Linear mixed-effects models for the analysis of high-density electromyography with application to diabetic peripheral neuropathy

Andreis, Felipe Rettore; Favretto, Mateus Andre; Cossul, Sandra; Nakamura, Luiz Ricardo; Barbetta, Pedro Alberto; Marques, Jefferson Luiz Brum

Published in:
Medical & Biological Engineering & Computing

DOI (link to publication from Publisher):
10.1007/s11517-020-02181-1

Publication date:
2020

Document Version
Accepted author manuscript, peer reviewed version

This is a post-peer-review, pre-copyedit version of an article published in Medical & Biological Engineering & Computing. The final authenticated version is available online at: https://doi.org/10.1007/s11517-020-02181-1

Linear Mixed-Effects Models for the Analysis of High-Density Electromyography with Application to Diabetic Peripheral Neuropathy

Felipe Rettore Andreis[1,2], Mateus Andre Favretto[1], Sandra Cossul[1], Luiz Ricardo Nakamura[3], Pedro Alberto Barbetta[3], Jefferson Luiz Brum Marques[1].

[1] Institute of Biomedical Engineering, Department of Electrical and Electronic Engineering, Federal University of Santa Catarina, Florianópolis, Brazil
[2] Center for Neuroplasticity and Pain (CNAP), SMI, Department of Health Science and Technology, Aalborg University, Aalborg, Denmark
[3] Department of Informatics and Statistics, Federal University of Santa Catarina, Florianópolis, Brazil

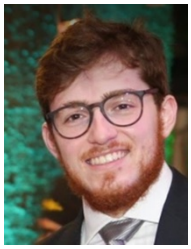
E-mail address: feliperandreis@gmail.com
Telephone and fax numbers: +45 91724244
Total number of words of the manuscript: 7246
Number of words of the abstract: 211
Number of figures: 4
Number of tables: 4

Abstract:

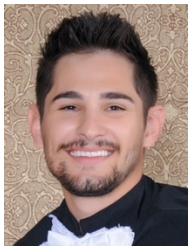
This article demonstrates the power and flexibility of linear mixed-effects models (LMEM) to investigate high-density surface electromyography (HD-sEMG) signals. The potentiality of the model is illustrated by investigating the root mean squared value of HD-sEMG signals in the tibialis anterior muscle of healthy (n = 11) and individuals with diabetic peripheral neuropathy (n = 12). We started by presenting the limitations of traditional approaches by building a linear model with only fixed-effects. Then, we showed how the model adequacy could be increased by including random-effects, as well as by adding alternative correlation structures. The models were compared with the Akaike information criterion and the Bayesian information criterion, as well as the likelihood ratio test. The results showed that the inclusion of the random-effects of intercept and slope, along with an autoregressive moving average correlation structure is the one that best describes the data ($p < .01$). Furthermore, we demonstrate how the inclusion of additional variance structures can accommodate heterogeneity in the residual analysis and therefore increase model adequacy ($p < .01$). Thus, in conclusion, we suggest that adopting LMEM to repeated measures such as electromyography can provide additional information from the data (*e.g.*, test for alternative correlation structures of the RMS value), and hence provide new insights into HD-sEMG related work.

Keywords: Linear mixed-effects models; High-density surface electromyography; Diabetic peripheral neuropathy.

Autor biography

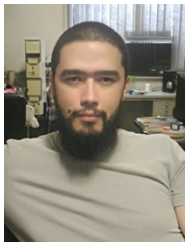


Felipe Rettore Andreis BEng and M.Sc. Electrical Engineering is a PhD Fellow at the Department of Health Science and Technology, Aalborg University.



Mateus André Favretto M.Sc. in Electrical Engineering, sub-area Biomedical Engineering (UFSC, Brazil, 2017); B.Sc. Electrical Engineering (UPF, Brazil, 2014). He is currently a PhD student.

Sandra Cossul M.Sc. Electrical Engineering, sub-area Biomedical Engineering (UFSC, 2018); B.Sc. Computer Engineering (UFSM, 2015). Currently, she is a PhD Student within the biomedical engineering research area.



Luiz R. Nakamura is a BSc., MSc. and PhD in Statistics. He is a senior lecturer in Statistics at the Federal University of Santa Catarina, Brazil. Dr. Nakamura does research in statistical modelling.

Pedro Alberto Barbetta is graduated in Statistics, doctor in Production Engineering and works at Federal University of Santa Catarina, Brazil. Special interest in educational assessment, hierarchical models and design of experiments.

Jefferson Luiz Brum Marques BEng and MSc Electrical Engineering, and PhD Medical Physics & Clinical Engineering (Sheffield-UK), is a Professor in Biomedical Engineering at the Federal University of Santa Catarina, Department of Electrical and Electronic Engineering (Florianópolis-Brazil).

1 Introduction

Diabetic peripheral neuropathy (DPN) is the most common type of neuropathic syndrome associated with diabetes mellitus. DPN results from a combination of various metabolic and vascular factors, and it is associated with several complications, such as distal weakness, premature fatigue, reduced muscle strength, foot ulcers and in severe cases lower limb amputation [1, 2]. DPN is a distal, length-dependent condition, affecting mainly muscles with a higher proportion of type I fibres, such as the tibialis anterior [3, 4].

The high-density surface electromyography (HD-sEMG) technique is a useful tool to provide information of an individual neuromuscular system, with the main advantages of being non-invasive and having a high spatial resolution due to a large number of closely spaced electrodes. The HD-sEMG reflects important physiological mechanisms such as the number of active motor units and their firing rate, as well as muscle fibre conduction velocity [5]. The technique has recently demonstrated to be a promising tool to assess diabetes mellitus and its complications [6–8].

The relationship between EMG variables and muscle fibre conduction velocity is well established during an isometric voluntary contraction. A reduction in the muscle fibre conduction velocity results in a positive linear trend in the amplitude variables (*e.g.*, the root mean square), and a

negative trend in the spectral variables (e.g., median frequency), reflecting a phenomena termed myoelectric manifestations of muscle fatigue [9]. Therefore, this relationship may be an important biomarker for assessing DPN individuals considering that premature muscle fatigue is strongly associated with a decrease in the muscle fibre conduction velocity. Therefore, DPN individuals may present the effects of myoelectric manifestations of muscle fatigue sooner than healthy individuals [10].

In high-density surface electromyography (HD-sEMG) experiments, the signal is recorded when an individual performs a specific task for a period that varies from a few seconds up until a few minutes [11]. This signal can be considered as a case of repeated measures data, where multiple observations are collected in each research unit (*i.e.*, subject). Therefore, such research design imposes a hierarchical or multilevel structure in the data, where multiple measurements are nested within each subject [12, 13].

Observations from the same individual tend to be "more alike" than observations from different subjects, which naturally induces a correlation between measures within the same individual. Consequently, all observations within a subject do not contribute to completely independent information. This fact results in the violation of independence assumption inherent in traditional statistical methods such as analysis of variance (ANOVA) and standard regression analysis. If the observations are incorrectly treated as independent, it may result in underestimated standard errors and inflation of the type I error rate (*i.e.*, false positives) [14–16].

