

ΙΟΝΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ
ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ

Ανάπτυξη αλγορίθμων για την
επεξεργασία
ηλεκτροεγκεφαλογραφικών
σημάτων.



Στυλιανός Γαβράς

Επιβλέπων: Αυλωνίτης Μάρκος, Αναπληρωτής Καθηγητής, Ιόνιο

Πανεπιστήμιο

Συν-Επιβλέπων: Έξαρχος Θεμιστοκλής, Επίκουρος Καθηγητής, Ιόνιο

Πανεπιστήμιο

Συν-Επιβλέπων: Βλάμος Παναγιώτης, Αναπληρωτής Καθηγητής

Η παρούσα πτυχιακή εργασία εκπονήθηκε στα πλαίσια του προπτυχιακού προγράμματος του Τμήματος Πληροφορικής του Ιονίου Πανεπιστημίου.

24 Σεπτεμβρίου 2020

Declaration

Αυτή η πτυχιακή έχει άδεια χρήσης Creative Commons CC0 1.0 Universal (CC0 1.0) Public Domain Dedication.



Ευχαριστίες

Σε αυτό το σημείο θα ήθελα αρχικά να ευχαριστήσω τους γονείς που με στηρίζουν σε κάθε μου βήμα και απόφαση σε όλη την πορεία μου.

Ένα μεγάλο ευχαριστώ στους : καθηγητή Μάρκο Αυλωνίτη με τον οποίο και μαζί με τους υποψ. διδάκτορες του έκανα τα πρώτα μου επιστημονικά βήματα στο εργαστήριο του και με εμπιστεύτηκε από νωρίς, καθηγητή Θέμη Έξαρχο για όλη την αμέριστη εμπιστοσύνη που μου έδειξε σε ακαδημαϊκό και προσωπικό επίπεδο και τέλος τον καθηγητή Παναγιώτη Βλάμο ως συνεπιβλέπων αυτής της εργασίας.

Ακόμη, ευχαριστώ πολύ τον Δημήτρη Κοτινά για όλη την καθοδήγηση και το ενδιαφέρον το οποίο έδειξε όλα αυτά τα χρόνια στην Κέρκυρα για εμένα και τις ατελειώτες ώρες που περάσαμε παρέα με το Mathematica!

Φυσικά δεν θα μπορούσα να παραλείψω όλους τους φίλους μου στην Κέρκυρα για τις ωραίες στιγμές που περάσαμε μαζί όλα αυτά τα χρόνια στο νησί και ιδιαίτερα τον Σπύρο, για όλη την υποστήριξη του τεχνικά όσον αφορά αυτή την εργασία αλλά και σε προσωπικό επίπεδο.

Τέλος, θα ήθελα να ευχαριστήσω τους Quants, από το τμήμα των αλγορίθμων της Stoiximan : Βασίλη, Δημήτρη, Γιάννη και Λευτέρη για την καθοδήγηση τους τον τελευταίο χρόνο στη δουλειά και την υποστήριξή τους ώστε να γίνομαι κάθε μέρα καλύτερος επαγγελματίας.

Αυτή η εργασία είναι αφιερωμένη στους γονείς μου, Σπύρο και Γεωργία.

Περίληψη

Η παρούσα πτυχιακή εργασία πρόκειται για μια ερευνητική μελέτη αλγορίθμων και υπολογιστικών τεχνικών με τη βοήθεια των οποίων δύναται να αντληθεί χρήσιμη πληροφορία σε πολύπλοκες χρονοσειρές όπως αυτές των ηλεκτροεγκεφαλογραφημάτων. Βιβλιογραφικά, έχουν προταθεί αρκετοί αλγόριθμοι οι οποίοι ποσοτικοποιούν με στατιστικό τρόπο τη μνήμη ενός σήματος παρέχοντας μετρικές και χαρακτηριστικά που δύναται να χρησιμεύσουν σε ποικίλες εφαρμογές. Αντικείμενο της παρούσας εργασίας είναι να αναπτυχθούν και να αξιολογηθούν με προσομοιώσεις και συνθετικά δεδομένα τρεις αλγόριθμοι που έχουν προταθεί στη βιβλιογραφία και έπειτα ο σχεδιασμός και υλοποίηση ενός νέου προτεινόμενου αλγορίθμου. Για την ανάλυση σε πραγματικά δεδομένα γίνεται χρήση βιοσημάτων (EEG) από ασθενείς με PTSD και υγιείς (control group). Έτσι, έπειτα από τον υπολογισμό του εκθέτη Hurst - μέτρο που ποσοτικοποιεί την ύπαρξη μνήμης σε μια χρονοσειρά, τα αποτελέσματα δείχνουν ότι το κανάλι F3 είναι αρκετά χρήσιμο για διαγνωστικές ιατρικές εφαρμογές μεταξύ των δύο ομάδων.

Περιεχόμενα

1	Κίνητρο	1
2	Χρονοσειρές, Στοχαστικά και Ντετερμινιστικά Συστήματα	2
2.1	Μελέτη Χρονοσειρών	2
2.1.1	Στασιμότητα	3
2.2	Δυναμικά Συστήματα	4
2.2.1	Λογιστική απεικόνιση	4
2.2.2	Mackey-Glass equation	5
2.2.3	Chirikov standard map	6
2.3	Στοχαστικές Διαδικασίες	6
2.3.1	Τυχαίος περίπατος	7
2.3.2	Λευκός θορυβος	8
2.3.3	Διαδικασία Wiener	9
3	Χαρακτηριστικά Χρονοσειρών	10
3.1	Περιγραφικά Στατιστικά	10
3.2	Αυτοδιακύμανση και Αυτοσυσχέτιση	10
3.3	Ανάλυση συχνοτήτων	11
3.4	Turning Points	11
3.5	Ratio of the mean square successive difference to the variance	12
3.6	Hjorth Parameters	12
3.7	Συσχετίσεις μακράς διάρκειας	13
3.7.1	Ιστορικό πλαίσιο	13
3.7.2	Fractals	14
3.7.3	Fractional Brownian Motion	15
3.7.4	Fractal Gaussian Noise	15
3.7.5	Σύγχρονη εφαρμογή στις βιολογικές διαδικασίες	16
4	Αλγόριθμοι υπολογισμού του εκθέτη H	18
4.1	Αλγόριθμος της αυξανόμενης διασπορας	18
4.2	Αλγόριθμος ανάλυσης αλλαγής κλίμακας του εύρους	18
4.3	Αλγόριθμος βασισμένος στον Haar Wavelet	19
4.4	Εφαρμογή αλγορίθμων σε προσομοιωτικά δεδομένα	20

4.5	Προτεινόμενος Αλγόριθμος	25
5	Ηλεκτροεγκεφαλογραφία	29
5.1	Φυσιολογία	29
5.2	Βιβλιογραφική ανασκόπηση για την εφαρμογή του Η σε ΗΕΓ δεδομένα .	33
6	Εφαρμογή σε πραγματικά δεδομένα	34
7	Συμπεράσματα	38

List of Figures

2.1	Η χρονοσειρά της τιμής του κρυπτονομίσματος Bitcoin για τη χρονιά αυτή.	3
2.2	Η λογιστική απεικόνιση και η ευαισθησία αρχικών συνθηκών για διαφορετική παράμετρο x_0 .	5
2.3	Η διαφορική εξίσωση με υστέρηση Mackey Glass.	5
2.4	Απεικόνιση Chirikov	6
2.5	Μονοπάτια τυχαίου περίπατου σε μία διάσταση.	7
2.6	Γκαουσιανός περήπατος και η το σχήμα της κανονικής κατανομής που περιγράφει μοντελοποιεί τα βήματα.	8
2.7	Λευκός θόρυβος (αριστερά) και ροζ θόρυβος (δεξιά).	8
2.8	Τρία μονοπάτια διαδικασίας Wiener.	9
3.1	Η συνεχής αλλά πουθενά παραγωγίσιμη σύναρτηση Weistrass και τα μορφοκλασματικά χαρακτηριστικά της.	14
3.2	Τυχαίες διαδικασίες FBM με διαφορετικούς εκθέτες Hurst σε ακραίες τιμές.	15
3.3	Μορφοκλασματική διαδικασία Gauss (FGN) για διαφορετικές παραμέτρους εκθέτη Hurst.(0.25, 0.5, 0.75, από κάτω προς τα πάνω)	16
4.1	Ιστογράμματα από προσομοιώσεις διαδικασίες FBM για διαφορετικούς εκθέτης Η όσον αφορά την μέθοδο την αυξανόμενης διασποράς.	21
4.2	Διάγραμμα θηκογραμμάτων για μια ποικιλία εκθετών Η όσον αφορά την μέθοδο Ανάλυσης Κλίμακας Εύρους.	22
4.3	Αποτελέσματα αλγορίθμου στα δεδομένα προσομοίωσης για διαδικασίες FBM και τρεις διαφορετικούς εκθέτες.	23
4.4	Αποτελέσματα αλγορίθμου στα δεδομένα προσομοίωσης για διαδικασίες FGN και τρεις διαφορετικούς εκθέτες.	24
4.5	Αριστερά η αξιολόγηση του αλγορίθμου Haar Wavelet σε διαδικασίες DFGN. Δεξιά η ίδια αξιολόγηση αλλά μετά από τον μετασχηματισμό των χρονοσειρών για την μετατροπή τους σε FGN.	24
4.6	Η μαθηματικές σχέσεις ανάμεσα στην τιμή Abbe και Η που προκύπτουν από προσομοιώσεις και στατιστική μοντελοποίηση [58].	25
4.7	Ο Confusion Matrix της ταξινόμησης.	26

4.8	Πάνω: Δύο τυχαίες διαδικασίες FBM και FGN με εκθέτη $H = 0.05$. Κάτω: Ραβδόγραμμα αλγορίθμων με βάση τον προτεινόμενο αλγόριθμο για τις περιπτώσεις μικρού εκθέτη σε διαδικασίες FBM.	27
4.9	Διαφορά κατανομής των δύο αλγορίθμων σε προσομοιωτικά δεδομένα DFGN με $Hurst = 0.5$	27
4.10	Το A/T πλανε για διαδικασίες μήκους 1024 στοιχείων που προτάθηκε για πρώτη φορά από τον M. Ταρνοπολσκι στο [58].	28
5.1	Δομή ενός πραγματικού νευρώνα. [18]	29
5.2	Περιοχές εγκεφάλου και αντίστοιχες λειτουργίες τους.[1]	30
5.3	Το διεθνές σύστημα τοποθέτησης 10-20	31
5.4	Εγκεφαλικοί ρυθμοί. [45]	31
6.1	Κατανομές σημάτων για control group και PTSD οσον αφορά το κανάλι F3.	35
6.2	Τυπική απόκλιση σημάτων για κάθε υπολογιστικό παράθυρο.	35
6.3	Κατανομές εκθέτη Hurst για τα δύο group.	36

List of Tables

4.1	Πίνακας βασικών χαρακτηριστικών όσον αφορά την αξιολόγηση του αλ- γορίθμου Wavelet σε διάφορα είδη χρονοσειρών και εκθετών.	22
-----	--	----

Κεφάλαιο 1

Κίνητρο

Κίνητρο για αυτήν την εργασία αποτελεί σε μεγάλο βαθμό το ενδιαφέρον μου όσον αφορά τις επιστημονικές περιοχές των εφαρμοσμένων μαθηματικών, της στατιστικής επεξεργασίας σήματος και την ανάπτυξη αλγορίθμων καθώς και την ερευνητική μου περιέργεια για την ηλεκτροεγκεφαλογραφία και την επεξεργασία αντίστοιχων δεδομένων.

Στο κεφάλαιο 2 γίνεται μία εισαγωγή στις χρονοσειρές και τα συστήματα που τις παράγουν - στοχαστικά και ντετερμινιστικά. Στη συνέχεια, στο κεφάλαιο 3 αναφέρονται και αναλύονται μέθοδοι εξαγωγής χαρακτηριστικών σε σήματα ενώ γίνεται μια εκτενής αναφορά στις συσχετίσεις μαρκάς διάρκειας. Στο επόμενο κεφάλαιο περιγράφεται η μεθοδολογία αλγορίθμων υπολογισμού του εκθέτη Hurst, αξιολογούνται μέσω προσομοιώσεων και συνθετικών δεδομένων, γνωστοί στη βιβλιογραφία αλγόριθμοι και τέλος προτείνεται ένας νέος αλγόριθμος βασισμένος σε ταξινόμηση. Παράλληλα, στο κεφάλαιο 5 με τίτλο "Ηλεκτροεγκεφαλογραφία" παρουσιάζονται βασικές πληροφορίες για την φυσιολογία του ανθρώπου, την ηλεκτροεγκεφαλογραφία ενώ γίνεται μια πρόσφατη ανασκόπηση της βιβλιογραφίας για εφαρμογές αλγορίθμων μαρκάς διάρκειας σε βιοσήματα. Τέλος στο κεφάλαιο 6 αναπτύσσεται μια μεθοδολογία και γίνεται εφαρμογή αλγορίθμου σε πραγματικά δεδομένα EEG από ασθενείς με διάγνωση μετατραυματικού στρες και control group.

Για την έρευνα, την ανάπτυξη των αλγορίθμων και την ανάλυση των δεδομένων έγινε χρήση της γλώσσας προγραμματισμού Mathematica.

Κεφάλαιο 2

Χρονοσειρές, Στοχαστικά και Ντετερμινιστικά Συστήματα

2.1 Μελέτη Χρονοσειρών

Μια χρονοσειρά $\{x_n\}_{n=1}^N$ είναι μια συλλογή από παρατηρήσεις ή οποία δημιουργήθηκε διαδοχικά στο χρόνο [11]. Μπορεί κανείς να βρει αρκετά τέτοια παραδείγματα σημάτων σε πεδία όπως τη μηχανική, τα οικονομικά, τη βιοϊατρική και άλλα.

Ο μηχανισμός που παράγει τέτοιες χρονοσειρές είναι πρακτικά ένα δυναμικό σύστημα, δηλαδή ένα σύστημα μεταβλητών που αλλάζουν στο χρόνο και αρκετές φορές δύναται να αλληλεπιδρούν μεταξύ τους. Άλλες φορές σε τέτοια συστήματα εμπεριέχεται η τυχαιότητα και οι τιμές των παρατηρήσεων προέρχονται με μία πιθανότητα και έτσι αυτό που μετράμε είναι μια τυχαία μεταβλητή. Ένα τέτοιο σύστημα ονομάζεται στοχαστικό ενώ όταν δεν εμπεριέχεται κάποιου είδους τυχαιότητα και είναι πλήρως κατανοητή η διαδικασία, δηλαδή οι εξισώσεις του συστήματος είναι γνωστές, το σύστημα αυτό ονομάζεται ντετερμινιστικό [3]. Παραδείγματα διαφοροποίησης των δύο συστημάτων θα δοθούν στα παρακάτω υποκεφάλαια.

Συχνό φαινόμενο στην επεξεργασία σημάτων είναι η ύπαρξη θορύβου. Ο τελευταίος αν και “εχθρός” στην ανάλυση των δεδομένων, στις περιπτώσεις που είναι γνωστός μπορεί να αποτελέσει ένας πολύ έμπιστος “σύμμαχος” παρέχοντας μας κρίσιμες πληροφορίες για το σύστημα και τις μετρήσεις μας. Τα είδη των θορύβων χωρίζονται σε δύο κατηγορίες, τον δυναμικό θόρυβο ή αλλιώς θόρυβο συστήματος ο οποίος κυρίως προέρχεται από εξωγενείς παράγοντες και τον θόρυβο παρατήρησης που οφείλεται στο σφάλμα μέτρησης (π.χ θόρυβος ενός αισθητήρα θερμοκρασίας). [3]

Στην παρακάτω εικόνα βλέπουμε ένα αρκετά γνωστό παράδειγμα χρονοσειράς από τα οικονομικά. Πρόκειται για τιμή του γνωστού κρυπτονομίσματος Bitcoin. Η συγκεκριμένη επιλογή για παράδειγμα χρονοσειράς δεν είναι τυχαία, καθώς οι αγορές των κρυπτονομισμάτων έχουν τραβήξει αρκετά βλέμματα τα τελευταία χρόνια.

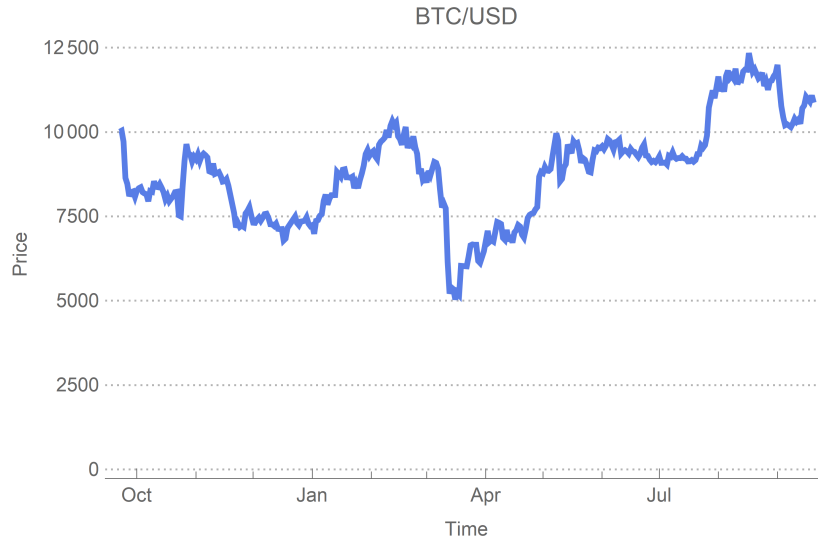


Figure 2.1: Η χρονοσειρά της τιμής του κρυπτονομίσματος Bitcoin για τη χρονιά αυτή.

