SimTech
Cluster of Excellence

Universität Stuttgart
Germany

Recent advances in optimizing field campaigns

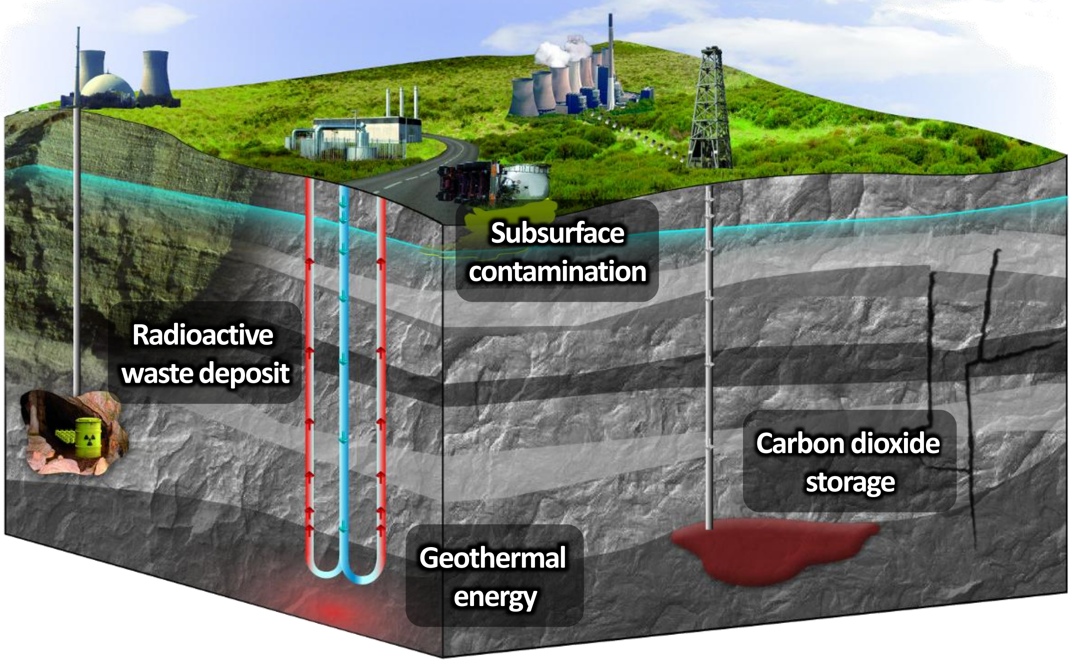
Jun.-Prof. Dr.-Ing. Wolfgang Nowak, M.Sc.
Stochastic Modeling of Hydrosystems Group
Chair for Hydromechanics and Modelling of Hydrosystems
Institute for Modelling Hydraulic and Environmental Systems
University of Stuttgart
Germany

cc BY

Prediction Uncertainty

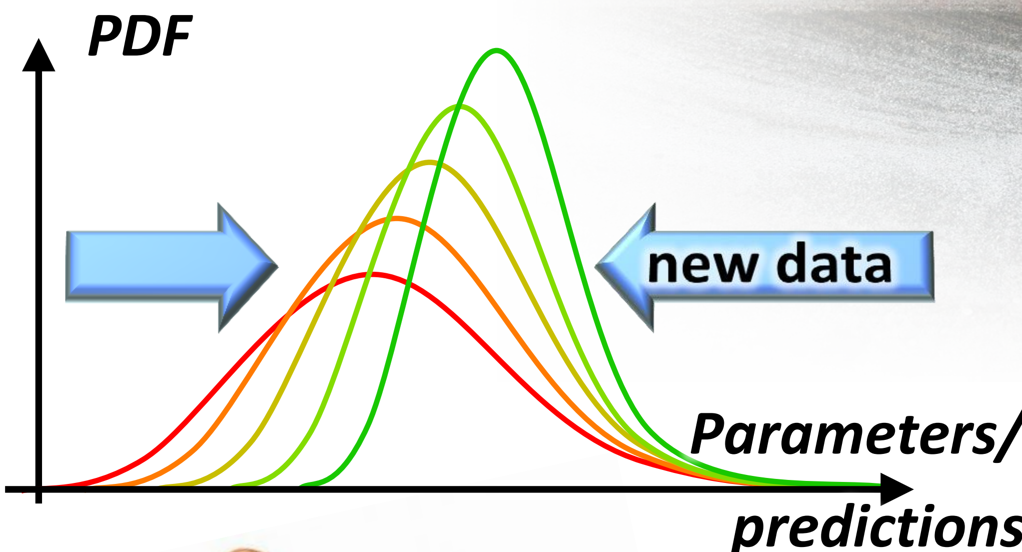
…an omnipresent challenge:

- Open system
  (boundaries, forcing)?
- Model concepts
- Scales of
  heterogeneity
- Scarce data
- Measurement
  errors



CC BY

Field Investigation?



SimTech Cluster of Excellence

Universität Stuttgart Germany

Experimental design

…is blueprint of experiment/campaign:

- Apparatus/system
- Conditions
  - System forcing (active/passive)
  - Boundary/initial conditions
- Observation type (what data)
- Sampling pattern (where)
- Sampling schedule (when)

„d“

CC BY

Optimal design: challenges

Prior enve-lope?