A common approach to handle such data structure is to use summary statistics to condense the repeated measured information into a single measurement, eliminating the within-subject correlation and allowing for a comparison of the summary statistics between groups using standard statistical techniques. A classic example of this approach is to perform a regression analysis on each individual and then perform standard statistical techniques to compare the mean intercept or slope of subjects within different groups. Although this approach is statistically valid, it does not take full advantage of the data, resulting in a significant loss of information and statistical power, given that the variability of the single-subject regression is not taken into account for further analysis [15, 17].

Another common approach is to analyse HD-sEMG data through repeated-measures ANOVA. Besides the normality of the residuals, repeated-measures ANOVA poses restrictive assumptions about the correlation patterns (*i.e.*, sphericity), since it requires equal variances at each time point (homoscedasticity) and an equal correlation between measurements at different times (*e.g.*, the correlation between measurements at time t_1 and t_2 is assumed to be the same as the correlation between measurements at time t_1 and t_n). This correlation structure is commonly not the case in repeated measures experiments since measures that are taken close together in time are usually more highly correlated than measurements taken far apart in time. If these assumptions are violated, it may lead to incorrect decisions. To be correctly applied, repeated-measures ANOVA also requires balanced data (*i.e.*, it does not allow unequally spaced time intervals) and complete observations on every subject. Hence, if a single observation is missing, the individual must be excluded from the analysis [18, 19].

To deal with the limitations of the previously described methods, Laird and Ware [20] introduced the linear mixed-effects models (LMEMs), also known as multilevel models, hierarchical linear models, variance component models, among others. This class of model provides a flexible and powerful tool for the analysis of grouped data, which includes a range class of data structures such as longitudinal data, repeated measures, blocked designs and hierarchical data [20–22]. The basic principle of LMEM is that individuals in the population are assumed to have their specific trajectories over time (*i.e.*, individuals within a population are heterogeneous). Therefore, it combines individual growth curves and a sample average curve into a single model. Besides that, LMEM permits several ways to deal with heteroscedastic data and also provides many alternative correlation structures, enabling the researcher to investigate with much more detail possible patterns of the observed data, also allowing to model unbalanced and missing data. Finally, using LMEMs result in higher statistical power to detect group-level differences as well as narrower confidence intervals in the parameter estimates [18, 19, 22].

A drawback of LMEM is that setting up the models is not as straightforward as traditional methods. LMEM requires not only the specification of main effects (*i.e.*, fixed effects) and their interactions but also the random-effects as well as the additional correlation structures, therefore the

name mixed-effects [23, 24]. Fixed effects represent the expected (*or mean*) values of the observations, shared by all observations in the sample. On the contrary, random-effects define individual departures from the intercept and slope of the average model. The random-effects result from the deviation between subjects and within-subjects, which is a natural way to account for the biological variability in the parameters across subjects [13].

In typical repeated measures experiments, changes are related to covariates, such as subjects assigned to active drug and placebo, and multiple observations are performed on each subject. This specific case results in two fixed effects, treatment and time. The LMEM allows the inclusion of time-varying or time-invariant covariates in the same way as conventional regression; examples of potential covariates are gender or age. Despite the complexity of such models, recently developed statistical software [25] makes such models accessible to researchers, and they are being used in a wide range of problems, including medicine [26], athletic training [27] and education [28]. That said, the LMEM can also be a useful tool to investigate HD-sEMG signals, given that it consists of repeated measures, where multiple measurements are nested within each subject.

The objective of this work was to outline the LMEM approach to analyse the root mean square (RMS) value of HD-sEMG signals in an application regarding healthy and subjects with DPN. The data consist of two groups of subjects: (i) healthy subjects (control) and (ii) subjects with DPN. Due to physiological alterations occasioned by the DPN like premature fatigue, the RMS value is expected to reflect such differences between healthy and DPN individuals. Thus, the research questions we ask are: (i) Do RMS values change over time? (ii) Do RMS values vary between the two groups? (iii) Do RMS values differ in the rate of change between the two groups? (iv) Is it necessary to account for random-effects to model such data? (v) Does the inclusion of alternative correlation structures increase the model adequacy? Furthermore, (vi) Is it necessary to adopt complex variance structures?

2 Methods

2.1 Subjects and Experimental Protocol

Data were collected from adults of both genders, 12 healthy subjects (Control) and 11 subjects with DPN. The mean age was 49.8 years (SD 9.11) for the healthy subjects and 55.7 for the DPN group (SD 7.01). Participants were instructed to be seated, with the hip and knee flexed at 90° angle and the ankle in a neutral position at a 90° angle to the leg. The participants performed three maximum voluntary isometric contractions of ankle dorsiflexion with a 5-s duration, with a 2-minute rest interval between the tests. The highest value of the three contractions was used as a reference for a submaximal voluntary isometric contraction of ankle dorsiflexion in 50% of the maximum value. The contractions were maintained for 30-s.

The HD-sEMG signals were recorded in the tibialis anterior muscle. Before attaching the electrode array, the skin was cleaned with alcohol and gauze. It was used a two-dimensional array of 64 electrodes (ELSCH064NM4, OT Bioelettronica, Turin, Italy) with 13 rows and 5 columns with 1-mm diameter and inter-electrode distance of 4 mm. Of those, it was used a subset of 9 rows and 4 columns in a differential lead, resulting in 32 signals (Fig. 1). The signals were sampled at 2 kHz. The electrode array was fixed using fixation adhesive, and the cavities of the electrode were filled with conductive cream (CC1, OT Bioelettronica, Turin, Italy). The reference electrode was placed on the tuberosity of the tibia. The details of the experimental setup and the acquisition system have been previously reported in [29, 30].

Fig. 1 Sampling in a differential configuration column-wise. There are nine rows (R) and four columns (C) resulting in 32 signals [(R-1) x C]

The signals were digitally filtered in MATLAB 2018 (MathWorks, MA, USA) by an eight order Butterworth bandpass filter with a frequency range of 10-400 Hz. A second-order notch filter was also used, with frequency of 60 Hz and the next five harmonics. The window size for signal processing was 0.5-s without overlap. The individuals were sampled for 30 s, and the first 2-s and the

last 3-s of the signal were discarded, resulting in a signal with 25-s duration. Thus, we have 50-time points for each subject. This study was approved by the Human Research Ethics committee of the Federal University of Santa Catarina (Protocol Number: 2.390.994), and all participants signed the informed consent before participation.

2.2 Model Specification and Analysis

Initially, we examined which was the kind of relationship (*i.e.*, linear, quadratic, cubic) between the response variable and subjects. Fig. 2 reproduces the plots for all 23 individuals in the study, which indicates that a linear pattern is applicable to model how RMS value varies over time.

Fig. 2 Individual trajectories with fitted regression lines for the root mean square variable (F = Female, M = Male)

From Fig. 2, there is evidence of heterogeneity in the intercepts and slopes between the individuals, since the response variable starts at different values for different subjects and it appears to have different rates of change as well. Therefore, we tested the inclusion of a random intercept and a random slope for the subjects. After deciding for a linear trajectory, we followed the top-down strategy [31–33]. The first step of this approach is to model a saturated parameter specification for the fixed effects; in this case including time, group and group-by-time interaction (*i.e.*, whether the groups follow the same trend over time), resulting in *Model A*, as follows

Model A

$$y_{ij} = (\beta_0 + \beta_2 G_{2i}) + (\beta_1 + \beta_3 G_{2i})\text{Time}_{ij} + \epsilon_{ij}, \quad \epsilon_{ij} \sim N(0, \sigma_e^2),$$

where y_{ij} denotes the value of outcome measure at time j for individual i; G_{2i} is a binary variable taking the value 1 if the ith subject belongs to the DPN group; β_0 is the average intercept and β_1 is the average slope for subjects in the control group; β_2 is the average difference in the intercept between subjects in the DPN group compared to subjects in the control group, and β_3 is the average difference in slope between subjects in the DPN group when compared to subjects in the control group. Finally, ϵ denotes the residual errors.

