MAM Intro
Scaling correction of sensory data

Per Bruun Brockhoff, ***Professor***

DTU Compute,

Danish Technical University

www.compute.dtu.dk/staff/perbb

Taking individual scaling differences into account by analyzing profile data with the Mixed Assessor Model.

Per Bruun Brockhoff, DTU Comput2, Technical University, Denmark

Pascal Schlich, INRA, Cente des Sciences du Goût et de l'Alimentation, France

Ib M. Skovgaard, Department of Basic Sciences and Environment, University of Copenhagen, Denmark

Brockhoff, P. B., Schlich, P., & Skovgaard, I. (2015). Taking individual scaling differences into account by analyzing profile data with the Mixed Assessor Model. FQP, 39, 156-166.

MAM Basics Summary

- **Individual differences in the use of the scale range in sensory profiling is inherent!**
- **The usual mixed model ANOVA for sensory profile data does NOT fully account for this!**
- **An easy method (MAM) exist for mending this!**
- **It gives more powerfull analysis and improved insight!**

Analysis of Variance

Data from:

Brockhoff and Skovgaard (1994): Modelling individual differences between assessors in sensory evaluations. *Food Quality and Preference* 5, 215-224.

Source	SS	DF	MS	F	P Value
Assessor	1636.8	6	272.8	9.60	<0.001
Product	1329.5	4	332.4	11.70	<0.001
Interaction	681.8	24	28.41	5.42	<0.001
Error	569.57	105	5.24		

"Usual" mixed model ANOVA

Analysis of Variance

Data from:

Brockhoff and Skovgaard (1994): Modelling individual differences between assessors in sensory evaluations. *Food Quality and Preference* 5, 215-224.

Source	SS	DF	MS	F	P Value
Assessor	1636.8	6	272.8	9.60	<0.001
Product	1329.5	4	332.4	11.70	<0.001
Interaction	681.8	24	28.41	5.42	<0.001
Error	569.57	105	5.24		

"Usual" mixed model ANOVA

Analysis of Variance

Data from:

Brockhoff and Skovgaard (1994): Modelling individual differences between assessors in sensory evaluations. *Food Quality and Preference* 5, 215-224.

Source	SS	DF	MS	F	P Value
Assessor	1636.8	6	272.8	9.60	<0.001
Product	1329.5	4	332.4	11.70	<0.001
Interaction	681.8	24	28.41	5.42	<0.001
Error	569.57	105	5.24		

"Usual" mixed model ANOVA

Analysis of Variance

Data from:

Brockhoff and Skovgaard (1994): Modelling individual differences between assessors in sensory evaluations. *Food Quality and Preference* 5, 215-224.

Source	SS	DF	MS	F	P Value
Assessor	1636.8	6	272.8	9.60	<0.001
Product	1329.5	4	332.4	11.70	<0.001
Interaction	681.8	24	28.41	5.42	<0.001
Error	569.57	105	5.24		

"Usual" mixed model ANOVA

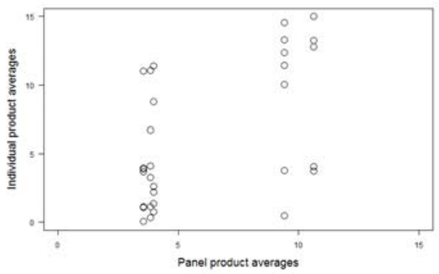
Analysis of Variance (Mixed)

$$F = \frac{MS_{Prod}}{MS_{Prod*Assessor}} = \frac{332.4}{28.41} = 11.70$$

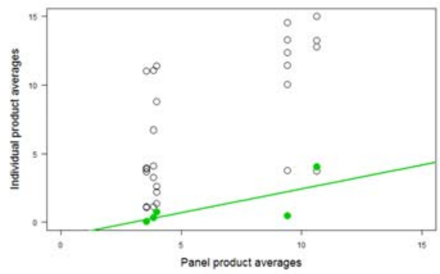
Uncertainty of product difference:
(for post hoc analysis)

$$SE = \sqrt{2MS_{Prod*Assessor}/(7\cdot 4)} = 1.42$$

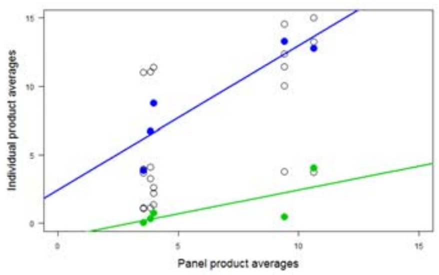
Individual product averages vs panel product averages:



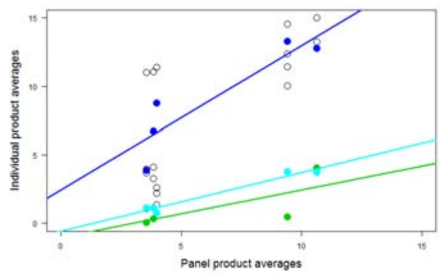
Individual product averages vs panel product averages:



