From Industrial Projects to Research Topics

Alessandro Di Bucchianico

SMSA Dresden March 6, 2019

TU/e Technische Universiteit Eindhoven University of Technology

Where innovation starts

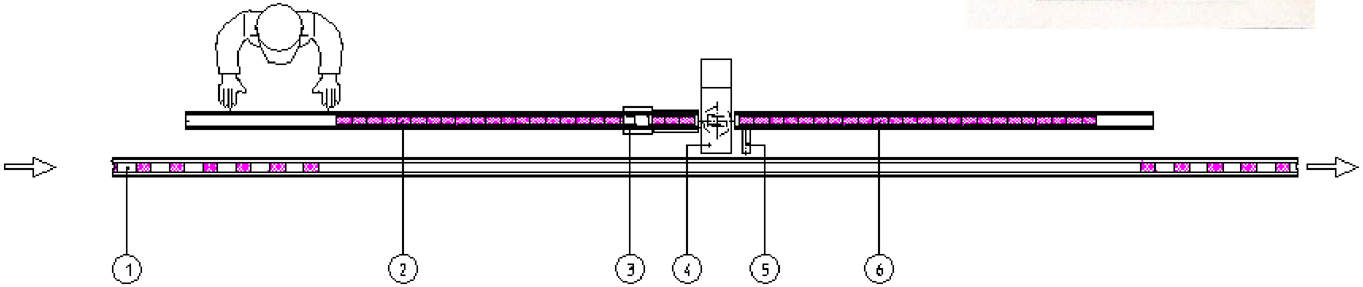
What should a keynote offer?

My choice:

Share my experiences with getting research inspiration from industrial projects

Industrial projects: sneak preview

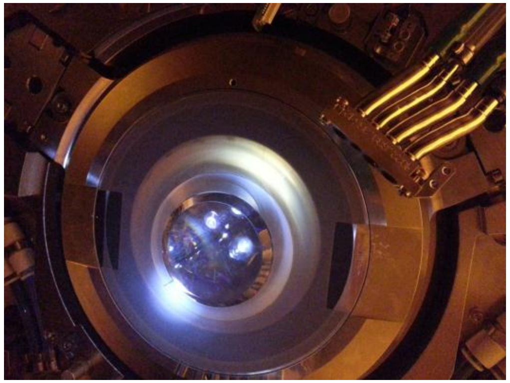
Leaks in Vacuum Coffee Packs



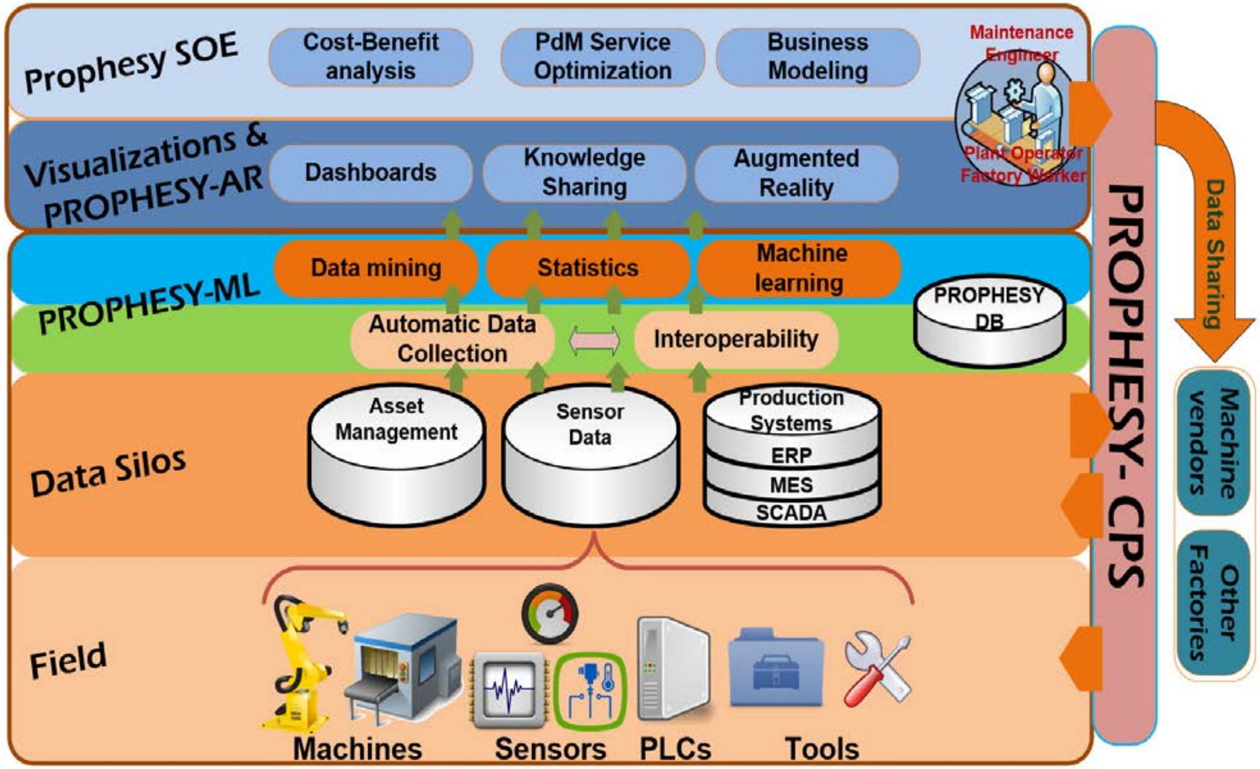
Wind Turbines

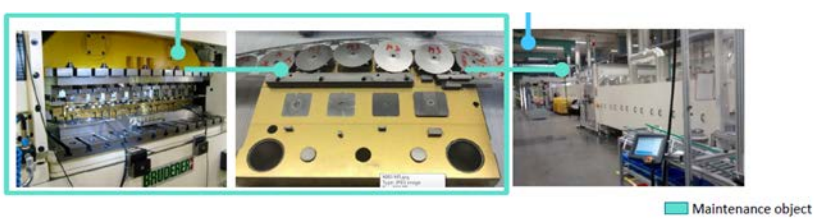
Technische Universiteit
Eindhoven
University of Technology

ASML – Wafer Stepper Machines

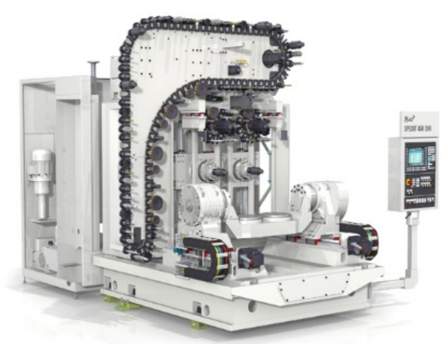


Prophesy – Predictive Maintenance





PROPHESY



TU/e Technische Universiteit Eindhoven University of Technology

Project Goals

Timely detect changes in percentages of leaks

Timely predict upcoming failures

Adaptive and self-configuring predictive maintenance system

Detection of degradation of optical system

Common topics

Projects share the following challenges

1. **timely detect changes over time**
2. **predict upcoming events/conditions**

We will treat these challenges from the viewpoint of SPC (Statistical Process Control), i.e. statistical techniques for monitoring changes over time (control is a historical misnomer).

Basics of SPC

Terminology

- **Statistical process control (SPC)**
  - **monitoring, no control**
  - **monitoring with intervention**
- **Changepoint detection**
  - **retrospective analysis**
- **Surveillance**
  - **monitoring without intervention**
- **Automatic Process Control (APC) / Engineering Process Control (EPC)**
  - **feedback control ("continuous intervention")**

Terminology in Different Communities

- **Statistical Process Control (industrial statistics, it is a historical misnomer)**
- **Anomaly detection (data mining)**
- **Concept drift (data mining)**
- **Surveillance (public health, usually no intervention possible)**
- **Changepoint analysis (econometrics/mathematical statistics, usually for the retrospective/off-line case)**

The Beginning – Shewhart's 1924 memo

Case 18013

WAS-724-5-16-24-FQ

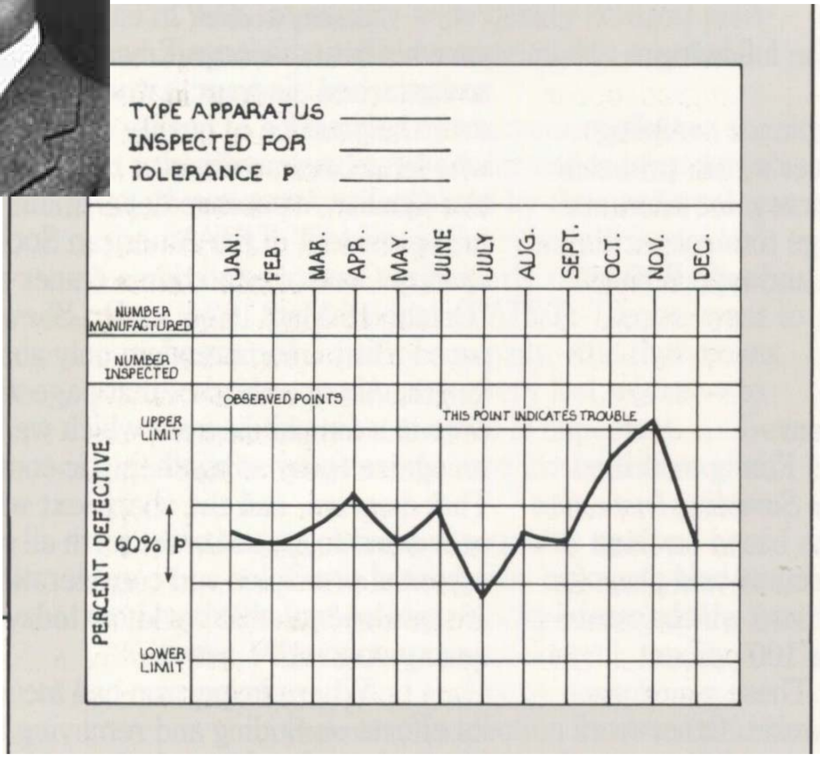
MR. R. L. JONES:-

A few days ago, you mentioned some of the problems connected with the development of an acceptable form of inspection report which might be modified from time to time, in order to give at a glance the greatest amount of accurate information.

