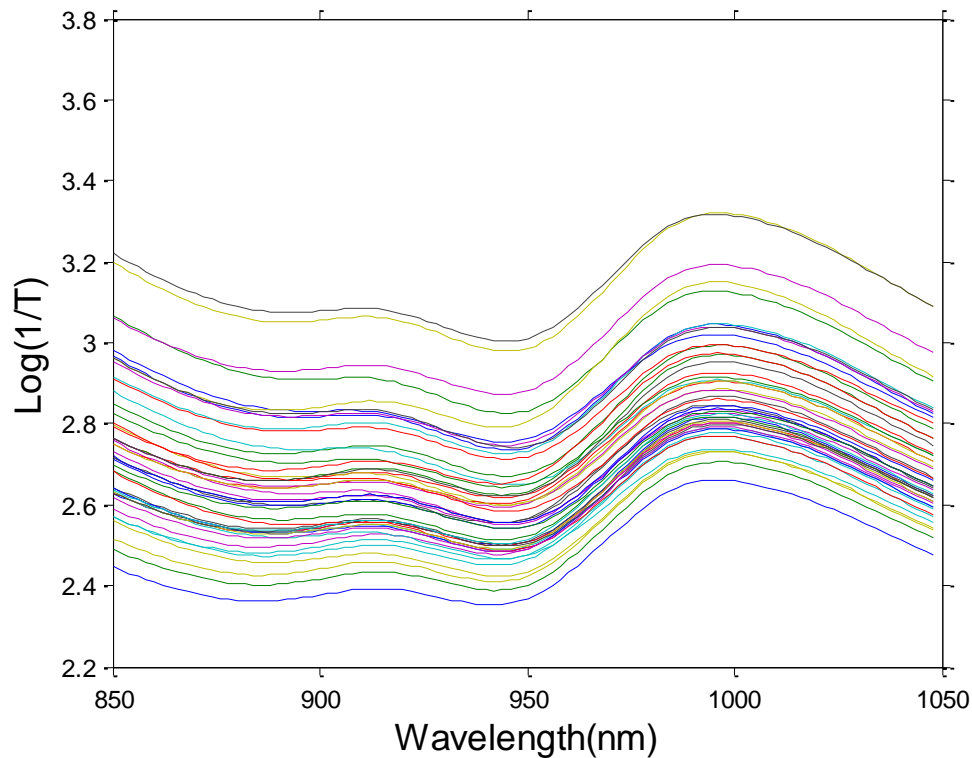


The Pros and Cons of Pre-treating Spectra

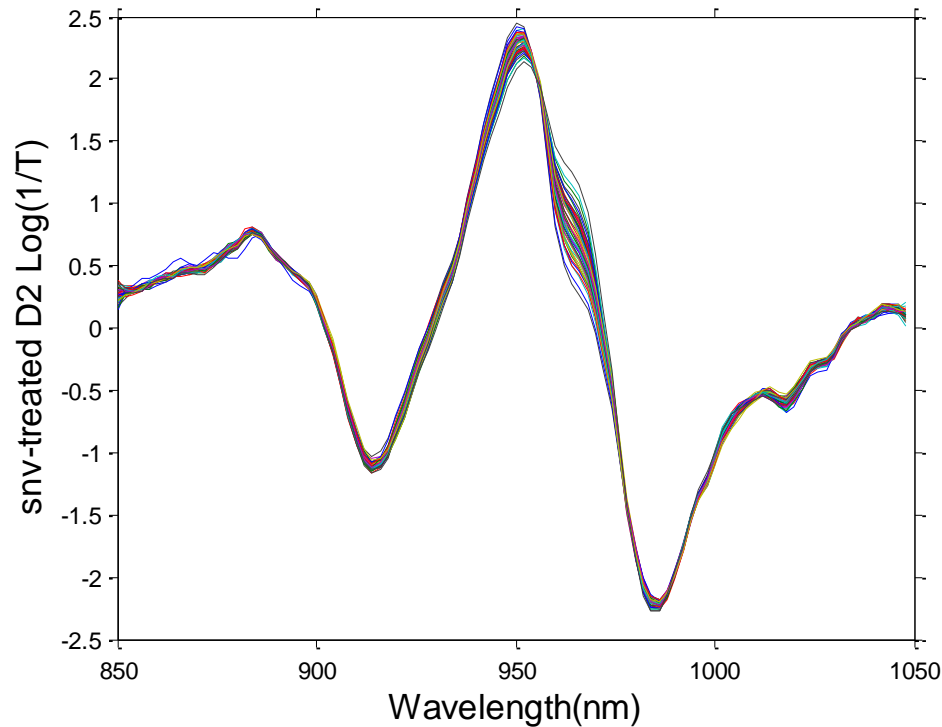
Tom Fearn, UCL
t.fearn@ucl.ac.uk

Pre-treatment of near infrared spectra



This plot shows 49 NIR transmission spectra of wheat measured in a Tecator instrument. Most of the variability in the spectra reflects physical rather than chemical properties of the wheat.

