
Rapport de Stage assistant ingénieur

Caractérisation des apnées du sommeil

AILLA Sana

22 Mai 2023 – 25 Août 2023

Encadrants de stage :

FORBES Florence (INRIA)

BAILLY Sebastien (HP2)

GANNAZ Irene (G-SCOP)

Tuteur de l'école :

GAUDOIN Olivier

Établissement d'enseignement : ENSIMAG - 2A MMIS

Structure d'accueil : INRIA

Inovallée 655 avenue de l'Europe

Montbonnot, FRANCE

Contents

1	Introduction	3
2	Contexte du stage	3
2.1	INRIA - CHU	3
3	Problématique	4
3.1	Présentation des jeux de données	4
3.1.1	La polysomnographie (PSG)	4
3.1.2	Les types des événements	5
3.1.3	Le jeu de données	5
3.2	Objectifs	6
3.3	État de l'art	7
4	Pré-traitement des données	7
4.1	Difficultés	8
4.2	Identification des événements	9
4.2.1	PhysioZoo et pobm	9
4.2.2	Code Guillemette et détection des hyperventilations	10
5	Analyse de données	12
5.1	Description	12
5.2	Modélisation	13
5.2.1	Modèle logistique	13
5.2.2	Sélection des Variables et Choix du Modèle	13
5.2.3	Qualité du modèle	14
5.2.4	SpaceM3	16
5.2.5	HDclassif	16
6	Bilan personnel	17
7	Conclusion et perspectives	18
A	Annexe	20

List of Figures

1	Patient dormant en polysomnographie [1].	5
2	Obstruction partielle ou totale du passage de l'air conduisant respectivement à des hypopnées ou à des apnées [2].	5
3	Exemple d'un signal de saturation d'oxygène de deux patients différents.	6
4	Exemple des annotations fournies (en bleu) pour des portions du signal, avec une durée de 30 secondes avant et après le label.	8
5	Indicateurs extraits du repérage des évènements de désaturation-resaturation.	10
6	Exemple du processus de ré annotation du signal	11
7	Représentation des données en 2D a l'aide de l'ACP.	12
8	BoxPlots de trois différentes variables en fonction des types d'événements.	13
9	Illustration du décalage entre la fin de l'apnée ou de l'hypopnée et la SaO2 la plus basse	22
10	Les différents channel des signaux issues de la PSG	22

1 Introduction

Le syndrome d'apnées-hypopnées du sommeil (SAHS) est un trouble respiratoire du sommeil qui se caractérise par des interruptions ou des diminutions des flux respiratoires de 10 à 30 secondes, voire plus, pendant le sommeil, pouvant se répéter plus de 10 fois par heure. Ces apnées peuvent avoir de sérieuses conséquences sur la santé. Elles altèrent considérablement la qualité de vie en provoquant une fatigue extrême en journée et exposent à un risque accru de maladies cardio-vasculaires et métaboliques.

Près d'un milliard de personnes dans le monde ont un syndrome d'apnées du sommeil, et à peine 20% sont diagnostiquées à ce jour. La difficulté principale est liée à la complexité du processus de diagnostic, qui nécessite la mise en œuvre d'un appareillage relativement compliqué et coûteux, la polysomnographie (PSG). Cet examen complet consiste à enregistrer de nombreux paramètres physiologiques pendant le sommeil, tels que les ondes cérébrales (EEG), les mouvements oculaires (EOG), ainsi que le signal de saturation sanguine en oxygène (SaO_2), qui est un signal crucial pour déterminer la sévérité et le pronostic de la maladie.

Mon stage d'assistant ingénieur s'est déroulé au sein de l'Institut National de Recherche en Sciences et Technologies du Numérique (INRIA) en collaboration avec le laboratoire INSERM Hypoxie et Physiopathologie du CHU de Grenoble Alpes. L'objectif était d'analyser les signaux de saturation d'oxygène (SaO_2) provenant de la polysomnographie, qui varient en fonction de l'origine de l'événement observé (central ou obstructif) ainsi que de la pathologie (hypoventilation alvéolaire). Mon travail consistait à mettre en place des indicateurs permettant d'identifier les patterns spécifiques du signal de saturation en oxygène en fonction du type d'événement observé. Pour cela, j'ai analysé les signaux SaO_2 ainsi que les annotations associées, afin de déterminer de nouveaux indicateurs liés aux événements. Ensuite, j'ai procédé à des tests avec différentes méthodes de classification supervisée et de clustering non supervisé pour déterminer si nous pouvons effectivement identifier des patterns spécifiques du signal de saturation en oxygène en fonction du type d'événement observé.

2 Contexte du stage

2.1 INRIA - CHU

L'INRIA (Institut National de Recherche en Sciences et Technologies du Numérique) est un établissement public à caractère scientifique et technologique français, spécialisé en mathématiques et informatique, placé sous la double tutelle du ministère de l'Éducation nationale, de la Recherche et de la Technologie, ainsi que du ministère de l'Économie, des Finances et de l'Industrie. Fondé en 1976 à Rocquencourt près de Paris, l'institut a depuis étendu ses activités et est désormais présent dans 10 centres

à travers la France, notamment à Paris, Saclay, Bordeaux, Rennes, Sophia Antipolis, Grenoble, Lyon, Montpellier, Lille et Rennes.

J'ai eu l'opportunité de réaliser mon stage au sein de L'INRIA Grenoble - Rhône-Alpes en collaboration avec le laboratoire INSERM Hypoxie et Physiopathologie du CHU de Grenoble Alpes. **(Parler un peu du projet et ou le HP2 .. !)**

Je faisais partie de l'équipe STATIFY créée en 2020 pour succéder à l'équipe MISTIS. IL s'agit d'une équipe conjointe de l'Inria et du Laboratoire Jean Kuntzmann qui s'intéresse au développement des méthodes statistiques mathématiquement fondées pour modéliser des systèmes complexes basés sur des données structurées. Ces méthodes visent à capturer la variabilité des systèmes étudiés, en particulier dans des domaines tels que l'imagerie cérébrale, la médecine personnalisée, l'analyse des risques environnementaux et les géosciences. Pour atteindre ces objectifs, l'équipe se concentre sur cinq défis majeurs liés aux données de haute dimension, multimodales et hétérogènes, à la science des données spatiales, aux modèles bayésiens évolutifs et aux problèmes de petites données. Les recherches de STATIFY sont organisées autour de quatre axes majeurs, notamment les modèles pour les graphes et les réseaux, la réduction de dimension et la modélisation de variables latentes, la modélisation bayésienne, et enfin, la modélisation et la quantification des risques extrêmes.

