

Analyse quantitative de l'électroencéphalogramme en
sommeil paradoxal

par

Samuel POULIN

MÉMOIRE PRÉSENTÉ À L'ÉCOLE DE TECHNOLOGIE SUPÉRIEURE
COMME EXIGENCE PARTIELLE À L'OBTENTION DE LA MAÎTRISE
AVEC MÉMOIRE EN GÉNIE DES TECHNOLOGIES DE LA SANTÉ
M. Sc. A

MONTRÉAL, LE 21 DÉCEMBRE 2022

ÉCOLE DE TECHNOLOGIE SUPÉRIEURE
UNIVERSITÉ DU QUÉBEC

©Tous droits réservés, Samuel Poulin, 2022

©Tous droits réservés

Il est interdit de reproduire, d'enregistrer ou de diffuser en tout ou en partie, le présent document.
Le lecteur qui désire imprimer ou conserver sur un autre média une partie importante de ce document, doit obligatoirement demander l'autorisation à l'auteur.

PRÉSENTATION DU JURY

CE MÉMOIRE A ÉTÉ ÉVALUÉ

PAR UN JURY COMPOSÉ DE :

M. Jean-Marc Lina, directeur de mémoire
Département de génie électrique à l'École de technologie supérieure

Mme Nicola Hagemeister, présidente du jury
Département de génie des systèmes à l'École de technologie supérieure

M. Mohamad Forouzanfar, membre du jury
Département de génie des systèmes à l'École de technologie supérieure

IL A FAIT L'OBJET D'UNE SOUTENANCE DEVANT JURY ET PUBLIC

LE 19 DÉCEMBRE 2022

À L'ÉCOLE DE TECHNOLOGIE SUPÉRIEURE

REMERCIEMENTS

Je tiens à remercier toutes les personnes qui m'ont appuyé tout au long de ce mémoire afin de persévérer et accomplir ce travail.

Je souhaite tout d'abord remercier mon directeur de recherche M. Jean Marc-Lina qui a toujours été disponible pour répondre aux questions et alimenter la discussion afin de pousser le raisonnement toujours plus loin.

Je remercie Karine Lacourse et David Lévesque pour leur support du côté technique pour l'application du CÉAMS et me permettre d'améliorer mes compétences de programmeur.

Un grand merci à Sonia Frenette, William Yao et Giuseppe Fiamingo pour leur précieux temps d'annotations de mouvement oculaire rapide sur plus de 50 nuits de sommeil.

Finalement, merci à toute l'équipe du CÉAMS pour les commentaires constructifs aux rencontres du vendredi et les échanges très enrichissants sur le sommeil.

Analyse quantitative de l'électroencéphalogramme en sommeil paradoxal

Samuel POULIN

RÉSUMÉ

L'analyse de l'électroencéphalogramme (EEG) durant le sommeil paradoxal (REM) est une problématique dans tous les domaines nécessitant cette analyse que ce soit pour l'étude du trouble comportemental en sommeil paradoxal (TCSP) ou l'étude des micro-états par exemple. Le problème majeur rencontré en REM demeure l'omniprésence des mouvements oculaires rapides (MOR) dans l'EEG et ceux-ci masquent l'activité cérébrale sous-jacente. Ce mémoire offre une méthode afin de corriger automatiquement les artéfacts oculaires (AO) présents sur l'EEG à l'aide d'une analyse par composantes indépendantes (ICA). De plus, un détecteur de mouvement oculaire rapide semi-automatique sera implémenté en sommeil paradoxal afin d'augmenter la précision du correcteur ainsi qu'offrir un outil essentiel pour l'analyse du REM pour les patients RBD (REM Sleep Behaviour Disorder). Finalement, l'analyse des techniques permettant d'identifier les micro-états durant le sommeil paradoxal complétera cette analyse quantitative. Ces topographies caractéristiques visibles sur le scalpe sont une représentation de l'activité cérébrale sous-jacente, donc il est primordial de nettoyer la trace d'activité oculaire sur l'EEG de REM afin de ne pas biaiser ces micro-états. Ce mémoire offre donc un algorithme permettant de corriger tout en préservant l'activité thêta et même l'augmenter un peu en REM ainsi que réduire l'activité delta associée aux mouvements oculaires. Un détecteur semi-automatique permet de faire une première passe très rapidement et ainsi grandement accélérer la détection des experts pour ensuite présenter les micro-états en sommeil paradoxal. Ces outils permettront d'analyser l'intégralité du sommeil paradoxal, mais surtout découvrir le procédé réalisé par le cerveau lors de ces micro-états. Tous ces outils seront disponibles via l'application Snooz développée au Centre d'études avancées en médecine du sommeil (CÉAMS) permettant aux chercheurs et aux cliniciens d'effectuer des analyses numériques sur le sommeil grâce aux outils qui seront accessibles. Cette plateforme permettra aussi aux intéressés d'implémenter leurs propres outils sur le sommeil.

Mots-clés : Détection des artéfacts oculaires (AO), Électroencéphalogramme (EEG), Sommeil paradoxal (REM), Analyse par composante indépendante (ICA), Micro-états, Snooz

Quantitative Analysis of EEG in REM Sleep

Samuel POULIN

ABSTRACT

The analysis of the electroencephalogram (EEG) during paradoxical sleep (REM) is a problem in all areas requiring this analysis, whether for the study of behavioural disorder in paradoxical sleep (RBD) or the study of microstates, for example. The major problem encountered in REM remains the omnipresence of rapid eye movements (REM) in the EEG and these mask the underlying cerebral activity. This thesis offers a method to automatically correct ocular artifacts (OA) present on the EEG using independent component analysis (ICA). In addition, a semi-automatic rapid eye movement detector will be implemented in REM sleep to increase the accuracy of the corrector as well as providing an essential tool for REM analysis for RBD patients (REM Sleep Behaviour Disorder). Finally, the analysis of the existing methods for identifying microstates during REM sleep will complete this quantitative analysis. Since these characteristic topographies on the scalp are a representation of the underlying brain activity, it is essential to clean up the trace of ocular activity on the REM EEG to not bias these microstates. This thesis therefore offers an algorithm to correct while preserving theta activity and even increase it a little in REM as well as reduce the delta activity associated with eye movements. A semi-automatic detector makes it possible to make a first pass very quickly and thus greatly accelerate the detection of experts to then present the microstates in paradoxical sleep. These tools will make it possible to analyze all the paradoxical sleep, but principally discover the brain process during these microstates. All these tools will be available via the Snooz application developed at the Center for Advanced Research in Sleep Medicine (CARSM), allowing researchers and clinicians to perform digital analyzes on sleep using the tools that will be accessible. This platform will also allow interested parties to implement their own sleep tools.

Keywords: Ocular Artifact (OA) Detection, Electroencephalogram (EEG), REM Sleep, Independent Component Analysis (ICA), Microstates, Snooz

TABLE DES MATIÈRES

	Page
INTRODUCTION	1
CHAPITRE 1 CORRECTION DES MOUVEMENTS OCULAIRES SUR L'EEG	3
1.1 Revue de littérature	3
1.2 La base de données	5
1.3 Méthode	7
1.3.1 Soustraction des électrodes d'EOG	7
1.3.2 Analyse par composante indépendante	9
1.3.3 Subdivision des époques	11
1.3.4 Information mutuelle	12
1.3.5 Normalisation et encodage	14
1.3.6 Valeur Z	18
1.3.7 Information mutuelle accidentelle	19
1.3.8 Filtrer les composantes	21
1.3.9 Reconstruction du signal	22
1.4 Résultats	22
1.5 Discussion	24
CHAPITRE 2 DÉTECTEUR DE MOUVEMENT OCULAIRE RAPIDE	27
2.1 Revue de littérature	27
2.2 Méthode	29
2.2.1 Base de données	29
2.2.2 Soustraction des électrodes d'EOG	29
2.2.3 Subdivision des époques	29
2.2.4 Approximation polynomiale et dérivée instantanée	29
2.2.5 Sélection par percentile	30
2.3 Résultats	33
2.4 Discussion	34
CHAPITRE 3 ÉTUDE DES MICRO-ÉTATS	37
3.1 Revue de littérature	37
3.2 Méthode	39
3.2.1 Prétraitement	39
3.2.2 Algorithmes de segmentation	41
3.2.3 Critère de sélection du nombre de micro-états	44
3.2.4 Assignation d'un micro-état pour chaque échantillon temporel	46
3.3 Résultats	46
3.4 Discussion	48
CHAPITRE 4 IMPLÉMENTATION NUMÉRIQUE	49
4.1 Description	49

LISTE DES FIGURES

	Page
Figure 1.1	Techniques pour retirer les artéfacts de l'EEG3
Figure 1.2	La position des électrodes EEG sur le scalpe6
Figure 1.3	EOGs contenant un MOR sur l'époque E18
Figure 1.4	dEOG de l'époque E18
Figure 1.5	Composantes générées pour l'époque E111
Figure 1.6	Échelles temporelles utilisées.....12
Figure 1.7	Échelle d'entropie en fonction des probabilités d'observer une valeur unique13
Figure 1.8	Information mutuelle entre dEOG et chaque composante de l'époque E114
Figure 1.9	Signal dEOG normalisé sur l'époque E115
Figure 1.10	Exemple du fonctionnement d'encodage d'un signal sur 90 bins16
Figure 1.11	Signal dEOG encodé sur l'époque E117
Figure 1.12	Statistiques Z entre dEOG et chacune des composantes pour l'époque E119
Figure 1.13	Représentation de la méthode pour calculer l'hypothèse nulle entre les composantes d'EEG et du signal dEOG.....20
Figure 1.14	Distribution des valeurs d'information mutuelle pour l'hypothèse nulle21
Figure 1.15	Différences des puissances spectrales relatives par région pour les sujets MASS23
Figure 1.16	Différences des puissances spectrales relatives par région pour les patients RBD24
Figure 2.1	Performance des détecteurs de MORs disponible28
Figure 2.2	F1 score en modifiant le percentile des 55 patients RBD.....31
Figure 2.3	Distribution de la précision en fonction de la sensibilité du détecteur.....32

Figure 2.4	Distribution des variations de puissance spectrale relative par région pour les sujets MASS avec le détecteur	33
Figure 2.5	Distribution des variations de puissance spectrale relative par région pour les patients RBD avec le détecteur	34
Figure 3.1	Hypnogramme d'une nuit avec identification des cycles de REM.....	40
Figure 3.2	Schématisation de l'algorithme K-means	41
Figure 3.3	Schématisation d'un algorithme de classification hiérarchique.....	43
Figure 3.4	GEV obtenu pour chaque algorithme sur les 50 sujets MASS.....	47
Figure 3.5	Exemple de micro-états trouvés à partir de l'algorithme K-means modifié sur quatre sujets MASS.....	48
Figure 4.1	Exemple d'interface plus avancé	49
Figure 4.2	Exemple d'interface des paramètres	50
Figure 4.3	Exemple d'interface des résultats.....	51
Figure 4.4	Exemple de procédé	51

LISTE DES ABRÉVIATIONS, SIGLES ET ACRONYMES

AAHC	Atomize and Agglomerate Hierarchical Clustering
AASM	American Academy of Sleep Medicine
AO	Artéfact oculaire
CÉAMS	Centre d'études avancées en médecine du sommeil
CV	Cross-validation criterion
dEOG	Différence d'EOG
EDF	European data format
EEG	Électroencéphalogramme
EOG	Électrooculogramme
GEV	Global explained variance
GFP	Global field power
GMD	Global map dissimilarity
ICA	Independent component analysis
IM	Information mutuelle
IMa	Information mutuelle accidentelle
MASS	The montreal archive of sleep studies
MO	Mouvement oculaire
MOR	Mouvement oculaire rapide
NREM	Non-REM
R&K	Rechtschaffen & Kales
REM	Rapid eye movement (Sommeil paradoxal)
RBD	REM Sleep Behaviour Disorder

RMSD	Root mean square difference
RMSE	Root mean square error
TCSP	Trouble comportemental en sommeil paradoxal
TAAHC	Topographic Atomize and Agglomerate Hierarchical Clustering

LISTE DES SYMBOLES ET UNITÉS DE MESURE

A	Matrice de décomposition
a_k	Centre du k^e micro-état
B	Nombre de bin
K	Nombre de micro-état
l_k	Indices appartenant au k^e micro-état
M	Nombre total d'échantillon temporel d'un signal
m	m^e échantillon temporel d'un signal
μV	Microvolt
N	Nombre d'électrodes
q	q^e échantillon temporel de la sélection faites pour les maxima locaux GFP
q_{max}	Nombre de maxima locaux GFP sélectionné
S	Signal EEG
S_n	Signal de la n^e électrode
$S_{n,m}$	Signal du m^e échantillon de la n^e électrode
S_q	q^e échantillon temporel du signal EEG de la sélection
ω_m	Bruit de l'EEG
$\widehat{\omega}_m$	Estimé de l'EEG
X	Composantes de l'EEG
$z_{k,m}$	Activation des échantillons m pour les k^e micro-état

INTRODUCTION

Le sommeil paradoxal ou REM (Rapid Eye Movement) est un des cinq stades de sommeil parmi N1, N2, N3, l'éveil et le REM (DORMEZ LÀ-DESSUS, 2019) et se caractérise par la présence de mouvements oculaires (MO) comme son nom l'indique. Les yeux se situant près de la zone frontale, laisse une trace de cette activité sur l'EEG principalement sur les électrodes F7 et F8. Ces mouvements oculaires se combinent souvent avec l'activité cérébrale sous-jacente lors de la prise de donnée en électroencéphalographie. Cette interférence empêche l'analyse exhaustive de ce stade de sommeil, donc d'un point de vue clinique pour l'étude du trouble comportemental en sommeil paradoxal (TCSP) ou l'étude des micro-états, l'EEG ne peut pas être exploité dans son entièreté. Actuellement, il existe très peu d'outils permettant de corriger l'EEG de REM et la plupart des chercheurs/cliniciens doivent sélectionner manuellement des sections qui ne contiennent aucune trace de mouvements oculaires, ce qui réduit considérablement la quantité de données disponibles. Le stade de sommeil paradoxal constitue par ailleurs 20% à 25% d'une nuit (Simor et al., 2020), on ne peut donc pas faire l'économie d'époques de REM puisque nous en avons déjà peu par nuit comparativement au non-REM (NREM). La sélection manuelle de données n'est pas acceptable pour des études sur de larges cohortes puisque le temps d'annotation pour une nuit de sommeil est relativement élevé (et sujet à interprétation si on ne dispose que d'un seul 'scoreur'). Pour ces raisons, à l'heure actuelle, il est impossible d'analyser exhaustivement et quantitativement l'EEG de REM à des fins d'analyse sur des cohortes pour répondre à une question clinique.

L'objectif de ce mémoire est de développer et d'implémenter des outils numériques permettant d'analyser plus efficacement tout le sommeil paradoxal d'un individu. Le premier outil sera donc un correcteur automatique d'artéfacts oculaires utilisant l'information mutuelle entre l'EOG et l'ensemble des canaux EEG afin de supprimer la composante portant la trace du mouvement oculaire. De plus, afin de rendre plus spécifique la correction, un détecteur semi-automatique de mouvement oculaire rapide sera développé afin d'être spécifique au sommeil paradoxal d'un individu. Finalement, comme application immédiate du nettoyage de l'EEG de

REM, un inventaire exhaustif des micro-états sera réalisé. Ces micro-états sont des topographies EEG quasi stables de l'activité cérébrale globale: l'inventaire complet nécessite l'absence de traces du mouvement oculaire sur l'EEG. Afin de rendre ces outils disponibles pour la communauté des chercheurs et des cliniciens, ces derniers ont été implémentés dans l'application Snooz en cours de développement au CÉAMS. Cette application sera disponible auprès des cliniciens et chercheurs afin qu'ils puissent utiliser des outils numériques innovants pour étudier le sommeil. Cette plateforme numérique permet de d'implémenter, de valider et de partager de nouveaux outils pour l'analyse du le sommeil chez l'humain.

Pour chacun des chapitres abordés, les sujets seront présentés dans l'ordre suivant : La revue de la littérature, la méthode développée, les résultats obtenus et finalement la discussion. Le premier chapitre abordera la correction de l'EEG de REM. Le deuxième chapitre présentera le détecteur de mouvement oculaire rapide et le troisième chapitre l'identification des micro-états en REM. Le quatrième chapitre portera sur l'implémentation numérique dans l'application Snooz.

CHAPITRE 1

CORRECTION DES MOUVEMENTS OCULAIRES SUR L'EEG

1.1 Revue de littérature

Tel que mentionné précédemment, le stade REM occupe 20-25% de la nuit (Simor et al., 2020) et l'activité oculaire rapide masque le signal de l'activité cérébrale sous-jacente. L'utilisation d'enregistrements où l'activité oculaire est marginale réduit considérablement la quantité de données disponibles. La plupart des articles s'attaquent à la correction des artéfacts en stade NREM et peu de travaux ont considéré les artéfacts causés par les mouvements oculaires rapides en REM.

Tout d'abord, un état des lieux est de mise afin de voir ce qui existe comme technique pour retirer les différents artéfacts présents sur l'EEG. L'étude '*Removal of Artifacts from EEG Signals: A Review*' (Jiang et al., 2019) présente les différentes méthodes en soulignant leurs avantages et leurs limites. La Figure 1.1 présente les principales techniques utilisées dans la littérature par leur pourcentage d'occurrences dans les études.

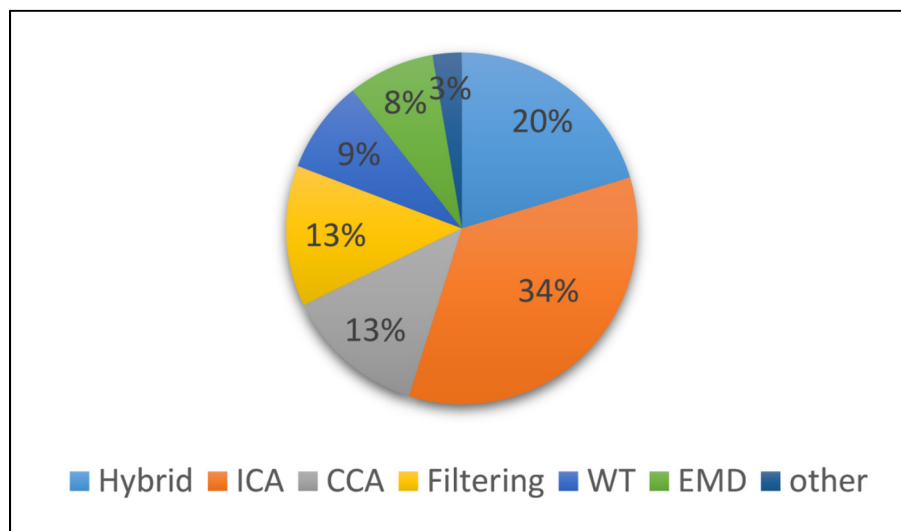


Figure 1.1 Techniques pour retirer les artéfacts de l'EEG
Tiré de Jiang et al. (2019, p. 2)

Comme l'on peut voir, la majorité des études utilisent l'analyse par composantes indépendantes (ICA) pour cibler et corriger les artéfacts. Cette méthode consiste à trouver une matrice de poids permettant de décomposer les canaux EEGs en maximisant l'indépendance des composantes entre elles. Les méthodes hybrides utilisent aussi l'ICA conjointement avec d'autres outils afin de pallier les différents problèmes engendrés par cet outil. Étant donné que le mouvement oculaire est censé être indépendant de l'activité cérébrale pendant le sommeil, il va de soi que l'ICA est la technique la plus utilisée puisque les composantes indépendantes ont davantage de chance de porter l'information d'un mouvement oculaire c'est pourquoi l'ICA sera la méthode proposée dans cette thèse utilisée conjointement avec une autre approche statistique qui seront présentées à la section 1.3.

En second lieu, la validation des différents outils est un point important puisqu'il est difficile d'affirmer avec certitude si les artéfacts ont bien été corrigés sur l'EEG. En effet, la 'vérité' n'est pas connue, puisque le signal 'nettoyé' n'existe pas à priori. Deux méthodes sont souvent utilisées dans la littérature. Dans la première, il s'agit d'évaluer la puissance spectrale dans les différentes bandes de fréquences afin de voir l'impact de la correction sur la puissance spectrale en fonction de la région du cerveau. Cette technique se retrouve dans la majorité des articles sur le sujet (Betta et al., 2013; Liu et al., 2018; Waterman et al., 1992). La puissance spectrale devrait diminuer pour les régions les plus proches du lobe frontal puisque les artéfacts oculaires se propagent d'abord dans cette région tandis que les autres régions devraient demeurer sensiblement les mêmes si l'activité cérébrale n'a pas été retirée (Romero et al., 2003). La seconde méthode consiste à calculer la RMSD (Root Mean Square Difference) (Ibrahim et al., 2015) ou RMSE (Root Mean Square Error) (Nassif et al., 2020) afin d'analyser si la valeur de RMSD est élevée ou si la valeur de la RMSE est petite pour voir l'efficacité de l'estimateur.

$$RMSD = \left(\frac{1}{M}\right) \sqrt{\sum_{m=1}^M (S_m - S'_m)^2} \quad (1.1)$$

Où S_m est l'EEG original, S'_m est l'EEG corrigé et M est le nombre d'échantillons.

$$RMSE = \sqrt{\frac{\sum_{m=1}^M (\widehat{\omega}_m - \omega_m)^2}{M}} \quad (1.2)$$

Où $\widehat{\omega}_m$ est l'EEG estimé, ω_m est le bruit de l'EEG et M est le nombre d'échantillons.

L'évaluation de la puissance spectrale ou de l'écart de puissance associé au traitement des données EEG en combinaison avec l'avis d'experts sur la qualité du signal obtenu semble la validation la plus appropriée et fiable pour mesurer l'efficacité d'un outil de correction des mouvements oculaires sur l'EEG en sommeil paradoxal.

1.2 La base de données

Deux types de données EEG ont été étudiées dans ce projet de développement méthodologique. Le premier consiste en données provenant de « The Montreal Archive of Sleep Studies » (MASS) (O'Reilly et al., 2014). Il contient 200 nuits de sommeil sous forme de fichiers sous forme « European Data Format » (EDF) accompagnés de leurs annotations sous forme d'EDF+. Le premier sous-ensemble SS1 sera utilisé et il est composé de 50 sujets en santé répartis en 31 hommes et 19 femmes ayant une moyenne d'âge de 63.5 ans avec un écart-type de 5.4. Les canaux EEG utilisés sont ceux d'un montage 10-20 représentés sur la Figure 1.2, mais les électrodes Fp1 et Fp2 et celles de références A1 et A2 ne seront pas prises en compte.

