

Vorlesung Diagnose- und Vorhersagesysteme - Lehrblätter -

Prof. Dr.-Ing. habil. Thomas Rauschenbach

Inhalt

1. Einführung

- Anwendungsbeispiele
- Aufgaben der Diagnose und Vorhersage in Entscheidungssystemen
- Definitionen

2. Diagnosestrategien

- Grundlegende Verfahren der Diagnose
 - signal- und zustandsbasierte Verfahren
 - Verwendung von Systemmodellen
- Diagnosekonzepte
 - Klassifikationsverfahren
 - Modellreferenzverfahren
 - wissensbasierte Verfahren

3. Vorhersagestrategien

- Grundlagen
- Vorhersageprinzipien
- Vorhersagbarkeit
- Prognoseprozess
- Primärdatenaufbereitung
- Vorhersage mit Signalmodellen
 - Vorhersage mit deterministischen Signalmodellen
 - Vorhersage mit stochastischen Signalmodellen
 - musterbasierte Vorhersage
 - Vorhersage mit neuronalen Netzen

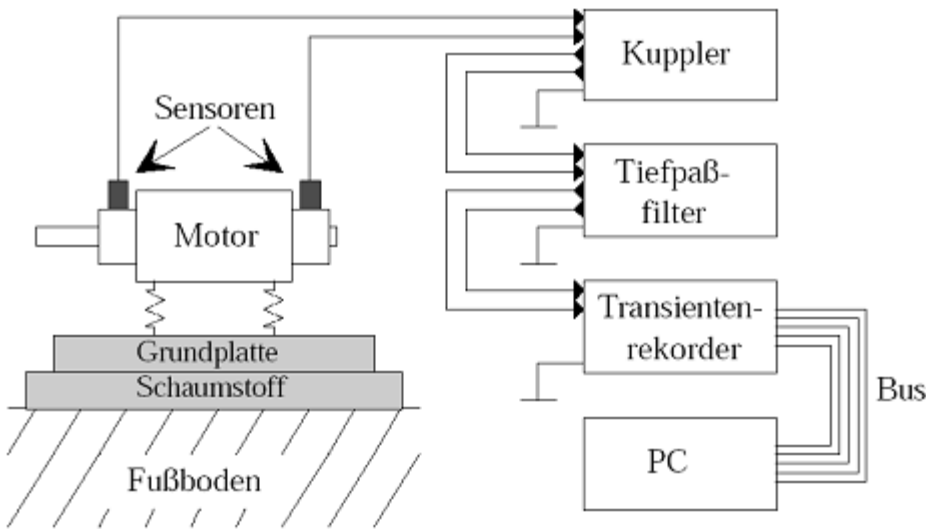
Literatur

- [BD96] Brockwell, P. J.; Davis, R. A.: *Introduction to Time Series and Forecasting*. New York: Springer-Verlag, 1996
- [Ise94] Isermann, Rolf: *Überwachung und Fehlerdiagnose*. VDI Verlag, 1994
- [JS93] Janacek, Gareth; Swift, Louise: *Time Series: Forecasting, Simulation, Applications*. New York, London, Toronto, Sydney, Tokyo, Singapore: Ellis Horwood, 1993
- [R+96] Romberg, T. [u. a.]: *Signal processing for industrial diagnostics*. Wiley, 1996
- [Sch01] Schlittgen, Rainer: *Angewandte Zeitreihenanalyse*. München, Wien: Oldenbourg Wissenschaftsverlag, 2001
- [SS01] Schlittgen, Rainer; Streitberg, Bernd H. J.: *Zeitreihenanalyse*. 9. Auflage. München, Wien: Oldenbourg Wissenschaftsverlag, 2001
- [Wer89] Wernstedt, Jürgen: *Experimentelle Prozessanalyse*. 1. Auflage. Berlin: Verlag Technik, 1989

1. Einführung

1.1 Anwendungsbeispiele

Vibrationsanalyse zur technischen Diagnose von Universalmotoren

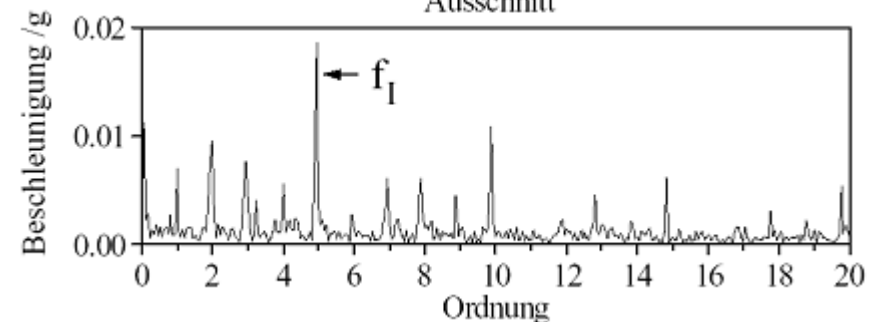
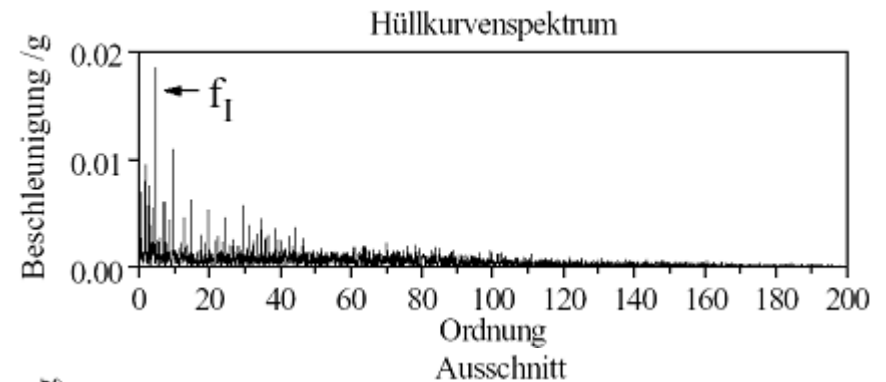


Typische Fehler können anhand der auftretenden Vibrationen erkannt werden:

$$\text{Außenringschaden: } f_A = 3,04 \cdot f_n$$

$$\text{Innenringschaden: } f_I = 4,96 \cdot f_n$$

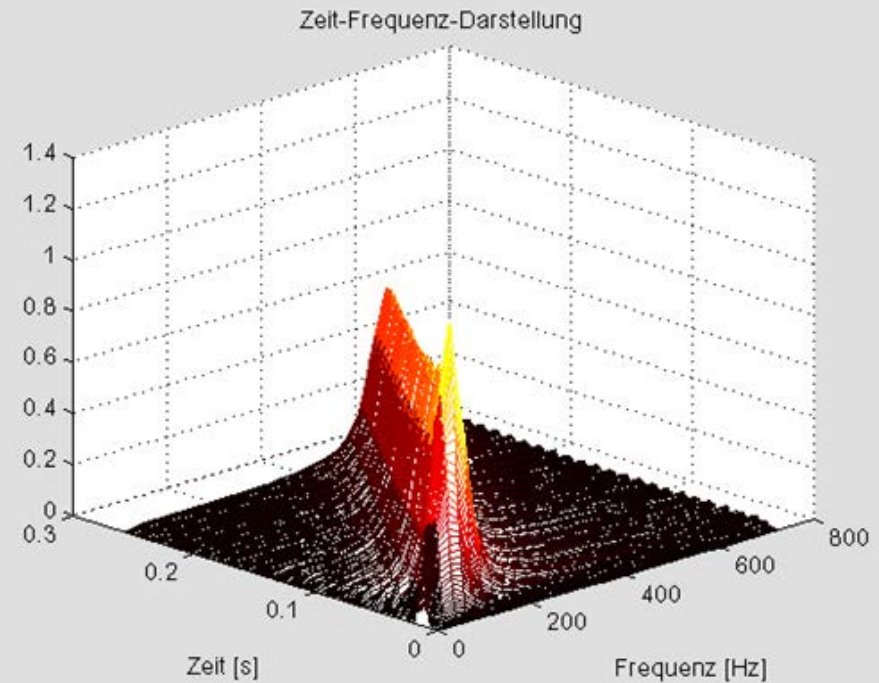
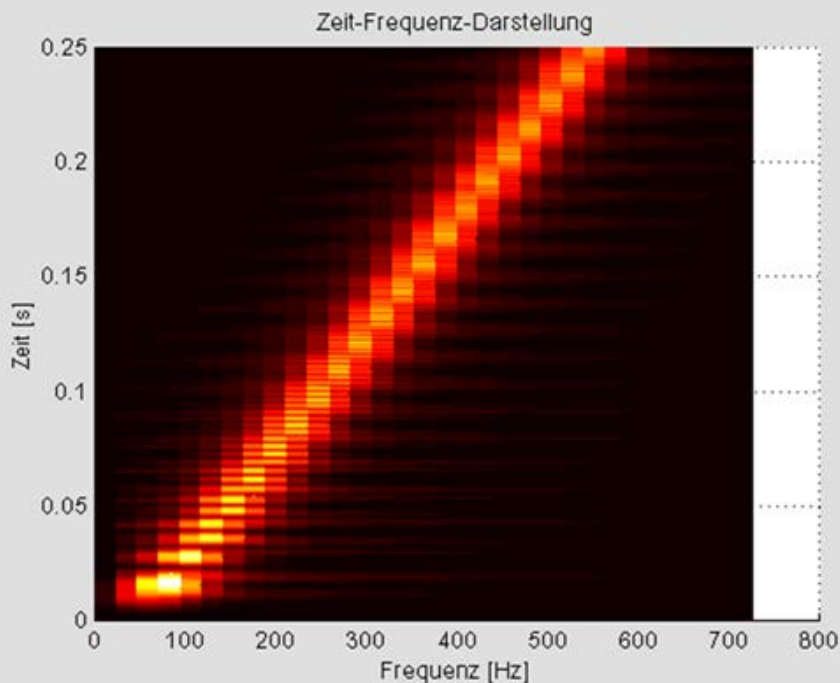
$$\text{Kugelschaden: } f_K = 3,94 \cdot f_n$$



1. Einführung

1.1 Anwendungsbeispiele

Diagnose mit Zeit-Frequenz-Darstellungen

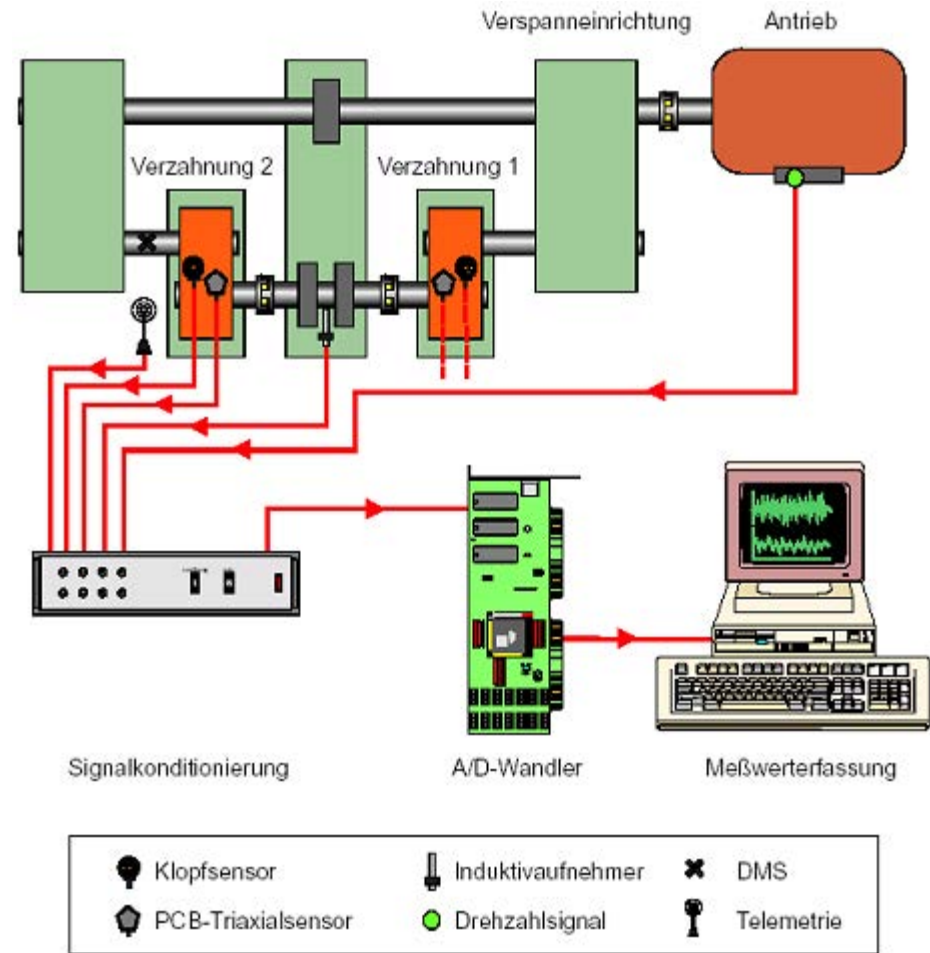


1. Einführung

1.1 Anwendungsbeispiele

Vorausschauende zustandsorientierte Instandhaltung

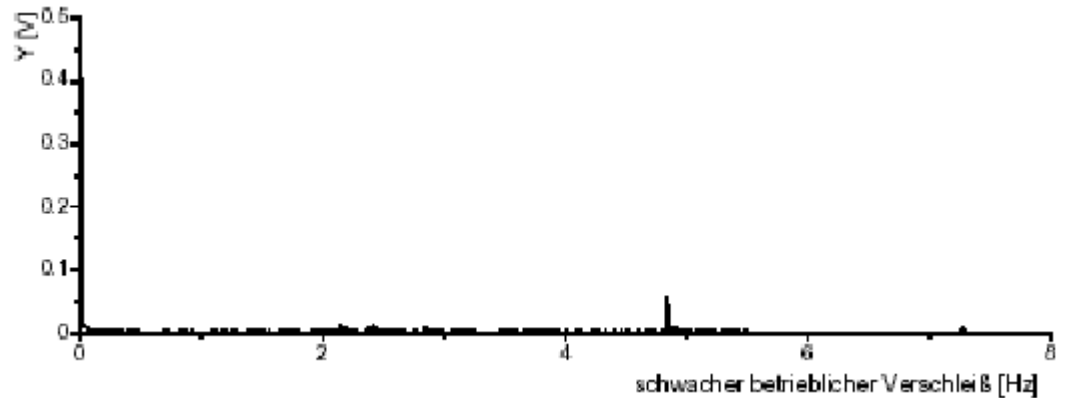
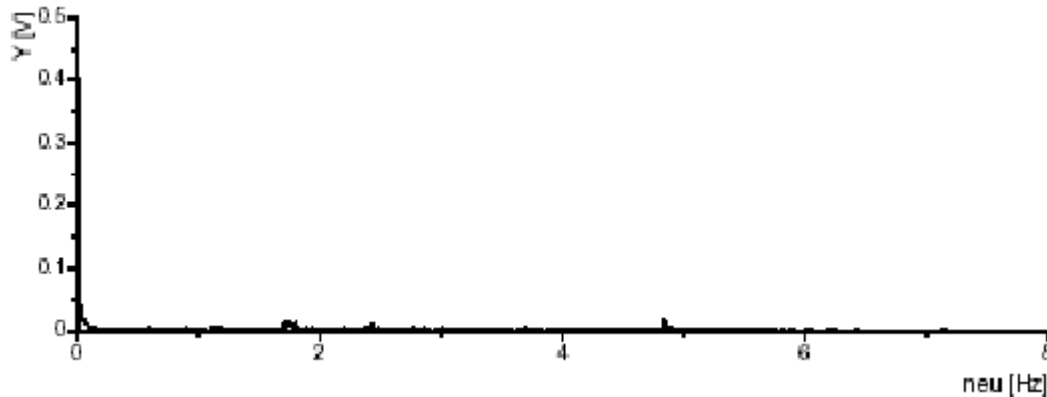
Um die Verfügbarkeit von Anlagen zu maximieren und die Ausfallzeiten und die Kosten der Instandhaltung auf ein Minimum zu reduzieren, wird heute immer mehr versucht, eine vorausschauende, zustandsorientierte Instandhaltung einzusetzen. Im Gegensatz zu ausfallorientierten oder rein vorbeugenden Maßnahmen birgt diese ein erhebliches Rationalisierungspotential und findet deshalb bei geplanten Betriebsabläufen immer häufiger ihre Anwendung.



1. Einführung

1.1 Anwendungsbeispiele

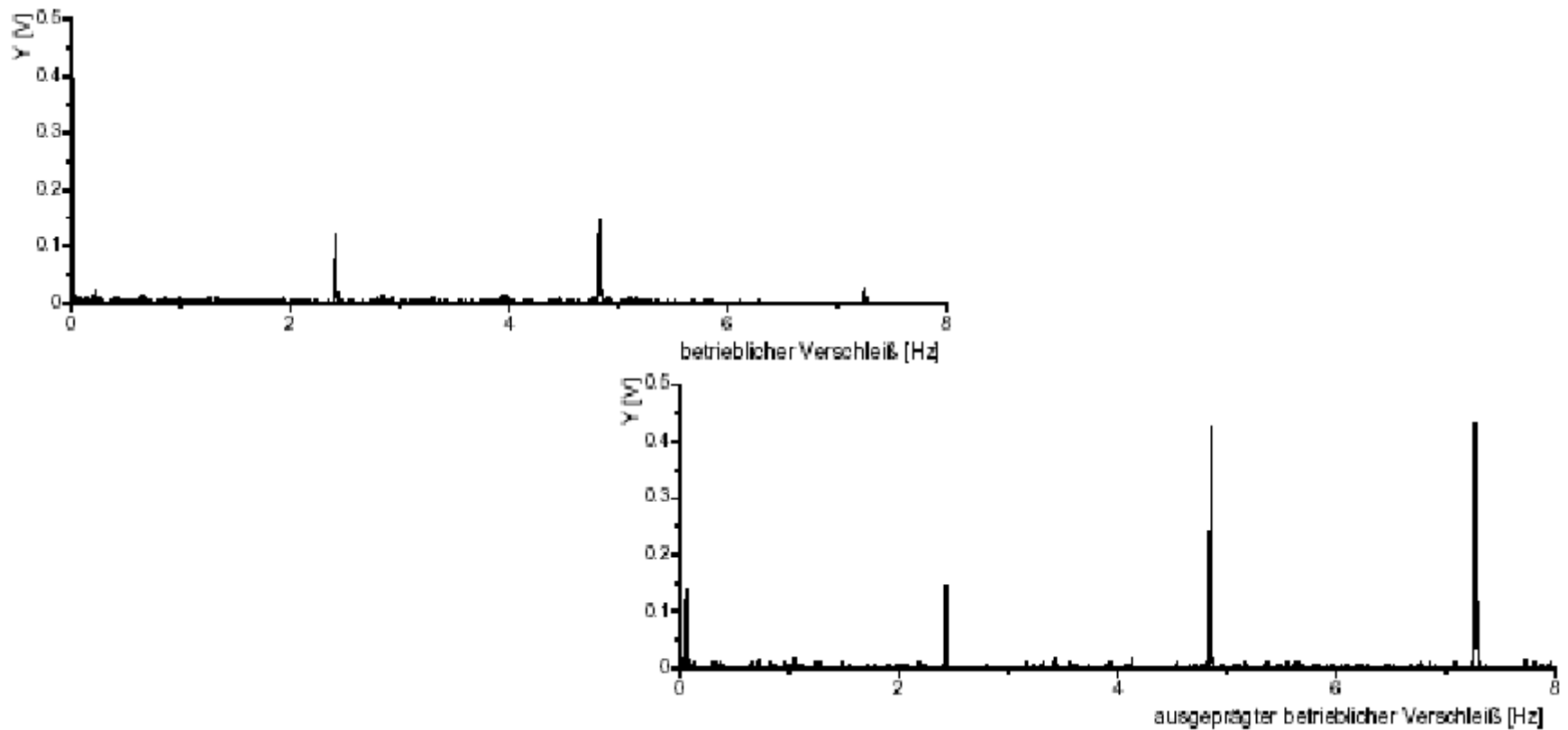
Erkennen von Verschleißzuständen im Spektrum



1. Einführung

1.1 Anwendungsbeispiele

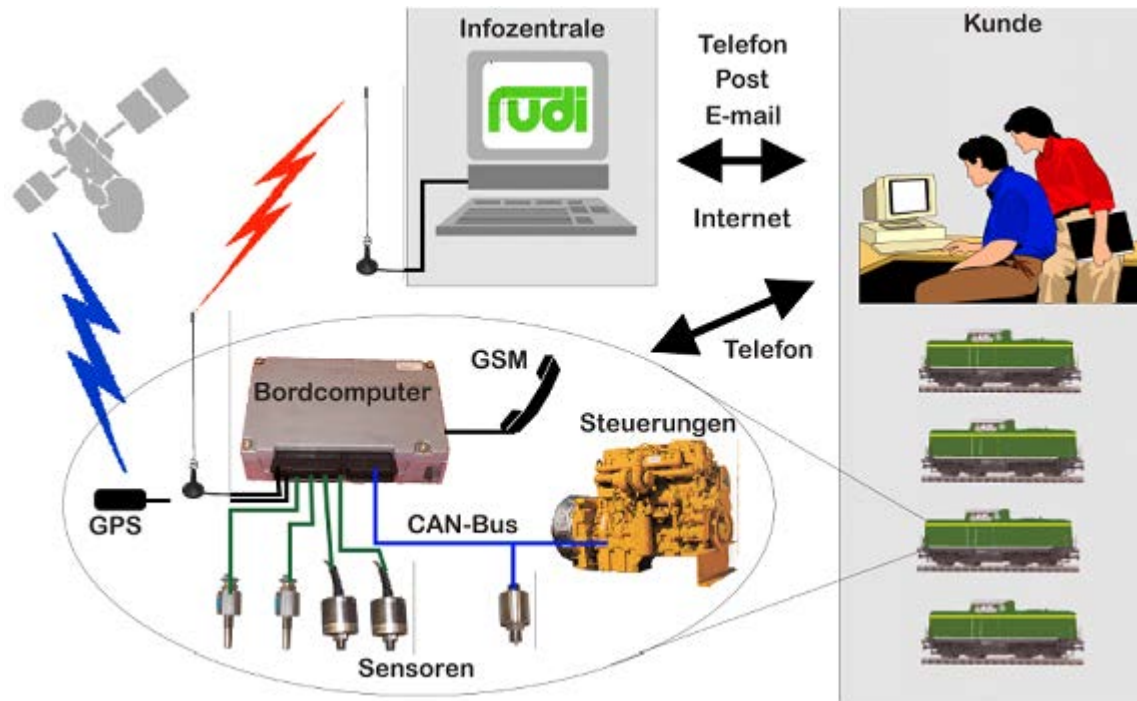
Erkennen von Verschleißzuständen im Spektrum



1. Einführung

1.1 Anwendungsbeispiele

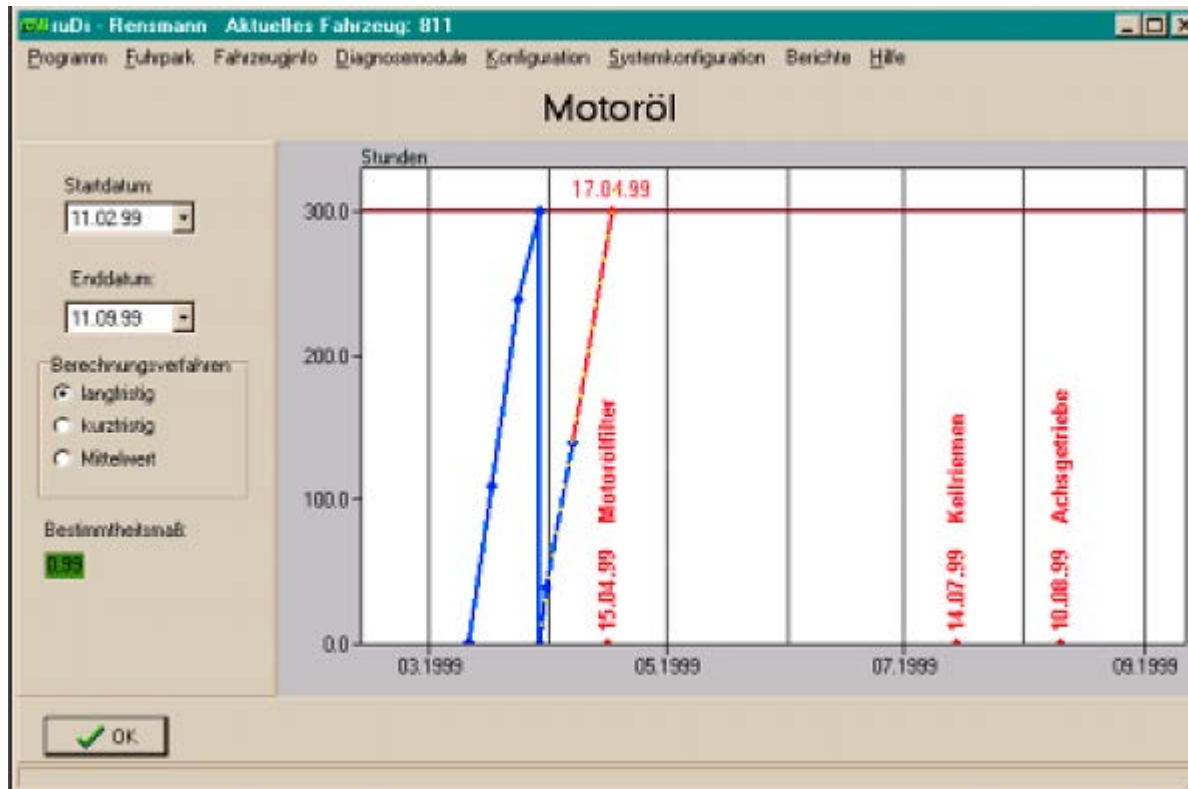
Fuhrparkinstandhaltung mit Telediagnose



1. Einführung

1.1 Anwendungsbeispiele

Fuhrparkinstandhaltung mit Telediagnose



1. Einführung

1.2 Aufgaben der Diagnose und Vorhersage in der Automatisierung

I. Prozessüberwachung

- Erfassung, Anzeige und Melden der aktuellen Prozesszustände
- Bildung, Anzeige und Melden von Abweichungen / Fehlern von Normzuständen

II. Prozesssicherung

- Ableitung und Realisierung von Gegenmaßnahmen, damit aus erkannten Fehlern keine Störung oder Ausfall entsteht

III. Prozessstabilisierung

- Ableitung und Realisierung von Steuerungen / Regelungen zur Beseitigung von Fehlern bei Störungen oder Umsteuerung (prädiktive Regelungen)

IV. Prozessoptimierung

- Ableitung und Realisierung optimaler Führungsgrößen

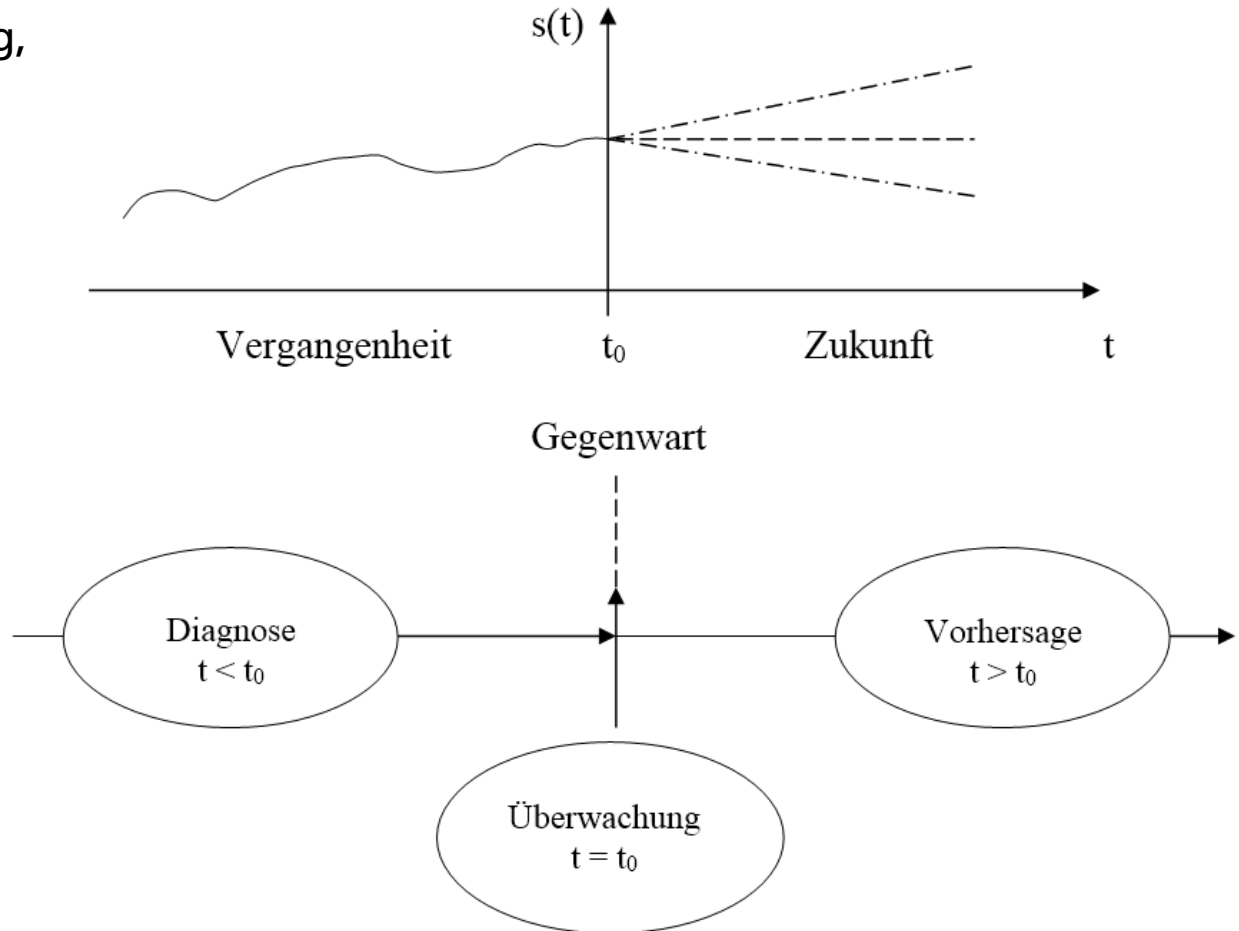
1. Einführung: 1.2 Aufgaben der Diagnose und Vorhersage

Aufgabenzerlegung

a) zeitliche Zerlegung
(Diagnose, Überwachung,
Vorhersage)

b) tätigkeitsorientierte
Zerlegung
(Erkennen: Diagnose,
Überwachung;
Handeln: Überwachung,
Steuerung)

a) zeitliche Zerlegung



1. Einführung: 1.2 Aufgaben der Diagnose und Vorhersage

Aufgabenzerlegung b) tätigkeitsorientierte Zerlegung

Stufe I: Erkennen

aus Bewertung von $\underline{s}(t)$ für $t < t_0$
-> Diagnose

aus Bewertung von $\underline{s}(t)$ für $t = t_0$
-> Überwachung

aus Bewertung von $\underline{s}(t)$ für $t > t_0$
-> Vorhersage

Stufe II: Handeln

auf Grundlage der Diagnose
-> Therapie
(nachträgliches Handeln)

auf Grundlage der Überwachung
-> Prozesssicherung /- stabilisierung

auf Grundlage der Vorhersage
-> Prädiktive Therapie / Prozess-
stabilisierung

1. Einführung: 1.3 Definitionen

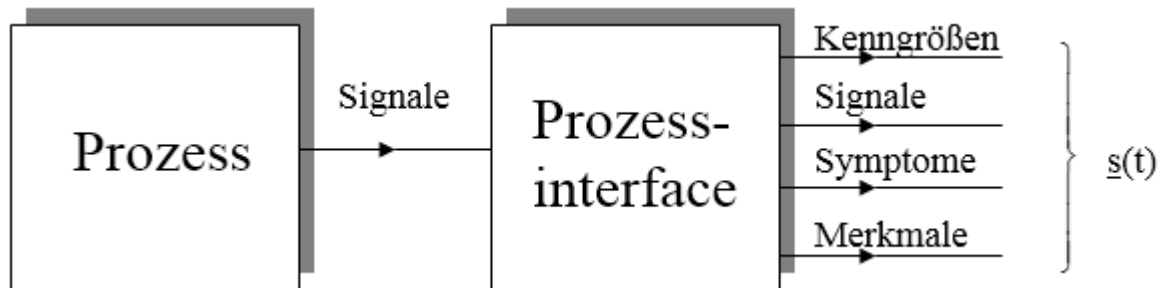
Prozesszustand, Fehler, Störungen, Ausfälle

1. Prozesszustand $\underline{s}(t)$:

wird aus den erfassten Signalen und Zuständen des Prozesses über ein Prozessinterface gebildet

kann enthalten: Signale ($u(t)$, $y(t)$, $z(t)$, $x(t)$) / Kenngrößen / Symptome / Merkmale

Allgemeine Prozessinterfacestruktur:



Hinweis:

Begriffe in Anlehnung VDI/VDE Richtlinie Nr. 3542 und DIN 40042

1. Einführung: 1.3 Definitionen

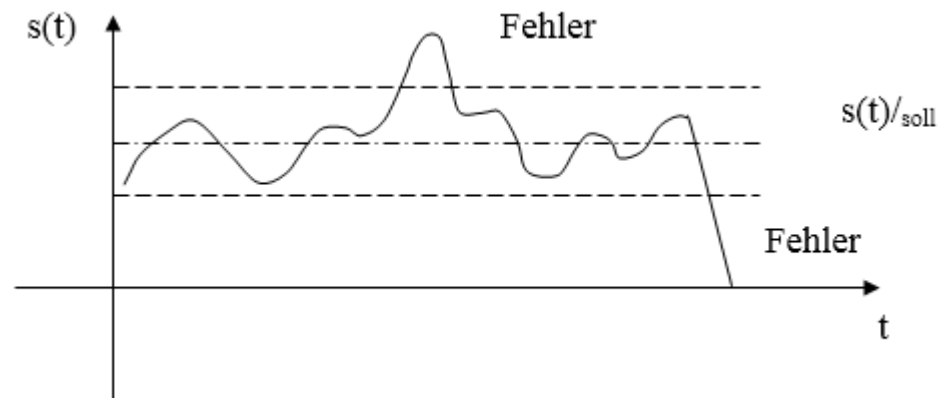
Prozesszustand, Fehler, Störungen, Ausfälle

2. Fehler:

= unzulässige Abweichung mindestens eines Elementes des Prozesszustandes $\underline{s}(t)$

Eigenschaften:

1. der Fehler ist ein Zustand,
2. die unzulässige Abweichung ist der über einen Toleranzbereich zwischen Ist- und Sollwert hinausgehende Wert,
3. der Fehler kann eine Störung oder einen Ausfall zur Folge haben.



1. Einführung: 1.3 Definitionen

Prozesszustand, Fehler, Störungen, Ausfälle

3. Störung:

= nach Beanspruchungsbeginn (Inbetriebnahme) entstandener vorübergehender Fehler

Eigenschaften:

1. vorübergehende Unterbrechung oder Beeinträchtigung der Funktion
2. Beanspruchungsbeginn: Betriebsbeginn, Abnahmeprüfung

1. Einführung: 1.3 Definitionen

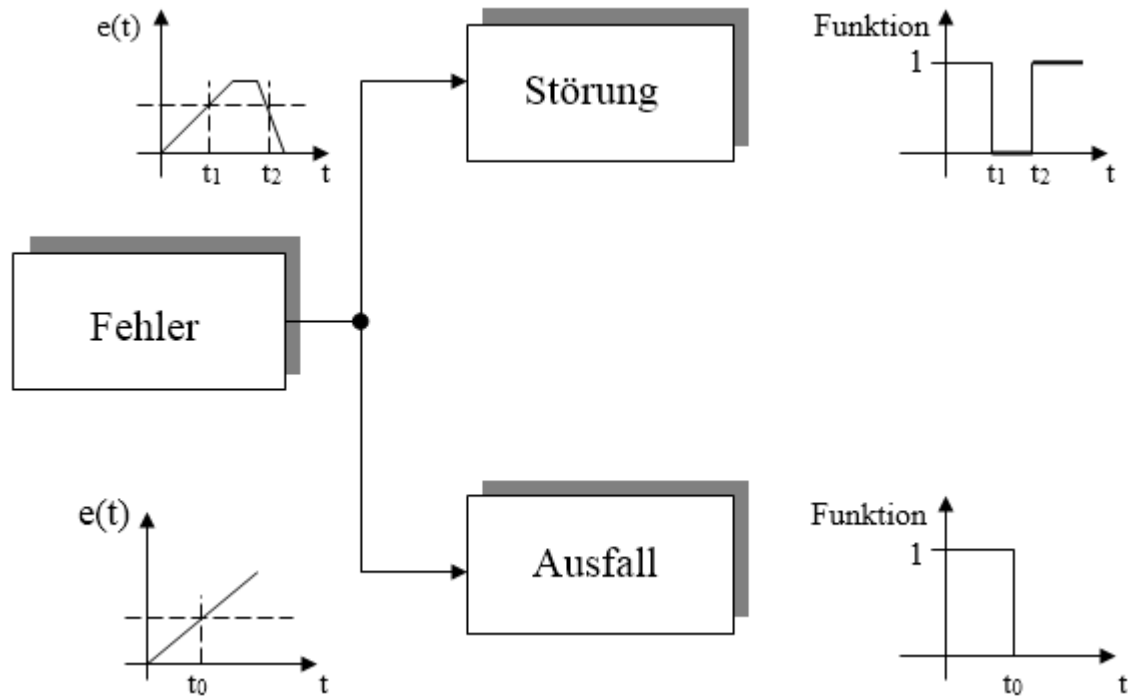
Prozesszustand, Fehler, Störungen, Ausfälle

4. Ausfall:

= nach Beanspruchungsbeginn entstandenes Aussetzen der Ausführung einer Aufgabe der Prozesseinheit aufgrund einer in ihr selbst liegenden Ursache

Eigenschaften:

1. der Ausfall ist ein Ereignis
2. der Ausfall entsteht durch einen oder mehrere Fehler
3. Unterscheidung in Zufallsausfall (nicht vorhersagbar) und deterministischer Ausfall (vorhersagbar)



-> Gesamtzusammenhang:
Fehler, Störung, Ausfall

1. Einführung: 1.3 Definitionen

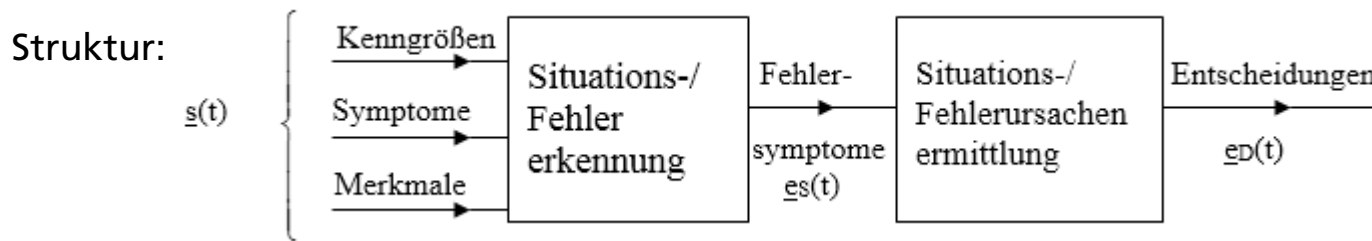
Diagnose und Vorhersage

1.3.1 Diagnose

= das Erkennen / Bewerten der Abweichung (Fehler) zwischen dem Zustand / Verhalten eines Systems und dem Sollzustand / -verhalten für den Zeitraum $t \leq t_0$

Grundlage:

Informationen zu Mengen- und / oder Qualitätsparametern für den Zeitraum $t \leq t_0$ in Form von Daten und Modellen



mit $\underline{s}(t) \triangleq$ Situationsvektor
 $e_s(t) \triangleq$ Fehlersymptomvektor
 $e_D(t) \triangleq$ Diagnoseentscheidungsvektor

1. Einführung: 1.3 Definitionen

Diagnose und Vorhersage

1.3.2 Vorhersage

= Vorhersage ist die Bestimmung / Schätzung von Mengen- und / oder Qualitätsparameter von Systemen für den Zeitraum $t > t_0$ (Zukunft) auf der Basis von Modellen

Grundlage:

Daten und Modelle von Signalen und Systemen für den Zeitraum $t > t_0$

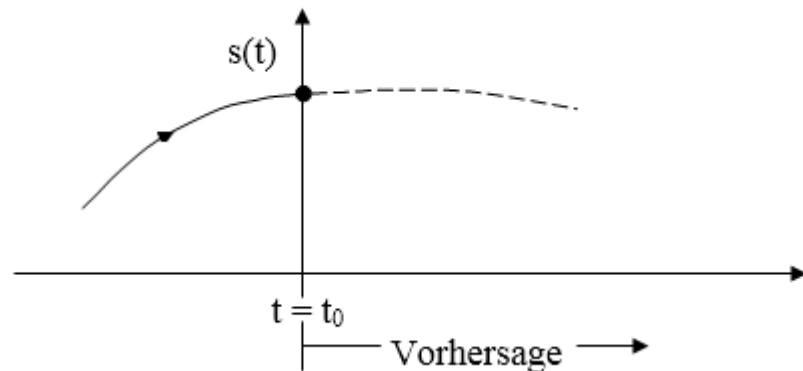
a) deterministische Situation:

gegeben:

$$\underline{s}(t) = \underline{f}[u(t), \underline{x}(t), y(t)]$$

$u(t)$ für $t \geq t_0$

$\underline{x}(t_0)$



gesucht:

$$\underline{\hat{s}}(t) = \underline{f}[u(t), \underline{x}(t_0), y(t)]$$

1. Einführung: 1.3 Definitionen

Diagnose und Vorhersage

b) stochastische Situation:

gegeben:

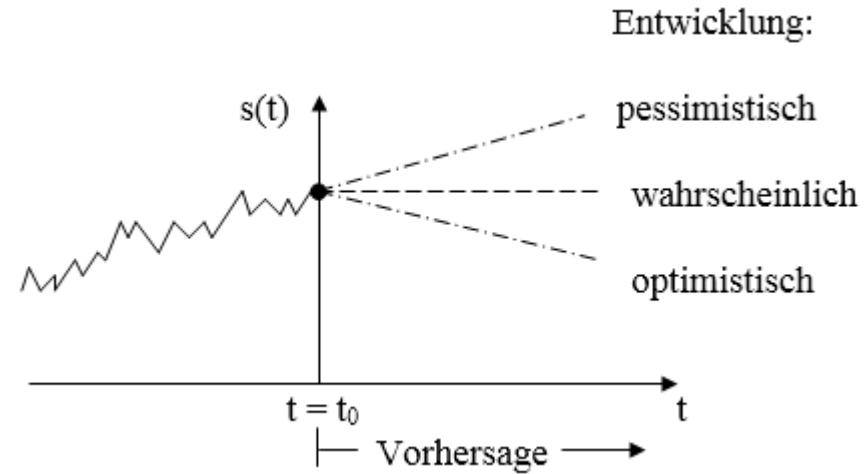
$$\underline{s}(t) = f[u(t), \underline{x}(t), y(t)]$$

$u(t)$ für $t \geq t_0$

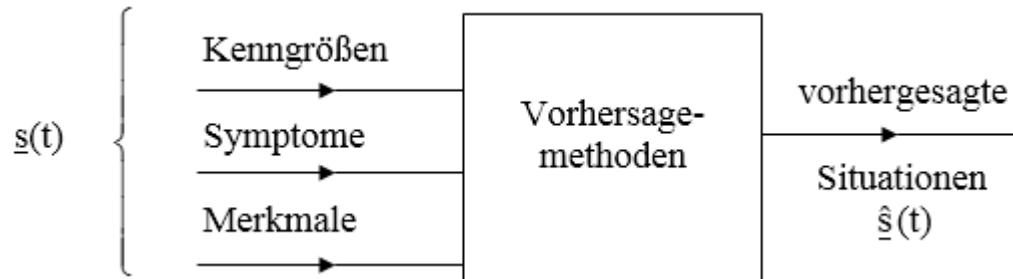
$\underline{x}(t_0)$

gesucht:

$\hat{\underline{s}}(t)$ für $t \geq t_0$ als Szenario



Struktur:

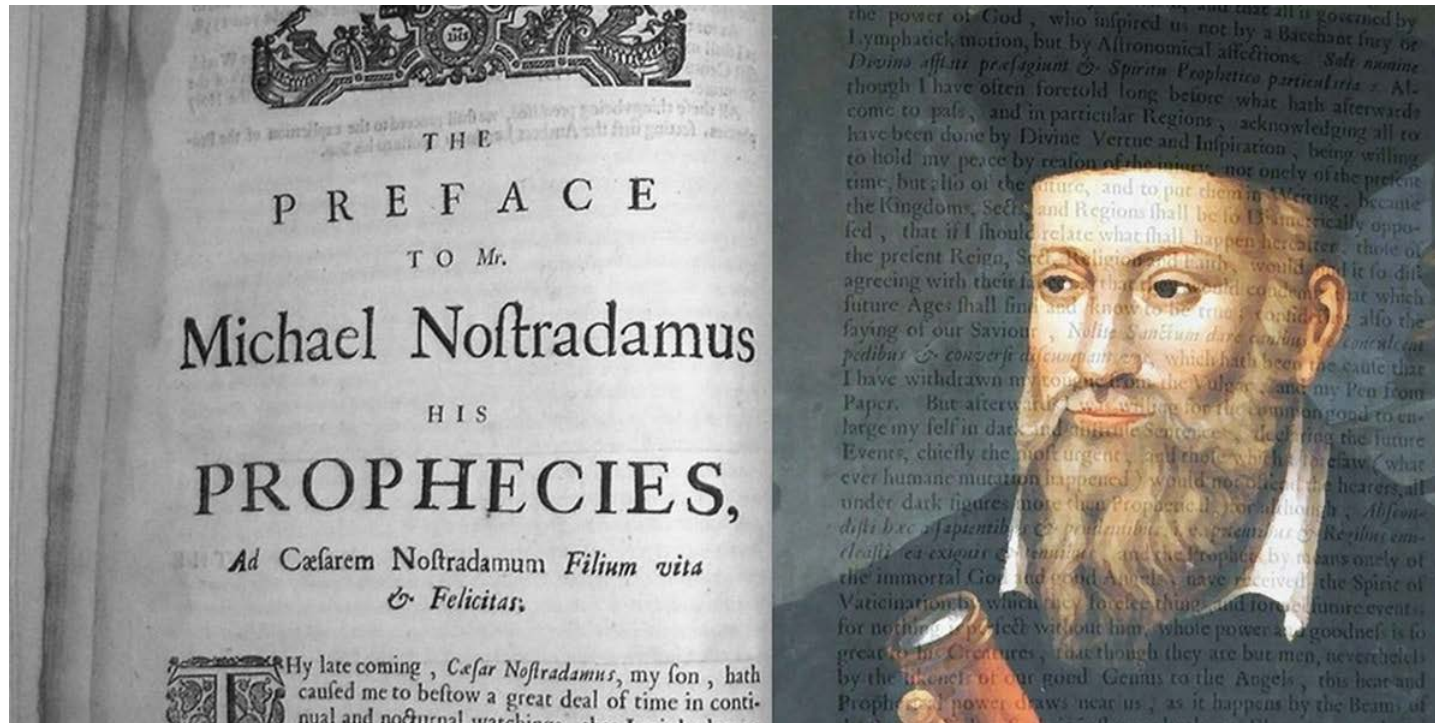


mit $\underline{s}(t) \triangleq$ Situationsvektor für $t \leq t_0$

$\hat{\underline{s}}(t) \triangleq$ vorhergesagter Situationsvektor für $t > t_0$

1. Einführung: 1.3 Definitionen

Diagnose und Vorhersage



Was bringt uns die Zukunft?

1. Einführung: 1.3 Definitionen

Diagnose und Vorhersage



Bauernregeln

2. Diagnosestrategien

2.1. Diagnosegrundlagen

Ausgangssituation

- Überwachung von Systemen bzw. Prozessen
- Modelle von Signalen- und / oder Systemen
- Informationen zur Funktionen (normal / unnormale)

Grundlagen - Methoden der Fehlererkennung und -bewertung

1. Überwachung direkt messbarer Zustände / Signale
2. Analyse von Signalmodellen
3. Analyse von Systemmodellen
4. Ermittlung von Kennwerten

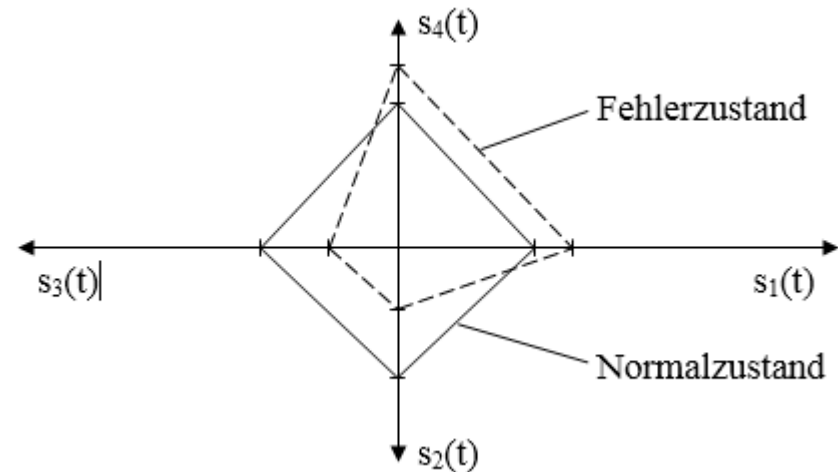
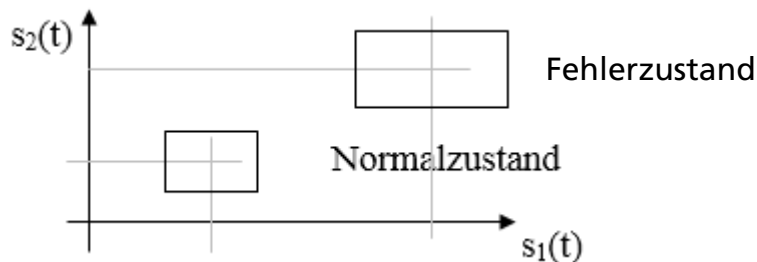
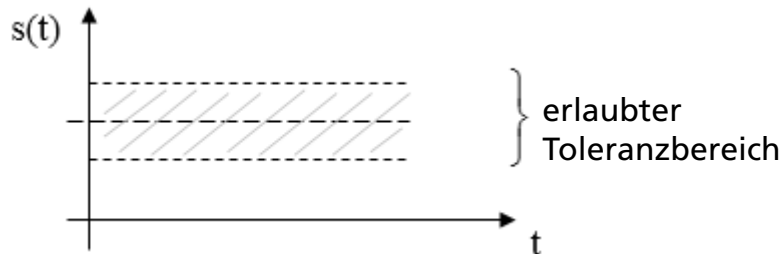
-> direkte Anwendung der Modelle und Methoden der
Modellbildung und
Systemanalyse

2. Diagnosestrategien

2.1.1. Auswertung von Signalen und Zuständen

Grundlagen: gemessene Signale / Zustände im Zeitraum $t \leq t_0$
Vorgabe von Normbereichen bzw. Normklassen

Darstellungen:



2. Diagnosestrategien

2.1.1. Auswertung von Signalen und Zuständen

Signal- / Zustandsparameter:

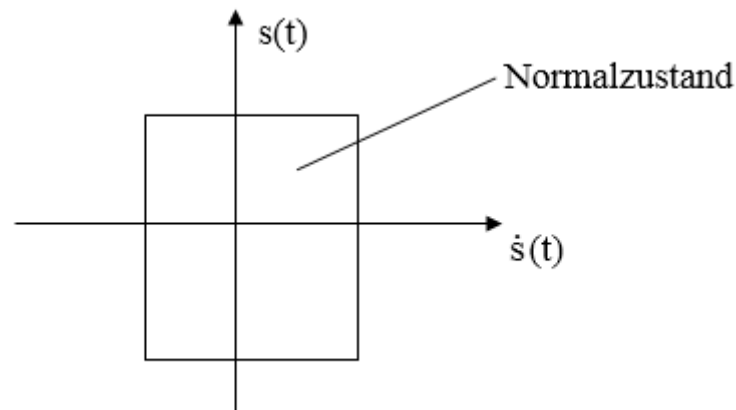
Absolutwerte von $\underline{s}(t)$ $s_{min} < s(t) < s_{max}$ d. h. Toleranzbänder

Trendwerte von $\underline{s}(t)$ $\dot{s}_{min} < \dot{s}(t) < \dot{s}_{max}$ d. h. Toleranzbänder

Kombinatorische Absolutwert- und Trendwertanalyse

$$s_{min} < s(t) < s_{max}$$

$$\dot{s}_{min} < \dot{s}(t) < \dot{s}_{max}$$



2. Diagnosestrategien

2.1.1. Auswertung von Signalen und Zuständen

Beschreibung von $\underline{s}(t)$:

1. Amplitudenkennwerte - Statik: Verteilung und Momente
Korrelationskoeffizienten

2. Dynamische Kennfunktionen: Korrelationsfunktion
Trendbeschreibung
Spektralfunktion
Ceptrumfunktion (Frequenz – Zeit – Analysen)

- Darstellungen:
- Fenstertechniken
 - MAP'-Techniken
 - GIS-Techniken, d. h. koordinatengestützt

2. Diagnosestrategien

2.1.2. Verwendung von Signalmodellen

Ausgangspunkt: Aufbau von Signalen aus Superpositionen von Teilkomponenten

Ansatz: $x(t) = \sum_{i=1}^4 x_i(t)$ mit $x_1(t) \triangleq x_T(t)$ Trendanteil
 $x_2(t) \triangleq x_P(t)$ periodischer Anteil
 $x_3(t) \triangleq x_M(t)$ Musteranteil
 $x_4(t) \triangleq x_S(t)$ stochastischer Anteil

und den Grundstrukturen:

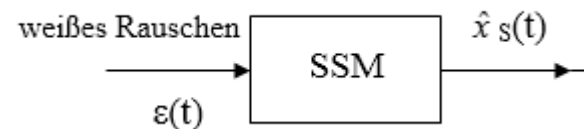
Deterministische Signalmodelle (DSM)



Mustermodell (MuM)



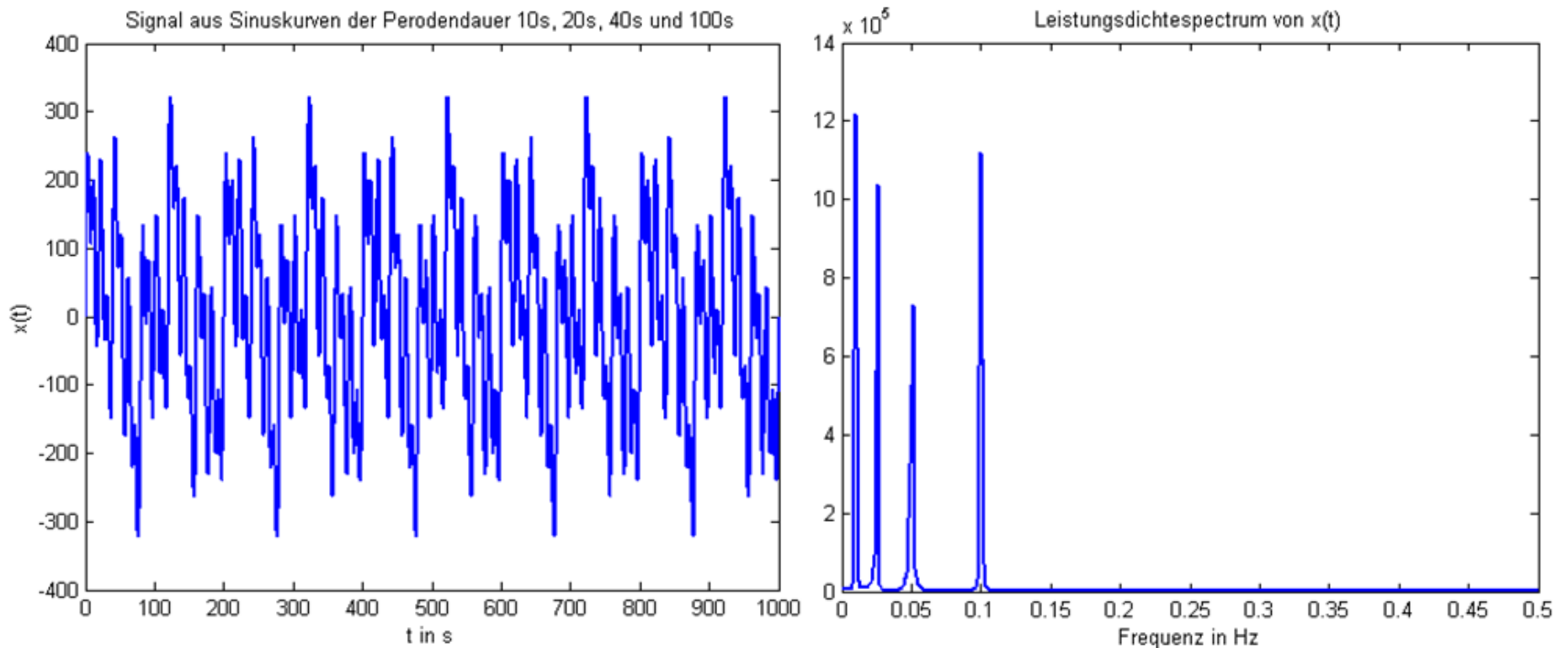
Stochastische Signalmodelle (SSM)



2. Diagnosestrategien

2.1.2. Verwendung von Signalmodellen

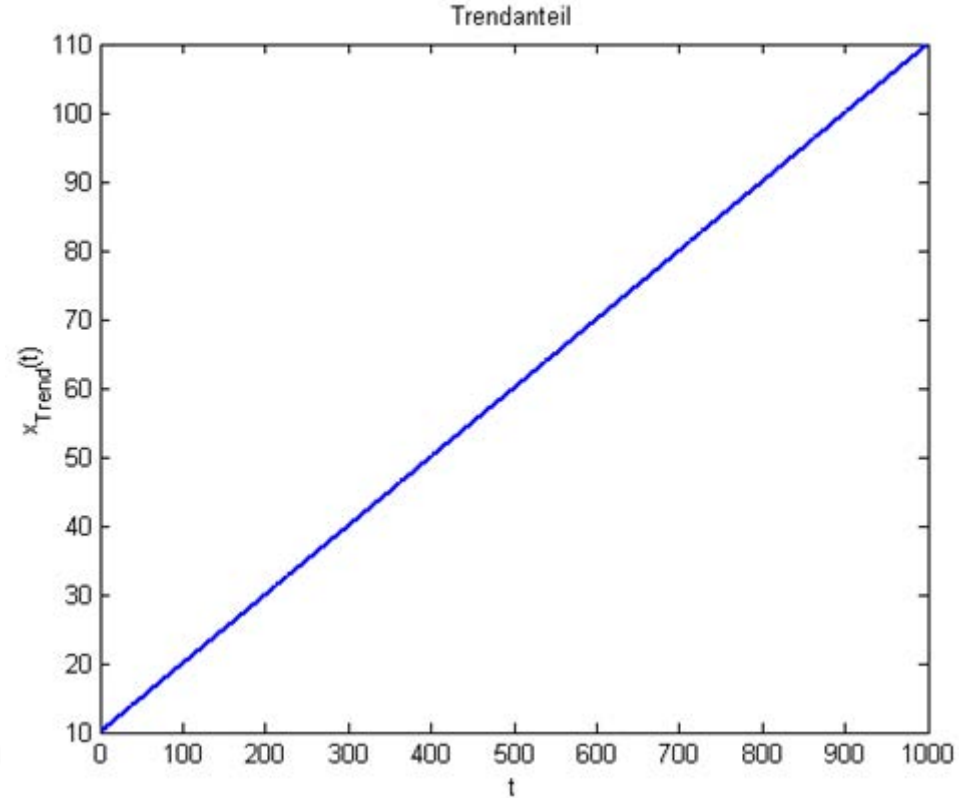
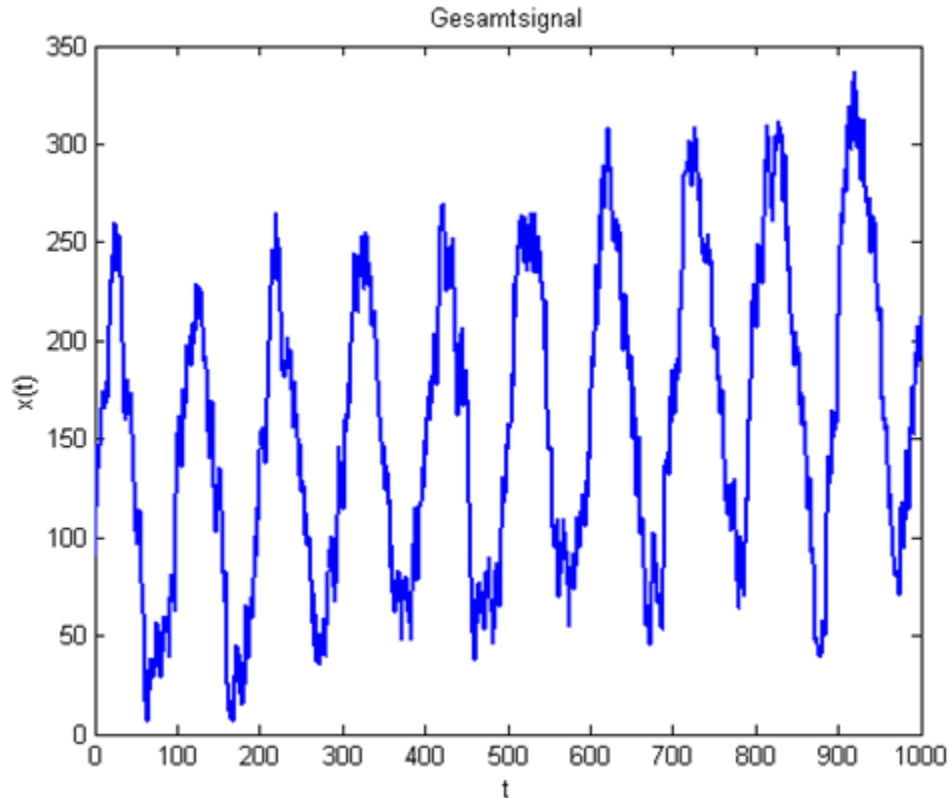
Leistungsdichtespektrum



2. Diagnosestrategien

2.1.2. Verwendung von Signalmodellen

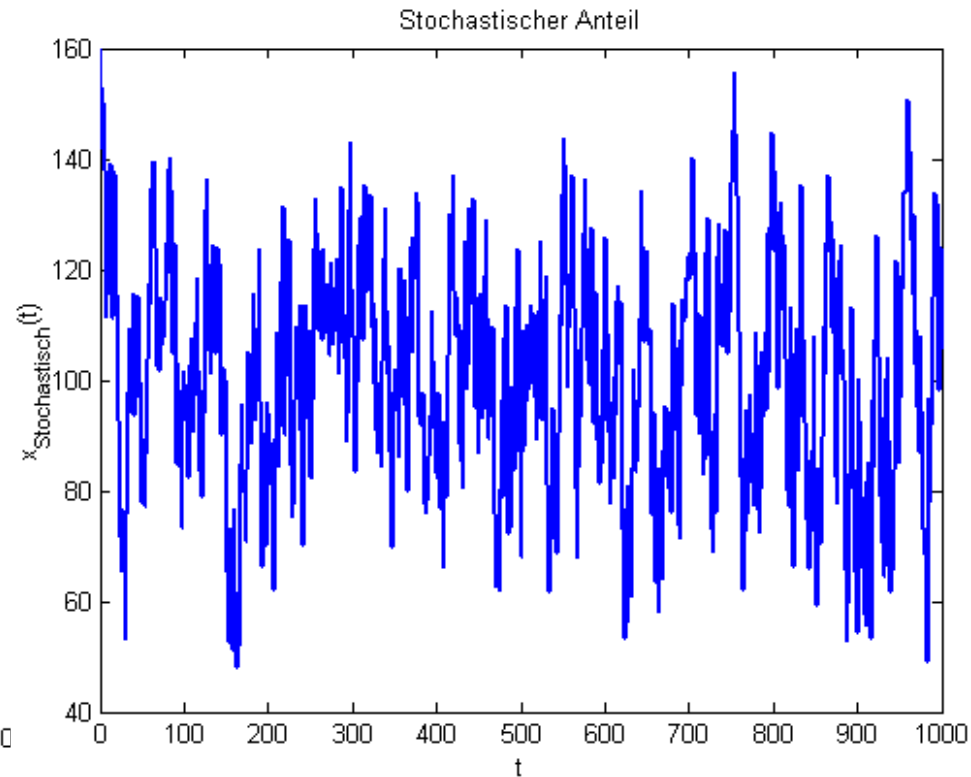
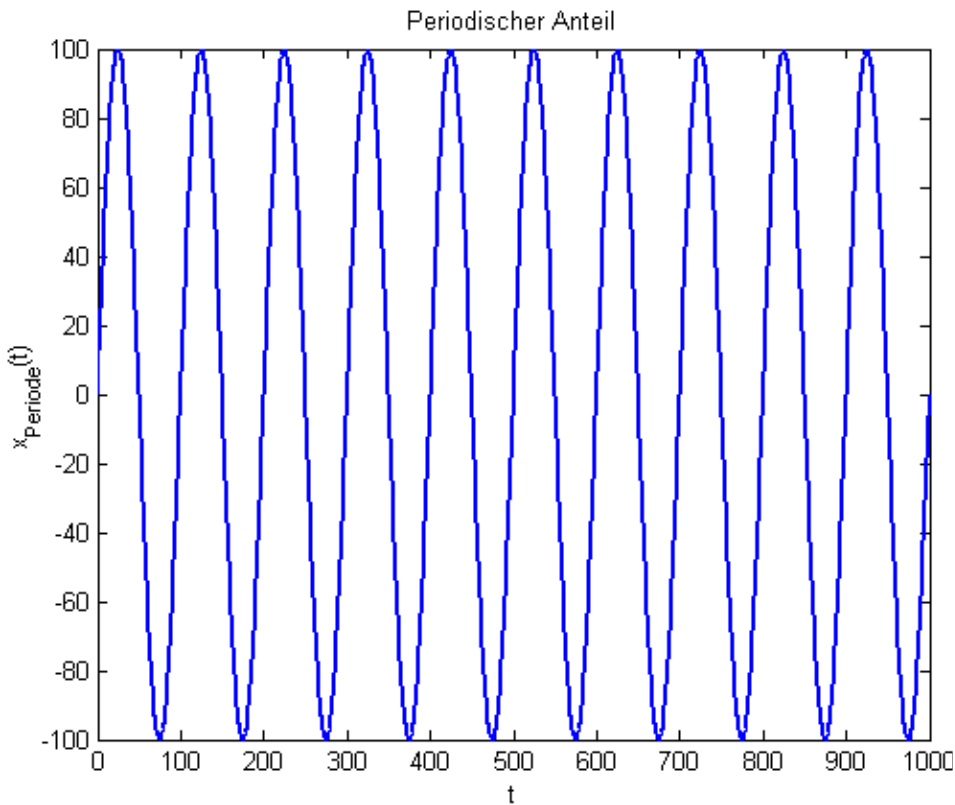
Komponenten eines Signals (1)



2. Diagnosestrategien

2.1.2. Verwendung von Signalmodellen

Komponenten eines Signals (2)



2. Diagnosestrategien

2.1.3. Verwendung von Systemmodellen

Ausgangspunkt: Bildung von Systemmodellen (Struktur- und Verhaltensmodelle) mit dem Ziel, einen Fehler zwischen Realität und Modell zu bilden und zu bewerten

Ansätze:

a) analytische Modelle (SISO-Typ)

Ein- / Ausgangsmodelle

$\hat{y} = f[u]$ Statik

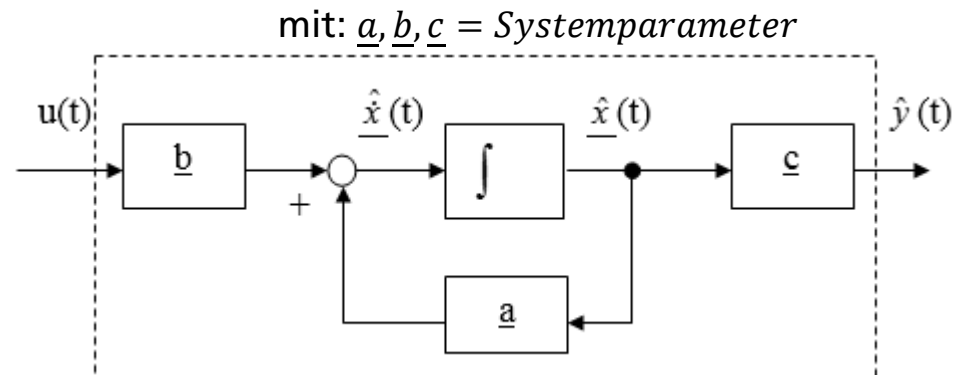
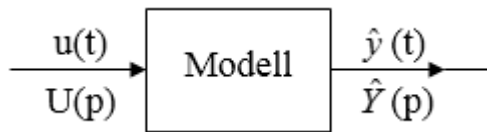
$\hat{y}(t) = f[u(t), y(t)]$ Dynamik / Zeitbereich

$\hat{y}(p) = f[U(p), Y(p)]$ Dynamik / Frequenzbereich

b) Zustandsmodelle

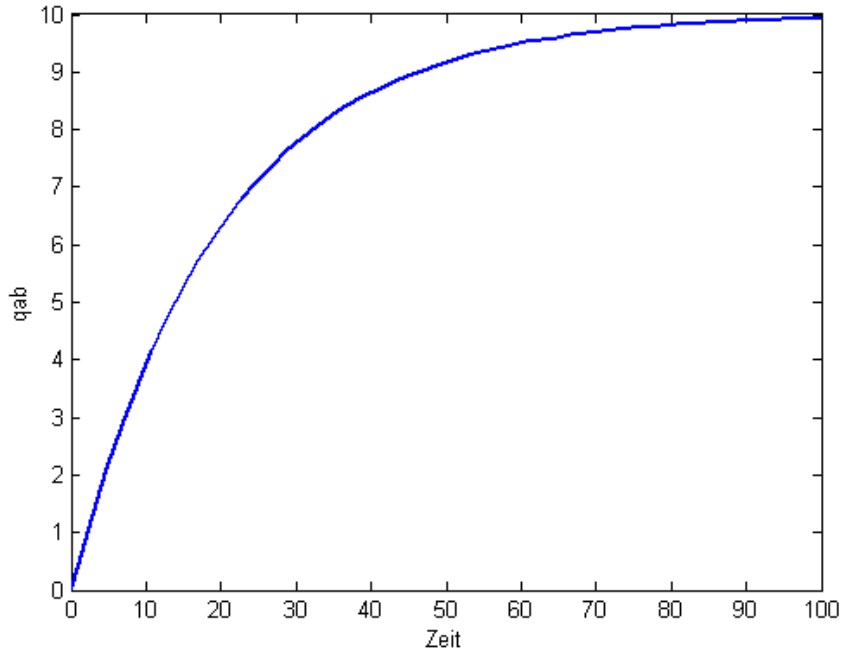
$$\dot{\hat{x}}(t) = \underline{a}^T \cdot \hat{x}(t) + \underline{b}^T \cdot u(t)$$

$$\hat{y}(t) = \underline{c}^T \cdot \hat{x}(t)$$

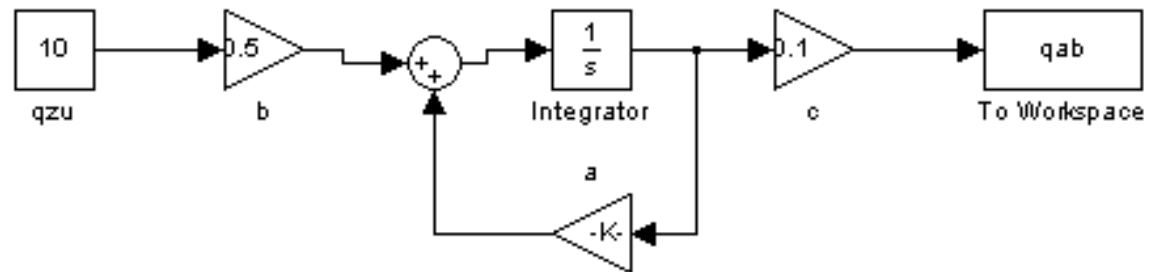


2. Diagnosestrategien

2.1.3. Verwendung von Systemmodellen



Zustandsraummodell eines Behälters



2. Diagnosestrategien

2.1.3. Verwendung von Systemmodellen

c) linguistische Modelle

Produktionssysteme / Regelwerke

-> zweiwertige Logik (scharf)

-> mehrwertige Logik (unscharf, fuzzy)

Wenn $\langle B_i \rangle$
Dann $\langle A_j \rangle$

mit $B_i \triangleq$ scharfe Bedingungen
 $A_j \triangleq$ scharfe Aktionen

mit $B_i \triangleq$ unscharfe Bedingungen
 $A_j \triangleq$ unscharfe Aktionen

d) konnektionistische Modelle

Petri-Netze für ereignisdiskrete
Beschreibung

künstliche Neuronale Netze für
kontinuierliche Beschreibung

2. Diagnosestrategien

2.1.4. Berechnung von Kennwerten

Ausgangspunkt: Berechnung von Kennwerten aus Prozessgrößen, die Aussagen über den Gesamtzustand des Prozesses ermöglichen

Beispiele:

- Wirkungsgrad
- Ressourcenverbrauch (Energie, Wasser, Luft ...)
- Verschleiß

Ansätze:

Kennwertberechnung $\eta = f[\underline{u}, \underline{y}, \underline{x}]$

Hinweis: meistens statisches Modell

Anwendung:

Kennwertänderung $\Delta\eta = \eta_{IST} - \eta_{SOLL}$

-> gibt Hinweis (meist global) für die Prozessentwicklung und die Fehlerentwicklung

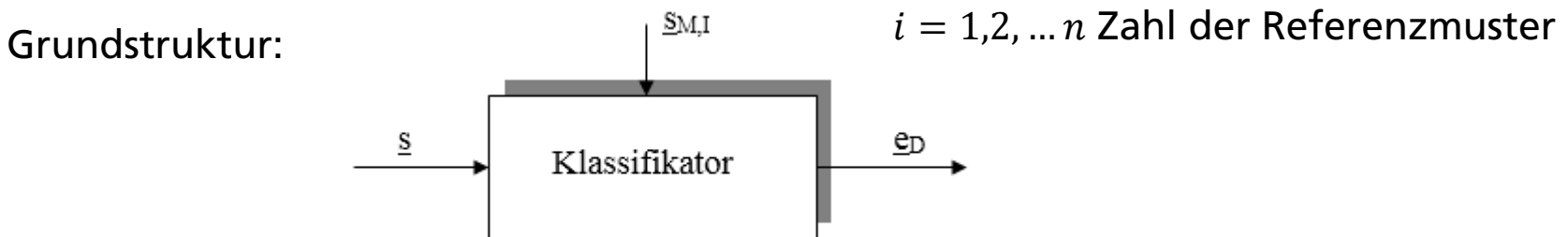
2. Diagnosestrategien

2.2. Diagnosekonzepte

- Grundlegende Konzepte:
- Klassifikationsverfahren
 - Modellreferenzverfahren
 - wissensbasierte Verfahren

2.2.1. Klassifikationsverfahren

- Basis:
1. Referenzsituationen / Muster s_M für jede Fehlerklasse
 2. Aktuelle Fehlersituation \underline{s}
 3. Klassifikator (deterministisch, statistisch, unscharf, Neuronales Netz)



2. Diagnosestrategien

2.2.1. Klassifikationsverfahren

Abstandsmaße: Ähnlichkeit, Sympathie

Konzept	Abstandsmaß	Wert
deterministisch	$d = f(\ \underline{s} - \underline{s}_M\)$	min
statistisch	$d \triangleq \hat{P}(i/\underline{s})$	max
unscharf	$d \triangleq \mu(i/\underline{s})$	max
Künstliche Neuronale Netze	$d \triangleq \mu(i/\underline{s})$	max

Entwurf:

1. Lernphase:

- Entwurf des Klassifikators anhand einer erfolgsbewerteten

Lernprobe LP (\underline{s}, e_D)

- Evaluierung des Klassifikators durch Teilung der LP (50/50)

2. Kannphase: Zuordnung neuer \underline{s} über den Klassifikator zu den Diagnosevorschlägen

2. Diagnosestrategien

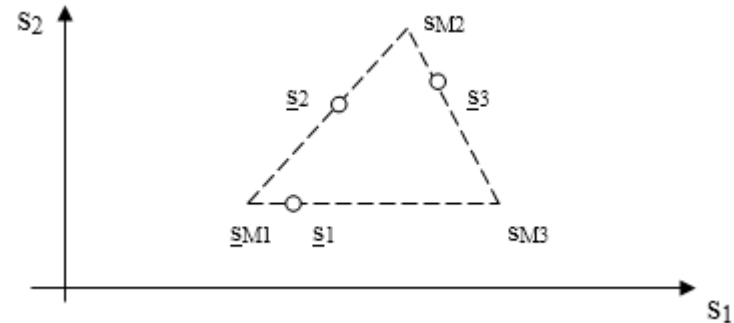
2.2.1. Klassifikationsverfahren

Berechnungen / Darstellungen

I. Deterministisches Klassifikatorkonzept

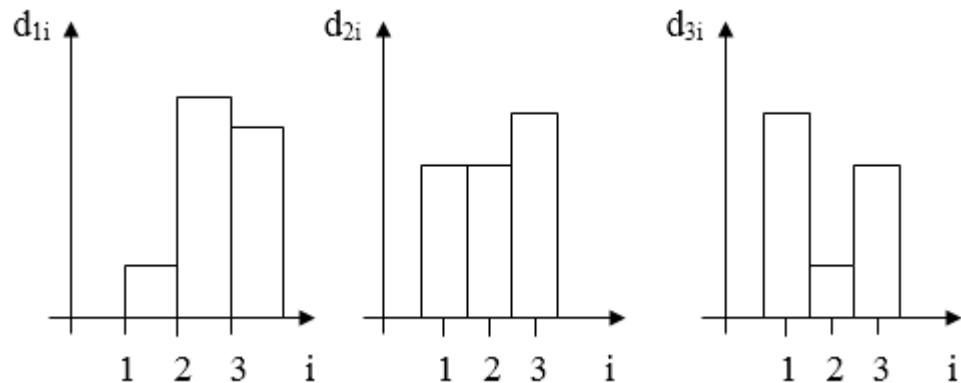
Ähnlichkeitsmaß: $d = f(\|\underline{s} - \underline{s}_{Mi}\|)$

mit $d = \|\underline{s} - \underline{s}_{Mi}\|^2$ d. h. euklidischer Abstand



Darstellungen:

für die Situationen $\underline{s}_1, \underline{s}_2, \underline{s}_3$



2. Diagnosestrategien

2.2.1. Klassifikationsverfahren

Berechnungen / Darstellungen

II. Statistisches Klassifikatorkonzept

Ähnlichkeitsmaß:

$$\rightarrow d = f[P(i/\underline{s})]$$

mit $P(i/s)$ bedingte Wahrscheinlichkeit der Alternative i unter der Bedingung \underline{s}

$$\rightarrow d = f(\underline{\mu}, \underline{N}) \text{ (Abstand zum Erwartungswert)}$$

aus $P(i/s)$ abgeleitete Ähnlichkeitsmaße für mehrdimensionale Normalverteilung, z. B.

