

Introduction: Spiked samples from an anaerobic digestion process were measured with an NIR and a Mid-IR instrument in order to compare the capability of the instruments for quantitative measurements of four constituents in two concentration ranges, with or without sample preparation.

Materials and Methods: The NIR transmission spectra were collected with a tec5 AG diode array instrument (305-2000 nm), and the mid-IR ATR spectra with an ASI ReactIR 1000 instrument (4000-650 cm^{-1}). All calculations were done in Matlab, utilising the PLS Toolbox.

The samples were collected from 4 anaerobic batch digestions, carried out in laboratory scale reactors using a mixture of pig manure, fish, and greenhouse waste as substrate. A total of 33 samples were withdrawn from the reactors, split into 14 subsamples and spiked with solutions of salts of ammonium, acetate and propionate according to a three factor central composite design. Measurements were carried out on both the unprepared samples and the supernatant of centrifuged samples. Two spectra were collected for each sample and sample preparation. (Figure 1)

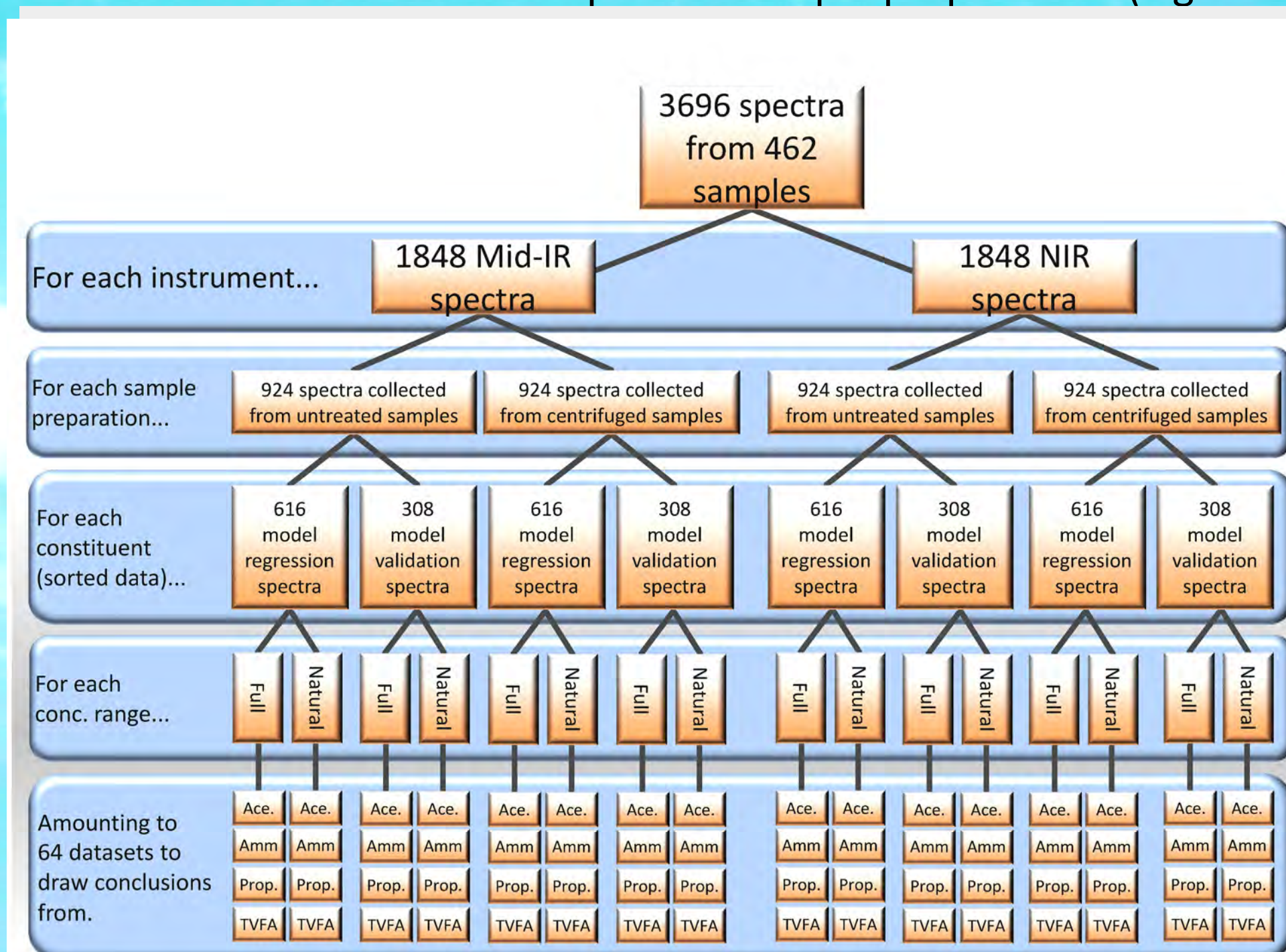


Figure 2. Spectral and constituent data sets and splits.