Pre-treatment of near infrared spectra (2)



The same 49 spectra after pre-treatment with second derivative and then SNV. These spectra are not only nicer to look at, but they give a better PLS calibration for protein, with RMSECV down from 0.35 (6 factors) to 0.28 (2 factors).

Pros

Spectra look nicer. Probably more chance of

- seeing where (at least some of) the information is,
- detecting bad spectra

Removing irrelevant variability can give simpler, and possibly better, calibrations.

An extra benefit may be that the calibrations are easier to transfer to another instrument.

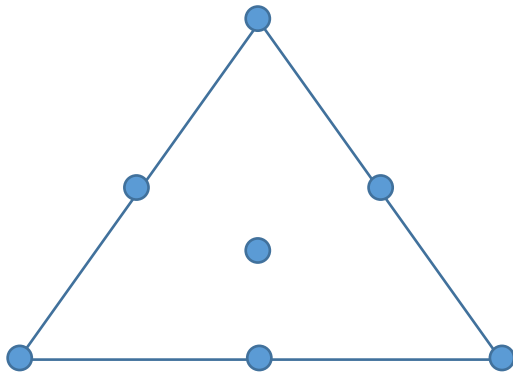
Cons

All (?) pre-treatments remove information. It might have been useful information.

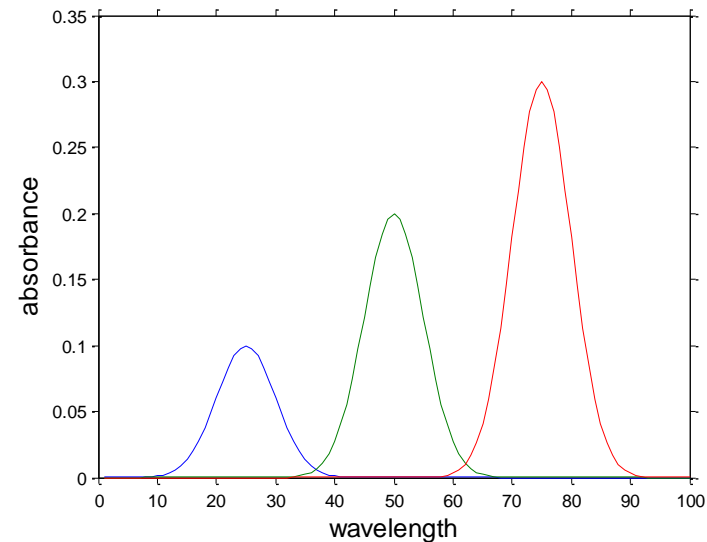
Multiplicative pre-treatments in particular

- will destroy linearity and additivity if they are present,
- move spectral information around
- introduce artefacts into the spectra

