Statistical Modelling and Design for Quality Control and Reliability Analysis in Power Semiconductor Manufacturing Processes

Jürgen Pilz[1] Daniel Kurz[2] Natalie Vollert[3]

[1]Institut für Statistik, Universität Klagenfurt
Universitätsstr. 65-67, 9020 Klagenfurt, Austria
juergen.pilz@aau.at

[2]Infineon Austria, Villach
daniel.kurz@infineon.com

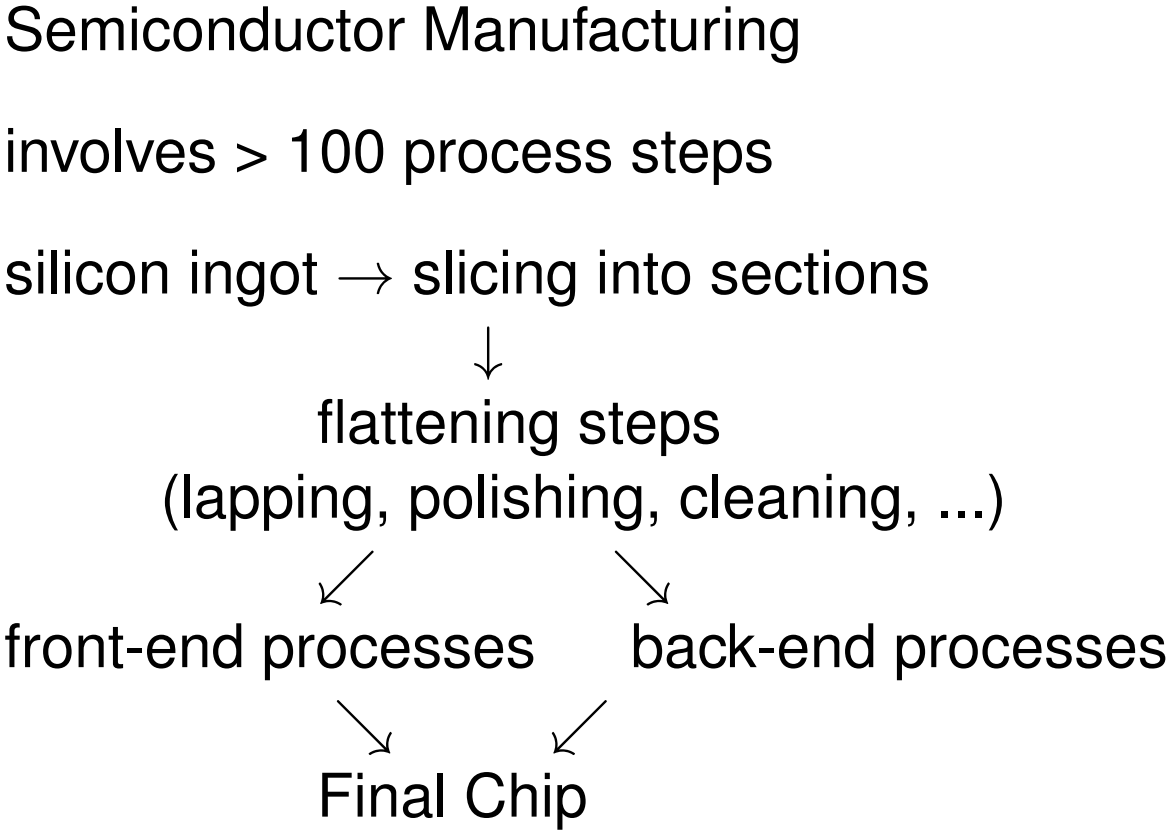
[3]Carinthian TechResearch AG, Villach
natalie.vollert@ctr.at

DAGSTAT
March 19-22, 2019 / Munich, Germany

Introduction

Semiconductor Manufacturing

involves > 100 process steps



Introduction

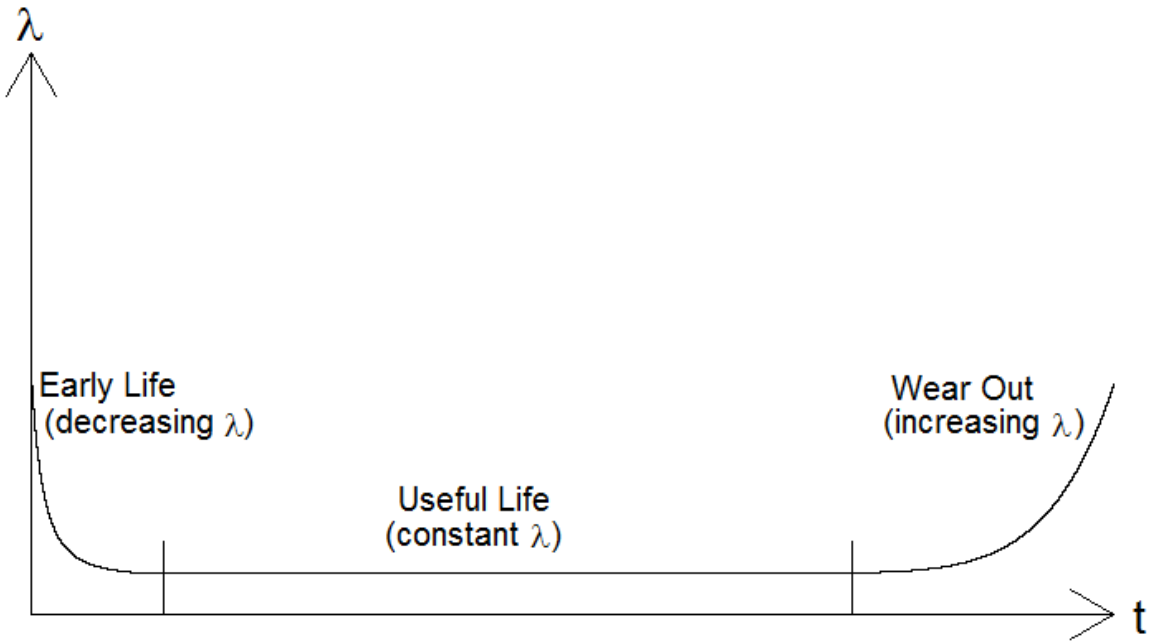
Key issue in semiconductor manufacturing:

Reliability

especially in safety-critical applications
Burn-in-study: most commonly applied failure screening technique.

Basis: bathtub curve describing hazard rate

Introduction



Introduction

Testing under accelerated stress conditions (increased temperature & voltage stress)

Burn-in: independently selected number of devices is investigated for early failures

Model for early failures: Weibull distribution $Wb(a, b), b < 1$.

Current ppm-requirement: 21ppm (Infineon)

Burn-in schemes different for logic and power devices. Here we focus on power devices.

Reasons for early failure: oxide particles, metallization defects.

Introduction

Problem: only very few failures

$\Rightarrow$ it's impossible to efficiently fit a Weibull DFR distribution to burn-in data.

Way out: prove that early life failure probability $p \in$ target confidence area

Burn-in read-outs at discrete time points t_1, t_2, t_3

Report statistics: $k_j = \#$ failures in $(t_{j-1}, t_j]$

$$j = 1, 2, 3;\ t_0 = 0$$

Introduction

Goal: P (early life failure after t_3 hours) $\leq 21 ppm$

Successful burn-in where $k = k_1 + k_2 + k_3 = 0$

(zero defect strategy)

Usually: Burn-in is re-started whenever a failure occurs

Current standard: introduction of countermeasures (CM) (ink out, design measures, optical inspection, ...) to reduce the failure probability p

Crucial: assessment of effectiveness of CM by experts.

Introduction

Our aim:

- development of a statistical model for taking account of CM's
- avoid re-start of burn-in by planning additional number of items to be burnt for zero defects.