1.) Stochastic/geostatistical model:
realistic envelope to uncertainty p(θ)

↓

2.) Define utility of desired data:
task-driven information needs U=f(d)

↓

3.) For a given design:
Quantify utility of future data U(d)

↓

4.) Optimize design: d_{opt}=arg max U(d)
maximize utility (e.g., cost-benefit)

(loop back from 4.) to 3.))

cc BY

Prior envelope?

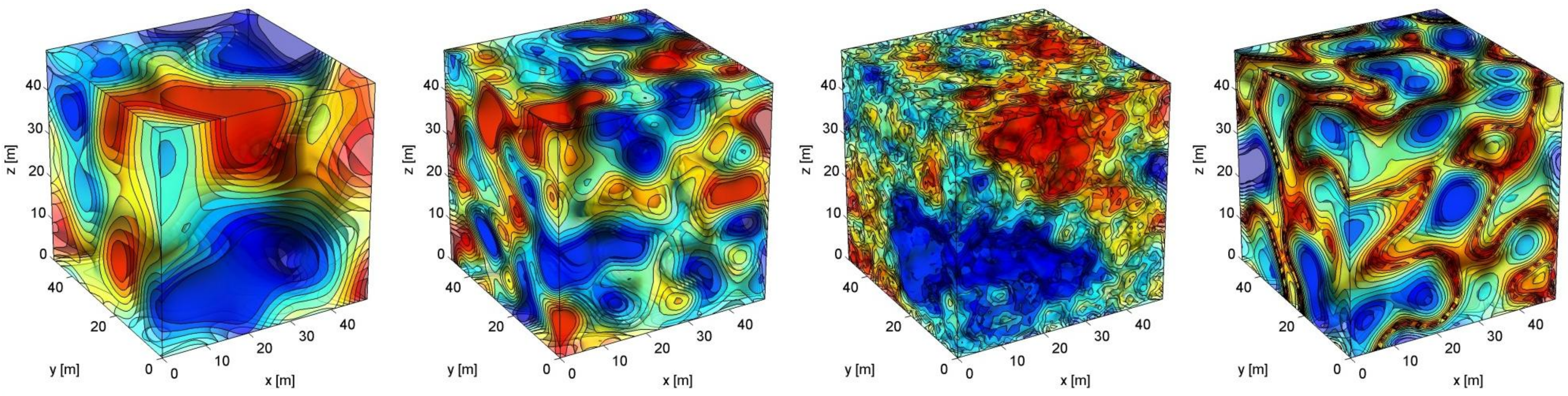
Example: geostatistical model uncertain!

– Bayesian geostatistics *(Kitanidis, WRR 1986)*

What could be uncertain?

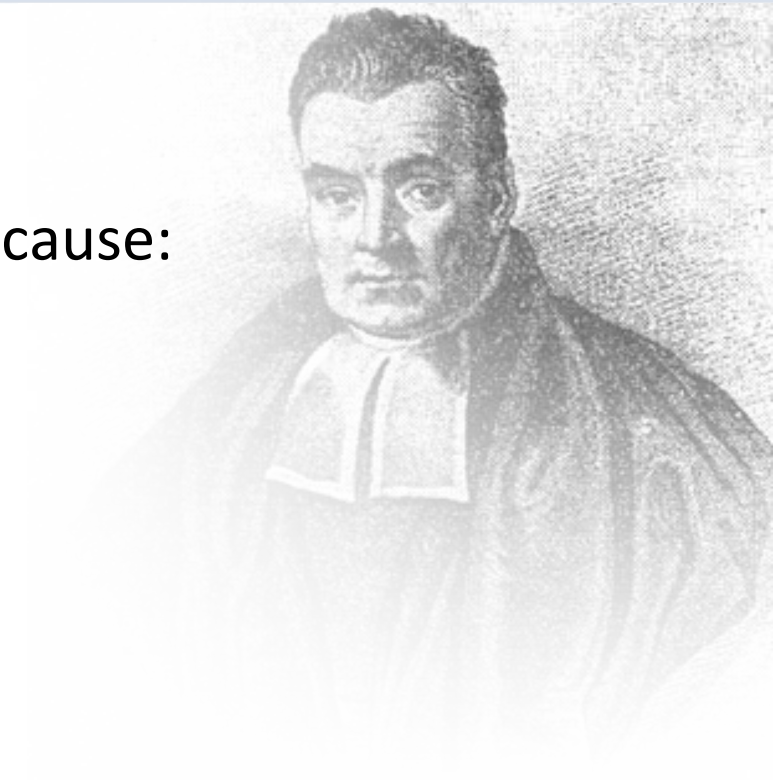
– Variance, length scales, anisotropy, smoothness
(Matérn covariance) *(Nowak & al, WRR 2010,*
(BMA) *S. Neumann & al, AWR 2012)*

– Trend/zonation models *(M. Fienen & al, WRR 2008)*



BY

What to do?

