i

#### HEC MONTRÉAL

#### Détermination de la structure de corrélation entre les différentes électrodes analysées conjointement lors de l'étude de données électroencéphalographiques (EEG) par Zi Su Zhao

Sciences de la gestion (Spécialisation Intelligence d'affaires)

Mémoire présenté en vue de l'obtention du grade de maîtrise ès sciences en gestion (M. Sc.)

> Octobre 2022 © Zi Su Zhao, 2022



Comité d'éthique de la recherche

#### CERTIFICAT D'APPROBATION ÉTHIQUE

La présente atteste que le projet de recherche décrit ci-dessous a fait l'objet d'une évaluation en matière d'éthique de la recherche avec des êtres humains et qu'il satisfait aux exigences de notre politique en cette matière.

Projet #: 2018-2904

Titre du projet de recherche : Recherche et développement collaborative en UX

Chercheur principal : Marc Fredette, Professeur titulaire, Département de sciences de la décision, HEC Montréal

Cochercheurs : Pierre-Majorique Léger; François Courtemanche; Shang-Lin Chen

Date d'approbation du projet : 11 janvier 2018

Date d'entrée en vigueur du certificat : 11 janvier 2018

Date d'échéance du certificat : 01 janvier 2019

Mu M

Maurice Lemelin Président du CER de HEC Montréal

NAGANO Approbation du projet par le comité d'éthique Comité d'éthique de la recherche - HEC Montréal

## HEC MONTREAL

Comité d'éthique de la recherche

#### ATTESTATION D'APPROBATION ÉTHIQUE COMPLÉTÉE

La présente atteste que le projet de recherche décrit ci-dessous a fait l'objet des approbations en matière d'éthique de la recherche avec des êtres humains nécessaires selon les exigences de HEC Montréal.

La période de validité du certificat d'approbation éthique émis pour ce projet est maintenant terminée. Si vous devez reprendre contact avec les participants ou reprendre une collecte de données pour ce projet, la certification éthique doit être réactivée préalablement. Vous devez alors prendre contact avec le secrétariat du CER de HEC Montréal.

Nom de l'étudiant(e) : Zhao, Zi Su

#### Titre du projet supervisé/mémoire/thèse :

Détermination de la structure de corrélation entre les différentes électrodes analysées conjointement lors de l'étude de données électroencéphalographiques (EEG).

Titre du projet sur le certificat : Recherche et développement collaborative en UX

Projet # : 2018-2904

Chercheur principal / directeur de recherche : Marc Fredette

#### **Cochercheurs :**

Pierre-Majorique Léger; François Courtemanche; Shang Lin Chen; Elise Boissonneault; Florian Coustures; Melissa Turp-Yonezawa; Amay Cheam; Pierre-Jude Souverain; Omar Chikhar

#### Date d'approbation initiale du projet : 5 décembre 2017

Date de fermeture de l'approbation éthique pour l'étudiant(e) : 10 août 2022

11 /1

Maurice Lemelin Président CER de HEC Montréal

Signé le 2022-08-10 à 10:35

## Résumé

Les données EEG sont les signaux électriques enregistrés par les électrodes fixées sur la surface du cuir chevelu des êtres humains. Les relations entre elles sont intéressantes à être modélisées afin d'avoir une prédictivité sur les activités du cerveau. Plusieurs relations sont existantes dans le cerveau, tout comme celles entre les différents moments et celles entre les différentes zones de celui-ci. Pour mieux comprendre leurs tendances, nous proposons de faire des modélisations de structures de covariances pour déterminer les modèles qui ajustent les mieux aux corrélations entre les différentes relations. La méthode proposée est d'analyser séparément la relation entre les modélisant des structures de covariance univariées. Nous tentons finalement de trouver un modèle multivarié (MVAR) permettant d'analyser les relations entre les différentes temps. Pour les analyses des relations entre les électrodes et pour les analyses de relations entre les temps, l'ajustement est basé sur la distance entre les électrodes et points de temps.

Il s'avère que les modèles qui modélisent les mieux les corrélations entre les électrodes sont « exponentielle spatiale », « exponentielle anisotrope spatiale », « puissance spatiale » et « linéaire autorégressif exposant spatial » et les modèles qui modélisent les mieux les corrélations entre les temps sont « modèle autorégressif » et « moyenne mobile autorégressive ». En raison des limites de la recherche, aucun modèle de covariance multivarié n'a été trouvé pour modéliser les relations entre les différentes électrodes aux différents temps. Cette recherche permettra donc de donner une piste pour les prochaines recherches futures pour mieux comprendre quels sont les meilleurs modèles qui permettent de prédire les activités du cerveau en fonction des données EEG existantes.

Mots clés : EEG, structure de covariance, corrélation, cerveau

## Abstract

EEG data are the electrical signals recorded by the designated electrodes on the surface of the scalp of human beings. The relationships between them are interesting to be modeled, in order to be able to predict the activities of the brain. Several relationships exist in the brain, just like those between the different moments and those between the different areas of it. To better understand their trends, we plan to model covariance structures to determine the models that best fit the correlations between the different relationships. The proposed method is to separately analyze the relation between the moments for each electrode and the relation between the electrodes for each selected time by modeling univariate covariance structures. We are finally trying to find a multivariate model (MVAR) to analyze the relationships between the different areas of the brain at different times. For electrode relationship analyses, the fit of the models is based on the distance between the time points.

It turns out that the models that model the best correlations between the electrodes are "spatial exponential", "spatial anisotropic exponential", "spatial power" and "linear autoregressive spatial exponent" and the models that model the best correlations between the times are "autoregressive model" and "autoregressive moving average". Due to research limitations, no multivariate covariance model was found to model the relationships between different electrodes at different times. This research will therefore provide an avenue for future research to better understand which are the best models that can predict brain activities based on existing EEG data.

Keywords: EEG, covariance structure, correlation, brain

## Table des matières

Résumé	vii
Abstract	viii
Table des matières	ix
Liste des tables	xii
Liste des figures	xvi
Liste des abréviations	xviii
Avant-propos	xix
Remerciement	XX
Introduction	1
Revue de littératures	3
Structure du cerveau	3
Les zones du cerveau	3
Les états du cerveau	5
EEG	7
Notions générales	8
Positions des électrodes	9
Mesures	9
Méthodes utilisées pour les analyses de corrélations des données EEG	10
Modèle autorégressif multivarié (MVAR)	10
Corrélation de Pearson	12
Analyse de cohérence	13
Chapitre 1 : Présentation des données	15
Collecte des données	15
Manipulation des données	16
Nettoyage des données EEG collectées	16

Manipulation des données nettoyées17
Données des positions d'électrodes18
Choix des données
Chapitre 2 : Analyse descriptive
Analyse de corrélation de Pearson24
Analyse des relations entre les électrodes24
Analyse des distances entre les électrodes
Analyse des relations entre les différents points de temps40
Analyse des relations entre les électrodes au cours des différents moments dans le temps
Recherche de solution pour l'analyse d'un grand nombre de données
Solution 1 : Agrégation des données par moyenne52
Solution 2 : Saut dans le temps
Chapitre 3 : Méthodologie
Tests de structures de covariance
Analyse des corrélations entre les électrodes58
Analyse des corrélations entre les différents points de temps64
Analyse des corrélations entre les électrodes aux différents moments dans le temps69
Critères de choix de structures70
Chapitre 4 : Analyse des modèles mixtes
Analyses des structures modélisant les corrélations entre les électrodes72
Analyse des paramètres des meilleurs modèles spatiaux74
Analyses des structures modélisant les corrélations entre les temps
Analyse des paramètres des meilleurs modèles
Analyse du modèle multivarié106
Chapitre 5 : Limites
Les conséquences causées par le temps de recherche limité107

Les conséquences causées par la limite des modèles	108
Les conséquences causées par la limite des bases de données utilisées	
Conclusion	
Bibliographie	111

## Liste des tables

Tableau 1 - Nombre de participants ayant les électrodes non éliminées selon les règles de
nettoyages des données EEG16
Tableau 2 - Coordonnées angulaires selon les positionnements fournis par Khosla et al. (1999)
Tableau 3 - Coordonnées angulaires selon les positionnements fournis par Böcker et al. (1994)
Tableau 4 - Coordonnées tridimensionnelles selon les positionnements fournis par Khosla et al.(1999)20
Tableau 5 Coordonnées tridimensionnelles selon les positionnements fournis par Böcker et
-1 (1004)
al. (1994)
Tableau 6 - Coordonnées bidimensionnelles selon les positionnements fournis par Khosla et al.         (1000)
(1999)
Tableau 7 - Coordonnées bidimensionnelles selon les positionnements fournis par Böcker et al.
(1994)
Tableau 8 - Mise en forme conditionnelle des corrélations dans la matrice (tableau 9)25
Tableau 9 - Matrice de corrélations de Pearson des paires d'électrodes       26
Tableau 10 - Matrice des corrélations de Pearson significatives entre les paires d'électrodes 28
Tableau 11 - Matrice des nombres de corrélations significatives
Tableau 12 - Mise en forme conditionnelle de la matrice des nombres de corrélations
significatives (tableau 13)
Tableau 13 - Les paramètres estimés pour la régression linéaire modélisée pour la corrélation
en fonction de la distance euclidienne selon les positionnements bidimensionnels de Böcker et
al. (1994)
Tableau 14 - Les paramètres estimés pour la régression linéaire modélisée pour la corrélation
en fonction de la distance euclidienne selon les positionnements bidimensionnels de Khosla et
al. (1999)
Tableau 15 - Les paramètres estimés pour la régression linéaire modélisée pour la corrélation
en fonction de la distance euclidienne selon les positionnements tridimensionnels de Böcker et
al. (1994)

Tableau 16 - Les paramètres estimés pour la régression linéaire modélisée pour la corrélation en fonction de la distance euclidienne selon les positionnements tridimensionnels de Khosla et Tableau 17 - Nombres de corrélations significatives, triés en ordre croissant, entre les temps Tableau 18 - Partie de matrice de moyennes de corrélations entre les électrodes aux différents Tableau 19 - Partie de matrice de moyennes de corrélations entre les électrodes aux différents temps (temps 1 à temps 9) pour les électrodes Fp1 et Fpz ......47 Tableau 20 - Partie de matrice de moyennes de corrélations entre les électrodes aux différents temps (temps 1 à temps 9) pour les électrodes Fp2 et Fpz ......47 Tableau 21 - Partie de matrice de moyennes de corrélations entre les électrodes aux différents Tableau 22 - Partie de matrice de moyennes de corrélations entre les électrodes aux différents Tableau 23 - Partie de matrice de moyennes de corrélations entre les électrodes aux différents temps (temps 1 à temps 9) pour les électrodes Oz et Pz......48 Tableau 24 - Les rangs moyens des modèles pour les deux types de coordonnées selon les positionnements d'électrodes fournis par les deux articles de Bocker et al. (1994) et Bocker et Tableau 25 - Nombre de fois que les modèles du tableau 24 reviennent parmi les meilleurs Tableau 26 - Comparaison des moyennes de BIC du modèle SP(POW) selon les deux types de coordonnées des positionnements provenant des deux articles de Böcker et al. (1994) et Khosla Tableau 27 - Comparaison des moyennes de BIC du modèle SP(EXP) selon les deux types de coordonnées des positionnements provenant des deux articles de Böcker et al. (1994) et Khosla Tableau 28 - Comparaison des moyennes de BIC du modèle SP(EXPA) selon les deux types de coordonnées des positionnements provenant des deux articles de Böcker et al. (1994) et Tableau 29 - Comparaison des moyennes de BIC du modèle SP(LEAR) selon les deux types de coordonnées des positionnements provenant des deux articles de Böcker et al. (1994) et 

Tableau 30 – Les moyennes des paramètres estimés par le modèle puissance spatiale ..........77 Tableau 31 - Corrélations entre les paires d'électrodes en lien avec FCz calculées selon les moyennes des paramètres estimés par le modèle puissance spatiale pour les positionnements tridimensionnels des électrodes selon l'article de Böcker et al. (1994) ......77 Tableau 32 - Corrélations entre les paires d'électrodes en lien avec FCz calculées selon les moyennes des paramètres estimés par le modèle puissance spatiale pour les positionnements Tableau 33 – Les moyennes des paramètres estimés par le modèle exponentiel spatial .........79 Tableau 34 - Corrélations entre les paires d'électrodes en lien avec FCz calculées selon les moyennes des paramètres estimés par le modèle exponentiel spatial pour les positionnements Tableau 35 - Corrélations entre les paires d'électrodes en lien avec FCz calculées selon les moyennes des paramètres estimés par le modèle exponentiel spatial pour les positionnements tridimensionnels des électrodes selon l'article de Khosla et al. (1999) ......80 Tableau 36 – Les moyennes des paramètres estimés par le modèle exponentiel anisotrope 

Tableau 37 - Corrélations entre les paires d'électrodes en lien avec FCz calculées selon les moyennes des paramètres estimés par le modèle exponentiel anisotrope spatial pour les positionnements tridimensionnels des électrodes selon l'article de Böcker et al. (1994) ......81 Tableau 38 - Corrélations entre les paires d'électrodes en lien avec FCz calculées selon les moyennes des paramètres estimés par le modèle exponentiel anisotrope spatial pour les positionnements tridimensionnels des électrodes selon l'article de Khosla et al. (1999) ......82 Tableau 39 – Les moyennes des paramètres estimés par le modèle linéaire exposant Tableau 40 - Corrélations entre les paires d'électrodes en lien avec FCz calculées selon les moyennes des paramètres estimés par le modèle linéaire exposant autorégressif spatial pour les positionnements tridimensionnels des électrodes selon l'article de Böcker et al. (1994) .......83 Tableau 41 - Corrélations entre les paires d'électrodes en lien avec FCz calculées selon les moyennes des paramètres estimés par le modèle linéaire exposant autorégressif spatial pour les positionnements tridimensionnels des électrodes selon l'article de Khosla et al. (1999) .......84 Tableau 42 - Moyennes et écart-types des rangs de BIC des trois meilleurs modèles pour modéliser les activités cérébrales à travers le temps sur 20 points dans le temps pour chacune 

Tableau 43 - Nombre de fois que les modèles reviennent comme les trois meilleurs modèles
pour chacune des électrodes pour modéliser les activités cérébrales à travers les durées de 20
points dans le temps91
Tableau 44 - Moyennes et écart-types des rangs de BIC des trois meilleurs modèles pour
modéliser les activités cérébrales à travers le temps sur 50 points dans le temps pour chacune
des électrodes
Tableau 45 - Nombre de fois que les modèles reviennent comme les trois meilleurs modèles
pour chacune des électrodes pour modéliser les activités cérébrales à travers les durées de 50
points dans le temps
Tableau 46 - Liste des modèles qui n'ont pas converti pour les 42 modélisations94
Tableau 47 - Moyennes et écart-types des rangs de BIC des trois meilleurs modèles pour
modéliser les activités cérébrales à travers le temps sur 100 points dans le temps pour chacune
des électrodes95
Tableau 48 - Nombre de fois que les modèles reviennent comme les trois meilleurs modèles
pour chacune des électrodes pour modéliser les activités cérébrales à travers les durées de 100
points dans le temps97
Tableau 49 - Liste des modèles qui n'ont pas converti pour les 5 modélisations97
Tableau 50 - Les paramètres estimés du modèle AR(1) pour les quatre moments avec une durée
de 500 points de temps
Tableau 51 - Les paramètres estimés du modèle AR(1) pour les quatre moments avec une durée
de 1000 points de temps
Tableau 52 - Les paramètres estimés du modèle AR(1) pour les quatre moments avec une durée
de 2000 points de temps
Tableau 53 - Moyennes de tous les paramètres estimés pour le modèle AR(1) à 1000 points de
temps
Tableau 54 - Moyennes des paramètres estimés pour chacune des électrodes pour le modèle
AR(1) à 1000 points de temps
Tableau 55 - Moyennes de tous les paramètres estimés pour le modèle ARMA(1,1) à 1000
points de temps
Tableau 56 - Moyennes des paramètres estimés pour chacune des électrodes pour le modèle
ARMA(1,1) à 1000 points de temps

## Liste des figures

Figure 1 – Les différents lobes du cerveau Note: Tirée du site de McGill (s.d)
(https://thebrain.mcgill.ca), utilisation libre sous copyleft
Figure 2 - Les différentes ondes du cerveau Note : Tirée du site de McGill (s.d)
(https://thebrain.mcgill.ca), utilisation libre sous copyleft
Figure 3 - Les positions des électrodes9
Figure 4 - Les corrélations de Pearson pour les paires d'électrodes associées à FCz illustrées
selon les positionnements d'électrodes Note : Les couleurs en ordre croissant de corrélations :
rouge (négatif), orange, jaune, jaune vert, vert, vert peu foncé et vert foncé
Figure 5 - Courbe d'ajustement de la corrélation de Pearson moyenne en fonction des distances
selon les positionnements bidimensionnels de Böcker et al. (1999)
Figure 6 - Courbe d'ajustement de la corrélation de Pearson moyenne en fonction des distances
selon les positionnements bidimensionnels de Khosla et al. (1994)
Figure 7 - Courbe d'ajustement de la corrélation de Pearson moyenne en fonction des distances
selon les positionnements tridimensionnels de Böcker et al. (1999)
Figure 8 - Courbe d'ajustement de la corrélation de Pearson moyenne en fonction des distances
selon les positionnements tridimensionnels de Khosla et al. (1994)
Figure 9 - Nuage de points des proportions des corrélations de Pearson significatives en
fonction des moyennes de corrélation pour les différentes distances (selon les positionnements
de Khosla et al. (1999))
Figure 10 - Comparaison des courbes de tendance pour les corrélations significatives moyennes
selon les distances calculées en fonction des positionnements bidimensionnels et
tridimensionnels des deux articles40
Figure 11 - Moyenne de toutes les corrélations en fonction du nombre de points de temps
séparant deux points de temps41
Figure 12 - Moyenne de toutes les corrélations en fonction du nombre de points de temps
séparant deux points de temps pour chacune des électrodes
Figure 13 - Moyenne de corrélations significatives en fonction du nombre de corrélations
significatives pour chacun des nombres de points de temps séparant deux temps43
Figure 14 - Moyenne de toutes les corrélations significatives en fonction du nombre de points
de temps séparant deux points de temps43

Figure 15 - Moyenne de toutes les corrélations significatives en fonction du nombre de points Figure 16 - Graphiques à bande des moyennes de corrélations agrégées en fonction des mêmes nombres de points de temps séparant deux points de temps. Note : les 19 électrodes sont Figure 17 - Graphique des moyennes de corrélations entre les activités cérébrales des différents Figure 18 - Graphique des moyennes de corrélations entre les activités cérébrales des différents Figure 19 - Graphique des moyennes de corrélations entre les activités cérébrales des différents Figure 20 - Graphique à bande des corrélations entre les paires d'électrodes en lien avec FCz calculée selon les moyennes des paramètres des quatre modèles (SP(POW), SP(EXP), SP(EXPA) et SP(LEAR)) pour les positionnements tridimensionnels des électrodes fournis par Figure 21 - Graphique à bande des corrélations entre les paires d'électrodes en lien avec FCz calculée selon les moyennes des paramètres des quatre modèles (SP(POW), SP(EXP), SP(EXPA) et SP(LEAR)) pour les positionnements tridimensionnels des électrodes fournis par Figure 22 – Graphique à bande pour fin de comparaison des corrélations calculées selon les positionnements tridimensionnels provenant des deux articles (Böcker et al (1994) et Khosla et al. (1999)) et des corrélations de Pearson pour les quatre modèles différents (SP(POW), Figure 23 - Boîte à moustache des paramètres  $\rho$  estimés pour le modèle AR(1) ......101 Figure 24 - Boîte à moustache des paramètres  $\rho$  estimés pour le modèle ARMA(1,1)......103 Figure 25 - Boîte à moustache des paramètres  $\gamma$  estimés pour le modèle ARMA(1,1)......104

## Liste des abréviations

- $EEG: \acute{E} lectroenc\acute{e} phalogramme/\acute{E} lectroenc\acute{e} phalographie/\acute{E} lectroenc\acute{e} phalographique$
- MVAR : Modèle multivarié
- BIC : Critère d'information bayésien
- LRT : Test de ratio de vraisemblance
- UV : coordonnées bidimensionnelles (U,V)
- XYZ : coordonnées tridimensionnelles (X,Y,Z)

## **Avant-propos**

L'autorisation de rédiger ce mémoire a été obtenue par la direction du programme de M.Sc. de HEC Montréal.

## Remerciement

Tout d'abord, je tiens à remercier mon directeur Marc Fredette pour son support tout le long de mon projet de mémoire. Il m'a beaucoup aidé au niveau des théories statistiques et au niveau des méthodes pratiques qui permettent de mieux analyser les résultats. Malgré les impacts de la COVID-19, il reste quand même disponible pour répondre à toutes mes questions, ce qui est très responsable de sa part.

J'aimerais aussi remercier mon ancienne co-directrice Élise Labonte-LeMoyne. Elle a quitté ses fonctions au cours de mon projet, mais avant son départ, elle a quand même fait tout son possible pour m'aider à comprendre et à analyser les données EEG. Elle a aussi été présente lorsque j'ai des incompréhensions face aux fonctionnements du cerveau et les résultats inhabituels douteux obtenus avec les modélisations des données EEG.

De plus, j'aimerais remercier ma famille qui m'a beaucoup supporté émotionnellement et matériellement. Durant les moments difficiles, ma mère m'a beaucoup réconforté en cuisinant des plats délicieux. Mon père était aussi à mes côtés lorsque je n'avais pas de revenu salarial, ce qui a fait en sorte que je n'ai pas eu besoin de me stresser à cause des problèmes financiers. Je tiens à remercier mon minou de m'avoir accompagné tout le long de mon mémoire. Lorsque j'étais désespérée, les bisous qu'il me donne avec son petit nez me redonnent de la force.

Enfin, j'aimerais remercier mes amis et mes collègues de travail qui m'ont aussi supporté émotionnellement. Je tiens à surligner la disponibilité de Melissa Turp-Yonezawa lorsque j'ai besoin de prendre un moment pour changer mes idées. Il y a aussi Ching Hua Huang et Ryan Chen qui m'ont donné beaucoup réconfort lors de différentes activités en groupe. Également, j'aimerais remercier mon ancien petit copain Kaiyao Duan pour sa patience envers mes émotions changeantes lorsque j'étais stressée pour ce projet. Il a pris du temps à organiser des voyages avec moi pour me donner l'occasion de profiter de ma vie. Malgré la séparation, je n'oublierais pas les supports qu'il m'a offerts.

## Introduction

Le domaine d'utilisation des données EEG est très large. Elles peuvent être utilisées pour des recherches cliniques tout comme pour les recherches en marketing (Lai et al., 2018). Ces recherches s'intéressent souvent aux activités du cerveau en fonction des stimulus dont les participants reçoivent lors des différentes expériences. Dans notre recherche actuelle, contrairement aux types de recherches nommées ci-dessus, nous nous intéressons purement à la relation des différentes zones du cerveau lorsque le cerveau est en état de repos. Par repos, on considère que les participants n'ont aucune charge cognitive et ne reçoivent aucun stimulus. Les zones du cerveau sont représentées par les différentes électrodes utilisées dans la recherche. Pour évaluer les relations entre les zones du cerveau, une revue de littérature a été faite pour étudier les méthodes qui ont été utilisées auparavant à cette fin. Cette revue, qui est présentée dans la prochaine section, montre trois méthodes les plus utilisées pour ce genre d'analyse : le modèle autorégressif multivarié, la corrélation de Pearson et la cohérence. Parmi ces trois méthodes, le modèle autorégressif multivarié nous semble le plus intéressant, car ce modèle fait partie des modèles mixtes, des modèles qui permettent de modéliser les données à mesures répétées. Ces modèles mixtes sont souvent univariés, ce qui veut dire qu'il mesure seulement une mesure répétée. Tandis que les modèles mixtes multivariés permettent de modéliser plus qu'une mesure répétée. On aimerait donc en découvrir davantage à propos des modèles mixtes univariés étant donné que plusieurs études ont déjà été faites sur les modèles mixtes multivariés. Cette recherche a pour but de trouver les modèles mixtes les plus appropriés qui permettent de modéliser les relations entre les électrodes en utilisant les données EEG. Les données EEG ont deux aspects à analyser, la première étant le temps et la deuxième étant les électrodes. On a donc deux mesures répétées : temps et électrodes. Si on veut examiner les modèles univariés, on doit modéliser les deux mesures répétées, temps et électrodes, une à la fois. Pour analyser l'aspect du temps, les modèles mixtes permettant de modéliser les données répétées à distance uniforme seront utilisés, comme AR(1), ARMA(1), ARH(1), etc. Pour analyser l'aspect des électrodes, on aimerait utiliser la distance entre les électrodes afin de déterminer leurs relations, parce qu'une étude dirigée par Bhavsar et al. (2018) a démontré un lien entre la distance physique et la corrélation entre les électrodes. Les modèles mixtes spatiaux seront donc utilisés. Ces modèles permettent de modéliser les données répétées à distances variées et de tenir compte de la distance entre les électrodes. Afin de mieux faire les analyses, nous allons

commencer par faire une revue de littérature, suivi de la description des données. Ensuite les analyses descriptives seront effectuées pour mieux comprendre les données. La méthodologie sera décrite par la suite. Pour finir, les analyses et résultats des modèles mixtes sont présentés.

## **Revue de littératures**

Plusieurs recherches ont été menées, dans le passé, concernant le cerveau et les analyses des données EEG. Dans cette partie, une revue de celles-ci sera faite afin de mieux comprendre la structure du cerveau et la fonctionnalité des différentes zones cérébrales. De plus, les recherches ont été effectuées pour découvrir les méthodes qui ont été proposées par les études antérieures pour analyser la relation entre les différentes zones du cerveau selon les données EEG. Ces études passées nous permettront d'avoir une piste pour établir la méthodologie de notre recherche. Les prochaines sous-sections présenteront un sommaire de la revue de littératures qui a été effectuée pour l'étude actuelle.

### Structure du cerveau

Le cerveau humain est le système le plus complexe qu'on peut retrouver dans cet univers (Russel, 1979). Il est aussi le système nerveux central du corps humain (Javaid and al., 2020). C'est aussi la raison pour laquelle plusieurs études ont été menées pour mieux comprendre cet organe. Dans cette présente sous-section, les différentes caractéristiques du cerveau seront présentées pour mieux comprendre les fonctions du cerveau. Nous allons voir quelles sont les différentes zones qui composent le cerveau humain et les différents états du cerveau ainsi que ce qui les caractérise.

#### Les zones du cerveau

Le cerveau humain est composé de trois parties : le cerebrum, le cervelet et le tronc cérébral. Dans cette étude, on s'intéresse au cerebrum et plus précisément aux lobes qui sont à l'intérieur du cortex cérébral qui compose le cerebrum. En fait, le cerebrum est composé de deux hémisphères, l'hémisphère droit et l'hémisphère gauche. Ces hémisphères ont deux couches, la couche de l'extérieur qui est appelée cortex et la couche de l'intérieur. Il y a quatre lobes qui composent l'intérieur du cortex, le lobe frontal, le lobe pariétal, le lobe temporal et le lobe occipital (Jawabri & Sharma, 2021). Ces quatre lobes occupent des fonctions différentes et les positions de ceux-ci sont présentées dans la Figure 1. Les détails de ces quatre lobes seront expliqués dans cette sous-section.



Figure 1 – Les différents lobes du cerveau Note : Tirée du site de McGill (s.d) (<u>https://thebrain.mcgill.ca</u>), utilisation libre sous copyleft.

#### Lobe frontal

Le lobe frontal est le lobe le plus grand du cerveau (Jawabri & Sharma, 2019) et est situé derrière le front de l'être humain (Javaid and al.). Il occupe aussi des fonctionnalités importantes sur l'être humain. Entre autres les sens de planification, comme les plans de tous les jours et les plans de la vie future. Il s'occupe également du langage, du discours, de la personnalité, du contrôle de mouvement (Jawabri & Sharma, 2019), du mémoire et du contrôle d'impulsion et de jugement (Javaid and al.). Ce lobe est aussi celui qui permet aux êtres humains de prendre des décisions (Jawabri & Sharma, 2019) et de résoudre des problèmes (Javaid and al.). Selon une étude faite en 2012, un modèle nommé PROBE a été proposé pour décrire comment ce lobe opère lorsqu'il prend une décision. Le modèle PROBE montre que lorsque l'humain fait face à une situation, le cerveau peut penser à une solution d'une situation passée, créer une solution selon une solution utilisée dans le passé ou créer une solution complètement nouvelle. Ce modèle affirme que le cerveau est capable de faire la comparaison de trois à quatre solutions en même temps (Jawabri & Sharma, 2019).

#### Lobe pariétal

Le lobe pariétal se situe derrière le lobe frontal et par-dessus le lobe temporal. Ce lobe est séparé en deux régions différentes. La région du lobe antérieur permet de recevoir les signaux somatosensoriels et la région du lobe postérieur qui est lui-même composé de deux parties, le lobule pariétal supérieur qui est impliqué dans des fonctions d'ordre supérieur et le lobule pariétal inférieur contenant le cortex somatosensoriel secondaire qui reçoit des entrées somatosensorielles et transforment ces entrées avec d'autres modalités majeures pour former des fonctions d'ordre supérieur comme l'apprentissage, le langage, la capacité de comprendre les espaces en deux ou trois dimensions, la capacité de faire la différence entre les objets selon leurs caractéristiques (Jawabri & Sharma, 2019).

#### Lobe temporal

Le lobe temporal est situé derrière le lobe frontal et en dessous du lobe pariétal (Jawabri & Sharma, 2019). Cette partie du cerveau permet entre autres d'organiser les informations, du mémoire et des discours (Javaid and al., 2020). Ce lobe est composé de deux surfaces, la surface latérale et la surface médiale. La surface latérale du lobe s'occupe d'interpréter les informations sonores, comme donner du sens à des mots, et d'interpréter les informations visuelles, comme reconnaître un visage ou certains patterns. Il a aussi un système très important qui y situe, c'en est le réseau en mode par défaut. Lorsqu'on est en repos, ce réseau est activé et au contraire, lorsqu'on est en train de faire des activités qui requièrent de l'effort intellectuel, celui-ci est désactivé. Ce concept est important pour cette recherche et sera expliqué en détail plus loin dans le rapport. La couche médiale s'en charge de la mémoire déclarative qui est séparée en trois types : la mémoire sémantique, la mémoire de reconnaissance et la mémoire épisodique (Jawabri & Sharma, 2019).

#### Lobe occipital

Le lobe occipital est le lobe le plus petit du cerveau qui est situé en arrière de tous les lobes. Cette partie du cerveau permet principalement de traiter les informations visuelles (Jawabri & Sharma, 2019).

