Supplementary material: Using the Growth Curve Model in Classification of Repeated Measurements

Dietrich von Rosen · Martin Singull

Comparisons of classification approaches using well known data sets

In this supplementary material we will consider five data sets comprising repeated measurements and compare the performance of the five different classification procedures which also were discussed in the previous section. We will do it by a leave-one-out cross-validation procedure, i.e., removing one vector of repeated measurements at time, which then should be classified. The six data sets all consists of data from two populations with n_1 and n_2 independent observations. Hence, for each data set, we will classify $n = n_1 + n_2$ observations and find the proportions of correctly, misclassified and unknown classified measurements, respectively. For each data set we will assume a quadratic growth.

The first data set is the classical set of dental measurements (Potthoff & Roy 1964). Each measurement is the distance, in millimeters, from the center of the pituitary to the pterygo-maxillary fissure, and were made on $n_1 = 11$ females and $n_2 = 16$ males at four different ages 8, 10, 12 and 14. The plot of the data is given in Figure 1. In Table 1, based on the cross-validation procedure, we can find the proportions of correctly, misclassified and unknown unclassified children, respectively.

The second data set is a glucose data set (Pan & Fang 2002), where the patients are divided into two groups: a control group ($n_1 = 13$ patients) and an obese group ($n_2 = 20$ patients). The number of repeated measurements (of plasma inorganic phosphate) in this medical study is eight, taken at time points 0, 0.5, 1, 1.5, 2, 3, 4 and 5 hours after a standard dose of oral glucose has been administrated. The data is plotted in Figure 2 and a second degree polynomial seems to be resonable growth model for both groups. In Table 2 we can find the proportions of correctly, misclassified and unclassified patients.

Dietrich von Rosen
Department of Energy and Technology, Swedish University of Agricultural Science, Uppsala, Sweden, and Department of Mathematics, Linköping University, Linköping, Sweden

Martin Singull
Department of Mathematics, Linköping University, Linköping, Sweden
E-mail: martin.singull@liu.se

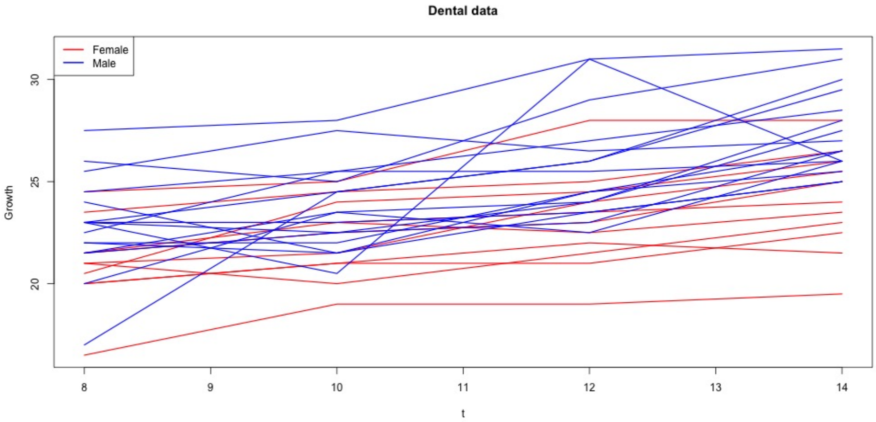

Fig. 1 Plot of dental data. Four repeated measurements at time 8, 10, 12 and 14 years, for the two groups, females ($n_1 = 11$) and males ($n_2 = 16$) are presented.

Table 1 The performance of the classification functions D, D_{MK}, D_{ML}, $(C_1, ..., C_6)$ is presented, i.e., the proportions of the children in the dental data set that are correctly, misclassified or classified as belonging to an unknown population, respectively.

	D	D_{MK}	D_{ML}	(C_1, C_2)	$(C_1, ..., C_6)$
correctly	63%	67%	63%	59%	33%
misclassified	37%	33%	37%	33%	7%
unknown	-	-	-	7%	59%

Table 2 The performance of the classification functions D, D_{MK}, D_{ML}, $(C_1, ..., C_6)$ is presented, i.e., the proportions of the patients based on the glucose data set that are correctly, misclassified or classified as belonging to an unknown population, respectively.

