Many small data can make *a* big data *base*.
The SensoBase story

Pascal Schlich[1], Caroline Peltier[1], Nadra Mammasse[1,2], Nicolas Pineau[1,3]

[1]INRA, Centre des Sciences du Goût et de l'Alimentation, Dijon, France
[2]AB Science, Paris, France
[3]Nestlé Research Centre, Lausanne, Swisszerland

pascal.schlich@dijon.inra.fr

Plan

INTRODUCTION

- The SensoBase project
- Meta-analysis methodology

META-ANALYSIS RESULTS

- Documenting sensory practices & benchmarking panel performances
- Optimizing sensory panels
- Comparing sensometric techniques

CONCLUSION

- Difficulties and limits
- Conclusion

The SensoBase project

- How to valorize numerous SAS macros developed in 1988-2003
- Exchanging raw QDA data for free statistical analysis!
- The first team : Sylvie Cordelle, Delphine Brajon and Nicolas Pineau (2003-2006)
- Nicolas Pineau's Ph.D (2003-2006)
- Nadra Mammasse's Ph.D (2009-2012):
- Caroline Peltier's Ph.D (2012-2015):
- 3 main objectives:
  1. Documenting practices & benchmarking panel performances
  2. Optimizing sensory panels
  3. Comparing sensometric techniques

The former sensoBase website now replaced by…

SENSOBASE Program

www.sensobase.fr

Applications | Météo heure par heur | Biblio | Voyages | ChemoSens - CSGA | Perso | Stat Links | Groupes | Imprimante | Restauration hors do

LABORATOIRE D'INTERFACE RECHERCHE INDUSTRIE SENSOMETRIE

SENSOBASE

Accueil

ESPACE INSCRITS

Identifiant :

Mot de passe :

Entrer

Le projet SENSOBASE

Déroulement de l'échange

Extrait de résultats statistiques et d'aide à l'interprétation

Vous inscrire

Envoyez nous vos suggestions

- communications Pangborn 2005
- Drinktec 2005 P.Schlich
- Atelier DTS, RSIG 2005, P.Schlich

SENSOBASE

Une base de données de profils sensoriels

Au **Centre Européen des Sciences du Goût**, l'équipe SENSOBASE du LIRIS vous accueille dans le cadre de son programme d'échange

"Jeux de données contre statistiques".

Avis important concernant l'avenir de la SensoBase :

Vous contribuez régulièrement à la croissance de la base de données SensoBase et nous vous en remercions.

Au 12 janvier 2015, le site internet de la SensoBase, qui jusqu'à présent vous permettait de déposer vos jeux de données dans la base et de demander des analyses statistiques en retour, sera fermé définitivement.

Le programme SensoBase ne disparaît pas pour autant. Vous pourrez continuer à déposer des jeux de données et à les analyser, en utilisant le logiciel TimeSens que nous développons actuellement (voir le lien http://www.timesens.com pour plus de précisions). Toutes les analyses que vous obteniez jusqu'à présent sur le site de la SensoBase seront disponibles dans le logiciel TimeSens.

Pour vous remercier d'avoir contribué à l'essor de la SensoBase, vous bénéficierez d'un accès gratuit au module d'analyses statistiques du logiciel TimeSens, jusqu'à fin juin 2015. Par la suite, vous pourrez continuer à utiliser TimeSens et à bénéficier des analyses de vos jeux de données aux conditions d'utilisation du logiciel TimeSens.

Nous restons à votre disposition pour tout renseignement complémentaire,

L'équipe de la Sensobase.

Important notice regarding the future of the Sensobase program:

You have regularly contributed to the development of the Sensobase database and we thank you for that.

On January 12th 2015, the Sensobase website, which until now allowed you to put your sensory datasets in the database and to request statistical analyses in return, will be closed permanently.

However, the Sensobase program does not disappear. You will be able to file your datasets and to analyze them, using the TimeSens software that we are developing (see that link for more details: http://www.timesens.com/). All analyzes that you are used to get on the Sensobase website will be available in the TimeSens software.

