

# Infrared Spectrometric Purity Control of Chemical Substances using R

Fayaz Kondagula<sup>a</sup> and Karl Molt<sup>b</sup>

## 1 Introduction

A typical Infrared Spectrum contains about 4000 data points and computer software is needed for processing the spectral data. Instrument manufacturers normally provide proprietary software for this purpose. But this is often limited to a restricted number of common applications. To enable a more flexible and universal numerical and statistical evaluation of spectral data, we have developed methods to directly read the spectra into R. If the spectrum of its pure reference is known, there are two ways for determining the purity of a potentially contaminated substance. The first way is to calculate the correlation coefficient between the spectrum of the substance whose purity has to be controlled and the corresponding reference spectrum and the second one is using difference spectroscopy.

## 2 Spectral Purity by spectral comparison with linear regression

### 2.1 Simulated example (Fig.1)

A simple way of calculating the spectral purity of a substance is regressing its spectrum on the spectrum of the corresponding pure reference substance [2] using the R function `lm(substance~reference)`. The resulting correlation coefficient is used as a measure for the spectral purity:

$$SPR_1 = 100 \times R \quad (1)$$

The correlation coefficient however has the disadvantage that with increasing impurity it only slowly moves away from one. This "convex" behaviour is shown in Fig. 2(a).