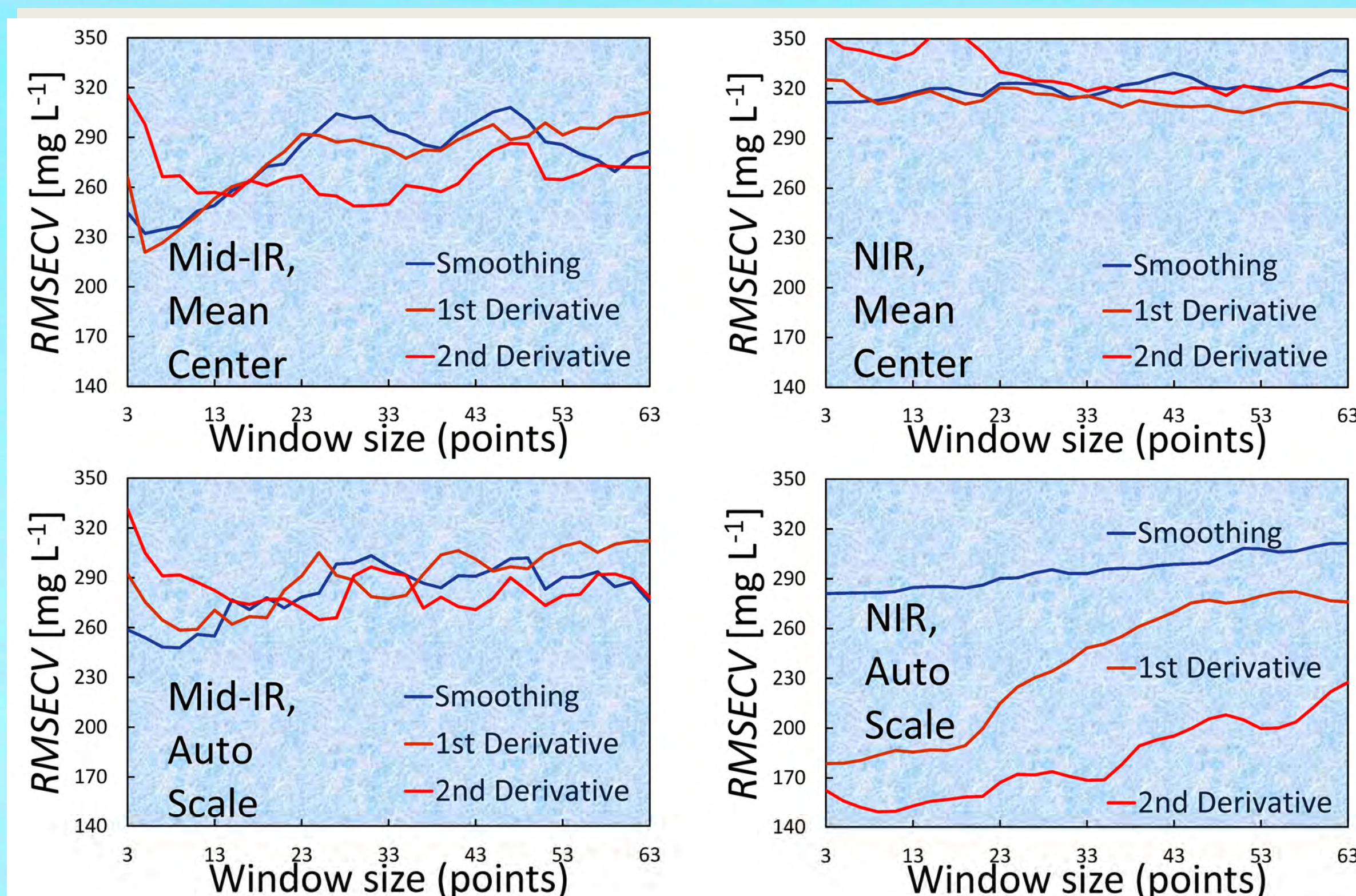


Figure 4. Impact of spectral pre-processing methods.

including the higher (noisy) NIR wavelengths when using spectral mean centering, whereas this effect is not noticeable when auto scaling is used. The actual interval selection was done manually, taking into account the results from calculations with both spectral auto scaling and mean centering.

The result from evaluating the pre-processing alternatives on the same data can be seen in figure 4. In general there was no dramatic effect from using smoothing or derivatives, as long as a wide enough window was used for the second order derivative. However, in this particular case, the impact of using a derivative with a narrow window in combination with spectral auto scaling was very significant for the NIR measurement. The *RMSEP* values for the 32 validation data sets are shown in figure 5. The accuracy was considerably higher with mid-IR compared to NIR for acetate, and the impact of the concentration range was very significant for ammonium. For total volatile fatty acids the relative error was very low, which can be seen in figure 6.