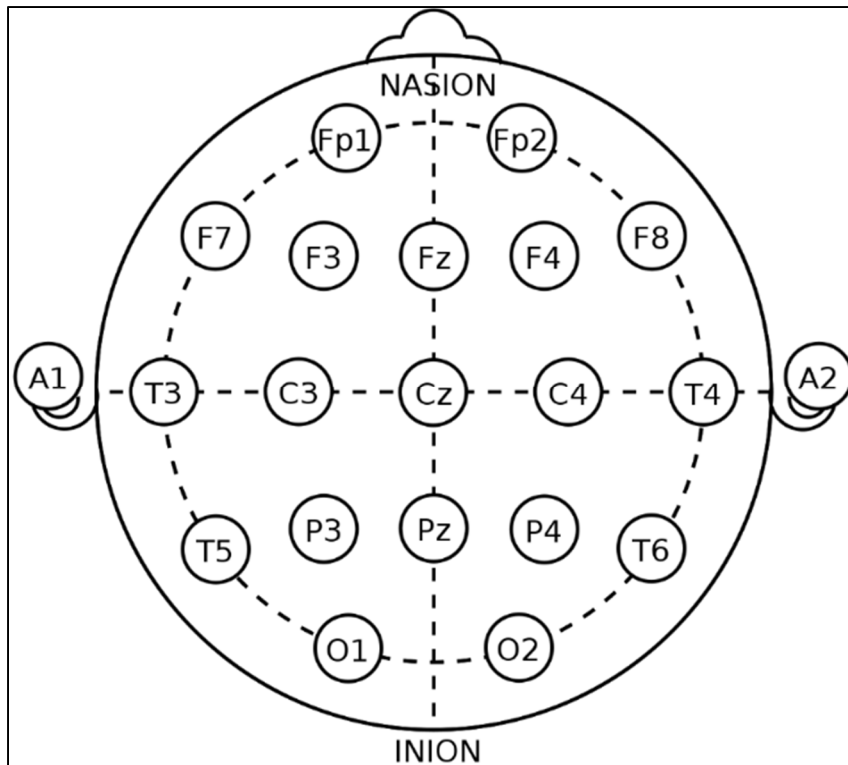


Figure 1.2 La position des électrodes EEG sur le scalpe
Tiré de Wikipedia (2010)

Le montage polysomnographique compte également deux électrodes EOG pour l'œil droit et gauche. L'enregistrement se fait pendant la nuit à 256Hz et l'EDF+ contient les stades de sommeil de la nuit par tranches de 30 secondes selon les règles de *l'American Academy of Sleep Medicine* (AASM). Puisque le but de ce mémoire est d'étudier le signal EEG de REM, seules les périodes du sommeil paradoxal seront étudiées.

Le second type de données analysées dans ce travail a été étudié en collaboration avec le Dr. Yao dont le programme de recherche porte sur des patients dits RBD. Ce trouble du sommeil correspond à une parasomnie accompagnée d'une perte d'atonie musculaire, ce qui peut engendrer des blessures (Dauvilliers et al., 2018). La cohorte du Dr. Yao comporte 154 patients (Fereshtehnejad et al., 2019), donc 55 ont été sélectionnés pour ce détecteur répartis en 41 hommes et 14 femmes ayant une moyenne d'âge de 65.7 ans avec un écart-type de 8.2 ans. Les mêmes 17 électrodes d'EEG sont utilisées ainsi que les deux mêmes électrodes d'EOG

tous enregistré à 256 Hz. Les stades de sommeil sont indiqués avec les règles de Rechtschaffen & Kales (R&K). Il est à noter que le présent travail visait à développer un outil de traitement du signal EEG qui puisse répondre ultimement aux objectifs cliniques du Dr. Yao.

1.3 Méthode

En supposant que le mouvement oculaire soit indépendant de l'activité cérébrale pendant le sommeil, nous proposons une procédure utilisant l'analyse par composantes indépendantes des signaux EEG. Partant de l'hypothèse que s'il y a une trace du mouvement oculaire, celle-ci se retrouvera dans au moins une composante des canaux EEG. L'objectif est donc de décomposer le signal en composantes statistiquement indépendantes, de calculer la corrélation entre l'EOG et ces dites composantes, de mesurer la significativité de cette corrélation pour déterminer les composantes à éliminer avant de reconstruire les canaux EEG. Certaines étapes qui ont été ajoutées afin d'amplifier cette corrélation entre l'EOG font l'objet des sections qui suivent.

1.3.1 Soustraction des électrodes d'EOG

Les mouvements oculaires sont toujours en opposition de phase sur les électrodes contralatérales puisque les électrodes étant placées en haut d'un œil et en bas de l'autre, un mouvement oculaire qui devrait être simultanément dans une même direction sera perçu en opposition sur les EOGs comme sur la Figure 1.3. Ces signaux sont enregistrés en microvolts (μV). Nommons cette époque temporelle de 10 secondes E1, car elle sera utilisée pour illustrer les différentes étapes.

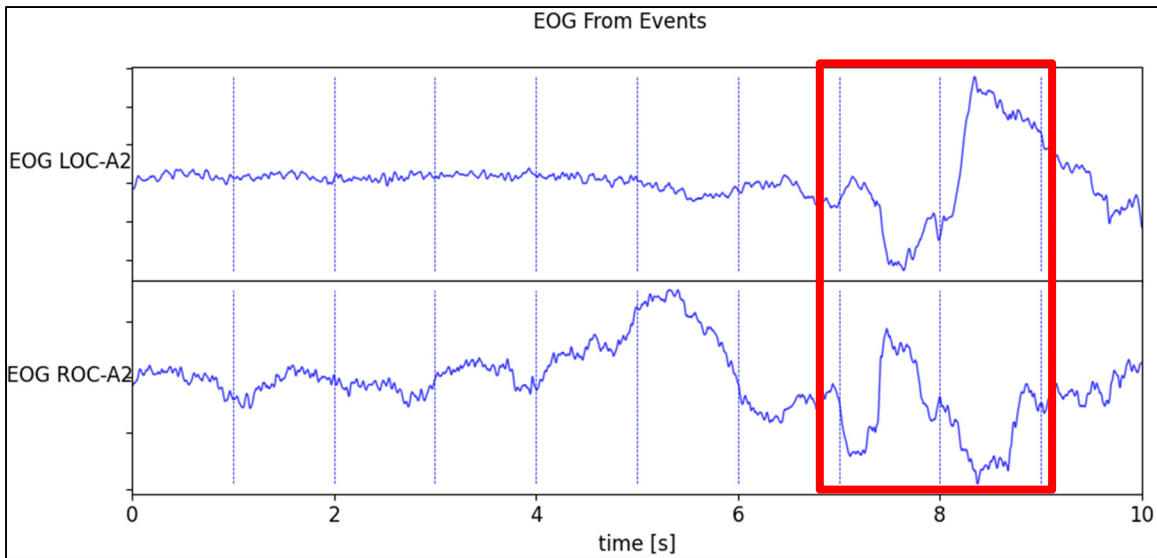


Figure 1.3 EOGs contenant un MOR sur l'époque E1

Les deux électrodes d'EOG sont donc soustraites (dEOG) afin d'amplifier le mouvement oculaire et de réduire toute activité visible sur l'EOG autre que celle recherchée telles que les complexes-K puisque ces signaux sont toujours dans la même orientation sur l'EOG. La représentation du signal résultant est présentée à la Figure 1.4.

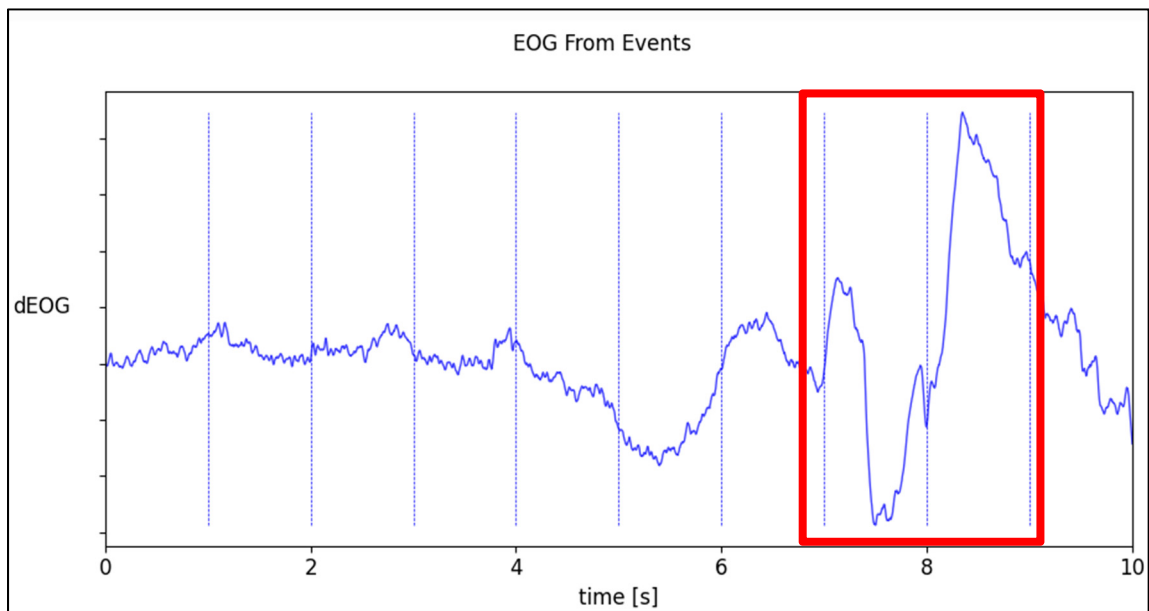


Figure 1.4 dEOG de l'époque E1

1.3.2 Analyse par composante indépendante

L'analyse par composantes indépendantes permet, comme son nom l'indique, de décomposer un signal en composantes temporelles statistiquement indépendantes. Pour l'outil proposé dans cette thèse, l'hypothèse est qu'une des composantes calculées portera l'information du mouvement oculaire si celui-ci est présent dans les canaux EEG. L'ICA a été sélectionné puisque le mouvement oculaire est censé être indépendant de l'activité cérébrale pendant le sommeil, donc logiquement une des composantes indépendantes portera l'information du MO. Chaque technique comporte ses avantages et ses inconvénients. L'ICA ne permet pas d'identifier automatiquement la composante portant l'information du MO, donc pour utiliser cet outil de correction de l'EEG de REM de façon automatique, il faut combiner d'autres techniques qui permettront de faire cette sélection automatiquement. L'ICA nécessite aussi assez d'électrodes pour capturer suffisamment d'information pour donner la possibilité de dégager des 'sources' indépendantes. Cet algorithme étant un processus itératif ayant besoin de converger pour trouver les composantes, (B. Zhou et al., 2019) a déterminé qu'il faut au minimum 8 électrodes pour avoir des résultats concluants. Notre base de données comportant 17 électrodes d'EEG nous permet d'utiliser l'ICA sans faire face à ce problème, mais il faut tenir en compte que l'outil développé nécessite au minimum 8 électrodes pour fonctionner correctement.

L'implémentation de l'algorithme provient de la librairie `sklearn.decomposition.FastICA` (Lafaye de Micheaux et al., 2022) procurant un grand nombre de paramètres permettant d'ajuster l'algorithme selon les besoins (Voir Annexe I, Tableau-A I-1 pour les paramètres utilisés). L'un des plus importants à discuter est le type d'algorithme choisi. En effet, il est possible d'utiliser le `FastIca` en parallèle, ce qui permet de calculer toutes les composantes en même temps, ou en mode 'déflation' qui permet de calculer les composantes une à la fois. La méthode par déflation a été sélectionnée malgré que la méthode parallèle est considérée supérieure puisque dans notre cas particulier, nous supposons que s'il y a du mouvement oculaire dans notre signal, celui-ci se retrouvera dans une seule composante, donc il est favorable d'accumuler l'information une composante à la fois (Miettinen et al., 2018).

Sans entrer dans les détails de l'algorithme, les composantes indépendantes sont un mélange des canaux exprimé par une relation linéaire :

$$X = AS \quad (1.3)$$

où X est la matrice de composantes recherchée [N électrodes \times M échantillons], A est une matrice de décomposition inconnue [N électrodes \times N électrodes] et S est la matrice de signaux initiaux [N électrodes \times M échantillons] (Tharwat, 2021). L'algorithme modifie de façon itérative la matrice de décomposition en maximisant l'indépendance entre les composantes jusqu'à converger. Cette matrice sera utilisée pour reconstruire le signal à partir d'une sélection de composantes dans la section 1.3.9.

Il y a plusieurs aspects de cet algorithme à connaître. Aucune des composantes n'est plus indépendante des autres. De plus, il y a une indétermination sur le signe puisque seules la direction et l'amplitude sont prises en compte. Ces aspects sont à ne pas oublier lors de l'analyse des composantes puisqu'en exécutant l'algorithme à plusieurs reprises, les résultats peuvent être légèrement différents. La prochaine étape consiste à trouver laquelle de ces composantes comprend l'information de la trace du mouvement oculaire. Pour ce faire, l'information mutuelle permet de quantifier les similitudes entre le signal dEOG et les composantes. Toutefois, les échelles temporelles seront d'abord présentées afin de clarifier la durée des époques de calculs.

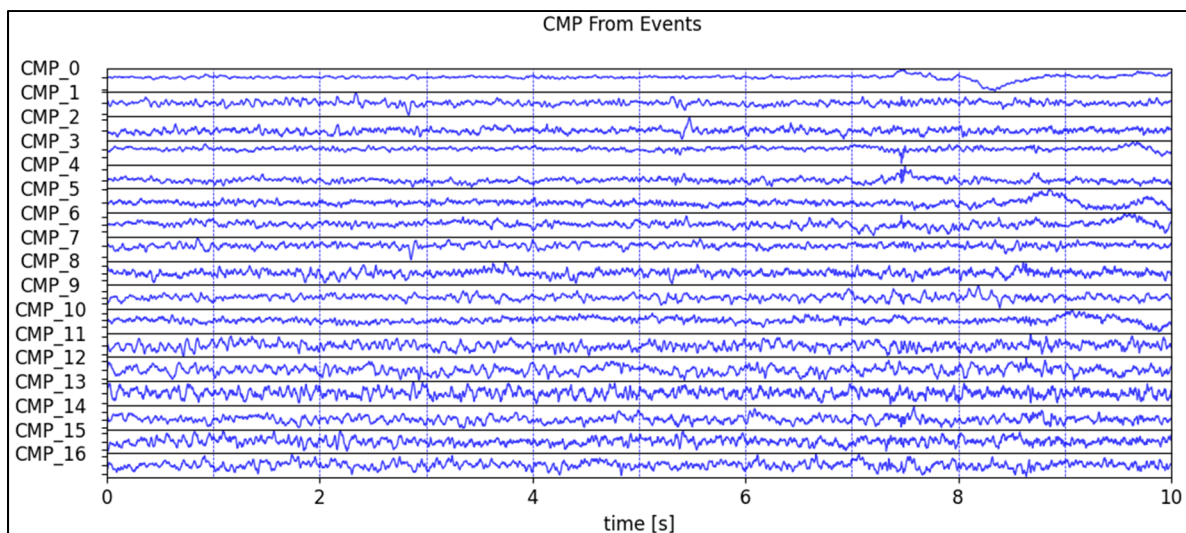


Figure 1.5 Composantes générées pour l'époque E1

1.3.3 Subdivision des époques

La durée de l'époque pour les calculs sera de dix secondes alors que les stades de sommeil sont définis en tranche de 30 secondes comme mentionnés à la section 1.2. La durée de dix secondes a été choisie, car l'ICA nécessite assez d'information provenant des signaux afin de permettre à l'algorithme de converger et ainsi d'extraire les composantes tendant vers le minimum de corrélations entre elles. Les résultats peuvent effectivement varier selon la quantité de données fournies tel que présenté par (Korats et al., 2013) et 10 secondes à 256 Hz représente 2560 valeurs ce qui est entre les deux règles de quantité minimale d'échantillons proposées par la littérature, soit supérieures à $5n^2$ (1445 données pour 17 électrodes) (Särelä & Vigário, 2003) et inférieures à $25n^2$ (7225 données pour 17 électrodes) (Onton & Makeig, 2006). De plus, 10 est un diviseur de 30 secondes, ce qui permet de découper le cycle de sommeil sans avoir de chevauchement. Il est important de retenir que les prochaines étapes seront effectuées sur des époques de dix secondes. Plusieurs échelles temporelles sont utilisées dans cette thèse c'est pourquoi la Figure 1.6 présente un résumé des différentes échelles qui seront utilisées. La détection des MOR serait davantage expliquée au CHAPITRE 2, mais les époques seront de 3 secondes.

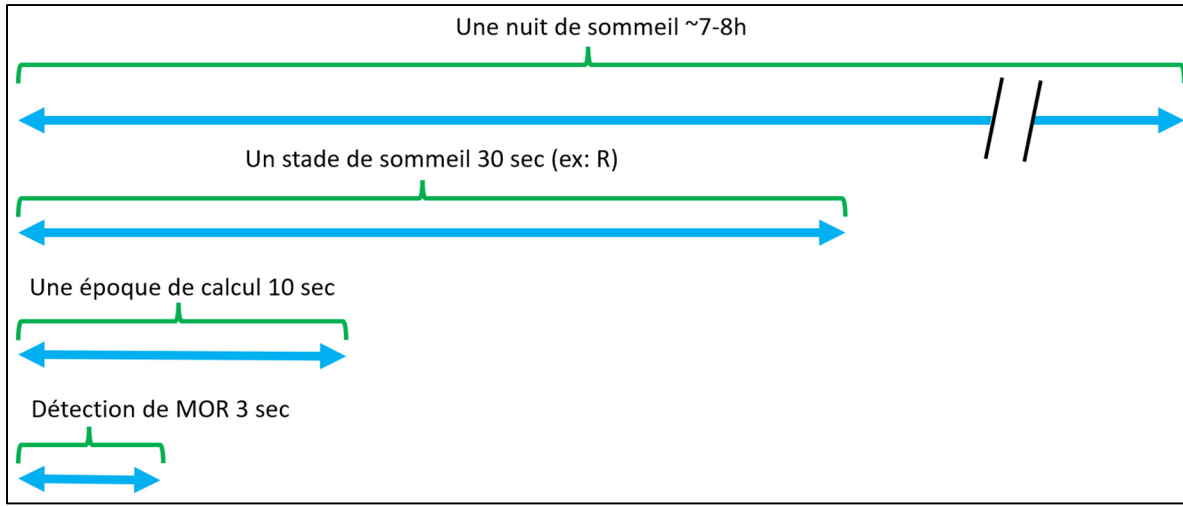


Figure 1.6 Échelles temporelles utilisées

1.3.4 Information mutuelle

L'information mutuelle (IM) est une métrique permettant d'établir, dans notre cas, la similarité entre deux signaux. L'objectif de cette section est de calculer cette métrique entre chacune des composantes et le signal dEOG. La définition de l'information mutuelle dans notre cas est définie comme suit¹ (Pluim et al., 2003) :

$$IM(\text{CMP}, \text{dEOG}) = H(\text{CMP}) + H(\text{dEOG}) - H(\text{CMP}, \text{dEOG}) \quad (1.4)$$

Où $H(\text{CMP})$ est l'entropie d'une composante, $H(\text{dEOG})$ est l'entropie du signal dEOG et $H(\text{CMP}, \text{dEOG})$ est l'entropie jointe entre une composante et dEOG. L'entropie correspond au désordre d'un système et la cohérence de ce désordre entre les deux signaux correspond à l'information mutuelle.

¹ À des fins d'homogénéité dans l'approche, la normalisation a été réalisée à l'aide de la librairie `sklearn.metrics.mutual_info_score` (Grisel et al., 2022) (Voir Annexe I, Tableau-A I-1 pour les paramètres utilisés).

L'expression de l'entropie est (Pluim et al., 2003):

$$H(\text{CMP}) = - \sum_m p(\text{CMP}) * \log_2 p(\text{CMP}) \quad (1.5)$$

$$H(\text{dEOG}) = - \sum_m p(\text{dEOG}) * \log_2 p(\text{dEOG}) \quad (1.6)$$

$$H(\text{CMP}, \text{dEOG}) = - \sum_m p(\text{CMP}, \text{dEOG}) * \log_2 p(\text{CMP}, \text{dEOG}) \quad (1.7)$$

Où $m = (1, 2, 3, \dots, 2560)$ pour chaque échantillon temporel, $p(\text{CMP})$ et $p(\text{dEOG})$ représente la probabilité d'observer chacune des valeurs, pour chaque composante ou pour le signal dEOG. Toutefois, chacune des valeurs étant unique dû à la précision des données, nous auront une probabilité uniforme pour les données 'brutes' (1/2560 étant donné le signal de 10 secondes enregistré à 256 Hz). Ce point sera abordé et résolu dans la section 1.3.5. La probabilité $p(\text{CMP}, \text{dEOG})$ représente la probabilité jointe $p(\text{CMP}|\text{dEOG}) * p(\text{dEOG})$. La Figure 1.7 présente l'échelle d'entropie possible selon les probabilités observées. Une probabilité qui tend vers 0% signifie qu'il est pratiquement impossible d'observer une valeur tandis qu'une probabilité de 100% signifie que l'ensemble des données contient une seule valeur qui sera observée à chaque instant.



Figure 1.7 Échelle d'entropie en fonction des probabilités d'observer une valeur unique

Cette métrique en est une de similarité entre le signal dEOG et chaque composante indépendante : elle permet ainsi de déterminer laquelle des composantes partage le plus d'information avec le mouvement oculaire.

La Figure 1.8 montre un exemple de l'information mutuelle obtenue pour une époque (époque E1 illustrée). Il y a 17 composantes numérotées de 0 à 16. Pour chaque composante, l'information mutuelle a été calculée entre celle-ci et le signal dEOG simultané. La valeur obtenue par cette métrique est indiquée pour chaque composante.

