SEE
Measurement as Process

Paul G. Plöger

Björn Kahl

Outline

General: Target

Learn how to **design** experiments decently

Learn how to **conduct** experiments decently

Learn how to **describe the results** decently (a.k.a. in a scientifically sound way)

1 Some recent BAD Examples

4,8376 +/- 0.1

	Precision	Recall	Time taken (s)
Instantanuous diagnosis			
Control run	-	-	1
Exogenous Intervention run	1	1	1
Progressive diagnosis			
Control run	-	-	1
Exogenous Intervention run	1	1	1

TABLE 6.1: Calculated precision and recall of exogenous intervention diagnosis in a simulated youbot. The calculation was done on a specific time interval when the exogenous interventions was occurring.

	Assigned positive	Assigned negative
Actual positive	60	12
Actual negative	48	0

Table 5.2: Error matrix for PADI dataset

	Assigned positive	Assigned negative
Actual positive	103	4
Actual negative	13	0

Table 5.3: Error matrix for NUS dataset-I

$$s = \frac{1}{2} a * t^2 \Rightarrow s = \frac{1}{2a} * a^2 * t^2 \xRightarrow{v=a*t} s = \frac{1}{2a} * v^2 \Rightarrow v = \sqrt{s * 2a}$$

Abbildung 30: Herleitung der Formel zur Geschwindigkeitsberechnung

Da diese Geschwindigkeit, je nach zufahrender Distanz, die maximale Geschwindigkeit des Motors um ein Vielfaches übersteigen kann, ist ein Vergleich mit dieser sinnvoll. Um ein zu abruptes Anfahren des Motors zu verhindern, wird dieser mit einer konstanten Beschleunigung geregelt. Um diesen Faktor nicht auszuhebeln, wird die Geschwindigkeit auch mit diesem Wert verglichen (siehe Abbildung 31).

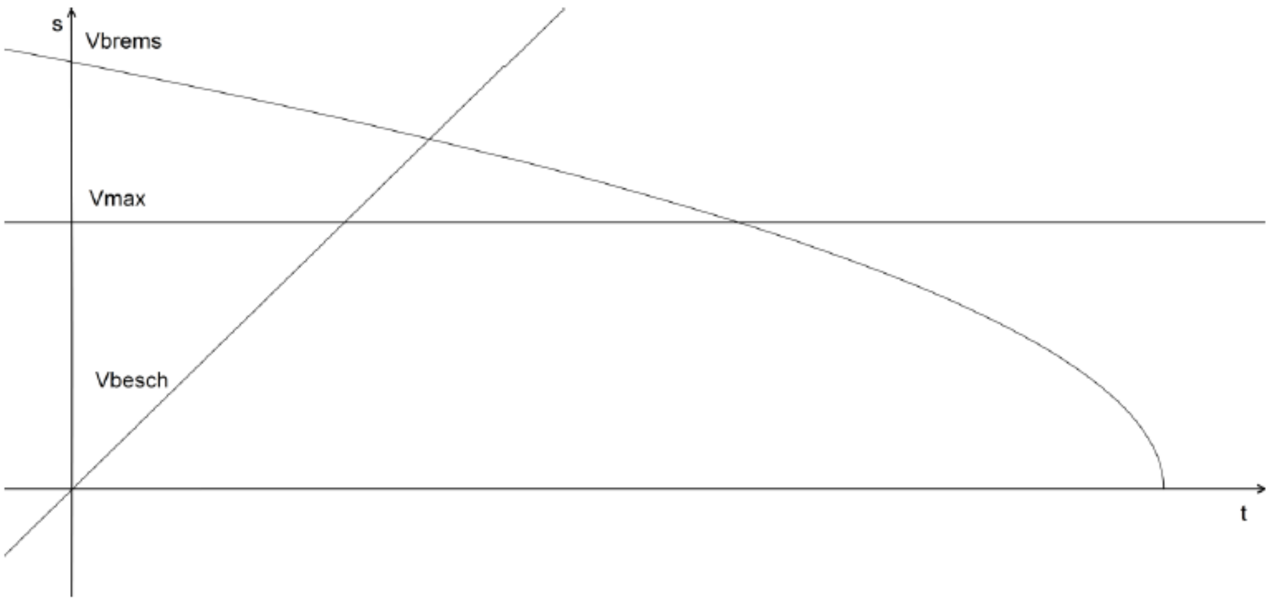


Abbildung 31: Ermittlung der zu fahrenden Geschwindigkeit

Daraus ergibt sich eine Formel für die zufahrende Geschwindigkeit (siehe Abbildung 32).

$$v = \min(Vmax, Vbeschleunigt, Vgebremst)$$

Abbildung 32: Formel zur Ermittlung der zufahrenden Geschwindigkeit

1 General: Knowledge areas

Theory of Measurement
("Messwerterfassung")

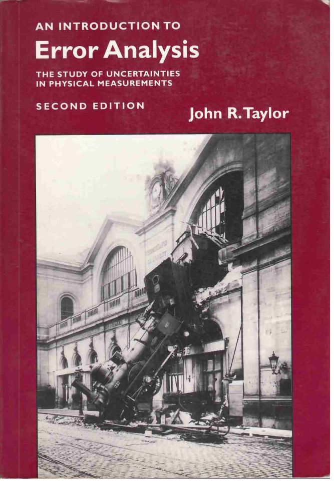
Mathematics

Probability theory

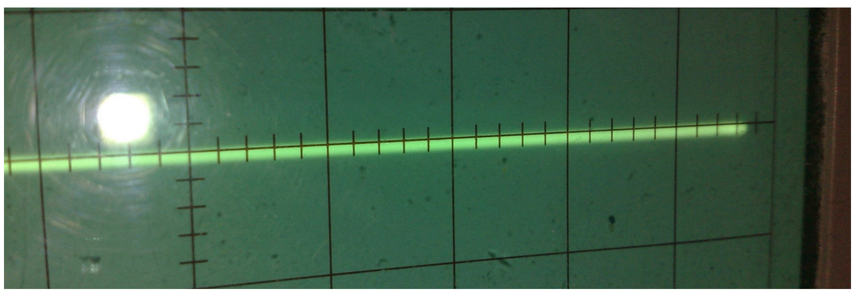
Statistics

Physics

Instrumentation



1 Reading results from experiments



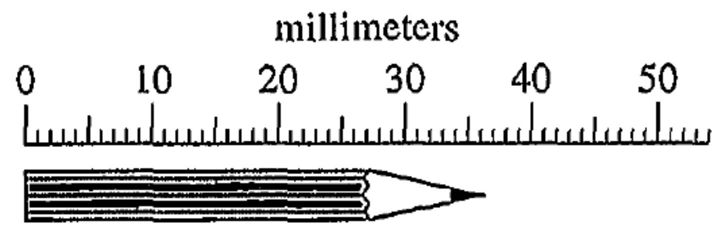
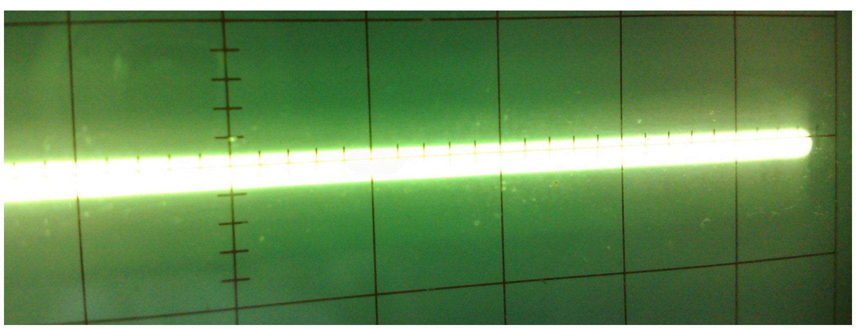


Figure 1.2. Measuring a length with a ruler.



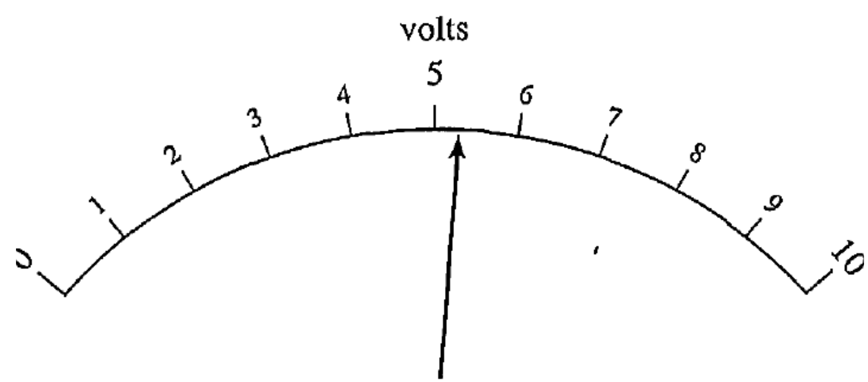
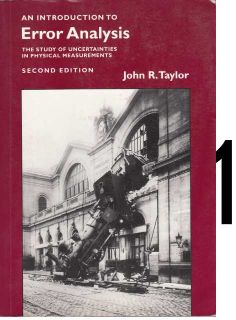


Figure 1.3. A reading on a voltmeter.



1 Best Estimate ± Uncertainty

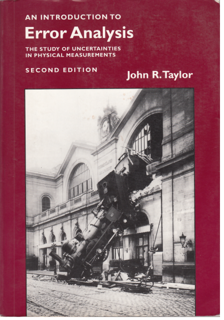
E.g.: measured value of time = 2.4 ± 0.1 s.

$$(\text{measured value of } x) = x_b \pm \delta x$$

x_b : experimenter‘s best estimate

δx: uncertainty, error, margin of error (it‘s > 0)

i.e. confidence, that x lies in range
$x_b - \delta x$ and $x_b + \delta x$



1 Rules for Stating the Uncertainties

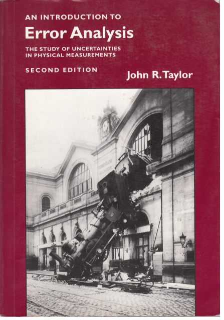
e.g. measure the acceleration of gravity g:

(measured g) = 9.82 ± 0.02$\cancel{385}$ m/s^2

?

