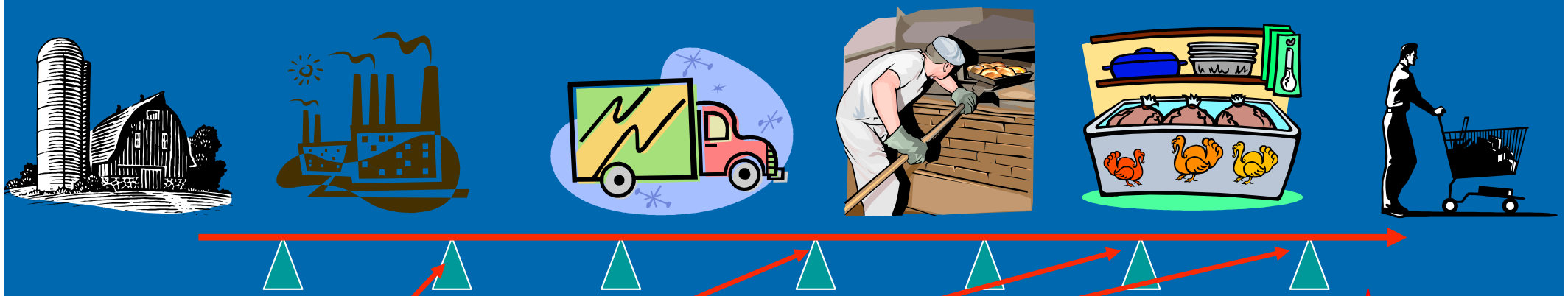
Factory door screening of food products: the potential and limitations of profiling techniques such as NIR

Dr Gerard Downey
Teagasc
Ashtown Food Research Centre, Dublin 15

Traceability System Overview



Transfer Point 1

Transfer Point 2

Transfer Point 3

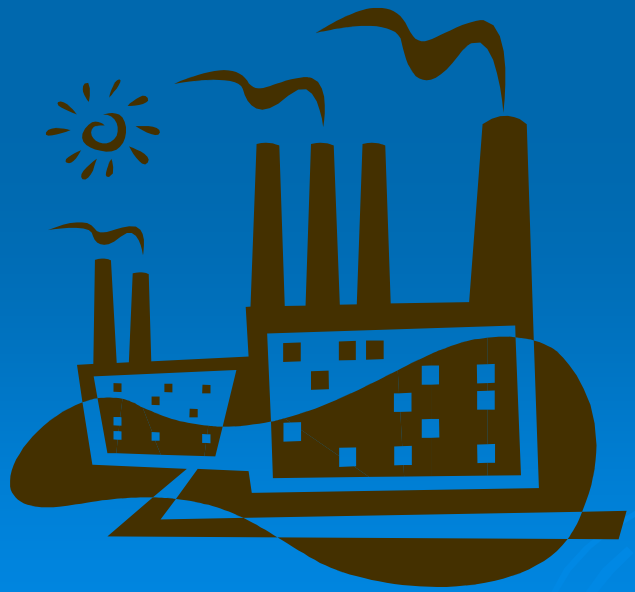
Transfer Point 4

Transfer Point 5

Paper Trail

Confirmatory Analysis For Authenticity

Factory Door Movements



Raw materials

Intermediate or finished foodstuffs

Distribution etc

What do we want to know about these materials *vis-à-vis* authenticity?

- **Conformance to specification** on purchase order or delivery documentation in the case of incoming material

- **Conformance to specification** on food label in the case of outgoing material

How do we address these knowledge requirements?

- Chemical or physical analysis
  - Time-consuming, expensive, destructive etc.

