

CANONICAL ANALYSIS OF INDIVIDUAL VOCABULARY PROFILING DATA

Tapio Lokki and Kai Puolamäki

Aalto University School of Science and Technology, Department of Media Technology,
P.O.Box 15400, FI-00076 Aalto, Finland

ABSTRACT

Canonical analysis is an ordination technique which allows a direct comparison of two data matrices. It is often applied in ecological studies and here it is shown how canonical analysis, in particular redundancy analysis, can be used to analyze sensory evaluation data. The example data are obtained with individual vocabulary profiling study of concert hall acoustics. The redundancy analysis is compared with hierarchical multiple factor analysis. In addition, it is shown how redundancy analysis can be used to explain the subjective data with different objective data, in this case room acoustical parameters and physical measures of the studied concert halls.

Index Terms— subjective quality assessment, architectural acoustics, statistics

1. INTRODUCTION

The context of this paper is sensory evaluation of concert hall acoustics. The applied data are obtained from individual vocabulary profiling study in which 20 assessors rated concert halls with their own descriptive attributes. The resulting data are multivariate by nature and such data are often analyzed with some ordination method. Many studies have applied Multiple Factor Analysis (MFA) [1] and Hierarchical MFA (HMFA) [2], since they derive an integrated picture of the observations and of the relationships between the groups of variables. The basis of (H)MFA is in the Principal Component Analysis (PCA). (H)MFA enables of grouping of attributes by analyzing each group individually with PCA. Then it merges all PCA basis together to a unique matrix and performs a global PCA on it before projecting individual data sets onto global analysis.

(H)MFA is very useful in analyzing the results of subjective studies where assessors are applying individual vocabulary in sensory evaluation. The applied attributes can be grouped by assessors, and (H)MFA creates a global view of the data. (H)MFA is also a convenient tool to link objective and subjective data together, when some objective metrics of the data is available. (H)MFA rotates the subspaces

The research leading to these results has received funding from the Academy of Finland, project no. [119092] and the European Research Council under the European Community's Seventh Framework Programme (FP7/2007-2013) / ERC grant agreement no. [203636].

into one global space and ordination of objective and subjective attributes can be compared.

In addition to (H)MFA, there exists plenty of other methods to study multidimensional data. For example, in ecological studies canonical analysis is quite often applied [3]. Canonical analysis combines the concepts of ordination and regression by involving a response matrix and an explanatory matrix in the analysis. This is convenient in ecology where the data tables often contain sampling sites on rows and variables on columns as well as a matrix of explanatory variables related to environment, such as temperature, precipitation etc. Sensory evaluation data resemble the ecological data, since data matrices have samples on rows and variables on columns. In addition, explanatory data can be created by analyzing signals with objective metrics. However, the authors are aware of only one sensory evaluation study which applied canonical analysis in texture assessment of cheese [4].

Here, it is shown how canonical analysis can be applied to data obtained with individual vocabulary profiling (IVP) [5] of concert hall acoustics. The data collection and grouping of attributes are briefly reviewed to introduce the context of the study. The canonical analysis is explained briefly, and (H)MFA and canonical analysis results are combined. Finally, some features of canonical analysis are explained with examples.

2. CONTEXT AND DATA FROM INDIVIDUAL VOCABULARY PROFILING

The dataset applied in this study was obtained from a subjective evaluation of concert hall acoustics with IVP. The detailed description of the whole process is out of the scope of this paper, but is reported in [6, 7]. One of the main motivations in IVP is parallel comparison of samples, which has been earlier almost impossible when comparing concert halls. In this study the problem was solved by recording the concert halls with a virtual orchestra and playing back the recordings with spatial sound reproduction in an anechoic chamber.

In each concert hall, a symphony orchestra was simulated with 34 loudspeakers on the stage in the layout of a real orchestra, see Fig. 1. The loudspeakers were calibrated and they reproduced anechoic symphony orchestra recordings [8] of four different musical pieces. The sound capturing in three

