Ansätze zur datengetriebenen Formulierung von Strukturhypothesen für dynamische Systeme

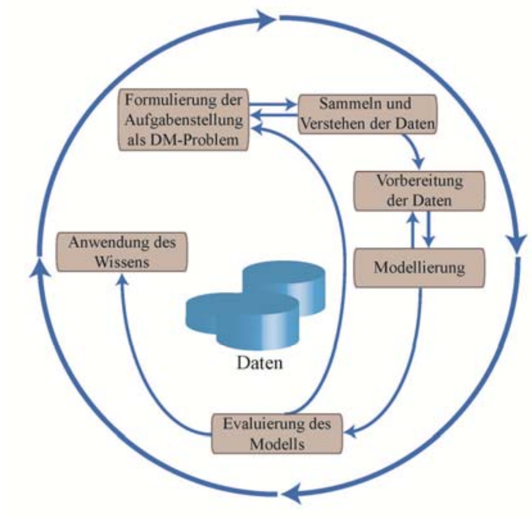
Christian Kühnert, Lutz Gröll, Michael Heizmann, Ralf Mikut
30.11.2011

Karlsruher Institut für Technologie

Fraunhofer IOSB

Motivation

CRISP
Cross-Industry Standard Process for Data Mining



- **„Klassisches“ Data-Mining**
  - Generierung von verborgenem Wissen aus vorhandenen Daten
  - CRISP: Wird häufig im Produktions- und Fertigungsbereich eingesetzt
  - Beispiele: Klassifikation in Gut/Schlecht-Produktionen, Online-Monitoring,…

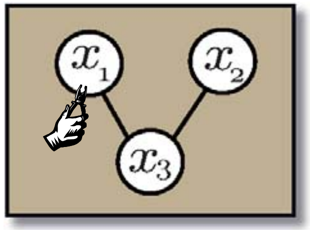
- **Eigenschaften**
  - Es wird von einer „natürlichen“ Prozessumgebung ausgegangen (Beobachtungsdaten)
  - Lernverfahren ist oftmals eine Black-Box
  - Konsequenz eines Eingriffs in den Prozess ist nicht vorhersagbar
  - „Was wäre, wenn“ kann nicht beantwortet werden

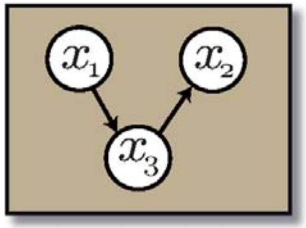
Oft ist es notwendig die kausale Struktur des Prozesses zu kennen

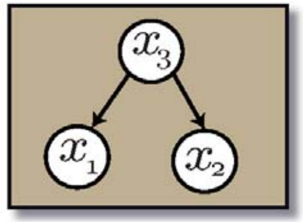
Einleitung

Probabilistisches Kausalprinzip (statische Systeme)

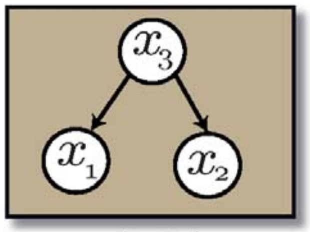
- Lernen kausaler Zusammenhänge mittels Wahrscheinlichkeitstheorie
  - Die Ursache erhöht die Wahrscheinlichkeit, dass ein Effekt eintritt
  - Die Ursache wird durch eine Intervention von außen herbeigeführt



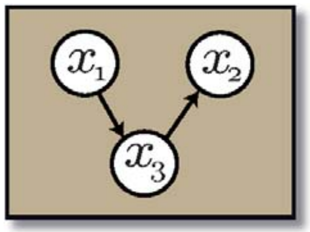




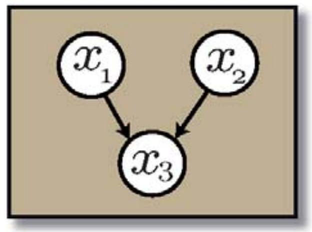
- Faithfulness Assumption
  - Erkennen kausaler Abhängigkeiten durch Beobachtungen
  - Gefundene Struktur enthält alle (un-) bedingten Unabhängigkeiten



$x_1 \perp x_2 | x_3$



$x_1 \perp x_2 | x_3$



$x_1 \not\perp x_2 | x_3$
$x_1 \perp x_2$

Einleitung

Kausalität und Zeit

- In der Regel werden dynamische Prozesse betrachtet
- Definition kausales System (Systemtheorie)
  - Werte der Eingangssignale $t > t_1$ haben keinen Einfluss auf Ausgangssignale bis $t < t_1$
  - Die Impulsantwort $t < 0$ ist 0.
  - Bei statischen Systemen liegt Kausalität vor, da das Ausgangssignal vom aktuellen Eingangssignal abhängt

→ Nutzen von Zeitverschiebungen zum Erkennen kausaler Abhängigkeiten in Daten

Beispiele:

- Tiefpass
- Feder-Dämpfer System
- Wachstumskurven in der Biologie
- Chemische Reaktionen

Einleitung

$$u(t) \longrightarrow \boxed{\text{PT1-Glied}} \longrightarrow y(t)$$

$$\dot{y}(t) = \frac{K}{T}\,u(t) - \frac{1}{T}\,y(t)$$

Der Datenpunkt $\dot{y}(t)$ hängt von $y(t)$ und $u(t)$ ab

Ausgewählte Methoden:

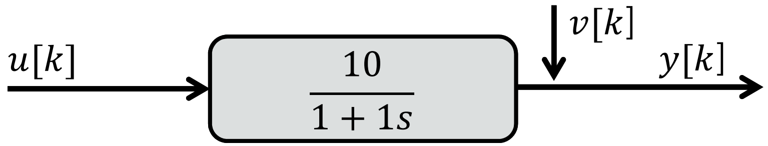
- Kreuzkovarianzfunktion
- Granger-Kausalität
- N4SID

Benchmark: Tiefpass

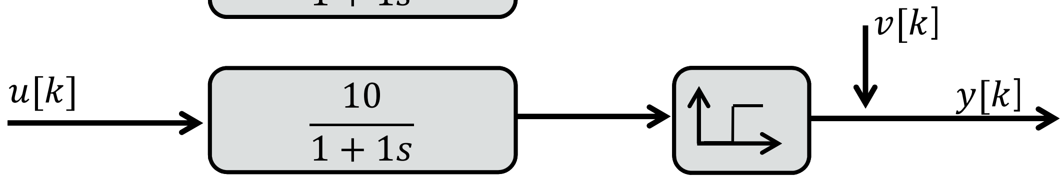
- $u(t)$ Eingangssignal - Weißes Rauschen
- $y(t)$ Ausgangssignal des Systems
- $v(t)$ Überlagertes Rauschen am Ausgang mit 10% der Stärke des Eingangssignals

