

Quantifizierung der Unsicherheit für einen dünnbesetzten maschinellen Lerndatensatz (ML) in der zerstörungsfreien Prüfung im Bauwesen (NDT-CE)

Christoph VÖLKER¹, Philipp BENNER¹, Sabine KRUSCHWITZ²

¹ Bundesanstalt für Materialforschung und -prüfung (BAM), Berlin

² Bundesanstalt für Materialforschung und -prüfung (BAM) und TU Berlin, Berlin

Kontakt E-Mail: christoph.voelker@bam.de

Kurzfassung

Das maschinelle Lernen (ML) wurde erfolgreich zur Lösung vieler Aufgaben in der zerstörungsfreien Prüfung im Bauwesen (ZfPBau) eingesetzt. Allerdings ist die Erstellung von Referenzdaten in den meisten Fällen extrem teuer und daher viel knapper als in anderen Forschungsbereichen. Auch decken die verfügbaren Daten mitunter nur ein einziges Szenario ab, so dass die Leistungsindikatoren oft nicht die tatsächliche Leistung des ML-Modells in der praktischen Anwendung widerspiegeln. Schätzungen, die die Übertragbarkeit von einem Szenario auf ein anderes quantifizieren, sind erforderlich, um dieser Herausforderung gerecht zu werden und den Weg für Anwendungen in der Praxis zu ebnet.

In diesem Beitrag stellen wir Werkzeuge zur Beschreibung der Unsicherheit von ML in neuen ZfPBau-Szenarien vor. Zu diesem Zweck haben wir einen bestehenden Trainingsdatensatz zur Klassifizierung von Korrosionsschäden der Bewehrung in Beton um eine neue Fallstudie erweitert. Die Messungen wurden an großflächigen Betonproben mit eingebauter chloridinduzierter Korrosion des Bewehrungsstahls durchgeführt. Das Experiment simulierte den gesamten Lebenszyklus von chloridinduzierten Sichtbetonbauteilen im Labor. Unser Datensatz umfasst Potenzialfeld- und Radarmessungen. Die einzigartige Fähigkeit, die Schädigung zu überwachen und eine gezielte Korrosion einzuleiten, ermöglichte es, die Daten zu labeln - was für die Konstruktion von ML-Modellen entscheidend ist. Um die Übertragbarkeit zu untersuchen, erweitern wir unser Modell um Metadaten - wie etwa Konstruktionsmerkmale des Prüfkörpers und Umweltbedingungen. Dies erlaubt es, die Veränderung dieser Merkmale in neuen Szenarien mit statistischen Methoden als Unsicherheiten auszudrücken. Wir vergleichen verschiedene auf Stichproben und statistischer Verteilung basierende Ansätze und zeigen, wie diese Methoden eingesetzt werden können, um Wissenslücken von ML-Modellen in der ZfP zu schließen.



Sicherheit in Technik und Chemie

10. Mai 2021

QUANTIFIZIERUNG DER UNSICHERHEIT FÜR EINEN DÜNNBESETZTEN MASCHINELLEN LERNDATENSATZ IN DER ZERSTÖRUNGSFREIEN PRÜFUNG IM BAUWESEN

Dr. Christoph Völker¹, Dr. Philipp Benner¹, Prof. Sabine Kruschwitz²

1) BAM Federal Institute for Materials Research and Testing, Berlin, Germany
Tel.: +49 (30) 8104-1440; Christoph.Voelker@Bam.de