> Experimental uncertainties should almost always be rounded to ONE significant figure

e.g.: **(measured g) = 9.82 ± 0.02 m/s^2**



1 Rules for Stating the Answers

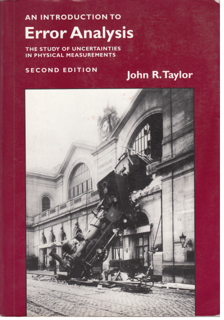
After uncertainty in measurement => significant figures in the measured value must also be considered.

e.g.: **measured speed = 60~~517~~8 ± 30 m/s**

The uncertainty of 30 means that the digit 5 might really be as small as 2 or as large as 8. Clearly the trailing digits 1, 7, and 8 have no significance at all and should be rounded

The **last significant figure** in any stated answer should usually be of then **same order of magnitude** (in the same decimal position) as the uncertainties

measured speed = 6050 ± 30 m/s



1 Examples

92.81 with an uncertainty of **0.3** $\Rightarrow$ **92.8** ± 0.3

92.81 with an uncertainty of **3** $\Rightarrow$ **93** ± 3

92.81 with an uncertainty of **30** $\Rightarrow$ **90** ± 30

Observe:

To reduce inaccuracies caused by rounding, **any numbers to be used in subsequent calculations** should normally **retain at least one significant figure more** than is finally justified. At the end of the calculations, the **final answer should be rounded** to remove these extra, insignificant figures.

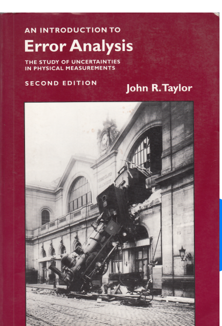
1 Quick check (10min)

Rewrite each of the following measurements in its most appropriate form:

a) $v = 8.123456 \pm 0.0312$ m/s

b) $X = 3.1234 * 10^4 \pm 2$ m

c) $m = 5.6789 * 10^{-7} \pm 3 * 10^{-9}$ kg.



PGP	30.09.13	lecture	Measurement as process
		Testat:	-none-
PR+PGP		Lab class	Run robots on sheet of paper, record distributions
you		Homework	Evaluate data, Do the fit
PR	07.10.13	lecture	Sample solution for 30.09.13
PGP+BK		Testat:	successfull Reprodution a normal distribution in 2D and proof that what you observed is actually a normal distribution, error analysis
PGP+BK		Lab class	read Zhang paper, explain it to newbies in Vision and projective geometry
you		Homework	finish undestanding of Zhang paper
PGP+BK	14.10.13	lecture	Intrinsic / extrinsic camera parameters, radial distortions, calibration
		Testat:	-none
PGP+BK		Lab class	run the CALTECH sw, do the calibration images, find intrisic / extrinsic params for your camera
you		Homework	Read Thrun / Probabal. Robotics / chapter 5.1, 5.2, 5.3
BK	21.10.13	lecture	Sampling of distributions and forward / backward model (Use of Prob. Models in mobile robotibs)
PR+BK		Testat:	correct intrinsic / extrinsic camera matrixes (compare on class level)
PR		Lab class	run the V,\omega experiemtent
		Homework	read 2 articles: On the repseentation and estimation of spatial uncertainty // Location estimation und uncertainty analysis for mobile robots
PGP	28.10.13	lecture	Covariance matrices and error propagation
		Testat:	-none
PR		Lab class	complete the experiment
you		Homework	Read Thrun / Probabal. Robotics / chapter 5.4, 5.5, 5.6
	04.11.13	lecture	-none
PGP+BK		Testat:	good v,\omega model für robot
PR		Lab class	do the ODOmetry Modion model // ticks_left ticks_right
you		Homework	Read Thrun / Probabal. Robotics / chapter 6

General: Overview

General: Overview

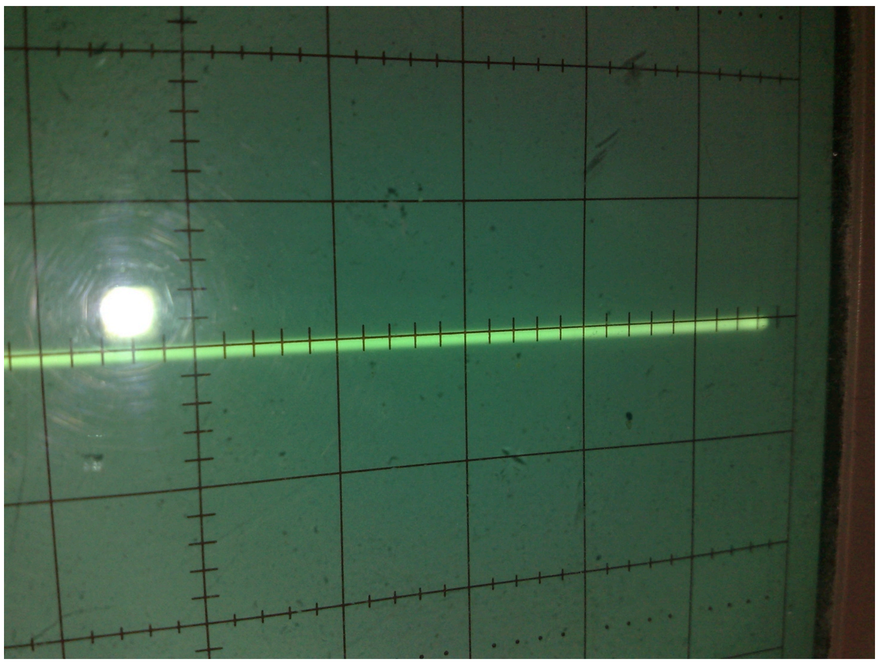
PGP	11.11.13	lecture	EM and Xpero Simulator, explanations
PGP+BK		Testat:	good ODO model
PR		Lab class	Build sonar radar on NXT robot, start experiemtens in cage
you		Homework	Refactor my matlab code in python (class project)
PGP	18.11.13	lecture	-none
PGP+BK		Testat:	-none
PR		Lab class	finish experiments on sensor model
you		Homework	read article: ROC Graphs: Notes and Practical Considerations for Researchers
BK	25.11.13	lecture	Confusion matrix and ROCs
PGP+BK		Testat:	good sensor model in XPEROSim
PR		Lab class	setup SVNs, learn hand-written digit classifier
you		Homework	read CT articles, hand in data piles from your R&D experiments
PGP	02.12.13	lecture	Data Visualization // based on Tufts book
PGP+BK		Testat:	good ROCS and explanation / comparisiosn: which classifier is best
PR		Lab class	get your own data from R&D, start visualize them
you		Homework	make nice diagrams
you	09.12.13	lecture	Explain your R&D data representation problem
PGP+BK		Testat:	on nice diagrams
		Lab class	TDB
you		Homework	TDB
you	16.12.13	lecture	Explain your R&D data representation problem
PGP+BK		Testat:	on nice diagrams
		Lab class	TDB
you		Homework	TDB

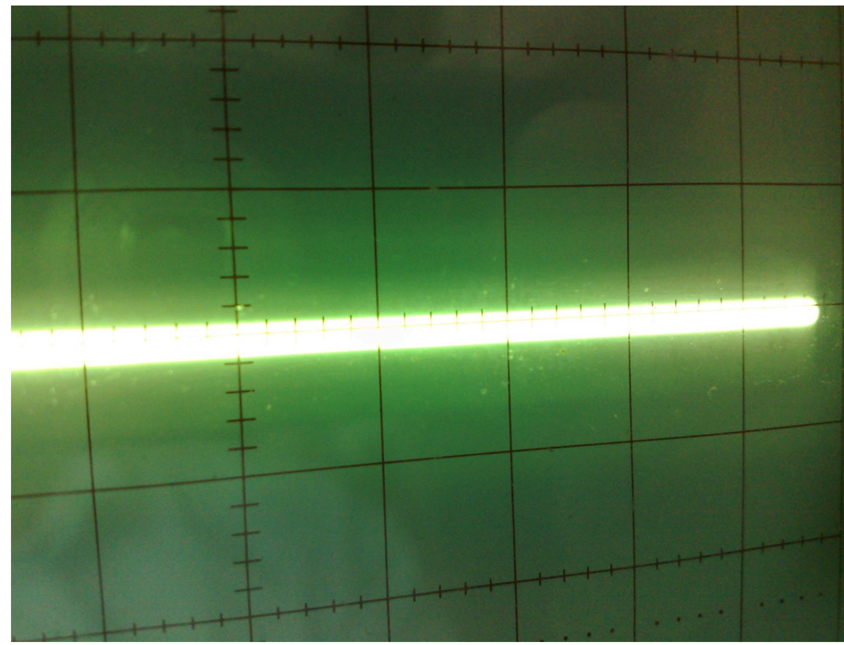
General: Overview

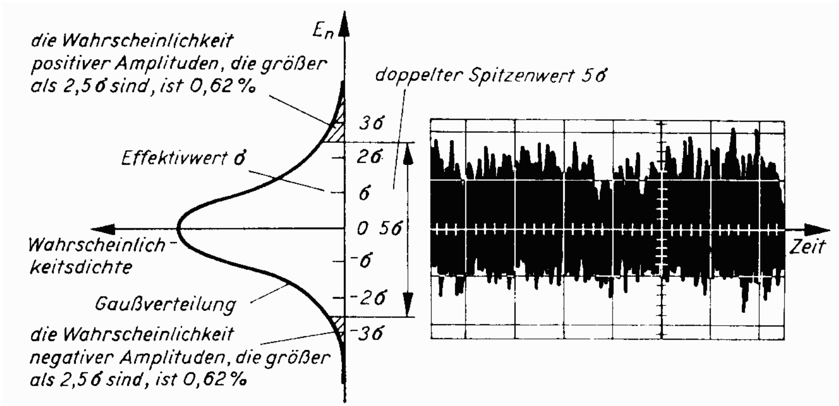
BK	06.01.14	lecture	you bot as a learning tool
		Testat:	-none
PR		Lab class	run youbot to get circle / square in air, record sensor traces
you		Homework	compare true ground to MEMS values and vision
BK	13.01.14	lecture	you bot as a learning tool
		Testat:	-none
PR		Lab class	run youbot to get circle / square in air, record sensor traces
you		Homework	compare true ground to MEMS values and vision
BK+PGP	20.01.14	lecture	resume / spare lecture
BK+PGP		Testat:	closing the gap between the sensor traces
PR		Lab class	none
you		Homework	none

- 75% must be solved and be attested (i.e. 6 out of 8)

1: Oscilloscope experiment







1: Questions

What is the true value (x_t) ?

