Improving the neutrality of simulation studies through open science practices

Sabine Hoffmann, Anne-Laure Boulesteix

Ludwig-Maximilians-Universität München

10.12.2024

Conflict of Interest: I have no current or past relationships with commercial entities.

Overview

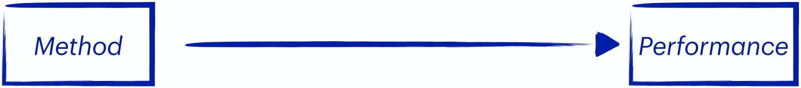
1. Why do we need open science practices in the design and analysis of simulation studies?

2. Illustration: Phase IV simulation study on the correction of measurement error in occupational epidemiology

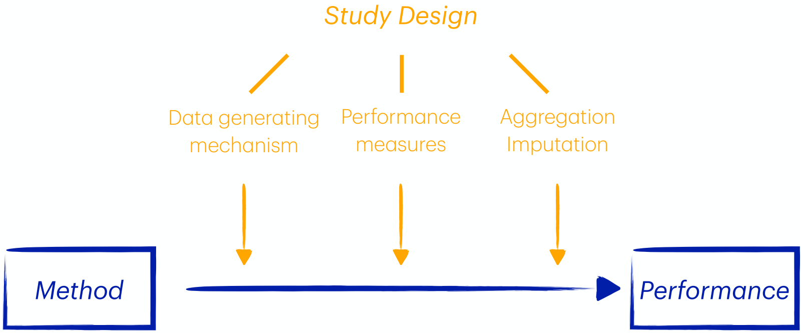
3. Outlook and discussion

Why do we need open science practices in the design and analysis of simulation studies?

Elements influencing the performance of a statistical method



Elements influencing the performance of a statistical method



Study Design

Data generating mechanism

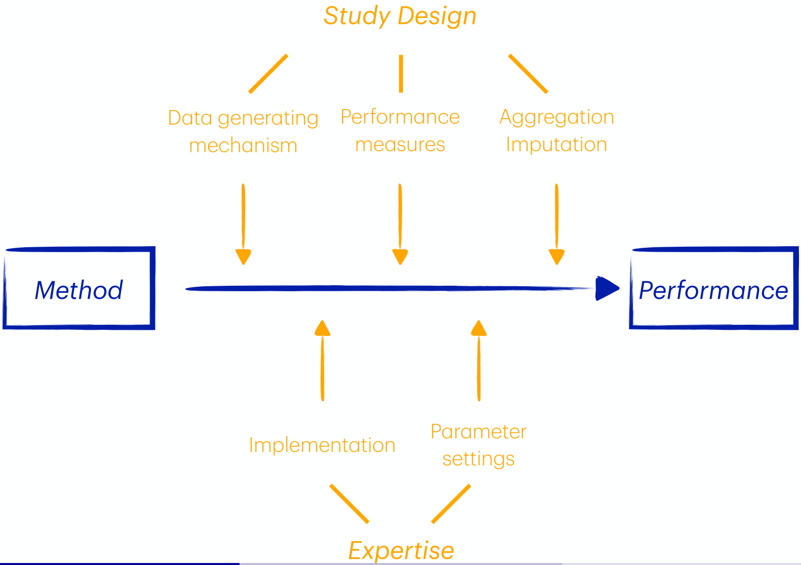
Performance measures

Aggregation
Imputation

Method

Performance

Elements influencing the performance of a statistical method



Study Design

Data generating mechanism

Performance measures

Aggregation
Imputation

Method

Performance

Implementation

Parameter settings

Expertise

Overoptimism in methodological research

Received: 1 June 2021 | Revised: 16 September 2021 | Accepted: 4 November 2021

DOI: 10.1002/widm.1441

FOCUS ARTICLE

WILEY

Over-optimism in benchmark studies and the multiplicity of design and analysis options when interpreting their results

Christina Nießl[1] | Moritz Herrmann[2] | Chiara Wiedemann[1] | Giuseppe Casalicchio[2] | Anne-Laure Boulesteix[1]

[1]Institute for Medical Information Processing, Biometry and Epidemiology, Ludwig Maximilians University Munich, Munich, Germany

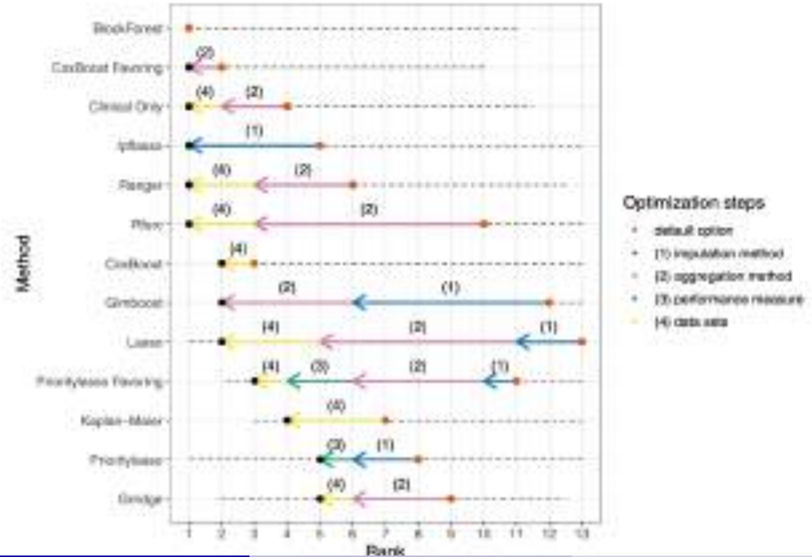
[2]Department of Statistics, Ludwig Maximilians University Munich, Munich, Germany

Correspondence
Christina Nießl, Institute for Medical Information Processing, Biometry and Epidemiology, Ludwig Maximilians University Munich, Marchioninistr. 15, D-81377 Munich, Germany.
Email: niessl@ibe.med.uni-muenchen.de

