

**KfK 3872**  
**Februar 1985**

# **Prognoseverfahren in parametrischen Signalmodellen zur Fehlerfrühdiagnose**

**K.P. Scherer**  
**Institut für Datenverarbeitung in der Technik**  
**Projekt Schneller Brüter**

**Kernforschungszentrum Karlsruhe**



KERNFORSCHUNGSZENTRUM KARLSRUHE

Institut für Datenverarbeitung in der Technik

Projekt Schneller Brüter

KfK 3872

Prognoseverfahren in parametrischen Signalmodellen  
zur Fehlerfrühdiagnose

K. P. Scherer

Kernforschungszentrum Karlsruhe GmbH, Karlsruhe

Als Manuskript vervielfältigt  
Für diesen Bericht behalten wir uns alle Rechte vor

Kernforschungszentrum Karlsruhe GmbH  
ISSN 0303-4003

## Zusammenfassung

---

Zur frühzeitigen Schadensdetektion, etwa zur Detektion kritischer Siedezustände mit akustischen Signalen ist es wichtig, typische Muster im Signal schnell und sicher zu erkennen. Über eine Überwachung mit parametrischen Signalmodellen ist noch recht wenig bekannt. In dieser Arbeit wird gezeigt, wie man das akustische Signal in den Merkmalsraum transformieren und in diesem eine mathematische Modellbildung vornehmen kann. Die Möglichkeit, den physikalischen Prozeß auf diese Weise zu beschreiben, ermöglicht die Anwendung von Prognosemodellen im Merkmalsraum. Somit kann eine aufwendige Klassifikation bereits vorgezogen werden.

Prediction methods in parametric signal models for early failure detection

## Abstract

---

For early detection of failures and critical phases in reactors by acoustic signals, it is important to recognize typical features very fast. A surveillance by parametric signal models is rather unknown. In this proposal it is shown, how to transform the acoustic signal in the feature space and how to generate several mathematical models in the feature space for the process. The ability to describe the physical process in such a way gives the possibility to apply forecast methods and different mathematical models and so the classification step can be preceded.

## Inhaltsverzeichnis

1. Einleitung.....	1
2. Globales Konzept des adaptiven Systems .....	2
3. Generierung statistischer Primärmerkmale.....	5
4. Prognoseverfahren im Merkmalsraum.....	8
4.1 Das Linear Prediction Verfahren.....	8
4.2 Schätzmethoden mit einem gewöhnlichen Dgl.system.....	11
4.2.1 Stabilitätsaussagen.....	12
4.2.2 Modellbildung im Merkmalsraum.....	14
4.3 Kalmanfilter für zeitdiskrete stochastische Prozesse.....	17
4.4 Das Berliner Verfahren.....	24
4.5 Arima-Modellierung im Merkmalsraum.....	26
5. Auswertung und Ergebnisse.....	34
6. Zusammenfassung und Ausblick.....	37
7. Literatur.....	45
8. Anhang.....	47

Abbildungsverzeichnis:

<u>Abb 1:</u>	Konzept eines adaptiven Systems.....	4
<u>Abb 2:</u>	ARIMA - Modellierung im Merkmalsraum.....	32
<u>Abb 3:</u>	Iterative Dimensionierung des ARIMA Modells.....	33
<u>Abb 4:</u>	Überwachungskonzept.....	37
<u>Abb 5:</u>	Tabelle der angewandten Prognoseverfahren.....	42
<u>Abb 6-7:</u>	Rohdatenplott.....	47-48
<u>Abb 8:</u>	Geglättetes Powerspektrum.....	49
	a) mit Rechteckfenster	
	b) mit Hanningfenster	
<u>Abb 9-14:</u>	Merkmalsplott.....	50-55

## 1. Einleitung

Im Sinne der Schadensfrüherkennung, insbesondere im Bereich gefährlicher Anlagen, ist es wichtig, eventuell auftretende Störungen schon im Ansatz zu erkennen, d.h. bevor die Auswirkung ein signifikantes Maß überschreitet. Im Falle der vorliegenden Studie bedeutet dies, den Siedezustand bereits in einem frühen Stadium abzuschätzen, sodaß noch rechtzeitig Korrektur- bzw. Abhilfemaßnahmen ergriffen werden können.

Die Entscheidungskriterien basieren auf wahrscheinlichkeitstheoretischen Aussagen. Zeitkritische Aspekte geben gewisse Randbedingungen vor, innerhalb denen die Entscheidungen zu treffen sind. Eine weitere wichtige Komponente zur Entscheidungsfindung ist die Vorgabe von a priori Wissen, welches qualitativ und quantitativ ausreichend zur Verfügung stehen muß, bzw. welches man erst geeignet transformieren muß, um eine günstige Ausgangsbasis zur Entscheidungsfindung zu haben.

Vorrangig in dieser Arbeit ist die Untersuchung folgender Probleme:

- geeignete Transformation des Datenmaterials in eine Form, die der Entscheidungsstufe direkt zugänglich ist
- Festlegung statistischer Kenngrößen als Primärmerkmale
- Entwicklung verschiedener Fortsetzungs- und Prognoseverfahren der Parameter im Merkmalsraum
- Anwendung der entwickelten Prognoseverfahren auf ein Signal
- Generierung von Sekundärmerkmalen aus den Prognoseparametern

Es werden aus den unterschiedlichen Prognoseverfahren Modellparameter extrahiert, die als Sekundärmerkmale zusammen mit den prognostizierten Primärmerkmalen eine frühzeitige Klassifikation ermöglichen.



## 2. Globales Konzept des adaptiven Systems

Im Rahmen der frühen Schadensdetektion wird ein System entwickelt, welches ausgehend von einem unkritischen Normalzustand Fehlzustände, insbesondere das Aufkommen von Sieden schnell und sicher erkennt. Im Sinne der Mustererkennung wird eine Klassifikation stattfinden in normale und anomale (kritische) Zustände. Im folgenden wird näher eingegangen auf die Punkte

1. Merkmalsgenerierung
2. Modellanpassung
3. Entscheidung

ad 1 Merkmalsgenerierung:

Diese Phase umfaßt sowohl die Extraktion wie auch die Merkmalsreduktion. Im Unterschied zur herkömmlichen Merkmalsgenerierung ist der Prozeß der Merkmalsbildung nicht mit einem festen Satz von Merkmalen abgeschlossen. Es können im Laufe des Prozesses neue Modellparameter als weitere Merkmale hinzukommen.

Zunächst wird ein Satz von Primärmerkmalen bereitgestellt. Er enthält eine Menge statistischer Parameter und Kenngrößen, die das typische von stochastischen Schwankungen überlagerte Signal darstellen. Es kommen die in Kapitel 3 gewählten Parameter zur Geltung.

Ein Satz von Sekundärmerkmalen wird im Laufe der Prognose generiert. Während Primärmerkmale aus dem Signal zu berechnen sind, basieren die Sekundärmerkmale auf dem Modell der Fortsetzbarkeit des Signals bzw. seiner Kenngrößen und sind mit entsprechenden stochastischen Fehlern behaftet. Trotzdem kann der Informationsgehalt gesteigert werden.

ad 2 Modellanpassung:

Durch die fortlaufende Merkmalsgenerierung und die damit verbundene neue Parametrisierung des Modells wird das System stets aktualisiert und angepaßt. Die Prognose und Entscheidungsfindung benutzt jeweils den letzten aktuellen Stand des Prozesses als Wissensbasis. In diesem Sinne ist auch die Trainingsphase nicht an feste Muster gebunden.

ad 3 Entscheidungsphase:

Die Klassifikationsstufe wird sich vom herkömmlichen Klassifikationsprozeß dadurch unterscheiden, daß der kritische Zustand 'Sieden' über mehrere Entscheidungsprozesse gemäß verschiedener Merkmalsstufen detektiert wird. Voraus geht eine Lernphase über beobachtete Referenzdaten. Gegenüber herkömmlichen Methoden hat man durch die Prognose und Schätzverfahren bereits eine Menge von Vorentscheidungen, die eine endgültige Aussage zuverlässiger und schneller machen.

Einen Überblick über das Konzept gibt Abb. 1.

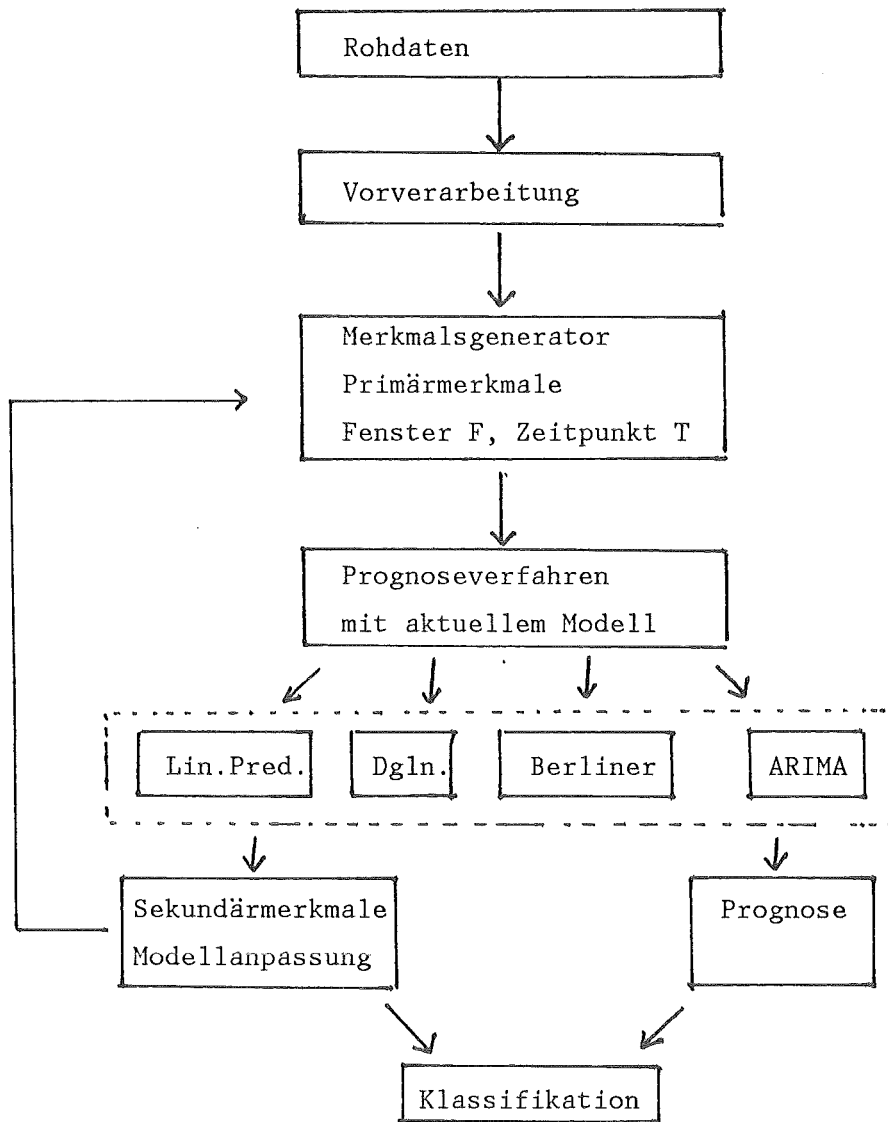


Abb. 1: Konzept des adaptiven Systems

## 2. Generierung statistischer Primärmerkmale

Es wird zunächst eine Grundmenge von Merkmalen generiert, ohne Beachtung irgendwelcher Dimensionsreduktionen oder Optimierungskriterien. Da es sich bei dem Signal um einen stochastischen Prozeß handelt, werden im wesentlichen statistische Kenngrößen als Merkmale verwendet. Der Berechnung aller folgenden Parameter  $p_1, \dots, p_n$  liegt ein Intervall  $I$  der momentanen Beobachtung zugrunde.

I besitze  $N$  Punkte des RMS Signals  $(x_1, \dots, x_N)$  mit  $x_i \geq 0$  für alle  $i$ .

Es werden Kenngrößen im Zeit- und im Frequenzbereich generiert.

p1:  $\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i$  zeitlicher Mittelwert

p2:  $\sigma^2 = \frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})^2$  Varianz

p3:  $\bar{x}^2 = \frac{1}{N} \sum_{i=1}^N x_i^2$  Energie

Für eine normierte Zufallsvariable ( $E\{y\} = 0; \sigma^2 = 1$ )  $y$  sei  $\varphi(y)$  die Normalverteilungsdichte; dann kann man für eine beliebige Wahrscheinlichkeitsdichte  $w(y)$  folgende Reihenentwicklung angeben

$$w(y) = \varphi(y) \left( 1 + \sum_n \frac{c_n}{n!} H_n(y) \right), \quad \text{wobei die } c_n \text{ Quasimomente}$$

heißen. Die  $c_n$  charakterisieren die Form der Verteilungsdichte; Die geraden geben ein Maß für die Flachheit an, die ungeraden drücken die Symmetrie bzw. die Asymmetrie aus.

Mit  $\mu_n = E\{(x - E\{x\})^n\}$  als Zentralmomente gilt

p4:	$c_3 = \mu_3$	}	Quasimomente
p5:	$c_4 = \mu_4 - 3\mu_2^2$		
p6:	$c_5 = \mu_5 - 10\mu_2\mu_3$		

Weitere Merkmale werden im Frequenzbereich generiert. Nach einer FFT steht folgender positiver reeller Vektor zur Verfügung.

$$(F_m)_m : |F(m)|^2 = \left| \sum_{n=1}^N x_n \cdot w(n) \cdot e^{-2\pi i m n / N} \right|^2 \quad m = 1, \dots, N/2 + 1 =: N_p$$

mit  $w(n)$  als Rechteckfenster bzw. Hanningfenster

$$p7: \quad \bar{F} = \frac{1}{N_p} \sum_1^{N_p} F_i$$

$$p8: \quad G_F^2 = \frac{1}{N_p} \sum_1^{N_p} (F_i - \bar{F})^2$$

$$p9: \quad \text{Quasimomente analog zum Zeitbereich} : \mu_n = E \{ (F_i - E(F_i))^n \}$$

$$p10: \quad "$$

$$p11: \quad "$$

$$p12: \quad \log \left( \frac{\sum_1^{N_p} F_v}{\sum_1^{N_p} F_v^*} \right) \text{ wobei sich * auf ein Referenzfenster bezieht}$$

$$p13: \quad \min_v \log (F_v / F_v^*)$$

$$p14: \quad \max_v \log (F_v / F_v^*)$$

$$p15: \quad \frac{1}{N_p} \sum_1^{N_p} \log (F_v / F_v^*)$$

$$p16: \quad \frac{1}{N_p} \sum_1^{N_p} [\log (F_v / F_v^*)]^2$$

$$\text{Mit } r_v = \text{sign} [\log (F_v / F_v^*)] :$$

$$p17: \quad \max(\text{Anzahl } r_v > 0, \text{Anzahl } r_v < 0)$$

$$p18: \quad \text{Anzahl von Intervallen mit gleichem Vorzeichen}$$

p19: maximale Länge gleichen Vorzeichens

Die Liste dieser Primärmerkmale kann beliebig fortgesetzt werden.  
In dieser Merkmalssammlung ist zunächst noch keine Beurteilung und  
Bewertung der einzelnen Kenngrößen enthalten.

