t} \sum_{i=1}^K (x_j(t_k + i\Delta t) - x_j(t_k + (i-1)\Delta t)) \quad (9.4)$$

En réduisant de termes en termes on obtient :

$$\bar{v}_j(t_k, D) = \frac{1}{D} (x_j(t_k + K\Delta t) - x_j(t_k)) = \frac{1}{D} (x_j(t_k + D) - x_j(t_k)) \quad (9.5)$$

soit :

$$\bar{v}_j(t_k, D) = \hat{v}_j(t_k, D) \quad (9.6)$$

Le calcul de la vitesse instantanée moyenne de cette façon revient donc à considérer uniquement la caractéristique finale $t_k + D$ et celle initiale à l'instant t_k . Il suffit donc d'étudier la *phase* de la caractéristique x_j à l'instant D , sans ajouter de calculs supplémentaires :

$$phase(t_k, D) = x(t_k + D) \quad (9.7)$$

Pour obtenir une moyenne qui a un sens, on peut étudier la vitesse quadratique moyenne obtenue sur chaque caractéristique retenue¹. Pour cela, on peut calculer la moyenne des vitesses au carré. On a alors en prenant $D = K\Delta t$ et en notant $vqm_j(t_k, D)$, la vitesse quadratique moyenne de la caractéristique j à l'instant t_k sur un temps d'observation D :

$$vqm_j(t_k, D) = \sqrt{\frac{1}{K} \sum_{i=1}^K v_j^2(t_k + i\Delta t, \Delta t)} \quad (9.8)$$

Enfin, on peut calculer une pente de régression sur l'ensemble des points sur l'intervalle $[t_k, t_k + D]$ pour obtenir une estimation de la tendance moyenne du signal. La minimisation de l'écart quadratique moyen par la méthode des moindres carrés donne :

$$regress_j(t_k, D) = \frac{12}{K(K+1)(K-1)} \sum_{i=1}^K ix_j(t_k + i\Delta t) - \frac{6(K+1)}{K(K+1)(K-1)} \sum_{i=1}^K x_j(t_k + i\Delta t) \quad (9.9)$$

9.2.1 Transformations

Nous avons effectué des transformations vers des distributions normales de façon à obtenir une distribution qui se rapproche le plus d'une distribution de Laplace-Gauss de façon à travailler dans un espace normé le plus homogène possible. Les effets d'une telle transformation sur la vitesse quadratique moyenne de la fréquence cardiaque sont représentés Fig. 9.2.

Les transformations retenues ont été sélectionnées visuellement, en observant l'effet des fonctions présentées sec. 5.1.2, pour chaque dimension de plongement variant de 0 à 20 minutes. L'ensemble des transformations retenues est proposé dans le tableau Tab. 9.1. Les caractéristiques testées sont celles obtenues précédemment associées à leur caractéristique dynamique (p. ex. pour la caractéristique dynamique *phase* : $\mathcal{S}_* = \{PAM, SpO_2, FC, PAS, phasePAM, phaseSpO_2, phaseFC, phasePAS\}$, $\mathcal{S}_P = \{FC, T_{cut}, RR_{Abd}, SpO_2, phaseFC, phaseT_{cut}, phaseRR_{Abd}, phaseSpO_2\}$.)

¹Il ne faudra pas confondre avec la vitesse quadratique moyenne du vecteur d'état (ou du vecteur forme) dans l'espace de représentation car celle ci utilise une mesure multidimensionnelle

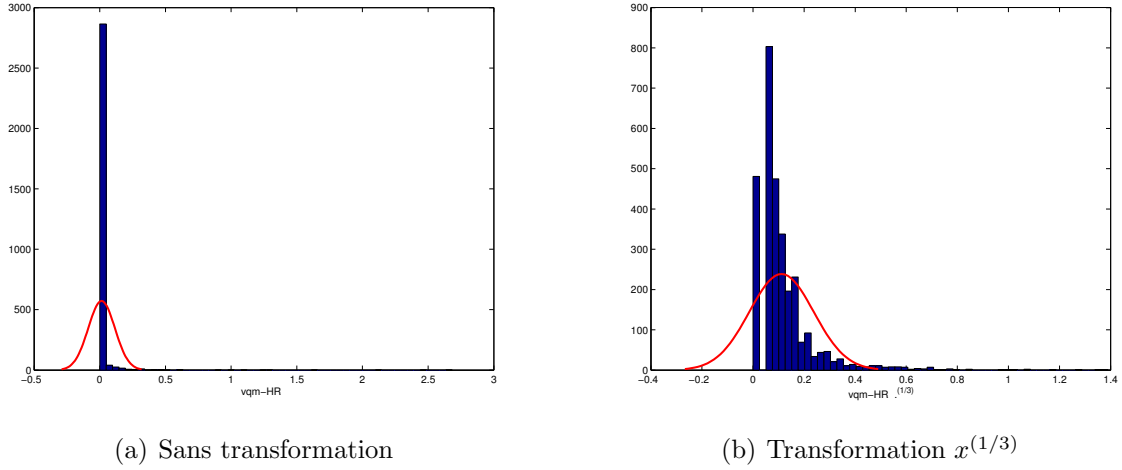


FIG. 9.2 – Effet de la transformation $x^{(1/3)}$ vers une distribution normale pour la vitesse quadratique moyenne de la fréquence cardiaque ($vqmFC$) : a) la caractéristique suit une loi de type exponentielle qui ne s'ajuste pas à une loi de Laplace-Gauss (courbe en rouge) b) une transformation $x^{(1/3)}$ permet d'ajuster au mieux cette variable sur une loi de Laplace-Gauss.

TAB. 9.1 – Transformation pour les caractéristiques dynamiques. La *phase* n'a pas subi de transformation.

	<i>vqm</i>	<i>regress</i>
<i>PAM</i>	$\log(x)$	$\text{sign}(x)(\text{abs}(x))^{(1/3)}$
<i>PAS</i>	$x^{(1/10)}$	$\text{sign}(x)(\text{abs}(x))^{(1/10)}$
<i>FC</i>	$x^{(1/3)}$	$\text{sign}(x)(\text{abs}(x))^{(1/3)}$
<i>RR_{Abd}</i>	$x^{(1/3)}$	x
<i>SpO₂</i>	$x^{(1/10)}$	$\text{sign}(x)(\text{abs}(x))^{(1/3)}$
<i>T_{cut}</i>	$x^{(1/3)}$	$\text{sign}(x)(\text{abs}(x))^{(1/3)}$

9.2.2 Performance

Nous avons ajouté les variables dynamiques transformées aux ensembles $\mathcal{S}_*$ et $\mathcal{S}_P$ et avons utilisé un classifieur quadratique pour évaluer les performances. Les résultats obtenus sont présentés (Fig. 9.3) en fonction de la durée D du calcul des caractéristiques dynamiques. L'analyse de ces résultats indique que l'ajout de variables dynamiques dans notre modèle n'améliore pas les performances obtenues ou très légèrement (de façon non significative, test de rang signé de Wilcoxon) pour tous les délais variant de 0 à 20 *min*.

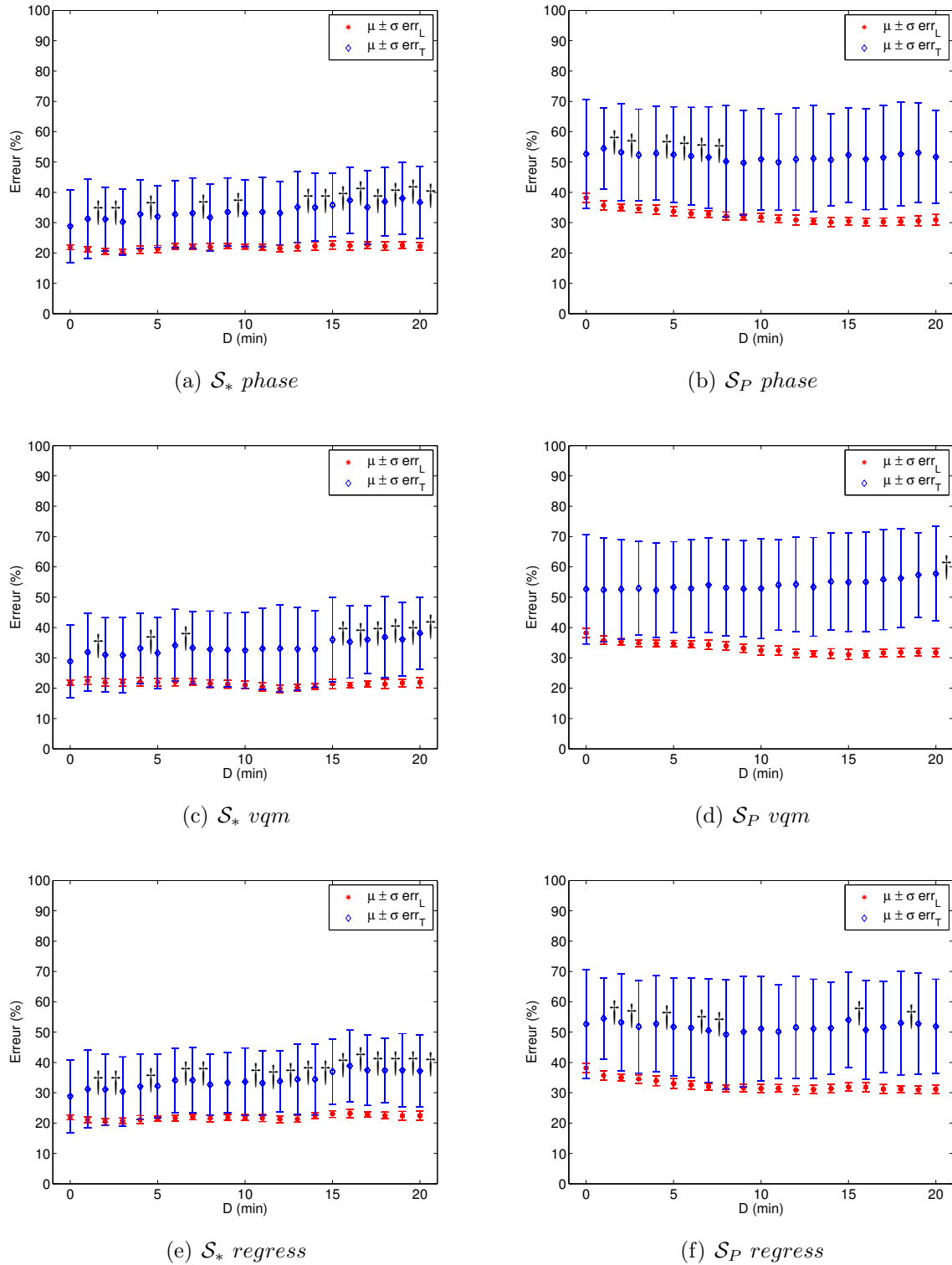


FIG. 9.3 – Performance des caractéristiques dynamiques : Effet du délai sur la classification en utilisant les caractéristiques dynamiques : a) b) de *phase* correspondant à la valeur déphasée de D ; c) d) de *vqm* correspondant à la vitesse quadratique moyenne sur D ; e) f) de *regress* correspondant à la pente de régression calculée sur D . L'emploi de celles-ci ne semblent pas améliorer les performances sur des durées (D) variant de 1 à 20 min. Une différence significative au seuil $p = 0.05$ entre les données traitées et celles obtenant l'erreur en test err_T la plus faible, est représenté par le symbole †.

9.3 Conclusion du chapitre

Nous avons montré que les résultats de la classification obtenue en changeant la granularité temporelle (durée de la fenêtre sur laquelle l'extraction des caractéristiques est réalisée) préconisaient l'emploi d'une fenêtre de 60 s.

Nous avons montré que l'ajout de caractéristiques dynamiques calculées sur des périodes variant de 0 à 20 minutes aux ensembles $\mathcal{S}_* = \{PAM, PAS, FC, T_{cut}\}$ ou $\mathcal{S}_P = \{FC, RR_{Abd}, SpO_2, T_{cut}\}$ ne permettait pas d'améliorer les performances obtenues pour un classifieur quadratique.

Ces résultats sont en accord avec les observations réalisées au chapitre précédent, sur le comportement des trajectoires des individus. Il semblerait que les régions de classification suffisent pour déterminer l'évolution des variables, aussi les informations dynamiques sembleraient redondantes par rapport aux informations apportées par les caractéristiques statiques. Il se peut encore que certaines caractéristiques dynamiques très fluctuantes n'apportent rien à la classification.

Pour conclure sur l'intérêt de certaines caractéristiques dynamiques, une recherche des sous-ensembles obtenant une bonne performance devrait être effectuée sur un ensemble incluant les caractéristiques statiques et dynamiques. L'espace de recherche pour cette sélection de caractéristiques devient alors très grand. Nous n'avons pu réaliser cette étude par manque de temps, mais nous pensons que certains modèles trouvés pourraient permettre d'utiliser au mieux certains capteurs et permettraient peut-être d'en exclure d'autres. Ceci permettrait de diminuer les coûts d'équipement (et le risque de panne) en utilisant par exemple deux capteurs au lieu de trois, en trouvant des combinaisons de caractéristiques qui conduisent aux mêmes performances.

Quatrième partie

Comparaison avec les travaux de
Glass et coll.

Chapitre 10

Article de Glass détaillé

Dans ce chapitre nous développons et commentons l'article de Glass et coll. paru en mars 2004 qui présente un système capable de classer un choc hémorragique en différents paliers de pertes de sang. La démarche adoptée par cette équipe est très proche de la nôtre c'est pourquoi nous présentons celle-ci en commentant les points qui nous paraissent obscurs. Leur classification a l'avantage d'être réalisée à partir de caractéristiques extraites sur un unique électrocardiogramme. La principale critique repose sur leur méthode de validation.

10.1 Présentation de l'article et commentaires associés

L'article de Glass [49] a été publié en Mars 2003 au cours de la troisième année de thèse. Étant donné la forte similarité de nos études, nous avons repris une grande partie de cet article que nous avons commenté (commentaires en italique dans le texte). Des questions ont été posées à son équipe malheureusement sans réponses pour le moment, malgré un premier contact encourageant (cf. Annexe D, dernier envoi datant du 20 mars 2004).

Les résultats indiquent qu'en utilisant uniquement l'ECG, et en calculant sur celui-ci les caractéristiques d'intervalle RR, d'amplitude QRS et d'intervalle RS, on peut créer un modèle capable de diagnostiquer la classe correspondant à la perte de sang.

19 animaux de 25 à 35 kg mâles et femelles ont été utilisés après splénectomie. Une heure de stabilisation a été respectée avant une hémorragie contrôlée réalisée par ponction en palier (aliquots) de 10, 20, 30, 35, 40 45, 50% du volume sanguin total. *La mise en place d'une expérimentation de ce type, avec paliers, a été envisagée lors de la mise en place du protocole mais nous ne l'avons pas retenue. En effet, il nous semble que la succession de ces paliers fait partir l'individu sur une base nouvelle à chaque étape. Il serait intéressant d'étudier ces différents paliers, sur différents individus, en partant de l'état stable de base uniquement. Nous avons retenu une perturbation de type rampe car elle correspond à une réalité physiologique, la phlébotomie ou la perte de sang continue non garrottée d'un individu isolé. Le paramètre de débit de la perte est constant (entrée de type échelon).*

La performance d'adéquation obtenue est en moyenne de 91% (84% à 96% pour un intervalle de confiance de 95%). *Ces résultats paraissent impressionnants comparés aux nôtres. Aucune indication n'est donnée pour le calcul de l'intervalle de confiance à 95%..*

Les auteurs pointent sur les études réalisées avec uniquement des scores d'évaluation à des instants ponctuels et indiquent qu'ils ne connaissent pas d'autres publications sur le suivi de choc hémorragique par des méthodes de classification ou de régression.

Les auteurs ont de suite utilisés les 3 caractéristiques de l'ECG décrites ci dessus. Les ensembles d'apprentissage et de test ont été générés en prenant, pour chaque animal un échantillon aléatoire pour l'ensemble d'apprentissage et un échantillon aléatoire pour l'ensemble de test. Chaque observation correspond à un complexe QRS détecté. L'ensemble d'apprentissage correspond à 50% des données et l'ensemble de test à l'autre moitié des données. Des classifieurs de type kPPV et quadratiques ont été utilisés. La classification a été réalisé sur des blocs de 9 complexes QRS. *Est ce que des segments plus grands ont été recomposés ? Est-ce ceux ci qui servent pour l'apprentissage et le test ou bien les blocs d'un seul complexe ?*

La préparation de l'animal est quasiment identique à la procédure que nous avons utilisée. L'animal est mis à jeun la veille. L'anesthésie est induite par ketamine (20 mg/kg), xylazine (2 mg/kg), et atropine (0.05 mg/kg). L'anesthésie générale lors de la mise en place des instruments se fait sous isoflurane. La ventilation mécanique est effectuée après intubation endotrachéale. Le CO_2 est monitoré. Un cathéter flottant à plusieurs voies (angl. *multilumen pulmonary artery catheter*) est introduit dans la veine jugulaire droite après incision chirurgicale, pour le monitoring de la pression veineuse centrale, du cœur droit, et de la vascularisation pulmonaire (angl. *pulmonary vascular monitoring*). Un cathéter 3 voies de 14–18 cm est placé dans l'artère carotide droite, pour le suivi de la pression artérielle continue et comme accès à l'hémorragie. Un capteur de température est introduite au niveau du rectum. Un électrocardiogramme et un électroencéphalogramme ont été monitoré en continu par des électrodes sous cutanées. Après cette procédure chirurgicale préparatoire, le protocole anesthésique est remplacé par une perfusion (angl. *infusion*) continue de ketamine (2–3 mg/kg/hr) et de fentanyl (15–20 µg/kg/hr) avec un apport d'isoflurane réduit à un taux inférieur ou égal à 0.5% (mélange Air - Oxygène avec $FiO_2 = 35\%$) pour éviter les effets cardiovasculaires dus à l'isoflurane et pour réduire le risque d'hyperthermie maligne. Ensuite, une période d'une heure s'écoule pour permettre une stabilisation physiologique. *On ne sait pas si le choc hémorragique est effectué sous ventilation assistée, mais ceci semble peu probable.*

L'hémorragie est débutée après une période de récupération de 1 heure et 15 minutes pendant laquelle une ligne de base pour la mesure des données physiologiques est réalisée. Une hémorragie artérielle incrémentale a été réalisée via le cathéter carotidien. A chaque palier correspondant à des prélèvements de 10%, 20%, 30%, 35%, 40%, 45% et 50% du volume sanguin total, une période de récupération de 10 minutes a été respectée. Le volume sanguin total a été calculé d'après les normes publiées pour le porc immature (67.3 ml/kg, étendue 58.1–73.9) [124]. Les animaux ont été observés pendant une heure après l'hémorragie complète ou jusqu'à une mort spontanée. Lors de la survie des animaux après une heure, une injection de solution euthanasique commerciale a été administrée.

Les données ont été récupérées en continu par l'intermédiaire d'un appareil de monitoring (Philips, Viridia CMS, M1176A). Les données enregistrées sont les suivantes : fréquence cardiaque, rythme respiratoire, électrocardiogramme, température centrale et rectale, pression veineuse centrale, CO_2 de fin d'expiration, oxymétrie de pouls, saturation en oxygène du sang veineux mêlé (angl. *mixed venous oxygenation saturation*), et électroencéphalogramme. L'ECG a été échantillonné à 500 Hz et toutes les autres données à 125 Hz. La résistance vasculaire systémique, le débit cardiaque par thermodilution, les

gaz sanguins artériels et veineux, ont été réalisés et enregistrés pour déterminer les lignes de base avant l'hémorragie. Ces mesures ont été répétées avant chaque palier au cours de l'hémorragie contrôlée. La mesure du débit cardiaque a été effectuée par un appareil commercial (Philips, Viridia Monitoring System, Thermodilution Module).

Ensuite une analyse par ondelettes a été réalisée (mesure, segmentation et extraction de caractéristique) pour extraire les complexes QRS et découper le signal en segments de longueur variables. La sortie désirée est la classe de perte de sang associée à ce segment. *La manière dont ils reconstituent les segments n'est pas très explicite. Rien n'est dit sur le nombre de voisins utilisés pour développer et tester chaque classifieur de type kPPV.* Les données de chaque animal ont été équitablement divisées, 50% des données pour l'apprentissage et 50% des données pour le test. L'état de la perte de sang a alors été déterminé par une banque (angl. *bank*) de classifieurs, chacun étant réglé sur le profil physiologique de l'animal à un volume hémorragique spécifique. *Il semble que 19 classifieurs construits sur chaque animal, soient mis en compétition (en parallèle), et que l'attribution à une classe soit réalisée sur celui obtenant la probabilité la plus grande.*

L'analyse statistique a été réalisée avec SPSS, SAS, GraphPad. Les mesures de performances standards de sensibilité et de spécificité ont été utilisées pour comparer l'hémorragie désirée et celle obtenue. Les données sont exprimées en pourcentage de bien classés.

Les valeurs obtenues au début et à 50% de la spoliation sont données Tab. 10.1. *les données de température interne seraient intéressantes.*

TAB. 10.1 – Valeurs physiologiques au cours d'une hémorragie contrôlée (Glass et coll.)

	Données (mean $\pm$ SD)	
	Ligne de base	Hémorragie de 50%
Pression Artérielle Moyenne (mmHg)	76 $\pm$ 10	25 $\pm$ 7
Fréquence Cardiaque (bpm)	92 $\pm$ 19	146 $\pm$ 25
Débit Cardiaque (l/min)	3.85 $\pm$ 1	1.4 $\pm$ 0.6
Saturation en oxygène du sang veineux mêlé (%)	60 $\pm$ 16	30 $\pm$ 16

Le système (Cognitive System Technology (CST) suivant leur termes) compare les 3 caractéristiques d'une séquence de 9 P-QRS-T, aux mêmes caractéristiques d'un échantillon qui a servi pour l'apprentissage. 9 valeurs sont donc obtenues pour chaque caractéristique (l'intervalle RR du premier battement est calculé en prenant le dernier battement détecté). La valeur médiane pour chaque caractéristique est alors présentée au classifieur (cette technique permet d'éliminer les artefacts). La sortie du classifieur est utilisée pour étiqueter les 9 valeurs des battements consécutifs. La concordance du système varie de 79% à 96%, selon l'animal, pour un classifieur quadratique et de 81% à 96% pour un classifieur kPPV. En permettant une erreur de 5% sur le volume estimé, la concordance varie de 98 à 100%.

Les différents points suivants sont développés dans la discussion.

Le choc hémorragique n'est pas un événement binaire qui commence à un instant précis. Il est bien connu que le système cardiovasculaire met en place une réponse bien avant que celle ci soit observable. En particulier au niveau cellulaire, il existe très tôt, un large éventail de réponses initiales, qui sont difficilement observables en médecine d'urgence. Les

études sur les mécanismes cellulaires, systémiques, ou moléculaires apportent de nouvelles connaissances mais elles ne permettent pas de réduire considérablement la mortalité en médecins d'urgence (référence sur [125]).

L'intervalle RR (un indicateur de la variabilité cardiaque) a été sélectionné parce que la fréquence cardiaque est régulée par le système nerveux autonome et est spécifiquement sensible à d'important changements de volume de sang. De nombreuses études montrent que les changements de fréquence cardiaque et de régularité sont des reflets d'états mortels. L'amplitude du QRS et l'intervalle RS ont été sélectionnés comme des indices d'énergie électrique totale impliquée dans la dépolarisation du ventricule. Au cours du choc, une ischémie myocardiale s'installe progressivement. Ainsi, l'énergie totale électrique du ventricule diminue. Ou bien, comme alternative, les changements de l'amplitude du QRS seraient dus à un changement de position par rapport aux électrodes alors que le volume intravasculaire diminue. *Nous avons observé ces mêmes phénomènes, ce qui expliquerait les changements de polarisation observé sur les électrocardiogrammes. A un moment, le cœur lâche et revient dans une position normale.*

Les auteurs insistent aussi sur le fait que la plupart des études considèrent les variables comme indépendantes ou utilisent des valeurs moyennes. Ces études font l'hypothèse d'un état physiologique stable (angl. *steady*) ou homéostatique (angl. *homestatic*) qui montre un comportement linéaire lors de stimulation et déclenche des alarmes dès qu'un seuil est dépassé. Par opposition, les auteurs indiquent que leur traitement multidimensionnel et l'étude de composantes non linéaires du signal électrocardiographique sont mieux adaptés aux phénomènes étudiés. *Même si nous sommes d'accord avec ces constatations d'échec des méthodes linéaires ou naïves, leur référence à em nonlinear analysis et "second-order" linear (power spectral) analysis n'est pas claire, nous ne voyons pas comment les non linéarités du phénomène sont mises en évidence par cet article.*

Le protocole anesthésique a été mis en place de façon à réduire l'impact sur la réponse cardiovasculaire en maintenant autant que possible une anesthésie optimale. Mais il existe toujours une action sur le système nerveux autonome. Il a été observé, dans des études préliminaires que l'isoflurane réduisait la réponse tachycardique. La Pression Artérielle de base a été de 15 mmHg plus basse que celle publiée pour des animaux conscients d'âge et de taille similaire. Néanmoins les autres variables ont été dans des limites acceptables. Des profils de réponse à l'hémorragie sont observés sur différents animaux et suivent des tendances similaires dans le temps, pas forcément quantitativement mais qualitativement. Des recherches futures devraient porter sur l'analyse des points d'inflexion et l'étude des trajectoires.

10.2 Conclusion du chapitre

Il est réconfortant de trouver une démarche identique à la nôtre, chez une autre équipe de recherche, pour l'étude d'une problématique qui peut être abordée de diverses façons.

Nous avons vu que de nombreux points, en particulier au niveau du modèle animal et des techniques d'intelligence artificielle utilisée pour créer le système (apprentissage supervisé de classifieur) étaient semblables.

Néanmoins, un grand nombre de différences apparaissent à la lecture de l'article et suscite quelques interrogations.

Chapitre 11

Influence de la validation sur les résultats

Dans le chapitre 7, nous avons présenté les résultats obtenus en utilisant une technique de validation qui semblait bien adaptée à notre problème. À la suite de la lecture de l'article de Glass et coll., nous pouvons nous interroger sur l'influence de la technique de validation sur la présentation des résultats. L'étude de cette influence fait l'objet de ce chapitre qui présente : dans un premier temps, une technique de sous échantillonnage aléatoire ; dans un second temps, une validation animal par animal telle qu'utilisée par l'équipe de Glass.

11.1 Validation par sous échantillonnage aléatoire

Pour évaluer les performances des différents classifieurs sur nos données, on peut réaliser une statistique en effectuant différents tirages aléatoires pour créer l'ensemble d'apprentissage et l'ensemble de test. Cette méthode nous permet d'obtenir une estimation pessimiste de l'erreur de classement des classifieurs (estimation de l'erreur réelle) mais aussi une erreur optimiste (estimation de l'erreur empirique) [126] (cf. appendice A).

En faisant varier le nombre d'échantillons présents dans l'ensemble d'apprentissage et l'ensemble de test, on obtient la figure représentée Fig 11.1 pour l'ensemble des caractéristiques portables. Les échantillons ont été tirés aléatoirement dans la base de données constituées de tous les individus mélangés. Les échantillons ont été choisis tels que ceux présents dans l'ensemble d'apprentissage ne puissent pas être présents dans l'ensemble de test. Le nombre de tirages a été fixé à 30 et ceux ci ont été réalisés de façon à obtenir une répartition équiprobable dans les différentes classes de l'étude, de façon à ne privilégier aucune classe par rapport à l'autre. On remarque qu'après une sélection de 300 échantillons pour l'ensemble d'apprentissage et 300 échantillons pour l'ensemble de test, l'estimation des performances n'est pas améliorée. On retiendra donc ces paramètres de façon à obtenir une évaluation rapide des performances (on évite ainsi de faire trop de calcul).

