Original article

A Nonlinear Identification Approach for Gait Analysis of Amyotrophic Lateral Sclerosis Patients

Enfoque de identificación no lineal para análisis de la marcha en pacientes con esclerosis lateral amiotrófica

Tania Yadira Aznielle[1] 0000-0002-3345-352X
Jose Luis Hernandez Caceres[1*] 0000-0002-4406-444X

[1] Centro de Neurociencias de Cuba. La Habana, Cuba.

[*] Autor para la correspondencia: jose.caceres@cneuro.edu.cu

ABSTRACT

Gait data analysis, is giving mixing results regarding locomotion changes associated to Amyotrophic Lateral Sclerosis (ALS) development; the need has been claimed for new tools. We applied a nonlinear identification approach to the study of gait data from both healthy and ALS patients, available from Physionet.org. Kernel nonparametric nonlinear autoregression allowed to obtain noise-free realizations (NFR) that mimicked original traces, though correlation between original data and corresponding NFR was lower among ALS patients (p=0.03), suggesting a higher contribution of stochastic influences. Visual inspection of phase portraits, reconstructed from NFR via Takens theorem application, suggested dynamics differences between control subjects and patients. This was confirmed when phase portrait features were quantified and submitted to discriminant analysis (89% of correct classifications; 24/28). Application of a nonlinear dissimilarity measure for comparing pairs gait recordings, defined as a distance between underlying nonlinear autoregressive functions allowed an excellent separation between ALS and controls, via multidimensional scaling. Obtained projection map clearly suggested that ALS traces lay in a narrower dynamical space. This might reflect the known fact about neuronal degeneration accompanying ALS progression. When dissimilarity matrix principal components were introduced as predicting variables, discriminant analysis yielded an 82% of correct classifications (23/28). Overall, our results suggest that a nonlinear identification approach,

centered in the characterization of the dynamics of the gait process can bring new insights to gait data interpretation.
Palabras clave: gait data analysis; non linear approach; Takens' theorem.

RESUMEN

El análisis de datos de la marcha, está dando resultados mixtos con respecto a los cambios de locomoción asociados con el desarrollo de la esclerosis lateral amiotrófica (ELA). Se ha reivindicado la necesidad de nuevas herramientas de análisis de datos de la marcha. Aplicamos un enfoque de identificación no lineal al estudio de los datos de la marcha de pacientes sanos y con ELA, disponibles en Physionet.org. La auto-regresión no lineal no paramétrica del núcleo, permitió obtener realizaciones libres de ruido (NFR) que imitaban las trazas originales, aunque la correlación entre los datos originales y la NFR correspondiente fue menor entre los pacientes con ELA (p = 0,03), lo que sugiere una mayor contribución de las influencias estocásticas. La inspección visual de los retratos de fase, reconstruidos a partir de NFR mediante la aplicación del teorema de Takens, sugirió diferencias dinámicas entre los sujetos de control y los pacientes. Esto se confirmó cuando se cuantificaron las características del retrato de fase y se sometieron a un análisis discriminante (89 % de clasificaciones correctas; 24/28). La aplicación de una medida de disimilitud no lineal para comparar registros de marcha de pares, definida como una distancia entre funciones auto-regresivas no lineales subyacentes, permitió una excelente separación entre ALS y controles, a través de una escala multidimensional. El mapa de proyección obtenido sugirió claramente que las huellas de ALS se encuentran en un espacio dinámico más estrecho. Esto podría reflejar el hecho conocido sobre la degeneración neuronal que acompaña a la progresión de la ELA. Cuando se introdujeron los componentes principales de la matriz de disimilitud como variables predictoras, el análisis discriminante arrojó un 82% de clasificaciones correctas (23/28). En general, nuestros resultados sugieren que un enfoque de identificación no lineal, centrado en la caracterización de la dinámica del proceso de la marcha, puede aportar nuevos conocimientos a la interpretación de los datos de la marcha.
Palabras clave: análisis de la marcha; enfoque no lineal; teorema de Takens.

