

Vergleich von Identifikationsverfahren für konventionelle und Fuzzy-Modelle am Beispiel der Box-Jenkins-Daten¹

H. Reuter

Forschungsbericht Nr. 12/95

Meß-, Steuer- und Regelungstechnik

Übersicht: Im vorliegenden Bericht wird am Beispiel der Box-Jenkins-Daten ein Vergleich von Identifikationsverfahren für konventionelle und Fuzzy-Modelle durchgeführt. Es wird gezeigt, daß neben der geforderten Modellgüte, der Modellkomplexität und der Komplexität des Algorithmus bei einer Anwendung außerdem noch der Anwendungszweck des erhaltenen Modells von wesentlicher Bedeutung ist. Bei einer prädiktiven Modellauswertung (Einschrittprädiktion) ergeben sich nämlich bei den verschiedenen Identifikationsverfahren zum Teil sehr unterschiedliche Modellgüten im Vergleich zur rekursiven Modellauswertung (Mehrschrittvorhersage).

Gerhard-Mercator-Universität - GH Duisburg
Meß-, Steuer- und Regelungstechnik
Prof. Dr.-Ing. H. Schwarz

¹Dieser Bericht entstand im Rahmen des vom Ministerium für Wissenschaft und Forschung des Landes Nordrhein-Westfalen (MWF) geförderten Forschungsprojektes „Echtzeitsimulation mittels Fuzzy-Systemmodellen“.

Der Autor möchte sich außerdem bei den Herren Dipl.-Ing. A. Kroll und Dipl.-Ing. K. Küpper für die Diskussionen bedanken.

Inhaltsverzeichnis

1	Einleitung	1
2	Der Gasofenprozeß	4
	Lineares Modell nach Box und Jenkins	4
3	Berücksichtigte Verfahren und eigene Identifikationsergebnisse	6
	Lineares Modell mit MATLAB	6
	Bilineare Modelle	8
	Einfaches Hammerstein-Modell	8
	Einfaches Wiener-Modell	9
	Polynommodell nach Kortmann	9
	Letzter Meßwert	11
	Relational-Modell nach Tong	11
	Relational-Modelle nach Pedrycz	12
	Relational-Modelle nach Xu und Lu	13
	Funktional-Modelle nach Sugeno und Tanaka	14
	Positions-Gradienten-Modell nach Sugeno und Yasukawa	16
	Relational-Modell nach Küpper	18
	Arithmetik-Modell nach Bertram	19
	Funktional-Modelle nach Kroll	20
	Fuzzifiziertes Polynommodell	22
4	Vergleich der Ergebnisse und Bewertung	23
5	Zusammenfassung und Ausblick	26
6	Literaturverzeichnis	28

1 Einleitung

Geeignete Modelle, die das statische und dynamische Verhalten von technischen, chemischen, ökologischen, ökonomischen und anderen Prozessen hinreichend gut beschreiben, sind bei Simulation, Systemanalyse und Reglersynthese nicht nur in der Regelungstechnik unverzichtbar. Zum Erhalt solcher Prozeßmodelle führen zwei Vorgehensweisen, die theoretische und die experimentelle Modellbildung. Bei der theoretischen Modellbildung entsteht, ausgehend von den notwendigerweise bekannten Gesetzmäßigkeiten, denen der Prozeß unterliegt, ein Modell, das neben dem Ein-/Ausgangsverhalten auch die innere Struktur des zu modellierenden Prozesses abbildet. Solche Modelle werden häufig sehr komplex und ihre Erarbeitung ist sehr schwierig und zeitaufwendig. Bei der experimentellen Modellbildung, der Identifikation, wird aufgrund von Meßdaten ein mathematisches Modell berechnet. Hierfür müssen nicht alle Gesetzmäßigkeiten bekannt sein, deren Herausarbeitung vielfach sehr aufwendig ist. In der Praxis wird in vielen Fällen eine sich ergänzende Anwendung von theoretischer und experimenteller Prozeßanalyse durchgeführt, wobei man die Modellstruktur des geeigneten Modells aus bekanntem theoretischem und/oder heuristischem Wissen ableitet und die optimalen Werte unsicherer oder unbekannter Prozeßparameter mittels Parameterschätzverfahren ermittelt.

Mit zunehmenden Anforderungen an die Modellgenauigkeit und wachsender Komplexität der untersuchten Systeme werden in den letzten Jahren in starkem Maße nichtlineare Modelle verwendet (Schwarz 1991):

Sogenannte konventionelle Modelle werden üblicherweise mit Hilfe von Differential- und Differenzgleichungen dargestellt. Da es ein allgemeines nichtlineares Modell nicht gibt, muß auf spezielle nichtlineare Modellstrukturen wie z. B. Wiener- und Hammerstein-Modelle, bilineare Systeme oder Polynommodelle zurückgegriffen werden. Für solche nichtlineare Modelle fanden in den letzten Jahren in zunehmendem Maße Parameterschätzverfahren wie z. B. die Prädiktionsfehlerverfahren (Ljung 1987, Ljung und Söderström 1987) auch für reale Systeme Anwendung.

Eine andere Darstellungsart nichtlinearer Modelle ist die der Fuzzy-Systeme, die auf der unscharfen Logik nach Zadeh (1965) basiert. Auch hier existieren mehrere Modelltypen (wie z. B. relationale und funktionale Fuzzy-Modelle). Die Fuzzy-Funktional-Modelle beruhen auf der mathematischen Beschreibung von linguistischen WENN–DANN-Regeln, die als Folgerung eine funktionale Beziehung besitzen. Somit resultiert aus jeder Regel ein scharfer Zahlenwert, der mit den Schlußfolgerungen der anderen Regeln zum Gesamtausgang verknüpft wird. Die Fuzzy-Relational-Modelle besitzen demgegenüber als Schlußfolgerung unscharfe Werte, die im allgemeinen durch eine linguistische Größe dargestellt sind.

Wie die konventionelle Identifikation besteht auch die Fuzzy-Modellbildung aus den drei Phasen Voridentifikation (dazu zählt vor allem die Struktursuche), Parameter- oder Modellschätzung und Modellverifikation (Sugeno und Yasukawa 1993, Reuter 1995). Das bei

weitem größte Problem hierbei stellt für beide Identifikationsrichtungen die Wahl einer geeigneten Modellstruktur dar. Zwar existieren Struktursuchverfahren, die bei Anwendung auf „gutmütige“ Prozesse auch gute Ergebnisse liefern, in der Regel jedoch ist hierbei viel analytisches Prozeßwissen, ein sehr großer Aufwand und letztendlich längeres „Herumprobieren“ unumgänglich. Obgleich gerade den Fuzzy-Modellen als universellen Approximatoren nachgesagt wird, daß sie beliebig nichtlineares Verhalten abbilden können, trifft auch sie das Problem der Struktursuche genauso wie die konventionellen Modelle.

- Dieser Bericht beschäftigt sich deshalb mit dem Vergleich von Identifikationsverfahren für konventionelle und Fuzzy-Modelle. Als Vergleichsprozeß werden die Meßdaten eines Gasofenprozesses (Box und Jenkins 1976) verwendet, die bei der Identifikation eine Art Benchmarkproblem darstellen (Abschnitt 2).
- In Abschnitt 3 werden anschließend die verwendeten 22 Verfahren und Modelle und die ihnen zugrunde liegenden Ideen kurz skizziert. Für eine genauere Beschreibung der Algorithmen sei an dieser Stelle auf die entsprechende Originalliteratur verwiesen.
- Abschnitt 4 vergleicht die Ergebnisse der einzelnen Verfahren und bewertet sie. Einen solchen Vergleichsparameter für die Modellkomplexität stellt die Anzahl der Modellparameter dar. Während bei konventionellen Modellen (bis auf Ausnahmen) die Parameter global gültig sind, wirken Zugehörigkeitsfunktionen für Fuzzy-Modelle (in den meisten Fällen) signifikant nur lokal. Zur Beschreibung eines Systemverhaltens sind wesentlich mehr freie Parameter erforderlich. Diese lokale Gültigkeit besitzt jedoch den Vorteil, daß das Modell lokal an einzelnen Stellen an reale Gegebenheiten angepaßt werden kann, ohne das globale Verhalten zu verändern (ggf. zu verschlechtern). Aus diesem Grund wird bei den Fuzzy-Modellen auch die Anzahl der verwendeten Regeln angegeben. Die hohe Anzahl der Parameter bei Fuzzy-Modellen wirkt sich jedoch erfahrungsgemäß auch auf das Konvergenzverhalten der meisten Schätzverfahren aus.

Ein anderes wichtiges Kriterium ist die Modellgüte. Bei dem in dieser Arbeit zur Beurteilung herangezogenen mittleren quadratischen Fehler wird zudem unterschieden zwischen prädiktiver und rekursiver Auswertung des Modells, also dem Anwendungszweck des Modells. Bei der prädiktiven Modellauswertung (Einschrittprädiktion; d. h., vorausgesagt wird nur der nächste Ausgangswert), wie sie beispielsweise bei adaptiven Regelungen (Isermann, Lachmann und Matko 1992, Behmenburg 1995) durchgeführt wird, sind die Anforderungen an die Modellgenauigkeit naturgemäß nicht so hoch wie bei der rekursiven Auswertung (Vorhersage des Prozeßverhaltens für einen längeren Zeitraum).

Bei der Verwendung der Modelle können Fuzzy-Modelle gerade bei zeitkritischen Echtzeitanwendungen einen Vorteil bedeuten, da sie als Kennfeld abgelegt werden

können und somit zu ihrer Auswertung nur ein minimaler Rechenzeitbedarf erforderlich ist.

Ein direkter Vergleich des Schätzaufwandes (z. B. Komplexität des Algorithmus, Programmieraufwand, Rechenzeitbedarf usw.) ist jedoch nicht möglich, da bei einem Teil der Verfahren auf die Literatur zurückgegriffen werden mußte.

- Zusammenfassung und Ausblick schließen diesen Bericht ab.

2 Der Gasofenprozeß

Bei den sog. Box-Jenkins-Daten handelt es sich um Meßdaten eines Gasofenprozesses (Box und Jenkins 1976), die insbesondere bei der Fuzzy-Identifikation als Benchmarkproblem und zum Vergleich der Genauigkeit einzelner Verfahren herangezogen werden. Die 296 Datenpaare (Bilder 2.1 und 2.2) beschreiben das Klemmenverhalten eines Teils eines Gasofens bei einer Abtastzeit $T_a = 9$ s. Über die Stellgröße U wird der Brennstoffstrom

$$\dot{V}_{\text{Methan}}(t) = (0,6 - 0,04 U(t)) \frac{\text{cu. ft}}{\text{min}} \quad (2.1)$$

des Gasofens – ein Methanvolumenstrom – variiert. Hierbei nimmt die Stellgröße in der Regel Werte im Bereich $U \in [-2,5; 2,5]$ an. Der sich einstellende Brennstoffvolumenstrom ergibt sich dann zu $Y \in [0,5; 0,7]$ cu. ft/min bzw. $Y \in [14,2; 19,8]$ l/min. Der CO_2 -Gehalt des Abgases in % stellt die Ausgangsgröße Y dieses Verbrennungsprozesses dar.

