Prognoseverfahren für Leistungstrends in der PKV

Dr. Markus Knappitsch (comma soft AG)
Alexander Krauskopf (Central Krankenversicherung AG, Köln)

DAV
DEUTSCHE AKTUARVEREINIGUNG e.V.

qx-Club, 02.05.2017

Agenda

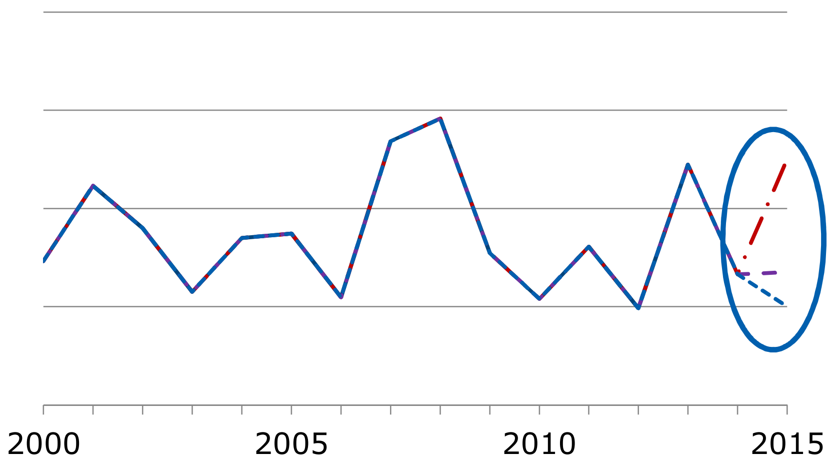
- Motivation: Bedeutung der Leistungstrends in der PKV
- Machine Learning mit künstlichen neuronalen Netzen
- ARMA-Modell aus der Zeitreihenanalyse
- Gegenüberstellung der Ergebnisse der unterschiedlichen Verfahren
- Stochastische Modellierung und Anwendungen in der Praxis
- Fazit

Motivation: Leistungstrends und deren Prognose

DAV
DEUTSCHE AKTUARVEREINIGUNG e.V.

Bedeutung der Leistungstrends in der PKV

Möglicher Verlauf eines Leistungstrends in einem PKV-Portfolio



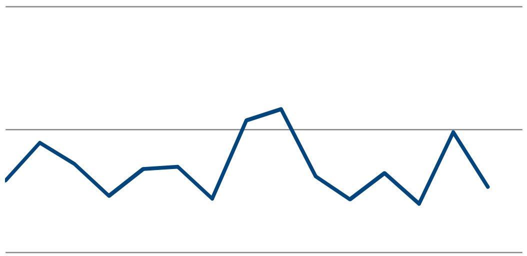
- Leistungstrends in der PKV können sehr volatil sein
- entscheidend für die Durchführung einer Beitragsanpassung (→ Schaden-AF)
- Festlegung des Anpassungsmaßes von großer Bedeutung für zukünftige Überschüsse und späteren Anpassungsbedarf

→Ziel: „Gute" Prognose der Leistungstrends für die Folgejahre

DAV DEUTSCHE AKTUARVEREINIGUNG e.V.

Maschinelles Lernen zur Prognose von Leistungstrends?

Maschinelles Lernen



Klassische Statistik

Statistisches Lernen und klassische Statistik bieten verschiedene Sichtweisen auf unser Zeitreihenanalyseproblem.

Leistungstrends in der PKV: ein Zeitreihenanalyseproblem?

Untersuchte Verfahren zur Prognose der Leistungstrends:

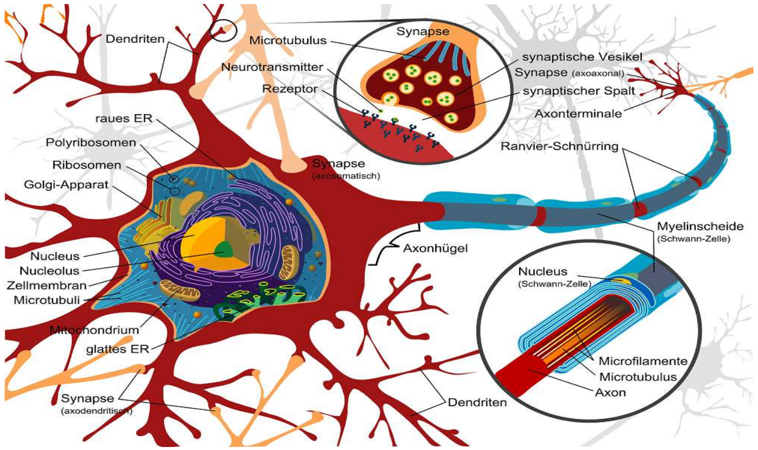
- Klassische Schätzverfahren (lineare Regression)
- **Zeitreihenanalyse mit künstlichen neuronalen Netzen (KNN)**
- Fuzzy-Zeitreihenanalyse (Fuzzy-ARMA)
- **Autoregressive Prozesse (ARMA)**