Multiplicative pre-treatments spoil Beer-Lambert

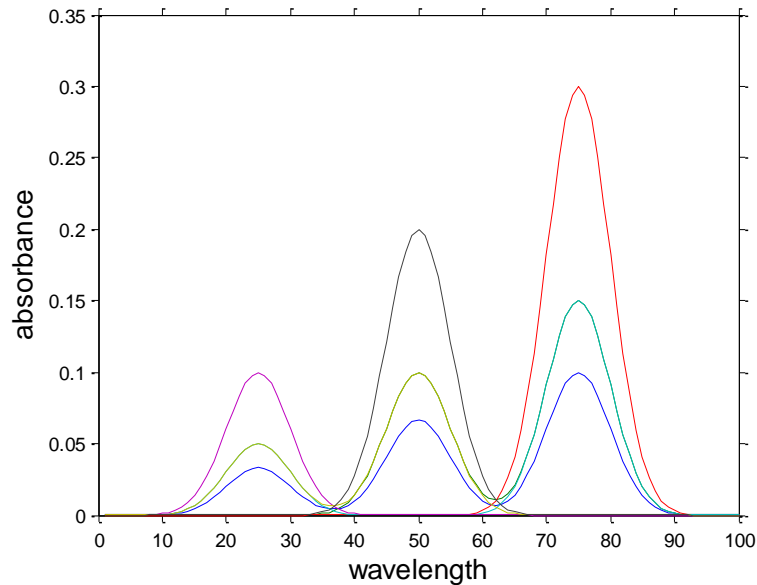


Design: 3 pures, 3 equal binary mixtures, 1 equal ternary mixture

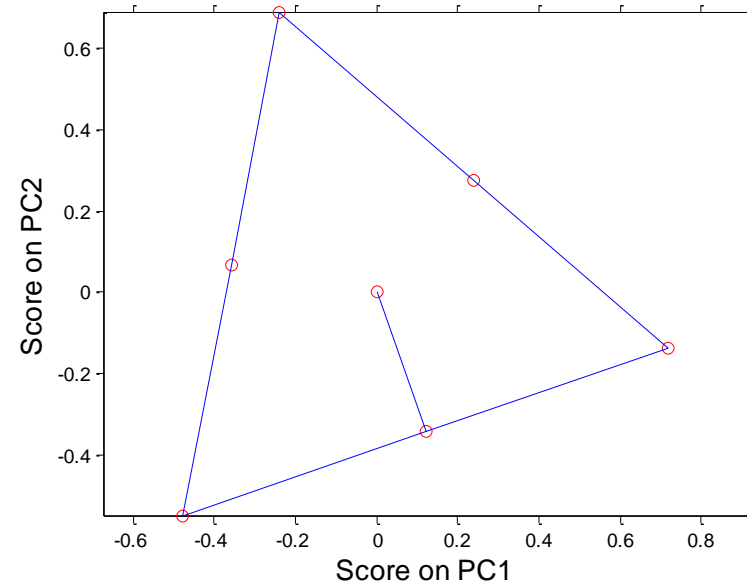


Spectra of the three ingredients in the mixtures, one peak per ingredient

Multiplicative pre-treatments spoil Beer-Lambert (2)

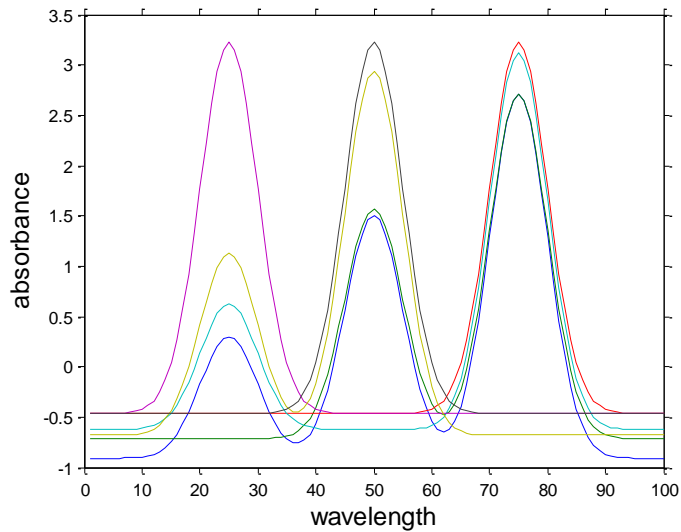


7 mixture spectra

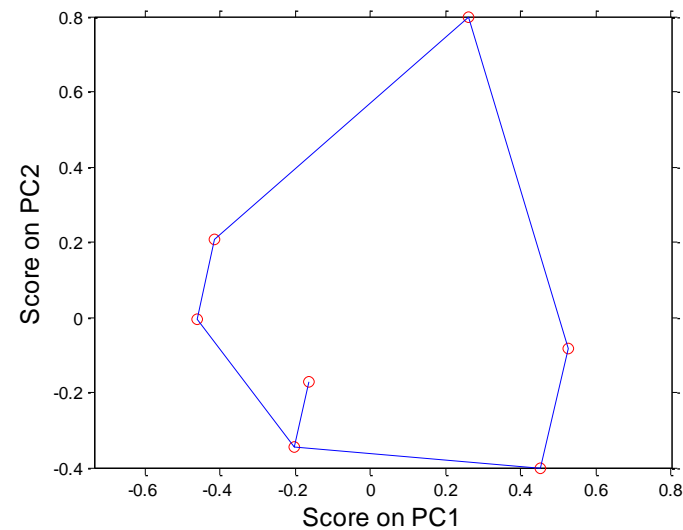


Perfect recovery of design by PCA, with 100% of variability in 2 PCs. Three wavelength MLR calibration is perfect.

