



Paolo Conti^a, Silvia Zamponi^a, David Ranganathan^a,
Mario Berrettoni^b, Marco Giorgetti^b,
Fabio Bruzzechesse^c, Mirko Marangoni^c

^aScuola di Scienze e Tecnologie

Sezione Chimica

Università di Camerino

^bDipartimento di Chimica Fisica ed Inorganica

Università di Bologna

e Unità di Ricerca INSTM di Bologna

^cAnalisi Control Srl

Corridonia (MC)

paolo.conti@unicam.it

TIPIZZARE LA LENTICCHIA CON XRF

Nel presente lavoro si è sviluppata una potenziale tecnica analitica per la caratterizzazione territoriale della lenticchia IGP di Castelluccio di Norcia. Si sono raccolti dei campioni di semi di lenticchia della stessa varietà, direttamente presso i coltivatori, in località ubicate entro i 50 km da Castelluccio di Norcia, cioè in zone, ancorché limitrofe, non comprese nel disciplinare IGP. I campioni sono stati macinati e la farina ottenuta, pressata in pasticche, è stata analizzata sotto vuoto in ED-XRF. Gli spettri acquisiti sono stati elaborati con varie tecniche di classificazione. Le elaborazioni hanno consentito di definire una modalità di elaborazione che consente di distinguere i campioni di Castelluccio di Norcia e Norcia da tutti gli altri. Un po' meno definita è risultata la separazione dei campioni provenienti da Norcia e Castelluccio, comunque superiore al 75% utilizzando la tecnica LDA.

I commercio di prodotti agroalimentari di alta qualità si basa su beni ben identificati e ben caratterizzati dal punto di vista organolettico. L'autenticità del prodotto, ossia la determinazione delle caratteristiche specifiche, ne determina la qualità e quindi il valore commerciale. Il riconoscimento delle caratteristiche del cibo è operazione di routine per assaggiatori professionisti addestrati, tuttavia le anzidette caratteristiche dei prodotti alimentari non sono facilmente misurabili con le più comuni metodologie analitiche. Sarebbe, tuttavia, desiderabile poter disporre di metodiche analitiche capaci di sostituire i gruppi di assaggiatori professionisti addestrati (panel test) che richiedono alti

costi di addestramento ed esercizio e, per la loro stessa natura, forniscono risultati non facilmente confrontabili e ripetibili. Le caratteristiche organolettiche e nutrizionali di un prodotto sono il frutto della specifica combinazione delle sostanze chimiche in esso presenti; solo eccezionalmente è possibile stabilire una relazione diretta fra una caratteristica sensoriale e la presenza o concentrazione di una specifica sostanza. Le caratteristiche sensoriali di un prodotto agroalimentare sono generalmente legate alla specie vegetale [1] ed alle sue declinazioni varietali [2, 3], alle condizioni meteorologiche di coltura [4], alle tecniche agronomiche, ai trattamenti di conservazione/trasformazione sub-

* L'articolo è stato presentato al convegno ISA 2010, Incontro di Spettroscopia Analitica, Rimini, 16-18 giugno 2010

iti [5, 6], altresì dipendono anche dall'origine geografica della materia prima. Identificare l'origine geografica significa stabilire una combinazione di fattori microclimatici e pedologici che influenzano lo sviluppo vegetativo della pianta e quindi la sua produzione di sostanze chimiche in proporzioni caratteristiche; inoltre l'identificazione geografica è necessaria per fregiarsi dei loghi IGP. La capacità di discriminare [7], in modo semplice e veloce, i prodotti sulla base della loro provenienza è essenziale nella protezione contro le frodi commerciali. Sia la spettroscopia di fluorescenza ED-XRF [8-12] che la determinazione della presenza di metalli [13-15] sono stati utilizzati per questo scopo. In questo studio abbiamo utilizzato la spettroscopia ED-XRF accoppiata all'elaborazione chemiometrica degli spettri per tentare di caratterizzare le lenticchie coltivate in un'area ristretta sull'altopiano di Castelluccio di Norcia (area del disciplinare IGP) rispetto a quelle coltivate in zone limitrofe.