#### Les états du cerveau

Le cerveau humain a plusieurs états qui sont représentés par différents types d'onde. Cinq ondes cérébrales sont reconnues dans le monde scientifique. Ces ondes sont des mesures des activités électriques du cerveau. Ces ondes étant gamma, bêta, alpha, thêta et delta (Abhang & al., 2016). Les ondes sont mesurées en hertz, une mesure qui représente le nombre cycles d'oscillations par seconde (Zhang, 2019). La Figure 2 montre la forme des ondes des différentes ondes. Les états représentés par chacun des types d'ondes seront décrits dans les paragraphes suivants. Comme les fréquences et les caractéristiques des ondes émises par les différentes

régions du cerveau ne sont pas les mêmes (Abhang & al., 2016), les informations présentées dans les prochaines sous-sections sont généralisées.

#### **Ondes** gamma

Lorsqu'on observe des ondes gamma, le cerveau se retrouve dans un état de concentration (Abhang & al., 2016). Plus précisément, il apparaît lorsqu'on apprend, mémorise ou procède au traitement du langage (Ismail et al., 2016) . Les ondes gamma ont une fréquence qui est souvent plus grande que 27 Hz.

#### Ondes bêta

Lorsqu'on observe des ondes bêta, le cerveau humain est actif. Il est aussi dans un état d'anxiété dominante, d'attention externe (Abhang & al., 2016) ou de méfiance (Woaswi et Hamzah, 2018). Dans cet état, le cerveau est en train de chercher des solutions. Cet état est ce dont on a besoin pour étudier ou faire des examens (Woaswi et Hamzah, 2018). Généralement, cette onde prend la fréquence entre 12 Hz et 27 Hz (Woaswi et Hamzah, 2018).

#### Ondes alpha

Lorsqu'on observe des ondes alpha, le cerveau humain est très relaxe ou est dans un état d'attention passif (Abhang & al., 2016). Dans cet état, le cerveau n'est pas concentré sur certains événements en particulier, mais peut être attiré involontairement vers des événements dans l'environnement comme les bruits (Thorne et Thomas, 2021). Cet état du cerveau permet souvent à faire de la réflexion et est idéal pour de la créativité. L'onde alpha prend les fréquences entre 8 Hz et 12 Hz (Woaswi et Hamzah, 2018).

#### Ondes thêta

Lorsqu'on observe les ondes thêta, le cerveau est dans un état de relaxation profonde ou de méditation. Dans cet état, l'humain est en train de rêver dans un sommeil léger ou est en mode de soulagement de stress. Cet état permet d'augmenter la créativité et l'apprentissage. L'onde thêta prend les fréquences entre 3 Hz et 8 Hz (Woaswi et Hamzah, 2018).

#### **Ondes** delta

Lorsqu'on observe les ondes delta, le cerveau est en état de réparation et de sommeil profond. Cet état permet de libérer des hormones de croissance qui permet la guérison du corps humain. L'onde thêta prend les fréquences entre 0.5 Hz et 3 Hz (Woaswi et Hamzah, 2018).



Figure 2 - Les différentes ondes du cerveau Note : Tirée du site de McGill (s.d) (<u>https://thebrain.mcgill.ca</u>), utilisation libre sous copyleft.

## EEG

Le EEG est utile dans plusieurs domaines d'expertise. Par exemple, il est utilisé pour capter les maladies du cerveau dans le domaine médical. Dans les recherches de l'interface d'ordinateurcerveau, il est utilisé pour le contrôle des applications et la reconnaissance des objets. Il est aussi utilisé pour la réhabilitation de l'humain afin d'aider les patients qui rencontrent des difficultés à faire des mouvements. Son usage est également répandu en neuromarketing pour reconnaître les différentes réactions des consommateurs afin d'établir les meilleures stratégies marketing (Lai et al., 2018). Dans le but de mieux comprendre le fonctionnement de l'EEG, les prochaines sous-sections permettront d'étudier les notions générales de l'EEG, les positions des électrodes et la mesure utilisée par l'EEG.

#### Notions générales

L'EEG permet de collecter des données sur activités du cerveau à l'aide des électrodes déposées sur le cuir chevelu. Ces électrodes permettent de capter les flux électriques émis par le système nerveux du cortex. Les signaux électriques captés proviennent principalement des neurones de la corticale pyramidale. Les signaux peuvent être transmis du générateur de ceuxci aux électrodes grâce au principe de conduction volumique, le principe de faire circuler les courants électriques des tissus cérébraux vers le cuir chevelu par un dipôle de courant (Xia et Hu, 2019).

Il est assez difficile de déterminer la source du flux électrique, car l'EEG estime une réalité tridimensionnelle par une projection en deux dimensions (Olejniczak, 2006), ce qu'on appelle le problème inverse de l'EEG (Xia et Hu, 2019; Olejniczak, 2006). Les simulations computationnelles permettent d'estimer la source du flux, mais ne sont toujours pas efficaces pour représenter le cerveau humain (Xia et Hu, 2019). Pour mieux estimer la source, il est préférable de combiner l'analyse des données EEG avec d'autres sources d'informations comme les données captées par le MRI (Xia et Hu, 2019; Olejniczak, 2006).

Les données EEG sont plus complexes à analyser, car lors de la collecte, il y a plusieurs sources d'artefacts qui peuvent interférer dans l'analyse. Les artefacts étant les signaux électriques collectés par les électrodes, mais qui ne proviennent pas du cerveau. Il y a deux types d'artefacts, le premier type étant les artefacts physiologiques et le deuxième type étant les artefacts non physiologiques. Les exemples d'artefacts physiologiques sont les clignements des yeux, le battement de cœur, les respirations, etc. Les artefacts non physiologiques sont les signaux électriques dans l'environnement capter par les électrodes (Britton et al., 2016). Les artefacts sont donc des bruits qui font en sorte que ces données deviennent difficiles à analyser, car certaines ondes observées sont probablement dues aux artefacts. Les processus de nettoyages sont donc nécessaires pour faire des analyses pertinentes. On peut supprimer les époques ayant des artefacts ou pratiquer, dès le départ, les bonnes pratiques de collecte pour baisser la présence d'artefacts en mettant plus d'attention sur l'environnement de collecte, du système d'enregistrement et des sujets (Xia et Hu, 2019). Ici, par « époques », on veut dire les périodes définies en tant que les moments où il y a des stimulus pendant l'expérience.

#### Positions des électrodes

Les électrodes sont posées sur le cuir chevelu pour capter les signaux électriques émis par le cerveau. En observant la Figure 3, on retrouve les électrodes qui permettent de capter les activités de la zone anteriofrontale (Fp1 et Fp2), la zone frontale (F3, F4, F7 et F8), la zone centrale (C3 et C4), la zone pariétale (P3, P4, P7 et P8) et la zone centrale (Fpz, Fz, FCz, Pz et Oz). Les casques EEG peuvent contenir entre 4 et 256 électrodes pour capter les activités du cerveau (Lau, Gwin et Ferris, 2012). Dans notre expérience, les électrodes qui ont été utilisées sont : C3, C4, F3, F4, F7, F8, FCz, Fp1, Fp2, Fpz, Fz, O1, O2, Oz, P3, P4, P7, P8 et Pz. La Figure 3 représente la position de toutes les électrodes utilisées dans notre expérience.



Figure 3 - Les positions des électrodes

#### Mesures

Les signaux électriques EEG sont collectés en microvolt et ensuite transformés par des amplificateurs en voltages appropriés (Xia et Hu, 2019). En effet, les oscillations des activités électriques présentent deux informations importantes, l'intensité et la fréquence électrique. Le voltage mesure l'intensité des signaux électriques qui sont captés par les électrodes. Dans le cadre de notre recherche, c'est la fréquence qui est analysée. La fréquence définie

précédemment est mesurée en hertz. C'est donc le nombre d'oscillations par seconde de l'activité électrique qui nous intéresse (« EEG : Introduction - The McGill Physiology Virtual Lab », s.d.).

# Méthodes utilisées pour les analyses de corrélations des données EEG

En étudiant plusieurs articles qui ont analysé la connectivité entre les zones du cerveau, trois méthodes d'analyse ont été identifiées. La première étant les analyses autorégressives multivariées. La deuxième étant les calculs des corrélations de Pearson. La troisième étant les analyses de cohérence. Dans les prochaines sous-sections, celles-ci seront décrites en détail pour mieux comprendre leurs fonctionnements.

#### Modèle autorégressif multivarié (MVAR)

#### Théorie

Le modèle autorégressif multivarié est un variant de la structure autorégressive. Il est défini par l'équation suivante :

$$y_n = \sum_{i=1}^m y_{n-i} A(i) + e_n (1)$$

Dans cette équation,  $y_n$  représente les *n* échantillons de *d* série de temps. A(i) est une matrice de coefficient (tout comme les  $\beta$  qu'on utilise dans les modèles linéaires) de *d* par *d* dimensions.  $e_n$  représente le bruit de Gaussian additive. Ce modèle est souvent utilisé pour estimer la connectivité fonctionnelle du réseau de neurones du cerveau. La structure autorégressive est utilisée pour les cas où la mesure longitudinale est univariée. Dans le cas des analyses EEG, il y a plusieurs électrodes qui captent les activités des différentes zones du cerveau aux différents temps, ce qui fait que ces données captées par les électrodes deviennent dépendantes entre elles. Lorsque plusieurs variables ayant des observations dépendantes et répétées sont présentes dans les analyses, les modèles autorégressifs univariés ne suffissent plus de modéliser ce genre de données. Le modèle autorégressif multivarié permet alors de modéliser les données EEG contenant *d* séries de temps qui sont associés à *d* nombre de variables. Dans notre analyse, on peut considérer que les d variables sont celles qui identifient les d différentes électrodes. Il est aussi important de trouver l'ordre m qui modélise le mieux les différentes données. En fait, l'ordre du modèle indique le nombre de données passées qui est nécessaire pour prédire la prochaine valeur. Par exemple, dans notre cas, on aimerait savoir quelle est l'activité du cerveau au temps t pour une électrode donnée et on présume qu'il faut avoir au moins les données du cerveau des cinq temps avant t, l'ordre du modèle est de m = 5. Le MVAR considère que chacune des séries de temps est dépendante de toutes les autres. Lorsque le test indique une connectivité significative, on peut inférer un sous-réseau. Ces sous-réseaux suivent le principe de causalité de Granger qui est défini comme suit : les activités de la région X cause les activités de l'Y lorsque toutes les connectivités de X à Y, à travers le temps, sont significativement non nulles (Penny et Harrison, 2007).

#### **Recherches antérieures**

Une partie de la méthodologie d'une étude sur la connectivité fonctionnelle frontopariétal lors de la mémoire à court terme menée par Babiloni et al. (2004) a utilisé le modèle MVAR afin d'estimer la direction du flux d'information entre les électrodes frontopariétal F3 – P3 et F4 – P4. Le modèle utilisé dans cette recherche est d'ordre 7 et celui-ci est calculé à l'aide du logiciel MATLAB. Seulement les ondes bêta et gamma sont utilisés dans cette étude. Les résultats, obtenus en combinant plusieurs autres méthodes d'analyse de connectivité, prouvent que la zone frontale permet d'augmenter l'influence sur la zone pariétale lors de la mémorisation à court terme.

Cette méthode est aussi utilisée dans une recherche, menée par Omidvarnia et al. (2011), qui s'intéressent à la connectivité entre les différentes régions du cortex des nouveau-nés. Cet article a pour but de comparer deux méthodes de connectivité (cohérence pariétale dirigée et fonction de transfert dirigée) à base du modèle MVAR. En bref, une partie des résultats démontre que le MVAR variant dans le temps pour estimer la connectivité des différentes régions du cortex permet de mieux comprendre les relations non symétriques entre les électrodes EEG.

Enfin, selon un article écrit par Greenblatt et al. (2012) portant sur les différentes méthodes d'analyse de connectivité entre les zones du cerveau, le MVAR est fréquemment utilisé dans différentes recherches impliquant les données EEG et la connectivité entre les électrodes. Par

exemple dans les applications cliniques, comme l'étude menée par Dauwels et al. (2010) sur le diagnostic précoce que la maladie Alzheimer. Greenblatt et al. (2012) a aussi cité un exemple d'analyse de données continues menée par Zhao et al. (2011) sur les algorithmes d'apprentissage du noyau de classification de la fatigue mentale au volant basé sur l'analyse des données EEG par la modélisation du MVAR. Un exemple en lien avec l'analyse des données EEG relatives aux événements à l'aide du MVAR a également été cité dans l'article de Greenblatt et al. (2012).

#### Corrélation de Pearson

#### Théorie

Les corrélations de Pearson entre deux variables sont calculées par cette équation (Berman, 2018) :

$$r = \frac{\sum_{i=1}^{n} (x_i - \bar{x}) (y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n} (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^{n} (y_i - \bar{y})^2}}$$
(2)

Dans cette équation,  $x_i$  représente  $i^{jeme}$  valeur de la première variable et  $\bar{x}$  représente la moyenne de l'ensemble des valeurs de la première variable évaluée et  $y_i$  représente  $i^{jeme}$  valeur de la deuxième variable et  $\bar{y}$  représente la moyenne de l'ensemble des valeurs de la deuxième variable évaluée. Elles sont souvent utilisées pour estimer la relation entre deux variables numériques. Ces corrélations prennent les valeurs entre -1 et 1. Lorsque la corrélation est de 0, cela signifie qu'il n'y a pas de lien linéaire entre les deux variables. Plus que la valeur de la corrélation est proche de -1 ou de 1, plus que la relation entre les deux variables est élevée. Si la valeur est significativement différente de 0 et plus grande que 0, on a donc une corrélation positive, ce qui signifie que lorsque la première variable augmente, la deuxième variable augmente, aussi. Si la valeur est significativement différente de 0 et plus petite que 0, on a donc une corrélation négative, ce qui signifie que lorsque la première variable augmente, la deuxième variable augmente, la deuxième variable diminue (Nettleton, 2014). Dans notre cas, les deux variables d'intérêt pour l'évaluation de la corrélation seraient des paires d'électrodes qui ont collecté des données du cerveau.

#### **Recherches antérieures**

Dans un article écrit par Choong et al. (2021), on retrouve une évaluation de la relation linéaire entre les paires d'électrodes à l'aide des coefficients de Pearson pour les fréquences associées aux ondes alpha, bêta et gamma. Cette analyse est faite dans le but de déterminer le lien entre ces relations et les différentes émotions évoquées par les sujets. Les corrélations trouvées pour les fréquences alpha sont plus élevées que pour celles pour les fréquences bêta et gamma. Les paires d'électrodes entre lesquelles les corrélations sont les plus élevées pour les fréquences alpha sont FC6 – F4 et F4 – AF4. Pour les fréquences bêta et gamma, les paires d'électrodes ayant les corrélations les plus élevées sont FC6 – F4 et O2 – P8.

Une recherche conduite par Chen, Ros et Gruzelier (2013) a aussi utilisé la corrélation de Pearson pour déterminer la relation entre les différentes zones du cerveau pour le réseau de mode de repos. Pour cette recherche, les électrodes ont été regroupées ensemble pour représenter les différentes zones. Les résultats démontrent que les corrélations de Pearson sont plus élevées dans les mêmes zones.

Une autre étude a aussi utilisé les corrélations de Pearson dans une partie de l'analyse. En effet, Islam et al. (2021) ont conduit une étude pour produire un modèle de reconnaissance d'émotions basé sur la corrélation entre les électrodes. Dans cette expérience, la première étape de l'analyse est de faire une analyse de corrélation de Pearson pour toutes les paires possibles des 32 électrodes afin de produire une matrice de corrélation.

#### Analyse de cohérence

#### Théorie

Dans la section précédente, les corrélations de Pearson ont été introduites comme une méthode qui est souvent utilisée dans les analyses des connectivités entre les électrodes EEG. L'analyse de cohérence est relativement similaire à la corrélation de Pearson et celle-ci est calculée selon l'équation suivante :

$$C_{xy}(\lambda) = \frac{|P_{xy}(\lambda)|^2}{P_{xx}(\lambda)P_{yy}(\lambda)}$$
(3)

Cette équation calcule la cohérence entre les spectres où  $\lambda$  représente les différents canaux de fréquence. Plus précisément,  $\lambda$  est la fréquence que l'on choisit dans chacun des deux signaux, x et y, pour calculer leur cohérence.  $P_{xy}(\lambda)$  représente la densité spectrale croisée entre les deux signaux x et y.  $P_{xx}(\lambda)$  ou  $P_{yy}(\lambda)$  représentent les densités autospectrales (Meyer Carson et al., 2014). On peut dire que la densité spectrale croisée est la corrélation entre deux spectres et la densité autospectrale est la corrélation d'un spectre avec lui-même (Welch, 1967). En effet, la formule mathématique de la cohérence est une extension de la corrélation de Pearson. La valeur de cohérence prend la valeur entre 0 et 1 et permet de prédire linéairement la fréquence d'une électrode à partir d'une autre électrode. Si la valeur de la cohérence se rapproche de 1, cela signifie qu'il y a une relation linéaire entre deux électrodes étudiées et si elle se rapproche de 0, cela signifie que les deux électrodes étudiées n'ont pas de relation linéaire. Comme les données de la cohérence ne sont pas normalement distribuées, le coefficient de corrélation est donc normalisé pour obtenir le coefficient de cohérence (Meyer Carson et al., 2014). La cohérence, comparée à la corrélation, est qu'elle dépend du domaine de la fréquence, ce qui est équivalent à la corrélation croisée dans le domaine du temps (Bastos et Schoffelen, 2016).

#### **Recherches antérieures**

Une étude, conduite par Lorist et al. (2009), sur l'influence de la fatigue mentale et de la motivation sur le réseau dynamique neutre, a utilisé l'analyse de cohérence comme la méthode d'analyse principale. La cohérence est calculée entre toutes les paires possibles d'électrodes. Les résultats démontrent qu'en général, lorsque la fréquence est basse, la cohérence est significativement plus élevée. On retrouve aussi que la cohérence entre O1 et O2 est relativement plus élevée pour toutes les bandes de fréquence. De plus, la cohérence interhémisphérique au niveau fronto-central est relativement basse. Il y a plusieurs autres résultats importants, mais ne seront pas reportés dans ce rapport dans le but de garder la consistance.

Une recherche sur la comparaison de la cohérence des zones du cerveau entre deux types de TDAH mené par Barry et al. (2002) a aussi choisi cette méthode. Huit paires d'électrodes intrahémisphériques et huit paires d'électrodes interhémisphériques sont analysées pour estimer la cohérence. Les résultats démontrent des différences de cohérences significatives entre les enfants ayant le TDAH et les sujets contrôlent. Les détails des résultats sont massifs et ne seront pas reportés dans ce rapport.

# Chapitre 1 Présentation des données

Les données EEG font partie de celles qui sont les plus difficiles à être traitées. Il est donc important de bien faire le nettoyage des données et le traitement des données afin que les données initialement collectées soient utilisables et puissent offrir des informations fiables lors de l'analyse. Dans les prochaines sous-sections, la description de la collecte et le nettoyage des données EEG seront présentés. Ensuite, les traitements des données nettoyées effectués particulièrement pour cette recherche seront décrits. Enfin, les critères de choix de données utilisées pour cette recherche seront expliqués dans la dernière sous-section.

#### Collecte des données

Les données proviennent d'une expérience qui s'est déroulée au Tech3Lab. L'expérience est composée de deux parties et les données utilisées sont celles de la première tâche de la première partie. Cette première partie permet d'entraîner un système AI. On utilise des tâches standardisées pour mesurer les réactions dans le cerveau. Cette première partie contient trois tâches. La première tâche étant le N-back. Cette tâche étant d'identifier si la lettre qui est présentée à l'écran de l'ordinateur est la même que la *n* lettre précédente (Kirchner, 1958). Au fur et à mesure que *n* augmente, la difficulté de la tâche augmente, ce qui augmente la charge cognitive du participant. En fait, cette première tâche contient quatre niveaux de difficulté. Au premier niveau, le participant doit identifier la lettre qui est la même que la précédente (n = 1). Au deuxième niveau, le participant doit identifier la lettre qui est la même que la deuxième lettre précédente (n = 2). Au troisième niveau, le participant doit identifier la lettre qui est la même que la troisième lettre précédente (n = 3). Au quatrième niveau, le participant doit identifier la lettre précédente (n = 4). En faisant cela, le participant doit faire un plus gros effort à chaque niveau de difficulté, ce qui augmente la charge cognitive de celui-ci.

## Manipulation des données

#### Nettoyage des données EEG collectées

Après la collecte des données, les assistants de recherche du Tech3Lab ont procédé à un nettoyage de données afin qu'elles soient utilisables dans les recherches futures. Les données brutes des signaux EEG ont été filtrées par deux types de filtres. Le premier étant un filtre passe-bande qui laisse seulement passer les signaux entre 0.5 Hz et 100 Hz. Le deuxième étant un filtre coupe-bande qui permet d'atténuer les signaux qui sont à 60 Hz, car les signaux électriques captés provenant de l'environnement sont de 60 Hz en Amérique du Nord. Les signaux enregistrés qui sont considérés comme défectueux et qui ont été rejetés correspondent aux critères suivants : si l'intensité dépasse une différence potentielle de 100uV/200ms ou si elle est inférieure à 0.5uV/100ms. Les électrodes qui ont 10% de signaux qui correspondent au critère ci-dessus seront considérées défectueuses et sont rejetées. C'est la raison pour laquelle on ne retrouve pas toutes les électrodes pour tous les participants. Voici une liste de nombre de participants par électrode présentée dans le Tableau 1, sachant qu'il y a en tout 25 participants dont l'électrode F8 n'était pas rejetée par les critères, car cette électrode est particulièrement sensible aux bruits et est facilement rejetée.

Electrode	Nombre de participants ayant cette électrode
C3	23
C4	24
F3	23
F4	22
F7	17
F8	5
FCz	21
Fp1	25
Fp2	25
Fpz	25

 Tableau 1 - Nombre de participants ayant les électrodes non éliminées selon les règles de nettoyages des données EEG
Fz	25
01	24
O2	25
Oz	24
P3	25
P4	25
P7	25
P8	23
Pz	23

#### Manipulation des données nettoyées

Cette section permet d'expliquer les manipulations des données qui ont été faites pour cette recherche. Au départ, les données reçues sont séparées par participant dans des fichiers CSV différents. Comme on a 25 participants, on a donc 25 fichiers CSV différents pour chacun des participants. Pour chaque participant, on a les fréquences sur des lignes différentes pour chacune des électrodes. On a donc, dans chaque colonne, les fréquences pour chaque point de temps. Chaque point de temps représente 1/256 secondes. Comme chaque participant passe environ une demi-heure sur la tâche, il y a donc environ 460 000 colonnes de données lors de l'importation des données dans le logiciel SAS. Cependant, on ne peut pas utiliser ce genre de format pour faire les analyses. En effet, les analyses de cette recherche ont besoin que la variable représentant les données à modéliser soit contenue dans une seule colonne. On doit donc faire une transposition de ces données de fréquences de manière verticale. Suite à cette transformation, toutes les fréquences sont contenues dans une colonne nommée « fréquence ». Les noms des électrodes sont contenus dans une autre colonne nommée « électrode » et ils se répètent tout au long de la variable « fréquence ». Les points de temps forment aussi une variable et ils se répètent tout au long de la variable « fréquence » et de la variable « électrode ». Lorsque toutes les tables de tous les participants sont transposées verticalement, on a concaténé tous les participants pour former une table complète de toutes les données tout en identifiant les données de chacun des participants par leurs numéros dans une nouvelle colonne.

### Données des positions d'électrodes

Lors des analyses des activités entre les zones du cerveau, on aura à utiliser les distances entre les électrodes pour modéliser les modèles spatiaux. Donc, il est nécessaire de connaître les positions des différentes électrodes. Après avoir fait les recherches, les positions des électrodes fournies dans les deux articles de Khosla & al. (1999) et de Böcker & al. (1994) sont utilisées dans nos analyses. Les positions qui sont fournies dans ces deux articles sont en coordonnées angulaires, qui est également reconnue comme coordonnées polaires dont les coordonnées sont mesurées par la déviation de l'emplacement de l'électrode en angle en termes d'azimut et de polaire. Dans notre cas, l'angle azimutal est l'angle au niveau horizontal. Donc, on peut dire entre le nasion et l'oreille de droite, on a un angle azimutal de 90° et entre le nasion et l'oreille de gauche, on a un angle azimutal de 270°. L'angle polaire est celui au niveau vertical. Donc, on peut dire qu'entre le centre du crâne et l'oreille de droite on a un angle polaire de 90°. Il est à noter que la définition de l'angle azimutal ou polaire est relative à la perception de différents chercheurs. Il est donc possible que les deux angles soient inversés, mais cela ne change pas le résultat de l'analyse. Cependant, le type de coordonnées que nous avons besoin pour faire les modélisations dans le logiciel SAS sont des coordonnées cartésiennes. Pour ce faire, il faut faire quelques transformations. Même si les positions en coordonnées cartésiennes tridimensionnelles sont déjà fournies dans l'article de Böcker et al. (1994), on veut quand même être conforme dans les transformations des données. Donc, ces transformations sont toutes basées sur les coordonnées angulaires.

Les transformations ont été faites en deux étapes. La première est la transformation de coordonnées angulaires en coordonnées cartésiennes tridimensionnelles. Les coordonnées angulaires provenant des deux articles sont transcrites dans les Tableau 2 et Tableau 3. Il est à noter que les positions des électrodes non présentes dans les deux articles ont été estimées selon la méthode internationale 10-20 de positionnement des électrodes, car les positions fournies par les deux articles ont été estimées de cette manière. Cette méthode consiste à placer les électrodes à une distance de 10% ou 20% de la longueur totale entre le front et l'arrière du crâne ou de la longueur totale entre le côté droit et le côté gauche du crâne (Jasper, 1958). Les électrodes identifiées en rouge sont celles dont les positions ne sont pas présentes dans l'article et estimées selon la règle de positionnement 10-20.

Électrode	Polaire	Azimut
C3	45	180
C4	45	0
Fz	45	90
FCz	22.5	90
Fpz	90	90
Pz	45	270
Oz	90	270
F3	59	129
F4	59	51
F7	90	144
F8	90	36
Fp1	90	108
Fp2	90	72
P3	59	231
P4	59	309
P7	90	216
P8	90	324
01	90	252
02	90	288

### Tableau 2 - Coordonnées angulaires selon les positionnements fournis par Khosla et al. (1999)

Tableau 3 - Coordonnées angulaires selon les	positionnements fournis par Böcker et al.
(1994	)

Électrode	Polaire	Azimut
C3	36	90
C4	36	270
Fz	36	0
FCz	18	0
Fpz	72	0
Pz	36	180
Oz	72	180
F3	47.7	38.36
F4	47.7	321.64
F7	72	54
F8	72	306
Fp1	72	18
Fp2	72	342
P3	47.7	141.64
P4	47.7	218.36
P7	72	126
P8	72	234
01	72	162
O2	72	198

La méthode de transformation en coordonnées tridimensionnelles provient de la section « I. Spherical Coordinates » de l'article de Deserno (2004). Les Tableau 4 et Tableau 5 présentent les positionnements cartésiens tridimensionnels calculés selon le positionnement angulaire provenant des deux articles.

Électrode	X	Y	Ζ
C3	-0.70711	0	0.707107
C4	0.707107	0	0.707107
Fz	0	0.707107	0.707107
FCz	0	0.382683	0.92388
Fpz	0	1	0
Pz	0	-0.70711	0.707107
Oz	0	-1	0
F3	-0.53943	0.666144	0.515038
F4	0.539433	0.666144	0.515038
F7	-0.80902	0.587785	0
F8	0.809017	0.587785	0
Fp1	-0.30902	0.951057	0
Fp2	0.309017	0.951057	0
P3	-0.53943	-0.66614	0.515038
P4	0.539433	-0.66614	0.515038
P7	-0.80902	-0.58779	0
P8	0.809017	-0.58779	0
01	-0.30902	-0.95106	0
O2	0.309017	-0.95106	0

Tableau 4 - Coordonnées tridimensionnelles selon les positionnements fournis par Khosla et al. (1999)

#### Tableau 5 - Coordonnées tridimensionnelles selon les positionnements fournis par Böcker et al. (1994)

Électrode	X	Y	Ζ
C3	0	0.587785	0.809017
C4	0	-0.58779	0.809017
Fz	0.587785	0	0.809017
FCz	0.309017	0	0.951057
Fpz	0.951057	0	0.309017
Pz	-0.58779	0	0.809017
Oz	-0.95106	0	0.309017
F3	0.579965	0.459015	0.673013
F4	0.579965	-0.45902	0.673013
F7	0.559017	0.769421	0.309017
F8	0.559017	-0.76942	0.309017

Fp1	0.904508	0.293893	0.309017
Fp2	0.904508	-0.29389	0.309017
P3	-0.57996	0.459015	0.673013
P4	-0.57996	-0.45902	0.673013
P7	-0.55902	0.769421	0.309017
P8	-0.55902	-0.76942	0.309017
O1	-0.90451	0.293893	0.309017
02	-0.90451	-0.29389	0.309017

Après avoir fait la transformation de coordonnées angulaires aux coordonnées cartésiennes tridimensionnelles, une transformation en coordonnées bidimensionnelles (les coordonnées sur un plan 2D, celles-ci sont nommées U et V) a été faite. Comme les électrodes se situent sur la surface du crâne et que l'on considère que le crâne est sphérique, il est donc possible d'effectuer une projection sur un plan ovale grâce à la méthode de positionnement UV en utilisant les formules 1 et 2 de l'article de Yavartanoo et al. (2019). Cette méthode est la même que celle utilisée pour mettre le globe terrestre sur un plan bidimensionnel.

Les positionnements UV transformés selon les positionnements cartésiens tridimensionnels sont présentés dans les Tableau 6 et Tableau 7. Il est à noter que les électrodes Fpz et Oz pour les positionnements UV basés sur l'article de Khosla et al. ont été réajustées étant donné que la transformation à l'aide des formules entre les coordonnées tridimensionnelle et bidimensionnelle pour les points où x et z sont à 0 et y est à 1 ou -1 (donc les points centraux du pôle nord et sud) ne les alignent pas totalement au centre. On a donc remis ces positions au centre pour s'aligner avec les électrodes FCz et Fz, puisque les électrodes Fpz et Oz sont sur la même ligne que FCz et Fz. Pour mieux comprendre, on peut aller voir un globe sur un plan ovale, les points centraux du globe au niveau du pôle Sud et au niveau du pôle Nord sont dispersés en plusieurs points au niveau horizontal. C'est la raison pour laquelle ces deux électrodes ont été replacées aux bonnes places.

