



UNIVERSITÀ POLITECNICA DELLE MARCHE
FACOLTÀ DI INGEGNERIA

Corso di Laurea in Ingegneria Biomedica
TESI DI LAUREA TRIENNALE

**Analisi della maturazione del cammino tramite
misure non lineari della variabilità del passo.**

**Gait maturation analysis through
non-linear measures of stride
variability.**

Relatore

Prof.ssa FEDERICA VERDINI

Tesi di Laurea di

SEMPLICI CLAUDIA

Correlatori

Ing. ALESSANDRO MENGARELLI

ANDREA TIGRINI

A.A. 2020/2021

Ringraziamenti

Un ringraziamento speciale va a tutte le persone che hanno fatto parte di questo meraviglioso viaggio.

Ringrazio mia madre per avermi insegnato l'importanza della responsabilità e dell'impegno quotidiano, non solo nello studio ma anche nella vita.

Ringrazio mio padre per la sua estrema bontà e per non aver mai dubitato delle mie capacità.

Grazie alle mie amiche.

A Silvia, per avermi fatto guardare il mondo con occhi diversi e aver abbattuto insieme a me alcuni miei limiti.

A Chiara, per dimostrarmi ogni giorno il valore della vera amicizia che dura nel tempo ed è incondizionata.

A Natalia, per avermi capito in ogni situazione e avermi aiutato sempre senza voler ricevere nulla in cambio.

Grazie alle mie compagne di università.

A Costanza, per la sua estrema disponibilità e per avermi accompagnato nell'ultimo anno mano nella mano.

A Veronica, per avermi trasmesso la sua forza, la sua sincerità e il suo spirito di intraprendenza.

Grazie alla Claudia del passato e buona fortuna alla Claudia del futuro.

Le auguro di trovare la felicità che ha sempre cercato.

SOMMARIO

L'obiettivo di questa tesi è quello di esaminare la dinamica del cammino a partire da dati cinematici di individui e come essa vari in funzione della maturazione con l'età. Per comprendere l'approccio su cui si basa la presente indagine, vanno considerati gli studi pubblicati negli ultimi decenni. Da questi si evince quali siano le proprietà significative di un intervallo di stride e come esse varino nel caso in cui i soggetti analizzati siano sani o affetti da malattie neurodegenerative. Nel presente studio, si analizzeranno serie temporali composte da intervalli di stride di bambini sani con metodi non-lineari. Mediante la valutazione di tali segnali, si estrapoleranno dei parametri caratteristici dell'analisi stride-to-stride che ci daranno informazioni riguardo la variabilità, la complessità e la regolarità del cammino in funzione dell'aumento di età. Si divideranno i soggetti in tre gruppi in ordine crescente di età e si osserverà quanto questi siano più o meno caratterizzabili tra loro in termini di regolarità del cammino. I metodi non-lineari impiegati sono la Detrended Fluctuations Analysis (DFA) e la Recurrence Quantification Analysis (RQA). La prima ci permette di ricavare un indice di scala frattale che dipende dalle fluttuazioni degli intervalli di stride e che indica la correlazione più o meno marcata delle serie temporali. Il secondo tipo di analisi, invece, è basata sull'individuazione di parametri che quantificano la regolarità delle serie temporali. Tali parametri vengono estrapolati da Recurrence Plots, ovvero grafici che stimano la ricorrenza di determinati comportamenti nelle serie temporali. Tramite tali metodi, si esaminerà la distribuzione dei parametri per ogni gruppo in relazione agli altri. Infine, si confronteranno i risultati ottenuti con altri studi presenti in letteratura.

INDICE DELLE TABELLE E DELLE FIGURE

Figura 1 - Stride interval rappresentativo prima (A) e dopo (B) l'operazione di shuffling.....	8
Figura 2 - Risultato della DFA per le serie originali e ordinate in maniera casuale.....	9
Figura 3 - Dinamica del ritmo del passo di un paziente con PD.....	10
Figura 4 - PCA dei parametri selezionati per il gruppo dei sani (sinistra) e patologici (destra). I colori freddi indicano la bassa influenza dei parametri nelle serie temporali, i colori caldi indicano alta influenza.....	14
Figura 5 - Analisi MSE durante un'andatura incontrollata. Confronto tra le serie originali e quelle surrogate.....	17
Figura 6 - Analisi mediante MSE di serie temporali regolate da metronomo. Confronto tra serie originali e surrogate.....	18
Figura 7 - Confronto tra serie temporali di un soggetto sano di 78 anni (a) e uno affetto da Parkinson di 68 anni (b).....	19
Figura 8 - Grafici di ApEn (a), NSE (b) e STC (c) per soggetti sani (CO) e pazienti con Parkinson (PD).....	20
Figura 9 - Esempio di Recurrence Plot.....	33
Tabella 1 - Valori medi \pm SD degli esponenti di scala frattale divisi nei tre gruppi in base all'età.....	39
Figura 10 - Distribuzione dei valori di esponente di scala frattale secondo i tre gruppi di età differente.....	39
Figura 11 - Distribuzione dei valori di esponente di scala frattale rispetto al gruppo 1 e al gruppo 3.....	40
Tabella 2 - Valori medi \pm SD del parametro RR secondo la divisione in tre gruppi.....	42
Figura 12 - Distribuzione del parametro RR in funzione dei tre gruppi.....	42
Tabella 3 - Valori medi \pm SD del parametro DET secondo la divisione in tre gruppi.....	43
Figura 13 - Distribuzione del parametro DET in funzione dei tre gruppi.....	43
Tabella 4 - Valori medi \pm SD del parametro ENT secondo la divisione in tre gruppi.....	44
Figura 14 - Distribuzione del parametro ENT in funzione dei tre gruppi.....	45
Tabella 5 - Valori medi \pm SD del parametro MLL secondo la divisione in tre	

gruppi.....	45
Figura 15 - Distribuzione del parametro MLL in funzione dei tre gruppi.....	45
Tabella 6 - Valori medi \pm SD del parametro LAM secondo la divisione in tre gruppi.....	47
Figura 16 - Distribuzione dei valori del parametro LAM in funzione dei tre gruppi.....	47
Figura 17 - Distribuzione del parametro LAM del primo e del terzo gruppo.....	48
Tabella 7 - Valori di p-value registrati confrontando i valori dell'esponente di scala frattale per i tre gruppi.....	50
Tabella 8 - Valori di p-value registrati confrontando i valori di RR per i tre gruppi.....	50
Tabella 9 - Valori di p-value registrati confrontando i valori di DET per i tre gruppi.....	50
Tabella 10 - Valori di p-value registrati confrontando i valori di ENT per i tre gruppi.....	51
Tabella 11 - Valori di p-value registrati confrontando i valori di MLL per i tre gruppi.....	51
Tabella 12 - Valori di p-value registrati confrontando i valori di LAM per i tre gruppi.....	51
Tabella 13 - Confronto della divisione in gruppi tra la ricerca presente e studio di V. Alcan.....	53
Tabella 14 - Risultati ottenuti dallo studio di V. Alcan secondo gli algoritmi ApEn e SampEn.....	54
Tabella 15 - Risultati ottenuti dallo studio di V. Alcan secondo l'algoritmo DistEn.....	55

INDICE

RINGRAZIAMENTI

SOMMARIO

INDICE DELLE TABELLE E DELLE FIGURE

INDICE

CAPITOLO 1 : INTRODUZIONE.....	7
CAPITOLO 2 : MATERIALI E METODI.....	21
2.1 Detrended Fluctuations Analysis (DFA).....	23
2.2 Metodi fondati sulla misura dell'entropia.....	26
2.2.1 Approximate Entropy (ApEn).....	26
2.2.2 Sample Entropy (SampEn).....	28
2.2.3 Distribution Entropy (DistEn).....	30
2.2.4 Multiscale Entropy Analysis (MSE).....	32
2.3 Recurrence Quantification Analysis (RQA).....	33
2.4 Materiali.....	37
CAPITOLO 3 : RISULTATI OTTENUTI.....	38
3.1 Analisi mediante DFA.....	38
3.2 Analisi mediante RQA.....	41
3.3 Test statistici.....	49
CAPITOLO 4 : DISCUSSIONE.....	53
CAPITOLO 5 : CONCLUSIONE.....	58
BIBLIOGRAFIA.....	60

CAPITOLO 1

Introduzione

Alla base dell'analisi del cammino, c'è la consapevolezza del fatto che l'andatura dell'uomo è un processo estremamente complesso e caratterizzato da proprietà ben precise, definite frattali. In particolare, il termine frattale indica un insieme, apparentemente caotico, costituito in realtà da proprietà autosimilari che si ripetono all'infinito in scala sempre più ridotta [8]. Quando si vuole esaminare questo tipo di processi, si fa riferimento ad un intervallo di tempo che indica la durata di un ciclo del passo chiamato "stride interval". Un intervallo di stride è un intervallo di tempo che intercorre tra due appoggi consecutivi del calcagno dello stesso piede. In uno studio del 1995 [1], si pone l'attenzione sulle caratteristiche dell'intervallo di stride. Si nota, infatti, che tale lasso di tempo presenta delle fluttuazioni di segnale in condizioni di attività muscolare, ovvero nel momento in cui il soggetto è chiamato a compiere il gesto di stride. Una possibile iniziale ipotesi della presenza di tali fluttuazioni è che sul segnale sia presente rumore bianco e quindi il processo è non correlato [1],[5]. Secondo un'altra linea interpretativa, la serie temporale potrebbe presentare una correlazione a corto raggio, ovvero che ogni intervallo di stride sia influenzato solamente dal più recente [1]. Infine, una possibilità meno

intuitiva è che tali fluttuazioni esprimono una correlazione a lungo termine tra gli intervalli temporali. In questo caso, l'intervallo di stride dipenderà in

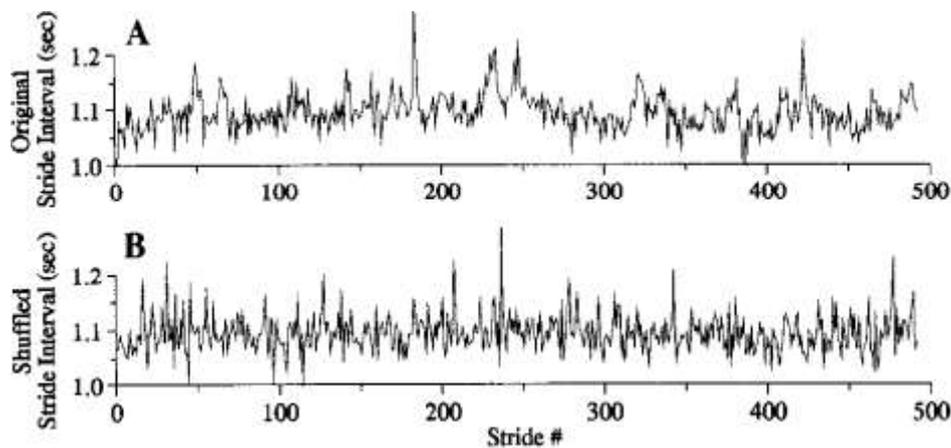


Figura 1. Stride interval rappresentativo prima (A) e dopo (B) l'operazione di shuffling.

ogni istante dagli intervalli precedenti e tale dipendenza decadrà in maniera frattale, ovvero gli intervalli saranno

caratterizzati da proprietà simili tra loro che si ripetono in scala sempre minore [8]. Per dimostrare la validità di quest'ultima ipotesi, Hausdorff [1] ha esaminato degli intervalli di stride di dieci uomini sani di età pari a 26 anni. Nella figura 1 A è presente un intervallo di stride rappresentativo e possiamo notare la sua stabilità, in quanto si è calcolato [1] un coefficiente di variazione è pari al 4%, ma allo stesso tempo osserviamo una struttura complessa. Tale struttura però varia considerevolmente quando la sequenza dei campioni viene riordinata in maniera casuale, come si evince dalla figura 1 B. Il metodo utilizzato per analizzare tali segnali è la Detrended Fluctuations Analysis (DFA) che ci permette di stimare le fluttuazioni F attorno al valore medio dell'intervallo di stride. L'entità della fluttuazione F viene calcolata integrando la serie temporale spostata del suo valore medio e divisa in finestre di osservazione di lunghezza n . La figura 2 mostra il risultato ottenuto dalla DFA, in particolare essa indica la pendenza del grafico della serie temporale originale e riordinata in maniera casuale. Il

valore della pendenza del grafico fornisce il valore dell'indice di scala frattale α il quale ci permette di trarre delle considerazioni. Si osserva quindi che la pendenza della linea ottenuta mettendo in relazione $\log_{10}F(n)$ e $\log_{10}n$, è pari a 0.83 per le serie temporali originali mentre per le serie temporali mescolate casualmente è di 0.5. Per confermare l'ipotesi iniziale di correlazione a lungo raggio delle serie temporali, l'esponente α deve avere un valore compreso tra 0.5 e 1. Se al contrario, α ha un valore pari a 0.5 significa che il segnale analizzato è casuale e non correlato. [1], [2]

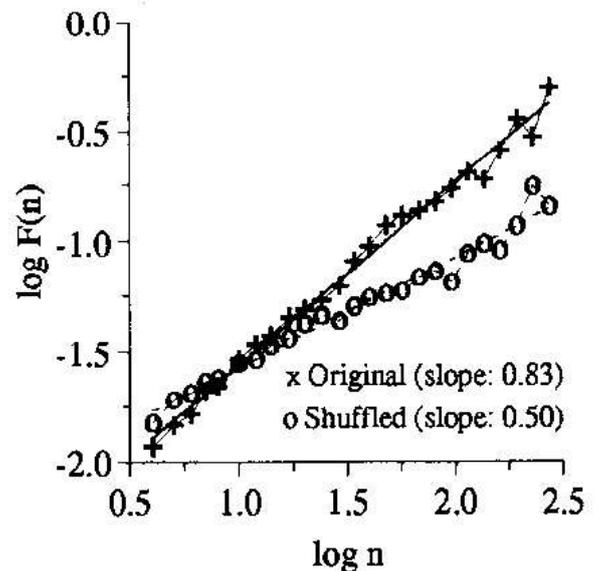


Figura 2. Risultato della DFA per le serie originali e ordinate in maniera casuale

Per i dieci soggetti considerati per effettuare l'esperimento, Hausdorff [1] ha ottenuto un valore di indice di scala frattale di 0.76 ± 0.11 per le serie temporali originali, mentre un valore pari a 0.50 ± 0.03 per le serie temporali ordinate casualmente. Questi risultati confermano l'ipotesi di correlazione a lungo termine tra gli intervalli temporali originali, mentre le serie temporali ordinate casualmente hanno la caratteristica di essere casuali e non correlate.