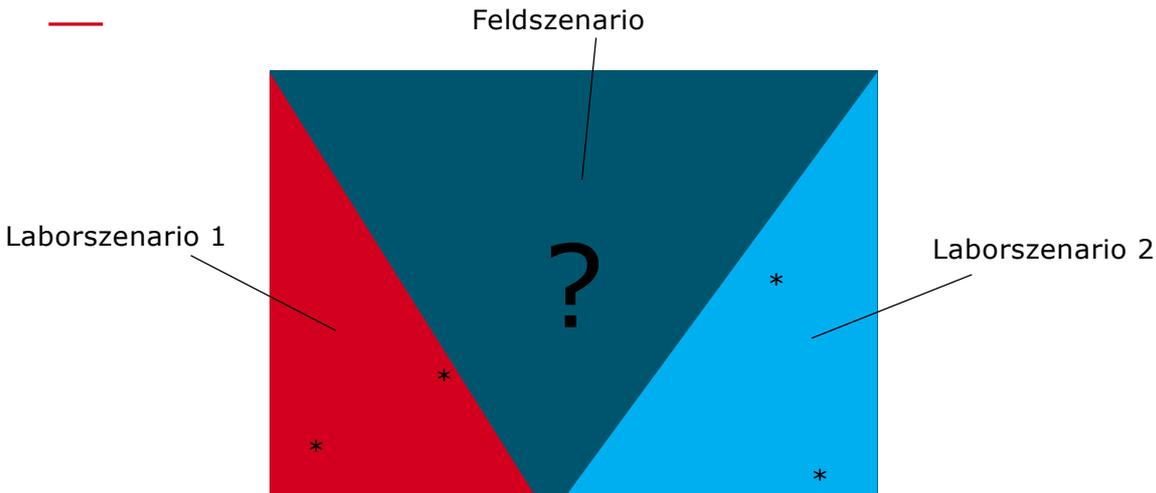
2) Technische Universität Berlin, Berlin
Tel.: +49 30 314 721 109; kruschwitz@tu-berlin.de



1

Problem

Vorhersagen in neuen Szenarien mit Machine Learning



Feldszenario

Laborszenario 1

Laborszenario 2



2

Problem
Vorhersagen in neuen Szenarien mit Machine Learning

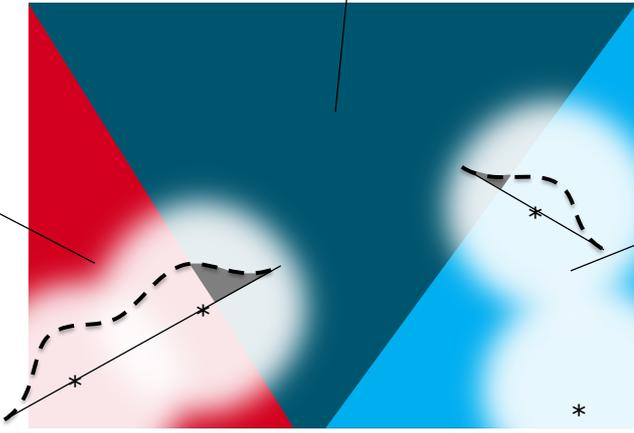


—

Feldszenario

Laborszenario 1

Laborszenario 2



3

3

Anwendungsbeispiel chloridinduzierte Lochkorrosion



—

Betonquerschnitt mit Lochkorrosion an Bewehrung



4

4

Anwendungsbeispiel chloridinduzierte Lochkorrosion

Surveys in Geophysics
<https://doi.org/10.1007/s10712-019-09558-4>

A Machine Learning-Based Data Fusion Approach for Improved Corrosion Testing

Christoph Völker¹ Sabine Kruschwitz^{1,2} · Gino Ebell¹

Received: 26 September 2018 / Accepted: 15 July 2019
 © Springer Nature B.V. 2019

Abstract
 This work presents machine learning-inspired data fusion approaches to improve the non-destructive testing of reinforced concrete. The principal effects that are used for data fusion are shown theoretically. Their effectiveness is tested in case studies carried out on large-scale concrete specimens with built-in chloride-induced rebar corrosion. The dataset consists of half-cell potential mapping, Wenner resistivity, microwave moisture and ground penetrating radar measurements. Data fusion is based on the logistic regression algorithm.

