IEEE SW Test Workshop
Semiconductor Wafer Test Workshop
June 9 - 12, 2013 | San Diego, California

probing@hot temperature
a new approach to accuracy

BOSCH

Wafer Test Center
Reutlingen, Germany

Harald Berger

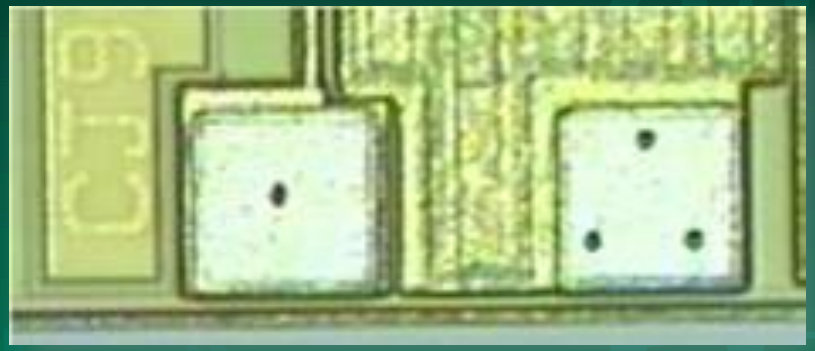
Walter Seitz

Robert Bosch GmbH

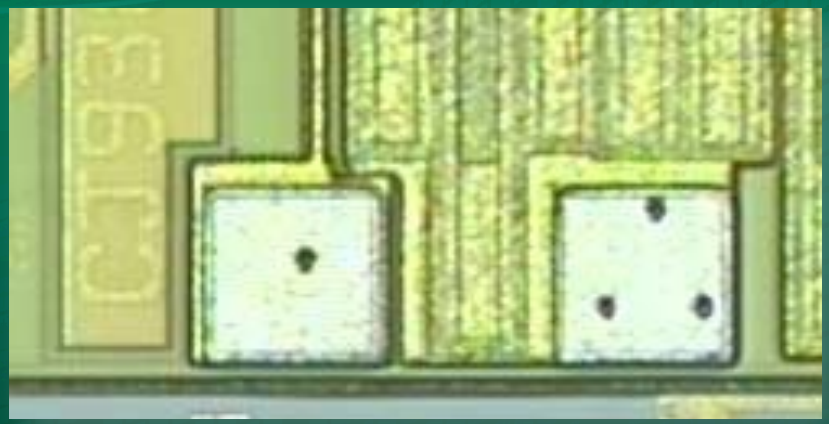
Thermal Expansion

$\Delta L/L=\alpha\ \Delta T$

Heat expansion cannot be defeated by any force



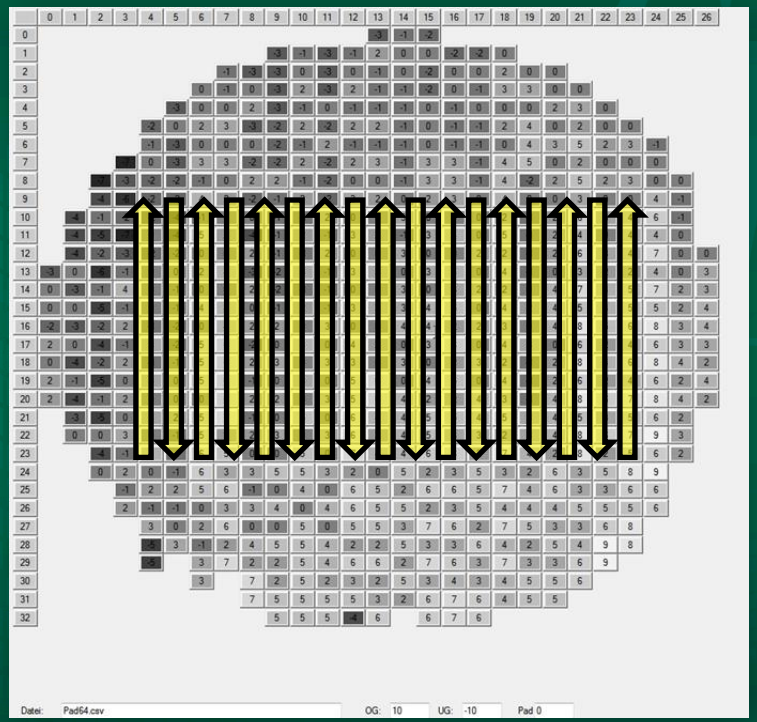
Scrub Mark Shift, caused by High Temperature EWS



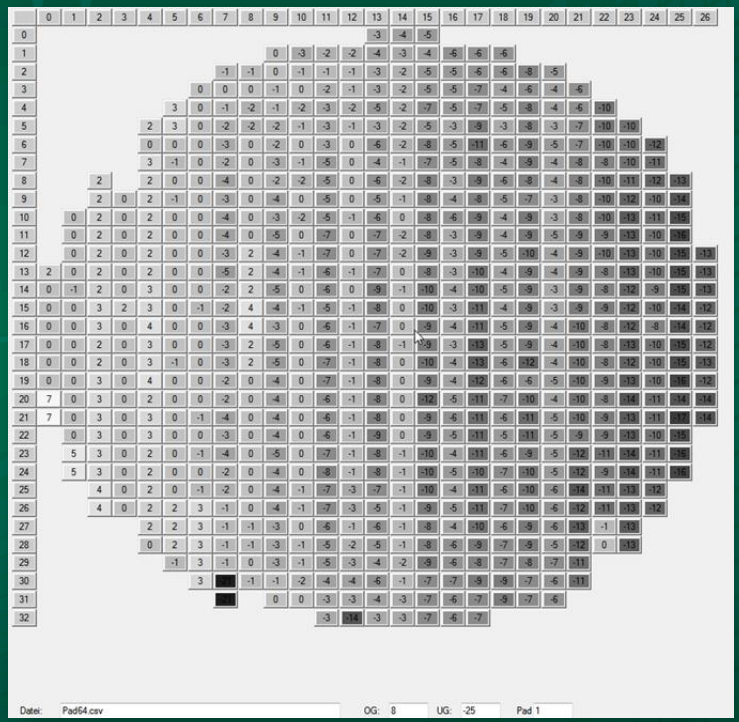
- **Standard Methods provided by Prober Vendor:**
  - Preheat / Re-Preheat
  - Realignment of Needle Tips and Wafer

- **Innovative Methods (mainly by Customers)**
  - Probecard Shielding / Heating / Cooling
  - Head plate Heating
  - Special kind of Stepping Patterns (Circle, Random, etc.)

Scrub Mark Postition
(WaferWoRx Result)