Untersuchte Systeme und Eingangssignale

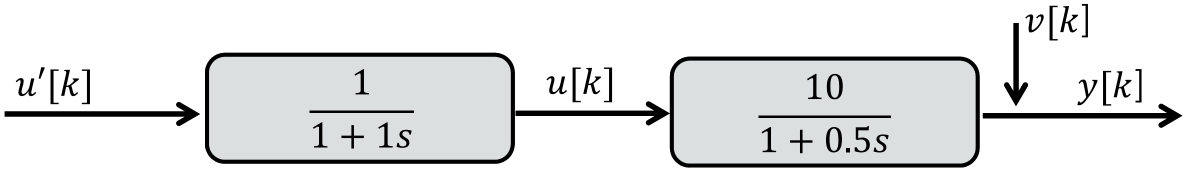
- Basissystem



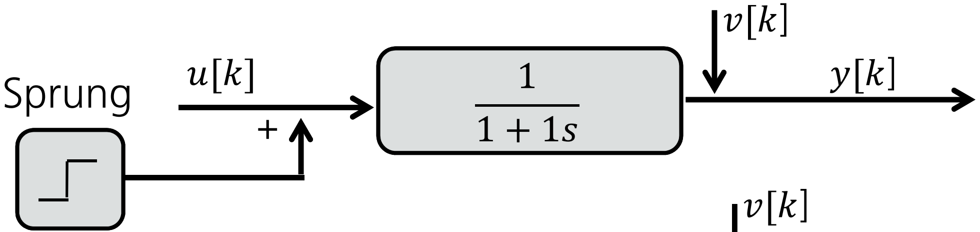
- Totzeit



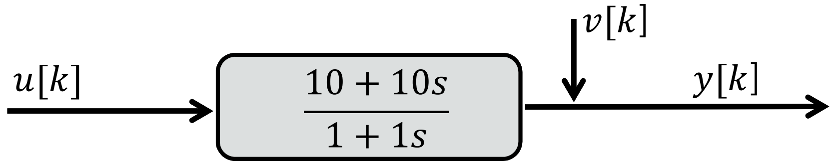
- Farbiges Rauschen



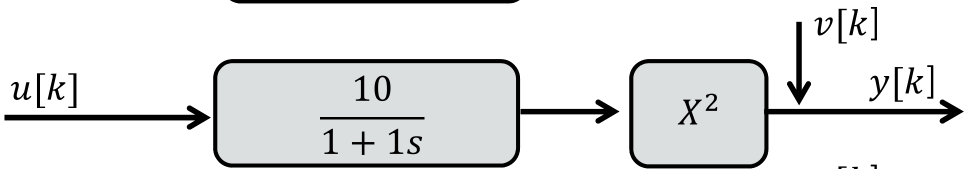
- Überlagerter Sprung



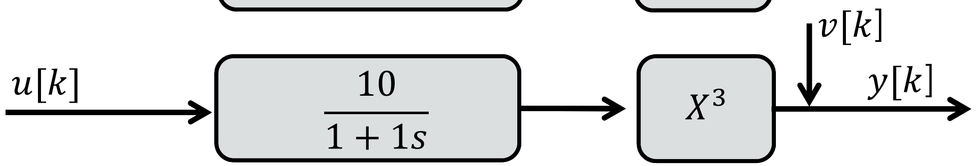
- Sprungfähig



- Quadrierer



- Kubierer



Kreuzkovarianzfunktion

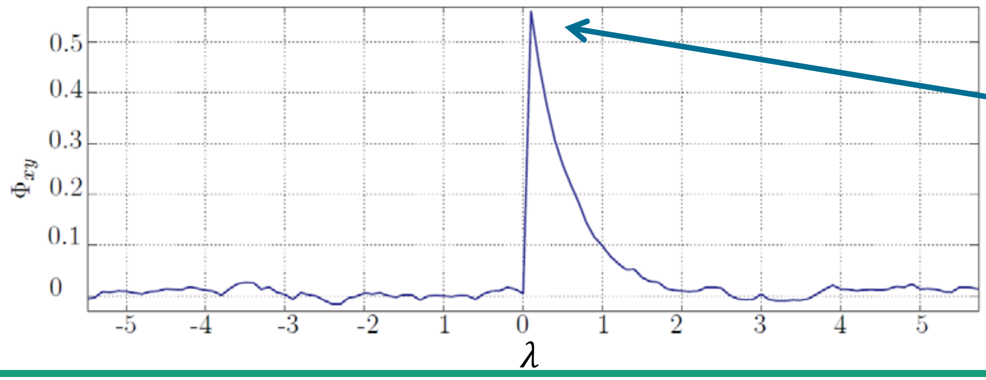
- Klassisches Maß zur Beschreibung der Abhängigkeit zweier Signale bei unterschiedlicher Zeitverschiebung
- Rein deterministisch

Voraussetzungen

- u ist eine Realisierung eines stationären weißen Rauschprozesses
- Das u und y verbindende System ist linear und stabil

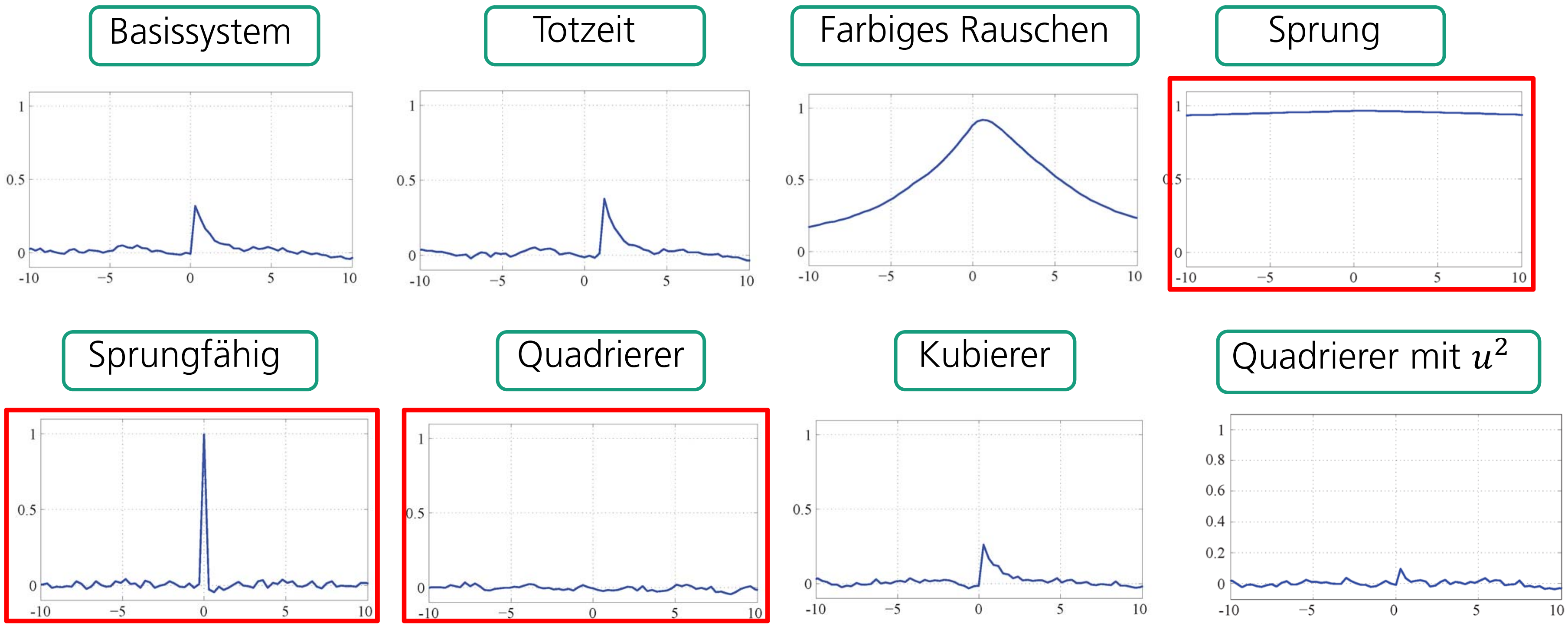
Erkennen kausales Verhalten

- Betrachtung der Zeitverschiebung λ
- Höhe/Position des Maximum
- Vergleich der Flächen für $\lambda < 0$ und $\lambda > 0$



- Peak bei $\lambda = 1$
- Abfallende e-Funktion (proportional Gewichtsfunktion)

Kreuzkorrelation – PT1-System



- Eigenschaften:
  - Für lineare, nicht-sprungfähige Systeme
  - Für stationäres Eingangssignal
  - Für große Totzeiten