#### 4. Prognoseverfahren im Merkmalsraum

##### 4.1 Das Linear Prediction Verfahren

In diesem Kapitel wird zunächst unter deterministischem Gesichtspunkt ein Verfahren für die Merkmalsprognose entwickelt. Das Linear Prediction Verfahren wird auf eine noch näher zu spezifizierende Zeitreihe angewandt. Es kann z.B. die um stochastische Schwankungen bereinigte Funktion eines statistischen Parameters sein.

Aus der Beobachtung früherer Daten wird ein linearer Modellansatz gemacht. Der momentane Wert läßt sich als Linearkombination aus  $m$  früheren Werten berechnen. Die Modellparameter werden so bestimmt, daß der mittlere zu erwartende Fehler bei der vorgegebenen Stichprobenmenge minimal ist.

##### Parameterschätzung:

Gegeben sei eine digitalisierte diskrete eindimensionale Funktion  $f$  mit den Abtastwerten  $f_j$ . Es stehen nur endlich viele Werte zur Verfügung;  $f_1, \dots, f_{n-1}$

Gesucht ist ein Schätzwert für  $f$  zum Zeitpunkt  $n$ .

Linearer Ansatz: 
$$\hat{f}_n = \sum_{\mu=1}^m \alpha_{\mu} f_{n-\mu} \quad (4.1.1)$$

Interpretation des Ansatzes:

Der Wert  $\hat{f}_n$  läßt sich als Linearkombination aus den  $m$  vorangehenden Werten mit den Gewichten  $\alpha_{\mu}$  berechnen. Die Koeffizienten  $\alpha_{\mu}$  sind unbekannt und müssen aus den Nebenbedingungen berechnet werden.

Sei  $f_n$  der wahre Wert zum Zeitpunkt  $n$ ,  $\hat{f}_n$  der geschätzte, dann gilt für den Fehler zum Zeitpunkt  $n$

$$e_n = f_n - \hat{f}_n = f_n - \sum_{\mu=1}^m \alpha_{\mu} f_{n-\mu} \quad (4.1.2)$$

Unter der Nebenbedingung für den mittleren quadratischen Fehler

$$E = \sum_n (e_n)^2 \rightarrow \text{Min} ! \quad (4.1.3)$$

zur Bestimmung der Koeffizienten ist folgende Extremwertaufgabe zu lösen

$$E = \sum_n (f_n + \sum \alpha_\mu f_{n-\mu})^2 \rightarrow \text{Min}!$$

$$\frac{\partial E}{\partial \alpha_i} = 0 \quad i=1, \dots, m \quad \Leftrightarrow \quad - \sum_n f_n f_{n-i} = \sum_{\mu=1}^m \alpha_\mu \sum_n f_{n-\mu} f_{n-i}$$

Mit  $R(\mu, i) := \sum_n f_{n-\mu} \cdot f_{n-i}$  und  $R(0, i) := \sum_n f_n \cdot f_{n-i}$

erhält man folgendes zu lösendes Gleichungssystem in den  $\alpha_i$ .

$$\sum_{\mu=1}^m \alpha_\mu R(\mu, i) = - R(0, i) \quad 1 \leq i \leq m \quad (4.1.4)$$

Bemerkung:

Die Größe R ist eine Summation über n.

1. Fall: Summation über  $-\infty < n < \infty$

Dann gilt:  $\sum_n f_{n-\mu} f_{n-i}$  hängt nur von  $|\mu-i|$  ab.

Somit lautet die Normalgleichung

$$\sum_{\mu=1}^m \alpha_\mu R(i-\mu) = - R(i) \quad (4.1.5)$$

Es gilt  $R$  ist symmetrisch  $R(i, \mu) = R(\mu, i)$

$R$  ist konstant entlang der Diagonalen  $R(0) = R(A, A)$

2. Fall: Summation über  $1 \leq n \leq N-1$

Dann gilt:  $\sum_n f_{n-\mu} f_{n-i}$  hängt von  $\mu$  und von  $i$  ab

Somit lautet die Normalgleichung

$$\sum_{\mu=1}^m \alpha_\mu K(\mu, i) = - K(0, i) \quad (4.1.6)$$

Da in unserem Falle endlich viele Stützpunkte vorhanden sind, wird der

2. Fall benutzt.

Abschätzung des mittleren quadratischen Fehlers

$$E = \sum_n f_n^2 + \sum_{\mu=1}^m \alpha_\mu \sum_n f_n f_{n-\mu} \quad (4.1.7)$$



Im 1. Fall erhält man 
$$E = R(0) + \sum_{\mu=1}^m \alpha_{\mu} \cdot R(\mu)$$

Im 2. Fall erhält man 
$$E = K(0,0) + \sum_{\mu=1}^m \alpha_{\mu} \cdot K(0, \mu)$$

Der Lösungsvektor  $a$  des Gl.systems  $K(i,j) * a = K(0,i)$  wird als Sekundärmerkmalsvektor benutzt. Des weiteren dienen diese Merkmale  $\alpha_i$  zur Prognose des weiteren Kurvenverlaufes der Primärmerkmale.

#### Schätzung zukünftiger Systemzustände

Aus der momentanen Merkmalsgenerierung resultieren die Modellparameter  $(\alpha_{\mu})_{\mu=1, \dots, m}$ . Seien also  $f_n, f_{n+1}, \dots$  zum gegenwärtigen Zeitpunkt gegeben. Für  $t_{n+j}$  gilt dann

$$f_{n+1} = - \sum_{\mu=1}^m \alpha_{\mu} \cdot f_{n+1-\mu}$$

für  $t_{n+j}$  entsprechend

$$f_{n+j} = - \sum_{\mu=1}^m \alpha_{\mu} \cdot f_{n+j-\mu}$$

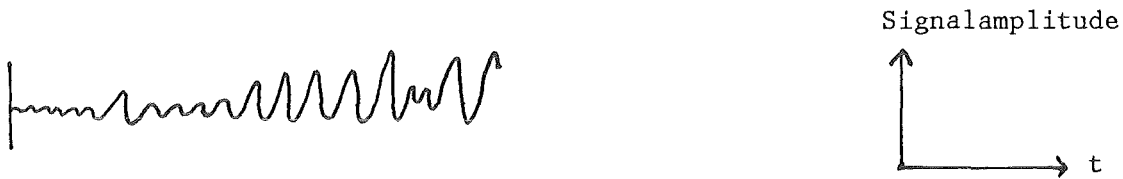
(4.1.8)

Die Parameter  $\alpha_{\mu}$  sind fest durch das bisherige Verhalten des Systems bestimmt und werden für die Abschätzung in den Zeitpunkten als fix betrachtet. Die Größe  $j$  für die Länge des Prognoseintervalls sei noch offen aber fest gewählt. Mit wachsendem  $j$  nimmt die Sicherheit der Prognose ab. Man beachte, die  $f_j$  sind nicht etwa Funktionswerte des ursprünglichen Signals, sondern Schätzwerte für statistische Parameter, die jeder für sich ein ganzes Intervall beschreiben. Der 2. Fall (endliches Datenmaterial) findet im Programm seine Anwendung.

4.2 Schätzmethoden mit einem gewöhnlichen Dgl.system

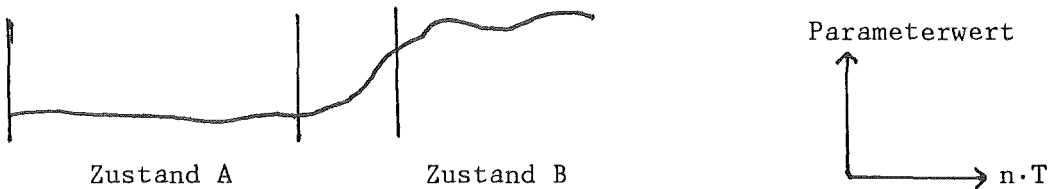
Unabhängig vom Linear Prediction Verfahren wird eine weitere Methode entwickelt, wie man aus der Beobachtung gegenwärtiger Muster auf zukünftige schließen kann.

Bild 1 zeigt ein ursprünglich digitalisiertes Signal



Zu diesem Signal sei eine Funktion  $f$  erzeugt, etwa eine statistische Kenngröße, die im wesentlichen die Information des Signals enthält. Stochastische Schwankungen dieser Parametergröße seien zunächst unberücksichtigt.

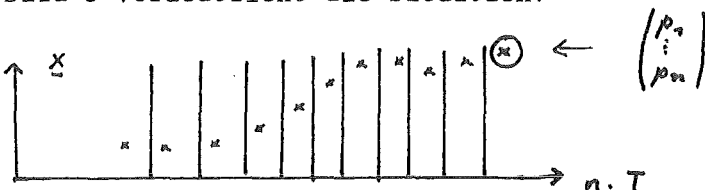
Bild 2 zeigt eine Parameterfunktion. Die Funktionswerte beziehen sich auf ein Intervall. Es ist eine Funktion der Zeit im Merkmalsraum.



Die ursprüngliche Signalfunktion sei äquidistant unterteilt in Intervalle der Länge  $T$ . In der Merkmalsextraktionsstufe werden für jedes  $T$  die Primärmerkmale gemäß Kapitel 3 berechnet und in einem Vektor der Dimension  $n$  zusammengefaßt.

Wir bezeichnen diesen Vektor mit  $\underline{x} = (p_1 \dots p_n)^T$ , wobei  $p_i \quad i \in \{1, \dots, n\}$  die einzelnen Parameter sind.

Bild 3 verdeutlicht die Situation.



Betrachtet man die Parameterfunktion als Lösung einer Dgl., so erwartet man im Normalzustand ein gewisses 'stabiles' Verhalten. Kommt man in den Bereich der Übergangsphase, wird das Verhalten gewissermaßen instabil. Die Stabilitätsbegriffe werden noch genauer definiert.

Ziel der folgenden Untersuchung ist es nun, aus den berechneten Merkmalen ein Dgl.system zu erstellen, um direkt aus dem System Stabilitätsaussagen machen zu können. Somit kann die Lösung lokal fortgesetzt in ihrem Verhalten abgeschätzt werden.

Bemerkung:

1. Dieses Verfahren steht in Ergänzung zum Linear Prediction Verfahren; Hier wie dort werden im Merkmalsraum statistische Parameter als Primärmerkmale prognostiziert und Modellparameter als Sekundärmerkmale verwendet.
2. Durch Betrachtung des linearen Systems kann man, was stabiles und asymptotisches Verhalten angeht, auf das ursprüngliche System schließen. Insofern genügt die Untersuchung der linearisierten Dgl..
3. Das Ergebnis ist in beiden Fällen abhängig von der Merkmalsauswahl.

4.2.1 Stabilitätsaussagen

Im folgenden werden die Stabilitätsbegriffe und die Auswirkungen am System näher erläutert. Für die exakte mathematische Definition von Stabilität, Attraktivität und asymptotischer Stabilität sei auf die einschlägige Literatur verwiesen. ( /7/ )

Man betrachte das Anfangswertproblem (AWP)  $\dot{x} = f(t, x)$ ,  $f(t, 0) = 0$  mit gewissen Stetigkeitsvoraussetzungen an  $f$ .

Def 1: Die Lösung  $x = 0$  des Systems heißt stabil  $\Leftrightarrow$

Jede Lösung des AWP mit einem Startwert  $x_0 \in (-\delta, +\delta)$  existiert für alle Zeit und bleibt in einer  $\epsilon$ -Umgebung der Nulllösung.

Def 2: Die Lösung heißt instabil, wenn sie nicht stabil ist.

Def 3: Die Lösung heißt attraktiv, wenn sie sich asymptotisch der Nulllösung nähert für  $t \rightarrow \infty$  .(i.a. nicht stabil !)

Def 4: Die Lösung heißt asymptotisch stabil, wenn die Bedingungen in Definition 1 und 3 erfüllt sind.

Es werden einige Aussagen gemacht, die später bei unserem Modell Anwendung finden werden.

Satz 1:

Für autonome (d.h. nicht explizit zeitabhängige) Systeme kann stabil bzw. asymptotisch stabil durch gleichmäßig stabil bzw. gleichmäßig asymptotisch stabil ersetzt werden.

Man betrachte nun den Spezialfall eines linearen autonomen Systems.

Satz 2:

- a) Haben alle Eigenwerte von A negativen Realteil, dann ist die Nulllösung  $\underline{x} \equiv 0$  glm. asymptotisch stabil.
- b) Haben alle Eigenwerte von A nichtpositiven Realteil und gibt es zu  $\lambda$  mit  $\text{Re}(\lambda) = 0$  der Vielfachheit k genau k linear unabhängige Eigenvektoren, dann ist  $\underline{x} \equiv 0$  glm. stabil.

Das Verhalten der Lösungen ist vollständig charakterisiert durch die Eigenwerte der Matrix A.

Hat A keine rein imaginären Eigenwerte, so ist A ähnlich einer

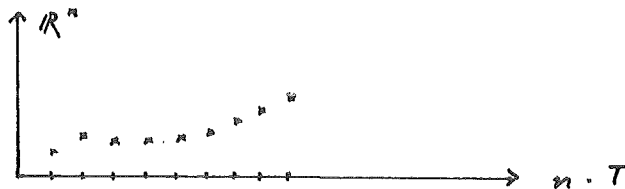
Matrix  $\begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix}$ , wobei  $A_1$  nur Eigenwerte mit negativem Realteil,

$A_2$  nur Eigenwerte mit positivem Realteil besitzt.

Dies bedeutet, daß der  $n$  dimensionale Lösungsraum  $X$  in 2 Komplementärräume aufgespalten wird. Man kann zeigen, daß Lösungen, die in  $X_1$  bzw. in  $X_2$  starten, auch in  $X_1$  bzw. in  $X_2$  bleiben. Man spricht dann auch von einer stabilen bzw. instabilen Mannigfaltigkeit. (S 247 /7/). Die Dimension der Mannigfaltigkeiten bzw. charakteristische Unterschiede in den Eigenwerten können zur Charakterisierung der Muster benutzt werden.

#### 4.2.2 Modellbildung im Merkmalsraum

Im  $n$  dimensionalen Merkmalsraum sei der Merkmalsvektor  $(p_1, \dots, p_n)^T$  im folgenden mit  $\underline{x}$  bezeichnet. Über die diskrete Zeitachse stellt  $\underline{x}(\cdot)$  die Trajektorie des Merkmalsvektors dar. Für jedes feste  $t_i = i \cdot T$  bezeichnet  $\underline{x}(t_i)$  den speziellen Merkmalsatz des Zustandes zum Zeitpunkt  $t_i$ . Folgendes Bild verdeutlicht die Situation.