ASTM-C876

Data Fusion

1 korrekt
11 unsicher
6 falsch

15 korrekt
3 falsch

5

5

Referenzdaten

7

Merkmale

2016

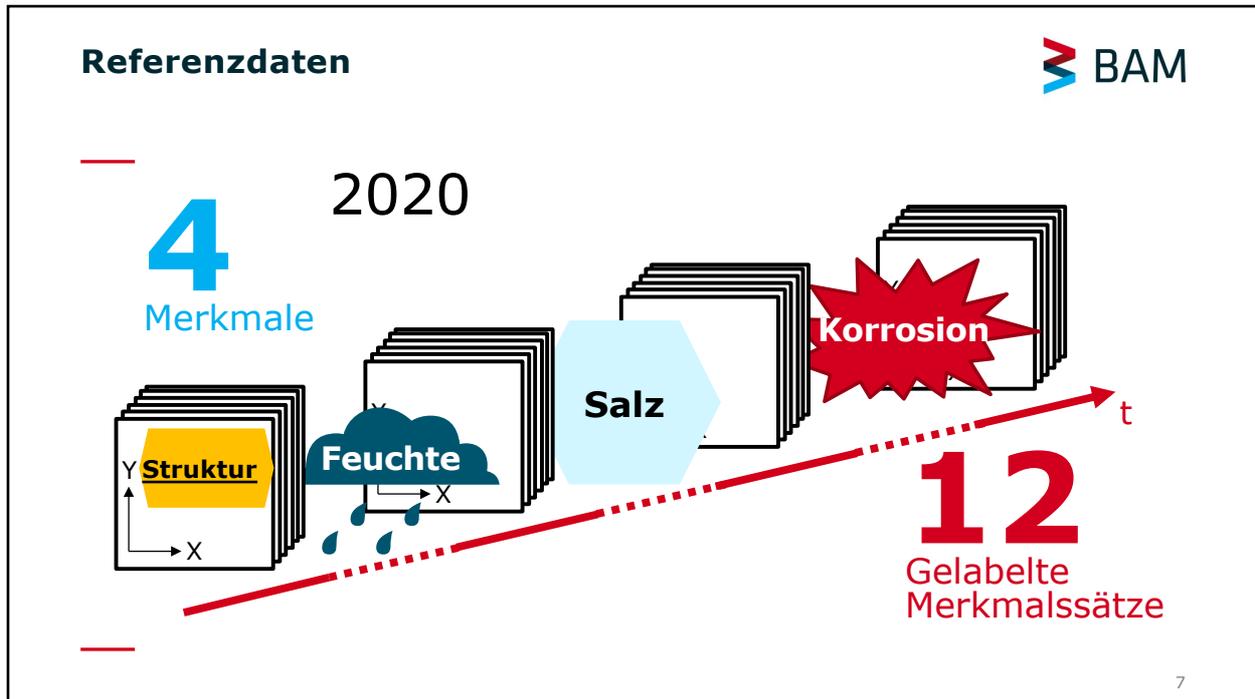
18

Gelabelte
Merkmalsätze

Data set available online: DOI
 10.13140/RG.2.2.24881.89448

6

6



7

Variabilität der Einflussparameter auf die ZfP Messung 

Einflussparameter	2016	2020 (hinzugefügte Szenarien)
Feuchtigkeitszustand	3	2 (0)
Chloridmigration in Betondeckung	2	2 (0)
Betonverdichtung	2	1 (0)
Bewehrungsdichte	1	2 (1)
Betondeckung	1	3 (3)
Korrosionsposition (obere/untere Bewehrungslage)	1	2 (1)
Max. Anz. der Korrosionsstellen je Messung	1	2 (1)
Alter der Probe	<5 a	<5 a (0)
Lastwechsel	0	0 (0)
...