3 Problématique

3.1 Présentation des jeux de données

3.1.1 La polysomnographie (PSG)

Il s'agit d'un examen médical qui permet d'étudier la physiologie du sommeil et d'identifier différents troubles du sommeil grâce à l'enregistrement de plusieurs signaux de quantification des stades du Sommeil. La PSG (Figure 1) consiste en une nuit d'enregistrements de différents paramètres pendant une durée d'au moins 6 heures à l'hôpital ou au domicile du patient. Des électrodes placées sur la tête, le visage, la poitrine et les jambes permettent d'évaluer plusieurs paramètres neurophysiologiques et respiratoires tels que :

- la mesure du débit aérien oro-nasal,
- la mesure du CO₂ expiré,
- la saturation sanguine en oxygène SaO₂,
- la fréquence cardiaque (ECG).

L'ensemble de ces signaux permet de repérer des événements physiologiques symptomatiques tels que les différents stades de sommeil, les événements respiratoires (apnées et hypopnées) et les événements de désaturation en oxygène.

3.1.2 Les types des évènements

Le syndrome d'apnées-hypopnées du sommeil est caractérisé par des interruptions cycliques de la ventilation, totales ou partielles, pendant le sommeil. Nous allons détailler ici les définitions des différents événements respiratoires qui nous intéressent.

Apnée : arrêt de la respiration durant au moins 10 secondes pendant le sommeil (Figure 2).

Apnée obstructive : arrêt du flux d'air mais persistance des efforts ventilatoires pendant l'apnée.

Apnée centrale : arrêt du flux d'air et absence des efforts ventilatoires pendant l'apnée.

Hypopnée : diminution de l'amplitude de la respiration associée à une chute du taux d'oxygène (désaturation) ou un bref réveil (Figure 2).

Hyperventilation : respiration rapide et profonde qui peut laisser une sensation d'essoufflement.



Figure 1: Patient dormant en polysomnographie [1].

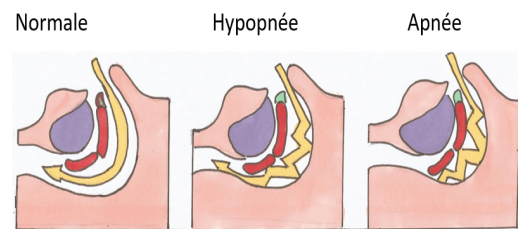


Figure 2: Obstruction partielle ou totale du passage de l'air conduisant respectivement à des hypopnées ou à des apnées [2].

3.1.3 Le jeu de données

Le jeu de données au début était sous forme d'un fichier CSV contenant l'enregistrement de la saturation SaO₂ en fonction du temps pour 4 patients différents pendant une nuit (Figure 3), ainsi que des annotations sur ce même signal indiquant les types d'événements, l'heure et la durée.

Par la suite, une partie de la base de données "Mars" a été fournie, comprenant un dossier pour chaque patient (un total de 10 dossiers). Chaque dossier contenait

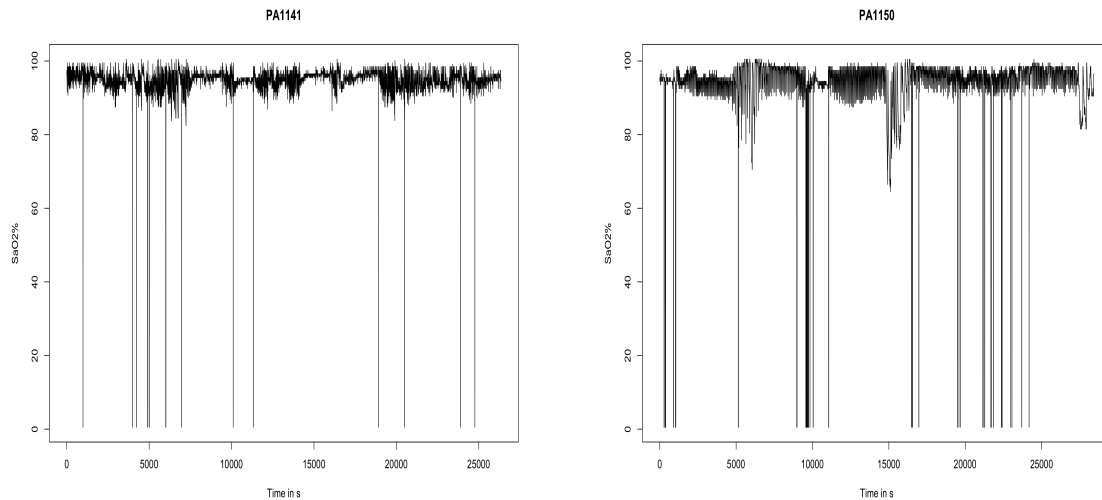


Figure 3: Exemple d'un signal de saturation d'oxygène de deux patients différents.

les signaux bruts issus directement de la polysomnographie, soit un fichier EDF contenant au total 26 channel pour 26 signaux différents (voir Annexe 10) tels que l'EEG¹, L'EMG², L'EOG³, L'ECG⁴ et d'autres. Parmi ces signaux, l'un d'eux était le signal de saturation d'oxygène SaO2.

3.2 Objectifs

Le stage avait pour objectif de caractériser les différents types d'événements mentionnés dans la section précédente (Section 3.1.2), à savoir les Apnées-Hypopnées centrales ou obstructives et de détecter les hyperventilations. Les étapes principales pour atteindre ces objectifs étaient les suivantes :

1. **Bibliographie.** Une première étape consistait à se documenter sur le sujet, notamment en se mettant en contexte concernant les apnées, en lisant des articles de recherche et en s'appropriant les connaissances nécessaires.
2. **Pré-traitement des signaux bruts.** Avant d'extraire les indicateurs, une analyse des fichiers fournis pour les 4 patients a été réalisée. Il était essentiel de visualiser les signaux et les annotations fournies pour mieux comprendre leur nature et leur apparence. La suppression des artefacts⁵ a également été effectuée dans cette étape.

¹L'électroencéphalogramme : Mesure l'activité électrique du cerveau

²L'électromyogramme : Enregistre l'activité électrique des muscles.

³L'électrooculogramme : Enregistre les mouvements oculaires.

⁴L'électrocardiogramme : Mesure l'activité électrique du cœur

⁵Artefacts : Des baisses soudaines au zéro du signal SaO2 qui peuvent être causées par des mouvements du patient, le détachement de l'oxymètre ou d'autres perturbations physiologiques.

3. **Détection et segmentation des événements.** À partir des annotations, l'objectif était de détecter et de segmenter les événements d'Apnées-Hypopnées centrales ou obstructives, ainsi que les hyperventilations.
4. **Extraction d'indicateurs.** Une fois les événements détectés, l'étape suivante consistait à extraire les indicateurs pertinents à partir des signaux bruts associés à chaque événement.
5. **Classification/Clustering des événements.** En utilisant les indicateurs extraits, l'objectif était de mettre en place un modèle de Classification/Clustering pour déterminer le type de chaque événement (Apnée-Hypopnée centrale, Apnée-Hypopnée obstructive) et les indicateurs qui caractérisent chaque groupe d'événements.
6. **Évaluation des résultats.** Des métriques appropriées, telles que la précision et la matrice de confusion, ont été utilisées pour évaluer la performance des modèles.
7. **Interprétation des résultats.** Enfin, il fallait interpréter les résultats obtenus pour comprendre la capacité du modèle à identifier les différents types d'événements et éventuellement indiquer des indicateurs spécifiques permettant une distinction claire entre les différents événements.