2.1.1 Στασιμότητα

Μια χρονοσειρά ονομάζεται αυστηρώς στάσιμη όταν τα πιθανοθεωρητικά της χαρακτηριστικά παραμένουν αναλλοίωτα στο χρόνο [3]. Ωστόσο, ο συγκεκριμένος ορισμός είναι υπερβολικά αυστηρός για να εφαρμοστεί σε πραγματικά δεδομένα συχνά γίνεται χρήση της ασθενούς στασιμότητας. Έτσι, για να μπορέσει να θεωρηθεί μια χρονοσειρά ασθενώς στάσιμη θα πρέπει τα στατιστικά μέτρα - μέσος όρος και διακύμανση - να μην μεταβάλλονται χρονικά.

Αναφέροντας στον όρο στασιμότητα ή στα αγγλικά stationarity αξίζει να υπογραμμιστεί ότι ο συγκεκριμένος έλεγχος είναι ένα από τα σημαντικότερα βήματα πριν προβεί κανείς σε μια ανάλυση χρονοσειρών/σημάτων. Αρκετοί αλγόριθμοι/υπολογιστικά εργαλεία, όπως θα διαπιστώσουμε παρακάτω, κάνουν υπόθεση στασιμότητας. Αυτό πρακτικά σημαίνει ότι αν μια χρονοσειρά δεν είναι στάσιμη και υπολογιστικά της συμπεριφερθούμε σαν να είναι το αποτέλεσμα θα είναι λανθασμένο.

Μια πολύ γνωστή [3] και αρκετές φορές αποδοτική μεθοδολογία για να μετατρέψει κανείς μια διαδικασία από μη στάσιμη σε στάσιμη είναι υπολογίζοντας τις διαφορές της αρχικής χρονοσειράς Y_t . Ο συγκεκριμένος μετασχηματισμός απαλείφει την τάση στη χρονοσειρά και ορίζεται ως:

$$x_t = y_t - y_{t-1} \quad (2.1)$$

Μια παραλλαγή του παραπάνω μετασχηματισμού είναι ο μετασχηματισμός της μεταβολής του λογαρίθμου ο οποίος απαλείφει την τάση και ταυτόχρονα ελαττώνει τις μεγάλες διακυμάνσεις. Ορίζεται ως:

$$x_t = \ln y_t - \ln y_{t-1} \quad (2.2)$$

Τέλος υπάρχει και ο μετασχηματισμός της σχετικής μεταβολής που έχει παρόμοια χαρακτηριστικά με τον παραπάνω και ορίζεται ως:

$$x_t = \frac{y_t - y_{t-1}}{y_t} \quad (2.3)$$

Ένας αντικειμενικός τρόπος για να αποφασιστεί αν μια χρονοσειρά χρειάζεται μετασχηματισμό διαφορών είναι αυτός του στατιστικού ελέγχου μοναδιαίας ρίζας (unit root) [48]. Σε γενικότερα πλαίσια, υπάρχουν διαθέσιμα αρκετά τέτοια στατιστικά test στασιμότητας τα οποία είναι σχεδιασμένα για να μας παρέχουν μία αντικειμενική στατιστική προσέγγιση για το αν τελικά χρειάζεται μετασχηματισμό ή όχι, η χρονοσειρά.

2.2 Δυναμικά Συστήματα

Στο προηγούμενο κεφάλαιο έγινε αναφορά στις χρονοσειρές και στα χαρακτηριστικά που δύναται να εμφανίσουν. Σε αυτή την ενότητα θα γίνει λόγος για ορισμένα γνωστά παραδείγματα ντετερμινιστικών δυναμικών συστημάτων που μπορεί να συναντήσει κάποιος στη βιβλιογραφία.

2.2.1 Λογιστική απεικόνιση

Η λογιστική απεικόνιση (logistic map) έγινε διάσημη από τον Βιολόγο Robert May με την δημοσίευσή του το 1976 στο περιοδικό Nature [43] και είναι ένα κλασσικό παράδειγμα όσον αφορά την πολυπλοκότητα που δύναται να πηγάζει από ένα απλό μη γραμμικό δυναμικό σύστημα. Η συγκεκριμένη εξίσωση βασίζεται στη λογιστική εξίσωση η οποία δημιουργήθηκε από τον Βέλγο μαθηματικό Pierre Francois Verhulst και χρησιμοποιείται για μοντελοποίηση πληθυσμών. Ο τύπος της λογιστικής απεικόνισης ορίζεται ως

$$x_{n+1} = rx_n(1 - x_n) \quad \mu\epsilon \quad n = 0, 1, 2, 3, \dots \quad (2.4)$$

Δίνοντας έτσι μια αρχική τιμή $0 \leq x_0 \leq 1$ και έναν θετικό αριθμό $0 < r < 4$ η απεικόνιση παράγει τιμές. Στο παρακάτω διάγραμμα φαίνεται η ευεσθησία στις αρχικές συνθήκες και πιο συγκεκριμένα για διαφορετικές παραμέτρους αρχικής τιμής $x_0 = 0.3$ και $x_0 = 0.3001$ πως αλλάζει το αποτέλεσμα.

Τέλος, η συγκεκριμένη λογιστική απεικόνιση για συγκεκριμένες τιμές της παραμέτρου r δύναται να χρησιμοποιηθεί για την παραγωγή ψευδοτυχαίων αριθμών [53, 61].

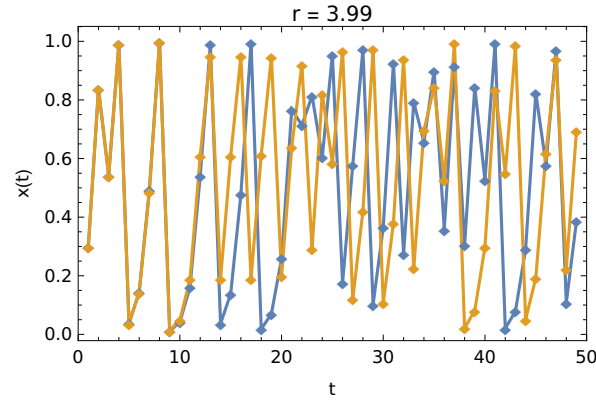


Figure 2.2: Η λογιστική απεικόνιση και η ευαισθησία αρχικών συνθηκών για διαφορετική παράμετρο $x\theta$.

2.2.2 Mackey-Glass equation

Το δυναμικό σύστημα της διαφορικής εξίσωσης με υστέρηση Mackey και Glass [36] ένα αρκετά διαδεδομένα δυναμικό σύστημα το οποίο για διαφορετικές καταστάσεις ανάλογα με την τιμή της παράμετρου Δ μπορεί να παράγει χάος. Το συγκεκριμένο σύστημα έχει οριστεί με σκοπό να περιγράψει συστήματα με την φυσιολογία του ανθρώπου και η εξίσωση που το περιγράφει είναι η εξής:

$$\frac{dx}{dt} = \frac{\beta_0 \theta^n x(t - \Delta)}{\theta^n + x(t - \Delta)^n} - \gamma x(t) \quad (2.5)$$

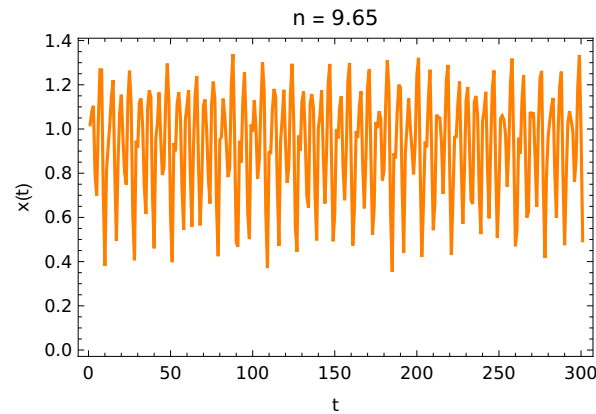


Figure 2.3: Η διαφορική εξίσωση με υστέρηση Mackey Glass.

Η πολυπλοκότητα του συγκεκριμένου δυναμικού συστήματος έχει άμεση σχέση με την παράμετρο Δ η οποία για διαφορετικούς παραμέτρους εμφανίζει διαφορετική μορφολογιαστική διάσταση [21].

2.2.3 Chirikov standard map

Η απεικόνιση του Chirikov ή αλλιώς standard map [12] δημιουργήθηκε από τον Boris Chirikov το 1969. Πρόκειται για ένα δυναμικό σύστημα, που αποτελείται από 2 επαναληπτικές σχέσεις στις οποίες η μόνη ορισμένη παράμετρος είναι η K και είναι της μορφής

$$\begin{cases} p_{n+1} = p_n + \frac{K}{2\pi} \sin(2\pi x_n) \\ x_{n+1} = x_n + p_{n+1} \end{cases} \quad (2.6)$$

Η παράμετρος K πρόκειται για μια μη γραμμική παράμετρο η οποία σε ορισμένες τιμές, όπως $K > K_c \simeq 0.97163540631$ [23, 35]. Το εντυπωσιακό είναι ότι το συγκεκριμένο σύστημα, εμφανίζει αρκετές φορές μια συμπεριφορά η οποία θα μπορούσε να χαρακτηριστεί "κολλημένη" για ένα μεγάλο αριθμό βημάτων και τα χαοτικά χαρακτηριστικά γίνονται ορατά.

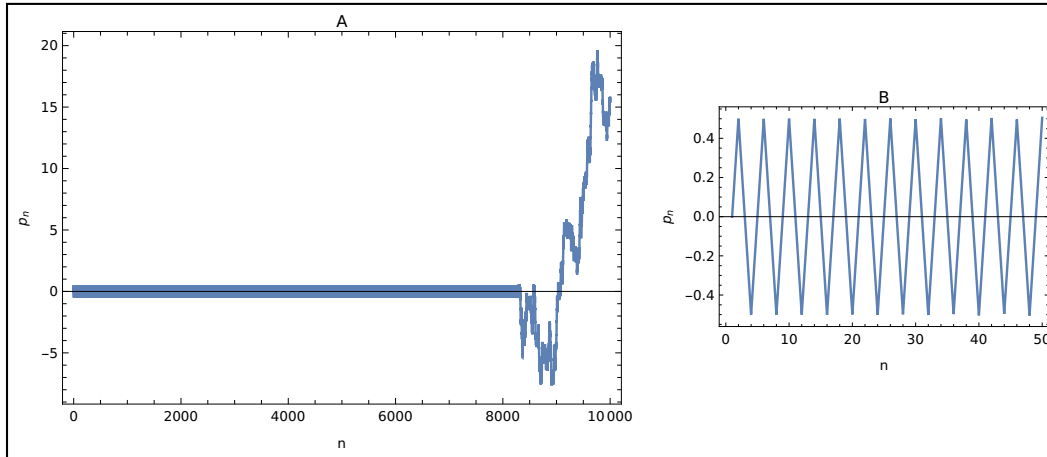


Figure 2.4: Απεικόνιση Chirikov

2.3 Στοχαστικές Διαδικασίες

Στο συγκεκριμένο κεφάλαιο αναφέρονται κάποιες από τις σημαντικότερες στοχαστικές διαδικασίες καθώς ένα υποκεφάλαιο όσον αφορά τις συσχετίσεις μακράς εμβέλειας και την συνάρτηση αυτοσυσχέτισης.

2.3.1 Τυχαίος περίπατος

Έστω κινητό σώμα σε στον άξονα της πραγματικής ευθείας $\mathbb{R}$ το οποίο με πιθανότητα p πηγαίνει ένα βήμα δεξιά και με πιθανότητα $q = 1 - p$ πηγαίνει ένα βήμα αριστερά [2].

Το άθροισμα αυτής της ακολουθίας $\{Z_t\}_{t=1, \dots, \infty}$ μέχρι τη χρονική στιγμή t αποτελεί μια στοχαστική διαδικασία -ονομάζεται τυχαίος περίπατος και μπορεί οριστεί ως :

$$X_t = \sum_{i=1}^t Z_i \quad (2.7)$$

Η παραπάνω στοχαστική διαδικασία είναι μη-στάσιμη, με μέση τιμή $E[X_t] = 0$ και διασπορά $\sigma_X^2 = t\sigma_Z^2$ [51].

Στο παρακάτω διάγραμμα φαίνεται η εξέλιξη 100 τυχαίων περιπάτων καθώς και η κατανομή του βήματος στο οποίο θα καταλήξουν. Η πιθανότητα p και q είναι αμφότερες στο 0.5. Παράλληλα, υπάρχουν αρκετές παραλλαγές του τυχαίου περιπατητή και μια από τις σημαντικότερες είναι η γενίκευση σε περισσότερες διαστάσεις.

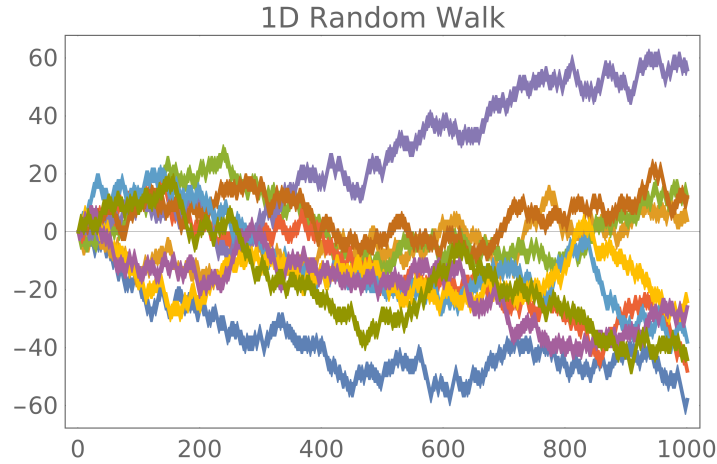


Figure 2.5: Μονοπάτια τυχαίου περιπάτου σε μία διάσταση.

Ακόμη μια ενδιαφέρουσα παραλλαγή είναι αυτή όπου το βήμα μετάβασης του τυχαίου περιπατητή υπακούει σε μια γνωστή κατανομή όπως αυτή του Gauss. Έτσι στο παρακάτω διάγραμμα βλέπουμε συγκριτικά δύο τυχαίους περιπατητές, έναν κλασσικό και έναν με γκαουσιανό βήμα $N(0, 1)$

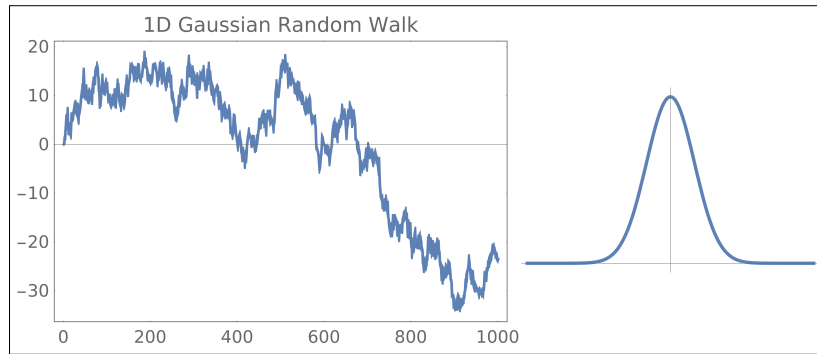


Figure 2.6: Γκαουσιανός περήπατος και η το σχήμα της κανονικής κατανομής που περιγράφει μοντελοποιεί τα βήματα.

2.3.2 Λευκός θόρυβος

Μια άλλη γνωστή βιβλιογραφικά στοχαστική διαδικασία είναι αυτή του λευκού θορύβου. Λευκός θόρυβος είναι ο θόρυβος στον οποίο η συχνότητα και το φάσμα ισχύος είναι σταθερά και ανεξάρτητα από τη συχνότητα [38].

Αν τα στοιχεία του θορύβου ακολουθούν την κανονική κατανομή τότε ονομάζεται Γκαουσιανός λευκός θόρυβος. Στο συγκεκριμένο σημείο αξίζει να αναφερθεί ότι υπάρχουν αρκετά είδη θορύβων τα οποία μπορεί κανείς να συναντήσει στην επεξεργασία σήματος. Συνιθίζεται στην βιβλιογραφία να αναφέρονται με διαφορετικά ονόματα χρωμάτων ανάλογα με τα χαρακτηριστικά τους.

