Comparison between different standardization methods for the quality estimation of forages in portable NIR spectrometers

Pierantonio Facco[1], F. Benozzo[2], M. Barolo[1] and P. Berzaghi[2]

[1]CAPE-Lab, Computer Aided Process Engineering Laboratory

Department of Chemical Engineering Principles and Practice

[2]Department of Animal Science

Università di Padova (Italy)

"NIR on the go" 2010 – AGRIPOLIS, Legnaro - May 27th, 2010

Summary

- Introduction to the standardization of NIR spectrometers

- State-of-the-art transfer methods

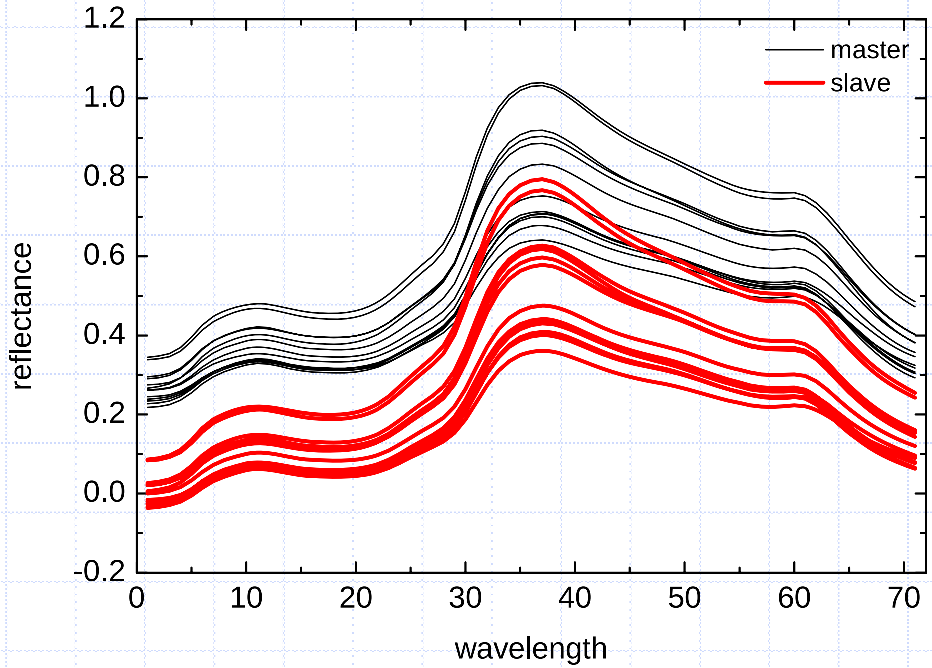
- Cross-validatory automatic technique for the selection of proper standardization methodologies
  - case study: portable NIR spectrometer for the estimation of the quality of bovine foodstuff

- Conclusions

Standardization of NIR instrumentation

- Model transfer is always required in NIRS:
  - sampling differences
  - accuracy of the wavelengths
  - electronic and mechanical tolerances
  - environmental differences (temperature, humidity, etc...)

- Aim of the standardization:
  - having the same responses from different instruments
  - accuracy ~ limit of repeatability of the measurements



"master" = reference

model transfer

"slaves"

Case study: portable NIR spectrometers

- Issue of model transfer is particularly difficult in **portable instrumentation**:
  - reduced number of calibration samples
  - heavy calibration burden for the analyst
  - the same accuracy as lab instrumentation is required
- Instrumentation: **AgriNIR®** (Dinamica Generale, Mantova - Italy):
  - diode array technology
  - reflectance spectra
  - wavelength range between 1100-1800 nm
- Analysis of forages for bovine feeding

AgriNIR® - Dinamica Generale

Data of forages for bovine foodstuff

- **Forages**: high moisture corn and grass silage
- **Spectra**:
  - spectra of 1 master: $\mathbf{X}_M$ (13 samples×71 wavelengths)
  - spectra of 3 slaves: $\mathbf{X}_S$ (13×71)
- **Quality parameters R** (13×6)

Quality variable	Acronym
dry matter (%)	DM
crude protein (%)	CP
acid detergent fiber (%)	ADF
neutral detergent fiber (%)	NDF
ash (%)	ASH
crude fat (%)	CF

- Adopted standardization strategy → transformation of $\mathbf{X}_S$ in $\mathbf{X}_M$

State-of-the-art transfer techniques

- Several transfer methods are available in the Literature:
  - patented method **PM** (Shenk and Westerhaus, 1985), window w=2 wavelength
  - direct standardization **DS** (Wang *et al.*, 1991)
  - piecewise direct standardization **PDS** (Wang *et al.*, 1991), w=5
  - double window piecewise direct standardization **DWPDS** (Wold *et al.*, 1998), w=5, 6
  - partial least squares **PLS** (Forina *et al.*, 1995)
  - orthogonal signal correction **OSC** (Wold *et al.*, 1998)

PM

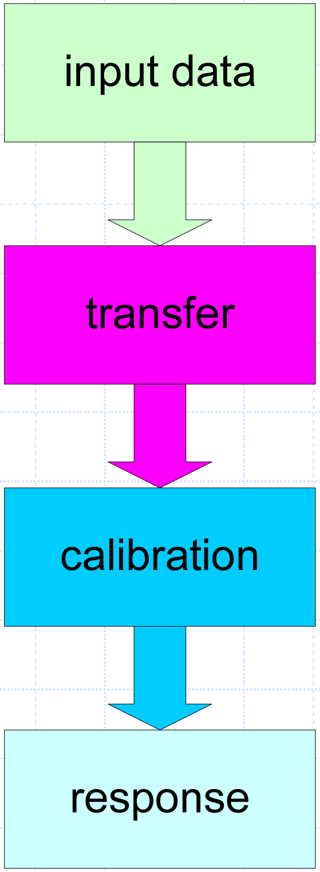
$$\begin{cases} l = \max(\mathrm{corr}(\mathbf{r}_{\mathrm{M},i}, \mathbf{r}_{\mathrm{S},i-w}), \mathrm{corr}(\mathbf{r}_{\mathrm{M},i}, \mathbf{r}_{\mathrm{S},i-w+1}), \ldots, \mathrm{corr}(\mathbf{r}_{\mathrm{M},i}, \mathbf{r}_{\mathrm{S},i+w})) \\ \lambda_{\mathrm{S},l} = a + b \cdot \lambda_i \\ \mathbf{r}_{\mathrm{M},i} = c + d \cdot \mathbf{r}_{\mathrm{S},l} \end{cases}$$

DS

$$\begin{cases} \mathbf{T} = \mathbf{X}_{\mathrm{S}}^{+} \mathbf{X}_{\mathrm{M}} \\ \mathbf{X}_{\mathrm{S}}^{\mathrm{T}} = \mathbf{X}_{\mathrm{S}} \mathbf{T} \end{cases}$$

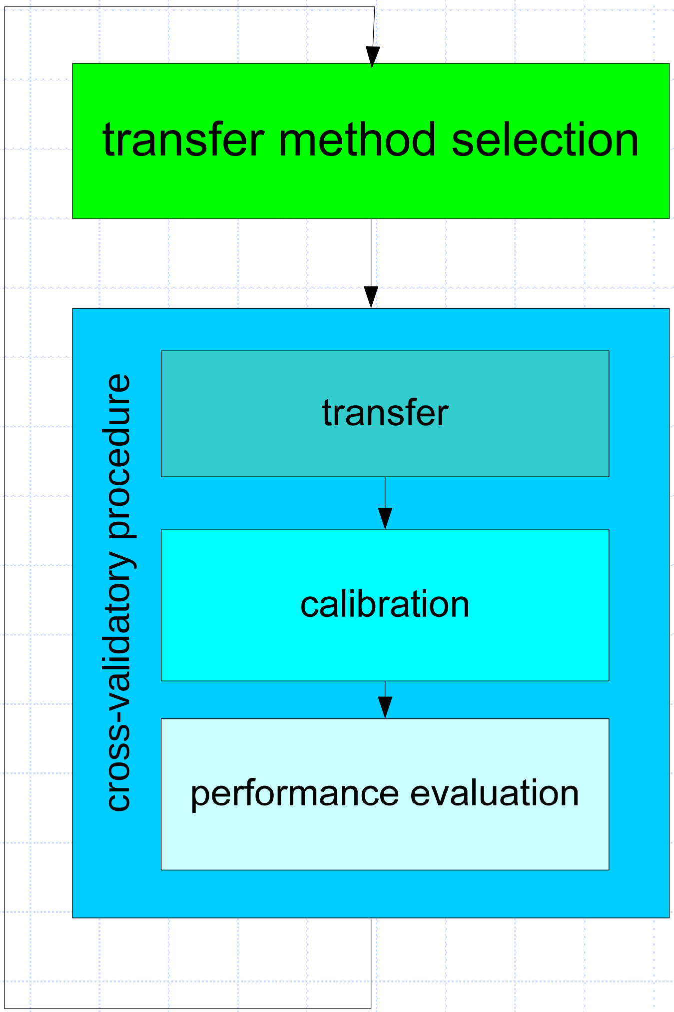
PDS

$$\begin{cases} \mathbf{X}_{\mathrm{S},i} = \left[\mathbf{x}_{\mathrm{S},i\text{-}w} \quad \mathbf{x}_{\mathrm{S},i\text{-}w+1} \quad \ldots \quad \mathbf{x}_{\mathrm{S},i+w} \right] \\ \mathbf{b}_i = \mathbf{X}_{\mathrm{S},i}^{+} \mathbf{x}_{\mathrm{M},i} \\ T = \mathrm{diag}(\mathbf{b}_1^{\mathrm{T}}, \mathbf{b}_2^{\mathrm{T}}, \ldots, \mathbf{b}_p^{\mathrm{T}}) \end{cases}$$

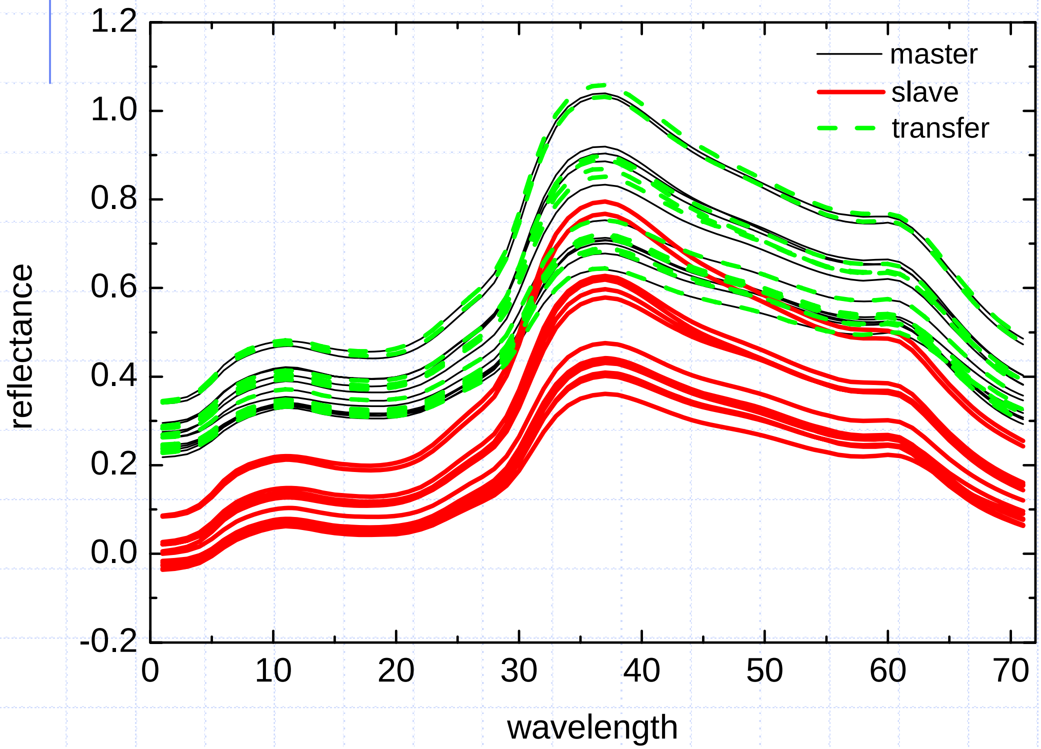


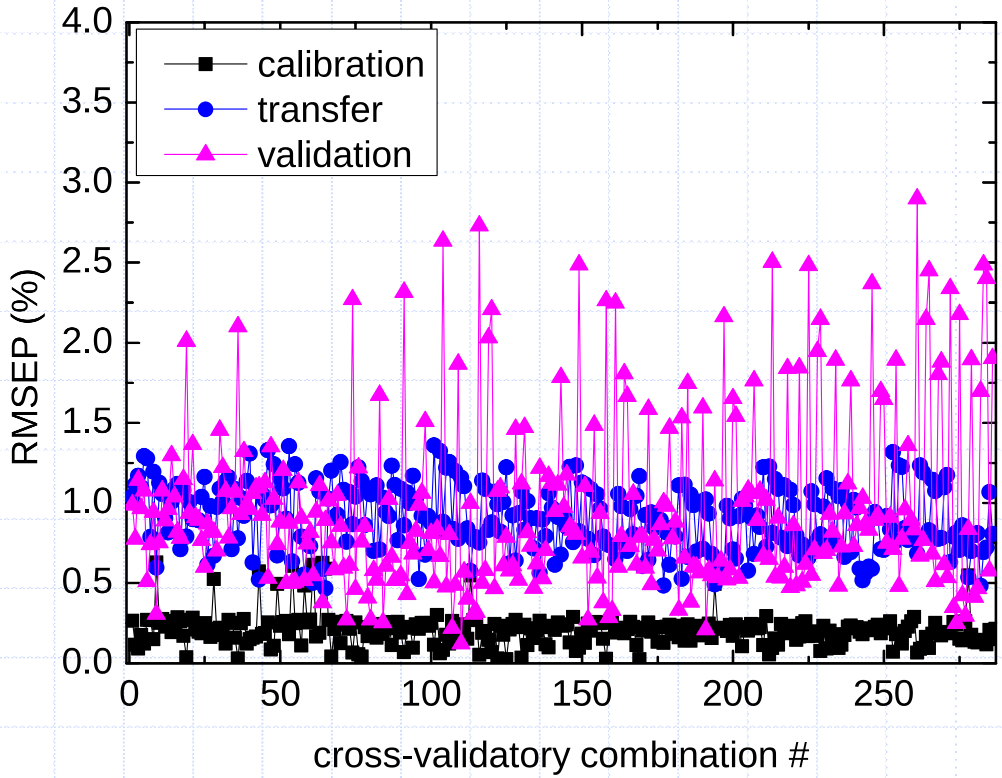
Selection of a proper standardization method

- Cross-validatory (automatic) procedure for forages analyzed both in master and slave instruments:
  - dataset split in: 10 reference samples
    3 validation samples
  - transfer using state-of-the-art methods
  - model calibration
  - performance evaluation:
    - **average root mean square error of prediction, ARMSEP**
    - **average mean relative error, AMRE**
  - change of the calibration/validation set
  - evaluation of the transfer method



Example of standardization with PDS





$$MRE = \frac{100}{N} \sum_{n=1}^{N} \frac{\sqrt{\left(y_n - \hat{y}_n\right)^2}}{\hat{y}_n}$$

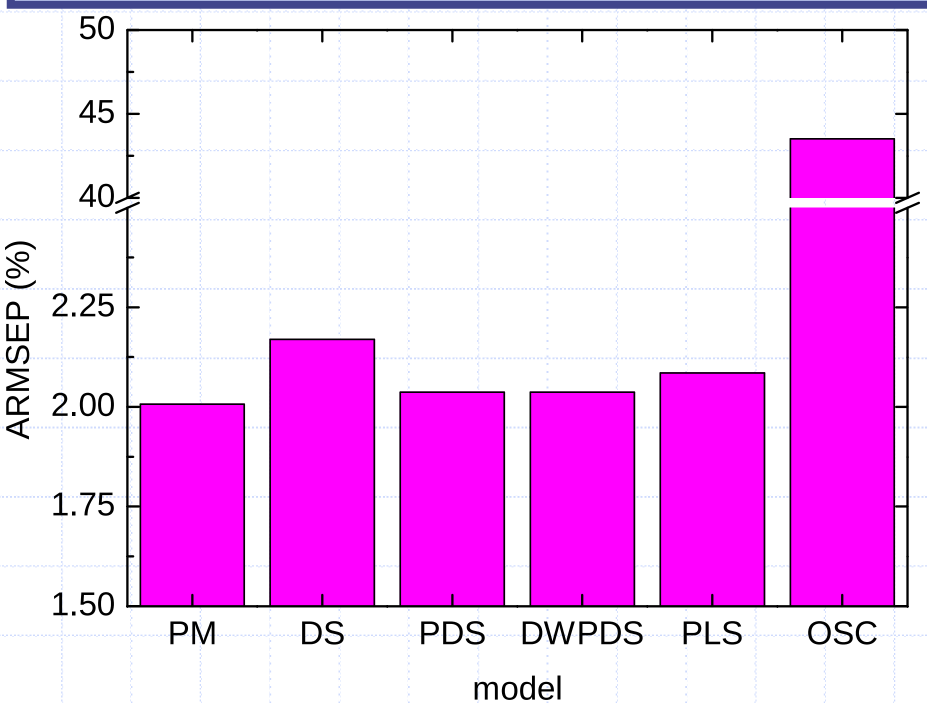
$$RMSE = \sqrt{\frac{\sum_{n=1}^{N} \left(y_n - \hat{y}_n\right)}{N}}$$

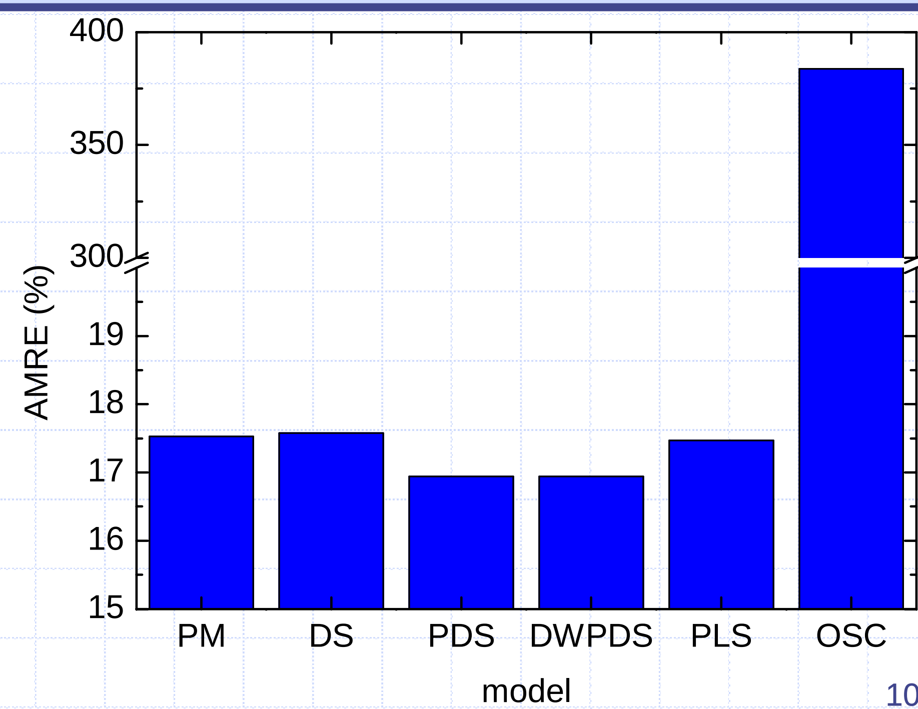
variable DM	ARMSEP (%)	AMRE (%)
calibration	0.21	0.28
transfer	0.91	1.05
validation	0.99	1.32

General results of the standardization

- Results are averaged throughout all the directions of variability:
  - forages
  - slaves
  - quality variables
  - reduction of the number of calibration samples

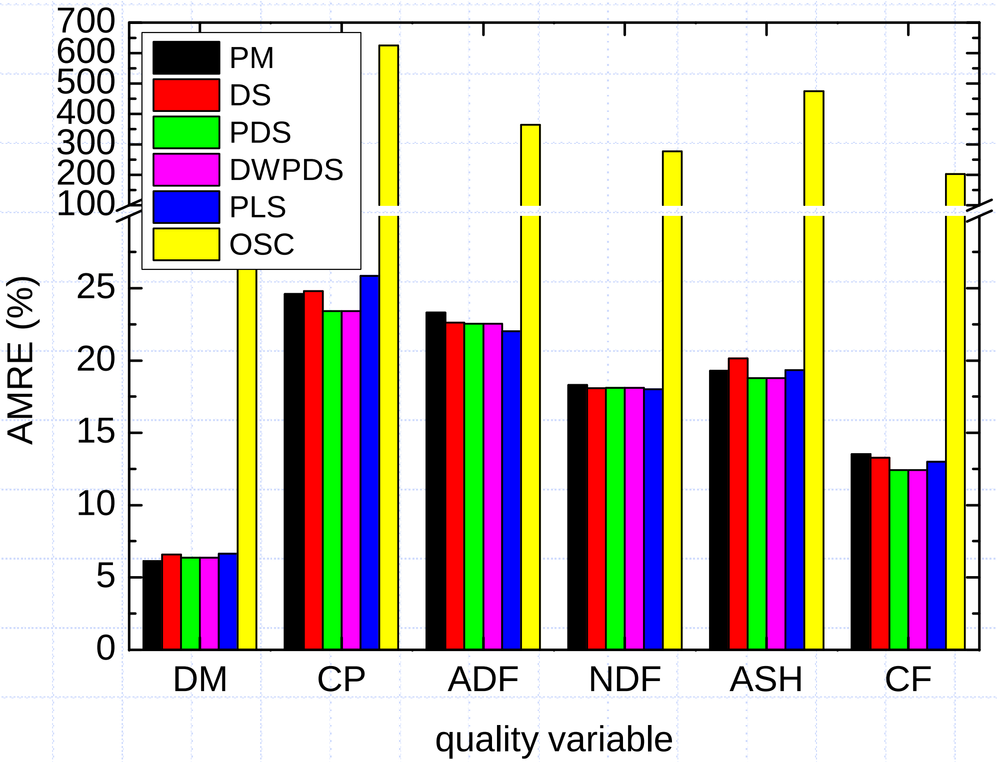
- The most accurate methods are **PDS and DWPDS**
  - DS and PM are slightly outperformed
  - OSC results to be inadequate

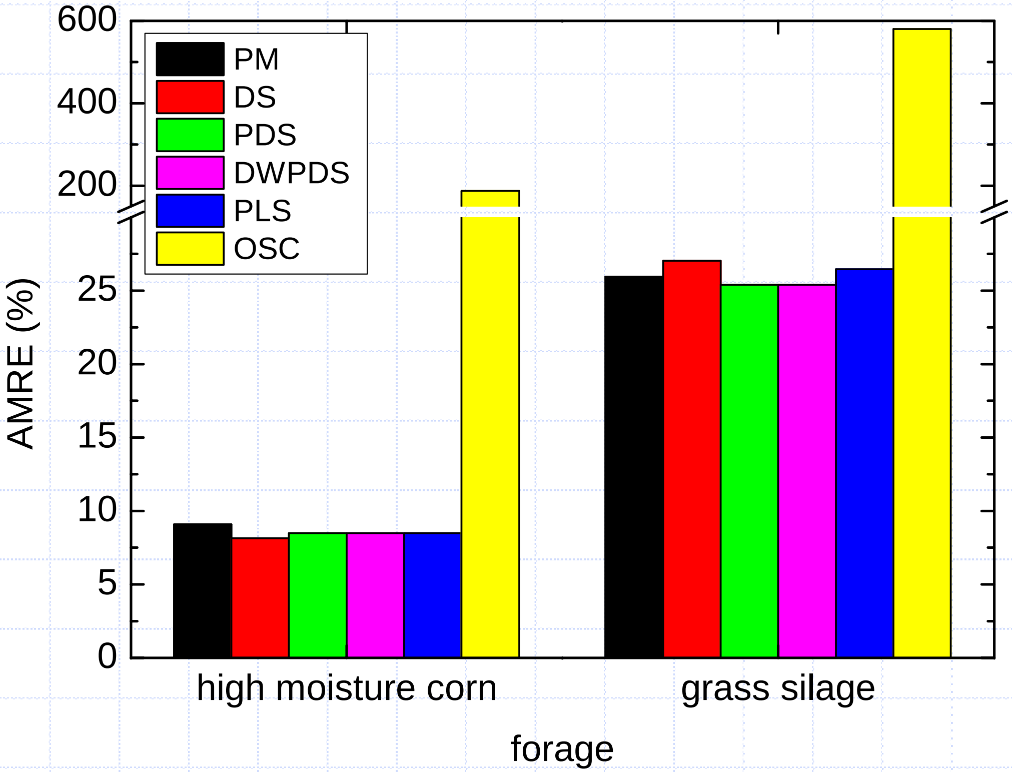




General results of the standardization

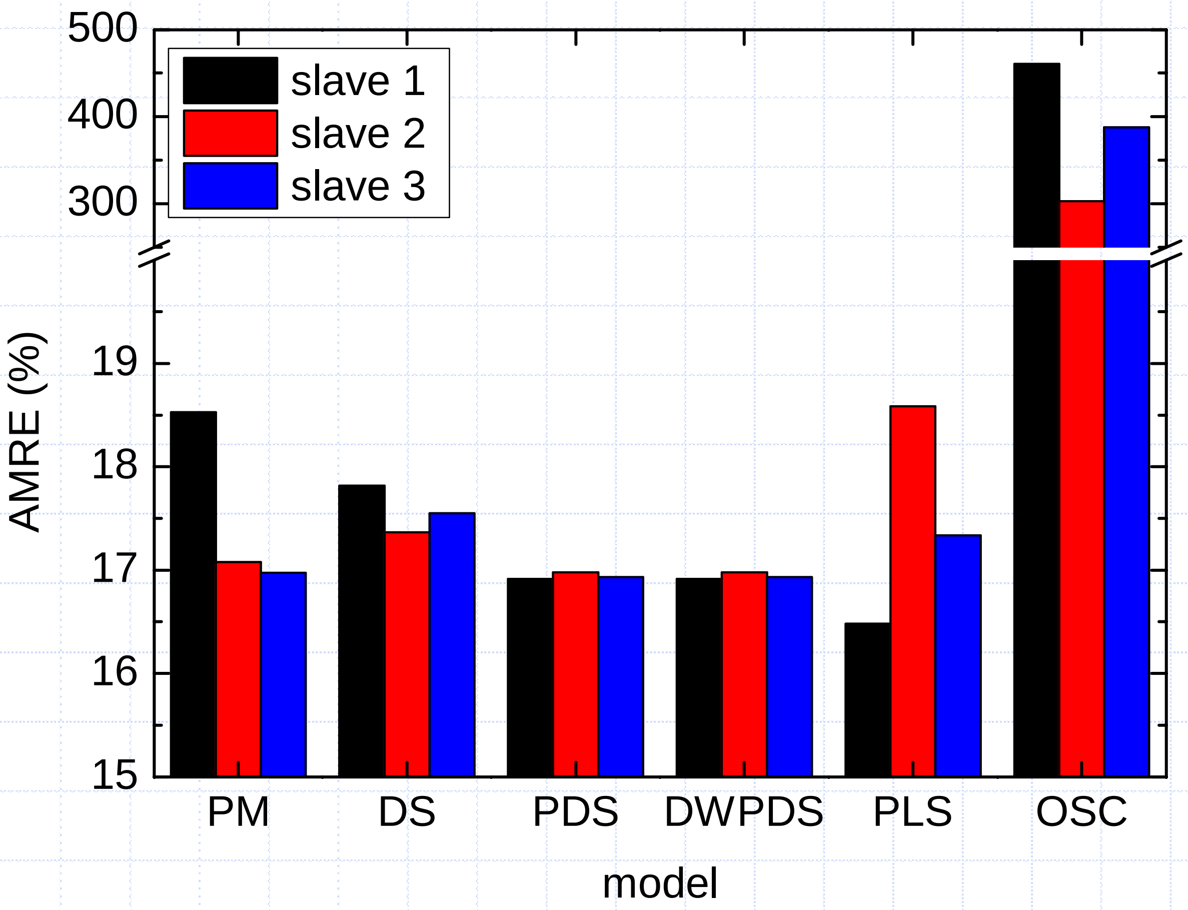
- The results are similar for the estimation of different quality variables
- Quality of the high moisture grass is easier to estimate than quality of grass silage





Transfer to different slaves

- Transfer to different slaves does not determine substantial differences
- PDS and DWPDS are the most robust methods for the change of slave

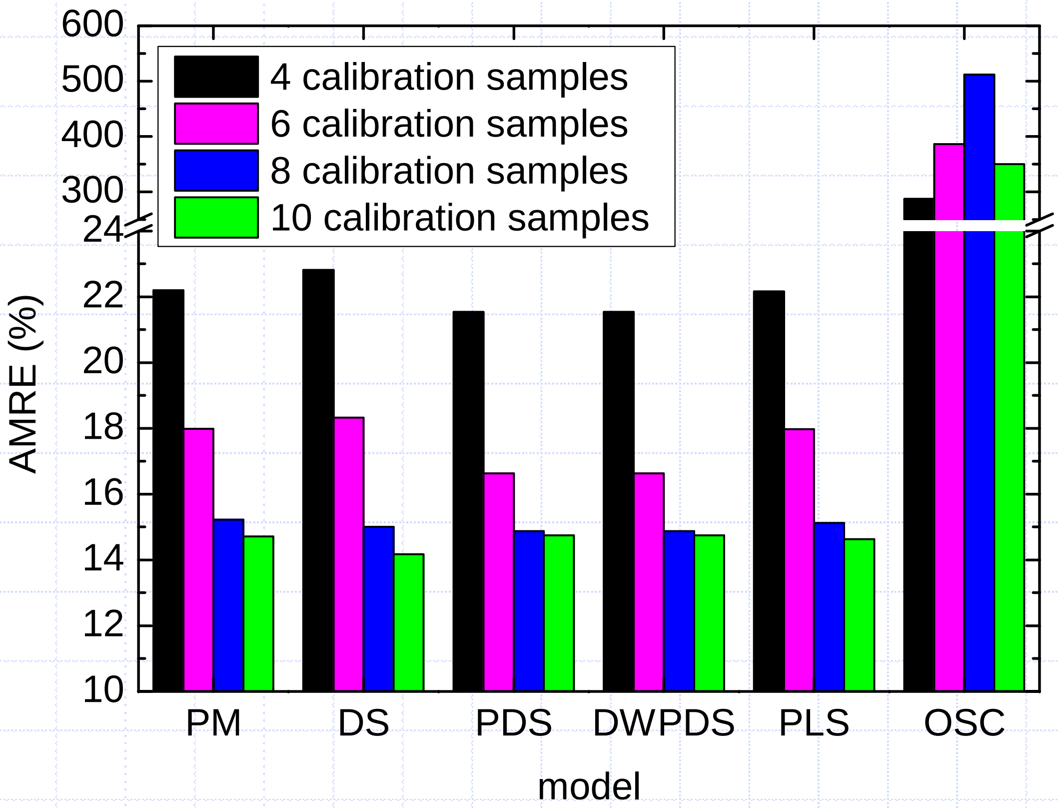


Reduction of the calibration set

- Subset of the calibration dataset can be selected based on the leverage of the samples

$$\text{leverage} = \text{diag}\left(\mathbf{X}_S \mathbf{X}_S^+\right)$$

- PDS and DWPDS are the most robust top the reduction of the calibration dataset



Concluding remarks

- A fully automatic methodology for the evaluation of different standardization techniques was developed:
  - PM, DS, PDS, DWPDS, PLS, OSC
  - the recommended method is **PDS** with a window **w=5** wavelength

- PDS did not outperform dramatically the other methods, but demonstrated:
  - greater **accuracy** (lower errors)
  - higher **robustness** to the reduction of the number of calibration samples
  - relative **simplicity** of implementation

- Future perspectives:
  - wider number of case studies and further generalization of the results
    - new forages
    - new slaves
    - different instrumentation
  - development of faster methods for the selection of reliable standardization models

Thank you for your kind attention!

pierantonio.facco@unipd.it

http://www.capelab.dipic.unipd.it/