Recibido: 13/02/2023
Aprobado: 25/04/2023

Introduction

It is generally accepted that the analysis of human gait rhythm provide important diagnostic information that can be used to distinguish particular disorders of motor or sensory function, or even specific neurodegenerative diseases. [1]

Besides clinical assessment of patients by experts, model-based approaches might contribute to a numerical allocation of patients. However, our insufficient knowledge about physiological mechanisms of human/vertebrate locomotion, as well as certain technical aspects appear as drawbacks on this endeavor. [2], [3]

Most authors have found increases in stride variability among Amyotrophic Lateral Sclerosis (ALS) patients as well as with mouse models for the disease. [4-9] At the same time, some reports have revealed no changes in gait dynamics of the SOD1 G93A transgenic model for ALS compared to wild-type control mice. [10-11] Contrastingly, Hampton found the gait to be more "athletic" in SOD1 G93A mice at 12 weeks and 13 weeks of age, as compared to saline-treated mice, with the stride length significantly longer for the ALS model. [12]

The nonlinear identification approach combines advanced statistical methods (such as kernel nonparametric estimation and minimal cross validation criteria), with the theoretical advantages of nonlinear dynamics theory. In particular, Takens´ theorem states the possibility to reconstruct a complex dynamical system from discrete observations of only one of the several participating variables. [13], [14] This is principally suited for the case when little is known about the explicit mechanisms of a given process, or when not all involved variables can be traced. In particular, it allows plausible reconstruction of phase portraits corresponding to putative underlying dynamics. Phase portraits are a convenient way of representing time series data that is suited for both qualitative and quantitative assessment of time series data. Specially, phase portraits have been used for representing complex gait data. [15]

This work is an attempt to apply a nonlinear identification approach to the analysis of gait data obtained from ALS patients.

Methods

Data.Human gait time series data, were downloaded from www.physionet.org [6] and corresponded to 15 control subjects and 13 persons with ALS, as they walked at their normal pace along a 77-m-long hallway for 5 min. The mean age (standard deviation) of the control and ALS subjects were 39 (18.5) and 56 (12.8) years, respectively. Haussdorf et al.

describe details about the recording protocol. [17] From the recorded force applied to the ground during walking, the time series of the right foot was sampled at 300Hz. For each subject, data corresponding to the first sixty seconds of walk were resampled to 30 Hz. The obtained 1440 data points were subdivided into three non-overlapping windows, 16 seconds each. Each 16-s segment was submitted to nonlinear kernel autoregressive analysis and dissimilarity function estimation. Data from the processed 16-s segment that yielded better mimicry respect to the original gait recording were taken for further analysis. Given the low incidence and short survival time of ALS (around 2.5 per 100 000 in industrialized countries and 0.8 in Cuba, with 36 months for median survival time), publicly available experimental data on ALS patients´ gait are scarce. [18] That´s why many studies on ALS gait analysis among ALS patients have been conducted with the database from Physionet. Even when age differences between controls and ALS patients may impose some limitations to the study, different authors consider that these are not disquieting enough as for invalidating studies with this database. [19]

Nonlinear identification

Here, we conceive a gait recording as a realization of the following non-linear stochastic model,

$$\frac{dx}{dt} = f(x(t)) + \varepsilon(t)$$
$$U_t = h(x_t) + n_t \qquad (1)$$

In (1), $x(t)$ is the activation level of a subpopulation of motor units involved in locomotion as a function of time (t). U_t represents the force exerted by the foot on the ground at time t and is understood as a transformed projection of x. System noise and observation noise are $\varepsilon(t)$ and n_t respectively. It has been shown theoretically that a system as (1) does generate the following non-linear stochastic autoregressive model: [20-22]

$$U_t = F(U_{t-1}, U_{t-2}, \ldots, U_{t-k}) + e_t \qquad (2)$$

The function F in equation represents the relation of exerted force at the moment t as depending on force values at preceeding moments (t-1, t-2, ...). Function F in (2) was estimated from the gait data time series. For that, a nonlinear nonparametric kernel autoregressive Nadaraya-Watson estimator with a Gaussian kernel was used.

$$\hat{F}(z_{t-1}, z_{t-2}, \ldots, z_{t-k}) = \frac{\sum_{i=k+1}^{N} U_i \prod_{j=1}^{k} K(|z_{i-j} - U_{i-j}|, w)}{\sum_{i=k+1}^{N} \prod_{j=1}^{k} K(|z_{i-j} - U_{i-j}|, w)} \quad (3)$$

Where $\hat{F}(z_{t-1}, z_{t-2}, \ldots, z_{t-k})$ is the estimator of the theoretical function, and $K(|z_{i-j} - U_{i-j}|, w)$ is a Gaussian kernel. The order *k* and the bandwidth parameter *w* were obtained by cross validation.