Can we catch it "exactly"? Why, why not?

If not catchable, how can we sure there *IS* a true value?

How to approximate it best, what is it?

How can we be sure that the measuring instrument did not add (too much) distortion? (called „burden")

How to validate that the observed values actually follow a distribution? Which?

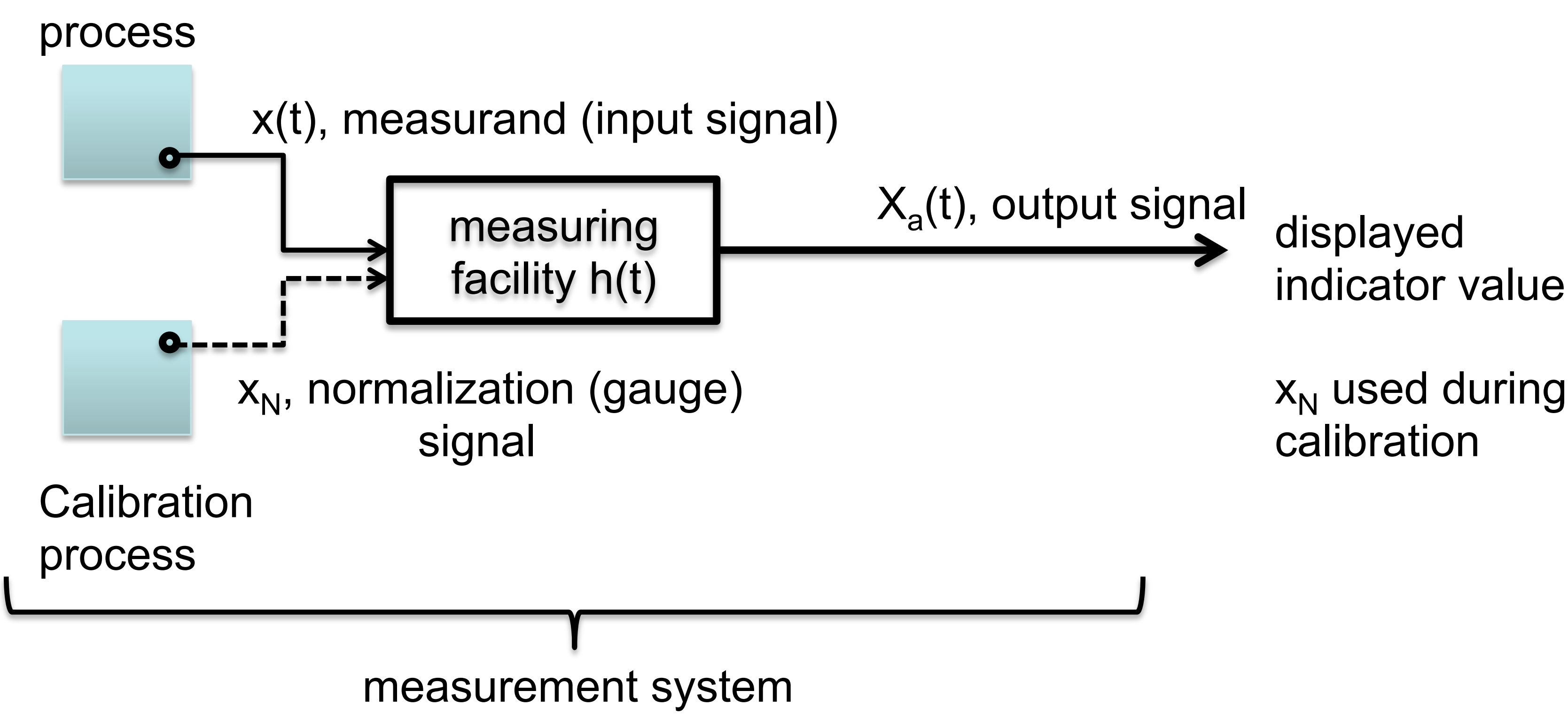
How to formalize the process as such?

How to interpret the measurement action in physical terms?

1: needed Sciences

- Mathematics
  - errors: absolute, relative, systematic, random, round off
  - (Random variables)
  - (Probability distributions)
- Statistics
  - Empirical mean value
  - (Maximum likelihood estimation)
  - Statistical tests
- Physics / System Theory
  - (Burden of instrument)
  - (Equivalent circuit diagram)
  - Measurement as signal flow graph

2B: System theory: Measuring System



$x(t)$ == measurand
$h(t)$ == transfer function of measuring facility (linear) =>
$x_a(t) = h(t) * x(t)$

2B: Formalization general terms

Measurand:

The measurand is **physical quantity, which is to be quantified** by the measurement (e.g. length, pressure, electrical resistance, etc.).

Measurement:

To „measure" in the **strict sense** is, to **determine the size of a variable as a multiple** of a to be determined generally recognized unit size of the same physical dimension through **experimental comparison** with a solid measure of this unit.

To „measure" **in wider sense** is, to **determine experimentally** a quantitative information of a particular property of a test object, which **may be derived from one or more measurable quantities of this test object**. This property may be a size (for example of a force or efficiency), or a mapping (for example a development over time). This property can also describe a system state, characterized via several variables which are recorded in parallel.

e.g. Special case of Measurement Count

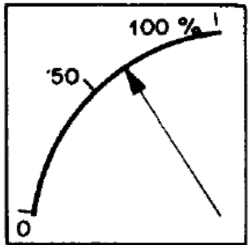
Counting is to **determine the number of elements** (e.g. piece count) or events (e.g. RPM). The count represents a special case of measurement.

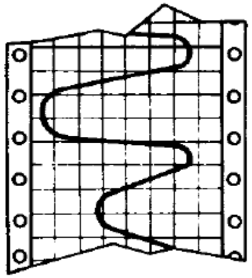
e.g. Special case of Measurem. Inspection:

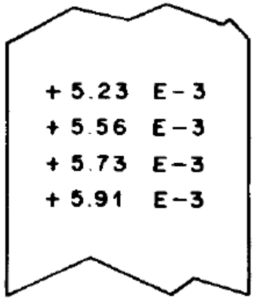
Inspecting means to determine experimentally whether a **particular property of a test piece corresponds to the specified requirements**. If the check is carried out via measuring, it can be considered a special form of measurement .

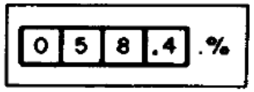
Display:

The **indicator shows the numerical value** of the to be measured variable. In analog displays one has to determine the **state of the indicator (arrow) on the scale by „reading"**. In digital displays it can be read directly as a number.









(Display) Range:

The display range of a measuring device is that range of **measured values, which can be read from the scale of the display (full scale, upper limit)**

Measurement range:

The range is that part of the display area for which the **error remains within a guaranteed or prescribed limit of error.**

Suppression area:

The suppression area is that **area of measured values above which the display of the measuring device starts**.

2B: Formalization general terms

Measured (Quantity) Value :
This is the **value of a specific variable. It will be determined from the displaying measuring instruments by reading this display and building the product of measured value and the unit of the measurand** (eg 3m). It can also be output in the form of a transferable measurement signal and may be sent for further automatic processing (storage, value processing, control, etc.).

Measurement result:
The measurement result is obtained **in general from several measured values by the help of a predetermined relationship**. In the simplest case, a single measurement is already the measurement result.

Device Under Test (DUT):
The DUT is **that part of a physical system, which carries the measurement variable** .

Measurement facility:
The measuring facility is **the entirety of device components used for the purpose of the measurement**. This includes sensor for detection of the measured quantity, amplifiers, computing devices and the output devices to display the observed value and possibly other components.

Measurement System:
The measurement system includes not only the measuring facility but also **those areas of the physical system which contains the DUT, which do affect the measurement process,** in particular the data acquisition .

Meter:
The meter is a either a part or the whole of the measuring facility.

Measuring Principle:
The principle of measurement is **defined as the characteristic physical phenomenon which is used during the measurement** (for example, measurand: temperature, measurement principle: linear expansion, or thermoelectric effects, etc.).

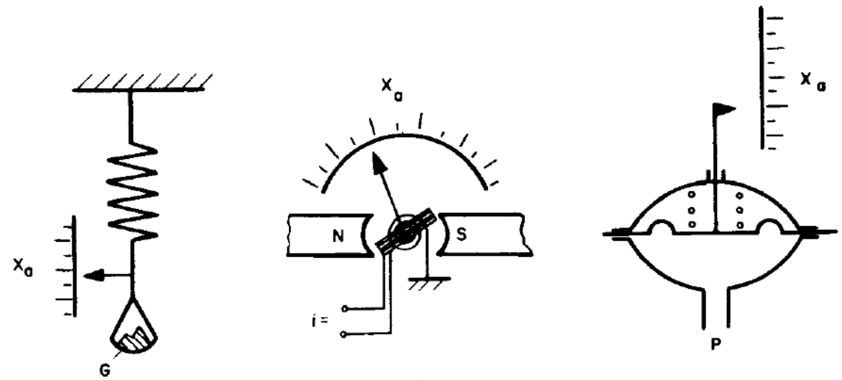
Measuring method:
Understand by this the kind of application or **implementation of the Measurement Principle**. In addition it may mean the function of the measuring device .

