Regularization approaches in sensory descriptive analysis

Applications to Kokumi-product development

Eduard Derks, Kae Morita, Tetsuo Aishima

Sensometrics
Juli 12th, 2012

DSM
BRIGHT SCIENCE. BRIGHTER LIVING.

NPD-study コク味

- **Goal of project**
    1. **Understand Japanese concept**
    2. **Identify koku-drivers in savoury/cullinary applications**
    3. **Support NPD to developed a koku-enhancing compound**

- **Sub-goal (method development)**
  Develop strategies to
    - **link koku to sensory data (difference tests, QDA)**
    - **deal with large datasets**
   - **improve interpretability**

Kokumi

Koku = コク ＝ 濃香 ＝ thickness, complexity, depth and/or width
mi   = 味 ＝ 味道 ＝ taste

Project steps

- **Evaluation in expert teams and QDA-panels (Japanese & Dutch)**

- **Interviews with Japanese cooks (n=80)**

- **Interviews with APAC consumers (Japan, China, Korea, n=600)**

- **Meat stew sensomics (NMR, GPC, LCMS, GCMS, QDA, etc.)**

- **YE-based reconstruction of koku-targeted flavour & mouthfeel**

- **Consumer tests: testing NPD samples against competition**

- **Lots of data – difficult to interpret – need for sparse methods both for regression and explorative analysis**

Data & models

Explain kokumi in terms of
sensory attributes (flavour & mouthfeel)

K

1) ANOVA
2) GLM with psychometric
link function

1) Penalized OLS
(regularized regression)

2) Penalized GLM with
psychometric link function

DOE

QDA

ANOVA
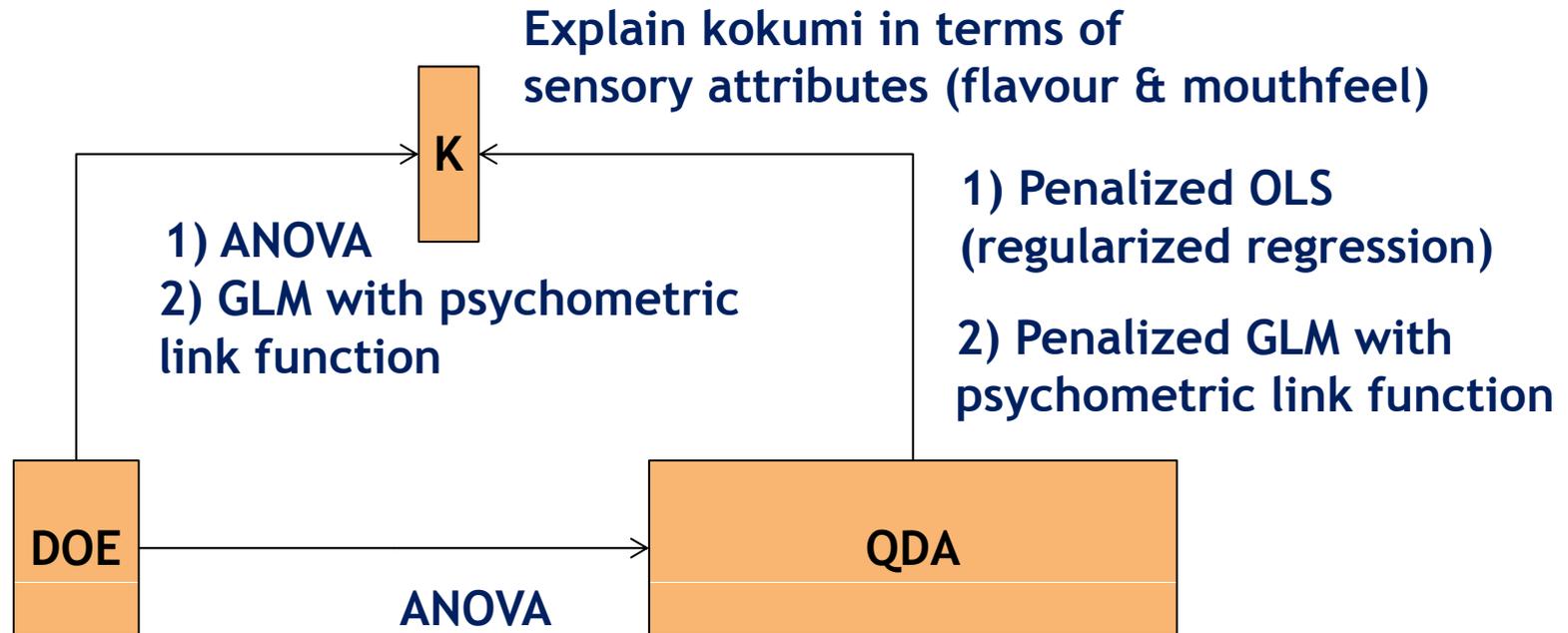
Alternatives
1) PLS/RR
- greedy
- developed for smooth spectra
2) GLM with psycho-lf. on PCA scores
- more difficult to interpret

Regularization
in multivariate regression

- **Multivariate regression**

$$y = \mathbf{X}\beta + \varepsilon$$

- **OLS minimizes**

$$\left\| y - \mathbf{X}\beta \right\|^2$$

- **Solution**

$$\beta = \left(\mathbf{X}^t \mathbf{X}\right)^{-1} \mathbf{X}^t y$$

- **Inversion problems for collinear X**

- **Solution:** restrict beta's

$$\left\| y - \mathbf{X}\beta \right\|^2 + \lambda \left\| \beta \right\|^n$$
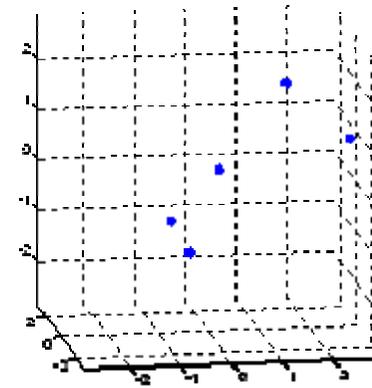

L2, L1 and L0 penalty function
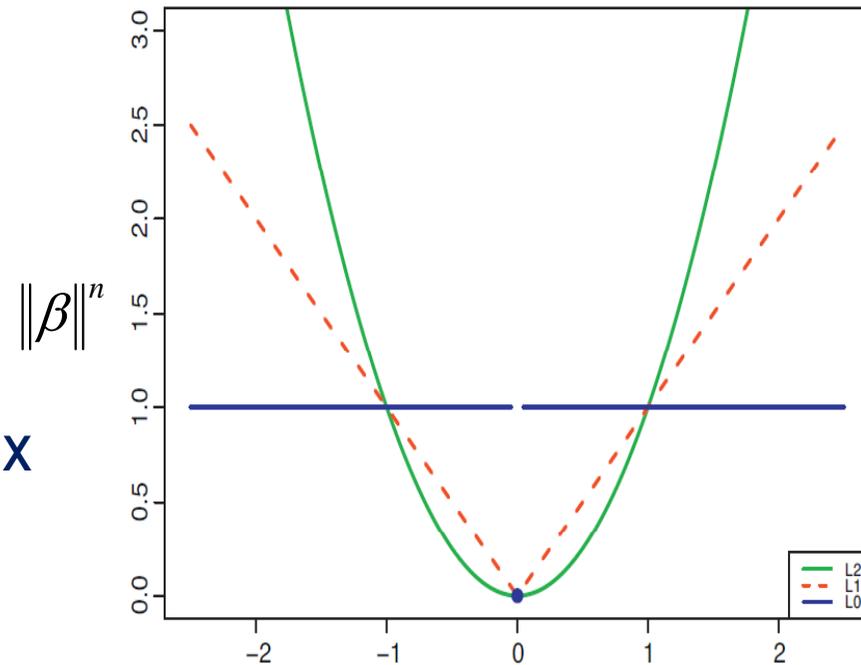
Regularization

in multivariate regression

**Heavy shrinkage
brings variable selection
(Sparse OLS, Sparse PCA, etc.)**

$$Q_{L1} = \left\| y - \mathbf{X}\beta \right\|^2 + \lambda \sum_k \left| \beta_k \right|$$