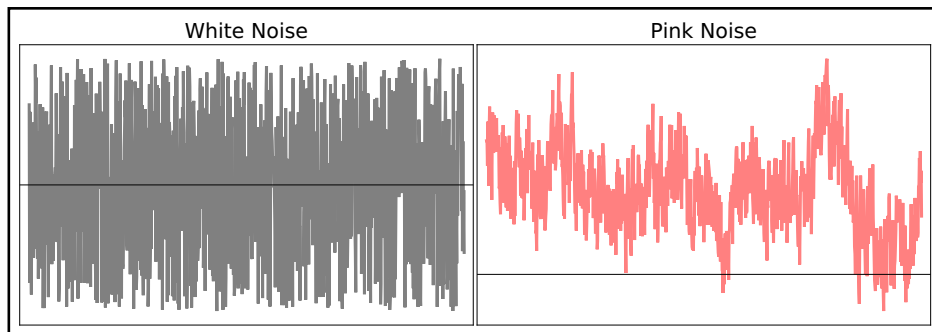


Figure 2.7: Λευκός θόρυβος (αριστερά) και ροζ θόρυβος (δεξιά).

2.3.3 Διαδικασία Wiener

Αρκετά φαινόμενα μπορούν να μοντελοποιηθούν με τη διαδικασία Wiener η οποία έχει πάρει το όνομα της από τον μαθηματικό Norbert Wiener. Ο τελευταίος μαζί με τον Albert Einstein θεμελίωσαν την κίνηση Brown την οποία είχε παρατηρήσει σε φυσικά πειράματά του ο Robert Brown. Ο ίδιος ο Einstein ερμήνευσε αυτή την κίνηση στον τομέα της στατιστικής μηχανικής σαν αποτέλεσμα συνεχών συγκρούσεων του σωματιδίου με τα μόρια ενός ρευστού.

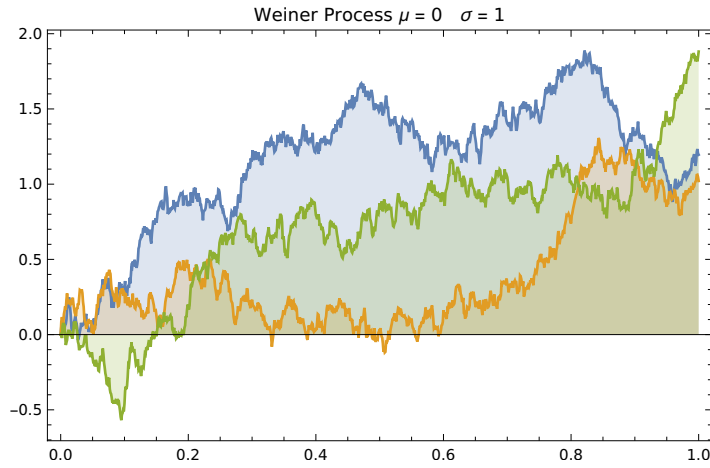


Figure 2.8: Τρία μονοπάτια διαδικασίας Wiener.

Μια κανονική διαδικασία Wiener, στο διάστημα $[0, T]$ είναι μια τυχαία μεταβλητή $W(t)$ όπου $t \in [0, T]$ και ισχύουν ότι : $W(0) = 0$ με πιθανότητα 1, για κάθε $0 \leq s < t \leq T$ τυχαία μεταβλητή που δίνεται από τη μεταβολή $W(t) - W(s)$ ακολουθεί κανονική κατανομή με μέσο όρο μηδέν και διακύμανση $t-s$, για $0 \leq s < t < u < v \leq T$ η διαφορές $W(t) - W(s)$ και $W(u) - W(s)$ είναι ανεξάρτητες.

Κεφάλαιο 3

Χαρακτηριστικά Χρονοσειρών

Υπάρχει ένα αρκετά μεγάλο πλήθος χαρακτηριστικών που μπορεί κανείς να επιλέξει για να περιγράψει μια χρονοσειρά. Αυτά μπορεί να είναι από πολύ απλά έως και αρκετά πολύπλοκοι αλγόριθμοι. Στο συγκεκριμένο κεφάλαιο θα γίνει αναφορά σε μερικά χαρακτηριστικά και μέτρα που αφορούν επεξεργασία σήματος και θα δοθεί έμφαση σε χαρακτηριστικά που η χρήση τους είναι ευρέως γνωστή σε επιστημονικές εργασίες συσχετισμένες με EEG.

3.1 Περιγραφικά Στατιστικά

Τα πιο απλά χαρακτηριστικά που μπορεί κανείς να χρησιμοποιήσει στην ανάλυσή του είναι αυτά που βασίζονται σε βασικά περιγραφικά στατιστικά. Είναι αρκετά συνηθισμένο στην επεξεργασία σήματος και ειδικότερα βιολογικών σημάτων να γίνεται χρήση τέτοιων χαρακτηριστικών. Δίνουν σημαντική πληροφορία για τη μορφή και τη συμπεριφορά του σήματος ενώ απαιτούν λιγότερη υπολογιστική ισχύ λόγω της απλότητάς τους. Τέτοια χαρακτηριστικά μπορούν να είναι ο μέσος όρος, η διακύμανση-διασπορά αλλά και οι ροπές μεγαλύτερου βαθμού όπως κύρτωση και ασυμμετρία. Κάθε χαρακτηριστικό δύναται να υπολογιστεί σε ολόκληρο το σήμα με τεχνικές υπολογιστικών παραθύρων τα οποία μπορεί να είναι επικαλυπτόμενα και μη.

3.2 Αυτοδιακύμανση και Αυτοσυσχέτιση

Η αυτοδιακύμανση και η αυτοσυσχέτιση είναι μέτρα εξάρτησης ανάμεσα σε μεταβλητές στις χρονοσειρές [17]. Ας υποθέσουμε ότι $Y_1, \dots, Y_n$ τυχαιές μεταβλητές με την ιδιότητα ότι η αυτοδιακύμανση

$$Cov(Y_{t+k}, Y_t) = E((Y_{t+k} - E(Y_{t+k}))(Y_t - E(Y_t))) \quad (3.1)$$

των παρατηρήσεων με υστέρηση k δεν εξαρτάται στο t . Τότε

$$\gamma(k) := \text{Cov}(Y_{k+1}, Y_1) = \text{Cov}(Y_{k+2}, Y_2) = \dots \quad (3.2)$$

ονομάζεται συνάρτηση αυτοδιακύμανσης (autocovariance function) και

$$\rho(k) := \frac{\gamma(k)}{\gamma(0)}, \quad k = 0, 1, \dots \quad (3.3)$$

ονομάζεται συνάρτηση αυτοσυσχέτισης (autocorrelation function). Με αριθμητικό τρόπο η συνάρτηση αυτοδιακύμανσης υπολογίζεται ως

$$c(k) := \frac{1}{n} \sum_{t=1}^{n-k} (y_{t+k} - \bar{y})(y_t - \bar{y}) \quad (3.4)$$

όπου $\bar{y} = \frac{1}{n} \sum_{t=1}^n y_t$ ενώ η εμπειρική αυτοσυσχέτιση από τον τύπο

$$r(k) := \frac{c(k)}{c(0)} = \frac{\sum_{t=1}^{n-k} (y_{t+k} - \bar{y})(y_t - \bar{y})}{\sum_{t=1}^n (y_t - \bar{y})^2} \quad (3.5)$$

3.3 Ανάλυση συχνοτήτων

Η ανάλυση συχνοτήτων και πιο συγκεκριμένα η ανάλυση Fourier είναι αρκετά χρήσιμη στην ανάλυση χρονοσειρών, και ειδικότερα εγκεφαλικών σημάτων καθώς περιγράφει το σήμα σε ημιτονοειδή σήματα διαφορετικών συχνοτήτων. Ο γρήγορος μετασχηματισμός Fourier (FFT) είναι ένας διαδοδόμενος και αποτελεσματικός αλγόριθμος [13] για τον υπολογισμό του Διακριτού μετασχηματισμού Φουριερ. Αξίζει να σημειωθεί ότι αμέσως μετά την ανάπτυξη του FFT, δοκιμάστηκε και εφαρμόστηκε σε ηλεκτρικά σήματα εγκεφάλου [15, 16].

3.4 Turning Points

Για τρεις συνεχόμενες μετρήσεις σε μια χρονοσειρά x_{i-1}, x_i, x_{i+1} και κάνοντας την υπόθεση ότι $x_j \neq x_k$ για οποιαδήποτε δύο από αυτά, οι συνδιασμοί που μπορούν αυτά τα τρία να τοποθετηθούν είναι 6. Το συγκεκριμένο χαρακτηριστικό ονομάζεται turning point και η πιθανότητα να βρεθεί ένα τέτοιο σημείο σε χρόνο για μια τυχαία χρονοσειρά είναι $2/3$. Έτσι ο αριθμός των turning point σε μια χρονοσειρά ορίζεται ως μια τυχαία μεταβλητή T και μια χρονοσειρά η οποία είναι τυχαία η αναμενόμενη τιμή T δίνεται από τον τύπο $\mu_T = \frac{2}{3}(n-2)$. Αξίζει να σημειωθεί ότι το $n-2$ υπάρχει καθώς στο πρώτο και στο δεύτερο σημείο μετρήσεων σε ένα σήμα δεν μπορούμε να βρούμε turning point. [58]

3.5 Ratio of the mean square successive difference to the variance

Το στατιστικό χαρακτηριστικό το οποίο περιγράφεται ως δύο φορές η Abbe value [28], το οποίο αναλύεται στην παρακάτω δημοσίευση του John Von Neumann [60] και ορίζεται ως εξής:

$$\mathcal{A} = \frac{\delta^2}{2s^2} = \frac{\frac{1}{n-1} \sum_{i=1}^{n-1} (x_{i+1} - x_i)^2}{\frac{2}{n} \sum_{i=1}^n (x_i - \bar{x})^2} \quad (3.6)$$

όπου $\bar{X}$ είναι η μέση τιμή ένα σήμα $\{x_i\}$. Η τιμή Abbe ποσοτικοποιεί το πόσο ομαλή είναι μία χρονοσειρά συγκρίνοντάς τη με το άθροισμα των διαφορών σε ρίζα ανάμεσα σε δύο μετρήσεις με την τυπική απόκλιση της χρονοσειράς. Για χρονοσειρές οι οποίες είναι αρκετά "ομαλές" η μεταβλητή πέφτει στο 0. Τέλος η ποσότητα $\mathcal{A}$ είναι μία τυχαία μεταβλητή. [58]

3.6 Hjorth Parameters

Ένα από τα πιο διάσημα χαρακτηριστικά για την ανάλυση κυρίως σημάτων εγκεφάλου εισαγάγεται από τον Hjorth [25, 24]. Ο τελευταίος, προτίνει τρεις παραμέτρους με ονόματα activity, mobility και complexity. Τα τρία αυτά χαρακτηριστικά περιγράφονται με τις παρακάτω εξισώσεις

$$Activity(x(t)) = var(x(t)) \quad (3.7)$$

$$Mobility(x(t)) = \sqrt{\frac{var\left(\frac{dx(t)}{dt}\right)}{var(x(t))}} \quad (3.8)$$

$$complexity(x(t)) = \frac{Mobility\left(\frac{dx(t)}{dt}\right)}{Mobility(x(t))} \quad (3.9)$$

Η πρώτη παράμετρος όπως παρατηρούμε και από την πρώτη εξίσωση υπολογίζει την αλλαγή του σήματος και χαρακτηρίζει την δύναμή του. Η κινητικότητα ή αλλιώς mobility υπολογίζεται παίρνοντας τη ρίζα της διακύμανσης της πρώτης παραγώγου του σήματος διαιρώντας ξανά με την παράμετρο activity του σήματος και έτσι περιγράφεται μια μέση συχνότητα του φάσματος ισχύος. Τέλος, η πολυπλοκότητα περιγράφει την αλλαγή στη συχνότητα, συγκρίνοντας την κινητικότητα της πρώτης παραγώγου του σήματος με την κινητικότητα του ίδιου του σήματος [46].

Αξίζει να σημειωθεί ότι η χρήση τους είναι αρκετά συχνή σε εργασίες που έχουν να κάνουν με εξαγωγή χαρακτηριστικών για EEG, ενώ πολύ σημαντικό είναι το γεγονός ότι υπολογίζονται σε στάσιμες χρονοσειρές για κάθε κανάλι ηλεκτροδίων [26, 46].

3.7 Συσχετίσεις μακράς διάρκειας

Σε πολλά πεδία υπάρχει υποψία ενός φαινομένου το οποίο καλείται βραχιά μνήμη ή στα αγγλικά συνήθως μπορεί κανείς να το συναντήσει “long memory” και παίζει αρκετά σπουδαίο ρόλο στην ανάλυση σημάτων και γενικότερα χρονοσειρών διαφόρων πεδίων.

3.7.1 Ιστορικό πλαίσιο

Ποιος όμως ήταν αυτός ο οποίος θεωρείται πατέρας του συγκεκριμένου μέτρου; Το 1960 ο Μπενουά Μαντεμπορντ καθώς δούλευε πάνω στο διδακτορικό του και πιο συγκεκριμένα στις στοχαστικές διαδικασίες και στις παχιές “ουρές” ήταν ο πρώτος που άρχισε να ανακαλύπτει τέτοιου είδους χαρακτηριστικά. Κίνητρο για τη δουλειά του Μαντελμπροτ αποτέλεσε η εργασία του Harold Edwin Hurst και η εργασία του πάνω στην υδρολογία. Ο τελευταίος πέρασε όλη του επαγγελματική ζωή στην Αίγυπτο και στο τέλος κατέληξε Γενικός Διευθυντής του Τμήματος Φυσικής καθώς ήταν υπεύθυνος για όλες τις υδρολογικές εργασίες όσον αφορά τον Νείλο. Ο Hurst λοιπόν, μαζί με το αντίστοιχο τμήμα υδρολογίας είχε αναλάβει μια μεγάλη υδρολογική μελέτη και εργαζόμενος πάνω σε αυτή επινόησε τη γνωστή μέθοδο R/S Analysis για την ποσοτικοποίηση της μνήμης στις χρονοσειρές την οποία και εφάρμοσε σε 690 διαφορετικές χρονοσειρές. Πιο συγκεκριμένα ο Hurst εντόπισε ότι στην περίπτωση του Νείλου οι διαφορές δεν ήταν ανεξάρτητες και χαρακτηρίζονται από μια σχέση νόμου δύναμης και εκθετη μεγαλύτερο του 0.5. Αυτό μας κατευθύνει στο σενάριο της Fractional Brownian motion. [22]

Ιδιαίτερη μνεία αξίζει η εφαρμογή του εκθέτη Hurst και γενικότερα των fractals στον οικονομικό τομέα και στις αγορές. Σύμφωνα με την αποτελεσματική υπόθεση της αγοράς (υπόθεση του τυχαίου περιπατητή) οι αγορές δεν έχουν καθόλου μνήμη και οι μεταβολές της τιμής τη συγκεκριμένη χρονική στιγμή είναι άκρως ανεξάρτητες από την ιστορία των τιμών. Αν και αρκετά ελκυστική και απλή η συγκεκριμένη θεωρία εκτιμάει αρκετά μικρές πιθανότητες σε οικονομικές κρίσεις και γεγονότα που συμβαίνουν σπάνια αλλά μπορεί να είναι καταστροφικά (black swan). Έτσι τα μοντέλα που βασίζονται στην θεωρία των fractal παρέχουν μια διαφορετική οπτική γωνία και ανοίγουν νέους δρόμους. Ήδη από το 1930 οι Frost and Prechter, 1998 ξεκίνησαν να μελετούν εναλλακτικές υποθέσεις θεωρώντας ότι οι αγορές παρουσιάζουν κάποιου είδους μνήμη και εμφανίζονται συγκεκριμένα patterns τα οποία ονομασαν “κύματα”. Από την άλλη, όπως αναφέρθηκε και παραπάνω ο Mandelbrot με την δημοσίευση [40] απέδειξε ότι κάποιες αγορές δεν ακολουθούν το random walk model καταρρίπτοντας την υπόθεση και δείχνοντας ότι υπάρχει μια στατιστική αυτο ομοιότητα στη συμπεριφορά.

Στην παρακάτω εργασία θα αναλυθούν εκτενώς διάφοροι αλγόριθμοι υπολογισμού του συγκεκριμένου εκθέτη καθώς τυχόν πλεονεκτήματα ή μειονεκτήματα του που μπορεί να έχει ο κάθε ένας. Αξίζει να αναφερθεί ότι ο συγκεκριμένος εκθέτης εμφανίζεται σε διάφορους τομείς των εφαρμοσμένων μαθηματικών, συμπεριλαμβανομένων των fractal καθώς και της θεωρίας του χάους. Αν και το μέτρο αναπτύχθηκε αρχικά στην υδρολογία οι σύγχρονοι αλγόριθμοι για τον υπολογισμό του προέρχονται από τα fractals στα μαθηματικά.

3.7.2 Fractals

Ένα Φράκταλ (Fractal) ορίζεται ως ένα μαθηματικό σύνολο με κλασματική διάσταση, δηλαδή ένα μοτίβο το οποίο παραμένει το ίδιο από όποια κλίμακα και αν παρατηρηθεί. Ο παρπάνω όρος χρησιμοποιήθηκε για πρώτη φορά από τον Benoit Mandelbrot το 1975.

