Supplementary Material

Table A. Results of calibration and prediction PLS models using the test set. Values are averaged over 100 runs each with associated standard deviations. (LOOCV = Leave-One-Out Cross Validation, RanCV = Random Cross Validation)

nLV_{min}	RMSEC (g/L)	RMSEP (g/L)	R^2 Calibration	R^2 Prediction
Average intersampled Dataset (n = 850, 25 spectra per sample)				
Polynomial Baseline,Vector Normalised LOOCV				
20.	0.38 ± 0.04	0.75 ± 0.05	$0.98 \pm 5E\text{-}3$	0.90 ± 0.01
Polynomial Baseline,Vector Normalised RanCV				
20	0.41 ± 0.03	0.75 ± 0.05	$0.97 \pm 4E\text{-}3$	0.90 ± 0.01
Polynomial Baseline,EMSC LOOCV				
20	0.40 ± 0.08	0.78 ± 0.06	0.97 ± 0.01	0.90 ± 0.02
Polynomial Baseline,EMSC RanCV				
18	0.44 ± 0.07	0.78 ± 0.05	0.97 ± 0.01	0.89 ± 0.01
Linear-Nodal, Vector Normalised LOOCV				
20	0.39 ± 0.06	0.72 ± 0.05	$0.97 \pm 6E\text{-}3$	0.91 ± 0.01
Linear-Nodal, Vector Normalised RanCV				
20	0.42 ± 0.02	0.73 ± 0.06	$0.97 \pm 4E\text{-}3$	0.91 ± 0.01
Linear-Nodal, EMSC LOOCV				
21	0.38 ± 0.05	0.73 ± 0.05	$0.98 \pm 6E\text{-}3$	0.91 ± 0.01
Linear-Nodal, EMSC RanCV				
21	0.40 ± 0.02	0.73 ± 0.06	$0.97 \pm 2E\text{-}3$	0.91 ± 0.02

nLV_{min}	RMSEC	RMSEP	R^2 Calibration	R^2 Prediction
Full Dataset (n = 3400, 100 spectra per sample)				
Polynomial Baseline,Vector Normalised LOOCV				
17	0.78 ± 0.03	0.98 ± 0.02	$0.89 \pm 8E\text{-}3$	$0.83 \pm 8E\text{-}3$
Polynomial Baseline,Vector Normalised RanCV				
16	0.81 ± 0.03	0.99 ± 0.02	$0.88 \pm 9E\text{-}3$	$0.82 \pm 9E\text{-}3$
Polynomial Baseline,EMSC LOOCV				
17	0.80 ± 0.02	0.98 ± 0.02	$0.89 \pm 7E\text{-}3$	$0.83 \pm 9E\text{-}3$
Polynomial Baseline,EMSC RanCV				
16	0.82 ± 0.02	0.98 ± 0.03	$0.88 \pm 6E\text{-}3$	$0.83 \pm 9E\text{-}3$
Linear-Nodal, Vector Normalised LOOCV				
17	0.77 ± 0.02	0.97 ± 0.02	$0.90 \pm 6E\text{-}3$	$0.83 \pm 9E\text{-}3$
Linear-Nodal, Vector Normalised RanCV				
16	0.80 ± 0.03	0.98 ± 0.02	0.89 ± 0.01	$0.83 \pm 8E\text{-}3$
Linear-Nodal, EMSC LOOCV				
18	0.78 ± 0.02	0.98 ± 0.02	$0.89 \pm 6E\text{-}3$	$0.83 \pm 9E\text{-}3$
Linear-Nodal, EMSC RanCV				
17	0.80 ± 0.01	0.98 ± 0.02	$0.89 \pm 4E\text{-}3$	$0.83 \pm 9E\text{-}3$