Multiplicative pre-treatments spoil Beer-Lambert (3)

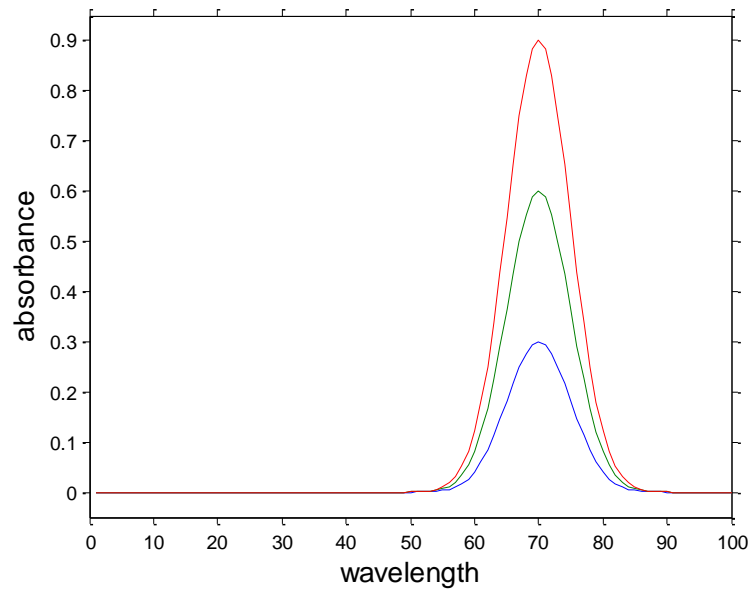


7 mixture spectra, treated with SNV

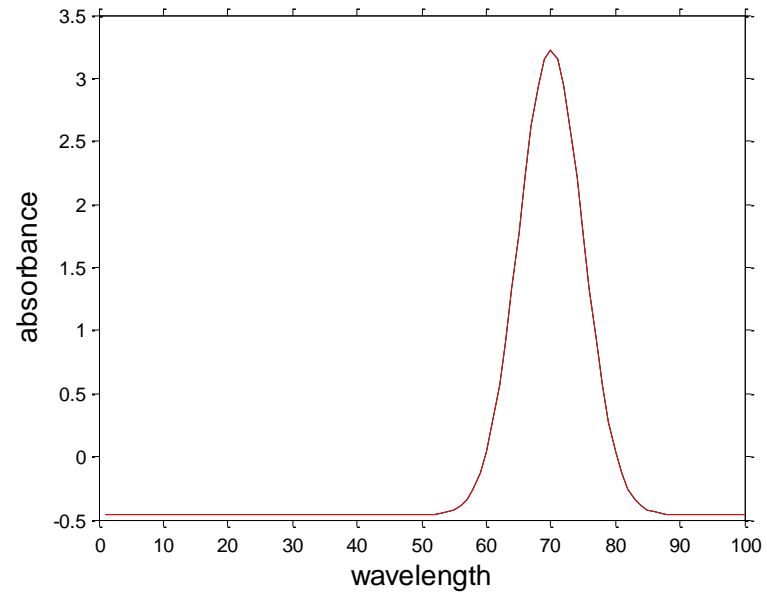


Imperfect recovery of design by PCA, and now only 97% of variability in 2 PCs. Three wavelength MLR calibration has errors of up to 40%.