	D	D_{MK}	D_{ML}	(C_1, C_2)	$(C_1, ..., C_6)$
correctly	76%	79%	58%	52%	33%
misclassified	24%	21%	42%	12%	6%
unknown	-	-	-	36%	61%

The third data set is about the growth of Sitka spruce trees (Diggle et al. 1994; Pan & Fang 2002). A pollution of the ozone layer on the tree affects the growth. This study has two groups, $n_1 = 54$ trees grown with ozone exposure at 70 ppb and $n_2 = 25$ grown under controlled conditions. In the original data set, the response log tree size was taken over two seasons, but here, as Pan & Fang (2002), we only consider the data from the second growing season. The data is plotted in Figure 3 and in Table 3 we can find the proportions of correctly, misclassified and unknown classified trees.

The fourth data set is a potassium intake study (Dreier et al. 2021). With potassium intake one can lower blood pressure for hypertensive patients. However, the underlying mechanism is not fully understood, as increased potassium intake also raise the concentration of the hormone aldosterone, which is a

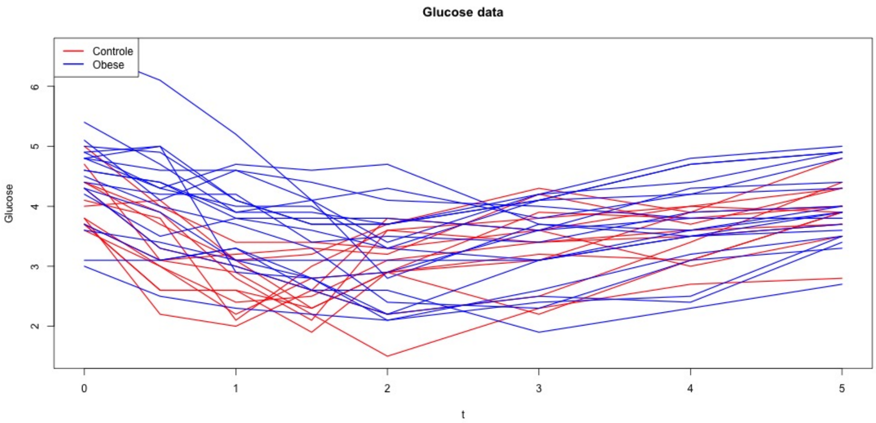

Fig. 2 Plot of glucose data. Eight repeated measurements at time 0, 0.5, 1, 1.5, 2, 3, 4 and 5 hours after the oral glucose intake for the two groups, control ($n_1 = 13$) and obese ($n_2 = 20$) are presented.

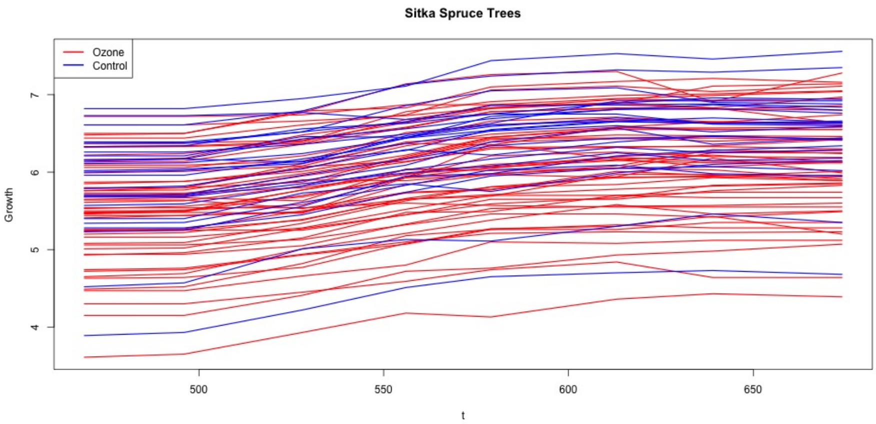

Fig. 3 Plot of Sitka spruce tree data. Eight repeated measurements at 469, 496, 528, 556, 579, 613, 639 and 674 days after start for the two groups, ozone ($n_1 = 54$) and control ($n_2 = 25$) are presented.