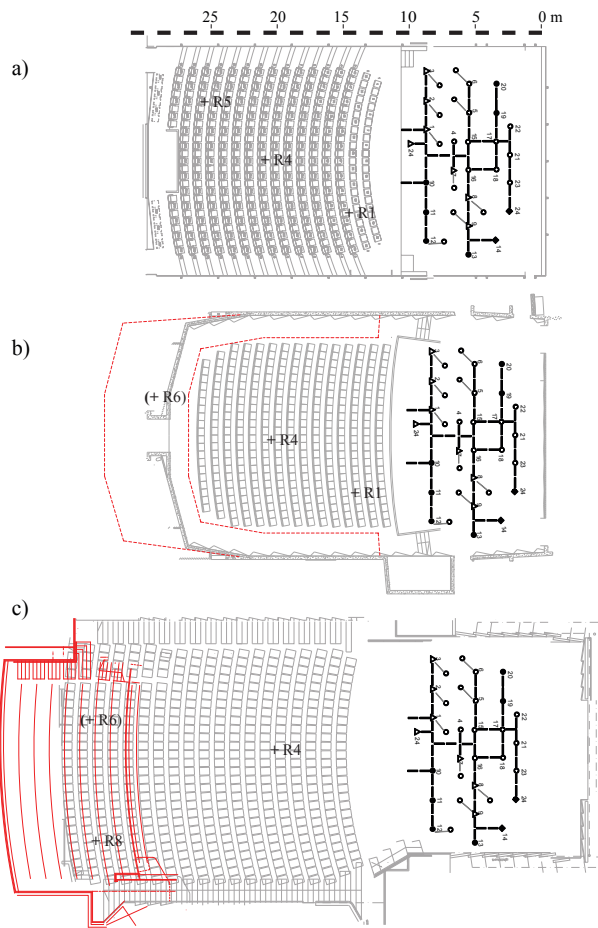


Fig. 1. Concert halls and recording positions. a) Sello hall (397 seats), recording positions R1, R4, and R5, b) Konservatorio hall (470 seats), recording positions R1, R4, and R6 (on balcony), c) Tapiola hall (690 seats), recording positions R4, R6 (on balcony), and R8.

receiver positions in each hall were made with a 3-D intensity probe microphone. Later, the recorded stimuli were processed with directional audio coding [9] and reproduced in the anechoic chamber with a 3-D loudspeaker setup consisting of 16 loudspeakers. This whole process enabled the creation of stimuli in which the only changed variables were a concert hall and a listening position, while the orchestra and music (excerpts of Mozart, Beethoven, Bruckner, and Mahler) as well as spatial sound recording and reproduction techniques remained constant.

Twenty screened assessors took part in the individual attribute elicitation process and they developed each from 4 to 6 attributes applied in the final rating of samples. In total 102 attributes were used in final rating where assessors compared 9 recordings by applying continuous scales which had values

Table 1. All collected 102 attributes grouped in 9 subgroups with hierarchical clustering and LDA.

Group	Attributes	N
Reverberance_1 (size of space)	X41, X77, X34, X103, X94, X60, X105	7
Reverberance_2 (envelopment)	X26, X3, X67, X86, X50, X106, X55, X61, X46, X5	10
Apparent Source Width	X39, X13, X80, X92, X71, X10, X7, X20, X95, X83, X107, X109, X99	13
Loudness	X37, X2, X43, X96, X8, X85, X57, X47, X69, X91	10
Distance	X82, X24, X28, X48, X44, X100, X88, X108, X97, X76, X18, X65	12
Ungrouped	X17, X74, X78, X66, X104, X64, X98, X9, X90, X23, X62	11
Balance	X31, X52, X11, X38, X6, X16, X111, X36	8
Openness	X15, X14, X73, X84, X70, X30, X79, X81, X4, X54, X75, X87, X89, X40, X32, X72	16
Definition (separability, clarity)	X27, X35, X102, X53, X59, X63, X101, X58, X110, X22, X56, X12, X33, X42, X19	15

from 0 to 120. The result of all ratings is collected to a data table containing 36 rows (9 positions \times 4 signals) and 102 columns, as depicted in Fig. 2. The grouping of attributes shown in Table 1 were performed with hierarchical agglomerative clustering, based on Euclidean distances, with Ward's minimum variance method. In addition, the groups were further identified with Linear Discriminant Analysis (LDA) [7]. In the following analysis, this grouping is used only for colors of attributes for better and more informative visualization.

