

March 08th 2018

PANIC 2018 Validation Workshop  
La Jolla, CA, USA

---

# **WORKSHOP SESSION 3**

## **(E) UNCERTAINTY ANALYSIS**

Michael Maiwald

Bundesanstalt für Materialforschung und -prüfung (BAM)  
Division 1.4 Process Analytical Technology

Richard-Willstätter-Str. 11  
D-12489 Berlin, Germany

---

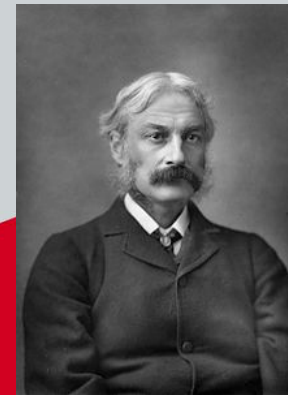
# Part 1: Basic Terms and Definitions

## Practical Statistics

*"Politicians use statistics in the same way  
that a drunk uses lamp-posts —  
for support rather than illumination"*

**Andrew Lang (1844–1912)**  
was a Scottish poet, novelist, and literary critic,  
and contributor to anthropology

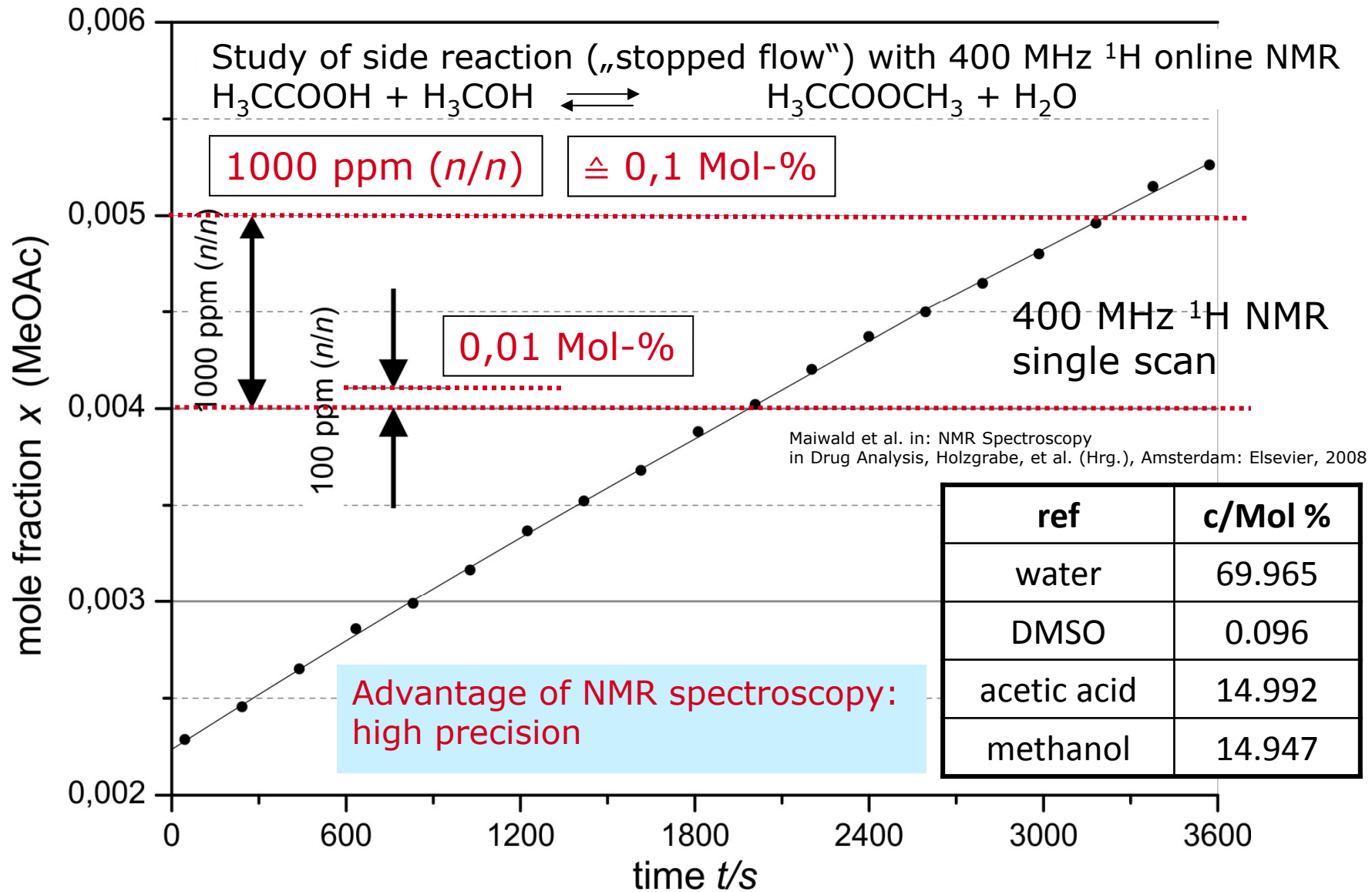
---



Source: [https://en.wikiquote.org/wiki/Andrew\\_Lang](https://en.wikiquote.org/wiki/Andrew_Lang)

# Example: Esterification

## Precision



- 
- In general, the **result of a measurement** is only an approximation or **estimate** of the value of the measurand and thus is complete only when accompanied by a statement of the **uncertainty** of that estimate.
  - In practice, the required specification or definition of the measurand is dictated by the required **accuracy of measurement**
  - In many cases, the result of a measurement is determined on the basis of series of observations obtained under **repeatability conditions**
  - Variations in repeated observations are assumed to arise because influence **quantities** that can affect the measurement result are not held completely constant

# Errors, Effects, and Corrections



- 
- In general, a measurement has imperfections that give rise to an error in the measurement result. Traditionally, an error is viewed as having two components, namely, a **random component** and a **systematic component**.
  - **Random error** presumably arises from unpredictable or stochastic temporal and spatial variations of influence quantities.
    - The effects of such variations, hereafter termed random effects, give rise to variations in **repeated observations** of the measurand.
  - **Systematic error** cannot be eliminated but it too can often be reduced.
    - If a systematic error arises from a recognized effect of an influence quantity on a measurement result the **effect can be quantified** and a **correction** or **correction factor** can be applied to compensate for the effect.



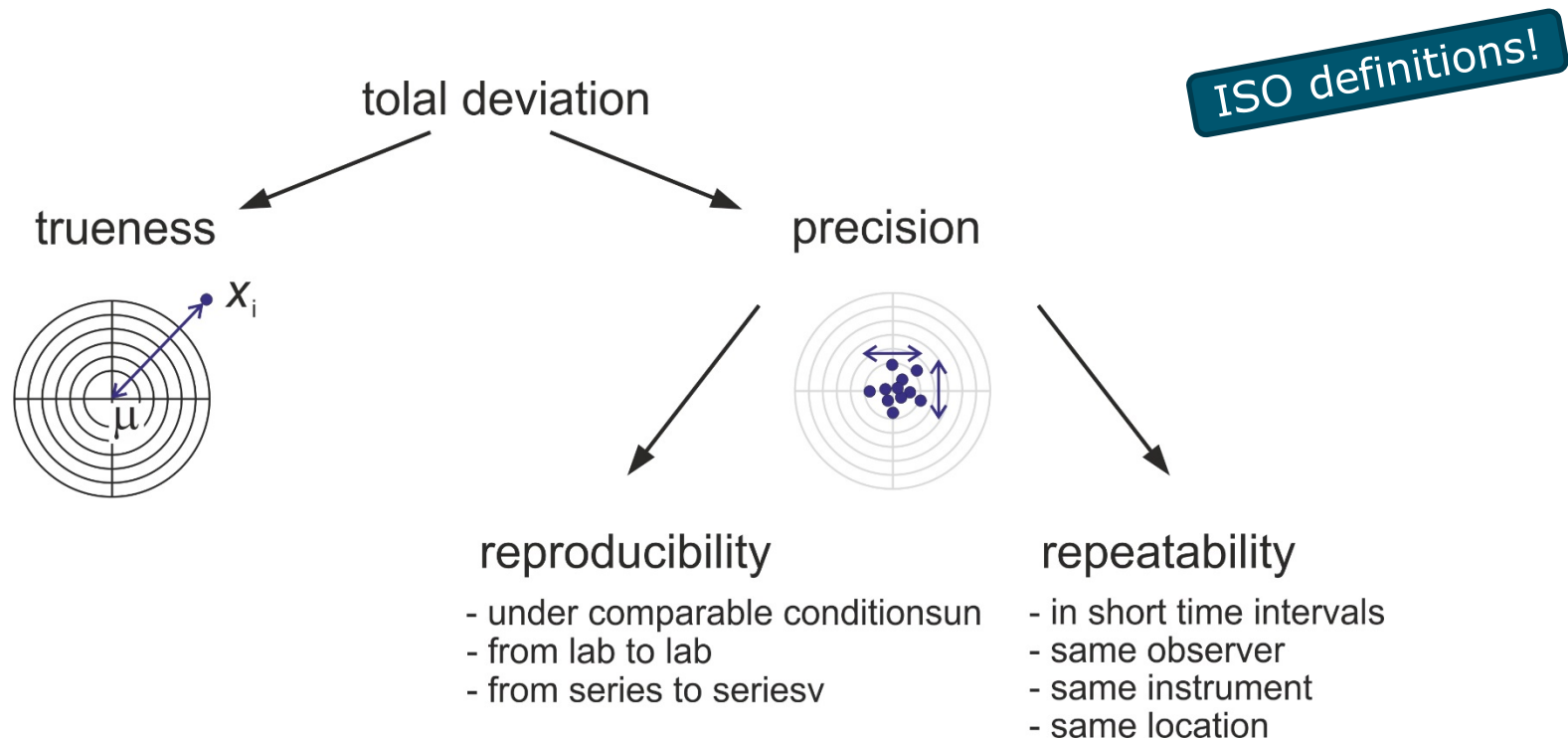
# Uncertainty

- 
- The uncertainty of the result of a measurement reflects the lack of exact knowledge of the value of the measurand. The result of a measurement after correction for recognized systematic effects is still only an estimate of the value of the measurand because of the uncertainty arising from random effects and from imperfect correction of the result for systematic effects.
  - In practice, there are many possible sources of uncertainty in a measurement, among others including:
    - incomplete definition of the measurand;
    - imperfect realization of the definition of the measurand;
    - non-representative sampling — the sample measured may not represent the defined measurand;
    - inadequate knowledge of the effects of environmental conditions on the measurement or imperfect
    - measurement of environmental conditions;
    - finite instrument resolution or discrimination threshold;
    - inexact values of measurement standards and reference materials; ...



# Accuracy, Trueness, and Precision Adjustment

The accuracy of (instrumental) analytical measurements (i.e., its trueness and precision) can be influenced (c.f., optimised) by adjustment of the instrument as well as by selection of the analytical method.



---

## Part 2: Why statistics?

- Distributions:
  - Normal Distribution
  - Student's  $t$ -distribution



# Normal Distribution

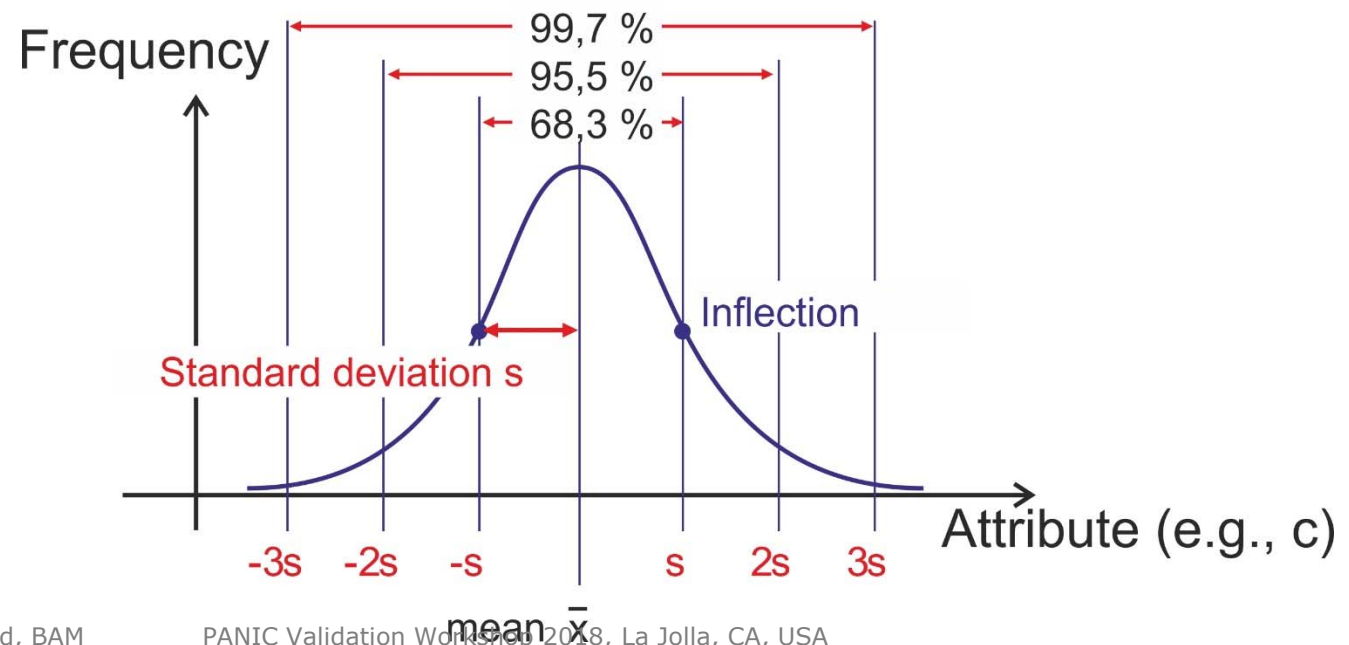


Central Limit Theorem:

numerous sources of arbitrary distribution are (equally) contributing to the measurement result

=> Gauss distribution

Most often we can assume that measurements are normally distributed



# Normal Distribution

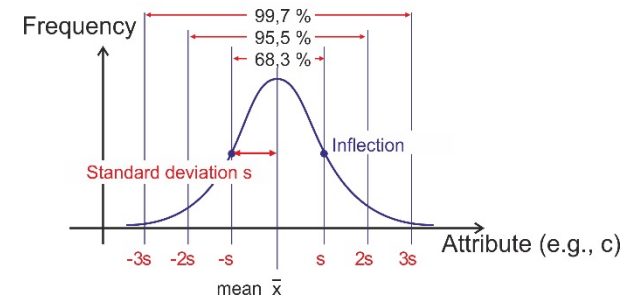


The **normal distribution**, also called **Gaussian distribution** is a **continuous probability distribution** and can be closely approximated by a curve called the “*normal distribution curve*”

$$p = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

The function can be described by  $\mu$  (arithmetic mean of a population) and  $\sigma$  (standard deviation of a population)

$\sigma$  characterise the dispersion of the values



# Sample & Population



---

## Sample:

only subset of population is known, the number of samples  $N$  is limited

$N$  limited

so  $x$  and  $s$  can be determined as **estimates** of  $\mu$  and  $\sigma$

- Rules of statistics
- Declarations as probabilities

## Population:

all possible data are available,  $N$  is infinite

$(N \rightarrow \infty)$

so  $\mu$  and  $\sigma$  can be calculated

# Sample & Population



—  $N$  limited

$(N \rightarrow \infty)$

**Sample:**

Estimated mean value of the sample

$$\bar{x} = \frac{\sum_{i=1}^N x_i}{N}$$

Estimated value of the standard deviation  $s$  and variance  $s^2$  of the sample

$$s = \sqrt{\frac{\sum_{i=1}^N (x_i - \bar{x})^2}{N - 1}}$$

**Population:**

Mean value of the population  $\mu$

$$\mu = \lim_{N \rightarrow \infty} \frac{\sum_{i=1}^N x_i}{N}$$

Standard deviation  $\sigma$  and variance  $\sigma^2$  of the population

$$\sigma = \sqrt{\lim_{N \rightarrow \infty} \frac{\sum_{i=1}^N (x_i - \mu)^2}{N}}$$

$$\sigma_t^2 = \sigma_1^2 + \sigma_2^2 + \dots + \sigma_n^2$$

# From Now On Only **Samples**

## Description of sets of repeated measurements

For a set of  $n$  values  $x_i$

Standard Deviation

$$s(x_i) = \sqrt{\frac{1}{n-1} \cdot \sum_{i=1}^n (x_i - \bar{x})^2}$$

Mean value

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n (x_i)$$

Standard Deviation of the Mean

$$s(\bar{x}_i) = \frac{s(x_i)}{\sqrt{n}}$$

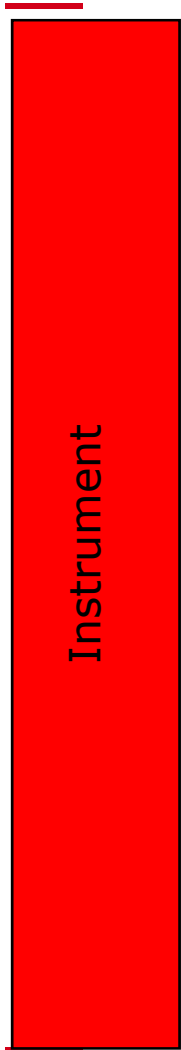
Variance

$$V(x_i) = s^2(x_i)$$

Relative Standard Deviation or Coefficient of Variation

$$\underline{RSD} = \frac{s(x_i)}{\bar{x}} \quad \text{or} \quad RSD (\%) = CV \% = \frac{s(x_i)}{\bar{x}} \cdot 100$$

# Example: Mean Values and Standard Deviations



$$\bar{x}_i = 5.65$$

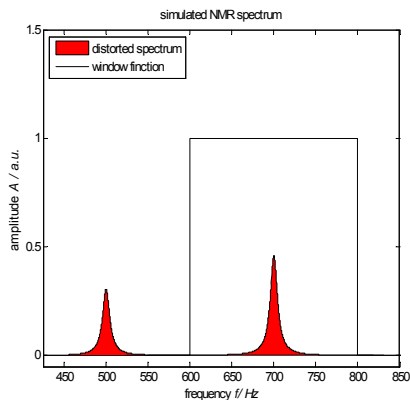
mean value

$$s(x_i) = 0.74$$

standard deviation

$$s(\bar{x}_i) = \frac{s(x_i)}{\sqrt{n}} = 0.17$$

standard deviation of the mean



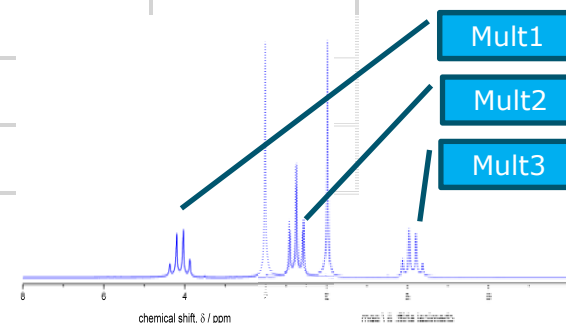
$n = 20$

5.70
5.94
5.82
5.80
5.08
6.73
6.30
4.92
5.75
5.32
6.82
5.79
5.80
3.66
5.04
5.51
6.46
6.28
5.46
4.74

# Example: Mean Values and Standard Deviations of NMR Measurements

Three multiplets and triplicate measurements

Data						
Sample	Mult1	Mult2	Mult3	Ave	Std dev	%RSD
1	98,462	96,83	97,95	97,7473	0,83466	0,8539
2	98,529	97,129	98,193	97,9503	0,73087	0,74616
3	98,475	96,829	98,052	97,7853	0,85479	0,87415
Average	98,4887	96,9293	98,065	97,8277		
Std dev	0,03553	0,17292	0,12202	0,70648		
%RSD	0,03607	0,1784	0,12443	0,68086		



EXCEL:

- **STDEV.S** is called the "N-1" method and assumes that its arguments are a sample of the population
- **STDEV.P** is the "N" method and is based on the entire population given as arguments