11.1.1 Sélection ascendante de caractéristiques

En utilisant les critères définis précédemment pour l'évaluation des performances, on procède à la sélection de caractéristiques en utilisant la méthode de Sequential Forward

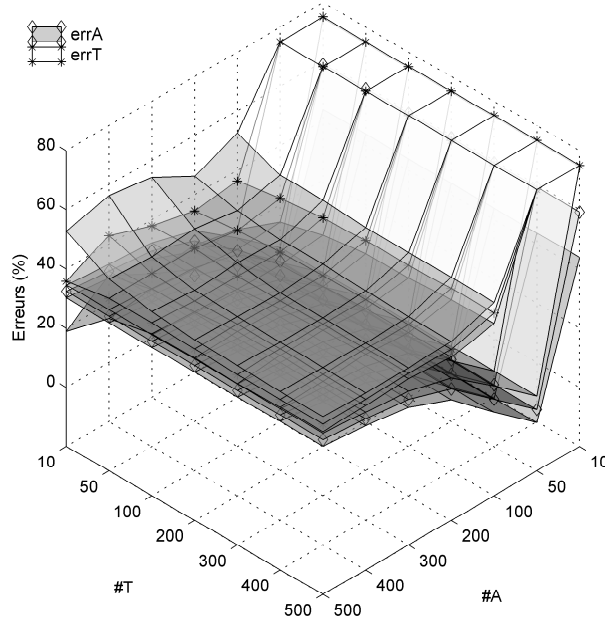


FIG. 11.1 – Évolution de la performance d'un classifieur quadratique en fonction du nombre de représentants pour l'ensemble d'apprentissage ($\sharp\{\mathcal{A}\}$) et du nombre de représentants dans l'ensemble de test ($\sharp\{\mathcal{T}\}$). Les 3 feuillets supérieurs correspondent aux moyennes (*) et moyennes $\pm$ écarts-types obtenues pour l'erreur de test $err_{\mathcal{T}}$. Les 3 feuillets du bas correspondent aux moyennes (◇) et moyennes $\pm$ écarts-types obtenues pour l'erreur d'apprentissage $err_{\mathcal{A}}$.

Selection sur un classifieur quadratique. Les résultats obtenus lorsque toutes les caractéristiques sont disponibles sont proposés Fig. 11.2(a). La première caractéristique retenue est la PAM avec une erreur réelle de $23.7 \pm 2.7\%$. Après l'ajout des caractéristiques de FC et T_{cut} on obtient une erreur de $18.4 \pm 2\%$. Les performances ne sont pas améliorées en ajoutant d'autres caractéristiques.

Les résultats obtenus en ne retenant que les caractéristiques calculées sur des capteurs portables sont représentés Fig. 11.2(b). La première caractéristique retenue est la SpO_2 avec une performance de $54.4 \pm 1.4\%$ d'erreur réelle estimée. L'ajout de la FC et de la T_{cut} fait chuter l'erreur à $42.1 \pm 3.1\%$. En augmentant le nombre de caractéristiques du modèle, l'erreur réelle atteint un minimum de $33.5 \pm 2.7\%$ avec 9 caractéristiques (après les 7 premières, on ne constate pas d'amélioration).

11.1.2 Comparaison des classifieurs

Les comparaisons ont été réalisées en ne retenant que les 3 meilleures caractéristiques invasives et portables retenues par la sélection SFS, car au delà de 3 caractéristiques, la performance n'est pas améliorée ou peu : $\mathcal{S}_* = \{PAM, FC \text{ et } T_{cut}\}$, $\mathcal{S}_P = \{SpO_2, FC \text{ et } T_{cut}\}$. Les résultats sont proposés Fig. 11.3. Le meilleur classifieur obtenu en terme d'erreur réelle est celui des 1PPV (kPPV(1) sur la figure) (le classifieur flou obtenu se comporte identiquement). Quand on regarde l'erreur empirique sur ce classifieur, celle-ci est de 0%. Ceci indique que l'on a effectué un apprentissage par cœur et que nos performances en généralisation peuvent être mises en défaut. On préférera donc les

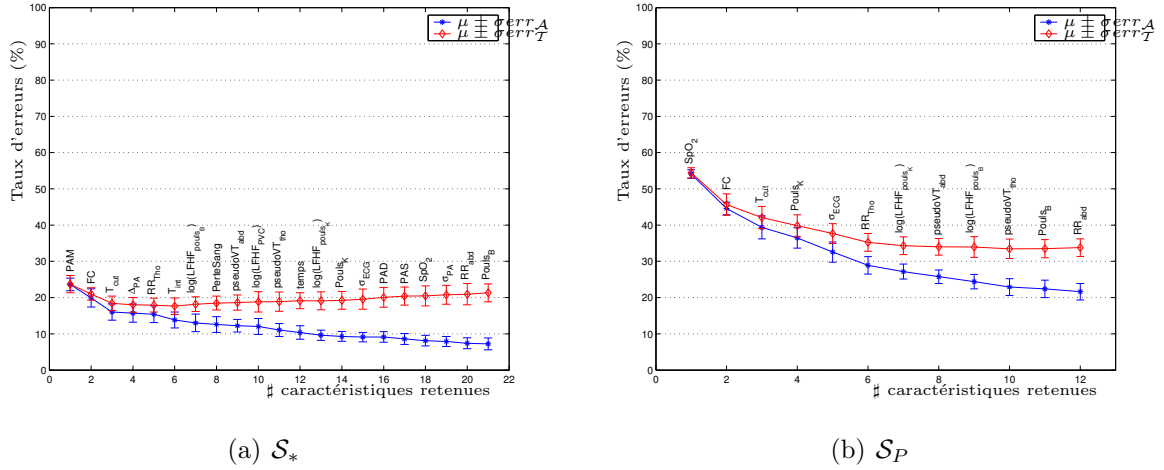


FIG. 11.2 – Sélection de caractéristiques par Sequential Forward Selection (SFS) en utilisant 300 échantillons pour l'ensemble d'apprentissage et 300 échantillons pour l'ensemble de test pour un classifieur quadratique.

classifieurs dont les performances obtenues sur l'erreur d'apprentissage et l'erreur de test sont proches l'une de l'autre qui permettent de prédire le comportement en généralisation. En considérant ce principe, le réseau de neurones permet d'obtenir une erreur empirique de $37.3 \pm 2.4\%$ et une erreur réelle de $34.6 \pm 3.1\%$ sur l'ensemble $\mathcal{S}_P$ par exemple.

En comparant cette étude à celle réalisée avec une méthode de leave one (animal) out, on remarque que les performances obtenues sont meilleures. L'utilisation de données provenant d'un même individu ayant servi pour l'apprentissage et le test permet d'expliquer cette diminution de l'erreur. La proximité des données (provenant du même animal à des instants très proches) est alors mise en évidence par le classifieur 1PPV.

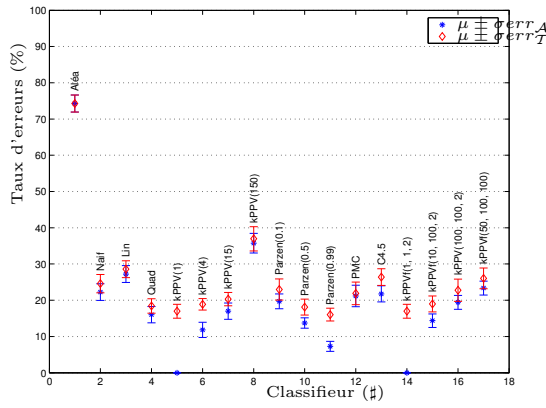
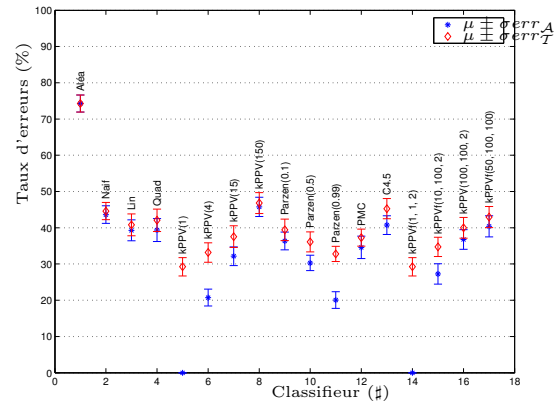
(a) $\mathcal{S}_*$ (b) $\mathcal{S}_P$

FIG. 11.3 – Comparaison des classifieurs lorsque la méthode d'évaluation est réalisée par sous échantillonnage aléatoire de 300 échantillons pour l'ensemble d'apprentissage et l'ensemble de test.

11.2 Validation animal par animal

On peut procéder en utilisant la même méthode que Glass pour la création et la validation des classifieurs. Chaque classifieur est créé à partir d'une moitié des données d'un individu et évalué sur les données restantes du même individu. On utilise donc 50% pour constituer l'ensemble d'apprentissage $\mathcal{A}$ et 50% des données pour l'ensemble de test $\mathcal{T}$, et ceci pour chaque animal. Les résultats obtenus en utilisant les ensembles de caractéristiques invasives et portables $\mathcal{S}_* = \{PAM, PAS, FC, SpO_2\}$, et $\mathcal{S}_P = \{FC, T_{cut}, SpO_2, RR_{Abd}\}$, sont disponibles Tab. 11.1.

TAB. 11.1 – Classification par animal. Problème à 4 Classes suivant le choc. $\mathcal{S}_* = \{PAM, PAS, FC, SpO_2\}$; $\mathcal{S}_P = \{FC, T_{cut}, SpO_2, RR_{Abd}\}$. Les résultats correspondent à la proportion de mal classés sur l'ensemble de test ($err_{\mathcal{T}}$). Cette méthode fournit des résultats optimistes mais elle est biaisée. Elle ne permet pas de conclure sur une généralisation du système.

Animal	$\mathcal{S}_*$		$\mathcal{S}_P$	
	1PPV (%)	Quad (%)	1PPV (%)	Quad (%)
#1	2.61 ± 1.62	7.88 ± 3.46	7.70 ± 2.71	12.21 ± 7.55
#2	8.87 ± 2.19	13.05 ± 3.77	11.13 ± 3.61	17.18 ± 4.02
#3	9.33 ± 2.27	16.80 ± 3.10	18.40 ± 3.92	25.29 ± 3.81
#4	9.63 ± 2.92	25.03 ± 7.14	12.54 ± 3.52	13.44 ± 3.58
#5	18.66 ± 3.45	17.27 ± 3.82	20.65 ± 4.55	15.02 ± 3.67
#7	14.11 ± 3.13	15.00 ± 2.83	20.08 ± 3.50	23.82 ± 3.74
#8	16.36 ± 2.93	21.84 ± 4.35	20.61 ± 3.91	27.11 ± 4.13
#9	12.26 ± 3.95	20.68 ± 5.32	19.36 ± 3.87	26.67 ± 4.78
#10	12.44 ± 2.59	23.78 ± 8.27	16.57 ± 2.59	21.71 ± 5.18
#12	9.24 ± 2.70	12.76 ± 4.53	10.67 ± 3.73	8.38 ± 3.51
#13	12.59 ± 3.68	12.87 ± 4.31	25.06 ± 4.91	15.11 ± 6.00
#14	16.25 ± 3.71	22.99 ± 4.84	23.37 ± 3.26	19.38 ± 4.61
#15	3.07 ± 2.27	4.25 ± 3.84	6.10 ± 2.86	7.06 ± 4.87
#16	4.92 ± 2.15	6.01 ± 2.54	14.92 ± 3.09	8.41 ± 3.69
#17	4.42 ± 2.08	3.90 ± 3.00	6.99 ± 2.18	4.10 ± 3.97
#18	28.11 ± 3.67	28.75 ± 3.38	34.24 ± 4.21	28.41 ± 3.38
#19	15.36 ± 4.85	15.12 ± 3.54	17.38 ± 2.88	14.60 ± 3.69
Moyenne	11.66 ± 6.43	15.76 ± 7.47	16.81 ± 7.29	16.93 ± 7.63

En procédant de cette façon, les échantillons des ensembles d'apprentissage et de tests ne sont pas indépendants l'un de l'autre, car les données sont liées par le temps et l'individu. L'estimation des moyennes et des écart-types a été réalisé selon 30 tirages aléatoires pour la création de $\mathcal{A}$ et $\mathcal{T}$ pour chaque individu et les résultats correspondent à la moyenne et aux écart-types réalisés sur les différents tirages. On constate qu'en utilisant l'algorithme du plus proche voisin, les performances obtenues sont les meilleures avec une erreur moyenne de $11.66 \pm 6.43\%$ lorsqu'on utilise des caractéristiques invasives et $16.81 \pm 7.29\%$ lorsqu'on utilise des caractéristiques portables.

Les résultats obtenus avec le classifieur quadratique sont légèrement moins bons avec une performance de $15.76 \pm 7.47\%$ pour les caractéristiques invasives et $16.93 \pm 7.63\%$ pour les mesures portables. Ils restent néanmoins du même ordre de grandeur. Ceci peut en partie être expliqué par le fait que le classifieur quadratique généralise mieux que le classifieur 1PPV. Le classifieur quadratique utilise plusieurs échantillons d'un individu

propre pour créer sa base de connaissance et généralise mieux que le classifieur par 1PPV qui va fournir la classe de l'échantillon le plus proche qu'il garde en mémoire (par cœur).

Enfin, on constate que les résultats obtenus lorsqu'on a accès à toutes les données (du moins celles retenues) et lorsqu'on utilise un ensemble de mesures portables, sont très comparables en particulier dès que l'algorithme généralise un minimum sa base de connaissance (en utilisant le classifieur quadratique p. ex.). Ceci nous indique que pour chaque individu, l'utilisation d'une classification basée sur des caractéristiques invasives ou portables conduit à la même performance. Ceci joue en faveur de notre hypothèse de départ concernant la validité d'un modèle portable. Néanmoins, l'utilisation d'une telle technique ne vérifie pas l'hypothèse d'indépendance des données pour l'ensemble d'apprentissage et l'ensemble de test, car le même animal est son propre référentiel et les échantillons temporels restent proche entre eux. On peut encore se demander quelle est l'influence de l'ensemble des mesures dans ce cas, puisque les résultats obtenus avec des ensembles différents conduisent aux mêmes résultats. Ceci nous conduit encore à expliquer l'erreur obtenue relativement à la nature même de la classification réalisée par l'expert.

11.3 Conclusion du chapitre

Nous remarquons que suivant la méthode que nous utilisons pour valider le modèle, nous obtenons des valeurs remarquables, mais qui ne correspondent qu'à une réalité limitée. D'une part un tirage aléatoire pour créer l'ensemble d'apprentissage et l'ensemble de test nous conduit à évaluer l'erreur d'une façon légèrement biaisée pour notre base de données constituée d'un faible nombre d'individus (17 animaux). D'autre part, le calcul d'une erreur par rapport à son propre référentiel individuel nous conduit à une estimation fortement biaisée, puisque celle-ci est effectuée dans le meilleur cas, c.-à-d. lorsque les données de test sont proches des données ayant servi pour l'apprentissage.

Il semble que ce soit néanmoins la méthode retenue par Glass et coll. pour évaluer son système. On peut alors se demander ce que nous obtenons avec nos données et les caractéristiques retenues en utilisant une classification par perte de sang, comme l'ont fait Glass et coll.

Chapitre 12

Classification par perte de sang

Dans ce chapitre, nous avons essayé de comparer nos résultats à ceux de Glass en utilisant une classification par portion de volume sanguin prélevé. La section 12.1 présente les résultats obtenus lorsqu'on utilise les mêmes techniques de validation que celles de Glass. La section 12.2 propose une évaluation par leave one animal out qui nous paraît plus juste pour ne pas rentrer dans un modèle individu dépendant.

12.1 Validation Animal par Animal

La classification a été définie selon des intervalles de définition obtenue sur la caractéristique *PerteSang* c.-à-d., définie selon le volume sanguin prélevé rapporté au volume sanguin total. Les intervalles retenus et leurs correspondances en terme de classes sont représentés Tab. 12.1.

TAB. 12.1 – Définition des classes suivant la perte de sang : les intervalles de définition sont calculés sur la caractéristique *PerteSang* de façon à obtenir la classification réalisée par Glass et coll.

Classe	Glass	Becq
ω_1	0%	$PerteSang < 5\%$
ω_2	10%	$5\% \leq PerteSang < 15\%$
ω_3	20%	$15\% \leq PerteSang < 25\%$
ω_4	30%	$25\% \leq PerteSang < 32.5\%$
ω_5	35%	$32.5\% \leq PerteSang < 37.5\%$
ω_6	40%	$37.5\% \leq PerteSang < 42.5\%$
ω_7	45%	$42.5\% \leq PerteSang < 47.5\%$
ω_8	50%	$47.5\% \leq PerteSang$

Les résultats obtenus en utilisant une validation animal par animal, comme présentée en 12.1 pour une classification suivant quatre classes, sont rapporté Tab. 12.2 pour la classification suivant huit classes directement calculées sur la perte de sang.

12.1.1 Analyse de l'erreur sur l'ensemble de test

L'erreur obtenue, en utilisant l'ensemble des caractéristiques invasives retenues dans notre étude ($\mathcal{S}_* = \{PAM, PAS, FC, SpO_2\}$), est de $14.04 \pm 5.58\%$ en moyenne ($\mu \pm \sigma$) pour le classifieur du plus proche voisin mais s'élève à $26.58 \pm 3.05\%$ dès qu'on utilise un

classifieur qui généralise l'apprentissage comme le classifieur quadratique. Avec le 1PPV, On est donc en présence d'un apprentissage par cœur dont la performance sera difficile à améliorer.

L'erreur obtenue en utilisant l'ensemble des caractéristiques portables retenues dans notre étude ($\mathcal{S}_P = \{FC, T_{cut}, SpO_2, RR_{Abd}\}$), est de $21.22 \pm 6.56\%$ en moyenne pour le classifieur 1PPV et de $26.98 \pm 4.43\%$ en moyenne pour le classifieur quadratique. Les meilleures performances individuelles présentent une erreur de 11% et les moins bonnes une erreur de 36%. Ces résultats sont proches de ceux obtenus avec le classifieur quadratique sur les données invasives et indiquent que les modèles obtenus avec des caractéristiques portables peuvent être compétitifs avec des modèles obtenus sur des caractéristiques invasives, en considérant une analyse au cas par cas individuel.

TAB. 12.2 – Classification par animal. Problème à 8 Classes suivant la perte de sang

Animal	$\mathcal{S}_*$		$\mathcal{S}_P$	
	1PPV (%)	Quad (%)	1PPV (%)	Quad (%)
#1	07.03 $\pm$ 2.42	24.91 $\pm$ 5.86	16.71 $\pm$ 3.50	32.93 $\pm$ 5.85
#2	11.74 $\pm$ 4.74	26.62 $\pm$ 7.48	11.22 $\pm$ 4.16	29.15 $\pm$ 7.68
#3	14.49 $\pm$ 2.94	27.07 $\pm$ 7.25	17.16 $\pm$ 4.08	28.00 $\pm$ 5.97
#4	12.01 $\pm$ 4.01	31.90 $\pm$ 7.80	17.04 $\pm$ 4.08	31.06 $\pm$ 7.87
#5	13.38 $\pm$ 3.64	27.27 $\pm$ 6.83	16.19 $\pm$ 3.52	24.94 $\pm$ 7.71
#7	12.36 $\pm$ 3.50	24.51 $\pm$ 8.38	18.70 $\pm$ 4.63	22.20 $\pm$ 6.95
#8	11.14 $\pm$ 3.01	28.55 $\pm$ 5.49	16.93 $\pm$ 4.08	28.60 $\pm$ 7.71
#9	12.14 $\pm$ 2.97	26.75 $\pm$ 6.24	32.82 $\pm$ 4.42	35.38 $\pm$ 6.34
#10	16.10 $\pm$ 2.65	24.16 $\pm$ 4.04	17.90 $\pm$ 3.31	24.76 $\pm$ 5.01
#12	18.57 $\pm$ 4.29	28.00 $\pm$ 7.92	24.76 $\pm$ 5.11	23.10 $\pm$ 7.73
#13	12.18 $\pm$ 4.68	29.89 $\pm$ 8.07	17.47 $\pm$ 5.81	29.02 $\pm$ 5.94
#14	15.36 $\pm$ 3.94	21.24 $\pm$ 6.85	21.82 $\pm$ 3.87	21.51 $\pm$ 8.40
#15	12.02 $\pm$ 3.69	26.18 $\pm$ 7.92	18.95 $\pm$ 3.16	26.67 $\pm$ 6.10
#16	20.78 $\pm$ 4.59	25.54 $\pm$ 5.41	29.84 $\pm$ 3.50	24.73 $\pm$ 5.03
#17	07.91 $\pm$ 3.52	22.37 $\pm$ 4.55	24.38 $\pm$ 4.19	21.12 $\pm$ 6.39
#18	31.14 $\pm$ 5.10	32.65 $\pm$ 5.63	36.02 $\pm$ 3.62	33.30 $\pm$ 5.73
#19	10.28 $\pm$ 3.97	24.25 $\pm$ 5.90	22.78 $\pm$ 3.90	22.26 $\pm$ 5.02
Moyenne	14.04 $\pm$ 5.58	26.58 $\pm$ 3.05	21.22 $\pm$ 6.56	26.98 $\pm$ 4.43

12.1.2 Matrice de confusion - Perte de sang

Un exemple de matrice de confusion obtenue après apprentissage de la moitié des données et testé sur la seconde moitié (ensembles d'évaluation obtenus par tirage aléatoire, la classification étant réalisée en utilisant le classifieur 1PPV), est présenté Tab. 12.3 et indique que les principales erreurs sont obtenues pour les classes intermédiaires, mais que la classification est bonne pour les classes marginales correspondant à 0% et 50% avec 94% de concordance pour cette dernière classe. On constate que les classes intermédiaires sont constituées de moins de données. De plus, notre spoliation par perte de sang à débit contrôlé va à l'encontre de cette classification par paliers de prélèvement, (ce qui n'est pas le cas dans le modèle de choc de Glass). Ainsi, les mauvaises classifications sont pour la plupart due aux ambiguïté des classes adjacentes (données figurant juste en dessous ou au dessus de la diagonale). On reste néanmoins assez fidèle pour un diagnostic. On constate une rupture intéressante qui se situe au niveau de 40% de *PerteSang* qui

pourrait indiquer un comportement particulier au delà de cette valeur et ceci pour chaque animal pris individuellement. Ce comportement semble être confirmé généralement par cette matrice.

TAB. 12.3 – Exemple de table de confusion obtenue pour un modèle quadratique pour une classification selon huit classes correspondant à des intervalles de volume sanguin prélevé. La classification a été obtenue sur l'ensemble de caractéristiques portables $\mathcal{S}_*$. L'évaluation a été réalisée animal sur animal.

<i>PerteSang</i>	Machine								Acc (%)
	0%	10%	20%	30%	35%	40%	45%	50%	
0%	278	37	13	3	0	0	0	0	83.99%
10%	25	106	20	1	1	0	0	0	69.28%
20%	7	15	118	15	3	0	0	0	74.68%
30%	0	4	13	96	15	2	1	0	73.28%
35%	0	0	1	17	52	4	2	0	68.42%
40%	0	0	0	2	9	56	10	3	70%
45%	0	0	0	1	2	8	58	11	72.5%
50%	2	0	0	1	1	2	14	314	94%

12.2 Évaluation par Leave One (Animal) Out

Les classes sont définies telles que présentées Tab. 12.1. Une classification selon la technique du leave one (animal) out a été réalisée sur l'ensemble des données $\mathcal{S}_*$ et $\mathcal{S}_P$ après sélection d'un ensemble équiprobable de classes pour la création de l'ensemble d'apprentissage.

12.2.1 Estimation de l'erreur réelle

Les résultats obtenus sur les ensembles de test uniquement (err_T) sont représentés Tab. 12.4. L'erreur obtenue, en utilisant l'ensemble des caractéristiques invasives retenues $\mathcal{S}_*$, est de $61.26 \pm 13.82\%$ en moyenne pour le classifieur du plus proche voisin. L'erreur diminue pour le classifieur quadratique et chute à $58 \pm 16.22\%$ en moyenne, mais un test de Wilcoxon ne conclut pas sur la significativité de ce résultat ($\alpha = 0.7$).

L'erreur obtenue en utilisant l'ensemble des caractéristiques portables retenues, est de $73.02 \pm 12.93\%$ en moyenne pour le classifieur 1PPV et de $65.19 \pm 12.82\%$ en moyenne pour le classifieur quadratique. Un test de Wilcoxon devient significatif pour un seuil $\alpha > 0.07$. On note encore que les meilleures performances individuelles présentent une erreur de 39% et les moins bonnes une erreur de 87% (un classement aléatoire produirait une erreur de $(K - 1)/K = 7/8$ soit 87.50%) pour les données portables.

Nous pouvons conclure de ces résultats qu'il n'est pas possible de réaliser une classification selon la perte de sang subie ou que celle-ci n'est pas bien adaptée au problème d'hypovolémie, même si la prise en considération de plusieurs individus semble améliorer les performances. L'étude de la matrice de confusion permettra de mieux comprendre les erreurs réalisées.

TAB. 12.4 – Classification par animal. Problème à 8 Classes suivant la perte de sang

Animal	$\mathcal{S}_*$		$\mathcal{S}_P$	
	1PPV (%)	Quad (%)	1PPV (%)	Quad (%)
#1	43.92	54.73	62.16	77.03
#2	35.71	36.43	90	59.29
#3	44.9	43.54	68.03	48.3
#4	75	64.52	84.68	58.06
#5	83.44	58.28	81.46	71.52
#7	57.41	37.65	60.49	65.43
#8	48	27.33	67.33	60.67
#9	66.23	61.69	61.04	61.69
#10	66.35	58.17	40.87	39.42
#12	56	35.33	81.33	51.33
#13	84.35	68.7	85.22	84.35
#14	66.84	74.09	69.43	74.61
#15	55.7	79.87	88.59	86.58
#16	70.59	70.59	72.35	72.94
#17	68.71	68.71	68.1	55.83
#18	50.87	70.52	86.13	64.74
#19	67.47	75.9	74.1	76.51
Moyenne	61.26 $\pm$ 13.82	58 $\pm$ 16.22	73.02 $\pm$ 12.93	65.19 $\pm$ 12.82

12.2.2 Matrice de confusion - Perte de sang

La matrice de confusion moyenne réalisée sur l'ensemble des animaux en considérant les caractéristiques portables retenues et obtenu avec un classifieur quadratique est représenté Tab. 12.5.

On constate que la table de confusion peut être séparée en 3 parties. La première zone correspond aux données dont la perte de sang est inférieure à 20% où l'on obtient une concordance (Acc) de 57% et 39% en moyenne. La seconde partition correspond à l'intervalle de perte de sang compris entre 20% et 40% où l'on obtient quasiment toujours une mauvaise classification. Enfin, la troisième partition correspond à une perte de sang supérieure à 40%, on l'on constate une concordance de 2%, 13% et une très bonne concordance pour des pertes de sang supérieures à 65%. Même si notre spoliation a été réalisée de façon continue, les résultats obtenus montrent que l'intérêt d'une classification par ce type de modèle est douteuse.

On peut se demander alors ce que produit une classification en prenant en compte les observations réalisées sur la matrice de confusion précédente, c.-à-d., en prenant 3 classes : $\omega_1 : PerteSang < 15\%$, $\omega_2 : 15\% \leq PerteSang < 37.5\%$, $\omega_3 : 37.5\% \leq PerteSang$.

On obtient alors la matrice de confusion représentée Tab. 12.6 pour un classifieur quadratique et les caractéristiques portables retenues. L'erreur obtenue est alors de 46.3% et l'on constate une concordance intéressante de 77.1% pour la classe ω_3 .

12.3 Conclusion du chapitre

L'intérêt d'une classification par portion de volume sanguin prélevé, ne peut être mise en évidence en utilisant une technique de validation animal par animal sur nos données (obtenues en effectuant une perte de sang à débit contrôlé), car elle est mise en défaut lors d'une évaluation plus pessimiste de type leave one (animal) out.

TAB. 12.5 – Exemple de table de confusion obtenue pour un modèle quadratique pour une classification selon huit classes correspondant à des plages de volume sanguin prélevé. La classification a été obtenue sur des données non invasives. L'évaluation a été réalisée selon la procédure de leave one (animal) out.

<i>PerteSang</i>	Machine								Acc (%)
	0%	10%	20%	30%	35%	40%	45%	50%	
0%	13 ± 10	8 ± 9	0 ± 0	1 ± 1	0 ± 0	1 ± 4	0 ± 1	0 ± 0	57.1%
10%	6 ± 5	5 ± 6	0 ± 0	0 ± 1	0 ± 0	0 ± 1	0 ± 1	1 ± 2	38.6%
20%	6 ± 6	4 ± 5	0 ± 0	1 ± 1	0 ± 1	0 ± 1	1 ± 4	1 ± 3	0.4%
30%	4 ± 4	2 ± 4	0 ± 0	0 ± 0	0 ± 1	0 ± 0	2 ± 3	1 ± 3	0.4%
35%	1 ± 2	2 ± 3	0 ± 0	0 ± 1	0 ± 0	0 ± 0	1 ± 2	2 ± 3	0.0%
40%	1 ± 1	2 ± 2	0 ± 1	0 ± 0	0 ± 0	0 ± 1	1 ± 1	2 ± 3	2.34%
45%	0 ± 1	1 ± 3	0 ± 1	0 ± 1	0 ± 0	0 ± 1	1 ± 1	3 ± 3	12.9%
50%	0 ± 0	2 ± 4	0 ± 1	1 ± 1	0 ± 1	1 ± 3	5 ± 5	16 ± 11	64.9%

TAB. 12.6 – Exemple de table de confusion obtenue pour un modèle quadratique pour une classification selon 3 classes correspondant à des intervalles de volume sanguin prélevé. La classification a été obtenue sur des données portables. L'évaluation a été réalisé selon la procédure de leave one (animal) out.