$\underline{x}(\cdot)$  wird aufgefaßt als Lösung eines gewöhnlichen deterministischen Dgl.systems

$$\dot{\underline{x}} = f(t, \underline{x}) \quad (*)$$

$\underline{x}$  hänge also nicht von höheren Ableitungen ab.

#### Systemlinearisierung

Um Aussagen über das stabile Verhalten der Lösung  $\underline{x} = 0$  des Systems zu machen, genügt die Untersuchung des linearisierten Systems. Auch der Bezug zur  $0$  Lösung ist keine Einschränkung, da das System mit einer additiven Konstanten transformiert werden kann.

Mit einer Taylorentwicklung von  $f(t, \cdot)$  im Punkt  $\underline{x} = 0$  erhält man:

$$f(t, \underline{x}) = f(t, 0) + A(t) \cdot \underline{x} + r(t, \underline{x}) ; \quad A(t) = \left( \frac{\partial f_i}{\partial x_j} (t, 0) \right)_{ij}$$

Somit gilt für die j-te Komponente der Dgl.

$$f_j(t, \underline{x}) = \sum_{i=1}^n \frac{\partial f_j(t, 0)}{\partial x_i} x_i + r_j(t, \underline{x})$$

In der linearisierten Form lautet die Dgl. ( \* ) nun:  $( f(t, 0) = 0 )$

$$\dot{\underline{x}} = A(t) \underline{x} + r(t, \underline{x}) \quad \text{und in erster Näherung} \quad \dot{\underline{x}} = A(t) \cdot \underline{x} \quad (4.2.1)$$

Im folgenden gelte generell die Voraussetzung:  $f(t, \underline{x}) = f(\underline{x})$

Wir betrachten somit ein zeitinvariantes autonomes System.

Für das autonome System schreiben wir die linearisierte Form

$$\dot{\underline{x}} = f_x(0) \cdot \underline{x} \quad \text{mit} \quad f_x(0) = \left( \frac{\partial f_j}{\partial x_i} \right)_{ij} \quad (4.2.2)$$

Stabilitätsaussagen werden durch Untersuchung der Matrix  $A = f_x(0)$  gewonnen.

### Modellanpassung an die Problematik

Wir betrachten eine deterministische Lösung der noch aufzustellenden Dgl. in den Fenstern  $F_i$  ( $i=1, \dots, m$ ).

Die berechneten Merkmalsvektoren seien  $x^1 \dots x^m$ , wobei  $x^i = (p_1^i \dots p_n^i)^T$  der Merkmalsvektor des i-ten Fensterpunktes ist.

Annahme: Die Änderung hänge nur von dem beobachteten Wert  $\underline{x}$  selbst ab. Die Gleichung ist autonom, zeitliche exogene Größen gehen nicht explizit ein.

Ansatz für die linearisierte Form: 
$$\dot{\underline{x}} = f(\underline{x}) = A \underline{x} + B \quad (4.2.3)$$

A und B sind die zu bestimmenden Modellparameter.

Das Dgl.system lautet ausgeschrieben in den einzelnen Komponenten:

$$\begin{aligned} \dot{x}_1 &= A_{11} x_1 + A_{12} x_2 \dots + A_{1n} x_n + B_1 \\ &\vdots \\ \dot{x}_n &= A_{n1} x_1 + A_{n2} x_2 \dots + A_{nn} x_n + B_n \end{aligned}$$

Zur Lösung des Gl.systems nach den Parametern A und B sind  $n \cdot (n+1)$  Gleichungen nötig. In jeder Matrixzeile werden die Koeffizienten  $A_{i_1} \dots A_{i_n}$  und  $B_i$  durch ein Linear-Least-Squares-Fit berechnet. Gegeben sind dazu m Merkmalsvektoren  $x^{(m)}$  mit  $m \geq n+1$ . Nach Berechnung der Differenzvektoren  $\dot{x}^{(m)} = x^{(m+1)} - x^{(m)}$  werden für die j-te Zeile des Systems die Parameter  $A_{1j} \dots A_{nj}, B_j$  berechnet durch Lösen des Gl.systems

$$\begin{pmatrix} x_1^1 & x_n^1 & 1 \\ x_1^2 & x_n^2 & 1 \\ \vdots & \vdots & \vdots \\ x_1^{m-1} & x_n^{m-1} & 1 \\ x_1^m & x_n^m & 1 \end{pmatrix} \cdot \begin{pmatrix} A_{j1} \\ A_{j2} \\ \vdots \\ A_{jn} \\ B_j \end{pmatrix} = \begin{pmatrix} \dot{x}_1^1 \\ \dot{x}_1^2 \\ \vdots \\ \dot{x}_1^{m-1} \\ \dot{x}_1^m \end{pmatrix} \quad (4.2.4)$$

Nach n maligem Anwenden dieses Verfahrens sind die Matrix A und der Vektor B bestimmt.

Stabilitätsaussagen folgen alleine aus der Matrix A.

Zur Prognose werden A und B benötigt.

#### 4.3 Kalmanfilter für zeitdiskrete stochastische Prozesse

Im Gegensatz zu den beiden bisherigen Verfahren -Linear prediction und die Modellierung durch ein Dglsystem- wird in diesem Ansatz der stochastische Aspekt betont. Bevor das Modell auf die spezielle Situation angepaßt wird, wird zuerst allgemein ein solches Filter angegeben und spezifiziert.

Man betrachte den stochastischen Prozeß  $X(t) = \{x(t), t \in T_0\}$   
 Für den zeitlich diskreten Fall mit einer äquidistanten Zeitunterteilung  $T$  betrachten wir  $X\{t_A\} = \{x(t+A \cdot T), k \in \mathbb{N}\} =: \{x_A\}$

Der stochastische Prozeß braucht nicht stationär zu sein. Das Kalmanfilter ist insbesondere eine Erweiterung der Wiener Kolmogoroffschen Theorie auf lineare Systeme instationärer Prozesse bei endlicher Beobachtungszeit.

Das mathematische Modell lautet:

$$\begin{aligned} x(k+1) &= A(k) x(k) + v(k) & (I) \\ y(k) &= C(k) x(k) + w(k) & (II) \end{aligned} \quad (4.3.1)$$

(I) beschreibt die Zustandsgleichung, (II) die Beobachtungsgleichung. Es sei  $x \in \mathbb{R}^n$  der gesuchte unbekannte Zustandsvektor mit Anfangswert  $x(k_0)$ . Der Erwartungswert  $E\{x(k_0)\}$  sei 0 und die Kovarianzmatrix sei  $P(k_0)$   $y \in \mathbb{R}^m$  ist die beobachtete Meßgröße.

Für die Störgröße  $v$  und den Meßfehler  $w$  des Systems gelten:

$$\begin{aligned} E\{v(k)\} &= E\{w(k)\} = 0 \\ E\{v(k) \cdot v^T(l)\} &= Q(k) \delta_{k,l} \\ E\{w(k) \cdot w^T(l)\} &= R(k) \delta_{k,l} \end{aligned} \quad (4.3.2)$$

Weiterhin seien der Anfangszustand  $x(k_0)$ ,  $v$  und  $w$  gegenseitig unkorreliert:

$$E\{x(k_0) v^T\} = E\{x(k_0) w^T\} = E\{v w^T\} = 0 \quad (4.3.3)$$



Aufgabe:

Ausgehend von den Beobachtungsgrößen  $y(0), y(1), \dots, y(k)$  soll der Zustand  $x(k)$  linear, erwartungstreu und mit minimaler Varianz geschätzt werden (beim reinen Filterproblem). Bei der Prognose ist  $x(K)$ ,  $K > k$  zu schätzen.

Als Lösung der oben gestellten Aufgabe ergeben sich folgende rekursive Filtergleichungen ( $\hat{x}$  sind die geschätzten Werte). ( / 3/ )

$$\begin{aligned} \hat{x}^*(k+1) &= A(k) \hat{x}^*(k), & \hat{x}^*(k_0) &= 0 \\ \hat{x}^*(k) &= \hat{x}^*(k) + K(k) \{ y(k) - C(k) \hat{x}^*(k) \} \end{aligned} \quad (4.3.4)$$

Als Fehlerkovarianzmatrizen und Verstärkungsmatrix  $K$  erhält man:

$$\begin{aligned} \hat{P}^*(k+1) &= A(k) \tilde{P}^*(k) A^T(k) + Q(k) \\ K(k) &= \hat{P}^*(k) C^T(k) \{ C(k) \hat{P}^*(k) C^T(k) + R(k) \}^{-1} \\ \tilde{P}^*(k) &= \hat{P}^*(k) - K(k) C(k) \hat{P}^*(k) \end{aligned} \quad (4.3.5)$$

Anfangsbedingung ist  $\hat{P}^*(k_0) = P(k_0)$

Bemerkung:

Dieser Algorithmus bezieht sich auf zentrierte Anfangswerte. Für  $\hat{y} := E\{x(k_0)\} \neq 0$  ist  $x - \hat{y}$  zu filtern und der Schätzwert ist zu  $\hat{y}$  zu addieren.

### Anwendung des Kalmanfilters bei der Schadensdetektion

In den beiden bisherigen Verfahren wurden die berechneten statistischen Kenngrößen (Merkmale) als deterministische Daten zur Modellbeschreibung und zur Prognose benutzt. In diesem Modell wird der stochastische Aspekt der Meß- und Störgrößen berücksichtigt.

Der folgende Abschnitt gliedert sich in 3 Unterabschnitte:

In einem ersten Schritt werden beim herkömmlichen Filterproblem die Situation und Bedeutung der gegebenen und abzuschätzenden Parameter erläutert. Danach wird die spezielle Situation im Falle der Siededetektion und die damit verbundene Problematik angeschnitten. Im dritten Teil wird versucht, zum Teil aus dem deterministischen Bereich stammende Parameter zur Modellbildung des stochastischen Systems zu benutzen, um so fehlende Systemkomponenten zu ersetzen und das Modell an ein Kalmanfilter anzugleichen.

ad 1)

Wir gehen aus von den Modellgleichungen (4.3.1)

$$\begin{aligned}x(k+1) &= A(k) x(k) + v(k) \\y(k) &= C(k) x(k) + w(k)\end{aligned}$$

In diesem Modell wird ein lineares System vorausgesetzt, welches aus dem wahren Zustandsvektor  $x$  zum Zeitpunkt  $k$  einen Zustandsvektor zum Zeitpunkt  $k+1$  erzeugt. Im System, welches durch die Matrix  $A$  beschrieben ist, wirken noch stochastische Störungen  $v$  als weißes Rauschen auf das System und die unbekannte Zustandsgröße  $x$ .

Die gemessene Größe ist  $y$ , eine verauschte Linearkombination von  $x$ . Zudem geht ein Meßfehler  $w$  in Form weißen Rauschens in die Beobachtungsgleichung ein. Wesentlich ist, daß neben den Korrelationen zwischen  $x(k)$ ,  $v$  und  $w$  insbesondere die Matrizen  $A$  und  $C$  apriori gegeben sind, um die wahre Zustandsgröße  $x$  mit Hilfe der Meßwerte  $y$  abzuschätzen.

ad 2)

Im speziellen Fall der Siedegeräusche hat man zunächst ein mit solchen Parametern bestücktes Modell nicht zur Verfügung. Eine Matrix A, welche die lineare Auswirkung des Systems auf den Zustand x beschreibt, fehlt ebenso wie der Parameter C, der die Darstellung des Wertes x über die Meßwerte angibt. Das Reaktorsystem ist von der Komplexität her nicht geeignet, durch solch ein Modell beschrieben zu werden. Man kann aber über die Meßwerte y und andere Verfahren ein Kalmanfilter simulieren wie folgt.

ad 3)

Um eine Möglichkeit der Prädiktion des stochastischen Prozesses zu haben, wird ein Modell konzipiert, welches formal dem Kalmanfilter gleich ist. Dies erlaubt dann, Aussagen über das zukünftige Verhalten des Zustandes machen zu können, wobei zum Teil stochastische Eigenschaften der Funktion berücksichtigt werden können. Das formale Modell bedient sich jedoch aus den in ad 2) angeführten Gründen solcher Parameter A und C, die aus dem deterministischen Modell resultieren. Durch das Dgl.verfahren wird ein System simuliert, welches als spezielles Kalmanfilter angesehen werden kann.

#### Parameteridentifikation:

1. Die Meßgrößen y beschreiben den wahren Zustand x identisch. die Matrix C sei die Identität.
2. Als wirksames System auf den Zustand x bzw.wegen 1 auf y werde die Übergangsmatrix benutzt, die im Dgl.verfahren in linearisierter Form ermittelt wurde.

Aus  $\dot{x} = Ax + B$  erhält man  $x(k+1) = x(k) + A(k)x(k) + B(k) = D(k)x(k) + B(k)$   
mit  $D(k) = I(k) + A(k)$

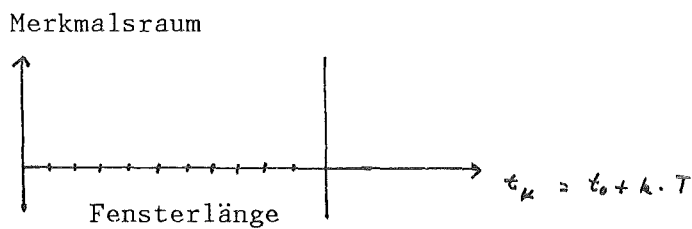
Mit dieser Matrix D wird das Zustandsmodell simuliert. Der konstante Vektor B hat die Bedeutung einer konstanten bekannten Stellgröße.

3. Als Störgrößen im Zustandsmodell werden die Residuals über das gesamte letzte Fenster der Merkmalsvektoren genommen. Es wird die Annahme gemacht: der stochastische Störprozeß sei im Fenster stationär und ergodisch mit  $E\{v(k) v^T(k)\}$  konstant für alle  $k$  im Fenster.

$$E\{v(k) \cdot v^T(k)\}_{ij} = Q_{ij} = \sum_{k=1}^K v_i(k) \cdot v_j(k) / K \quad (4.3.6)$$

4. Die Meßfehler seien nicht angebar  $R(k) = 0$  Matrix.

Zur Erläuterung diene folgende Skizze:



Im Fenster der Länge  $k$  sind zu den  $k$  Intervallen, dargestellt als diskrete Zeiten, die  $k$  statistischen Merkmalsvektoren berechnet. Jeder Merkmalsvektor besteht aus  $n$  statistischen Kenngrößen. Die Parameter des Kalmanfilters werden aus den Daten dieses Fensters berechnet. Die Prognose wird aus den so gewonnenen Parametern für die Merkmalsvektoren außerhalb des Fensters durchgeführt.

