Identification and prediction of robust non linear relationships between recipes, instrumental measurements and sensory evaluation of cake texture and flavour with xtractis® fuzzy models. Application to the discovery of optimal recipes.

Z. Zalila [1,7], E. Mehinagic [2], J. Cuquemelle [1], A.-V. Profizi [3], F. Texier [4], S. Toulgoat [5], A. Saint-Eve [6], B. Lorentz [1]

Data collection: 21 cake formulations

Objective data

Recipe

- 11 variables
- Storage conditions

Instrumental measurements

- Cake density
- Cake color (measured 3 times)
- Water activity (measured twice)
- Water content (measured twice)
- Double compression (measured 8 times)

Data pre-processing

Intra filtering

- 15% of Definition Range
➔ 2% of data discarded

Intra filtering

- 38% of Definition Range
➔ 12.3% of data discarded

Inter filtering

- 45% of Definition Range
➔ 17.4% of data discarded

Sensory data

Attributes

- Sensory texture attributes
  - between fingers: **springiness, fracturability, fatness**
  - in mouth: **melt in mouth, adhesiveness, granularity, dryness**
- Flavour attributes: **sweet taste, flavour**

➔ 12 assessors, 9 points scale, 3 sessions

Variable selection

Exclusive use of original variables (no PCA)

- Maintains interpretability: semantic validation of models by experts

Rank	Var. ID	Label	Contribution	Missing value
1	2	% wheat flour	1.000	0.0%
2	10	% vanilla flav.	0.383	0.0%
3	1	% sugar	0.377	0.0%
4	3	% sunflower oil	0.343	0.0%
5	8	% E471 Emulsifier	0.341	0.0%

Rank	Var. ID	Label	Contribution	Missing value
1	2	% wheat flour	1.000	0.0%
2	8	% E471 Emulsifier	0.792	0.0%
3	1	% sugar	0.562	0.0%
4	3	% sunflower oil	0.455	0.0%
5	5	% sodium bicarbonate	0.336	0.0%

Experts pointed out a statistical artefact: **vanilla flavouring** *has no influence on* **springiness**
➔ *a second robust model with* **% sodium bicarbonate** *was validated instead*

- Allows research of exact values of optimal recipes
- Latent variables:
  - are data driven and may be prone to overfitting
  - may hide non linear relationships between variables

Learning Base

Fuzzy modelling

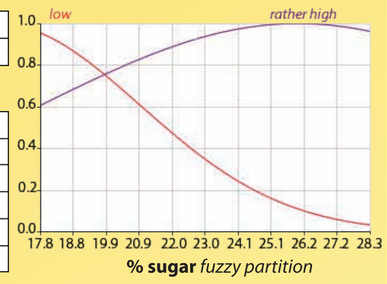
- Fuzzy model: collection of linguistic rules

Rule 1:

If	% sugar	is	*rather high*
Then	sweet taste	is	*6.99*

Rule 2:

If	% sugar	is	*low*
and	% wheat flour	is	*high*
and	% vanilla flav.	is	*low*
and	% sugar flav.	is	*low*
and	storage cond.	is	*Ambiant T°*
Then	sweet taste	is	*2.46*



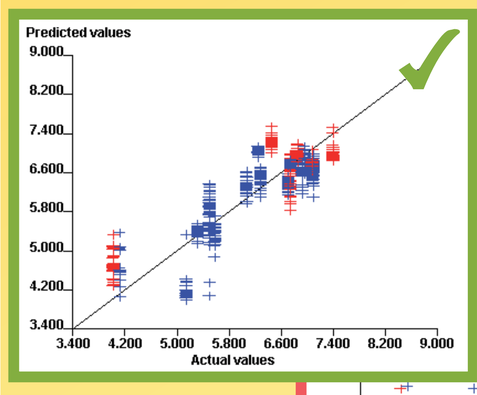
% sugar *fuzzy partition*

sweet taste *can be explained by 5 variables combined into 2 fuzzy rules*

- Easy to interpret when low number of explanatory variables or rules
- Interface between qualitative and quantitative data
- Universal approximator of multidimensional non linear functions [Kosko, 1992]
- x**tractis**®: extraction of robust fuzzy models