Tableau 6 - Coordonnées bidimensionnelles selon les positionnements fournis pa	r
Khosla et al. (1999)	

Électrode	U	V
C3	0.875	0.5
C4	0.625	0.5
Fz	0.75	0.25
FCz	0.75	0.375
Fpz	0.75	0
Pz	0.75	0.75

0.75	1
0.878681	0.267944
0.621319	0.267944
1	0.3
0.5	0.3
1	0.1
0.5	0.1
0.878681	0.732056
0.621319	0.732056
1	0.7
0.5	0.7
1	0.9
0.5	0.9
	0.75 0.878681 0.621319 1 0.5 0.878681 0.621319 1 0.5 1 0.5 1 0.5

#### Tableau 7 - Coordonnées bidimensionnelles selon les positionnements fournis par Böcker et al. (1994)

Électrodes	U	V
C3	0.75	0.3
C4	0.75	0.7
Fz	0.65	0.5
FCz	0.7	0.5
Fpz	0.55	0.5
Pz	0.85	0.5
Oz	0.95	0.5
F3	0.636797	0.348202
F4	0.636797	0.651798
F7	0.58037	0.220545
F8	0.58037	0.779455
Fp1	0.552395	0.405049
Fp2	0.552395	0.594951
P3	0.863203	0.348202
P4	0.863203	0.651798
P7	0.91963	0.220545
P8	0.91963	0.779455
O1	0.947605	0.405049
O2	0.947605	0.594951

Le fait d'avoir choisi les deux types de coordonnées tridimensionnelles et bidimensionnelles a pour but de tester deux types de distances différents afin de comprendre comment les relations entre les électrodes sont affectées par les distances entre elles. En effet, si on prend les coordonnées cartésiennes tridimensionnelles, on aura les distances entre les électrodes tout en traversant le crâne. Dans le cas des coordonnées cartésiennes bidimensionnelles, on aura donc les distances entre les électrodes qui passent par la surface du crâne.

En ayant ces deux types de coordonnées provenant des positions d'électrodes fournies dans les deux articles présentés précédemment, on a la possibilité d'intégrer les informations sur les positions des électrodes dans les analyses des modèles mixtes tout en modélisant les distances entre les électrodes.

### Choix des données

Pour les analyses de cette recherche, on ne veut pas prendre les données collectées durant la réalisation des tâches, car on ne veut pas que celles-ci aient un effet sur notre analyse étant donné que notre but est d'analyser, de manière généralisée, les relations entre les différentes zones du cerveau et entre les différents moments du cerveau. Dans ce cas, il faut seulement utiliser les données aux moments des pauses entre les différentes tâches de l'expérience. En effet, durant la réalisation des tâches, il y a des stimulus qui sont communs à tous les participants, ce qui fait que les réactions et les activités du cerveau peuvent être similaires pour chaque participant. Contrairement aux moments des réalisations des tâches, les moments de pause n'ont pas de stimulus particuliers, ce qui fait en sorte qu'il n'y a rien de commun entre les activités du cerveau de chacun des participants. Ces moments de pause sont connus comme étant l'état de repos du cerveau. Au moment de l'état de repos du cerveau, le réseau de mode par défaut est actif (Ekhtiari et al. 2016). C'est donc ce qui nous intéresse, car à l'état de repos, aucune activité du cerveau n'est déclenchée par une tâche particulière (Banz et al., 2016).

Entre chaque tâche, les assistants de recherche demandent aux participants de prendre une pause et de relaxer. Il y a trois pauses pour chaque participant. On a donc supprimé toutes les fréquences qui ne sont pas collectées durant ces périodes de pause. Ces pauses sont identifiées par 1, 2 ou 3 dans notre base de données et les temps de ces pauses sont reclassés à partir de temps 1 pour le premier point de temps de chacune des pauses pour faciliter les analyses. La première pause a une durée entre 57 secondes et 2 minutes 31 secondes selon les participants. La troisième pause a une durée entre 41 secondes et 2 minutes 26 secondes selon les participants. La troisième pause a une durée entre 44 secondes et 2 minutes et 1 seconde selon les participants.

### Chapitre 2 Analyse descriptive

Dans cette partie, nous allons explorer les données et en faire une analyse descriptive pour mieux comprendre la structure des données et pour avoir une idée sur les tendances de cellesci. De cette manière, ces analyses descriptives permettront aussi d'appuyer les résultats finaux. Pour ce faire, les corrélations de Pearson ont été utilisées. En effet, comme vu précédemment, la corrélation de Pearson est souvent utilisée comme méthode d'analyse. Il est donc un bon choix de l'utiliser pour explorer les données. Il y a trois aspects à cette analyse : l'analyse des relations entre les électrodes, l'analyse des relations entre les points de temps et l'analyse des relations entre les différentes électrodes aux différents moments dans le temps. Tous les résultats ont été produits par le logiciel SAS.

### Analyse de corrélation de Pearson

### Analyse des relations entre les électrodes

Cette partie de l'analyse descriptive porte sur les corrélations entre les électrodes. On aimerait savoir quelles sont les relations, positives ou négatives, entre les différentes parties du cerveau en utilisant les données sur les activités captées par les électrodes. Pour pouvoir faire cette analyse, les corrélations de Pearson doivent être générées pour des points de temps fixes, c'està-dire que pour chaque matrice de corrélation de Pearson entre les électrodes, il faut choisir les données des fréquences des électrodes à un temps donné.

Il y a deux parties dans cette analyse. Pour la première partie de cette analyse, on ne tiendra pas compte de la significativité de ces corrélations, car nous voulons avoir le maximum de corrélations pour analyser les tendances de la relation entre les différentes électrodes. Si on tenait compte de la significativité, on risque d'éliminer une grande partie des corrélations générées, ce qui ne permettra pas d'analyser les tendances de manière complète. Dix moments dans chaque pause ont été fixés pour l'analyse. Donc, trente moments ont été choisis. En fait, il se peut que la relation entre les activités des différentes zones du cerveau soit légèrement différente pour les différents moments de temps. En prenant tous ces différents moments, on aura une vue plus globale de différents moments dans le temps. La variable « fréquence » a été utilisée pour produire les matrices de corrélations de Pearson pour ces trente moments et ces matrices sont moyennées en une seule matrice.

Pour mieux illustrer les corrélations moyennes en fonction des paires d'électrodes, une mise en forme conditionnelle a été faite sur ces valeurs. La règle de mise en forme est la suivante :

ſ							(	1				
Si la valeur	>= ~	-0.99	Nombre 🗸 🤤	et	<	0	Nombre 🗸	alors	$\sim$	$\uparrow$	$\downarrow$	×
Si la valeur 🛛	>= ~	0	Nombre 🗸 e	et	< ~	0.2	Nombre 🗸	alors	$\sim$	$\uparrow$	$\downarrow$	×
Si la valeur	>= ~	0.2	Nombre 🗸 e	et	< ×	0.4	Nombre 🗸	alors	$\sim$	$\uparrow$	$\downarrow$	×
Si la valeur 🛛	>= ~	0.4	Nombre 🗸 e	et	< ~	0.6	Nombre 🗸	alors	$\sim$	$\uparrow$	$\downarrow$	×
Si la valeur	>= ~	0.6	Nombre 🗸 e	et	< ~	0.8	Nombre 🗸	alors	$\sim$	$\uparrow$	$\downarrow$	×
Si la valeur	>= ~	0.8	Nombre 🗸 e	et	< ~	1	Nombre 🗸	alors	$\sim$	$\uparrow$	$\checkmark$	×

Tableau 8 - Mise en forme conditionnelle des corrélations dans la matrice (tableau 9)

À titre d'explication, on a la couleur mauve pour les valeurs de corrélation moyenne négatives, la couleur rouge pour la moyenne de corrélation plus grande ou égal à 0 et plus petit que 0.2 et ainsi de suite.

21 	C4	F3	F4	F7	F8	FCz	Fp1	Fp2	Fpz	Fz	01	02	Oz	P3	P4	P7	P8	Pz
C3	0,08	0,46	0,14	0,46	0,10	0,23	0,28	0,18	0,22	0,26	0,31	0,31	0,30	0,66	0,20	0,44	0,25	0,27
C4		0,10	0,37	0,21	0,37	0,23	0,11	0,24	0,18	0,21	0,21	0,25	0,23	0,16	0,61	0,20	0,42	0,25
F3			0,24	0,61	0,34	0,49	0,44	0,41	0,53	0,55	0,10	0,15	0,10	0,21	0,06	0,19	0,14	0,08
F4				0,30	0,74	0,53	0,28	0,44	0,41	0,47	0,11	0,14	0,12	0,03	0,20	0,10	0,28	0,00
F7					0,43	0,35	0,57	0,38	0,54	0,44	0,28	0,31	0,29	0,41	0,26	0,40	0,27	0,20
F8						0,39	0,46	0,51	0,49	0,49	0,13	0,43	0,31	0,05	0,50	0,28	0,51	0,20
FCz							0,44	0,48	0,57	0,78	0,01	0,09	0,03	0,01	0,09	0,05	0,16	-0,16
Fp1								0,43	0,72	0,53	0,10	0,16	0,12	0,13	0,11	0,14	0,23	0,06
Fp2									0,65	0,55	0,07	0,21	0,13	0,08	0,20	0,15	0,30	0,10
Fpz										0,63	0,11	0,18	0,15	0,09	0,16	0,14	0,27	0,05
Fz											0,07	0,15	0,10	0,08	0,12	0,12	0,21	-0,05
01												0,68	0,87	0,58	0,45	0,71	0,52	0,62
O2													0,86	0,44	0,61	0,53	0,71	0,62
Oz														0,51	0,57	0,63	0,64	0,69
P3															0,33	0,73	0,36	0,56
P4																0,39	0,75	0,59
P7																	0,45	0,55
P8																		0,50

Tableau 9 - Matrice de corrélations de Pearson des paires d'électrodes

Comme on peut voir dans cette matrice (Tableau 9), les corrélations moyennes positives les plus élevées (vert et bleu fluo) sont associées aux paires d'électrodes qui sont relativement proches entre elles et que les corrélations moyennes positives les plus faibles (rouge) sont associées aux paires d'électrodes qui sont relativement loin entre elles. Dans la Figure 4, on a les corrélations entre l'électrode FCz, l'électrode la plus au centre, et les autres électrodes pour mieux visualiser la relation entre les électrodes ainsi que l'effet de la distance. Un code de couleur est utilisé : vert foncé étant la corrélation la plus élevée et rouge étant la corrélation la plus faible.



Figure 4 - Les corrélations de Pearson pour les paires d'électrodes associées à FCz illustrées selon les positionnements d'électrodes Note : Les couleurs en ordre croissant de corrélations : rouge (négatif), orange, jaune, jaune vert, vert, vert peu foncé et vert foncé.

Ce qui est remarquable est le fait que les électrodes qui sont situées dans le même hémisphère sont plus corrélées en termes de fréquence et que les électrodes qui ne sont pas dans le même hémisphère sont beaucoup moins corrélées. De plus, les activités de la zone frontale (électrodes commençant par F) du cerveau ont moins tendance à être corrélées avec celles de la partie occipitale (électrodes commençant par O) et de la partie pariétale (électrodes commençant par P) du cerveau. Cette tendance est particulièrement remarquable pour l'électrode Pz. On voit que les moyennes des corrélations sont très faibles entre Pz et les électrodes dans la zone frontale en plus des moyennes de corrélations avec FCz et Fpz négative. Aussi, selon cette matrice, il semble que les électrodes de la partie du cervelet (C3 et C4) ne sont pas très corrélées avec toutes les autres parties du cerveau (voir les deux premières lignes de la matrice) sauf avec P3 et P4.

Pour la deuxième partie de l'analyse des relations des zones du cerveau, les mêmes analyses ont été faites avec les corrélations qui sont seulement significatives. On considère que les corrélations ayant un p-value de plus petit ou égal à 5% sont significatives. Les deux matrices ont été produites pour mieux visualiser les valeurs moyennes de corrélations qui ont été significatives (Tableau 10) et le nombre de fois que les corrélations ont été significatives entre les différentes paires d'électrodes (Tableau 11). La matrice du Tableau 10 contient toutes les moyennes de corrélations qui ont été significatives en utilisant les mêmes codes de couleurs que ceux utilisés précédemment. La matrice du Tableau 11 montre le nombre fois où les corrélations ont été significatives pour chacune des paires d'électrodes. Par exemple, pour la paire d'électrodes P7 et P8, on voit afficher le nombre 30, ce qui signifie que tous les tests de corrélations de Pearson ont été significatifs, car il y a 30 matrices de corrélations qui ont été générées pour 30 différents moments dans le temps. On voit aussi que le nombre afficher pour la paire d'électrodes F8 et C3, les corrélations ont seulement été significatives une seule fois.

Tableau 10 - Matrice des corrélations de Pearson significatives entre les paires d'électrodes

	C4	F3	F4	F7	F8	FCz	Fp1	Fp2	Fpz	Fz	01	02	Oz	P3	P4	P7	P8	Pz
C3	0,37	0,63	0,51	0,68	0,88	0,63	0,61	0,56	0,53	0,51	0,57	0,60	0,57	0,70	0,54	0,59	0,59	0,51
C4		0,36	0,58	0,48	0,95	0,60	0,47	0,51	0,56	0,51	0,58	0,56	0,62	0,51	0,66	0,54	0,64	0,43
F3			0,45	0,71	0,61	0,64	0,64	0,61	0,62	0,60	0,51	0,58	0,50	0,52	0,58	0,61	0,50	0,20
F4				0,54	0,98	0,63	0,43	0,59	0,53	0,60	0,54	0,52	0,53	-0,48	0,60	0,56	0,56	-0,01
F7					0,64	0,67	0,66	0,64	0,68	0,68	0,66	0,60	0,66	0,63	0,59	0,70	0,60	0,61
F8						0,90	0,57	0,93	0,46	0,94	0,58	0,92	0,99	-0,54	0,94	0,55	0,95	0,91
FCz							0,60	0,65	0,66	0,78	0,29	0,33	0,18	0,22	0,32	0,19	0,38	-0,33
Fp1								0,62	0,75	0,60	0,19	0,60	0,60	0,32	0,45	0,51	0,50	-0,58
Fp2									0,67	0,62	0,18	0,33	0,48	0,51	0,44	0,32	0,55	0,52
Fpz										0,68	0,20	0,54	0,50	0,39	0,47	0,38	0,54	0,20
Fz											0,30	0,47	0,34	0,28	0,53	0,35	0,57	-0,53
01												0,73	0,87	0,61	0,55	0,71	0,64	0,66
O2													0,86	0,56	0,62	0,59	0,71	0,65
Oz														0,63	0,59	0,65	0,66	0,72
P3															0,52	0,73	0,53	0,63
P4																0,53	0,76	0,63
P7																	0,56	0,61
P8																		0,60

	C4	F3	F4	F7	F8	FCz	Fp1	Fp2	Fpz	Fz	01	02	Oz	P3	P4	P7	P8	Pz
C3	6	15	2	13	1	5	5	5	7	9	10	8	10	27	7	18	6	9
C4		6	15	8	4	7	3	6	5	6	7	9	7	6	26	7	14	8
F3			9	22	6	18	19	15	24	25	2	3	4	- 7	2	4	2	3
F4				9	8	21	14	21	20	20	4	5	4	1	5	3	9	2
F7					6	10	22	13	18	12	7	10	8	12	6	10	5	3
F8						3	5	8	4	- 7	5	5	2	5	6	5	- 5	1
FCz							16	16	21	30	5	6	3	6	5	6	6	4
Fp1								19	28	25	6	5	4	6	1	5	6	1
Fp2									29	25	3	6	2	3	4	6	12	2
Fpz										27	- 7	5	5	6	3	- 7	8	4
Fz											5	5	5	5	3	6	5	2
01												27	30	26	20	30	20	27
O2													30	19	29	24	30	27
Oz														18	27	28	28	28
P3															13	30	13	25
P4																14	29	26
P7																	18	25
P8																		21

Tableau 11 - Matrice des nombres de corrélations significatives

La matrice présentant les nombres de fois où les corrélations ont été significatives (Tableau 11) suit les codes de couleurs suivants pour identifier les différentes valeurs de nombre de corrélations significatives pour chaque paire d'électrodes (Tableau 12).

Tableau 12 - Mise en forme conditionnelle de la matrice des nombres de corrélations<br/>significatives (tableau 13)

Si la valeur	>= ``	<ul><li>0</li></ul>	Nombre 🗸 e	et < 🗸	5	Nombre 🗸 alors		$\land \downarrow \times$
Si la valeur	>= ``	<b>~</b> 5	Nombre 🗸 e	et < 🗸	10	Nombre 🗸 alors		$\uparrow \ \downarrow \ \times$
Si la valeur	>=	/ 10	Nombre 🗸 e	et < 🗸	15	Nombre 🗸 alors	<ul> <li></li> </ul>	$\uparrow \ \downarrow \ \times$
Si la valeur	>=	/ 15	Nombre 🗸 e	et < ~	20	Nombre 🗸 alors	<ul> <li></li> </ul>	$\uparrow \ \downarrow \ \times$
Si la valeur	>= ``	20	Nombre 🗸 e	et < 🗸 🗸	25	Nombre 🗸 alors	<ul> <li></li> </ul>	$\uparrow \ \lor \ \times$
Si la valeur	>= ``	25	Nombre 🗸 e	et <= ~	30	Nombre 🗸 alors	$\sim$	$\wedge$ $\downarrow$ $\times$

Comme on peut voir, la matrice des moyennes de corrélations significatives (Tableau 10) ont les couleurs qui semblent avoir changées comparée à la matrice avec toutes les corrélations (Tableau 9). Les parties ayant des valeurs de corrélations moyennes les plus proches de zéro dans la matrice contenant toutes les corrélations (Tableau 9) semblent avoir les valeurs de corrélation moins proche de zéro dans cette matrice qui contient seulement les corrélations qui ont été significatives (Tableau 10) et ce, parce que les valeurs de corrélations non significatives sont des valeurs très basses. Cependant, on remarque que les parties en orange, rouge et mauve de la matrice illustrant le nombre de corrélations significatives (Tableau 11) se retrouvent presque aux mêmes endroits que les valeurs les plus proches de zéro dans la matrice de moyennes de toutes les corrélations (Tableau 9), ce qui rejoint aussi le fait que les électrodes qui ont une distance plus grande entre elles ont une valeur absolue de corrélation moins élevée étant donné que la non significativité de la corrélation équivaut à dire que les corrélations ne sont pas significativement différentes de zéro. De plus, on peut dire que les activités de la zone frontale (électrodes commençant par F) du cerveau ont non seulement moins tendance à être corrélées avec celles de la partie occipitale (électrodes commençant par O) et de la partie pariétale (électrodes commençant par P) du cerveau, mais ont aussi tendance à avoir des corrélations non significatives. Si on regarde la partie de la matrice avec le nombre de corrélations les plus significatives, on peut voir que les électrodes de partie occipitale et la partie pariétale ont des corrélations plus significatives entre elles et ces corrélations sont relativement élevées pour la majorité des paires. Il y a, cependant, des exceptions comme les paires P3 et P4, P4 et P7 et P3 et P8. Ces paires sont plus loin entre-elles et ne sont pas dans le même hémisphère, c'est probablement ce qui explique ces exceptions. Une autre partie de la matrice qui contient le plus grand nombre de corrélations significatives est celle avant toutes les électrodes de la zone frontale et les corrélations associées à cette zone dans la matrice de corrélation moyenne (Tableau 9) sont aussi relativement plus loin de zéro. De plus, F3 - F4 et F4 - F7 sont aussi deux paires d'électrodes qui ne sont pas très corrélées et ces corrélations ne sont pas souvent significatives. Tel que mentionné précédemment, on observe encore une fois que les électrodes de la partie du cervelet ne semblent être corrélées avec aucune des autres électrodes des autres zones du cerveau de manière significative sauf avec les électrodes P3 et P4.

L'exception est l'électrode F8, qui a des corrélations presque jamais significatives entre lui et les électrodes, ce qui n'est pas étonnant, car F8 n'est pas une électrode qui est présente pour beaucoup de participants. En effet, lorsqu'on fait des analyses de corrélations de Pearson dans SAS, le logiciel élimine les valeurs manquantes par paire (*pairwise deletion*), ce qui veut dire que pour une paire d'électrodes dont une des deux électrodes est absente pour le participant, cette paire est éliminée de l'analyse (« The CORR Procedure », 2020). Il est donc normal d'observer des résultats qui sont inattendus pour l'électrode F8, car les corrélations entre F8 et les autres électrodes sont basées sur moins d'observations que les corrélations entre les autres paires d'électrodes. En fait, la baisse de la taille d'échantillon peut affecter la significativité (Janvier, 2001).

#### Analyse des distances entre les électrodes

Dans cette partie, nous aimerions vérifier les corrélations entre les électrodes selon leurs distances qui les séparent. Dans cette analyse, les différents points de temps sont fixés pour seulement avoir les fréquences à un moment donné afin de pouvoir analyser si les activités entre les électrodes sont affectées par les distances. La variable « fréquence » des électrodes est utilisée pour calculer la corrélation. On a aussi choisi dix différents moments à l'intérieur de chaque moment de pause, ce qui représente trente différents moments analysés. Ce nombre de moments choisis est pour mieux analyser les relations entre les points de temps pour différents moments dans le temps afin de ne pas négliger la divergence des activités aux différents moments. Après l'obtention des corrélations, la moyenne des corrélations a été calculée pour obtenir une matrice de moyennes des corrélations entre les électrodes afin de synthétiser les résultats obtenus.

La matrice de corrélation moyenne obtenue est utilisée pour analyser les corrélations en fonction des distances entre chaque électrode. Les distances euclidiennes sont calculées à partir des coordonnées des positions des électrodes présentées dans le dernier chapitre. Comme nous avons produit les deux types de coordonnées pour les positions des électrodes données dans les deux articles (Böcker et al. (1994) et Khosla et al. (1999)), ces coordonnées ont donc été utilisées pour l'analyse. Pour mieux comprendre la tendance, des nuages de points ont été employés pour vérifier l'effet de la distance entre les électrodes sur les activités des différentes parties du cerveau. Ces graphiques ont été produits à l'aide du logiciel SAS. Les distances euclidiennes pour les différentes positions provenant des deux articles ne sont pas sur la même échelle, donc pour chacun des positionnements (UV ou XYZ selon Böcker et al. (1994) et UV ou XYZ selon Khosla et al. (1999)), nous avons divisé chacune des distances par la distance maximale pour savoir à quel point une électrode est loin d'une autre comparativement à la distance la plus loin. La distance la plus loin entre deux électrodes pour tous les types de positionnements. De cette manière, les distances seront comparables pour tous les types de positionnement que nous avons.

Pour commencer, nous allons analyser la tendance des moyennes des corrélations selon les distances calculées à partir des coordonnées UV. Les résultats du Tableau 13 montre les paramètres estimés pour la régression linéaire modélisée pour les moyennes de corrélations associées à chacune des 114 distances différentes en fonction des distances euclidiennes

calculées à partir des coordonnées UV calculées pour les positions d'électrodes fournies dans l'article de Böcker et al. (1994) et le graphique de la Figure 5 présenté ci-dessous illustre la courbe de tendance. Comme on peut le voir, il y a une tendance de la moyenne de corrélation descendante et qui se rapproche de zéro lorsque la distance augmente, car la pente, calculée à partir d'une régression linéaire, qu'on voit sur le nuage de point est de -0.598 et les corrélations sont majoritairement positives. Cependant, le coefficient de détermination (r-carré) est de 0.32, ce qui signifie que seulement 32% de la variabilité observée pour la corrélation moyenne entre les électrodes est expliquée par la distance entre les électrodes. Les distances calculées à partir des coordonnées UV selon les positions d'électrodes fournies dans l'article de Böcker et al. (1994) ne sont donc pas le meilleur choix pour expliquer la variation des corrélations entre les électrodes.

Tableau 13 - Les paramètres estimés pour la régression linéaire modélisée pour la corrélation en fonction de la distance euclidienne selon les positionnements bidimensionnels de Böcker et al. (1994)

	Résu	Iltats estimés des	paramèti	res	
Variable	DDL	Valeur estimée des paramètres	Erreur type	Valeur du test t	Pr >  t
Intercept	1	0.61504	0.04035	15.24	<.0001
distance_uv_bocker	1	-0.59803	0.08174	-7.32	<.0001



Figure 5 - Courbe d'ajustement de la corrélation de Pearson moyenne en fonction des distances selon les positionnements bidimensionnels de Böcker et al. (1999)

Les résultats du Tableau 14 présente montre les paramètres estimés pour la régression linéaire modélisée pour les moyennes des corrélations associées à chacune des 96 distances différentes en fonction des distances euclidiennes calculées à partir des coordonnées UV selon les positions d'électrodes fournies dans l'article de Khosla et al. (1999) et le graphique de la Figure 6 montre la courbe de tendance. Selon la pente, on peut voir qu'il y a une tendance de la moyenne de corrélation descendante et se rapprochant de zéro lorsque la distance augmente, car la pente, calculée à partir d'une régression linéaire, qu'on voit sur le nuage de point est de -0.642 et le coefficient de détermination (r-carré) est de 0.53, ce qui signifie que seulement 53% de la variabilité observée dans la corrélation moyenne entre les électrodes est expliquée par la distance entre les électrodes. Même si la pente est plus accrue et le coefficient de détermination est plus élevé, on ne peut pas dire que les distances calculées à partir des coordonnées UV selon les positions d'électrodes fournies dans l'article de Khosla et al. (1999) sont le meilleur choix pour expliquer la variation des corrélations entre les électrodes.

## Tableau 14 - Les paramètres estimés pour la régression linéaire modélisée pour la corrélation en fonction de la distance euclidienne selon les positionnements bidimensionnels de Khosla et al. (1999)

	Résu	ıltats estimés des	paramèt	res	
Variable	DDL	Valeur estimée des paramètres	Erreur type	Valeur du test t	Pr >  t
Intercept	1	0.64153	0.03383	18.96	<.0001
distance_uv_khosla	1	-0.69057	0.06644	-10.39	<.0001



Figure 6 - Courbe d'ajustement de la corrélation de Pearson moyenne en fonction des distances selon les positionnements bidimensionnels de Khosla et al. (1994)

Lorsqu'on arrive à analyser les distances calculées à partir des coordonnées XYZ, on observe des tendances qui sont relativement différentes de celles observées dans les quatre derniers graphiques qui utilisent les distances calculées à partir des coordonnées UV.

Tout d'abord, regardons les tendances des moyennes de corrélation en fonction de chacune des 91 distances euclidiennes calculées à partir des coordonnées XYZ générées selon les positions d'électrodes fournies dans l'article de Böcker et al (1994). Nous voyons ici que la tendance descendante de la corrélation moyenne au fur et à mesure que la distance augmente devient plus claire et plus puissante, car la valeur absolue de la pente (|-.076|) est beaucoup plus élevée que la valeur absolue de celles observées précédemment. De plus, en se fiant au coefficient de détermination de 0.64, ce qui est aussi plus élevé que ceux observés précédemment. Donc, 64% de la variabilité des corrélations moyennes sont expliqués par le modèle. On peut dire que les points se gravitent plus autour de la pente pour expliquer la variabilité de la corrélation moyenne par la distance puisque cette valeur est relativement plus élevée que les deux dernières que nous avions vues précédemment.

Tableau 15 - Les paramètres estimés pour la régression linéaire modélisée pour la corrélation en fonction de la distance euclidienne selon les positionnements tridimensionnels de Böcker et al. (1994)

	Résu	ltats estimés des	paramètr	es	
Variable	DDL	Valeur estimée des paramètres	Erreur type	Valeur du test t	Pr >  t
Intercept	1	0.76928	0.03673	20.95	<.0001
distance_xyz_bocker	1	-0.75680	0.05971	-12.68	<.0001



Figure 7 - Courbe d'ajustement de la corrélation de Pearson moyenne en fonction des distances selon les positionnements tridimensionnels de Böcker et al. (1999)

Les résultats du **Tableau** *16* - Les paramètres estimés pour la régression linéaire modélisée pour la corrélation en fonction de la distance euclidienne selon les positionnements tridimensionnels de Khosla et al. (1999)Tableau 16 présentent les paramètres estimés pour la régression linéaire modélisée pour les moyennes de corrélations selon les 88 différentes distances euclidiennes calculées à partir des coordonnées XYZ selon les positions d'électrodes fournies dans l'article de Khosla et al. (1999) et le graphique de la Figure 8 montre la courbe de tendance. On remarque presque les mêmes tendances que celle observée dans le graphique de la Figure 7, car la pente est relativement similaire (-0.75). Puisque le R-carré (0.71) qui est le plus élevé parmi celles observées précédemment, on peut dire que les distances calculées à partir des coordonnées XYZ en fonction des positions fournies d'électrode fournies dans l'article de Khosla et al. (1999) peuvent mieux expliquer la variabilité de la moyenne de corrélation.

### Tableau 16 - Les paramètres estimés pour la régression linéaire modélisée pour la corrélation en fonction de la distance euclidienne selon les positionnements tridimensionnels de Khosla et al. (1999)

	Résu	ltats estimés des	paramètr	es	
Variable	DDL	Valeur estimée des paramètres	Erreur type	Valeur du test t	Pr >  t
Intercept	1	0.78872	0.03353	23.52	<.0001
distance_xyz_khosla	1	-0.75424	0.05176	-14.57	<.0001



Figure 8 - Courbe d'ajustement de la corrélation de Pearson moyenne en fonction des distances selon les positionnements tridimensionnels de Khosla et al. (1994)

D'après ces observations, on peut dire qu'il est probable que la relation entre les activités enregistrées par les électrodes dans les différentes zones du cerveau est affectée par les distances qui les séparent. Plus précisément, il semble avoir une tendance descendante et qui se rapprochant de zéro de manière plus accrue lorsqu'on fait l'analyse avec les distances calculées selon les coordonnées XYZ. De ce fait, on peut déduire que les coordonnées XYZ sont un meilleur choix pour le calcul des distances pour analyser les tendances de corrélation

entre les électrodes. Il y a une raison possible qui peut expliquer ce phénomène. En fait, les coordonnées UV sont calculées lorsqu'on veut retrouver les positions des électrodes sur une surface de crâne de manière « aplatie ». C'est comme si on avait transformé les positions des électrodes sur un crâne 3D aux positions des électrodes sur un crâne 2D. Les distances ont donc été calculées comme étant la distance absolue entre deux points à la surface du crâne. Contrairement aux coordonnées UV, les coordonnées XYZ considèrent que le crâne est en 3D, ce qui fait que les distances calculées peuvent « traverser » le cerveau sans avoir la contrainte de seulement pouvoir passer par la surface du crâne. On peut croire que la distance calculée à partir des coordonnées XYZ est meilleure pour analyser l'effet de la distance sur les activités entre les différentes zones du cerveau.

L'analyse qui vient d'être faite n'a pas tenu compte de la significativité des corrélations, car nous voulons voir la tendance générale des estimations des corrélations pour observer les relations. Si on élimine les corrélations n'étant pas significatives, on risque de ne pas avoir toutes les corrélations pour vérifier les relations. Comme les corrélations de Pearson doivent être significatives pour dire qu'elles sont différentes de zéro, une autre analyse a été faite pour les corrélations qui ont seulement été significatives. Cette analyse a pour but de vérifier quelles sont les distances qui ont été les plus souvent significatives et quelles sont ces corrélations. Les corrélations ayant un p-value plus petit que 5% sont considérées comme étant significativement différentes de zéro et sont sélectionnées pour les prochaines analyses.