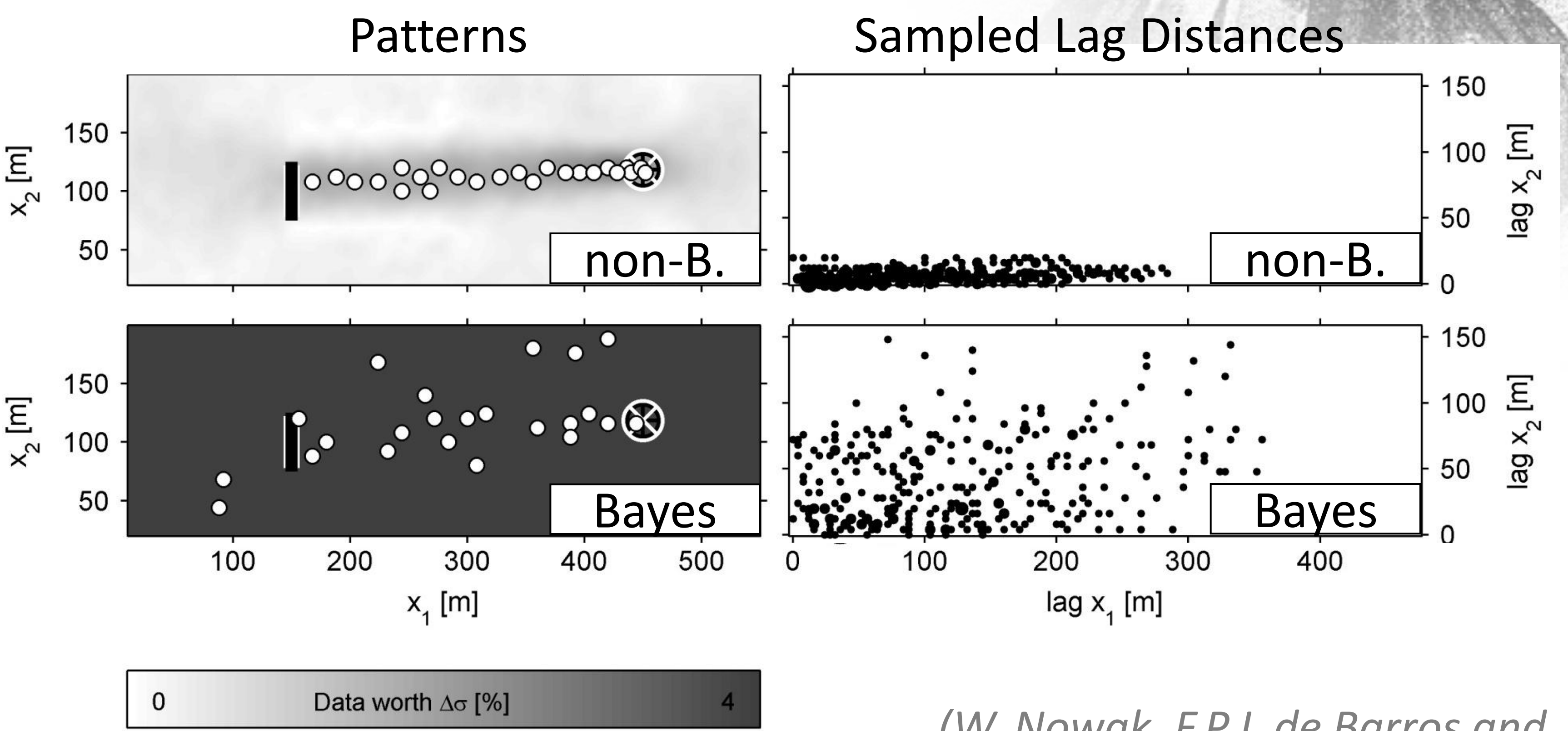


- Accept structural uncertainty, because:
  - Uncertain even after sampling
  - High impact on prediction goal
  - You will learn nothing about what you claim to know already
- Include & tackle this uncertainty in optimal design!
- Method: “Bayesian geostatistical design”:
  - Model identification as (implicit/secondary) objective / MOO
  - Robustness against too narrow priors

  (W. Nowak, F.P.J. de Barros and Y. Rubin, WRR 2010)

cc BY

Bayesian geostat. Design



(W. Nowak, F.P.J. de Barros and Y. Rubin, WRR 2010)

CC BY

Optimal design: challenges

Information needs?

1.) Stochastic/geostatistical model:
realistic envelope to uncertainty

↓

2.) Define utility of desired data:
task-driven information needs

↓

3.) For a given design:
Quantify utility of future data

↓

4.) Optimize design:
maximize utility (e.g., cost-benefit)

(loop back from 4.) to 3.))