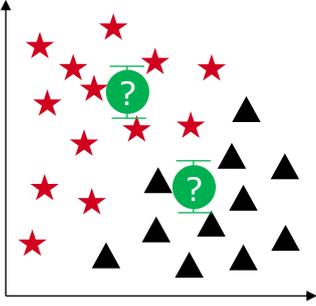
8

8

Unsicherheiten kleiner Referenzdatensätze BAM

Klassische Data Science*

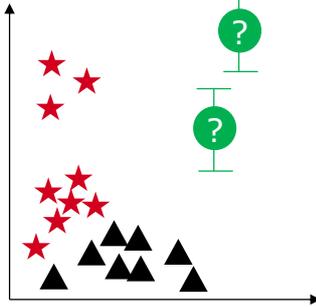
- Aleatorische Unsicherheit



*Ansatz: Independent sampling from Identical Distributed data (IID)

Small Data (Extrapolation)

- Aleatorischer Unsicherheit
- Epistemischer Unsicherheit



9

9

Adaptive Modelselektion auf Basis epistemischer Unsicherheitsberechnungen BAM

Trainingsdaten ? **Model** ➔ **Vorhersage**

Modell mit höchster Erfolgswahrscheinlichkeit
 $\min(\text{Fehler} + \text{Unsicherheit})$

