Non linear mixed-effects model for Sound Intensity In Area X & High Vocal cord (HVC)

Course: Topics In Advanced Modeling Techniques
Assignment: Non Linear Models
2019-2020

2nd year Master of Statistics
Hasselt University

Kubam Ivo (1850637)

Submission Date: 08/01/2020

Lecturers:
Prof. Dr. Geert Molenberghs
Prof. Dr. Geert Verbeke
Prof. Dr. Tomasz Burzykowski

1 Introduction

In order to understand the functional characteristics of specific neuronal populations in a strongly connected brain circuitry, the so-called song control system in the songbird brain,Van der Linden et al. 2002 established a novel in vivo magnetic resonance imaging (MRI) approach. Nucleus HVC (formerly called high vocal center) of songbirds contains two types of projecting neurons connecting HVC respectively to the nucleus robustus archistriatalis, RA, or to area X. Van der Linden et al. 2002 analyzed the effect of testosterone on the dynamics of Mn^{2+} accumulation in RA and area X in individual birds injected with manganese in their HVC.

1.1 Scientific Question

This study was aimed at modeling a non-linear mixed-effects model for SI in area X considering both periods simultaneously in order to answer the question if there is a significant difference in SI in areax for control vs treated birds. It was of interest to answer the question if there is an effect in adding volume area X as covariate to the model. Secondgly, it was of interest to build a model for high vocal center signal intensity SI HVC in the second period only to answer the question if there is a difference in SI in HVC for treated vs control birds.

2 Methodology

2.1 Data Description And Summary Statistics

With 10 variables measured, the dataset consist of 640 observations (320 each for both periods) from 10 song birds. The response variable of interest; sound intensity in area X and HVC, both had an average value of 0.101 and 2.138 for the treated birds in period two respectively. These values were higher than the mean values for the control birds in period two; 0.043 and 1.617 respectively. Slight difference in SI for Area X for the control birds in the initial and active period.

Table 1: Summary Statistics

Period	Treatment	Variable	Mean	Min	Max
Initial	Control 1	SI Area X	0.056	-0.026	0.187
Active	Control 2	SI Area X	0.043	-0.008	0.122
Active	Control 2	SI HVC	1.617	0.652	3.055
Active	Treatment 2	SI Area X	0.101	-0.025	0.297
Active	Treatment 2	SI HVC	2.138	0.999	4.017

2.2 Exploratory Analysis

Plotting the sound intensity at HVC for period two only (Figure 1) shows the SI initially starts high with a small increase for some birds and a general decrease in SI for all birds. SI for area X for both periods shows a different pattern as all birds starts low and steadily increase with time as seen in figure 2.

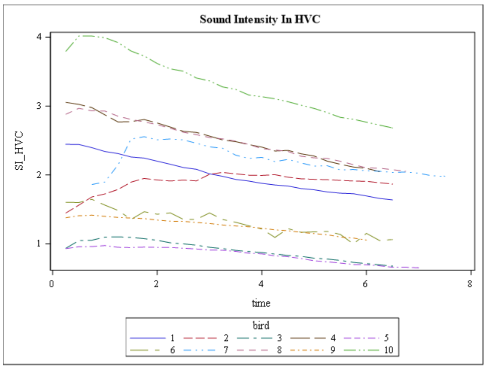

Figure 1: SI HVC

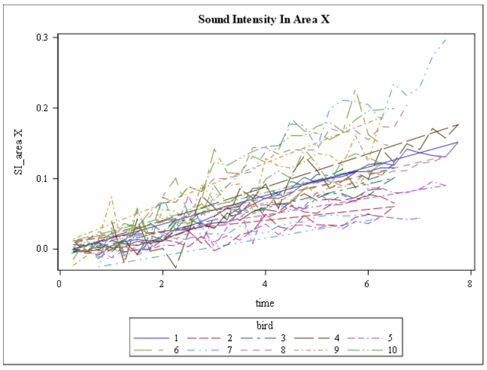

Figure 2: SI Area X

2.3 Statistical Analysis

A hierarchical form of the model first formulated by Van der Linden et al. 2002 was considered to model the sound intensity in Area X for all 10 birds.