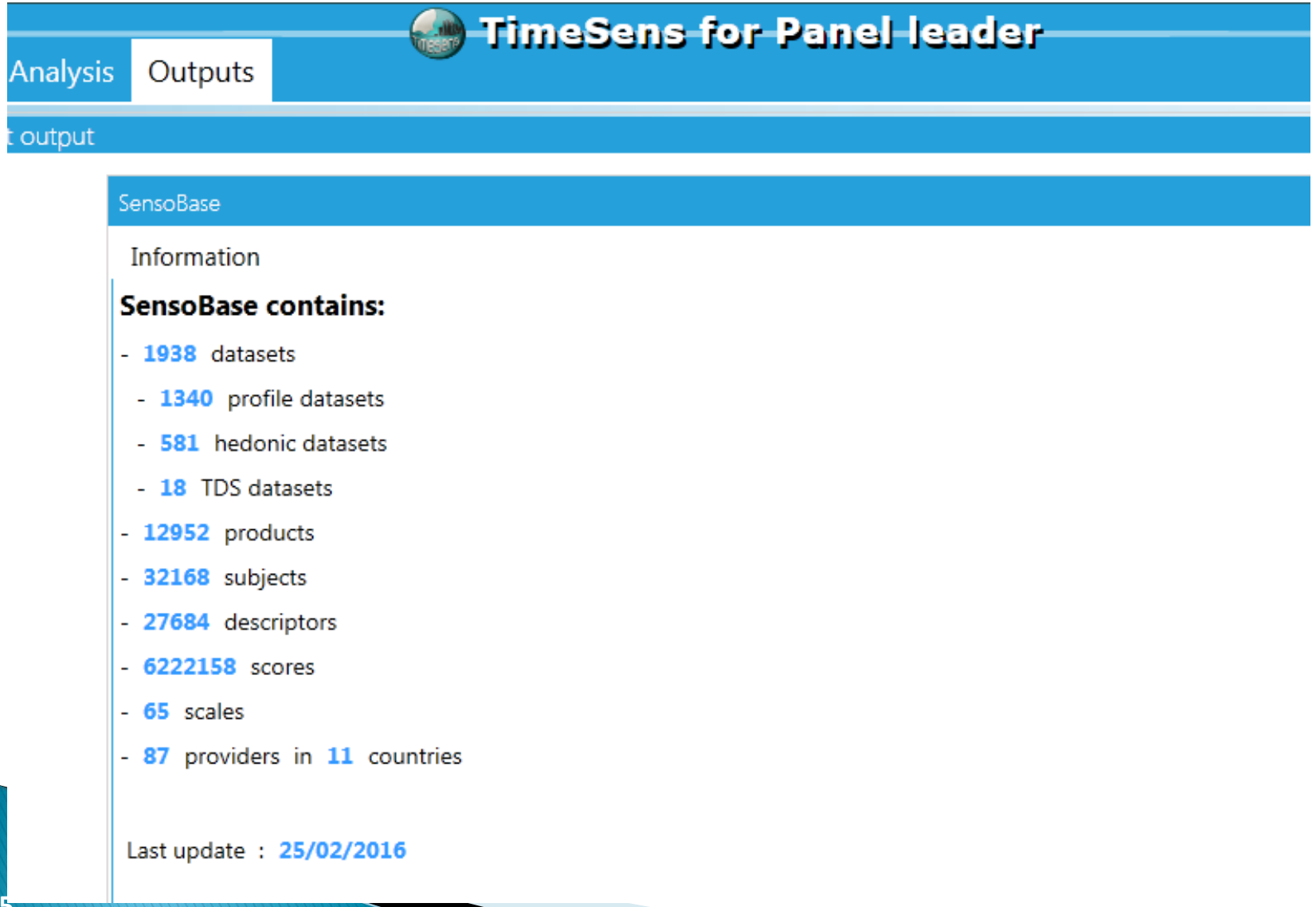
To thank you for having contributed to the development of the Sensobase database, you will have free access to the statistical analysis module in the TimeSens software until the end of June 2015. Thereafter, you will be allowed to use TimeSens and to obtain the analysis of your datasets under the conditions of use of the TimeSens software.