Campionamento e preparazione del campione

I campioni per lo studio sono stati prelevati direttamente presso i produttori nell'autunno del 2007 e riguardano il raccolto estivo dello stesso anno. Le lenticchie analizzate sono tutte della stessa varietà di quella seminata a Castelluccio di Norcia con l'eccezione del campione proveniente da Ustica. Nella mappa di Fig. 1 sono indicate le località di campionamento, Ustica esclusa, e, tra parentesi, la numerosità dei campioni. Tutte le località di campionamento insistono nella stessa area geografica entro 50 km dall'altopiano di Castelluccio. Il campione proveniente dall'isola di Ustica è stato analizzato come campione test per verificare la "sensibilità" delle tecniche utilizzate.

I singoli campioni sono stati così trattati: 60 g di semi di lenticchia, senza pretrattamenti, sono stati macinati con un macinacaffè. La farina è stata setacciata con setaccio Analysensieb da 60 mesh; pressando la farina, così separata, a 10 t/cm² con una pressa idraulica; la pasticca così preparata è stata analizzata in fluorescenza X con uno strumento ED-XRF Shimadzu 800HS2, con collimatore di 10 mm, in vuoto. La fluorescenza X degli elementi dal sodio all'uranio è stata registrata in due spettri separati di 2048 punti ciascuno; uno spettro è acquisito, in condizioni ottimali, per coprire dal sodio allo scandio (Na-Sc) e l'altro dal titanio all'uranio (Ti-U). Entrambi gli spettri ED-XRF acquisiti sono stati utilizzati nelle successive elaborazioni.

Trattamento dati

Si analizzano solo le parti degli spettri con conteggi superiori alla linea di base; in particolare per gli analiti Na-Sc si considerano le letture da 0,11 a 15,22 keV (1512 punti) e per gli analiti Ti-U si prendono in

considerazione le letture fra 0,22 e 37,50 keV (1864 punti). Gli spettri così ridotti vengono normalizzati con due diversi trattamenti [16, 17]: a) gli spettri sono separatamente sottoposti al pretrattamento di Standard Normal Variate (SNV) seguito da autoscaling sulle colonne; b) gli spettri sono stati solo autoscalati sulle colonne.

La matrice dati è costituita dagli spettri relativi ad un campione disposti su di una riga, vale a dire lo spettro ottenuto per Na-Sc precede sulla riga quello di Ti-U dello stesso campione. Ogni riga descrive un campione, ogni colonna contiene i conteggi di tutti i campioni all'energia corrispondente.

La prima elaborazione effettuata è stata l'analisi dei componenti principali (PCA) [16] dopo pretrattamento con il metodo b). L'autoscaling attenua le differenze fra gli spettri sullo stesso campione dovute alle diverse condizioni di misura. Lo score plot di Fig. 2 mostra come, pur con un semplice trattamento, gli spettri XRF siano ottimi candidati per l'identificazione territoriale della lenticchia. In questa figura spicca la separazione dei campioni di Arcofiato e Collevocchio da quelli di Norcia, Castelluccio e Cupi; nel primo gruppo infatti sono raccolti i campioni coltivati su terreni più argillosi e con meno scheletro che nelle altre zone.

Si è provato poi ad usare PLS2 per classificare i dati (PLS-LDA) [17]. Le elaborazioni sono eseguite con The Unscrambler ver. 9, creando