Spostando l'attenzione su individui affetti da malattie neurodegenerative, è stata svolta lo stesso tipo di analisi descritta precedentemente. Ancora lo stesso Hausdorff [2] si concentra sullo studio delle caratteristiche frattali del

passo in soggetti affetti da Parkinson e come esse variano sotto certe condizioni. Come si può facilmente dedurre, in pazienti con malattia di Parkinson (Parkinson Disease, PD) la variabilità del cammino è maggiore rispetto a quella dei controlli. Questo avviene perché i soggetti affetti da PD hanno un deficit del controllo del passo e ridotte capacità motorie, perciò tendono a non avere un'andatura stabile. Per analizzare le andature di pazienti con Parkinson si fa riferimento alla proprietà di correlazione a lungo raggio delle serie temporali, come visto precedentemente. Le serie temporali sono analizzate ancora una volta mediante DFA, in modo da discriminare le serie temporali di soggetti sani da quelle di pazienti patologici. Più nello specifico, il risultato ottenuto da questa analisi è che i soggetti sani hanno un esponente di scala frattale compreso tra 0.8 e 1, mentre quello dei

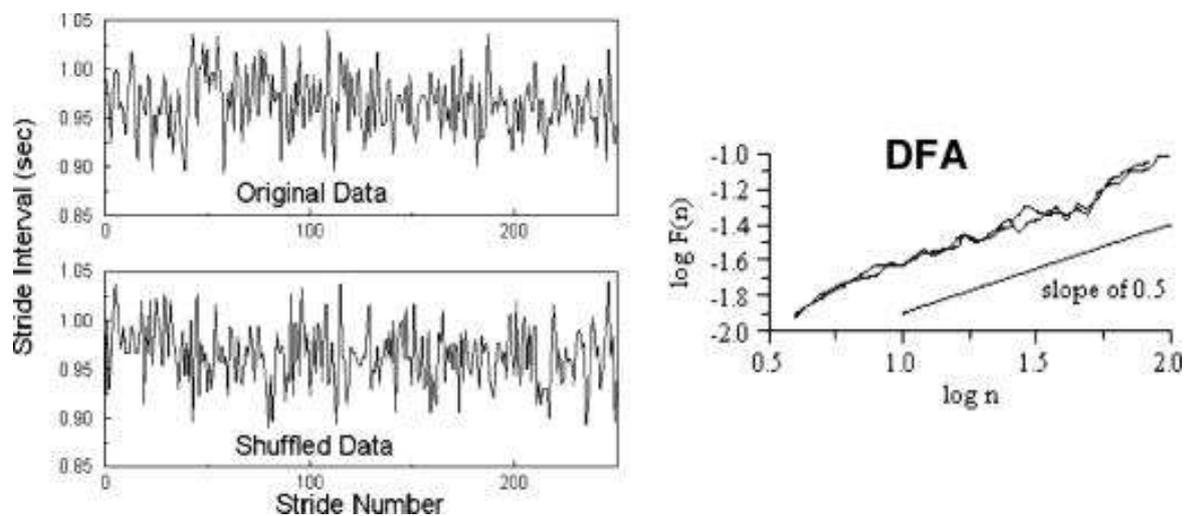


Figura 3. Dinamica del ritmo del passo di un paziente con PD

patologici è molto più vicino a 0.5. Una semplice spiegazione di questa conclusione è che nei pazienti con PD, il passo perde la sua automaticità e fluidità ed è come se stride, e di conseguenza la sua durata, fosse un processo

a sé stante. Un'ulteriore conferma di questa conclusione è riportata nella figura 3, dove si può notare che la serie temporale originale di un paziente con PD non è considerevolmente differente dalla serie temporale riordinata in maniera random ($\alpha = 0.5$ come visto nello studio precedente). Ciò è evidenziato anche dall'analisi mediante DFA.

Per meglio comprendere la relazione tra variabilità del passo e caratteristiche frattali dello stesso in pazienti con PD, Hausdorff [2] esamina gli effetti di diverse condizioni e interventi durante l'andatura di questi pazienti per osservare come ognuno reagisce sotto stati differenti. In particolare, si analizzano gli effetti causati da stimolazione auditiva ritmata, da regolazione mediante tapis roulant, da dual tasking e da somministrazione farmacologica (metilfenidato).

La stimolazione auditiva ritmata è un mezzo molto utile per migliorare l'andatura in pazienti affetti da PD. In particolare, utilizzando un metronomo, la velocità del passo e la lunghezza dell'intervallo di stride aumentano avvicinandosi a quelli caratteristici dei soggetti non-patologici. Ciò avviene in maniera considerevole nel momento in cui tale stimolazione è più veloce rispetto al ritmo usuale del passo o viene utilizzata per un periodo più lungo di tempo (tre settimane). Sorprendentemente, ciò che non cambia in maniera considerevole è l'indice di scala frattale il quale è indifferente a tale tipo di stimolazione. Per quanto riguarda invece la regolazione dell'andatura mediante tapis roulant, si nota che essa riduce la variabilità dell'intervallo di stride e dell'oscillazione durante tale intervallo, sia in pazienti sani che in soggetti affetti da PD. Nonostante tale miglioramento della variabilità, l'indice di scala frattale non varia in

maniera considerevole in pazienti malati come nel caso della stimolazione mediante metronomo. Questi risultati portano alla conclusione che il tapis roulant può essere utilizzato come pacemaker per scandire il ritmo degli individui in modo che essi possano avere un'andatura più stabile e meno variabile. Tuttavia, la velocità del passo fissata dal tapis roulant non condiziona positivamente l'esponente frattale. Ciò dimostra che la variabilità, la velocità e le correlazioni frattali del cammino hanno una natura indipendente dalle stimolazioni esterne. La situazione cambia nel momento in cui il paziente affetto da Parkinson è costretto a svolgere un compito mentre è intento a camminare. In adulti sani e giovani si ha una diminuzione della velocità del passo, ma la variabilità del passo non è affetta in maniera rilevante. In individui sani ma più anziani è stata osservata invece una riduzione della velocità del passo, un aumento del tempo di supporto ma la variabilità dell'andatura non viene influenzata particolarmente. Al contrario in risposta a dual tasking, i pazienti con PD camminano molto più lentamente e, inoltre, si ha un aumento dell'asimmetria tra destra e sinistra e una riduzione della coordinazione. Ne consegue che l'indice di scala frattale diventi sempre più vicino a 0.5 e quindi il processo è da considerarsi più random e simile a rumore bianco. Infine, quando si interviene tramite un medicinale adibito a contrastare la malattia come il metilfenidato, si osserva un miglioramento delle capacità motorie in pazienti con Parkinson, in particolare la velocità del passo aumenta e la variabilità del tempo di stride diventerà più regolare. L'indice di scala, infatti, diventerà più alto.

Considerando nuovamente la caratteristica di variabilità del cammino e come essa varia rispetto al manifestarsi di malattia neurodegenerativa, si possono confrontare gli effetti di diverse malattie confrontando serie temporali di adulti sani e di pazienti patologici. In particolare, F. Dierick e altri [3] fanno riferimento oltre alla malattia del Parkinson, alla sindrome di Huntington e alla sclerosi laterale amiotrofica (SLA). Per analizzare segnali di soggetti affetti da queste tre malattie hanno valutato tre parametri fondamentali, ovvero il coefficiente di variazione (CV), l'Hurst exponent (H) calcolato mediante DFA, e la dimensione frattale (D). Quest'ultimo parametro può variare tra 1 e 2 e più D è vicino a 2, più le oscillazioni dell'intervallo di stride sono elevate tra un ciclo e l'altro quindi il segnale sarà più variabile. Più specificatamente, sono esaminate le correlazioni (positive o negative) di questi tre parametri rispetto a varie condizioni come l'età e le patologie considerate precedentemente. I risultati ottenuti dallo studio svolto da Dierick portano a concludere che lo stride interval (SI) e il CV sono significativamente correlati con l'età nei soggetti sani, mentre in individui patologici non si ha una forte correlazione con l'età. Al contrario, quando si confrontano le scale di disabilità in soggetti con Parkinson e sindrome di Huntington, si nota un legame significativo con il coefficiente di variazione (p-value di circa 0.01/0.02). Inoltre, si nota che in soggetti sani e soggetti affetti da SLA l'intervallo di stride SI ha una correlazione positiva rispetto all'età, ma in pazienti con SLA tale lasso di tempo è considerevolmente più lungo e l'inizio del ciclo di stride è spostato di 0.214 secondi in avanti rispetto all'inizio del ciclo in soggetti sani. Per quanto riguarda invece i parametri H e D, si osserva una forte correlazione tra il

parametro H e l'età nei soggetti sani, mentre il parametro D è correlato in maniera evidente con la scala di disabilità TFC (Total Functional Capacity Score) del gruppo affetto da Huntington. Si può concludere perciò che il gruppo che mostra meno correlazioni con gli altri parametri e le relative condizioni è quello di pazienti soggetti da sindrome laterale amiotrofica. Tramite Principal Component Analysis (PCA) è stato rilevato [3] che l'insieme dei parametri utilizzati (SI, CV, D e H) mantiene il 70.1% della variabilità totale nel caso del gruppo che comprende individui sani. Inoltre, gli indici standard (SI e CV) e quelli frattali (D e H) sono non correlati tra loro, dimostrando l'importanza di includere entrambi i tipi di indici (standard e frattali) quando si vogliono caratterizzare questo tipo di analisi serie temporali. Si è arrivato ad analoghe conclusioni tramite la PCA del gruppo formato da soggetti patologici. Si nota che CV e SI sono

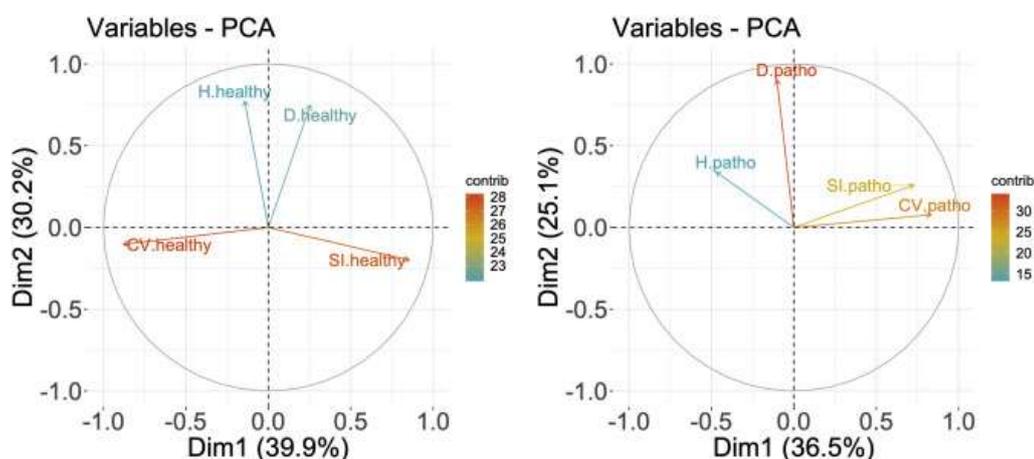


Figura 4. PCA dei parametri selezionati per il gruppo dei sani (sinistra) e patologici (destra). I colori freddi indicano la bassa influenza dei parametri nelle serie temporali, i colori caldi indicano alta influenza.

negativamente correlati nel “gruppo sano”, mentre sono positivamente

correlati nel “gruppo patologico”. Ciò che è importante e che si nota dalla figura 4, è che il parametro D è molto importante per caratterizzare il gruppo degli individui patologici. Infatti, il valore di D in soggetti affetti da malattie neurodegenerative diminuisce poiché le serie temporali sono molto meno complesse rispetto a quelle dei controlli sani proprio perché risultano avere caratteristiche più randomiche. Il parametro H , invece, non ci aiuta a caratterizzare sufficientemente entrambi i due gruppi, malgrado il valore di H sia minore in soggetti patologici, come visto in precedenza.