We remain at your disposal for further information,

The team of the Sensobase program.

FR 08:18 17/03/2016

TimeSens: current content of SensoBase datasets in Timesens



Meta-analysis methodology

Two methods to test the effect of one individual characteristic on performances:

Method A (Pineau, 2006): subject as experimental unit

Dataset analysed

Dataset	Sub	Char	Perf
D1	S1	Male	0.8
D1	S2	Male	1.2
D1	S3	Female	1
D1	S4	Female	1.5
D1	S5	Female	2
D2	…	…	…

Applying the weighted ANOVA model:

perf ~ char + dataset + char*dataset

with dataset as a random effect

→ ***Many observations but unbalanced design***

Method B (Peltier, 2015): dataset by characteristic level as experimental unit

Dataset analysed

Dataset	Char	Perf
D1	Male	1
D1	Female	1.5
D2	Male	2
D2	Female	0.9
…	…	…

Applying the weighted ANOVA model:

perf ~ char + dataset

→ ***Few observations but balanced design***

Sensobase contents

- 1316 « clean » datasets from 66 data providers, 38 from France and 28 from 16 foreign countries
- The « median dataset » includes 5 products, 11 panelists and 22 attributes
- Number of replicates: none 31% - 2 reps 52% - 3 reps 10%

Examples of two tables describing SensoBase extracted from Peltier's thesis

Echelle utilisée	Effectif	Pourcentage
Echelle inconnue ou non renseignée	5 271	27.11 %
Echelle continue rectangulaire	493	2.54 %
Echelles continues linéaires fermées	7 670	39.44 %
Echelles continues linéaires ouvertes	502	2.58 %
Echelles discrètes	5 458	28.07 %
Autre	51	0.27 %
Total	19 445	100 %

TABLEAU 15 : ECHELLES UTILISEES DANS LA SENSOBASE

Famille de descripteur	Effectif	Pourcentage
Non renseignée	92	0.47 %
Apparence	2 518	12.95 %
Odeur	4 045	20.80 %
Saveur	3 129	16.09 %
Texture	3 509	18.05 %
Arôme/flaveur	4 531	23.30 %
Arrière-gout	1 006	5.17 %
Trigéminal	611	3.14 %
Hédonique	4	0.02 %
Total	19 445	100 %

TABLEAU 16 : FAMILLES DE DESCRIPTEURS DANS LA SENSOBASE

Benchmarking repeatability

Average standard deviations over replicates on a 0-10 scale (number of panelist x product x attribute)