- Profiling methods
  - Generally fast, may be inexpensive, may be non-destructive, limited or no reagent disposal issues

Profiling Methods

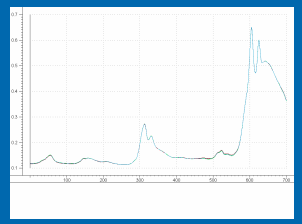
- Vibrational spectroscopy
  - Near infrared, mid-infrared, Raman
- Separation methods
  - Combinations of LC, GC with mass spectroscopy (MS); IRMS
- Nuclear magnetic resonance (NMR)

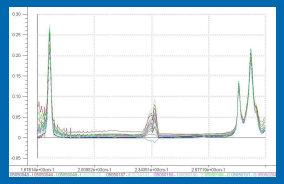
TRACE:
Tracing Food Commodities in Europe

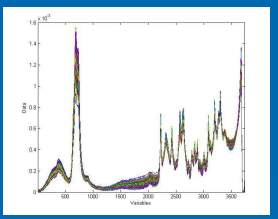
www.trace.eu.org trace.enquiries@trace.eu.org

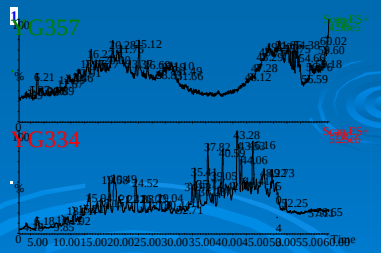
Information Type and Specificity of Profiling Methods

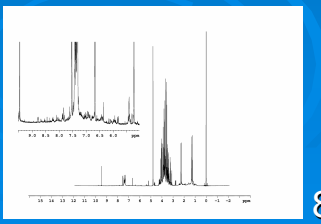
Method	Information	Specificity
NIR	Molecular bond vibration – overtones + combinations	Low to medium
FT-IR	Molecular bond vibration – fundamental	Medium
Raman	Light scatter	Low to medium
*-MS	Molecular fragment mass	High
NMR	Chemical shift	High









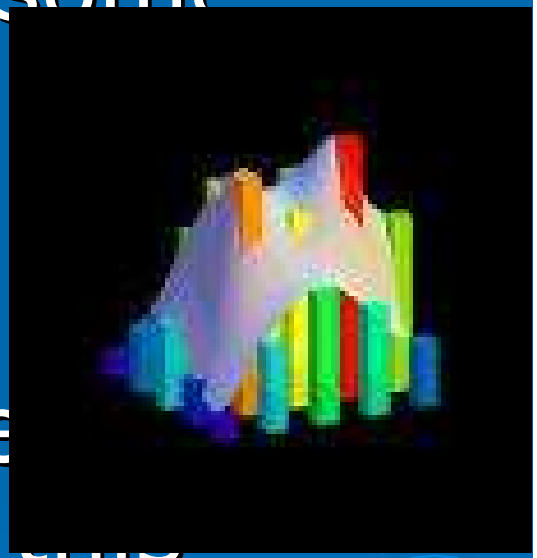


Speed of Analysis and Ease of Industrial Deployment of Profiling Methods

Method	Speed	Ease of Industrial Deployment
NIR	Fast - seconds	Very easy
FT-IR	Fast – seconds (ATR)	Easy
Raman	Fast - seconds	Moderate
*-MS	Slow	Low
NMR	Slow	Low

What type of questions do we ask of the analytical data?

- Is this sample the same as some specified material?
  - Classification
- To which of a limited number of possible sample types does this sample belong?
  - Discrimination



Classification and Discrimination

- Need a comprehensive collection of representative samples of material to develop robust and well-defined models
- For naturally-occurring samples, several harvest seasons need to be included
- Whether samples for test are natural or processed, sample collection period can be prolonged

Classification and Discrimination

- Correct use requires significant knowledge of and familiarity with chemometric procedures
- How to describe confidence regarding outputs from predictions of real, unknown samples?

