

Neue Methoden für die dynamische Identifikation mit Neuro-Fuzzy Netzen

New Methods for Dynamic Identification with Neuro-Fuzzy Networks

Stefan Jakubek, Christoph Hametner, Nikolaus Keuth, Andreas Voigt

Dieser Artikel behandelt die Identifikation nichtlinearer dynamischer Prozesse mit lokalen Neuro-Fuzzy Netzen. Diese bieten den Vorteil, dass aufgrund ihrer Architektur die Möglichkeit besteht, Prozesswissen bei der Modellbildung miteinzubinden. Es werden Lösungsansätze für zwei wesentliche Problemstellungen der dynamischen Identifikation präsentiert: Einerseits wird die Problematik verrauschter Ein- und Ausgangsdaten behandelt, die bei der Parameterbestimmung mit herkömmlichen Regressionsmethoden zu biasbehafteten Resultaten führt. Als Lösungsweg wird die Total Least Squares Methode vorgestellt, und für die Anwendung in lokalen Neuro-Fuzzy Netzen adaptiert. Andererseits wird eine Methode vorgestellt, mit der die Einhaltung von Stationärpunkten erzwungen werden kann, was wesentlich zur Verbesserung der stationären Genauigkeit des Modells beiträgt. Resultate aus einem Praxisbeispiel illustrieren die Anwendbarkeit beider Konzepte.

This article deals with the identification of nonlinear dynamic processes with local Neuro-Fuzzy networks. These networks have the advantage that their architecture offers the possibility to incorporate in-depth process know-how into the modeling procedure. The article presents solution approaches for two major issues in dynamic identification: First, the problem of noisy input- and output data is treated, which causes biased parameters when conventional regression techniques are applied. As a possible solution, the concept of Total Least Squares is presented and adapted for application in local Neuro Fuzzy Networks. Second, a method for the enforcement of stationary gains is presented that significantly improves the model precision during steady-state phases. Results from a practical example illustrate the applicability of both concepts.

1 Einleitung

Die Anwendung komplexer Konzepte der Prozessautomatisierung wie zum Beispiel Regelung oder Fehlerdiagnose erfordert in zunehmendem Maß genaue mathematische Modelle für das statische und dynamische Verhalten von realen Systemen und Prozessen. Aufgrund der hohen Komplexität der Vorgänge in vielen realen Systemen kommen physikalische Modelle nur selten zum Einsatz. Ihr hoher Rechenaufwand in Verbindung mit den Schwierigkeiten bei der Parametrierung machen eine praktische Anwendung oft unmöglich. Auf der anderen Seite stehen reine Black-Box Modelle, die ohne jegliche strukturelle Information über den Prozess auskommen. Diese Eigenschaft, die meist als Vorteil der Black-Box Modelle gesehen wird, kann sich jedoch dann nachteilig auswirken, wenn durch den Versuchsingenieur qualitati-

ves Wissen über die Nichtlinearität des Prozesses eingebracht werden könnte. In diesem Fall ist die Anwendung eines sogenannten *Grey-Box* Modells sinnvoll, wobei sich besonders die Architektur der lokalen Neuro-Fuzzy Netze bewährt hat. Diese interpolieren zwischen verschiedenen lokalen Modellen, die jeweils in unterschiedlichen Bereichen eines durch den Benutzer definierbaren Partitionierungsraums gültig sind. Dadurch kann sich das Modell automatisch und effizient an die Komplexität des Systems anpassen, [1, 2, 3, 4, 5].

Ein prinzipielles Problem bei der Identifikation dynamischer Modelle ist das in der Praxis fast immer vorhandene Messrauschen in Ein- und Ausgängen. Herkömmliche Methoden zur Minimierung sowohl des Prädiktionsfehlers als auch des Simulationsfehlers haben den Nachteil, dass die Parameterschätzung durch das Rauschen biasbehaftet ist. Das bedeutet, dass auch bei be-

liebig großer Anzahl der Datensätze ein systematischer Schätzfehler vorliegt. Besonders gravierend wirkt sich dieser Umstand bei starkem Messrauschen sowie bei Systemen mit sehr unterschiedlich stark verrauschten Ein- und Ausgangsgrößen aus. Ein Alternative, die auch bei Neuro-Fuzzy Netzwerken gut angewendet werden kann, ist das *Total Least Squares Verfahren* (TLS). In Abschnitt 3 wird die Anwendung dieses Verfahrens für Neuro-Fuzzy Netzwerke sowie die Adaption für unterschiedlich stark verrauschte bzw. untereinander korrelierte Ein- und Ausgänge beleuchtet.

Ein anderes grundsätzliches Problem bei dynamischen Modellen ist deren Präzision in Stationärphasen. Während statische Modelle hier grundsätzlich eine hohe Genauigkeit erzielen, müssen bei dynamischen Modellen stationäre Haltepunkte in die Vermessungssequenz eingefügt werden, um deren stationäre Genauigkeit zu verbessern. Eine alternative und sehr effiziente Strategie besteht darin, beim Training neben der Optimierung eines Gütekriteriums die Einhaltung ausgewählter Stationärpunkte zu *erzwingen*. Dadurch kann einerseits Vermessungszeit eingespart werden, andererseits ist die stationäre Genauigkeit an diesen Punkten dann garantiert. Aus mathematischer Sicht führt diese Strategie vom reinen Extremwertproblem hin zu einer Optimierung mit Nebenbedingungen. Abschnitt 4 beschreibt, wie Stationärpunkte in einem Neuro-Fuzzy Netz erzwingen werden können und wie der Aufbau der lokalen Modelle an diesen ausgerichtet werden kann.

Insgesamt kommt es sowohl durch die Anwendung des Total Least Squares Verfahrens als auch durch die erzwungenen Stationärpunkte zu einer wesentlichen Verbesserung der Performance. In Abschnitt 5 werden beide Konzepte anhand eines Praxisbeispiels beleuchtet und mit herkömmlicher Optimierung verglichen.