$$d_{ij} = \|\underline{s}_j - \underline{s}_{Mi}\|^2 \text{ Euklidischer Abstand}$$

$$d_{ij} = (\underline{s}_j - \underline{s}_{Mi})^T \underline{N}^{-1} (\underline{s}_j - \underline{s}_{Mi}) \text{ Mahalanobis Distanz}$$

$$\rightarrow d = f[P(i/s)]$$

Schätzung der $P(i/s)$ mit Annahme: $P(i/s) = \hat{w}_{0i} + \hat{w}_{1i}s + w_{2i}s^2$ aus Lernprobe und Regressionsverfahren

2. Diagnosestrategien

2.2.1. Klassifikationsverfahren

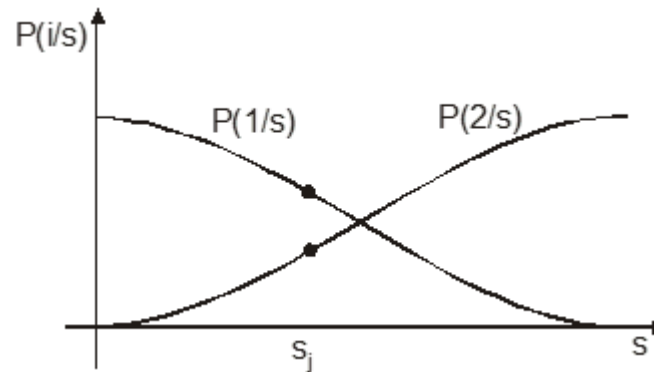
Berechnungen / Darstellungen

$i \rightarrow$ Alternativen
 $j \rightarrow$ Situationsvektor

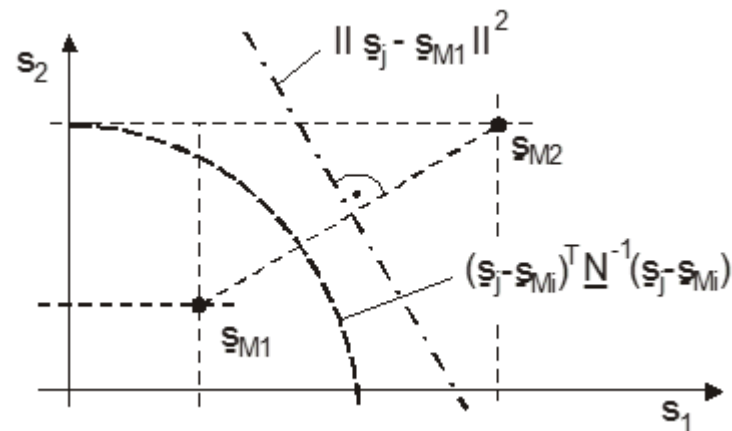
II. Statistisches Klassifikatorkonzept

Realisierungen:

$$\rightarrow d_{ij} = P(i/s_j) \quad \text{mit } i = 2, j = 1$$



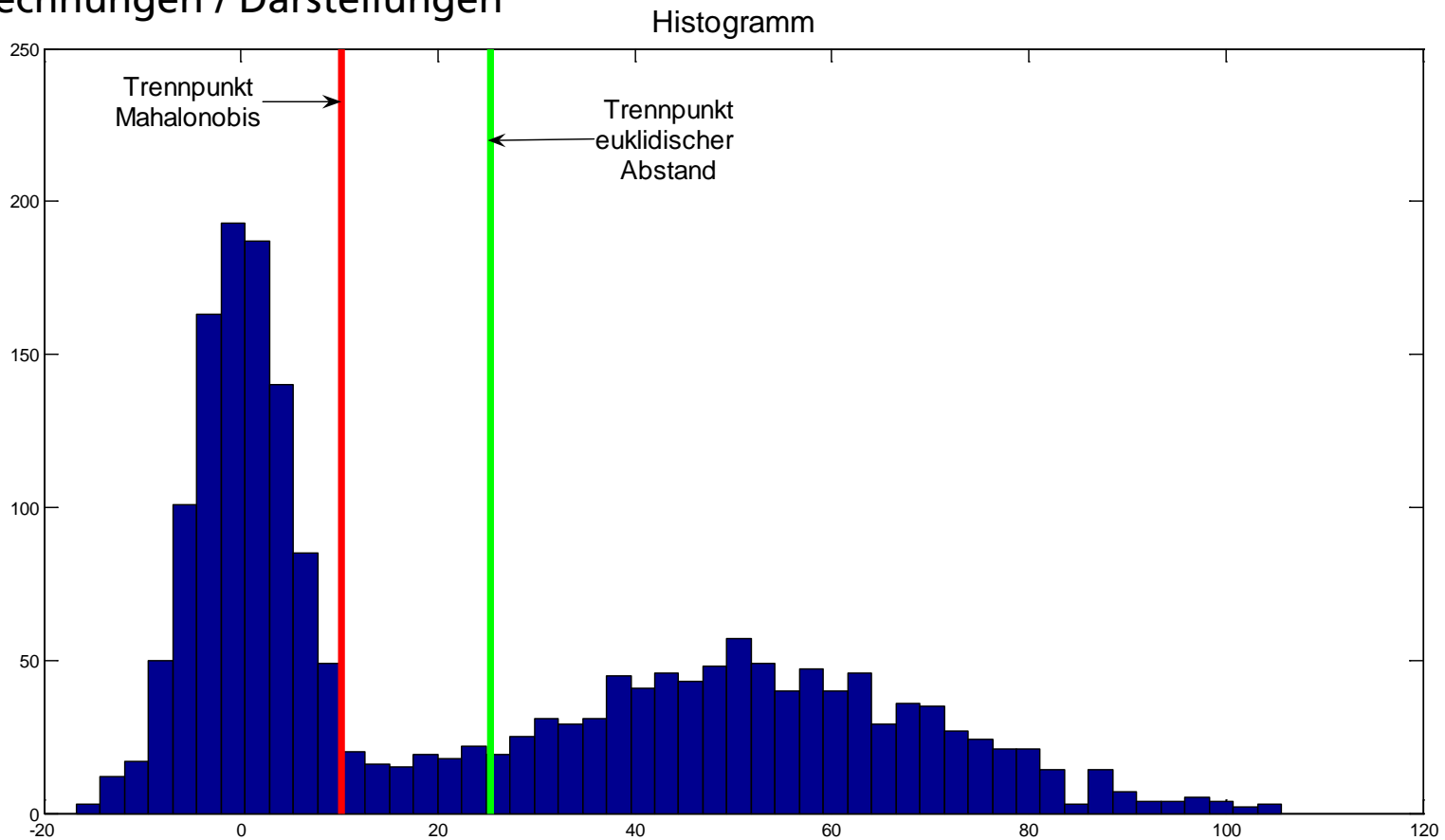
$$\rightarrow d_{ij} = f(\underline{\mu}, \underline{N}_i) \quad \text{mit } i = 2, j = 2$$



2. Diagnosestrategien

2.2.1. Klassifikationsverfahren

Berechnungen / Darstellungen



2. Diagnosestrategien

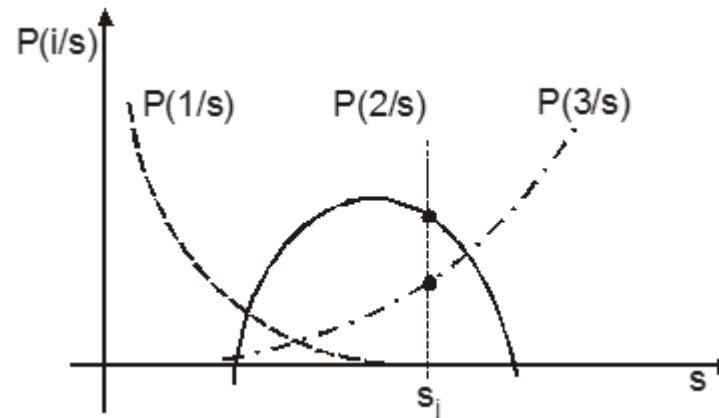
2.2.1. Klassifikationsverfahren

Berechnungen / Darstellungen

II. Statistisches Klassifikatorkonzept

Realisierungen:

$$\rightarrow d_{ij} = \hat{P}(i/\underline{s}) \quad \text{mit } i = 3, j = 1$$



2. Diagnosestrategien

2.2.1. Klassifikationsverfahren

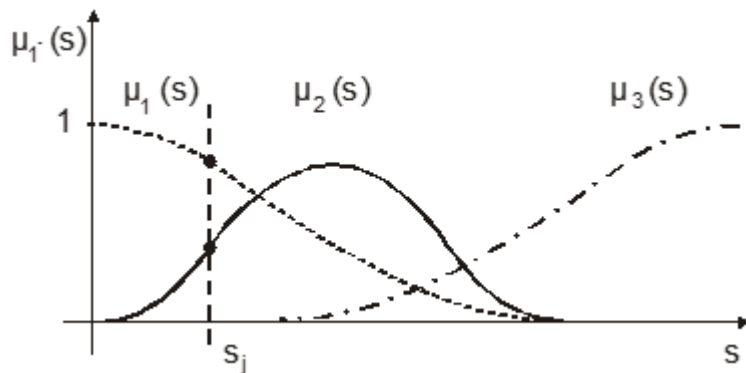
Berechnungen / Darstellungen

III. Unscharfes Klassifikatorkonzept

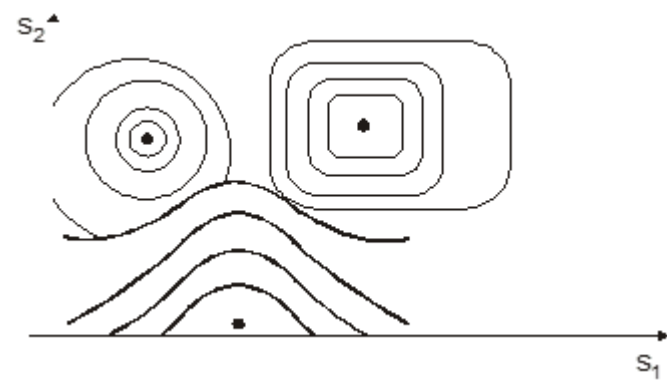
Ähnlichkeitsmaß:

$$d = \hat{f}[\mu_i(\underline{s})] \quad \text{mit } \mu_i(\underline{s}) \begin{array}{l} \text{Zugehörigkeitsfunktion} \\ \text{Wahrheitsfunktion} \\ \text{Sympathiefunktion} \\ \text{der Alternativ } i \text{ zur Situation } \underline{s} \end{array}$$

Realisierung: $i = 3, j = 1$



$i = 3, j = 2$



2. Diagnosestrategien

2.2.1. Klassifikationsverfahren

Berechnungen / Darstellungen

IV. Künstliche Neuronale Netze

Ähnlichkeitsmaß: $d = \hat{f}[\mu_i(\underline{s})]$ mit $\mu_i(\underline{s})$

Zugehörigkeitsfunktion
Wahrheitsfunktion
Sympathiefunktion
der Alternativ i zur Situation \underline{s}

Annahme:

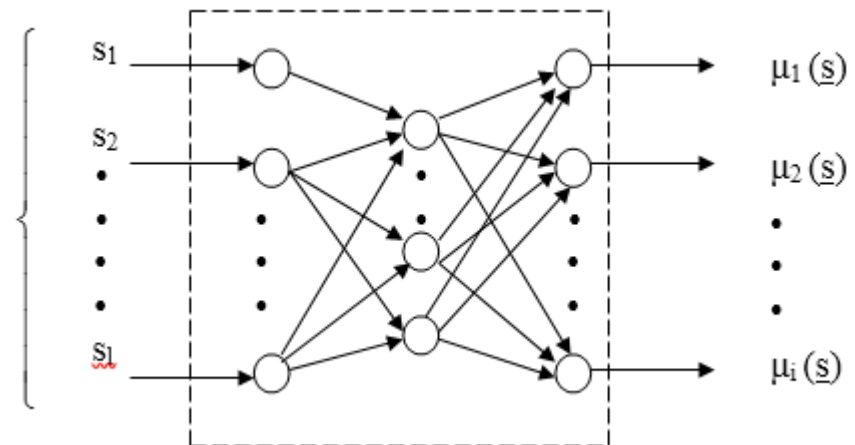
Auswahl Backpropagation Netze

mit: $\rightarrow \bigcirc \rightarrow$ Neuron

s_l Situationskomponenten

$\mu_i(\underline{s})$ Wahrheitsfunktion der i -ten Alternative i

$$0 \leq \mu_i(\underline{s}) \leq 1$$



2. Diagnosestrategien

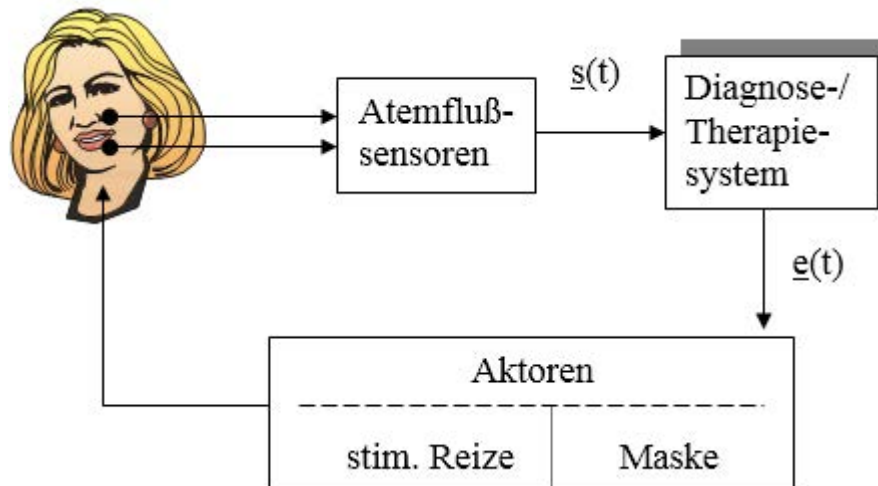
2.2.1. Klassifikationsverfahren

Projekt: Erkennen von Atemstillstand

Aufgabe: Rechtzeitiges Erkennen des sich entwickelnden Atemstillstandes beim Menschen und Ergreifen von Maßnahmen zur Abwehr / Überbrückung des zu erwartenden Sauerstoffmangels im Blut

Lösung: Entwurf und Realisierung eines automatischen Diagnose- und Therapiesystems unter Verwendung eines Klassifikators und einer Klassensteuerung

Schema:

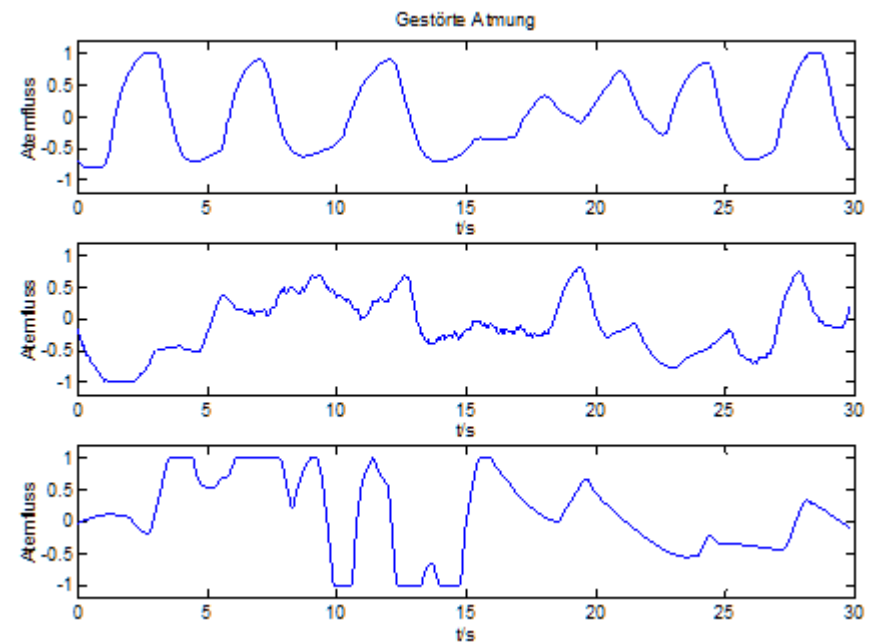
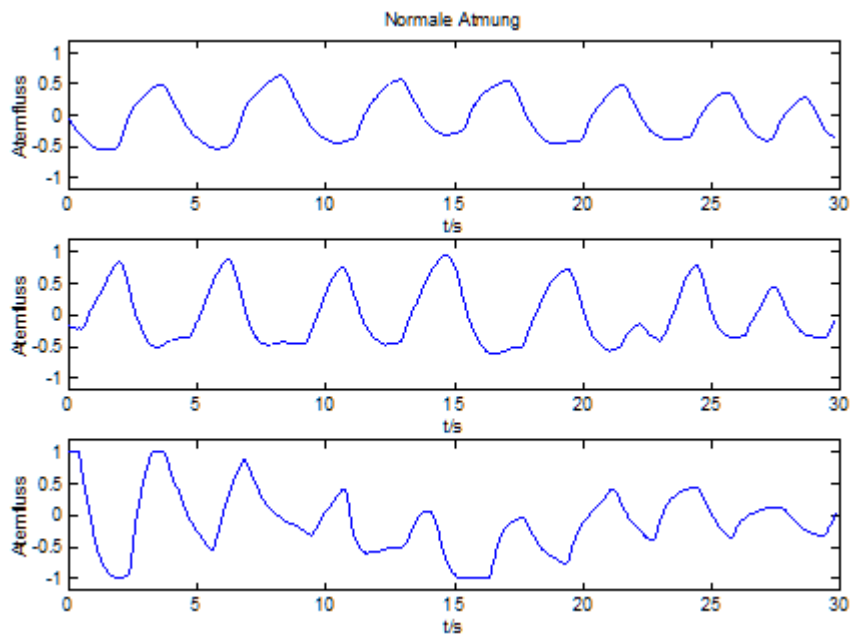


2. Diagnosestrategien

2.2.1. Klassifikationsverfahren

Projekt: Erkennen von Atemstillstand

Situationsverläufe: Atemfluss



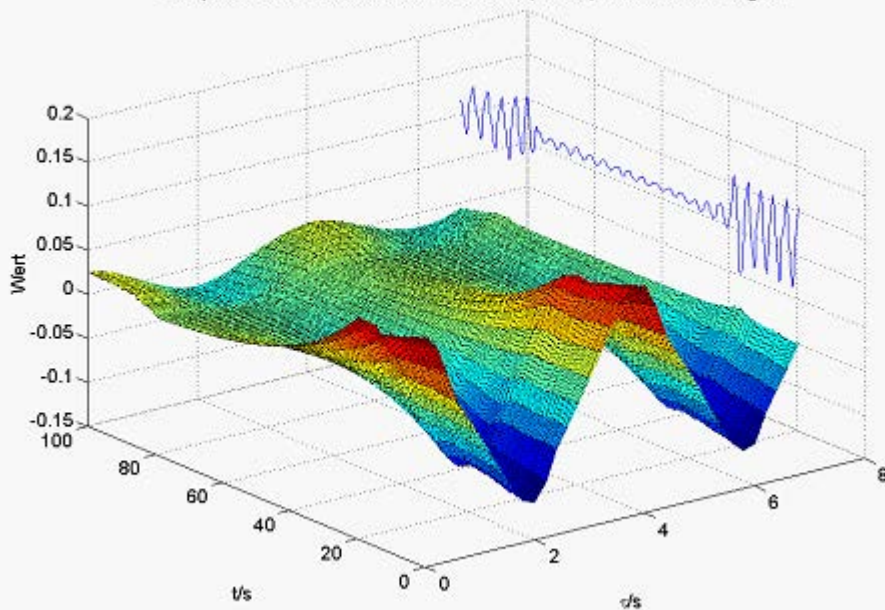
2. Diagnosestrategien

2.2.1. Klassifikationsverfahren

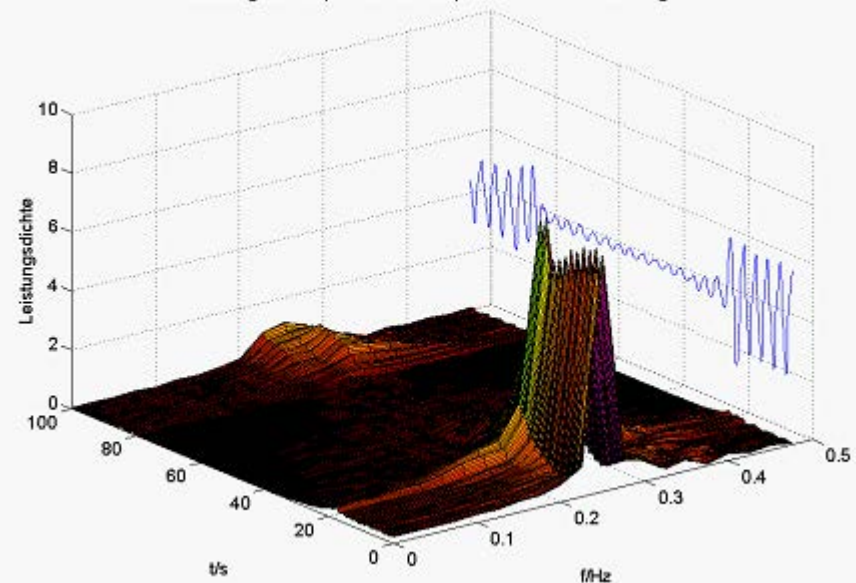
Projekt: Erkennen von Atemstillstand

Merkmalsverläufe:
rekursive Autokorrelationsfunktion
rekursives Leistungsdichtespektrum

Adaptive rekursive Autokorrelationsfunktion ($c=0.01$) und Atemfluss- Signal



Leistungsdichtespektrum aus adapt. AKF und Atemfluss- Signal

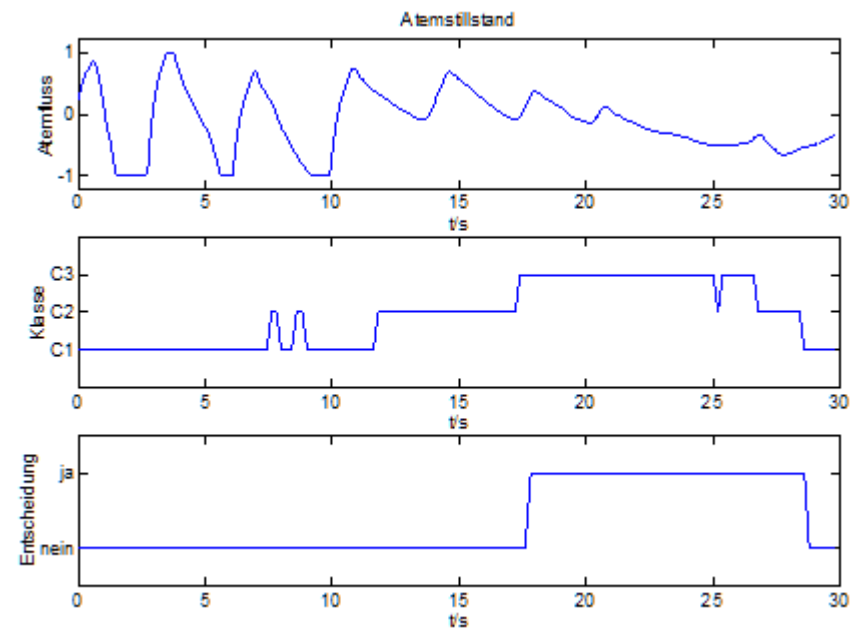
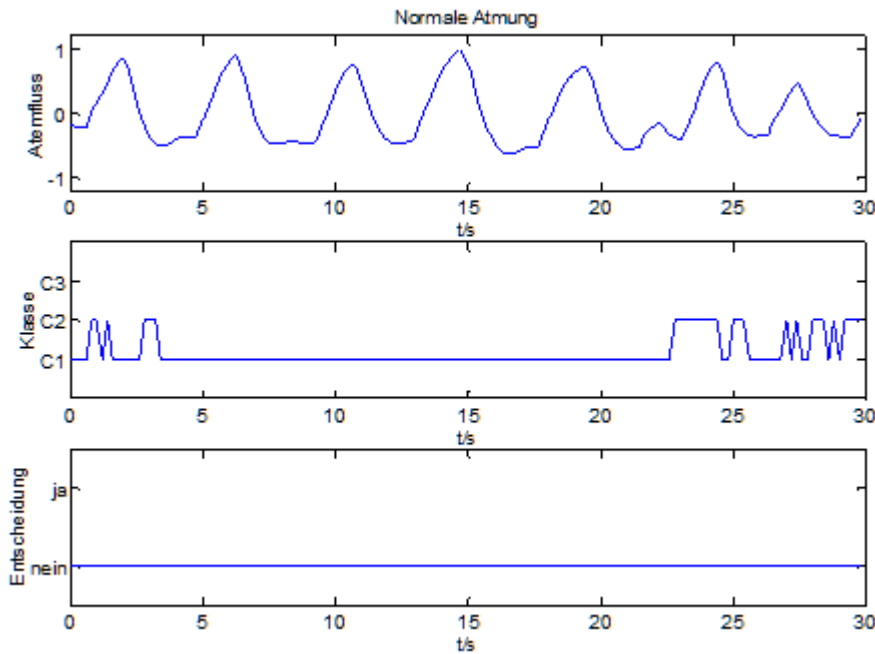


2. Diagnosestrategien

2.2.1. Klassifikationsverfahren

Projekt: Erkennen von Atemstillstand

Entscheidungsvorschläge: Reiz / Maske

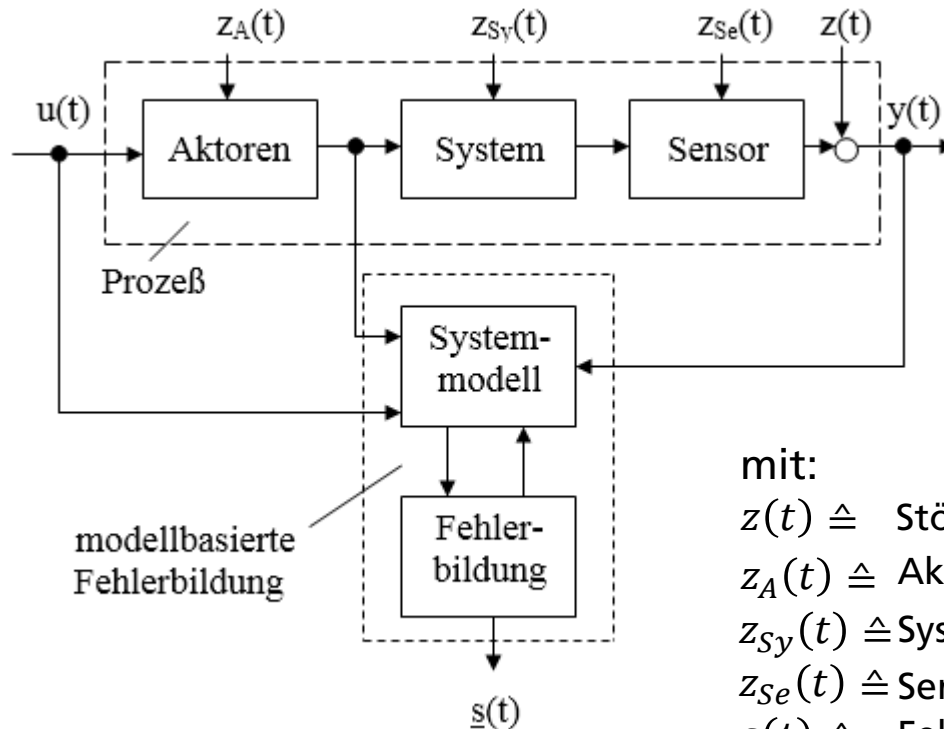


2. Diagnosestrategien

2.2.2. Modellreferenzverfahren

Ausgangsgedanken: Vergleich des Verhaltens des realen Prozesses mit einem Prozessmodell und Bewertung des Fehlers

Grundstruktur:



mit:

$z(t) \triangleq$ Störung auf Prozessausgang

$z_A(t) \triangleq$ Aktorstörung

$z_{Sy}(t) \triangleq$ Systemstörung

$z_{Se}(t) \triangleq$ Sensorstörung

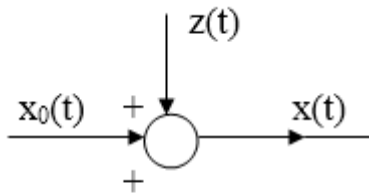
$\underline{s}(t) \triangleq$ Fehlervektor

2. Diagnosestrategien

2.2.2. Modellreferenzverfahren

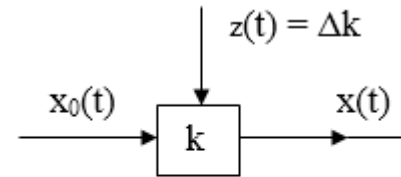
Fehlermodelle:

α) additive Fehler $z(t)$



$$x(t) = x_0(t) + z(t)$$

β) multiplikativer Fehler $Z(t)$



$$x(t) = (k + \Delta k) x_0(t)$$

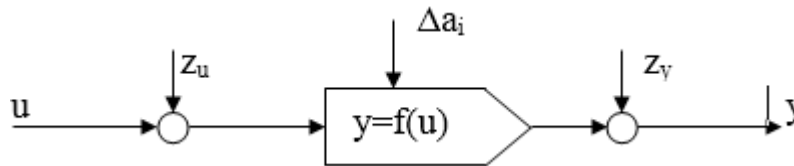
2. Diagnosestrategien

2.2.2. Modellreferenzverfahren

Systemfehlermodelle

I. Statisches nichtlineares Systemfehlermodell

Struktur:



Grundgleichungen:

für $z_u = \Delta a_i = z_y = 0$

$$y = a_0 + \sum_{i=1}^n a_i u^i$$

für $z_u, z_y, \Delta a_i$

$$y = a_0^* + \sum a_i^* \tilde{u}^i + z_y$$

mit $a_i^* = (a_i + \Delta a_i),$
 $\tilde{u} = u + z_u$

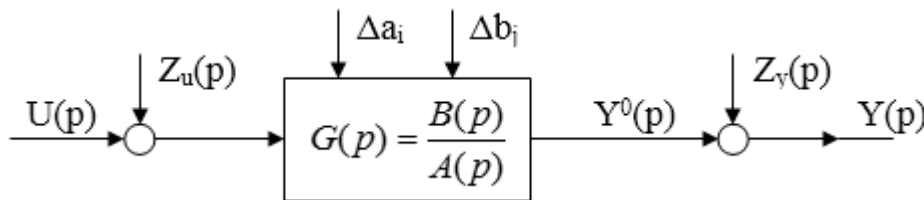
2. Diagnosestrategien

2.2.2. Modellreferenzverfahren

Systemfehlermodelle

II. Parametrisches Ein- / Ausgangsmodell

Struktur:



mit

$$B(p) \triangleq [b_0 + b_1 p + \dots + b_m p^m]$$
$$A(p) \triangleq [a_0 + a_1 p + \dots + a_n p^n]$$

Grundgleichungen:

$$\text{für } z_u(t) = z_y(t) = \Delta a_i = \Delta b_j = 0$$

$$a_0 y(t) + a_1 y^{(1)}(t) + \dots + a_n y^{(n)}(t) = b_0 u(t) + b_1 u^{(1)}(t) + \dots + b_m u^{(m)}(t)$$

für $z_u(t), z_y(t), \Delta a_i, \Delta b_j$

$$a_0^* y(t) + a_1^* y^{(1)}(t) + \dots + a_n^* y^{(n)}(t) = b_0^* \tilde{u}(t) + b_1^* \tilde{u}^{(1)}(t) + \dots + b_m \tilde{u}^{(m)}(t)$$

$$\text{mit } a_i^* = a_i + \Delta a_i; b_j^* = b_j + \Delta b_j$$

$$\tilde{u}(t) = u(t) + z_u(t)$$

$$y(t) = y^0(t) + z_y(t)$$

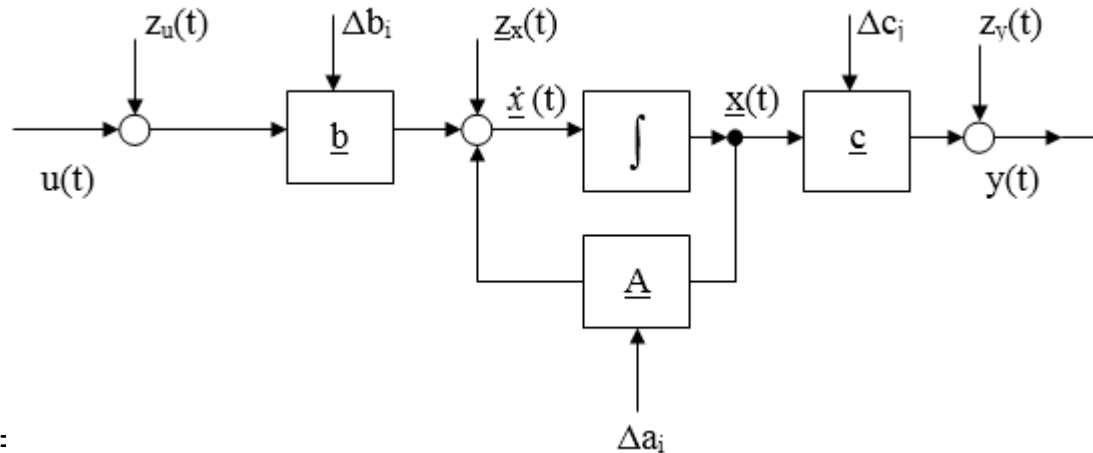
2. Diagnosestrategien

2.2.2. Modellreferenzverfahren

Systemfehlermodelle

III. Zustandsmodell

Struktur:



Grundgleichungen:

für $z_u(t) = z_x(t) = z_y(t) :$

$$\Delta a_i = \Delta b_j = \Delta c_j = 0$$

$$\dot{\underline{x}}(t) = \underline{A} \cdot \underline{x}(t) + \underline{b} \cdot u(t)$$

$$y(t) = \underline{c}^T \cdot \underline{x}(t)$$

für $z_u(t), z_x(t), z_y(t) \neq 0$

$$\Delta a_i, \Delta b_j, \Delta c_j \neq 0$$

$$\dot{\underline{x}}(t) = \underline{A}^* \cdot \underline{x}(t) + \underline{b}^* \cdot u^*(t)$$

$$y(t) = \underline{c}^* \cdot \underline{x}(t) + z_y(t)$$

$$\text{mit } \underline{A}^* = \underline{A} + \Delta \underline{A}; \underline{b}^* = \underline{b} + \Delta \underline{b}; \underline{c}^* = \underline{c} + \Delta \underline{c}$$

$$u^*(t) = u(t) + z_u(t)$$

2. Diagnosestrategien

2.2.2. Modellreferenzverfahren

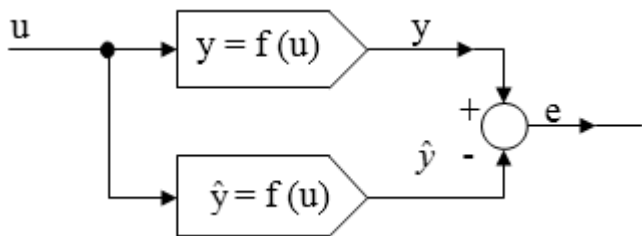
Fehlerbildungsstrategien

I. Anwendung von Paritätsgleichungen

- Voraussetzungen:
- (1) Prozessparameter bekannt
 - (2) Prozessverhalten zeitinvariant
 - (3) Geringe Störungen

Modellstrukturen:

(1) Ein- / Ausgangsmodell - Statik



Paritätsgleichungen

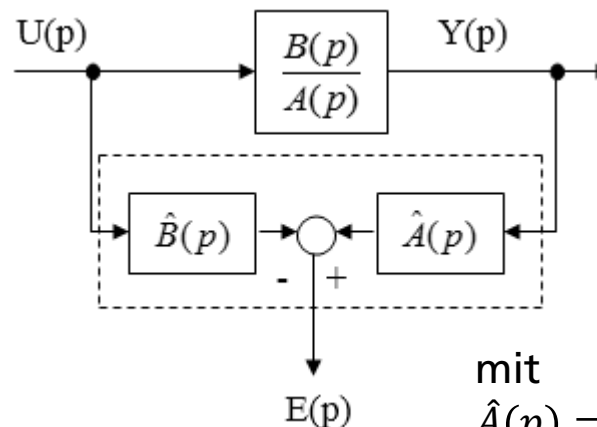
$$e = y - \hat{y}$$

$$e = y - \underline{m}^T \cdot \underline{\hat{s}} \quad \text{mit}$$

$$\underline{m}^T = [1 \ u \ u^2 \ \dots]$$

$$\underline{\hat{s}}^T = [\hat{a}_0 \ \hat{a}_1 \ \hat{a}_2]$$

(2) Ein- / Ausgangsmodell - Dynamik



Paritätsgleichungen

$$E(p) = \hat{A}(p) \cdot Y(p) - \hat{B}(p) \cdot U(p)$$

$$\text{mit}$$

$$\hat{A}(p) = [\hat{a}_0 + \hat{a}_1 p + \dots + \hat{a}_m p^m]$$

$$\hat{B}(p) = [\hat{b}_0 + \hat{b}_1 p + \dots + \hat{b}_n p^n]$$

2. Diagnosestrategien

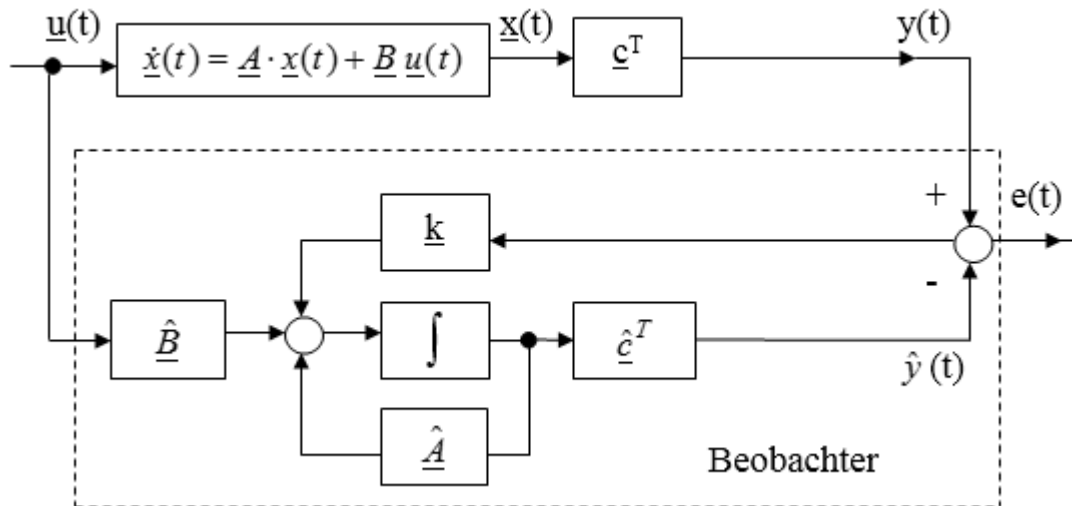
2.2.2. Modellreferenzverfahren

Fehlerbildungsstragien

I. Anwendung von Paritätsgleichungen

Modellstrukturen:

(3) Zustandsbeobachter



Paritätsgleichungen

$$e(t) = y(t) - \underline{\hat{c}}^T \cdot \underline{\hat{x}}(t)$$

mit

$$\dot{\hat{\underline{x}}}(t) = \hat{\underline{A}} \cdot \hat{\underline{x}}(t) + \hat{\underline{B}} \cdot u(t) + \underline{k} \cdot e(t)$$

$$\text{Symptome: } \Delta \underline{x}(t) = \hat{\underline{x}}(t) - \underline{x}(t)$$

$$e(t)$$

Anwendungen

Erkennen additiver

Fehler, z. B. Lecks,

Kurzschlüsse

2. Diagnosestrategien

2.2.2. Modellreferenzverfahren

Fehlerbildungsstragien

II. Anwendung von Parameterschätzverfahren

Voraussetzungen: (1) Prozessparameter unbekannt
(2) Prozessverhalten zeitvariant
(3) Störungen

α) Struktur für Ausgangsfehler

Annahmen: Modellstruktur $\hat{y} = \underline{m}^T \cdot \hat{\underline{s}}$

Gütekriterium $Q_1 = \underline{e}^T \underline{e}$

$$Q_2 = \frac{1}{2} \underline{e}_i^2$$

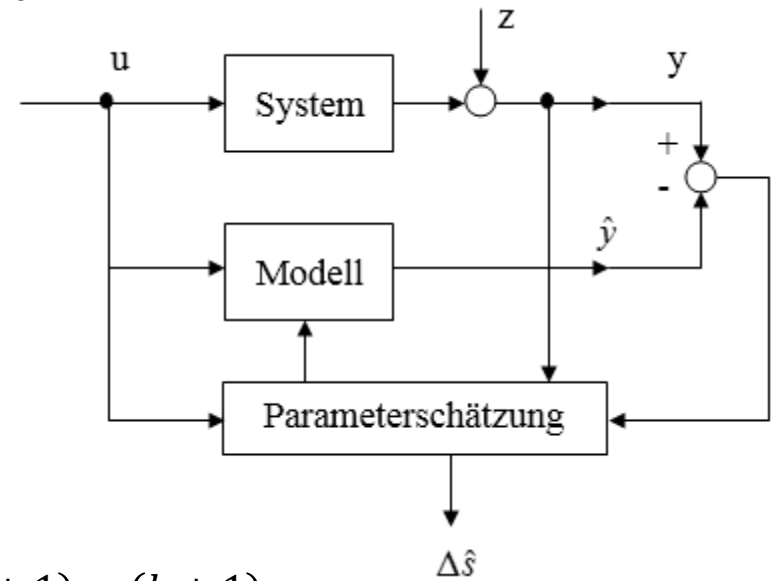
$$\underline{e} = \underline{y} - \hat{\underline{y}}$$

Merkmale $\Delta \hat{\underline{s}}$

Schätzverfahren: direkt: $\hat{\underline{s}} = [\underline{M}^T \underline{M}]^{-1} \underline{M}^T \cdot \underline{y}$

rekursiv: $\hat{\underline{s}}(k+1) = \hat{\underline{s}}(k) + \underline{y}(k+1) \cdot \underline{e}(k+1)$

iterativ: $\hat{\underline{s}}^{i+1} = \hat{\underline{s}}^i + \hat{\underline{y}}^{i+1} \cdot [\nabla \nabla Q]^{-1}$



2. Diagnosestrategien

2.2.2. Modellreferenzverfahren

Fehlerbildungsstrategien

II. Anwendung von Parameterschätzverfahren

β) Struktur für Gleichungsfehler

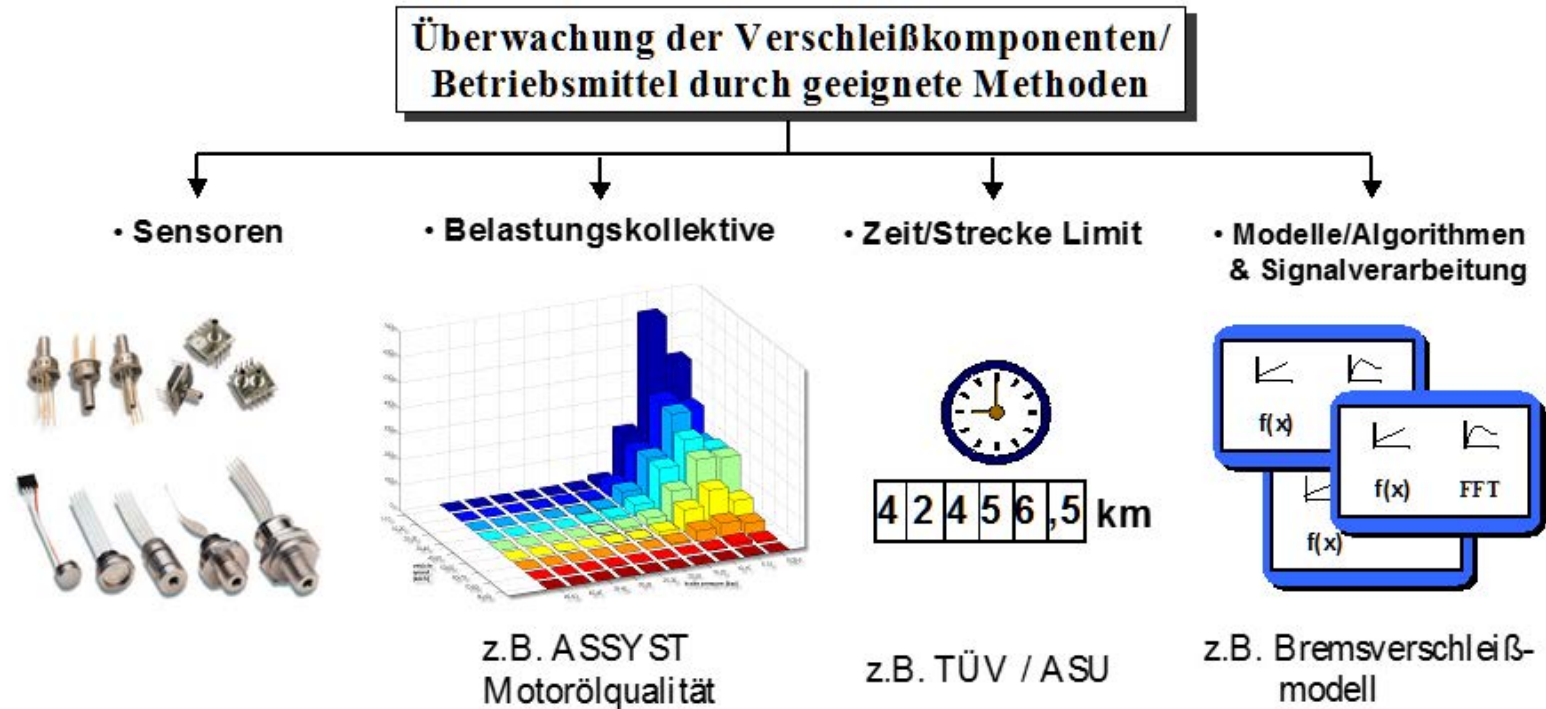
Annahmen: Modellstruktur $\hat{y} = \underline{m}^T \cdot \hat{\underline{s}}$
 $\underline{m}^T = [y(k-1), \dots, y(k-n) | u(k-m)]$
 $\hat{\underline{s}}^T = [-\hat{a}_1, \dots, -\hat{a}_n | \hat{b}_0 \dots \hat{b}_m]$
Gleichungsfehler $e = A(q^{-1}) \cdot y(k) - B(q^{-1}) \cdot U(k)$
Gütekriterien $Q_1 = \underline{e}^T \underline{e}$ und $Q_2 = \frac{1}{2} e_i^2$

Schätzverfahren: direkt $\hat{\underline{s}} = [\underline{H}^T \underline{M}]^{-1} \underline{H}^T \cdot \underline{y}$
 $\underline{h}^T \triangleq$ Hilfsvariablenvektor
 $\underline{h}^T = [h(k-1) \dots h(k-n) | u(k) \dots u(k-m)]$
 rekursiv $\hat{\underline{s}}(k+1) = \hat{\underline{s}}(k) + \underline{y}(k+1)e(k+1)$
 mit $e(k+1) = \hat{A}(q^{-1}) \cdot h(k) - \hat{B}(q^{-1}) \cdot u(k)$
 iterativ $\hat{\underline{s}}^{i+1} = \hat{\underline{s}}^i + y^{i+1} \cdot [\nabla \nabla Q]^{-1}$

2. Diagnosestrategien

2.2.2. Modellreferenzverfahren

Zustandsabhängige Wartung



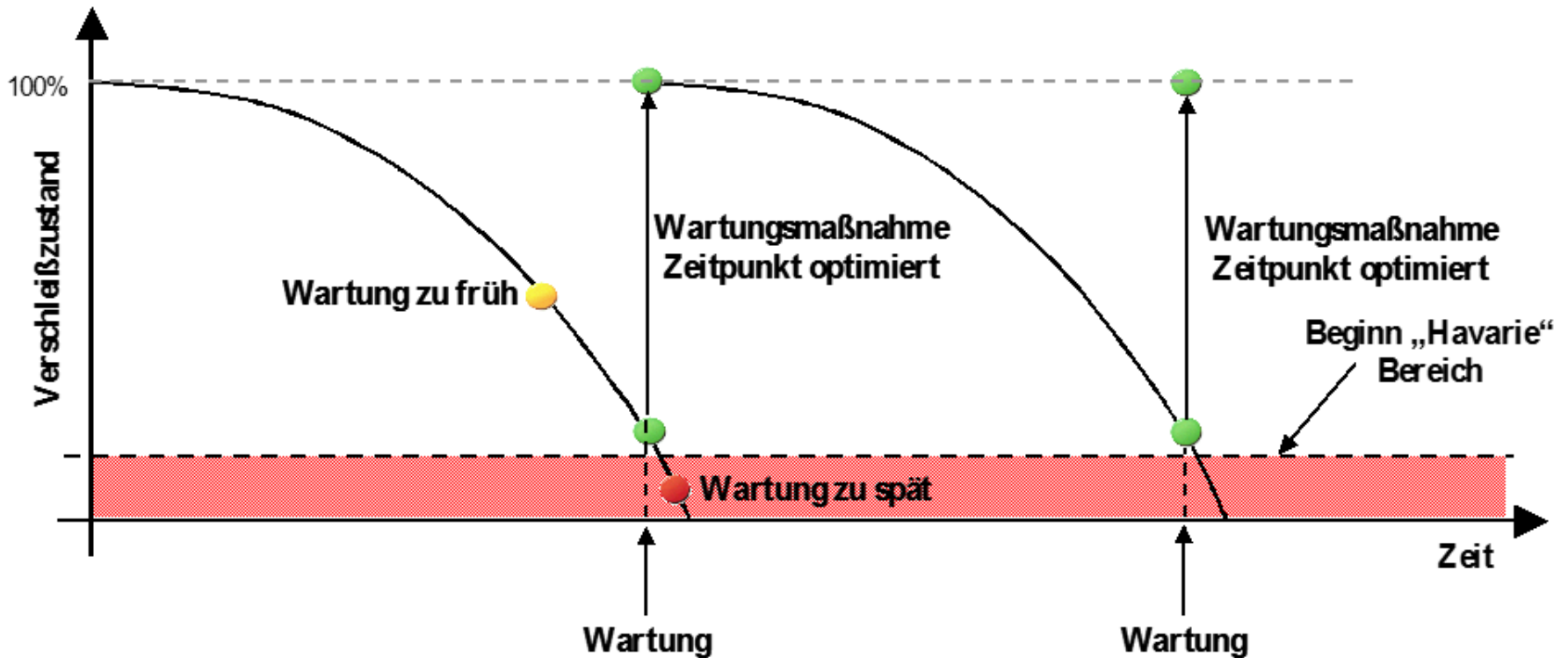
$$\text{Zustand_WP_i} = f(e, a, \dot{x}, s, t) \quad i = 1 \dots n \text{ Wartungspositionen}$$

(DaimlerChrysler)

2. Diagnosestrategien

2.2.2. Modellreferenzverfahren

Zustandsabhängige Wartung



(DaimlerChrysler)

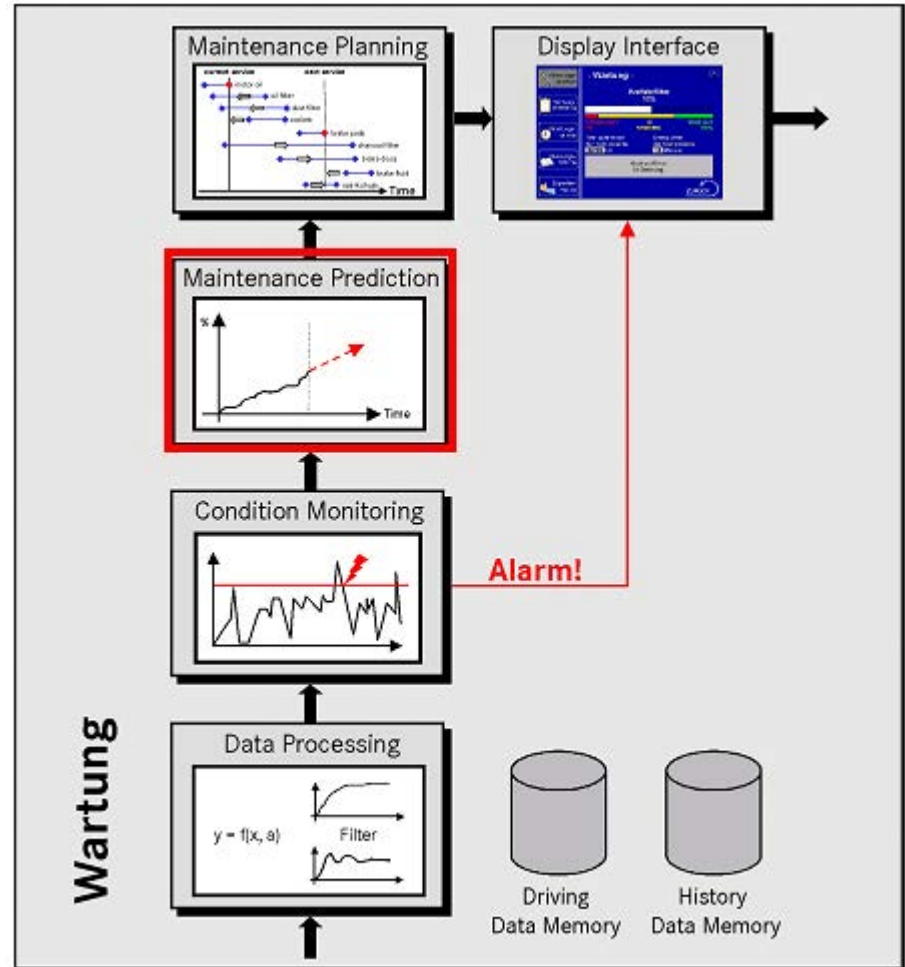
2. Diagnosestrategien

2.2.2. Modellreferenzverfahren

Ablauf der Wartungsfunktion

Maintenance Prediction

- Prognose der Verschleißentwicklung einzelner Komponenten auf Grundlage der Belastung in der Vergangenheit

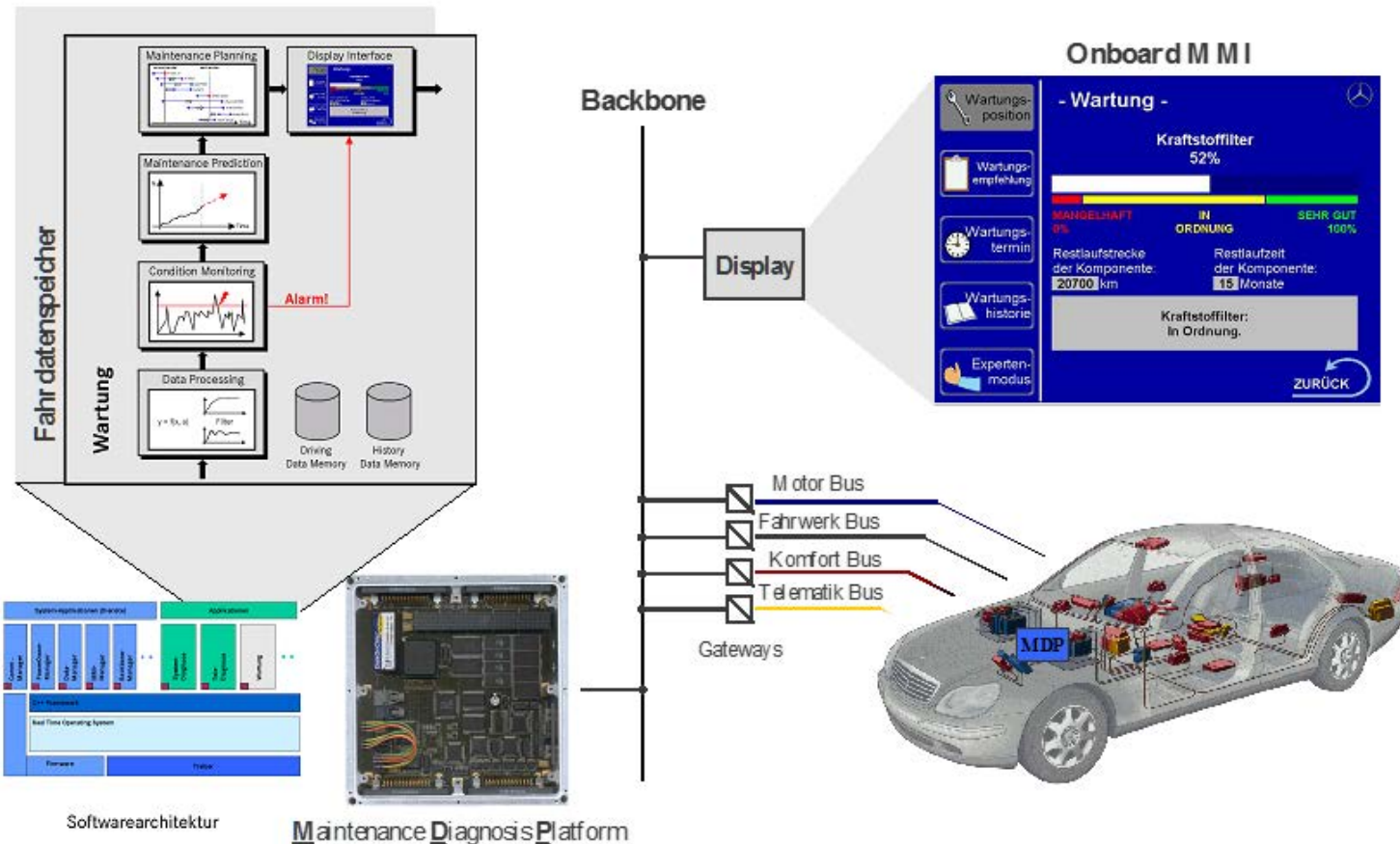


(DaimlerChrysler)

2. Diagnosestrategien

2.2.2. Modellreferenzverfahren

Integration in das Fahrzeug



(DaimlerChrysler)

Softwarearchitektur

Maintenance Diagnosis Platform

2. Diagnosestrategien

2.2.2. Modellreferenzverfahren

Wartungsanzeige Onboard MMI

Lister aller Wartungspositionen

- Wartung -

Zustand	Bezeichnung
●	Bremsflüssigkeit
●	Kühlmittel
●	Motoröl
●	Zündkerzen
●	Ölfilter
●	Keilrippenriemen
●	Kraftstofffilter
●	Luftfilter
●	Wischergummi FS
●	Wischergummi HS
●	Bremsbeläge HA

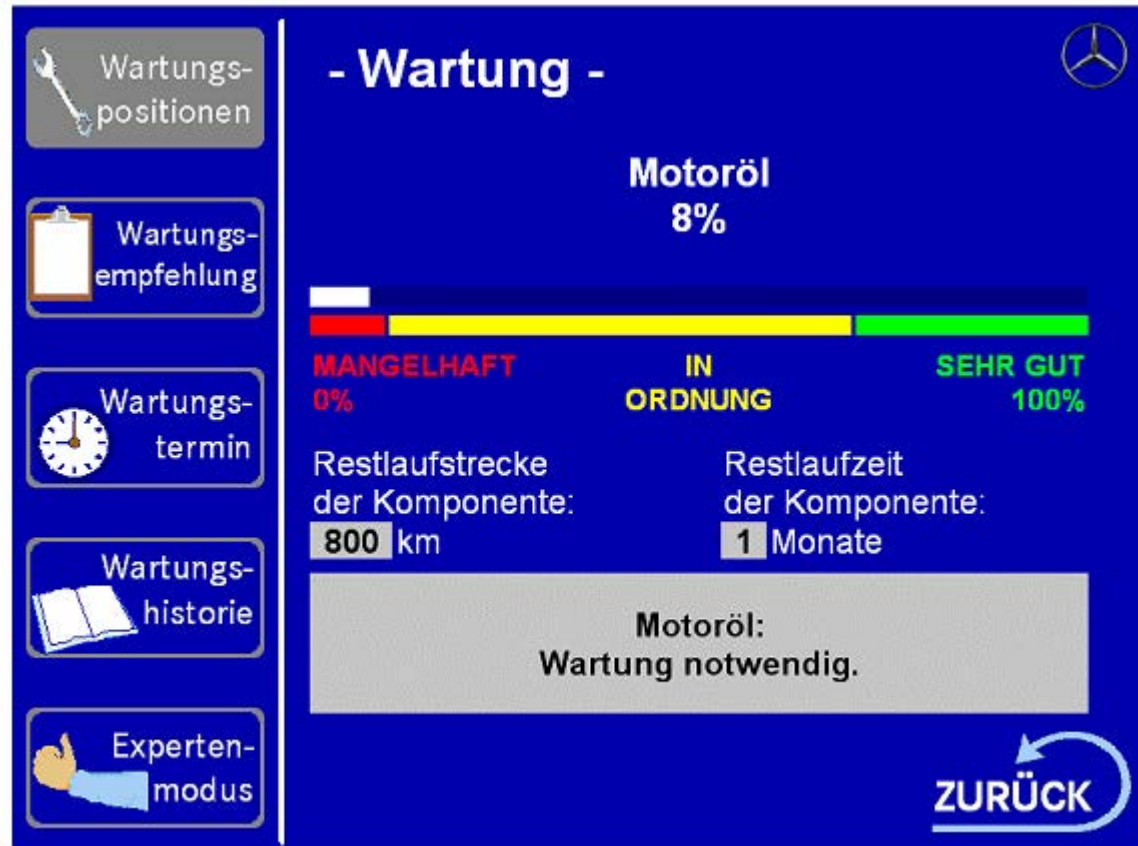
(DaimlerChrysler)

2. Diagnosestrategien

2.2.2. Modellreferenzverfahren

Wartungsanzeige Onboard MMI

Zustand einer Wartungspositionen



(DaimlerChrysler)

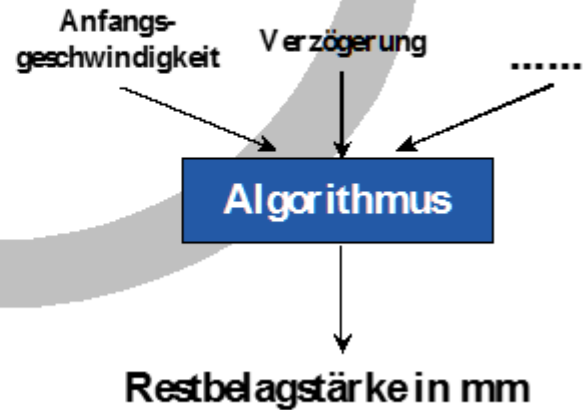
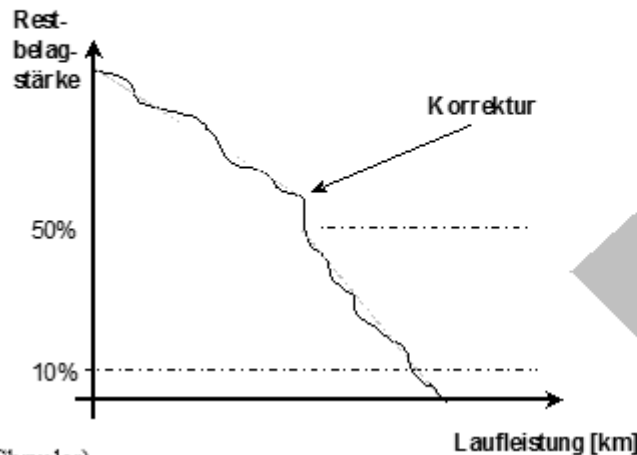
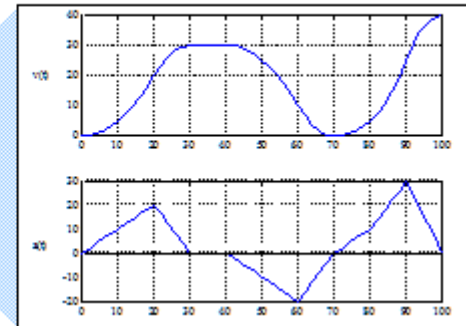
2. Diagnosestrategien

2.2.2. Modellreferenzverfahren

Ziel



kont. Fahrdaten



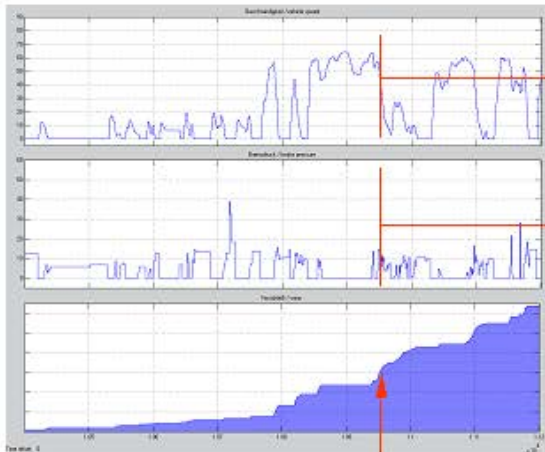
(DaimlerChrysler)

2. Diagnosestrategien

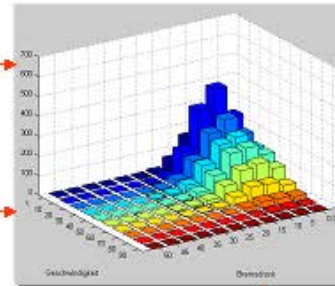
2.2.2. Modellreferenzverfahren

Lastkollektiv

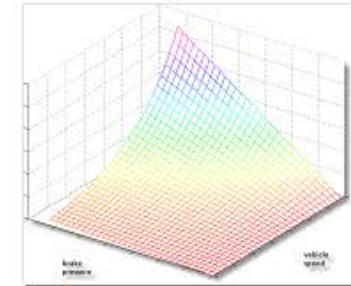
1 Sensorabtastung



2 Einordnung der Eingangsdaten in das Lastkollektiv



Wichtungskennfeld



4 Aussage über den Verschleiß

$$B = \frac{1}{c_T} \sum_i \sum_j (n_{ij} * w_{ij})$$

3 Wichtung und Aufsummierung

- B: vorhergesagter Verschleiß
- n_{ij} : aufsummierte Zustandsklasse (i,j)
- w_{ij} : Wichtungsfaktor für die Zustandsklasse (i,j)
- c_T : Verschleißfaktor

(DaimlerChrysler)

2. Diagnosestrategien

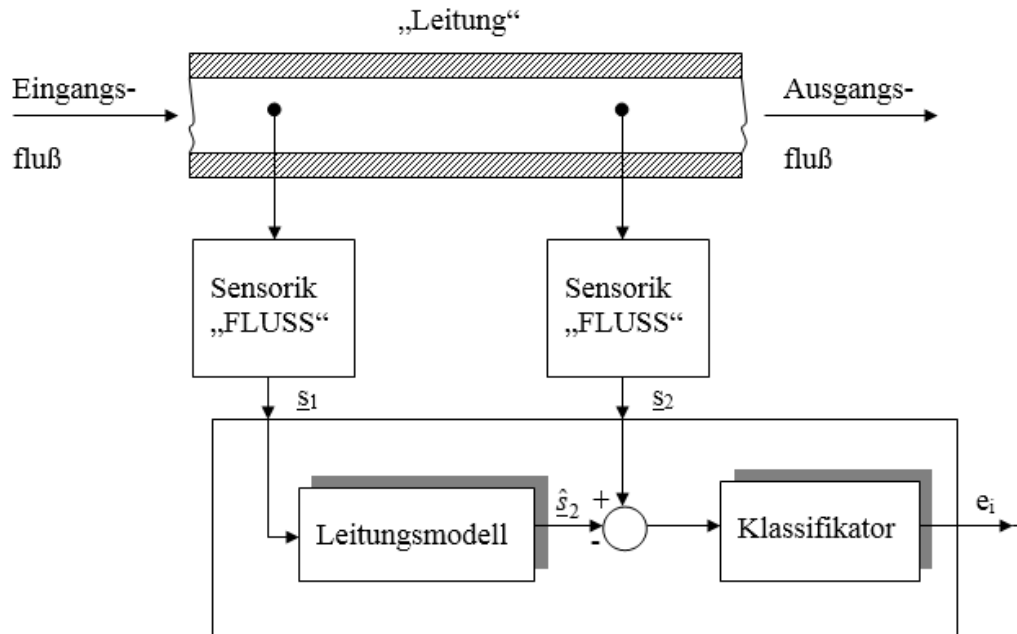
2.2.2. Modellreferenzverfahren

Projekt: Erkennen von Durchflussänderungen (Lecks, Staus u. ä.)

Aufgabe: rechtzeitiges Erkennen von Veränderungen von „Flüssen“ in Leitungen, Gerinnen, Adern u. ä.

Lösung: Entwurf und Realisierung eines automatischen Diagnose- und Therapie-Systems unter Verwendung von Modellreferenzverfahren

Schema:



2. Diagnosestrategien

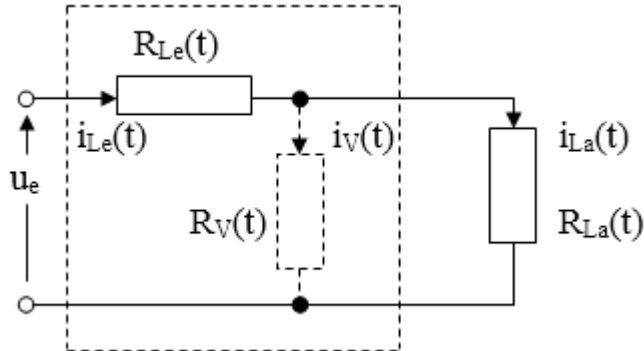
2.2.2. Modellreferenzverfahren

Projekt: Erkennen von Durchflussänderungen (Lecks, Staus u. ä.)

Beispiel: Leckerkennung

allgemeines elektrisches Ersatzschaltbild

$$\dot{m}(t) \triangleq i(t)$$



mit:

$R_{Le}(t) \triangleq$ Leitungswiderstand

$R_V(t) \triangleq$ Leckwiderstand

$R_{La}(t) \triangleq$ Lastwiderstand

$\dot{m}(t) \triangleq$ „Fluss“

Fälle: Normalfall $R_{Le}(t) \ll R_V(t); R_{La}(t) > R_{Le}(t)$

Beispiel $R_{Le}(t) = 1; R_V(t) = 1 \cdot 10^6; R_{La}(t) = 1 \cdot 10^1; u_e = 1$

$$i_{La}(t) = i_{Le}(t) - i_V(t) \quad \rightarrow \quad i_{La}^*(t) = i_{Le}(t); \text{ *Näherung}$$

Fehler: $\Delta i(t) = i_{La}^*(t) - i_{La}(t)$

Klassifikator: $\Delta i(t) \geq \varepsilon$ $\varepsilon \triangleq$ Schranke

2. Diagnosestrategien

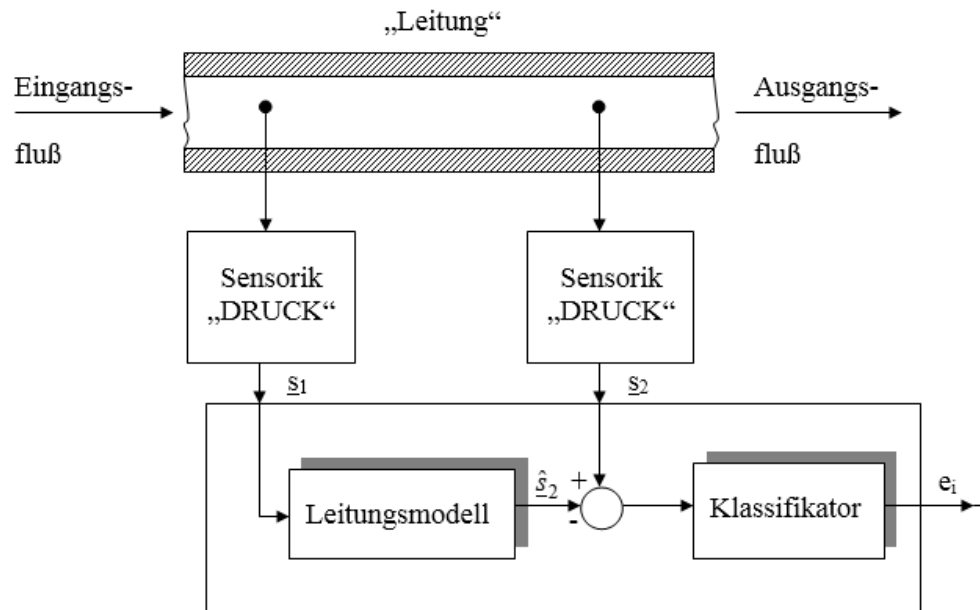
2.2.2. Modellreferenzverfahren

Projekt: Erkennen von Druckänderungen (Lecks u. ä.)

Aufgabe: rechtzeitiges Erkennen von Veränderungen von „Drücken“ in Leitungen, Gerinnen, Adern u. ä.

Lösung: Entwurf und Realisierung eines automatischen Diagnose- und Therapie-Systems unter Verwendung von Modellreferenzverfahren

Schema:



2. Diagnosestrategien

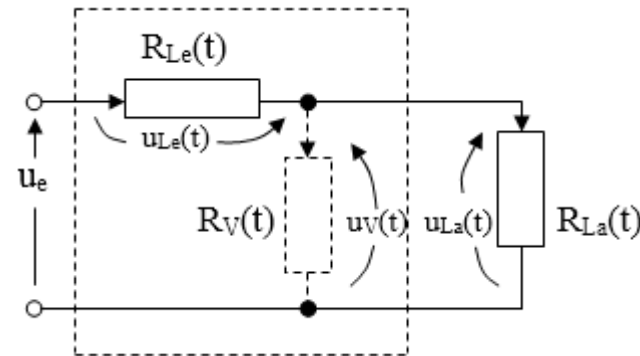
2.2.2. Modellreferenzverfahren

Projekt: Erkennen von Druckänderungen (Lecks u. ä.)

Beispiel: Leckerkennung

allgemeines elektrisches Ersatzschaltbild

$$p(t) \triangleq u(t)$$



Fälle: Normalfall $R_{Le}(t) \ll R_V(t); R_{La}(t) > R_{Le}(t)$

Beispiel $R_{Le}(t) = 1; R_V(t) = 1 \times 10^6; R_{La}(t) = 1 \times 10^1; u_e = 1$

$$u_{La}(t) = \frac{R_V || R_{La}}{R_{Le} + R_V || R_{La}} \cdot u_e(t)$$

$$\rightarrow u_{La}^*(t) = \frac{R_{La}}{R_{Le} + R_{La}} \cdot u_e(t); \text{*Näherung}$$

Fehler: $\Delta u(t) = u_{La}^*(t) - u_{La}(t)$

Klassifikator: $\Delta u(t) \geq \varepsilon$ $\varepsilon \triangleq$ Schranke

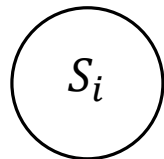
2. Diagnosestrategien

2.2.3. Wissensbasierte Verfahren

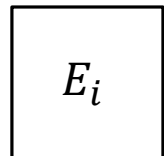
I. Gesamtkonzept

Grundlage: A-priori Wissen über Fehler-Ursache-Wirkungen und deren Abbildung in Kausale Netze / Bäume

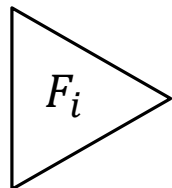
Darstellung: *Symptome \rightarrow Ereignisse \rightarrow Fehler



Symptome: quantitative und qualitative Größen und deren Änderungen (Situationen / Merkmale)

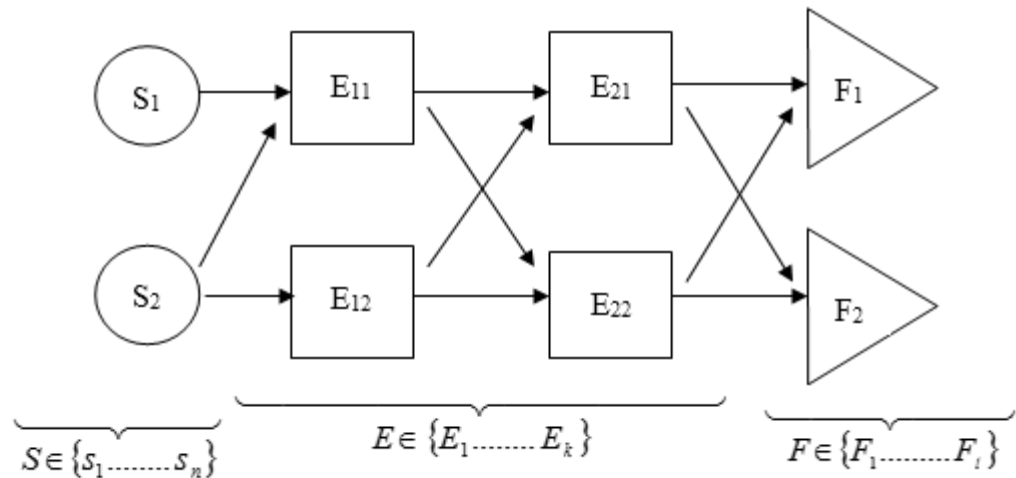


Ereignisse: Symptome mit Eingangs- und Ausgangsgrößen



Fehler: Ergebnis

z. B. 3-Ebenen Kausales Netz



2. Diagnosestrategien

2.2.3. Wissensbasierte Verfahren

I. Gesamtkonzept  Netz- / Baumdarstellung:

Knoten: Symptome S_i
Ereignisse E_i
Fehler F_i

Kanten: Darstellung der
Wirkungsketten (gerichtet)

Diagnoseaufgabe: ausgehend von aktuellen Symptomen wird schrittweise auf die zugrundeliegenden Ursachen / Fehler geschlossen



Verknüpfung von Diagnoseregeln
Wenn < Bedingung >
Dann < Schlussfolgerung >

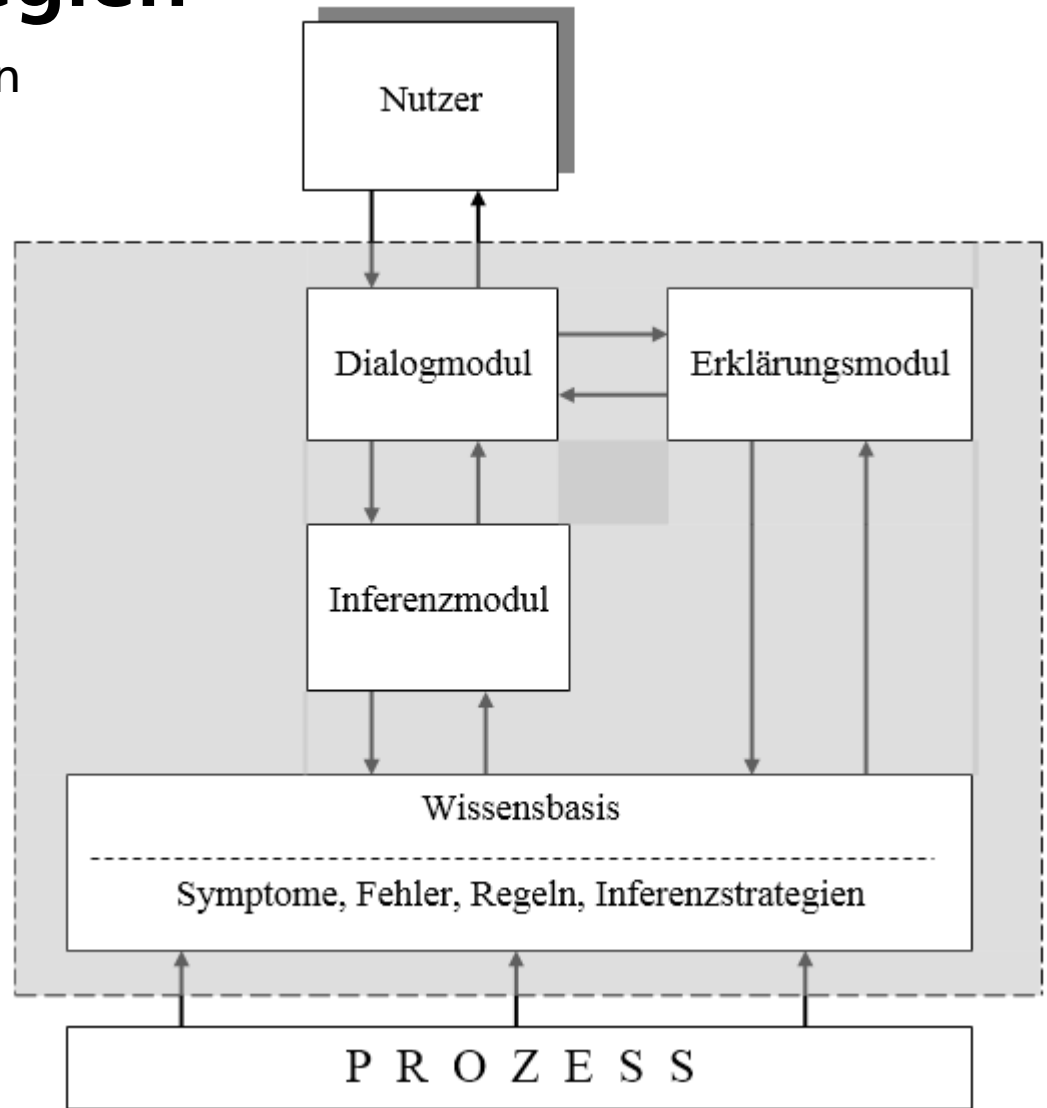
Arten: Boole'sche Logik
Bays Konzept
Fuzzy Konzept

2. Diagnosestrategien

2.2.3. Wissensbasierte Verfahren

I. Gesamtkonzept

Konzept: Beratungsstruktur



2. Diagnosestrategien

2.2.3. Wissensbasierte Verfahren

II. Inferenzverfahren

- Ausgangspunkt: Fehlerbaumanalyse (DIN 25 424)
 - Ereignisbaumanalyse (DIN 25 419)
- für Wenn-Dann-Regeln der Form:
WENN < Bedingung >
DANN < Schlussfolgerung >

Fehlerbäume auf Basis der Boole'schen Logik

Grundlagen:

Fehlerbaum mit binären Variablen

Auswertung der Regeln mit der Boole'schen Algebra



Binäre Schließen

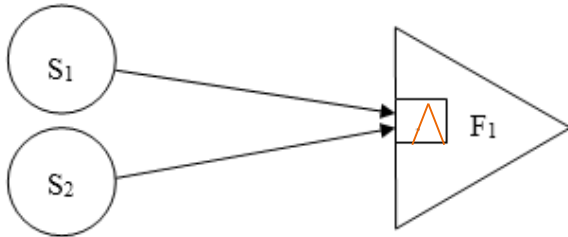
2. Diagnosestrategien

2.2.3. Wissensbasierte Verfahren

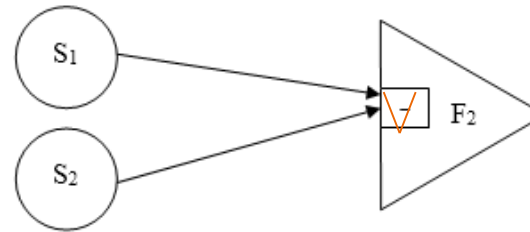
II. Inferenzverfahren

Binäre Schließen, z. B.