L'objectif ultime était de parvenir à une caractérisation précise des événements et à une détection efficace des types d'apnées et d'hypopnées à partir des signaux et des annotations disponibles dans la base de données "Mars".

3.3 État de l'art

L'état de l'art dans ce domaine trouve ses racines dans l'étude fondamentale menée par Levy et son équipe de collaborateurs, [Lev+21]. Cette étude approfondie explore les séries temporelles de saturation en oxygène (SaO₂) en corrélation avec les événements obstructifs tels que l'apnée obstructive du sommeil (OSA), dans le but de caractériser la gravité de chaque type d'événement (léger, modéré et sévère). Cet article met en évidence la sélection et l'extraction de plusieurs biomarqueurs issus de diverses catégories : les mesures statistiques générales (comme la moyenne, le passage par zéro, etc.), les mesures de complexité (telles que l'entropie qui quantifie l'irrégularité des signaux) ainsi que les mesures liées à la désaturation (comme l'indice de désaturation en oxygène, l'aire de désaturation, etc.). Plus de détails concernant la bibliothèque implémentée dans cette recherche sont présentés ici (voir Section 4.2.1).

4 Pré-traitement des données

Nous allons étudier les signaux de saturation en oxygène et les annotations fournies pour identifier les caractéristiques des différents types d'événements. Cependant, ces

annotations ne sont pas réalisées directement sur ce signal. Elles proviennent d'un autre signal.

4.1 Difficultés

En analysant les premières données fournies, nous avons observé que les annotations ne correspondaient pas aux désaturations, qui indiquent normalement les événements se produisant dans le signal (voir Figure 4).

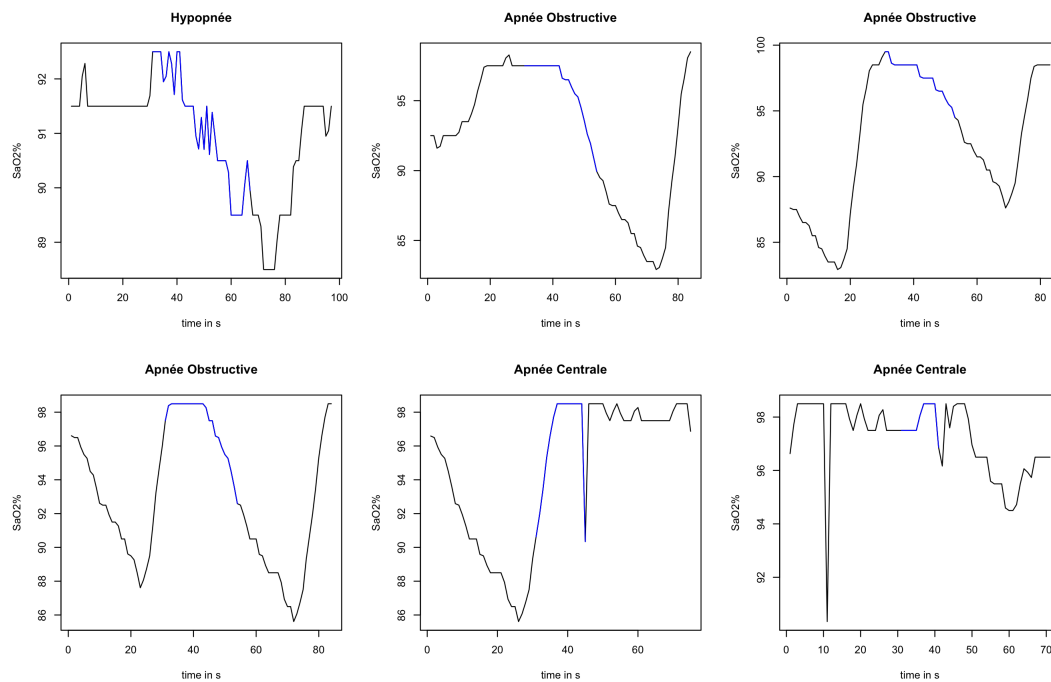


Figure 4: Exemple des annotations fournies (en bleu) pour des portions du signal, avec une durée de 30 secondes avant et après le label.

Comme précisé avant, les annotations ont été faites sur un autre signal (voir Annexe 9), ce qui a entraîné un décalage entre les annotations fournies et les vrais événements (sous forme de désaturations). On peut bien remarquer que les labels (en bleu dans la Figure 3) sont suivis des désaturations. J'ai d'abord essayé d'utiliser des décalages fixes pour chaque patient, mais les résultats n'étaient pas entièrement satisfaisants. Ainsi, les décalages fixes ne fonctionnaient pas pour tous les événements.

Par la suite, j'ai découvert que le temps entre la fin du label et le minimum de la désaturation varie selon le patient, en fonction de facteurs tels que la présence ou non d'insuffisances cardiaques, ainsi que le nombre d'événements spécifiques à chaque individu. En effet, la médiane du temps de circulation, connu sous le nom de Lung-to-Finger Circulation Time (LFCT) ⁶ (voir Figure 9), est de 18 secondes chez les patients sans insuffisance cardiaque (souvent associés à l'apnée obstructive du sommeil, OSA),

⁶LFCT : décalage entre la fin de l'apnée ou de l'hypopnée et la SaO2 la plus basse

et de 26 secondes chez les patients atteints d'insuffisance cardiaque (plus souvent liée à l'apnée centrale du sommeil, CSA). J'ai testé différentes valeurs pour les 4 patients en prenant en compte les intervalles de confiance indiqués dans l'article [Kwo+21], mais même visuellement, cela n'a pas toujours correspondu parfaitement. De plus, les données n'étaient pas suffisantes pour déterminer le bon décalage (la présence ou non de l'insuffisance cardiaque) pour chaque patient.

4.2 Identification des événements

Vu les annotations décalées, il était nécessaire de trouver une méthode pour détecter les événements qui se produisent. Plusieurs hypothèses concernant chaque type d'événement ont été faites concernant des indicateurs tels que la durée, la pente de restauration et de désaturation, etc. Cependant, ce n'étaient que des hypothèses, et nous ne pouvions pas déterminer avec certitude le type de chaque désaturation détectée. Nous aurons toujours besoin des annotations fournies comme référence pour une validation précise.

L'idée était alors de détecter les désaturations en associant à chaque événement la désaturation la plus proche en termes de temps, tout en prenant en compte un seuil comme décalage maximum entre le début des annotations et le début des désaturations associées (90 secondes au maximum). Cette approche nous a permis d'établir un lien entre les annotations réelles et les désaturations détectées, malgré les décalages des labels.