Following the top-down strategy for model selection, the next step is to model between subjects' variation, as well as the correlation between measures at different times on the same subject. Therefore, we started by including a random intercept (b_{0i}) to *Model A*, resulting in *Model B*, given by

Model B

$$y_{ij} = (\beta_0 + \beta_2 G_{2i} + b_{0i}) + (\beta_1 + \beta_3 G_{2i})\text{Time}_{ij} + \epsilon_{ij}, \quad b_{0i} \sim N(0, \tau_{b0}^2), \quad \epsilon_{ij} \sim N(0, \sigma_e^2),$$

where b_{0i} accounts for the random intercept (*i.e.*, mean differences across subjects). By adding the random intercept, each person gets his or her deviation from the fixed intercept. With this model, the outcome variance, or error, is partitioned in two: the between-person mean differences (b_{0i}) and the within-person deviation (ϵ_{ij}).

The intraclass correlation coefficient (ICC) is often used to assess the necessity of including a random intercept. The ICC is an indicator of the amount of correlation in the data, the proportion of between-person variance to the sum of the between and within-subject variances, calculated as

$$ICC = \frac{\tau_{b0}^2}{\tau_{b0}^2 + \sigma_e^2},$$

and it ranges from zero to one, where zero represents no correlation between subjects and one represents a perfect correlation between subjects.

Adding only a random intercept implies that the residuals from the same person would have a constant correlation over time. This assumption may not be the case for repeated measures data, given that some individuals may change the RMS values at a higher or lower rate over time. Hence, the model can be further improved by adding a random effect of time in the model (b_{1i}). In this case, each individual is allowed to have his or her time slope deviation, resulting in *Model C*, that can be written as

$$\text{Model C}$$

$$y_{ij} = (\beta_0 + \beta_2 G_{2i} + b_{0i}) + (\beta_1 + \beta_3 G_{2i} + b_{1i})\text{Time}_{ij} + \epsilon_{ij}$$

$$b_i = \begin{bmatrix} b_{0i} \\ b_{1i} \end{bmatrix} \sim N(0, \psi), \epsilon_{ij} \sim N(0, \sigma_e^2),$$

where b_i is the vector of random-effects, assumed independent for different *i*; ϵ_{ij} is the within-group error, assumed independent for different *i*, *j* and independent of the random-effects. Residuals from b_{0i} and b_1 are assumed to be normally distributed with variances τ_{b0}^2 and τ_{b1}^2, respectively. The random intercept and random slope are often correlated; therefore, the model also estimates one more parameter, the correlation between b_{0i} and b_1, denoted by τ_{b01}.

After the inclusion of the random-effects, the assumptions of homoscedastic within-group errors have to be verified. If such an assumption does not hold, the LMEM can be extended by fitting models with different variance structures, such as different variances for each level of a stratification variable [22]. Furthermore, under the appropriate model, residuals must behave like uncorrelated noise. If the residuals do not follow this requirement, model predictions may be incorrect [22]. In this case, after fitting the models with random-effects, some degree of correlation among the residuals may remain, suggesting that the model did not fully account for the within-individual dependencies in the data. To solve this, LMEM can be extended by combining random-effects with time-specific patterns of correlations.

In the model selection process, we observed the empirical autocorrelation function for the residuals. If in the autocorrelation function, the residuals were significantly different from zero, additional correlation structures were taken into account to model the within-individual correlation and thus ensure accurate tests of the fixed effects [31].

In this work, we tested the models for three of the most commonly used alternative correlation structures for repeated measures and longitudinal studies, which are: unstructured (UN), first-order autoregressive (AR1) and first-order autoregressive-moving average (ARMA (1,1)). The UN makes no assumptions about the residual covariance, estimating the covariance directly from the data. Therefore, it is the most complex covariance structure and can be used as a baseline to compare to other simpler covariance structures. The AR1 is used for stronger correlation among adjacent measurements and can be extended to the ARMA structure model by including one more parameter. In the ARMA model, the correlation between observations declines at different rates compared to the AR structure. An example of UN, AR1 and ARMA (1,1) with four repeated measures is shown below.

$$\textbf{AR1} = \sigma_e^2 \begin{bmatrix} 1 & \rho & \rho^2 & \rho^3 \\ \rho & 1 & \rho & \rho^2 \\ \rho^2 & \rho & 1 & \rho \\ \rho^3 & \rho^2 & \rho & 1 \end{bmatrix}$$

$$\textbf{ARMA (1,1)} = \sigma_e^2 \begin{bmatrix} 1 & \gamma & \gamma\rho & \gamma\rho^2 \\ \gamma & 1 & \gamma & \gamma\rho \\ \gamma\rho & \gamma & 1 & \gamma \\ \gamma\rho^2 & \gamma\rho & \gamma & 1 \end{bmatrix}$$

$$\mathbf{UN} = \begin{bmatrix} \sigma_{e1}^2 & \sigma_{e21} & \sigma_{e31} & \sigma_{e41} \\ \sigma_{e21} & \sigma_{e2}^2 & \sigma_{e32} & \sigma_{e42} \\ \sigma_{e31} & \sigma_{e32} & \sigma_{e3}^2 & \sigma_{e43} \\ \sigma_{e41} & \sigma_{e42} & \sigma_{e43} & \sigma_{e4}^2 \end{bmatrix}$$

where ρ is the autoregressive parameter, λ is the moving average component and σ_e^2 the residual variance. A glossary of the terms used in all models presented is listed in Table 1.

Table 1 Glossary of symbols for the LMEMs

Symbol	Definition
y_{ij}	Value of outcome measures at time *j* for subject *i.*
β_0	Average intercept across all subjects from the control group.
β_1	Average slope across all subjects from the control group.
β_2	Average difference in the intercept between subjects in the DPN group compared to subjects in the control group.
G_{2i}	Binary variable taking the value 1 if the *i*th subject belongs to the DPN group.
β_3	Average difference in slope between subjects in the DPN group when compared to subjects in the control group.
Time_{ij}	Timing of the outcome measure on the *i*th subject at the *j*th measurement.
b_{0i}	Difference between the population-averaged intercept and the intercept for individual *i.*
b_{1i}	Difference between the population-averaged slope and the slope for individual *i.*
ϵ_{ij}	Residual error in outcome measures at time *j* for subject *i.*
σ_e^2	Variance of residuals ϵ_{ij}.
τ_{b0}^2	Variance of the random intercept.
τ_{b1}^2	Variance of the random slope.
τ_{01}	Correlation between the random intercept and the random slope.
ρ	Autoregressive parameter.
λ	Moving average component.