Modell mit niedrigsten Fehlererwartungswert
 $\min(\text{Fehler})$

2 Jahre (~200k€)
4 ZfP-Verfahren

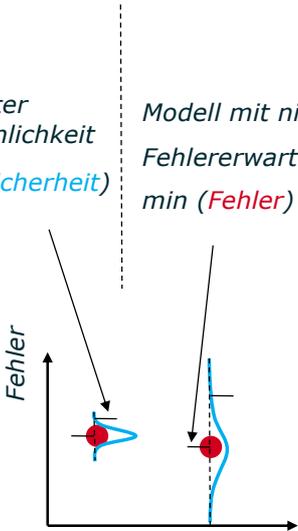


3 Jahre (~200k€)
2 ZfP-Verfahren



Felddaten

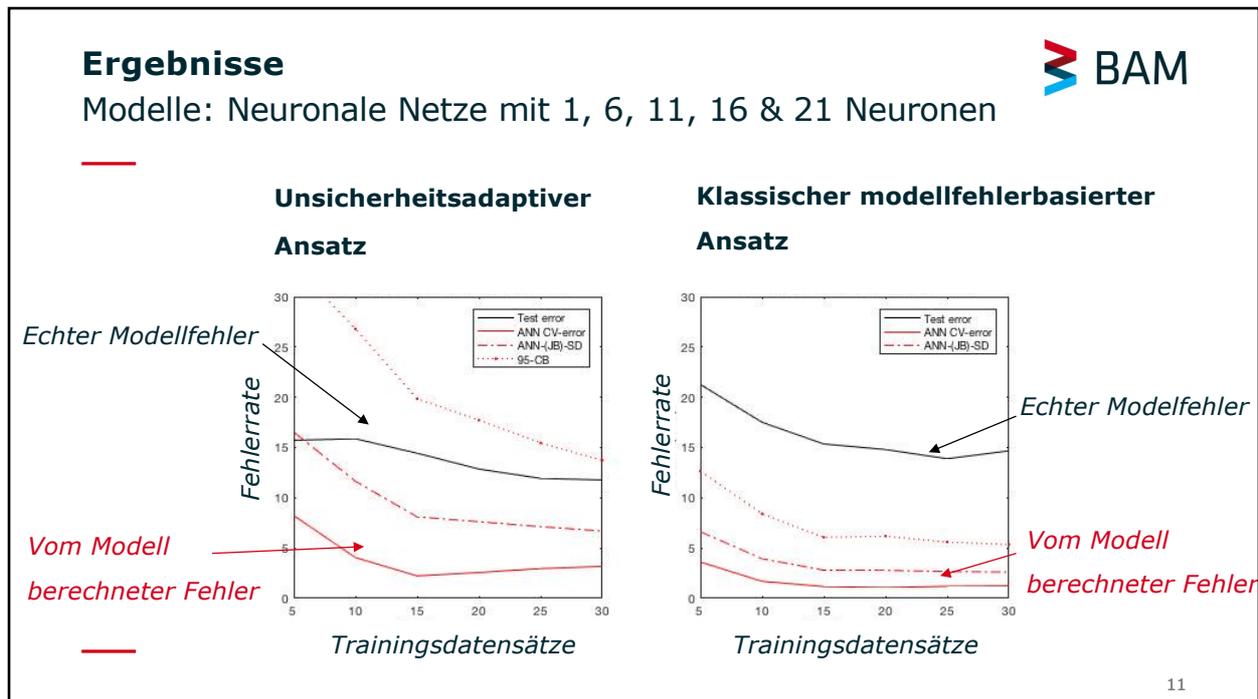




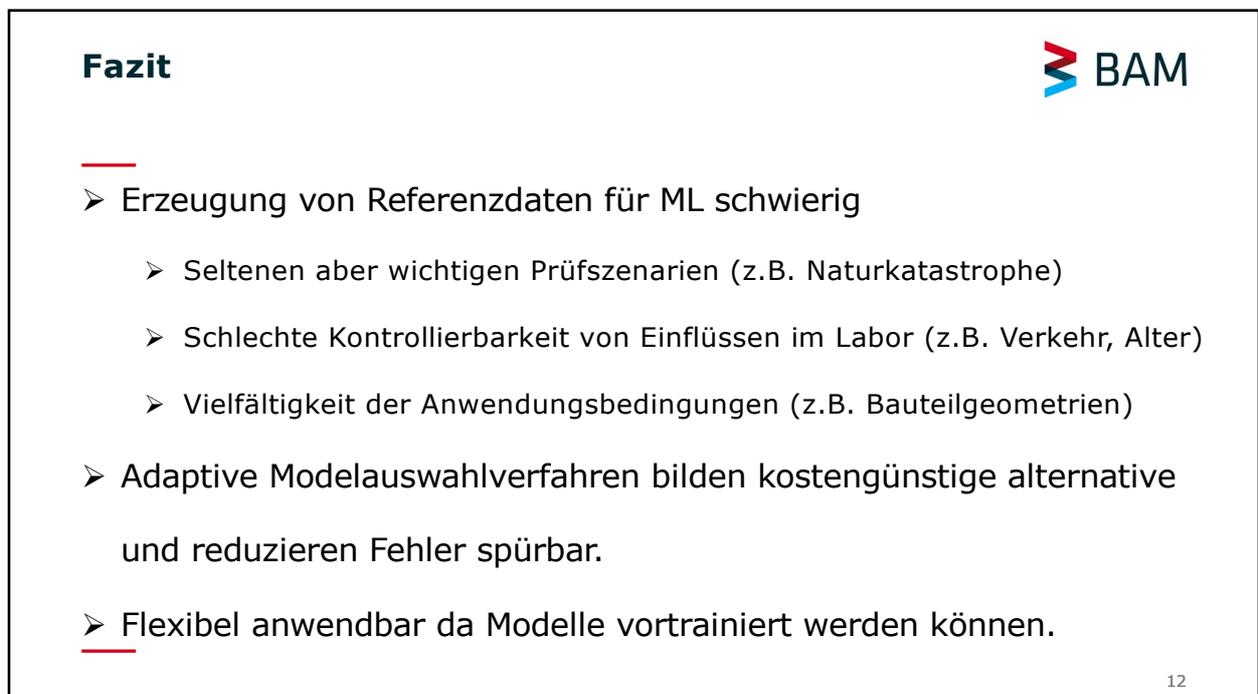
10

10

10



11



12



Vielen Dank für Ihre Aufmerksamkeit

13

The slide features a background with abstract geometric shapes in shades of gray, blue, and red. Two small red horizontal lines are positioned on the left side of the slide.