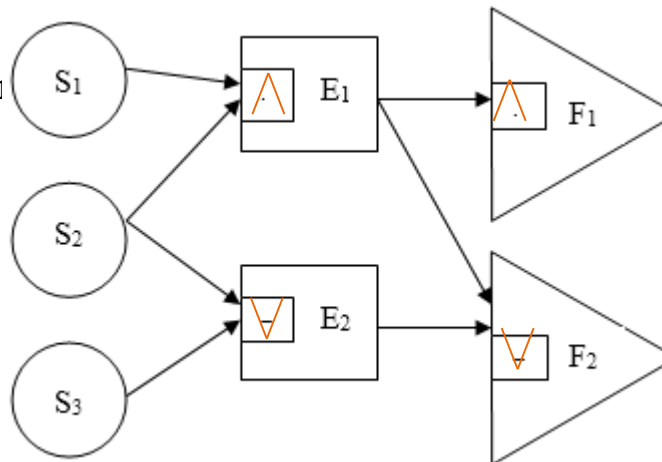
R1: wenn S_1 und S_2 dann F_1



R2: wenn S_1 oder S_2 dann F_2



R3: wenn S_1 und S_2 dann E_1
wenn S_2 oder S_3 dann E_2
wenn E_1 dann F_1
wenn E_1 oder E_2 dann F_2



2. Diagnosestrategien

2.2.3. Wissensbasierte Verfahren

Fehlerbäume auf Basis des Bayes-Konzeptes

Grundlagen: Wahrscheinlichkeiten für $S_i \Rightarrow P(S_i)$ und $F_j \Rightarrow P(F_j)$

bedingte Wahrscheinlichkeit für kausale Zusammenhänge von S_i und F_j
im Netz / Baum $\Rightarrow P(F_j/S_i)$

Lösung: Probabilistische Schließen


d. h. Berechnung der bedingten Wahrscheinlichkeiten der Fehler mit dem Bayes-Theorem

$$P(F_j/S_i) = \frac{P(F_j, S_i)}{P(S_i)}$$


$$P(F_j/S_i) = \frac{P(F_j \cap S_i)}{P(S_i)}$$

Beispiel:

$$P(S_i) = 0,5; P(F_j, S_i) = 0,5$$


$$P(F_j/S_i) = 1$$

$$P(S_i) = 0,5; P(F_j, S_i) = 0$$


$$P(F_j/S_i) = 0$$

2. Diagnosestrategien

2.2.3. Wissensbasierte Verfahren

Fehlerbäume auf Basis des Fuzzy-Konzeptes

Grundlagen:

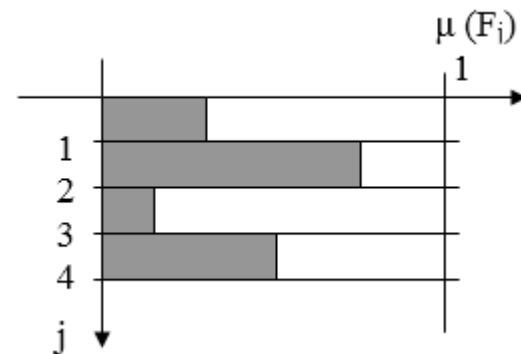
- Zugehörigkeitsfunktionen für $S_i \Rightarrow \mu(S_i)$ mit $0 \leq [\mu(s_i)] \leq 1$
Vorteil: Abbildung heuristischer und / oder analytischer Symptome möglich
- Regelwerke der Form
Wenn < Bedingung >
Dann < Schlussfolgerung >
- Inferenzstrategien
Max-Min-Strategie
Max-Prod-Strategie

Lösung: Entscheidungsstrategie

$$\max\{\mu(F_j)\}$$

$$j = 1, 2, \dots, m$$

Beispiel:



2. Diagnosestrategien

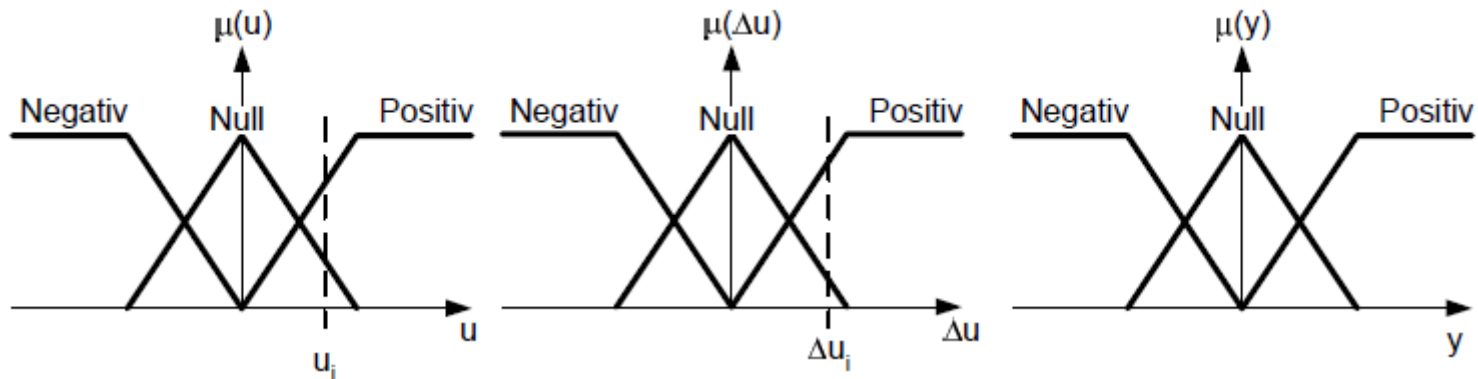
2.2.3. Wissensbasierte Verfahren

Inferenz: Auswertung der Regelbasis, wodurch aus fuzzifizierten Eingangsgrößen eine Fuzzy-Menge der Ausgangsgrößen erzeugt wird.

Teilschritte der Inferenz:

- a) Prämisseauswertung,
- b) Aktivierung und
- c) Aggregation.

Beispiel: R1: Wenn u Null und Δu Null, dann y Null.
R2: Wenn u Positiv oder Δu Positiv, dann y Negativ.



2. Diagnosestrategien

2.2.3. Wissensbasierte Verfahren

a) **Prämissenauswertung:** Bestimmung des Zugehörigkeitsgrades der Prämisse (WENN-Teil) einer linguistischen Regel durch Verknüpfung der Zugehörigkeitsgrade aller linguistischer Teilprämissen mittels Fuzzy-Operatoren.

UND -> Minimum

ODER -> Maximum

b) **Aktivierung:** Bestimmung des Zugehörigkeitsgrades der Konklusion (DANN-Teil) einer linguistischen Regel aus dem Zugehörigkeitsgrad der Prämisse.

MIN-Methode

PROD-Methode

c) **Aggregation:** Zusammensetzen der einzelnen durch die Aktivierung festgelegten Zugehörigkeitsgrade der Konklusion

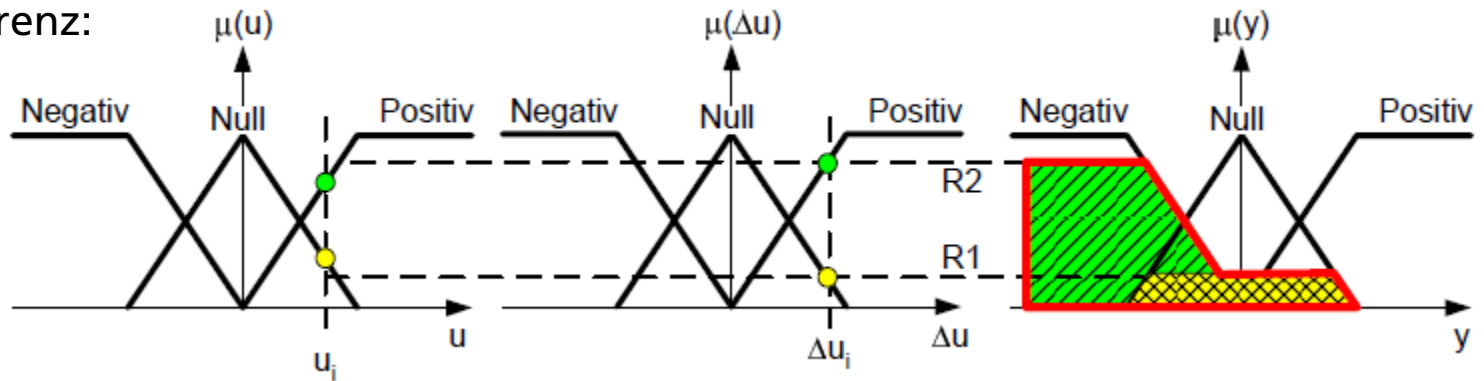
MAX-Methode

-> MAX-MIN-Inferenz
MAX-PROD-Inferenz

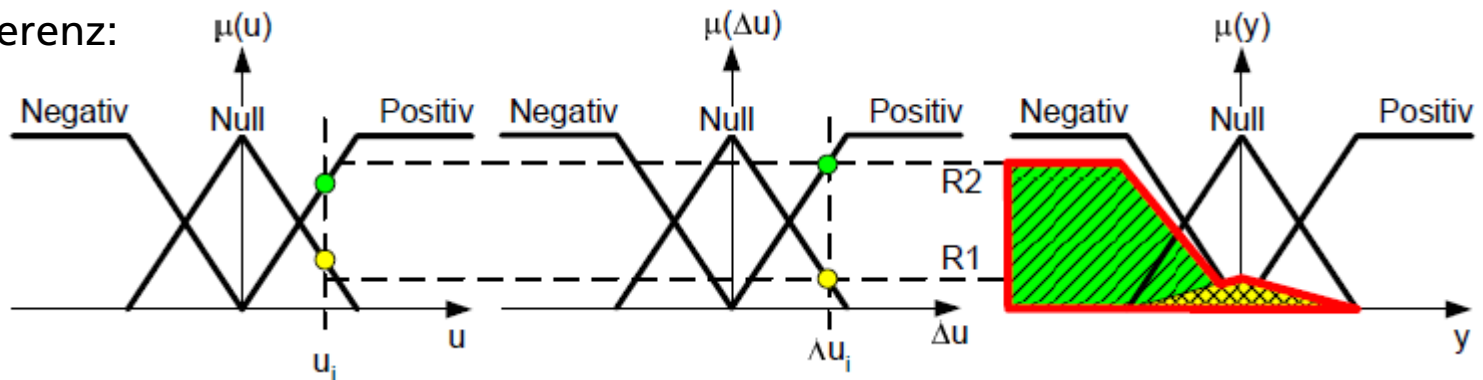
2. Diagnosestrategien

2.2.3. Wissensbasierte Verfahren

MAX-MIN-Inferenz:



MAX-PROD-Inferenz:



2. Diagnosestrategien

2.2.3. Wissensbasierte Verfahren

Defuzzifizierung: z. B. Schwerpunktmethod

$$y_D = \frac{\int_{-\infty}^{+\infty} y \cdot \mu(y) \cdot dy}{\int_{-\infty}^{+\infty} \mu(y) \cdot dy}$$

2. Diagnosestrategien

2.2.3. Wissensbasierte Verfahren

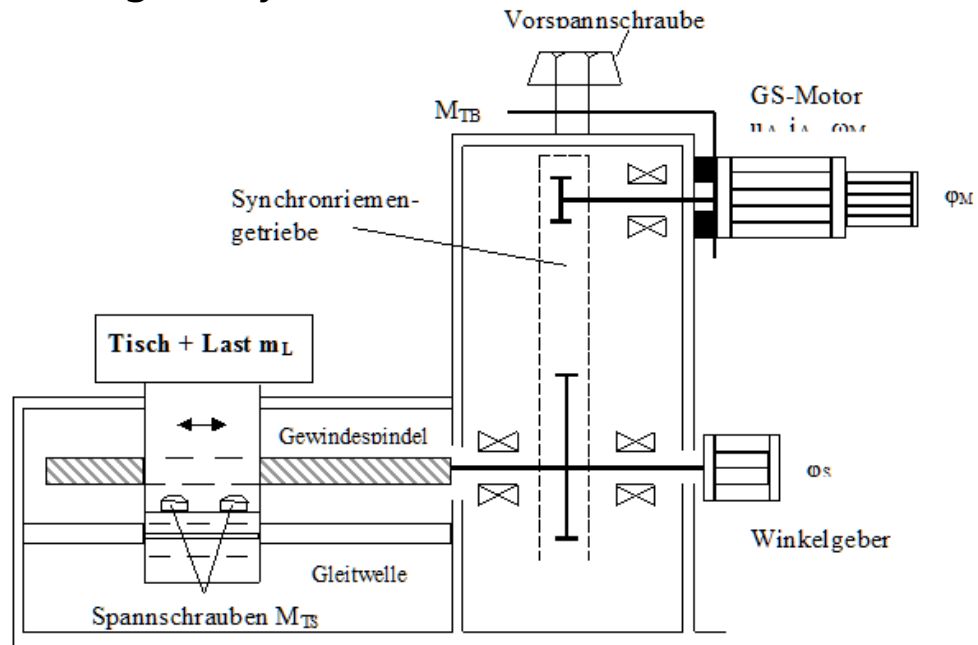
Projekt: Werkzeugmaschinenvorschub (Riemenschaden)

Aufgabe: rechtzeitiges Erkennen eines geschädigten Riemens (z. B. gerissener Cordfaden)

Lösung: Entwurf eines Diagnosesystems auf der Basis unscharfer Konzepte

Technologisches Schema:

Tisch und Last der Maschine werden über die Getriebespindel durch den gezahnten Riemen von einem drehzahlregulierten Gleichstrommotor angetrieben



[Isermann, 1996]

2. Diagnosestrategien

2.2.3. Wissensbasierte Verfahren

Projekt: Werkzeugmaschinenvorschub (Riemenschaden)

Gemessene Größen: $u_A(t)$ Ankerspannung
 $i_A(t)$ Ankerstrom
 $\omega_M(t)$ Motordrehzahl
 $\omega_S(t)$ Spindeldrehzahl
über Modellreferenzverfahren
10 Parameter geschätzt, z. B. Trägheitsmoment I_{sm}
Coulomb-Reibungskoeffizient M_{FC}
Riemensteifigkeit c_{TR}
Riemendämpfung d_{TR}

Lösung: Erstellung eines Fehler-Symptom-Baums mit
Zugehörigkeitsfunktionen der verschiedenen Symptome

Arten der Symptome: $S_a \triangleq$ analytische Symptome
 $S_h \triangleq$ heuristische Symptome
 $S_i \triangleq$ Symptome der Prozessgeschichte
 $\bar{S} \triangleq$ analytische Symptome

[Isermann, 1996]

2. Diagnosestrategien

2.2.3. Wissensbasierte Verfahren

Projekt: Werkzeugmaschinenenvorschub (Riemenschaden)

Fehler-Symptom-Baum:

Beispiele:

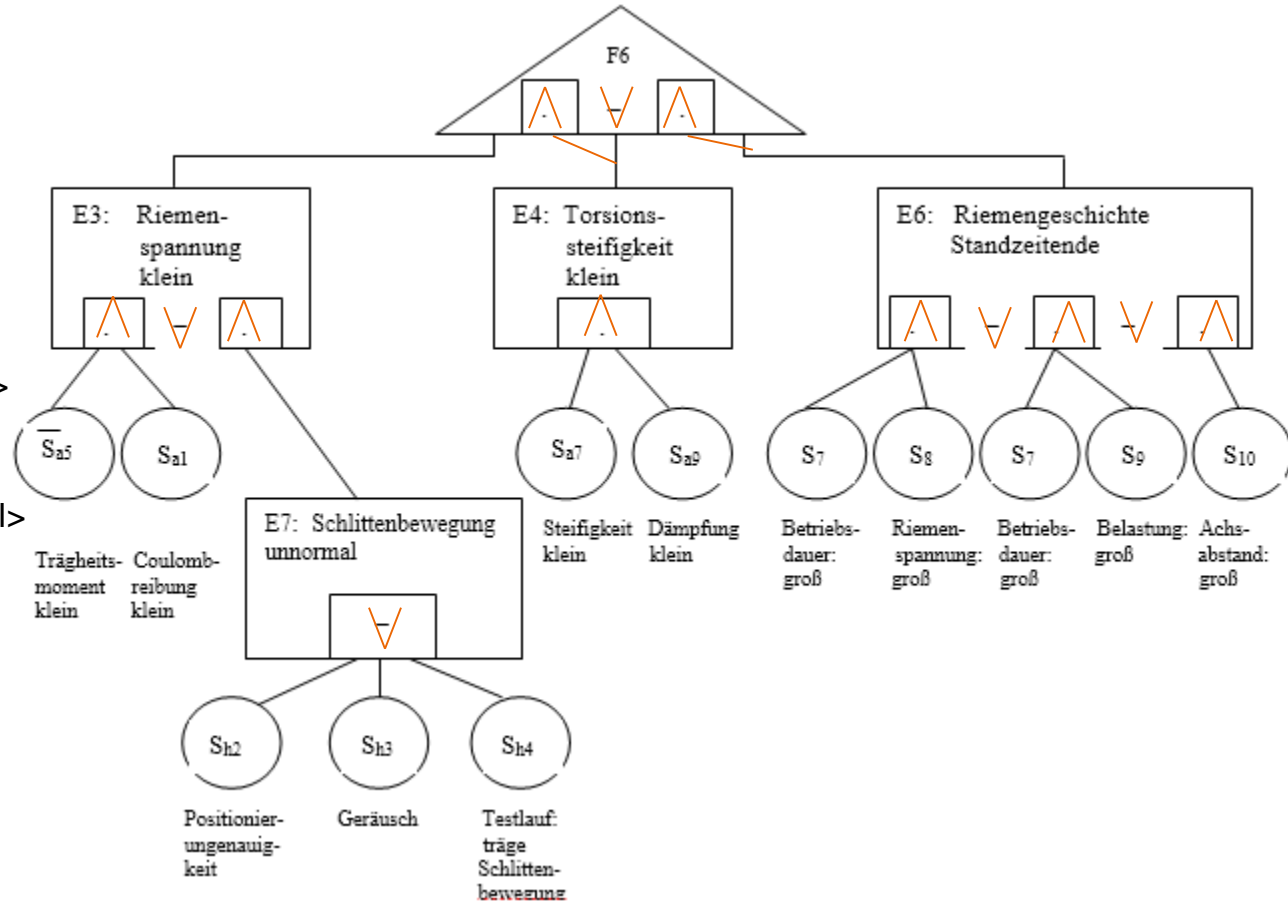
R: Wenn Torsionssteifigkeit <klein>
UND Riemenspannung <klein>
Dann <Riemenschaden>

Wenn Trägheitsmoment nicht <klein>
UND Coulombreibung <klein>
Dann Riemenspannung <klein>

Wenn Schlittenbewegung <unnormal>
Dann Riemenspannung <klein>

R: Wenn Torsionssteifigkeit <klein>
Dann <Riemenschaden>

Wenn Steifigkeit <klein>
UND Dämpfung <klein>
Dann Torsionssteifigkeit <klein>



[Isermann, 1996]

Studiengang Ingenieurinformatik - Lehrblätter -

Prof. Dr.-Ing. habil. Thomas Rauschenbach

Inhalt

1. Einführung

- Anwendungsbeispiele
- Aufgaben der Diagnose und Vorhersage in Entscheidungssystemen
- Definitionen

2. Diagnosestrategien

- Grundlegende Verfahren der Diagnose
 - signal- und zustandsbasierte Verfahren
 - Verwendung von Systemmodellen
- Diagnosekonzepte
 - Klassifikationsverfahren
 - Modellreferenzverfahren
 - wissensbasierte Verfahren

3. Vorhersagestrategien

- Grundlagen
- Vorhersageprinzipien
- Vorhersagbarkeit
- Prognoseprozess
- Primärdatenaufbereitung
- Vorhersage mit Signalmodellen
 - Vorhersage mit deterministischen Signalmodellen
 - Vorhersage mit stochastischen Signalmodellen
 - musterbasierte Vorhersage
 - Vorhersage mit neuronalen Netzen

3. Vorhersagestrategien

3.1 Grundlagen

Definition

Prognose:

Bestimmung zukünftiger Ereignisse auf Grund von Vergangenheitsinformationen. Von der Wahrsagerei unterscheidet sich die Prognose durch die explizite Unterstellung eines Modells für den beobachteten Prozess. Die Vorhersage verliert daher ihre Gültigkeit, wenn das postulierte Modell nicht zutrifft oder sich ändert.

3. Vorhersagestrategien

3.1 Grundlagen

Gesetz der Vorhersage (P. Lemesurier, 1981)

1. Das Gesetz der überraschenden Erfüllung
Das wahrscheinlichste Ergebnis ist das, was am wenigsten erwartet wird.
2. Das Gesetz der vereitelten Erwartung
Die öffentliche Deutung ist am ehesten falsch.
3. Das Gesetz des beeinträchtigenden Vorurteils
Vorurteile und Prophezeiungen passen nicht zusammen.
4. Das Gesetz der Selbstbewahrheitung
Prophezeiungen neigen dazu, sich selbst zu bewahrheiten.
5. Das Gesetz der abnehmenden Genauigkeit
Die Genauigkeit einer Prophezeiung nimmt im Quadrat der Zeit zu ihrer Erfüllung ab.
6. Das Gesetz der getrennten Funktionen
Prophezeiungen und Interpretationen sind als Tätigkeiten unvereinbar.
7. Das Gesetz der prophetischen Perspektivierung
Hellseherei zoomt die Zukunft heran.
8. Das Gesetz der nicht existenten Unmöglichkeit
Was passieren kann, passiert auch. Was nicht passieren kann, passiert vielleicht.

3. Vorhersagestrategien

3.1 Grundlagen

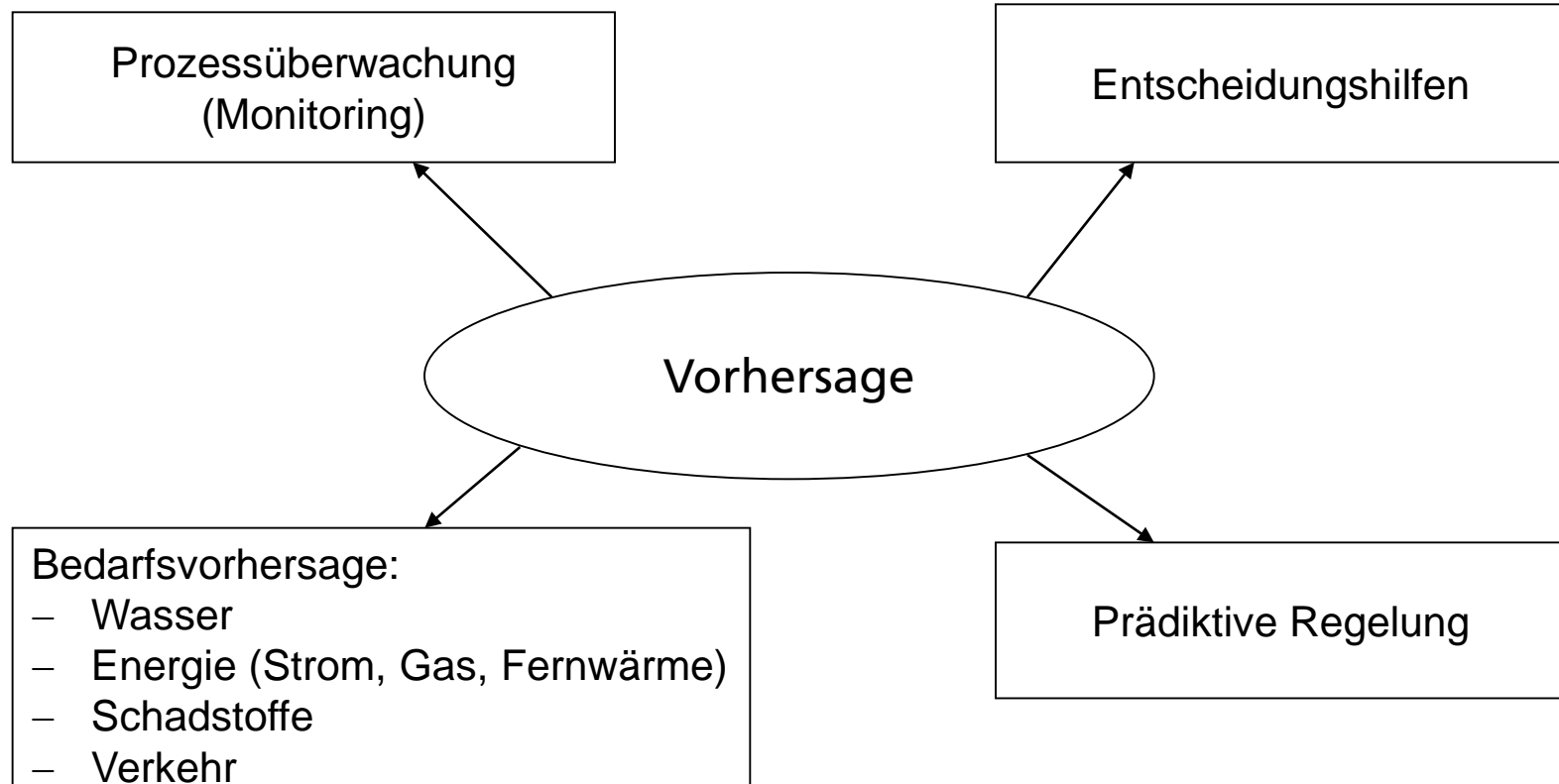
Vorhersagehorizonte / Entscheidungsebenen

Horizont	Ebenen
Kurzfrist 1 min ⋮ 15 min ⋮ 1 h ⋮ 24 h	operative
Mittelfrist 1 Tag ⋮ 7 Tage ⋮ 30 Tage	taktische
Langfrist 1 Monat ⋮ 6 Monaten ⋮ 12 Monate	strategische

3. Vorhersagestrategien

3.1 Grundlagen

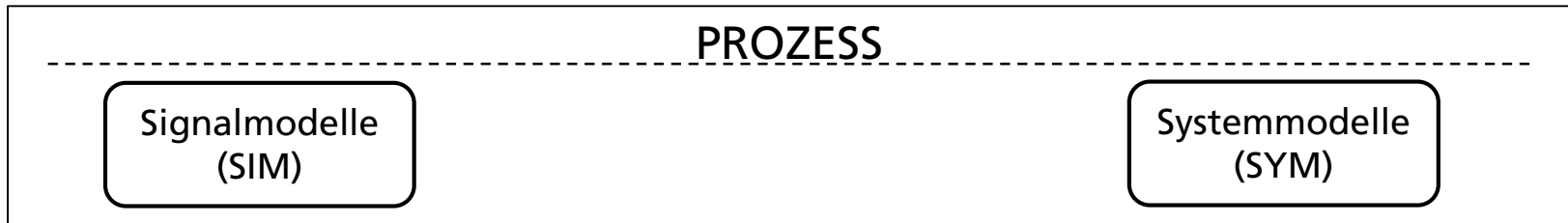
Einsatzgebiete der Vorhersage



3. Vorhersagestrategien

3.2 Vorhersageprinzipien

Vorhersagemodelle



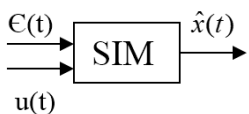
- Deterministische



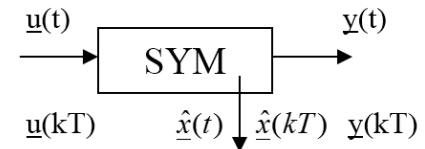
- Stochastische



- Stochastische



- Klassifikatorkonzepte
- Künstliche Neuronale Netze
- Fuzzy Modelle

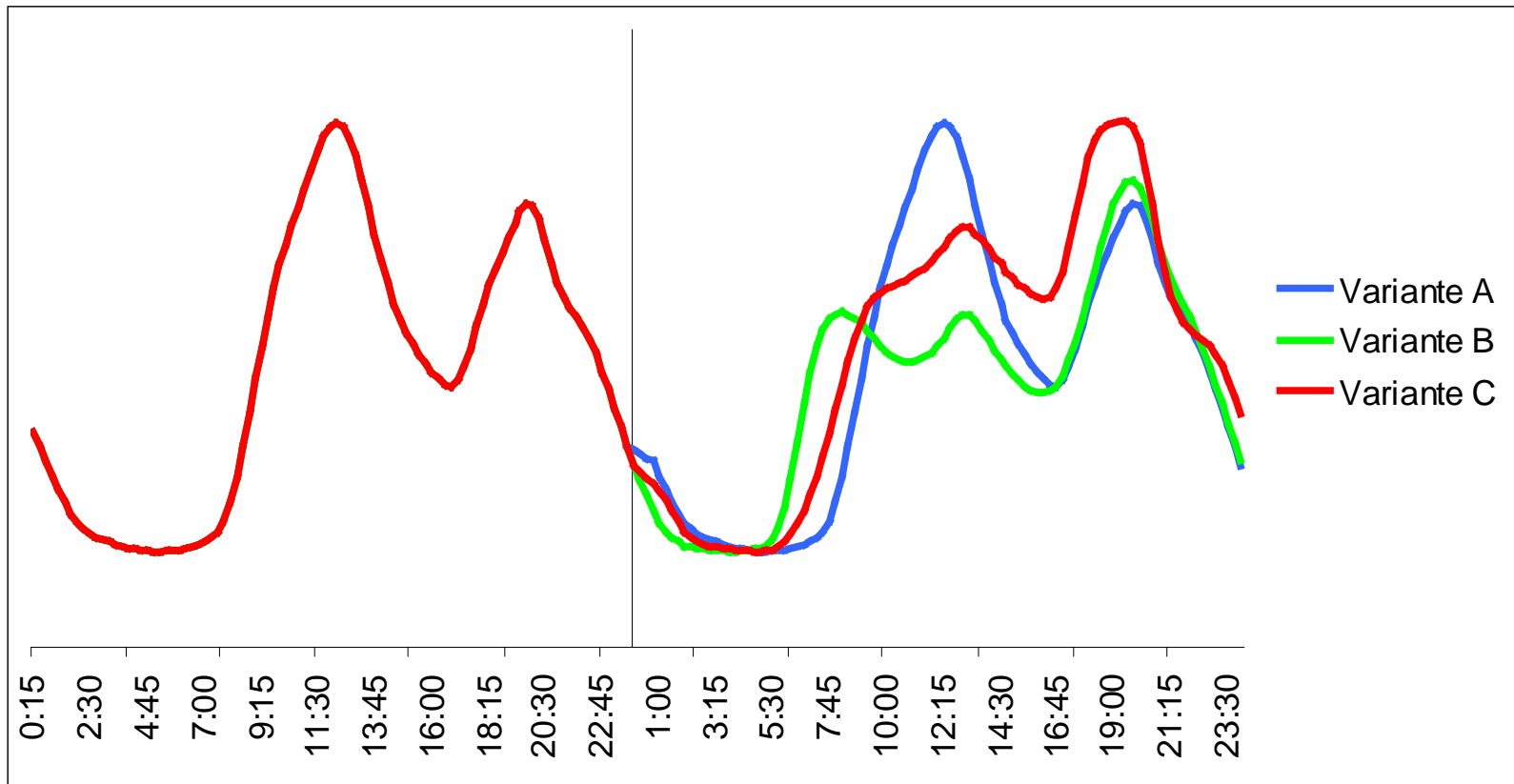


- Ein-/ Ausgangsmodell
 - Differenzgleichungen
- Zustandsmodell
 - Beobachter
 - Filter
- Fuzzy Modell
- Künstliche Neuronale Netze

3. Vorhersagestrategien

3.2 Vorhersageprinzipien

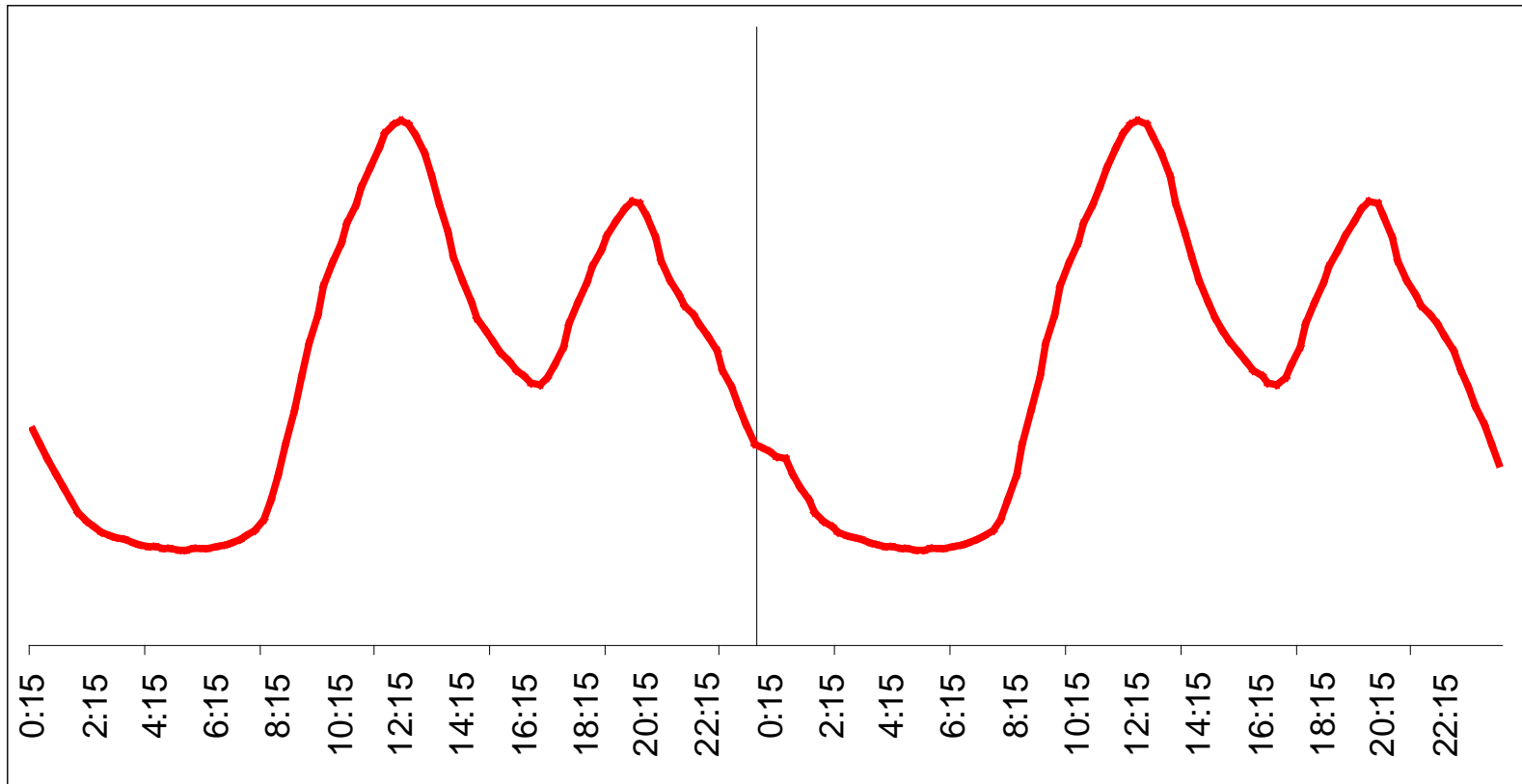
Aufgaben der Vorhersage



Abschätzung der zukünftigen Entwicklung von Signalen
Berücksichtigung von Zuständen

3. Vorhersagestrategien

3.2 Vorhersageprinzipien (1)

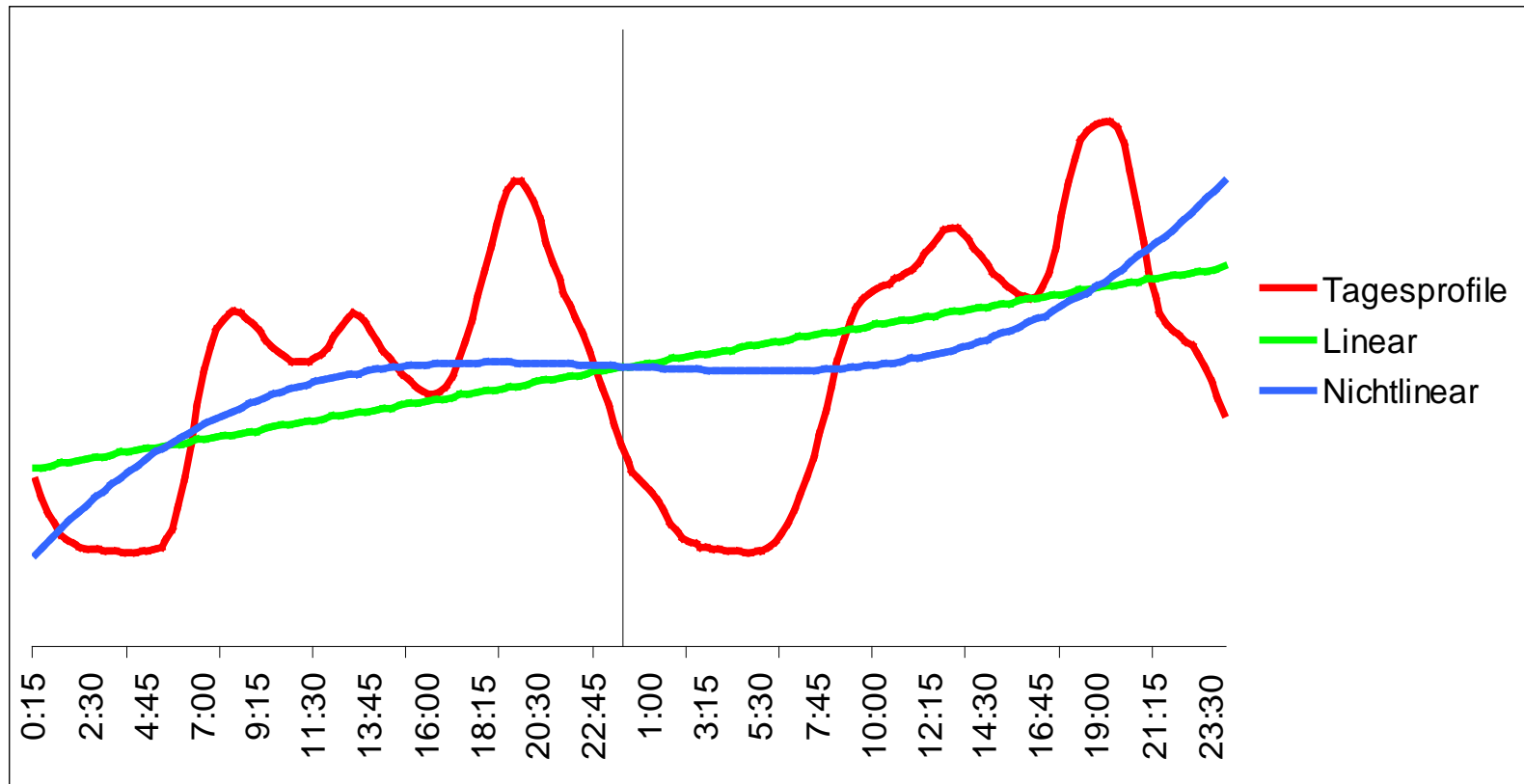


Zukunft = Vergangenheit

keine Vergangenheit = keine Prognose

3. Vorhersagestrategien

3.2 Vorhersageprinzipien (2)

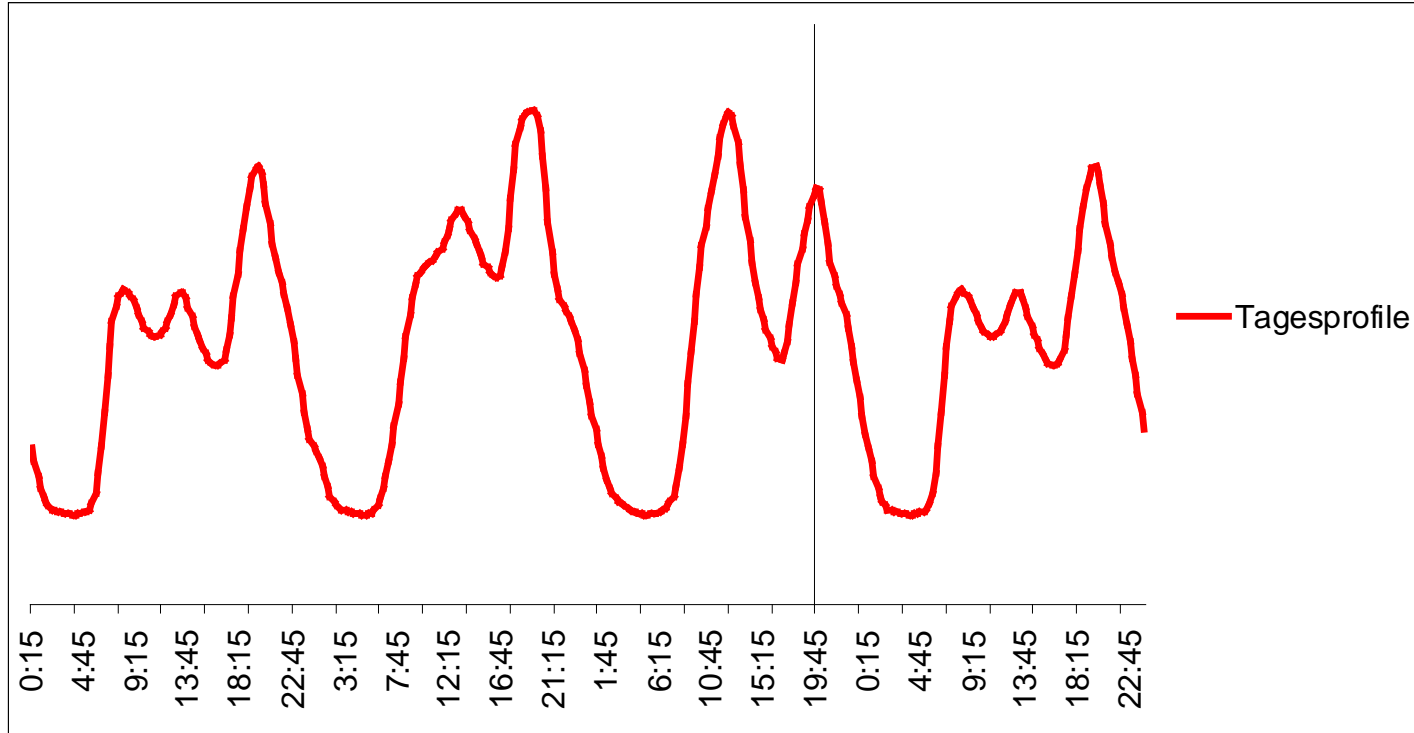


Zukunft = Vergangenheit + linearer Trend

Zukunft = Vergangenheit + nichtlinearer Trend

3. Vorhersagestrategien

3.2 Vorhersageprinzipien (3)

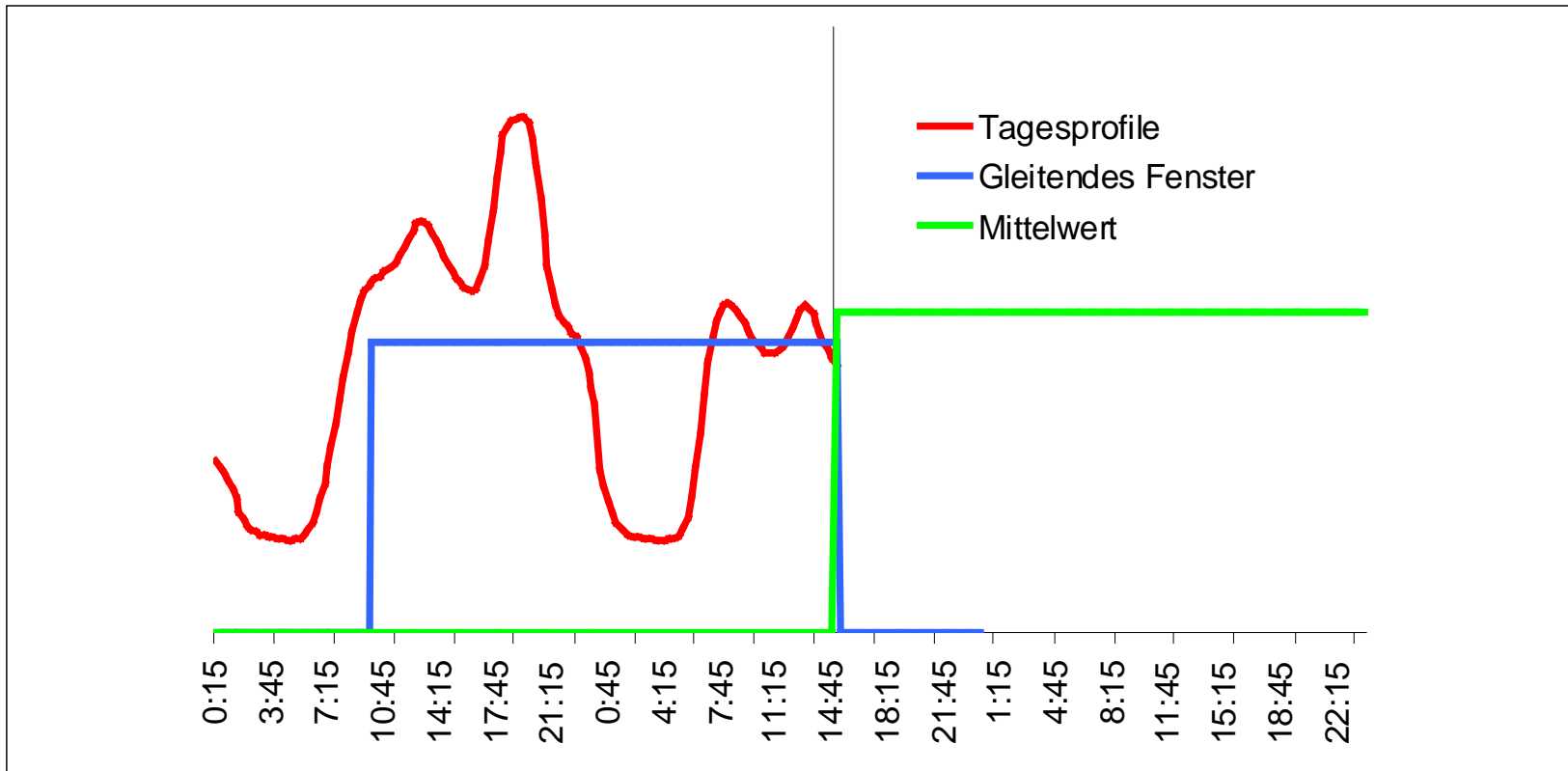


Zukunft = Vergangenheit bei periodischer Fortsetzung

Zukunft = Vergangenheit bei periodischer Fortsetzung & Phasenverschiebung

3. Vorhersagestrategien

3.2 Vorhersageprinzipien (4)

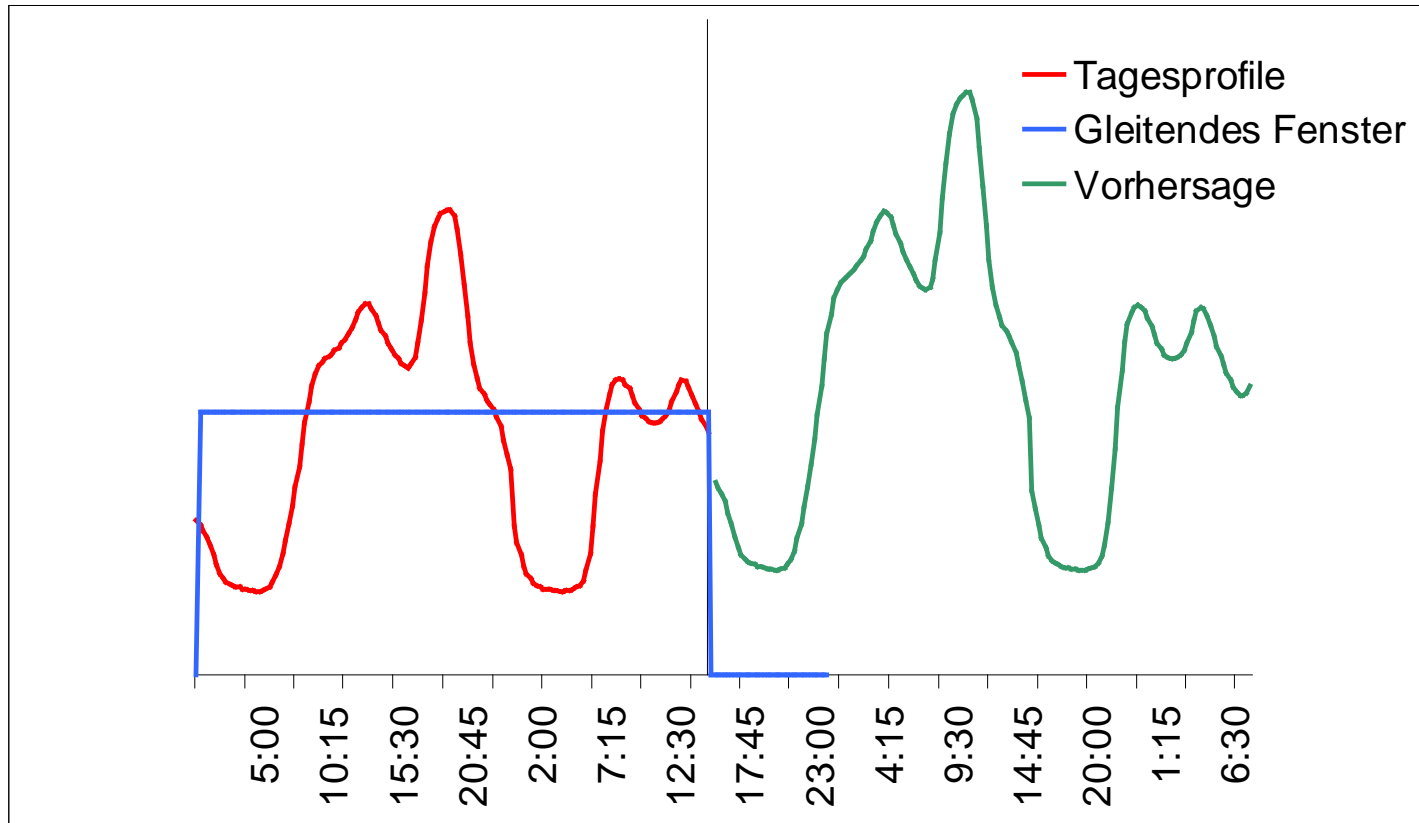


Zukunft = gleitender Mittelwert (Moving Average)

Gleitendes Zeitfenster bestimmt relevanten Zeitbereich

3. Vorhersagestrategien

3.2 Vorhersageprinzipien (5)

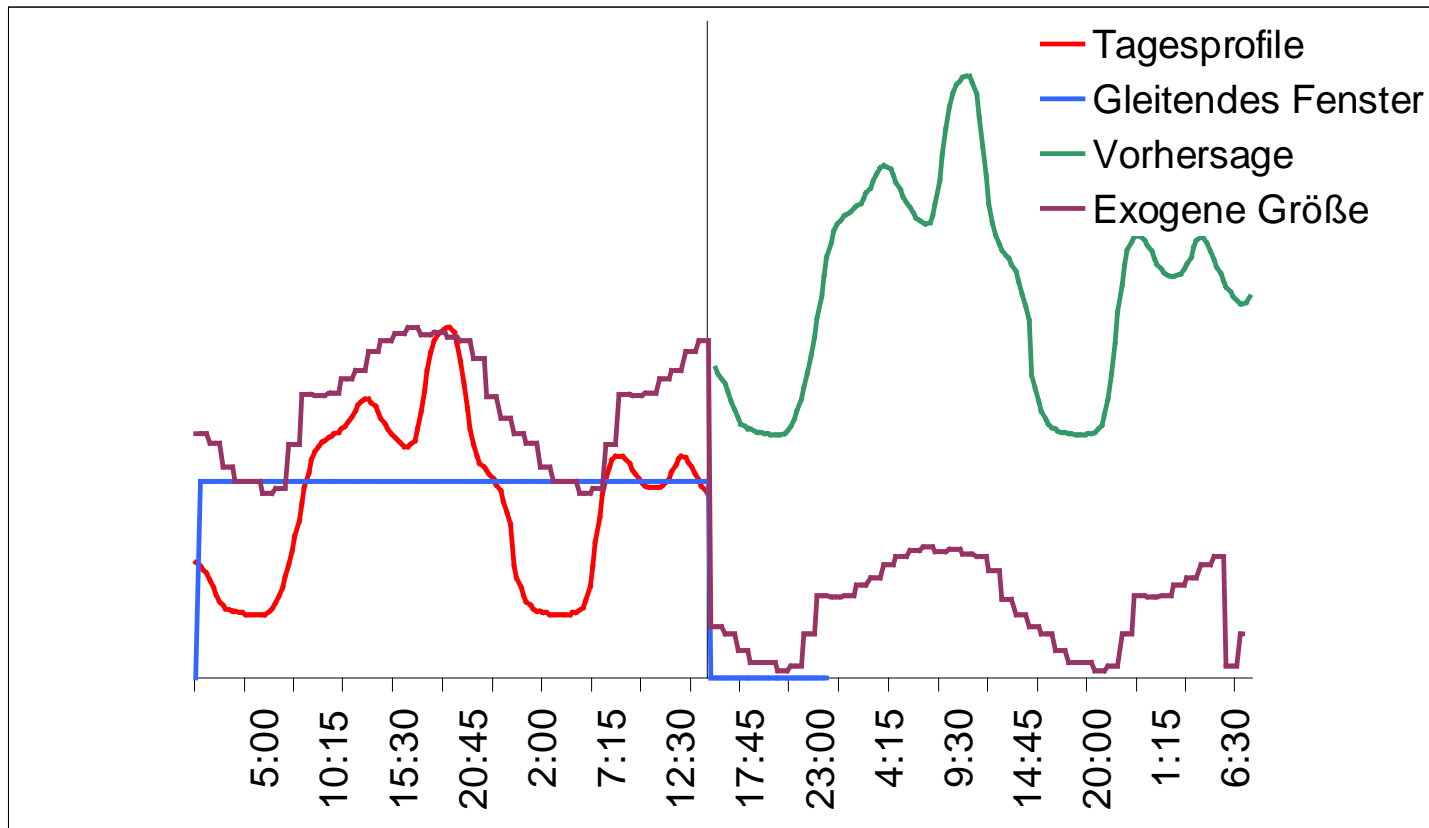


Zukunft = gewichtete Vergangenheit (Auto Regressive)

Gleitendes Zeitfenster bestimmt relevanten Zeitbereich

3. Vorhersagestrategien

3.2 Vorhersageprinzipien (6)



Zukunft = gewichtete Vergangenheit + exogene Einflussgröße (ARX)

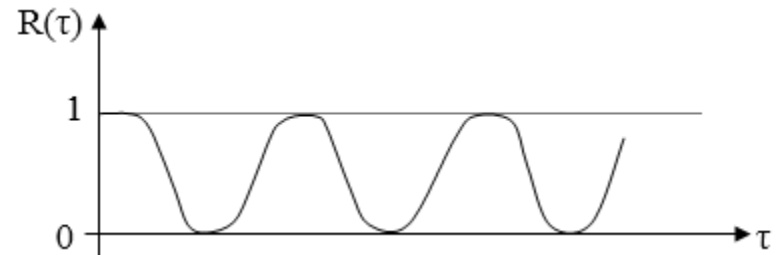
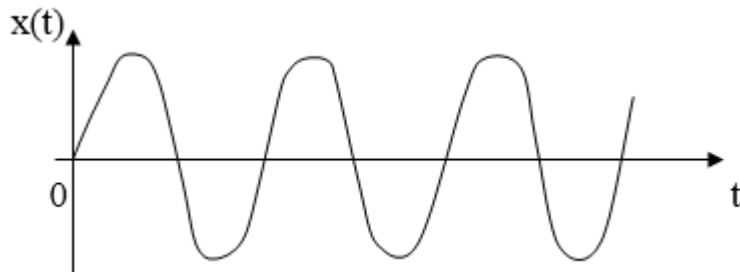
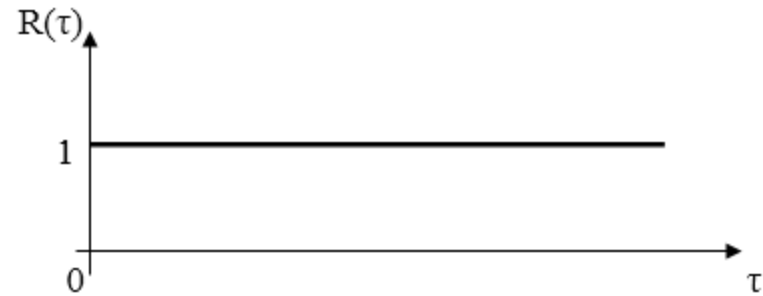
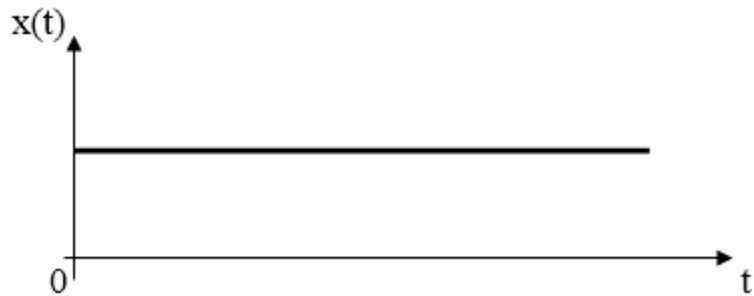
3. Vorhersagestrategien

3.3 Vorhersagbarkeit

3.3.1 Lineare Vorhersagbarkeit (1)

- Bestimmung mit Hilfe der Autokorrelationsfunktion

a) ohne Fehler vorhersagbar



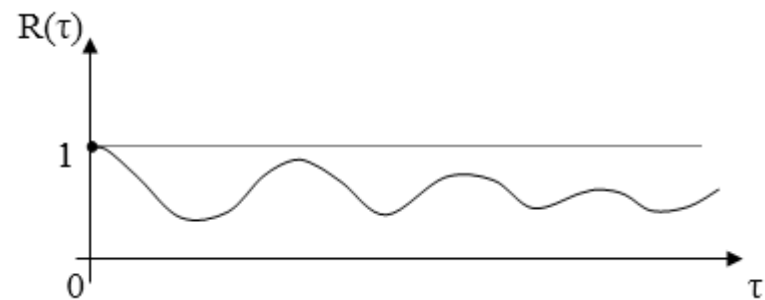
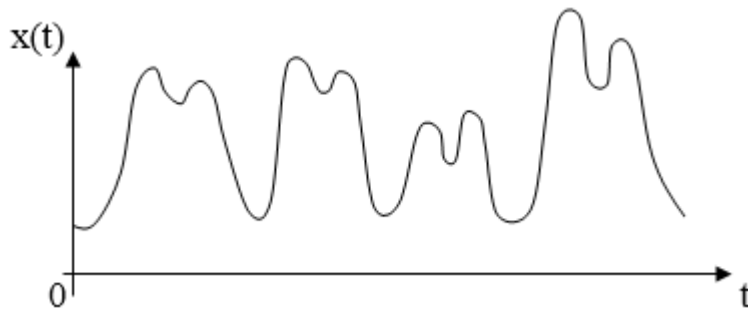
3. Vorhersagestrategien

3.3 Vorhersagbarkeit

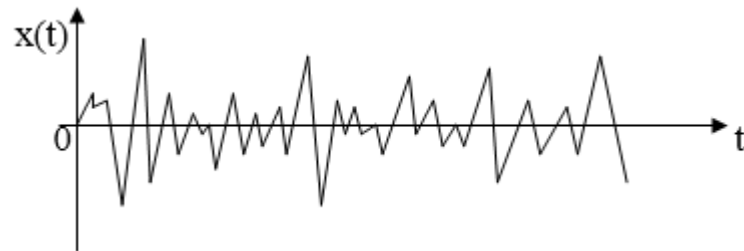
3.3.1 Lineare Vorhersagbarkeit (2)

- Bestimmung mit Hilfe der Autokorrelationsfunktion

b) vorhersagbar, aber mit Fehler



c) nicht vorhersagbar



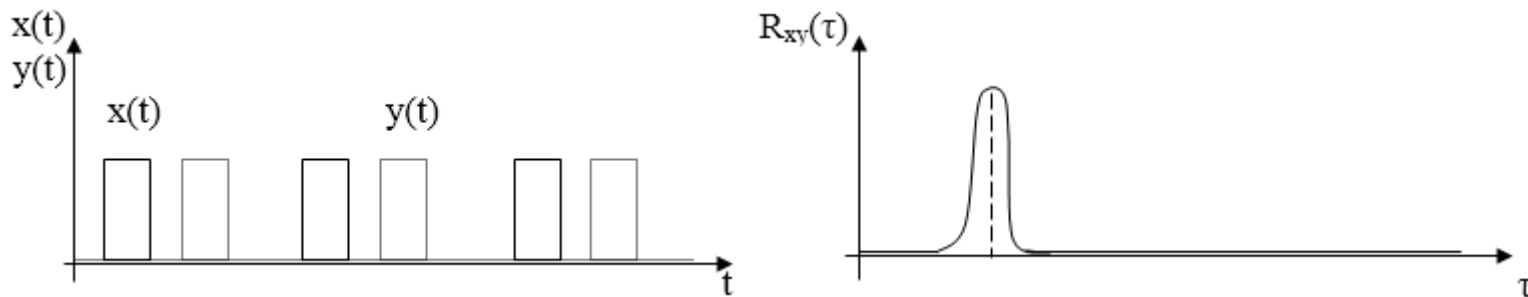
3. Vorhersagestrategien

3.3 Vorhersagbarkeit

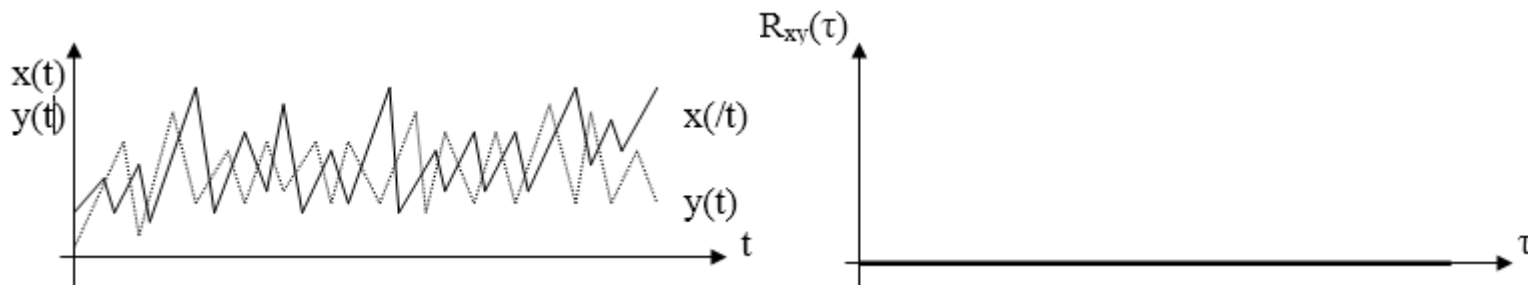
3.3.1 Lineare Vorhersagbarkeit (3)

- Bestimmung mit Hilfe der Kreuzkorrelationsfunktion

a) vorhersagbar



b) nicht vorhersagbar



3. Vorhersagestrategien

3.3 Vorhersagbarkeit

3.3.2 Nichtlineare Vorhersagbarkeit

- Ausgangspunkt: Definition der stochastischen Unabhängigkeit von zwei Zufallsvariablen (Zufallsvektoren) \underline{x} und \underline{y} . Diese sind unabhängig und nur dann unabhängig, wenn gilt:

$$p(\underline{x} \cdot \cap \underline{y}) = p(\underline{x}) \cdot p(\underline{y}) \quad \text{p-Wahrscheinlichkeit}$$

- Berechnung der „gegenseitigen Informationen“ zwischen zwei Signalen:

$$I(\underline{x}, \underline{y}) = \iint p(\underline{x} \cap \underline{y}) \ln \frac{p(\underline{x} \cap \underline{y})}{p(\underline{x}) \cdot p(\underline{y})} d\underline{x} d\underline{y}$$

$$0 \leq I \leq \infty$$

- Nominierung zwischen 0 und 1: $\rho = \sqrt{1 - e^{-2I}}$

$p = 0$ \underline{x} enthält keine Informationen über \underline{y}
 \underline{y} kann nicht mit Hilfe von \underline{x} vorhergesagt werden

$p = 1$ Vollständige Determiniertheit
 \underline{y} kann durch \underline{x} vorhergesagt werden

3. Vorhersagestrategien

3.3 Vorhersagbarkeit

3.3.2 Nichtlineare Vorhersagbarkeit - Beispiel 1

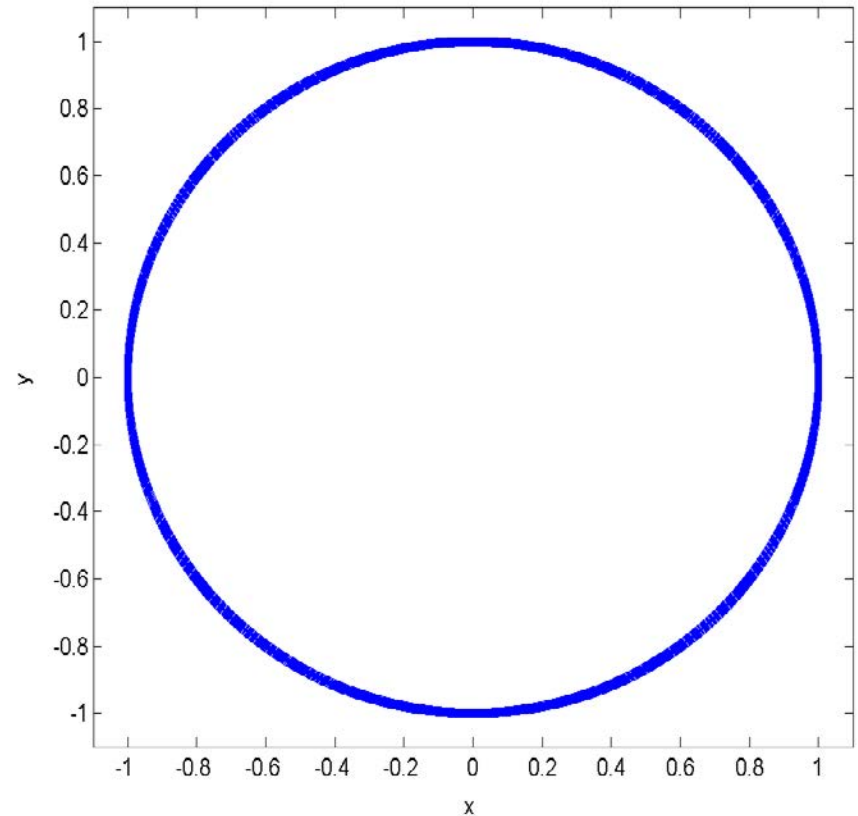
Beispiel zum Testen des Algorithmus für die Berechnung der nichtlinearen Vorhersagbarkeit. Die Daten sind auf einem Kreisumfang verteilt.

Berechnung aus ca. 10.000 Datenpunkten

Ergebnisse:

$\rho = 0,9999$ (nichtlineare Vorhersagbarkeit)

$\lambda = 0,0$ (lineare Vorhersagbarkeit)



3. Vorhersagestrategien

3.3 Vorhersagbarkeit

3.3.2 Nichtlineare Vorhersagbarkeit - Beispiel 2

Berechnung aus ca. 10.000 Datenpunkten

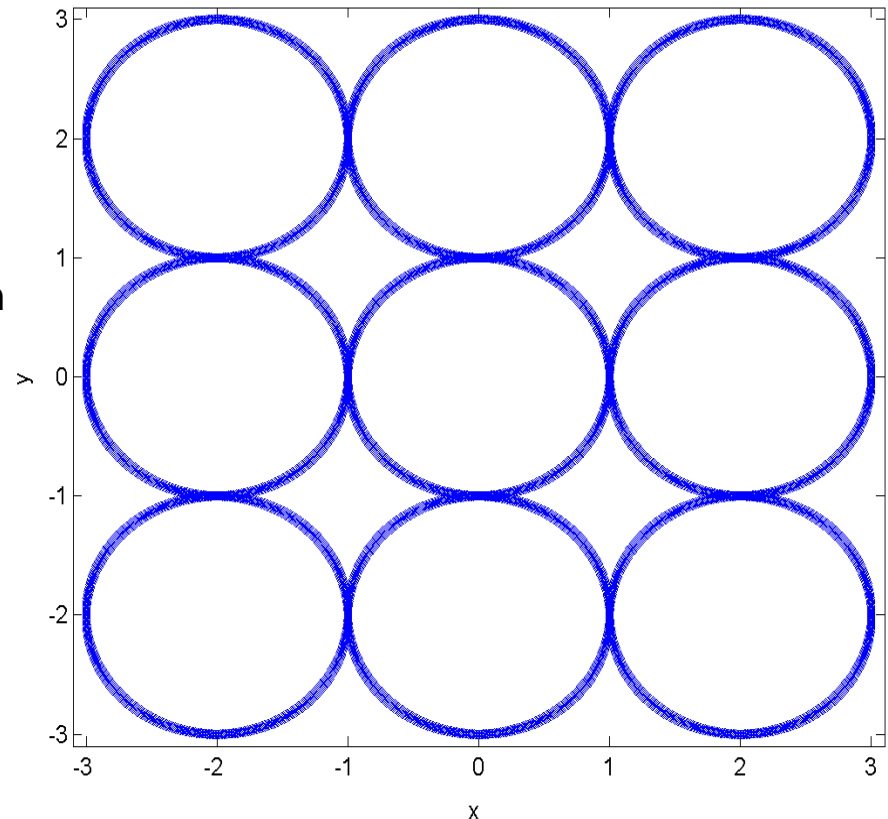
Ergebnisse:

$$\rho = 0,9988$$

(nichtlinear)

$$\lambda = 1,4 \cdot 10^{-7}$$

(linear)



3. Vorhersagestrategien

3.4 Der Prognoseprozess

Teilschritte für die Erstellung eines Prognosemodells

1. Primärdatenaufbereitung

- Beseitigung von Messfehlern und Ausreißern
- Plausibilitätsprüfung des Datenmaterials

2. Modellspezifikation

- Auswahl des Modellansatzes (Modellstruktur)
- Bestimmung der Modellordnung usw.

3. Modellschätzung

- Schätzung der Parameter des Modells bei gegebener Struktur

4. Modellverifikation

- Überprüfung der Gültigkeit und Genauigkeit des angepassten Modells

5. Modellinterpretation

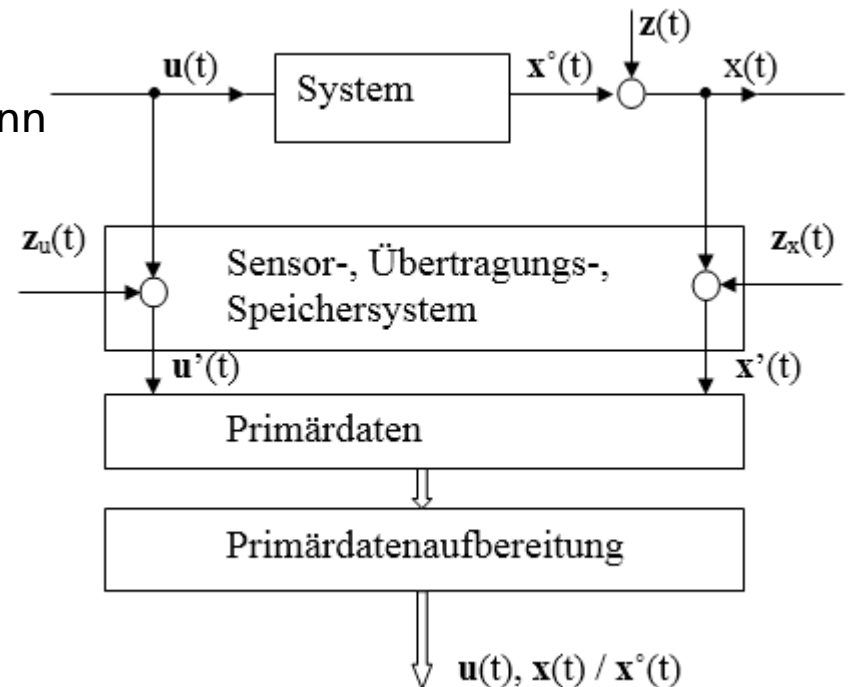
- Substanzwissenschaftliche Deutung des Modells

3. Vorhersagestrategien

3.5 Primärdatenaufbereitung (1)

Warum ist die Primärdatenaufbereitung notwendig?