N4SID

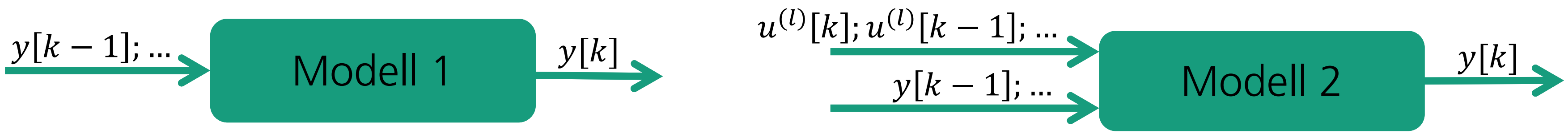
- Algorithmus gehört zur Klasse der Identifikationsverfahren im Unterraum
  - Ein-/Ausgangsdaten werden aufeinander projiziert
  - Über die Berechnung von Singulärwerten wird die Dimension des Raums geschätzt (wichtig für Totzeiten)

Störquelle (MISO-System)

- $z[k]$ Weißes Rauschen
- $z_s[k]$ Weißes Rauschen mit überlagertem Eingangssprung

Erkennen kausales Verhalten

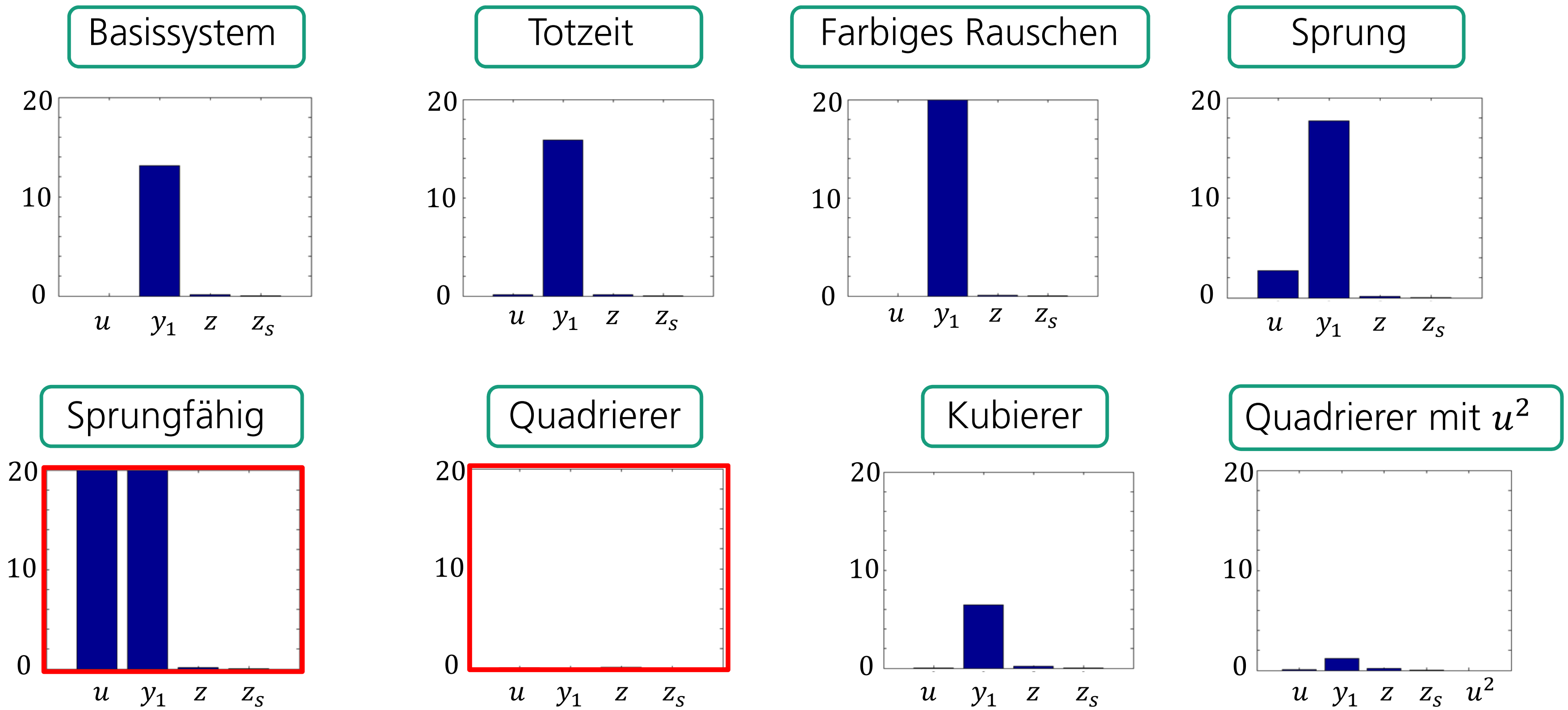
- Hypothetische Zuordnung der Ein- Ausgangsgrößen
- Berechnung des Prognosefehlers für beide Modelle
- Ergebnis liefert Aussage darüber ob $y[k]$ Ausgangssignal ist



$$fit_y = VAF(y, \hat{y}) \qquad fit_{uy} = VAF(y, \hat{y})$$

Information der Eingangsdaten über $y[k]$: $\boxed{100 \cdot \frac{fit_{uy} - fit_y}{100 - fit_y}}$

