



2nd MS Wine Day 9-10 Maggio 2017

RUFFINO
LA VITA RUFFINO



Constellation Brands

Ruffino – Progetto Vigneto Italia

2nd MS Wine Day 9-10 Maggio 2017

CREA, Conegliano (TV)

2nd MS Wine Day 9-10 Maggio 2017

RUFFINO
LA VITA RUFFINO



Constellation Brands

In collaborazione con :

- Agilent Technologies
- SRA Instruments
- Dipartimento di Scienze dell'Ambiente e del Territorio e di Scienze della Terra – Milano Chemometrics and QSAR Research group - *Università degli Studi di Milano Bicocca*.
- Semeion, Research Center of Sciences of Communication, Rome, Italy
- Dept. of Mathematical and Statistical Sciences, University of Colorado Denver, CO, USA
- Università Cattolica Sacro Cuore, Istituto di Chimica agraria ed ambientale



Agilent Technologies



CU COLLEGE OF LIBERAL ARTS AND SCIENCES
Mathematical and Statistical Sciences
UNIVERSITY OF COLORADO DENVER

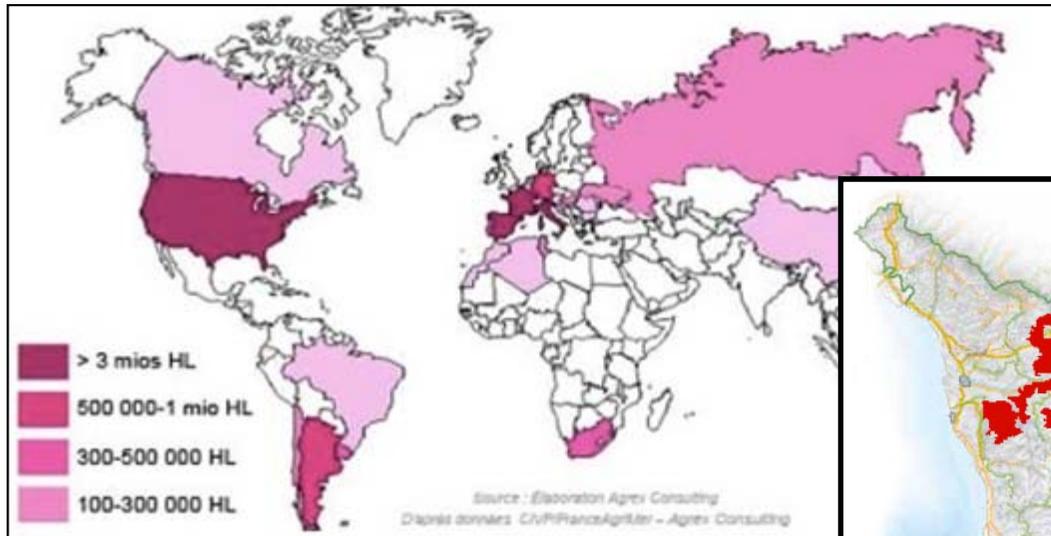
2nd MS Wine Day 9-10 Maggio 2017

RUFFINO
LA VITA RUFFINO

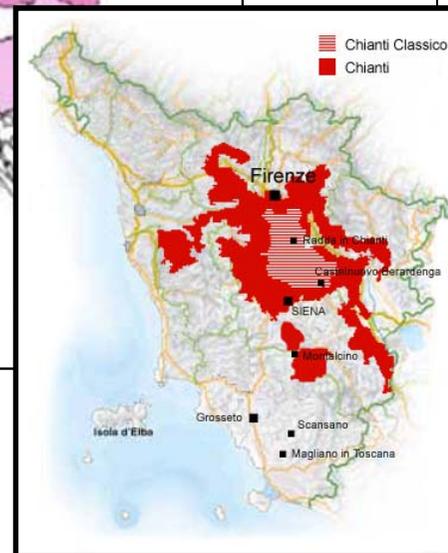
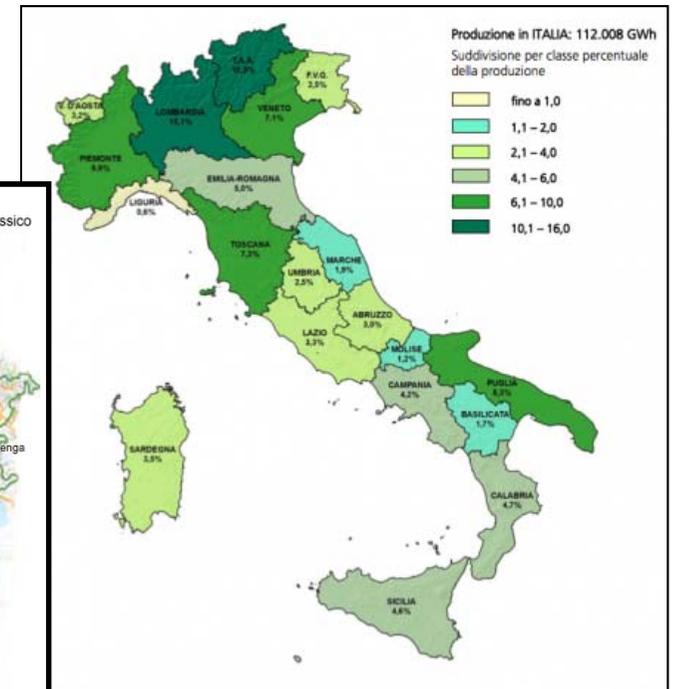
OBIETTIVO

Trovare una correlazione fra **elementi** e **origine geografica** di vini provenienti da zone molto vicine tra loro

Zone di produzione di uva nel Mondo



Zone di produzione di uva in Italia



Zone di produzione di Chianti-Chianti Classico

ELEMENTI MINERALI NEL VINO



IA 0
 26 — Numero atomico
 Simbolo chimico — **Fe**
 55.8 — Massa atomica relativa (uma)

IIIA	IVA	VA	VIA	VIIA	0
3	6	7	8	9	10
B	C	N	O	F	Ne

IVB	V	VI	VII	VIII	IB
22	23	24	25	26	27
Ti	V	Cr	Mn	Fe	Co
47.88	50.94	51.99	54.94	55.85	58.93
40	41	42	43	44	45
Zr	Nb	Mo	Tc	Ru	Rh
91.22	92.91	95.94	(98.906)	101.07	101.07
72	73	74	75	76	77
Hf	Ta	W	Re	Os	Ir
178.49	180.95	183.85	186.21	190.23	192.22
104	105	106	107	108	109
Rf	Db	Sg	Bh	Hs	Mt
(261)	(262)	(263)	(264)	(265)	(266)

58	59	60	61	62	63	64	65	66	67	68	69	70	71
Ce	Pr	Nd	Pm	Sm	Eu	Gd	Tb	Dy	Ho	Er	Tm	Yb	Lu
140.12	140.91	144.24	(144.913)	150.36	151.96	157.25	158.93	162.50	164.93	167.26	168.93	173.04	174.97
90	91	92	93	94	95	96	97	98	99	100	101	102	103
Th	Pa	U	Np	Pu	Am	Cm	Bk	Cf	Es	Fm	Md	No	Lr
232.04	231.04	238.03	(237.048)	(244)	(243)	(247)	(247)	(251)	(252)	(257)	(258)	(259)	(260)

IIIA IVA VA VIA VIIA 0
 5 6 7 8 9 10
 B C N O F Ne

IVB V VI VII VIII IB
 22 23 24 25 26 27 28 29
 Ti V Cr Mn Fe Co Ni Cu
 47.88 50.94 51.99 54.94 55.85 58.93 58.93 63.55
 40 41 42 43 44 45 46 47
 Zr Nb Mo Tc Ru Rh Pd Ag
 91.22 92.91 95.94 (98.906) 101.07 101.07 106.42 107.87
 72 73 74 75 76 77 78 79
 Hf Ta W Re Os Ir Pt Au
 178.49 180.95 183.85 186.21 190.23 192.22 195.08 196.97
 104 105 106 107 108 109 110 111 112
 Rf Db Sg Bh Hs Mt S S S
 (261) (262) (263) (264) (265) (266) (269) (272) (277)

58 59 60 61 62 63 64 65 66 67 68 69 70 71
 Ce Pr Nd Pm Sm Eu Gd Tb Dy Ho Er Tm Yb Lu
 140.12 140.91 144.24 (144.913) 150.36 151.96 157.25 158.93 162.50 164.93 167.26 168.93 173.04 174.97
 90 91 92 93 94 95 96 97 98 99 100 101 102 103
 Th Pa U Np Pu Am Cm Bk Cf Es Fm Md No Lr
 232.04 231.04 238.03 237.05 (244) (243) (247) (247) (251) (252) (257) (258) (259) (260)