- Eine fehlerfreie Erfassung von Messwerten und Informationen ist selten gegeben.
- In den Stufen der Signalgewinnung (Erfassung, Übertragung, Speicherung) wird das ursprüngliche Signal verfälscht.
- Die vom Sensor erfasste Prozessgröße kann bereits gestört sein.
- Ursachen sind:
 - Prinzip des Sensors oder
 - Einfluss von „Umweltfaktoren“ und andere nicht messbaren Größen auf den Sensor



3. Vorhersagestrategien

3.5 Primärdatenaufbereitung (2)

Ziel der Primärdatenaufbereitung

Erkennen der Verfälschung der Prozessgrößen und diese durch Transformation, Korrekturen oder Filterung möglichst zu beseitigen

Methoden zur Erkennung von Signalverfälschungen:

- *Grundlage sind A-priori-Informationen über das unverfälschte Verhalten. Diese können gewonnen werden durch:*

- theoretische und / oder experimentelle Analyse des statischen und dynamischen Verhaltens,
 - Vergleichsmessungen und
 - Erfahrungen / Intuition der Prozessspezialisten.
- *Prüfung des statischen Verhaltens*
- Bereichstest (Maximum, Minimum)
 - Differenzentest
 - Bilanztest (z. B. Massenbilanz)
 - Ausreißertest nach Nalimov ($x \cdot \sigma$ – Bereich)

3. Vorhersagestrategien

3.5 Primärdatenaufbereitung (3)

weiter Methoden zur Erkennung von Signalverfälschungen:

- Prüfung des dynamischen Verhaltens
 - Test der Verteilungsmomente eines Signals innerhalb verschiedener Serien von Messungen, z. B.:
 - Test auf Gleichheit der Erwartungswerte (t-Test)
 - Vergleich der mittleren Streuung der Serien (F-Test)

Methoden zur Korrektur der gemessenen Signale:

- Eliminierung von Signalwerten (Problem: zeitliche Kontinuität innerhalb des Signals geht verloren)
- Statische Korrektur von Signalen (z. B. mit Hilfe von Kennlinien)
- Dynamische Korrektur von Signalen (z. B. Endwertvorhersage bei beobachteten Anfangsverläufen)
- Filterung von Signalen
- Splines

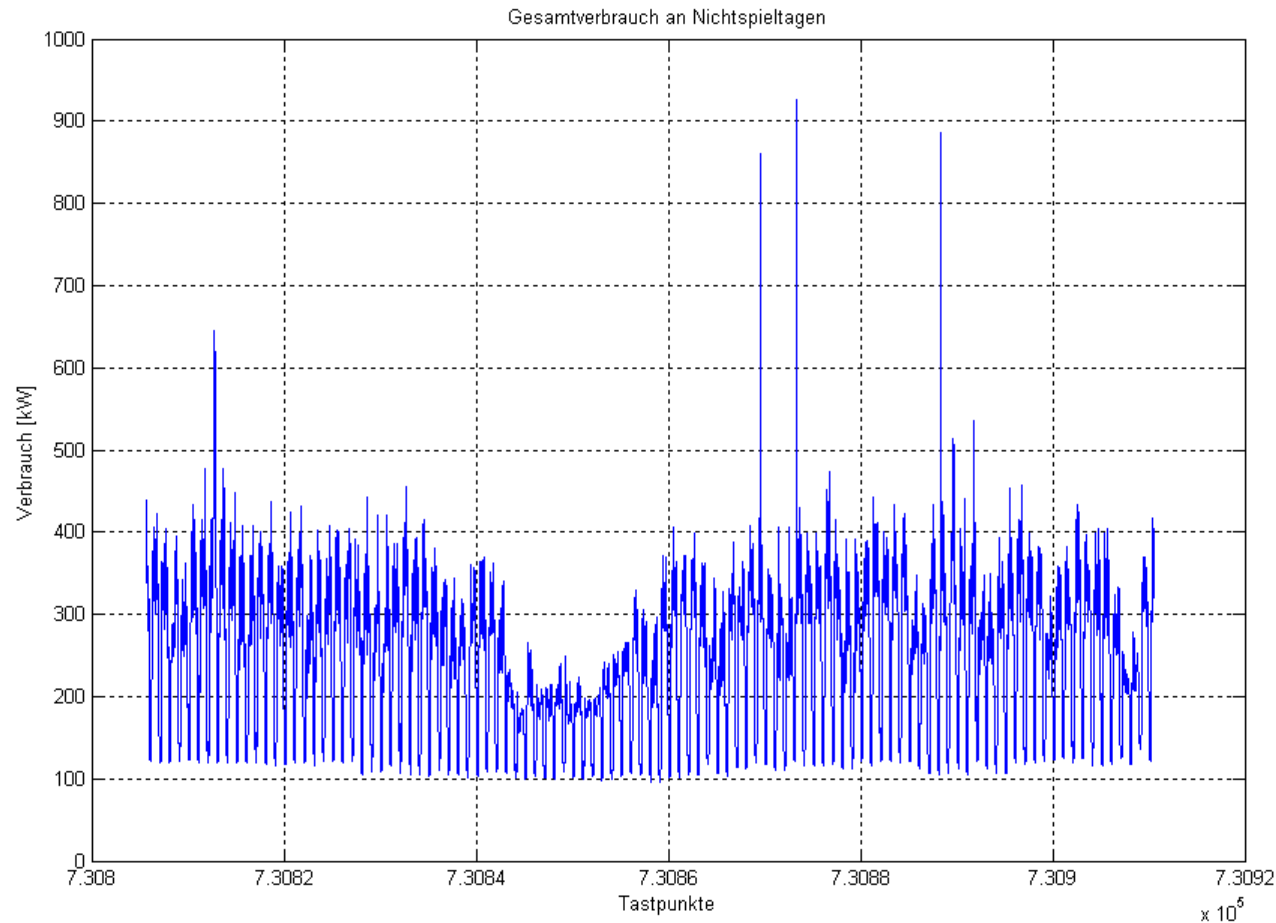
3. Vorhersagestrategien

3.5 Primärdatenaufbereitung (3) – Filter- und Glättungsmethoden

Bezeichnung	Vorschrift
Nichtrekursives Filter (auch als Glättung oder Mittelwertbildung bezeichnet)	$\tilde{x}(k) = \sum_{i=-n}^n b_i x(k-i)$ mit $b_i \geq 0$ n endlich
Rekursives Filter	$\tilde{x}(k) = \sum_{i=0}^n b_i x(k-i) + \sum_{j=0}^m a_j \tilde{x}(k-j)$ mit $b_i \geq 0$ $a_j \geq 0$ n, m endlich
Backshiftoperator (Filter B)	$B \cdot x(k) = x(k-1)$
Backshiftoperator mit Rückwärtsverschiebung τ	$B^\tau \cdot x(k) = x(k-\tau)$
Differenzfilter 1. Ordnung	$\Delta x(k) = x(k) - x(k-1) = (1-B) \cdot x(k)$ mit $k = 2, 3, \dots, n$ $\Delta = 1 - B$
Differenzfilter p -ter Ordnung	(rekursive Berechnungsvorschrift) $\Delta^p x(k) = \Delta^{p-1} x(k) - \Delta^{p-1} x(k-1) = (1-B)^p x(k)$ mit $k = p+1, \dots, n$
Differenzfilter 1. Ordnung mit der Verschiebung τ	$\Delta_\tau x(k) = x(k) - x(k-\tau)$ mit $k = \tau, \dots, n$ $\tau < n$

3. Vorhersagestrategien

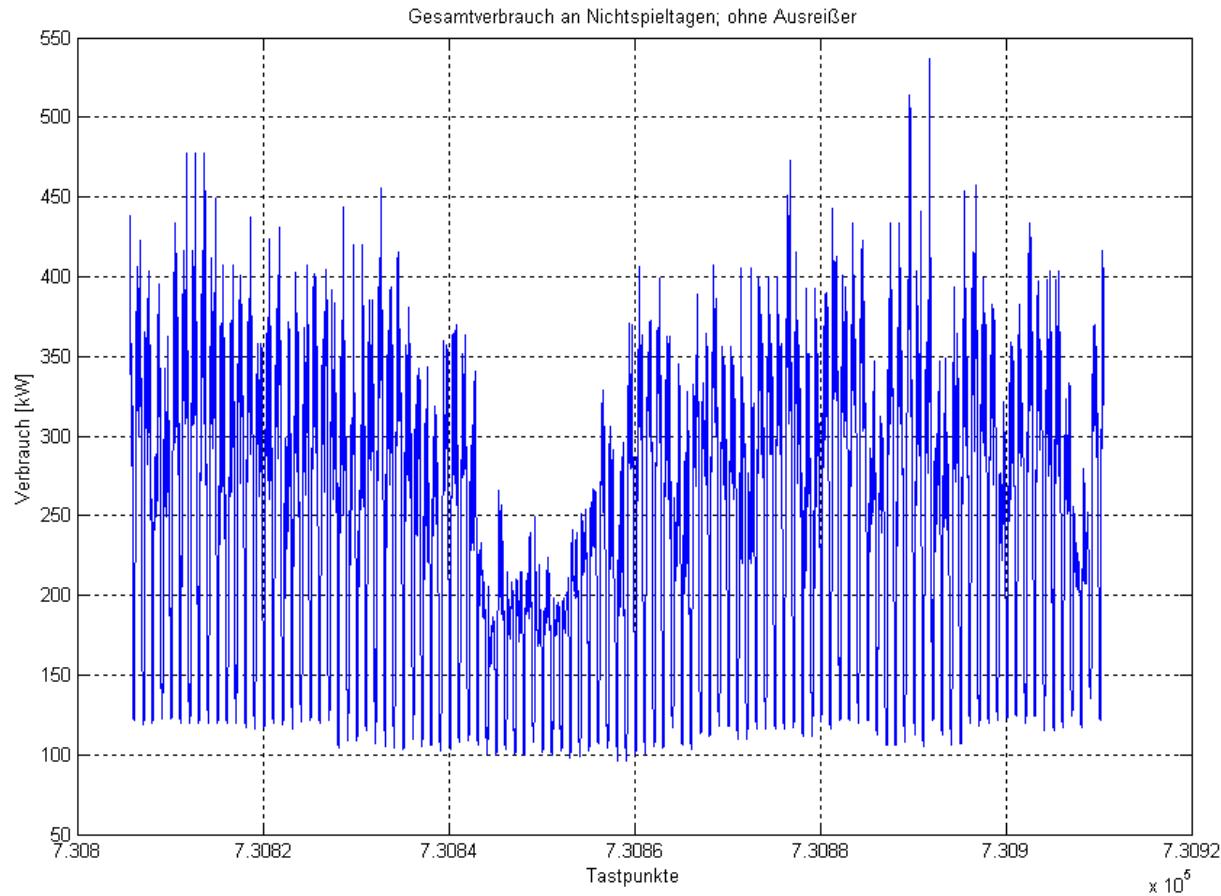
3.5 Primärdatenaufbereitung (3) – Lastdaten der Bay Arena Leverkusen



3. Vorhersagestrategien

3.5 Primärdatenaufbereitung (3) – Lastdaten der Bay Arena Leverkusen

nach
Primärdatenauf-
bereitung



3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.1 Allgemeiner Ansatz zur Beschreibung von Signalen

Klassische Zerlegung eines Signals – Komponentenmodell (1)

Trendkomponente(Polynomkomponente) $x_T(k)$

-> beschreibt die langfristige systematische Veränderung des mittleren Niveaus der Zeitreihe

Saisonale Komponente (periodische Komponente) $x_P(t)$

-> Schwankungskomponente, die sich in regelmäßigen Abständen wiederholt, Beschreibung durch Sinus- oder Kosinusfunktionen

Musterkomponente $x_M(t)$

-> situationsabhängige Wiederkehr von typischen Mustern im Signal

Stochastische Komponente $x_S(t)$

-> zusammenfassende Beschreibung der nicht zu erklärenden Einflüsse oder Störungen

3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.1 Allgemeiner Ansatz zur Beschreibung von Signalen

Klassische Zerlegung eines Signals – Komponentenmodell (2)

Kombinationsmöglichkeiten für Komponenten:

- a) Additive Verknüpfung $x(k) = x_T(k) + x_P(k) + x_M(k) + x_S(k)$
- b) Multiplikative Verknüpfung $x(k) = x_T(k) \cdot x_P(k) \cdot x_M(k) \cdot x_S(k)$

Für Variante a) ist die Schätzung der einzelnen Signalkomponenten einfacher, da das Verstärkungs- und Überlagerungsprinzip gilt. Die Ergebnisse sind außerdem besser interpretierbar. Daher wird Variante b) häufig in ein additives Modell überführt.

Es gilt dann:

$$\begin{aligned}\log(x(k)) &= \log(x_T(k) \cdot x_P(k) \cdot x_M(k) \cdot x_S(k)) \\ &= \log(x_T(k)) + \log(x_P(k)) + \log(x_M(k)) + \log(x_S(k))\end{aligned}$$

Das Komponentenmodell kann angewendet werden:

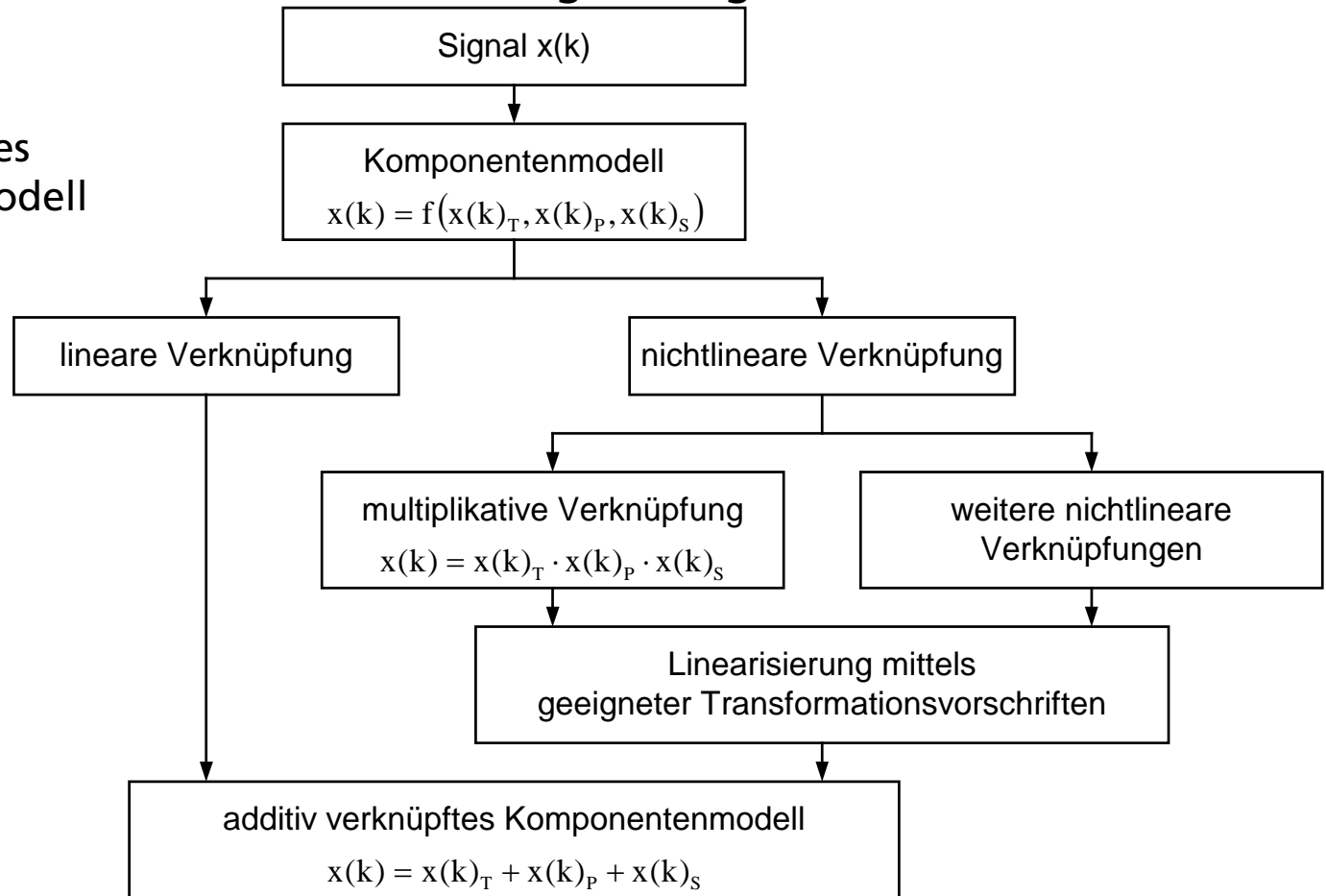
- auf das gesamte Signal oder
- auf Teilstücke des Signals

3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.1 Allgemeiner Ansatz zur Beschreibung von Signalen

Zerlegung eines
Signals in ein
linear verknüpftes
Komponentenmodell



3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.2 Vorhersage mit deterministischen Signalmodellen

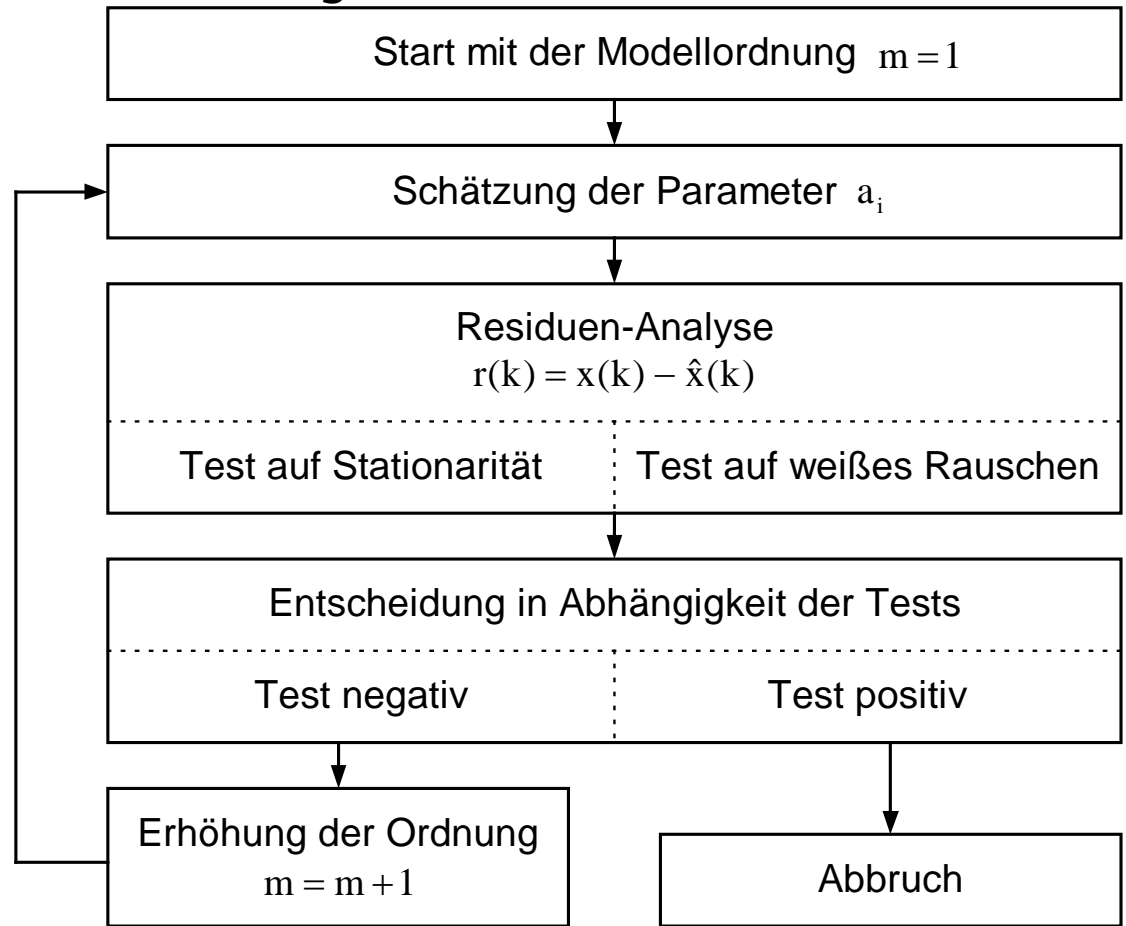
Bezeichnung	Modellgleichung
Polynommodell	$x(k) = a_0 + \sum_{i=1}^m a_i \cdot (kT)^i$ <p>mit $k = 1, \dots, n$ n Anzahl der Messpunkte T Abtastzeit m Ordnung des Modells a_0, a_i Modellparameter</p>
Periodisches Modell	$x(k) = a_0 + \sum_{i=1}^m (a_i \sin \omega_i(kT) + b_i \cos \omega_i(kT))$ <p>mit $k = 1, \dots, n$ n Anzahl der Messpunkte T Abtastzeit m Ordnung des Modells ω_i Frequenz a_0, a_i, b_i Modellparameter</p>

3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.2 Vorhersage mit deterministischen Signalmodellen

Ermittlung der
Ordnung des
Polynommodells

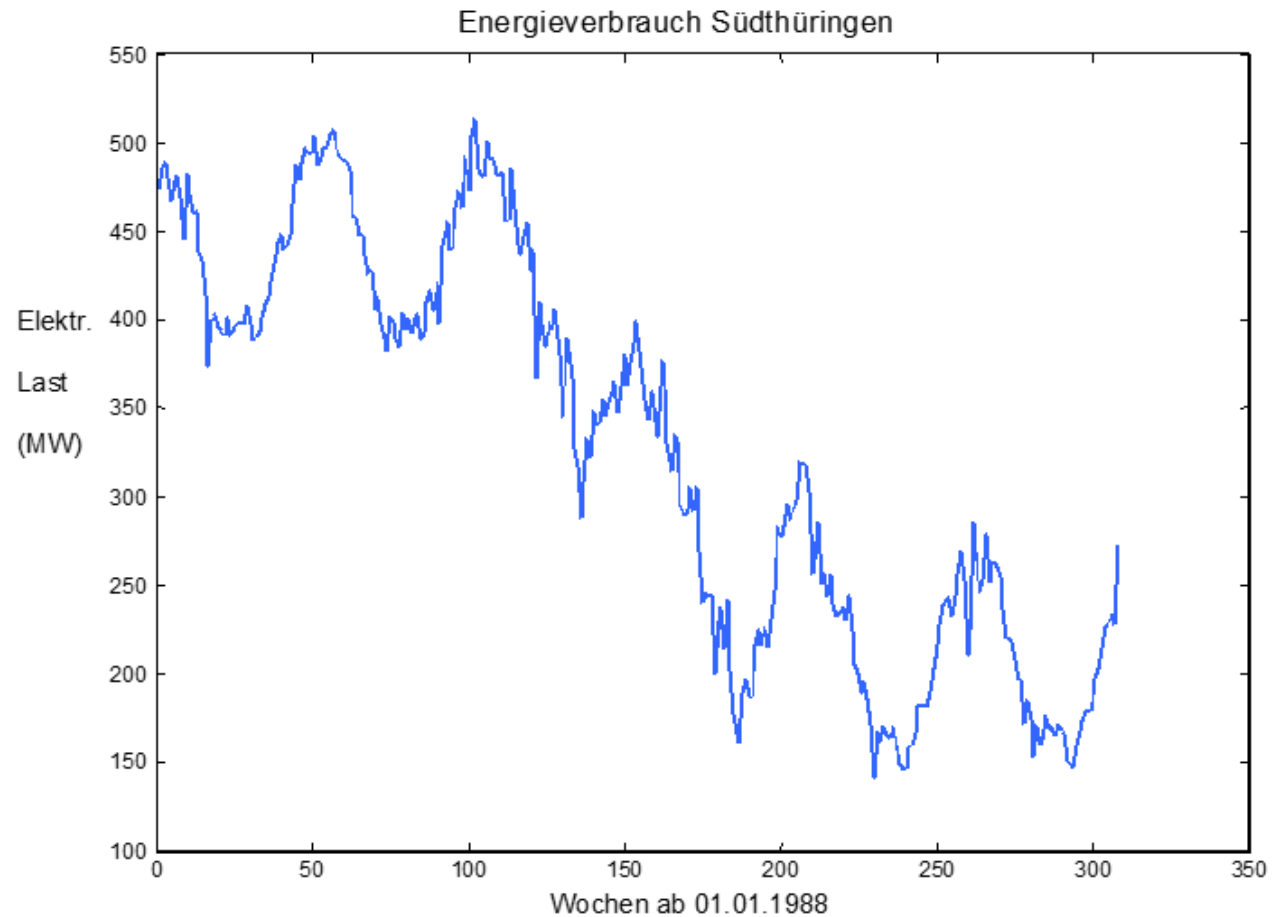


3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.2 Vorhersage mit deterministischen Signalmodellen

Elektrische Last in
Südthüringen

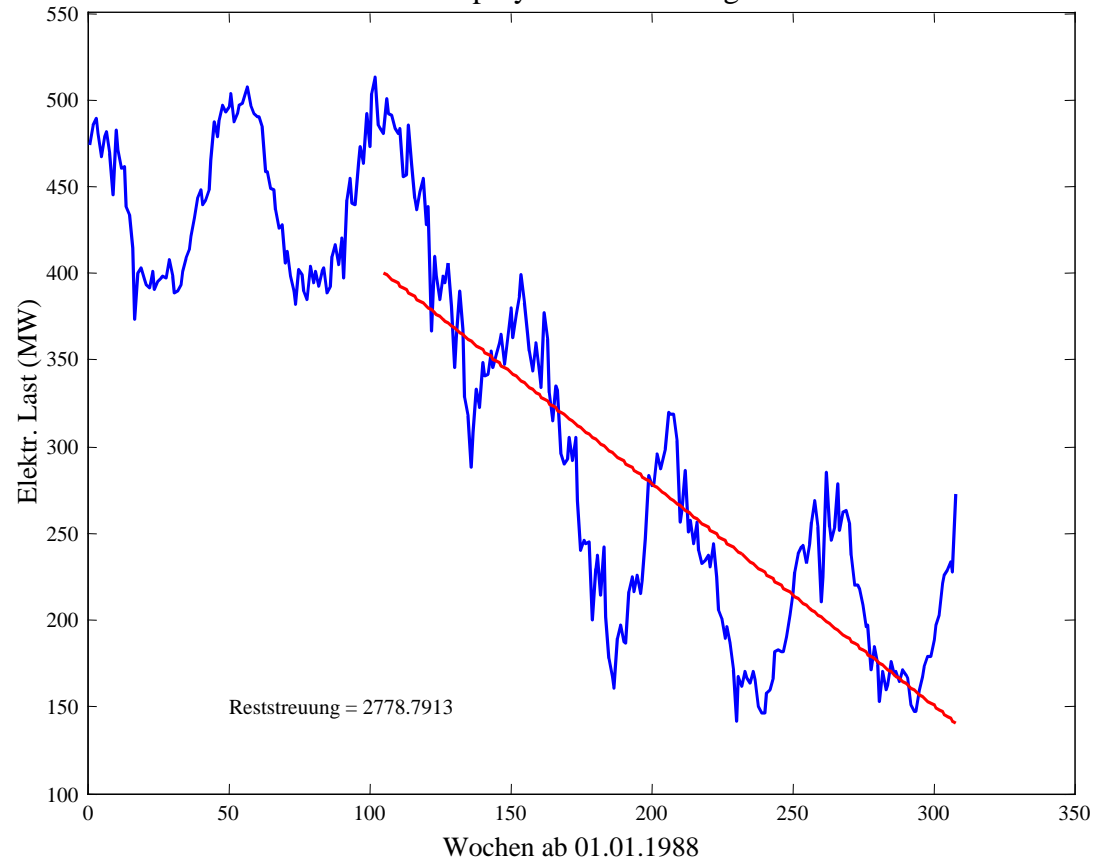


3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.2 Vorhersage mit deterministischen Signalmodellen

Trendpolynom 1. Ordnung



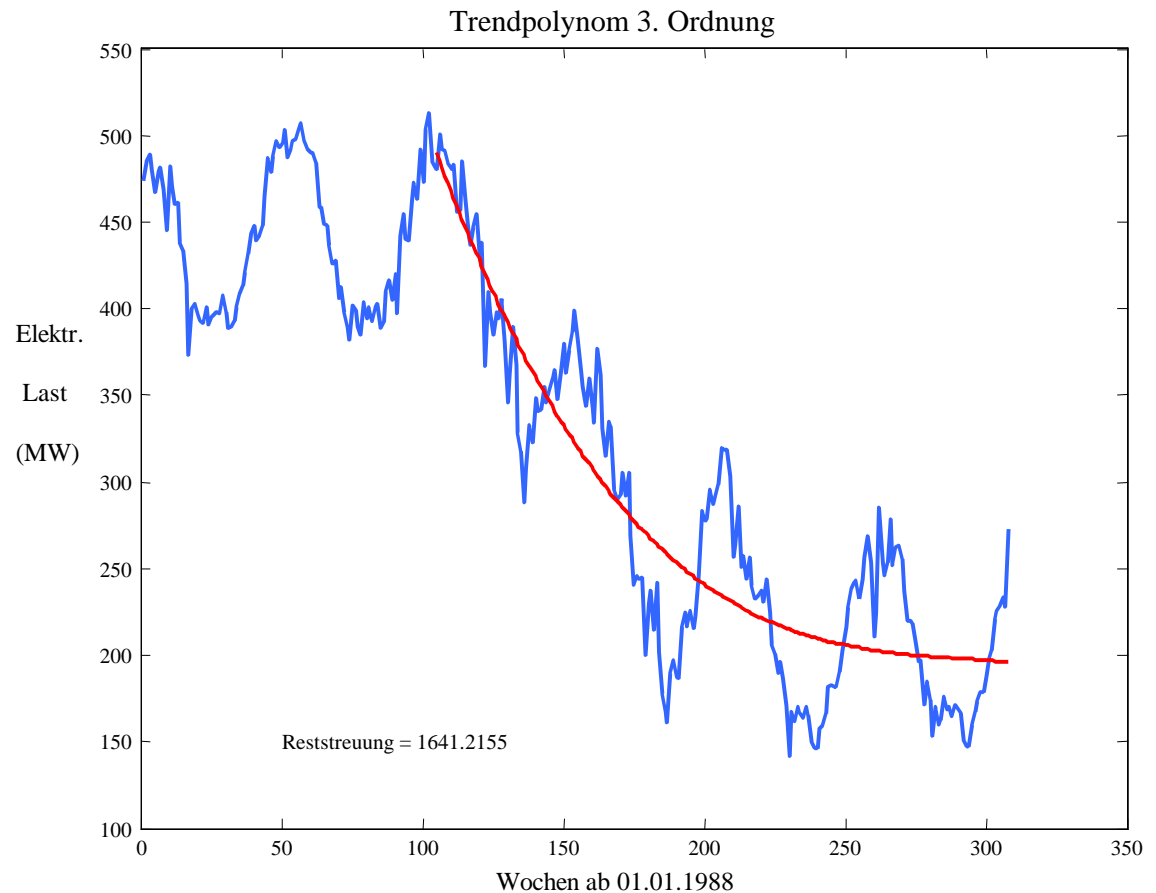
Trendbestimmung
(1)

3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.2 Vorhersage mit deterministischen Signalmodellen

Trendbestimmung
(3)

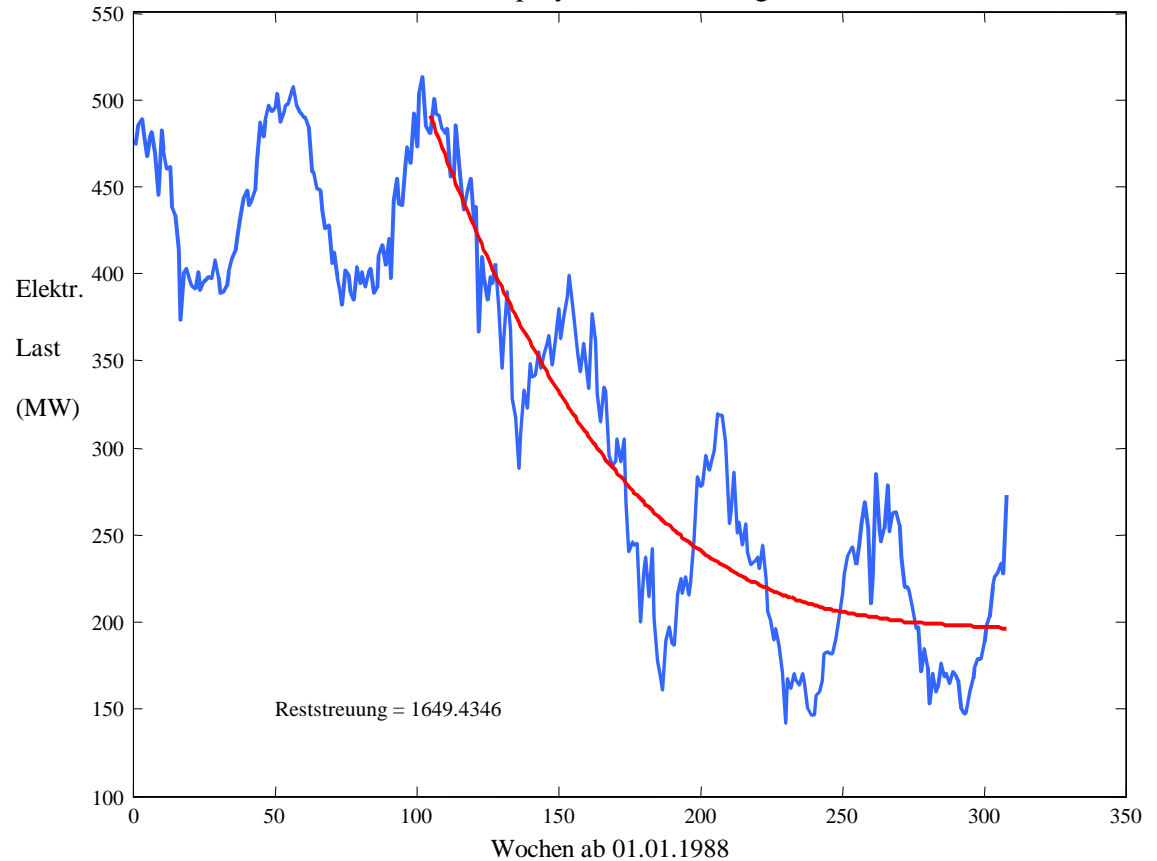


3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.2 Vorhersage mit deterministischen Signalmodellen

Trendpolynom 4. Ordnung



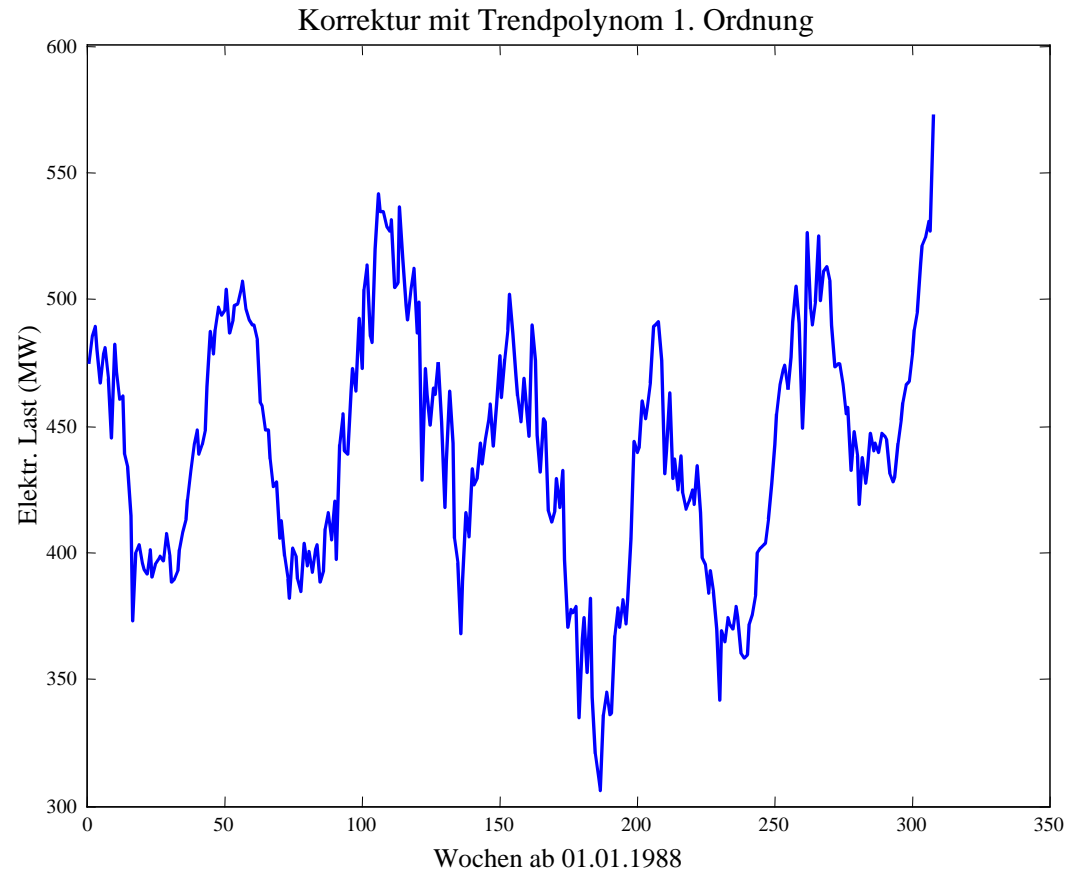
Trendbestimmung
(4)

3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.2 Vorhersage mit deterministischen Signalmodellen

Trendkorrektur (1)

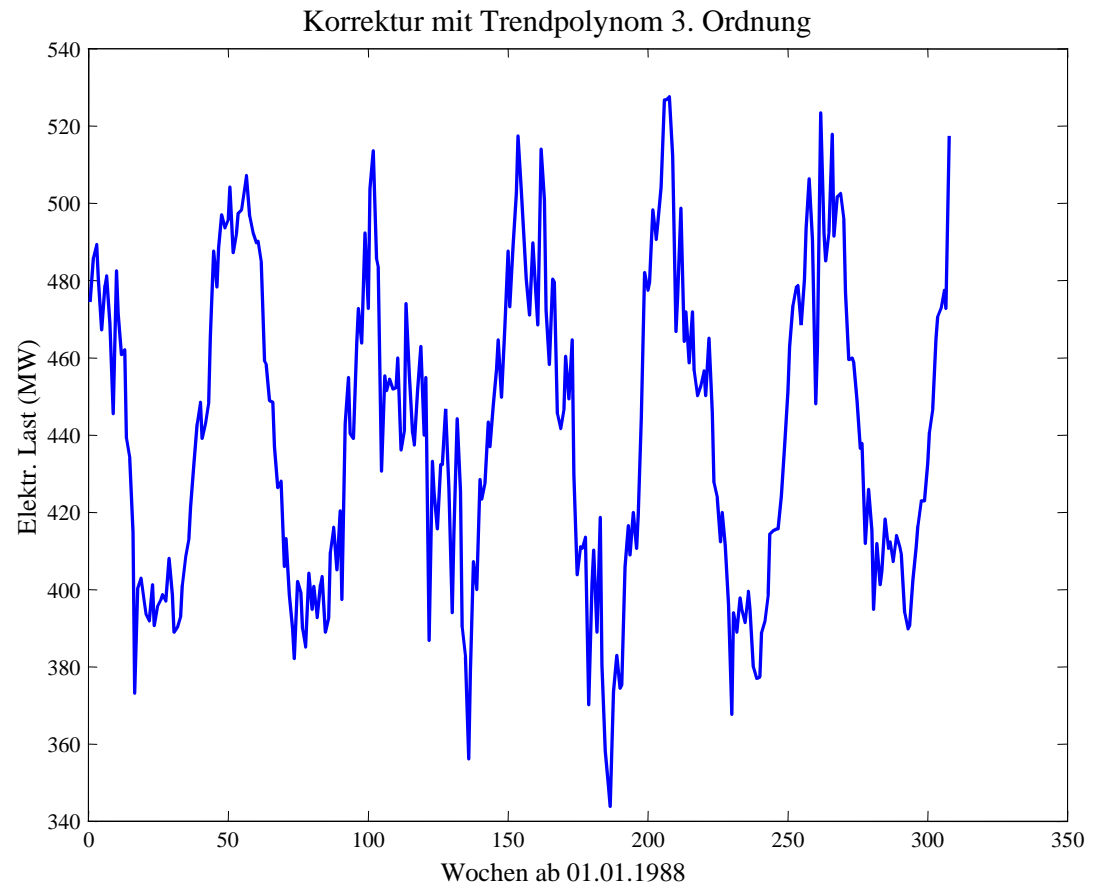


3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.2 Vorhersage mit deterministischen Signalmodellen

Trendkorrektur (2)

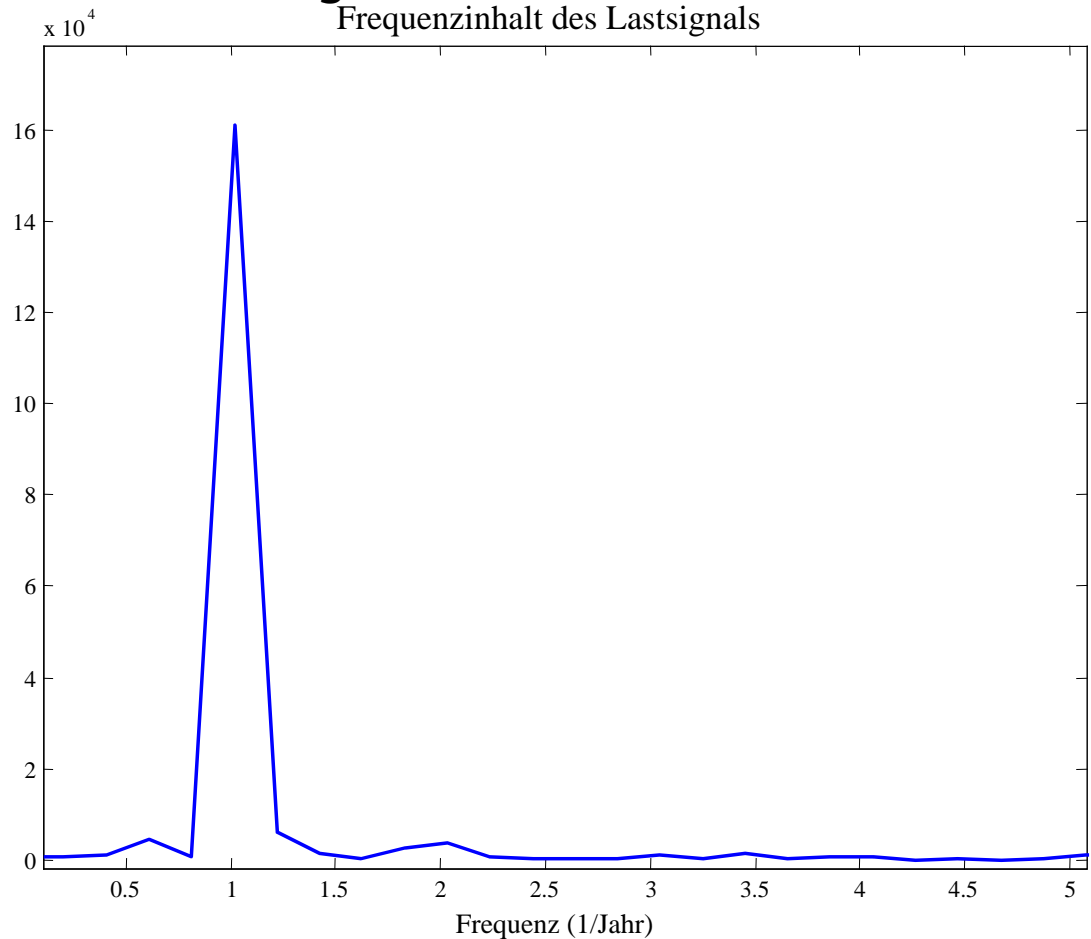


3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.2 Vorhersage mit deterministischen Signalmodellen

Leistungsdichtespektrum
des Lastsignals



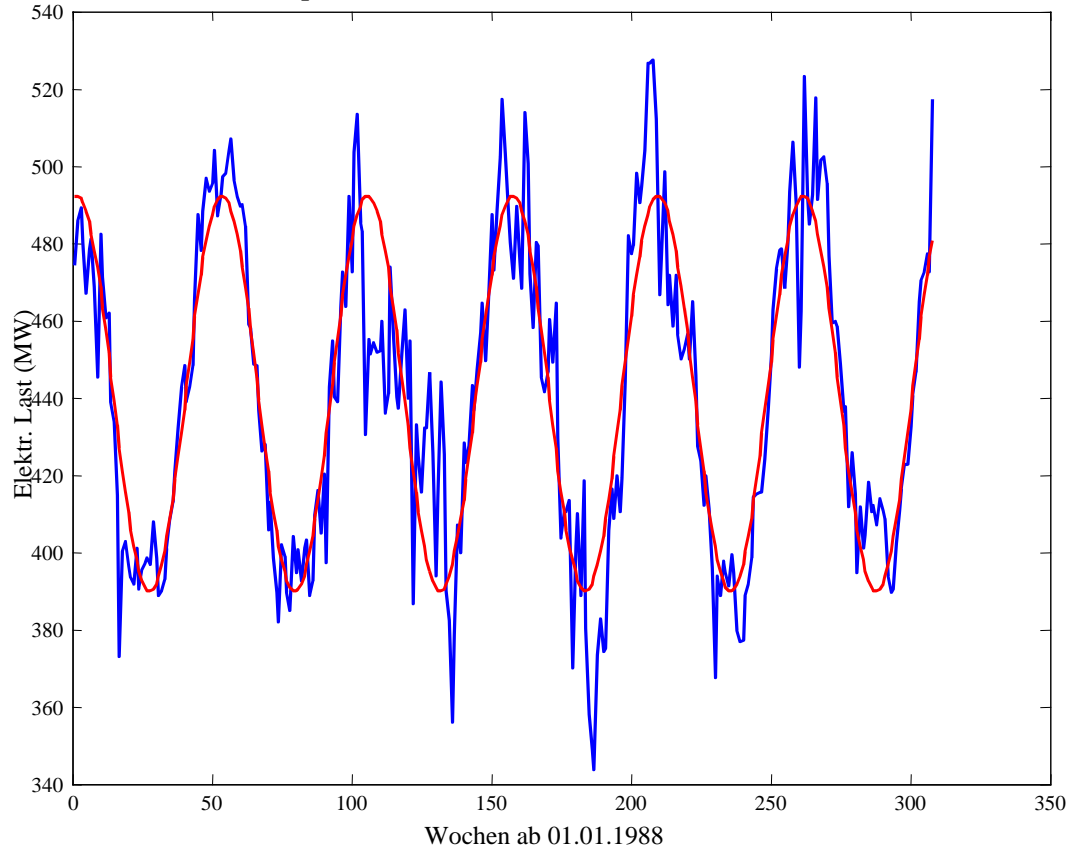
3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.2 Vorhersage mit deterministischen Signalmodellen

Periodischer Ansatz
für Beschreibung des
Lastsignals

Periodischer Modellansatz $x_p(kT) = 441.1206 + 10.4351 * \sin(\omega kT) + 50.1798 * \cos(\omega kT)$

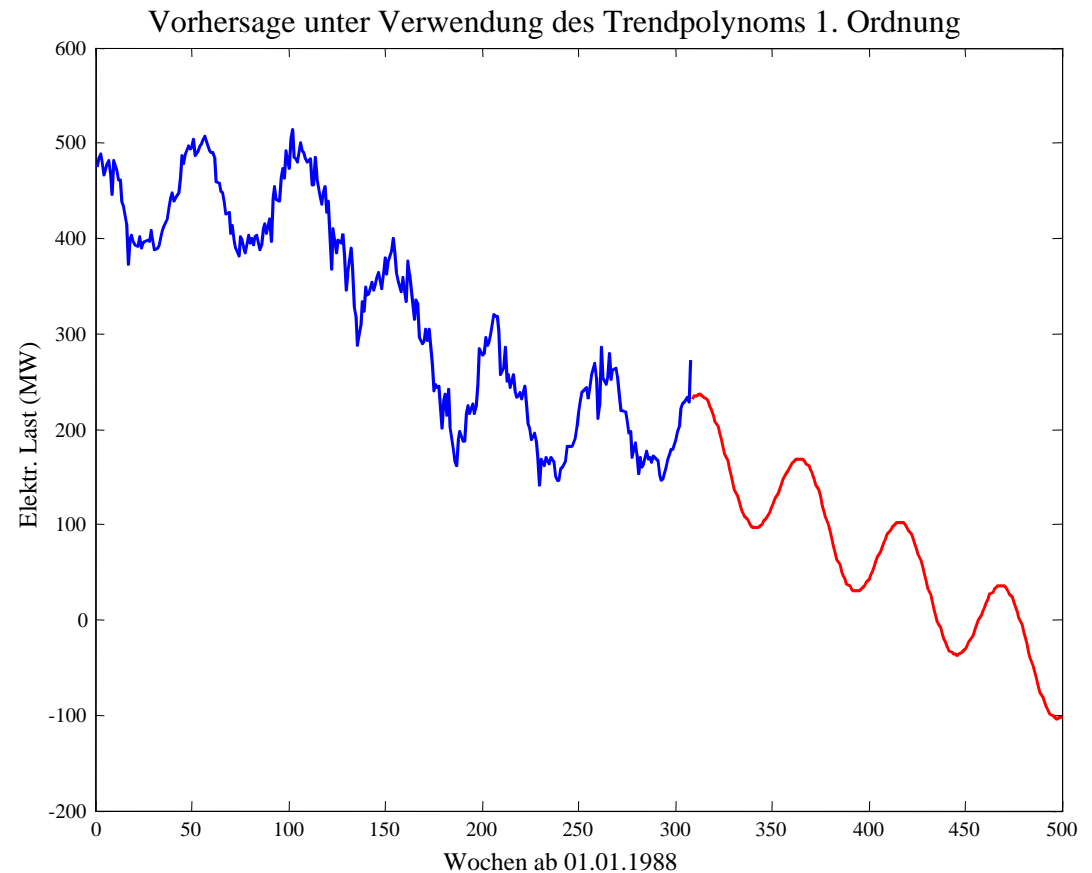


3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.2 Vorhersage mit deterministischen Signalmodellen

Vorhersage der Last
(1)

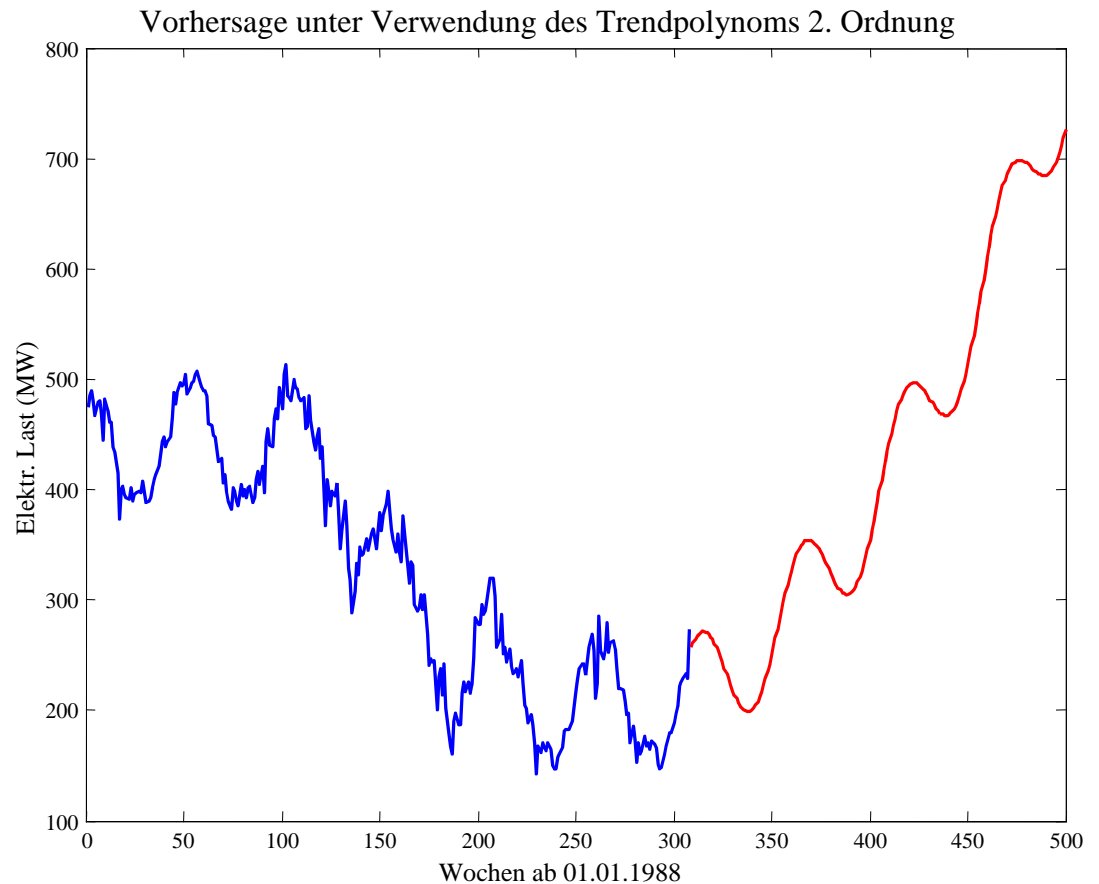


3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.2 Vorhersage mit deterministischen Signalmodellen

Vorhersage der Last
(2)

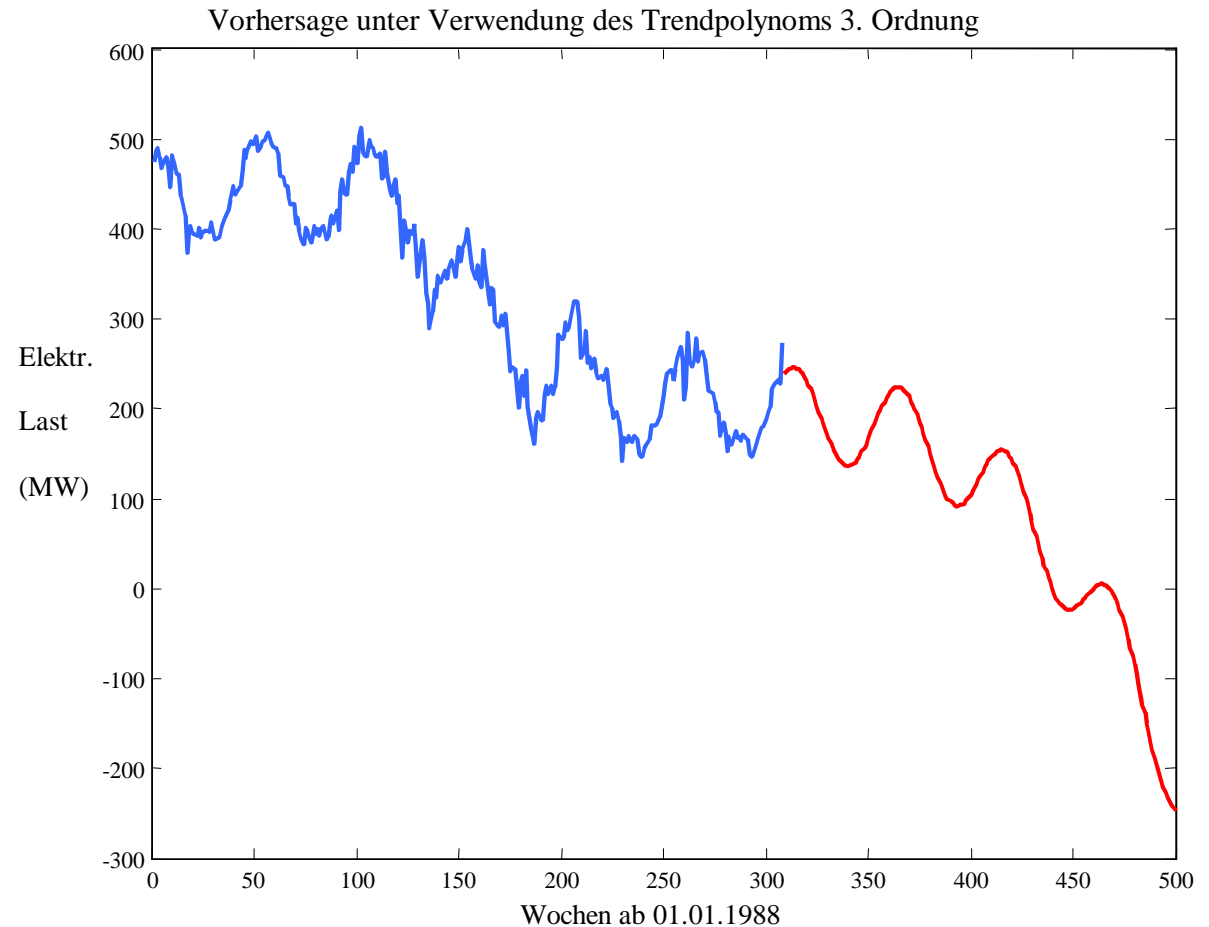


3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.2 Vorhersage mit deterministischen Signalmodellen

Vorhersage der Last
(3)



3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.3 Vorhersage mit stochastischen Signalmodellen – Optimale Prognose (1)

$\{x(k)\}$: Diese Zeitreihe ist ein Ausschnitt aus der Realisierung des stochastischen Prozesses $\{X(k)\}$ mit $k = 1 \dots N$.

- Prognose ist eine auf $\{x(k)\}$ basierende Schätzung $\hat{x}(k + h)$ des Wertes von $X(k + h)$, der zu selben Trajektorie (Realisierung des Prozesses) gehört.
- Die Güte einer Prognose wird stochastisch beurteilt, d. h. es erfolgt eine Mitteilung über alle Realisierungen des Prozesses.
- $\hat{x}(k + h)$ kann als Realisierung der Prognosefunktion $\hat{X}(k + h)$ betrachtet werden.
- $\hat{x}(k + h)$ ist eine Zufallsvariable, die von den Werten $x(1), x(2), \dots, x(N)$ abhängt.
- Als Gütemaß wird der mittlere quadratische Fehler (mean squared error) verwendet:

$$MSE[\hat{x}(k + h)] = E[(x(k + h) - \hat{x}(k + h))^2]$$

- **Optimal ist eine Prognosefunktion, die den MSE minimiert.**

3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.3 Vorhersage mit stochastischen Signalmodellen – Optimale Prognose (2)

- Die optimale Prognose von $x(k + h)$ durch einen (i. a. nichtlineare) Funktion $\hat{x}(k + h)$ von $x(1), x(2), \dots, x(k)$ ist der bedingte Erwartungswert:

$$\hat{x}(k + h) = E[x(k + h) | x(1), x(2), \dots, x(k)]$$

- Die konkrete Berechnung bedingter Erwartungswerte ist im Allgemeinen sehr schwierig, da man dazu die gemeinsame Verteilung der Prozessvariablen kennen muss.
- Es folgt daher die Beschränkung auf lineare Prognosen mit Prognosefunktion der Form

$$\hat{x}(k + h) = a_{k-1} \cdot x(1) + a_{k-2} \cdot x(2) + \dots + a_0 \cdot x(k)$$

- **Prognoseproblem: Lineare Prognose mit dem kleinsten MSE bestimmen; d. h. die Parameter a_i zu suchen, für die gilt:**

$$E \left[\left(x(k + h) - \sum_{u=0}^{k-1} a_u \cdot x(k - u) \right)^2 \right] = \min$$

3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.3 Vorhersage mit stochastischen Signalmodellen – Modellansatz des ARMA-Prozesses

Allgemeine Beschreibungsmöglichkeit für Signalprozesse in Form von ARMA (m,n)-Prozessen:

$$x(k) + a_1 \cdot x(k-1) + \dots + a_m x(k-m) = \varepsilon(k) + c_1 \cdot \varepsilon(k-1) + c_n \cdot \varepsilon(k-n)$$

Vereinfachte Darstellung durch Einführung des Verschiebeoperators q :

$$x(k) \left(1 + a_1 \cdot q^{-1} + \dots + a_m \cdot q^{-m} \right) = \varepsilon(k) \left(1 + c_1 \cdot q^{-1} + \dots + c_n \cdot q^{-n} \right)$$

$$x(k) \cdot A(q^{-1}) = \varepsilon(k) \cdot C(q^{-1})$$

$\varepsilon(k)$: Orthogonale stochastische Folge von Zufallszahlen $\{\varepsilon(k)\}$ mit den Eigenschaften:

a) $E\{\varepsilon(k)\} = 0$

b) $\text{cov}(\varepsilon(k), \varepsilon\{k+l\}) = E\{\varepsilon(k)\varepsilon(k+l)\}$

$$= \begin{cases} \sigma_\varepsilon^2 & \text{für } l = 0 \\ 0 & \text{für } l \neq 0 \end{cases}$$

$\{\varepsilon(k)\}$ wird als schwach stationär angenommen, d.

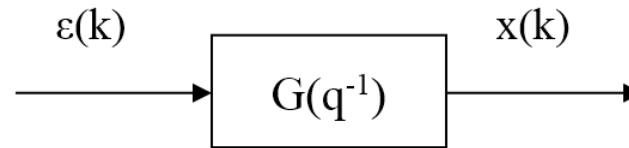
h. $\sigma_\varepsilon^2 < \infty$

3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.3 Vorhersage mit stochastischen Signalmodellen – MA-Darstellung eines ARMA-Prozesses (1)

$$x(k) = \frac{C(q^{-1})}{A(q^{-1})} \cdot \varepsilon(k) = G(q^{-1}) \cdot \varepsilon(k)$$



$$x(k) = \sum_{i=-\infty}^{+\infty} (g_i \cdot \varepsilon(k-i))$$

Liegen alle Nullstellen von $A(q^{-1})$ innerhalb des Einheitskreises, dann gilt:

a) $g_i = 0$ für $i < 0$

b)
$$x(k) = \lim_{l \rightarrow \infty} \left(\sum_{i=0}^l g_i \cdot \varepsilon(k-i) \right)$$

c) $x(k)$ ist schwach stationär.

Es gilt:

$$A(q^{-1}) \cdot G(q^{-1}) = C(q^{-1})$$

$$(1 + a_1 \cdot q^{-1} + \dots + a_m q^{-m})(g_0 + g_1 \cdot q^{-1} + \dots + g_r \cdot q^{-r} + \dots) = (1 + c_1 \cdot q^{-1} + \dots + c_n \cdot q^{-n})$$

Über Koeffizientenvergleich erhält man:

$$q^0: g_0 = 1$$

$$q^{-1}: g_1 + g_0 \cdot a_1 = c_1$$

$$q^{-2}: g_2 + g_1 + a_1 + g_0 \cdot a_2 = c_2$$

3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.3 Vorhersage mit stochastischen Signalmodellen – MA-Darstellung eines ARMA-Prozesses (2)

Allgemein gilt:

$$\sum_{i=0}^l g_{r-i} \cdot a_i = \begin{cases} c_r & \text{für } r \leq n \\ 0 & \text{für } r > n \end{cases}$$

mit $l = \min(m, r)$ und $a_0 = c_0 = 1$

Für MA-Prozess gilt: $m = 0$ und $A(q^{-1})$

$$g_r = \begin{cases} c_k & \text{für } r \leq n, c_0 = 1 \\ 0 & \text{für } r > n \end{cases}$$

Für AR-Prozess gilt: $n = 0$ und $C(q^{-1}) = 1$

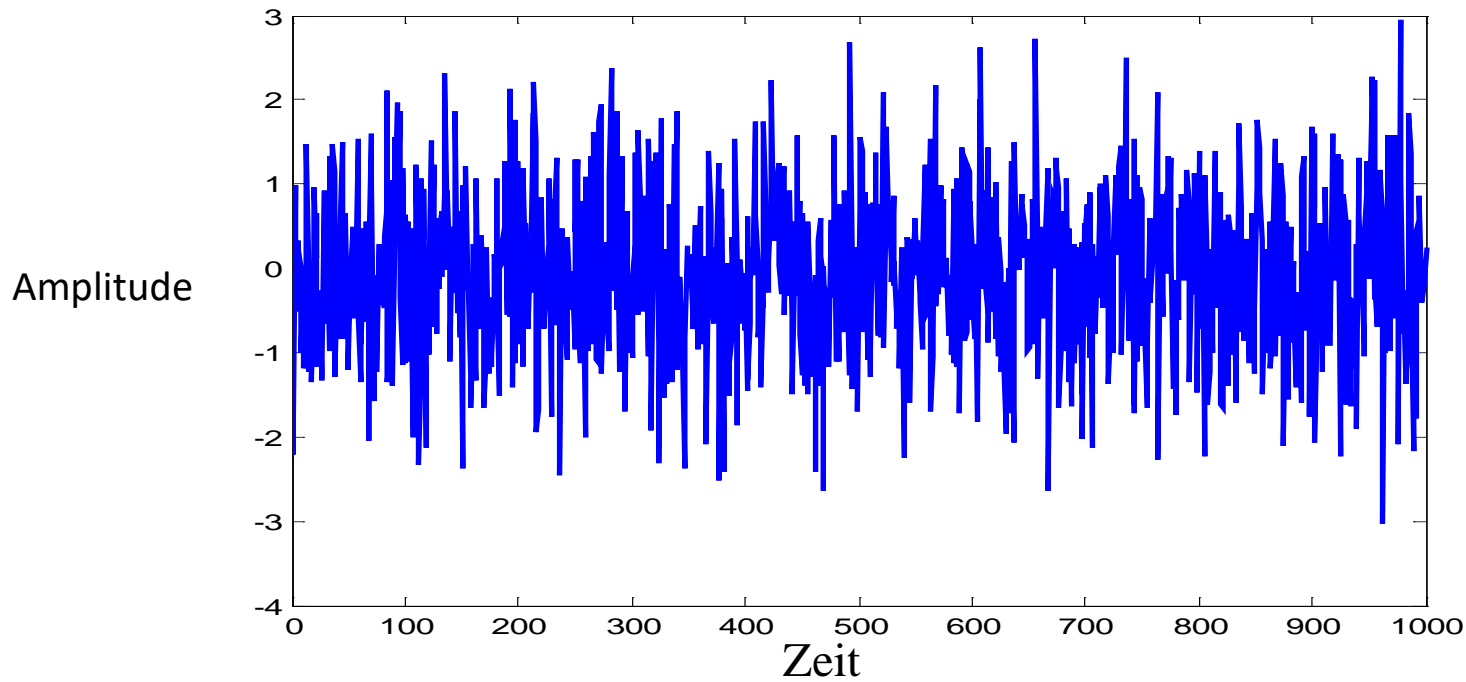
$$\sum_{i=0}^l g_{r-i} \cdot a_i = \begin{cases} 1 & \text{für } r = 0 \\ 0 & \text{für } r > 0 \end{cases}$$

3. VORHERSAGESTRATEGIEN

3.6 Vorhersage mit Signalmodellen

3.6.3 Vorhersage mit stochastischen Signalmodellen – MA-Darstellung eines ARMA-Prozesses (AR – Prozess mit Nullstellen im und außerhalb des Einheitskreises)

Weisses Rauschen

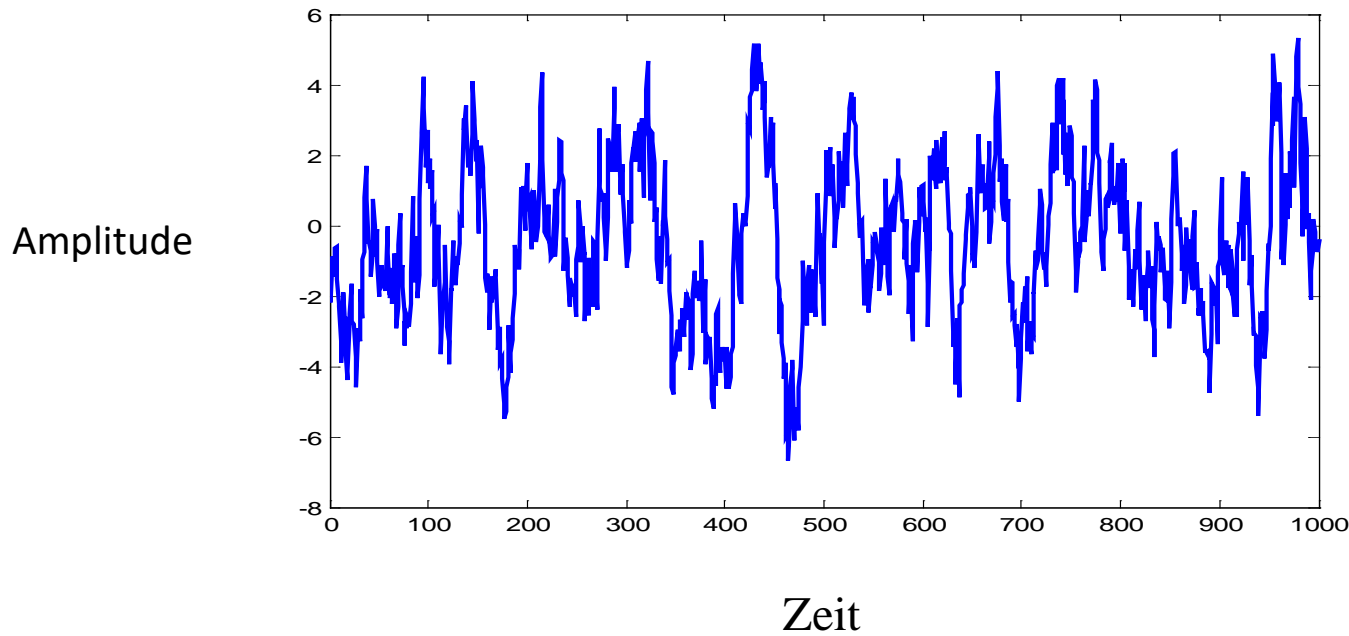


3. VORHERSAGESTRATEGIEN

3.6 Vorhersage mit Signalmodellen

3.6.3 Vorhersage mit stochastischen Signalmodellen –
MA-Darstellung eines ARMA-Prozesses
(AR – Prozess mit Nullstellen im und außerhalb des
Einheitskreises)

$$x(k) = 0.8x(k-1) + 0.2x(k-2) + \varepsilon(k) \text{ – Nullstellen bei } 1 \text{ und } -0.2$$



3. VORHERSAGESTRATEGIEN

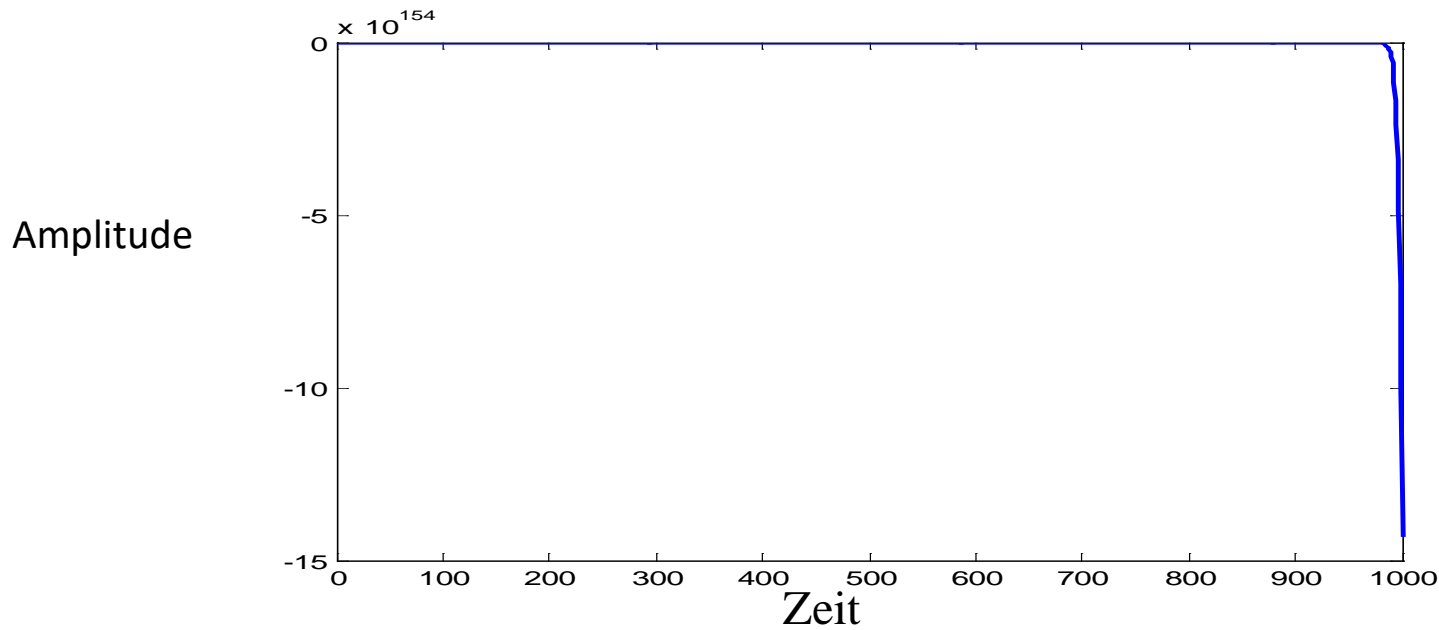
3.6 Vorhersage mit Signalmodellen

3.6.3 Vorhersage mit stochastischen Signalmodellen –

MA-Darstellung eines ARMA-Prozesses

(AR – Prozess mit Nullstellen im und außerhalb des Einheitskreises)

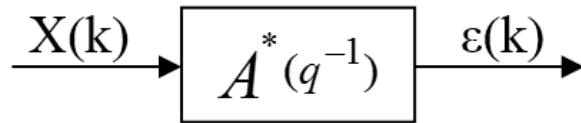
$$x(k) = 0.8x(k-1) + 0.9x(k-2) + \varepsilon(k) \text{ – Nullstellen bei } 1.42 \text{ und } -0.63$$



3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.3 Vorhersage mit stochastischen Signalmodellen – AR-Darstellung eines ARMA-Prozesses (1)



$$\varepsilon(k) = A^*(q^{-1}) \cdot x(k)$$

$$\varepsilon(k) = \sum_{i=-\infty}^{+\infty} a_i^* \cdot x_{k-i}$$

Unter Verwendung von $x(k) = G(q^{-1}) \cdot \varepsilon(k)$ gilt:

$$x(k) = G(q^{-1}) \cdot A^*(q^{-1}) \cdot x(k)$$

$$G(q^{-1}) \cdot A^*(q^{-1}) = 1$$

$A^*(q^{-1})$ ist die inverse Folge von $G(q^{-1})$.

$$G(q^{-1}) = \frac{C(q^{-1})}{A(q^{-1})} \quad \curvearrowright \quad A(q^{-1}) = \frac{C(q^{-1})}{G(q^{-1})} = C(q^{-1}) \cdot A^*(q^{-1})$$

$$(1 + a_1 \cdot q^{-1} + \dots + a_m q^{-m}) = (1 + c_1 q^{-1} + \dots + c_n q^{-n}) \cdot (1 + a_1^* q^1 + \dots + a_r^* q^{-r} + \dots)$$

Über Koeffizientenvergleich erhält man:

$$q^0: 1 = 1$$

$$q^{-1}: a_1 = c_1 + a_1^*$$

$$q^{-2}: a_2 = c_2 + c_1 \cdot a_1^* + a_2^*$$

3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.3 Vorhersage mit stochastischen Signalmodellen – AR-Darstellung eines ARMA-Prozesses (2)

Allgemein gilt:

$$\sum_{i=0}^l c_i \cdot \mathbf{a}_{r-i}^* = \begin{cases} a_k & \text{für } r \leq m \\ 0 & \text{für } r > m \end{cases}$$

mit $l = \min(n, r)$ und $a_0 = c_0 = 1$

Für MA(n)-Prozess gilt:

$$m = 0 \text{ und } A(q^{-1}) = 1$$

$$\sum_{i=0}^l c_i \cdot \mathbf{a}_{r-i}^* = \begin{cases} 1 & \text{für } r = 0 \\ 0 & \text{für } r > 0 \end{cases}$$

Für AR(m)-Prozess gilt:

$$n = 0 \text{ und } C(q^{-1}) = 1$$

$$\mathbf{a}_r^* = \begin{cases} a_k & \text{für } r \leq m, a_0 = 1 \\ 0 & \text{für } r > m \end{cases}$$

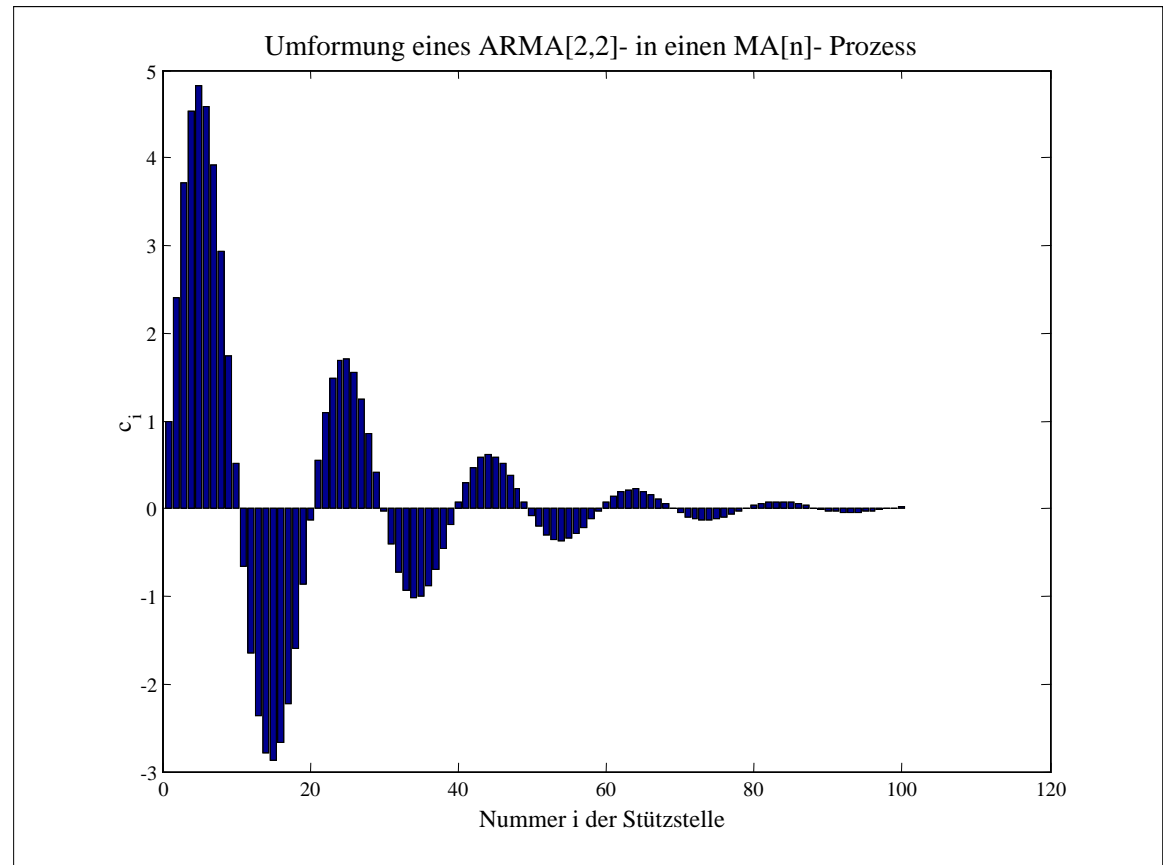
3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.3 Vorhersage mit stochastischen Signalmodellen – Umformung eines ARMA-Prozesses in einen MA-Prozess

ARMA [2,2]-Prozess:

$$x(k) - 1,8x(k-1) + 0,9x(k-2) = \varepsilon(k) + 0,6\varepsilon(k-1) + 0,3\varepsilon(k-2)$$



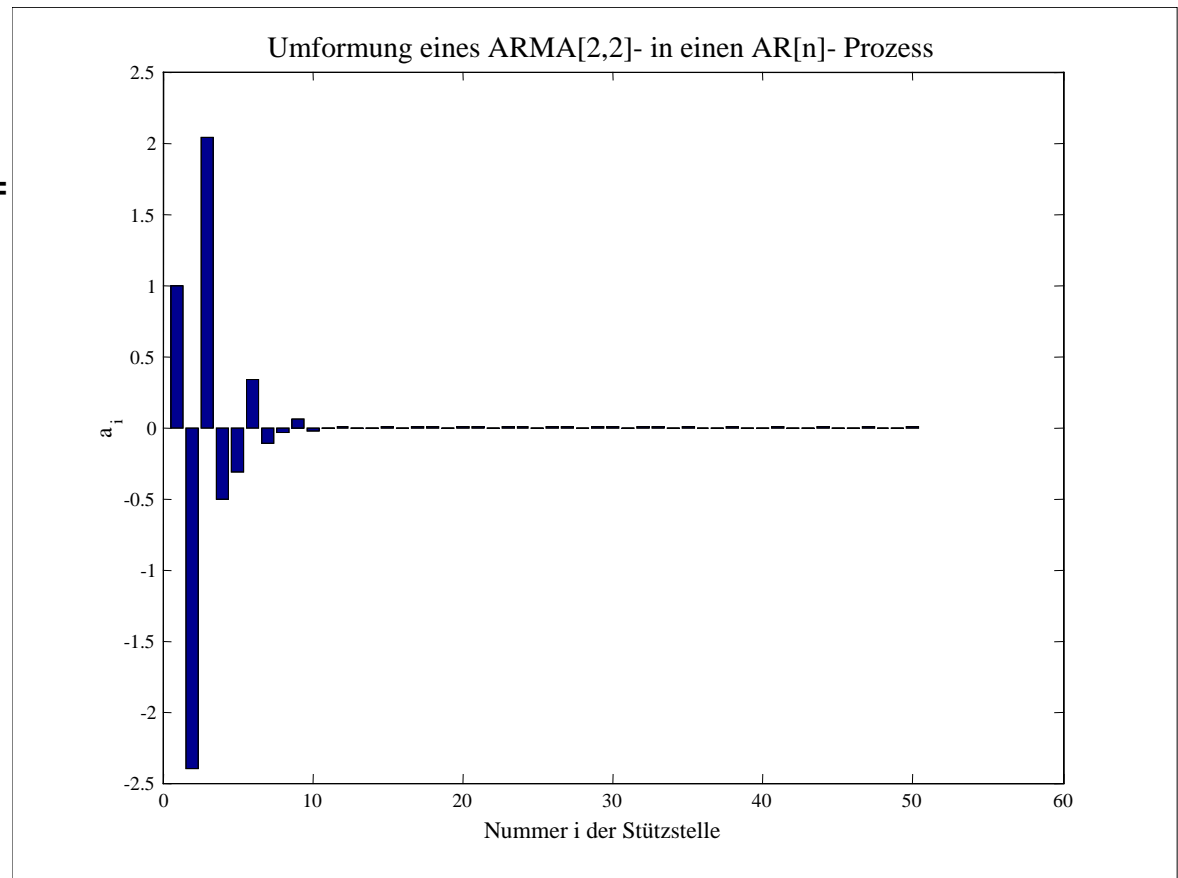
3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.3 Vorhersage mit stochastischen Signalmodellen – Umformung eines ARMA-Prozesses in einen AR-Prozess

ARMA [2,2]-Prozess:

$$x(k) - 1,8x(k-1) + 0,9x(k-2) = \varepsilon(k) + 0,6\varepsilon(k-1) + 0,3\varepsilon(k-2)$$



3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.3 Vorhersage mit stochastischen Signalmodellen – Stochastischer Signalprozess und Autokorrelationsfunktion (1.1)

ARMA(m,n)-Prozess: $x(k) + a_1 \cdot x(k-1) + \dots + a_m x(k-m) = \varepsilon(k) + c_1 \cdot \varepsilon(k-1) + \dots + c_n \cdot \varepsilon(k-n)$

Multiplikation beider Seiten mit $x(k-s)$ $s \in [0, m]$

$$x(k)x(k-s) + a_1 x(k)x(k-s) + \dots + a_m x(k-m)x(k-s) =$$

$$\varepsilon(k) \cdot x(k-s) + c_1 \cdot \varepsilon(k-1)x(k-s) + \dots + c_n \cdot \varepsilon(k-n)x(k-s)$$

Wenn $\{x(k)\}$ schwach stationär, gilt für die Erwartungswerte:

Linker Term:

$$\sum_{i=0}^m a_i \cdot \text{cov}(x(k-i), x(k-s)) = \sum_{i=0}^m a_i \cdot r(i-s)$$

$$E\{x(k-i)x(k-s)\} = \text{cov}(x(k-i), x(k-s))$$

3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.3 Vorhersage mit stochastischen Signalmodellen – Stochastischer Signalprozess und Autokorrelationsfunktion (1.2)


Rechter Term:

$$x(k) = \frac{C(q^{-1})}{A(q^{-1})} \cdot \varepsilon(k) = G(q^{-1}) \cdot \varepsilon(k)$$

$$E\{x(k-s)\varepsilon(k-l)\} = E\left\{\sum_{i=0}^n (g_i \cdot \varepsilon(k-s-i))\varepsilon(k-l)\right\}$$

$$\text{mit } E\{\varepsilon(s)\varepsilon(l)\} = \begin{cases} \sigma_\varepsilon^2 & \text{für } s = l \\ 0 & \text{für } s \neq l \end{cases} \text{ gilt:}$$

$$E\{x(k-s)\varepsilon(k-l)\} = g_{l-s} \cdot \sigma_\varepsilon^2$$

 für rechten Term: $\sigma_\varepsilon^2 \cdot \sum_{i=0}^n c_i \cdot g_{i-s}$

3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.3 Vorhersage mit stochastischen Signalmodellen – Stochastischer Signalprozess und Autokorrelationsfunktion (2)

Insgesamt gilt somit:

$$\sum_{i=0}^m a_i \cdot r(s-i) = \begin{cases} \sigma_\varepsilon^2 \cdot \sum_{i=0}^n c_i \cdot g_{i-s} & \text{für } s \leq n \\ 0 & \text{für } s > n \end{cases}$$

Für $s > n$ erhält man die Yule-Walker-Gleichungen:

$$r(s) + a_1 \cdot r(s-1) + \dots + a_m r(s-m) = 0 \quad (\text{Autokovarianzfunktion})$$

bzw. normiert

$$R(s) + q_1 \cdot R(s-1) + \dots + a_m R(s-m) = 0 \quad (\text{Autokorrelationsfunktion})$$

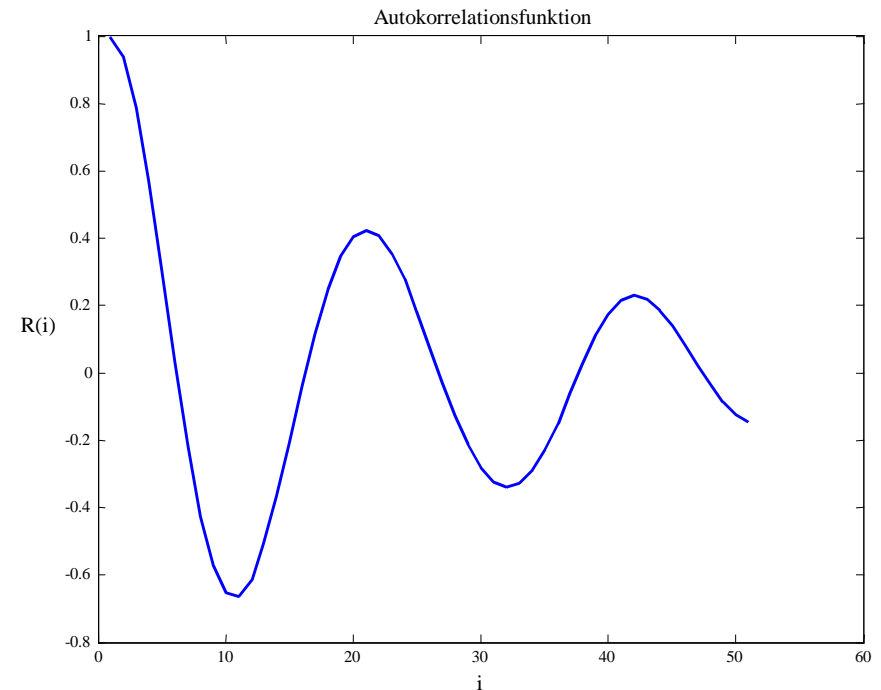
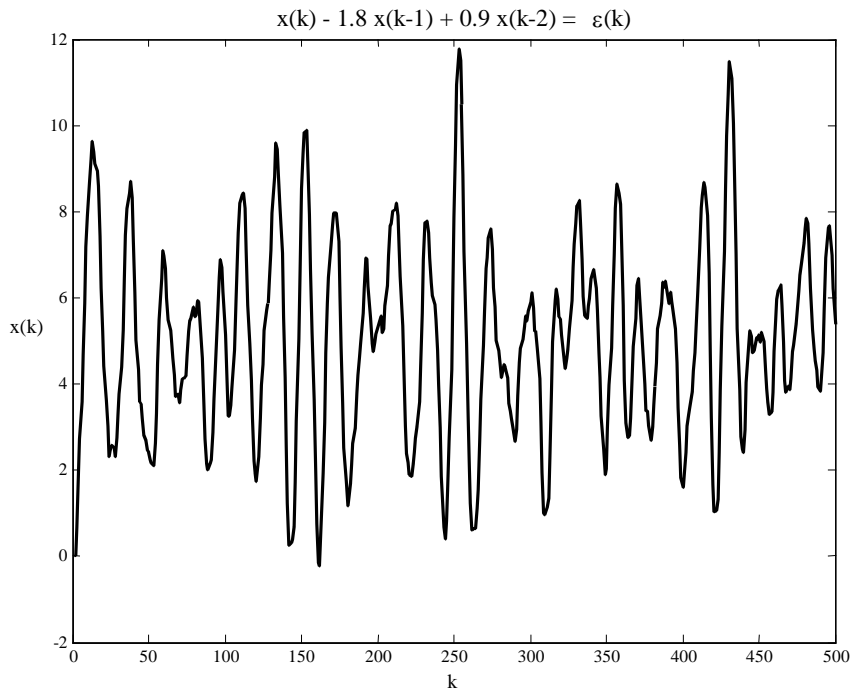
Folgende wichtige Aussagen sind daraus ableitbar:

1. Bei einem AR(m)-Prozess ergeben die Lösungen der Yule-Walker-Gleichungen je nach Art der Nullstellen gedämpfte reine Exponentialfunktionen oder mit Sinus- und / oder Potenzfunktionen gemischte. Ist der AR(m)-Prozess stabil, sind die Autokovarianz- bzw. die Autokorrelationsfunktion gedämpft, also abklingend.
2. Bei einem MA(n)-Prozess bricht die Autokovarianz- bzw. Autokorrelationsfunktion nach dem n-ten Schritt ab.

