

Multivariate Characterization of Olive Oils using $^1\text{H-NMR}$ and $^{13}\text{C-IRMS}$ Techniques

*Rosa Maria Alonso Salces, Fabiano Reniero,
Jose Manuel Moreno Rojas, Claude Guillou,
Károly Héberger[†]*


*Institute for Health and Consumer Protection, Physical and Chemical
Exposure Unit, BEVABS T.P. 281, I-21020 Ispra (VA), Italy*

*[†] Chemical Research Center, Hungarian Academy of Sciences,
P.O. Box 17, H-1525, Budapest, Hungary*





Geographical characterization of olive oils by $^1\text{H-NMR}$ and $^{13}\text{C-IRMS}$

1. **Trace 2005*** 
2. **Italian olive oils 2005 (regions)**
3. **Olive oils 2005 (countries)**
4. **Unsaponifiable fraction of olive oils
(countries)**

* TRACE project is funded by the EU through the Sixth Framework Programme under the Food Quality and Safety Priority (<http://www.trace.eu.org>). WP2: Fingerprinting and Profiling Methods.





1. Trace 2005

TRACE (<http://www.trace.eu.org>) is project funded by the EU through the Sixth Framework Programme under the Food Quality and Safety Priority, which aims to develop cost effective analytical methods, that should allow the determination of authenticity and detection of fraud in several food products, olive oil being one of them.

Objective:

Classification of olive oils as Ligurian or non-Ligurian

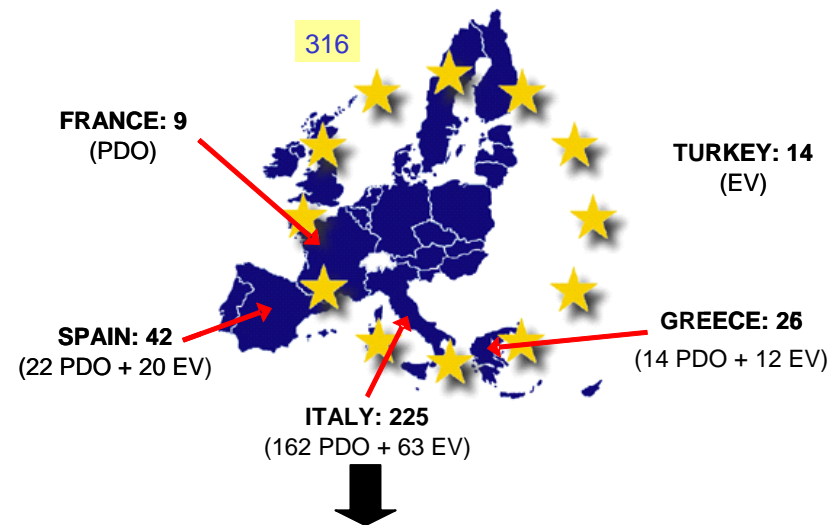
Samples:

316 olive oils

{ 63 Ligurian
253 non-Ligurian

Variables:

175 buckets of ¹H-NMR spectra



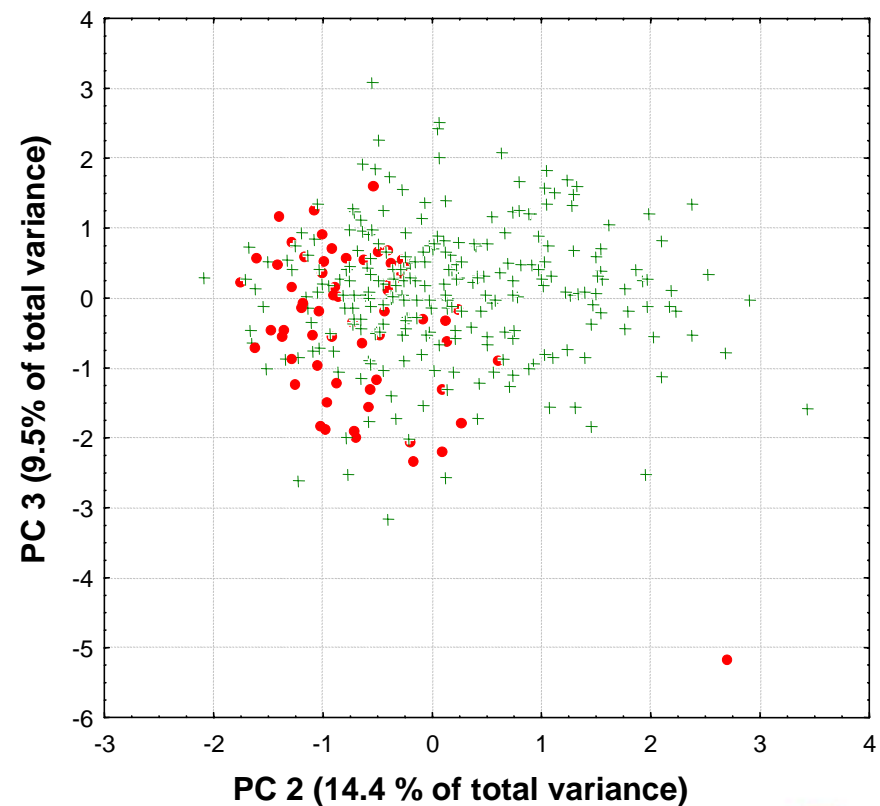
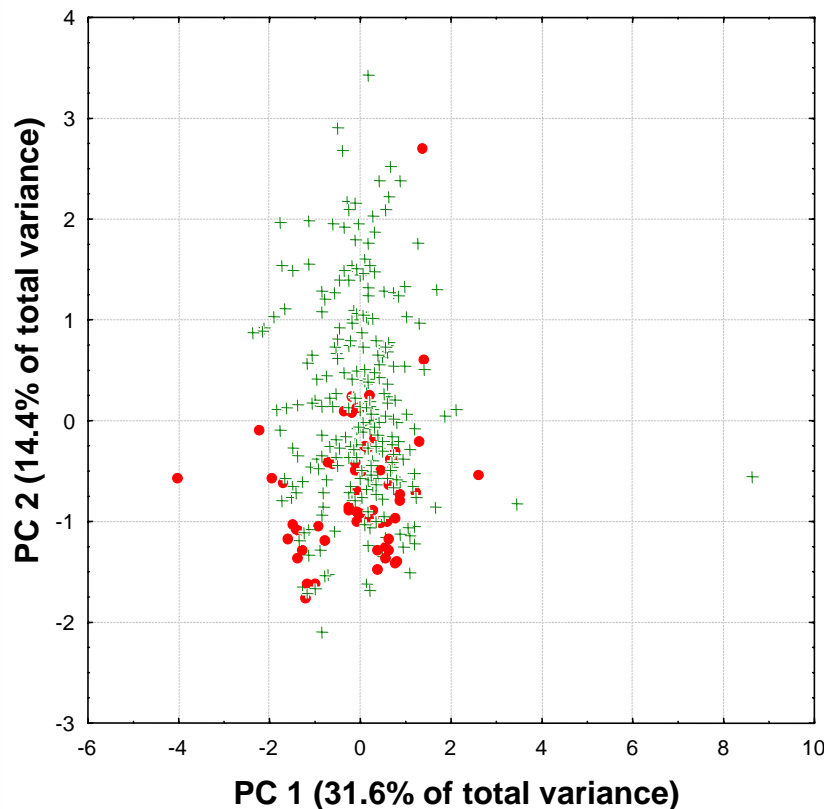
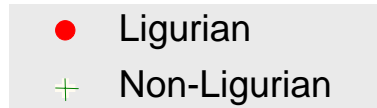
Origin	samples	Classes
Italy - Liguria	63	Ligurian
Italy - other regions	162	Non-Ligurian
Spain	42	Non-Ligurian
Greece	26	Non-Ligurian
Turkey	14	Non-Ligurian
France	9	Non-Ligurian