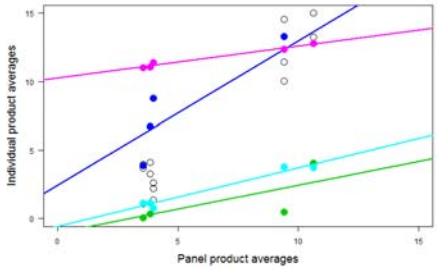
Individual product averages vs panel product averages:



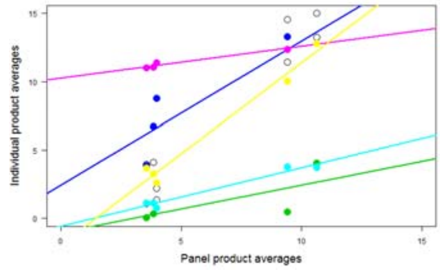
Individual product averages vs panel product averages:



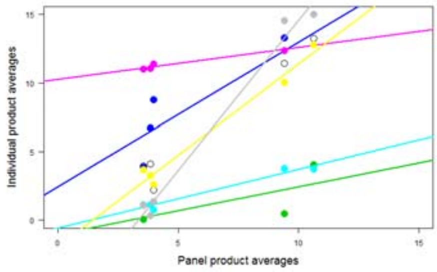
Individual product averages vs panel product averages:



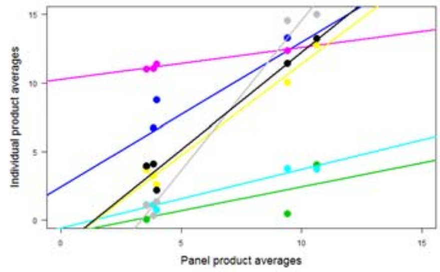
Individual product averages vs panel product averages:



Individual product averages vs panel product averages:



Individual product averages vs panel product averages:



The original Assessor Model, 1994

- Result: A ”pure scaling difference” model fits to these data!
- Random interaction model is WRONG!
- In general:

Interaction = Scaling differences + Disagreement

A simple method of analysis

Usual ANOVA model:

$$Y_{ijk} = \mu + \alpha_i + \upsilon_j + \delta_{ij} + \varepsilon_{ijk}$$

A simple method of analysis

Use the product (centered) averages as a (fixed) covariate in the model:

$$Y_{ijk} = \mu + \alpha_i + \upsilon_j + \underbrace{\beta_i x_j}_{\text{Scaling}} + \underbrace{d_{ij}}_{\text{Disagreement}} + \varepsilon_{ijk}$$

Interaction ($\beta_i x_j + d_{ij}$)

Scaling — Disagreement

Assessor model approximation + random interaction term

Linear mixed model (use Type I tests)

NEW Analysis of Variance

Source	SS	DF	MS	F	P Value
Assessor	1636.8	6	272.8	9.60	<0.001
Product	1329.5	4	332.4	57.69	<0.001
Scaling	578.10	6	96.35	16.72	<0.001
Disagree	103.71	18	5.76	1.06	0.4003
Error	569.57	105	5.24		

DTU Compute
Department of Applied Mathematics and Computer Science

NEW Analysis of Variance

Source	SS	DF	MS	F	P Value
Assessor	1636.8	6	272.8	9.60	<0.001
Product	1329.5	4	332.4	57.69	<0.001
Scaling	578.10	6	96.35	16.72	<0.001
Disagree	103.71	18	5.76	1.06	0.4003
Error	569.57	105	5.24		

DTU Compute
Department of Applied Mathematics and Computer Science

NEW Analysis of Variance

Source	SS	DF	MS	F	P Value
Assessor	1636.8	6	272.8	9.60	<0.001
Product	1329.5	4	332.4	57.69	<0.001
Scaling	578.10	6	96.35	16.72	<0.001
Disagree	103.71	18	5.76	1.06	0.4003
Error	569.57	105	5.24		

DTU Compute
Department of Applied Mathematics and Computer Science

NEW Analysis of Variance

$$F = \frac{MS_{Prod}}{MS_{Disagreement}} = \frac{332.2}{5.76} = 57.69$$

Uncertainty of product difference: (for post hoc hypothesis testing)

$$SE = \sqrt{2MS_{Disagreement} / (7 \cdot 4)} = 0.64$$

DTU Compute
Department of Applied Mathematics and Computer Science

Sensobase Investigation
www.sensobase.fr

- **477 data sets**
- **8091 attributes**

At Centre Européen des Sciences du Goût, a project is conducted to build a database of sensory profiling datasets in which the data providers can exchange their sensory profiling data for statistical analyses.