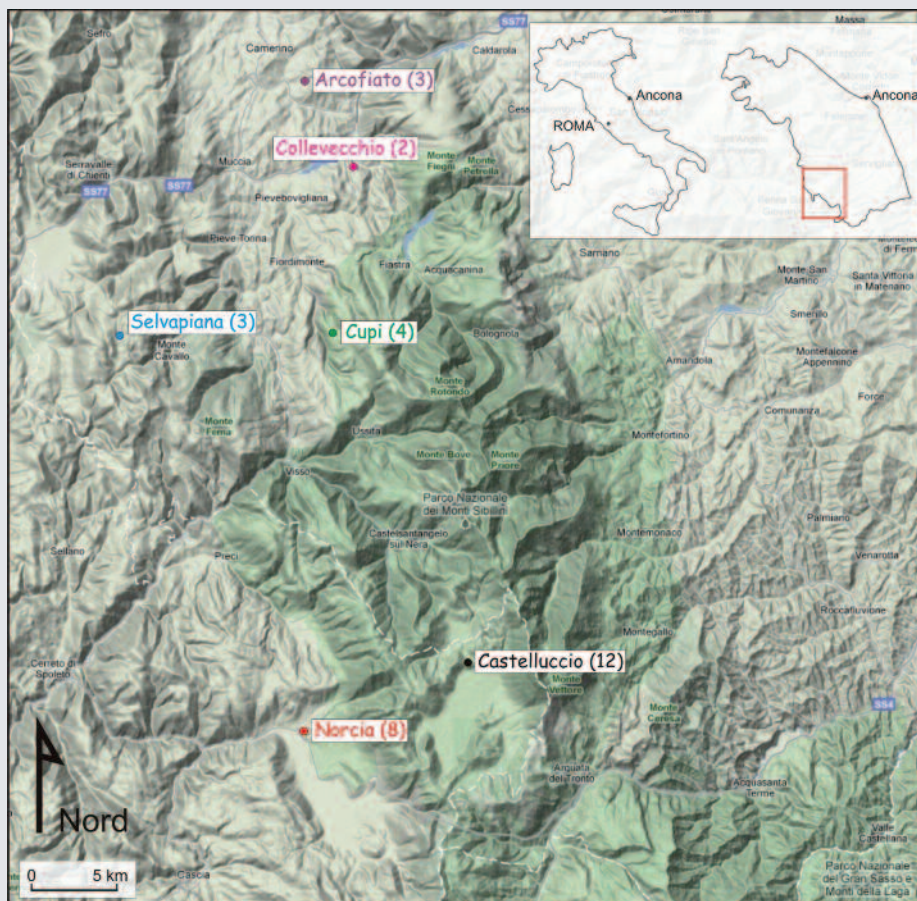


Fig. 1 - Mappa con evidenziate le località di campionamento, tra parentesi il numero di campioni. Le aree verdi delimitano i confini del parco nazionale dei monti Sibillini e del parco nazionale dei monti della Laga

prima delle variabili fittizie per le diverse provenienze. Per ogni categoria è definita una variabile fittizia codificata: 0 assenza, 1 presenza. Poiché la numerosità, soprattutto per alcune campionature, è estremamente ridotta, si è cercato di individuare solo la capacità di discriminazione rispetto a Castelluccio. Per questo motivo si sono considerati solo tre gruppi (categorie): campioni di Castelluccio (■), campioni di Norcia (●) ed una categoria Altri (☆ vari colori) che raccoglie tutte le altre provenienze. L'elaborazione consente, da un punto di vista grafico (vedi Fig. 3), di discriminare distintamente le tre categorie. In predizione il metodo è capace solo di distinguere campioni di Castelluccio o Norcia da tutti gli altri.

Di seguito parleremo di uso dei modelli in classificazione ed in predizione. Queste espressioni sono correlate alla tecnica di cross-validation [16] con cui abbiamo valutato la capacità del modello, calcolato dai dati, di predire il comportamento di un campione incognito. In breve possiamo dire che la capacità di "classificazione" esprime quanto il modello calcolato sia capace di descrivere i campioni noti, mentre la "predizione" è una stima della capacità di inferenza del modello con campioni la cui provenienza è incognita.

La capacità sia di classificazione sia di predizione migliora, pur restando la difficoltà a distinguere Castelluccio da Norcia, quando si applica il pretrattamento di SNV più autoscaling di colonna; risultati simili si ottengono con il solo SNV. La Fig. 4 mostra la capacità di raggruppamento di questa regressione, in essa sono ben distinti i campioni di Castelluccio, Altri e Norcia sia fra loro che rispetto agli altri, ma in predizione non si hanno risultati altrettanto chiari. L'elevato numero di variabili in confronto al numero di campioni usati può dar luogo facilmente alla rappresentazione, nel modello, di una parte della variabilità dei dati che non ha significato ai fini della soluzione del nostro problema. Una riduzione delle variabili usate nell'elaborazione, escludendo quelle meno significative in rapporto al problema, è auspicabile per ridurre il rischio predetto e migliorare la capacità di classificazione/predizione delle elaborazioni. Sulla base di queste considerazioni ed usando i coefficienti

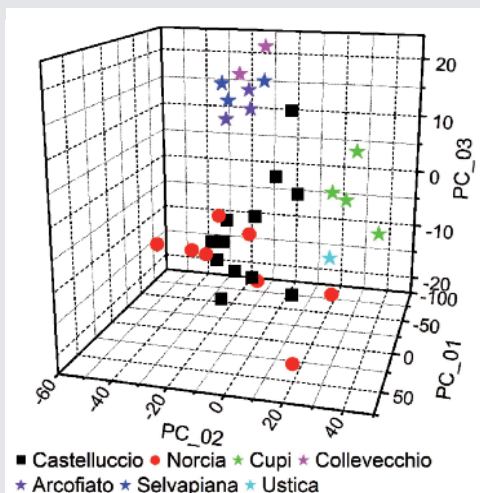


Fig. 2 - Score plot (primi tre componenti principali) dell'analisi PCA sulla matrice dei dati dopo autoscaling

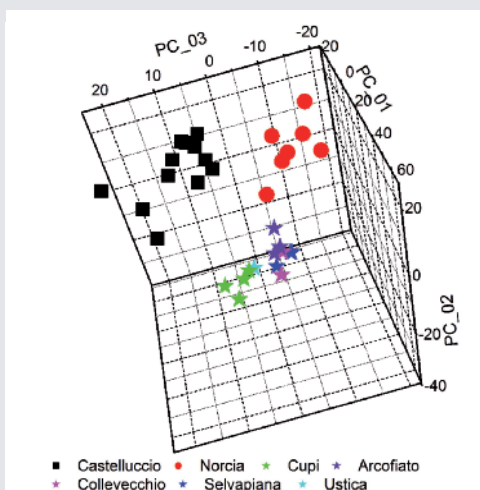


Fig. 3 - Rappresentazione dei campioni sulle prime tre variabili latenti (39% della varianza degli spettri ritenuta) della PLS-DA su dati autoscalati

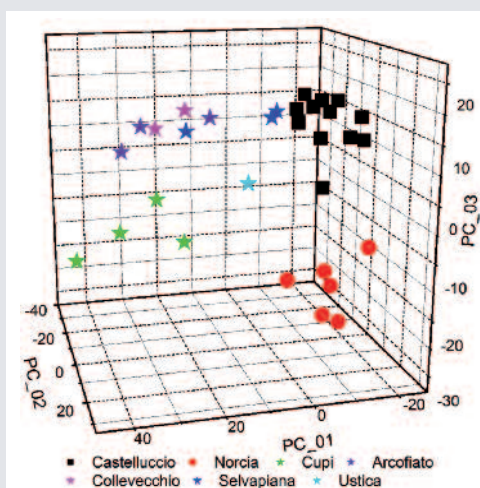


Fig. 4 - Rappresentazione dei campioni sulle prime tre variabili latenti della PLS-DA (40% della varianza degli spettri ritenuta); dati pretrattati con SNV e autoscalati per colonna

(loadings) delle variabili sui componenti più significativi, precedentemente calcolati (vedi Fig. 2), come indicatore dell'importanza delle variabili, si sono individuate le energie più significative (216) per il calcolo dei modelli di categoria. Eseguito l'elaborazione PLS-DA con le sole variabili selezionate si vede (Fig. 5) un lieve miglioramento delle capacità di classificazione e predizione in special modo per la categoria Norcia. In ognuno dei quadri di Fig. 5 sono confrontati i valori delle rispettive variabili categoriche calcolati coi rispettivi modelli di categoria, in classificazione (▲) ed in predizione (▼), e con il valore atteso (◆) per ogni campione. I risultati ottenuti sono stati confermati usando [16, 17] vari metodi di classificazione e modellazione del pacchetto v-parvus [18]. I diversi metodi di classificazione e modellazione devono essere applicati riducendo ulteriormente le variabili con l'utilizzo dei componenti principali più significativi valutati sulla base della cross-validation. In generale i metodi di classificazione, in particolar modo LDA, danno risultati migliori delle tecniche di modellazione. Questo può essere dovuto al fatto che nella modellazione si calcolano modelli per ciascuna categoria, quindi, con un numero ridotto di campioni che vincola anche il numero di componenti principali utilizzabili.

Il pretrattamento SNV con autoscaling di colonna è il più efficace, in questo caso la tecnica di classificazione LDA consente di ottenere sia una classificazione che una predizione superiore al 75% per tutte le categorie mentre sia SIMCA che PLS-DA mostrano un evidente overfitting, infatti pur classificando perfettamente tutti i dati hanno una scarsa capacità predittiva.

La tecnica di modellazione QDA-UNEQ permette di avere una capacità di classificazione simile a quella di LDA ma la sua capacità predittiva scade notevolmente specialmente per le categorie Norcia e Castelluccio.

La categoria "Altri" è molto ben classificata con tutte le tecniche, infatti, anche in predizione, si ottengono, per essa, risultati simili alla classificazione evidenziando così la possibilità di riconoscere dei campioni estranei a quelli di Norcia e Castelluccio. Meno efficiente è la discriminazione dei campioni di queste due cate-

gorie tra loro, solo LDA sembrerebbe idoneo allo scopo. Risultati peggiori si hanno quando i dati sono pretrattati con la sola SNV.

Conclusioni

Nel presente lavoro si è dimostrata la possibilità di usare la spettroscopia ED-XRF accoppiata all'elaborazione chemiometrica per ottenere una buona attribuzione del luogo d'origine della lenticchia:

- la classificazione è in tutte le classi elevata;
- è sempre possibile distinguere semi provenienti da località diverse da Castelluccio o Norcia;
- la predizione è scarsa specialmente nel discriminare campioni di Castelluccio da quelli di Norcia, tuttavia è possibile discriminare questi da quelli di altre località;

- la categoria Norcia è la meno definita;

Pur necessitando di verifiche ulteriori, una miglior selezione e riduzione delle variabili ed un incremento del numero di campioni potrebbero migliorare la classificazione.

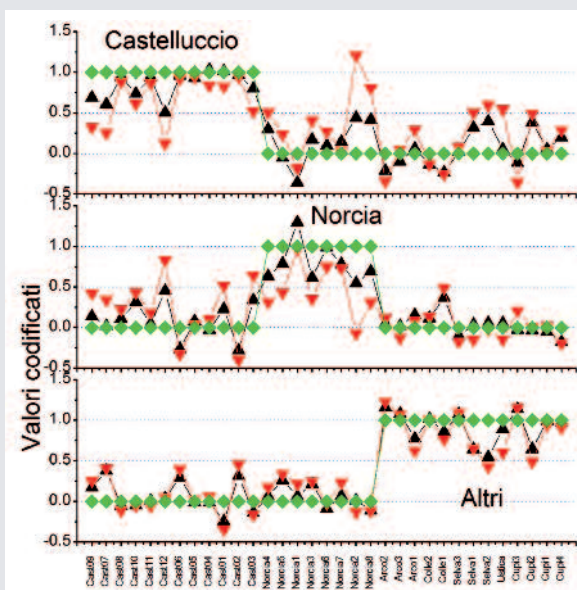


Fig. 5 - PLS-DA sui dati con sole 216 variabili. Risultati della classificazione per i tre modelli (Castelluccio, Norcia, Altri). I simboli indicano: ◆ valori attesi, ▲ risultati ottenuti in classificazione, ▼ risultati ottenuti in predizione.