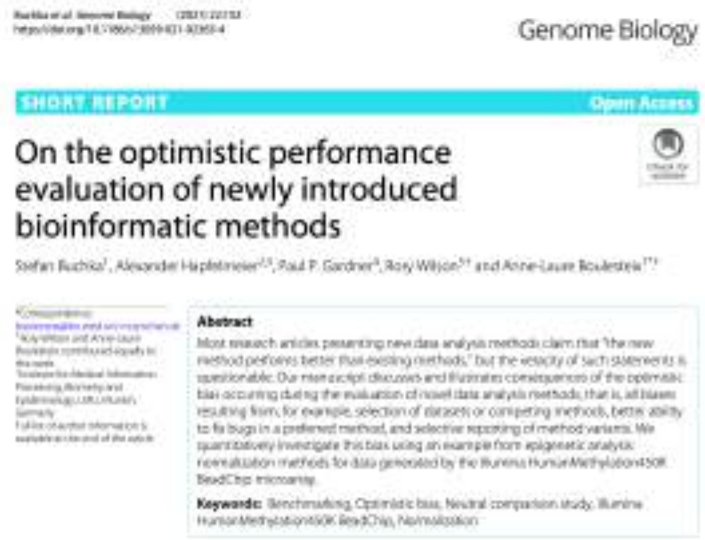
Abstract

In recent years, the need for neutral benchmark studies that focus on the comparison of methods coming from computational sciences has been increasingly recognized by the scientific community. While general advice on the design and analysis of neutral benchmark studies can be found in recent literature, a certain flexibility always exists. This includes the choice of data sets and performance measures, the handling of missing performance values, and the way the performance values are aggregated over the data sets. As a consequence of this flexibility, researchers may be concerned about how their choices affect the results or, in the worst case, may be tempted to engage in questionable research practices (e.g., the selective reporting of results or the post hoc modifi-

Overoptimism in methodological research



Overoptimism in methodological research



Buchka et al. Genome Biology (2021) 22:152
https://doi.org/10.1186/s13059-021-02365-4

Genome Biology

SHORT REPORT — Open Access

On the optimistic performance evaluation of newly introduced bioinformatic methods

Stefan Buchka[1], Alexander Hapfelmeier[1,2], Paul P. Gardner[3], Rory Wilson[4†] and Anne-Laure Boulesteix[1*†]

*Correspondence: boulesteix@ibe.med.uni-muenchen.de
†Rory Wilson and Anne-Laure Boulesteix contributed equally to this work
[1] Institute for Medical Information Processing, Biometry and Epidemiology, LMU Munich, Germany
Full list of author information is available at the end of the article

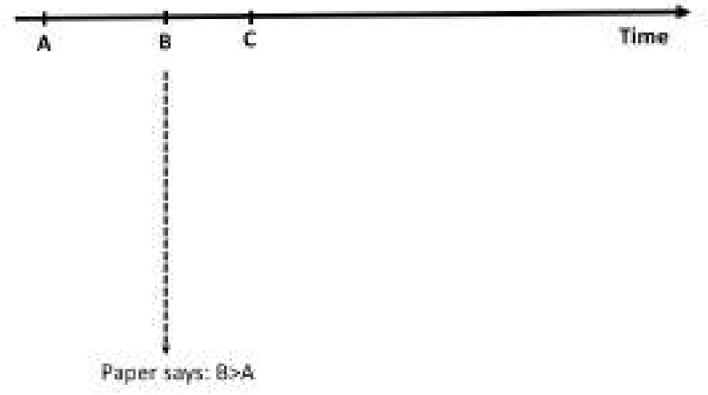
Abstract

Most research articles presenting new data analysis methods claim that "the new method performs better than existing methods," but the veracity of such statements is questionable. Our manuscript discusses and illustrates consequences of the optimistic bias occurring during the evaluation of novel data analysis methods, that is, all biases resulting from, for example, selection of datasets or competing methods, better ability to fix bugs in a preferred method, and selective reporting of method variants. We quantitatively investigate this bias using an example from epigenetic analysis: normalization methods for data generated by the Illumina HumanMethylation450K BeadChip microarray.

Keywords: Benchmarking, Optimistic bias, Neutral comparison study, Illumina HumanMethylation450K BeadChip, Normalization

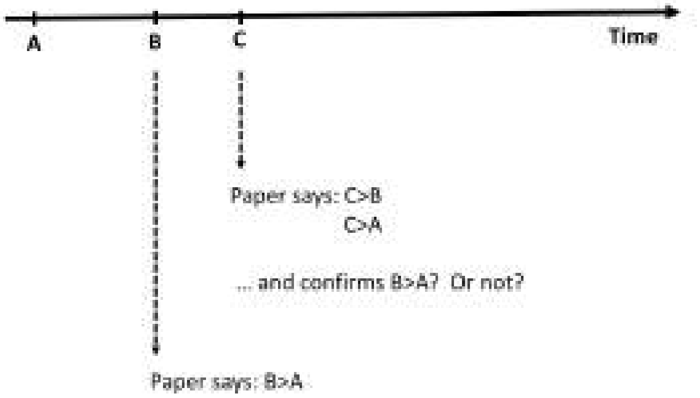
Overoptimism in methodological research

"The New Method Performed Better Than Existing Ones"

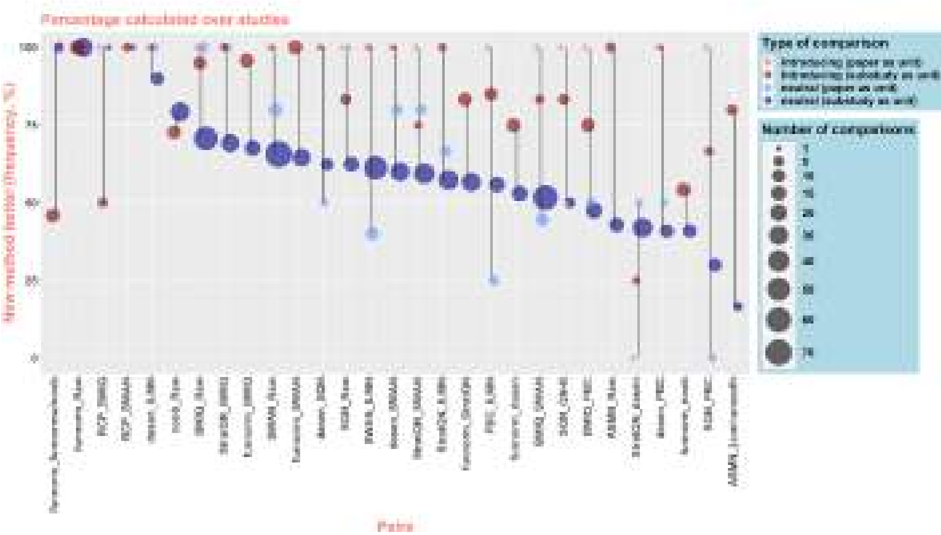


Overoptimism in methodological research

“The New Method Performed Better Than Existing Ones”



Overoptimism in methodological research



A replication crisis in methodological research?