3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

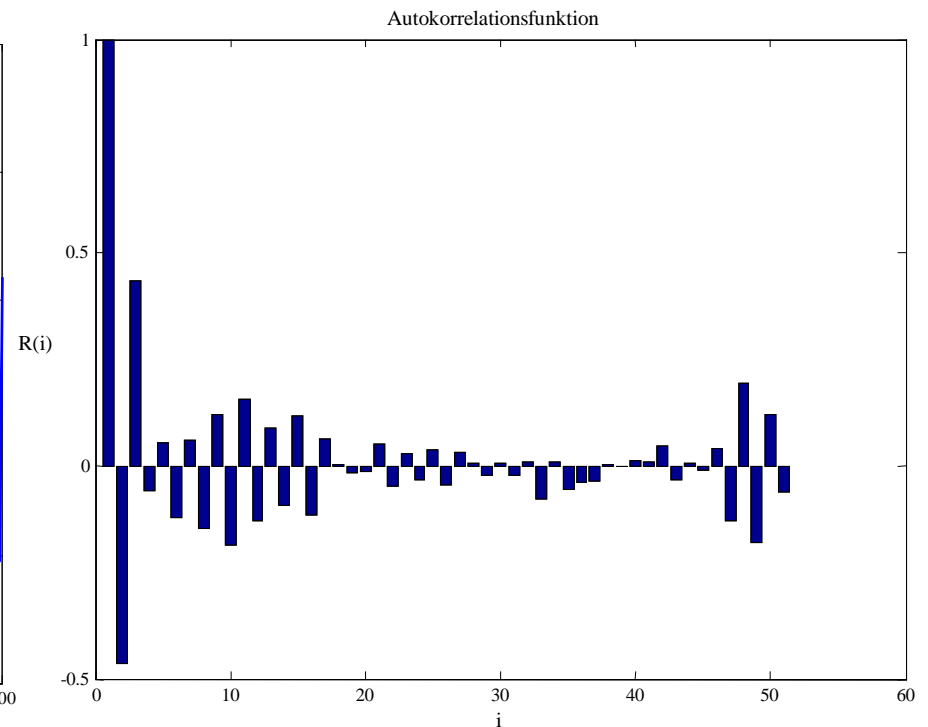
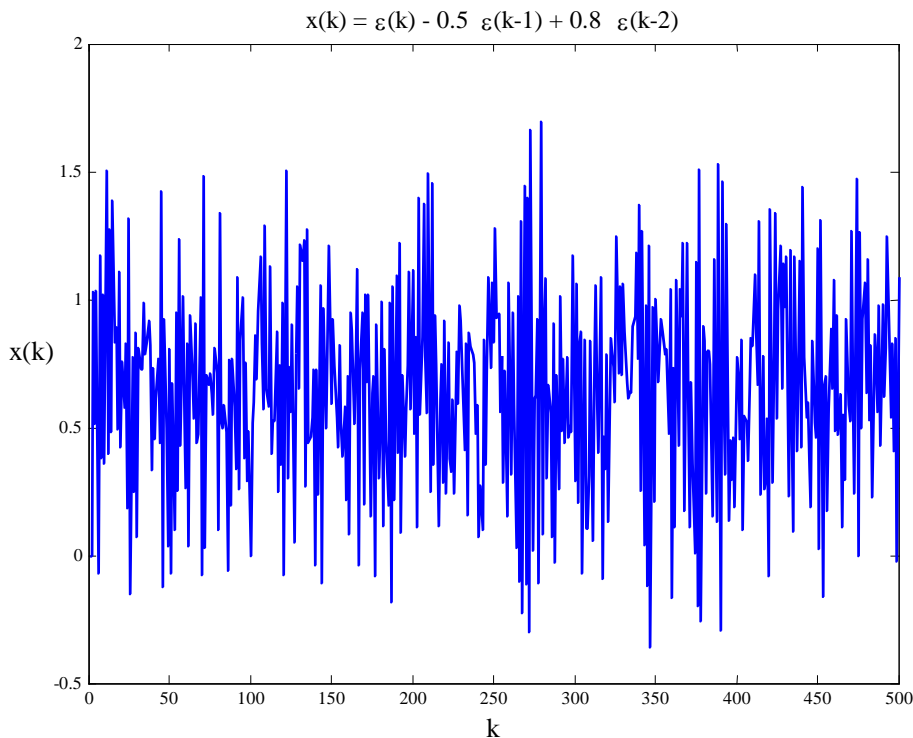
3.6.3 Vorhersage mit stochastischen Signalmodellen – AR (2) – Prozess und seine Autokorrelationsfunktion



3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.3 Vorhersage mit stochastischen Signalmodellen – MA (2) – Prozess und seine Autokorrelationsfunktion



3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.3 Vorhersage mit stochastischen Signalmodellen – Stochastischer Signalprozess und partielle Autokorrelationsfunktion (1)

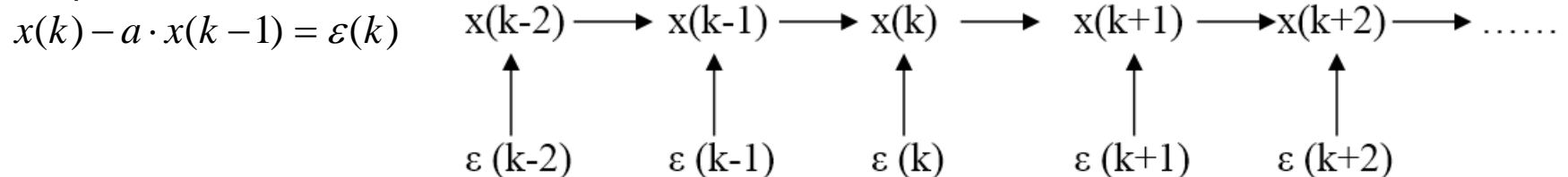
Partielle Autokorrelationen sind auf das engste mit dem Komplex der Prognose verknüpft. Dadurch wird ermittelt, wie gut $x(k)$ durch eine lineare Funktion der vorausgehenden Zufallsvariablen $x(k-1), x(k-2), \dots$ beschrieben werden kann.

Definition der partiellen Autokorrelation:

$\{x(k)\}$ sei ein stationärer Prozess. Die partielle Autokorrelation $PR(\tau)$ ist die partielle Korrelation von $x(k)$ und $x(k-\tau)$ unter Konstanthaltung der dazwischen liegenden Zufallsvariablen $x(u)$ mit $t-\tau < u < t$.

Es gilt: $PR(0) = 1$ und $PR(-\tau) = PR(\tau)$

Beispiel: AR(1)-Prozess



Wird $PR(2)$ von $x(k)$ und $x(k-2)$ unter Konstanthaltung von $x(k-1)$ bestimmt, verschwindet diese Korrelation, da alle Beziehungen über $x(k-1)$ vermittelt werden. Gleiches gilt für $PR(\tau)$ mit $\tau > 2$. Somit erhält man:

$$PR(0) = 1$$
$$PR(1) = a \text{ und } PR(\tau) = 0 \text{ f\u00fcr } \tau > 1$$

3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.3 Vorhersage mit stochastischen Signalmodellen – Stochastischer Signalprozess und partielle Autokorrelationsfunktion (2)

Berechnung der partiellen Autokorrelationsfunktion (PAKF):

$$PR(s, j) = PR(s-1, j) + PR(s-1, s-j) \cdot PR(s, s) \quad \begin{array}{l} R - \text{Autokorrelationsfunktion} \\ j = 1, 2, \dots, s-1 \\ s = 1, 2, \dots \text{ (Berechnungstiefe)} \\ PR(s, 0) = 1 \end{array}$$

$$PR(s, s) = \frac{\sum_{i=0}^{s-1} R(s-i) \cdot PR(s-1, i)}{\sum_{i=0}^{s-1} R(i) \cdot PR(s-1, i)}$$

Autoregressiven Modellansatz:

$$\sum_{i=0}^m a_i x(k-i) = \varepsilon(k) \quad \text{mit } a_0 = 1$$

Beziehung zwischen a_i und PR :

$$a_i = PR(m, i) \quad (\text{m-Ordnung des AR-Ansatzes})$$

$PR(m, m)^2$ stellt ein relatives Maß dar, um den die Unbestimmtheit eines autoregressiven Modells verringert wird, falls von einem AR(m-1)-Modell zu einem AR(m)-Modell übergegangen wird.

3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.3 Vorhersage mit stochastischen Signalmodellen – Stochastischer Signalprozess und partielle Autokorrelationsfunktion (3)

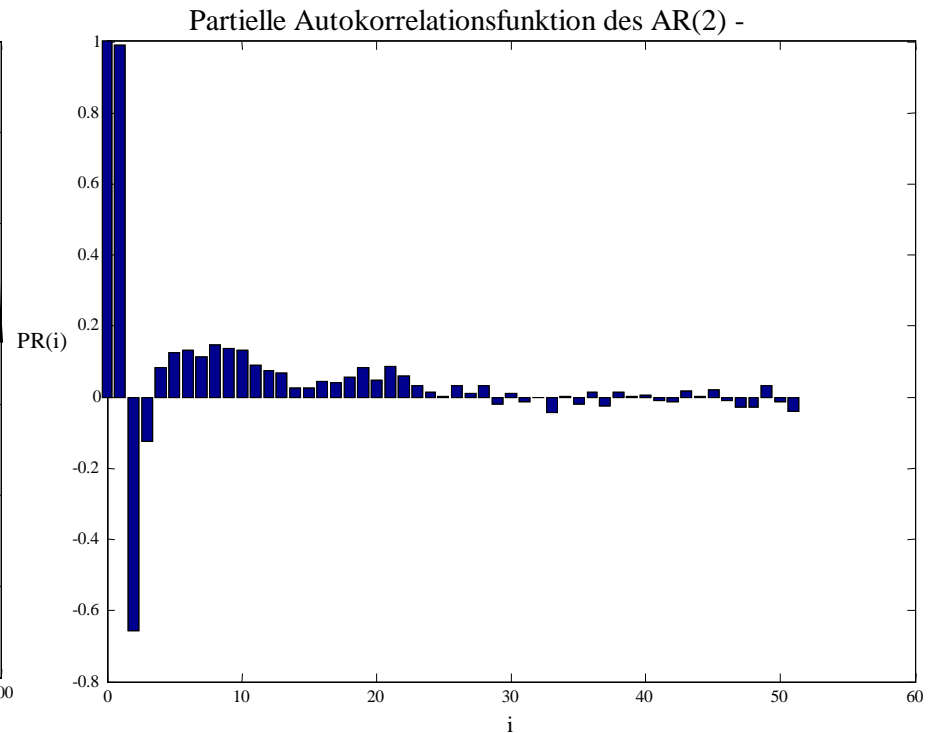
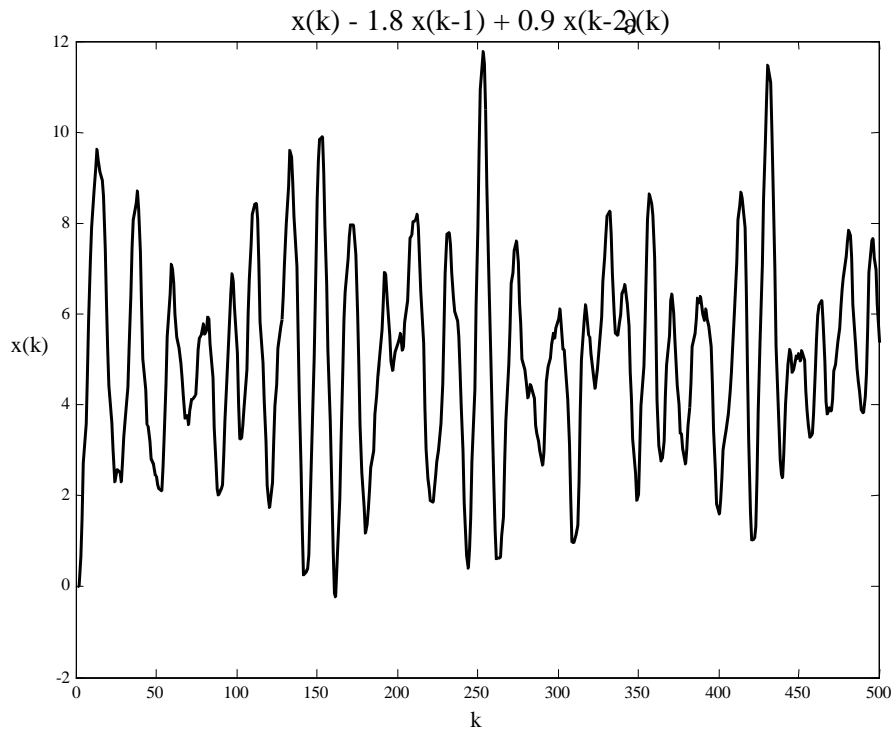
Daraus leiten sich folgende wichtige Aussagen für die Identifikation eines Vorhersagemodells ab:

- Bei einem $AR(m)$ -Prozess bricht die partielle Autokorrelationsfunktion nach genau m Schritten ab.
- Für einen $MA(n)$ -Prozess entstehen gedämpfte Exponentialfunktionen oder mit Sinus- und / oder Potenzfunktionen gemischte Funktionen.

3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

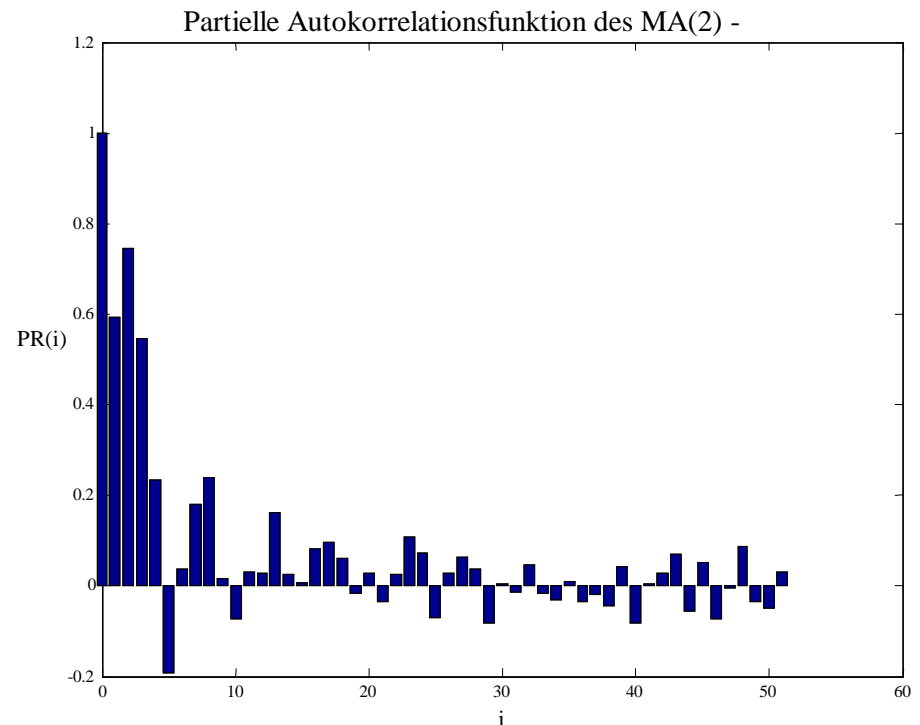
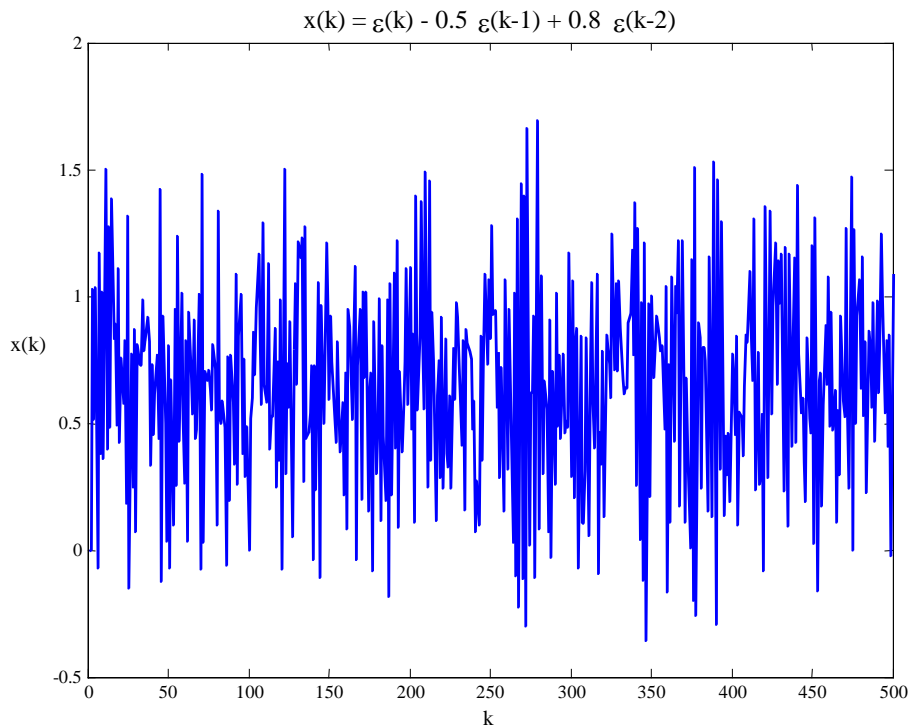
3.6.3 Vorhersage mit stochastischen Signalmodellen – AR(2)-Prozess und seine partielle Autokorrelationsfunktion



3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.3 Vorhersage mit stochastischen Signalmodellen – MA(2)-Prozess und seine partielle Autokorrelationsfunktion



3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.3 Vorhersage mit stochastischen Signalmodellen – Stochastischer Signalprozess und inverse Autokorrelationsfunktion(1)

Darstellung von $\{x(k)\}$ als AR(∞)-Prozess: $A(q^{-1})x(k) = \varepsilon(k)$

Der dazu inverse Prozess $\{x_{inv}(k)\}$ wird folgendermaßen beschrieben:

$$\begin{aligned}x_{inv}(k) &= A(q^{-1}) \cdot \varepsilon(k) \\ &= A(q^{-1}) \cdot A(q^{-1}) \cdot x(k)\end{aligned}$$

Beispiel: AR(1)-Prozess

$$x(k) + a_1x(k-1) = \varepsilon(k)$$

$$A(q^{-1}) = 1 + a_1 \cdot q^{-1}$$

$$A(q^{-1}) \cdot A(q^{-1}) = (1 + a_1q^{-1})(1 + a_1q^{-1})$$

$$= 1 + 2a_1q^{-1} + a_1^2q^{-2}$$

$$x_{inv}(k) = (1 + 2a_1q^{-1} + a_1^2q^{-2})x(k)$$

$$= x(k) + 2a_1x(k-1) + a_1^2x(k-2)$$

$$= \varepsilon(k) + a_1x(k-1) + a_1^2x(k-2)$$

$$= \varepsilon(k) + a_1(x(k-1) + a_1x(-2))$$

$$= \varepsilon(k) + a_1\varepsilon(k-1)$$

Die inverse Autokorrelationsfunktion (IAKF) $IR(\tau)$ von $\{x(k)\}$ ist die Autokorrelationsfunktion von $\{x_{inv}(k)\}$.

3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.3 Vorhersage mit stochastischen Signalmodellen – Stochastischer Signalprozess und inverse Autokorrelationsfunktion(2)

Die Berechnung der IAKF unter Verwendung der PAKF:

$$IR(s, l) \begin{cases} \frac{\sum_{j=l}^s PR(s, j) \cdot PR(s, j-l)}{\sum_{j=0}^s PR(s, j)^2} & \text{für } l = 0, 1, \dots, s \\ 0 & \text{für } l > s \end{cases}$$

Vorteile der IAKF: Exakte Dualität zur Autokorrelationsfunktion.

Daraus lassen sich besser theoretische Aussagen herleiten als aus der partiellen Autokorrelationsfunktion.

Wichtige Aussagen aus IAKF zur Identifikation von Vorhersagemodellen:

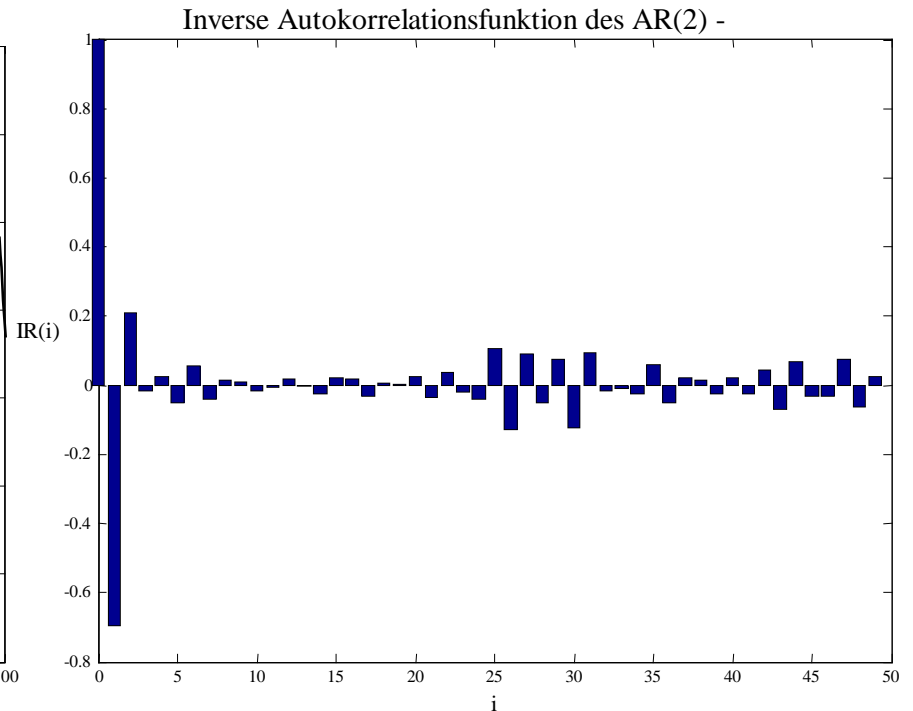
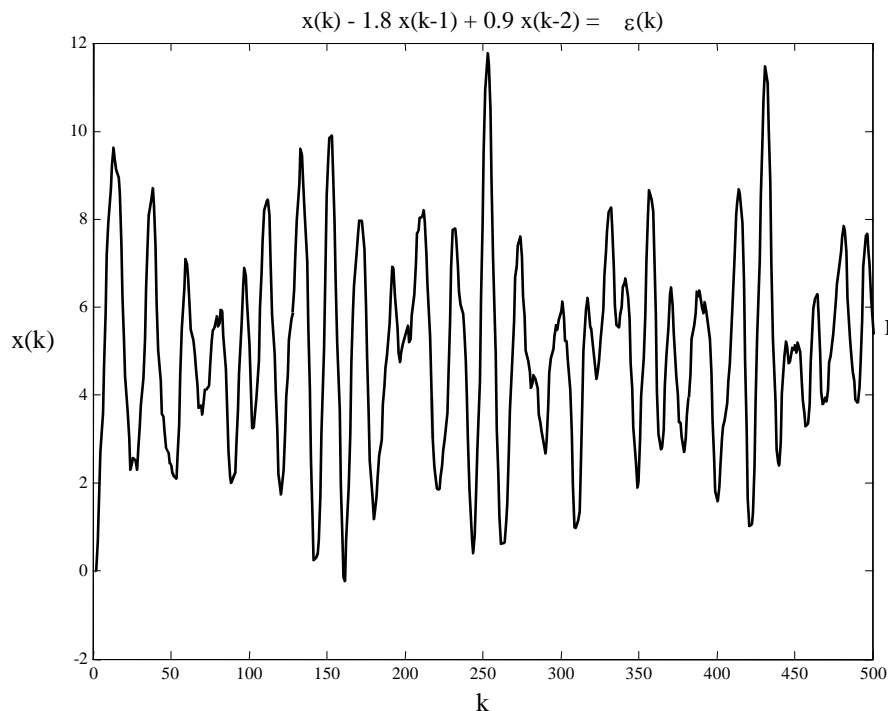
Bei einem MA(n)-Prozess entstehen für die IAKF gedämpfte reine Exponentialfunktionen oder mit Sinus- und / oder Potenzfunktionen gemischte Funktionen.

Die IAKF bricht bei einem AR(m)-Prozess nach dem m-ten Glied ab.

3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.3 Vorhersage mit stochastischen Signalmodellen – AR(2)-Prozess und seine inverse Autokorrelationsfunktion

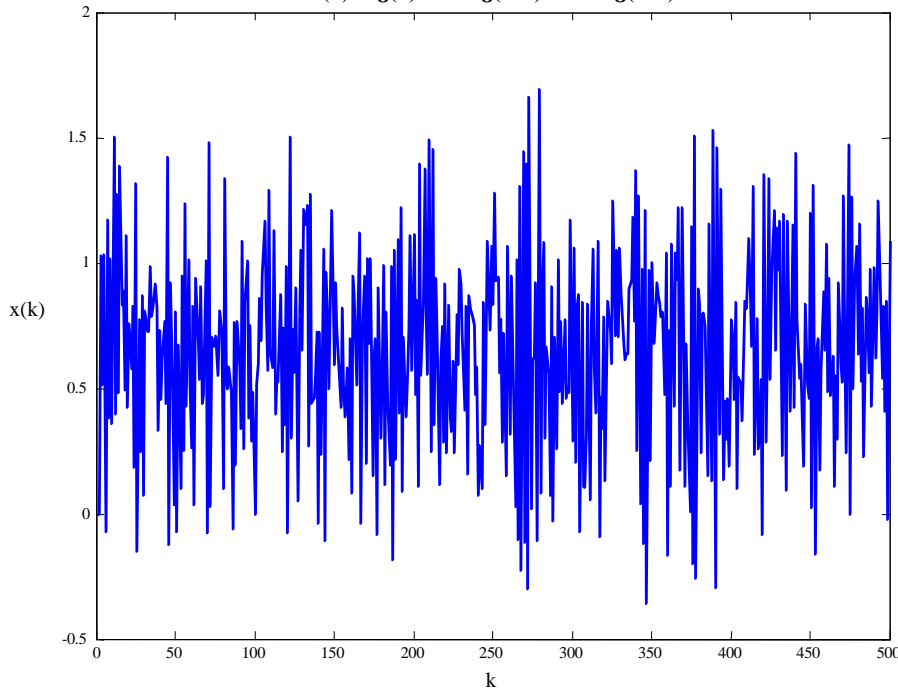


3. Vorhersagestrategien

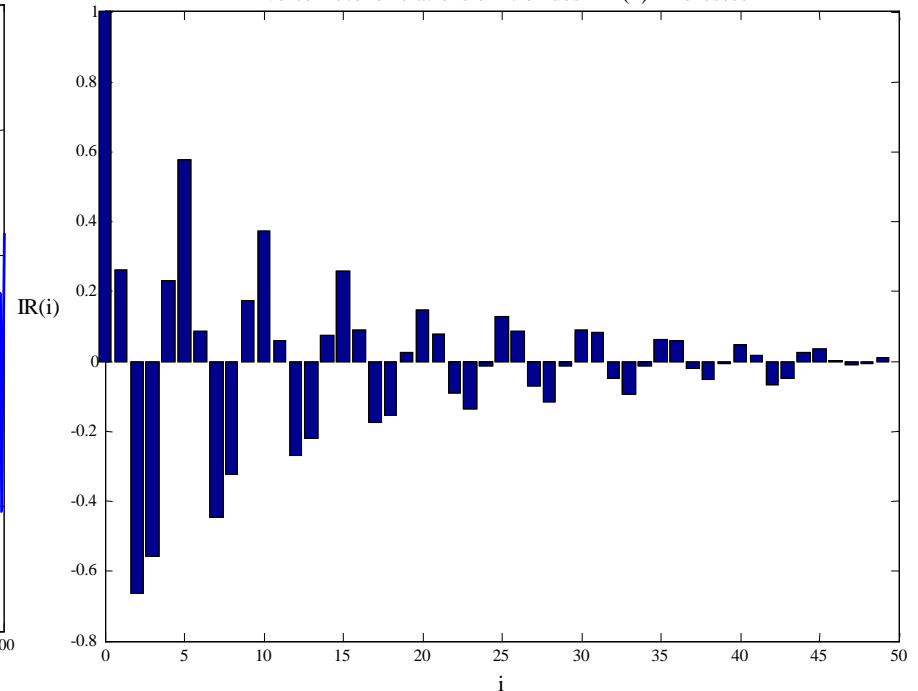
3.6 Vorhersage mit Signalmodellen

3.6.3 Vorhersage mit stochastischen Signalmodellen – MA(2)-Prozess und seine inverse Autokorrelationsfunktion

$$x(k) = \varepsilon(k) - 0.5 \varepsilon(k-1) + 0.8 \varepsilon(k-2)$$



Inverse Autokorrelationsfunktion des MA(2)- Prozesses



3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.3 Vorhersage mit stochastischen Signalmodellen – Zusammenfassende Bewertung von Autokorrelationsfunktion, partieller und inverser Autokorrelationsfunktion

	AR(p)	MA(q)	ARMA(p, q)
$R(\tau)$	abklingende Exponentialfunktion oder gedämpfte Sinuswellen	Abbruch bei Korrelationstiefe q $R(\tau) \neq 0$ für $\tau \leq q$ $R(\tau) = 0$ für $\tau > q$	wie AR(p) ab $\tau > q$
$PR(\tau)$	Abbruch bei Korrelationstiefe p $PR(\tau) \neq 0$ für $\tau \leq p$ $PR(\tau) = 0$ für $\tau > p$	Abklingende Exponentialfunktion oder gedämpfte Sinuswellen	wie MA(q) ab $\tau > p$
$IR(\tau)$	Abbruch bei Korrelationstiefe p $IR(\tau) \neq 0$ für $\tau \leq p$ $IR(\tau) = 0$ für $\tau > p$	ähnlich $PR(\tau)$	ähnlich $PR(\tau)$

3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.3 Vorhersage mit stochastischen Signalmodellen – ARIMA-Prozess (1)

Problem: In praktischen Anwendungsfällen ist die Stationarität des Prozesses nicht immer gewährleistet.

Lösung: Einführung der Klasse der ARIMA-Modelle

Einfachste Form des ARIMA-Modells: $A(q^{-1})\nabla x(k) = A(q^{-1})(1 - q^{-1})x(k) = C(q^{-1}) \cdot \varepsilon(k)$

Der integrierte Charakter bzgl. ε zeigt sich bei Umformung:

$$A(q^{-1})x(k) = C(q^{-1})\varepsilon(k) \cdot \frac{1}{1 - q^{-1}}$$

Allgemeine Darstellung eines ARIMA (m,p,n)-Prozesses:

$$A(q^{-1}) \prod_{i=1}^l \nabla_{p_i}^{q_i} x(k) = C(q^{-1}) \cdot \varepsilon(k)$$

$$\text{mit } \nabla_{p_i}^{q_i} = (1 - q^{-p_i})^{q_i}.$$

p_i = Periode (zur Verarbeitung von Saisonanteilen)

q_i = Ordnung des i-ten Differenzenfilters (häufig beschränkt auf ≤ 2)

3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.3 Vorhersage mit stochastischen Signalmodellen – ARIMA-Prozess (2)

Beispiel 1: ARIMA (1,1,1)-Prozess

$$A(q^{-1})\nabla_1^1 x(k) = C(q^{-1}) \cdot \varepsilon(k)$$

$$a_0(x(k) - x(k-1)) + a_1(x(k-1) - x(k-2)) = c_0\varepsilon(k) + c_1 \cdot \varepsilon(k-1)$$

$$a_0 = c_0 = 1$$

$$x(k) + (a_1 - 1)x(k-1) - a_1x(k-2) = \varepsilon(k) + c_1 \cdot \varepsilon(k-1)$$

$$x(k) = (1 - a_1)x(k-1) + a_1 \cdot x(k-2) + \varepsilon(k) + c_1 \cdot \varepsilon(k-1)$$

$$a_1 = 0,8$$

$$c_1 = 0,1$$

Beispiel 2: ARIMA (1,10,1)-Prozess

$$A(q^{-1})\nabla_{10}^1 x(k) = C(q^{-1}) \cdot \varepsilon(k)$$

$$a_0(x(k) - x(k-10)) + a_1(x(k-10) - x(k-20)) = c_0\varepsilon(k) + c_1 \cdot \varepsilon(k-1)$$

$$a_0 = c_0 = 1$$

$$x(k) + (a_1 - 1)x(k-10) - a_1x(k-20) = \varepsilon(k) + c_1 \cdot \varepsilon(k-1)$$

$$x(k) = (1 - a_1)x(k-10) + a_1 \cdot x(k-20) + \varepsilon(k) + c_1 \cdot \varepsilon(k-1)$$

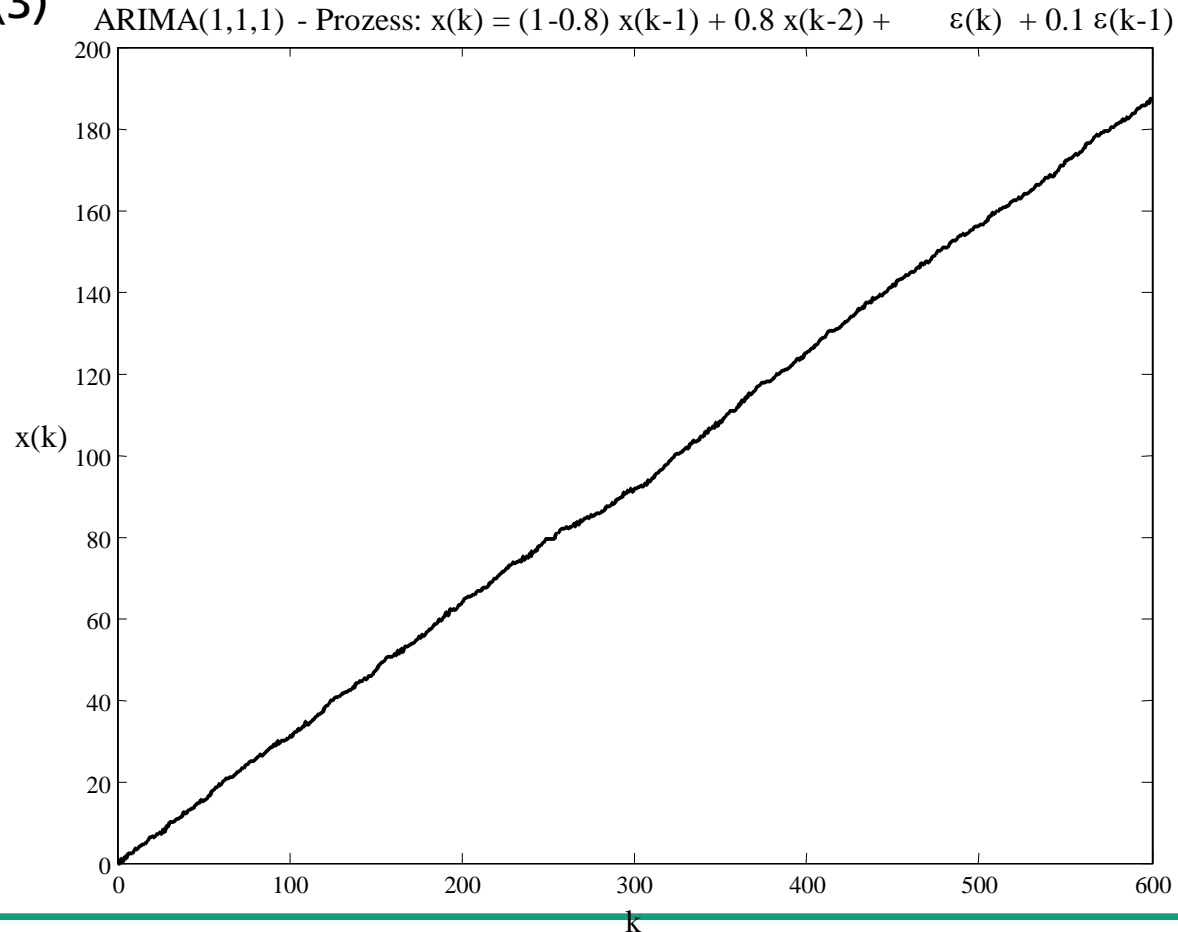
$$a_1 = 0,8$$

$$c_1 = 0,1$$

3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.3 Vorhersage mit stochastischen Signalmodellen – ARIMA-Prozess (3)



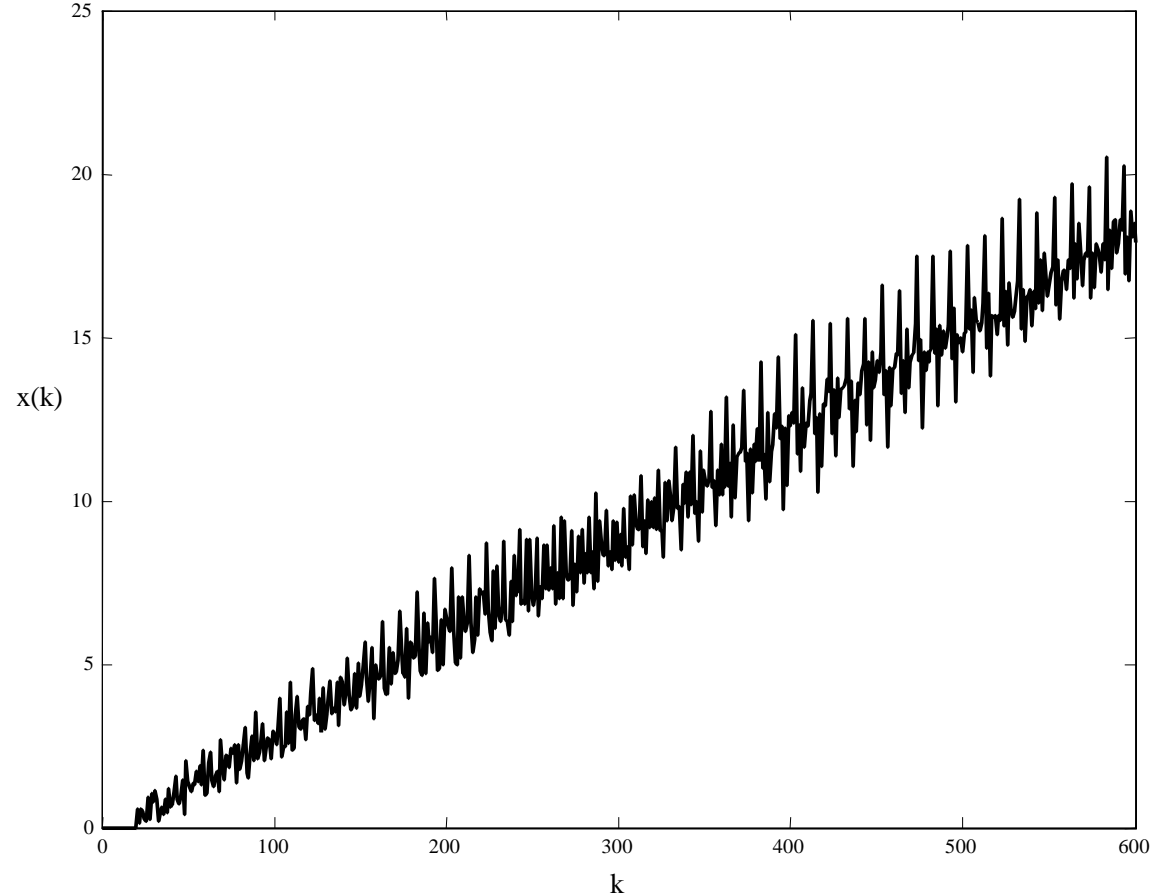
3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.3 Vorhersage mit stochastischen Signalmodellen –

ARIMA-Prozess (4)

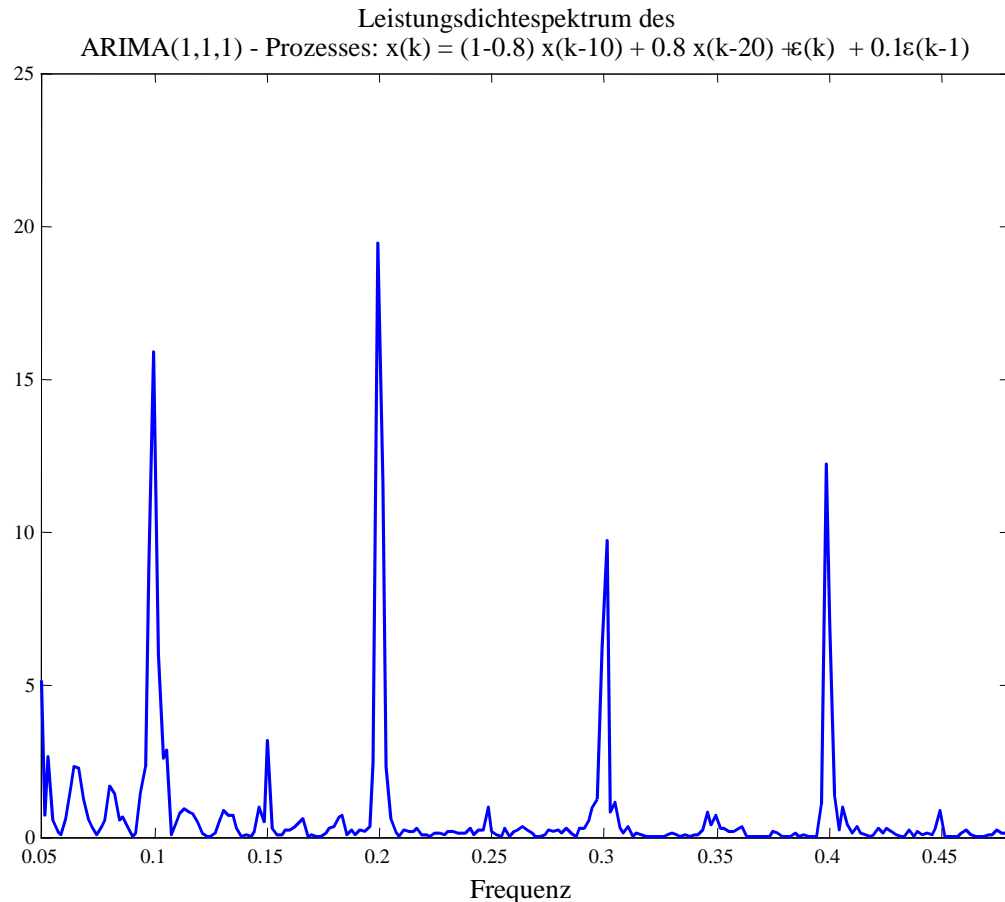
$$\text{ARIMA}(1,1,1) \text{ - Prozess: } x(k) = (1-0.8) x(k-10) + 0.8 x(k-20) + \varepsilon(k) + 0.1 \varepsilon(k-1)$$



3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.3 Vorhersage mit stochastischen Signalmodellen – ARIMA-Prozess (5)



3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.3 Vorhersage mit stochastischen Signalmodellen – Rekursive Prognoseverfahren

Wichtige Anforderungen für praktischen Einsatz von Prognoseverfahren:

- Für neu hinzukommende Daten werden die weitesten in der Vergangenheit liegenden Daten gelöscht.
- Bei Bestimmung der h -Schrittprognose kann auf die $(h-1)$ -Schritt-Prognose zurück gegriffen werden.

Geeignete Verfahren:

- Box-Jenkins-Ansatz
- Kalman-Filter

Wichtig:

Alle Verfahren sind nur unterschiedliche Berechnungsmethoden zur Ermittlung von ein und derselben Qualität, nämlich der Vorhersage eines zukünftigen Wertes aus einer Linearkombination von Vergangenheitsdaten durch Minimierung des mittleren quadratischen Fehlers (MSE).

3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.3 Vorhersage mit stochastischen Signalmodellen – Box-Jenkins-Ansatz (1)

Optimale Prognosen für AR(p)-Prozesse:

Gegeben: $A(q^{-1})x(k) = \varepsilon(k)$ mit $E\{\varepsilon(k)\} = 0$

Gesucht: Optimale 1-Schritt-Prognose $\hat{x}(k+1)$

Lösung:

- $\hat{x}(k+1)$ ist eine Linearkombination vergangener Werte des Prozesses bis einschließlich zum Zeitpunkt k .
- $\hat{x}(k+1)$ und $x(k+1) - \hat{x}(k+1)$ sind unkorreliert, wenn der Prognosefehler nur von $\varepsilon(k+1)$ abhängt.

• Es gilt: $x(k+1) + a_1x(k) + \dots + a_px(k+1-p) = \varepsilon(k+1)$

$$x(k+1) = -a_1x(k) - \dots - a_px(k+1-p) + \varepsilon(k+1)$$

Wenn $x(k+1) - \hat{x}(k+1) = \varepsilon(k+1)$,

dann $\hat{x}(k+1) = x(k+1) - \varepsilon(k+1)$

und somit ist die optimale 1-Schritt-Prognose:

$$\hat{x}(k+1) = -a_1x(k) - \dots - a_px(k+1-p)$$

3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.3 Vorhersage mit stochastischen Signalmodellen – Box-Jenkins-Ansatz (2.1)

Gegeben: $A(q^{-1})x(k) = \varepsilon(k) \quad E\{\varepsilon(k)\} = 0$

Gesucht: Optimale 2-Schritt-Prognose $\hat{x}(k+2)$

Lösung:

- Voraussetzungen wie bei 1-Schritt-Prognose
- Der Prognosefehler darf nur von $\varepsilon(k+1)$ und $\varepsilon(k+2)$ abhängen.
- Es gilt: $x(k+2) + a_1x(k+1) + a_2x(k) + \dots + a_px(k+1-p) = \varepsilon(k+2)$

$$x(k+2) = -a_1x(k+1) - a_2x(k) - \dots - a_px(k+1-p) + \varepsilon(k+2)$$

Wenn $x(k+2) - \hat{x}(k+2) = \varepsilon(k+2)$

dann $\hat{x}(k+2) = x(k+2) - \varepsilon(k+2)$.

3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.3 Vorhersage mit stochastischen Signalmodellen – Box-Jenkins-Ansatz (2.2)

Für die Prognose erhält man: $\hat{x}(k+2) = -a_1x(k+1) - a_2x(k) - \dots - a_px(k+2-p)$

$x(k+1)$ wird mit $\hat{x}(k+1)$ ersetzt :

$$\hat{x}(k+2) = -a_1\hat{x}(k+1) - a_2x(k) - \dots - a_px(k+2-p)$$

Wenn aus $\hat{x}(k+1)$ einer optimalen 1-Schritt-Prognose stammt, gilt:

$$\hat{x}(k+2) = -a_1(x(k+1) - \varepsilon(k+1)) - a_2x(k) - \dots$$

$$\hat{x}(k+2) = -a_1x(k+1) - a_2x(k) - \dots - a_px(k+2-p) + a_1\varepsilon(k+1)$$

Somit gilt :

$$x(k+2) - \hat{x}(k+2) = \varepsilon(k+2) - a_1\varepsilon(k+1)$$

Prognosefunktion und Prognosefehler sind unkorreliert. Die optimale 2-Schritt-Prognose ist:

$$\hat{x}(k+2) = -a_1\hat{x}(k+1) - a_2x(k) - \dots - a_px(k+2-p)$$

3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.3 Vorhersage mit stochastischen Signalmodellen – Box-Jenkins-Ansatz (3)

Optimale Prognose von MA(q)-Prozessen:

Gegeben: $x(k) = \varepsilon(k) + c_1\varepsilon(k-1) + \dots + c_q\varepsilon(k-q)$

Gesucht: optimale h-Schritt-Prognose

Lösung:

$$\hat{x}(k+h) = \begin{cases} c_0\varepsilon(k) + c_1\varepsilon(k-1) + \dots + c_q \cdot \varepsilon(k-q) & \text{für } h \leq q \\ 0 & \text{für } h > q \end{cases}$$

Interpretation:

- Die Definitionsgleichung von $\{x(k)\}$ wird bis zum Zeitpunkt $k+h$ fortgeschrieben.
- Für die nicht realisierten $\varepsilon(k+s)$ wird der Erwartungswert (=0) eingesetzt.

3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.3 Vorhersage mit stochastischen Signalmodellen – Box-Jenkins-Ansatz (4)

Optimale Prognose für ARMA(p,q)-Prozesse

Die optimale h-Schritt-Prognose $\hat{x}(k+h)$ eines stationären ARMA(p,q)-Prozesses $A(q^{-1})x(k) = C(q^{-1})\varepsilon(k)$ erhält man rekursiv, indem die Definitionsgleichung

$$x(k+h) = -a_1x(k+h-1) - \dots - a_px(k+h-p) + \varepsilon(k+h) + c_1\varepsilon(k+h-1) + \dots + c_q\varepsilon(k+h-q)$$

wie folgt interpretiert wird:

1. Die Variablen $x(k), x(k-1), \dots$ entsprechen den tatsächlichen Beobachtungen.
2. Die unbeobachteten Variablen $x(k+h), x(k+h-1), \dots, x(k+1)$, werden durch die optimalen Prognosen $\hat{x}(k+h), \hat{x}(k+h-1), \dots, \hat{x}(k+1)$ ersetzt.
3. Die $\varepsilon(k), \varepsilon(k-1), \dots$ entsprechen den Prognosefehlern $x(k) - \hat{x}(k), x(k-1) - \hat{x}(k-1), \dots$ der optimalen 1-Schritt-Prognosen.
4. Die $\varepsilon(k+h), \varepsilon(k+h-1), \dots, \varepsilon(k+1)$ werden durch ihren Erwartungswert Null ersetzt.

3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.3 Vorhersage mit stochastischen Signalmodellen – Box-Jenkins-Ansatz (5)

Anwendung der Prognoseformel auf einen ARMA(1,1)-Prozess

$$x(k) - 0,4x(k-1) = \varepsilon(k) - 0,6\varepsilon(k-1)$$

Berechnung der optimalen 1-Schritt-Prognosen $\hat{x}(s+1)$ für $s = 0, 1, 2, \dots, k-1$.

Unbekannte Werte durch Nullstellen vor Beginn der Reihe initialisieren

Man erhält: $\hat{x}(0+1) = 0$ mit Prognosefehler $0 - x(1)$

$$\hat{x}(1+1) = 0,4x(1) + 0 - 0,6(0 - x(1))$$

= $x(1)$ mit Prognosefehler $x(1) - x(2)$

$$\hat{x}(2+1) = 0,4x(2) + 0 - 0,6(x(1) - x(2))$$

= $x(2) - 0,6x(1)$ mit Prognosefehler $\hat{x}(2+1) - x(3)$

Schließlich liegen Approximationen für alle Prognosefehler $(x(s) - \hat{x}(s-1+1))$ vor.

Damit beginnt die

Tatsächlich gewünschte

Prognose:

$$\hat{x}(k+1) = 0,4x(k) + 0 - 0,6(\hat{x}(k) - x(k))$$

$$\hat{x}(k+2) = 0,4\hat{x}(k+1)$$

⋮

$$\hat{x}(k+h) = 0,4\hat{x}(k+h-1)$$

3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.3 Vorhersage mit stochastischen Signalmodellen – Box-Jenkins-Ansatz (6)

Wichtig:

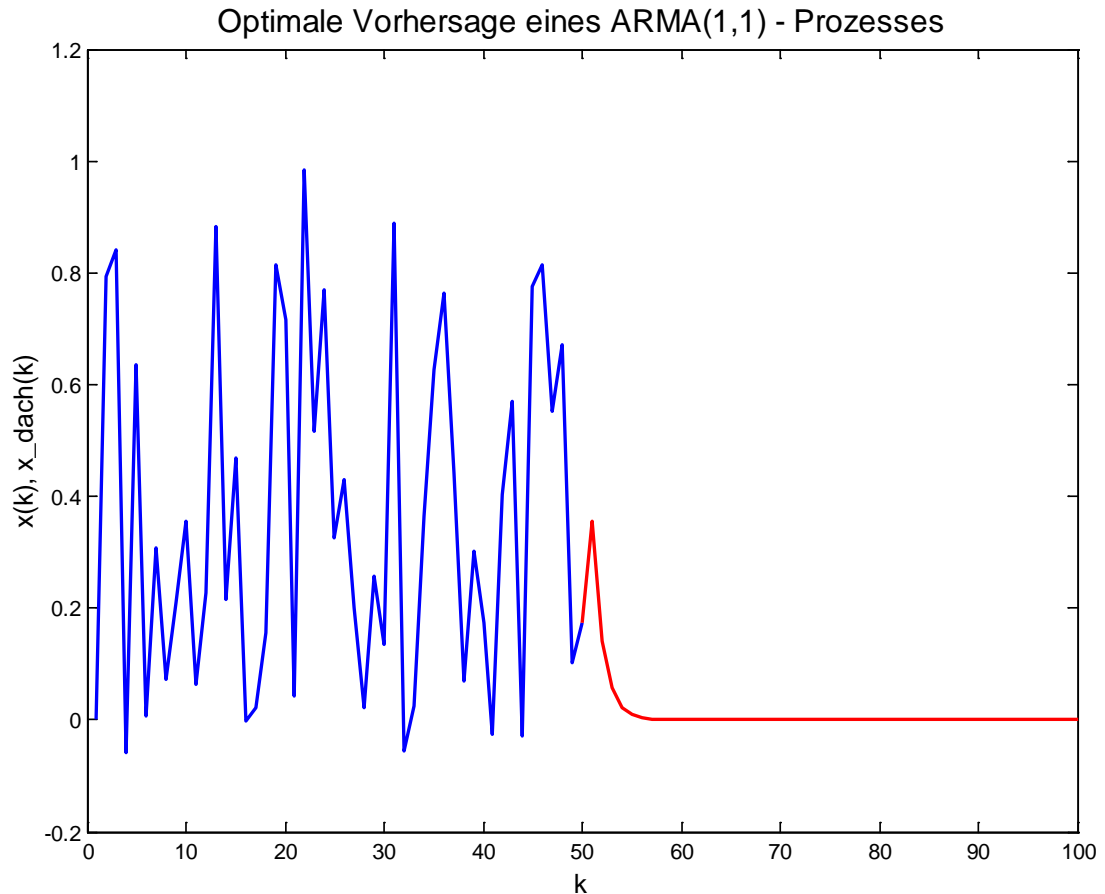
In einem ARMA(p,q)-Prozess werden die Prognosen für $h > q$ nur noch vom AR-Teil des Modells gesteuert.

Der MA-Teil geht über die Anfangswerte $\hat{x}(k + 1), \dots, \hat{x}(k + q)$ ein. Für $h \rightarrow \infty$ geht $\hat{x}(k + h) \rightarrow 0$, d. h. für große Prognosehorizonte werden die Prognosen zunehmend unbrauchbar.

3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.3 Vorhersage mit stochastischen Signalmodellen – Box-Jenkins-Ansatz (7)



3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.3 Vorhersage mit stochastischen Signalmodellen – Box-Jenkins-Ansatz (8) → nur zur Information

Ableitung der Prognoseformel für ARMA-Prozesse mit dem Konzept der bedingten Erwartungen:

Ausgangspunkt: Optimale Prognose $\hat{x}(k+h)$ ist der bedingte Erwartungswert von $x(k+h)$ bei gegebenen $x(k), x(k-1), \dots$

Für einen stationären und invertierbaren ARMA-Prozess $A(q^{-1})x(k) = B(q^{-1})\varepsilon(k)$ gilt mit der Unabhängigkeit von $\varepsilon(k)$:

$$\begin{aligned} E\{\varepsilon(k+1)|x(k), x(k-1), \dots\} &= E\{\varepsilon(k+1)|\varepsilon(k), \varepsilon(k-1), \dots\} \\ &= E\{\varepsilon(k+1)\} \\ &= 0 \end{aligned}$$

Für $s \leq k$ ist

$$\begin{aligned} E\{\varepsilon(k+1)|x(k), x(k-1), \dots\} &= E\{\varepsilon(k+1)|\varepsilon(k), \varepsilon(k-1), \dots\} \\ &= E\{\varepsilon(k+1)\} \\ &= 0 \end{aligned}$$

Daraus folgt:

$$\begin{aligned} \hat{x}(k+1) &= E\{x(k+1)|x(k), x(k-1), \dots\} \\ &= E\{-a_1x(k) - \dots - a_mx(k+1-m) + \varepsilon(k+1) + b_1\varepsilon(k) + \dots + b_p\varepsilon(k+1-n)|x(k), x(k-1), \dots\} \\ &= -a_1x(k) - \dots - a_mx(k+1-m) + b_1\varepsilon(k) + \dots + b_p\varepsilon(k+1-n) \end{aligned}$$

Analog dazu erhält man die h-Schritt-Prognose $\hat{x}(k+h)$.

3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.3 Vorhersage mit stochastischen Signalmodellen – Box-Jenkins-Ansatz (9)

Die Varianz des Prognosefehlers ist: $\text{var} \{x(k+h) - \hat{x}(k+h)\} = \sigma_\varepsilon^2 \sum_{u=0}^{h-1} c_u^2$

mit σ_ε^2 Varianz des White-Noise-Prozesses $\varepsilon(k)$
 c_u Koeffizienten der MA-Darstellung von $\{x(k)\}$ $x(k) = \frac{B(q^{-1})}{A(q^{-1})} \varepsilon(k)$

c_u wird praktisch durch Koeffizientenvergleich bestimmt.

$$C(q^{-1}) = \frac{B(q^{-1})}{A(q^{-1})}$$

allgemein:

$$c_k + \sum_{j=1}^m \alpha_j \cdot c_{k-j} = b_k$$

$$A(q^{-1}) \cdot C(q^{-1}) = B(q^{-1})$$

$$(1 + a_1 q^{-1} + \dots + a_m q^{-m}) \cdot (c_0 + c_1 q^{-1} + c_2 q^{-2} + \dots) = (1 + b_1 q^{-1} + \dots + b_n \cdot q^{-n})$$

$$c_0 = 1$$

$$c_1 + a_1 c_0 = b_1$$

⋮

$$c_m + a_1 c_{m-1} + \dots + a_m c_0 = b_m$$

mit $b_k = 0$ für $k > n$

$c_k = 0$ für $k > 0$

3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.3 Vorhersage mit stochastischen Signalmodellen – Box-Jenkins-Ansatz (10)

Die Varianz des Prognosefehlers strebt mit wachsendem Prognosehorizont h gegen die Prozessvarianz.

Die Prognose geht mit wachsendem h immer mehr gegen den Erwartungswert des Prozesses.

-> Daher ist die Prognose über einen vom jeweiligen Prozess abhängigen Prognosehorizont hinaus nicht sinnvoll.

Hinweis:

Momente eines MA(n)-Prozesses $x(k) = B(q^{-1})\varepsilon(k)$

Wenn $E\{\varepsilon(k)\} = \mu_\varepsilon$ und $\text{var}\{\varepsilon(k)\} = \sigma_\varepsilon^2$,

dann $\mu_x = E\{x(k)\} = \mu_\varepsilon \sum_{u=0}^n b_u$

$$\sigma_x^2 = \text{var}\{x(k)\} = \sigma_\varepsilon^2 \cdot \sum_{u=0}^n b_u^2$$

3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.3 Vorhersage mit stochastischen Signalmodellen – Box-Jenkins-Ansatz (11)

Prognoseintervall der h -Schritt-Prognose für einen Normalprozess $X(k)$:
Intervall, in dem der Wert von $X(k + h)$ mit einer vorgegebenen Wahrscheinlichkeit $(1 - \alpha)$ liegt.

Bei Normalprozess ist auch der Prognosefehler normal verteilt. Es gilt:

$$P\left[\hat{x}(k+h) - z \cdot \sigma_F \leq \hat{x}(k+h) \leq \hat{x}(k+h) + z \cdot \sigma_F\right]$$

z : $\left(1 - \frac{\alpha}{2}\right)$ Quantil der Standardnormalverteilung

$$= P\left[-z \leq \frac{x(k+h) - \hat{x}(k+h)}{\sigma_F} \leq z\right]$$

σ_F^2 : Varianz der Prognosefehler

$$= 1 - \alpha$$

Damit gilt für das Prognoseintervall $\hat{x}(k+h) \pm z \cdot \sigma_\varepsilon \sqrt{\sum_{u=0}^{h-1} b_u^2}$

Gewichte b_u gehen für große u gegen Null



Prognoseintervall steigt ab einem bestimmten Prognosehorizont h^* kaum noch an.

3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.3 Vorhersage mit stochastischen Signalmodellen – Box-Jenkins-Ansatz (12)

Beispiel:

ARMA(1,1)-Prozess $x(k) - 0,4x(k-1) = \varepsilon(k) - 0,6\varepsilon(k-1)$
mit $\varepsilon(k) \sim N[0, \sigma^2]$

Koeffizienten c_u :

$$\begin{aligned}c_0 &= 1 \\c_1 &= b_1 - a_1 \cdot c_0 \\c_1 &= -0,6 + 0,4 = -0,2 \\c_2 &= b_2 - a_1 \cdot c_1 \\c_2 &= 0 - (-0,4 \cdot -0,2) = -0,08 \\&\vdots\end{aligned}$$

Für Prognoseintervall, in dem 75% aller Werte liegen, erhält man für z den Tafelwert $z = 1,15$.



$$\hat{x}(k+1)_{\pm 1,15 \cdot \sigma}$$

$$\hat{x}(k+2)_{\pm 1,15 \cdot \sigma} \cdot \sqrt{1 + 0,2^2} = \hat{x}(k+2)_{\pm 1,173 \cdot \sigma}$$

$$\begin{aligned}\hat{x}(k+3)_{\pm 1,15 \cdot \sigma} \cdot \sqrt{1 + 0,2^2 + 0,08^2} &= \hat{x}(k+3)_{\pm 1,203} \\&= \hat{x}(k+4)_{\pm 1,206 \cdot \sigma}\end{aligned}$$

3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.3 Vorhersage mit stochastischen Signalmodellen – Kalman-Filter (1)

Zustandsraummodelle:

- wurden im Bereich der Regelungstechnik entwickelt, erst in letzter Zeit werden sie zur Modellierung und Prognose empirischer Zeitreihen verstärkt verwendet.
- beruhen auf der Markov-Eigenschaft. Das bedeutet, die gesamte für die Zukunft (\triangleq Prognose) relevante Information ist in der Gegenwart enthalten. Wie es zu dem in der Gegenwart erreichten Zustand kam, ist nicht wesentlich.

besteht aus Systemgleichung: $\underline{x}(k+1) = \underline{A}(k) \cdot \underline{x}(k) + \underline{B}(k) \cdot \underline{\varepsilon}(k+1)$

und Beobachtungsgleichung $\underline{y}(k) = \underline{C}(k) \cdot \underline{x}(k) + \underline{\eta}(k)$

mit $\underline{A}(k)$ - Übertragungsmatrix

$\underline{B}(k)$ - Eingangsmatrix

$\underline{C}(k)$ - Beobachtungsmatrix

$\underline{x}(k)$ - Zustandsvektor

$\underline{y}(k)$ - Ausgangs- oder Messgrößenvektor

$\underline{\varepsilon}(k)$ - Modellfehlervektor

$\underline{\eta}(k)$ - Messfehlervektor

} unabhängige
Zufallsvektoren

3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.3 Vorhersage mit stochastischen Signalmodellen – Kalman-Filter (2)

Für $\underline{\varepsilon}(k)$ und $\underline{\eta}(k)$ gilt: $E\{\underline{\varepsilon}(k)\} = E\{\underline{\eta}(k)\} = 0$,
 $Var\{\underline{\varepsilon}(k)\} = \underline{R}(k)$ und
 $Var\{\underline{\eta}(k)\} = \underline{Q}(k)$.

Darstellung eines AR(p)-Prozesses im Zustandsraum: $y(k) = -a_1y(k-1) - \dots - a_p y(k-p) + \varepsilon(k)$

Zustandsvektor $\underline{x}(k) = (y(k), y(k-1), \dots, y(k-p+1))'$,



Systemgleichung:

$$\begin{bmatrix} y(k+1) \\ y(k) \\ y(k-1) \\ \vdots \\ y(k-p+2) \end{bmatrix} = \begin{bmatrix} -a_1 - a_2 \dots - a_{p-1} - a_p & & & & \\ 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 \end{bmatrix} \begin{bmatrix} y(k) \\ y(k-1) \\ y(k-2) \\ \vdots \\ y(k-p+1) \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Beobachtungsgleichung: $y(k) = [1 \ 0 \ \dots \ 0 \ 0] \begin{bmatrix} y(k) \\ y(k-1) \\ \vdots \\ y(k-p+1) \end{bmatrix}$

3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.3 Vorhersage mit stochastischen Signalmodellen – Kalman-Filter (3)

Weitere Möglichkeit:

Zustandsvektor: $\underline{x}(k) = (-a_1(k), -a_2(k), \dots, -a_p(k-p+1))'$



Systemgleichung:

$$\begin{bmatrix} -a_1(k+1) \\ -a_2(k+1) \\ \vdots \\ -a_p(k+1) \end{bmatrix} = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix} \begin{bmatrix} -a_1(k) \\ -a_2(k) \\ \vdots \\ -a_p(k) \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \underline{\varepsilon}(k)$$

Beobachtungsgleichung:

$$y(k) = [y(k-1) \dots y(k-p)] \begin{bmatrix} -a_1(k) \\ -a_2(k) \\ \vdots \\ -a_p(k) \end{bmatrix} + \underline{\eta}(k)$$

3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.3 Vorhersage mit stochastischen Signalmodellen – Kalman-Filter (4)

Zustandsraum-Darstellung eines MA(q)-Prozesses: $y(k) = \varepsilon(k) + b_1\varepsilon(k-1) + b_q\varepsilon(k-q)$

Zustandsvektor: $x(k) = [\varepsilon(k), \varepsilon(k-1), \dots, \varepsilon(k-q)]'$



Systemgleichung:

$$\begin{bmatrix} \varepsilon(k+1) \\ \varepsilon(k) \\ \varepsilon(k-1) \\ \vdots \\ \varepsilon(k-q+1) \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & \dots & 0 & 0 \\ 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & 0 \end{bmatrix} \begin{bmatrix} \varepsilon(k) \\ \varepsilon(k-1) \\ \varepsilon(k-2) \\ \vdots \\ \varepsilon(k-q) \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \varepsilon(k+1)$$

Beobachtungsgleichung: $y(k) = [1 \ b_1 \ b_2 \ \dots \ b_q] \begin{bmatrix} \varepsilon(k) \\ \varepsilon(k-1) \\ \varepsilon(k-2) \\ \vdots \\ \varepsilon(k-q) \end{bmatrix}$

3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.3 Vorhersage mit stochastischen Signalmodellen – Kalman-Filter (5)

Weitere Möglichkeit:

Parameter als Zustandsvektor: $x(k) = [1 \ b_1(k) \ b_2(k) \dots b_q(k)]'$



Systemgleichung:

$$\begin{bmatrix} 1 \\ b_1(k+1) \\ b_2(k+1) \\ \vdots \\ b_q(k+1) \end{bmatrix} = \begin{bmatrix} 0 & 0 & \dots & 0 \\ 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix} \begin{bmatrix} 1 \\ b_1(k) \\ b_2(k) \\ \vdots \\ b_q(k) \end{bmatrix}$$

Beobachtungsgleichung:

$$y(k) = [\varepsilon(k), \varepsilon(k-1), \dots, \varepsilon(k-q)] \cdot \begin{bmatrix} 1 \\ b_1(k) \\ b_2(k) \\ \vdots \\ b_q(k) \end{bmatrix} + \eta(k)$$

3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.3 Vorhersage mit stochastischen Signalmodellen – Kalman-Filter (6)

Zustandsraumbeschreibung eines ARMA(p,q)-Prozesses:

$$y(k) = -a_1 y(k-1) - \dots - a_p y(k-p) + \varepsilon(k) + b_1 \varepsilon(k-1) + \dots + b_q \varepsilon(k-q)$$

Zustandsvektor: $\underline{x}(k) = [y(k), y(k-1), \dots, y(k-p+1), \varepsilon(k), \varepsilon(k-1), \dots, \varepsilon(k-q+1)]'$



Systemgleichung: $[y(k+1), y(k), \dots, y(k-p+2), \varepsilon(k+1), \varepsilon(k), \dots, \varepsilon(k-q+2)]' =$

$$\begin{bmatrix} -a_1 & -a_2 & \dots & -a_{p-1} & -a_p & b_1 & b_2 & \dots & b_{q-1} & b_q \\ 1 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 & 1 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots & & & & \vdots & \vdots \\ 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & 1 & 0 \end{bmatrix} \begin{bmatrix} y(k) \\ y(k-1) \\ \vdots \\ y(k-p+1) \\ \varepsilon(k) \\ \varepsilon(k-1) \\ \vdots \\ \varepsilon(k-q+1) \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \varepsilon(k+1)$$

Beobachtungsgleichung:



$$y(k) = [1 \ 0 \ \dots \ 0 \ \dots \ 0] \begin{bmatrix} y(k) \\ y(k-1) \\ \vdots \\ \varepsilon(k-q+2) \\ \varepsilon(k-q+1) \end{bmatrix}$$

3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.3 Vorhersage mit stochastischen Signalmodellen – Kalman-Filter (7.1)

Weitere Möglichkeit:

Parameter als Zustandsvektor: $\underline{x}(k) = [-a_1(k), -a_2(k), \dots, -a_p(k), 1, b_1(k), b_2(k), \dots, b_q(k)]$



Systemgleichung:

$$\begin{bmatrix} -a_1(k+1) \\ -a_2(k+1) \\ \vdots \\ -a_p(k+1) \\ 1 \\ b_1(k+1) \\ b_2(k+1) \\ \vdots \\ b_q(k+1) \end{bmatrix} = \begin{bmatrix} 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \dots & 0 & 1 \end{bmatrix} \begin{bmatrix} -a_1(k) \\ -a_2(k) \\ \vdots \\ -a_p(k) \\ 1 \\ b_1(k) \\ b_2(k) \\ \vdots \\ b_q(k) \end{bmatrix}$$

3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.3 Vorhersage mit stochastischen Signalmodellen – Kalman-Filter (7.2)

Beobachtungsgleichung:

$$y(k) = [y(k) \ y(k-1), \dots, y(k-p), \varepsilon(k), \dots, \varepsilon(k-q)] \begin{bmatrix} -a_1(k) \\ -a_2(k) \\ \vdots \\ 1 \\ b_1(k) \\ b_2(k) \\ \vdots \\ b_q(k) \end{bmatrix} + \eta(k)$$

3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.3 Vorhersage mit stochastischen Signalmodellen – Kalman-Filter (8)

Schätzung des Zustandes $\underline{x}(k)$ mit dem Kalman-Filter:

Ausgangspunkt: multivariates Zustandsraummodell $\underline{x}(k+1) = \underline{A}(k) \cdot \underline{x}(k) + \underline{B}(k) \cdot \underline{\varepsilon}(k+1)$

$$\underline{y}(k) = \underline{C}(k)\underline{x}(k) + \underline{\eta}(k)$$

$$E\{\underline{\varepsilon}\} = E\{\underline{\eta}\} = 0$$

$$\text{Var}\{\underline{\varepsilon}\} = \underline{Q}(k)$$

$$\text{Var}\{\underline{\eta}\} = \underline{R}(k)$$

Grundlage: Schätzung von $\underline{x}(k)$ rekursiv aus der Schätzung von $\underline{x}(k-1)$ und der letzten Beobachtung $\underline{y}(k)$

Vorteil: Für jedes Modell, das sich in Zustandsform darstellen lässt, erhält man mit dem Kalman-Filter eine einheitliche Prognosemethode.