Noise free realizations (NFR) were obtained from the estimated function via recursive estimation of future values given past values. After visual inspection for correspondence between observed gait traces and NFR, correlation coefficients between NFR and corresponding gait data segment were estimated.

For phase portrait reconstruction, two-dimensional projections of attractors from each NFR were obtained using Takens´ theorem. Phase portraits´ features were quantified (main axis length, minor axis length, slope at the basis, angle at the apex, second loop length, axis ratio, and loop ratio) and used for data stratification via discriminant analysis.

Nonlinear Dissimilarity Measure

Hernandez Cáceres et al. proposed a nonlinear dissimilarity measure for comparing a pair of time segments. [23] This distance measure is defined by manipulating global models to quantify dynamic similarity of time series. The dissimilarity between two traces is conceived as the dissimilarity between the underlying dynamics. [24] With the aim of expressing this quantitatively, it is interpreted as dissimilarity between those autoregressive functions that best fit the processes. For that, the concept of distance between functions is commonly used. We apply this concept to time series of gait data. This approach avoids the severe loss of information, which happens if a time series is summarized by a few characteristic numbers. [25]

If F_1 and F_2 are, respectively, the estimated nonlinear autoregressive functions corresponding to the gait data segments U^1 and U^2, then a measure of dissimilarity between these segments is defined as:

$$d(U^1, U^2) = \sum_{t=k+1}^{N1} \left| F_1(U^1_{t-1}, U^1_{t-2}, ..., U^1_{t-k}) - F_2(U^1_{t-1}, U^1_{t-2}, ..., U^1_{t-k}) \right| +$$

$$\sum_{t=k+1}^{N2} \left| F_1(U^2_{t-1}, U^2_{t-2}, ..., U^2_{t-k}) - F_2(U^2_{t-1}, U^2_{t-2}, ..., U^2_{t-k}) \right|$$

The two terms on the right side are, the distances between two sequences of one-step predictions. Nonlinear autoregressive estimation here was obtained by approximating the current value with a second-degree polynomial of the previous value (k=1).
All possible pairs of traces were compared, and a 28 by 28-dissimilarity matrix was obtained. For visualizing respective positions of individual gait recordings in the multidimensional function space, a 3D plot was obtained via multidimensional scaling.

For numerical classification, the eight largest principal components from the dissimilarity matrix were introduced as predictors for discriminant analysis.

Results and Discussion

Noise-free realizations (NFR)

All analyzed traces, either from controls or from ALS patients, yielded NFR resembling the original recordings (Fig. 1). Quantitatively, correlations between NFR and original traces ranged from r=0.17 (p=0.018) in one ALS recording to r=0.99 (p= ≤ 0.00005) in one healthy control. The median correlations were 0.94 for controls and 0.84 for ALS patients (p=0.033, Mann-Whitney's "U" test). The lower correlation between NFR and original traces can suggest a higher contribution of stochastic components to gait dynamics among ALS patients.

In figure 2, a set of 24 phase portraits corresponding to 12 controls and 12 ALS patients is shown. Visual inspection suggests different patterns for controls and patients.