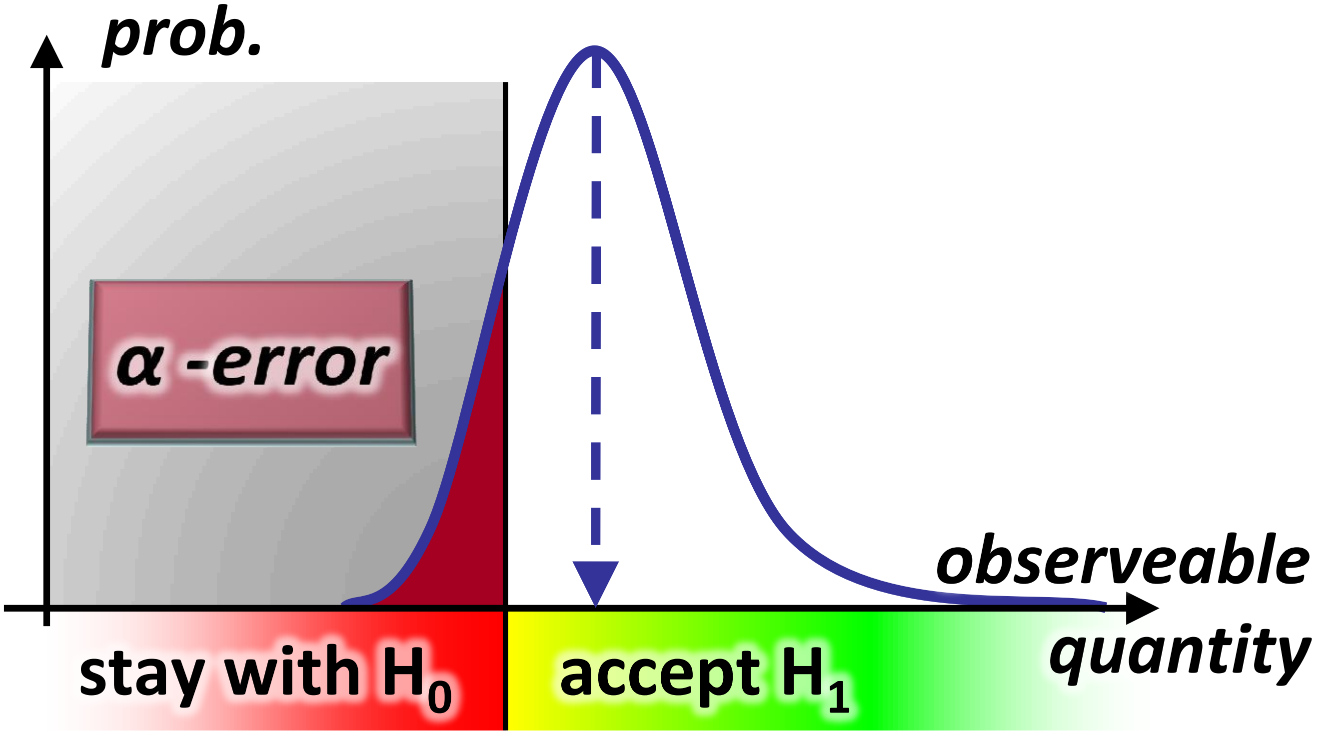
cc BY

Information needs?

- Shut down well after contaminant spill?
  (e.g., Frind et al., 2006; Enzenhöfer et al., 2010)
- Is a contaminant release rate critical?
  (e.g., Troldborg et al., 2010)
- Is natural attenuation occurring?
  (e.g., Schwede und Cirpka, 2010)
- Is a proposed remediation design safe?
  (e.g., Cirpka et al., 2004)
- Is a radioactive waste site safe?
  (e.g., Andricevic und Cvetkovic, 1996)
- Is human health risk above critical level?
  (e.g., de Barros und Rubin, 2008; de Barros et al., 2009)
- Is a model adequate to describe a site?
  (e.g., Neuman, 2003; Refsgaard et al., 2006)