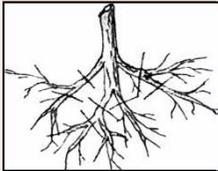


2nd MS Wine Day 9-10 Maggio 2017

RUFFINO
LA VITA RUFFINO

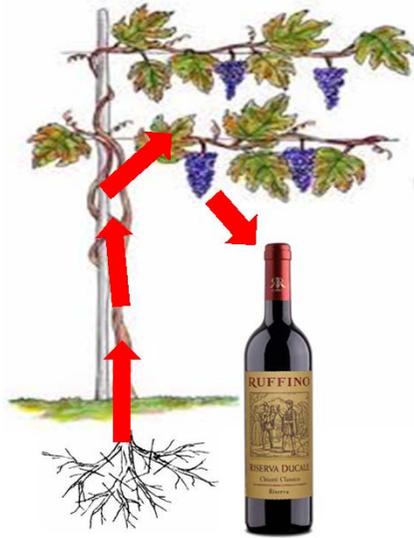
ELEMENTI MINERALI NEL VINO

Il contenuto di elementi nel vino è il risultato di una somma di fattori naturali e artificiali:

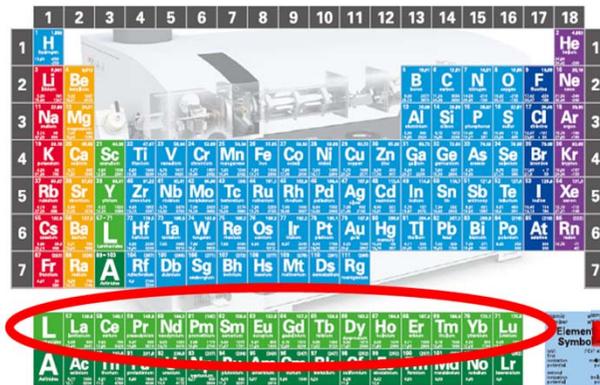
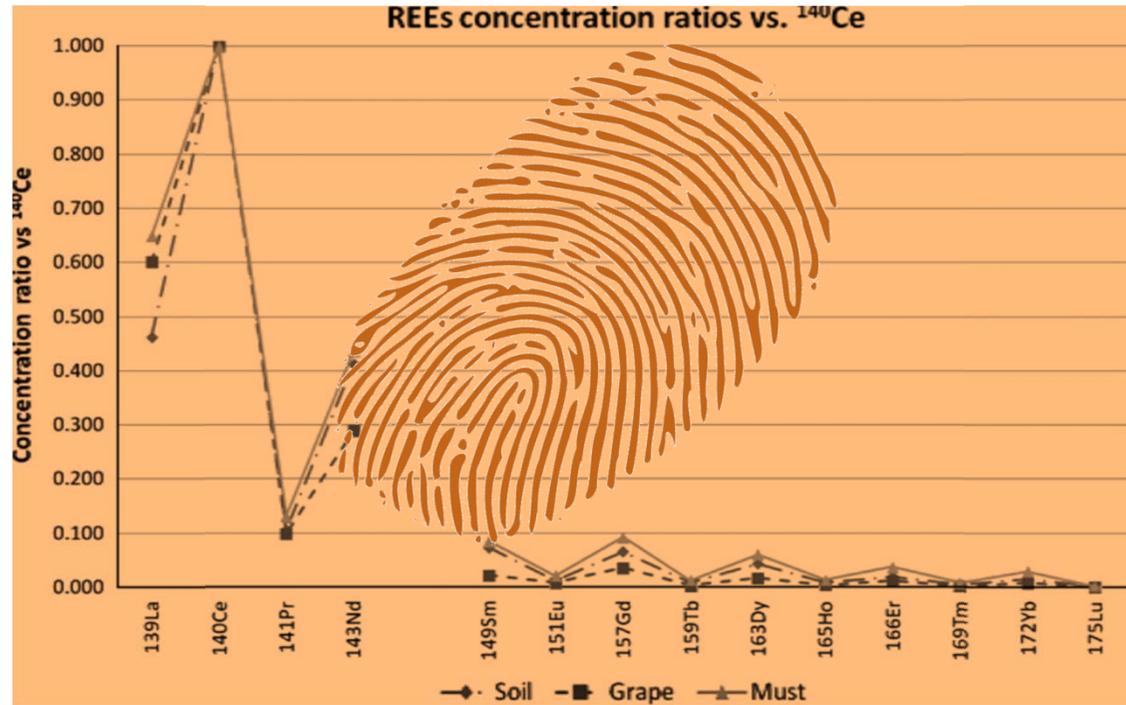
	Al, B, Ba, Li, Mn, Mo, Rb, Si, Sr, Ti	Ca e Mg, Cu, Zn, Fe, K, Na, P	Pb, Co, Cr, Zn, Ni, V
	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	
	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	
		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>

Maggiore potere discriminante

ELEMENTI MINERALI NEL VINO



La distribuzione di concentrazione delle Terre Rare si mantiene inalterata dal suolo al mosto (Aceto et al., 2013)



ottimi marker di provenienza geografica

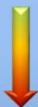
FASI DEL PROGETTO

sito



1. CAMPIONAMENTO

uva



2. MICROVINIFICAZIONE

vino



3. ANALISI ICP-MS

dati



4. ELABORAZIONE DATI

modello



5. VALIDAZIONE

risultati

Il progetto si è articolato in 5 fasi:

- campionamento delle uve
- micro-vinificazione
- analisi ICP-MS dei vini
- elaborazione statistica dei risultati analitici
- validazione del modello di tracciabilità



1. CAMPIONAMENTO



2013

- 4 squadre campionamento
- 16 Regioni
- 41 Province
- 246 Campioni uve



2014

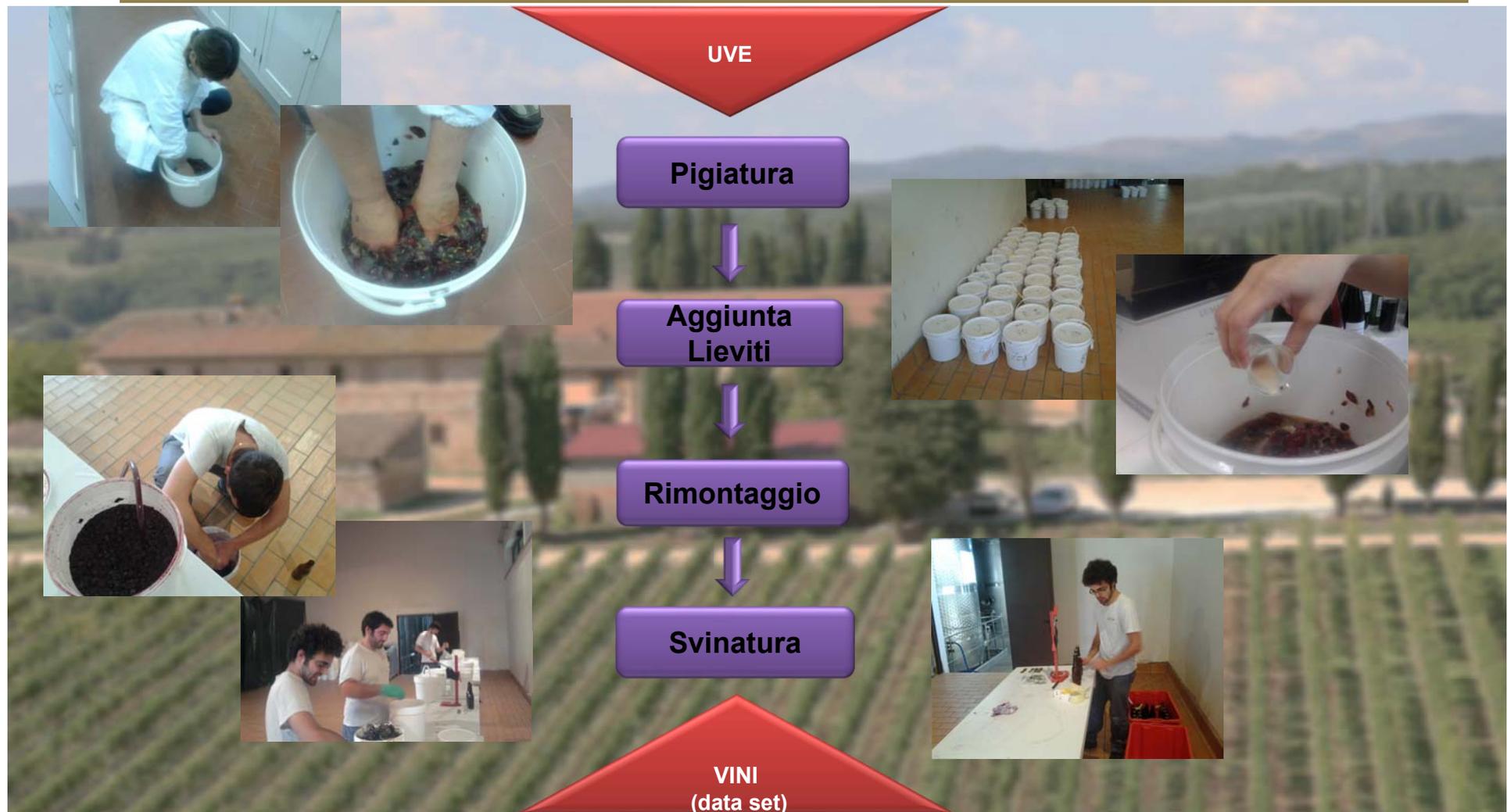
- 4 squadre campionamento
- 17 Regioni
- 47 Province
- 393 Campioni uve