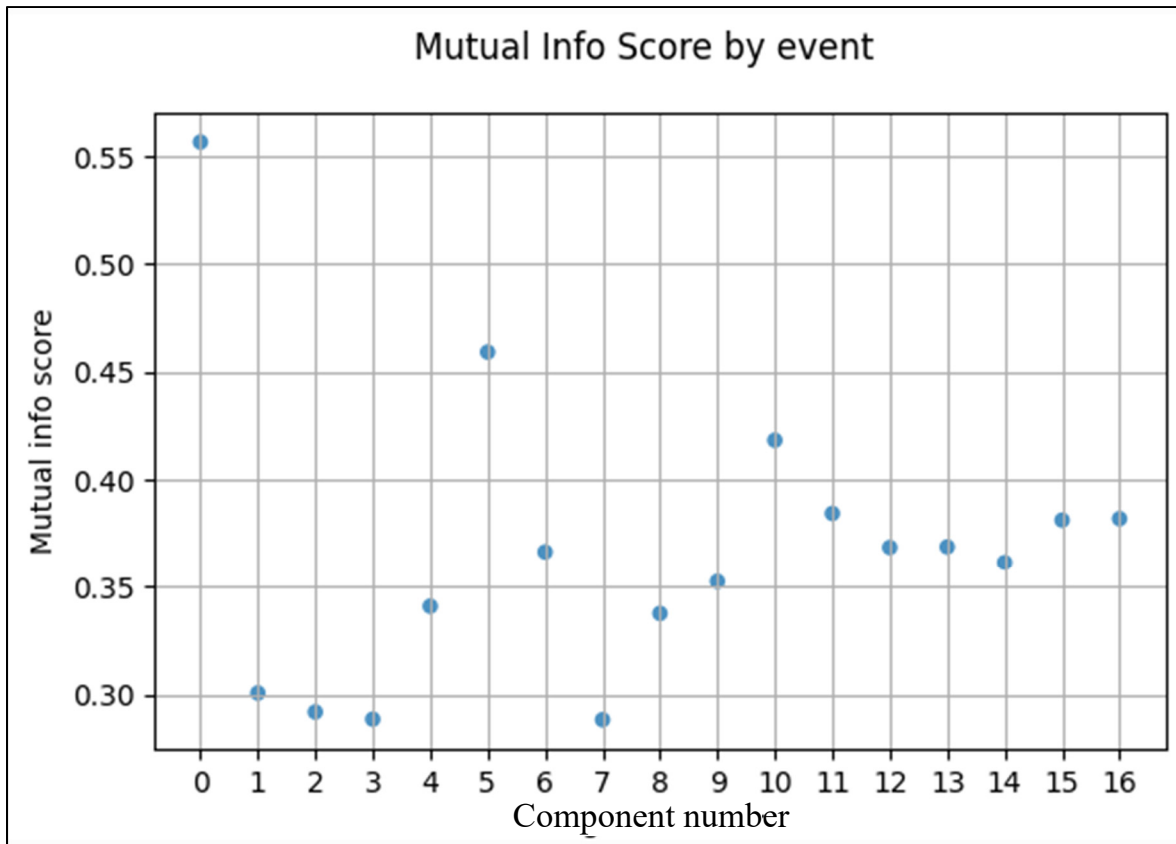


Figure 1.8 Information mutuelle entre dEOG et chaque composante de l'époque E1

1.3.5 Normalisation et encodage

Deux opérations sur les signaux sont nécessaires avant de calculer l'information mutuelle abordée à la section 1.3.4. Les concepts de normalisation et d'encodage doivent être utilisés afin de préparer les signaux dEOG et les composantes.

- La normalisation consiste à modifier l'amplitude de ceux-ci entre -1 et 1 afin de s'affranchir des amplitudes et de ne garder que les fluctuations dans les signaux. Cette modification des signaux ne vient pas changer les variations dans le signal, c'est pour cette raison qu'il est possible d'effectuer cette opération. L'époque de calcul de la normalisation sera de 10

secondes pour le dEOG et les composantes. La formule permettant de calculer la normalisation entre -1 et 1 est la suivante² :

$$S_{n,m,normalisé} = 2 * \frac{S_{n,m} - \min(S_n)}{\max(S_n) - \min(S_n)} - 1 \quad (1.8)$$

Où $S_{n,m}$ correspond au m^e échantillon de la n^e électrode d'un signal (dEOG ou une des composantes) dans une époque de 10 secondes et S_n correspond au signal de la n^e électrode. La Figure 1.9 démontre l'exemple du signal dEOG normalisé, mais la même opération est réalisée sur chacune des composantes. Le signal est identique au signal dEOG, mais entre 1 et -1 d'amplitude et les fluctuations sont toujours inchangées.

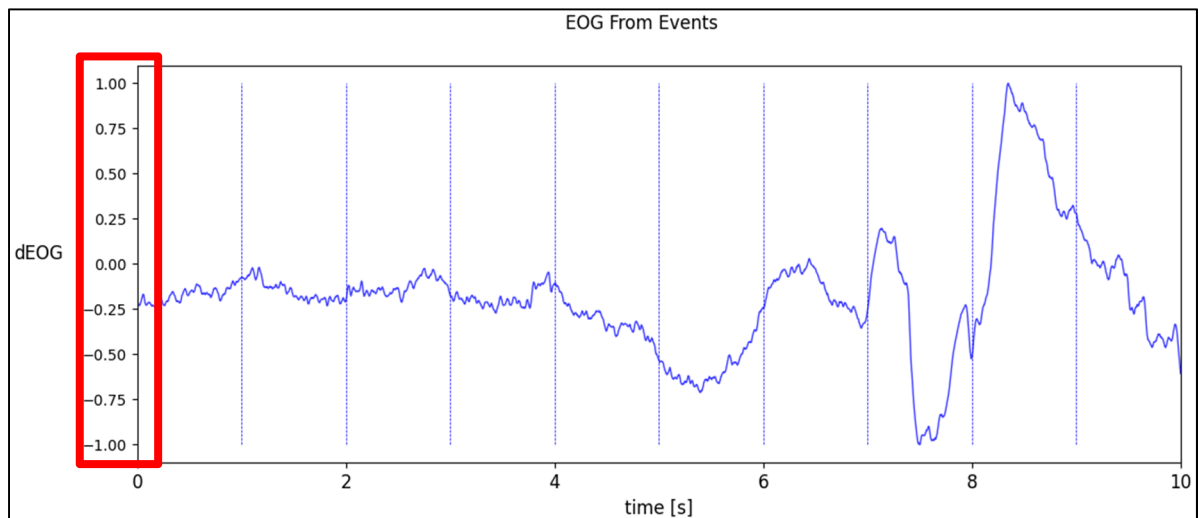


Figure 1.9 Signal dEOG normalisé sur l'époque E1

- Une dernière étape est primordiale pour être en mesure de calculer l'information mutuelle entre le signal dEOG et les composantes. Il s'agit de l'encodage. Comme vu précédemment, l'information mutuelle repose sur la probabilité d'observer chaque valeur

² À des fins d'homogénéité dans l'approche, la normalisation a été réalisée à l'aide de la librairie `sklearn.preprocessing.MinMaxScaler` (Blondel et al., 2022) (Voir Annexe I, Tableau-A I-1 pour les paramètres utilisés).

d'un signal. Cependant, si nous gardons ces valeurs à la précision initiale, la probabilité de retrouver une valeur telle que $0.834526 \mu\text{V}$, par exemple, est quasi nulle. Pour estimer une probabilité sur les amplitudes qui prend en compte de la variabilité, l'encodage des signaux permet de préparer les données au calcul d'information mutuelle. Cet encodage consiste à réduire la précision en assignant une valeur selon la 'bin' dans laquelle chaque valeur temporelle se situe. La Figure 1.10 représente l'encodage d'un signal avec 90 bins. Chaque valeur temporelle se verra attribuée l'étiquette de la bin dans laquelle la valeur se trouve et ainsi augmenter la probabilité de retrouver les valeurs uniques dans ce nouveau signal encodé. L'équation (1.9) présente la formule d'attribution des étiquettes³.

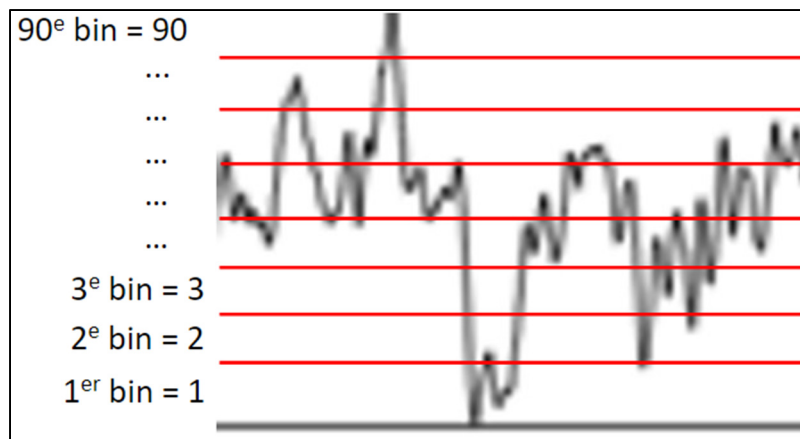


Figure 1.10 Exemple du fonctionnement d'encodage d'un signal sur 90 bins

$$S_{m,n,\text{encodé}} = \begin{cases} 1, & 0 * \frac{\Delta S_n}{B} + \min(S_n) \leq S_{n,m} < 1 * \frac{\Delta S_n}{B} + \min(S_n) \\ 2, & 1 * \frac{\Delta S_n}{B} + \min(S_n) \leq S_{n,m} < 2 * \frac{\Delta S_n}{B} + \min(S_n) \\ \dots, & \dots \\ B, & (B - 1) * \frac{\Delta S_n}{B} + \min(S_n) \leq S_{n,m} \leq B * \frac{\Delta S_n}{B} + \min(S_n) \end{cases} \quad (1.9)$$

³ À des fins d'homogénéité dans l'approche, l'encodage a été réalisée à l'aide de la librairie `sklearn.preprocessing.KBinsDiscretizer` (Lin & Dupré la Tour, 2022) (Voir Annexe I, Tableau-A I-1 pour les paramètres utilisés).

Où ΔS_n correspond à la variation du signal (dans notre cas $1 - -1 = 2$) pour la n^{e} électrodes, B est le nombre de bin désiré, S_n correspond au signal de la n^{e} électrode et $S_{m,n}$ correspond au m^{e} échantillon de la n^{e} électrode du signal (dEOG ou une des composantes) dans une époque de 10 secondes. Le nombre de bin utilisé permet de modifier la complexité des fluctuations. À l'extrême, un encodage avec une seule bin équivaldrait à une ligne droite tandis que 10 000 bins ne modifieraient pas le signal initial. Pour le signal dEOG, le nombre de 30 bins a été déterminé comme optimal afin de réduire les variations ne correspondant pas aux mouvements oculaires qui sont de plus basses fréquences. Pour les composantes indépendantes de l'EEG, le nombre de 90 bins permet de conserver les fluctuations de plus hautes fréquences associées à l'activité cérébrale. La Figure 1.11 illustre l'encodage sur 30 bins de la dEOG. Toutefois la même opération est réalisée sur 90 bins pour chacune des composantes. Le signal ressemble bien au signal initial avec moins de précision dans les hautes fréquences.

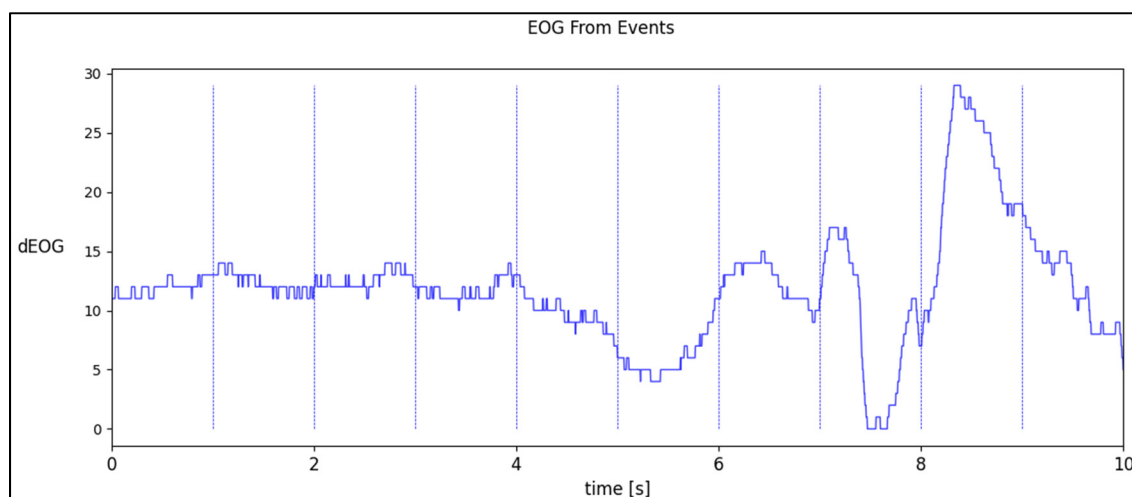


Figure 1.11 Signal dEOG encodé sur l'époque E1

Afin de simplifier les explications, il est possible de faire une analogie avec des dictionnaires. Supposons que le signal provenant de dEOG est un dictionnaire d'une langue avec 30 mots à l'intérieur qui correspond au nombre de bin dans lequel le signal a été classé. Les composantes, quant à elles, sont un dictionnaire de 90 mots d'une autre langue. Avec plus de mots, il est possible de faire des phrases plus complexes et variées. Certains mots auront la même signification dans les deux dictionnaires, d'autres auront un sens proche, mais non identique

et certains voudront dire des choses complètement différentes. Les mots signifiant la même chose correspondent à une IM de 1 et ceux qui ne signifient pas du tout la même chose auront un IM de 0 tandis que les autres seront entre les deux dépendamment de la proximité de leur signification. Les phrases produites à l'aide de ces deux dictionnaires correspondent aux fluctuations dans les deux signaux et la IM quantifie à quel point les deux phrases sont une traduction exacte.

1.3.6 Valeur Z

À partir de l'information mutuelle calculée pour chaque composante, il serait possible de simplement retirer systématiquement la composante ayant la valeur d'IM la plus élevée avant de reconstruire les signaux des canaux. Cependant, cette unique valeur ne permet pas d'affirmer que la composante ayant la plus haute valeur d'IM est effectivement porteuse de l'information du signal dEOG. Pour s'en assurer, une statistique Z peut être calculée afin de faire porter la décision de retirer une composante non pas sur la valeur de IM, mais sur la statistique attachée. Nous définissons une statistique Z,

$$Z = \frac{IM - \overline{IMa}}{\sigma IMa} \quad (1.10)$$

où IM est l'information mutuelle entre une composante et le signal dEOG telle que calculée à la section 1.3.4, alors que \overline{IMa} et σIMa sont respectivement la moyenne et l'écart-type de l'information mutuelle 'accidentelle' (IMa). L'IMa sera expliquée dans la prochaine section, mais elle donnera le critère permettant d'affirmer que les valeurs au-dessus de Z_{seuil} (équation (1.11)) portent effectivement la trace du mouvement oculaire. La Figure 1.12 présente les statistiques Z calculées pour chaque composante ainsi que le critère relatif à l'IMa indiqué par la ligne pointillée rouge : les composantes avec des valeurs qui sont au-dessus de cette limite comportent avec une p-value contrôlée la trace d'un mouvement oculaire.

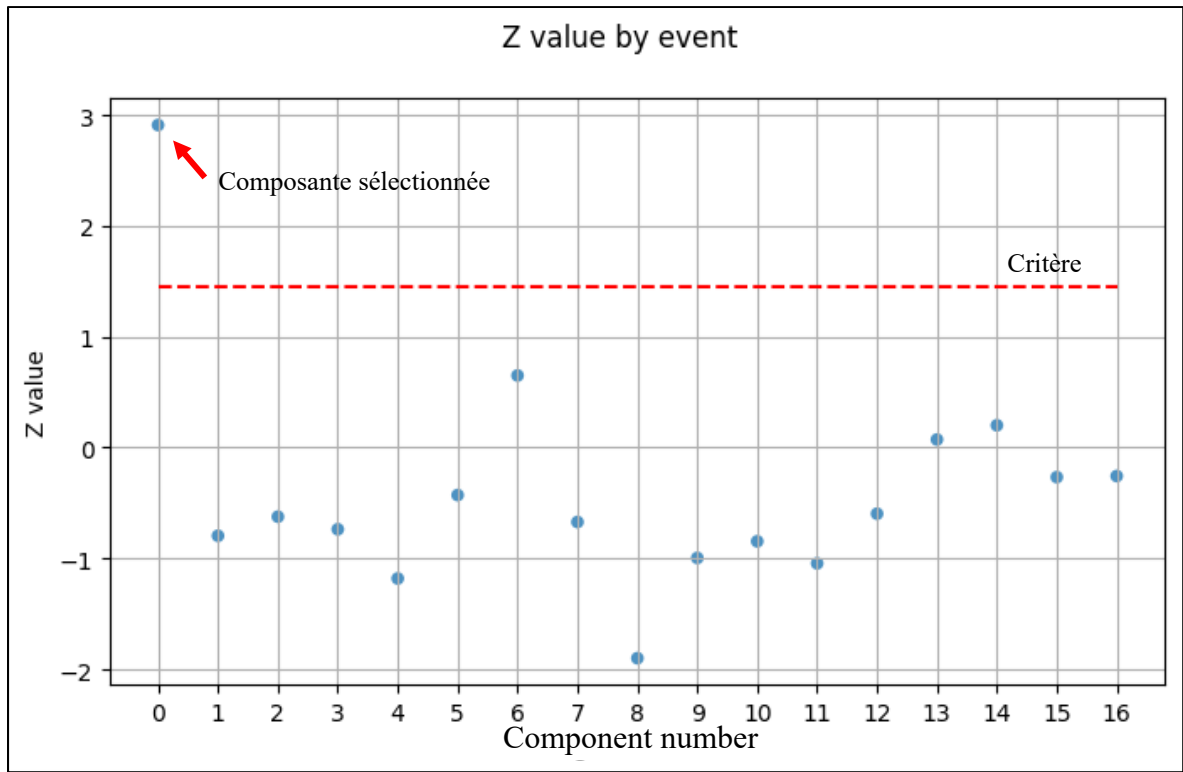


Figure 1.12 Statistiques Z entre dEOG et chacune des composantes pour l'époque E1

1.3.7 Information mutuelle accidentelle

Le critère d'information mutuelle accidentelle permettra de définir le seuil correspondant à la valeur limite d'IM permettant d'affirmer qu'une composante comporte la trace du mouvement oculaire. Pour ce faire, la modélisation de l'hypothèse nulle établira la valeur correspondant à l'absence d'IM. Normalement, il ne devrait pas y avoir d'information partagée entre les composantes indépendantes d'une époque d'EEG et le signal dEOG d'une autre époque temporelle (Époque E1 et une époque E2 par exemple). Cependant, il y aura toujours une information mutuelle résiduelle qui ne correspond pas à la présence d'un MOR, autrement dit, une l'information mutuelle 'accidentelle'. La méthode consiste donc à choisir aléatoirement des paires d'époques de composantes d'EEG et de dEOG asynchrones et d'en calculer l'IM comme le montre la Figure 1.13.

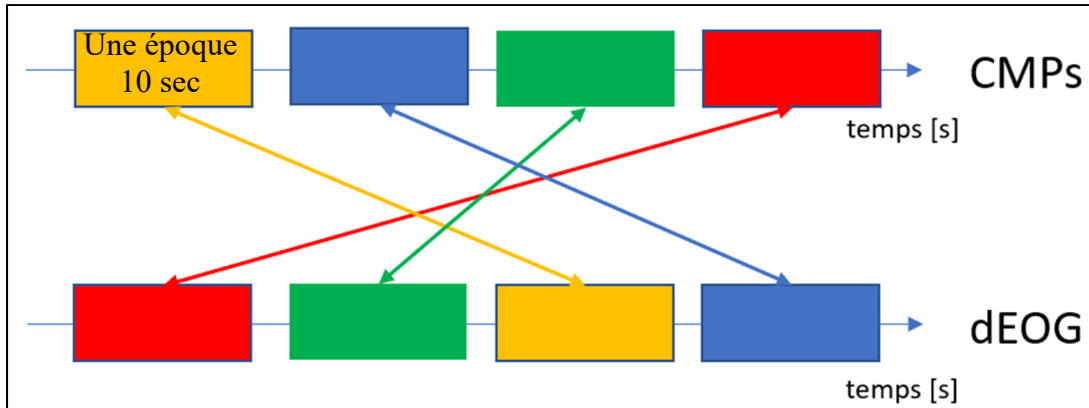


Figure 1.13 Représentation de la méthode pour calculer l'hypothèse nulle entre les composantes d'EEG et du signal dEOG

En répétant ce processus 500 fois, l'histogramme des valeurs IM tel que présenté à la Figure 1.14, décrit une distribution de l'IMa qui correspond à l'hypothèse nulle. La moyenne ainsi que l'écart-type de cette distribution d'IMa sont finalement calculés pour obtenir les statistiques Z de la formule (1.10). Cette statistique se prête à une définition d'un seuil correspondant à la p -value de l'absence d'IM. Une p -value de 0.1 a été utilisée pour définir ce seuil tel qu'indiqué par la ligne pointillée rouge sur la Figure 1.14. La valeur Z_{seuil} est donc obtenue par l'équation :

$$Z_{seuil} = \frac{IM_{p-value} - \overline{IMa}}{\sigma IMa} \quad (1.11)$$

Où $IM_{p-value}$ est la valeur de IMa à la p -value (dans notre cas p -value = 0.1). Cette valeur est spécifique à l'individu, ce qui permet de voir cette stratégie. Ce critère permet au correcteur de mouvement oculaire sur l'EEG de sélectionner automatiquement les composantes à retirer ce qui n'était pas le cas précédemment. La stratégie de sélection sera discutée à la prochaine section.