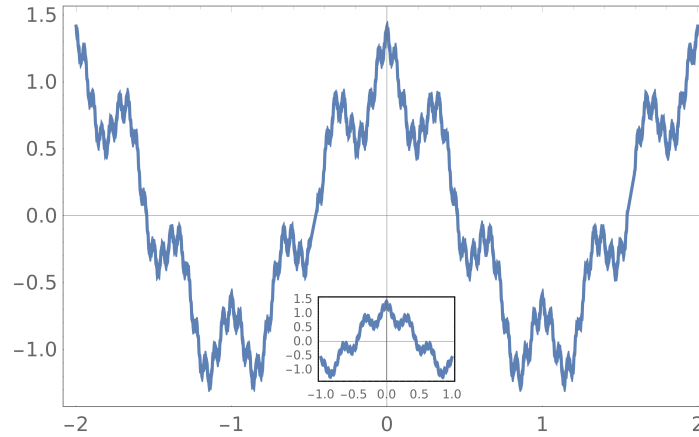


Figure 3.1: *H συνεχής αλλά πουθενά παραγωγίσιμη συνάρτηση Weistrass και τα μορφοκλασματικά χαρακτηριστικά της.*

Ενδιαφέρον αποκτά η σύνδεση του με τον εκθέτη Hurst για μια διάσταση με τη σχέση $D = 2 - H$ όπου $D \in (1, 2)$ [39]. Η συγκεκριμένη σχέση γενικεύεται [9] και για διαδικασίες σε περισσότερες διαστάσεις d : $D = d + 1 - H$ όπου H ο εκθέτης του Hurst.

Αξίζει να σημειωθεί ότι τα fractals μπορούν να βρεθούν στη φύση ή ακόμη να παραχθούν με μαθηματικό τρόπο. Το όνομα fractal προέρχεται από το λατινικό fractus που παραπέμπει στο frangere το οποίο σημαίνει "σπάζω" [39]. Παράλληλα, χωρίζονται σε δύο μεγάλες κατηγορίες, τα deterministic και not deterministic. Για παράδειγμα η διαδικασία Fractional Brownian Motion είναι not deterministic ενώ το γνωστό Mandelbrot set είναι μη γραμμικό και ντετερμινιστικό.

3.7.3 Fractional Brownian Motion

Μια συνεχής με μηδενικό μέσο όρο Γκαουσιανή διαδικασία $\{W_t, t \geq 0\}$ με συνάρτηση συνδιακρήνσης

$$\text{cov}(W_t, W_s) = \frac{1}{2} (|t|^\alpha + |s|^\alpha - |t-s|^\alpha), \quad t, s \geq 0 \quad (3.10)$$

ονομάζεται κλασματική κίνηση brown ή Fractional Brownian Motion με παράμετρο τραχύτητας (roughness) $\alpha \in (0, 2)$. Η συγκεκριμένη διαδικασία συνήθως συναντάται στη βιβλιογραφία με την παράμετρο $H = \alpha/2$ όπου H είναι ο Εκθέτης Hurst και λαμβάνει τιμές ανάμεσα στο 0 και στο 1. Η αυτόμοιότητα στην συγκεκριμένη διαδικασία εμφανίζεται καθώς η FBM υπακούει στην ιδιότητα ότι η re scaled process $\{c^{-H}W_{ct}, t \geq 0\}$ έχει την ίδια κατανομή $\{W_t, t \geq 0\}$ για κάθε c μεγαλύτερο του μηδενός.

Αξίζει να σημειωθεί ότι η κλίμακα (scale) σε μια τέτοια διαδικασία είναι $\sigma \propto n^H$ ενώ χαρακτηρίζεται από την ιδιότητα της μη στασιμότητας. Όσον αφορά τις τιμές του εκθέτη H , για $H = 1$ η διαδικασία είναι Brownian motion (random walk), $H > 1/2$ είναι μια επίμονη με μαρκά μνήμη και συσχέτιση διαδικασία ενώ με $H < 1/2$ είναι μη επίμονη με μικρή μνήμη και αντισυσχετιζόμενη διαδικασία.



Figure 3.2: Τυχαίες διαδικασίες FBM με διαφορετικούς εκθέτες Hurst σε ακραίες τιμές.

3.7.4 Fractal Gaussian Noise

Τα προαναφερθέντα δεν ισχύουν μόνο για μία μη στάσιμη διαδικασία όπως την FBM αλλά εμφανίζονται και στην Fractional Gaussian Noise [41] η οποία ορίζεται ως η μεταβολές

μια FBM

$$fGn_H(t) = B_H(t+1) - B_H(t) \quad (3.11)$$

Με τον ίδιο τρόπο μπορούμε να παράξουμε και την DFGN [58] η οποία ορίζεται ως η μεταβολή μιας FGN

$$DfGn_H(t) = fGn_H(t+1) - fGn_H(t) \quad (3.12)$$

Και οι δύο παραπάνω διαδικασίες είναι στάσιμες, σε αντίθεση με την FBM.

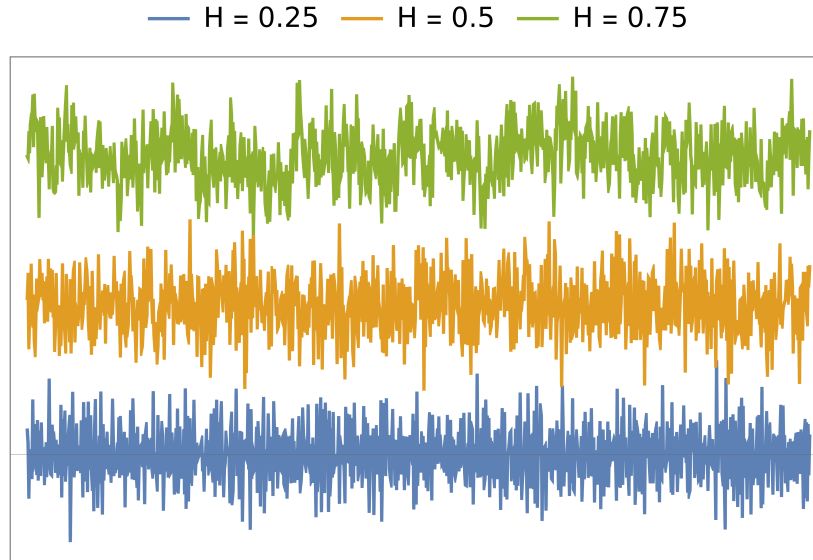


Figure 3.3: Μορφοκλασματική διαδικασία Gauss (FGN) για διαφορετικές παραμέτρους εκθέτη Hurst. (0.25, 0.5, 0.75, από κάτω προς τα πάνω)

3.7.5 Σύγχρονη εφαρμογή στις βιολογικές διαδικασίες

Ακόμη μια πρόσφατη και αρκετά διαδεδομένη εφαρμογή του εκθέτη Hurst και των κλασματικών διαδικασιών στον χαρακτηρισμό φυσικών διαδικασιών παρουσιάζεται στην δημοσίευση Single particle diffusion characterization by deep learning [20]. Στην προαναφερθείσα δημοσίευση γίνεται η χρήση ενός τεχνητού νευρωνικού δικτύου το οποίο εκπαιδεύεται με σκοπό να μπορεί να διακρίνει παραμέτρους διαφόρων διαδικασιών σύμφωνα με το τύπο της

διάχυσης και της συμπεριφοράς του. Ακόμη γίνεται χρήση του νευρωνικού για της εκτίμηση του εκθέτη Hurst σε μια διαδικασία FBM, του συντελεστή διάχυσης σε μια κίνηση Brown βάση συνθετικών δεδομένων που παράγονται με τεχνικές προσομοίωσης. Τέλος γίνεται δοκιμή του μοντέλου σε βιολογικά εργαστηριακά πειραματικά δεδομένα ανοίγοντας νέους δρόμους για το πεδίο της ακριβούς αξιολόγηση της διαδικασίας διάχυσης.

Κεφάλαιο 4

Αλγόριθμοι υπολογισμού του εκθέτη H

Στο προηγούμενο κεφάλαιο αναφερθήκαμε αρκετά στο ιστορικό αλλά και τεχνικό πλαίσιο χρήσης του Hurst exponent. Εδώ παρουσιάζονται οι υλοποιήσεις κάποιων από τον πιο σημαντικών αλγορίθμων υπολογισμού του εκθέτη καθώς και σχόλια όσον αφορά τον κάθε ένα.

4.1 Αλγόριθμος της αυξανόμενης διασπορας

Για μια στοχαστική διαδικασία με μαρκές συσχετίσεις αναμένουμε για το στατιστικό μέτρο της τυπικής απόκλισης να ισχύει ο νόμος της κλίμακας $\sigma_Y(n) \sim n^H$ και αντίστοιχα για την διασπορά του $\sigma_Y^2(n) \sim n^{2H}$. Ο συγκεκριμένος αλγόριθμος εκμεταλεύεται τους παραπάνω τύπους για τον υπολογισμό του Εκθέτη H . Πιο συγκεκριμένα, για μια χρονοσειρά Y , ο υπολογισμός του εκθέτη δίνεται από την εκτίμηση του νόμου κλιμάκωσης [3] $\sigma_Y^2(n) \sim n^{2H}$ για αυξανόμενα υπολογιστικά παράθυρα n . Έτσι πρακτικά σε κάθε τμήμα μήκους n υπολογίζονται οι διασπορές, και έπειτα ο μέσος όρος τους. Τέλος ο υπολογισμός του εκθέτη γίνεται με μια προσαρμογή ευθείας στο $s^2(n)$ έναντι του $\log n$.

Έτσι με τον παραπάνω αλγόριθμο, δύναται να πάρουμε μια γρήγορα εκτίμηση του εκθέτη H σε ένα σήμα χωρίς πολύ δύσκολους υπολογισμούς.

4.2 Αλγόριθμος ανάλυσης αλλαγής κλίμακας του εύρους

Ο συγκεκριμένος αλγόριθμος είναι σίγουρα ο πιο διαδεδομένος αλγόριθμος για τον υπολογισμό του εκθέτη H . Έχει χρησιμοποιηθεί σε μια ποικιλία εφαρμογών από οικονομικές χρονοσειρές [54, 30], ανάλυση εγκεφαλικών σημάτων [4] μέχρι φυσικές εφαρμογές [19].

Παρακάτω θα αναλυθεί ο αλγόριθμος για υπολογισμό σε μία διάσταση αλλά αξίζει να σημειωθεί ότι έχουν προταθεί στη βιβλιογραφία, γενικεύσεις [5] σε περισσότερες από μία διαστάσεις.

Πρώτο βήμα είναι ο χωρισμός του σήματος μελέτης $\{x_1, x_2, \dots, x_N\}$ σε τμήματα μήκους n . Σε κάθε τμήμα αφαιρούμε από το αρχικό τμήμα τον μέσο όρο και έπειτα υπολογίζουμε το προφίλ αθροίζοντας με βάση τον τύπο

$$Y_v(j) = \sum_{i=1}^j (x_{vn+i} - \bar{x}_{vn}) \quad (4.1)$$

Έπειτα σε κάθε τμήμα ομοίως με πριν, υπολογίζουμε το εύρος αφαιρώντας την μεγαλύτερη από την μικρότερη τιμή, και τέλος την τυπική απόκλιση του προφίλ που δημιουργήθηκε πριν.

$$R_v(n) = \max_{j=1}^n Y_v(j) - \min_{j=1}^n Y_v(j) \quad (4.2)$$

$$S_v(n) = \sqrt{\frac{1}{n} \sum_{j=1}^n Y_v^2(j)} \quad (4.3)$$

Τέλος, υπολογίζουμε το μέσο όρο από τον λόγο της εξίσωσης 4.3 και 4.2 για όλα τα τμήματα φυσικά που δημιουργήθηκαν.

$$F_{RS}(n) = \frac{1}{N_n} \sum_{v=0}^{N_n-1} \frac{R_v(n)}{S_v(n)} \quad (4.4)$$

Τέλος για μεγάλα υπολογιστικά παράθυρα μήκους n αναμένεται να ισχύει ο νόμος της κλίμακας $F_{RS}(n) \sim n^H$ όπως και στον αλγόριθμο της αυξανόμενης διασποράς.

4.3 Αλγόριθμος βασισμένος στον Haar Wavelet

Υπάρχουν αρκετοί αλγόριθμοι [37, 57, 42] στη βιβλιογραφία η οποίοι κάνουν χρήση wavelets για την εκτίμηση του παραμέτρου H . Ο συγκεκριμένος αλγόριθμος [59] που θα παρουσιαστεί εδώ βασίζεται στον Haar wavelet και φαίνεται να είναι αρκετά αποδοτικός και ακριβείς σε σχέση με άλλους.

Πιο συγκεκριμένα, υποθέτουμε ότι έχουμε μια χρονοσειρά $x(t)$ η οποία αποτελείται από n τιμές θεωρούμε πως χαρακτηρίζεται από αυτόμοιότητα όταν ο εκθέτης Hurst αναλογικά $t \rightarrow \lambda t$ ικανοποιεί τη συνθήκη της παρακάτω εξίσωσης

$$x(t) \doteq \lambda^{-H} x(\lambda t) \quad (4.5)$$

όπου $\doteq$ δηλώνει ομοιότητα στην κατανομή. Η βάση υπολογίζεται από τον λεγόμενο mother wavelet μέσω της εξίσωσης [42]

$$\psi(t) \rightarrow \psi_{j,k}(t) = 2^{-j/2} \psi(2^{-j}t - k) \quad (4.6)$$

όπου το j αναπαριστά το κλίμακα χρόνου ή αλλιώς οκτάβα (octave) στην αντίστοιχη θέση k του wavelet. Ακόμη ο συντελεστής μετασχηματισμού του wavelet ορίζεται ως

$$d_{j,k} = \langle x, \psi_{j,k} \rangle \quad (4.7)$$

και τέλος η μέση δύναμη της χρονοσειράς σε κλίμακα χρόνου j εκφράζεται μέσω της διακύμανσης των συντελεστών του wavelet [14]

$$\text{var}(d_{j,k}) = \frac{1}{2^j} \sum_{k=0}^{2^j-1} |d_{j,k}|^2 \quad (4.8)$$

Υποθέτοντας ότι ένα σήμα μήκους $n = 2^m$, $m \in N$ τότε 2^j βάση του παραπάνω τύπου είναι ο αριθμός των coefficients σε κλίμακα j . Επιπλέον πρέπει να ισχύει $0 \leq j \leq m - 1$ όπως και $0 \leq k \leq 2^j - 1$ [52]. Τέλος, τις δύο παραπάνω αρχικές εξισώσεις παρατρούμε την σχέση ανάμεσα στη διακύμανση των coefficient wavelets και της κλίμακας j

$$\log_2 \text{var}(d_{j,k}) = \alpha \cdot j + \text{const.} \quad (4.9)$$

Τέλος, ο εκθέτης Hurst υπολογίζεται ως η παράμετρος α , μέσω της προσαρμογής ευθείας στο $\log_2 \text{var}(d_{j,k})$ έναντι του j . Εδώ αξίζει να σημειωθεί ότι το πλεονέκτημα της συγκεκριμένης μεθόδου είναι ότι από την τιμή α , γίνεται διακριτή η τιμή του H αλλά και κατηγοροποίηση του σήματος σε FBM, FGN και DFGN με βάση το παρακάτω τυπολόγιο. Για διαδικασίες FBM ο εκθέτης είναι $H = \frac{\beta-1}{2}$, για FGN $H = \frac{\beta+1}{2}$ και τέλος DFGN $H = \frac{\beta+3}{2}$.

4.4 Εφαρμογή αλγορίθμων σε προσομοιωτικά δεδομένα

Για διευρευνητικούς λόγους όπως και για λόγους αξιολόγησης θα γίνει χρήση προσομοιωτικών δεδομένων από τις διαδικασίες FBM FGN και DFDN και εφαρμογή των αλγορίθμων που αναφέρθηκαν προηγουμένως.

Ο αλγόριθμος Αυξανόμενης διασποράς είναι ένας υπολογιστικά γρήγορος αλγόριθμος, καθώς το κύριο κομμάτι του είναι να υπολογίσει τις διασπορές από τα παράθυρα που χωρίζουν το σήμα και επείτα να βρει τον κάθε μέσο όρο αντίστοιχα. Στο παρακάτω διάγραμμα βλέπουμε 3 ιστογράμματα από την εφαρμογή του αλγορίθμου σε διαδικασίες FBM. Πιο συγκεκριμένα, προσομοιώθηκαν 10000 χρονοσειρές μήκους 1024 σημείων, για εκθέτες $H = 0.25$ - $H = 0.5$ - $H = 0.75$ και εφαρμόστηκαν στον αλγόριθμο. Αν και γίνεται μια καλή προσέγγιση, παρατηρούμε ότι οι μέσοι όροι των ιστογραμμάτων (0.24, 0.48, 0.71) αντίστοιχα για κάθε εκθέτη διαφέρουν από τους εκθέτες που δόθηκαν ως είσοδο (0.25, 0.5, 0.75) στις χρονοσειρές ενώ υπάρχει αρκετά μεγάλη διασπορά στις τιμές.