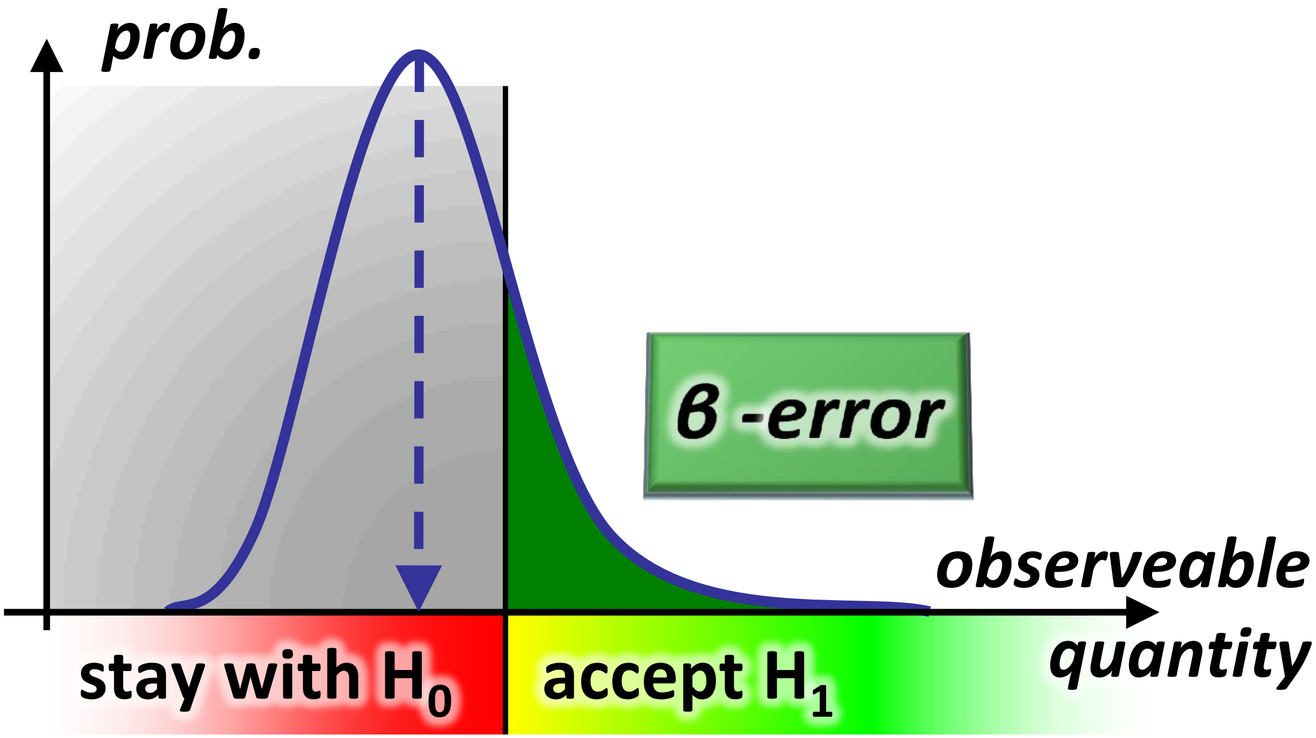
CC BY

Failure of decisions



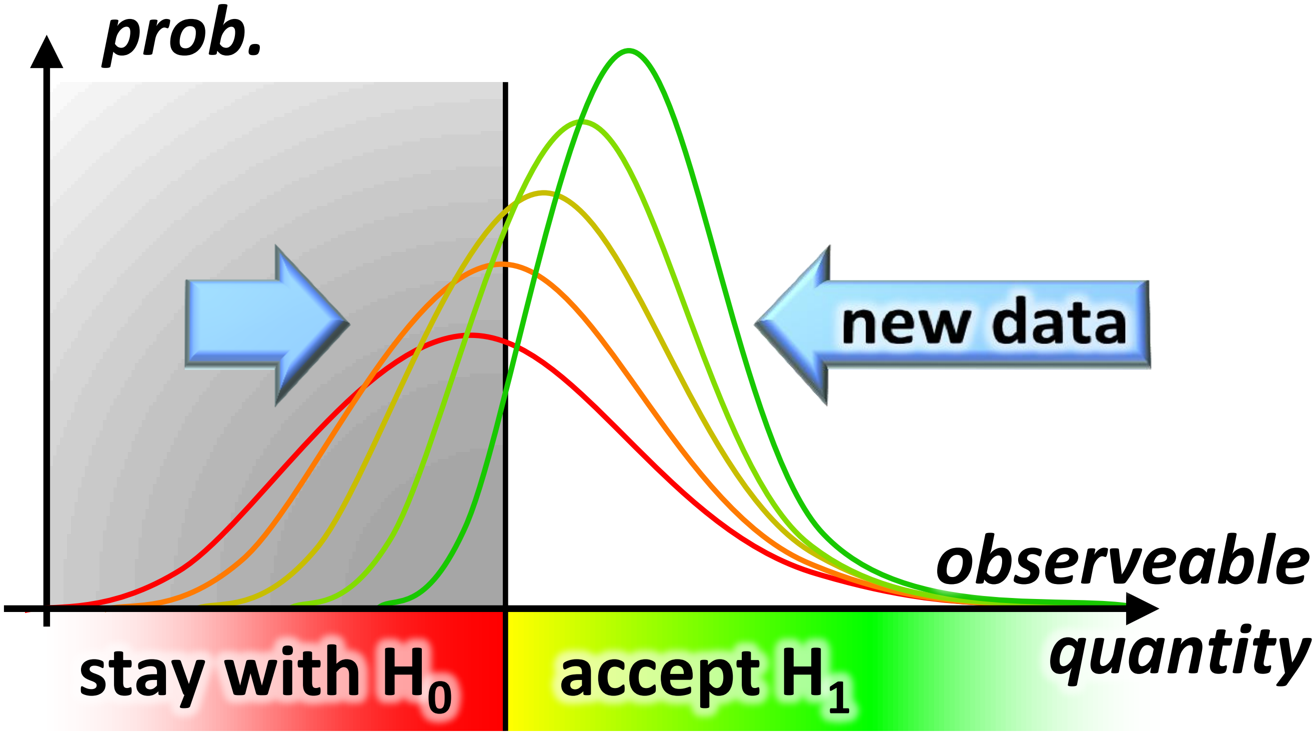
CC BY

Failure of decisions



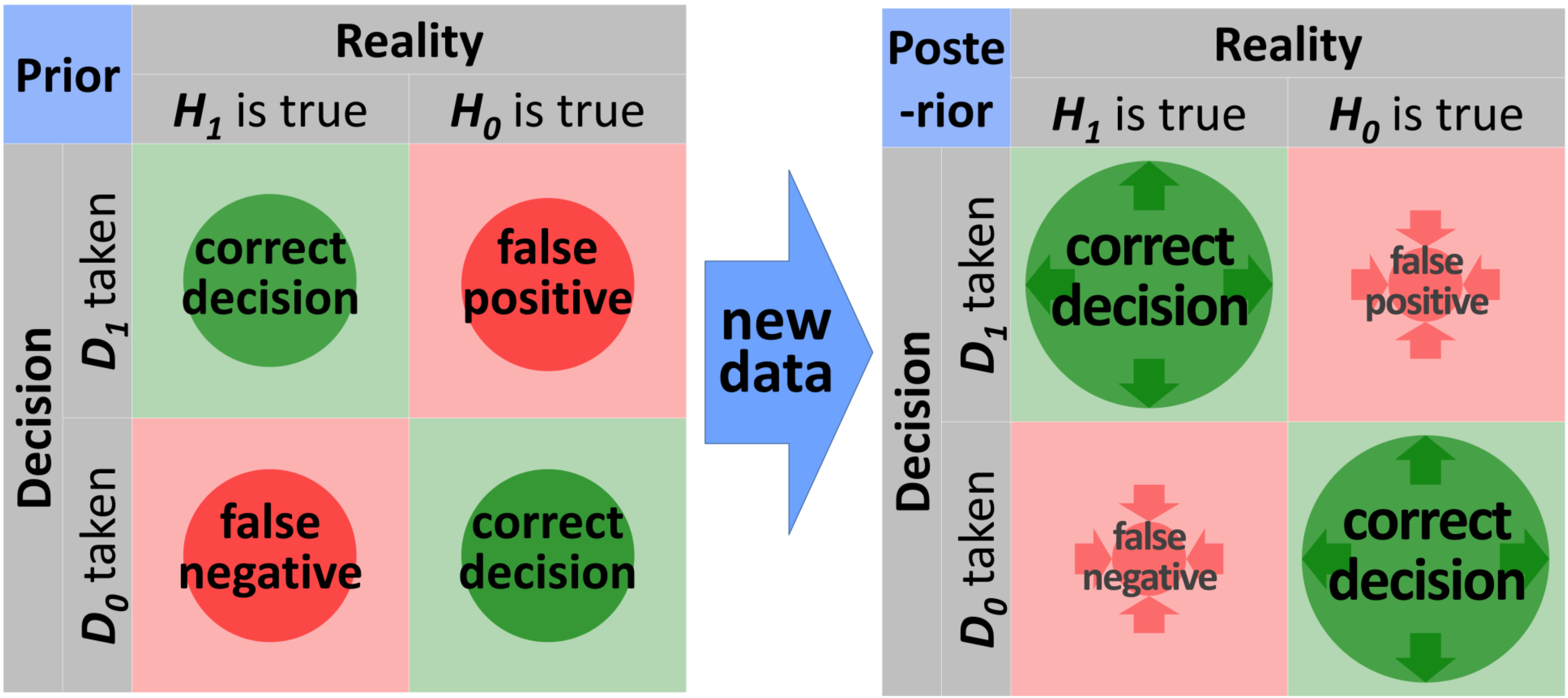
CC BY

Failure of decisions



CC BY

Hypothesis-driven design



- Uses **actual** goals (no surrogates) as objective function
- Searching “scientific truth” avoids “monetization” etc.
- Model assumptions / identification, decision support…

W. Nowak, F.P.J. de Barros and Y. Rubin, WRR2012 (submitted)

CC BY

Optimal design: challenges

1.) Stochastic/geostatistical model:
realistic envelope to uncertainty

↓

2.) Define utility of desired data:
task-driven information needs

↓

3.) For a given design:
Quantify ***expected*** utility

↓

pre-posterior analysis →

**3b.) Generate random data sets,
condition & assess utility for all**

↓

4.) Optimize design:
maximize utility (e.g., cost-benefit)