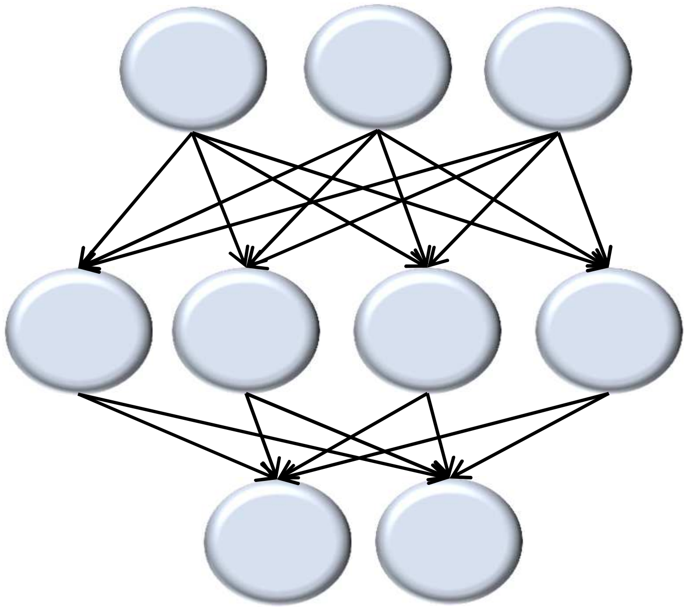
Künstliche Neuronale Netze

Künstliche neuronale Netze

- Netze künstlicher Neuronen haben Ihre Anfänge in den 1940er Jahren

- Werden zur Mustererkennung genutzt

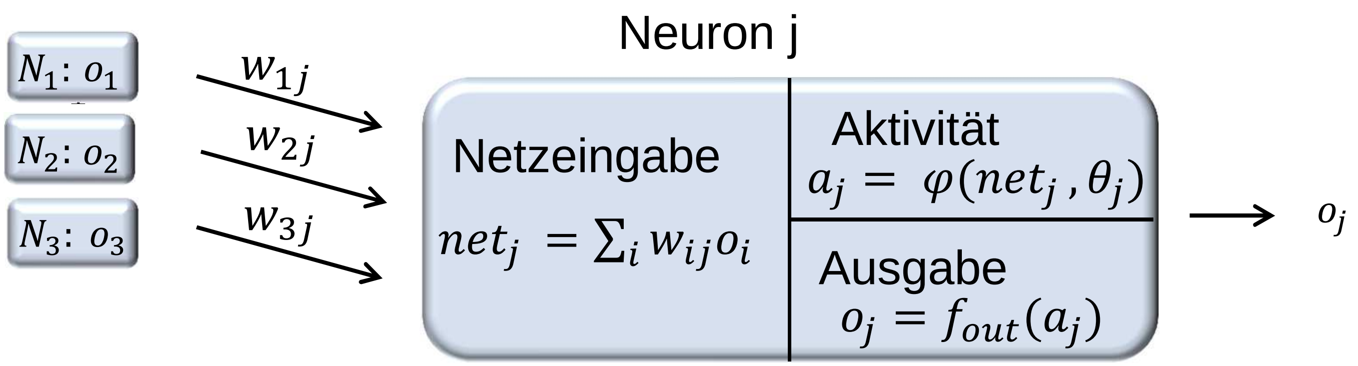




https://upload.wikimedia.org/wikipedia/commons/thumb/d/d6/Complete_neuron_cell_diagram_de.svg/2000px-Complete_neuron_cell_diagram_de.svg.png

Schema eines künstlichen Neurons

Jedes Neuron bündelt eingehende Signale und gibt ggf. ein Output Signal weiter.



Aktivierungsfunktionen künstlicher Neuronen

Beispiele für die Aktivierungsfunktionen:

Sprungfunktion

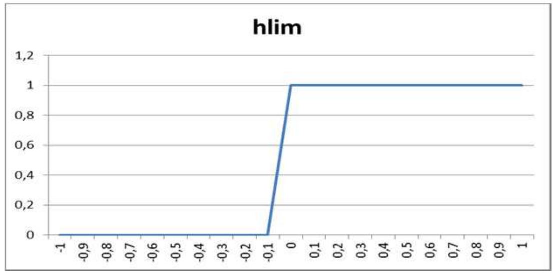
$$\varphi^{\mathrm{hlim}}(x) = \begin{cases} 0, \text{wenn } x \le 0 \\ 1, \text{wenn } x > 0 \end{cases}$$

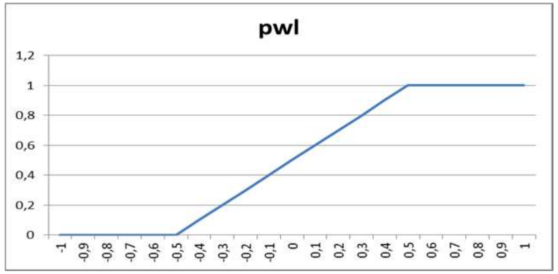
Stückweise linear

$$\varphi^{pwl}(x) = \begin{cases} 1, \text{für } x \ge 0{,}5 \\ x + 0{,}5, \text{für } -0{,}5 < x < 0{,}5 \\ 0, \text{für } x \le -0{,}5 \end{cases}$$

Sigmoidal

$$\varphi^{sig}(x) = \frac{1}{1 + e^{(-ax)}}$$





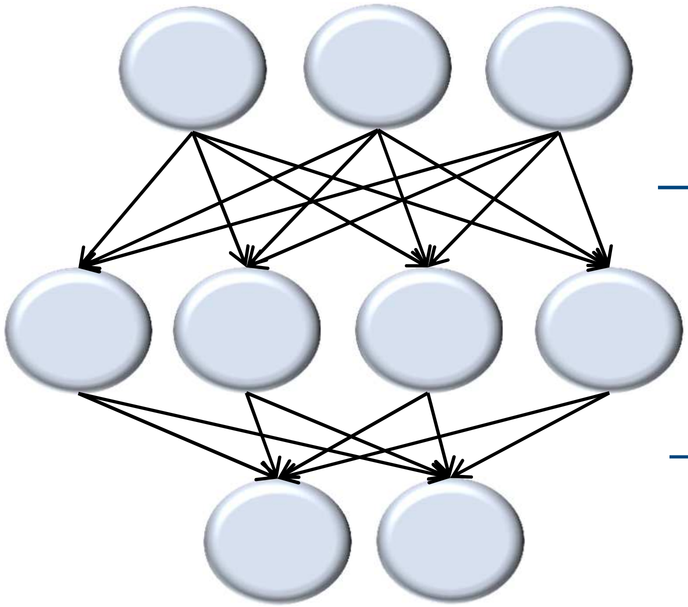
Netztopologien: vom einzelnen Neuron zum Netz

Wie werden Neurone zu einem Netz verschaltet?
Hier: mehrschichtiges Feed-Forward Netz

Eingabeschicht

Verborgene Schicht

Ausgabeschicht



Gewichte

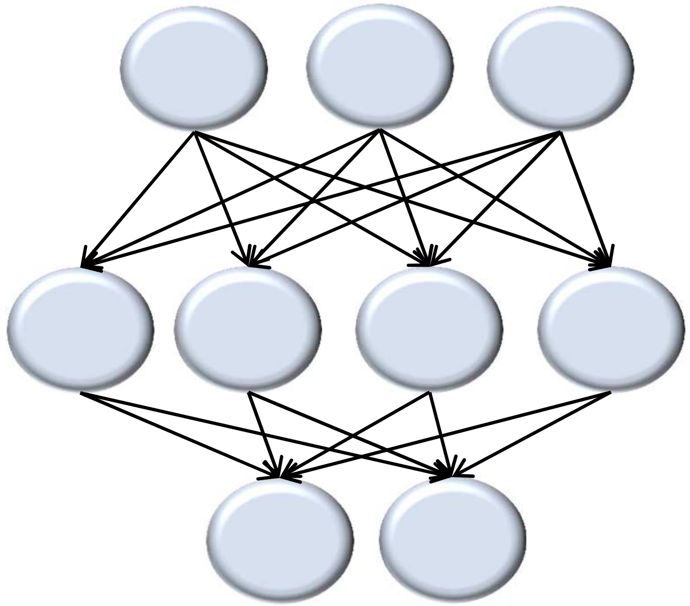
$$\begin{pmatrix} {w^1}_{11} & {w^1}_{12} & \cdots \\ {w^1}_{21} & {w^1}_{22} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}$$

$$\begin{pmatrix} {w^2}_{11} & {w^2}_{12} & \cdots \\ {w^2}_{21} & {w^2}_{22} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}$$

Wie lernt ein neuronales Netz?