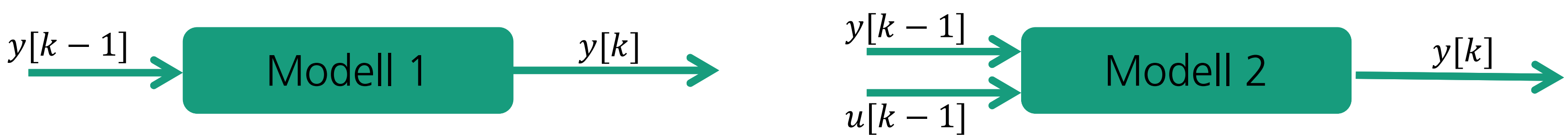
N4Sid



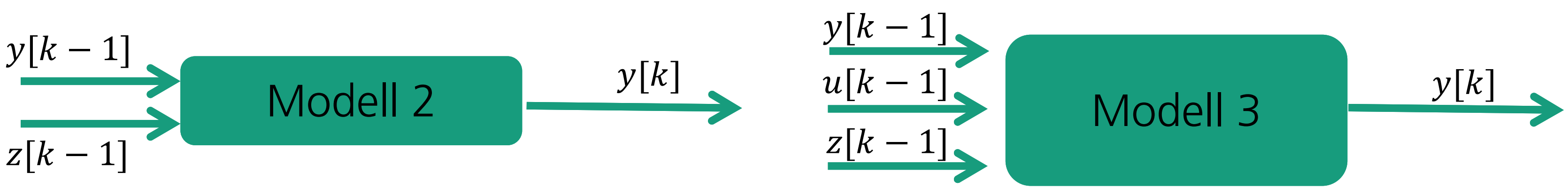
- Für lineare Systeme
- Anwendbar für Mehrgrößensysteme
- Implizite Schätzung der Totzeit → wird möglicherweise nicht erkannt

Granger-Kausalität

- Ursprüngliche Idee stammt aus der Ökonometrie (Clive Granger)
- Ein Signal u hat kausalen Einfluss auf y, wenn u hilft Werte von y besser vorherzusagen, als nur vergangene Werte von y



- Ein Signal u hat direkten kausalen Einfluss auf y, wenn es kein Signal z gibt, dass hilft, Werte von y besser vorherzusagen → Vergleich gegen alle weiteren Zeitreihen



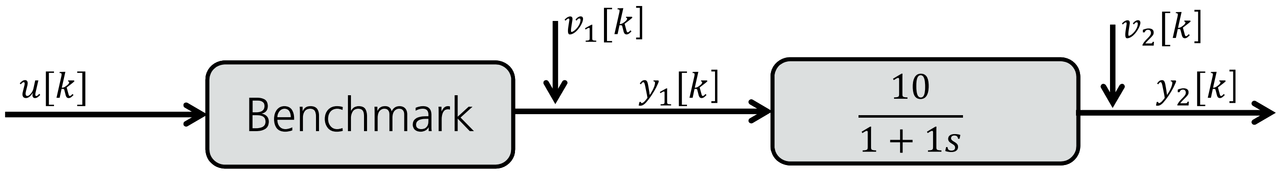
- Vergleich der Residuen
  - 2 Signifikanztests

$\boxed{Res_{uy} < Res_y} \rightarrow \boxed{Res_{uy} < Res_{uyz}} \rightarrow \boxed{\text{Kante } u \rightarrow y}$

- Bewertung des kausalen Einflusses: $Res_y - Res_{uy}$

Granger-Kausalität

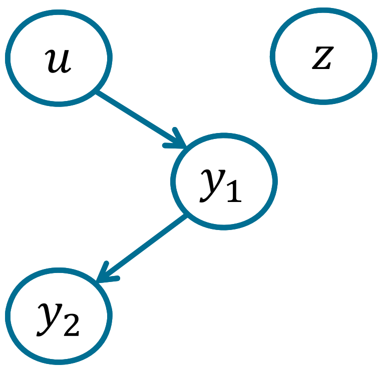
- Erweiterung des Benchmarks durch Hinzufügen eines zweiten Tiefpasses



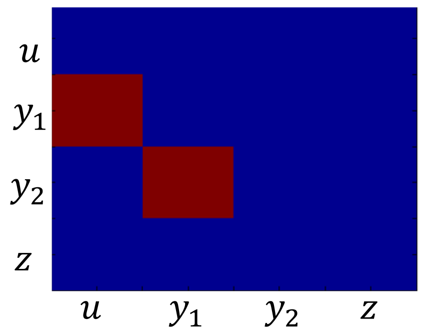
- Störquelle
  - $z[k]$ ▪ Weißes Rauschen
  - $z_s[k]$ ▪ Weißes Rauschen mit überlagertem Eingangssprung
  - ⇨ In Abbildung zusammengefasst zu einer Störgröße z

- Gesuchte Struktur

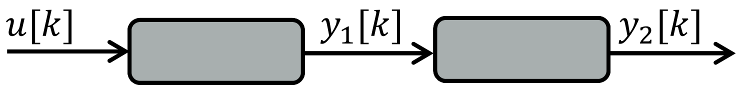
Graph



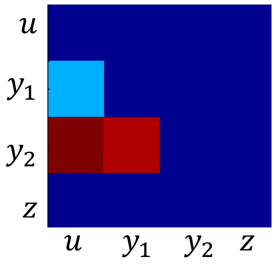
Heat-Map



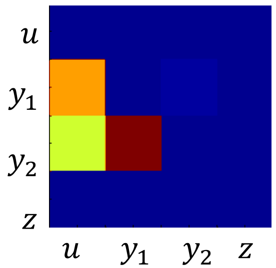
Granger Kausalität (MIMO)