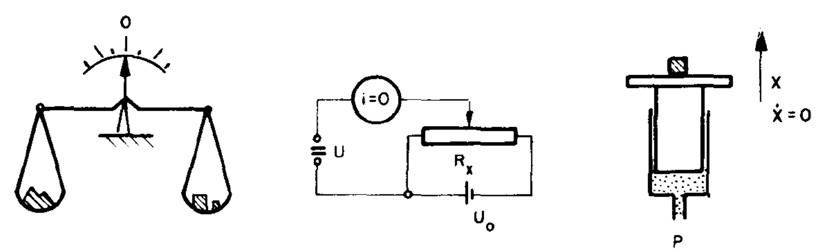
Sensitivity:
The Sensitivity is the **indicator pathway on the scale (in mm) per unit of the variable**. In light pointer devices, the vector length is defined as 1m. In digital instruments, the sensitivity is equal to the number of digits per unit of measured quantity. For non-linear meters the sensitivity is a function of the measured value or the display respectively.

2B:Example Methods

Methods:

- deflection of armatur
- compensation





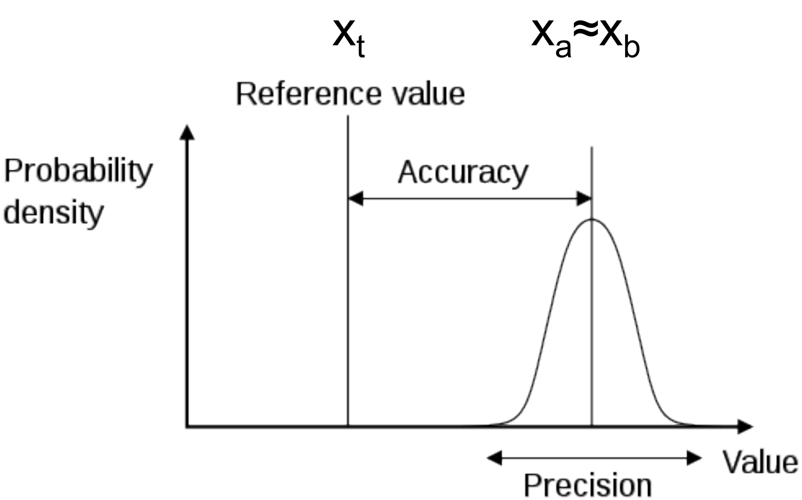
2B Quick check (10min)

Assume: it‘s a nice and very hot summer day and you like to measure the temperature of your desk at home while working.

- Apply the general formalized terms to this experiment
- What is measurand, what is DUT?
- ....

2B: Accuracy and Precision

x_t true value (reference, true grd), i.g. unknown, e.g. distance too wall
x_b "best" value, as close as we can get (assume: error free measurement, approximated by expected value $\mu \approx \overline{x}$, „right" value)
x_a "angezeigter" value, as currently displayed



Remark: earlier we bound this distribution by the term uncertainity δ.

(Measurement) accuracy

closeness of agreement between a measured quantity value and a true quantity value of a measurand.

(Measurement) precision (distribution)

closeness of agreement between indications or measured quantity values obtained by replicate measurements on the same or similar objects under specified conditions.

Precision is used to define measurement repeatability, intermediate measurement precision, and measurement reproducibility. Sometimes "measurement precision" is erroneously used to indicate measurement accuracy.

2E: Systematic and Random Errors Absolut error and relative Error

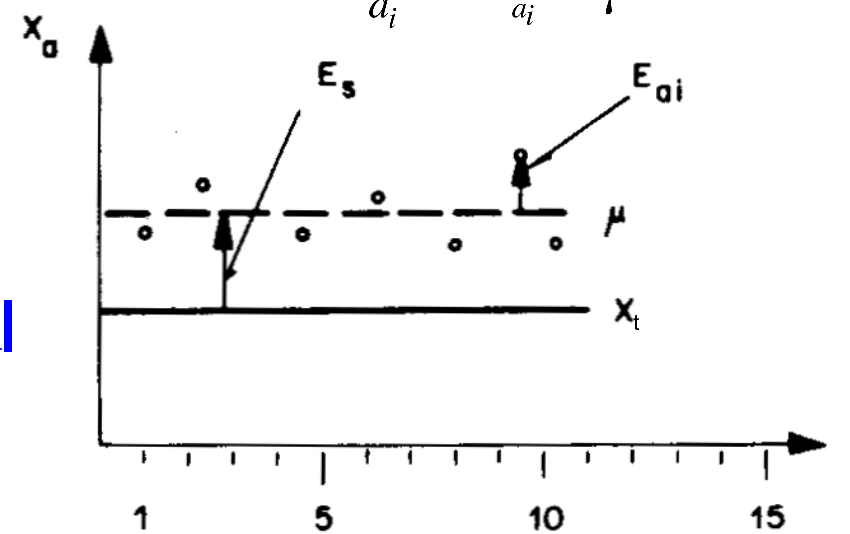
Any repetitive measurement shows two types of errors:

E_s := systematic error

E_{ai} := random error.

$$E_s = \mu - x_t$$

$$E_{a_i} = x_{a_i} - \mu$$



absolute error: $\Delta_{ai} := |x_{ai} - x_t|$

relative error : $\rho_{ai} := \Delta_{ai} / x_t$

2E Significant Figures notation

$c = (2.99792 \pm 0.00030) \cdot 10^8$ m/s

(extended notation)

$c = (0.299792 \pm 0.000030)$ Gm/s

(extended notation with unit prefix)

$c = 2.99792(30) \cdot 10^8$ m/s

(concise notation)

All intervals relate to the **absolute error Δ_{ai}** of the measurement

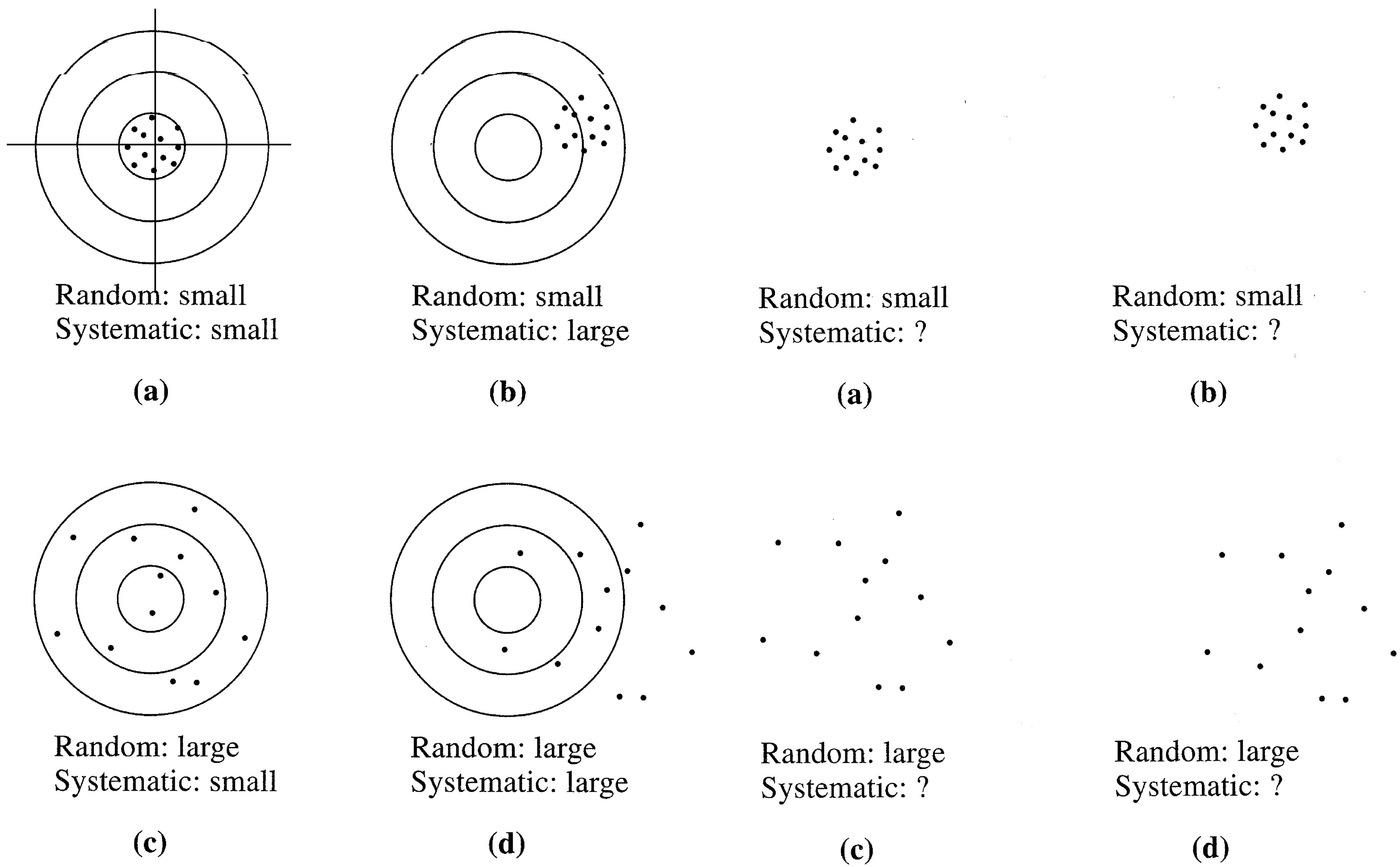
2E Quick check (10min)

Assume: it‘s a nice and very hot summer day and you happen to have a tape measure from metal. You a trying to measure the length of the classroom and the tape measure stayed on the window sill all day.

- When you now actually measure the length which kind error will you observe?
- Will it be positive or negative? Why?
- Can we cure it? How?
- What SIZES do you expect?

Example

2E: Interplay: Systematic vs. Random



Target known => can judge systematic errors
target unknown => only random errors can be judged

shoots to a known /
unknown target

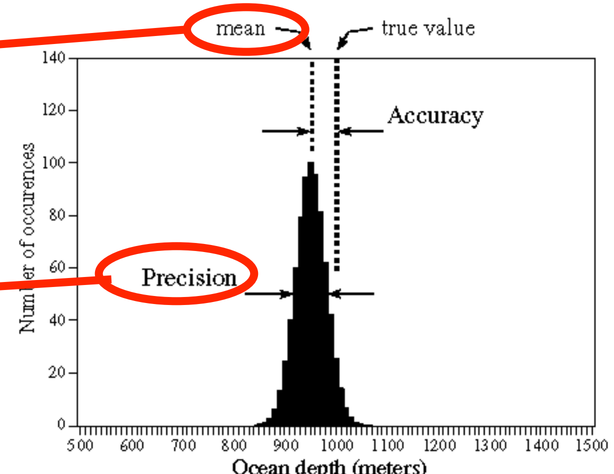
2E: How to estimate x_b?