Interval estimation for early life failure probabilities

n independently selected devices are stressed

$$X_i = \begin{cases} 0 & \text{if device } i \text{ passes the burn-in} \\ 1 & \text{if device } i \text{ fails within burn-in} \end{cases}$$

$$X = \sum_{i=1}^{n} X_i \sim Bi(n, p)$$

$$\boldsymbol{x} = (x_1, \ldots, x_n) \in \{0,1\}^n; \quad k = \boldsymbol{x}^T \boldsymbol{x} \in \{0, 1, \ldots, n\}$$

failures

Clopper-Pearson interval estimation

$$I_{CP} = (\hat{p}_l, \hat{p}_u) \qquad \text{where}$$

$$P(X \geq k|\hat{p}_l) = \alpha/2 \quad \text{and}$$

$$P(X \leq k|\hat{p}_u) = \alpha/2$$

To obtain $\hat{p}_l$ and $\hat{p}_u$, we use the well-known relationship with the Beta distribution

$\hat{p}_l = F_{Z_l}^{-1}(\alpha/2)$ with $Z_l \sim Be(k, n-k+1)$

$\hat{p}_u = F_{Z_u}^{-1}(1-\alpha/2)$ with $Z_u \sim Be(k+1, n-k)$

90% one-sided interval $I_p = [0, \hat{p}_u]; \quad \alpha = 0.1$

Bayesian equal-tail interval for p

In a Bayesian framework, this relationship comes in naturally observing that the conjugate prior for p is the Beta distribution:

$$p \sim Be(a, b); \quad a, b > 0$$

$$\Rightarrow f(p|\boldsymbol{x}) \propto l(p; \boldsymbol{x}) f(p) = p^{a+k-1}(1-p)^{b+n-k-1}$$

$$\text{i.e. } p|\boldsymbol{x} \sim Be(a^* = a + k, b^* = b + n - k)$$

Bayesian equal-tail interval for p

Bayesian equal-tail credible interval

$$C_e = (\hat{p}_l^*, \hat{p}_u^*) \text{ where } \hat{p}_l^* = F_{p|\boldsymbol{x}}^{-1}(\alpha/2), \hat{p}_u^* = F_{p|\boldsymbol{x}}^{-1}(1-\alpha/2)$$

Jeffreys' prior: $a = b = 1/2$

Choosing $a = 1, b = 0$ we have

$$p|\boldsymbol{x} = Be(k+1, n-k)$$

$$\hat{p}_u^* = \hat{p}_u$$

coincidence of one-sided Bayesian interval with Clopper-Pearson interval

Assessing *ppm*-levels using CM's

Repair is impossible for semiconductor devices; they either pass or fail within the burn-in. If a burn-in related failure occurs, then a CM is introduced (optical inspection, process improvement, ...) aiming to reduce p to $\pi \leq p$.

Experts assess the CM's effectiveness $\vartheta \in [0, 1]$ = probability for correcting the failure.

Single CM failure probability model

Consider k failures for which a single CM with effectiveness $\vartheta \in [0, 1]$ is implemented in the process

Interpretation: There is a likelihood ξ_j that $j \leq k$ failures would have occured or, equiv., $k - j$ failures would have been solved (corrected) if the CM would have already been introduced before the burn-in study.

Let $K_l = \begin{cases} 1 & \text{failure } l \text{ solved} \\ 0 & \text{else} \end{cases}$

Clearly: $K = \sum_{l=1}^{k} K_l \sim Bi(k, \vartheta)$

$\downarrow$

unknown number of failures that would have been caught by the CM

$\Rightarrow (*)\xi_j = P(K = k - j); \quad j \in \{0, \ldots, k\}$

Clopper-Pearson model for single CM

after the CM: $X' \sim Bi(n, \pi)$

Weighting of Clopper-Pearson upper limits according to $(*)$ leads to assessing $\hat{\pi}$ as

$$\sum_{j=0}^{k} \xi_j P(X' \leq j|\hat{\pi}) = \alpha$$

Equivalently: using $P(X' \leq j|\pi) = 1 - P(Z_j < \pi)$

$$\text{with } Z_j \sim Be(j+1, n-j); \quad j = 0, \ldots, k$$

$\Rightarrow \hat{\pi} = F_{Z'}^{-1}(1-\alpha)$ quantile of

$$Z' \sim \sum_{j=0}^{k} \xi_j Be(j+1, n-j) \text{ Beta mixture}$$

Multiple CM failure model

now consider $r \geq 1$ different CM's and denote $\boldsymbol{\vartheta} = (\vartheta_1, \ldots, \vartheta_r) =$ vector of effectivenesses; $r \leq k$

$\boldsymbol{k} = (k_1 \ldots, k_r)$; $k_i = \#$ failures tackled by CM$_i$

$$\text{with } \sum_{i=1}^{r} k_i = k$$

Now: $K = \sum_{l=1}^{k} K_l \sim GBi(k, \boldsymbol{\vartheta}_k)$ generalized binomial, where

$$\boldsymbol{\vartheta}_k = (\underbrace{\vartheta_1, \ldots, \vartheta_1}_{k_1 \text{ times}}, \underbrace{\vartheta_2, \ldots, \vartheta_2}_{k_2 \text{ times}}, \ldots, \underbrace{\vartheta_r, \ldots, \vartheta_r}_{k_r \text{ times}})$$

We have developed an **efficient method** for computing generalized binomial probabilities employing sequential convolution.

CM's with uncertain effectivenesses

So far: $\vartheta_i; \quad i = 1, \ldots, r$; were fixed

often: process experts are uncertain about the effectivenesses of the applied CM's.

a) Beta-Binomial model for a single uncertain effectiveness, $r = 1$

$\vartheta \sim Be(u, v)$

$K|\vartheta \sim Bi(k, \vartheta)$

$$\Rightarrow \xi_j = P(K = k - j) = \int_0^1 P(K = k - j|\vartheta) f(\vartheta) d\vartheta$$

$$= \binom{k}{k-j} \frac{\Gamma(u+k-j)\Gamma(v+j)}{\Gamma(u+k+v)} \frac{\Gamma(u+v)}{\Gamma(u)\Gamma(v)}$$

$K \sim BeBi(k, u, v)$ Beta-Binomial

CM's with uncertain effectivenesses

b) Generalized Beta-Binomial model for more than a single uncertain effectiveness

$K|\vartheta \sim GBi(k, \vartheta_k)$

$\vartheta_i \sim Be(u_i, v_i); \quad i = 1, \ldots, r$

$$\Rightarrow P(K = k - j) = \int_{[0,1]^r} P(K = k - j|\vartheta) f(\vartheta) d\vartheta$$

$K \sim GBeBi(k, u_1, \ldots, u_r, v_1, \ldots, v_r, k_1, \ldots, k_r)$ no closed form solution, MC-integration

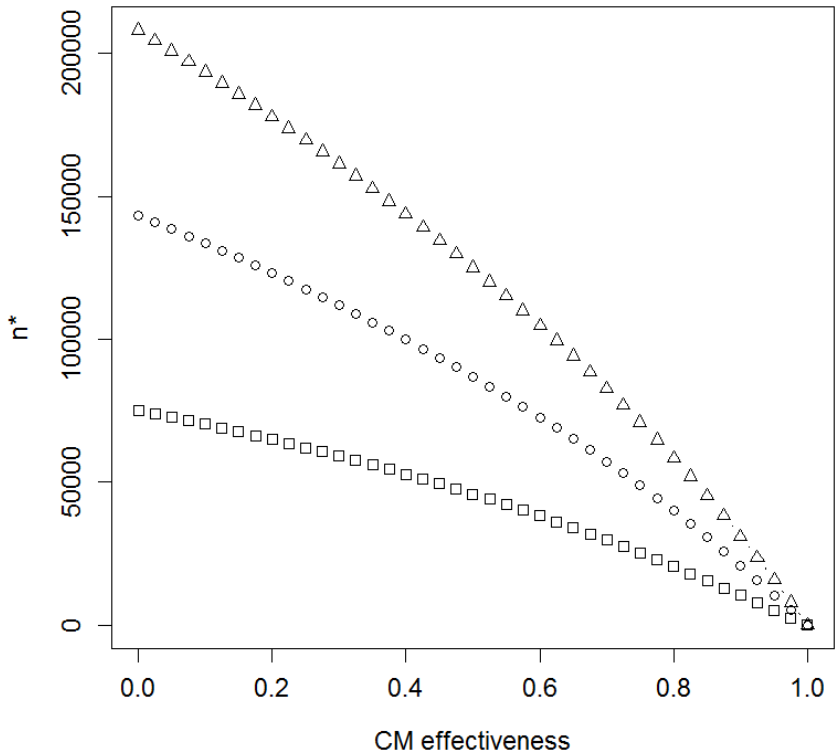
Rationale:

Take $n^* < n$ (even $n^* \ll n$) additional devices and prove that the target $ppm-$level is still guaranteed on the basis of the CM failure probability model

Efficiency of the new approach: illustration for single CM case (different degrees of effectiveness) and $k = 1, 2, 3$.

Significant reduction of n^* in case of high CM effectiveness!

Rationale:



Required inspections in addition to $n = 110000$ for reaching target $p \leq 21$ ppm at 90% CL with implemented CM(s) and $k = 3, 2, 1$ fails.