Thanks to Mike Bernstein

# Variance and Covariance



---

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n (x_i)$$

$$\text{var}(x_i) = \frac{1}{n-1} \cdot \sum_{i=1}^n (x_i - \bar{x})^2$$

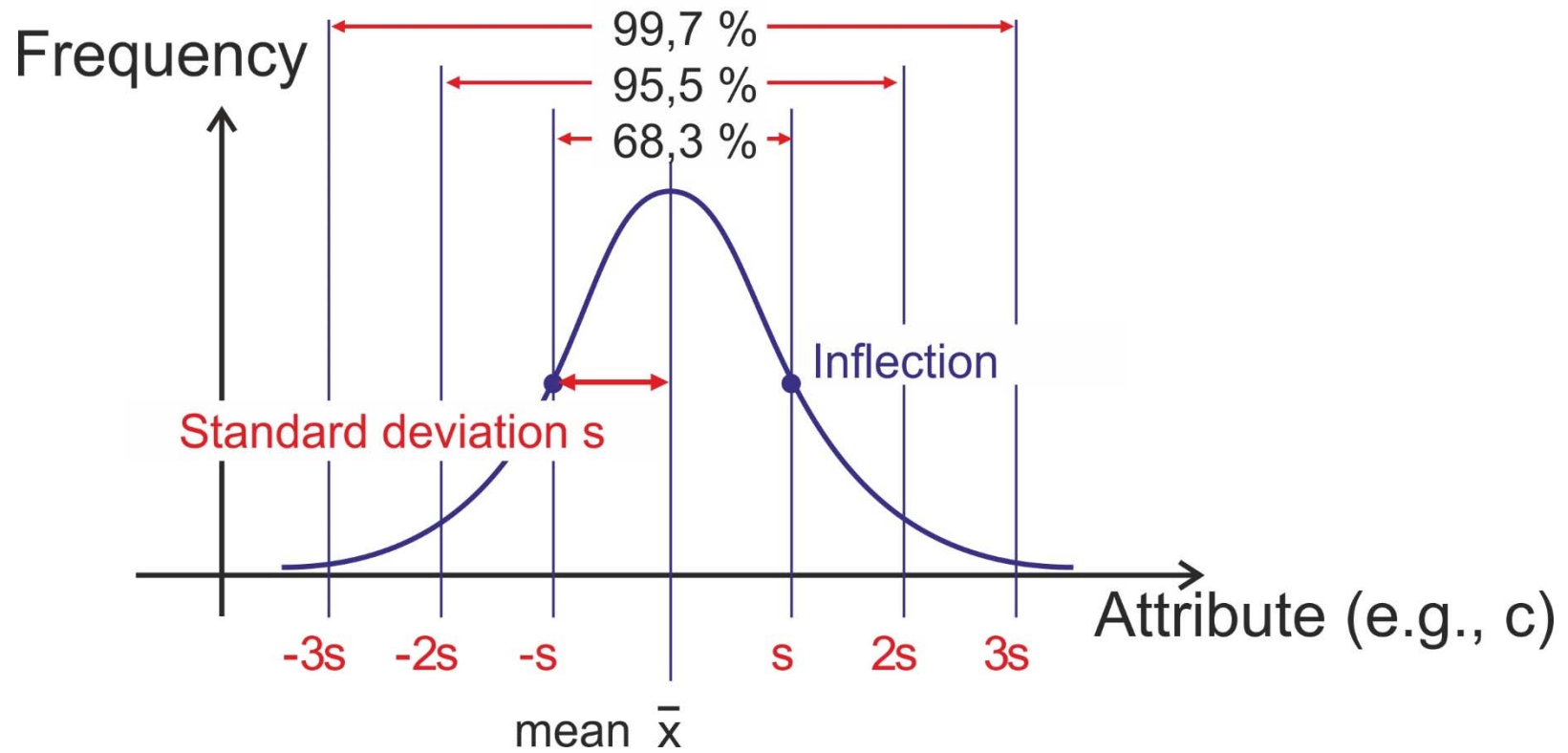
$$\text{cov}(x_i, y_i) = \frac{1}{n-1} \cdot \sum_{i=1}^n (x_i - \bar{x}) \cdot (y_i - \bar{y})$$

$$r = \frac{\sum (x_i - \bar{x}) \cdot (y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \cdot \sum (y_i - \bar{y})^2}}$$





# Confidence Interval



# Student's $t$ -Distribution



Probability distribution used when the population is normally distributed, but the sample size is small.

When the sample mean is  $m$  and  $s(x)$  is the sample standard deviation, then the quantity


$$t = \frac{\bar{x} - \mu}{s(x_i) / \sqrt{n}}$$

has a  $t$  distribution with  $\nu = n - 1$  degrees of freedom.



# Student's *t*-Distribution

**William Sealy Gosset**

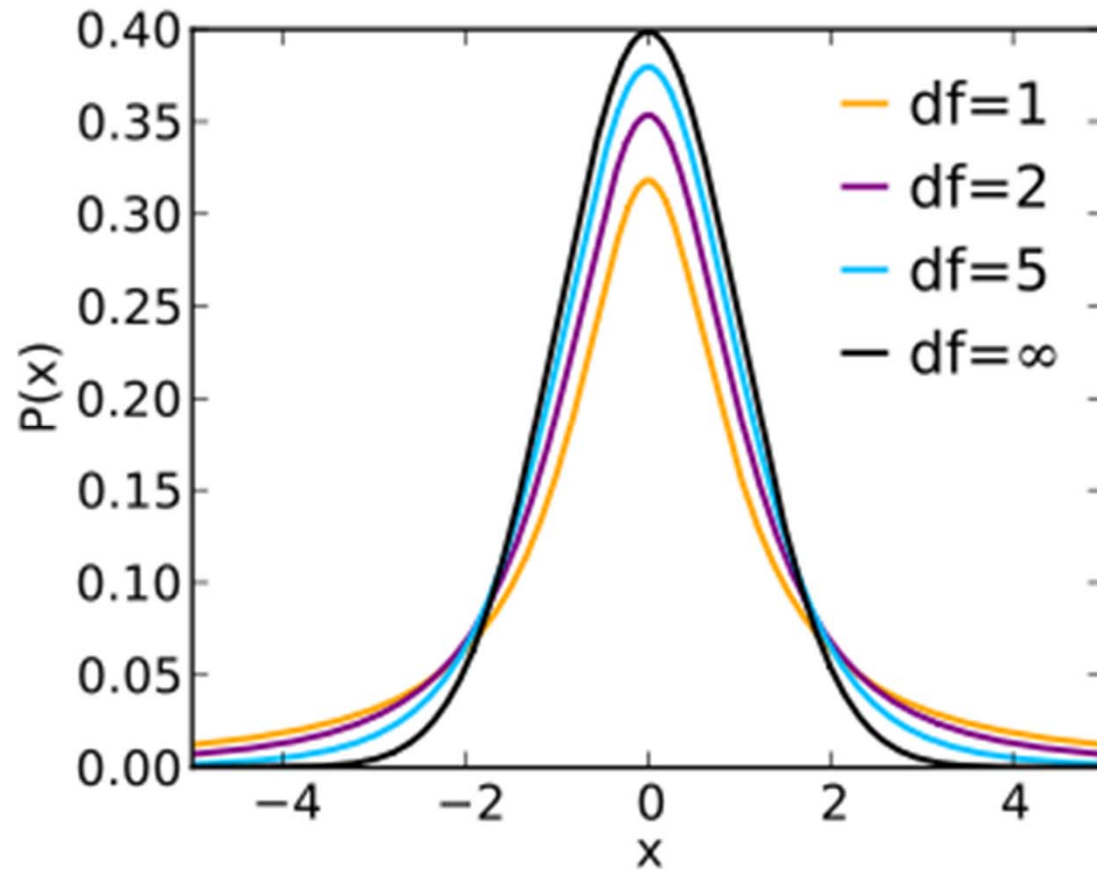



*Student in 1908*

**Born** June 13, 1876  
Canterbury, Kent, England

**Died** October 16, 1937 (aged 61)  
Beaconsfield, Buckinghamshire, England

**Known for** Student's *t*-distribution



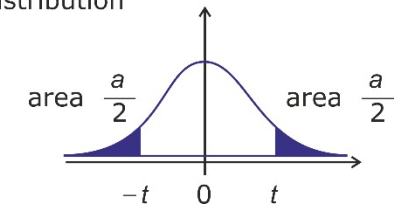
$$f(t) = \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\nu\pi} \Gamma(\frac{\nu}{2})} \left(1 + \frac{t^2}{\nu}\right)^{-\frac{\nu+1}{2}},$$


# Student's *t*-Distribution



$m \backslash \alpha$	0,10	0,05	0,025	0,020	0,010	0,005	0,003	0,002	0,001
1	6,314	12,706	25,452	31,821	63,657	127,3	212,2	318,3	636,6
2	2,920	4,303	6,205	6,965	9,925	14,089	18,216	22,327	31,600
3	2,353	3,182	4,177	4,541	5,841	7,453	8,891	10,214	12,922
4	2,132	2,776	3,495	3,747	4,604	5,597	6,435	7,173	8,610
5	2,015	2,571	3,163	3,365	4,032	4,773	5,376	5,893	6,869
6	1,943	2,447	2,969	3,143	3,707	4,317	4,800	5,208	5,959
7	1,895	2,365	2,841	2,998	3,499	4,029	4,442	4,785	5,408
8	1,860	2,306	2,752	2,896	3,355	3,833	4,199	4,501	5,041
9	1,833	2,262	2,685	2,821	3,250	3,690	4,024	4,297	4,781
10	1,812	2,228	2,634	2,764	3,169	3,581	3,892	4,144	4,587
12	1,782	2,179	2,560	2,681	3,055	3,428	3,706	3,930	4,318
14	1,761	2,145	2,510	2,624	2,977	3,326	3,583	3,787	4,140
16	1,746	2,120	2,473	2,583	2,921	3,252	3,494	3,686	4,015
18	1,734	2,101	2,445	2,552	2,878	3,193	3,428	3,610	3,922
20	1,725	2,086	2,423	2,528	2,845	3,153	3,376	3,552	3,849
22	1,717	2,074	2,405	2,508	2,819	3,119	3,335	3,505	3,792
24	1,711	2,064	2,391	2,492	2,797	3,092	3,302	3,467	3,745
26	1,706	2,056	2,379	2,479	2,779	3,067	3,274	3,435	3,704
28	1,701	2,048	2,369	2,467	2,763	3,047	3,250	3,408	3,674
30	1,697	2,042	2,360	2,457	2,750	3,030	3,230	3,386	3,646
$\infty$	1,645	1,960	2,241	2,326	2,576	2,807	2,968	3,090	3,291

*t* distribution



---

## Part 3: Conventional Linear Calibration

- Examples for univariate and linear calibrations
- Regression analysis
- Examples from NMR spectroscopy

# Agenda Part 3.



- 
- Adjustment, calibration and other terms
  - Calibration using a linear function
    - Basics and requirements for linear regressions
    - Estimation of the regression line
    - Confidence interval for slope and the linear function
    - Examples
  - Tests on linearity
    - Plotting the residues
  
  - Norms and literature for calibration with linear and non-linear functions
  
  - Examples
-

# Calibration



A calibration determines the **functional relationship** between the measurand and the analytical property to be determined.

This will also set the permissible **operating range** of the measurements.

The quality of an analytical measurement is not associated with the (financial) cost of a measurement nor of the instrument.

**Linearity** of the calibration refers to the concrete analysis method and the preferable proportionality between the signal and the feature (concentration), after carrying out all steps of the method.

However, it is only necessary to have a mathematically describable connection between the signal and the feature (concentration).



# Sensitivity and Analytical Sensitivity

- 
- (a) „Ability of an examination procedure to differentiate adjacent values“;
  - (b) IUPAC\*: „Proportionality between the signal and the measured value of a measuring channel (to be determined)“;
  - (c) Slope of the calibration line.

Simplest case: calibration function is a straight line

$$S = m \cdot c + S_{bl}$$



almost always  
fulfilled?

\*IUPAC: „International Union of Pure and Applied Chemistry“, 1919 as subsequent organisation of „International Association of Chemical Societies“  
Webpage: [chemistry.rsc.org/rsc/iupac.htm](http://chemistry.rsc.org/rsc/iupac.htm)

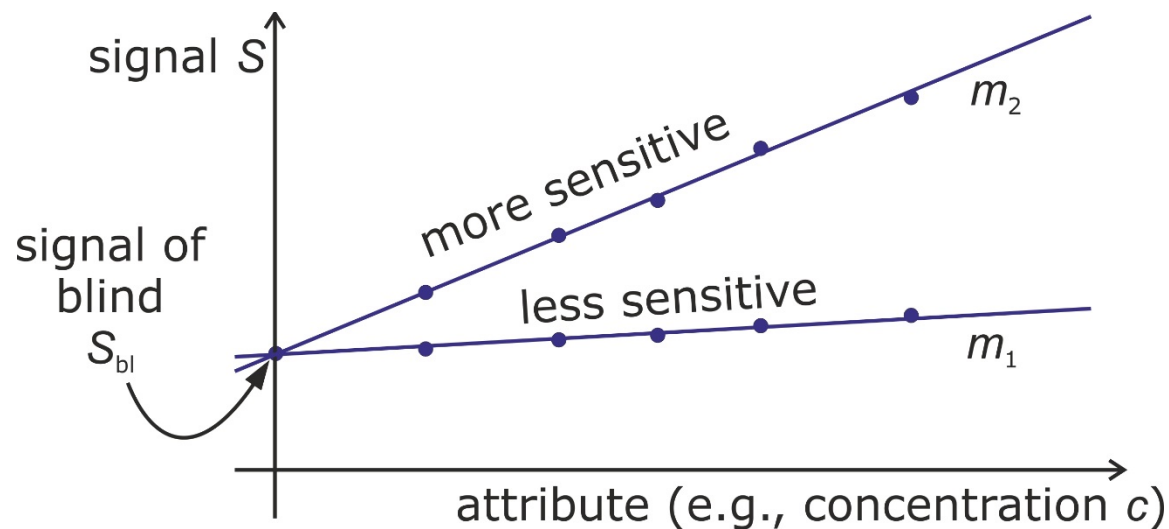


# Analytical Sensitivity

Since the definition of sensitivity does not take into account the precision, the analytical sensitivity is defined

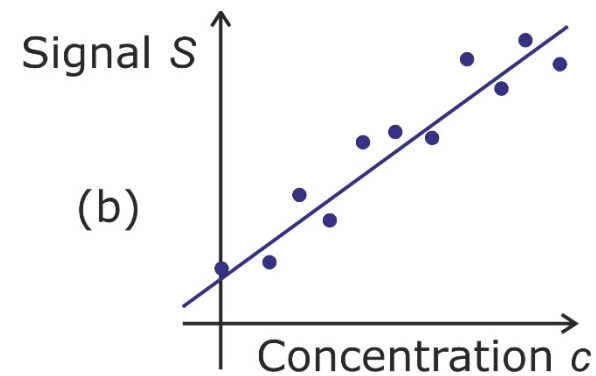
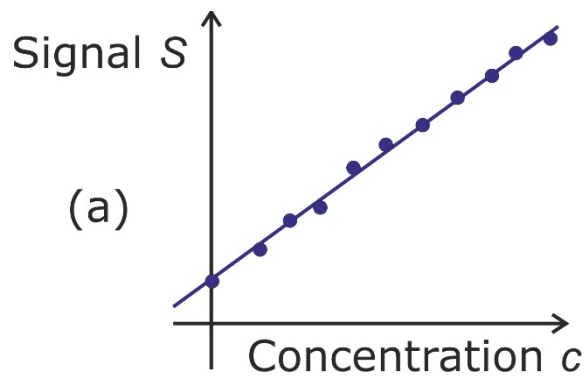
$$\gamma = \frac{m}{s_s}$$

Mean standard deviation of the signal



# Analytical Sensitivity

The method (a, left) has a higher analytical sensitivity than method (b, right), although the calibration function has the same slope



# Example: Analytical Sensitivity

## Comparison HF/LF (5 mm) and

## Analytical Sensitivities $^1\text{H}$ , $^{19}\text{F}$

Signal-to-noise ratios [35] for differently concentrated ethyl benzene solutions in acetone-d6 for MR-NMR with different sample geometries compared to HR-NMR spectrometer.

Ethylbenzene in $\text{CDCl}_3$ : c/% (mass)	MR NMR		HR NMR
	5 mm glass <sup>a</sup> ID 4.2 mm	1/16" polymer tubing	5 mm glass <sup>a</sup> ID 4.2 mm
1	303	–	23,727
10	1939	60	n. a.

<sup>a</sup> NMR tube: Deutero, Economic ID  $4.20 \pm 0.05$  mm.

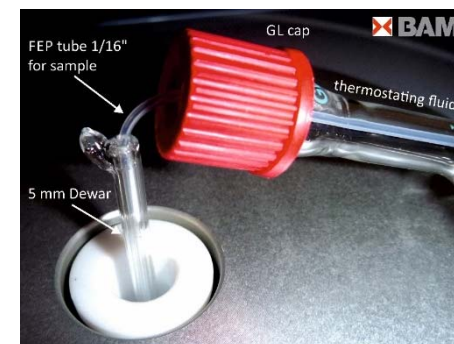
Detection limits (LOD) and quantification limits (LOQ) of 2,2,2-trifluoroethanol in deuterated water according to [36] in  $\text{mol L}^{-1}$  of a MR-NMR with different sample geometries.

TFE in $\text{D}_2\text{O}$	Polymer tubing 0.04" $^1\text{H}$ 43.5 MHz	Polymer tubing 0.04" $^{19}\text{F}$ 40.2 MHz	5 mm glass <sup>a</sup> ID 4.2 mm $^1\text{H}$ 43.5 MHz	5 mm glass <sup>a</sup> ID 4.2 mm $^{19}\text{F}$ 40.2 MHz
LOD/ $\text{mol L}^{-1}$	0.043	0.112	0.019	0.016
LOQ/ $\text{mol L}^{-1}$	0.130	0.335	0.058	0.049
Analytical sensitivity/ $\text{L mol}^{-1}$	74.6	57.0	1242	892

<sup>a</sup> NMR tube: Deutero, Economic ID  $4.20 \pm 0.05$  mm.