blood pressure-raising hormone. The data is a randomized placebo-controlled longitudinal crossover study where the effect of potassium supplement (90 mmol/day) on the aldosterone hormone was measured six repeated times at 0, 10, 20, 30, 45, 60. Since this study is a crossover study, for our purpose we only consider the data for the first intake for each individual, i.e., potassium ($n_1 = 13$) or placebo ($n_2 = 12$), and the second intakes are omitted. The repeated measurements are plotted in Figure 4, and in Table 4 we can find the proportions of correctly, misclassified and unclassified patients.

Table 3 The performance of the classification functions D, D_{MK}, D_{ML}, $(C_1, ..., C_6)$ is presented, i.e., the proportions of the Sitka spruce tree measurements that are correctly, misclassified or classified as belonging to an unknown population, respectively.

	D	D_{MK}	D_{ML}	(C_1, C_2)	$(C_1, ..., C_6)$
correctly	71%	62%	58%	47%	3%
misclassified	29%	38%	42%	19%	4%
unknown	-	-	-	34%	94%

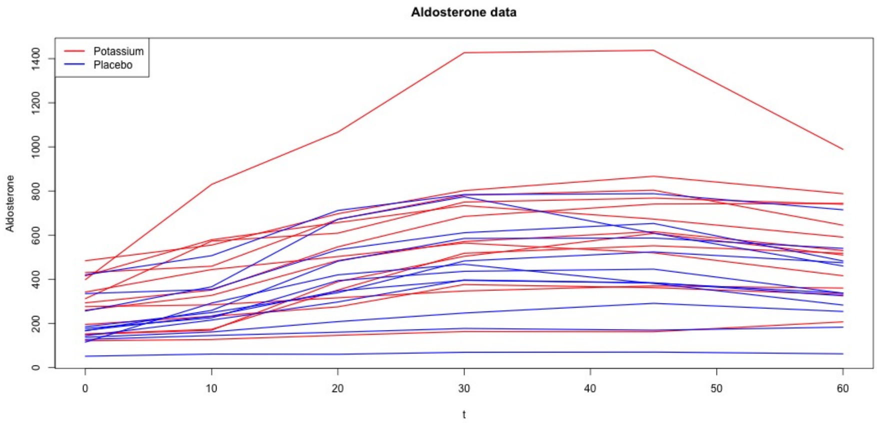

Fig. 4 Plot of aldosterone data. Six repeated measurements at 0, 10, 20, 30, 45 and 60 for the two groups, potassium intake ($n_1 = 13$) and placebo ($n_2 = 12$) are presented.

Table 4 The performance of the classification functions D, D_{MK}, D_{ML}, $(C_1, ..., C_6)$ is presented, i.e., the proportions of the patients based on the aldosterone measurements that are correctly, misclassified or classified as unknown, respectively.

Population	D	D_{MK}	D_{ML}	(C_1, C_2)	$(C_1, ..., C_6)$
correctly	72%	68%	72%	68%	36%
misclassified	28%	32%	28%	24%	8%
unknown	-	-	-	8%	56%

The final data set is calcium measurements (loss of calcium over time) of the dominant ulna bone in older women, at four time points 0, 1, 2 and 3 years (Johnson & Wichern 2007). Two groups are considered, a control group ($n_1 = 15$) and a group under treatment ($n_2 = 16$). The data is plotted in Figure 5, and in Table 5 we can find the proportions of correctly, misclassified and unclassified women.