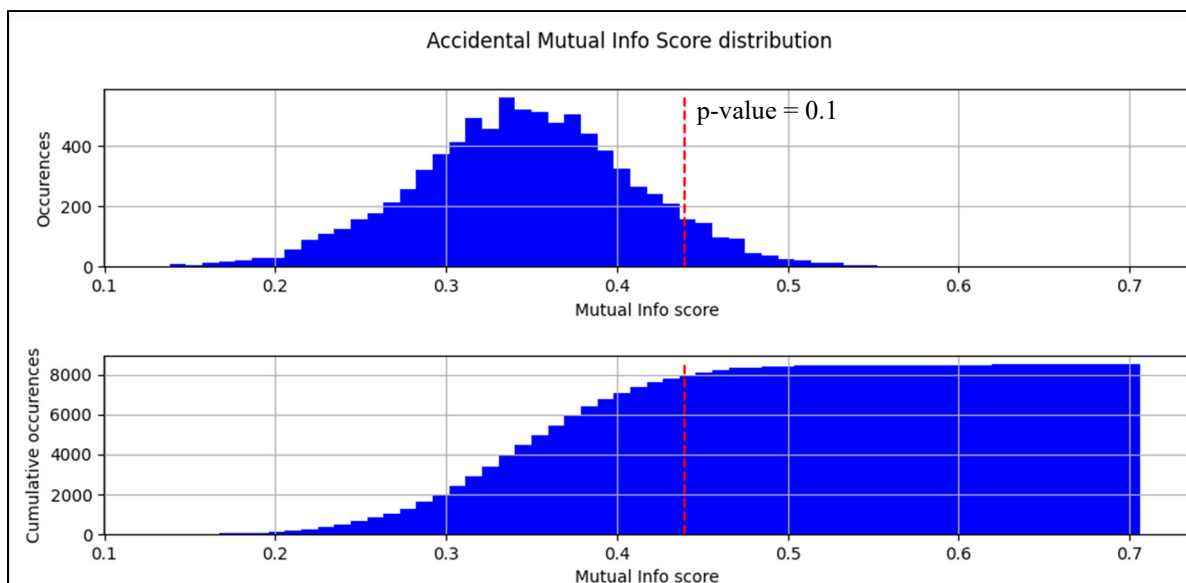


Figure 1.14 Distribution des valeurs d'information mutuelle pour l'hypothèse nulle

1.3.8 Filtrer les composantes

L'information mutuelle accidentelle permet de définir la valeur seuil associée à l'absence de mouvement oculaire dans une composante. L'hypothèse est que si un MOR est présent dans l'EEG, il ne sera présent au plus que dans une des composantes indépendantes. Cela signifie que même s'il y a une ou plusieurs composantes supérieures à la valeur seuil, seule la composante avec la plus grande valeur d'information mutuelle sera retirée. Il est à noter que dans le cas où aucune composante ne se situe au-dessus du critère, aucune composante ne sera supprimée. Pour retirer une composante, il suffit de remplacer toutes ses valeurs temporelles de cette composante par 0 avant de procéder à la reconstruction du signal. En effet, l'objectif de cet outil est de corriger avec parcimonie l'EEG de REM. Pour augmenter la spécificité de l'outil, i.e. d'être plus précis quant à la présence réelle d'un MOR dans une époque, nous introduisons un détecteur automatique de mouvement oculaire dans le CHAPITRE 2. Nous distinguons ainsi la procédure de filtrage telle que présentée dans cette section de celle relative à la détection de la présence d'un mouvement oculaire dans les époques à corriger.

1.3.9 Reconstruction du signal

Afin de reconstruire les signaux décomposés, la méthode ‘inverse_transform’ de la librairie FastICA (Lafaye de Micheaux et al., 2022) permet de les reconstruire les canaux à partir des composantes et de la matrice de décomposition. L’équation est la suivante :

$$S = X \cdot A^T \quad (1.12)$$

où X est la matrice de composantes [N électrodes \times M échantillons] et A^T est la matrice de décomposition transposée [N électrodes \times N électrodes] obtenues à la suite de l’ICA réalisé à la section 1.3.2.

1.4 Résultats

Pour évaluer l’approche numérique proposée, nous devrions comparer le résultat avec ‘la vérité’ en vérifiant ainsi que l’information introduite par un MOR a bien été retirée du signal EEG. Tel que présenté à la section 1.1, une validation indirecte via la puissance spectrale palie l’absence de la ‘vérité’. En REM, la bande de fréquences delta (0.3 à 4 Hz) est surtout dominée par les mouvements oculaires tandis que la bande de fréquences thêta (4 à 8 Hz) comporte majoritairement l’activité cérébrale de ce stade (Blumberg et al., 2020). Ainsi, en considérant la différence de la puissance spectrale relative de l’EEG sur l’ensemble de la nuit avant et après la correction sur l’EEG, celle-ci devrait montrer une réduction en delta (sur les électrodes frontales puisque l’activité oculaire aurait été retirée) et une augmentation en thêta (sur les mêmes électrodes). Ce comportement devrait être aussi observé sur les électrodes temporales, mais pas de façon systématique. Les autres électrodes devraient rester sensiblement identiques puisque les artéfacts sont davantage présents sur les signaux EEG près des yeux. La Figure 1.15 montre la variation de puissance spectrale relative par région en utilisant les segments d’enregistrements en REM de 50 sujets de la base de données Mass. La Figure 1.16 donne la même évaluation pour les 55 patients RBD du Dr. Yao.

En termes de pourcentage de correction des nuits, en moyenne 68.8% du REM par nuit a été corrigé avec un écart-type de 6.2% pour les sujets MASS et en moyenne 70.7% du REM par nuit avec un écart-type de 8% pour les patients RBD.

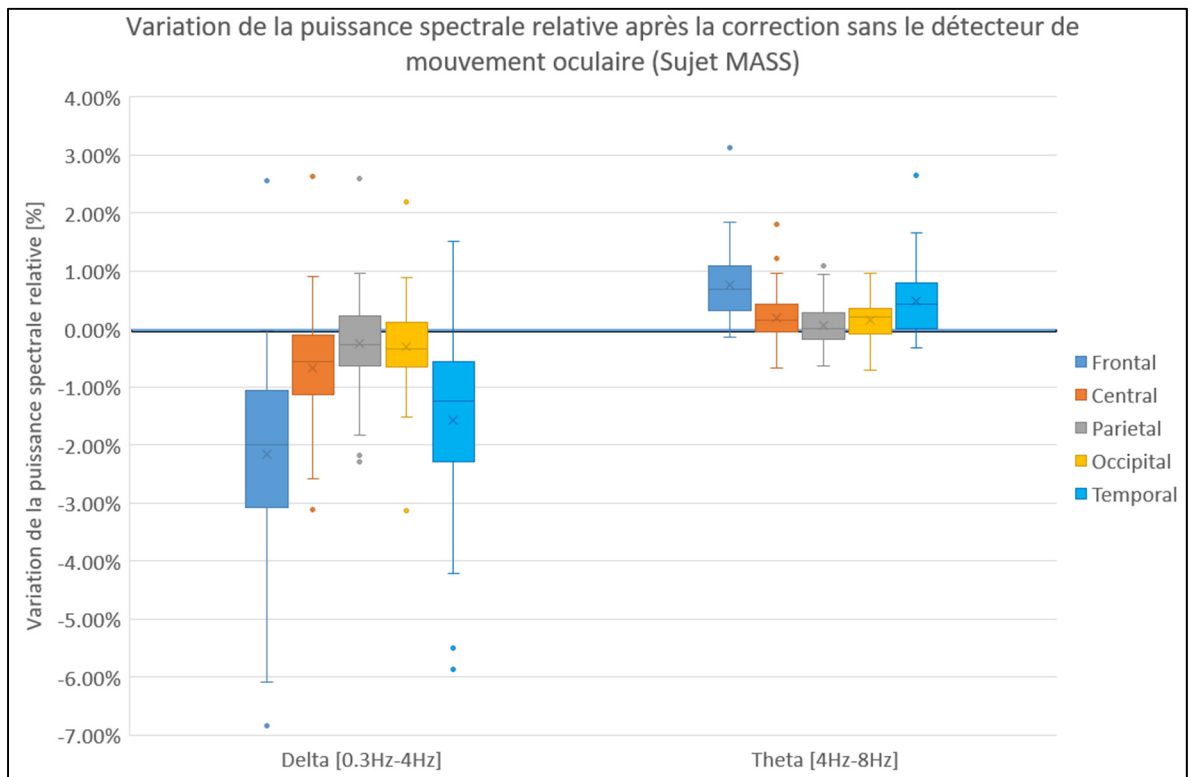


Figure 1.15 Différences des puissances spectrales relatives par région pour les sujets MASS

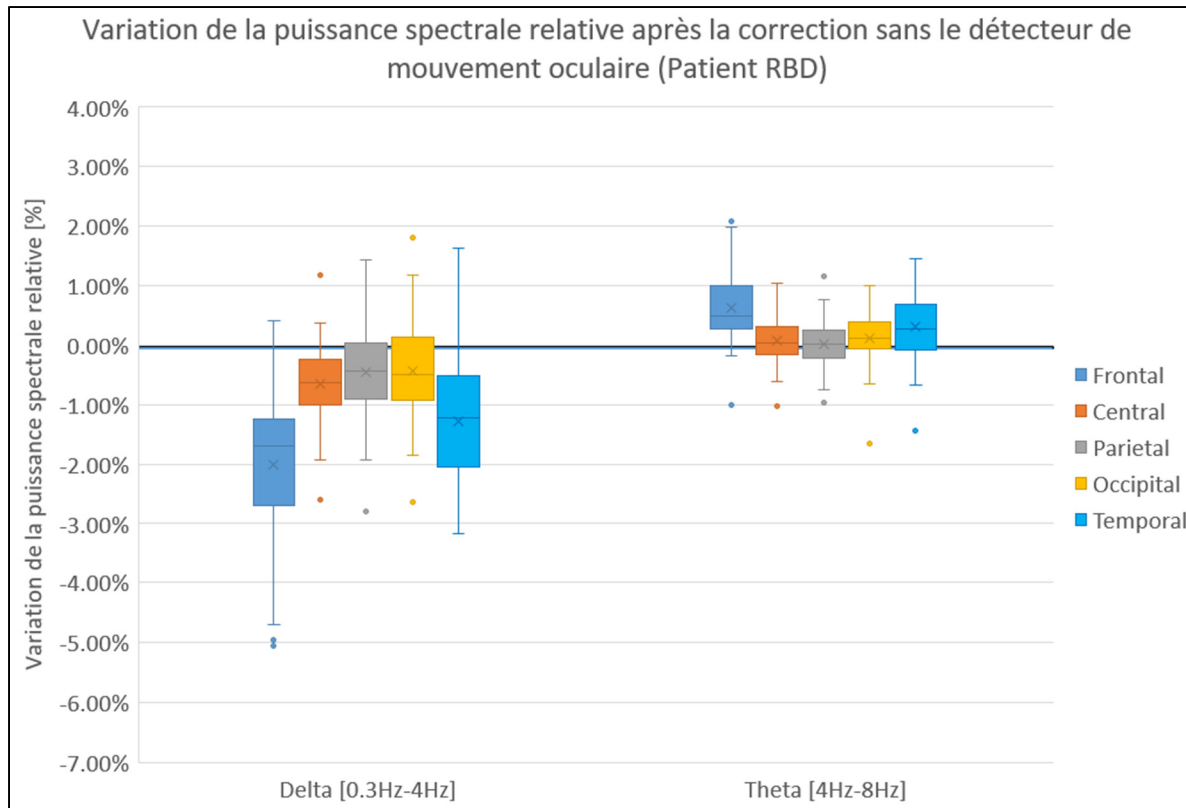


Figure 1.16 Différences des puissances spectrales relatives par région pour les patients RBD

1.5 Discussion

La Figure 1.15 et la Figure 1.16 montrent la même tendance en termes de variation relative de la puissance spectrale par région pour les 50 sujets MASS et les 55 patients RBD, ce qui représente un total de 105 nuits. Comme mentionné à la section précédente, la variation devrait être négative dans la bande de fréquence delta puisque le mouvement oculaire se trouve dans cette bande spectrale, ce qui est majoritairement le cas des boîtes à moustaches. Pour ce qui est des patients RBD, les moustaches supérieures en delta sont légèrement plus élevées que les sujets MASS, mais cette différence peut être attribuable à la pathologie des RBD. En thêta, les modifications apportées sont visiblement de moins grande amplitude, ce qui appuie la correction parcimonieuse recherchée par cet outil en ne retirant qu'une seule composante. Il est important de noter que la correction apportée pour MASS et les RBD en thêta est pratiquement identique.

La prochaine étape consiste à ajouter un détecteur de mouvement oculaire pour contraindre la correction de l'algorithme précédent à la présence effective d'un mouvement oculaire. L'hypothèse est que les boîtes à moustache en delta se concentreraient plus dans les variations de puissance spectrale négative puisque seuls les 'vrais' MORs seraient alors corrigés en amplifiant ainsi la variation dans cette bande de fréquence tandis qu'en thêta, les variations demeureraient sensiblement identiques.

Finalement, rappelons que pour utiliser cet outil, il faut disposer de $n = 8$ électrodes EEG au minimum pour pouvoir utiliser l'ICA. En modifiant le nombre d'électrodes, s'assurer de rester dans le nombre d'échantillons requis entre $5n^2$ et $30n^2$. L'ICA est un algorithme 'glouton', donc l'ordre des composantes est aléatoire, mais la reconstruction du signal demeure la même malgré tout.

CHAPITRE 2

DÉTECTEUR DE MOUVEMENT OCULAIRE RAPIDE

La détection de mouvement oculaire est présentée à part puisque cet outil peut être utilisé indépendamment de la correction des mouvements oculaires sur l'EEG. L'objectif est de cibler les mouvements oculaires dans les périodes de REM afin d'être plus précis dans la correction des artéfacts oculaires sur l'EEG de REM.

2.1 Revue de littérature

Il existe différents types de mouvement oculaire. Les clignements d'yeux présents à l'éveil puis les mouvements oculaires lents ou rapides dépendamment de la rapidité du mouvement. Ce détecteur a comme objectif de détecter les mouvements oculaires rapides qui nous seront utiles afin d'améliorer la précision du correcteur de mouvement oculaire sur l'EEG de REM. Plusieurs études ont abordé le sujet afin de trouver un outil de détection automatique des MORs. La détection se fait encore dans plusieurs endroits par des experts à la main, mais cela n'est pas viable sur de larges cohortes. Divers algorithmes ont été développés afin de résoudre ce problème. (Yetton et al., 2016) a revu l'ensemble des techniques développées ainsi que leur propre technique pour la détection automatique des MORs dans le sommeil paradoxal. La Figure 2.1 présente les différentes performances des outils disponibles. L'approche par classification proposée par (Yetton et al., 2016) ainsi que les algorithmes combinatoires présentent les meilleurs résultats pour les outils automatiques.

L'objectif de l'outil est avant tout de détecter les MORs pour des patients RBD ce qui n'a pas été réalisé précédemment. De plus, cet outil sera testé sur une plus grande base de données (55 RBD avec en moyenne 88 minutes de REM) que les algorithmes précédemment créés tels que 5 sujets pour l'approche par classification, par critère limite et pour les deux approches par algorithme combinatoire (4 utilisés pour entraîner et 1 pour tester) (Yetton et al., 2016), 8 sujets pour la méthode détectant les débuts de MORs (Soler et al., 2021), 10 sujets pour la méthode utilisant l'amplitude, la corrélation et les angles de déflexion (Agarwal et al., 2005), 18 sujets

pour la méthode utilisant l'amplitude et la synchronie entre les EOGs (Hatzilabrou et al., 1994) et 24 sujets pour la méthode utilisant les angles de déflexion (Doman et al., 1995). Une nouvelle méthode sera présentée en utilisant la dérivée de la polynomiale du signal dEOG au lieu d'utiliser les EOGs séparément, le produit négatif des deux EOGs (Agarwal et al., 2005) ou une moyenne mobile du signal dEOG (Soler et al., 2021). Il est important de prendre en compte que les meilleurs résultats des détecteurs existants ont été testés sur de petites bases de données.

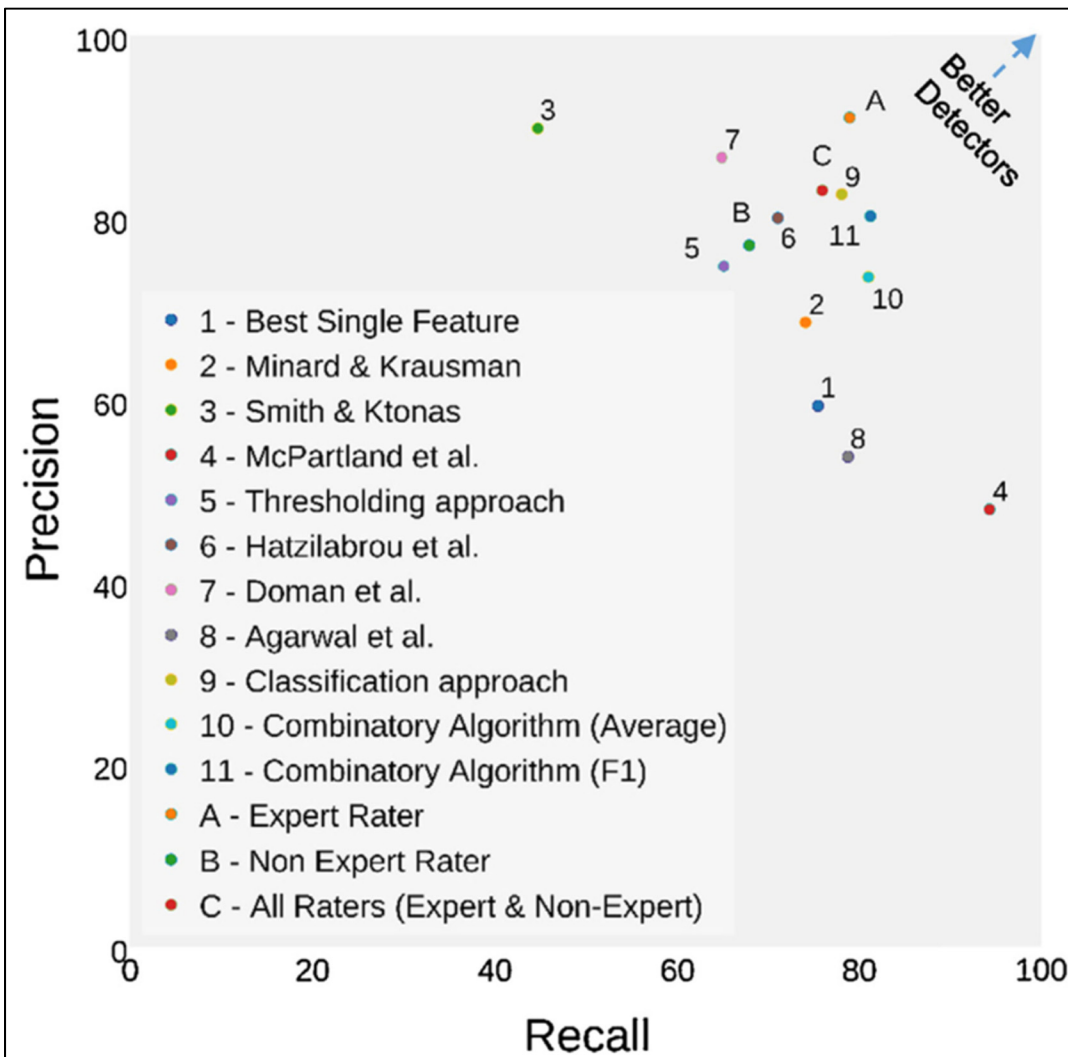


Figure 2.1 Performance des détecteurs de MORs disponible
Tiré de Yetton et al. (2016, p. 24)

2.2 Méthode

2.2.1 Base de données

Pour ce chapitre, la détection des MORs a seulement été effectuée sur les 55 patients RBD puisque les deux experts devaient réaliser la détection pour leur recherche sur les RBDs et chaque nuit prend plusieurs heures à annoter. Durant les stades de sommeil paradoxal, chaque époque de 3 secondes a été identifiée comme contenant au moins un MOR ou non. Le kappa moyen est de 0.82 avec un écart-type de 0.08.

2.2.2 Soustraction des électrodes d'EOG

Afin d'amplifier le mouvement oculaire ainsi que de réduire les éléments de l'EOG qui ne sont pas associés à ce type de mouvement, un EOG a été soustrait à l'autre tel qu'expliqué à la section 1.3.1.

2.2.3 Subdivision des époques

L'objectif principal de cet outil est d'ajouter de la spécificité au correcteur du mouvement oculaire et pour ce faire, il faut détecter la présence d'un MOR dans les époques de 10 secondes. Cette détection doit être plus précise que la densité de mouvement oculaire. Pour cette raison, même si les détecteurs présentés utilisent des époques d'une seconde pour identifier les MORs, des époques de 3 secondes ont été choisies afin de garder une certaine précision temporelle sans pour autant être d'une grande précision dans la détection. Ceci permet d'accélérer le processus et de réduire la tâche d'annotation des experts.

2.2.4 Approximation polynomiale et dérivée instantanée

Pour s'assurer que le signal dEOG contient uniquement le mouvement oculaire, l'approximation polynomiale du signal sera calculée afin d'en retirer les variations de hautes fréquences. Cette approximation peut être vue comme un filtre passe-bas sur le signal dEOG.

Les MOR se manifestent par une transition rapide, donc les pentes les plus élevées de chaque époque sont plus susceptibles de contenir un MOR. Les équations (2.1) et (2.2) permettent de calculer l'approximation polynomiale ainsi que sa dérivée et les équations (2.3) à (2.5) permettent de trouver les coefficients de l'approximation polynomiale.

$$p(t) = c(K) + \sum_{k=1}^{K-1} c(k) * t^{K-k} \quad (2.1)$$

$$q(t) = \sum_{k=1}^{K-1} (K - k) * c(k) * t^{K-k-1} \quad (2.2)$$

$$A(i, j) = \sum t^{2*K-i-j} \quad (2.3)$$

$$B(i) = \sum S * t^{K-i} \quad (2.4)$$

$$c = A^{-1} * B \quad (2.5)$$

Où $p(t)$ sera l'approximation polynomiale du signal dEOG pour une époque de 3 secondes, $q(t)$ sera la dérivée à chaque instant de cette même époque et $k \in \{1, 2, \dots, K - 1\}$. $K-1$ représente l'ordre du polynôme, t est le vecteur temps, i et j vont de 1 à K et S correspond au signal.