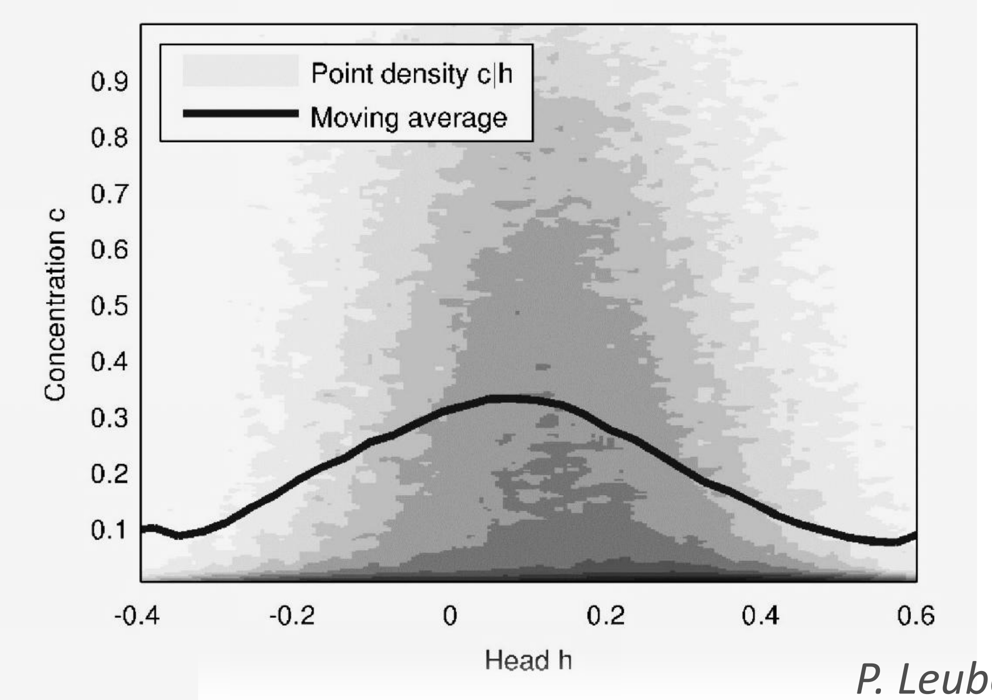
(loop back from 4.) to 3.))

CC BY

Preposterior Analysis

Preposterior **D**ata **I**mpact **A**ssessment (PreDIA):

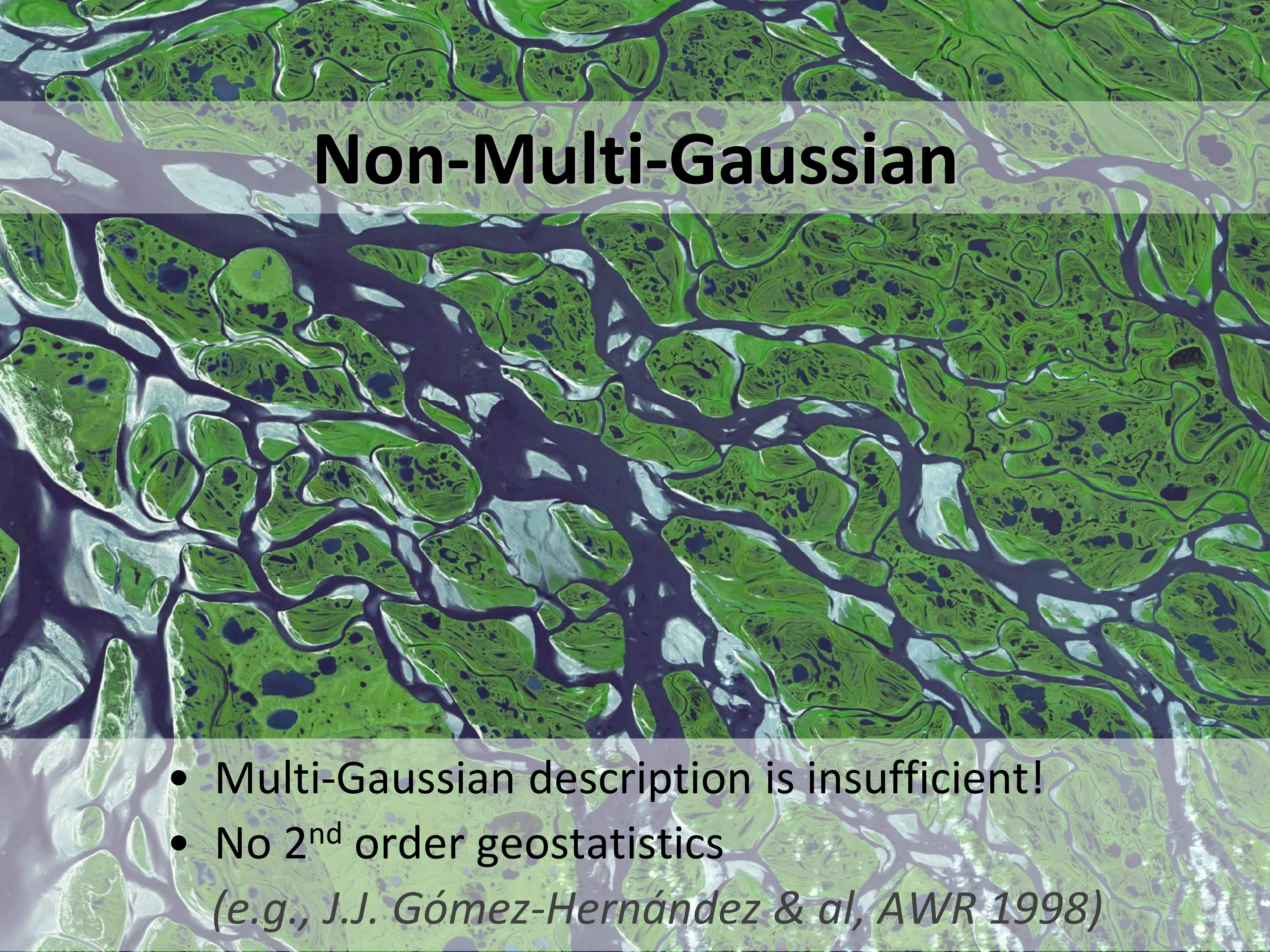
- Any task-driven criteria possible
- Fully non-linear method (non-lin. info hidden for linear tools)