The attached form of report is designed to indicate whether or not the observed variations in the percent of defective apparatus of a given type are significant; that is, to indicate whether or not the product is satisfactory. The theory underlying the method of determining the significance of the variations in the value of p is somewhat involved when considered in such a form as to cover practically all types of problems. I have already started the preparation of a series of memoranda covering these points in detail. Should it be found desirable, however, to make use of this form of chart in any of the studies now being conducted within the Inspection Department, it will be possible to indicate the method to be followed in the particular examples.

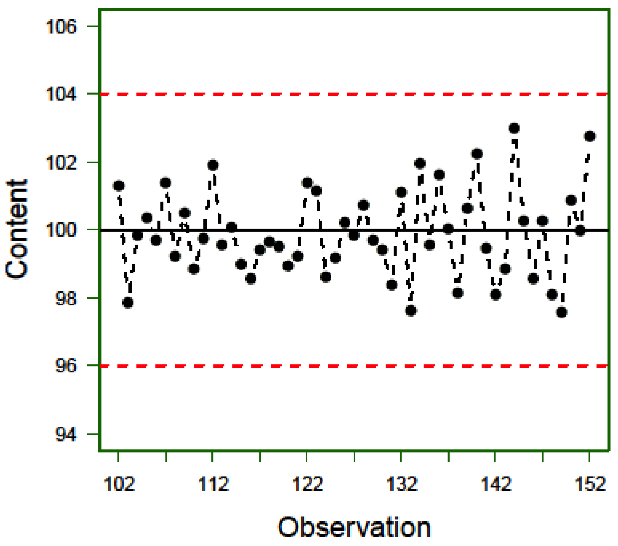
W. A. SHEWHART.

Enc.:
Form of Report.



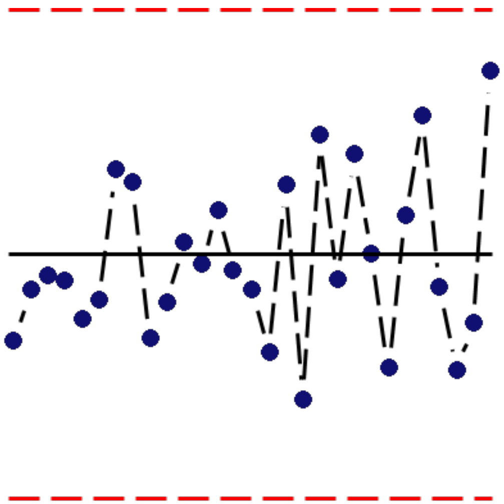
Shewhart Chart Basics

- **observations/computed statistics come in one by one**
- **decision at every observation/statistic**
- **red lines are "control" limits**
  (unfortunate historical name SPC = Statistical Process Control)
- **stop when observation/statistics is above UCL or below LCL (and then ...)**
- **this is not quite standard hypothesis testing ("*t*-test")**



Common control charts

Shewhart charts



EWMA = Exponentially Weighted Moving Average

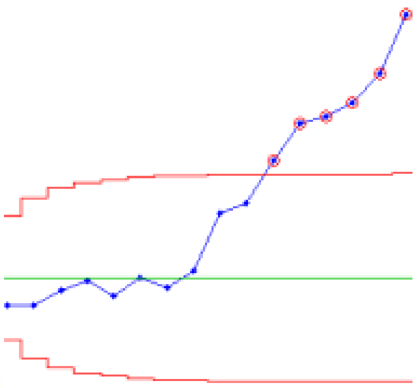
$$Y_i = \lambda X_i + (1-\lambda) Y_{i-1}$$

originated from Bayesian setting (Shiryaev-Roberts) , robust against deviations from normality

CUSUM = Cumulative Sums Chart cumulative sums with reflecting boundary at 0,

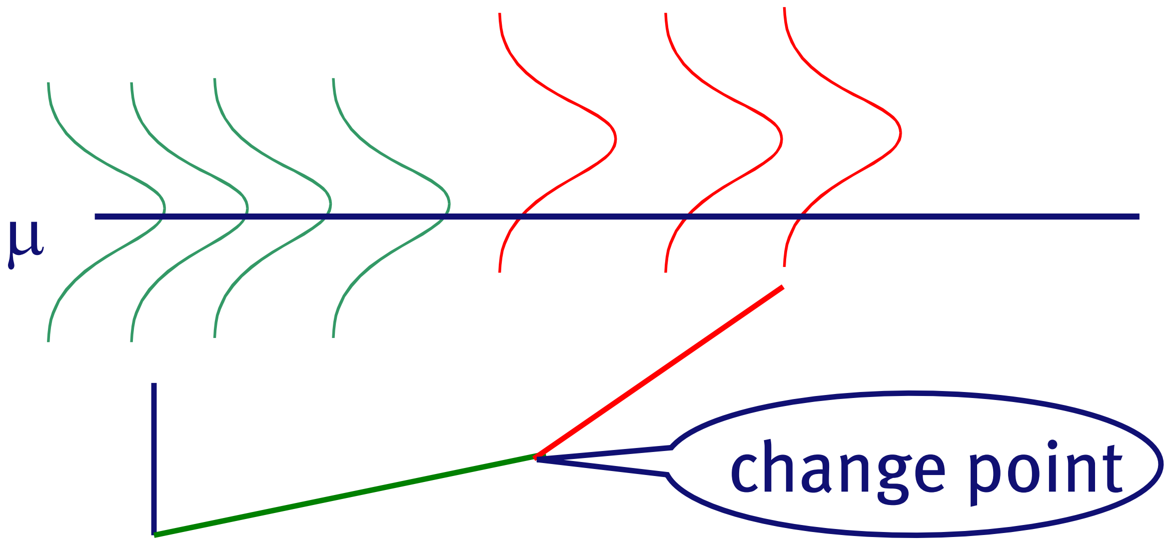
GLR = Generalized Likelihood Ratio Chart sequential form of likelihood ratio tests

CCC = Cumulative Count of Conforming Chart originally for high-yield processes (time between events or several events)



TU/e Technische Universiteit Eindhoven University of Technology

CUSUM Charts



- better at small persistent shifts than Shewhart charts
- does have memory
- asymptotically optimal (in some technical sense)
- ARL computations involve solving Fredholm integral equations ; simple discretizations lead to Markov chain approaches

TU/e Technische Universiteit Eindhoven University of Technology

SPC Toy model

Grouped observations (manufacturing context)

- $Y_{ij} \sim N(\mu_i, \sigma^2)$ **independent** , $i = 1, 2, \ldots$, $j = 1, \ldots, n_i$
- **index *i* corresponds to equi-distant time intervals**
- **important special case:** $n_i = 1$

"In-control situations"

$$\mu_0 = \mu_1 = \ldots$$

versus "out of-control situations"

$$\mu_0 = \mu_1 = \ldots = \mu_{\tau-1} , \mu_\tau = \mu_{\tau+1} = \ldots = \mu_0 + \delta$$

Technische Universiteit Eindhoven University of Technology

Control Charts and Hypothesis Testing

Let τ be a parameter (if you are a frequentist) or a random variable (if you are a Bayesian)

Standard changepoint model (persistent change of the mean)

$$H_0 : \mu_0 = \mu_1 = \ldots$$

versus

$$H_1 : \mu_0 = \mu_1 = \ldots = \mu_{\tau-1}, \mu_\tau = \mu_{\tau+1} = \ldots = \mu_0 + \delta$$

See Does and Schriever (1992) for more examples

These settings are not realistic.

Change Scenarios

There are more realistic changes than the persistent change of mean, e.g.

- **Gradual changes**
- **Chemical processes that cannot be kept a setpoint**
- **Trend reversals in business cycles**
- **Processes with feedback controllers (epidemic changes/window of opportunity)**
- **…**

See Di Bucchianico and Van den Heuvel (2015).

