

Big Data und Machine Learning
bei Schadenzahlprozessen:
Möglichkeiten und Grenzen

Dr. Olaf Kruse

DAV/DGVFM Jahrestagung 2023 - 27.04.2023

1 Agenda

Einführung

1. Einführung

2. Motivation

3. Modellbewertung

4. Ergebnisanalyse

5. Zusammenfassung

... keine Agenda

... lediglich ein Reisebericht
durch die Untiefen von
Big Data und ML

... mit ungewissem Ausgang
und einer offenen Frage !

1 Schadenzahlprozesse

Einführung

Zweistufiger Zufalls-Prozess

- ⇒ 1. Stufe: Es tritt ein (Schaden-) Ereignis ein (oder nicht) ⇒ z.B. Poisson-Prozess
- ⇒ 2. Wie „teuer“ ist das Ereignis, wenn es eintritt ⇒ z.B. Gamma-Verteilung

Nicht untypisch

- ⇒ Online-Shopping: Kaufentscheidung ja/nein, Größe des Warenkorb
- ⇒ Kreditrisiko: Kreditausfall ja/nein, Ausfallvolumen
- ⇒ Versicherungen: Schadenereignisse



1 Big Data Einführung

Big Data is the frontier of a firm's ability to store, process, and access all the data it needs to operate effectively, make decisions, reduce risks, and serve customers.
[Forrester Research]



1 Machine Learning

Einführung



Mithilfe **statistischer Verfahren** werden **Statistiker** in die Lage versetzt, auf Basis vorhandener Datenbestände und Algorithmen Muster und Gesetzmäßigkeiten zu erkennen und Lösungen zu entwickeln.

Die aus den Daten gewonnenen Erkenntnisse lassen sich verallgemeinern und für neue Problemlösungen oder für die Analyse von bisher unbekanntem Daten verwenden.

1 Machine Learning

Einführung

Mithilfe des **maschinellen Lernens** werden **IT-Systeme** in die Lage versetzt, auf Basis vorhandener Datenbestände und Algorithmen Muster und Gesetzmäßigkeiten zu erkennen und Lösungen zu entwickeln.

Die aus den Daten gewonnenen Erkenntnisse lassen sich verallgemeinern und für neue Problemlösungen oder für die Analyse von bisher unbekanntem Daten verwenden.



⇒ **Machine Learning:** Summe aller statistischen Optimierungsalgorithmen

... nur ohne Statistik

2 Daten- und Methoden-Setting

Motivation

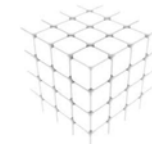
Tarifikalkulation in der Kraftfahrtversicherung

- ⇒ Wie hoch ist der Schadenbedarf (Prämie) für ein zu versicherndes Risiko
- ⇒ Datenbestand von ca. 10 Mio. Risiken p.a.
- ⇒ ⇒ ⇒ Viele Daten und große / komplexe Modelle



Methodischer Ansatz

- ⇒ Ordinalskaliert Merkmale, die einen n-Dimensionalen Tarifwürfel aufspannen
- ⇒ Verallgemeinerte Lineare Modelle in allen Facetten



$$P_{\lambda}(k) = \frac{\lambda^k}{k!} e^{-\lambda}$$

2 Entwicklung der Kraftfahrt-Modelle

Motivation

„Erstes“ Tarif-Modell (nach Tarif-Freigabe 1992) **100.000 Tarifzellen**
(SF-Klasse – Tarif-Gruppe – Typ-Klasse – Regionalklasse – Garage – Fahrleistung)

„Normales“ GDV-Tarif-Modell **0,35 Mrd. Tarifzellen**
(+ Nutzer-Alter + Kfz.-Alter bei Erwerb + Nutzerkreis + Wohngebäude + Halter)

„Fortgeschrittenes“ Modell **45 Mrd. Tarifzellen**
(+ Rabattschutz + Zahlweise + Mahnverfahren + Antriebsart)

Theoretisches XXL-Modell **9.000 Mrd. Tarifzellen**
(+ Dauer der Kundenbeziehung + Erst/Zweitwagen + Fahrzeugtyp (Cabrio, SUV, ...) + ...)

VST Kraftfahrt-Bestand **0,01 Mrd. Risiken**
Anzahl Schäden p.a. **0,001 Mrd. Schäden**

2 Telematik-Projekt

Motivation

Masterarbeit:
**Foundations of modern regression analysis and
application to the analysis of telematics data**



Ergebnis: relativ stabiles Modell auf Basis von 6 Telematik-Parametern
„Vielfahrer“ – „Kurzstreckenpendler“ – „Gelegenheitsfahrer“

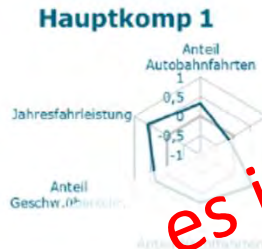
Wie geht das mit wenigen 1.000 Telematik-Fahrern, die keine 100 Schäden produzieren?

- ⇒ Es bestand gar **kein** Zugriff auf Schadendaten
- ⇒ Substituierung durch die Versicherungsprämie aus dem „normalen“ Tarif
- ⇒ „Selbsterfüllende Prophezeiung“

2 Telematik-Projekt

Motivation

Masterarbeit: Foundations of modern regression analysis and application to the analysis of telematics data



Ergebnis:

relativ stabiles Modell auf Basis von 6 Telematik-Parametern

„Vielfahrer“ – „Kurzstreckenpendler“ – „Gelegenheitsfahrer“

Wie geht das mit wenigen 1.000 Telematik-Fahrern, die keine 100 Schäden produzieren?