Le graphique de la Figure 9 a été produit par le logiciel PowerBI dans le but de visualiser et de résumer le pourcentage de corrélations qui ont été significatives en fonction des moyennes de corrélations pour chaque distance différente. Comme pour chaque distance, le nombre de corrélations associé à chacune des distances n'est pas la même, le pourcentage est utilisé pour les analyses suivantes. En effet, il faut comprendre que chacune des distances n'est pas unique à chacune des paires d'électrodes et que certaines paires d'électrodes ont la même distance que d'autres. Seules les distances calculées à partir des coordonnées XYZ produites selon les positions présentées dans l'article de Khosla et al. (1999) sont utilisées dans ce graphique, car selon la dernière analyse, les coordonnées XYZ sont mieux pour analyser l'effet de la distance. Les distances ont été identifiées par des couleurs différentes : bleu représente la distance la plus faible, rouge représente la distance la plus élevée et jaune représente la distance médiane. Tel qu'observé dans ce graphique, on peut voir que le nombre de corrélations significatives augmente avec la moyenne des corrélations significatives qui croît. De plus, on observe que

plus que c'est proche en termes de distance, plus que la moyenne des corrélations associée est élevée et est loin de zéro et plus qu'il y a un grand nombre de corrélations significatives.



Figure 9 - Nuage de points des proportions des corrélations de Pearson significatives en fonction des moyennes de corrélation pour les différentes distances (selon les positionnements de Khosla et al. (1999))

Les deux graphiques de moyenne de corrélations en fonction des distances calculées à partir des coordonnées UV et les deux graphiques de moyenne de corrélations en fonction des distances calculées à partir des coordonnées XYZ ont été reproduits pour les corrélations significatives dans la Figure 10. On peut voir que la tendance de la corrélation vers le bas lors de l'augmentation de la distance est moins accrue étant donné qu'on a éliminé les corrélations qui sont associées aux distances plus élevées, car elles sont moins significatives.



Figure 10 - Comparaison des courbes de tendance pour les corrélations significatives moyennes selon les distances calculées en fonction des positionnements bidimensionnels et tridimensionnels des deux articles.

### Analyse des relations entre les différents points de temps

Dans cette partie nous allons fixer chacune des 19 électrodes pour observer la tendance de la corrélation entre les activités du cerveau à travers temps en utilisant encore une fois des matrices de Pearson. Comme on veut seulement vérifier la tendance de la corrélation dans le

temps et non de faire une analyse complète du temps, seulement des laps de vingt points dans le temps ont été choisis. Quatre durées de temps, 1000 à 1020, 3000 à 3020, 5000 à 5020 et 8000 à 8020, ont été choisies parmi les trois pauses sélectionnées pour chaque électrode, ce qui en fait douze durées de vingt points dans le temps pour chaque électrode, donc 0.078 seconde par durée (20/256). Le fait de choisir des durées de 20 points de temps est dans le but de minimiser le volume de données et en même temps d'avoir assez de données pour voir la tendance de la corrélation en lien avec le temps. Après avoir obtenu toutes les 228 matrices de corrélation de Pearson (19 électrodes \* 12 durées dans le temps), celles-ci ont été transposées verticalement pour obtenir chaque paire de temps et leur distance dans le temps. Par exemple, si on a la paire de temps 1002 et 1014, la distance dans le temps est de 12. En faisant ceci, on peut voir comment les corrélations évoluent au fur et à mesure que le temps avance.

Pour commencer, on va voir toutes les corrélations sans distinguer celles qui sont significatives et celles qui ne le sont pas. On a produit deux graphiques pour voir la tendance.

Le graphique de la Figure 11 montre la moyenne de toutes les corrélations pour toutes les électrodes fixées selon chacune des distances dans le temps. On voit clairement une tendance vers le bas, c'est-à-dire que la corrélation diminue et s'approche de zéro lorsque la distance entre les temps augmente. Plus que c'est loin dans le temps, moins que c'est corrélé.



Figure 11 - Moyenne de toutes les corrélations en fonction du nombre de points de temps séparant deux points de temps

Le graphique de la Figure 12 contient les courbes de tendance de la corrélation moyenne en fonction de la distance du temps pour chacune des électrodes séparément. Ce graphique est

nécessaire pour vérifier si toutes les électrodes suivent la même tendance ou certaines d'entre elles s'écartent de la tendance générale. Selon l'observation de ce graphique, les courbes de tendance de corrélation moyenne de l'électrode qui s'écartent plus que les autres sont celles qui représente les électrodes F8, P7 et Pz, mais ce ne sont pas des écarts vraiment extrêmes. On peut donc dire que sans considérer la significativité des corrélations, la tendance de la moyenne de corrélation par distance du temps et par électrode est descendante et s'approche de zéro même si la tendance de quelques-unes des électrodes s'écarte de la majorité.



Figure 12 - Moyenne de toutes les corrélations en fonction du nombre de points de temps séparant deux points de temps pour chacune des électrodes

Ensuite, une analyse a été faite pour observer les moyennes des corrélations qui ont été seulement significatives. On considère que les corrélations ayant un p-value plus petit que 0.05 est significativement différent de zéro. L'ensemble de toutes les corrélations entre les temps pour toutes les électrodes ont été agrégées par la moyenne pour chacune des distances du temps. De plus, le nombre de corrélations significatives a aussi été calculé pour chaque distance du temps. Un graphique a été produit à cet effet. Un code de couleur a été mis en place pour identifier les différentes distances. Comme on peut voir dans le graphique de la Figure 13, plus que le nombre de corrélations significatives est élevé, plus que la moyenne de corrélation est élevée et loin de zéro et plus que la distance entre les temps est faible. Ce fait n'est pas étonnant, puisqu'on a observé précédemment le fait que la moyenne de corrélation est descendante lorsque la distance augmente.



Moyenne de corrélations significatives en fonction du nombre de corrélations significatives

Figure 13 - Moyenne de corrélations significatives en fonction du nombre de corrélations significatives pour chacun des nombres de points de temps séparant deux temps

Le graphique de la Figure 14 présente la courbe de tendance des corrélations significatives moyennes en fonction de la distance entre les temps a été produit. On peut voir que la tendance est toujours descendante.



Figure 14 - Moyenne de toutes les corrélations significatives en fonction du nombre de points de temps séparant deux points de temps

Voyons si les tendances changent en fonction des électrodes. Comme on peut voir dans la Figure 15, les courbes ne sont plus aussi regroupées lorsqu'on sélectionne seulement les corrélations qui ont été significatives. La courbe représentant l'électrode F8 s'écarte le plus des autres courbes.



Figure 15 - Moyenne de toutes les corrélations significatives en fonction du nombre de points de temps séparant deux points de temps pour chacune des électrodes

Si on regarde la liste du Tableau 17, on peut voir que le nombre de corrélations significatives pour F8 est le plus faible et s'écarte beaucoup des autres électrodes. Sachant que le nombre de corrélations différentes obtenues par matrice est de 190 et qu'il y ait 12 durées de temps différentes, il y a donc 2280 valeurs de corrélation générées par électrode. On voit que le nombre des corrélations significatives pour les autres électrodes est environ entre 600 et 1100, mais celui de l'électrode F8 est de 363, ce qui est assez faible comparé aux autres. De plus, tel qu'expliqué précédemment, comme F8 n'est pas présent pour tous les participants et que la taille d'échantillon est faible pour l'électrode F8, il est probable que c'est pour cela que F8 ait moins de corrélations significatives. Selon la courbe de tendance représentant l'électrode F8, elle est placée plus vers le haut comparé aux autres courbes et comparée à sa courbe contenant toutes les corrélations, ce qui veut dire que les corrélations qui ont été significatives sont celles qui sont plus élevées.

### Tableau 17 - Nombres de corrélations significatives, triés en ordre croissant, entre les temps pour chacune des électrodes

F8	363
01	591
Oz	597
O2	615
C4	700
F4	711
P8	721
F3	735
P7	736
C3	771
Fp2	773
P4	796
Pz	800
F7	826
P3	843
Fz	897
FCz	954
Fp1	972
Fp2	1119

Électrode Nombre de corrélations significatives

# Analyse des relations entre les électrodes au cours des différents moments dans le temps

Dans cette partie, les corrélations de Pearson ont été produites pour analyser les relations entre les activités des électrodes entre-elles et entre les différents points de temps. Pour pouvoir faire cette analyse, 24 durées de 10 points de temps ont été choisies. Une table de données a été créée pour chaque durée choisie et toutes les électrodes sont gardées dans chacune des tables des données. La variable contenant les noms d'électrodes et la variable contenant les temps ont été concaténées pour créer une nouvelle variable afin d'identifier de manière unique chaque ligne de fréquence de chaque participant. En fait, pour pouvoir générer les matrices de corrélation avec le logiciel SAS, il est essentiel d'avoir une et une seule variable pour identifier les catégories entre lesquelles on veut analyser leurs corrélations. Dans ce cas, on veut analyser, par exemple, la corrélation entre la fréquence de l'électrode C3 au temps 1 et la fréquence de l'électrode C4 au temps 2. Comme les différentes électrodes et les différents temps sont identifiés par deux variables distinctes, c'est la raison pour laquelle les deux variables sont concaténées en une seule variable pour ensuite l'utiliser pour générer les matrices de corrélation de Pearson. La moyenne de ces matrices a été calculée pour faire la suite de l'analyse.

Nous voyons avec les analyses précédentes que plus on est loin dans le temps, moins la corrélation est élevée et plus que c'est loin en termes de distance, moins la corrélation est élevée. En nous basant sur ces informations, nous présumons que plus on est loin en termes de distance et en termes de temps, moins la corrélation est élevée. Par exemple, si on prend le temps 1 de l'électrode C3 et le temps 8 de l'électrode P3, la corrélation sera moins élevée que si on prend le temps 1 de l'électrode C3 et le temps 2 de l'électrode P3.

La matrice qui a été produite est très grande et ne peut pas être affichée dans ce rapport. On a donc vérifié si la tendance des corrélations est semblable à notre hypothèse. D'après les vérifications, nous avons remarqué que pour 14 électrodes sur les 19 électrodes, notre hypothèse est confirmée. Seules les électrodes Fp1, Fp2, O1, O2 et Oz présentent des exceptions.

Dans la majorité des cas, les corrélations les plus élevées sont les corrélations retrouvant sans sur la diagonale secondaire des « sous-matrices » de chacun des électrodes. On parle donc ici de la matrice de temps à l'intérieur d'une électrode comme dans la partie de la matrice que l'on voit dans le Tableau 18.

Tableau 18 - Partie de matrice de moyennes de corrélations entre les électrodes au	X
différents temps (temps 1 à temps 9) pour les électrodes C3 et C4	

	C3_T2	C3_T3	C3_T4	C3_T5	C3_T6	C3_T7	C3_T8	C3_T9	C4_T1	C4_T2	C4_T3	C4_T4	C4_T5	C4_T6	C4_T7	C4_T8	C4_T9
C3_T1	0,69	0,59	0,45	0,30	0,23	0,13	0,15	0,08	0,11	0,15	0,09	0,08	0,05	0,07	0,06	0,02	0,04
C3_T2		0,68	0,51	0,33	0,23	0,20	0,11	0,03	0,16	0,06	0,09	0,10	0,11	0,06	0,07	0,10	0,07
C3_T3			0,73	0,50	0,38	0,34	0,26	0,16	0,14	0,12	0,12	0,12	0,10	0,02	0,07	0,15	0,12
C3_T4				0,69	0,47	0,36	0,31	0,21	0,08	0,10	0,11	0,09	0,07	0,02	0,10	0,16	0,08
C3_T5					0,66	0,47	0,42	0,33	0,06	0,09	0,09	0,10	0,02	0,03	0,09	0,13	0,09
C3_T6						0,68	0,53	0,45	0,02	-0,01	0,03	0,06	0,08	0,05	0,08	0,12	0,10
C3_T7							0,66	0,47	0,02	-0,02	0,00	0,02	0,07	0,09	0,07	0,15	0,16
C3_T8								0,69	0,05	0,06	0,05	0,09	0,13	0,13	0,14	0,12	0,12
C3_T9									0,06	0,09	0,03	0,08	0,07	0,05	0,08	0,13	0,03

On voit que dans cette partie de la grande matrice, il y a deux « sous-matrices » celle de la C3 et celle de la C4. Celle de la C3 a des corrélations sur la diagonale secondaire qui sont très élevées qui représente les corrélations entre les points de temps adjacents. Ces résultats sont normaux étant donné que ce sont des corrélations entre les points de temps à l'intérieur d'une électrode sans être affecté par le saut à d'autres électrodes qui les écarte. Malgré ces observations, quelques exceptions ont été observées pour les électrodes Fp1, Fp2, O1, O2 et Oz. On a remarqué que les corrélations les plus élevées associées à ces électrodes ne sont pas celles sur la diagonale secondaire intraélectrode, mais plutôt celles qui y sont associées se trouvent sur la diagonale de la « sous-matrice » Fpz. Pour O1 et O2, les corrélations les plus élevées qui y sont associées se trouvent sur la diagonale de la « sous-matrice » Fpz. Pour O1 et O2, les corrélations les plus élevées qui y sont associées se trouvent sur la diagonale de la « sous-matrice » Fpz. Pour O1 et O2, les corrélations les plus élevées qui y sont associées se trouvent sur la diagonale de la « sous-matrice » Fpz. Pour O1 et O2, les corrélations les plus élevées qui y sont associées se trouvent sur la diagonale de la « sous-matrice » Fpz. Pour O1 et O2, les corrélations les plus élevées qui y sont associées se trouvent sur la diagonale de la « sous-matrice » Pz.

Tableau 19 - Partie de matrice de moyennes de corrélations entre les électrodes auxdifférents temps (temps 1 à temps 9) pour les électrodes Fp1 et Fpz

IDte_1	Fp1_T2	Fp1_T3	Fp1_T4	Fp1_T5	Fp1_T6	Fp1_T7	Fp1_T8	Fp1_T9	Fpz_T1	Fpz_T2	Fpz_T3	Fpz_T4	Fpz_T5	Fpz_T6	Fpz_T7	Fpz_T8	Fpz_T9
Fp1_T1	0,67	0,37	0,15	0,06	0,01	-0,03	0,07	0,15	0,72	0,56	0,36	0,24	0,15	0,07	-0,04	0,04	0,09
Fp1_T2		0,62	0,27	0,11	0,05	0,07	0,05	0,11	0,53	0,73	0,51	0,32	0,23	0,13	0,02	0,08	0,11
Fp1_T3			0,61	0,16	0,09	0,12	0,05	0,07	0,37	0,58	0,74	0,54	0,33	0,18	0,01	0,03	0,10
Fp1_T4				0,53	0,26	0,20	0,09	-0,01	0,22	0,30	0,57	0,71	0,53	0,34	0,07	0,07	0,06
Fp1_T5					0,61	0,24	0,09	-0,02	0,10	0,06	0,17	0,39	0,68	0,51	0,21	0,10	0,06
Fp1_T6						0,58	0,20	0,14	0,07	-0,01	0,14	0,24	0,47	0,74	0,49	0,24	0,19
Fp1_T7							0,57	0,23	0,02	0,05	0,12	0,19	0,23	0,50	0,64	0,52	0,38
Fp1_T8								0,55	0,03	0,09	0,05	0,03	0,04	0,21	0,50	0,73	0,50
Fp1 T9									0.08	0.15	0.11	0.04	-0.02	0.13	0.37	0.49	0.64

Tableau 20 - Partie de matrice de moyennes de corrélations entre les électrodes aux différents temps (temps 1 à temps 9) pour les électrodes Fp2 et Fpz

IDte_1	Fp2_T2	Fp2_T3	Fp2_T4	Fp2_T5	Fp2_T6	Fp2_T7	Fp2_T8	Fp2_T9	Fpz_T1	Fpz_T2	Fpz_T3	Fpz_T4	Fpz_T5	Fpz_T6	Fpz_T7	Fpz_T8	Fpz_T9
Fp2_T1	0,64	0,28	0,09	0,07	0,21	0,27	0,22	0,16	0,64	0,37	0,17	0,12	0,22	0,29	0,26	0,10	0,09
Fp2_T2		0,63	0,24	0,18	0,25	0,28	0,17	0,12	0,47	0,60	0,48	0,27	0,26	0,29	0,23	0,04	0,11
Fp2_T3			0,55	0,35	0,30	0,21	0,17	0,09	0,38	0,47	0,72	0,55	0,40	0,33	0,16	0,09	0,09
Fp2_T4				0,69	0,32	0,10	0,01	-0,06	0,29	0,45	0,58	0,75	0,48	0,23	0,01	0,07	0,08
Fp2_T5					0,55	0,13	-0,01	0,04	0,20	0,36	0,44	0,54	0,65	0,39	0,04	0,07	0,12
Fp2_T6						0,57	0,29	0,33	0,21	0,15	0,19	0,27	0,51	0,70	0,39	0,22	0,26
Fp2_T7							0,68	0,42	0,10	0,07	0,06	0,09	0,21	0,46	0,70	0,51	0,36
Fp2_T8								0,63	0,10	0,04	-0,03	-0,03	0,03	0,27	0,55	0,67	0,49
Fp2_T9									0,14	0,00	-0,02	-0,03	0,07	0,29	0,38	0,45	0,60

### Tableau 21 - Partie de matrice de moyennes de corrélations entre les électrodes auxdifférents temps (temps 1 à temps 9) pour les électrodes O1 et Oz

IDte_1	01_T2	O1_T3	01_T4	O1_T5	O1_T6	01_T7	O1_T8	O1_T9	Oz_T1	Oz_T2	Oz_T3	Oz_T4	Oz_T5	Oz_T6	Oz_T7	Oz_T8	Oz_T9
01_T1	0,48	0,28	0,18	0,24	0,20	0,15	0,02	0,05	0,87	0,47	0,29	0,23	0,25	0,24	0,16	0,03	0,02
O1_T2		0,56	0,27	0,27	0,22	0,12	0,06	0,05	0,52	0,88	0,56	0,32	0,30	0,22	0,14	0,04	0,02
O1_T3			0,53	0,31	0,28	0,23	0,12	0,05	0,34	0,55	0,86	0,51	0,29	0,28	0,22	0,10	0,06
O1_T4				0,60	0,35	0,27	0,17	0,11	0,22	0,27	0,48	0,85	0,53	0,31	0,26	0,19	0,16
O1_T5					0,56	0,22	0,13	0,11	0,29	0,30	0,30	0,57	0,84	0,51	0,25	0,18	0,16
O1_T6						0,54	0,24	0,16	0,21	0,20	0,28	0,36	0,52	0,86	0,50	0,23	0,19
O1_T7							0,48	0,22	0,11	0,10	0,21	0,25	0,22	0,50	0,84	0,42	0,21
O1_T8								0,56	0,04	0,06	0,11	0,19	0,18	0,27	0,47	0,87	0,54
O1_T9									0,03	0,04	0,06	0,11	0,16	0,18	0,26	0,54	0,88

Tableau 22 - Partie de matrice de moyennes de corrélations entre les électrodes aux différents temps (temps 1 à temps 9) pour les électrodes O2 et Oz

IDte_1	O2_T2	O2_T3	O2_T4	O2_T5	O2_T6	O2_T7	O2_T8	O2_T9	Oz_T1	Oz_T2	Oz_T3	Oz_T4	Oz_T5	Oz_T6	Oz_T7	Oz_T8	Oz_T9
O2_T1	0,63	0,34	0,32	0,23	0,28	0,18	0,12	0,06	0,86	0,54	0,35	0,29	0,29	0,24	0,13	0,08	0,01
O2_T2		0,59	0,37	0,25	0,27	0,26	0,12	0,10	0,56	0,81	0,54	0,33	0,31	0,26	0,16	0,05	0,02
O2_T3			0,63	0,31	0,26	0,30	0,11	0,05	0,35	0,56	0,85	0,52	0,31	0,27	0,24	0,07	0,03
O2_T4				0,60	0,40	0,29	0,19	0,12	0,35	0,39	0,59	0,85	0,57	0,40	0,28	0,18	0,11
O2_T5					0,60	0,27	0,23	0,23	0,27	0,25	0,32	0,56	0,82	0,52	0,27	0,26	0,20
O2_T6						0,47	0,30	0,29	0,28	0,23	0,27	0,36	0,52	0,79	0,44	0,27	0,22
O2_T7							0,54	0,29	0,19	0,22	0,24	0,29	0,29	0,43	0,80	0,43	0,26
O2_T8								0,56	0,09	0,07	0,06	0,21	0,27	0,28	0,46	0,81	0,51
O2_T9									0,04	0,04	0,04	0,16	0,23	0,25	0,28	0,52	0,80

Tableau 23 - Partie de matrice de moyennes de corrélations entre les électrodes auxdifférents temps (temps 1 à temps 9) pour les électrodes Oz et Pz

IDte_1	Oz_T2	Oz_T3	Oz_T4	Oz_T5	Oz_T6	Oz_T7	Oz_T8	Oz_T9	Pz_T1	Pz_T2	Pz_T3	Pz_T4	Pz_T5	Pz_T6	Pz_T7	Pz_T8	Pz_T9
Oz_T1	0,60	0,40	0,32	0,33	0,29	0,18	0,08	0,04	0,69	0,54	0,41	0,29	0,28	0,21	0,13	0,07	0,02
Oz_T2		0,63	0,39	0,37	0,27	0,19	0,09	0,05	0,51	0,70	0,53	0,37	0,35	0,26	0,20	0,12	0,06
Oz_T3			0,59	0,34	0,32	0,25	0,11	0,07	0,41	0,54	0,72	0,52	0,42	0,32	0,27	0,16	0,06
Oz_T4				0,62	0,41	0,33	0,23	0,17	0,36	0,40	0,58	0,71	0,58	0,37	0,33	0,24	0,16
Oz_T5					0,60	0,32	0,28	0,22	0,33	0,36	0,41	0,52	0,72	0,49	0,32	0,25	0,20
Oz_T6						0,59	0,32	0,24	0,28	0,31	0,36	0,39	0,53	0,71	0,49	0,36	0,27
Oz_T7							0,53	0,29	0,20	0,22	0,27	0,34	0,34	0,46	0,67	0,52	0,37
Oz_T8								0,62	0,04	0,09	0,16	0,26	0,29	0,32	0,46	0,65	0,47
Oz_T9									0,04	0,08	0,09	0,16	0,23	0,27	0,31	0,47	0,67

Comme on peut voir, les corrélations qui sont sur la diagonale de la matrice ont la même signification. C'est-à-dire que le fait de calculer la corrélation entre le temps 1 et le temps 2 est la même que le fait de calculer la corrélation entre le temps 2 et le temps 3, car entre la différence entre le temps 1 et le temps 2 et la différence le temps 2 et le temps 3 sont tous d'un point de temps. Sachant ce fait, nous allons donc faire la moyenne des diagonales pour résumer les résultats. La comparaison à l'aide des graphiques à barres sont présentés la Figure 16 pour mieux présenter les résultats. Chaque graphique représente les paires d'électrodes liées à une électrode donnée.







Figure 16 - Graphiques à bande des moyennes de corrélations agrégées en fonction des mêmes nombres de points de temps séparant deux points de temps.

Note : les 19 électrodes sont séparées dans 19 cadrans différents, leurs paires associées sont sur l'axe des x

Comme on peut voir dans ces multiples graphiques, il est évident que pour chaque électrode, la corrélation la plus élevée est celle avec lui-même et que la corrélation baisse au fur et à mesure que le temps est long entre deux points de temps. Il se présente quelques exceptions comme montre dans les graphiques, mais pas assez pour contredire cette évidence. On remarque aussi que les électrodes ont tendance à être plus corrélées avec les électrodes qui sont dans le même hémisphère et en fonction des zones qui sont plus proches d'elles. Par exemple, si on prend l'électrode C3, on voit que les électrodes les plus corrélées avec elle sont, en ordre décroissant de corrélations, P3, F7, F3 et P7. Comme on peut voir, ces électrodes sont toutes

dans l'hémisphère gauche, le même hémisphère que C3. Encore pour C3, les électrodes les moins corrélées avec elle sont, en ordre croissant de corrélation, P4, Fp2, C4 et F4. Ces électrodes sont situées toutes dans l'hémisphère droit du cerveau. Les électrodes qui sont plus ou moins corrélées avec C3 sont majoritairement les électrodes sur la ligne verticale du centre : Fpz, Pz, Fz, Oz et FCz (en ordre décroissant de corrélations). Bien certainement, les exceptions se présentent comme l'électrode F8 est plus corrélée avec C3 comparativement aux électrodes qui sont sur la ligne centrale. Ce résultat peut être dû au manque de données pour l'électrode F8 qui ne permettent pas d'estimer la corrélation comme avec les autres électrodes. Après avoir obtenu ces résultats, quelques recherches ont été faites pour mieux comprendre ce phénomène. Il s'avère que la recherche dirigée par Deligianni et al. (2014) a aussi remarqué une plus grande connectivité entre intrahémisphérique, ce qui appuie le résultat obtenu dans cette partie de l'analyse.

Enfin, on voit que la corrélation entre FCz et Pz sont, en moyenne, négative, ce qui veut dire que, dans la majorité des cas, l'activité des deux électrodes vont à l'encontre de l'un et l'autre. D'après les recherches, il est plausible que ce résultat peut être expliqué par les dipôles existants sur la surface du crâne. En fait, les dipôles sont présents sur le crâne du cerveau humain. Ces dipôles sont nommés, plus spécifiquement, les dipôles équivalents, dont la définition est un peu différente de la définition des dipôles qu'on entend souvent dans le domaine de la physique. En physique les dipôles sont présents dans un espace infini, tandis que les dipôles équivalents sont présents seulement sur le crâne humain à espace limité. Ce qu'ils ont en commun c'est le fait que les dipôles sont des charges électriques qui sont de signes contraires (Shipton, 1991). Comme les activités du cerveau sont principalement des activités électriques du système nerveux central, on retrouve donc des dipôles de charges électriques sur la surface du cuir chevelu. C'est donc une explication possible de la corrélation moyenne négative entre les deux électrodes FCz et Pz.

# Recherche de solution pour l'analyse d'un grand nombre de données

Comme il y a environ 50 000 (environ 3 minutes de pause par participant) enregistrements pour chaque électrode pour chacun des participants, il n'est clairement pas représentatif si on utilise seulement quelques dizaines de points dans le temps de chaque électrode pour les analyses. Pour essayer de résoudre ce problème, il y a deux solutions possibles, le premier étant d'agréger les données en faisant la moyenne des fréquences de plusieurs temps et le deuxième étant de prendre seulement les fréquences de quelques points dans le temps en faisant des sauts. Par exemple on prend seulement les valeurs de fréquence correspondant aux temps 100, 200, 300, 400 ... , etc.

### Solution 1 : Agrégation des données par moyenne

La condition pour appliquer la première solution (agrégation des données) est que les données des différents temps aient une haute corrélation. Donc, si on veut agréger les fréquences en agrégeant, par exemple, par moyennes de chaque 20 points dans le temps, il faut que la corrélation de la fréquence des électrodes entre le temps 1 et le temps 2 soit aussi élevée que celle entre les temps 1 et 20. On aimerait donc que ces corrélations soient stables dans une durée de temps où on veut agréger les fréquences en moyenne. Comme on a vu précédemment, les corrélations de Pearson produites pour chaque électrode à travers le temps résultent à une courbe descendante, ce qui veut dire que la corrélation à travers le temps n'est pas stable. On conclut que ce ne serait pas possible d'agréger les fréquences à travers le temps. On veut alors vérifier si on obtient le même type de tendance de la corrélation si on fait des sauts dans le temps pour savoir si le fait de faire des sauts est l'une de solution possible.

### Solution 2 : Saut dans le temps

Par saut, on entend le fait prendre les données de fréquences des électrodes à chaque X nombres de points de temps. Par exemple si on dit qu'on va faire des sauts de 5 points dans le temps, on va aller chercher les fréquences aux temps 5, 10, 15, 20, etc. Ici, on a choisi trois types de sauts, des sauts de 5 points dans le temps, des sauts de 20 points dans le temps et des sauts de 100 points dans le temps. Pour chaque type de sauts, quatre périodes de 20 sauts de temps ont été
choisies pour chacune des trois pauses. Pour être plus clair, on parle ici de 20 sauts de 5 points dans le temps débutant aux temps 1000, 3000, 5000 et 7000 à chacune des trois pauses pour chacune des électrodes, 20 sauts de 20 points dans le temps débutant aux temps 1000, 3000, 5000 et 7000 à chacune des trois pauses pour chacune des électrodes et enfin 20 sauts de 100 points dans le temps débutant aux temps 1000, 3000, 5000 et 7000 à chacune des trois pauses pour chacune des électrodes et enfin 20 sauts de 100 points dans le temps débutant aux temps 1000, 3000, 5000 et 7000 à chacune des trois pauses pour chacune des électrodes et enfin 20 sauts de 100 points dans le temps débutant aux temps 1000, 3000, 5000 et 7000 à chacune des trois pauses pour chacune des électrodes pour chacune des électrodes. On a donc douze périodes sélectionnées pour chaque électrode. Les corrélations de Pearson entre les temps sautés sont générées pour chaque type de saut, pour chaque période et pour chaque électrode. Les graphiques de tendance des moyennes de corrélations de Pearson pour ces trois types de saut ont été générés. Pour savoir si le saut peut être la solution au gros volume de points de temps, il faut voir si la tendance de la corrélation à travers le temps est la même que lorsqu'on ne faisait pas de saut.

Pour commencer, on va regarder le graphique de la Figure 17 représentant les moyennes de corrélations entre les activités cérébrales des différents points de temps lorsqu'on fait des sauts de 5 points dans le temps. En regardant le graphique ci-dessous, on remarque que la tendance est descendante et qui s'approche de zéro lorsque la distance dans le temps est entre 0 et 40, mais par la suite, la tendance est plutôt montante, ce qui diffère de la tendance qu'on a vue précédemment qui est strictement descendante. Cette tendance n'est donc pas stable pour pouvoir faire des sauts de 5 points dans le temps.



Figure 17 - Graphique des moyennes de corrélations entre les activités cérébrales des différents points de temps lorsqu'on fait des sauts de 5 points dans le temps

Ensuite, regardons la courbe de tendance illustrée dans la Figure 18 pour les moyennes de corrélations des activités des électrodes entre les différents temps lorsqu'on fait des sauts de 20 points dans le temps. Le graphique ci-dessous montre une descente autour de 40 et remonte pour ensuite avoir une tendance qui est assez aléatoire. Cette instabilité fait en sorte que le saut de 20 points de temps ne soit pas un bon choix pour réduire le nombre de points dans le temps, car elle n'est pas du tout similaire à la tendance strictement descendante observée pour les corrélations de Pearson sans saut dans le temps.