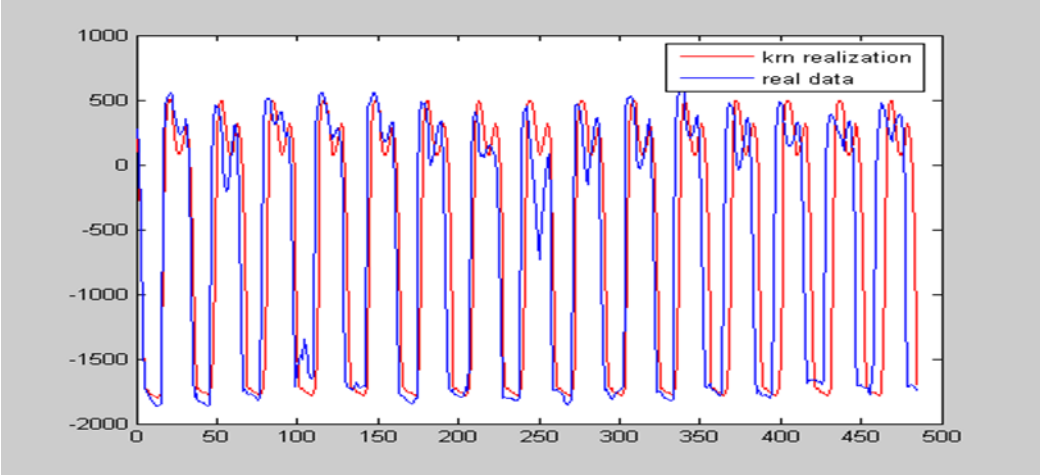

Fig.1- Time series of an original 16-s gait trace (blue) and corresponding NFR (red). Note that the NFR retained the amplitude, approached the frequency and showed the general pattern of the original trace.

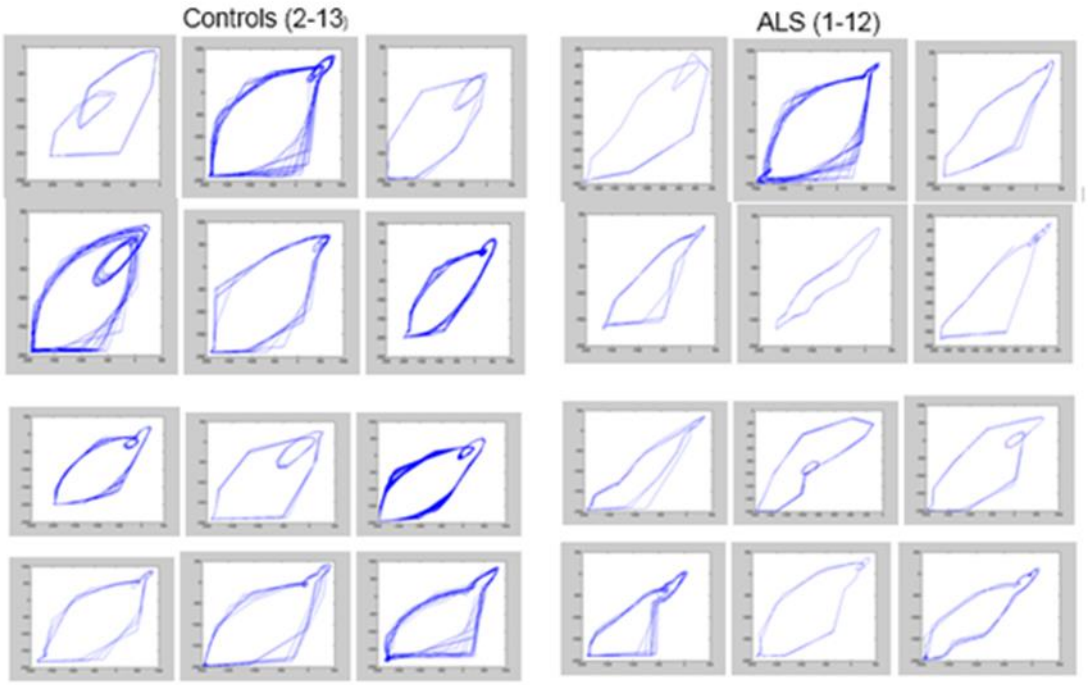

Fig. 2- Phase portraits obtained from 12 controls (left panel) and 12 ALS patients.

We attempted a classification based on phase portraits. For that, a group of phase portrait features were introduced as predictive variables for discriminant analysis: Major Axis Length, Minor Axis Length, Major Axis of the Second Loop, Slope at the Basis, Angle at the Apex, Main-to-Minor Axes Ratio, and First-to-Second Loop Ratio. The following classification matrix was obtained (Table 1). The obtained classification function is shown in table 2.

Table 1- Classification Matrix for Gait data phase portraits via discriminant analysis with phase portrait features as predicting variables; Rows: Observed classifications; Columns: Predicted classifications

	Percent Correct	**Controls p=.53571**	**ALS p=.46429**
Controls	93.33	14	1
ALS	84.62	2	11
Total	89.29	16	12

Table 2- Classification Functions based on discrimination of phase portrait with feature indices as predicting variables, and classification (Control/ALS) as grouping variable.