Hypothesis Tests for Simple Hypotheses

Neyman-Pearson Lemma (hypothesis testing)

$$H_0:\ \theta = \theta_0 \text{ versus } H_a:\ \theta = \theta_1$$

Optimal test using likelihood ratio:

$$\Lambda_k = \frac{f_{\theta_1}(x_1)\cdots f_{\theta_1}(x_k)}{f_{\theta_0}(x_1)\cdots f_{\theta_0}(x_k)}$$

Adapting this setup to a sequential version leads to CUSUM charts in the case of testing of means for normal distributions (see Frisén and De Maré (1991)).

General Hypotheses - Example

$$H_0: \delta_1 \le \mu_i \le \delta_2 \text{ for all } i$$

versus

$H_1: \delta_1 \le \mu_i \le \delta_2 \ (i = 1, \cdots, \tau - 1)$ and $\mu_i < \delta_1$ or $\mu_i > \delta_2$ for $i = \tau, \tau + 1, \cdots$

Versions possible for several distributions.

Other versions possible (e.g., tool wear may be described with monotonicity constraints).

Tailor-made procedures may be derived using statistical theory

Generalized Likelihood Ratio Tests

Procedures for the general scenarios with their complex null and alternative hypothesis structures may be obtained using GLR (with slight abuse of notation to indicate the hypotheses:

$$\Lambda_k = \max_{1 \le i \le k} \frac{\prod_{j=1}^{i} \sup_{\theta_j | H_0} f_{\theta_j}(x_j) \prod_{j=i+1}^{k} \sup_{\theta_j | H_1} f_{\theta_j}(x_j)}{\prod_{j=1}^{k} \sup_{\theta_j | H_0} f_{\theta_j}(x_j)}$$

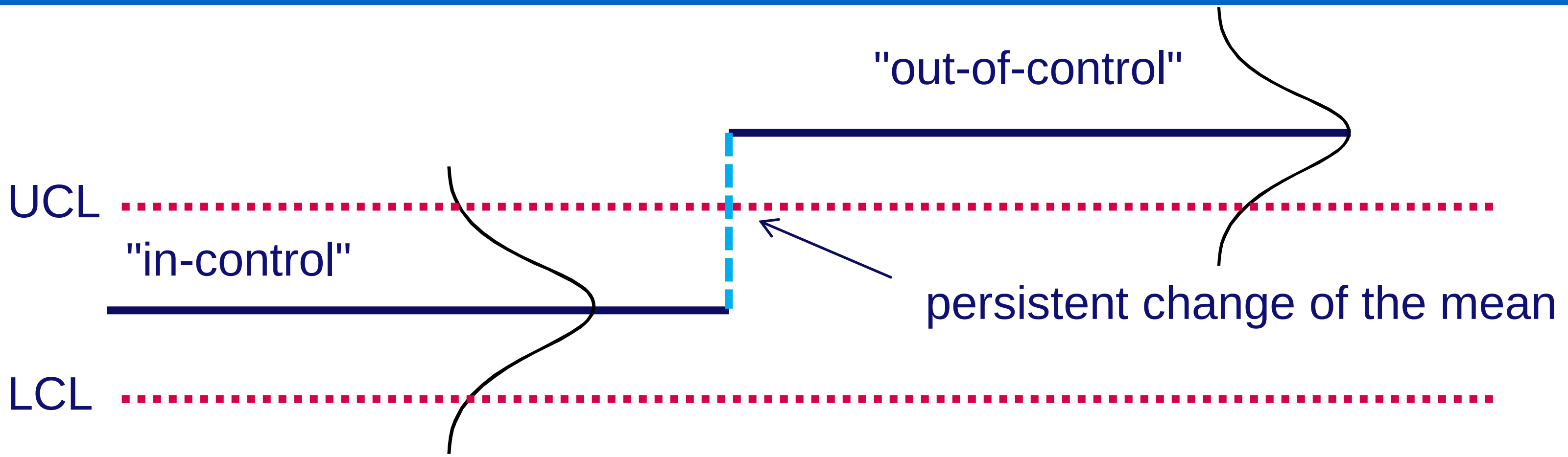
Explicit expression may be obtained in several cases.

For hypotheses with monotonicity constraints one needs to use isotonic regression.

Practical evaluation by restricting "max" in Λ_k to a fixed window. See Di Bucchianico et al (2004).

TU/e Technische Universiteit Eindhoven University of Technology

Detection Performance



Desired performance:

- **as quick as possible "alarm" when out-of-control**
- **as few as possible "false alarms" when in-control**

Type I/II errors do not capture this. An appropriate description is the "run length distribution".

Run Length Distribution

For Shewhart charts: both in-control and out-of-control run lengths have a geometric distribution.

Standard 3σ limits lead to ARL = 1/(2*0.00135) = 370.

Run length distributions are usually skewed, so looking only at ARL (Average Run Length) may be misleading:

- **SRL (Standard Deviation of Run Length)**
- **quantiles**

There also issues with starting the run length at 0 (conditional run lengths): see e.g. Kenett & Pollak (2012).

Phase I - Phase II

Traditional setup from manufacturing:

- **Phase I (pilot/training phase; retrospective change detection – setting of control limits)**
- **Phase II (on-line monitoring)**

But many cases do not naturally have such a division.

Self-starting control charts have been developed that recursively estimate control limits.

Coffee packs

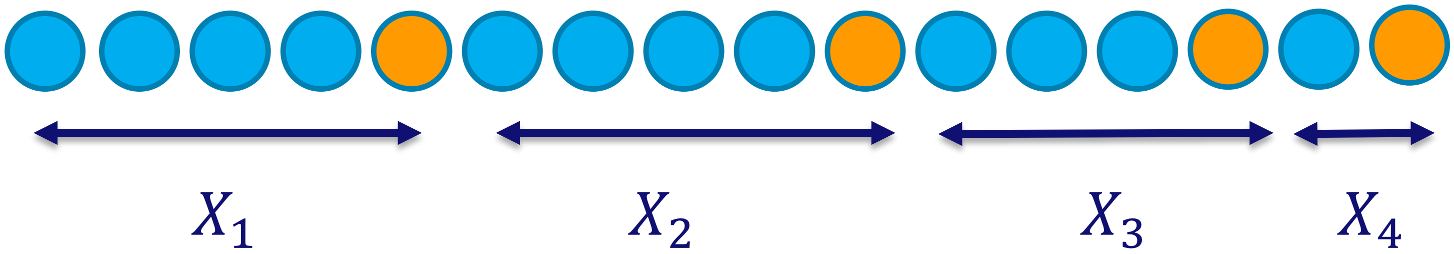
Coffee packs: industrial problem

- small leaks in vacuum seals may manifest themselves much later
- percentage of packs with leaks is very low (0.2 %)
- high volume production of coffee packs
- reduce time needed to detect increased percentage of leakages
  - increase customer satisfaction
  - reduce failure costs

Coffee packs: statistical challenges

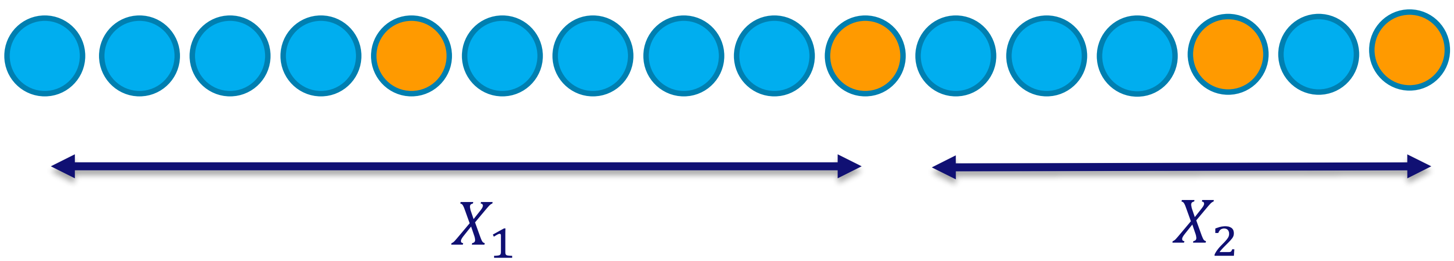
- **leak detection device yields only yes/no result, so binary data (yes/no)**
- **standard Shewhart control charts are not appropriate because $p \approx 0.002$:**
  - **3σ-limits incorrect since normal approximation is poor**
  - **LCL for 3 σ-limits is negative unless *n* is large**
  - **UCL is such that any leaking pack causes an alarm**

CCC = Cumulative Count of Conforming chart

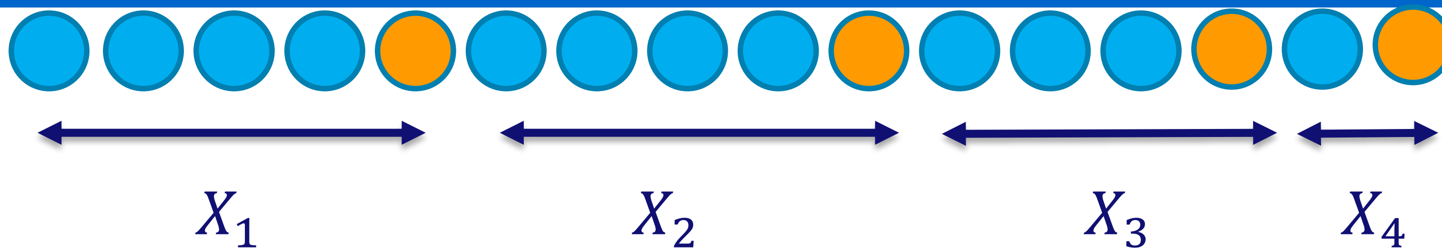


CCC rule: alarm if $X_i < LCL$
(too frequent failures indicate deterioration)

A variant of this is the CCC-*r* chart (below *r*=2)



Run Lengths and Stopping Times



$$N := \min\{i \mid X_i < \mathrm{LCL}\}$$

Note the two different time scales:

1. **the original time scale (number of coffee packs)**
2. **the index *i* of the control chart statistics X_i**

ARL = E(*N*) cannot be used since different CCC-*r* charts cannot be compared directly.