Figure 18 - Graphique des moyennes de corrélations entre les activités cérébrales des différents points de temps lorsqu'on fait des sauts de 20 points dans le temps

Finalement, voyons la tendance des moyennes de corrélations des activités des électrodes à travers le temps lorsqu'on fait des sauts de 100 points dans le temps. Si on regarde le graphique de la Figure 19, on remarque qu'il ne semble pas y avoir une tendance claire et remarquable. Il y a des montées et des descentes assez imprévisibles. Encore une fois, cette tendance ne ressemble aucunement à la tendance descendante des corrélations générées que lorsqu'on ne faisait pas de saut dans le temps. Les sauts de 100 points dans le temps ne sont donc pas plus un bon choix pour réduire le nombre de points dans le temps.



Figure 19 - Graphique des moyennes de corrélations entre les activités cérébrales des différents points de temps lorsqu'on fait des sauts de 100 points dans le temps

Après avoir testé ces trois types de sauts, il s'avère qu'on ne voit pas du tout la même tendance lorsqu'on compare avec le graphique qu'on a généré pour les tendances des moyennes de corrélations sans saut dans le temps. Lorsqu'on ne fait pas de saut dans le temps, on voit une tendance de la corrélation qui s'approche de zéro au fur et à mesure que le temps s'éloigne, mais avec les sauts on ne voit aucune tendance particulière. On peut même dire que ces tendances semblent être assez instables et aléatoires. La deuxième solution n'est donc pas applicable. Les problèmes causés par un grand volume de données ne peuvent donc pas être résolus.

# Chapitre 3 Méthodologie

Pour faire l'analyse des relations entre les différentes zones du cerveau, une méthodologie est établie en fonction des méthodes d'analyses déjà existantes. Dans le chapitre de la revue de littérature, nous avons vu trois principales méthodes d'analyses de relations entre les électrodes. Selon les résultats des recherches antérieurs, les modélisations des structures de covariance nous semblent les plus appropriées pour notre étude actuelle. La réflexion sur la méthode choisie et la procédure de modélisation seront décrites en détail dans les prochaines soussections.

# Tests de structures de covariance

Pour découvrir les différentes activités du cerveau, les structures de covariance des modèles mixtes seront analysées. Comme l'intérêt de cette recherche est d'analyser les relations entre les activités cérébrales, les modèles mixtes, qui vont être utilisés dans l'analyse principale de cette recherche, permettent justement de générer des corrélations à l'aide des différentes structures de covariance. En effet, comme les données de fréquence sont des mesures répétées, il est possible d'utiliser les modèles mixtes pour analyser les relations. On a choisi de faire l'analyse par des modèles mixtes, car selon les revues de littératures, les modèles multivariés autorégressifs sont souvent utilisés pour analyser ce type de données et ces modèles sont similaires aux modèles mixtes univariés. De plus, les modèles mixtes permettent plus de possibilités selon les différents modèles existants. On a donc plus de variabilités au niveau des analyses. Les corrélations de Pearson et les analyses de cohérences n'ont pas beaucoup de variabilités et elles ont été beaucoup utilisées dans les recherches antécédentes. C'est la raison pour laquelle elles ne sont pas sélectionnées comme la méthode principale de cette recherche. Au contraire, les différents modèles mixtes n'ont pas tous été retrouvés dans la revue de la littérature. Il serait donc plus intéressant de découvrir quels sont les modèles mixtes qui ajustent les mieux les données EEG. Il y a trois aspects à l'analyse, celui du temps, celui des zones du cerveau et les deux en même temps. La variable « fréquence » est sujette à notre analyse et se répète tout au long des points de temps et des électrodes. Dans ce cas, pour faire les analyses des modèles mixtes univariés, nous allons donc fixer les temps ou les électrodes une à la fois. Finalement, nous allons laisser les électrodes et temps de manière variables pour analyser le modèle multivarié. Les modèles utilisés pour chacune des situations, temps fixes, électrodes fixes ou les modèles multivariés, proviennent du logiciel SAS et les explications de ceux-ci qui sont présentés dans les prochaines sous-sections proviennent du SAS Help Center (« The Mixed Procedure », 2019).

# Analyse des corrélations entre les électrodes

La première analyse consiste à fixer les points de temps pour analyser les relations entre les électrodes. On a fixé sept points de temps différents de chacune des trois pauses pour analyser les tendances des activités intraélectrodes pour un temps donné. Comme ce qui peut quantifier les relations entre les électrodes sont les distances entre elles, on a donc utilisé les coordonnées cartésiennes, produites précédemment, dans les modélisations de modèles mixtes. Les modélisations ont été faites sur les modèles « puissance spatiale » (SP(POW)), « puissance anisotrope spatiale » (SP(POWA)), « exponentielle spatiale » (SP(EXP)), « exponentielle anisotrope spatiale » (SP(EXPA)), « linéaire spatial » (SP(LIN)), « Gaussian spatial » (SP(GAU)), « Matérn spatial » (SP(MATERN)), « sphérique spatial » (SP(SPH)) et « linéaire exposant autorégressif spatial » (SP(LEAR)). Étant donné que la variable « fréquence » permet de quantifier les activités des différentes zones du cerveau, elle est donc utilisée comme la variable réponse sans aucune variable explicative, car nous n'avons pas besoin de variable pour expliquer les variations de la fréquence. Les variables permettant d'identifier la distance entre les électrodes sont aussi utilisées. Ils sont nommés soit en tant que X, Y et Z pour les coordonnées tridimensionnelles, soit en tant qu'U et V pour les coordonnées bidimensionnelles.

Les sous-sections ci-dessous expliqueront les détails des structures de covariance utilisées et modélisées. Les formules et les informations en lien avec les structures proviennent du site web officiel de SAS pour la procédure « mixed » qui a été utilisée pour les analyses (« The Mixed Procedure », 2019).

Pour diminuer les répétitions dans les explications, les définitions des variables qui sont communes à toutes les structures sont les suivantes :

- $\sigma^2$  étant la variance résiduelle
- $\rho$  étant le paramètre autorégressif
- d<sub>ij</sub> étant la distance absolue entre l'électrode sur les points i et l'électrode sur les points j du plan cartésien. Noter que d<sub>ii</sub> et d<sub>jj</sub> prennent la valeur de 0 puisque c'est la distance entre le même point donné.
- *c-list* étant la liste des variables représentant les coordonnées cartésiennes. Dans notre cas, nous aurons soit « X Y Z », soit « U V » pour *c-list*. Dans toutes les formules qui suivent, *c* représente le nombre de variables représentant les coordonnées. Donc pour nos modélisations tridimensionnelles, *c* = 3 (pour les trois variables « X », « Y » et « Z » ) et pour nos modélisations bidimensionnelles, *c* = 2 (pour les deux variables « U » et « V » ).

# Puissance spatiale (SP(POW))

La structure de covariance de puissance spatiale a les caractéristiques suivantes :

Structure	Description	Nombre de paramètres	Éléments (i,j)
SP(POW)( <i>c</i> -list)	Puissance	2	$\sigma^2  ho^{d_{ij}}$ (4)

Comme on peut voir, cette structure a une variance homogène et les corrélations hétérogènes. Il y a deux paramètres à estimer :  $\sigma^2$  et  $\rho$ . Selon la formulation de cette structure, lorsqu'on éloigne en termes de distance, la corrélation devrait s'approcher de zéro, puisque  $\rho$  étant entre -1 et 1, plus que la distance entre deux électrodes  $d_{ij}$  augmente et mise en exposant sur  $\rho$  plus que le  $\rho$  se rapproche de zéro.

# Puissance anisotrope spatiale (SP(POWA))

La structure de covariance de puissance anisotrope spatiale a les caractéristiques suivantes :

Structure	Description	Nombre de	Éléments (i,j)	
$SP(POWA)(c_{-}list)$	Puissance	c+1	$\sigma^2 \rho^{d(i,j,1)} \rho^{d(i,j,2)} = \rho^{d(i,j,c)}$ (5)	
	anisotrope	0 71		

La formulation de cette structure nous montre que la variance est homogène et que les corrélations sont hétérogènes. Il y a c + 1 nombre de paramètres à estimer :  $\sigma^2$  et  $\rho^{d(i,j,c)}$ , selon le nombre de variables *c* utilisées pour représenter les coordonnées. Cette structure est similaire à la structure puissance spatiale, mais la  $\sigma^2$  est multipliée, cette fois, par le produit entre les  $\rho^{d(i,j,k)}$ . On note ici que *c* est le nombre de coordonnées et que d(i, j, k) est la distance absolue entre la  $k^{i \acute{e}me}$  coordonnée (k = 1, ..., c). Par exemple, si on utilise les coordonnées U et V pour modéliser ce modèle, on a donc l'équation suivante qui représente chacune des covariances entre les électrodes i et j :  $\sigma^2 \rho^{d(i,j,U)} \rho^{d(i,j,V)}$ . Donc, d(i, j, U) est la distance absolue entre la coordonnée U de l'électrode située au point i et la coordonnée V de l'électrode située au point *i* et la coordonnée V de l'électrode située au point *j*. La corrélation est donc de  $\rho^{d(i,j,1)}\rho^{d(i,j,2)} \dots \rho^{d(i,j,c)}$  et devrait se rapprocher de zéro au fur et à mesure que la distance augmente.

Pour les deux structures, lorsque les corrélations sont négatives, elles sont multipliées par  $\cos(\pi d_{ij})$  pour prendre en compte la négativité.

# Exponentielle spatiale (SP(EXP))

La structure de covariance d'exponentielle spatial a les caractéristiques suivantes :

Structure	Description	Nombre de paramètres	Éléments (i,j)
SP(EXP)( <i>c</i> - <i>list</i> )	Exponentiel	2	$\sigma^2 \exp\{-d_{ij}/\theta\} \ (6)$

Selon la formulation de cette structure, la variance est homogène et les corrélations sont hétérogènes. Il y a deux paramètres à estimer :  $\sigma^2$  et  $\theta$ . Les corrélations sont de plus en plus proches de zéro au fur et à mesure que la distance augmente. La corrélation suit une fonction exponentielle ayant comme paramètre  $-d_{ij}/\theta$ .

### Exponentielle anisotrope spatiale (SP(EXPA))

La structure de covariance d'exponentielle anisotrope spatiale a les caractéristiques suivantes :

Structure	Description	Nombre de paramètres	Éléments (i,j)
SP(EXPA)( <i>c</i> -list)	Exponentielle anisotrope	2c+1	$\sigma^2 \prod_{k=1}^c \exp\left\{-\theta_k d(i,j,k)^{p_k}\right\} (7)$

Ici, la formulation de la structure montre que la variance est homogène et que les corrélations sont hétérogènes. Il y a 2c+1 paramètres à estimer :  $\sigma^2$ ,  $\theta_c$  et  $p_c$ , selon le nombre de variables *c* utilisées pour représenter les coordonnées. Comparée à la structure exponentielle spatiale, la corrélation de celle-ci est la multiplication des corrélations calculées selon les distances absolues entre les  $k^{ième}$  coordonnées des observations *i* et *j*. On note ici que *c* est le nombre de coordonnées et que d(i, j, k) est la distance absolue entre les kième coordonnées (k = 1, ..., c). *P* est la puissance que l'on peut configurer lors du codage. La corrélation est donc égale à  $\prod_{k=1}^{c} \exp \{-\theta_k d(i, j, k)^{p_k}\}$ .

# Linéaire spatial (SP(LIN))

La structure de covariance de linéaire spatial a les caractéristiques suivantes :

Structure	Description	Nombre de paramètres	Éléments (i,j)
SP(LIN)( <i>c</i> -list)	Linéaire	2	$\sigma^{2}(1 - \rho d_{ij}) 1(\rho d_{ij} \le 1) $ (8)

La formulation de cette structure montre que la variance est homogène et que les corrélations sont hétérogènes. Il y a 2 paramètres à estimer :  $\sigma^2$  et  $\rho$ . Les corrélations sont égales à la différence de 1 et  $\rho d_{ij}$ . La corrélation devrait se rapprocher de zéro au fur et à mesure que la distance augmente. Le tout est multiplié par 1 si  $\rho d_{ij}$  est plus petit ou égal à 1 sinon le tout est multiplié par 0. Cela implique que si  $1 - \rho d_{ij}$  estimée est négatif, la covariance et la corrélation entre l'électrode au point *i* et l'électrode au point *j* est automatiquement égale à 0. Cette structure de covariance ne permet donc pas les corrélations négatives.

# Gaussian spatial (SP(GAU))

La structure de covariance de Gaussian spatial a les caractéristiques suivantes :

Structure	Description	Nombre de paramètres	Éléments (i,j)
SP(GAU)( <i>c</i> -list)	Gaussian	2	$\sigma^2 \exp\{-d_{ij}^2/\rho^2\}$ (9)

Selon la formulation de cette structure, la variance est homogène et les corrélations sont hétérogènes. Deux paramètres sont estimés :  $\sigma^2$  et  $\rho^2$ . Les corrélations sont de plus en plus proches de zéro au fur et à mesure que la distance augmente. Cette structure ressemble beaucoup à la SP(EXP), car la corrélation s'approche aussi de zéro selon une fonction exponentielle, mais selon cette structure, la corrélation diminue plus rapidement que la SP(EXP).

# Matérn spatial (SP(MATERN))

La structure de covariance de Matérn spatial a les caractéristiques suivantes :

Structure	Description	Nombre de paramètres	Éléments (i,j)
SP(MATERN)( <i>c</i> -list)	Matérn	3	$\sigma^{2} \frac{1}{\Gamma(\nu)} \left(\frac{d_{ij}}{2\rho}\right)^{\nu} 2K_{\nu}(d_{ij}/\rho)  (10)$

Selon la formulation de cette structure, la variance est homogène et les corrélations sont hétérogènes. Dans cette équation  $K_{\nu}$  est la fonction Bessel de deuxième type d'ordre, ce qui veut dire que  $\nu > 0$ . L'équation qui définit  $K_{\nu}(x)$  est de (« Bessel Function of the Second Kind », 2022) :

$$-\frac{2(\frac{1}{2}x)^{-\nu}}{\sqrt{\pi}\Gamma(\frac{1}{2}-\nu)}\int_{1}^{\infty}\frac{\cos(xt)dt}{(t^{2}-1)^{\nu+1/2}}(11)$$

Lorsque  $\nu = 0.5$ , ce modèle est équivalent au modèle exponentiel spatial. Lorsque  $\nu \to \infty$ , ce modèle se rapproche du modèle gaussian spatial. La corrélation est de  $\frac{1}{\Gamma(\nu)} \left(\frac{d_{ij}}{2\rho}\right)^{\nu} 2K_{\nu}(d_{ij}/\rho)$ .

# Sphérique spatial (SP(SPH))

La structure de covariance de sphérique spatial a les caractéristiques suivantes :

Structure	Description	Nombre de paramètres	Éléments (i,j)
SP(SPH)( <i>c</i> -list)	Sphérique	2	$\sigma^2 \left[ 1 - \left( \frac{{}^{3d}_{ij}}{2\rho} \right) + \left( \frac{d_{ij}^3}{2\rho^3} \right) \right] 1(d_{ij} \le \rho)  (12)$

Cette structure a une variance homogène et les corrélations hétérogènes. Deux paramètres sont estimés :  $\sigma^2$  et  $\rho$ . La corrélation est égale à  $1 - \left(\frac{{}^{3d}{ij}}{{}^{2}\rho}\right) + \left(\frac{{}^{d}{i}{}^{3}}{{}^{2}\rho^3}\right)$  et devrait se rapprocher de zéro au fur et à mesure que la distance augmente. Le tout est multiplié par 1 lorsque  $d_{ij} \le \rho$ , sinon le tout est multiplié par 0. Cela implique que seules les corrélations positives sont acceptées sinon celles-ci sont directement évaluées à 0, car la condition  $d_{ij} \le \rho$  fait en sorte que le résultat de  $1 - \left(\frac{{}^{3d}{ij}}{{}^{2}\rho}\right) + \left(\frac{{}^{d}{ij}}{{}^{2}\rho^3}\right)$  soit positif.

# Exposant linéaire autorégressif spatial (SP(LEAR))

La structure de covariance d'exposant linéaire autorégressif spatial a les caractéristiques suivantes :

Structure	Description	Nombre de paramètres	Éléments (i,j)
SP(LEAR)( <i>c</i> -list)	Exposant linéaire autorégressif	3	$\sigma^2 \rho^{d_{min} + \delta[(d_{ij} - d_{min})/(d_{max} - d_{min})]} $ (13)

Cette structure a une variance homogène et les corrélations hétérogènes. Trois paramètres sont estimés :  $\sigma^2$ ,  $\rho$  et  $\delta$ . La corrélation est de  $\rho^{d_{min}+\delta[(d_{ij}-d_{min})/(d_{max}-d_{min})]}$  et devrait s'approcher de zéro au fur et à mesure que la distance augmente, car  $\rho$  étant entre 0 et 1, plus que la puissance augmente, plus que la valeur est proche de zéro. Ici,  $\delta$ , qui est plus grand ou égal à 0, représente la vitesse de décroissance. Donc, plus que cette vitesse estimée est grande plus que la corrélation s'approche rapidement de zéro. La formulation de cette structure suit les règles suivantes :

$$\mathrm{Cov}\left[\xi_i,\xi_j
ight]=\sigma^2 egin{cases} 
ho^{d_{\min}+\delta\left[(d_{ij}-d_{\min})/(d_{\max}-d_{\min})
ight]}&i
eq j ext{ and } d_{\min}
eq d_{\max}\ 
ho^{d_{\min}}&i
eq j ext{ and } d_{\min}=d_{\max}\ 1&i=j \end{cases}$$

Ici, il est à noter que  $Cov[\xi_i, \xi_i]$  représente la covariance entre l'électrode au situé au point *i* du plan cartésien et l'électrode située au point *j* du plan cartésien.

# Analyse des corrélations entre les différents points de temps

La deuxième analyse consiste à fixer les électrodes pour analyser les relations entre les différents points de temps. Malgré que le logiciel SAS est l'un des logiciels le plus puissants pour faire des analyses statistiques, mais le grand volume de points de temps que nous avons ne peut pas toutes être prise en compte pour faire la modélisation des modèles mixtes. On doit donc prendre un petit nombre de points de temps pour les modélisations. On a choisi de modéliser sur 20, 50 et 100 points dans le temps. Ces trois nombres de points de temps représentent respectivement 0.08 seconde, 0.20 seconde et 0.39 seconde. Comme le fait de modéliser sur 20 points dans le temps prend moins de temps d'exécution, on a donc décidé de tester plus de modèles afin de sélectionner ceux qui conviennent mieux aux données. Ensuite, ces modèles sélectionnés seront modélisés sur les cas de 50 et 100 points dans le temps dont les modélisations sont plus longues à être exécutés. Pour les modélisations sur 20 points dans le temps, on a choisi 14 périodes dans chacune des 3 pauses pour chacune des 19 électrodes. Les modèles analysés à cette étape sont « moyenne mobile autorégressive de premier ordre » (ARMA(1,1)), « autorégressif de premier ordre » (AR(1)), « autorégressif de premiers ordres hétérogènes » (ARH(1)), « Toeplitz » (TOEP), « antédépendance de premier ordre » (ANTE(1)), « non structuré » (UN), « symétrie composée » (CS), « symétrie composée hétérogène » (CSH), « analyse factorielle de premier ordre » FA(1) et « Huynh-Feldt » (HF). Lorsque les meilleurs modèles sont sélectionnés, on a choisi 14 périodes de 50 points dans le temps dans chaque pause pour chacune des électrodes. Ensuite, ces modèles sont modélisés sur 5 périodes de 100 points dans le temps pour chacune des électrodes.

Les sections ci-dessous expliqueront les détails des structures de covariance utilisées. Les formules et les informations en lien avec les structures proviennent du site web officiel de SAS pour la procédure « mixed » qui a été utilisée pour les analyses (« The Mixed Procedure », 2019).

Pour diminuer les répétitions dans les explications, les définitions des variables qui sont communes à toutes les structures sont les suivantes :

- $\sigma^2$  étant la variance résiduelle
- $\rho$  étant le paramètre autorégressif

- *i et j* étant le point de temps *i* et le point de temps *j*.
- t étant l'intervalle du nombre de points de temps qui sépare les points de temps i et j

### Moyenne mobile autorégressive de premier ordre (ARMA(1,1))

La structure de covariance de ARMA(1,1) a les caractéristiques suivantes :

Structure	Description	Nombre de paramètres	Éléments (i,j)
ARMA(1,1)	ARMA(1,1)	3	$\sigma^{2}[\gamma \rho^{ i-j -1} 1(i \neq j) + 1(i = j)] \ (14)$

Celle-ci est une structure à moyenne mobile autorégressive de premier ordre dont  $\gamma$  représente le composant de moyenne mobile. Il y a trois paramètres à estimer :  $\sigma^2$ ,  $\gamma$  et  $\rho$ . La variance est homogène et les corrélations sont hétérogènes. La corrélation est de  $\gamma \rho^{|i-j|-1} 1 (i \neq j) + 1 (i = j)$  et s'approche de zéro au fur et à mesure que les points de temps s'éloignent du premier. On note que  $|\rho| < 1$  et  $|\gamma| < 1$ . Pour un point de temps *i* donné, la corrélation du premier point de temps adjacent est de  $\gamma$ , celle du second point de temps est de  $\gamma \rho$ , celle du troisième est de  $\gamma \rho^2$ , etc.

# Processus autorégressif de premier ordre (AR(1))

La structure de covariance d'AR(1) a les caractéristiques suivantes :

Structure	Description	Nombre de paramètres	Éléments (i,j)
AR(1)	Autorégressif d'ordre 1	2	$\sigma^2 \rho^{ i-j } (15)$

Cette structure a une variance qui est homogène et les corrélations hétérogènes. Il y a 2 paramètres à estimer :  $\sigma^2$  et  $\rho$ . Similairement à ARMA(1,1), on peut voir que la corrélation a une puissance qui représente le nombre d'éléments qui sépare le premier élément et l'élément lui-même. La corrélation devrait donc s'approcher de zéro au fur et à mesure qu'on avance dans le temps. On note ici une contrainte de  $|\rho| < 1$ . Pour un point de temps i donné, la corrélation du premier point de temps adjacent est de  $\rho$ , celle du second point de temps est de  $\rho^2$ , celle du troisième est de  $\rho^3$ , etc.

# Processus autorégressif hétérogène de premier ordre (ARH(1))

La structure de covariance de ARH(1) a les caractéristiques suivantes :

Structure	Description	Nombre de paramètres	Éléments (i,j)
ARH(1)	AR(1) hétérogène	t+1	$\sigma_i \sigma_j \rho^{ i-j }$ (16)

Cette structure est similaire à AR(1) à la différence que les variances sont hétérogènes et que le nombre de paramètres à estimer est plus grand et est variable. On doit estimer t + 1 paramètres, car la variance ici est variable selon les différents nombres de temps *t*, c'est ce qui fait que cette structure est hétérogène. Les paramètres à estimer sont :  $\sigma_i$ ,  $\sigma_j$  et  $\rho$ . Pour un point de temps i donné, la corrélation du premier point de temps adjacent est de  $\rho$ , celle du second point de temps est de  $\rho^2$ , celle du troisième est de  $\rho^3$ , etc.

# Toeplitz (TOEP)

La structure de covariance de Toeplitz a les caractéristiques suivantes :

Structure	Description	Nombre de paramètres	Éléments (i,j)
TOEP	Toeplitz	t	$\sigma_{ i-j +1}  (17)$

Selon la formulation de cette structure de covariance, on peut voir que les variances sont hétérogènes et que les corrélations sont aussi hétérogènes. En fait, la structure Toeplitz est similaire à la structure AR(1), car l'AR(1) est un cas spécial du Toeplitz (Kincaid, 2005). Elle estime *t* nombre de paramètres :  $\sigma_{|i-j|+1}$  selon le nombre de points de temps *t* utilisé dans la structure. Les corrélations entre un temps donné et les temps adjacents sont déterminées par :  $\sigma_{|i-j|+1}/\sigma_2$ . Donc la corrélation entre un temps donné et le premier temps adjacent est donc de  $\sigma_2/\sigma_2$ , la corrélation entre un temps donné et le second temps adjacent est donc de  $\sigma_3/\sigma_2$ , la corrélation entre un temps donné et le troisième temps adjacent est donc de  $\sigma_4/\sigma_2$  et ainsi de suite...

# Antedépendance de premier ordre (ANTE(1))

La structure de covariance de ANTE(1) a les caractéristiques suivantes :

Structure	Description	Nombre de paramètres	Éléments (i,j)
ANTE(1)	Antédépendance	2t-1	$\sigma_j \sigma_i \prod_{k=i}^{j-1} \rho_k  (18)$

Comme on peut le voir dans la formulation de cette structure, les variances sont hétérogènes et les corrélations le sont aussi. Il y a 2t-1 nombre de paramètres à estimer :  $\sigma_j$ ,  $\sigma_i$  et  $\rho_k$  selon le nombre temps *t* utilisé dans le modèle. On a la racine de la variance de l'élément étudié (i) multipliée avec celle de l'élément d'intérêt (j) et tout cela est multiplié par le produit de toutes les corrélations à compter du premier élément jusqu'à l'élément d'intérêt. Donc la corrélation entre un temps donné et le premier temps adjacent est de  $\rho_1$ , la corrélation entre un temps donné et le second temps adjacent est de  $\rho_1\rho_2$ , la corrélation entre un temps donné et le troisième temps adjacent est de  $\rho_1\rho_2\rho_3$ , etc. La corrélation s'approche donc de plus en plus de zéro en fonction de l'avancement dans le temps.

# Unstructured (UN)

La structure de covariance UN a les caractéristiques suivantes :

Structure	Description	Nombre de paramètres	Éléments (i,j)
UN	Non structuré	t(t+1)/2	σ <sub>ij</sub> (19)

La structure non structurée, comme le nom le décrit, n'a pas une forme spécifique. Il y a t(t+1)/2 paramètres à estimer :  $\sigma_{ij}$ , selon le nombre de temps *t* utilisés dans le modèle. Chacune des covariances et des corrélations des éléments sont hétérogènes et sont spécifiques à chaque élément. Donc la corrélation entre un temps donné et le premier temps adjacent est de  $\rho_{12}$ , la corrélation entre un temps donné et le second temps adjacent est de  $\rho_{13}$ , la corrélation entre un temps donné et le troisième temps adjacent est de  $\rho_{14}$ , etc. Noter ici que  $\rho_{ij}$  représente la corrélation entre les activités du point de temps *i* et du point de temps *j*.

# Compound symmetry (CS)

La structure de covariance CS a les caractéristiques suivantes :

Structure	Description	Nombre de paramètres	Éléments (i,j)
CS	Symétrie composée	2	$\sigma_1 + \sigma^2 1(i=j) \ (20)$

Cette structure a la variance et la corrélation homogènes pour tous les éléments. Il y a deux paramètres qui sont estimés :  $\sigma_1$  et  $\sigma^2$ . Les corrélations entre les activités des temps i et j sont toutes de  $\sigma_1/(\sigma_1 + \sigma^2)$ .

# Heterogenerous compound symmetry (CSH)

La structure de covariance CSH a les caractéristiques suivantes :

Structure	Description	Nombre de paramètres	Éléments (i,j)
CSH	Symétrie composée hétérogène	t+1	$\sigma_i \sigma_j [\rho 1(i \neq j) + 1(i = j)]  (21)$

Cette structure a des variances hétérogènes et la corrélation homogène. Il y a t + 1 paramètres qui sont estimés :  $\sigma_i$ ,  $\sigma_j$  et  $\rho$  selon le nombre de temps *t* utilisé dans le modèle. On a la racine de la variance de l'élément étudié (*i*) multipliée avec celle de l'élément d'intérêt (*j*) et tout cela multiplié par  $\rho$ . Les corrélations entre les activités des temps *i* et *j* sont toutes de  $\rho$ .

# Factor analytic (FA)

La structure de covariance FA a les caractéristiques suivantes :

Structure	Description	Nombre de paramètres	Éléments (i,j)
FA(q)	Analyse factorielle	$\frac{q}{2}(2t-q+1)+t$	$\sum_{k=1}^{\min(i,j,q)} \lambda_{ik} \lambda_{jk} + \sigma_i^2  1(i=j) \ (22)$

La fonction de facteur analytique a  $\frac{q}{2}(2t - q + 1) + t$  nombre de paramètres à estimer où q est le nombre de facteurs. Cette fonction additionne la multiplication de  $\lambda_{ik}\lambda_{jk} + \sigma_i^2 \mathbf{1}(i = j)$ ,

où les valeurs de  $\lambda_{ik}$  et  $\lambda_{jk}$  sont estimé pour chacune des observations *i* et *j* pour les  $k^{i \partial m e}$ addition. Dans ce cas, *k* prend la valeur de 1 jusqu'au nombre minimum entre *i*, *j* et *q*. Donc, on additionne min(i,j,q) nombre de fois  $\lambda_{ik}\lambda_{jk} + \sigma_i^2 \mathbf{1}(i = j)$ . L'addition de  $\sigma_i^2$  est à condition que i = j.

## Huynh-Feldt (HF)

La structure de covariance HF a les caractéristiques suivantes :

Structure	Description	Nombre de paramètres	Éléments (i,j)
HF	Huynh-Feldt	t+1	$(\sigma_i^2 + \sigma_j^2)/2 + \lambda 1(i \neq j) \ (23)$

Cette structure a les variances hétérogènes. Il y a t + 1 paramètres à estimer :  $\sigma_i^2$ ,  $\sigma_j^2$  et  $\lambda$  selon le nombre de temps *t* utilisé dans le modèle. Les covariances sont équivalentes à la moyenne de leurs variances plus une constante .

# Analyse des corrélations entre les électrodes aux différents moments dans le temps

Comme mentionné dans la revue de littérature, pour modéliser les données contenant deux variables à mesures répétées, les modèles MVAR ont souvent été utilisés. Comme le logiciel SAS est principalement utilisé pour les analyses de ce projet, les modèles seuls qui sont conçus pour les mesures répétées multivariées sont UN@AR(1), UN@CS, UN@UN. Parmi ces trois modèles, seul UN@AR(1) correspond au modèle autorégressif multivarié. Les informations de ce modèle sont expliquées ci-dessous.