Type of Food	Taste	Appearance	Aftertaste	Aroma	Trigeminal	Texture	Odor	TOTAL
Dairy products	**0.61 (65931)**	0.83 (19913)	0.72 (12572)	0.85 (118091)	0.83 (3657)	0.97 (69738)	**0.69 (34602)**	**0.80 (a)**
Others	0.77 (1191)	0.96 (618)	NA	0.96 (1369)	0.95 (93)	0.93 (3772)	**0.78 (284)**	0.90 (b)
Sweets	0.87 (12458)	1.13 (955)	**0.29 (1788)**	1.22 (1616)	0.94 (9833)	0.93 (1854)	1.27 (1546)	0.90 (b)
Drink without alcohol	0.86 (6001)	**0.55 (1460)**	0.9 (3260)	0.94 (15996)	0.97 (1461)	**0.75 (2062)**	1.16 (4547)	0.92 (b)
Fruits	0.93 (2876)	1.08 (1994)	0.72 (65)	0.94 (2198)	**0.77 (657)**	1.08 (3503)	0.94 (1659)	0.99 (bc)
Oils	NA	NA	NA	1.02 (498)	NA	NA	NA	1.02 (bc)
Delicatessen	0.97 (1308)	1.22 (2595)	1.57 (58)	**0.81 (1996)**	1.14 (264)	1.01 (2746)	1.11 (281)	1.03 (c)
Fish	0.89 (204)	1.16 (427)	NA	1.06 (258)	0.95 (43)	1.12 (249)	1.03 (41)	1.07 (cd)
Condiments	0.86 (9632)	0.89 (20448)	**1.32 (13492)**	1.18 (20970)	1.06 (5714)	1.06 (20064)	1.17 (28632)	1.09 (d)
Feculents	1.00 (2299)	**1.38 (447)**	1.22 (168)	1.09 (7924)	0.29 (12)	**1.39 (6294)**	**1.39 (244)**	1.20 (e)
Meats	1.03 (189)	1.26 (1205)	NA	1.15 (195)	NA	1.13 (613)	1.25 (545)	1.20 (e)
Bread	0.92 (1386)	1.16 (2979)	NA	1.02 (3218)	0.88 (133)	**1.47 (7046)**	1.06 (4631)	1.21 (e)
Prepared dishes	1.13 (3585)	1.24 (8597)	NA	1.26 (4684)	1.11 (2336)	**1.46 (4187)**	1.07 (5511)	1.22 (e)
Vegetables	1.19 (1818)	1.01 (3389)	1.28 (120)	1.26 (3763)	NA	**1.43 (3853)**	1.09 (448)	1.23 (e)
Drinks with alcohol	**1.41 (26132)**	**0.86 (8267)**	1.00 (680)	**1.38 (40237)**	**1.27 (5145)**	**1.38 (3422)**	**1.41** (55841)	**1.36 (f)**
TOTAL	**0.86 (f)**	0.95 (e)	0.98 (d)	1.02 (c)	1.02 (c)	1.07 (b)	**1.14 (a)**	**1.01**

808 datasets and 761 083 standard deviations used. **Best** repeatabilities in **green** and **worst** in **red.** **Median** repetability is **0.71.**
Two means with the same letters in line or column margins are not significantly (HSD, p=0.05) different.

Optimizing sensory panels

- Which individual characteristics relate to performances?
- How many subjects should a panel include?
- Do we really need replicates?

Age effect on performance?

	Analyse 1					Analyse 2				
	Stat F	Pvalue	Modalités	Groupe	n	Stat F	Pvalue	Modalités	Groupes	n
Discrimination	2.26	0.105	-40	0.776	328	2.20	0.11	-40	0.792	70
			40-60	0.775	843			40-60	0.789	70
			+60	0.770	548			+60	0.781	70
Désaccord	1.45	0.23	-40	0.391	328	**4.93**	**0.008**	**-40**	**0.472 a**	**70**
			40-60	0.427	843			**40-60**	**0.442 b**	**70**
			+60	0.463	548			**+60**	**0.438 b**	**70**
Scaling	**20.31**	**<0.001**	**-40**	**-0.0001a**	**328**	**12.22**	**<0.001**	**-40**	**0.023 a**	**70**
			40-60	**0.078 a**	**843**			**40-60**	**0.06 a**	**70**
			+60	**-0.146 b**	**548**			**+60**	**-0.311 b**	**70**
Repetabilité	**21.26**	**<0.001**	**-40**	**1.253 a**	**328**	**18.8**	**<0.001**	**-40**	**1.009 a**	**70**
			40-60	**1.079 b**	**843**			**40-60**	**1.034 a**	**70**
			+60	**0.917 c**	**548**			**+60**	**0.840 b**	**70**
Niveau	1.8	0.165	-40	-0.089	328	0.52	0.99	-40	-0.10	70
			40-60	0.028	843			40-60	0.051	70
			+60	0.002	548			+60	0.002	70

Elderly people have a smaller scaling, resulting in a better repeatability, though no better discrimination. The two methods disagree on the « Disagreement » performance only.

Caroline Peltier Thesis (2015)

How many subjects should a descriptive panel include?