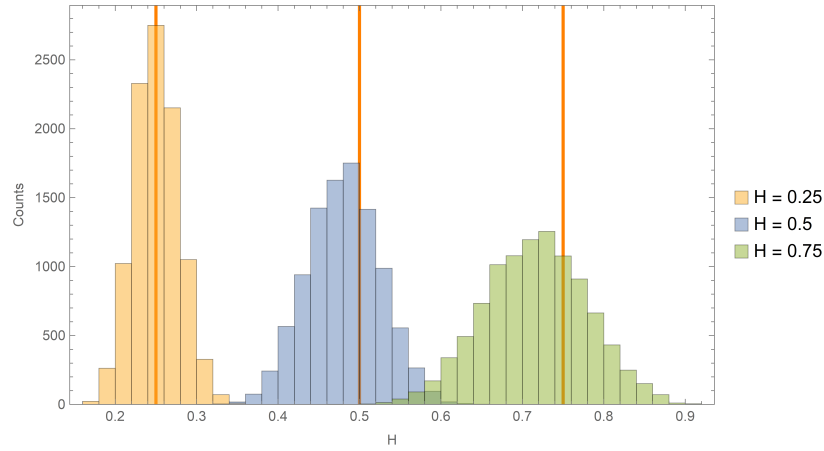


Figure 4.1: Ιστογράμματα από προσομιώσεις διαδικασίες FBM για διαφορετικούς εκθέτες H όσον αφορά την μέθοδο την αυξανόμενης διασποράς.

Ο Rescale Range είναι ίσως ο πιο διαδομένος για τον υπολογισμό του εκθέτη Hurst και ίσως και ο πιο αμφιλεγόμενος στα αποτελέσματα του. Στο συγκεκριμένο πείραμα, προσομοιώθηκαν χρονοσειρές μήκους 1024 στοιχείων με εκθέτη Hurst από 0.05 έως 0.95 με βήμα 0.05 και για κάθε H το δείγμα ήταν 1000 διαφορετικές. Αυτό μας δίνει τη δυνατότητα να μελετήσουμε την συμπεριφορά του αλγόριθμου σε όλο το φάσμα υπολογισμού και να εντοπιστούν τυχόν αδυναμίες. Όπως παρατηρούμε και από το παρακάτω διάγραμμα ο αλγόριθμος άλλοτε υπερεκτιμάει τον υπολογισμό του εκθέτη και άλλοτε προσφέρει μικρότερες τιμές από αυτές που πραγματικά η χρονοσειρά έχει. Ακόμη, αξίζει να σημειωθεί ότι για μη στάσιμες διαδικασίες όπως την FBM ο αλγόριθμος υπολογίζει τιμές κοντά στο 1, γεγονός που φανερώνει ένδειξη μη στασιμότητας [49].

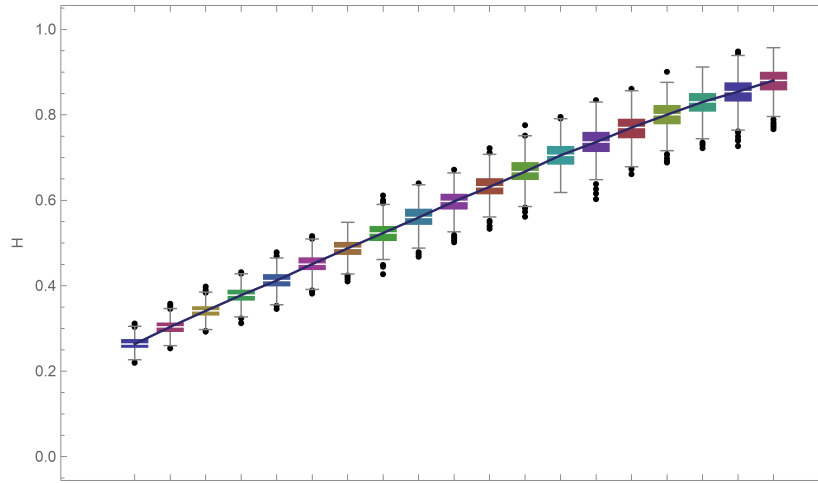


Figure 4.2: Διάγραμμα θηκογραμμάτων για μια ποικιλία εκθετών H όσον αφορά την μέθοδο Ανάλυσης Κλίμακας Εύρους.

Ο τρίτος αλγόριθμος που παρουσιάζεται είναι αυτός που βασίζεται στον Haar Wavelet. Υπολογιστικά, ο αλγόριθμος είναι γρήγορος ενώ αυτό που τον κάνει να ξεχωρίζει είναι η δυνατότητα μέσω των τύπων που παρουσιάστηκαν στο κεφάλαιο 2, να ξεχωρίζει τις διαδικασίες FBM, FGN και DFGN. Μπορεί να χρησιμοποιηθεί για στάσιμες και μη στάσιμες χρονοσειρές ενώ είναι αρκετά ακριβής και απλός. Για την αξιολόγησή του, αρχικά προσομοιώνονται 1000 χρονοσειρές FBM μήκους 16384 σημείων για διαφορετικούς εκθέτες (0.25, 0.5, 0.75). Οι κατανομές των εκτιμήσεων φαίνονται στο διάγραμμα της επόμενης σελίδας όπου παρατηρούμε ότι όσο πιο μικρότερος είναι ο εκθέτης τόσο μεγαλύτερη είναι η "υποτίμηση" της εκτίμησης. Τέλος, αυτό φαίνεται πιο ξεκάθαρα στον παρακάτω πίνακα βασικών στατιστικών της προαναφερθείσας προσομοίωσης.

Διαδικασία	Μέγεθος Σήματος (n)	Στατιστικά
FBM $H = 0.25$	16384	$\mu = 0.20, \sigma = 0.02$
FBM $H = 0.5$	16384	$\mu = 0.47, \sigma = 0.03$
FBM $H = 0.75$	16384	$\mu = 0.72, \sigma = 0.04$
FGN $H = 0.25$	16384	$\mu = 0.23, \sigma = 0.02$
FGN $H = 0.5$	16384	$\mu = 0.48, \sigma = 0.03$
FGN $H = 0.75$	16384	$\mu = 0.73, \sigma = 0.04$

Table 4.1: Πίνακας βασικών χαρακτηριστικών όσον αφορά την αξιολόγηση του αλγορίθμου Wavelet σε διάφορα είδη χρονοσειρών και εκθετών.

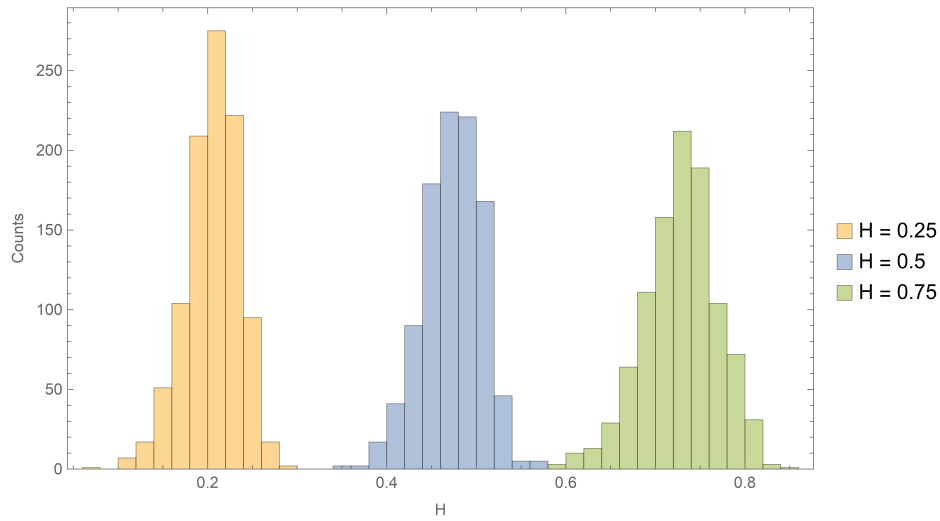


Figure 4.3: Αποτελέσματα αλγορίθμου στα δεδομένα προσομοίωσης για διαδικασίες FBM και τρεις διαφορετικούς εκθέτες.

Έτσι παρατηρώντας τα ιστογράμματα για τις προσομοιώσεις FBM και FGN και σε συνδιασμό με το πίνακα στατιστικών καταλήγουμε ότι ο συγκεκριμένος αλγόριθμος είναι αρκετά έμπιστος για να εφαρμοστεί σε εκτιμήσεις του εκθέτη H . Ακόμη τα παραπάνω αποτελέσματα αναπαράγουν ακριβώς τη δημοσίευση του Mariuz Tarnopolski [58]. Ακόμη οπτικά οι κατανομές για τις δοκιμές στη διαδικασία FGN είναι πιο στενές σε σχέση με την FBM γεγονός που επικουρικά με τους μέσους όρους του table μας δείχνει ότι γίνεται μια πιο ακριβής εκτίμηση στις διαδικασίες FGN.

Αν και η συμπεριφορά του αλγορίθμου είναι σε γενικά πλαίσια ακριβής για διαδικασίες FBM και FGN, δεν μπορεί να υποθεί το ίδιο και για την διαδικασία DFGN. Όπως γίνεται αντιλείπτο στο διάγραμμα 4.5 και πιο συγκεκριμένα στο πρώτο υποδιάγραμμα. Στην συγκεκριμένη περίπτωση, προσομοιώθηκαν διαδικασίες FGN και έπειτα παράχθηκαν μέσω διαφορών οι DFGN. Ο αλγόριθμος δεν είναι καθόλου ακριβής στην συγκεκριμένη περίπτωση δίνοντας άλλοτε μεγαλύτερους ή μικρότερους εκθέτες ενώ αρκετές φορές όπως αναφέρεται και στην προαναφερθήσα δημοσίευση [58] κατατάσει την διαδικασία σε άλλη κλάση και όχι αυτή της DFGN. Αν και το συγκεκριμένο πρόβλημα κάνει τον αλγόριθμο ακατάλληλο για την συγκεκριμένη περίπτωση το πρόβλημα λύνεται αθροίζοντας (cumulative sums) το προφίλ της χρονοσειράς και μετατρέποντας την σε FGN βάση της εξίσωσης του ορισμού. Έτσι με τον παρακάτω μετασχηματισμό η εκτίμηση γίνεται σαφώς πιο ακριβής αν και υπάρχουν ακόμη λίγες ακραίες τιμές οι οποίες εκτιμούν διαφορετικό εκθέτη από τον δοθέν (διάγραμμα 4.5, δεύτερο υποδιάγραμμα).

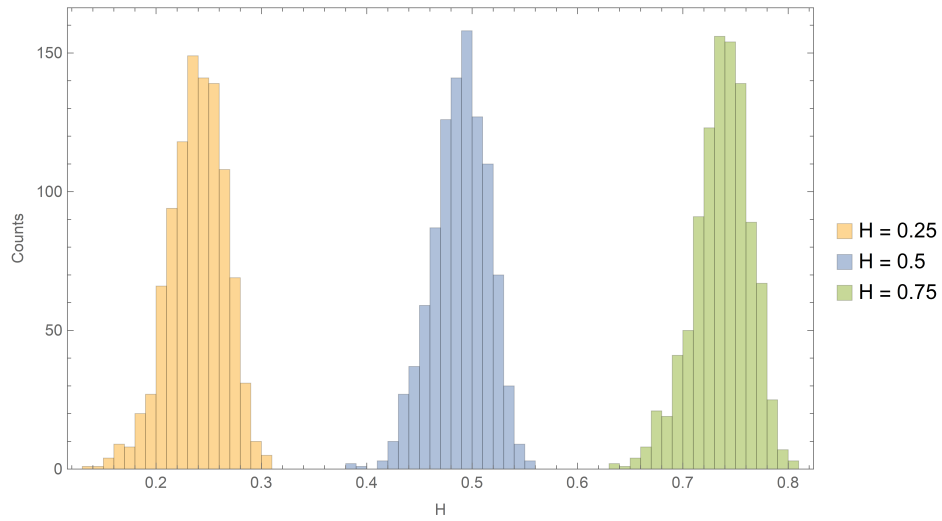


Figure 4.4: Αποτελέσματα αλγορίθμου στα δεδομένα προσομοίωσης για διαδικασίες FGN και τρεις διαφορετικούς εκθέτες.

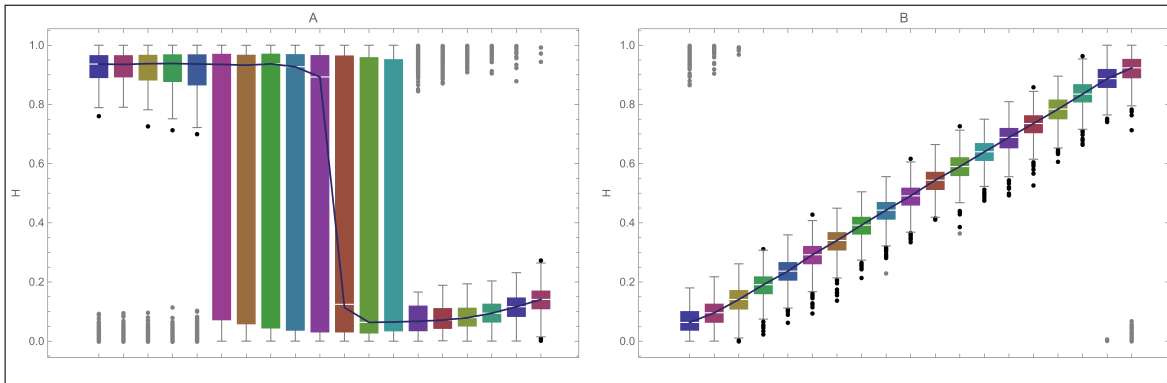


Figure 4.5: Αριστερά η αξιολόγηση του αλγορίθμου Haar Wavelet σε διαδικασίες DFGN. Δεξιά η ίδια αξιολόγηση αλλά μετά από τον μετασχηματισμό των χρονοσειρών για την μετατροπή τους σε FGN.

4.5 Προτεινόμενος Αλγόριθμος

Στο συγκεκριμένο υποκεφάλαιο προτείνεται ένας νέος αλγόριθμος ο οποίος διαφέρει από τους παραπάνω αλγορίθμους που μπορεί κανείς να αναζητήσει στη βιβλιογραφία. Συνδυάζει τη μέθοδο της ταξινόμησης καθώς και τις στατιστικές σχέσεις που προέκυψαν από προσομοιώσεις στην δημοσίευση [58].

Όπως παρουσιάζεται εκτενώς στην παραπάνω δημοσίευση, υπάρχει άμεση σχέση της τιμής του Hurst με την τιμή Abbe που περιγράφεται στο δεύτερο κεφάλαιο. Μελετώντας τη σχέση που έχουν οι παράμετροι Hurst και Abbe μέσω στατιστικών μοντέλων προκύπτουν εξισώσεις που περιγράφουν την μεταβολή του εκθέτη H βάση του Abbe value. Αυτό μας δίνει τη δυνατότητα να υπολογίσουμε με μικρό υπολογιστικό κόστος την τιμή Abbe και απευθείας να έχουμε μία σχετικά καλή προσέγγιση του εκθέτη Hurst. Παρ'όλα αυτά όμως, πρέπει να είναι γνωστή η διαδικασία από πριν για να μπορεί να εφαρμοστεί η αντίστοιχη εξίσωση. Το συγκεκριμένο πρόβλημα λύνεται με τη μέθοδο της ταξινόμησης. Έχοντας έναν ταξινομητή ο οποίος μπορεί να διαλέξει με αρκετά μεγάλη ακρίβεια τη σωστή διαδικασία δύναται να γίνει η προσέγγιση του H με έναν απλό υπολογισμό της τιμής Abbe στο αντίστοιχο σήμα και έπειτα εφαρμογή της αντίστοιχης εξίσωσης.

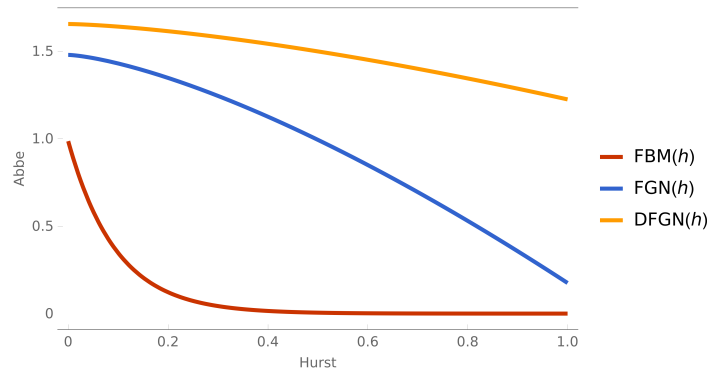


Figure 4.6: Η μαθηματικές σχέσεις ανάμεσα στην τιμή Abbe και H που προκύπτουν από προσομοιώσεις και στατιστική μοντελοποίηση [58].