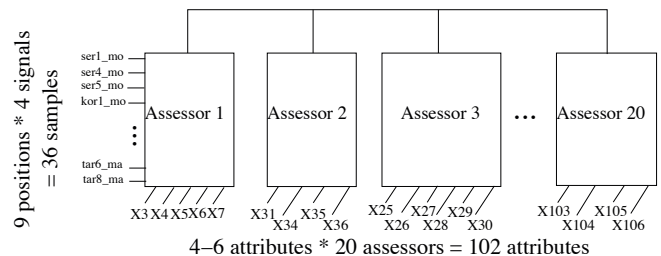


Fig. 2. The rows of data matrix consists of 4 signals \times 9 received positions (3 positions in 3 halls). The 102 columns contain 4, 5, or 6 attributes from 20 assessors.

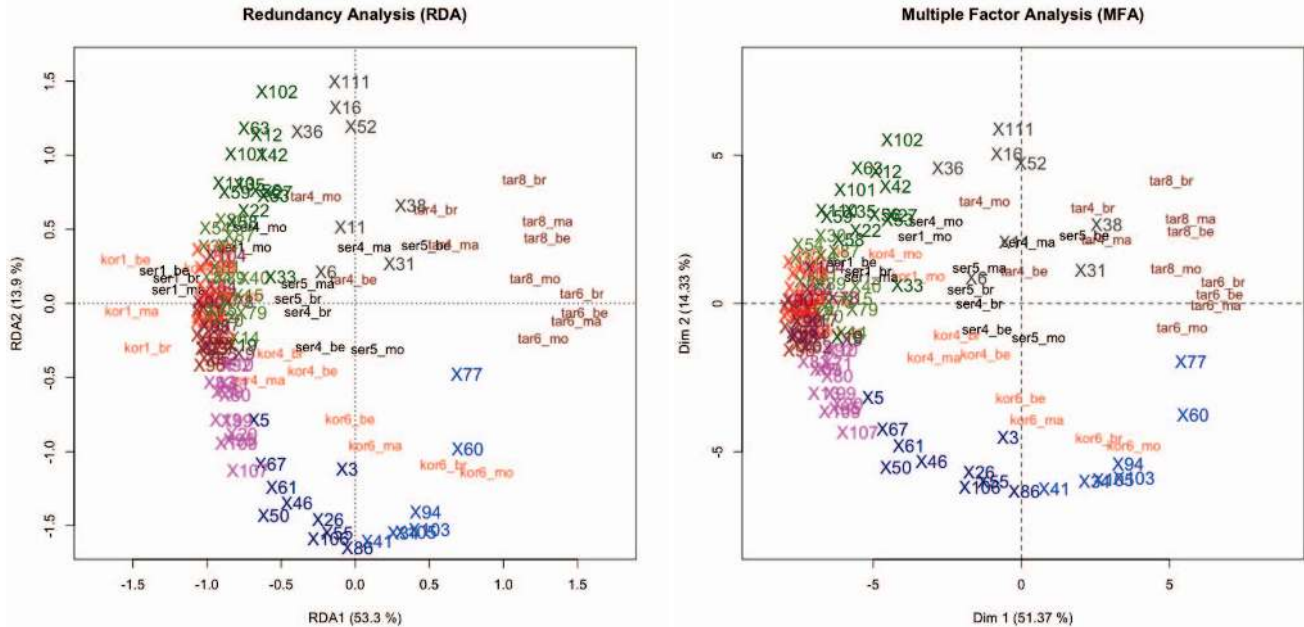


Fig. 3. RDA and MFA for the whole data. Attributes (X2 ... X111) are marked with colors introduced in Table 1. The listening positions are colored by concert halls and different signals are abbreviated with mo = Mozart, be = Beethoven, br = Bruckner, and ma = Mahler. For example, tar8_br means Tapiola hall (ta) at position R8 (r8) with excitation signal Bruckner (_br).

3. CANONICAL ANALYSIS AND EXAMPLES

Canonical analysis produces ordinations of matrix \mathbf{X} that is constrained to be related in some way to a second matrix \mathbf{Y} . The way in which the relationship between \mathbf{Y} and \mathbf{X} is established differs among methods of canonical analysis. Two most common variants are Redundancy Analysis (RDA) and Canonical Correspondence Analysis (CCA). The main difference between these two methods is the same as between PCA and correspondence analysis; RDA assumes multidimensional normal distribution of the data, while CCA assumes the data forms a contingency table. RDA is based on Euclidean distances and rotation of principal component basis, while CCA is based on χ^2 distances and weighted rotation. The more detailed description of canonical analysis and the mathematics is presented by Legendre and Legendre [3].