<i>PerteSang</i>	Machine			Acc (%)
	ω_1	ω_2	ω_3	
ω_1	19.4 ± 9.9	12.0 ± 7.7	3.9 ± 8.2	55.0%
ω_2	11.1 ± 8.5	6.2 ± 5.2	10.9 ± 12.8	22.0%
ω_3	1.8 ± 3.6	6.6 ± 7.2	28.2 ± 9.6	77.1%

Nous pouvons de cette façon porter un doute sur les résultats proposés par Glass et coll. Cependant il semble qu'une classification du choc hémorragique par perte de sang pourrait être possible comme l'indique une analyse du problème à trois classes ($\omega_1 : PerteSang < 15\%$, $\omega_2 : 15\% \leq PerteSang < 37.5\%$, $\omega_3 : 37.5\% \leq PerteSang$) qui montre une concordance non négligeable pour une perte de sang importante ($\omega_3 : 37.5\% \leq PerteSang$). Une recherche des valeurs seuils pourrait permettre d'affiner les résultats. On peut encore se demander si une régression par morceaux ne serait pas plus appropriée.

Chapitre 13

Discussion

En essayant les méthodes décrites par l'équipe de Glass, sur la création et la validation des classifieurs, qui consiste en un apprentissage et une validation animal par animal, on obtient des résultats qui paraissent bons et semblent peu entachés d'erreur. Nous avons vu qu'en utilisant cette méthode, notre système était très compétitif vis à vis du leur, mais au prix d'une grande erreur d'évaluation du risque empirique associé au modèle. Aussi, la validation apportée par Glass et coll. sur leur système peut être mise en défaut. Néanmoins l'article reste vague sur la démarche adoptée et nous ne saurions conclure sans doutes, sur la validité de leurs résultats en terme de généralisation.

L'utilisation d'une classification par perte de sang peut paraître intéressante puisqu'elle provient d'une donnée mesurable et non entachée d'un raisonnement humain telle que la classification en classe de choc hémorragique par un expert. Néanmoins les problèmes de classification dans les classes voisines permettent de préciser qu'un modèle de régression est plus justifié. Cependant, il semble que le comportement des données soit différent suivant différents niveaux de choc, ce qui implique plusieurs modèles de régression. La figure Fig. 13.1 montre que suivant la quantité de sang prélevé, les variables ont des comportements linéaires puis, à partir de 40% il commence à y avoir une courbure. Les conclusions de Glass et coll. indiquent que ces courbures sont aussi observées sur leurs données et que leurs prochaines études porteront sur l'étude de ces courbures et des trajectoires prises par les variables au cours du choc hémorragique.

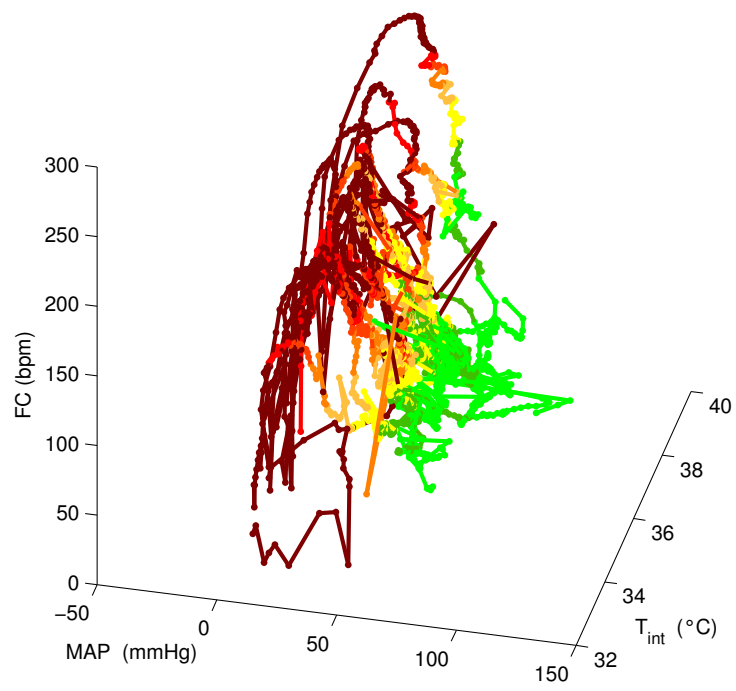


FIG. 13.1 – Classification selon perte de sang

Cinquième partie

Conclusion

Chapitre 14

Conclusion

L'objectif de ce travail de thèse était de mettre en évidence si un système portable pouvait être envisagé pour diagnostiquer l'état de santé d'un blessé polytraumatisé.

A cette fin, nous avons effectuée une étude bibliographique très vaste de façon à explorer les différents domaines mis en cause dans la conception et la réalisation d'une telle étude. Nous avons vu que lors des situations de crise, l'utilisation d'un tel appareil pourrait permettre d'aider les médecins débordés par l'afflux des blessés dans la phase critique de triage. La connaissance médicale agrégée des différents travaux antérieurs nous a montré que l'état du blessé polytraumatisé dépendait de la perte de sang subie (en excluant les traumatismes crâniens et les lésions cérébrales) et que cette perte engendrait un choc hémorragique qui pouvait être décomposée selon trois classes (choc compensé, choc décompensé, choc irréversible). Nous avons vu que les scores proposés par la recherche médicale étaient difficilement applicables pour notre système puisqu'ils nécessitaient la présence d'une personne capable de réaliser un diagnostic (parfois anatomique ou neuropsychologique).

Nous avons donc cherché à comprendre la démarche diagnostique utilisée par les médecins pour essayer d'implanter celle-ci dans notre système. Nous avons confronté celle-ci à la démarche utilisée en automatique, traitement du signal et informatique, pour la surveillance et le diagnostic des systèmes. Nous avons finalement conclu que ces deux approches présentaient des similitudes dont l'étude faisait l'objet des sciences cognitives. Nous avons alors choisi de réaliser un système qui associe les deux démarches en retenant une approche par modèle externe créé par apprentissage supervisé d'un score réalisé par un expert.

Nous nous sommes donc portés sur l'expérimentation animale pour pouvoir recueillir des données nécessaires à l'évaluation de ce type de modèle. Pour cela, nous avons utilisé un modèle animal : le modèle du jeune porc subissant une perte de sang à débit contrôlé sous anesthésie la plus faible possible. Nous avons alors mis au point l'expérimentation associée à ce modèle pour nous permettre l'acquisition des mesures nécessaires à la réalisation du scoring par l'expert et à la réalisation du système portable. L'utilisation en routine du protocole expérimental résultant de cette mise au point, nous a permis d'effectuer une collecte de données sur 19 animaux.

Nous avons alors utilisé les enregistrements pour réaliser une extraction de caractéristiques, pour la plupart couramment utilisées en anesthésie réanimation, en proposant des algorithmes simples mais robustes aux modifications engendrées sur des signaux évoluant au cours de l'expérimentation. La visualisation de ces caractéristiques a permis à un expert

de classer ces observations suivant quatre classes correspondant à une classe d'état normal et trois classes de choc hémorragique définies précédemment.

Nous avons alors effectué une normalisation des données par zscore (données centrées réduites) et nous avons utilisé différentes transformations pour essayer de se rapprocher le plus d'une loi de type Laplace Gauss centrée en 0 et de variance 1.

Un apprentissage artificiel de ces données a alors été amorcé et nous a conduit à définir les règles, principes et techniques de validation associées. Nous avons précisé notre démarche en nous focalisant sur l'apprentissage supervisé, la reconnaissance de formes et la description des classifieurs utilisés dans notre étude.

Nous avons alors présenté les résultats obtenus, la visualisation des régions de décisions, l'influence de la prise en compte de la granularité temporel pour l'extraction des caractéristiques et l'influence de l'introduction de caractéristiques dynamiques.

Enfin nous avons comparé nos résultats à ceux obtenus par l'équipe de Glass et coll. car il nous semblait important de préciser notre démarche, par rapport à une étude comportant de grandes similarités expérimentales et techniques.

Nous pouvons conclure de cette étude qu'une chaîne de traitement de l'information d'un système capable d'indiquer l'état de santé d'un blessé polytraumatisé en temps réel, chaque minute, semble réalisable. La mise en place et l'évaluation d'un tel système sur le modèle animal du porc subissant une lourde perte de sang continue à débit contrôlé indique une performance de 80% de concordance (20% d'erreur) avec un scoring expert lorsque la pression artérielle est disponible. La performance chute à 60% de concordance (40% d'erreur) si l'on utilise des capteurs non invasifs.

L'utilisation d'une méthode d'estimation de cette erreur par une technique non dépendante des individus constituant la population servant à créer le système (leave one (animal) out) a été retenue pour obtenir une estimation la plus réaliste possible. Des estimations plus optimistes permettent de prédire une erreur de l'ordre de 30% entre l'homme et la machine lorsque des capteurs non invasifs sont utilisés.

Le changement de granularité temporelle pour la détermination des caractéristiques du modèle indique qu'une fenêtre temporelle de 60 s est un bon compromis entre précision et vitesse, bien que des durées plus courtes soient compétitives. L'apport d'une information dynamique ne semble pas améliorer la performance des classifieurs utilisés, et ceci quel que soit le mode de calcul de cette information dynamique (introduction d'une phase, calcul de vitesse, pente de régression).

L'utilisation d'une classification par un seul expert peut prêter à confusion dans l'évaluation des performances du système. Cette contrainte forte peut mettre en échec la plausibilité du modèle. Il serait meilleur d'utiliser les classes résultantes d'une combinaison de plusieurs experts. L'analyse du rejet permet de mettre en évidence que les classes ω_2 et ω_3 sont particulièrement touchées par une ambiguïté et nous pensons que celle-ci provient en grande partie de la classification de l'expert. L'analyse du rejet des transitoires entre classes n'améliore pas les résultats, ce qui montre que l'erreur ne provient pas de ces transitoires observés sur chaque individu, mais pourrait indiquer que l'ambiguïté provient de la mise en commun des classifications réalisées sur les différents individus. Nous pouvons conclure en visualisant les classifications réalisées par la machine qui semble plus robustes que celles de l'expert.

L'utilisation et la comparaison de plusieurs classifieurs de natures différentes aboutit à des performances comparables. Cette comparaison a été réalisée après transformation

des données dans un espace de représentation approprié de façon à réaliser une analyse multidimensionnelle sur des caractéristiques normées. Cette procédure permet de préciser que la nature du problème dépend bien des données et non de la méthode utilisée. Elle met aussi en évidence l'influence des caractéristiques les plus performantes telles que la pression artérielle, la fréquence cardiaque, la température cutanée, le pouls périphérique et la saturation pulsée, le rythme respiratoire.

La vision de l'évolution des variables du modèles et des régions au cours du choc hémorragique permet d'apprécier l'apport de l'apprentissage artificiel effectué par la machine et guidé par l'expert dans la prise en compte du phénomène. La mise en place d'un système de visualisation de ce type lors de la prise en charge du blessé ou au chevet de celui ci nous semble une voie intéressante pour le monitoring de l'état de santé. Elle permet en outre d'apprécier l'aspect multidimensionnel de cet état de santé et met en évidence la possibilité de déterminer des régions de l'espace indépendantes des individus mais dépendantes uniquement des valeurs prises par les observations réalisées. Si la représentation dans un espace de mesures non invasives est moins discriminante que dans un espace comprenant des mesures invasives, le comportement global de la population étudiée indique qu'il existe des régions de l'espace correspondant à différents comportements typiques pouvant correspondre aux classes de choc hémorragique.

La comparaison du système avec les résultats obtenus par l'équipe de Glass montre que notre méthode pour calculer l'erreur est plus réaliste que leur méthode pour créer et évaluer le système. En utilisant leur méthode, on obtient des résultats qui favorisent un apprentissage par cœur sur chaque individu, qui sont très optimistes pour le système et ne permettent pas de discriminer les caractéristiques importantes pour la création du modèle le plus adéquat.

L'ensemble de ces résultats permet de confirmer l'hypothèse selon laquelle l'état de santé d'un blessé polytraumatisé peut être représenté dans un espace de paramètres portables dans lequel les classes de choc hémorragique sont déterminées par différentes régions.

Il est utile de noter que la restitution d'une valeur telle que la classe de choc hémorragique, si elle synthétise l'information, peut être mal acceptée par les spécialistes. On peut se demander quel crédit sera porté par les spécialistes sur cette mesure à partir du moment où elle n'aura pas fait ses preuves en pratique, clinique ou sur le terrain. Il est donc important de fournir, dans un premier temps pour les personnes compétentes, un ensemble non réduit à une seule information mais un ensemble contenant plusieurs valeurs. L'information sur la classe de choc sera alors une aide à la décision pour le médecin et lui permettra d'ajouter un critère pour sa prise de décision. C'est dans ce sens que les visualisations obtenues permettent de prendre en considération plus facilement, l'état de l'individu et son évolution (amélioration, l'espace devient plus « vert », ou aggravation, l'espace devient plus « rouge foncé »).

Une indication sur la nature des données et en particulier sur leur fiabilité, devrait être envisagée pour permettre de différencier une erreur de mesure, d'une valeur réelle. Ceci constitue une recherche à part entière sur les capteurs à utiliser et permettrait de cerner quelles sont les mesures portables robustes, utilisables sur le terrain et celles que l'on ne peut envisager en conditions extrêmes. Par exemple, le capteur de SpO_2 permet cette différenciation assez facilement : la différence entre un *débranchement capteur* et une *perte de pouls périphérique* peut être obtenue par simple détection de la puissance circulant entre l'émission et la réception du signal lumineux, le premier événement correspondant

à une absorption nulle et une transmission intégrale du signal, le second correspondant à une absorption importante. Dans ce même ordre d'idée, il serait important de faire une recherche spécifique sur les rapport qualité prix des matériels possibles, ainsi que leur dérive dans le temps et leur consommabilité. Ces études pourraient encore préciser les capteurs cibles potentiels.

Bien sûr, on pourrait augmenter le nombre d'animaux pour l'étude pour obtenir une statistique plus précise. Néanmoins, cette solution ne paraît pas éthique au sens où la recherche réalisée ne fait que partie d'une étude amont et que les résultats générés par celle-ci ne sont pas là pour créer un modèle précis mais fournir des pistes pour décider ou créer des voies à aborder. Ainsi pour cette étude, le nombre de cochons peut paraître suffisant car il met en évidence les limites de notre démarche. On peut bien entendu extrapoler les modèles à l'homme mais il faut se méfier de ce genre de démarche, l'animal utilisé ayant ses propres particularités physiologiques et anatomiques (par exemple, réflexe splanchnique important dans la phase compensée du choc hémorragique pour le cochon).

De nombreux progrès pourraient être réalisés et ceci à tous les étages de la chaîne complète de traitement de l'information. En particulier, l'utilisation de fenêtres temporelles adaptatives permettrait d'affiner le calcul des caractéristiques utilisées en entrée de la classification.

Les principales améliorations du système pourraient venir de la détermination de nouvelles caractéristiques non évaluées dans notre système : temps de transit du pouls, variabilité cardiaque en utilisant des fenêtres de 5 min, etc. L'utilisation d'autres capteurs pourrait améliorer les performances pour le scoring réalisé par l'expert : prise en compte de l'acido-basicité du sang (acidémie du système central) (cf. Cannon) ainsi que d'autres paramètres, corrélés à un indice de la dette en oxygène (angl. *oxygen debt*), d'un excès de base, pression partielle de dioxyde de carbone ou encore un indice permettant d'évaluer la fonction neurophysiologique. Un article récent sur l'évaluation de nouveaux paramètres indique que l'utilisation d'autres mesures, comme les valeurs de précharge (angl. *preload values*) et de résistance vasculaire systémique, peut se révéler plus pertinente pour le pronostic de survie au choc hémorragique [127]. En ce qui concerne la mesure de pression artérielle non invasive par photopléthysmographie (de type finapres) on peut noter l'article de James Davis [102] qui indique que la mesure oscillométrique de la pression artérielle n'est pas adaptée (elle donne des valeurs supérieures à une mesure manuelle) pour les patients traumatisés. Ce qui implique que le système n'est pas applicable avec ce type de mesures de pression artérielle à l'heure actuelle.

Au terme de cette étude, nous pouvons lister un ensemble des principales erreurs réalisées, faire ressortir quelques points critiques et pouvons fournir quelques notes pour l'avenir...

D'un point de vue expérimental, nous devons affiner la maîtrise de l'anesthésie et minimiser ses effets délétères sur la réaction induite par le choc hémorragique. Le travail sur l'animal devant être réalisé en accord avec une certaine éthique, l'anesthésie ne doit cependant pas nuire à l'observation des phénomènes étudiés et venir biaiser les résultats. Il serait de même préférable d'aborder par voies inférieures (introduction des cathéters dans l'artère ou la veine fémorale) pour éviter de toucher aux barorécepteurs carotidiens et ainsi préserver un des mécanismes afférents de la réponse sympathoexcitatrice.

Du point de vue du traitement des données et de la mise en place des valeurs de référence, on constate que l'expert et le travail visuel et manuel sont nécessaire. Avant de passer au tout automatique, une phase de scoring manuel est indispensable et sert en

général de référence. De façon à bien évaluer la détermination automatique des caractéristiques, un marquage et une détermination manuelle de celles-ci semble inévitable. Aussi, on peut remarquer qu'il n'y a pas besoin de scorer les données sur une fenêtre de 7 min en faisant glisser celle-ci au cours de l'expérimentation comme l'a fait l'expert, puisque nous recherchons un modèle correspondant à la réalité et non un mimétisme de la pratique experte. Aussi, il serait intéressant de travailler avec une classification réalisée par l'expert en visualisant la globalité des données.

On pourrait aussi rechercher une classification selon 2 phases, ce qui amènerait le système à 3 classes : ω_1 - RAS, ω_2 - Phase sympathoexcitatrice, ω_3 - Phase sympathoinhibitrice, comme il semble être précisé dans la description de Schadt. Les travaux de Glass et coll. orientent la recherche sur différentes classes correspondant aux paliers de choc hémorragique. Nous avons vu que la classification était alors difficile sur nos données puisque celles-ci avaient été obtenues suivant une perte de sang continue à débit contrôlé. Néanmoins, nous avons noté l'influence de certaines valeurs sur la diminution du nombre d'erreurs (en particulier pour une perte de sang supérieure à 40% du volume sanguin total). Ceci semble en faveur de l'utilisation d'un modèle de type régression par morceaux, selon les régions précédemment déterminées par classification et qui ressortent de l'étude. L'étude du comportement d'un point de vue dynamique fait lui aussi ressortir cette notion de linéarité par morceaux et de comportement selon des régions de l'espace. Il semble donc intéressant de poursuivre ces recherches dans cette direction.

On pourrait encore utiliser une classification non supervisée mais on remarque en particulier sur les figures (p. ex. Fig. 8.1 ou 13.1) que la courbure des points implique la définition d'une distance particulière. Pour cela on peut se reporter à Nakache [128] et Celeux [109] ou d'autres ouvrages de classification non supervisée pour le choix d'une ultramétrie adéquate [106, 107]. L'utilisation d'une méthode d'analyse en composante curviligne peut être intéressante pour ce type de données afin de déplier celles-ci et permettre une classification plus facile dans un espace mieux adapté aux outils mathématiques utilisés. Là encore, une régression pourrait être réalisée.

Ces constatations font penser aux modèles de régression utilisés pour créer les scores présentés en 3.3 et particulièrement aux travaux de Champion et coll. en triage [129, 130, 5]. Trouver des relations analytiques entre les variables et la perte de sang, par une régression par morceaux pourrait permettre d'obtenir un modèle plus fidèle de la réalité et permettrait d'améliorer la prédictibilité du système. Le système permettrait de prédire plus précisément le temps de survie du patient et ainsi permettrait une aide plus importante dans le choix du malade à traiter en priorité. Pour la création du modèle, il faudrait tenir compte, comme pour l'établissement des scores d'évaluation, de la contrainte de survie, mais celle-ci reste très discutable, puisqu'à partir du moment où on interagit sur le modèle pour essayer de récupérer l'organisme, l'aide extérieure est difficile à modéliser. Les études qui étudient ces phénomènes de récupération par injection de différents fluides sont importantes et mettent en évidence la difficulté de tels modèles et en particulier des réactions individuelles non expliquées par les variables de leurs modèles.

Nous avons utilisé dans cette étude un mélange de signaux provenant de différents individus observés à des instants différents de façon à réaliser un modèle qui puisse s'appliquer à n'importe quel nouvel individu traité. Cette approche se rapproche d'un raisonnement de type physique statistique avec mélange de particules observées individuellement dont on cherche une loi générale. Aussi, comme il est fait en physique, la mise en évidence des processus dynamiques générés au cours du choc hémorragique suivant la perte de sang à

débit contrôlé pourrait être une voie de recherche intéressante. Elle pourrait permettre de conclure sur l'influence de la perturbation en expérimentant différents débits contrôlés. Nous pensons que la mise en place de ces deux dernières approches (statistique physique et dynamique), permettrait de mieux prédire l'évolution de l'état de santé au cours d'un choc hémorragique et ceci pour un ensemble de caractéristiques observables de natures différentes (invasifs ou non invasifs).

Annexe A

Article Chapitre Sommeil

A paraître dans [131]

A.1 Introduction

In biology, taxonomy has been the source of numerous studies and still remains one of the predominant fields of research (genome studies). The development of multidimensional exploratory analyses, computing power and numerical solutions can explain the growth of such studies. However, in the case of time series, one can notice that relative few works have been developed to deal with clustering or classification techniques. One interesting source of such studies is the study of sleep, where several classification techniques have been tested to determine structures on real temporal data [132, 133].

The starting point of sleep studies has been the observation of the electrical activity of the brain measured by electrodes fixed on the scalp, during all night recordings. First observations showed that several patterns were similar from one individual to another, their distributions fluctuating throughout the night. In its beginning (about 1940) [134, 135], analog signals were plotted on pages of paper. At that time, sleep recordings consisted of huge blocks of paper. With the first discoveries and the evidence of different phases of electroencephalic activity during the night [136, 137, 138, 139], a growing of several techniques for electrodes placement were applied and various practices for classifying these activities sprang up. In order to extract the different patterns of such recordings, one expert was assigned to read signals page by page, and give a score corresponding to specific processes of the sleep activity of the brain. The result of this reading has been called an hypnogram and consists in a succession of stages through the night (see Fig. A.1).

The advantages of working with hypnogram are : an extraction of information from raw data generated by polysomnographic (PSG : multi-channel sleep) recordings, an easier interpretation of the architecture of the night and a better vision of the organization of long term biological processes. It is then easier to discriminate strange charts from normal ones. Therefore, the hypnogram becomes a powerful tool for the diagnosis of sleep pathologies. Besides, the hypnogram, as a summary of the night, considerably reduces the storage of data and allows different laboratories to exchange their results and share their knowledge. For that reason, a consensus for a standardization of the rules used to score PSG recordings was held in 1968 bringing together the different leaders in electroencephalography. It lead to the creation of a reference known as the manual by Rechtschaffen and Kales (R&K)

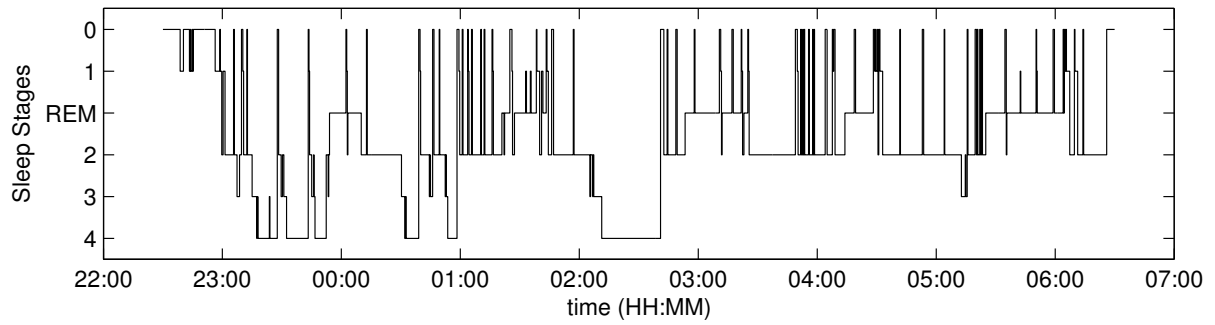


FIG. A.1 – A human hypnogram. Sleep-wake stage scoring has been realized by an expert into 6 different sleep stages from 22 h 30 min to 06 h 30 min over epochs of 20 s : 0–Wake and Movement Time, 1–stage 1 (transition from waking to sleeping), 2–stage 2, 3–stage 3, 4–stage 4 (stage 2, 3 and 4 are part of the orthodox sleep with more and more slow waves observed on the recording), REM–Rapid Eye Movement (or paradoxical) sleep (with rapid brain activity with or without rapid eye movements and muscle atonia)

[83] currently applied in the different sleep laboratories where pathologies, sleep disorders and untypical hypnograms are studied.

Since 1970, and the growth of computerized methods, interests have been initiated in order to score automatically polysomnographic recordings [140, 141, 142], avoiding the expert to spend too much time to this time-consuming work. But studies are still in progress and improvements have to be performed. For a complete review of the history of sleep, the reader is referred to [143] where the author notes, speaking about automatic sleep staging : ‘The task turned out to be much more difficult because of ambiguities, artifacts and variations in human scoring’.

This study has been lead in order to understand what could be the different difficulties encountered with real biological data, while confronting expert practices and machine learning algorithms. For that purpose, a comparative study of five classifiers for automatic analysis of human sleep recordings is presented where temporal data coming from different individuals are mixed together. The interest of transformations toward normal distribution is emphasized since they lead to homogeneous representations for the different selected features. In the first paragraph, the database, the different classifiers, the method chosen to evaluate the performances of the classifiers and the transformations toward normal distribution are presented. In the second paragraph, the results obtained are discussed.

A.2 Materials and Methods

A.2.1 Presentation of the Database

The study has been realized over $N = 11$ polysomnographic recordings available in our database (from 4 healthy subjects). Features were extracted from one EEG (electroencephalogram, differential lead C3–A2) and one bipolar EMG (electromyogram, position chin), sampled at 200 Hz. The choice of these features has been made in accord with experts in an effort to test a minimal set of electrodes considered necessary for the scoring of sleep.

TAB. A.1 – Description of the features used in the study and their statistical values for a) raw data, b) with z-score normalisation and c) with transformations toward normal distribution

		a)			b)		c)			
Feature	μ	σ	min	max	min	max	Transformation	min	max	
F_1	$\sigma(EEG)$	16.87	13.67	4.69	227.31	-0.89	15.40	$\log(1+x)$	-1.97	5.38
F_2	$P_{rel}(EEG, \delta)$	0.69	0.16	0.01	0.99	-4.27	1.88	$\arcsin(\sqrt{x})$	-4.94	2.58
F_3	$P_{rel}(EEG, \theta)$	0.14	0.07	0.00	0.68	-1.92	7.46	$\arcsin(\sqrt{x})$	-2.93	5.60
F_4	$P_{rel}(EEG, \alpha)$	0.05	0.04	0.00	0.46	-1.34	10.65	$\log(\frac{x}{1-x})$	-4.91	2.95
F_5	$P_{rel}(EEG, \sigma)$	0.05	0.04	0.00	0.50	-1.20	11.54	$\log(\frac{x}{1-x})$	-4.53	2.87
F_6	$P_{rel}(EEG, \beta)$	0.04	0.04	0.00	0.94	-0.93	23.31	$\log(\frac{x}{1-x})$	-3.62	2.95
F_7	$P_{rel}(EEG, \gamma)$	0.06	0.10	0.00	1.35	-0.59	13.44	$\log(\frac{x}{1-x})$	-2.95	2.45
F_8	$\sigma(EMG)$	21.42	39.54	0.00	394.97	-0.54	9.45	$\log(1+x)$	-2.10	3.25

Eight features thought to represent important physiological processes calculated over epochs of 20s have been considered and are reported table A.1, where σ , denotes the standard deviation and P_{rel} the relative power in a given frequency band. The different bands are : δ (0.5–4.5 Hz), θ (4.5–8.5 Hz), α (8.5–11.5 Hz), σ (11.5–15.5 Hz), β (15.5–22.0 Hz), γ (22.0–45.0 Hz) and corresponds to the ones generally employed in sleep and waking EEG spectral studies [144].