$$SI_{ij}(AreaX) = \frac{(\phi_0 + \phi_1 G_{1i} + \phi_2 G_{2i} + v_i) t_{ij}^{\eta_0 + \eta_1 G_{1i} + \eta_2 G_{2i} + n_i}}{(\tau_0 + \tau_1 G_{1i} + \tau_2 G_{2i} + t_i)^{\eta_0 + \eta_1 G_{1i} + \eta_2 G_{2i}} + T_{ij}^{\eta_0 + \eta_1 G_{1i} + \eta_2 G_{2i}}} + \gamma_0 + \gamma_1 G_1 + \gamma_2 G_{2i} + \varepsilon_{ij}$$

$(\mathrm{v}_i,\, n_i,\, t_i) \sim N(0, D)$
$SI_{ij}(AreaX)$ is the measurement at time j for bird i. G_{1i} and G_{2i} were indicators for control birds in period one vs control birds in period two and testosterone treated birds in period two vs control birds in period two respectively. t_{ij} is the measurement time. These indicators were considered to account for the two periods taken into consideration simultaneously. ϕ_0 is the maximal signal intensity for control bird in period two while ϕ_1 + ϕ_2 were the maximal signal intensity for untreated bird in period one and treated birds in period two respectively. η_0 + η_1 + η_2 governs the shape of the curve. τ_0+ τ_1 + τ_2 are the required time to reach 50% of the maximum intensity. ε_{ij} is a measurement error term assumed to be normally distributed.

As explained by Van der Linden et al. 2002, the HVC region can be regarded as the central compartment from which manganese is dispersed to area X and SA. So a two-compartment model is a reasonable choice to model the sound intensity in HVC. The model considered is non linear in its mean structure but can assumes different variance structure (Power model).

$$Y_{ij} = e^{(\beta_1 + \gamma_1 G_i + b_{1i})} exp[-e^{(-\beta_2 + \gamma_2 G_i + b_{2i}) t_{ij}}] - e^{(\beta_3 + \gamma_3 G_i + b_{3i})} exp[-e^{(-\beta_4 + \gamma_4 G_i + b_{4i}) t_{ij}}] + \varepsilon_{ij}$$

$\varepsilon_{ij} \sim N(0, \sigma^2 \mu_{ij}^{2\theta})$ and $Var(y_{ij}) = \sigma^2 g^2(\mu_{ij}, \theta)$. $g^2()$ is is a function of the mean, additional variance parameters, and perhaps other covariates.
Y_{ij} is the SI measurement at time j for bird i, G_i is 1 for testosterone treated birds and 0 otherwise. $(\mathrm{b}_{1i},\, b_{2i},\, b_{3i}, b_{4i}) \sim N(0, D)$ representing bird's random slopes and intercepts for the two compartments. e^{β_2} and e^{β_4} are the sound intensity rates corresponding to the two compartments. Details about the above two model formulations can be found in Verbeke and Molenberghs 2009

3 Results

3.1 Modeling SI Area X

In order to build this model, a simpler no random effects model was considered and random effects added progressively in order to get good initial values. Due to computational difficulties, the covariances between the random effects were considered as zero.This constraint still does not destroy the correlation within the obervations. Fitting the full model resulted in the following parameter estimates with -2loglikilihood value of -2892.

Table 2: Parameter Estimates for the SI Area X

Effect	Parameter	Estimate(Std Error)	Effect	Parameter	Estimate(Std Error)
	ϕ_0	0.2366(0.05975)*		γ_0	-0.00056(0.009562)
	ϕ_1	1.2778(0.4116)*		γ_1	-0.00236 (0.004574)
	ϕ_2	0.03260(0.07549)		γ_2	0.003422 (0.006125)
	η_0	1.5048(0.2217)*		β_1	-0.00174(0.003917)
	η_1	-0.1822(0.2062)	$Var(v_{1i})$	d_{11}	0.003439(0.002358)
	η_2	0.4993(0.3263)	$Var(t_{1i})$	d_{22}	17.5255(0.003422)
	τ_0	8.8867(2.8974)	$Var(n_{1i})$	d_{33}	0.03034 (0.01753)
	τ_1	34.4305(11.5109) *	$Var(\epsilon_{ij})$	σ^2	0.000190(0.000013)*
	τ_2	-3.9708(3.5206)			

The estimates presented in the table 2 shows no significant difference in SI for Areax for treated birds in period two vs control birds in period two as ϕ_2, η_2, τ_2 and γ_2 were all found to be insignificant. At least ϕ_1 was found to be significant which implies significant difference between control birds in period two and control birds in period one. Backward selection technique using the likelihood ratio test was conducted in order to simplify the model.