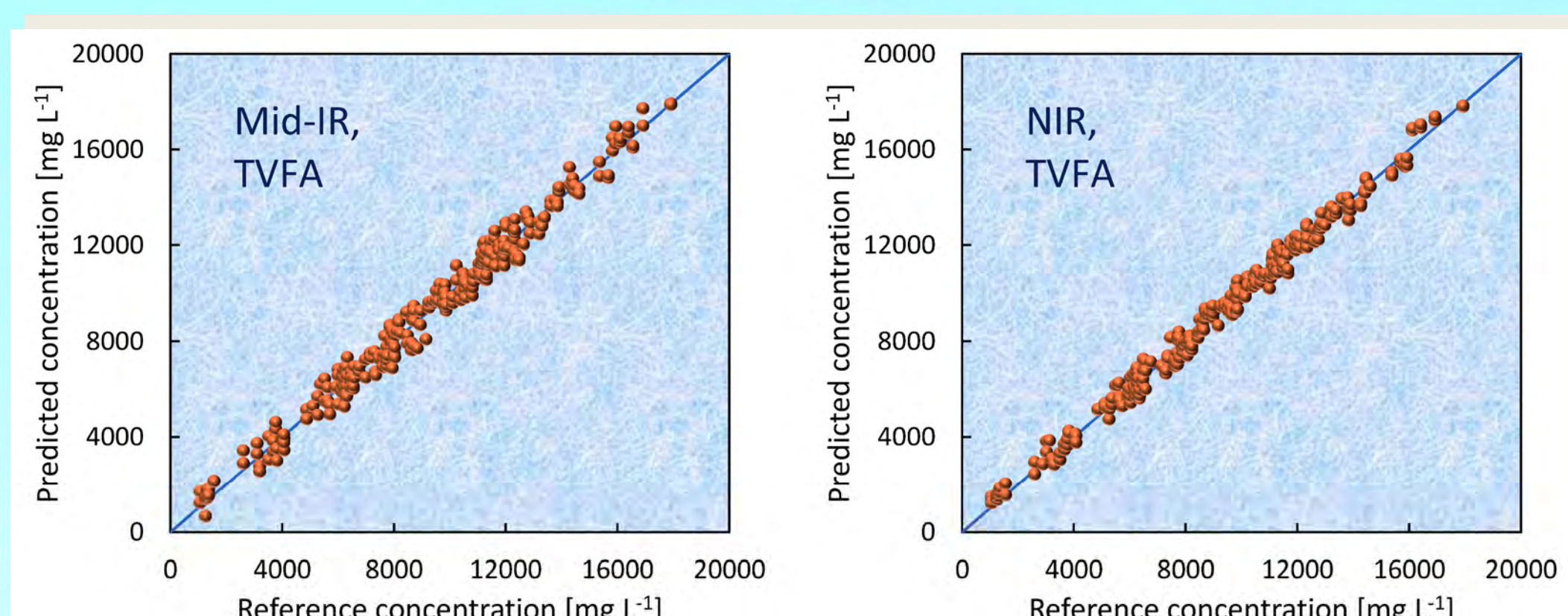


Figure 6. Reference concentration vs. mid-IR and NIR predicted concentrations of total volatile fatty acids.

The data was sorted according to constituent concentration and the spectra from every third sample was thereafter selected to the validation data set and the remaining data used for model regression. From these data sets, natural concentration range data sets were also extracted. (Figure 2)

The wavelength or wavenumber interval to be used in the PLS models was determined for each data set by selecting a narrow region based on the highest amplitude in the PLS component with the highest constituent variance explained, and thereafter evaluate the impact of a widening interval by computing a *RMSECV* value for every second point in the spectra. Thereafter the interval ends were moved to the minimum *RMSECV* value, and the procedure repeated again for 3 times. (Figure 3)

The spectral pre-processing to be used in the models was determined for each data set by recording the *RMSECV* values for models with the spectral pre-processing options of no derivative or smoothing, Savitzky-Golay smoothing or first and second order derivative, based on second order polynomials, with a window size of from 3-63 at a step size of 2, followed by either mean centering or auto scaling (138 settings in total). This step was repeated after the removal of 10% of the regression data. (Figure 4)

In order to avoid any influence of outliers and any need for outlier detection, the 10% of the calibration and validation data with the highest residuals were removed before re-evaluating the pre-processing alternatives and computing the final models, and before computing the validation statistics. In the case of NIR measurements on untreated samples, 131 spectra were also removed prior to the wavelength interval selection due to very low spectral quality.

Results and Discussion: The result from the wavelength and wavenumber calculations for propionate in the natural concentration range is shown in figure 3. For the mid-IR measurements in general, the calculations favoured using essentially the whole recorded wavenumber interval. For the NIR measurements, no improvements were generally obtained by including the information from the visible region. Also, figure 3 shows that the *RMSECV* value increases significantly by



Figure 1. Experimental procedure and equipment.