During these epochs of fixed temporal interval ($\Delta t = 20$ s), EEG can be considered approximately stationary [145]. This assumption is fundamental for estimation of the different retained features, both in time domain and in spectral domain. To each epoch, a score has been attributed by an expert. This score is assigned from a set constituted of $K = 6$ classes representing the 6 different stages encountered during human sleep defined in regards with the conventional criteria of R&K [83] : 0 - Wake and Movement Time, 1 - stage 1, 2 - stage 2, 3 - stage 3, 4 - stage 4, 5 - Rapid Eye Movement sleep (or Paradoxical sleep). The different aspects of EEG and EMG signals are represented Fig. A.2, in order to appreciate the variations of the different signals throughout human sleep.

Once all the signals have been segmented into epochs, preprocessed and their features extracted, we can represent any observation $\mathbf{x}$ by a state representation in an R^d space ($d = 8$ for our study) where $(.)^t$ denotes the transpose operator :

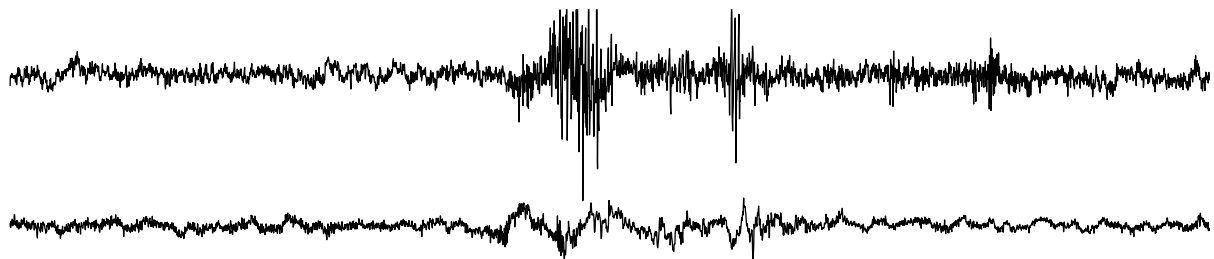
$$\mathbf{x} = (F_1, F_2, \dots, F_d)^t \quad (\text{A.1})$$

When regrouping both the temporal instants and the different individuals, we obtain an array of observations or statistical units [107] representing the database over which the classification study is done :

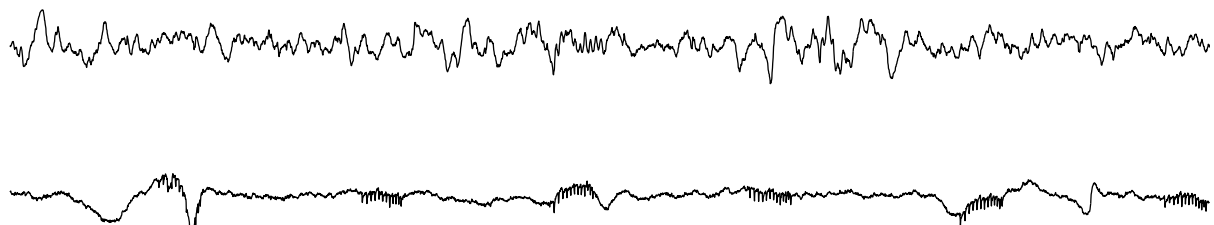
$$M = \begin{pmatrix} F_1(t_0(1)) & F_2(t_0(1)) & \dots & F_d(t_0(1)) & C(t_0(1)) \\ \dots & \dots & \dots & \dots & \dots \\ F_1(t_f(1)) & F_2(t_f(1)) & \dots & F_d(t_f(1)) & C(t_f(1)) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ F_1(t_k(i)) & F_2(t_k(i)) & \dots & F_d(t_k(i)) & C(t_k(i)) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ F_1(t_1(N)) & F_2(t_1(N)) & \dots & F_d(t_1(N)) & C(t_1(N)) \\ \dots & \dots & \dots & \dots & \dots \\ F_1(t_f(N)) & F_2(t_f(N)) & \dots & F_d(t_f(N)) & C(t_f(N)) \end{pmatrix} \quad (\text{A.2})$$

Where $t_k(i) = t_0(i) + k\Delta t$ for individual i .

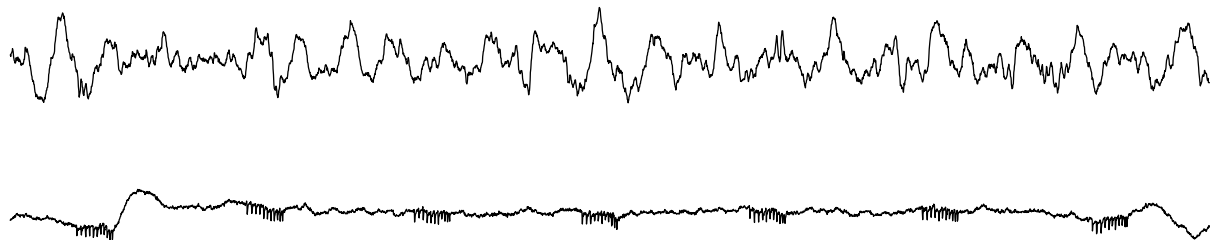
The database was then constituted of 17265 observations over 8 parameters in which we introduced into the last column the expert's classification. A visual part of such a



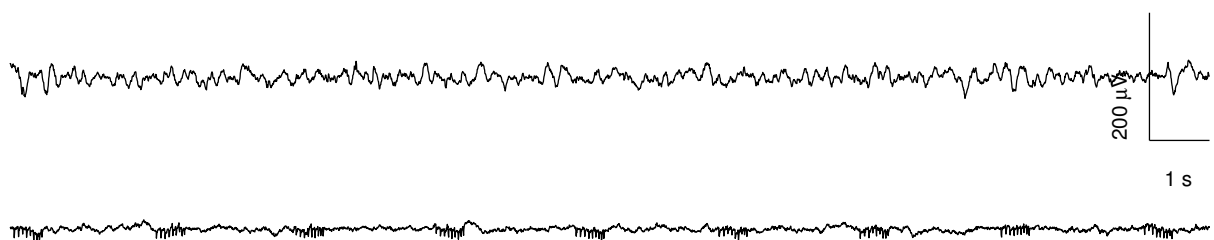
a) Stage 0 : Waking and Movement Times, noisy signals.



b) Stage 2 : Apparition of short bursting events and slow waves



c) Stage 3-4 : Large slow waves.



d) Stage 5 : Rhythmic θ activity and flat EMG.

FIG. A.2 – Electrophysiological behaviour during principal sleep stages. Each figure represents an epoch of 20 s. Same scale has been used for all figures.

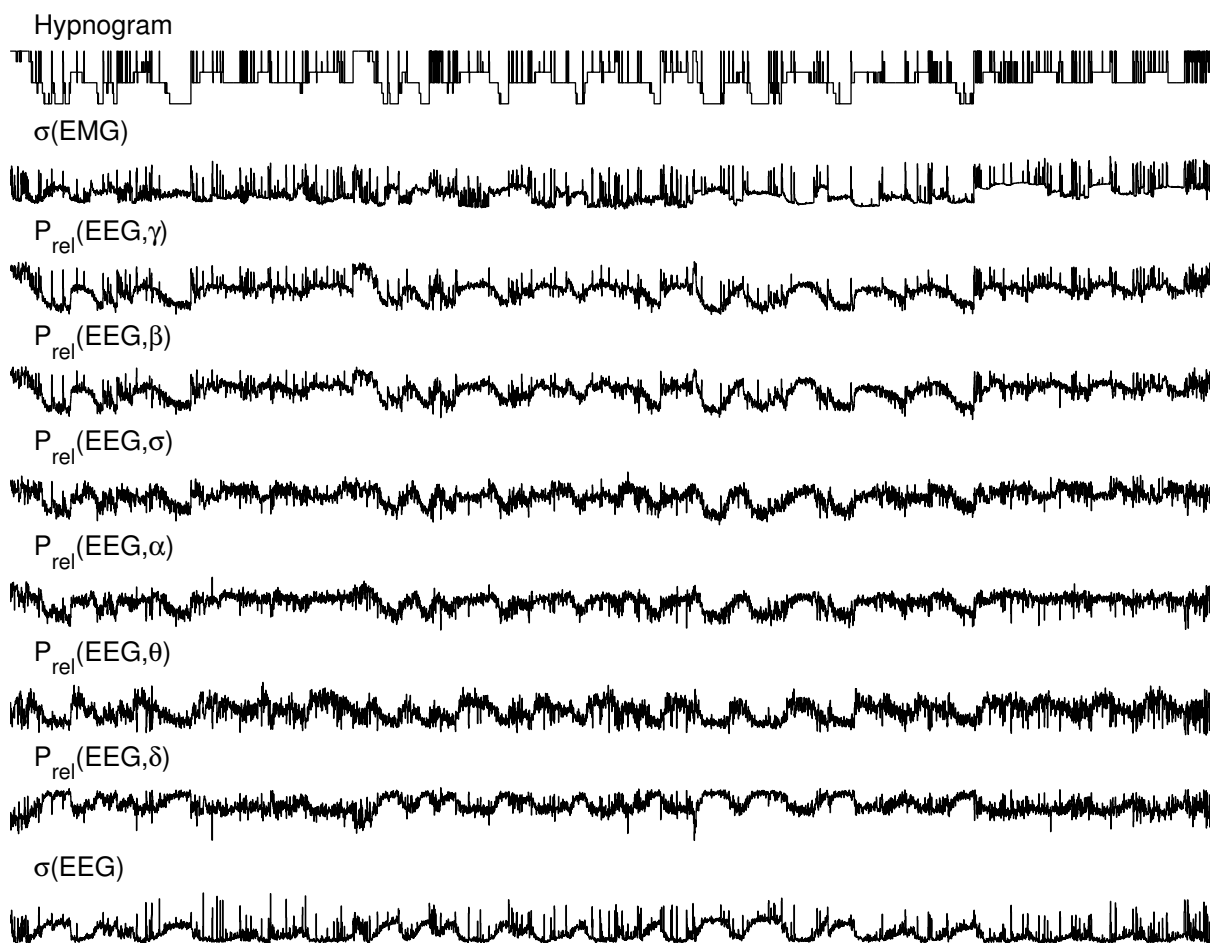


FIG. A.3 – Representation of 5000 elements of the database after transformation. Time and Individuals are grouped together. When data is homogeneous, the influence of each feature is directly observable over the classification of the expert represented at the top of the figure

database is given Fig. A.3, with features transformed into normal distribution (detailed in A.2.4) for a better homogeneity of representation.

A.2.2 Learning and Testing Sets

In machine learning, the supervised learning approach tries to learn rules, statistics, mathematical models, with a computer, from a desired result. A database containing both the different features used to solve the problem and the corresponding desired results are used. The aim is to find the model that minimizes a criteria which is a function of the difference between the results calculated by the machine and the desired results.

For this reason, it is common to separate the database into 2 sets : the first is used to induce the machine in a so called learning (or training) phase ; the second is used during a phase of validation (or test) for evaluating the performance with data that has not been used during the learning process. For a review of the different technique for evaluating and preparing the data into learning set (LSet) and testing set (TSet) the reader is referred to [146, 121, 147].

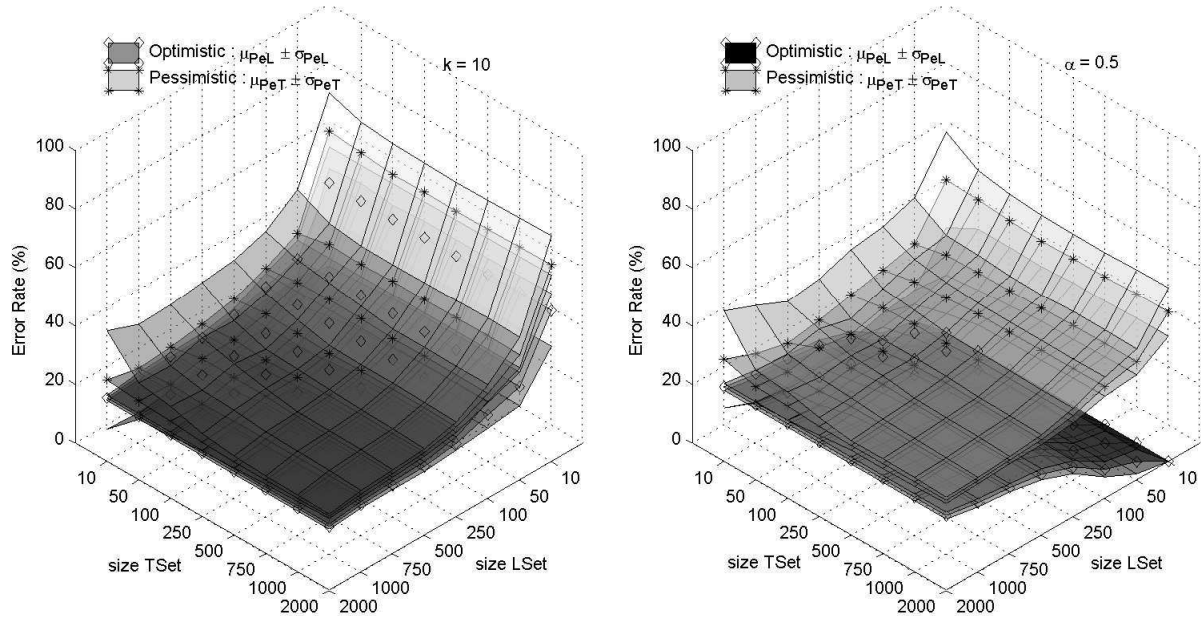


FIG. A.4 – Misclassification percentage in function of the size of the samples drawn from the database using the kNN classifier with $k = 10$ or the Parzen estimator with $\alpha = 0.5$. After 500 samples in the learning set and 500 samples in the testing set, performances are not improved.

Leave one out (N-1 vectors for learning and 1 vector for testing) or classical cross-validation (N/2 vectors for learning, and N/2 vectors for testing), can not be applied when working with large database, such as ours, without exploding calculus. We decided to randomly select a fixed number of data for the learning set and for the testing set, as it is done in Bootstrap techniques. The learning set will serve to train the classifiers, but also to calculate an optimist estimation (resubstitution techniques, empirical error) of the convergence of them. The testing set is used to obtain a pessimist estimation (cross validation techniques, real error) of them.

The difference with Bootstrap techniques is that we do not reset the drawings after each drawing. This is done in order to obtain independence between estimation from learning set and estimation from testing set.

The choice of the number of data for the learning and testing set can be obtained by looking over the stability of the performances of the classifiers. For a given size of the learning and testing sets, we trained a kNN classifier and a Parzen estimator. An estimation of the performance was realized over 30 subsets for both the optimistic and pessimistic error that are represented given with their standard deviation on Fig. A.4. Classification errors reaches $\approx 30\%$ and do not improve when the number of data in the sets increase over 500 samples when using both the kNN classifier or the Parzen estimator.

A.2.3 The Different classifiers

Five common classifiers have been evaluated that can be regrouped into two distinct categories : the first category corresponds to the set of classifiers using probabilistic com-

putations based upon the Bayes' rule to assign a class to a feature vector. The second category corresponds to classifiers delimiting regions into the representation space by direct computation of frontiers. Explanation of the behavior of the different classifiers and the learning hypothesis can be found in [148, 26, 37]. Here we provide a short description of them.

Our study learning problem is to induce a classifier able to assign to a vector $\mathbf{x}$ of the representation space, a class $C \in \{\omega_i\}_{i=1}^K$ with respect to the knowledge database constituted of the data present in the learning set. We use the following notation : P for the probability, p the probability density, $E[.]$ the expectation operator, $|\cdot|$ the determinant.

Bayes Rule-based classifiers

The attribution of a vector $\mathbf{x}$ to a class is made using the Bayes' rule (A.3). The posterior conditional probability $P(\omega_i|\mathbf{x})$ is calculated for each of the K classes and the vector is given the class ω_i for which $P(\omega_i|\mathbf{x})$ is maximal (maximum a posteriori).

$$P(\omega_i|\mathbf{x}) = \frac{p(\mathbf{x}|\omega_i)P(\omega_i)}{p(\mathbf{x})} \quad (\text{A.3})$$

$$p(\mathbf{x}) = \sum_{i=1}^K p(\mathbf{x}|\omega_i)P(\omega_i) \quad (\text{A.4})$$

The learning problem consists in estimating the conditional density function $p(\mathbf{x}|\omega_i)$ from the different samples of the learning set. The different classifiers depend on the hypotheses made on this density function.

Parametric models The probability density function is assumed to be a multidimensional Gaussian model.

$$p(\mathbf{x}|\omega_i) = \frac{1}{(\sqrt{2\pi})^d \sqrt{|\Sigma|}} \exp\left(-\frac{1}{2}(\mathbf{x} - \mu)^t \Sigma^{-1}(\mathbf{x} - \mu)\right) \quad (\text{A.5})$$

Its parameters (mean μ and covariance matrices Σ) are estimated with samples drawn from the learning set :

$$\mu = E[\mathbf{x}] = (\mu_1, \mu_2, \dots, \mu_d)^t \quad (\text{A.6})$$

$$\Sigma = E[(\mathbf{x} - \mu)(\mathbf{x} - \mu)^t] \quad (\text{A.7})$$

Linear classifier : the covariance matrix $\Sigma = \Sigma_i$ is assumed to be the same for all classes.

The resulting boundaries delimiting the classes are linear functions.

Quadratic classifier : the covariance matrix Σ_i is assumed to be different for each class and is estimated with representatives of each class in the learning set. The resulting boundaries delimiting the classes are quadratic functions.

Non-Parametric models : The density function is described with

$$p(\mathbf{x}|\omega_i) = \frac{k_{n_i}}{n_i V_{n_i}} \quad (\text{A.8})$$

With n_i number of representatives of class ω_i in the volume V_{n_i} .

k Nearest Neighbor (kNN) classifier : the probability density function is estimated by the volume occupied by a fixed number of neighbors (search of V_n with fixed k_n). It is simple to show that the decision obtained with the Bayes' rule maximization is equivalent to a voting kNN procedure. This procedure is a majority vote over the classes of the k nearest neighbors (present in the learning set) of the feature vector to classify.

Parzen estimator with Gaussian kernels : The probability density function is estimated by the sum of density kernels given a fixed volume V_n . To each sample $\mathbf{x}_{i,j}$ representative of class ω_i in the learning set, a density kernel $K(\cdot)$ is associated. The sum over j of these n_i kernels gives the density of that class in that region and the probability density function is then

$$p(\mathbf{x}|\omega_i) = \frac{1}{n_i V_{n_i}} \sum_{j=1}^{n_i} K\left(\frac{\mathbf{x} - \mathbf{x}_{i,j}}{h_{n_i}}\right) \quad (\text{A.9})$$

With

$$V_{n_i} = h_{n_i}^d = n_i^{-\alpha} \quad (\text{A.10})$$

and the Gaussian kernel

$$K(u) = \frac{1}{(\sqrt{2\pi})^d} \exp\left(-\frac{1}{2}(u^2)\right) \quad (\text{A.11})$$

This two methods require the tuning of a parameter : k , the number of neighbors for the k nearest neighbor estimator and α , for the Parzen estimator. The number of neighbors has been chosen equal to 10 and the parameter α has been set to 0.5, after evaluating the performance of the classifiers when incrementing the values of these parameters as it is shown Fig. A.5.

Frontiers based classifiers

The frontiers of the classes in the multidimensional space are directly calculated from the data present in the learning set :

A multi layer perceptron (MLP) : with 3 layers fully connected composed with 8 neurons in the input layer (hyperbolic tangent transfer function $y = 2/(1 + \exp(-2x)) - 1$), 6 neurons in the hidden layer (linear transfer function $y = x$) and 6 in the output layer (logarithmic sigmoid transfer function $y = 1/(1 + \exp(-x))$), trained by the feedforward backpropagation gradient algorithm. Weights were initiated randomly at the beginning of the learning phase. This structure is often used in discrimination [114] with an input layer connected to the representation space of $\mathbf{x}$ with $d = 8$ and the output layer connected to the desired class with $K = 6$. The choice of the number of neurons in the hidden layer has not been optimized in this study.

A.2.4 Transformations Towards Normal Distribution

Means, standard deviations, maximal and minimal values for the different retained features are given in Table A.1. Inhomogeneity in raw data can be observed, as well as

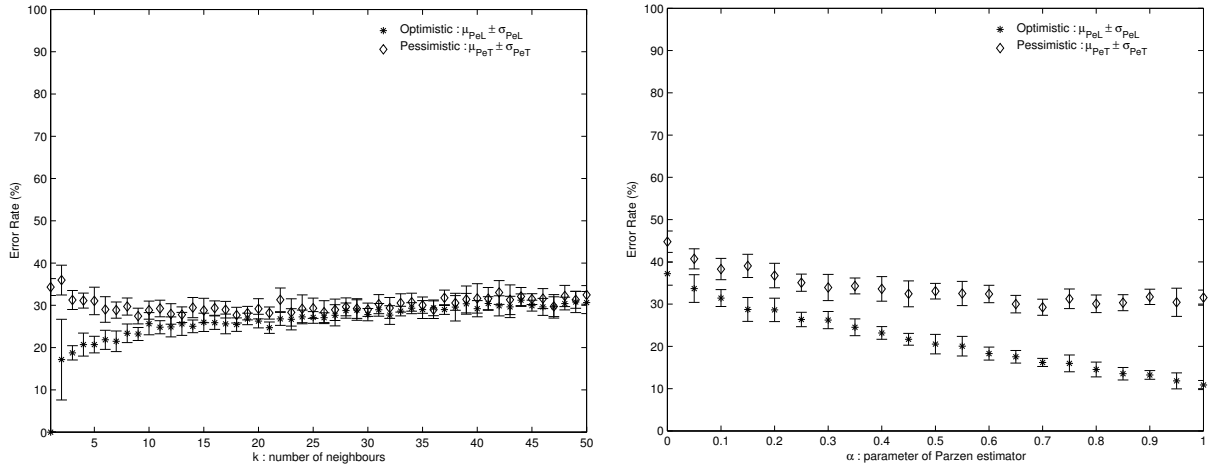


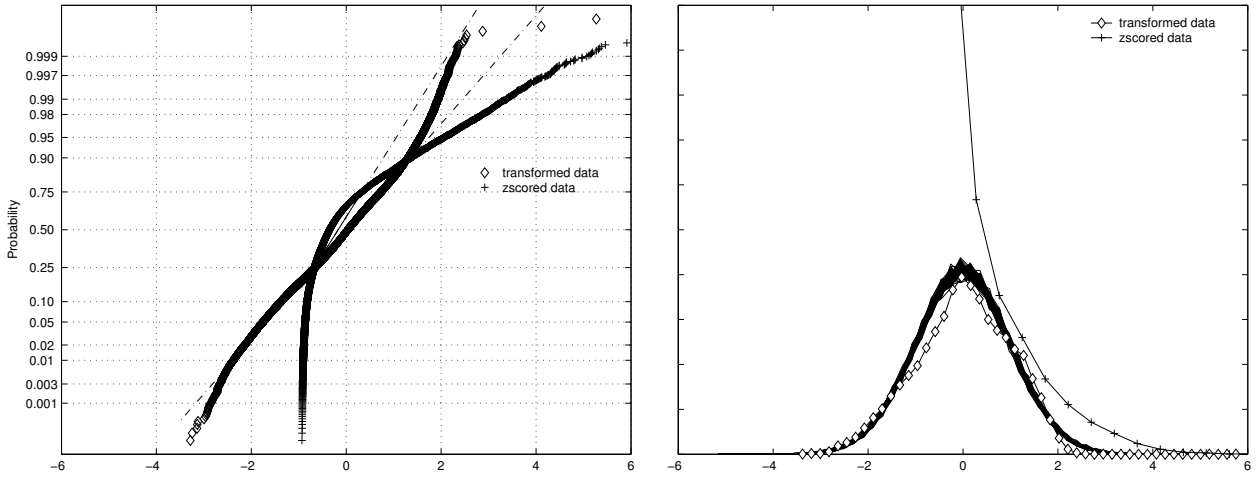
FIG. A.5 – Selection of the parameters for the k Nearest Neighbor classifier and for the Parzen estimator with Gaussian kernels. Size of the learning set and size of the test set have been set to $n_{LSet} = 500$ and $n_{TSet} = 500$. We retained $k = 10$ and $\alpha = 0.5$ for the comparison of the classifier because there is no improvement of the pessimistic error with greater values.

a wide spread of the data, which is typical with biological data. For example, even after doing a z-score defined by the transformation $\mathbf{z} = (\mathbf{x} - \mu)/(\sigma)$ (where μ is the mean of $\mathbf{x}$ and σ is its standard deviation), the maximal value of the sixth parameter is twenty three times the standard deviation.

In order to reduce the influence of extreme values, we applied transformations towards normal distribution on the whole set of the data, and for each parameter. These transformations are either $\log(x)$, $\log(1+x)$, $\sqrt{x}$, $\sqrt[3]{x}$, $\log((x)/(1-x))$, $1/(\sqrt{x})$, $\arcsin(\sqrt{x})$ depending on their effect over the different features. This was induced by Theo Gasser [111] for normalization of EEG spectral parameters. These transformations perform a better effect over the data, better than doing the simple z-score : the inter-individual variability is reduced with the advantage to reduce tails in distributions. The effects of these transformations can be seen in Table A.1. The maximal value of the eight parameters after transformation is no more than 6 times the standard deviation. An example of such transformations over one feature can be seen in the Henry plot shown in Fig. A.6(a) or from the density plot Fig. A.6(b).

A.3 Results

A training set and a validation set, each made up of 500 vectors randomly chosen, was built. Each classifier was trained on the first set and applied on the validation set. The performance of the classifier is given by the classification error expressed in percentage on the training set (which is optimistic) and on the validation set (which is pessimistic). This procedure was achieved ten times, which provides two times ten values for each classifier and enables the estimation of mean and variance. The results from one classifier to the other is said to be different if means are statistically different.



(a) Effect of the transformations for $P_{rel}(EEG, \beta)$ plotted in a lognormal axis system, known as Henry plot (normal probability plot). The $\log(\frac{x}{1-x})$ transformation gives a better approximation to a normal distribution represented by a line than a simple z-score.

(b) Effect of the transformation for $P_{rel}(EEG, \beta)$ over the distribution of the data as compared to zscored data and data obtained from Gaussian distributions simulated with the same number of realisations.

FIG. A.6 –

A.3.1 Results with Raw Data

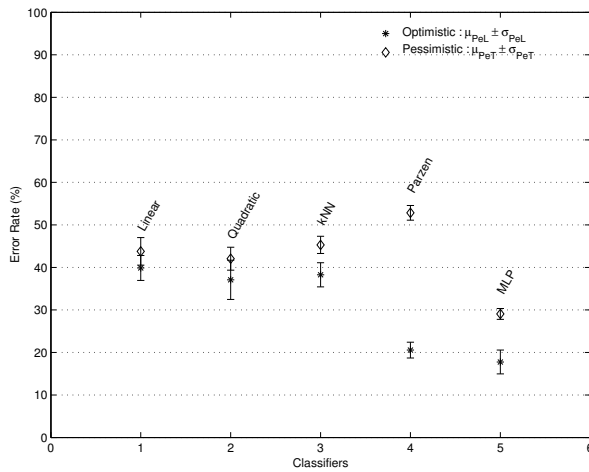
Results from raw data are presented in Fig. A.7(a). It corresponds to the mean value of the classification error on the training set and on the validation set obtained over 10 trials. For MLP, at each trial, 10 classifiers were trained with a different initialization of the weights, and the network with the minimal classification error was selected, in order to insure the convergence of the network.

In Fig. A.7(a), the small standard deviation of the results tends to prove that the technique we use for the evaluation is appropriate. The best result is obtained with the neural network with $29 \pm 1\%$ of misclassification error on the validation set. It is significantly different from others ($p < 0.01$, using a Wilcoxon sign rank test for paired samples). Their results vary from $53 \pm 2\%$ for the Parzen estimator to $42 \pm 3\%$ for the quadratic classifier.

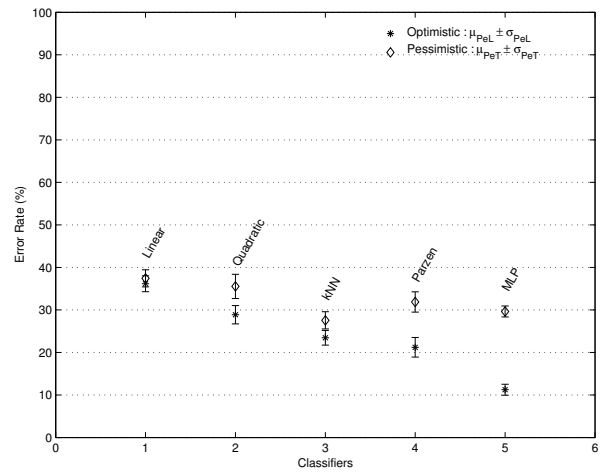
The large difference between optimistic and pessimistic estimation of the percentage for Parzen estimator (classifier d) ($p < 0.01$ Wilcoxon sign rank test for paired samples) shows that the error on the training set can definitely not be used to evaluate the performance of a classifier. Indeed, using this classifier, the vector from the training set participates too much in the decision for its classification. The high percentages obtained for optimistic estimation for the classifiers a) b) and c) shows that those classifiers do not perform well on the data. This can be explained by the large tails in the distribution and by the fact that the density probability functions of the classes cannot be fitted correctly by a multidimensional Gaussian model.