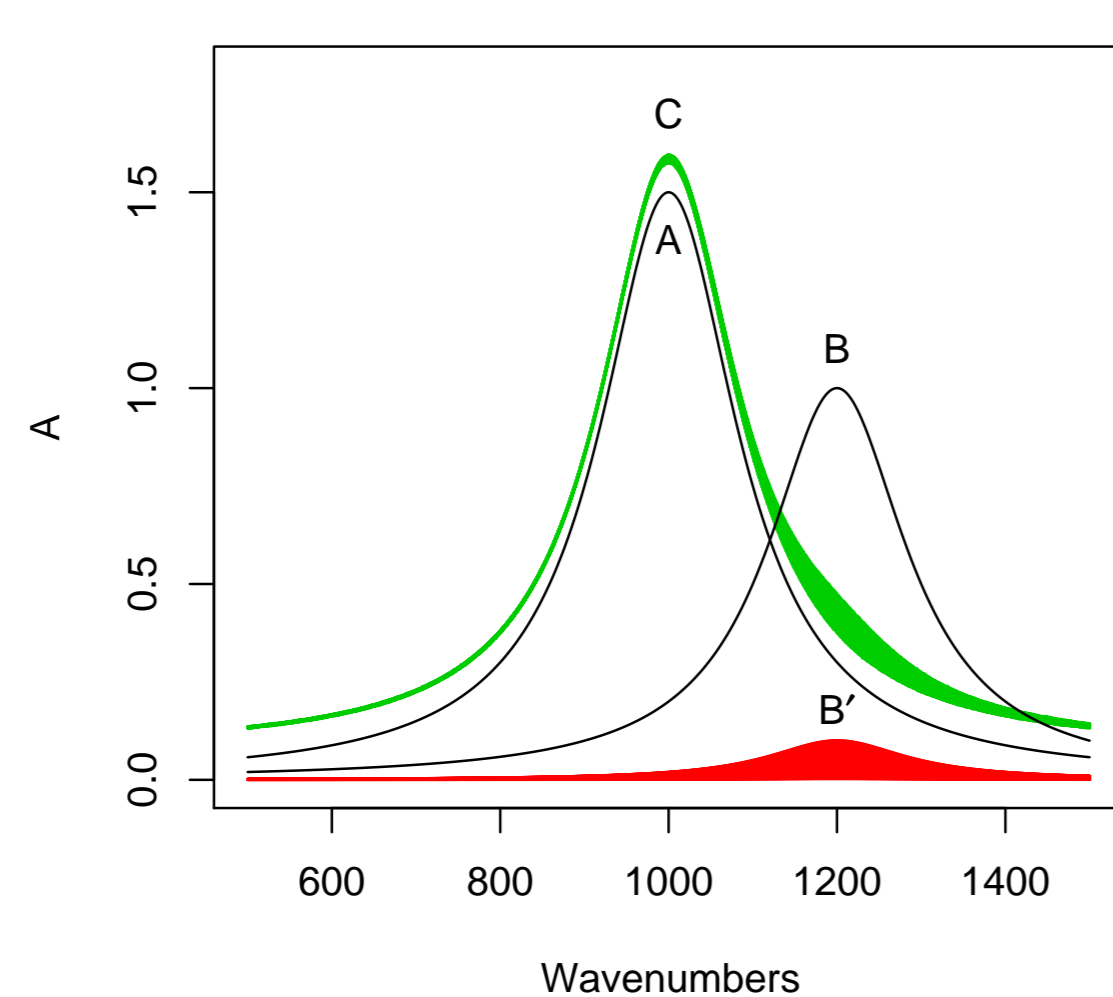
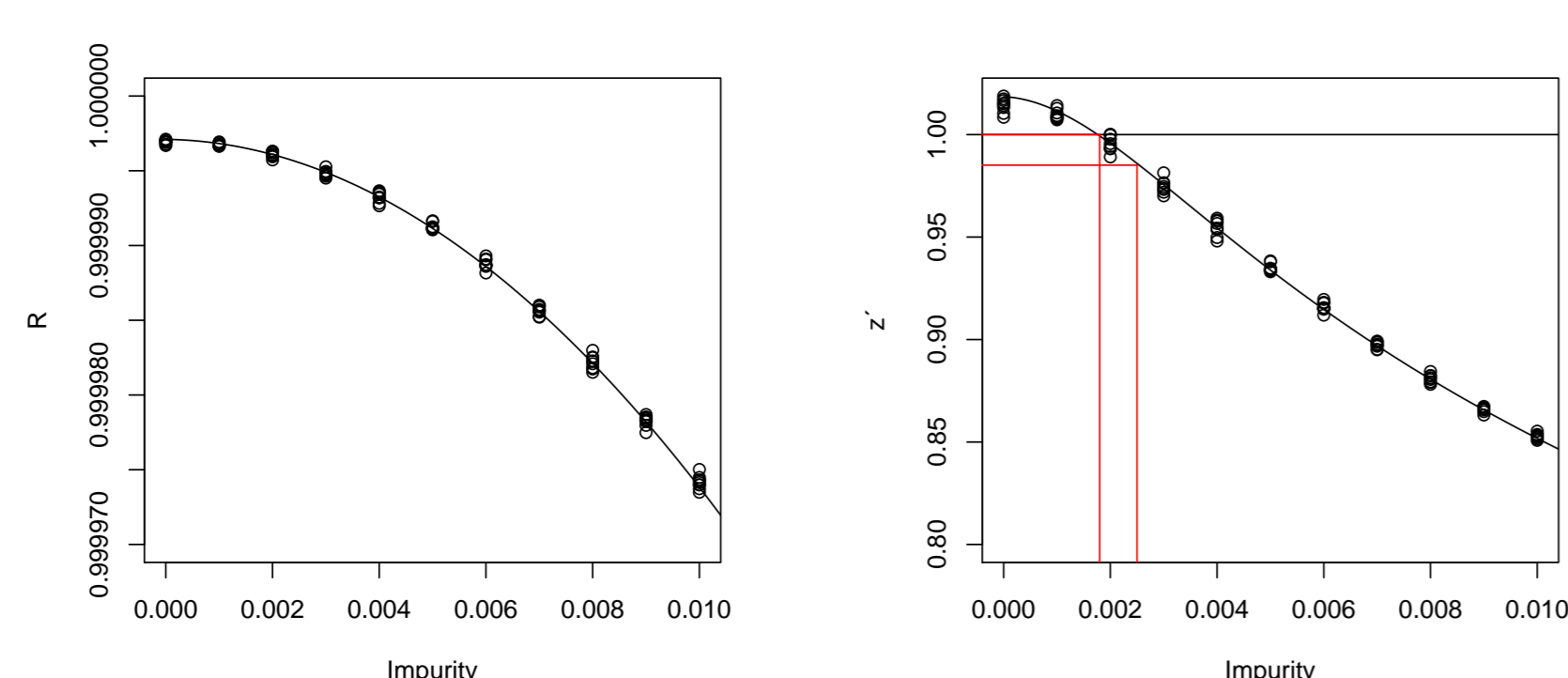