- ⇒ Es bestand gar kein Zugriff auf Schadendaten
- ⇒ Substituierung durch die Versicherungsprämie aus dem „normalen“ Tarif
- ⇒ „Selbsterfüllende Prophezeiung“

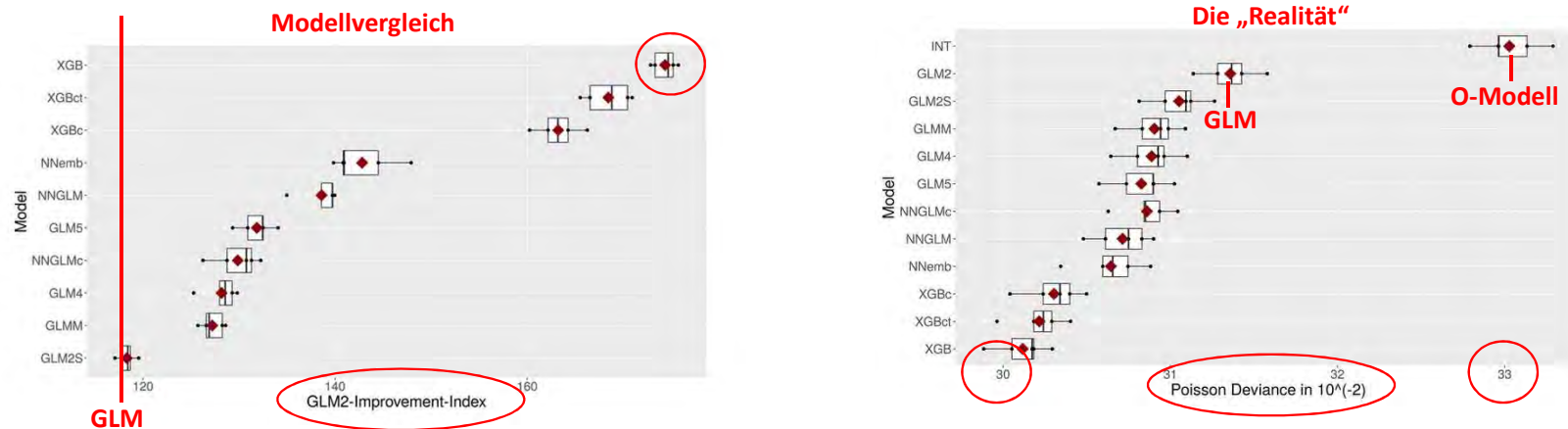
2 Machine Learning-Projekt

Motivation

GLM, Neural Nets and XGBoost for Insurance Pricing

Verfahrensvergleich an einem frei zugänglichen kleinen KFZ-Datensatz

⇒ Klassisches Big-Data Setting: Trainings- und Test-Daten für Modellbildung und -bewertung



⇒ „Moderne“ Verfahren passen die Daten bis zu 80% besser an als klassische GLMs

⇒ Aber: Alle Modelle performen sehr schlecht (Erklärung von ca. 10% der Gesamt-Devianz)

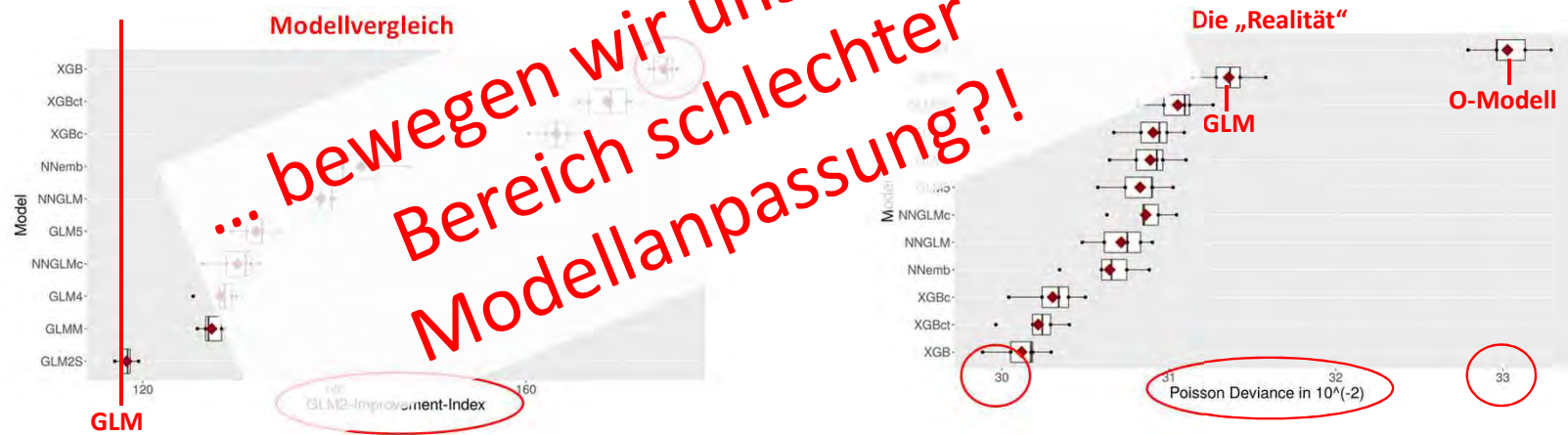
2 Machine Learning-Projekt

Motivation

GLM, Neural Nets and XGBoost for Insurance Pricing

Verfahrensvergleich an einem frei zugänglichen kleinen KFZ-Datensatz

⇒ Klassisches Big-Data Setting: Trainings- und Test-Daten für Modellbildung und -bewertung