**Minor shrinkage reduces variance
at the cost of a bias (RR,PLS, PCR)**

$$Q_{RR} = \left\| y - \mathbf{X}\beta \right\|^2 + \lambda \left\| \beta \right\|^2$$
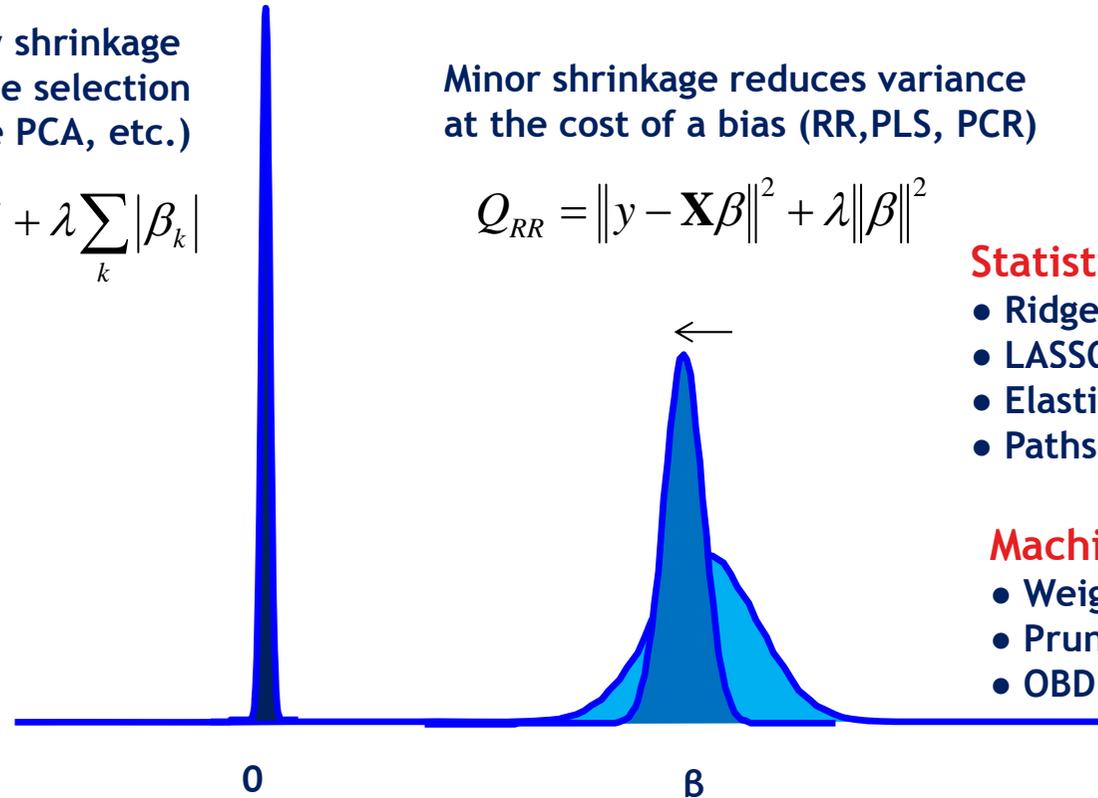
Statistics
- Ridge regression
- LASSO/ LARS
- Elastic net
- Pathseeker

Machine learning
- Weight decay
- Pruning
- OBD

0

β

DSM
BRIGHT SCIENCE. BRIGHTER LIVING.

Regularization

in principal component analysis

L2, L1 and L0 penalty function

- **PCA**

$$\hat{\mathbf{X}} = tp^{\mathrm{t}}$$

$$Q_{PCA} = \left\| \mathbf{X} - tp^t \right\|^2$$

- **Issues:**

  messy biplots
  many variables (e.g. internal prefmapping, higher order QDA)
 many objects (e.g. consumer studies)

- **Sparse PCA**

  Shrink loadings (automatic attribute selection)

$$Q_{L2(p)} = \left\| \mathbf{X} - tp^t \right\|^2 + \lambda \| p \|^2$$

$$Q_{L1(p)} = \left\| \mathbf{X} - tp^t \right\|^2 + \lambda \sum_j \left| p_j \right|$$

$$Q_{L0(p)} = \left\| \mathbf{X} - tp^t \right\|^2 + \lambda \sum_j \left(\left| p_j \right| > 0 \right)$$

DSM
BRIGHT SCIENCE. BRIGHTER LIVING.
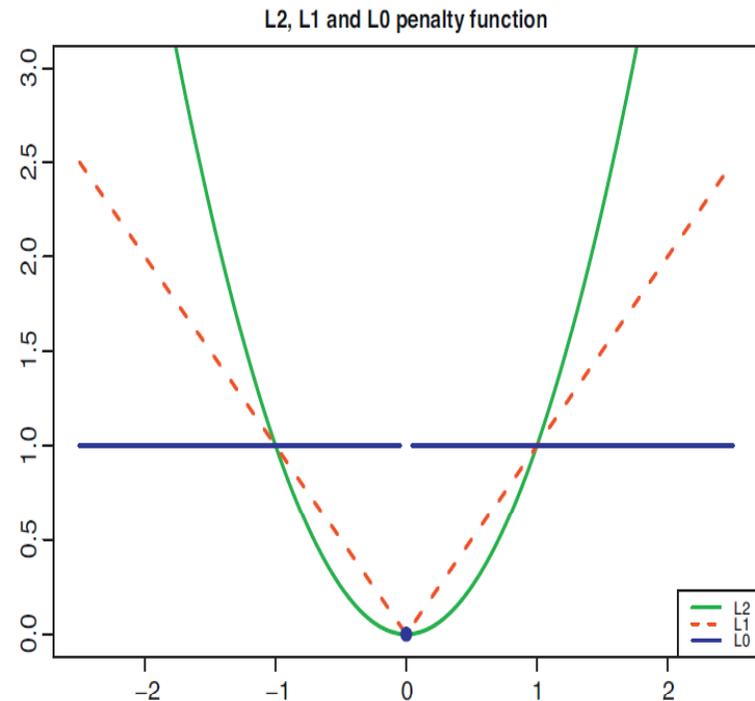
Regularization
in principal component analysis

Easier to interpret
- due to sparseness
- due to the fact that PC's get a meaning
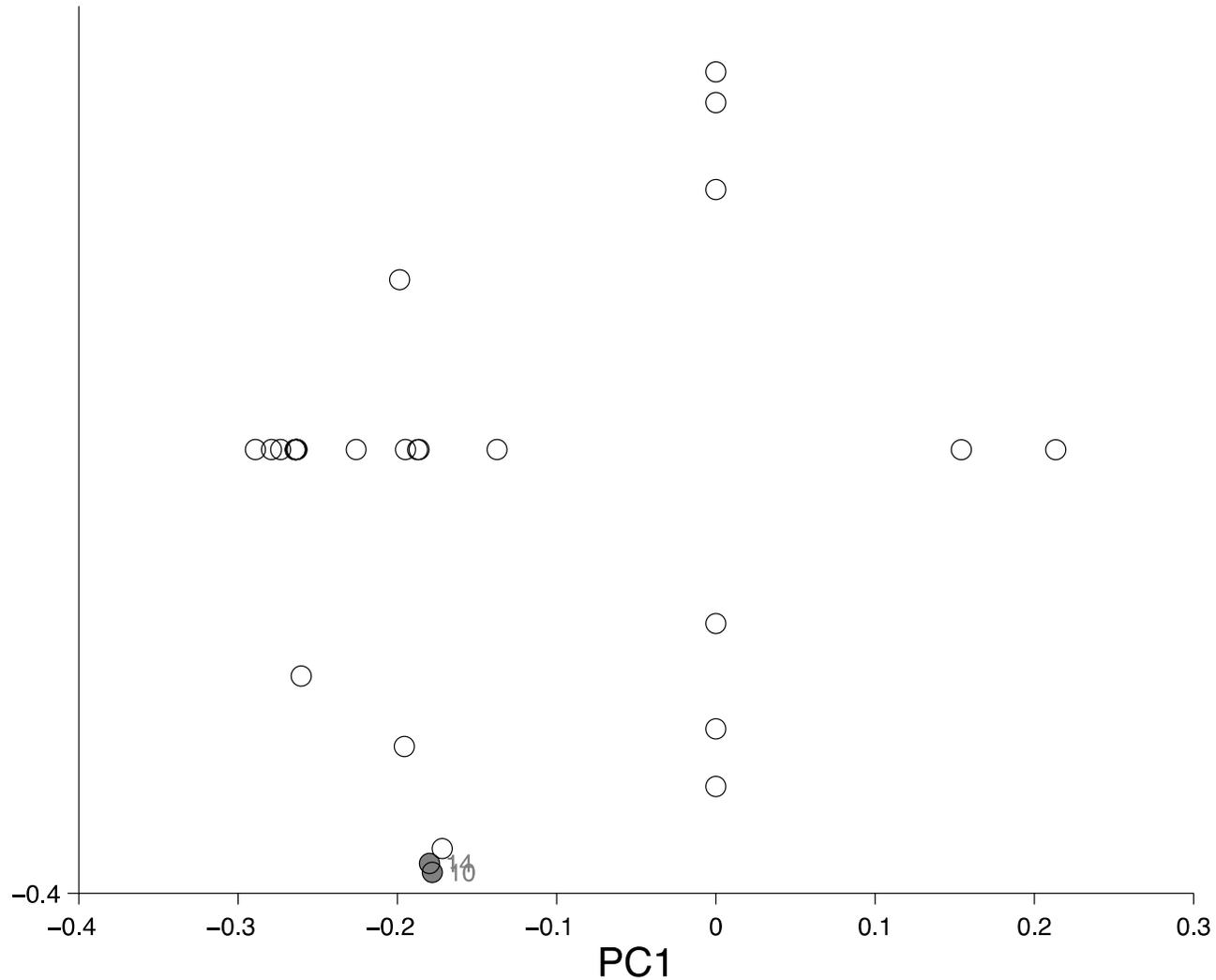
Regularization