Instead we use the distribution of $\sum_{i=1}^{N} X_i$.

Expectation of Run Length Distribution

$N := \min\{i | X_i < LCL\}$ **is a stopping time w.r.t.** $X_1, X_2, \ldots$

Wald Formula **(version for stopping times)**

$$E\left(\sum_{i=1}^{N} X_i\right) = E(N)E(X)$$

Distribution is skewed, so only computing mean is not sufficient.

Variance of Run Length Distribution

Blackwell-Girshick Formula

If *N* is independent of iid $X_1, X_2, \ldots$ with $\sigma^2 = Var(X_i) < \infty$, then

$$Var\left(\sum_{i=1}^{N} X_i\right) = E(N)Var(X) + Var(N) + E^2(X)$$

This does not apply here because $N = \min\{i | X_i < LCL\}$

Correct formula for this specific choice of *N*

$$Var\left(\sum_{i=1}^{N} X_i\right) = \frac{\sigma^2}{p} - \mu^2\left(\frac{3-p}{p^2}\right) + 2\mu E\left(N\sum_{i=1}^{N} X_i\right)$$

where $p = P(X_i < LCL)$

Technische Universiteit Eindhoven University of Technology

Results CCC-*r* charts coffee packs

We compared several scenarios with $p_{in} = 0.002$ and p_{out}=0.008 .

Comparison of CCC-*r* charts with 3 leak detection devices

r	LCL	ALIin	SDLIin	ALIout	SDLIout
1	8	99.8	101.2	4.1	4.3
2	129	100.9	102.7	1.5	1.7
3	380	100.6	102.6	**1.1**	1.1
4	709	100.3	102.3	1.2	**0.8**
5	1087	100.2	102.0	1.4	0.7

Coffee packs: further research topics

- other stopping times:
  - CCC-EWMA
  - CCC-CUSUM
  - self-starting versions
  - …
- incorporate incomplete inspection
- continuous time analogues
- overdispersed count data
- …

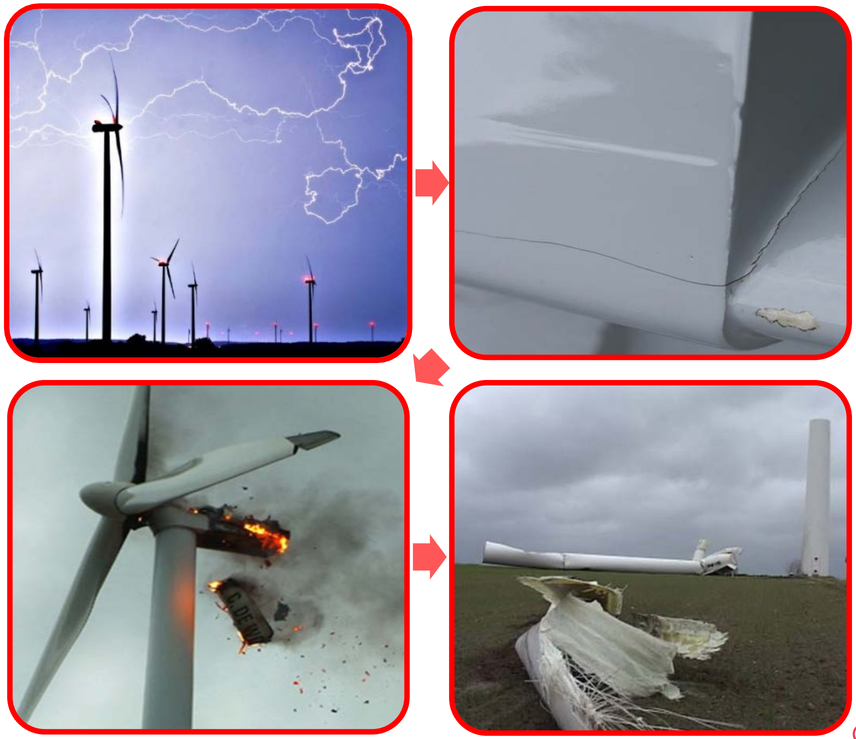
An extensive overview paper on control charts in this setting is Ali et al (2016).

TU/e Technische Universiteit Eindhoven University of Technology

Wind turbines

Wind turbines: industrial problem

Timely detection of tiny cracks that much later lead to severe damage

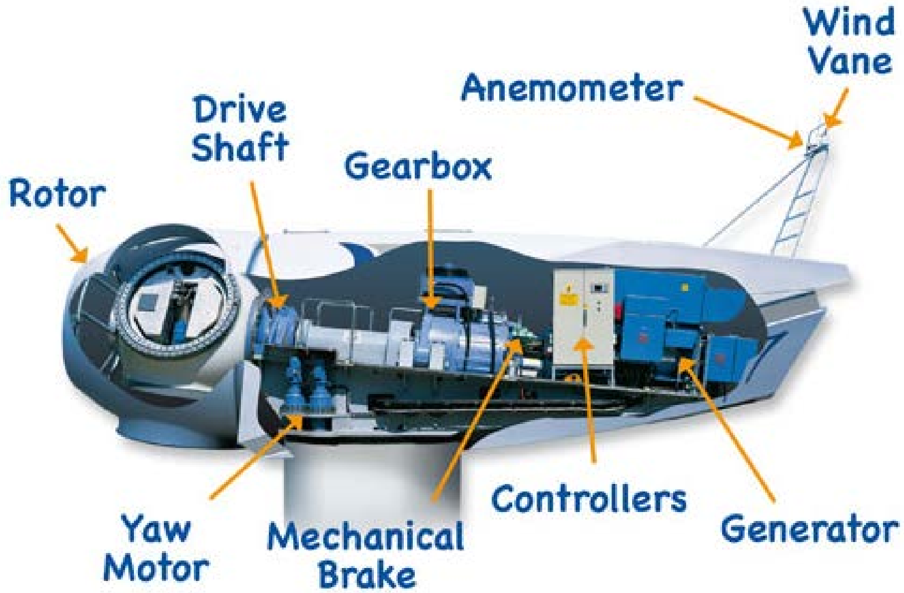


Google images

Wind turbines: available sensor data

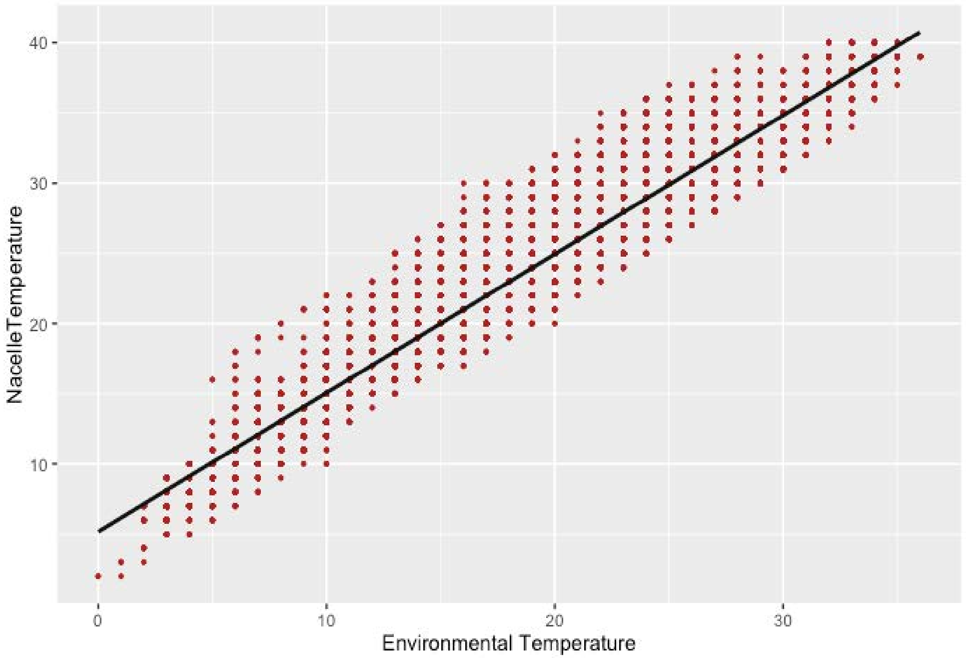
Monitoring of wind turbine using 4 min data

