

BACTERIAL COUNTS IN AGED BEEF ANALYZED WITH MIXED GENERALIZED LINEAR MODELS IN R

Ales Dufek^{1*}, Marcela Vyletelova¹, and Miroslav Homola²

¹Agriresearch Rapotin Ltd. Vyzkumniku 267, 788 13 Vikyrovce, Czech Republic

²Research Institute for Cattle Breeding, Ltd., Vyzkumniku 267, 788 13 Vikyrovce, Czech Republic

*Corresponding author e-mail: ales.dufek@vuchs.cz

Abstract— We used the package lme4 in R for fitting of GLMM. The fixed effect on ratio between coliform bacteria and total bacteria can be expressed as following mean values: 0.00073, 0.00059, for fluid, surface on the 16th d of aging, and 0.002629, 0.003133 for fluid, surface on the 44th d of aging, resp.

Index Terms— bacteria, counts, GLMM, lme4.

I. INTRODUCTION

The purpose of our short paper is to present procedure and results of fitting of the generalized linear mixed effects models for a microbial meat quality trait – bacterial counts in long-term aging period. During statistical analysis we had to tackle with problems associated with transforming data and data from a field research: unbalanced data, many covariates, fixed and random effects. We found the package lme4 (Bates and Maechler 2010) in R software environment as the great tool for data diagnostics and fitting of generalized linear mixed models (GLMM). The lme4 package allows fitting of LMM and GLMM with nested or cross-classified factor design. Nested effects was in our research that is focused on achievement of beef quality from the point of palatability and microbial safety. Besides, the beef was obtained from the animals reared at very extensive breeding conditions.

II. MATERIALS AND METHODS

Extensively fattened crossbreeds (11 bulls, 12 heifers) were bred in grass-based fattening system in low favored area. Animals were grazing in vegetative season and were fed grass silage in winter. Before slaughter, 23 animals were from technical reasons divided into 4 groups (A, B, C, D). Each group (the first random effect) was slaughtered in different days. The average age at time of slaughter was 622.3 days (s.d.=37.9) and average weight of carcass 289 kg (s.d.=58.4). A part of *Musculus longissimus lumborum et thoracis* was removed from every carcass at 24 h post-mortem. The part (second random effect) was divided to 3 samples and they were individually vacuum packed. One sample was analyzed 48 hours post-mortem. The surface of each part was sampled by swabbing an area of 10 cm² by swabs. The sponge was put into 1 ml of physiological solution. The other two samples were stored at 2-4 °C for aging period on following 16 and 44 days. The surfaces of these samples were also sampled as described above. Besides, 1 ml of meat fluid was obtained from the plastic bags in which the samples were stored during aging. In both methods, ("surface", "fluid") the solution (or its relevant dilution) was inoculated onto Petri dishes and the total count of bacteria (TCB) and coliform bacteria (CB) were determined according to norms (CSN EN ISO 4833; CSN ISO 4832). Counts were expressed as the CFU per 1 ml of solution.