Bayesian assessment of Weibull early life failure distributions

Burn-in settings (read-outs, burn-in time, ...) are typically assessed using a Weibull DFR distribution $Wb(a, b)$ with

scale $a > 0$ and **shape** $b \in (0, 1)$

crucial point: joint prior $p(a, b)$

- There is no continuous conjugate joint prior
- Conjugate continuous-discrete joint prior: Gamma dist. for a, categorical distr. for b (Soland 1969)
- Jeffreys' prior: $p_J(a, b) \propto 1/ab$ (Sinha 1986)

Bayesian assessment of Weibull early life failure distributions

We propose two alternatives:

- Histogram prior (remains still challenging)
- Dirichlet prior

Let $T \sim Wb(a, b)$ with density

$$f(t|a,b) \propto \begin{cases} t^{b-1} \exp(-(\frac{t}{a})^b) & t > 0 \\ 0 & \text{else} \end{cases}$$

with $a > 0, 0 < b < 1$

Burn-in read outs at fixed time points $t_1^*, \ldots, t_m^* > 0$

Example

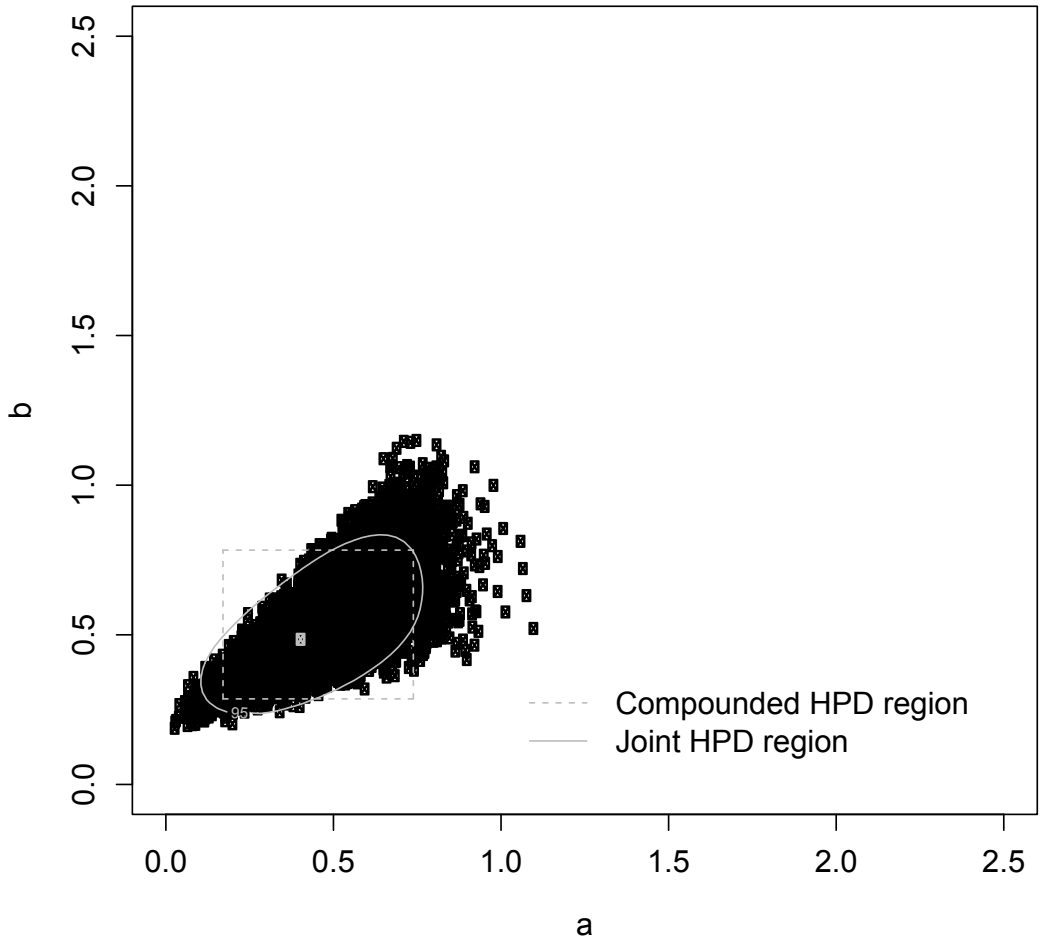
Data: $\boldsymbol{k} = (k_1 = 20, k_2 = 2, k_3 = 1, k_4 = 7)^T$

$k_4 = 7$ failures not detected within $t_3^* = 4$ hours (burn-in time not adequate)

$\Rightarrow$ posterior: $\boldsymbol{p}|\boldsymbol{k} \sim Dir(101, 15, 6, 8)$

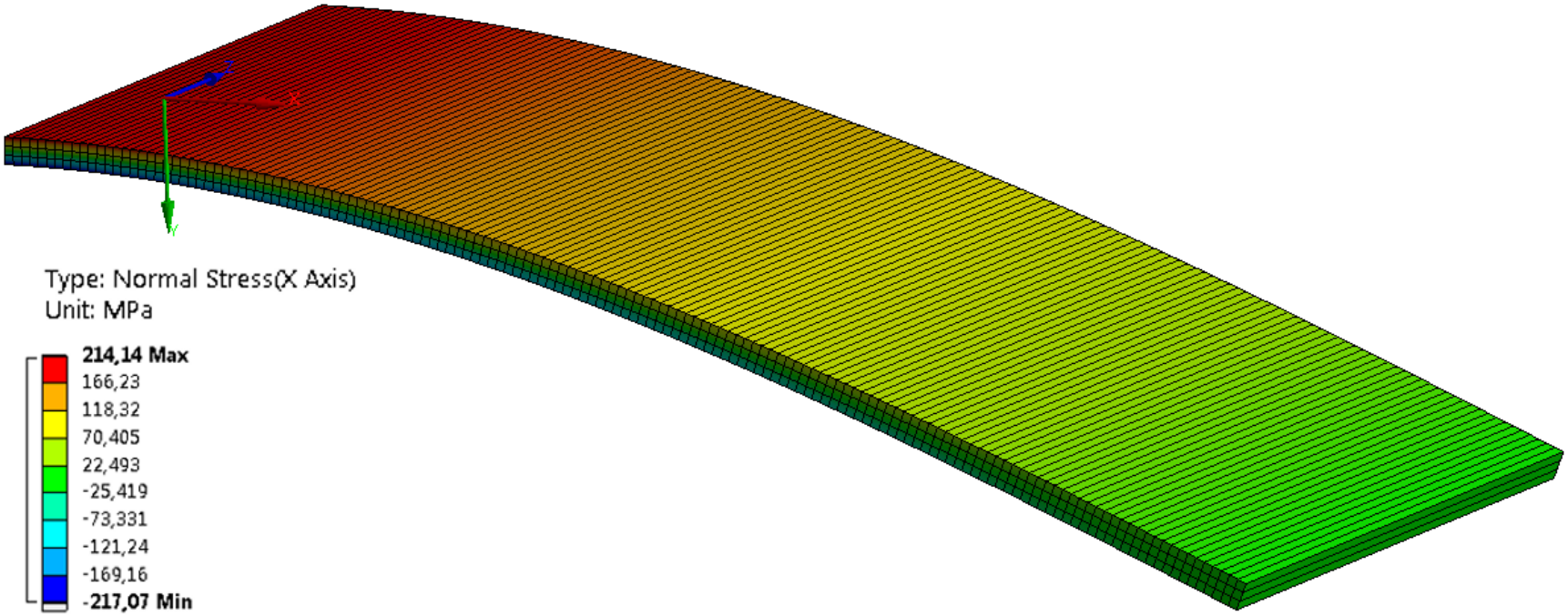
and $Wb(0.5, 0.75)$ shifted to $Wb(\hat{a}^* = 0.409, \hat{b}^* = 0.485)$

Example



Part II: Gaussian Process Regression

Stress testing in semiconductor processing for **thin wafers** (thickness$\leq 40\mu m$)
Kriging metamodel for stress prediction validated against Ramann spectroscopy measurements, FEM simulations
+ Modelling of electrical parameters (signals)



- **Experiments**
  - physical experiments
  - Computer experiments (Computer-based simulations like FEM)

Which simulations to run?
Main difference: Computer Models are **deterministic**
Modification of classical DOE $\Rightarrow$ DOCE

Math. model: $y = \tilde{f}(x_1 \ldots, x_k)$, e.g. solution of ODE/PDE system

$$\mathbf{x} = (x_1, \ldots, x_k)^T \in \mathcal{X} = \text{experimental domain}$$

replaced by *meta-model*

$$\mathbb{E}\, Y(\mathbf{x}) = f(x_1, \ldots, x_k), f \text{ "close" to } \tilde{f}$$

Requirements for good designs:

- space filling property
- projective property
- computational efficiency

Compromise: **LHD**= Latin Hypercube Designs

w.l.o.g. experimental domain $\mathcal{X} = [0, 1]^k$

Designs: $d_n = (\mathbf{x}_1, \ldots, \mathbf{x}_n) \in \mathcal{X}^n$,

n runs, k factors

Part II: Gaussian Process Regression

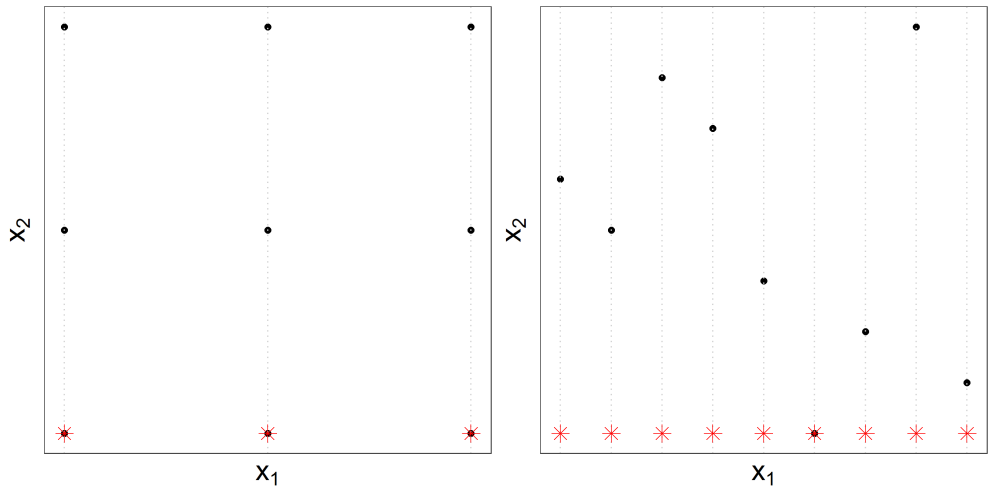


Figure: Regular (left) and latin hypercube design (right)