Bild 2.1: Stellsignal $U(k)$

Modell 1: Lineares System (Box und Jenkins 1976)

Box und Jenkins (1976) geben hierfür ein mittels eines iterativen, nichtlinearen Least-Squares-Verfahrens identifiziertes Modell mit

$$y(k) = 0,57 y(k-1) - 0,53 u(k-3) - 0,37 u(k-4) - 0,51 u(k-5) \quad (2.2)$$

an, das eine Verstärkung von $K \approx -3,3$ besitzt. Hierbei verwenden sie

$$u(k) = U(k) - \bar{U} \quad \text{und} \quad y(k) = Y(k) - \bar{Y} \quad (2.3)$$

Bild 2.2: Ausgangssignal $Y(k)$

mit dem Arbeitspunkt $\bar{U} = 0,057$ und $\bar{Y} = 53,51$. Anhand von Bild 2.2 ist zu sehen, daß dieses lineare Modell für die ersten 220 Meßwertpaare eine gute Approximation darstellt. Die anschließend auftretenden großen Abweichungen rühren von nicht im Modell enthaltenen Signalanteilen. Da dieses Phänomen auch bei allen anderen identifizierten Modellen (auch den nichtlinearen) auftritt, liegt die Vermutung nahe, daß es sich hierbei um die Auswirkungen einer unbekanntes und somit nicht berücksichtigten zusätzlichen Eingangsgröße handeln dürfte.

□

Außerdem wird in Bild 2.2 deutlich, daß eine rekursive Modellauswertung mit

$$\hat{y}(k) = f(u(k), \dots, u(k-n), \hat{y}(k-1), \dots, \hat{y}(k-n)) \quad , \quad (2.4)$$

bei der die geschätzten zeitlich zurückliegenden Ausgänge $\hat{y}(k-x)$ ($x \in \{1, \dots, n\}$ mit $n =$ Modellordnung) verwendet werden, naturgemäß nicht so gut sein kann wie eine prädiktive Auswertung mit

$$\hat{y}(k) = f(u(k), \dots, u(k-n), y(k-1), \dots, y(k-n)) \quad , \quad (2.5)$$

bei der die gemessenen Werte zur Verfügung stehen.

3 Berücksichtigte Verfahren und eigene Identifikationsergebnisse

Im folgenden werden die verwendeten Identifikationsverfahren kurz beschrieben. Allen im Rahmen dieses Berichtes durchgeführten eigenen Rechnungen liegt ein Arbeitspunkt

$$\bar{U} = 0 \quad \text{und} \quad \bar{Y} = 53,5 \quad (3.1)$$

(vgl. Gl. (2.3)) zugrunde. Zudem wird als Schreibweise für lineare Blöcke in der Regel die Darstellung

$$A(q^{-1}) y(k) = B(q^{-1}) q^{-d} u(k) \quad \text{mit} \quad (3.2)$$

$$B(q^{-1}) = b_0 + b_1 q^{-1} + b_2 q^{-2} + \dots + b_n q^{-n} \quad \text{und} \quad (3.3)$$

$$A(q^{-1}) = 1 + a_1 q^{-1} + a_2 q^{-2} + \dots + a_n q^{-n} \quad (3.4)$$

gewählt. Hierbei stellt d die diskrete Modelltotzeit, n die Modellordnung und q^{-1} den Zeitverschiebeoperator mit

$$q^{-m} x(k) = x(k - m) \quad (3.5)$$

(Wernstedt 1989) dar.

Modell 2: Lineares Modell mit MATLAB (Ljung 1988)

Zusätzlich zum linearen Modell (2.2) wird ein weiteres lineares Modell mit einem Prädiktionsfehlerverfahren nach Ljung (1987) aus der MATLAB System Identification Toolbox identifiziert. Diesem Prädiktionsfehlerverfahren liegt ein iterativer Gauß-Newton-Algorithmus zugrunde.

Der verwendete MATLAB-Algorithmus `pem.m` identifizierte das beste lineare Modell der Struktur

$$A(q^{-1}) y(k) = \frac{B(q^{-1})}{F(q^{-1})} q^{-d} u(k) + \frac{C(q^{-1})}{D(q^{-1})} \epsilon(k) \quad (3.6)$$

mit dem MATLAB-Befehl

```
[A,B,C,D,F]=polyform(pem([y,u],[5,6,0,0,0,0]));
```

Dies ergibt ein lineares Modell der Ordnung $n = 5$ in der Darstellung (3.2) mit

$$\begin{aligned} A(q^{-1}) &= 1 - 1,5691 q^{-1} + 0,6232 q^{-2} \\ &\quad + 0,1906 q^{-3} - 0,1706 q^{-4} + 0,0105 q^{-5} \\ B(q^{-1}) &= -0,0496 + 0,1526 q^{-1} - 0,1890 q^{-2} \\ &\quad - 0,4134 q^{-3} + 0,0240 q^{-4} + 0,1931 q^{-5} \\ \text{sowie} \quad C(q^{-1}) &= D(q^{-1}) = F(q^{-1}) = 1 \end{aligned} \quad (3.7)$$

und einer zeitdiskreten Modelltotzeit $d = 0$.

□

Als Alternative zu einem linearen Modell besteht natürlich die Möglichkeit, eine Approximation des Prozeßverhaltens mit (verschiedenen) nichtlinearen Modellstrukturen zu erreichen. Hierbei kann man zwischen Modellen unterscheiden, die dynamische (z. B. bilineare Systeme) oder statische (z. B. einfache Wiener- und Hammerstein-Modelle) Nichtlinearitäten besitzen oder beides (z. B. Polynommodelle).

Ein iteratives Prädiktionsfehlerverfahren

Zur Identifikation von Modellen, die nicht linear in ihren Parametern darzustellen sind, können „einfache“, d. h. lineare Parameterschätzverfahren wie die Least-Squares-Verfahren nicht mehr verwendet werden. Es muß auf nichtlineare Optimierungsalgorithmen zurückgegriffen werden. Nach Erfahrungen des Autors (Reuter 1995) bieten sich hierfür die Prädiktionsfehlerverfahren nach Ljung (1987) besonders an. Das im Rahmen dieser Arbeit verwendete iterative Prädiktionsfehlerverfahren verwendet einen gedämpften Gauß-Newton-Algorithmus

$$\hat{\Theta}_{i+1} = \hat{\Theta}_i + \alpha_i \left[\sum_{k=1}^N \Psi(k) \Psi^T(k) \right]^{-1} \sum_{k=1}^N \Psi(k) \quad (3.8)$$

(Reuter 1995), der mit einer Approximation der normalerweise erforderlichen Hesse-Matrix durch

$$\frac{\partial^2 V_G}{\partial \Theta \partial \Theta^T} \approx \sum_{k=1}^N \Psi(k) \Psi^T(k) \quad (3.9)$$

die Positiv-Semidefinitheit garantiert (Ljung und Söderström 1987). Als Gütefunktional wird

$$V_G = \frac{1}{2} \sum_{k=1}^N e^2(k) \quad (3.10)$$

mit dem Vorhersage- oder Prädiktionsfehler

$$\hat{e}(k) = y(k) - \hat{y}(k) \quad (3.11)$$

über alle N zur Verfügung stehenden Meßwerte verwendet ($\hat{y}(k)$ bezeichnet das geschätzte Ausgangssignal und $y(k)$ das gemessene). Zur Berechnung des geschätzten Parametervektors $\hat{\Theta}_{i+1}$ beim $(i + 1)$ -ten Iterationsschritt benötigt man neben einem geeignet zur wählenden Dämpfungsfaktor α_i – Hinweise dazu, zu günstigen Startparametern für den Algorithmus und zur Beeinflussung des Konvergenzverhaltens finden sich in (Reuter 1995) – noch den Prädiktorgradienten

$$\Psi(k) = \frac{\partial \hat{y}(k)}{\partial \Theta} \quad (3.12)$$

mit der Ableitung des geschätzten Modellausganges $\hat{y}(k)$ nach den einzelnen Modellparametern. Für alle hier durchgeführten Rechnungen gilt $\alpha_i = 0, 1$.

Modell 3: Bilineare Modelle (Reuter 1995)

Als geeignet zur Approximation dynamisch nichtlinearen Verhaltens haben sich bilineare Modelle mit der Struktur

$$\begin{aligned} \mathbf{x}(k+1) &= \mathbf{A} \mathbf{x}(k) + \mathbf{N} \mathbf{x}(k) u(k-d) + \mathbf{b} u(k-d) \\ y(k) &= \mathbf{c}^T \mathbf{x}(k) \end{aligned} \quad (3.13)$$

herausgestellt (Schwarz 1991, Mohler 1973). Für die Parameteridentifikation muß aus mehreren Gründen (z. B. Konvergenzverhalten, Identifizierbarkeitsproblem) auf kanonische Formen zurückgegriffen werden. In (Reuter 1995) finden sich die Identifikationsgleichungen für vier kanonische Modelle, die eine bilineare Erweiterung der vier linearen Frobeniuskanonischen Formen darstellen.

Eine Konvergenz der Parameter gegen ein Gütefunktionsminimum trat nur für $n = 3$ ein. Bei größeren Ordnungen $n > 3$ konvergierten die Parameterverläufe nicht, die Modelle waren für diese Modellstruktur wohl überparametrisiert. Daran änderte auch die Berücksichtigung einer Modelltotzeit $d > 0$ nichts. Als beste bilineare beobachtbarkeitskanonische Darstellung (BLS-OYCF) ergab sich mit vollem Differenzegrad (Schwarz 1991)

Modell 3a zu

$$\begin{aligned} \mathbf{A} &= \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0,2979 & -1,1554 & 1,7307 \end{bmatrix}; & \mathbf{N} &= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ -0,0035 & 0,0035 & 0,0069 \end{bmatrix}; \\ \mathbf{b} &= \begin{bmatrix} 0 & 0 & -0,4090 \end{bmatrix}^T; & \mathbf{c} &= \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}^T, \end{aligned} \quad (3.14)$$

die beobachterkanonische Form (BLS-ORCF) mit **Modell 3b** zu

$$\begin{aligned} \mathbf{A} &= \begin{bmatrix} 0 & 0 & 0,1417 \\ 1 & 0 & -0,6511 \\ 0 & 1 & 1,2931 \end{bmatrix}; & \mathbf{N} &= \begin{bmatrix} 0 & 0 & -0,0193 \\ 0 & 0 & 0,0520 \\ 0 & 0 & -0,0197 \end{bmatrix}; \\ \mathbf{b} &= \begin{bmatrix} -1,4262 & 1,1940 & -0,4622 \end{bmatrix}^T; & \mathbf{c} &= \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}. \end{aligned} \quad (3.15)$$

Die bilinearen Anteile \mathbf{N} besitzen keinen signifikanten Beitrag zum Übertragungsverhalten, so daß die Aussage gerechtfertigt ist, daß der Gasofenprozeß wohl keine dynamische sondern eine statische Nichtlinearität besitzt.