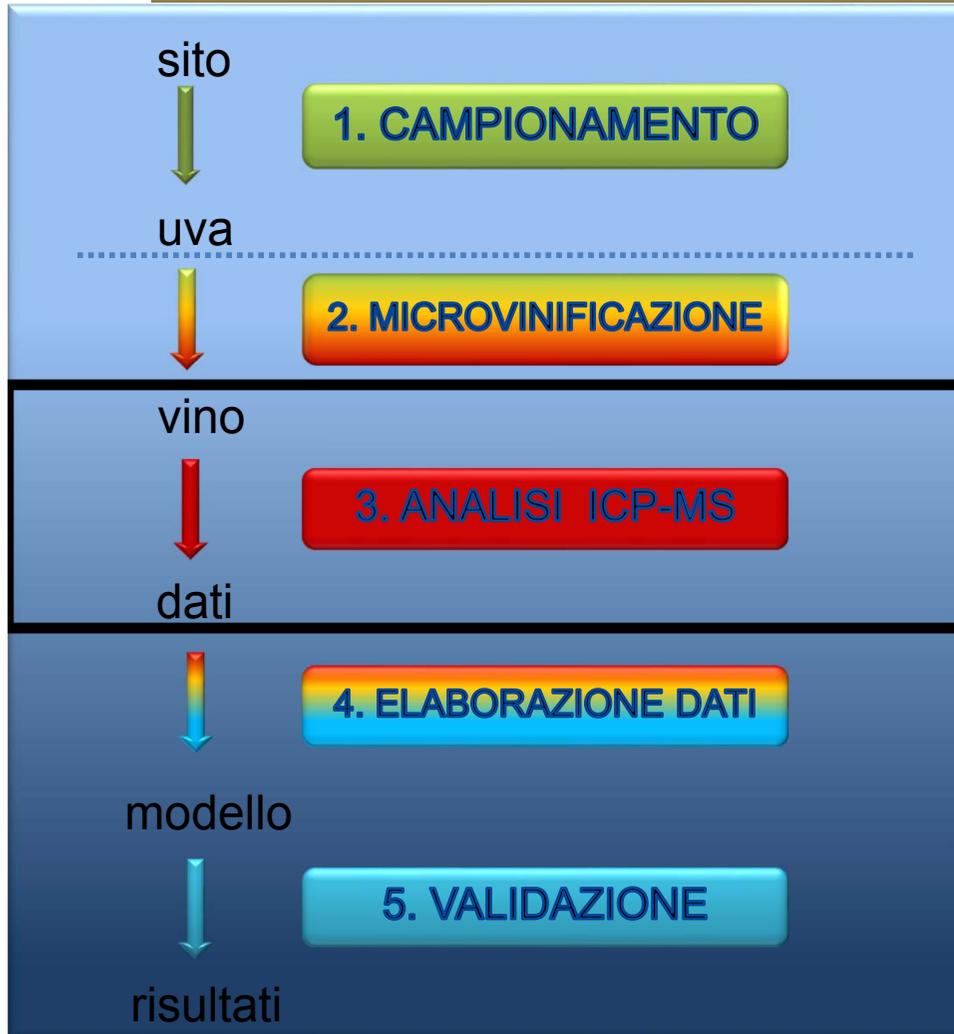
2nd MS Wine Day 9-10 Maggio 2017

RUFFINO
LA VITA RUFFINO

2. MICRO-VINIFICAZIONE

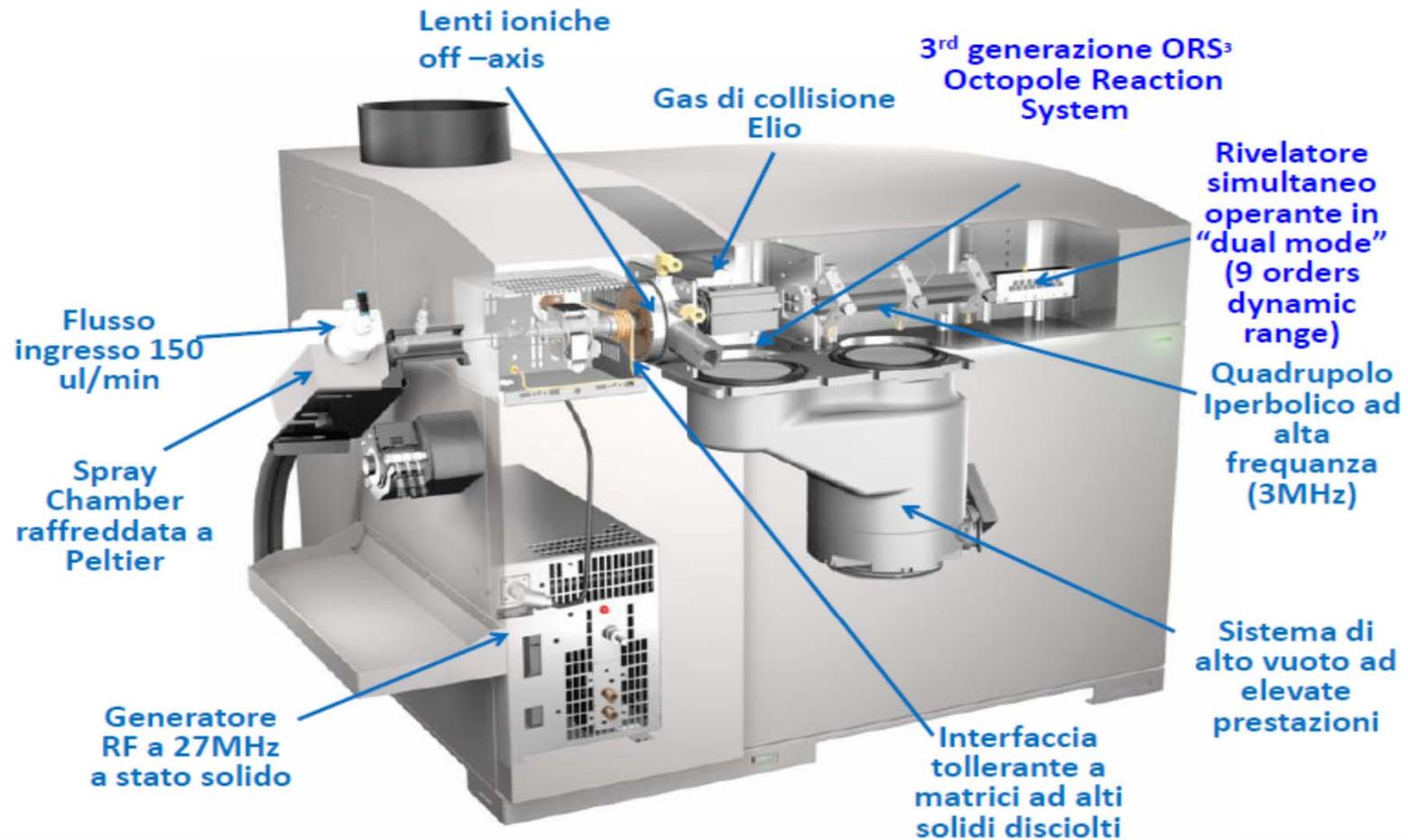


3. FASE ANALITICA



- Strumentazione
- Preparazione del campione
- Metodo analitico

3.a. STRUMENTAZIONE



ICP MS Agilent 7700



2nd MS Wine Day 9-10 Maggio 2017

RUFFINO
LA VITA RUFFINO

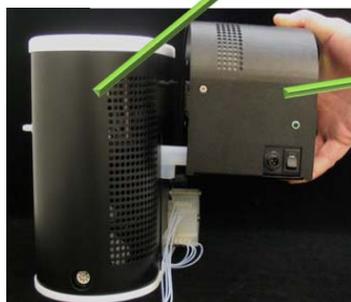
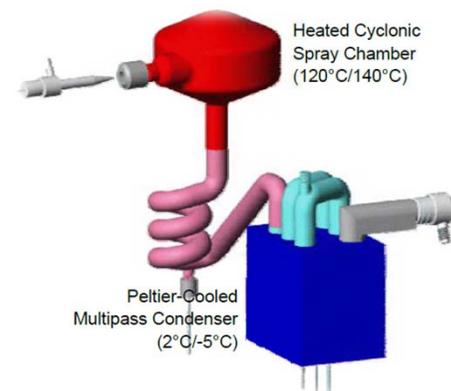
3.a. STRUMENTAZIONE

Sistema di introduzione del campione



APEX Q

nebulizzatore con spray chamber ciclonica riscaldata a 140° C, condensatore e sistema Peltier di raffreddamento a 2° C che consente la desolvatazione del campione.