in principal component analysis

- Can also be applied for sparse scores

- or penalization on scores and loadings simultaneously

This is called CoClustering*

$$Q_{coclustering} = \left\| \mathbf{X} - tp^t \right\|^2 + \lambda \sum_i \left| t_i \right| + \mu \sum_j \left| p_j \right|$$

- method of choice when specific variables are related
by specific groups of objects
  - contingency (COUNT) tables
  - CATA biplots (consumers x checked attributes)
  - sparse scoring data

*Bro, R., Papalexakis, E. E., Acar, E. and Sidiropoulos, N. D. (2012), Coclustering—a useful tool for chemometrics. J. Chemometrics, 26: 256-263

DSM
BRIGHT SCIENCE. BRIGHTER LIVING.

	Has eyes	Number of	Carnivor	Feather	Wings	Domesticia	Eaten by (	>100kg	>2m	Breathe un	Extinct	Dangerous	Life expect
Giraffe	1	4	0	0	0	0	0	1	1	0	0	0	30
Cow	1	4	0	0	0	1	1	1	1	0	0	0	15
Lion	1	4	1	0	0	0	0	1	0	0	0	1	15
Gorilla	1	4	0	0	0	0	0	1	0	0	0	1	30
Fly	1										0	0	0.1
Spider	1										0	0	1
Shark	1										0	1	50
House	0										0	0	100
Horse	1										0	0	15
Elephant	1										0	0	35
Mammoth	1										1	0	35
Sabre Tige	1										1	1	15
Pig	1										0	0	25
Cod	1										0	0	40
Eel	1										0	0	55
Jellyfish	1										0	0	0.7
Dolphin	1										0	0	30
Nemo	1										0	0	1
Shrimp	1										0	0	1
Dog	1										0	0	13
Cat	1										0	0	25
Fox	1										0	0	14
Wolf	1										0	1	18
Rabbit	1										0	0	9
Chicken	1										0	0	15
Eagle	1										0	0	55
Seagull	1										0	0	10
Blackbird	1	2	1	1	1	0	0	0	0	0	0	0	18
Bat	1	2	1	0	1	0	0	0	0	0	0	0	24
T. Rex.	1	4	1	0	0	0	0	1	1	0	1	1	40
Neandertha	1	4	1	0	0	0	0	0	0	0	1	0	50
Triceratops	1	4	1	0	0	0	0	1	1	0	1	1	30

Cluster 1
Penguin

Blackbird Walk on two legs

Seagull Has a beak

Eagle Wings

Chicken Feather

Cluster 2
Triceratops

Neanderthal

T. Rex.

Sabre Tiger

Mammoth Extinct

C
Tricera

T. Rex.

Dolph

Mamm

Eleph

Horse

House

Cow

Giraff

Cluster 4
Chicken
Rabbit

Cluster 5
Shrimp

C
Penguin
Man
Tricerato

Now some applications ...

LASSO

koku versus QDA

Cross-validated MSE of Lasso fit

Trace Plot of coefficients fit by Lasso

K

QDA

DSM
BRIGHT SCIENCE. BRIGHTER LIVING.

