Confronto tra diverse strategie per il calcolo di modelli di calibrazione: esempio di analisi del latte

Stefania Barzaghi1\*, Giusi Poerio1

1 CREA- Centro di Zootecnia e Acquacoltura, Via A. Lombardo, 11, 26900 Lodi, Italia;

[stefania.barzaghi@crea.gov.it](mailto:stefania.barzaghi@crea.gov.it); giusi.poerio@gmail.com

\*

Nel 2009, Filtzmoser et al., hanno proposto la “Repeated double cross validation” (rdCV), per il calcolo di modelli di calibrazione, per fare una stima realistica degli errori di predizione di eventuali campioni futuri.

Scopo del presente studio è confrontare le performance di modelli di calibrazione, da spettri NIR di campioni di latte, calcolati mediante diverse strategie: A) il numero ottimale di variabili latenti è stato ottenuto con la rdCV, utilizzando il package “Chemometrics” di R; B) il modello di calibrazione è stato calcolato dividendo i campioni in due set, uno di calibrazione ed uno di validazione, scelti con l’algoritmo di Kennard-Stone, mediante il software PLS Toolbox (Eigenvector Research, USA)

105 campioni di latte prelevati sia da singoli animali che da latte di massa, sono stati raccolti nei mesi di Marzo e Aprile 2021 per essere utilizzati nel calcolo dei modelli di calibrazione.

Le performance dei modelli ottenuti sono state valutate su 35 campioni di latte di massa, provenienti da diverse stalle, raccolti in un periodo successivo, tra Maggio e Settembre 2021.

I campioni di latte sono stati analizzati con Milkoscan (Foss Italia, Padova) per quantificare il contenuto in grasso, proteine, lattosio e solidi totali, e scansiti con spettrometro Proxi­Mate TM (Büchi Labortechnik AG, Flawil (CH)) nel range spettrale da 400 a 1700nm. Per le letture in transflettanza, i campioni (spessore di 1mm) sono stati posti in capsule Petri ruotanti.

Applicando entrambe le strategie, nel quantificare i diversi componenti del latte, il numero ottimale di variabili latenti è risultato lo stesso, tranne che nel caso del lattosio. Nella predizione di campioni indipendenti, le migliori performances sono state ottenute per i modelli calcolati con la strategia B. Per il contenuto in grasso l’RMSEP del modello da A è risultato 0.55 mentre per il modello da B 0.13. Per il contenuto in proteine gli RMSEP sono stati rispettivamente pari a 0.39 per A e 0.29 per B.

**Keywords:** Repeated double cross validation, calibrazioni, latte, grasso, proteine, lattosio

**Acknowledgements:** Lo studio è stato svolto nell’ambito del progetto AGRI HUBSviluppo ed integrazione tecnologica di una piattaforma high-throughput per il miglioramento sostenibile dei processi produttivi delle filiere dell’agroalimentare – Sub-task 3.2 Influenza delle innovazioni sul benessere animale e sulla qualità del latte.

REFERENCES

P. Filzmoser, B. Liebmann, K. Varmuza. 2009 Repeated double cross validation, J. Chemom. 23, 160-171. https://doi.org/10.1002/cem.1225