Schätzung: Beste lineare Approximation $\hat{\underline{x}}(k)$ von $\underline{x}(k)$ durch die bis einschließlich zum Zeitpunkt k erhaltenen Beobachtungen $\underline{y}(1), \underline{y}(2), \dots, \underline{y}(k)$.

3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.3 Vorhersage mit stochastischen Signalmodellen – Kalman-Filter (9)

$\hat{x}(k)$ ist dadurch charakterisiert, dass die Schätzfehler $\underline{x}(k) - \hat{x}(k)$ und die Beobachtungen $\underline{y}(1), \underline{y}(2), \dots, \underline{y}(k)$ unkorreliert sind:

$$\text{cov}\{\underline{x}(k) - \hat{x}(k), \underline{y}(s)\} = 0; \quad s \leq k$$

Daraus Ableitung der Rekursivformel. Die Rekursion verläuft in zwei Schritten:

- 1) Prädiktionsschritt (vor der Beobachtung von $\underline{y}(k)$)
- 2) Korrekturschritt (nach der Beobachtung von $\underline{y}(k)$)

Annahme: $\hat{x}(k-1)$ bereits bekannt

Neue Beobachtung $\underline{y}(k)$ noch nicht bekannt.



Beste lineare Approximation $\tilde{x}(k)$ von $\underline{x}(k)$ durch $\underline{y}(k-1), \underline{y}(k-2), \dots, \underline{y}(1)$ erhält man, wenn die unbekannte $\varepsilon(k)$ durch die optimale Prognose, d. h. durch den Erwartungswert $\underline{0}$ ersetzt wird: $\tilde{x}(k) = \underline{A}(k-1) \cdot \hat{x}(k-1)$

analog: $\tilde{y}(k) = \underline{C}(k) \cdot \tilde{x}(k)$

3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.3 Vorhersage mit stochastischen Signalmodellen – Kalman-Filter (10)

Abschätzung der Größe des Prognosefehlers, d. h. der Kovarianzmatrix ist möglich, wenn Kovarianzmatrix von $\underline{x}(k-1) - \underline{\hat{x}}(k-1)$ bekannt ist:

$$\underline{P}(k) = \text{Var}\{\underline{x}(k) - \underline{\hat{x}}(k)\} = E\{(\underline{x}(k) - \underline{\hat{x}}(k))(\underline{x}(k) - \underline{\hat{x}}(k))'\}$$

Prognosefehler:
$$\begin{aligned}\underline{x}(k) - \underline{\tilde{x}}(k) &= \underline{A}(k-1) \cdot \underline{x}(k-1) + \underline{B}(k-1) \cdot \underline{\varepsilon}(k) - \underline{A}(k-1) \cdot \underline{\hat{x}}(k-1) \\ &= \underline{A}(k-1)(\underline{x}(k-1) - \underline{\hat{x}}(k-1)) + \underline{B}(k-1) \cdot \underline{\varepsilon}(k)\end{aligned}$$

Setzen von:

$$\underline{\tilde{P}}(k) = \text{var}\{\underline{x}(k) - \underline{\tilde{x}}(k)\}$$



$$\underline{\tilde{P}}(k) = \text{Var}\{\underline{A}(k-1)(\underline{x}(k-1) - \underline{\hat{x}}(k-1)) + \underline{B}(k-1) \cdot \underline{\varepsilon}(k)\}$$

da $(\underline{x}(k-1) - \underline{\hat{x}}(k-1))$ und $\underline{\varepsilon}(k)$ unkorreliert, gilt:

$$\underline{\tilde{P}}(k) = \text{Var}\{\underline{A}(k-1)(\underline{x}(k-1) - \underline{\hat{x}}(k-1))\} + \text{Var}\{\underline{B}(k-1) \cdot \underline{\varepsilon}(k)\}$$

mit $\text{Var}\{\underline{\varepsilon}(k)\} = \underline{Q}(k)$ kann man

ableiten:

$$\underline{\tilde{P}}(k) = \underline{A}(k-1) \underline{P}(k-1) \underline{A}'(k-1) + \underline{B}(k-1) \underline{Q}(k-1) \underline{B}'(k-1)$$

Analog erhält man:
$$\text{Var}\{\underline{y}(k) - \underline{\tilde{y}}(k)\} = \underline{C}(k) \underline{\tilde{P}}(k) \underline{C}'(k) + \underline{R}(k)$$

3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.3 Vorhersage mit stochastischen Signalmodellen – Kalman-Filter (11)

Wenn die Beobachtung $\underline{y}(k)$ eintritt, kann der Prognosefehler $\underline{y}(k) - \underline{\tilde{y}}(k)$ berechnet werden.



Korrekturschritt für Zustand:

$$\hat{\underline{x}}(k) = \underline{\tilde{x}}(k) + \underline{K}(k)(\underline{y}(k) - \underline{\tilde{y}}(k))$$

$\underline{K}(k)$ - Kalman'sche Verstärkung

Berechnung von $\underline{K}(k)$:
$$\underline{K}(k) = \underline{\tilde{P}}(k) \underline{C}'(k) \cdot (\underline{C}(k) \underline{\tilde{P}}(k) \underline{C}'(k) + \underline{R}(k))^{-1}$$

Berechnung der Kovarianzmatrix $\underline{P}(k)$ von $\underline{x}(k) - \underline{\tilde{x}}(k)$:

$$\underline{P}(k) = \hat{\underline{P}}(k) - \underline{K}(k) \underline{C}(k) \underline{\tilde{P}}(k)$$

3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.3 Vorhersage mit stochastischen Signalmodellen – Kalman-Filter (12)

Zusammenfassung der Berechnungsvorschriften für den Kalman-Filter:

(1) Prädiktionsschritt:

a) $\tilde{\underline{x}}(k) = \underline{A}(k-1)\hat{\underline{x}}(k-1)$

b) $\tilde{\underline{y}}(k) = \underline{C}(k)\tilde{\underline{x}}(k)$

c) $\tilde{\underline{P}}(k) = \underline{A}(k-1)\underline{P}(k-1)\underline{A}'(k-1) + \underline{B}(k-1)\underline{Q}(k-1)\underline{B}'(k-1)$

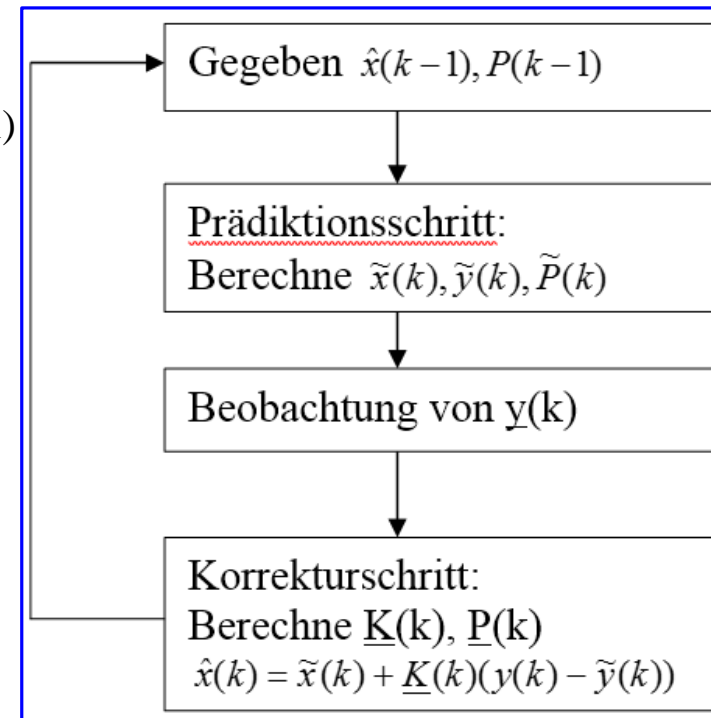
(2) Korrekturschritt:

a) $\underline{K}(k) = \tilde{\underline{P}}(k)\underline{C}'(k)\left(\underline{C}(k)\tilde{\underline{P}}(k)\underline{C}'(k) + \underline{R}(k)\right)^{-1}$

b) $\hat{\underline{x}}(k) = \tilde{\underline{x}}(k) + \underline{K}(k)(\underline{y}(k) - \tilde{\underline{y}}(k))$

c) $\underline{P}(k) = \tilde{\underline{P}}(k) - \underline{K}(k)\cdot\underline{C}(k)\cdot\tilde{\underline{P}}(k)$

(3) Ablaufdiagramm:



3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.3 Vorhersage mit stochastischen Signalmodellen – Anwendung des Kalman-Filters (1)

Vorhersage für einen AR(2)-Prozess: $y(k) - 0,55y(k-1) - 0,23y(k-2) = \varepsilon(k)$

Darstellung im Zustandsraum: $\underline{x}(k) = [y(k) \quad y(k-1)]'$;



Systemgleichung:
$$\begin{bmatrix} y(k+1) \\ y(k) \end{bmatrix} = \begin{bmatrix} -a_1 - a_2 \\ 1 \quad 0 \end{bmatrix} \begin{bmatrix} y(k) \\ y(k-1) \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} \cdot \varepsilon(k)$$



Beobachtungsgleichung:
$$y(k) = [1 \quad 0] \begin{bmatrix} y(k) \\ y(k-1) \end{bmatrix}$$

Somit gilt:

$$A = \begin{bmatrix} -a_1 - a_2 \\ 1 \quad 0 \end{bmatrix} = \begin{bmatrix} 0,55 & 0,23 \\ 1 & 0 \end{bmatrix}$$

$$B = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$C = [1 \quad 0]$$

3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.3 Vorhersage mit stochastischen Signalmodellen – Anwendung des Kalman-Filters (2)

Problem: Bestimmung von

- $Var\{\underline{\varepsilon}(k)\} = Q$
- $Var\{\underline{\eta}(k)\} = R$
- Startwert von $P = Var\{x(k) - \hat{x}(k)\}$

R :

- Maß für Genauigkeit der Messung
- wird für Zeitreihenvorhersage häufig auf 0 gesetzt (oder sehr kleine Werte)
- je größer \underline{R} , um so kleiner die Kalman-Verstärkung, d. h. Anpassung an Prozessänderung erfolgt langsamer

Q :

- Maß für Genauigkeit der Zustandsermittlung
- geht in die Berechnung von $\underline{\tilde{P}}$ und damit in Kalman-Verstärkung ein
- je größer \underline{Q} , um so größer \underline{K} , Anpassung an Prozessänderung erfolgt schneller
- Q kann z. B. durch eine Optimierung unter Einbeziehung von Messdaten erfolgen.

P :

- wird als Diagonalmatrix mit großen Werten (z. B. 10^4) angenommen.
- Werte sollen um so höher sein, je ungenauer die Vorabinformationen sind.

3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.3 Vorhersage mit stochastischen Signalmodellen – Anwendung des Kalman-Filters (3.1)

Beispielrechnung: $Q = 10$ $\underline{A}, \underline{B}, \underline{C}$ sind zeitinvariant

$$R = 0$$

$$P(0) = \begin{bmatrix} 10000 & 0 \\ 0 & 10000 \end{bmatrix}$$

Angenommene Messreihe des AR(2)-Prozesses: $\underline{Y} = [y(k) \quad y(k-1) \quad y(k-2) \quad y(k-3), \dots]'$
 $= [-0,169 \quad -0,0094 \quad 0,314 \quad 0,5285, \dots]'$

$y(k)$ sei noch nicht bekannt und soll vorhergesagt werden.

1) Prädiktionsschritt:

$$a) \underline{\tilde{x}}(k) = \underline{A} \cdot \underline{\hat{x}}(k-1) \quad \underline{\hat{x}}(k-1) = \begin{bmatrix} y(k-1) \\ y(k-2) \end{bmatrix} = \begin{bmatrix} -0,0094 \\ 0,3140 \end{bmatrix}$$

$$\underline{\hat{x}}(k) = \begin{bmatrix} 0,55 & 0,23 \\ 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} -0,0094 \\ 0,3140 \end{bmatrix}$$

$$\underline{\tilde{x}}(k) = \begin{bmatrix} 0,0670 \\ -0,0094 \end{bmatrix}$$

3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.3 Vorhersage mit stochastischen Signalmodellen – Anwendung des Kalman-Filters (3.2)

1) Prädiktionsschritt:

$$b) \tilde{y}(k) = \underline{C} \cdot \tilde{x}(k) = [1 \ 0] \cdot \begin{bmatrix} 0,0670 \\ -0,0094 \end{bmatrix}$$

$$\underline{\tilde{y}(k)} = 0,0670$$

$$c) \underline{\tilde{P}}(k) = \underline{A} \cdot P(k-1) \cdot \underline{A}' + \underline{B} \cdot Q \cdot \underline{B}'$$

$$= \begin{bmatrix} 0,55 & 0,23 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 10000 & 0 \\ 0 & 10000 \end{bmatrix} \cdot \begin{bmatrix} 0,53 & 1 \\ 0,23 & 0 \end{bmatrix} \\ + \begin{bmatrix} 1 \\ 0 \end{bmatrix} \cdot 10 \cdot [10] = \underline{\underline{\begin{bmatrix} 3564 & 5510 \\ 5510 & 10010 \end{bmatrix}}}$$

3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.3 Vorhersage mit stochastischen Signalmodellen – Anwendung des Kalman-Filters (4.1)

2) Korrekturschritt: a) $\underline{K}(k) = \underline{\tilde{P}}(k) \cdot \underline{c}' \cdot (\underline{C} \cdot \underline{\tilde{P}}(k) \cdot \underline{C}' + R)^{-1}$

$$\underline{K}(k) = \begin{bmatrix} 3564 & 5510 \\ 5510 & 10010 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix} \cdot \left(\begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} 3564 & 5510 \\ 5510 & 10010 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 0 \right)^{-1}$$

$$\underline{K}(k) = \begin{bmatrix} 3564 \\ 5510 \end{bmatrix} \cdot (3564)^{-1}$$

$$\underline{K}(k) = \begin{bmatrix} 1,00 \\ 1,55 \end{bmatrix}$$

b) $\underline{\hat{x}}(k) = \underline{\tilde{x}}(k) + \underline{K}(k)(y(k) - \tilde{y}(k))$

$$\underline{\hat{x}}(k) = \begin{bmatrix} 0,0670 \\ -0,0094 \end{bmatrix} + \begin{bmatrix} 1,00 \\ 1,55 \end{bmatrix} \cdot (0,169 - 0,067)$$

$$\underline{\hat{x}}(k) = \begin{bmatrix} -0,1690 \\ -0,3752 \end{bmatrix}$$

3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.3 Vorhersage mit stochastischen Signalmodellen – Anwendung des Kalman-Filters (4.2)

2) Korrekturschritt: c) $\underline{P}(k) = \tilde{\underline{P}}(k) - \underline{K}(k) \cdot C(k) \cdot \tilde{\underline{P}}(k)$

$$\underline{P}(k) = \begin{bmatrix} 3564 & 5510 \\ 5510 & 10010 \end{bmatrix} - \begin{bmatrix} 1,00 \\ 1,55 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} 3564 & 5510 \\ 5510 & 10010 \end{bmatrix}$$

$$\underline{P}(k) = \begin{bmatrix} 3564 & 5510 \\ 5510 & 10010 \end{bmatrix} - \begin{bmatrix} 3564 & 5510 \\ 5524 & 8540 \end{bmatrix}$$

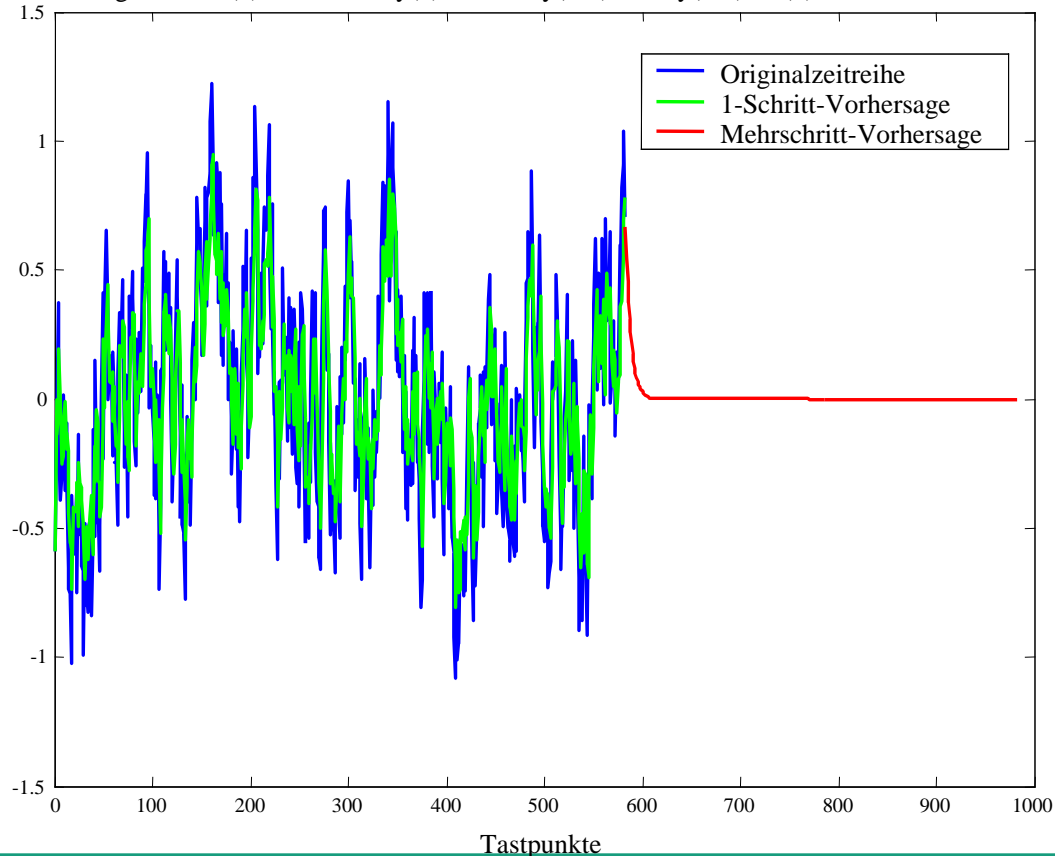
$$\underline{P}(k) = \begin{bmatrix} 0 & 0 \\ -14 & 1470 \end{bmatrix}$$

3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.3 Vorhersage mit stochastischen Signalmodellen – Vorhersage eines AR(2)-Prozesses mit dem Kalman-Filters (1)

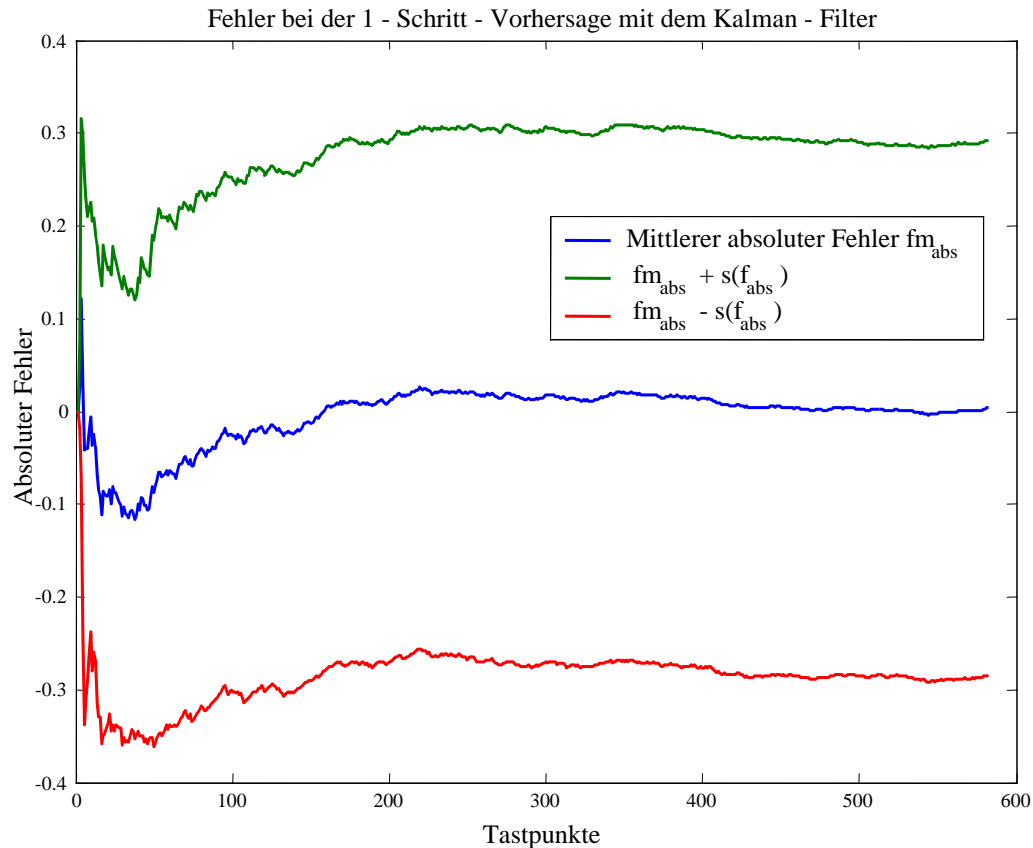
Vorhersage des AR(2) - Prozesses $y(k) - 0.55 \cdot y(k-1) - 0.23 \cdot y(k-2) = \varepsilon(k)$ mit dem Kalman-Filter



3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.3 Vorhersage mit stochastischen Signalmodellen – Vorhersage eines AR(2)-Prozesses mit dem Kalman-Filters (2)

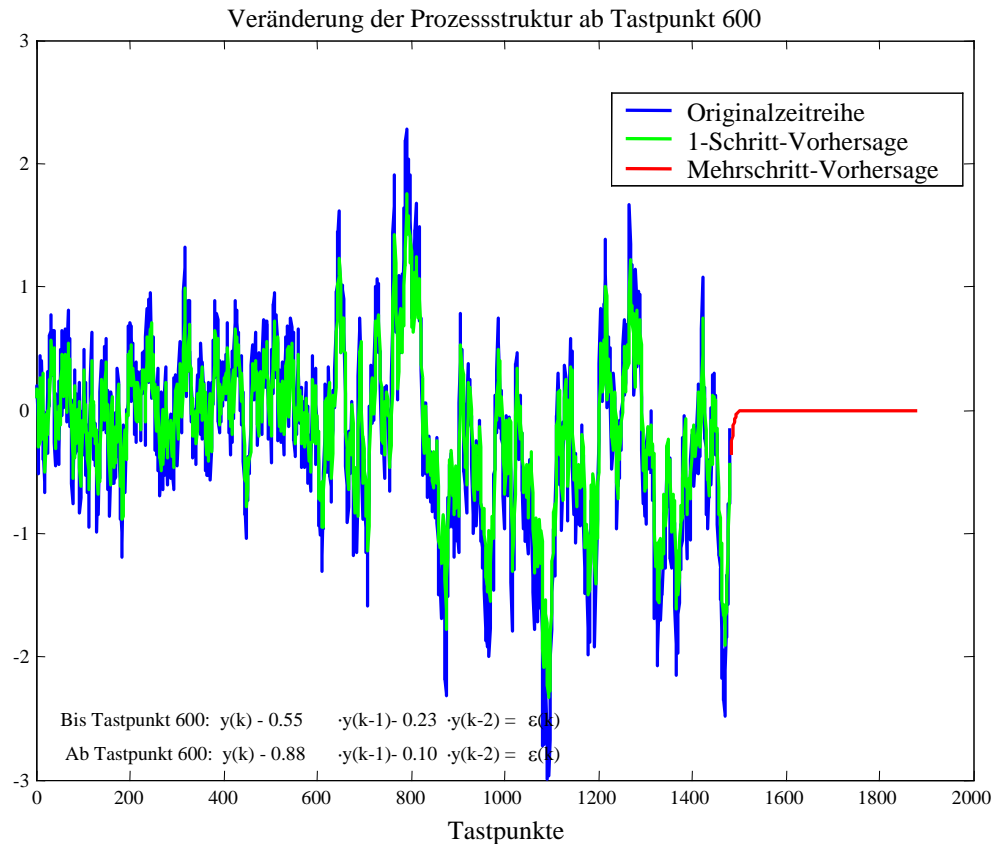


3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.3 Vorhersage mit stochastischen Signalmodellen –

Vorhersage mit dem Kalman-Filters bei sich verändernden Prozessverhalten (1)

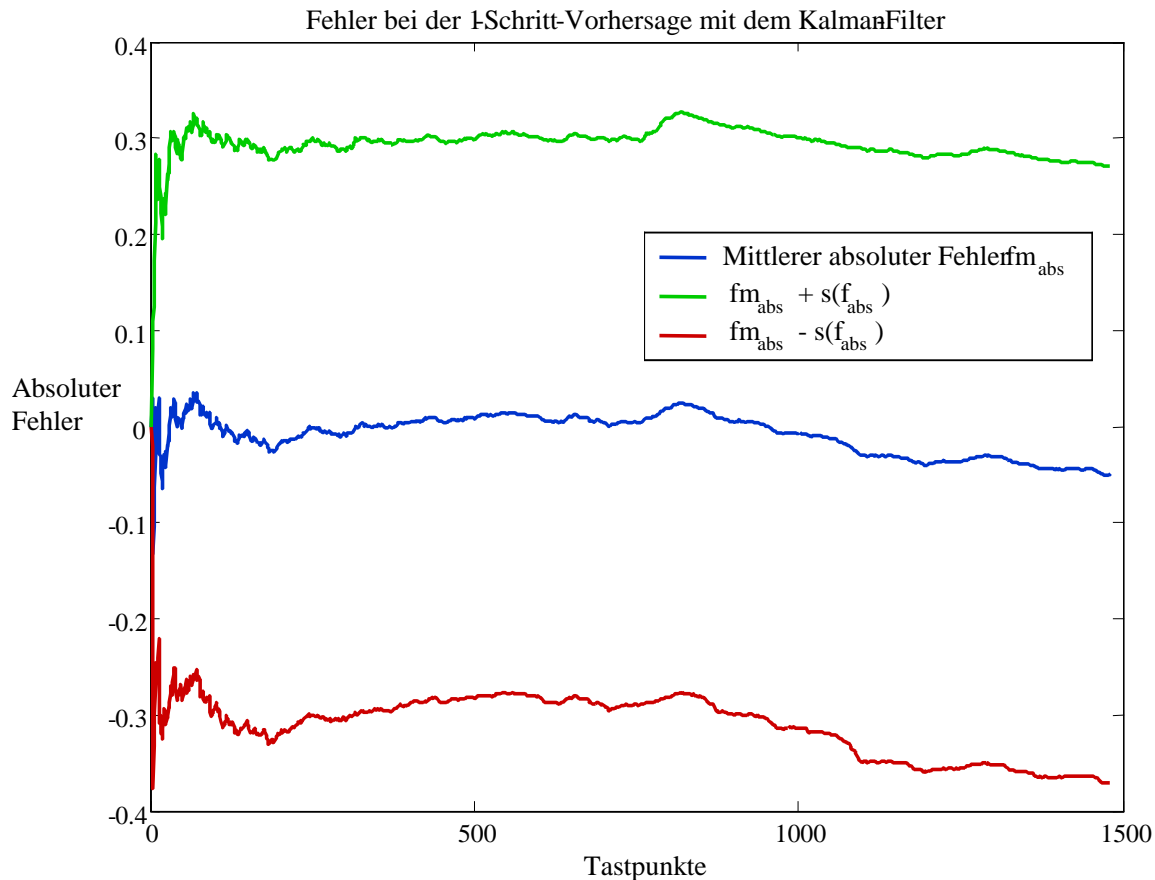


3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.3 Vorhersage mit stochastischen Signalmodellen –

Vorhersage mit dem Kalman-Filters bei sich verändernden Prozessverhalten (2)



3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.3 Vorhersage mit stochastischen Signalmodellen – Kalman-Filter mit Prozessparametern als Zustand (1)

Zustandsraumdarstellung des AR(2)-Prozesses mit den Parametern als Zustandsvektor:

$$y(k) - 0,55y(k-1) - 0,23y(k-2) = \eta(k)$$

Zustandsvektor: $x(k) = \begin{bmatrix} -a_1(k) \\ -a_2(k) \end{bmatrix}$



Systemgleichung: $\begin{bmatrix} -a_1(k+1) \\ -a_2(k+1) \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} -a_1(k) \\ -a_2(k) \end{bmatrix}$



Beobachtungsgleichung: $y(k) = [y(k-1) \quad y(k-2)] \begin{bmatrix} -a_1(k) \\ -a_2(k) \end{bmatrix} + \eta(k)$

Somit gilt: $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

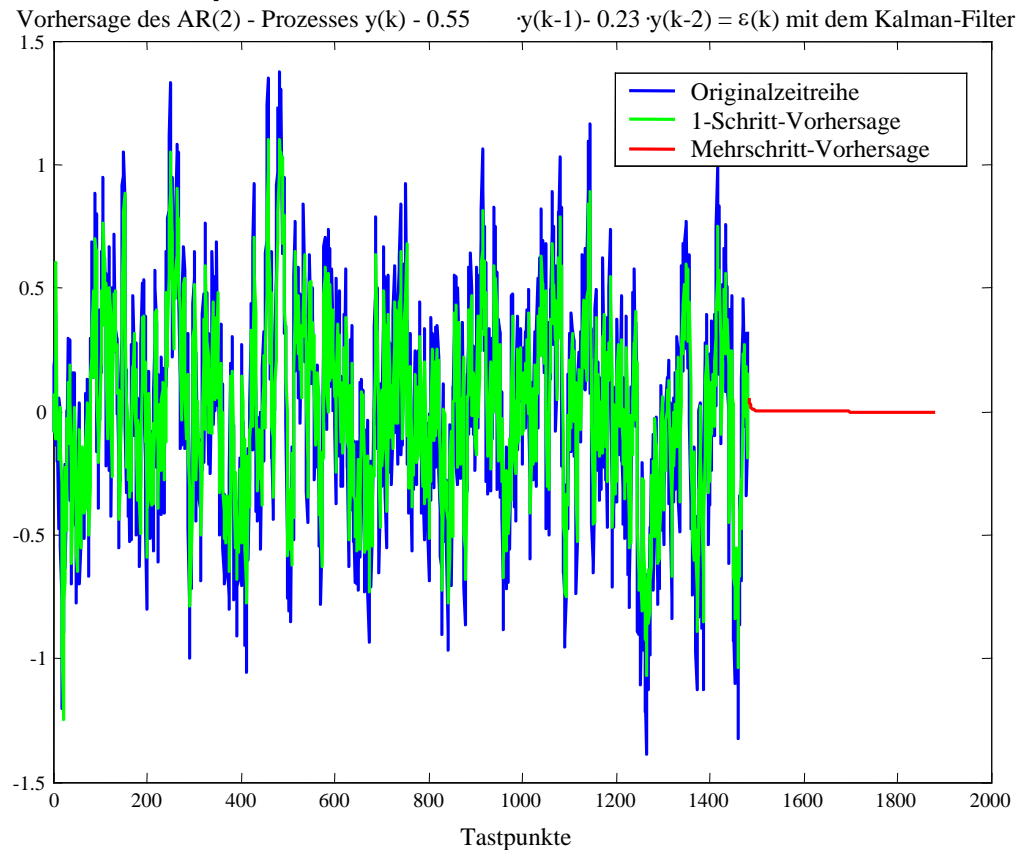
$$B = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$C = [y(k-1) \quad y(k-2)]$$

3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

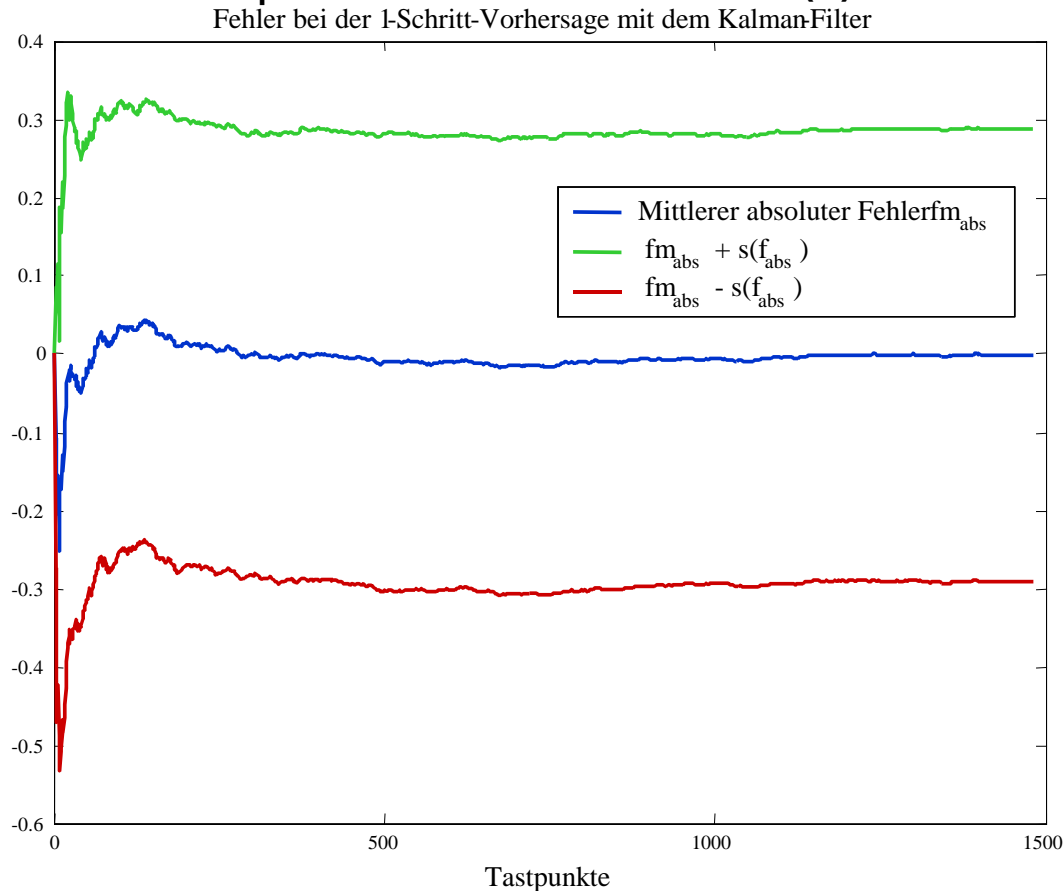
3.6.3 Vorhersage mit stochastischen Signalmodellen – Kalman-Filter mit Prozessparametern als Zustand (2)



3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.3 Vorhersage mit stochastischen Signalmodellen – Kalman-Filter mit Prozessparametern als Zustand (3)

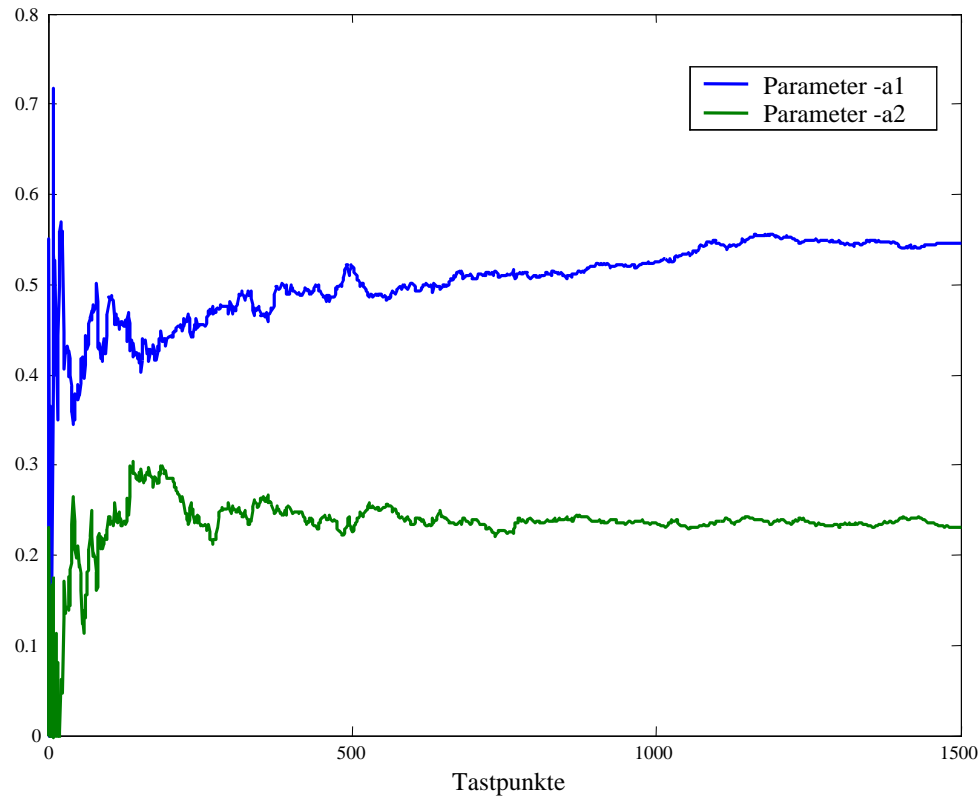


3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.3 Vorhersage mit stochastischen Signalmodellen – Kalman-Filter mit Prozessparametern als Zustand (4)

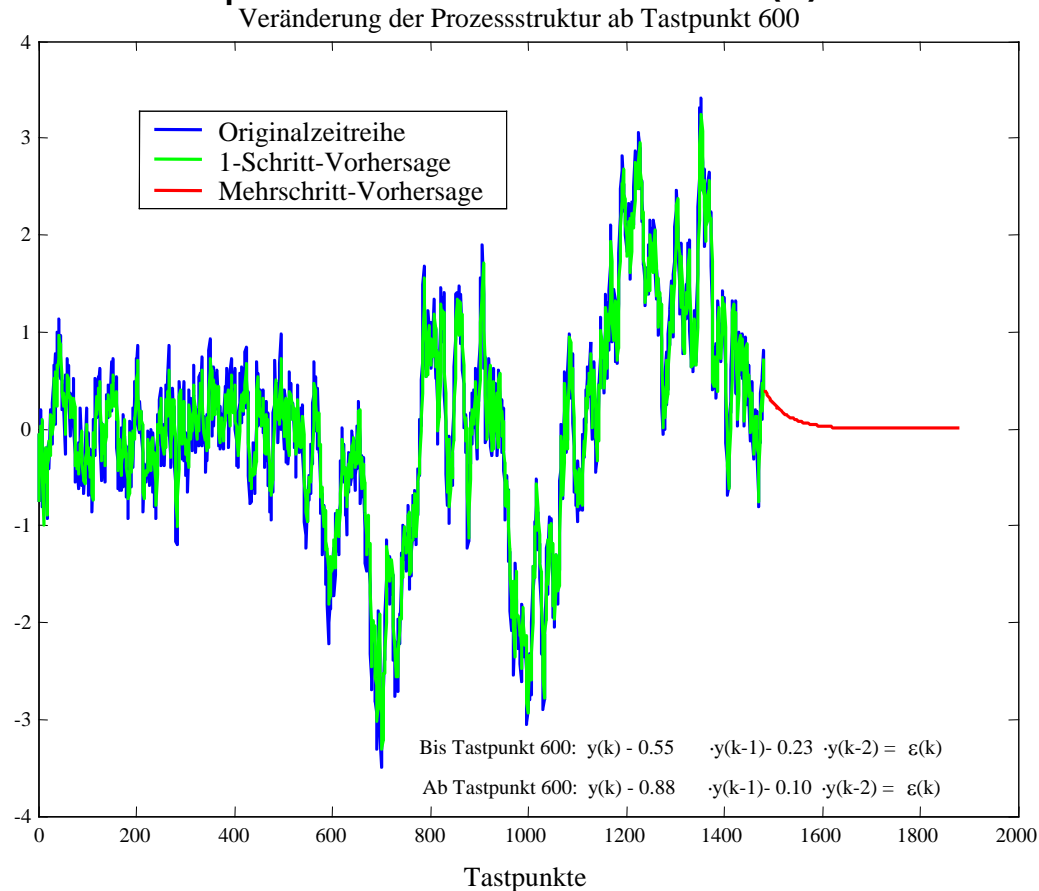
Verlauf der mit dem Kalman-Filter geschätzten Parameter



3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

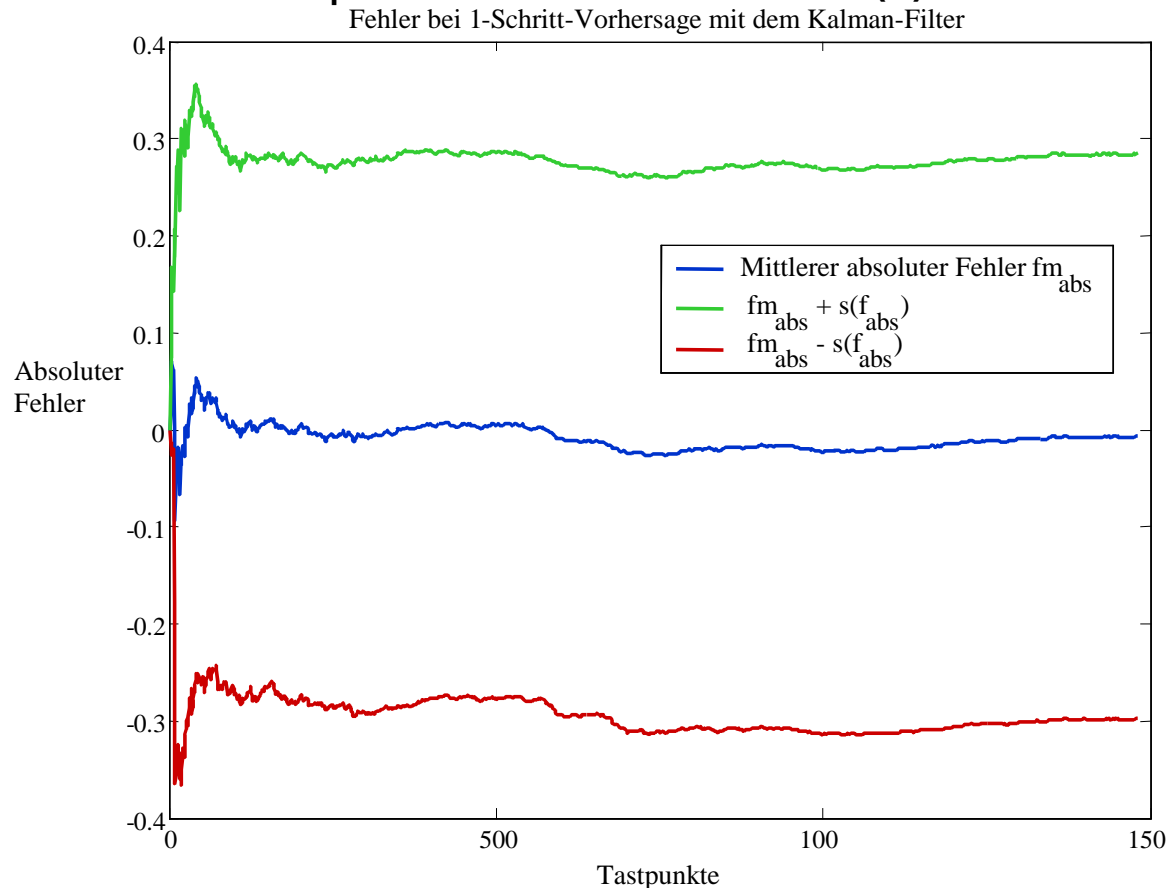
3.6.3 Vorhersage mit stochastischen Signalmodellen – Kalman-Filter mit Prozessparametern als Zustand (5)



3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.3 Vorhersage mit stochastischen Signalmodellen – Kalman-Filter mit Prozessparametern als Zustand (6)

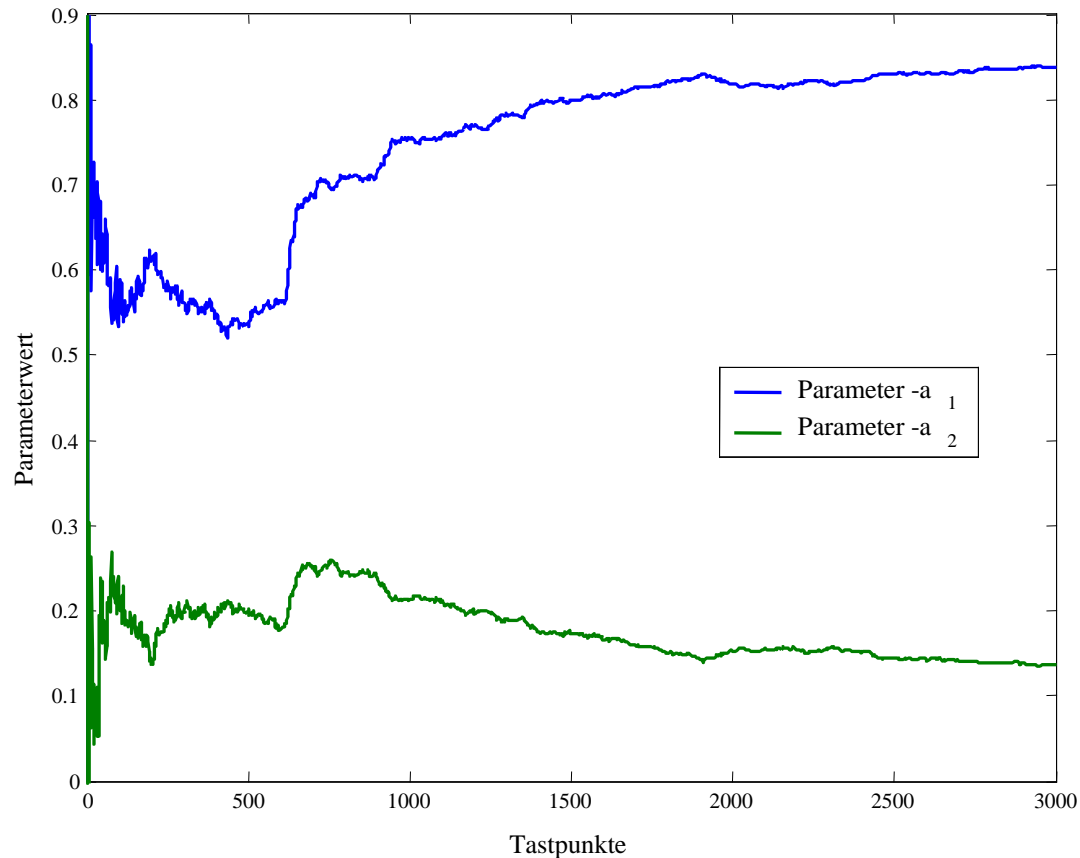


3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.3 Vorhersage mit stochastischen Signalmodellen – Kalman-Filter mit Prozessparametern als Zustand (7)

Verlauf der mit dem Kalman-Filter geschätzten Parameter



3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen


3.6.3 Vorhersage mit stochastischen Signalmodellen – Kalman-Filter mit Prozessparametern als Zustand (8)

Ausgangspunkt: Zustandsraumdarstellung mit Prozessparametern als Zustand.

$$\begin{aligned}\underline{\tilde{x}}(k+1) &= \underline{A}(k)\underline{\hat{x}}(k) + \underline{B} \cdot \underline{\varepsilon}(k+1) & \text{mit } E\{\underline{\varepsilon}(k)\} &= 0 & E\{\underline{\eta}(k)\} &= 0 \\ \underline{\tilde{y}}(k) &= \underline{C}(k)\underline{\tilde{x}}(k) + \underline{\eta}(k) & \text{Var}\{\underline{\varepsilon}(k)\} &= \underline{Q} = 0 & \text{Var}\{\underline{\eta}(k)\} &= R = 1 \\ & & \underline{A}(k) &= \underline{I}\end{aligned}$$

folgt: $\tilde{x}(k+1) = \hat{x}(k)$

weiterhin erhält man für $\tilde{P}(k) = \underline{A}(k-1)\underline{P}(k-1)\underline{A}'(k-1) + \underline{B}(k-1) \cdot \underline{Q}(k-1) \cdot \underline{B}'(k-1)$
 $\tilde{P}(k) = \underline{P}(k-1)$

 aus Korrekturschritt (1) $\underline{P}(k) = \underline{P}(k-1) - \underline{K}(k) \cdot \underline{C}(k) \cdot \underline{P}(k-1)$

$$(2) \quad \underline{K}(k) = \underline{P}(k-1) \cdot \underline{C}'(k) (\underline{C}(k) \cdot \underline{P}(k-1) \cdot \underline{C}'(k) + 1)^{-1}$$

$$(3) \quad \hat{x}(k) = \hat{x}(k-1) + \underline{K}(k)(y(k) - \underline{C}(k) \cdot \hat{x}(k-1))$$

Da $\underline{C}(k)$ die Messwerte enthält, kann sie als Messwertmatrix $x \underline{m}(k)$ aufgefasst werden. Dann entsprechen die Gleichungen (1) – (3) exakt der Methode der rekursiven Regression.

3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.4 Musterbasierte Vorhersage

Grundlagen der musterbasierten Vorhersage

Einsatzgebiet: Vorhersage nichtlinearer stochastischer Signale, die periodische oder nichtperiodische Muster enthalten. Ziel ist die Linearisierung des Signals zur Erzeugung approximativer stationärer Signaleigenschaften. Damit ist die Vorhersage des Signalverlaufs mit klassischen Methoden möglich.

Definition „musterbasiertes Signal“:

$\{x(k)\}$ ist ein musterbasiertes Signal, wenn Teilbereiche des Signals durch typische, wiederkehrende Teilverläufe klassifizierbar oder beschreibbar sind. Diese Teilverläufe werden als Muster oder auch Musterobjekt bezeichnet und können linearer und nichtlinearer Gestalt sein. Ihr Auftreten erfolgt entweder periodisch oder nichtperiodisch.

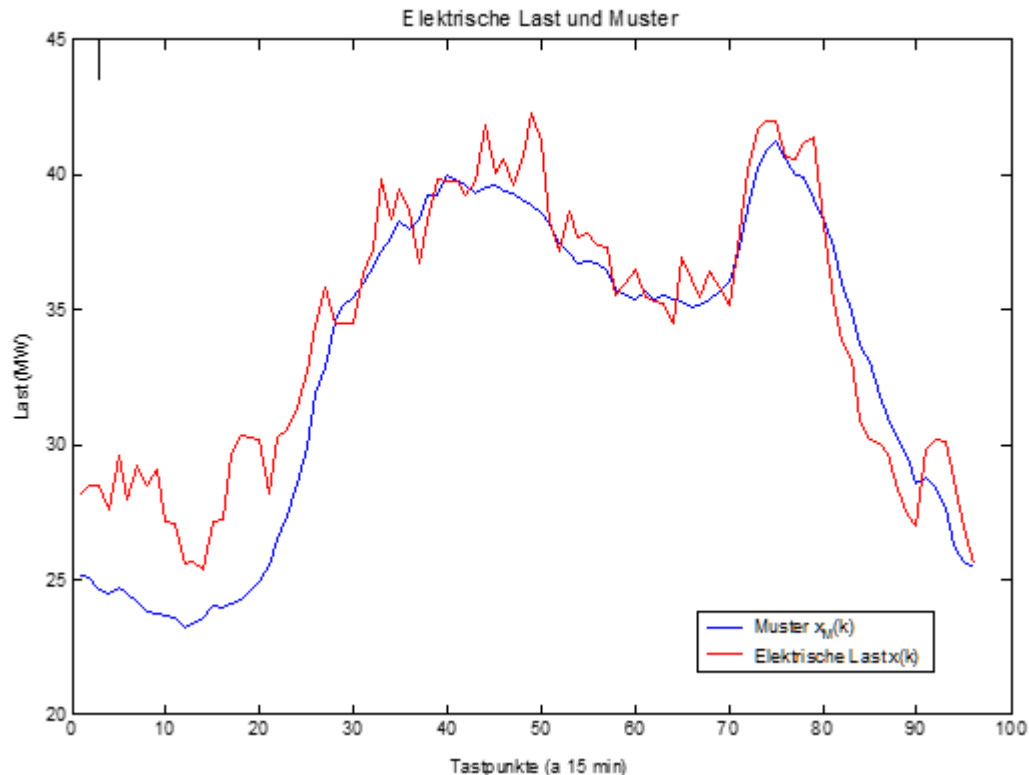
Daraus ergibt sich das erweiterte Komponentenmodell:

$$x(k) = x_T(k) + x_P(k) + x_M(k) + x_S(k)$$

3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

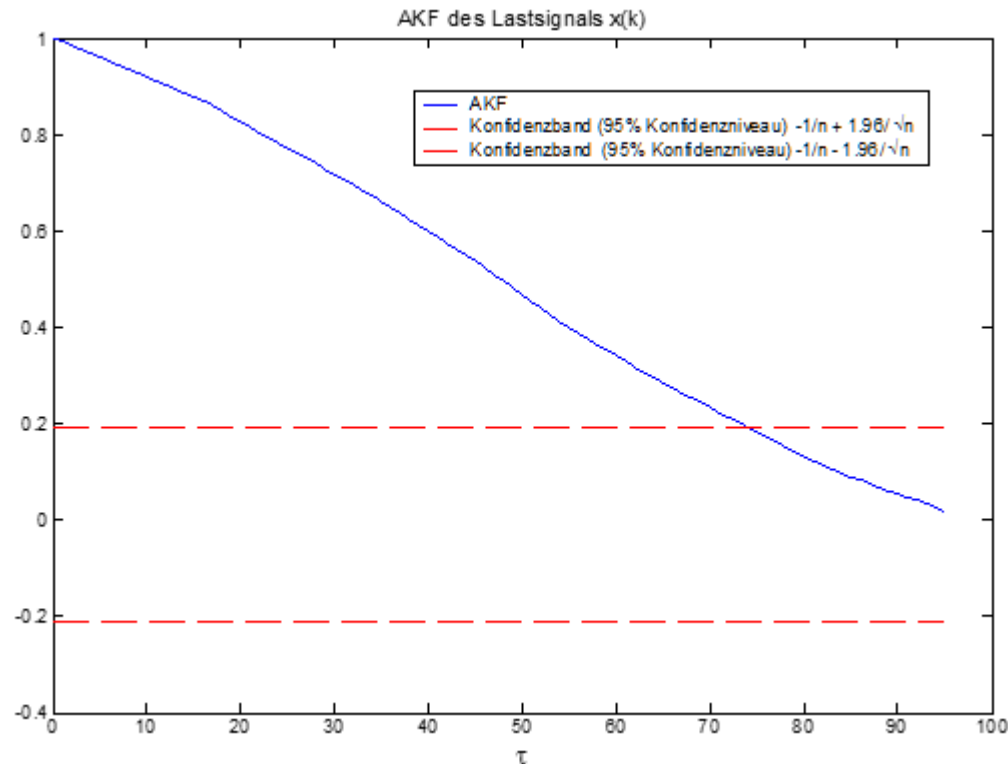
3.6.4 Musterbasierte Vorhersage – Signaldekomposition mit Musterkomponente (1.1)



3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

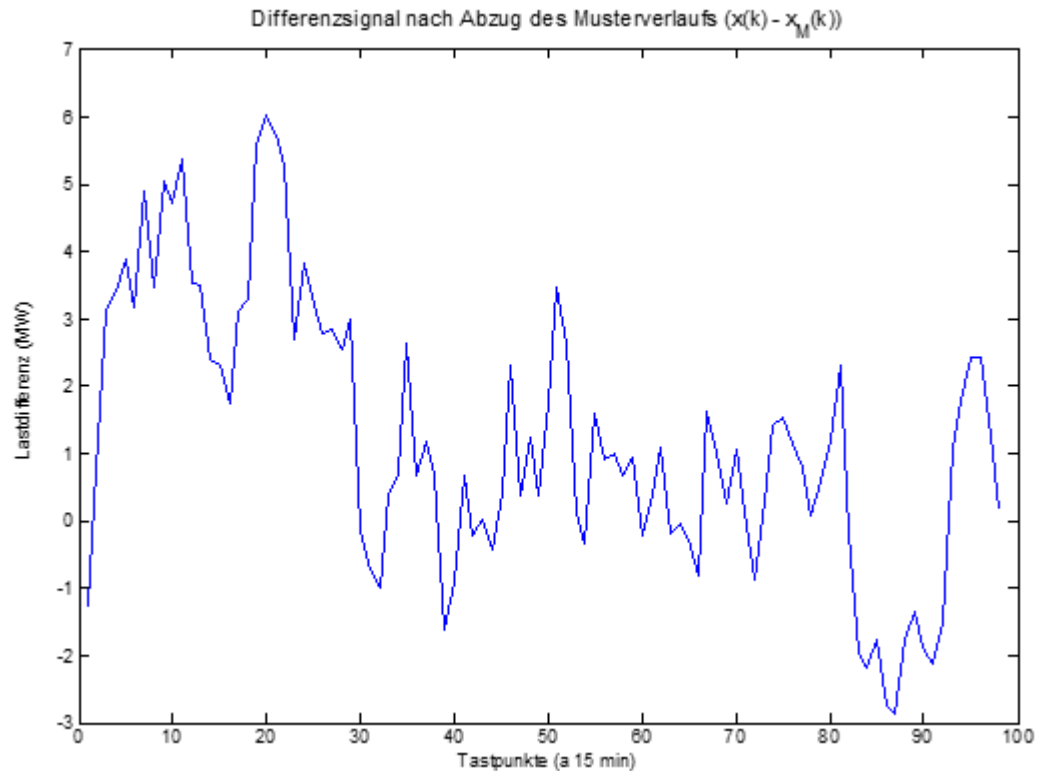
3.6.4 Musterbasierte Vorhersage – Signaldekomposition mit Musterkomponente (1.2)



3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

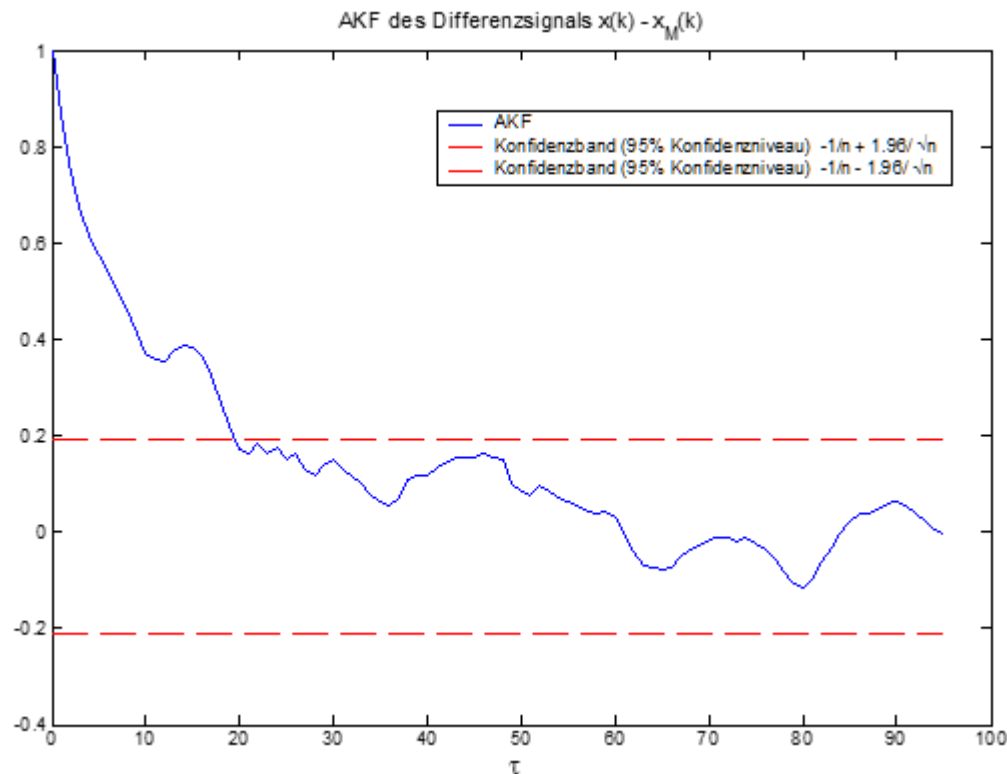
3.6.4 Musterbasierte Vorhersage – Signaldekomposition mit Musterkomponente (2.1)



3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

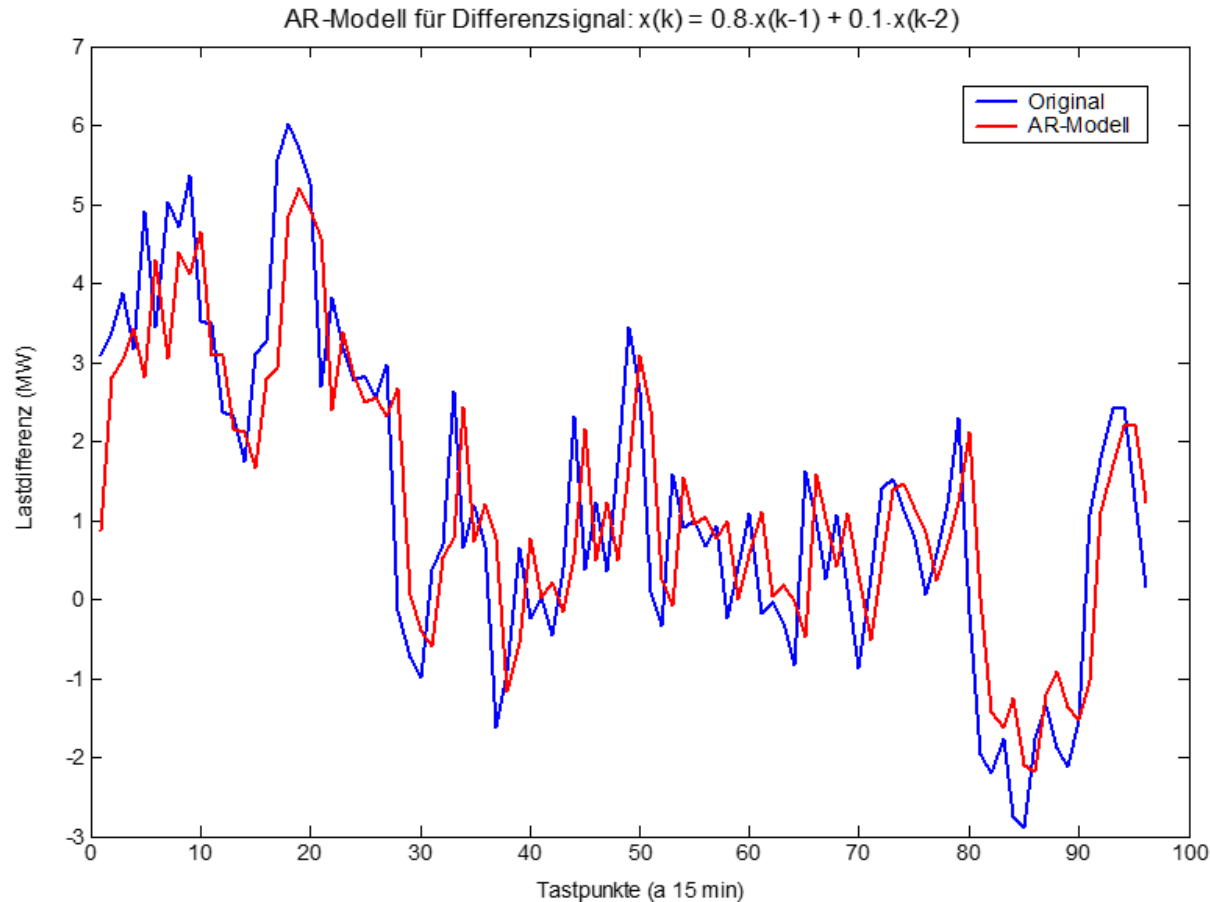
3.6.4 Musterbasierte Vorhersage – Signaldekomposition mit Musterkomponente (2.2)



3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

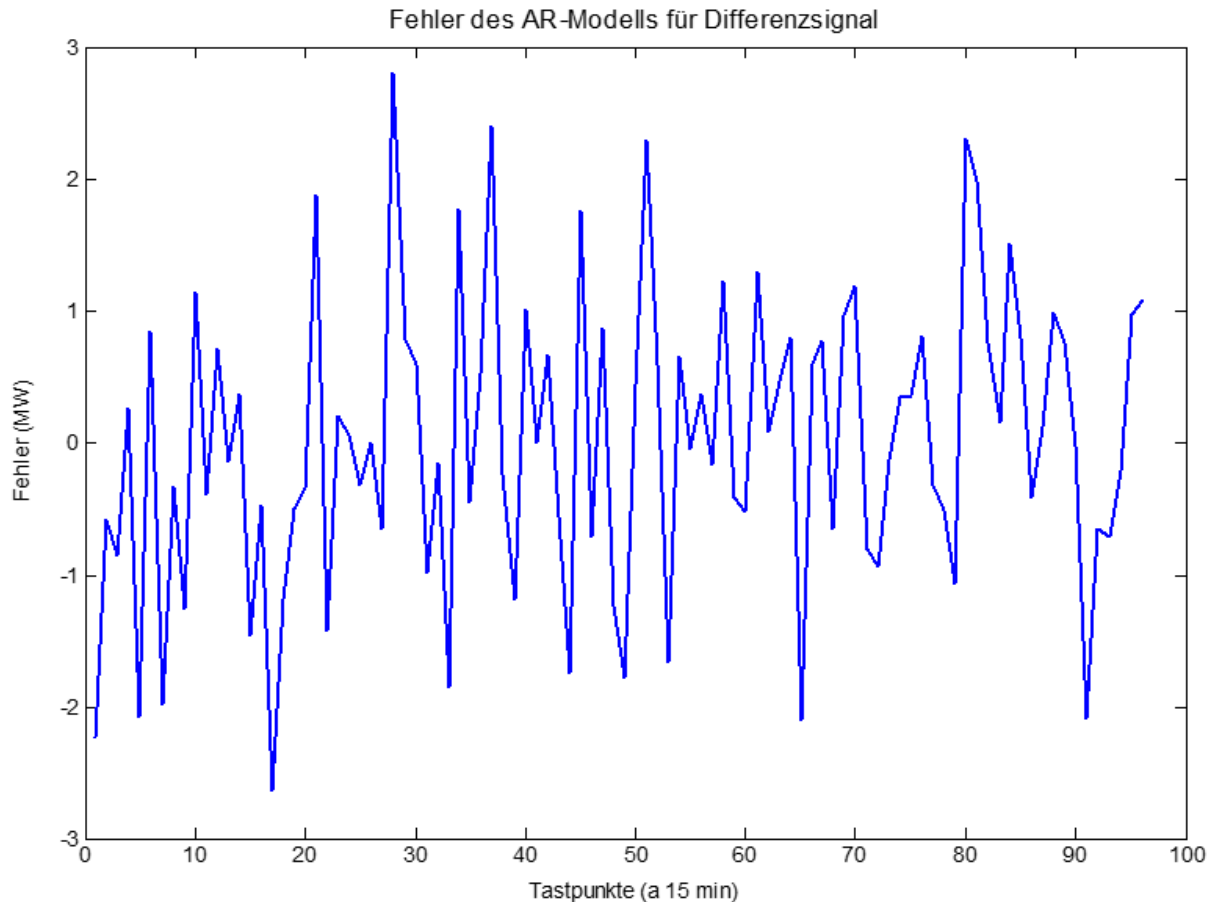
3.6.4 Musterbasierte Vorhersage – Signaldekomposition mit Musterkomponente (3.1)



3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

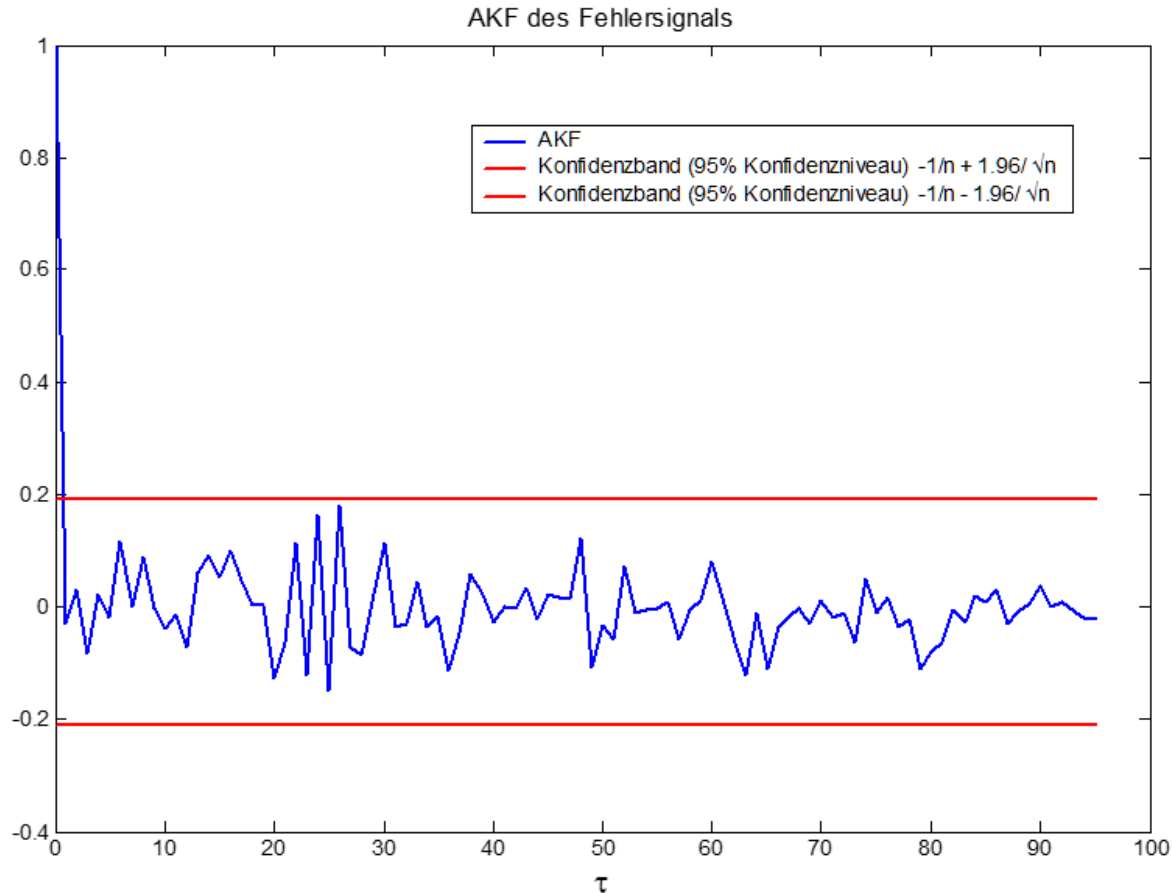
3.6.4 Musterbasierte Vorhersage – Signaldekomposition mit Musterkomponente (3.2)



3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.4 Musterbasierte Vorhersage – Signaldekomposition mit Musterkomponente (3.3)

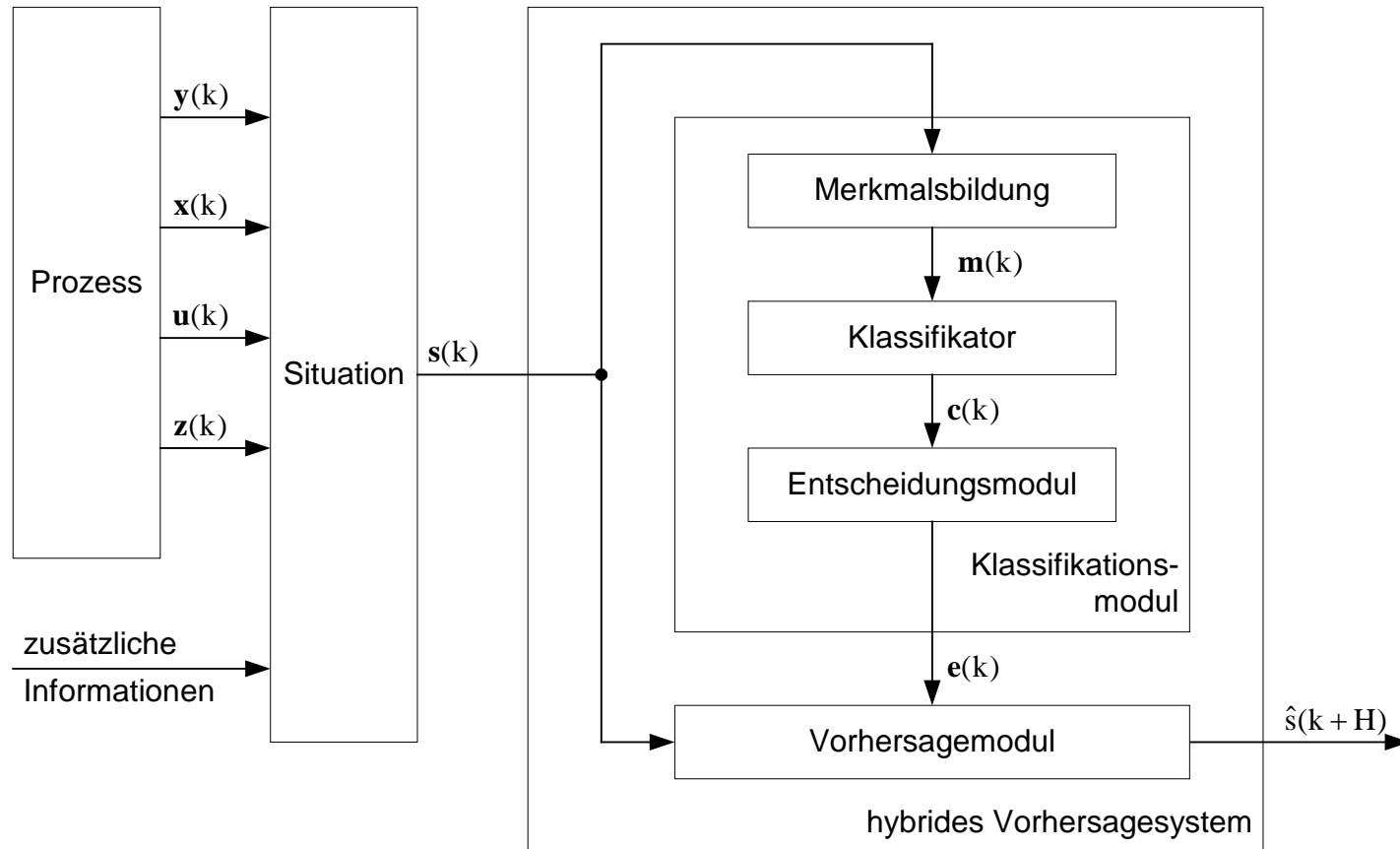


3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.4 Musterbasierte Vorhersage

Struktur des hybriden Vorhersagesystems zur Prognose musterbasierter Signale



3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

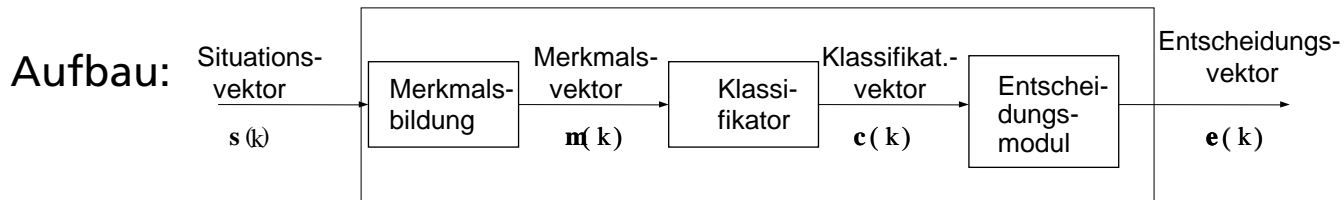
3.6.4 Musterbasierte Vorhersage

Klassifikationsmodul

Aufgabe: Einordnung der aktuellen Prozesssituation zu einer hinterlegten Klassenstruktur und anschließende Generierung eines Entscheidungsvektors als zusätzliche Eingangsgröße für das nachfolgende Vorhersagemodul.

$$\underline{e}(k) = f(s(k))$$

mit $\underline{e}(k)$ – Entscheidungsvektor
 $\underline{s}(k)$ – Situationsvektor



Klassifikator:

- Fuzzy-Klassifikator
- Abstandsklassifikator

Entscheidungsmodul: Regelwerk zur Verarbeitung des Klassifikationsvektors $\underline{c}(k)$.

3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.4 Musterbasierte Vorhersage Merkmalsbildung (1)

Aufgabe:

Ermittlung und Auswahl der Merkmale, die eine maximale Objektseparation unter Beachtung der aufgabenspezifischen Randbedingungen gewährleisten.

Bei der Merkmalsbildung wird zwischen Merkmalsauswahl und Merkmalstransformation unterschieden.

Merkmalsauswahl:

Nur eine Teilmenge der Elemente des Situationsvektors wird zum Merkmalsvektor zusammengefasst. Die Auswahl kann z. B. nach folgenden Gesichtspunkten erfolgen:

- Rangfolge nach den Gütewerten der Einzelmerkmale,
- Rangfolge nach dem Beitrag eines Merkmals unter Beachtung des mit den bereits ausgewählten Merkmalen erreichten Gütewertes (Anstiegsverfahren),
- Rangfolge nach dem Güteverlust, der sich bei Weglassen eines Merkmals ergibt (Abstiegsverfahren).

Ziel ist die Auswahl der Untermengen m aus n Merkmalen, die für die Klassifikation eine maximale Objektrennung ermöglichen.

