Assessment of Dynamic Soil Carbon Pools Using Visible/Near-Infrared Diffuse Reflectance Spectroscopy (VNIRS) and Various Multivariate Methods

Gustavo M. Vasques[1], Sabine Grunwald[1], and James O. Sickman[2]
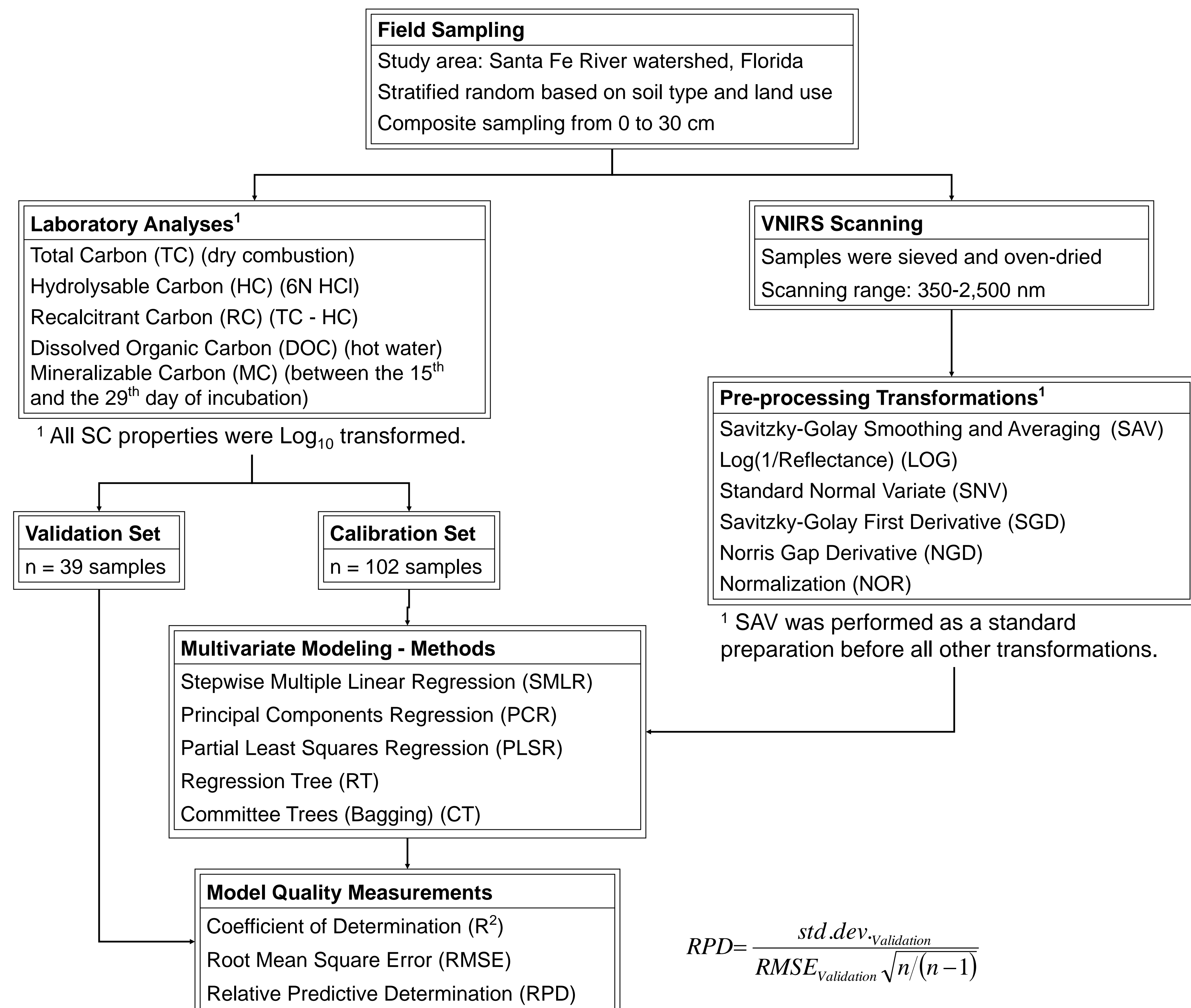
[1]Soil and Water Science Department, University of Florida; [2]Department of Environmental Sciences, University of California, Riverside