2 Training von dynamischen Neuro-Fuzzy Netzen

Abb. 1 zeigt die Architektur eines dynamischen lokalen Fuzzy-Netzwerks. Jedes lokale Modell (bezeichnet mit LM_j) bildet vergangene Ein- und Ausgangswerte in \mathbf{u} gemäß

$$\mathbf{u} = [\mathbf{u}_1(k), \mathbf{u}_2(k), \dots, \mathbf{u}_r(k), \hat{\mathbf{y}}(k)]^T \quad (1)$$

auf eine lokale Schätzung $\hat{y}_j(k)$ von $y(k)$ ab. Dabei enthält $\mathbf{u}_i(k)$ vergangene Werte des i -ten Eingangs gemäß

$$\mathbf{u}_i(k) = [u_i(k - d_i - 1), u_i(k - d_i - 2), \dots, \dots, u_i(k - d_i - m_i)]$$

und $\hat{\mathbf{y}}(k)$ enthält vergangene Netzwerkausgänge:

$$\hat{\mathbf{y}}(k) = [\hat{y}(k - 1) \hat{y}(k - 2) \dots \hat{y}(k - n)].$$

In obigen Gleichungen bezeichnet m_i ($i = 1, \dots, r$) die Zählerordnung der i -ten Eingangsgröße, d_i ist die zugeordnete Totzeit und n die Nennerordnung. Über die Zugehörigkeitsfunktionen Φ_j werden diese lokalen Schätzungen \hat{y}_j zum globalen Modellausgang $\hat{y}(k)$ aggregiert:

$$\hat{y}(k) = \sum_{j=1}^m \Phi_j(\mathbf{u}_\Phi) \hat{y}_j(\mathbf{u}, \theta_j). \quad (2)$$

Dabei bezeichnet θ_j einen Vektor, der die Parameter des j -ten lokalen Modells enthält.

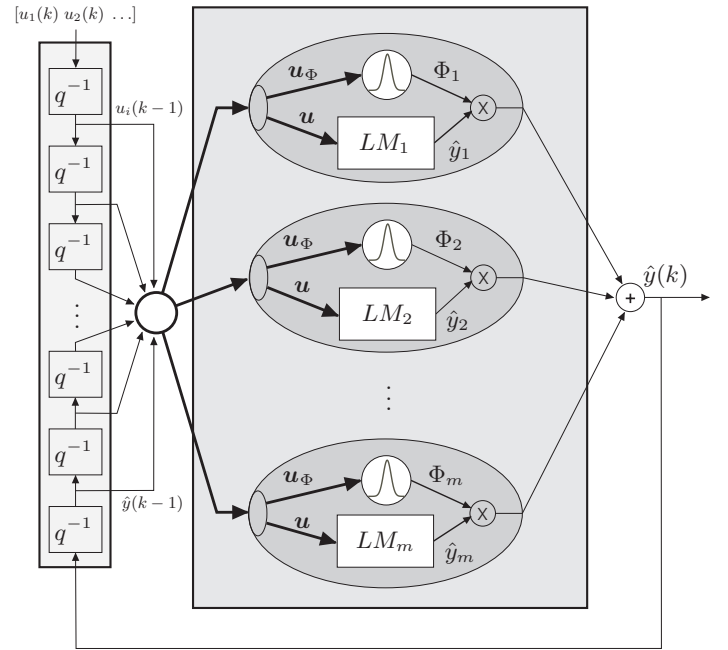


Bild 1: Architektur eines dynamischen lokalen Fuzzy-Netzwerks

Für dynamische Modelle ist es üblich, eine lineare lokale Modellstruktur gemäß

$$\hat{y}_j(\mathbf{u}, \theta_j) = \mathbf{u}^T \theta_j \quad (3)$$

vorzugeben. Die dadurch resultierenden linearen lokalen Übertragungsfunktionen können dann z.B. mit dem umfangreichen Fundus der linearen Modellanalyse untersucht werden. Auch andere modellbasierte Anwendungen der Automatisierungstechnik wie Reglerauslegung oder Fehlerdiagnoseverfahren profitieren von einer linearen Modellstruktur.

Wie aus Abb. 1 und Gl. (2) hervorgeht, haben die lokalen Modelle und die Zugehörigkeitsfunktionen unterschiedliche Argumente. Während die lokalen Modelle alle in (1) bezeichneten vergangenen Ein- und Ausgänge verarbeiten, ist es für die Partitionierung der lokalen

Modelle sinnvoll, nur jene Größen in \mathbf{u}_Φ zu inkludieren, welche die Nichtlinearität des Prozesses charakterisieren. Diese Möglichkeit der Unterscheidung zwischen *Eingangsraum* und *Partitionierungsraum* ist ein wesentlicher Vorteil der Neuro-Fuzzy Architektur, weil dadurch die Dimensionalität der Problemstellung drastisch reduziert werden kann, [3]. Da die Auswahl des Partitionierungsraums meist auf Prozesswissen beruht, liegt streng genommen ein *Grey-Box-Modell* vor. Wird der Partitionierungsraum aus allen Eingangsgrößen aus (1) sowie aus der Ausgangsgröße aufgespannt, spricht man von *Product-Space Clustering*, wobei dann wieder ein Black-Box Modell vorliegt. Hier ist das Gustafson-Kessel Clusteringverfahren [6] zu nennen, welches anhand der vorhandenen Daten linear modellierbare Teilbereiche auffindet. Bei höherer Eingangsdimension und einer großen Anzahl an Trainingsdaten steht den Vorteilen dieses Verfahrens der Nachteil des erheblichen Rechenaufwands gegenüber, [7].

Die vorliegende Arbeit baut auf dem in [5] für statische Modelle vorgestellten Trainingsalgorithmus auf. Die Zugehörigkeitsfunktionen werden iterativ anhand der vorhandenen Trainingsdaten bestimmt, siehe Abb. 2. Der Partitionierungsalgorithmus hat die Aufgabe, lokale Modelle bzw. deren Zugehörigkeitsfunktionen $\Phi_j(\mathbf{u}_\Phi)$ so zu bestimmen, dass diese eine vom Anwender gewünschte Genauigkeit erreichen. Bei dynamischen Identifikationsproblemen wird dazu gefordert, dass der mittlere quadratische Prädiktionsfehler (MSE) der Trainingsdaten in einem lokalen Modell ein bestimmtes Vielfaches der Rauschvarianz σ_y^2 des Ausgangs $y(k)$ nicht überschreitet:

$$\frac{1}{N_j - n_{\text{param}}} \sum_{k=1}^{N_j} [y(k) - \hat{y}_j(k, \mathbf{u}, \boldsymbol{\theta}_j)]^2 \leq \sigma_y^2 \alpha. \quad (4)$$

Der *Prädiktionsfehler* entsteht, indem man in \mathbf{u} aus (1) die Schätzungen in $\hat{\mathbf{y}}(k)$ durch *Messungen* ersetzt. Dabei bezeichnet n_{param} die Anzahl der Parameter des j -ten lokalen Modells, N_j ist die Anzahl der Datenpunkte, aus denen es gebildet wird und α ist ein Skalierungsparameter, der die relative Größe der lokalen Modelle bestimmt und zur Regularisierung dient.