A replication crisis in methodological research?

Statisticians have been keen to critique statistical aspects of the "replication crisis" in other scientific disciplines. But new statistical tools are often published and promoted without any thought to replicability. This needs to change, argue **Anne-Laure Boulesteix, Sabine Hoffmann, Alethea Charlton** and **Heidi Seibold**

How to improve neutrality through open science practices?

- Data generation:
  - *Pre-registration* of simulation setup including transparent reporting of pilot studies with feedback by experts
  - Data are generated by an independent team

How to improve neutrality through open science practices?

- Data generation:
    - *Pre-registration* of simulation setup including transparent reporting of pilot studies with feedback by experts
    - Data are generated by an independent team
- Expertise:
    - Involve independent experts for all methods

How to improve neutrality through open science practices?

- Data generation:
  - *Pre-registration* of simulation setup including transparent reporting of pilot studies with feedback by experts
  - Data are generated by an independent team
- Expertise:
  - Involve independent experts for all methods
- Reporting:
  - *Blinded reporting* of results by independent person who has little experience with any of the methods
  - Shiny app: *Comprehensive visualization* of complex simulation results may reduce selective reporting of results

How to improve neutrality through open science practices?

- Data generation:
  - *Pre-registration* of simulation setup including transparent reporting of pilot studies with feedback by experts
  - Data are generated by an independent team
- Expertise:
  - Involve independent experts for all methods
- Reporting:
  - *Blinded reporting* of results by independent person who has little experience with any of the methods
  - Shiny app: *Comprehensive visualization* of complex simulation results may reduce selective reporting of results
- Transparency:
  - *Code sharing* for methods and for simulation study

How can code sharing help?

[illegible] | [illegible] | [illegible]

[illegible]

RESEARCH ARTICLE

Biometrical Journal

Explaining the optimistic performance evaluation of newly proposed methods: A cross-design validation experiment

Christina Nießl[1,2] | Sabine Hoffmann[2,3] | Theresa Ullmann[1] | Anne-Laure Boulesteix[1]

[1] Institute for Medical Information Processing, Biometry and Epidemiology, LMU Munich, Munich, Germany

[2] Munich Center for Machine Learning (MCML), Munich, Germany

[3] Department of Statistics, LMU Munich, Munich, Germany

Correspondence
Christina Nießl, Institute for Medical Information Processing, Biometry and Epidemiology, LMU Munich, Marchioninistr. 15, 81377 Munich, Germany.
Email: [illegible]

Funding information
German Federal Ministry of Education and Research, Grant/Award Number: [illegible]; Deutsche Forschungsgemeinschaft, Grant/Award

Abstract
The constant development of new data analysis methods in many fields of research is accompanied by an increasing awareness that these new methods often perform better in their introductory paper than in subsequent comparison studies conducted by other researchers. We attempt to explain this discrepancy by conducting a systematic experiment that we call "cross-design validation of methods". In the experiment, we select two methods designed for the same data analysis task, reproduce the results shown in each paper, and then reevaluate each method based on the study design (i.e., datasets, competing methods, and evaluation criteria) that was used to show the abilities of the other method. We conduct the experiment for two data analysis tasks, namely cancer subtyping using multiomic data and differential gene expression analysis. Three of the four methods included in the experiment indeed perform worse when they are evaluated on the new study design, which is mainly caused by the different datasets. Apart from illustrating the many degrees of freedom existing in the assessment of a method and their effect on its performance, our experiment suggests that

How can code sharing help?

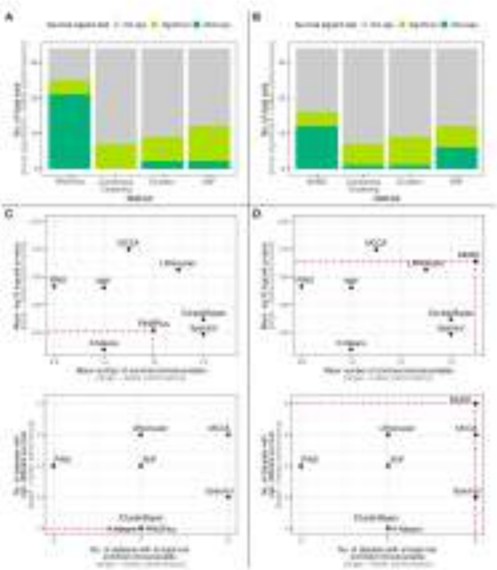


Illustration: Phase IV simulation study on the correction of measurement error in occupational epidemiology

Background

- Uncertainty in exposure assessment poses an important threat to the validity of statistical inference in occupational epidemiology

Background

- Uncertainty in exposure assessment poses an important threat to the validity of statistical inference in occupational epidemiology
- Exposure assessment in occupational epidemiology is often based on Job Exposure Matrices in which there are different sources of error [Greenland et al., 2016]:
    - Exposure information for each job is usually imprecise or incomplete
    - Exposures within a given job code may vary considerably from person to person due to differences in job conditions and worker practices

Classical measurement error

$Z_i(t) = X_i(t) \cdot U_i(t)$

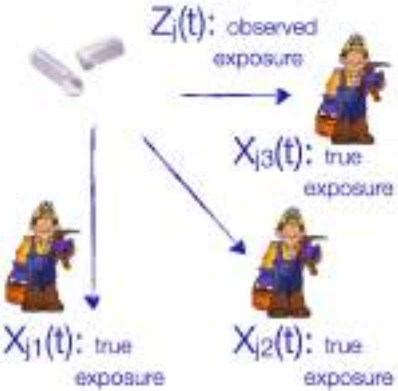
- $U_i(t) \perp X_i(t)$
- $Var(Z_i(t)) > Var(X_i(t))$