⇒ „Moderne“ Verfahren passen die Daten bis zu 80% besser an als klassische GLMs

⇒ Aber: Alle Modelle performen sehr schlecht (Erklärung von ca. 10% der Gesamt-Devianz)

2 Können wir Big-Data?!

Motivation

Entwicklung eines **maximalen Risikomodells**

- ⇒ Keine Rücksichtnahme auf **Umsetzbarkeit** in einen Tarif
(weder rechtlich, noch technisch, noch Marktgängigkeit oder Organik)
- ⇒ Zusammenfassung **mehrerer Jahre** als Grundlage der **Modellentwicklung**
(systematische Effekte als zeit- und zufallsstabil identifizieren)
- ⇒ Entwicklung effizienter **Modellbewertungsstrategien**

3 Modellierungsstrategien

Modellbewertung

Sukzessive Aufnahme neuer Merkmale

| Basis-Modell | Modell 1 | Modell 2 | Modell 3 | Modell 4 |
|------------------------|-------------|-------------|------------------|----------------------|
| Erweitertes VST-Modell | + Kfz-Alter | + Mahnverf. | + Erst-/Zweit-FZ | + Antrieb/Sitzplätze |

Modellierung über verschiedene Statistikjahre

| Modell 4 | Modell 4 | Modell 4 | Modell 4 | Modell 4 | Modell 4 |
|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| Stat.-Jahr 2018 | Stat.-Jahr 2017 | Stat.-Jahr 2016 | Stat.-Jahr 2015 | Stat.-Jahr 2014 | Stat.-Jahr 2013 |

⇒ Unterschiedliche Sichtweise, aber vergleichbare Problemstellung

3 Prinzipielle Ansätze

Modellbewertung

Vergleich der Modellparameter („Old School“)

- ⇒ Parameter eines Modell über verschiedenen Datensätze (Statistikjahre)
- ⇒ Auswirkung neuer Merkmale auf „alte“ Parameter
- ⇒ ggf. „klassische“ Inferenz-Statistiken („Statistische Tests“)

Bewertung der Modellanpassung („Scoring“)

- ⇒ Anpassung des Modells an Trainings-Datensatz
- ⇒ Anpassung des Modells an Test-Datensatz (z.B. neues Statistikjahr)
- ⇒ Veränderung der Anpassung durch Aufnahme neuer Merkmale



3 Anpassungsmasse

Modellbewertung

Typische Anpassungsmaße

| Anpassungsmaß | Trainingsdaten | Testdaten |
|-------------------------------------|----------------|---------------|
| -2 Log Likelihood | 80,549,463 | 90,746,279 |
| AIC (kleiner ist besser) | 80,549,597 | 90,746,413 |
| AICC (kleiner ist besser) | 80,549,597 | 90,746,413 |
| BIC (kleiner ist besser) | 80,550,630 | 90,747,442 |
| Pearson Chi-Quadrat | 3,517,758,153 | 3,693,232,867 |
| Pearson Chi-Quadrat / DF | 1,000.9 | 1,020.8 |
| Durchschnittlicher absoluter Fehler | 382.7 | 396.4 |

⇒ Berechnung über alle Zellen jeweils für Trainings- und Test-Datensatz

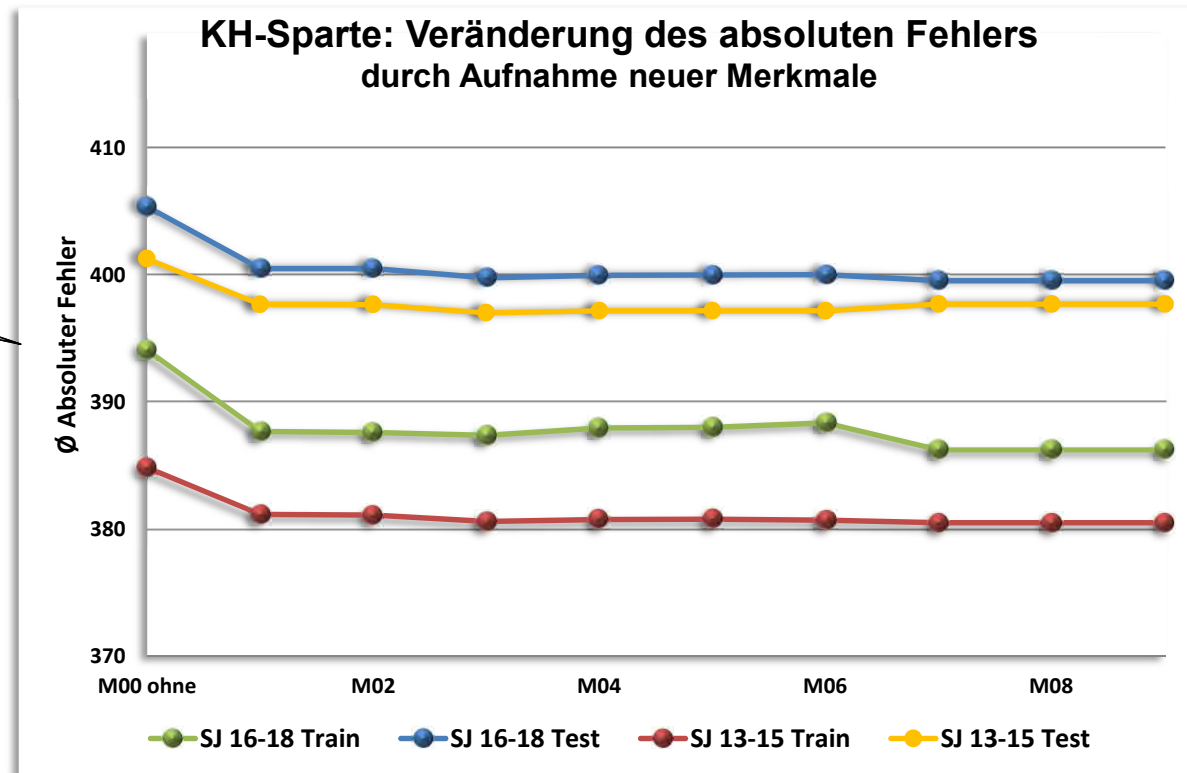
⇒ 50:50 Aufteilung Statistikjahre 2013-2015 und 2016-2018