LASSO

LASSO
koku versus QDA

Cross-validated MSE of Lasso fit

UNDERFIT   OVERFIT

K

QDA   +interactions

DSM
BRIGHT SCIENCE. BRIGHTER LIVING.
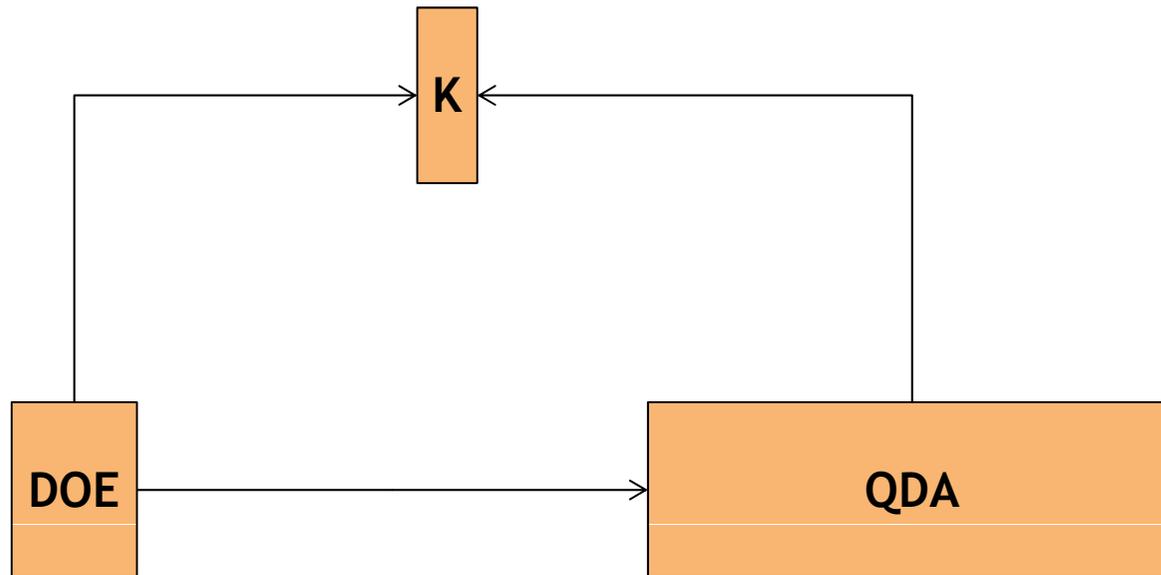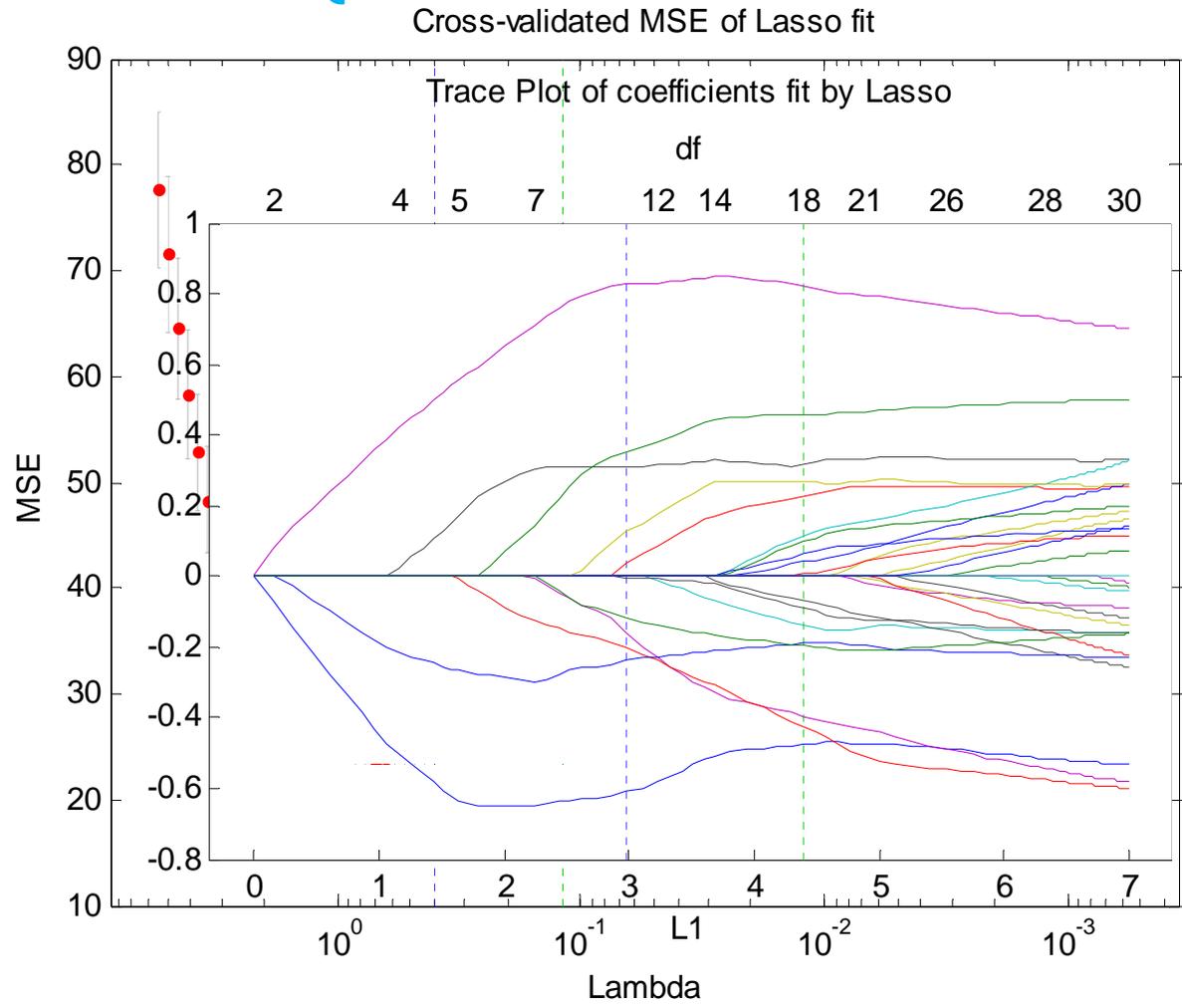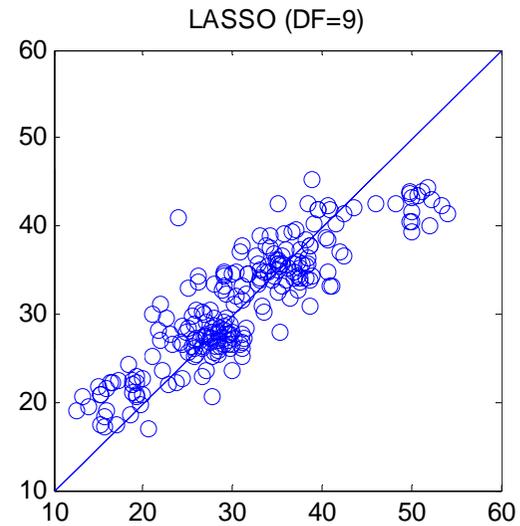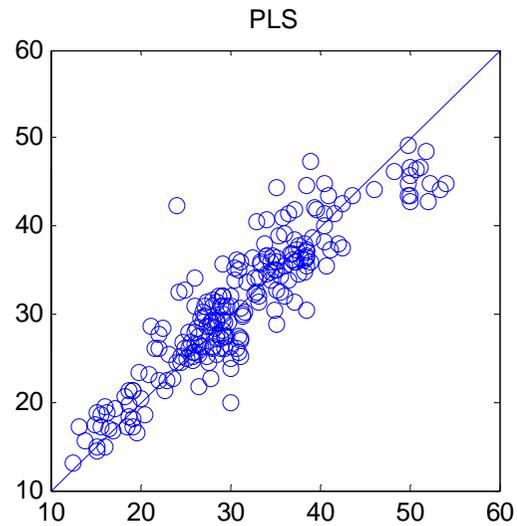