Do repeated empirical measurements => use samples

The empirical expectation is called “sample mean value”
similar: empirical variance S_x (“S”treuung or “S”catter):

$$\overline{x}_{an} = \frac{1}{n}\sum_{i=1}^{n} x_{ai}$$

$$S_{xa}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_{ai} - \overline{x}_n)^2$$



One can show: these are optimal estimators (maximum likelihood estimators)

Furthermore for the expected value μ:

$$\mu = \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} x_i$$

Precision depends of the meter (repeatability).

Accuracy depends on all components of the measurement facility.

2E: we may **cure** systematic errors E_s **by calibration**

E_s can be repeated, every time it is the same thus it is a deterministic error.

- caused by factors that can (in theory) be modeled -> prediction
- e.g. **calibration** of a laser sensor
- e.g. distortion caused by the optic of a camera

E_{ai} is a random error -> non-deterministic

- no prediction possible
- however, they can be described probabilistically
- e.g. Hue, instability of camera, black level noise of camera .

Calibration:

An illustrative example is the calibration of a self-indicating weighing by putting of normalized weights. Taking into account systematic effects (previously determined by calibration measurement deviations of the weights, air pressure, temperature, buoyancy forces) and random influences the display of the scale is compared with the launched mass and then we estimate the uncertainty of this deviation.

A simple calibration result is: The weight displays at a load of 200 g, a deviation of +0.12 g, this result has an uncertainty of 0.20 g with a confidence interval of 95%.

(wikipedia on Calibration)

2E: example calibration

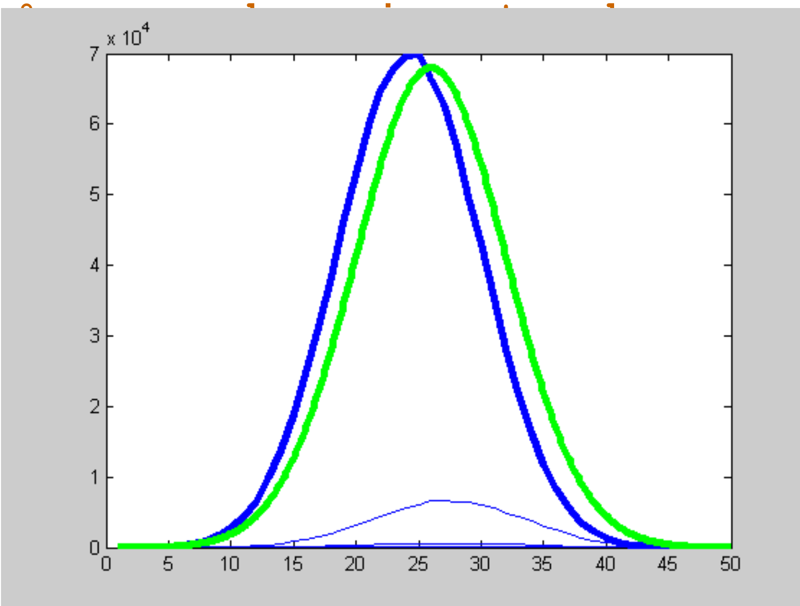
a b

Fig. 7.6. **Radial distortion correction.** *(a) The original image with lines which are straight in the world, but curved in the image. Several of these lines are annotated by dashed curves. (b) The image warped to remove the radial distortion. Note that the lines in the periphery of the image are now straight, but that the boundary of the image is curved.*

Observation

A theorem at the heart of nature: A Numerical Experiment on the Central Limit Theorem

```
close all; clear;
```




```
A(2,:))/2,'b'),plot(...
(A(1,:)+A(2,:)+A(3,:)+A(4,:))...
/4,'g')
pause
%use build in MATLAB function
plot(mean(A))
pause
```

MATLAB

```
%divide range into 50 equidistant
%classes and count the members
A=rand(12,1000);
plot(hist(mean(A),50))
pause
%make more precise by higher sample count
hold on
A=rand(12,10000);...
        plot(hist(mean(A),50));pause
A=rand(12,100000);...
        plot(hist(mean(A),50));pause
A=rand(12,1000000);...
        plot(hist(mean(A),50));pause
%compare to a very special functionm
t=1:0.1:50;
plot(t,6.8e4*exp(-(t-26).*(t-26)./
(2*6*6)),'g')
```

1) Noise is not completely arbitrary, rather it contains some hidden regularity i.e. the Gaussian bell shaped function
2) This is the actual reason why repetitions of measurements will eventually converge to the true values!
3) The math behind it: central limit theorem

2D: $\bar{x}_{an}$ is Normally distributed

By central limit theorem:

If we repeat measurements with same instrument very often (I.I.D.) AND sum and build averages, then we are arbitrary close to a normal distribution

$$\lim_{n\to\infty} \bar{x}_{an} = \frac{X_1+\ldots+X_n}{n} \to \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Example

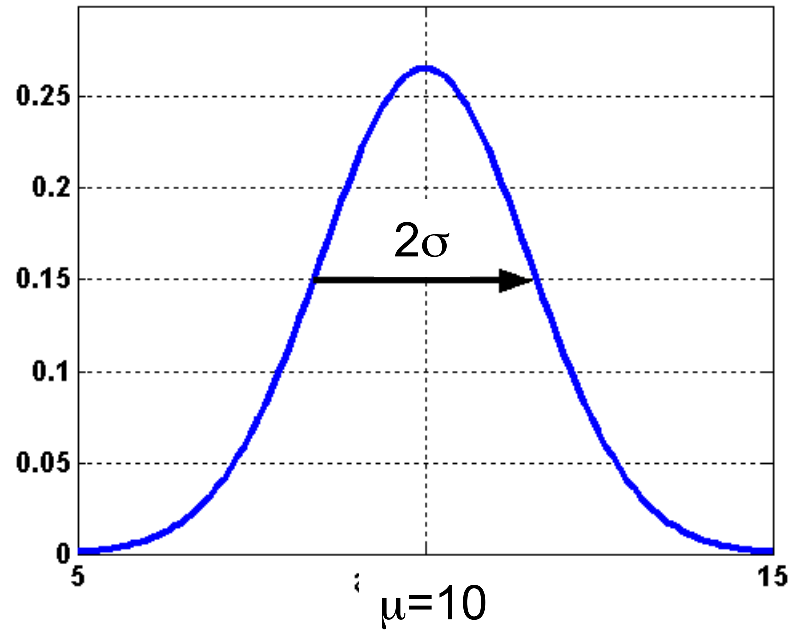
2D: Normal (Gaussian) distr.

$$f(t) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\frac{(t-\mu)^2}{\sigma^2}}$$