A.3.2 Results with Transformed Data

Results from data with transformations are presented in Fig. A.7(b). The transformations applied to the variables are given in Table A.1. classifiers were trained with new coordinates obtained after these transformations.



(a) Means and standard deviations for the misclassification percentage obtained by 10 classifiers of each type on raw data : a) linear discrimination, b) quadratic discrimination, c) kNN with $k=10$, d) Parzen estimator, with Gaussian kernel and e) Best MLP retained on each trial.



(b) same as Fig. A.7(a) with Transformed Data. All classifiers have improved their results (in terms of pessimistic error), except the neural network which obtains the same results.

All classifiers have increased their performances ($p < 0.01$, Wilcoxon rank sum test for independent samples) except the neural network which obtains the same results. The performance of the k nearest neighbor classifier and the Parzen classifier have been significantly improved. The pessimistic error decreases from $45 \pm 2\%$ to $28 \pm 2\%$ for the k nearest neighbor classifier and from $53 \pm 2\%$ to $32 \pm 2\%$ with the Parzen estimator. The results obtained by the k nearest neighbor classifier are then equivalent to the results obtained with the neural network. The results of the linear and quadratic classifiers are slightly better with the transformed data. The misclassification error decreased from $44 \pm 3\%$ to $37 \pm 2\%$ with the linear classifier and from $42 \pm 3\%$ to $36 \pm 3\%$ with the quadratic classifier.

These results can be explained by the fact that the neural network is not sensitive to the distributions of the data, transformations have no effects on its ability to separate space into subspaces [149]. But, the effect of the transformation leads to an improvement of the speed of convergence for the optimization of the backpropagation during the learning process explained by a better homogeneity in the distribution of the weights and features in the input layer.

On the contrary, both the Parzen estimator and the k nearest neighbor estimator use the concept of data proximity to classify a new vector. They are then penalized by extreme values. When the extreme values are moved closer by transformation, their performances equal those of the neural network. The linear and the quadratic classifier make an assumption on the shape of the classes which is still not completely verified even after data transformation.

A.4 Discussion

Disagreements between human scorers are known to vary from 10% to 20% [150]. The results obtained by these classifiers do not enter this interval, but they are not very far from them. Besides, this study enables us to compare different techniques of classification.

The advantage of the neural network is that it does not require any data transformation. The results obtained are the same with the raw data or with the transformed data. The neural network can deal with a non gaussian probability density function and with extreme values. However, the selection of the best neural network and the optimization of the structure of the layers are no easy tasks and can be time consuming. Though the results obtained by the nearest neighbor classifier applied to homogeneous data are the best, this method requires storing a large amount of learning vectors in memory. This can make its application difficult in practice. The main advantage of classical statistical techniques (linear and quadratic discrimination) is that the algorithms are fast.

Why results do not enter the inherent interval of disagreement under scorer? One answer is that the hypothesis of the independence of the temporal epochs is not completely true because when experts score a recording, they intrinsically know the preceding page and score the new one in consequence. A way to take into account this temporal causality is to add new columns in the database corresponding to the preceding data of the parameters (switching from a state representation to a phase one). Another way is to introduce an inference table at the end of the classification process allowing certain transitions or rejecting others.

Another answer is that the parameters retained are not as discriminative as the ones chosen visually by the expert. This is a problem which is generally encountered in automatic classification as a mean to replace human classifier. Moreover, the classifiers were trained to classify data recorded on different subjects. They had to deal not only with temporal dependence of the data as discussed before, but also with inter individual variability. In our study, we have constructed and evaluated classifiers with no adaptation to one particular individual.

Experts are not so strict and often adapt their mind to fit the problem. classifiers are built to fit optimized mathematical models from a learning database. The solutions proposed by these models can sometimes show the limits of the visual technique of human scoring and can be a way to refine the expert knowledge.

For example, one visual interesting dilemma in the R&K manual is when a transition occurred during an epoch from one stage to another : the rule is to assign the class to the predominant stage, i.e. the stage that last more than fifty percent of the epoch. When there is a lot of transitions, this results in a lot of problems for the scorer and a lower productivity. This also puts a doubt over the stationarity hypotheses for the computing of temporal or spectral parameters. For a more accurate precision, one can use recent segmentation techniques for temporal time series, where signal is segmented into non overlapping windows of variable lengths with respect to different criteria [151, 64, 152, 77]. But then, the estimation of the performance of the classifiers is not so easy.

Nowadays, R&K scoring proves its usefulness every days, but its limits are more and more admitted [153].

A.5 Conclusion

We have evaluated and compared the performance of five classifiers to automatically score polysomnographic data from various individuals into the six R&K sleep-wake stages. Though the results obtained (the misclassification percentage is about 30%) are not as good as the results obtained with the human scorers (the misclassification percentage is

inferior to 20%), the results are interesting considering the amount of work human scoring requires. Automatic scoring may lighten the doctor's burden.

We showed that extreme values, frequently present in biological data, were a problem for all the evaluated classifiers, except for the neural network. To apply a transformation toward normal distribution appeared to be an interesting way to improve the performance of the classifiers. Both the neural network and the k nearest neighbor algorithm using transformed data gave good results. However, considering the information required to implement the two methods, we would recommend the use of the neural network.

Annexe B

Indices de Gravités

Cet appendice fournit les principales tables utilisées pour le scoring et l'évaluation de l'état de santé d'un individu. Les principales mises à jour peuvent être consultées auprès de la Société Française de Anesthésie et de Réanimation, <http://www.sfar.org>

GCS Le score de Glasgow (angl. *Glasgow Coma Scale*) est un indice de gravité de l'atteinte neuro-physiologique. Créé en 1974 par Teasdale et Jennett [154], il est d'un emploi très courant et il rentre dans le calcul de nombreux autres scores dont l'IGS et le TRISS.

IGS L'indice de gravité simplifié (angl. *SAPS : Simplified acute physiology score*) a été publié par Le Gall. L'IGS I comprend 14 paramètres cliniques ou biologiques dont l'âge du patient. Chaque paramètre est coté de 0 à 4 selon son degré d'anormalité. L'IGS II a été établi à partir d'une base de données plus importante et propose un risque de décès hospitalier. Il comprend 17 variables (12 physiologiques, l'âge, le type d'admission et 3 maladies chroniques sous-jacentes).

APACHE Le score APACHE (angl. *Acute physiology and chronic health evaluation*) a été publié par Knaus et coll. Il comprend 3 parties : variables physiologiques, âge, et état de santé chronique. L'APACHE II propose une révision du score précédent.

Score ISS, AIS L'ISS (angl. *Injury Severity Score*) de Susan Baker [155], 1974, est basé sur un score plus ancien qui associe à une détermination anatomique, une gravité traumatique, l'AIS (angl. *Abbreviated Injury Scale*) (1971). Le calcul de l'ISS consiste à prendre une expression quadratique des 3 principales lésions AIS. Ceci permet d'avoir un meilleur pronostique pour la mortalité suivant l'âge du polytraumatisé.

Score TS, Score TRISS Le TS (angl. *Trauma Score*) de Champion [129], spécialiste incontournable du triage et des situations de catastrophes, prend en compte le rythme respiratoire, l'effort respiratoire, la *PAS* prise par brassard, le remplissage capillaire (angl. *capillary refill*) et le score de Glasgow GCS. Il propose la création d'un indice mixte par régression linéaire (algorithme de régression de Walker-Duncan) : le TRISS (TRauma Score, ISS, age combination index). Le TRISS permet d'obtenir la probabilité de survie suivant un modèle logistique (régression multiple, modèle linéaire général). Le calcul d'un score $S(\mathbf{x})$ ou statistique d'Anderson permet de calculer la probabilité de survie :

$$P = \frac{1}{1 + \exp(-S(\mathbf{x}))} = \frac{\exp(S(\mathbf{x}))}{1 + \exp(S(\mathbf{x}))} \quad (\text{B.1})$$

on dit encore que P est *fonction logistique* du *score* [69]. Pour le calcul du TRISS, on utilise une somme pondérée des scores TS, ISS et A (correspondant à un score lié à une classe d'âge), en notant plus couramment $b = S(\mathbf{x})$, on obtient :

$$b = b_0 + b_1(TS) + b_2(ISS) + b_3(A) \quad (\text{B.2})$$

et la probabilité de survie :

$$P = \frac{1}{1 + \exp(-b)} = \frac{\exp(b)}{1 + \exp(b)} \quad (\text{B.3})$$

TAB. B.1 – TS : Trauma Score [129].

Trauma score		Value	Points	Score
A. Respiratory Rate		10–24	4	A. _____
Number of respirations in 15sec, multiply by four		25–35	3	
		>35	2	
		<10	1	
		0	0	
B. Respiratory Effort				B. _____
Shallow—markedly decreased chest movement or air exchange		Normal	1	
Retractive—use of accessory muscles or intercostal retraction		Shallow, or retractive	0	
C. Systolic blood pressure		>90	4	C. _____
Systolic cuff pressure—either arm—auscultate or palpate		70–90	3	
		50–69	2	
		<50	1	
		0	0	
D. Capillary refill				D. _____
Normal—forehead, lip mucosa or nail bed color refill in 2 sec		Normal	2	
Delayed—more than 2 sec of capillary refill		Delayed	1	
None—no capillary refill		None	0	
E. Glasgow Coma Scale		Total		
		<u>GCS Points</u>	<u>Score</u>	
1. Eye opening				E. _____
Spontaneous _____4		14–15	5	
To voice _____3		11–13	4	
To pain _____2		8–10	3	
None _____1		5–7	2	
		3–4	1	
2. Verbal response				
Oriented _____5				
Confused _____4				
Inappropriate words _____3				
Incomprehensible words _____2				
None _____1				
3. Motor response				
Obeys commands _____6				
Purposeful movement (pain) _____5				
Withdraw (pain) _____4				
Flexion (pain) _____3				
Extension (pain) _____2				
None _____1				
Total GCS point (1+2+3)_____		Trauma score_____		
		(Total points A+B+C+D+E)		

Score RTS Le score RTS (angl. *Revised Trauma Score*) [130] est une révision du score précédent, il le simplifie et n'utilise que le GCS, la pression artérielle systolique et le rythme respiratoire.

TAB. B.2 – RTS : Revised Trauma Score

Glasgow Coma Scale	Systolic Blood Pressure (mmHg)	Respiratory Rate (min)	Coded Value
13–15	>89	10–29	4
9–12	76–89	>29	3
6–8	50–75	6–9	2
4–5	1–49	1–5	1
3	0	0	0

$$TS = 0.9368(GCS) + 0.7326(SBP) + 0.2908(RR)$$

Score PHI Le PHI (angl. *prehospital index*) [156, 157], propose une distinction entre trauma majeur et mineur pour le triage. Les variables ont été choisies suivant leur corrélation avec la gravité du traumatisme. Les caractéristiques de *PAS*, pouls (c.-à-d., la fréquence cardiaque (*FC*)), statut respiratoire (angl. *respiratory status*) et état de conscience (*consciousness*) ont été retenus. C’est le plus utilisé à ce jour en Amérique du Nord [44].

TAB. B.3 – PHI : Prehospital Index

Components	Value	Score
Blood Pressure	> 100	0
	86 to 100	1
	75 to 85	2
	0 to 74	5
Pulse	≥ 120	3
	51 to 119	0
	< 50	5
Respirations	Normal	0
	Labored/Shallow	3
	< 10/Min/Needs intubation	5
Consciousness	Normal	0
	Confused/Combative	3
	No intelligible words	5
Total		0–20

0 to 3 $\rightarrow$ minor trauma.

4 to 20 $\rightarrow$ major trauma.

(penetrating abdominal or chest injuries given four points in addition to the calculated PHI)

Score CRAMS L’échelle CRAMS (Circulation, Respiration, Abdomen, Motor, Speech) [158] propose une évaluation des principales fonctions qui constituent son acronyme et essaie de simplifier encore l’évaluation pour le triage du blessé. Il catégorise ceux ci en 2 groupes : les traumatisés mineurs et les traumatisés majeurs.

Score AP le score AP (angl. *Anatomic Profile*) de Copes et coll. [159] remet en question l’ISS car il sous estime la sévérité, pour des patients avec des blessures multiples dans une ou plusieurs régions ; il ne tient pas compte des différences suivant les combinaisons des blessures ; il attribue le même poids pour différentes régions. Le score AP tient compte de 4 catégories de blessures désignées par des lettres : A résume toutes les blessures « sérieuses » à la tête et au cou ; B et C sont des scores

TAB. B.4 – CRAMS : Circulation, Respiration, Abdomen, Motor, Speech Scale.

Components		Score
Circulation	Normal capillary refill and blood pressure > 100	2
	Delayed capillary refill or blood pressure > 85 < 100	1
	No capillary refill or blood pressure < 85	0
Respiration	Normal	2
	Abnormal (labored or shallow)	1
	Absent	0
Abdomen	Abdomen and thorax nontender	2
	Abdomen or thorax tender	1
	Abdomen rigid or flail chest	0
Motor	Normal	2
	Responds only to pain (other than decerebrate)	1
	No response (or decerebrate)	0
Speech	Normal	2
	Confused	1
	No intelligible word	0
Total		0–10

Score $\leq 8 \rightarrow$ major trauma.Score $\geq 9 \rightarrow$ minor trauma

analogues pour le thorax, l'abdomen, le bassin, et les autres régions du corps ; D résume toutes les blessures « non sérieuses ». La cotation dans l'article initial se fait à partir de l'AIS-85 ou l'ICD-9-CM (classification internationale des maladies) et des données de l'American College of Surgeons' Major Trauma Outcome Study (MTOS) (les calculs des coefficients du modèle ont été réalisés sur 120000 patients traités dont l'âge varié de 15 à 54 ans). Les calculs statistiques réalisés sur cette base de données importantes permettent de donner une bon crédit à son utilisation.

TAB. B.5 – AP : Anatomic Profile, Assignation des composantes suivant les blessures

Component	Injury	AIS Severity	ISS Body Regions	ICD-9-CM Codes
A	Head/brain	3–5	1	800, 801, 803, 850–854
	Spinal cord	3–5	1, 3, 4	806, 950, 952, 953
B	Thoracic	3–5	3	807, 839.61/.71, 860–862, 901
	Front of neck	3–5	1	807.5/.6, 874, 900
C	Abd./pelvis	3–5	4	863–868, 902
	Spine w/o cord	3	1, 3, 4	805, 839
	Pelvic fracture	4–5	5	808, 839.42/.52/.69/.79
	Femoral Artery	4–5	5	904.0/.1
	Crush above knee	4–5	5	928.00/.01, 928.8
	Amputation above knee	4–5	5	897.2/.3/.6/.7
	Popliteal artery	4	5	904.41
D	Face	1–4	2	802, 830
	All others	1–2	1–6	–

l'ASCOT (angl. *a severity characterization of trauma*) C'est une révision du TRISS et ne se fonde plus sur l'ISS mais sur l'AP. L'âge est séparé en plusieurs catégories.

Score de Lindsey d'après [7] qui référence sur [160], utilisent des composantes mixtes

et semble facile d'emploi.

TRISS-like On peut encore noter le récent score TRISS-like [161] qui a été réalisé de façon à évaluer l'état des patients directement intubés sur le terrain. En effet, pour ce type de personnes, l'évaluation de la réponse verbale n'est plus possible pour l'élaboration du score GCS et il y a impossibilité d'obtenir une estimation de la ventilation naturelle. Ce score utilise la pression artérielle systolique ainsi que le pouls (fréquence cardiaque).

TAB. B.6 – ASCOT : A Severity Characterization of Trauma

	AIS3	AIS4	AIS5		
Organes	1-6	1-6	1-6		
Tête, Encéphale, Moëlle	1-6	1-6	1-6		
Thorax, partie antérieure du cou	1-6	1-6	1-6		
Autres localisations	1-6	1-6	1-6		
RTS					
Fréquence Respiratoire (cpm)	0 0	1-5 1	6-9 2	10-29 4	≥30 3
Pression Artérielle Systolique (mmHg)	0 0	1-49 1	50-75 2	76-89 3	≥90 4
Score de Glasgow	3 0	4-5 1	6-8 2	9-12 3	13-15 4
Age	≤54 0	55-64 1	65-74 2	75-84 3	≥85 4

Pour chaque profil anatomique :

$$Points = \sqrt{3^2(nbre\ de\ lésions\ AIS\ 3)} + 4^2(nbre\ AIS\ 4) + 5^2(nbre\ AIS\ 5)$$

ASCOT (lésion fermée) : Mortalité prédite = $\frac{1}{1+e^K}$

$$K = -1.1570 + 0.2810(points\ FR) + 0.6583(points\ PAS) + 0.7705(points\ Glasgow) - 0.3002(points\ AIS\ lésions\ tête,\ encéphale,\ moelle) - 0.1961(points\ AIS\ lésions\ thorax,\ cou) - 0.2086(points\ AIS\ autres\ localisations) - 0.6355(points\ pour\ Age)$$

ASCOT (lésion pénétrante) : Mortalité prédite = $\frac{1}{1+e^K}$

$$K = -1.1350 + 0.3332(points\ FR) + 0.3638(points\ PAS) + 1.0626(points\ Glasgow) - 0.3702(points\ AIS\ lésions\ tête,\ encéphale,\ moelle) - 0.2053(points\ AIS\ lésions\ thorax,\ cou) - 0.3188(points\ AIS\ autres\ localisations) - 0.8365(points\ pour\ Age)$$

L'ASCOT utilise trois composants (A, B, C) pour les profils anatomiques, le quatrième (D) regroupant la somme de tous les scores pour lésions minimales (AIS 1 ou 2). Il est exclu de ce modèle. Les données physiologiques des patients sont colligées à l'admission aux urgences (NB : selon l'organisation sanitaire US). Les patients avec un très bon ou très mauvais pronostic sont exclus du modèle avec régression logistique. Les données concernant ces patients *A-part - Set-Aside* sont colligées dans le tableau ci dessous.

Set-Aside	Description du Patient	Lésion fermée (% mortalité)	Lésion pénétrante (% mortalité)
1	AIS 6 RTS = 0	100	100
2	MAX AIS < 6, RTS = 0	98,6	97,4
3	AIS 6, RTS > 0	77,1	77,8
4	MAX AIS = 1 or 2, RTS > 0	0,2	0,1

Lésion non pénétrante *Blunt* : Lésion non pénétrante, mais incluant *crush syndrome*, lacération, amputation et asphyxie.

Lésion pénétrante *Penetrating* : Lésion secondaire à un projectile d'arme à feu, couteau ou autre pointe.

TAB. B.7 – Score de Lindsey

Score	1	3	5	6
Région anatomique	Membres	Dos	Thorax	Tête, cou, abdomen
Type de lésion	Déchirure	Contusion	Arme blanche	Arme à feu
App. cardiovasculaire	Normotendu	PAS* : 60-100	PAS < 60	Rien
Respiration	Douloureuse normale	Dyspnée	Cyanose	Apnée
État de conscience	Obnubilation	Stupeur	Confusion	Coma
	A =	B =	C =	D =

* = Pression Artérielle Systolique

Score = A + B + C + D

Résultats :

— 2-9 Bénin

— 10-16 Modéré (hospitalisation obligatoire)

— 17-20 Sérieux mais pronostic vital bon

— 21 et plus Polytraumatisé, mortalité élevé

TAB. B.8 – Poids pour les modèle de régression logistique pour le calcul de la probabilité de survie d'après les score de TS, TRISS, AP suivant les études réalisées [129, 162, 159]. MTOS : Major Trauma Outcome Study.

	b_0	b_1	b_2	b_3	b_4
TS [129]					
Total Set	-6.9383	0.7124			
Critical Subset	-7.0710	0.7424			
TRISS [129]					
Total Set	-5.9583	1.8384	-0.4281	-9.5624	
Critical Subset	-3.5170	0.6650	-0.0808	-1.5938	
TRISS [162] MTOS coefficients					
Blunt	-1.6465	0.5175	-0.0739	-1.9261	
Penetrating	-0.8068	0.5442	-0.1159	-2.4782	
RTS et TRISS [162] MTOS coefficients					
Blunt	-1.2470	0.9544	-0.0768	-1.9052	
Penetrating	-0.6029	1.1430	-0.1516	-2.6676	
ISS [159] : $b_1 = ISS$					
	4.3705	-0.1064			
AP [159] : $b_1 = A, b_2 = B, b_3 = B^2, b_4 = C^2$					
	4.0801	-0.4914	-0.2066	-0.0161	-0.0351

Annexe C

Techniques de validation

C.1 Matrice de confusion

C.1.1 Problèmes à deux classes

Dans le cadre d'un problème à deux classes, on est réduit au tableau suivant (Tab. C.1).

L'expert est en ligne, la machine en colonne. On peut alors associer à cette matrice de confusion plusieurs définitions :

- 1 : Cas Positif
- -1 : Cas Négatif
- A : VP, Vrais Positifs
- B : FP, Faux Positifs
- C : FN, Faux Négatifs
- D : VN, Vrais Négatifs
- ACC , $ACC = \frac{(A+D)}{(A+B+C+D)}$, Concordance, Accord (angl. *Accuracy*)
- PPV , $PPV = \frac{A}{(A+C)}$, valeur prédictive positive (angl. *positive predictive value*)
- NPV , $NPV = \frac{D}{(B+D)}$, valeur prédictive négative (angl. *negative predictive value*)
- SEN , $SEN = \frac{A}{(A+B)}$, sensibilité (angl. *sensitivity*)
- SPE , $SPE = \frac{D}{(C+D)}$, spécificité (angl. *specificity*)
- PRE , $PRE = \frac{A}{(A+C)}$, précision (Precision = PPV) (angl. *precision*)
- REC , $REC = \frac{A}{(A+B)}$, rappel (Recall = SEN) (angl. *recall*)
- ROC curve, (angl. *Receiver Operating Curve*) : On trace la sensibilité en fonction de la spécificité.

TAB. C.1 – Matrice de confusion (cas 2 classes)

		Machine (prédiction du modèle)		
		1	-1	
Expert (sortie réelle)	1	A	B	$A + B$
	-1	C	D	$C + D$
		$A + C$	$B + D$	$A + B + C + D$

- *ROC area*, (angl. *Area Under Curve*) : Aire sous la courbe : si elle est proche de 1, la classification est optimale; si elle est proche de 0.5, on n'a pas de relation entre prédiction et réalité.

C.1.2 Problèmes à plusieurs classes

Pour un problème à plusieurs classes, les opérateurs de performances décrits plus haut pour un cas à deux classes ne sont plus valables¹. On préférera alors utiliser la concordance, l'accord entre l'expert et la machine Acc , ou encore l'erreur de classement Err , définie par :

$$Acc = \sum_{i=j} \frac{M_{ij}}{m} \quad (C.1)$$

$$Err = \sum_{i \neq j} \frac{M_{ij}}{m} = 1 - Acc \quad (C.2)$$

C.1.3 Techniques de Rééchantillonnage

De façon à évaluer les performances des différents classifieurs, on peut procéder à différentes méthodes pour créer les ensembles d'apprentissage et de test. Les méthodes qui permettront de sélectionner des ensembles d'apprentissage et de tests les plus indépendants possibles permettront d'obtenir des résultats les moins biaisés possibles. Toutes ces méthodes reposent sur des méthodes de rééchantillonnage de données. Une description des principales méthodes utilisées est la suivante, pour de plus amples informations on pourra se reporter à [120, 109, 107, 26]

Méthode de resubstitution Cette méthode consiste à apprendre et évaluer les performances sur les mêmes exemples d'un ensemble de données. On obtient alors une estimation très optimiste des performances des classifieurs testés. Pour des ensembles de données de faible taille où la distinction des exemples est bien réalisée, on aura une évaluation très optimiste. Pour des ensembles de grandes dimensions, où la richesse des exemples permet d'éviter un apprentissage par cœur et permet une généralisation, on se rapproche d'une évaluation plus réaliste des performances. En particulier, pour des classifieurs de type du 1 plus proche voisin et des arbres de décision non élagués, cette méthode conduit souvent à une estimation des performances trop optimiste ($Err = 0$, $Acc = 1$) puisque ces techniques conduisent à un apprentissage prototypique, un apprentissage par cœur. $err_{\mathcal{A}}$, erreur obtenue sur l'ensemble d'apprentissage introduite précédemment est obtenue par resubstitution sur l'ensemble d'apprentissage, c'est encore une évaluation de l'erreur empirique R_{emp} .

Méthode de holdout (angl. *holdout*) Cette méthode consiste à séparer l'ensemble des données disponibles (ensemble $\mathcal{S}$) en deux ensembles disjoints pour l'apprentissage $\mathcal{A}$ et le test $\mathcal{T}$ ($\mathcal{S} = \mathcal{A} \cup \mathcal{T}$). On obtient alors une estimation d'une erreur plus pessimiste obtenue sur l'ensemble de test $err_{\mathcal{T}}$ pour la performance des classifieurs et une erreur sur l'ensemble d'apprentissage $err_{\mathcal{A}}$ plus optimiste. Les exemples de la base sont utilisés une seule fois, soit pour l'apprentissage soit pour le test.

¹On peut encore utiliser des statistiques pour évaluer les performances de classement et comparer les classifications : coefficient Tau de concordance de Kendall, coefficient Kappa de Cohen, test de χ^2 (angl. *goodness of fit*), test de McNemar [26, 123]

Sous échantillonnage aléatoire (angl. *random subsampling*) On peut effectuer un tirage aléatoire pour la création des ensembles d'apprentissage et de test pour la méthode de holdout. On peut alors répéter l'opération plusieurs fois et réaliser une moyenne sur l'ensemble des erreurs calculées. On obtient alors, si k est le nombre de tirages réalisés :

$$\bar{Err} = \frac{1}{k} \sum_{i=1}^k Err_i \quad (C.3)$$

avec Err erreur sur l'ensemble généré par le $i^{ème}$ tirage aléatoire (en apprentissage ou en test, on obtient alors une estimation plus juste de $err_{\mathcal{A}}$ et $err_{\mathcal{T}}$). Le résultat est généralisable pour de grandes bases de données en utilisant des sous échantillons de façon à ne pas avoir à réaliser des calculs sur l'ensemble de la base de données et limiter ainsi le nombre d'opérations à effectuer par la machine. C'est une technique que nous avons utilisé pour l'évaluation de nos performances.

Validation Croisée (angl. *cross validation*) On divise l'ensemble des exemples disponibles en k divisions exclusives de tailles équivalentes (angl. *k-fold cross validation*). On effectue k procédures visant à réaliser un apprentissage sur $\mathcal{A}_i = \mathcal{S} \setminus \mathcal{T}_i$ et à tester sur $\mathcal{T}_i$ correspond à la $k^{ième}$ division réalisée. La permutation circulaire permet d'obtenir k estimations des erreurs d'apprentissage et de test. On utilise l'équation C.3, pour obtenir une estimation des différentes erreurs. Cette technique n'est pas réalisable pour des ensembles de petites tailles mais devient très compétitive pour des ensembles de grandes tailles. En effet, cette technique permet de lisser les effets des perturbations de la base de données et permet de réaliser une estimation sans biais de l'erreur pour les classifieurs stables. On parle encore de validation croisée complète lorsqu'on évalue les résultats de la validation croisée sur tous les choix possibles de k divisions de l'ensemble initial $\mathcal{S}$.