Denkbare Modifikationen, um ein KNN „lernen“ zu lassen:

- Neue Verbindungen schaffen
- Alte Verbindungen löschen
- Verbindungsgewichte ändern
- Neuronenfunktionen ändern
- Schwellenwerte ändern
- Neue Neuronen erzeugen
- Alte Neuronen löschen



DAV DEUTSCHE AKTUARVEREINIGUNG e.V.

Lernen über Backpropagation

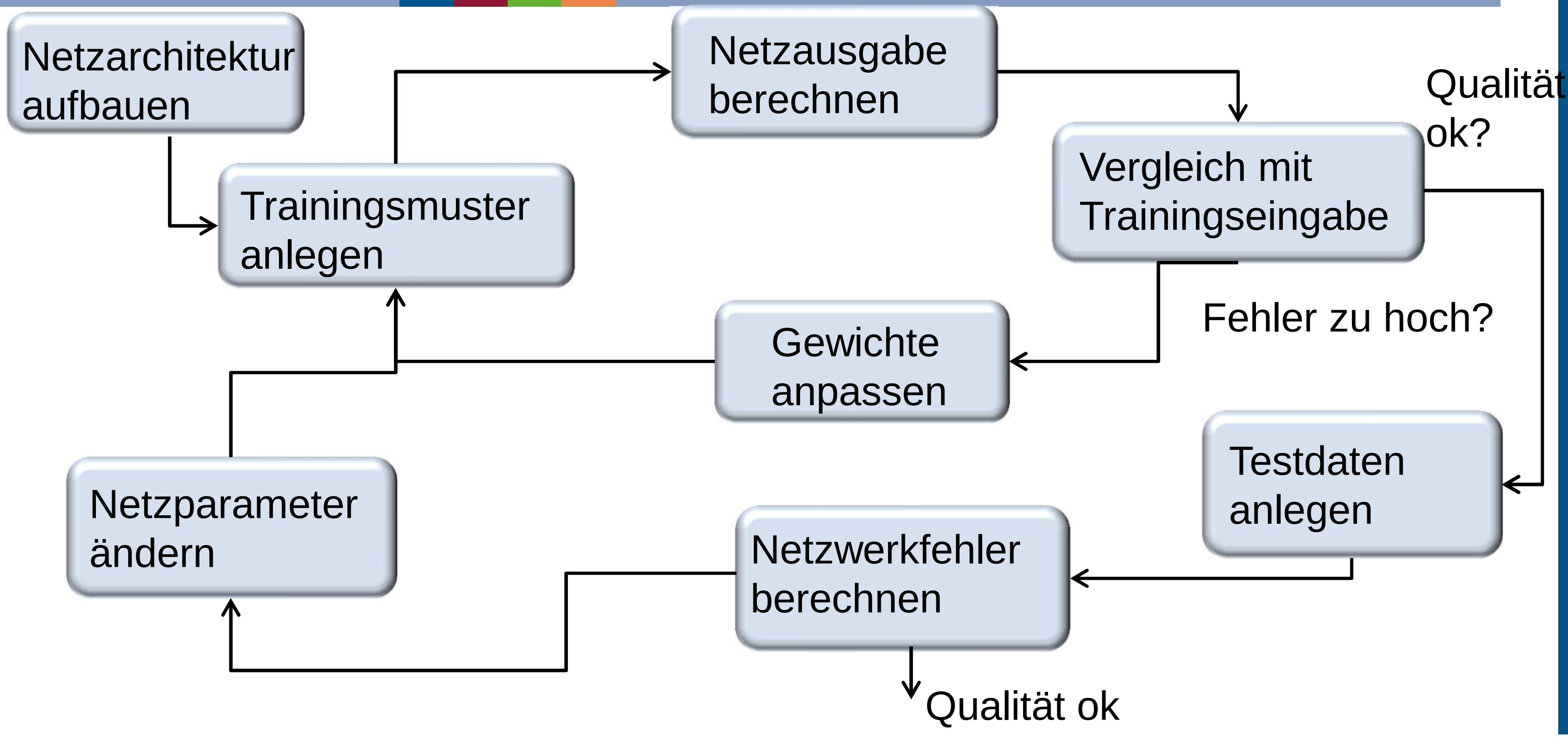
→ Idee: Beschreibe Qualität der I/O-Abbildung durch ein Fehlerfunktional und minimiere dies über ein Gradientenverfahren:

$$E = \frac{1}{2} \sum_i (t_i - o_i)^2 \xrightarrow{\text{Minimierung nach Gewichten}} \frac{\partial E}{\partial w_{i,j}} = \frac{\partial E}{\partial o_j} \frac{\partial o_j}{\partial net_j} \frac{\partial net_j}{\partial w_{i,j}}$$

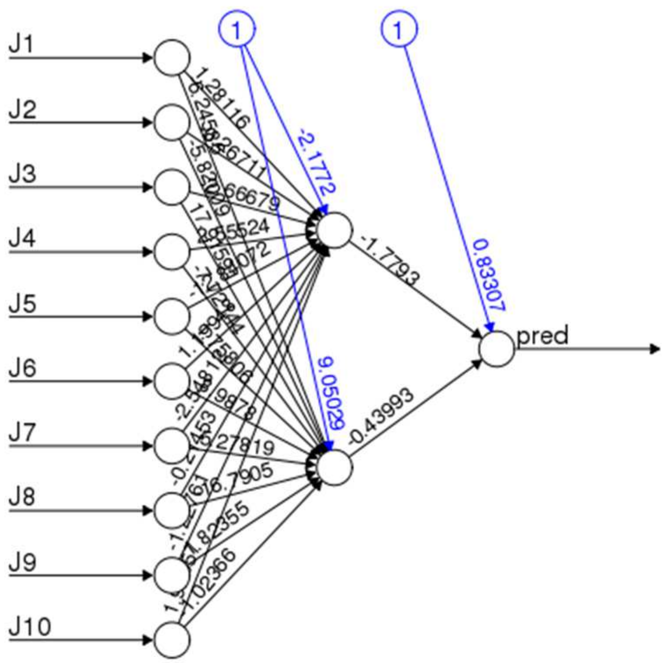
$$\Delta w_{i,j} = -\beta \frac{\partial E}{\partial w_{i,j}} = \beta \delta_j x_i \qquad \text{mit} \qquad \delta_j = \begin{cases} \varphi'(net_j)(t_j - o_j) \\ \varphi'(net_j) \sum_l \delta_l w_{i,j} \end{cases}$$

Update Formel: $\boxed{w_{i,j}^{(neu)} = w_{i,j}^{(alt)} + \Delta w_{i,j}}$

Trainings- und Testphase: wie lernt eine Maschine?