3 Modellanpassung

Modellbewertung

Ø absoluter Fehler
größer als
Ø Schadenbedarf



⇒ Geringe Unterschiede zwischen Trainings- und Test-Daten (++)

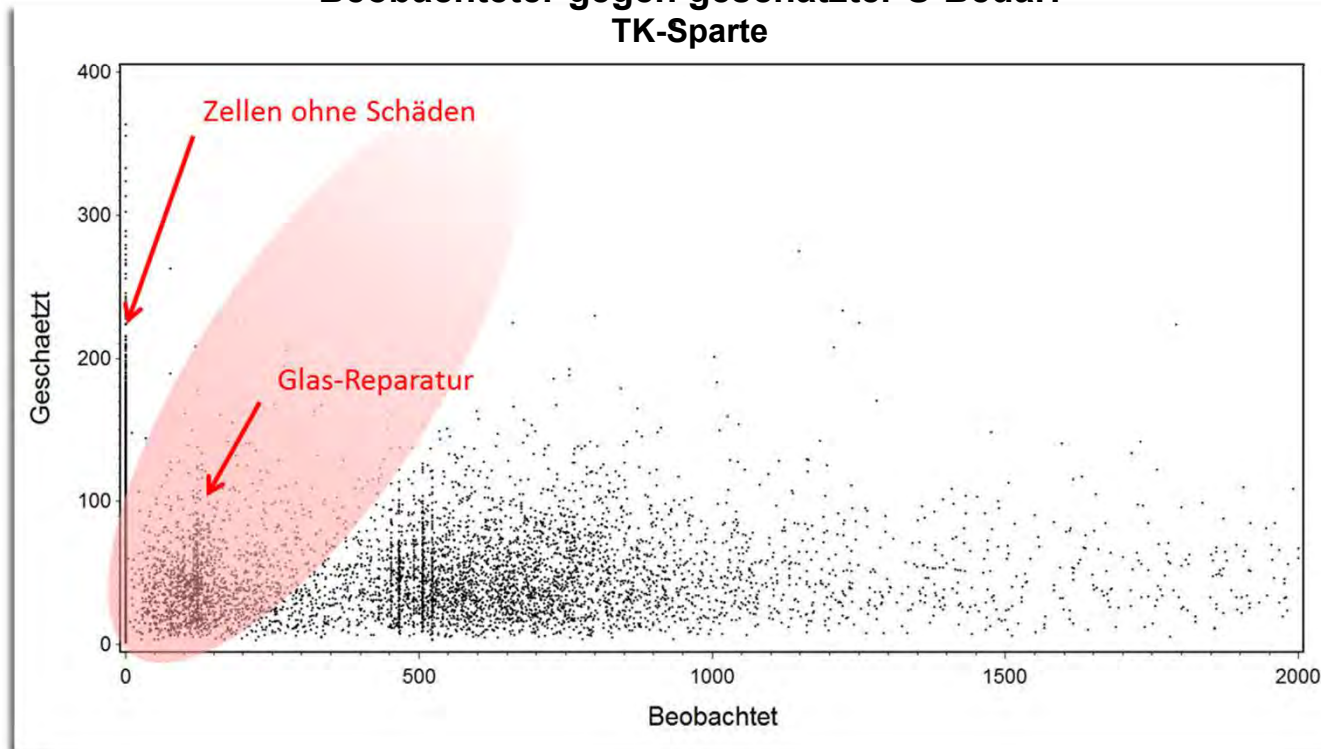
⇒ Marginale Modellverbesserung durch Aufnahme weiterer Merkmale (--)