	Controls p=.53571	**ALS p=.46429**
Major Axis	.0518	.0423
Angle at Apex	.2222	.1479
Axes Ratio	13.7915	12.7231
Constant	-57.1963	-40.4804

As shown, only three of the seven proposed features were incorporated into the discriminant model.

Assessing Dissimilarities

Comparisons between traces corresponding to each health condition revealed different degrees of inter-individual variability among groups, ranging from an average of 241 a. u. among ALS patients to 291 a. u. among controls. Nonparametric comparisons revealed significant differences between groups ($p \leq 0.0001$), suggesting a wider variability for controls.

Multidimensional scaling confirmed the presence of narrower region in the functions´ space for ALS patients. Controls appeared as scattered over a wider area around (Fig. 3). As it can be noticed, there is a clear separation between the two conditions.

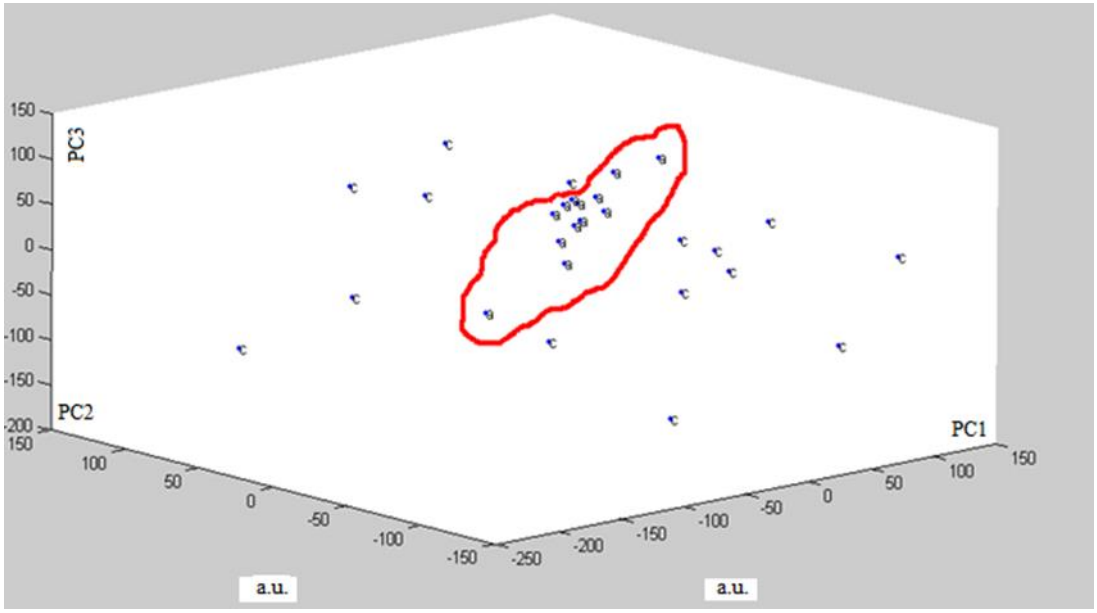

Fig. 3- Multidimensional scaling from the obtained dissimilarity matrix. ALS traces appear grouped in a smaller region, flanked by healthy controls.

Classification based on principal components

Results from multidimensional scaling suggest that it is possible to stratify between healthy subjects and patients suffering from ALS. Discriminant analysis with the first 8 principal components of the dissimilarity matrix as predictors showed an eighty-two percent of correct classifications (Table3).

Table 3- Classification Matrix with principal components from dissimilarity matrix as predicting variables. Rows: Observed classifications; Columns: Predicted classifications.

	Percent Correct	**Controls p=.53571**	**ALS p=.46429**
Controls	84.62	12	3
ALS	80.00	2	11
Total	82.14	14	14

Overall, our results suggest that a nonlinear identification approach, centered in the characterization of the dynamics of the gait process can bring new insights to gait data

interpretation. In particular, they suggest that amyotrophic lateral sclerosis is associated with a specific pattern of phase portraits as well as a restriction of choices in the repertory of dynamics. This approach might be a promising aid for the stratification of gait disorders associated to neurodegenerative diseases, as well as for implementing rehabilitation protocols.