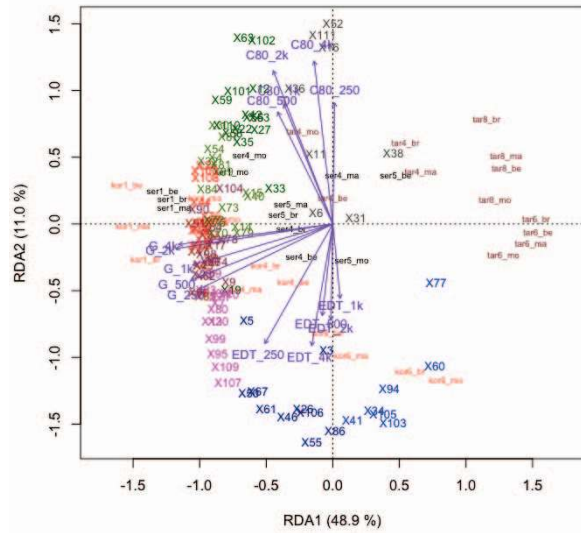
The data analyzed here are real valued and approximately normally distributed, thus it is reasonable to apply RDA. CCA was also tried, but the results did not change significantly. First, RDA was applied in its simplest form without any constraining explanatory matrix \mathbf{Y} . Thus, RDA reduces to a normal PCA for the whole data matrix. For comparison, the analysis was repeated with MFA. The analyses were performed with the FactoMineR [10] and vegan [11] packages to the centered and scaled data.

Figure 3 compares the RDA and MFA in which grouping was done by assessors (see Fig. 2). Both analyses provide

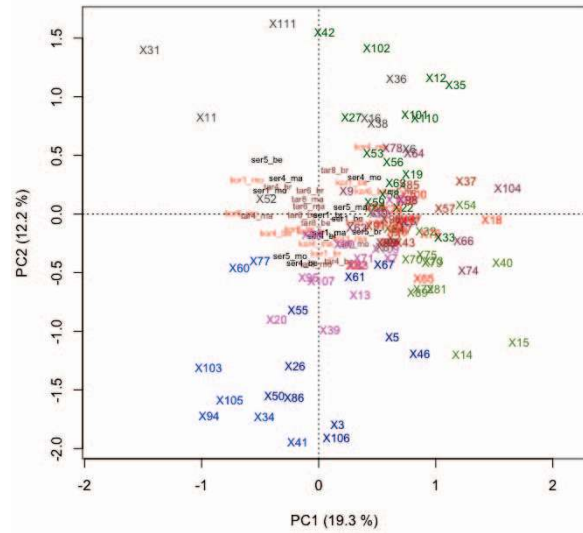
very similar results. The structure of subjective attributes is alike and 36 samples are positioned in the same manner on the plane defined by the first two principal components. In addition, the first two principal axes explain 67.2% of the total variance in RDA and 65.7% of the total variance in MFA.

When an explanatory matrix \mathbf{Y} is included to RDA, canonical analysis is related to multiple regression analysis. The analysis provides ordination of the linear subspace of the subjective data matrix \mathbf{X} which can be inferred via regression using the explanatory variables in \mathbf{Y} . For concert hall acoustics widely used objective metrics is standardized room acoustical parameters [12]. A set of parameters were computed from impulse responses, measured from 24 loudspeakers to each receiver position. The means of 24 values of these parameters were applied on five octave bands [250...4000 Hz]. The parameters are early decay time (EDT), clarity (C80), and strength (G). In total, the explanatory matrix \mathbf{Y} consists of 15 columns (3 parameters at 5 octave bands) and 36 rows (9 receiver positions replicated four times, since the number of rows has to be the same in \mathbf{X} and \mathbf{Y}).