[35] T.D.W. Claridge, J.E. Baldwin, F.R.S. Williams, R.M. Williams (Eds.), *High-Resolution NMR Techniques in Organic Chemistry*, Pergamon, Bd. 19. Amsterdam, 1999. ISBN: 0-08-042799-5

[36] J. Mocak, A.M. Bond, S. Mitchell, G. Scollary, A statistical overview of standard (IUPAC and ACS) and new procedures for determining the limits of detection and quantification: application to voltammetric and stripping techniques, *Pure Appl. Chem.* 69 (1997) 297–328, <http://dx.doi.org/10.1351/pac199769020297>



## Ethylbenzene in Acetone-d6

5 mm glass tube HF vs. LF: factor of 78.3 (for 1% ethyl benzene) theoretical value of 72.3 since  $S/N$  varies with  $B_0^{7/4}$

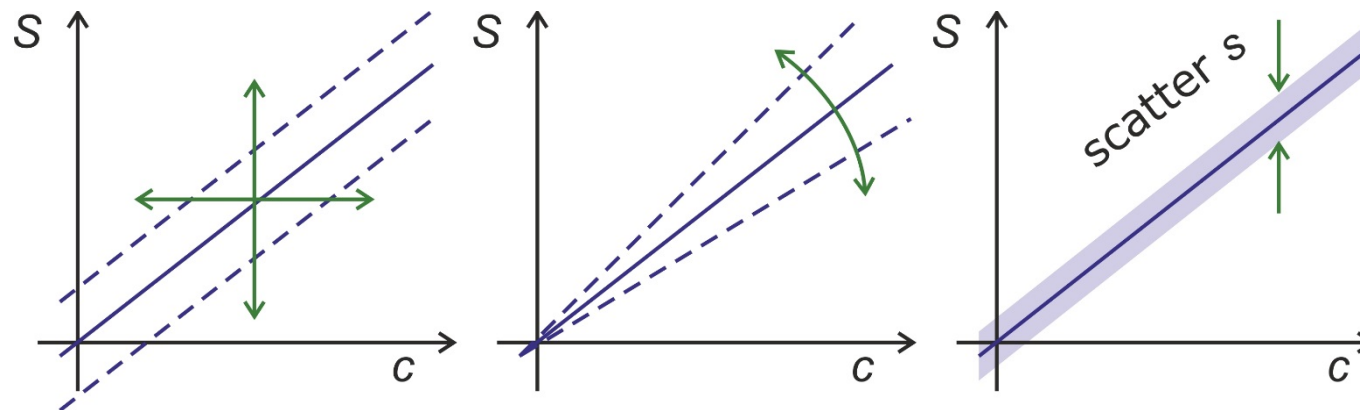
## 2,2,2-Trifluoroethanol in $\text{D}_2\text{O}$

Analytical sensitivities are reduced by a factor of 16.7 ( $^1\text{H}$ ) and 15.7 ( $^{19}\text{F}$ ) comparing 5 mm glass tube (4.2 mm ID) to polymer tubing flow cell (1.02 mm ID)

Zientek, N. et al. *J Mag Res*, 249 (2014) 53-62

# Other Common Definitions

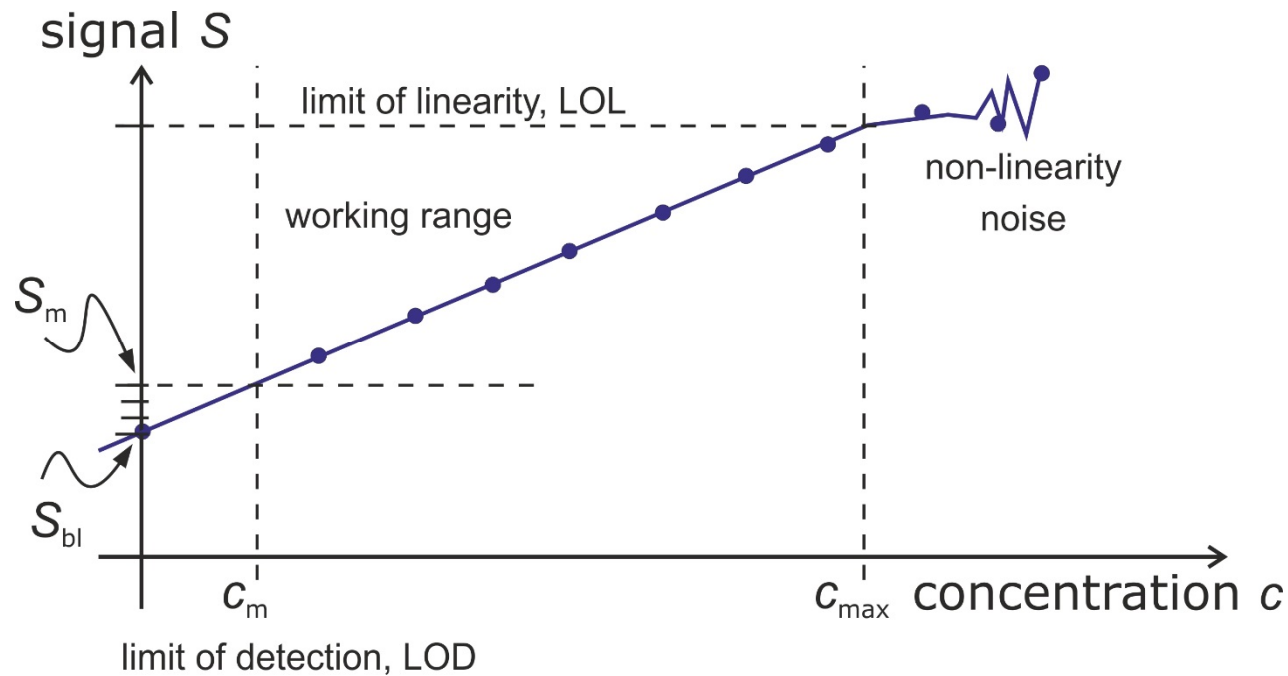
Adjustment: The measuring device is set so that measurement deviations are as small as possible or do not exceed the working range! This is a technical intervention in the measuring device, calibrations lose their validity.



# Other Common Definitions

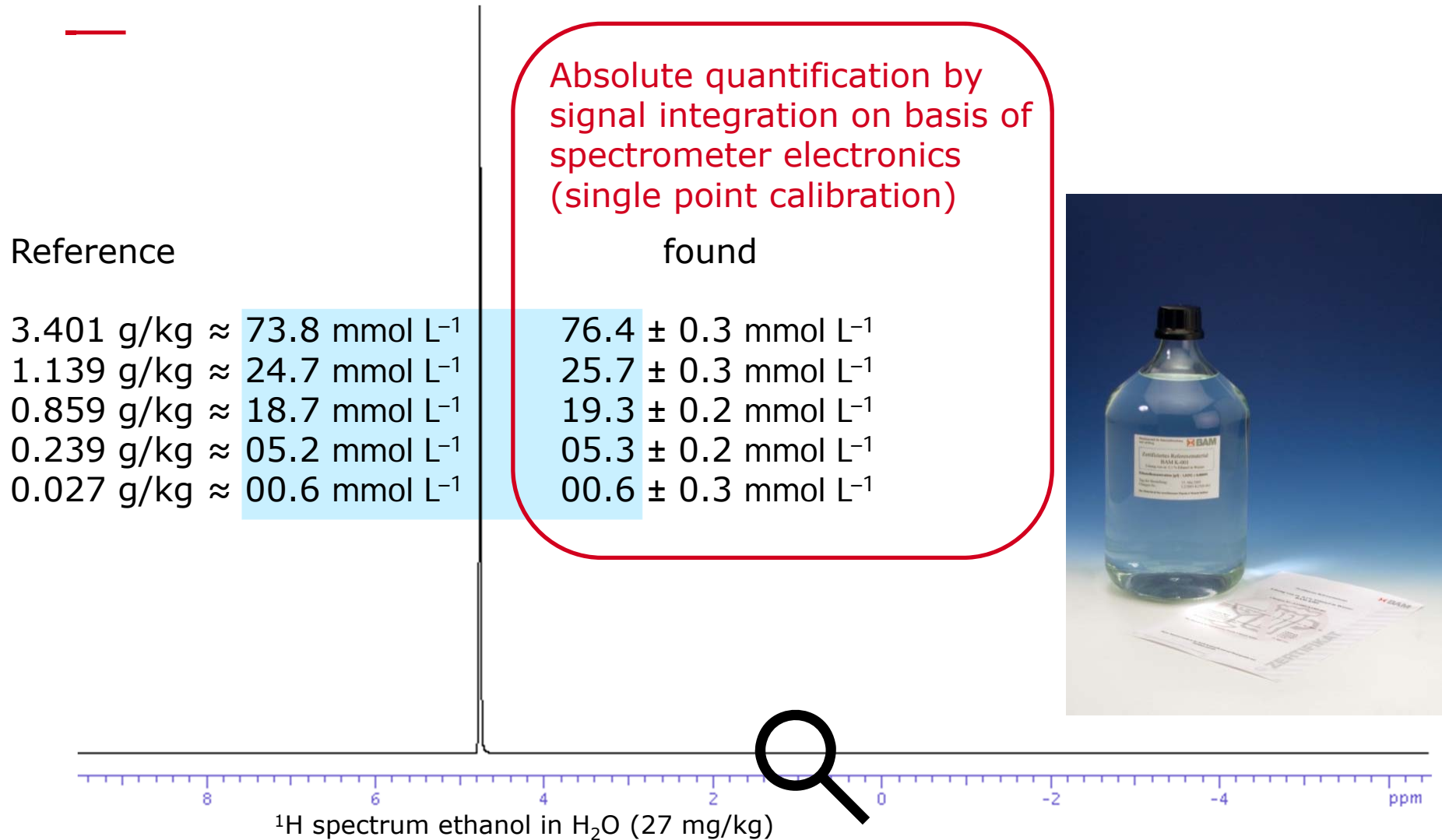


**Range (Working Range):** "analytically usable area between the limit of detection of a measuring channel and the linearity limit. Characteristic or concentration range for acceptable/confirmed information about the accuracy, precision, selectivity, linearity, and robustness."



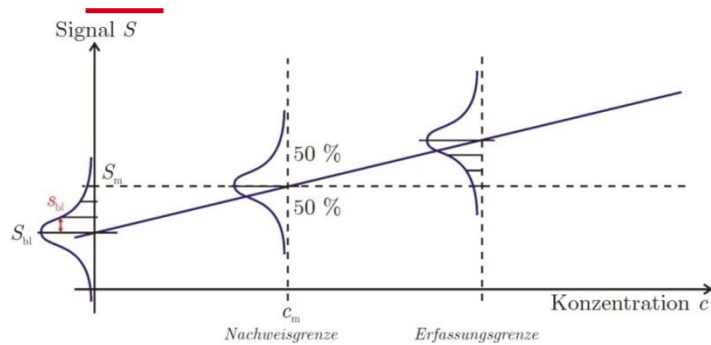
# Example: Limit of Detection

## Quantification of Ethanol in Pure Water @ 500 MHz



# Example: Limit of Detection

## Quantification of Ethanol in Water

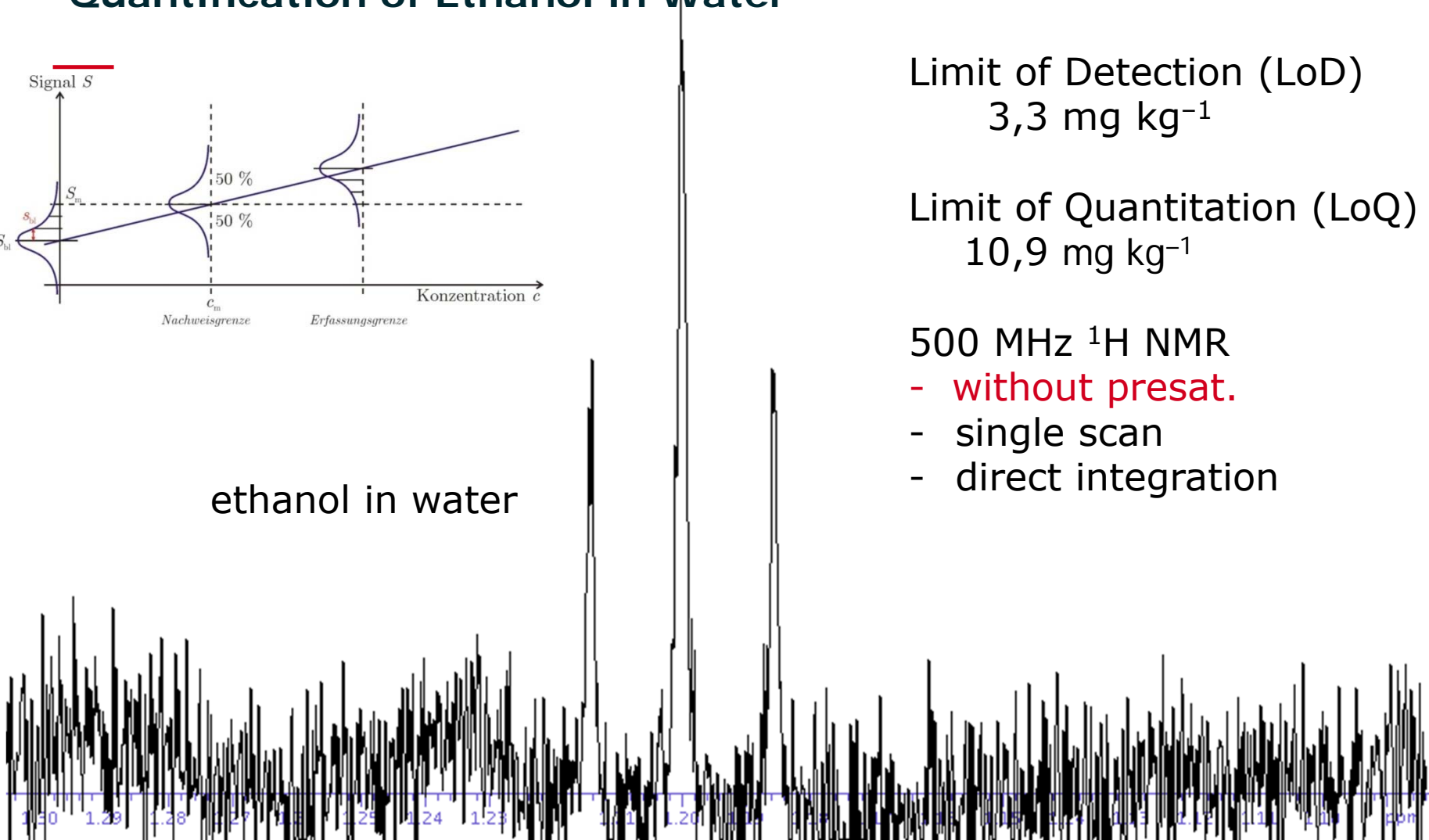


Limit of Detection (LoD)  
3,3 mg kg<sup>-1</sup>

Limit of Quantitation (LoQ)  
10,9 mg kg<sup>-1</sup>

- 500 MHz <sup>1</sup>H NMR
- without presat.
  - single scan
  - direct integration

ethanol in water



<sup>1</sup>H spectrum ethanol in H<sub>2</sub>O (27 mg/kg)

# Calibrating With Linear Functions



## Basics of linear regression

---

A calibration function (to be determined between the size (characteristic) and a measured signal) is usually a linear function within the analytical range (working range).

Both in the definition of this functional relationship, the calibration itself, as well as during all subsequent measurements, experimentally determined data points are linearly correlated.

These data points are limited in their number and by the experiment more or less prone to error. It is called a "linked sample".

At approximately linearly correlated variables  $x$  and  $y$  "regression lines" can be obtained following the rules of mathematical statistics and best possible predictions of  $y$  for a given  $x$  and vice versa can be made.

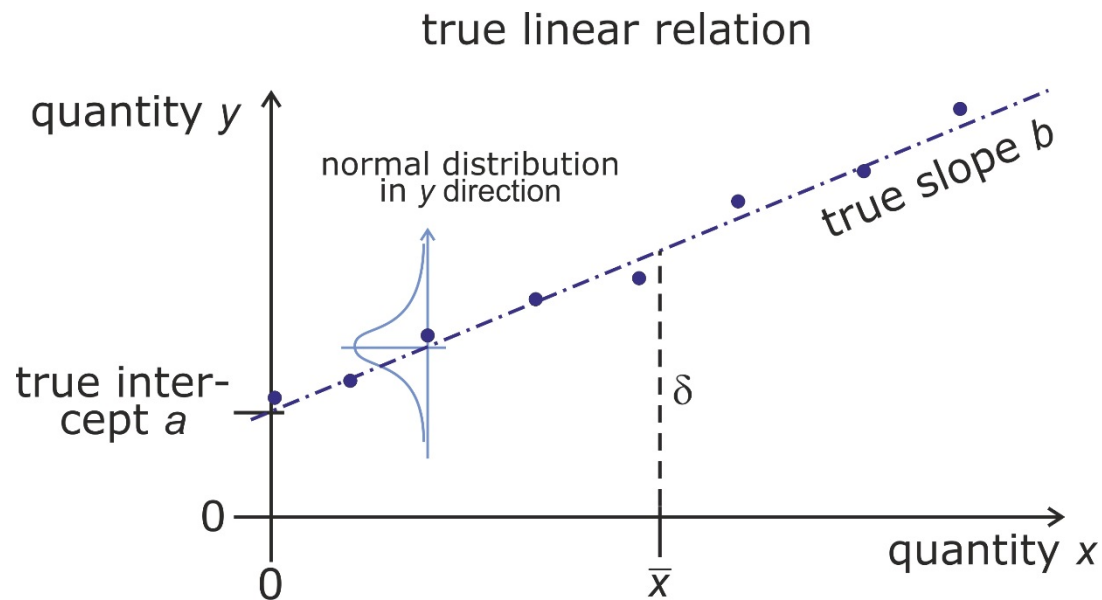
If the following **requirements** are met, one can generally use the procedures outlined below to find a regression line as well as important parameters upon its validity.



# Calibrating With Linear Functions

## Prerequisites

- The value for  $x$  can be measured accurately
- All statistical fluctuations are attributed to the variable  $y$
- The expected value of  $y(x)$  depends linearly from  $x$ :  $E(y(x)) = \alpha + \beta \cdot x$
- For every fixed  $x$  its  $y(x)$  is normally distributed (GAUSS distribution) with the expected value  $E(y(x))$  and constant, from  $x$  independant scattering\*  $\sigma$
- There are  $n$  given measuerment points  $(x_i; y_i)$  with  $i = 1, 2, \dots, n$ .



\*Note: The size of  $\sigma$  is unimportant, only  $\sigma$  may not depend on  $x$ . It is a constant. There are also approaches with  $\sigma$  proportional to  $x$  or  $x^2$ , etc.

# Calibrating With Linear Functions



## What to do, if the requirements are not fulfilled?

- 
- There are other methods, for example, whether there is a regression line, or not at all
    - F tests
  - There are also other statistical procedures if one or several of the above conditions are not fulfilled, as for example
    - The variances of the measured values behave inhomogeneous (heteroscedasticity)
      - weighted regression
    - Errors in the determination of the value  $x$  cannot be neglected
      - Generalised Least Squares, in particular orthogonal regression
    - Under the underlying calibration data are outliers, which must remain unconsidered in the regression function
      - outlier elimination or robust regression

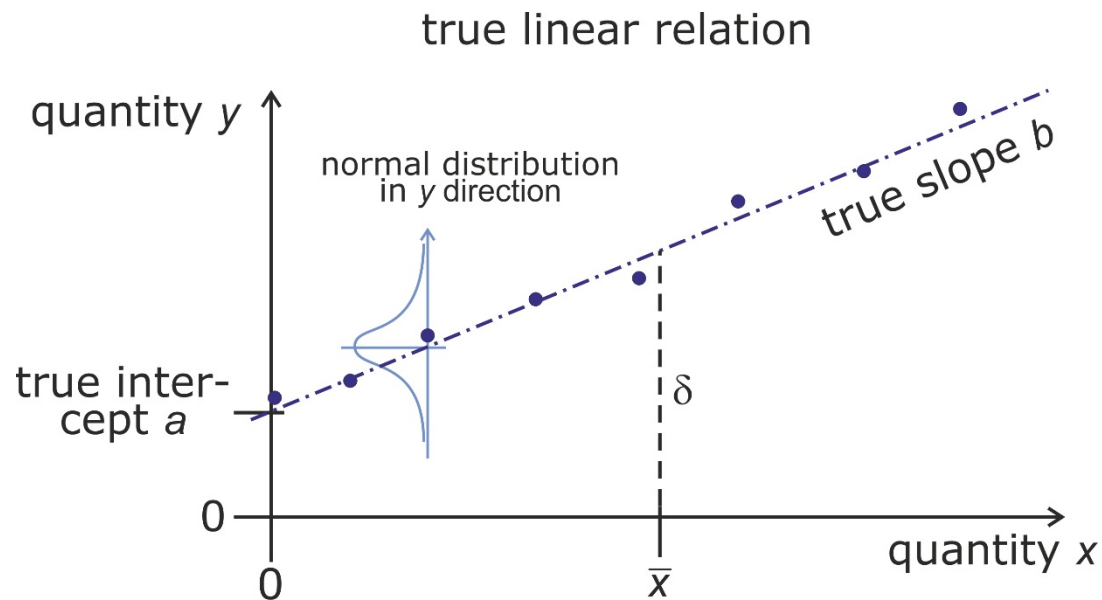
# Calibrating With Linear Functions

## Target and approach

Wanted

(1) Estimates for the slope  $b$  and the distance  $d$  above the mean value  $\bar{x}$ ,  
where  $E(y(x)) = \beta(x - \bar{x}) + \delta$  and  $\alpha = \delta - \beta \cdot \bar{x}$

(2) Confidence interval for the slope  $\beta$  and the linear function  $\alpha = \delta - \beta \cdot \bar{x}$



# Calibrating With Linear Functions

## (1) Estimate of the regression line

Following estimates give the best regression line:

- Estimate  $b$  for the slope  $\beta$

$$b = \frac{\sum_{i=1}^n (x_i - \bar{x})y}{\sum_{i=1}^n (x_i - \bar{x})^2} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2} = \frac{\sum_{i=1}^n x_i \cdot y_i - n \cdot \bar{x} \cdot \bar{y}}{(n - 1) \cdot s_x^2} = \frac{s_y \cdot r}{s_x}$$

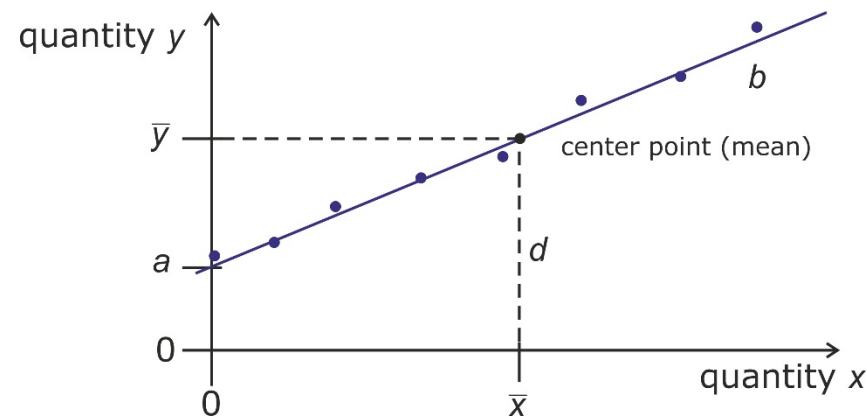
Pocket calculator formula

- Estimate  $d$  for the distance  $\delta$  over the  $x$  axis at the center point  $\bar{x}$

$$d = \bar{y}$$

- Estimate  $a$  for the  $y$  axis intercept  $\alpha$

$$a = d - b \cdot \bar{x}$$



# Calibrating With Linear Functions

## (1) Estimate for the regression line

The regression line is then

$$y = b(x - \bar{x}) + d \quad \text{or} \quad y = bx + \underbrace{d - b\bar{x}}_{=a}$$

For ist estimates one has to calculate as follows:

- Mean values („center points“) of the measured values in x and y direction

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad \bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$$

- Estimates for the standard deviations of the samples in x and y direction

$$s_x = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}} = \sqrt{\frac{(\sum_{i=1}^n \bar{x}_i^2) - n \cdot \bar{x}^2}{n - 1}} \quad s_y = \sqrt{\frac{\sum_{i=1}^n (y_i - \bar{y})^2}{n - 1}}$$

Pocket calculator formula

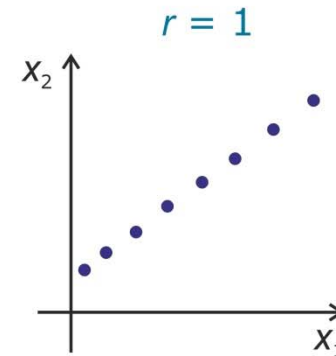
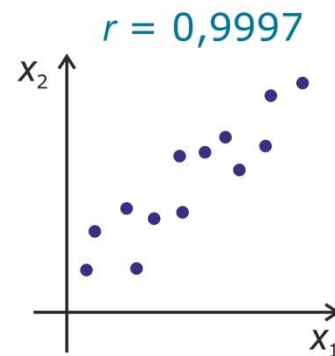
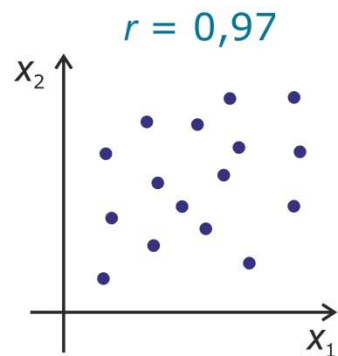
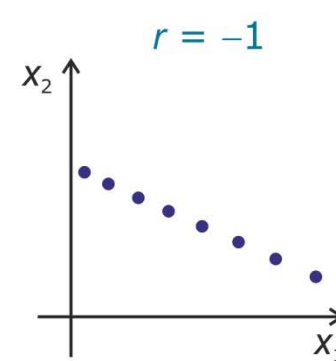
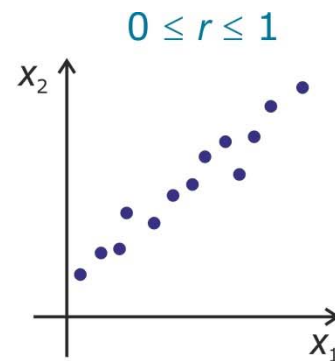
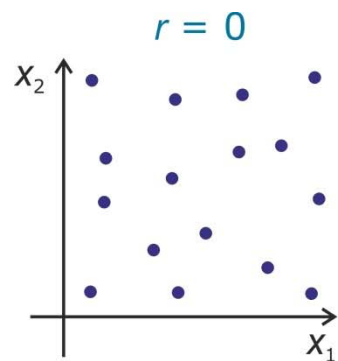


# Calibrating With Linear Functions

## (1) Estimate for the regression line

- Correlation coefficient

$$r = \frac{1}{s_x \cdot s_y} \cdot \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})$$



### Notes

- the correlation coefficient is a number between -1 and 1, which describes the correlation of two variables
- If the covariance (correlation) is  $\neq 0$ , the two quantities certainly are not independent. Unfortunately, the contrary does not apply!
- Only if the two variables are normally distributed from  $r = 0$  follows the independence of both quantities.
- $r = 1$  does not mean, the slope  $\beta = 1$ !
- In the context of "calibration quality" please forget  $r$ !**

# Calibrating With Linear Functions

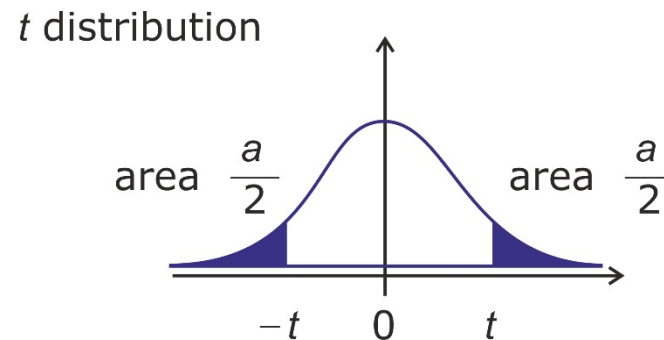
## (2) Confidence interval for slope and regression line

Go ahead with the following recipe:

- Select a probability of error  $\alpha$

„I would like to be no more than  $\alpha = 0.05 = 5\%$  (also  $\alpha = 1\%$ , etc.) of all cases wrong and reject the true values of  $\alpha$  and  $\beta$  - although they are correct"

- Read  $c = t_{\frac{\alpha}{2}; n-2}$  from the  $t$  distribution with  $n-2$  degrees of freedom



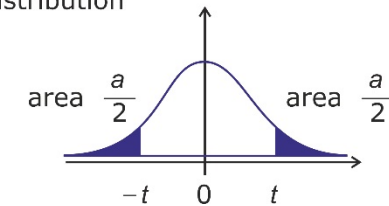
# Calibrating With Linear Functions

## *t* distribution



$m \backslash \alpha$	0,10	0,05	0,025	0,020	0,010	0,005	0,003	0,002	0,001
1	6,314	12,706	25,452	31,821	63,657	127,3	212,2	318,3	636,6
2	2,920	4,303	6,205	6,965	9,925	14,089	18,216	22,327	31,600
3	2,353	3,182	4,177	4,541	5,841	7,453	8,891	10,214	12,922
4	2,132	2,776	3,495	3,747	4,604	5,597	6,435	7,173	8,610
5	2,015	2,571	3,163	3,365	4,032	4,773	5,376	5,893	6,869
6	1,943	2,447	2,969	3,143	3,707	4,317	4,800	5,208	5,959
7	1,895	2,365	2,841	2,998	3,499	4,029	4,442	4,785	5,408
8	1,860	2,306	2,752	2,896	3,355	3,833	4,199	4,501	5,041
9	1,833	2,262	2,685	2,821	3,250	3,690	4,024	4,297	4,781
10	1,812	2,228	2,634	2,764	3,169	3,581	3,892	4,144	4,587
12	1,782	2,179	2,560	2,681	3,055	3,428	3,706	3,930	4,318
14	1,761	2,145	2,510	2,624	2,977	3,326	3,583	3,787	4,140
16	1,746	2,120	2,473	2,583	2,921	3,252	3,494	3,686	4,015
18	1,734	2,101	2,445	2,552	2,878	3,193	3,428	3,610	3,922
20	1,725	2,086	2,423	2,528	2,845	3,153	3,376	3,552	3,849
22	1,717	2,074	2,405	2,508	2,819	3,119	3,335	3,505	3,792
24	1,711	2,064	2,391	2,492	2,797	3,092	3,302	3,467	3,745
26	1,706	2,056	2,379	2,479	2,779	3,067	3,274	3,435	3,704
28	1,701	2,048	2,369	2,467	2,763	3,047	3,250	3,408	3,674
30	1,697	2,042	2,360	2,457	2,750	3,030	3,230	3,386	3,646
$\infty$	1,645	1,960	2,241	2,326	2,576	2,807	2,968	3,090	3,291

*t* distribution





# Calibrating With Linear Functions

## (2) Confidence interval for slope and regression line

- calculate

$$r = \frac{1}{s_x \cdot s_y} \cdot \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})$$

$$s = \sqrt{\frac{1}{n} \sum_{i=1}^n [b(x_i - \bar{x}) + d + y]^2} = \sqrt{\frac{n-1}{n} (s_y^2 - b^2 \cdot s_x^2)}$$

Taschenrechnerformel

$s$  is a measure, how far the measurement points deviate from the regression line.

- $b$ ,  $d$ ,  $s_x$  and  $s_y$  are also calculated

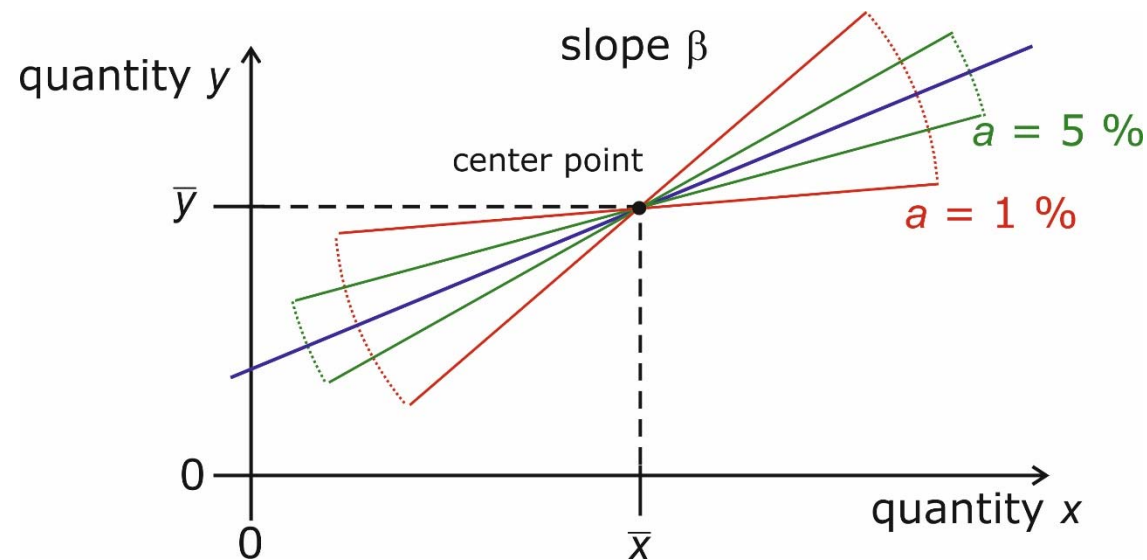


# Calibrating With Linear Functions

## (2) Confidence interval for slope and regression line

- Confidence interval for the slope  $\beta$

$$b - c \cdot \frac{s}{s_x} \sqrt{\frac{n}{(n-1) \cdot (n-2)}} \leq \beta \leq b + c \cdot \frac{s}{s_x} \sqrt{\frac{n}{(n-1) \cdot (n-2)}}$$



As an approximation also the following formula can be used

$$\frac{s}{s_x} \sqrt{(n-1) \cdot (n-2)} = \bar{r} \sqrt{n-2}$$



# Calibrating With Linear Functions

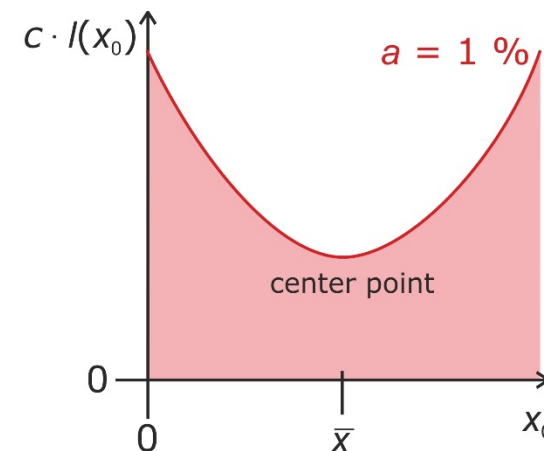
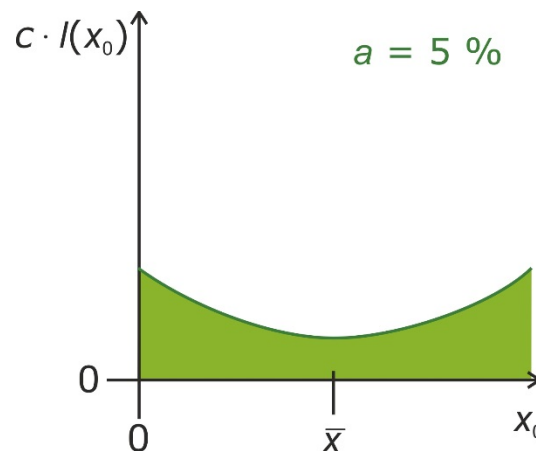
## (2) Confidence interval for slope and regression line

- Confidence interval for the  $y$  value at given  $x$  (and, e.g., the  $y$  axis intercept at  $x = 0$ )

Estimate of the scatter of  $y(x_0)$  for given  $x_0$ . Here, for a number of supporting points  $x_0$  the following function is computed.

$$l(x_0) = \frac{s}{\sqrt{n-2}} \sqrt{1 + \frac{n(x_0 - \bar{x})^2}{(n-1)s_x^2}}$$

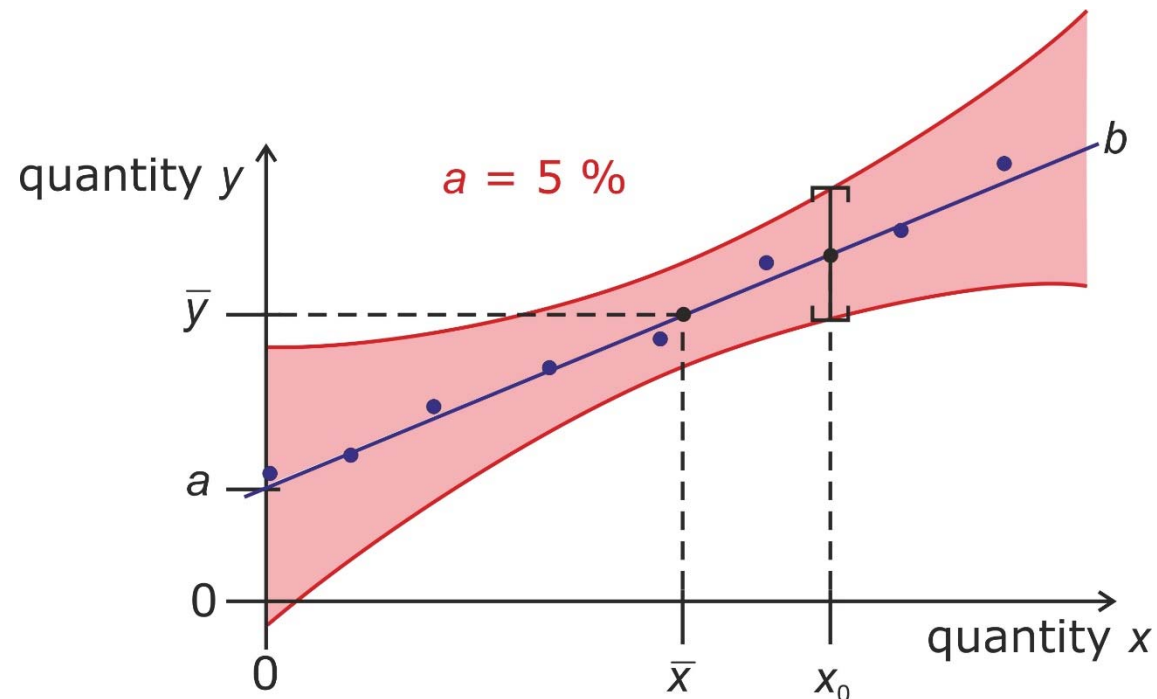
Für  $x_0 = \bar{x}$  gilt  $l(x_0) = \frac{s}{\sqrt{n-2}}$



# Calibrating With Linear Functions

## (2) Confidence interval for slope and regression line

$$\underbrace{b \cdot (x_0 - \bar{x}) + d}_{\text{line}} - c \cdot l(x_0) \leq E(y(x)) \leq \underbrace{b \cdot (x_0 - \bar{x}) + d}_{\text{line}} + c \cdot l(x_0)$$



Meaning: the calculated regression line is found to be within the red area to 95 %.