Dopo aver compreso questo tipo di approccio fondato sull’analisi di serie temporali mediante caratteristiche di tipo frattale, possiamo spostare la nostra attenzione su un altro tipo di metodo utilizzato per discriminare determinati fenomeni che si verificano sotto certe condizioni. In particolare, i parametri utilizzati per esaminare i segnali saranno parametri che derivano dalla caratterizzazione mediante entropia. Come descritto precedentemente, i metodi di analisi non-lineari permettono di studiare serie temporali del cammino sia per soggetti sani sia per soggetti patologici anche nel caso di parametri entropici. Nel 2003 M. Costa e i suoi collaboratori [5] hanno svolto uno studio fondato su un metodo chiamato Multiscale Entropy Analysis. Costa esamina delle serie temporali di dieci soggetti sani di età compresa tra i 18 e i 29 anni sotto diverse condizioni. Ai soggetti è stato chiesto di camminare per un’ora alla loro andatura usuale, lentamente e velocemente. Successivamente, sono stati costretti a camminare seguendo il ritmo di un metronomo alla loro andatura, lentamente e velocemente di nuovo. Dapprima si è applicato il metodo Multiscale Sample Entropy

(MSE) per confrontare la complessità del rumore bianco e del rumore $1/f$. In particolare, viene calcolato il valore di Sample Entropy (SampEn) che misura il grado di complessità di una serie temporale, e si osserva che tale valore diminuisce esponenzialmente all'aumentare del fattore di scala nel caso del rumore bianco, mentre nel caso del rumore $1/f$ il valore entropico rimane costante. Ciò ci porta a concludere che avendo il rumore $1/f$ una struttura ben definita a tutte le scale, avrà una struttura più complessa. Dai segnali raccolti dai dieci soggetti sani, sono state ottenute 20 serie temporali, di cui 10 sono quelle originali mentre le rimanenti 10 sono ottenute tramite l'operazione di shuffling. Si calcolano le curve MSE per tutte le 20 serie e poi, per ogni scala, si ottiene il valore medio di entropia più o meno il valore di deviazione standard ($\pm SD$). Successivamente, si calcola l'area compresa tra le curve inferiore (valore medio - SD) e superiore (valore medio + SD) in modo da ricavare un valore δ che ci permette di confrontare le curve ottenute dall'analisi dell'andatura a ritmo usuale, lento e veloce. Se l'area tra le curve è minore di δ , le curve non saranno considerate significativamente differenti. Nello studio svolto da Costa [5] si arriva alla conclusione che, nel caso di andatura non costretta da metronomo, l'entropia delle serie temporali dei 10 individui sani ordinate casualmente diminuisce con l'aumentare del fattore di scala e hanno quindi un comportamento analogo a quello del rumore bianco. Al contrario, l'entropia delle serie temporali fisiologiche si stabilizza tra le scale 2 e 4 nel caso di andatura usuale del soggetto, e tra le scale 3 e 5 nel caso di andatura lenta e veloce. Inoltre, per scale maggiori il valore di entropia delle serie temporali originali nelle tre condizioni è maggiore rispetto al valore entropico delle

corrispondenti serie mescolate casualmente. Il valore δ che misura l'area tra le serie originali e surrogate per le velocità normale, lenta e veloce è di 1.18,

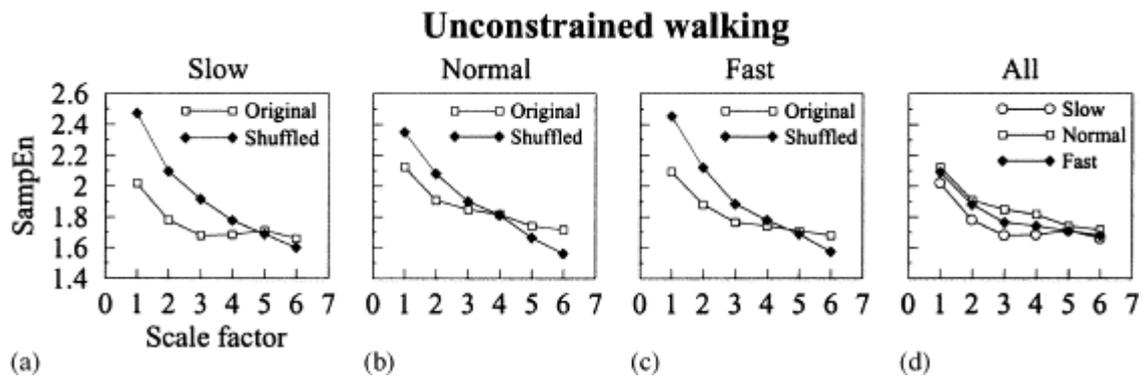


Figura 5. Analisi MSE durante un'andatura incontrollata. Confronto tra le serie originali e quelle surrogate.

0.68 e 0.89 rispettivamente. Questo risultato mostra che le serie temporali fisiologiche sono molto più complesse rispetto a quelle surrogate. Confrontando tutte e tre le curve, come in figura 5d, possiamo notare che l'andatura non controllata e usuale del soggetto in questione presenta una dinamica molto più complessa rispetto a quelle lenta e veloce. Se, invece, consideriamo un passo regolato da metronomo non si ha una netta differenza tra le serie temporali fisiologiche e surrogate ma saranno pressoché simili in tutte e tre le condizioni. Infatti, le curve ottenute mediante MSE decrescono all'aumentare del fattore di scala similmente al rumore bianco precedentemente esaminato. I valori di δ saranno minimi e ciò significa che le curve non saranno molto diverse tra loro, ma saranno considerate random. Dato che camminare seguendo il ritmo di un metronomo ha l'effetto di limitare la cadenza sovra-spinale durante l'andatura, questo risultato indica che i meccanismi di controllo al di sopra della corda spinale sono essenziali

quando si parla della struttura complessa delle serie temporali di soggetti che camminano liberamente.

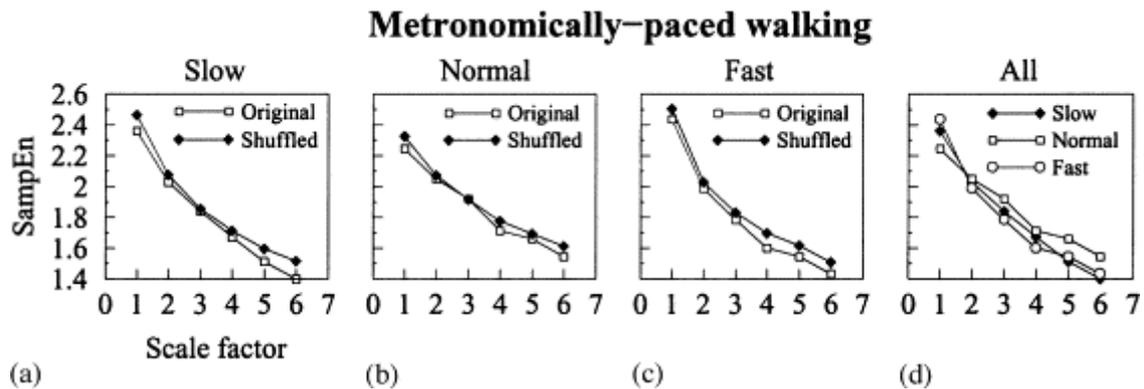


Figura 6. Analisi mediante MSE di serie temporali regolate da metronomo. Confronto tra serie originali e surrogate.

Facendo un confronto tra il metodo applicato da Costa nell'articolo descritto e i risultati ottenuti mediante DFA applicata alle serie temporali a differenti ritmi, si ottiene un coefficiente di correlazione tra i due metodi minore di 0.46. Ciò indica che i due metodi, MSE e DFA, non sono connessi tra loro e quindi evidenziano caratteristiche diverse delle serie temporali prese in considerazione. [5]

Per quanto riguarda soggetti malati di Parkinson, si fa riferimento a diversi parametri entropici. Nell'indagine svolta da Yunfeng Wu [4], si utilizzano

- Approximate Entropy (ApEn), il cui valore indica il grado di regolarità delle serie temporali. Un basso valore di ApEn è indice di un alto grado di regolarità.
- Normalized Symbolic Entropy (NSE), un ulteriore parametro che valuta la regolarità delle serie temporali.

- Signal Turns Count (STC), il quale specifica il numero di variazioni di ampiezza significativi.

Partendo dal confronto tra la serie temporale di un soggetto sano e la serie temporale di un altro soggetto affetto da PD, si osserva che il numero di

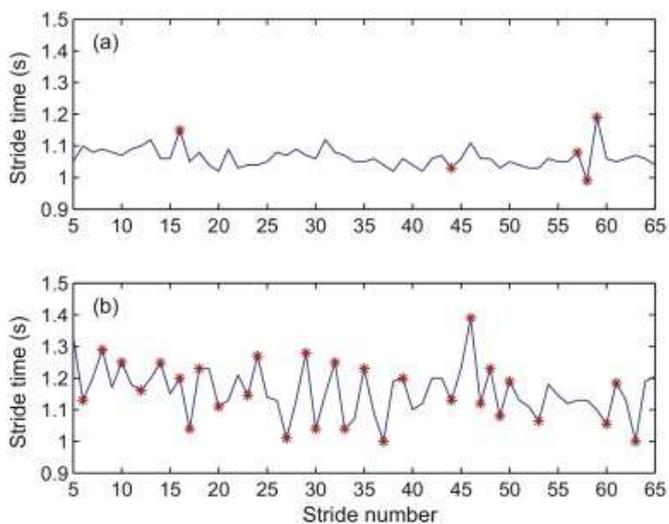


Figura 7. Confronto tra serie temporali di un soggetto sano di 78 anni (a) e uno affetto da Parkinson di 68 anni (b).

STC è estremamente maggiore nel secondo caso, come si evince dalla figura 7. In particolare, la figura 7 raffigura tramite asterischi rossi i cambi di segnale significativi di ampiezza al di sopra degli 0.05 s.

Per quanto riguarda, invece, l'analisi dei parametri dal punto di vista statistico sono state fatte diverse considerazioni. Il valore medio di ApEn nel caso di pazienti con Parkinson è maggiore di 1.3 volte rispetto al valore riscontrato nel gruppo dei controlli. Ciò implica che le serie temporali dei soggetti patologici sono caratterizzate da maggiore irregolarità. Alla stessa conclusione si arriva considerando il valore di NSE, il quale risulta maggiore esaminando i soggetti patologici. Il risultato fornito dal parametro NSE mostra come gli intervalli stride-to-stride degli individui con PD abbiano una dinamica con un grado maggiore

di irregolarità rispetto ai controlli. Per quanto riguarda il parametro STC, si avranno cambi molto più grandi di ampiezza del segnale nei pazienti con patologie, ovvero molte più fluttuazioni. Peraltro, i tre parametri descritti sono strettamente correlati tra di loro ($p\text{-value} < 0.01$) nel caso di pazienti

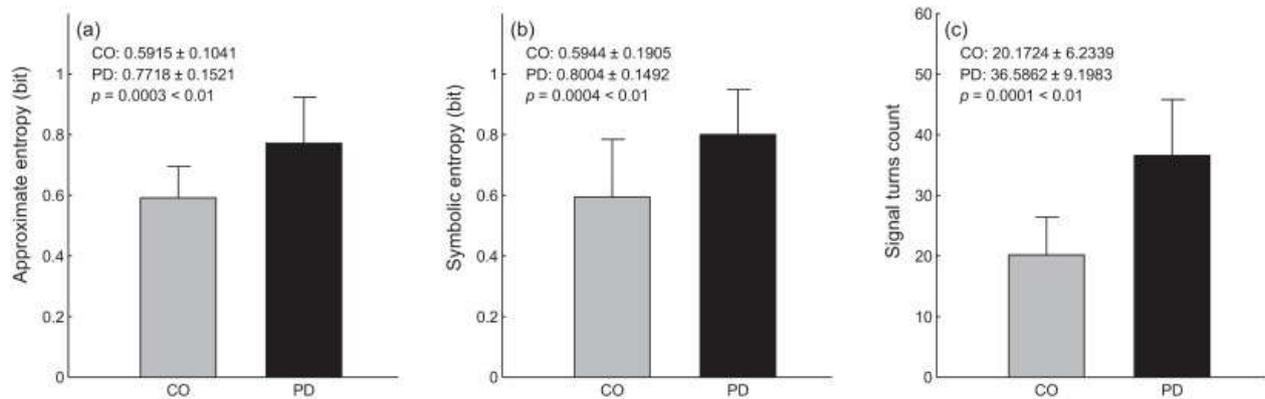


Figura 8. Grafici di ApEn (a), NSE (b) e STC (c) per soggetti sani (CO) e pazienti con Parkinson (PD).

con Parkinson e permettono di discriminare nettamente i due gruppi (controlli e parkinsoniani). I soggetti patologici, quindi, avranno una successione di stride intervals caratterizzata da una variabilità maggiore o strutture complesse.

A seguito dei riferimenti appena esaminati, nel presente studio si pone l'attenzione sulla correlazione che intercorre tra gli intervalli di stride e l'età. La maturazione in base all'età sarà il metro di giudizio mediante il quale si tenterà di caratterizzare gruppi di bambini sani in termini di complessità e regolarità delle serie temporali. L'obiettivo è quello di determinare se siano presenti differenze significative tra le serie temporali dei tre gruppi divisi per fasce di età.