abbreviated to

$$x \sim N(\mu, \sigma^2)$$

$$F(x) = \int_{-\infty}^{x} f(t)\,dt$$



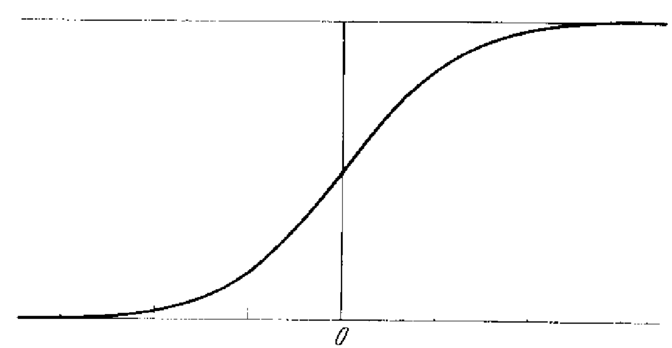


Fig. 2. Gaussian error function

2D: Examples of 2 Gaussians

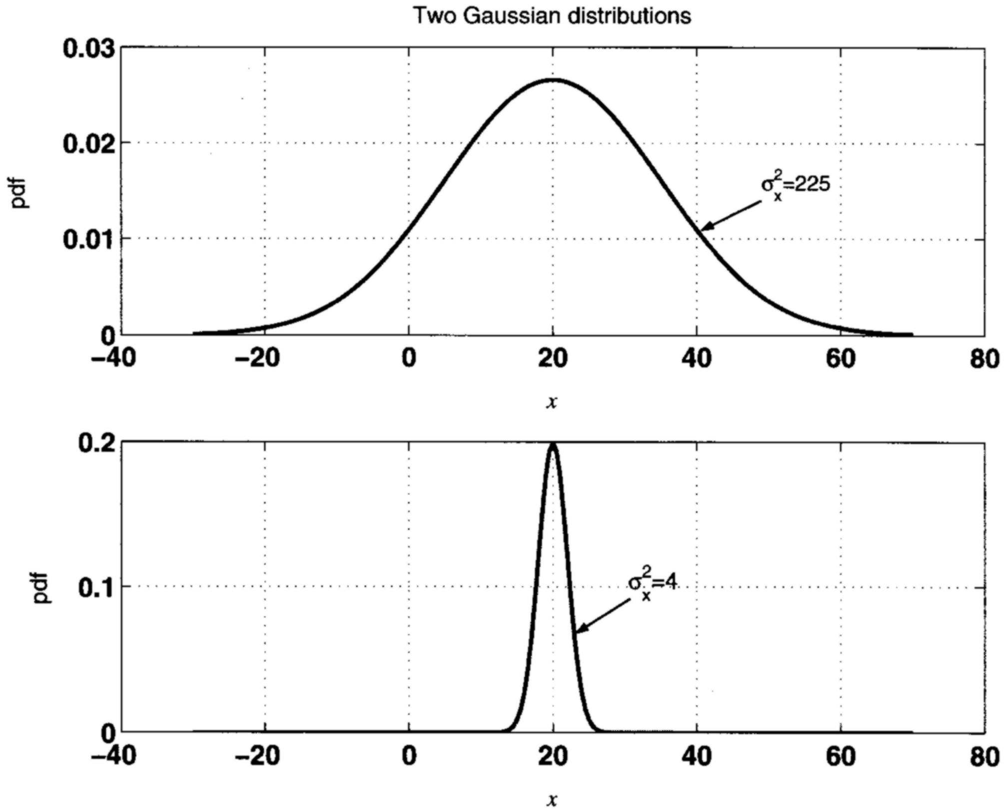
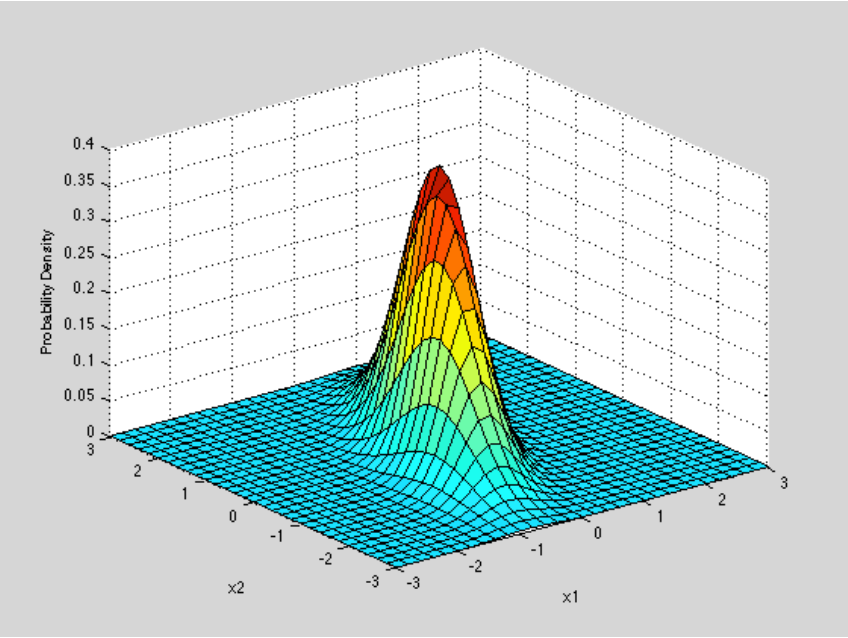


Figure 1.1. The figure shows the plots of the probability density functions of a Gaussian random variable $\boldsymbol{x}$ with mean $\bar{x} = 20$, variance $\sigma_x^2 = 225$ in the top plot, and variance $\sigma_x^2 = 4$ in the bottom plot.

2D: higher dimensional N(μ,σ)

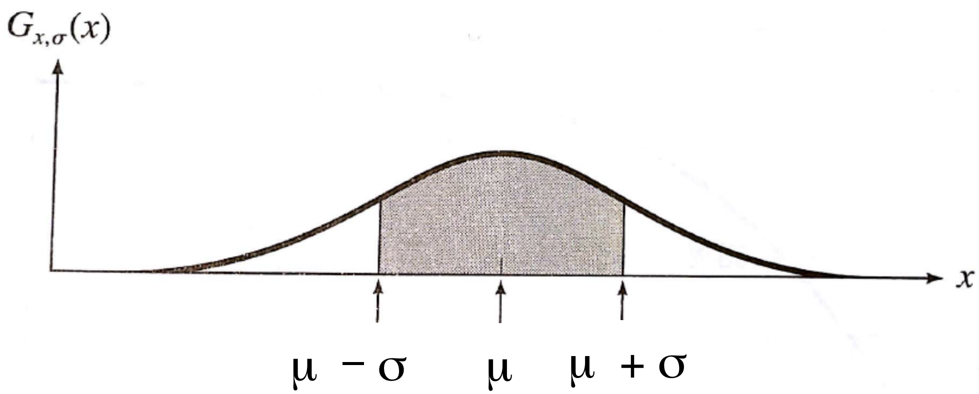
```
mu = [0 0];
Sigma = [.25 .3; .3 1];
x1 = -3:.2:3; x2 = -3:.2:3;
[X1,X2] = meshgrid(x1,x2);
F = mvnpdf([X1(:) X2(:)],mu,Sigma);
F = reshape(F,length(x2),length(x1));
surf(x1,x2,F);
caxis([min(F(:))-.5*range(F(:)),max(F(:))]);
axis([-3 3 -3 3 0 .4])
xlabel('x1'); ylabel('x2');
zlabel('Probability Density');
```



Example

UseC1: If N(μ,σ) is known ...

…. we then can apply a statistical test, to determine how likely a given measurement value x_i may actually be an outlier !



The shaded area between $\mu \pm \sigma$ is the probability of a measurement within one standard deviation of μ

E.g: if we want to be sure at level e.g. p = 95% that x_i is an inlier, what is the bound for x_i?

UseC1: If N(μ,σ) is known ...

Define the confidence level of x_i: $P(|x_i-\mu| < c)$

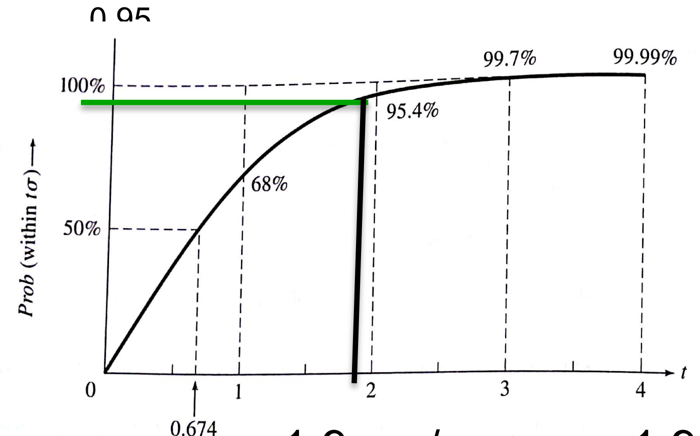
Read this as the likelihood, that distance from x_i to μ is less then c.

=> shaded area under Gauss distribution(integral)

Method: remap the given likelihood (0.95) back to multiples of σ,

here: 0.95 ≃≃ 1.9

Thus: an x_i which is farer away from μ then 1.9σ , then this x_i an outlier in 95% of all cases.



$1{,}9 = c/\sigma \Rightarrow c = 1.9\sigma$

t	0	0.25	0.5	0.75	1.0	1.25	1.5	1.75	2.0	2.5	3.0	3.5	4.0
Prob (%)	0	20	38	55	68	79	87	92	95.4	98.8	99.7	99.95	99.99

Figure 5.13. The probability *Prob*(within *tσ*) that a measurement of *x* will fall within *t* standard deviations of the true value *x* = *X*. Two common names for this function are the *normal error integral* and the *error function*, erf(*t*).

Example

UseC2: testing for N(μ,σ)

Chi square test:

a) From n sample values (from an underlying totality following some hidden distribution D) build estimators $\bar{x}_n$ and S_x !

b) Subdivide all n samples into K classes ($K >= 4$) s.t. in each class there are at least 5 samples (width of class may vary if necessary).

c) Get n_{ei}, i.e. the observed number of samples per class i.

d) Build $N(\bar{x}_n, S_x)$ and P_i, i.e. the likelihood that a sample lies in class i, then build $n_{oi}=P_i * n$, i.e. the number of to-be-expected samples in class i if the totality would be distributed according to $N(\bar{x}_n, S_x)$.

e) Build χ^2 and $n_f = K-1$.

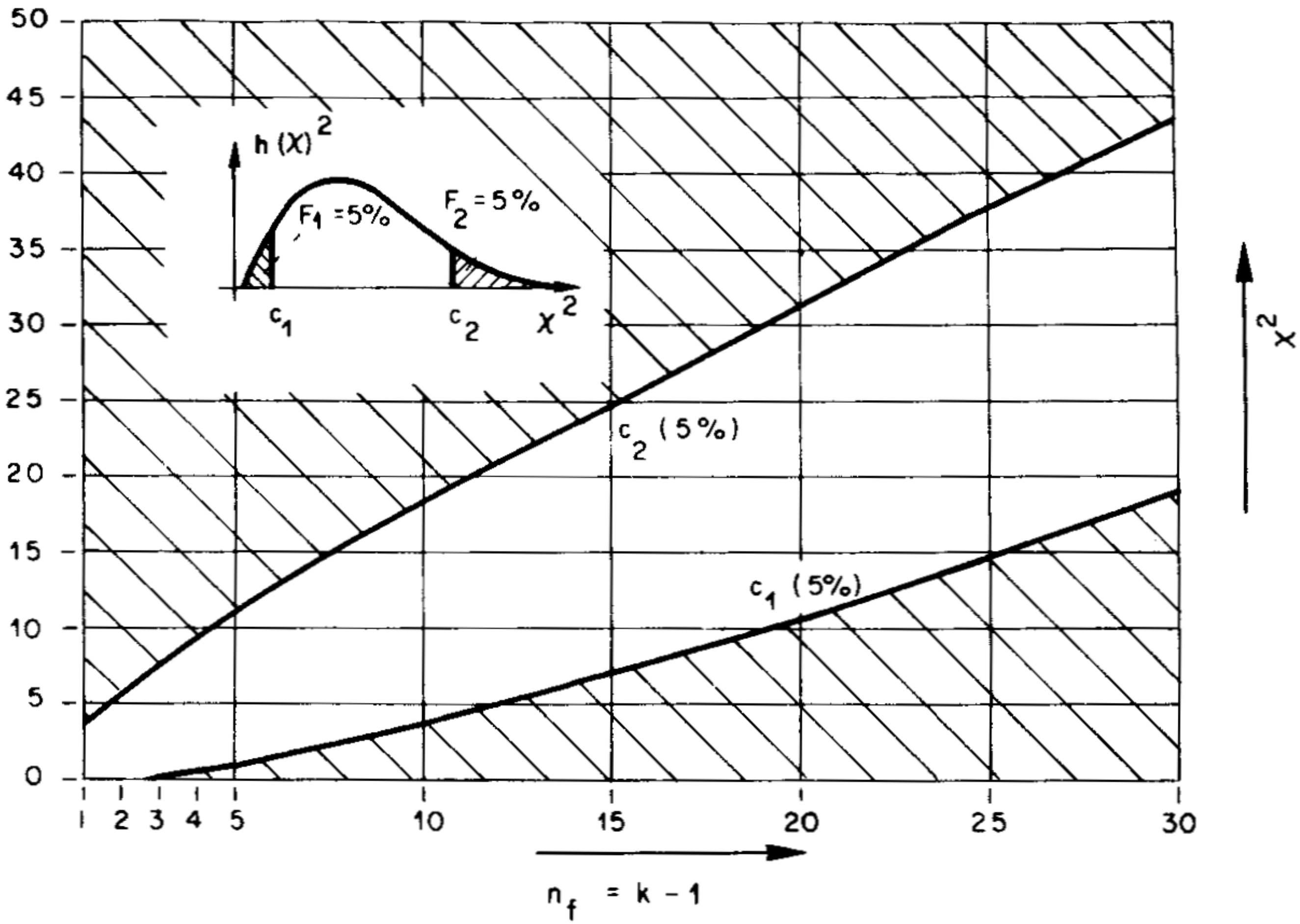
$$\bar{x}_n = \frac{1}{n}\sum_{i=1}^{n} x_i \quad S_x = \sqrt{\frac{1}{n-1}\sum_i (x_i - \bar{x}_n)^2} \quad \chi^2 = \sum_{i=1}^{K} \frac{(n_{ei}-n_{oi})^2}{n_{oi}}$$