Classification Options

- SIMCA (soft independent modelling of class analogy)
  - models a given type of *e.g.* raw material or finished product
  - projects unknown samples into model space
  - designates unknown as belonging or not belonging to model
  - probability of designation is statistically derived

SIMCA Deployment

- Example of honey adulteration
  - Irish artisanal honey (n=580)
  - adulterated by fully-inverted beet syrup (n=280), high fructose corn syrup (n=160), partial invert cane syrup (n=120), dextrose syrup (n=160) and beet sucrose (n=120)
  - all adjusted to 70 °Brix with distilled water; FT-IR ATR spectra collected
  - SIMCA model (using 6 PCs) was developed on 290 authentic honeys; test sample file contained 290 authentic and 800 adulterated samples

SIMCA Honey Results

Sample type	No. of samples	No. correctly identified	No. incorrectly identified	% correctly identified	% incorrectly identified	
Honey	290	279	11	**96.2**	**3.8**	
BS Adulterated	120	117	3	**97.5**	**2.5**	
CS adulterated	120	115	5	**95.8**	**4.2**	

SIMCA Honey Results

Sample type	No. of samples	No. correctly Identified	No incorrectly Identified	% correctly Identified	% incorrectly identified
Honey	290	279	11	**96.2**	3.8
PICS Adulterated	120	62	58	**51.2**	**48.3**
HFCS Adulterated	160	2	158	**1.3**	**98.8**
IB adulterated	280	19	241	**6.8**	**93.2**


Reasons for Failure

- Solutions are chemically identical, *or*
- Information on chemical differences is not contained in FT-IR ATR spectra, *or*
- Chemometric operations are insufficient to extract information

Discrimination

- Often stated as a limited class problem
- An example could be does this sample match one of a limited number of types?
- Some recent work involving NIR analysis of honey may serve as an example

Geographic origin of honey

- Preliminary feasibility study
- Sample set: 125 filtered honeys
  - 25 Irish
  - 25 Argentinean
  - 50 Czech
  - 25 Hungarian

Woodcock, T., Downey, G., Kelly, J.D.K. and O'Donnell, C.
J Ag. Fd Chem., 555(22), 9128-9134.

Four class PLS2 model results

		Argentinean (n=8)		Czech (n=16)		Hungarian (n=8)		Irish (n=8)	
Data	#L	% correct	% false	% correct	% false	% correct	% false	% correct	% false
Pretreatment		classification	positives	classification	positives	classification	positives	classification	positives
Raw Data	4	50	0	**100**	0	**100**	0	75	37.5
1st der, 5 pts	5	37.5	0	**100**	0	**100**	0	**100**	**25**
1st der, 9 pts	5	37.5	0	**100**	0	**100**	0	**100**	**25**
1st der, 13 pts	5	37.5	0	**100**	0	**100**	0	**100**	**25**
1st der, 21 pts	5	25	0	**100**	0	**100**	0	**100**	**25**
2nd der, 5 pts	5	62.5	0	**100**	0	**100**	0	87.5	50
2nd der, 9 pts	5	62.5	0	**100**	0	**100**	0	87.5	50
2nd der, 13 pts	5	50	0	**100**	0	**100**	0	100	37.5
2nd der, 21 pts	5	50	0	**100**	0	**100**	0	100	37.5
SNV	5	**75**	**0**	**100**	0	**100**	0	87.5	75