2169 McCarty Hall, P.O. Box 110290, Gainesville, FL 32611-0290
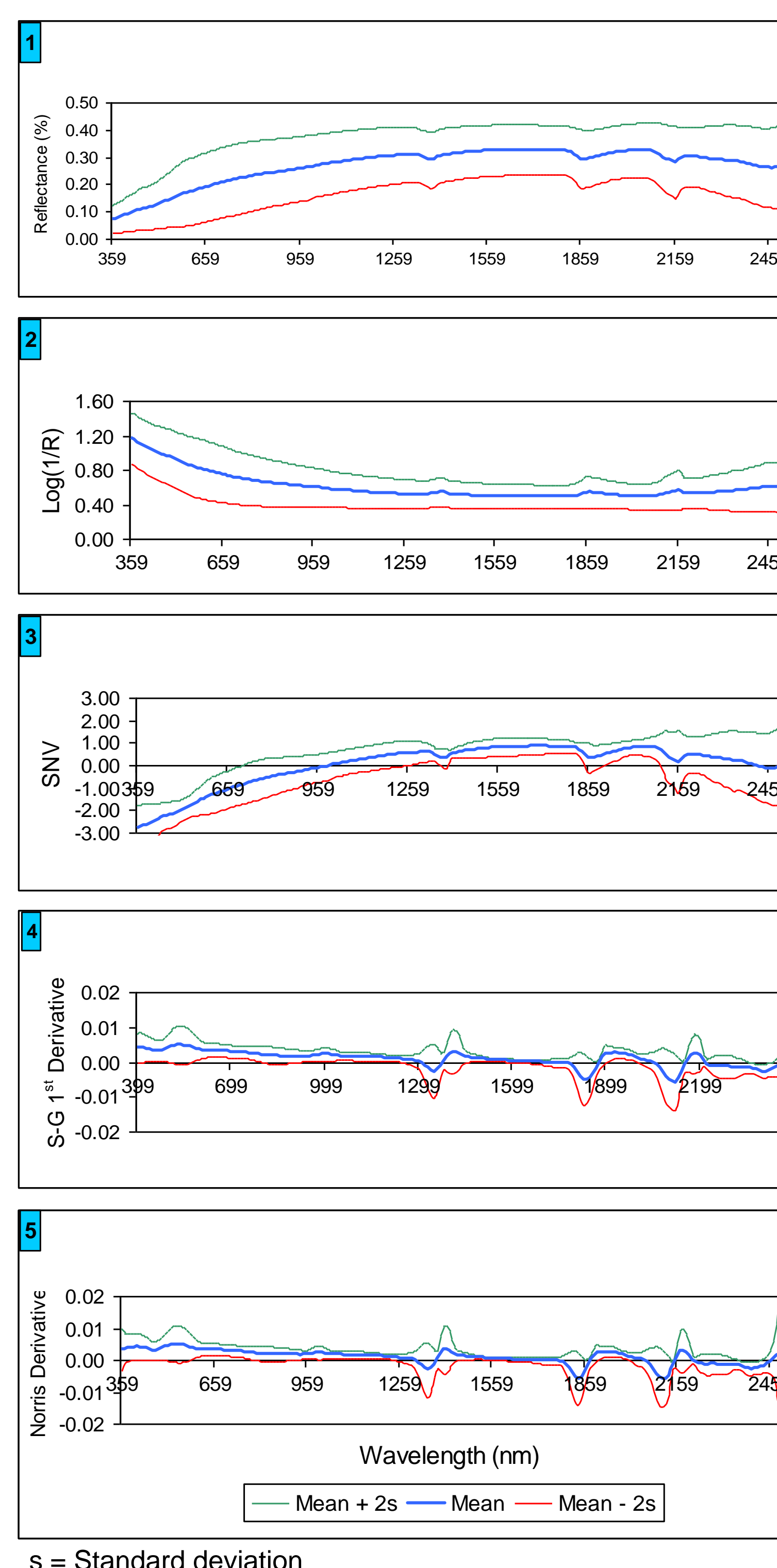Phone: 352-392-1951 ext. 233, Fax: 352-392-3902 E-mail: gvasques@ufl.edu

Introduction

Rapid, cost-effective and reliable methods are in need to assess the soil carbon (SC) pools and carbon sequestration potential at landscape scales. Visible/near-infrared diffuse reflectance spectroscopy (VNIRS) is a rapid and cost-effective method that provides inferences on multiple soil properties. Our aim was to estimate total SC and five SC chemical fractions using VNIRS comparing five multivariate parametric and non-parametric regression methods.

Methodology

Field Sampling
Study area: Santa Fe River watershed, Florida
Stratified random based on soil type and land use
Composite sampling from 0 to 30 cm

Laboratory Analyses[1]
Total Carbon (TC) (dry combustion)
Hydrolysable Carbon (HC) (6N HCl)
Recalcitrant Carbon (RC) (TC - HC)
Dissolved Organic Carbon (DOC) (hot water)
Mineralizable Carbon (MC) (between the 15th and the 29th day of incubation)

[1] All SC properties were Log10 transformed.