- Sampling of 100 datasets of size k from each dataset of size n (1<k<n)
- Comparison of sampled and complete datasets by:
  - ✓ Correlation coefficients between the two product means scores per attribute
  - ✓ RV coefficient between the two product configurations
  - ✓ Product F-values per attribute and MANOVA F-statistics

TAB. 4.2 – *Recommandations par type de descripteur*

Type d'attributs	Nombre d'attributs	Nombre de jeux de données	Nombre de sujets (N)		Recommandation (n)	
			Étendue	Moyenne- IC*	Étendue	Moyenne- IC*
Arôme	1933	284	6-32	12.71 [12.31; 13.10]	2-28	9.61 [9.21; 10.01]
Odeur	1481	162	4-24	12 [11.48; 12.52]	2-19	8.37 [7.81; 8.92]
Saveur	2082	286	4-27	12.40 [12.00; 12.80]	2-26	8.39 [7.97; 8.80]
Texture	1234	241	6-26	12.69 [12.28; 13.10]	2-21	7.75 [7.30; 8.20]
Visuel	708	177	6-26	12.77 [12.25; 13.30]	2-25	6.36 [5.82; 6.91]

(*) Intervalle de confiance à 95%

Substantial reduction of panelist numbers suggested!

Nadra Mammasse Thesis (2012)

Do we need to replicate in sensory profiling studies?

Using the first replicate only instead of two replicates resulted in (n=377 datasets):

- 60% of significant (p=0.10) attributes instead of 67%
- Average correlation coefficient between product mean scores of 0.94 (median 0.98, n=5467)
- Multivariate ratio of Hotteling-Lawley F-statistics is 2.02 (median of 1.21)
- Average RV value between the 1-2 CVA maps of 0.90
- Percentage of similar Hotteling test diagnosis 90%
- Similar interpretation of CVA biplots (average MaxAngle of 40°)
- 5% of product pairs no longer discriminated

The second replicate does not bring much more information !

Nadra Mammasse Thesis (2012) – Caroline Peltier Thesis (2015)

Comparing Sensometric techniques

TWO EXAMPLES:

- The Mixed Assessor Model (MAM) versus the usual mixed model of ANOVA
- The Canonical variate Analysis (CVA) versus the usual PCA

Caroline Peltier Thesis (2015)

The Mixed Assessor Model (MAM)

Brockhoff, Schlich & Skovgaard (2015)

The Mixed model

$$Y_{ijk} = \mu + \alpha_i + \gamma_j + c_{ij} + \varepsilon_{ijk}$$

α_i: subjet effect; γ_j: product effect;

c_{ij} : interaction

$$F = \frac{MS_{product}}{MS_{interaction}}$$

The Mixed Assessor Model

$$Y_{ijk} = \mu + \alpha_i + \gamma_j + \beta_i x_j + d_{ij} + \varepsilon_{ijk}$$

α_i: subject effect γ_j: product effect x_j: $(y_{.j.} - y_{...})$

β_i: scaling coefficient d_{ij}: pure disagreement

$$F = \frac{MS_{product}}{MS_{pure\ disagreement}}$$

Application of MAM to 236 datasets from the SensoBase

having at least 3 products and 2 replicates and being balanced

Scaling heterogeneity is definitely presents in our sensory data:

- **45 %** of the attributes exhibited a significant scaling effect
- **23 %** of the individual scaling coefficients were tested different than 1
- **92 %** of the panelists scaled at least one attribute differently than the group

Product by panelist interaction is a pessimistic view of disagreement:

- **29 %** of significant product by panelist interaction with MAM instead of **48%**
- **40 %** of the usual interaction significances were just due to scaling effect

MAM increases power moderately:

- **64 %** of significant product effect in MAM, compared to **58 %** in ANOVA
- **19 %** of the non significant attributes in ANOVA become significant in MAM, compared to **4 %** the other way round
- **10 %** of attributes with a different product diagnostic (in average **2.4** attribute per dataset)