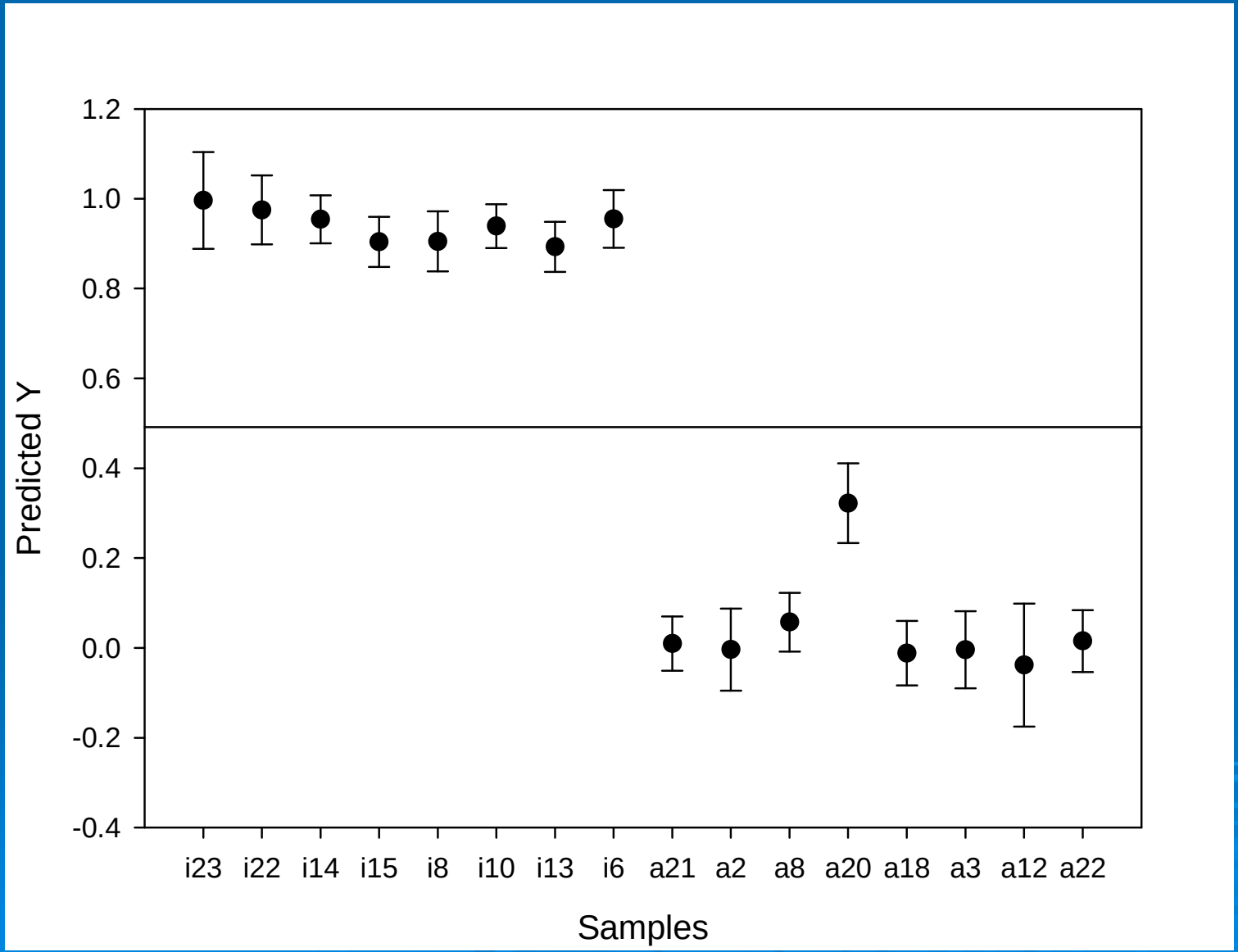
Argentine and Irish Separation

		Argentinean (n=8)		Irish (n=8)	
Data Pretreatment	#L	% correct classification	% false positives	% correct classification	% false positives
Raw Data	5	87.5	0	100	12.5
1st der, 5 pts	4	**100**	**0**	**100**	**0**
1st der, 9 pts	4	**100**	**0**	**100**	**0**
1st der, 13 pts	4	**100**	**0**	**100**	**0**
1st der, 21 pts	4	**100**	**0**	**100**	**0**
2nd der, 5 pts	4	**100**	**0**	**100**	**0**
2nd der, 9 pts	4	**100**	**0**	**100**	**0**
2nd der, 13 pts	4	**100**	**0**	**100**	**0**
2nd der, 21 pts	4	**100**	**0**	**100**	**0**
SNV	4	**100**	**0**	**100**	**0**