Der Prädiktionsfehler $[y(k) - \hat{y}_j(k, \mathbf{u}, \boldsymbol{\theta}_j)]$ setzt sich aus mehreren Komponenten zusammen:

1. Messrauschen in $y(k)$
2. Messrauschen in den Komponenten von \mathbf{u}
3. Statistische Unsicherheit des Parametervektors $\boldsymbol{\theta}_j$

Der Wertebereich $\alpha < 1$ würde demnach zu einem Overfit des Modells führen, während für steigendes α ein größerer Truncation-Fehler zugelassen wird, sodass prinzipiell immer größere Modelle gebildet werden. Abb. 2 verdeutlicht die Vorgangsweise bei der Bestimmung eines lokalen Modells in einem zweidimensionalen Partitionierungsraum: Schwarze Punkte (●) kennzeichnen jene Datensätze, die das Kriterium (4) bereits erfüllen. Die Erweiterung der Zugehörigkeitsfunktion auf einen

benachbarten Datenpunkt (★) bewirkt, dass diese nun auch weitere Datensätze umfasst (○). Der neue Datenpunkt (★) darf nur dann zum lokalen Modell hinzugefügt werden, wenn *alle* betroffenen Punkte (●, ★, ○) zusammen (4) erfüllen.

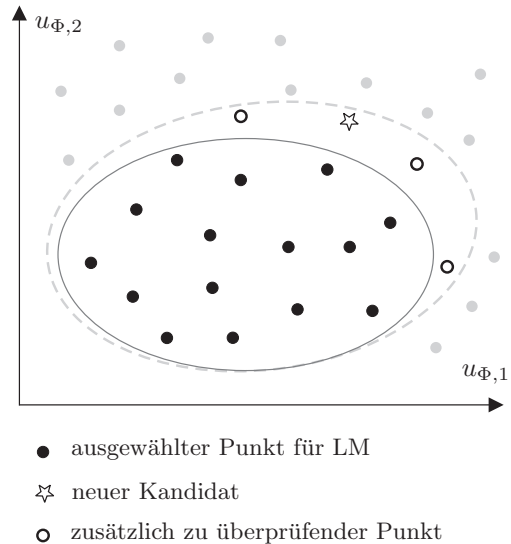


Bild 2: Auswahl der Datenpunkte für die Bestimmung der lokalen Modellparameter

3 Total Least Squares

3.1 Grundprinzipien

Die Schätzung der lokalen Modellparameter $\boldsymbol{\theta}_j$ in (2) durch Least-Squares beruht auf einer Minimierung des Prädiktionsfehlers an den Trainingsdatenpunkten. Bei der Identifikation dynamischer Prozessmodelle enthalten die Regressoren in (1) vergangene Ein- und Ausgänge des Prozesses, das Messrauschen beeinflusst also sowohl die Zielgröße $y(k)$, als auch den Regressor selbst. Die so entstehende Korrelation zwischen Regressor und Zielgröße bewirkt, dass die Parameter durch Least Squares nicht biasfrei geschätzt werden. Um im Fall verrauschter Ein- und Ausgangsdaten eine biasfreie Parameterschätzung zu erhalten, müssen sowohl Eingänge als auch Ausgänge rekonstruiert werden. Dies führt zum Total Least Squares-Verfahren (TLS), für welches gezeigt werden kann, dass die Schätzung der Parameter erwartungstreu ist, [8].

3.2 Methodik

Die Anwendung von Total Least Squares bei verrauschten Ein- und Ausgangsdaten wurde in der Vergangenheit schon öfters vorgeschlagen, [8, 9, 10]. Implizit liefert auch der Gustafson-Kessel-Algorithmus [6] im Product Space aus den Cluster-Parametern die TLS-Schätzung, jedoch mit dem weiter oben beschriebenen Nachteil.

Das Total Least-Squares Verfahren wird in Folge kurz für ein SISO-System vorgestellt, eine anführliche Darstellung liefert z.B. [11]. Danach wird in Abschnitt 3.3 auf die in der Praxis normalerweise auftretende Situation eingegangen, dass Ein- und Ausgänge unterschiedlich stark verrauscht sind, bzw. dass auch Korrelationen zwischen den Rauschsignalen vorliegen.

Seien $\mathbf{X} \in \mathbb{R}^{N \times M}$ die Regressormatrix und $\mathbf{y} \in \mathbb{R}^{N \times 1}$ der Zielgrößenvektor eines SISO-Identifikationsproblems. Die Zeilenvektoren \mathbf{x}^T in \mathbf{X} und die Elemente von \mathbf{y} sind also nach dem Schema

$$\left. \begin{aligned} \mathbf{x}^T(k) &= [y(k-1), y(k-2), \dots | \\ \mathbf{y}(k) &= [u(k-d-1), u(k-d-2), \dots] \end{aligned} \right\} (5)$$

aufgebaut. Total Least Squares zielt nun darauf ab, sowohl \mathbf{y} als auch \mathbf{X} so zu modifizieren, dass folgendes gilt:

$$\mathbf{y}, \mathbf{X} \Rightarrow \hat{\mathbf{y}}, \hat{\mathbf{X}}, \text{ wobei} (6)$$

$$\hat{\mathbf{y}} \in \text{Image}(\hat{\mathbf{X}}) \text{ und } \|\mathbf{y} - \hat{\mathbf{y}} | \mathbf{X} - \hat{\mathbf{X}}\|_F = \min.$$

Dazu definiert man eine erweiterte Regressormatrix:

$$\mathbf{W} = [\mathbf{y} | \mathbf{X}] (7)$$

Die TLS-Parameterschätzung bestimmt in dem durch die erweiterten Regressoren (7) aufgespannten Product Space eine Hyperebene durch Stützpunkt \mathbf{m} und Normaleneinheitsvektor \mathbf{b} , die (6) erfüllt. Der Stützpunkt \mathbf{m} ergibt sich als der Schwerpunkt (Zentroide) aller Datensätze in \mathbf{W} :

$$\mathbf{m} = \frac{1}{N} \mathbf{W}^T \mathbf{1} \quad \text{mit} \quad \mathbf{1} = [1 \ 1 \ \dots]^T. (8)$$

Bezieht man die Datensätze nun auf die Zentroide gemäß $\tilde{\mathbf{W}} = \mathbf{W} - \mathbf{1}\mathbf{m}^T$, dann wird der Normalenvektor \mathbf{b} jener Eigenvektor, der zum kleinsten Eigenwert von $(\tilde{\mathbf{W}}^T \tilde{\mathbf{W}})$ gehört.