□

Modell 4: Einfaches Hammerstein-Modell (Reuter 1995)

Mit den Identifikationsgleichungen des iterativen Prädiktionsfehlerverfahrens für ein einfaches Hammerstein-Modell (Reuter 1995)

$$\begin{aligned} x(k) &= \gamma_0 + \sum_{r=1}^p \gamma_r u^r(k) \\ A(q^{-1}) \hat{y}(k) &= B(q^{-1}) q^{-d} x(k) \end{aligned} \quad (3.16)$$

ergab sich das beste Modell zu

$$\begin{aligned} x(k) &= -0,0106 + u(k) + 0,0105 u^2(k) \quad , \\ B(q^{-1}) &= -0,1045 + 0,2249 q^{-1} + 0,1019 q^{-2} \\ &\quad -1,1529 q^{-3} + 0,5850 q^{-4} + 0,0688 q^{-5} \quad , \\ A(q^{-1}) &= 1 - 1,5592 q^{-1} + 0,6419 q^{-2} \\ &\quad +0,0914 q^{-3} - 0,0668 q^{-4} - 0,0243 q^{-5} \quad . \end{aligned} \quad (3.17)$$

Für Polynomordnungen $p > 2$ und Modellordnungen $n > 5$ ergab sich eine Verschlechterung der Modellgüte infolge von Überparametrisierung.

□

Modell 5: Einfaches Wiener-Modell (Reuter 1995)

Die Identifikation eines einfachen Wiener-Modells

$$\begin{aligned} A(q^{-1}) x(k) &= B(q^{-1}) q^{-d} u(k) \\ \hat{y}(k) &= \gamma_0 + \sum_{r=1}^N \gamma_r x^r(k) \end{aligned} \quad (3.18)$$

mit dem iterativen Prädiktionsfehlerverfahren (Gl. (3.8)) nach (Reuter 1995) lieferte als bestes Modell

$$\begin{aligned} B(q^{-1}) &= -0,2885 + 0,8580 q^{-1} - 0,5637 q^{-2} \\ &\quad -1,3211 q^{-3} + 2,2305 q^{-4} - 0,9559 q^{-5} \quad , \\ A(q^{-1}) &= 1 - 3,0922 q^{-1} + 3,7451 q^{-2} \\ &\quad -2,2068 q^{-3} + 0,6433 q^{-4} - 0,0781 q^{-5} \quad \text{und} \\ \hat{y}(k) &= 0,1130 + x(k) - 0,0450 x^2(k) - 0,0011 x^3(k) + 0,0007 x^4(k) \quad . \end{aligned}$$

Obwohl hier die Parameter γ_3 und γ_4 einen sehr kleinen Wert besitzen, erhält man durch ihr Weglassen doch eine leichte Modellverschlechterung ebenso wie bei einer Wahl von $n > 5$ und/oder $p > 4$.

□

Modell 6: Polynommodell (Kortmann 1989)

Der direkte Strukturelektionsalgorithmus von Kortmann (1989) sucht die bezüglich des Übertragungsverhaltens signifikanten Parameter aus einem Kolmogorov-Gabor-Polynommodell

$$\begin{aligned}
\hat{y}(k) = & \bar{y} + \sum_{i_1=0}^n b_{i_1} u(k-i_1) + \sum_{i_1=0}^n \sum_{i_2=i_1}^n b_{i_1,i_2} u(k-i_1) u(k-i_2) \\
& + \dots + \sum_{i_1=0}^n \dots \sum_{i_p=i_{p-1}}^n b_{i_1,\dots,i_p} u(k-i_1) \dots u(k-i_p) \\
& + \sum_{i_1=1}^m a_{i_1} y(k-i_1) + \sum_{i_1=1}^m \sum_{i_2=i_1}^m a_{i_1,i_2} y(k-i_1) y(k-i_2) \\
& + \dots + \sum_{i_1=1}^m \dots \sum_{i_p=i_{p-1}}^m a_{i_1,\dots,i_p} y(k-i_1) \dots y(k-i_p) \\
& + \sum_{i_1=0}^n \sum_{i_2=1}^m c_{i_1,i_2} u(k-i_1) y(k-i_2) \\
& + \dots + \sum_{i_1=0}^n \dots \sum_{i_{p-1}=i_p}^n \sum_{i_p=1}^m c_{i_1,\dots,i_{p-1},i_p} u(k-i_1) \dots u(k-i_{p-1}) y(k-i_p) \\
& + \dots + \sum_{i_1=0}^n \sum_{i_2=1}^m \dots \sum_{i_p=i_{p-1}}^m c_{i_1,i_2,\dots,i_p} u(k-i_1) y(k-i_2) \dots y(k-i_p) \quad (3.19)
\end{aligned}$$

heraus und berechnet dazu mittels des nichtrekursiven Least-Squares-Verfahrens (Isermann 1988)

$$\hat{\Theta} = [\varphi^T \varphi]^{-1} \varphi^T \mathbf{y} \quad (3.20)$$

mit dem Meßvektor φ die zugehörigen Parameter, wobei Linearität in den Parametern für diese Modellbeschreibung gegeben ist:

$$y(k) = \varphi^T(k) \Theta + \epsilon(k) \quad (3.21)$$

Dies geschieht in sieben Schritten:

Im 1. Schritt erfolgt die Vorgabe der maximalen Modellgröße (Parameter n , m , p in Gl. (3.19)). Anschließend werden alle möglichen Signalanteile des Modells zueinander orthogonalisiert und es wird ein Einflußwert jedes Signalanteils auf das Ein-/Ausgangsverhalten ermittelt (Schritt 2). Anhand dieser Einflußwerte erfolgt in Schritt 3 die Auswahl eines optimalen Signalanteils bzw. Parameters und seine Übernahme in ein vorläufiges Modell. Das so stufenweise aufgebaute vorläufige Modell wird nun bezüglich der Modellgüte untersucht: Verschiedene statistische Kriterien ermitteln die Güte des vorläufigen Modells (Schritt 4). In Schritt 5 werden die Fragen geklärt, ob das vorläufige Modell überhaupt statistisch signifikante Parameter enthält (wichtig für den ersten Durchlauf durch den Algorithmus) und ob gegenüber dem vorherigen Durchlauf eine Verbesserung erzielt wurde. Sind beide Antworten positiv, wird die Signifikanz aller Parameter im vorläufigen Modell

bestimmt (Schritt 6). In Schritt 7 folgt dann die Entscheidung, ob der in diesem Durchlauf im 2. Schritt ausgewählte Parameter im Modell enthalten bleibt, einen oder mehrere andere Parameter ersetzt oder keine signifikante Verbesserung erbringt und deshalb selber wieder daraus zu entfernen ist. Der Algorithmus wird beendet, wenn keine Modellverbesserung mehr eintritt.

Das bezüglich des mittleren quadratischen Fehlers beste Modell mit diesem Algorithmus besitzt die Darstellung

$$\begin{aligned}
 y(k) = & -0,0091 - 0,3401 u(k-3) \\
 & +1,7446 y(k-1) - 1,1290 y(k-2) + 0,2767 y(k-3) \\
 & -0,0293 u^2(k) - 0,1974 y(k-1) u^2(k) + 0,2369 y(k-2) u^2(k) \\
 & -0,0879 y(k-3) u^2(k) - 0,1160 u(k-3) u^2(k) - 0,0017 u^4(k) \quad . \quad (3.22)
 \end{aligned}$$

□

Modell 7: Letzter Meßwert

Als Vergleichswert und zur besseren Bewertung der übrigen Ergebnisse wird ein sehr einfaches Verfahren gegenübergestellt, das für einen unbekanntem Wert den letzten bekannten schätzt. Das bedeutet also

$$\hat{y}(k) = y(k-1) \quad (3.23)$$

für prädiktive Auswertung.

□

Modell 8: Relational-Modell (Tong 1978)

Tong (1978) schätzt die Regeln einer Fuzzy-Struktur, indem er eine Methode namens *Logische Prüfung* („logical examination“) verwendet. Dazu wird ein „Datenmuster“ (z. B. $\langle u(k-1); y(k-1); y(k) \rangle$), bestehend aus Modelleingangs- und Modellausgangsdaten, aufgestellt. Nach einer Quantisierung der Meßwerte (Einordnung in diskrete Klassen) ergibt sich ein Folge von Indizes (jeder Klasse wird ein Index zugeordnet), die die quantisierten Meßwerte auf vorher definierte Modellklassen (linguistische Werte wie z. B. *groß*, *mittel* oder *klein*) bezieht: Man erhält also für jeden Meßdatentupel eine Regel. Wenn der Wert der Zugehörigkeitsfunktion für diese Regel 1 beträgt, dann beschreibt sie den jeweiligen Meßdatentupel exakt. Diese Regel stellt eine sog. *elementare Regel* für die Meßdaten dar. Regeln mit Zugehörigkeitswerten kleiner 1 (nicht elementare Regeln) werden nicht berücksichtigt.

Ergeben sich aus den elementaren Regeln Konflikte, so kann dies an verrauschten Meßwerten, der Quantisierung (z. B. Anzahl der Klassen) oder einer falschen Modellstruktur (ungünstig gewähltes Datenmuster) liegen. Zur Auflösung dieser Konflikte muß dann qualitatives Prozeßwissen (also sog. Experten- und/oder Domänenwissen) herangezogen werden.

Tong (1978) schlägt das Datenmuster $\langle u(k-4); y(k-1); y(k) \rangle$ für die Box-Jenkins-Daten vor. Er verwendet zur Quantisierung jeweils 6 Klassen und erhält 19 elementare Regeln. \square

Modell 9: Relational-Modell (Pedrycz 1984), Pedrycz + (Küpper 1995)

Modell 9a

Beim Verfahren von Pedrycz (1984) wird jeweils eine eindimensionale Clusterung der Meßdaten (d. h. für jeden Eingangsraum einzeln) mit dem FCM-Algorithmus von Bezdek (1981) durchgeführt (FCM = Fuzzy-*c*-Means). Ergebnis der Clusterung sind die Fuzzy-Referenzmengen, wobei die Anzahl der Cluster vorzugeben ist. Über die Wahl der Anzahl der Fuzzy-Referenzmengen aus Eingangsraum \mathbb{U} und Zustandsraum \mathbb{X} kann keine allgemeine Aussage getroffen werden, sie ist abhängig vom jeweiligen Anwendungsfall.

Im nächsten Schritt ist die Struktur des Fuzzy-Modells mit der Relationalgleichung

$$Y(k) = X_1(k) \circ X_2(k) \circ \dots \circ X_p(k) \circ \mathbf{R}(k) \quad (3.24)$$

vorzugeben, wobei die $X_i(k)$ sowohl Stellgrößen $U_i(k-a-d)$ als auch rückgeführte verzögerte Ausgangssignale $Y_i(k-b)$ oder Kombinationen hiervon darstellen können. Ebenfalls sind hierbei Modelltotzeiten d zu berücksichtigen.

Die Relationalmatrix $\mathbf{R}(k)$ wird für jeden Meßdatensatz mittels des unscharfen kartesischen Produkts (Bertram 1991) der Vektoren der Möglichkeitsgrade bzw. der Gewichte der erfüllten Regeln nach

$$\mathbf{R}(k) = X_1(k) \otimes X_2(k) \otimes \dots \otimes X_p(k) \otimes Y(k) \quad (3.25)$$

bestimmt. Die Teilrelationalmatrizen werden mittel Disjunktion zur Gesamtrationalmatrix

$$\mathbf{R} = \bigvee_{h=1}^N \mathbf{R}(h) \quad (3.26)$$

zusammengesetzt.

Für die Box-Jenkins-Daten verwendet Pedrycz (1984) die Modellstruktur

$$Y(k) = Y(k-1) \circ U(k-3) \circ \mathbf{R} \quad . \quad (3.27)$$

Hierbei optimiert er iterativ durch Variation eine Euklidische Abstandsnorm bezüglich der Modelltotzeit d und des Kompositionsoperators \circ . Die besten Ergebnisse erhielt Pedrycz (1984) mit dem Produktoperator für die Konjunktion (3.25), dem Maximumoperator für die Disjunktion (3.26) sowie 9 Fuzzy-Referenzmengen für jede Fuzzy-Größe. Dabei beschränkt er sich auf eine Folgerung pro Regelsatz, nämlich die mit dem größten Möglichkeitsgrad bzw. Gewicht. So fallen Regelsätze mit einem Möglichkeitsgrad kleiner als dem Schwellwert 0,1 heraus. Das bedeutet, daß er in diesem Fall zwar 729 maximal mögliche Parameter schätzt, aber nur 81 davon für das Modell verwendet (vgl. Tabelle in Abschnitt 4).

□

Modell 9b

Eine Verbesserung dieses Verfahrens von Pedrycz (1984) nimmt Küpper (1995) vor, indem er mehrere Einschränkungen fallen läßt: Das Modell läßt mehrere Folgerungen je Regelsatz zu. Außerdem sind auch Möglichkeitsgrade im Bereich $]0; 0,1[$ zugelassen, was die Modellgüte deutlich verbessert. Weiterhin verwendet er Summe und Produkt für Disjunktion und Konjunktion. Dies erfordert, da der Summenoperator eine Randbedingung für s-Normen verletzt, eine zusätzliche Skalierung der Relationalmatrix, die aber ohne Auswirkungen auf den Ausgang bleibt (Küpper 1995). Zudem benutzt er über den Definitionsbereich der Fuzzy-Größen gleichförmig verteilte dreiecksförmige Fuzzy-Referenzmengen.

□

Modell 10: Relational-Modell (Xu und Lu 1987), Xu und Lu + (Küpper 1995)

Modell 10a

Während das Verfahren von Pedrycz (1984) keine Optimierung enthält, geben Xu und Lu (1987; auch Küpper 1995) eine Identifikationsmethode mit einer iterativ ablaufenden Optimierung an. Sie besteht aus den drei Schritten

1. Bestimmung des Startmodells,
2. Auswertung des Fuzzy-Modells und
3. Anwendung des Lernverfahrens,

wobei die Schritte 2 und 3 iterativ ablaufen.

Zur Ermittlung des Startmodells wird das Verfahren von Pedrycz (1984) verwendet. Anschließend werden im 2. Schritt zuerst die Partialprämissen der Fuzzy-Referenzmengen gesucht, deren Erfülltheitsgrade für die Eingangswerte des aktuellen Meßdatensatzes eine gewisse Schwelle $q \in]0; 1[$ überschreiten. Nur diese Regeln werden zur Modellauswertung herangezogen. Bei der Berechnung der scharfen Ausgangsgröße werden allerdings die Erfülltheitsgrade der Prämissen nicht berücksichtigt. Die scharfe Ausgangsgröße berechnet sich ausschließlich aus den Folgerungsteilen der Regeln.

Im dritten Schritt erfolgt, wenn der Prädiktionsfehler $\epsilon(k) = y(k) - \hat{y}(k)$ eine festzulegende Grenze überschreitet, eine Aktualisierung der Elemente der Relationalmatrix \mathbf{R} . Hierbei wird ein Lernfaktor verwendet, der sich aus drei Komponenten, nämlich dem Prädiktionsfehler, einem Gewichtungsfaktor und einem zusätzlichen Skalierungsfaktor, zusammensetzt. Der Gewichtungsfaktor beschreibt, welchen Anteil eine jeweilige Fuzzy-Referenzmenge des Ausgangs am Gesamtergebnis besitzt, und bestimmt sich aus der jeweiligen Fuzzy-Referenzmenge des Ausgangs und ihrem Möglichkeitsgrad. Der Lernvorgang erfolgt nach jedem Meßdatensatz, so daß auch eine On-line-Implementierung einfach möglich ist.

Für die Box-Jenkins-Daten verwenden Xu und Lu (1987) ein Modell

$$Y(k) = Y(k-1) \circ U(k-4) \circ \mathbf{R} \quad (3.28)$$

mit jeweils fünf Glockenkurven als Fuzzy-Referenzmengen für jede Fuzzy-Größe.

□

Modell 10b

Bei einer Auswertung des Fuzzy-Modells nach Xu und Lu (1987) werden viele Informationen vernachlässigt bzw. fallen gelassen. Küpper (1995) verbessert dieses Verfahren, indem er den Erfülltheitsgrad der Prämissen bei der Auswertung mit berücksichtigt. Dadurch wird eine deutlich komplexere Form des Kennfeldes des Fuzzy-Modells möglich. Bei den von ihm verwendeten dreiecksförmigen Fuzzy-Referenzmengen ist auch eine Verwendung von Schranken – zur Vermeidung von starken Überschneidungen der Fuzzy-Referenzmengen – für die Erfülltheitsgrade nicht sinnvoll. Zur Auswertung verwendet Küpper (1995) wieder die Sum-Prod-Komposition. Für die Ermittlung des Gewichtungsfaktors im Lernschritt gibt er zudem eine andere Berechnungsvorschrift an, die eine feinere Anpassung der Möglichkeitsgrade bewirkt. Hier gehen dann auch die Erfülltheitsgrade der Prämissen differenzierter ein.

□

Modell 11: Funktional-Modell (Sugeno und Tanaka 1991)

Sugeno und Tanaka (1991) verwenden ein Fuzzy-Funktional-Modell nach Takagi und Sugeno (1985) mit

$$R_i : \text{ WENN } (X_1 \text{ IST } A_{1,i}) \text{ UND } \dots \text{ UND } (X_n \text{ IST } A_{n,i}) \\ \text{ DANN } y_i = c_{0,i} + c_{1,i} x_1 + \dots + c_{n,i} x_n \quad , \quad (3.29)$$

wobei R_i die i -te Regel, $A_{j,i}$ die Fuzzy-Referenzmengen, $c_{j,i}$ die Funktionsparameter, y_i den jeweiligen Regelausgang y_i und die x_j ($j = 1, \dots, n$) die Modelleingänge bezeichnet. Die Identifikation geschieht in zwei Stufen. In der ersten Stufe werden die Modellstruktur und ein Initialmodell bestimmt, in der zweiten dann das eigentliche Fuzzy-Modell.

Zur Bestimmung des Initialmodells verwenden Sugeno und Tanaka (1991) das Verfahren von Sugeno und Kang (1988):

Hierbei wird jeweils für den Prämissen- oder Voraussetzungsteil und den Konsequenz- oder Folgerungsteil eine Struktur- und eine Parameteridentifikation vorgenommen: Die Regelprämissen folgen direkt aus der Fuzzy-Partitionierung des Eingangsraumes, die wiederum direkt mit der Anzahl der notwendigen Folgerungsteile verbunden ist. Zur Parameterschätzung für den Folgerungsteil findet ein Least-Squares-Verfahren Anwendung, da dieser linear in den Parametern ist. Für die Berechnung der Parameter des Voraussetzungsteils (die Parameter der Zugehörigkeitsfunktionen) optimieren Sugeno und Kang (1988) einen Güteindex mittels eines nichtlinearen Schätzalgorithmus.

Der eigentliche Identifikationsalgorithmus von Sugeno und Tanaka (1991) besteht aus zwei „Ebenen“: einer Überwachungs- und einer Einstellebene.

Die Überwachungsebene gibt mit sog. Fuzzy-Anpaß-Regeln (FAR) ein Konzept für die Parameteranpassung vor. Ausgehend von einer vorliegenden Fuzzy-Partitionierung (FP1) wird eine harte Partitionierung abgeleitet, indem zwischen Kern- und Flankenbereichen differenziert wird. Zur harten Partitionierung des Eingangsraums werden dann die Grundregeln aufgestellt. Die Idee der Grundregeln ist, in Kernbereichen die Konklusionsparameter und in den Flankenbereichen die Prämissenparameter nachzustellen. Diese Grundregeln werden dann zu den Fuzzy-Anpaß-Regeln erweitert, indem die harte Partitionierung in eine andere Fuzzy-Partitionierung überführt wird. Motivation hierzu ist die Berücksichtigung von Störungseinflüssen. Hierbei wird auch die Partitionierung des Eingangsraumes angepaßt. Die zweite Fuzzy-Partitionierung unterscheidet sich also in Anzahl und Lage der Partitionen von der Fuzzy-Partitionierung FP1.

In der Anpaßebene werden die Parameter entsprechend den Fuzzy-Anpaß-Regeln eingestellt. Hierbei wird zur Parameterschätzung sowohl für den Voraussetzungs- als auch für den Folgerungsteil ein gewichteter rekursiver Least-Squares-Algorithmus (WRLS) herangezogen. Während die Schätzung für den Folgerungsteil einfach ist, gestaltet sich die Be-

stimmung der Parameter der Zugehörigkeitsfunktionen $A_{j,i}$ problematisch, so daß Sugeno und Tanaka (1991) eine Restriktion einführen: Bei den trapezförmigen Zugehörigkeitsfunktionen besitzen die Flanken dieselben Stützpunkte, so daß hierfür also nur noch die Hälfte der Parameter zu bestimmen ist. Die Parameterschätzung läuft in 4 Schritten ab: Im ersten Schritt erfolgt die Schätzung für den Folgerungsteil mit dem WRLS. Im zweiten Schritt wird dann anhand eines Kriteriums geprüft, ob eine Nachstellung der Zugehörigkeitsfunktionen überhaupt möglich ist oder nicht. Ist sie nicht möglich, fallen die nächsten beiden Schritte weg. Zur Schätzung der Prämissenparameter müssen in Schritt 3 in einem Zwischenschritt aus den Ein- und Ausgangsdaten sog. Prämissendaten berechnet werden, da die Ein-/Ausgangsdaten nicht direkt zur Anpassung der Parameter der Fuzzy-Referenzmengen verwendet werden können. Im 4. Schritt findet hiermit die Nachstellung der Prämissenparameter der Fuzzy-Referenzmengen $A_{j,i}$ (also der Zugehörigkeitsfunktionen) mit dem WRLS statt.

Im Laufe von Untersuchungen hat sich herausgestellt (was auch durch brieflichen Kontakt mit den Entwicklern des Verfahrens bestätigt wurde), daß die allgemeine Anwendung dieses Verfahrens zumindest sehr problematisch ist (Hebisch 1992).

Für die Box-Jenkins-Daten geben Sugeno und Tanaka (1991) zwei Modelle an:

Modell 11a

mit 2 Regeln, 2 Zugehörigkeitsfunktionen, den Eingängen $u(k-1)$ und $y(k-1)$ (also 6 Parametern) und eine komplexeres

Modell 11b

mit ebenfalls 2 Regeln und 2 eindimensionalen Zugehörigkeitsfunktionen aber 6 (für die erste Regel) bzw. 7 (für die zweite) Eingangsgrößen $y(k-1)$, $y(k-2)$, $y(k-3)$, $u(k)$, $u(k-1)$, $u(k-2)$ und einem Gleichanteil C , wobei dann 15 Modellparameter zu bestimmen sind.

□

Modell 12: Positions-Gradienten-Modell (Sugeno und Yasukawa 1993)

Das bei weitem größte Problem bei der Identifikation liegt nicht in der Parameterschätzung – diese stellt mittlerweile aufgrund der bekannten mathematischen Schätzalgorithmen eine „eher einfache“ Aufgabe dar – sondern in der Vorgabe einer „richtigen“ bzw. geeigneten Modellstruktur: Für Fuzzy-Modelle bedeutet das, die dominierenden Modelleingänge zu finden, die Anzahl der wichtigen Regeln festzulegen und die geeigneten Fuzzy-Referenzmengen zu bestimmen. Hier setzt auch der iterative Algorithmus von Sugeno und

Yasukawa (1993) an, der aus den drei Hauptteilen

1. Aufteilung des Ausgangsraumes durch Clusterung,
2. Bestimmung geeigneter Modelleingänge und
3. Parameterschätzung

besteht. Im Gegensatz zu den meisten anderen bekannten Struktursuchverfahren (wie z. B. dem Strukturselektionsalgorithmus aus (Kortmann 1989)) sind bei dieser Methode Strukturidentifikation und Parameterschätzung nicht aneinander gekoppelt, sondern können auch getrennt und unabhängig voneinander durchgeführt werden (Sugeno und Yasukawa 1993).

Zur Clusterung wird ein FCM-Algorithmus verwendet, als Ergebnis wird jeder Ausgang y mit einem Zugehörigkeitsgrad zu einem Fuzzy-Cluster versehen. Die Anzahl der Cluster entspricht hierbei aber nicht immer der Anzahl der Fuzzy-Regeln, da Sugeno und Yasukawa (1993) Fälle angeben, für die ein Cluster im Eingangsraum in zwei Fuzzy-Cluster (und damit 2 Regeln) einzuteilen ist. Die Parameter aus der Fuzzy-Clusterung sind als Startwerte für die spätere Parameterschätzung sehr geeignet. Dazu werden die mehrdimensionalen Cluster auf eindimensionale Fuzzy-Mengen projiziert. Zum Erhalt einer optimalen Anzahl von Clustern/Regeln wird die Clusterung für eine steigende Anzahl von Clustern durchgeführt. Die optimale Clusteranzahl ist erreicht, wenn ein Gütekriterium, das sowohl die Varianz der Daten im jeweiligen Cluster als auch die Varianz der Daten für die Cluster untereinander verwendet, ein Minimum erreicht.

Zur Bestimmung der geeigneten Modelleingänge werden alle möglichen Modelleingänge stufenweise und aufeinander aufbauend miteinander gekoppelt. Die Auswahl der „richtigen“ Eingänge erfolgt dann mittels eines Güteindizes als Bewertungskriterium. Anschließend muß entschieden werden, ob als Fuzzy-Modell ein sog. Positions-Modell oder ein Positions-Gradienten-Modell ermittelt werden soll. Diese beiden Modelle unterscheiden sich dadurch, daß beim Positions-Gradienten-Modell zusätzlich zu einem absoluten Ausgangswert noch eine Gradienten- oder Steigungs-Variable ausgegeben wird:

$$R_i : \quad \text{WENN } (X_j \text{ IST } A_{i,j}) \quad \text{DANN } (y = B_i) \quad \text{UND } (\partial y / \partial x_j = C_{i,j}) \quad . \quad (3.30)$$

Die Ableitungen $\partial y / \partial x_j$ dienen, wenn beispielsweise infolge zu weniger Daten kein Fuzzy-Modell für den gesamten Eingangsraum erstellt werden kann, zur Extrapolation der Schätzung des Ausgangssignals über die lokalen Fuzzy-Regeln hinaus.

Bei der Parameteridentifikation werden die jeweils 4 Parameter der trapezförmigen Zugehörigkeitsfunktionen ggf. in immer kleiner werdenden Schritten bis zu einem Abbruchkriterium optimiert. Die Bewertung erfolgt hier entsprechend dem mittleren quadratischen

Fehler.

Bei den Box-Jenkins-Daten finden Sugeno und Yasukawa (1993) 6 Cluster, als Eingänge für das Positions-Gradienten-Modell $y(k-1)$, $u(k-4)$ und $u(k-3)$. Ihr Fuzzy-Modell besteht aus 6 Regeln, so daß $6 \cdot 3 \cdot 4 + 24 = 96$ Parameter für die Lage der Zugehörigkeitsfunktionen und die Modell-Ausgänge $y(k)$, $\partial y / \partial y(k-1)$, $\partial y / \partial u(k-4)$ und $\partial y / \partial u(k-3)$ zu optimieren sind.

□

Modell 13: Relational-Modell (Küpper 1995)

Ein weiteres Identifikationsverfahren für Fuzzy-Modelle schlägt Küpper (1994, 1995) vor. Für eine spezielle Form der Fuzzy-Relational-Modelle nach Pedrycz (1984) benutzt er zur Optimierung ein Gradientenverfahren, das auf dem der stochastischen Approximation (Tsytkin 1971, Ljung und Söderström 1987) beruht. Außerdem zeigt Küpper (1995), daß bei einer Fuzzifizierung über Fuzzy-Einermengen, einer Defuzzifizierung über die Summenmethode und mit der sum-prod-Komposition als Kompositionsoperator ein Fuzzy-Relational-System mit 2 symmetrischen dreiecksförmigen Fuzzy-Referenzmengen im Intervall $[-1; 1]$ mit Kern -1 bzw. 1 und Träger 2 das gleiche Ein-/Ausgangsverhalten besitzt wie ein Fuzzy-System mit beliebiger Zahl, Form und Anordnung der Fuzzy-Referenzmengen des Ausgangs. Zur Optimierung verwendet er ein (analytisches) quadratisches Gütekriterium für das Fuzzy-System mit der Relationalgleichung

$$Y = X_1 \circ X_2 \circ \dots \circ X_a \circ \mathbf{R} \quad (3.31)$$

und Regelsätzen der Form

$$\begin{aligned} \text{WENN } & (X_1 \text{ IST } X_{1,p_1}) \wedge (X_2 \text{ IST } X_{2,p_2}) \wedge \dots \wedge (X_a \text{ IST } X_{a,p_a}) \\ \text{DANN } & (Y_{1,l} = Y_1) \Big|_{g_{1,l}} \vee (Y_{2,l} = Y_2) \Big|_{g_{2,l}} \quad , \end{aligned} \quad (3.32)$$

mit

$$\begin{aligned} i &= 1, 2, \dots, a && (a = \text{Anzahl der Eingänge des Fuzzy-Systems}) \quad , \\ j &= 1, 2, \dots, b && (\text{hier } b = 2 = \text{Anzahl der Fuzzy-Referenzmengen für } Y) \quad , \\ p_i &= 1, 2, \dots, c_i && (c_i = \text{Anzahl der Fuzzy-Referenzmengen für die } X_i) \quad , \\ l &= 1, 2, \dots, d && (d = \text{Anzahl der Regelsätze}) \quad , \end{aligned}$$

wobei $g_{j,l}$ der Möglichkeitsgrad der Regel des Regelsatzes l mit der Fuzzy-Referenzmenge Y_j für den Ausgang, die Variable X_{i,p_i} die p_i -te Fuzzy-Referenzmenge der i -ten Fuzzy-Variablen ist und als Fuzzy-Zahl X_i für einen konkreten Eingangswert interpretiert wird.

Hierbei werden die Möglichkeitsgrade bzw. Gewichte g_{1l} und g_{2l} der l Regelsätze im Parametervektor Θ zusammengefaßt und mit

$$\Theta(k+1) = \Theta(k) + \gamma(k) (y(k) - \hat{y}(k)) \frac{\partial \hat{y}(k)}{\partial \Theta} \quad (3.33)$$

on-line geschätzt. Dies ist möglich, da Küpper (1994, 1995) eine geschlossene Beziehung $\hat{y} = f(\alpha_l, g_{j,l})$ zwischen der scharfen Ausgangsgröße \hat{y} und den Parametern des Fuzzy-Systems angibt ($\alpha_l =$ Erfülltheitsgrad der Prämisse des l -ten Regelsatzes). Diese Beziehung ist jedoch nicht linear in ihren Parametern, so daß ein nichtlineares Schätzproblem vorliegt und sog. „lineare“ Identifikationsalgorithmen wie z. B. die Least-Squares-Verfahren nicht anwendbar sind. Der Lernfaktor $\gamma(k)$ stellt einen wichtigen Parameter zur Beeinflussung des Konvergenzverhaltens der Schätzung dar (ähnlich wie der Vergessensfaktor $\lambda(k)$ bei den meisten konventionellen rekursiven Identifikationsalgorithmen; für konstante Werte gilt $\gamma_0 = 1 - \lambda_0$) (Ljung und Söderström 1987).

Küpper schlägt weiterhin eine Brücke zu den Neuro-Fuzzy-Systemen, indem er zeigt, daß sein Identifikationsverfahren der stochastischen Approximation und der Backpropagation-Algorithmus aus dem Bereich der Neuronalen Netze ineinander überführbar sind.

Auch Küpper (1994, 1995) verwendet bei den Box-Jenkins-Daten die Struktur

$$Y(k) = Y(k-1) \circ U(k-4) \circ \mathbf{R} \quad (3.34)$$

mit 5 dreiecksförmigen Fuzzy-Referenzmengen für jeden Eingang und 2 für den Ausgang. Zudem untersucht er an diesem Beispiel den Einfluß der Startwerte von \mathbf{R} und von $\gamma(k)$ auf Modellgüte und Konvergenzgeschwindigkeit (Küpper 1995). □

Modell 14: Arithmetik-Modell (Bertram 1995)

Eine völlig neue Darstellungsart für Fuzzy-Systeme führt Bertram (1995) mit den Fuzzy-Arithmetik-Modellen ein. Es handelt sich hierbei um gewichtete Fuzzy-Parametergleichungen der Struktur

$$Y_l = g_l (X_{1,p_1} \odot M_{1,l} \oplus X_{2,p_2} \odot M_{2,l} \oplus \dots \oplus X_{a,p_a} \odot M_{a,l}) \quad (3.35)$$

mit

$$\begin{aligned} i &= 1, 2, \dots, a && (a = \text{Anzahl der Eingänge des Fuzzy-Systems}) \quad , \\ p_i &= 1, 2, \dots, c_i && (c_i = \text{Anzahl der Fuzzy-Referenzmengen für die } X_i) \quad , \\ l &= 1, 2, \dots, d && (d = \text{Anzahl der Regelsätze}) \quad , \end{aligned}$$

wobei g_l den Möglichkeitsgrad des Regelsatzes l des Regelsatzes l mit der Fuzzy-Referenzmenge Y_l für den Ausgang, die Variable X_{i,p_i} die p_i -te Fuzzy-Referenzmenge der i -ten Fuzzy-Variablen ist und als Fuzzy-Zahl X_i für einen konkreten Eingangswert interpretiert wird. Die Variablen $M_{i,l}$ sind Parameter (Fuzzy-Zahlen), die die entsprechenden Fuzzy-Referenzmengen einer Fuzzy-Variablen mittels einer Modifikation von Kern und Träger der Fuzzy-Menge (Bertram u. a. 1994) über den ganzen Definitionsbereich gewichten. Zudem werden hier spezielle Fuzzy-Rechen-Operatoren \odot und \oplus (z. B. Dubois und Prade 1978, 1979, 1980) verwendet. Somit sind d Parameter für die g_l und $4 \cdot d \cdot a$ Parameter für

die $M_{i,l}$ zu identifizieren (Bertram 1995).

Dieser Fuzzy-Arithmetik-Ansatz besitzt u. a. folgende Vorteile (Bertram 1995):

Die subjektiven Einflüsse bei der Synthese der Zugehörigkeitsfunktionen werden minimiert (sie lassen sich über die Fuzzy-Parameter-Gleichungen berechnen), es liegt eine mathematisch einfach auszuwertende Beschreibungsform vor, es können mathematische Gütekriterien zur Optimierung verwendet werden, im Vergleich zu Fuzzy-Relational-Modellen liegen deutlich weniger (festzulegende oder zu bestimmende) Parameter vor, es kann zudem axiomatisches und heuristisches Wissen in die Fuzzy-Arithmetik-Gleichungen eingebunden werden. Unter anderem auf Grund ihres begrenzten Trägers sollten aber keine dreiecksförmigen Zugehörigkeitsfunktionen sondern glockenkurvenförmige verwendet werden, wodurch eine stärkere Nuancierung der Meßwerte möglich wird.

Bei der Identifikation eines Fuzzy-Arithmetik-Modells werden die Fuzzy-Zahlen im Sinne eines Gütekriteriums zur Gewichtung der Fuzzy-Eingangswerte und die Gleichungsgewichte optimiert. Bertram (1995) greift hierbei nicht auf eine analytische – dazu müßte die Zielfunktion explizit als Funktion der Entwurfsparameter vorliegen und noch spezielle andere Bedingungen erfüllen – sondern eine numerische Optimierung zurück, bei der beliebige Gütefunktionen verwendet werden können, sofern diese programmierbar sind. Außerdem lassen sich hierbei beliebig viele Randbedingungen für die Problemlösung in die Optimierung integrieren. Im Gegensatz zu den häufig verwendeten deterministischen Gradientenverfahren zieht Bertram (1995) ein stochastisches Optimierungsverfahren – eine Evolutionsstrategie (Rechenberg 1973) mit Mutation, Rekombination und Selektion – heran.

Zur Fuzzy-Identifikation der Box-Jenkins-Daten verwendet auch Bertram (1995) die Variablen $Y(k-1)$ und $U(k-4)$ als Modelleingang und erhält sein Ergebnis nach 5000 Iterationsschritten.

□

Modell 15: Funktional-Modell (Kroll 1994, 1995b)

Kroll (1994, 1995b) verwendet zur Identifikation von funktionalen Fuzzy-Modellen eine zweistufige Methode. In der ersten Stufe wird ein Startmodell geschätzt, für dessen Parameter in der zweiten Stufe eine Nachoptimierung erfolgt. Es hat sich hierbei an mehreren darunter auch industriellen Anwendungsbeispielen gezeigt, daß die Startmodelle schon so gut sind, daß sich in der Regel durch die Nachoptimierung nur noch eine geringfügige Verbesserung der Modellgüte ergibt.

Die hierbei verwendeten Fuzzy-Modelle bestehen aus c funktionalen Fuzzy-Regeln mit der

Struktur

$$\begin{aligned}
 R_i : \text{ WENN } & (\mathbf{x}(k) \text{ IST } \mathbf{v}_i) \\
 \text{ DANN } & y_i(\mathbf{x}(k)) = f_i(\mathbf{x}(k)) \quad , \quad i = 1, \dots, c \quad .
 \end{aligned}
 \tag{3.36}$$

Hierbei werden mehrdimensionale Fuzzy-Referenzmengen verwendet, die den Eingangsgrößenraum in c unscharfe Bereiche aufteilen (sog. Fuzzy- c -Partitionierung). In der Prämisse von Gl. (3.36) bezeichnet $\mathbf{x}(k)$ den scharfen Eingangsgrößenvektor und \mathbf{v}_i den i -ten Clusterschwerpunkt, der durch die Anwendung von Fuzzy-Clusterverfahren berechnet wird. Zuletzt werden die c scharfen Ausgangsgrößen y_i der c Regeln des Fuzzy-Systems mit den Erfülltheitsgraden der Prämissen gewichtet und zum resultierenden Ausgangswert gemittelt. Als Konklusionsfunktionen werden sehr häufig Polynome, die linear in den Parametern sind verwendet.

Modell 15a,b (1. Stufe)

Bei dieser Methode erfolgt zuerst die Identifikation der Zugehörigkeitsfunktionen durch Clusterung. Ein einfach anzuwendendes Verfahren hierfür ist der Fuzzy- c -Means-Algorithmus (FCM). Bezdek (1981) gibt auch andere Verfahren an. Der verwendete Zugehörigkeitsfunktionstyp folgt praktisch aus der Anwendung metrischer Fuzzy-Clusterverfahren mit Zielfunktion (Optimierungskriterium). Auch besitzt hierbei die festzulegende und zu verwendende Abstandsnorm einen wesentlichen Einfluß auf die Partitionierung, wobei eine geeignete Norm unter vielen möglichen auszuwählen ist. Kroll (1995a) verwendet u. a. mit der Mahalanobisnorm eine Norm, die die statistischen Eigenschaften der Meßdaten bewertet. Nach Abschluß der Identifikation der Partitionierung erfolgt die Schätzung der Parameter der Konklusionspolynome mittels eines gewichteten rekursiven Least-Squares-Verfahrens (GRLS). Da die Gewichtung des verwendeten GRLS-Verfahrens auch auf die Meßdaten umgerechnet werden kann, ist zudem jedes beliebige „ungewichtete“, d. h. praktische jedes andere Parameterschätzverfahren hiernach anwendbar.

Für die Box-Jenkins-Daten werden mit diesem Verfahren zwei Modelle identifiziert: Zum einen ein einfaches **Modell 15a** mit

$$y(k) = f(y(k-1), u(k-3)) \tag{3.37}$$

und zum anderen ein **Modell 15b** mit

$$y(k) = f(y(k-1), y(k-2), y(k-3), u(k-1), u(k-2), u(k-3)) \quad . \tag{3.38}$$

□

Modell 15c (2. Stufe)

In der zweiten Stufe der Fuzzy-Identifikation optimiert Kroll (1995a) die Parameter des in der ersten Stufe geschätzten Modells nach. Dies geschieht mit Hilfe eines iterativen

Gradientenverfahrens 1. Ordnung (Ljung und Söderström 1987, Ljung 1987), wobei ein quadratisches Gütefunktional analytisch minimiert wird. Alternativ kann als zweite Variante ein Hybridverfahren verwendet werden, bei dem zuerst mit dem Gradientenverfahren die Parameter für die Partitionierung – also für die Prämissen – und anschließend mit einem Least-Squares-Verfahren die Parameter des Konklusionsteils berechnet werden.

In der zweiten Stufe wurde das Hybrid-Verfahren mit der Mahalanobis-Norm auf das Modell (3.37) angewendet. Bemerkenswert an diesem Ergebnis ist, daß zwar die Modellgüte bzgl. der prädiktiven Auswertung verbessert wird, bei rekursiver Modellauswertung jedoch eine Modellverschlechterung eintritt. Grund hierfür ist, daß bei diesen Verfahren eine Optimierung nur bezüglich der Einschnittprädiktion durchgeführt wird (Kroll 1995a).

□

Modell 16: Fuzzifiziertes Polynommodell (Kroll u. a. 1995)

Die aus WENN–DANN-Produktionsregeln bestehenden Fuzzy-Modelle können durch *Verbalisierung* (natürlichsprachliche Erfahrungsregeln menschlicher Experten), *Fuzzifizierung* (Umsetzung einer scharfen mathematischen Beschreibung in ein unscharfes Fuzzy-Modell) oder *Identifikation* (wie bei allen anderen im Rahmen dieses Berichtes erwähnten Verfahren) bestimmt werden. Da das größte Problem der experimentellen Prozeßanalyse in der Bestimmung einer geeigneten Modellstruktur liegt (und diese bei den Box-Jenkins-Daten auch nicht bekannt ist), ist zur Verwendung als Modellstruktur das Ergebnis eines konventionellen Struktursuchalgorithmus sinnvoll. Nachdem nun die Struktur eines parametrischen Modells vorliegt, wird dieses Modell fuzzifiziert (Kroll, Reuter und Hebisch 1995) in dem Sinne, daß die nunmehr c Konklusionspolynome anstatt von einem die gleichen Argumente aufweisen. Hierfür werden jetzt die Koeffizienten geschätzt. Unangenehm kann sich allerdings die aufgrund der vielen Modelleingänge ggf. relativ hohe Komplexität des Fuzzy-Modells bemerkbar machen.

Bei dieser Vorgehensweise handelt es sich um eine zweistufige „gemischte“, also hybride Identifikation in dem Sinne, daß eine konventionelle Identifikationsmethode (Strukturselektionsalgorithmus und Polynommodell) und ein Fuzzy-Identifikationsverfahren gemeinsam herangezogen werden. Diese von Kroll stammende Idee wurde in (Kroll, Reuter und Hebisch 1995) erstmals skizziert.

Für die Box-Jenkins-Daten wird zur Fuzzifizierung das aus dem Strukturselektionsalgorithmus nach Kortmann (1989) erhaltene Polynommodell (3.22) verwendet. Hierbei zeigt sich, daß eine signifikante Verbesserung der Modellgüte erzielt werden kann.

□

4 Vergleich der Ergebnisse und Bewertung

In Tabelle 4.1 sind für die im vorigen Abschnitt kurz skizzierten Identifikationsverfahren und Modelle die mittleren quadratischen Fehler

$$V = \frac{1}{N - n - d} \sum_{k=1+d+n}^N \epsilon^2(k) \quad (4.1)$$

für prädiktive und (soweit vorhanden) rekursive Modellauswertung angegeben².

Modell-Nr.: Modelltyp: Verfahren	Anz. Regeln /Parameter	Auswertung		eig. Erg.
		prädiktiv	rekursiv	
1: LS: Box und Jenkins (1976)	-/4	0,1970	0,7180	ja
2: LS: PEM-MATLAB (Ljung 1988)	-/10	0,0785	0,6964	ja
3a: BLS-OYCF: PEM (Reuter 1995)	-/7	0,0661	0,7128	ja
3b: BLS-ORCF: PEM (Reuter 1995)	-/9	0,0913	0,7070	ja
4: Hammerstein: PEM (Reuter 1995)	-/13	0,0621	0,6834	ja
5: Wiener: PEM (Reuter 1995)	-/15	-	0,6029	ja
6: Kolmogorov-Gabor-Polynom: Kortmann (1989)	-/11	0,0519	0,6244	ja
7: $\hat{y}(k) = y(k-1)$: -	-/-	0,5582	-	ja
8: Relational: Tong (1978)	19/-	0,469	-	
9a: Relational: Pedrycz (1984)	81/729	0,3200	-	
9b: Relational: Pedrycz + Küpper (1995)	81/729	0,1754	0,8502	ja
10a: Relational: Xu und Lu (1987)	25/125	0,3280	-	
10b: Relational: Xu und Lu + Küpper (1995)	25/125	0,2541	-	ja
11a: Funktional: Sugeno und Tanaka (1991)	2/6	0,359	-	
11b: Funktional: Sugeno und Tanaka (1991)	2/15	0,068	-	
12: Positions-Gradienten-Modell: Sugeno und Yasukawa (1993)	6/96	0,190	-	
13: Relational: Küpper (1995)	25/50	0,1199	0,8149	ja
14: Arithmetik: Bertram (1995)	25/225	0,1170	-	
15a: Funktional: Kroll (1994) (1. Stufe)	2/10	0,1374	0,8585	ja
15b: Funktional: Kroll (1994) (1. Stufe)	2/36	0,0556	-	ja
15c: Funktional: Kroll (1995a) (2. Stufe)	2/10	0,1170	0,9316	ja
16: Konventionell und funktional gemischt: Kroll (Kroll, Reuter und Hebisch 1995)	8/168	0,0317	0,5641	ja

Tabelle 4.1: Modellgröße und Gütwerte V der berücksichtigten Modelle

²Der Autor dankt Herrn Dipl.-Ing. A. Kroll für seine Hilfe bei der Berechnung der Ergebnisse für die Modelle 15 und 16 sowie Herrn Dipl.-Ing. K. Küpper für die Ergebnisse der Modelle 9b und 10b.

Dabei ergeben sich bei der Identifikation der Box-Jenkins-Daten für die *prädiktive* Modellauswertung folgende Aussagen:

Wie in Tabelle 4.1 anhand der Regel- bzw. Parameteranzahl zu erkennen ist, benötigen die konventionellen Modelle zum Erreichen einer ähnlich guten Modellgüte wie die Fuzzy-Modelle deutlich weniger Parameter, wenngleich die Parameter der Fuzzy-Modelle signifikant nur lokal wirken. Dies trifft auch auf gaußförmige Zugehörigkeitsfunktionen zu, die zwar einen unendlich großen Träger besitzen, aber Zugehörigkeitsgrade deutlich größer als Null nur in der Nähe des Kerns aufweisen. Nichtsdestotrotz müssen diese vielen Parameter natürlich auch identifiziert werden, was ggf. eine entsprechend hohe Anzahl zur Verfügung stehender geeigneter Meßwerte voraussetzt.

Eine alleinige Betrachtung der Modellgüte für dieses spezielle Anwendungsbeispiel läßt leicht den (einseitigen) Eindruck entstehen, daß die konventionellen Verfahren (Modell 1 bis 6) im allgemeinen besser seien. Dies trifft für die Box-Jenkins-Daten in der Regel zwar auch zu, doch der Hauptgrund dürfte in diesem Fall darin zu suchen sein, daß sich diese Meßdaten schon sehr gut durch ein lineares Modell approximieren lassen und der Gasofenprozeß während des Großteils des betrachteten Zeitraumes annähernd lineares Verhalten besitzt. Die besonders bei den letzten ca. 35 Meßdatentupeln auftretenden Abweichungen sind wohl auf Auswirkungen unbekannter zusätzlicher Eingangs- bzw. Störgrößen zurückzuführen, da sie bei allen betrachteten Modelltypen in ähnlichem Maße auftreten.

Bei genauerer Betrachtung der Ergebnisse ergibt sich für die Fuzzy-Modelle ein differenzierteres Bild. Insbesondere die relationalen Fuzzy-Modelle besitzen sehr viele (zu bestimmende) Parameter – teilweise um 1 bis 2 Größenordnungen mehr als die konventionellen Modelle –, ohne daß sich dies in der Modellgüte niederschlägt. Zudem scheinen die fortschrittlicheren Verfahren für diese Modelle (nach Erfahrungen des Autors) sehr empfindlich gegenüber ungünstigen Startwerten zu sein (z. B. auch bei ungünstig gewählter oder komplexer Modellstruktur). Das identifizierte arithmetische Fuzzy-Modell besitzt ebenfalls relativ viele Parameter und liegt auch qualitativ im Bereich des besten relationalen Modells (vgl. Modell 13 und 14).

Für die Identifikation besonders geeignet zu sein scheinen jedoch die funktionalen Fuzzy-Modelle. Sie benötigen für Fuzzy-Modelle relativ wenig Regeln und Parameter für eine gute Approximation und erfüllen somit eine sehr wichtige Erfahrungsregel bei der Identifikation, nach der im Zweifelsfalle (d. h. vor allem bei ähnlicher Modellgüte) immer das weniger komplexe und einfachere Modell vorgezogen werden sollte (Ljung 1987, Reuter 1995). Zudem liegen die funktionalen Fuzzy-Modelle deutlich näher an der bewährten konventionellen Modelldarstellung als die Fuzzy-Modelle mit rein linguistischer Beschreibung. Sie stellen in dem Sinne eine vorteilhafte Verbindung von konventionellem Modell und linguistischem Fuzzy-Modell dar, da sie wegen ihrer semilinguistischen Darstellung

eine linguistische Beschreibung im Prämissenteil und häufig eine konventionelle Polynomdarstellung im Konklusionsteil besitzen.

Die Vorteile der funktionalen Fuzzy-Modelle mit ihrer relativ guten Approximationsfähigkeit gegenüber den in ihrer Modellstruktur relativ starren, konventionellen Modellen zeigen sich bei „weniger gutmütigen“ Anwendungen wie bei der Identifikation eines pneumatischen Translationsantriebes (Kroll, Reuter und Hebisch 1995). Infolge von sehr hohen Reibungsanteilen (bis zu 50 % der maximal möglichen Stellgröße) und anderen nichtlinearen Effekten war für dieses komplexe System keine geeignete Modellstruktur bekannt. Gute Ergebnisse ergaben sich bereits mit der ersten Stufe des Verfahrens von Kroll (1994) analog zu Modell 15a,b. Mit einem konventionellen Polynommodell nach Kortmann (1989) ergeben sich je nach Verwendungszweck des Modells gute oder schlechte Ergebnisse. Zwar besitzt das Polynommodell bei prädiktiver Auswertung eine etwas bessere Modellgüte als das identifizierte funktionale Fuzzy-Modell, es versagt aber bei der rekursiven Auswertung. Andere durchgeführte Identifikationen mit Modellen gemäß Nr. 2 bis 5 erbrachten deutlich schlechtere Ergebnisse.

Erfahrungen bzgl. der *rekursiven* Modellauswertung zeigen, daß hierbei in der Regel auch funktionale Fuzzy-Modelle nicht so gut abschneiden wie geeignete konventionelle Modelle. Der Grund hierfür liegt darin, daß bei der Fuzzy-Identifikation bei fast allen Verfahren (wenn überhaupt) eine Optimierung bzgl. der Einschnittprädiktion vorgenommen wird (Kroll 1995a). Dieser Nachteil spricht jedoch nicht prinzipiell zu Ungunsten der Fuzzy-Identifikation, da die Stärken der funktionalen Fuzzy-Identifikation vor allem dort zum Tragen kommen, wo eine Identifikation konventioneller Modelle nicht erfolgreich durchgeführt werden kann. Dies trifft in der Praxis häufig auf relativ komplexe Systeme und Prozesse zu.

Das beste Ergebnis für die Box-Jenkins-Daten ergab sich jedoch bei der „gemischten“ Identifikation, also der gemeinsamen Verwendung eines konventionellen und eines Fuzzy-Identifikationsalgorithmus. Hierbei scheinen sich die Vorteile beider Identifikationsarten zu verbinden. Weiterführende genauere Aussagen können jedoch zum jetzigen Zeitpunkt noch nicht getroffen werden, da sich dieses Verfahren noch am Anfang seiner Entwicklung befindet.

5 Zusammenfassung und Ausblick

Zusammenfassung

In den letzten Jahren kamen in immer stärkerem Maße Anwendungen der Fuzzy-Logik im Bereich der Automatisierungstechnik zum Einsatz. Ein wichtiger Bereich in der Regelungstechnik neben der Fuzzy-Regelung ist die sog. Fuzzy-Identifikation (ein Teilbereich der Fuzzy-Modellbildung), also die Berechnung von Fuzzy-Systemmodellen aus Meßdaten. Wichtig für die Auswertungsart dieser Systemmodelle ist der Anwendungszweck. Hierbei kann es sich um eine prädiktive Modellauswertung, wie sie beispielsweise bei adaptiven Regelungen erforderlich ist, oder um eine rekursive Modellauswertung, die zur Vorhersage weiter in der Zukunft liegender Ereignisse dient, handeln. Gerade bei adaptiven Regelungen ist eine Echtzeitfähigkeit unumgänglich, bei Vorhersageaufgaben zur Vermeidung unnötig hoher laufender Kosten sehr wünschenswert.

Im vorliegenden Bericht werden für unterschiedliche Verfahren Identifikationsergebnisse gegenübergestellt und verglichen. Es handelt sich hierbei sowohl um konventionelle Verfahren als auch um Fuzzy-Identifikationsmethoden unterschiedlichster Art. Solch eine Gegenüberstellung, bei der sowohl die Modellgröße (d. h. Anzahl von Regeln und/oder Parametern) als auch die Modellgüte bewertet wird, ist zur Einordnung der Vor- und Nachteile der einzelnen Verfahren wichtig und wurde in dieser Breite noch nicht vorgenommen. Dazu fand in diesem Bericht eine kurze Skizzierung der den betrachteten Verfahren zu Grunde liegenden Ideen statt. Als Anwendungsbeispiel wurden die sog. Box-Jenkins-Daten herangezogen, wobei es sich um Meßdaten eines realen Gasofenprozesses handelt. Diese Box-Jenkins-Daten haben sich in den letzten ca. 10 Jahren zu einem Benchmarktest für (Fuzzy-)Identifikationsverfahren entwickelt.

Ergebnis der Untersuchungen ist, daß von den Fuzzy-Modelltypen sich die funktionalen Fuzzy-Systeme bei weitem am besten für die Identifikation zu eignen scheinen. Die Identifikationsverfahren für funktionale Fuzzy-Modelle liefern beispielsweise im Vergleich zu relationalen Fuzzy-Modellen deutlich bessere Modellgüten bei wesentlich einfacheren Modellen und besitzen im Vergleich zu Arithmetik-Modellen deutlich weniger Parameter bei vergleichbaren Güten.

Handelt es sich bei dem zu untersuchenden Prozeß um einen „gutmütigen“ im Sinne der Identifikation, so lassen sich mit Hilfe von konventionellen Identifikationsverfahren in der Regel bessere und weniger komplexe Modelle bestimmen. Die Stärke der Identifikation funktionaler Fuzzy-Modelle liegt in den Anwendungen, bei denen konventionelle Verfahren aufgrund ihrer relativ starren Struktur und der relativ geringen Approximationsfähigkeit kein zufriedenstellendes Ergebnis liefern. Hierbei handelt es sich in der Praxis häufig um komplexe (industrielle) Prozesse und Anlagen, für die eine für konventionelle Modelle erforderliche günstige System- und/oder Modellstruktur nicht mit vertretbarem Aufwand

gefunden werden kann.

Ausblick

Das Gebiet von Fuzzy Control bietet noch viele Forschungsfelder. Direkt mit dem vorliegenden Bericht verbunden sind die folgenden Aufgaben:

- Ein wichtiger Punkt liegt in der Weiterentwicklung der Methode der „gemischten Identifikation“ analog zu Modell 16 von Kroll (Kroll u. a. 1995), bei der ein konventionelles Struktursuchverfahren und Identifikationsalgorithmen, wie sie auch für funktionale Fuzzy-Modelle Anwendung finden, zusammen herangezogen werden.
- Bei den Fuzzy-Identifikationsverfahren findet (wenn überhaupt) eine Optimierung nur bezüglich der Einschrittprädiktion statt. Im Zweifelsfalle erkaufte man sich hierbei die Verbesserung eines Fuzzy-Modells bei prädiktiver Auswertung mit seiner Verschlechterung bei der rekursiven Auswertung und umgekehrt (Kroll 1995a). Hier sollte untersucht werden, mit welchem Aufwand und Ergebnis eine Optimierung bezüglich des gesamten Meßdatensatzes erfolgen kann. Dies dürfte dann zu Fuzzy-Modellen führen, die auch für die rekursive Modellauswertung deutlich bessere Güten erreichen.

6 Literaturverzeichnis

- Behmenburg, C.** 1995. *Zur adaptiven Fuzzy-Regelung technischer Systeme*. Dissertation Gerhard-Mercator-Universität – GH Duisburg. Fortschritt-Bericht VDI Reihe 8 (zur Veröffentlichung angenommen).
- Bertram, T.** 1991. *Einführung in die Fuzzy-Regelung*. Forschungsbericht Nr. 04/91 MSRT. Universität –GH– Duisburg.
- Bertram, T.** 1995. *Zur systematischen Analyse und Synthese nichtlinearer Systeme mit Fuzzy-Logik*. Dissertation Gerhard-Mercator-Universität – GH Duisburg (zur Veröffentlichung freigegeben).
- Bertram, T., K. Küpper und H. Schwarz.** 1993. Fuzzy modelling for the control of mechatronic systems. *Proceedings of the 2nd Conference on Mechatronics and Robotics, Moers/Deutschland*. Hrsg. Hiller, M. und B. Fink. 86–106.
- Bertram, T., F. Svaricek, T. Bindel, R. Böhm, H. Kiendl, B.-M. Pfeiffer und M. Weber.** 1994. Fuzzy Control: Zusammenstellung und Beschreibung wichtiger Begriffe. *Automatisierungstechnik – at* 42. 322–326.
- Bezdek, J.C.** 1981. *Pattern recognition with fuzzy objective function algorithms*. New York, London: Plenum Press.
- Box, G.E.P. und G.M. Jenkins.** 1976. *Time Series Analysis – forecasting and control*. Überarbeitete Ausgabe (1. Ausgabe 1970). San Francisco: Holden-Day.
- Dubois, D. und H. Prade.** 1978. Operations on fuzzy numbers. *International Journal of Systems Science* 9. 613–626.
- Dubois, D. und H. Prade.** 1979. Fuzzy real algebra: Some results. *Fuzzy Sets and Systems* 2. 327–348.
- Dubois, D. und H. Prade.** 1980. *Fuzzy sets and systems: Theory and applications*. New York, London, Toronto: Academic Press.
- Hebisch, H.** 1992. *Identifikation eines Fuzzy-Modells am Beispiel einer hydraulischen Regelstrecke*. Diplomarbeit Fachgebiet Meß-, Steuer- und Regelungstechnik. Universität –GH– Duisburg.
- Isermann, R., K.-H. Lachmann und D. Matko.** 1992. *Adaptive Control Systems*. Englewood Cliffs/New Jersey: Prentice Hall.
- Kortmann, M.** 1989. *Die Identifikation nichtlinearer Ein- und Mehrgrößensysteme auf der Basis nichtlinearer Modellansätze*. Dissertation Ruhr-Universität Bochum. Fortschritt-Bericht VDI Reihe 8 Nr. 177. Düsseldorf: VDI-Verlag.

- Kroll, A.** 1994. *Identifikation funktionaler Fuzzy-Modelle*. Forschungsbericht Nr. 10/94 MSRT. Gerhard-Mercator-Universität – GH Duisburg.
- Kroll, A.** 1995a. *Zur analytischen Beschreibung von Fuzzy-Systemen*. Forschungsbericht Nr. 04/95 MSRT. Gerhard-Mercator-Universität – GH Duisburg.
- Kroll, A.** 1995b. Identification of functional fuzzy models using multidimensional reference fuzzy sets. *Fuzzy Sets and Systems* (zur Veröffentlichung angenommen).
- Kroll, A., H. Reuter und H. Hebisch.** 1995. *Zur experimentellen Fuzzy-Modellbildung eines pneumatischen Translationsantriebes* Forschungsbericht Nr. 13/95 MSRT. Gerhard-Mercator-Universität – GH Duisburg.
- Küpper, K.** 1994. Self Learning Fuzzy Models Using Stochastic Approximation. *Proceedings of the 3rd IEEE Conference on Control Applications, Glasgow/Schottland*. 1723–1728.
- Küpper, K.** 1995. *Echtzeitsimulation mittels Fuzzy-Systemmodellen*. Zwischenbericht zum gleichnamigen vom Ministerium für Wissenschaft und Forschung des Landes Nordrhein-Westfalen (MWF) geförderten Forschungsvorhaben (unveröffentlicht). Fachgebiet Meß-, Steuer- und Regelungstechnik, Gerhard-Mercator-Universität – GH Duisburg.
- Ljung, L.** 1987. *System Identification – Theory for the User*. Englewood Cliffs /New Jersey: Prentice Hall.
- Ljung, L.** 1988. *MATLAB: System Identification Toolbox User's Guide*. South Natick /Massachusetts: The MathWorks, Inc.
- Ljung, L. und T. Söderström.** 1987. *Theory and Practice of Recursive Identification*. 1st Paperback edition. Cambridge/Massachusetts: The MIT Press.
- Mohler, R.R.** 1973. *Bilinear Control Processes*. New York: Academic Press.
- Pedrycz, W.** 1984. An identification algorithm in fuzzy relational systems. *Fuzzy Sets and Systems* 13. 153–167.
- Rechenberg, I.** 1973. *Evolutionsstrategie: Optimierung technischer Systeme nach Prinzipien der biologischen Evolution*. Stuttgart: Frommann-Holsboog.
- Reuter, H.** 1995. *Zur Identifikation nichtlinearer Systemmodelle mit wenig A-priori-Informationen*. Dissertation Gerhard-Mercator-Universität – GH Duisburg. Fortschritt-Bericht VDI Reihe 8 Nr. 471. Düsseldorf: VDI-Verlag.
- Schwarz, H.** 1991. *Nichtlineare Regelungssysteme – Systemtheoretische Grundlagen*. München: Oldenbourg.

- Sugeno, M. und G.T. Kang.** 1985. Structure identification of fuzzy model. *Fuzzy Sets and Systems* 28. 15–33.
- Sugeno, M. und K. Tanaka.** 1991. Successive identification of a fuzzy model and its applications to prediction of a complex system. *Fuzzy Sets and Systems* 42. 315–334.
- Sugeno, M. und T. Yasukawa.** 1993. A fuzzy logic based approach to qualitative modeling. *IEEE Transactions of Fuzzy Systems* 1. 1–34.
- Takagi, T. und M. Sugeno.** 1985. Fuzzy identification of systems and its application to modeling and control. *IEEE Transactions on Systems, Man, and Cybernetic* 15. 116–132.
- Tong, R.M.** 1978. Synthesis of fuzzy models derived from experimental data. *Fuzzy Sets and Systems* 4. 1–12.
- Tsympkin, Y.Z.** 1971. *Adaption and Learning in Automatic Systems*. New York, London: Academic Press.
- Xu, C.-W. und Y.Z. Lu.** 1987. Fuzzy model identification and self-learning for dynamic systems. *IEEE Transactions on Systems, Man, and Cybernetics* 17. 683–689.
- Zadeh, L.A.** 1965. Fuzzy sets. *Information and Control* 8. 338–353.