- **power output**
- **temperatures**
- **vibrations**

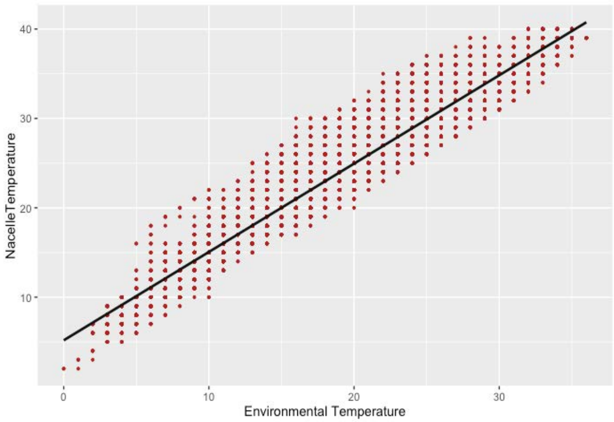


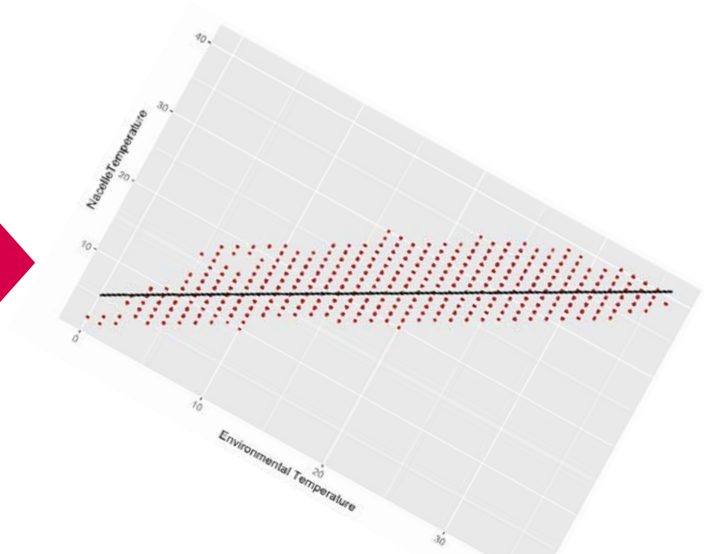
Wind turbines: statistical challenges

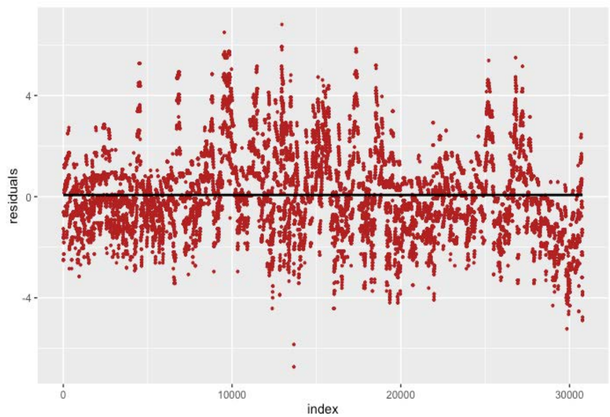
- no Phase I (pilot) data available
- static control limits are not appropriate

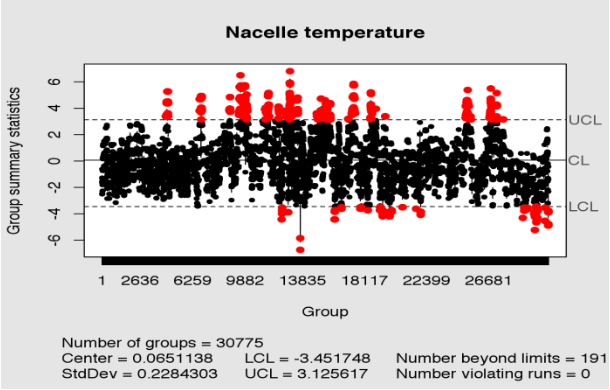


Regression Approach – Temperature









Monitoring approach 1

- **Linear regression model:** $Y = X\beta + \varepsilon$, $\varepsilon \sim N(0, \sigma^2 I)$
- Y **is vector of univariate observations**
- **In-control period (baseline, Phase I):** $Y_1, \ldots, Y_m$
- **On-line period (Phase II):** $Y_{m+1}, Y_{m+2}, \ldots$
- **Monitor (standardized) prediction residuals**

$$Y_{m+1} - \hat{Y}_{m+1}, Y_{m+2} - \hat{Y}_{m+2}, \ldots$$

where $\hat{Y}_{m+i} = x_{m+i}^{\boldsymbol{T}} \; \hat{\beta}_{(m)}$ **with** $\hat{\beta}_{(m)}$ **the LS estimator based on** $Y_1, \ldots, Y_m$

Technische Universiteit Eindhoven University of Technology

Monitoring approach 1: issues

1. How to determine *m* (baseline/Phase I)?
2. How to monitor prediction residuals
   a) residuals are dependent
   b) what change ("out-of-control situation") do we wish to detect?

Ad 2b) Possible setup for model $Y = X\beta + \varepsilon$

- null hypothesis: β does not change
- alternative hypothesis: β constant until τ , change to β^* afterwards

Technische Universiteit **Eindhoven** University of Technology

Monitoring approach 2

- **Linear regression model:** $Y = X\beta + \varepsilon$ **,** $\varepsilon \sim N(0, \sigma^2 I)$
- Y **is vector of univariate observations**
- **Monitor recursive residuals of Browne et al (1975), i.e. standardized versions of** $Y_i - x_i^T \hat{\beta}_{(i-1)}$

Issues:

- **no run length distribution / specification of change scenario (only type I error)**
- **might not be able to detect gradual changes (since regression parameters are re-estimated)**

Monitoring approach 3: fluctuation tests

- **Linear regression model:** $Y = X\beta + \varepsilon$, $\varepsilon \sim N(0, \sigma^2 I)$
- Y **is vector of univariate observations**
- **Monitor changes in parameter estimates ($\hat{\beta}_{(i)}$), rather than residuals**

- **Issues:**
  - **no run length distribution (only type I error; asymptotic distribution of "fluctuation process" by FCLT, see Chu et al (1996) or Zeileis et al (2005))**
  - **might not be able to detect gradual changes (since regression parameters are re-estimated)**

Prediction of 2014 failure

Model	Warnings for imminent failure
Nacelle temperature	Oct-Nov 2013
Oil temperature	Oct-Nov 2013
Bearing temperature	Oct-Nov 2013
Gearbox temperature	Oct-Nov 2013
Main gen. temperature	Oct-Nov 2013
Starting gen. temperature	Oct-Nov 2013
Power output main gen.	Oct-Nov 2013
Power output starting gen.	None

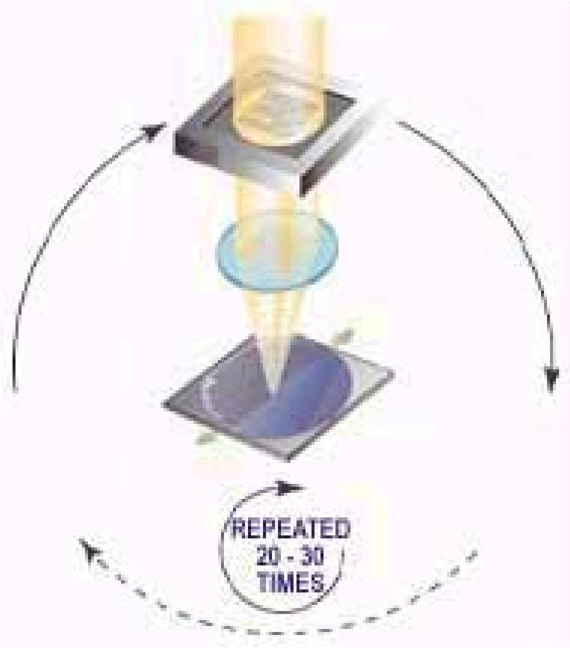
Wind turbines: future research topics

- perform run length distribution comparisons between monitoring approaches
- multivariate monitoring
- specification of out-of-control scenarios/alterative hypotheses
- combine control charts variable selection methods (lasso, least angle regression) see e.g. overview paper Capizzi (2015)
- develop online versions of Theil's 1968 BLUS residuals (BLUS = Best Linear Unbiased Scalar Covariance)
- use weighted regression (see. e.g Horvath et al (2004))

Technische Universiteit Eindhoven University of Technology

ASML – Wafer Stepper Machines

Optics



Industrial problems

- Standstill of machine extremely costly
- Large number of machines at customers to be monitored
- Specific case: gradual degradation of optical performance through specific part
- Measurements are costly
- Not clear what to monitor