Spiro TMD (Teflon Membran Desolvation)

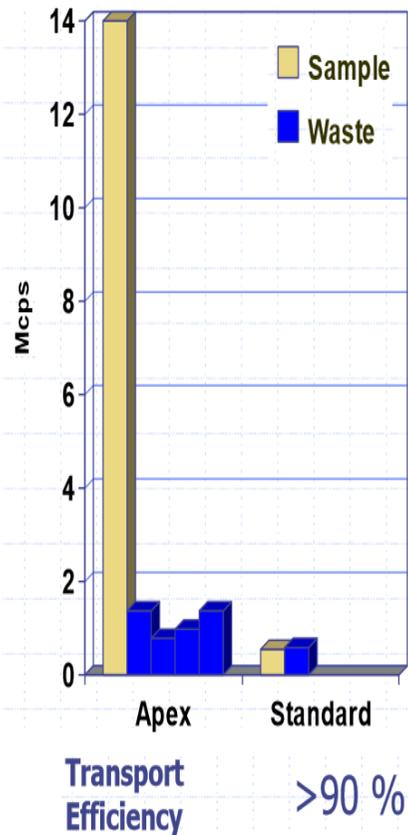
Sistema di desolvatazione con membrana di Teflon macroporosa riscaldata che permette di rimuovere i residui di vapore acqueo dal flusso di aerosol del campione.



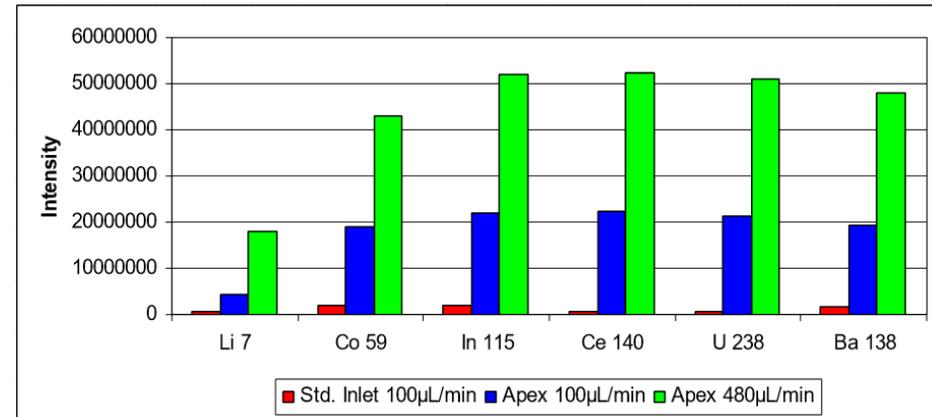
3.a. STRUMENTAZIONE

Sistema di introduzione del campione

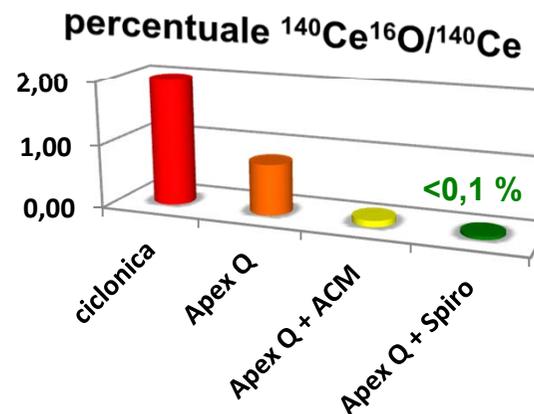
Vantaggi del sistema di desolvatazione:



➤ Preconcentrazione del campione
 ↓
 aumento del segnale da 3 a 10 volte

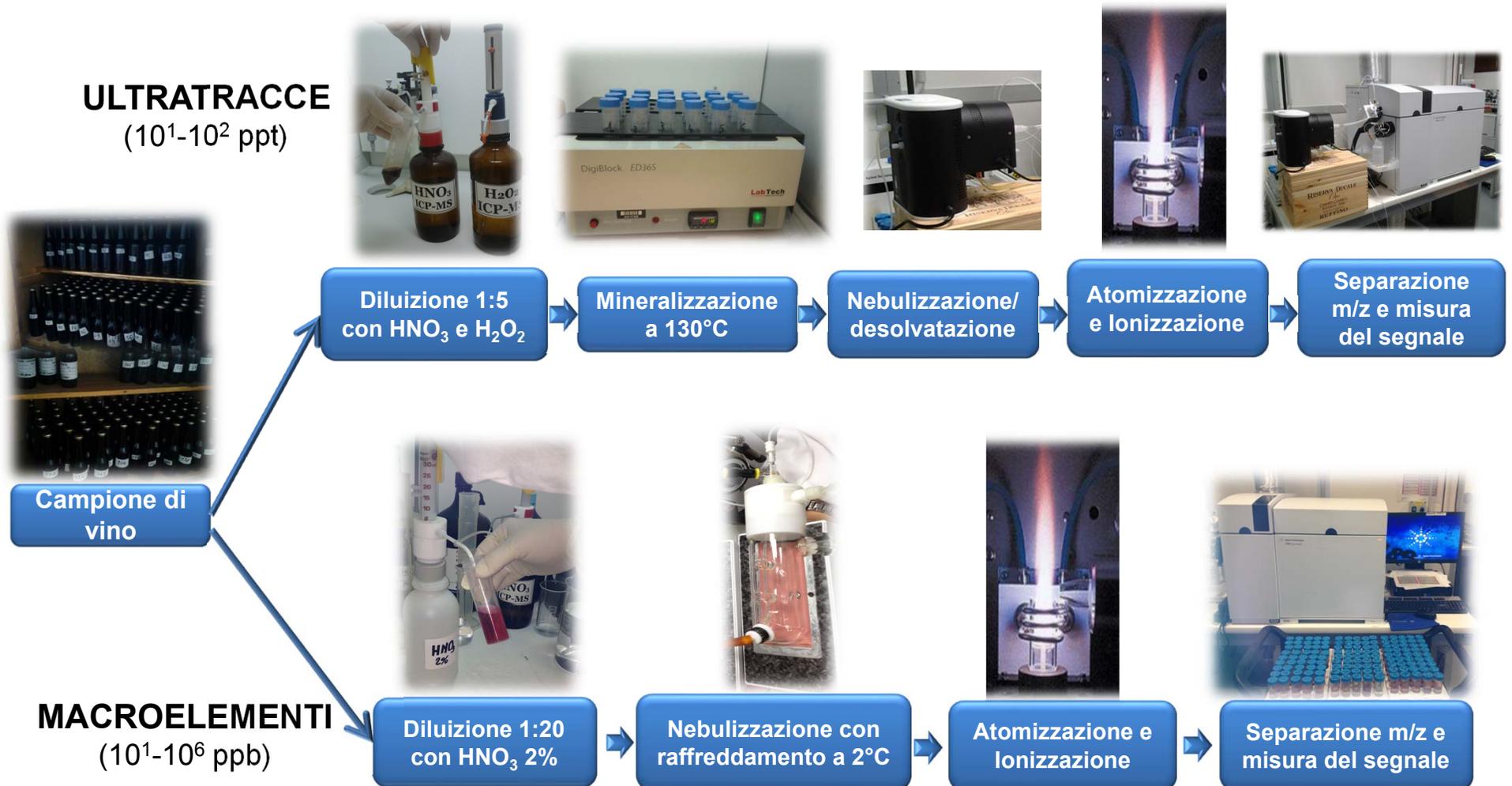


➤ Riduzione della formazione di ossidi
 ↓
 diminuzione delle specie interferenti



Elemento	Interferenza
^{155}Gd	$^{139}\text{La}^{16}\text{O}$
^{156}Gd	$^{140}\text{Ce}^{16}\text{O}$
^{157}Gd	$^{141}\text{Pr}^{16}\text{O}$
^{158}Gd	$^{142}\text{Ce}^{16}\text{O}$, $^{142}\text{Nd}^{16}\text{O}$
^{159}Gd	$^{143}\text{Nd}^{16}\text{O}$

3.b. PREPARAZIONE DEL CAMPIONE



3.c. METODO ANALITICO

Campione	Tipo	Vial
Lavaggio	sample	0
CAL 0	cal level 1	1
CAL 1	cal level 2	2
CAL 1'	cal level 3	3
CAL 1''	cal level 4	4
Lavaggio	sample	0
ICB 1	sample	5
ICB 2	sample	6
ICB 3	sample	7
ICB 4	sample	8
ICV	sample	9
ICV'	sample	10
CCV	QC1	25
CAMP 1	sample	11
CAMP 2	sample	12
CAMP 3	sample	13
CAMP 4	sample	14
CAMP 5	sample	15
CAMP 6	sample	16
CCB	QC2	5
CCV	QC1	25
Batch di circa 30-40 campioni		

UNI EN ISO 17294-2 (2005)