Unlike previous research with gait data, we centered ourselves on the crude, unprocessed traces, and limited our analysis to a 16-s segment taken from the first minute of recording. We expected that this might allow us to retain information about the global mechanisms of gait dynamics. At the same time, recording time reduction to only 1 minute can simplify test procedures and reduce discomfort to patients.

Comparing our data with other reports with this data base revealed that we obtained 89% of correct classifications with phase portrait features and 82% accuracy with dissimilarity matrix principal components. These is comparable to the obtained by Wa and Shi using a Support Vector Machine algorithm (93.1% accuracy with all-training-all-testing; and 82.8 with leave-one-out). [26] We interpret this as a support for nonlinear identification as a way to retain important information about gait dynamics and ALS associated changes.

As a paradigm in nonlinear identification, we conceive the gait process as emerging from a nonlinear dynamical system perturbed by innovation noise. This combination of deterministic and random factors result in the human gait. Ideally, if stochastic innovation were absent, the noise free realization of a periodic attractor should be identical to the original time series. The fact that correlations between NFR and original traces were lower than unity could suggest about stochastic contributions. In this regard, the lower correlations obtained among ALS patients might suggest that healthy persons do display a more robust gait in terms of constancy of pace, with lower influence from environment´s stochastic factors.

As stated by Wu and Shi, “Although a few nonlinear analysis methods have been used in previous studies to investigate the complexity of human locomotion process, further quantitative studies call for more computational tools that are suited for gait analysis.”[26] We hope that the approach proposed here can contribute to fill this gap.

Our results are in agreement with those of Haussdorf et al. They found a degradation of complexity in ALS patients. In particular, they found that gait of healthy humans remains self-similar at a broad range of time scales; this self-similarity is disrupted in ALS patients, and the result could be a higher sensitivity to external environmental disturbances. [17]

Studies revealed that in ALS, about half of neurons involved in locomotion become dysfunctional. Surprisingly, animals can still walk in these conditions. [15], [27] Among the

putative compensations allowing individuals with ALS to continue moving, the following strategies have been suggested: Increasing muscle co-contraction of lower leg muscles, circuit adjustment for maintaining posture and supporting gait, and a stronger initial activity of TA and GC muscle groups at step initiation. [28] These re-adjustments may result in a less robust dynamics, with higher influence of stochastic influences and corresponding increased gait variability. The lower correlation between NFR and the original signal among ALS patients might suggest an enhanced role for stochastic components, with corresponding deterioration of self-affine homeostatic mechanisms.

References

1. Reeder B, Whitehouse K. Sensor-Based Detection of Gait Speed in Older Adults. An Integrative Review. Res Gerontol Nurs. 2015; 8(1):12-27.
2. Kuo AD. The six determinants of gait and the inverted pendulum analogy: A dynamic 3. Buczek FL, Cooney KM, Walker MR, Rainbow MJ, Concha MC, Sanders JO. Performance of an inverted pendulum model directly applied to normal human gait. Clinical Biomechanics (Bristol, Avon) 2006; 21(3):288-296
4. Kiernan MC et al., Amyotrophic lateral sclerosis. The Lancet 2011; 377 (4): 645–646
5. Sreedharan J, Brown RH. Amyotrophic lateral sclerosis: Problems and prospects. Ann Neurol 2013; 74(3):309–316.
6. Harms MB, Baloh RH. Clinical neurogenetics: Amyotrophic lateral sclerosis. NeurolClin 2013; 31(4):929–950.
7. Liu Y, Pattamatta A, Zu T, Reid T, Bardhi O, Borchelt DR, Yachnis AT, Ranum LP. C9orf72 BAC mouse model with motor deficits and neurodegenerative features of ALS/FTD. Neuron 2016; 90:521–534.
8. Barnéoud P, Lolivier J, Sanger DJ, Scatton B, Moser P (). Quantitative motor assessment in FALS mice: a longitudinal study. Neuroreport 1997; 8(13):2861-2865.
9. Fischer LR., Culver DG, Tennant P, Davis AA, Wang M, Castellano-Sanchez A, Khan J, Polak MA, Glass JD. Amyotrophic lateral sclerosis is a distal axonopathy: evidence in mice and man. Experimental neurology 2004; 185(2):232-240.
10. Guillot TS, Asress SA, Richardson JR, Glass JD, Miller GW. Treadmill Gait Analysis Does Not Detect Motor Deficits in Animal Models of Parkinson's Disease or Amyotrophic Lateral Sclerosis. J Mot Behav 2008; 40(6): 568–577
11. Amende I, Kale A, McCue S, Glazier S, Morgan JP, Hampton TG. Gait dynamics in mouse models of Parkinson's disease and Huntington's disease. Journal of Neuro Engineering and Rehabilitation 2005, 2:20-28