Figure 1: Simulated example of a spectral impurity. A is the spectrum of the pure component (average from 9 spectra replicates) and B the spectrum of the contaminating substance. B is added to the pure component A stepwise from 0.001 - 0.01 concentration units (spectra B'). C shows the spectra of the corresponding mixtures, on which some noise (amplitude 0.001) was superimposed.

Compared to  $R$  the  $z$  coefficient due to Fisher [3] reacts more sensitive to impurities. The  $z$ -transformation was performed using the R function `fisherz` included in the package `survcomp`:

$$z = \frac{1}{2} \times \ln \frac{1+R}{1-R} \quad (2)$$

Fig. 2(b) demonstrates that  $z$  shows a "concave" behavior which is favorable for purity control because even small impurities have a large effect.



(a) Dependency of the correlation coefficient (R) from the degree of contamination. (b)  $z'$  values calculated by transforming the R values in the left figure.

Figure 2: Results obtained by regressing the spectra of the mixtures C in Fig. 1 on the reference spectrum A. The impurity concentration belonging to  $z_c$  resp.  $z_{cl}$  corresponds to the limit of detection (LOD) resp. the limit of capture (LOC).

Repetitively measuring the spectra of the pure reference substance and then comparing the single spectra with the mean spectrum allows the calculation of a one-sided confidence interval for  $z$ . Its lower limit is as follows:

$$z_c = \bar{z} - ts_z$$

with  $t = \Phi_t^{-1}(1 - \alpha)$  and  $\alpha = 0.01$ .  $z_c$  determines the limit of detection (LOD) and  $z_{cl} = \bar{z} - 2ts_z$  the limit of capture (LOC). A normalization of  $z$  is then performed

<sup>a</sup>Fayaz Kondagula, M.Sc., Universität Duisburg-Essen, Fachgebiet Instrumentelle Analytik, 47048 Duisburg, fayazpharm@yahoo.co.in  
<sup>b</sup>Prof. Dr. Karl Molt, Universität Duisburg-Essen, Fachgebiet Instrumentelle Analytik, 47048 Duisburg, karl.molt@uni-due.de

as follows

$$z' = \frac{z}{z_c} \quad (3)$$

The corresponding spectral purity parameter is calculated as

$$SPR_2 = 100 \times z' \quad (4)$$

## 3 Spectral Purity based on Difference Spectroscopy

Another way of determining the spectral purity is by calculating the difference between the sample spectrum ( $C$ ) and the reference spectrum ( $A$ ):

$$D = C - f \times A \quad (5)$$

$f$  should be chosen in such a way that an optimal compensation of the reference spectrum is achieved. A special algorithm was developed for determining the optimal difference factor  $f_{opt}$ . This works by calculating a whole series of difference spectra with decreasing difference factors and determining the integral of the positive part of the difference spectra ( $\chi = \int D^+$ ). From this the so called "Difference Operating Characteristic" (DOC) is calculated by applying a lagged difference of second order on  $\chi$  using the R function `diff(Chi, differences=2)`. Fig 3(a) shows a plot of the DOC against the factor  $f$ . The optimal factor is where the DOC exceeds a certain threshold set by the user.