2.2.5 Sélection par percentile

Les valeurs absolues maximales de tous les $q(t)$ à travers les époques sont regroupées et classées en ordre croissant de dérivée maximale. Pour déterminer quelles époques seront identifiées comme comportant un mouvement oculaire, un percentile sera appliqué pour définir si les époques contiennent au moins un MOR. La Figure 2.2 présente l'impact du percentile sur le F1 score pour les 55 patients RBD.

$$F_1 = 2 * \frac{\text{précision} * \text{sensibilité}}{\text{précision} + \text{sensibilité}} \quad (2.6)$$

$$\text{Précision} = \frac{\text{Vrai positif}}{\text{Vrai positif} + \text{Faux positif}} \quad (2.7)$$

$$\text{Sensibilité} = \frac{\text{Vrai positif}}{\text{Vrai positif} + \text{Faux négatif}} \quad (2.8)$$

Où les vrais positifs sont les évènements comportant un MOR dans le détecteur ainsi que l'annotation des experts. Les faux positifs sont ceux identifiés comme comportant un MOR par le détecteur, mais non par les experts. Les faux négatifs sont les évènements qui ne sont pas identifiés comme comportant un MOR par le détecteur, mais le sont par les experts.

La majorité des patients ont la même tendance sans pour autant avoir leur maximum au même percentile. Ceci complique la sélection d'un percentile automatique par patient. Jusqu'à présent, aucun paramètre ou biomarqueur n'a permis de déterminer un percentile optimal par nuit. Cependant, il est important d'observer la distribution de la précision et la sensibilité des résultats afin d'identifier les performances du détecteur.

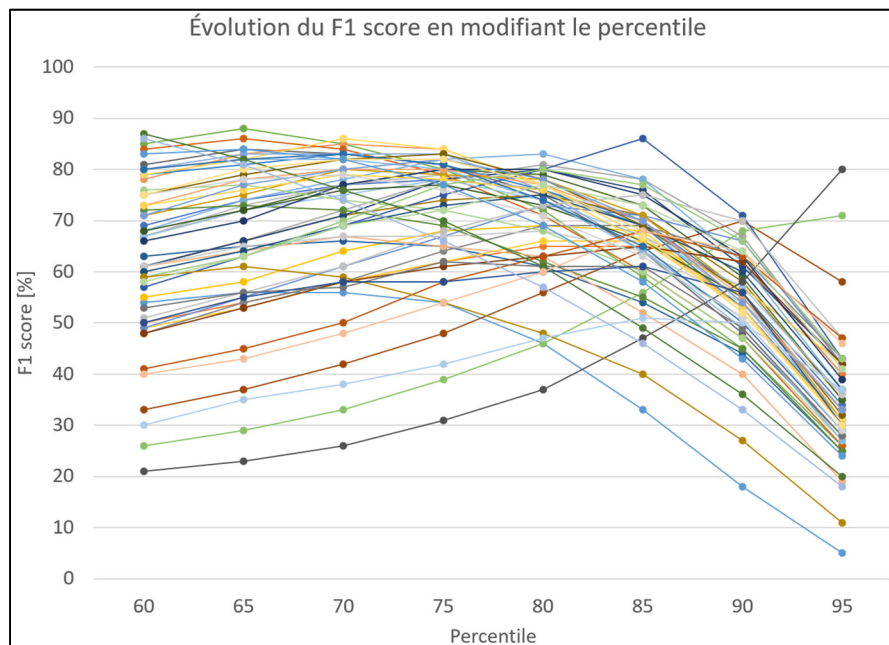


Figure 2.2 F1 score en modifiant le percentile des 55 patients RBD

La Figure 2.3 présente la précision et la sensibilité du détecteur de MOR pour les RBD comparativement aux meilleurs algorithmes présentés par (Yetton et al., 2016). Les performances pour les percentiles 65 et 85 sont présentées pour les 55 patients. En modifiant le percentile, il est possible de ‘naviguer’ entre précision et sensibilité. Comme l’on peut voir, plus le percentile est bas, plus le détecteur est précis et plus le percentile est haut, plus la sensibilité est élevée. Dans la recherche d’un meilleur compromis, le percentile 65 a été choisi pour les raisons suivantes :

- 1) les données sont moins dispersées;
- 2) des résultats précis à défaut d’être sensibles permettent aux experts de faire une première passe très rapide et de se concentrer sur des MORs qui n’ont pas été détectés en premier lieu.

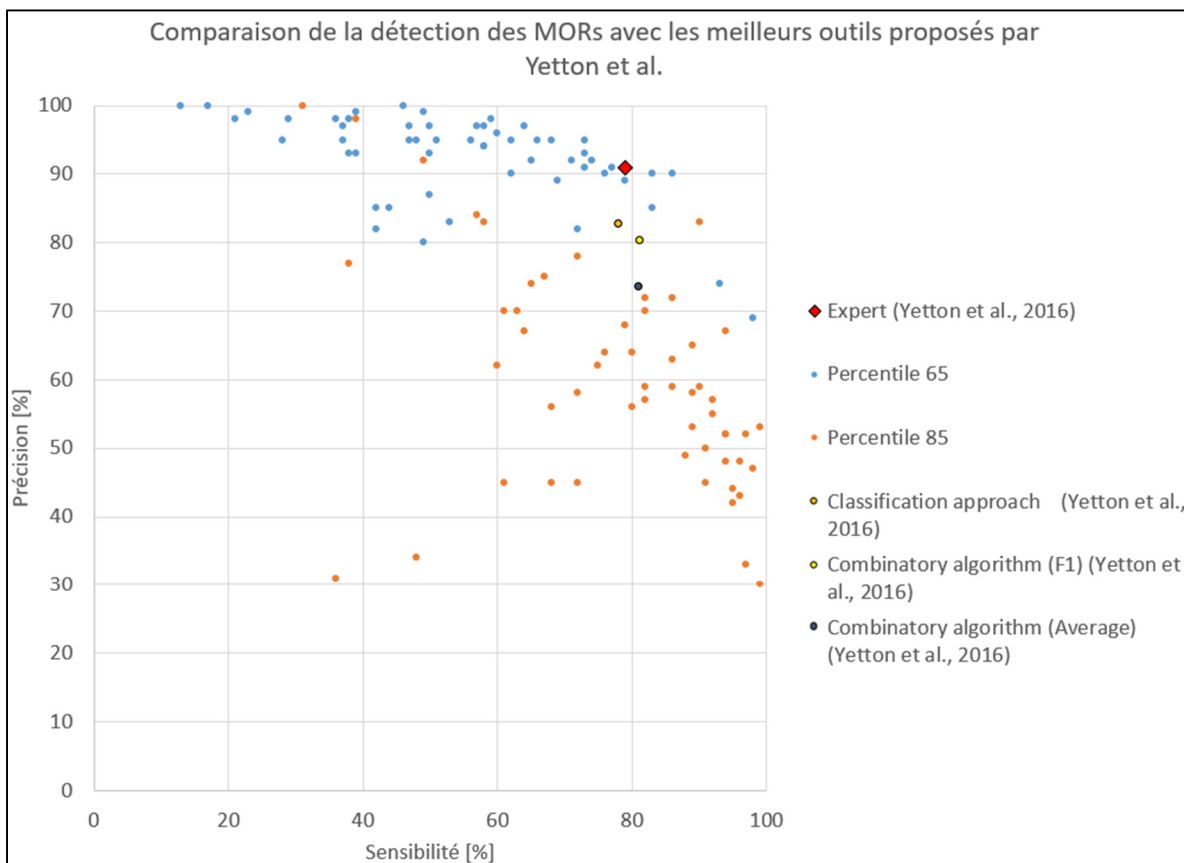


Figure 2.3 Distribution de la précision en fonction de la sensibilité du détecteur

2.3 Résultats

Les résultats de la correction du REM ont été obtenus en ajoutant le détecteur de mouvement oculaire avec 65 percentiles. Les époques de 10 secondes contenant au moins une composante au-dessus du critère devaient aussi contenir au moins une époque de 3 secondes avec un MOR détecté. Dans le cas contraire, aucune correction n'était effectuée sur l'époque. Cela devrait 'condenser' les variations de puissance relative en delta puisque les corrections seraient spécifiquement sur un plus petit nombre d'époques. La Figure 2.4 et la Figure 2.5 présentent la variation de puissance relative dans la bande de fréquence delta et thêta pour les 50 sujets MASS et 55 patients RBD respectivement. En termes de pourcentage de correction des nuits avec le correcteur, en moyenne 31.5% du REM par nuit a été corrigé avec un écart-type de 2.6% pour les sujets MASS et en moyenne 31.7% du REM par nuit avec un écart-type de 2.4% pour les patients RBD ce qui représente un peu plus de la moitié moins d'époques corrigées que les résultats sans détecteur.

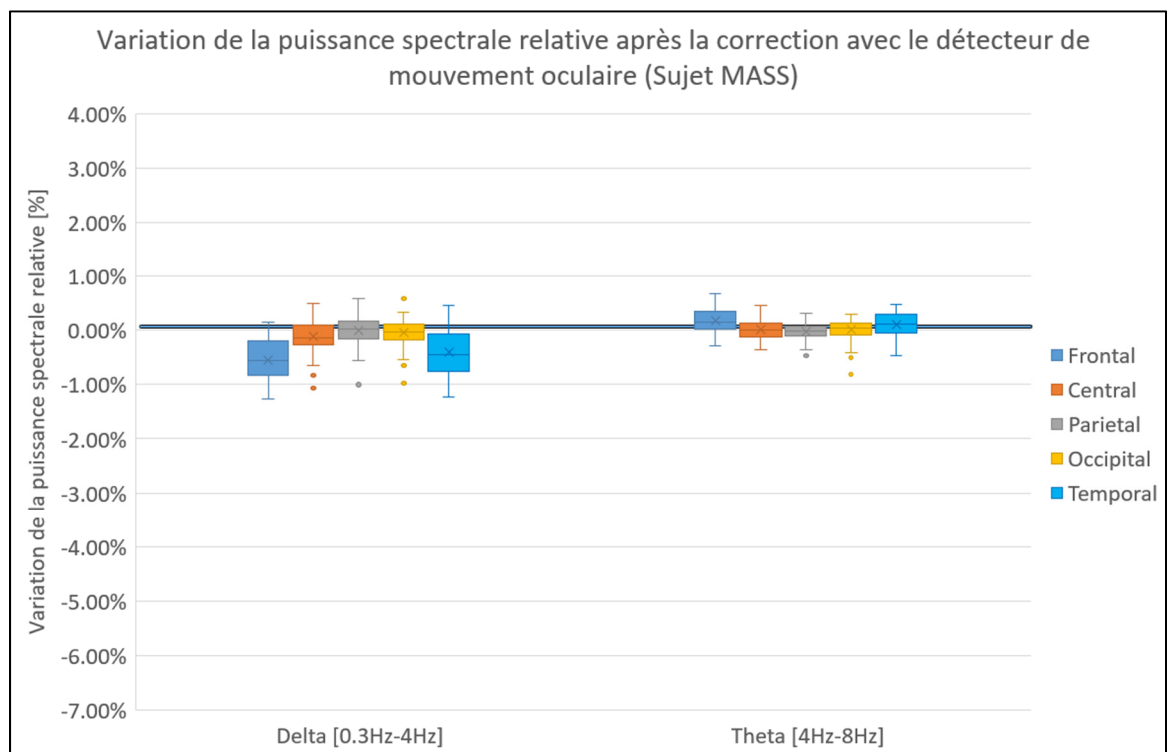


Figure 2.4 Distribution des variations de puissance spectrale relative par région pour les sujets MASS avec le détecteur

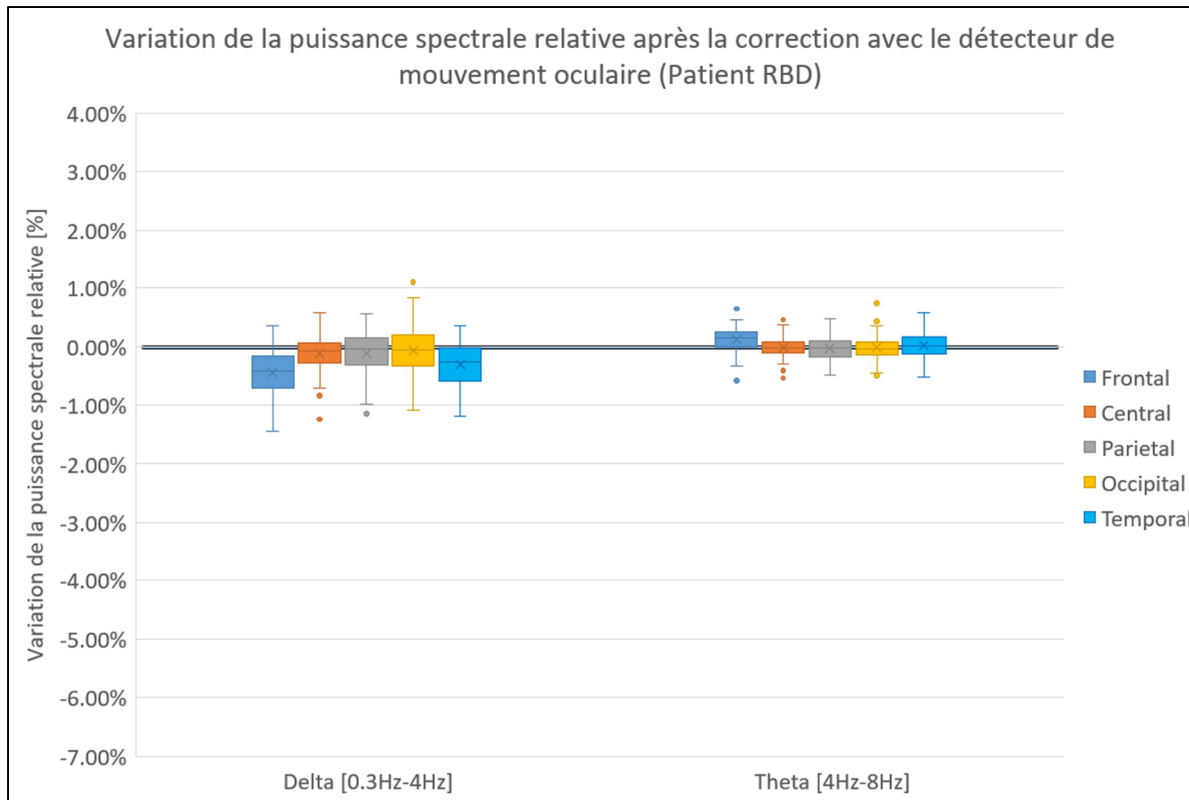


Figure 2.5 Distribution des variations de puissance spectrale relative par région pour les patients RBD avec le détecteur

2.4 Discussion

La performance du détecteur est donc variable dépendamment du percentile, mais il est possible de rendre ce détecteur semi-automatique afin d'aider les experts à annoter les MORs. Cet outil prend en moyenne une trentaine de secondes de calcul pour une nuit avec en moyenne 88 minutes de REM, donc ceci est beaucoup plus rapide qu'une annotation visuelle et permet ainsi de soi de faire une première passe ou de sélectionner des données de façon à rester spécifique. Le percentile 65 a été choisi pour avoir une plus grande précision, mais il est possible d'utiliser celui qui répond davantage aux besoins puisque ce détecteur est rapide et donc propice à quelques essais pour l'adapter aux données.

En comparant la Figure 1.15 et la Figure 2.4 pour les sujets MASS et la Figure 1.16 et la Figure 2.5 pour les patients RBD, on remarque qu'avec et sans RBD, les nuits ont à peu près les

mêmes performances. Le détecteur a grandement réduit la variation de la puissance spectrale en concentrant les valeurs, mais en gardant une variation plus importante en frontale et temporelle. Cela signifie dépendamment du résultat désirer, il est possible d'avoir un correcteur ne réduisant pas l'activité cérébrale en étant plus ou moins spécifique en ajoutant ou retirant ce détecteur.

Cet outil permet donc de faire une rapide détection pour faciliter la tâche des experts et accélérer leur travail. Il est surtout utile en complément du correcteur pour être plus spécifique sur les époques corrigées en réduisant de plus de la moitié les époques notées comme comprenant au moins un MOR. Les paramètres pour cet outil sont présentés dans l'Annexe I, Tableau-A I-2.

CHAPITRE 3

ÉTUDE DES MICRO-ÉTATS

3.1 Revue de littérature

Un signal EEG comportant plusieurs électrodes peut être représenté comme une topographie variant à chaque instant temporel. Les micro-états sont des topographies de potentiels EEG quasi stables pendant une centaine de millisecondes (Koenig et al., 2002). Introduites à la fin des années 90, ces topographies ‘dominantes’ en termes d’énergie déployées sur le scalp sont devenues un des outils d’analyse de l’EEG parmi les plus intéressants. Une des conjectures importantes suggère que ces topographies reflètent sur le scalp l’occurrence transitoire de réseaux fonctionnels au niveau des sources neurales. Les maxima locaux *Global Field Power* (GFP) permettent à des algorithmes de segmentation d’identifier les micro-états et correspondent à la variance topographique à chaque instant se définissant comme suit :

$$\text{GFP}_m = \sqrt{\frac{\sum_{n=1}^N (S_{n,m} - \overline{S}_m)^2}{N}} \quad (3.1)$$

Où N est le nombre d’électrodes, $S_{n,m}$ est la valeur du m^{e} échantillon à la n^{e} électrode et \overline{S}_m est la moyenne de toutes les électrodes au m^{e} échantillon où $m \in \{1, 2, \dots, 256 \text{ Hz} * \text{Durée cycle}\}$. Les maxima locaux GFP sont utilisés pour identifier les micro-états plutôt que les valeurs d’EEG puisque les micro-états demeurent relativement constants entre les minimums de GFP (Khanna et al., 2014). Il est important de souligner que seule la distribution des énergies déployées sur le scalp est observée plutôt que l’intensité ou la polarité des différentes topographies, i.e. des topographies ayant une même distribution, mais des intensités ou des polarités différentes seront associées au même micro-état.

Les micro-états sont étudiés dans plusieurs sujets afin d’identifier des marqueurs potentiels pour les différentes pathologies. Cela inclut l’Alzheimer (Tait et al., 2020), le Parkinson (Chu et al., 2020), la dépression (Murphy et al., 2020), la maladie à corps de Lewy (Schumacher et

al., 2019), la schizophrénie (Lehmann et al., 2005), la sclérose en plaques (Gschwind et al., 2016), l'apnée obstructive du sommeil (Xiong et al., 2021) et même les patients RBD (Peng et al., 2021). Des études ont aussi été réalisées afin d'observer la variation des micro-états selon l'âge chez les adultes (Jabès et al., 2021; Koenig et al., 2002) et les nouveau-nés (Khazaei et al., 2021) ainsi que les différentes techniques de calcul et de sélection des micro-états (Khanna et al., 2014; Michel & Koenig, 2018; Pascual-Marqui et al., 1995). La plupart de ces études ont été réalisées à l'état de repos excepté quelques-unes qui ont regardé les topographies à l'éveil, N1, N2, N3 (Michel & Koenig, 2018) ainsi que le REM (Khazaei et al., 2021; Xiong et al., 2021). Toutefois, aucune étude n'a porté spécifiquement sur l'extraction des micro-états en REM chez l'adulte qui constitue un stade important du développement de diverses pathologies. Cela sera donc le sujet de ce chapitre et puisque l'inventaire de ces topographies nécessite des données REM sans trace du mouvement oculaire sur l'EEG, les outils présentés aux chapitres 1 et 2 seront utilisés pour corriger l'EEG. Dans la majorité des études réalisées sur des sujets sains, 4 micro-états sont identifiés : A, B, C et D. Dans d'autres articles, il est question d'un cinquième micro-état E (Murphy et al., 2020; Pascual-Marqui et al., 1995; Schumacher et al., 2019; Xiong et al., 2021) et chez les nouveau-nés, le nombre serait de sept (A, B, C, D, E, F, G) (Khazaei et al., 2021). Chez les patients avec diverses pathologies, les micro-états peuvent varier ainsi que lorsqu'on regarde des stades spécifiques.

Deux éléments sont à choisir pour segmenter les micro-états : l'algorithme permettant de regrouper les maxima locaux GFP en classe (les micro-états) ainsi que la mesure permettant de sélectionner le nombre optimal de micro-états. Quatre algorithmes de segmentation sont généralement utilisés : K-means, K-means modifié, *The Topographic Atomize and Agglomerate Hierarchical Clustering* (TAAHC) et *Atomize and Agglomerate Hierarchical Clustering* (AAHC). Une des mesures suivantes permet de déterminer le bon nombre de micro-états: la *global explained variance* (GEV), le *cross-validation criterion* (CV), la dispersion et le critère de Krzanowski-Lai. Chaque élément sera expliqué en détail dans la prochaine section.

Il existe plusieurs outils permettant d'extraire les micro-états. Ce chapitre consistera à définir la méthode la plus appropriée au contexte du REM. Le logiciel CARTOOL (Brunet et al.,

2011) est parmi les plus utilisés. Les deux méthodes de segmentation disponibles sont K-means modifié et le AAHC. Les deux critères offerts permettant de sélectionner le bon nombre de micro-états sont le CV et le critère de Krzanowski-Lai. Une boîte à outils nommée Microstate-EEGlab-toolbox (Poulsen et al., 2018) est disponible sur MATLAB et offre comme algorithme le K-means, le K-means modifié, le TAAHC ainsi que le AAHC. Les critères de sélection sont le GEV, le CV, la dispersion, le critère de Krzanowski-Lai et le critère de Krzanowski-Lai normalisé. *The Chicago Electrical Neuroimaging Analytics* ont aussi un plugiciel disponible dans brainstorm sur MATLAB (Cacioppo et al., 2014) offrant les algorithmes utilisés dans les autres outils ainsi que leur propre méthode de segmentation. Les critères disponibles sont le CV et le critère de Krzanowski-Lai. Une méthode en python a aussi été développée pour la méthode k-means (von Wegner & Laufs, 2018). L'objectif de ce chapitre est d'identifier la méthode optimale pour segmenter les micro-états en REM. Cette analyse sera réalisée dans le Microstate-EEGlab-toolbox implémenté sur MATLAB. Les paramètres utilisés sont décrits dans l'Annexe I, Tableau-A I-3.