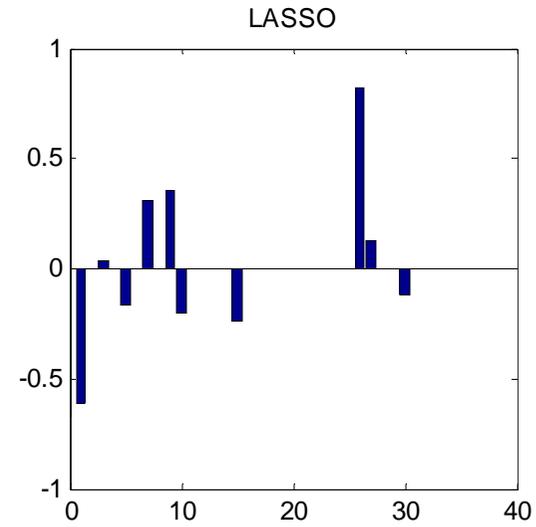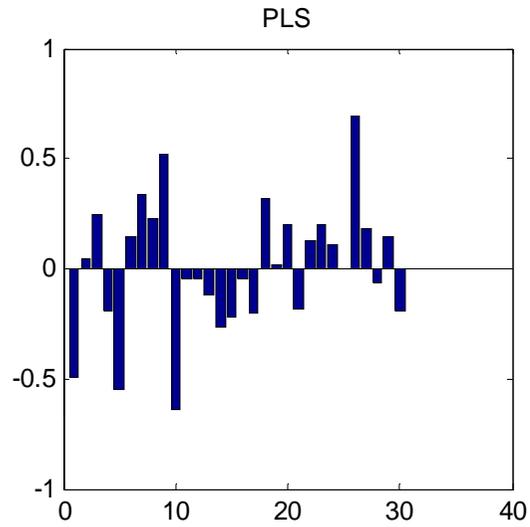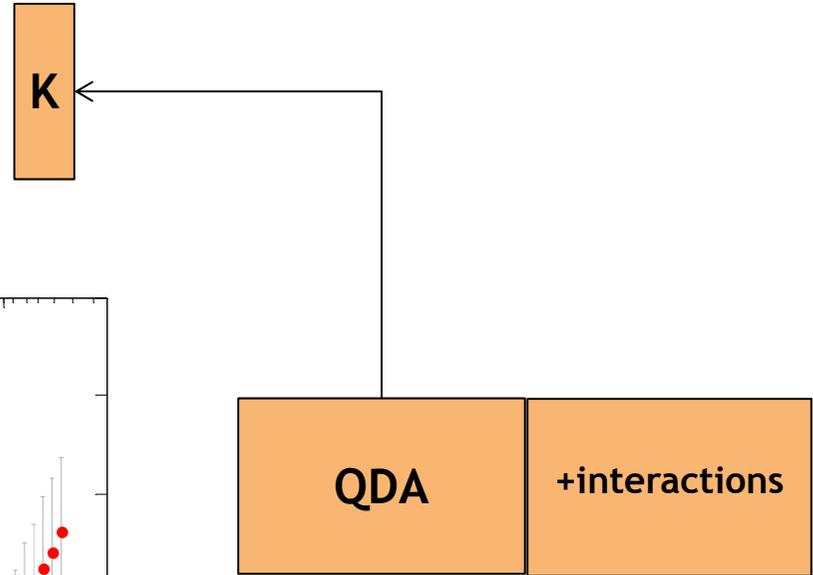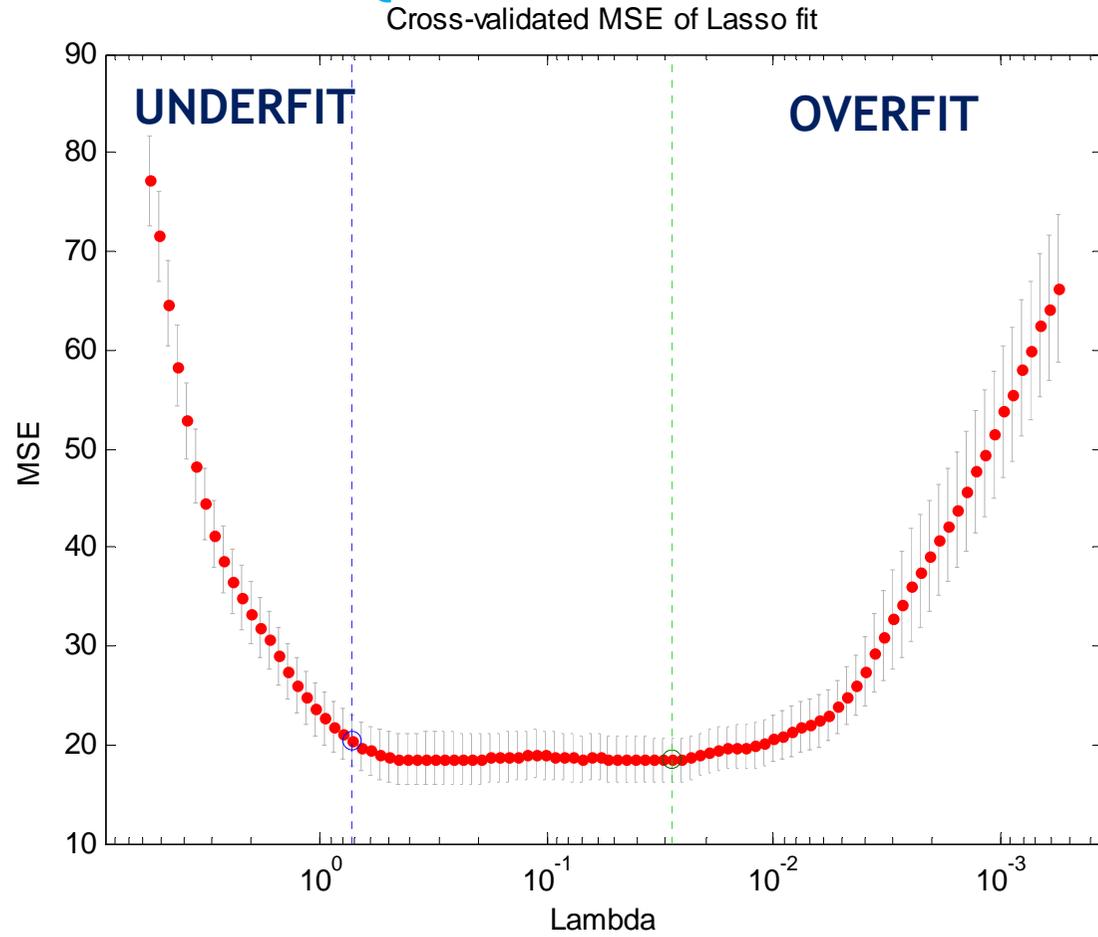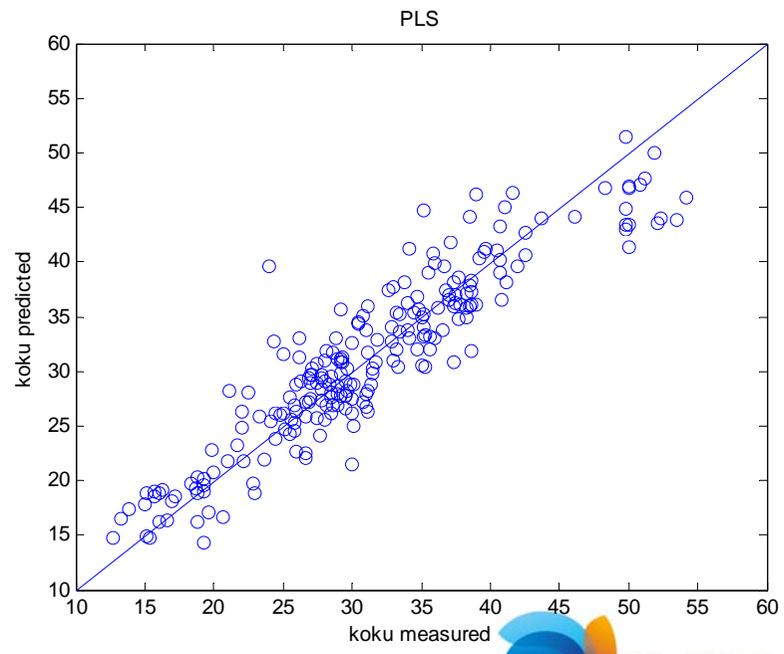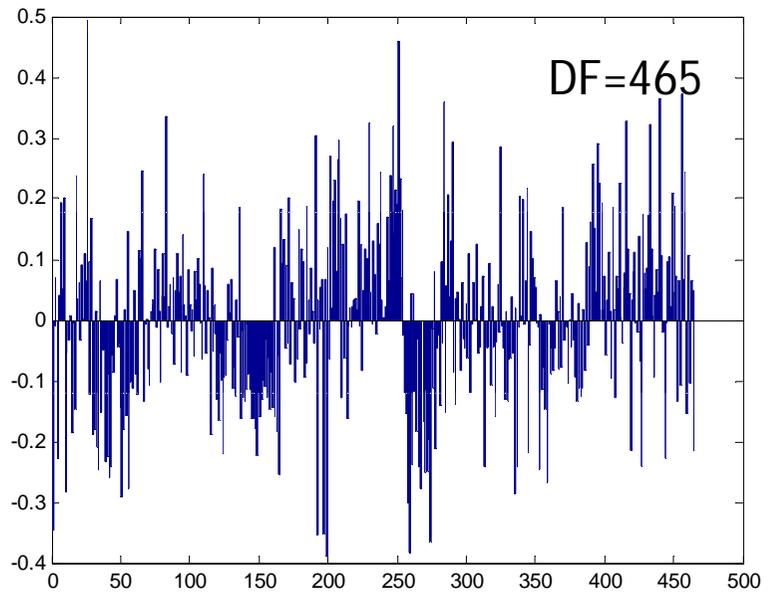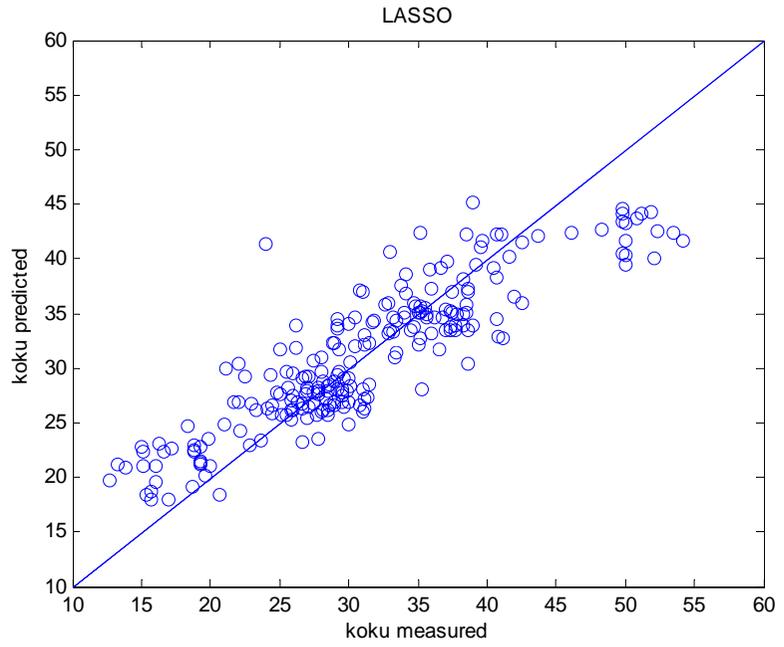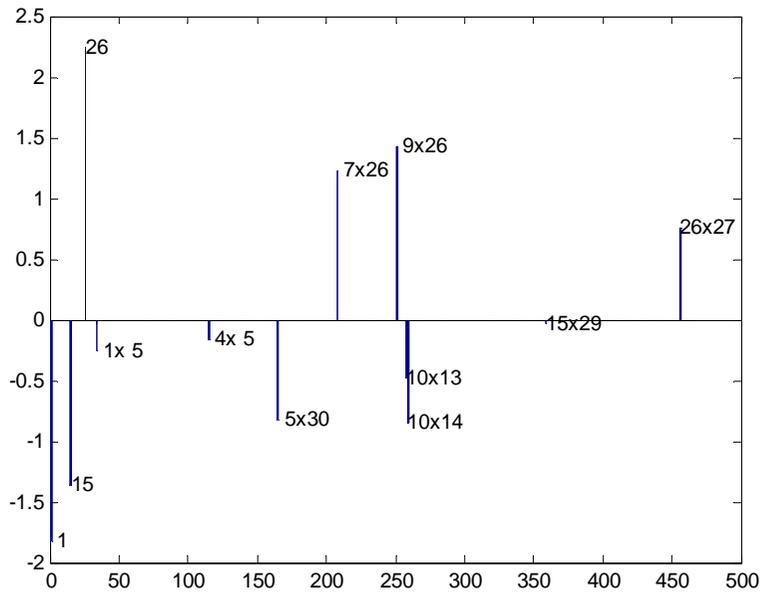
First conclusions