#### Parametrisierung des Kalmanfilters

Geht man aus vom Modell

$$\begin{aligned} x(k+1) &= D(k) \cdot x(k) + v(k) \\ y(k) &= C(k) \cdot x(k) + w(k) \end{aligned} \quad (4.3.7)$$

mit den Modellparametern

$$\begin{aligned} D &= A + I, \quad C = I, \quad R(k) = 0 \\ Q &= E\{v \cdot v^T\} \\ P(k_0) &= \hat{P}(k_0) = E\{x(k_0) \cdot x^T(k_0)\} \end{aligned}$$

dann erhält man aus (4.3.4)

$$\hat{x}^*(k+1) = D(k) \hat{x}^*(k) \quad , \quad \hat{x}^*(k_0) = 0 \quad (4.3.8)$$

$$\hat{x}^*(k) = \hat{x}^*(k) + K(k) \{ y(k) - C(k) \hat{x}^*(k) \}$$

mit der Verstärkungsmatrix  $K(k) = \hat{P}^*(k) \cdot \{ \hat{P}^*(k) \}^{-1} = I$

Daraus folgt unmittelbar  $\tilde{P}^*(k) = 0$  und  $\tilde{P}^*(k+1) = Q(k)$

Also lauten die Filtergleichungen

$$\hat{x}^*(k+1) = D(k) \hat{x}^*(k) \quad (1) \quad (4.3.9)$$

$$\hat{x}^*(k) = y(k) \quad (2)$$

Setzt man Gl. 2 in 1 ein, so hat man für die Abschätzung von x zum Zeitpunkt k+1

$$\hat{x}^*(k+1) = D(k) y(k) = D(k) \cdot \hat{x}^*(k)$$

beim zentrierten Anfangswertproblem.

In diesem Falle hat man die gleiche Aussage wie beim Dgl.verfahren. Dies resultiert im wesentlichen aus der Bedingung

$$C = I \quad , \quad R(k) = 0$$

und ist unabhängig von  $Q(k)$  und  $P(k_0)$ .

Man betrachte nun eine Erweiterung des Modells in folgender Form:

Die Rauschkomponenten im Beobachtungssystem seien als maximale relative Fehler gegeben (z.B.  $w(k) = \frac{1}{10}$ ) Dann ist  $R = E \{ w_k w_k^T \} \neq 0$   
C sei weiterhin die Identität.

Dann gelten für die Matrizen folgende Gleichungen

$$K(k) = \hat{P}^*(k) \{ \tilde{P}^*(k) + R(k) \}^{-1}$$

$$\tilde{P}^*(k) = \tilde{P}^*(k) - K(k) \tilde{P}^*(k) \quad (4.3.10)$$

$$\tilde{P}^*(k+1) = A(k) \tilde{P}^*(k) A^T(k) + Q(k)$$

In diesem Modell ist also die Verstärkungsmatrix  $K$  von  $I$  verschieden. Erst mit der Bedingung  $R \neq 0$  geht die Größe  $\hat{P}(k)$  in die Filtergleichungen ein.

Bemerkung:

Mindestens eine der Bedingungen  $R \neq 0$  oder  $C \neq I$  muß erfüllt sein, damit das Kalmanfilter eine echte Erweiterung des deterministischen Modells ist.

Über die Beobachtungsmatrix  $C$  lassen sich vom Reaktorsystem her keine Aussagen machen. Man identifiziert quasi die Zustandsgrößen als Beobachtungsgrößen. Erst  $R \neq 0$  bewirkt eine optimale Filterung gegenüber dem deterministischen Modell. Ist die Kovarianzmatrix  $R = 0$ , so sind beide Modelle identisch. Insbesondere bei der reinen Prognose sind nicht nur die Aussagen, sondern auch die Verfahren gleich, da keine Korrektur durch die Beobachtungsgrößen angesetzt wird. Ist nun  $R \neq 0$ , so liefern die Filtergleichungen eine Korrektur, und durch die Überwachung des Fensters ergibt sich ein anderer Startwert für die reine Prognose. Damit ändern sich auch die Abweichungen für spätere Zeitpunkte.

Für die reine Prognose hat man:  $\hat{x}(K) = \Phi(K, k) \cdot \hat{x}(k)$

mit  $K > k$  und  $\Phi(K, k) = A(k) \cdots \cdots A(k+1) \cdot A(k)$  als Übergangsmatrix

#### 4.4 Das Berliner Verfahren:

Bei dem Berliner Verfahren handelt es sich um einen Regressionsansatz für instationäre stochastische Prozesse. Das Modell des instationären Prozesses bestehe aus einem stationären Anteil, der von einer systematischen additiven Störgröße überlagert ist. Die systematische Komponente werde durch einen KQ (Kleinst Quadrat) Schätzer approximiert. Es werden jeweils die Daten aus einem gleitenden Stützbereich zur Berechnung verwendet. Eine weitere Möglichkeit wäre die exponentielle Gewichtung der Daten bei einer langsamen Veränderung des Modells. Dies führt dann zum sogenannten Exponential smoothing. Jedoch ist der Glättungsparameter eine unbekannte Größe.

Es wird hier das Berliner Verfahren angewendet.

Modellansatz 
$$Y_t = T_t + X_t \quad (4.4.1)$$

Die beobachtete Zeitreihe  $Y_t$  ist additiv zusammengesetzt aus einer Trendkomponente  $T_t$  und einem Rauschanteil  $X_t$ .  $X_t$  sei schwach stationär.

Das begleitende Intervall sei  $I_t$  mit  $m$  Stützstellen.

Für  $t \in \mathbb{N}$  sei  $I_t = \{t+1, \dots, t+m\}$  mit  $t = 0, 1, \dots$

Für die Trendkomponente werde ein  $d$  polynomialer Ansatz gemacht

$$T_t = \sum_{i=0}^d a_i t^i = a_0 + a_1 t + \dots + a_d t^d \quad (4.4.2)$$

Hierbei sind die stochastischen Koeffizienten  $a_i$   $i=0, \dots, d$  noch zu bestimmen.

Mit

$$F_t = \begin{pmatrix} 1 & t+1 & \dots & (t+1)^d \\ \vdots & & & \\ 1 & t+m & \dots & (t+m)^d \end{pmatrix} \quad (4.4.3)$$

ist der KQ Schätzer zu berechnen für

$$Y_t = F_t \cdot \underline{a} \quad \text{mit} \quad \underline{a} = (a_0, a_1, \dots, a_d)^T$$

Im Falle  $d = 1$  hat man

$$F_t = \begin{pmatrix} 1 & t+1 \\ \vdots & \\ 1 & t+m \end{pmatrix} \Rightarrow \begin{matrix} Y_1 = a_0 + a_1 \\ \vdots \\ Y_m = a_0 + m \cdot a_1 \end{matrix}$$

Bemerkung 1:

Für ein festes  $t$  ist  $I_t$  mit den  $m$  Stützstellen fest vorgegeben und für alle  $I_t$  konstant., da es sich gemäß den Stützstellen  $1, \dots, m$  jeweils auf den Anfangsstützpunkt von  $I_t$  bezieht.  $I_t$  ist zeitinvariant.

Der beste Schätzer für  $\underline{a}$  ist nun 
$$\hat{\underline{a}} = (F^T \cdot F)^{-1} F^T Y \quad (4.4.4)$$

Die Dimension des Polynomansatzes geht in  $F$  ein. Er ist zunächst frei wählbar. Die so gewonnene Regressorkoeffizienten können als Sekundärmerkmale in die Klassifikation eingehen.

Bemerkung 2:

Der Regressionsansatz beim Berliner Verfahren kann formal wie ein Kalmanfilter behandelt werden, wenn die Matrizen entsprechend gesetzt werden. In der Beobachtungsgleichung erhält man für  $x = b$  und die Matrix  $C = F_t$  ein Modell für die Regressionsparameter  $b$ . Die Zustandsgleichung mit  $b$  als  $x$  modelliert dann die stochastischen Schwankungen von  $b$ .

Für  $D = I$  ( $b = b_t$ ) erhält man das klassische Regressionsmodell ( /3/ S 277)

Im Programm ist das Berliner Verfahren implementiert mit  $d = 2$ ; eine Erweiterung ist jedoch ohne weiteres möglich.



#### 4.5 ARIMA-Modellierung im Merkmalsraum:

Ein weiteres und sehr mächtiges Verfahren bietet die Beschreibung des stochastischen Prozesses durch ein autoregressives integriertes moving average (ARIMA) Modell, in dem auch Instationaritäten erfaßt werden. In diesem Kapitel werden zuerst AR-, MA- und ARMA-Modelle vorgestellt. Anschließend werden die Probleme der Stationarität und Instationarität erläutert und der Übergang zum ARIMA-Modell vollzogen. Eigenschaften der Modellparameter und Kriterien zur Modellidentifikation werden angesprochen. In einem Diagramm wird schließlich der Programmablauf logisch festgelegt. Die Anwendung auf die Testsignale und Ergebnisse finden sich im Kapitel 5 wieder.

Im folgenden seien mit  $t, t-1, \dots$  die diskreten Zeitintervalle bezeichnet.  $x_t, x_{t-1}, \dots$  sind die dazugehörigen Funktionswerte des Prozesses, in unserem Falle die eindimensionalen Komponenten der Merkmalsvektoren im Merkmalsraum der statistischen Kenngrößen.  $\tilde{x}_t = x_t - \mu$  seien die zentrierten Werte. Mit  $\varepsilon_t$  sei stets weißes Rauschen bezeichnet.

#### Autoregressives Modell der Ordnung p: AR(p)

Def 1: Ein AR(p) Prozeß wird dargestellt

$$\tilde{x}_t = \sum_{i=1}^p \phi_i \tilde{x}_{t-i} + \varepsilon_t \quad (4.5.1)$$

mit  $\varepsilon_t$  weißes Rauschen mit  $E(\varepsilon_t) = 0$  und  $\sigma^2 > 0$

Durch Einführung des Operators  $B := B^m \tilde{x}_t = \tilde{x}_{t-m}$

und  $\phi(B) = 1 - \phi_1 B - \dots - \phi_p B^p$  erhält man  $\phi(B) \tilde{x}_t = \varepsilon_t$ .

In diesem AR Modell der Ordnung p sind p+2 Unbekannte.  $(\phi_1, \dots, \phi_p, \sigma^2, \mu)$  Dieser unbekannte Parametersatz ist aus der Vergangenheit zu schätzen.

Moving average Modell der Ordnung q MA(q)

Def 2: Ein MA(q) Prozeß hat die Darstellung

$$\tilde{X}_t = \varepsilon_t + \sum_{j=1}^q \theta_j \varepsilon_{t-j} \quad (4.5.2)$$

mit  $\varepsilon_t$  weißes Rauschen

Mit der Operatordarstellung hat man

$$\tilde{X}_t = \Theta(B) \varepsilon_t, \quad \Theta(B) = 1 + \theta_1 B + \dots + \theta_q B^q$$

Autoregressives moving average Modell der Ordnung p,q ARMA(p,q)

Def 3: Ein ARMA(p,q) Prozeß hat die Darstellung

$$\tilde{X}_t = \sum_{i=1}^p \phi_i \tilde{X}_{t-i} + \varepsilon_t + \sum_{j=1}^q \theta_j \varepsilon_{t-j} \quad \Leftrightarrow$$

$$\Phi(B) \tilde{X}_t = \Theta(B) \varepsilon_t \quad (4.5.3)$$

In diesem Modell müssen die p+q+2 Parameter  $\phi_1 \dots \phi_p, \theta_1 \dots \theta_q, \mu, \sigma^2$  aus dem gegebenen Datenmaterial geschätzt werden.

Bemerkung:

Ein MA Prozeß der endlichen Ordnung q ist stets stationär mit der Autokorrelationsfunktion (im folgenden mit ACF bezeichnet)

$$\rho(h) = \begin{cases} \frac{\sum_{j=0}^{q-h} \theta_j \theta_{j+h}}{\sum_{j=0}^q \theta_j^2} & 0 \leq h \leq q \\ 0 & h > q \end{cases} \quad (4.5.4)$$

Die Stationarität im Sinne von schwach stationär folgt unmittelbar aus der Definition, da  $\varepsilon_t$  weißes Rauschen ist. Die ACF ist 0 für  $h > q$ .

Parameterschätzung aus den Yule Walker Gleichungen

Ausgangsmodell ist 
$$X_t = \sum_{j=1}^p \phi_j X_{t-j} + \varepsilon_t$$

Durch Multiplikation mit  $X_{t-h}$  und Erwartungswertbildung erhält man

$$\begin{aligned} \gamma(h) &= \sum_{j=1}^p \phi_j \gamma(h-j); \quad h > 0 \\ \gamma(0) &= \sum_{j=1}^p \phi_j \gamma(j) + \sigma^2; \quad h = 0 \end{aligned}$$

- a) Aus  $\gamma$  bzw.  $\rho$  lassen sich die  $\phi_j$  berechnen.
- b) Aus  $\phi_j$  kann man  $\gamma$  bzw.  $\rho$  berechnen gemäß (\*) (4.5.5)

$$\begin{pmatrix} 1 & \rho(1) & \dots & \rho(p-1) \\ \rho(1) & 1 & & \rho(p-2) \\ \vdots & & & \vdots \\ \rho(p-1) & \rho(p-2) & \dots & 1 \end{pmatrix} \cdot \begin{pmatrix} \phi_1 \\ \phi_2 \\ \vdots \\ \phi_p \end{pmatrix} = \begin{pmatrix} \rho(1) \\ \rho(2) \\ \vdots \\ \rho(p) \end{pmatrix} \Leftrightarrow R_p \cdot \Phi_p = \rho_p$$

Allgemeine Bedeutung der ARMA(p,q) Prozesse

Der AR(p) Prozeß modelliert eine Zeitreihe, in der ein momentaner Wert funktional linear aus p früheren Zeitreihenwerten, mit gewissen Faktoren  $\phi_i$  gewichtet, hervorgeht.

Der MA(q) Prozeß (Prozeß der gleitenden Durchschnitte) findet oft Verwendung, wenn ein im Gleichgewicht befindlicher Zustand von additiven Störgrößen überlagert ist, die als weißes Rauschen mit der Zeit abklingen.

Ein ARMA Prozeß beinhaltet beide Modellierungen: wenn z.B. bei einem AR Prozeß die Störung selbst ein MA Prozeß ist.

Da ein MA Prozeß stets stationär (schwach stationär) ist, richtet sich die Frage nach der Stationarität eines ARMA Prozesses nur nach dem AR-Anteil. Um Instationaritäten zu beseitigen, werden ARIMA Modelle eingeführt.

Autoregr. integrierter moving average Prozeß der Ordnung (p,d,q)  
ARIMA(p,d,q)

Def 4 Ein stochastischer Prozeß  $Y_t$  heißt ARIMA(p,d,q) Prozeß,

wenn gilt:  $\Phi(B) \nabla^d Y_t = \Theta(B) \varepsilon_t$  ; (4.5.6)

Hierbei :  $\nabla^1 Y_t = Y_t - Y_{t-1}$   
 $\nabla^2 Y_t = (Y_t - Y_{t-1}) - (Y_{t-1} - Y_{t-2})$   
 $\vdots$

Es gilt folgende Äquivalenz:

$Y_t$  ist ein ARIMA(p,d,q) Prozeß  $\Leftrightarrow X_t = \nabla^d Y_t$  ist ein ARMA(p,q) Prozeß

Durch den Operator  $\nabla^d$  wird ein polynomialer Trend der Ordnung d-1 ausgeglichen um Stationarität zu erreichen. In der Praxis genügt meist  $d < 3$ .