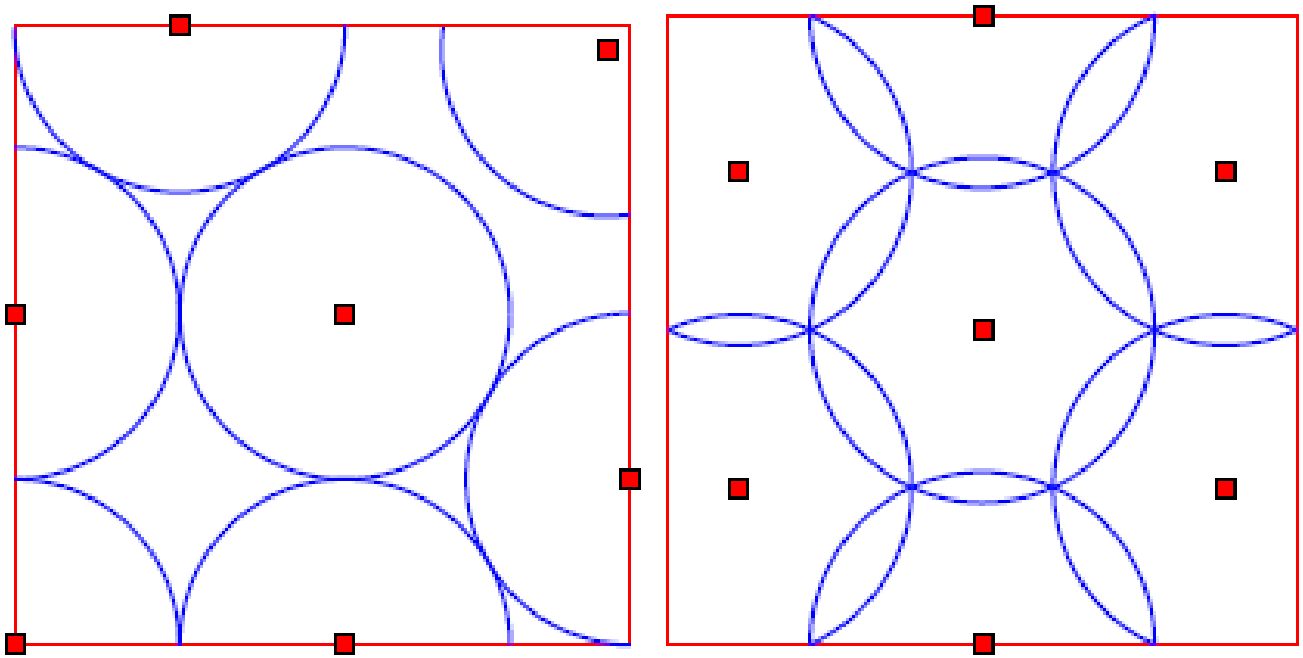
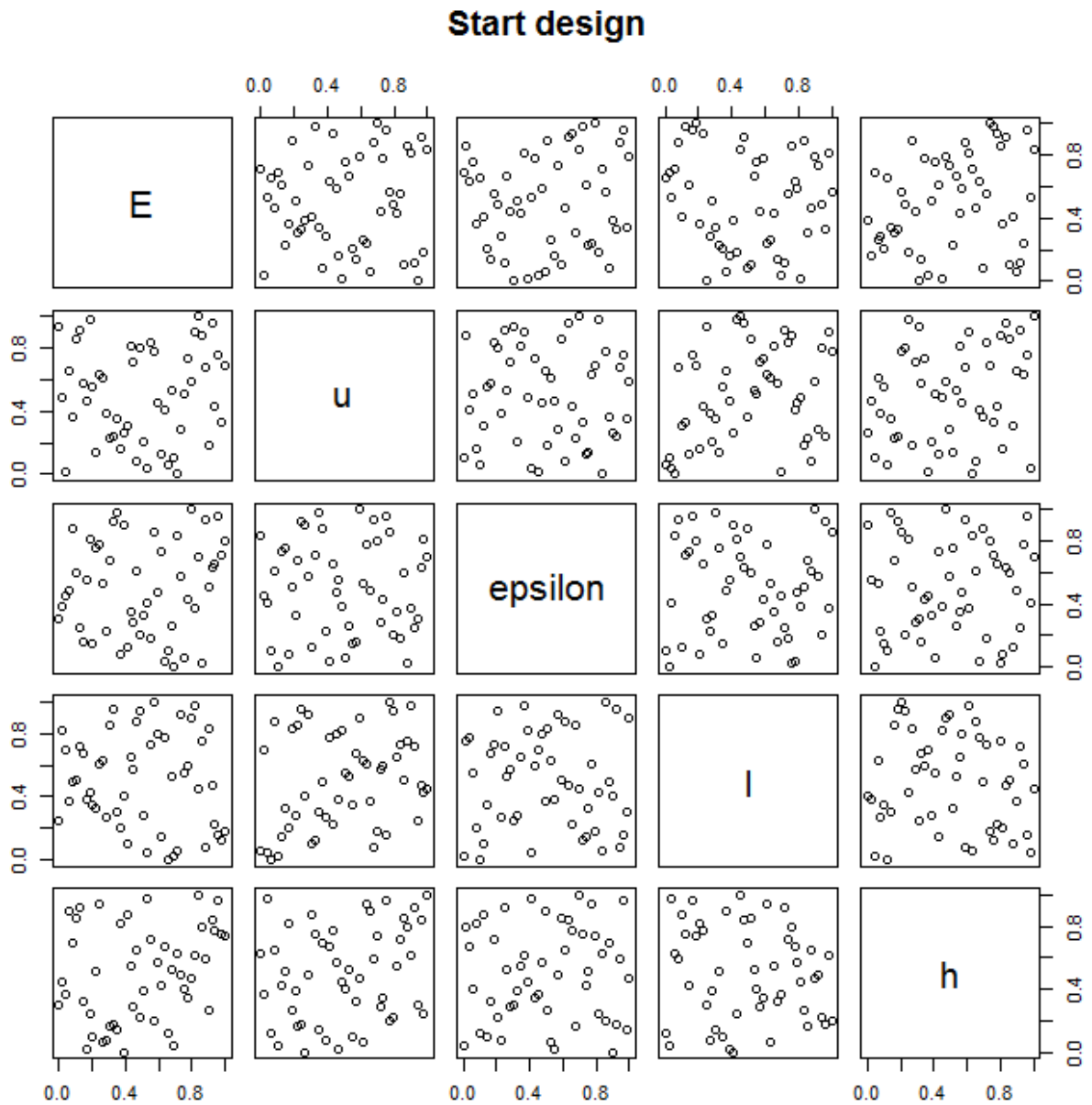
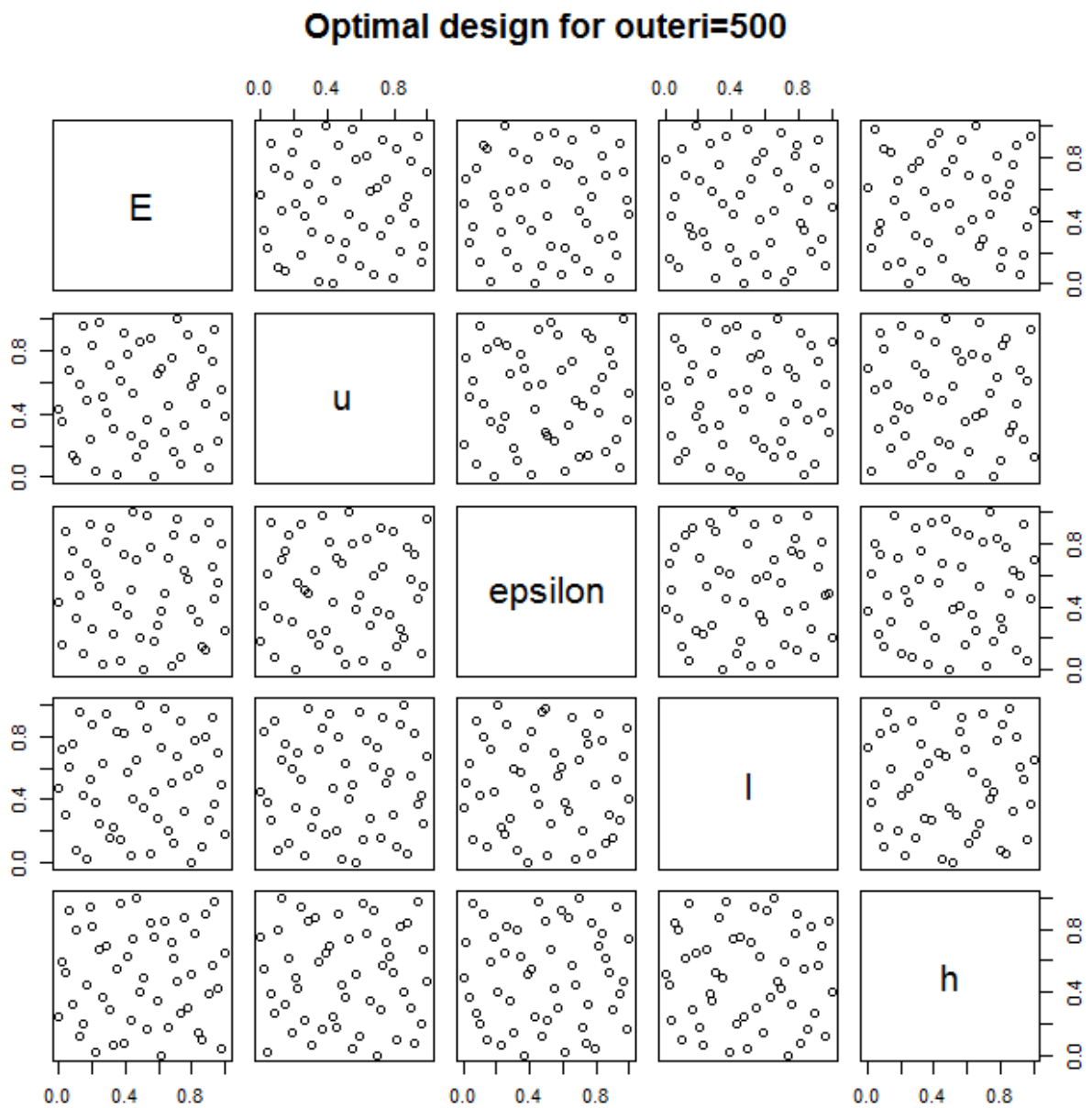