CAPITOLO 2

Materiali e metodi

Quando si parla di sistema dinamico si intende un sistema il cui comportamento varia nel tempo. Tutto ciò che concerne il movimento umano può essere considerato come un sistema di tipo dinamico. Per studiare l'evoluzione del movimento, si è fatto uso per molto tempo di metodi biomeccanici tradizionali basati su relazioni matematiche lineari. Più nello specifico, tali metodi lineari interpretano le strutture dei dati raccolti attraverso correlazioni lineari. L'ipotesi che si ha alla base è che la dinamica intrinseca di un sistema è regolata dal fatto che per piccoli cambiamenti di segnale si avranno piccoli effetti. Le equazioni di tipo lineare, infatti, portano ad una soluzione tale per cui lo stato del sistema può solamente decadere, incrementare o mantenersi costante. Come è illustrato nell'analisi svolta da Richard E.A. van Emmerik e altri [9], questo tipo di analisi ha dei grossi limiti in quanto non permette di esaminare la dinamica di un sistema in maniera completa. Negli ultimi decenni, molti ricercatori, infatti, hanno compiuto passi in avanti per trovare dei metodi adeguati a meglio caratterizzare la dinamica del movimento. Alla base di questi metodi, vi è la consapevolezza che il movimento umano non è caratterizzato da processi ben determinati bensì da comportamenti talvolta stocastici. Ciò

significa che un sistema dinamico è caratterizzato da processi non-lineari ed è per tale motivo che per meglio caratterizzare questi aspetti è utile servirsi di analisi di tipo non-lineare. In particolare, tramite tali metodi si vuole studiare come i sistemi dinamici

1. mantengono il loro stato corrente esaminando la stabilità del sistema stesso. In generale, per stabilità si intende la capacità di mantenere la propria condizione a fronte di sollecitazioni. Quando, invece, si parla di controllo posturale, il termine stabilità implica il mantenimento della posizione eretta in risposta a perturbazioni esterne e inaspettate o generate dall'individuo stesso.
2. transitano tra uno stato e l'altro in termini di variabilità e adattabilità. La variabilità è definita da un punto di vista statistico ed è espressa come la varianza (o la radice quadrata della varianza ovvero la deviazione standard) attorno al valore medio. Per adattabilità, invece, si intende la capacità di un sistema di passare da uno stato all'altro studiando le fluttuazioni scaturite da questo passaggio.
3. regolano la complessità o dinamica frattale caratterizzata da interazioni tra scale spazio-temporali diverse. La complessità di un sistema viene studiata considerando le correlazioni a lungo termine in un determinato segnale. In altre parole, ciò significa che un segnale presenta proprietà simili e strutture correlate tra loro anche a scale differenti.

Più nello specifico, si valutano dei segnali che identificano l'andamento temporale del cammino, ovvero si può analizzare proprio l'intervallo di stride. Tramite questo intervallo di tempo verranno estrapolati dei parametri dai quali trarre delle informazioni in termini di stabilità, variabilità, adattabilità e complessità.

Di seguito verranno descritti in maniera dettagliata alcuni dei metodi non-lineari impiegati durante l'analisi del cammino.

2.1 Detrended Fluctuations Analysis (DFA)

Quando si vuole analizzare la variabilità del passo, ci si può avvalere di indici che sono strettamente correlati al valore medio dell'intervallo di stride. In particolare, due parametri che quantificano tale media sono la deviazione standard e il coefficiente di variazione. Sebbene entrambi siano importanti e descrivano molti degli evidenti cambiamenti che si incontrano durante un'analisi del passo, essi non tengono conto delle alterazioni più sottili. Precisamente, essi non esaminano la dinamica del cammino ovvero come il passo varia nel corso del tempo a partire da un intervallo di stride al successivo. Ciò è di vitale importanza quando si valutano due serie temporali. Tali serie possono avere identiche medie e varianze ma mostrare persino delle dinamiche molto differenti ed è per questo motivo che si concentra lo studio su proprietà del passo definite frattali. Infatti, notiamo che durante un intervallo di stride, la durata del ciclo fluttua attorno al valore medio della serie temporale [1], [2], [9]. Nell'ultimo decennio, si è arrivati alla conclusione che tali fluttuazioni indicano che ogni intervallo di stride è

strettamente correlato al successivo. In altre parole, le fluttuazioni presenti nell'intervallo stride-to-stride non sono repentini cambiamenti intorno al valor medio. Al contrario, esse mostrano una “memoria a lungo termine” e che quindi è presente una relazione tra gli stride di una serie temporale.

Questa proprietà frattale può essere studiata tramite tecniche diverse. Una tra le più popolari è basata sull'analisi di andature random alterate e va sotto il nome di Detrended Fluctuations Analysis (DFA), già definita in precedenza. Questo tipo di analisi è molto utilizzata in quanto mostra notevoli vantaggi rispetto ai metodi convenzionali, come l'analisi spettrale o l'analisi di Hurst. Essa, infatti, permette di identificare la correlazione presente tra una serie temporale e l'altra e di non considerare gli intervalli che sono apparentemente simili, ma in realtà dovuti ad artefatti di fattori esterni.

L'algoritmo della DFA è fondato su due step:

- (1) la serie di dati denominata $B(k)$ e di lunghezza N è privata del suo valore medio B_{ave} per eliminarne il trend e viene calcolata la somma cumulativa al campione k . Se il profilo ottenuto è

$$y(k) = \sum_{i=1}^k [B(i) - B_{ave}]$$

esso viene successivamente diviso in finestre di osservazione di lunghezza N .

(2) in ogni finestra di osservazione il dato appena integrato è adattato ad un polinomio tipicamente lineare $y_n(k)$. In seguito si calcola l'entità della fluttuazione $F(n)$ (mean-squared residual) come

$$F(n) = \sqrt{\frac{1}{N} \sum_{k=1}^N [y(k) - y_n]^2}$$

Si può notare come la dimensione della fluttuazione dipenda dalla grandezza della finestra di osservazione n . Il principale obiettivo dell'analisi basata su DFA è infatti quello di calcolare un esponente frattale α che quantifica come la dimensione della fluttuazione $F(n)$ dipenda dalla finestra di osservazione (n). Tale esponente è il coefficiente angolare del grafico che si crea graficando la relazione tra $F(n)$ e (n) mediante il logaritmo in base 10. Sull'asse delle ascisse avremo il valore $\log_{10}(n)$, mentre sull'asse delle ordinate avremo $\log_{10}F(n)$. Dopo aver trovato il valore dell'esponente di scala frattale possiamo trarre alcune conclusioni. In particolare, se α è compreso tra un valore pari a 0.5 e 1, le serie temporali sono correlate e simili tra loro per un lungo periodo di tempo. D'altro canto, se tale esponente è pari a 0.5 significa che i processi che si stanno studiando sono random e quindi associati a rumore bianco (rumore caratterizzato dall'assenza di periodicità nel tempo e da ampiezza costante su tutto lo spettro delle frequenze).

2.2 Metodi fondati sulla misura dell'entropia

Negli ultimi decenni, l'entropia ha ottenuto una considerevole attenzione per analizzare la complessità e la non-linearità delle serie temporali. Sono presenti, infatti, diversi algoritmi entropici i quali forniscono utili informazioni riguardo l'irregolarità delle diverse strutture del passo in differenti individui. Tra questi ricordiamo

2.2.1 Approximate Entropy (ApEn)

Ha proposto un algoritmo S.M. Pincus [10] per il calcolo di entropia che va sotto il nome di Approximate Entropy ed è un metodo basato su un approccio statistico che valuta l'irregolarità e le lievi fluttuazioni presenti in un processo fisiologico. Per serie temporali di intervalli di stride di N campioni, il modello ApEn è espresso come $ApEn(m,r,N)$, dove m è un numero intero che denota la lunghezza della finestra in cui si hanno confronti simili, mentre il numero positivo e reale r è il parametro di tolleranza per far sì che tali confronti possano essere ritenuti validi. In seguito, si definisce una sequenza di vettori $\{x^1(i), x^2(i), \dots, x^{(N-m+1)}(i)\}$ dove ogni vettore è composto da m campioni consecutivi delle serie temporali $x^m(i) = [s(i), s(i+1), \dots, s(i+m-1)]$. La distanza $d[x^m(i), x^m(j)]$ è definita come la differenza massima assoluta tra gli elementi corrispondenti dal vettore $x^m(i)$ e $x^m(j)$ dove quest'ultimo viene considerato come vettore modello. Per ogni i tale per cui $1 < i < N-m+1$, definiamo $C_i^m(r)$ come il parametro che

quantifica la probabilità che j possa soddisfare la condizione tale per cui la distanza tra $x^m(i)$ e il modello $x^m(j)$ è più piccola di r ovvero

$$C_i^m(r) = \frac{\text{valore di } j \text{ tale per cui } d[x^m(i), x^m(j)] < r}{N - m + 1}$$

Il valore di $C_i^m(r)$ rappresenta, quindi, la frequenza di matches simili entro il valore soglia r tra il vettore $x^m(i)$ con una finestra temporale lunga m , e un modello dato. Successivamente, si stabilisce una funzione $\phi^m(r)$ come la media dei logaritmi naturali di $C_i^m(r)$ ovvero

$$\phi^m(r) = \frac{\sum_{i=1}^N \ln C_i^m(r)}{N - m + 1}$$

Infine, si definisce $ApEn(m,r)$ come

$$ApEn(m,r) = \lim_{N \rightarrow \infty} [\phi^m(r) - \phi^{m+1}(r)]$$

Se consideriamo delle serie temporali finite, $ApEn$ dipenderà anche dal parametro N e quindi otteniamo

$$\begin{aligned} ApEn(m,r,N) &= [\phi^m(r) - \phi^{m+1}(r)] \\ &= \frac{\sum_{i=1}^{N-m+1} \ln C_i^m(r)}{N - m + 1} - \frac{\sum_{i=1}^{N-m} \ln C_i^{m+1}(r)}{N - m} \end{aligned}$$

Nel caso in cui N è molto maggiore di m , possiamo approssimare il valore di $ApEn(m,r,N)$ esclusivamente al secondo membro dell'equazione precedente.

In analisi del cammino, tipicamente, un valore basso di ApEn indica che le serie temporali hanno un alto grado di regolarità.

2.2.2 *Sample Entropy (SampEn)*

Come nell'Approximate Entropy, anche l'algoritmo di Sample Entropy è un metodo che ci permette di quantificare la complessità delle serie temporali. In realtà, l'algoritmo basato su SampEn, introdotto da J.S. Richman e J.R. Moorman [11], è una modifica dell'algoritmo descritto precedentemente in cui si calcola il valore di ApEn. Tuttavia, SampEn presenta una piccola differenza di calcolo tra i due algoritmi e offre il vantaggio di essere indipendente dalla lunghezza dei dati. In ApEn, infatti, il confronto tra il vettore modello e il resto dei vettori include anche il confronto con sé stesso e ciò significa che le probabilità $C_i^m(r)$ non sono mai zero. Di conseguenza, è sempre possibile ottenere il logaritmo di probabilità. Dato che il modello confronta con sé stesso valori di ApEn inferiori, i segnali da analizzare vengono interpretati come più regolari di quanto non siano in realtà. Nell'algoritmo di SampEn questi auto-confronti non sono inclusi. Anche in questo caso, definiamo un numero di N campioni, una dimensione di inclusione m e un valore di soglia r. Si definisce un valore che indica la percentuale di distanze tra il vettore campione x_j^m e il vettore x_i^m delle serie temporali che ricadono entro il valore di soglia r, escludendo gli auto-confronti.

$$A_i^{(m)}(r) = \frac{N_i^{(m)}(r)}{N - m + 1}$$

In tale formula, m rappresenta la dimensione di integrazione, N rappresenta la lunghezza dei dati, r rappresenta il parametro di soglia, e, infine, $N_i^{(m)}(r)$ mostra il numero di j nell'intervallo tale per cui si verifica che la distanza tra la serie i e la serie j sia minore del valore soglia con $i \neq j$.

Successivamente, si calcola la percentuale media tramite il valore percentuale precedentemente trovato.

$$\Psi^m(r) = \frac{1}{N - m + 1} \sum_{i=1}^{N-m+1} A_i^{(m)}(r)$$

Allo stesso modo, $\Psi^{(m+1)}(r)$ è calcolato aumentando la dimensione di inclusione a $m+1$. SampEn è quindi data dalla seguente formula

$$SampEn(m, r, N) = \ln \frac{\Psi^{(m+1)}(r)}{\Psi^m(r)}$$

SampEn sarà, quindi, definita come il logaritmo naturale della probabilità che se due insiemi di punti simultanei di lunghezza m si trovano ad una distanza minore del valore di soglia r , allora anche due serie di punti simultanei di lunghezza $m+1$ si trovano anche loro ad una distanza minore del valore r .

2.2.3 Distribution Entropy (DistEn)

Un ulteriore metodo impiegato per misurare la complessità e l'irregolarità del cammino è quello fondato sulla Distribution Entropy, messo a punto da P. Li e altri [12]. Tale algoritmo quantifica la distribuzione delle distanze tra vettori invece di quantificare unicamente la probabilità che tali distanze ricadano nello spazio di stato limitato da un valore di soglia. Per ricostruire lo spazio di stato, consideriamo un insieme di campioni $\{u(i), 1 \leq i \leq N\}$ tali per cui si hanno dei vettori così formati $X(i) = \{u(i), u(i + \tau), \dots, u(i + (m - 1)\tau)\}, 1 \leq i \leq N - (m - 1)\tau$. In questo caso, il parametro τ indica la costante di tempo. Successivamente, si calcola una matrice che contiene le distanze tra vettori, ovvero tutte le possibili combinazioni tra $X(i)$ e $X(j)$. L'equazione che ci permette di calcolare la distanza tra le possibili combinazioni è la seguente

$$d[i, j] = \max(|u(i + k) - u(j + k)|, 0 \leq k \leq m - 1)$$

per ogni $1 \leq i, j \leq N - m$

Notiamo che l'algoritmo DistEn usa un numero di parametri M al posto di un valore di soglia r . Per determinare M , sfruttiamo la funzione di densità di probabilità empirica (ePDF) della matrice D di distanze la quale è stimata mediante istogrammi con un numero selezionato di bins. La scelta di M viene attuata in modo tale per cui il valore di M possa ripartire opportunamente la distribuzione della matrice $D = \{d[i, j]\}$, di dimensioni $(N-m)(N-m-1)$. Anche in

questo caso, gli elementi con $i = j$ sono esclusi dalla stima di ePDF. Infine, DistEn è calcolata come

$$DistEn(m, \tau, M) = \frac{1}{\log_2(M)} \sum_{i=1}^M P_t \log_2(P_t)$$

dove P_t rappresenta la probabilità di ogni bin.