3.2 Méthode

3.2.1 Prétraitement

Les micro-états en REM seront évalués sur les 50 sujets MASS ayant chacun 17 électrodes tel que présenté à la section 1.2. Le correcteur de REM vu au chapitre 1 ainsi que le détecteur du chapitre 2 ont été utilisés afin de nettoyer les signaux. Ceux-ci ont ensuite été regroupés par cycle de REM. Il y a en moyenne entre 4 et 6 cycles dans une nuit de sommeil (Suni & Vyas, 2022). Un cycle de REM dure en moyenne entre 10 et 60 minutes. La Figure 3.1 montre 3 cycles de REM lors d'une nuit qui sont encadrés en rouge.

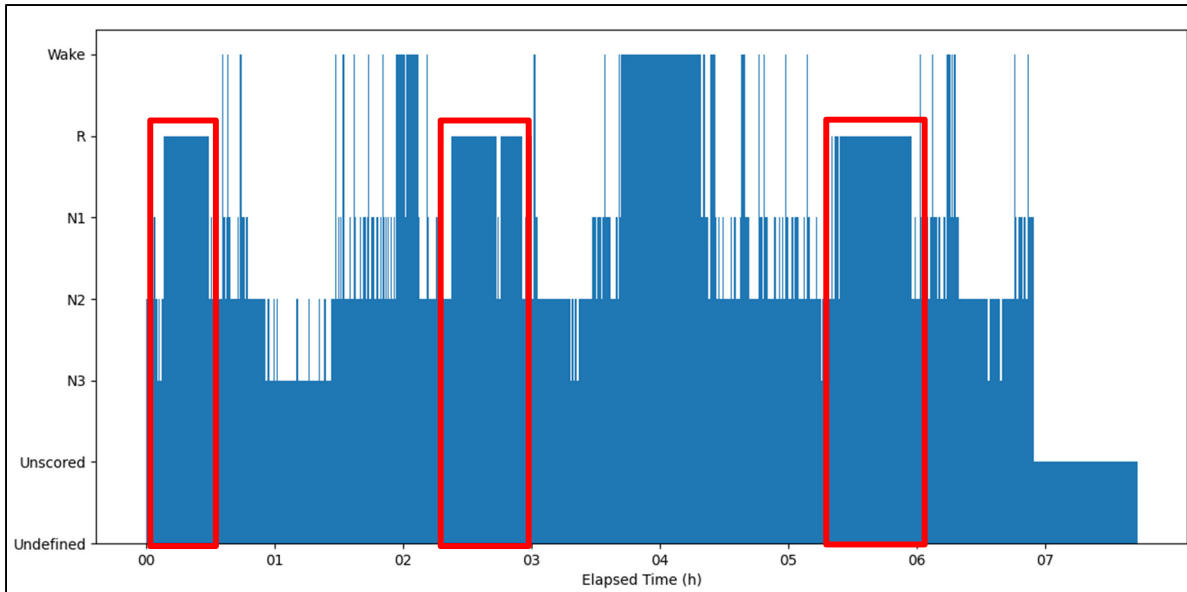


Figure 3.1 Hypnogramme d'une nuit avec identification des cycles de REM

Une fois les signaux regroupés par cycle, un filtre passe-bande entre 1Hz et 30Hz permet de retirer les artéfacts se retrouvant aux hautes fréquences. Ce filtrage est communément utilisé dans les outils de détection des micro-états. De plus, la moyenne est retirée du signal et chaque électrode est normalisée grâce à la formule suivante :

$$S_{n,m,normalisé} = \frac{S_{n,m} - \bar{S}}{\sigma_m(S)} \quad (3.2)$$

Où S est le signal d'un cycle REM [N électrodes \times M échantillons], \bar{S} est la moyenne du signal et $\sigma_m(S)$ est la moyenne de l'écart-type du m^e échantillon sur toutes les électrodes du signal S . Les maxima locaux GFP sont ensuite calculés tel que présentés précédemment à l'équation (3.1). Un maximum de 10 000 maxima locaux sera choisi aux hasards afin de constituer notre nouvel ensemble de données pour la segmentation ($GFP_m \rightarrow GFP_q$ où $q \in \{1, 2, \dots, 10\ 000\}$). Il est inutile de garder tous les maxima locaux puisque 17 minutes de REM peut représenter jusqu'à 25 000 maxima locaux GFP ce qui ralentira l'analyse sans apporter de gain considérable dans les résultats.

3.2.2 Algorithmes de segmentation

L'algorithme de K-means et K-means modifié sont pratiquement identiques. Ils consistent à initialement créer K classe où K est le nombre de micro-états souhaités. Les différentes valeurs sont les maxima locaux GFP sélectionnés. Chaque point est associé à une classe puis la moyenne de cette classe est calculée et le centre de celle-ci est déplacé avant de réassigner les points (équation (3.3)) et ainsi de suite jusqu'à atteindre un critère de convergence (équation (3.4)). Un procédé avec quatre itérations est présenté à la Figure 3.2. En pratique, le nombre d'itérations est de l'ordre de 1000 dans notre cas ou jusqu'au critère de convergence de l'algorithme de l'ordre de 1e-6 dans notre cas (les valeurs peuvent être définies par l'utilisateur, mais celles-ci sont les valeurs par défaut).

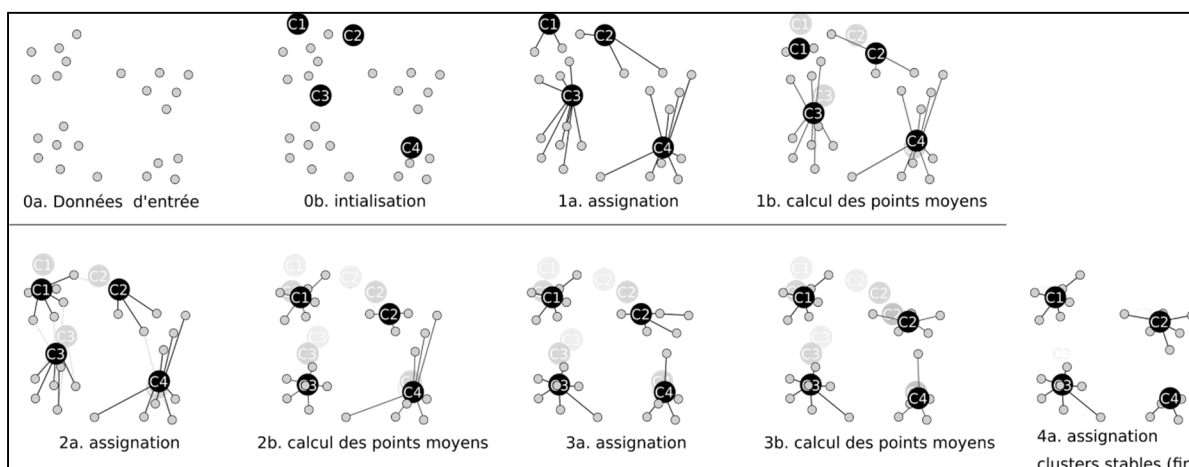


Figure 3.2 Schématisation de l'algorithme K-means

Tiré de Quantin (2017)

L'assignation se fait grâce à la plus petite distance euclidienne au carré (Poulsen et al., 2018) :

$$d_{k,m}^2 = S_m^T * S_m - (S_m^T * a_k)^2 \quad (3.3)$$

Où S_m est le signal EEG du m^e échantillon et a_k correspond au centre du k^e micro-état.

Le critère de convergence se calcule grâce à la fonction de perte :

$$E = \sum_{m=1}^M \left\| S_m - \sum_{k=1}^K a_k * z_{k,m} \right\|^2 \quad (3.4)$$

Où $z_{k,m}$ est une matrice identifiant le micro-état actif pour le m^e échantillon. Il y a deux différences entre le K-means et le K-means modifié. Il a déjà été mentionné que les micro-états ayant une même distribution d'énergie, mais avec des topographies différentes devraient être associés au même micro-état. La première différence est donc que pour les échantillons de même topographie, mais ayant des polarités différentes, le K-means modifié les associera à la même classe ce qui n'est pas le cas du K-means. L'algorithme de segmentation de K-means modifié permet donc de présenter des micro-états respectant ce critère. La seconde différence est que le nouvel algorithme vient pondérer l'activation des micro-états à chaque instant temporel par l'amplitude. Dans d'autres mots, un échantillon temporel activera davantage un micro-état si celui-ci est associé à un signal EEG fort que s'il est associé à un signal EEG faible (Poulsen et al., 2018). Cela permet donc de favoriser des micro-états ayant des membres ayant une plus grande contribution ce qui n'est pas le cas du K-means.

La version optimisée de segmentation a été utilisée pour le K-means modifié. Cela consiste à modifier la fonction de perte (3.4) par (3.5), ce qui a permis d'améliorer la vitesse de convergence (Poulsen et al., 2018).

$$E = \sum_{m=1}^M (\|S_m\|^2 - (a_k^T * S_m)^2) \quad (3.5)$$

Les algorithmes AAHC et TAAHC sont très similaires. Ils font partie des algorithmes de classification hiérarchique. L'algorithme débute par assigner à chaque échantillon temporel d'EEG une classe (de a à f sur la Figure 3.3). À chaque itération, la 'pire' classe ou micro-état est supprimé et les membres de cette classe sont redistribués à la classe avec le plus de similarité (Poulsen et al., 2018) jusqu'à atteindre le nombre de micro-états définis par

l'utilisateur (entre 3 et 8 micro-états dans notre cas avant que le critère de sélection du nombre de micro-états à la section 3.2.3 ne définisse le nombre de micro-état optimal). La différence entre les deux algorithmes est la définition de la 'pire' classe. Le TAAHC est se sert de la plus petite somme de corrélation tandis que le AAHC se sert de la somme des GEV.

$$\text{CorrS} \sum_k = \sum_q \frac{|S_q * a_k|}{\|S_q\| * \|a_k\|} \Big|_{q \in \{l_k\}} \quad (3.6)$$

$$\text{GEV}_k = \sum_q \text{GEV}_q \Big|_{q \in \{l_k\}} \quad (3.7)$$

Où k correspond au micro-état, q correspond à un échantillon appartenant à k , S_q correspond au signal EEG du q^{e} maxima locaux GFP, a_k correspond au centre du k^{e} micro-état, l_k correspond aux indices des topographies de q appartenant à la k^{e} classe et GEV_q correspond à la GEV du q^{e} maxima locaux GFP.

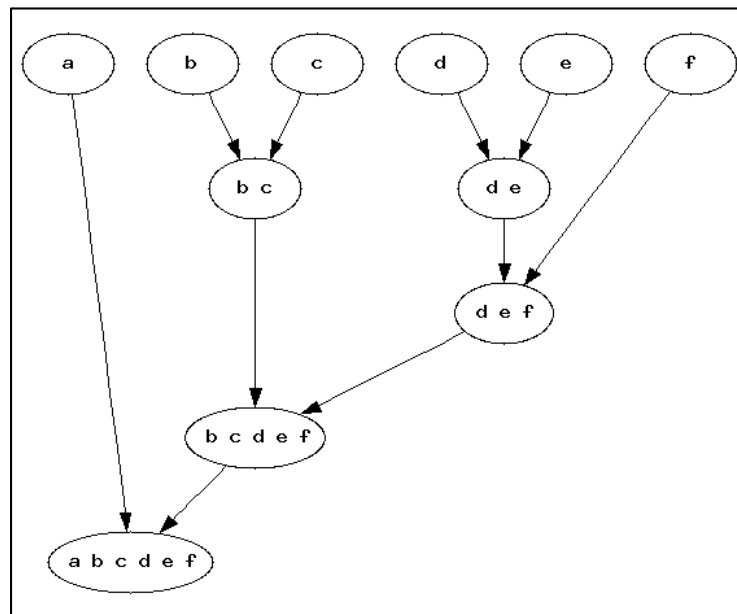


Figure 3.3 Schématisation d'un algorithme de classification hiérarchique
Tiré de Barnard (2004)

3.2.3 Critère de sélection du nombre de micro-états

Il existe 4 critères souvent utilisés afin de choisir le nombre de micro-état optimal. La GEV permet de quantifier la corrélation entre chaque valeur EEG pour un maxima locaux et son micro-état. Cette corrélation est pondérée par la puissance de sa GFP par rapport à la GFP de son micro-état. Plus la valeur de la GEV est élevée, plus le maxima locaux influence son micro-état (Poulsen et al., 2018). Cette métrique ce calcul par maxima locaux (équation (3.8)), par micro-état (équation (3.9)) ou pour l'ensemble des micro-états (équation (3.10)) selon la comparaison désirée. Une valeur de GEV élevé est donc favorable puisque cela signifie que la distribution des maxima locaux GFP ainsi que le choix des micro-états est optimal. Pour $q \in \{l_k\}$,

$$GEV_q = \left(\frac{|a_k * S_q|}{\|a_k\| * \|S_q\|} \right)^2 * \frac{GFP_q^2}{\sum_q GFP_q^2} \quad (3.8)$$

$$GEV_k = \sum_{q \in \{l_k\}} GEV_q \quad (3.9)$$

$$GEV = \sum_k^K GEV_k \quad (3.10)$$

Où S_q correspond au signal EEG du q^e maxima locaux GFP, a_k correspond au centre du k^e micro-état et l_k correspond à l'ensemble des indices des topographies q appartenant à la k^e classe.

Le CV est une mesure permettant d'estimer la variance résiduelle (Poulsen et al., 2018), donc l'objectif est de diminuer cette valeur afin d'avoir la plus petite variance possible pour que chaque maxima locaux GFP soit très près de son micro-état.

Ce paramètre se calcule comme suit :

$$\hat{\sigma}_k^2 = \frac{\sum_{q=1}^{q_{max}} (S_q^T * S_q - (a_{q,k}^T * S_q)^2)}{q_{max}(N-1)} \quad (3.11)$$

$$CV = \hat{\sigma}_k^2 * \left(\frac{N-1}{N-K-1} \right)^2 \quad (3.12)$$

Où $\hat{\sigma}_k^2$ est une mesure permettant de calculer le bruit résiduel, S_q^T correspond au signal EEG transposé du q^e maxima locaux GFP, $a_{q,k}^T$ correspond au centre transposé du k^e micro-état attribué au q^e maxima locaux GFP, q_{max} est le nombre de maxima locaux GFP ($q_{max} = 1000$), N est le nombre d'électrodes et K est le nombre de micro-états.

La dispersion permet de quantifier la variabilité intra classe en comparant les différentes valeurs EEG des membres d'un même micro-état. L'objectif est donc de réduire la dispersion pour que les membres d'un micro-état soient le plus près du centre.

$$D_k = \sum_{q \in \{l_k\}} \sum_{q' \in \{l_k\}} \|S_q - S_{q'}\|^2 \quad (3.13)$$

$$W_K = \sum_k^K \frac{D_k}{2q_{k,max}} \quad (3.14)$$

Où S_q correspond au signal EEG du q^e maxima locaux GFP, l_k correspond aux indices q appartenant à la k^e classe et $q_{k,max}$ est le nombre de maxima locaux pour le k^e micro-état.

Le critère de Krzanowski-Lai permet de trouver le nombre de micro-état optimal en observant l'impact de la dispersion sur l'ajout de classe. Dans un cas extrême, s'il y avait autant de classe que d'échantillon d'EEG, la dispersion serait de 0 puisque la distance serait nulle entre tous les points d'une classe. Le critère de Krzanowski-Lai trouve le nombre de micro-état optimal auquel l'ajout de classe ne changerait plus significativement la dispersion (Poulsen et al., 2018).

L'objectif est d'obtenir la valeur la plus élevée selon l'équation :

$$\text{DIFF}(K) = (K - 1)^{2/N} * W_{K-1} - (K)^{2/N} * W_K \quad (3.15)$$

$$\text{KL}(K) = \left| \frac{\text{DIFF}(K)}{\text{DIFF}(K+1)} \right| \quad (3.16)$$

Où W_K correspond à la dispersion calculée à l'équation (3.14) et N le nombre d'électrodes.

Le critère de Krzanowski-Lai sera utilisé afin de définir le nombre de micro-état optimal. La GEV permettra de comparer quelle méthode est la meilleure afin d'identifier les micro-états.

3.2.4 Assignment d'un micro-état pour chaque échantillon temporel

Comme mentionné à la fin de la section 3.2.1, les calculs des micro-états sont réalisés à partir de 10 000 échantillons temporels d'EEG. Afin d'attribuer un micro-état pour les échantillons n'ayant pas été utilisés dans la segmentation des micro-états, il est possible d'utiliser le *global map dissimilarity* (GMD) permettant de trouver les distances entre les points et les micro-états existants (Poulsen et al., 2018).

$$\text{GMD} = \frac{\left\| \frac{S_m}{\text{GFP}_m} - \frac{S_{m'}}{\text{GFP}_{m'}} \right\|}{\sqrt{N}} \quad (3.17)$$

Où GFP_m est le maxima locaux GFP du m^{e} échantillon et $\text{GFP}_{m'}$ est le maxima locaux GFP d'un autre échantillon faisant parti du même micro-état que l'échantillon m .

3.3 Résultats

Les 50 sujets MASS ont été testés avec chacune des méthodes de segmentation. Pour chaque sujet, le premier cycle de REM a été sélectionné pour une durée moyenne de 13 minutes. Pour chaque sujet, les algorithmes trouvaient les résultats pour 3, 4, 5, 6, 7 et 8 micro-états et le critère de Krzanowski-Lai a permis de définir le nombre optimal de micro-états. La Figure 3.4

présente la distribution des GEV calculés à partir de l'équation (3.10) pour le nombre de micro-état optimal en fonction des quatre différents algorithmes de segmentation. Chaque boîte à moustache englobe les résultats des 50 sujets MASS.

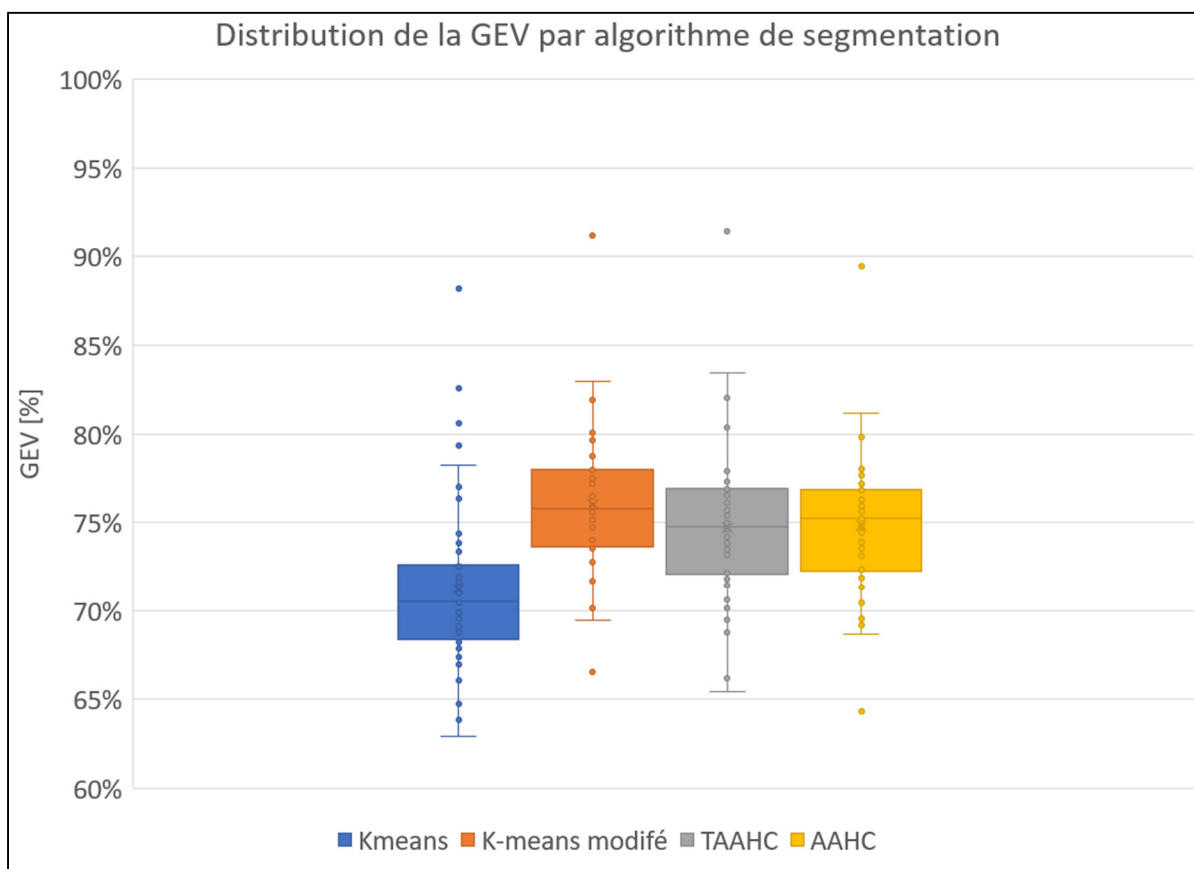


Figure 3.4 GEV obtenu pour chaque algorithme sur les 50 sujets MASS

Il est important de mentionner que les algorithmes TAAHC et AAHC prennent beaucoup plus de temps (de l'ordre de 40 minutes de différence en moyenne) pour trouver les micro-états que les algorithmes K-means pour les segments observés. À titre indicatif, quelques topographies de divers sujets MASS trouvées à partir de l'algorithme K-means modifié sont présentées à la Figure 3.5.

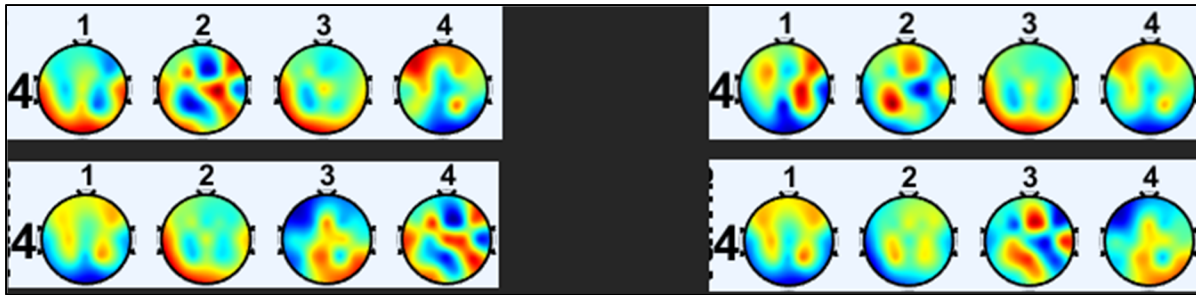


Figure 3.5 Exemple de micro-états trouvés à partir de l'algorithme K-means modifié sur quatre sujets MASS

3.4 Discussion

L'objectif de cette étude est de trouver la meilleure technique afin d'identifier les micro-états en REM. En testant les quatre algorithmes couramment utilisés pour l'identification de ces derniers, le K-means modifié est la méthode ayant obtenue le meilleur GEV de 75.8% comparativement au K-means avec un GEV de 70.5%, au TAAHC avec 74.8% et au AAHC avec 75.2%. Malgré la proximité des résultats entre le K-means modifié et le TAAHC, ce dernier obtient des résultats plus étendus, ce qui nuit à la reproductibilité des résultats. Le temps de calcul est aussi beaucoup plus long pour le TAAHC. L'approche par une classification K-means modifié est donc celle que l'on devrait privilégier. Une implémentation de l'algorithme 'K-means modifié' a été réalisé en python afin d'avoir un outil disponible sur l'application développée au CÉAMS qui sera présentée au chapitre suivant.