Operation of PLS1

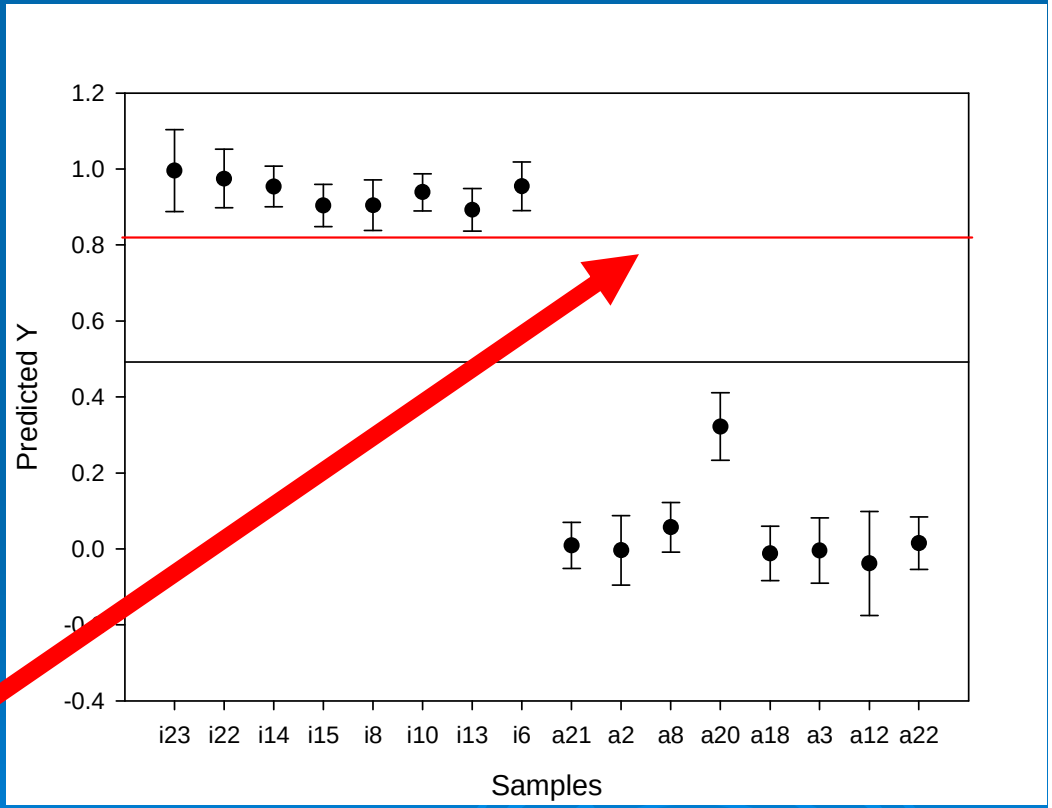
- Ascribe dummy variables to each sample type e.g.
  - 1 for honey belonging to a given class, &
  - 0 for other honeys
- Predict unknown honeys
- Select cut-off value for class ascription *e.g.* 0.5