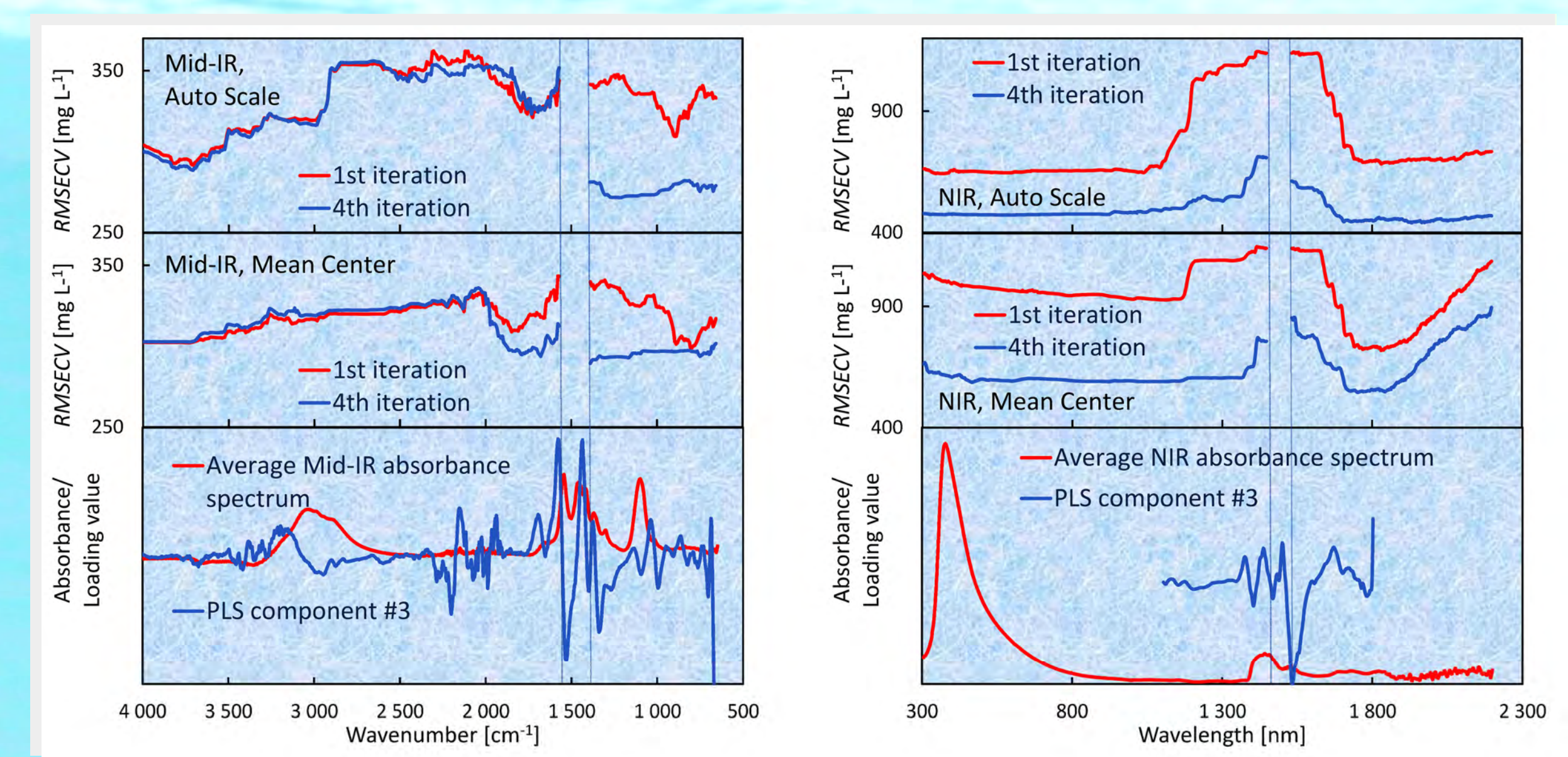


Figure 3. *RMSECV* calculations performed to determine the spectral interval to be used in the PLS models.