VST

4 Prognosevergleich

Ergebnisanalyse

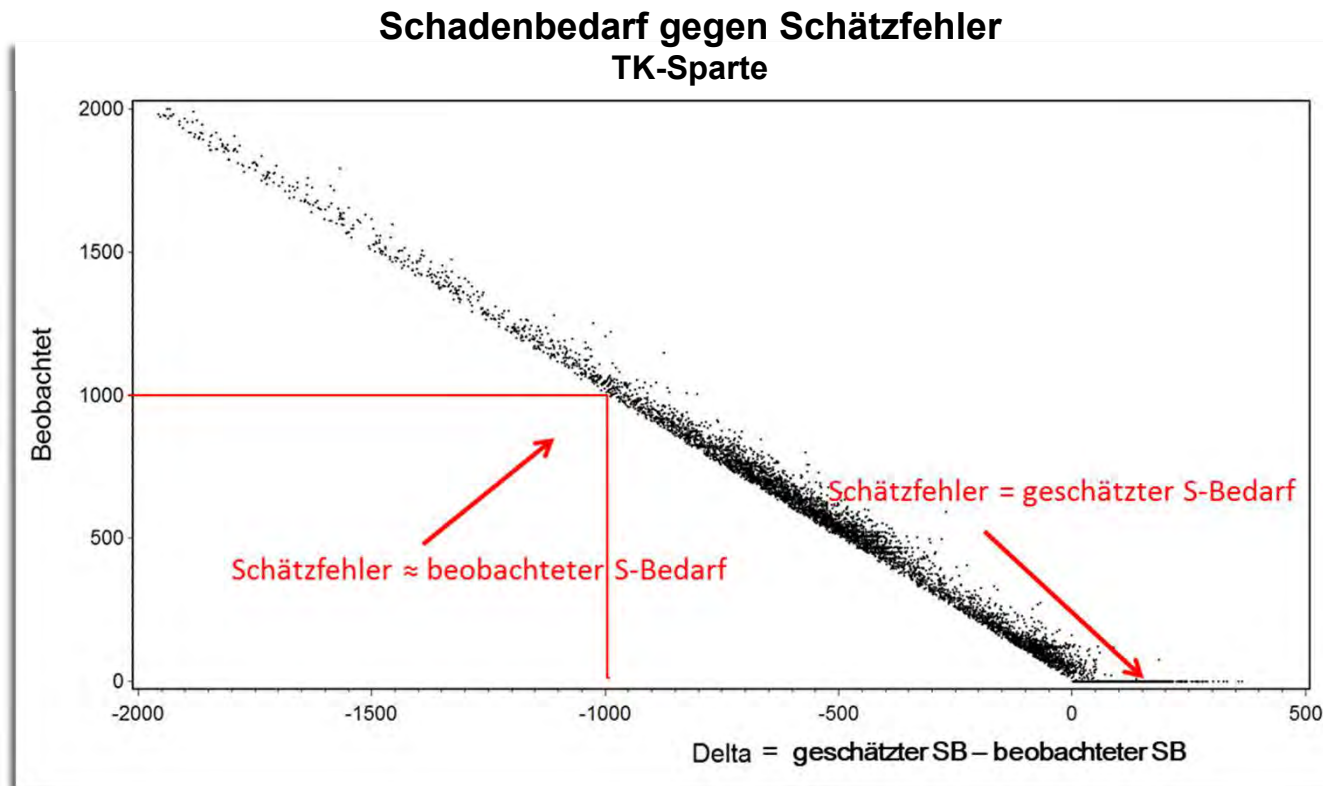
Beobachteter gegen geschätzter S-Bedarf
TK-Sparte