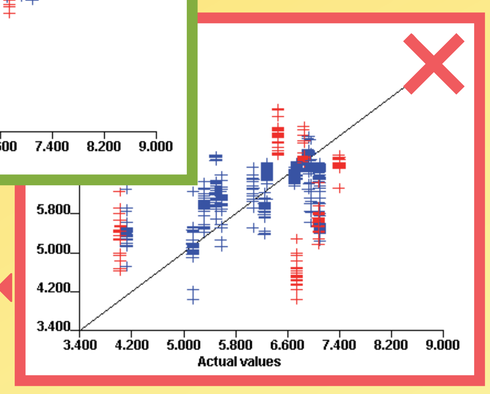
Generalisation capacity

- Hold-Out-Set method: estimator with high variance
- Cross Validation method: best estimator [Zalila, Cuquemelle & al., 2008]
  ➔ Monte Carlo (MC): 15% of data in each split, 150 splits



A robust **sweet taste** *model, with* **storage condition** *variable:*
- *MC correlation =* ***0.852***
- *MC Hamming's error =* ***5.07%***

Without **storage condition**, **sweet taste** *cannot be explained:*
- *MC correlation =* ***0.413***
- *MC Hamming's error =* ***8.87%***

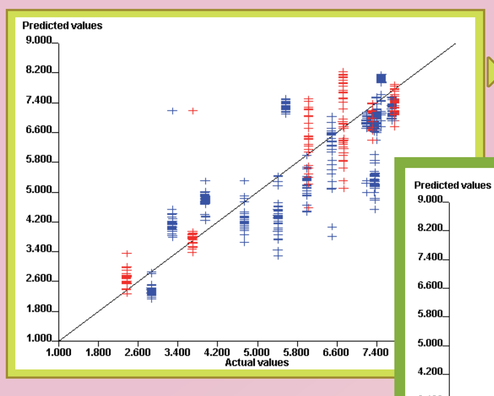


sweet taste *robusness evaluation*

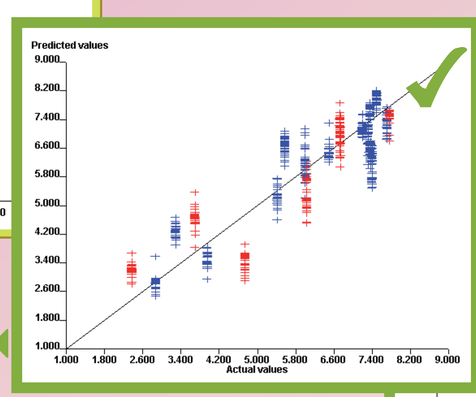
Blue crosses represent predictions. Red crosses represent predictions of points lying on boundaries of LB input/output range.

Modelling of cake's sensory profile: melt in mouth

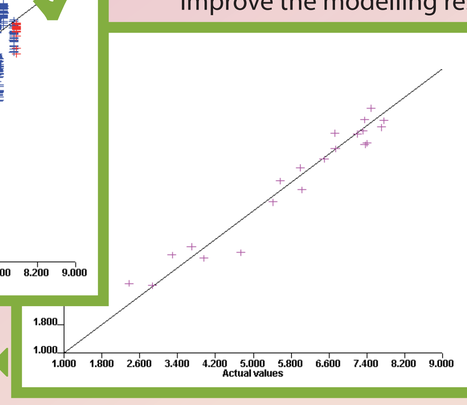
Influence of instrumental measurements



Recipe only:
- ***6*** *variables,* ***3*** *rules*
- *MC correlation =* ***0.855***
- *MC Hamming's error =* ***9.04%***



Recipe and instrumental measurements:
- ***5*** *variables,* ***2*** *rules*
- *MC correlation =* ***0.898***
- *MC Hamming's error =* ***7.22%***



- *Accuracy correlation =* ***0.975***
- *Accuracy Hamming's error =* ***3.86%***

- Robust model from recipe variables
- Less robust model from instrumental variables
- Recipe + instrumental (*e.g.* **Grad AB**):
  - MC robustness increased
  - complexity reduced