2.2.4 Multiscale Entropy Analysis (MSE)

Come descritto in precedenza, i metodi di analisi del cammino basati su entropia forniscono una stima del grado di regolarità del processo che si vuole analizzare. Tuttavia, è molto difficile dare una definizione esatta di complessità ed inoltre può risultare complicato calcolarla. In altre parole, non vi è diretta corrispondenza tra entropia e complessità. Infatti, alcuni metodi entropici tradizionali possono condurre a risultati ingannevoli in quanto un aumento di entropia di un sistema è solitamente ma non sempre associato ad un aumento di complessità. Una possibile ragione per cui gli algoritmi tradizionali basati su entropia possono fallire nel quantificare il corretto grado di complessità di una serie temporale è che queste misure sono fondate su scale spazio-temporali singole. Come visto in precedenza, ciò è ben evidente nello studio condotto da M.Costa [5], dove si evince quanto sia vantaggioso servirsi di più scale spazio-temporali a causa della complessità dei segnali in esame. È per tale motivo che viene introdotto un metodo chiamato Multiscale

Entropy Analysis che permette di calcolare l'entropia considerando scale temporali multiple. Data una serie temporale $\{x_1, \dots, x_i, \dots, x_N\}$, si costruiscono delle serie grezze consecutive facendo una media di un numero sempre maggiore di dati in finestre non sovrapposte. Ogni elemento della serie temporale grezza è calcolato nel seguente modo

$$y_j^{(\tau)} = 1/\tau \sum_{i=(j-1)\tau+1}^{j\tau} x_i$$

dove τ rappresenta il fattore di scala e j è tale per cui $1 \leq j \leq N/\tau$. Per la scala corrispondente a 1, la serie è semplicemente quella originale. In seguito, si calcola la Sample Entropy per ogni serie come una funzione dipendente dal fattore di scala. Come abbiamo già detto in precedenza, SampEn quantifica la regolarità di una serie temporale. In particolare, essa riflette la probabilità che due sequenze di m dati consecutivi che sono simili tra loro rimarranno simili persino quando un ulteriore dato è incluso. Essere simili vuol dire che il valore di distanza che si ha tra quei due dati è minore di un valore soglia pari a r , come descritto precedentemente.

2.3 Recurrence Quantification Analysis (RQA)

Come detto precedentemente, i segnali legati al cammino sono più utili per individuare disturbi causati del sistema nervoso centrale. Tradizionalmente, per identificare tali disordini viene utilizzata l'analisi mediante la trasformata di Fourier. Tuttavia, questo tipo di analisi non sempre risulta adeguata. In particolare, Fourier ipotizza che il segnale sia stazionario, lineare e privo di rumore, ma è evidente che queste supposizioni non

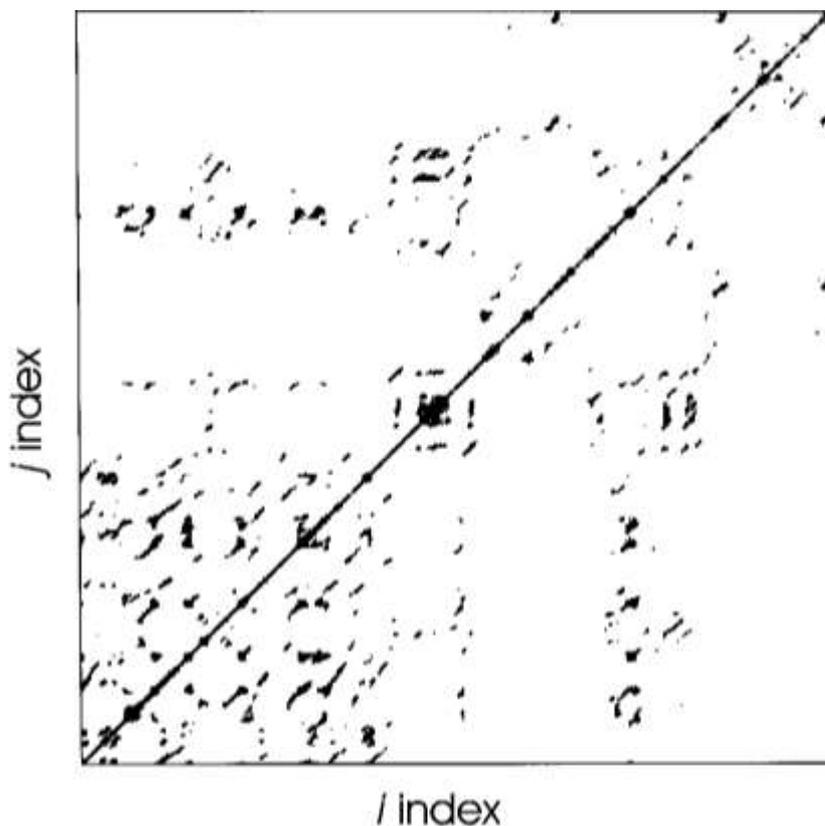


Figura 9. Esempio di Recurrence Plot

rispecchiano i segnali reali. Per risolvere questo problema, si introduce il metodo Recurrence Quantification Analysis per quantificare parametri tipici dell'andatura dell'uomo. È un metodo adatto allo studio di segnali non lineari, non stazionari e di lunghezza breve. Più nello specifico, l'analisi

RQA si concentra nel quantificare le caratteristiche geometriche che si osservano in strutture a scala ridotta e graficate come Recurrence Plots (RP). I Recurrence Plots contengono principalmente singoli punti e linee parallele alla diagonale media (linea di identità) o diagonali o verticali. Le linee parallele alla linea di identità sono le linee diagonali e le strutture verticali sono le linee verticali. Tali “textures” identificano il numero e la durata delle ricorrenze di un sistema dinamico. La recurrence matrix è calcolata come segue

$$R_{i,j}(\varepsilon) = H(\varepsilon - \|x_i - y_j\|)$$

dove ε è un valore soglia definito da chi opera, x_i e y_j sono i vettori di embedding e H è la funzione di Heaviside. Valutando l’architettura delle textures, si estraggono tipicamente otto parametri tipici dell’analisi RQA.

Recurrence Rate (RR)

RR quantifica la percentuale di ricorrenza di punti presenti in RP con l’eccezione di quelli presenti nelle linee di identità. Ciò implica che i segnali periodici hanno un *RR* più elevato, poiché la periodicità è più presente in segnali di andature più disordinate. Il valore di *RR* più basso si avrà quindi in soggetti sani. [6]

Determinism (DET)

DET misura l'indice di ricorrenza di punti che formano linee diagonali di lunghezza minima l_{min} . Le linee diagonali più lunghe segnalano un segnale di tipo periodico, mentre quelle più corte sono indice di segnali più caotici. Perciò, linee diagonali più lunghe sono presenti in segnali di soggetti patologici rispetto a segnali di soggetti sani. [6]

Average diagonal line (L)

Average diagonal line è un parametro ottenuto facendo la media di tutti gli 1 consecutivi nell'RP i quali sono paralleli alla linea di identità. Tale indice misura la media temporale in cui le traiettorie associate a due serie temporali sono vicine tra di loro.

Maximal diagonal line (Lmax)

L_{max} determina la lunghezza della linea diagonale più lunga quando confrontata con le altre linee diagonali parallele alla linea di identità. I segnali meno stabili o quelli più caotici hanno un basso valore di L_{max} . È per questo motivo che segnali di soggetti sani, i quali sono caotici per natura, avranno un valore di L_{max} piccolo, considerando che segnali più caotici danno origine a linee diagonali più lunghe. Al contrario, L_{max} sarà tanto maggiore quanto più il segnale sarà periodico e caratteristico di un soggetto malato. [6]

Entropy (ENT)

ENT misura la complessità del segnale determinando la probabilità di occorrenze di tutte le linee diagonali di lunghezza al di sopra di l_{min} . In particolare, nel caso di soggetti sani le linee diagonali del segnale corrispondente saranno di lunghezza variabile e questo sarà indice di un valore di entropia basso e quindi di elevata complessità. Ciò è dovuto dal fatto che soggetti normali hanno vari gradi di libertà. Al contrario, nel caso di soggetti patologici, il valore di *ENT* sarà elevato e quindi la complessità del segnale sarà minore. [6]

Laminarity (LAM)

LAM misura la percentuale di punti ricorrenti che formano linee verticali di lunghezza minima v_{min} .

Trapping Time (TT)

TT viene trovato mediante la media di tutte le lunghezze delle linee verticali al di sopra del valore di soglia. Ad esempio, nel caso della sindrome di Huntington, il valore di *TT* è più alto quando confrontato con le altre classi, dal momento che la persona affetta da tale malattia mostra dei movimenti improvvisi seguiti da tremori a riposo.

2.4 Materiali

I dati utilizzati nella presente indagine sono formati da intervalli di stride espressi in secondi. Questi intervalli sono relativi a un dataset presente in letteratura [7] formato da 50 bambini sani di età compresa tra i 40 mesi e i 163 mesi. Lo scopo dello studio è quello di caratterizzare la variabilità e le caratteristiche frattali in base all'età. Si sono divisi i 50 bambini in tre gruppi in base a tre fasce di età.

- 1) Il primo gruppo comprende 16 bambini di età compresa tra 40 e 79 mesi
- 2) Il secondo gruppo è formato da 18 bambini di età tra 80 e 94 mesi
- 3) Il terzo gruppo è costituito da 16 bambini di età tra 101 e 163 mesi

CAPITOLO 3

Risultati ottenuti

L'obiettivo del presente studio è quello di osservare che relazione ci sia tra gli intervalli di stride di soggetti sani e l'avanzamento dell'età. I metodi utilizzati per l'analisi delle serie temporali sono la Detrended Fluctuations Analysis (DFA) e la Recurrence Quantification Analysis (RQA).

3.1 Analisi mediante DFA

Per ogni serie temporale si è calcolato l'indice di scala frattale per verificare se le serie temporali dei dati in esame sono correlate tra di loro a lungo raggio.

Si sono ottenuti i seguenti valori di esponente di scala frattale secondo la divisione in gruppi precedentemente descritta.

EXP GRUPPO 1	EXP GRUPPO 2	EXP GRUPPO 3
0.8387 ± 0.1519	0.9150 ± 0.1077	0.9284 ± 0.1334

Tabella 1 – Valori medi \pm SD degli esponenti di scala frattale divisi nei tre gruppi in base all'età

Dai dati riportati nella Tabella 1, possiamo notare che i valori medi degli esponenti di scala frattale per tutti e tre i gruppi sono compresi tra 0.5 e 1. Infatti i soggetti analizzati sono tutti sani e quindi il comportamento atteso, alla luce di quanto detto in precedenza, è che le serie temporali non presentino caratteristiche random come nel caso di pazienti patologici, bensì

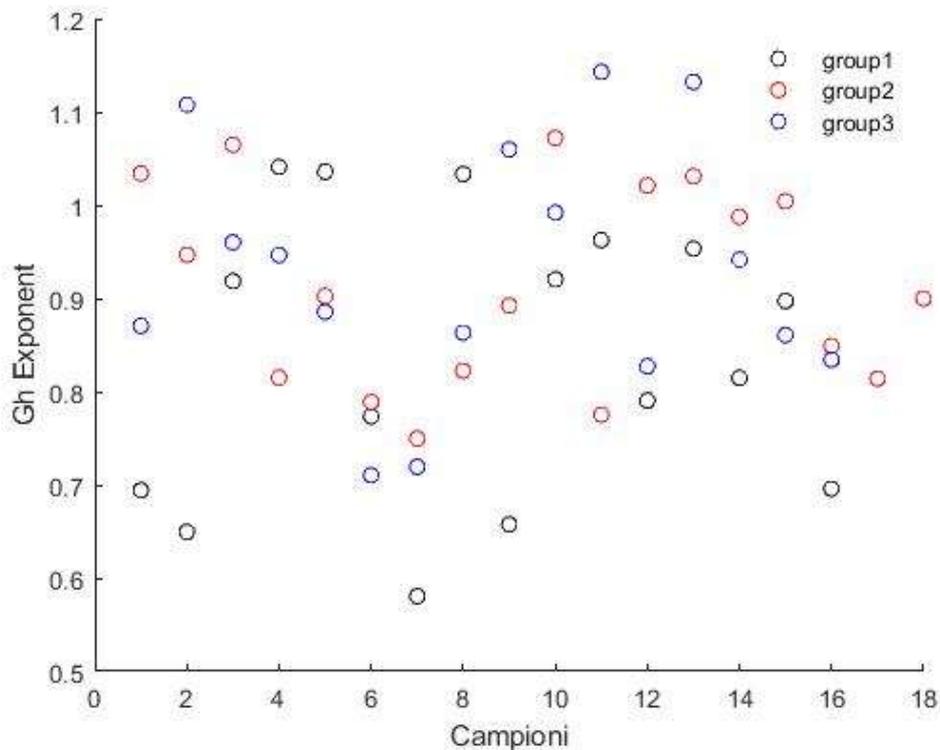


Figura 10. Distribuzione dei valori di esponente di scala frattale secondo i tre gruppi di età differente.