In the model selection process, we compared the models through the Akaike's information criterion (AIC) [34] and the Bayesian information criterion (BIC) [35]. Both AIC and BIC decrease with the goodness of fit.

As models were nested, we used a likelihood ratio test (LRT), where a reduced model is nested within a larger model. The two models differ only by the removal of variables from the more complex model. Using the *p*-value resulted from the LRT we decide if the inclusion of a specific term is reasonable (*i.e.*, if the *p*-value is significant, the variable in question improves the model fit and must be accounted for in the model). The LRT takes -2 times the difference in log-likelihoods which follows a chi-squared distribution with degrees of freedom equal to the difference in the number of estimated parameters between the models [22].

We considered here the restricted maximum likelihood (REML) method to estimate the model parameters, which was achieved through the *nlme* package [36] in the R software [37]. The considered significance level was 0.05. The validity of the models was examined using visual inspection of the model residuals.

3 Results

To answer whether it was necessary to account for the random-effects to model the data, we initially compared a random intercept model (*Model B*) and random intercept and slope model (*Model C*) to the model considering only fixed effects (*Model A*). *Model C* was further tested with *Model C1*, which considers random effects in intercept and slope; however, without the intercept and slope correlation. The results are listed in Table 2.

Table 2 Evaluation indices of each LMEM for the RMS value

Model	**AIC**	**BIC**	**Test**	**LRT**	***p*-value**

A	9840	9866	–	–	–
B	7244	7274	A vs B	2598	< .01
C	7050	7091	B vs C	197	< .01
C1	7051	7087	C vs C1	3.12	0.07

To assess the necessity of a random intercept, *Model A* was compared to *Model B*. The ICC of 0.91 indicates that the between-subjects' variability in the intercept needs to be accounted for in the model; thus, *Model B* was selected. This result is confirmed by the lower AIC and BIC for Model B when compared to *Model A*. As *Model A* is nested within *Model B*, this result was further confirmed by the result of the LRT, χ^2 (1) = 2598 (p < .01).

The necessity of including a random slope was performed by the comparison of *Model B* with *Model C*. *Model C* was preferred, due to the lower AIC and BIC values. As model B is nested within *Model C*, the LRT was also applied, resulting in a significant p-value, χ^2 (2) = 197 (p < .01). Thus, there is evidence for significant between-subjects' variability in the intercept. Also, in the rate of change of RMS value over time, as represented by *Model C*. Therefore, it was necessary to include random-effects to model such data, as hypothesized by item (iv) of the introduction section.

In *Model C*, the correlation between the random intercept and random slope was positive (τ_{01}= 0.39), indicating that individuals with higher levels of intercept also have higher rates of change. We tested the necessity of adding a parameter to control this feature by fitting a model that constrained the correlation between the random effects (τ_{b01}) to zero (named *Model C1*). Thus, *Model C1* is nested within *Model C* (corresponding to τ_{01} = 0). Table 2 shows that both AIC and BIC values disagree in their results because AIC is smaller for *Model C* whereas BIC is smaller for *Model C1*. The LRT favours the simpler model, *Model C1*, χ^2 (1) = 3.12 (p = 0.07). Therefore, we chose the most parsimonious model (*Model C1*) to our subsequent analysis.

We assessed the empirical autocorrelations to investigate the within-group correlation. We observed significant empirical autocorrelation in the first few lags, suggesting that a correlation structure must be accounted to model time-series dependence [38]. Therefore, we compared three standard options of alternative correlation structures. The results are presented in Table 3.

Table 3 Evaluation indices of alternative correlation structures for the RMS value models

Model	AIC	BIC	Test	LRT	*p*-value
C1	7051	7087	–	–	–
AR1	6889	6929	C1 vs AR1	164	< .01
ARMA (1,1)	6881	6927	AR1 vs ARMA (1,1)	9.23	< .01
ARMAH (1,1)	6440	6495	ARMA (1,1) vs ARMAH (1,1)	445	< .01

The unstructured (UN) model was not estimated due to a large number of parameters which caused convergence problems. Thus, we have assessed the correlation structures in ascending order (*i.e.*, from the least complex to the most complex model). The first comparison is between *Model C1* with a first-order autoregressive correlation structure (*Model AR1*). *Model AR1* estimates the random-effects of intercept and slope, as well as an autoregressive parameter (ρ), estimated as 0.38. The smaller values of AIC and BIC confirm that *Model AR1* is a substantially better fit of the data than the independent errors model. Since *Model C1* is nested within *Model AR1* (corresponding to ρ = 0), the LRT was also used, corroborating this result, χ^2 (1) = 164 (p < .01).

The comparison of *Model AR1* with *Model ARMA (1,1)* was then performed. There were smaller AIC and BIC values for the ARMA (1,1). Model AR1 is nested within *Model ARMA (1,1)* (corresponding to λ = 0), so the LRT was performed and confirmed this result, χ^2 (1) = 9.23 (p < .01). Hence, we shall select the ARMA (1,1) model over the AR1 model, confirming that the inclusion of alternative correlation structures increases the model adequacy, as hypothesized by item (v).

To assess the model adequacy, we checked *Model ARMA (1,1)* residuals and found that the variability presented a reasonably homogeneous pattern through time, indicating that there was no need to include more complex variance structures for the covariate time. However, we observed two problems: (i) the variability of the residuals was higher among the DPN group than among the control

group and (ii) a pattern of increasing variability for the within-group errors. Therefore, we used a more general model to account different variances for within-group errors [22].

We specified a heteroscedastic model (*Model ARMAH (1,1)*) corresponding to a variance covariate given by the fitted values and the group variable, simultaneously. Such variance model is represented by

$$Var(\epsilon_{ij}) = \sigma^2 |v_{ij}|^{2\delta_{si_j}}$$

where v_{ij} is a vector of fitted values, δ is a vector of variance parameters and s is the group variable. Thus, the variance changes with a power of the absolute fitted values, but at different rates for each group.

Model ARMAH (1,1) showed better model adequacy, according to the AIC and BIC, as well as the LRT, χ^2 (2) = 445 (p < .01). The residuals analysis also indicates a better fit, with similar variances for each group and with a homogeneous pattern of variability for the standardised residuals, demonstrating that the chosen variance model successfully described the within-group variance. Therefore, showing that it is necessary to adopt complex variance structures for the data, answering question (vi). The parameters of the final RMS value model are presented in Table 4.

Table 4 Parameter estimates for the *Model ARMAH (1,1)* for the RMS value

Model Parameter	**Estimate (95% CI)**	***p*-value**
Fixed Effects		
Intercept RMS value for the Control group (β_0)	43.88 (34.89 – 52.88)	<.01
Change in RMS value for a unit increase in time (β_1)	0.07 (-0.08 – 0.24)	0.33
Difference in the intercept for the DPN group (β_2)	-11.06 (-24.87 – 2.75)	0.11
Difference in change in the RMS value for a unit increase in time (β_3)	0.24 (0.04 – 0.48)	0.04
Random Effects		
SD of between subjects' intercept (b_{0i})	15.82 (11.64 – 21.51)	
SD of between subjects' slope (b_{1i})	0.27 (0.17 – 0.41)	
Autoregressive parameter (ρ)	0.25 (0.09 – 0.41)	
Moving average component (λ)	0.11 (-0.04 – 0.27)	
SD of residual errors (σ_e)	0.23 (0.16 – 0.32)	
Variance function Control (δ_C)	0.71 (0.61 – 0.80)	
Variance function DPN (δ_{DPN})	0.86 (0.77 – 0.96)	

From the final model (ARMAH (1,1)), we observed that there is a non-significant increase in the RMS values for the control group (p = 0.33). As our research question involves investigating different rates of change for the RMS value among the two groups, the model included an interaction term (β_3). The interaction term is marginally significant (p = 0.04). The interaction term indicates that the RMS values for the DPN group increase at a rate of 0.24 units higher than the control group [95 % confidence interval (CI) 0.04 – 0.48], thereby answering question (i), (ii) and (iii) from introduction; showing that (i) the RMS values change over time, but only for the DPN group (ii), therefore having a different rate of change across the two groups (iii).