Die Bestimmung der Modellparameter geschieht im wesentlichen über die Berechnung der ACF und der partiellen Autokorrelationsfunktion (PACF).

Bemerkung zur ACF und PACF

Als Schätzfunktion für die ACF dient  $\hat{\rho}(h)$  mit  $\hat{\rho}(h) = \frac{\hat{\gamma}(h)}{\hat{\gamma}(0)}$  und  $\hat{\gamma}(h)$  wie

folgt:  $\hat{\gamma}(h) = \frac{1}{n} \sum_{t=1}^{n-h} X_t X_{t+h}$

Da bei sehr großem h die Schätzung ungenau wird, sollte h  $n/3$  nicht überschreiten.

Das Yule Walker GL.system

$$R_h \cdot \Phi_h = \rho_h \Leftrightarrow \begin{pmatrix} 1 & \rho(1) & \dots & \rho(h-1) \\ \rho(1) & 1 & & \rho(h-2) \\ \vdots & & & \vdots \\ \rho(h-1) & \dots & & 1 \end{pmatrix} \cdot \begin{pmatrix} \Phi_{h1} \\ \vdots \\ \Phi_{hh} \end{pmatrix} = \begin{pmatrix} \rho(h) \\ \vdots \\ \rho(1) \end{pmatrix} \quad (4.5.7)$$

hängt vom Parameter  $h$  ab und die Lösung des Systems liefert eine Folge  $\phi_{11}, \dots, \phi_{hh}$  ( $\phi_{hh}$ ) heißt PACF. Der Koeffizient  $\phi_{hh}$  gibt die Korrelation zwischen  $x_t$  und  $x_{t-h}$  nach Bereinigung des linearen Zusammenhanges in den anderen  $x_{t-i}$  ( $i < h$ ) an. ( /3/ S 237)

Kriterien zur Modellidentifizierung

Ein Vergleich von (4.5.7)  $R_h \cdot \phi_h = \rho_h$

mit (4.5.5)  $R_p \cdot \phi = \rho_p$  zeigt, daß für einen AR(p) Prozeß gilt:

$$\phi_p = \phi_{pp}$$

Außerdem ist für  $h > p$  ein AR(p) Prozeß ein AR(h) Prozeß mit  $\phi_{p+h} = \dots = \phi_h = 0$ .

⇒ Ein AR(p) Prozeß ist charakterisiert durch Abbrechen der PACF.

$$\phi_{hh} = 0 \quad (h > p)$$

Bricht die PACF nicht ab, so handelt es sich um einen MA bzw einen ARMA Prozeß, da ein MA Anteil vorhanden ist.

Es gilt: (/3/ S 239)

- a)  $X \sim \text{AR}(p)$  :  $\rho(h)$  ist abklingend und  $\phi_{hh} = 0$  für  $h > p$
- b)  $X \sim \text{MA}(q)$  :  $\phi_{hh}$  ist abklingend und  $\rho(h) = 0$  für  $h > q$
- c)  $X \sim \text{ARMA}(p,q)$  : Es gilt a) und b)

Anhand von  $\rho(h)$  und  $\phi_{hh}$  wird auf das entsprechende Modell geschlossen.

### Kriterien zur Stationarität, Instationarität

Die Frage nach der Stationarität des Prozesses kann alleine mit dem AR Anteil entschieden werden.

Betrachten wir das Modell  $\phi(B) X_t = \theta(B) \varepsilon_t$   
mit  $\varepsilon_t$  weißes Rauschen und  $X_t$  sei zentriert.

$$\phi(B) = 1 - \phi_1 B - \dots - \phi_p B^p$$

$$\theta(B) = 1 + \theta_1 B + \dots + \theta_q B^q$$

Stationaritätsbedingung für ARMA Prozesse:

$X_t$  ist stationär  $\Leftrightarrow$  Alle Nullstellen des Polynoms  $\phi(z)$  liegen außerhalb des Einheitskreises (Insbesondere ist bei AR(1) Prozeß  $|\phi_1| < 1$ )  $\Leftrightarrow X_t$  ist lineares Filter und kann geschrieben werden

$$X_t = \sum_{i=0}^{\infty} \psi_i \varepsilon_{t-i} = \psi(B) \varepsilon_t, \quad \sum_{i=0}^{\infty} \psi_i^2 < \infty$$

Koeffizientenberechnung:

a) Für einen reinen AR Prozeß  $\phi(B) X_t = \varepsilon_t$  ist somit  $\phi(B) \psi(B) \varepsilon_t = \varepsilon_t$   
also  $\phi(B) \cdot \psi(B) = 1 \Rightarrow$  Berechnung der  $\psi_i$ .

b) Für einen ARMA Prozeß werden die  $\psi_i$  berechnet durch Koeffizientenvergleich  $\phi(B) \cdot \psi(B) = \theta(B)$

In beiden Fällen sind die  $\phi_i$  und die  $\theta_i$  gegeben.

Eine für die Praxis wichtige Aussage ist folgende:

### Invertierbarkeitsbedingung

Sei  $X_t$  stationär und alle Nullstellen von  $\theta$  liegen außerhalb des Einheitskreises, dann ist  $X_t$  darstellbar als AR( $\infty$ ) Prozeß (/3/ S 242)

Diese Darstellung erlaubt es, im Merkmalsraum über die ACF und PACF signifikant von 0 verschiedene Werte zur Modellbildung zu verwenden.

Programmstruktur zur ARIMA Modellierung mit Parametern (p,d,q)

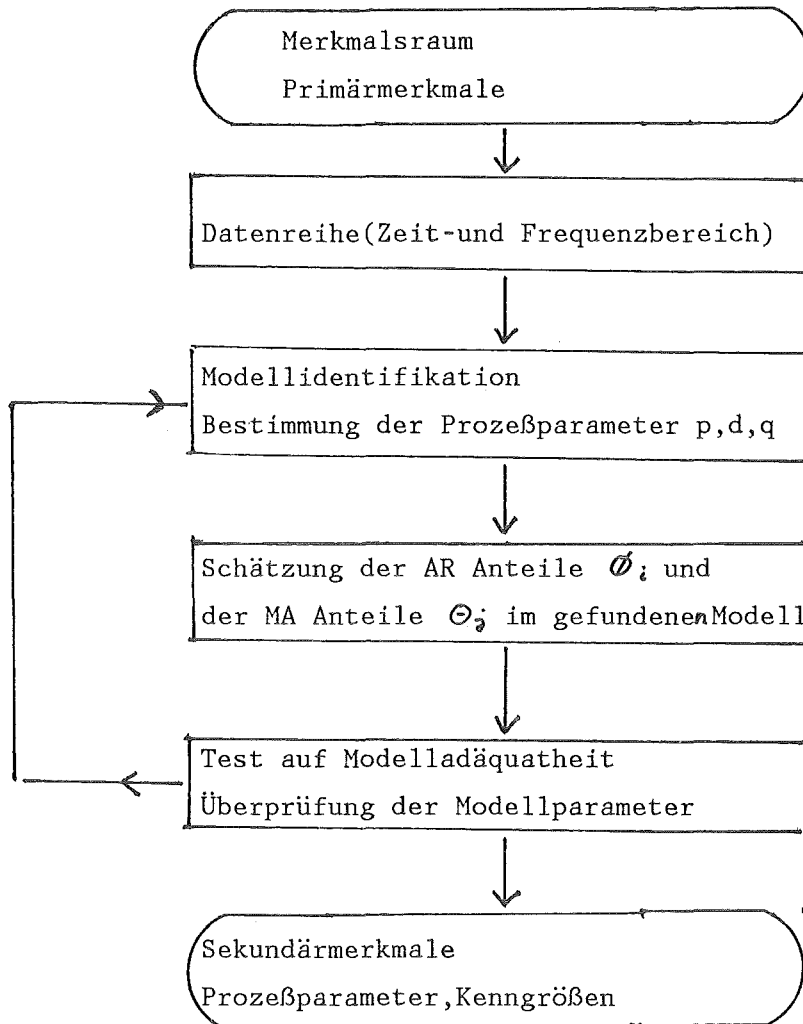


Abb.2 ARIMA-Modellierung

Ausgehend von den Merkmalen (Primärmerkmale), die die Zeitreihe darstellen, werden Modell- und Prozeßparameter konstruiert und geschätzt, welche die Struktur des Signals wiedergeben. (Sekundärmerkmale). Diese sollen später als Zusatzinformation in den Klassifikationsprozeß einfließen. Die Größen  $p, q, d$  werden iterativ bestimmt.

Nachfolgendes Diagramm zeigt die Iterationsschleifen zur Dimensionierung

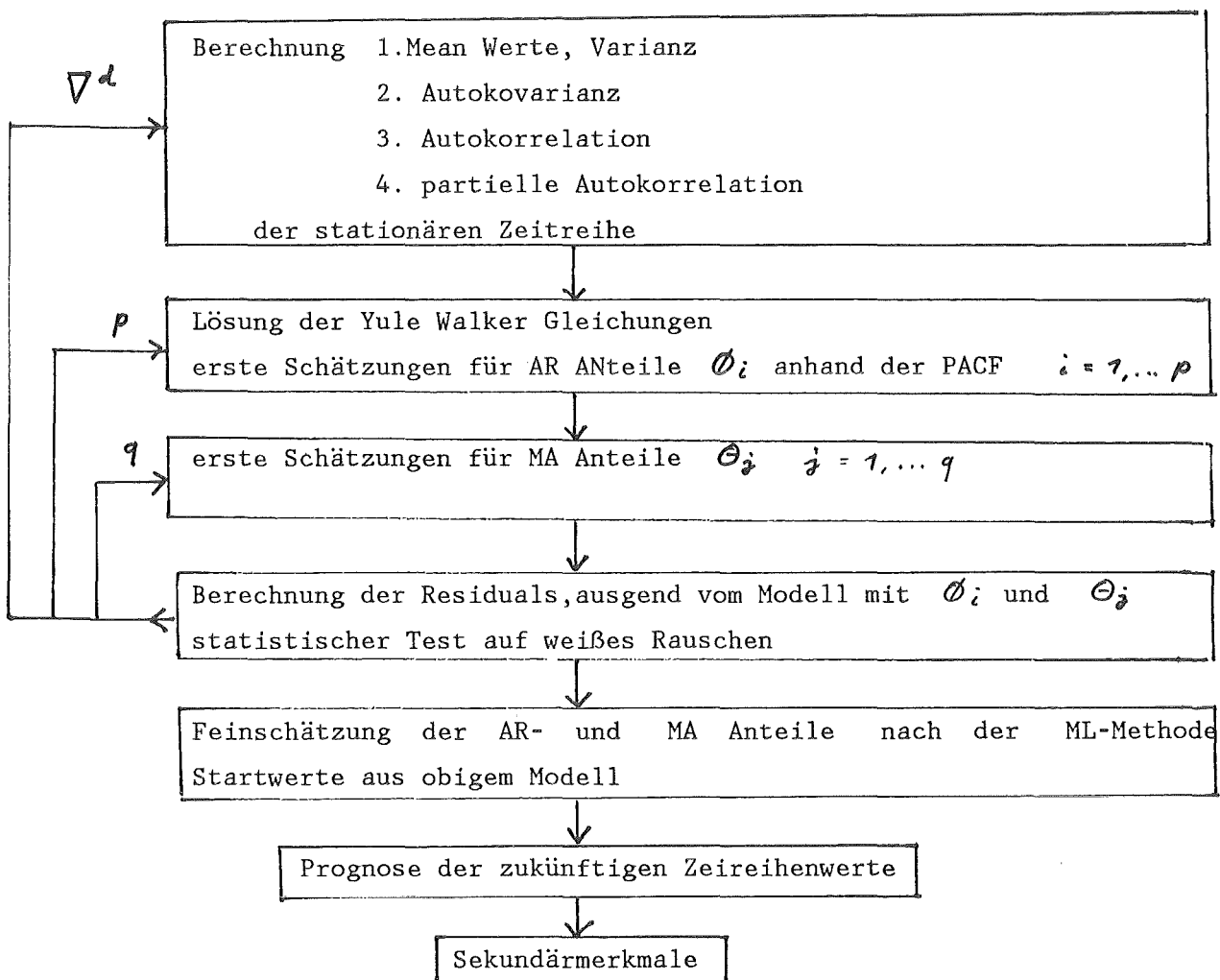


Abb. 3 : Iterative Dimensionierung des ARIMA-Modells



## 5. Auswertung und Ergebnisse:

Für ein Signal aus der Reihe gelte die folgende Auswertungsfolge:

### 1. Schritt: Rohdatenspeicherung

Das gesamte Signal (bzw. der interessierende Teil) wird digitalisiert unter einer Versuchsnummer NR abgespeichert.

### 2. Schritt: Rohdatenplott

Das Signal der Rohdaten wird pro Block mit 16 Werten ausgeplottet. Ein Block besteht aus 512 Abtastwerten, die zu je 32 zusammengefaßt sind. ( $16 * 32 = 512$ )

Programm `sieden.fort(plotroh)`

Abb. 6,7

### 3. Schritt: FFT

Die Rohdaten werden vom File gelesen und es wird blockweise eine Fouriertransformation in den Frequenzbereich gemacht. Pro Block werden die FFT Koeffizienten gemäß ihrer Fensterfunktion berechnet und in einem File abgelegt.

Programm `Sieden.fort(gen)`

### 4. Schritt: Ausplotten des geglätteten Powerspektrums

Es werden die in Schritt 3 berechneten Koeffizienten gelesen, geglättet und in einem 3-D Plot der Größe 50x50 geplottet.

Programm `Sieden.fort(plot3d)`

Abb. 8

### 5. Schritt: Merkmalsgenerierung

Es werden im Zeitbereich die Rohdaten aus ihrem File und im Frequenzbereich die FFT Koeffizienten gelesen. Es geschieht eine Merkmalsberechnung in beiden Räumen gemäß der Programmparameter. Die Merkmale werden dann zusammen mit den prognostizierten auf ein File geschrieben und stehen dort zum Aufruf des Plotprogrammes bereit.

Programm `Sieden.fort(merk)`

## 6. Schritt: Plot der Merkmale

Vom File werden die Merkmale gelesen und geplottet. Die Anzahl ist aus plotechnischen Gründen begrenzt.

Programm Sieden.fort(plotmerk)

Abb. 9-14

## 7. Schritt: Klassifikation

Mit den Merkmalen wird der Klassifikator trainiert. Anschließend erfolgt die Klassifikation sowohl der prognostizierten Primärmerkmale als auch der Sekundärmerkmale.

Programm Sieden.fort(redklas) (noch in Entwicklung)

## Auswertung der Ergebnisse (Abb. 6 - Abb. 14)

Abb. 6 zeigt das aufgenommene RMS Signal.

Die Abtastrate beträgt 100 kHz, das gesamte Signal enthält 500 logische Blöcke zu je 512 Abtastwerten. Ein Block wird in 16 Einheiten dargestellt (Skalierung). Abb. 6 zeigt den Ausschnitt Block 100 - 200. Die kritische Phase beginnt etwa bei Block 160 (  $16 \cdot 60 = 960$  ). Abb. 7 zeigt die Fortsetzung des Signals: Block 200 - 300.

Abb. 8 zeigt das Autoleistungsdichtespektrum in geglätteter Form, einmal mit Rechteckfenster, zum andern mit Hanningfenster.

In Abb. 9 - 14 werden die Resultate der verschiedenen Prognoseverfahren im Merkmalsraum dargestellt. Als X-Achse wird die Zeit in Anzahl von Blöcken (ein Block umfaßt 512 RMS Werte) dargestellt. Die Ordinate gibt die skalare Größe des Merkmals an.

Zuerst wird ein Datenfenster von FLEN Merkmalsvektoren als Ausgangsbasis zugrunde gelegt. Dann werden das Fenster um VERSCH Einheiten verschoben, jeweils für die einzelnen Prognosemodelle die Parameter geschätzt und dann weitere VERSCH Merkmale prognostiziert.

Die verschiedenen Prognosewerte sind gemäß den einzelnen Verfahren eingetragen.

Die Prognoseverfahren beziehen sich hier auf zwei Merkmale (p7 und p17 ). Es wird hier nur der Ausschnitt des Siedebeginns dargestellt.

Abb. 9 zeigt die Auswertung mit 4 Prognosemodellen für das Merkmal p7.  
Aus Übersichtlichkeitsgründen sind pro Schaubild nur zwei Verfahren eingezeichnet. Für die beiden anderen Verfahren wurde die Null-Linie mit der eingetragenen Merkmalskurve versetzt  
Die Parameter sind: Fensterlänge FLEN = 60

Prognoselänge VERSCH = 5

Abb.10 zeigt das gleiche für das Merkmal p17.

Abb.11 zeigt die Anwendung zweier Verfahren für das Merkmal p7.

Die Parameter sind: FLEN = 50

VERSCH = 5

Abb.12 zeigt das gleiche für das Merkmal p17.

Abb.13 zeigt die Anwendung zweier Verfahren für das Merkmal p7.

Die Parameter sind: FLEN = 30

VERSCH = 2

Abb.14 zeigt das gleiche für das Merkmal p17.

Es gibt Fälle, wo das ARIMA Modell versagt. (Block 75, 140, 345)  
Dies hängt vom vorgegebenen Datenmaterial ab, dh. von den Stützpunkten im Merkmalsraum. Über einen Vergleich der Verfahren untereinander sollte in einem Entscheidungsprozeß das Modell akzeptiert werden oder nicht. Im Falle der Akzeptanz können dann die Modellparameter als Sekundärmerkmale verwendet werden. Die Art der Merkmale spielt hier mit eine Rolle, da die Auswahl der Merkmale die zu modellierende Zeitreihe liefert.

## 6. Zusammenfassung und Ausblick

In dieser Arbeit wird die Schadensfrüherkennung unter dem Aspekt des Prognostizierens betrachtet. Als Beispiel einer Störung dient ein Experiment, in dem lokales Sieden simuliert wird. Da es Prozesse gibt, die sehr schnell propagieren, ist es wichtig, Fehlverhalten und erste Anzeichen schon früh richtig zu interpretieren, sodaß eine entsprechende Maßnahme getroffen werden kann. Es wird aufgezeigt, wie dieser Interpretations- und Entscheidungsprozeß durch Anwendung von Prognoseverfahren und Trendanalyse im Merkmalsraum durchgeführt werden kann.

Betrachtet man ein herkömmliches Überwachungssystem, so läßt sich der Informationsfluß wie folgt darstellen.

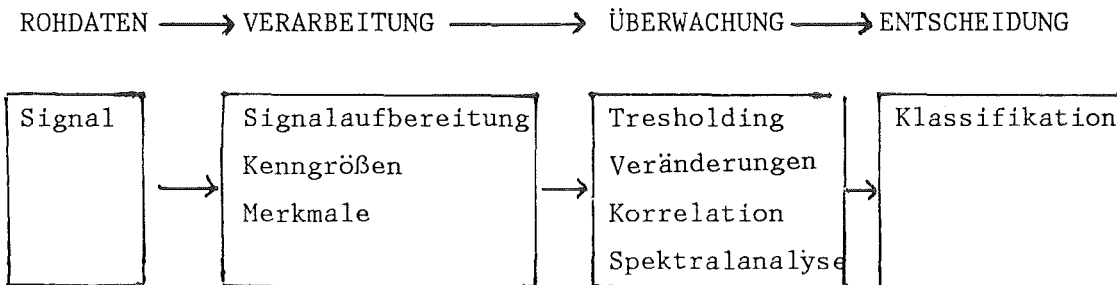


Abb. 4: Überwachungskonzept

Die Zeit, die vom Auftreten des Schadens bis zu seiner Erkennung vergeht, hängt von mehreren Randbedingungen ab, z.B. von der Aufbereitung und Art der Merkmale. Auch der Klassifikationsprozeß ist nicht trivial und zeitintensiv. Eine a priori Ausnutzung der im Voraus abgeschätzten Information kann man mit Hilfe von Prognoseverfahren und Trendanalyse erreichen.

Nach einer Berechnung der für den jeweiligen Anwendungsfall relevanten Kenngrößen (Auf das Problem der Merkmalsextraktion wird hier nicht eingegangen) werden im Merkmalsraum mathematische Modelle erstellt, die den beobachteten Prozeß beschreiben. Es werden nicht nur die Kenngrößen überwacht, sondern das Modell, repräsentiert durch seine Parameter, ist charakteristisch für den Prozeßzustand. Man hat eine parametrische Prozeß bzw. Signalbeschreibung. Diese Prozeßparameter haben direkten Einfluß auf das zukünftige Verhalten. In diesem Sinne kann, ausgehend vom parametrischen Modell eine Prognose des zeitlichen Verhaltens der Kenngrößen angegeben werden. Die Prognose vollzieht sich im Merkmalsraum, nicht auf Signalebene. Diese Tatsache ermöglicht die Ausdehnung des Klassifikationsprozesses auf die prognostizierten Merkmale. Es kann also bereits eine Entscheidungsfindung vorweggenommen werden, und die aus dem überwachten Signal berechneten Merkmale gehen nachträglich nur noch in eine schnelle ja-nein Entscheidung ein.

In einer ersten Ausbaustufe werden die Verfahren vorgestellt und in ihren Leistungen am Experiment des lokalen Siedens verglichen. Zur Klassifikation werden zwei Typen von Merkmalen verwendet, zum einen prognostizierte Merkmale- dies sind die gemäß dem mathematischen Modell vorausberechneten statistischen Kenngrößen-; zum andern die Modellparameter selbst als Sekundärmerkmale.

#### Vorstellung der einzelnen Verfahren:

Alle im folgenden vorgestellten Methoden operieren im Merkmalsraum. Die Merkmalsextraktionsstufe ist abgeschlossen und es steht der für die Klassifikation relevante dimensionsreduzierte Merkmalsraum im Zeit- und Frequenzbereich zur Verfügung.

1. Beim Linear Prediction Verfahren wird mit dem Ansatz

$$f_n = - \sum_{\nu=1}^M \alpha_\nu f_{n-\nu}$$

die momentane Merkmalsvektorkomponente  $f_{x_i}$  aus den M vorangehenden mit den linearen Gewichten  $\alpha_{\nu}$  berechnet. In einer Einstellungsphase werden die  $\alpha_{\nu}$  durch ein Optimierungskriterium ermittelt: Der erwartete Fehler für die Summe der quadratischen Abweichungen sei minimal. Diese so berechneten  $\alpha_{\nu}$  beschreiben dann das Modell. Ausgehend von ihnen werden die Prognosewerte berechnet. (M ist zunächst frei wählbar, beeinflusst aber das Ergebnis)

2. Ein weiteres Prognoseverfahren wird über die Modellbildung einer Dgl. erzielt. Ausgehend von einem Satz Beobachtungswerten im n-dimensionalen Merkmalsraum wird ein n-dimensionales gewöhnliches Dgl.system erster Ordnung angesetzt. Unter der Annahme, daß in einem lokalen Zeitraum keine exogenen Größen ins System einwirken und daß lokal das System autonom ist, ergibt eine Taylorentwicklung der Dgl. um den Punkt  $x = 0$  die linearisierte Form

$$\dot{x} = Ax + B, \quad A = \left( \frac{\partial f_i}{\partial x_j} (x=0) \right)_{ij}, \quad B = f(x=0)$$

mit noch unbekanntem Modellparametern A und B.

Durch Berechnung der Merkmalsvektoren und ihrer Differenzen in einem Beobachtungsintervall werden A und B in einem Linear-Least-Squares-Fit Verfahren geschätzt. Bei einer ständigen Überwachung werden durch eine fortlaufende Verschiebung des Beobachtungsfensters im Merkmalsraum die Parameter A und B stets angepaßt. Somit beziehen sich die aus dem parametrischen Modell prognostizierten Werte immer auf den letzten aktuellen Prozeßzustand.

Dieses Verfahren bietet neben der Möglichkeit zur Prognose zukünftiger Merkmale über das parametrische Modell mit den Parametern A und B noch folgende wichtige Eigenschaften:

- Das Stabilitätsverhalten der Lösungen der aufgestellten Dgl. ist allein abhängig von der Matrix A der linearisierten Form. Die Dimensionen der komplementären Räume, die von Eigenvektoren, zugehörig zu Eigenwerten mit positivem und negativem Realteil, aufgespannt werden, bzw. die Eigenwerte selbst, geben Zusatzinformationen über die Stabilität des Systems. Bei einer "Explosion" der Merkmale wird das System äußerst instabil. Bei der Überwachung des Systems ist ein Siedebeginn oder überhaupt eine kriische Situation mit einer Instabilität im Merkmalsraum verknüpft. Durch eine Beurteilung dieser Modellparameter zusätzlich zur Prognoseklassifikation erhält man einen gestaffelten Klassifikationsprozeß. Die Prognosewerte werden gemäß dem Modellzuwachs  $\dot{x} = Ax + B$  berechnet.

- Weiterhin hat man mit der linearisierten Form der Dgl. die Möglichkeit ein Kalmanfilter zu simulieren.

Obwohl vom eigentlichen physikalischen Vorgang kein mathematisches Modell erstellt werden kann, gelingt es mit dieser Methode über Erstellung eines deterministischen Dgl.systems ein Kalmanfilter einzusetzen, das bei bestimmten Nebenbedingungen ( C, R ) als Prognoseverfahren mit dem Dgl.verfahren identisch ist. Mit Berücksichtigung und Kenntnis der Störungen wird das Kalmanfilter dann zu einer echten Erweiterung und das Dgl.verfahren kann als Spezialfall darin eingebettet werden.

Die Zustandgleichung  $x(k+1) = D(k) x(k) + v(k)$

und die Beobachtungsgleichung  $y(k) = C(k) x(k) + w(k)$

können bei entsprechender Parametrisierung von A, D, v und w das DGL.verfahren beschreiben.

- Ein dritter Ansatz zur Prognose von Merkmalsvektoren wird im Berliner Verfahren modelliert.

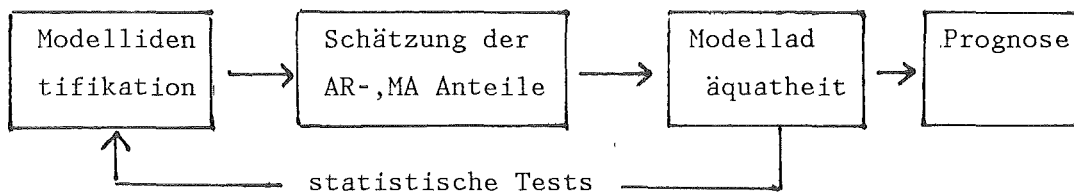
Mit einer Rauschkomponente  $X_t$  und einer Trendkomponente  $T_t$  des stochastischen Prozesses wird durch den Ansatz  $Y_t = X_t + T_t$  (mit  $T_t = \sum_{i=0}^d a_i t^i$ ) ein Polynom vom Grad  $d$  angesetzt, um den Merkmalsverlauf zu approximieren. Die Modellparameter  $a_i$  werden nach der Methode der kleinsten Quadrate geschätzt.

Einen Schätzer für  $a$  in  $Y_t = F_t * a$  erhält man mit  $\hat{a} = (F^T F)^{-1} F^T Y$

- Ein mächtiges Werkzeug zur Modellierung stochastischer Prozesse

bieten die ARIMA Modelle. 
$$x_t = \sum_{i=1}^p \phi_i x_{t-i} + \sum_{j=0}^q \theta_j \varepsilon_{t-j}$$

Nach Elimination der Instationaritäten mit einem Differenzenoperator  $\nabla^d$  werden die autoregressiven Anteile (AR-Anteile) und die Moving average Anteile (MA-Anteile) berechnet. Als Beobachtungsgrößen dienen die berechneten Merkmale. Die Modellbildung wird nach folgendem Schema vollzogen:



Folgende Tabelle gibt eine Zusammenfassung der Verfahren, der den Prozeß im Merkmalsraum darstellenden Rechenmodelle und der für eine zweite Klassifikationsebene relevanten Sekundärmerkmale.



Verfahren	Rechenmodell	Sekundärmerkmale
Lin. pred.	$f_n = - \sum_{v=1}^M \alpha_v f_{n-v}$	lineare Gewichte $\alpha_v$
Dgl	$\dot{x} = Ax + B; \quad A = \frac{\partial f_i}{\partial x_j}$	Matrix A
Berliner	$Y_t = X_t + T_t = \sum_{i=1}^d a_i t^i$	Polynomkoeff. $a_i$
ARIMA	$x_t = \sum_{i=1}^p \phi_i x_{t-i} + \sum_{j=0}^q \theta_j \epsilon_{t-j}$	AR-,MA Anteile $\phi_i, \theta_j$ p,q

Abb. 5: Tabelle der angewandten Prognoseverfahren

Alle vier Verfahren werden zunächst in ihrer Leistungsfähigkeit untersucht. Dabei sind im wesentlichen zwei Aspekte zu berücksichtigen.

1. die Güte der Modellanpassung an die gegebenen Daten und somit eine gute Prozeßdarstellung
2. die Güte der aus dem Modell abgeleiteten Prognosen

Die Verfahren basieren auf unterschiedlichen Ansätzen und bieten bei Paralleleinsatz sichere Entscheidungen. Im Sinne der Schadensfrüherkennung ist vor allem Punkt 2 von Bedeutung.

Bei einer "explosiven" Propagation des Störfalles wird im Merkmalsraum die Zustandsdarstellung über das DGL.system den Prozeß sehr gut beschreiben mit dem Ansatz  $\dot{x} = Ax + B$ , also einen exponentiellen Ansatz gegenüber einem linear gewichteten. Für die Darstellung eines lokal linearen Trends sorgt der konstante Vektor B.

Beim ARIMA Modell hat man eine Palette von Möglichkeiten, auch Konfidenzbereiche für die Prognosewerte anzugeben. Es zeigt sich, daß alle Prognoseverfahren in einem gewissen Toleranzbereich ähnliche Werte voraussagen.

Die Prognosewerte sind noch abhängig von folgenden Parametern:

- Fensterlänge FLEN
- Verschiebungslänge = Prognoselänge VERSCH

### Ausblick

In dieser Studie sind verschiedene Methoden gezeigt, wie man ein parametrisches Signalmodell im Merkmalsraum simulieren kann, welches das vorgegebene Datenmaterial in seiner Struktur zum einen beschreibt und auch Aussagen über den künftigen Verlauf im Merkmalsraum macht. Die Merkmale werden prognostiziert und können in einen Klassifikator schon vor Eintreten der Ereignisse eingegeben werden.

Als weitere Wissensbasis über das System können die Modellparameter selbst dienen. Es wäre zu untersuchen, ob und wie man mit einer Clusteranalyse aus diesen Modellparametern gewisse Klassen trennen kann die dann mit zu einer Entscheidungsfindung beitragen. Das Problem dieser sogenannten Sekundärmerkmale ist hier nur im Konzept eingebettet zwecks einer besseren und redundanten Entscheidungsfindung. Auf einen so erzielten gestaffelten Klassifikationsprozeß wird hier nicht eingegangen.

7: Literaturverzeichnis:

- /1/ J.Aberle, Potential and Problems of Acoustic  
R.Barholomay, Boiling Detection  
H.Rohrbacher, 10th Meeting of the Liquid Metal  
K.Schleisiek Boiling Working Group  
Karlsruhe, Oct. 27 - 29, 1982
- /2/ G.E.P. Box Time Series Analysis, forecasting and control  
G.M. Jenkins Holden Day, San Francisco, 1970
- /3/ L.Fahrmeir, Stochastische Prozesse  
H.Kaufmann, Carl Hanser Verlag München, Wien, 1981  
F.Ost
- /4/ R.Lewandowski Prognose- und Informationssysteme  
De Gruyter Verlag, Berlin, New York, 1974
- /5/ W.G.Schneeweiß Zufallsprozesse in dynamischen Systemen  
Springer Verlag, Berlin, Heidelberg, 1974
- /6/ G.Winkler Stochastische Systeme: Analyse und Synthese  
Akademische Verlagsgesellschaft Wiesbaden  
1977
- /7/ H.W.Knobloch Gewöhnliche Differentialgleichungen  
F.Kappel Mathematische Leitfäden  
Teubner Verlag Stuttgart, 1974
- /8/ P.Mertens Prognoserechnung  
Physica Verlag Würzburg, Wie, 1981
- /9/ Brown Smoothing, Forecasting and Prediction of  
Discrete Time Series  
Prentice Hall, 1963

- /10/ K.P. Scherer      unveröffentlichter Bericht,  
Mai 1984
- /11/ D. Schulz        Lineare und nichtlineare Verarbeitung von ampli-  
tudenmodulierten stochastischen Prozessen zur  
akustischen Qualitätskontrolle und Schadenfrüh-  
erkennung  
Dissertation, TU Karlsruhe, 1976

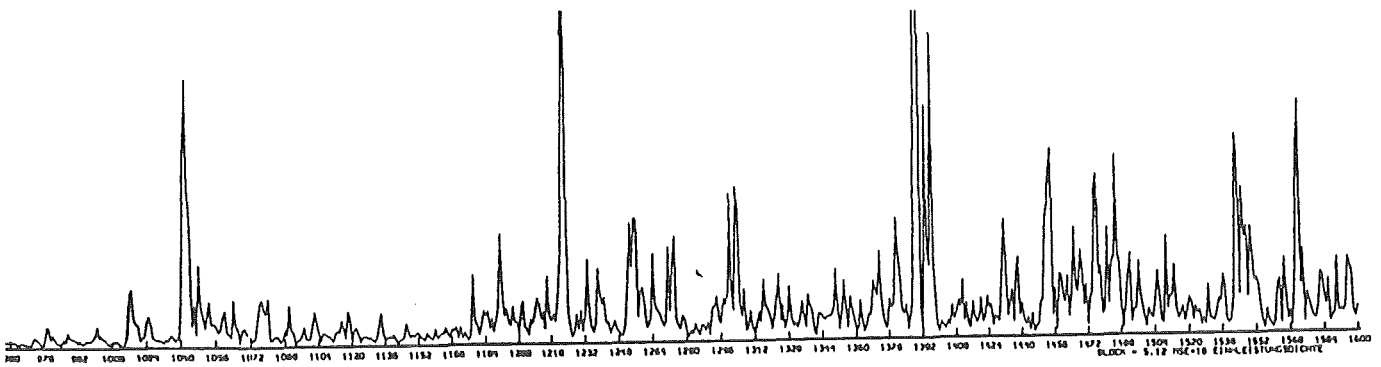
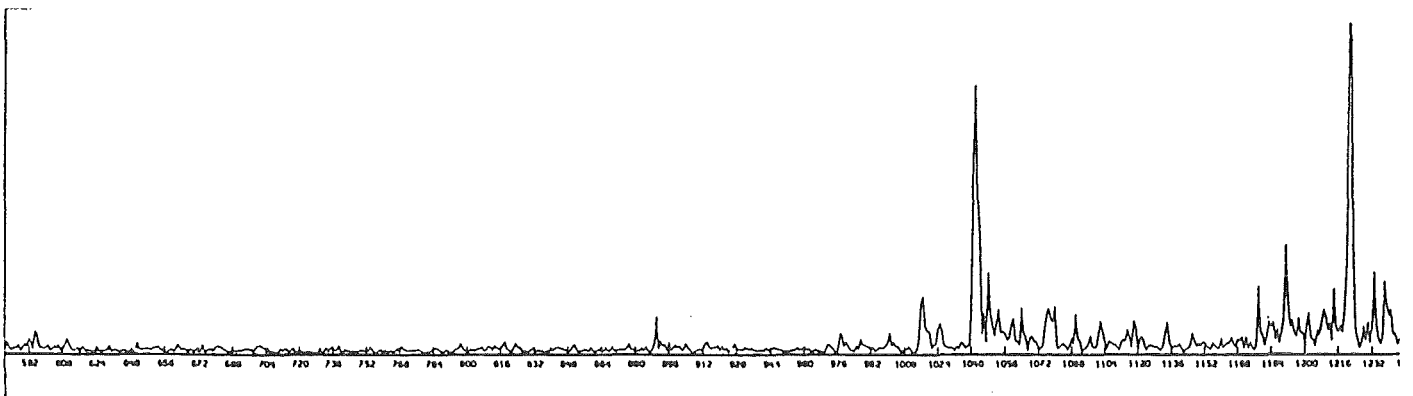
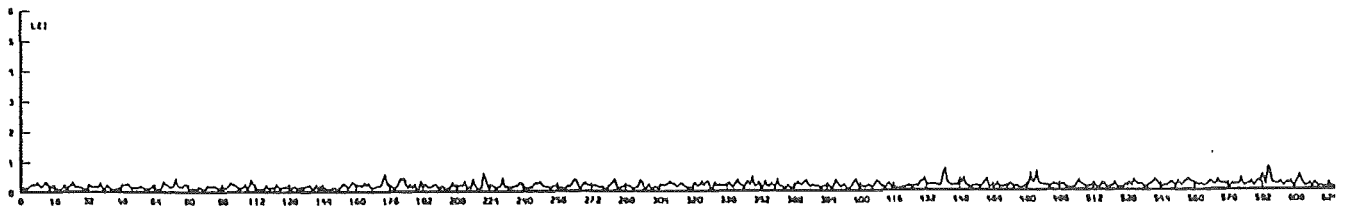


Abb. 6

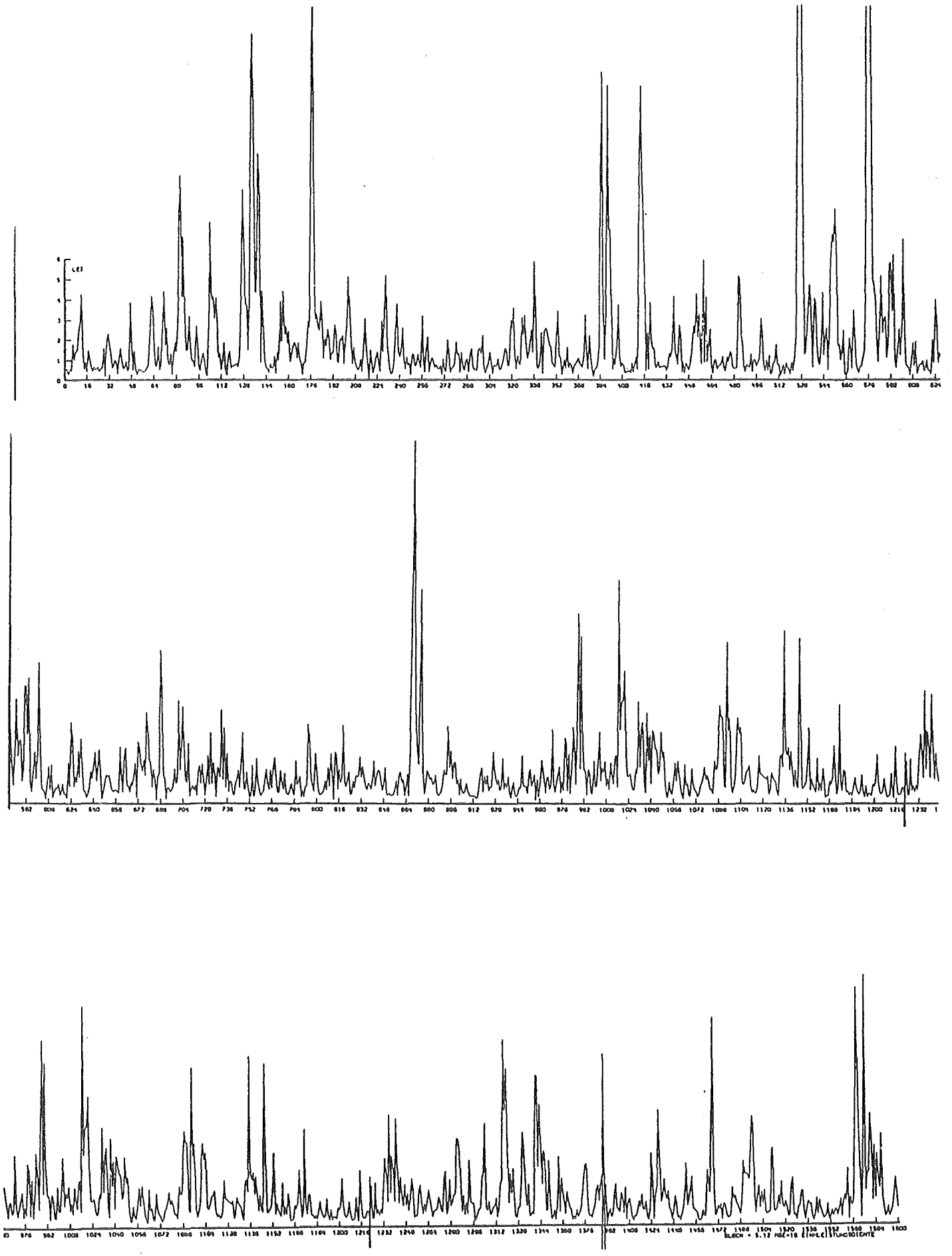


Abb. 7

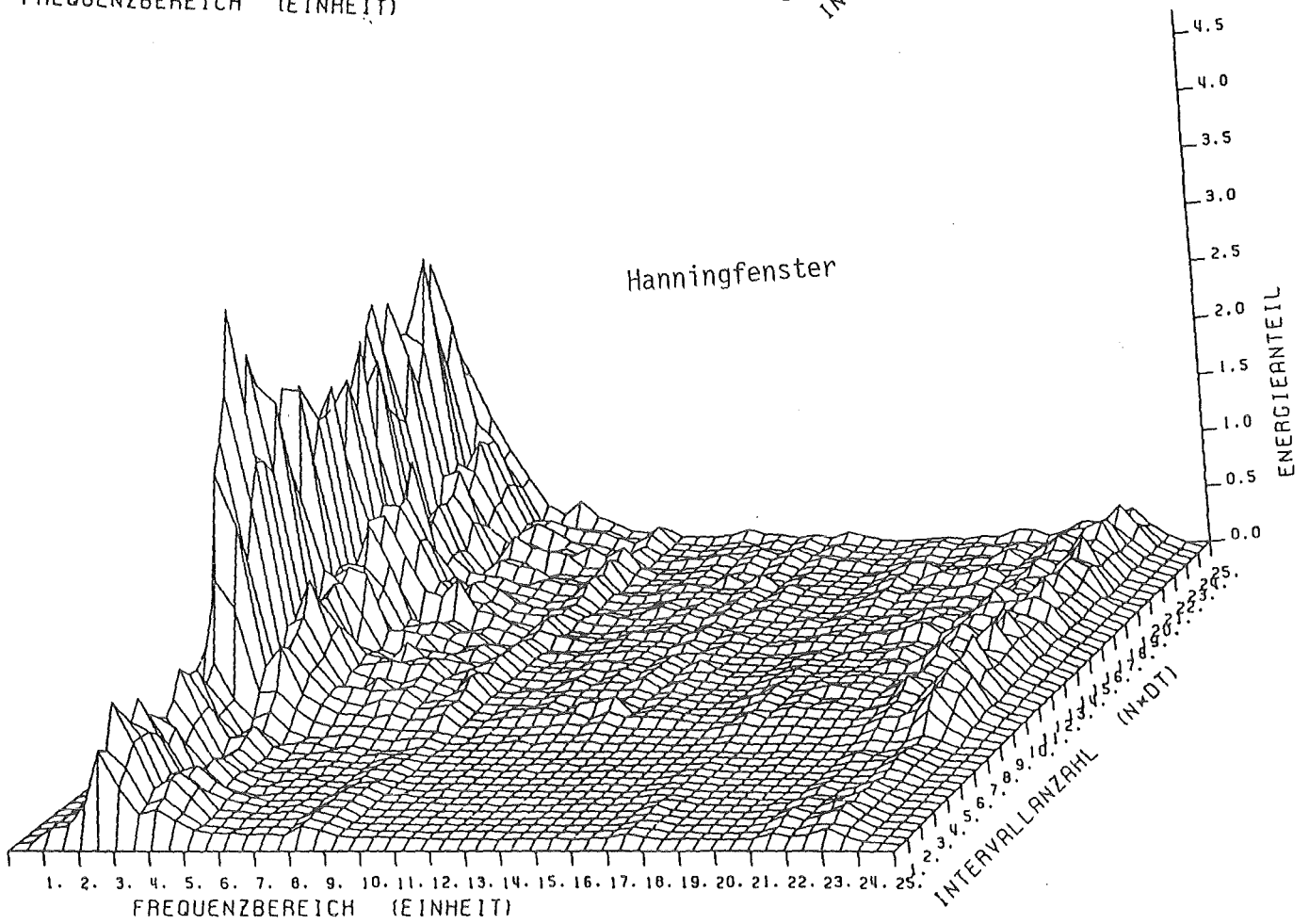
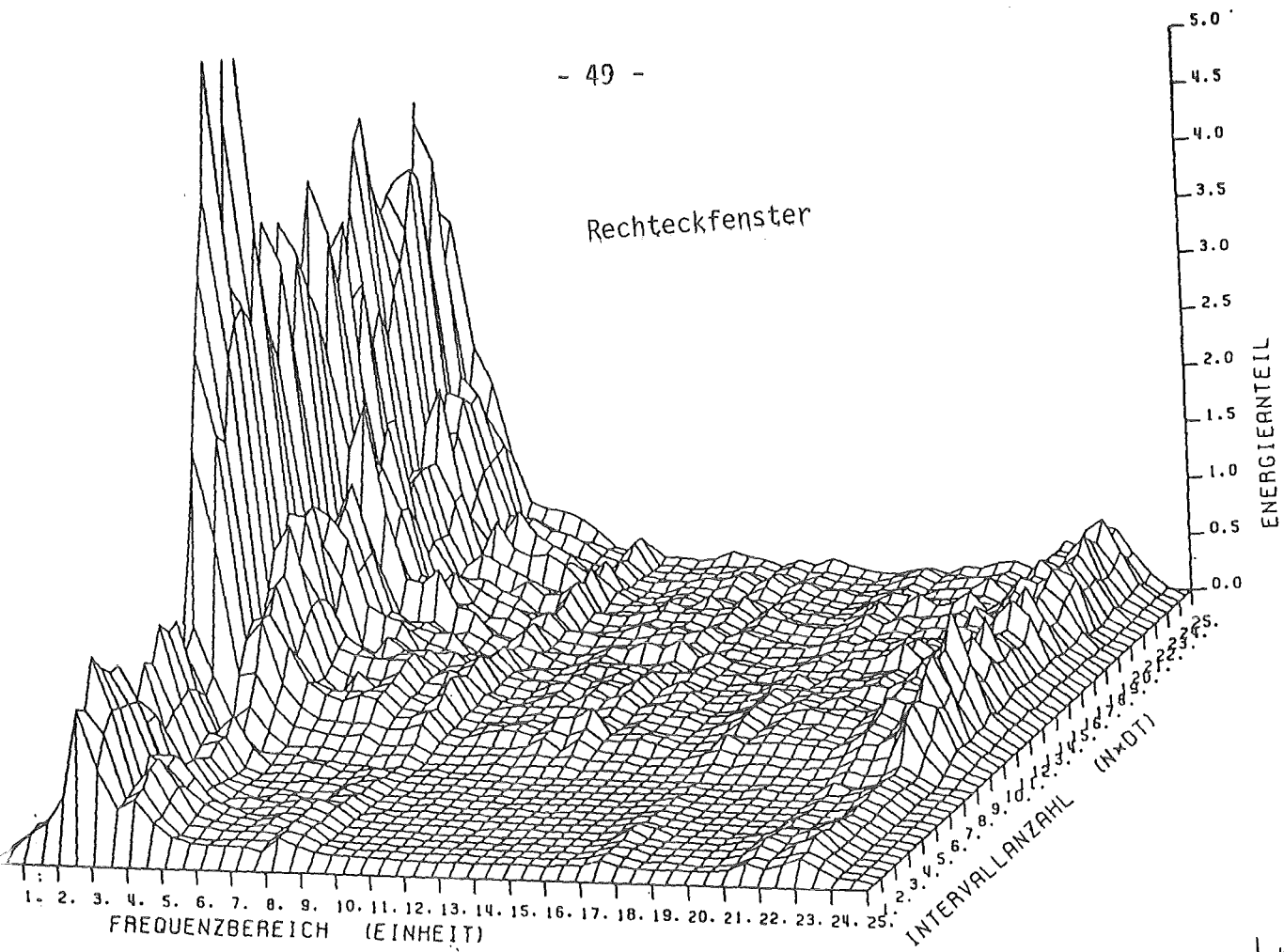


Abb. 8



- ↑ Linear Prediction
- × Dg. I. system
- z ARIMA Modellierung
- + Berliner Verfahren

Fensterlänge = FLEN = 60

Merkmal  $p^7$

Prognoselänge = VERSCH = 5

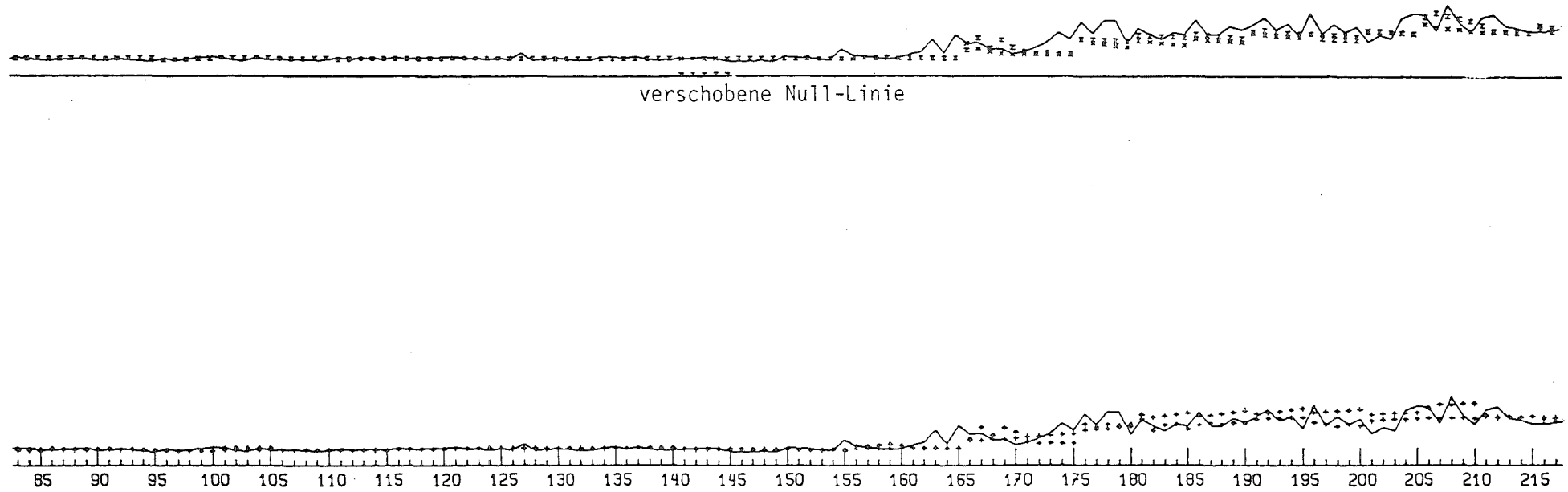
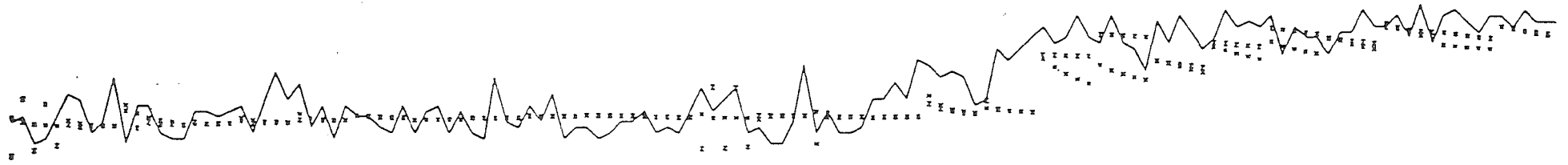


Abb. 9

Merkmal p<sup>17</sup>



verschobene Null-Linie

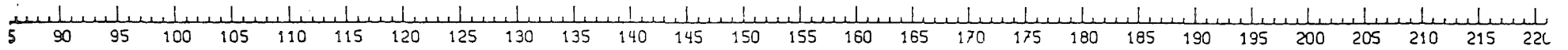
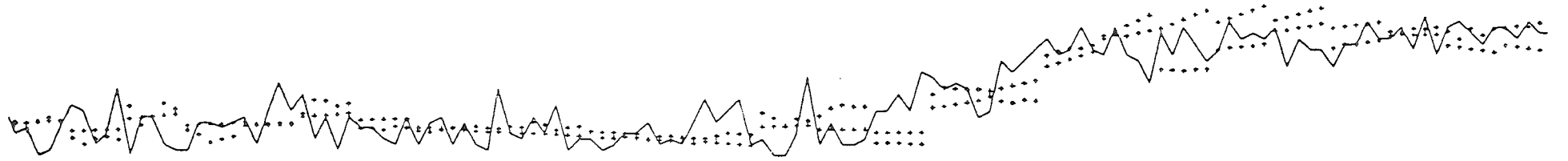


Abb. 10

+ Linear Prediction

o Dg. I. system

Fensterlänge = FLEN = 30

Merkmal  $p^7$

Prognoselänge = VERSCH = 5

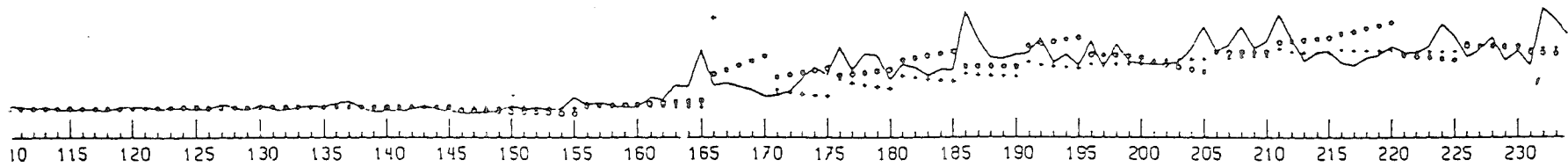


Abb. 11

Merkmal p<sup>17</sup>

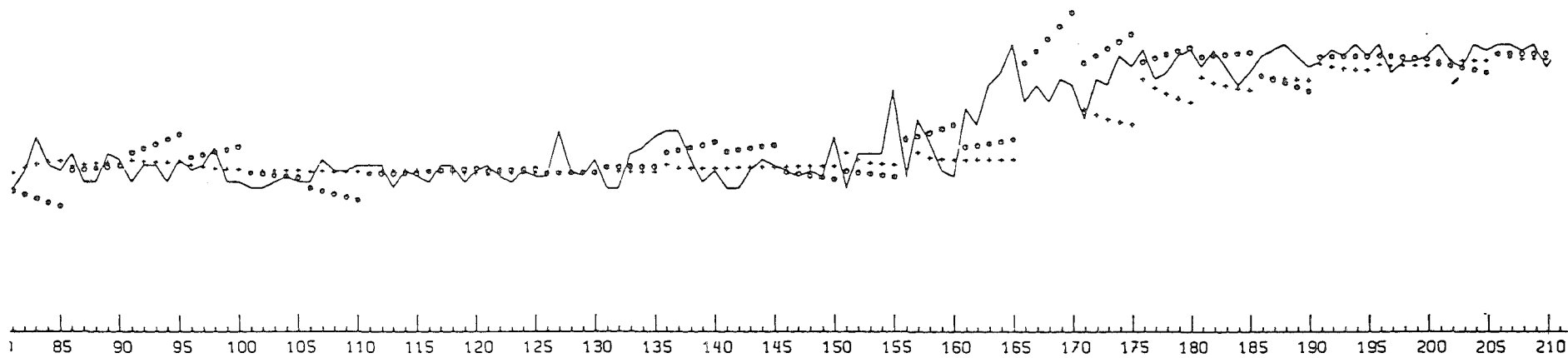


Abb. 12

+ Linear Prediction  
○ Dg. I. system

Fensterlänge = FLEN = 50  
Prognoselänge = VERSCH = 2

Merkmal  $p^7$

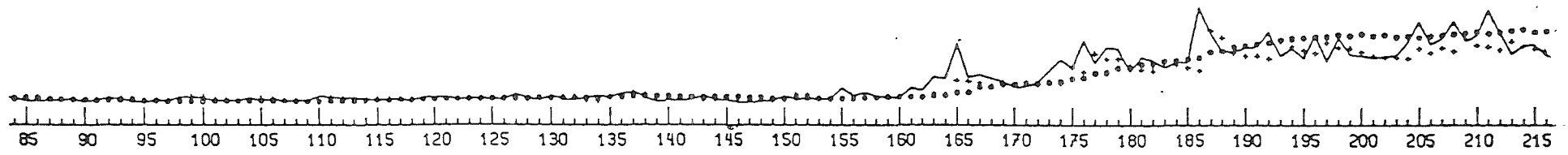


Abb. 13

Merkmal p<sup>17</sup>

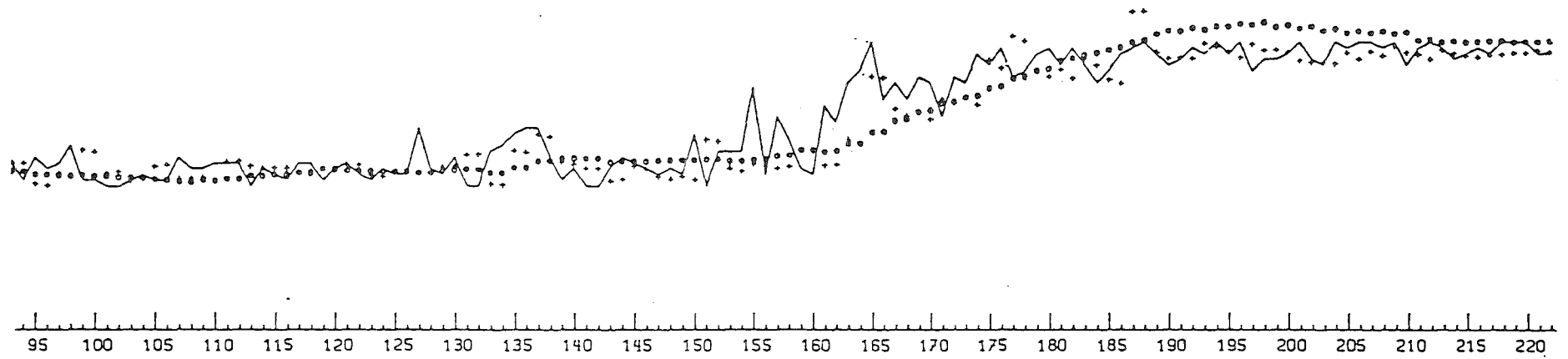


Abb. 14