Ringraziamenti:

Si ringraziano Az. Agr. Coccia Camillo - Norcia (PG); Az. Il Casale nel Parco dei Monti Sibillini Loc. Fontevena - Norcia (PG); Az. Agr. Di Giovambattista Vittorio Norcia (PG); Az. Agr. Perla Emilia loc. Castelluccio - Norcia (PG); Az. Agr. Lu Stazzu loc. Castelluccio - Norcia (PG); Az. Agr. Il Margine - Norcia (PG); Az. Agr. Il Quadrifoglio - Norcia (PG); Az. Agr. Salvatori Vincenzo loc. Opaco - Norcia (PG); Az. Agr. Hibiscus c.da Tramontana Ustica (PA); Di Mattia Ida e Sensi Antonio Loc. Villa S. Antonio - Visso (MC); Az. Agr. Rinozzi Vincenzo Fraz. Selvapiana - Montecavallo (MC); Az. Agr. Misici Mario Loc. Arcofiato - Camerino (MC); Testa Nicolina (Loc. Castelluccio) - Norcia (PG); Az. Agr. La Fattoria Loc. Vocabolo Fontevena - Norcia (PG); Az. Agr. La Casina di Opaco

Loc. Opaco - Norcia (PG); Az. Agr. Brandimarte Renzo loc. Castelluccio - Norcia (PG); Az. Agr. Liberti Giampietro Fraz. Campobonomo - Fiastra (MC); Az. Agr. Piccinini Domenico fraz. Gualdo - Castelsantangelo sul Nera (MC); che hanno cortesemente fornito i campioni e ci hanno supportato con utili consigli ed informazioni.

References

- [1] S. Jood *et al.*, *Nahrung*, 1998, **42**, 71.
- [2] M. Dueñas *et al.*, *Europ. Food Res. Tech.*, 2002, **215**, 478.
- [3] N. Wang, J.K. Daun, *Food Chem.*, 2006, **95**, 493.
- [4] Dil. Thavarajah, P. Thavarajah, see, C. Thiam, A. Vandenberg, *Food Chem.*, 2010, **122**, 254.
- [5] R.G. Ruiz, Raquel *et al.*, *J. Agric. Food Chem.*, 1996, **44**, 1526.
- [6] N. Wang *et al.*, *LWT Sci. Food Technol.*, 2009, **42**, 842.
- [7] R. Karoui, J. De Baerdemaeker, *Food Chemistry*, 2007, **102**, 621.
- [8] G.G. Bortoleto *et al.*, *Anal. Chim. Acta*, 2005, **539**, 283.
- [9] L. Perring *et al.*, *J. Agric. Food Chem.*, 2005, **53**, 4696.
- [10] F. Manhas Verbi Pereira *et al.*, *J. Agric. Food Chem.*, 2006, **54**, 5723.
- [11] M. NeCemer *et al.*, *J. Agric. Food Chem.*, 2009, **57**, 4409.
- [12] E. Marguá *et al.*, *Trends in Analytical Chemistry*, 2009, **28**, 362.
- [13] A. Pisani *et al.*, *Food Chemistry*, 2008, **107**, 1553.
- [14] R. Koplík *et al.*, *J. Chrom. B. Anal. Tech. Biomed. Life Sci.*, 2002, **775**, 179.
- [15] A. Gonzalez *et al.*, *TRAC. Trends in Analytical Chemistry*, 2009, **28**, 1295.
- [16] R.G. Brereton, *Chemometrics Data Analysis for the Laboratory and Chemical Plant*, John Wiley & Sons, 2003.
- [17] B.G.M. Vandeginste *et al.*, *Handbook of Qualimetrics*, Part B, Elsevier, 1998.
- [18] M. Forina *et al.*, *PARVUS Release 2008*, Dip. Chimica e Tecnologie Farmaceutiche, University of Genova, available (free, with manual and examples) from authors or at www.pavvus.unige.it.

ABSTRACT

Typify Lentils with XRF

In this paper we try to develop an analytical method able to characterize the geographical origin of lentils, especially, we want to classify the Castelluccio di Norcia IGP lentil. Samples of the same variety of lentil seeds were collected directly from farmers in an area within 50 km from Castelluccio di Norcia. The samples were ground and flours, compressed into pellets, were analyzed in a vacuum with ED-XRF. The acquired spectra were processed using various classification techniques. Performed calculations identified a processing mode that allows to distinguish samples coming from Norcia and Castelluccio di Norcia with respect to all the others. The separation of samples from Norcia and Castelluccio, still above 75% using the LDA technique, results slightly less defined.