1. Trace 2005

316c x 175v

Principal Component Analysis: Ligurian olive oils are grouped,
but overlapping with non-Ligurian class.





1. Trace 2005

Comparison of supervised pattern recognition techniques: PLS DA and LDA

Cross-validation (x3):

- 2/3 of samples in training set
- 1/3 of samples in test set

LDA:

- Variable selection, **175v**:

Modified best-subset and forward stepwise

- 6 selected variables:

6.90, 6.02, 5.38, 4.70, 1.62, 0.90 ppm

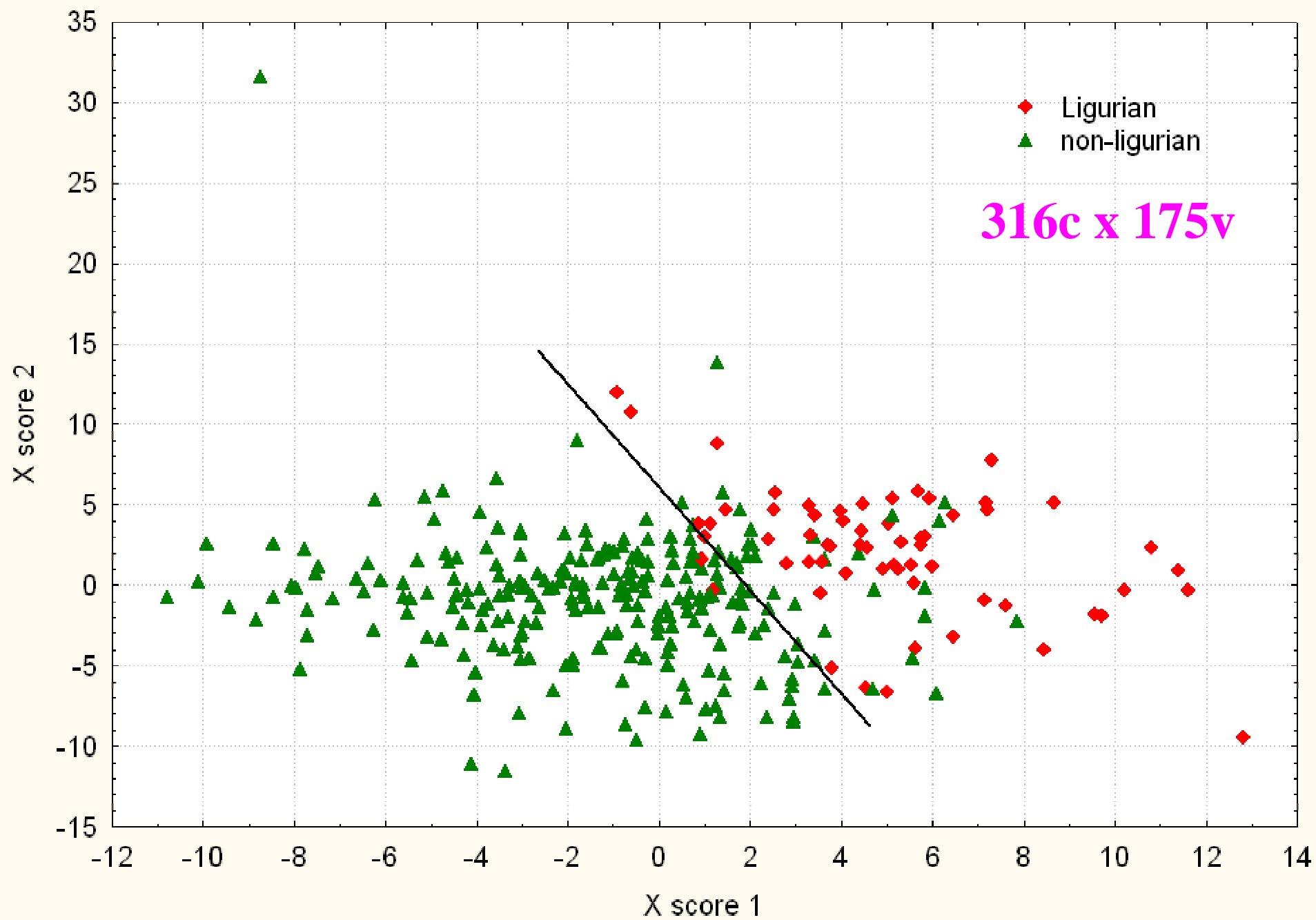
PLS DA

175v, no variable selection