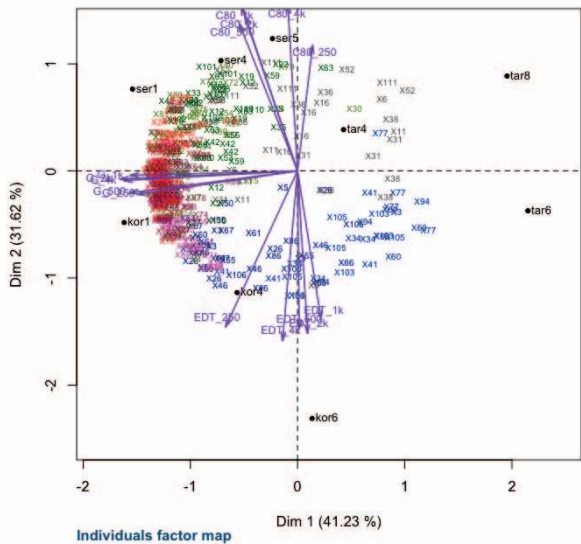
The result of RDA with subjective matrix \mathbf{X} and objective data \mathbf{Y} is shown in Fig. 4(a). The biplot can be interpreted so that a projection of a sample at right angle on a response variable approximates the value of the sample along that variable. The constraining matrix \mathbf{Y} explains 67.6% of the total variance of matrix \mathbf{X} and the first two RDA axes explains



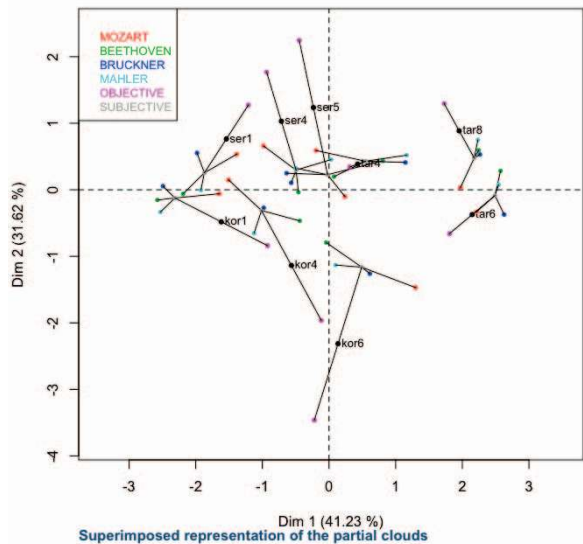
(a) RDA with constrained 15 objective parameters.



(b) RDA residuals.



(c) HMFA with subjective and objective data.



(d) Superimposed representation of subjective and objective data.

Fig. 4. RDA and HMFA analyses for subjective and objective data

59.9% of the total variance (88.6% of the explained variance). Figure 4(a) shows that constraint does not break the general structure of the data visible in Fig. 3 and therefore this structure is mostly explained by the objective variables. The best projections to this plane (= the longest vectors) seem to be with parameter G at all octave bands. They are also pointing to the same direction as Loudness which make perfect sense since G measures total energy of an impulse response. Other

ordinations are not so perfectly linked since EDT is pointing to enveloping Reverberance_2 direction, but not in Reverberance_1 direction which describes the perception of size of space through reverberation. In addition, C80 is considered to be a measure of clarity, but here C80 vectors are slightly deviating from directions of attributes in group Definition.

Since RDA could only explain 67.6% of the total variance of matrix X, the unconstrained 32.4% remains as RDA resid-

uals. The nice feature of RDA is that these residuals are the original attributes from which the effect of constraining data is removed. Therefore, by plotting RDA residuals gives information of the unexplained variance, see Fig. 4(b). In this case the first two PC axes explain only 31.5% of the unconstrained variance, meaning that the shown two-dimensional plane does not represent the data very well. However, it can be seen that there is still some structure left in subjective attributes. Some of *Reverberance_1* and *Reverberance_2* attributes are still clustered to a certain direction far from origo, meaning that some unknown objective parameter might explain the variance of these attributes. A few *Balance* attributes have large variance and attributes are scattered all over, suggesting that no objective attribute could explain their variance. Such structure is not readily available in HMFA analysis which can also link objective and subjective data.

For comparison, HMFA analysis with two hierarchical levels is done to matrices \mathbf{X} and \mathbf{Y} . First level consists of four MFAs (different signals) with data of 102 attributes (9 rows in each). The second level then links the PCA of 15 objective parameters (EDT, G, C80 at octave bands [250...4000]) with the combined MFA of four signals. Fig. 4(c) shows the HMFA result and it can be seen that the overall structure of the data is similar than in 4(a). There are now 4×102 subjective attributes due to the applied hierarchy. The plot shows that individual attributes have moved slightly, which is due to different variance explaining percentages of the main dimensions, in total 72.8%. HMFA could not extract the residuals as RDA, but it has some other nice features, such as superimposed representation of partial clouds, as depicted in Fig. 4(d). There the mapping of individual samples on the main plane is visualized so that contribution of different musical signals is seen. It is also seen that objective parameters are loaded more to the second dimension than the subjective data.