- **Verifica della linearità di risposta:** 5 livelli, criterio di accettabilità: $R^2 > 0.998$
- **Taratura giornaliera:** aggiunte in matrice, 2 punti (livello alto x3)
- **Controlli di accettabilità:**
 - ✓ ICB: \leq MDL
 - ✓ ICV: $\pm 10\%$
 - ✓ CCB: $<$ MDL (effetto memoria)
 - ✓ CCV: $\pm 10\%$ (deriva strumentale)
- **Standard interno:** aggiunto in concentrazione costante a tutte le soluzioni (standard e campioni); criterio di accettabilità sul recupero: $\pm 20\%$

→ x6

Soluzioni standard	
CAL 0	Matrice vino tal quale
CAL 1, 1', 1''	Matrice vino + aggiunta standard
ICB 1 (CCB), 2, 3, 4	Bianco analitico
ICV, ICV'	Matrice vino + aggiunta standard (livello intermedio)
CCV	Matrice vino + aggiunta standard (livello alto)

3.c. METODO ANALITICO

Ripetibilità e riproducibilità

Calcolo della ripetibilità del metodo analitico su 5 campioni di vino analizzati 15 volte (5 letture per 3 batch)



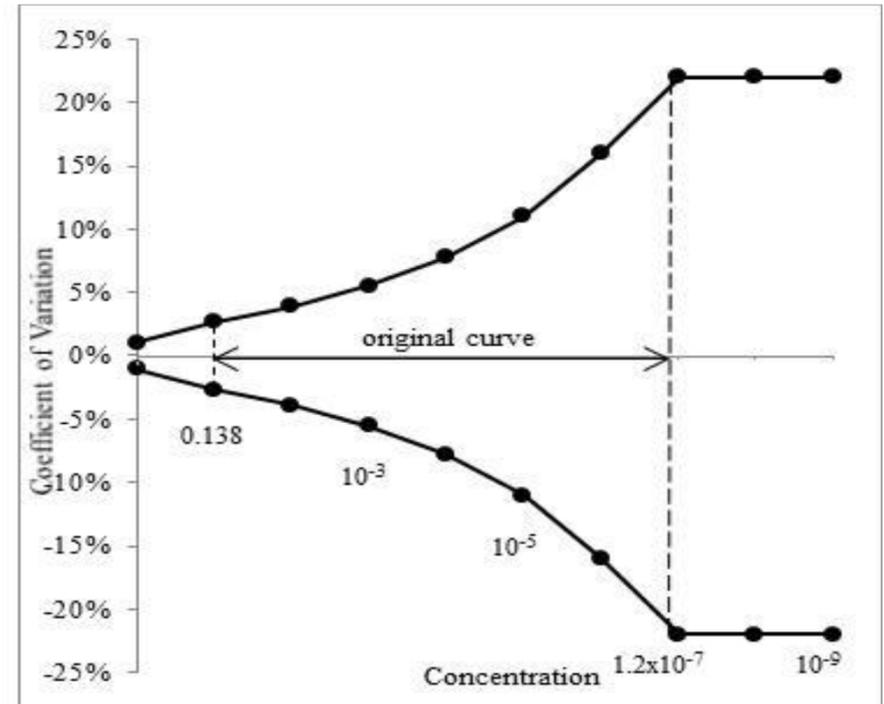
Criterio di accettabilità: $RSD \leq \sigma_{Horwitz-Thomson}$

INCERTEZZA DI MISURAZIONE (EQUAZIONE DI HORWITZ E CORREZIONI DI THOMPSON)

Equazione	Campo di applicazione
$\sigma = 0,22C$	$C < 1,2 * 10^{-7}$
$\sigma = 2^{(1-0,5\log C)}$	$1,2 * 10^{-7} \leq C \leq 1,38 * 10^{-1}$
$\sigma = 0,01\sqrt{C}$	$C > 1,38 * 10^{-1}$

C = concentrazione dell'analita
 σ = deviazione standard di riproducibilità

Nel caso di misurazioni eseguite all'interno del laboratorio (in condizioni di ripetibilità) σ può essere ridotta al 65%.



3.c. METODO ANALITICO

Selezione analiti (macroelementi)

Isotopo	Elemento	u mis	VINO 1			VINO 2			VINO 3			VINO 4			VINO 5		
			CV (%)	R(%)	r(%)												
7	Li	ppb	4,5	22,0	14,3	2,5	22,0	14,3	3,0	22,0	14,3	3,2	22,0	14,3	4,8	22,0	14,3
23	Na	ppm	4,2	10,4	6,8	5,5	11,3	7,4	4,5	10,7	6,9	3,1	10,5	6,8	3,6	10,9	7,1
26	Mg	ppm	2,5	7,8	5,1	2,7	7,6	5,0	2,4	7,7	5,0	1,7	7,6	5,0	2,0	7,7	5,0
48	Ti	ppb	3,0	22,0	14,3	3,3	22,0	14,3	3,0	22,0	14,3	2,4	22,0	14,3	2,6	22,0	14,3
51	V	ppb	8,4	22,0	14,3	20,3	22,0	14,3	22,1	22,0	14,3	21,7	22,0	14,3	16,1	22,0	14,3
55	Mn	ppb	2,3	14,8	9,6	2,3	14,0	9,1	2,2	13,7	8,9	1,5	15,0	9,8	2,4	15,4	10,0
59	Co	ppb	2,9	22,0	14,3	2,4	22,0	14,3	2,9	22,0	14,3	3,1	22,0	14,3	3,3	22,0	14,3
60	Ni	ppb	3,0	22,0	14,3	2,6	22,0	14,3	2,7	22,0	14,3	3,6	22,0	14,3	2,4	22,0	14,3
75	As	ppb	21,9	22,0	14,3	16,8	22,0	14,3	20,2	22,0	14,3	21,2	22,0	14,3	21,1	22,0	14,3
85	Rb	ppb	2,2	14,2	9,2	2,2	14,4	9,4	2,0	14,4	9,4	2,0	14,6	9,5	2,0	14,9	9,7
88	Sr	ppb	1,9	16,3	10,6	2,1	16,4	10,6	1,9	16,4	10,6	1,9	16,6	10,8	7,1	17,7	11,5
135	Ba	ppb	3,1	20,5	13,4	2,3	20,0	13,0	2,5	18,6	12,1	13,8	22,0	14,3	2,8	19,8	12,8

R%: stima riproducibilità % (Horwitz-Thompson)

r%: stima ripetibilità % (Horwitz-Thompson)

3.c. METODO ANALITICO

Selezione analiti (ultratracce)

Isotopo	Elemento	Vino 1			Vino 2			Vino 3			Vino 4			Vino 5		
		CV (%)	R %	r %	CV (%)	R %	r %	CV (%)	R %	r %	CV (%)	R %	r %	CV (%)	R %	r %
89	Y	1,5	22	14,3	3,25	22	14,3	2	22	14,3	0,9	22	14,3	1,7	22	14,3
140	Ce	3	22	14,3	6,4	22	14,3	6,4	22	14,3	4,7	22	14,3	6,5	22	14,3
151	Eu	4,6	22	14,3	5,6	22	14,3	6,1	22	14,3	5,9	22	14,3	8,8	22	14,3
166	Er	6	22	14,3	6,4	22	14,3	10,3	22	14,3	4,8	22	14,3	6,2	22	14,3
172	Yb	5,7	22	14,3	7,4	22	14,3	9,7	22	14,3	7,1	22	14,3	7,6	22	14,3
175	Lu	5,1	22	14,3	7,9	22	14,3	5,8	22	14,3	5,7	22	14,3	6,6	22	14,3
181	Ta	10,9	22	14,3	12,2	22	14,3	9,8	22	14,3	10,6	22	14,3	8,8	22	14,3
185	Re	6,6	22	14,3	7,6	22	14,3	5,5	22	14,3	7,9	22	14,3	6	22	14,3

R%: stima riproducibilità % (Horwitz-Thompson)

r%: stima ripetibilità % (Horwitz-Thompson)

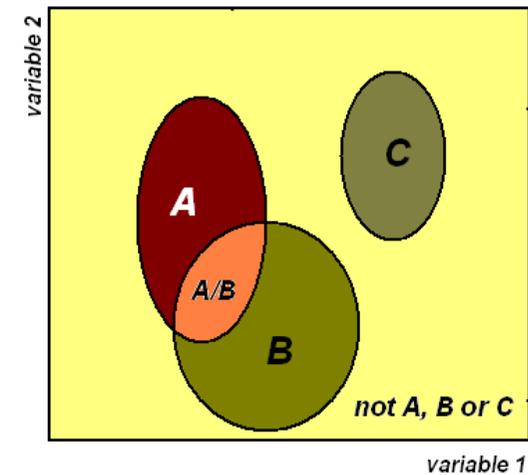
4. ELABORAZIONE DEI RISULTATI

I metodi di classificazione hanno l'obiettivo di costruire, sulla base di un certo numero di variabili indipendenti, un modello capace di individuare la classe cui appartiene ciascun campione. Quindi, i metodi di classificazione cercano una relazione tra le variabili indipendenti ed una risposta qualitativa (la classe).