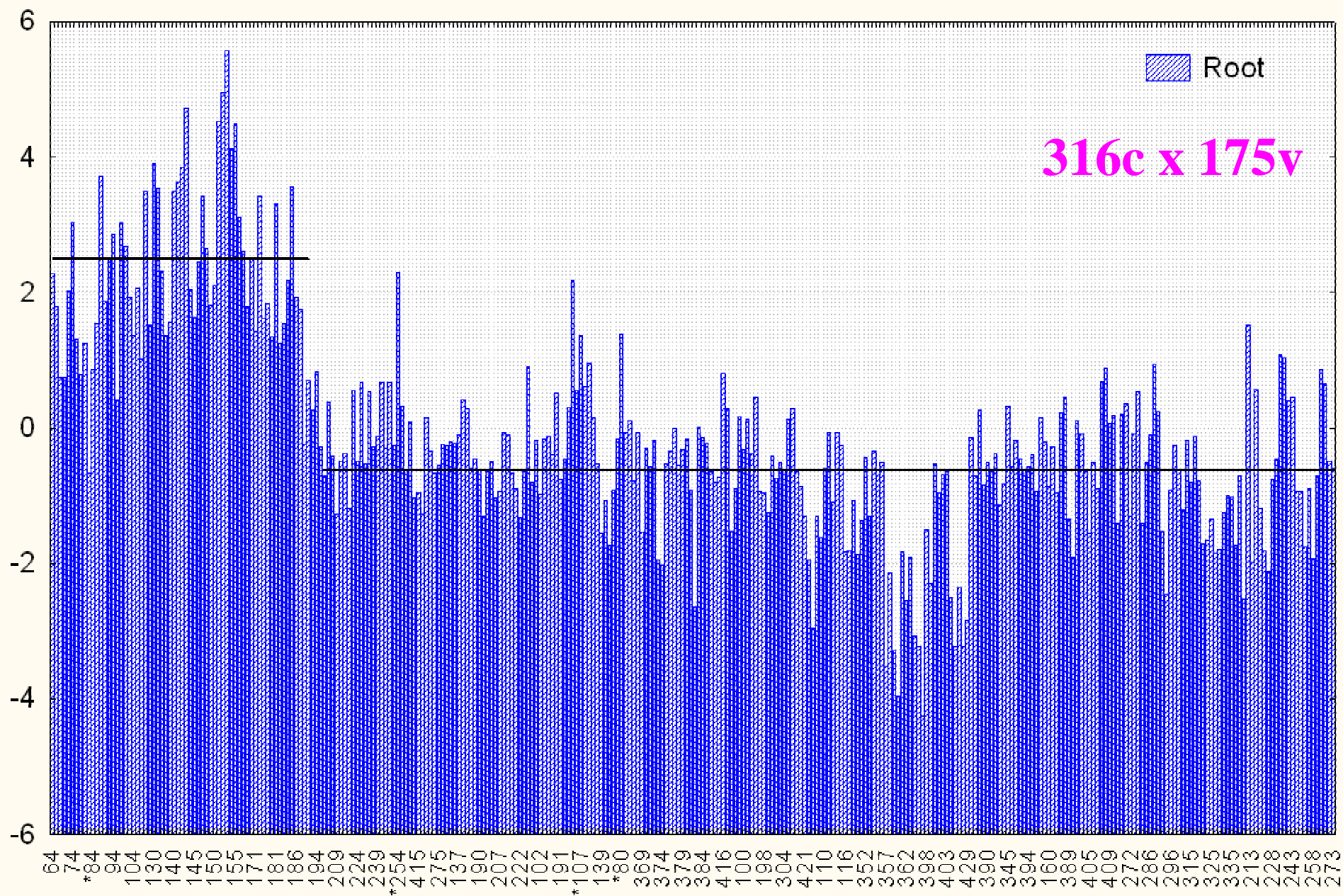
3500v, no variable selection,

Uncertainty test (Martens)

PLS DA results



LDA results (six best NMR shifts selected)
6.90, 6.02, 5.38, 4.70, 1.62, 0.90 ppm





1. Trace 2005

Supervised pattern recognition techniques: PLS DA and LDA

316c x 175v

Abilities (%)	PLS DA	LDA
Recognition	98.6	95.3
Ligurian	99.4	82.3
Non-Ligurian	95.2	98.4
Prediction	96.5	95.3
Ligurian	98.4	82.3
Non-Ligurian	88.7	98.4
Classification	97.9	95.3
Ligurian	99.1	82.3
Non-Ligurian	93.0	98.4