Technische Universiteit **Eindhoven** University of Technology

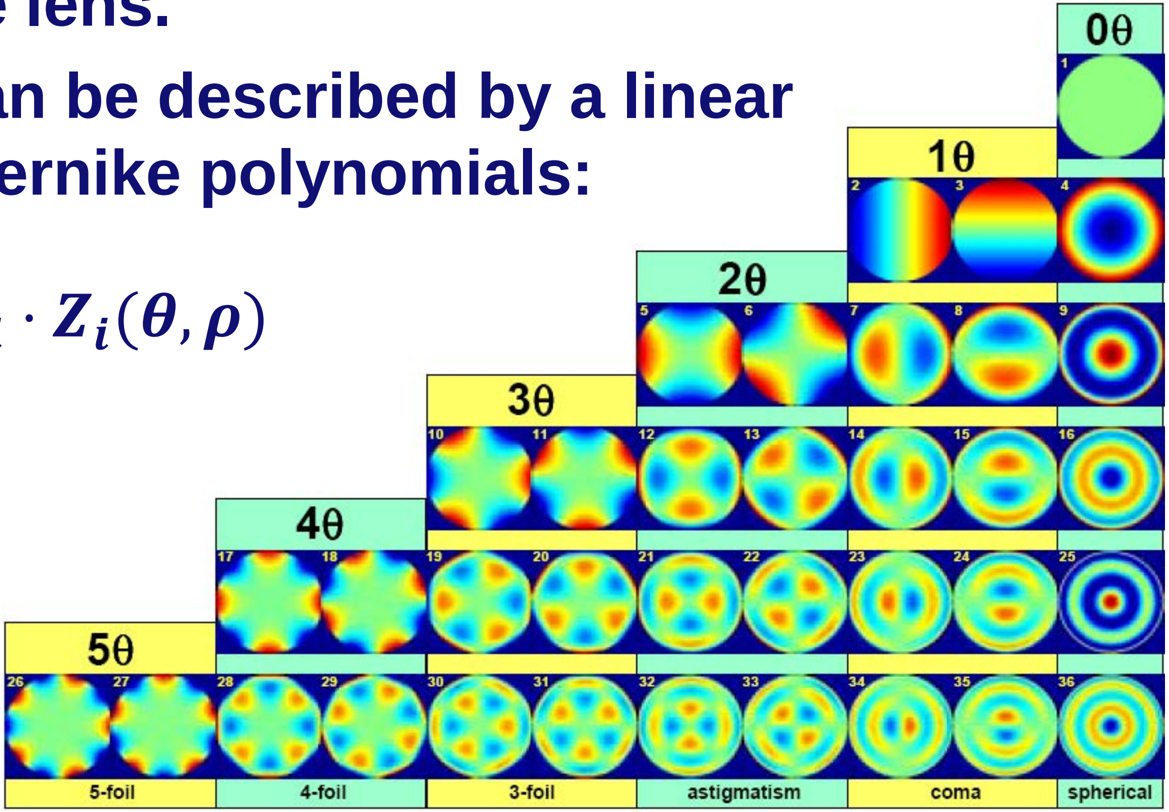
Statistical challenges

- compete with machine learning approaches
- no Phase I/in-control
- high dimensional data in the form of Zernike polynomials (measured in some way)
- lead time for detection
- optical system undergoes several feedback control actions

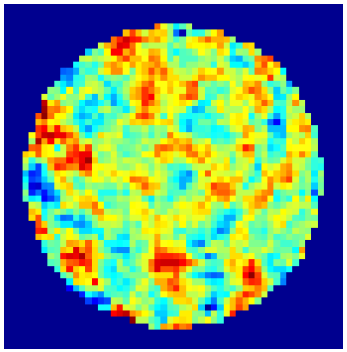
Zernike polynomials

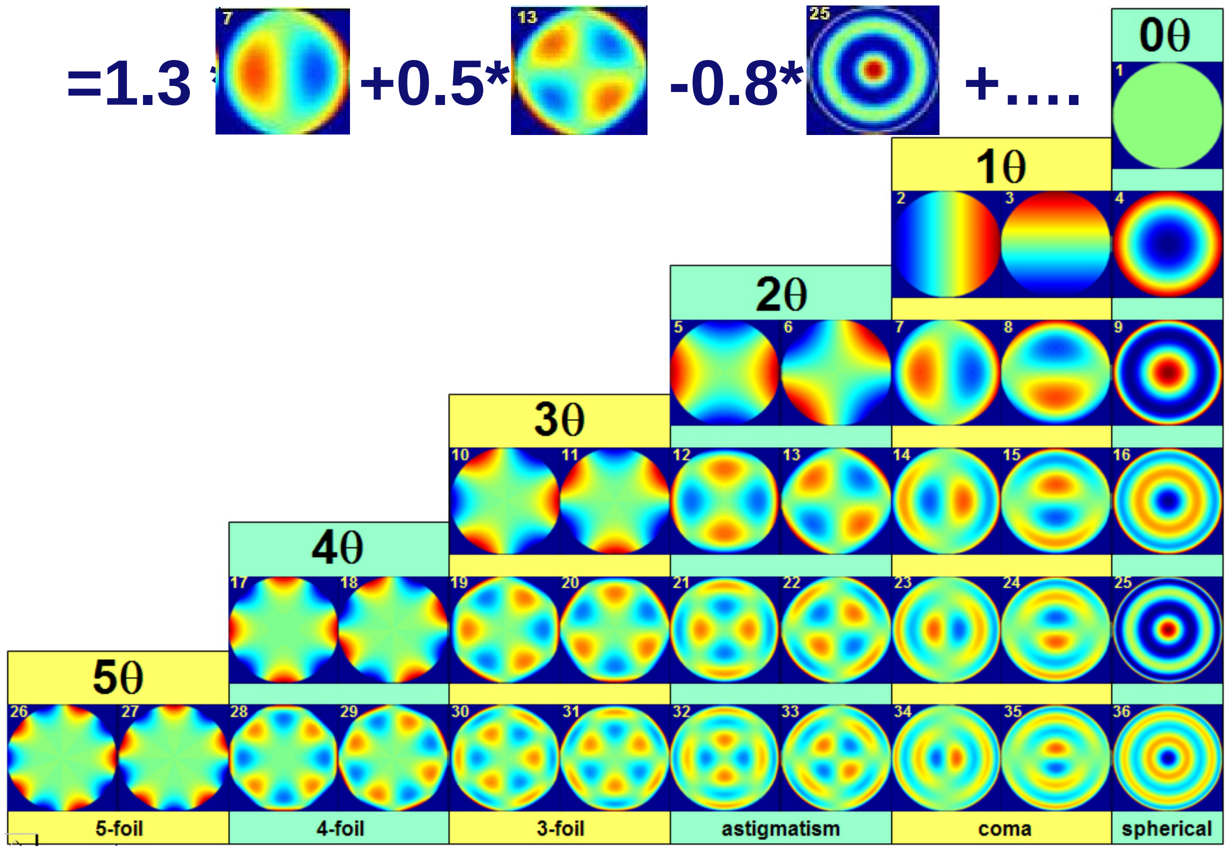
- **Zernike polynomials describe the aberrations of wavefronts of the lens.**
- **Any wavefront can be described by a linear combination of Zernike polynomials:**

$$W(\theta,\rho) = \sum_{i=1}^{\infty} a_i \cdot Z_i(\theta,\rho)$$



Zernike expansion



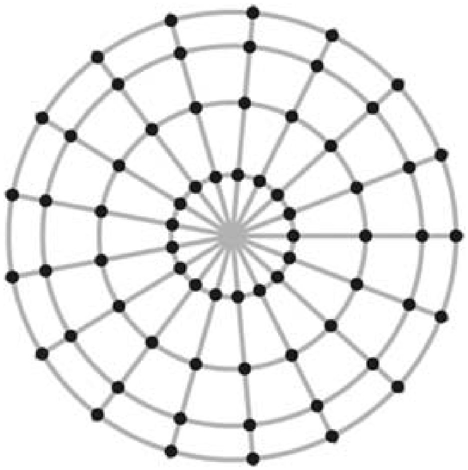


Zernike polynomials: orthogonality

Zernike polynomials are orthogonal (useful for statistical purposes) in a continuous sense.