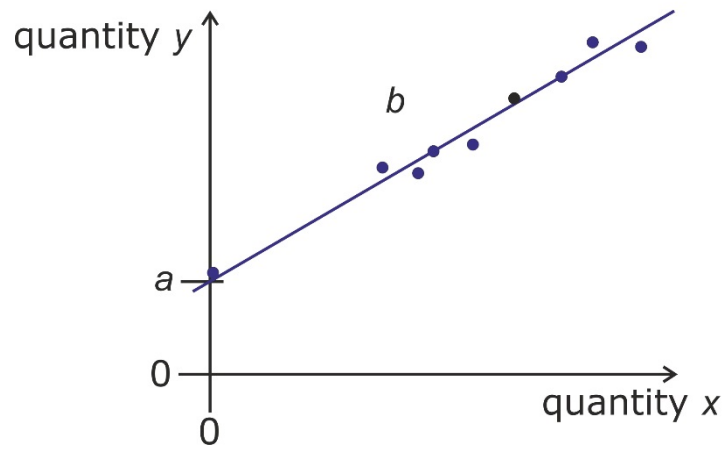
# Calibrating With Linear Functions



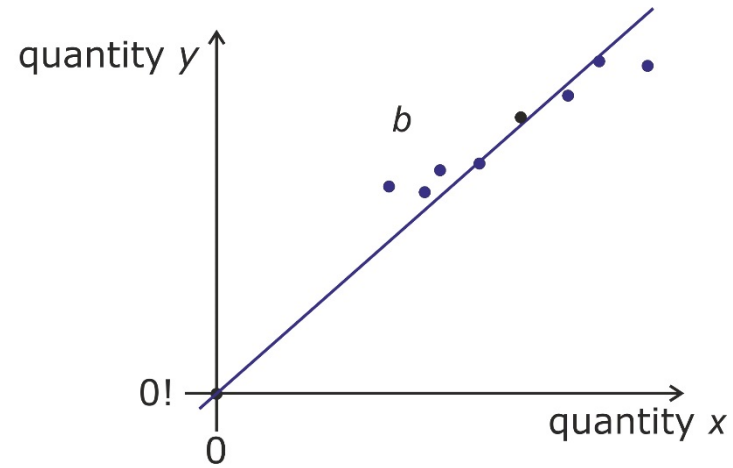
## Test on linearity



$$y = a + b \cdot x$$



$$y = b \cdot x$$

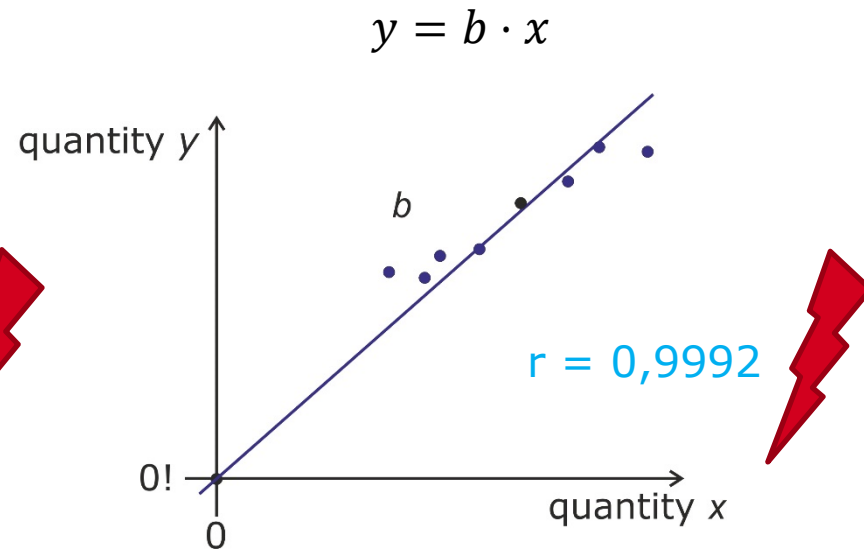
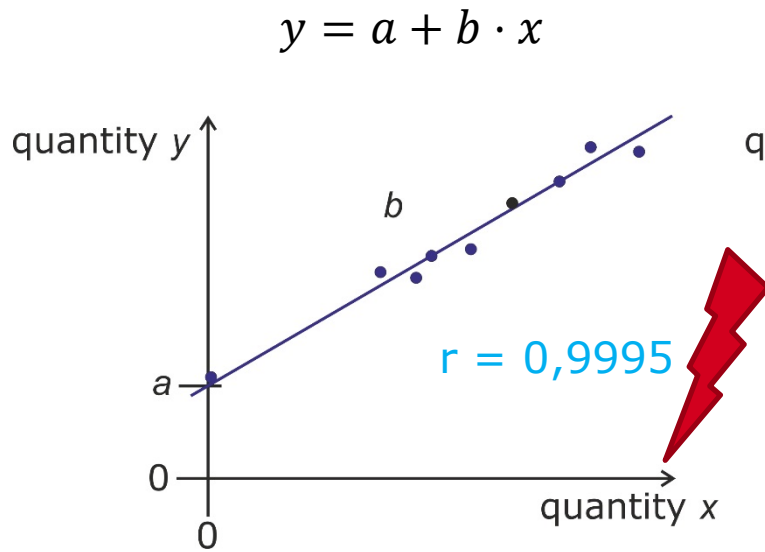


Enforced proportionality?



# Calibrating With Linear Functions

## Test on linearity



Enforced proportionality?

Proof by correlation coefficient?

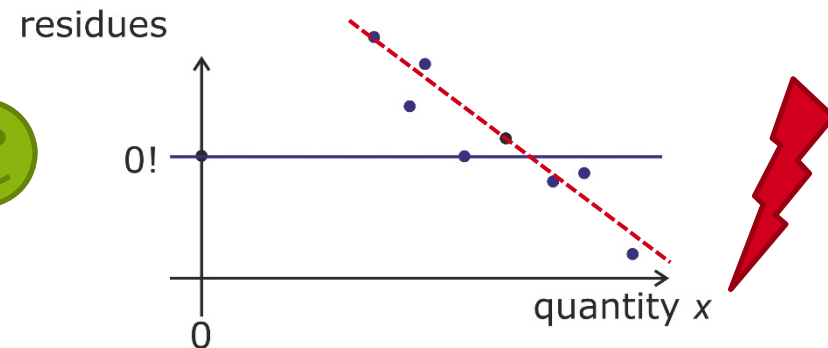
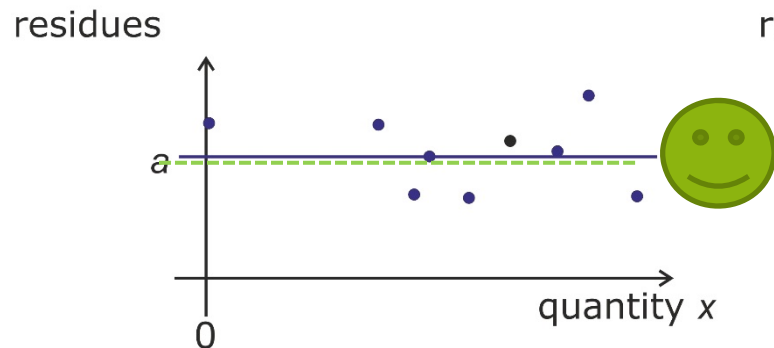
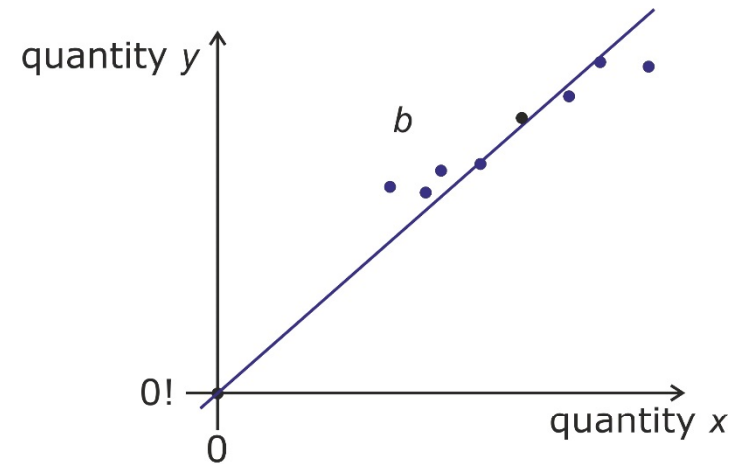
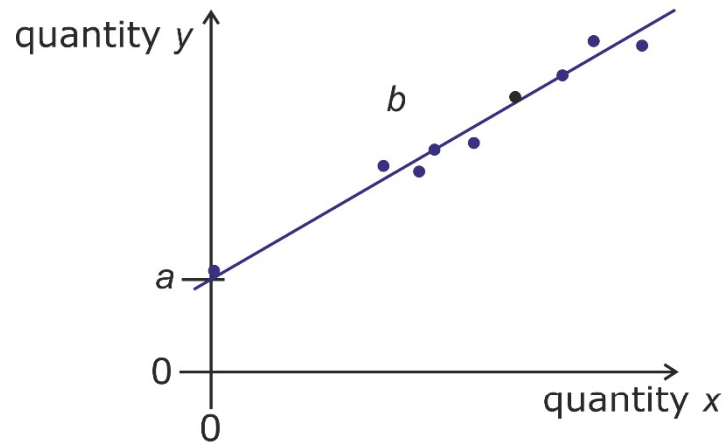
# Calibrating With Linear Functions

## Test on linearity – plotting the residues

→ Difference between a fitted function and its measured data points

$$y = a + b \cdot x$$

$$y = b \cdot x$$



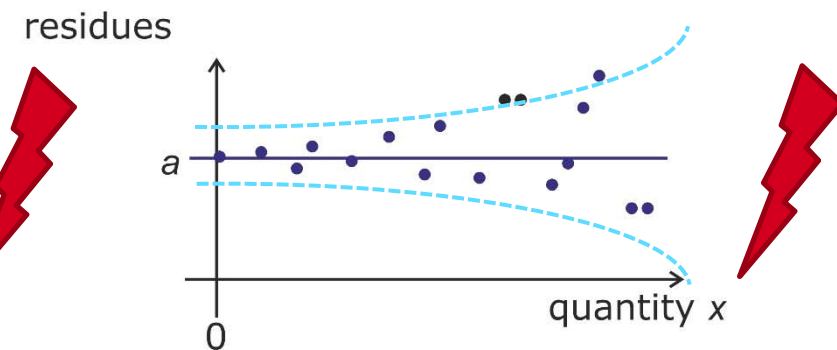
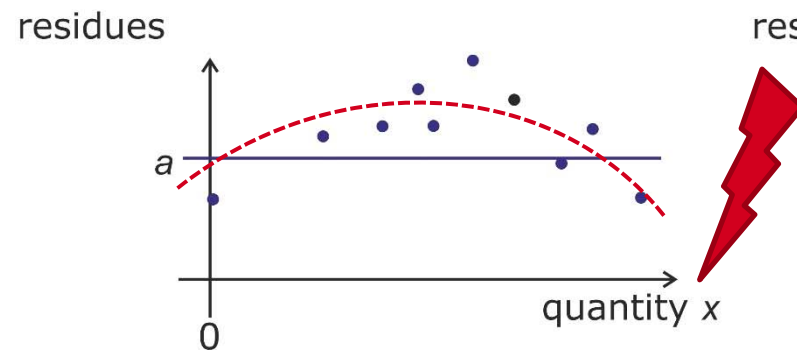
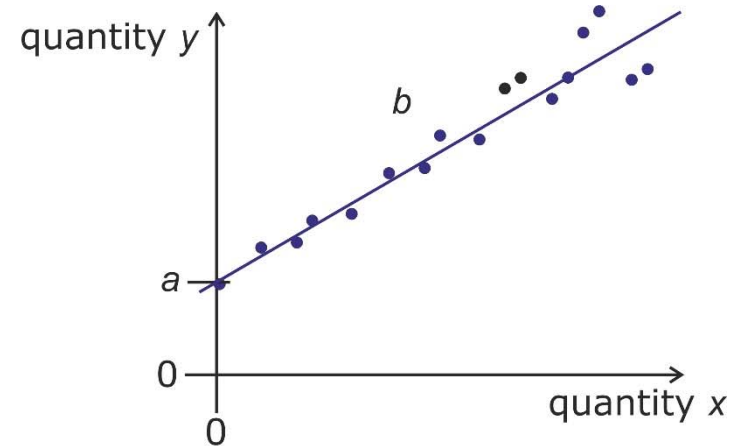
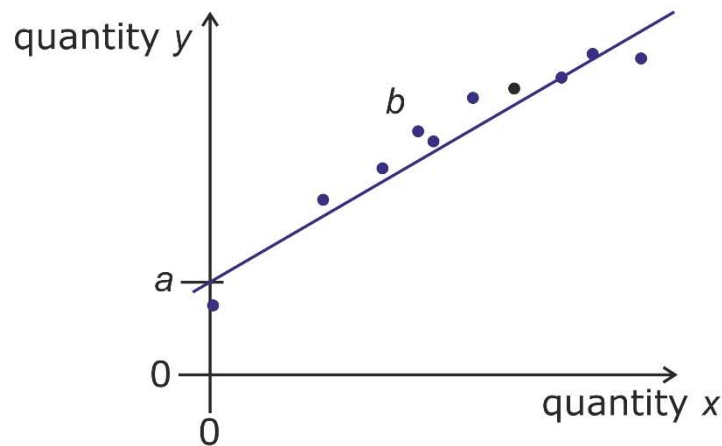
# Hoping to Calibrate With Linear Functions BAM

Test on linearity – recognition of trends



$$y = a + b \cdot x + c \cdot x^2$$

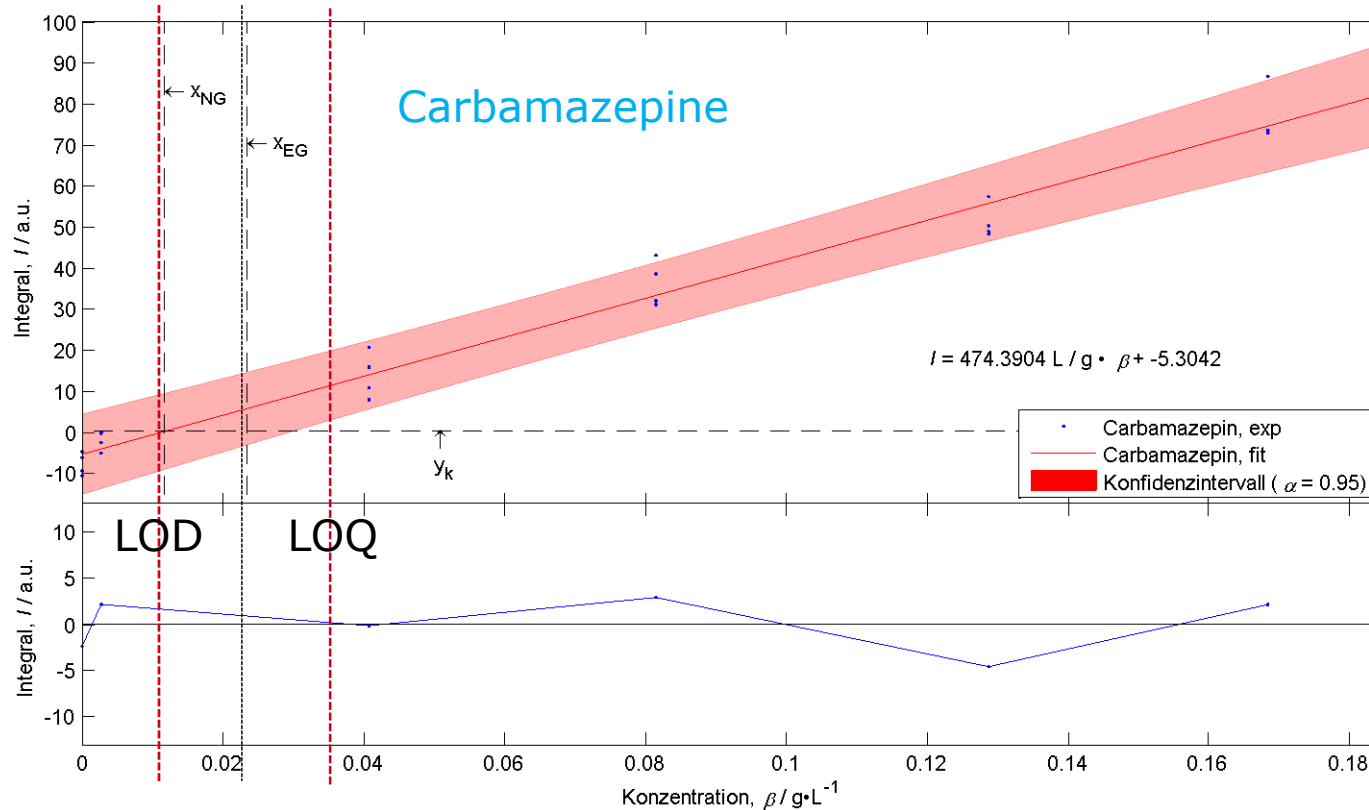
$$y = a + b \cdot x$$





# Calibrating With Linear Functions

## Example 1: Carbamazepine im Water (500 MHz)



Limit of detection

$$x_{LOD} = 11.7 \text{ mg/L}$$

Limit of quantification

$$x_{LOQ} = 35.1 \text{ mg/L}$$

Standard deviation of analytical method

$$s_{x0} = 6.44 \text{ mg/L}$$

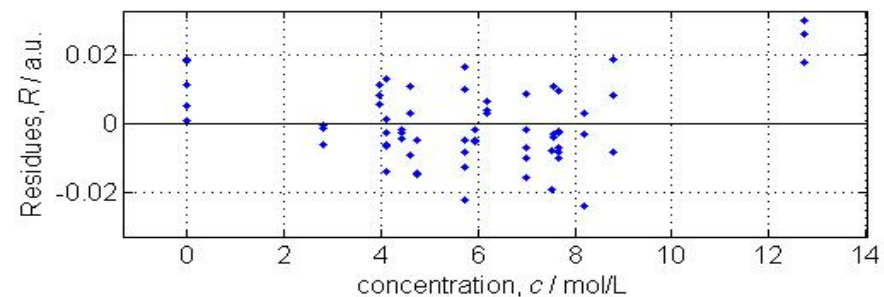
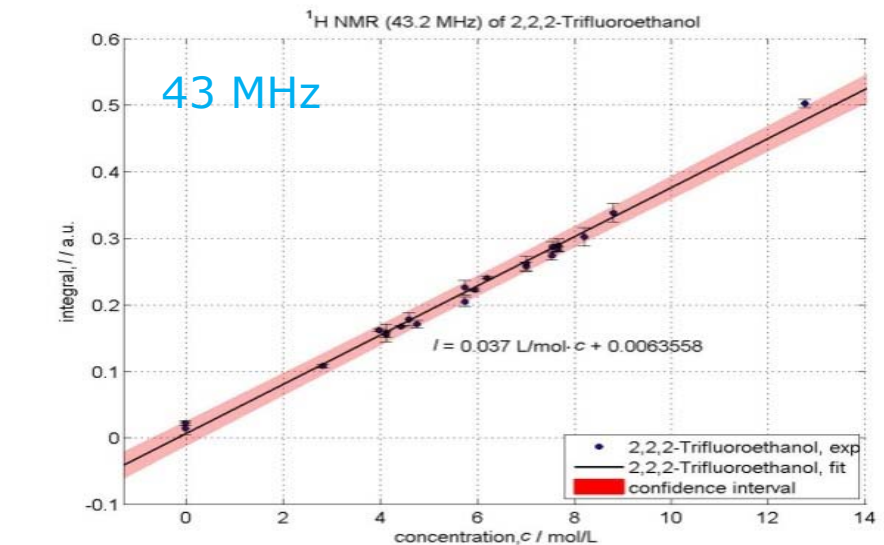
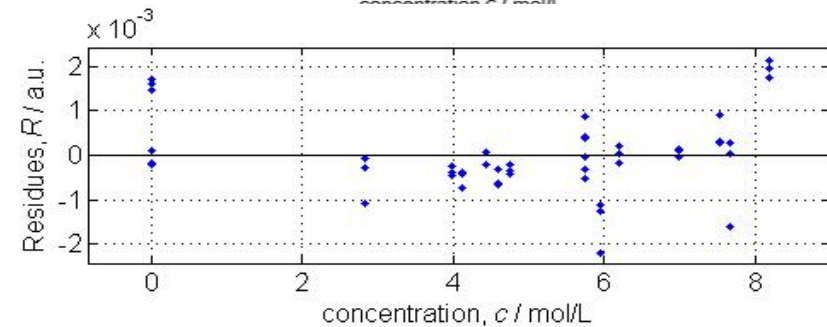
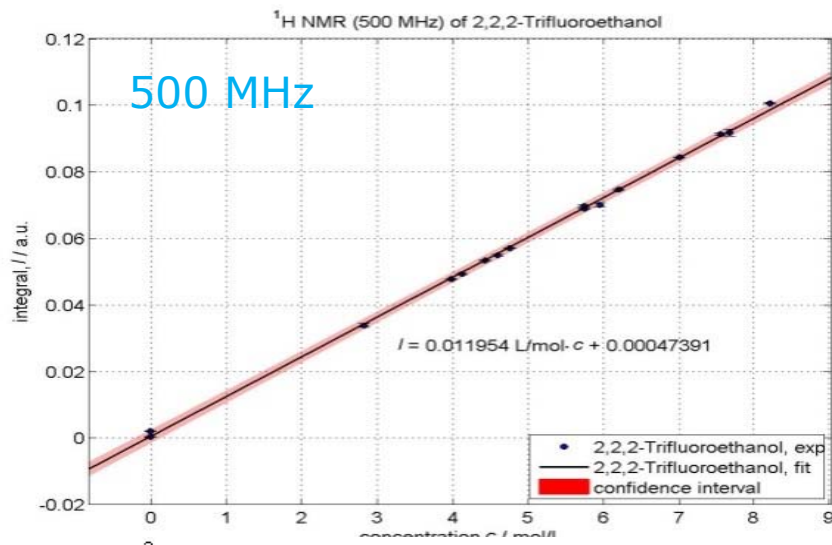
Critical value of the measured quantity

$$y_k = 0.249 \text{ a.u.}$$

# Calibrating With Linear Functions

## Example 2: 2,2,2-trifluoroethanol (43 and 500 MHz)

Calibration of 2,2,2-trifluoroethanol with  $^1\text{H}$  NMR spectroscopy:  
Plot of NMR signal areas versus amount of substance concentrations  
and plot of residues to linear function