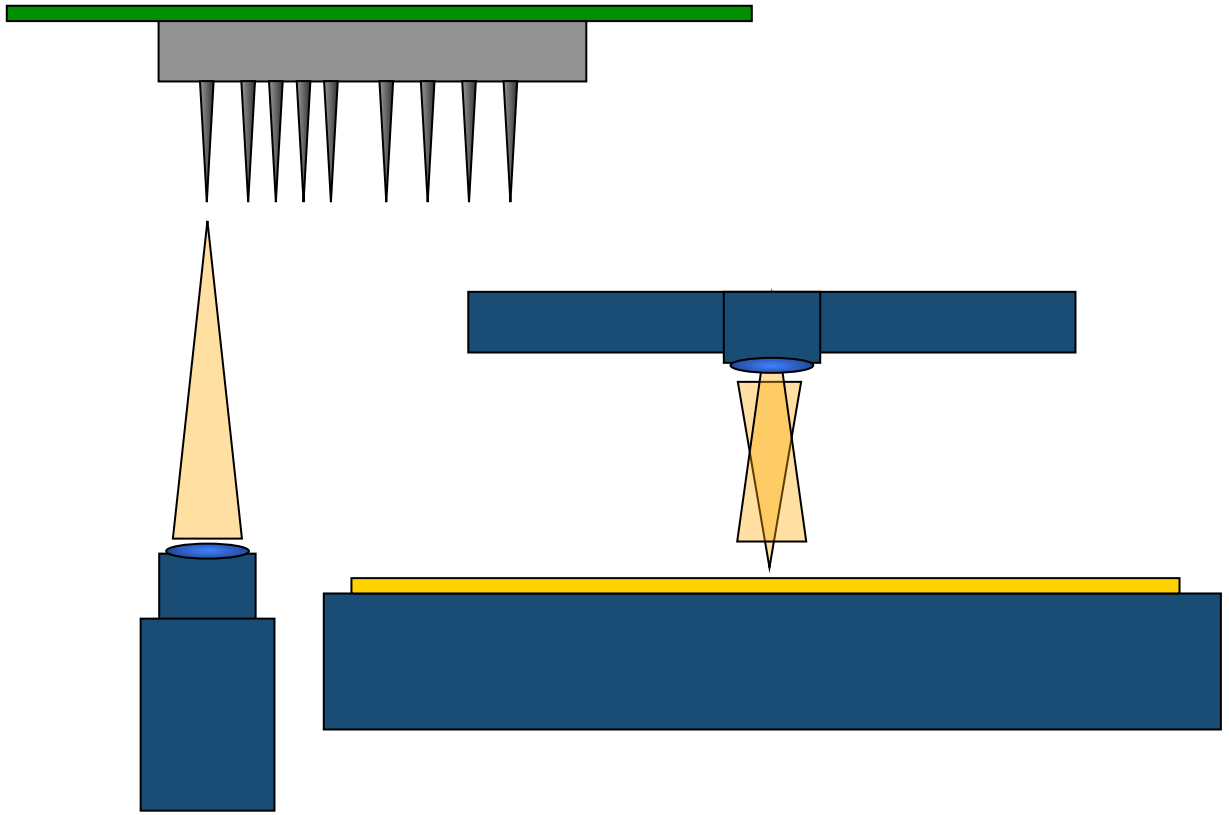


X-Direction

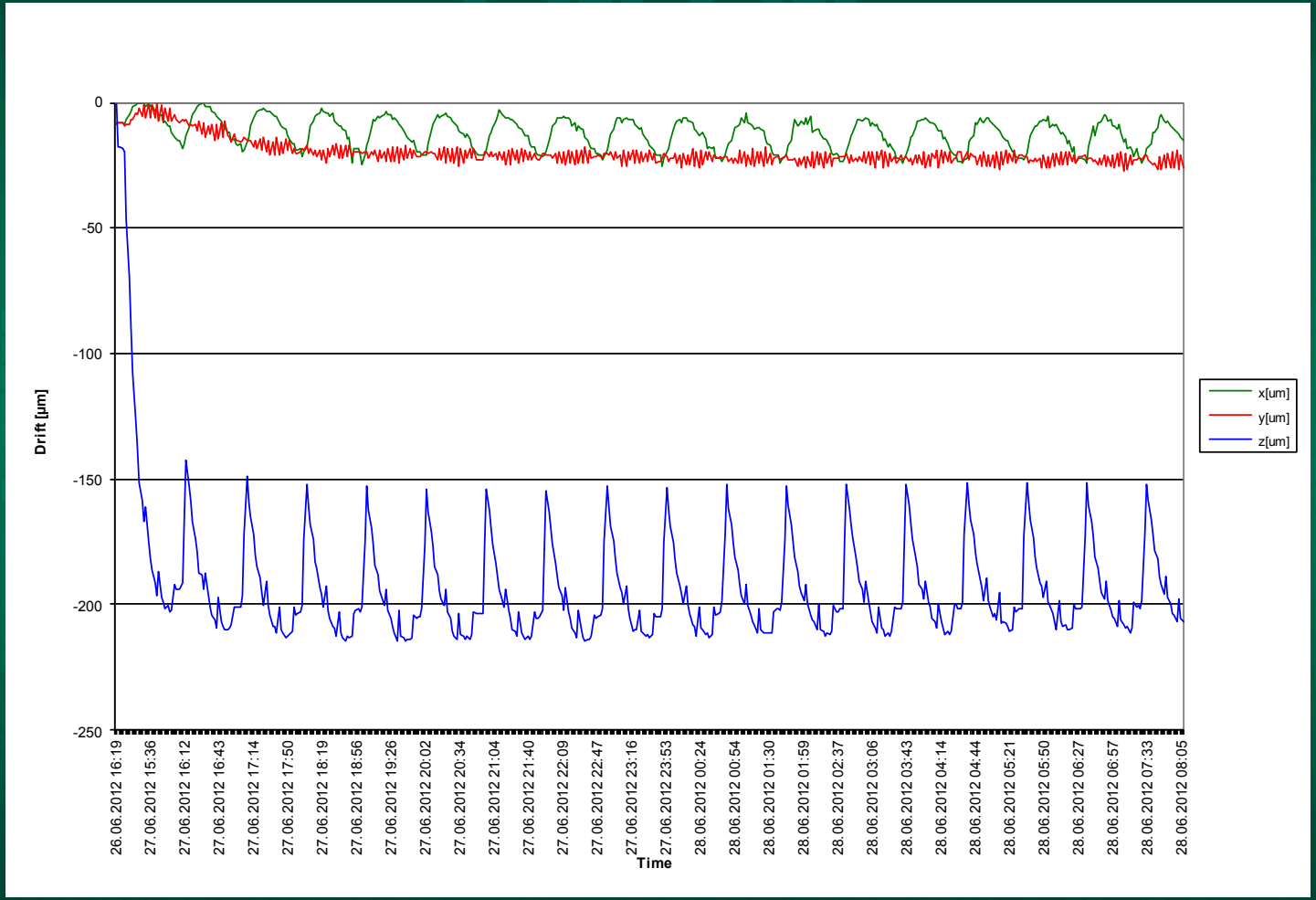


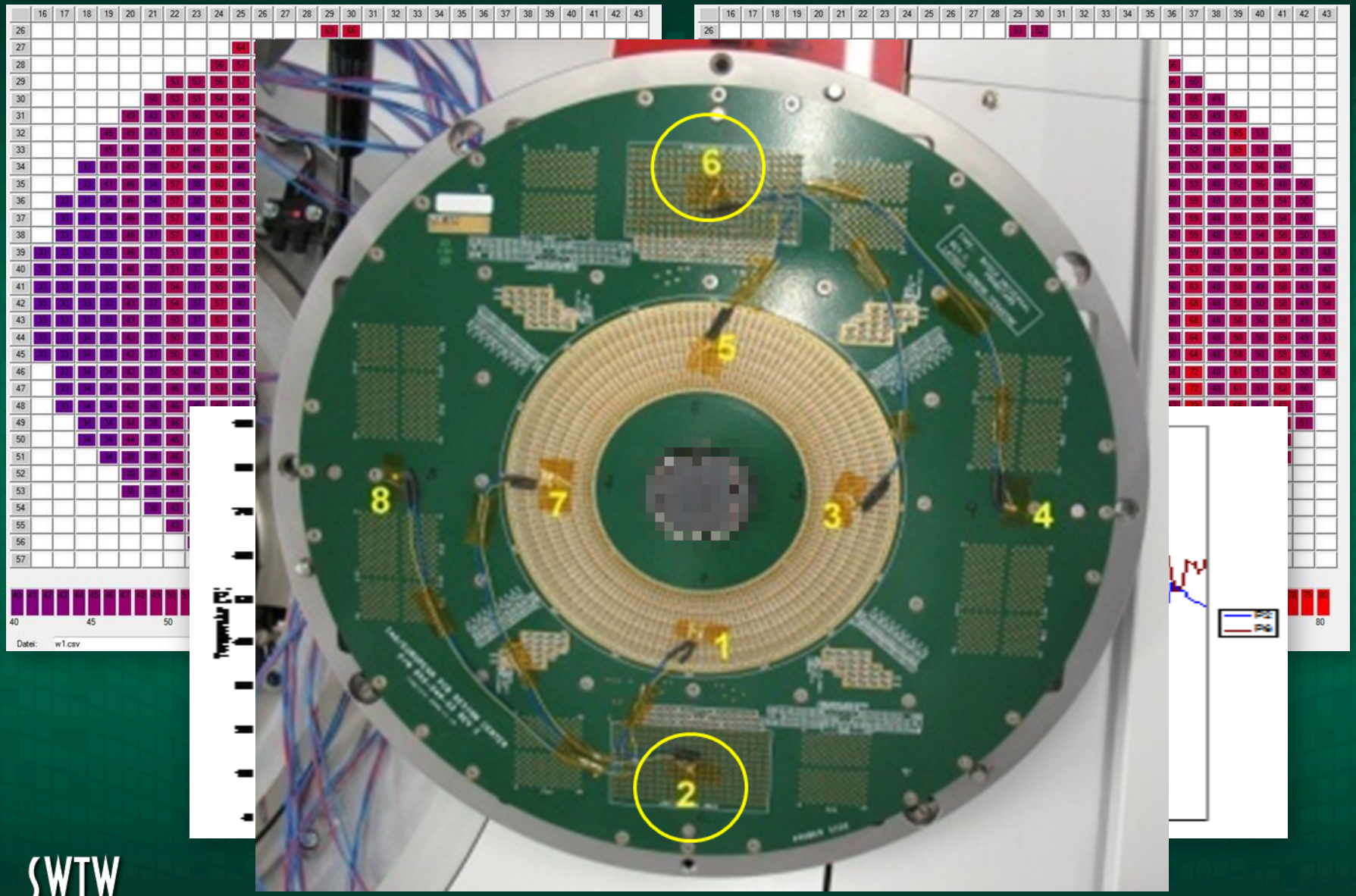
Y-Direction

Probe Centre X=108730,8 µm
Y=228750,0 µm
Z=67795,7 µm

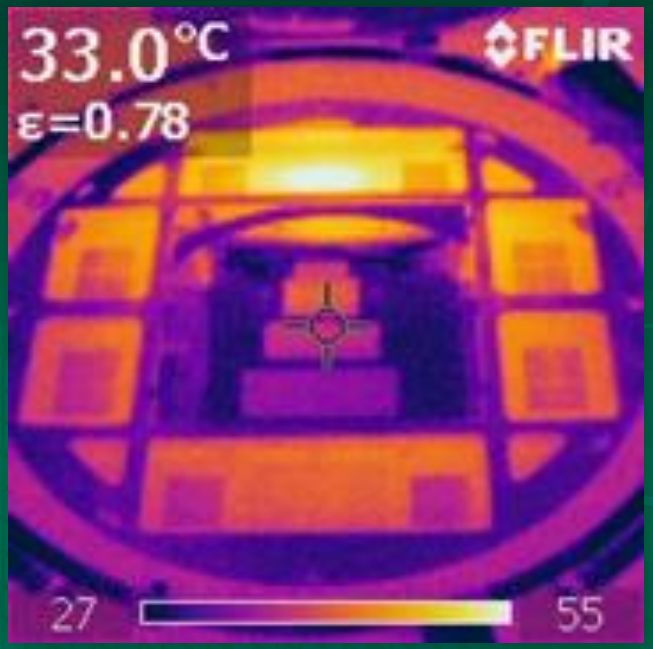


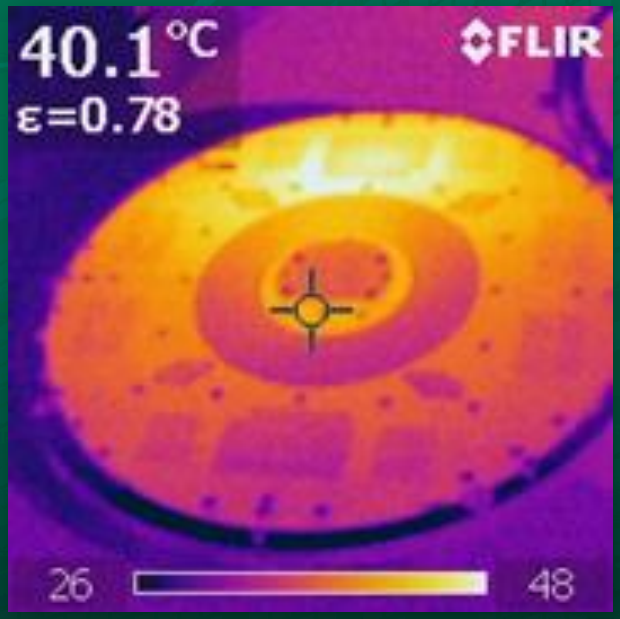
Realignment Data of one Lot



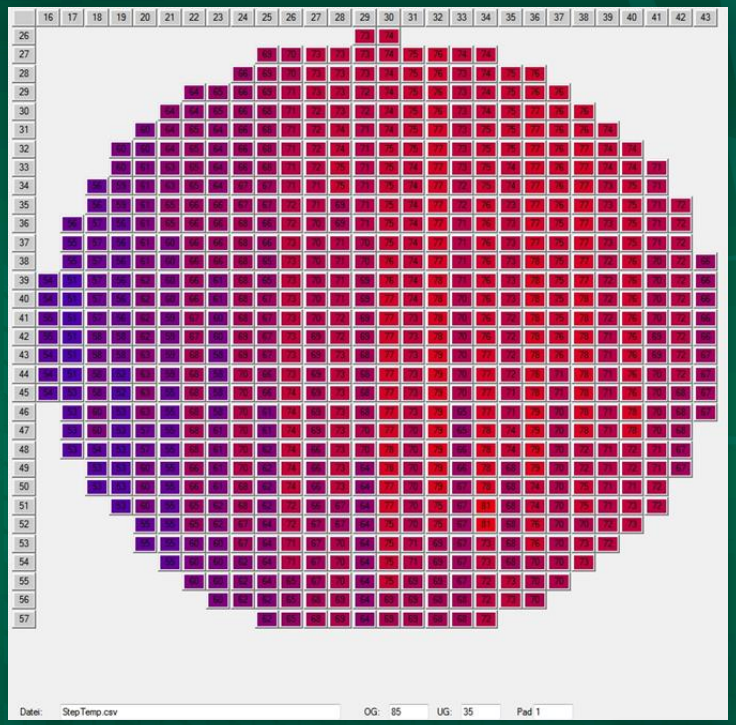

6
5
8
7
3
4
1
2

IR Camera Image