- Removal of the random effect v_i gave $G^2 = 18$ and null distribution $\chi_{2:3}$. Test statistic was found to be significant, so not possible to remove v_i. The same conclusion was found for the other two random effects.
- The fixed effect ϕ_2 was found to be highly significant ($G^2 = 19$, 1df, p value= 0.0001) so could not be removed from the model. All other parameters were tested and non was found to non significant.
- Even though the covariate volume of area X was found to be insignificant, it could not be removed from the full model.

The figure 3 and 4 below show two randomly selected individual curves, showing data points as well as empirical Bayes predictions for bird specific curves. The plots therefore signifies how well the model fitted the data.

3.2 Modeling SI HVC

Using the bi-exponential model as formulated in the methodology, individual simpler models with no random effects were fitted for each bird in order to get good starting values for each parameters. Homoscedastic hierarchical model was then fitted with four random effects.The variance covariance matrix was reduced not to contain covariances due to computational difficulties.

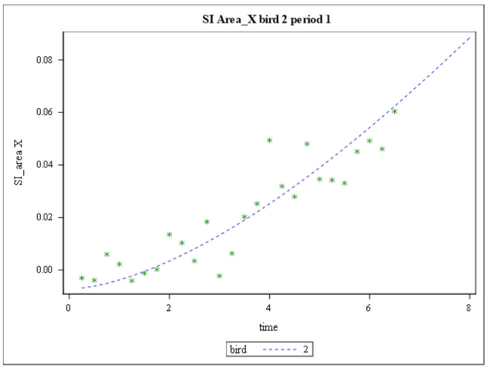

Figure 3

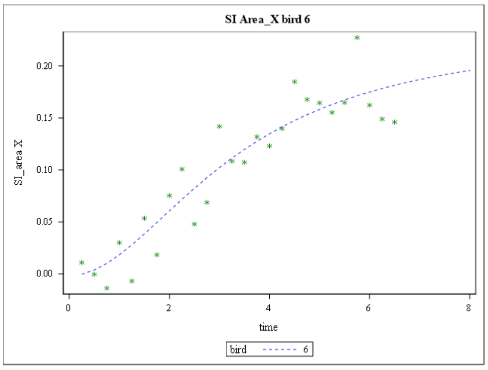

Figure 4

Table 3: Parameter Estimates for the SI HVC

Parameter	Estimate(Std Error)	Effect	Parameter	Estimate(Std Error)
β_1	1.1719(0.02125)	$Var(b_{1i})$	d_{11}	0.4689(0.03411)
β_2	0.3965(0.1655)	$Var(b_{2i})$	d_{22}	1.5751(0.2254)
β_3	0.7584(0.04161)	$Var(b_{3i})$	d_{33}	0.8508(0.05897)
β_4	2.5014(0.1287)	$Var(b_{4i})$	d_{44}	1.0410 (0.1467)
γ_1	0.951(0.1560)	$Var(b_{4i})$	d_{44}	0.1622 (0.007096)
γ_2	-3.1068(0.2613)	$Var(\epsilon_{ij})$	σ^2	0.0026(0.007096)
γ_3	0.8551(0.2640)			
γ_4	0.8551(0.2640)			

The -2log likelihood was found to be -124.1. From the estimates in table 3, all parameters were found to be significant. $e^{0.3965}$ and $e^{2.5014}$ are the sound intensity rates corresponding to the two compartments. From the values of the random effects there is a strong indication for between-bird variability. In order to extend this model, a power model were the variance is assumed to depend on the mean was tested. But to computational difficulties, this model was not implemented further. The figure 5 and 6 below show two randomly selected individual curves, showing data points as well as empirical Bayes predictions for bird specific curves. The plots therefore signifies how well the model fitted the data.