DTU Compute
Department of Applied Mathematics and Computer Science

Sensobase results (8091 attributes) Interaction structure

- 46% show scaling heterogeneity
- 47% show (usual) interaction
- 30% show significant disagreement
- Among the (usual) interaction cases:
  - 58% show scaling heterogeneity
  - 38% NO further significant disagreement

DTU Compute
Department of Applied Mathematics and Computer Science

Sensobase results (8091 attributes) Product structure

- 60% show product differences by ORIGINAL approach
- 66% show product differences by NEW approach
- Among the scaling heterogeneous cases:
  - 43% of the NS cases (1034) become SIGNIFICANT

DTU Compute
Department of Applied Mathematics and Computer Science

P-value Classes

Classes
>0.20
0.10-0.20
0.05-0.10
0.01-0.05
0.001-0.01
<0.001

DTU Compute
Department of Applied Mathematics and Computer Science

P-value Classes Product differences

Classes	Original
>0.20	45%
0.10-0.20	34%
0.05-0.10	
0.01-0.05	
0.001-0.01	
<0.001	22%

DTU Compute
Department of Applied Mathematics and Computer Science

Sensobase results (8091 attributes) Product structure

- 34% have P-values between 0.001 and 0.20 by ORIGINAL approach
  - 43% of these by move DOWN in P-value class by NEW approach
  - 11% of these by move DOWN MORE than ONE P-value class by NEW approach

DTU Compute
Department of Applied Mathematics and Computer Science

Confidence Intervals (CIs) for pairwise differences

- Simple correction method OK for overall F test AND post-hoc pairwise null hypotheses.
- BUT for CIs a NEW method is necessary – presented in the paper – no details here.
- A potential large difference value induces larger variance than a smaller one – due to the ”random scaling effect”

DTU Compute
Department of Applied Mathematics and Computer Science

MAM BASIC

– One-way product structure
– Complete sensory attribute data
– WITH replications

SO:

– Simple ANOVA decompositions
– Simple links to performance measures

MAM BASICS in R:
CONDITIONAL and *ADJUSTED* MAM

CONDITIONAL:

ONLY scale correct IF Scaling is significant

ADJUSTED:

ONLY scale correct for the positive scalings
(leave the negatives as part of disagreement)

MAM analysis in R

1. Simple data structure implementation (based on ANOVA decompositions)

2. General covariate-based

Individual performance measures

Overall ANOVA table:

Source	SS	DF
Assessor	SS(Ass)	I-1
Product	SS(Prod)	J-1
Scaling	SS(Scal)	I-1
Disagreement	SS(Dis)	(I-1)(J-2)
Error	SS(Error)	IJ(K-1)

Individual performance measures

Decompose further into individual contribututions:

Source	Ass 1	Ass 2	....	Ass 7	SS
Assessor					SS(Ass)
Product					SS(Prod)
Scaling					SS(Scal)
Disagree					SS(Dis)
Error					SS(Error)

Individual performance measures

	A 1	A 2	A 3	A 4	A 5	A 6	A 7
LEVEL	-5.15	2.8	-4.17	5.43	0.19	0.17	0.71
PRODUCT	1.66	4	1.51	0.81	4.63	7.6	4.98
SCALING	0.35	1.05	0.43	0.23	1.33	2.19	1.42
CORRELLATION	0.72	0.91	0.98	0.99	0.99	0.99	0.98
DISAGREEMENT	2.66	3.89	0.68	0.3	1.62	2.15	2.13
REPEATABILITY	2.98	3.19	1.27	0.85	2.56	0.69	3.1

P-value:	<0.001	<0.01	<0.05

MAM-CAP Table

C. Peltier, M. Visali, P. B. Brockhoff & Schlich, P. (2014). The mam-cap table: a new tool for monitoring panel performances.
Food Quality and Preference, Vol. 32, 24-27.

DTU Compute
Department of Applied Mathematics and Computer Science

Take home

- Individual differences in the use of the scale in sensory profiling is inherent!
- The usual mixed model ANOVA for sensory profile data does NOT fully account for this!
- An easy method (MAM) exist for mending this!
- It gives more powerfull analysis and improved insight!

DTU Compute
Department of Applied Mathematics and Computer Science

MAM analysis in R (basic)

Produces 6 results data structures

↓

result[[1]]
result[[2]]
result[[3]]
result[[4]]
result[[5]]
result[[6]]

DTU Compute
Department of Applied Mathematics and Computer Science

MAM analysis in R (basic)

result[[1]]: *Individually split ANOVA tables*
result[[2]]: *Performance tests (MAM-CAP-like)*
result[[3]]: *MAM ANOVA tables*
result[[4]]: *MAM post hoc I: ”pair-wise”*
result[[5]]: *MAM post hoc II: ”diff from mean”*
result[[6]]: *MAM post hoc III: NEW CLs (pair-wise)*

DTU Compute
Department of Applied Mathematics and Computer Science