Avec cette méthode, nous avons pu mieux aligner les événements réels de désaturation avec les annotations fournies (voir Figure 6), et ainsi obtenir des résultats plus précis dans la détection des événements. Cependant, il est important de noter que la validation avec les annotations fournies reste essentielle pour confirmer la nature de chaque désaturation détectée et pour garantir la fiabilité des résultats obtenues par la suite.

Pour détecter les désaturations, deux codes ont été utilisés dans deux temps différents:

4.2.1 PhysioZoo et pobm

PhysioZoo est une plateforme collaborative implémentée à l'issue de l'article de Levy et al [Lev+21] dédiée à l'analyse continue des séries temporelles physiologiques. Elle propose des outils spécialisés pour diverses analyses, notamment la variabilité du rythme cardiaque (HRV) à partir d'enregistrements électrophysiologiques, les biomarqueurs numériques d'oxymétrie (SPO2) à partir de séries temporelles continues d'oxymétrie, et les biomarqueurs d'électrocardiographie (ECG). Dans notre cas, nous avons utilisé cette plateforme pour analyser des séries temporelles continues de SaO2. Cependant, le logiciel lui-même n'a pas fonctionné lors de l'installation, ce qui nous a obligé à utiliser directement le code sous forme d'une bibliothèque Python (POBM) pour détecter les

désaturations. Nous avons cherché à détecter les désaturations ainsi que les différents indicateurs caractéristiques de chaque désaturation. Cependant, pour l'utilisation des trois méthodes implémentées, il était nécessaire de régler des paramètres qui dépendent eux-mêmes des désaturations (par exemple L'ODI ⁷) ainsi que d'autres paramètres. En utilisant ces différentes méthodes (échangeant les différents paramètres), une grande majorité des désaturations détectées présentaient des pentes indéfinies et des allures incohérentes **Exemple !!** et des paramètres NaN (exemple des pentes de désaturation et resaturation). En éliminant ces données problématiques, seul un très petit nombre de données est resté exploitable.

4.2.2 Code Guillemette et détection des hyperventilations

Le code de Guillemette Boissier (Guillemette est une ancienne stagiaire PFE au sein du laboratoire HP2) a été développé pour traiter les données brutes de la polysomnographie, y compris le signal SaO₂. Elle a mis en place un code permettant la détection des désaturations ainsi que certains indicateurs liés aux pentes et à la durée des désaturations. En plus, le travail réalisé a été validé par les gens du CHU.

A partir des données brutes et à l'aide du code de Guillemette, on a filtré les signaux et supprimé les artefacts, puis détecté les désaturations avec les caractéristiques indiquées sur la Figure 5.

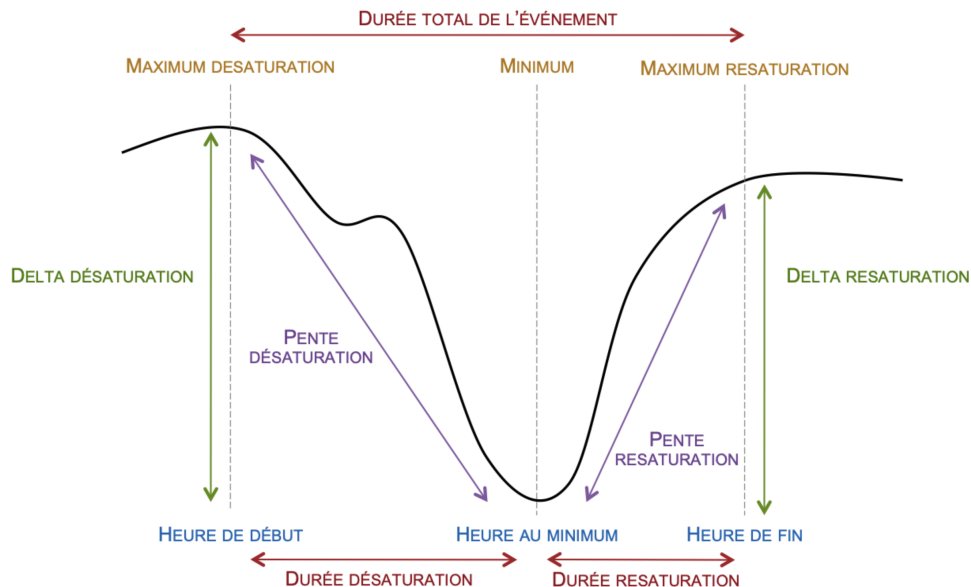


Figure 5: Indicateurs extraits du repérage des évènements de désaturation-resaturation.