Για την ταξινόμηση έγινε χρήση ενός Δένδρου Απόφασης (Decision Tree) [55] το οποίο εκπαιδεύτηκε χρονοσειρές τύπου FBM, FGN και DFGN. Για να καλυφθεί όλο το φάσμα των διαφόρων εκθετών, δημιουργούνται συνθετικά δεδομένα για κάθε τύπο process. Συνολικά δημιουργείται δείγμα 3000 χρονοσειρών για κάθε βήμα εκθέτη (0.025, 0.95, 0.025) και το μήκος του κάθε σήματος αποτελείται από 1024 τιμές. Έτσι, το σετ εκπαίδευσης αποτελείται από 342.000 δείγματα. Όσον αφορά τον χρόνο εκπαίδευσης είναι άμεσα συνδεδεμένος στο μήκος της χρονοσειράς όπως και στον αριθμό του δείγματος. Στην προκειμένη περίπτωση το δένδρο απόφασης εκπαιδεύεται στα 20 λεπτά με

παραμέτρους Distribution Smoothing = 1 και Feature Fraction = 1. Τέλος, για να αξιολογηθεί το δένδρο απόφασης δημιουργήθηκε ένα δοκιμαστικό σετ δεδομένων βασισμένο στην ίδια νοοτροπία με πριν. Πιο συγκεκριμένα, για να επιτευχθεί η αναλογία του δοκιμαστικού σετ δεδομένων περίπου 30 τις εκατό του συνολικού αρχικού dataset, προσομοιώνονται σε κάθε βήμα 1000 διαφορετικές διαδικασίες ίδιου φυσικά μήκους (1024) -σύνολο 114.000 διαφορετικές χρονοσειρές.

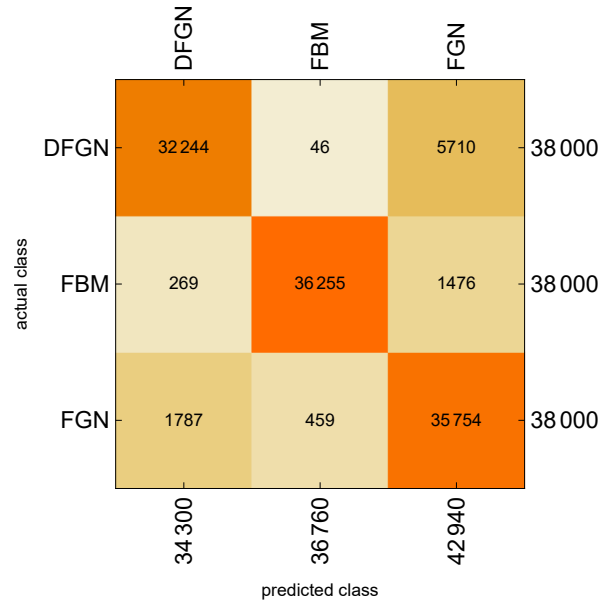


Figure 4.7: *Ο Confusion Matrix της ταξινόμησης.*

Όπως γίνεται αντιληπτό και από το confusion matrix, η ταξινόμηση είναι αρκετά ακριβής (0.91) στο δοκιμαστικό σετ. Αξίζει να σημειωθεί ότι ο αλγόριθμος έχει ένα κύριο μειονέκτημα. Για πολύ μικρές παραμέτρους H (0.05- 0.10) σε διαδικασίες FBM, εσφαλμένα τις ταξινομεί σε FGN. Με σκοπό την ανάδειξη του μειονεκτήματος, δημιουργούνται συνθετικά δεδομένα για διαδικασία FBM με πολύ μικρό εκθέτη $H = 0.05$, δείγματος 5000 και μήκος χρονοσειράς 1024. Όπως φαίνεται και στο παρακάτω διάγραμμα λίγο λιγότερο από τις μισές φορές ο ταξινομητής λανθασμένα ταξινομεί τη διαδικασία σε FGN.

Όπως διακρίνεται και στο πάνω μέρος της εικόνας η συγκεκριμένη διαδικασία FBM αναγνωρίζεται από τον ταξινομητή με σημαντική πιθανότητα 0.89 για FGN. Οι διαδικασίες οπτικά μοιάζουν αρκετά, παρ' όλα αυτά προκειται για δύο τελείως διαφορετικούς μηχανισμούς.

Το μεγαλύτερο πλεονέκτημα του αλγορίθμου αποτελεί η επίλυση του ζητήματος που εμφανίζεται στον αλγόριθμο Haar Wavelet και αφορά τις λανθασμένες ταξινομήσεις των

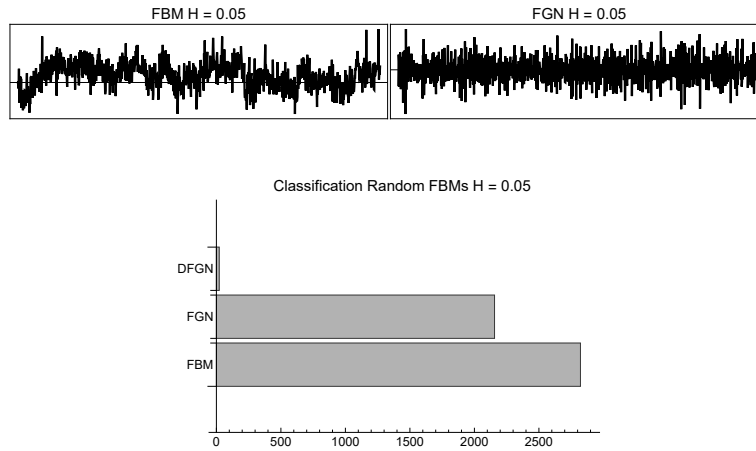


Figure 4.8: Πάνω: Δύο τυχαίες διαδικασίες FBM και FGN με εκθέτη $H = 0.05$. Κάτω: Ραβδόγραμμα αλγορίθμων με βάση τον προτεινόμενο αλγόριθμο για τις περιπτώσεις μικρού εκθέτη σε διαδικασίες FBM .

διαδικασιών $DFGN$. Αυτό φαίνεται ξεκάθαρα στην παρακάτω σύγκριση των δύο αλγορίθμων για προσομοιωτικά δεδομένα $DFGN$ μεγέθους δείγματος 5000 και μήκος χρονοσειράς 1024.

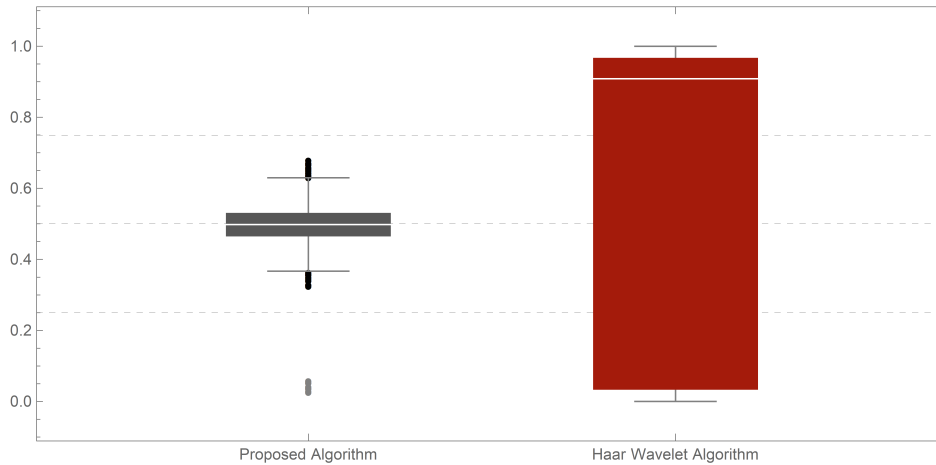


Figure 4.9: Διαφορά κατανομής των δύο αλγορίθμων σε προσομοιωτικά δεδομένα $DFGN$ με $Hurst = 0.5$.

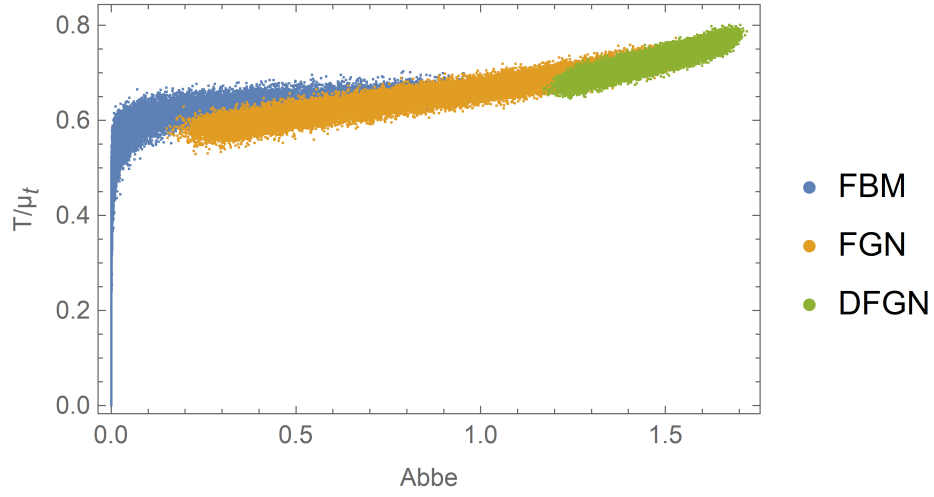


Figure 4.10: *To A/T plane για διαδικασίες μήκους 1024 στοιχείων που προτάθηκε για πρώτη φορά από τον M. Tarnopolski στο [58].*

Algorithm 1 Proposed Algorithm

```

0: procedure COMPUTE HURST ( $signal$ )

0:    $abbevalue \leftarrow \mathcal{A}(signal)$ 
0:    $findclass \leftarrow Classifier(signal)$ 

1: if  $findclass = FBM$  then
1:    $hurst \leftarrow Solve \mathcal{HFBM}(abbevalue)(1)$  [58]

2: else if  $findclass = FGN$  then
2:    $hurst \leftarrow Solve \mathcal{HFGN}(abbevalue)(2)$  [58]

3: else
3:    $hurst \leftarrow Solve \mathcal{HDFGN}(abbevalue)(3)$  [58]

4: end if

return  $hurst$ 

```

Κεφάλαιο 5

Ηλεκτροεγκεφαλογραφία

5.1 Φυσιολογία

Ο ανθρώπινος εγκέφαλος είναι το κύριο όργανο του κεντρικού νευρικού συστήματος του ανθρώπου βασίζεται στη λειτουργία νευρώνων οι οποίοι μεταφέρουν πληροφορία μέσω των συνάψεων. Στον ανθρώπινο εγκέφαλο υπάρχουν περίπου 10^{10} νευρώνες και αυτοί με τη σειρά τους είναι συνδεδεμένοι με τρισεκατομμυρια συνάψεις. Οι νευρώνες είναι κύτταρα τα οποία είναι ευερέθιστα στον ηλεκτρισμό και πρακτικά επεξεργάζονται και διαβιβάζουν πληροφορίες. Οι τελευταίες είναι σήματα τα οποία ενισχύονται κατά μήκος της πορείας του νευρώνα και μεταδίδουν αυτές τις πληροφορίες από και προς το Κεντρικό Νευρικό Σύστημα. [8]

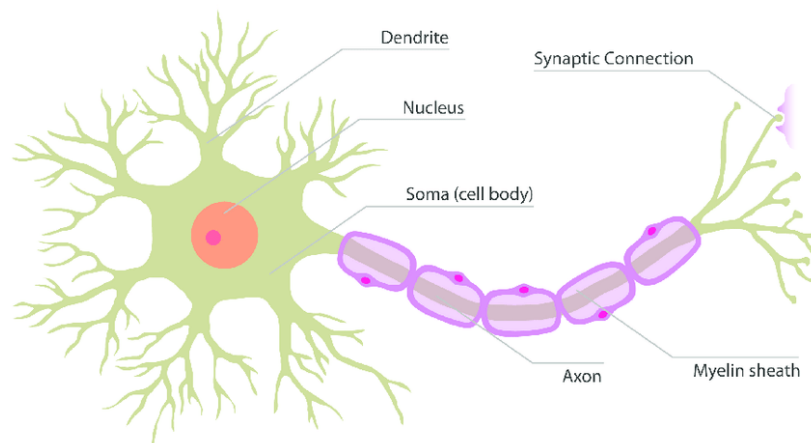


Figure 5.1: Δομή ενός πραγματικού νευρώνα. [18]

Ο ανθρώπινος εγκέφαλος χωρίζεται σε διαφορετικές περιοχές οι οποίες αφορούν διαφορετικές λειτουργίες. Μια πλάγια όψη του εγκεφάλου στην οποία γίνονται διακριτοί οι λοβοί του εγκεφαλικού φλοιού στο αριστερό ημισφαίριο φαίνονται στην παρακάτω εικόνα.

Ακόμη ο εγκέφαλος μπορεί να χωριστεί σε 3 βασικά μέρη : Το cerebrum, cerebellum, και brainstem.

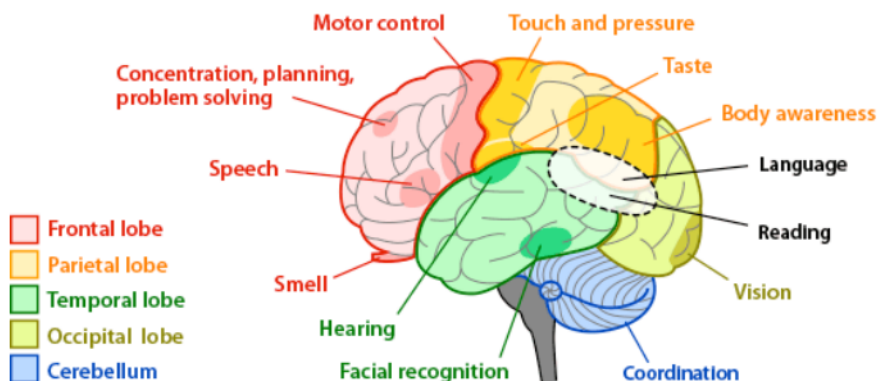


Figure 5.2: Περιοχές εγκεφάλου και αντίστοιχες λειτουργίες τους.[1]

Το cerebrum αποτελεί το μεγαλύτερο μέρος και αποτελείται από το δεξί και το αριστερό ημισφαίριο, ενώ κάθε ημισφαίριο χωρίζεται σε broad regions και lobes. Κάθε ημισφαίριο από το cerebrum χωρίζεται σε 4 κύριους λοβούς και ένα κρύφο μικρό λοβό που ονομάζεται insula. Ο μετωπιαίος λοβός ο οποίος είναι υπεύθυνος για την ευθύνη και αφαιρετική σκέψη, την κρίση και την κίνηση. Ο βρεγματικός λοβός ο οποίος είναι συνδεδεμένος με τις αισθήσεις. Ακόμη ο κροταφικός λοβός σχετίζεται με τη μνήμη και την ακοή και τέλος ο ινικός λοβός που είναι υπεύθυνος για την ανθρώπινη όραση.

Για την μέτρηση της ηλεκτρικής δραστηριότητας του εγκεφάλου γίνεται χρήση αισθητήρων οι οποίοι τοποθετούνται εξωτερικά του κρανίου με τη μέθοδο της ηλεκτροεγκεφαλογραφίας ή αλλιώς EEG. Η τελευταία προτάθηκε για πρώτη φορά από τον Berger το 1929 και τα μεγέθη μετρώνται με μV . Οι μετρήσεις γίνονται μέσω της τοποθέτησης ηλεκτροδίων σε συγκεκριμένες θέσεις ώστε να καλύπτεται όσο το δυνατόν περισσότερο μέρος στο ανθρώπινο κρανίο. Αρχικά συνιθισμένη είναι η χρήση του διεθνούς συστήματος 10/20 το οποίο περιλαμβάνει 21 ηλεκτρόδια καταγραφής [27]. Μπορεί κανείς να συναντήσει ακόμη το σύστημα 10/10 [50] το οποίο περιλαμβάνει 81 ηλεκτρόδια ενώ δεν είναι λίγες οι φορές που γίνεται χρήση πυκνότερων συστημάτων.

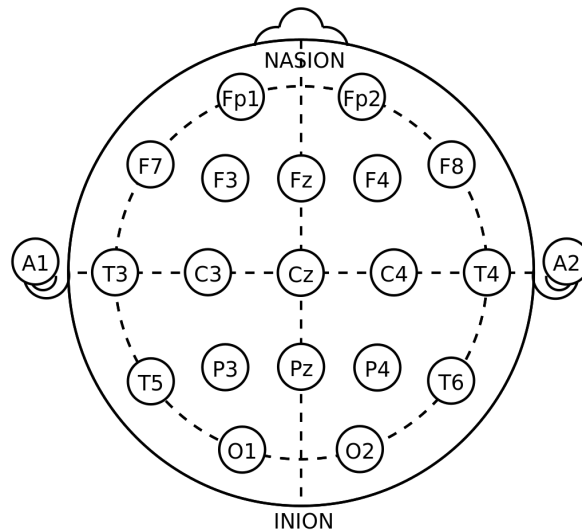


Figure 5.3: Το διεθνές σύστημα τοποθέτησης 10-20

Οι κύριες συχνότητες των ηλεκτροδίων είναι τα σήματα δέλτα τα οποία βρίσκονται κάτω από 4 Hz, τα theta 4-8 Hz, τα alpha (9-13), τα β (14-30Hz) και τέλος τα gamma κάτω από 30. [47, 44].