Figure: Maximin (left) and Minimax (right) designs

Part II: Gaussian Process Regression



Part II: Gaussian Process Regression



RSM modelling

Classical approach: Regression (response surface) modelling $\Rightarrow$

prediction reduces to interpolation problem

e.g. quadratic RSM

$$y(\mathbf{x}) = \beta_0 + \sum_{i=1}^{k} \beta_{ii} x_i^2 + \sum_{i<j}^{k} \sum_{j=1}^{k} \beta_{ij} x_i x_j$$

For complex responses, LSE $\hat{\beta} = (X^T X)^{-1} X^T \mathbf{y}_d$

yields bad interpolations

Remedy

Kriging models

$\Rightarrow$ response = realization of stochastic process

$$Y(\mathbf{x}) = \underbrace{\mu(\mathbf{x})}_{\text{trend}} + \underbrace{Z(\mathbf{x})}_{\substack{\text{Gaussian Process} \\ \text{(zero mean)}}}$$

Effect: good approx. over a wide range of different designs and sample sizes and well-defined basis for statistical framework

$$Y(\cdot) \sim GP(\mu(\mathbf{x}), \sigma^2 R(\cdot, \cdot))$$

Main difference to geostatistical settings:
– $\mathbf{x}$ is no spatial coordinate vector
– usually, higher dimensional settings: $K > 3$

Covariance structure

covariance function: $Cov(Z(\mathbf{x}_i), Z(\mathbf{x}_j)) = \sigma^2 \underbrace{R(\mathbf{x}_i, \mathbf{x}_j)}_{\text{correlation function}}$

$\downarrow$

process variance

Common assumptions:

1. covariance-stationarity, i.e.

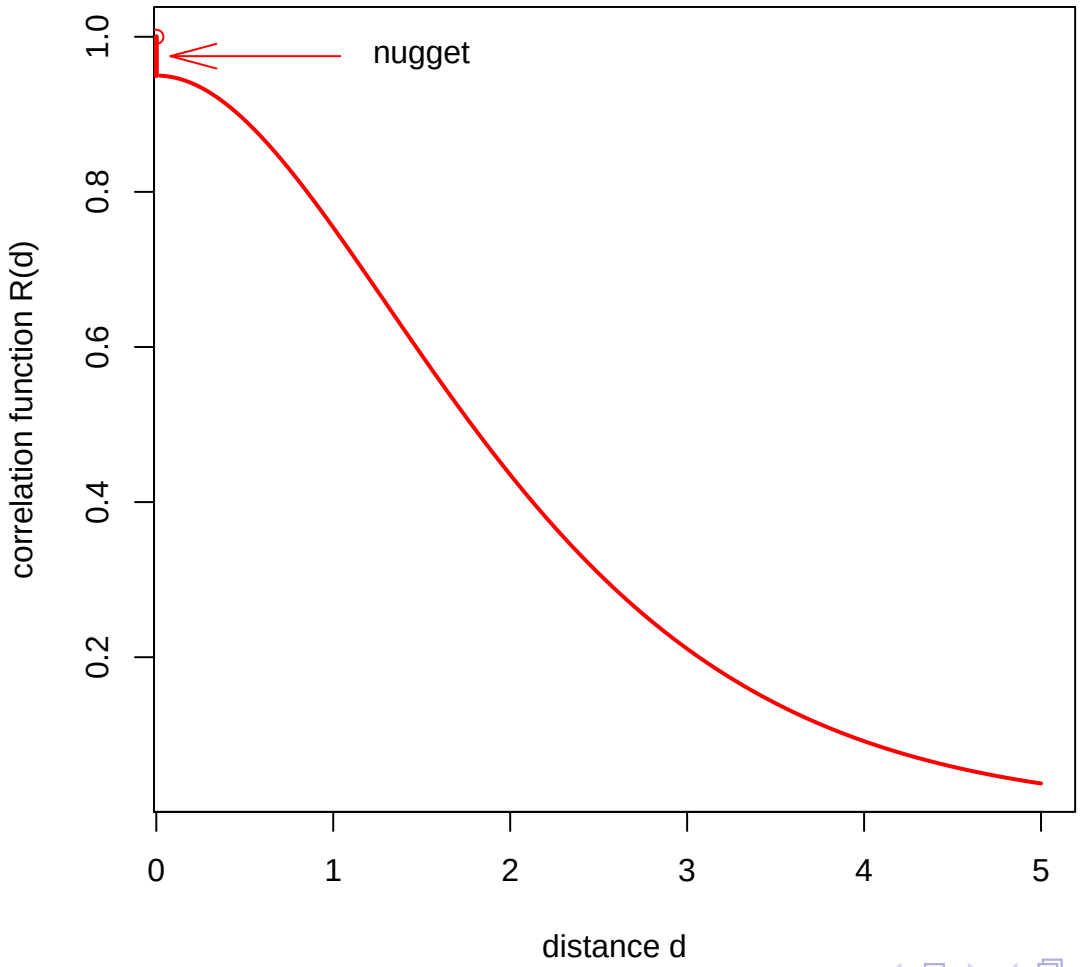
$$R(\mathbf{x}_i, \mathbf{x}_j) = R(\mathbf{x}_i - \mathbf{x}_j)$$

2. (tensor-)product correlation structure

$$R(\mathbf{x}_i, \mathbf{x}_j) = \prod_{m=1}^{k} \underbrace{R_m(|x_{im} - x_{jm}|)}_{\text{univariate Matérn c.f.}}$$

Matérn c.f. $\nu = \frac{5}{2}$

$$R(d) = \left(1 - \frac{\tau^2}{\sigma^2}\right) * \left(1 + \frac{\sqrt{5}d}{\theta} + \frac{5d^2}{3\theta^2}\right) \exp\left(-\frac{\sqrt{5}d}{\theta}\right), \; d > 0$$



Estimation and Prediction

MLE: available for β and σ^2

$\hat{\beta} = (X^T R_n^{-1}(\theta) X)^{-1} X^T R_n^{-1}(\theta) \mathbf{y}_d$

$\hat{\sigma}^2 = \frac{1}{n}(\mathbf{y}_d - X\hat{\beta})^T R_n^{-1}(\theta)(\mathbf{y}_d - X\hat{\beta})$

$\hat{\boldsymbol{\theta}} = (\hat{\theta}_1, \ldots, \hat{\theta}_k)^T$ Gauss-Newton (or genetic optimiz.)