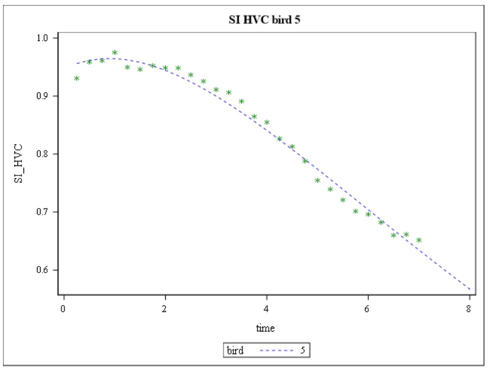

Figure 5

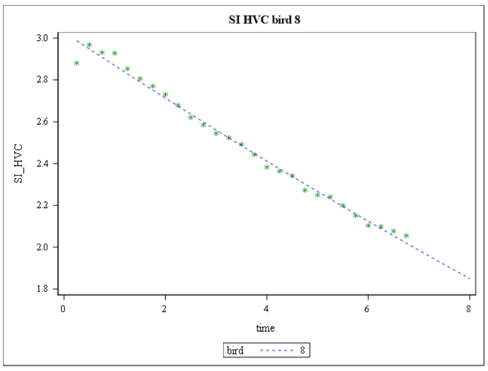

Figure 6

4 Discussion And Conclusion

As was describe in the methodology, two indicators variables were used in order to model SI in Area X for the two periods simultaneously. The second period was really the period of interest as we had control and treated birds. From the results showed in table 2, no significant difference was found between control and treated birds in period two. Including the covariate volume of area x made the model to be stable as trying to exclude this variable cause a lot of problems. On the other hand modelling SI in HVC presented significant results for treated birds vs control birds for both compartments. A homoscedastic model was the best choice as trying to extend the model to a heteroscedastic type led to unwanted results.

References

Van der Linden, Annemie et al. (2002). “Applications of manganese-enhanced magnetic resonance imaging (MEMRI) to image brain plasticity in song birds”. In: *NMR in Biomedicine: An International Journal Devoted to the Development and Application of Magnetic Resonance In Vivo* 17.8, pp. 602–612.

Verbeke, G. and G. Molenberghs (2009). *Linear Mixed Models for Longitudinal Data*. Springer Series in Statistics. Springer New York. ISBN: 9781441902993. URL: `https://books.google.be/books?id=jmPkX4VU7h0C`.

5 Appendix

```
proc sgplot data = nl.songbird2;
styleattrs datalinepatterns=(solid);
title "Sound Intensity In HVC";
series x = time y =  SI_HVC/group=bird;
run;
*Homoscedastic model;
proc nlmixed data=nl.songbird2 noad qpoints=3;
parms beta1=2.7505 beta2=-1.5 beta3=2.5855
beta4=3.9273 sigma=0.9 d11=0.052 d22=0.246 d33=0.509 d44=0.937
eta1=2.5 eta2=-1.3 eta3=2.3 eta4=-3.2;
avg = exp(beta1+ eta1*group + b1)*exp(-exp(-beta2 +
eta2*group +b2)*time) - exp(beta3 + eta3*group + b3)*exp(-exp(-beta4 +eta4*group+b4)*time);
model  SI_HVC~normal(avg,sigma**2);
random b1 b2 b3 b4 ~ normal([0,0,0,0],[d11,0,d22,0,0,d33,0,0,0,d44])
subject = bird;
predict avg out=nl.hvc;
run;

proc sgplot data = nl.hvc;
title "SI HVC bird 8";
scatter x = time y =  SI_HVC /markerattrs=(symbol=asterisk color=green);
series x = time y =  pred/group=bird lineattrs=(pattern=2);
where bird=8;
run;

*General Equation;
*Homoscedastic model;
proc nlmixed data=nl.songbird2 noad qpoints=3;
parms beta1=2.7505 beta2=-1.5 beta3=2.5855
beta4=3.9273 sigma=0.9 d11=0.052 d22=0.246 d33=0.509 d44=0.937
eta1=2.5 eta2=-1.3 eta3=2.3 eta4=-3.2;
avg = exp(beta1+ eta1*group + b1)*exp(-exp(-beta2 +
eta2*group +b2)*time) - exp(beta3 + eta3*group + b3)*exp(-exp(-beta4 +eta4*group+b4)*time);
model  SI_HVC~normal(avg,sigma**2);
random b1 b2 b3 b4 ~ normal([0,0,0,0],[d11,0,d22,0,0,d33,0,0,0,d44])
subject = bird;
predict avg out=nl.hvc;
run;

proc sgplot data = nl.hvc;
title "Sound Intensity In Area X";
scatter x = time y =  SI_HVC /markerattrs=(symbol=asterisk color=green);
series x = time y =  pred/group=bird lineattrs=(pattern=2);
*where bird=9;
run;
```