siano correlate tra di loro. Se si considerano i dati della tabella 1 in base alla

suddivisione nei tre gruppi, notiamo che non si ha una distinzione definita tra gli indici frattali. Tali valori sono molto vicini tra di loro evidenziando che tutti e tre i gruppi hanno pressoché indici di scala frattale molto simili. Confrontando, infatti, la distribuzione dei valori degli esponenti trovati tra i tre gruppi, non riusciamo a evidenziare differenze tra le serie temporali. La figura 10 mostra come i valori dell'esponente di scala frattale siano molto vicini tra di loro. Ciò non ci permette di concludere in maniera chiara come

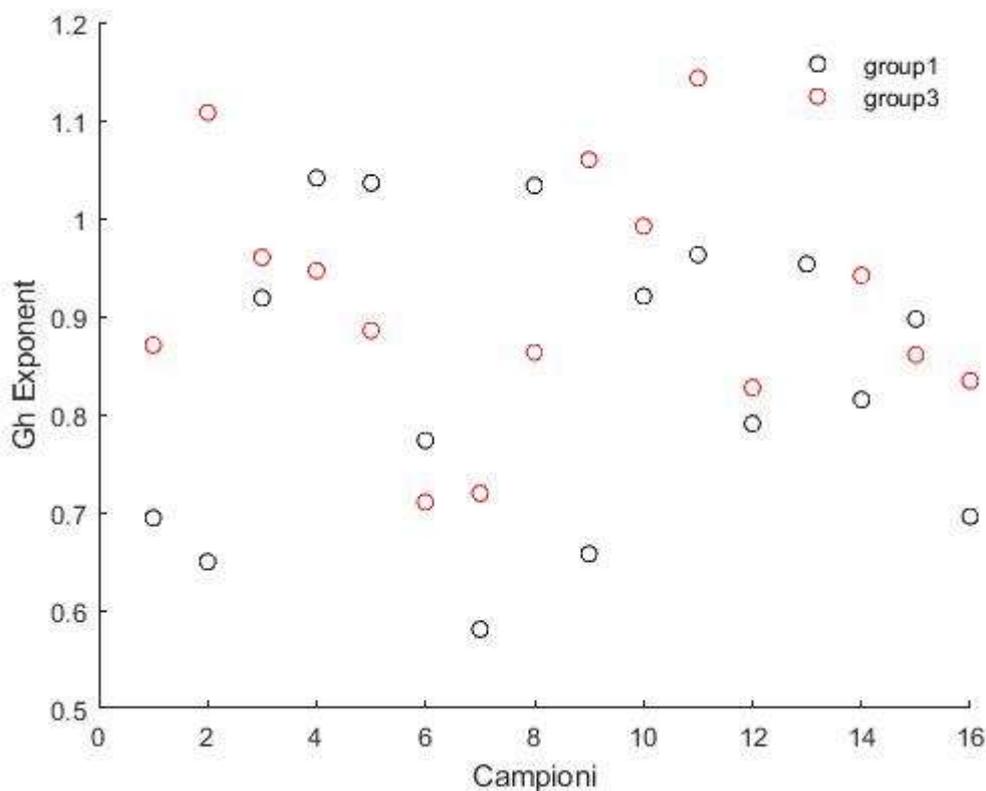


Figura 11. Distribuzione dei valori di esponente di scala frattale rispetto al gruppo 1 e al gruppo 3

gli intervalli stride-to-stride varino in funzione della maturazione dell'individuo.

Un'ulteriore conferma di questa conclusione è data dal fatto che, eliminando dal grafico la distribuzione dei valori dell'esponente di scala del secondo

gruppo e lasciando a confronto i valori del primo e del terzo gruppo i quali sono più separati come range di età, questi risultano essere molto vicini tra loro (figura 11).

Si è eliminato il secondo gruppo dal grafico in modo tale da riuscire a vedere se la distinzione tra gli esponenti è facilmente visibile dal momento che la differenza di età tra primo e terzo gruppo è elevata.

In conclusione, possiamo dedurre che il metodo applicato, ovvero la DFA, non ci permette di determinare in maniera esaustiva come variano le caratteristiche di tipo frattale in funzione della maturazione del cammino degli individui analizzati nel presente studio.

3.2 Analisi mediante RQA

Il secondo metodo utilizzato è la Recurrence Quantification Analysis (RQA), già definita in precedenza, per mezzo della quale sono stati calcolati 10 parametri per tutte le serie temporali esaminate. Anche in questo caso si sono confrontate le medie dei parametri ottenuti tra tutti e tre i gruppi di bambini. Si è esclusa dalla seguente analisi la serie temporale di un soggetto appartenente al secondo gruppo poiché caratterizzata da un'anomalia.

In primo luogo, osserviamo il valore medio di Recurrence Rate (RR) trovato per ogni gruppo.

RR GRUPPO 1	RR GRUPPO 2	RR GRUPPO 3
1.9873 ± 0.5928	1.6454 ± 1.0497	1.4193 ± 0.6318

Tabella 2 – Valori medi ±SD del parametro RR secondo la divisione in tre gruppi

Il gruppo numero 3 ha un valor medio di RR più basso rispetto agli altri, perciò, limitatamente al parametro RR, gli intervalli di stride di bambini più grandi sembrano più regolari. Se, invece, ci si limitasse al risultato ottenuto dal grafico in figura 12, si osserva che la distribuzione del parametro RR

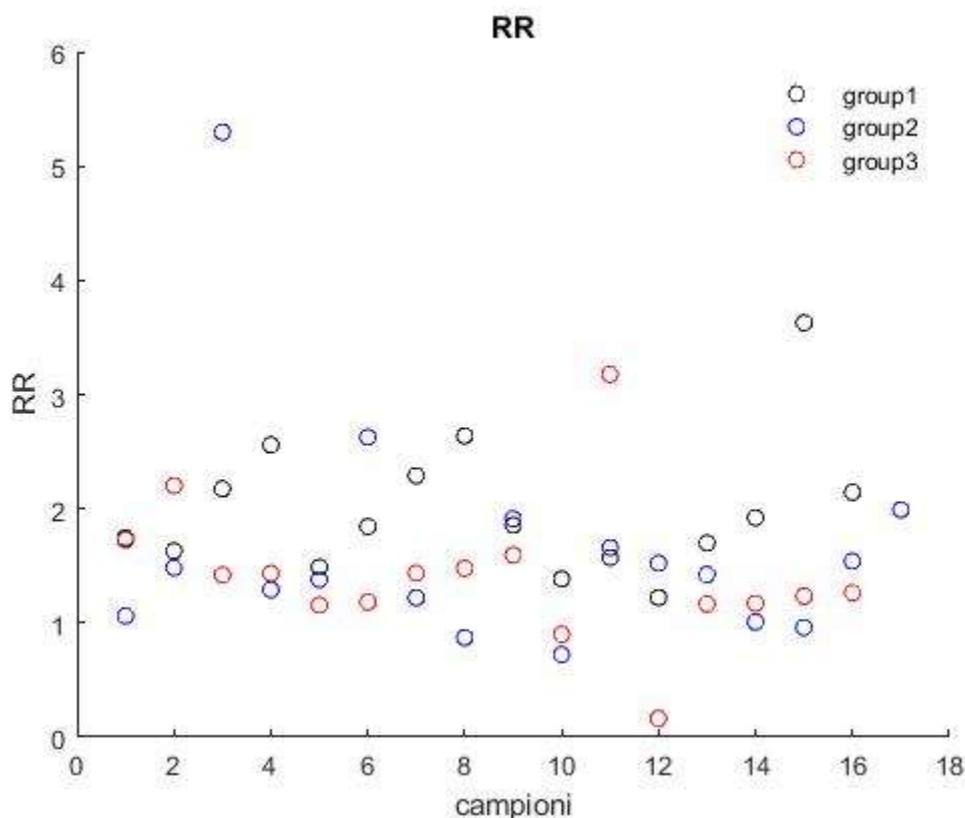


Figura 12. Distribuzione del parametro RR in funzione dei tre gruppi

non è ben definita tra i tre gruppi, come per l'esponente di scala frattale.

Considerando, invece, il valore medio del parametro Determinism (DET) per ogni gruppo otterremo i seguenti valori

DET GRUPPO 1	DET GRUPPO 2	DET GRUPPO 3
56.5408 ± 5.4057	51.1525 ± 9.7987	47.5818 ± 11.9755

Tabella 3 – Valori medi ±SD del parametro DET secondo la divisione in tre gruppi

Anche in questo caso, come nel caso dell'indice RR, il valore medio del parametro DET risulta minore per il gruppo 3. Questo sembra ad indicare che il gruppo 3 sia formato da bambini con un'andatura più regolare. Inoltre,

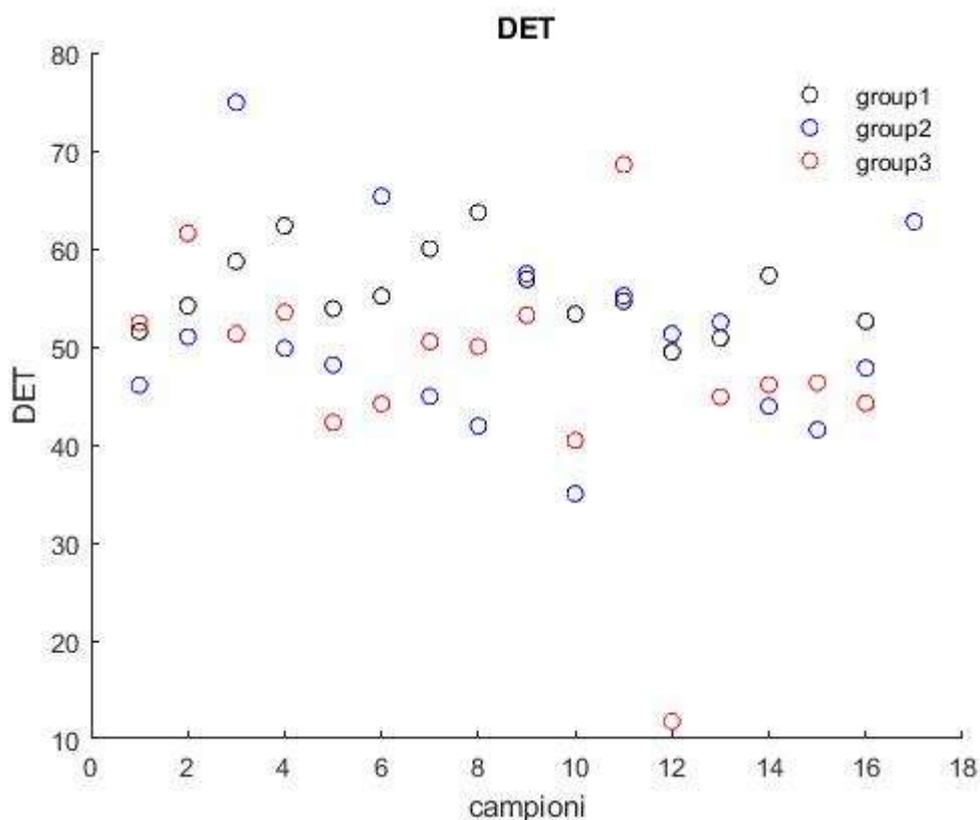


Figura 13. Distribuzione del parametro DET in funzione dei tre gruppi

si può notare che il parametro DET presenta un trend decrescente tra il primo

e il terzo gruppo. Perciò, l'indice DET ci permette di supporre che le serie temporali che stiamo analizzando saranno tanto più regolari quanto più si avanza con l'età. Nonostante ciò, se esaminiamo i valori dei parametri DET per ogni gruppo e come essi si distribuiscono tra loro come in figura 13, non riscontriamo una netta separazione tra i vari gruppi sebbene i valori medi del parametro siano diversi tra loro.