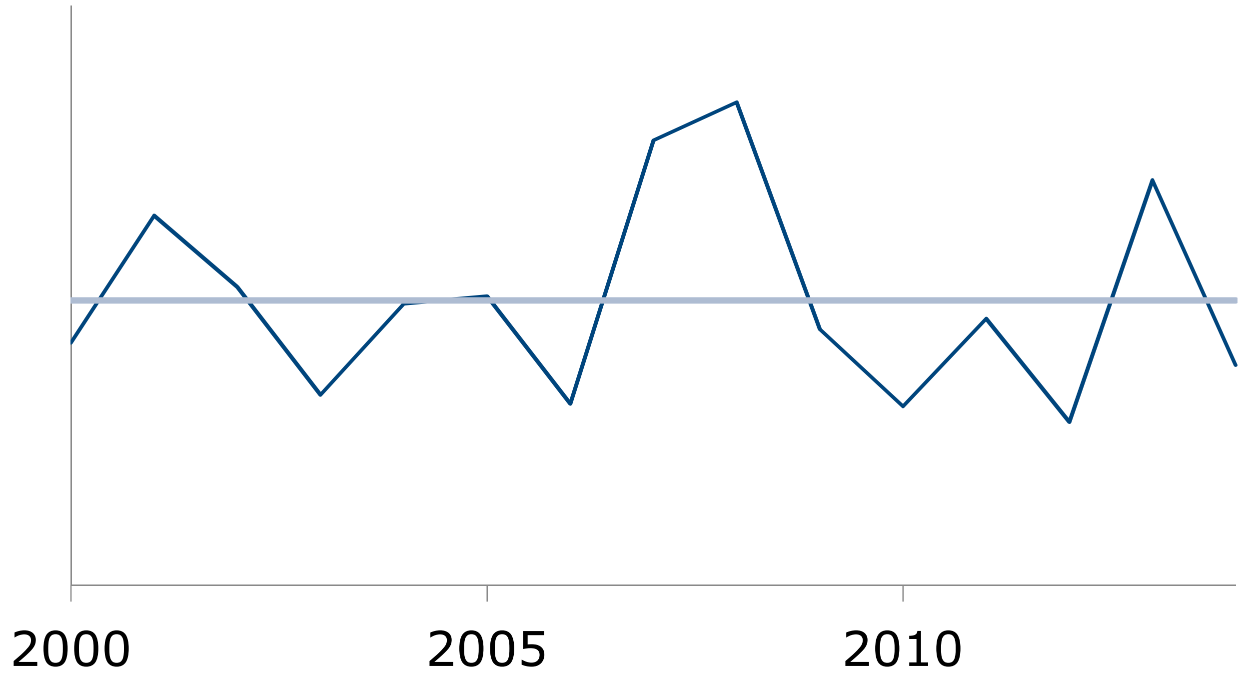
Implementierung mit dem „neuralnet“- Package in R




```
# Trainigs- Testdaten zufaellig waehlen
index <- sample(1:nrow(data),round(0.75*nrow(data)))
train <- data[index,]
test <- data[-index,]

# Daten normalisieren
maxs <- apply(data, 2, max)
mins <- apply(data, 2, min)
scaled <- as.data.frame(scale(data, center = mins, scale = maxs - mins))
train_ <- scaled[index,]test_ <- scaled[-index,]

#NN trainieren
library(neuralnet)n <- names(train_)
f <- as.formula(paste("pred ~", paste(n[!n %in% "pred"], collapse = " + ")))
nn <- neuralnet(f,data=train_,hidden=2,linear.output=T)

#NN visualisieren
plot(nn)

# Model fit
pr.nn <- compute(nn,test_[,1:10])
pr.nn_ <- pr.nn$net.result*(max(data$pred)-min(data$pred))+min(data$pred)
test.r <- (test_$pred)*(max(data$pred)-min(data$pred))+min(data$pred)
MSE.nn <- sum((test.r - pr.nn_)^2)/nrow(test_)

# Fehler vergleichen
print(paste(MSE.lm,MSE.nn))
```

ARMA-Modelle

Mögliche Struktur eines jährlichen Leistungstrends



- Beobachtungswerte schwanken um langfristigen Mittelwert
- Schätzungen mit Mittelwert und linearen Trends können zu z. T. erheblichen Schätzfehlern führen
- starke Ausreißer lassen sich nur über Verteilungsaussagen prognostizieren

DAV DEUTSCHE AKTUARVEREINIGUNG e.V.

Leistungstrend als stochastischer Prozess

- Leistungstrend wird als **stochastischer Prozess** angesehen $\{I_t\}_{t \in \mathbb{Z}}$
- Annahme sei (zunächst), dass I_t **stationär** ist, d. h. es gelten:

  (i) $E[I_t] = \mu \; konstant$

  (ii) $Var[I_t] < \infty \; für \; alle \; t \in \mathbb{Z}$

  (iii) $Cov[I_s, I_t] = Cov[I_{s+h}, I_{t+h}] \;\; für \; alle \; s, t, h \in \mathbb{Z}$ (**Autokovarianz**)

- Leistungstrend streut um einen langfristigen konstanten Mittelwert μ
- durch die Autokovarianz wird die „*innere Abhängigkeitsstruktur*“ des Prozesses beschrieben („Wie hängt der Wert im Jahr t vom Wert in t-1, t-2, ... ab?“)
- die Autokovarianz des Prozesses hängt nur vom <u>Abstand</u> der Beobachtungszeiträume, aber nicht vom konkreten Beobachtungszeitpunkt ab

Autoregressiver Prozess AR(1)

- **Idee:** Leistungstrend als **Mean-Reversion-Prozess** ansehen
- Modellierung der jährlichen Veränderung des Leistungstrends I_t - I_{t-1}

$$\boldsymbol{I_t - I_{t-1} = \underbrace{a \cdot (\mu - I_{t-1})}_{\textbf{Drift-Komponente}} + \underbrace{Z_t}_{\textbf{Diffusion}}}$$

Drift-Komponente:

μ = langfristiger (beobachteter) Mittelwert des Leistungstrends

a = Faktor für die „Rückkehrgeschwindigkeit“ zum langfristigen Mittelwert

Diffusion:

allgemein: $Z_t \sim WN(0,\sigma^2)$ (White-Noise-Prozess)

→ stochastische Komponente des Leistungstrends (repräsentiert die Volatilität der Werte)

DAV
DEUTSCHE AKTUARVEREINIGUNG e.V.