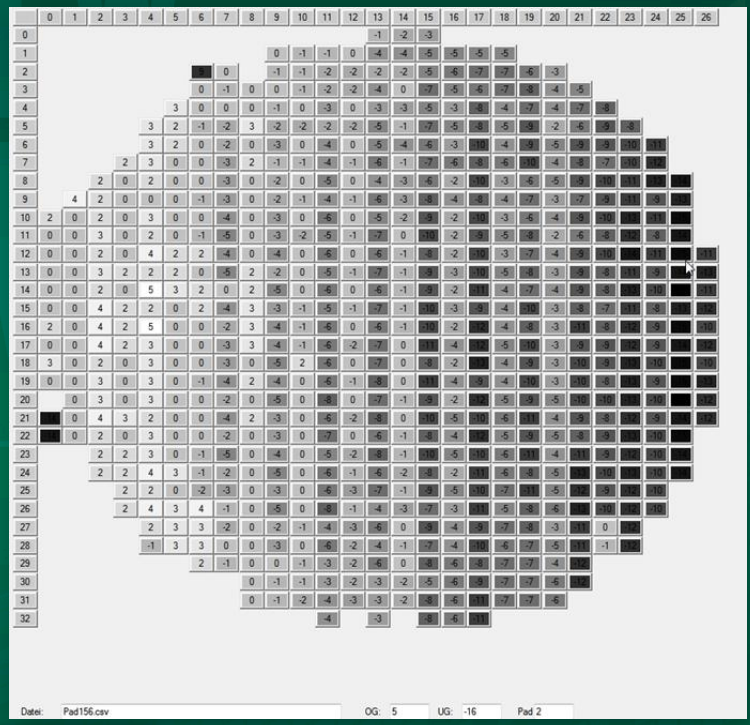


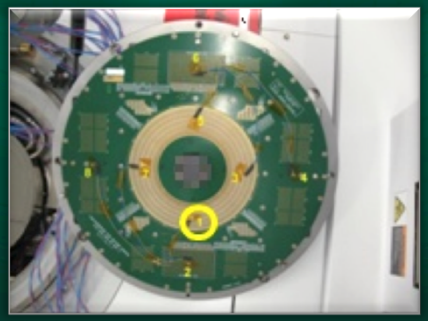


Temperature deviation while stepping

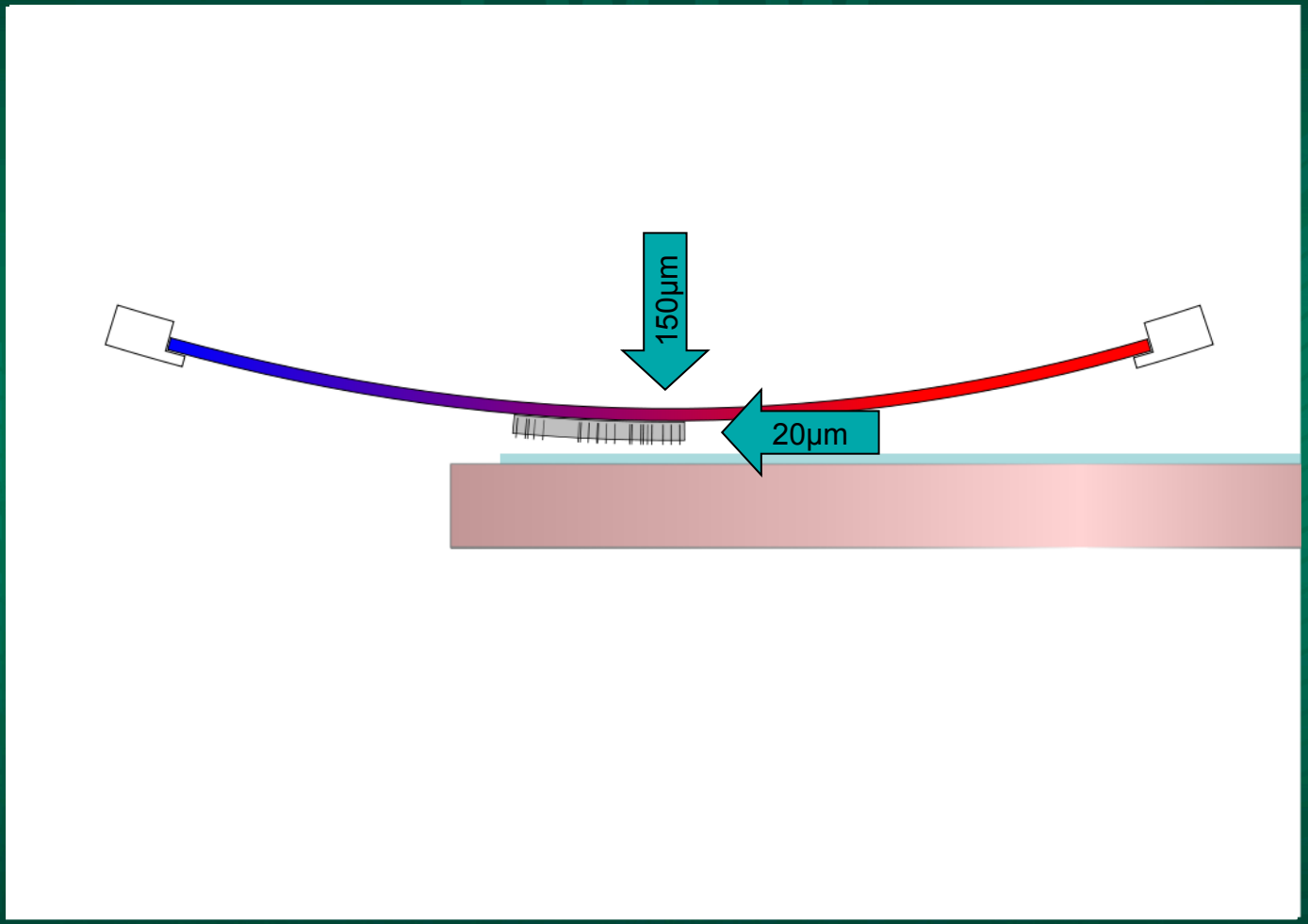


Scrub Mark Position



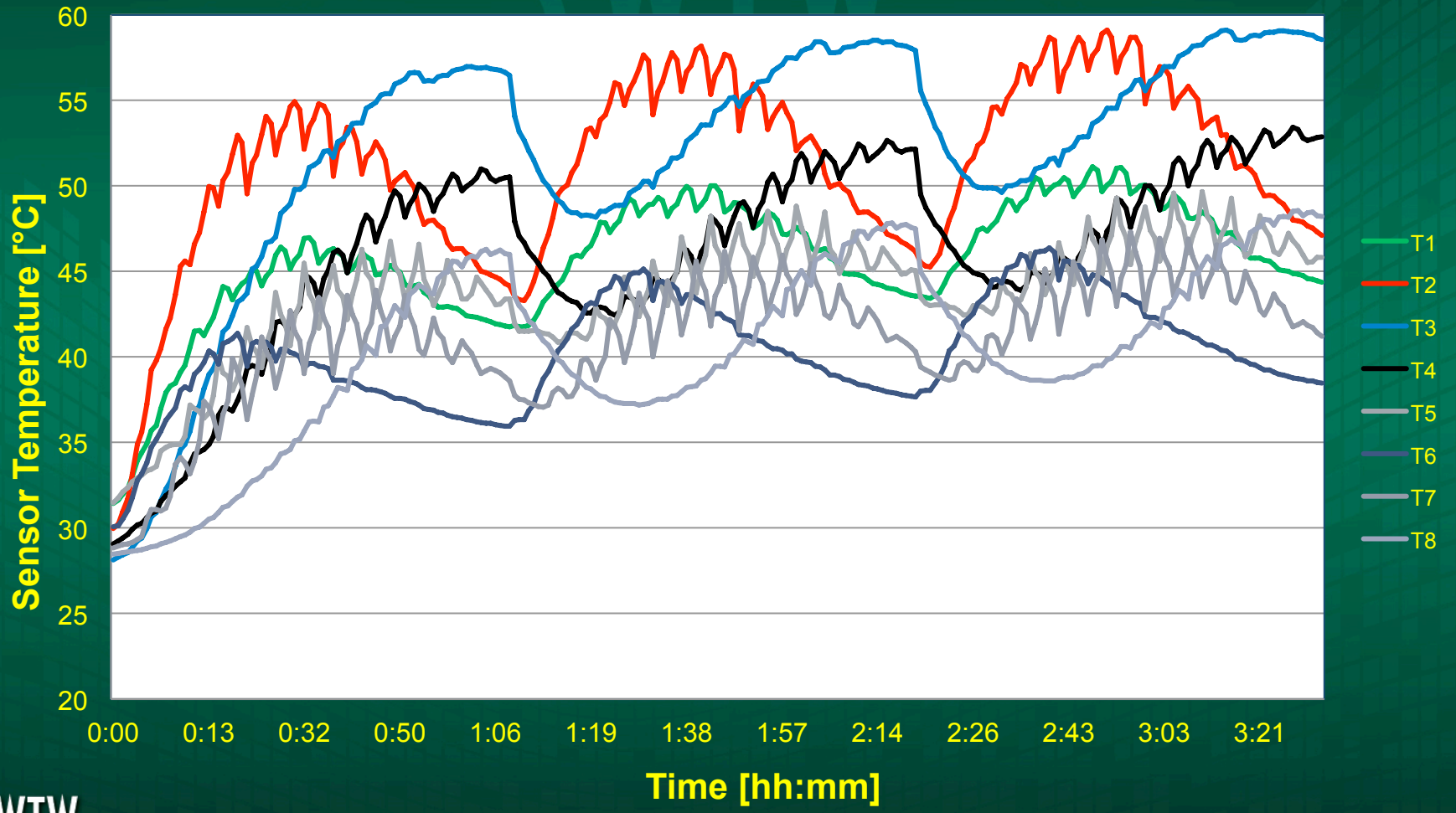


Thermal Expansion Model

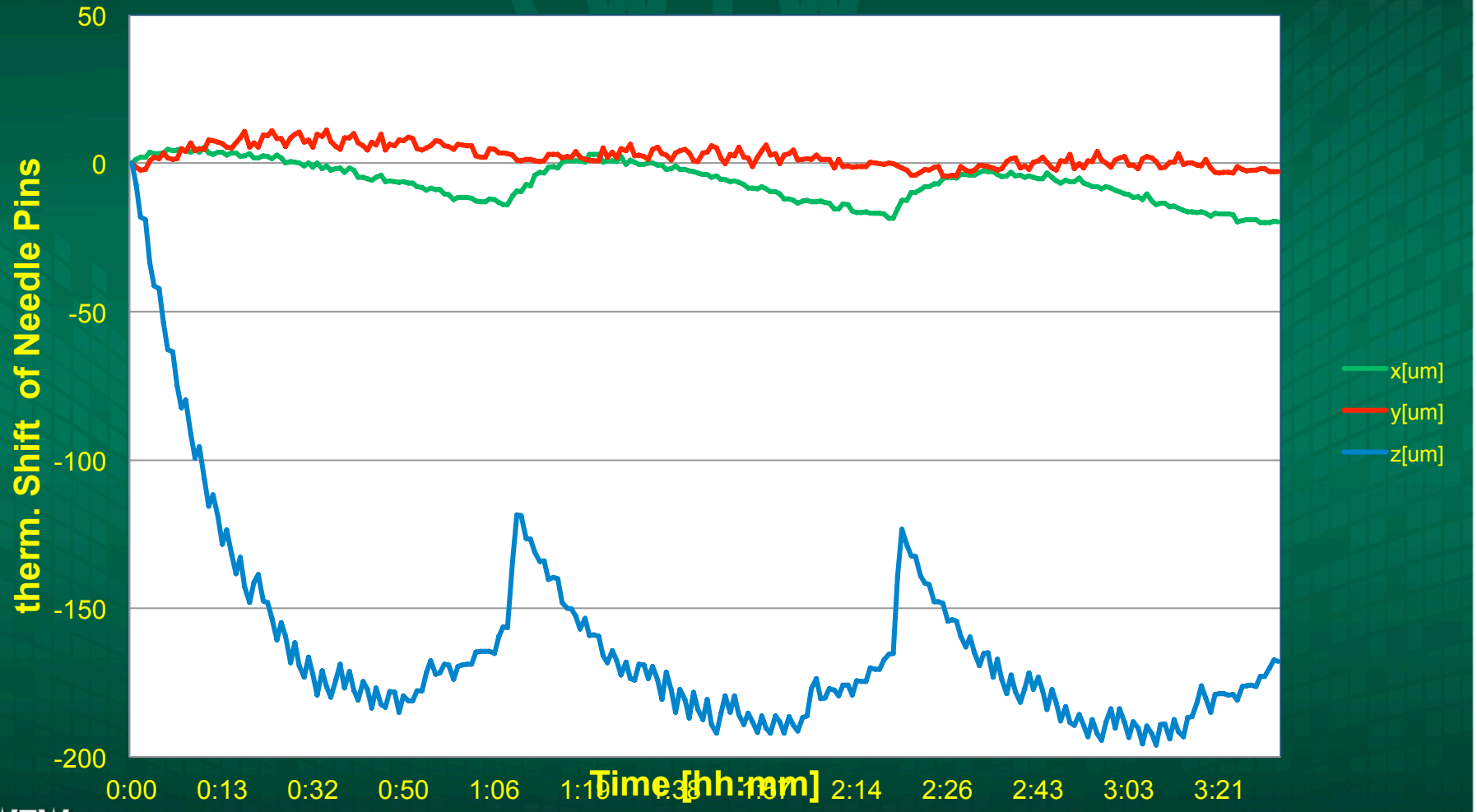


Sensor Position on Top of Probe Card

Data of 8 Sensors on Top of Probecard (3 Wafers)