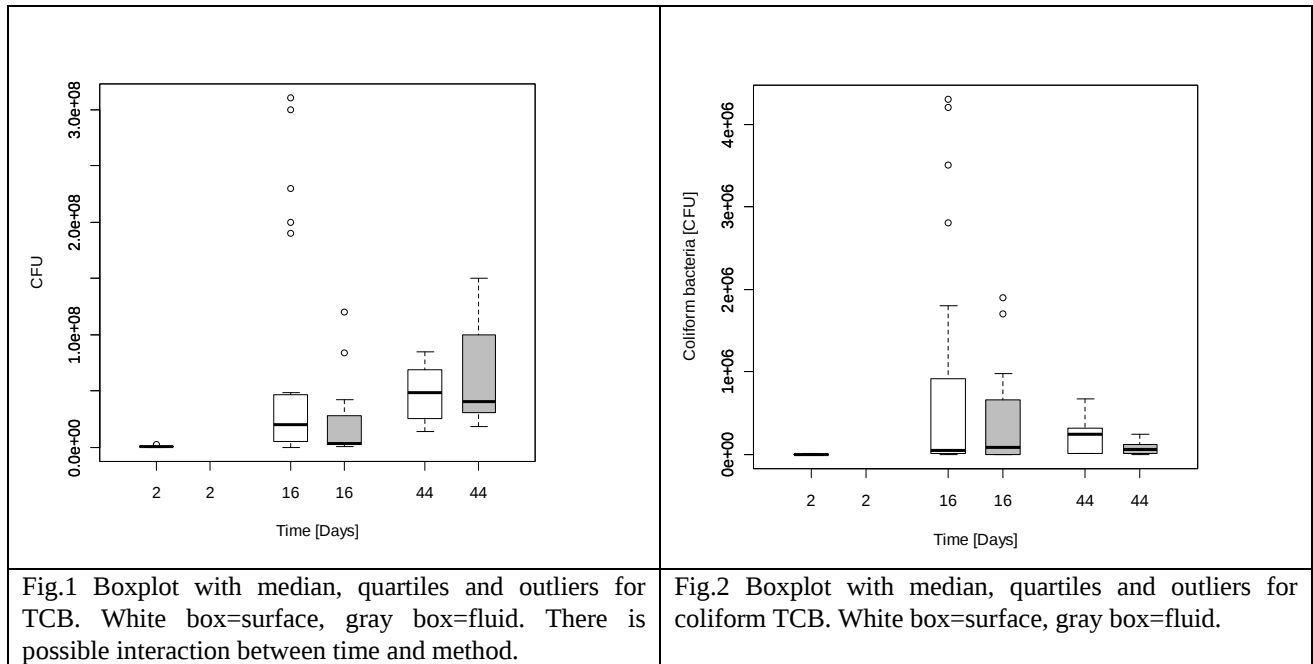
There was unbalanced design (46 cases from surface analysis, 22 cases from liquid analysis performed only in groups B and D). We computed GLMM only for cases from 16th and 44th d, because of computing of interaction between aging time and method ("surface", "fluid"). Distribution of cases in subgroups and effect of group as a random factor is presented in the Fig. 3.

We used package lme4 version 0.999375-33 in R version 2.11.0. Dependent variables were not transformed, because we built GLMM models with the Poisson error distribution for evaluating of TCB and binomial distribution for evaluating of ratio between counts of coliform bacteria and TCB.

III. RESULTS AND DISCUSSION

TCB

Total counts of bacteria are presented in the Fig. 1. Counts of coliform bacteria are presented in Fig 2.



The first model obtained with `TCB.1<-glmer(TCB~time*method+(1|group/subject)),data=mic4,family=poisson)` showed significant effect of aging time, method and interaction on TCB.

Fixed effects:TCB.1					
	Estimate	Std. Error	z value	Pr(> z)	
Intercept=time16, fluid	1.57E+01	6.03E-01	26	<2e-16	***
time44	3.54E+00	1.16E-04	30550	<2e-16	***
methodsurface	1.32E+00	5.53E-05	23797	<2e-16	***
time44:methodsurface	-1.60E+00	7.93E-05	-20165	<2e-16	***

However the estimates of fixed effects are on the transformed scale, the relative differences between estimates are in according to the obvious differences between boxplots in Fig. 1. Also the significant interaction between time and method is obvious in the Fig. 1. TCB are higher on the surface then in fluid on the 16th d of aging. On the contrary, TCB are lower on the meat surface then in fluid on the 44th d of aging. The interaction is very significant.

Fitted values by the model were obtained with `mic4$f.TCB.1<-fitted(TCB.1)` and the quality of fitting was assessed with correlation coefficient (Baayen, R.H. 2008) between fitted and real values with `cor(mic4$TCB,mic4$f.TCB.1)` and with coefficient of determination `cor(mic4$TCB,mic4$f.TCB.1)^2` ($r = 0.9137761$, $r^2 = 0.8349868$).

Ratio between coliform bacteria and TCB

The counts of coliform bacteria were not analyzed with a model with parameter for Poisson error distribution as the model for TCB. Instead of such a model we analyzed ratio between counts of coli form bacteria and TCB. The ratio between coliform bacteria and TCB in aging time and methods are presented in the Fig. 3. The left plot presents ratio counted from real data. The right plot presents ratio from values fitted with GLLM. The colifmCFU/TCB CFU ratio was evaluated with GLLM with binomial error distribution obtained with

```
mb1.1<-glmer(cbind(Colifm, TCB - Colifm)~time*method+(1|group/subject),data=mic4,family="binomial")
print(mb1.1,cor=F)
```