Berkson error

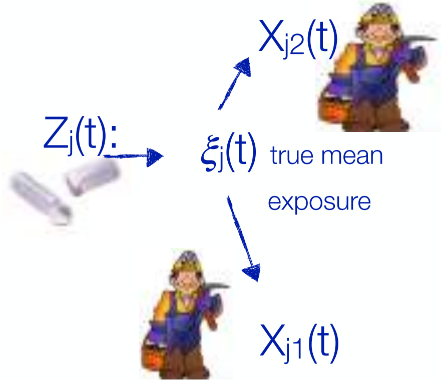
$X_{ji}(t) = Z_j(t) \cdot U_{ji}(t)$

- $U_i(t) \perp Z(t)$
- $Var(X_{ij}(t)) > Var(Z(t))$

$Z_i(t)$: observed exposure
$X_i(t)$: true exposure


$Z_j(t)$: observed exposure
$X_{j3}(t)$: true exposure
$X_{j1}(t)$: true exposure
$X_{j2}(t)$: true exposure

Shared measurement error



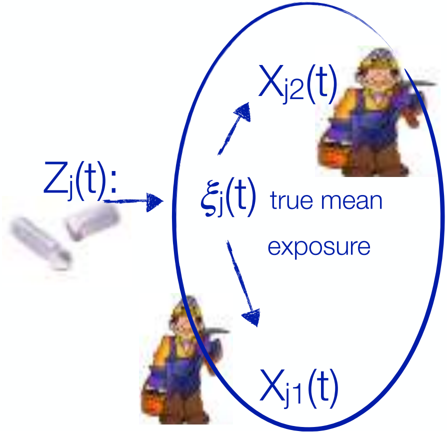
$X_{j2}(t)$

$Z_j(t)$:

$\xi_j(t)$ true mean exposure

$X_{j1}(t)$

Shared measurement error



$Z_j(t)$:

$X_{j2}(t)$

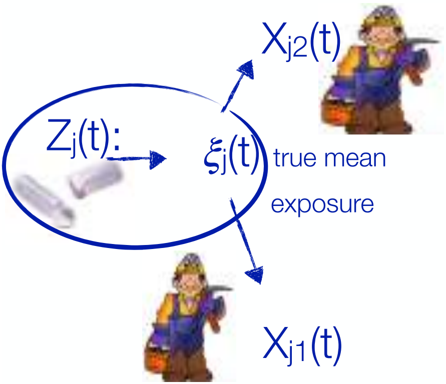
$\xi_j(t)$ true mean exposure

$X_{j1}(t)$

Berkson error

Shared measurement error

Classical measurement error
shared between miners



$Z_j(t)$: $\xi_j(t)$ true mean exposure

$X_{j2}(t)$

$X_{j1}(t)$

Exposure assessment in the second exposure period

$$E(t,o,j) = C_{Rn}(p_{to}) \cdot 12 \cdot g(p_{to}) \cdot w(p_t) \cdot f(p_{oj})$$

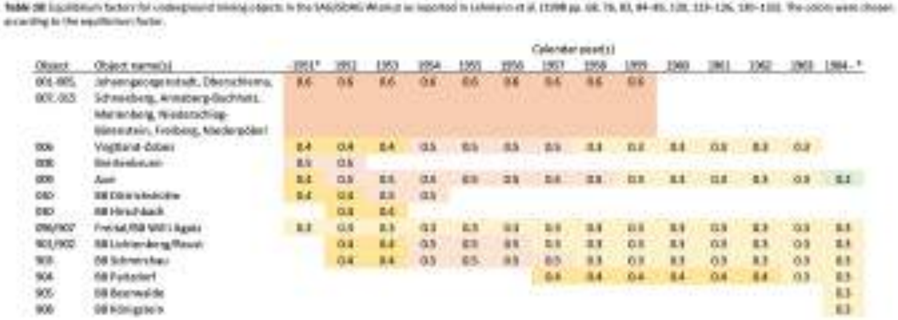
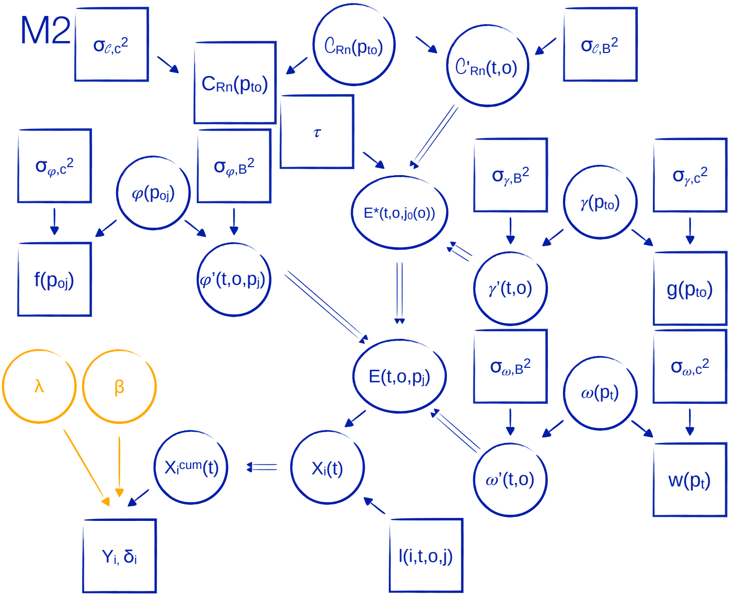


Table 38 Equilibrium factors for underground mining objects in the SAG/SDAG Wismut as reported in Lehmann et al. (1998 pp. 68, 76, 83, 84–85, 120, 123–126, 130–133). The colors were chosen according to the equilibrium factor.