⁷Oxygen Desaturation Index : le nombre de desaturation par heure

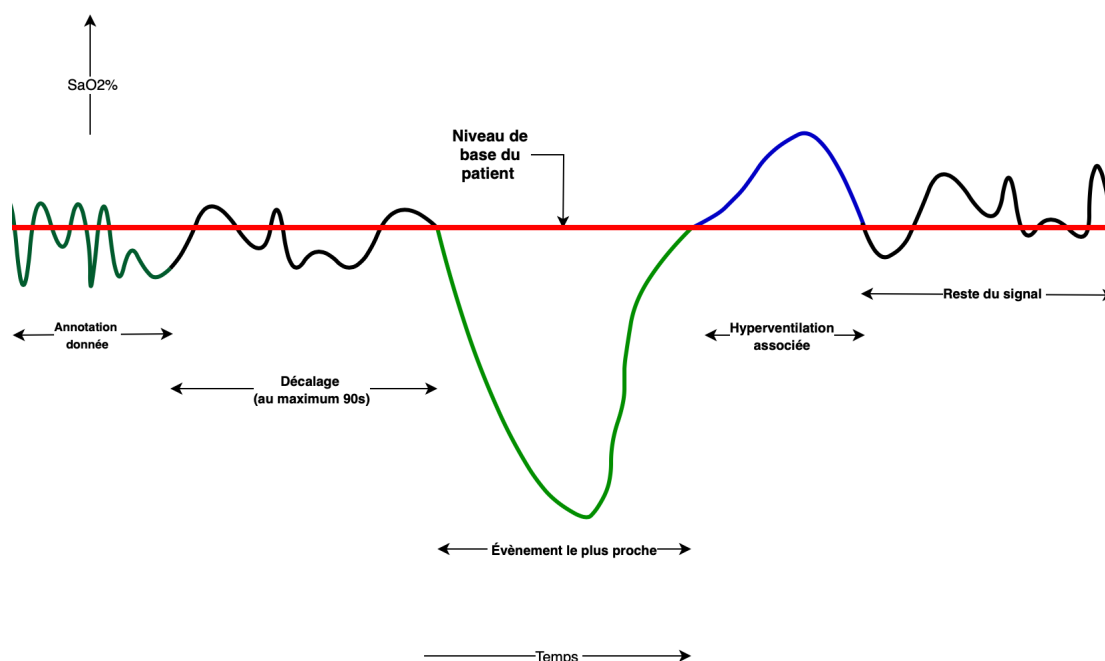


Figure 6: Exemple du processus de ré annotation du signal

Un niveau de base est utilisé comme référence pour chaque patient⁸. Ainsi, les désaturations conservées sont celles dont le minimum est forcément inférieur au niveau de base (sinon on ne les considère pas comme des désaturations). Par la suite, un premier indicateur est ajouté : il s'agit de la surface sous le niveau de base, calculée à l'aide de la méthode des trapèzes. Pour les événements d'hyperventilations (Section 3.1.2), une approche consiste à récupérer le signal SaO2 symétrique par rapport au niveau de base. Ainsi, les désaturations liées aux hyperventilations peuvent être détectées en reprenant le traitement déjà effectué pour récupérer les désaturations mais sur le signal symétrique. Cette approche nous a permis de détecter les remontées par rapport au niveau de base et l'ajout d'un autre indicateur, qui est l'aire sous la courbe du niveau de base avant le retour au niveau normal (voir Figure 6). Alors, en utilisant les désaturations détectées par le code de Guillemette avec les annotations données, on obtient les événements en attribuant à chaque désaturation l'annotation la plus proche (afin d'obtenir le même nombre d'événements fournis). De même, nous appliquons la même démarche entre les désaturations annotées et les hyperventilations associées, c'est-à-dire que chaque hyperventilation est associée à l'évènement le plus proche en terme de temps). D'autres indicateurs ont été ajoutés : DD100, DDbase, TimeRat, PenteRat (voir Annexe A). A ce stade, nous avons un jeu de données contenant les différents événements annotés ainsi qu'un ensemble d'indicateurs associés. (au total 16 indicateurs, voir Table 1).

⁸le niveau de base d'un patient est choisi étant la valeur de saturation d'oxygène à l'instant initial dans notre cas

5 Analyse de données

5.1 Description

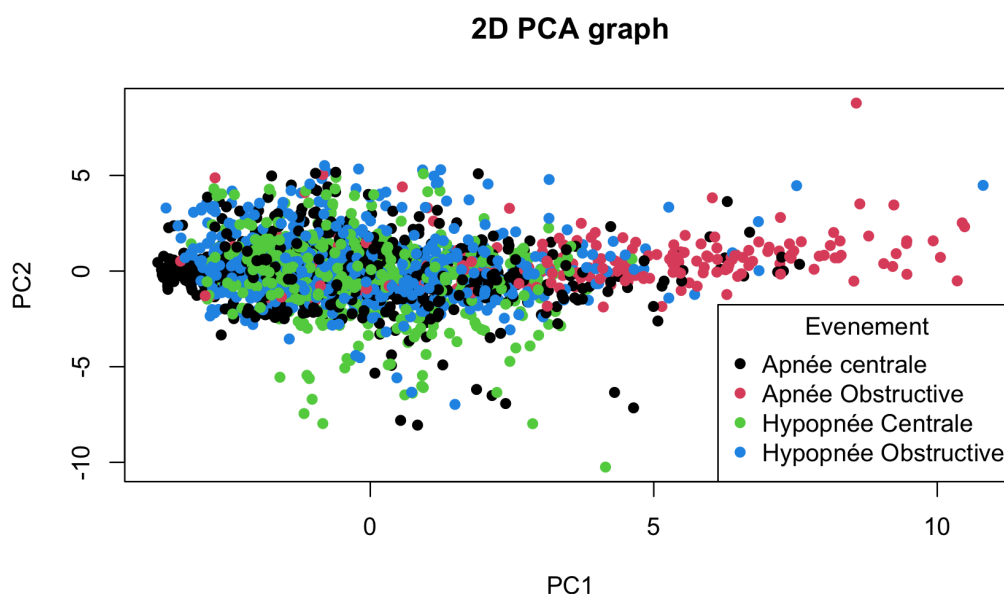


Figure 7: Représentation des données en 2D a l'aide de l'ACP.

Le graphique en 2D (Figure 7) généré à partir de l'analyse en composantes principales (PCA) offre une représentation visuelle de la réduction des 16 indicateurs à deux dimensions principales. Cette technique vise à capturer la variance maximale dans les données en projetant les points dans un nouvel espace plus petit. En observant le graphique résultant, chaque point représente un événement d'un patient, et sa position dans le graphique est déterminée par ses valeurs réduites en deux dimensions a l'aide de l'analyse de composante principale.

Visuellement, on peut dire que le graphique en 2D (Figure 7) ne révèle pas une distinction évidente entre les quatre types d'événements. Les données ne semblent pas être séparées facilement et linéairement dans cet espace réduit, soulignant ainsi la complexité et la variabilité des relations entre les indicateurs et les types d'événements.

Pour analyser en détail les caractéristiques de chaque événement, nous avons utilisé des boxplots. Dans la Figure 8, nous avons présenté des exemples de boxplots représentant les quatre événements distincts. Chaque boxplot inclut un indicateur spécifique (les autres sont donnés en annexe). L'observation attentive de ces boxplots permet d'analyser la capacité discriminatoire concernant des indicateurs pour différencier entre les types d'événements.

Les résultats présentés dans la Figure 8 démontrent clairement que les deux premiers

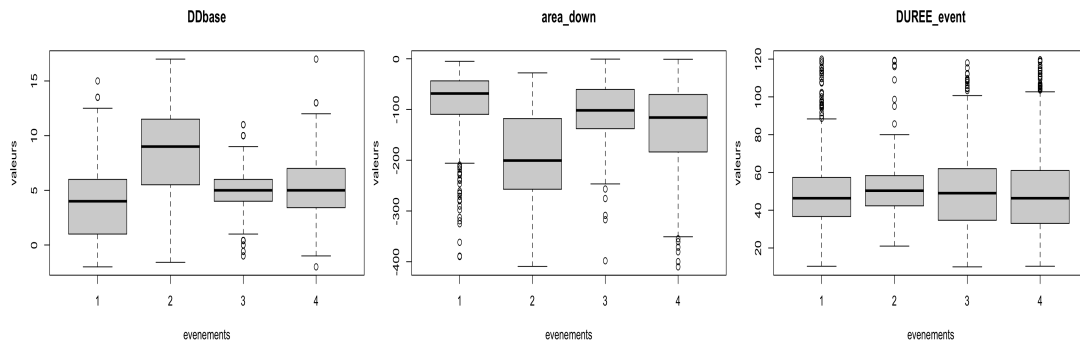


Figure 8: BoxPlots de trois différentes variables en fonction des types d'événements.

indicateurs (DDbase et area_down [A](#)) possèdent un pouvoir discriminatoire marqué entre les deux types d'apnées : centrale et obstructive. En effet, ces indicateurs varient de manière significative car les données affichent des médianes nettement décalées et des étendues interquartiles non superposées, ce qui suggère que ces deux indicateurs possèdent effectivement un pouvoir discriminatoire pour distinguer entre les deux types d'apnées, ainsi ils peuvent jouer un rôle crucial dans la distinction entre ces situations spécifiques. Cependant, aucune conclusion probante ne peut être tirée pour les événements d'hypopnées.