Durch \mathbf{m} erhält man zusätzlich zu den Regressoren in $\mathbf{x}^T(k)$ noch einen Bias-Term. Partitioniert man \mathbf{b} gemäß $\mathbf{b} = [b_1, \boldsymbol{\beta}]^T$, dann liefert TLS mit den Regressoren aus (5) folgende Differenzgleichung:

$$\hat{y}(k) = \frac{1}{b_1} [-\mathbf{x}^T(k)\boldsymbol{\beta} + \mathbf{m}^T \mathbf{b}]. (9)$$

3.3 Dekorrelation der Messdaten

Die Parameterberechnung durch TLS setzt voraus, dass alle Messgrößen gleich stark verrauscht und nicht miteinander korreliert sind. In der Praxis sind diese beiden Voraussetzungen meist nicht erfüllt, weshalb dann die Ein- und Ausgangssignale zuvor dekorreliert werden müssen. Es bezeichne $\nu(k)$ das Störgeräusch, das dem wahren Ausgang $y_0(k)$ überlagert ist, weiters sei $\mu(k)$ das Störgeräusch zum wahren Eingang $u_0(k)$.

Ein Datensatz $\mathbf{w}(k)$ aus \mathbf{W} ergibt sich damit aus seinem ungestörten Äquivalent $\mathbf{w}_0(k)$ durch

$$\mathbf{w}^T(k) = \mathbf{w}_0^T(k) + \mathbf{n}^T(k),$$

mit dem Rauschvektor

$$\mathbf{n}(k) = [\nu(k), \nu(k-1), \dots | \mu(k-d-1), \mu(k-d-2), \dots]^T. (10)$$

Die Kovarianzmatrix $\mathbf{R}_n = E\{\mathbf{n}(k)\mathbf{n}^T(k)\}$ enthält alle oben angesprochenen Korrelationen und wird hier als bekannt vorausgesetzt. Sie kann in der Praxis mit Datensätzen aus Stationärphasen gemäß

$$\mathbf{R}_n \approx \frac{1}{N-1} (\tilde{\mathbf{W}}^T \tilde{\mathbf{W}}) (11)$$

geschätzt werden, wobei (11) für $N \rightarrow \infty$ gegen den Erwartungswert $E\{\mathbf{n}(k)\mathbf{n}^T(k)\}$ konvergiert. In den durchgeführten Praxiserprobungen hat die Anwendung von (11) stets zu befriedigenden Ergebnissen geführt. Für eine korrekte Anwendung von TLS müssten die einzelnen Elemente des Rauschvektors (10) unkorreliert sein. Ist dies nicht der Fall, werden die Regressoren in (7) einer Transformation unterworfen, indem \mathbf{b} durch $\tilde{\mathbf{b}} = \mathbf{T}\tilde{\mathbf{b}}$ ersetzt wird. Der für die Bestimmung von $\tilde{\mathbf{b}}$ relevante Rauschvektor $\tilde{\mathbf{n}}$ ergibt sich dann zu $\tilde{\mathbf{n}}^T(k) = \mathbf{n}^T(k)\mathbf{T}$. Die korrekte Optimierung von $\tilde{\mathbf{b}}$ durch TLS kann sichergestellt werden, indem \mathbf{T} derart gewählt wird, dass die Kovarianzmatrix $\mathbf{R}_{\tilde{\mathbf{n}}} = E\{\tilde{\mathbf{n}}(k)\tilde{\mathbf{n}}^T(k)\}$ gleich der Einheitsmatrix wird:

$$\mathbf{R}_{\tilde{\mathbf{n}}} = \mathbf{T}^T \mathbf{R}_n \mathbf{T} = \mathbf{I} (12)$$

Der Vektor $\tilde{\mathbf{b}}$ ergibt sich dann als Eigenvektor zum kleinsten Eigenwert der Matrix $\mathbf{T}^T \tilde{\mathbf{W}}^T \tilde{\mathbf{W}} \mathbf{T}$. Im Fall reinen Messrauschens ist \mathbf{R}_n eine Diagonalmatrix, sodass $\mathbf{T} = \mathbf{R}_n^{-1/2}$ gewählt werden kann. In komplexeren Fällen muß auf eine Eigenwert-Eigenvektorzerlegung von \mathbf{R}_n gemäß $\mathbf{R}_n = \mathbf{S}\mathbf{D}\mathbf{S}^T$ zurückgegriffen werden. Gl. (12) wird dann erfüllt durch $\mathbf{T} = \mathbf{S}\mathbf{D}^{-1/2}$.

3.4 Weighted TLS

Für die Parameterschätzung an lokalen Modellen werden die Schätzfehler an den einzelnen Datenpunkten über den Wert der Zugehörigkeitsfunktion Φ_j gewichtet. Es bezeichne \mathbf{Q}_j eine diagonale Gewichtsmatrix für das j -te lokale Modell. Ihre Diagonalelemente q_{ji} entsprechen den Werten der Zugehörigkeitsfunktion $\Phi_j(\mathbf{u}_\Phi(i))$ an den Trainingsdatenpunkten $\mathbf{u}_\Phi(i)$.

Anstatt der Frobeniusnorm in (6) wird nun folgendes Kriterium minimiert:

$$\|\mathbf{Q}_j^{1/2}(\mathbf{W} - \hat{\mathbf{W}})\|_F^2 = \min. (13)$$

An die Stelle der Zentroide in (8) tritt nun eine gewichtete Variante:

$$\mathbf{m}_j = \mathbf{W}^T \mathbf{q}_j / s_q. \quad (14)$$

Dabei bezeichnet \mathbf{q}_j den Vektor der Hauptdiagonalen von \mathbf{Q}_j und s_q ist die Summe aller seiner Elemente: $s_q = \mathbf{1}^T \mathbf{q}_j$. Für die Bestimmung von \mathbf{b} bzw. $\tilde{\mathbf{b}}$ wird nun $\tilde{\mathbf{W}}^T \mathbf{Q}_j \tilde{\mathbf{W}}$ herangezogen.