Pre-treatment may destroy useful information

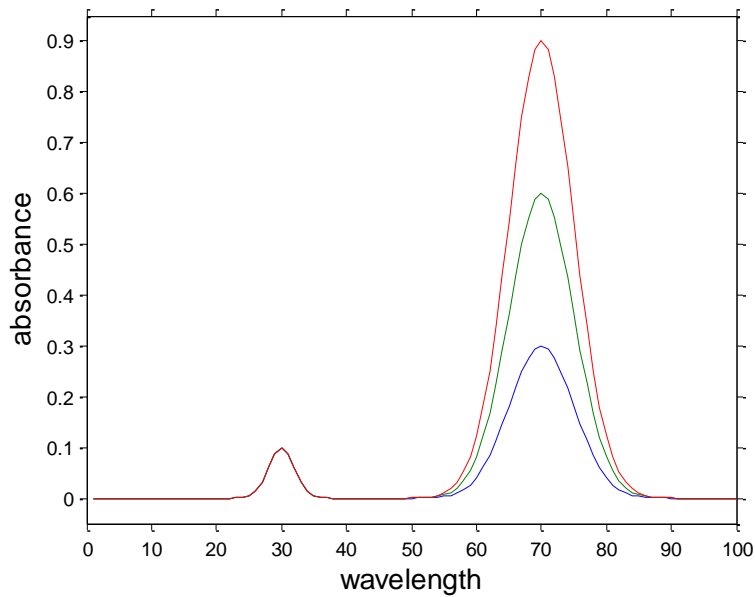


Before SNV

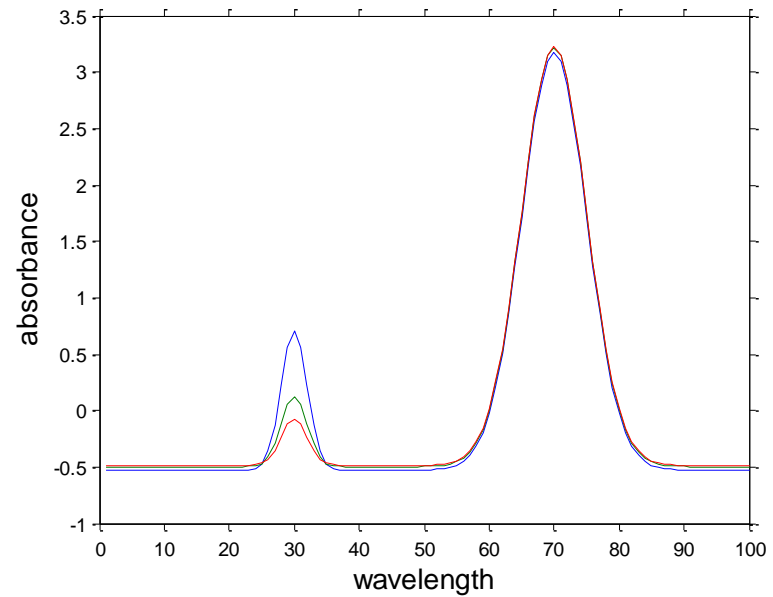


After SNV

Pre-treatment may shift information