Practical Outputs



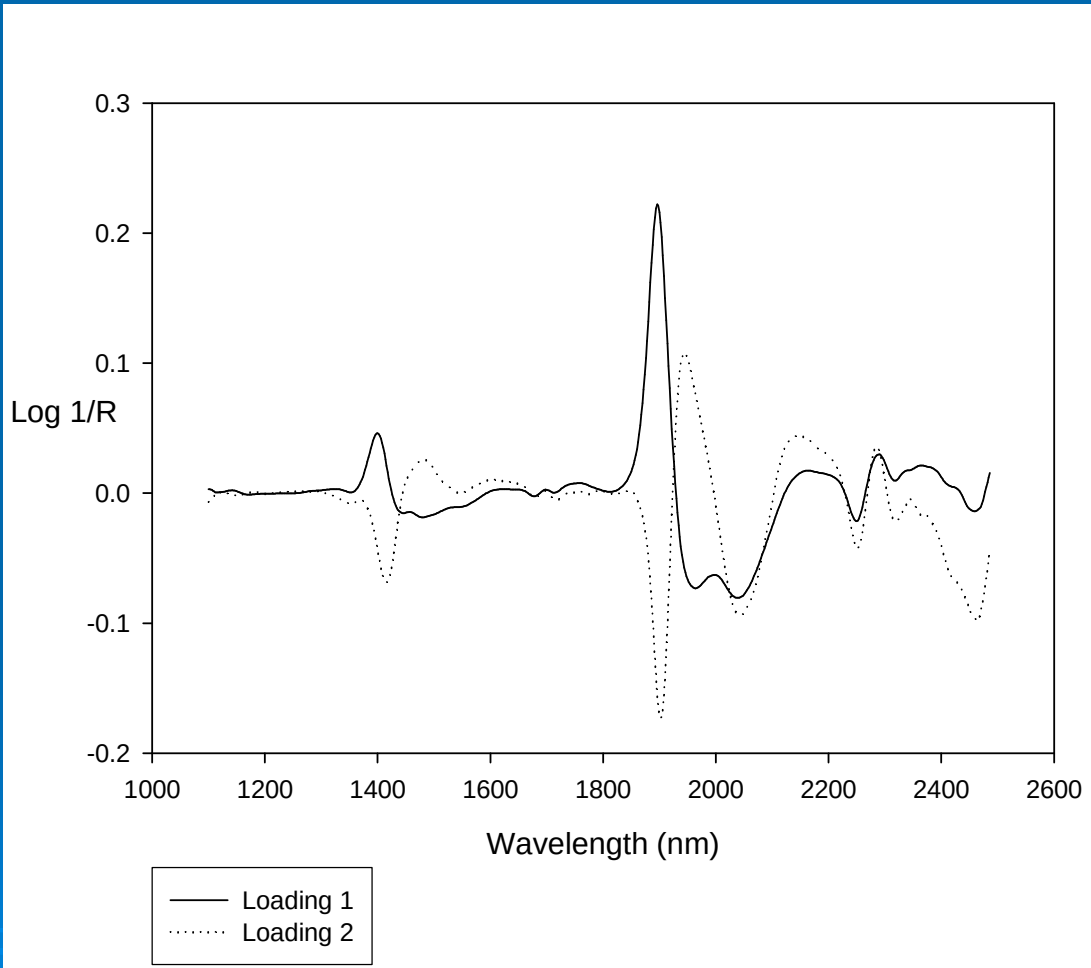
Utilisation

- Such a model may be used to generate a specification derived from multivariate data
- How to decide on value of cut-off for specification?



How does the model work?