Par contre, l'analyse du troisième indicateur (DUREE_event [A](#)) ne montre pas de variations significatives entre les différentes catégories d'événements. Les valeurs médianes et les étendues interquartiles des boîtes dans le boxplot ne présentent pas de différences prononcées, ce qui suggère que cet indicateur ne contribue pas à la distinction entre les divers types d'événements.

5.2 Modélisation

5.2.1 Modèle logistique

L'objectif est de construire un modèle logistique pour prédire le type d'événement à partir de l'ensemble des indicateurs. Le modèle logistique multinomial est une approche appropriée pour modéliser des relations entre une variable catégorielle (le type d'événement) et des variables explicatives (les indicateurs). Nous utiliserons cette méthode pour déterminer quelles variables contribuent le plus à la prédiction du type d'événement.

5.2.2 Sélection des Variables et Choix du Modèle

Nous avons opté pour une approche basée sur le critère d'information d'Akaike (AIC) pour la sélection des variables. L'AIC est un outil statistique qui nous permettra de comparer différents modèles et de choisir celui qui décrit au mieux les données, tout en évitant la sur-complexité.

Les étapes de la sélection des variables sont :

1. Modèle 0 : Nous avons d'abord créé un modèle sans aucune variable explicative. Cela nous donne une base de référence pour évaluer l'amélioration apportée par les variables dans les modèles suivants.
2. Modèle avec toutes les variables : Ensuite, nous avons construit un modèle incluant toutes les variables disponibles dans notre jeu de données.
3. Comparaison des modèles : Nous avons comparé les deux modèles à l'aide d'analyses de variance (ANOVA) et de tests du chi deux. Les hypothèses basées sur les p-values nous ont permis de déterminer si le modèle avec toutes les variables était significativement meilleur que le modèle vide.
4. Sélection des variables à l'aide de la Méthode Step : Nous avons appliqué la méthode step (Voir [A](#)) pour sélectionner un sous-ensemble de variables optimisé. Cette procédure nous a permis d'éliminer quelques indicateurs dans le modèle afin d'améliorer sa performance en se basant sur l'AIC.
5. Comparaison des modèles finalisés : Enfin, nous avons comparé le modèle simplifié résultant de la sélection des variables (troisième modèle) avec le modèle complet (deuxième modèle) pour évaluer si la sélection des variables a conduit à une amélioration significative du modèle.

Nous avons répété cette démarche en changeant les types d'événements dans nos données pour évaluer la capacité de notre modèle à distinguer entre différents types d'événements. Cette itération nous a permis d'identifier les variables les plus pertinentes pour chaque distinction.

5.2.3 Qualité du modèle

Pour évaluer la qualité des modèles résultants, on a utilisé plusieurs métriques, notamment le train-test split (70% et 30% respectivement), la matrice de confusion et l'accuracy (Voir [A](#)) .

1. **Pour toutes les données (l'ensemble des évènements) :**

Indicateurs enlevés par la sélection des variables : le rapport entre les pentes de désaturation et resaturation.

Test accuracy : 0.5268657 Train accuracy : 0.5704225

	1	2	3	4
1	143	22	56	64
2	4	41	0	14
3	39	1	88	30
4	43	15	29	81

Matrice de confusion (test)

	1	2	3	4
1	339	36	106	145
2	21	118	6	16
3	70	2	223	66
4	100	38	65	211

Matrice de confusion (train)

Les perfor-

mances du modèle sont relativement modestes, avec une accuracy de test de seulement 52,69%. Les résultats de la matrice de confusion montrent que le modèle a des difficultés à bien classer les événements dans les différentes catégories. Ces difficultés à bien classer les événements est cohérent avec les observations sur les boxplots (Figure 8).

2. **Pour les apnées centrales, obstructives et les hypopnées centrales :**

Indicateurs enlevés par la sélection des variables : le rapport entre les pentes de désaturation et resaturation.

Test accuracy : 0.6701245

Train accuracy : 0.7105966

	1	2	3
1	175	27	80
2	10	50	3
3	37	2	98

	1	2	3
1	436	68	144
2	26	122	8
3	75	4	240

Lorsque nous

Matrice de confusion (test)

Matrice de confusion (train)

nous sommes concentrés sur un sous-ensemble spécifique d'événements, à savoir les apnées centrales, obstructives et les hypopnées centrales, les performances du modèle se sont améliorées. L'accuracy de test est passée à 67,01%, mais elle reste encore modeste.

3. **Pour les apnées centrales et les apnées obstructives :**

Indicateurs enlevés dans la sélection des variables : la pente de resaturation, le rapport entre les pentes de désaturation et resaturation et le rapport entre les durées de désaturation et de resaturation.

Test accuracy : 0.8516129

Train accuracy : 0.867036

	1	2
1	211	36
2	10	53

	1	2
1	516	74
2	22	110

Matrice de confusion (test)

Matrice de confusion (train)

En se concentrant uniquement sur les apnées centrales et les apnées obstructives, les performances du modèle atteignent leur meilleur niveau, avec une accuracy de test de 85,16%. Cependant, bien que ces performances soient meilleures, elles ne sont toujours pas excellentes.

Conclusion :

L'ensemble des résultats indique que les différents modèles ont des difficultés à bien classer les événements, en particulier lorsque les événements sont plus diversifiés. La non-excellente performance pourrait être attribuée à la complexité intrinsèque des données médicales et aux hypothèses posées pour la ré-annotation des signaux. Il est essentiel de vérifier et de valider les différentes annotations et signaux pour construire des modèles sur une base solide et valide. Une analyse plus approfondie des données et une réévaluation des hypothèses pourraient améliorer les performances du modèle.

5.2.4 SpaceM3

SpaCEM3 (Spatial Clustering with EM and Markov Models) est un logiciel développé par l'équipe Mistis de l'Inria Grenoble Alpes. Ce logiciel propose une variété d'algorithmes pour la classification, supervisée ou non supervisée, de données uni ou multidimensionnelles en interaction, certaines de ces données pouvant être manquantes. Les structures de dépendances prises en compte sont celles pouvant être décrites par un graphe fini quelconque. Elles incluent le cas particulier des grilles régulières utilisées notamment en segmentation d'images. L'approche principale se fonde sur l'utilisation de l'algorithme EM pour une classification et sur les modèles de champs de Markov pour la modélisation des dépendances.

Dans le cadre du stage et avec l'aide des données fournies, un graphe a été utilisé pour modéliser les données et permettre l'utilisation du logiciel. L'ensemble de données a été représenté par un graphe contenant 10 composantes connexes complètes (un arc entre tous ses sommets), chaque composante connexe représentant un patient différent. Chaque événement est représenté par un sommet, et les événements du même patient sont liés par des arcs. Le graphe résultant a été utilisé pour effectuer la classification/clustering dans ce logiciel. Après avoir créé les fichiers nécessaires et après le lancement du logiciel à l'aide de docker, nous n'avons pas pu utiliser le logiciel comme il se doit en raison d'une erreur que nous n'avons pas réussi à résoudre (plusieurs essais ont été faits).