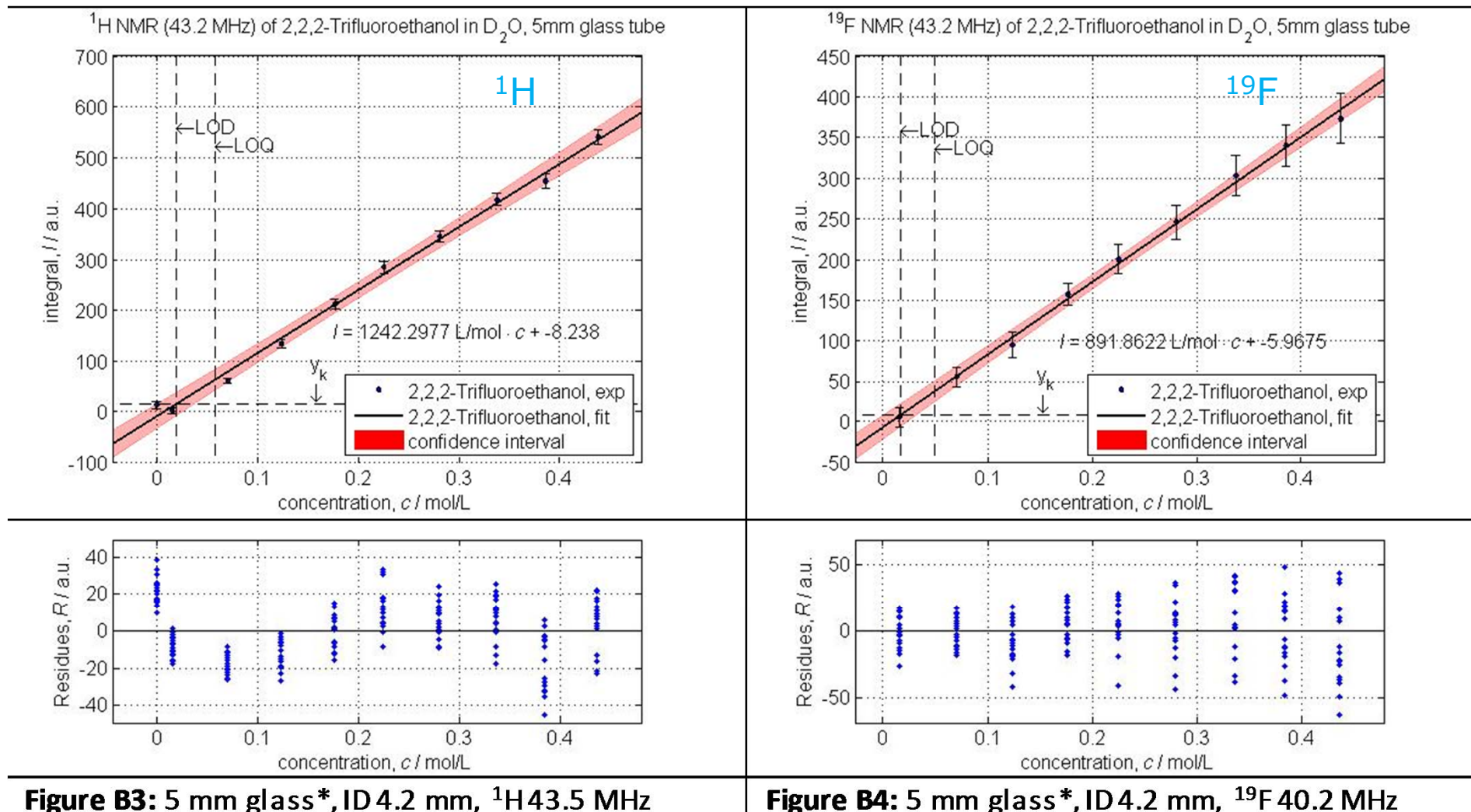
Zientek, N. et al. *Magn. Reson. Chem.* 54 (2016) 513–520

# LoD 43 MHz

## 5 mm Glass ID 4.2 mm (\*)



## 2,2,2-Trifluoroethanol in D<sub>2</sub>O



(\*)NMR tube: Deutero, Economic ID 4.20 ± 0.05 mm

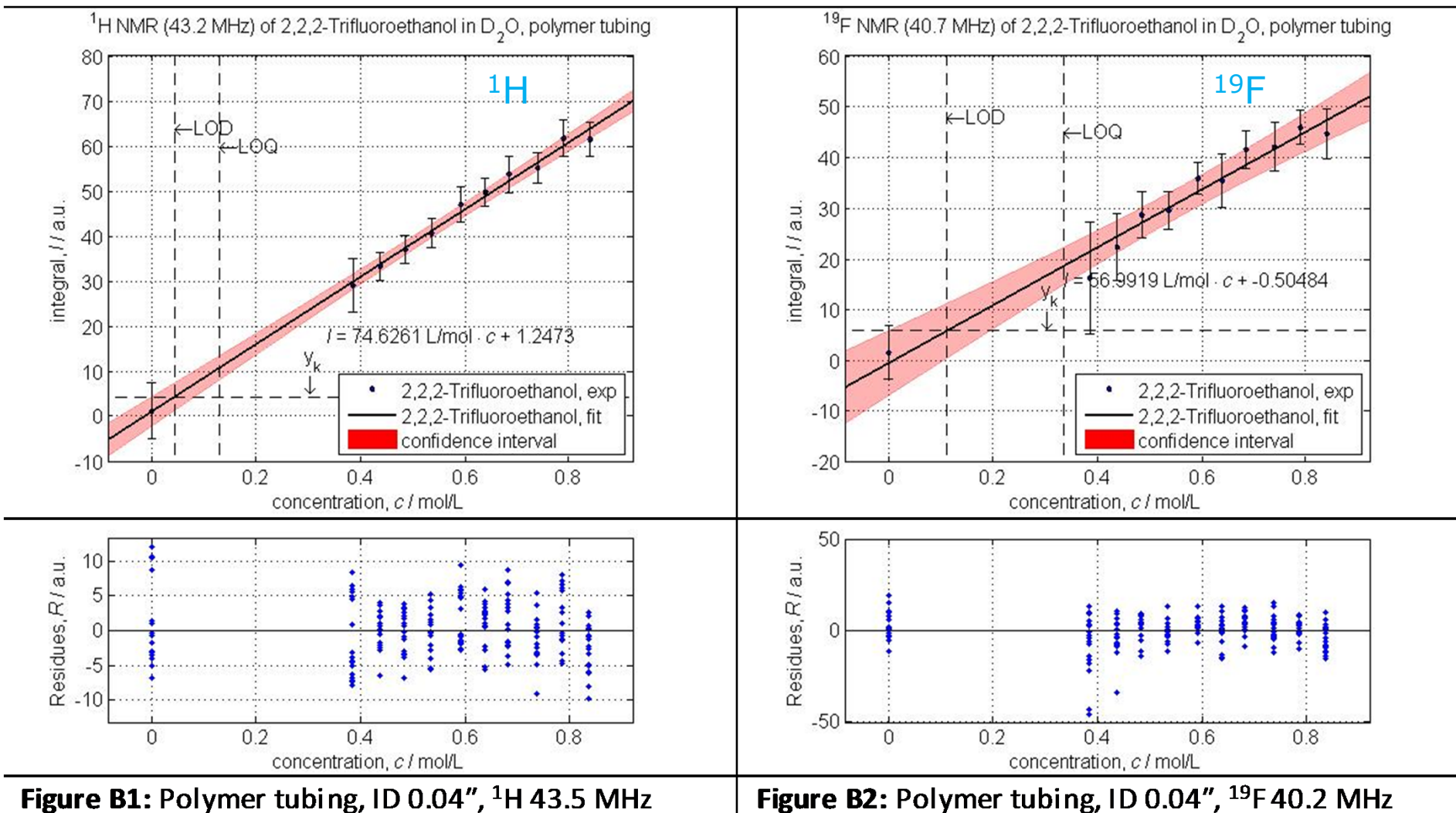
Zientek, N. et al. J Mag Res, 249 (2014) 53-62

# LoD 43 MHz

## Polymer Tubing ID 1 mm (0.04")



2,2,2-Trifluoroethanol in D<sub>2</sub>O



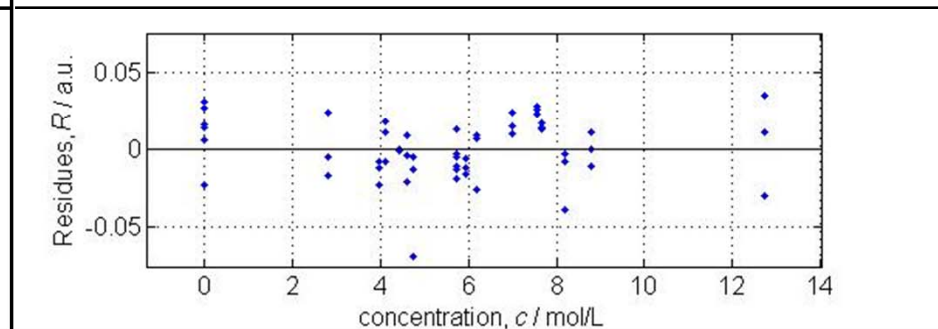
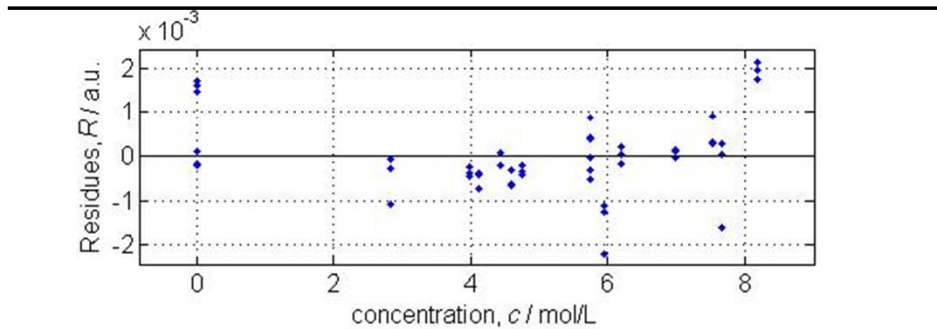
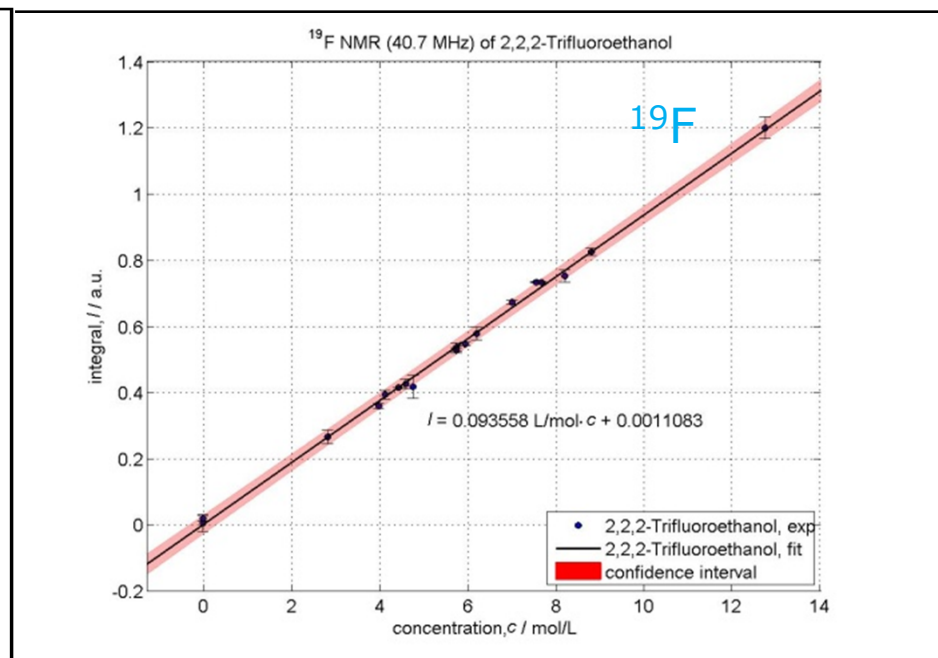
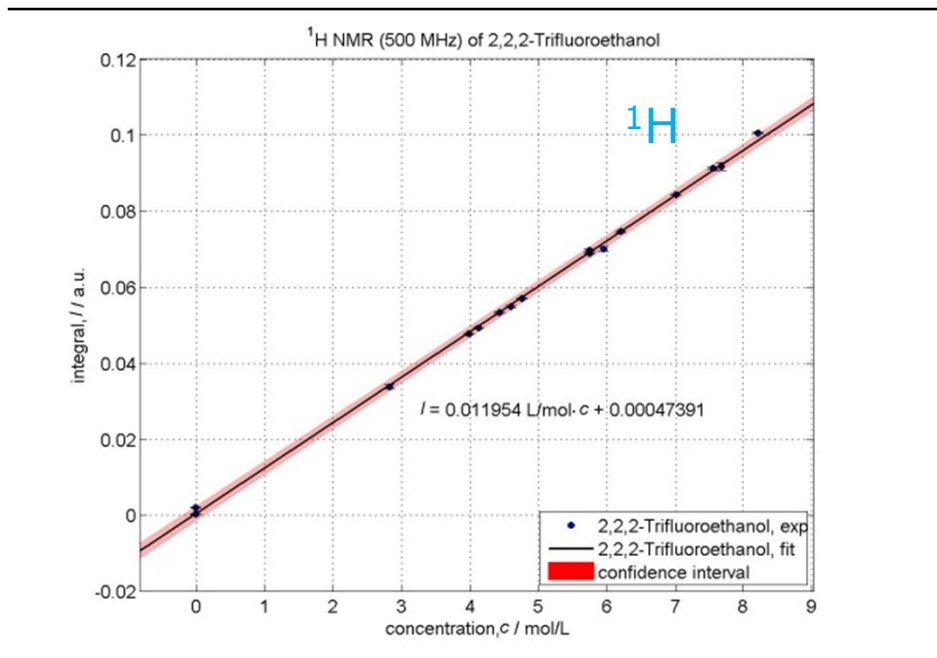
Zientek, N. et al. J Mag Res, 249 (2014) 53-62

# LoD 500 MHz

## 5 mm Glass ID 4.2 mm (\*)



2,2,2-Trifluoroethanol in D<sub>2</sub>O



**Figure C1: 2,2,2-Trifluoroethanol, <sup>1</sup>H NMR (500 MHz)**

**Figure C7: 2,2,2-Trifluoroethanol, <sup>19</sup>F NMR (40.7 MHz)**

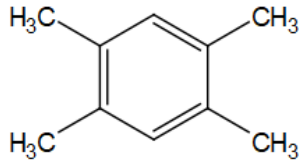
(\*)NMR tube: Deutero, Economic ID 4.20 ± 0.05 mm

# Calibrating With Linear Functions

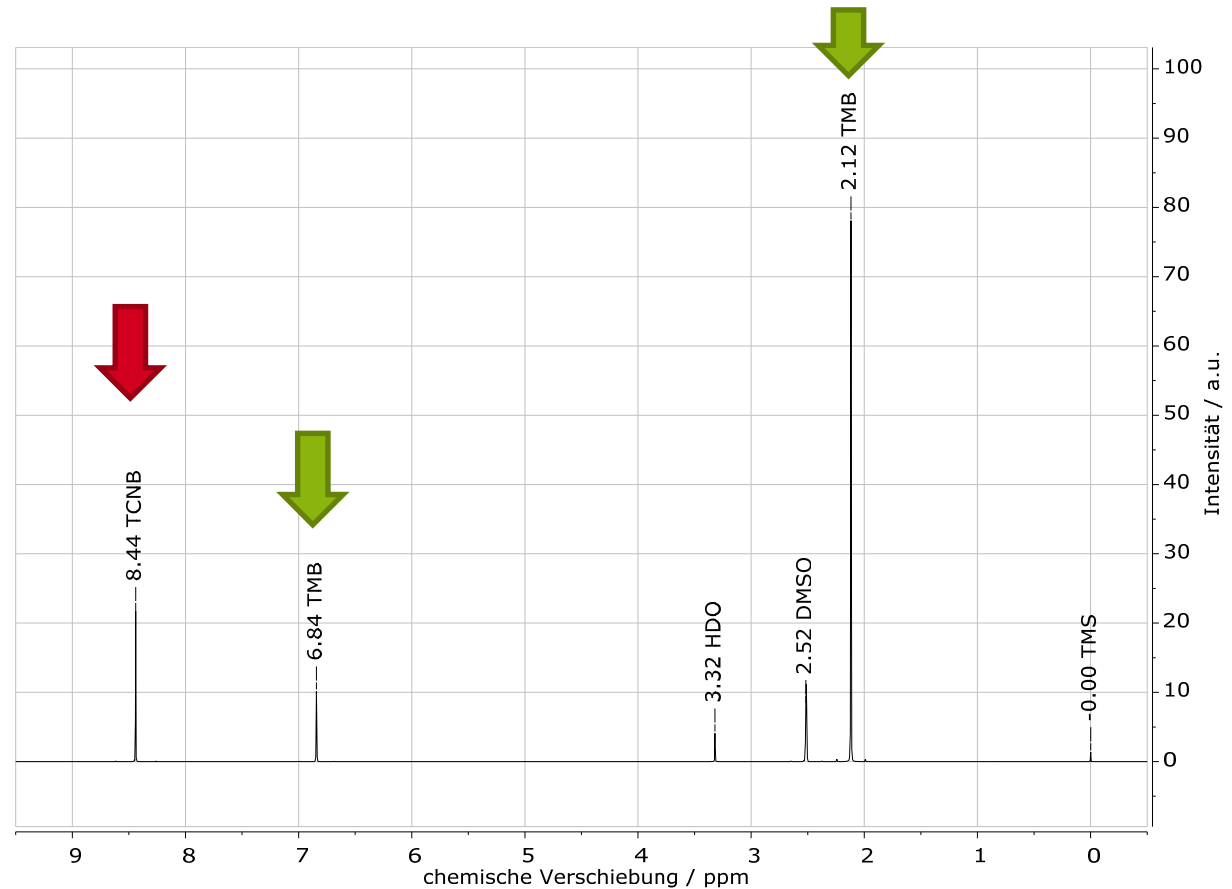
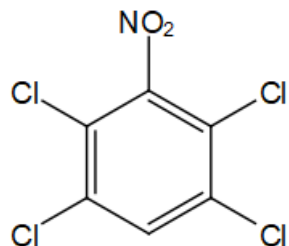
## Example 3: Proposal for a low-field NMR spectroscopy comparison study (43 and 500 MHz)