The estimated population mean RMS value at baseline for the control group is 43.88 (95% CI: 34.89 – 52.88) and the estimated population mean for the DPN group is 11.06 (95% CI: -24.87 – 2.75) lower for the DPN group compared to the control group. This result is non-significant (p = 0.11). As for the random-effects, there was significant individual heterogeneity in initial levels (*i.e.*, intercepts) and slopes. The estimated standard deviation of the random intercept is 15.82, and the estimated standard deviation of the random slope is 0.27.

Fig. 3 shows the final model, with the observed values (dots), the estimated curves for each subject (grey line) and the estimated overall mean for the two groups (red line).

Fig. 3 Observed RMS values (dots), individual predictions (black line) and predictions for the group value (red) over the observed range of time (0 – 25 s) for the two groups. Healthy (Con) and diabetic peripheral neuropathy (DPN)

The adequacy of *Model ARMAH (1,1)* was verified through residuals analysis. The residuals did not reveal any systematic pattern for the covariate time. Furthermore, the normal probability plot of the residuals (Fig. 4) suggests that the residuals follow a normal distribution, except for a long tail at the upper end. As the residuals are symmetrically distributed around zero, and the errors present heavier tails than expected under normality, a possible solution is to model the errors with a heavier-tailed distribution *e.g.*, a Student's *t* distribution. However, this will not lead to inflated type I error rates, given that under normality assumptions, the heavier tails lead to more conservative tests for the fixed effects [22].

Fig. 4 A normal plot of the residuals for the ARMAH (1,1) model

4 Discussion

This paper was aimed to provide a description of LMEM and its use in a real-world application of HD-sEMG data analysis. The results showed a significant increase in the RMS value during the isometric contraction for the DPN group. This positive trend in the RMS value in EMG research has been named myoelectric manifestation of muscle fatigue because the trend occurs before the subjects can sustain the desired force [9]. Thus, as subjects with DPN are known to experience complications such as muscle weakness and higher fatigability [39], the positive linear trend in the RMS value observed only for the DPN group may be associated with a display of premature fatigue for these subjects. However, as the HD-sEMG indicate a global measure of the motor unit activity, confounding factors such as volume conductor properties may affect the interpretation of the results. Consequently, the change in the amplitude variables cannot be solely attributed to the changes in the muscle fibre conduction velocity, as they are also related to recruitment and firing rate properties of a motor unit. Therefore, more HD-sEMG parameters can be used to improve the characterisation of muscle fatigue [40].

LMEM provides several advantages over traditional methods for the analysis of repeated measures. One of the most important is the use of all available data. Often, some of the HD-sEMG recordings are too noisy, requiring the subject to be excluded for the analysis if using traditional repeated-measures ANOVA. On the contrary, LMEM can use all available data, also treating time as a continuous variable instead of a discrete variable, resulting in increased statistical power to detect growth effects [41].

Like any statistical analysis, the LMEM depends on the fulfilment of certain assumptions. Contrary to standard approaches, where the assumptions are too strong to hold in realistic applications, LMEM allows the use of realistic yet parsimonious assumptions for repeated measures experiments. More specifically, it provides the researcher with a wide range of correlation and variance structures to choose the best-fitting one and then obtain additional information regarding the observed data. It is also straightforward to include time-dependent or time-independent covariates in the LMEM, as long as the sample size is large enough, which could be used to investigate important research questions, such as the influence of age, gender, among other factors [22].

In this work, we fitted several models to the observed data in ascending order (*i.e.*, from the least complex to the most complex model), getting the optimal yet parsimonious model to describe the data. We observed that including the random-effects of intercept and slope increased the model adequacy significantly. However, even after specifying the random-effects, there was still a within-group serial correlation present in our data. Therefore, we considered in the model some commonly used correlation structures available in the *nlme* package [36].

We adopted a model with ARMA (1,1) correlation structure to our data, which resulted in non-significant autocorrelations in the residual analysis, as expected under the proper correlation model. Moreover, we had to include a more complex variance structure (*i.e.*, a larger variance for the DPN group), resulting in *Model ARMAH (1,1)*. We used a linear function to describe the data; however, if necessary, it would be straightforward to include quadratic, cubic, or higher-order polynomials to model the HD-sEMG variables. These steps elucidate the flexibility of working with LMEM. Such a detailed description of the data would be impossible when working with traditional repeated-measures ANOVA. Moreover, by using all available data, we did not have to use an averaging procedure to reduce the data to a single measurement, which would undoubtedly result in loss of information.

Our interest in the optimal random-effects and correlation structure is not only to obtain a reasonable model to the data. The estimates of the random-effects itself are of interest, given that they provide estimates of heterogeneity that might be useful for several purposes, such as reliability analysis [42] and also to provide insights into physiological phenomena. For instance, in this work, we had a larger variance for the DPN group, which might be a result of people within different stages of the DPN condition. Using standard statistical analysis, it would not be possible to quantify a different pattern of variance for the DPN group.

The high flexibility offered by the LMEM comes at a cost. While widely-used statistical software packages such as SAS (SAS Institute, Cary, NC, USA), SPSS (SPSS Inc., Chicago, IL, USA) and R (R Core Team, Vienna, AT) have LMEM readily available to applied researchers, there is no standard approach to the LMEM construction. Thus, several decision-making steps must be made by the investigator regarding model development and specification, requiring a deeper understanding of the data and the experimental design. In this case, researchers can look for recent recommendations for model building strategies [24]. Besides, LMEM has more computational complexity relative to standard techniques, which can result in convergence problems, such as the one we encountered when trying to fit an unstructured correlation matrix to our data, which occurred because the unstructured model requires the estimation of a large number of parameters.

The model adequacy was verified through residual analysis. We observed that there were no substantial problems with the assumptions of normality and homoscedasticity of the residuals. However, the residuals showed slightly heavier tails. Although not approached in this work, LMEM can be easily extended to generalised linear mixed models to use in the case of non-Gaussian error distributions [22]. In our case, the heavy tails could be fixed by adopting another distribution of the within-group errors, such as the Students *t* or generalized *t* distributions.