Optimal prediction:

$$\begin{aligned} \hat{Y}(\mathbf{x}^*) &= \mathbf{f}(x^*)^T \hat{\beta} + \mathbf{r}_0^T R_n^{-1}(\mathbf{y}_d - X\hat{\beta}) \\ &= \text{GLSE + smoothed residual} \end{aligned}$$

where $\mathbf{r}_0^T = (R(\mathbf{x}_0 - \mathbf{x}_1), \ldots, R(\mathbf{x}_0 - \mathbf{x}_n))$, $R_n =$ correl. matrix

Implementation in R: **DiceKriging**

Robust Additive Gaussian Processes

Aims

- higher flexibility in meta-modelling
- numerical stability: robustness of parameter estimates, esp. for correlation parameters

Solution: Bayesian approach using additive models and (objective) reference priors

Side effect: high-dimensional optimization problems reduced to a few sub-routines of ≤ 3 dimensions

Additive model:

$$\mathbb{E}\,Y(\mathbf{x}) = f_0 + \sum_{i=1}^{k} f_i(x_i) + \sum_{i<j} f_{ij}(x_i, x_j) + \ldots + f_{12\ldots k}(x_{1,\ldots}, x_k)$$

Functional ANOVA Representation

Robust Additive Gaussian Processes

Special case: first order GAM

$$\mathbb{E}\, Y(x_1, \ldots, x_k) = f_0 + \sum_{i=1}^{k} f_i(x_i)$$

$f_1, ..., f_k$: smooth basis functions

$\Rightarrow$ non-parametric modelling of main effects

Goal: Extension of classical GAM regression

Fo a good overview of the advantages of additive structures compared to fully parametric GP models in high dimensions see Dourante, Ginsbourger, Roustant (2012)

Robust Additive Gaussian Processes

Novelty of our recently proposed concept: Combination of AGP with robust reference priors proposed by Gu, Wang and Berger (submitted to AS 2017) + new sampling design scheme

Our new model: Second order Kriging AGP with

$$f_i \sim N(\mu_i, \sigma^2 R_i)$$

$$f_{ij} \sim N(\mu_{ij}, \sigma^2 R_i R_j)$$

Result: AGP $Y(\mathbf{x}) \sim N(\mu, \sigma^2 R(\cdot,\cdot))$, locally constant trend

and $R(\mathbf{x}, \mathbf{x}') = \sum_{i=1}^{k} R_i(x_i, x_i') + \sum_{i=1}^{k} \sum_{j=i+1}^{k} R_i(x_i, x_i') R_j(x_j, x_j') + \delta_{\mathbf{x}\mathbf{x}'} \tau^2$

Robust Additive Gaussian Processes

Renormalization such that

$f^* := \frac{1}{\sqrt{m}} f \sim N\left(\frac{1}{\sqrt{m}}\mu, \sigma^2 R^*\right)$

$R^* := \frac{1}{m} R$ valid correlation matrix, $m = \#$ correlation components

Each component function has parameters

$\Theta_i = (\mu, \sigma^2, \theta_i, \tau_i^2)$ for 1st order terms

$\Theta_{ij} = (\mu, \sigma^2, \theta_i, \theta_j, \tau_{ij}^2)$ for 2nd order terms

Profile likelihood approach often fails, results in estimates $\hat{\theta}$ for which

$(*)$ $\quad R \approx I_n \quad (\hat{\theta} \approx 0)$ $\quad \text{or} \quad R \approx \mathbf{1}_n \mathbf{1}_n{}^T$ singular corr.m.

$\downarrow$

bad ("impulse") prediction

Robust Additive Gaussian Processes

Remedy: robust Bayes prediction using reference priors of the form

$$\pi^R(\mu, \sigma^2, \theta^*) = \frac{\pi^R(\theta^*)}{\sigma^2}$$

$\downarrow$

correl. parameters

where $\pi^R(\theta^*) \propto (det \;\; I_F(\theta^*))^{1/2}$

$\downarrow$

exp. Fisher information

Explicit representations for $I_F(\theta^*)$ available in Kazianka & Pilz (2012)

Result: proper posteriors $p(\theta^*|\mathbf{y}_d)$

Simplified estimate: $\hat{\theta}^* = arg \;\; \max_{\theta^*} \;\; p(\theta^*|\mathbf{y}_d)$

posterior mode (to avoid MCMC)

Bayes Prediction

Bayes predictor of $Y(\mathbf{x}_0)$ for untried input $\mathbf{x}_0$ is based on the predictive distribution

$$p(Y_0|\mathbf{y}_d) = \int \underbrace{p(Y_0|\mathbf{y}_d, \theta^*)}_{Student-t} p(\theta^*|\mathbf{y}_d) d\theta^*$$

Simplification: Use plug-in predictor

$$\begin{aligned} \mu^* &:= E(Y_0|\mathbf{y}_d, \hat{\theta}^*) \\ &= \hat{\mu} + \mathbf{r}_0^T R_{\hat{\theta}^*}^{-1} (\mathbf{y}_d - \hat{\mu}\mathbf{1}_n) \\ \text{where } \hat{\mu} &= (\mathbf{1}_n{}^T R_{\hat{\theta}^*}^{-1} \mathbf{1}_n)^{-1} \mathbf{1}_n{}^T R_{\hat{\theta}^*}^{-1} \mathbf{y}_d \text{ GLSE} \end{aligned}$$

R-implementation fully described in Vollert, Ortner & Pilz (2019) is based on an iterative estimation scheme, using reparametrizations

Note: increasing nugget with increasing dimension k of input space

Sampling design

Due to additive structure, space-filling is important (for all variable projections)

Need **compromise** between LHD and regular grid designs: **Cut-FD** combines *HDMR Designs* based on a cut-center with *Factorial Designs*

$\Rightarrow n_0 = 2^k + 2k$ boundary points $+1$ cut-point

$< 10 \cdot k = n^*$ (recommended `min.size` for DOCE)

whenever $k \leq 5$

Add $(n^* - n_0)$ points along (ij)-planes of cut point $\mathbf{x}_c$

For $k > 5$ we recommend to use *Fractional Factorial Designs* instead of Full Factorials.

Example

3 commonly used test functions

Pepelyshev function: $x_i \in [0,1]; i = 1,2,3$

$$f_1(\mathbf{x}) = 4(x_1 - 2 + 8x_2 - 8x_2^2)^2 + (3 - 4x_2)^2 + 16\sqrt{x_3 + 1}(2x_3 - 1)^2$$

$n_1 = 31$ samples

Park function: $x_i \in [0,1]; i = 1,\ldots,4$

$$f_2(\mathbf{x}) = \frac{2}{3}\exp(x_1 + x_2) - x_4 \sin(x_3) + x_3$$

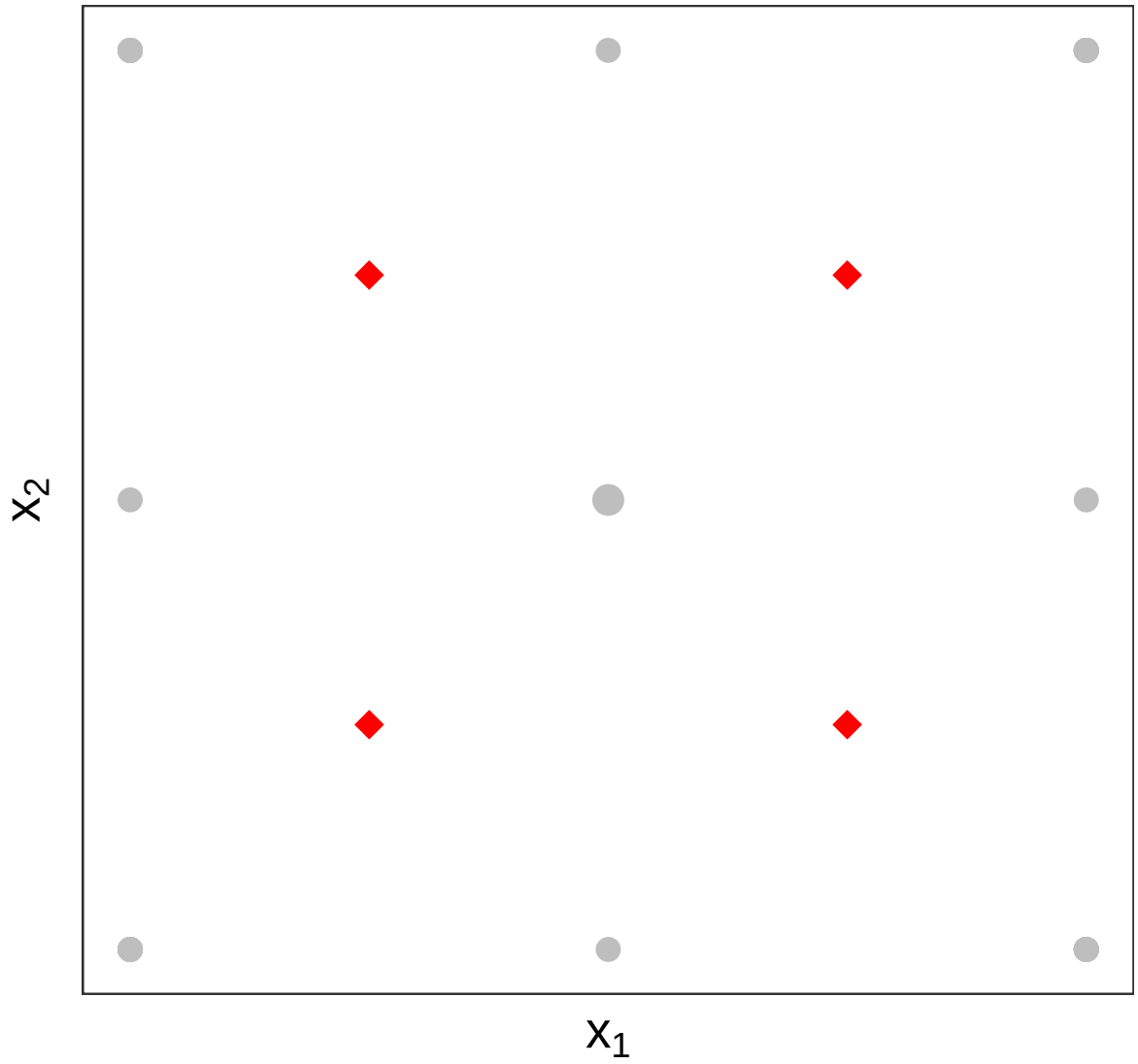
$n_2 = 41$ samples

Friedman function: $x_i \in [0,1]; i = 1,\ldots,5$

$$f_3(\mathbf{x}) = 10\sin(x_1 x_2) + 20(x_3 - 0.5)^2 + 10x_4 + 5x_5$$

$n_3 = 47$ samples

Example



Implementation

All calculations in R,
lhs package for constructing a maximin LHD,
DiceKriging package for constructing GP models