3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.4 Musterbasierte Vorhersage Merkmalsbildung (2)

Merkmalsstransformation:

Grundidee ist eine Transformation des Situationsvektors \underline{s} in einen anderen Vektorraum \underline{m} , um eine bessere Darstellung für die Klassifikation der Objekte zu erreichen. Mögliche Methoden sind:

- Normierung
- Standardisierung
- Wichtung
- lineare und nichtlineare Transformationen
- Betrag
- Trend
- Maximum
- Minimum

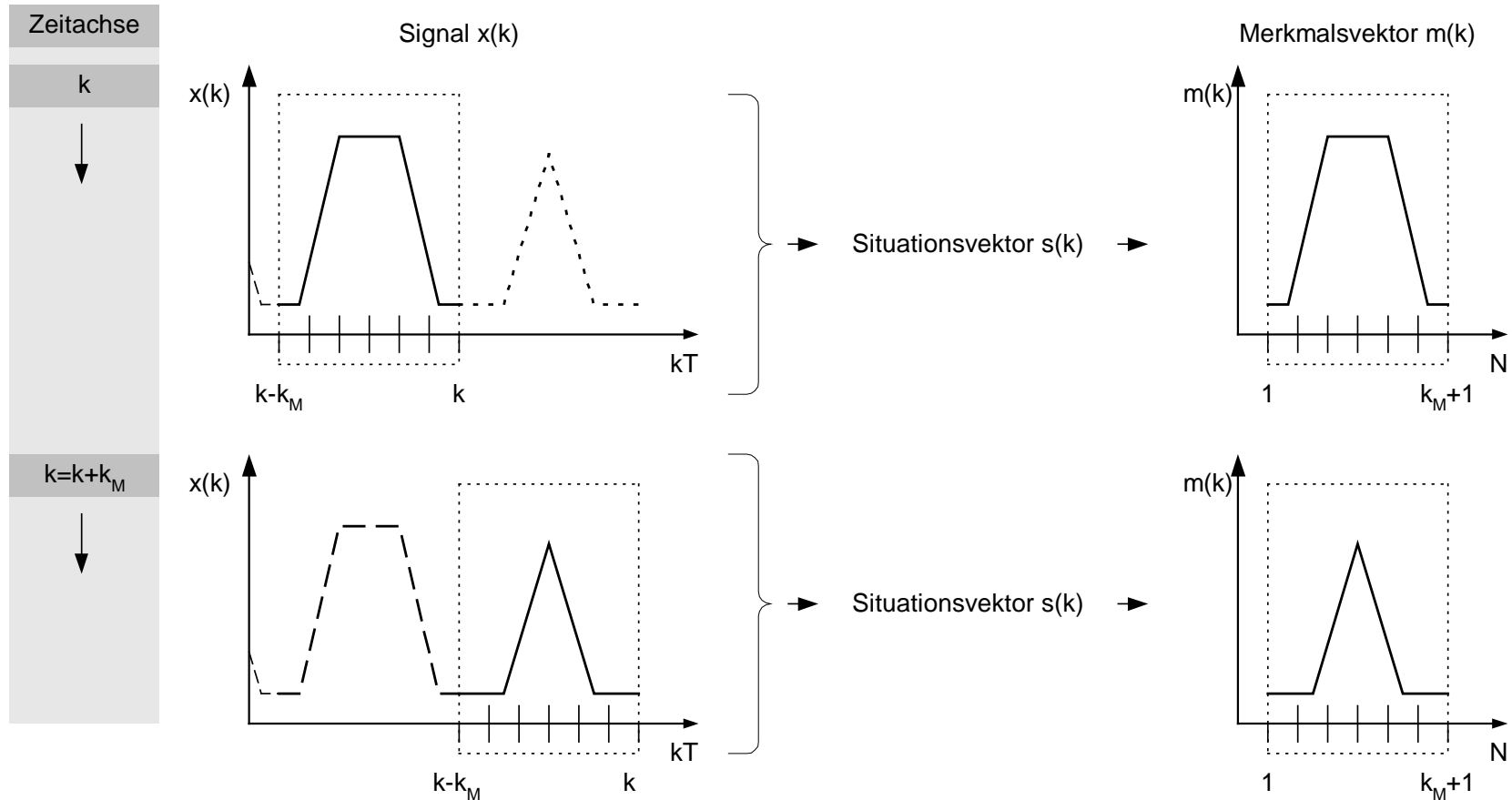
Kriterien zur Auswahl der Verfahren:

- Merkmalsauswahl: Wenn a-priori-Informationen vorliegen oder auf Grund der Objektdarstellung intuitiv Merkmale ausgeschlossen werden können.
- Merkmalsstransformation: Sehr rechenintensives Verfahren. Wird eingesetzt, wenn zur Reduktion der Dimensionalität keine anderen begründeten Vorschriften vorhanden sind.

3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.4 Musterbasierte Vorhersage - Merkmalsbildung (3)

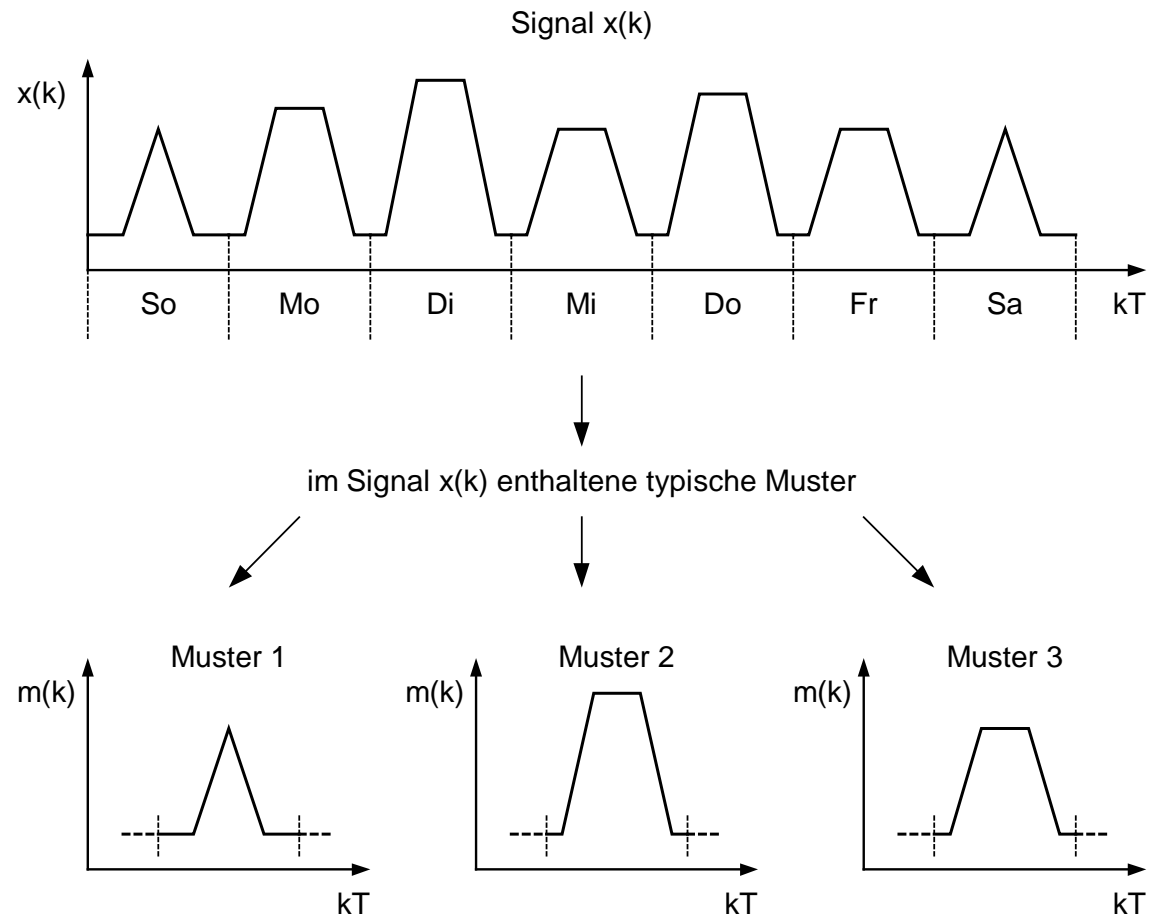


3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.4 Musterbasierte Vorhersage - Merkmalsbildung (4)

Wochenverlauf eines musterbasierten Signals inklusive typischer Muster:

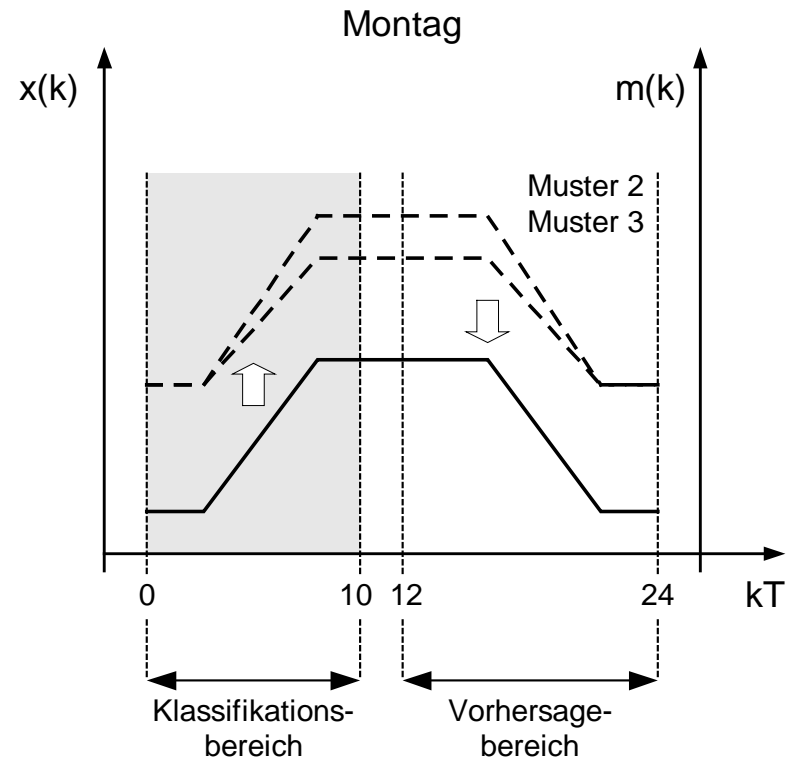
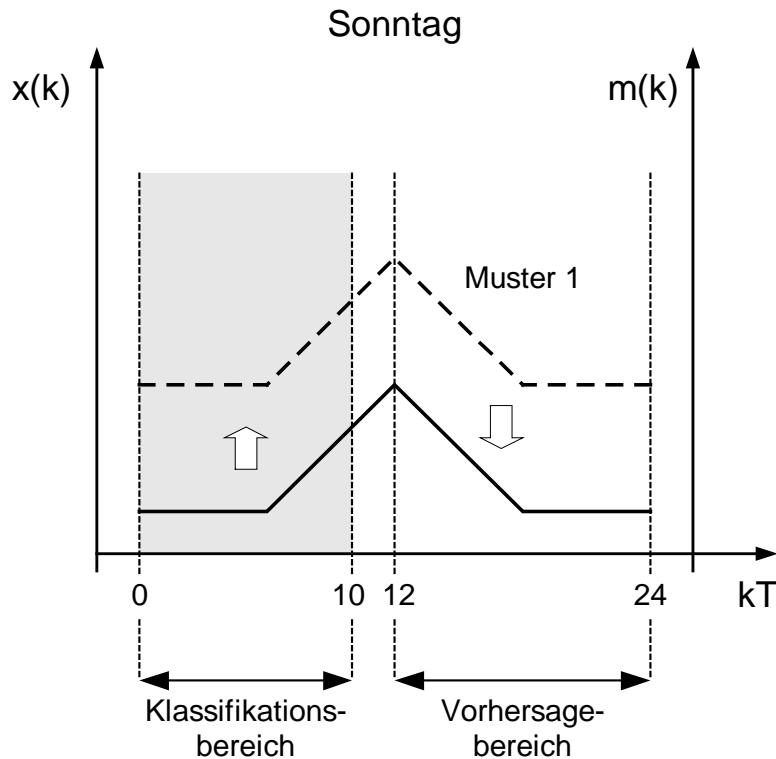


3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.4 Musterbasierte Vorhersage - Merkmalsbildung (5)

Beispiel der Musterklassifikation für den Vorhersagebereich 12:00 bis 24:00 Uhr



3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.4 Musterbasierte Vorhersage Merkmalsbildung (6)

Methoden für Untersuchungen und Analysen zur Merkmalsbildung:

- 1) Grafische Darstellung des Signals, insbesondere übersichtlicher Teilbereiche.
- 2) Bestimmung typischer Rhythmen mittels der Autokorrelationsfunktion.
- 3) Anwendung von Hauptkomponenten- und Diskriminanzanalyse zur Ermittlung typischer Klassenkonfigurationen.
- 4) Anwendung hierarchischer und partitionierender Clusterverfahren zur Ermittlung von Musterobjekten sowie deren Gruppierung zu Klassen als Repräsentanten typischer im Signalverlauf vorkommender Muster.

3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.4 Musterbasierte Vorhersage - Klassifikator (1)

Aufgabe: Mit Hilfe einer Vorschrift bzw. eines Algorithmus wird die aktuelle Prozesssituation einer oder mehreren Klassen zugeordnet. Der Merkmalsvektor \underline{m} ist die Eingangsgröße des Klassifikators, der über eine festgelegte Vorschrift den Klassifikationsvektor \underline{c} als Ausgangsgröße generiert: $\underline{c}(k) = f(m(k))$

Klassifikatortypen:

- Bayes-Klassifikator
- Abstands-Klassifikator
- Trennfunktions-Klassifikator
- Deterministischer Klassifikator
- Fuzzy-Klassifikator
- Neuronale Netze

Aufgabe des Klassifikators bei musterbasierter Vorhersage:

- 1) Beschreibung der Signalmuster durch Musterobjekte und Gruppierung zu typischen Klassen.
- 2) Einordnung der aktuellen Prozesssituation (Musterobjekt) zu den Klassen der Klassenkonfiguration in Form des Klassifikationsvektors.

3. Vorhersagestrategien

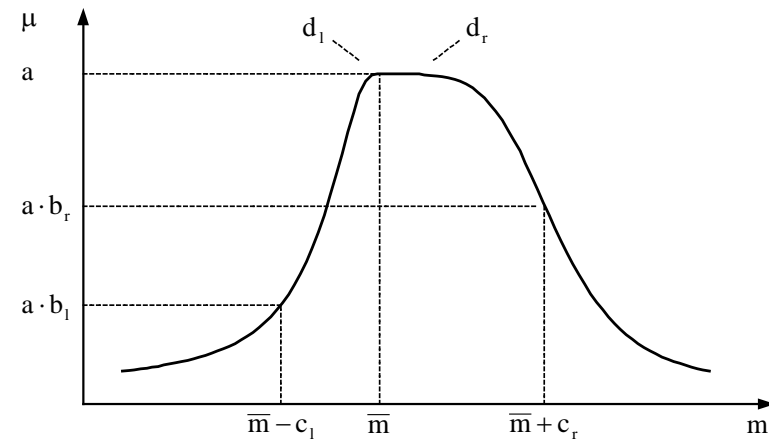
3.6 Vorhersage mit Signalmodellen

3.6.4 Musterbasierte Vorhersage - Klassifikator (2)

Fuzzy- Klassifikator:

mit $m'_i = m_i - \bar{m}_i$
 m_i
 \bar{m}_i
 M
 $\underline{p} = \{a, \underline{b}_l, \underline{b}_r, \underline{c}_l, \underline{c}_r, \underline{d}_l, \underline{d}_r\}$

- i -tes zentriertes Merkmal,
- i -tes Merkmal des Merkmalsvektor \underline{m} ,
- i -tes Merkmal einer Klasse, (auch i -ter Klassenschwerpunkt),
- Merkmalsanzahl,
- Parametervektor, fasst alle Parameter der Zugehörigkeitsfunktion zusammen.



Das Ergebnis der parametrischen Zugehörigkeitsfunktion $\mu(\underline{m})$ ist der Sympathiewert. Dieser beschreibt die Ähnlichkeit des Musterobjektes \underline{m} zu der Klasse C_K .

$$\mu(\underline{m}) = a \cdot \left\{ 1 + \sum_{i=1}^M \frac{1}{2} \cdot (1 + \operatorname{sgn}(-m'_i)) \cdot \left(\frac{1}{b_{l_i}} - 1 \right) \cdot \left| \frac{m'_i}{c_{l_i}} \right|^{d_{l_i}} + \sum_{i=1}^M \frac{1}{2} \cdot (1 + \operatorname{sgn}(m'_i)) \cdot \left(\frac{1}{b_{r_i}} - 1 \right) \cdot \left| \frac{m'_i}{c_{r_i}} \right|^{d_{r_i}} \right\}^{-1}$$

3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.4 Musterbasierte Vorhersage - Klassifikator (3)

Vorteile des Fuzzy-Klassifikators bei der musterbasierten Vorhersage:

- 1) Klassen werden als unscharfe Mengen beschrieben und ermöglichen somit die Behandlung unsicherer und stochastischer Daten.
- 2) Fehlende Merkmale bereiten dem Fuzzy-Klassifikator im Rahmen der Mustererkennung keine Probleme. Die betreffenden Merkmale können ohne weiteres aus dem Ansatz gestrichen werden, da die Parameter zur Beschreibung der Merkmale unabhängig voneinander bestimmt wurden und somit das Entwurfskriterium des Klassifikators erfüllt bleibt.
- 3) Erfassung nichtlinearer Musterobjekte ist problemlos und wird durch den Merkmalsvektor in Form einer nichtlinearen Kennlinie abgebildet.
- 4) Die Verarbeitung hochdimensionaler Vektorräume und somit hochdimensionaler Muster ist möglich.
- 5) Anpassung des Klassifikators an aktuelle Prozesssituationen ist auf Grund des parametrischen Konzepts gegeben. Es existieren rekursive Algorithmen zur Beeinflussung der klassenbeschreibenden Parameter.

3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.4 Musterbasierte Vorhersage Entscheidungsmodul

Aufgabe: Ermittlung des Entscheidungsvektors \underline{e} aus dem Klassifikationsvektor \underline{c} auf Grundlage eines festgelegten funktionellen Zusammenhangs: $\underline{e}(k) = f(c(k))$

z. B. $e(k) = \arg \max_{k'} c_{k'}(\underline{m}(k)) \quad k' = l \dots k$

mit $c_{k'}$ - k' tes Element des Klassifikationsvektors \underline{c} (Ähnlichkeit zur k' -ten Klasse)

Klassifikation und Entscheidung wird getrennt bearbeitet. Da verschiedenartige Entscheidungsvektoren generiert werden sollen, die z. B.

- Referenzmuster
 - Modellparameter oder
 - Auswahl zwischen unterschiedlichen Vorhersagestrategien
- enthalten.

Bei der musterbasierten Vorhersage entspricht der Entscheidungsvektor \underline{e} dem Referenzmuster.

Das Referenzmuster wird gebildet aus:

- dem Klassenschwerpunkt der Klasse mit dem höchsten Sympathiewert oder
- Gewichteter Mischung aller vorhandenen Klassenschwerpunkte (Wichtungsfaktor entspricht dem jeweiligen Sympathiewert)

3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.4 Musterbasierte Vorhersage

Vorhersagemodul (1)

Aufgabe: Im Rahmen der Strategie von globaler und lokaler Vorhersage dient es der Umsetzung der lokalen Vorhersage. Der Situationsvektor \underline{s} und der Entscheidungsvektor \underline{e} werden zum Vorhersagevektor \hat{s} verknüpft.

Ausprägungen: In Abhängigkeit vom Entscheidungsvektor \underline{e} sind folgende Ausprägungen möglich:

- a) Vorhersage mit Referenzmustervorgabe,
- b) Vorhersage mit Parameterumschaltung und
- c) Vorhersage mit Modellumschaltung.

Vorhersagemodelle:

- Deterministische Signalmodelle
- Stochastische Signalmodelle
- Neuronale Netze
- Fuzzy-Modelle

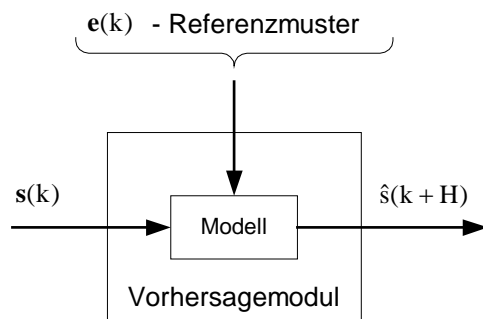
3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

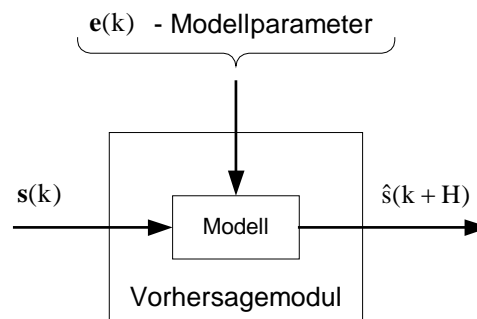
3.6.4 Musterbasierte Vorhersage

Vorhersagemodul (2)

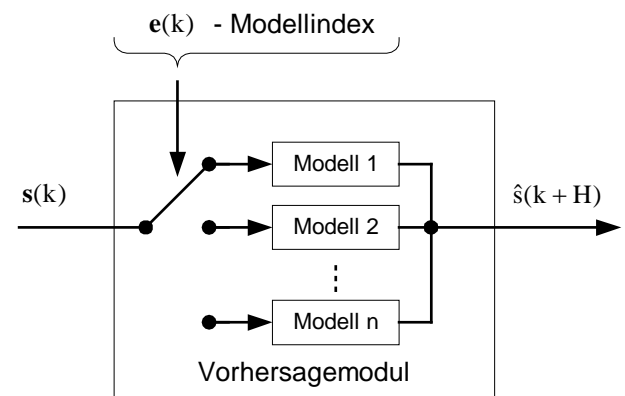
Strukturen des Vorhersagemoduls in Abhängigkeit zum Entscheidungsvektor



a) Referenzmustervorgabe



b) Parameterumschaltung



c) Modellumschaltung

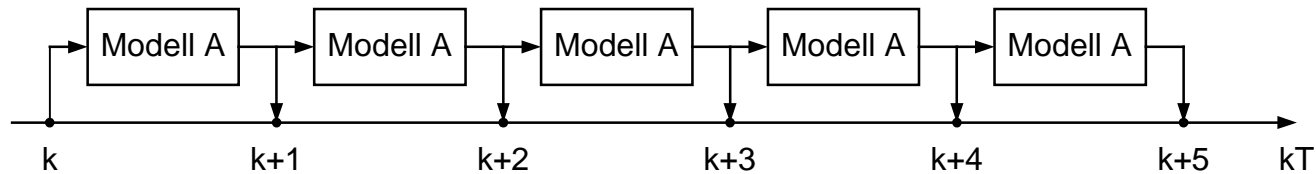
3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

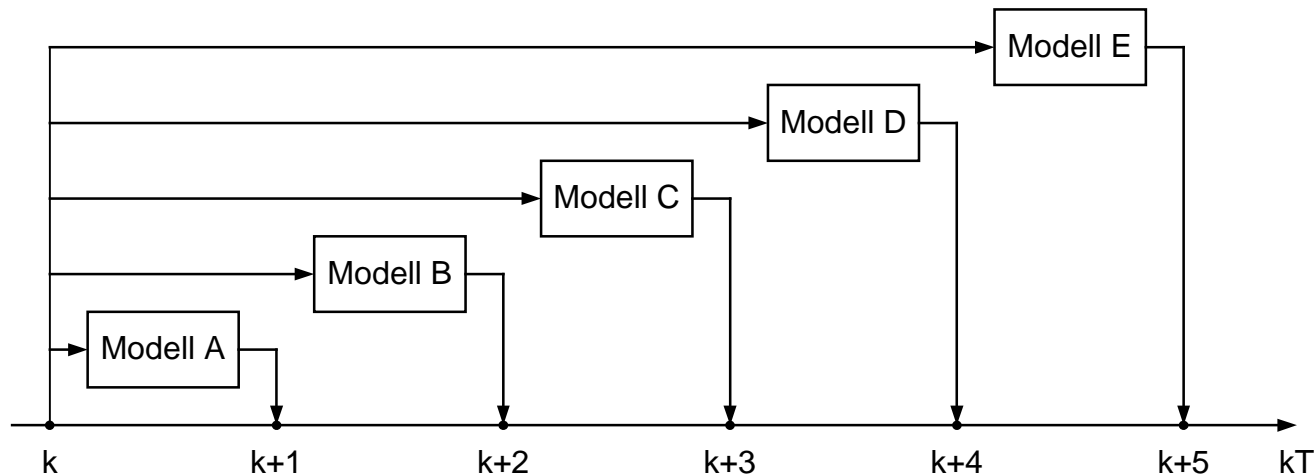
3.6.4 Musterbasierte Vorhersage

Vorhersagemodul (3)

Einschrittvorhersage



Mehrschrittvorhersage

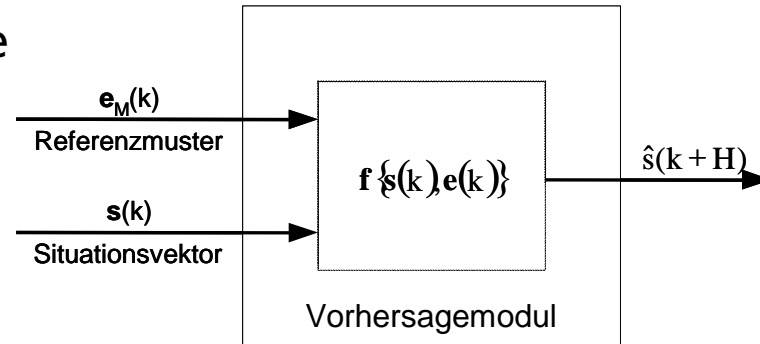


3. Vorhersagestrategien

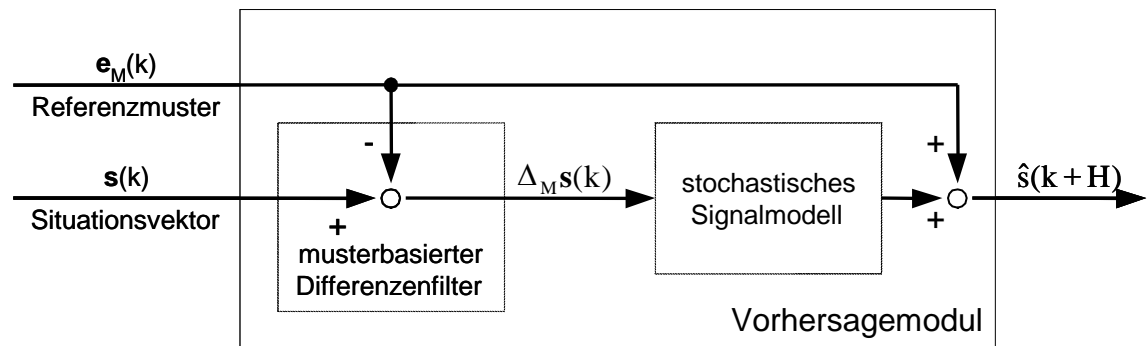
3.6 Vorhersage mit Signalmodellen

3.6.4 Musterbasierte Vorhersage

Vorhersagemodul (4)



a) Modell ohne Linearisierungskomponente



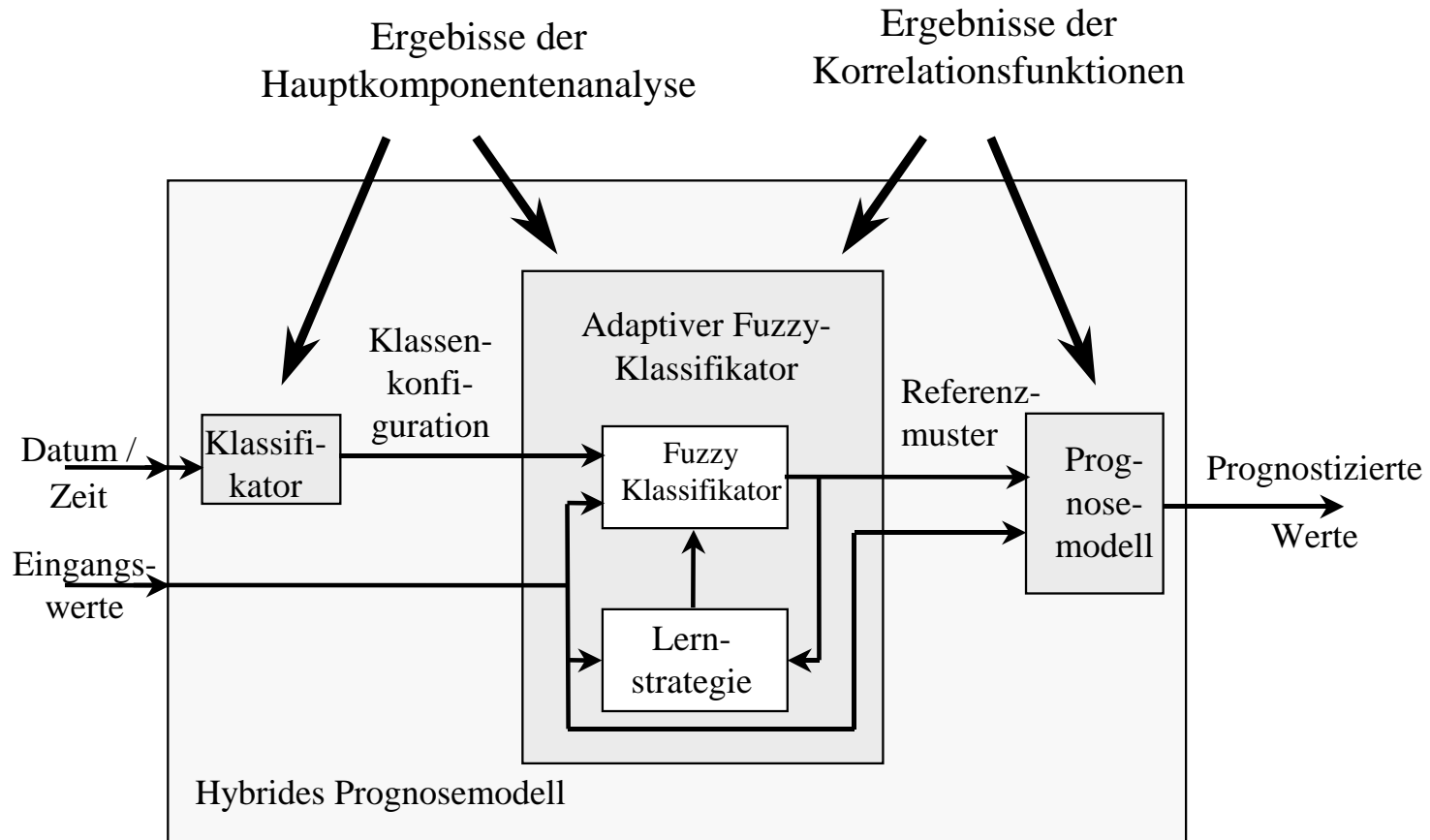
b) Modell mit Linearisierungskomponente

3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.4 Musterbasierte Vorhersage

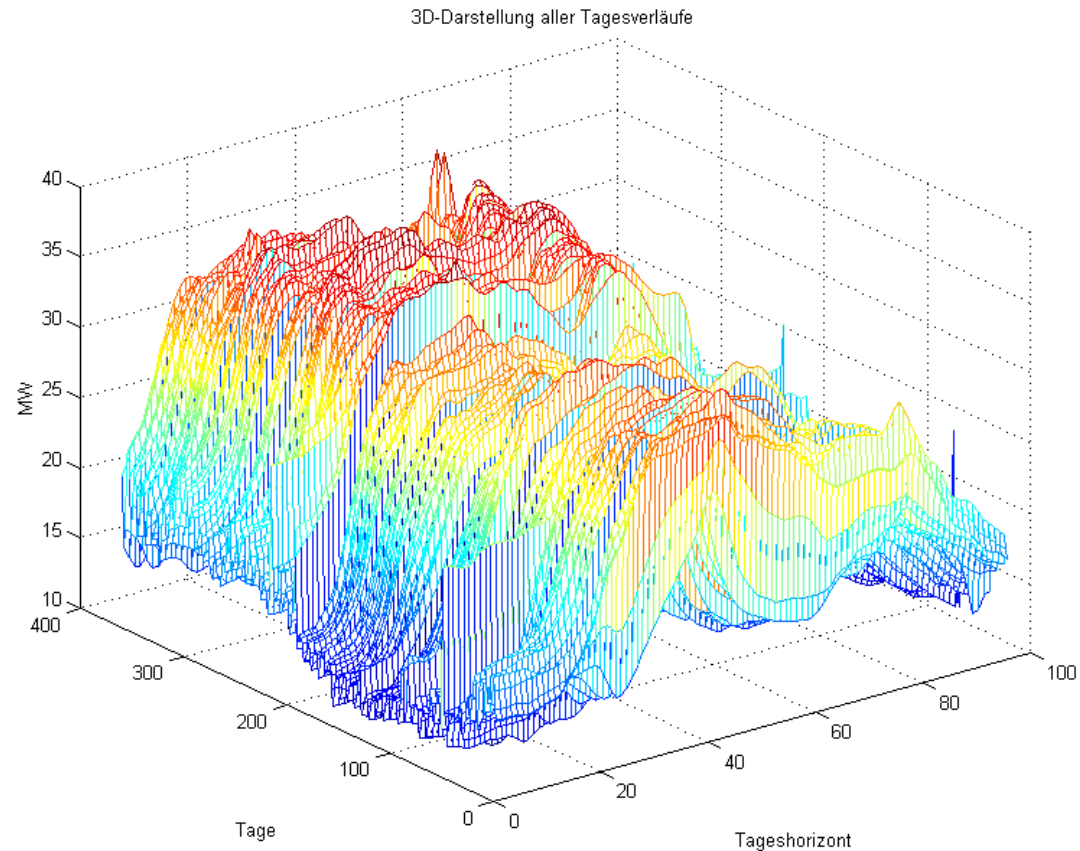
Musterbasierte Prognose für Verbrauchsvorhersagen



3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.4 Musterbasierte Vorhersage Elektrische Last

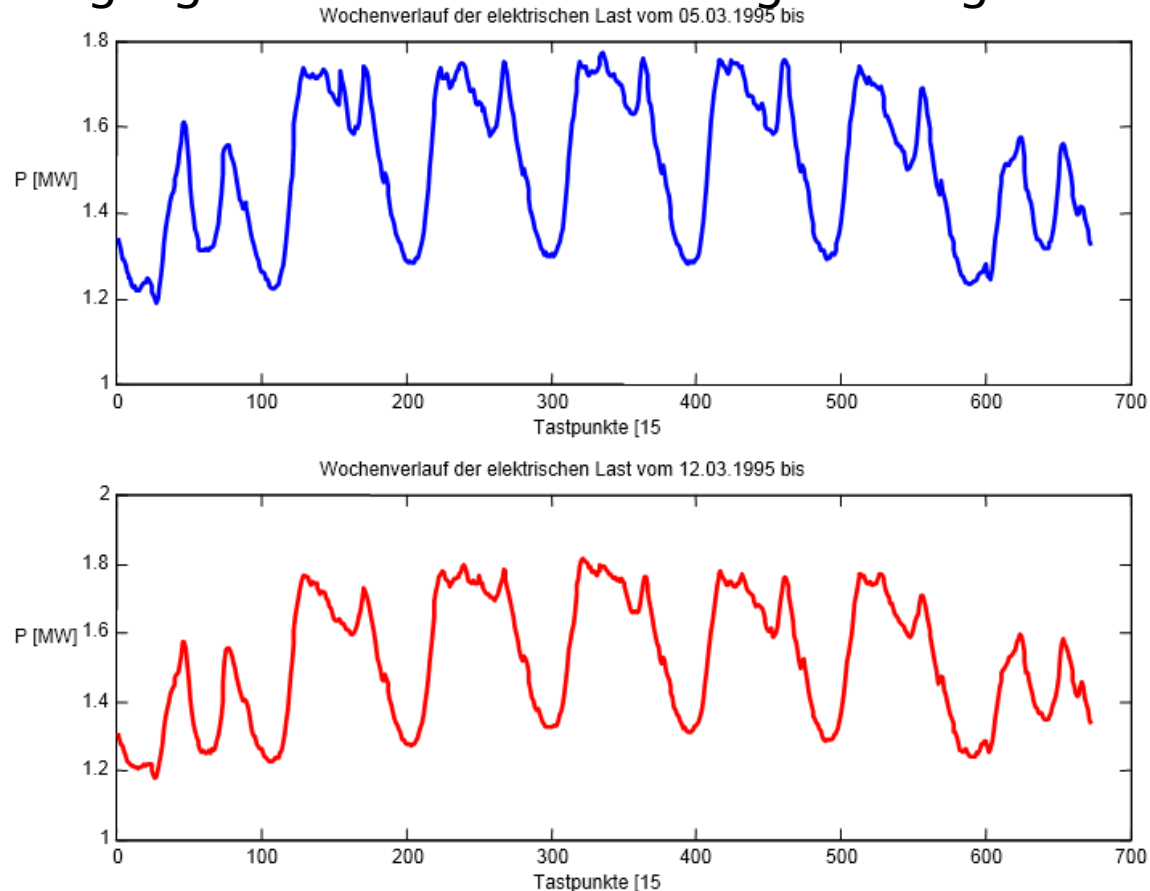


3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.4 Musterbasierte Vorhersage

Elektrischer Lastgang eines kommunalen Energieversorgers

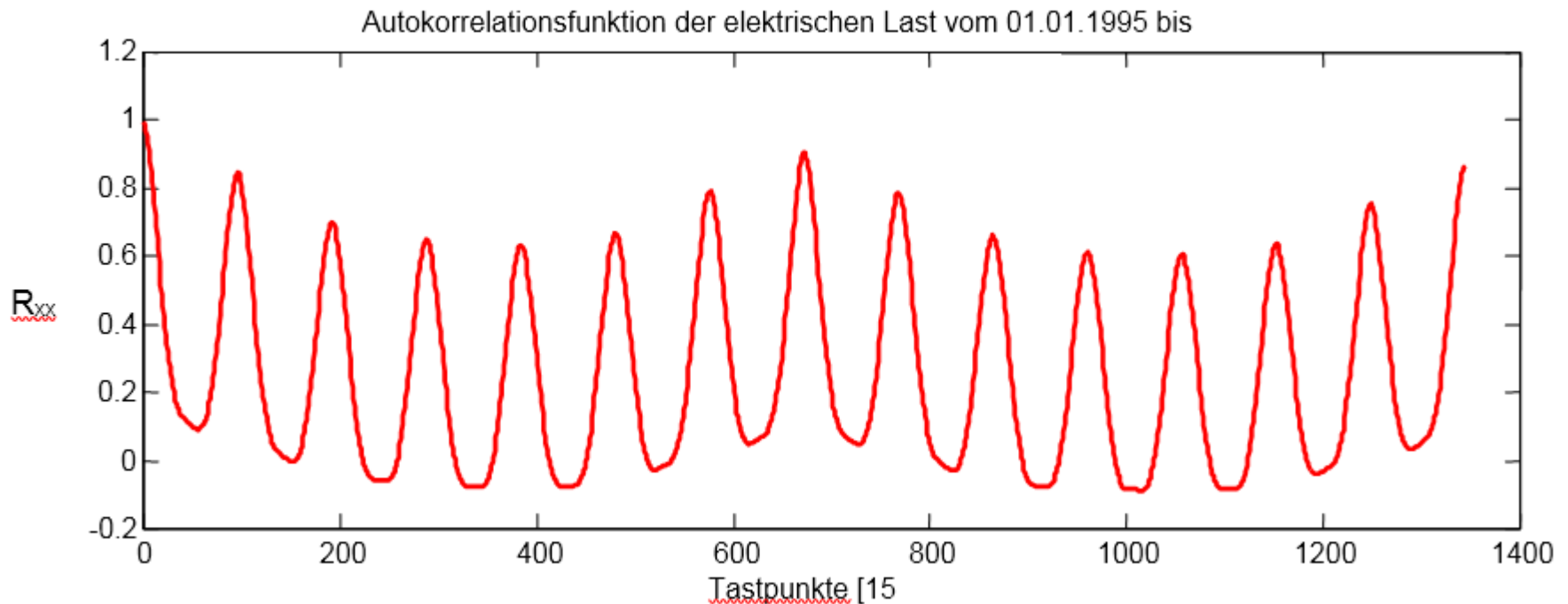


3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.4 Musterbasierte Vorhersage

Autokorrelationsfunktion des elektrischen Lastgangs



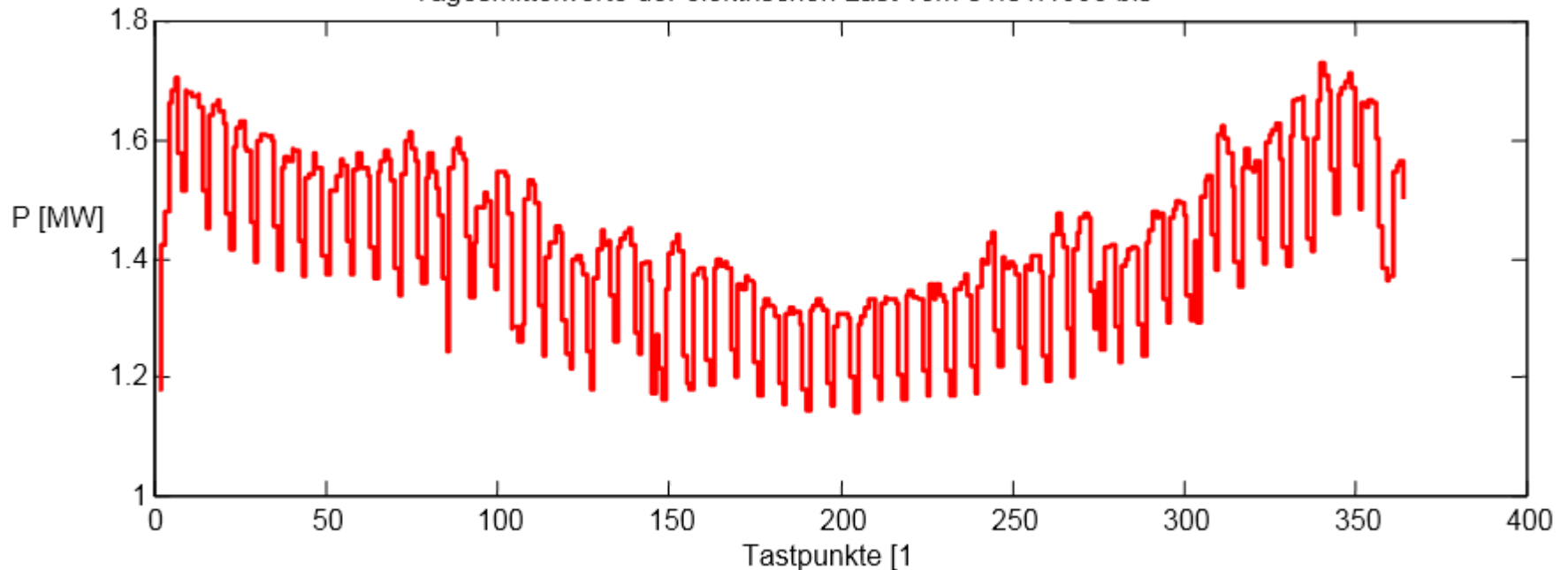
3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.4 Musterbasierte Vorhersage

Jahresverlauf der elektrischen Energie

Tagesmittelwerte der elektrischen Last vom 01.01.1995 bis

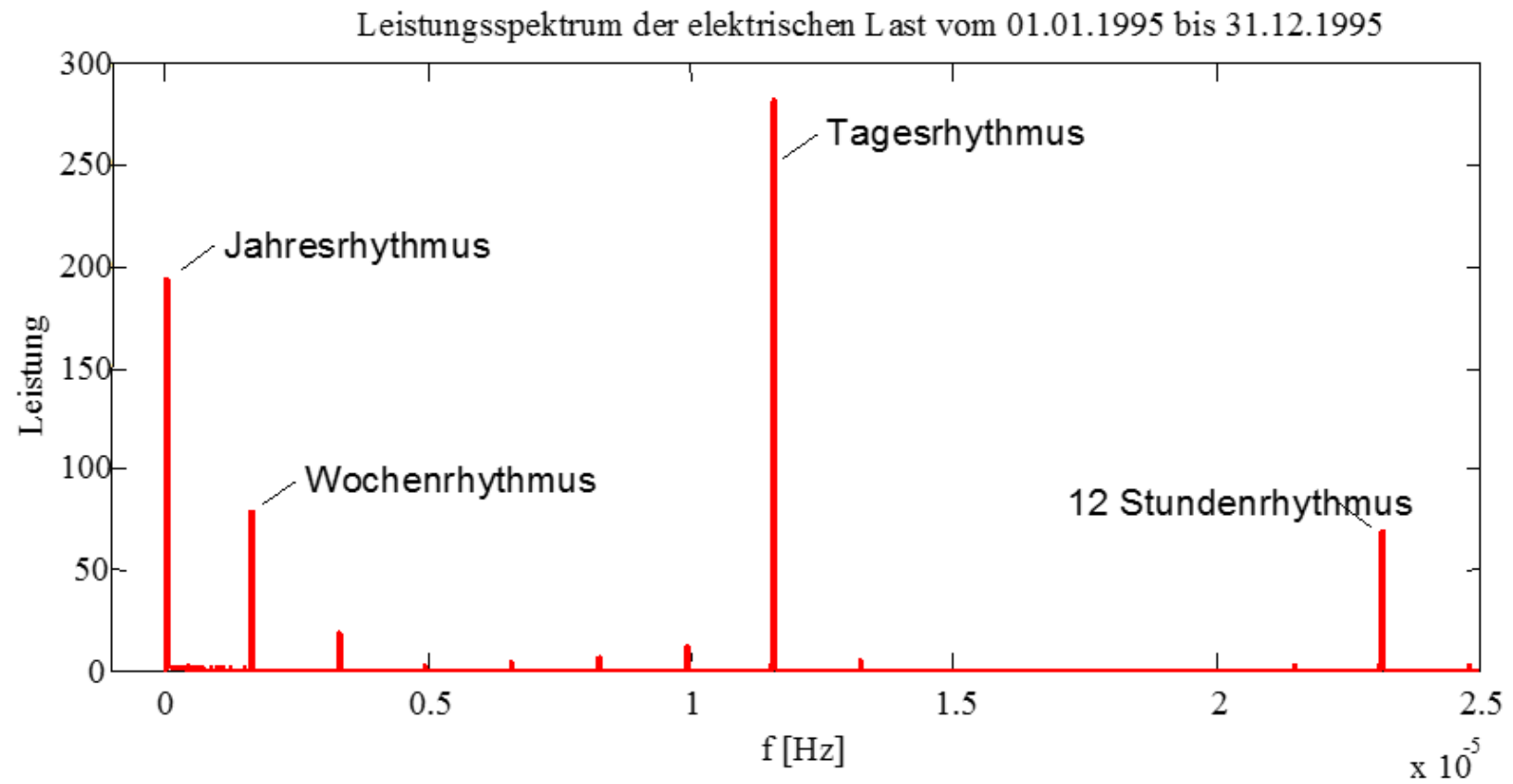


3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.4 Musterbasierte Vorhersage

Leistungsdichtespektrum der elektrischen Last

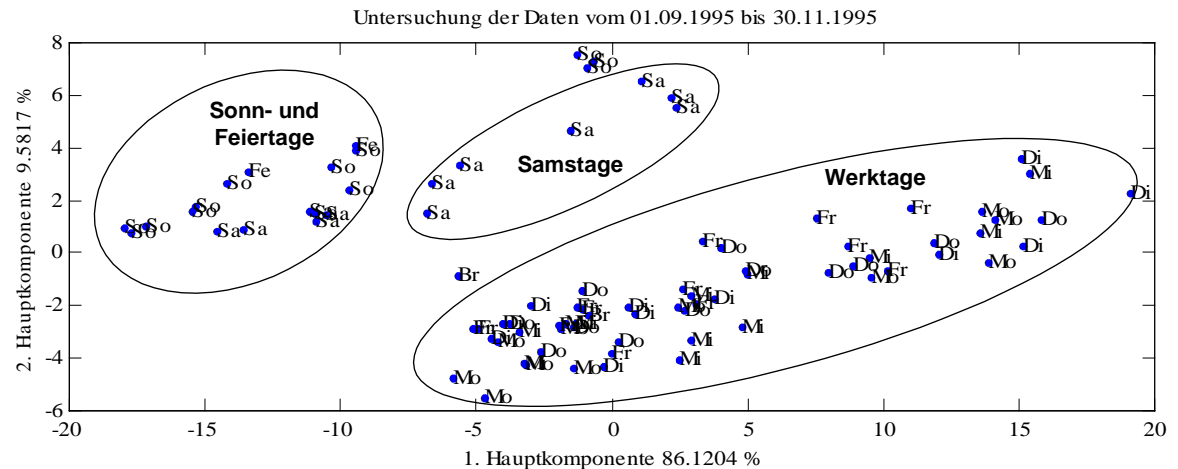
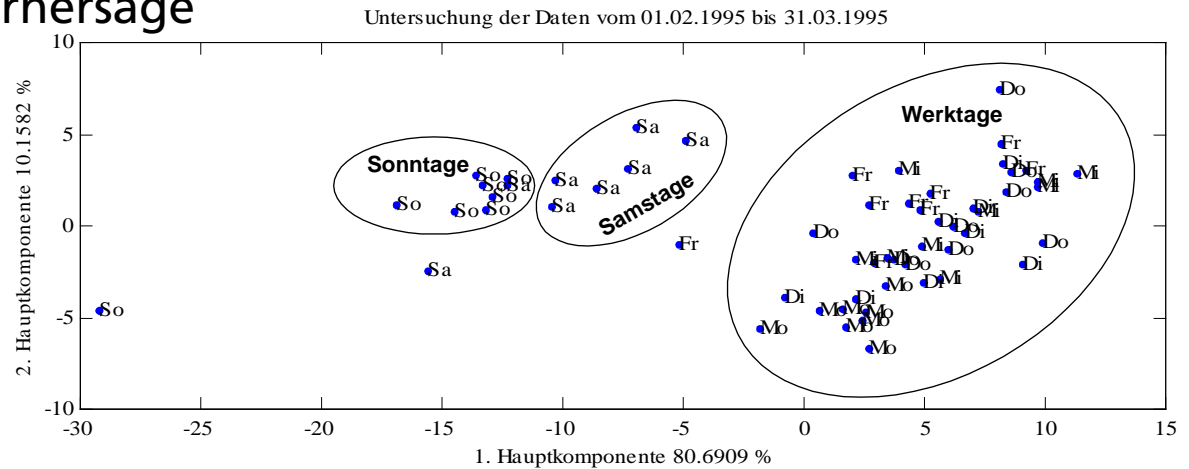


3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.4 Musterbasierte Vorhersage

Hauptkomponentenanalyse der Lastdaten



3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.4 Musterbasierte Vorhersage - Erste Ergebnisse der Lastdaten-Analyse (1.1)

- Der elektrische Lastgang enthält typische nichtlineare Tagesprofile, die als Musterobjekte aufgefasst werden können.
- Der Vorwochentag zeigt in den meisten Fällen die größte Ähnlichkeit zum Tagesprofil des aktuellen Tages.
- Die Untersuchung des elektrischen Lastganges mittels der Hauptkomponentenanalyse ergibt Werktags-, Samstags- und Sonntagsklassen, die aufgrund ihres globalen Charakters zukünftig als Klassenkonfiguration definiert werden. Diese Klassenkonfigurationen sind über das Merkmal „Datum“ erkennbar.
- Feiertage besitzen eine eigene Charakteristik und sollten in der Vorhersage separat berücksichtigt werden.
- Die bisher genutzten Verfahren eignen sich nur für die Generierung einer auf den Zeitraum der Lernphase begrenzten Klassenstruktur.
- Das Verhalten der Tageslastgänge ist durch sich täglich ändernde Tagesprofile charakterisiert und erfordert die Anwendung adaptiver Klassifikationsmethoden.

3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.4 Musterbasierte Vorhersage - Erste Ergebnisse der Lastdaten-Analyse (1.2)

Damit ergeben sich folgende Klassenkonfigurationen:

Werktag – Werktag
Sonntag-Werktag
Werktag-Samstag
Samstag-Sonntag sowie
die Feiertagskombinationen

3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.4 Musterbasierte Vorhersage - Erzeugung stationärer Signaleigenschaften

Als Autokorrelationsfunktion ist hohe Korrelation zum Vorwochenwert zu erkennen ($\zeta = 0,91$).

➡ Linearisierung über Differenzenfilter 1. Ordnung mit Verschiebung $r=672$.

$$\Delta_r P(k) = P(k) - P(k - 672)$$

mit P – elektrische Last.

Problem: Kalenderabhängiges Verhalten, z. B. Feiertage

➡ Neue Lösungsansätze:

- Regelwerkbasierter Differenzenfilter Δ_R $\Delta_R P(k) = P(k) - P(k - R)$

$$\text{mit } R = \begin{cases} 96 \cdot t_i & \text{falls aktueller Wochentag, Feiertag und nicht Sonntag} \\ 672 & \text{sonst} \end{cases}$$

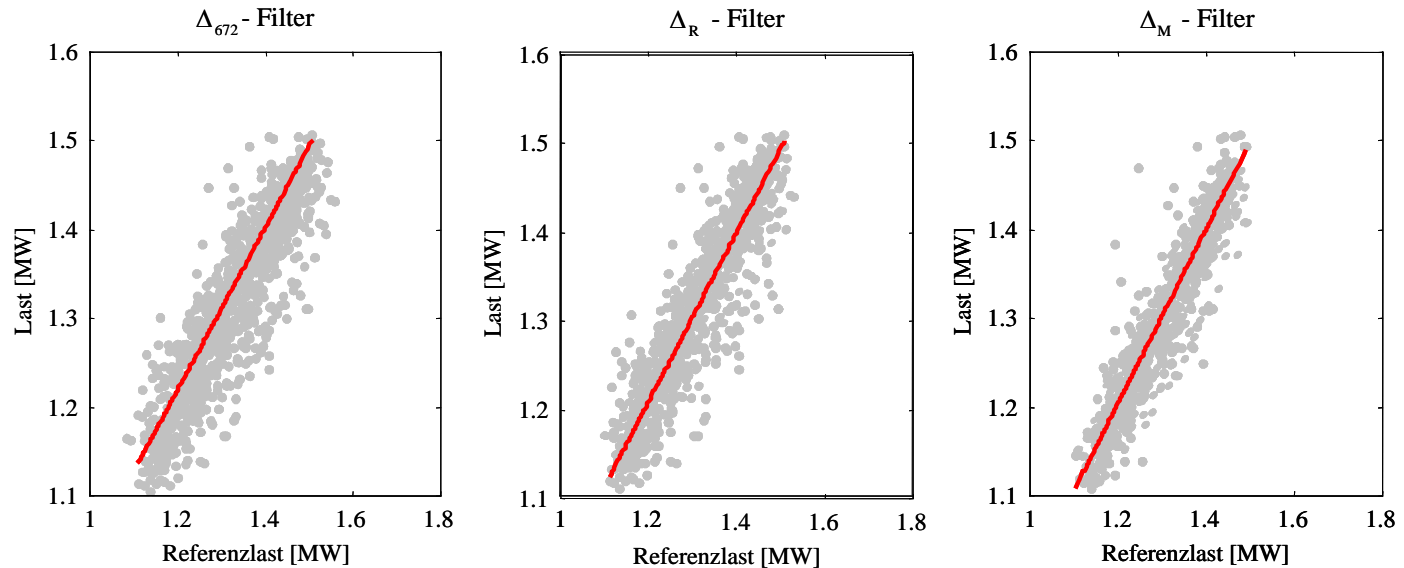
t_i – Tagesindex beginnend mit Sonntag gleich Null

- Musterbasierter Differenzenfilter Δ_M (entspricht Fuzzy-Klassifikator)

3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.4 Musterbasierte Vorhersage - Gegenüberstellung von Last und Referenzlast

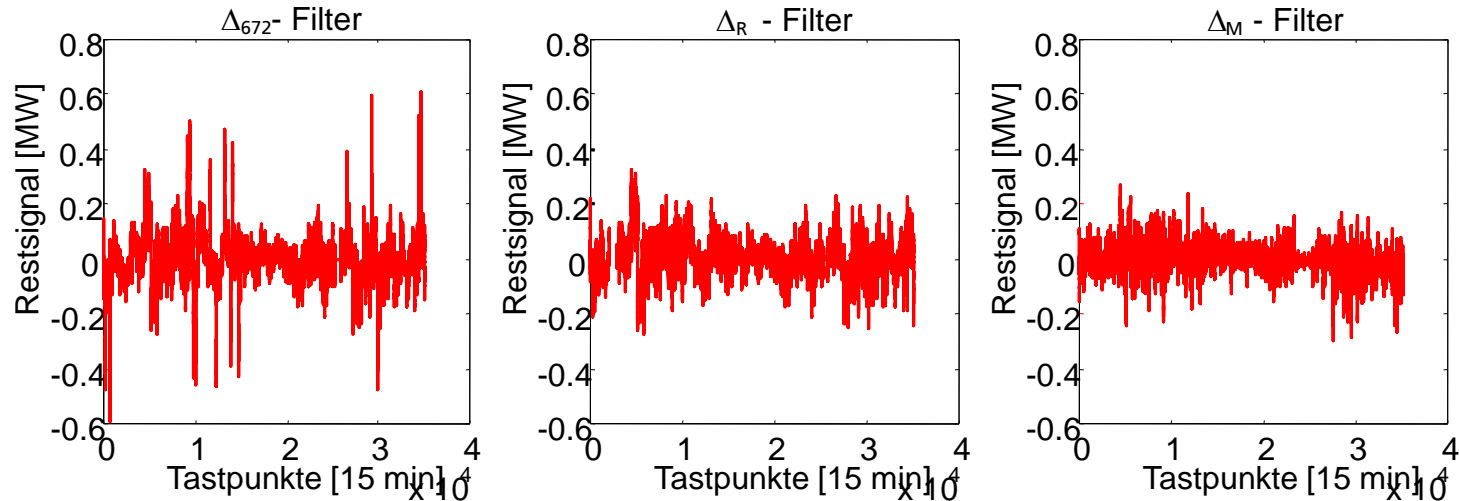


Differenzenfilter	Korrelationskoeffizient
Δ_{672}	0.911
Δ_R	0.949
Δ_M	0.973

3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.4 Musterbasierte Vorhersage - Gefilterte Signale



Differenzenfilter	Mittelwert [MW]	Relativer Mittelwert [%]	Standardabweichung [MW]	rel. Standardabweichung [%]
Δ_{672}	0.0007	0.23	0.0875	6.11
Δ_R	0.0008	0.17	0.0659	4.58
Δ_M	-0.0011	-0.02	0.0480	3.26

3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

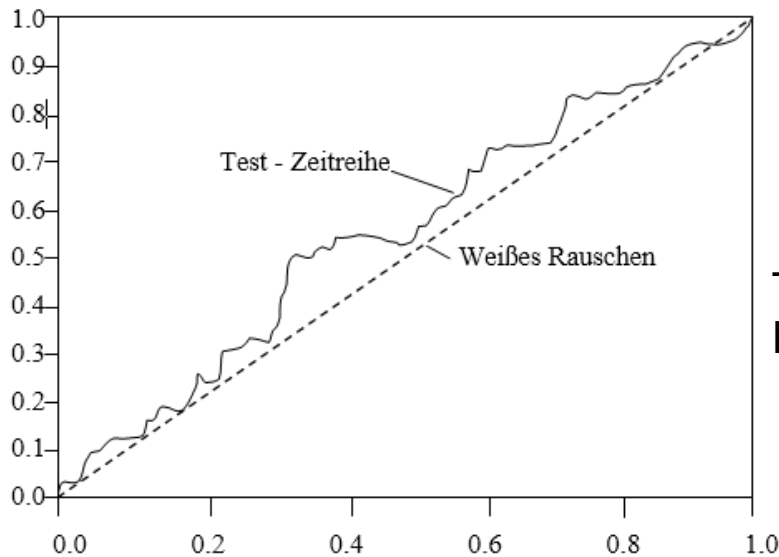
3.6.4 Musterbasierte Vorhersage - Periodogramm

Periodogramm: Aufgabe ist die Erkennung von manifesten oder verborgenen Periodizitäten in Zeitreihen.

Berechnung des Periodogramms $I(f)$:

mit N – Anzahl der Messwerte
 F – Frequenz
 k – Tastpunkt

$$I(f) = N \cdot [C(f)^2 + S(f)^2]$$
$$= N \cdot \left[\frac{1}{N} \sum_{k=1}^N (x(k) - \bar{x}) \cdot \cos 2\pi f \cdot k \right]^2 +$$
$$+ N \cdot \left[\frac{1}{N} \sum_{k=1}^N (x(k) - \bar{x}) \cdot \sin 2\pi f \cdot k \right]^2$$

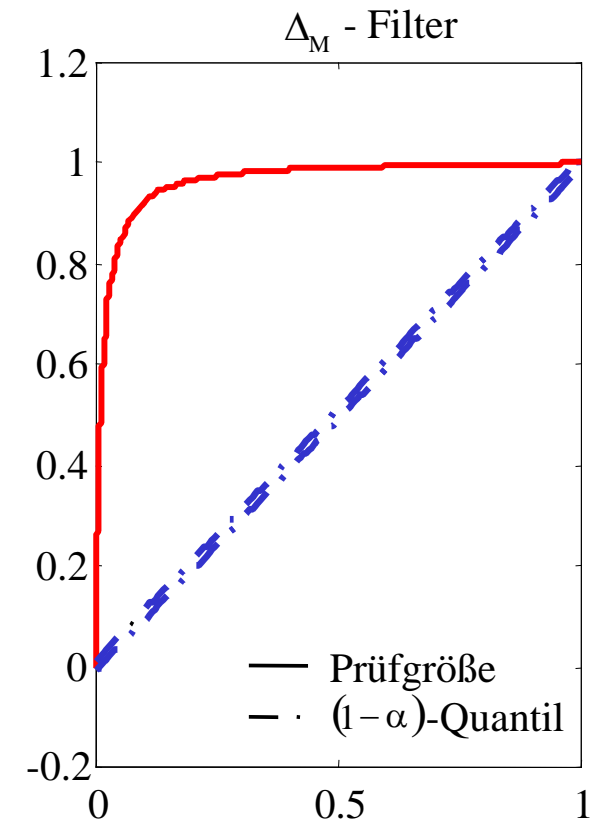
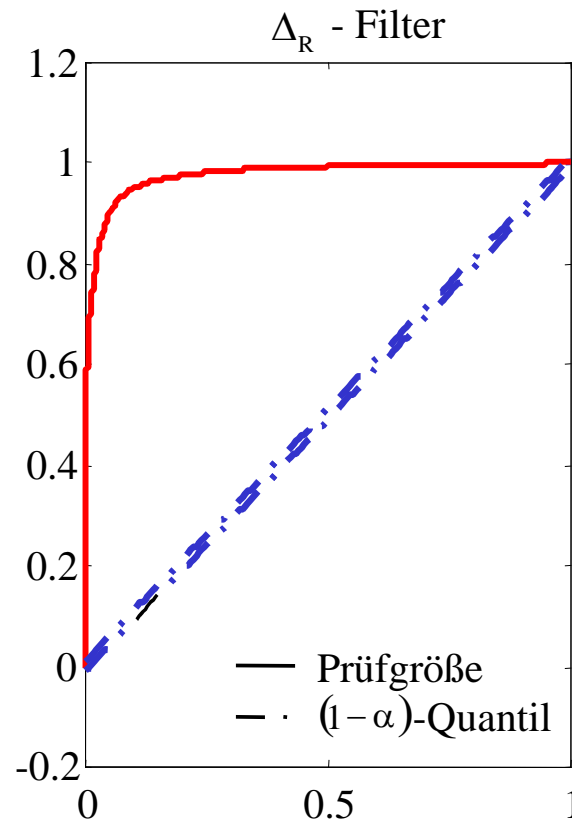
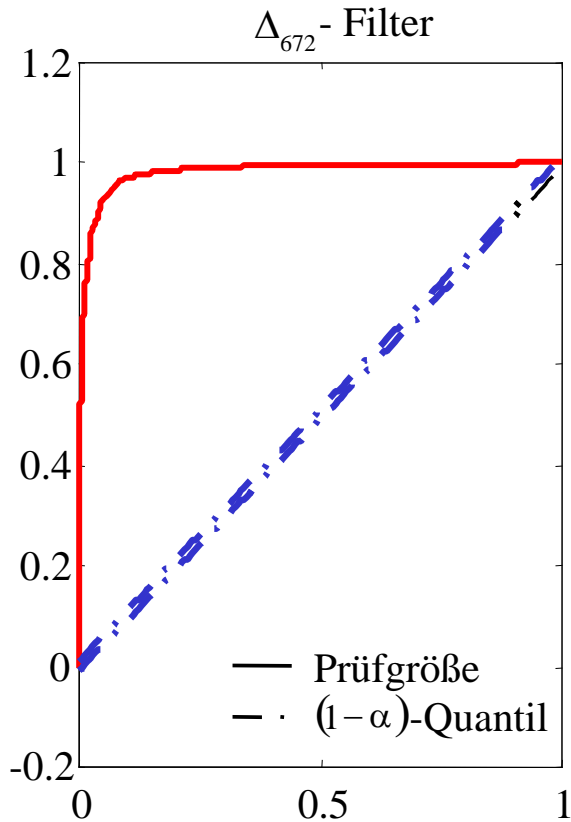


Test auf weißes Rauschen mit kumuliertem Periodogramm

3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

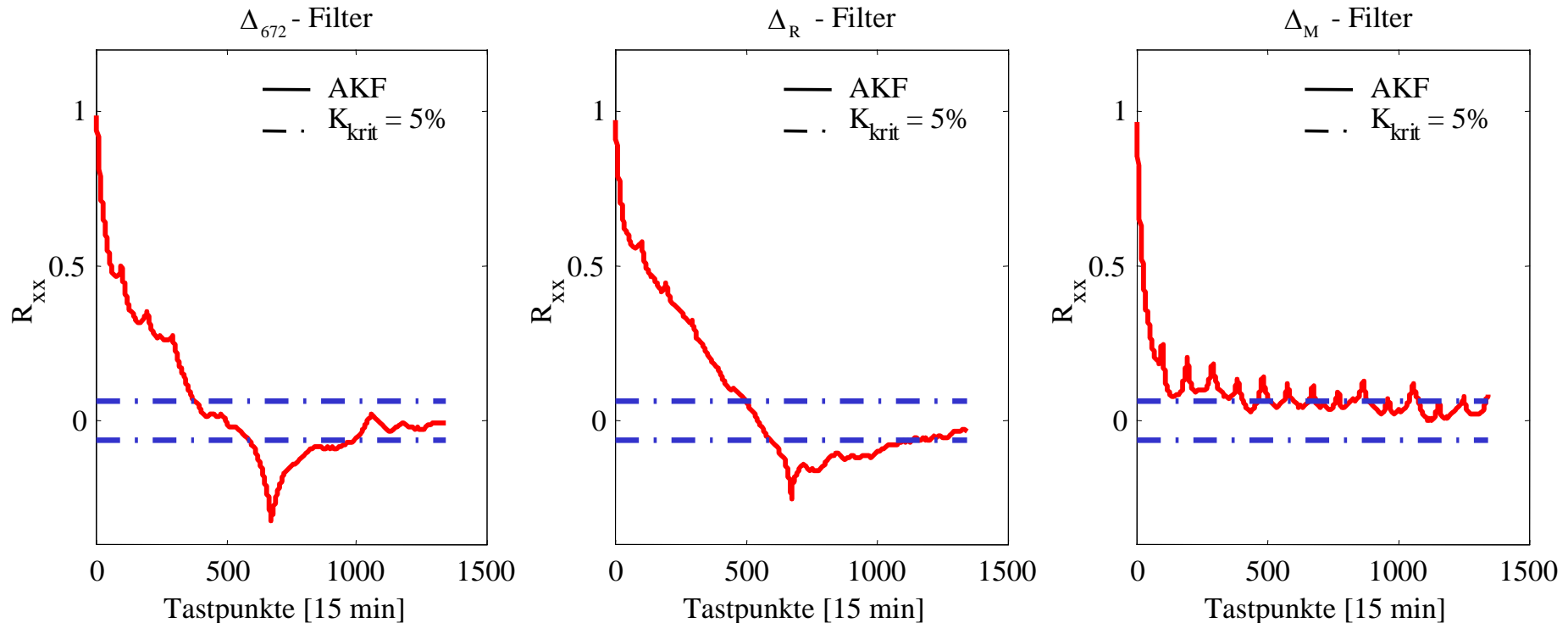
3.6.4 Musterbasierte Vorhersage - Periodogramm Test auf weißes Rauschen



3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.4 Musterbasierte Vorhersage - Test auf weißes Rauschen mit der Autokorrelationsfunktion



3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

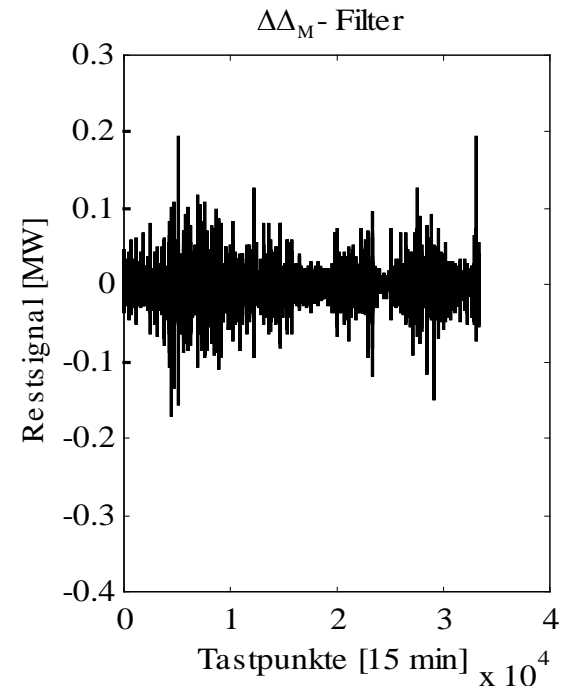
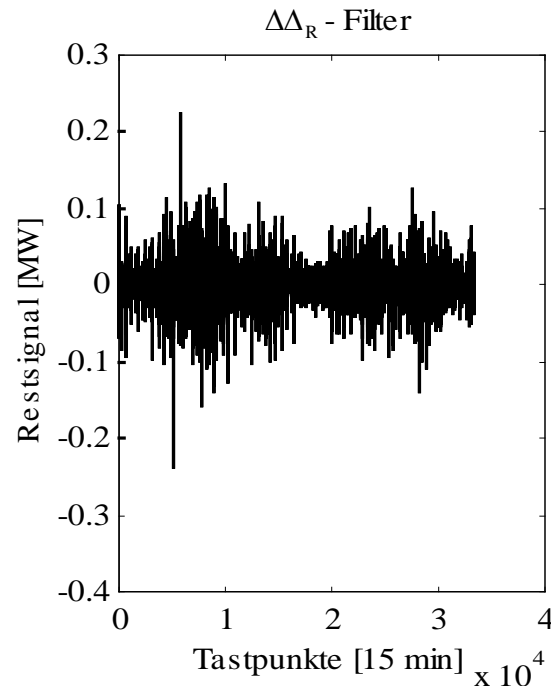
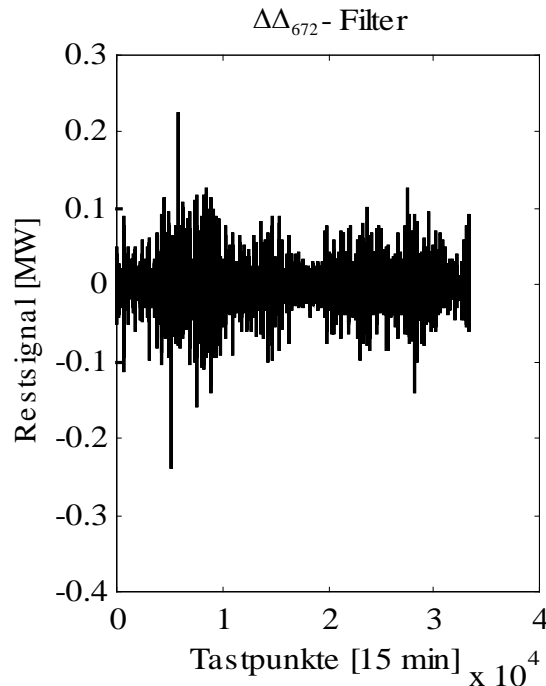
3.6.4 Musterbasierte Vorhersage

Zusätzliche Filterung des Signals

$$\Delta_H \Delta_{672} P(k) = \Delta_{672} P(k) - \Delta_{672} P(k - H)$$

$$\Delta_H \Delta_R P(k) = \Delta_R P(k) - \Delta_R P(k - H)$$

$$\Delta_H \Delta_M P(k) = \Delta_M P(k) - \Delta_M P(k - H)$$

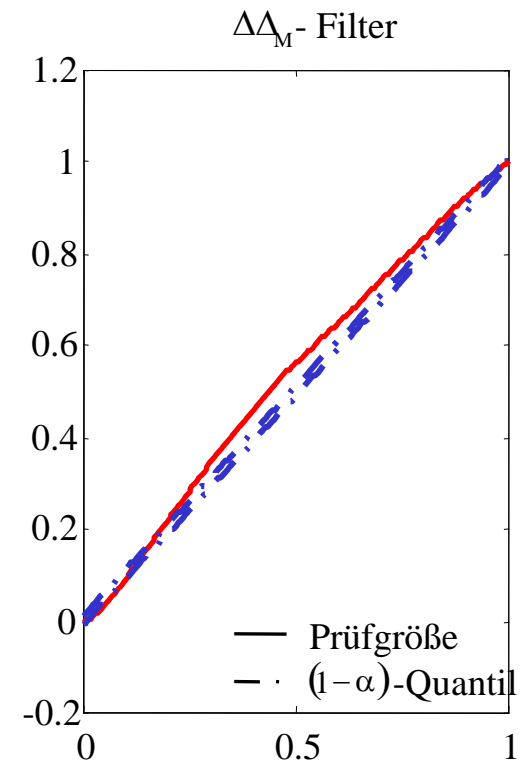
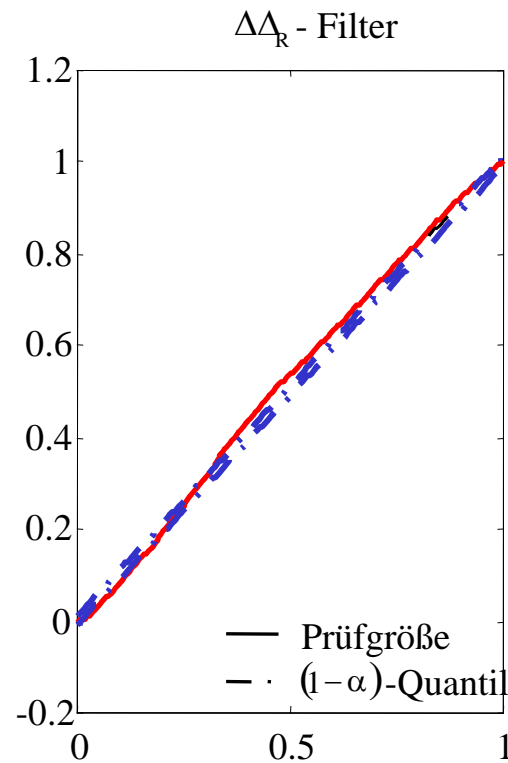
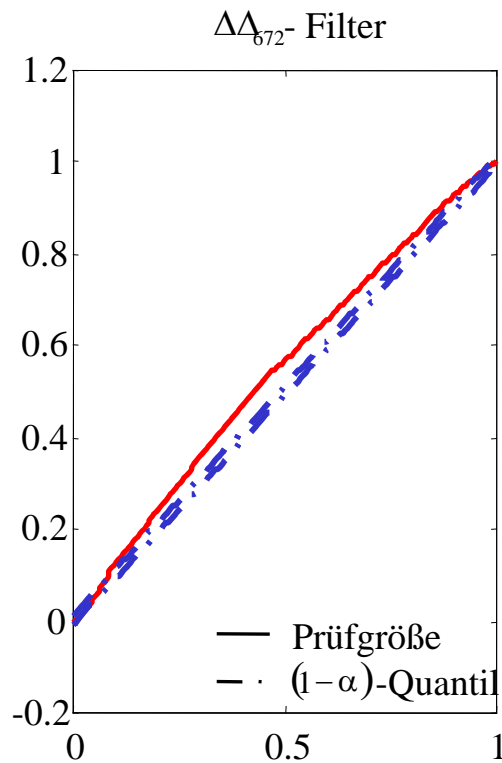


3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.4 Musterbasierte Vorhersage

Periodogramm – Test auf weißes Rauschen für $H = 1$



3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.4 Musterbasierte Vorhersage - Ergebnisse nach den bisherigen Untersuchungen

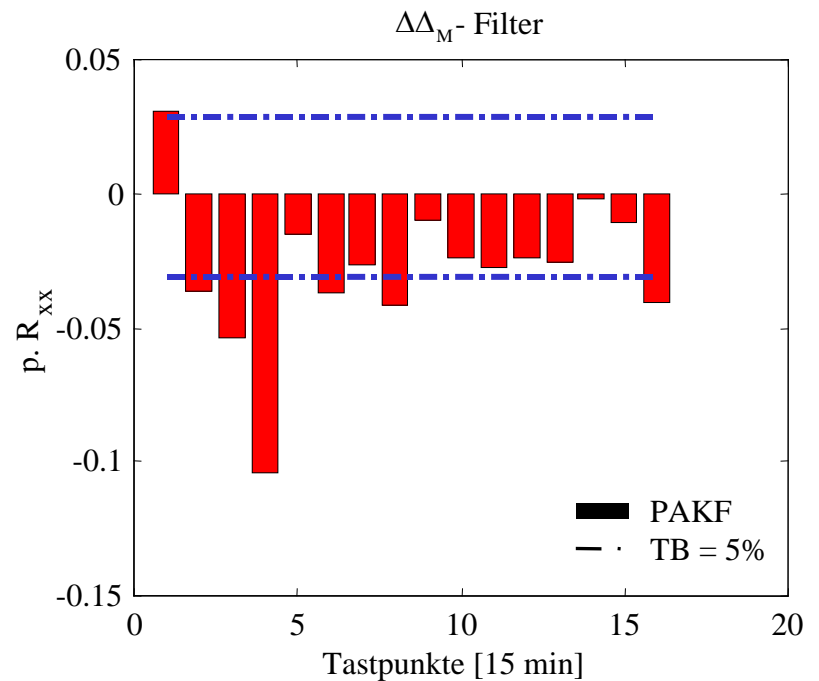
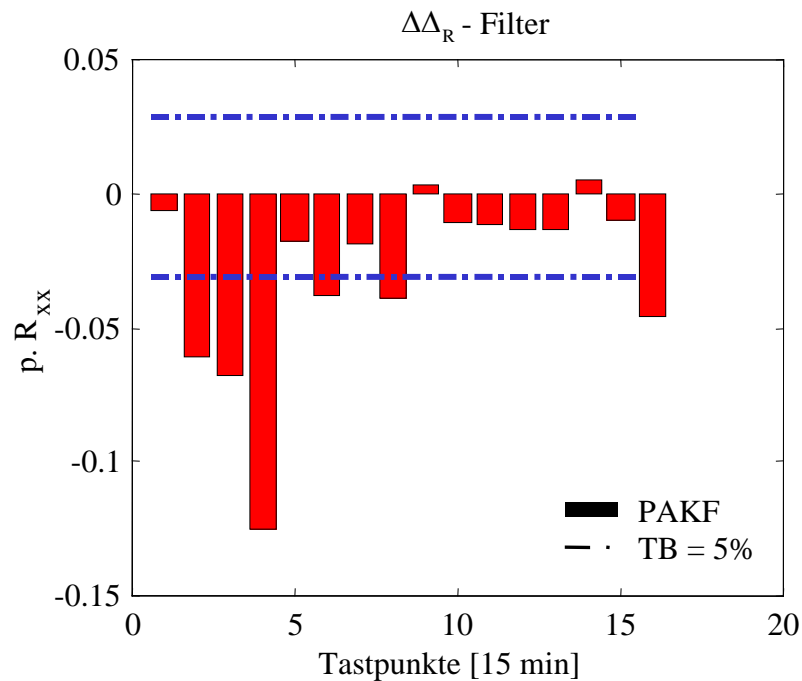
1. Die Differenzenfilter $\Delta_H\Delta_{672}$, $\Delta_H\Delta_R$ und $\Delta_H\Delta_M$ erfassen die wesentlichen deterministischen Signalanteile des elektrischen Lastgangs und bewirken somit eine sehr gute Linearisierung des Signalverlaufs.
2. Die gefilterten Signale $\Delta_H\Delta_{672}$, $\Delta_H\Delta_R$ und $\Delta_H\Delta_M$ weisen instationäres Verhalten auf und erfordern adaptive Modelle.
3. Für den Horizont $H = 1$ entsprechen die gefilterten Signale fast dem weißen Rauschen (siehe Periodogrammtest).
4. Mittelwert und Standardabweichung der gefilterten Signale $\Delta_H\Delta_{672}$, $\Delta_H\Delta_R$ und $\Delta_H\Delta_M$ unterscheiden sich untereinander und sind vom Horizont H abhängig.
5. Der Vergleich von Mittelwert und Standardabweichung offenbart für alle Horizonte H ($H = 1, \dots, 96$) deutliche Vorteile der Filtervarianten $\Delta_H\Delta_M$ und $\Delta_H\Delta_R$, wobei die Variante $\Delta_H\Delta_M$ die besten Ergebnisse erzielt.
6. Eine grobe Vorhersage des elektrischen Lastganges ist mittels der kombinierten Differenzenfilter bereits möglich, wobei aufgrund der ermittelten Ergebnisse der musterbasierte Differenzenfilter $\Delta_H\Delta_M$ zu favorisieren ist.
7. Der Differenzenfilter $\Delta_H\Delta_{672}$ wird aufgrund der bisherigen Ergebnisse nicht weiter betrachtet und findet im Rahmen der Vorhersage keine Anwendung.