4 Stationärpunkte

Dynamische Modelle streben an, gleichermaßen die stationären Genauigkeitsanforderungen und die Anforderungen hinsichtlich einer präzisen Abbildung der Prozessdynamik zu erfüllen, [12]. Besonders bei Modellen, die für Diagnosekonzepte eingesetzt werden sollen, werden hohe stationäre Genauigkeitsanforderungen gestellt. Üblicherweise werden deshalb innerhalb der dynamischen Anregung stationäre Haltepunkte eingefügt. Stärkere Gewichtung dieser Haltepunkte bei der LS- oder TLS-Optimierung bewirkt generell eine bessere stationäre Genauigkeit.

Eine andere Strategie besteht darin, die Einhaltung gewisser stationärer Punkte durch das Modell zu *erzwingen*. Die verbleibenden Freiheitsgrade des Modells werden dann durch LS bzw. TLS zur Optimierung der dynamischen Abbildungsqualität des Modells eingesetzt. Dadurch ist eine exakte Trennung der beiden Modellanforderungen gegeben. Prinzipiell ist es möglich, pro lokalem Modell und unabhängiger Eingangsgröße einen Stationärpunkt zu erzwingen.

Während stationärer Phasen gilt für den Regressor (5)

$$\mathbf{x}_s^T(i) = [y_s(i) \ y_s(i) \ \dots \ | \ u_s(i) \ u_s(i) \ \dots]. \quad (15)$$

bzw. für den erweiterten TLS-Regressor

$$\mathbf{w}_s^T(i) = [y_s(i) \ | \ \mathbf{x}_s^T(i)] \quad (16)$$

und nach Bezug auf den Stützpunkt \mathbf{m}

$$\tilde{\mathbf{w}}_s^T(i) = \mathbf{w}_s^T(i) - \mathbf{m}^T. \quad (17)$$

Der Subskript s deutet stationäre Regressoren an, der Index i unterscheidet die Stationärphasen, falls es mehrere geben sollte.

4.1 Stationärpunkte bei LS-Optimierung

Das Least Squares Verfahren zielt darauf ab, den Prädiktionsfehler an den Trainingsdaten zu minimieren. Mit \mathbf{y} aus (5) und der Regressionsgleichung $\hat{\mathbf{y}}_j = \mathbf{X}\boldsymbol{\theta}_j$ wird diese Forderung zu

$$J = (\mathbf{y} - \mathbf{X}\boldsymbol{\theta}_j)^T (\mathbf{y} - \mathbf{X}\boldsymbol{\theta}_j) = \min. \quad (18)$$

Für stationäre Punkte gilt mit (15)

$$\mathbf{x}_s^T(i)\boldsymbol{\theta}_j = y_s(i).$$

Fasst man \mathbf{x}_s^T und y_s aus n_s Stationärpunkten zusammen, so wird aus obiger Gleichung

$$\mathbf{X}_s \boldsymbol{\theta}_j = \mathbf{y}_s. \quad (19)$$

Die Minimierung von (18) unter den Nebenbedingungen (19) erfolgt mit dem Prinzip der Lagrange'schen Multiplikatoren:

$$\mathcal{L}(\boldsymbol{\theta}_j, \boldsymbol{\lambda}) = \frac{1}{2}(\mathbf{y} - \mathbf{X}\boldsymbol{\theta}_j)^T (\mathbf{y} - \mathbf{X}\boldsymbol{\theta}_j) + \boldsymbol{\lambda}^T (\mathbf{X}_s \boldsymbol{\theta}_j - \mathbf{y}_s) \quad (20)$$

In [13] findet sich folgende kompakte Lösung für das Problem (18) & (19): Die Minimierung von (18) alleine würde mit $\mathbf{P} = (\mathbf{X}^T \mathbf{X})^{-1}$ die Least-Squares Lösung $\boldsymbol{\theta}_{LS,j} = \mathbf{P}\mathbf{X}^T \mathbf{y}$ ergeben. Mit der Abkürzung $\mathbf{P}_s = (\mathbf{X}_s \mathbf{P} \mathbf{X}_s^T)^{-1}$ erhält man für die endgültige Lösung

$$\boldsymbol{\theta}_j = \boldsymbol{\theta}_{LS,j} - \mathbf{P} \mathbf{X}_s^T \mathbf{P}_s (\mathbf{X}_s \boldsymbol{\theta}_{LS,j} - \mathbf{y}_s). \quad (21)$$

Mit (21) ist es also möglich, einen mit Least-Squares ermittelten Parametervektor $\boldsymbol{\theta}_{LS,j}$ *nachträglich* an Stationärbedingungen anzupassen.

4.2 Stationärpunkte bei TLS-Optimierung

Damit ein durch das TLS-Verfahren optimiertes Modell die i -te Stationärphase korrekt abbildet, muß für den Vektor \mathbf{b} (bzw. $\tilde{\mathbf{b}}$) zusätzlich gelten

$$\tilde{\mathbf{w}}_s^T(i)\mathbf{b} = 0. \quad (22)$$

bzw.

$$\tilde{\mathbf{S}}_s \mathbf{b} = 0 \quad \text{mit} \quad \tilde{\mathbf{S}}_s = \begin{bmatrix} \tilde{\mathbf{w}}_s^T(1) \\ \tilde{\mathbf{w}}_s^T(2) \\ \vdots \end{bmatrix} \quad (23)$$

wenn mehrere Stationärpunkte $1, 2, \dots$ erzwungen werden sollen. Die Erfüllung von (23) bei der TLS-Optimierung kann sichergestellt werden, indem \mathbf{b} (bzw. $\tilde{\mathbf{b}}$) auf den Null-Raum von $\tilde{\mathbf{S}}_s$ beschränkt wird:

$$\mathbf{b} \in \mathcal{N}(\tilde{\mathbf{S}}_s) \quad \text{bzw.} \quad \mathbf{b} = \mathbf{N}(\tilde{\mathbf{S}}_s)\boldsymbol{\xi}. \quad (24)$$

Dabei spannen die Spaltenvektoren von $\mathbf{N}(\tilde{\mathbf{S}}_s)$ den Nullraum von $\tilde{\mathbf{S}}_s$ auf. Setzt man für den Normaleneinheitsvektor $\mathbf{b} = \mathbf{N}(\tilde{\mathbf{S}}_s)\boldsymbol{\xi}$, dann zeigt sich, dass $\boldsymbol{\xi}$ sich nun als Eigenvektor zum kleinsten Eigenwert von $\mathbf{N}^T \tilde{\mathbf{W}}^T \tilde{\mathbf{W}} \mathbf{N}$ ergibt. Nach der Bestimmung von \mathbf{b} bzw. $\tilde{\mathbf{b}}$ aus $\mathbf{b} = \mathbf{N}(\tilde{\mathbf{S}}_s)\boldsymbol{\xi}$ kann die Übertragungsfunktion wieder aus (9) ermittelt werden.