3.1. Testing different constraining parameters

The constraining data can also be binary vectors, e.g., Fig. 5(a) shows analysis where different concert halls are indicated with binary data. RDA explains 35.9% of the total variance and halls are clearly separated. In addition, directions agree well with general opinions of these three halls. Tapiola is usually considered too quiet, it does not support the musicians, thus, samples are located in the opposite direction than *Loudness*. Konservatorio is too reverberant and quite loud. Sello has good clarity, the analysis shows the *Definition* and *Openness* attributes in the same direction than the Sello vector. The unexplained variance (64.1%) is plotted in Fig. 5(b), but the interpretation of the plot is not obvious.

Another example is given with ad-hoc physical parameters, such as distance to orchestra and distances to front, back, left, and right walls from each listening position. Such constraining data explain 66.4% of the total variance, as shown in Fig. 5(c). The unconstrained 33.6% variance (Fig. 5(d)) is

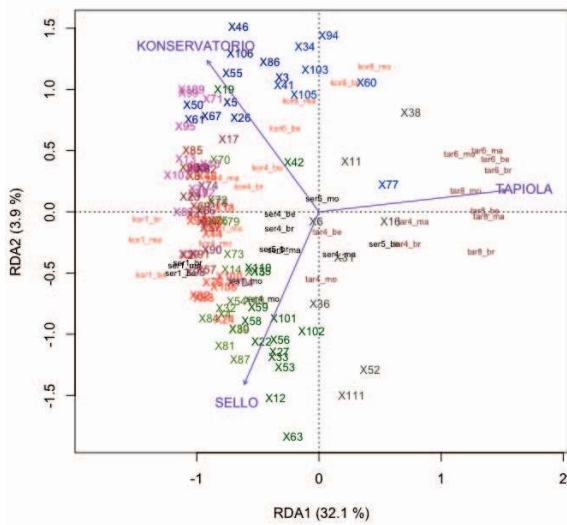
interestingly almost identical to RDA residuals of objective room acoustical parameters (Fig. 4(b)). This final example suggests that any objective data related to halls can explain only 2/3 of the variance of the subjective data. To explain the rest some signal dependent parameters have to be found.

4. CONCLUSIONS

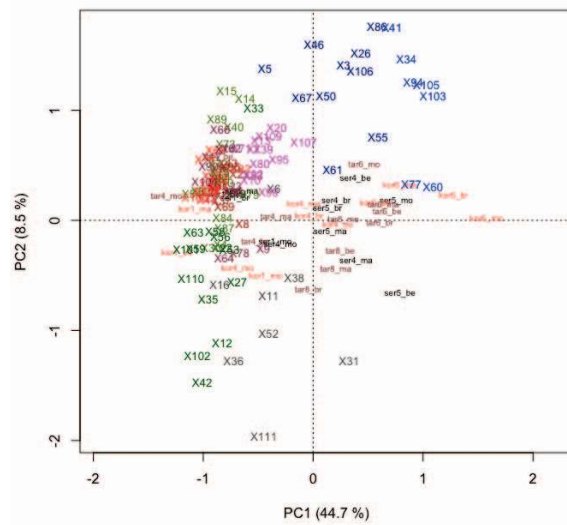
Canonical analysis is applied in the context of sensory evaluation. With various examples, it is shown that canonical analysis is a useful tool and complements the more often applied analysis methods, such as (H)MFA. In particular, canonical analysis gives estimates how well some objective data explains the variance of multidimensional subjective data. In addition, RDA residuals are shown to give additional information of the data structures that are not explained with the constraining data.