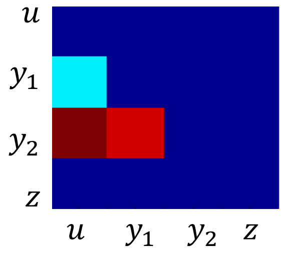
Basissystem



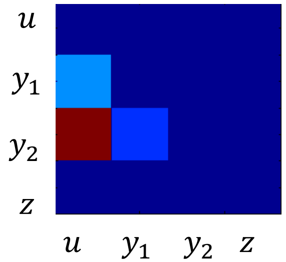
Totzeit



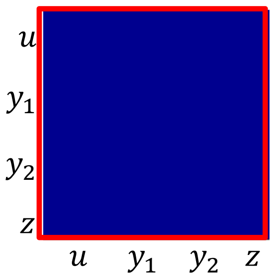
Farbiges Rauschen



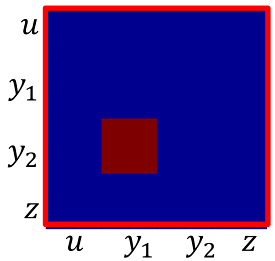
Sprung



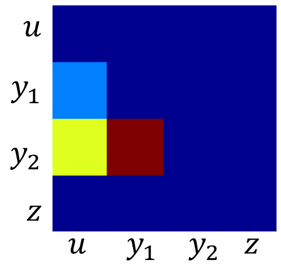
Sprungfähig



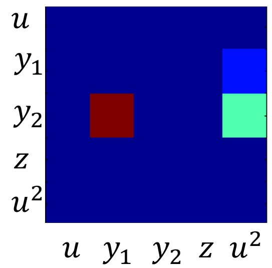
Quadrierer



Kubierer

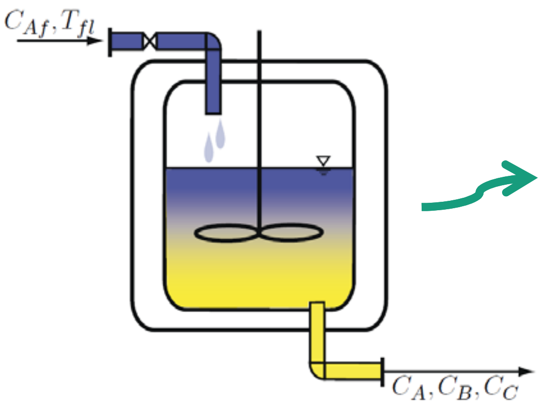


Quadrierer mit u^2

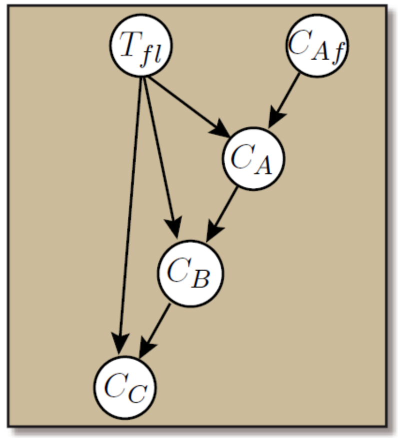


- Für lineare Systeme
- Bedingte Abhängigkeit $u \rightarrow y_1 \rightarrow y_2$ wegen Störrauschen nicht erkannt
- Implizite Schätzung der Totzeit → wird möglicherweise nicht erkannt

Beispiel: Chemiereaktor



$$\dot{C}_A = \frac{F}{V}(C_{Af} - C_A) - k_1 C_A e^{-E_1/RT}$$
$$\dot{C}_B = k_1 C_A e^{-E_1/RT} - k_2 C_B e^{-E_2/RT} - \frac{F}{V} C_B$$
$$\dot{C}_C = k_2 C_B - \frac{F}{V} k_1 C_C$$



- Kontinuierlicher Prozess*
  - $C_{Af} \rightarrow C_A \rightarrow C_B \rightarrow C_c$
  - T_{fl} hat Einfluss auf alle Reaktionen über Exponentialfunktion
  - Anregung mit weißem Rauschen für C_{Af} , T_{fl}

Lernen der unterliegenden Reaktionsstruktur

M. Tenny: *Closed-loop Behavior of Nonlinear Model Predictive Control, Technical report 2002-04*

Chemiereaktor

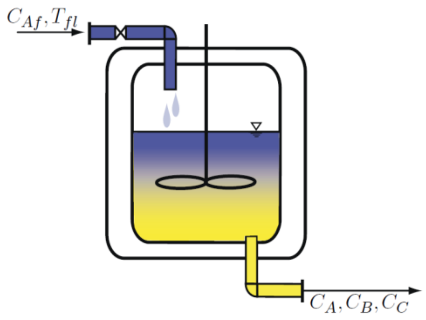
Kreuzkorrelationsfunktion

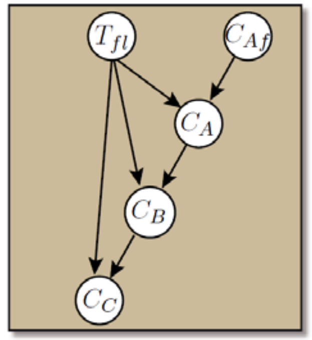
Erkennen kausales Verhalten

- Betrachtung der Zeitverschiebung λ
- Höhe des Maximum
- Vergleich der Flächen für $\lambda < 0$ und $\lambda > 0$