- **LASSO selects a limited set of sensory attributes and interaction terms to explain kokumi with similar predictability compared to PLS**

- **But, LASSO improves interpretation (variable selection)**

- **PLS overfits (not demonstrated here)**

- **PLS (and RR) do not handle mega/multivariate sensory data well**
    - **eager to overfit**
   - **seeks correlated coefficients***

* Hastie, T., Tibshirani, R., Friedman, J. 'The elements of statistical learning'. Springer series in Statistics. Page 83
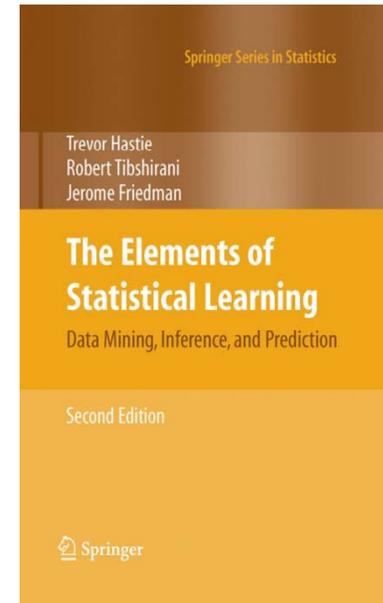
LASSO-GLM

## MSG dose-response analysis

U

QDA

- Purpose: model umami perception from an EXPERT team as a function of QDA attributes with trained panelists

- EXPERT team (n=6) used 3AFC (2reps)

- QDA panel (n=14) (no umami in attribute list)

- MSG series (0, 0.12, 0.24, 0.36, 0.48, 0.60 percent)

- Applied in beef bouillon

- Proportion correct odds were modeled with GLM as function of MSG dosage and QDA attributes

- GLM model was organized with a 3AFC-psychometric link function and L1 penalty

DSM
BRIGHT SCIENCE. BRIGHTER LIVING.
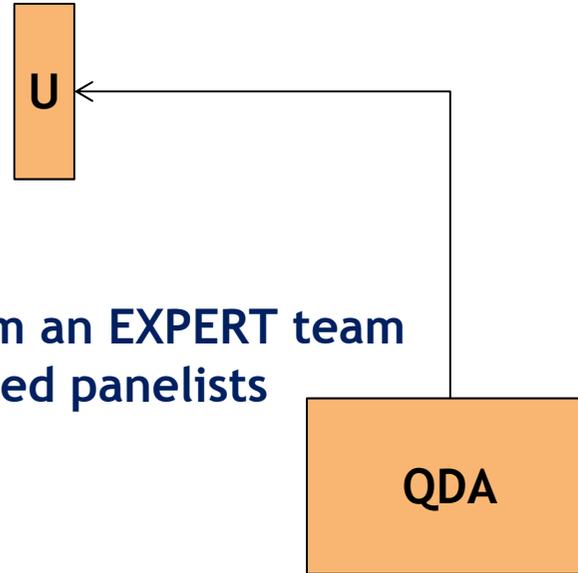
What about the panel?
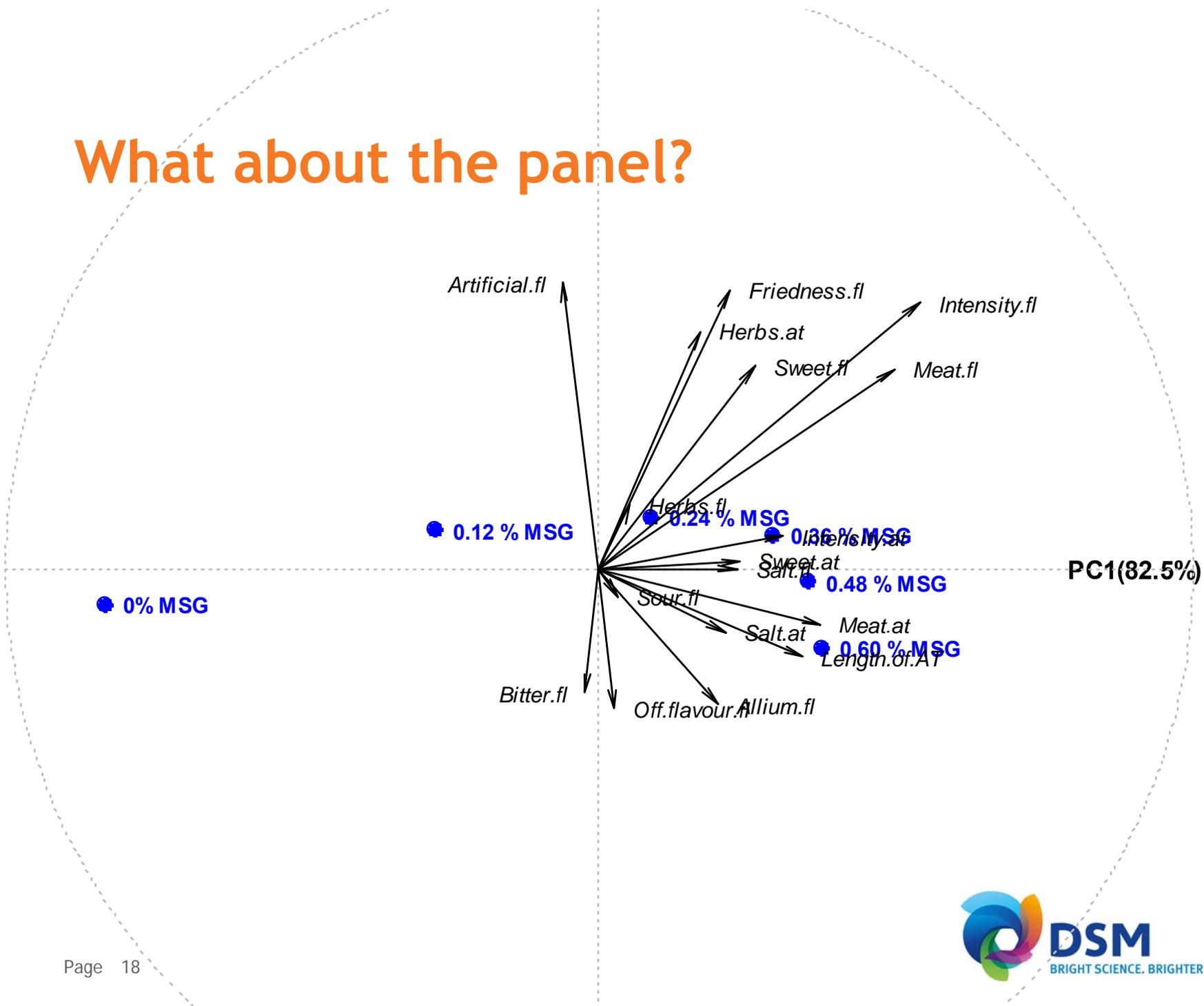
What about the experts?