Realignment Data of Probecard (3 Wafers)



facts:

probecard is heated by the chuck on different positions

heat generates expansion of the probecard

expansion of the probecard generates movement of x,y & z of the beams

idea:

anticipate the movement of the beams in x,y and z as a function of the sensor signals

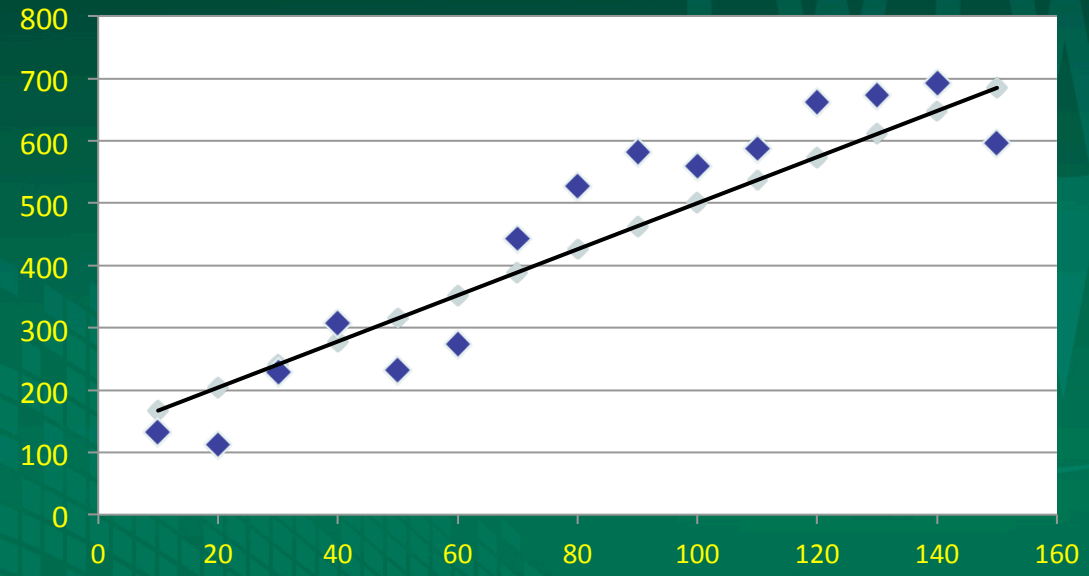
advantage:

real time correction of beam movement in x, y & z possible