La calibrazione dei modelli di classificazione è stata effettuata sul training set costituito dai campioni di micro-vinificazione.



Training set



Modello

4. ELABORAZIONE DEI RISULTATI

Pretrattamento dei dati

In fase preliminare sono state applicate diverse tecniche di scalatura e trasformazioni delle variabili al fine di poter enucleare dai dati informazione utile e non banale ed eliminare informazione ovvia (cioè intrinseca nella singola variabile) e non interessante:

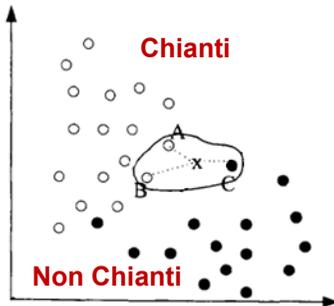
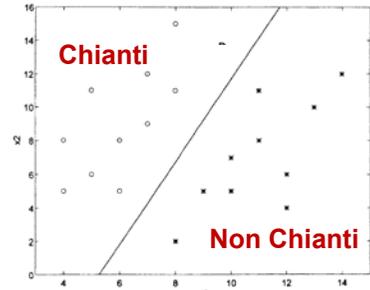
- **rapporti di concentrazione:** sulla base di alcuni studi pubblicati in letteratura e sulla base dei risultati dell'analisi di tutti i possibili fattori di normalizzazione, si è deciso di calcolare i seguenti 23 rapporti di concentrazione, che sono stati aggiunti alle variabili chimiche originali: $^{140}\text{Ce}/\text{La}$, $^{141}\text{Pr}/\text{La}$, $^{144}\text{Nd}/\text{La}$, $^{147}\text{Sm}/\text{La}$, $^{151}\text{Eu}/\text{La}$, $^{152}\text{Sm}/\text{La}$, $^{159}\text{Tb}/\text{La}$, $^{160}\text{Gd}/\text{La}$, $^{163}\text{Dy}/\text{La}$, $^{164}\text{Dy}/\text{La}$, $^{165}\text{Ho}/\text{La}$, $^{166}\text{Er}/\text{La}$, $^{169}\text{Tm}/\text{La}$, $^{172}\text{Yb}/\text{La}$, $^{175}\text{Lu}/\text{La}$, $^{160}\text{Gd}/\text{Er}$, $^{166}\text{Er}/\text{Yb}$, $^{141}\text{Pr}/\text{Lu}$, $^{140}\text{Ce}/\text{Y}$, $^{141}\text{Pr}/\text{Th}$, $^{172}\text{Yb}/\text{Te}$, $^{172}\text{Yb}/\text{Eu}$, $^{125}\text{Te}/\text{Er}$
- **trasformazione logaritmica:** tutte le variabili chimiche, rapporti inclusi, sono state trasformate con il logaritmo allo scopo di rendere le distribuzioni più vicine alla normalità e mitigare l'effetto dei dati anomali.
- **autoscalatura (autoscaling):** questa scalatura è stata applicata ai dati, precedentemente trattati con la trasformazione logaritmica, prima di eseguire la calibrazione del modello PLS-DA.
- **scalatura di intervallo (range scaling):** questa scalatura è stata applicata ai dati, precedentemente trattati con la trasformazione logaritmica, prima di eseguire la calibrazione dei modelli FF_Sin e N3_JT.

4. ELABORAZIONE DEI RISULTATI

Modelli di classificazione

1. Modello Lineare (PLSDA)

Il modello PLSDA calcola delle direzioni (combinazioni lineari delle variabili originali), chiamate Variabili Latenti (LV), che contemporaneamente spiegano la variazione dei dati e separano le classi in analisi. La selezione delle variabili del modello è stata effettuata con la tecnica evolutiva iterativa degli Algoritmi Genetici (GA).



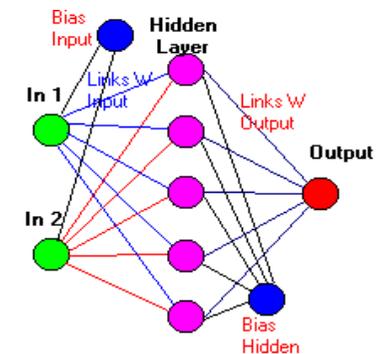
2. Modello Similarità (N3_JT)

Il modello N3_JT si basa su una strategia di similarità locale tra il campione target e i campioni del training set, ispirata dal metodo k-NN. Le similarità tra le coppie di campioni sono state calcolate usando la funzione di distanza media di Jaccard-Tanimoto. Tutti i campioni del training set contribuiscono alla stima della probabilità di appartenenza alla classe Chianti con un contributo che è inversamente proporzionale alla loro similarità con il target.

3. Modello Non Lineare (FF_Sn)

Le reti neurali artificiali (ANNs) sono algoritmi di elaborazione di dati in grado di scoprire dinamicamente quali regole “sfumate” descrivono più efficacemente l’interdipendenza tra i dati stessi. Man mano che nuovi dati vengono aggiunti, le ANNs aggiustano le proprie regole di conseguenza, senza bisogno di alcun intervento esterno.

Il modello FF_Sn è stato sviluppato con l’algoritmo Sine Net, una famiglia di reti neurali artificiali in cui ciascun nodo opera una trasformazione non-lineare (tangente iperbolica) della trasformazione non-lineare degli input (somma degli input pesati e trasformati con la funzione sinusoidale).

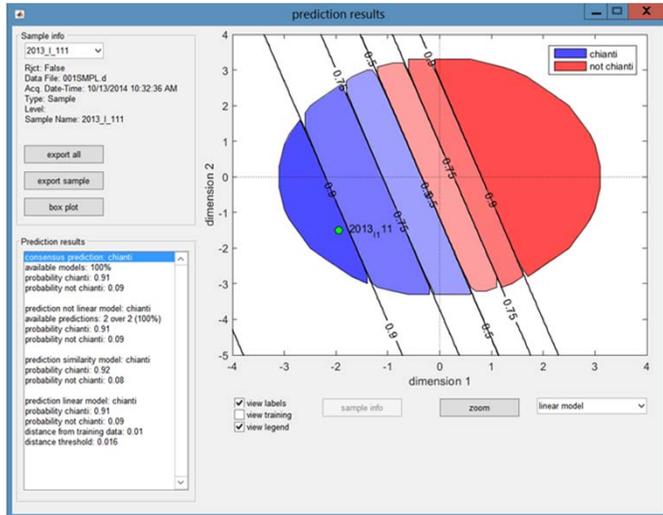


4. Modello Consensus

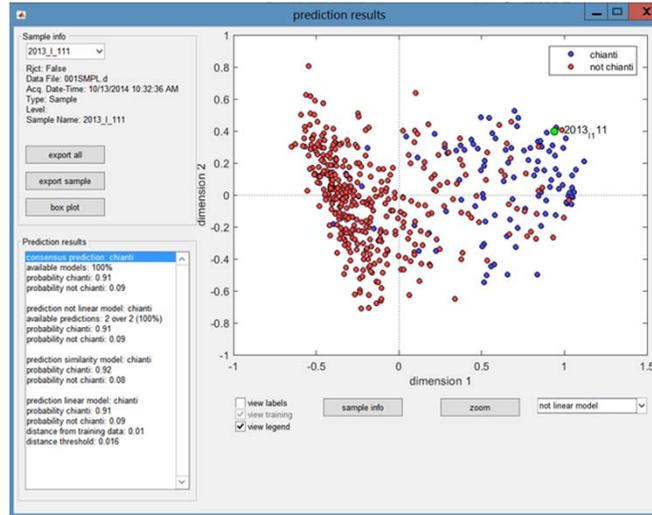
Il modello Consensus utilizza i valori della probabilità di appartenenza alla classe Chianti predetti dai modelli PLSDA, N3_JT e FF_Sn e ne calcola la media aritmetica.