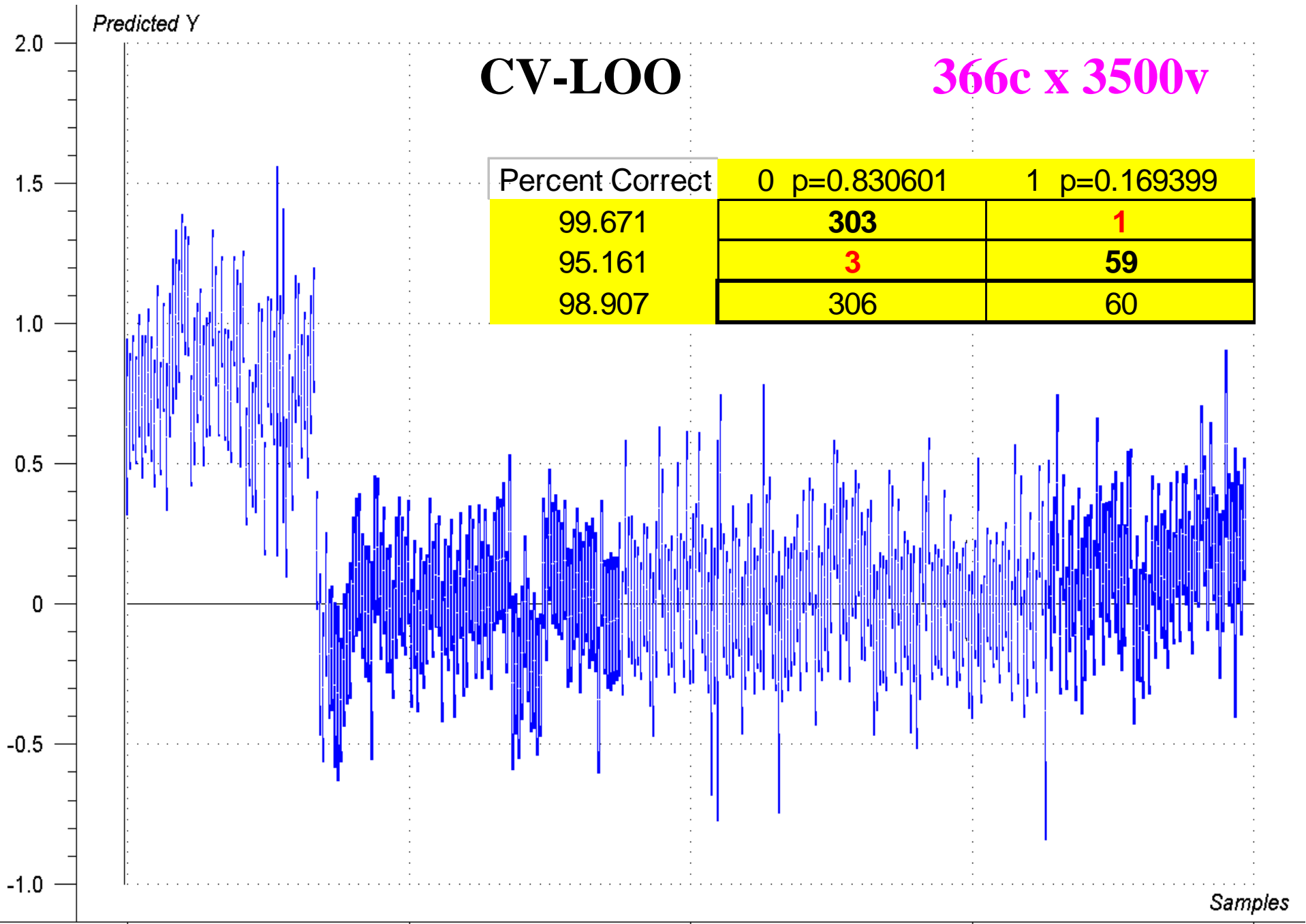
- ✓ ¹H-NMR spectra contains **useful information** for the classification of olive oils according to their origin: Ligurian or non-Ligurian olive oils.
- ✓ The supervised pattern recognition techniques PLS DA and LDA achieve **complementary** results, allowing the correct classification (>98% of hits) of both classes.
- ✓ ¹³C-IRMS **does not contain** significant information for the classification of olive oils in Ligurian or non-Ligurian.