challenges & questions

- Is there a deterministic correlation between the sensor signals and the movement of the beams ?
- What about unknown variables influencing the movement as well ?
- How reproducible are the results ?

let's start simple:



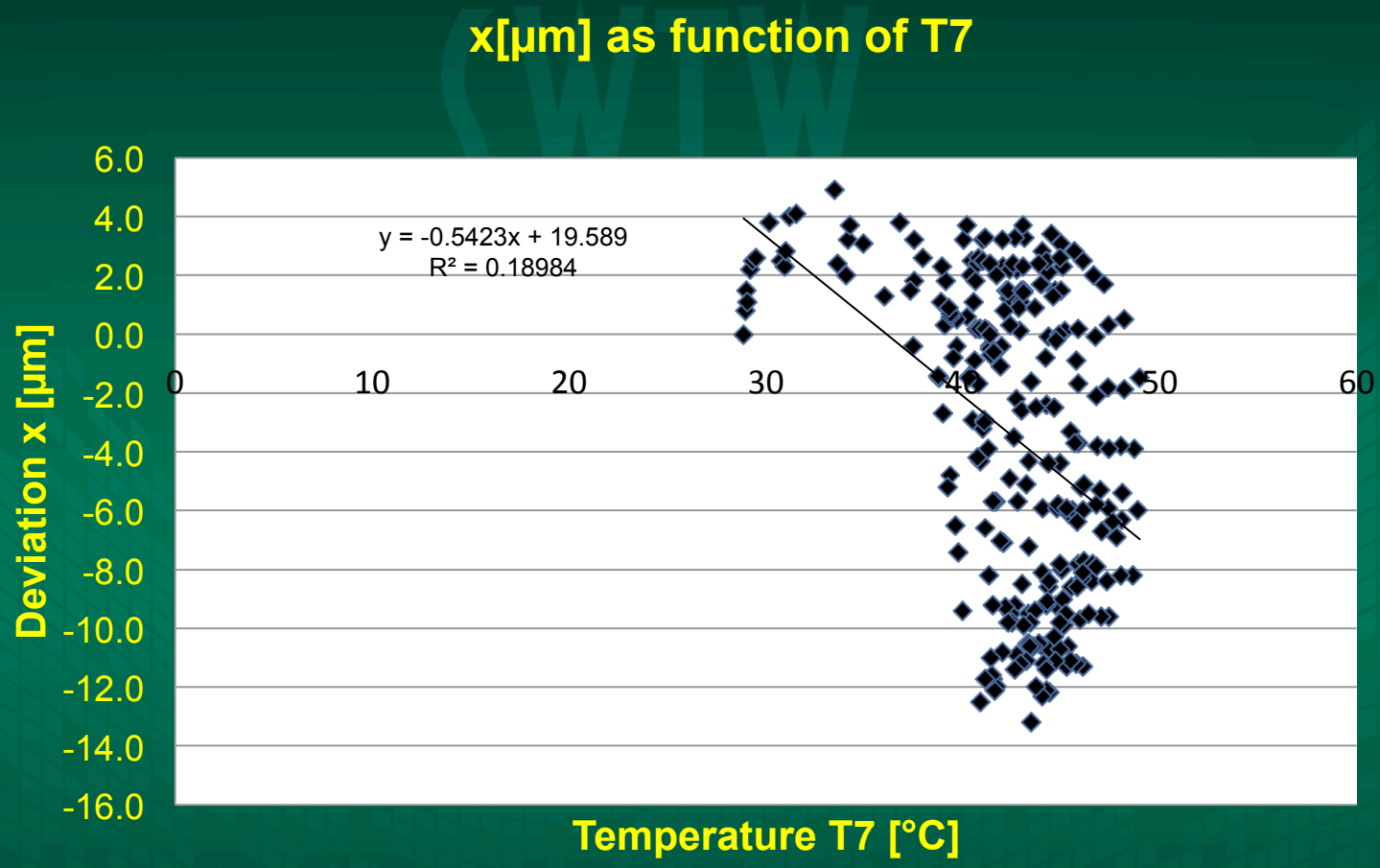
supposed correlation between y and x

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

challenges:

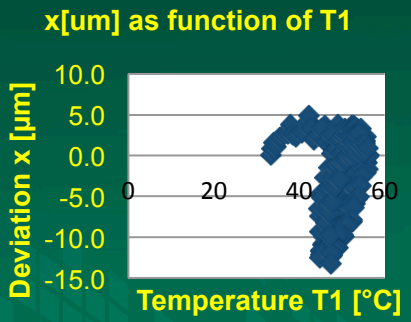
1. not solveable (more equations then variables)
2. what about higher correlations (Square,more then one variable…)

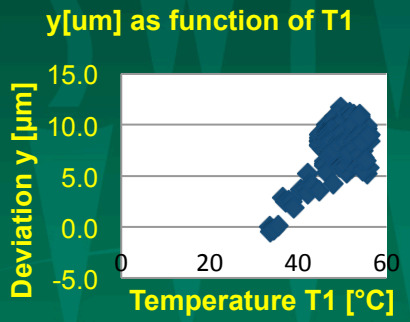
trying to apply it on our problem:

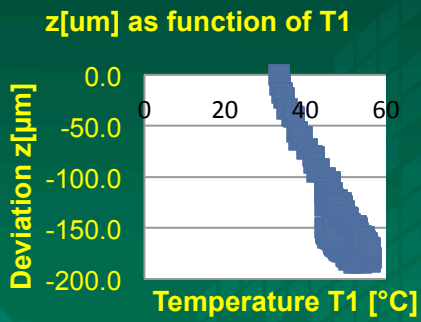


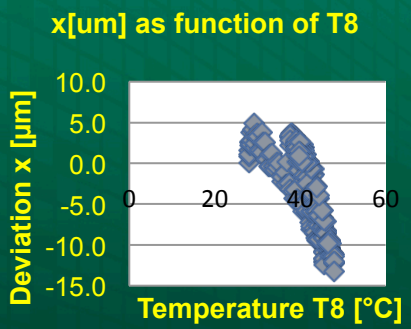
poor quality, no function for a given signal

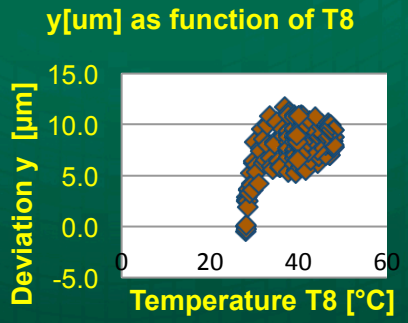
make it more complicate:
taking into account more variables

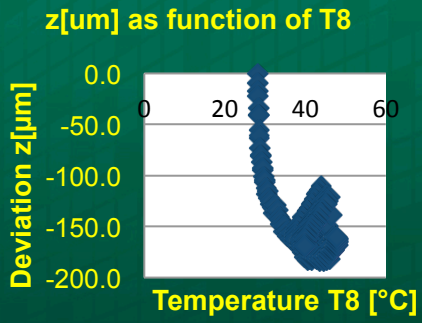












generalization:

multi – dimensional non linear correlation

$y = Ax$ (least square, Carl Friedrich Gauss, 1795)

with $A = \begin{pmatrix} 1 & x_1 & x_1^2 \\ \vdots & \ddots & \vdots \\ 1 & x_m & x_m^2 \end{pmatrix}$; $x = \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix}$ $y = \begin{pmatrix} y_1 \\ . \\ . \\ y_m \end{pmatrix}$

search of lowest discrepancy results in condition

$$\|Ax - y\|^2 = \min$$

$$\Leftrightarrow \frac{d}{dx_j}\left(\sum_{i=1}^{m}(\sum_{k=1}^{n} a_{ik}\,x_k - y_i)^2\right) = 0$$

$$= 2\sum_{i=1}^{m} a_{ij} \sum_{k=1}^{n}(a_{ik}\,x_k - y_i)$$

$$= 2\sum_{k=1}^{n} x_k \sum_{i=1}^{m} a_{ij}a_{ik} - 2\sum_{i=1}^{m} a_{ij}\,y_i$$

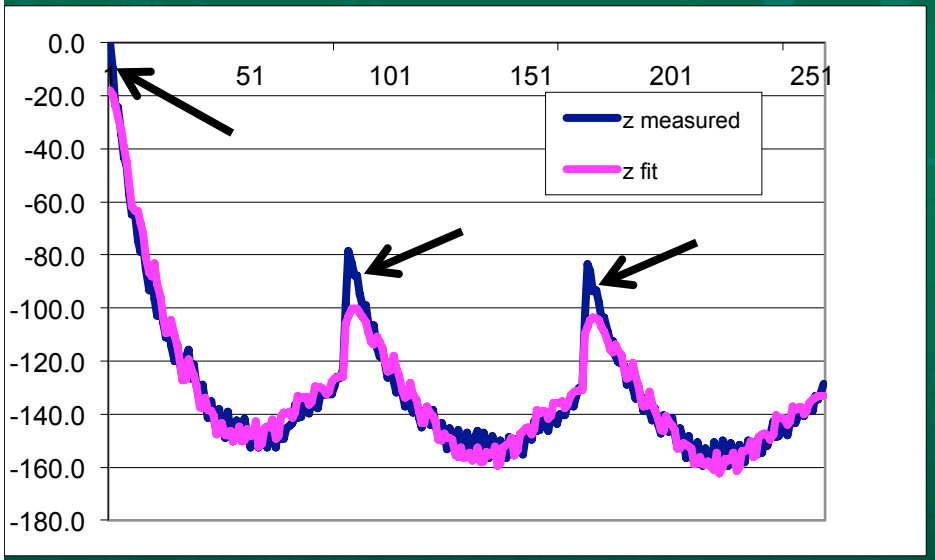
$$A^T Ax = A^T y$$

$$x = (A^T A)^{-1} A^T y$$

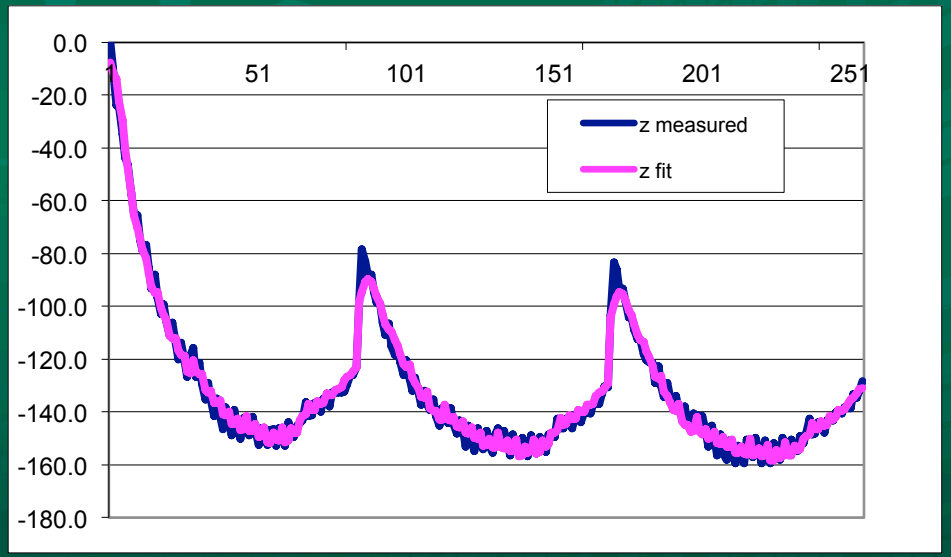
[Fischer, Kaul: „Mathematik für Physiker“, Teubner, 1988]

programming was done in Python

important question:
Which functional correlation has to be taken ?