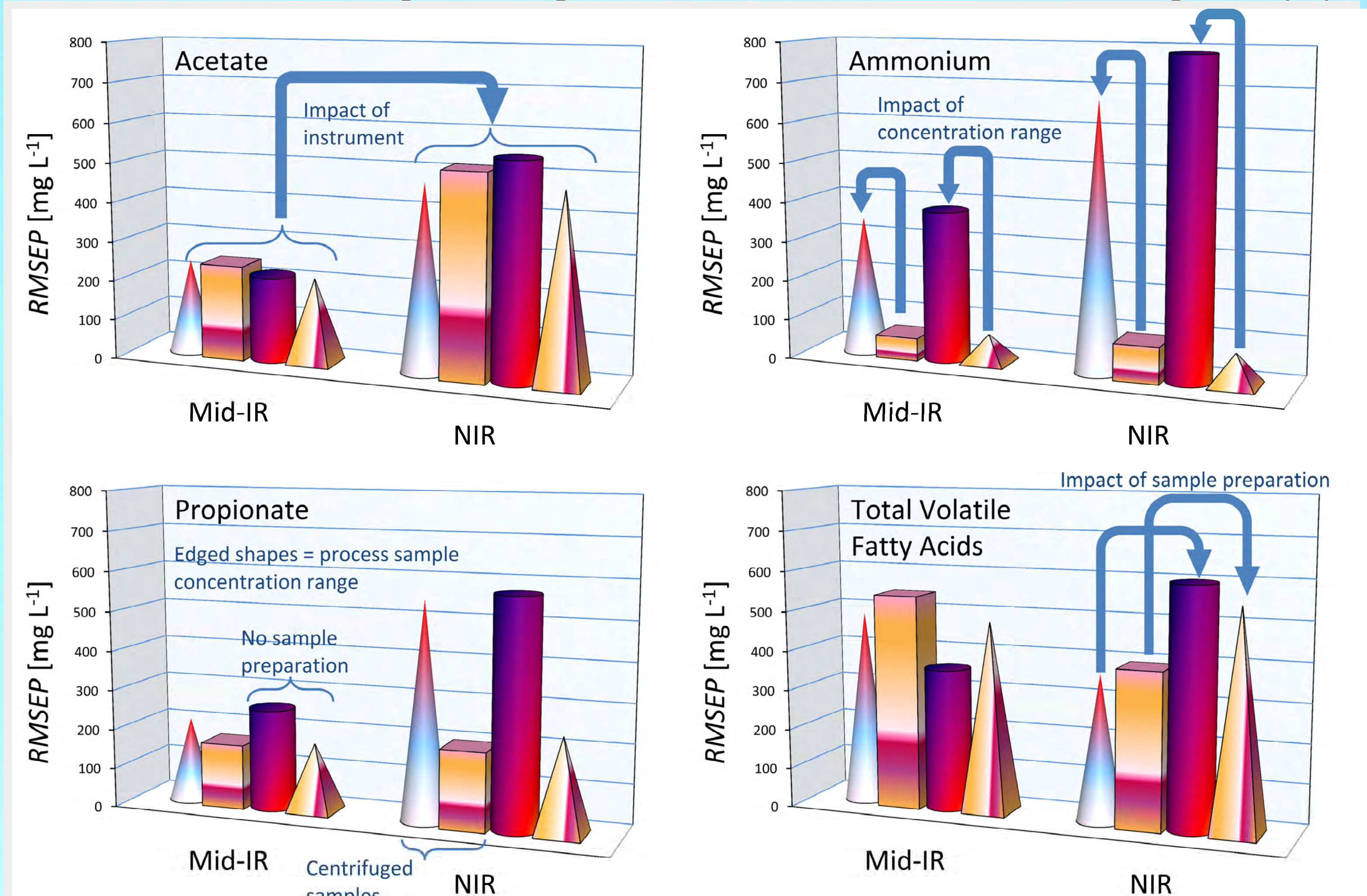


Figure 5. *RMSEP* for the 32 validation data sets.

Conclusions: On the samples measured, with the measurement setup used, and the constituents studied, mid-IR attenuated total reflection spectroscopy generally outperformed

near infrared transmittance spectroscopy. However, practical and economic considerations could still favour the use of NIR spectroscopy. It can also be pointed out that no attempt was made to optimise the measurement accuracy. The approach was instead to create all calibration models using exactly the same methodology in order to enable a comparison between the instruments.

Acknowledgements: This work was made possible through the Mare Purum project, funded by the European territorial cooperation programme Botnia-Atlantica.

Matlab scripts will be uploaded to: www.mare-purum.eu