GLM-fit
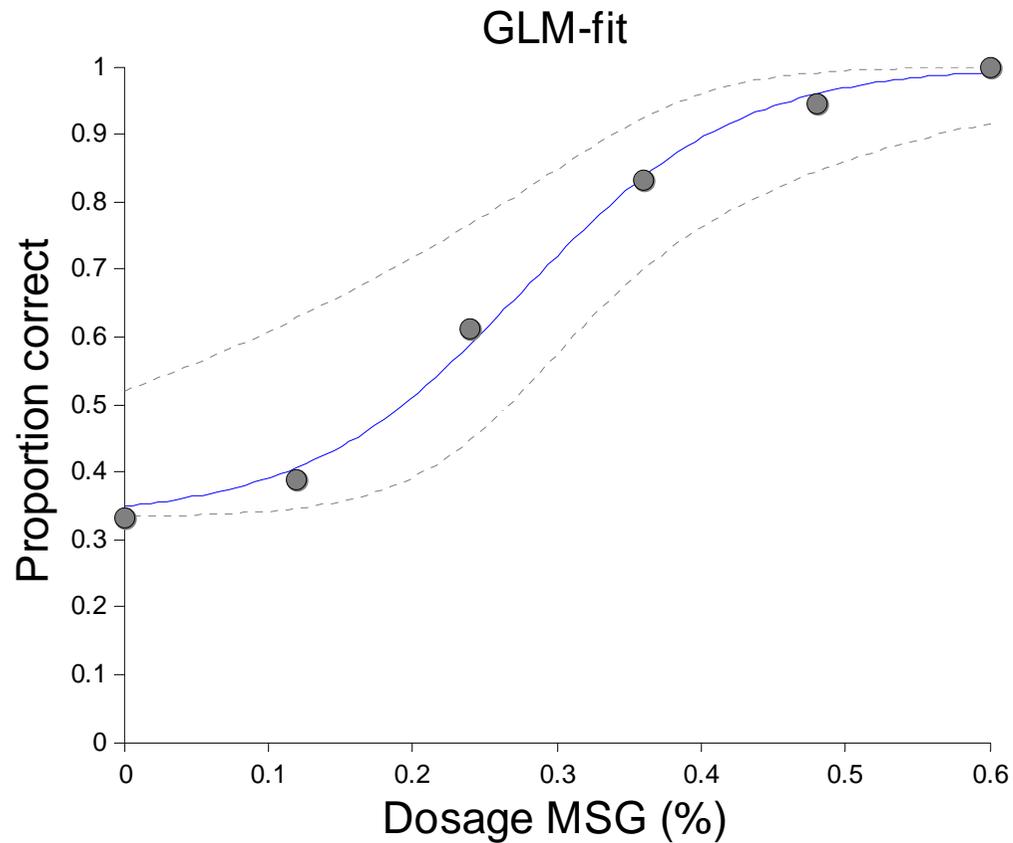
Treatments in beef bouillon	Trials	Successes
'0.24 % MSG'	18	8
'0% MSG'	18	6
'0.36 % MSG'	18	15
'0.12 % MSG'	18	6
'0.60 % MSG'	18	18
'0.48 % MSG'	18	16

LASSO-GLM

LASSO-GLM

K

QDA | +interactions

DSM
BRIGHT SCIENCE. BRIGHTER LIVING.

LASSO-GLM

QDA

PC1(82.5%)

Artificial.fl
Friedness.fl
Intensity.fl
Herbs.at
Sweet.fl
Meat.fl
Herbs.fl
0.12 % MSG
0.24 % MSG
0.36 % MSG
Intensity.at
Sweet.at
Salt.fl
0.48 % MSG
0% MSG
Sour.fl
Salt.at
Meat.at
0.60 % MSG
Length.of.AT
Bitter.fl
Off.flavour.fl
Allium.fl