CHAPITRE 4

IMPLÉMENTATION NUMÉRIQUE

4.1 Description

L'entièreté de ce mémoire a été réalisée sur l'application Snooz développée au Centre d'études avancées en médecine du sommeil (CÉAMS) par David Lévesque et Karine Lacourse. L'objectif de cette application est de regrouper un ensemble d'outils pour analyser le sommeil et d'offrir aux chercheurs et aux cliniciens la possibilité de faire des études sans avoir à programmer leurs propres outils. La Figure 4.1 montre un exemple d'analyse un peu plus avancée permettant de connecter les divers éléments entre eux en liant les entrées et sorties des différents plugiciels. Ceci est une partie de mon correcteur automatique. Des pipelines peuvent être échangés afin d'envoyer ce regroupement de plugiciels à d'autres chercheurs. L'application est bientôt en distribution gratuite sous licence privée pour le moment, donc il est possible que cette interface change, mais le principe reste le même.

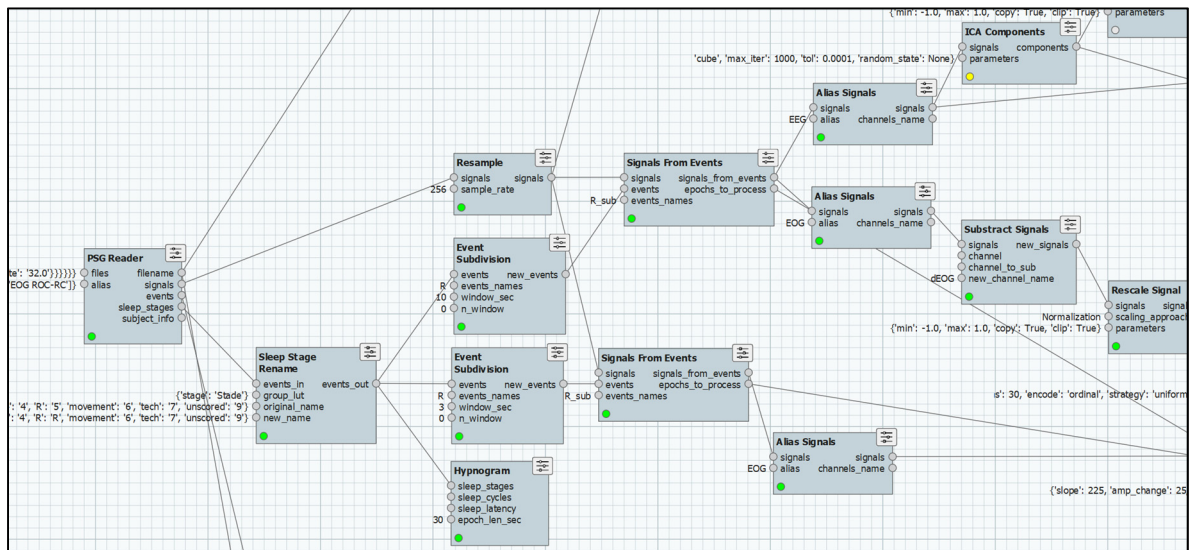


Figure 4.1 Exemple d'interface plus avancé
Tiré de Lévesque et Lacourse (2022)

Pour chaque plugiciel, une interface pour les paramètres et les résultats est disponible afin d'ajuster les différents éléments souhaités. Tous les graphiques présentés dans ce mémoire proviennent de cette application. La Figure 4.2 et la Figure 4.3 sont des exemples des différentes interfaces disponibles pour chaque plugiciel.

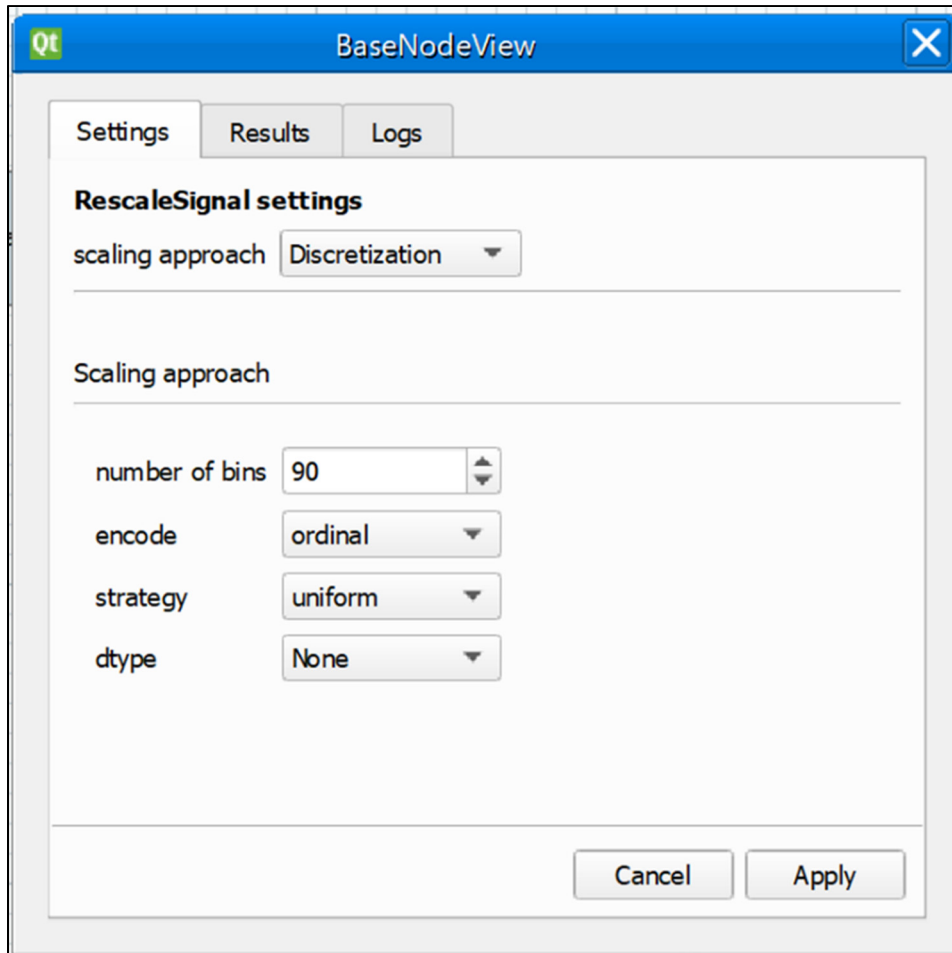


Figure 4.2 Exemple d'interface des paramètres

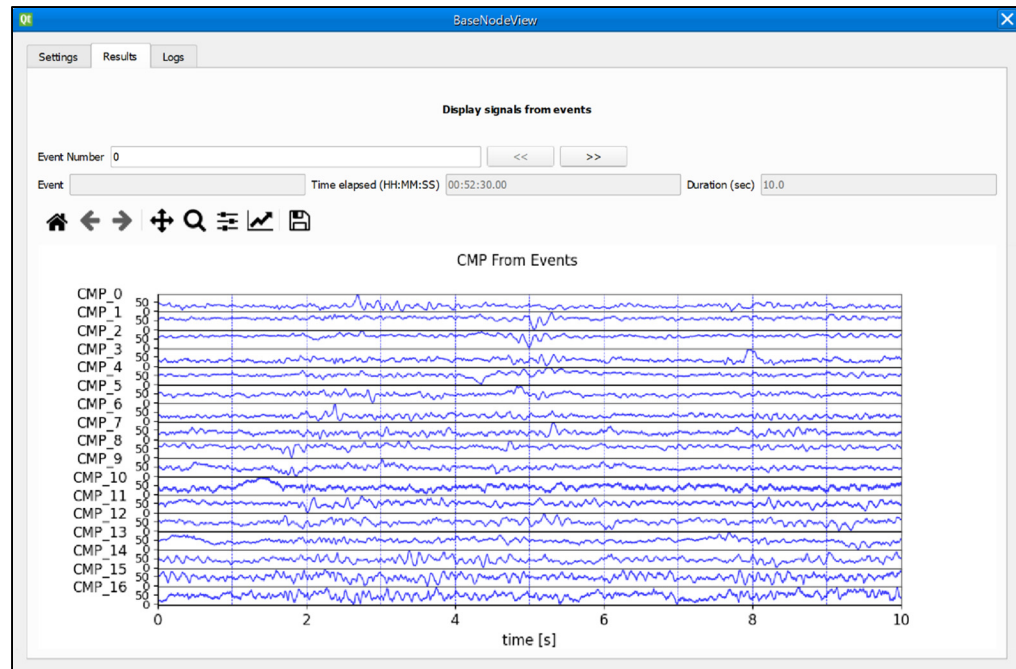


Figure 4.3 Exemple d'interface des résultats

Pour les personnes souhaitant simplement exécuter un procédé rapide sans avoir à connecter les plugiciels, une liste de pipeline préconnectée avec une interface plus familière est aussi disponible. Un exemple d'un détecteur d'artéfacts est démontré à la Figure 4.4.

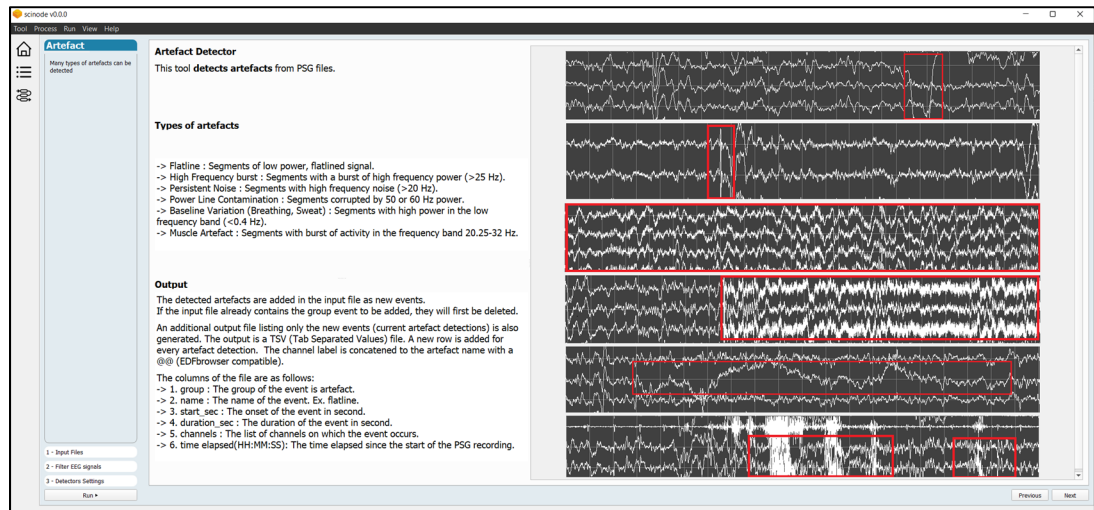


Figure 4.4 Exemple de procédé
Tiré de Lévesque et Lacourse (2022)

Dans le cadre de ce mémoire, je suis le premier programmeur à ajouter des plugiciels pour répondre à mes besoins, mais aussi au besoin des autres puisque chaque plugiciel est réalisé de façon universelle afin de les réutiliser dans d'autres études. Éventuellement, d'autres analystes du sommeil pourront rajouter de leurs outils et ainsi apporter une application facilitant l'analyse du sommeil.

CONCLUSION

L'interférence causée par l'activité oculaire venant se combiner avec l'activité cérébrale sous-jacente empêche l'analyse exhaustive du sommeil paradoxal. En effet, l'étude du REM se fait présentement en sélectionnant certaines époques sans artéfact, ce qui ne permet pas d'observer l'entièreté du signal. Pour résoudre ce problème, plusieurs outils numériques ont été développés afin de résoudre cette problématique.

Un correcteur automatique de mouvement oculaire sur l'EEG a été implémenté afin de nettoyer avec parcimonie le signal. La méthode utilise principalement l'ICA pour décomposer le signal puis calculer l'information mutuelle entre l'EOG et les composantes d'EEG. Les MORs, si présents, sont ainsi contenus dans une composante qui est ensuite retirée. Cette technique permet de nettoyer l'activité oculaire de l'EEG sans retirer l'activité cérébrale. En effet, la différence d'activité dans la bande de fréquence delta, qui est relative à la présence oculaire, a diminué démontrant ainsi une réduction de l'activité oculaire sur les électrodes frontales (la principale région affectée par la présence d'artéfacts oculaires). De plus, la différence d'activité dans la bande de fréquence thêta, fréquences associées à l'activité cérébrale en REM, a augmenté démontrant ainsi un nettoyage de l'EEG de REM. Cet outil permet donc de nettoyer automatiquement en seulement 10 à 15 minutes les périodes de REM d'une nuit de sommeil de 8h.

Afin d'assurer la parcimonie de la correction, un outil de détection des mouvements oculaires rapides a été développé afin de cibler la correction. Il permet aussi de détecter les MORs en REM de patients RBD. La méthode utilise un percentile des dérivées maximales de chaque époque afin de cibler les MORs ce qui permet à l'algorithme d'obtenir un temps de calcul moyen de 34.2 secondes pour une nuit avec en moyenne 88 minutes de REM. Cette méthode semi-automatique offre une détection rapide permettant de varier la précision ou la sensibilité en modifiant le percentile. En combinant ce détecteur avec le correcteur de mouvement oculaire, cela réduit l'amplitude des variations dans la bande de fréquence delta et thêta tout

en gardant une variation négative en delta après la correction ainsi qu'une variation positive en θ .

Les micro-états sont des topographies quasi stables représentant l'activité cérébrale. Ceux-ci ont besoin d'un signal nettoyé afin d'identifier les micro-états ce qu'il est maintenant possible d'obtenir automatiquement grâce au correcteur de mouvement oculaire. Cette étude a permis d'établir la meilleure méthode pour segmenter le signal en micro-états en REM. Quatre algorithmes de segmentation sont utilisés : K-means, K-means modifié, AAHC et TAAHC. Le nombre optimal de micro-état a été sélectionné grâce au critère de Krzanowski-Lai et les différents algorithmes ont été comparés avec la GEV. Le K-means modifié est la méthode ayant obtenue le meilleur GEV de 75.78% comparativement au K-means avec un GEV de 70.51%, au TAAHC avec 74.77% et au AAHC avec 75.24%.

Chaque outil a été développé sur l'application du CÉAMS qui sera disponible pour tous. Ces outils ainsi que plusieurs autres seront offerts afin d'étudier le sommeil et combiner les efforts de chacun afin de faciliter l'étude du sommeil. Les outils présentés dans ce mémoire permettront d'analyser le REM plus en profondeur. Un article portant sur les deux outils permettant la détection et la correction des canaux EEG est en cours de rédaction.

RECOMMANDATIONS

Le correcteur de mouvement oculaire permet de corriger le REM ce qui offre l'opportunité d'étudier le sommeil paradoxal dans son ensemble. Plusieurs domaines tels que l'étude du trouble comportemental en sommeil paradoxal pourront approfondir leur étude avec plus d'époques de REM nettoyées. Pour améliorer le détecteur de mouvement oculaire, un biomarqueur permettant d'identifier le percentile optimal par patient automatiserait la méthode. En ce qui concerne les micro-états, la démarche pour les segmenter a été définie, il ne reste plus qu'à identifier les micro-états en REM. Comme démontré à la Figure 3.5, plusieurs topographies reviennent d'un sujet à l'autre, il faudrait donc déterminer lesquelles sont spécifiques au REM. De plus, celles-ci sont les topographies identifiées durant le premier cycle REM, mais les autres cycles comportent éventuellement leurs propres micro-états associés à des procédés neuronaux réalisés plus tard au cours de la nuit. Ces micro-états permettront de faire progresser plusieurs sujets tels que les RBD, le parkinson ou l'Alzheimer.

ANNEXE I

PARAMÈTRES UTILISÉS DANS LES DIFFÉRENTES FONCTIONS

Tableau-A I-1 Résumé des paramètres utilisés pour le correcteur des mouvements oculaires

<u>Section</u>	<u>Fonction</u>	<u>Nom du paramètre</u>	<u>Valeur</u>
Analyse par composante indépendante	sklearn.decomposition.FastICA (Lafaye de Micheaux et al., 2022)	n_components	None
		algorithm	deflation
		whiten	True
		fun	cube
		fun_args	None
		max_iter	1000
		tol	0.0001
		whiten_solver	svd
		random_state	None
Normalisation	sklearn.preprocessing.MinMaxScaler (Blondel et al., 2022)	feature_range	(-1,1)
		copy	True
		clip	True
Encodage	sklearn.preprocessing.KBinsDiscretizer (Lin & Dupré la Tour, 2022)	n_bins	30 or 90
		encode	ordinal
		strategy	uniform
		dtype	None
		subsample	warn
		random_state	None
Information mutuelle accidentelle	Aucune fonction existante utilisée	max_iter	500
		p-val	0.1

Tableau-A I-2 Résumé des paramètres utilisés pour le détecteur des mouvements oculaires

<u>Section</u>	<u>Fonction</u>	<u>Nom du paramètre</u>	<u>Valeur</u>
Approximation polynomiale	Aucune fonction existante utilisée	order	11
Dérivée instantanée	Aucune fonction existante utilisée	border_effect	10
Sélection par percentile	Aucune fonction existante utilisée	percentile	65

Tableau-A I-3 Résumé des paramètres utilisés pour les micro-états

<u>Section</u>	<u>Fonction</u>	<u>Nom du paramètre</u>	<u>Valeur</u>
Prétraitement	Pop_micro_selectdata (Poulsen et al., 2018)	Data type	Spontaneous – GFP peaks
		Select and aggregate data from other datasets	Not check
		Calculate average reference	Check
		Normalise dataset	Check
		Minimum peak distance	10
		No. of GFP peaks per subjects that enter the segmentation	1000
		Reject peaks over threshold	0
Algorithme de segmentation : k-means	pop_micro_segment (Poulsen et al., 2018)	algorithm	kmeans
		sorting	Global explained variance
		Nmicrostates	3:8
		verbose	1
		normalise	1
		Nrepetitions	50
		max_iterations	1000
		threshold	1e-06

<u>Section</u>	<u>Fonction</u>	<u>Nom du paramètre</u>	<u>Valeur</u>
Algorithme de segmentation : modified k-means	pop_micro_segment (Poulsen et al., 2018)	algorithm	modkmeans
		sorting	Global explained variance
		Nmicrostates	3:8
		verbose	1
		normalise	1
		Nrepetitions	50
		max_iterations	1000
		threshold	1e-06
		fitmeas	CV
optimised	1		
Algorithme de segmentation : aahc	pop_micro_segment (Poulsen et al., 2018)	algorithm	aahc
		sorting	Global explained variance
		Nmicrostates	3:8
		verbose	1
		normalise	1
		polarity	0
		determinism	1
Algorithme de segmentation : taahc	pop_micro_segment (Poulsen et al., 2018)	algorithm	taahc
		sorting	Global explained variance
		Nmicrostates	3:8
		verbose	1
		normalise	1
		polarity	0
		determinism	1

<u>Section</u>	<u>Fonction</u>	<u>Nom du paramètre</u>	<u>Valeur</u>
Assignment d'un micro-état à chaque échantillon	pop_micro_fit (Poulsen et al., 2018)	polarity	0
Assignment d'un micro-état à chaque échantillon	pop_micro_smooth (Poulsen et al., 2018)	label_type	backfit
		smooth_type	reject segments
		minTF	30
		polarity	0
Permet de produire les résultats	Pop_micro_stats (Poulsen et al., 2018)	label_type	backfit
		polarity	0

LISTE DE RÉFÉRENCES BIBLIOGRAPHIQUES

- Agarwal, R., Takeuchi, T., Laroche, S., & Gotman, J. (2005). Detection of Rapid-Eye Movements in Sleep Studies. *IEEE Transactions on Biomedical Engineering*, 52(8), 1390-1396. <https://doi.org/10.1109/TBME.2005.851512>
- Alvarez-Estevez, D., van Velzen, I., Ottolini-Capellen, T., & Kemp, B. (2017). Derivation and modeling of two new features for the characterization of rapid and slow eye movements in electrooculographic sleep recordings. *Biomedical Signal Processing and Control*, 35, 87-99. <https://doi.org/10.1016/j.bspc.2017.02.014>
- Arthur, D., & Vassilvitskii, S. (2007, janvier 7). Principles—K-means versus k-means ++ clustering technique. *Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms*. SODA 2007, New Orleans, Louisiana, USA. https://www.researchgate.net/publication/220778887_K-Means_The_Advantages_of_Careful_Seeding
- Barnard, B. (2004). Hierarchical clustering. Dans *Wikipedia commons*. https://commons.wikimedia.org/wiki/File:Hierarchical_clustering.png
- Betta, M., Gemignani, A., Landi, A., Laurino, M., Piaggi, P., & Menicucci, D. (2013). Detection and removal of ocular artifacts from EEG signals for an automated REM sleep analysis. *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 5079-5082. <https://doi.org/10.1109/EMBC.2013.6610690>
- Bian, N.-Y., Wang, B., Cao, Y., & Zhang, L. (2006). Automatic Removal of Artifacts from EEG Data Using ICA and Exponential Analysis. Dans J. Wang, Z. Yi, J. M. Zurada, B.-L. Lu, & H. Yin (Éds.), *Advances in Neural Networks—ISNN 2006* (Vol. 3972, p. 719-726). Springer Berlin Heidelberg. https://doi.org/10.1007/11760023_106
- Blondel, M., Grisel, O., Mueller, A., Martin, E., Patrini, G., & Chang, E. (2022, juillet 6). *MinMaxScaler* [Algorithm documentation]. scikit learn. <https://scikit-learn.org/stable/modules/generated/sklearn.preprocessing.MinMaxScaler.html>
- Blumberg, M. S., Lesku, J. A., Libourel, P.-A., Schmidt, M. H., & Rattenborg, N. C. (2020). What Is REM Sleep? *Current Biology*, 30(1), R38-R49. <https://doi.org/10.1016/j.cub.2019.11.045>
- Britz, J., Van De Ville, D., & Michel, C. M. (2010). BOLD correlates of EEG topography reveal rapid resting-state network dynamics. *NeuroImage*, 52(4), 1162-1170. <https://doi.org/10.1016/j.neuroimage.2010.02.052>