Bootstrap Les techniques de bootstrap² consistent à générer b ensembles rééchantillonnés de même taille n que l'ensemble initial ($\mathcal{S}$) par tirage uniforme avec remise (on permet donc de tirer plusieurs fois le même échantillon). Pour chaque ensemble rééchantillonné $\mathcal{S}_i$ (échantillon bootstrap), on apprend sur l'ensemble $\mathcal{A}_i$ (avec $\mathcal{A}_i = \mathcal{S}_i$) et on évalue la performance sur un premier ensemble de test que l'on choisi égal à l'ensemble initial $\mathcal{T}_{1i} = \mathcal{S}$. On effectue une seconde évaluation des performances sur un ensemble de test $\mathcal{T}_{2i}$ constitué des échantillons non présents dans l'ensemble bootstrap $\mathcal{A}_i$ et restant dans l'ensemble initial $\mathcal{S}$. Pour un ensemble de taille n , la probabilité pour qu'un échantillon ne soit pas choisi après n tirages avec remise est $(1 - 1/n)^n \approx 0.368$ pour n grand, et la probabilité pour que celui ci appartienne à l'ensemble de test $\mathcal{A}_i$ est alors de $1 - (1 - 1/n)^n \approx 0.632$ pour n grand. On obtient ainsi une estimation pour l'erreur par la méthode de bootstrap (estimation de $R_{réel}$

²Le terme de bootstrap fait référence au roman fantastique « les aventures du baron de Munchausen » de Rudolph Eric Raspe (1785) dans lequel le baron de Münchhausen réussit à voler par lui même, simplement en tirant fortement sur les lanières de ses chaussures (angl. *bootstraps*). Ce terme est encore utilisé pour la séquence bootstrap ou *boot* d'un ordinateur, qui indique la séquence d'initialisation qui permet à la machine de charger des logiciels. De la même façon que le baron s'auto-élève dans le ciel avec ses lacets, la machine charge ses logiciels avec ses propres logiciels. Pour l'apprentissage artificiel, la machine utilise les mêmes données pour apprendre les modèles et les évaluer. Cyrano de Bergerac utiliserait cette même technique pour parvenir aux étoiles...

pour une fonction de coût $\{0, 1\}$:

$$Err_{bootstrap} = \frac{1}{b} \sum_{i=1}^b (0.632Err_{T_{1i}} + 0.368Err_{T_{2i}}) \quad (C.4)$$

Jackknife C'est une validation croisée à n blocs, n étant le nombre d'exemples disponibles dans l'ensemble initial $\mathcal{S}$. On apprend sur $n - 1$ exemples et on teste sur le dernier. L'estimateur Jackknife³ est défini ainsi :

$$Err_{jackknife} = nErr_{\mathcal{S}} - (n - 1) \times \frac{1}{n} \sum_{i=1}^n err_{\mathcal{A}_i} \quad (C.5)$$

avec $Err_{\mathcal{S}}$ taux d'erreur obtenu par resubstitution sur l'ensemble initial $\mathcal{S}$ (on a alors $\mathcal{S} = \mathcal{A} = \mathcal{T}$) et $err_{\mathcal{A}_i}$ erreur obtenue sur l'ensemble d'apprentissage $\mathcal{A}$ constitué de $n - 1$ exemples avec le $i^{ème}$ exemple mis de côté. L'estimateur Jackknife est efficace pour des estimations connues pour être biaisée en $1/n$.

Leave one out On laisse un exemple de côté pour l'apprentissage et on teste sur ce dernier (méthode du « laissé-pour-compte »). C'est une méthode de Jackknife mais l'erreur n'est pas calculée de la même façon. On estime l'erreur de la même façon qu'une validation croisée.

Partition stratifié (angl. *stratified partitioning*) On parle de stratification lorsqu'on constitue des ensembles de validation croisé en sélectionnant des groupes d'individus de façon à obtenir une statistique voulue. Par exemple, on choisit un nombre identiques de représentants de chaque classe pour chaque divisions lors de l'évaluation d'une classification par validation croisée stratifiée.

Ces différentes méthodes constituent la base des techniques disponibles et nous orientons le lecteur intéressé vers les ouvrages de statistiques, théorie de l'estimation ou autres traités d'analyse de données pour plus de renseignements.

³Jackknife fait référence au couteau suisse qui sert à tout faire

Annexe D

Correspondances avec le Dr Glass

D.1 Courrier du 18 mars 2004 adressé à Glass

From: "Guillaume Becq" <guillaume.becq@crssa.net>
To: "Todd Glass" <todd.glass@chmcc.org>
Subject: G. Becq, questions about : Use of AI to identify cardiovascular compromise in a model of hemorrhagic shock
Date: Thu, 18 Mar 2004 17:47:42 +0100
MIME-Version: 1.0
Content-Type: multipart/alternative;
boundary="-----_NextPart_000_0038_01C40D11.1CB97DA0"
X-Priority: 3
X-MSMail-Priority: Normal
X-Mailer: Microsoft Outlook Express 6.00.2800.1106
X-MimeOLE: Produced By Microsoft MimeOLE V6.00.2800.1106

Dear M. Glass,

I am currently a PhD student in the Centre de Recherches du Service de Santé des Armées, (French Military Health Service Research Center) in Grenoble, France. I am doing my third year of PhD thesis on the "Evaluation to the realizability of a portable device scoring trauma patient's emergency state". In order to evaluate this device, we decided to use an animal model of swine enduring a continuous blood loss (50% blood loss in 1h30min). From 2002 to 2003, we experimented this model and since then, we are working on the results made upon features extraction and machine learning techniques.

We were very surprised (and enjoyed) to read the article from you and your partners : Use of artificial intelligence to identify cardiovascular compromise in a model of hemorrhagic shock (Critical Care Medicine 2004, 32 :450-456), dealing with approximately the same experimentation and multivariate analysis approach. As you noticed 'There are no reports of using cognitive system technology to characterize circulatory dysfunction following traumatic shock', we thought in the same direction.

The work you present is very impressive (especially the results, since we can obtain such performances with other features and evaluation techniques), but it rises some questions :

1) Are the animals in natural ventilation during the blood loss protocol or are they assisted by a mechanical ventilation (air mixture control or/and Respiratory Rate set point) ?

2) How many neighbors did you use for the KNN classifier (the value of K) ?

3) Are 9 P-QRS-T sequences used for learning and testing ? or did you learn models from individual P-QRS-T segment then test models with the decision obtain with median filtering applied to the 9 P-QRS-T outputs obtained with the learned model ?

4) How do you compute the 95% confidence interval for error estimation (what is the statistical distribution model for the error) ?

5) You are talking about non linear analysis (second order linear analysis) in the discussion, but it seems that you do not use any power spectral analysis or variability measure. Can you precise that point ? In fact, when you look Fig.3, there are sort of piecewise linear response between variables and it is observed in approximately all presented animals. Then, the quadratic classifier can give good information about the correlation between variables in these classes and can precise piecewise linear responses.

6) The more critical point : Do you learn 19 classifiers, each one corresponding to one animal (i.e. learned with 50% random samples drawn from one animal data), and test the battery of classifiers with the remaining data of each animal ? This means that it exists a strong temporal dependence between samples and between individual, resulting in a biased evaluation of the error of your model. What is your opinion about this point ?

I am waiting for your answer with interest. If you can not answer some questions due to interdisciplinary specialization, can you connect me to the people in charge of the CST development.

Yours, faithfully
Guillaume

```
-----
Guillaume Becq
CRSSA
Centre de Recherches du
Service de Santé des Armées
24, Avenue des Maquis du Grésivaudan
BP 87
38702 La Tronche cedex
FRANCE
tél. : 04 76 63 69 75
por. : 06 63 70 93 48
fax : 04 76 63 69 45
guillaume.becq@crssa.net
-----
```

D.2 Réponse du 19 mars 2004 adressée à Becq

```
X-From_: Todd.Glass@cchmc.org Fri Mar 19 19:05:52 2004
Return-Path: <Todd.Glass@cchmc.org>
Received: from n6mcgw16.cchmc.org ([205.142.197.62])
by smtp3.clb.oleane.net with ESMTP id i2JI5pCx025878
for <guillaume.becq@crssa.net>; Fri, 19 Mar 2004 19:05:52 +0100
Received: from DOMSVC03-MTA by n6mcgw16.cchmc.org
with Novell_GroupWise; Fri, 19 Mar 2004 13:05:50 -0500
Message-Id: <s05af02e.079@n6mcgw16.cchmc.org>
X-Mailer: Novell GroupWise Internet Agent 6.5.1
Date: Fri, 19 Mar 2004 13:05:32 -0500
From: "Todd Glass" <Todd.Glass@cchmc.org>
To: <guillaume.becq@crssa.net>
Subject: Re: G. Becq, questions about : Use of AI to identify
cardiovascular compromise in a model of hemorrhagic
Mime-Version: 1.0
Content-Type: text/plain; charset=ISO-8859-1
Content-Transfer-Encoding: quoted-printable
Content-Disposition: inline
```

I appreciate your comments and interest in our paper. I was also interested to learn that you have been working on essentially the same problem with a very similar approach. Have you published any of your work ? What specific features are your group pursuing ?

Regarding your questions, I will respond to some and forward others to engineers with whom I am working.

thank you very much,

Todd F. Glass, MD
Assistant Professor
Pediatric Emergency Medicine
Cincinnati Children's Hospital Medical Center
Cincinnati, OH, USA

D.3 Réponse du 19 mars adressée à Glass

```
From: "Guillaume Becq" <guillaume.becq@crssa.net>
To: "Todd Glass" <Todd.Glass@cchmc.org>
References: <s05af02e.079@n6mcgw16.cchmc.org>
Subject: Re: G. Becq, questions about : Use of AI to identifycardiovascular
compromise in a model of hemorrhagic
Date: Fri, 19 Mar 2004 23:06:26 +0100
MIME-Version: 1.0
Content-Type: text/plain;
charset="iso-8859-1"
Content-Transfer-Encoding: 8bit
X-Priority: 3
X-MSMail-Priority: Normal
X-Mailer: Microsoft Outlook Express 6.00.2800.1106
X-MimeOLE: Produced By Microsoft MimeOLE V6.00.2800.1106
```

We haven't published yet but articles are in preparation (for ex, one for IEEE TBME04 conference in San Francisco). Models with invasive features (PAS, PAD, PAM, pulsedPA, PVC, Tint, EtCO2) were evaluated and models with non invasive ones (pulseOximeter, HR, SDECG, TCut, RRabd, RRtho, Tamb...). The target class was the different states of hemorrhagic shock described by Guyton or other physiologist (compensated, uncompensated, irreversal state). Yours, sincerely,

Guillaume

Annexe E

Liens

Quelques adresses internet utiles :

<http://www.sfar.org>
<http://www.splf.fr/>
<http://www.srlf.org/s/>
<http://www.sante.gouv.fr>
<http://www.defense.gouv.fr>
<http://www-sante.ujf-grenoble.fr/SANTE/>

<http://www.univ-st-etienne.fr:1042/sfc/>

<http://www.nato.int>
<http://www.trauma.org>
<http://www.physionet.org>

<http://www.ecglibrary.com/>
<http://www.medal.org>
<http://www.chu-rouen.fr/pneumobg/ventiweb.htm>
<http://www.wealthy-ist.com/>
<http://kdd.ics.uci.edu/>
<http://www.ncbi.nlm.nih.gov/pubmed>

<http://citeseer.ist.psu.edu/cis>

Société française d'anesthésie et réanimation
Société de pneumologie de langue française
Société de réanimation de langue française
Ministère de la santé et de la protection sociale
Ministère de la défense
Corpus médical de la faculté

de médecine de Grenoble
Société francophone de chronobiologie

Organisation du traité de l'atlantique nord
Trauma.org
The research resource for complex physiologic signals
ECG library
The medical algorithms project
VentiWeb : le site francophone de la ventilation non invasive (VNI)
Projet européen Wealthy
The UCI KDD archive
PubMed, a service of the national library of medicine
CiteSeer : scientific literature digital library

Glossaire

Cette sélection est proposée comme une recherche d'une langue stable et compréhensible par la plus grande partie des disciplines concernées. Pour cela, il est souvent utile de comprendre l'étymologie du mot pour en extraire sa substance, sa signification lors de son utilisation dans un contexte particulier. Ce glossaire nous permettra de faire référence à une base commune, pour ne pas jouer sur les mots mais comprendre l'essence des idées que nous voulons faire passer, se rendre compréhensible.

Afférent adj. (lat. *ad*, vers ; *ferre*, porter) (angl. *afferent*) [28] Qui s'approche de ; centripète ; qui amène un fluide ou un influx nerveux vers un organe.

Aléa n. m. (lat. *alea*, coup de dé) : [25] Hasard, favorable ou non. ◇ (**souvent pl.**). Risque d'incidents défavorables, d'inconvénients. *cette affaire présente bien des aléas*.

Aléatoire adj. [25] **1.** Qui relève du hasard ; qui dépend d'un événement incertain ; hasardeux, problématique. *Bénéfices aléatoires*. **2.** MATH. *Variable aléatoire* : variable dont la variation dépend d'une loi de probabilité.

Anamnèse n. f. ou Anamnétiques (gr. *ana*, derechef, *mnasthai*, se souvenir) (angl. *anamnesia*) [28] Renseignement que fournit le malade lui-même ou son entourage sur le début de sa maladie jusqu'au moment où il se trouve soumis à l'observation du médecin. *V. antécédents*.

Anesthésie n. f. (gr. *an-* priv. , *aïsthêsis*, sensibilité) (angl. *anaesthesia*, amér. *anesthesia*) [28] ■ privation générale ou partielle de la faculté de sentir ■ (Littré). Elle peut être due à un état morbide, ou provoquée par un médicament. *a. artificielle*. *A.* tantôt limitée à une région du corps et due à l'action d'un agent anesthésiant sur les terminaisons des nerfs. (*a. locale* [*v. stabilisateur de membrane*] *a. régionale*) sur leurs troncs (*a. tronculaire*) ou sur leurs racines sensibles (rachianesthésie) ; tantôt générale, s'accompagnant de narcose et due à l'absorption de l'anesthésique par l'appareil respiratoire ou l'appareil digestif (voie buccale ou rectale) ou bien due à son introduction directe dans la circulation par voie intraveineuse. L'*a.* enfin peut être obtenue par l'emploi simultané de plusieurs de ces méthodes (*v. a. de base et anocie-associée*).

Anoxémie n. f. (gr. *an-* priv. , *oxus*, oxygène, *haïma*, sang) (angl. *anoxaemia*) (Jourdanet, 1861) [28] Diminution de la quantité d'oxygène contenu dans le sang (si elle est faible, on l'appelle hypoxémie). Cette diminution peut avoir différentes causes : dépression atmosphérique (mal d'altitude), anémie, anomalies ou altérations de l'hémoglobine par certains toxiques (aniline, nitrites, sulfamides), intoxication par le *CO*, pneumopathies chroniques, cardiopathies congénitales avec shunt veino-artériel, insuffisance cardiaque, etc. Elle provoque l'anoxie.

Anoxie n. f. (gr. *an-* priv. , *oxus*, oxygène) (angl. *anoxia*) [28] Diminution de la quantité d'oxygène distribuée aux tissus par le sang dans l'unité de temps ; elle est la conséquence de l'anoxémie. Lorsque cette diminution est faible elle est appelée hypoxie.

Anurie n. f. (gr. *an-* priv. , *ouron*, urine) (angl. *anuria*) [28] Absence d'urine dans la vessie. Elle est due à l'arrêt de la sécrétion rénale (*anurie vraie* ou *sécrétoire*) ou à un obstacle au cours de l'urine entre le rein et la vessie (*fausse anurie* ou *a. excrétoire*).

Archétype n. m. (gr. *arkhetupos*, modèle primitif) [25] **1.** Modèle sur lequel sont construits un ouvrage, une œuvre **2.** PHILOS. **a.** Idée, forme du monde intelligible sur laquelle sont construits les objets du monde sensible, chez Platon. **b.** Idée qui sert de modèle à une autre, pour les empiristes.

Artefact n. m. (lat. *ars facta*, effet de l'art) [25] Phénomène d'origine artificielle ou accidentelle, rencontré au cours d'une observation ou d'une expérience.

- Bruire** v. i. (lat. pop. *brugere*, braire, croisé avec *rugire*, rugir) Litt. Faire entendre un son, un murmure confus.
- Bruit** n. m. (de *bruire*) (angl. *noise*) [25] **5.** Perturbation indésirable qui se superpose au signal et aux données utiles, dans un canal de transmission, dans un système de traitement de l'information.
- Capnigramme** n. m. (gr. *kapnos*, vapeur, *gramma*, caractère d'écriture) (angl. *capnogram*) [28] Courbe indiquant la concentration d'un gaz en CO_2 (p. ex. de l'air expiré) en fonction du temps.
- Capnimétrie** n. m. (gr. *kapnos*, vapeur, *métron*, mesure) (angl. *capnometry*) [28] Mesure de la concentration en CO_2 dans un fluide.
- Caractère** n. m. (gr. *kharaktêr*, signe gravé) [25] **II.1.** Marque distinctive de qqch ou de qqn ; état ou qualité propre de qqn ou qqch.
- Caractéristique** n. f. (angl. *feature*) [25] Ce qui constitue la particularité, le caractère distinctif de qqn ou qqch.
- Cathéter** n. m. (gr. *kathiénai*, plonger) (angl. *catheter*) [28] Tube long et mince (sonde), flexible ou rigide, en métal, verre, gomme, caoutchouc ou matière plastique etc., destiné à être introduit dans un canal, un conduit, un vaisseau ou un organe creux pour l'explorer, injecter un liquide ou vider une cavité.
- Choc** n. m. (angl. *shock*) [28] Sidération brusque du système nerveux caractérisée par une stupeur, une hypovolémie, un collapsus cardiovasculaire (v. ce terme), parfois même oar des convulsions avec perte de connaissance, aboutissant soit à la mort, soit à la guérison rapide. Il résulte toujours « d'une diminution du volume du sang circulant par rapport à la capacité fonctionnellement active du lit vasculaire » (Jean Hamburger). Cette diminution entraîne une insuffisance brutale de l'irrigation d'organes d'importance vitale, privés d'oxygène et encombrés de déchets métaboliques acides qui ne sont plus évacués. Le *c.* peut relever de deux mécanismes : *1. une brusque réduction du retour du sang veineux au cœur* : par effondrement de la masse sanguine à la suite de pertes de sang ou de plasma (*c.* hémorragique, *c.* des brûlés, *c.* traumatiques, *c.* postopératoire) ; par déshydratation aiguë ou perte de sodium à la suite de diarrhée ou de vomissements abondants, au cours de la maladie d'Addison ou de l'acidose diabétique ; par vasodilatation généralisée avec stockage du sang dans les artérioles et les capillaires (*c.* anaphylactique, *c.* réflexe, *c.* toxique, *c.* infectieux ou bactériémique) ; dans tous ces cas il s'agit de *c. hypovolémique* (v. ce terme). *2. un fonctionnement cardiaque défectueux.* C'est alors le *c. cardiogénique* (v. ce terme) dû à un remplissage cardiaque insuffisant, gêné par une tamponade ou provoqué par une tachycardie extrême, ou bien dû à une insuffisance cardiaque aiguë au cours de l'infarctus du myocarde, des myocardites, de l'embolie pulmonaire massive, des sténoses aortiques et mitrales très serrées. V. *poumon de choc*.
- Circonstance** n. f. (lat. *circumstere*, se tenir autour) [25] **1.** Particularité qui accompagne un fait, une situation. **3.** Ce qui détermine un moment donné.
- Classe** n. f. (lat. *classis*) [25] **I. 1.** Groupe, ensemble de choses, de personnes ayant des traits communs ; catégories. MATH. Classe d'équivalence : dans un ensemble muni d'une loi d'équivalence, chacun des sous ensembles formés par les éléments équivalents entre eux deux à deux. **3.** Catégorie, rang attribué à qqn, à qqch, selon un ordre d'importance, de valeur, de qualité.
- Classificateur** adj. et n. m. [25] qui classe
- Classification** n. f. [25] **1.** distribution par classes, par catégories, selon un certain ordre et une certaines méthodes ; le résultat de cette opération.
- Classifier** [25] Procéder à la classification (d'un ensemble de données).
- Classifieur** Ne figure pas dans le dictionnaire. De la même façon qu'un multiplieur et un multiplicateur, ou de la même façon qu'un additionneur... Le choix du suffixe *eur* et *ateur*, signifie *agent* [25].
- Collapsus** n. m. (Cullen) (lat. *cum*, avec, *lapsus*, chute) (angl. *collapse*) [28] **1.** « Chute rapide des forces, par suite de laquelle les mouvements deviennent pénibles, la parole faible, le pouls dépressible ; c'est une sorte d'intermédiaire entre la syncope et l'adynamie » (Dechambre). Aujourd'hui pris souvent dans le sens de *collapsus cardiaque*. V. ce terme. **2.** Affaissement d'un organe (*c.* pulmonaire)
- Collapsus cardiaque ou cardiovasculaire** n. m. (angl. *cardiac shock*) [28] Syndrome, d'apparition brutale, caractérisé par un refroidissement des extrémités avec prostration considérable, sueurs profuses, cyanose, pouls rapide et imperceptible, chute de la tension artérielle systolique à 8 cm de mercure ou au dessous, oligurie ou anurie. Il peut être dû à l'effondrement subit de l'énergie

cardiaque ou à une hypotonie vasculaire primitive. On réserve souvent le nom de *c.* à un accident de courte durée, au cours duquel prédominent les réactions vagales et guérissant sans séquelle. V. *choc*, *choc cardiogénique* et *choc d'origine vasculaire*.

Consistance n. f. (du lat. *consistere*, se tenir ensemble) [25] **1.** État d'un corps considéré du point de vue de la cohésion de ses parties. *L'argile séchée a une consistance dure* ◇ *Prendre de la consistance* : devenir épais, solide. – **Fig.** Solidité, force. *Cet argument manque de consistance.* ◇ *Personne sans consistance*, qui manque de caractère, de personnalité. **2.** LOG. Non-contradiction.

Consistant, e adj. [25] **4.** LOG. Autosuffisant.

Constant, e adj. (lat. *constans*) [25] **1.** Résolu, persévérant, dans ses actes, ses sentiments, etc. **2.** Qui est continu, qui dure ou se répète de façon identique. *Bonheur constant* ◇ **litt.** *Il est constant que* : Il est évident que. **3.** MATH. *Fonction constante*, qui donne la même image de tous les éléments de son ensemble de définition.

Contexte n. m. (du lat. *contexere*, tisser ensemble) [25] **1. a.** Texte à l'intérieur duquel se situe un élément linguistique (phonème, mot, phrase, etc.) et dont il tire sa signification ou sa valeur. **b.** Condition d'élocution d'un discours, oral ou écrit **2.** Circonstances, situation globale où se situe un évènement. *Remplacer un fait dans son contexte historique.*

Contingence n. f. (du lat. *contingere*, arriver par hasard) Caractère de ce qui est contingent ; éventualité, possibilité que qqch arrive ou non.

Contingent adj. Qui peut se produire ou non (par opp. à *nécessaire*).

Contractilité n. f. (angl. *contractility*) [28], Propriété vitale que possèdent certaines cellules et surtout la fibre musculaire, de réduire une ou plusieurs de leurs dimensions en effectuant un travail actif. – la *c. du myocarde*, « régit », avec les conditions de travail du cœur que sont la précharge et la postcharge, la performance ventriculaire » (J. F. Landau).

Cybernétique n. f. (gr. *kubernân*, gouverner) (angl. *cybernetics*) [25] Étude des processus de commande et de communication chez les êtres vivants, dans les machines et les systèmes sociologiques et économiques. [163] Cybernetics is a science in which control systems in electronic and mechanical devices are studied and compared to biological systems.

Déduction n. f. (lat. *deducere*, extraire) [25] **2.** Conséquence tirée d'un raisonnement, conclusion. ◇ Raisonnement qui conclut, à partir de prémisses, d'hypothèses, à la vérité d'une proposition en usant de règles d'inférence.

Déterminisme n. m. (all. *determinismus*, lat. *deteminare*) (angl. *determinism*) [25] Conception philosophique selon laquelle il existe des rapports de cause à effet entre les phénomènes physiques, les actes humains, etc.

Diagnose n. f. (gr. *diagnôsis*, discernement) (angl. *diagnosis*) [28] « Connaissance qui s'acquiert par l'observation des signes diagnostiques »

Diagnostic n. m. (gr. *diagnôsis*, connaissance) (angl. *diagnosis*) [25] **1.** Identification d'une maladie par ses symptômes. **2.** Jugement porté sur une situation, sur un état. [163] Diagnostic is the discovery and identification of what is wrong with someone who is ill or with something that is not working properly.

Diagnostic n. m. (gr. *dia*, à travers, *gnômaï*, connaître) (angl. *diagnosis*) [28] Acte par lequel le médecin, groupant les symptômes morbides qu'offre son patient, les rattache à une maladie ayant sa place dans le cadre nosologique.

Diastole n. f. (gr. *diastellô*, je dilate) (angl. *diastole*) [28] Relâchement du cœur ou des artères au moment de l'afflux sanguin. Au niveau du cœur, la d. des oreillettes précède celle des ventricules. la d. des ventricules succède à leur systole au moment de la fermeture des valvules sigmoïdes aortiques et pulmonaires, elle débute par la relaxation isovolumétrique, se continue par les temps de remplissage rapide puis lent et se termine par la période correspondant à la systole suivante, marquée par la fermeture des valvules auriculo-ventriculaires. Elle correspond au grand silence.

Didactique adj. (du gr. *didaskein*, enseigner) [25] Qui a pour objet d'instruire ; pédagogique. – *Poésie didactique* Qui se propose l'exposé d'une doctrine philosophique ou de connaissances scientifiques ou techniques.

Didactique n. f. [25] Théorie et méthode de l'enseignement (d'une spécialité).

- Dioxyde de carbone** (angl. *carbon dioxide*) [28] syn. ancien gaz carbonique. Oxyde de carbone de formule CO_2 , gazeux dans les conditions ordinaires et inodore. Contrairement au monoxyde, le d. de c. n'est pas toxique et constitue le gaz des boissons dites « gazeuses » (eaux minérales, limonades, bières). Le d. de c. se forme lors de l'oxydation des composés organiques. C'est dire qu'il joue un rôle important dans de nombreuses réactions biologiques (respiration, synthèse chlorophyllienne, fermentation alcoolique, etc.).
- Efférent** adj. (lat. *ex*, hors de; *ferre*, porter) (angl. *effluent*) [28] Qui s'éloigne de; qui est issu d'un organe, centrifuge.
- Endogène** adj. (gr. *endon*, dedans, *genos*, origine) [25] **1.** DIDACT. Qui est produit par qqch en dehors de tout apport extérieur. CONTR. : *exogène*
- Épistémologie** n. f. (gr. *epistêmê*, science, et *logos* étude) [25] Partie de la philosophie qui étudie l'histoire, les méthodes, les principes des sciences.
- Ergodique** adj. (angl. *ergodic*) Pas dans le dico. (gr. *ergon*, action, travail) ?
- État** n. m. (lat. *status*) (angl. *state*) [25] **II. 1.** Manière d'être d'une chose à un moment donné. **2.** PHYS. CHIM. **a.** Manière d'être d'un corps relativement à sa cohésion, l'arrangement ou l'ionisation de ses atomes, etc. PHYS. *équation d'état* : relation existant à l'équilibre entre les grandeurs qui définissent l'état d'un corps pur. **b.** Ensemble des données caractéristiques d'un système thermodynamique ou cybernétique.
- Éthique** adj. (gr. *êthikos*, morale) [25] Qui concerne les principes de la morale.
- Éthique** n. f. [25] **1.** Partie de la philosophie qui étudie les fondements de la morale. **2.** Ensemble de règles de conduite **3.** *Éthique médicale* : *bioéthique*
- Éthologie** n. f. (gr. *êthos*, mœurs, et *logos*, science) [25] Étude scientifique du comportement des animaux dans leur milieu naturel.
- Êthos** n. m. (gr. *êthos*) [25] Caractère commun à un groupe d'individus appartenant à une même société.
- Étiologie** n. f. (gr. *aitia*, cause, et *logos*, science) [25] MÉD. Recherche des causes d'une maladie.
- Étiologie** n. f. (gr. *aîtia*, cause, *logos*, discours) (angl. *aetiology*) (amér. *etiology*) [28] Étude des causes des maladies. Celles-ci peuvent être d'origine traumatique, infectieuse, parasitaire, inflammatoire, immunologique, allergique, toxique, iatrogénique, carencielle, métabolique, génétique, tumorale, dégénérative, vasculaire... ou inconnue.
- Étymologie** n. f. (gr. *etumos*, vrai, *logos*, science) [25] **1.** Étude scientifique de l'origine des mots. **2.** Origine ou filiation d'un mot.
- Évènement** n. m. (du lat. *evenire*, arriver) (angl. *event*) [25] **1.** Ce qui se produit, arrive ou apparaît; fait, circonstance. **2.** fait important, marquant. **3.** STAT. éventualité qui se réalise dans un univers donné.
- Évident** adj. (lat. *evidens*) [25] **1** Qui s'impose à l'esprit, d'une certitude absolue; manifeste, indiscutable.
- Exogène** adj. (gr. *exô*, au-dehors, et *gennân*, engendrer) [25] **1.** DIDACT. Qui est provient du dehors, de l'extérieur, par opp. à *endogène* **2.** GÉOL. *Roche exogène*, formée à la surface de la terre.
- Explicite** (lat. *explicitus*) [25] **1.** Clair, qui ne prête à aucune contestation. *Réponse explicite*. **2.** DR. Énoncé formellement, complètement. *Clause explicite*.
- Forme** n.f. (lat. *forma*) [25] **I. 1.** Manière d'être extérieure, configuration des corps, des objets; aspect particulier. $\diamond$ *En forme de* : avec l'aspect de. – *Prendre forme* : commencer à avoir une apparence reconnaissable. **2.** Structure expressive, plastique de l'œuvre d'art. **II. 1.** Mode, modalité selon lesquels qqch d'abstrait se présente, peut exister. *Il y a plusieurs formes d'intelligence*. **2.** LING. Aspect sous lequel se présente un mot, une construction; unité linguistique (morphème, syntagme, etc.). *Forme interrogative, négative. Les formes du futur*. **3.** Manière dont une idée est présentée. *Juger sur la forme. Le fond et la forme*. **V. 1.** *Théorie de la forme* : gestaltisme **2.** MATH. Application associant à un, deux ou n vecteurs un élément du corps des scalaires de l'espace vectoriel. *Forme linéaire, quadratique, multilinéaire, etc.*
- Hémodynamique** adj. (gr. *haîma*, sang, *dunamis*, force) (angl. *haemodynamic*) [28] Qui se rapporte aux conditions mécaniques de la circulation du sang : pression, débit, vitesse, vasomotricité, résistance vasculaire etc. - n. f. (angl. *haemodynamics*) Étude des lois qui règlent l'écoulement et le débit du sang dans les vaisseaux.