Prognose des Leistungstrends

Schätzer für den Leistungstrend im Folgejahr:

Wegen $E[Z_t]=0$ ergibt sich folgende **Best-Estimate-Schätzung**:

$$E[I_t|I_{t-1} = i_{t-1}] = i_{t-1} + a \cdot (\mu - i_{t-1})$$

D. h. nach Festlegung der Modellparameter kann der erwartete Leistungstrend im Folgejahr geschätzt werden.

Einbindung der stochastischen Komponente ermöglicht auch die Angabe der gesamten Verteilungsfunktion des Leistungstrends im Folgejahr / in den Folgejahren.

DAV DEUTSCHE AKTUARVEREINIGUNG e.V.

Schätzung der Modellparameter

$$\left.\begin{aligned} \mu &= \hat{\mu} \\ a &= 1 - \hat{\rho}(1) \\ \sigma^2 &= \hat{\gamma}(0) \cdot (1 - (1 - a)^2) \end{aligned}\right\} \quad \textbf{Yule-Walker-Schätzer}$$

– gegeben seien T Beobachtungswerte $I_1, ..., I_T$ des Leistungstrends

– Schätzer für Mittelwert: $\hat{\mu} = \frac{1}{T} \cdot \sum_{t=1}^{T} I_t$

– Schätzer für die Autokovarianz: $\hat{\gamma}(h) = \frac{1}{T} \cdot \sum_{t=1}^{T-h} (I_t - \hat{\mu}) \cdot (I_{t+h} - \hat{\mu})$

– Schätzer für die Autokorrelation: $\hat{\rho}(h) = \frac{\hat{\gamma}(h)}{\hat{\gamma}(0)}$

Vergleich der Prognosegüte

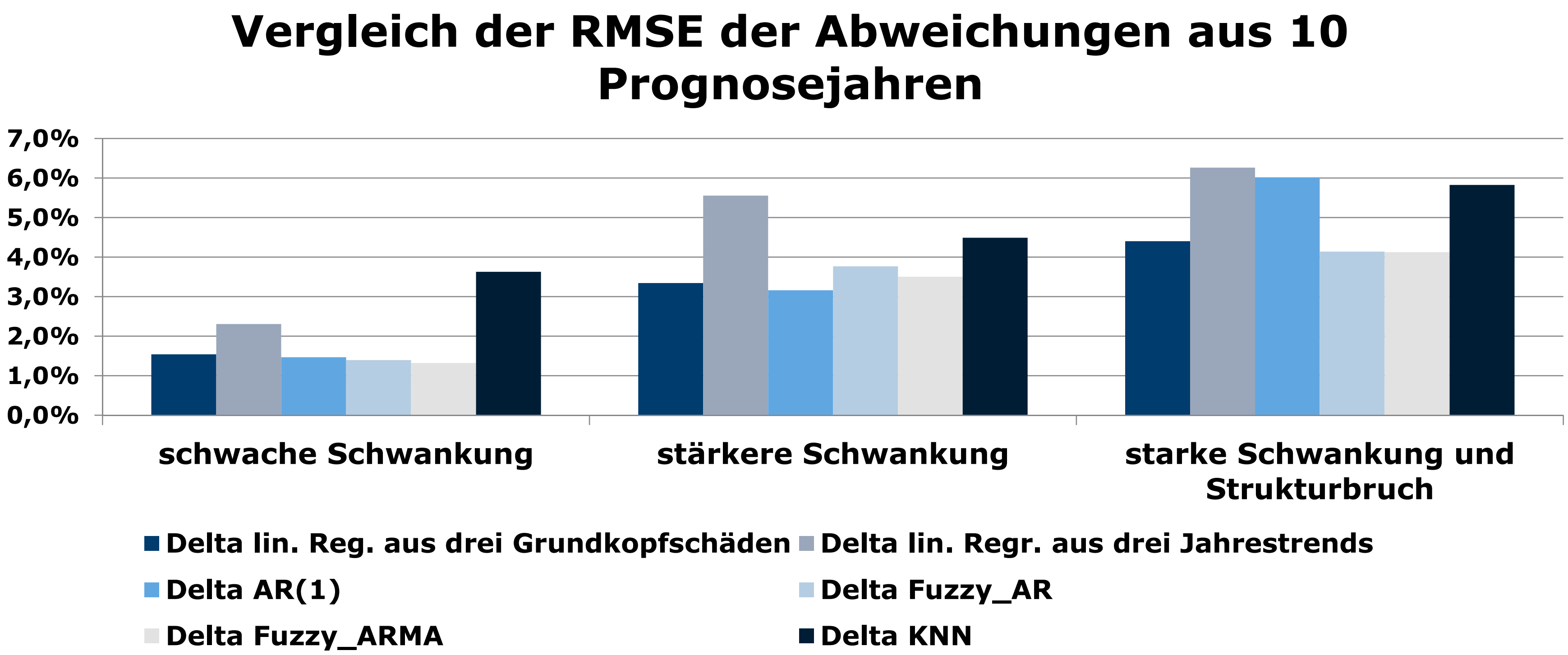
Vergleich der Prognosegüte

Methodik zur Feststellung der Prognosegüte:
Schätzung der Jahrestrends der letzten zehn Jahre unter Verwendung der im jeweiligen Prognosejahr verfügbaren Daten

Kennzahlen:
- jährliche Abweichung der Prognosewerte vom tatsächlichen Leistungstrend
- Mittelwert und RMSE (= Root-Mean-Square-Error) der Schätzfehler in den zehn Prognosejahren

Prognosegüte hängt stark von der Struktur der Zeitreihe ab
- Leistungstrends mit schwacher Schwankung
- Leistungstrends mit höherer Schwankung
- Leistungstrends mit starker Schwankung und Strukturbruch

Vergleich der Prognosequalität der Methoden



Stochastische Modellierung und ihre Anwendungen in der Praxis

Wozu stochastische Modellierung?