5 Conclusion

In this paper, we outlined the power and flexibility of using LMEM to analyse HD-sEMG signals. We observed that for the RMS value, it was necessary to include not only the random intercept but also a random effect in the slope. Moreover, we investigated several alternative correlation structures as well as more complex variance models. The best fit was found when adopting an ARMA (1,1) correlation matrix together with a heteroscedastic model for the variance, with an increasing variance for the fitted values, more accentuated for the DPN group. Such detailed results would not be possible to obtain by the usual repeated-measures ANOVA. Thus, in conclusion, we believe that adopting LMEM to repeated measures such as electromyography can provide additional information from the data, such as the possibility to investigate alternative correlation structures, and therefore provide new insights into HD-sEMG signals at group and individual levels.

Acknowledgements

The authors thank the Brazilian Government Funding Agencies Coordination for the Improvement of Higher Education Personnel (CAPES) and National Council for Scientific and Technological Development (CNPq) for MAF, SC and FRA postgrad scholarships and JLBM Research Productivity scholarship.

Compliance with Ethical Standards

Ethical approval All procedures of this study followed the principles of the Declaration of Helsinki and were approved by the Human Research Ethics Committee of the Federal University of Santa Catarina (Protocol Number: 2.390.994). Written informed consent was obtained from all participants before the experiment.

Conflict of Interest The authors declare that there is no conflict of interest.

References

1. Tesfaye S (2011) Recent advances in the management of diabetic distal symmetrical polyneuropathy. J Diabetes Investig 2:33–42. https://doi.org/10.1111/j.2040-1124.2010.00083.x
2. Parasoglou P, Rao S, Slade JM (2017) Declining Skeletal Muscle Function in Diabetic Peripheral Neuropathy. Clin Ther 1–19. https://doi.org/10.1016/j.clinthera.2017.05.001
3. Butugan MK, Sartor CD, Watari R, et al (2014) Multichannel EMG-based estimation of fiber conduction velocity during isometric contraction of patients with different stages of diabetic neuropathy. J Electromyogr Kinesiol 24:465–472. https://doi.org/10.1016/j.jelekin.2014.04.007
4. Oberbach A, Bossenz Y, Lehmann S, et al (2006) Altered Fiber Distribution and Fiber-Specific Glycolytic and Oxidative Enzyme Activity in Skeletal Muscle of Patients With Type 2 Diabetes. Diabetes Care 29:895–900. https://doi.org/10.2337/diacare.29.04.06.dc05-1854
5. Merletti R, Aventaggiato M, Botter A, et al (2010) Advances in Surface EMG: Recent Progress in Detection and Processing Techniques. Crit Rev Biomed Eng 38:305–345. https://doi.org/10.1615/CritRevBiomedEng.v38.i4.10
6. Suda EY, Gomes AA, Butugan MK, Sacco ICN (2016) Muscle fiber conduction velocity in different gait phases of early and late-stage diabetic neuropathy. J Electromyogr Kinesiol 30:263–271. https://doi.org/10.1016/j.jelekin.2016.08.006
7. Watanabe K, Gazzoni M, Holobar A, et al (2013) Motor unit firing pattern of vastus lateralis muscle in type 2 diabetes mellitus patients. Muscle Nerve 48:806–13. https://doi.org/10.1002/mus.23828
8. Suda EY, Madeleine P, Hirata RP, et al (2017) Reduced complexity of force and muscle activity during low level isometric contractions of the ankle in diabetic individuals. Clin Biomech 42:38–46. https://doi.org/10.1016/j.clinbiomech.2017.01.001
9. Merletti R, Knaflitz M, De Luca CJ (1990) Myoelectric manifestations of fatigue in voluntary and electrically elicited contractions. J Appl Physiol 69:1810–1820. https://doi.org/10.1152/jappl.1990.69.5.1810
10. Almeida S, Riddell MC, Cafarelli E (2008) Slower conduction velocity and motor unit discharge frequency are associated with muscle fatigue during isometric exercise in type 1 diabetes mellitus. Muscle and Nerve 37:231–240. https://doi.org/10.1002/mus.20919
11. Merletti R, Farina D (2016) Surface Electromyography: Physiology, Engineering, and Applications, 1st ed. IEEE Press Series on Biomedical Engineering, Hoboken
12. Matthews JN, Altman DG, Campbell MJ, Royston P (1990) Analysis of serial measurements in medical research. BMJ 300:230–235. https://doi.org/10.1136/bmj.300.6719.230
13. Littell RC, Pendergast J, Natarajan R (2000) Modelling covariance structure in the analysis of repeated measures data. Stat Med 19:1793–1819. https://doi.org/10.1002/1097-0258(20000715)19:13<1793::AID-SIM482>3.0.CO;2-Q
14. Lazic SE (2010) The problem of pseudoreplication in neuroscientific studies: is it affecting your analysis? BMC Neurosci 11:5. https://doi.org/10.1186/1471-2202-11-5
15. Galbraith S, Daniel JA, Vissel B (2010) A Study of Clustered Data and Approaches to Its Analysis. J Neurosci 30:10601–10608. https://doi.org/10.1523/JNEUROSCI.0362-10.2010
16. Aarts E, Dolan C V., Verhage M, Van der Sluis S (2015) Multilevel analysis quantifies variation in the experimental effect while optimizing power and preventing false positives. BMC Neurosci 16:1–15. https://doi.org/10.1186/s12868-015-0228-5
17. Schober P, Vetter TR (2018) Repeated measures designs and analysis of longitudinal data: If