UseC2: testing for N(μ,σ)

If point (χ^2,n_f) lies in non-hashed area, there is no indication, that D would NOT be normal (up to a confidence level of 5% below and above)

NB:

This test works for ANY distribution! E.g. uniform etc



3 Summary

- Measurements suffer from many effects (**random** and/or **systematic**).
- **Systematic errors** may be cured by **calibration**.
- **Random errors** are not totally random, there is distribution underneath (**mostly: Gaussian**, also in higher dimensions)
- **Sampling and averaging** helps to read „richtige“ **values x_r** (empirical mean and spread)
- **Distributions** may be checked versus model distributions **using χ^2 test.**

Example and Homework

The following is a recent published example for the description of an experimental setup.

How detailed is the example decribed, can we rerun the experiment with equal results?

What is missing (if so...) ?

Apply the **Formalized General terms** to this setup!

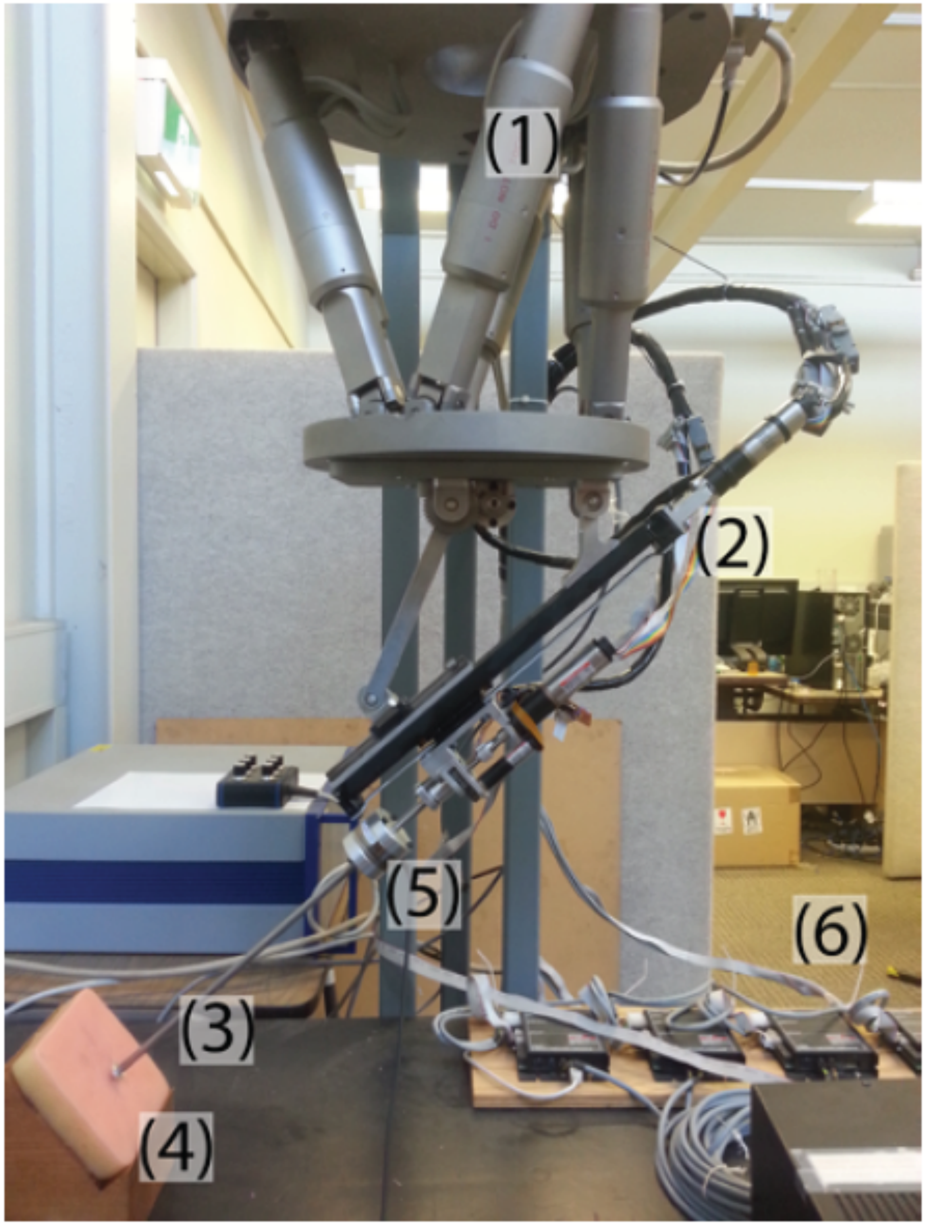


Fig. 2. The experimental research facility for measuring tissue forces. This facility combines the following components: (1) PI Hexapod, (2) monocarrier drive, (3) indentation tool, (4) Simulab complex tissue model, (5) ATI Mini40 force sensor and (6) EPOS2 motor controllers.

Title: Modelling the Indentation Force Response of Non-Uniform Soft Tissue using a Recurrent Neural Network

Rohan Nowell1, Bijan Shirinzadeh2, Julian Smith3 and Yongmin Zhong4

II. Experimental Facility

The research facility used to generate the experimental reference and test data is shown in figure 2. The facility contains a 10 DOF robotic system designed and developed for laparoscopic surgery research [24] consisting of a Physik Instrumente Hexapod (1) with an attached monocarrier drive (2). The Hexapod is used to move the indenter (3) over a grid of locations on the tissue analogue (4) while the monocarrier is used to provide the motion in the indentation direction. The indenter contains a force torque (FT) sensor (5) which provides axial force measurements.

The FT sensor is an ATI Mini40 which has a specified unfiltered axial force resolution of 0.04 N. The force measurements are read using a NI DAQ which operates at a rate of 1 kHz. Software filtering is then performed to reduce the high frequency noise and sensor drift is mitigated by periodically zeroing the FT sensor between indentation test locations. The monocarrier is controlled using an EPOS2 motor controller (6) which provides motor position and velocity readings every 4 ms.

The tissue analogue is a Simulab complex tissue model. Damage to the tissue sample from penetration testing has caused it to have a very non-uniform force response over its surface, making it ideal for testing the soft tissue models.

III. Experimental Procedure

The force response data to be modelled was taken by pressing the indenter (6.5 mm hemispherical tip) into the tissue model in the surface normal direction up to a depth of 11.5 mm (38% of the sample depth).

The process for taking the experimental data was as follows: The indenter was moved towards the first grid location above the sample, then an indentation depth is chosen between 0 and 11.5 mm. The indenter is moved to this indentation depth before another depth is randomly chosen. 800 random depth indentations were performed in succession at each of the 25 grid locations on the tissue sample with no material preconditioning. This was done to ensure that a wide array of conditions exist in the training and test data for the neural network models to develop a model which can generalise well. The axial force and the indentation depth were recorded at approximately 4 millisecond intervals. The measured force is negative for compression in the force/torque (FT) sensor's co-ordinate system and this convention was preserved when modelling.

IV. RESULTS AND DISCUSSIONS

A. *Experimental Results*

Figure 3 shows the results of one of the indentation tests. It can be seen that the indentation depth alone is insufficient to accurately model the force observed. The amplitude of the noise in the FT sensor was analysed to determine a lower limit for model performance, since neural networks are often prone to overfitting noise in training data. Figure 4 shows the filtered force response from the FT sensor in a 5 minute period of no load with an approximate sample period of 4 ms.

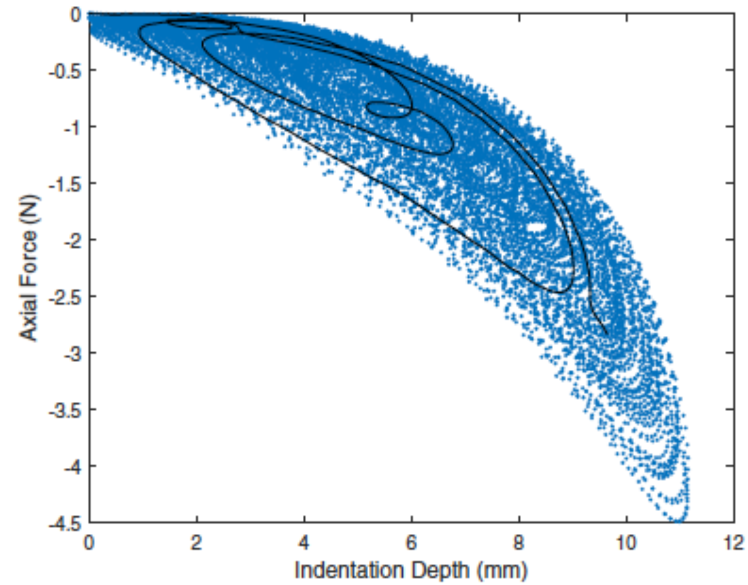


Fig. 3. Experimental force-indentation data showing the wide spread of forces at each indentation depth. The black line shows the results of the first 10 indentations.