Setup:

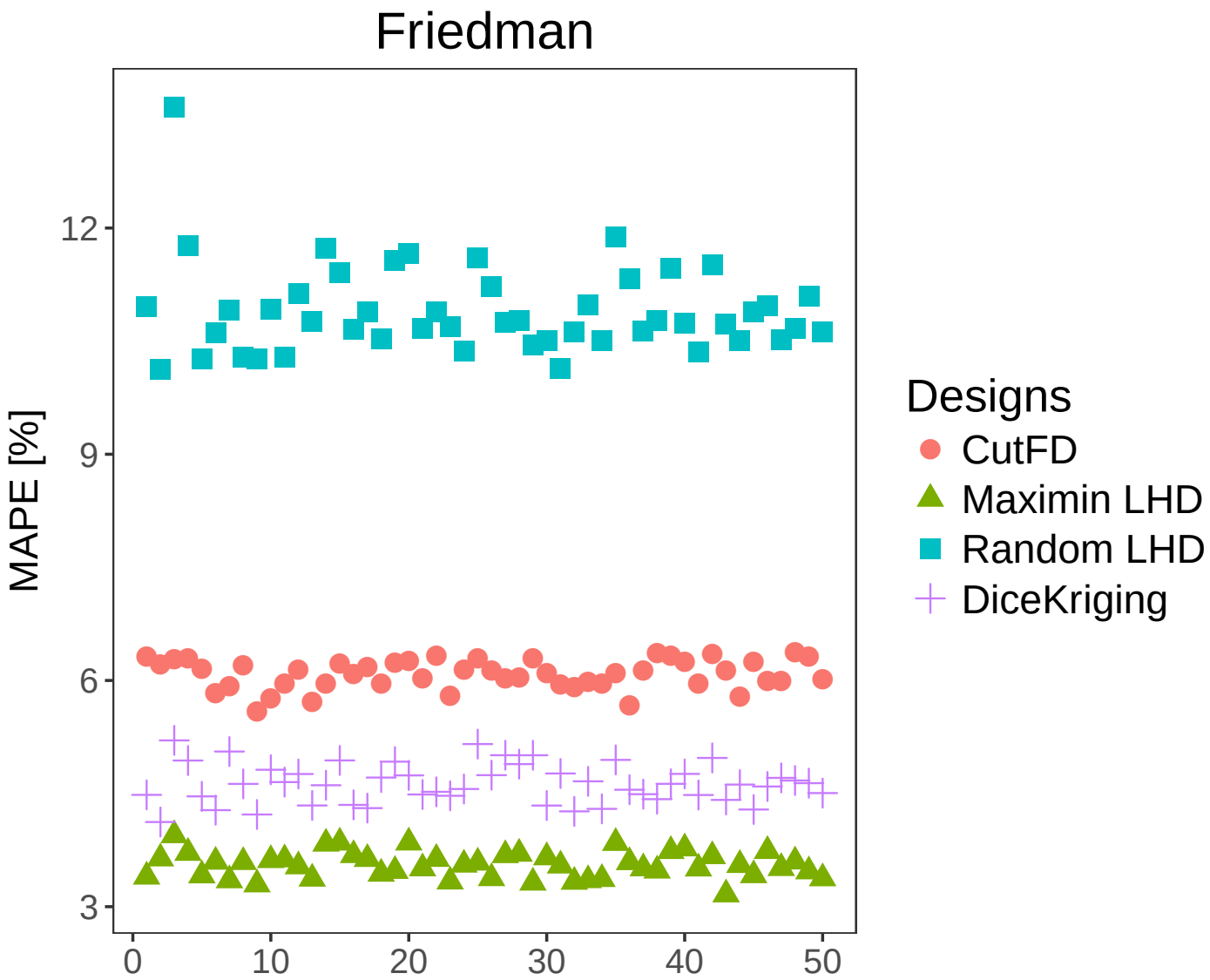
- Matérn correlation with $\nu = 5/2$ for all components
- comparison for 3 designs: random LHD, maximin LHD and Cut-FD

Criterion for comparisons: MAPE = mean absolute prediction error, measured (in %) at 25000 positions (generated by 50 random designs each containing 500 points)

Results

1. Cut-FD can better determine the actual structure of the test functions than maximin and other LHDs (found exact set of components for Pepelyshev and Friedman functions, maximin did not)
2. Maximin LHD design was best with regard to MAPE (pred. power): For 5D-Friedman function MAPE $< 4\%$ (based on only $n_3 = 47$ sample points!)
3. **Robust AGP** model **outperforms** commonly used **GP** models for all three test functions
4. Simple random LHDs are least appropriate for approximation, getting even worse with increasing dimension

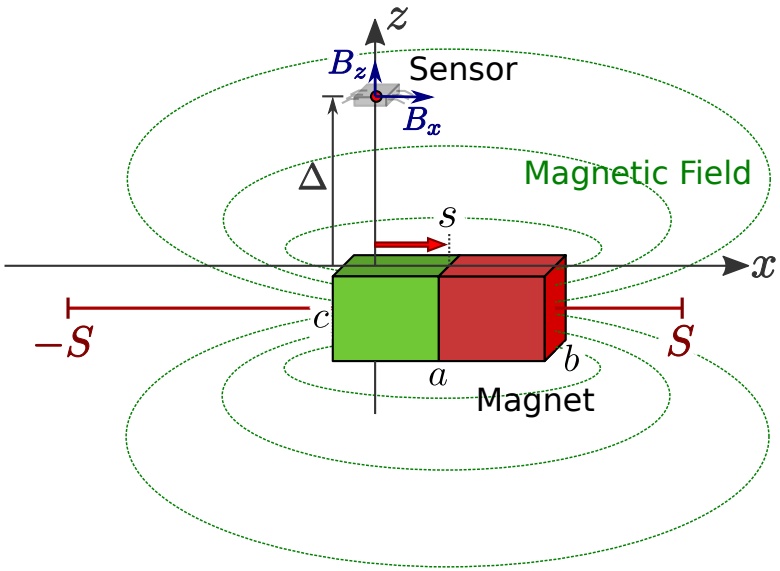
Results


Friedman
MAPE [%]
12
9
6
3
0
10
20
30
40
50
Designs
CutFD
Maximin LHD
Random LHD
DiceKriging

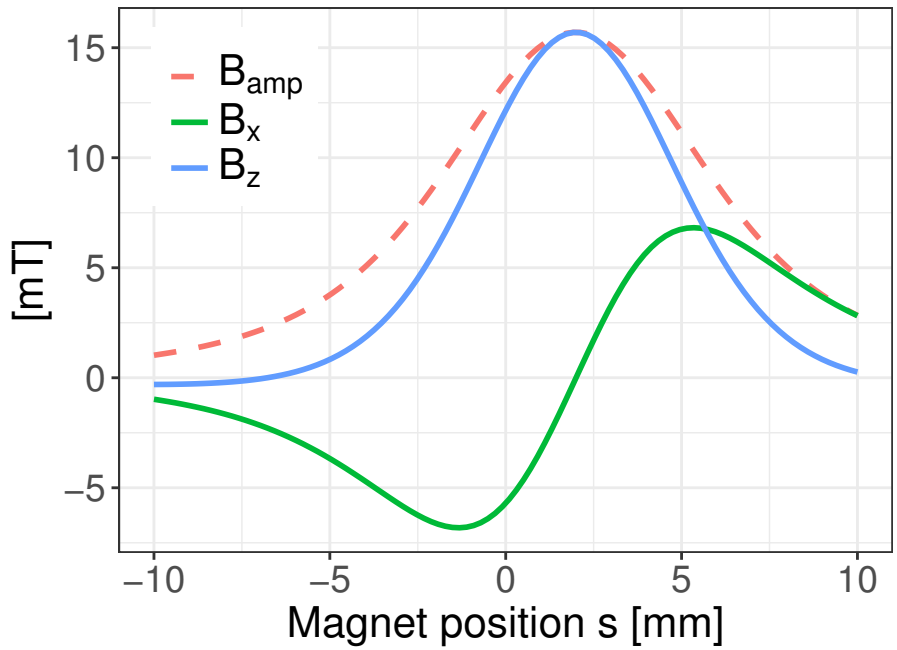
Work in progress:

AGP modelling for real DOCE applications based on FEMs for geometric and material parameter optimization problems, e.g. Magnetic field shaping for position and orientation detection systems

a)



b)

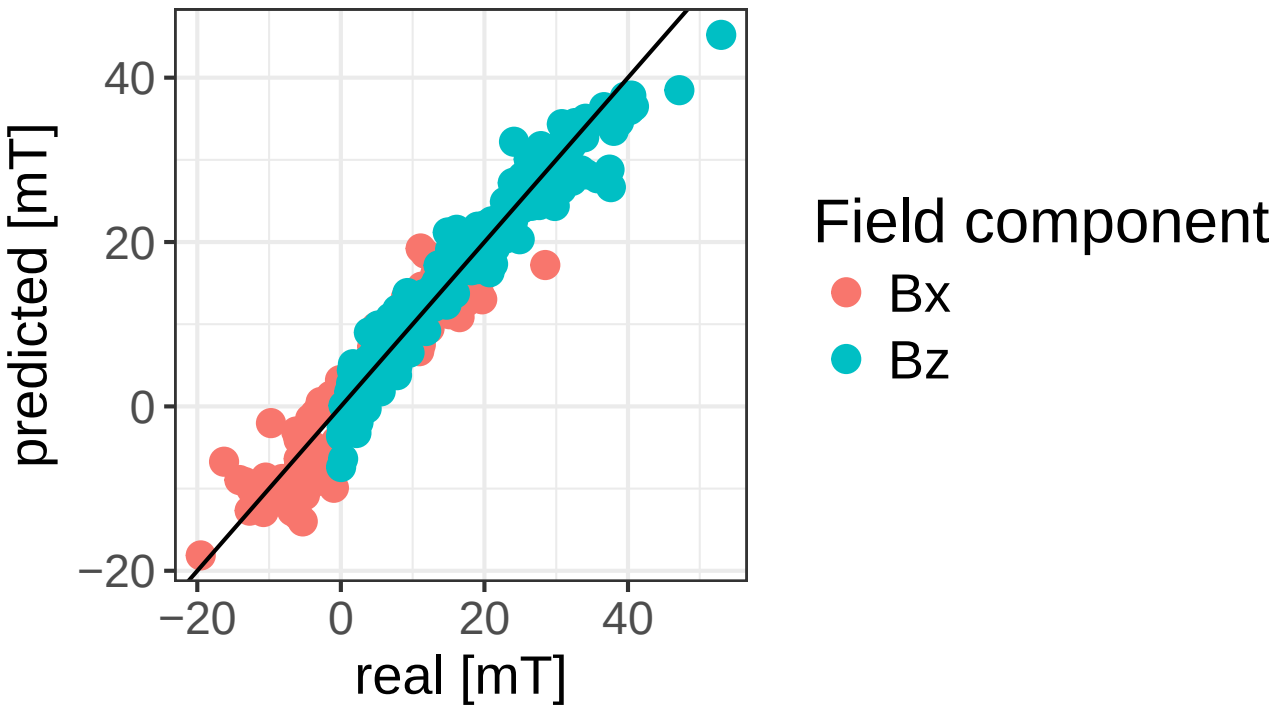


Model Performance based on random LHD

Component functions chosen by our algorithm:

– for B_x : f_s, f_b, f_{cs}, f_{as}

– for B_z : $f_s, f_a, f_b, f_c, f_{as}, f_{bM}, f_{ab}, f_{bs}, f_{cs}$



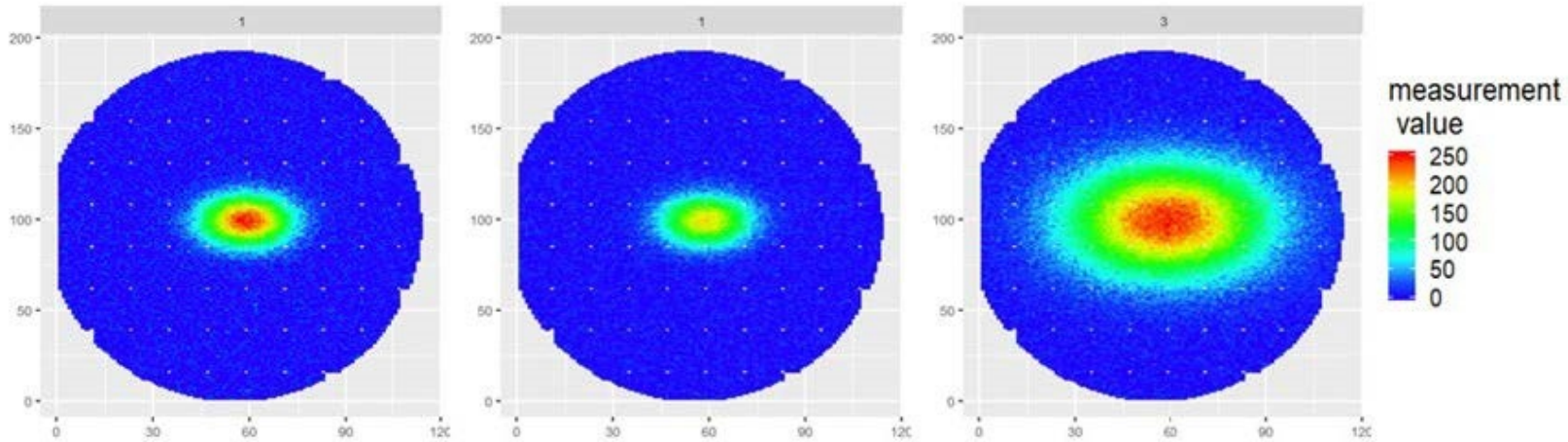
Current and Future Work with Infineon

HORIZON 2020 & ECSEL Joint Undertaking Project „iDev40“:
Integrated Development 4.0 for Semiconductor Manufacturing

Project start: May 2018, End date: April 2021
Our project part will focus on Statistical Root-Cause Failure Analysis and Advanced Statistical and Bayesian Deep Learning Methodology.
The project includes industrial partners Infineon Austria/Germany/Rom., AVL, ELMOS, Yazzoom, ETC, IBERMATICA, . . . and quite a few European university research partners

Current and Future Work with Infineon

Spatiotemporal Random Field Modeling and Bayesian Deep Learning Methods for Signal/Image Processing of wafer maps



Extension of AGP to **Nearest Neighbour AGP**

References

D. Kurz, H. Lewitschnig and J. Pilz: Bayesian Assessment of Weibull Early Life Failure Distributions. Int. J. Quality and Reliability Engineering 30 (2014) 3, 363-373

D. Kurz, H. Lewitschnig and J. Pilz: Decision- theoretical model for failures which are tackled by countermeasures. IEEE Transactions on Reliability 63 (2014) 2, 583 – 592 DOI:10.1109/TR.2014.2315952

D. Kurz, H. Lewitschnig and J. Pilz: Failure Probability Estimation For Differently Sized Products in Semiconductor Manufacturing. Applied Stochastic Models in Business and Industry 31 (2015) 5, 732-744

D. Kurz, H. Lewitschnig and J. Pilz: An advanced area scaling approach for semiconductor burn-in. Microelectronics Reliability 55(2015)1, 129-137, DOI: 10.1016/ j.microrel.2014. 09.007

References

D. Kurz, H. Lewitschnig and J. Pilz: Modeling of chip synergies for failure probability estimation in semiconductor manufacturing. J. Applied Statistics 44 (2017) 6, 955-967 DOI: 10.1080/02664763.2016.1189522

O. Roustant, D. Ginsbourger, Y. Deville, DiceKriging, DiceOptim: Two R packages for the analysis of computer experiments by kriging-based metamodeling and optimization, Journal of Statistical Software 51 (2012) 1, 1–55

H. Kazianka, J. Pilz, Objective Bayesian analysis of spatial data with uncertain nugget and range parameters, The Canadian Journal of Statistics 40 (2012), 304–327

M. Gu, X. Wang, J. O. Berger, Robust Gaussian stochastic process emulation, submitted to the Annals of Statistics (2018)

References

N. Vollert, M. Ortner and J. Pilz, Robust Additive Gaussian Process Models Using Reference Priors and Cut-Off-Designs, J. Applied Mathematical Modelling 65 (2019), 586-596

K. Posch, J. Steinbrener and J. Pilz :Variational Inference to Measure Model Uncertainty in Deep Neural Networks. Submitted to: International Journal of Computer Vision, Springer, May 2018

K. Posch, M. Arbeiter and J. Pilz. A novel Bayesian approach for variable selection in linear regression models. Submitted to: Computational Statistics and Data Analysis, February 2019