at first you do not succeed-try, try again. Anesth Analg 127:569–575. https://doi.org/10.1213/ANE.0000000000003511
18. Quené H, Van Den Bergh H (2004) On multi-level modeling of data from repeated measures designs: A tutorial. Speech Commun 43:103–121. https://doi.org/10.1016/j.specom.2004.02.004
19. Gueorguieva R, Krystal JH (2004) Move over ANOVA: Progress in Analyzing Repeated-Measures Data and Its Reflection in Papers Published in the Archives of General Psychiatry. Arch Gen Psychiatry 61:310–317. https://doi.org/10.1001/archpsyc.61.3.310
20. Laird NM, Ware JH (1982) Random-Effeects Models for Longitudinal Data. Biometrics 38:963–974
21. Demidenko E (2013) Mixed models: theory and applications with R. John Wiley & Sons
22. Pinheiro J, Bates D (2006) Mixed-effects models in S and S-PLUS. Springer Science & Business Media
23. Matuschek H, Kliegl R, Vasishth S, et al (2017) Balancing Type I error and power in linear mixed models. J Mem Lang 94:305–315. https://doi.org/10.1016/j.jml.2017.01.001
24. Cheng J, Edwards LJ, Maldonado-Molina MM, et al (2009) Real longitudinal data analysis for real people: Building a good enough mixed model. Stat Med 29:504–520. https://doi.org/10.1002/sim.3775
25. West BT, Galecki AT (2011) An overview of current software procedures for fitting linear mixed models. Am Stat 65:274–282. https://doi.org/10.1198/tas.2011.11077
26. De Livera AlM, Zaloumis SO, Simpson JUA (2014) Models for the analysis of repeated continuous outcome measures in clinical trials. Respirology 19:155–161. https://doi.org/10.1111/resp.12217
27. Lininger M, Spybrook J, Cheatham CC (2015) Hierarchical linear model: Thinking outside the traditional repeated-measures analysis-of-variance box. J Athl Train 50:438–441. https://doi.org/10.4085/1062-6050-49.5.09
28. Murphy D, Pituch K (2009) The performance of multilevel growth curve models under an autoregressive moving average process. J Exp Educ 77:255–282. https://doi.org/10.3200/JEXE.77.3.255-284
29. Favretto MA, Cossul S, Andreis FR, et al (2018) High density surface EMG system based on ADS1298-front end. IEEE Lat Am Trans 16:1616–1622. https://doi.org/10.1109/TLA.2018.8444157
30. Favretto MA, Cossul S, Andreis FR, Marques JLB (2019) Evaluation of Rate of Muscular Force Development in Type 2 Diabetic Individuals with and without Diabetic Peripheral Neuropathy. In: XXVI Brazilian Congress on Biomedical Engineering. pp 31–36
31. Diggle P, Diggle PJ, Heagerty P, et al (2002) Analysis of longitudinal data. Oxford University Press
32. Verbeke G, Molenberghs G (2009) Linear mixed models for longitudinal data. Springer Science & Business Media
33. West BT, Welch KB, Galecki AT (2014) Linear mixed models: a practical guide using statistical software. Chapman and Hall/CRC
34. Akaike H (1998) Information theory and an extension of the maximum likelihood principle: Selected papers of Hirotugu Akaike. Springer, New York, NY
35. Schwarz G (1978) Estimating the dimension of a model. Ann Stat 6:461–464
36. Pinheiro J, Bates D, DegRoy S, et al (2018) nlme: Linear and Nonlinear Mixed Effects Models
37. R Core Team (2018) R: A Language and Environment for Statistical Computing
38. Kwok OM, West SG, Green SB (2007) The impact of misspecifying the within-subject covariance structure in multiwave longitudinal multilevel models: A Monte Carlo study. Multivariate Behav Res 42:557–592. https://doi.org/10.1080/00273170701540537
39. Allen MD, Kimpinski K, Doherty TJ, Rice CL (2015) Decreased muscle endurance associated with diabetic neuropathy may be attributed partially to neuromuscular transmission failure. J Appl Physiol 118:1014–1022. https://doi.org/10.1152/japplphysiol.00441.2014
40. González-Izal M, Malanda A, Gorostiaga E, Izquierdo M (2012) Electromyographic models to assess muscle fatigue. J Electromyogr Kinesiol 22:501–512. https://doi.org/10.1016/j.jelekin.2012.02.019

41. Muthén BO, Curran PJ (1997) General Longitudinal Modeling of Individual Differences in Experimental Designs: A Latent Variable Framework for Analysis and Power Estimation. Psychol Methods 2:371–402. https://doi.org/10.1037/1082-989X.2.4.371
42. Andreis FR, Favretto MA, Cossul S, et al (2019) Reliability of Maximal Voluntary Isometric Contraction of Ankle Dorsiflexion in Male Subjects. In: Costa-Felix R, Machado J, Alvarenga A (eds) XXVI Brazilian Congress on Biomedical Engineering. Springer, Singapore, pp 353–357

Figures

Fig 1:

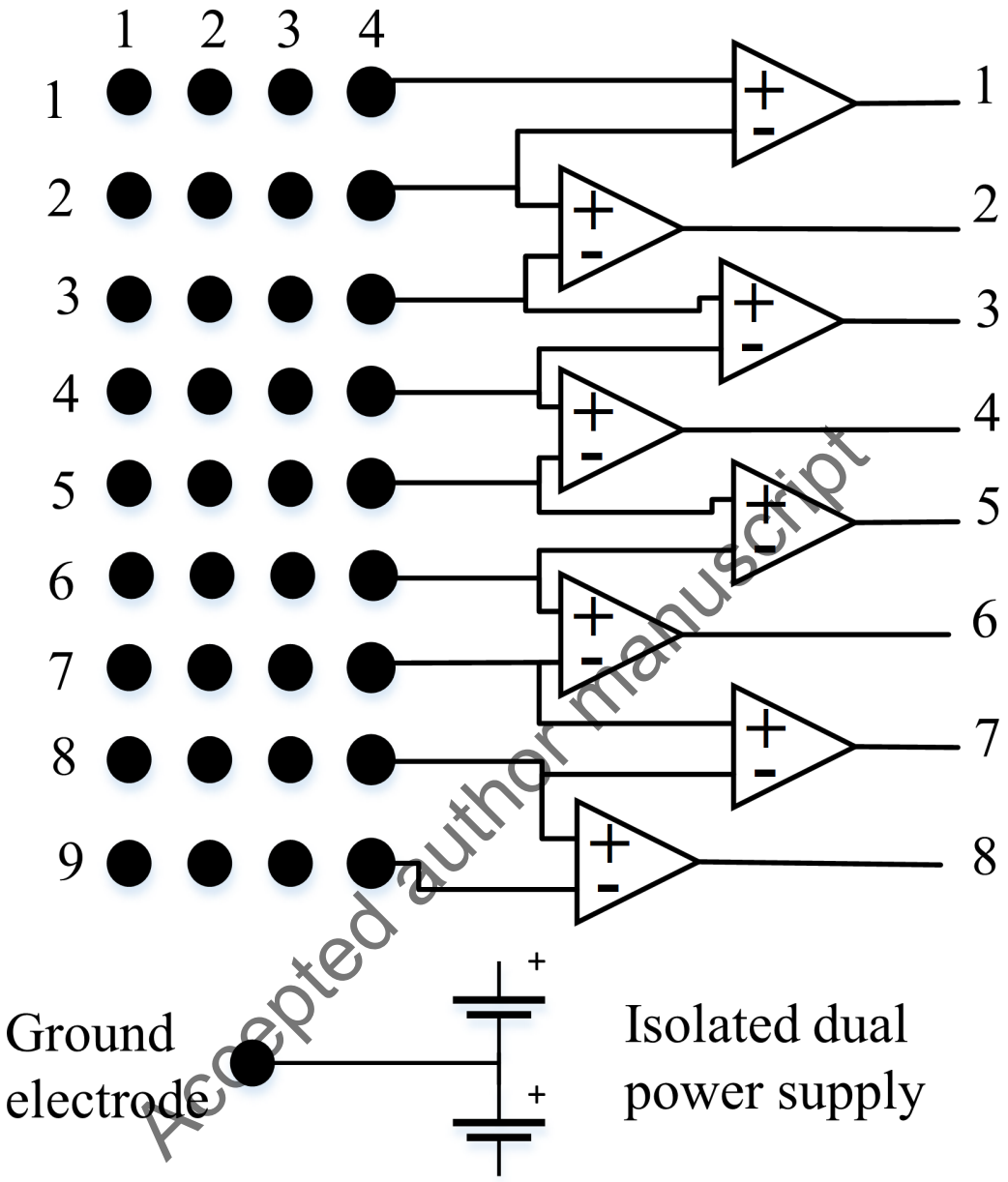


Fig 2:

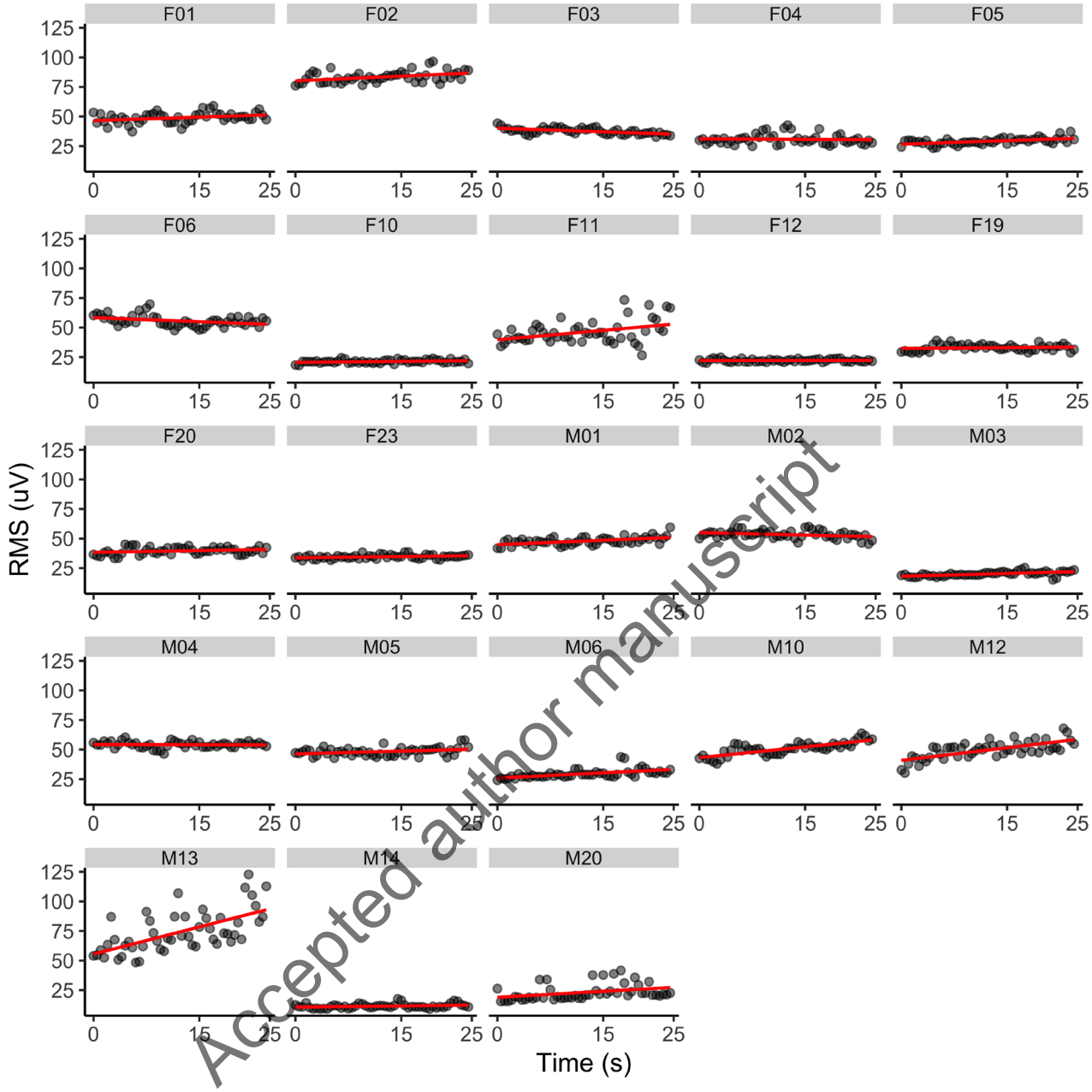


Fig 3:

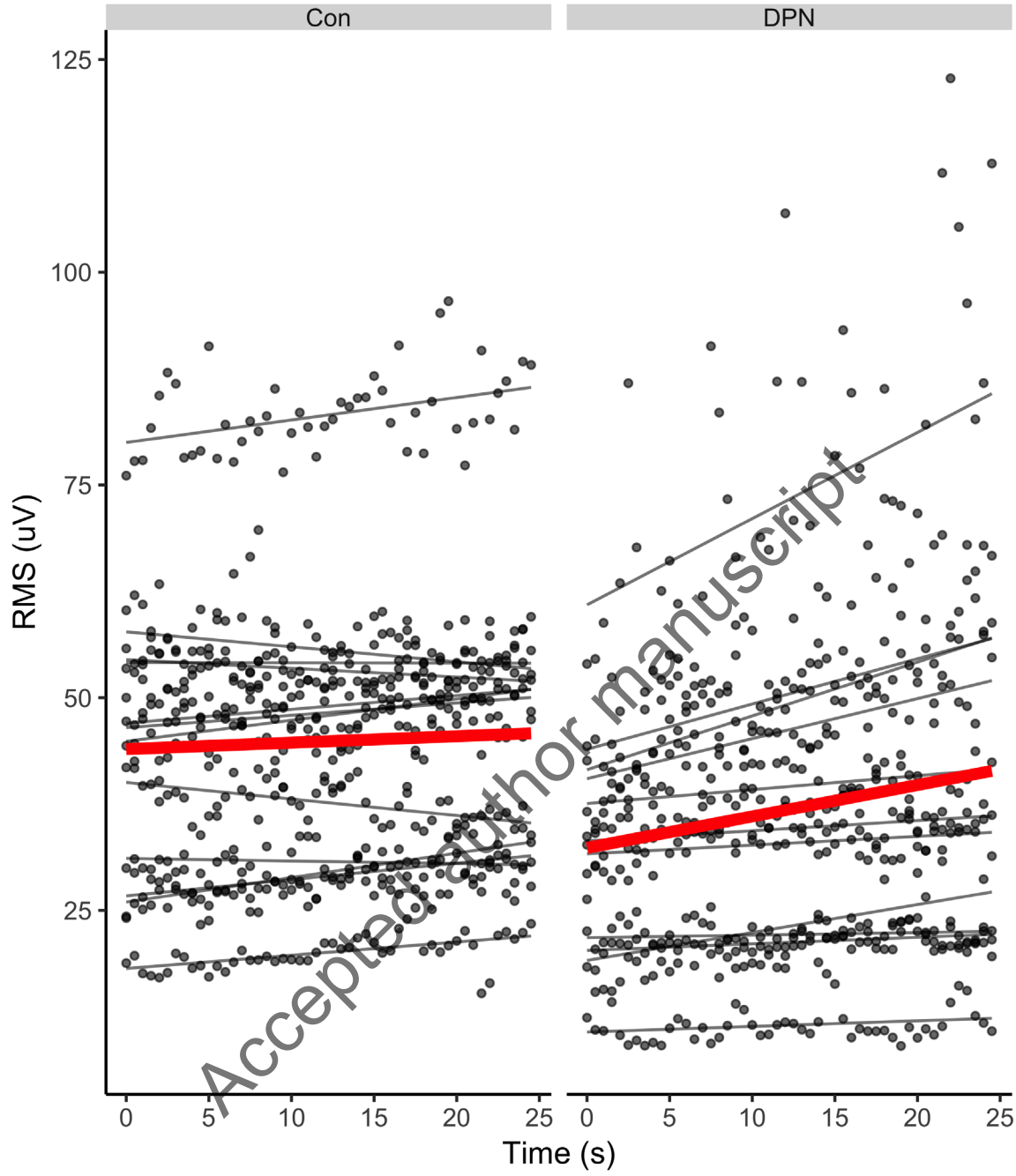


Fig 4: