A REVIEW OF CALIBRATION TRANSFER AND MAINTENANCE IN AN UNSUPERVISED WAY

Valeria Fonseca Diaz

Supervisors: Jean-Michel Roger, Wouter Saeys

KU Leuven, Belgium

INRAE, Montpellier, France

HélioSPIR - June 2022

Nantes, France

WHAT IS CALIBRATION TRANSFER?

Motivated by transfer learning *(Pan, S. et al. 2010)*

Calibration transfer may refer to retaining and reusing a source calibration model and/or the source calibration data to achieve an adaptation obtaining a target calibration model.

Methodology	Types of samples		
Category	**Standards**	**Target spectra and reference values**	**Only target spectral measurements**
Linear standardization	DS, PDS		Spectral mean correction
Prediction correction		SBC	
Concatenation	Joint PCA	Joint PLS	
Orthogonalization	TOP	DOP	uDOP
Variable selection	CAWS		di-CovSel
Domain invariant		Supervised di-PLS	Unsupervised di-PLS

Methodology	Types of samples		
Category	**Standards**	**Target spectra and reference values**	**Only target spectral measurements**
Linear standardization	DS, PDS		Spectral mean correction
Prediction correction		SBC	
Concatenation	Joint PCA	Joint PLS	
Orthogonalization	TOP	DOP	uDOP
Variable selection	CAWS		di-CovSel
Domain invariant		Supervised di-PLS	Unsupervised di-PLS

1. SPECTRAL MEAN CORRECTION

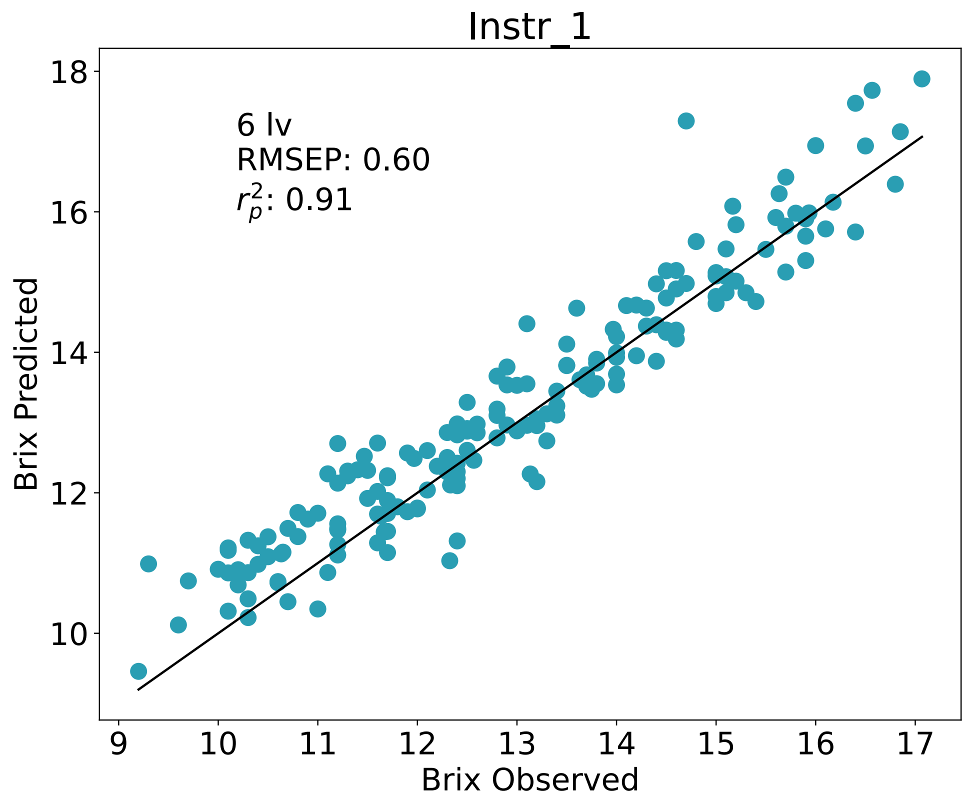
$$y_s = X_s B_s + \beta_s$$

$$\beta_s = \bar{y} - B'_s \bar{X}_s$$

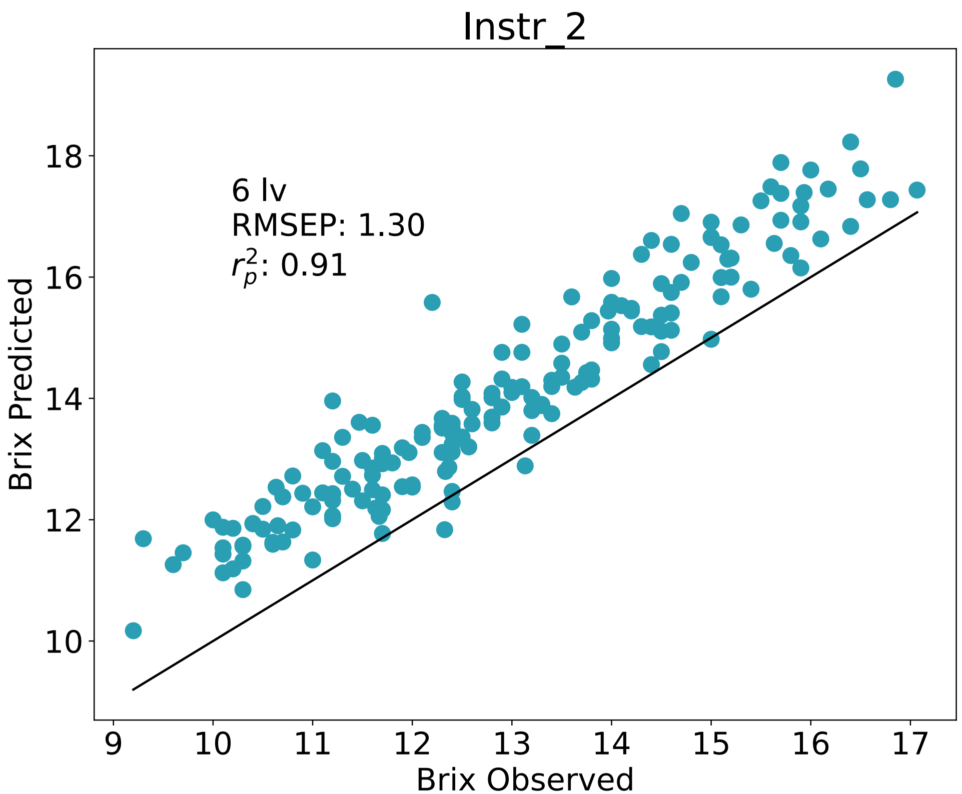
$$\beta_t = \bar{y} - B'_s \bar{X}_t$$

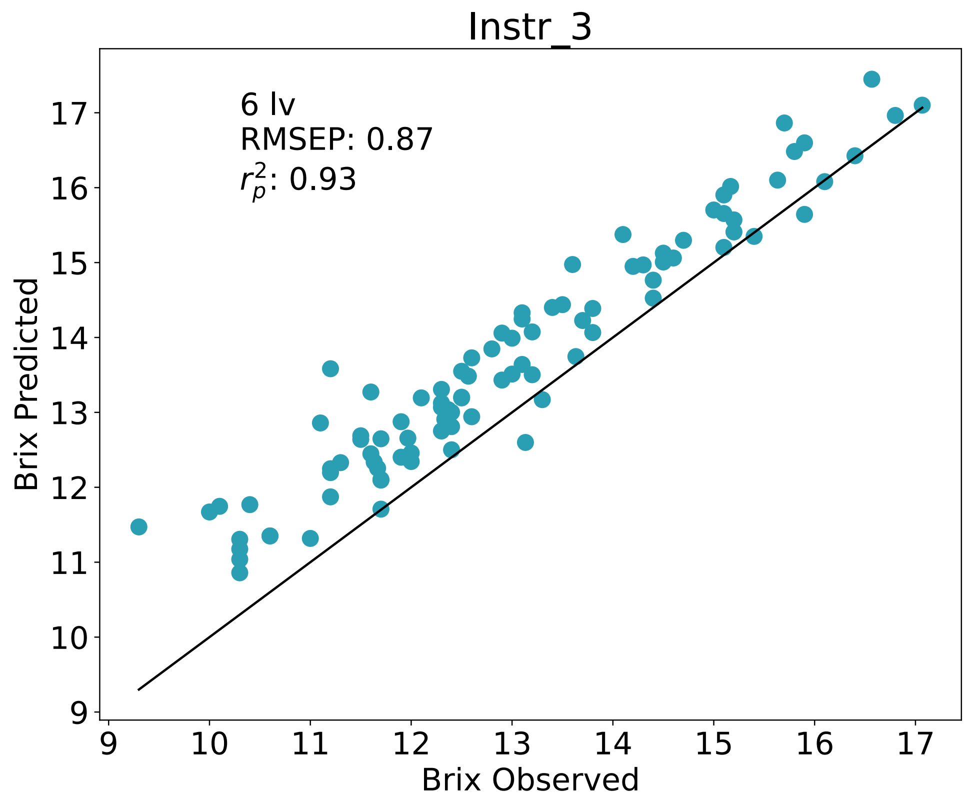
$$y_t = X_t B_s + \beta_t$$

BRIX PREDICTION IN PEARS FOR MULTIPLE INSTRUMENTS

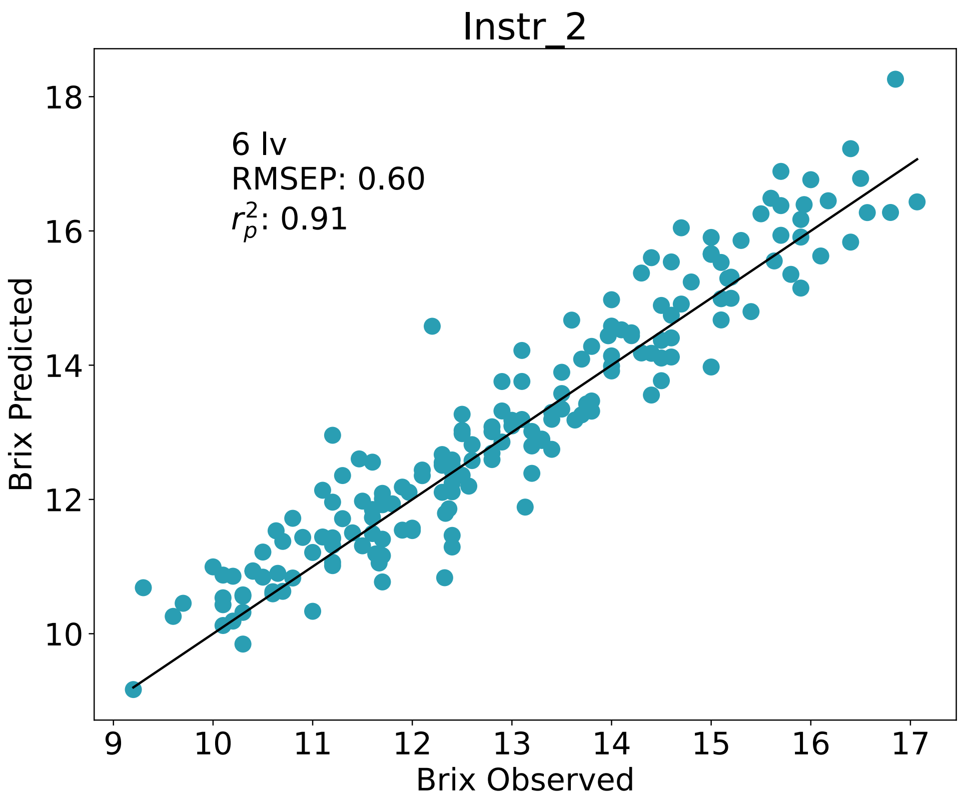


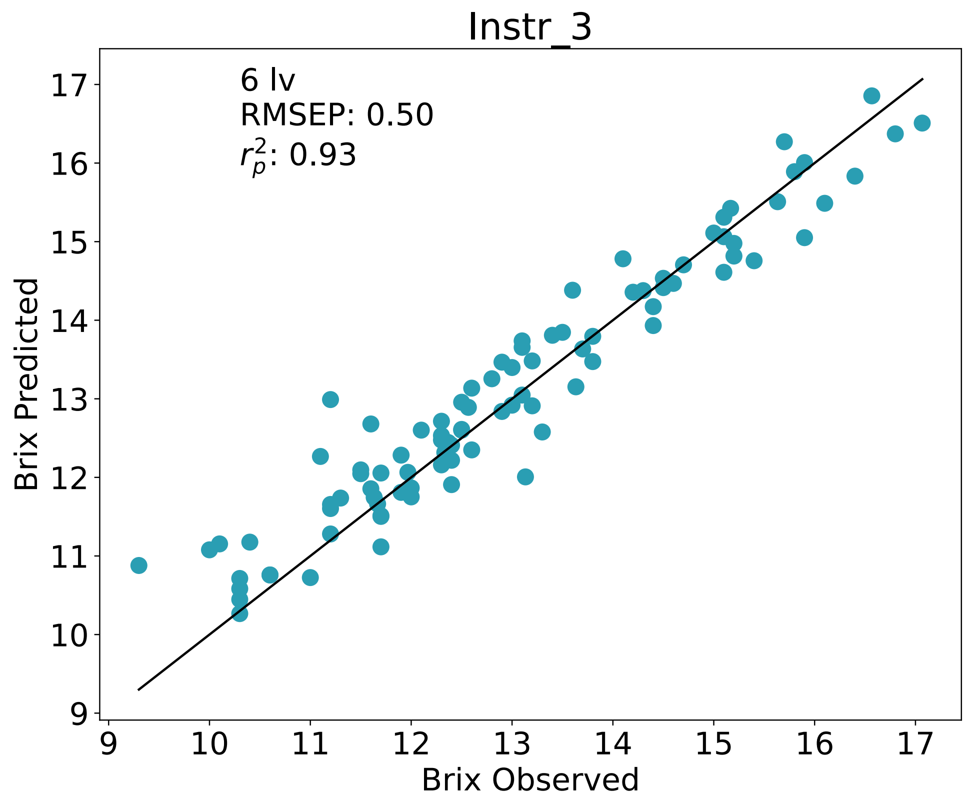
BRIX PREDICTION IN PEARS FOR MULTIPLE INSTRUMENTS