Structure	Description	Nombre de paramètres	Éléments (i,j)
UN@AR(1)	Produit direct AR(1)	$t_1(t_1+1)/2 + 1$	$\sigma_{i_1 j_1} \rho^{ i_2 - j_2 }  (24)$

Comme on peut remarquer, les variances et les corrélations sont homogènes. Cette structure est construite à partir du produit entre la matrice de covariance UN qui modélise la covariance à travers les observations multivariées représentant les électrodes et la matrice de covariance AR(1) qui modélise le temps dans notre cas. Pour mieux visualiser comment la matrice de cette structure est construite, voici une représentation plus concrète de la structure :

$$\begin{bmatrix} \sigma_1^2 & \sigma_{21} \\ \sigma_{21} & \sigma_2^2 \end{bmatrix} \otimes \begin{bmatrix} 1 & \rho & \rho^2 \\ \rho & 1 & \rho \\ \rho^2 & \rho & 1 \end{bmatrix} = \begin{bmatrix} \sigma_1^2 & \sigma_1^2 \rho & \sigma_1^2 \rho^2 & \sigma_{21} & \sigma_{21} \rho & \sigma_{21} \rho^2 \\ \sigma_1^2 \rho & \sigma_1^2 & \sigma_1^2 \rho & \sigma_{21} \rho & \sigma_{21} & \sigma_{21} \rho \\ \sigma_1^2 \rho^2 & \sigma_1^2 \rho & \sigma_1^2 & \sigma_{21} \rho^2 & \sigma_{21} \rho & \sigma_{21} \\ \sigma_{21} & \sigma_{21} \rho & \sigma_{21} \rho^2 & \sigma_2^2 & \sigma_2^2 \rho & \sigma_2^2 \rho^2 \\ \sigma_{21} \rho^2 & \sigma_{21} \rho & \sigma_{21} & \sigma_{21}^2 \rho^2 & \sigma_2^2 \rho & \sigma_2^2 \end{bmatrix}$$

Toutes ces structures de covariances nommées dans les dernières sections sont modélisées pour les différentes situations (temps fixe, électrodes fixes et temps et électrodes non fixes). Ensuite, les meilleures seront choisies. Les méthodes de choix de structures de covariances sont décrites dans la section suivante.

# Critères de choix de structures

Le choix de structure de covariance est basé sur les BIC. En effet, comme il y a beaucoup de points de temps à faire modéliser, il y a la possibilité que certaines structures de covariance génèrent beaucoup de paramètres selon le volume de données. Lorsqu'il y a beaucoup de paramètres de covariance qui sont générés, il est possible que le problème de surajustement soit présent. Selon Scott I. Vrieze (2012), le critère de sélection AIC a tendance à surajuster les modèles lorsque la taille de l'échantillon est grande. Pour éviter ces problèmes de surajustement, les BIC sont donc choisis pour la sélection des meilleurs modèles. Pour que les analyses soient cohérentes pour l'entièreté de la recherche, les BIC sont donc utilisés pour la sélection des modèles pour toutes les analyses. Un plus petit BIC est préféré. Les BIC sont classés en ordre croissant de rang. Le plus petit BIC a donc le rang 1. Les détails des différentes analyses faites sont expliqués dans le chapitre 4.

En plus de sélectionner les structures de covariance par la comparaison des BIC, il est aussi important de voir la significativité de celles-ci. Le p-value des tests de vraisemblance est donc utilisé pour évaluer la significativité des structures de corrélation. Lorsque le p-value est plus petit que 5%, on considère que le modèle avec la structure de covariance est utile.

# Chapitre 4 Analyse des modèles mixtes

Dans cette partie, des analyses de régression linéaire à mesures répétées ont été faites pour obtenir les matrices de corrélations selon les paramètres de covariance. Cette analyse a été faite afin de déterminer quels modèles permettent de mieux modéliser les données EEG.

# Analyses des structures modélisant les corrélations entre les électrodes

On commence par analyser les modèles mixtes qui modélisent mieux les relations entre les électrodes. Comme on l'a vu dans les analyses descriptives, les distances ont un effet sur la corrélation. On aimerait donc considérer ceci pour modéliser les relations entre les électrodes. Pour ce faire, les différents points dans le temps ont été choisis pour être fixés. De cette façon, les électrodes deviennent la seule variable selon laquelle se répètent les activités du cerveau à l'intérieur de chaque participant. La variable qui identifie les électrodes est donc utilisée pour définir nos structures de covariance. Les modèles spatiaux sont utilisés pour modéliser les activités des différentes zones du cerveau (les fréquences) selon les distances entre les électrodes pour chaque participant. Dans le logiciel SAS, pour utiliser ces modèles, il faut connaître les coordonnées des électrodes. C'est la raison pour laquelle nous avions fait des calculs pour trouver celles-ci.

Parmi les différents modèles spatiaux, on a choisi d'analyser « puissance spatiale » (SP(POW)), « puissance anisotrope spatiale » (SP(POWA)), « exponentielle spatiale » (SP(EXP)), « exponentielle anisotrope spatiale » (SP(EXPA)), « linéaire spatial » (SP(LIN)), « Gaussian spatial » (SP(GAU)), « Matérn spatial » (SP(MATERN)), « sphérique spatial » (SP(SPH)) et « linéaire exposant autorégressif spatial » (SP(LEAR)). Pour faire les modélisations de ces modèles, on a choisi de prendre sept moments dans les trois pauses, ce qui en fait 21 moments différents. Ces modélisations sont faites en utilisant les quatre types de coordonnées (UV selon les positionnements de Böcker et al. (1994), UV de Khosla et al. (1999), XYZ de Böcker et al. (1994) et XYZ de Khosla et al. (1999)) pour identifier les positionnements des électrodes. On a donc fait 756 modélisations représentant tous les moments, tous les deux types de positionnement fournis par les deux articles et tous les modèles de covariance.

Pour mieux justifier le choix de meilleurs modèles, le BIC est utilisé pour évaluer la qualité de chacun d'eux, car le BIC permet de mieux corriger le surajustement des modèles. Les modèles ayant le plus petit BIC ont une meilleure qualité que les autres pour modéliser la relation entre les électrodes. Les p-values des tests de vraisemblance ont été utilisés pour éliminer les modèles dont la structure de corrélation n'est pas utile. Lorsque les modèles dont les p-values est plus grande que 0.05, ces modélisations sont automatiquement éliminées de la sélection des meilleurs modèles, car les structures modéles, les rangs selon la croissance des valeurs de BIC de chaque modèle ont été produits pour chaque moment sélectionné de chaque pause pour chaque type de coordonnées et pour chaque article. Donc, nous allons avoir des rangs allant de 1 à 9, car nous avons modélisé en tout neuf modèles. Ces rangs sont ensuite moyennés selon chaque modèle, chaque type de coordonnées et chaque article. Les rangs des moyennes sont ensuite classés en ordre croissant par article pour trouver les meilleurs modèles. Les trois meilleurs modèles sont sélectionnés (rangs de moyenne de rangs inférieurs à 4). Ces rangs de moyennes sont encore une fois moyennés selon chaque type de coordonnées et chaque type de coordonnées et chaque type de coordonnées et chaque type de coordonnées (rangs de moyenne de rangs inférieurs à 4). Ces rangs de moyennes sont encore une fois moyennés selon chaque type de coordonnées et chaque type de coordonnées (rangs de moyenne de rangs inférieurs à 4). Ces rangs de moyennes sont encore une fois moyennés selon chaque type de coordonnées et chaque article.

Comme on peut le voir dans le Tableau 24, on a obtenu les moyennes des trois meilleurs rangs moyens pour chacun des modèles et types de coordonnées. Ces moyennes sont ensuite ordonnées de manière croissante pour obtenir une meilleure vue sur les qualités des modèles.

Tableau 24 - Les rangs moyens des modèles pour les deux types de coordonnées selon les positionnements d'électrodes fournis par les deux articles de Bocker et al. (1994) et Bocker et al. (1999)

Article	Type de coordonnées	Modèle	Rang moyen
Böcker et al. (1994)	UV	SP(EXPA)	1.5
Böcker et al. (1994)	UV	SP(LEAR)	1.67
Böcker et al. (1994)	UV	SP(EXP)	2.5
Böcker et al. (1994)	UV	SP(POWA)	3
Böcker et al. (1994)	XYZ	SP(POW)	1.5
Böcker et al. (1994)	XYZ	SP(EXPA)	1.83
Böcker et al. (1994)	XYZ	SP(EXP)	2.67
Böcker et al. (1994)	XYZ	SP(LEAR)	2.75
Khosla et al. (1999)	UV	SP(POW)	1.67

Khosla et al. (1999)	UV	SP(EXP)	2
Khosla et al. (1999)	UV	SP(EXPA)	2.33
Khosla et al. (1999)	XYZ	SP(EXPA)	1
Khosla et al. (1999)	XYZ	SP(LEAR)	2
Khosla et al. (1999)	XYZ	SP(EXP)	2.33
Khosla et al. (1999)	XYZ	SP(POW)	3

En fonction de ce dernier tableau, nous avons compté le nombre de fois que les modèles reviennent parmi les modèles ayant les trois meilleurs rangs. Selon le Tableau 25, les modèles « exponentielle spatiale » (SP(EXP)) et « exponentielle anisotrope spatiale » (SP(EXPA)) sont souvent placés en premier rang. Ces deux modèles peuvent être les meilleurs modèles pour modéliser les distances entre les électrodes. Les deux autres modèles qui sont placés les plus souvent parmi les trois premiers sont « puissance spatiale » (SP(POW)) et « linéaire exposant autorégressif » (SP(LEAR)).

Tableau 25 - Nombre de fois que les modèles du tableau 24 reviennent parmi les meilleurs modèles

Modèle	Compte du nombre de fois que les modèles reviennent comme les meilleurs
SP(EXP)	4
SP(EXPA)	4
SP(LEAR)	3
SP(POW)	3
SP(POWA)	1

# Analyse des paramètres des meilleurs modèles spatiaux

Dans la situation où le temps est fixe, selon les choix des modèles effectués par l'analyse des valeurs de BIC, quatre modèles ont été ressortis comme étant ceux qui modélisent les mieux les distances entre les électrodes : SP(EXP), SP(EXPA), SP(LEAR) et SP(POW).

Pour commencer, on aimerait déterminer lequel des deux types de coordonnées (bidimensionnelles ou tridimensionnelles) est mieux pour modéliser les modèles spatiaux. Encore une fois, on a comparé la moyenne des BIC des différents modèles selon les coordonnées. Une plus petite valeur de BIC préférée. Voici les tableaux qui permettent de comparer les BIC moyens de chacun des modèles.

### Tableau 26 - Comparaison des moyennes de BIC du modèle SP(POW) selon les deux types de coordonnées des positionnements provenant des deux articles de Böcker et al. (1994) et Khosla et al. (1999)

Article	Coordonnées	Moyenne de BIC de SP(POW)
Böcker et al. (1994)	UV	2589.324
Böcker et al. (1994)	XYZ	2569.851
Khosla et al. (1999)	UV	2591.387
Khosla et al. (1999)	XYZ	2571.113

#### Tableau 27 - Comparaison des moyennes de BIC du modèle SP(EXP) selon les deux types de coordonnées des positionnements provenant des deux articles de Böcker et al. (1994) et Khosla et al. (1999)

Article	Coordonnées	Moyenne de BIC de SP(EXP)
Böcker et al. (1994)	UV	2587.698
Böcker et al. (1994)	XYZ	2569.851
Khosla et al. (1999)	UV	2573.291
Khosla et al. (1999)	XYZ	2572.349

#### Tableau 28 - Comparaison des moyennes de BIC du modèle SP(EXPA) selon les deux types de coordonnées des positionnements provenant des deux articles de Böcker et al. (1994) et Khosla et al. (1999)

Article	Coordonnées	Moyenne de BIC de SP(EXPA)
Böcker et al. (1994)	UV	2578.194
Böcker et al. (1994)	XYZ	2566.76
Khosla et al. (1999)	UV	2578.633
Khosla et al. (1999)	XYZ	2563.708

#### Tableau 29 - Comparaison des moyennes de BIC du modèle SP(LEAR) selon les deux types de coordonnées des positionnements provenant des deux articles de Böcker et al. (1994) et Khosla et al. (1999)

Article	Coordonnées	Moyenne de BIC de SP(LEAR)
Böcker et al. (1994)	UV	2567.286
Böcker et al. (1994)	XYZ	2571.259
Khosla et al. (1999)	UV	2579.173

### Khosla et al. (1999) XYZ 2572.772

Comme on peut voir, pour les positionnements d'électrodes fournis par les deux articles : Böcker et al. (1994) et Khosla et al. (1999), il s'agit toujours des modèles utilisant les coordonnées tridimensionnelles des positions des électrodes qui ont les BIC moyens les moins élevés pour chacun des modèles spatiaux. Il y a juste une seule exception pour le modèle exposant linéaire autorégressif dont le BIC moyen pour le positionnement bidimensionnel provenant de l'article de Böcker et al. (1994) est plus faible que celui du positionnement tridimensionnel du même article. On peut donc conclure que les positionnements tridimensionnels sont mieux pour modéliser les activités des électrodes selon les distances. Dans les prochaines analyses des paramètres, seuls les résultats des positionnements tridimensionnels sont analysés.

Pour obtenir et analyser les paramètres, plusieurs modélisations ont été faites pour chacune des quatre modèles. Pour chacun d'entre eux, les modélisations ont été effectuées pour les coordonnées tridimensionnelles provenant des positionnements fournis par les deux articles. Ces modélisations ont été faites sur trois moments (temps 1000, 3000 et 5000) de chacune des trois pauses. Donc, dix-huit modélisations sont faites pour chaque modèle, dont neuf pour chacun des deux articles.

Après l'obtention des paramètres, nous allons prendre seulement les paires d'électrodes contenant l'électrode FCz pour mieux visualiser la tendance des corrélations calculées selon les paramètres. De cette manière, il est possible de faire la comparaison des corrélations des mêmes paires d'électrodes selon estimées par les différents modèles. La raison pour laquelle FCz est choisie est parce qu'elle est l'électrode qui est située le plus au centre parmi les électrodes que nous avons. Aussi, comme il y a au total 19 électrodes, ce qui en fait 171 paires d'électrode différentes, ce qui en fait donc beaucoup à analyser. Nous allons donc choisir une seule électrode et ses paires pour les analyses de paramètres.

Pour mieux résumer les paramètres, les modèles qui ne correspondaient pas totalement aux critères de convergence sont éliminés. Par la suite, pour ceux qui n'ont pas été éliminés, leurs paramètres estimés sont moyennés par article. Ces préparations permettent d'avoir une vue sommaire lors de l'analyse des paramètres.

### Analyse des paramètres du modèle puissance spatiale

Dans cette sous-section les paramètres du modèle puissance spatiale seront analysés. Ce modèle a deux paramètres :  $\sigma^2$  et  $\rho$ . Dans le contexte de notre analyse seul  $\rho$  serait intéressant à analyser, car celui-ci permet de calculer la corrélation entre les différentes électrodes en fonction de leur distance. Selon la description du modèle puissance spatiale dans le chapitre 3, la corrélation est de  $\rho d_{ij}$ . Comme mentionné précédemment, les paramètres provenant de chacune des modélisations sont moyennés en fonction de chacun des positionnements tridimensionnels fournis par les deux articles. Dans le Tableau 30, on retrouve les moyennes des paramètres.

Article	Moyenne de $\rho$	Moyenne de $\sigma^2$
Böcker et al. (1994)	0.383256	45.2596
Khosla et al. (1999)	0.430623	44.47408

Tableau 30 – Les moyennes des paramètres estimés par le modèle puissance spatiale

Pour mieux visualiser les corrélations entre les différentes paires d'électrodes, nous avons reporté, dans le Tableau 31 et le Tableau 32, toutes les corrélations entre toutes les paires d'électrodes associées à l'électrode FCz pour les positionnements d'électrodes fournis dans les deux articles. Voici les tableaux :

Tableau 31 - Corrélations entre les paires d'électrodes en lien avec FCz calculées selon les moyennes des paramètres estimés par le modèle puissance spatiale pour les positionnements tridimensionnels des électrodes selon l'article de Böcker et al. (1994)

Paire d'électrodes avec	Corrélations estimées selon les moyennes des paramètres
FCz	
C3	0.52
C4	0.52
F3	0.56
F4	0.56
F7	0.37
F8	0.37
Fp1	0.41
Fp2	0.41

Fpz	0.42
Fz	0.74
01	0.26
O2	0.26
Oz	0.26
Р3	0.37
P4	0.37
P7	0.28
P8	0.28
Pz	0.42

### Tableau 32 - Corrélations entre les paires d'électrodes en lien avec FCz calculées selon les moyennes des paramètres estimés par le modèle puissance spatiale pour les positionnements tridimensionnels des électrodes selon l'article de Khosla et al. (1999)

Paire d'électrodes avec	Corrélations estimées selon les moyennes des paramètres
FCz	
C3	0.50
C4	0.50
F3	0.54
F4	0.54
F7	0.35
F8	0.35
Fp1	0.39
Fp2	0.39
Fpz	0.39
Fz	0.72
01	0.25
02	0.25
Oz	0.25
P3	0.35
P4	0.35
P7	0.27
P8	0.27

Pz

0.39

### Analyse des paramètres du modèle exponentiel spatial

Dans cette section, on veut analyser les paramètres estimés par le modèle exponentiel spatial. Ce modèle a deux paramètres à modéliser :  $\sigma^2$  et  $\theta$ . Comme la corrélation est plus intéressante pour notre recherche actuelle, seul le paramètre  $\theta$ , qui compose la corrélation, sera considéré dans les prochaines analyses. Selon la formule du modèle, la corrélation est de exp $\{-d_{ij}/\theta\}$ . On va donc analyser les corrélations, selon cette formule, pour les coordonnées tridimensionnelles en fonctions des positionnements fournis dans les deux articles. Dans le Tableau 33, il y a les moyennes des paramètres.

Tableau 33 – Les moyennes des paramètres estimés par le modèle exponentiel spatial

Article	Moyenne de p	Moyenne de $\sigma^2$
Böcker et al. (1994)	1.100903	45.25769
Khosla et al. (1999)	1.162486	42.86967

Pour mieux visualiser les corrélations entre les différentes paires d'électrodes, nous avons reporté, dans le Tableau 34 et le Tableau 35, toutes les corrélations entre toutes les paires d'électrodes associées à l'électrode FCz pour les positionnements d'électrodes fournis dans les deux articles. Voici les tableaux :

### Tableau 34 - Corrélations entre les paires d'électrodes en lien avec FCz calculées selon les moyennes des paramètres estimés par le modèle exponentiel spatial pour les positionnements tridimensionnels des électrodes selon l'article de Böcker et al. (1994)

Paire d'électrodes avec	Corrélations estimées selon les moyennes des paramètres
FCz	
C3	0.54
C4	0.54
F3	0.58
F4	0.58
F7	0.39
F8	0.39

Fp1	0.43
Fp2	0.43
Fpz	0.44
Fz	0.75
01	0.28
02	0.28
Oz	0.28
P3	0.39
P4	0.39
P7	0.30
P8	0.30
Pz	0.44

Tableau 35 - Corrélations entre les paires d'électrodes en lien avec FCz calculées selon les moyennes des paramètres estimés par le modèle exponentiel spatial pour les positionnements tridimensionnels des électrodes selon l'article de Khosla et al. (1999)

Paire d'électrodes avec	Corrélations estimées selon les moyennes des paramètres
FCz	
C3	0.49
C4	0.49
F3	0.53
F4	0.53
F7	0.34
F8	0.34
Fp1	0.38
Fp2	0.38
Fpz	0.38
Fz	0.71
01	0.24
02	0.24
Oz	0.24
P3	0.34
P4	0.34

P7	0.26
P8	0.26
Pz	0.38

### Analyse des paramètres du modèle exponentiel anisotrope spatial

Dans cette section, les paramètres du modèle exponentiel anisotrope spatial. Le nombre de paramètres est égal à 2c+1. Dans le cas des coordonnées tridimensionnelles, il y a trois composants X, Y et Z et les modèles générés par ce type de coordonnées comportera  $2^*3+1 = 7$  paramètres :  $\sigma^2$ ,  $\theta_X$ ,  $\theta_Y$ ,  $\theta_Z$ ,  $P_X$ ,  $P_Y$  et P<sub>Z</sub>. Comme la corrélation est plus intéressante pour notre recherche, seuls les paramètres  $\theta_X$ ,  $\theta_Y$ ,  $\theta_Z$ ,  $P_X$ ,  $P_Y$  et P<sub>Z</sub> seront analysés. Selon la formule du modèle, la corrélation équivaut à  $\exp\{-\theta_X d(i,j,X)^{p_X}\} * \exp\{-\theta_Y d(i,j,Y)^{p_Y}\} * \exp\{-\theta_Z d(i,j,Z)^{p_Z}\}$ . On va donc analyser les corrélations, pour les coordonnées tridimensionnelles en fonction des positionnements fournis par les deux articles. Dans le Tableau 36, il y a les moyennes des paramètres.

 

 Tableau 36 – Les moyennes des paramètres estimés par le modèle exponentiel anisotrope spatial

Article	Moy. $\theta_x$	Moy. $\theta_x$	Moy. $\theta_x$	Moy. $P_x$	Moy. $P_x$	Moy. $P_x$	Moyenne de $\sigma^2$
Böcker et al. (1994)	1.66877	1.11793	6.01199	3.21233	0.60539	2.20568	40.6024
Khosla et al. (1999)	1.01599	0.67635	2.09247	0.60564	2.45348	2.51406	42.18377

Pour mieux visualiser les corrélations entre les différentes paires d'électrodes, nous avons reporté, dans le Tableau 37 et le Tableau 38, toutes les corrélations entre toutes les paires d'électrodes associées à l'électrode FCz pour les positionnements d'électrodes fournis dans les deux articles. Voici les tableaux :

Tableau 37 - Corrélations entre les paires d'électrodes en lien avec FCz calculées selon les moyennes des paramètres estimés par le modèle exponentiel anisotrope spatial pour les positionnements tridimensionnels des électrodes selon l'article de Böcker et al. (1994)

Paire d'électrodes avec	Corrélations estimées selon les moyennes des paramètres
FCz	
C3	0.39
C4	0.39

F3	0.34
F4	0.34
F7	0.04
F8	0.04
Fp1	0.04
Fp2	0.04
Fpz	0.07
Fz	0.90
01	0.00
02	0.00
Oz	0.00
P3	0.11
P4	0.11
P7	0.01
P8	0.01
Pz	0.28

### Tableau 38 - Corrélations entre les paires d'électrodes en lien avec FCz calculées selon les moyennes des paramètres estimés par le modèle exponentiel anisotrope spatial pour les positionnements tridimensionnels des électrodes selon l'article de Khosla et al. (1999)

Paire d'électrodes avec	Corrélations estimées selon les moyennes des paramètres	
FCz		
C3	0.39	
C4	0.39	
F3	0.39	
F4	0.39	
F7	0.07	
F8	0.07	
Fp1	0.09	
Fp2	0.09	
Fpz	0.15	
Fz	0.92	
01	0.03	

02	0.03
Oz	0.04
P3	0.19
P4	0.19
P7	0.04
P8	0.04
Pz	0.41

# Analyse des paramètres du modèle exposant linéaire autorégressif spatial

Dans cette section, on veut analyser les paramètres estimés par le modèle exponentiel spatial. Ce modèle a trois paramètres à modéliser :  $\sigma^2$ ,  $\rho$  et  $\delta$ . Comme la corrélation est plus intéressantes pour notre recherche actuelle, seuls les paramètres  $\rho$  et  $\delta$ , qui compose la corrélation, sera considéré dans les prochaines analyses. Selon la formule du modèle, la corrélation est de  $\rho^{d_{min}+\delta[(d_{ij}-d_{min})/(d_{max}-d_{min})]}$ . On va donc analyser les corrélations pour les coordonnées tridimensionnelles selon les positionnements fournis dans les deux articles. Dans le Tableau 39, il y a les moyennes des paramètres.

Tableau 39 – Les moyennes des paramètres estimés par le modèle linéaire exposant autorégressif spatial

Article	Moyenne de p	Moyenne de $\delta$	Moyenne de $\sigma^2$
Böcker et al. (1994)	0.379386	1.441315	46.27232
Khosla et al. (1999)	0.39277	1.40768	44.20859

Pour mieux visualiser les corrélations entre les différentes paires d'électrodes, nous avons reporté, dans le Tableau 40 et le Tableau 41, toutes les corrélations entre toutes les paires d'électrodes associées à l'électrode FCz pour les positionnements d'électrodes fournis dans les deux articles. Voici les tableaux :

#### Tableau 40 - Corrélations entre les paires d'électrodes en lien avec FCz calculées selon les moyennes des paramètres estimés par le modèle linéaire exposant autorégressif

### spatial pour les positionnements tridimensionnels des électrodes selon l'article de Böcker et al. (1994)

Paire d'électrodes avec	Corrélations estimées selon les moyennes des paramètres
FCz	
C3	0.54
C4	0.54
F3	0.58
F4	0.58
F7	0.40
F8	0.40
Fp1	0.43
Fp2	0.43
Fpz	0.44
Fz	0.74
01	0.29
02	0.29
Oz	0.28
P3	0.39
P4	0.39
P7	0.31
P8	0.31
Pz	0.44

### Tableau 41 - Corrélations entre les paires d'électrodes en lien avec FCz calculées selon les moyennes des paramètres estimés par le modèle linéaire exposant autorégressif spatial pour les positionnements tridimensionnels des électrodes selon l'article de Khosla et al. (1999)

Paire d'électrodes avec	Corrélations estimées selon les moyennes des paramètres
FCz	
C3	0.50
C4	0.50
F3	0.54
F4	0.54
F7	0.36

F8	0.36
Fp1	0.40
Fp2	0.40
Fpz	0.40
Fz	0.70
01	0.26
02	0.26
Oz	0.26
P3	0.36
P4	0.36
P7	0.28
P8	0.28
Pz	0.40

# Sommaire des analyses de l'effet de la distance

Pour faire un résumé des analyses des relations entre les paires, deux graphiques à bandes pour chaque article (coordonnées tridimensionnelles selon les positionnements de Böcker et al. (1994) et selon les positionnements de Khosla et al. (1999)) ont aussi été générés à cet effet. On peut voir, dans la Figure 20 et dans la Figure 21, qu'en général les bandes du modèle SP(EXPA) s'écartent beaucoup des autres bandes. De plus, en général, il semble que la corrélation entre FCz et les électrodes P7, P8, Oz, O1, O2 et Oz sont particulièrement faibles et que la corrélation entre FCz et Fz est toujours la plus élevée.



Figure 20 - Graphique à bande des corrélations entre les paires d'électrodes en lien avec FCz calculée selon les moyennes des paramètres des quatre modèles (SP(POW), SP(EXP), SP(EXPA) et SP(LEAR)) pour les positionnements tridimensionnels des électrodes fournis par l'article de Böcker et al. (1994)



Figure 21 - Graphique à bande des corrélations entre les paires d'électrodes en lien avec FCz calculée selon les moyennes des paramètres des quatre modèles (SP(POW), SP(EXP), SP(EXPA) et SP(LEAR)) pour les positionnements tridimensionnels des électrodes fournis par l'article de Khosla et al. (1999)

Ensuite, on a comparé les corrélations générées par les positionnements des deux articles pour chaque modèle dans Figure 22. Il semble qu'en général, les corrélations estimées selon les positionnements de Böcker et al. (1994) est souvent plus élevé que les corrélations estimées

selon les positionnements de Khosla et al. (1999). Cependant, pour le modèle exponentiel anisotrope spatial, les corrélations estimées selon les positionnements de Khosla et al. (1999) est plus élevé que celles estimées selon les positionnements de Böcker et al. (1994). En général, la différence n'est pas très grande entre les corrélations estimées selon les positionnements des deux articles.

Dans cette même figure, les corrélations de Pearson sont aussi mises à disposition afin de comparer les différents modèles avec les moyennes de corrélations de Pearson obtenu dans les analyses descriptives. Il s'avère que, pour les modèles exponentiels spatiaux, puissance spatiale et linéaire exposant autorégressif spatial, les corrélations obtenues par les modèles se rapprochent beaucoup de la corrélation de Pearson pour les paires d'électrodes FCz – F3, FCz – F4, FCz – F7, FCz – F8, FCz – Fp1, FCz – Fp2, FCz – Fpz et FCz – Fz. Les corrélations estimées par le modèle exponentiel anisotrope spatial semblent ne pas correspondre aux corrélations de Pearson estimées précédemment. Comme toutes les paires d'électrodes qui ont les corrélations similaires aux corrélations de Pearson sont dans la même zone que FCz (la zone frontale) il est probable que les trois modèles relevés ci-dessous sont mieux pour estimer les corrélations entre les électrodes dans la même zone du cerveau.



Figure 22 – Graphique à bande pour fin de comparaison des corrélations calculées selon les positionnements tridimensionnels provenant des deux articles (Böcker et al (1994) et Khosla et al. (1999)) et des corrélations de Pearson pour les quatre modèles différents (SP(POW), SP(EXP), SP(EXPA) et SP(LEAR))

Enfin, après avoir fait toutes ces analyses, on se demande si ces modèles ont des caractéristiques en commun qui font qu'ils modélisent le mieux la distance des électrodes. Dans cette section, les modèles puissance spatiale, exponentielle spatiale, exponentielle anisotrope spatiale et exposant linéaire autorégressif spatial ont été choisis comme ceux qui modélisent le mieux l'effet de la distance entre les électrodes. Premièrement, on peut voir que ces quatre modèles ont tous la distance en puissance sur une constante pour déterminer la corrélation entre deux électrodes à une distance données. Cela nous laisse croire que les corrélations entre les électrodes à des distances données suivent une fonction exponentielle avec la distance comme exposant et la constante estimée comme base. Deuxièmement, on observe que le modèle exponentiel et le modèle exponentiel anisotrope spatial reviennent le plus souvent comme le meilleur modèle. On peut croire que les fonctions exponentielles naturelles (à base e) soient plus convenables pour modéliser les distances afin de déterminer les corrélations.

# Analyses des structures modélisant les corrélations entre les temps

Pour mieux comprendre les tendances des activités des électrodes, on aimerait savoir quels sont les modèles qui permettent de mieux modéliser celles-ci en fonction de l'avancement dans le temps. Une analyse a donc été faite en fixant chacune des électrodes. En fixant chacune des électrodes, le temps devient la seule variable selon laquelle se répètent les activités du cerveau à l'intérieur de chaque participant. Pour comprendre les structures de covariance qui modélisent mieux les activités des électrodes à travers le temps, on a utilisé la variable « fréquence » qui représente les activités du cerveau captées par les électrodes comme la variable dépendante. Ici, le temps est la variable qui définit la structure de covariance. Pour ce faire, il y a trois parties d'analyse qui a été faite. Comme il y a un grand nombre de points dans le temps et qu'il n'y a pas de solution à cela, on a choisi de modéliser sur 20 points dans le temps, 50 points dans le temps et 100 points dans le temps pour voir si les modèles ont des comportements différents pour les différents nombres de temps choisis.

Pour commencer, nous avons modélisé les modèles de covariance sur 20 points dans le temps pour trouver les modèles qui sont plus convenables pour modéliser les activités du cerveau dans le temps afin de faire les modélisations sur les plus grands nombres de temps. Pour ces
premières analyses, on a choisi de modéliser les modèles : « moyenne mobile autorégressive » (ARMA(1,1)), « autorégressif d'ordre 1 » (AR(1)), « Toeplitz » (TOEP), « autorégressif d'ordre 1 hétérogène » (ARH(1)), « antédependence » (ANTE(1)), « non structuré » (UN), « symétrie composée » (CS), « symétrie composée hétérogène » (CSH), « analyse de facteur » FA(1) et « Huynh-Feldt » (HF). 14 périodes de 20 points dans le temps ont été choisies dans chaque pause et pour chaque électrode. Il y a donc 42 modèles qui ont été créés pour chacune des 19 électrodes.