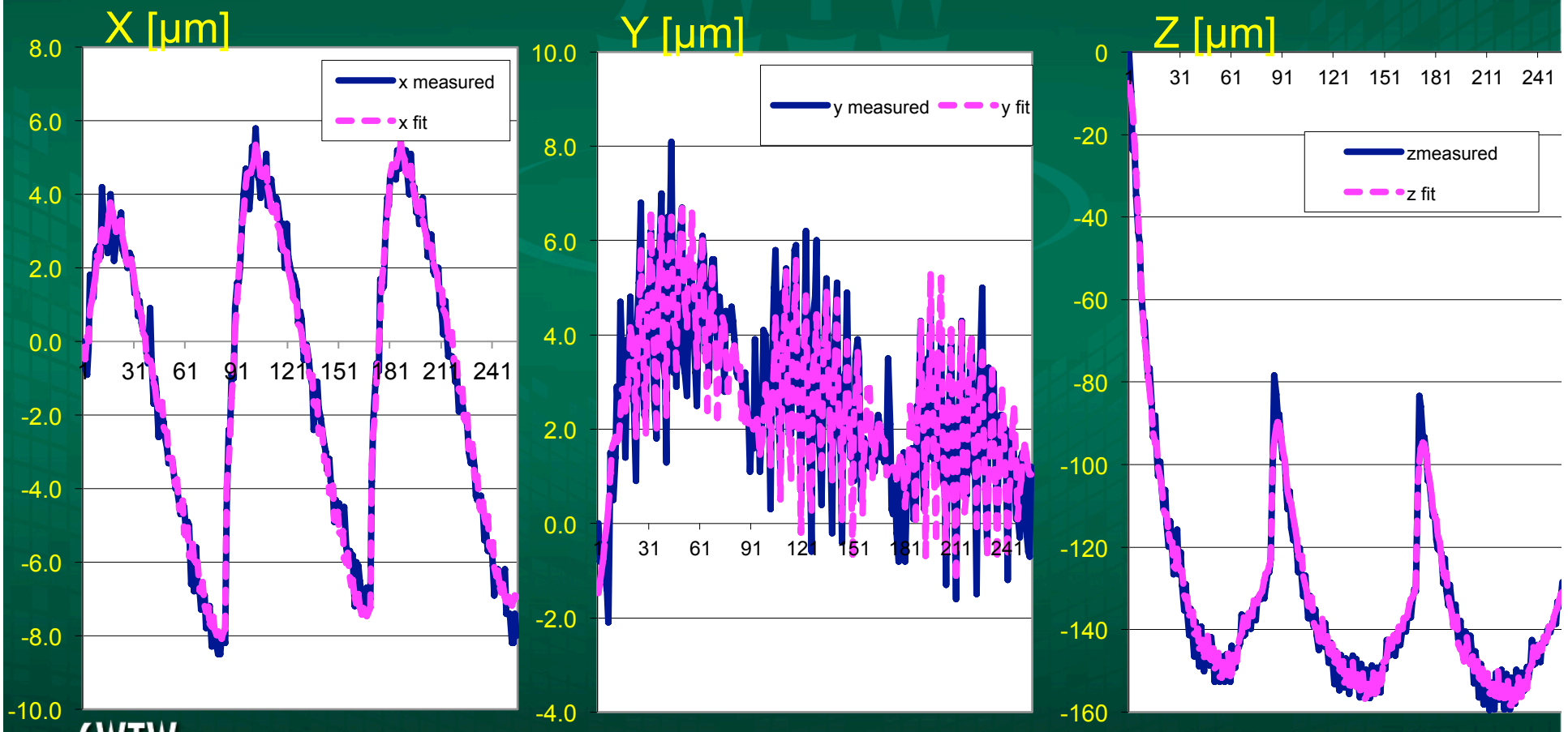
linear



linear and square

aproximation function (e.g. for x):

$$x = a_0 + a_1 T_1 + a_2 T_2 + \ldots\ldots + a_8 T_8 + a_9 T_1^2 + \cdots + a_{16} T_8^2$$



up to now:

fit of a given data set of sensorsignals (find coefficients)

note: this will mathematically always fit with many variables

now vice versa - anticipate:

compare calculated (with „old“ coefficients) values with the x/y/z – values („new“)

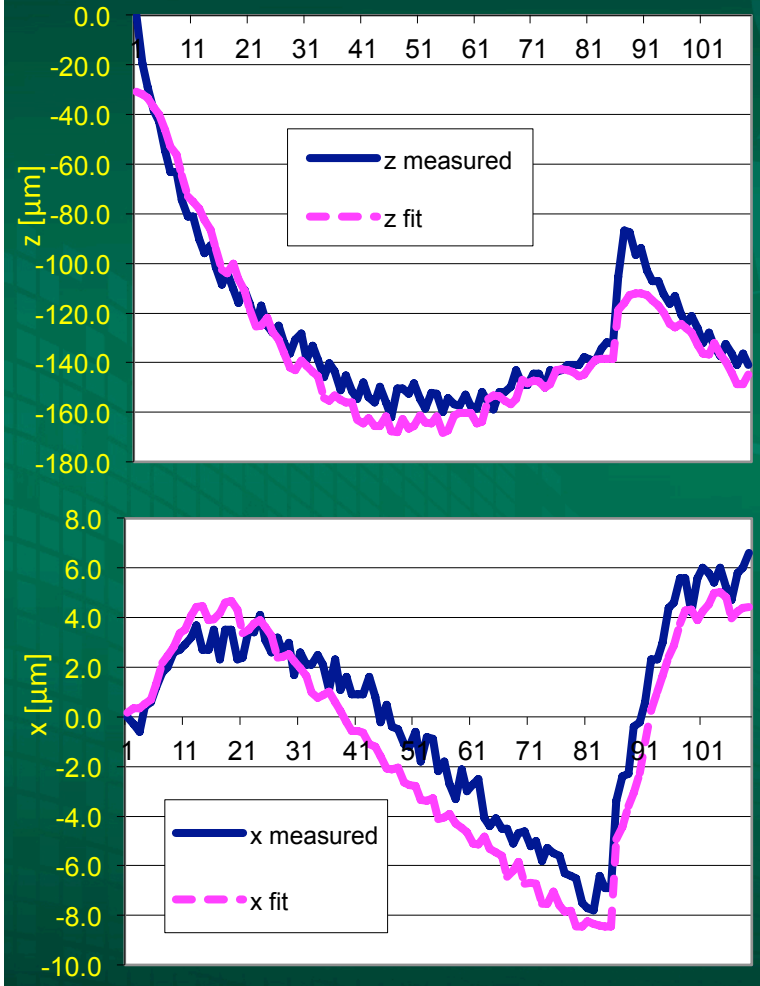
what do we require for a method in production:

minimum: it must be reproducible (same card/prober)

better: it must be transferable (same card, different prober)

optimum: it‘s universal for a given type of probecard

same probecard – same prober – different time comparison anticipated vs. real values

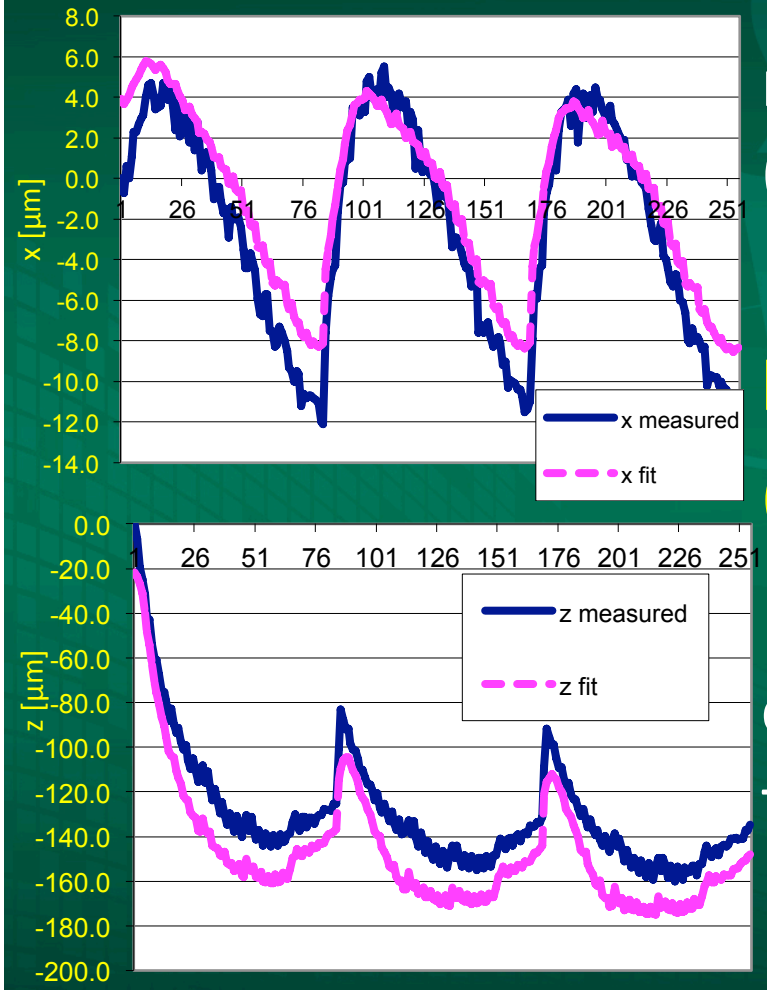


minimum: it must be reproducible (same probecard, same prober)

better: it must be transferable (same probecard, different prober)

optimum: it‘s universal for a given type of probecard

same probecard – different prober – different time comparison anticipated vs. real values