BRIX PREDICTION IN PEARS FOR MULTIPLE INSTRUMENTS





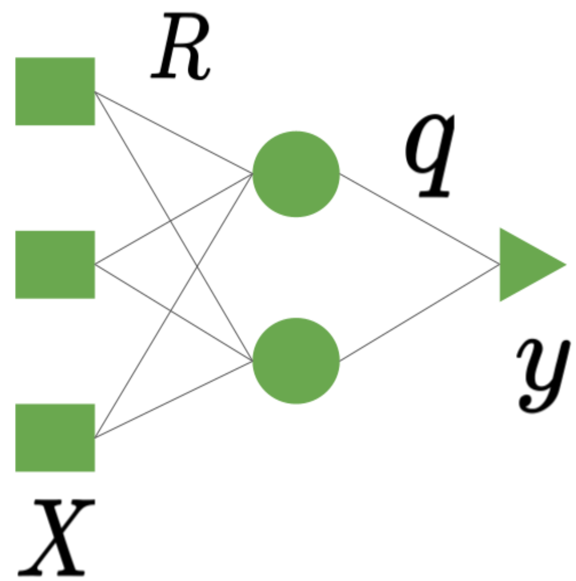
2. CALIBRATION USING DOMAIN INVARIANT PLS *(Nikzad-Langerodi, R. et al. 2018)*

Optimize

$$\min_{w} ||X - yw^T||$$

subject to:

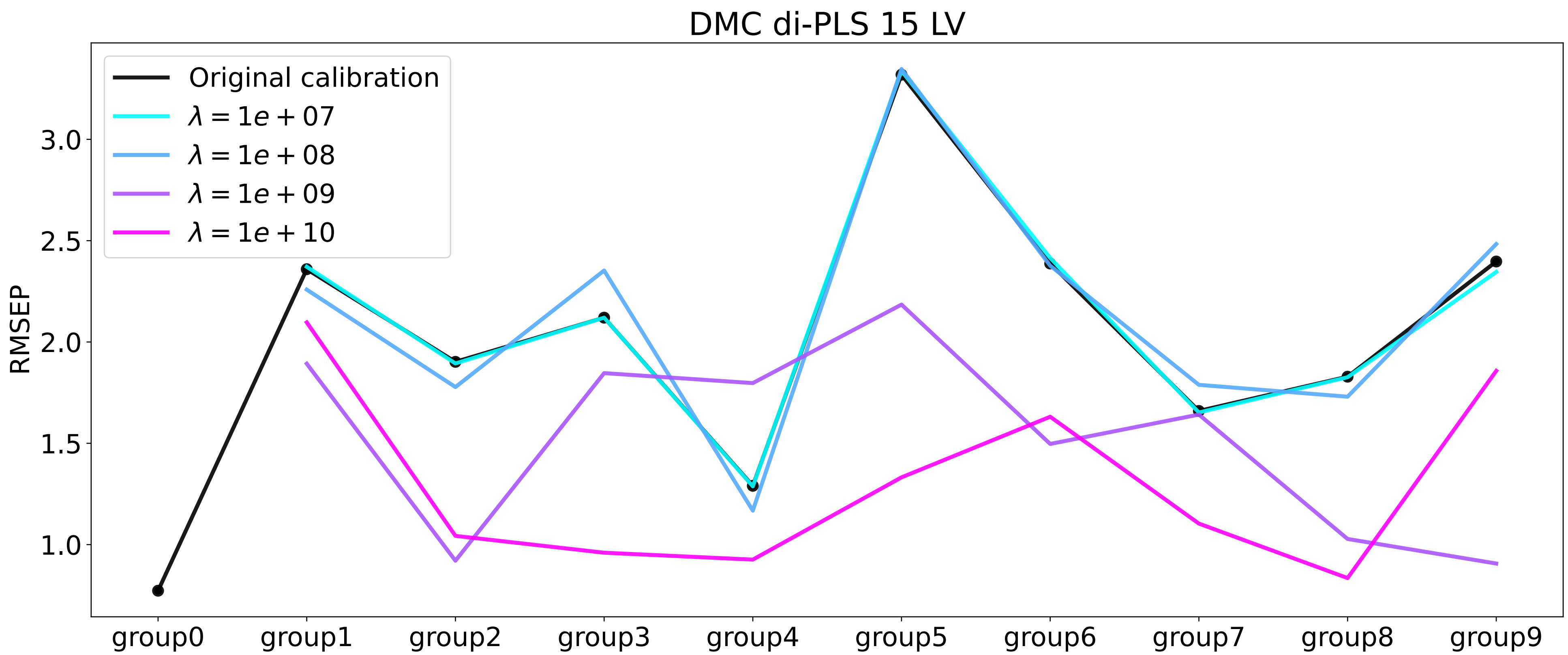
$$|var(\mathbf{t}_s) - var(\mathbf{t}_t)| = 0$$



DRY MATTER CONTENT PREDICTION IN MANGO SAMPLES

Domain	Type	Cultivar	Season	n
Source	Hard Green	Caly	1	1277
Target 1	Hard Green	KP	1	1059
Target 2	Hard Green	HG	1	293
Target 3	Hard Green	KP	2	340
Target 4	Hard Green	HG	2	88
Target 5	Hard Green	R2E2	2	276
Target 6	Hard Green	KP	3	216
Target 7	Hard Green	HG	3	80
Target 8	Hard Green	R2E2	3	156
Target 9	Ripen	KP	3	78

DRY MATTER CONTENT PREDICTION IN MANGO SAMPLES FROM SEVERAL DOMAINS



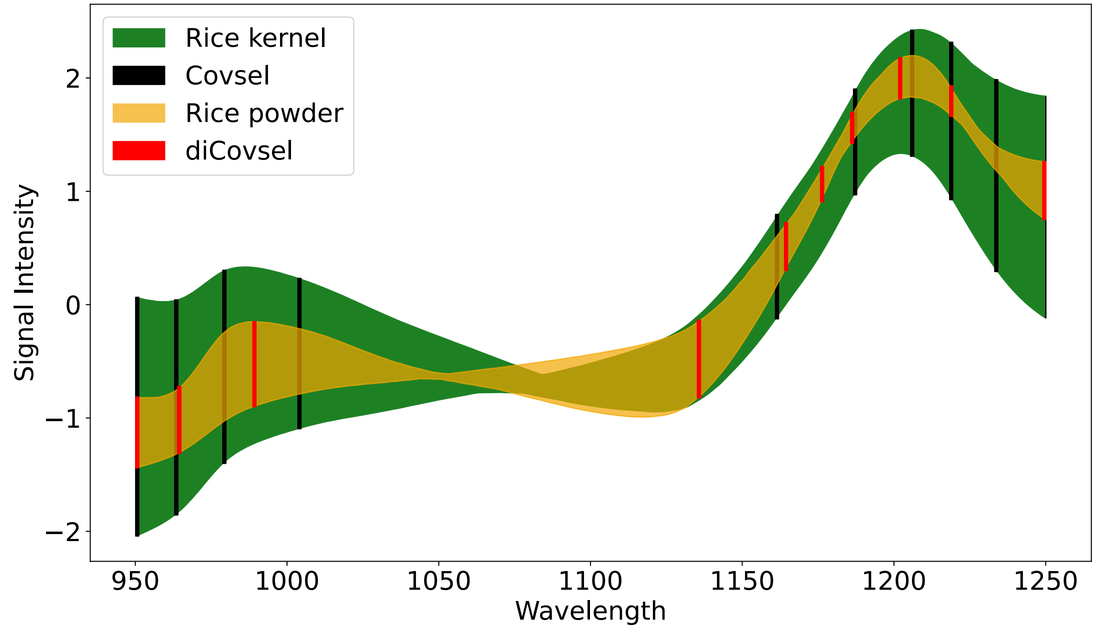
3. VARIABLE SELECTION USING DOMAIN INVARIABILITY (DI-COVSEL) *(Diaz, V.F.,et al. 2022)*

Optimize

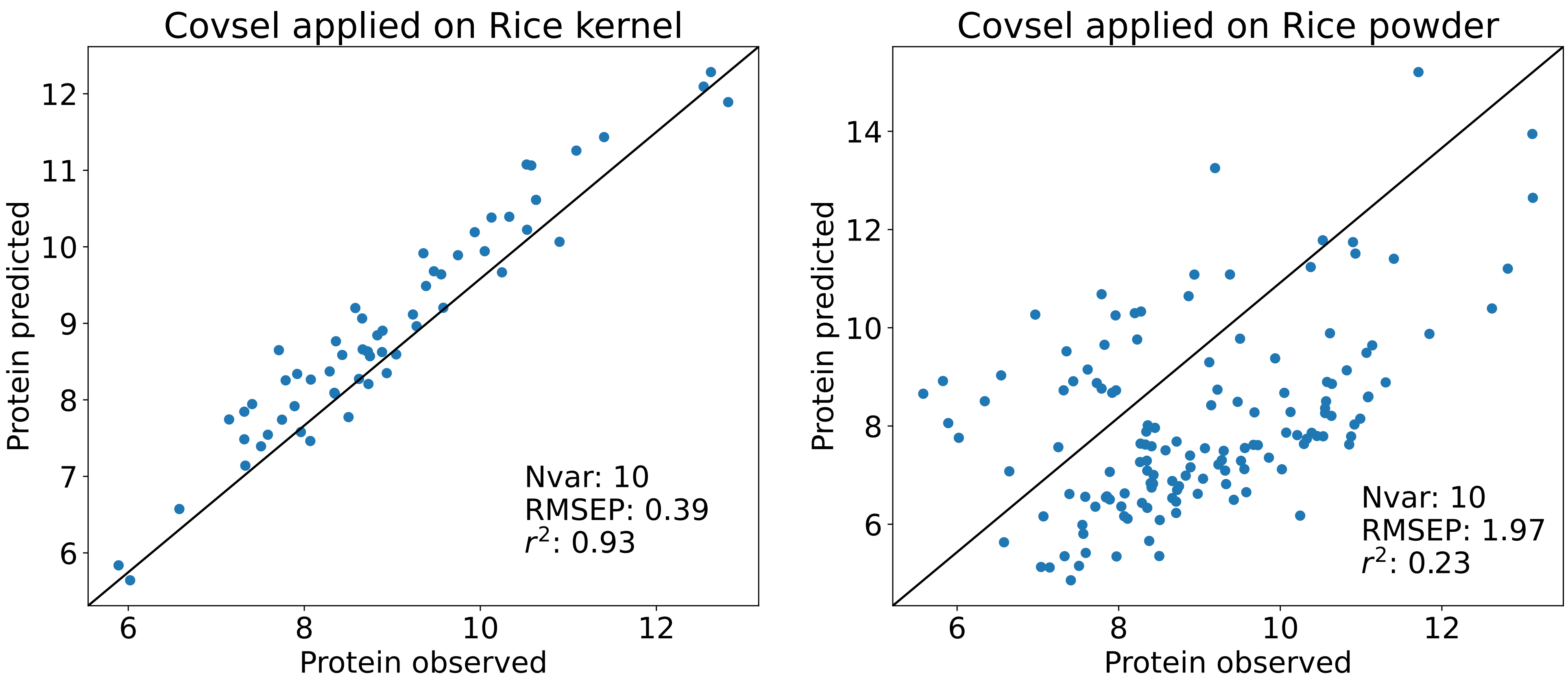
$$\max_j cov^2(a_j x_j, y)$$

where:

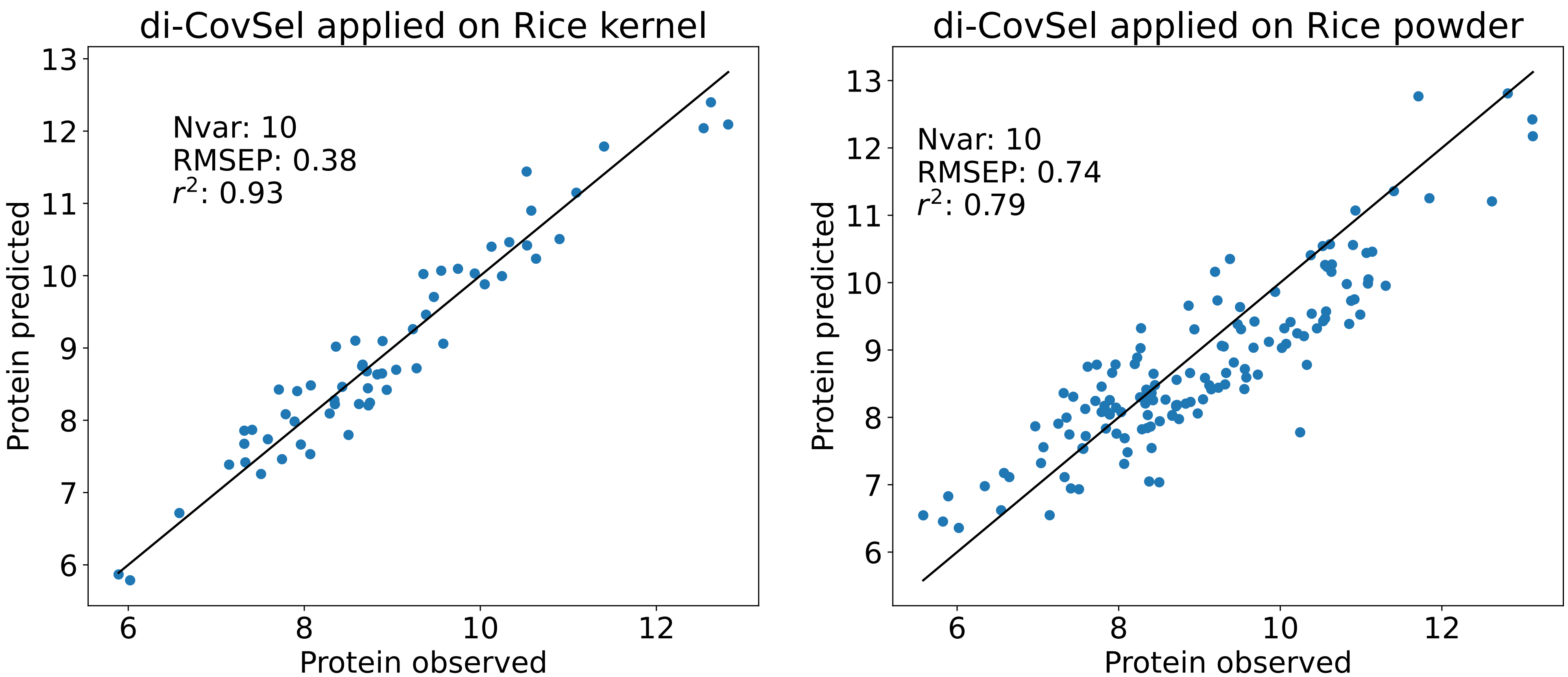
$$A = f(cov(X_s) - cov(X_t))$$



PROTEIN CONTENT PREDICTION IN TWO RICE FORMS



PROTEIN CONTENT PREDICTION IN TWO RICE FORMS



RECOMMENDATIONS

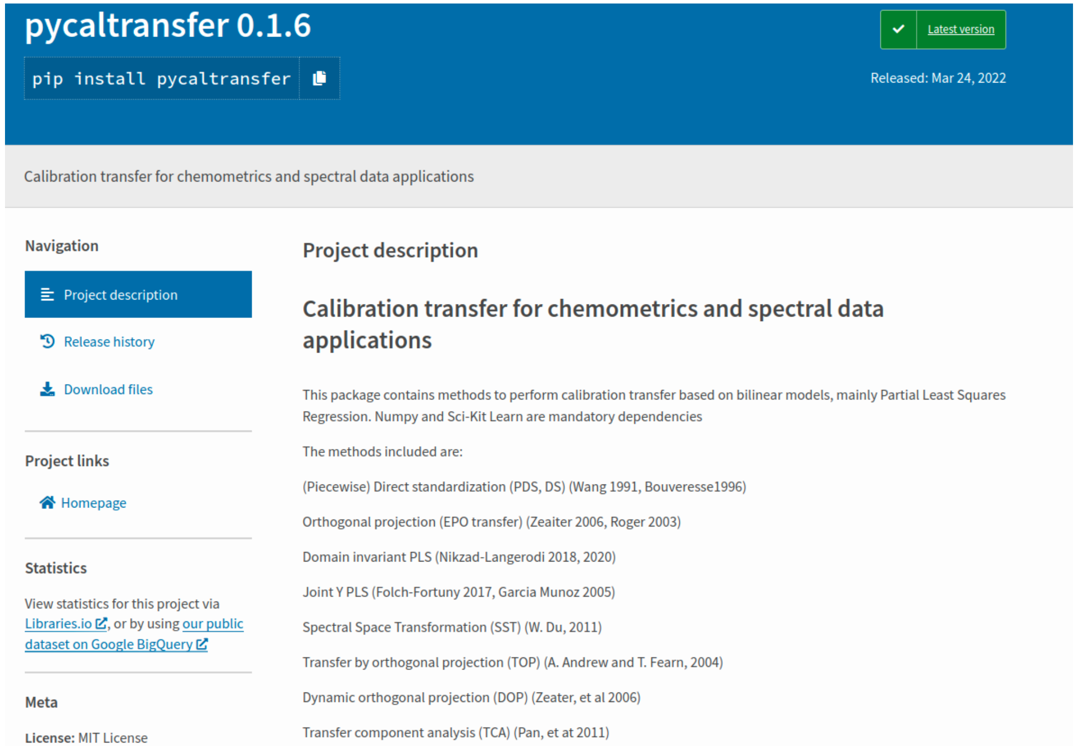
Think of calibration transfer as an adaptation of the model rather than an equalization of domains (instruments)

Study the type of degradation experienced by the source model in the target domain

Consider the different methodology categories rather than specific methods to decide on the types of samples needed

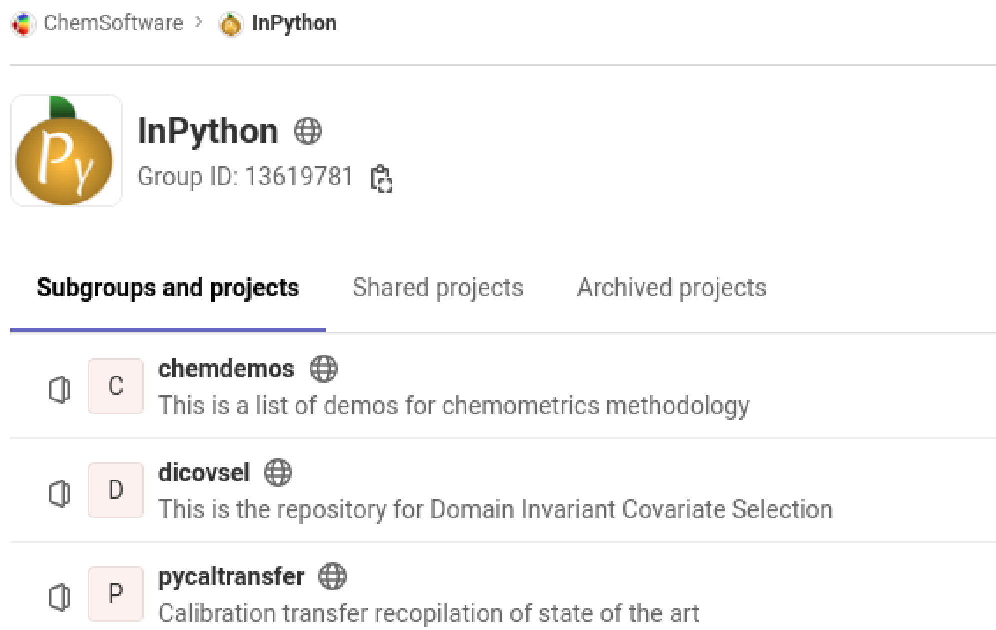
SOFTWARE AVAILABILITY

pycaltransfer



https://gitlab.com/chemsoftware/python/pycaltransfer

chemdemos



https://gitlab.com/chemsoftware/python/chemdemos

Special thanks to

FWO Research Foundation – Flanders (Belgium)

Gilles Chaix and HélioSPIR

Jean-Michel Roger and ChemHouse