Durchführung von 2 Signifikanztests

Grundstruktur



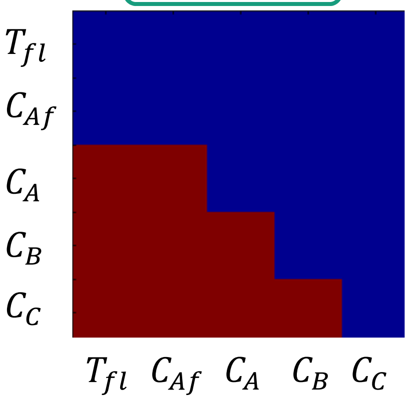


Granger-Kausalität

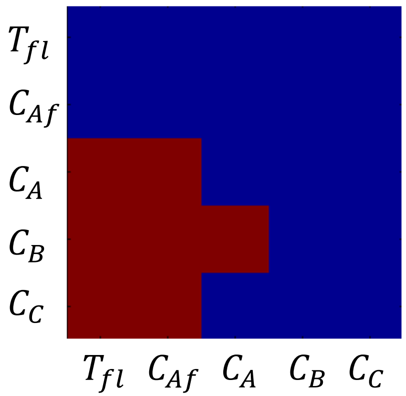
- Darstellung der bestandenen Signifikanztest

Kreuzkorrelationsfunktion

Erwartet

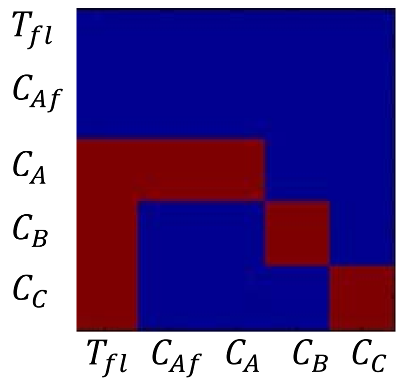


Gefunden

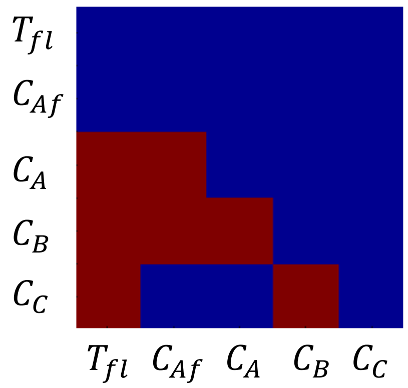


Granger-Kausalität

Erwartet

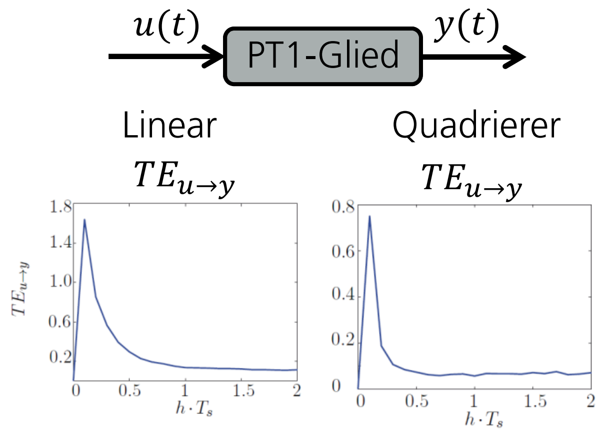


Gefunden

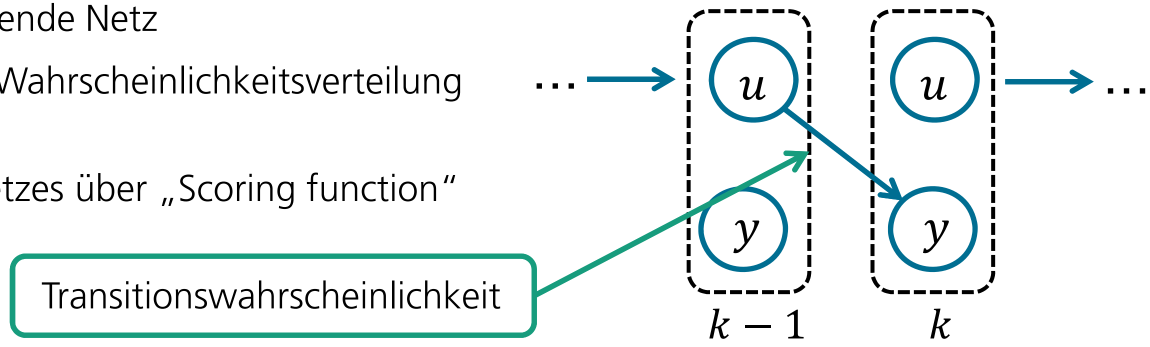


Weitere Methoden

- **Transfer Entropy**
  - Messdaten sind in gewisser Weise stochastisch
  - Berechnung bedingter Wahrscheinlichkeiten für zeitlich verschobene Zeitreihen.
  - Vergleich des Informationsgehaltes für zeitverschobene Zeitreihen