5.2.5 HDclassif

HDclassif est un package R dédié au regroupement et à l'analyse discriminante de données de grande dimension. Les méthodes de classification proposées dans le package résultent d'une nouvelle paramétrisation du modèle de mélange gaussien, combinant l'idée de réduction de dimension et des contraintes sur les matrices de covariance. La méthode de classification supervisée utilisant cette paramétrisation est appelée analyse discriminante en grande dimension (HDDA). De manière similaire, la méthode de regroupement associée est appelée regroupement de données en grande dimension (HDDC) et utilise l'algorithme d'espérance-maximisation pour l'inférence.

Nous avons utilisé les deux méthodes de classification et de regroupement sur les données en changeant à chaque fois le modèle utilisé (les contraintes sur les matrices de covariance), mais les résultats dans la phase de validation ne dépassaient pas 50% de précision.

6 Bilan personnel

J'ai eu l'occasion de réaliser mon stage en tant qu'assistant ingénieur au sein du laboratoire INRIA Grenoble Alpes, en collaboration avec le laboratoire HP2 du CHU de Grenoble. Cette expérience m'a ouvert les portes du monde de la recherche scientifique, où j'ai travaillé sur des données médicales et effectué des analyses visant à confirmer les hypothèses formulées au début du projet, notamment en ce qui concerne les types d'événements et leur caractérisation. Cependant, tout au long de ce stage, j'ai été confronté à plusieurs défis liés aux données, en raison de leur complexité et de la manière dont les annotations avaient été effectuées.

J'ai beaucoup apprécié la liberté dont j'ai bénéficié dans mon travail, tant sur le plan technique que dans le choix de mon approche et de mon code. J'ai également eu une grande flexibilité dans l'organisation de mon temps et de mes tâches, ainsi qu'un encadrement très constructif, notamment à travers nos réunions hebdomadaires avec l'un des encadrants, et même avec l'ensemble de mes encadrants. J'ai eu l'opportunité d'assister aux réunions de l'équipe et aux présentations effectuées par ses membres, ainsi qu'à la soutenance de thèse d'un de mes collègues de la même équipe.

Ce stage revêt également une grande importance dans mon parcours professionnel. Initialement, je n'avais pas envisagé de poursuivre mes études en effectuant une thèse après l'obtention de mon diplôme à l'ENSIMAG. Mon intention était de me lancer directement dans le domaine industriel. Cependant, cette expérience m'a permis de faire mes premiers pas au sein d'un laboratoire de recherche, me donnant un aperçu de la réalité du travail d'un chercheur ou d'un doctorant. Elle m'a également offert une opportunité rare d'observer le fonctionnement interne d'un laboratoire de recherche prestigieux tel que l'INRIA.

Travailler aux côtés de chercheurs expérimentés, de doctorants et d'autres stagiaires a été une expérience particulièrement enrichissante. Les échanges et les discussions avec ces collègues m'ont permis de bénéficier de leurs connaissances et de leurs expériences, élargissant ainsi ma propre compréhension du domaine de la recherche en informatique et en santé.

En conclusion, ce stage a non seulement enrichi mes compétences professionnelles, mais il a également transformé ma vision de l'avenir. Il m'a ouvert de nouvelles perspectives, notamment celle de poursuivre des études supérieures et de m'engager dans la recherche scientifique. Cette expérience a été une étape essentielle dans mon développement académique et professionnel, m'apportant une précieuse compréhension du monde de la recherche et du rôle important qu'elle joue dans l'innovation et l'avancement des connaissances.

7 Conclusion et perspectives

L'objectif de mon stage au sein de l'INRIA était de caractériser les différents types d'événements à l'aide des indicateurs extraits des saturations d'oxygène SaO2 mesurées lors de la polysomnographie, sous forme de séries temporelles ([lien vers le Répertoire Git](#)). J'ai commencé par identifier les différents événements pour les patients et procédé à une reannotation des saturations en me basant sur les annotations existantes. Ensuite, j'ai extrait un ensemble d'indicateurs pour chaque événement chez tous les patients, créant ainsi une table de données regroupant l'ensemble des événements, leurs caractéristiques, ainsi qu'une variable indiquant leur type.

Afin d'atteindre cet objectif, j'ai testé divers algorithmes de classification et de clustering, en passant ensuite par une étape de validation pour évaluer les résultats. Cependant, comme mentionné précédemment, les modèles construits n'ont pas montré des performances exceptionnelles dans la caractérisation des différents types d'événements à partir des caractéristiques extraites. Il est possible que cela soit dû au traitement préalable des données.

Par la suite, l'objectif sera d'explorer des méthodes de classification et de clustering fonctionnels, c'est-à-dire directement sur les courbes sans extraire de caractéristiques, mais en utilisant les formes des signaux des événements. Par exemple, l'utilisation d'un modèle de régression multinomiale fonctionnelle pourrait être envisagée.

([lien vers le Répertoire Git](#))

References

- [1] 1. *Polysomnography: definition, indication and interpretation of results*. URL: <https://back2sleep.eu/blogs/news/polysomnographie-definition-indication-interpretation>.
- [2] 2. *L'apnée du sommeil est-elle incurable?* URL: <https://dorconfort.com/l-apnee-du-sommeil-est-elle-incurable>.
- [Kwo+21] Younghoon Kwon et al. “Prolonged circulation time is associated with mortality among older men with sleep-disordered breathing”. In: *Chest* 159.4 (2021), pp. 1610–1620.
- [Lev+21] Jeremy Levy et al. “Digital oximetry biomarkers for assessing respiratory function: standards of measurement, physiological interpretation, and clinical use”. In: *NPJ digital medicine* 4.1 (2021), p. 1.
- [Myr] Frédéric Bertrand et Myriam Maumy. *Choix du modèle*. URL: https://irma.math.unistra.fr/~fbertran/enseignement/Master1_2010_2/Master2Cours3.pdf.
- [Wik] Wikipedia. *Critère d'information d'Akaike* — *Wikipedia, The Free Encyclopedia*. URL: https://fr.wikipedia.org/wiki/Crit%C3%A8re_d%27information_d%27Akaike.