Fixed effects:					Without random effects	
	Estimate	Std. Error	z value	Pr(> z)		Estimate
Intercept=time16, fluid	-6.08E+00	1.49E+00	-4.1	4.32E-05 ***		-2.89E+00
time44	1.23E+00	4.71E-03	260.7	< 2e-16 ***		-3.92E+00
methodsurface	-1.04E+00	3.60E-04	-2889.1	< 2e-16 ***		-1.46E+00
time44:methodsurface	1.36E+00	1.40E-03	971	< 2e-16 ***		2.82E+00

Estimates in the very right column represents generalized model without any random effects. Differences correspond with graphical presentation of real and also fitted data from GLMM. However, the interpretation according to estimates obtained from "only fixed model" or from the graph with fitted values from mixed model is misleading. Neither estimates of "only fixed model" nor widespread-used box-plot graphical presentation of fitted values from mixed-model do not correspond with real studied design.

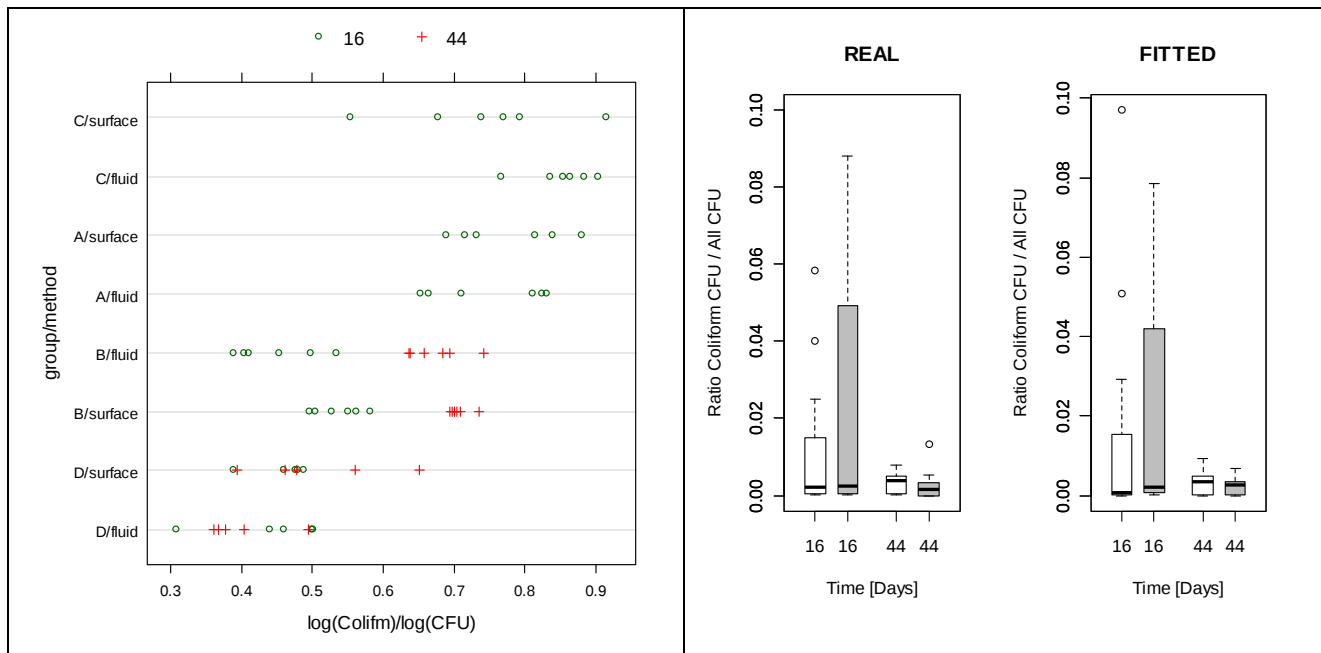


Fig.3 Random effect (groups A, B, C, D) on the ratio between counts of coliform bacteria and TCB. Unbalanced data for effect of aging are also obvious.

Fig.4 Boxplot with median, quartiles and outliers for ratio between coliform bacteria and TCB. White box=surface, gray box=fluid.