Orthogonality with data is discrete orthogonality

Optimal designs for LS and Fourier estimation have been developed (Dette et al. (2007))



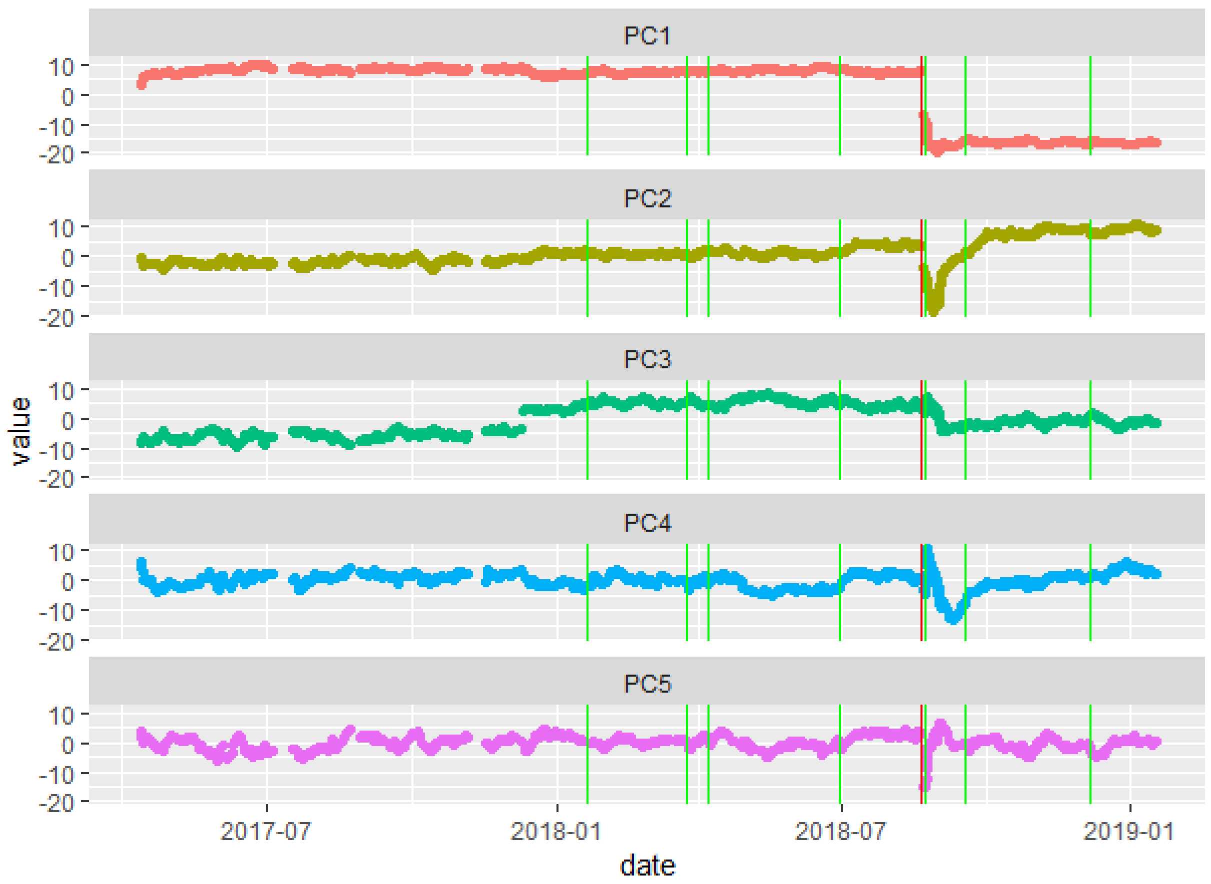
But: - circular designs are not practical

- company measures several wavefronts

Main approach

- **Apply PCA to find linear combinations of Zernike coefficients**
- **Add feedback control actions to variables to be monitored**
- **Apply self-starting control charts**

Results



PC1
PC2
PC3
PC4
PC5
10
0
-10
-20
value
2017-07
2018-01
2018-07
2019-01
date

Statistics versus machine learning

Statistics:

- **good**
  - **“guaranteed performance”**
  - **incorporate time**
- **not so good:**
  - **flexibility (usually requires modelling)**
  - **scalability**

Machine learning has opposite (dis)advantages.

Models

A common distinction is made between

1. predictive models
2. explanatory models

I feel that there is a need for third type of model:

3. benchmark models

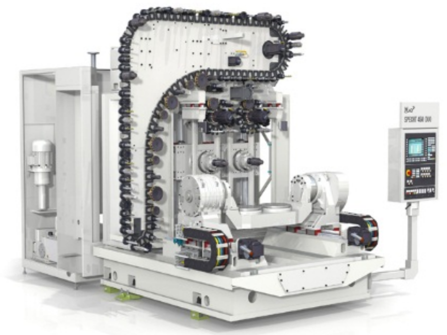
Lessons learned

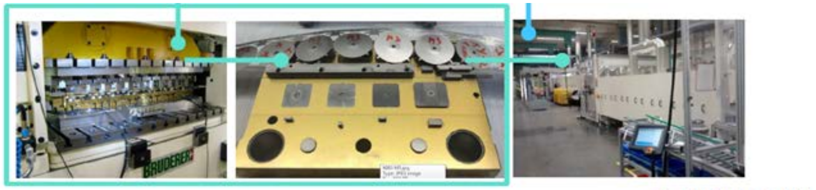
1. Control actions contain useful information about degradation and should thus be used in monitoring
2. Dimension reduction should involve “linear” combinations of all parameters rather than choosing subsets
3. Prediction procedures should be tuned to classes of machines

Wafer steppers: future research topics

- self-starting high-dimensional control charts
- hybrid approach:
  - use statistics to find onset of gradual deterioration
  - use machine learning to find patterns / thresholds
- optimal grid within practical restrictions for Zernike coefficients
- estimation of RUL (remaining useful life)

Prophesy





Prophesy questions

- **Estimate RUL (Remaining Useful Life) with uncertainty margins**
- **Automaticaly update maintenance plans based on continuous streams of sensor data (“predictive maintenance”/ “condition based maintenance”)**

Statistical challenges

- **compete with machine learning approaches**
- **no Phase I/in-control period**
- **high dimensional, high frequency data**
- **few failures**
- **combine with maintenance optimization**

Comparison of SPC and CBM

- **SPC**
  - **has no failure**
  - **has minimal maintenance to go from out-of-control to in-control**
  - **process continues even when out-of-control**

- **CBM**
  - **usually no distinction between operating states**
  - **several types of maintenance**
  - **shut down during maintenance**

TU/e Technische Universiteit Eindhoven University of Technology

Role of monitoring

Level 1: Change Detection

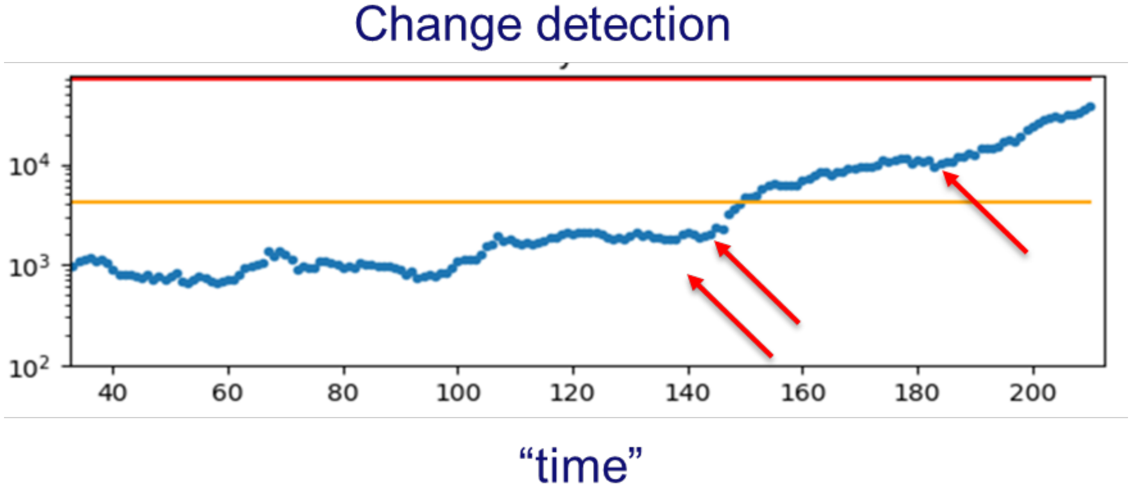
- Detect if everything is working well

Level 2: RCA – Root Cause Analysis

- Analyze where the failure comes from

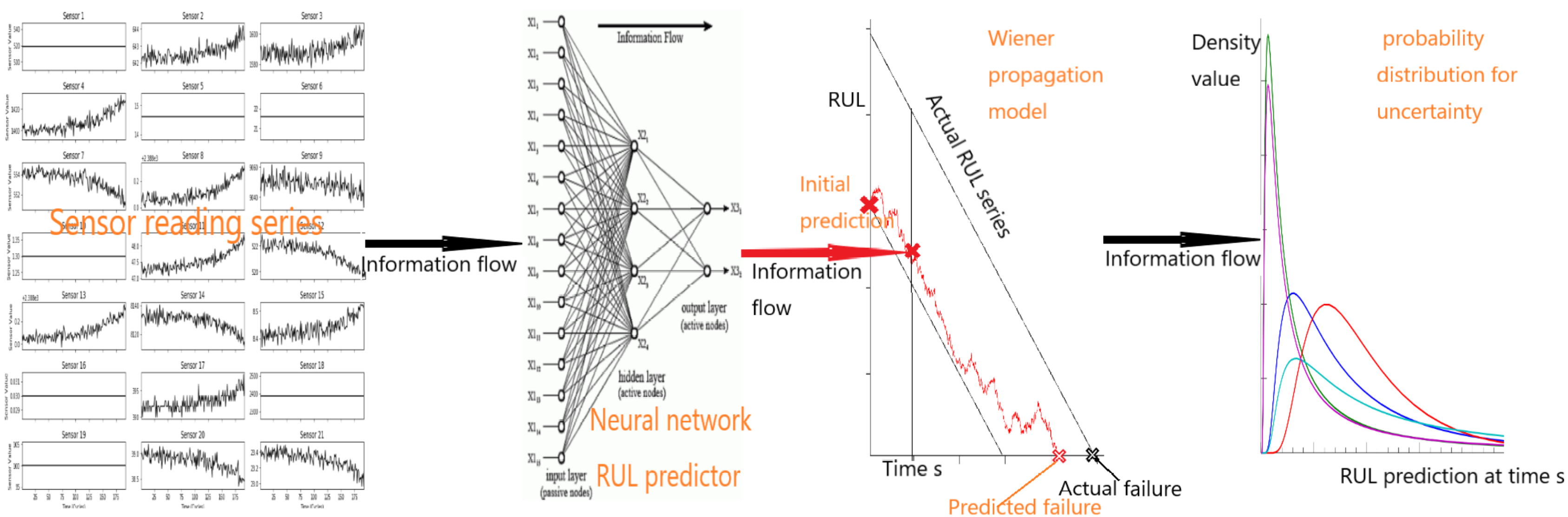
Level 3: RUL – Remaining Useful Life

- Calculate the RUL based on wear data



Hybrid approach: uncertaintly

- **Machine learning to estimate RUL**
- **Degradation modelling with Wiener process with drift to estimate uncertainty**