- Hémorragie** n. f. (gr. *haïma*, sang, *rhêgnumi*, je jaillis) (angl. *haemorrhage*) (amér. *hemorrhage*) [28] Effusion d'une quantité plus ou moins considérable de sang hors d'un vaisseau sanguin.
- Heuristique** adj. (du gr. *heuriskein*, trouver) [25] DIDACT. Qui a une utilité dans la recherche, notamm. la recherche scientifique et épistémologique; qui aide à la découverte. *Hypothèse heuristique*.
- Heuristique** n. f. [25] Discipline qui se propose de dégager les règles de la recherche scientifique et de la découverte.
- Homéostasie** n. f. (W. B. Cannon, 1929-32) (gr. *homoïos*, semblable, *stasis*, position) (angl. *homeostasis*) [28] Équilibre du milieu intérieur de l'individu.
- Homéostasie** n. f. (W. B. Cannon, 1929-32) (gr. *homoïos*, semblable, *stasis*, position) (angl. *homeostasis*) [28] Maintien à leur valeur normale des différentes constantes physiologiques de l'individu (température, tonus cardiovasculaire, composition du sang etc.) l'*h.* est réglée par le système nerveux végétatif et les glandes endocrines.
- Hypoxémie** (gr. *hupo*, sous, *oxus*, oxygène, *haïma*, sang) (angl. *hypoxemia*) [28] V. *anoxémie*.
- Hypoxie** (gr. *hupo*, sous, *oxus*, oxygène) (angl. *hypoxia*) [28] V. *anoxie*.
- Idiosyncrasie** n. f. (gr. *idios*, particulier, et *sugkrasis*, mélange) [25] **Didact.** Manière d'être particulière à chaque individu qui l'amène à avoir des réactions, des comportements qui lui sont propres.
- Implicite** (lat. *implicitus*) [25] qui est contenu dans une proposition sans être exprimé en terme précis, formels; qui est la conséquence nécessaire. *Clause, condition, volonté implicite*.
- Individu** n. m. (lat. *individuum*, ce qui est indivisible) (angl. *individual*) [25] **1.** Chaque spécimen vivant d'une espèce animale ou végétale, issu d'une cellule unique. *Le genre, l'espèce et l'individu* **2.** Être humain, personne, par opp. à la collectivité, à la société.
- Induction** n. f. (lat. *inductio*) [25] **I. 1.** Généralisation d'une observation ou d'un raisonnement établis à partir de cas singulier. **2.** MATH. Raisonnement par récurrence. **II.1** *Induction magnétique* : vecteur caractérisant la densité du flux magnétique qui traverse une substance. **2.** EMBRYOL. Processus qui commande la différenciation des cellules de l'embryon et contrôle la constitution de celui-ci.
- Inférence** n.f. **1.** LOG. [25] Opération intellectuelle par laquelle on passe d'une vérité à une autre vérité, jugée telle en raison de son lien avec la première. *La déduction est une inférence*. $\diamond$ *Règles d'inférence*, celles qui permettent, dans une théorie déductive, de conclure à la vérité d'une proposition à partir d'une ou plusieurs propositions, prises comme hypothèses. **2.** INFORM. *Moteur d'inférence* : programme qui, dans un système expert, interprète les données de la base de connaissances et assure, suivant des stratégies générales ou particulières, l'enchaînement des étapes de la résolution d'un problème donné.
- Inférer** v. t. (lat. *inferre*, alléguer) **Litt.** [25] Tirer comme conséquence d'un fait, d'un principe.
- Instance** n. f. (lat. *instantia*, de *instare*, presser vivement) [25] **1.** (au pl) Prière, demande pressante sollicitation. *céder aux instances de qqn.* **2.** Avec *instance* : avec insistance **3.** DR. série des actes d'une procédure depuis la demande en justice jusqu'au jugement. *introduire une instance* **4.** Organisme, service qui exerce le pouvoir de décision. *les instances du parti.* **5.** PSYCHAN. Chacune des structures de l'appareil psychique (le ça, le moi, le surmoi). SYN. : *système*.
- Instant** n. m. (lat. *instans*) [25] Moment très court.
- Métabolisme** n. m. (gr. *metabolê*, changement) [25] PHYSIOL. Ensemble des processus complexes et incessants de transformation de matière et d'énergie par la cellule ou l'organisme, au cours des phénomènes d'édification et de dégradation organiques (anabolisme et catabolisme).
- Moment** (lat. *momentum*) [25] **I.1.** Espace de temps considéré dans sa durée plus ou moins brève. **2.** Espace de temps considéré du point de vue de son contenu, des événements qui s'y situent. **II. 1.** PHYS. **a.** *Moment cinétique* : vecteur égal au moment du vecteur quantité de mouvement. **b** *Moment d'un couple de forces* produit de l'une des forces du couple par le bras de levier de ce couple. **c.** *Moment électrique, magnétique d'un dipôle*, produit de la charge (électrique, magnétique) d'un des 2 pôles par la distance qui les sépare. **d.** *Moment d'une force par rapport à un point* : vecteur égal au moment du vecteur qui représente la force. **2.** ALG. Moment d'un bipoint (A, B) par rapport à un point O de l'espace : produit vectoriel de $\overrightarrow{OA}$ et $\overrightarrow{OB}$. (Ce vecteur ne varie pas si on fait glisser le vecteur $\overrightarrow{AB}$ sur la droite A et B .) **3.** STAT. *Moment (d'ordre n) d'une variable statistique* : moyenne des puissances $n^{ièmes}$ de ces valeurs pondérées par les effectifs de leurs classes respectives.

Nosologie n. f. (gr. *nosos*, maladie, *graphein*, écrire) (angl. *nosology*) [28] Étude des caractères distinctifs qui permettent de définir les maladies et de les classer.

Oligo-anurie n. f. (gr. oligo, peu) [28] Diminution extrême de la diurèse confinant à sa suppression.

Oxygène (gr. *oxus*, acide, *gennan*, engendrer) (angl. *oxygen*) (Lavoisier, 1783) [28] **1.** Élément chimique de numéro atomique 8 (huit électrons entourant le noyau de l'atome). Symbole O. L'oxygène est un des éléments les plus abondants dans la nature (oxydes, eau, air). Extrêmement réactif, il se combine à peu près à tous les éléments chimiques. L'élément o. est le constituant des corps purs *oxygène* (O_2) et *ozone* (O_3) **2.** Corps pur de formule O_2 . Gazeux dans les conditions ordinaires, l'o. est un des constituants principaux de l'air. À ce titre, il joue un rôle capital dans les phénomènes liés à la respiration des êtres vivants.

Oxygène (Saturation du sang en) (angl. *oxygen saturation*) [28] Rapport de la contenance en oxygène du sang à sa capacité en oxygène. Le sang artériel est normalement saturé en O_2 (So_2) à 97%, et le sang veineux mêlé (SvO_2) à 73%. La saturation en oxygène du sang artériel mesure la valeur de la fonction pulmonaire d'oxygénation du sang.

Oxylogie (gr. *oxus*, aigu, *logos*, discours) (angl. *emergency medicine*) [28] Branche de la médecine qui traite des secours d'urgence. V. *SAMU et réanimation*.

Oxymétrie (angl. *oximetry*) [28] Dosage de la quantité d'oxygène contenue dans un gaz ou un liquide (p. ex sang). Il peut être effectué par une méthode chimique (dosage volumétrique) ou par procédé physique (mesure spectrophotométrique). Par ce dernier procédé, fondé sur la différence des spectres d'absorption de l'oxyhémoglobine et de l'hémoglobine réduite, le taux d'oxygène peut être apprécié sans prise de sang, à travers la peau. V. *capnimétrie*.

Oxylogie (Gabor, 1970) [28] Terme mal formé mais consacré par l'usage. On devrait dire *oxylogie* (v. ce terme).

Paramètre n. m. (du lat. *para*, voisin de, *metron*, mesure) (angl. *parameter*) [25] : **1.** MATH. Élément autre que la variable ou l'inconnue, désignant un coefficient en fonction duquel on peut exprimer une proposition ou les solutions d'un problème. **2.** STAT. Grandeur mesurable permettant de présenter de façon plus simple, les caractéristiques principales d'un ensemble statistique. **3.** INFORM. Variable dont la valeur, l'adresse ou le nom ne sont précisés qu'à l'exécution du programme. **4.** Élément important à prendre en compte pour évaluer une situation, comprendre un phénomène dans le détail.

Partition (lat. *partitio*, partage) (angl. *cluster*) [25] **1.** Division, Séparation. **3.** MATH. Partition d'un ensemble : famille de parties non vides de cet ensemble, deux à deux disjointes et dont la réunion est égale à l'ensemble.

Pathognomie (gr. *pathos*, souffrance, *gnômon*, signe indicateur) (angl. **pathognomy**) [28] Étude des signes caractéristiques d'une maladie.

Pathognomonique (angl. *pathognomonic*) [28] Syn. désuet *diacritique* Qui est spécifique d'une maladie. **signe pathognomonique** Signe qui ne se rencontrant que dans une maladie bien définie et suffisant à lui seul à la caractériser et à poser le diagnostic. P. ex. le roulement diastolique est p. du rétrécissement mitral.

Périodique adj. (lat. *periodicus*, du gr.) (angl. *periodic*) [25] **1.** Qui revient, qui se reproduit à intervalles fixes. *une publication périodique*. **2.** CHIM. *classification périodique des éléments* : tableau des éléments d'après l'ordre croissant de leurs numéros atomiques, qui groupe par colonnes les éléments présentant des propriétés réactionnelles voisines. **3.** MATH **a.** *Fonction périodique*, qui reprend la même valeur lorsque la variable subit un accroissement égal à un multiple quelconque d'une quantité fixe (dite *périodique*). **b.** *Fraction périodique*, dont le développement décimal comporte une période, une tranche de chiffres qui se répètent indéfiniment.

Phase n. f. (gr. *phasis*) [25] **1.** Chacun des changements, des aspects successifs d'un phénomène en évolution ; chacun des intervalles de temps marqués par ces changements. **3.** CHIM. Partie homogène d'un système. *L'eau et la glace sont deux phases d'un même corps pur*. **4.** PHYS. **a.** Constante angulaire d'un mouvement vibratoire. **b.** Quantité $\omega t + \phi$ (ω étant la pulsation, t le temps, ϕ la phase à l'origine) dont le cosinus représente la variation d'une grandeur sinusoïdale. - *Phénomènes périodiques en phase* : phénomènes périodiques de même fréquence qui varient de la même façon et qui présentent des maximums et des minimums simultanés. ♦ **Fig.** *Être en phase avec qqn, qqch* : être en accord, en harmonie avec eux.

- Phénomène** n. m. (gr. *phainomenon*, ce qui apparaît) [25] **1.** Fait observable, événement. *Chercher les causes d'un phénomène.* ◇ PHILOS. Pour Kant, ce qui est perçu par les sens, ce qui apparaît et se manifeste à la conscience (par opp. à *noumène*). **2.** Fait, événement qui frappe par sa nouveauté ou son caractère exceptionnel.
- Phlébotomie** (gr. *phleps*, veine *tomê*, section) (angl. *phlebotomy*, *venesection*) [28] Incision d'une veine pratiquée pour extraire un caillot, introduire un cathéter ou évacuer une certaine quantité de sang (saignée veineuse).
- Polytraumatisme** [25] MÉD. Se dit d'un blessé qui présente simultanément plusieurs lésions traumatiques.
- Pression Artérielle** (angl. *blood pressure*) (syn. tension artérielle) [28] Force élastique exercée par les parois artérielles sur leur contenu sanguin. Elle s'équilibre, en pratique, avec la force contractile du cœur transmise par le sang (pression artérielle); et les termes de tension artérielle et de pression artérielle, bien que correspondant à des notions physiques différentes, sont en clinique, devenus synonymes.
- Pression Diastolique** (angl. *diastolic pressure*) (syn. tension minima) [28] Valeur de la pression qui existe dans les artères au moment de la diastole, c'est-à-dire entre deux contractions cardiaques. Elle est à l'état normal de 7 à 9 *cm* de mercure.
- Pression Moyenne** (angl. *mean pressure*) (syn. tension moyenne) [28] Terme désignant « la pression constante qui assurerait le même débit dans les vaisseaux que la pression variable qui y règne » (Marey). En clinique, la *t.m.* correspond à la plus grande oscillation observée sur le cadran oscillométrique du sphygmomanomètre. Normalement, elle est de 11 à 9 *cm* de mercure.
- Pression Systolique** (angl. *systolic pressure*) (syn. tension maxima) [28] Valeur de la pression existant dans le système artériel au moment même de la systole cardiaque. Elle est à l'état normal de 12.5 *cm* à 14 *cm* de mercure.
- Pronostic** (gr. *prognôstikein*, connaître d'avance) (angl. *prognosis*) [25] **1.** Prévision, supposition sur ce qui doit arriver. **2.** MÉD. Jugement porté à l'avance sur l'évolution d'une maladie.
- Science** à votre avis;)
- Sémiologie** ou (à tort) Séméiologie (gr. *sêméion*, signe, *logos*, discours) (angl. *semeiology*) [28] Syn. *sémiotique* Partie de la médecine qui étudie les signes des maladies.
- Signal** n. m. (lat. de *signalis*, de *signum*, signe) (angl. *signal*) [25] **1.** Signe convenu pour avertir, donner un ordre. **4.** CYBERN., INFORM. et TÉLÉCOMM. Variation d'une grandeur physique de nature quelconque porteuse d'information.
- Signe** n. m. (lat. *signum*) [25] **1.** Ce qui permet de connaître, de deviner, de prévoir; indice, marque.
- Stabilité** n. f. **1.** [25] Caractère de ce qui est stable, de ce qui tend à conserver son équilibre. *Vérifier la stabilité d'un pont.* ◇ MÉCA. Propriété qu'a un système dynamique de revenir à son régime établi après en avoir été écarté par une perturbation. - PHYS. Propriété d'un système en équilibre stable. **2.** Caractère de ce qui se maintient durablement sans profondes variations. *Stabilité de la monnaie, du pouvoir.*
- Stable** adj. (lat. *stabilis* de *stare*, se tenir debout) [25] **1.** Qui est dans un état, une situation ferme, une situation ferme, solide, qui ne risque pas de tomber. *Édifice stable.* ◇ SPÉCIALT. Qui a une bonne position d'équilibre. *Bateau, voiture stable* **2.** Qui se maintient, reste dans le même état; durable, permanent. *Situation stable.* **3.** dont la conduite est marquée par la constance, la permanence; équilibré. *Garçon stable. Humeur stable* **4.** CHIM. *Composé stable* Qui résiste à la décomposition. **5.** MÉCA. *Équilibre stable*, qui n'est pas détruit par une faible variation des conditions. **6.** MATH. *Partie stable d'un ensemble (muni d'une opération)*, telle que tout couple d'éléments de cette partie a son composé appartenant à cette même partie.
- Stationnaire** adj. (lat. *stare*, se tenir debout) [25] **1.** Qui ne subit aucune évolution, reste dans le même état. *L'état du malade est stationnaire.* **2.** Qui conserve la même valeur ou les mêmes propriétés. **3.** MATH. *Suite stationnaire*: suite (a_n) telle qu'il existe un nombre naturel p tel que $a_n = a_p$ pour tout $n > p$. **4.** PHYS. *Ondes stationnaires*, dans lesquelles les phénomènes d'oscillation sont, en tout point, soit en concordance, soit en opposition de phase.
- Statique** adj. (gr. *statikos*) [25] **1.** Qui demeure au même point, qui est sans mouvement, par opp. à *dynamique*. **2.** PHYS Qui a rapport à l'équilibre des forces.

Statique n. f. [25] Branche de la mécanique qui a pour objet l'équilibre des systèmes de forces. $\diamond$ *Statique des gaz* : aérostatique.

Statistique (all. *Statistik*, du lat. *status*, état) [25] **1.** Ensemble de méthodes mathématiques qui, à partir du recueil et de l'analyse de données réelles, permettent l'élaboration de modèles probabilistes autorisant les prévisions. **2.** Tableau numérique d'un phénomène se prêtant à l'analyse statistique. *Statistique de la natalité.*

Stochastique adj. (gr. *stokhastikos* de *stokhazein*, viser) (angl. *stochastic*) [25] Qui est de nature aléatoire. *Processus Stochastique.*

Stochastique n. f. [25] Calcul des probabilités appliqué à l'analyse des données statistiques.

Système (gr. *sustêma*, ensemble) (angl. *system*) [25] **I. 1.** Ensemble ordonné d'idées scientifiques ou philosophiques. **II. 1.** Combinaison d'éléments réunis de manière à former un ensemble. **II. 2.** Ensemble d'organes ou de tissus de même nature et destinés à des fonctions analogues. **II. 3.** Ensemble de termes définis par les relations qu'ils entretiennent entre eux. *Théorie des systèmes* : théorie générale et interdisciplinaire qui étudie les systèmes en tant qu'ensembles d'éléments, matériels ou non, en relation les uns avec les autres et formant un tout. **III. 1.** Ensemble de méthodes, de procédés destinés à assurer une fonction définie ou à produire un résultat. *Système expert* : programme élaboré pour résoudre des problèmes spécifiques en exploitant les connaissances accumulées dans un domaine spécialisé et en canalisant la recherche des solutions.

Sympathique (gr. *sun*, avec, *pathê*, affection) (angl. *sympathetic*) [28], **1.** Se dit des retentissements des troubles morbides d'un organe sur un ou plusieurs autres organes avec ou sans lésion de ceux-ci. **2.** Qui se rapporte au système nerveux sympathique. *récepteur s. V. récepteurs adrénergiques ou sympathique.* s. m. Système nerveux sympathique. *V. ce terme.*

Syncope (gr. *sun*, avec, *koptein*, couper) [28] « Perte de connaissance brutale et complète liée à une soudaine anorexie cérébrale » (J. Hamburger). Elle s'accompagne de pâleur extrême et généralement d'arrêt respiratoire. Réversible et de brève durée, elle est provoquée par une pause cardiaque, une bradycardie ou une tachycardie excessives ou bien par une subite hypotension artérielle. *V. lypothymie.* – On donne parfois le nom de *s. locale* à l'arrêt de la circulation dans une partie du corps bien limitée, tel qu'on l'observe dans la maladie de Raynaud.

Système Nerveux Autonome (angl. *autonomic nervous system*) [28] syn. système nerveux neurovégétatif, système végétatif, SNA. Ensemble des systèmes nerveux qui régissent le fonctionnement des viscères et entretiennent les fonctions vitales de base : respiration, circulation, digestion, excrétion ; le SNA est divisé en système sympathique et parasympathique. Il est complémentaire du système nerveux de la vie de relation.

Système Nerveux Central (angl. *central nervous system*) [28], Ensemble constitué par l'encéphale et la moelle épinière.

Système Nerveux Parasympathique (angl. *parasympathetic nervous system*) [28] L'un des constituants du système nerveux autonome. Il agit par l'intermédiaire d'un médiateur chimique, l'acétylcholine (v. ce terme). Le nerf vague (ou classiquement pneumogastrique, *X^e* paire crânienne) contient de nombreuses fibres parasympathiques.

Système Nerveux Sympathique (angl. *sympathetic nervous system*) [28] syn. système orthosympathique. L'un des constituants du système nerveux autonome. Il agit par l'intermédiaire de deux médiateurs chimiques, l'adrénaline et la noradrénaline (v. ces termes).

Systémique (angl. *systemic*) [28] **1.** Qui se rapporte à un système. – *maladie s. V. ce terme* **2.** CARDIOLOGIE Employé, à la suite des auteurs anglo-américains, dans le sens de : qui se rapporte à la grande circulation. – *cavités, ventricules s.* Cavités, ventricule du cœur qui reçoivent le sang des veines pulmonaires et l'envoient dans l'aorte : normalement, cavités et ventricules gauches.

Systole (gr. *sustolê*, resserrement) (angl. *systole*) [28] Contraction du muscle cardiaque. la *s.* simultanée des 2 oreillettes précède celle, également simultanée, des 2 ventricules. Le début de la *s.* des ventricules est marquée par la fermeture des valvules auriculo-ventriculaires ; sa fin par celle des valvules sigmoïdes aortiques et pulmonaires. La *s.* ventriculaire passe par deux phases, isométrique et isotonique (v. ces termes), elle correspond au petit silence. Elle provoque le premier bruit du cœur, le choc de pointe et la pulsation des artères.

Température n. f. (lat. *temperare*, adoucir) [25] **1.** Grandeur physique qui caractérise de façon objective la sensation subjective de chaleur ou de froid laissée par le contact d'un corps.

Temps n. m. (lat. *tempus*) [25] **I. 1.** Notion fondamentale conçue comme un milieu infini dans lequel se succèdent les événements et considérée souvent comme une force agissant sur le monde, les êtres.

Trauma [25] Traumatisme psychique.

Traumatisme (gr. *trauma*, blessure) (angl. *trauma*) [25] **1.** Ensemble des lésions locales intéressant les tissus et les organes, provoquées par un agent extérieur; troubles généraux qui en résultent. **2.** événement qui, pour un sujet, a une forte portée émotionnelle et qui entraîne chez lui des troubles psychiques ou somatiques par suite de son incapacité à y répondre adéquatement sur-le-champs.

Vague (lat. *vagus*, errant) (NA *nervus vagus*) (angl. *vagus nerve*) [28] syn. nerf pneumogastrique. Dixième paire crânienne; nerf très long, mixte, à destination cervicale, thoracique et abdominale, comportant un nombre importants de rameaux parasympathiques.

Variable adj. (lat. *varius*, varié) (angl. *variable*) [25] **1.** Qui varie, peut varier. *Humeur variable.* ◇ **gramm.** *Mot variable*, dont la forme varie selon le genre, le nombre, la fonction. **2.** Divers. *Résultats variables.*

Variable n. f. [25] MATH. terme indéterminé dont l'ensemble des valeurs possibles est déterminé.

Voie n. f. (lat. *via*) (angl. *channel*) [25] **4.** ANAT. Canal, organe, etc., permettant la circulation d'un liquide, d'un gaz, d'un influx nerveux; trajet suivi par ce liquide, ce gaz, etc.

Bibliographie

- [1] Guillaume Becq. Gestion des alarmes d'oxymétrie de pouls en unités de soins intensifs. Rapport de DEA, modèles et instruments en médecine et biologie, Lag, UJF, Grenoble, September 2000.
- [2] Gilles Zwingelstein. *Diagnostic des défaillances, théorie et pratique pour les systèmes industriels*. série Diagnostic et Maintenance. Hermès, Paris, 1995.
- [3] Jacky Montmain. *Interprétation qualitative de simulations pour le diagnostic en ligne de procédés continus*. Thèse d'automatique, Institut National Polytechnique de Grenoble, 1992.
- [4] Guillaume Becq, Sarah Mercier, Lionel Bourdon, Raymond Cespuglio, Alain Buguet, and Florian Chapotot. Analyse automatique du sommeil humain et animal à l'aide de réseaux de neurones configurables. In *XVIème Congrès de la Société Française de Recherche sur le Sommeil*, Nice, November 29–30, December 1, 2001. Poster.
- [5] Howard R. Champion, Ronald F. Bellamy, Colonel P. Roberts, and Ari Leppaniemi. A profile of combat injury. *The Journal of TRAUMA® Injury, Infection, and Critical Care*, 54(5) :S13–S17, May 2003.
- [6] Luc Quintin. Diagnostic précoce et secours primaire sur le champ de bataille. Technical report, Faculté de médecine de Lyon, Anesthésie cardiovasculaire, CHU Nord St Etienne et Physiologie, faculté de médecine 69373 Lyon cedex, 2000.
- [7] E. Tissot, L. J. Courbil, and R. Dorne. *Médecine en situation de catastrophes*, chapter Les techniques médicales du triage, pages 123–129. In *Service de Santé des Armées* [8], 1987.
- [8] L. J. Courbil. *Médecine en situation de catastrophe*. Service de santé des armées. Masson, Paris, 1987.
- [9] Frank H. Wilhelm, Walton T. Roth, and Marvin A. Sackner. The lifeshirt : An advance system for ambulatory measurement of respiratory and cardiac function. *Behavior Modification*, 27 :1–21, 2003.
- [10] A. Dittmar. Smart clothes and gloves for the study in real conditions : vital signs, emotional, sensorial, and comfort responses, vigilance and tasks. *special topic ITAB, Conference on Information Technology Applications in Biomedicine, 4th Annual IEEE-EMBS*, April 2003.
- [11] Nicola Taccini, Giannicola Loriga, André Dittmar, and Rita Paradiso. Knitted bioclothes for health monitoring. In *26th Annual International Conference*, San Francisco, California, September 1–5 2004. IEEE Engineering in Medicine and Biology Society (EMBS).
- [12] R. M. Ferro, Y. Merrien, M. Aubert, and F. Coursange. *Médecine en situation de catastrophe*, chapter Les polytraumatisés, pages 403–408. In *Service de santé des armées* [8], 1987.
- [13] Jean-Claude Otteni. *Le polytraumatisé : diagnostic, réanimation, chirurgie*, volume 6 of *Collection d'anesthésiologie et de réanimation*. Masson, Paris, 1986.
- [14] Jacques Albanèse. *Le Polytraumatisé*. Springer-Verlag, France, 2003.
- [15] R. Dorne, B. Palmier, A. Beaulaton, and P. Leroy. *Médecine en situation de catastrophes*, chapter Les États de choc, pages 154–160. In *Service de santé des armées* [8], 1987.
- [16] R. Vialet. *Aspects hémodynamique de la prise en charge du traumatisé*, pages 21–32. In [14], 2003.
- [17] Arthur C. Guyton. *Circulatory Shock and Physiology of Its Treatment*, chapter 24, pages 263–271. In Wonsiewicz [85], 8 edition, 1991.
- [18] James C. Schadt and John Ludbrook. Hemodynamic and neurohumoral responses to acute hypovolemia in conscious mammals. *American Journal of Physiology*, 260(Heart Circ. Physiol. 29) :H305–H318, 1991.