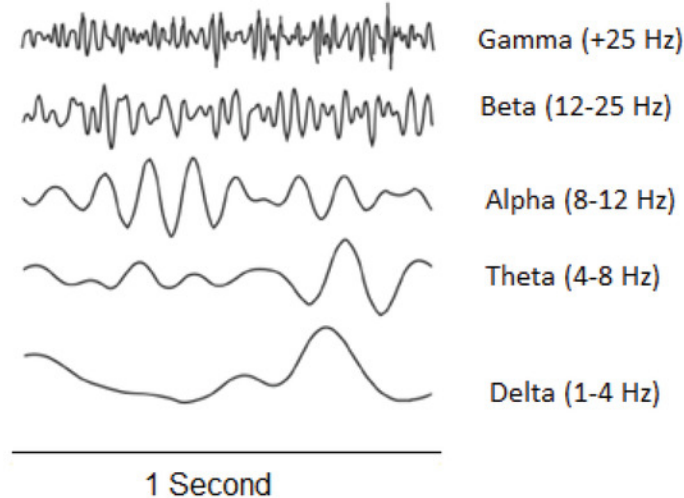


Figure 5.4: Εγκεφαλικοί ρυθμοί. [45]

Παράλληλα, με τα σήματα καταγραφής δύναται να υπάρξουν και σήματα μη εγκεφαλικής προέλευσης τα οποία καταγράφονται κατά τη διάρκεια ενός ΗΕΓ και συνήθως συναντώνται στην ξένη βιβλιογραφία ως artifacts [32]. Ένα κλασσικό παράδειγμα αποτελούν τα οφθαλμικά artifacts τα οποία είναι αποτέλεσμα μικρών ή μεγαλύτερων κινήσεων των ματιών είτε ακόμη και των βλεφάρων. Ακόμη κάποια artifacts προέρχονται και από το περιβάλλον

- ως εξωτερικός θόρυβος κατά τη διάρκεια μίας καταγραφής για αυτό και αρκετές φορές γίνεται χρήση φίλτρων για την εξάλειψή τους.

5.2 Βιβλιογραφική ανασκόπηση για την εφαρμογή του H σε ΗΕΓ δεδομένα

Όπως έχει αναφερθεί και στα προηγούμενα κεφάλαια υπάρχει αρκετά πλούσια δημοσιευμένη βιβλιογραφία με εφαρμογές και προτεινόμενους αλγόριθμους όσον αφορά των υπολογισμό του Hurst Exponent και ειδικότερα αρκετές εφαρμογές σε ηλεκτροεγκεφαλογραφικά δεδομένα.

Αρκετά πρόσφατα, με την δημοσίευση "TTA, a new approach to estimate Hurst exponent with less estimation error and computational time" [34] το πεδίο συνεχίζει να εξελίσσεται και οι συγγραφείς της προαναφερθείσας δημοσίευσης προτείνουν έναν καινούργιο αλγόριθμο βασισμένο σε υπολογισμό χώρου σε τρίγωνα που προκύπτουν από τιμές της χρονοσειράς εμπνευσμένοι από την φιλοσοφία του αλγόριθμου DFA (detrended fluctuation analysis). Ο προτεινόμενος αλγόριθμος δοκιμάζεται σε συνθετικά δεδομένα αλλά και δεδομένα EEG από την βάση δεδομένων του πανεπιστημίου του Bonn με στόχο την ανίχνευση της επιληψίας [6]. Τα αποτελέσματα δείχνουν ότι η νέα προσσέγγιση είναι αρκετά πιο γρήγορη και ευέλικτη ενώ υπάρχουν λιγότερα σφάλματα συσχετιζόμενα φυσικά με το μήκος του σήματος.

Παράλληλα με τις δημοσιευμένες εργασίες, υπάρχουν και αρκετά πιο συνεκτικές δουλειές όπως το διδακτορικό "Long Range Analysis of EEG Data" [7] στο οποίο γίνεται μία αρκετά εξειδικευμένη μελέτη σε σήματα EEG και στις μακράς διάρκειας διακυμάνσεις που εμφανίζονται. Στην διατριβή αναπτύσσονται αλγόριθμοι και στατιστικά εργαλεία με σκοπό να απαντηθούν ανοιχτά ερωτήματα για την πολύπλοκη δυναμική των σημάτων.

Αρκετά συχνά σε ερευνητικές προσεγγίσεις νευροεπιστήμης μπορεί κανείς να συναντήσει την εμφάνιση δύο ομάδων, για παράδειγμα ηγιών ατομών (control group) και ασθενών με συγκεκριμένη ασθένεια (πχ επιληψία) με σκοπό τη διαφοροποίηση των εγκεφαλικών σημάτων τους και την έρευνα διαφόρων μοτίβων (patterns) που δύναται να τους διαφοροποιούν χρονικά ή χωρικά. Υπάρχουν αρκετές δημοσιευμένες δουλειές όσον αφορά επιληπτικούς ασθενείς και αλγόριθμους βασισμένους στον εκθέτη H με πιο πρόσφατες τις [33, 63]. Παρ' όλα αυτά, δύσκολα μπορεί κανείς να βρει δημοσιεύσεις με σύγκριση σημάτων EEG σε ασθενείς με μετατραυματικό στρες και υγιή άτομα. Αυτό οφείλεται κυρίως στην πολυπλοκότητα της ασθένειας του PTSD [62] και στη δυσκολία του να γίνει μια τέτοια έρευνα με ανθρώπους που έχουν το σύνδρομο καθώς η ιατρική διάγνωση είναι αρκετά πιο πολύπλοκη σε σχέση με άλλες όπως αυτή της επιληψίας.

Έτσι, σύμφωνα με τα παραπάνω αρκετό ενδιαφέρον παρουσιάζει η ανάλυση και περαιτέρω μελέτη της αρκετά πρόσφατης δημοσίευσης [56] η οποία αποτελεί από τις λίγες δημοσιεύσεις στο πεδίο σύγκρισης εγκεφαλικών κυμάτων σε ασθενείς με PTSD και ασθενείς η οποίοι δεν έχουν τον σύνδρομο. Για την παραπάνω ανάλυση επιστρατεύεται η μέθοδος R/S analysis για την ποσοτικοποίηση του εκθέτη H ενώ μετά από ανάλυση των δεδομένων και εφαρμογή του αλγόριθμου βρίσκεται σημαντική διαφορά στο κανάλι (electrode) F3.

Κεφάλαιο 6

Εφαρμογή σε πραγματικά δεδομένα

Στο συγκεκριμένο κεφάλαιο λαμβάνει χώρα μία πρακτική εφαρμογή των προαναφερθέντων τεχνικών στα δεδομένα που παρέχονται ελεύθερα από την δημοσίευση [56] και αφορούν ασθενείς με PTSD και το control group.

Στο προηγούμενο κεφάλαιο έγινε αναφορά σε εργασίες οι οποίες προτείνουν και εφαρμόζουν αλγορίθμους βασισμένους στον Hurst Exponent σε ΗΗΓ δεδομένα. Στην δημοσίευση "Dynamical Hurst analysis identifies EEG channel differences between PTSD and healthy controls" οι συγγραφείς εμπνευσμένοι από την ανάλυση του Jeong-Ho Chae [10] ο οποίος βρήκε μειωμένη πολυπλοκότητα σε συγκεκριμένα κανάλια σε ασθενείς με μετατραυματικό στρες, διερευνούν τα σήματα μεταξύ 6 ανεξάρτητων υγιών ασθενών και 6 οι οποίοι έχουν διαγνωστεί με σύνδρομο μετατραυματικού στρες. Συγκεκριμένα, η ανάλυση περιλαμβάνει 31 κανάλια (Fp1, Fp2, F3, F4, C3, C4, P3, P4, O1, O2, F7, F8, T7, T8, P7, P8, Fz, Cz, Pz, Oz, FC1, FC2, CP1, CP2, FC5, FC6, CP5, CP6, TP9, TP10, POz). Όπως σε όλες τις καταγραφές ηλεκτροεγκεφαλογραφήματος, ένα από τα σημαντικότερα και πιο χρονοβόρα βήματα είναι η προεξεργασία των δεδομένων. Στην συγκεκριμένη περίπτωση τονίζεται πως η καταγραφή έγινε με sampling rate 5000 samples/s και το resolution στα 0.1μV. Έπειτα, αφαιρέθηκε το gradient artifact και κάθε σήμα μεταθέρθηκε στα 250 samples/s ενώ εφαρμόστηκε ένα low pass φίλτρο των 40Hz. Έπειτα, διάφορων ειδών artifacts αφαιρέθηκαν με την διαδομένη τεχνική του Independent Component Analysis (ICA). Τέλος, αξίζει να σημειωθεί ότι τα πρώτα 6 δευτερόλεπτα αφαιρέθηκαν, πιο συγκεκριμένα από τα 130.000 σημεία στο σήμα (σε κάθε κανάλι), έγινε επιλογή των πρώτων 50.000 χωρίς κίνηση.

Πριν την εφαρμογή αλγορίθμων είναι πάντα θεμιτό και αρκετά σημαντικό να γίνεται μια εξερεύνηση βασικών στατιστικών στα δεδομένα μας. Έτσι, στο παρακάτω διάγραμμα, μπορεί κανείς να δει 12 θηκογράμματα που αφορούν τα σήματα του πειράματος. Γίνεται αντιληπτό αρκετά γρήγορα, ότι και μόνο οπτικά υπάρχει μια διαφοροποίηση στις δύο ομάδες - τα σήματα των ασθενών με PTSD παρουσιάζουν μικρότερες διακυμάνσεις.

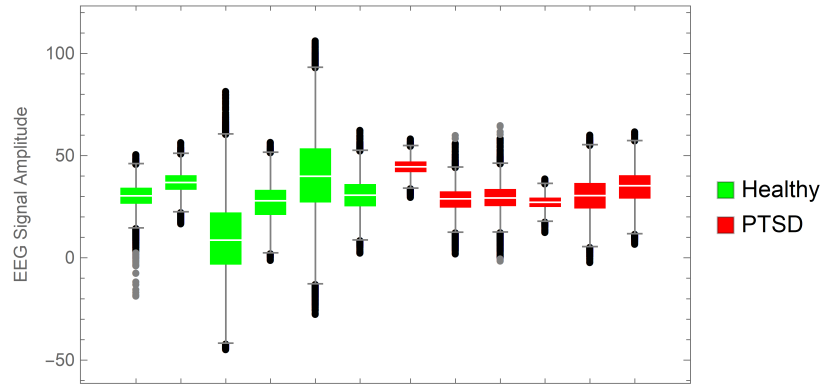


Figure 6.1: Κατανομές σημάτων για *control group* και *PTSD* οσον αφορά το κανάλι *F3*.

Ίσως και το κρίσιμότερο σημείο της ανάλυσης της προαναφερθείσας δημοσίευσης είναι ο έλεγχος της στασιμότητας των δεδομένων. Οι συγγραφείς προτείνουν να μετρηθεί η τυπική απόκλιση των δεδομένων με διαφορετικά κινούμενα υπολογιστικά παράθυρα για την έρευνα του βέλτιστου παραθύρου όπου το σήμα θα συμπεριφέρεται ως στάσιμο. Η παραπάνω τεχνική έχει ως στόχο την έπειτα ανάλυση του σήματος με τον αλγόριθμο R/S Analysis ο οποίος όπως έχουμε μελετήσει παραπάνω χρησιμοποιείται μόνο σε στάσιμες χρονοσειρές. Ακόμη σημαντικό μέρος της ανάλυσης που περιγράφεται στην μεθοδολογία του αλγορίθμου είναι ο μετασχηματισμός του σήματος (log ratio) πρώτου τρέξει ο αλγόριθμος διασφαλίζοντας την παραπάνω υπόθεση.

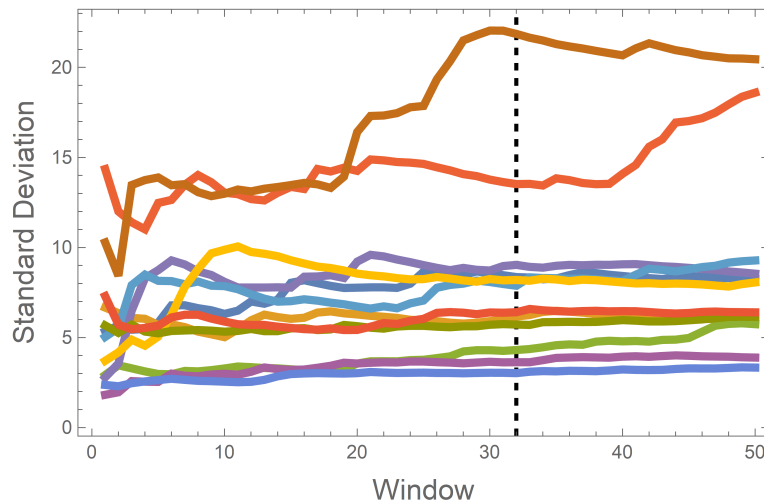


Figure 6.2: Τυπική απόκλιση σημάτων για κάθε υπολογιστικό παράθυρο.

Παρ'όλα αυτά, στην συγκεκριμένη προσέγγιση, η ανάλυση διαφέρει αρκετά. Παρατηρώντας την συμπεριφορά του αλγορίθμου R/S Analysis στα συνθετικά δεδομένα, την βιβλιογραφία και λαμβάνοντας υπόψη την πολυπλοκότητα του ΗΗΓ σήματος ο αλγόριθμος που θα δοκιμαστεί για τον διαχωρισμό των σημάτων είναι αυτός που βασίζεται στον Haar Wavelet. Κάθε σήμα λοιπόν χωρίζεται σε μικρότερα κομμάτια, συγκεκριμένα 1000 τιμές και ο αλγόριθμος τρέχει σε κάθε ένα σετ από αυτές. Αξίζει να σημειωθεί πως το δείγμα για κάθε group είναι 6 σήματα, άρα θα βρεθούν συνολικά 600 Hurst Exponents.

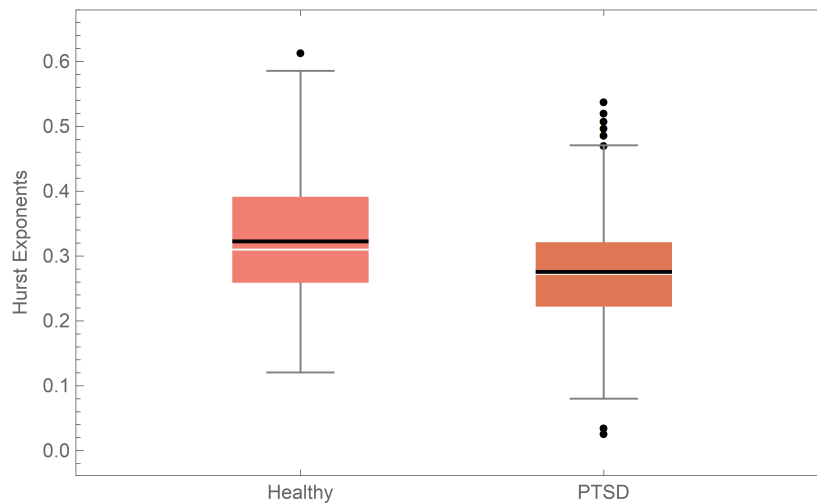


Figure 6.3: Κατανομές εκθέτη Hurst για τα δύο group.

Παρατηρώντας το πάνω θηκόγραμμα μεταξύ των δύο group βλέπουμε πως τα αποτελέσματα συμφωνούν σε κάποιο βαθμό με τους συγγραφείς. Φαίνεται να υπάρχει μια διαφορά στο μέσους όρους (Healthy = 0.33, PTSD = 0.27) αλλά και στις κατανομές των εκθετών H. Τέτοιες διαφορές έχουν ξανά αναφερθεί στη βιβλιογραφία [29] σε σήματα από ασθενείς με μετατραυματικό στρες και πιο συγκεκριμένα σε κανάλια όπως το F3 που εξετάζεται. Οι διαφοροποιήσεις του εκθέτη στην κλίμακα, είναι αρκετό πιθανό να οφείλονται στη διαφορά της διαχείρισης του σήματος αλλά και του αλγορίθμου επιλογής. Όπως αναφέρθηκε και πιο πάνω στην συγκεκριμένη προσέγγιση γίνεται επιλογή ενός διαφορετικού αλγορίθμου από τον R/S Analysis που προτείνεται στην δημοσίευση. Αυτό μας δίνει το πλεονέκτημα μιας καλύτερης προσέγγισης του εκθέτη χωρίς να επέμβουμε με κάποιον μετασχηματισμό στο σήμα. Βιβλιογραφικά, αλλά και πειραματικά στα προηγούμενα κεφάλαια έχει γίνει εκτενής ανφορά στα μειονεκτήματα του R/S Analysis.