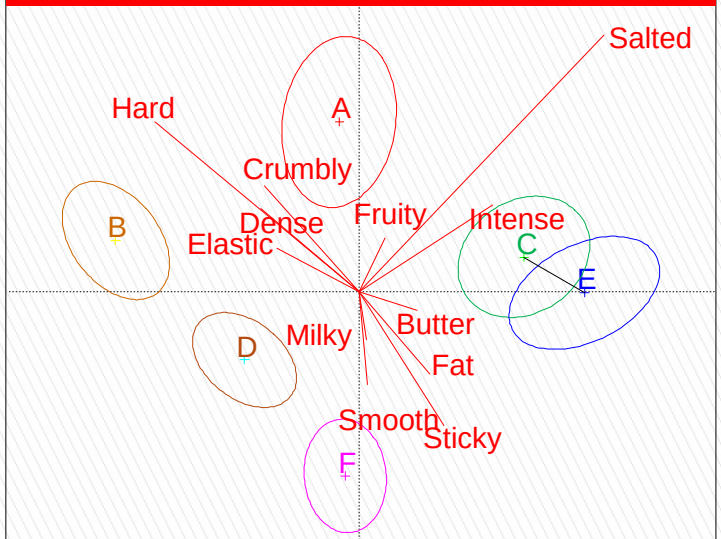
A meta-analysis of 379 datasets to compare PCA to CVA

Comparison criteria	PCA vs CVA	PCA vs CVA with no subject effect	PCA vs overall MAM-CVA	PCA vs multivariate MAM-CVA
RV coef	0.91	0.92	0.95	0.92
Max Angle	33	26	23	25
Product pairs discriminated in PCA but not in CVA	5%	0.6%	0.3%	0.3%
Product pairs discriminated in CVA but not in PCA	4%	17%	18%	19%

Accounting for individual differences in both level (subject effect) and range (scaling effect) of scores results in an improved multivariate product discrimination

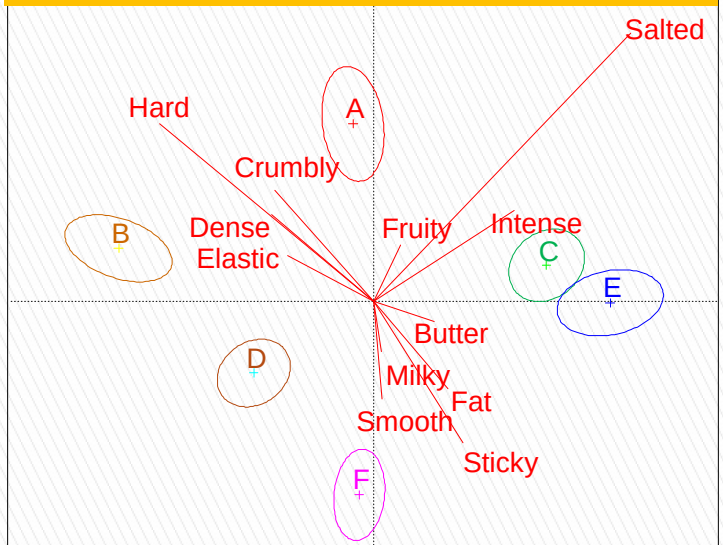
Caroline Peltier Thesis (2015)