- untersuchte statistische Verfahren lieferten mehr oder weniger gute Ergebnisse, jedoch keine „exakte“ Prognose

- Problem liegt nicht an den Verfahren, sondern an den mitunter starken Schwankungen der jährlichen Leistungstrends

→ Punktschätzer ergänzen um eine Aussage zu einer möglichen **Bandbreite** der zukünftigen Leistungstrends

→ **Idee:** stochastische Modellierung der zukünftigen Leistungstrends und Ableitung der Verteilungsfunktion für unterschiedliche Unternehmenskennzahlen

Ansatz für eine stochastische Modellierung

- **Idee:** Leistungstrend als **Mean-Reversion-Prozess** ansehen
- Modellierung der jährlichen Veränderung des Leistungstrends $I_t - I_{t-1}$

$$I_t - I_{t-1} = \underbrace{a \cdot (\mu - I_{t-1})}_{\text{Drift-Komponente}} + \underbrace{Z_t}_{\text{Diffusion}}$$

Drift-Komponente:

μ = langfristiger (beobachteter) Mittelwert des Leistungstrends

a = Faktor für die „Rückkehrgeschwindigkeit" zum langfristigen Mittelwert

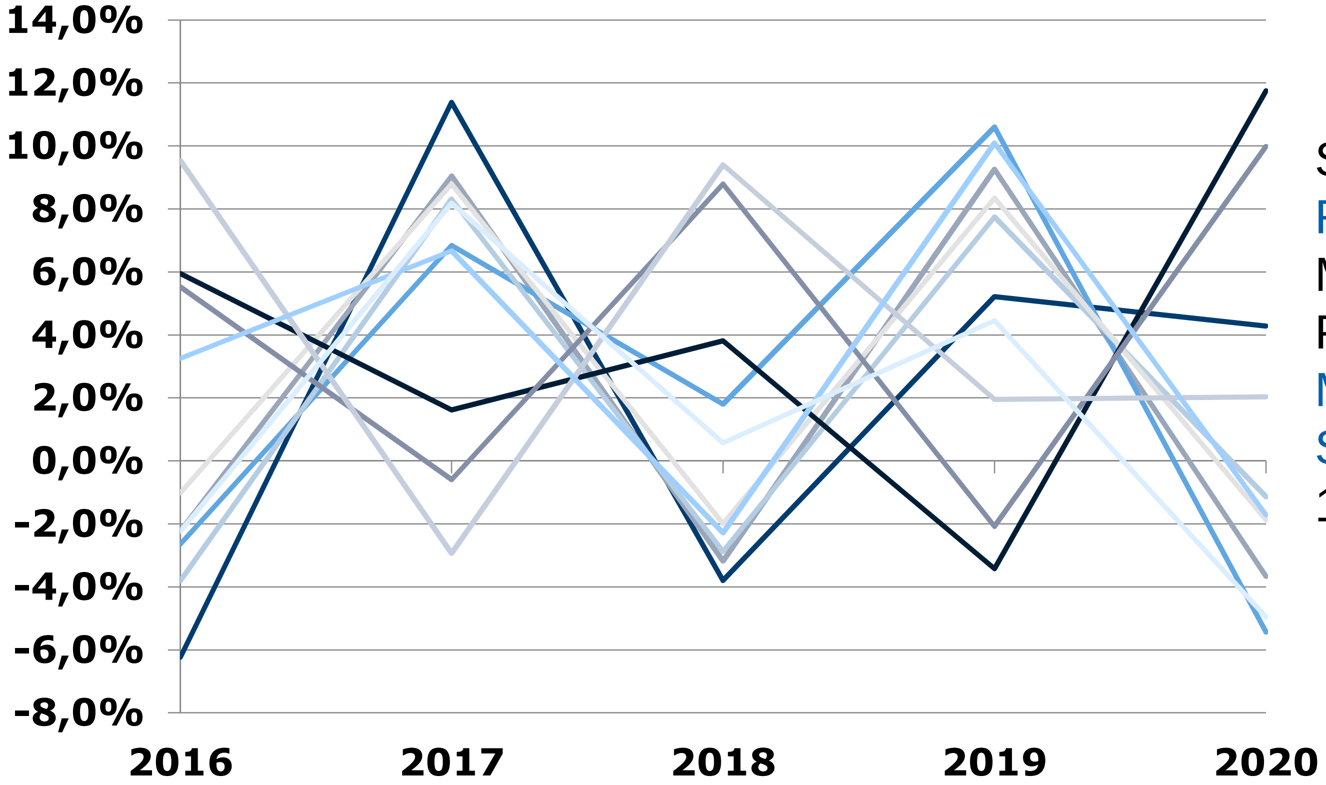
Diffusion:

allgemein: $Z_t \sim WN(0,\sigma^2)$ (White-Noise-Prozess)

→ stochastische Komponente des Leistungstrends (repräsentiert die Volatilität der Werte)

→ **hier: $Z_t \sim N(0,\sigma^2)$**

Simulation möglicher zukünftiger Leistungstrends



Simulation möglicher Realisierungen des Mean-Reversion-Prozesses mittels Monte-Carlo-Simulation (z. B. 1.000 Pfade)

Mögliche Anwendungen in der Praxis - Kalkulation

- Im Kalkulationsjahr t sind die tatsächlichen Grundkopfschäden G_i der Beobachtungsjahre i = t-3, t-2 und t-1 sowie der aktuelle rechnungsmäßige Grundkopfschaden G^{rm} bekannt

- Simulation von 1.000 Realisierungen des unbekannten Grundkopfschadens im Jahr t+1 gemäß