Andando avanti nell'analisi, un parametro molto importante che ci permette di determinare con certezza quali serie temporali siano più complesse e regolari rispetto alle altre ENT, ovvero l'entropia. Per quanto riguarda i valori di entropia calcolati per ogni gruppo, avremo i seguenti valori medi

ENT GRUPPO 1	ENT GRUPPO 2	ENT GRUPPO 3
1.136 ± 0.1198	1.0688 ± 0.2393	0.9911 ± 0.1958

Tabella 4 – Valori medi ±SD del parametro ENT secondo la divisione in tre gruppi

Più il valore di entropia sarà basso, più le serie temporali saranno regolari. Anche in questo caso, il valore medio delle entropie registrate per il terzo gruppo risulta minore rispetto agli altri due. Facendo un confronto tra gruppi a due a due, osserviamo che la differenza tra il gruppo 1 e il gruppo 2 del valore entropico medio sarà minima, così come tra gruppo 2 e gruppo 3. Nel caso in cui, invece, viene confrontato il valore di ENT medio tra gruppo 1 e gruppo 3, si riesce a determinare quali serie temporali sono più regolari. In questo caso, dato che il valore di entropia del gruppo 3 è minore, i bambini

di questo gruppo avranno degli stride-to-stride più regolari rispetto al gruppo 1. Se, invece, si osserva il grafico in figura 14, si ha una situazione analoga a quella descritta nelle figure precedenti, ovvero i dati risulteranno

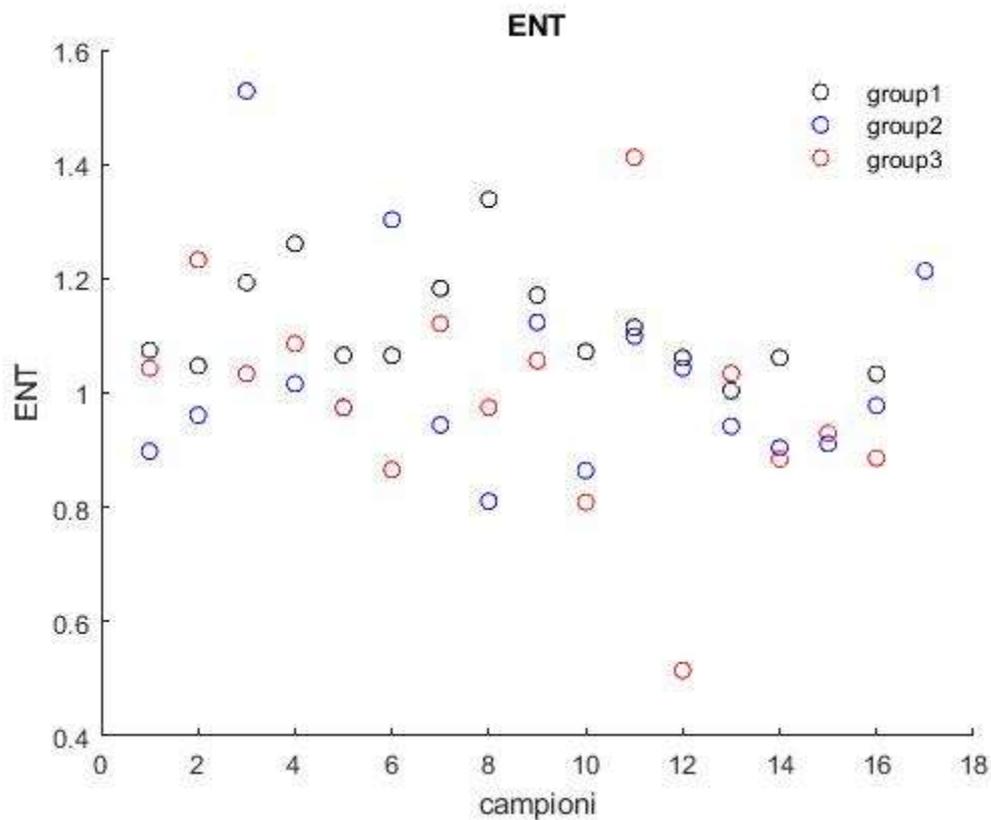


Figura 14. Distribuzione del parametro ENT in funzione dei tre gruppi

scarsamente distinguibili tra i tre gruppi.

Un ulteriore parametro da considerare è Maximal Line Length (MLL). Tenendo conto anche in questo caso dei valori medi ottenuti per i tre gruppi

MLL GRUPPO 1	MLL GRUPPO 2	MLL GRUPPO 3
9.9375 ± 2.6450	8.1176 ± 2.4208	7.3750 ± 1.9958

Tabella 5 – Valori medi ±SD del parametro MLL secondo la divisione in tre gruppi

si nota una maggiore differenza tra i valori ottenuti per il primo e quelli per il terzo gruppo. Infatti, il valore di MLL del terzo gruppo è minore rispetto

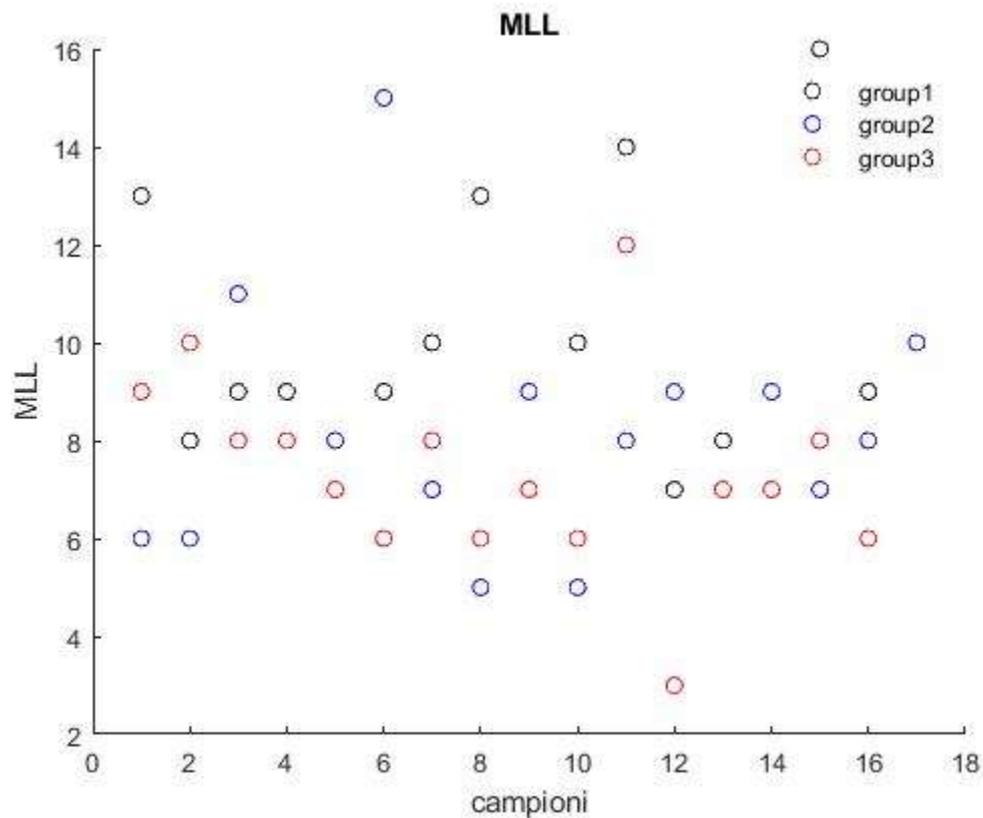


Figura 15. Distribuzione del parametro MLL in funzione dei tre gruppi

ai valori medi di MLL dei primi due gruppi. Ciò conferma ulteriormente quanto concluso prima, ovvero che le serie temporali di intervalli di stride del gruppo 3 formato da bambini di età maggiore saranno più controllate e meno caotiche. Al contrario, dalla figura 14, non emerge chiaramente questa distinzione.

Spostando l'attenzione sul parametro LAM, si sono registrati i seguenti valori medi

LAM GRUPPO 1	LAM GRUPPO 2	LAM GRUPPO 3
30.8354 ± 9.3314	24.0195 ± 12.8273	20.4182 ± 11.2740

Tabella 6 – Valori medi ±SD del parametro LAM secondo la divisione in tre gruppi

La differenza tra di essi è evidente, in particolar modo confrontando i valori tra gruppo 1 e gruppo 3. LAM indica quanto una serie temporale è regolare e più il suo valore è basso più la serie sarà regolare, come nel caso degli intervalli temporali del gruppo 3. Come detto per il parametro ENT, anche

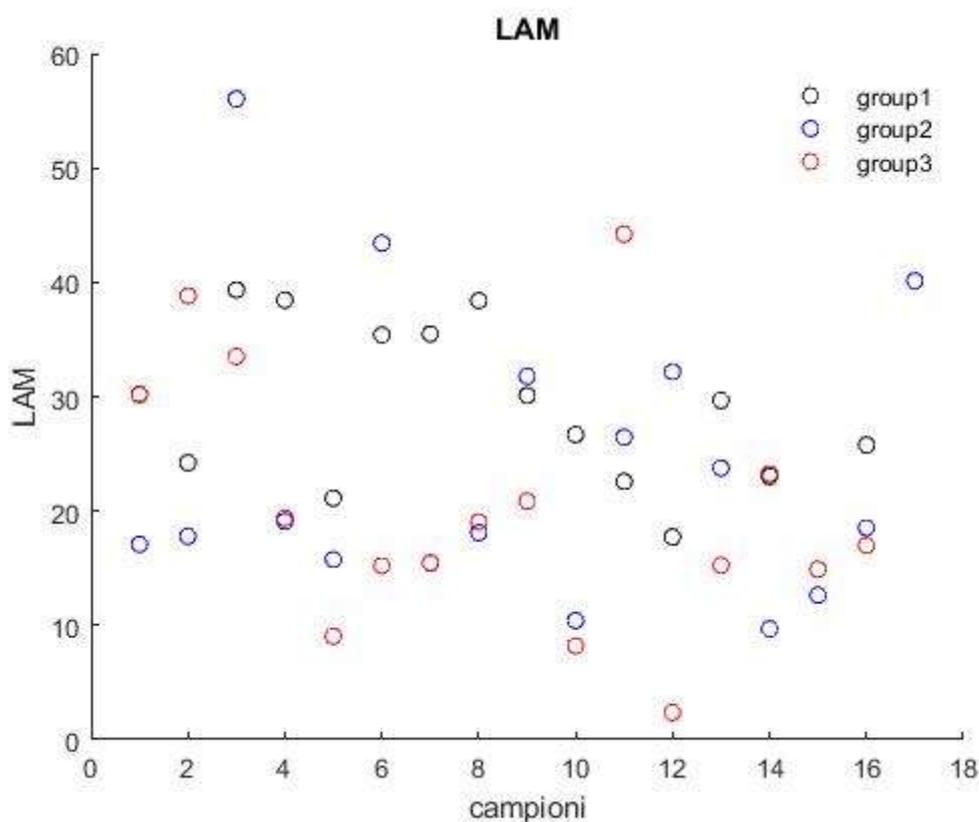


Figura 16. Distribuzione del parametro LAM in funzione dei tre gruppi

nel caso di LAM le serie temporali del gruppo 3 mostrano maggiore regolarità rispetto alle altre. Dalla figura 16, però, possiamo notare che la

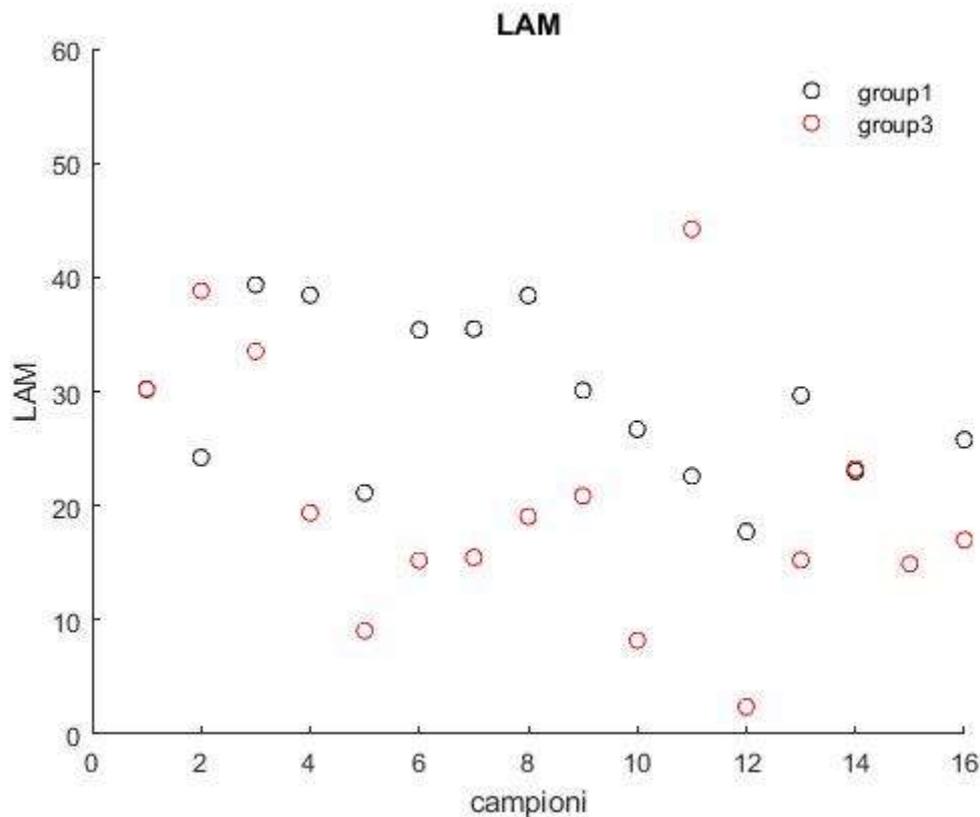


Figura 17. Distribuzione del parametro LAM del primo e del terzo gruppo

distribuzione dei valori del parametro LAM non evidenzia una separazione tra i gruppi.

Se consideriamo la figura 17, la distinzione dei valori del parametro LAM sarà leggermente più marcata quando si confrontano i gruppi 1 e 3. Ciò avviene poiché per il terzo gruppo si hanno valori minori di laminarità, mentre per il primo gruppo saranno maggiori.

In conclusione, possiamo dire che tramite il metodo della Recurrence Quantification Analysis (RQA) si è in grado di discriminare i dati presi in

esame nel caso in cui vengano confrontati i due gruppi con maggior differenza di età. In questo caso, alcuni parametri (RR, DET, LAM) dimostrano come le serie temporali di bambini più adulti sono più regolari rispetto a quelle di bambini più piccoli. Perciò, parametri che evidenziano regolarità nelle serie temporali permettono di caratterizzare lo sviluppo del cammino in funzione della maturazione. Come si evince dal presente studio, i metodi della DFA e della RQA non sono riusciti a fornire risultati chiari riguardo la distinzione delle serie temporali secondo la fascia di età.

3.3 Test statistici

Vengono poi applicate delle valutazioni statistiche utilizzate per misurare la probabilità che le differenze osservate dai parametri considerati e relative ai tre gruppi siano statisticamente significative. La probabilità di essere distinti gli uni dagli altri viene identificata dal p-value il quale, per far sì che due gruppi abbiano maggiore probabilità di essere differenti, deve avere un valore al di sotto di 0.05.

Per quanto riguarda i risultati ottenuti nel presente studio, si è applicata l'analisi statistica ad ogni parametro trovato mediante DFA e RQA.