1,2,4,5-tetramethyl-benzene and 1,2,4,5-tetrachloro-3-nitrobenzene

1,2,4,5-tetramethyl-  
benzen



1,2,4,5-tetrachloro-3-  
nitrobenzen



# Calibrating With Linear Functions

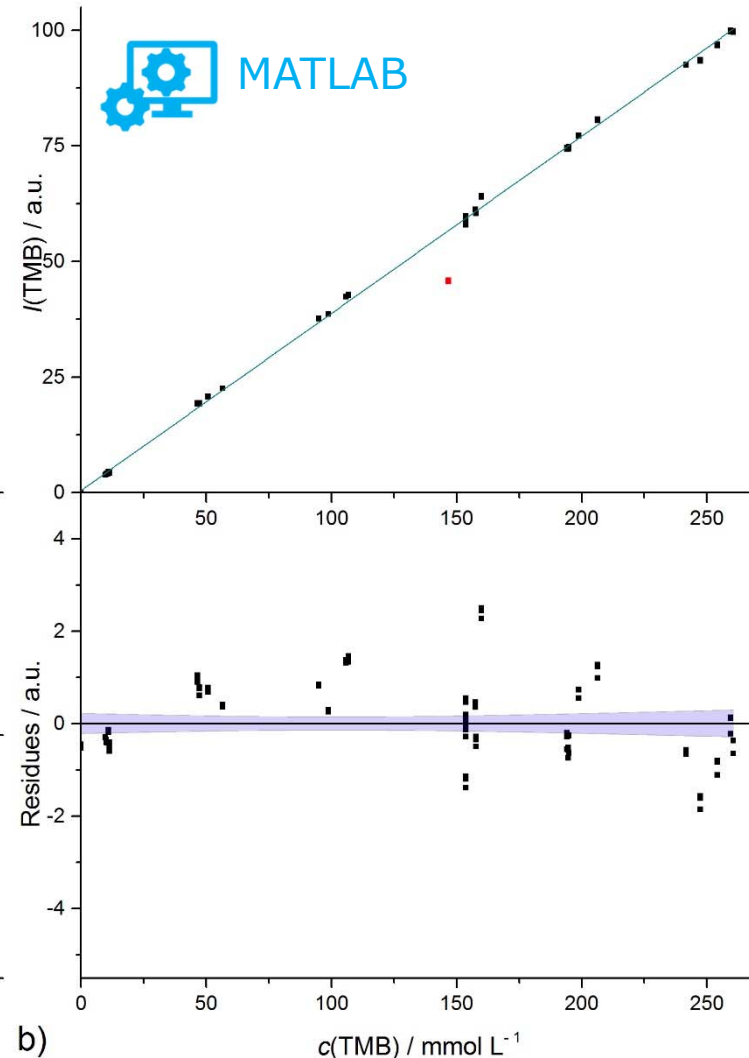
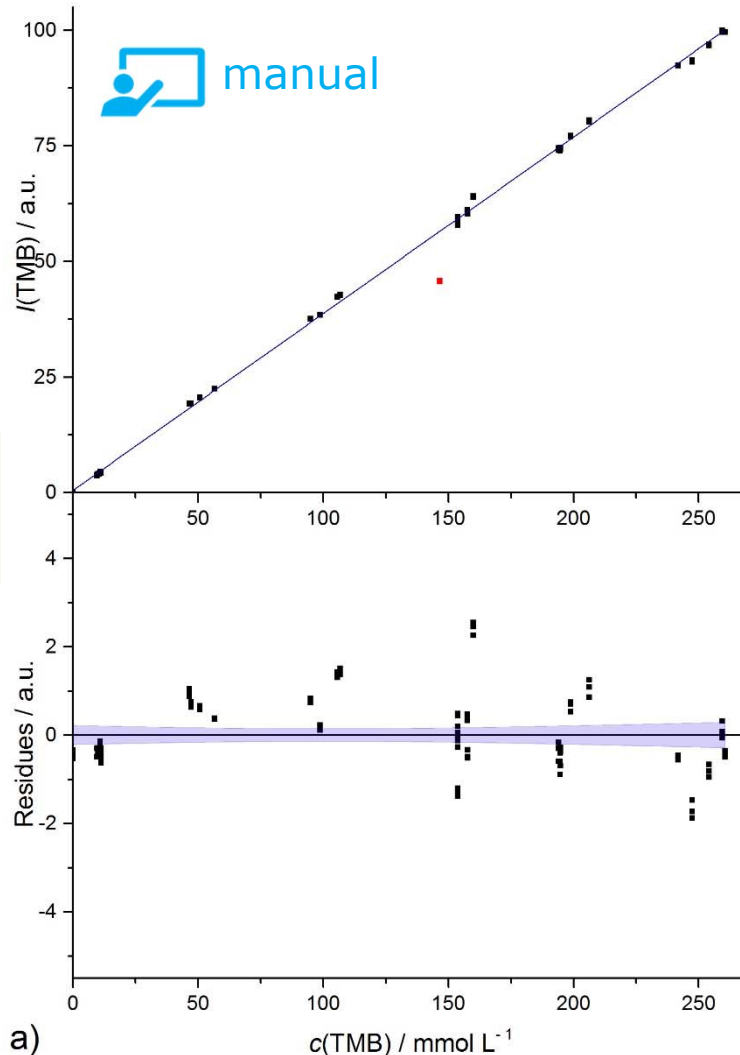
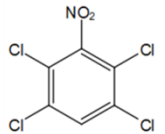
## Example 3: Proposal for a low-field NMR spectroscopy comparison study (500 MHz)

Linearity  
at 500 MHz

 manual

 MATLAB

1,2,4,5-tetrachloro-3-nitrobenzen



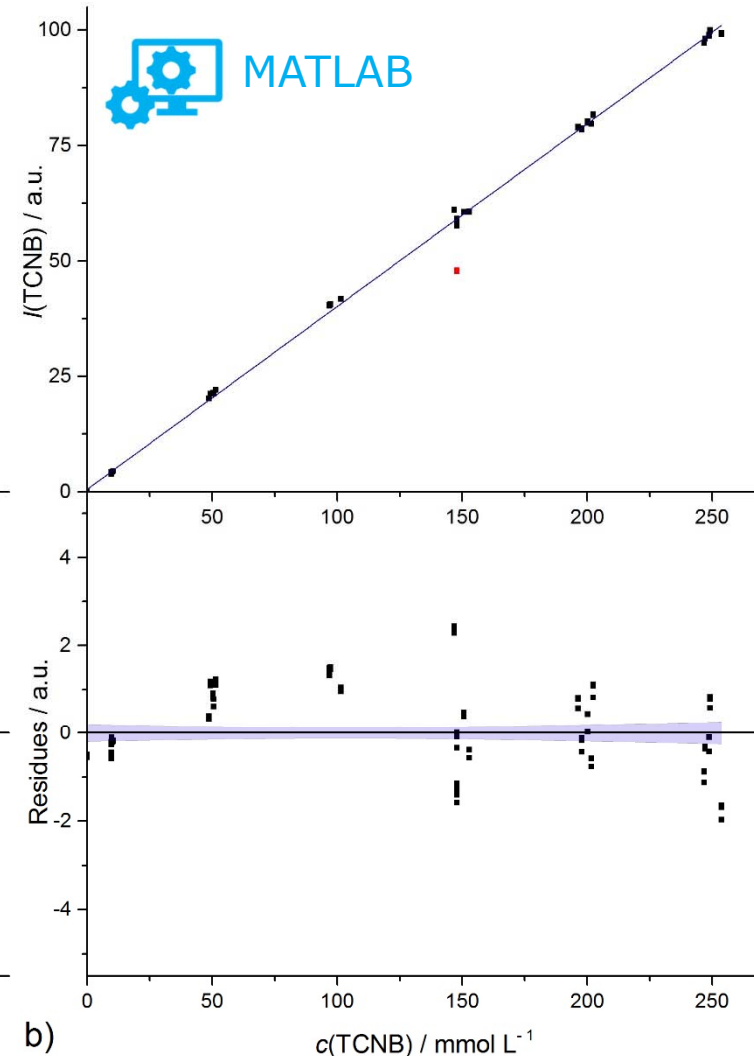
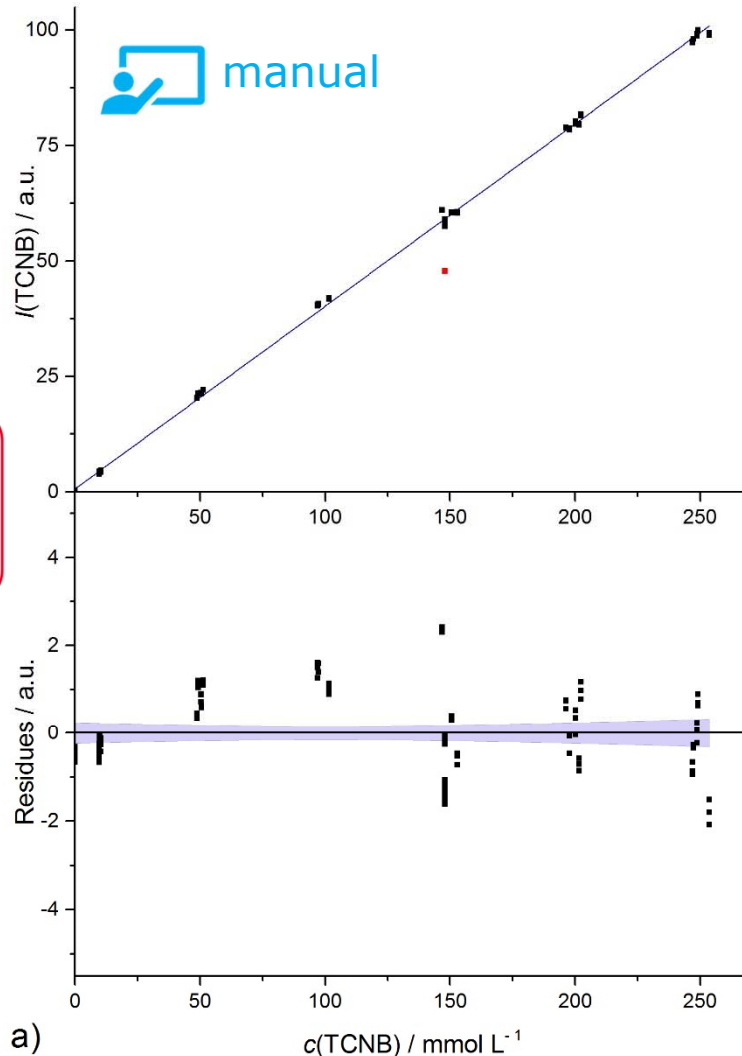
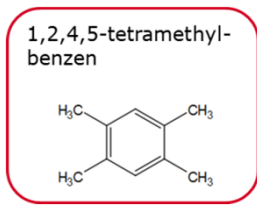
# Calibrating With Linear Functions

## Example 3: Proposal for a low-field NMR spectroscopy comparison study (500 MHz)

Linearity  
at 500 MHz

 manual

 MATLAB





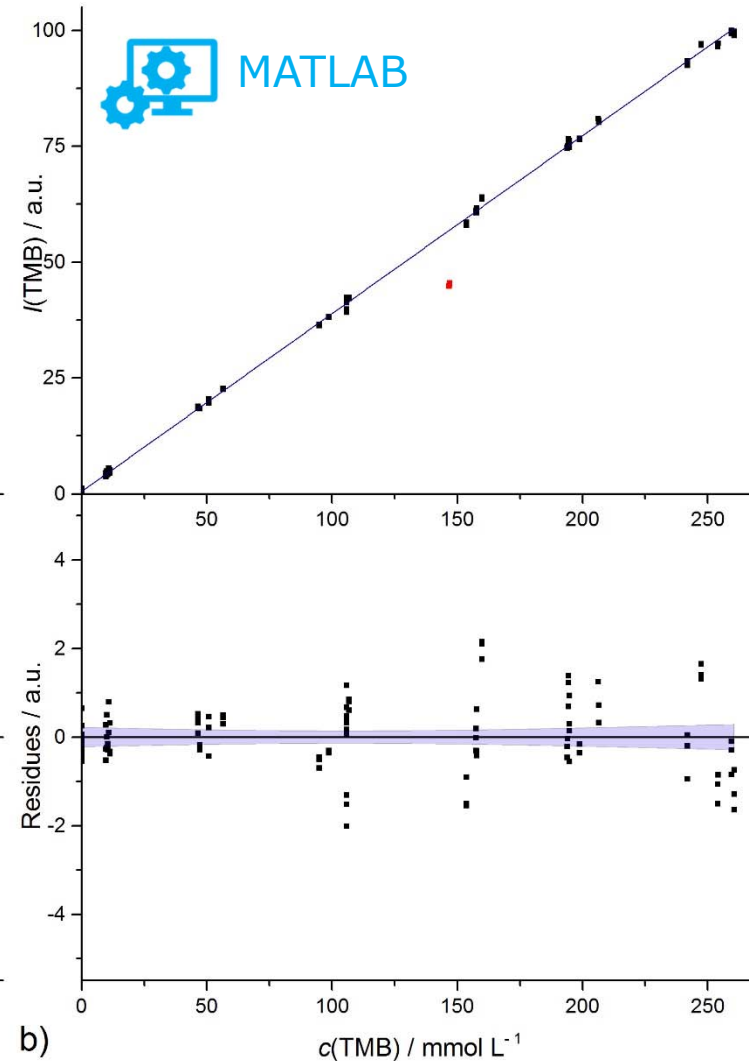
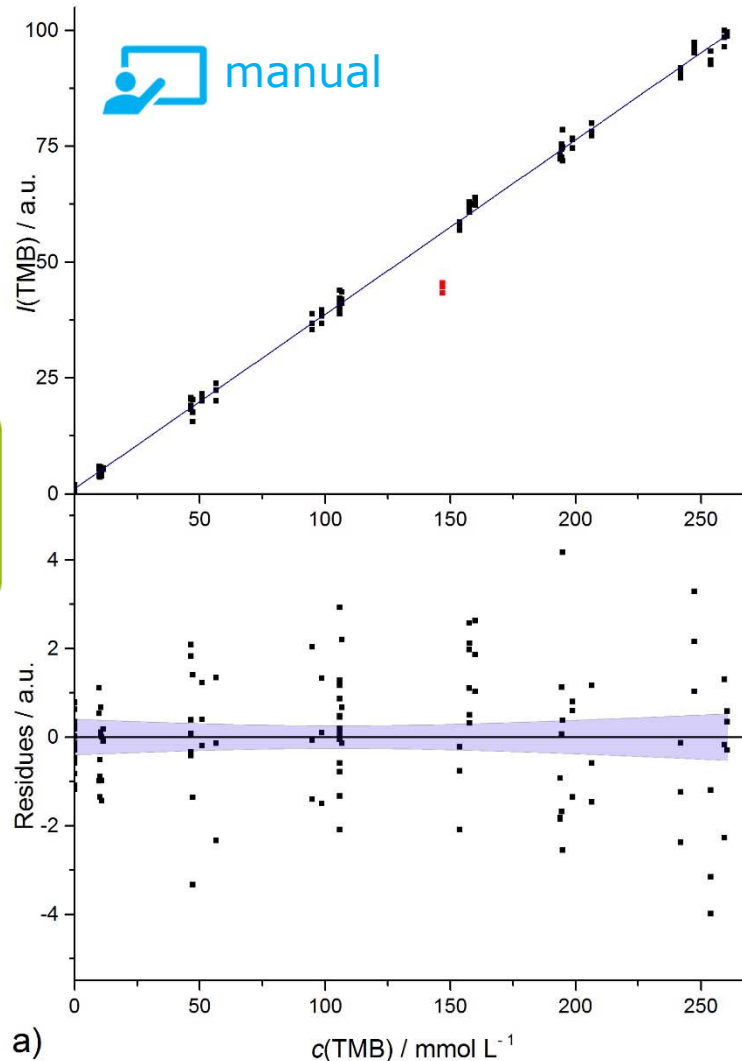
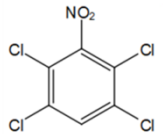
# Calibrating With Linear Functions

## Example 3: Proposal for a low-field NMR

### spectroscopy comparison study (43 MHz)

Linearity  
at 43 MHz

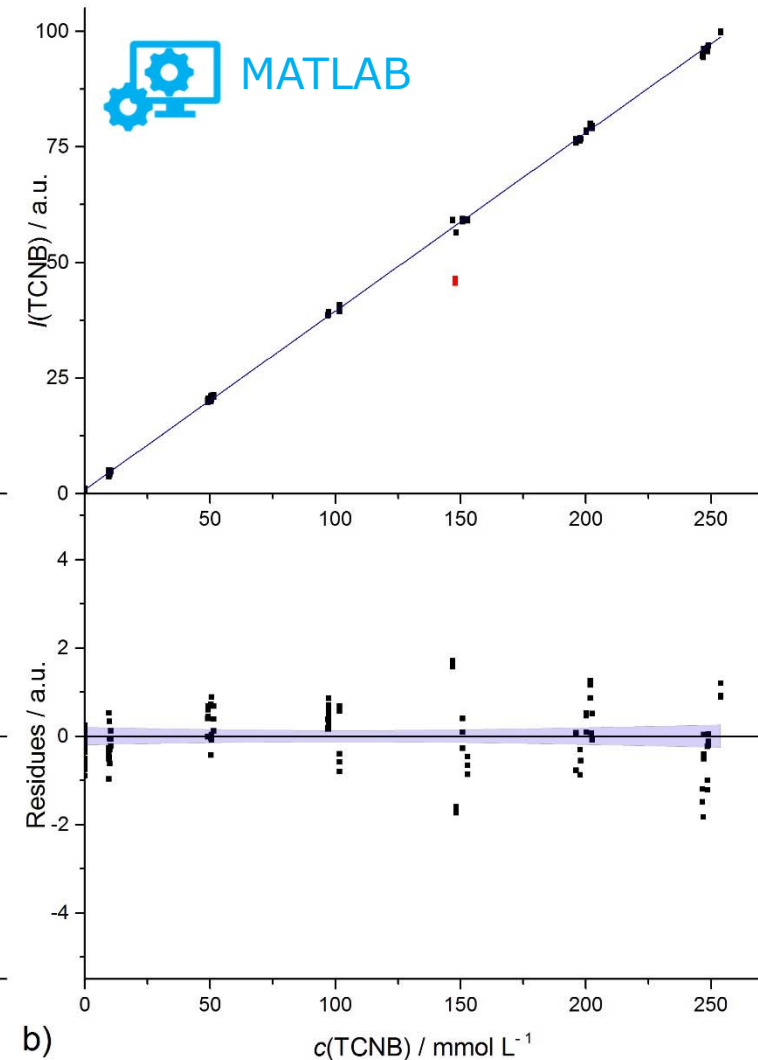
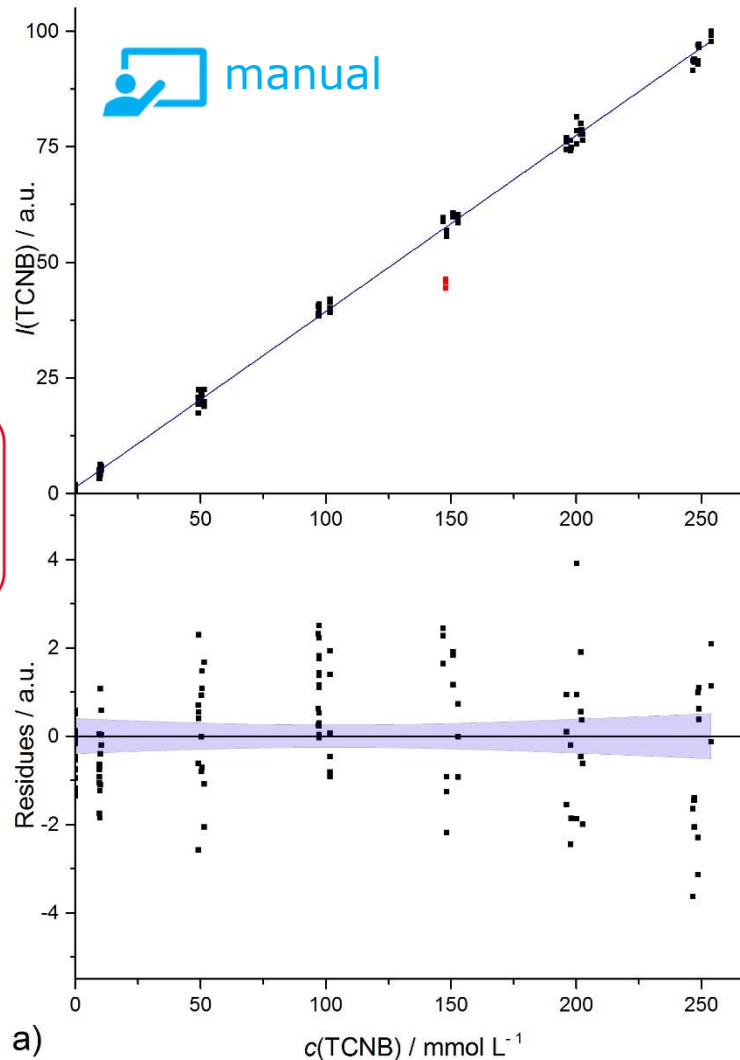
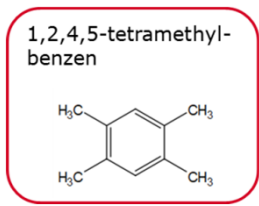
1,2,4,5-tetrachloro-3-nitrobenzen



# Calibrating With Linear Functions

## Example 3: Proposal for a low-field NMR spectroscopy comparison study (43 MHz)

Linearity at 43 MHz



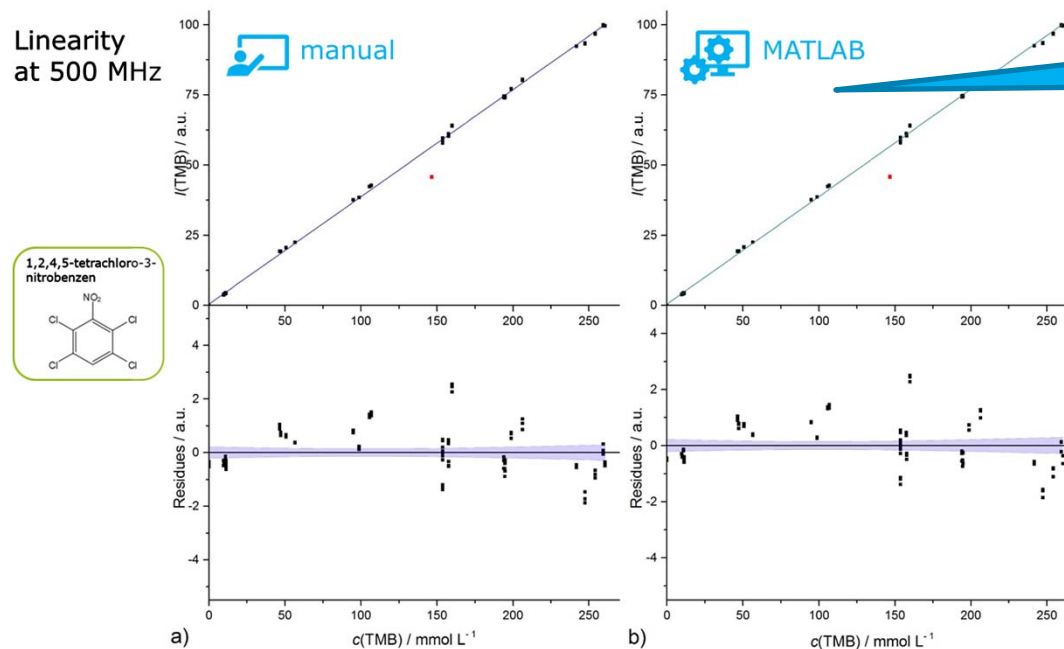
# Significance Tests – Example 1

## Significance of Precision of two Data Analysis

### Methods

Example: Comparison study (500 MHz)

- Two different data analysis methods (manual vs. MATLAB)



Are they different?