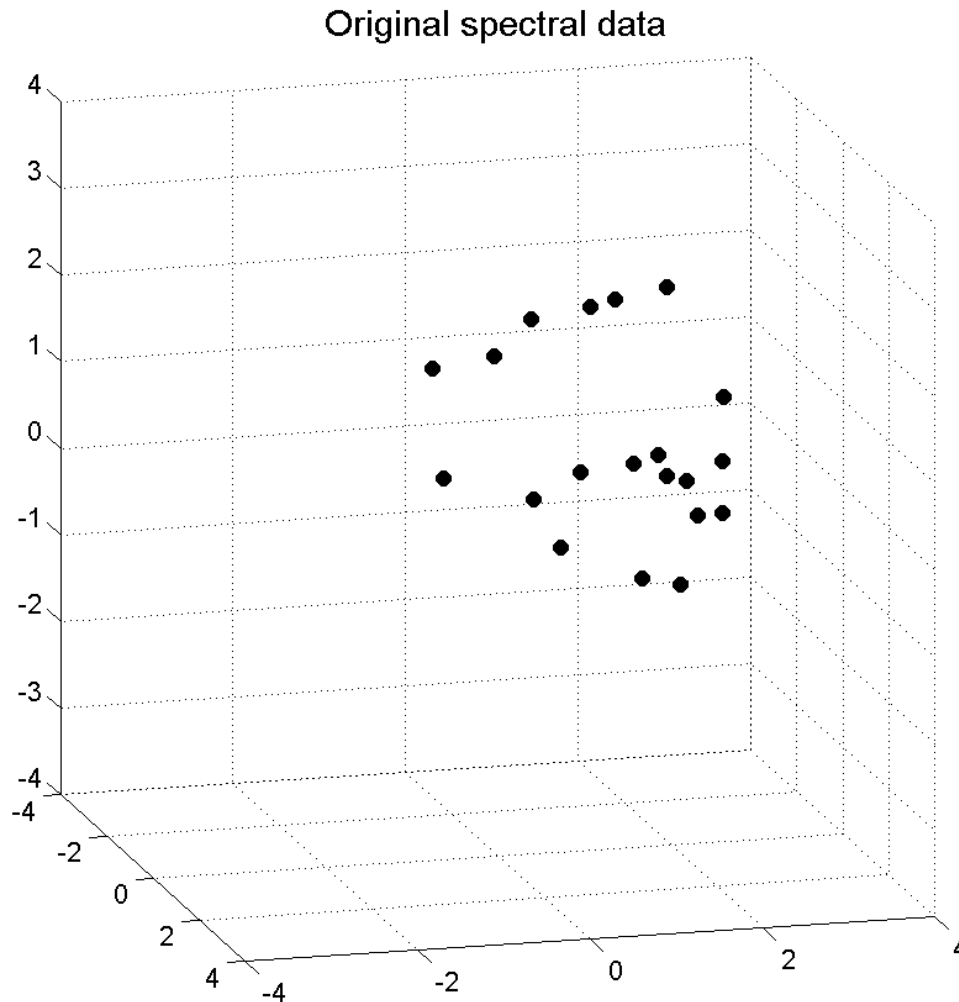


Before SNV

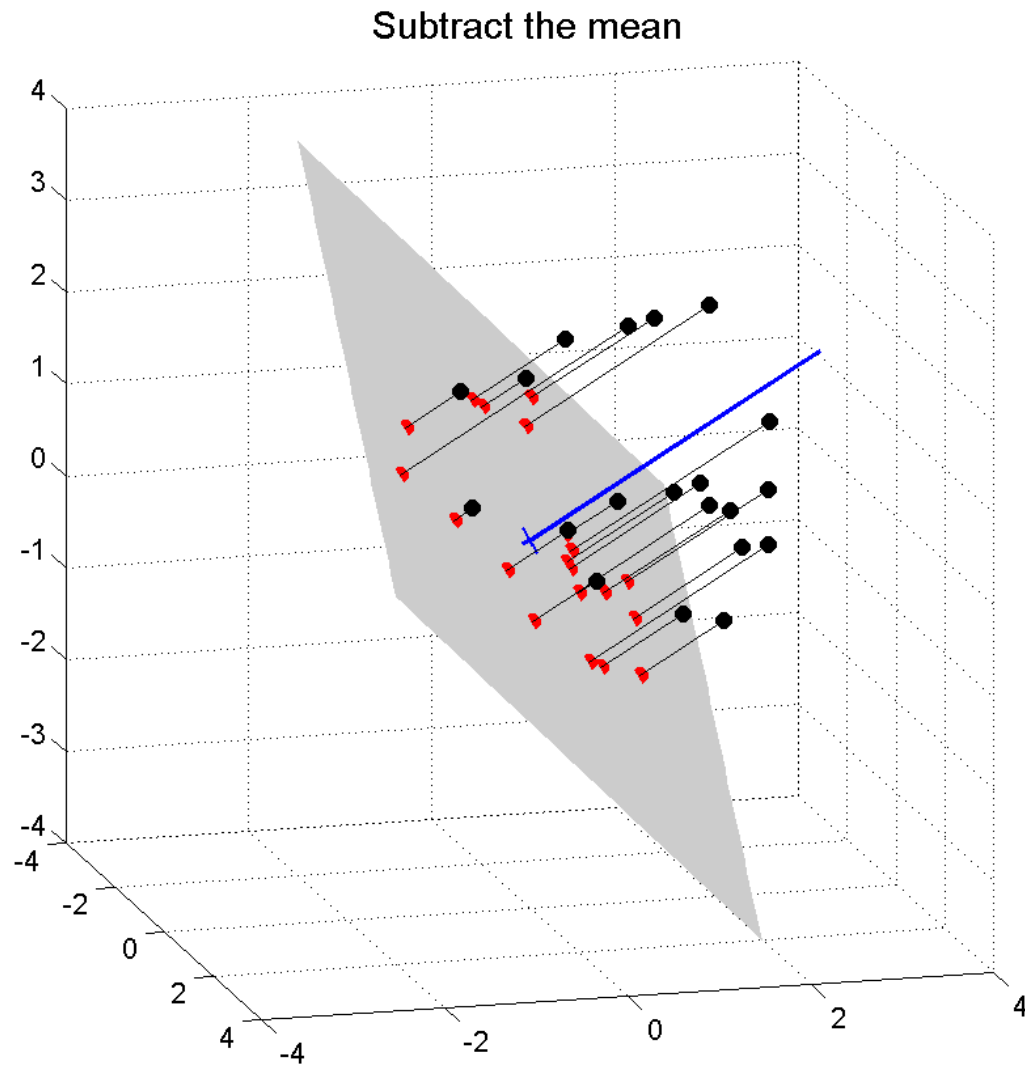


After SNV

Pre-treatment may create artefacts in the data