12. Hampton TG. Measurement of gait dynamics and use of beta-blockers to detect, prognose, prevent and treat amyotrophic lateral sclerosis. Patent US 20070021421, Publication date Jan 25, 2007
13. Takens, F. Detecting strange attractors in turbulence. Lecture notes in Mathematics 1981; 898: 366-381
14. Noakes L. The Takens embedding theorem. Int. J. Bifurcation Chaos 1991; 867:867-872
15. Hadzipasic M, Ni W, Nagy M, Steenrod N, McGinley MJ, Kaushal A, Thomas E, McCormick DA, Horwich AL. Reduced high-frequency motor neuron firing, EMG fractionation, and gait variability in awake walking ALS mice. Proc Natl Acad Sci USA 2016 ????, E7600–E7609
16. Goldberger AL et al., PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. Circulation 2000; 101 (23): E215–E220
17. Hausdorff JM, Lertratanakul A, Cudkowicz ME, Peterson AL, KalitonD, Goldberger AL. Dynamic markers of altered gait rhythm in amyotrophyic lateral sclerosis. J Appl Physiol 2000; 88 (6):2045–53
18. Zaldivar T, Gutierrez J, Lara G, Carbonara M, Logroscino G, Hardiman O. Reduced frequency of ALS in an ethnically mixed population. A population-based mortality study. Neurology 2009; 72(19):1640-1645
19. Ran P, Tang Z, Fang F, Lou L, Zu L, Bringas-Vega ML, Yao D, Kendrick KM, Valdes-Sosa PA. Gait Rhythm Fluctuation Analysis for Neurodegenerative Diseases by Empirical Mode Decomposition. IEEE Transactions on Biomedical Engineering 2017,
20. Valdés-Sosa PA, Bosch J, Jimenez JC, Trujillo-Barreto NJ, Biscay-Lirio RJ, Morales F, Hernández JL, Ozaki T: The statistical identification of nonlinear brain dynamics: A progress report. Nova Science Lots, 01/1999;
21. Hernández Cáceres JL, Hernández Martínez L, Pérez Monzón M, García Domínguez L: Nonlinear Properties of Measles Epidemic Data Assessed With A Kernel Nonparametric Identification Approach. Electronic Journal of Biomedicine 2006;
22. Hernandez JL, Valdes PA, Vila P. EEG spike and wave modelled by a stochastic limit cycle. Neuroreport 1996; 7(2):2246-2250
23. Hernández JL, Biscay R, Jimenez JC, Valdes P, Grave de Peralta R. Measuring the dissimilarity between EEG recordings through a non-linear dynamical system approach. International Journal of Bio-Medical Computing 1995; 38(2):121-129
24. Liu X, Povinelli RJ, Johnson MT Detecting determinism in speech phonemes, proceedings of IEEE Signal Processing Society 10th Digital Signal Processing Workshop, 2002.
25. Schreiber T, Schmitz A. Classification of Time Series Data with Nonlinear Similarity Measures. Physical Review Letters 1997; 79 (8): 1475-1478
26.Wu WF, Krishnan S, Computer-aided analysis of gait rhythm fluctuations in amyotrophic lateral sclerosis. Medical Biol Eng Comput 2009; 47(11):1165–1171
27. Hadzipasic M, et al. Selective degeneration of a physiological subtype of spinal motor neuron in mice with SOD1-linked ALS. Proc Natl Acad Sci USA 2014; 111(47):16883–16888;

28. Gorassini MA, Norton JA, Nevett-Duchcherer J, Roy FD, Yang JF. Changes in locomotor muscle activity after treadmill training in subjects with incomplete spinal cord injury. J Neurophysiol 2009; 101 (2):969–979.

Interest's conflicts

Authors declare no interest's conflicts.

Contribution of authors

Both authors contributed equally to the paper.