5. REFERENCES

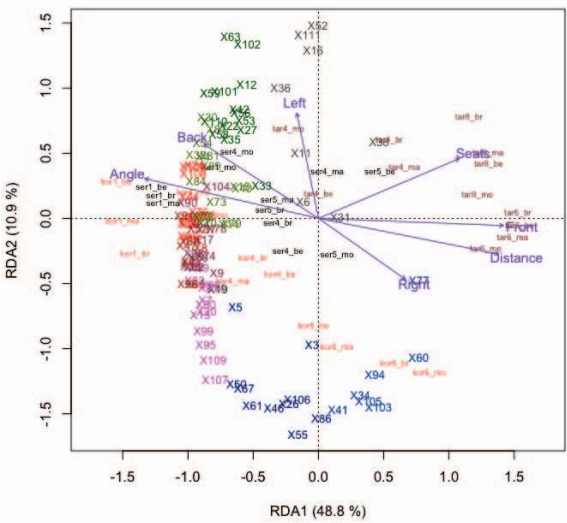
- [1] B. Escofier and J. Pagès, "Multiple factor analysis," *Computational Statistics & Data Analysis*, vol. 18, no. 121–140, 1990.
- [2] S. Le Dien and J. Pagès, "Hierarchical multiple factor analysis: application to the comparison of sensory profiles," *Food Quality and Preference*, vol. 14, pp. 397–403, 2003.
- [3] P. Legendre and L. Legendre, *Numerical Ecology*, Elsevier, 2 edition, 1998.
- [4] K.D. Antoniou, D. Petridis, S. Raphaelides, Z. Ben Omar, and R. Kesteloot, "Texture assessment of french cheeses," *Journal of Food Science*, vol. 65, no. 1, pp. 168–172, 2000.
- [5] G. Lorho, "Individual vocabulary profiling of spatial enhancement system for stereo headphone reproduction," in *119th AES Convention*, New York, NY, USA, October 7-10 2005, Paper 6629.
- [6] A. Kuusinen, H. Vertanen, and T. Lokki, "Assessor selection and behavior in individual vocabulary profiling of concert hall acoustics," in *AES 38th Int. Conf. on Sound Quality Evaluation*, Piteå, Sweden, June 13-15 2010.
- [7] T. Lokki, J. Pätynen, A. Kuusinen, H. Vertanen, and S. Tervo, "Concert hall acoustics assessment with individually elicited attributes," Unpublished manuscript, 2010.
- [8] J. Pätynen, V. Pulkki, and T. Lokki, "Anechoic recording system for symphony orchestra," *Acta Acustica united with Acustica*, vol. 94, no. 6, pp. 856–865, November/December 2008.



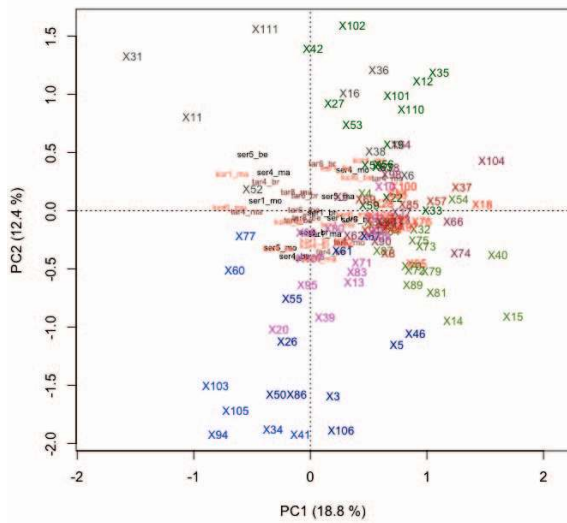
(a) RDA, X = subjective attributes, Y = binary data indicating concert halls.



(b) RDA residuals when conditioned “concert hall” removed.



(c) RDA, X = subjective attributes, Y = physical measures related to listening positions.



(d) RDA residuals when conditioned physical parameters removed.

Fig. 5. Example RDA analysis with different constraining data.

- [9] J. Vilkkamo, T. Lokki, and V. Pulkki, “Directional audio coding: Virtual microphone based synthesis and subjective evaluation,” *Journal of the Audio Engineering Society*, vol. 57, no. 9, pp. 709–724, September 2009.
- [10] Francois Husson, Julie Josse, Sebastien Le, and Jeremy Mazet, *FactoMineR: Factor Analysis and Data Mining with R*, 2009, R package version 1.12.
- [11] Jari Oksanen, Roeland Kindt, Pierre Legendre, Bob O’Hara, Gavin L. Simpson, Peter Solymos, M. Henry H. Stevens, and Helene Wagner, *vegan: Community Ecology Package*, 2009, R package version 1.15-4.
- [12] ISO 3382-1:2009, *Acoustics – Measurement of room acoustic parameters – Part 1: Performance spaces*, International Standards Organization, 2009.