- Brunet, D., Murray, M. M., & Michel, C. M. (2011). Spatiotemporal Analysis of Multichannel EEG: CARTOOL. *Computational Intelligence and Neuroscience*, 2011, 1-15. <https://doi.org/10.1155/2011/813870>
- Cacioppo, S., Weiss, R. M., Runesha, H. B., & Cacioppo, J. T. (2014). Dynamic spatiotemporal brain analyses using high performance electrical neuroimaging: Theoretical framework and validation. *Journal of Neuroscience Methods*, 238, 11-34. <https://doi.org/10.1016/j.jneumeth.2014.09.009>
- Chen, K.-S., Xu, M., Zhang, Z., Chang, W.-C., Gaj, T., Schaffer, D. V., & Dan, Y. (2018). A Hypothalamic Switch for REM and Non-REM Sleep. *Neuron*, 97(5), 1168-1176.e4. <https://doi.org/10.1016/j.neuron.2018.02.005>
- Chu, C., Wang, X., Cai, L., Zhang, L., Wang, J., Liu, C., & Zhu, X. (2020). Spatiotemporal EEG microstate analysis in drug-free patients with Parkinson's disease. *NeuroImage: Clinical*, 25, 102132. <https://doi.org/10.1016/j.nicl.2019.102132>
- Custo, A., Van De Ville, D., Wells, W. M., Tomescu, M. I., Brunet, D., & Michel, C. M. (2017). Electroencephalographic Resting-State Networks: Source Localization of Microstates. *Brain Connectivity*, 7(10), 671-682. <https://doi.org/10.1089/brain.2016.0476>
- Dauvilliers, Y., Schenck, C. H., Postuma, R. B., Iranzo, A., Luppi, P.-H., Plazzi, G., Montplaisir, J., & Boeve, B. (2018). REM sleep behaviour disorder. *Nature Reviews Disease Primers*, 4(1), 19. <https://doi.org/10.1038/s41572-018-0016-5>
- Degler, H. E., Smith, J. R., & Black, F. O. (1975). Automatic detection and resolution of synchronous rapid eye movements. *Computers and Biomedical Research*, 8(4), 393-404. [https://doi.org/10.1016/0010-4809\(75\)90015-4](https://doi.org/10.1016/0010-4809(75)90015-4)
- Doman, J., Detka, C., Hoffman, T., Kesicki, D., Monahan, J. P., Buysse, D. J., Reynolds, C. F., Coble, P. A., Matzzie, J., & Kupfer, D. J. (1995). Automating the sleep laboratory: Implementation and validation of digital recording and analysis. *International Journal of Bio-Medical Computing*, 38(3), 277-290. [https://doi.org/10.1016/S0020-7101\(05\)80010-8](https://doi.org/10.1016/S0020-7101(05)80010-8)
- DORMEZ LÀ-DESSUS. (2019). *Qu'est-ce que le sommeil* [Campagne de sensibilisation]. Dormez là-dessus. <https://dormezladessuscanada.ca/tout-sur-le-sommeil/qu-est-ce-que-le-sommeil/>
- Fereshtehnejad, S.-M., Yao, C., Pelletier, A., Montplaisir, J. Y., Gagnon, J.-F., & Postuma, R. B. (2019). Evolution of prodromal Parkinson's disease and dementia with Lewy

- bodies : A prospective study. *Brain*, 142(7), 2051-2067. <https://doi.org/10.1093/brain/awz111>
- Figorilli, M., Ferri, R., Zibetti, M., Beudin, P., Puligheddu, M., Lopiano, L., Cicolin, A., Durif, F., Marques, A., & Fantini, M. L. (2017). Comparison Between Automatic and Visual Scorings of REM Sleep Without Atonia for the Diagnosis of REM Sleep Behavior Disorder in Parkinson Disease. *Sleep*, 40(2). <https://doi.org/10.1093/sleep/zsw060>
- Flexer, A., Bauer, H., Pripfl, J., & Dorffner, G. (2005). Using ICA for removal of ocular artifacts in EEG recorded from blind subjects. *Neural Networks*, 18(7), 998-1005. <https://doi.org/10.1016/j.neunet.2005.03.012>
- Grisel, O., Li, W., Molla, D., Fouchet, A., Guillemot, T., Stupp, G., Nothman, J., McCarthy, A., & F Mayer, U. (2022, juillet 6). *Mutual_info_score* [Algorithm documentation]. scikit learn. https://scikit-learn.org/stable/modules/generated/sklearn.metrics.mutual_info_score.html
- Gschwind, M., Hardmeier, M., Van De Ville, D., Tomescu, M. I., Penner, I.-K., Naegelin, Y., Fuhr, P., Michel, C. M., & Seeck, M. (2016). Fluctuations of spontaneous EEG topographies predict disease state in relapsing-remitting multiple sclerosis. *NeuroImage: Clinical*, 12, 466-477. <https://doi.org/10.1016/j.nicl.2016.08.008>
- Hatzilabrou, G. M., Greenberg, N., Scwabassi, R. J., Carroll, T., Guthrie, R. D., & Scher, M. S. (1994). A comparison of conventional and matched filtering techniques for rapid eye movement detection of the newborn [electro-oculography]. *IEEE Transactions on Biomedical Engineering*, 41(10), 990-995. <https://doi.org/10.1109/10.324532>
- Herrmann, J. M., & Theis, F. J. (2007). Statistical Analysis of Sample-Size Effects in ICA. Dans H. Yin, P. Tino, E. Corchado, W. Byrne, & X. Yao (Éds.), *Intelligent Data Engineering and Automated Learning—IDEAL 2007* (Vol. 4881, p. 416-425). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-77226-2_43
- Ibrahim, I. A., Santhosh, J., & Moghavvemi, M. (2015). A new approach for an effective eye movement artifact elimination from EEG signal. *2015 IEEE International Conference on Signal and Image Processing Applications (ICSIPA)*, 557-562. <https://doi.org/10.1109/ICSIPA.2015.7412253>
- Jabès, A., Klencklen, G., Ruggeri, P., Michel, C. M., Banta Lavenex, P., & Lavenex, P. (2021). Resting-State EEG Microstates Parallel Age-Related Differences in Allocentric Spatial Working Memory Performance. *Brain Topography*, 34(4), 442-460. <https://doi.org/10.1007/s10548-021-00835-3>

- Jiang, X., Bian, G.-B., & Tian, Z. (2019). Removal of Artifacts from EEG Signals : A Review. *Sensors*, *19*(5), 987. <https://doi.org/10.3390/s19050987>
- Khanna, A., Pascual-Leone, A., & Farzan, F. (2014). Reliability of Resting-State Microstate Features in Electroencephalography. *PLoS ONE*, *9*(12), e114163. <https://doi.org/10.1371/journal.pone.0114163>
- Khazaei, M., Raeisi, K., Croce, P., Tamburro, G., Tokariev, A., Vanhatalo, S., Zappasodi, F., & Comani, S. (2021). Characterization of the Functional Dynamics in the Neonatal Brain during REM and NREM Sleep States by means of Microstate Analysis. *Brain Topography*, *34*(5), 555-567. <https://doi.org/10.1007/s10548-021-00861-1>
- Koenig, T., Prichep, L., Lehmann, D., Sosa, P. V., Braecker, E., Kleinlogel, H., Isenhardt, R., & John, E. R. (2002). Millisecond by Millisecond, Year by Year: Normative EEG Microstates and Developmental Stages. *NeuroImage*, *16*(1), 41-48. <https://doi.org/10.1006/nimg.2002.1070>
- Koenig, T., Studer, D., Hubl, D., Melie, L., & Strik, W. K. (2005). Brain connectivity at different time-scales measured with EEG. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *360*(1457), 1015-1024. <https://doi.org/10.1098/rstb.2005.1649>
- Korats, G., Le Cam, S., Ranta, R., & Hamid, M. (2013). Applying ICA in EEG : Choice of the Window Length and of the Decorrelation Method. Dans J. Gabriel, J. Schier, S. Van Huffel, E. Conchon, C. Correia, A. Fred, & H. Gamboa (Éds.), *Biomedical Engineering Systems and Technologies* (Vol. 357, p. 269-286). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-38256-7_18
- Krzanowski, W. J., & Lai, Y. T. (1988). A Criterion for Determining the Number of Groups in a Data Set Using Sum-of-Squares Clustering. *Biometrics*, *44*(1), 23. <https://doi.org/10.2307/2531893>
- Ktonas, P. Y., & Smith, J. R. (1978). Automatic rem detection : Modifications on an existing system and preliminary normative data. *International Journal of Bio-Medical Computing*, *9*(6), 445-464. [https://doi.org/10.1016/0020-7101\(78\)90052-1](https://doi.org/10.1016/0020-7101(78)90052-1)
- Lafaye de Micheaux, P., van der Walt, S., Varoquaux, G., Thirion, B., Gramfort, A., & A. Engemann, D. (2022, avril 9). *FastICA* [Algorithm documentation]. scikit learn. <https://scikit-learn.org/stable/modules/generated/sklearn.decomposition.FastICA.html>
- Lehmann, D., Faber, P. L., Galderisi, S., Herrmann, W. M., Kinoshita, T., Koukkou, M., Mucci, A., Pascual-Marqui, R. D., Saito, N., Wackermann, J., Winterer, G., & Koenig,

- T. (2005). EEG microstate duration and syntax in acute, medication-naïve, first-episode schizophrenia : A multi-center study. *Psychiatry Research: Neuroimaging*, 138(2), 141-156. <https://doi.org/10.1016/j.psychresns.2004.05.007>
- Lévesque, D., Lacourse, K. (2022). Snooz (Version 0) [Logiciel]. Montréal, QC: CIUSS NIM
- Li, Y., Ma, Z., Lu, W., & Li, Y. (2006). Automatic removal of the eye blink artifact from EEG using an ICA-based template matching approach. *Physiological Measurement*, 27(4), 425-436. <https://doi.org/10.1088/0967-3334/27/4/008>
- Lin, H., & Dupré la Tour, T. (2022, juillet 6). *KBinsDiscretizer* [Algorithm documentation]. scikit learn. <https://scikit-learn.org/stable/modules/generated/sklearn.preprocessing.KBinsDiscretizer.html>
- Liu, J., Ramakrishnan, S., Laxminarayan, S., Neal, M., Cashmere, D. J., Germain, A., & Reifman, J. (2018). Effects of signal artefacts on electroencephalography spectral power during sleep : Quantifying the effectiveness of automated artefact-rejection algorithms. *Journal of Sleep Research*, 27(1), 98-102. <https://doi.org/10.1111/jsr.12576>
- Marks, G. A., Shaffery, J. P., Oksenberg, A., Speciale, S. G., & Roffwarg, H. P. (1995). A functional role for REM sleep in brain maturation. *Behavioural Brain Research*, 69(1-2), 1-11. [https://doi.org/10.1016/0166-4328\(95\)00018-0](https://doi.org/10.1016/0166-4328(95)00018-0)
- McPartland, R. J., Kupfer, D. J., & Gordon Foster, F. (1973). Rapid eye movement analyzer. *Electroencephalography and Clinical Neurophysiology*, 34(3), 317-320. [https://doi.org/10.1016/0013-4694\(73\)90259-9](https://doi.org/10.1016/0013-4694(73)90259-9)
- Michel, C. M., & Koenig, T. (2018). EEG microstates as a tool for studying the temporal dynamics of whole-brain neuronal networks : A review. *NeuroImage*, 180, 577-593. <https://doi.org/10.1016/j.neuroimage.2017.11.062>
- Miettinen, J., Nordhausen, K., & Taskinen, S. (2018). *FICA: FastICA Algorithms and Their Improved Variants*. 10, 148-158.
- Minard, J. G., & Krausman, D. (1971). Rapid eye movement definition and count : An on-line detector. *Electroencephalography and Clinical Neurophysiology*, 31(1), 99-102. [https://doi.org/10.1016/0013-4694\(71\)90295-1](https://doi.org/10.1016/0013-4694(71)90295-1)
- Morissette, L., & Chartier, S. (2013). The k-means clustering technique : General considerations and implementation in Mathematica. *Tutorials in Quantitative Methods for Psychology*, 9(1), 15-24. <https://doi.org/10.20982/tqmp.09.1.p015>

- Murphy, M., Whitton, A. E., Deccy, S., Ironside, M. L., Rutherford, A., Beltzer, M., Sacchet, M., & Pizzagalli, D. A. (2020). Abnormalities in electroencephalographic microstates are state and trait markers of major depressive disorder. *Neuropsychopharmacology*, *45*(12), 2030-2037. <https://doi.org/10.1038/s41386-020-0749-1>
- Murray, M. M., Brunet, D., & Michel, C. M. (2008). Topographic ERP Analyses : A Step-by-Step Tutorial Review. *Brain Topography*, *20*(4), 249-264. <https://doi.org/10.1007/s10548-008-0054-5>
- Nassif, F., Rezaii, T. Y., & Beheshti, S. (2020). Pole-Zero REM Modeling with Application in EEG Artifact Removal. *2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, 402-405. <https://doi.org/10.1109/EMBC44109.2020.9176489>
- Onton, J., & Makeig, S. (2006). Information-based modeling of event-related brain dynamics. Dans *Progress in Brain Research* (Vol. 159, p. 99-120). Elsevier. [https://doi.org/10.1016/S0079-6123\(06\)59007-7](https://doi.org/10.1016/S0079-6123(06)59007-7)
- O'Reilly, C., Gosselin, N., Carrier, J., & Nielsen, T. (2014). Montreal Archive of Sleep Studies : An open-access resource for instrument benchmarking and exploratory research. *Journal of Sleep Research*, *23*(6), 628-635. <https://doi.org/10.1111/jsr.12169>
- Pascual-Marqui, R. D., Michel, C. M., & Lehmann, D. (1995). Segmentation of brain electrical activity into microstates : Model estimation and validation. *IEEE Transactions on Biomedical Engineering*, *42*(7), 658-665. <https://doi.org/10.1109/10.391164>
- Peng, A., Wang, R., Huang, J., Wu, H., & Chen, L. (2021). Abnormalities of Resting-State Electroencephalographic Microstate in Rapid Eye Movement Sleep Behavior Disorder. *Frontiers in Human Neuroscience*, *15*, 728405. <https://doi.org/10.3389/fnhum.2021.728405>
- Pluim, J. P. W., Maintz, J. B. A., & Viergever, M. A. (2003). Mutual-information-based registration of medical images : A survey. *IEEE Transactions on Medical Imaging*, *22*(8), 986-1004. <https://doi.org/10.1109/TMI.2003.815867>
- Poulsen, A. T., Pedroni, A., Langer, N., & Hansen, L. K. (2018). *Microstate EEGlab toolbox : An introductory guide* [Preprint]. Neuroscience. <https://doi.org/10.1101/289850>

- Quantin, M. (2017, juillet). *K-moyennes* [Encyclopédie]. Wikipedia. <https://fr.wikipedia.org/wiki/K-moyennes>
- Romero, S., Mananas, M. A., Clos, S., Gimenez, S., & Barbanoj, M. J. (2003). Reduction of EEG artifacts by ICA in different sleep stages. *Proceedings of the 25th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (IEEE Cat. No.03CH37439)*, 2675-2678. <https://doi.org/10.1109/IEMBS.2003.1280467>
- Särelä, J., & Vigário, R. (2003, Décembre). Journal of Machine Learning. *Overlearning in Marginal Distribution-Based ICA: Analysis and Solutions*, 1447-1469.
- Schumacher, J., Peraza, L. R., Firbank, M., Thomas, A. J., Kaiser, M., Gallagher, P., O'Brien, J. T., Blamire, A. M., & Taylor, J.-P. (2019). Dysfunctional brain dynamics and their origin in Lewy body dementia. *Brain*, 142(6), 1767-1782. <https://doi.org/10.1093/brain/awz069>
- Shannon, C. E. (1948). A Mathematical Theory of Communication. *Bell System Technical Journal*, 27(3), 379-423. <https://doi.org/10.1002/j.1538-7305.1948.tb01338.x>
- Simor, P., van der Wijk, G., Nobili, L., & Peigneux, P. (2020). The microstructure of REM sleep: Why phasic and tonic? *Sleep Medicine Reviews*, 52, 101305. <https://doi.org/10.1016/j.smr.2020.101305>
- Skrandies, W. (1990). Global field power and topographic similarity. *Brain Topography*, 3(1), 137-141. <https://doi.org/10.1007/BF01128870>
- Smailovic, U., Ferreira, D., Ausén, B., Ashton, N. J., Koenig, T., Zetterberg, H., Blennow, K., & Jelic, V. (2022). Decreased Electroencephalography Global Field Synchronization in Slow-Frequency Bands Characterizes Synaptic Dysfunction in Amnesic Subtypes of Mild Cognitive Impairment. *Frontiers in Aging Neuroscience*, 14, 755454. <https://doi.org/10.3389/fnagi.2022.755454>
- Smith, J. R., Cronin, M. J., & Karacan, I. (1971). A multichannel hybrid system for rapid eye movement detection (REM detection). *Computers and Biomedical Research*, 4(3), 275-290. [https://doi.org/10.1016/0010-4809\(71\)90032-2](https://doi.org/10.1016/0010-4809(71)90032-2)
- Soler, A., Drange, O., Furuki, J., Abe, T., & Molinas, M. (2021). Automatic Onset Detection of Rapid Eye Movements in REM Sleep EEG Data. *IFAC-PapersOnLine*, 54(15), 257-262. <https://doi.org/10.1016/j.ifacol.2021.10.265>

- Suni, E., & Vyas, Dr. N. (2022, octobre 7). *Stages of Sleep*. Sleep Foundation. <https://www.sleepfoundation.org/stages-of-sleep>
- Tait, L., Tamagnini, F., Stothart, G., Barvas, E., Monaldini, C., Frusciante, R., Volpini, M., Guttmann, S., Coulthard, E., Brown, J. T., Kazanina, N., & Goodfellow, M. (2020). EEG microstate complexity for aiding early diagnosis of Alzheimer's disease. *Scientific Reports*, *10*(1), 17627. <https://doi.org/10.1038/s41598-020-74790-7>
- Tharwat, A. (2021). Independent component analysis : An introduction. *Applied Computing and Informatics*, *17*(2), 222-249. <https://doi.org/10.1016/j.aci.2018.08.006>
- Ting, K. H., Fung, P. C. W., Chang, C. Q., & Chan, F. H. Y. (2006). Automatic correction of artifact from single-trial event-related potentials by blind source separation using second order statistics only. *Medical Engineering & Physics*, *28*(8), 780-794. <https://doi.org/10.1016/j.medengphy.2005.11.006>
- Tsuji, Y., Satoh, H., Itoh, N., Sekiguchi, Y., & Nagasawa, K. (2000). Automatic detection of rapid eye movements by discrete wavelet transform. *Psychiatry and Clinical Neurosciences*, *54*(3), 276-277. <https://doi.org/10.1046/j.1440-1819.2000.00676.x>
- Uddin, Z., Ahmad, A., Iqbal, M., & Naeem, M. (2016). Modified Infomax Algorithm for Smaller Data Block Lengths. *Wireless Personal Communications*, *87*(1), 245-267. <https://doi.org/10.1007/s11277-015-3041-7>
- von Wegner, F., & Laufs, H. (2018). Information-Theoretical Analysis of EEG Microstate Sequences in Python. *Frontiers in Neuroinformatics*, *12*, 30. <https://doi.org/10.3389/fninf.2018.00030>
- Vyazovskiy, V. V., & Delogu, A. (2014). NREM and REM Sleep : Complementary Roles in Recovery after Wakefulness. *The Neuroscientist*, *20*(3), 203-219. <https://doi.org/10.1177/1073858413518152>
- Waterman, D., Woestenburg, J. C., Elton, M., Hofman, W., & Kok, A. (1992). Removal of Ocular Artifacts From the REM Sleep EEG. *Sleep*, *15*(4), 371-375. <https://doi.org/10.1093/sleep/15.4.371>
- Wikipedia. (2010). 10–20 system (EEG). Dans *Wikipedia*. [https://en.wikipedia.org/w/index.php?title=10%E2%80%9320_system_\(EEG\)&oldid=1076814653](https://en.wikipedia.org/w/index.php?title=10%E2%80%9320_system_(EEG)&oldid=1076814653)
- Xiong, X., Ren, Y., Gao, S., Luo, J., Liao, J., Wang, C., Yi, S., Liu, R., Xiang, Y., & He, J. (2021). EEG microstate in obstructive sleep apnea patients. *Scientific Reports*, *11*(1), 17178. <https://doi.org/10.1038/s41598-021-95749-2>

- Yetton, B. D., Niknazar, M., Duggan, K. A., McDevitt, E. A., Whitehurst, L. N., Sattari, N., & Mednick, S. C. (2016). Automatic detection of rapid eye movements (REMs): A machine learning approach. *Journal of Neuroscience Methods*, 259, 72-82. <https://doi.org/10.1016/j.jneumeth.2015.11.015>
- Zhou, B., Wu, X., Ruan, J., Lv, Z., & Zhang, L. (2019). How many channels are suitable for independent component analysis in motor imagery brain-computer interface. *Biomedical Signal Processing and Control*, 50, 103-120. <https://doi.org/10.1016/j.bspc.2019.01.017>
- Zhou, W., & Gotman, J. (2009). Automatic removal of eye movement artifacts from the EEG using ICA and the dipole model. *Progress in Natural Science*, 19(9), 1165-1170. <https://doi.org/10.1016/j.pnsc.2008.11.013>