		Calendar year(s)													
Object	Object name(s)	1951*	1952	1953	1954	1955	1956	1957	1958	1959	1960	1961	1962	1963	1964–*
001-005, 007-015	Johanngeorgenstadt, Oberschlema, Schneeberg, Annaberg-Buchholz, Marienberg, Niederschlag-Bärenstein, Freiberg, Niederpöbel	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6					
006	Vogtland-Zobes	0.4	0.4	0.4	0.5	0.5	0.5	0.5	0.3	0.3	0.3	0.3	0.3	0.3	
008	Breitenbrunn	0.5	0.5												
009	Aue	0.4	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.3	0.3	0.3	0.3	0.3	0.2
010	BB Dittrichshütte	0.4	0.4	0.5	0.5										
030	BB Hirschbach		0.4	0.4											
096/907	Freital, BB Willi Agatz	0.3	0.5	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3
901/902	BB Lichtenberg/Reust		0.4	0.4	0.5	0.5	0.5	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3
903	BB Schmirchau		0.4	0.4	0.5	0.5	0.5	0.5	0.3	0.3	0.3	0.3	0.3	0.3	0.3
904	BB Paitzdorf							0.4	0.4	0.4	0.4	0.4	0.4	0.3	0.3
905	BB Beerwalde														0.3
906	BB Königstein														0.3


M2
$\sigma_{\mathcal{C},c}^2$
$C_{Rn}(p_{to})$
$\mathcal{C}_{Rn}(p_{to})$
$\mathcal{C}'_{Rn}(t,o)$
$\sigma_{\mathcal{C},B}^2$
τ
$\sigma_{\varphi,c}^2$
$\varphi(p_{oj})$
$\sigma_{\varphi,B}^2$
$E^*(t,o,j_0(o))$
$\sigma_{\gamma,B}^2$
$\gamma(p_{to})$
$\sigma_{\gamma,c}^2$
$f(p_{oj})$
$\varphi'(t,o,p_j)$
$\gamma'(t,o)$
$g(p_{to})$
λ
β
$E(t,o,p_j)$
$\sigma_{\omega,B}^2$
$\omega(p_t)$
$\sigma_{\omega,c}^2$
$X_i^{cum}(t)$
$X_i(t)$
$\omega'(t,o)$
$w(p_t)$
Y_i, δ_i
$I(i,t,o,j)$

Measurement models for the second exposure period

$$C_{Rn}(p_{to}) = \mathcal{C}_{Rn}(p_{to}) + U_{\mathcal{C},c}(p_{to})$$
$$\mathcal{C}'_{Rn}(t,o) = \mathcal{C}_{Rn}(p_{to}) \cdot U_{\mathcal{C},B}(t,o)$$

$$f(p_{oj}) = \varphi(p_{oj}) \cdot U_{\varphi,c}(p_{oj})$$
$$\varphi'(t,o,p_j) = \varphi(p_{oj}) \cdot U_{\varphi,B}(t,o,p_j)$$

$$w(p_t) = \omega(p_t) \cdot U_{\omega,c}(p_t)$$
$$\omega'(t,o) = \omega(p_t) \cdot U_{\omega,B}(t,o)$$

$$g(p_{to}) = \gamma(p_{to}) \cdot U_{\gamma,c}(p_{to})$$
$$\gamma'(t,o) = \gamma(p_{to}) \cdot U_{\gamma,B}(t,o)$$

Aims of the simulation study

- Assess the overall impact of measurement error on risk estimation with a naive estimate which does not assume any measurement error

Aims of the simulation study

- Assess the overall impact of measurement error on risk estimation with a naive estimate which does not assume any measurement error
- Assess the performance of a Bayesian hierarchical approach and compare it with SIMEX and regression calibration

Aims of the simulation study

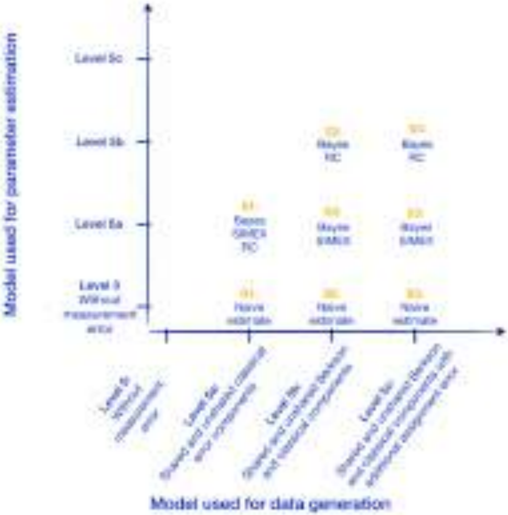
- Assess the overall impact of measurement error on risk estimation with a naive estimate which does not assume any measurement error
- Assess the performance of a Bayesian hierarchical approach and compare it with SIMEX and regression calibration
- Assess to what extent the complex structures of measurement error can be accounted for with simplified measurement models by considering the results under model misspecification

How to choose a neutral data generating mechanism?

Figure: “Climb the tree”. Drawing from Alexandra Kalberer, published in [Strobl and Leisch, 2024]

How to choose a neutral data generating mechanism?

Figure: "Climb the tree". Drawing from Alexandra Kalberer, published in [Strobl and Leisch, 2024]



How to address inventor bias and differences in expertise?

- Independence: Person A responsible for the implementation of the Bayesian hierarchical model, person B responsible for data generation and the implementation of SIMEX and regression calibration

How to address inventor bias and differences in expertise?

- Independence: Person A responsible for the implementation of the Bayesian hierarchical model, person B responsible for data generation and the implementation of SIMEX and regression calibration
- Expertise: Involve two experts on frequentist methods for measurement error correction

Preliminary results - Scenario 1

	coverage	beta		bias of the mean	
	rate	mean	median	absolute	relative in %
naive (frequentist)	0.31	0.27	0.25	-0.03	-11.32
naive (Bayes)	0.31	0.26	0.25	-0.04	-12.76
RC	0.39	0.32	0.27	0.02	5.96
Bayes	0.94	0.29	0.29	-0.01	-2.98
SIMEX	0.57	0.29	0.28	-0.01	-4.24

Preliminary results - Scenario 2

	coverage	beta		bias of the mean	
	rate	mean	median	absolute	relative in %
naive (frequentist)	0.25	0.25	0.24	-0.05	-17.36
naive (Bayes)	0.27	0.24	0.24	-0.06	-18.65
RC	0.29	0.29	0.25	-0.01	-2.57
Bayes	0.93	0.32	0.32	0.02	6.76
adjustment for classical error					
Bayes Level a	0.60	0.31	0.31	0.01	4.88
SIMEX	0.61	0.27	0.25	-0.03	-11.47