⇒ Keine positive Korrelation zwischen Schätzung und Beobachtung

4 Struktur der Schätzfehler

Ergebnisanalyse

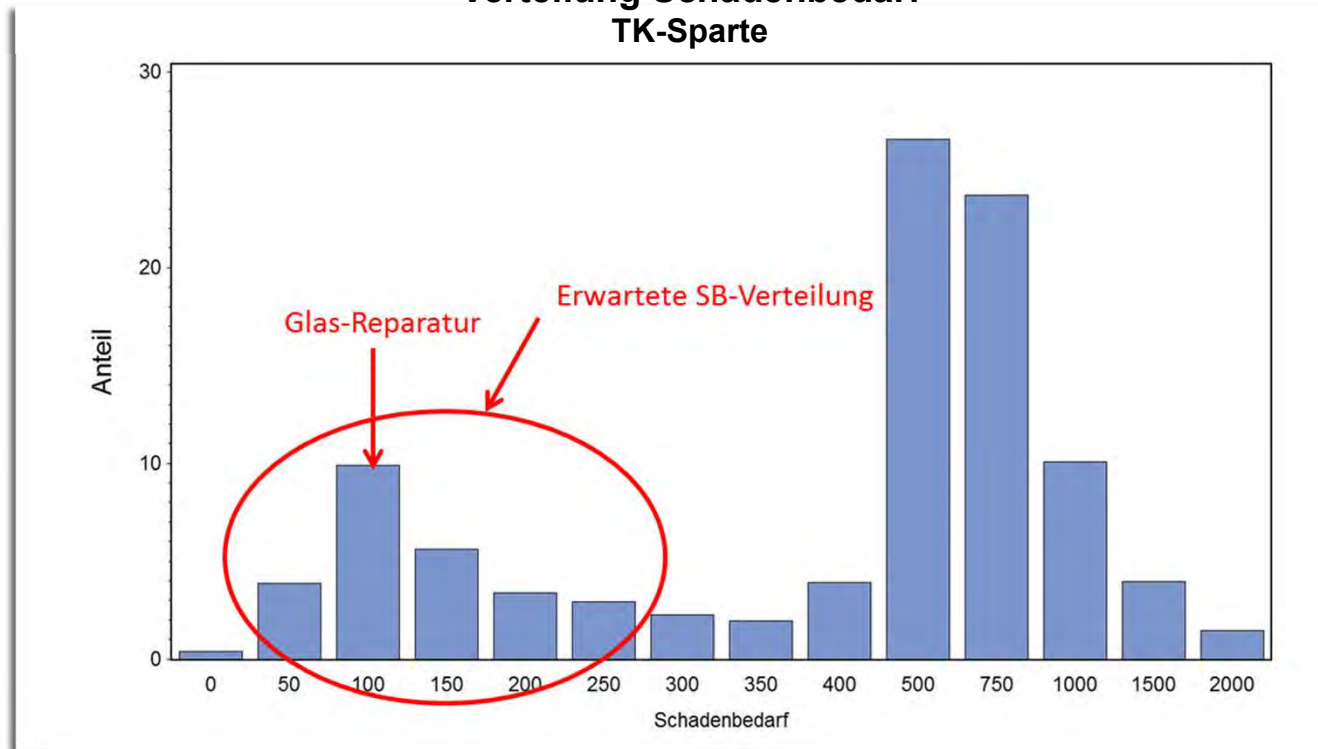


⇒ Kaum Beobachtungen mit geringem Schätzfehler

VST

4 Verteilung Ergebnisanalyse

Verteilung Schadenbedarf TK-Sparte

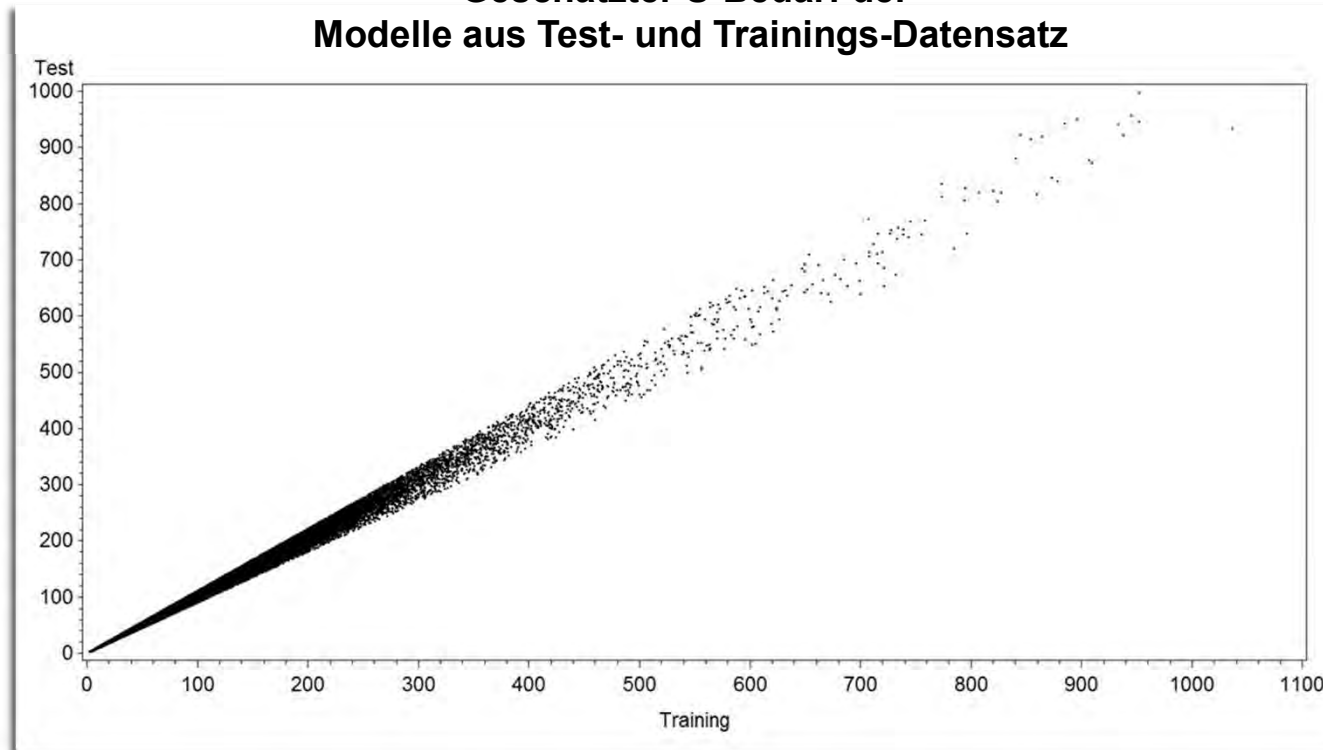


⇒ S-Bedarfs-Verteilung entspricht der S-Höhenverteilung

4 Test- und Trainingsdaten

Ergebnisanalyse

**Geschätzter S-Bedarf der
Modelle aus Test- und Trainings-Datensatz**

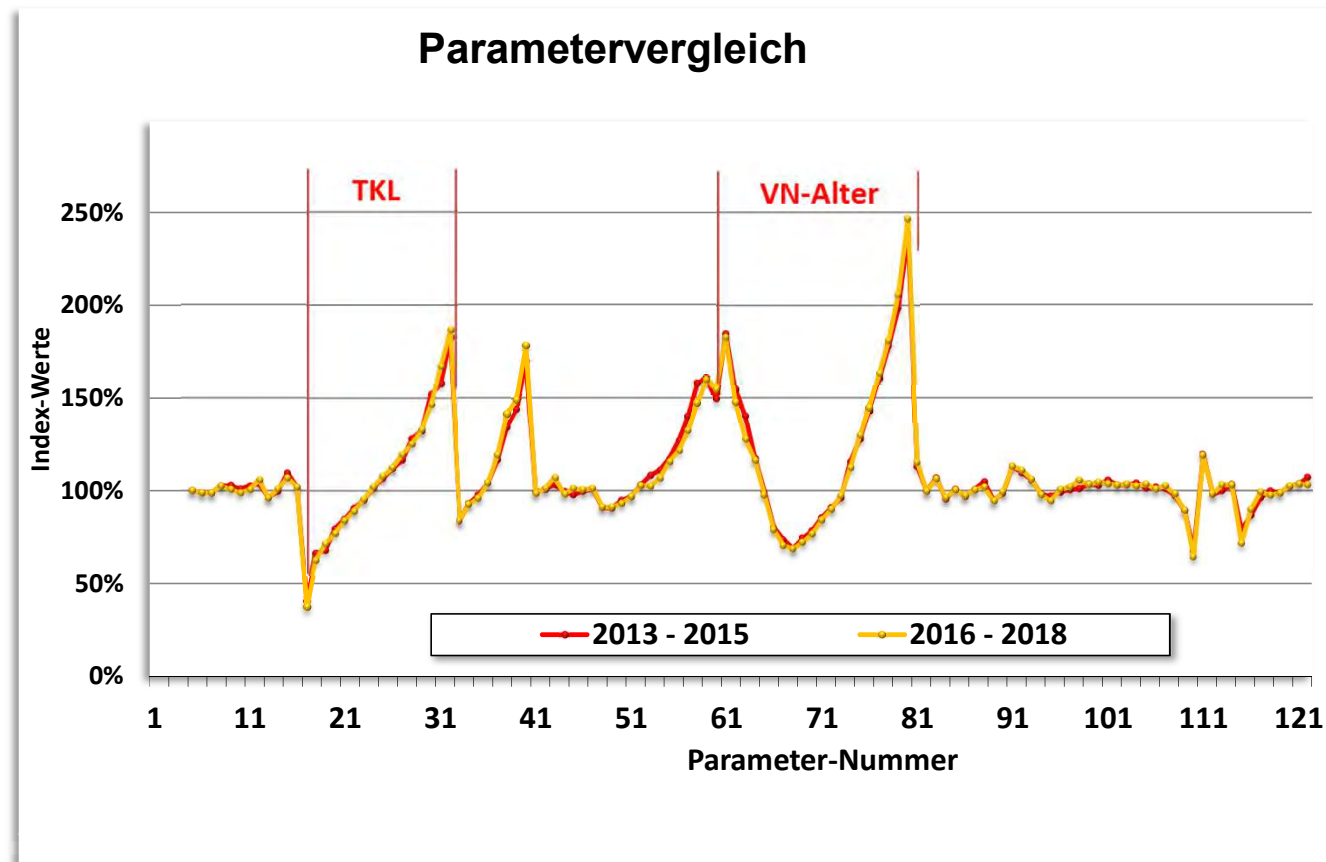


⇒ Stabile Schätzergebnisse zeigen sich auch im Parameter-Vergleich

VST

4 Modellparameter

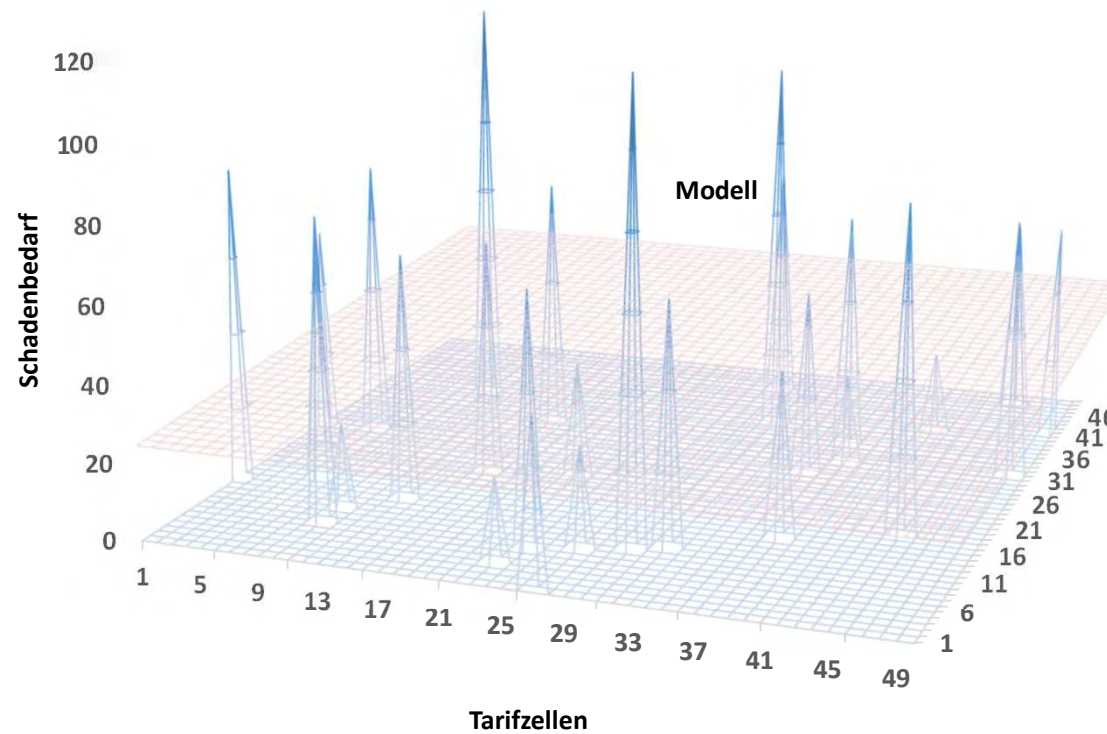
Ergebnisanalyse



⇒ Trotz alledem: „gute“ Modelle mit stabilen Modellparametern

5 Schlussfolgerung

Zusammenfassung