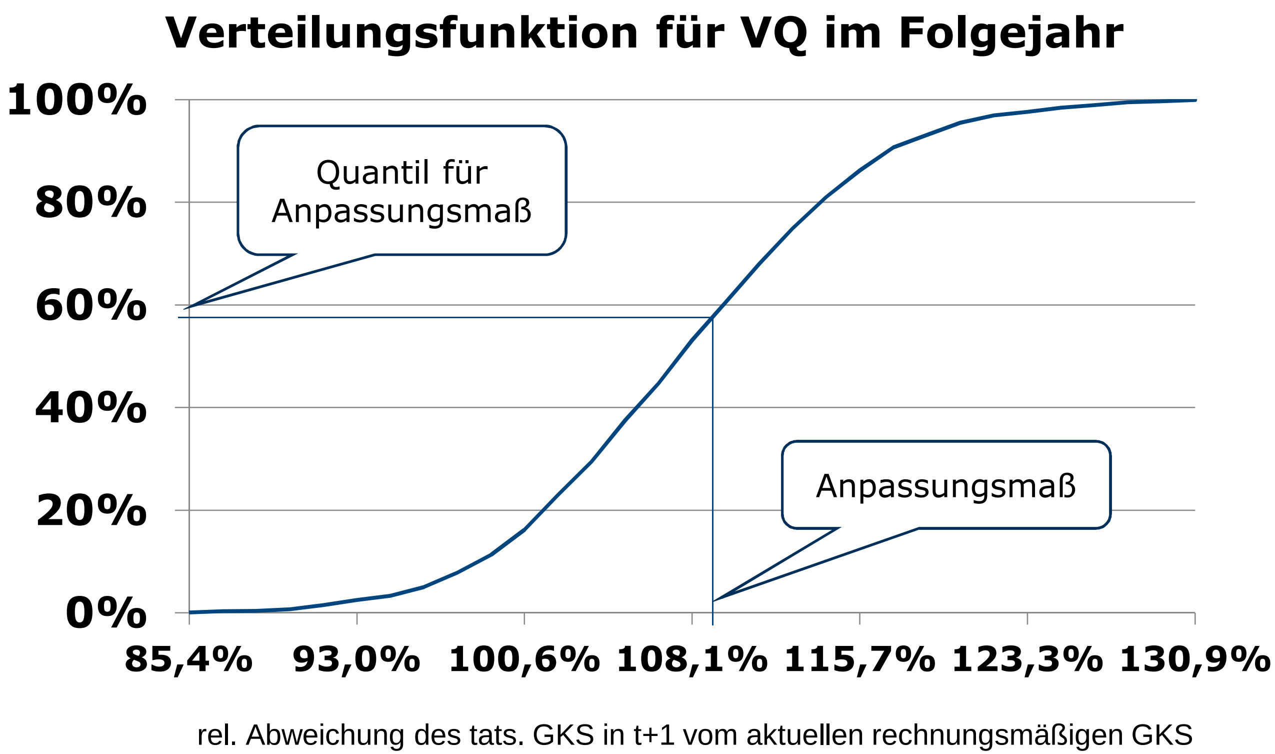
$$\hat{G}_{t+1,Pfad} = G_{t-1} \cdot \left(1 + i_{t,Pfad}\right) \cdot \left(1 + i_{t+1,Pfad}\right)$$

- Erstellung der empirischen Verteilungsfunktion für die relative Abweichung

$$\widehat{VQ}_{t+1,Pfad} := \frac{\hat{G}_{t+1,Pfad}}{G^{rm}}$$

→ Quantil als Maß für die Sicherheit des aktuellen Anpassungsmaßes

Mögliche Anwendungen in der Praxis - Kalkulation



Perspektive 1:

Festlegung des Anpassungsmaßes nach den üblichen Verfahren und Betrachtung der gewählten Sicherheit

Perspektive 2:

Auswahl des Anpassungsmaßes als Quantil abhängig von gewünschtem Sicherheitsniveau

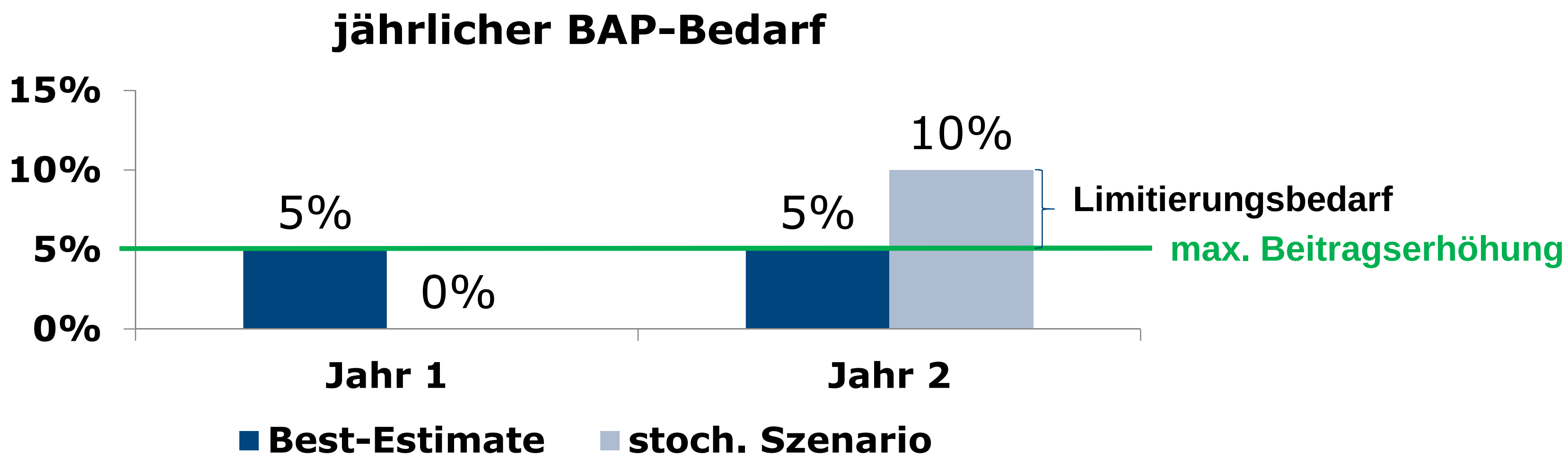
Mögliche Anwendungen in der Praxis - Unternehmensplanung

- Unternehmensplanung beschreibt i. d. R. die erwartete Entwicklung der wesentlichen Unternehmenskennzahlen unter Verwendung von Best-Estimate-Schätzern für die Berechnungsparameter

- die zukünftige Leistungsentwicklung hat wesentlichen Einfluss auf
  - Anpassungsausmaß (Zeitpunkt und Höhe)
  - Überschuss-Situation (insb. Möglichkeit einer Rechnungszins-Senkung)
  - Entwicklung der RfB (Entnahmebedarf und Zuführung)

→ **Reicht die Betrachtung eines Best-Estimate-Szenarios für die zukünftige Leistungsentwicklung aus?**