Nous avions choisi le BIC pour évaluer la qualité de différents modèles. Un plus petit BIC représente une meilleure qualité de modèle. Les rangs des BIC sont produits pour chaque période de chaque pause et pour chaque électrode de manière croissante. Les rangs vont de 1 à 9, car à cette étape, nous avons utilisé 9 modèles pour faire les modélisations. Les modèles dont les p-value de leurs tests de vraisemblance qui sont plus grands que 0.05 sont éliminés. On considère seulement les modélisations dont la structure de covariance est significativement utile. Ensuite, les moyennes des rangs sont calculées en agrégeant tous les rangs pour chaque électrode. Les moyennes des rangs sont classées en ordre croissant pour déterminer quels sont les meilleurs modèles pour chacune des électrodes et les trois meilleurs rangs moyens pour chaque électrode sont reportés.

Selon le Tableau 42 - Moyennes et écart-types des rangs de BIC des trois meilleurs modèles pour modéliser les activités cérébrales à travers le temps sur 20 points dans le temps pour chacune des électrodes, les modèles ARMA(1,1) et AR(1) sont les modèles dont la moyenne des rangs sont placés en premier (le meilleur rang moyen) pour les électrodes placées sur les zones centrales (C3 et C4), pariétales (P3, P4, P7 et P8) et occipitales (O1, O2 et Oz). Tandis que le meilleur modèle qui revient le plus souvent pour les électrodes la zone frontale (Fp1, Fp2, Fpz, FCz, Fz, F3, F4, F7 et F8) est assez instable : quatre sur neuf électrodes ont TOEP comme meilleur modèle, deux sur neuf électrodes ont ARMA(1,1) comme meilleur modèle et deux fois sur neuf électrodes ont ANTE(1) comme meilleur modèle.

#### Tableau 42 - Moyennes et écart-types des rangs de BIC des trois meilleurs modèles pour modéliser les activités cérébrales à travers le temps sur 20 points dans le temps pour chacune des électrodes

Électrode	Modèle	Moyenne de rangs de BIC	Écart-type de rangs de BIC
C3	ARMA(1,1)	1.97619	0.896826

C3	AR(1)	2.071429	1.090823
C3	TOEP	2.880952	1.193527
C4	AR(1)	1.857143	1.523338
C4	ARMA(1,1)	1.952381	0.986553
C4	TOEP	3	0.883452
F3	TOEP	3.238095	1.664691
F3	ARMA(1.1)	3.285714	1.502611
F3	AR(1)	3.547619	2.050608
F4	TOEP	2.833333	1.591389
F4	ARMA(1,1)	3.261905	1.767726
F4	ARH(1)	3 452381	1 365112
F7	ARMA(1 1)	1 809524	0.862161
F7	AR(1)	2 190476	1 194256
F7	TOEP	2 738095	1.03734
F8	ANTE(1)	2 238095	1 358075
F8	$\Delta RH(1)$	2.256075	1 1189/3
F8	$\Delta R(1)$	2 738095	1 210916
EC <sub>7</sub>	$\frac{1}{\text{ARMA}(1,1)}$	1 285714	0.45723
FCz	$\Delta \mathbf{R}(1)$	1.203714	0.43723
FCz	TOFP	3 1190/18	0.739228
Fn1	TOEP	2 52381	1 13133
Ep1	$\frac{10EI}{ANTE(1)}$	2.52561	1.15155
Fp1	$\frac{\text{ANIL}(1)}{\text{APMA}(1,1)}$	2.880932	1.40039
Ep2	$\frac{\text{ARWA}(1,1)}{\text{TOEP}}$	2 285714	1.009508
<u> </u>	$\frac{10L1}{APMA(1,1)}$	2.265714	1.040000
<u>гр2</u> Ер2	$\frac{\text{AKWIA}(1,1)}{\text{ANTE}(1)}$	3 404762	1.497900
<u> </u>	ANTE(1)	2 261905	1.449298
<u> </u>	TOEP	2.201905	1.103029
<u> </u>		2.101905	2.041407
<u>грг</u>	$\frac{\text{UN}}{\text{ADMA}(1,1)}$	1 261005	0.406706
<u> </u>	$\frac{AKWIA(1,1)}{AD(1)}$	1.201903	0.490790
	TOED	2.07610	0.382338
	$\frac{10EP}{ADMA(1,1)}$	2.97019	1 420640
01	$\frac{\text{AKMA}(1,1)}{\text{TOED}}$	2.97019	1.439049
	$\frac{10EP}{ADMA(1,1)}$	2 222005	1.317991
02	ARMA(1,1)	2.238093	1.339993
02	AR(1) TOED	2.763714	1.//4284
02	$\frac{10EP}{10EP}$	2.228005	1.1/0808
Oz	$\frac{\text{ARMA}(1,1)}{\text{ARM}(1,1)}$	2.238095	1.122052
Oz	AK(1) TOED	2.019048	1.3/844/
UZ D2		2.928571	1.454897
P3	AKMA(1,1)	1.833333	0.095514
P3	AR(1)	1.833333	1.033976
P3	TOEP	5.04/619	0.986553
P4	AR(1)	1.285/14	0.553/33
P4	AKMA(1,1)	1.880952	0.452763
P4	TOEP	3.119048	0.705462
P7	AR(1)	1.619048	0.660834
<b>P</b> 7	ARMA(1,1)	1.666667	0.611542

P7	TOEP	3.166667	0.85302
P8	AR(1)	1.642857	0.72655
P8	ARMA(1,1)	1.666667	0.686691
P8	TOEP	3.428571	0.703401
Pz	AR(1)	1.714286	0.834784
Pz	ARMA(1,1)	1.97619	0.604378
Pz	TOEP	2.52381	1.06469

Par la suite, les modèles faisant partie des trois meilleurs rangs moyens (les modèles dans le Tableau 42) sont comptés et reportés dans le Tableau 43. En regardant les résultats de ce tableau, il s'avère que TOEP est souvent le meilleur modèle parmi tous les modèles, qu'ARMA(1,1) est deuxièmes modèles qui revient le plus souvent parmi les trois meilleurs modèles et que le modèle AR(1) est le troisième qui revient le plus souvent parmi les trois meilleurs modèles. Les modèles ANTE(1), ARH(1) et UN apparaissent aussi quelques fois parmi les trois meilleures moyennes.

Tableau 43 - Nombre de fois que les modèles reviennent comme les trois meilleurs modèles pour chacune des électrodes pour modéliser les activités cérébrales à travers les durées de 20 points dans le temps

Modèle	Nombre de fois que le modèle revient parmi les meilleurs modèles
TOEP	18
ARMA(1,1)	17
AR(1)	14
ANTE(1)	4
ARH(1)	2
UN	1

Les analyses sur un plus grand nombre de temps seraient souhaitables. Cependant, augmenter les points de temps sur les différents modèles prend beaucoup de temps au logiciel SAS d'effectuer les modélisations. On va donc sélectionner les modèles qui sont les plus utiles à modéliser. D'après les premiers résultats des modélisations sur 20 points dans le temps, on peut éliminer les modèles « unstructured » (UN), « compound symmetry » (CS), « heterogeneous CS » (CSH), « factor analytic » FA(1) et « Huynh-Feldt » (HF). Malgré qu'UN est un modèle qui n'est pas si mauvais, on aperçoit qu'il ne permet pas de modéliser plus de 22 points dans le temps en raison du nombre de paramètres à modéliser, ce qui n'est pas souhaitable dans notre cas étant donné le grand nombre de points de temps que nous avons dans notre jeu de données. De plus, CS, CSH et FA(1) ont souvent un rang de BIC très élevé, donc des BIC très élevés, ce qui nous prouve que ce ne sont pas de meilleurs modèles pour nos données. Le modèle HF ne converge pas souvent, ce qui n'est pas un bon modèle à modéliser prochainement étant donné que la majorité des modélisations ne résulte à rien. On garde seulement ARMA(1,1), AR(1), TOEP, ANTE(1) et ARH(1) pour les prochaines modélisations contenant de plus grands nombres de points dans le temps.

En utilisant les cinq modèles sélectionnés précédemment (ARMA(1,1), AR(1), TOEP, ANTE(1) et ARH(1)), les analyses ont été faites sur 50 points dans le temps. Les mêmes 14 périodes dans le temps ont été choisies pour chaque électrode et dans chaque pause. 42 modèles ont été créés pour chacune des 19 électrodes. Encore une fois, le BIC est utilisé pour évaluer la qualité des modèles. Un plus petit BIC est préféré. Les rangs des BIC ont été produits par pause, période et électrode. Les rangs vont de 1 à 5, car cinq modèles sont analysés à cette étape. Les tests dont les p-value de vraisemblance sont plus grands que 0.05 sont éliminés. On garde seulement les modèles dont la structure de covariance est significativement utile. Les rangs sont ensuite moyennés par électrode. Ces moyennes de rangs ont ensuite été classées ordre croissant et reportées dans le Tableau 44. Seulement les trois meilleures moyennes de rangs de BIC ont été sélectionnées.

En regardant le Tableau 44, on obtient la même conclusion que pour les modèles avec 20 points de temps.

Électrode	Modèle	Moyenne de rangs de BIC	Écart-type de rangs de BIC
C3	ARMA(1,1)	2.714286	0.596155
C3	TOEP	3.190476	1.13133
C3	AR(1)	3.357143	0.932383
C4	ARMA(1,1)	2.642857	0.692166
C4	AR(1)	3.02381	1.070402
C4	TOEP	3.690476	0.923622
F3	TOEP	3.243243	1.064722
F3	ARMA(1,1)	3.714286	1.254955
F3	ARH(1)	4.071429	1.237481
F4	TOEP	2.947368	1.113732
F4	ANTE(1)	3.857143	1.690309

#### Tableau 44 - Moyennes et écart-types des rangs de BIC des trois meilleurs modèles pour modéliser les activités cérébrales à travers le temps sur 50 points dans le temps pour chacune des électrodes

F4	ARH(1)	3.928571	1.155957
F7	ARMA(1,1)	2.309524	0.562577
F7	AR(1)	3.095238	0.878178
F7	TOEP	3.725	0.876693
F8	ANTE(1)	2.119048	1.086556
F8	ARH(1)	2.857143	1.159719
F8	AR(1)	3.071429	1.155957
FCz	ARMA(1,1)	1.785714	0.564637
FCz	AR(1)	2.761905	0.576344
FCz	TOEP	3.6	0.590523
Fp1	TOEP	2.72	0.613732
Fp1	ANTE(1)	2.785714	1.179825
Fp1	ARMA(1,1)	3.595238	1.03734
Fp2	TOEP	2.516129	0.676805
Fp2	ARMA(1,1)	3.357143	0.958185
Fp2	ANTE(1)	3.452381	1.4517
Fpz	ANTE(1)	2.452381	0.802508
Fpz	TOEP	3.08	1.187434
Fpz	ARH(1)	3.857143	0.977089
Fz	ARMA(1,1)	2.119048	0.32777
Fz	AR(1)	3.071429	0.406823
Fz	TOEP	3.853659	0.614857
01	TOEP	3	1.322876
01	ARMA(1,1)	3.642857	1.055102
01	ANTE(1)	4.166667	1.606643
O2	ARMA(1,1)	3.452381	1.40039
O2	TOEP	3.619048	1.08093
O2	AR(1)	3.97619	1.522003
Oz	ARMA(1,1)	3.095238	1.143583
Oz	TOEP	3.309524	1.27811
Oz	AR(1)	3.738095	1.060592
P3	ARMA(1,1)	2.571429	0.703401
P3	AR(1)	2.952381	0.961513
P3	TOEP	3.666667	0.754391
P4	AR(1)	2.47619	0.633923
P4	ARMA(1,1)	2.690476	0.517409
P4	TOEP	3.880952	0.592736
P7	AR(1)	2.809524	0.916997
P7	ARMA(1.1)	2.928571	0.921104
P7	TOEP	3.880952	1.086556
P8	ARMA(1.1)	2.404762	0.496796
 P8	AR(1)	2.666667	0.570266
 P8	TOEP	4.261905	0.496796
Pz	ARMA(1.1)	2.690476	0.64347
Pz	AR(1)	2.97619	0.780497
Pz	TOEP	3.275	1.012423

Nous obtenons encore une fois, en ordre du nombre de fois que le modèle revient parmi les meilleurs modèles, TOEP comme le modèle revenant le plus souvent, ARMA(1,1) le deuxième et AR(1) le troisième parmi les trois meilleurs modèles. ANTE(1) apparaît seulement six fois et ARH(1) apparaît seulement quatre fois parmi les trois meilleurs modèles.

#### Tableau 45 - Nombre de fois que les modèles reviennent comme les trois meilleurs modèles pour chacune des électrodes pour modéliser les activités cérébrales à travers les durées de 50 points dans le temps

Modèle	Nombre de fois que le modèle revient parmi les meilleurs modèles
TOEP	18
ARMA(1,1)	16
AR(1)	13
ANTE(1)	6
ARH(1)	4

Bien que TOEP est le modèle qui revient le plus souvent parmi les meilleurs modèles, mais il ne converge pas toujours. Dans le Tableau 46, on montre les électrodes pour lesquelles TOEP n'a pas convergé pour tous les 42 modélisations. En fait, la requête effectuée pour obtenir ce tableau a recherché les modèles faisant partie des trois meilleurs et qui n'ont pas tous convergé. Sachant cela, on voit que TOEP est le seul modèle qui ne converge pas toujours comparativement aux autres modèles comme ARMA(1,1), AR(1), ANTE(1), etc.

#### Tableau 46 - Liste des modèles qui n'ont pas converti pour les 42 modélisations

Électrode	Modèle	Nombre de fois que le modèle a convergé
F3	TOEP	37
F4	TOEP	38
F7	TOEP	40
FCz	TOEP	40
Fp1	TOEP	25
Fp2	TOEP	31
Fpz	TOEP	25
Fz	TOEP	41
O1	TOEP	41

Les mêmes modélisations ont été faites sur cinq périodes de 100 points dans le temps de la première pause. Seulement cinq périodes ont été choisies, car le temps d'exécution des modèles sur 100 points de temps est vraiment long et car nous avons déjà modélisé ces modèles plusieurs fois pour 20 et 50 points dans le temps. Le fait de modéliser sur 100 points dans le temps est seulement à titre de vérification dans ce cas-ci. Encore une fois, le BIC est utilisé pour évaluer la qualité des modèles. Un plus petit BIC est préféré. Les rangs des BIC ont été produits par électrode. Les rangs vont de 1 à 5, car cinq modèles sont analysés à cette étape.

Les résultats des analyses des modèles sur 100 points dans le temps sont un peu différents. Comme on peut voir dans le Tableau 47, cette fois-ci ARMA(1,1) devient le modèle qui revient le plus souvent comme étant le meilleur modèle. Sur les 19 électrodes, il y a 17 électrodes qui ont ARMA(1,1) comme le meilleur modèle. Cette fois, les électrodes frontales ne sont plus aussi instables en ce qui concerne leurs meilleurs modèles. Seules les électrodes Fpz et O1 n'ont pas ARMA(1,1) comme meilleur modèle. Cela nous laisse croire qu'en augmentant le nombre de points de temps, ARMA(1,1) est le modèle qui permet de mieux modéliser les données à travers le temps, peu importe l'électrode analysée. En effet, il se peut aussi que ce phénomène est causé par la diminution de modélisations faites pour les comparaisons.

Électrode	Modèle	Moyenne des rangs de BIC	Écart-type des rangs de BIC
C3	ARMA(1,1)	1.6	0.547723
C3	TOEP	2.333333	2.309401
C3	AR(1)	2.8	0.83666
C4	ARMA(1,1)	1.8	1.30384
C4	AR(1)	2.4	0.547723
C4	TOEP	2.4	0.894427
F3	ARMA(1,1)	2	1
F3	ARH(1)	2.8	0.83666
F3	AR(1)	2.8	1.643168
F4	ARMA(1,1)	1.8	0.447214
F4	AR(1)	2	1
F4	TOEP	2.4	1.341641

Tableau 47 - Moyennes et écart-types des rangs de BIC des trois meilleurs modèles pour modéliser les activités cérébrales à travers le temps sur 100 points dans le temps pour chacune des électrodes

F7	ARMA(1,1)	1.2	0.447214
F7	TOEP	2	1
F7	AR(1)	2.4	0.547723
F8	ARMA(1,1)	1.8	0.83666
F8	ANTE(1)	2.2	1.643168
F8	AR(1)	2.8	0.83666
FCz	ARMA(1,1)	1	0
FCz	AR(1)	2	0
FCz	TOEP	3	0
Fp1	ARMA(1,1)	1.6	0.547723
Fp1	TOEP	1.8	1.095445
Fp1	ANTE(1)	3.6	1.341641
Fp2	ARMA(1,1)	1.2	0.447214
Fp2	TOEP	2	0
Fp2	AR(1)	3.2	0.447214
Fpz	ANTE(1)	1.8	1.788854
Fpz	TOEP	2	0
Fpz	ARMA(1,1)	2.6	1.341641
Fz	ARMA(1,1)	1	0
Fz	AR(1)	2	0
Fz	TOEP	3	0
01	ANTE(1)	2.4	1.949359
01	ARMA(1,1)	2.6	1.516575
01	ARH(1)	2.6	0.894427
O2	ARMA(1,1)	2	0.707107
O2	AR(1)	2	1.414214
O2	TOEP	3.2	1.095445
Oz	ARMA(1,1)	1.6	0.547723
Oz	AR(1)	2.2	1.643168
Oz	TOEP	2.6	0.894427
P3	ARMA(1,1)	1.4	0.547723
P3	AR(1)	2	0.707107
P3	TOEP	2.6	0.894427
P4	ARMA(1,1)	1.4	0.547723
P4	AR(1)	1.6	0.547723
P4	TOEP	3	0
P7	ARMA(1,1)	1.6	0.547723
P7	AR(1)	1.6	0.894427
P7	ARH(1)	3.2	0.83666
P8	ARMA(1,1)	1.2	0.447214
P8	AR(1)	1.8	0.447214
P8	TOEP	3.2	0.447214
Pz	ARMA(1,1)	1.6	0.547723
Pz	TOEP	2.2	1.095445
Pz	AR(1)	2.2	0.83666

En général, les résultats ont eu peu de changements. Selon le Tableau 48, on a en ordre croissant du nombre de fois que le modèle est parmi les trois meilleurs modèles : ARMA(1,1), AR(1) et TOEP.

#### Tableau 48 - Nombre de fois que les modèles reviennent comme les trois meilleurs modèles pour chacune des électrodes pour modéliser les activités cérébrales à travers les durées de 100 points dans le temps

Modèle	Nombre de fois que le modèle revient parmi les meilleurs modèles
ARMA(1,1)	19
AR(1)	16
TOEP	15
ANTE(1)	4
ARH(1)	3

Encore une fois, TOEP est celui qui ne converge pas toujours. Selon le Tableau 49, sur les cinq modélisations qui ont été faits pour chaque électrode, il y a six électrodes pour lesquelles TOEP n'a pas convergé.

Électrode	Modèle	Nombre de fois que le modèle a convergé
C3	TOEP	3
F7	TOEP	3
FCz	TOEP	4
Fp2	TOEP	3
Fpz	TOEP	3
P4	TOEP	4

Tableau 49 - Liste des modèles qui n'ont pas converti pour les 5 modélisations

D'après toutes ces analyses, nous sommes confiants que les trois meilleurs modèles sont ARMA(1,1), AR(1) et TOEP. Cependant, on remarque que le modèle TOEP semble converger moins souvent que les deux autres modèles. Ce fait montre que TOEP n'est pas un modèle très stable pour être utilisé pour les futures analyses sur les paramètres.

#### Analyse des paramètres des meilleurs modèles

Dans le cas des électrodes fixes, deux modèles ont été ressortis comme étant ceux qui modélise le mieux les activités des zones du cerveau à travers le temps : AR(1) et ARMA(1,1).

#### Choix du nombre de points dans le temps

En augmentant le nombre de points dans le temps, le temps d'exécution augmentera aussi. Étant donné la limite du temps de cette recherche le meilleur nombre de points dans le temps devrait être choisi pour optimiser le temps d'exécution et les résultats obtenus. Il ne faut donc pas choisir le maximum de points que le logiciel peut supporter pour estimer les paramètres des différents modèles. Premièrement, cinq différents nombres de points de temps ont été choisis pour vérifier le temps d'exécution. Les cinq différents nombres de points de temps sont de 500, de 1000, de 2000, de 5000 et de 8000 points. Ces périodes ont été testées avec le modèle AR(1) sur la première pause de l'électrode C3. L'exécution avec ces cinq laps de temps montre que le logiciel SAS prend respectivement :

- 15 secondes pour exécuter les modèles sur 500 points de temps ;
- 1 minute 46 secondes pour exécuter les modèles sur 1000 points dans le temps ;
- 56 minutes pour exécuter les modèles sur 2000 points dans le temps ;
- 22 heures et 21 minutes pour exécuter les modèles sur 5000 points dans le temps ;
- 102 heures et 30 minutes pour exécuter les modèles sur 8000 points dans le temps.

Deuxièmement, les tests ont été effectués pour vérifier la stabilité des paramètres des modèles estimés pour différents nombres de temps. Comme le temps d'exécution est très long pour les nombres de points de temps 5000 et 8000, ces derniers ne sont donc pas tenus comptes à cette étape de l'analyse. Les nombres de points dans le temps choisi sont respectivement, 500, 1000 et 2000. Pour faire les tests, seulement quatre moments ont été choisis à la première pause pour l'électrode C3 afin d'exécuter le test avec le modèle AR(1), car à ce nombre de points de temps, le temps d'exécution est extrêmement long. De plus, cette étape est seulement à titre de vérification de la stabilité des paramètres générés par les données des activités du cerveau. Donc, ce n'est pas utile de faire beaucoup de tests comme ce qu'il a été fait dans les dernières sections. Les résultats sont les suivants :

Temps	LRT p-value	ρ	$\sigma^2$	AIC	BIC
1000	< 0.0001	0.6600	15.4044	57525.7	57528.0
3000	< 0.0001	0.6395	17.0181	59201.4	59203.6
5000	< 0.0001	0.6475	14.9170	57484.6	57486.9
7000	< 0.0001	0.6662	17.9480	59113.3	59115.6

Tableau 50 - Les paramètres estimés du modèle AR(1) pour les quatre moments avecune durée de 500 points de temps

 

 Tableau 51 - Les paramètres estimés du modèle AR(1) pour les quatre moments avec une durée de 1000 points de temps

Temps	LRT p-value	ρ	$\sigma^2$	AIC	BIC
1000	< 0.0001	0.6529	14.5000	114017.8	114020.1
3000	< 0.0001	0.6331	16.5144	118007.9	118010.2
5000	< 0.0001	0.6615	15.2173	114670.1	114672.3
7000	< 0.0001	0.6508	17.1850	118035.4	118037.7

Tableau 52 - Les paramètres estimés du modèle AR(1) pour les quatre moments avec<br/>une durée de 2000 points de temps

Temps	LRT p-value	ρ	$\sigma^2$	AIC	BIC
1000	< 0.0001	0.6506	14.6298	228664.8	228667.1
3000	< 0.0001	0.6577	16.2900	232862.3	232864.5
5000	< 0.0001	0.6500	14.7360	229064.3	229066.6
7000	< 0.0001	0.6545	16.1894	232911.7	232914.0

Comme on peut voir dans le Tableau 50, le Tableau 51 et le Tableau 52, les paramètres semblent être assez similaires lorsqu'on fait la comparaison entre les différents nombres de points de temps testés aux mêmes moments. Même en comparant les différents moments dans le temps, ceux-ci demeurent stables. Ils semblent tous prendre les valeurs autour de 0.65 pour la corrélation et entre 14 et 17 pour la variance. On peut donc conclure que nous n'avons pas besoin de prendre un très grand nombre de points dans le temps pour faire les analyses. Afin de tenir compte du temps d'exécutions des analyses et d'avoir une vue globale de la variation de l'activité du cerveau, le nombre de temps de 1000 points a été choisi pour l'analyse des paramètres.

Pour obtenir et analyser les paramètres des deux modèles, trois modélisations (aux temps 1000, 3000 et 5000) ont été effectuées pour chacune des trois pauses pour chacune des électrodes. Il y a donc en tout 171 modélisations (19 électrodes \* 3 pauses \* 3 temps) qui ont été effectuées pour chacun des deux modèles. Ensuite, l'analyse est faite en deux étapes. La première étant de calculer la moyenne de tous les paramètres des 171 modélisations et de faire une analyse sommaire pour ces deux modèles. La deuxième étant de faire le calcul de la moyenne par électrode pour vérifier la différence des comportements de ces modèles entre les différentes électrodes.

#### Analyse des paramètres du modèle autorégressif de premier ordre

Dans cette sous-section, les paramètres du modèle AR(1) seront analysés. Pour commencer, la moyenne de tous les paramètres a été calculée afin d'avoir une vue sommaire sur la variation de l'activité. Il est important de préciser que les tests de vraisemblance de toutes les modélisations du modèle AR(1) sont significatifs, ce qui signifie que le modèle de covariance est utile pour tous les modèles. La moyenne pour les deux paramètres du modèle est dans le tableau suivant :

Tableau 53 - Moyennes de tous les paramètres estimés pour le modèle AR(1) à 1000points de temps

Moyenne de $\rho$	Moyenne de $\sigma^2$
0.655968	44.90837

Dans le cadre de cette recherche, la valeur de la corrélation est plus intéressante à analyser. Une boîte à moustache est créée afin de mieux visualiser la dispersion des paramètres autorégressifs (rho ou  $\rho$ ). Comme on peut le voir dans la Figure 23, les paramètres autorégressifs prennent les valeurs environ entre 0.5 et 0.85. La moyenne ne s'écarte pas beaucoup de la médiane. Il semble y avoir une faible asymétrie, car la boîte s'approche plus de la gauche. La dispersion ne semble pas être énorme, mais elle n'est pas si faible non plus. Ceci peut être causé par les comportements différents entre les électrodes.



Figure 23 - Boîte à moustache des paramètres  $\rho$  estimés pour le modèle AR(1)

Pour ce modèle, la corrélation est représentée par le paramètre autorégressif (rho ou  $\rho$ ) à une certaine puissance selon l'avancement dans le temps. Comme on peut voir, la moyenne obtenue de celle-ci est de 0.655968. Cette valeur est relativement élevée, ce qui signifie que la corrélation de l'activité du cerveau entre le premier point de temps et le deuxième point de temps est en moyenne de 0.66. Comme on le sait, plus que la corrélation s'approche de |1|, plus que la dépendance entre deux variables est élevée. Donc, notre corrélation AR(1) est  $\rho^{|i-j|}$ , ce qui veut dire que plus qu'on s'éloigne dans le temps moins que la corrélation est élevée. Cela correspond bien à la tendance des corrélations données des activités du cerveau dans le temps, car les corrélations de Pearson montrent aussi cette tendance. Il est raisonnable que les activités soient de moins en moins corrélées avec le premier point de temps analysé. C'est probablement aussi la raison pour laquelle ce modèle est l'un des meilleurs parmi tous ceux qui ont été analysés précédemment.

Ensuite, les moyennes des paramètres de corrélation pour chacune des électrodes sont calculées et reportées dans le tableau suivant. Comme on peut le voir dans le Tableau 54, la majorité des électrodes ne s'écarte pas de la moyenne générale calculée précédemment. Cependant, les électrodes F8, FCz, Fz, O1, O2, Oz et Pz ont un écart important par rapport à la moyenne générale. Plus précisément, en moyenne, les paramètres autorégressifs des électrodes F8, FCz, Fz, Fz et Pz prennent des valeurs plus élevées que celles des autres électrodes. En effet, leurs corrélations sont toutes en haut de 0.75. On peut donc dire que les activités du cerveau sont plus corrélées dans le temps pour les zones qui sont représentées par ces électrodes. Au

contraire, les paramètres autorégressifs des électrodes O1, O2 et Oz sont en moyenne plus faible comparativement à la moyenne générale. Ils prennent une valeur inférieure à 0.57, ce qui signifie qu'en moyenne, les activités du cerveau sont moins corrélées dans le temps pour les zones qui sont représentées par ces électrodes.

Électrode	Moyenne de p	Moyenne de $\sigma^2$
C3	0.646314	14.96814
C4	0.598812	15.7097
F3	0.599366	42.25718
F4	0.582992	41.54795
F7	0.685898	48.17185
F8	0.753923	55.32416
FCz	0.822767	7.241861
Fp1	0.639868	114.8993
Fp2	0.637483	91.90071
Fpz	0.675228	87.37619
Fz	0.793782	23.5331
O1	0.517856	61.67333
O2	0.540585	51.65419
Oz	0.576879	44.93163
P3	0.671652	21.80445
P4	0.695724	22.27431
P7	0.618947	45.83661
P8	0.634025	45.39581
Pz	0.771286	16.75848

Tableau 54 - Moyennes des paramètres estimés pour chacune des électrodes pour le modèle AR(1) à 1000 points de temps

# Analyse des paramètres du modèle moyenne mobile autorégressive de premier ordre

Dans cette sous-section, les paramètres du modèle ARMA(1,1) seront analysés. Pour commencer, la moyenne de tous les moments de toutes les électrodes est calculée. En faisant une moyenne, on a une vue sommaire sur le modèle. Tous les tests effectués ont un p-value de test de vraisemblance significatif, ce qui signifie que ce modèle de covariance est toujours utile pour modéliser les effets du temps sur les activités du cerveau. Comme mentionné précédemment, ce qui est plus intéressant d'analyser est la corrélation. Pour ce modèle, la corrélation est représentée par  $\gamma \rho^{|i-j|-1} 1(i \neq j) + 1(i = j)$ . Donc, les paramètres qui sont

plus importants pour cette analyse sont la moyenne mobile (gamma ou  $\gamma$ ) et le paramètre autorégressifs (rho ou  $\rho$ ). Le Tableau 55 montre la moyenne générale de tous les paramètres obtenus de toutes les modélisations effectuées.

Tableau 55 - Moyennes de tous les paramètres estimés pour le modèle ARMA(1,1) à1000 points de temps

Moyenne de p	Moyenne de y	Moyenne de $\sigma^2$
0.609117	0.655897	44.90973

Dans le cadre de cette recherche, la valeur de la corrélation est plus intéressante à analyser. Deux boîtes à moustache sont créées afin de mieux visualiser la dispersion des paramètres autorégressifs (rho ou  $\rho$ ) et composante de moyenne mobile (gamma ou  $\gamma$ ). Comme on peut le voir dans la Figure 24, les paramètres autorégressifs prennent les valeurs environ entre 0.25 et 0.8. La moyenne a un certain écart avec la médiane. Il semble y avoir une asymétrie, car la boîte s'approche plus de la droite. Il y a une certaine dispersion, car l'étendue des valeurs est grande et la boîte est assez grande aussi. Ceci peut être causé par les comportements différents entre les électrodes. En ce qui concerne la boîte à moustache de la moyenne mobile (Figure 25), les valeurs sont environ entre 0.5 et 0.85. La moyenne a un écart moins élevé avec la médiane. Il y a une faible asymétrie, car la boîte s'approche plus de la gauche. La dispersion ne semble pas aussi élevée comparativement au paramètre autorégressif. Les dispersions pour ces deux paramètres peuvent être en raison de la différence entre les électrodes.