3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.4 Musterbasierte Vorhersage

Partielle Autokorrelationsfunktion des Restsignals

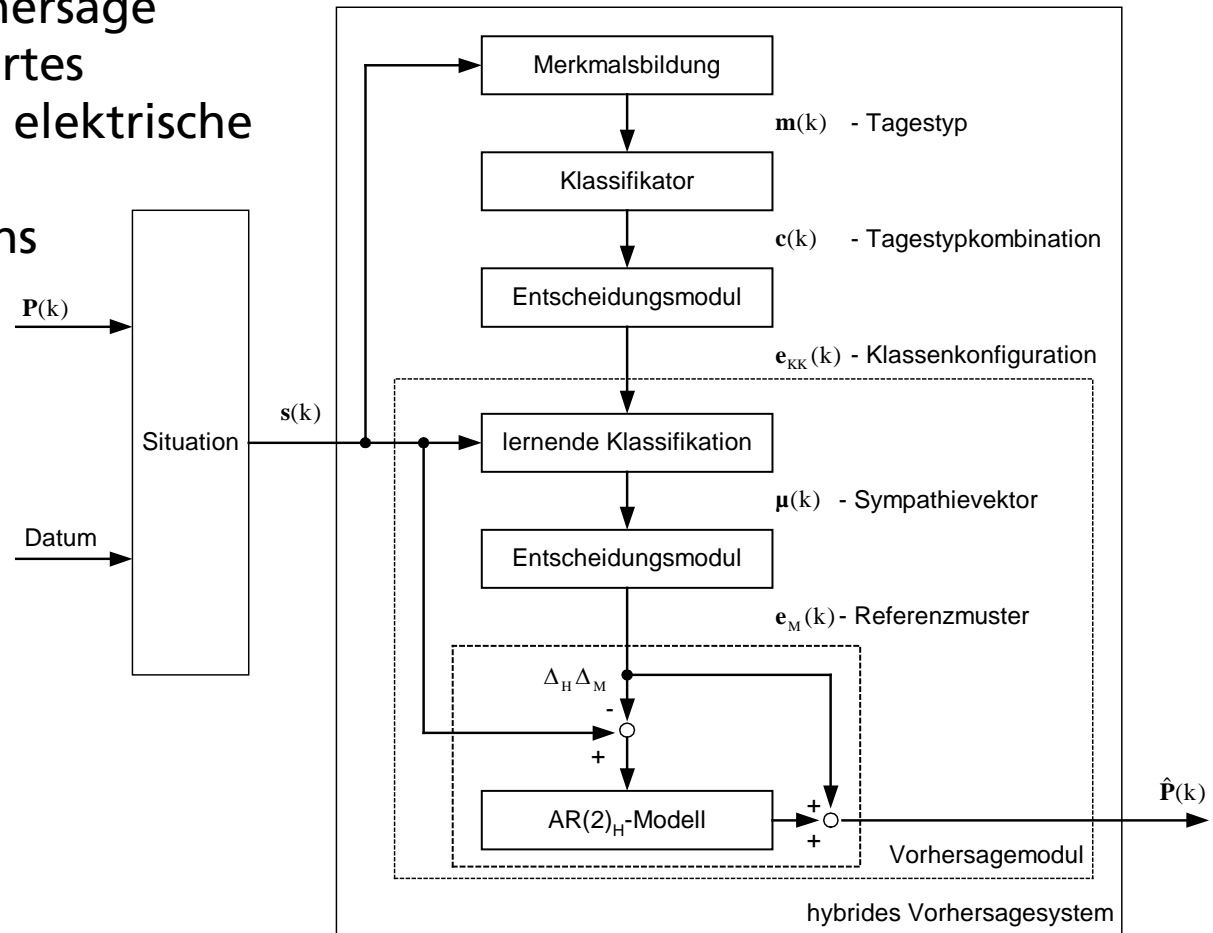


3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.4 Musterbasierte Vorhersage

Umgesetztes musterbasiertes Vorhersagesystem für die elektrische Last eines kommunalen Versorgungsunternehmens

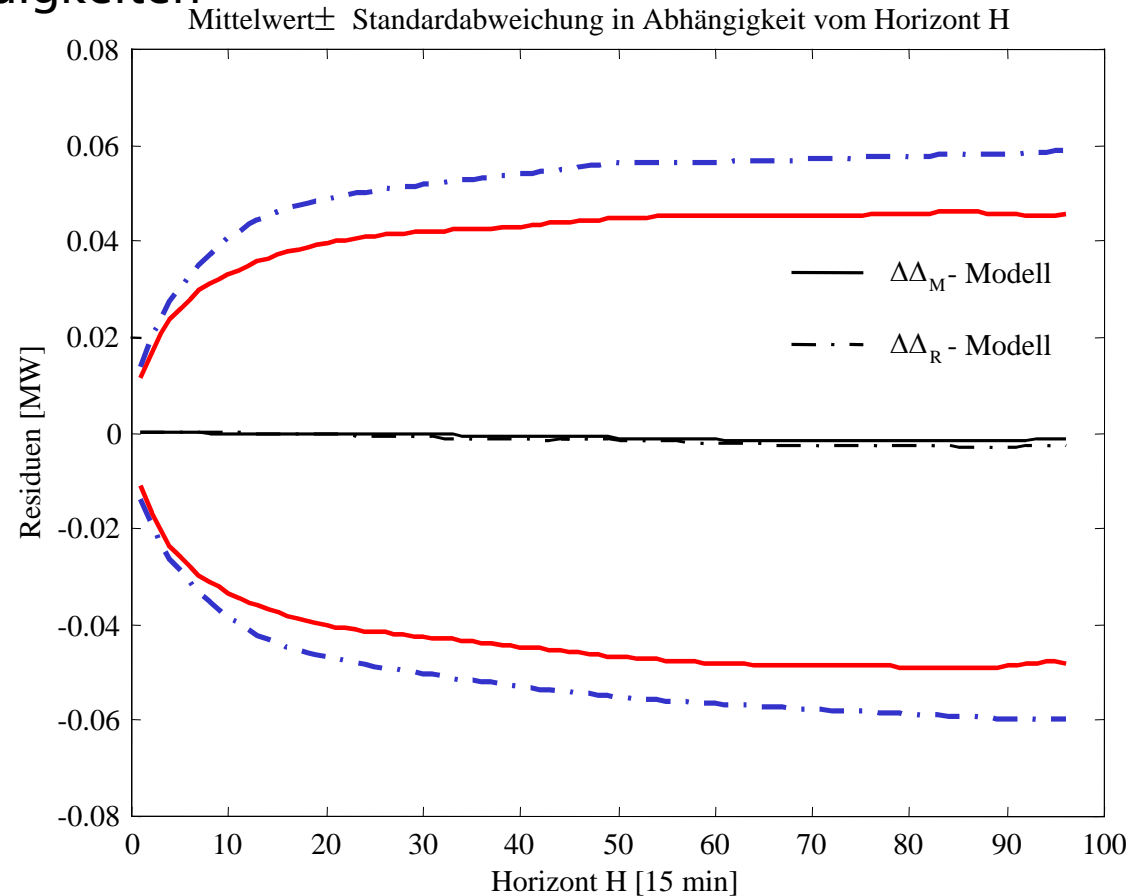


3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.4 Musterbasierte Vorhersage

Erreichte Vorhersagegenauigkeiten

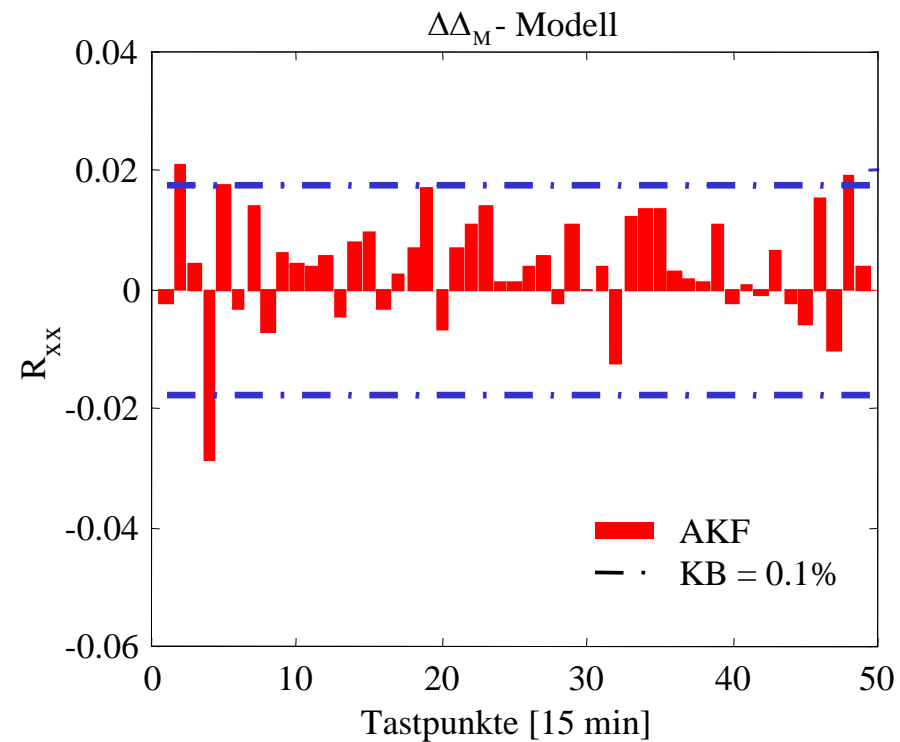
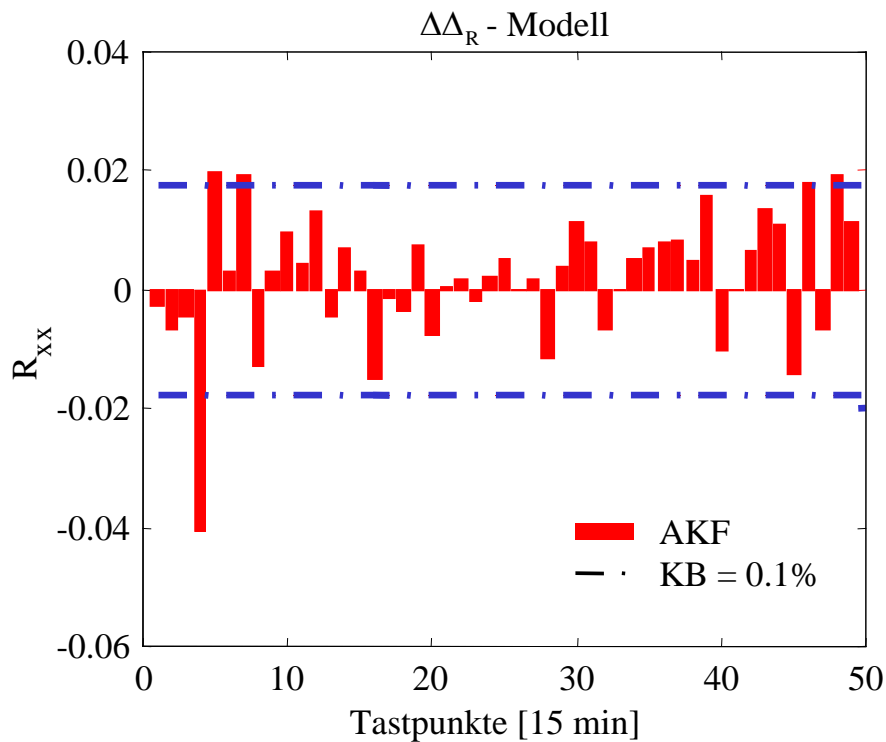


3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.4 Musterbasierte Vorhersage

Autokorrelationsfunktion des Fehlersignals

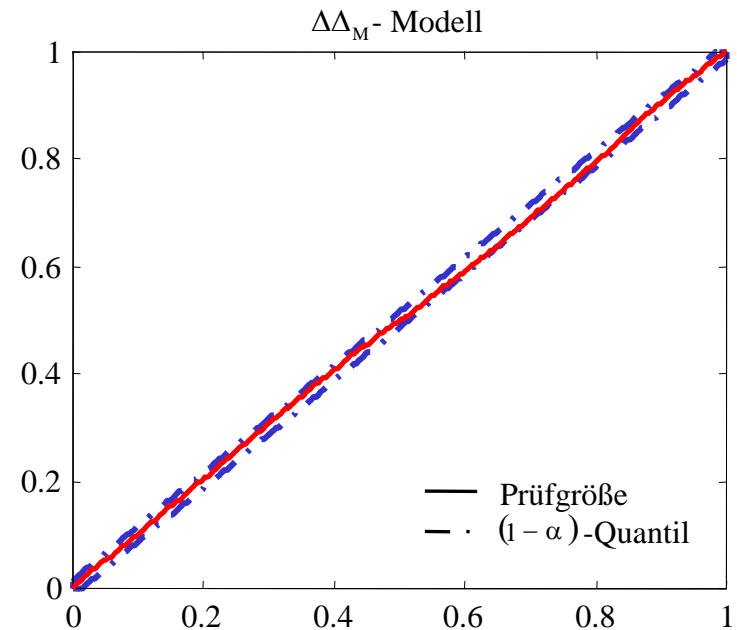
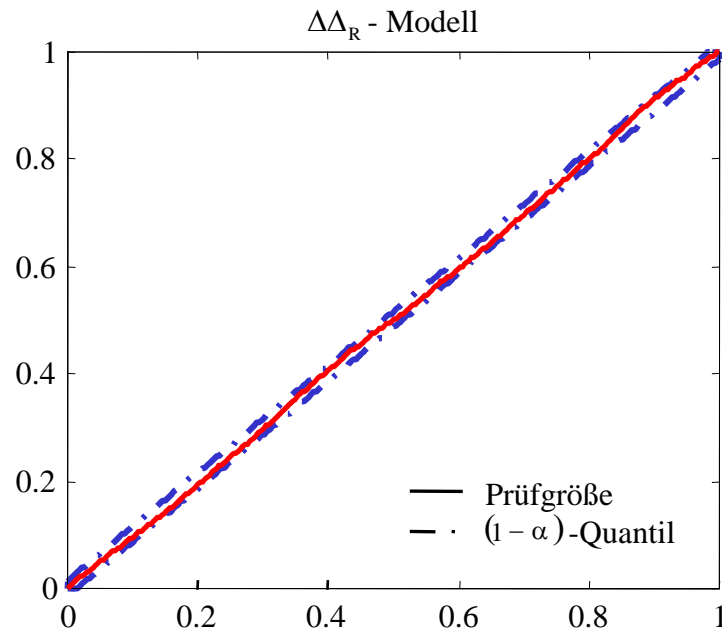


3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.4 Musterbasierte Vorhersage

Periodogramm – Test des Fehlersignals



3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.5 Vorhersage mit neuronalen Netzen

- Erste Ansätze und Ideen zur Verwendung von Neuronalen Netzen gab es ab 1940
- Zwischen 1980 und 1990 war die Entwicklung soweit voran geschritten, dass sie sinnvoll in den verschiedensten Anwendungsgebieten eingesetzt wurden, z. B.:
 - Bilderkennung
 - Spracherkennung
 - Modellbildung
 - Finanzvorhersage
 - Regelung / Steuerung
- Gerade auf dem Gebiet der Finanzvorhersage hat sich gezeigt, dass die Strategie „Gib dem Neuronalen Netz irgendwelche Daten und es wird gute Ergebnisse liefern.“ nicht funktioniert.

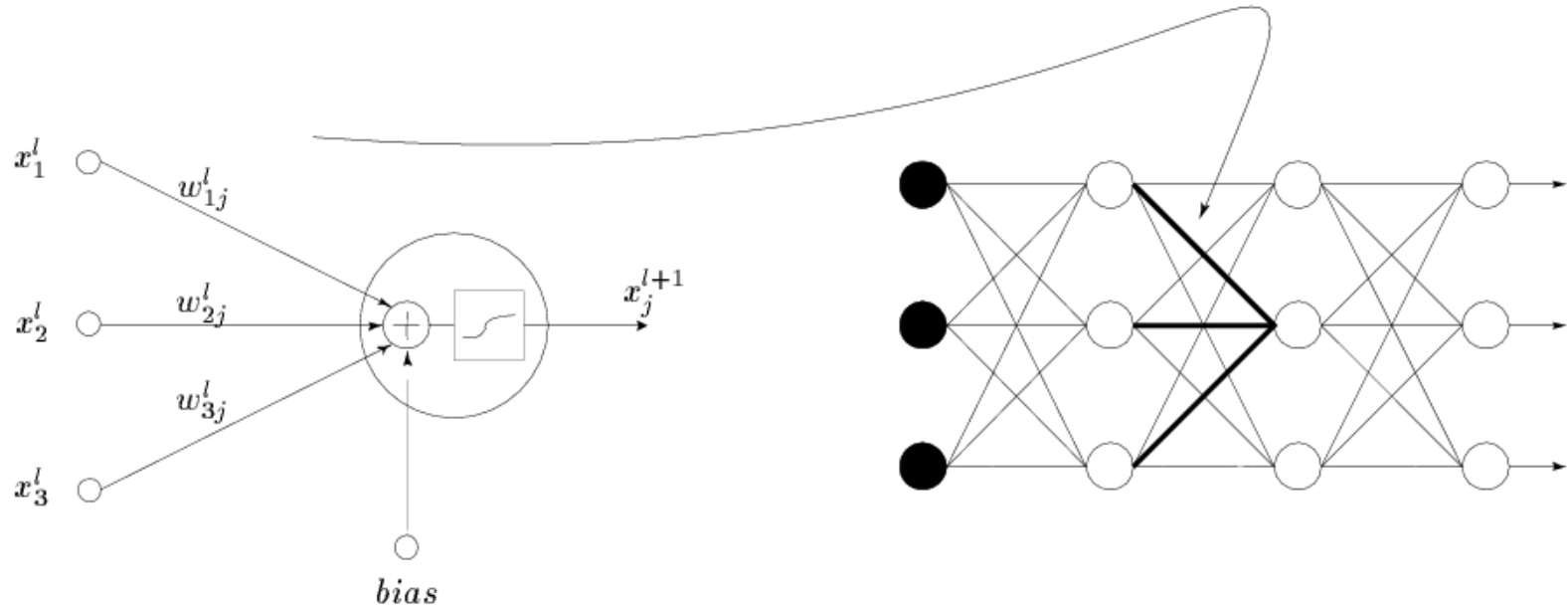
Eigenschaften:

- Bilden Ein- / Ausgangsverhalten allein aus Beobachtungen des Prozesses nach
- Kenntnis der inneren Prozessvorgänge ist nicht erforderlich
- Minimum an a-priori-Wissen ist notwendig
- durch innere Strukturen in der Lage innere Zusammenhänge abzubilden

3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.5 Vorhersage mit neuronalen Netzen - Struktur eines neuronalen Netzes

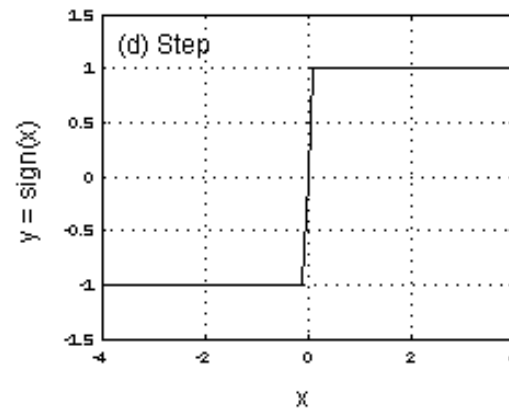
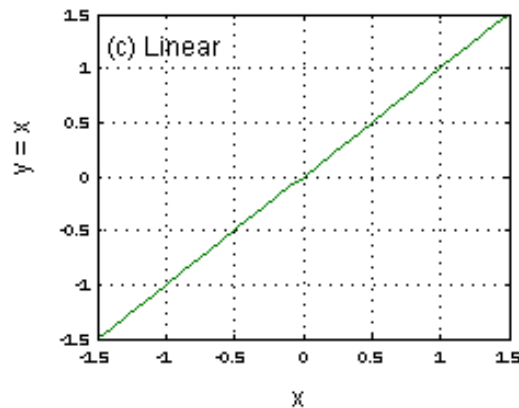
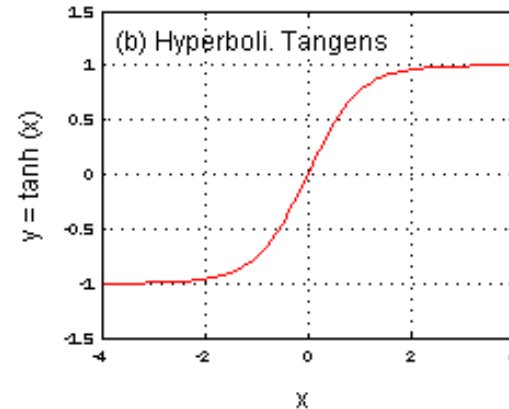
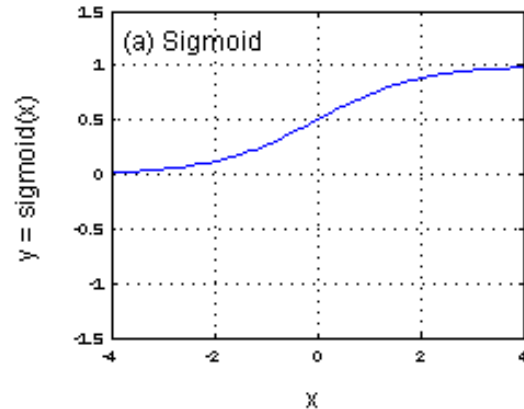


$$x_j^{l+1} = f \left(\sum_i w_{i,j}^l x_i^l \right)$$

3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.5 Vorhersage mit neuronalen Netzen - Aktivierungsfunktionen



3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.5 Vorhersage mit neuronalen Netzen (1)

Lineare stochastische Signalmodelle:

z. B. AR(m)-Prozess

$$y(k) + a_1 y(k-1) + \dots + a_m y(k-m) = \varepsilon(k)$$

Vorteil: Beschreibung der Dynamik eines Prozesses möglich

Nachteil: nicht geeignet für nichtlineare Prozesse

Neuronale Netze:

Vorteil: Abbildung nichtlinearer statischer Zusammenhänge

Nachteil: Dynamik eines Prozesses wird nicht berücksichtigt

Lösung:

- Dynamisierung des neuronalen Netzes durch Einbeziehung zurückliegender Werte als separater Eingang
- Verwendung von FIR-Neuronalen Netzen


3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.5 Vorhersage mit neuronalen Netzen (2)

Linearer AR(m)-Prozess:
$$y(k) = \sum_{i=1}^m a_i y(k-i) + \varepsilon(k)$$

Optimale 1-Schritt-Prognose:
$$\hat{y}(k) = \sum_{i=1}^m a_i y(k-i)$$


$$y(k) = \hat{y}(k) + \varepsilon(k)$$


Nichtlinearer AR(m)-Prozess:

$$y(k) = f(y(k-1), y(k-1), \dots, y(k-m)) + \varepsilon(k)$$

f – nichtlineare Funktion

Optimale 1-Schritt-Prognose:
$$\hat{y}(u) = f(y(k-1), y(k-2), \dots, y(k-m))$$

Mit Hilfe eines neuronalen Netzes wird die ideale Funktion f approximiert.


$$y(k) = \hat{y}(k) + \varepsilon(k)$$

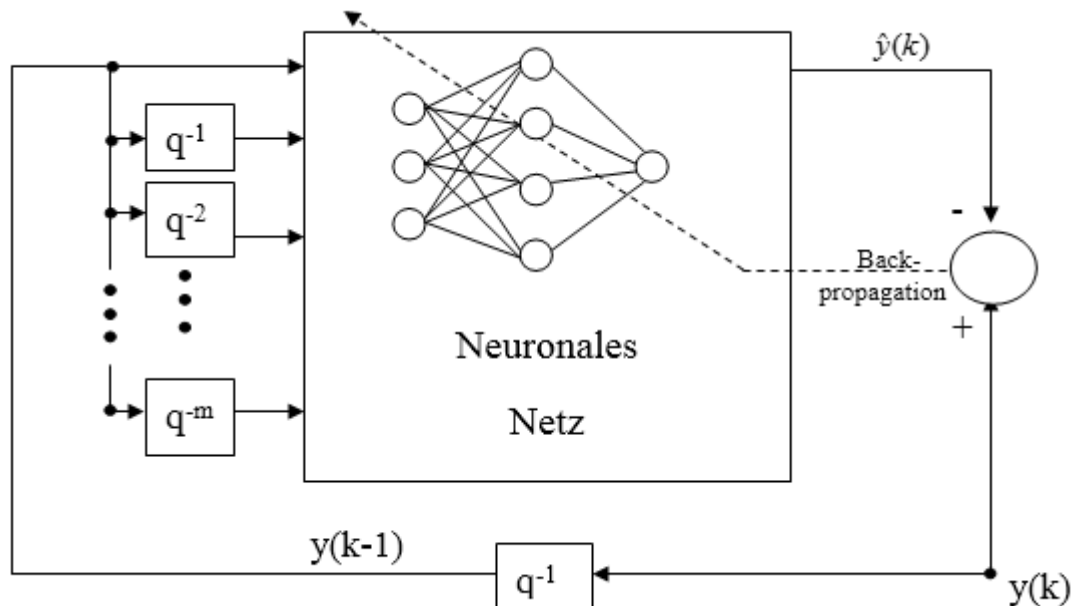
- Ein Feedforward Neuronales Netz ist in der Lage, jede gleichmäßige stetige Funktion zu approximieren.
- Daraus ergibt sich die Motivation, die neuronalen Netze für die Vorhersage einzusetzen.

3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.5 Vorhersage mit neuronalen Netzen - Neuronales Netz und Dynamik (1)

Training:

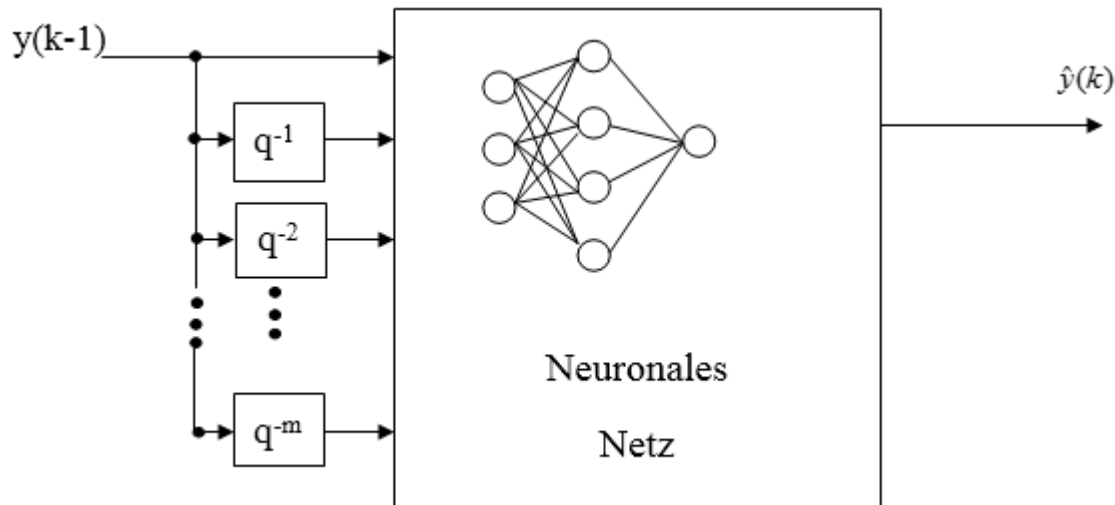


3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.5 Vorhersage mit neuronalen Netzen - Neuronales Netz und Dynamik (2)

Vorhersage:

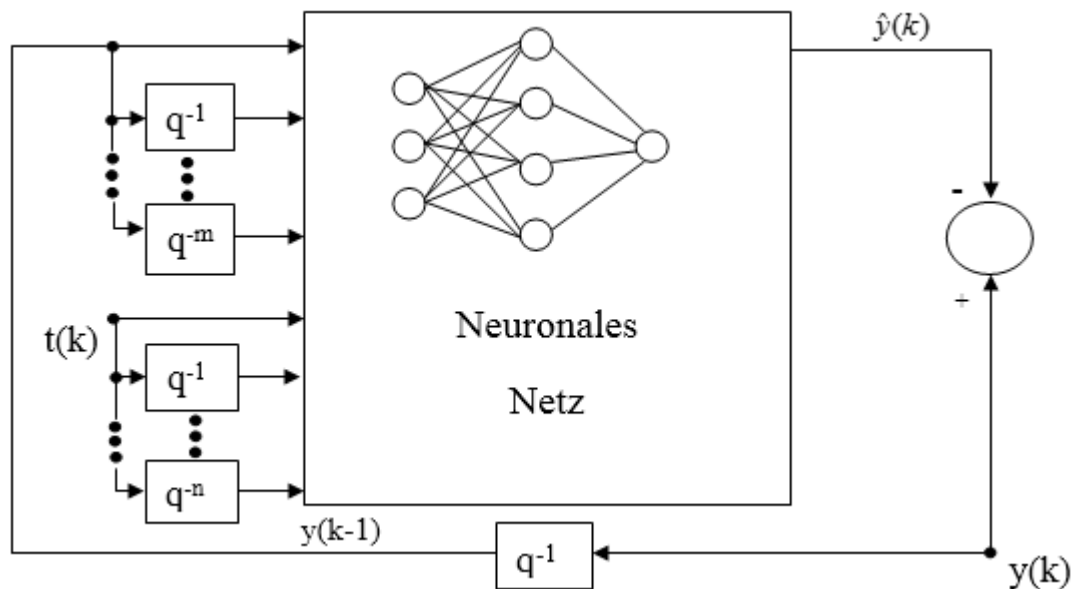


3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.5 Vorhersage mit neuronalen Netzen - Neuronales Netz mit exogenen Eingangsgrößen (1)

Training:

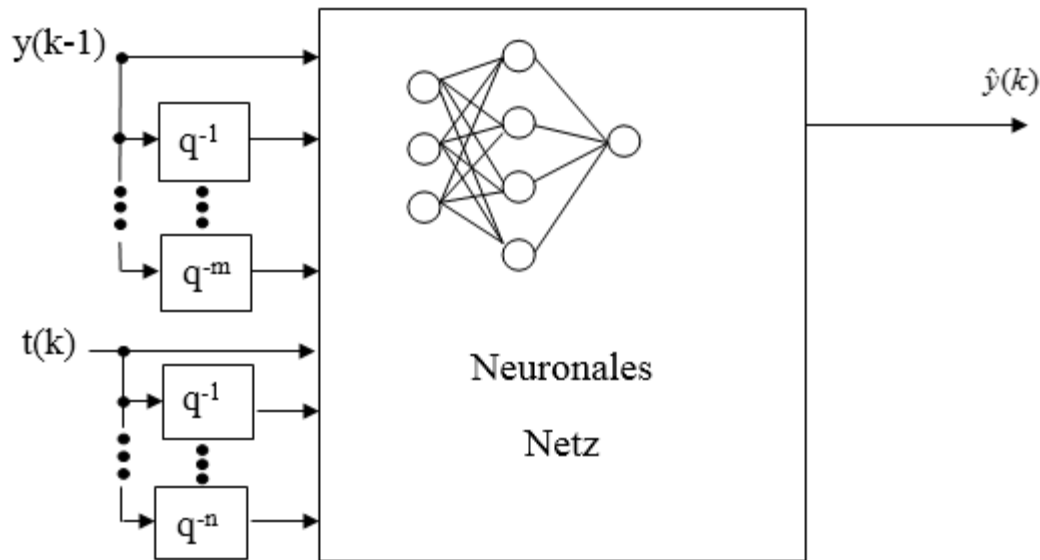


3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.5 Vorhersage mit neuronalen Netzen - Neuronales Netz mit exogenen Eingangsgrößen (2)

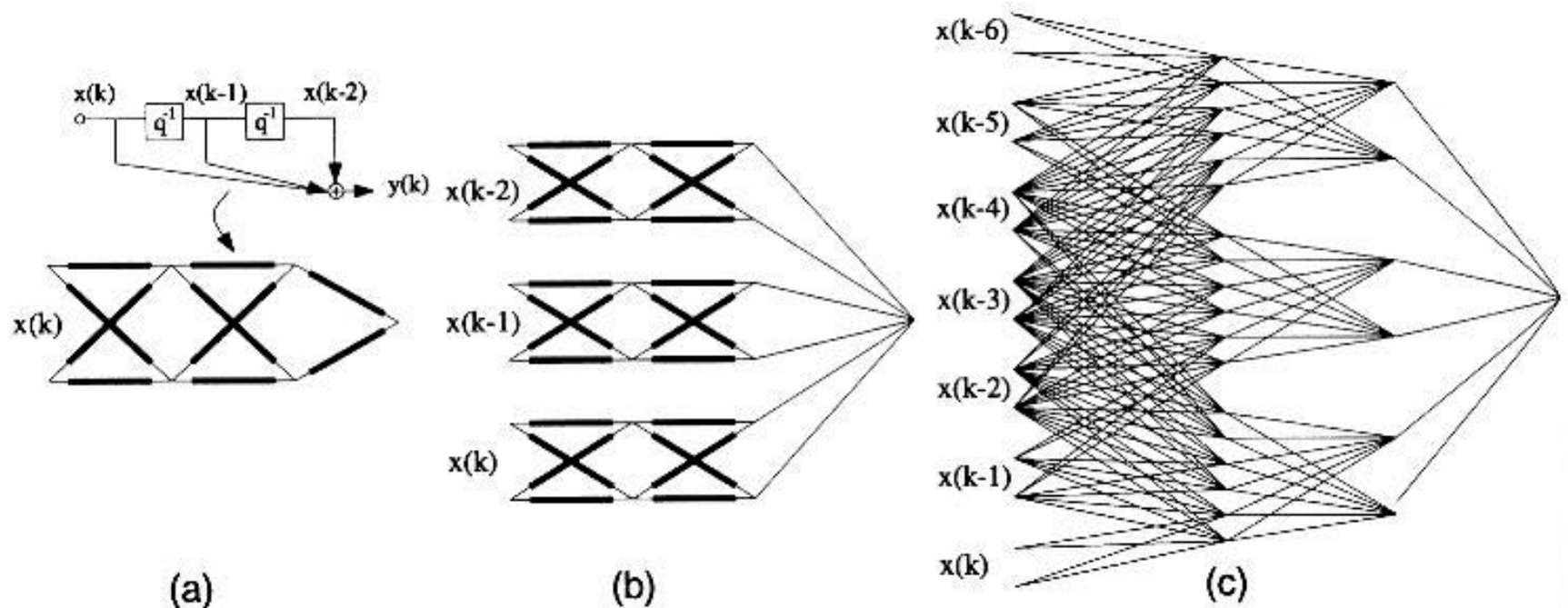
Vorhersage:



3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

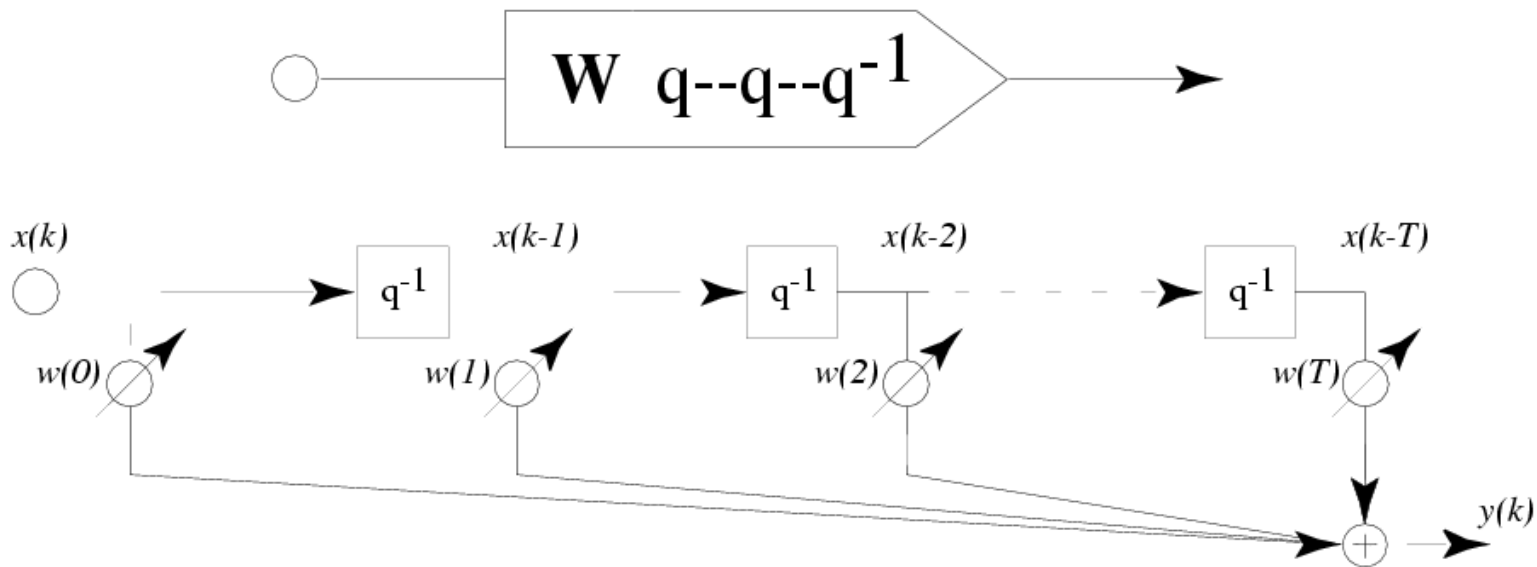
3.6.5 Vorhersage mit neuronalen Netzen - Übergang von einem FIR-Netzwerk zu einem statischen neuronalen Netz



3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

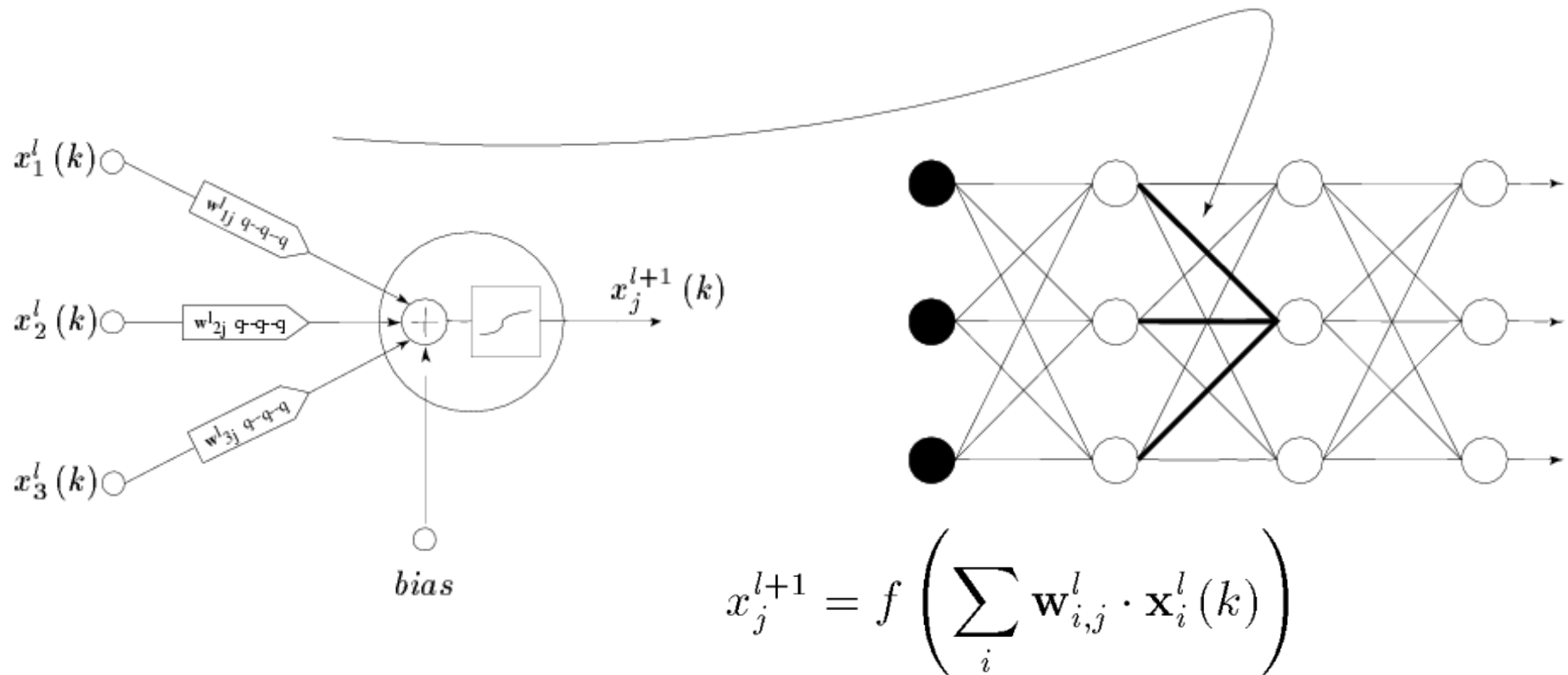
3.6.5 Vorhersage mit neuronalen Netzen - FIR-Filter als Synapse



3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.5 Vorhersage mit neuronalen Netzen - FIR-Neuron und Netzwerk



3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.5 Vorhersage mit neuronalen Netzen - Vergleich FIR-Netz mit statischem Netz

FIR - Netz			Statisches Netz
Neuronen	Filter-Ordnung	Variable Parameter	Variable Parameter
2 x 2 x 2 x 2	2:2:2	30	150
5 x 5 x 5 x 5	10:10:10	605	36355
3 x 3 x 3	9:9	180	990
3 x 3 x 3 x 3	9:9:9	270	9990
3^n	9^{n-1}	$(n-1)90$	$10^n - 10$

3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.5 Vorhersage mit neuronalen Netzen – Wahl der Struktur des Neuronalen Netzes für Vorhersagen

Anzahl der Hidden-Schichten:

Umfangreiche Untersuchungen haben gezeigt, dass eine Hidden-Schicht ausreichend ist.

Anzahl der Knoten in der Hiddenschicht:

- Bedingungen:
- voll vernetzt
 - Feedforward-Netzwerk

Regel 1: Je komplexer die Beziehungen zwischen Ein- und Ausgangsdaten, umso mehr Knoten muss die Hidden-Schicht enthalten.

Regel 2: Die obere Grenze für die Zahl der Knoten in der Hidden-Schicht ergibt sich nach folgender Vorschrift

$$h = \frac{\text{Anzahl der Trainingsdaten}}{10 \times (m+n)}$$

mit h = Anzahl der Knoten in der Hidden-Schicht
 m = Anzahl der Knoten in der Output-Schicht
 n = Anzahl der Knoten in der Input-Schicht

gewöhnlich: $5 \leq h \leq 10$

starke Störungen: $10 \leq h \leq 100$

3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.5 Vorhersage mit neuronalen Netzen – Hinweise für die Verwendung des neuronalen Netzes bei der Vorhersage (1)

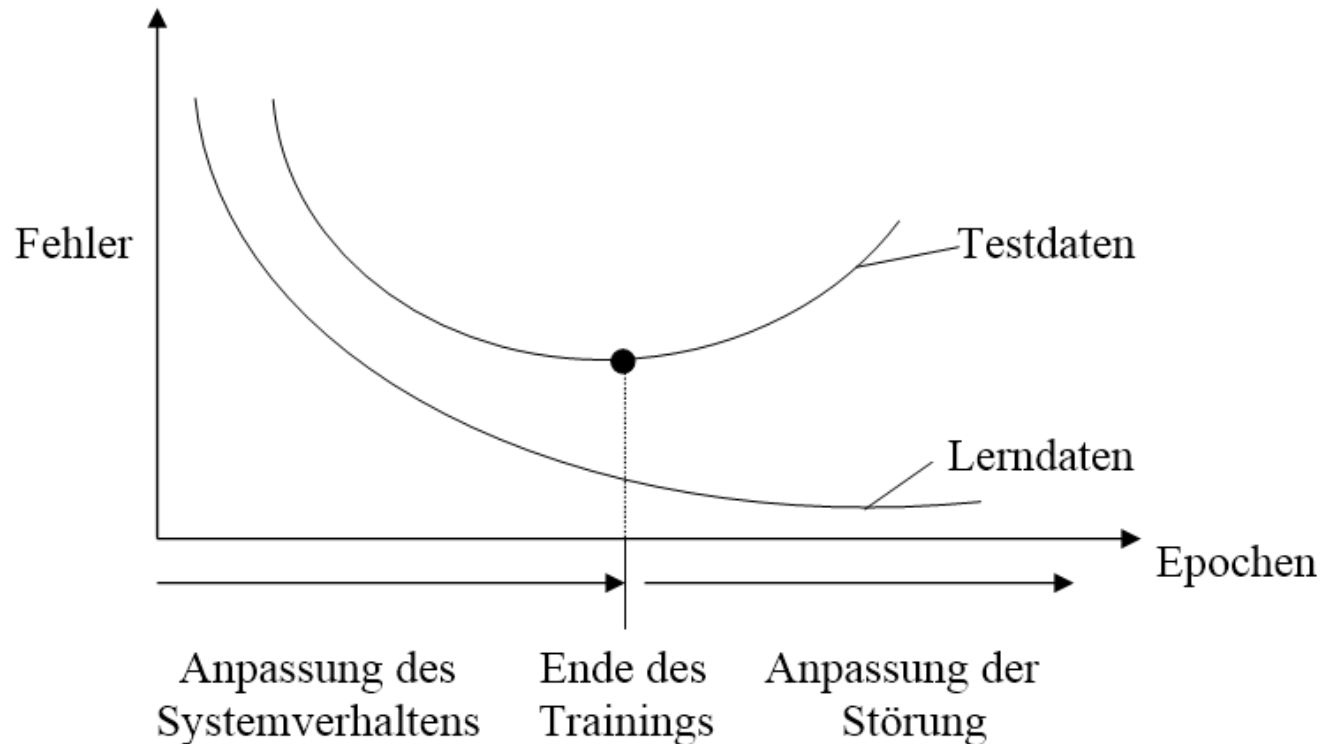
- Normierung der Daten zwischen 0 und 1 oder -1 und +1 (verwendete Aktivierungsfunktionen beachten).
- Die Anzahl der Epochen sollte klein gehalten werden. Dies erhöht die Generalisierungsfähigkeit des neuronalen Netzes.
- Bei Verbrauchsvorhersagen sollten Trainingsdaten von mindestens einer Periode, besser von zwei Perioden vorliegen, z. B. 1 Jahr oder 2 Jahre.
- Der Datensatz muss in Trainings- und Testdaten eingeteilt werden, z. B. 2/3 des Datensatzes werden zum Training verwendet und 1/3 zum Testen. Das Training sollte möglichst über eine volle Periode der Daten erfolgen. Liegen nicht genügend Daten vor, kann die Teilung des Datensatzes anders erfolgen, z. B. gerade Wochen werden zum Training genutzt, ungerade Wochen zum Testen.

3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.5 Vorhersage mit neuronalen Netzen – Hinweise für die Verwendung des neuronalen Netzes bei der Vorhersage (2)

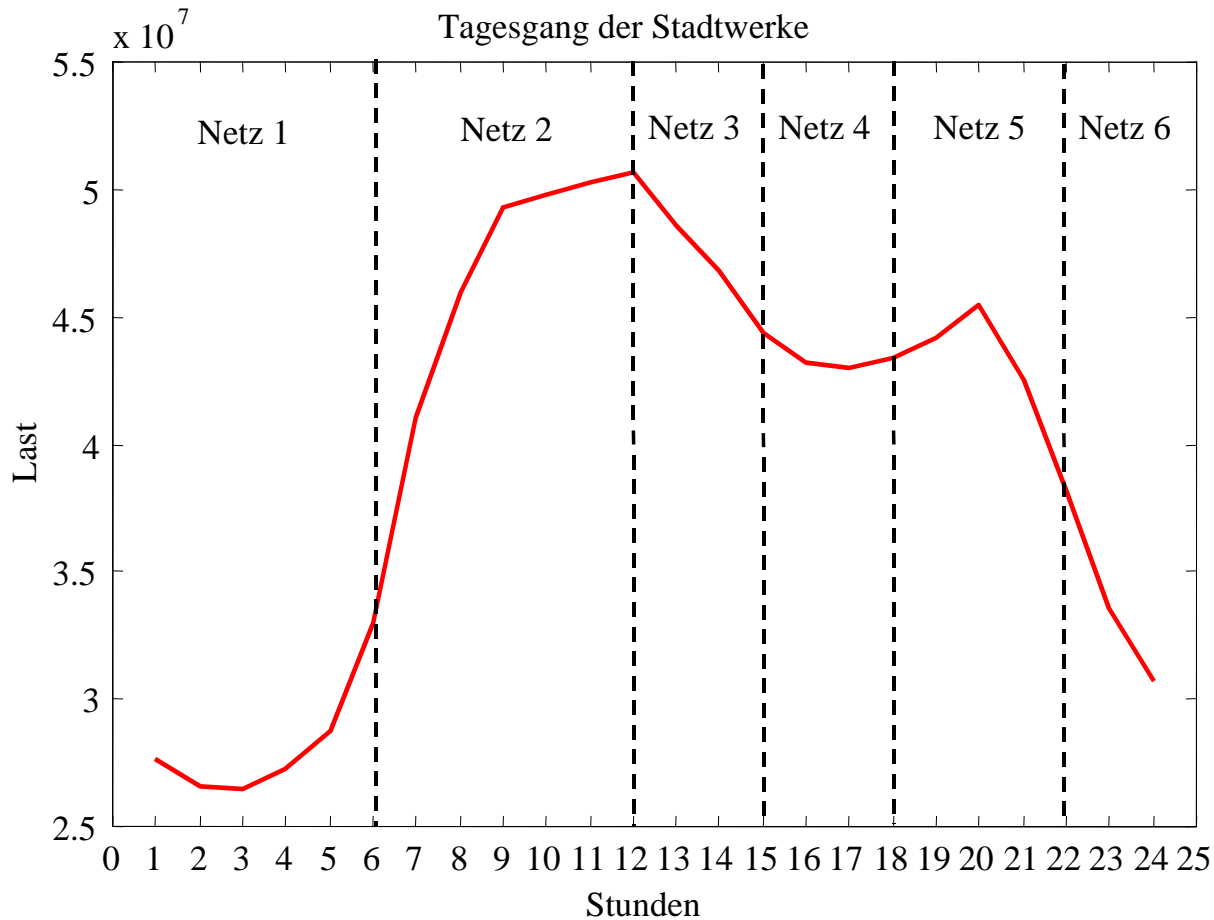
- Abbruch des Trainings



3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.5 Vorhersage mit neuronalen Netzen – Vorhersage der elektrischen Last (1)



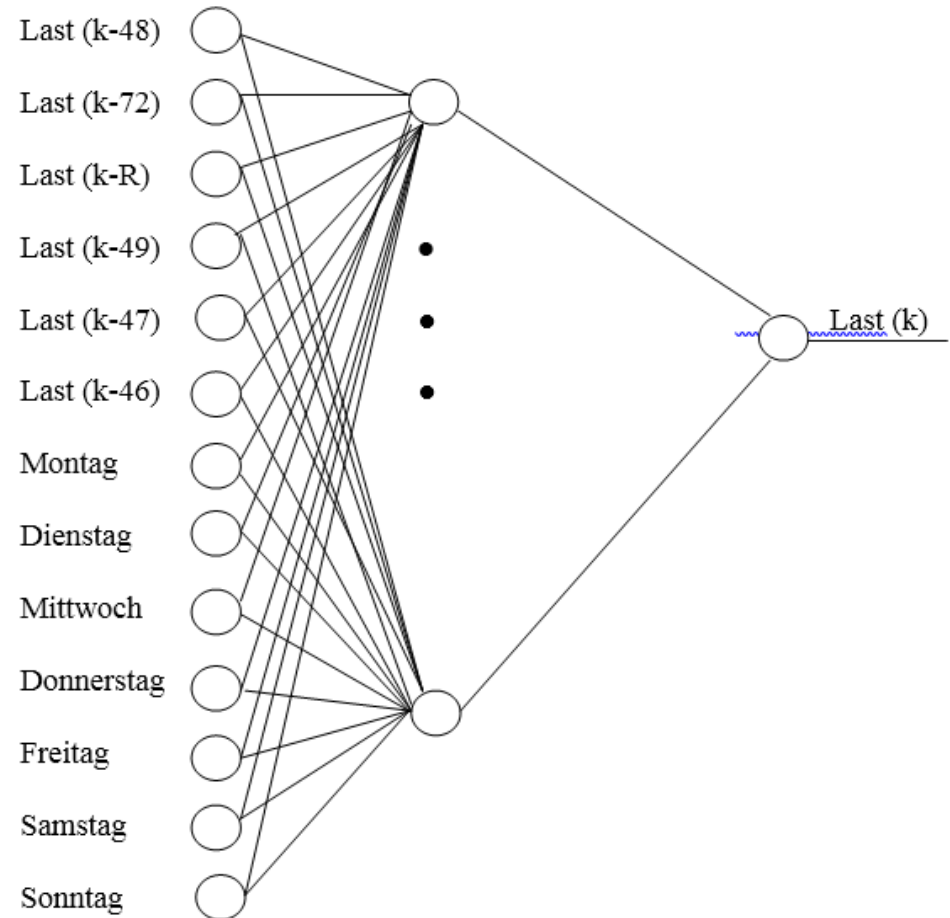
3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.5 Vorhersage mit neuronalen Netzen – Vorhersage der elektrischen Last (2)

Struktur des Netzes:

k = Stunde, für die Vorhersage erfolgt
 R = Referenzzeitraum
Anzahl Hidden-Knoten: 6 und 12
Epochen: 19

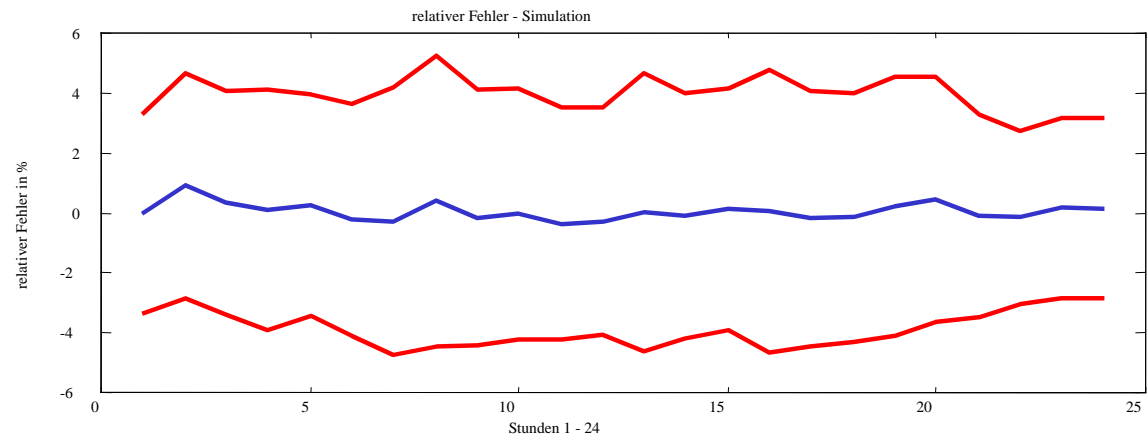
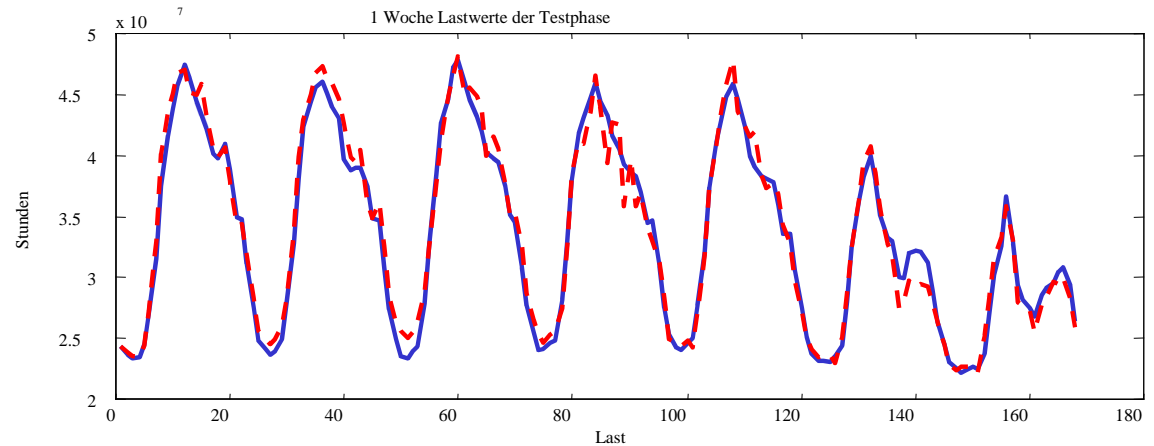


3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.5 Vorhersage mit neuronalen Netzen – Vorhersage der elektrischen Last (3)

Vorhersage mit
neuronaalem Netz mit 6
Hidden-Knoten

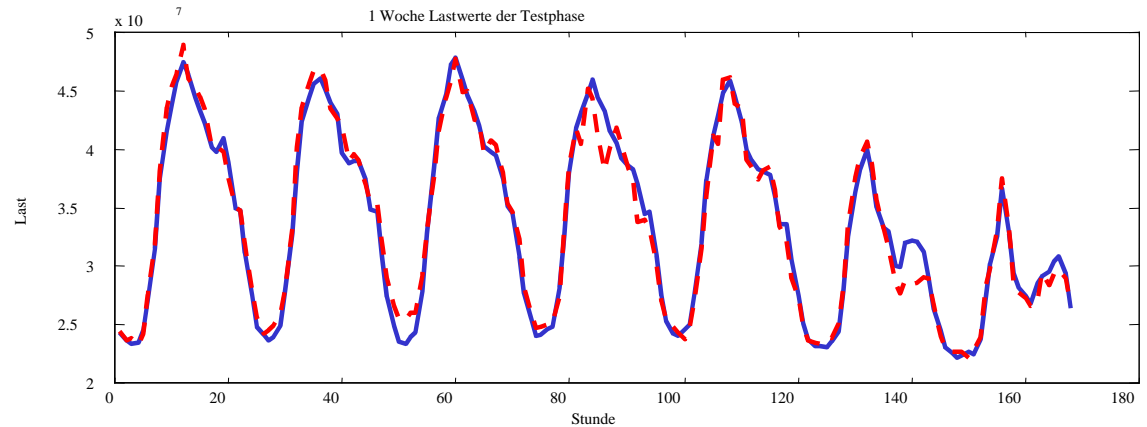


3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.5 Vorhersage mit neuronalen Netzen – Vorhersage der elektrischen Last (4)

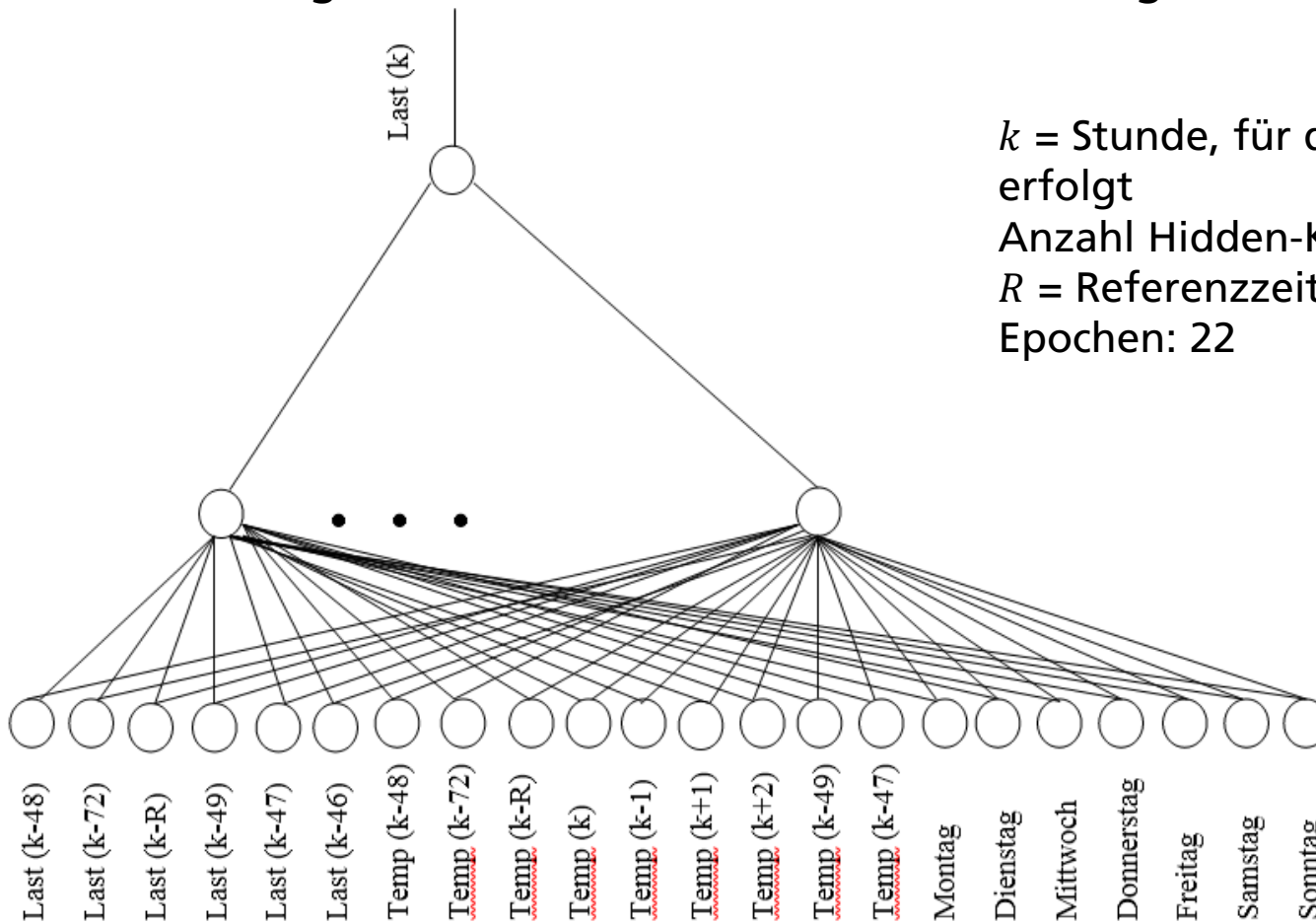
Vorhersage mit neuronalem Netz mit 12 Hidden-Knoten



3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.5 Vorhersage mit neuronalen Netzen – Vorhersage der elektrischen Last (5)



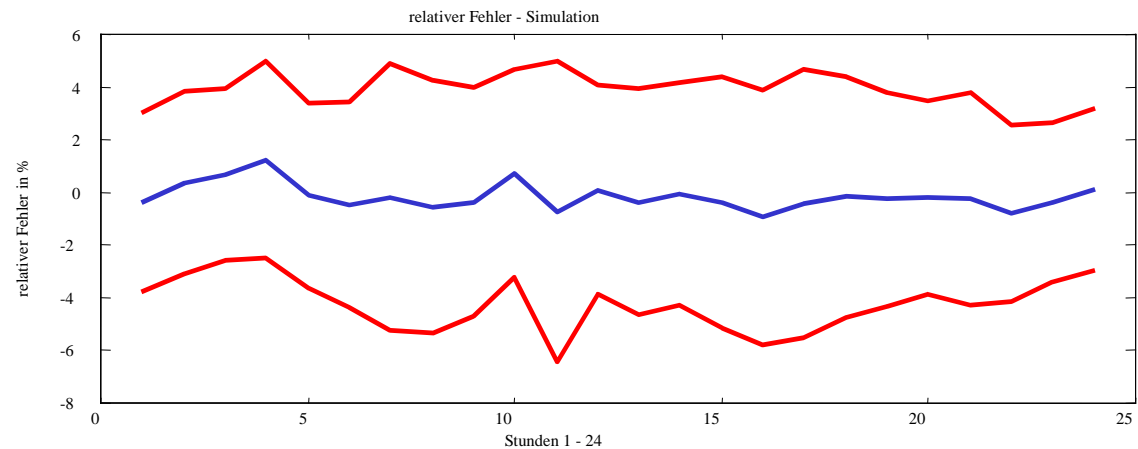
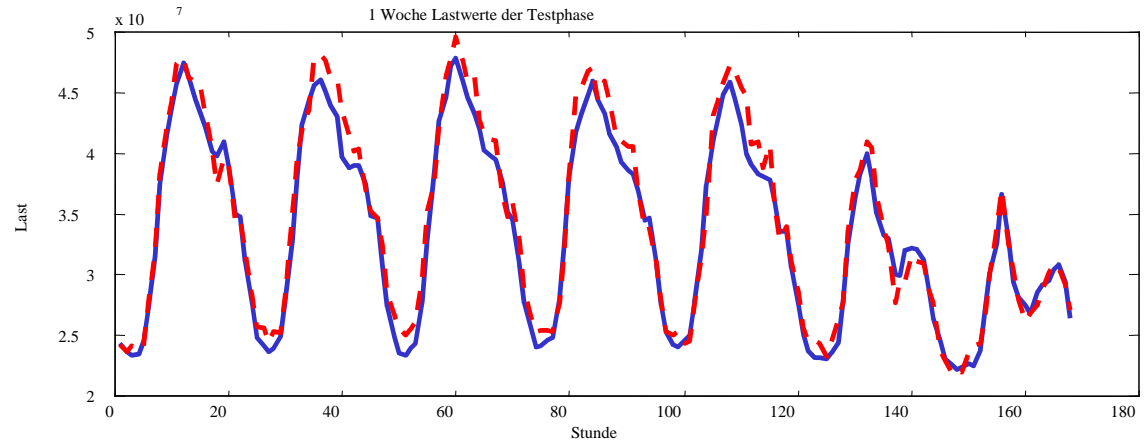
k = Stunde, für die Vorhersage erfolgt
Anzahl Hidden-Knoten: 6 und 12
 R = Referenzzeitraum
Epochen: 22

3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.5 Vorhersage mit neuronalen Netzen – Vorhersage der elektrischen Last (6)

Vorhersage mit neuronalem Netz mit 6 Hidden-Knoten und Einbeziehung der Temperatur

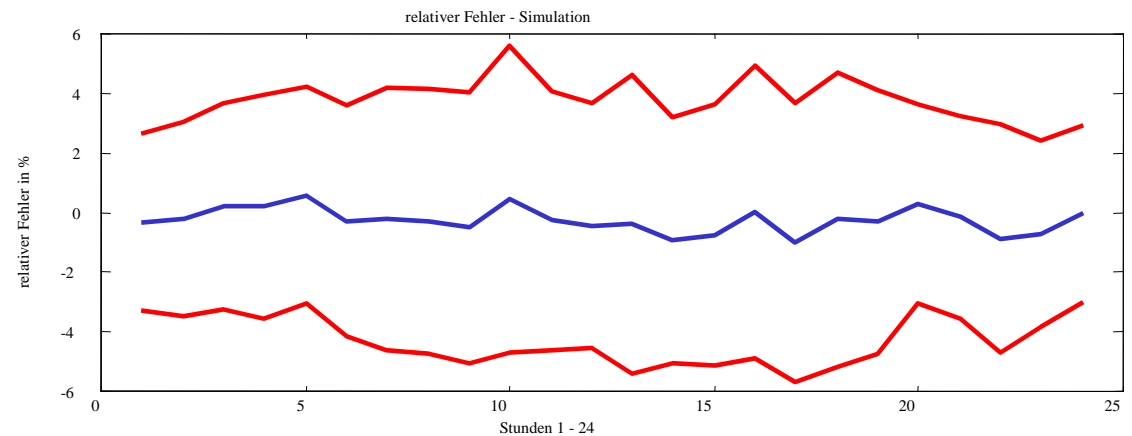
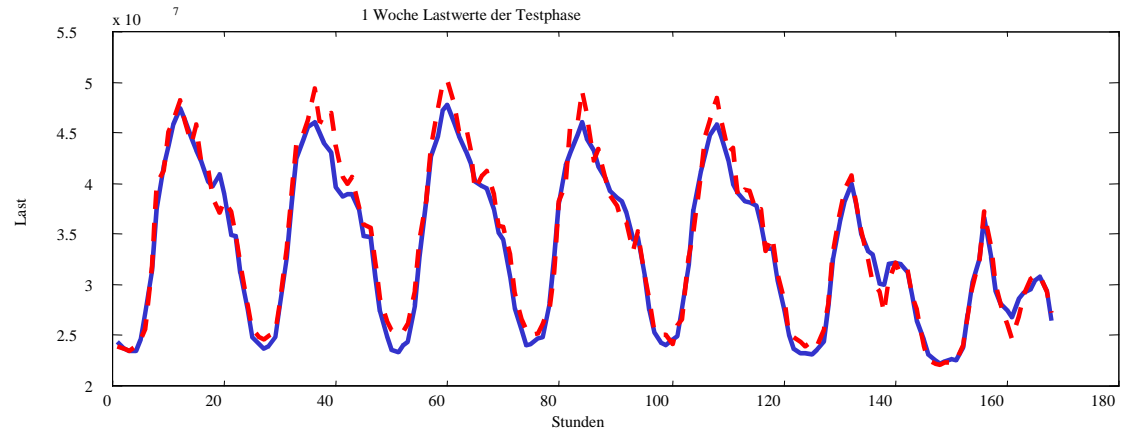


3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.5 Vorhersage mit neuronalen Netzen – Vorhersage der elektrischen Last (7)

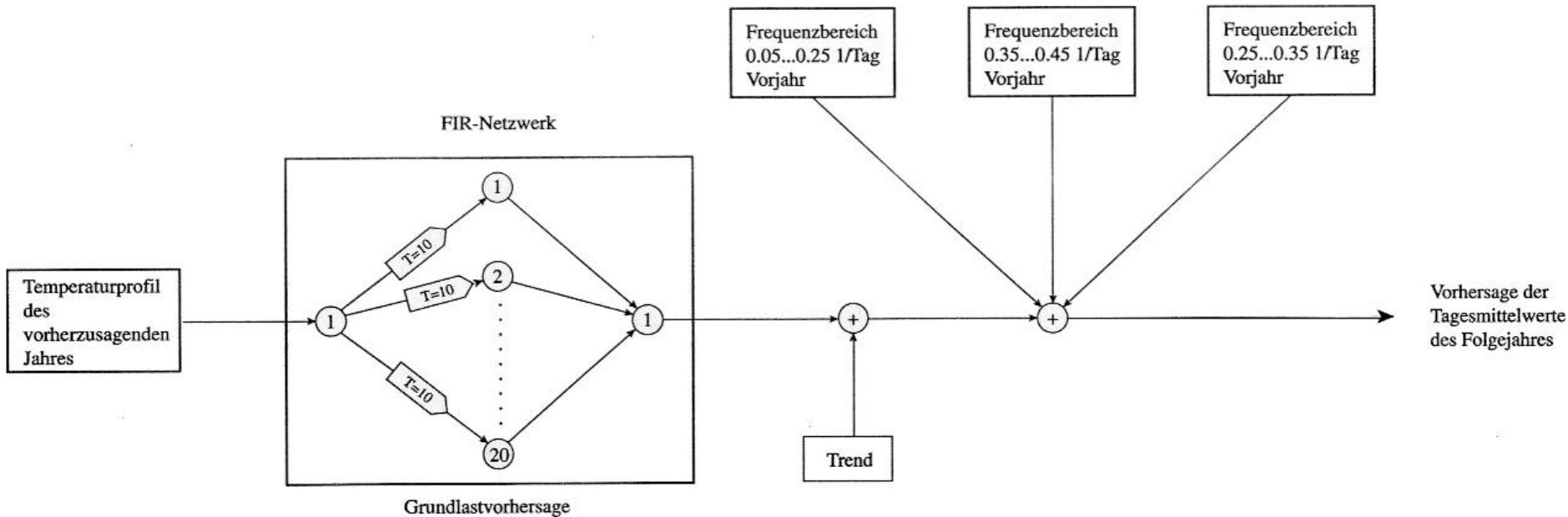
Vorhersage mit neuronalem Netz mit 12 Hidden-Knoten und Einbeziehung der Temperatur



3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

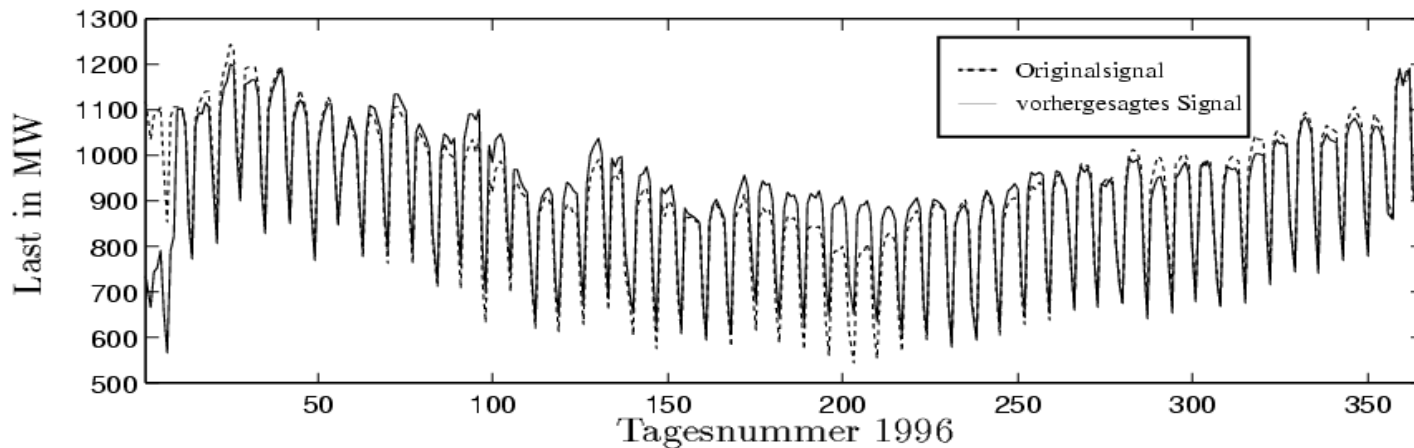
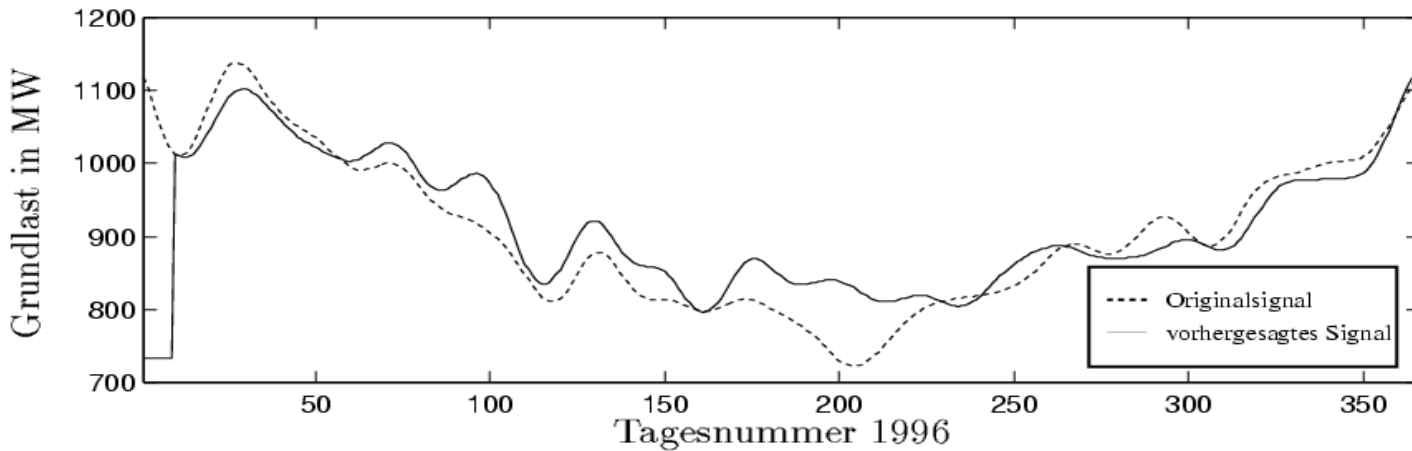
3.6.5 Vorhersage mit neuronalen Netzen – Jahresvorhersage der elektrischen Last mit FIR-Netzwerken



3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.5 Vorhersage mit neuronalen Netzen – Ergebnisse der Jahresvorhersage



3. Vorhersagestrategien

3.6 Vorhersage mit Signalmodellen

3.6.5 Vorhersage mit neuronalen Netzen – Zustandsraum mit neuronalem Netz

$$\underline{x}(k+1) = \phi(\underline{x}(k), \underline{u}(k))$$
$$\underline{y}(k) = \psi(\underline{x}(k))$$

Voraussetzung:

- alle Zustände messbar
- stabiler Prozess