As one can see in Tables 1-5, the performance of the classification functions D, D_{MK} and D_{ML} are rather similar and the correctly classified measurements are as low as 50% for some data sets. Non of the data sets are designed for classification and the profiles for each group are fairly similar, which can be depicted in Figures 1-5, hence the poor performance. Furthermore, when considering the two classifiers proposed in this paper, respectively based on

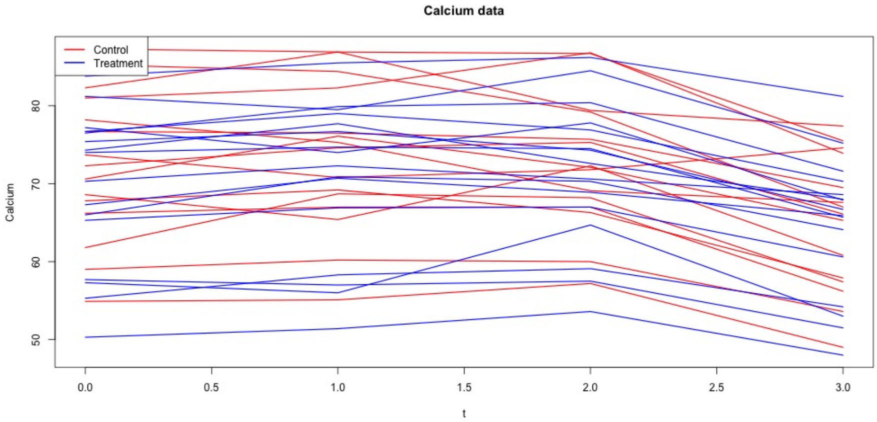

Fig. 5 Plot of calcium data. Four repeated measurements at 0, 1, 2 and 3 years for two groups, control ($n_1 = 15$) and under treatment ($n_2 = 16$) are presented.

Table 5 The performance of the classification functions D, D_{MK}, D_{ML}, $(C_1, ..., C_6)$ is presented, i.e., the proportions of the women in the calcium data that are correctly, misclassified or classified as belonging to an unknown population, respectively.

Population	D	D_{MK}	D_{ML}	(C_1, C_2)	$(C_1, ..., C_6)$
correctly	52%	52%	55%	52%	10%
misclassified	48%	48%	45%	45%	3%
unknown	-	-	-	3%	87%

(C_1, C_2) and $(C_1, ..., C_6)$, some of the measurements are classified as unknown, which seems to be reasonable due to the similar growths. Specifically, the final data set with calcium measurements one can see in Figure 5 that the two populations have very similar growths, and the performance given in Table 5 for the first four classification functions are almost like coin tossing (50-50). However, the proposed classifier based on $(C_1, ..., C_6)$, which is much more conservative, will only misclassify 3% of the measurements, but classify 87% as unknown.

We conclude that the proposed classifiers based on (C_1, C_2) and $(C_1, ..., C_6)$, have significant smaller number of misclassified observations compared to the other methods. However, $(C_1, ..., C_6)$ also deliver a smaller number of correctly classified object due to the vital number of observations that are classified as unknown.

References

Diggle, P.J., Liang, K.Y. & Zeger, S.L. (1994). *Analysis of Longitudinal Data*, Oxford University Press.

Dreier, R., Abdolalizadeh, B., Asferg, C.L., Hölmich, L.R., Buus, N.H., Forman, J.L., Andersen, U.B., Egfjord, M., Sheykhzade, M. & Jeppesen, J.L. (2021). Effect of increased potassium intake on the renin-angiotensin-aldosterone system and subcutaneous resistance arteries: a randomized crossover study. *Nephrology Dialysis Transplantation*, **36**, 1282–1291.

Johnson, R.A. & Wichern, D.W. (2007). *Applied Multivariate Statistical Analysis.* Prentice Hall, Upper Saddle River, New Jersey.

Pan, J.X. & Fang, K.T. (2002). *Growth Curve Models and Statistical Diagnostics.* Springer Science & Business Media, New York.

Potthoff, R.F. & Roy, S.N. (1964). A generalized multivariate analysis of variance model useful especially for growth curve problems. *Biometrika*, **51**, 313–326.