Outlook and discussion

- Outlook:
  - Evaluate performance on new data generation mechanism

Outlook and discussion

- Outlook:
  - Evaluate performance on new data generation mechanism
  - Pre-register simulation design and methods and ask for feedback of STRATOS experts on measurement error

Outlook and discussion

- Outlook:
  - Evaluate performance on new data generation mechanism
  - Pre-register simulation design and methods and ask for feedback of STRATOS experts on measurement error
  - Limit spin and selective reporting through blinded reporting of results

Outlook and discussion

- Outlook:
  - Evaluate performance on new data generation mechanism
  - Pre-register simulation design and methods and ask for feedback of STRATOS experts on measurement error
  - Limit spin and selective reporting through blinded reporting of results
- Discussion:
  - Is it really a phase IV study?

Outlook and discussion

- Outlook:
  - Evaluate performance on new data generation mechanism
  - Pre-register simulation design and methods and ask for feedback of STRATOS experts on measurement error
  - Limit spin and selective reporting through blinded reporting of results
- Discussion:
  - Is it really a phase IV study?
  - Is the performance of a method when implemented by experts (level 3) really of interest?

Thank you for your attention!

Greenland, S., Senn, S. J., Rothman, K. J., Carlin, J. B., Poole, C., Goodman, S. N., and Altman, D. G. (2016).
Statistical tests, p values, confidence intervals, and power: a guide to misinterpretations.
European Journal of Epidemiology, 31(4):337–50.

Küchenhoff, H., Deffner, V., Aßenmacher, M., Neppl, H., Kaiser, C., Güthlin, D., et al. (2018).
Ermittlung der Unsicherheiten der Strahlenexpositionsabschätzung in der Wismut-Kohorte - Teil I - Vorhaben 3616S12223.
Ressortforschungsberichte zum Strahlenschutz.
Bundesamt für Strahlenschutz (BfS).

Strobl, C. and Leisch, F. (2024).
Against the “one method fits all data sets” philosophy for comparison studies in methodological research.
Biometrical Journal, 66(1):2200104.

Simulation scenario S1

$$C_{Rn}(t,o) = \mathcal{C}_{Rn}(t,o) + U_c(t,o)$$

$$f(o,j) = \varphi(o,j) \cdot U_{\varphi,c}(o,j)$$

$$w(p_t) = \omega(p_t) \cdot U_{\omega,c}(p_t)$$

$$g(p_t,o) = \gamma(p_t,o) \cdot U_{\gamma,c}(p_t,o)$$

$$X_i(t,o) = \mathcal{C}_{Rn}(t,o) \cdot 12 \cdot \gamma(p_t,o) \cdot \omega(p_t) \cdot \varphi(o,j)$$

$$Z_i(t,o) = C_{Rn}(t,o) \cdot 12 \cdot g(p_t,o) \cdot w(p_t) \cdot f(o,j)$$

Simulation scenario S2

$$C_{Rn}(t,o) = \mathcal{C}_{Rn}(t,o) + U_c(t,o)$$

$$f(o,j) = \varphi(o,j) \cdot U_{\varphi,c}(o,j)$$
$$\varphi'(t,o,j) = \varphi(o,j) \cdot U_{\varphi',B}(t,o,j)$$

$$w(p_t) = \omega(p_t) \cdot U_{\omega,c}(p_t)$$
$$\omega'(t,o) = \omega(p_t) \cdot U_{\omega',B}(t,o)$$

$$g(p_t,o) = \gamma(p_t,o) \cdot U_{\gamma,c}(p_t,o)$$
$$\gamma'(t,o) = \gamma(p_t,o) \cdot U_{\gamma',B}(t,o)$$

$$X_i(t,o) = \mathcal{C}_{Rn}(t,o) \cdot 12 \cdot \gamma'(t,o) \cdot \omega'(t,o) \cdot \varphi'(t,o,j)$$

$$Z_i(t,o) = C_{Rn}(t,o) \cdot 12 \cdot g(p_t,o) \cdot w(p_t) \cdot f(o,j)$$

Preliminary results - Scenario 1

	coverage	beta		bias of the mean	
	rate	mean	median	absolute	relative in %
naive (frequentist)	0.31	0.27	0.25	-0.03	-11.32
naive (Bayes)	0.31	0.26	0.25	-0.04	-12.76
RC	0.39	0.32	0.27	0.02	5.96
Bayes	0.94	0.29	0.29	-0.01	-2.98
SIMEX	0.57	0.29	0.28	-0.01	-4.24

Simulation scenario S2

$$C_{Rn}(t,o) = \mathcal{C}_{Rn}(t,o) + U_c(t,o)$$

$$f(o,j) = \varphi(o,j) \cdot U_{\varphi,c}(o,j)$$
$$\varphi'(t,o,j) = \varphi(o,j) \cdot U_{\varphi',B}(t,o,j)$$

$$w(p_t) = \omega(p_t) \cdot U_{\omega,c}(p_t)$$
$$\omega'(t,o) = \omega(p_t) \cdot U_{\omega',B}(t,o)$$

$$g(p_t,o) = \gamma(p_t,o) \cdot U_{\gamma,c}(p_t,o)$$
$$\gamma'(t,o) = \gamma(p_t,o) \cdot U_{\gamma',B}(t,o)$$

$$X_i(t,o) = \mathcal{C}_{Rn}(t,o) \cdot 12 \cdot \gamma'(t,o) \cdot \omega'(t,o) \cdot \varphi'(t,o,j)$$

$$Z_i(t,o) = C_{Rn}(t,o) \cdot 12 \cdot g(p_t,o) \cdot w(p_t) \cdot f(o,j)$$

Preliminary results - Scenario 2

	coverage	beta		bias of the mean	
	rate	mean	median	absolute	relative in %
naive (frequentist)	0.25	0.25	0.24	-0.05	-17.36
naive (Bayes)	0.27	0.24	0.24	-0.06	-18.65
RC	0.29	0.29	0.25	-0.01	-2.57
Bayes	0.93	0.32	0.32	0.02	6.76
adjustment for classical error					
Bayes Level a	0.60	0.31	0.31	0.01	4.88
SIMEX	0.61	0.27	0.25	-0.03	-11.47