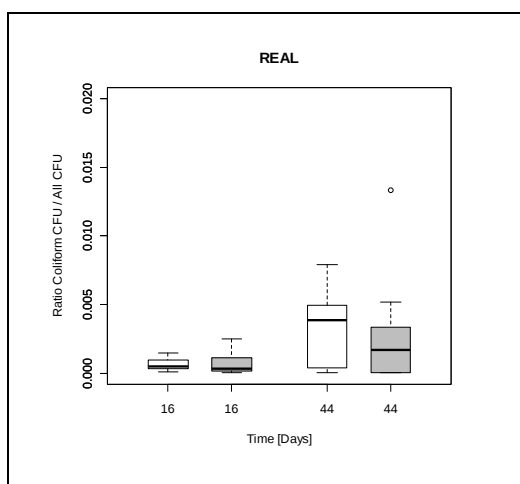


Fig.5 Boxplot with median, quartiles and outliers for ratio between coliform bacteria and TCB in groups B and D. The effect of groups A and C is removed. White box=surface, gray box=fluid.

Estimate for time44 in GLMM (1.23E+00) shows, that ratio is higher in the 44 d of aging. On the contrary, estimate from "only fixed model" (-3.92E+00) and boxplot from GLLM show, that ratio is higher in 16th d of aging. It is because of unbalanced design (Fig.3). There are missing values for 44th d of aging in the groups A and C. It covers effect of aging observed in groups B and D (Fig.5). In the Fig. 5, there is also obvious very significant interaction between aging time and method of sampling collection. On the 16th d the ratio for surface is lower then for fluid. On the 44th d the ratio for surface is higher then for fluid.

However there are z and corresponding p values in the table with estimates for GLLM, significance of random or fixed effect is obtained by comparison of the models with anova (Vazquez, Bates, Rosa, Gianola and Weigel., 2010). We evaluated the fixed effect of sex with

mb1.2<-update(mb1.1,~.+sex)
anova(mb1.1,mb1.2) #generates following table

Model	Df	AIC	logLik	Chisq	Chi Df	Pr(>Chisq)
mb1.1	6	18886598	-9443293			
mb1.2	7	18886598	-9443292	2.2656	1	0.1323

IV. CONCLUSION

Mainly because of unbalanced data there was a need to carefully build up and interpret statistical generalized linear mixed effects models to reveal the interaction between selected fixed effects. We found out significant effect of aging, method of sampling collections and interaction of these effects on the ratio between count of coliform bacteria and total count of bacteria. The significant interaction between these fixed effects is connected with following mean values (counted from groups B, D with completed design): 0.00073, 0.00059, for fluid, surface on the 16th d of aging, and 0.002629, 0.003133 for fluid, surface on the 44th d of aging, respectively. Mean values counted from all groups are 0.03646, 0.02729 for fluid and surface, respectively. We simply present commands for GLMM counting in package lme4 in R.

ACKNOWLEDGEMENT

This study was funded by the Ministry of Education, Youth and Sports of the Czech Republic (Projects LA08035, 2B08037 and MSM 267 884 6201)

REFERENCES

- Baayen, R.H. (2008). Analyzing Linguistic Data. A practical introduction to statistics. 7.4.:Generalized linear mixed models. Cambridge University Press.
- Bates, D. & Maechler, M. (2010). Matrix: Sparse and Dense Matrix Classes and Methods. R package version 0.999375-38. <http://CRAN.R-project.org/package=Matrix>
- Bates, D. & Maechler, M. (2010). lme4: Linear mixed-effects models using Eigen and Eigen interfaces. R package version 0.999375-33. <http://CRAN.R-project.org/package=lme4>
- CSN EN ISO 4833. (2003). Microbiology of food and animal feeding stuffs – Horizontal method for enumeration of micro-organisms – Colony-count technique at 30°C. CNI Prague.
- CSN ISO 4832. (1995). Microbiology – General guidance for enumeration of coliforms – Colony-count technique, CNI Prague.
- Pinheiro, J., Bates D., (2000). Mixed-Effects Models in S and S-PLUS. 2nd ed. Springer, 528 p., 172 illus.
- Pinheiro J., Bates D., DebRoy S. & Sarkar D. the R Core team. (2007). nlme: Linear and Nonlinear Mixed Effects Models. R package version 3.1-86.
- R Development Core Team (2010). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org>.
- Sarkar, D. (2010). lattice: Lattice Graphics. R package version 0.18-5. <http://r-forge.r-project.org/projects/lattice/>
- Vazquez, A.I., Bates, D. M., Rosa, G. J. M., Gianola, D. & Weigel, K. A. (2010). Technical note: An R package for fitting generalized linear mixed models in animal breeding. *Journal of Animal Science*, 88, 497-504.