A Annexe

Tableau 1: Statistiques des variables pour les différents types d'événements

Indicateures	std	min	mean	max	75%	50%	25%
MAX_desat	1.28	92.50	97.96	99.50	98.50	98.50	97.50
MIN_desat	2.69	77.50	92.17	96.50	94.50	92.50	90.50
MAX_resat	1.31	92.50	97.99	99.50	98.50	98.50	97.50
DUREE_event	20.38	10.00	49.53	119.99	59.98	47.36	35.66
DUREE_desat	8.32	0.01	19.47	115.66	26.01	20.01	13.35
DUREE_resat	18.68	0.99	30.06	108.33	37.66	27.00	17.33
DELTA_desat	2.75	3.00	5.79	18.00	7.00	5.00	4.00
DELTA_resat	2.83	3.00	5.82	21.99	6.99	4.99	3.99
PENTE_desat	41.66	-896.00	-3.35	-0.05	-0.20	-0.29	-0.43
PENTE_resat	0.42	0.03	0.32	5.04	0.39	0.20	0.12
area_down	113.81	-1738.39	-130.52	-0.40	-53.19	-102.08	-173.81
area_up	27.15	0.00	11.45	209.39	8.02	0.00	0.00
DD100	2.69	3.50	7.83	22.50	9.50	7.50	5.50
DDbase	3.33	-1.99	5.13	16.99	6.99	4.99	2.99
TRatio	2.58	0.00	1.14	73.67	1.08	0.71	0.44
PRatio	88.86	0.02	8.23	2095.87	2.33	1.42	0.91

Signification des indicateurs

Dans cette section, nous détaillons les significations des différents indicateurs utilisés dans notre analyse.

- **MAX_desat** : La valeur du signal au début de l'évènement.
- **MIN_desat** : La valeur minimale du signal tout au long de l'évènement.
- **MAX_resat** : La valeur à la fin de l'évènement (hors hyperventilation).
- **DUREE_event** : La durée totale de l'évènement (sans prendre en compte l'hyperventilation).
- **DUREE_desat** : La durée entre le début de l'évènement et l'atteinte du minimum du signal.
- **DUREE_resat** : La durée entre le minimum du signal SaO2 et la fin de l'évènement.

- **PENTE_desat** : La pente de la désaturation.
- **PENTE_resat** : La pente de la restauration.
- **area_down** : L'aire sous la courbe du signal par rapport au niveau de base du patient.
- **area_up** : En cas de présence d'une hyperventilation, il s'agit de l'aire au-dessus de la courbe du signal par rapport au niveau de base du patient.
- **DELTA_desat** : La profondeur de la désaturation par rapport au maximum de désaturation, c'est-à-dire la différence entre la valeur maximale et le minimum du signal pendant la désaturation.
- **DELTA_resat** : La profondeur de la résaturation par rapport au maximum de résaturation (la fin de l'évènement), c'est-à-dire la différence entre la valeur maximale et le minimum du signal pendant la résaturation.
- **DD100** : La profondeur de la désaturation par rapport à 100%, c'est-à-dire la différence entre 100% et le minimum du signal pendant la désaturation.
- **DDbase** : La profondeur de la désaturation par rapport au niveau de base du patient, c'est-à-dire la différence entre la valeur du niveau de base et le minimum du signal pendant la désaturation.
- **TRatio** : Le rapport entre la durée de la désaturation et la durée de la restauration.
- **PRatio** : Le rapport entre la pente de la désaturation et la pente de la restauration.

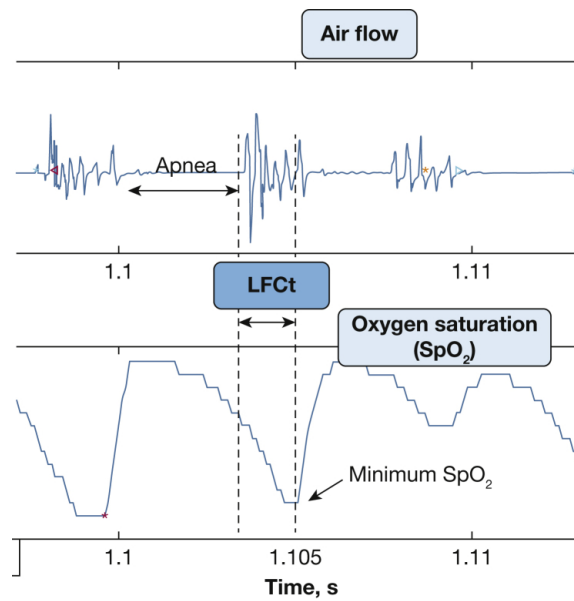


Figure 9: Illustration du décalage entre la fin de l'apnée ou de l'hypopnée et la SaO₂ la plus basse

Subject: PA1718

Recording: 170126H-A

Start: 26 jan 2017 23:02:48 End: 27 jan 2017 6:28:09

Signals in file: (25 selected)

1	EEG Cz	256 Hz
2	EEG O2	256 Hz
3	EMG	256 Hz
4	EOG1	256 Hz
5	EOG2	256 Hz
6	JMB1	256 Hz
7	JMB2	256 Hz
8	ECG	256 Hz
9	THER	256 Hz
10	THO	256 Hz
11	ABD	256 Hz
12	OES	256 Hz
13	POS	256 Hz
14	SAT	256 Hz
15	FC	256 Hz
16	Pouls	256 Hz
17	PTT	256 Hz
18	DEB	256 Hz
19	RONF	256 Hz
20	PCO2	256 Hz
21	TOSCA	256 Hz
22	PRES	256 Hz
23	EEG F4	256 Hz
24	EEG C4	256 Hz
25	EEG A1	256 Hz

Figure 10: Les différents channel des signaux issues de la PSG

Définitions

L'accuracy est une mesure de la performance d'un modèle, qui indique la proportion de prédictions correctes parmi l'ensemble des prédictions effectuées par ce modèle. En d'autres termes, c'est le nombre de prédictions correctes divisé par le nombre total d'exemples.

La matrice de confusion est un outil utilisé pour évaluer la performance d'un modèle de classification. Elle présente les résultats des prédictions du modèle sous forme d'une matrice, où chaque colonne représente les prédictions d'une classe particulière et chaque ligne représente les exemples réels appartenant à cette classe. La matrice de confusion permet de calculer diverses métriques de performance, telles que la précision, le rappel et le F1-score, en fonction des valeurs présentes dans cette matrice. Elle est utile pour comprendre comment un modèle se comporte lors de la classification des données dans différentes classes.

Procédure stepwise

Il s'agit d'examiner un modèle avec une seule variable explicative puis introduction une à une d'autres variables explicatives. À chaque étape, on réexamine toutes les variables introduites précédemment dans le modèle. Cette méthode réexamine les tests de Student pour chaque variable explicative anciennement admise dans le modèle. Après réexamen, si des variables ne sont plus significatives, alors retirer du modèle la moins significative d'entre elles. Le processus continue jusqu'à ce que plus aucune variable ne puisse être introduite ni retirée du modèle.[\[Myr\]](#)

Critère d'information d'Akaike AIC

Le critère d'information d'Akaike, est une mesure de la qualité d'un modèle statistique proposée par Hirotugu Akaike en 1973. Lorsque l'on estime un modèle statistique, il est possible d'augmenter la vraisemblance du modèle en ajoutant un paramètre. Le critère d'information d'Akaike, tout comme le critère d'information bayésien, permet de pénaliser les modèles en fonction du nombre de paramètres afin de satisfaire le critère de parcimonie. On choisit alors le modèle avec le critère d'information d'Akaike le plus faible. [\[Wik\]](#)