- Bootstrap filter
- Average over data Marginalized errors
- Cross-breeding

P. Leube, A. Geiges and W. Nowak, WRR 2012

cc BY

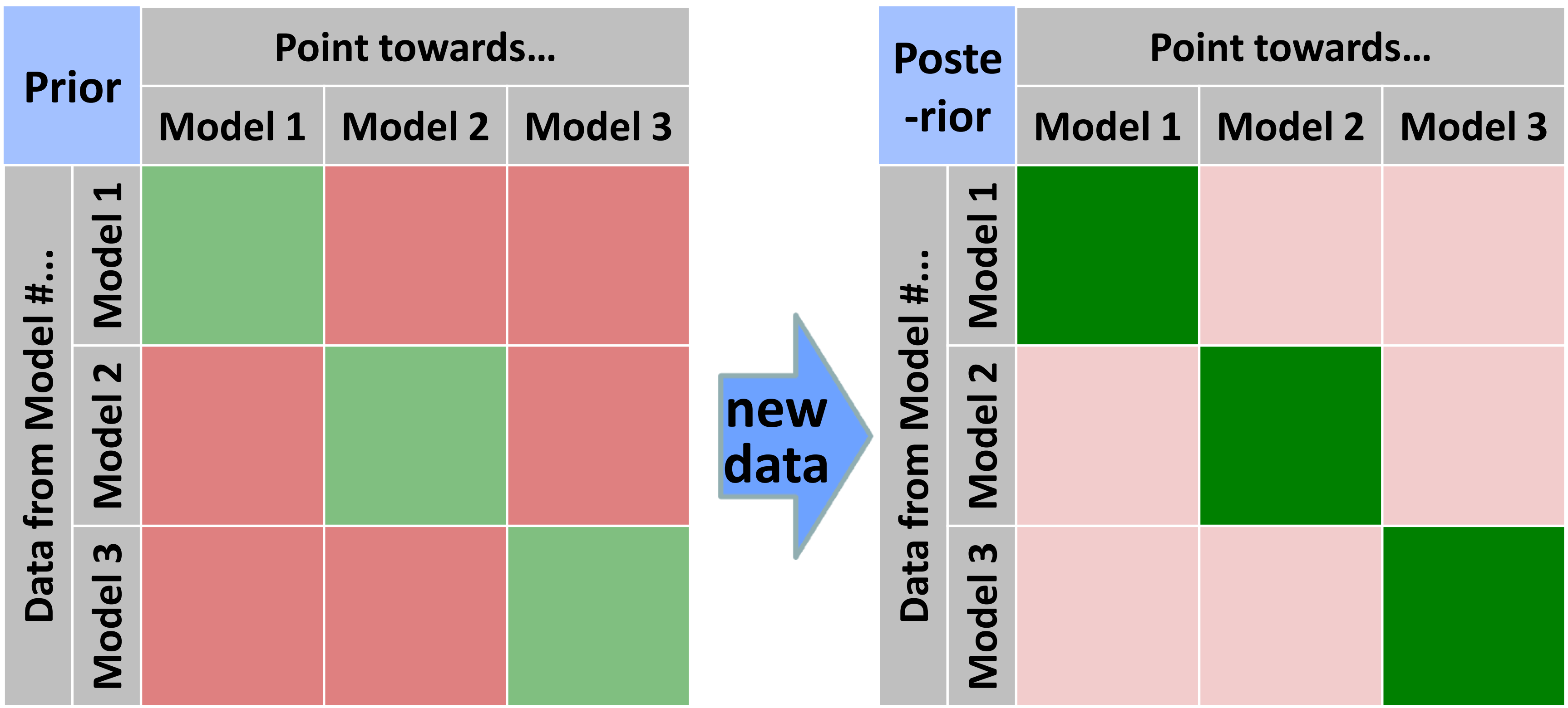


Non-Multi-Gaussian

- Multi-Gaussian description is insufficient!
- No 2nd order geostatistics
  (e.g., J.J. Gómez-Hernández & al, AWR 1998)

Outlook: Model Choice

- What will a good "model choice" design do?



- Hardly any analytical solutions
- Brute-Force (PreDIA), CPU-intensive

Summary & Conclusions

- OD: Quest for perfect information
  - Envelope for prior uncertainty (stay robust w.r.t. priors)
  - Align OD with task at hand (e.g., hypothesis-driven)
  - Fast tools becoming available (e.g., PreDIA)

- Major challenges: 1.) Multivariate structure
  2.) CPU time / real time

CC BY

Funding Acknowledgments

- EXC 310/1 SimTech
- IRTG 1398 (NUPUS)
- VW foundation
- State of BW