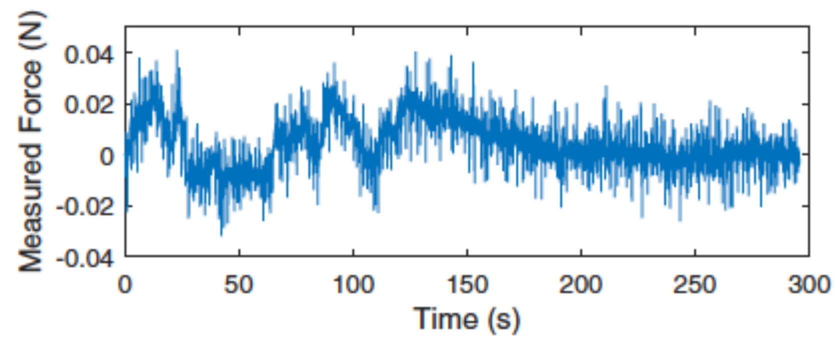


Fig. 4. Force reading from the ATI Force sensor under no external loading showing the extent of noise.

The amount of high frequency noise in the filtered measurements was calculated by taking a moving average with a 1001 sample window as a baseline and then taking the standard deviation of the difference. The standard deviation of the noise is 0.0042 N. This is important for neural network training as it sets a lower limit for acceptable error. If a neural network modelled training data with a root mean squared error below this it was considered overfitted, and training was halted. The maximum peak to peak values over this test was 0.0987 N.

Experimental results verify the non-uniformity of the tissue sample. The maximum measured axial force for 25 different locations on the sample is shown in figure 5. Because of the non linear force response of the tissue, a linear elasticity based parametrisation is not suitable. Developing a neural network model which can account for the indentation location would require the network to be trained on a case by case basis, using large sets of training data. As a compromise, a hybrid model of a neural network and the inverse strain model was developed. This model is described in section IV-C.

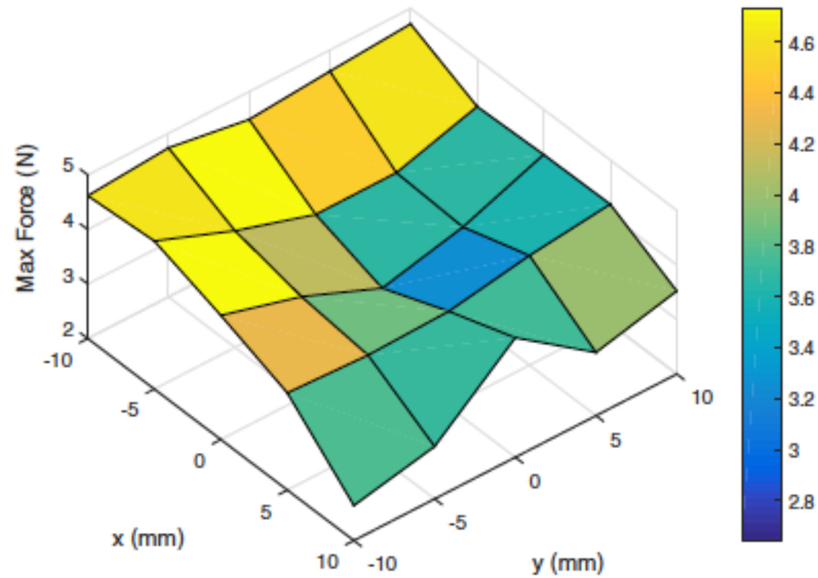
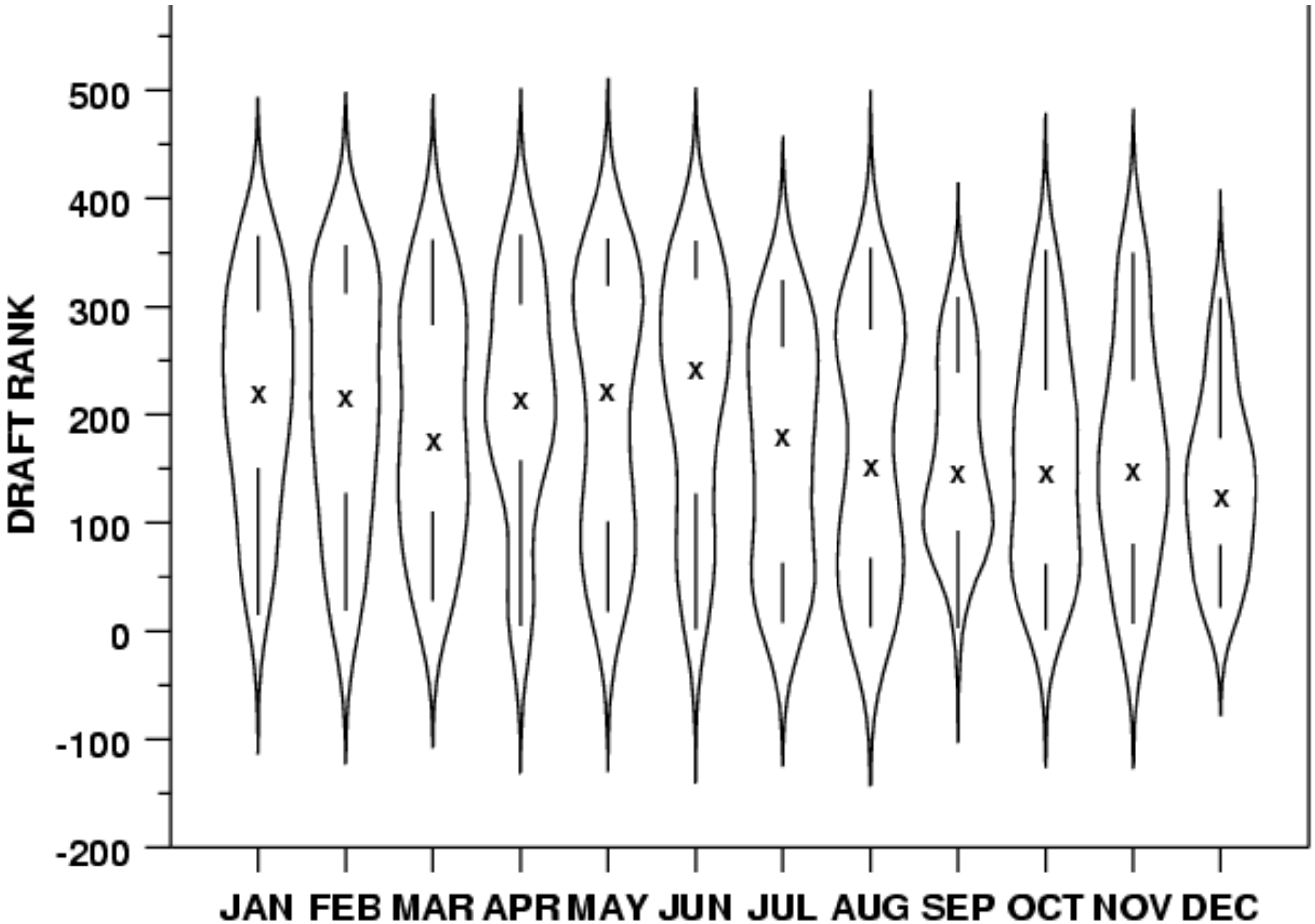


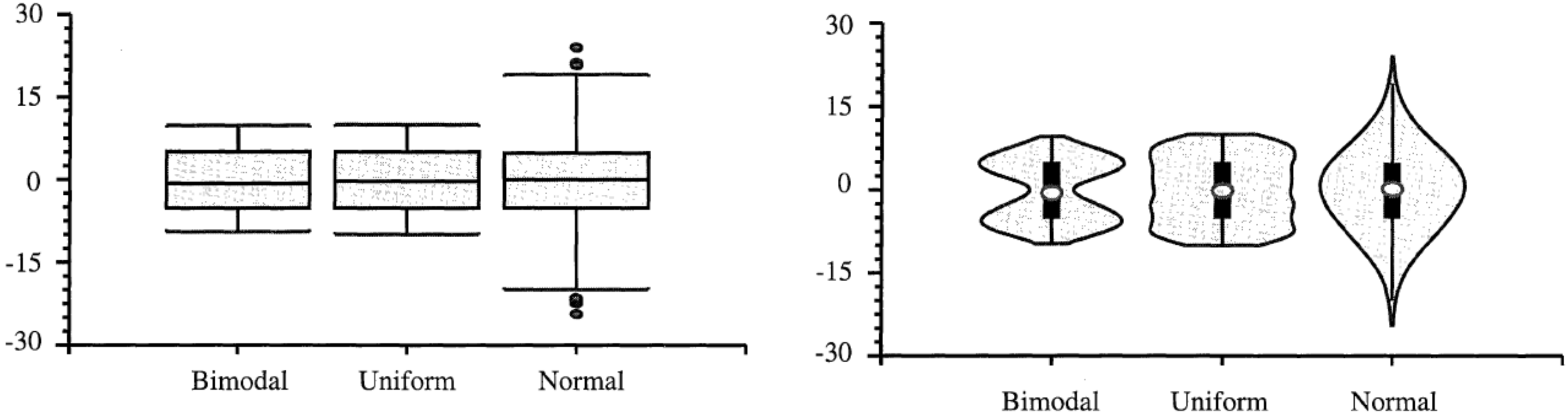
Fig. 5. Maximum force response over a 5 $\times$ 5 mm grid on the surface of the tissue analogue. The indentation test locations occur at the surface vertices.

Violin Plot



DRAFT RANK
500
400
300
200
100
0
-100
-200
JAN FEB MAR APR MAY JUN JUL AUG SEP OCT NOV DEC

Box plot versus Violin plot



(a) (b)