4.3 Einbindung in das Trainingsverfahren

Um den effizienten Einsatz der vermessenen Stationärpunkte im Neuro-Fuzzy Modell sicherzustellen, wurde der Trainingsalgorithmus aus [5] modifiziert: Zentrums-kandidaten $z_{i,cand}$ für neue lokale Modelle werden anfänglich auf Stationärpunkte gesetzt, sodass Modelle rund um diese gebildet werden. Erst wenn alle Stationärpunkte „aufgebraucht“ sind, werden Zentrums-kandidaten an die Stelle des größten Prädiktionsfehlers gesetzt.

In Verbindung mit dem vorher beschriebenen iterativen Konstruktionsalgorithmus für die Zugehörigkeitsfunktionen bewirkt die Zentrierung lokaler Modelle anhand der Stationärpunkte, dass der linear modellierbare Bereich rund um diese gut genutzt wird, siehe auch Abb. 6. Andererseits kann dann eine Lage der Stationärpunkte genau im Überlappungsbereich zweier Modelle vermieden werden. Ein Netzwerktraining ohne Berücksichtigung der Lage der Stationärpunkte im Partitionierungsraum kann darin resultieren, dass ein lokales Modell zu viele Stationärpunkte beinhaltet, sodass nicht alle durch (19) bzw. (23) berücksichtigt werden können.

5 Anwendungsbeispiel

5.1 Prozessbeschreibung

Die Erprobung der beiden Elemente Total Least Squares und Stationärpunkte in Verbindung mit dem für dynamische Probleme adaptierten Trainingsalgorithmus [5] erfolgte an einem turboaufgeladenen Großgasmotor der Firma GE-Jenbacher. Die Motoren liefern Nennleistungen von 300kW bis 3MW, werden zur Erzeugung von elektrischer Energie und Fernwärme eingesetzt und können mit einer großen Anzahl unterschiedlicher Biogase, Erdgas oder Industrieabgas betrieben werden.

Abb. 3 zeigt ein Prozessschaubild des Motors. Für die Online-Diagnose des Motors ist es besonders wichtig, die Abgastemperatur T_3 genau zu überwachen, da sie ein unmittelbares Maß für den emissionsarmen Betrieb des Motors darstellt. Online- NO_x -Emissionsmessung steht bei Serienmotoren üblicherweise nicht zur Verfügung. Im Temperaturband $T_3 = 470 - 490 \text{ }^\circ\text{C}$ werden gute Wirkungsgrade bei gleichzeitig niedriger NO_x -Emission erzielt. Ein gesetzmäßiger und wirtschaftlicher Motordauerbetrieb ist bei diesem Motor daher nur innerhalb dieses Abgastemperaturbereiches möglich.

5.2 Modellierung

Die Anregung des Motors zur Identifikation der Temperatur T_3 erfolgte durch die drei Variablen Generatorleistung (P_{El}), Gemischkonzentration (c_{Gas}) und Zündzeitpunkt (ZZP). Weiters wurden insgesamt sechs Stationärpunkte vorgeschlagen, von denen fünf am Prüfstand tatsächlich angefahren wurden, Abb. 4. Die

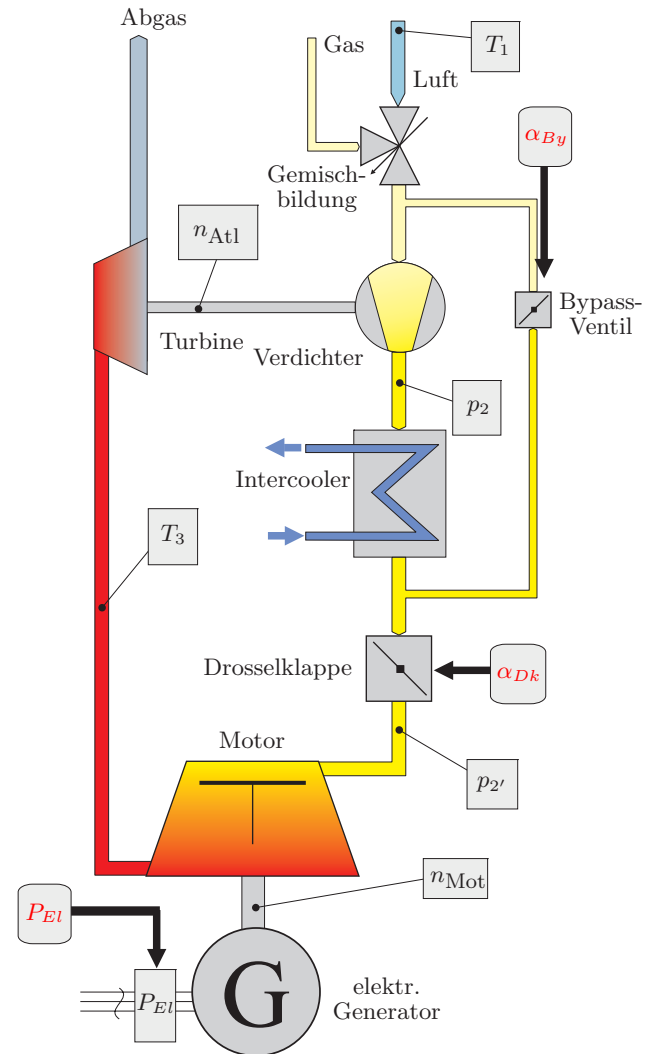


Bild 3: Prozessschaubild eines GE-Jenbacher Gasmotors

Wahl der Modellstruktur erfolgte anhand von Messdaten, die aus Voridentifikationen mit Einzelkanalanregung gewonnen wurden, sowie aus physikalischen Überlegungen. Die Temperatur T_3 reagiert mit unterschiedlich schnellem Ansprechverhalten erster Ordnung sowie mit unterschiedlichen Totzeiten auf die einzelnen Eingangskanäle. Die Reaktion auf Änderungen des ZZP ist z.B. wesentlich schneller als jene auf Laständerungen. Diesem Umstand muss durch eine minimale Systemordnung von drei Rechnung getragen werden, sodass die relevante Dynamik durch ein MISO-ARX-Modell abgebildet werden kann. Zur Wahl der passenden Systemordnung wurde weiters die Modellordnung schrittweise erhöht. Bei Verwendung eines Modells dritter Ordnung konnte jedoch bereits minimaler Generalisierungsfehler an den Validierungsdaten erzielt werden. Die unterschiedlichen Totzeiten der drei gewählten Eingänge wurden ebenfalls aus Sprungversuchen der Voridentifikation ermittelt und beim Aufbau des Regressors entsprechend berücksichtigt.