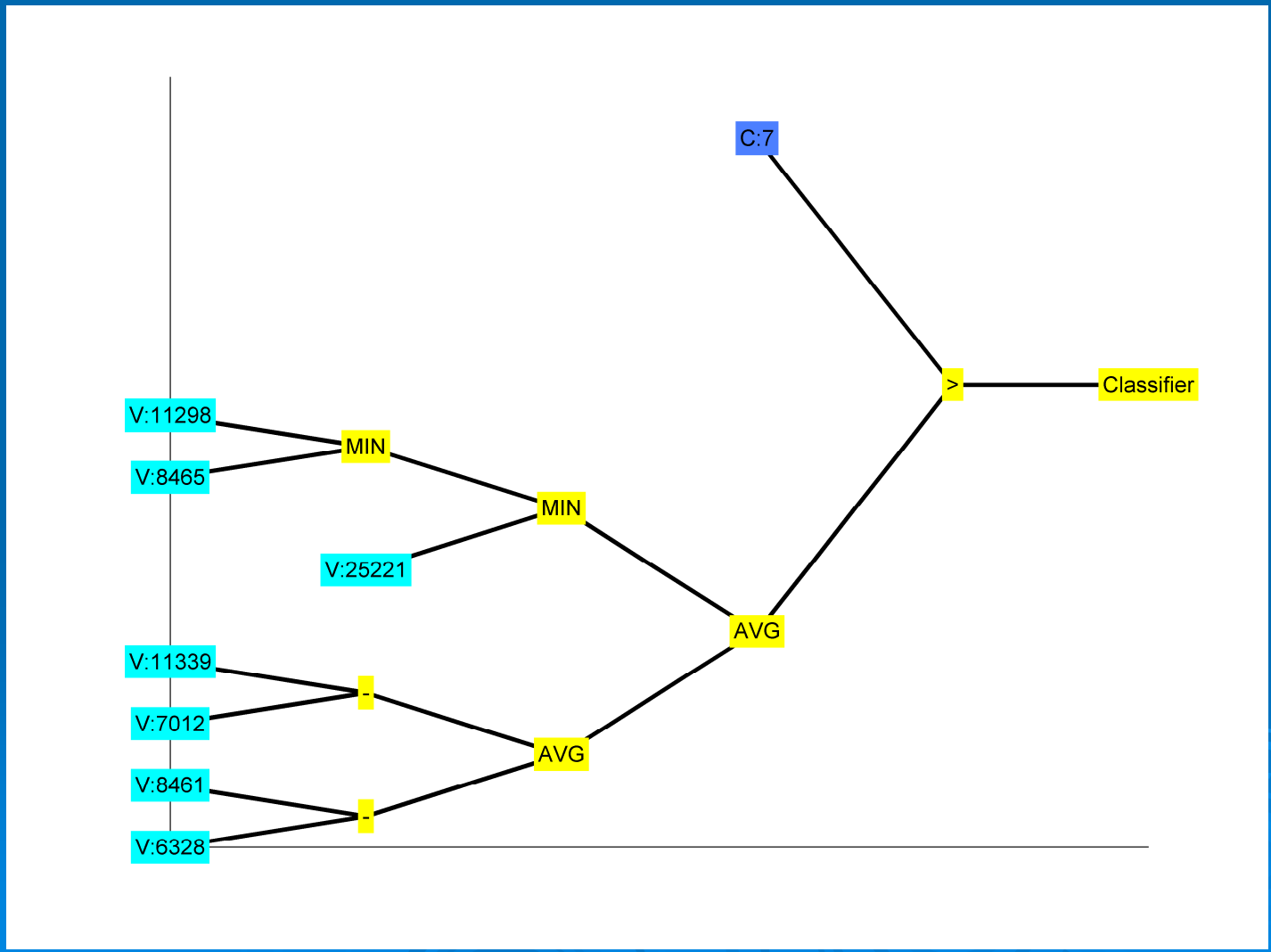
- Examination of e.g. PLS loading coefficients
- Example shown is from Argentine:Irish PLS1 model
- Difficulty in unambiguous identification of chemical species is apparent



Variable Selection

- Can often improve especially SIMCA models
- Permit selection of limited number of variables grouped in such a way as to provide a single number cut-off value
- Particularly applicable to */MS and NMR data

Variable Selection + Genetic Programming



Exploitation Scenario - 1

- Develop classification or discriminant model for material of interest using profiling technique
- Ensure robustness of model by creating extensive and representative spectral database
- Continually update as poorly predicted authentic samples are encountered

Exploitation Scenario - 2

- Each "producer" can generate fingerprint model of his/her product
- Production batches of this product may be checked using the model and guaranteed as meeting specification
- *Caveat*: In the absence of identifiable marker compounds such specifications may be used only to confirm conformance – they are not exclusive
- There is a trade-off between speed and accuracy; profiling methods can optimise this trade-off

Conclusions

- Fingerprinting and profiling methods can provide sufficient reassurance for efficient trading
- Industry needs to grasp the opportunity and invest in database collection by some technique
- Positive results from conformance tests may be used for marketing purposes as well as QC checks

C’est fini!

Merci pour votre attention.