Do they yield the same standard deviation?

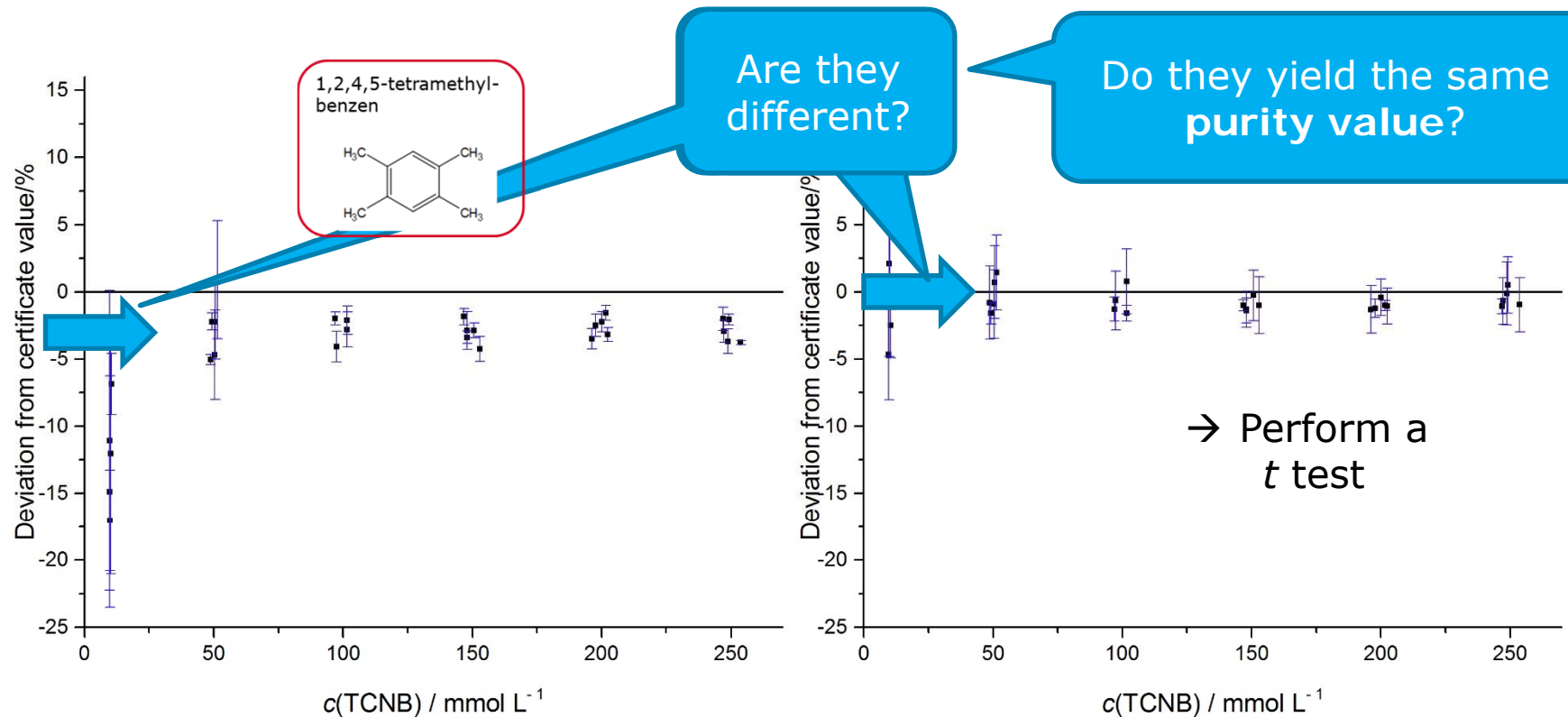
→ Perform a *F* test

# Significance Tests – Example 2

## Significance of Difference in Certificate Values from two Data Analysis Methods

Example: Comparison study (43 MHz)

- Two different data analysis methods (manual vs. MATLAB)

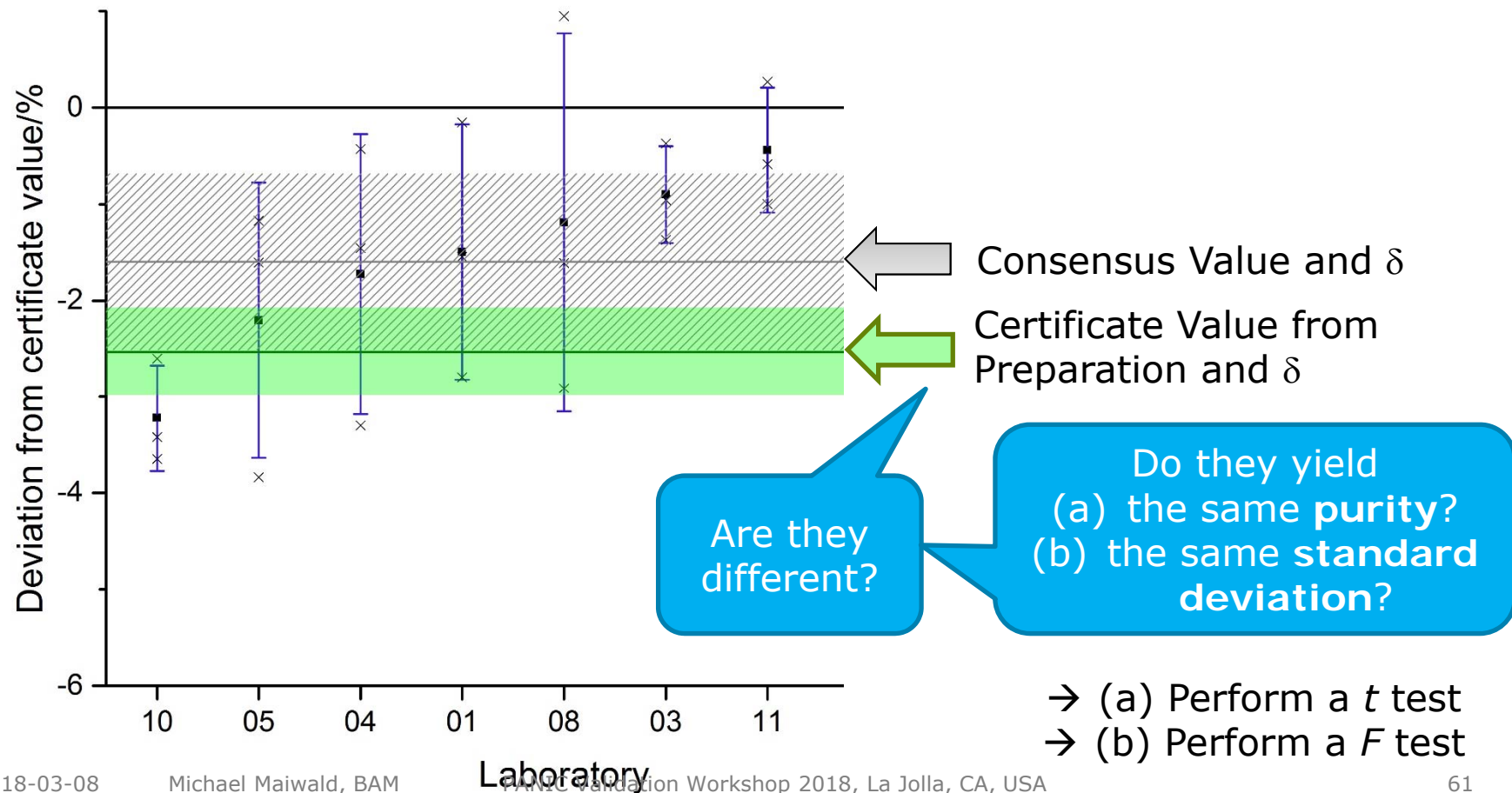


# Significance Tests – Example 3

## Significance of Difference in Certificate Values from Preparation and Consensus Values

Example: Comparison study (43 MHz)

- Two expected values along with their standard deviations



---

## Part 4: Significance Tests

- T-Test
- F-Test
- ANOVA

# Significance Tests – Overview



---

A decision at a given level of confidence about a population is based on observations from a sample of the population

## *t* test

- Testing for significant difference between the
  - means and a reference value
  - two data sets (difference of means) or
  - difference between pairs of measurement

## F test

- Testing for significant difference between the spreads of two data sets (difference of  $s$ )

# One-sided/Two-sided Probabilities



One Side

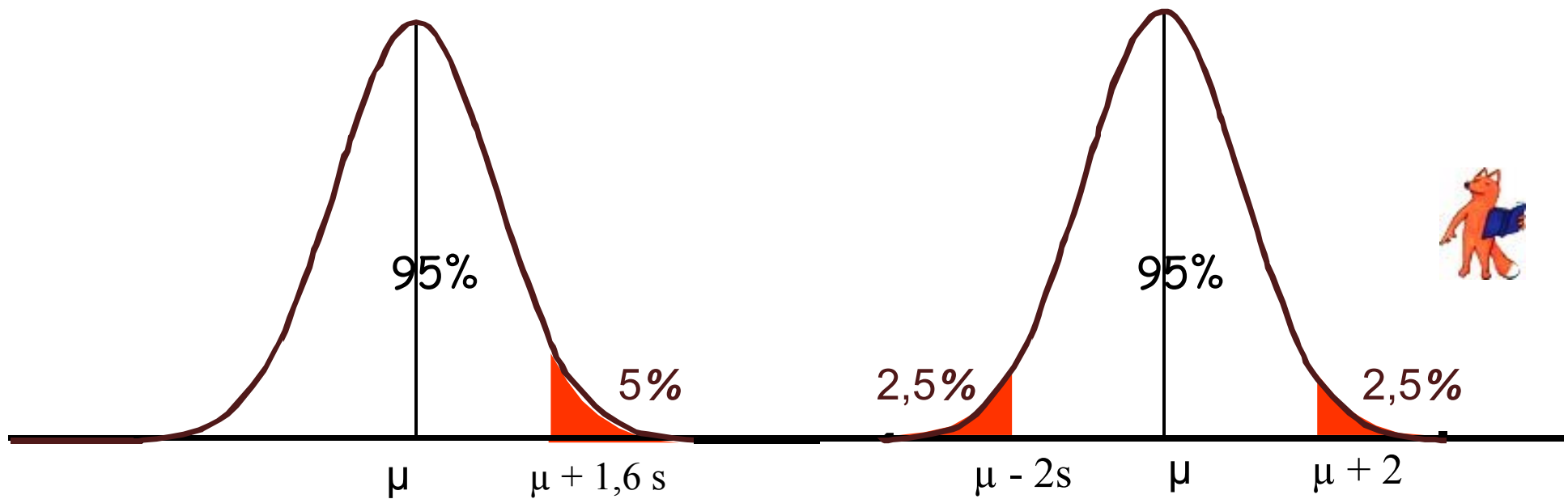
Two Sides

□ Probability that  $x$  is less than  $\mu + 1,6 s$

□ Probability that  $x$  is within the range  $\mu \pm 2s$

■ Probability that  $x$  is greater than  $\mu + 1,6 s$

■ Probability that  $x$  is outside the range  $\mu \pm 2s$





# Significance Test in 8 Steps



- 
1. Formulate the question
  2. Select the test
  3. Decide on one or two-sided test
  4. Choose the level of significance
  5. Define Null and Alternative hypothesis
  6. Determine the Critical value
  7. Evaluation of the test statistic using the appropriate equations
  8. Decisions and conclusions



# Significance Test – Step 5



- 5. Define Null and Alternative hypotheses

*Null hypothesis  $H_0$*

- The term null is used to imply that there **is no** difference between the observed and known value, other than that which can be attributed to random variation ( $\mu = x$ )

*Alternative hypothesis  $H_1$*

- the opposite of the null hypothesis - there **is** a difference ( $\mu \neq x$ ), where  $x$  is a sample mean,  $\mu$  is a true value



# Significance Test – Step 6



- 6. Determine the critical value

The *critical value* for a hypothesis test is a threshold to which the value of the test statistic in a sample is compared to determine whether or not the null hypothesis is rejected.

Critical value for any hypothesis is set by

- the level of significance required
- degrees of freedom
- whether the test is one-sided or two-sided

Critical values can be found in tables/textbooks.



# One Sample $t$ -Test



Comparison of experimental mean with a reference value or a nominal value

$$t_{calc.} = (\bar{x} - x_0) / \frac{s}{\sqrt{n}}$$

$s$  is a sample standard deviation,  $n$  sample size,  
 $\bar{x}$  sample mean, and  $x_0$  the stated value

$t_{crit}$  value for  $\alpha = 0.05$  and degree of freedom  $\nu = n-1$



# Two Samples *t*-Test



Comparison of two experimental means, when two samples are drawn from a population with different standard deviations

$$t_{\text{calc.}} = \frac{(\bar{x}_1 - \bar{x}_2)}{\sqrt{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)}}$$

$$v = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)}{\left(\frac{s_1^4}{n_1^2(n_1 - 1)} + \frac{s_2^4}{n_2^2(n_2 - 1)}\right)}$$

$v = n_1 + n_2 - 2$   
*for same sd*



# Two Samples *t*-Test



x_1	x_2				
			two-sample t test, same variance		
10.08795	10.3435			var 1	var 2
13	12		mean	14.5	13
15	15		variance	4.000004694	3.000014192
17	10.3		observations	16	14
18.2	13		pooled variance	3.535723389	
12	12		hypothetic difference of the mean	0	
15	11		dof	28	
16.6	15.6		t statistics	2.179794237	
14	14		P(T<=t) one-sided	0.018919954	
13	13		critical t value, one-sided	1.701130908	
15	13		P(T<=t) double-sided	0.037839908	
15	15		critical t value, double-sided	2.048407115	
14	14.7				
13.2	13.0565				
15			two-sample t test, different variances		
15.91205				var 1	var 2
16	14		mean	14.5	13
14.5	13		variance	4.000004694	3.000014192
4.000004694	3.000014192		observations	16	14
			hypothetic difference of the mean	0	
			dof	28	
			t statistics	2.201395058	
			P(T<=t) one-sided	0.018058168	
			critical t value, one-sided	1.701130908	
			P(T<=t) double-sided	0.036116337	
			critical t value, double-sided	2.048407115	



# F-Test



---

The  $F$  test establishes if there is a significant difference between variances.

$F$  test considers the ratio of two sample variances, i.e. ratio of the squares of the standard deviations,  $s_1^2/s_2^2$ .

Answer the question: are the spreads different, i.e., do the two sets of data come from two separate populations?

## Statistics for the $F$ Test

The  $F$ -value is calculated according to equation

$$F_{calc.} = s_1^2/s_2^2, \text{ where } s_1^2 > s_2^2$$

The ratio is compared with  $F_{critical}$  values from tables.

$F_{critical}$  is based on two values of degrees of freedom  $\nu_1 = n_1 - 1$  and  $\nu_2 = n_2 - 1$ .



# Significance Tests – Back to Example 2

## Significance of Difference in Certificate Values from two Data Analysis Methods



### Typical *t*-test in ORIGIN

Paired Sample *t* Test (09.01.2018 18:28:06)

Notes

X-Function	Paired Sample <i>t</i> Test
User Name	claud
Time	09.01.2018 18:28:06

Input Data

	Data	Range
1st Data Range	[Book60]Sheet1!B"Konsens automatisiert"	[1*:7*]
2nd Data Range	[Book60]Sheet1!C"Konsens manuell"	[1*:12*]

Descriptive Statistics

	N	Mean	SD	SEM
"Konsens automatisiert"	7	97,06	1,67	0,634
"Konsens manuell"	7	98,26	0,91	0,346
Difference		-1,195		

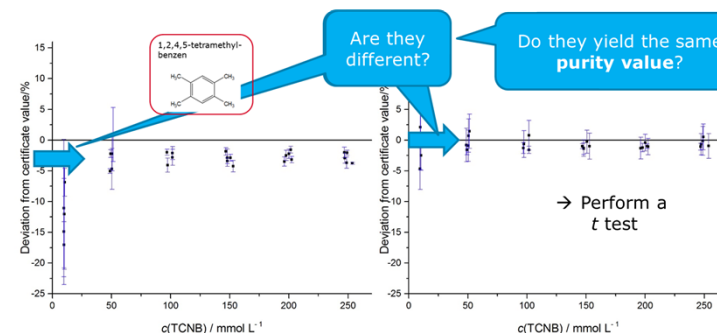
Test Statistics

t Statistic	DF	Prob> t
-3,88305	6	0,00814

Null Hypothesis: mean1-mean2 = 0  
 Alternative Hypothesis: mean1-mean2 <> 0  
 At the 0.05 level, the difference of the population means is significantly different from the test difference(0).

Example: Comparison study (43 MHz)

- Two different data analysis methods (manual vs. MATLAB)



Datasets:  
Are they different?

They are!





# $t$ and $F$ values from EXCEL



Mappe1 - Microsoft Excel nichtkommerzielle

Start Einfügen Seitenlayout Formeln Daten Überprüfen Ansicht Add-Ins

Einfügen Zwischenablage

Schriftart Ausrichtung Zahl Bedingte Formatierung

SUMME  $\times$   $\checkmark$   $f_x$  =TINV(0.05, 8)

	A	B	C	D
1				
2				
3		=TINV(0.05, 8)		2.3060041
4				

	A	B	C	D
1				
2				
3		=FINV(0.05, 4, 4)		6.3882329
4				



# ANOVA – What is it Good For?



---

ANOVA separates different sources of variation

- *Variation within a group*
- *Variations between groups*

Estimates of variance can be compared

ANOVA can be applied to any data, that can be grouped by a (many) particular factor (s)

ANOVA is used to compare sets of data



# Thank You!



---

Acknowledgements to Wolfram Bremser, BAM

Literature:

- GUM: Guide to the Expression of Uncertainty in Measurement  
<https://www.bipm.org/en/publications/guides/gum.html>

Contact: michael.maiwald -at- bam.de

---