Die Intensität des Messrauschens in den drei Eingängen und im Ausgang ist stark unterschiedlich: $\sigma_{Pel} = 3.5$,

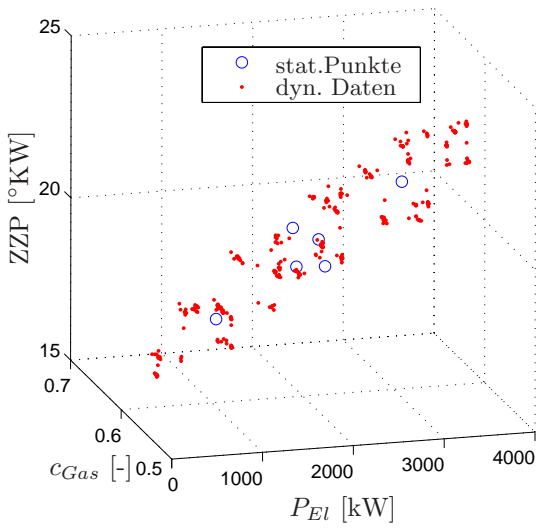


Bild 4: Eingangsraum mit dynamischer Anregung und Stationärpunkten

$\sigma_{c_{Gas}} = 10^{-3}$, $\sigma_{ZZP} = 10^{-1}$ und $\sigma_{T_3} = 4 \cdot 10^{-2}$. Dabei ist zu beachten, dass sich die σ -Werte nach Normierung der Einzelkanäle auf den Wertebereich $[0; 1]$ um bis zum nahezu Zehnfachen unterscheiden.

Die Anregung erfolgte über APRB-Signale, wobei Amplituden und Bandbreiten der Systemeingänge nach folgenden Gesichtspunkten gewählt wurden:

Aus der Voridentifikation wurde die dominante Prozesszeitkonstante ermittelt. Die Bandbreite der einzelnen APRB-Signale wurde so gewählt, dass der Ausgleich dieser Zeitkonstanten bei Wirkung aller drei Eingänge bis auf etwa zehn Prozent zugelassen wurde.

Die maximalen Amplituden der Eingangssignale wurden prinzipiell so groß wie möglich gewählt. Die Limits ergeben sich einerseits aus den physikalischen Grenzen für die Zündung Gas-Luft-Gemisches, da der Motor im Magerbetrieb ($\lambda \approx 1.8$) läuft und andererseits durch die geforderte Klopffestigkeit. Eine simultane Laststeigerung von Leerlauf bis Maximallast garantiert letztlich die Erfassung des gesamten zulässigen Betriebsbereichs des Gasmotors.

Abb. 5 zeigt die Trainingsdaten und das resultierende Ausgangssignal. Es ist zu beachten, dass für den Zündzeitpunkt (ZZP) im Gegensatz zu den restlichen Signalen der Sollwert dargestellt wurde.

Für die Partitionierung der lokalen Modelle wurden die Generatorlast $u_{\Phi,1}(k) = P_{El}(k-1)$ und die Abgastemperatur $u_{\Phi,2}(k) = T_3(k-1)$ gewählt. Der Partitionierungsraum wird idealerweise durch jene Größen aufgespannt, welche die maßgebliche Nichtlinearität des Systems enthalten. Die Motorlast, und damit der Betriebspunkt des Abgasturboladers, hat einen stark nichtlinearen Einfluss auf das Systemverhalten. Nachdem die Turboladerdrehzahl n_{Atl} beim Serienmodell nicht gemessen wird, stellt P_{El} die erste relevante Größe für den Parti-

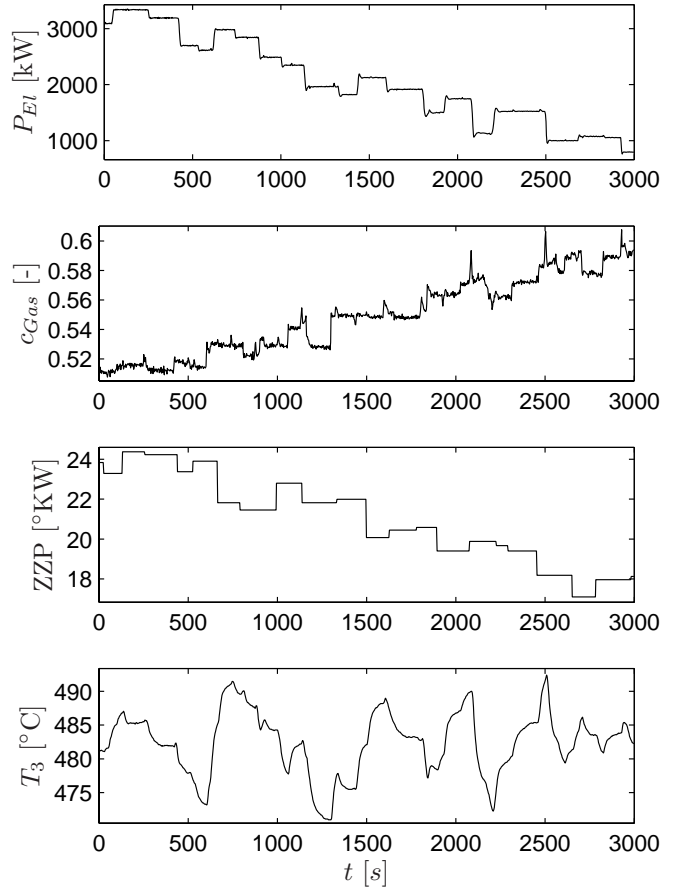


Bild 5: Trainingsdaten im Zeitbereich

tionierungsraum dar.