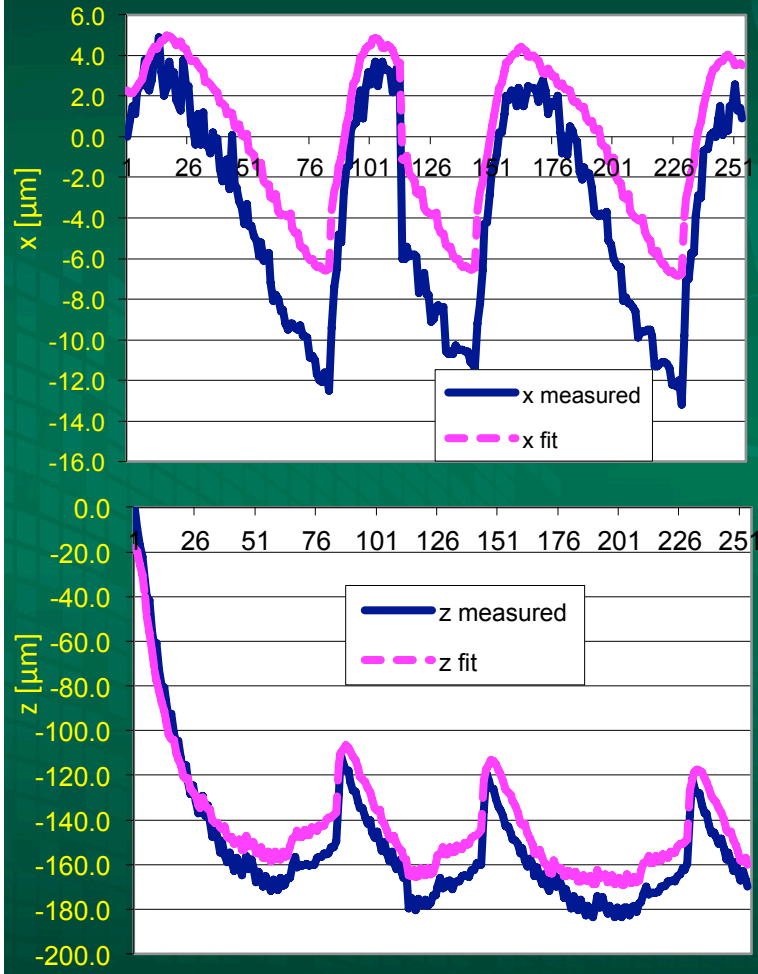


minimum: it must be reproducible (same probecard, same prober) ✓

better: it must be transferable (same probecard, different prober) ✓

optimum: it‘s universal for a given type of probecard

different probecard/prober – different time comparison anticipated vs. real values

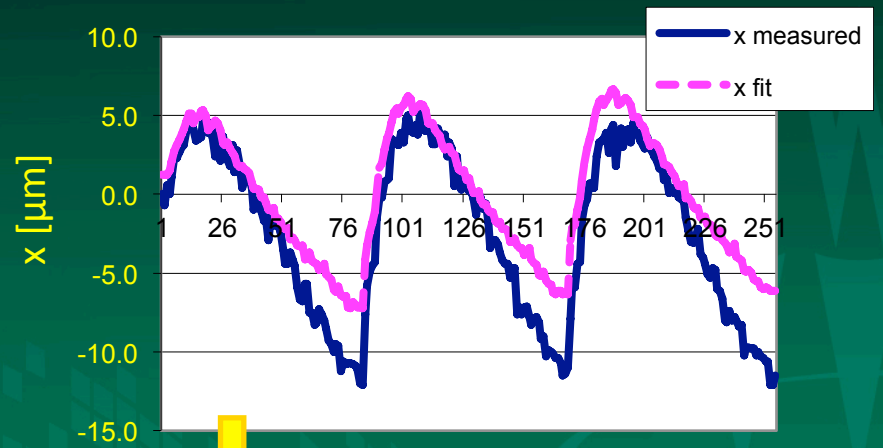


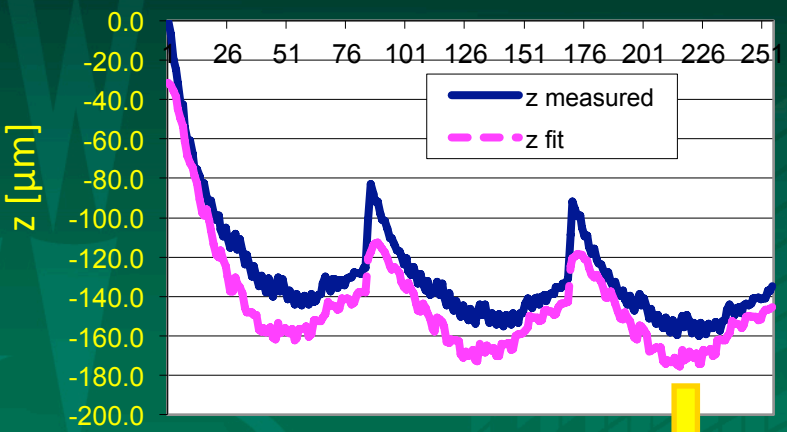
minimum: it must be reproducible (same probecard, same prober) ✓

better: it must be transferable (same probecard, different prober) ✓

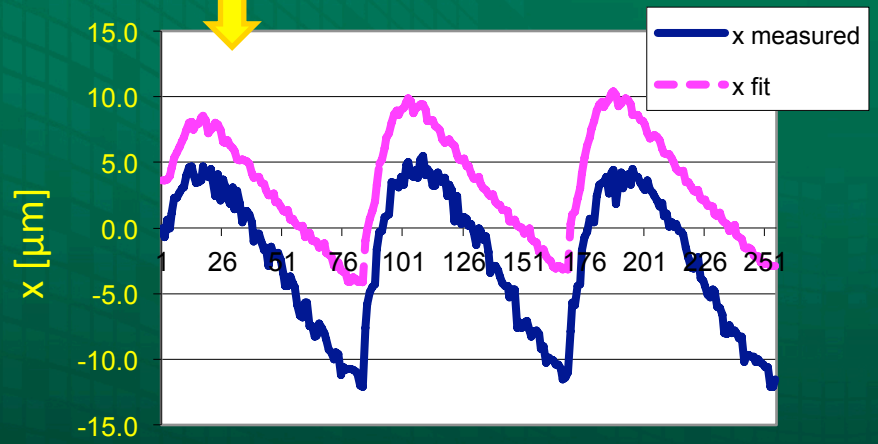
optimum: it's universal for a given type of probecard (✓)

calculation of perturbations

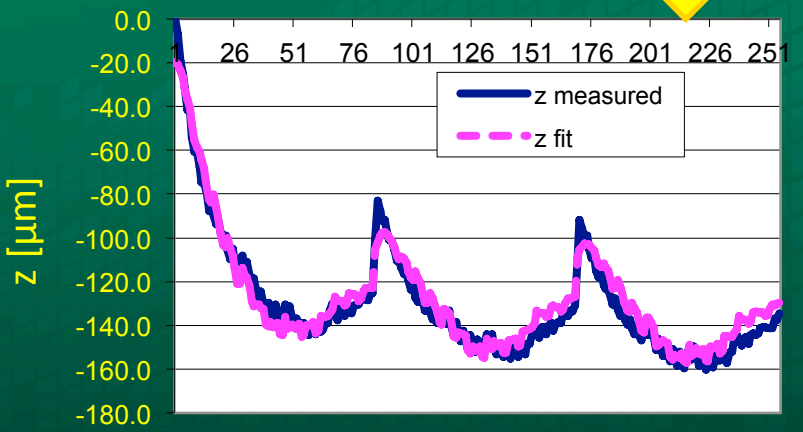




increase of only one sensor signal (T1) by 10 %



only half the accuracy in x



perfect (!) accuracy in z

Discussion of Results:

- **accuracy increased by 80 % in x,y & z**
- **independant of used type of probe card**
- **remaining inaccuracy is possibly due to**
  - non uniform application of sensors
  - mismatch of realignment in a decreasing temperature state
  - not having found the optimum fitting function
- **quality: extraordinary care is necessary**
  - control of chuck as function of sensor signals
  - checks are mandatory

Possible Follow-On for Improvements:

prober vendor:

open the prober for direct controling of x,y and z according to the model or implement the model direct in prober software

probecard vendor:

implement an array of precise and cheap sensors in the probecard itself.

algorithm:

addition of a retarded time function, higher power, fit piece by piece

we like to thank :

- Ph.D. Stefan Scholz for programming in Python
- the process team (especially Christos Roussou, Wolfgang Kirsammer and Peter Staigle) at EWS for all the trials and discussions

questions?