➔ Instrumental measurements can improve the modelling results

Non linear modelling

- Predictive rules validated by experts
- Non monotonic relationship: variable influence not only positive/negative

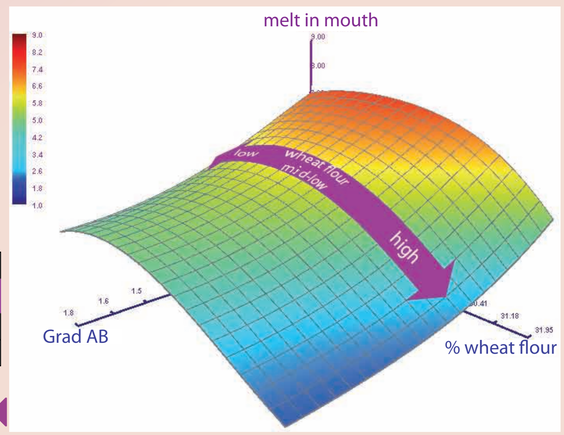
Rule 1:

If	% sugar	is	*high*
and	% vanilla flav.	is	*low*
and	Grad AB	is	*high*
Then	melt in mouth	is	*1.99*

Rule 2:

If	% sugar	is	*low*
and	% wheat flour	is	*med-low*
and	% butter flav.	is	*low*
Then	melt in mouth	is	*8.61*

Cross-section of 5D-decision surface:
% sugar *is 0.85-high / 0.65-low,* **% butter flav.** *is 1-low,* **% vanilla flav.** *is 0.88-low*



Relations between sensory descriptors

- **melt in mouth** model 1

Rule 1:

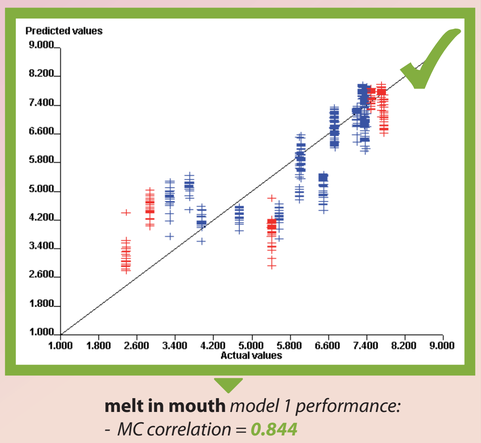
If	fatness	is	*low*
and	dryness	is	*very high*
Then	melt in mouth	is	*1.51*

Rule 2:

If	fatness	is	*high*
and	dryness	is	*low*
and	springiness	is	*high*
Then	melt in mouth	is	*8.43*

- Predictive rules approved by experts
- Accurate:
  - Accuracy correlation = 0.931
  - Accuracy Hamming's error = 6.31%

- **melt in mouth** model 2
  - 4 variables, 2 rules
  - Accuracy correlation = 0.982, Accuracy Hamming's error = 3.29%
  - MC correlation = 0.860, MC Hamming's error = 7.71%



melt in mouth *model 1 performance:*
- *MC correlation =* ***0.844***
- *MC Hamming's error =* ***8.34%***

Primary findings

- Evaluation of predictive model performance: Accuracy and Hold-Out Set are bad estimators, Cross Validation is mandatory to estimate the generalisation capacity.
- Sensory profiles cannot be modelled by using only compression data.
- Sensory profiles can be modelled by using only recipe data, but adding instrumental measurements may improve models.
- x**tractis**® OPTIMIZE allows discovery of optimal recipes thanks to models using recipe variables only.
- Sensory descriptors are not independent.
- Variables influence may be non linear and non monotonic.
- Extracted fuzzy rules are interpreted and validated by the interdisciplinary group of scientists.

8th Pangborn Sensory Science Symposium 2009

[1] **intelli**tech, France
[2] ESA, France
[3] NutriXo-Grands Moulins de Paris, France
[4] ENITIAA, France
[5] ISIPCA, France
[6] AgroParisTech-INRA, France
[7] University of Technology of Compiègne, France