Mögliche Anwendungen in der Praxis - Unternehmensplanung



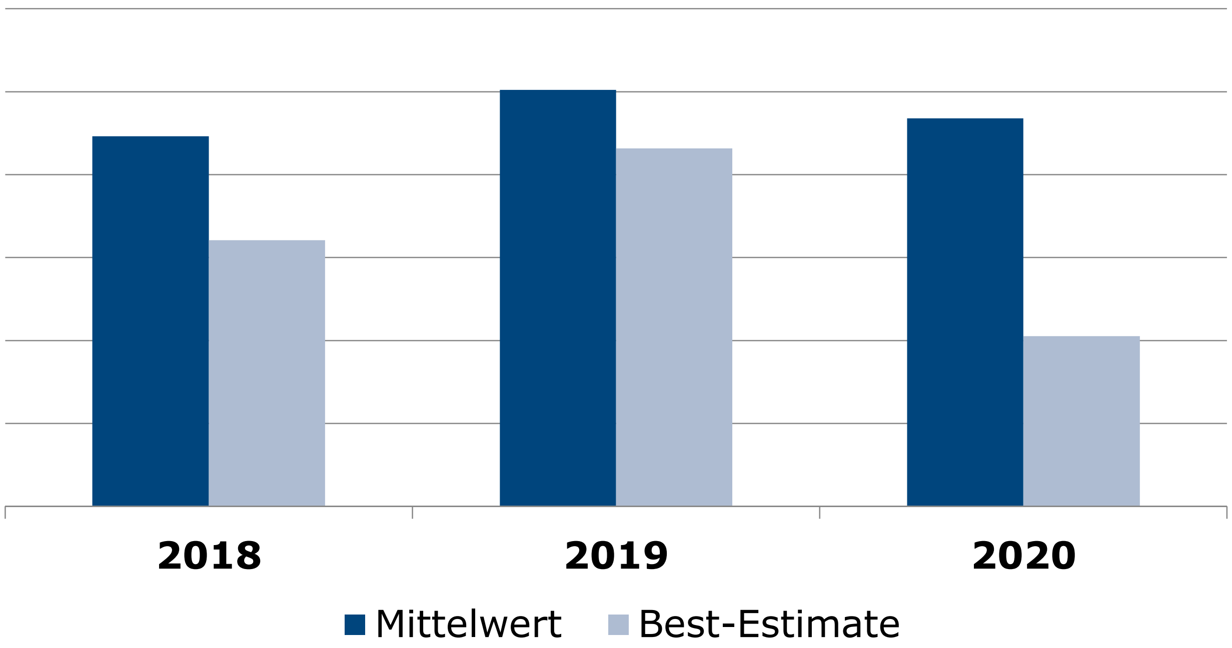
- in beiden Szenarien ergibt sich derselbe durchschnittliche BAP-Bedarf
- aber: ein Limitierungsbedarf wird erst durch die Berücksichtigung einer Schwankung des BAP-Bedarfs deutlich

→ ein reiner Best-Estimate-Ansatz kann zu einer Fehleinschätzung des RfB-Mittelbedarfs führen

DAV DEUTSCHE AKTUARVEREINIGUNG e.V.

Mögliche Anwendungen in der Praxis - Unternehmensplanung

Vergleich der RfB-Entnahme aus Mittelwert der Szenarien und Best-Estimate-Szenario

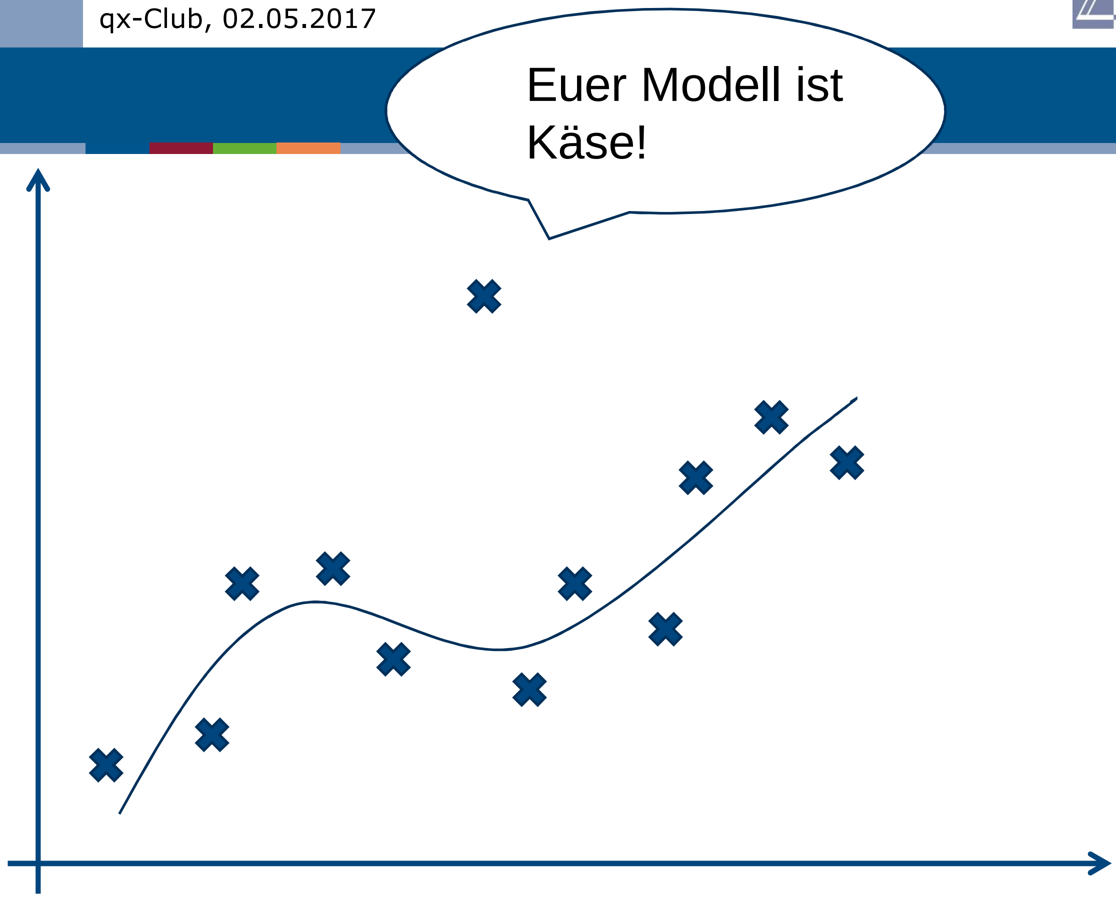


- mehrjährige Simulation der BAP und der zugehörigen RfB-Entnahme auf Basis stochastischer Szenarien für die Leistungsentwicklung und des Best-Estimate-Szenarios
- Mittelwert der stoch. Szenarien kann sich deutlich vom Best-Estimate unterscheiden

Fazit

Fazit

- Es wurden verschiedene Verfahren zur Prognose von Leistungstrends in Bezug auf Prognosequalität und Handhabbarkeit verprobt
- Stochastische Schwankungen in den Kostentrends überlagern strukturelle Effekte und machen Prognosen unabhängig vom Verfahren schwierig
- Unsere Erfahrungen bzgl.:
  - Lineare Regressionsmodelle
  - Künstliche neuronale Netze
  - ARMA- Modelle
  - Fuzzy-Forecast Ansätze
- Ergänzung der Punktschätzern durch stochastische Modellierung und Aussagen über die Verteilungsfunktion der Leistungstrends liefert zusätzliche Erkenntnisse



Vielen Dank für Ihre Aufmerksamkeit.