Simulation scenario S3

$$C_{Rn}(t,o) = \mathcal{C}_{Rn}(t,o) + U_c(t,o)$$

$$\begin{aligned} f(o,j) &= \varphi(o,j) \cdot U_{\varphi,c}(o,j) \\ \varphi'(t,o,j) &= \varphi(o,j) \cdot U_{\varphi',B}(t,o,j) \end{aligned}$$

$$\begin{aligned} w(p_t) &= \omega(p_t) \cdot U_{\omega,c}(p_t) \\ \omega'(t,o) &= \omega(p_t) \cdot U_{\omega',B}(t,o) \end{aligned}$$

$$\begin{aligned} g(p_t,o) &= \gamma(p_t,o) \cdot U_{\gamma,c}(p_t,o) \\ \gamma'(t,o) &= \gamma(p_t,o) \cdot U_{\gamma',B}(t,o) \end{aligned}$$

$$\begin{aligned} X_i(t,o) =& \mathcal{C}_{Rn}(t,o) \cdot 12 \cdot \gamma'(t,o) \cdot \omega'(t,o) \cdot \varphi'(t,o,j) \\ & + U_{E,B}(i,t,o,j) + U_{E,B}(i,o,j) \end{aligned}$$

$$Z_i(t,o) = C_{Rn}(t,o) \cdot 12 \cdot g(p_t,o) \cdot w(p_t) \cdot f(o,j)$$

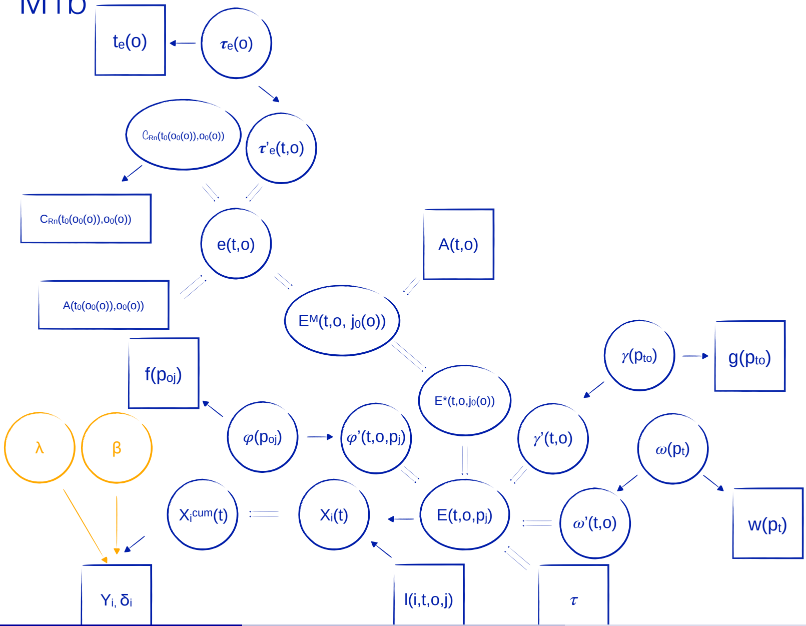
Preliminary results - Scenario 3

	coverage	beta		bias of the mean	
	rate	mean	median	absolute	relative in %
naive (frequentist)	0.28	0.24	0.24	-0.06	-19.27
naive (Bayes)	0.22	0.23	0.23	-0.06	-20.63
RC	0.37	0.29	0.25	-0.01	-3.91
Bayes	0.98	0.31	0.31	0.01	3.50
Bayes double size	0.28	0.84	0.80	0.54	178.76
Bayes half size	0.80	0.30	0.30	-0.00	-1.43
adjustment for classical error					
Bayes Level 5a	0.55	0.30	0.29	0.00	0.88
SIMEX	0.60	0.26	0.25	-0.04	-13.81

M1b: Measurement model to describe uncertain quantities in underground-mining objects in Thuringia in the first exposure period

$$\begin{gathered}E(t,o,j) = \frac{C_{Rn}(t_0(o_0(o)), o_0(o)) \cdot 12}{A(t_0(o_0(o)), o_0(o))} \cdot t_e(o) \cdot A(t,o)) \cdot \\ g(p_{to}) \cdot w(p_t) \cdot f(p_{oj})\end{gathered}$$