Nel caso dell'esponente di scala frattale, il test statistico conferma quanto concluso in precedenza, ovvero che tale indice non è statisticamente diverso nei tre gruppi e pertanto non sembra essere adeguato a descrivere la maturazione del cammino poiché i valori di p-value registrati confrontando i gruppi a due a due sono tutti e tre maggiori di 0.05.

GRUPPO		GRUPPO	P-VALUE
1		2	0.3779
1		3	0.2558
2		3	0.9543

Tabella 7 – Valori di p-value registrati confrontando i valori dell'esponente di scala frattale per i tre gruppi

Al contrario, per quanto riguarda i parametri precedentemente descritti ed estrapolati dalla RQA, possiamo notare una maggiore distinzione tra gruppi. In particolare, i valori di p-value ottenuti confrontando il primo e il terzo gruppo saranno minori di 0.05 per i parametri RR, DET, ENT, MLL e LAM.

RR	GRUPPO		P-VALUE
	1	2	0.0614
	1	3	0.0090
	2	3	0.7209

Tabella 8 – Valori di p-value registrati confrontando i valori di RR per i tre gruppi

DET	GRUPPO		P-VALUE
	1	2	0.0920
	1	3	0.0080
	2	3	0.5894

Tabella 9 – Valori di p-value registrati confrontando i valori di DET per i tre gruppi

ENT	GRUPPO	GRUPPO	P-VALUE
	1	2	0.0702
	1	3	0.0188
	2	3	0.8325

Tabella 10 – Valori di p-value registrati confrontando i valori di ENT per i tre gruppi

MLL	GRUPPO	GRUPPO	P-VALUE
	1	2	0.1501
	1	3	0.0088
	2	3	0.4669

Tabella 11 – Valori di p-value registrati confrontando i valori di MLL per i tre gruppi

LAM	GRUPPO	GRUPPO	P-VALUE
	1	2	0.2058
	1	3	0.0175
	2	3	0.5084

Tabella 12 – Valori di p-value registrati confrontando i valori di LAM per i tre gruppi

Si può notare che dal confronto tra gruppo 1 e gruppo 3 si ottiene per tutti i parametri un valore di p-value minore di 0.05. Ciò significa tutti i parametri calcolati mediante RQA risultano statisticamente differenti tra i due gruppi. Al contrario, confrontando il gruppo 1 con il 2 e il gruppo 2 con il 3, si ottiene un p-value maggiore di 0.05. Pertanto, si può dedurre che il gruppo

2, essendo composto da bambini di età intermedia tra il primo e il terzo gruppo, non si differenzia particolarmente rispetto agli altri due gruppi. Tramite queste relazioni statistiche, quindi, è emerso come il gruppo 2 sia poco caratterizzabile rispetto agli altri due. Infatti, la dinamica del cammino dei soggetti del secondo gruppo risulta meno regolare rispetto a quella che caratterizza il gruppo dei più grandi.

CAPITOLO 4

Discussione

L'approccio utilizzato e i risultati ottenuti riguardo il presente studio possono essere confrontati con indagini svolte in altri studi presenti in letteratura. In particolare, si pone l'attenzione sulla relazione che intercorre tra le serie temporali formate da intervalli stride-to-stride e la maturazione del cammino dei bambini. Considerando un'analisi svolta da V. Alcan [7] si possono paragonare i risultati ottenuti nel presente studio con quelli ricavati da Alcan, i cui dati sono gli stessi impiegati nell'indagine corrente. In particolare, si riscontrano delle differenze legate alla diversa suddivisione dei gruppi e ai metodi impiegati. Alcan suddivide i gruppi come riportato in tabella 13.

	PRESENTE STUDIO	STUDIO DI V. ALCAN
GRUPPO 1	16 soggetti, 3/6 anni	14 soggetti, 3/5 anni
GRUPPO 2	18 soggetti, 7/8 anni	21 soggetti, 6/8 anni
GRUPPO 3	16 soggetti, 9/14 anni	15 soggetti, 10/14 anni

Tabella 13 – Confronto della divisione in gruppi tra la ricerca presente e studio di V. Alcan

Si può notare che la scelta di Alcan rende i gruppi leggermente sbilanciati nella numerosità, in particolare il secondo gruppo. Alcan ricorre a tre metodi

che si basano sul calcolo di tre differenti entropie e confronta i risultati ottenuti in modo da stabilire quale tra i tre metodi è più adatto per discriminare i tre gruppi. Tali metodi sono tre algoritmi che vanno sotto il nome di Approximate Entropy (ApEn), Sample Entropy (SampEn) e Distribution Entropy (DistEn). Per ogni algoritmo si ottiene un valore specifico di entropia. Alcan applica i tre algoritmi per molteplici valori di r (valore soglia), m (lunghezza della finestra temporale) e M (nel caso di DistEn) in modo da osservare se i metodi sono strettamente dipendenti da tali valori. I risultati ottenuti da Alcan sono riportati di seguito nelle tabelle.

PARAMETRI IN INPUT	ENTROPIA	PRIMO GRUPPO	SECONDO GRUPPO	TERZO GRUPPO
m=2, r=0.15	ApEn	1.150±0.06	1.152±0.19	1.023±0.32
	SampEn	1.981±0.18	2.064±0.31	2.259±0.51
m=2, r=0.20	ApEn	1.299±0.02	1.294±0.05	1.270±0.08
	SampEn	1.715±0.14	1.831±0.25	1.868±0.21
m=2, r=0.25	ApEn	1.307±0.04	1.286±0.10	1.315±0.05
	SampEn	1.545±0.13	1.529±0.27	1.685±0.20
m=2, r=0.30	ApEn	1.264±0.05	1.283±0.10	1.301±0.08
	SampEn	1.348±0.11	1.427±0.23	1.529±0.23
m=3, r=0.15	ApEn	0.354±0.11	0.350±0.17	0.285±0.18
	SampEn	2.070±0.35	1.973±0.30	1.978±0.24
m=3, r=0.20	ApEn	0.591±0.09	0.530±0.14	0.480±0.13
	SampEn	1.717±0.18	1.828±0.34	1.868±0.30

m=3, r=0.25	ApEn	0.722±0.08	0.721±0.14	0.642±0.14
	SampEn	1.515±0.17	1.552±0.30	1.675±0.24
m=3, r=0.30	ApEn	0.839±0.04	0.813±0.08	0.758±0.12
	SampEn	1.326±0.11	1.384±0.29	1.525±0.25

Tabella 14 – Risultati ottenuti dallo studio di V. Alcan secondo gli algoritmi ApEn e SampEn

Dalla tabella 14, si nota come i valori di entropie ApEn e SampEn non siano in grado di discriminare in maniera significativa i tre gruppi. Infatti, i valori di entropie registrate non permettono di caratterizzare i tre gruppi in quanto essi sono molto vicini tra loro.

Un esito differente è ottenuto da Alcan con la Distribution Entropy

PARAMETRI IN INPUT	ENTROPIA	PRIMO GRUPPO	SECONDO GRUPPO	TERZO GRUPPO
m=2, M=256	DistEn	0.750±0.04	0.652±0.05	0.609±0.04
m=2, M=512	DistEn	0.750±0.04	0.651±0.05	0.609±0.04
m=3, M=256	DistEn	0.753±0.04	0.660±0.06	0.608±0.04
m=3, M=512	DistEn	0.753±0.04	0.660±0.06	0.608±0.04

Tabella 15 – Risultati ottenuti dallo studio di V. Alcan secondo l'algoritmo DistEn

L'analisi svolta da Alcan evidenzia come i valori di DistEn non siano dipendenti dal set di parametri in input. Tuttavia, se si confrontano i valori di entropia ottenuti con tali parametri, si nota una differenza statistica

significativa tra il primo e il terzo gruppo con p-value inferiori a 0.001. Inoltre, è possibile rilevare che il valore di entropia ottenuto diminuisce con l'aumentare dell'età. Tale comportamento si riscontra anche nell'analisi mediante RQA attuata nel presente studio quando si confrontano le differenze statistiche significative tra gruppo 1 e gruppo 3. In particolare, i valori del parametro ENT registrati tramite RQA del gruppo 1 e del gruppo 3 diminuiscono significativamente quando si ha un aumento così ampio dell'età (p-value pari a 0.0188). Nel presente studio, questo trend è osservato anche dagli altri parametri ottenuti mediante RQA.

Se, invece, si pone l'attenzione sul gruppo composto da bambini di età intermedia, Alcan con la DistEn riesce a renderlo caratterizzabile. Nel suo lavoro si nota una differenza statistica significativa tra i tre gruppi. In particolare, questa differenza è evidente dal confronto tra gruppo 1 e 2 (p-value inferiore a 0.01) e tra gruppo 2 e 3 (p-value inferiore a 0.05). Questo è un risultato significativamente differente rispetto a quello ottenuto nel presente studio. Mediante RQA si osserva che le dinamiche del cammino dei bambini di età intermedia non risultano regolari come quelle che caratterizzano il gruppo dei più grandi e non mostrano nemmeno maggiore regolarità rispetto al gruppo dei più piccoli. Al contrario, nello studio di Alcan, la DistEn indica che il gruppo di età intermedia risulta statisticamente distinto rispetto agli altri due, specialmente rispetto al primo gruppo. Ciò è indice di una maggiore regolarità nelle serie temporali del secondo gruppo.

Alcan sostiene, infine, che una ragione per cui si ha una riduzione di entropia in gruppi di età diverse può essere legata alla velocità del cammino. Infatti, i soggetti considerati sono stati chiamati a camminare a velocità libera

secondo la loro età e la loro struttura anatomica invece di una velocità fissata. Man mano che l'età aumenta dal gruppo 1 al gruppo 3, la velocità del passo potrebbe aumentare e abbiamo già constatato in altre indagini che gli intervalli di stride sono particolarmente sensibili ai cambiamenti di velocità [5]. Nello studio condotto da Alcan si arriva alla conclusione che la diminuzione di entropia dovuta all'aumento dell'età è conseguenza di un processo di sviluppo motorio. Infatti, tra i bambini tale crescita è caratterizzata da un aumento di agilità, adattabilità e abilità di mettere in atto sequenze motorie più complesse ed efficienti. In accordo con quanto appena detto, le serie temporali di soggetti più adulti saranno sempre più regolari.

CAPITOLO 5

Conclusione

In questo studio, le serie temporali composte da intervalli di stride rilevate da bambini sono state valutate mediante metodi non-lineari. In particolare, tramite DFA e RQA, si sono analizzate tali serie temporali e come esse variano in relazione alla maturazione dei bambini in termini di complessità e regolarità.

I metodi impiegati hanno portato a risultati che vanno discussi. L'indagine ha permesso di ritenere l'analisi con DFA poco utile ad individuare caratteristiche nelle serie temporali ottenute dalle durate degli stride dipendenti dall'età. Infatti, l'indice di scala frattale individuato nella DFA non differenziava in maniera netta i vari gruppi.

Al contrario, l'impiego dell'RQA ha condotto ad un esito confortante rispetto all'obiettivo. I valori dei parametri ottenuti diminuiscono con l'aumento dell'età e la diminuzione di tali indici segnala un aumento di regolarità delle serie temporali man mano che si avanza con l'età. Ciò è particolarmente evidente quando la differenza di età è elevata, mentre per fasce di età intermedie tale caratterizzazione non può essere considerata. In relazione ad altri studi, quindi, la maturazione di un individuo va di pari

passo con la maturazione motoria dell'individuo stesso e questo è visibile dall'incremento di regolarità delle serie temporali di stride-to-stride.

BIBLIOGRAFIA

[1] Jeffrey M. Hausdorff, C.-K. Peng, Zvi Ladin, Jeanne Y. Wei, Ary L. Goldberger (1995). *Is walking a random walk? Evidence for long-range correlations in stride interval of human gait.*

[2] Jeffrey M. Hausdorff (2009). *Gait dynamics in Parkinson's disease: Common and distinct behavior among stride length, gait variability, and fractal-like scaling.*

[3] Frédéric Dierick, Charlotte Vandevoorde, Frédéric Chantraine, Olivier White, Fabien Buisseret (2020). *Benefits of nonlinear analysis indices of walking stride interval in the evaluation of neurodegenerative diseases.*

[4] Yunfeng Wu, Pinnan Chen, Xin Luo, Meihong Wu, Lifang Liao, Shanshan Yang, Rangaraj M. Rangayyan (2015). *Measuring signal fluctuations in gait rhythm time series of patients with Parkinson's disease using entropy parameters.*

[5] M. Costa, C.-K. Peng, Ary L. Goldberger, Jeffrey M. Hausdorff (2003). *Multiscale entropy analysis of human gait dynamics.*

- [6] Pooja Prabhu, A.K. Karunakar, H. Anitha, N. Pradhan (2018). *Classification of gait signal into different neurodegenerative diseases using statistical analysis and recurrence quantification analysis.*
- [7] V. Alcan (2020). *Nonlinear Analysis of Stride Interval Time Series in Gait Maturation Using Distribution Entropy.*
- [8] B. Mandelbrot (1975). *Gli oggetti frattali: forma, caso e dimensione.*
- [9] Richard E.A. van Emmerik, Scott W. Ducharme, Avelino C. Amado, Joseph Hamill (2015). *Comparing dynamical systems concepts and techniques for biomechanical analysis.*
- [10] S. M. Pincus (1991). *Approximate Entropy as a measure of system complexity.*
- [11] Richman JS, Moorman JR (2000). *Physiological time series analysis using approximate entropy and sample entropy.*
- [12] Li P. Liu C. Li K. Zhe et al. (2008). *Assessing the complexity of short-term heartbeat interval series by distribution entropy.*