Figure 24 - Boîte à moustache des paramètres  $\rho$  estimés pour le modèle ARMA(1,1)



Figure 25 - Boîte à moustache des paramètres  $\gamma$  estimés pour le modèle ARMA(1,1)

On peut voir que la moyenne de rho est de 0.609117 et celle de gamma est de 0.655897. Selon l'équation du modèle, cela signifie que la corrélation entre deux temps adjacents équivaut à environ 0.61, que la corrélation avec le deuxième point de temps adjacent est de  $0.66 * 0.61^{1} \approx 0.40$ , que la corrélation avec le troisième point de temps adjacent est de  $0.66 * 0.61^{2} \approx 0.25$  et ainsi de suite. On observe aussi une diminution de la corrélation au fur et à mesure que le temps avance tout comme avec le modèle AR(1), ce qui est raisonnable tout comme expliquer dans la dernière sous-section.

Ensuite les paramètres des modélisations sont moyennés selon chacune des électrodes et cellesci sont résumées dans le tableau suivant. On peut observer que la majorité de la moyenne de la moyenne mobile n'a pas un grand écart avec la moyenne générale. Seulement la moyenne des paramètres de moyenne mobile associée aux électrodes F8, FCz, Fz et Pz prennent des valeurs plus élevées comparativement à la moyenne générale et celles associées aux électrodes O1, O2 et Oz sont en moyenne plus faible comparativement à la moyenne générale. Cependant, cette valeur est invariante en fonction du temps. La valeur du paramètre autorégressif prend plus d'importance, car elle varie en fonction du temps avec les différentes puissances qu'elle prend au fur et à mesure que le temps avance. On remarque que cette valeur est vraiment instable en fonction des différentes électrodes. Tenant compte de cette remarque, on peut déduire que pour certaines zones du cerveau, la corrélation des activités du cerveau se rapproche plus rapidement de zéro en fonction du temps comme pour les zones associées aux électrodes F3, F4, Fp1, Fp2, Fpz, O1, O2 et Oz. Pour d'autres zones du cerveau, la corrélation des activités du cerveau se rapproche moins rapidement de zéro en fonction du temps comme pour les zones associées aux électrodes C3, F7, F8, FCz, Fz, P3, P4, P8 et Pz. Seules les électrodes C4 et P7 ont une moyenne de paramètre autorégressif proche de la moyenne générale. Cette instabilité explique donc la dispersion observée dans la boîte a moustache associée à ce paramètre. On peut donc dire que, selon ce modèle, les activités du cerveau ont des comportements différents à travers le temps selon les différentes zones du cerveau.

Channel	Moyenne p	Moyenne y	Moyenne $\sigma^2$
C3	0.745517	0.645499	14.97038
C4	0.684723	0.597923	15.71122
F3	0.521322	0.598864	42.25756
F4	0.444425	0.581632	41.54879
F7	0.789904	0.684043	48.18216
F8	0.73715	0.753654	55.32284
FCz	0.772825	0.82242	7.240704
Fp1	0.408871	0.645736	114.9049
Fp2	0.402771	0.638587	91.89841
Fpz	0.461597	0.681849	87.37556
Fz	0.73623	0.792857	23.53641
01	0.355397	0.514947	61.67222
O2	0.483681	0.539426	51.65446
Oz	0.484539	0.575356	44.93128
P3	0.731744	0.67131	21.80762
P4	0.719322	0.695647	22.27628
P7	0.669044	0.618535	45.83792
P8	0.702196	0.633297	45.40015
Pz	0.721957	0.770463	16.75601

Tableau 56 - Moyennes des paramètres estimés pour chacune des électrodes pour le modèle ARMA(1,1) à 1000 points de temps

#### Sommaire de l'analyse de l'effet du temps

Dans cette section, les modèles ARMA(1,1) et AR(1) ont été choisis comme ceux qui modélisent le mieux l'effet du temps sur l'activité du cerveau. On se demande ce que ces deux modèles ont en commun pour mieux comprendre la tendance des données du cerveau à travers le temps. D'après les analyses, ARMA(1,1) et AR(1) se ressemble beaucoup en termes de structure, avec la différence que la structure d'AMRA(1,1) a une constante de plus. En fait, les deux structures ont les corrélations qui se rapprochent de zéro de façon exponentielle à travers le temps. Pour les deux modèles, l'exposant augmente uniformément en fonction de l'avancement dans le temps. Cette augmentation commence au premier temps adjacent à un temps donné pour le modèle AR(1) et au deuxième temps adjacent à un temps donné pour le

modèle ARMA(1,1). On peut donc croire que les corrélations des activités du cerveau ont tendance à se rapprocher uniformément de zéro et de façon exponentielle à travers le temps.

### Analyse du modèle multivarié

Comme expliqué dans la partie de présentation des structures de covariance, le seul modèle multivarié (MVAR) existant dans le logiciel SAS et qui est possiblement ajustable à nos données d'activités du cerveau est le produit direct autorégressif (UN@AR(1)). En fait, selon les analyses précédentes, AR(1) est l'un des modèles choisis pour modéliser les activités du cerveau à travers le temps, ce qui fait que le modèle MVAR UN@AR(1) semble être le plus approprié pour cette partie de l'analyse. Donc, dans ce cas, la partie UN (non structuré) modélisera les mesures répétées au niveau des électrodes et la partie AR(1) (autorégressif de premier ordre) permettra de modéliser les mesures répétées au niveau du temps. Les deux mesures répétées temps et électrodes auront la possibilité à être modélisés en même temps par un seul et même modèle.

Pour cette analyse, dix périodes de 10 points de temps ont été choisies pour être modélisées par le UN@AR(1). Cependant, après les dix essais, aucun modèle n'a convergé, ce qui fait que les analyses devront s'arrêter à cette étape, car aucun paramètre ne peut être généré dans ce cas.

À la suite de l'obtention de ce résultat, les tests ont été faits uniquement avec les électrodes comme mesures répétées (temps fixe) en utilisant le modèle UN. On a remarqué que ce modèle ne converge pas lorsqu'il essaye de modéliser les activités du cerveau en fonction des différentes électrodes, ce qui nous montre la raison pour laquelle le modèle multivarié UN@AR(1) ne convergeait pas.

Malheureusement, aucun résultat n'est obtenu à cette étape et dû au temps limité par la recherche, il n'a pas été possible de trouver d'autres solutions permettant de produire des modèles MVAR qui permettent de mieux modéliser les données du cerveau.

# Chapitre 5 Limites

Dans le cadre de cette recherche, il y a plusieurs analyses qui seraient intéressantes à être faites, mais qui n'étaient pas effectuées. En effet, la cause est le temps de recherche limité et le gros volume de données non traitable par les modèles de covariance. De plus, la base de données utilisées est basée sur une seule recherche, ce qui fait que les résultats ne peuvent pas être généralisés à tous les cas. Les sections suivantes expliqueront en détail ces trois principales limites et leurs conséquences liées à cette recherche.

# Les conséquences causées par le temps de recherche limité

Comme nous avons vu dans le chapitre 1, l'une des méthodes les plus utilisées pour analyser les données EEG est le modèle multivarié. Ce modèle permet de modéliser les données répétées sur deux variables tout en utilisant un premier modèle pour la première variable et un deuxième modèle pour la deuxième variable. Dans notre cas, la répétition se fait au niveau des électrodes et du temps. Le logiciel SAS contient trois types de modèles multivariés : UN@UN (le modèle non structuré pour la première et la deuxième variable sur lesquelles répètent les données), UN@CS (le modèle non structuré pour la première variable et le modèle de symétrie composée pour la deuxième variable sur lesquelles répètent les données) et UN@AR(1) (le modèle non structuré pour la première variable et le modèle autorégressif pour la deuxième variable sur lesquelles répètent les données). Comme expliqué dans le chapitre 3, seul UN@AR(1) semble être convenable pour notre analyse et il a été modélisé dans le chapitre 4, mais le résultat n'a pas été concluant. Cependant, il est possible de coder manuellement un modèle multivarié tout en associant les bons modèles. Toutefois, en raison de la limite du temps de recherche, cette étape n'a pas été effectuée, ce qui fait que cette analyse n'a pas pu être complétée. Ce serait donc important d'aborder cet aspect dans les recherches futures pour découvrir d'autres résultats intéressants.

### Les conséquences causées par la limite des modèles

En raison du grand volume de données, les modèles mixtes ne sont pas idéals pour modéliser celles-ci. Ceci cause donc un problème pour la modélisation du temps. Plus que le nombre de points de temps utilisé est grand plus que c'est long à modéliser. Comme discuter précédemment, nous avons essayé de trouver des solutions pour modéliser un plus grand nombre de temps possible, mais les solutions ne sont pas applicables pour les données que nous avons actuellement. Le maximum de nombre de points de temps utilisé est de 100 points de temps pour sélectionner les meilleurs modèles et le nombre approximatif de points de temps de chaque pause est de 15 000, ce qui fait que ce nombre 100 est très petit pour représenter les 15 000 points de temps. Il est donc possible que beaucoup d'activités du cerveau ne soient pas captées par les modèles, ce qui biaise nos choix de modèles.

# Les conséquences causées par la limite des bases de données utilisées

La recherche actuelle a utilisé les données d'une expérience passée. On utilise les moments de pause de la partie des tâches N-back. Ce contexte peut générer des activités du cerveau qui est unique à celui-ci. Dans ce cas, il est difficile de dire si les modèles choisis sont spécifiques à ce contexte de recherche ou s'ils peuvent aussi être utilisés pour modéliser les données collectées dans d'autres expériences. Aussi, un autre problème peut aussi constituer un biais pour cette recherche. En fait, lors des pauses des tâches, les professionnels de recherches demandent aux participants de regarder en avant et de relaxer le plus possible. Cependant, rien ne peut garantir que les participants soient réellement en train de relaxer. Personne ne saura s'ils sont en train de planifier leurs projets personnels ou d'essayer de résoudre un exercice de leurs cours de mathématique. Toutes ces situations peuvent demander des charges cognitives aux participants, ce qui fait qu'on ne peut pas garantir que le mode de repos est activé pour tous les participants lors des pauses. Les données EEG que nous analysons peuvent donc ne pas être toutes « neutres » comme nous le voulions.

## Conclusion

Cette recherche a pour but de déterminer les modèles les plus appropriés parmi les différents modèles mixtes qui permettent de modéliser la corrélation entre les électrodes. Nous avons commencé par une analyse descriptive qui a permis de découvrir des faits intéressants en lien avec la relation entre les électrodes. La méthode utilisée pour l'analyse descriptive a été la corrélation de Pearson. Tout d'abord, nous avons remarqué que la corrélation entre les activités des électrodes s'approche de plus en plus de zéro au fur et à mesure que le temps avance. Ensuite, lorsqu'on étudie la corrélation entre les électrodes, on remarque qu'il y a effectivement une relation entre les électrodes augmente plus que la corrélation entre elles s'approche de zéro. Par la suite, nous avons analysé les relations entre les électrodes à travers le temps. Nous avons remarqué, avec évidence, que la corrélation entre une électrode au même moment. Nous avons aussi détecté une corrélation plus élevée entre les électrodes qui sont dans le même hémisphère. Enfin, on a remarqué que la corrélation moyenne entre les électrodes FCz et Pz est négative, ce qui peut être expliqué par les dipôles électriques sur la surface du crâne.

Après avoir fait les analyses descriptives, les analyses des modèles mixtes ont été effectuées. Nous avons commencé par l'analyse des modèles mixtes pour les relations entre les électrodes lorsque le temps est fixe. Il s'avère que les meilleurs modèles spatiaux pour modéliser la distance entre les électrodes sont l'exponentielle spatiale, l'exponentielle anisotrope spatiale, la puissance spatiale et l'exposant linéaire autorégressif. Pour cette partie de l'analyse, nous avons utilisé deux types de coordonnées de positionnement d'électrodes pour l'estimation des distances, les coordonnées bidimensionnelles et les coordonnées tridimensionnelles. Après avoir analysé les résultats obtenus par les modélisations faites en fonction des deux types de coordonnées, il nous semble que les coordonnées tridimensionnelles permettent de mieux estimer les distances entre les électrodes et de modéliser la relation entre celle-ci. De plus, dans cette analyse, les positionnements sont basés sur les positionnements d'électrodes fournis par deux articles, celui de Böcker et al. (1994) et celui de Khosla et al. (1999). On a remarqué que, pour les modèles exponentiels spatiaux, puissance spatiale et exposant linéaire autorégressif, les corrélations estimées en fonction des positionnements de Böcker et al. (1994) sont plus élevées que celles estimées en fonction des positionnements de Khosla et al. (1999) et le modèle exponentiel anisotrope spatial démontre le contraire.

L'autre partie de l'analyse consiste en l'analyse des corrélations à travers le temps lorsque les électrodes sont fixes. Après avoir comparé les différents modèles, il s'avère que le modèle autorégressif de premier ordre et le modèle de moyenne mobile autorégressive de premier ordre sont les meilleurs modèles pour modéliser les activités des électrodes à travers le temps dans le cas où les électrodes sont fixes. En sommes, selon les deux modèles, il y a une différence supérieure à la moyenne des corrélations de toutes les électrodes pour les électrodes F8, FCz et Fz et il y a une différence inférieure à la moyenne des corrélations de toutes les électrodes F8, FCz et Fz et il y a une différence inférieure à la moyenne des corrélations de toutes les électrodes F8, FCz et Fz, la corrélation entre les activités dans le temps devrait se rapprocher moins rapidement de zéro au fur et à mesure que le temps avance et que pour les électrodes O1, O2 et Oz, la corrélation entre les activités dans le temps devrait se rapprocher moins rapidement de zéro au fur et à mesure que le temps avance et que pour les rapidement de zéro au fur et à mesure que le temps avance et que pour les rapidement de zéro au fur et à mesure que le temps devrait se rapprocher plus rapidement de zéro au fur et à mesure que le temps devrait se rapprocher plus rapidement de zéro au fur et à mesure que le temps devrait se rapprocher plus rapidement de zéro au fur et à mesure que le temps devrait se rapprocher plus rapidement de zéro au fur et à mesure que le temps devrait se rapprocher plus rapidement de zéro au fur et à mesure que le temps devrait se rapprocher plus rapidement de zéro au fur et à mesure que le temps devrait se rapprocher plus rapidement de zéro au fur et à mesure que le temps avance.

Dans la dernière partie de l'analyse, nous avions voulu modéliser les corrélations entre les électrodes en fonction du temps et des électrodes, c'est-à-dire qu'on aimerait savoir quelles sont les relations entre les électrodes pour les temps différents. Cependant, les modèles mixtes multivariés existants dans le logiciel SAS ne permettent pas de modéliser les données que nous détenons pour cette recherche. Comme expliqué dans le chapitre 5, à cause de la limite du temps de recherche, il n'a pas été possible de construire des modèles multivariés afin de mieux modéliser et d'adapter à nos données. Cette partie de l'analyse n'a donc pas été concluante.

En bref, les analyses qui ont été faites dans cette recherche ont permis de comprendre quels sont les meilleurs modèles pour modéliser les données EEG dans les différents types de situations. Aussi, on comprend mieux les relations entre les électrodes. Pour les analyses futures, il est fortement suggéré de faire l'analyse en utilisant les données de plusieurs sources d'expériences afin de pouvoir généraliser les résultats de l'analyse à l'ensemble des données EEG à l'état de repos du cerveau. De plus, si possible, il est aussi préférable de construire des modèles mixtes multivariés qui s'adaptent mieux aux données EEG afin de découvrir quels sont les meilleurs modèles qui permettent de modéliser les données EEG lorsque le temps et les électrodes ne sont pas fixés.

## Bibliographie

- ABHANG Priyanka A., GAWALI Bharti W. et MEHROTRA Supresh C. (2016).
   « Chapter 2 Technological Basics of EEG Recording and Operation of Apparatus », Introduction to EEG- and Speech-Based Emotion Recognition, 1<sup>ère</sup> édition, Massachussetts, Academic Press, p. 19-50.
- BABILONI Claudio, BABILONI Fabio, CARDUCCI Filippo, CINCOTTI Febo, VECCHIO Fabrizio, COLA Benedetta, ROSSI Simone, MINIUSSI Carlo et ROSSINI Paolo Maria (2004). « Functional Frontoparietal Connectivity During Short-Term Memory as Revealed by High-Resolution EEG Coherence Analysis », *Behavioral Neuroscience*, vol 118, no 4. Récupéré de https://psycnet.apa.org/doiLanding?doi=10.1037%2F0735-7044.118.4.687.
- BANZ Barbara C., YIP Sarah W., YAU Yvonne HC et Potenza Marc N. (2016).
   « Behavioral addictions in addiction medicine: from mechanisms to practical considerations », *Progress in brain research*, Amsterdam, Elsevier, vol. 223, p. 311-328.
- BASTOS André M. et SCHOFFELEN Jan-Mathijs (2016). « A tutorial review of functional connectivity analysis methods and their interpretational pitfalls », *Frontiers in systems neuroscience*, vol. 9, p. 175. Récupéré de https://www.frontiersin.org/articles/10.3389/fnsys.2015.00175/full.
- BARRY Robert J., CLARKE Adam R., MCCARTHY Rory et SELIKOWITZ Mark (2022). « EEG coherence in attention-deficit/hyperactivity disorder: a comparative study of two DSM-IV types », *Clinical Neurophysiology*, vol. 113, no 4. Récupéré de <u>https://www.sciencedirect.com/science/article/abs/pii/S1388245702000366</u>.
- BERMAN, Jules (2018). « Indispensable Tips for Fast and Simple Big Data Analysis », *Principles and Practice of Big Data*, 2<sup>ième</sup> édition, Cambridge, Academic Press, p. 231-257
- Bessel Function of the Second Kind (2022). Wolfram MathWorld. Récupéré le 27 juillet 2022 de <u>https://mathworld.wolfram.com/BesselFunctionoftheSecondKind.html</u>.
- BHAVSAR Ronakben, SUN Yi, HELIAN Na, DAVEY Neil, MAYOR David and STEFFERT Tony (2018). « The Correlation between EEG Signals as Measured in Different Positions on Scalp Varying with Distance », *Procedia computer science*, vol

123, p. 92-97. Récupéré de

https://www.sciencedirect.com/science/article/pii/S1877050918300164.

- BÖCKER Koen BE, VAN AVERMAETE Jurgen AG et VAN DEN BERG-LENSSEN Margaretha (1994). « The international 10–20 system revisited: Cartesian and spherical co-ordinates », *Brain Topography*, vol. 6, no 3. Récupéré de https://link.springer.com/article/10.1007/BF01187714.
- BRITTON Jeffrey W., FREY Lauren C., HOPP Jennifer L., KORB Pearce, KOUBEISSI Mohamad Z., LIEVENS William E., PESTANA-KNIGHT Elia M. et ST. LOUIS Erik K. (2016). « Appendix 4. Common Artifacts During EEG Recording », *Electroencephalography (EEG): An Introductory Text and Atlas of Normal and Abnormal Findings in Adults, Children, and Infants*, 1<sup>ière</sup> édition, Chicago, American Epilepsy Society, p. 84-94.
- 11. CARSON Audrey M., SALOWITZ Nicole M., SCHEIDT Robert A., DOLAN Bridget K. et VAN HECKE, Amy V. (2014). « Electroencephalogram coherence in children with and without autism spectrum disorders: decreased interhemispheric connectivity in autism ». *Autism Research*, vol 7, no 3. Récupéré de https://onlinelibrary.wiley.com/doi/abs/10.1002/aur.1367.
- CHEN Jean-Lon, ROS Tomas et GRUZELIER John H (2013). « Dynamic changes of ICA-derived EEG functional connectivity in the resting state », *Human brain mapping*, vol. 34, no 4. Récupéré de https://onlinelibrary.wiley.com/doi/abs/10.1002/hbm.21475
- 13. CHOONG W Y, KHAIRUNIZAM W, MUSTAFA W A, MURUGAPPAN M, HAMID A, BONG S Z, YUVARAJ R, OMAR M I, JUNOH A K, ALI HASIMAH, RAZLAN Z M et SHAHRIMAN A B (2021). « Correlation Analysis of Emotional EEG In Alpha, Beta and Gamma Frequency Bands », *Journal of Physics: Conference Series*, vol 1997, no 1. Récupéré de <u>https://iopscience.iop.org/article/10.1088/1742-6596/1997/1/012029/meta</u>.
- 14. DELIGIANNI Fani, CENTENO Maria, CARMICHAEL David W. and Clayden Jonathan D. (2014). « Relating resting-state fMRI and EEG whole brain connectomes across frequency bands », *Frontiers in Neuroscience*, vol 8, no 258. Récupéré de <u>https://www.frontiersin.org/articles/10.3389/fnins.2014.00258/full</u>.
- 15. DESERNO Markus (2004). « How to generate equidistributed points on the surface of a sphere », *If Polymerforshung (Ed.)*, vol. 99, no 2. Récupéré de <u>https://www.cmu.edu/biolphys/deserno/pdf/sphere\_equi.pdf</u>.

- 16. *EEG : Introduction* (s.d.). The McGill Physiology Virtual Lab. Récupéré le 10 juillet 2022 de <u>https://www.medicine.mcgill.ca/physio/vlab/biomed\_signals/eeg\_n.htm</u>.
- EKHTIARI Hamed, NASSERI Padideh, YAVARI Fatemeh, MOKRI Azarkhsh et MONTEROSSO John (2016). « Neuroscience of drug craving for addiction medicine: From circuits to therapies », *Progress in Brain Research*, vol. 223, p. 115-141.
- Foster, J. L., Shipstead, Z., Harrison, T. L., Hicks, K. L., Redick, T. S., & Engle, R. W. (2014). « Shortened complex span tasks can reliably measure working memory capacity », *Memory & Cognition*, vol 43, no 2. Récupéré de : <a href="https://link.springer.com/article/10.3758/s13421-014-0461-7">https://link.springer.com/article/10.3758/s13421-014-0461-7</a>.
- GRABNER Roland H. et DE SMEDT Bert (2012). « Oscillatory EEG Correlates of Arithmetic Atrategies: A Training Study », *Frontiers in Psychology*, vol 3, no 428. Récupéré de <u>https://doi.org/10.3389/fpsyg.2012.00428</u>.
- 20. GREENBLATT Robert E., PFLIEGER M. E. et OSSADTCHI A. E. (2012). « Connectivity measures applied to human brain electrophysiological data », *Journal of neuroscience methods*, vol. 207, no 1. Récupéré de <u>https://www.sciencedirect.com/science/article/abs/pii/S0165027012000817</u>
- 21. ISMAIL W Omar Ali Saifuddin Wan, HANIF M. et HAMZAH Noraini (2016). « Human Emotion Detection Via Brain Waves Study by Using Electroencephalogram (EEG) », *International Journal on Advanced Science, Engineering and Information Technology*, vol 6, no 6. Récupéré de

https://www.researchgate.net/publication/312422191\_Human\_Emotion\_Detection\_vi a\_Brain\_Waves\_Study\_by\_Using\_Electroencephalogram\_EEG.

- 22. JAVAID Hafiza Iqra, ROHEEN Arsa et AFZAL Farjad (2020). « Anatomy and Physiology of Brain in Context of Learning: A Review from Current Literature », *Biomedical Journal of Scientific & Technical Research*, vol 26, no 5. Récupéré de <u>https://biomedres.us/fulltexts/BJSTR.MS.ID.004415.php</u>.
- 23. JAWARBI Khalid H. et SHARMA Sandeep (2021). « Physiology, Cerebral Cortex Functions », *StatPearls [Internet]*, Récupéré de <u>https://www.ncbi.nlm.nih.gov/books/NBK538496/</u>.
- 24. KHOSLA Deepak, DON Manuel, et KWONG Betty (1999). « Spatial mislocalization of EEG electrodes–effects on accuracy of dipole estimation », *Clinical neurophysiology*, vol. 110, no 2. Récupéré de <u>https://www.sciencedirect.com/science/article/abs/pii/S0013469498001217</u>.

- 25. KINCAID Chuck (2005). « Guidelines for selecting the covariance structure in mixed model analysis », *Proceedings of the thirtieth annual SAS users group international conference*, vol 30 no 198. Récupéré de https://support.sas.com/resources/papers/proceedings/proceedings/sugi30/198-30.pdf.
- 26. Kirchner, W. K. (1958), « Age differences in short-term retention of rapidly changing information », *Journal of Experimental Psychology*, vol 55 no 4. Récupéré de <u>https://psycnet.apa.org/doiLanding?doi=10.1037%2Fh0043688</u>.
- 27. LAI Chi Qin, IBRAHIM Haidi, ABDULLAH Mohd Zaid, SUANDI Shahrel Azmin et AZMAN Azlinda (2018). « Literature survey on applications of electroencephalography (EEG) », AIP Conference Proceedings, vol 2016, no 1. Recupéré de <u>https://doi.org/10.1063/1.5055472</u>.
- 28. LAU Troy M., GWIN Joseph T. et FERRIS Daniel P. (2012). « How Many Electrodes Are Really Needed for EEG-Based Mobile Brain Imaging? », *Journal of Behavioral and Brain Science*, vol 2, no 3. Récupéré de <u>https://www.scirp.org/journal/PaperInformation.aspx?PaperID=22107</u>.
- 29. LORIST Monicque M., BEZDAN Eniko, TEN CAAT Michael, SPAN Mark M., ROERDINK Jos B.T.M. et MAURITSBD Natasha M. (2009). « The influence of mental fatigue and motivation on neural network dynamics; an EEG coherence study », *Brain research*, vol. 1270, p. 95-106. Récupéré de <u>https://www.sciencedirect.com/science/article/abs/pii/S0006899309005162</u>.
- 30. ISLAM M. R., ISLAM, M. M., RAHMAN, M. M., MONDAL, C., SINGHA, S. K., AHMAD, M., AWAL A., ISLAM S. et MONI, M. A. (2021). « EEG channel correlation based model for emotion recognition », *Computers in Biology and Medicine*, vol 136, no 104757. Récupéré de https://www.sciencedirect.com/science/article/abs/pii/S0010482521005515 .
- 31. McGill, (s.d). Les différents lobes du cerveau [Illustration]. <u>https://thebrain.mcgill.ca/flash/d/d\_06/d\_06\_cr/d\_06\_cr\_mou/d\_06\_cr\_mou.html</u>.
- 32. McGill, (s.d). Les différentes ondes du cerveau [Illustration]. https://thebrain.mcgill.ca/flash/a/a\_11/a\_11\_p/a\_11\_p\_cyc/a\_11\_p\_cyc.html.
- NETTLETON, David (2014). « Selection of Variables and Factor Derivation», *Commercial data mining*, 1<sup>ière</sup> édition, Waltham, Elsevier, p. 79-104.
- 34. OLEJNICZAK Piotr (2006). « Neurophysiologic basis of EEG », Journal of clinical neurophysiology, vol 23, no 3. Récupéré de <u>http://norwalk.braincoretherapy.com/wpcontent/uploads/2014/01/Neurophysiologic-Basis-of-EEG.pdf</u>.

35. OMIDVARNIA Amir, MESBAH Mostefa, O'TOOLE John M., COLDITZ Paul et BOASHASH Boualem (2011). « Analysis of the time-varying cortical neural connectivity in the newborn EEG: A time-frequency approach », 7th International Workshop on Systems, Signal Processing and their Applications WOSSPA. Récupéré de

https://ieeexplore.ieee.org/document/5931445/citations?tabFilter=papers#citations.

- 36. PENNY W. et HARRISON L. (2007), « Multivariate autoregressive models », Statistical Parametric Mapping: The Analysis of Functional Brain Images, 1<sup>ière</sup> édition, Amsterdam, Academic Press, p. 543-540.
- 37. Redick, T. S., Broadway, J. M., Meier, M. E., Kuriakose, P. S., Unsworth, N., Kane, M. J., & Engle, R. W. (2012). « Measuring Working Memory Capacity With Automated Complex Span Tasks », *European Journal of Psychological Assessment*, vol 28, no 3, Récupéré de <u>https://psycnet.apa.org/record/2012-15435-002</u>.
- RUSSEL Peter (2013). « Preface », *The Brain Book*, 15<sup>ième</sup> édition, New York, Routledge & Kegan Paul, p. vii-px.
- 39. SHIPTON Harold W. (1991). « Dipole Theory », American Journal of EEG Technology, vol 31, no 2. Récupéré de <u>https://www.tandfonline.com/doi/abs/10.1080/00029238.1991.11080364?journalCode</u> <u>=utnj18</u>.
- 40. The CORR Procedure (2020). SAS Help Center. Récupéré le 27 juillet 2022 de https://documentation.sas.com/doc/en/pgmsascdc/9.4\_3.5/procstat/procstat\_corr\_detai ls20.htm#:~:text=PROC%20CORR%20excludes%20observations%20with,variables %20in%20the%20statistical%20computations.
- 41. The MIXED Procedure (2019). SAS Help Center. Récupéré le 25 juillet 2022 de <u>https://documentation.sas.com/doc/en/pgmsascdc/9.4\_3.4/statug/statug\_mixed\_syntax</u> <u>14.htm#statug.mixed.repeatedstmt\_type</u>.
- 42. THORNE Glenda and THOMAS Alice (2021). What Is Attention?, The Center for Literacy and Learning. Récupéré le 10 juillet 2022 de <u>https://mycll.org/what-is-</u> <u>attention/#:~:text=Passive%20attention%20refers%20to%20the,or%20a%20sudden%</u> <u>20loud%20noise</u>.
- 43. VIDAL Jacques J. (1977). « Real-Time Detection of Brain Events in EEG », Proceedings of IEEE, vol 65, no 5. Récupéré de https://ieeexplore.ieee.org/abstract/document/1454811.

- 44. WELCH, Peter D. (1967). « The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms », *IEEE Transactions on Audio and Electroacoustics*, vol 15(2), p. 70-73. Récupéré de <u>https://ieeexplore.ieee.org/document/1161901</u>.
- 45. XIA XiaoLei et HU Li (2019). « EEG: neural basis and measurement », *EEG Signal Processing and Feature Extraction*, 1<sup>ière</sup> édition, New York, Springer Singapore, p. 7-21.
- 46. YAVARTANOO Mohsen, KIM Eu Young et LEE Kyoung Mu (2018). « Spnet: Deep 3d object classification and retrieval using stereographic projection », Asian conference on computer vision. Springer, Cham, p. 691-706. Récupéré de <u>https://link.springer.com/chapter/10.1007/978-3-030-20873-8\_44</u>.
- 47. ZHANG Zhiguo (2019). « Spectral and Time-Frequency Analysis », *EEG Signal Processing and Feature Extraction*, 1<sup>ère</sup> édition, New York, Springer Singapore, p. 89-116.