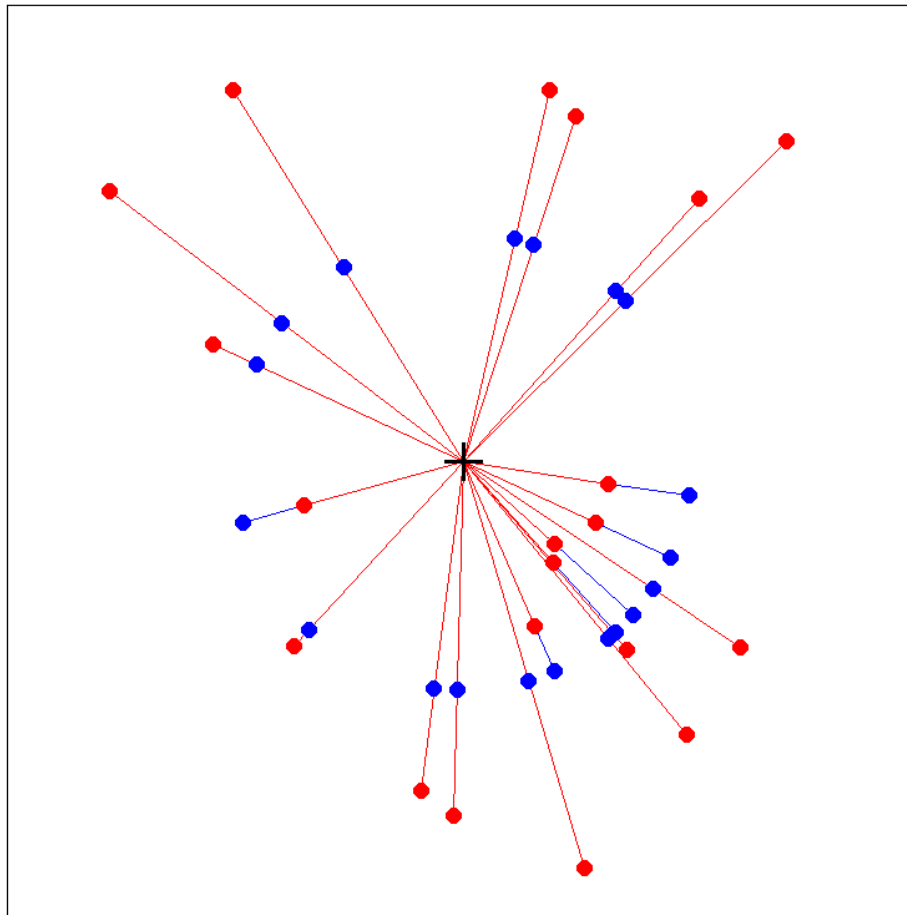


Pre-treatment may create artefacts in the data (2)

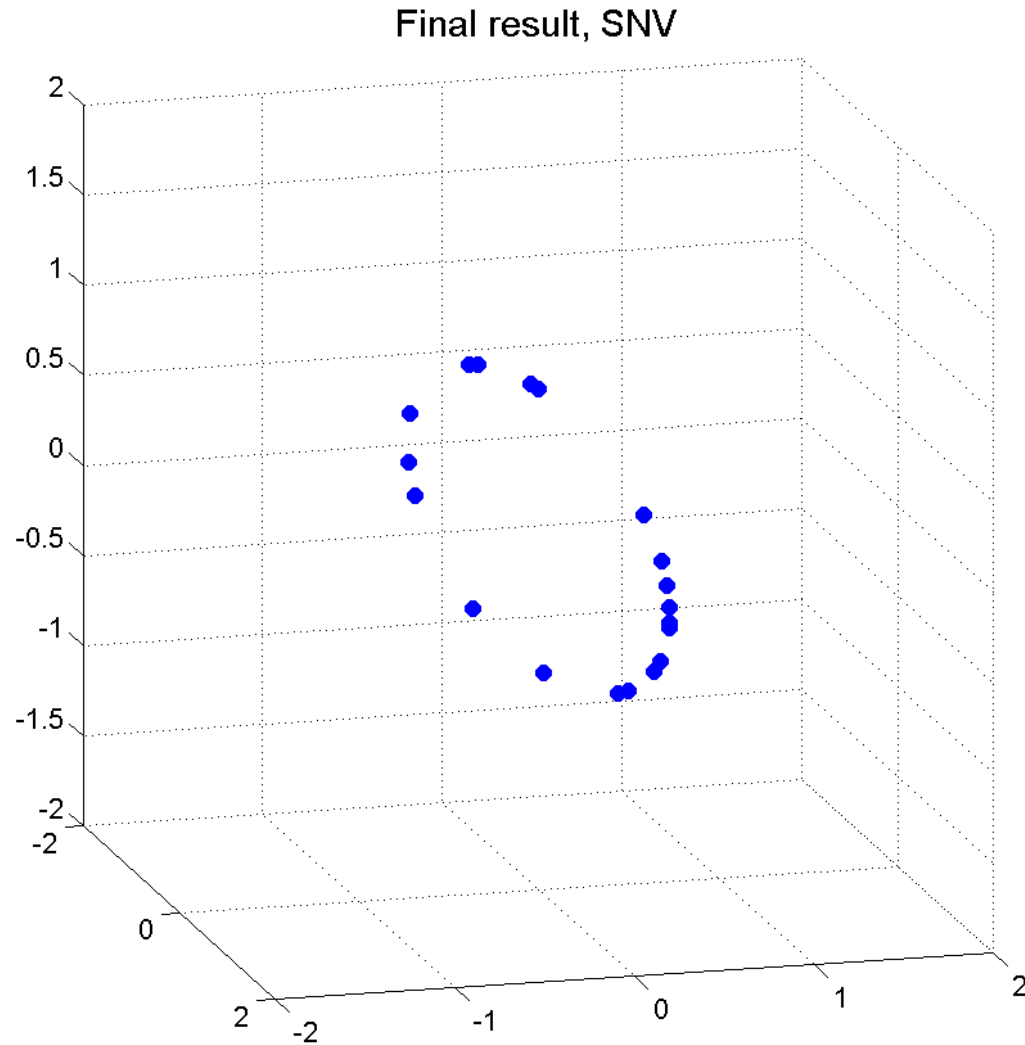


Pre-treatment may create artefacts in the data (3)