VNIRS Scanning
Samples were sieved and oven-dried
Scanning range: 350-2,500 nm

Pre-processing Transformations[1]
Savitzky-Golay Smoothing and Averaging (SAV)
Log(1/Reflectance) (LOG)
Standard Normal Variate (SNV)
Savitzky-Golay First Derivative (SGD)
Norris Gap Derivative (NGD)
Normalization (NOR)

[1] SAV was performed as a standard preparation before all other transformations.

Validation Set
n = 39 samples

Calibration Set
n = 102 samples

Multivariate Modeling - Methods
Stepwise Multiple Linear Regression (SMLR)
Principal Components Regression (PCR)
Partial Least Squares Regression (PLSR)
Regression Tree (RT)
Committee Trees (Bagging) (CT)

Model Quality Measurements
Coefficient of Determination (R^2)
Root Mean Square Error (RMSE)
Relative Predictive Determination (RPD)

$$RPD = \frac{std.dev_{Validation}}{RMSE_{Validation}\sqrt{n/(n-1)}}$$

Soil Spectra

s = Standard deviation

Mean + 2s — Mean — Mean - 2s

Results

Descriptive statistics of observed SC properties

Statistic	Observed Property (mg/kg)				
	TC	HC	RC	DOC	MC
Sample Size	141	141	141	141	141
Minimum	2,670	37	1,150	221	18
Maximum	201,988	29,399	181,738	8,995	1,036
Median	10,529	2,892	7,382	664	90
Mean	14,828	3,707	11,122	809	111
Std. Deviation	21,993	3,292	19,194	818	107
Skewness	6.35	4.58	6.64	7.57	5.39

Error statistics of the best VNIRS models

Property	Best Model	R^2		RMSE [log(mg/kg)]		RPD
		Calib.	Valid.	Calib.	Valid.	
TC	LOG-PLSR	0.93	0.86	0.082	0.078	2.71
HC	SAV-PLSR	0.49	0.40	0.218	0.285	1.29
RC	SAV-PLSR	0.90	0.82	0.109	0.108	2.23
DOC	SNV-PLSR	0.80	0.69	0.110	0.100	1.73
MC	SNV-PLSR	0.65	0.69	0.159	0.137	1.66

Graphs of estimated versus observed TC and SC fractions of the best VNIRS models
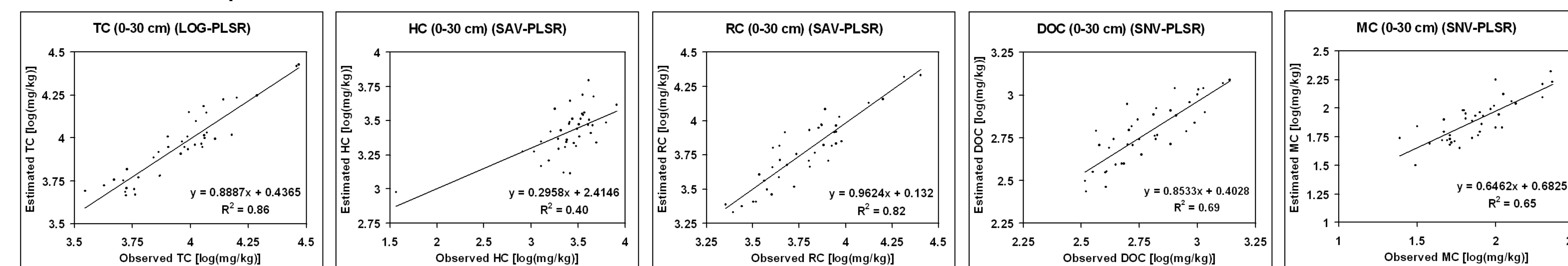
Discussion

Partial Least Squares Regression gave the most accurate estimations of SC properties at 0-30 cm. Some advantages of PLSR are: rapidness, ease of use, and flexibility to deal with correlated and missing data.

Non-parametric methods (RT and CT) are more flexible to deal with non-linearity; however RT estimated discontinuous values, thus was less suited for VNIRS modeling.

The type of multivariate method had a higher influence in the quality of model than the type of pre-processing transformation.

All the best VNIRS models of TC and SC fractions were sensitive to the regions of absorption features of C-H, O-H, N-H and H_2O (~900; 1,100-1,400; 1,600-1,800; and 2,000-2,500 nm).

Conclusions

Except for HC, VNIRS produced reliable models of TC and SC fractions.

Partial Least Squares Regression was the best method amongst all the multivariate methods tested.

Acknowledgements