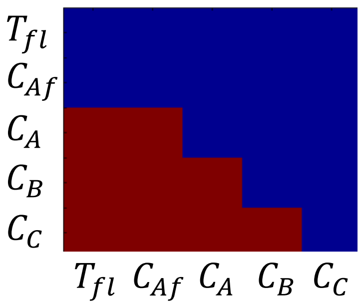
- **Dynamisches Bayes-Netz**
  - Zeitreihen werden um einen Zeitschritt verschoben
    → Verdopplung der Knoten für zu lernende Netz
  - Jeder Knoten bekommt eine bedingte Wahrscheinlichkeitsverteilung zugeordnet
  - Berechnung des wahrscheinlichsten Netzes über „Scoring function"



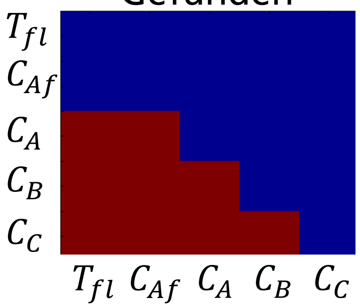
Chemiereaktor

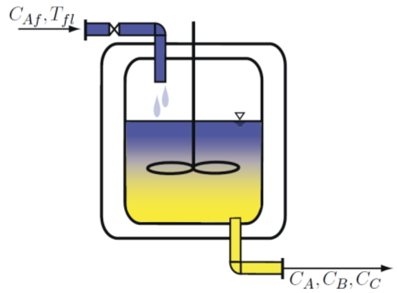
Transfer entropy

Erwartet



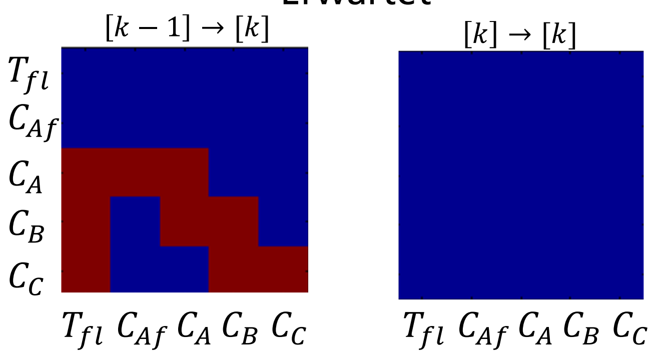
Gefunden



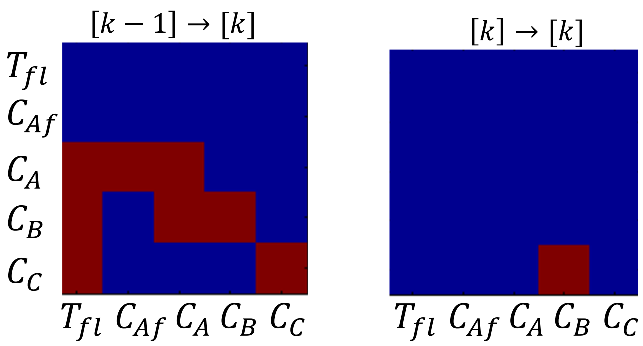


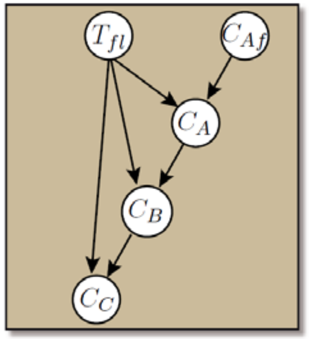
Bayes-Netze

Erwartet



Gefunden





Zusammenfassung

- Ziel: Erkennen kausaler Abhängigkeiten in dynamischen Systemen
  → Ausnutzen von Zeitverzögerungen

- Im Detail untersucht: Kreuzkorrelation, Granger-Kausalität, N4SID
- Als weitere Methoden vorgestellt: Transfer Entropy, Bayes-Netze

- Erstellen von Benchmarks (Tiefpass, Rührkessel) zum Vergleich und Bewerten der Verfahren
  - Je nach Aufgabenstellung bieten die einzelnen Verfahren jeweils Vor- und Nachteile

- Data-Mining kann mehr finden als reine Korrelationen!
- Es sind viele Methoden vorhanden, um kausale Strukturen in Daten dynamischer Systeme zu finden. Man muss sie nur nutzen.

Zusammenfassung

	KKF	Granger Kausalität	N4SID
Basissystem	√	√	√
Zus. Totzeit	√	√	√
Farbiges Eingangsrauschen	(√)	√	√
Überlagerter Eingangssprung	(√)	√	√
Zusätzlicher Vorhalt	(√)	-	-
Quadrierer	-	-	-

	KKF	Granger Kausalität	N4SID
Lineare Systeme	√	√	√
Integrierte Ordnungsschätzung	-	(√) (mit AIC)	√
Wiener/Hammerstein-Systeme	(√)	(√)	(√)
Umgang mit Mehrgrößensystemen	-	√	√
Schätzung der Totzeit	√	(√) (indirekt)	(√)(indirekt)
Einfache Parametrierung	√	(√)	√