M1b

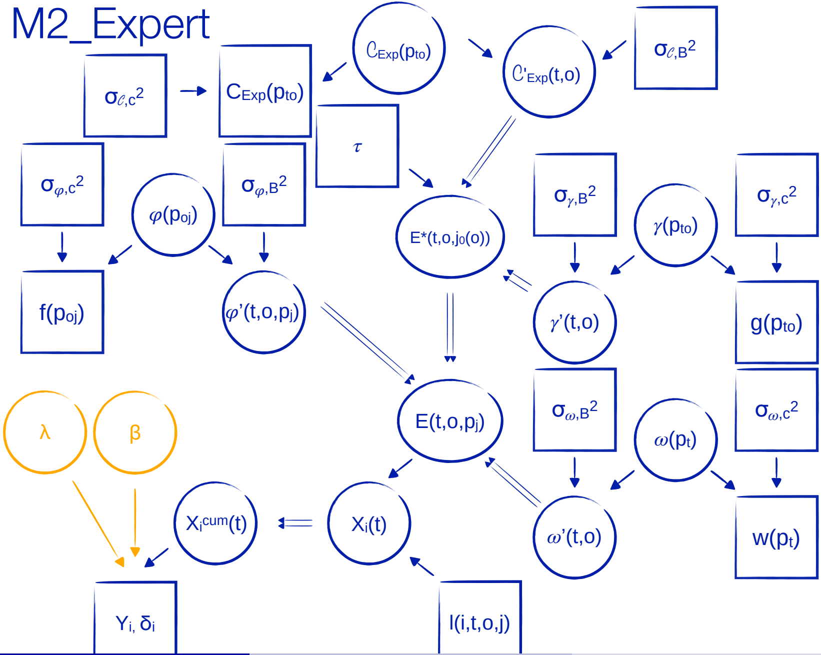


$t_e(o)$
$\tau_e(o)$
$C_{Rn}(t_0(o_0(o)),o_0(o))$
$\tau'_e(t,o)$
$C_{Rn}(t_0(o_0(o)),o_0(o))$
$e(t,o)$
$A(t,o)$
$A(t_0(o_0(o)),o_0(o))$
$E^M(t,o, j_0(o))$
$\gamma(p_{to})$
$g(p_{to})$
$f(p_{oj})$
$E^*(t,o,j_0(o))$
λ
β
$\varphi(p_{oj})$
$\varphi'(t,o,p_j)$
$\gamma'(t,o)$
$\omega(p_t)$
$X_i^{cum}(t)$
$X_i(t)$
$E(t,o,p_j)$
$\omega'(t,o)$
$w(p_t)$
Y_i, δ_i
$I(i,t,o,j)$
τ

M2_Expert: Measurement model to describe uncertain quantities in underground-mining objects in the second exposure period

$$E(t,o,j) = C_{Exp}(p_{to}) \cdot 12 \cdot g(p_{to}) \cdot w(p_t) \cdot f(p_{oj})$$

M2_Expert



$C_{Exp}(p_{to})$

$\mathcal{C}'_{Exp}(t,o)$

$\sigma_{\mathcal{C},B}{}^2$

$\sigma_{\mathcal{C},c}{}^2$

$C_{Exp}(p_{to})$

τ

$\sigma_{\varphi,c}{}^2$

$\varphi(p_{oj})$

$\sigma_{\varphi,B}{}^2$

$E^*(t,o,j_0(o))$

$\sigma_{\gamma,B}{}^2$

$\gamma(p_{to})$

$\sigma_{\gamma,c}{}^2$

$f(p_{oj})$

$\varphi'(t,o,p_j)$

$\gamma'(t,o)$

$g(p_{to})$

λ

β

$E(t,o,p_j)$

$\sigma_{\omega,B}{}^2$

$\omega(p_t)$

$\sigma_{\omega,c}{}^2$

$X_i^{cum}(t)$

$X_i(t)$

$\omega'(t,o)$

$w(p_t)$

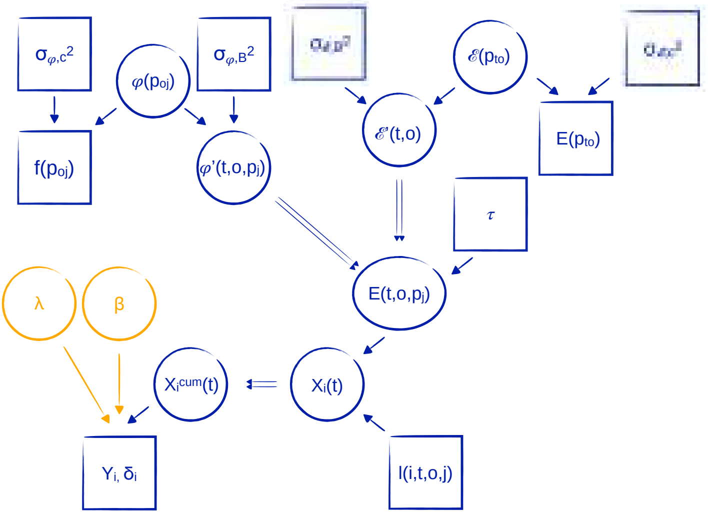
Y_i, δ_i

$I(i,t,o,j)$

M4: Measurement model to describe uncertain quantities in surface areas affiliated to mining and in exploration objects in Thuringia

$$E(t,o,j) = f(p_{oj}) \cdot E(p_{to})$$

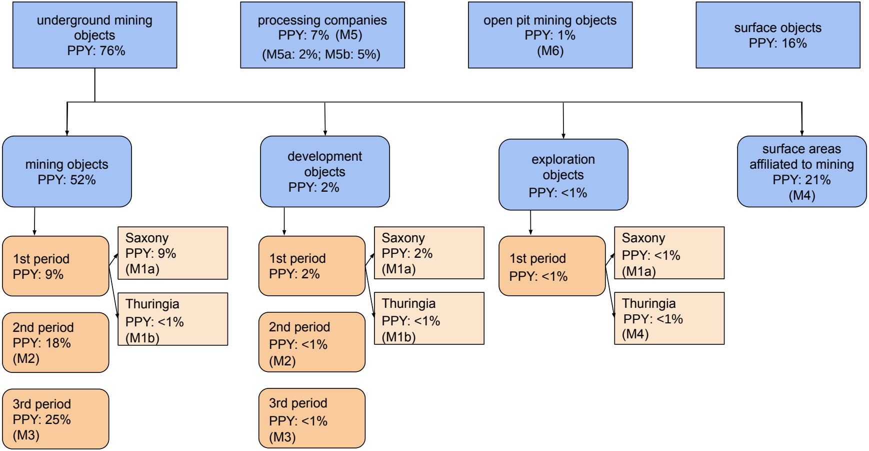
M4/MX_Expert_WLM



M6: Measurement model to describe uncertain quantities in open pit mining objects

$$
\begin{aligned}
E(t, o, j(o)) = & \frac{12}{3700} \left(C_{Rn,0}(1994/1995, 300) + \right. \\
& (C_{Rn,130}(1994/1995, 300) - C_{Rn,0}(1994/1995, 300)) \frac{d(t, o)}{130} \cdot \\
& \left. e(p_{to}) \cdot e_2(p_{to}) \right) \cdot \\
& g(p_{to}) \cdot w(p_t) \cdot f(p_{tj})
\end{aligned}
$$

Measurement models in the Wismut cohort



Exposure assessment in the Wismut cohort [Küchenhoff et al., 2018]

Küchenhoff, H., Deffner, V., Aßenmacher, M., Neppl, H., Kaiser, C., Güthlin, D. et al. (2018). Ermittlung der Unsicherhieten der Strahlenexpositionsabschätzung in der Wismut-Kohorte - Teil I - Vorhaben 3616S12223. Resssortforschungsberichte zum Strahlenschutz. Bundesamt für Strahlenschutz (BfS).