DSM
BRIGHT SCIENCE. BRIGHTER LIVING.
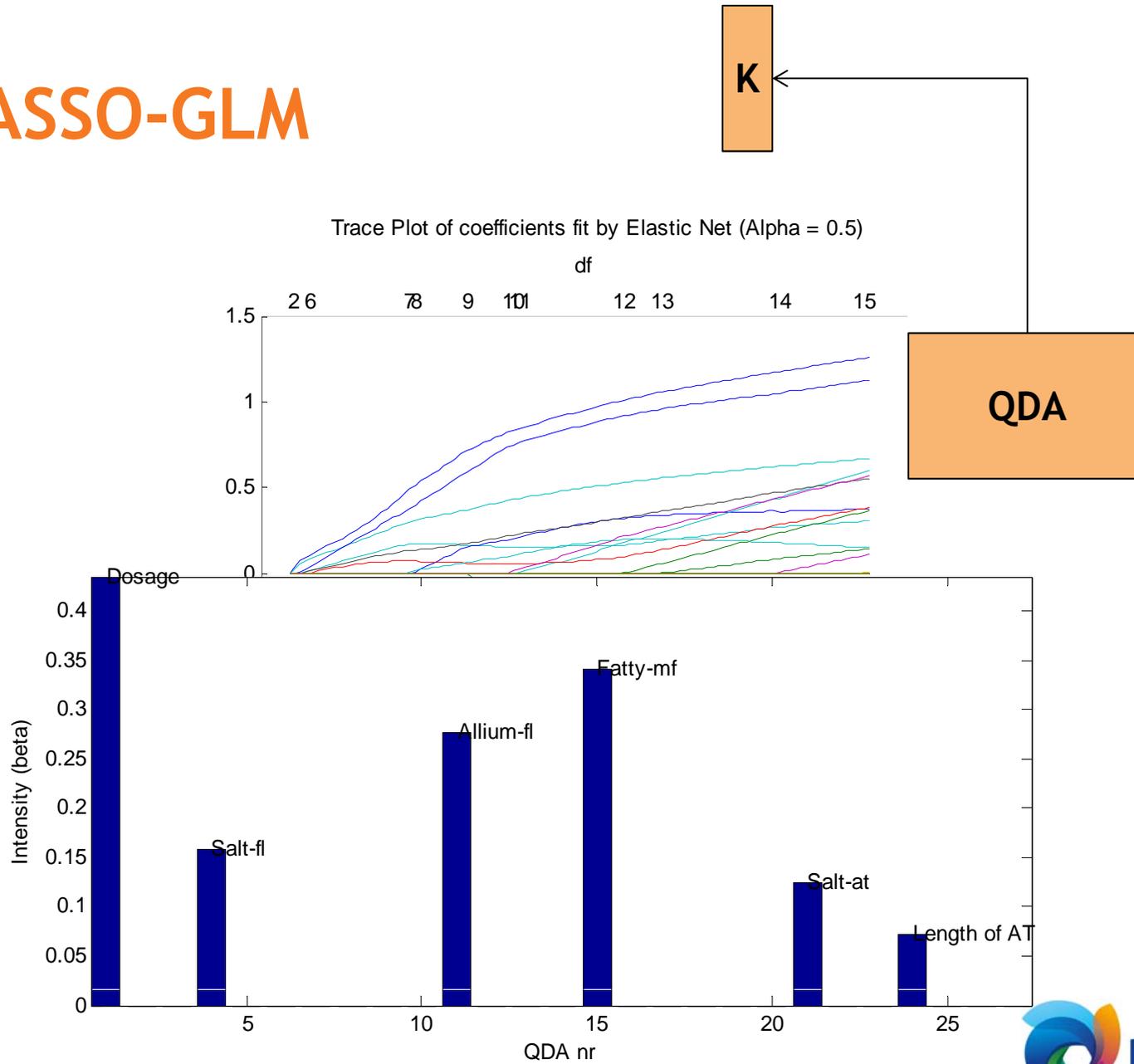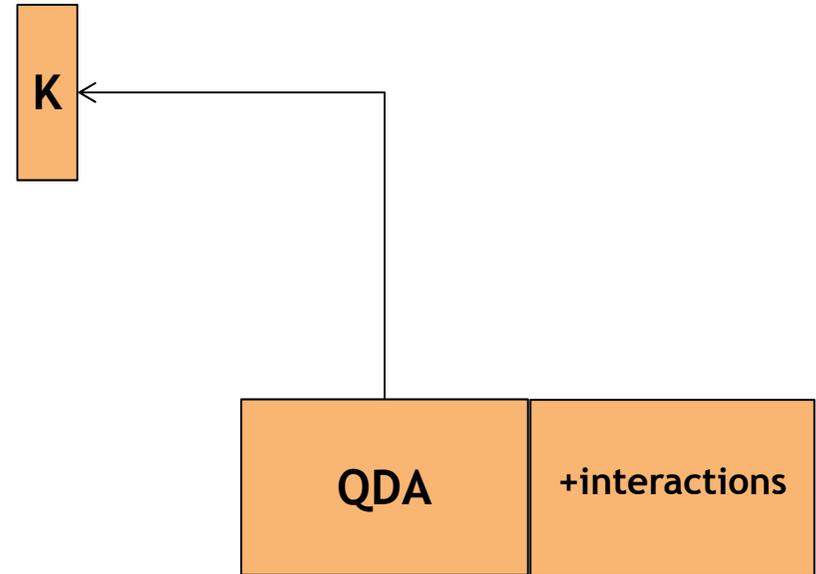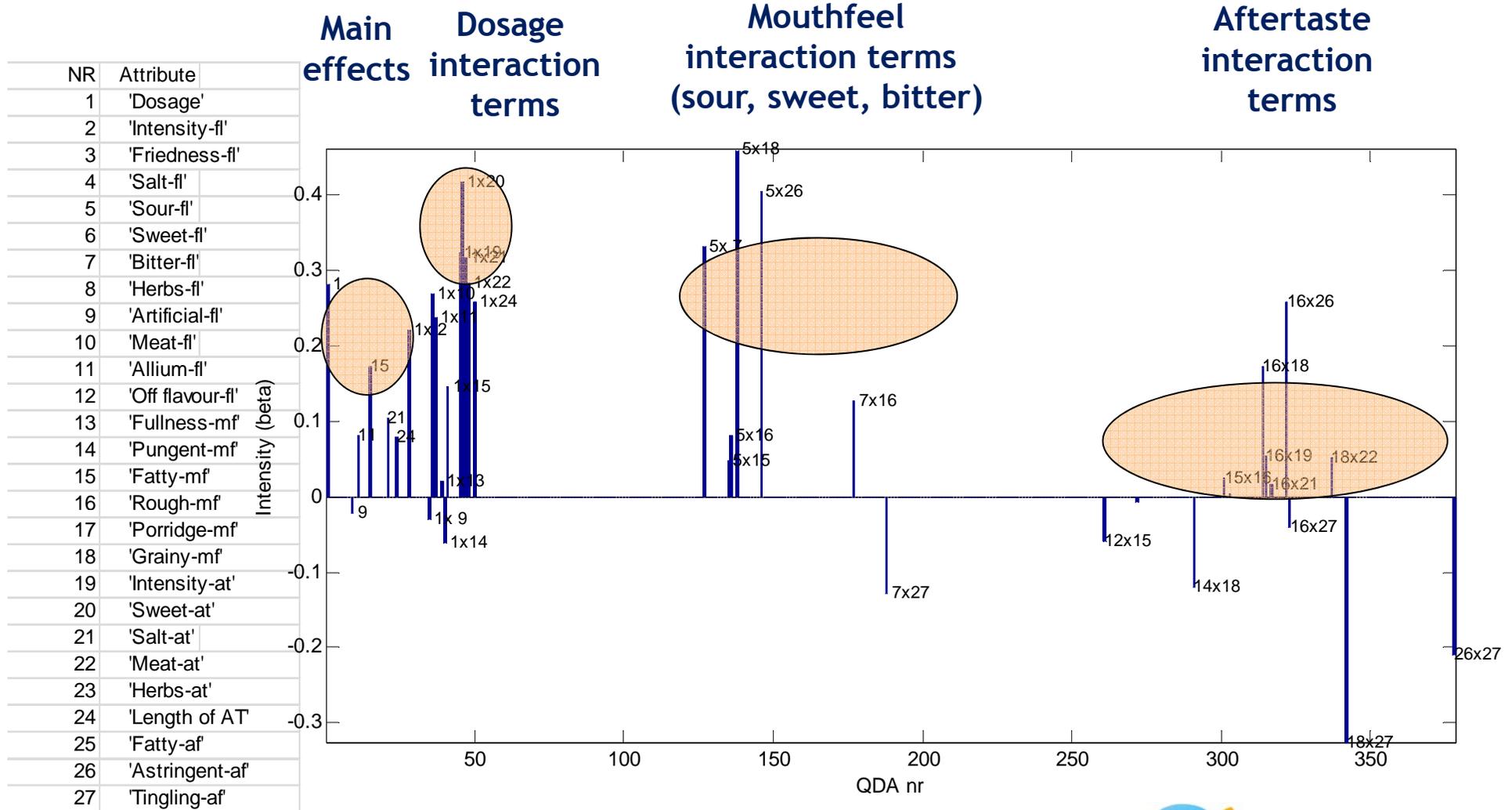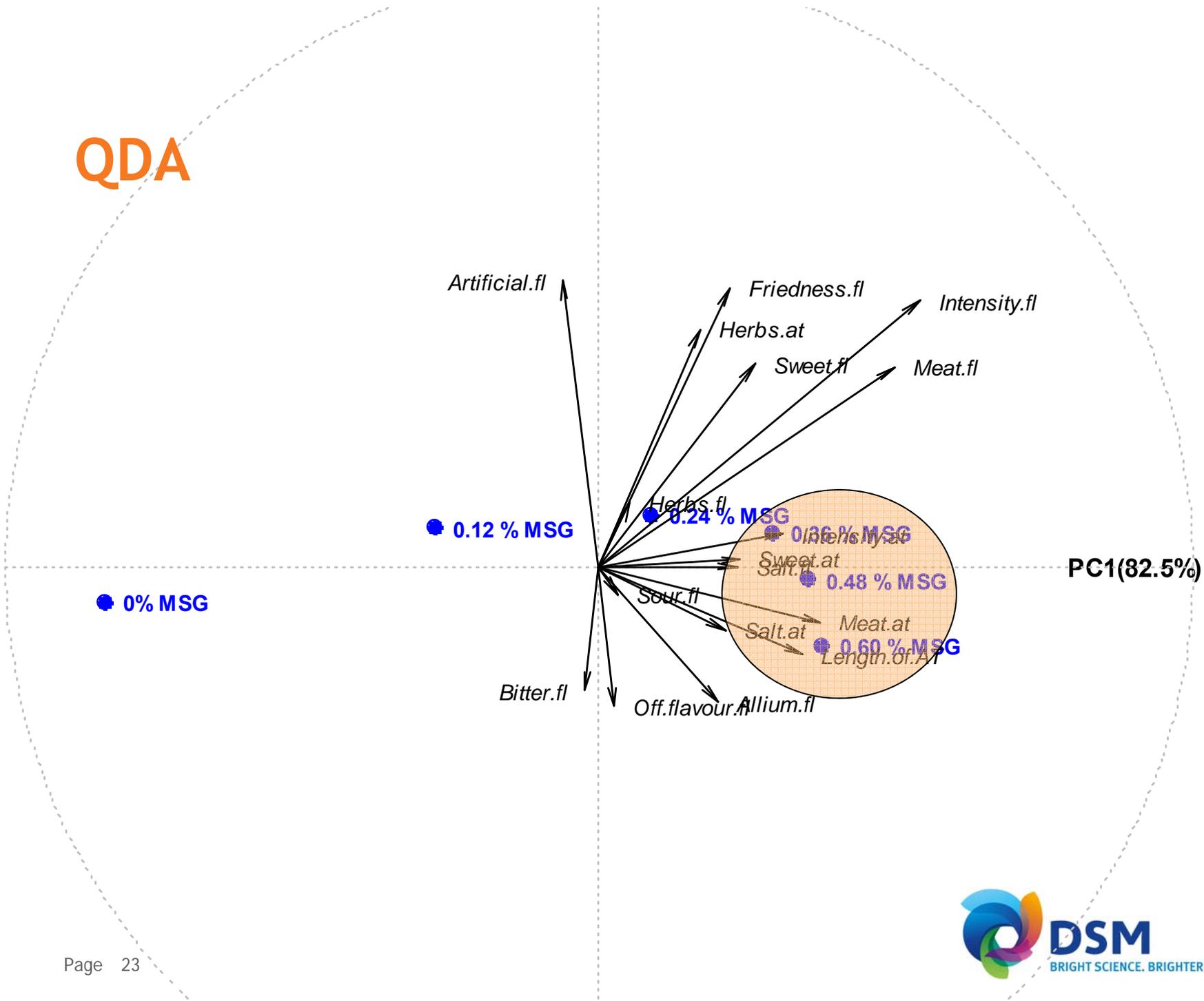
Take home messages / conclusion

- **Regularization is a powerful concept that enables**

    - stabilization (RR)
    - sparseness (variable selection in PCA in OLS)
    - improved interpretability
    - improved predictability (for unseen data)

- **Other types of restrictions (i.e. other than L_n) can be implemented as well**

    - roughness penalties for smoothness (Smooth PCA)
    - dispersion penalties (clustering)
    - m-univariate regression (using correlation in y) (applications in shelf life analysis)

DSM
BRIGHT SCIENCE. BRIGHTER LIVING.

Thank you for your attention!

BRIGHT SCIENCE. BRIGHTER LIVING.™