Τέλος αξίζει να σημειωθεί ότι τέτοιες διαφορές μεταξύ group σε συγκεκριμένα κανάλια είναι αρκετά σημαντικές για μελλοντικές ιατρικές εφαρμογές. Έτσι είναι αρκετά καίριο να υπάρχουν αλγόριθμοι οι οποίοι με σωστό τρόπο και γρήγορα θα μετράνε στατιστικές ποσότητες οι οποίες θα μπορούν να μας προσφέρουν περισσότερες πληροφορίες

για το σήμα. Σήματα σαν τα EEG χαρακτηρίζονται στη βιβλιογραφία χαοτικά και πολυ-μετάβλητα καθώς η καταγραφή μας προσφέρει πολλές χρονοσειρές ταυτόχρονα οι οποίες δύναται να αλληλεπιδρούν μεταξύ τους.

Κεφάλαιο 7

Συμπεράσματα

Στην συγκεκριμένη πτυχιακή εργασία προτείνεται μια διαφορετική προσέγγιση στην ανάλυση ηλεκτροεγκεφαλογραφικών σημάτων μέσω αλγορίθμων μαρκας διάρκειας. Αφού γίνεται η απαραίτητη εισαγωγή στο θεωρητικό υπόβαθρο όσον αφορά τις χρονοσειρές και διάφορους ντετερμινιστικούς και στοχαστικούς μηχανισμούς που τις παράγουν, περιγράφονται και αξιολογούνται διάφορα μέτρα που έχουν προταθεί κατά καιρούς στη βιβλιογραφία. Εκεί γίνεται για πρώτη φορά μνεία στο χαρακτηριστικό της μαρκάς μνήμης (long memory effect) στα σήματα. Έπειτα, ακολουθεί μια αναλυτική ανασκόπηση σε γνωστούς αλγορίθμους της βιβλιογραφίας, τονίζοντας τα πλεονεκτήματα και τα μειονεκτήματά τους. Μέσα από αυτή την ανάλυση, και τις προσομοιώσεις σε συνθετικά δεδομένα διάφορων διαδικασιών προτείνεται ένας νέος διαφορετικός αλγόριθμος και αξιολογείται.

Για την εφαρμογή τέτοιου είδους αλγορίθμων σε πραγματικά δεδομένα, μια πρόσφατη έρευνα στο χώρο την ηλεκτροεγκεφαλογραφίας επιλέγεται. Στόχος είναι η επεξεργασία των ηλεκτροεγκεφαλογραφικών δεδομένων με σκοπό της αναζήτησης διαφοροποίησης σε δύο ομάδες σημάτων. Για τον σκοπό αυτό επιλέχθηκε να δοκιμαστεί ο αλγόριθμος Haar Wavelet βάση των προσομοιώσεων και των συγκρίσεων στη βιβλιογραφία. Η παραπάνω επιλογή δίνει το πλεονέκτημα ότι ο αλγόριθμος συμπεριφέρεται σωστά και με ακρίβεια σε μη στάσιμες διαδικασίες χωρίς να χρειάζεται να γίνει κάποιος μετασχηματισμός στα δεδομένα.

Παράλληλα, βιβλιογραφικά γνωρίζουμε πως σήματα σαν τα EEG είναι μη στάσιμα [31] και χρησιμοποιώντας έναν αλγόριθμο σαν τον R/S Analysis ο οποίος προτείνεται από τους συγγραφείς της προαναφερθείσας δημοσίευσης ενδεχομένως να οδηγήσει σε εσφαλμένες μετρήσεις/ενδείξεις. Το βασικό συμπέρασμα αυτής της μελέτης είναι ότι υπάρχει μια μικρή αλλά σημαντική διαφορά στον εκθέτη Hurst που σχετίζεται με long memory effects στα σήματα των δύο ομάδων (PTSD και Healthy). Πιο συγκεκριμένα, αυτή η διαφορά εντοπίζεται στο κανάλι F3 ενώ αξίζει να σημειωθεί ότι έχουν βρεθεί αντίστοιχα αποτελέσματα σε σχετικές μελέτες [29] όσον αφορά ασθενείς με PTSD.

Μελλοντικές προεκτάσεις στην παραπάνω μελέτη σίγουρα αποτελεί η βελτίωση του προτεινόμενου αλγορίθμου για τον υπολογισμό του Hurst Exponent. Η χρήση ενός

νευρωνικού δικτύου προχωρημένης αρχιτεκτονικής σίγουρα θα μπορούσε να βελτιώσει την ακρίβεια ενώ θα έδινε μεγαλύτερη ασφάλεια σε ταξινομήσεις μικρών Hurst Exponent στις διαδικασίες FBM. Επίσης, περετέρω διερεύνηση χρήζουν οι εξισώσεις με τις οποίες υπολογίζεται ο εκθέτης Hurst μέσα από τον υπολογισμό της τιμής Abbe. Τέλος, όσον αφορά το πειραματικό στάδιο και την ανάλυση των δεδομένων EEG, ενδιαφέρον θα αποτελούσε ο αλγόριθμος να δοκιμαστεί σε περισσότερα σετ δεδομένων (πχ επιληπτικοί και υγιείς) και αντίστοιχες περιπτώσεις στις οποίες το μέτρο της μαρκάς διάρκειας δύναται να ξεχωρίσει δύο group σήματων παρέχοντας χρήσιμες πληροφορίες για το σήμα.

Bibliography

- [1] <https://askabiologist.asu.edu/brain-regions>.
- [2] Κοκολάκης Γ.Ε. Σημειώσεις στοχαστικών ανελίξεων, Σ.Ε.Μ.Φ.Ε. Ε.Μ.Π.
- [3] Κουγιουμτζής, Δ. Ανάλυση Χρονοσειρών, Σημειώσεις μεταπτυχιακού μαθήματος του τμήματος Μαθηματικών του Α.Π.Θ.
- [4] U Rajendra Acharya, S Vinitha Sree, G Swapna, Roshan Joy Martis, and Jasjit S Suri. Automated eeg analysis of epilepsy: a review. *Knowledge-Based Systems*, 45:147–165, 2013.
- [5] Jose Alvarez-Ramirez, Juan C Echeverria, and Eduardo Rodriguez. Performance of a high-dimensional r/s method for hurst exponent estimation. *Physica A: Statistical Mechanics and its Applications*, 387(26):6452–6462, 2008.
- [6] Ralph G Andrzejak, Klaus Lehnertz, Florian Mormann, Christoph Rieke, Peter David, and Christian E Elger. Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: Dependence on recording region and brain state. *Physical Review E*, 64(6):061907, 2001.
- [7] Duncan AJ Blythe. Long range analysis of eeg data. 2016.
- [8] Alan Geoffrey Brown. *Nerve cells and nervous systems: an introduction to neuroscience*. Springer Science & Business Media, 2001.
- [9] Anna Carbone. Algorithm to estimate the hurst exponent of high-dimensional fractals. *Physical Review E*, 76(5):056703, 2007.
- [10] Jeong-Ho Chae, Jaeseung Jeong, Bradley S Peterson, Dai-Jin Kim, Won-Myong Bahk, Tae-Youn Jun, Soo-Yong Kim, and Kwang-Soo Kim. Dimensional complexity of the eeg in patients with posttraumatic stress disorder. *Psychiatry Research: Neuroimaging*, 131(1):79–89, 2004.
- [11] Chris Chatfield. *The analysis of time series: an introduction*. CRC Press, Florida, US, 6th edition, 2004.

- [12] Boris V Chirikov. A universal instability of many-dimensional oscillator systems. *Physics reports*, 52(5):263–379, 1979.
- [13] James W Cooley and John W Tukey. An algorithm for the machine calculation of complex fourier series. *Mathematics of computation*, 19(90):297–301, 1965.
- [14] Paul Doukhan, George Oppenheim, and Murad Taqqu. *Theory and applications of long-range dependence*. Springer Science & Business Media, 2002.
- [15] G Dumermuth and H Flühler. Some modern aspects in numerical spectrum analysis of multichannel electroencephalographic data. *Medical and biological engineering*, 5(4):319–331, 1967.
- [16] G Dumermuth and L Molinari. Spectral analysis of eeg background activity. *Handbook of electroencephalography and clinical neurophysiology*, 1:85–130, 1987.
- [17] Michael Falk, Frank Marohn, René Michel, Daniel Hofmann, Maria Macke, Bernhard Tewes, and Peter Dinges. A first course on time series analysis: examples with sas. 2006.
- [18] Zafeirios Fountas et al. Spiking neural networks for human-like avatar control in a simulated environment. *Computing Science of Imperial College London*, 2011.
- [19] M Gilmore, CX Yu, TL Rhodes, and WA Peebles. Investigation of rescaled range analysis, the hurst exponent, and long-time correlations in plasma turbulence. *Physics of Plasmas*, 9(4):1312–1317, 2002.
- [20] Naor Granik, Lucien E Weiss, Elias Nehme, Maayan Levin, Michael Chein, Eran Perlson, Yael Roichman, and Yoav Shechtman. Single-particle diffusion characterization by deep learning. *Biophysical journal*, 117(2):185–192, 2019.
- [21] Peter Grassberger and Itamar Procaccia. Measuring the strangeness of strange attractors. In *The Theory of Chaotic Attractors*, pages 170–189. Springer, 2004.
- [22] Timothy Graves, Robert Gramacy, Nicholas Watkins, and Christian Franzke. A brief history of long memory: Hurst, mandelbrot and the road to arfima, 1951–1980. *Entropy*, 19(9):437, 2017.
- [23] John M Greene. A method for determining a stochastic transition. *Journal of Mathematical Physics*, 20(6):1183–1201, 1979.
- [24] B Hjorth. Time domain descriptors and their relation to a particular model for generation of eeg activity. *CEAN-Computerized EEG analysis*, pages 3–8, 1975.
- [25] Bo Hjorth. Eeg analysis based on time domain properties. *Electroencephalography and clinical neurophysiology*, 29(3):306–310, 1970.

- [26] Yvonne Höller, Eugen Trinka, Gudrun Kalss, Günter Schiepek, and Rosa Michaelis. Correlation of eeg spectra, connectivity, and information theoretical biomarkers with psychological states in the epilepsy monitoring unit—a pilot study. *Epilepsy & Behavior*, 99:106485, 2019.
- [27] Herbert H Jasper. The ten-twenty electrode system of the international federation. *Electroencephalogr. Clin. Neurophysiol.*, 10:370–375, 1958.
- [28] MG Kendall. Studies in the history of probability and statistics. xxvi: The work of ernst abbe. *Biometrika*, pages 369–373, 1971.
- [29] Jinho Kim, Jeong-Ho Chae, Hee-kyoung Ko, Charles-Francois Vincent Latchoumane, Aveek Banerjee, Donald J Mandell, Christina W Hoven, and Jae-seung Jeong. Hemispheric asymmetry in non-linear interdependence of eeg in post-traumatic stress disorder. *Psychiatry and clinical neurosciences*, 66(2):87–96, 2012.
- [30] Kyungsik Kim and Seong-Min Yoon. Multifractal features of financial markets. *Physica A: Statistical Mechanics and its Applications*, 344(1-2):272–278, 2004.
- [31] Wlodzimierz Klonowski. Everything you wanted to ask about eeg but were afraid to get the right answer. *Nonlinear biomedical physics*, 3(1):1–5, 2009.
- [32] James N Knight. *Signal fraction analysis and artifact removal in EEG*. PhD thesis, Colorado State University, 2003.
- [33] Salim Lahmiri. Generalized hurst exponent estimates differentiate eeg signals of healthy and epileptic patients. *Physica A: Statistical Mechanics and its Applications*, 490:378–385, 2018.
- [34] Hamze Lotfalinezhad and Ali Maleki. Tta, a new approach to estimate hurst exponent with less estimation error and computational time. *Physica A: Statistical Mechanics and its Applications*, page 124093, 2020.
- [35] Robert S MacKay. A renormalization approach to invariant circles in area-preserving maps. *Physica D: Nonlinear Phenomena*, 7(1-3):283–300, 1983.
- [36] Michael C Mackey and Leon Glass. Oscillation and chaos in physiological control systems. *Science*, 197(4300):287–289, 1977.
- [37] Glen MacLachlan, Ashwin Shenoy, Eda Sonbas, Rob Coyne, Kalvir Dhuga, Ali Eskandarian, Leonard Maximon, and William Parke. The hurst exponent of fermi grbs. *arXiv preprint arXiv:1209.2396*, 2012.
- [38] Ron Mancini. *Op amps for everyone: design reference*. Newnes, 2003.
- [39] Benoit B Mandelbrot. The fractal geometry of. *Nature*, pages 394–397, 1982.

- [40] Benoit B Mandelbrot and Richard L Hudson. *The (mis) behaviour of markets: a fractal view of risk, ruin and reward*. Profile books, 2010.
- [41] Benoit B Mandelbrot and John W Van Ness. Fractional brownian motions, fractional noises and applications. *SIAM review*, 10(4):422–437, 1968.
- [42] Elias Masry. The wavelet transform of stochastic processes with stationary increments and its application to fractional brownian motion. *IEEE Transactions on Information Theory*, 39(1):260–264, 1993.
- [43] Robert M May. Simple mathematical models with very complicated dynamics. *Nature*, 261(5560):459–467, 1976.
- [44] Christoph M Michel, Thomas Koenig, Daniel Brandeis, Jiří Wackermann, and Lorena RR Gianotti. *Electrical neuroimaging*. Cambridge University Press, 2009.
- [45] Jesús Jaime Moreno Escobar, Oswaldo Morales Matamoros, Ricardo Tejeida Padilla, Ixchel Lina Reyes, Liliana Chanona Hernández, and Ana Gabriela Ramírez Gutiérrez. Brain-inspired healthcare smart system based on perception-action cycle. *Applied Sciences*, 10(10):3532, 2020.
- [46] Tamanna TK Munia, Ali Haider, Charles Schneider, Mark Romanick, and Reza Fazel-Rezai. A novel eeg based spectral analysis of persistent brain function alteration in athletes with concussion history. *Scientific reports*, 7(1):1–13, 2017.
- [47] Luis Fernando Nicolas-Alonso and Jaime Gomez-Gil. Brain computer interfaces, a review. *sensors*, 12(2):1211–1279, 2012.
- [48] Heino Bohn Nielsen. Non-stationary time series and unit root tests. *Unpublished Lecture Notes for Econometrics*, 2, 2005.
- [49] Colin P North and David I Halliwell. Bias in estimating fractal dimension with the rescaled-range (r/s) technique. *Mathematical Geology*, 26(5):531–555, 1994.
- [50] Robert Oostenveld and Peter Praamstra. The five percent electrode system for high-resolution eeg and erp measurements. *Clinical neurophysiology*, 112(4):713–719, 2001.
- [51] Athanasios Papoulis and S Unnikrishna Pillai. *Probability, random variables, and stochastic processes*. Tata McGraw-Hill Education, 2002.
- [52] Donald B Percival and Andrew T Walden. *Wavelet methods for time series analysis*, volume 4. Cambridge university press, 2000.
- [53] SC Phatak and S Suresh Rao. Logistic map: A possible random-number generator. *Physical review E*, 51(4):3670, 1995.

- [54] Bo Qian and Khaled Rasheed. Hurst exponent and financial market predictability. In *IASTED conference on Financial Engineering and Applications*, pages 203–209, 2004.
- [55] J. Ross Quinlan. Induction of decision trees. *Machine learning*, 1(1):81–106, 1986.
- [56] Bahareh Rahmani, Chung Ki Wong, Payam Norouzzadeh, Jerzy Bodurka, and Brett McKinney. Dynamical hurst analysis identifies eeg channel differences between ptsd and healthy controls. *PloS one*, 13(7):e0199144, 2018.
- [57] Mariusz Tarnopolski. Distinguishing short and long fermi gamma-ray bursts. *Monthly Notices of the Royal Astronomical Society*, 454(1):1132–1139, 2015.
- [58] Mariusz Tarnopolski. On the relationship between the hurst exponent, the ratio of the mean square successive difference to the variance, and the number of turning points. *Physica A: Statistical Mechanics and its Applications*, 461:662–673, 2016.
- [59] Mariusz Tarnopolski, Natalia Żywucka, Volodymyr Marchenko, and Javier Pascual-Granado. A comprehensive power spectral density analysis of astronomical time series i: the fermi-lat gamma-ray light curves of selected blazars. *arXiv preprint arXiv:2006.03991*, 2020.
- [60] John Von Neumann. Distribution of the ratio of the mean square successive difference to the variance. *The Annals of Mathematical Statistics*, 12(4):367–395, 1941.
- [61] Luyao Wang and Hai Cheng. Pseudo-random number generator based on logistic chaotic system. *Entropy*, 21(10):960, 2019.
- [62] Rachel Yehuda. Post-traumatic stress disorder. *New England journal of medicine*, 346(2):108–114, 2002.
- [63] Luciano Zunino, Felipe Olivares, Aurelio F Bariviera, and Osvaldo A Rosso. A simple and fast representation space for classifying complex time series. *Physics Letters A*, 381(11):1021–1028, 2017.