Technische Universiteit Eindhoven University of Technology

Prophesy: future research topics

- **integration of SPC with CBM**
- **integrate SPC with RUL estimation**
- **estimate RUL with uncertainty margin**
- **use of knowledge of physical models**
- **…**

Conclusion

Tips to get into contact

- network
- listen to problems
- contact R&D department of companies
- Data Science Center
- join other groups (industrial engineering, computer science,…)
- alumni
- …

Tips for success

- find champion within company
- use Master students
- try to set up “road map”

Conclusions

Problems from industrial projects are a source for research topics in statistical monitoring that deviate from “standard settings”.

It seems worthwhile to investigate hybrid approaches combining statistics with machine learning

References

S. Ali, A. Pievatolo and R. Göb, An Overview of Control Charts for High-quality Processes, Qual. Rel. Engng. Int. *32*(7), pp.2171-2189, 2016

R.L. Brown, J. Durbin and J. M. Evans, Techniques for testing the constancy of regression relationships over time. Journal of the Royal Statistical Society, Series B 37 (1975), 149–92.

C.-S.J Chu, M. Stinchcombe and H. White, Monitoring structural change. Econometrica 64 (1996), 1045–1065.

G. Capizzi, Recent Advances in Process Monitoring: Nonparametric and Variable-Selection Methods for Phase I and Phase II. Quality Engineering 27 (2015), 44-67.

References (continued)

A. Di Bucchianico, M. Hušková, P. Klášterecký, and W.R. van Zwet. Performance of control charts for specific alternative hypotheses. In: J. Antoch (ed). COMPSTAT 2004 Symposium, 903-910.

A. Di Bucchianico, G.D. Mooiweer, and E.J.G. Moonen. Monitoring infrequent failures of high-volume production processes. Qual. Reliab. Engng. Int., 21:521–528, 2005.

A. Di Bucchianico and E.R. van den Heuvel. Shewhart's Idea of Predictability and Modern Statistics. In: S. Knoth and W. Schmid (eds.), Frontiers in Statistical Quality Control 11: 237-248, 2015.

M. Frisén and J. De Maré. Optimal surveillance. Biometrika, 78:271-280, 1991.

References (continued)

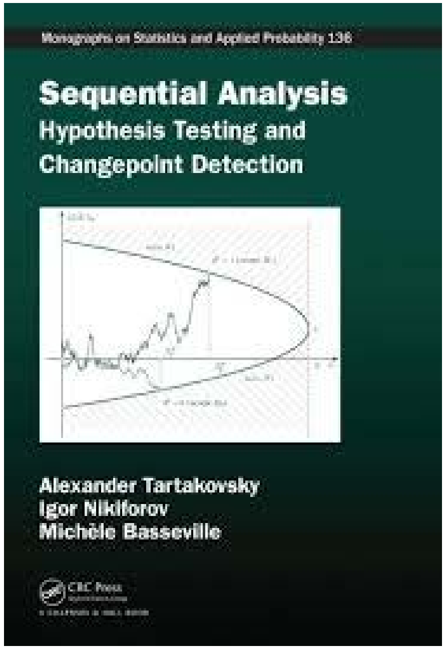
L. Horváth, M. Husková, P. Kokoszka and J. Steinebach, Monitoring changes in linear models, Journal of Statistical Planning and Inference 126 (2004), 225-251.

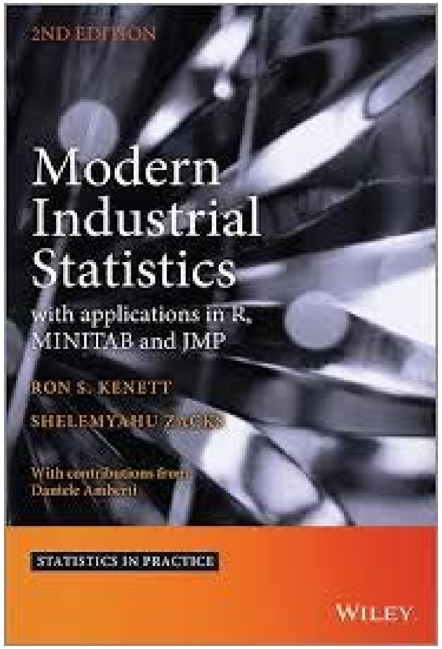
R.S. Kenett and M. Pollak, On Assessing the Performance of Sequential Procedures for Detecting a Change, Qual. Reliab. Engng. Int. 28:500-507, 2012.

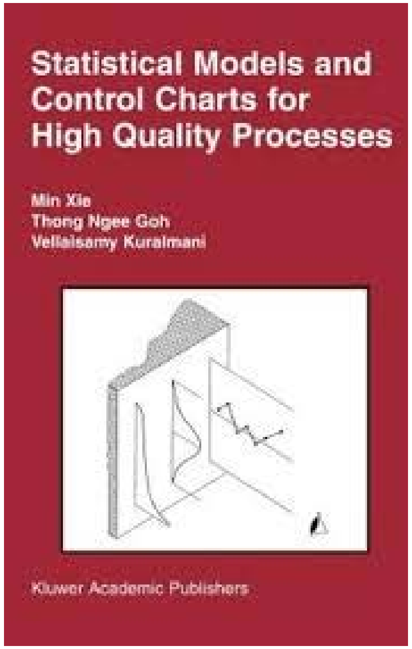
H. Theil, The analysis of disturbances in regression analysis. Journal of the American Statistical Association 60 (1965), 1067–79**.**

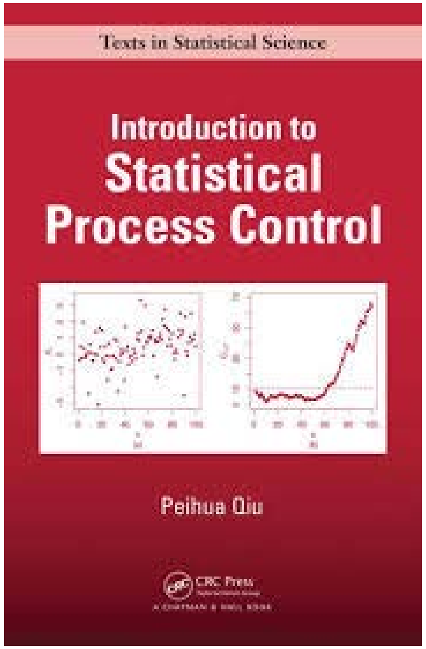
A. Zeileis, F. Leisch, C. Kleiber and K. Hornik, Monitoring structural changes in dynamic econometric models, Journal of Applied Econometrics 20 (2005), 99-121.

Books


Monographs on Statistics and Applied Probability 136
Sequential Analysis
Hypothesis Testing and Changepoint Detection
Alexander Tartakovsky
Igor Nikiforov
Michèle Basseville
CRC Press


2ND EDITION
Modern Industrial Statistics
with applications in R, MINITAB and JMP
RON S. KENETT
SHELEMYAHU ZACKS
With contributions from Daniele Amberti
STATISTICS IN PRACTICE
WILEY


Statistical Models and Control Charts for High Quality Processes
Min Xie
Thong Ngee Goh
Vellaisamy Kuralmani
Kluwer Academic Publishers


Texts in Statistical Science
Introduction to Statistical Process Control
Peihua Qiu
CRC Press

Future Directions

- network data
- high-dimensional data
- high-frequency data
- high volume data

Network Data

Research has just begun for network data (mainly social networks and computer networks).

Challenges:

- **model change scenarios (much more complicated than changes in the "mean")**
- **compute distribution of suitable statistics (e.g., degree/centrality measures)**
- **derive run length distributions under realistic network models (so that we know the performance)**
- **practical numerical evaluation of control charts**

High-Volume Data

- **hypothesis testing: everything is significant**
- **use general hypothesis (cf. equivalence testing)**
- **scalability of algorithms**