4. ELABORAZIONE DEI RISULTATI Software

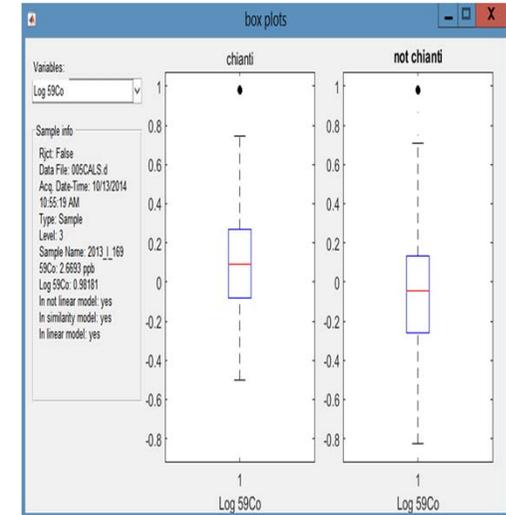
report grafico Modello Lineare



report grafico Modello Similarità



Box Plots



Report per esportazione Excel

Rjct	Data File	Acq. Date-Time	Type	Level	Sample Name	consensus prediction	available models	probability chianti	probability not chianti	prediction not-linear model	available predictions	probability chianti	probability not chianti	prediction similarity model	probability chianti	probability not chianti	prediction linear model	probability chianti	probability not chianti	distance from training data	distance threshold	prediction against marche	probability chianti	prediction against sicilia	probability chianti
False	001SMPL.d	10/13/2014 10:32:36 AM	Sample		2013_1_111	chianti	100%	0.91	0.09	chianti	2 over 2 (100%)	0.91	0.09	chianti	0.92	0.08	chianti	0.91	0.09	0.01	0.016	chianti	0.59	chianti	1
False	002SMPL.d	10/13/2014 10:37:47 AM	Sample		2013_1_162	chianti	100%	0.92	0.08	chianti	2 over 2 (100%)	0.96	0.04	chianti	0.9	0.1	chianti	0.91	0.09	0.009	0.016	chianti	0.77	chianti	1
False	003CALB.d	10/13/2014 10:42:50 AM	Sample		12013_1_165	chianti	100%	0.68	0.32	chianti	2 over 2 (100%)	0.64	0.36	chianti	0.53	0.47	unreliable	0.85	0.15	0.005	0.016	not chianti	0.26	chianti	1
False	004CALB.d	10/13/2014 10:50:00 AM	Sample		22013_1_166	chianti	100%	0.86	0.14	chianti	2 over 2 (100%)	0.78	0.22	chianti	0.89	0.11	chianti	0.91	0.09	0.01	0.016	chianti	0.77	chianti	1
False	005CALB.d	10/13/2014 10:55:19 AM	Sample		32013_1_169	not chianti	50%	0.14	0.86	not chianti	1 over 2 (50%)	0	1	not chianti	0.2	0.8	unreliable	unavailable	unavailable	0.02	0.016	chianti	0.77	unreliable	unavailable
False	006CALB.d	10/13/2014 11:00:37 AM	Sample		42014_1_002	chianti	100%	0.85	0.15	chianti	2 over 2 (100%)	0.76	0.24	chianti	0.89	0.11	chianti	0.91	0.09	0.01	0.016	chianti	0.99	chianti	0.99
False	007CALB.d	10/13/2014 11:06:01 AM	Sample		52014_1_003	chianti	67%	0.62	0.38	not chianti	2 over 2 (100%)	0.48	0.52	chianti	0.75	0.25	unreliable	unavailable	unavailable	0.018	0.016	chianti	0.99	chianti	0.99
False	003CALB.d	10/13/2014 10:42:50 AM	Sample		12014_1_004	chianti	100%	0.67	0.33	not chianti	2 over 2 (100%)	0.27	0.73	chianti	0.86	0.14	chianti	0.88	0.12	0.006	0.016	chianti	0.98	chianti	1
False	004CALB.d	10/13/2014 10:50:00 AM	Sample		22014_1_007	chianti	100%	0.65	0.35	not chianti	2 over 2 (100%)	0.36	0.64	chianti	0.92	0.08	chianti	0.65	0.35	0.001	0.016	chianti	0.68	chianti	0.98
False	005CALB.d	10/13/2014 10:55:19 AM	Sample		32014_1_167	not chianti	100%	0.33	0.67	not chianti	2 over 2 (100%)	0.06	0.94	not chianti	0.49	0.51	not chianti	0.44	0.56	0	0.016	not chianti	0.27	chianti	0.63
False	003CALB.d	10/13/2014 10:42:50 AM	Sample		12014_1_168	chianti	100%	0.74	0.26	chianti	2 over 2 (100%)	0.59	0.41	chianti	0.87	0.13	chianti	0.75	0.25	0.004	0.016	chianti	0.71	chianti	0.89
False	003CALB.d	10/13/2014 10:42:50 AM	Sample		12014_1_229	chianti	100%	0.86	0.14	chianti	2 over 2 (100%)	0.85	0.15	chianti	0.92	0.08	chianti	0.81	0.19	0.003	0.016	chianti	0.96	chianti	1
False	004CALB.d	10/13/2014 10:50:00 AM	Sample		22014_1_231	chianti	100%	0.91	0.09	chianti	2 over 2 (100%)	0.93	0.07	chianti	0.93	0.07	chianti	0.87	0.13	0.008	0.016	chianti	0.98	chianti	1
False	005CALB.d	10/13/2014 10:55:19 AM	Sample		32014_1_232	chianti	100%	0.88	0.12	chianti	2 over 2 (100%)	0.91	0.09	chianti	0.93	0.07	chianti	0.8	0.2	0.003	0.016	chianti	0.88	chianti	1
False	006CALB.d	10/13/2014 11:00:37 AM	Sample		42014_1_236	chianti	100%	0.83	0.17	chianti	2 over 2 (100%)	0.82	0.18	chianti	0.9	0.1	chianti	0.76	0.24	0.003	0.016	chianti	0.93	chianti	0.99
False	007CALB.d	10/13/2014 11:06:01 AM	Sample		52014_1_240	chianti	100%	0.7	0.3	chianti	2 over 2 (100%)	0.56	0.44	chianti	0.67	0.33	chianti	0.86	0.14	0.009	0.016	chianti	0.94	chianti	0.94

2nd MS Wine Day 9-10 Maggio 2017

RUFFINO
LA VITA RUFFINO

5. VALIDAZIONE DEI MODELLI

Analisi di 134 campioni microvinificati di provenienza certa (vendemmie 2013 e 2014) precedentemente esclusi dal training set.

Sono state considerate come affidabili predizioni con probabilità di assegnazione superiore al 60% (intervallo 40%-60% da considerarsi come non assegnato).

consensus	chianti	not chianti	not assigned
chianti	13	1	7
not chianti	10	85	18

similarity	chianti	not chianti	not assigned
chianti	14	3	4
not chianti	12	78	23

not linear	chianti	not chianti	not assigned
chianti	7	4	10
not chianti	10	96	7

linear	chianti	not chianti	not assigned
chianti	11	1	9
not chianti	9	71	33



	NER	Sn Chianti	Sn Not Chianti	% not assigned
consensus	90%	93%	89%	19%
similarity	86%	82%	87%	20%
not linear	88%	64%	91%	13%
linear	89%	92%	89%	31%

I primi risultati sono molto incoraggianti: la percentuale delle assegnazioni corrette del modello Consensus è intorno al 90%, mentre i campioni non assegnati si attestano intorno al 20%.

5. VALIDAZIONE DEI MODELLI

Assegnazioni dei 134 campioni del *validation set* divisi per regioni di appartenenza:

	Totale	Corrette	Sbagliate	Non Assegnate
Abruzzo	8	5	0	3
Basilicata	8	8	0	0
Calabria	2	1	0	1
Campania	9	7	1	1
Emilia romagna	3	0	1	2
Friuli venezia giulia	1	0	1	0
Lazio	2	2	0	0
Lombardia	6	3	1	2
Marche	9	8	1	0
Piemonte	10	8	0	2
Puglia	9	9	0	0
Sardegna	16	16	0	0
Sicilia	14	7	3	4
Toscana Chianti	21	13	3	5
Toscana non Chianti	3	2	0	1
Trentino alto adige	2	1	1	0
Umbria	10	8	1	1
Veneto	1	1	0	0
Totale	134	99	13	22

Aspetti positivi:

- campioni provenienti da regioni con alta produzione vinicola (Sardegna, Piemonte, Puglia, Marche, Campania, Basilicata, Umbria) classificati correttamente
- totale sovrapposizione delle vendemmie 2013 e 2014 (previa normalizzazione dei dati): i modelli non risentono della variabilità relativa alle diverse annate.

Note di cautela:

- alcune regioni (es. Sicilia) presentano una particolare difficoltà nell'essere differenziate dalla zona Chianti.