CVA



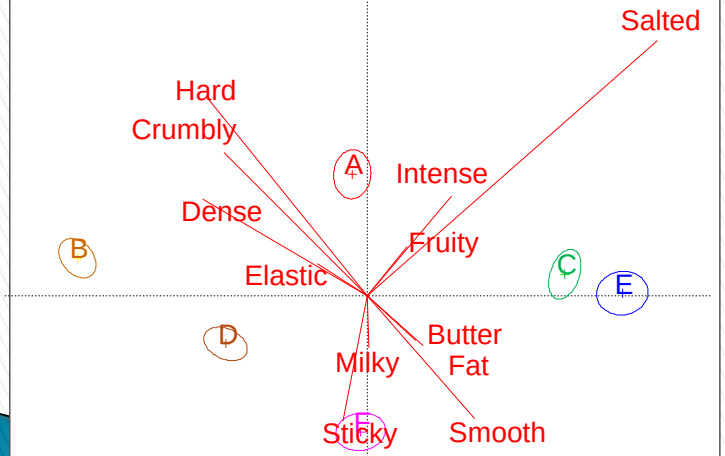
Stat= 7.808, F=8.149, p-value<0.001

CVA without subject effect



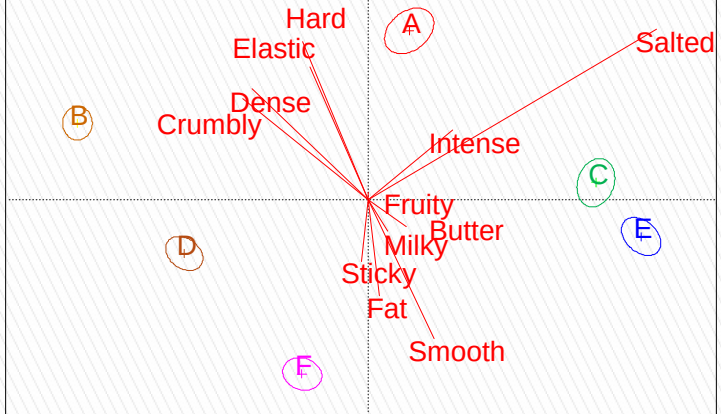
Stat= 7.808, F=8.149, p-value<0.001

MAM-CVA overall



Stat= 28.159, F=22.353, p-value<0.001

MAM-CVA multivariate



Stat= 43.434, F=35.382, p-value<0.001

Database and Meta-analysis : method, limits & cautions

DATABASE

- Is my base really representative of the field under interest?
- How to be sure that a dataset does not contain fictitious data or training data?
- Make impossible to enter the same data several times

META-ANALYSIS

1. State precisely the objective of the meta-analysis
2. Define the criteria to be computed
3. Selection the datasets suitable to these criteria
4. Validate the dataset selection (representativity, no outlier, …)
5. Run the computation
6. Possibly return to 2. due to errors in computation

Conclusion

- Sensobase: a unique database in the sensory field worldwide
- Benchmarking panel performances is highly expected by panel leaders
- Our recomendations about number of replicates and panelists can save a lot of money to companies
- A rational framework under which power computations can be done
- A reliable investigation of real impacts of Sensometrics progresses

Message to the major companies producing numerous sensory data routinely:

Pool all of your data and gain knowledge by analyzing them together!

References

Pineau, N. (2006), La performance en profil sensoriel : une approche base de données. Ph.D. thesis defended on December 13, 2006 at the University of Burgundy, France.

Pineau N. Chabanet C., Schlich P. (2007). Modeling the evolution of a sensory panel: a mixed-model and control chart approach. Journal of Sensory Studies, 22, 212-241.

Mammasse, N. (2012), Le nombre de sujets dans les panels d'analyse sensorielle: une approche base de données. Ph.D. thesis defended on March 22, 2012 at the University of Burgundy, France.

Mammasse N, Schlich P (2014). Adequate number of consumers in a liking test. Insights from resampling in seven studies. Food Quality and Preference 31, 124-128.

Peltier C., Visalli M., Brockhoff P.B., Schlich P. (2014). The MAM-CAP table: A new tool for monitoring panel performances. Food Quality and Preference 32, 24-27.

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

Peltier, C. (2015). L'analyse du profil sensoriel revisitée par une approche base de données. Ph.D. thesis defended on September 25, 2015 at the University of Burgundy, France.

Peltier C., Visalli M., Schlich P. (2015). Canonical Variate Analysis of Sensory Profiling Data. Journal of Sensory Studies 30 (2015) 316–328

Peltier C., Visalli M., Schlich P. (2015). Comparison of Canonical Variate Analysis and Principal Component Analysis applied to data from 422 descriptive sensory studies. Food Quality and Preference, 40, 326-333.

Peltier C., Visalli M., Schlich P. (2016). Multiplicative decomposition of the scaling effect in the Mixed Assessor Model into a descriptor-specific and an overall coefficients. Food Quality and Preference, 48, 268-273.

Conference ads

Brighton. July 26-29, 2016

Dijon. September 11-14, 2016