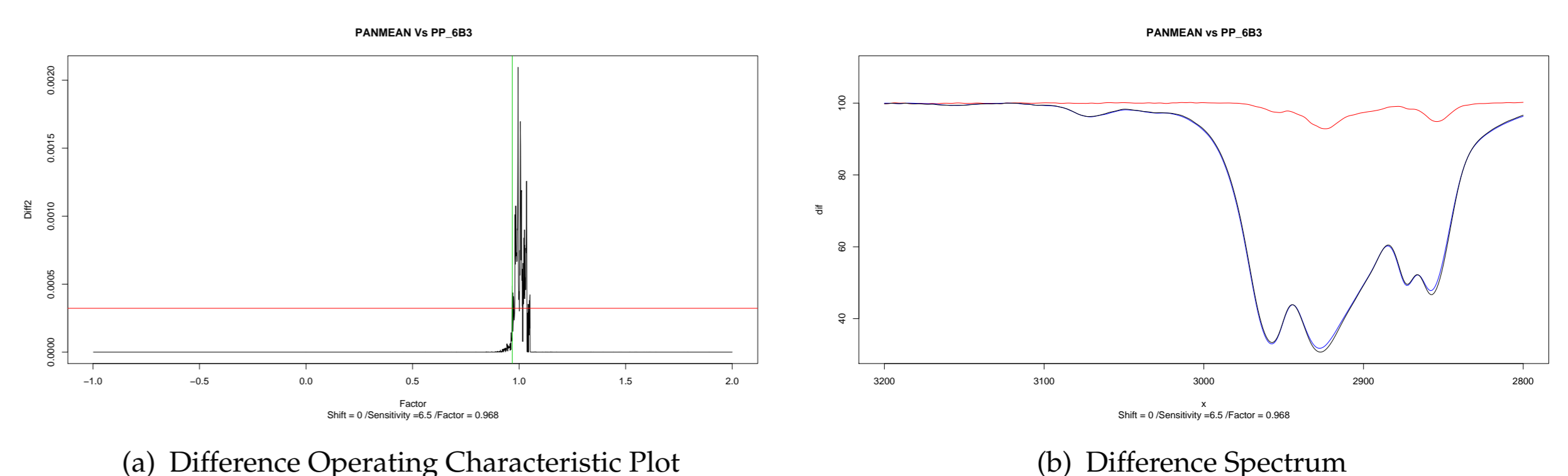


Figure 3: Purity control based on Difference Spectroscopy

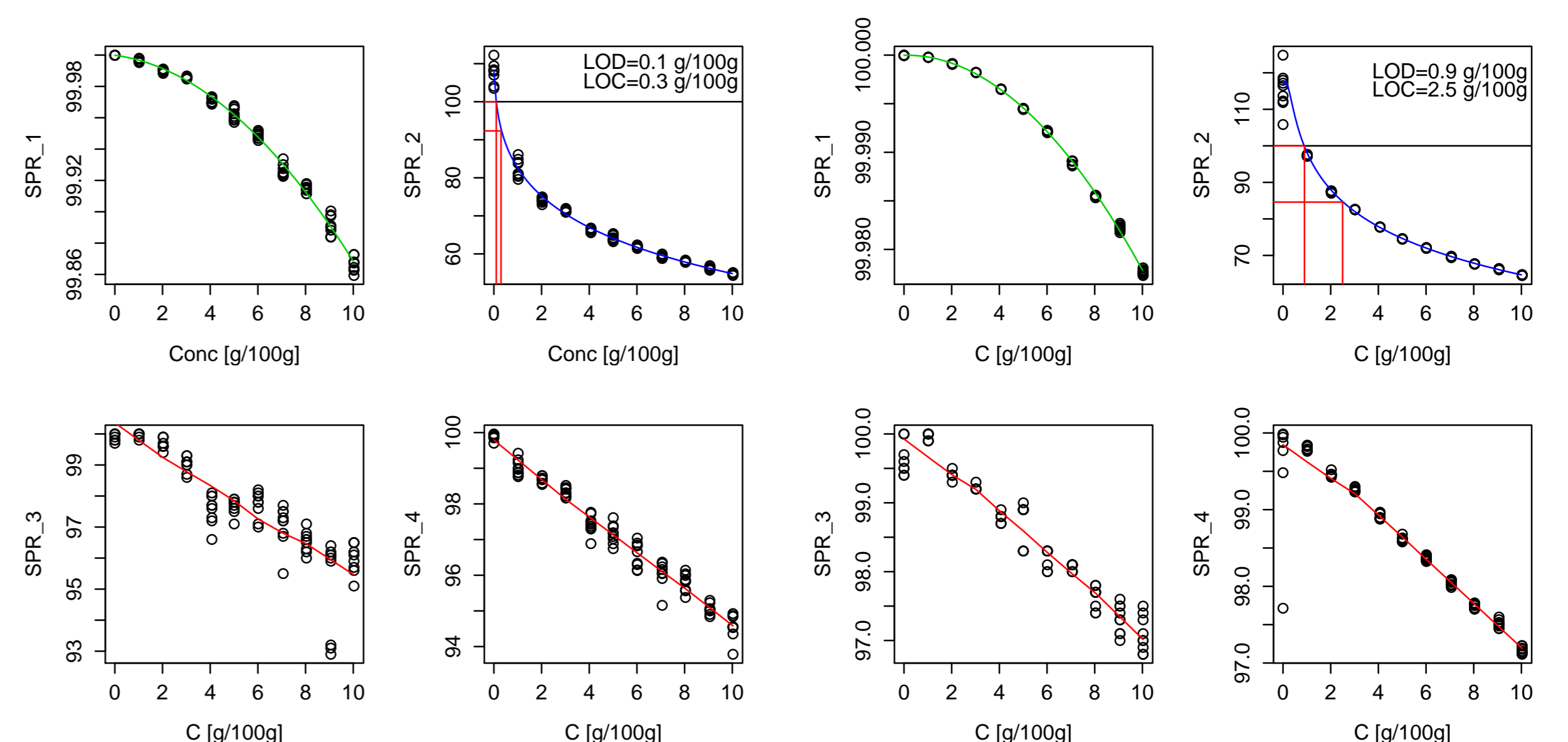
$\chi$  and  $f_{opt}$  are used as spectral purity parameters based on following equations:

$$SPR_3 = 100 \times f_{opt} \quad (6)$$

$$SPR_4 = 100 \times \left(1 - \frac{\int_{\bar{\nu}_1}^{\bar{\nu}_2} D^+}{\int_{\bar{\nu}_1}^{\bar{\nu}_2} A}\right) \quad (7)$$

## 4 Purity control of Palatinol N

With the example of a certain plasticizer (Palatinol N) it is shown that a contamination originating from another plasticizer (Palatinol 911P) can be detected down to about a concentration of 0.1 g/100g of this impurity by Diamond ATR Spectroscopy and down to 0.9 g/100g with Near Infrared Spectroscopy.



(a) Spectral Purity based on Correlation coefficient and Difference Spectroscopy (MIR) (b) Spectral Purity based on Correlation coefficient and Difference Spectroscopy (NIR)

Figure 4: Dependency of Spectral Purity grades of Palatinol N as function of the concentration of an impurity (Palatinol 911P)

## 5 Summary

The purity of chemical substances can be calculated either by difference spectroscopy or spectral comparison via the correlation coefficient. The necessary statistical computations can easily be performed with R.

## References

- [1] Karl Molt und Alfred Schlachter, VDI-Berichte Nr. 1959, 3-26, 2006
- [2] Horst Weitkamp und Dieter Wortig, Mikrochimica Acta 1983 II, 31-57
- [3] Jacob Cohen, Statistical Power Analysis for the Behavioral Sciences, Second Edition, S. 110 f., Lawrence Erlbaum Associates, Inc., Hove and London (1988)