⇒ Schadenzahlprozesse sind mit „modernen“ ML-/Big-Data-Methoden schwer greifbar

5 Schlussfolgerung

Zusammenfassung

(Big-) Datensituation

- ⇒ Anzahl Tarifzellen >> Anzahl Beobachtungen >> Anzahl Schäden
- ⇒ pro Zelle **nicht aussagekräftig**/repräsentativ
- ⇒ Perfekte Anpassung an diese Daten (-struktur) kann **nicht Ziel** der Modellierung sein

(ML-) Modellbewertung

- ⇒ Klassische ~ bewerten Anpassung an o.g. nichtrepräsentative Datenstruktur
- ⇒ Aufteilung in Test- und Trainingsdaten hilft **nicht**
- ⇒ Gefahr von Fehlinterpretationen

Algorithmen

- ⇒ Algorithmen konvergieren bzw. sind performant (CPU-Zeit)
- ⇒ Keine Probleme mit „Sparse Data“ bzw. „Quasi-Separation“

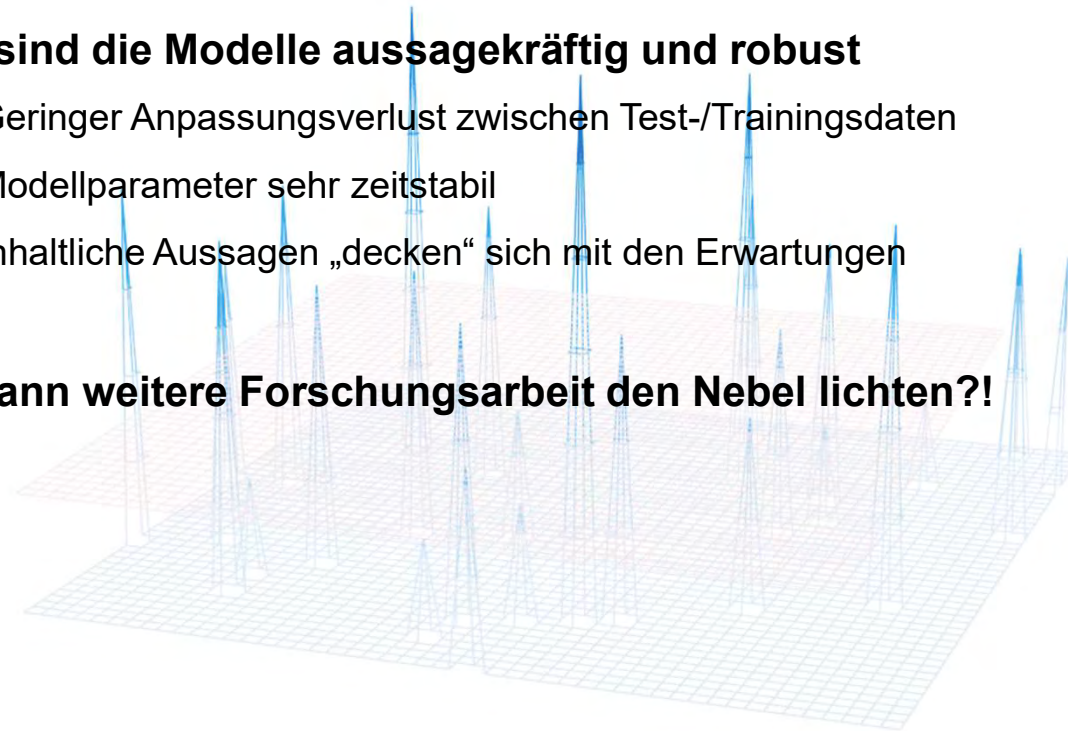
5 Positives Schlusswort

Zusammenfassung

...dennoch sind die Modelle aussagekräftig und robust

- ⇒ Geringer Anpassungsverlust zwischen Test-/Trainingsdaten
- ⇒ Modellparameter sehr zeitstabil
- ⇒ Inhaltliche Aussagen „decken“ sich mit den Erwartungen

⇒⇒⇒ **kann weitere Forschungsarbeit den Nebel lichten?!**



5 Positives Schlusswort

Zusammenfassung

...dennoch sind die Modelle aussagekräftig und robust

⇒ Geringer Anpassungsverbund zwischen Test-/Trainingsdaten

⇒ Modellparameter sehr zeitstabil

⇒ Inhaltliche Aussagen „drücken“ sich mit den Erwartungen

⇒⇒⇒ kann weitere Forschungsarbeit den Nebel lichten?!

... wir (Aktuare) können
„Big Data“,
wenn wir es nur richtig
machen!

5 Fragen bitte !!

Ende

Big Data und Machine Learning bei Schadenzahlprozessen: Möglichkeiten und Grenzen

Dr. Olaf Kruse

VST Gesellschaft für Versicherungsstatistik mbH

Roscherstr. 10

30161 Hannover

olaf.kruse@vst-gmbh.de