Predicted Y

CV-LOO

366c x 3500v

Percent Correct	0 p=0.830601	1 p=0.169399
99.671	303	1
95.161	3	59
98.907	306	60



olio_2.5mm_187

olio_2.5mm_400

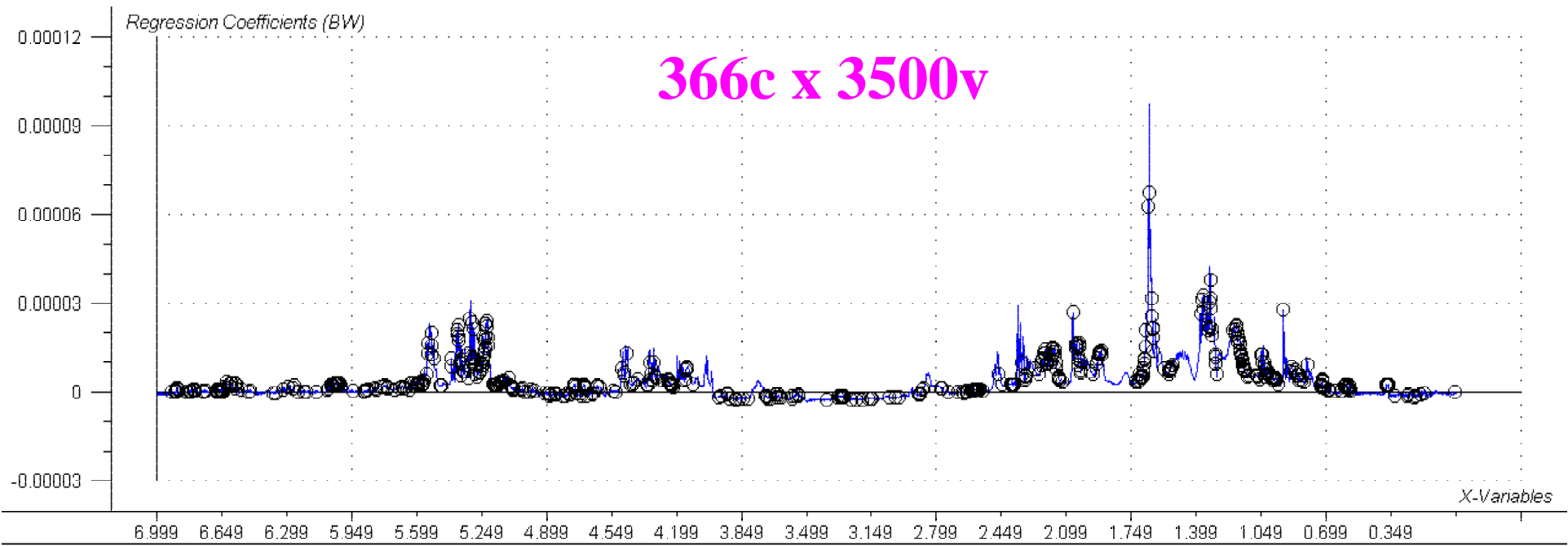
olio_2.5mm_308

olio_2.5mm_216

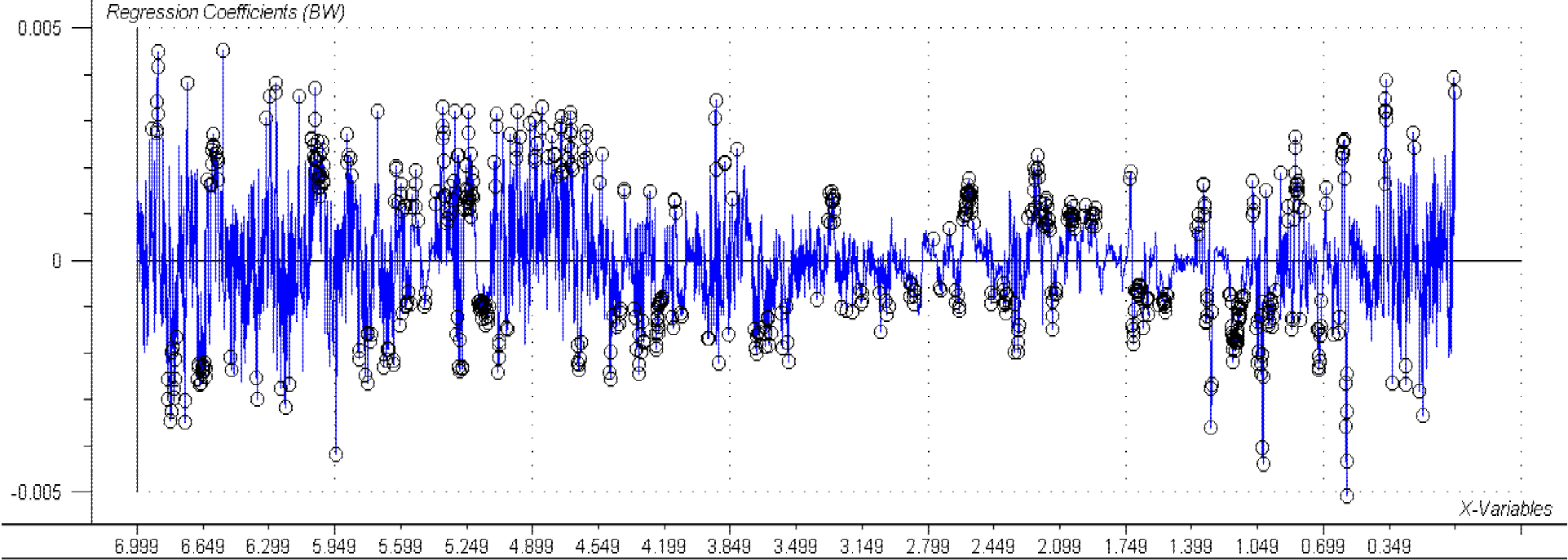
Samples

RESULT2, (Y-var, PC): (11,6)

366c x 3500v



PLS1LOOPLS7noMC..., (Y-var, PC): (II,1) BQV = 0.000000



PLS1LOOPLS7noMC..., (Y-var, PC): (II,7) BQV = 0.000000



2. ITALIAN OLIVE OILS 2005

Objective:

Classification of Italian olive oils by regions

Samples: 225 olive oils

Origin	Samples
Liguria	62
Garda	18
Centre	73
Lazio	29
Umbria	18
Abruzzo	6
Molise	13
Campania	7
South	72
Puglia	28
Calabria	13
Sicilia	31

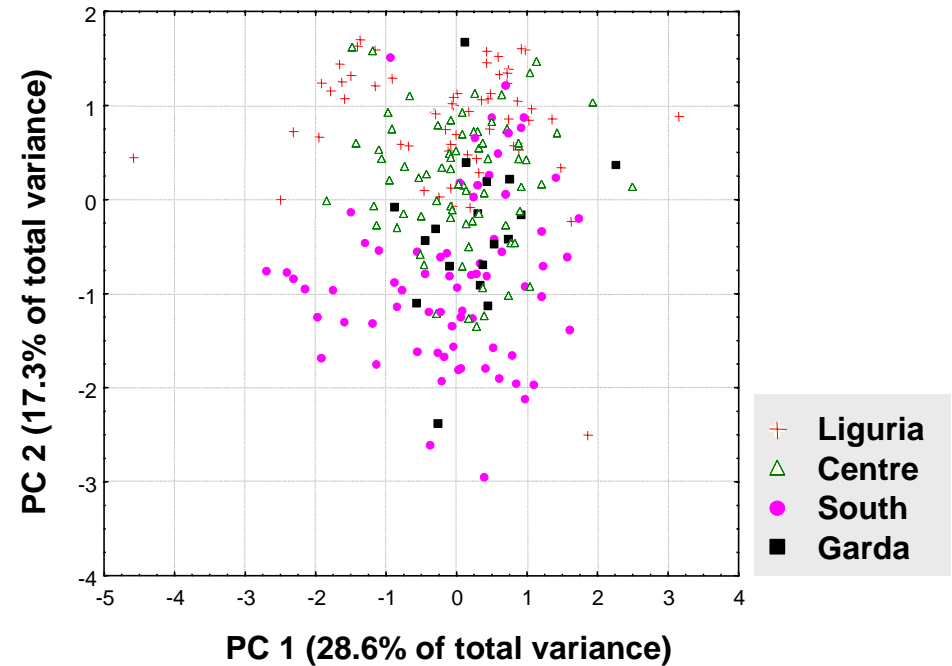
Variables:

175 buckets ¹H-NMR spectra

+

¹³C-IRMS

Principal Component Analysis:



- ❖ Samples of each class are grouped, but groups are overlapping.



2. ITALIAN OLIVE OILS 2005

Supervised pattern recognition technique: LDA

Crossvalidation (x3):

- 2/3 of samples in training set
- 1/3 of samples in test set

LDA:

- Variable selection:
Modified best-subset and forward stepwise
- 8 selected variables:
6.66, 5.10, 4.70, 4.62, 4.22, 2.82, 2.14, 2.10 ppm

LDA results using 9 variables:

8 NMR
+
¹³C-IRMS



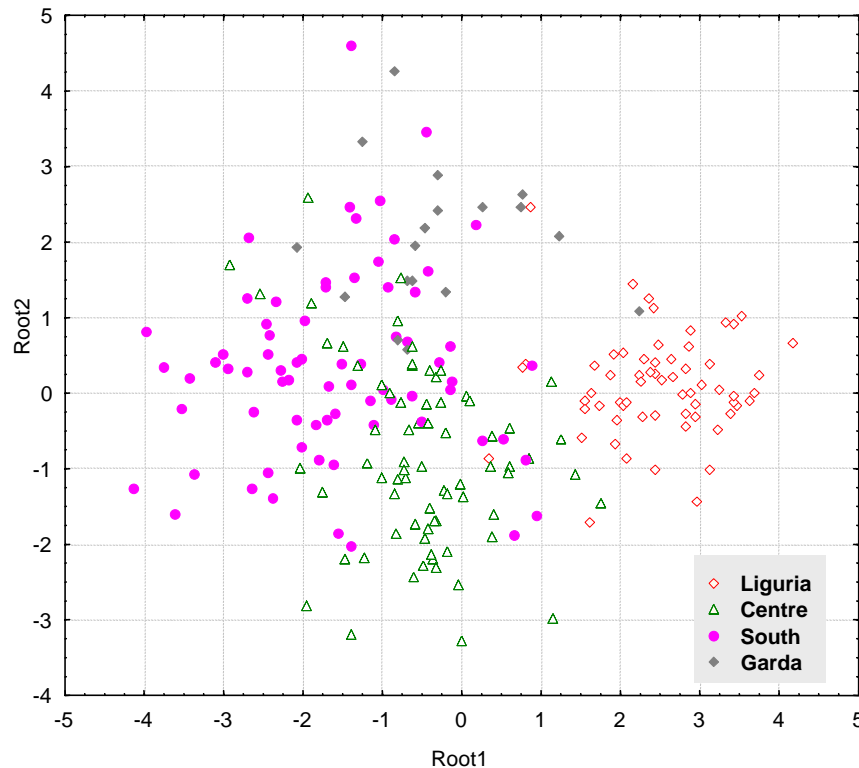
Abilities (%)	Recognition	Prediction	Classification
	82.4	79.1	81.3
Liguria	96.0	91.9	94.6
Centre	80.1	75.3	78.5
South	74.3	73.6	74.1
Garda	77.8	72.2	75.9



2. ITALIAN OLIVE OILS 2005

Supervised pattern recognition technique: LDA

LDA: 8 NMR variables + ^{13}C -IRMS



- ✓ ^1H -NMR spectra and ^{13}C -IRMS data contain **useful information** for the classification of Italian olive oils according to their origin: Liguria, Garda, Centre and South of Italy.
- ✓ LDA achieves a **satisfactory classification** for Ligurian olive oils, although better classification results are expected by increasing and equilibrating the number of samples.



3. OLIVE OILS 2005

Objective:
Classification of olive oils by country of origin.