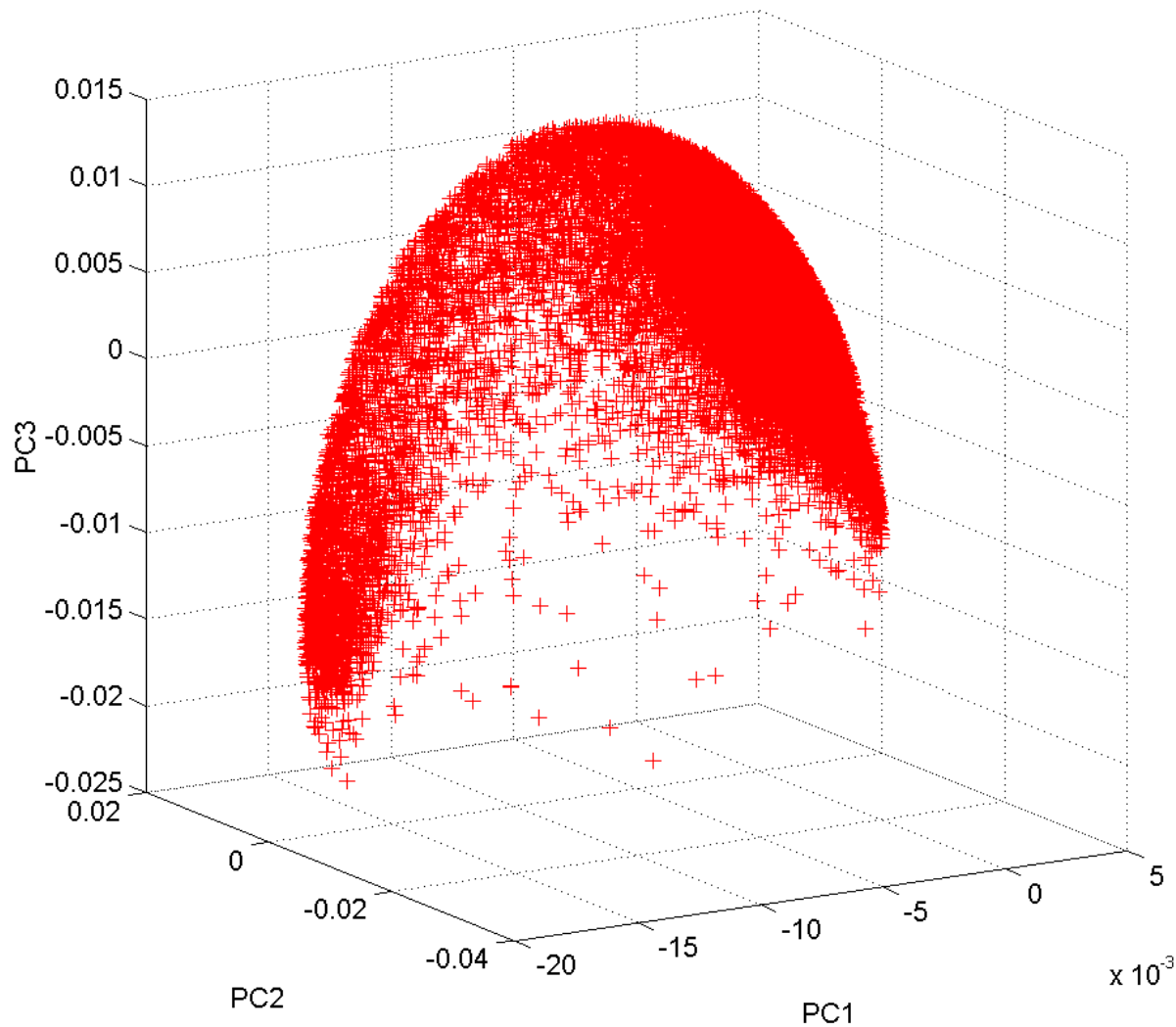
Geometry of SNV



Pre-treatment may create artefacts in the data (4)



Pre-treatment may create artefacts in the data (5)



Final comments

I'm not suggesting that we stop pre-treating spectra, but we need to remember the distortion caused by multiplicative pre-treatments in particular when it comes to interpreting the results.

A couple of references

Fearn – The effect of spectral pre-treatments on interpretation, *NIR news*, **20**, 6, 15-16 (2009).

Fearn, Riccioli, Garrido Varo, Guerrero Ginel, On the geometry of SNV and MSC, *Chemolab*, **96**, 22-26 (2009).