5. VALIDAZIONE DEI MODELLI

Tagli



Taglio di Chianti con vino Pugliese

Provenienza	Taglio	consensus prediction	probability chianti	probability not chianti
Toscana	Toscana 100%	Chianti	84%	16%
Puglia	Puglia 100%	not Chianti	10%	90%
Taglio Puglia	Toscana 25% + Puglia 75%	not Chianti	17%	83%
Taglio Puglia	Toscana 50% + Puglia 50%	not Chianti	27%	73%
Taglio Puglia	Toscana 75% + Puglia 25%	not Chianti	46%	54%



Taglio di Chianti con vino Siciliano

Provenienza	Taglio	consensus prediction	probability chianti	probability not chianti
Toscana	Toscana 100%	Chianti	84%	16%
Sicilia	Sicilia 100%	Non Chianti	13%	87%
Taglio Sicilia	Toscana 25% + Sicilia 75%	Non Chianti	18%	82%
Taglio Sicilia	Toscana 50% + Sicilia 50%	Non Chianti	29%	71%
Taglio Sicilia	Toscana 75% + Sicilia 25%	Chianti	59%	41%

Intervallo 40-60% non attendibile

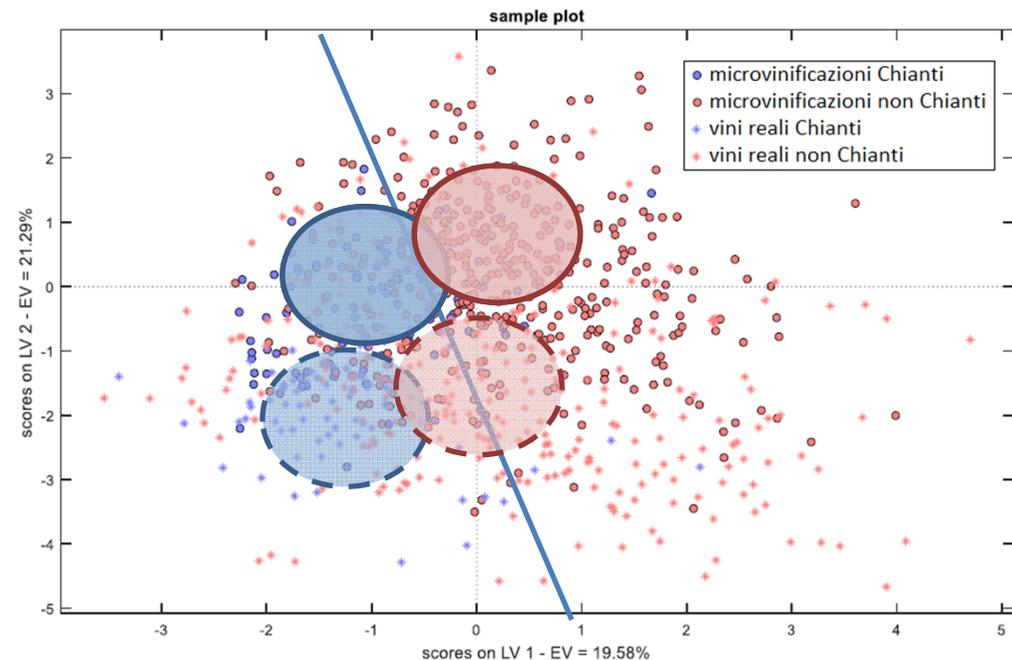
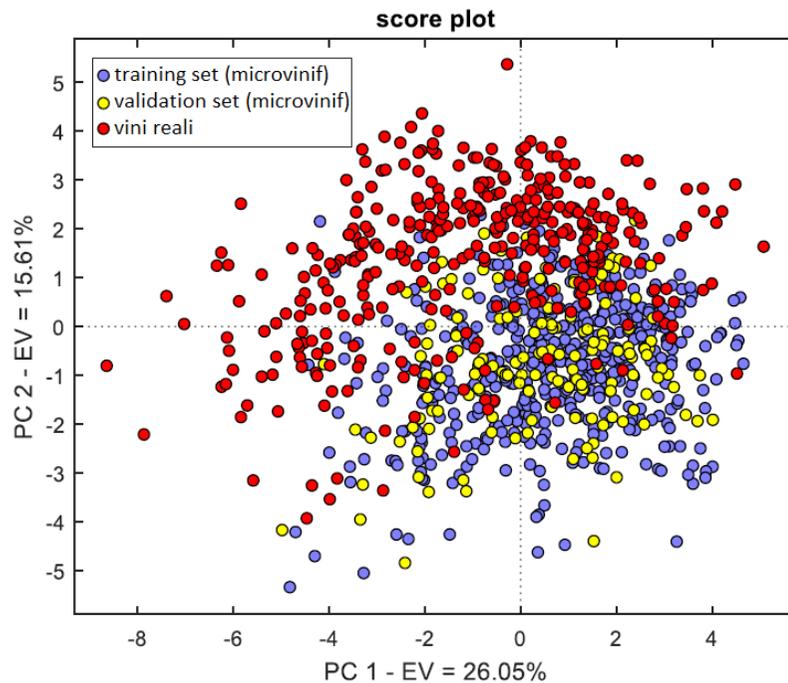


5. VALIDAZIONE DEI MODELLI

Vini reali

Un'ulteriore valutazione delle prestazioni dei modelli di classificazione è stata effettuata analizzando 364 campioni di vini reali di provenienza certa (annate 2014, 2015 e 2016).

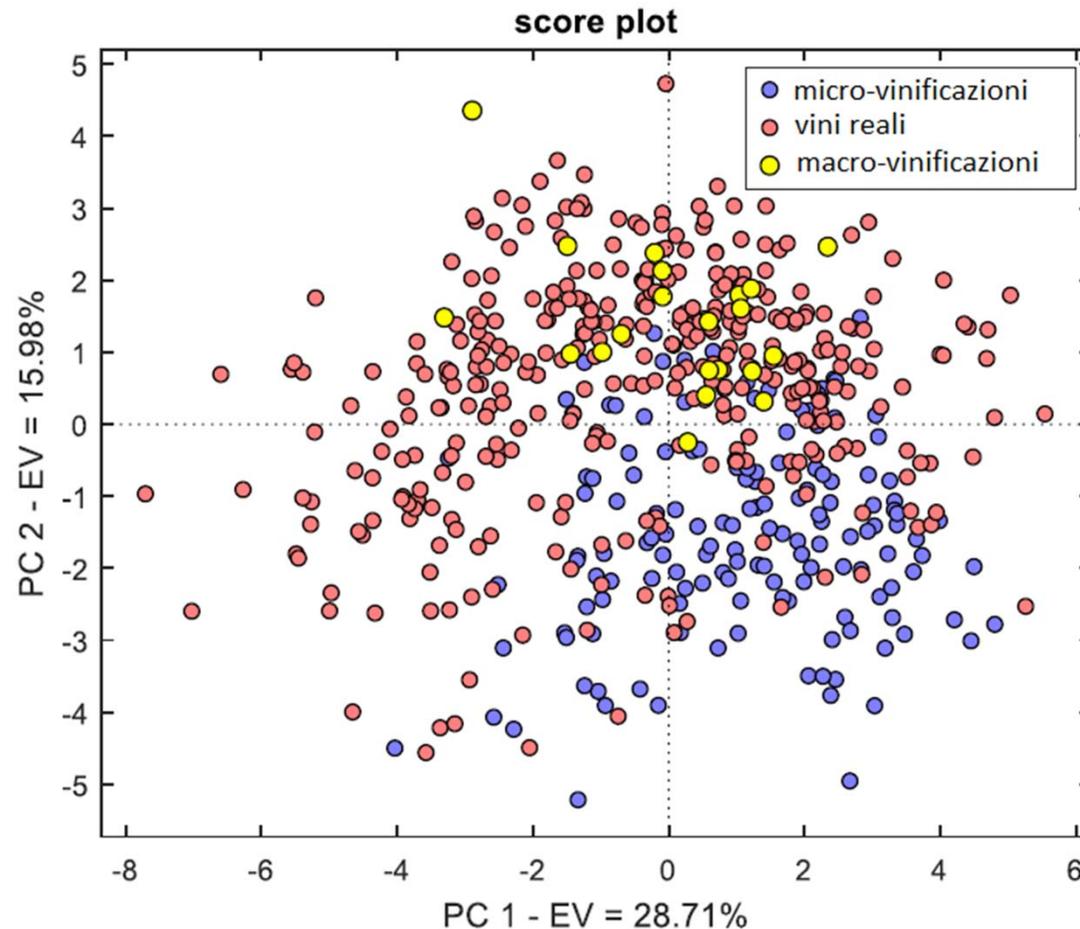
	NER	Sn Chianti	Sn Not Chianti	% not assigned
consensus	73%	86%	59%	-
similarity	67%	92%	43%	-
not linear	69%	60%	79%	9%
linear	78%	96%	59%	44%



5. VALIDAZIONE DEI MODELLI

Vini reali

Differente classificazione fra micro-vinificazione e macro-vinificazione delle stesse uve:



I campioni di macro-vinificazione sono più simili ai vini reali che ai corrispondenti campioni micro-vinificati.



Necessarie strategie per ridurre lo shift nella distribuzione elementare fra i diversi metodi di vinificazione

SVILUPPI FUTURI

- Normalizzazione rispetto ad un elemento del profilo elementale di campioni micro-vinificati e di vini reali
- Aggiornamento/ricalibrazione dei modelli di classificazione con i dati dei 134 campioni di provenienza certa precedentemente esclusi e/o dei vini reali usati per la validazione;
- Sviluppo di un metodo di screening rapido, basato solo sui macro-elementi.
- Ricerca di un modello di classificazione per vini bianchi (target: Prosecco e Pinot Grigio)



GRAZIE PER L'ATTENZIONE

Claudio Brilli (claudio.brilli@ruffino.it)

2nd MS Wine Day 9-10 Maggio 2017

RUFFINO
LA VITA RUFFINO