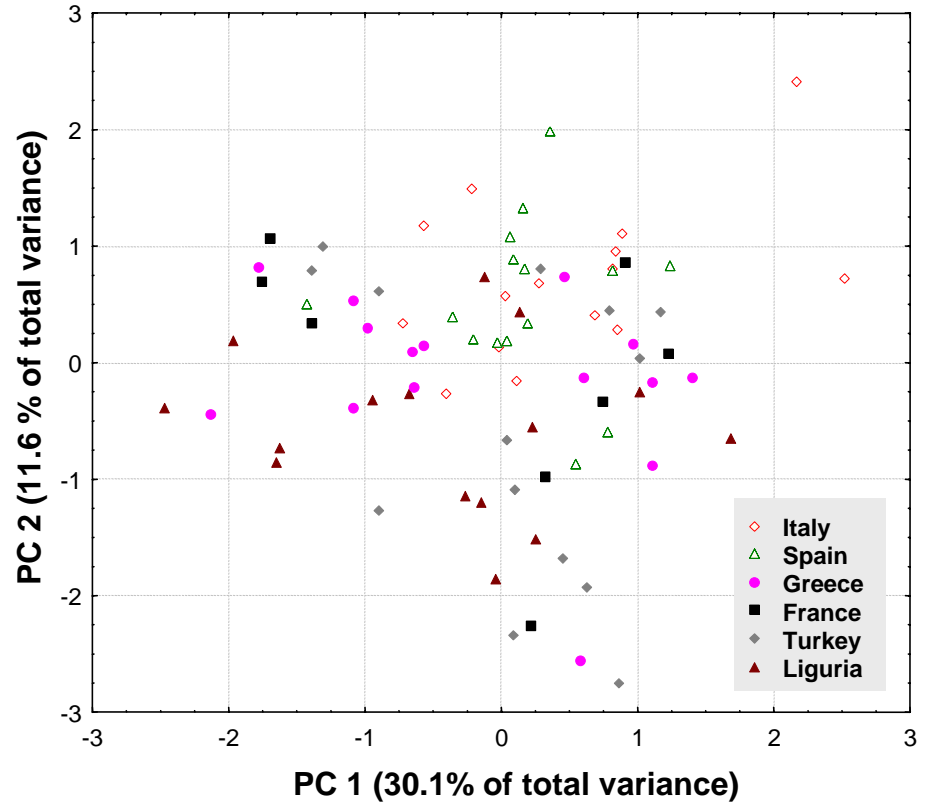
Samples: 364 olive oils

Origin	Samples
Italy *	163
Spain *	72
Greece *	44
Turkey	14
France	9
Liguria *	62

* 15 samples were selected randomly, the rest was used for external validation

Variables:
175 buckets ¹H-NMR spectra
+
¹³C-IRMS

Principal Component Analysis:





3. OLIVE OILS 2005

Supervised pattern recognition techniques: PLS DA and LDA

Crossvalidation (x3):

- 2/3 of samples in training set
- 1/3 of samples in test set

LDA:

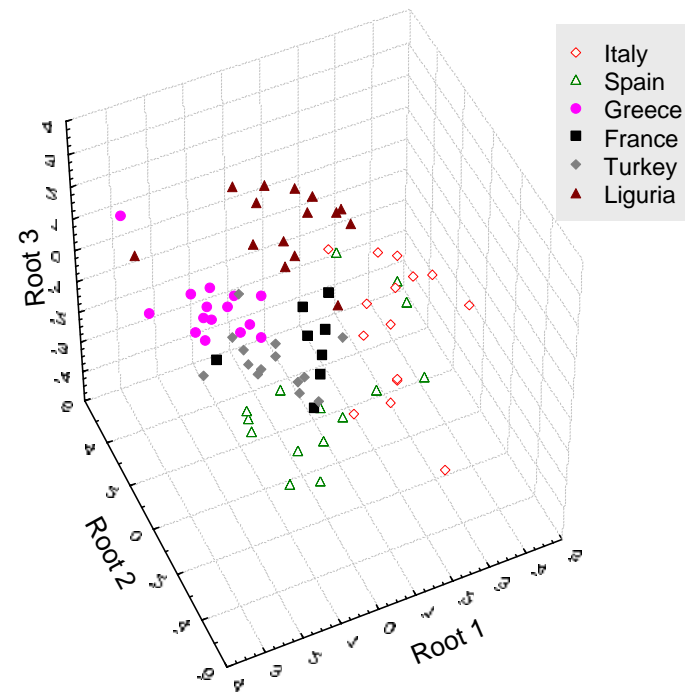
- Variable selection:

Modified best-subset and forward stepwise

- 7 selected NMR variables:

6.62, 5.18, 4.66, 4.26, 2.58, 2.18, 1.34 ppm

LDA: 7 NMR variables + ^{13}C -IRMS





3. OLIVE OILS 2005

Supervised pattern recognition techniques: PLS DA and LDA

Abilities (%)	CROSS-VALIDATION		EXTERNAL VALIDATION	
	PLS	LDA	PLS	LDA
Recognition	77.7	89.6		
Italy	66.7	93.3		
Spain	80.0	83.3		
Greece	80.0	96.7		
France	50.0	93.8		
Turkey	89.3	82.1		
Liguria	100.0	90.0		
Prediction	51.9	80.5		
Italy	66.7	80.0	54.3	66.9
Spain	40.0	66.7	36.3	52.6
Greece	53.3	93.3	63.1	84.5
France	0.0	75.0		
Turkey	71.4	78.6		
Liguria	80.0	86.7	64.5	75.2
Classification	69.1	86.6		
Italy	66.7	88.9		
Spain	66.7	77.8		
Greece	71.1	95.6		
France	33.3	87.5		
Turkey	83.3	81.0		
Liguria	93.3	88.9		

- ✓ ¹H-NMR spectra and C¹³-IRMS contains useful information for the classification of olive oils according to their country of origin.
- ✓ LDA achieves satisfactory classification of Greek oils (96%), and 78-89% for the other oil origins.
- ✓ Better classification are expected by increasing and equilibrating the number of samples.



4. UNSAPONIFIABLE FRACTION OF OLIVE OILS

Objective:

Classification of olive oils by country of origin.

Samples:

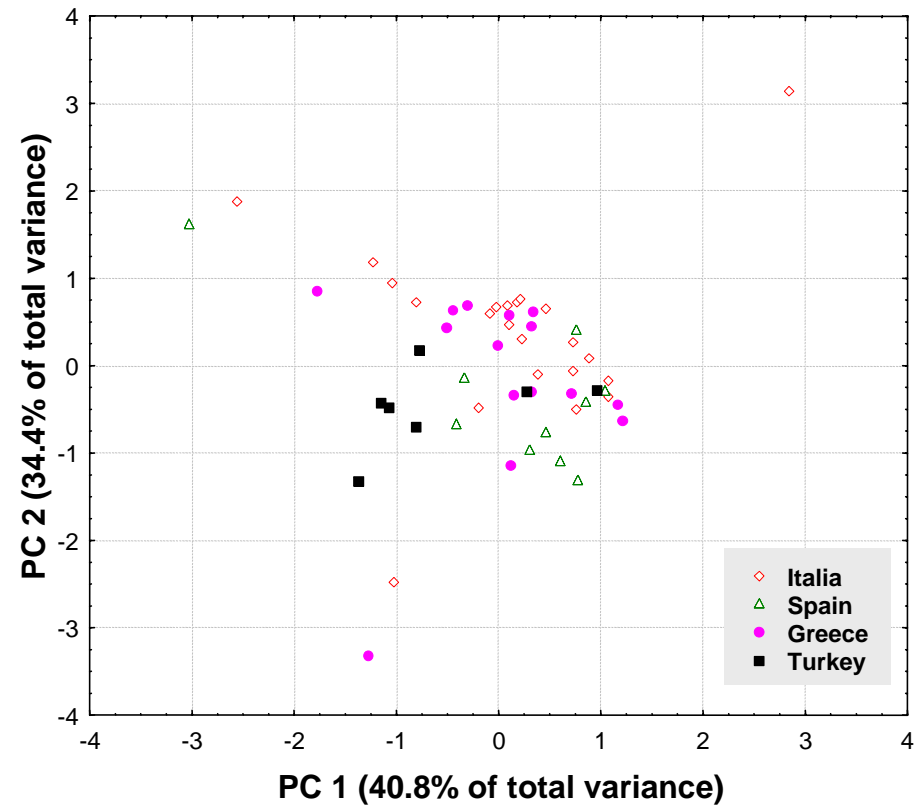
54 olive oil unsaponifiable fractions

Origin	Samples
Italy	22
Spain	10
Greece	15
Turkey	7

Variables:

209 buckets of $^1\text{H-NMR}$ spectra

Principal Component Analysis:





4. UNSAPONIFIABLE FRACTION OF OLIVE OILS

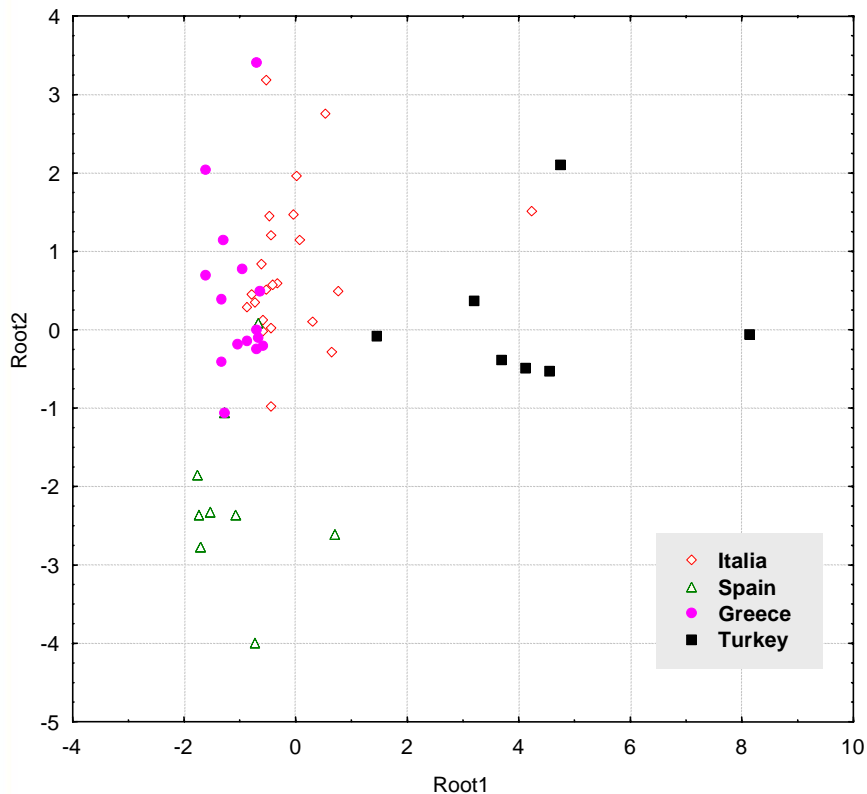
Supervised pattern recognition techniques: LDA

Crossvalidation (x3):

- 2/3 of samples in training set
- 1/3 of samples in test set

LDA:

- Variable selection: Modified best-subset and forward stepwise
- 5 selected variables: 5.28, 4.76, 3.72, 0.88, 0.60 ppm



Abilities (%)	LDA
Recognition	83.3
Italia	97.7
Spain	85.0
Greece	56.7
Turkey	92.9
Prediction	81.3
Italia	90.9
Spain	80.0
Greece	66.7
Turkey	85.7
Classification	82.7
Italia	95.5
Spain	83.3
Greece	60.0
Turkey	90.5

✓ This preliminary results of the analysis of the unsaponifiable fraction of olive oils by ¹H-NMR spectra indicate potential, although better classification results are expected by increasing the number of samples.



CONCLUSIONS

- ✓ ^1H -NMR spectra of olive oils and olive oil unsaponifiable fractions contain useful information for the classification of olive oils according to their geographical origin.
- ✓ ^{13}C -IRMS data of olive oils can provide 'complementary' information to NMR data for the geographical characterization of olive oils.
- ✓ Both analytical techniques, NMR and IRMS, together with multivariate data analysis have a 'potential' for the differentiation of olive oils from different geographical origins.
- ✓ Better classification results are expected by increasing the number of samples, and equilibrating the number of samples in each class.



Acknowledgements

Rosa Maria Alonso Salces,

Fabiano Reniero,

Jose Manuel Moreno Rojas,

Claude Guillou

JRC Institute for Health and Consumer
Protection, Physical and Chemical Exposure Unit
Chemical Research Center, Hungarian Academy
of Sciences

TRACE project (<http://www.trace.eu.org>)