- [19] P. Plaisance and D. Payen. *Le choc hémorragique*, volume 22 of *Collection d'anesthésiologie et de réanimation*, chapter Physiopathologie du choc hémorragique, pages 1–16. Masson, Paris, 1991.
- [20] Daniel Provost. Anesthésie du microporc en recherche biomédicale. *Sciences et Techniques de l'Animal de Laboratoire*, 19 :29–38, 1994.
- [21] C. J. Dickinson. Fainting precipitated by collapse-firing of venous baroreceptors. *Lancet*, 342 :970–972, October 1993.
- [22] Stefan Rose and Ingo Marzi. Mediators in polytrauma - pathophysiological significance and clinical relevance. *Langenbeck's Arch Surg*, 383 :199–208, 1998.
- [23] Jean-Charles Sournia. *Histoire du diagnostic en médecine*. Éditions de Santé, Paris, 1995.
- [24] Robert Hegglin and Walter Siegenthaler. *Le diagnostic en médecine interne. Diagnostics positif, différentiel, étiologique*. Masson, Paris, 2 edition, 1982.
- [25] Larousse, editor. *Le Petit Larousse Illustré*. Larousse, Paris, 1996.
- [26] Antoine Cornuéjols and Laurent Miclet. *Apprentissage artificiel, Concepts et algorithmes*. Eyrolles, France, 2002.
- [27] Marius Fieschi. *Intelligence Artificielle en Médecine des systèmes experts*. Masson, Paris, 1984.
- [28] Marcel Garnier, Valéry Delamare, Jean Delamare, and Thérèse Delamare, editors. *Dictionnaire des Termes de Médecine*. Maloine, Paris, 25 edition, 1999.
- [29] Bernard Dubuisson. *Automatique et statistiques pour le diagnostic*. Hermès Science Publications, Paris, 2001.
- [30] Alexandre Evsukoff. *Le raisonnement approché pour la surveillance de procédé*. Thèse, automatique – productique, INPG, Grenoble, November 1998.
- [31] Ahmed Nassim Boudaoud and Mylène Masson. Diagnosis of transient states using a pattern recognition approach. *Journal Européen des Systèmes Automatisés*, 34(5) :689–708, 2000.
- [32] Ahmed Nassim Boudaoud. *Conception d'un système de diagnostic adaptatif en ligne pour la surveillance des systèmes évolutifs*. PhD thesis, Université de technologie de Compiègne, 1998.
- [33] Bernard Dubuisson. *Diagnostic et reconnaissance des formes*. série Diagnostic et Maintenance. Hermès, Paris, 1990.
- [34] Jean Louis Amat and Gérard Yahiaoui. *Techniques avancées pour le Traitement de l'Information. réseaux de neurones, logique floue, algorithmes génétiques*. Cépadués Editions, Toulouse, 2 edition, 2002.
- [35] Nada Lavrac, E. Keravnou, and B. Zupan. *Encyclopedia of computer science and technology*, volume 42, chapter Intelligent data analysis in medicine, pages 113–157. Dekker, New York, 2000.
- [36] Werner Horn. AI in medicine on its way from knowledge-intensive to data-intensive systems. *Artificial Intelligence in Medicine*, 23(1) :5–12, 2001.
- [37] E. Artioli, G. Avanzolini, P. Barbini, G. Cevenini, and G. Gnudi. Classification of postoperative cardiac patients : comparative evaluation of four algorithms. *Int J Biomed Comput*, 29 :257–270, 1991.
- [38] Chuan Lu, Tony Van-Gestel, Johan A.K. Suykens, Sabine Van-Huffel, and Dirk Timmerman. Classification of ovarian tumors using bayesian least squares support vector machines. In Michel Dojat, E. Keravnou, and P. Barahona, editors, *AIME03*, pages 219–228. Springer-Verlag, 2003.
- [39] Aleks Jakulin, Ivan Bratko, Dragica Smrke, Janez Demsar, and Blaz Zupan. Attribute interactions in medical data analysis. In Michel Dojat, E. Keravnou, and P. Barahona, editors, *AIME03*, pages 229–238. Springer-Verlag, 2003.
- [40] Basilio Sierra, Elena Lazkano, Inaki Inza, Marisa Merino, Pedro Larranaga, and Jorge Quiroga. Prototype selection and feature subset selection by estimation of distribution algorithms. a case study in the survival of cirrhotic patients treated with TIPS. In *AIME2001*, pages 20–29, 2001.
- [41] Katharina Morik, Michael Imhoff, Peter Brockhausen, Thorsten Joachims, and Ursula Gather. Knowledge discovery and knowledge validation in intensive care unit. *Artificial Intelligence in Medicine*, 19 :225–249, 2000.
- [42] Basilio Sierra, Nicolas Serrano, Pedro Larranaga, Eliseo J. Plasencia, Inaki Inza, Juan Jose Jimenes, and Jose Maria De la Rosa ans Maria Luisa Mora. Machine learning inspired approaches to combine standard medical measures at an intensive care unit. In *AIMDM99*, pages 366–371, 1999.

- [43] J. Benichou and Chastang C. Méthodologie des études pronostiques, élaboration de scores et classifications pronostiques. In J.R. Le Gall and Ph. Loirat, editors, *évaluation en réanimation*, volume 18 of *Collection d'anesthésiologie et de réanimation*. Masson, Paris, 1990.
- [44] M. Leone and F. Antonini. *Scores de Gravité*, pages 9–19. In [14], 2003.
- [45] B. Bedock. Indices de gravité généraux en réanimation. In J.R. Le Gall and Ph. Loirat, editors, *Évaluation en réanimation*, volume 18 of *Collection d'anesthésiologie et de réanimation*, pages 20–39. Masson, Paris, 1990.
- [46] Gérard Bleichner. *Guide des outils d'évaluation en réanimation*. Arnette Blackwell, Paris, 1995.
- [47] Jeremy W. Cannon. A mathematical model of hemorrhagic shock : The future of trauma triage. *Military Medicine*, 167(4) :312–316, April 2002.
- [48] David A. Roberts, John B. Holcomb, B. Eugene Parker, Jill L. Sondeen, Anthony E. Pusateri, William J. Brady, David E. Sweenor, and Jeffrey S. Young. The use of polynomial neural network for mortality prediction in uncontrolled venous and arterial hemorrhage. *The Journal of TRAUMA® Injury, Infection, and Critical Care*, 52 :130–135, January 2002.
- [49] Todd F. Glass, Jason Knapp, Philip Ambrun, Bruce A. Clay, Matt Kabrisky, Steven K. Rogers, and Victor F. Garcia. Use of artificial intelligence to identify cardiovascular compromise in a model of hemorrhagic shock. *Critical Care Medicine*, 32(2) :450–456, February 2004.
- [50] R. Barone. *Anatomie comparée des mammifères domestiques*. Vigot, Paris, 4 edition, 2000.
- [51] Jeannine A. Majde. Animal models for hemorrhage and resuscitation research. *The Journal of TRAUMA® Injury, Infection, and Critical Care*, 54 :100–105, May 2003.
- [52] Mp system hardware guide. Technical report, Biopac systems, inc, 2002.
- [53] Datex-engstrom as/3 anaesthesia monitor user's guide. Technical Reference Manual 891 309 - 1, Datex-Engstrom, April 1997.
- [54] Bourdon Lionel, Philippe d'Aleo, and Guillaume Becq. Dossier d'expérimentation animale soumis au comité d'éthique du CRSSA. Technical report, CRSSA, Grenoble, France, December 2001.
- [55] G. Hoffmann. *Les animaux de laboratoire*. Vigot Frères éditeurs, Paris, 1963.
- [56] Alain Faure. *Perception et Reconnaissance des Formes*. Éditests, Paris, 1985.
- [57] Jacques Max. *Méthodes et techniques de traitement du signal et applications aux mesures physiques*, volume 1 et 2. Masson, Paris, 4 edition, 1987.
- [58] Maurice Bellanger. *Traitement numérique du signal*. Masson, Paris, 4 edition, 1990.
- [59] Francis Cottet. *Aide-mémoire de traitement du signal*. Dunod, Paris, 1 edition, 2000.
- [60] Kari Nieminen, Richard M. Langford, C.J. Morgan, J. Takala, and Aarno Kari. A clinical description of the IMPROVE data library. *IEEE Engineering in Medicine and Biology Magazine*, 16(6) :21–24 40, NovemberDecember 1997.
- [61] Ilkka Korhonen, Jyrki Ojaniemi, Kari Nieminen, Mark van Gils, Arno Heikelä, and Aarno Kari. Building the IMPROVE data library. *IEEE Engineering in Medicine and Biology Magazine*, 16(6) :25–32, NovemberDecember 1997.
- [62] Bob Kemp, Alpo Värri, Agostinho C. Rosa, Kim D. Nielsen, and John Gade. A simple format for exchange of digitized polygraphic recordings. *Electroencephalography and Clinical Neurophysiology*, 82 :391–393, 1992.
- [63] Michèle Basseville and Igor Nikiforov. *Detection of Abrupt Changes - Theory and Application*. Prentice-Hall, April 1993.
- [64] Sylvie Charbonnier, Guillaume Becq, and Loïc Biot. On-line segmentation algorithm for continuously monitored data in intensive care units. *IEEE Transactions on Biomedical Engineering*, 51(3) :484–492, March 2004.
- [65] Jim Hunter and Neil McIntosh. Knowledge-based event detection in complex time series data. In *AIMDM99*, pages 271–280. Springer-Verlag, 1999.
- [66] Jürgen Fell, Alexander Kaplan, Boris Darkhovsky, and Joachim Röschke. EEG analysis with non linear deterministic and stochastic methods : a combined strategy. *Acta Neurobiologiae Experimentalis*, 60 :87–108, 2000.

- [67] Rajeev Agarwal and Jean Gotman. Computer-assisted sleep staging. *IEEE Transactions on Biomedical Engineering*, 48(12) :1412–1423, December 2001.
- [68] Daniel Calvelo, Marie-Christine Chambrin, Denis Pomorski, and Pierre Ravaux. ICU patient state characterization using machine learning in a time series framework. In *AIMDM99*, pages 356–360. Springer-Verlag, 1999.
- [69] Gilbert Saporta. *Probabilités Analyse de Données et Statistique*, volume 1. Éditions Technip, Paris, 1 edition, 1990.
- [70] Elias Vyzas and Rosalind W. Picard. Affective pattern classification. In *AAAI 1998 fall symposium, emotional and intelligent : The tangled knot of cognition.*, pages 176–182, 1998.
- [71] Elias Vyzas. *Recognition of emotional and cognitive states using physiological data*. PhD thesis, Massachusetts Institute of Technology, Cambridge, June 1999.
- [72] Patrick Karasinski, Luis Stinus, Claude Robert, and Aimé Limoge. Real-time sleep-wake scoring in the rat using a single EEG channel. *Sleep*, 17(2) :113–119, 1994.
- [73] Bo Hjorth. EEG analysis based on time domain properties. *Electroencephalography and Clinical Neurophysiology*, 29 :306–310, 1970.
- [74] Bo Hjort. The physical significance of time domain descriptors in EEG analysis. *Electroencephalography and Clinical Neurophysiology*, 34 :321–325, 1973.
- [75] Patrick Wahlberg. *On methods for pattern recognition with application to epileptic electroencephalograms*. Phd thesis, Department of Applied Electronics, Lund University, Sweden, 1999.
- [76] Randy Keith Avent and John D. Charlton. A critical review of trend-detection methodologies for biomedical monitoring systems. *Critical Reviews in Biomedical Engineering*, 17(6) :621–659, 1990.
- [77] Eamonn Keogh, Selina Chu, David Hart, and Michael Pazzani. An online algorithm for segmenting time series. In *IEEE International Conference on Data Mining*, pages 289–296. IEEE International Conference on Data Mining, 2001.
- [78] Eamonn Keogh, Stefano Lonardi, and Bill Chiu. Finding surprising patterns in a time series database in linear time and space. In *Proceedings of the 8th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 550–556, Edmonton, Alberta, Canada, 2002. ACM Press.
- [79] Jessica Lin, Eamonn Keogh, Stefano Lonardi, and Bill Chiu. A symbolic representation of time series, with implications for streaming algorithms. In *Proceedings of the 8th ACM SIGMOD workshop on Research issues in data mining and knowledge discovery*, pages 2–11. ACM Press, 2003.
- [80] Bill Chiu, Eamonn Keogh, and Stefano Lonardi. Probabilistic discovery of time series motifs. In *Proceedings of the 9th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 493–498, Washington, D.C., 2003. ACM Press.
- [81] Robin P. Smith, Jérôme Argod, Jean-Louis Pépin, and Patrick A Lévy. Pulse transit time : an appraisal of potential clinical applications. *THORAX*, 54(5) :452–457, May 1999.
- [82] American Academy of Sleep Medicine Task Force. Sleep-related breathing disorders in adults : recommendations for syndrome definition and measurement techniques in clinical research. *Sleep*, 22(5) :667–689, 1999.
- [83] Allan Rechtschaffen and Anthony Kales. *A Manual of Standardized Terminology, Techniques and Scoring System for Sleep Stages of Human Subjects*. US Government Printing Office, Washington, 1968.
- [84] Hans-Holger Ebert. *Guide de l'ECG*. Maloine, Paris, 1 edition, 2003.
- [85] Arthur C. Guyton. *Textbook of Medical Physiology*. W. B. Saunders Company, 8 edition, 1991.
- [86] Philippe Meyer, editor. *Physiologie humaine*. Flammarion, Paris, 2 edition, 1983.
- [87] Klaus Schröder, Karl Wegscheider, Uwe Zeymer, Ulrich Tebbe, and Rolf Schröder. Extent of ST-segment deviation in a single electrocardiogram lead 90 min after thrombolysis as a predictor of medium-term mortality in acute myocardial infarction. *Lancet*, 358 :1479–1486, November 2001.
- [88] Willem Einthoven. Über die form des menschlichen electrocardiogramms. *Archiv für die gesamte Physiologie des Menschen und der Tiere*, 60 :101–123, 1895.

- [89] Solange Akselrod, David Gordon, F. Andrew Ubel, Daniel C. Shannon, A. Clifford Barger, and Richard J. Cohen. Power spectrum analysis of heart rate fluctuation : a quantitative probe of beat-to-beat cardiovascular control. *Science*, 213 :220–222, July 1981.
- [90] Anna M. Bianchi, Luca Mainardi, Ettore Petrucci, Maria G. Signorini, Mauro Mainardi, and Sergio Cerutti. Time-variant power spectrum analysis for the detection of transient episodes in HRV signal. *IEEE Transactions on Biomedical Engineering*, 40(2) :136–144, February 1993.
- [91] Laurence Keselbrener and Solange Akselrod. Selective discrete fourier transform algorithm for time-frequency analysis : Method and application on simulated and cardiovascular signals. *IEEE Transactions on Biomedical Engineering*, 43(8) :789–802, August 1997.
- [92] Task force of the european society of cardiology, the north american society of pacing, and electrophysiology. Heart rate variability. *European Heart Journal*, 17 :354–381, 1996.
- [93] Gary M. Friesen, Thomas C. Jannett, Manal Afify Jadallah, Stanford L. Yates, Stephen R. Quint, and H. Troy Nagle. A comparison of the noise sensitivity of nine QRS detection algorithms. *IEEE Transactions on Biomedical Engineering*, 37(1) :85–98, January 1990.
- [94] Laurence Keselbrener, Michel Keselbrener, and Solange Akselrod. Nonlinear high pass filter for R-wave detection in ECG signal. *Med. Eng. Phys.*, 19(5) :481–484, 1997.
- [95] Anton Bartolo, Bradley D. Clymer, Richard C. Burgess, John P. Turnbull, Joseph A. Golish, and Michael C. Perry. An arrhythmia detector and heart rate estimator for overnight polysomnography studies. *IEEE Transactions on Biomedical Engineering*, 48(5) :513–521, May 2001.
- [96] H. A. N. Dinh, D. K. Kumar, N. D. Pah, and P. Burton. Wavelets for QRS detection. In *IEEE/EMBS 23rd Annual Conference*, October 25–28 2001.
- [97] Jiapu Pan and Willis J. Tompkins. A real-time QRS detection algorithm. *IEEE Transactions on Biomedical Engineering*, 32(3) :230–236, March 1985.
- [98] Michael Small, Dejin Yu, and Robert G. Harrison. Variation in the dominant period during ventricular fibrillation. *IEEE Transactions on Biomedical Engineering*, 48 :1056–1061, 2001.
- [99] Martin J. Tobin. State of the art. respiratory monitoring in the intensive care unit. *American Revue of Respiratory Disease*, 138 :1625–1642, 1988.
- [100] Pierre F. Baconnier, Pierre-Yves Carry, André Eberhard, Jean-Pierre Perdrix, and Jean-Marc Fargnoli. A computer program for automatic measurement of respiratory mechanics in artificially ventilated patients. *Computer Methods and Programs in Biomedicine*, 47(3) :205–220, August 1995.
- [101] André Eberhard, Pierre-Yves Carry, Jean-Pierre Perdrix, Jean-Marc Fargnoli, Loïc Biot, and Pierre F. Baconnier. A program based on a ‘selective’ least-squares method for respiratory mechanics monitoring in ventilated patients. *Computer Methods and Programs in Biomedicine*, 71(1) :39–61, May 2003.
- [102] James W. Davis, Ivan C. Davis, Lynn D. Bennick, John F. Bilello, Krista L. Kaups, and Steven N. Parks. Are automated blood pressure measurements accurate in trauma patients. *The Journal of TRAUMA® Injury, Infection, and Critical Care*, 55(5) :860–863, November 2003.
- [103] François Clergue and G. Orliaguet. *Monitoring non invasif en réanimation*, chapter Monitoring de l’oxygénation par oxymétrie de pouls, pages 53–67. Masson, Paris, 1992.
- [104] Loïc Biot, Laurent Holzapfel, Guillaume Becq, C. Mélot, P. Bouletreau, and Pierre Baconnier. Selective activation of blood pressure monitoring alarms : effect of noise pollution in the intensive care unit. *Annales Françaises d’Anesthésiologie et de Réanimation*, 20(8) :677–685, October 2001.
- [105] Loïc Biot, Laurent Hopzapfel, Guillaume Becq, C. Mélot, and Pierre Baconnier. Do we need a systematic activation of alarm soundings for blood pressure monitoring for the safety of icu patients? *Journal of Critical Care*, 18(4) :212–216, December 2003.
- [106] Edwin Diday, Jacques Lemaire, Jean Pouget, and Françoise Testu. *Éléments d’analyse de données*. Dunod, Paris, 1982.
- [107] Ludovic Lebart, Alain Morineau, and Marie Piron. *Statistique exploratoire multidimensionnelle*. Dunod, Paris, 3 edition, 2000.
- [108] Centre d’Enseignement et de Recherche de Statistique Appliquée. *Aide-Mémoire Pratique des Techniques Statistiques. Pour Ingénieurs et Techniciens Supérieurs*, volume 1. Lavoisier TecDoc, Paris, 4 edition, 1994.

- [109] Gilles Celeux and Jean-Pierre Nakache, editors. *Analyse Discriminante sur Variables Qualitatives*, volume 1. Polytechnica, Paris, 1 edition, 1994.
- [110] Keinosuke Fukunaga. *Introduction to Statistical Pattern Recognition*. Academic Press, San Diego, USA, 2 edition, 1990.
- [111] Theo Gasser, P. Bächer, and J. Möchs. Transformations towards the normal distribution of broad band spectral parameters of the EEG. *Electroencephalography and Clinical Neurophysiology*, pages 119–124, 1982.
- [112] Bernard Widrow and Michael A. Lehr. 30 years of adaptative neural networks : perceptron, madaline and backpropagation. *Proceeding of the IEEE*, 78(9) :1415–1442, September 1990.
- [113] Jeanny Hérault and Christian Jutten. *Les réseaux de neurones et le traitement du signal*. Hermès, 1994.
- [114] Sylvie Thiria, Yves Lechevallier, Olivier Gascuel, and Stéphane Canu, editors. *Statistique et méthodes neuronales*. Dunod, Paris, 1997.
- [115] Gérard Dreyfus, Jean-Marc Martinez, Manuel Samuelides, Mirta B. Gordon, Fouad Badran, Sylvie Thiria, and Laurent Hérault. *Réseaux de neurones. Méthodologie et applications*. Eyrolles, Paris, 1 edition, 2002.
- [116] John Ross Quinlan. *C4. 5 : Programs for Machine Learning*. Morgan Kaufman Publishers, San Mateo, California, USA, 1993.
- [117] Lukas Zoubek. Development of diagnostic system using decision trees. Master report in electrical engineering : Measurement and control apparatus in biomedicine, Lag, INPG, July 2003.
- [118] Grégory Brouillet. Diagnostic automatique à base de logique floue de systèmes médicaux. Application au diagnostic en ligne de l'état de gravité d'un blessé. Dea, automatique des procédés industriels, Lag, Ensieg, INPG, Grenoble, France, September 2003.
- [119] James M. Keller, Michael R. Gray, and James A. Givens. A fuzzy k-nearest neighbor algorithm. *IEEE Transactions on Systems, Man and Cybernetics.*, 15(4) :580–585, August 1985.
- [120] Ron Kohavi. *Wrappers for performance enhancement and oblivious decision graphs*. Doctor of philosophy, Department of computer science and the committee on graduate studies of stanford university, USA, September 1995.
- [121] Ron Kohavi. A study of cross-validation and bootstrap for accuracy estimation and model selection. In *IJCAI*, pages 1137–1145, 1995.
- [122] Ron Kohavi and George H. John. Wrappers for feature subset selection. *Artificial Intelligence*, 97(1–2) :273–324, 1997.
- [123] T G. Dietterich. Approximate statistical tests for comparing supervised classification learning algorithms. *Neural Computation*, 10(7) :1895–1924, 1998.
- [124] J.P. Hannon, C.A. Bossone, and C.E. Wade. Normal physiological values for conscious pigs used in biomedical research. *Lab Anim Sci*, 40 :293–298, 1990.
- [125] W.C. Shoemaker, A.B. Peitzman, R. Bellamy, et al. Resuscitation from severe hemorrhage. *Critical Care Medicine*, 24(Suppl 2) :S12–S23, 1996.
- [126] Guillaume Becq, Sylvie Charbonnier, Florian Chapotot, Alain Buguet, Lionel Bourdon, and Pierre Baconnier. Comparison between five classifiers for automatic scoring of human sleep recordings. In *Fuzzy Systems and Knowledge Discovery*, Singapore, November 18–22, 2002.
- [127] Lucia Martini, Milena Fini, Gianluca Giavaresi, Stefano Faenza, Flavia Petrini, and Roberto Giardino. Haemodynamic and volumetric monitoring during haemorrhagic shock in swine. *Resuscitation*, 51 :69–76, 2001.
- [128] J.P. Nakache and J. Confais. *Méthodes de classification avec illustration SPAD et SAS*. Cisia Ceresta, Montreuil, 1 edition, 2000.
- [129] Howard R. Champion, William J. Sacco, Anthony J. Carnazzo, Wayne Copes, and William J. Fouty. Trauma score. *Critical Care Medicine*, 9(9) :672–676, September 1981.
- [130] Howard R. Champion and William J. Sacco. A revision of the trauma score. *The Journal of TRAUMA® Injury, Infection, and Critical Care*, 29 :623–629, 1989.

- [131] Guillaume Becq, Sylvie Charbonnier, Florian Chapotot, Alain Buguet, Lionel Bourdon, and Pierre Baconnier. *Softcomputing in Knowledge Discovery Methods and Applications*, chapter Comparison Between Five Classifiers for Automatic Scoring of Human Sleep Recordings. Springer Verlag, à paraître.
- [132] Claude Robert, Christian Guilpin, and Aymé Limoge. Automated sleep staging systems in rats. *Journal of Neuroscience Methods*, 88 :111–122, 1999.
- [133] T. Penzel and R. Conradt. Computer based sleep recording and analysis. *Sleep Medicine Reviews*, 4(2) :131–148, 2000.
- [134] A. L. Loomis, E. N. Harvey, and G. Hobart. Cerebral stages during sleep, as studied by human brain potentials. *J Exp Psychol*, 21 :127–144, 1937.
- [135] A. L. Loomis, E. N. Harvey, and G. Hobart. Distribution of disturbance patterns in the human electroencephalogram, with special reference to sleep. *J. Neurophysiol.*, 1 :413–418, 1938.
- [136] Michel Jouvet, F. Michel, and J. Courjon. Sur un stade d'activité électrique cérébrale rapide au cours du sommeil physiologique. *C. R. Soc. Biol.*, 153 :1024–1028, May 1959.
- [137] E. Aserinsky and N. Kleitman. Regularly occurring periods of eye motility, and concomitant phenomena, during sleep. *Science*, 118 :273–274, 1957.
- [138] W. C. Dement and N. Kleitman. Cyclic variations in EEG during sleep and their relation to eye movements, body motility and dreaming. *Electroencephalography and Clinical Neurophysiology*, 9 :673–690, 1957.
- [139] W. C. Dement. The occurrence of low voltage fast electroencephalogram patterns during behavioral sleep in the cat. *Electroencephalography and Clinical Neurophysiology*, 10 :291–296, 1958.
- [140] J.R. Smith, M. Negin, and A. H. Nevis. Automatic analysis of sleep electroencephalograms by hybrid computation. *IEEE Transactions on Syst. Sci. Cybern.*, 5 :278–284, October 1969.
- [141] J. D. Frost. An automatic sleep analyser. *Electroencephalography and Clinical Neurophysiology*, 29 :88–92, 1970.
- [142] J.R. Smith and I. Karacan. EEG sleep stage scoring by an automatic hybrid system. *Electroencephalography and Clinical Neurophysiology*, 31 :231–237, 1971.
- [143] Max Hirshkowitz. Standing on the shoulders of giants : the *standardized sleep manual* after 30 years. *Sleep Medicine Reviews*, 4(2) :169–179, 2000.
- [144] F. Chapotot, R. Pigeau, F. Canini, L. Bourdon, and A. Buguet. Distinctive effects of modafinil and d-amphetamine on the homeostatic and circadian modulation of the human waking EEG. *Psychopharmacology*, 166 :127–138, 2003.
- [145] J. Mocks and T. Gasser. How to select epochs of the EEG at rest for quantitative analysis. *Electroencephalography and Clinical Neurophysiology*, pages 89–92, 1984.
- [146] B. Efron. Estimating the error rate of a prediction rule : improvement on crossvalidation. *Journal of the American Statistical Association*, 78(382) :316–330, 1983.
- [147] B. Efron and R. Tibshirani. Crossvalidation and the bootstrap : Estimating the error rate of a prediction rule. Technical Report 477, Statistics department, Stanford University, 1995.
- [148] Bernard Dubuisson. *Diagnostic, intelligence artificielle et reconnaissance de formes*. Hermès science Europe, Paris, 2001.
- [149] Claude Robert, Christian Guilpin, and Aymé Limoge. Comparison between conventional and neural network classifiers for rat sleep-wake stage discrimination. *Neuropsychobiology*, 35 :221–225, 1997.
- [150] Nicolas Schaltenbrand, Régis Lengelle, M. Toussaint, R. Luthringer, G. Carelli, A. Jacqmin, E. Lainey, Alain Muzet, and Jean-Paul Macher. Sleep stage scoring using the neural network model : comparison between visual and automatic analysis in normal subjects and patients. *Sleep*, pages 27–35, 1996.
- [151] Katharina Morik. The representation race – preprocessing for handling time phenomena. In *Machine Learning : ECML 2000, 11th European Conference on Machine Learning, Barcelona, Catalonia, Spain*, volume 1810, pages 4–19, Berlin, May 31, June 1, 2000. Springer-Verlag.
- [152] Jens Kohlmorgen, K.-R. Müller, J. Rittweger, and K. Pawelzik. Identification of nonstationary dynamics in physiological recordings. *Biological Cybernetics*, 83(1) :73–84, 2000.

- [153] Sari-Leena Himanen and Joel Hasan. Limitations of rechtschaffen and kales. *Sleep Medicine Reviews*, 4(2) :149–167, 2000.
- [154] Graham Teasdale and Bryan Jennett. Assessment of coma and impaired consciousness. a practical scale. *Lancet*, 2(13) :81–84, July 1974.
- [155] Susan P. Baker, Brian O’Neill, William Haddon, and William B. Long. The injury severity score : a method for describing patients with multiple injuries and evaluating emergency care. *The Journal of Trauma*, 14(3) :187–196, March 1974.
- [156] John J. Koehler, Lawrence J. Baer, Stuart A. Malafa, M. S. Meindertsma, Nancy R. Navitskas, and James E. Huizenga. Prehospital index : A scoring system for field triage of trauma victims. *Annals of Emergency Medicine*, 15(2) :178–182, February 1986.
- [157] John J. Koehler, Stuart A. Malafa, Jeffrey Hillesland, Lawrence J. Baer, Ralph N. Rogers, Nancy R. Navitskas, Deborah Briggs, Donna Simpson, Benedicte Roller, Polly Lilleboe, Terrence Morton, Steven A Meador, Martin Harris, Michael S Mendertsma, Mark Czarnecki, and Greg Bjerke. A multicenter validation of the prehospital index. *Annals of Emergency Medicine*, 16(4) :380–385, April 1987.
- [158] S.P. Gormican. CRAMS scale-field triage of trauma victims. *Annals of Emergency Medicine*, 11 :132–135, 1982.
- [159] Wayne S. Copes, Howard R. Champion, William J. Sacco, Mary M. Lawnick, Donald S. Gann, Thomas Gennarelli, Ellen MacKenzie, and Steven Schwartzberg. Progress in characterizing anatomic injury. *The Journal of Trauma*, 30(10) :1200–1207, October 1990. AP.
- [160] D. Lindsey. Teaching the initial management of major multiple system trauma. *The Journal of TRAUMA® Injury, Infection, and Critical Care*, 20(2) :160–162, February 1980.
- [161] Edward G. Davis, Ellen J. MacKenzie, William J. Sacco, Lawrence W. Bain, Robert F. Buckman, Howard R. Champion, and Peter S. J. Lees. A new “TRISS-Like” probability of survival model for intuated patients. *The Journal of TRAUMA® Injury, Infection, and Critical Care*, 55(1) :53–61, July 2003.
- [162] Carl R. Boyd, Mary Ann Tolson, and Wayne S. Copes. Evaluating trauma care : The TRISS method. *The Journal of Trauma*, 27(4) :370–378, 1987.
- [163] John Sinclair, editor. *Collins COBUILD*. HarperCollins, London, 1987.