Exploring the potential of local partial least squares regression in forages analysis

A. Deryck[1], P. Vermeulen[1], V. Baeten[1] and J.A. Fernández Pierna[1]

[1]Walloon Agricultural Research Centre (**CRA-W**), Knowledge and valorization of agricultural products Department, Quality and authentification of agricultural products Unit, Gembloux, Belgium

Contact: Antoine Deryck (a.deryck@cra.wallonie.be)

1 INTRODUCTION

Feed, and particularly forage characterization is critical for optimal growth and health of livestock. However, the variability of nutrient concentrations in these products makes it challenging. Near-infrared spectroscopy (NIRS) combined with chemometrics appears to be a promising, rapid, and non-destructive technique for reliable prediction of forage composition. Among all existing chemometrics tools, partial least squares regression (PLS) has become widely used for analyzing the nutrient content of forages from their near-infrared spectra. In recent years, local frameworks, which involve selecting a subset of samples for PLS model calibration, have demonstrated their efficiency in improving prediction accuracy.

This research aims to compare the performances of traditional global PLS (general models capturing overall patterns in the dataset) with four different local PLS methods (specific models based on similar spectra) on a dataset to determine their potential for accurate forage characterization. The motivation behind exploring these variations lies in their potential to provide more accurate predictions by focusing on locally relevant data.

2 DATABASE

Feed database = 583 samples (Effort project, Belgium 2020-2022) from 4 products (figure 1).

Figure 1. Composition of the Feed database

Figure 2. AuroraNIR, GraiNit, Italy

- Spectra measured with the GraiNit AuroraNIR (figure 2). Range of 950-1650 nm (each 2 nm).
- Chemical composition in terms of acid detergent fiber, acid detergent lignin, ash, crude fiber, neutral detergent fiber, and protein obtained with historical FOSS XDS equations *(O. Minet et al. (2016). Requasud: la spectrométrie proche infrarouge, ASBL Requasud, 32p.)*
- Dry matter content measured with the oven method.

3 METHODS

PLS models, classical LPLS, and 3 variations of LPLS (figure 3) were built on those spectra to predict the samples' chemical composition.

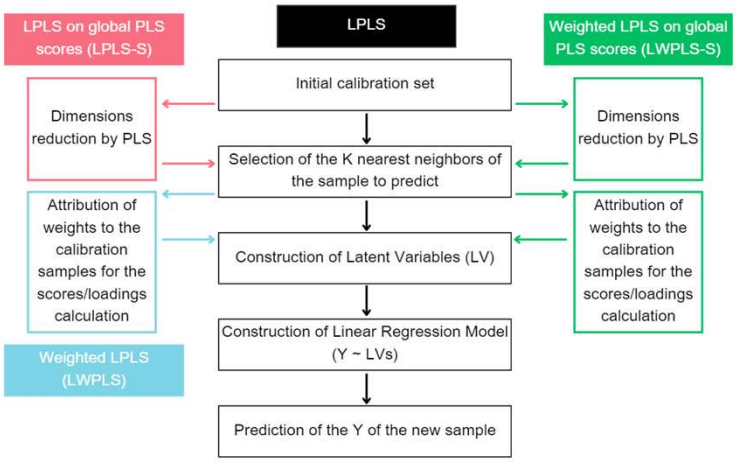

Figure 3. LPLS variations applied to the feed database

The local approach of the LPLS gives them some specific features:

- Better consideration of data complexity
- Robust to outliers
- Allow centralization of samples of different natures

- Increased computation time
- Technical constraints for implementation on portable instruments

- Weighted variations: Adjust the importance of samples according to spectral similarities.
- Global PLS scores: Preliminary dimensionality reduction, leading to reduced computation time.

4 RESULTS

The performances of these models were evaluated after optimization of the parameters on a calibration set (random split 70% calibration 30% validation, tuning based on the lowest RMSECV). Those performances, expressed in RMSEP (Table 1), were similar to or slightly better than Global PLS for all parameters with the local methods.

Table 1. Performances of the PLS, LPLS, LPLS-S, LWPLS, and LWPLS-S. The RMSEP is expressed in percentage of dry matter

	Acid detergent fiber (ADF)	Acid detergent lignin (ADL)	Ash	Crude fiber (CF)	Dry matter (DM)	Neutral detergent fiber (NDF)	Protein (Prot)
PLS	2.21	0.57	1.44	2.11	4.26	4.24	1.68
LPLS	1.97	0.50	1.58	1.92	4.10	4.11	1.82
LPLS-S	1.80	0.55	1.45	1.76	3.89	3.63	1.69
LWPLS	1.98	0.51	1.49	1.93	3.99	3.97	1.83
LWPLS-S	1.80	0.50	1.55	1.64	3.82	4.19	1.59
Parameter' Span	14.71	3.55	9.63	12.59	65.29	30.65	14.39

5 CONCLUSION AND PERSPECTIVES

As the results showed, there is a clear potential for local methods in feedstuffs characterization and this potential could be even more significant with a larger dataset with local variation within individual product samples (where Global PLS might show worse results). A database of more than 24,000 samples of feed products (formulations) is currently being built at the CRA-W to compare the performances of PLS, LPLS, and LPLS variations in this context.

Acknowledgments
The research leading to these results has received funding from the "Plan de relance de la Wallonie" project Mobilab. The authors wish also to thank the CRA-W teams from the Research Unit "*Quality and authentication of agricultural products*" for providing and processing the samples as well as performing NIR analyses.