Die Wahl der Ausgangsgröße T_3 als weitere Partitionierungsachse hat folgende Gründe: Es besteht eine wesentliche Nichtlinearität in der Wirkung der Eingänge auf die Abgastemperatur in Abhängigkeit von den Verbrennungsbedingungen, welche aber durch T_3 gut charakterisiert sind. Wie in [2] dargestellt wird, ist die Hinzunahme des Ausgangs in den Partitionierungsraum auch für die Unterscheidung von lokalen Gleichgewichtsmodellen („equilibrium models“) und transienten Modellen („off-equilibrium models“) wichtig.

5.3 Diskussion

Abb. 6 zeigt die Gültigkeitsfunktionen der lokalen Modelle im oben beschriebenen Partitionierungsraum sowie die Stationärpunkte. Deutlich ist zu erkennen, dass einige lokale Modelle rund um diese gebildet wurden. Die Form und Ausdehnung der einzelnen lokalen Modelle geben Aufschluss über den stark nichtlinearen Einfluss sowohl von T_3 als auch P_{El} . Tabelle 1 beleuchtet die Performance des Neuro-Fuzzy Modells unter Einsatz des TLS-Verfahrens. Es ist zu erkennen, dass der RMSE für die Prädiktion der Identifikationsdaten über σ_{T_3} liegt, was eine notwendige Bedingung für eine gute Generalisierungsfähigkeit ist, siehe die Erläuterungen zu Gl. (4) weiter oben. Tabelle 2 zeigt im Ver-

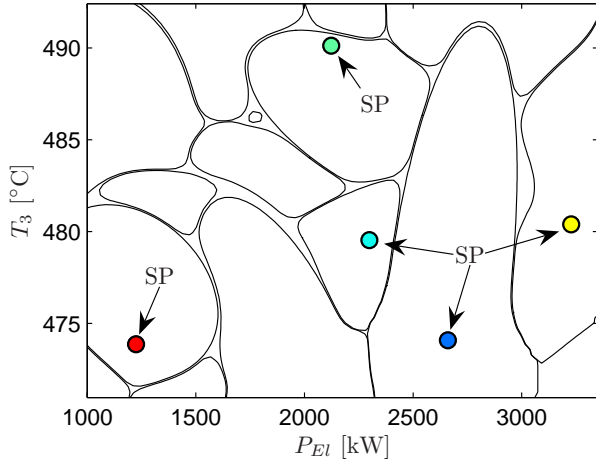


Bild 6: Zugehörigkeitsfunktionen und Stationärpunkte (SP) im Partitionierungsraum

gleich dazu die Resultate eines Modells, bei dem die lokalen Modellparameter herkömmlich bestimmt wurden. Es fällt auf, dass dieses Modell sowohl bei Prädiktion als auch bei Simulation an den Trainingsdaten bessere Werte liefert, als das TLS-Modell, während es sich bei den Validierungsdaten umgekehrt verhält. Dieses Merkmal ist darauf zurückzuführen, dass die Bestimmung der Modellparameter mittels LS auf die Minimierung des Prädiktionsfehlers der Trainingsdaten abzielt. Bei der Generalisierung mit Validierungsdaten zeigt TLS dafür einen wesentlich geringeren Fehler. Dies ist auch deutlich in Abb. 7 zu sehen, wo die Autokorrelationsfunktionen der Prädiktionsfehler beider Modelle dargestellt sind. Beim Simulationsfehler, welcher in einem linearen Modell durch Filterung des Prädiktionsfehlers mit dem geschätzten Nennerpolynom der Übertragungsfunktion entsteht, setzt sich diese Tendenz so wie im vorliegenden Beispiel meist fort. Es ist auch zu beachten, dass sich der Prädiktions-MSE bei Identifikations- und Validierungsdaten unter Verwendung von TLS nur um etwa 11% unterscheidet, während diese Differenz bei LS 66% beträgt.

Modus	Daten aus:	
	Identifikation	Validierung
Prädiktion	0.0608	0.0678
Simulation	0.2387	0.5090

Tabelle 1: Root Mean Squared Error (RMSE) in $^{\circ}\text{C}$ für das TLS-Training mit Stationärpunkten

Modus	Daten aus:	
	Identifikation	Validierung
Prädiktion	0.0490	0.0812
Simulation	0.2170	0.7856

Tabelle 2: Root Mean Squared Error (RMSE) in $^{\circ}\text{C}$ für LS-Training mit Stationärpunkten

Abb. 8 zeigt schließlich den Vergleich zwischen gemessener Abgastemperatur T_3 und simulierten Werten aus dem TLS-Modell bei einem Validierungsversuch mit dynamischer Anregung. Die Versuchsplanung wurde wieder so gestaltet, dass der gesamte Arbeitsbereich des Motors durchfahren wurde. Der maximale Fehler beträgt 1.4°C , der RMSE beträgt 0.5°C , siehe auch Tabelle 1.

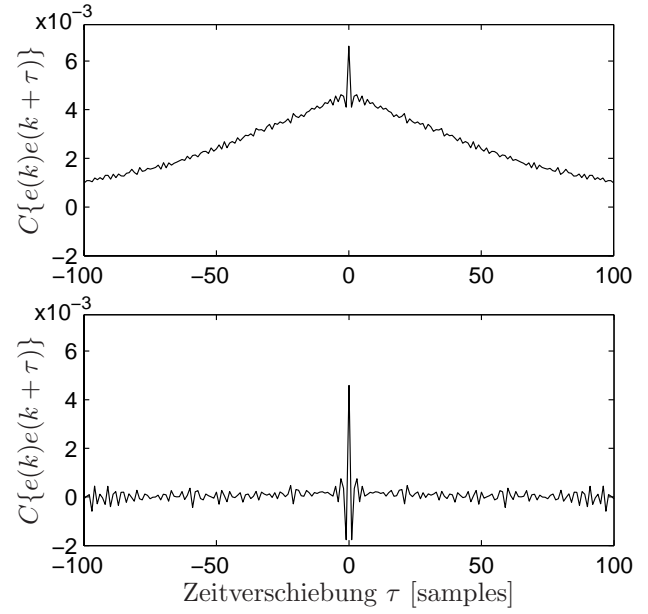


Bild 7: Autokorrelationsfunktion des Prädiktionsfehlers von LS (oben) und TLS(unten)

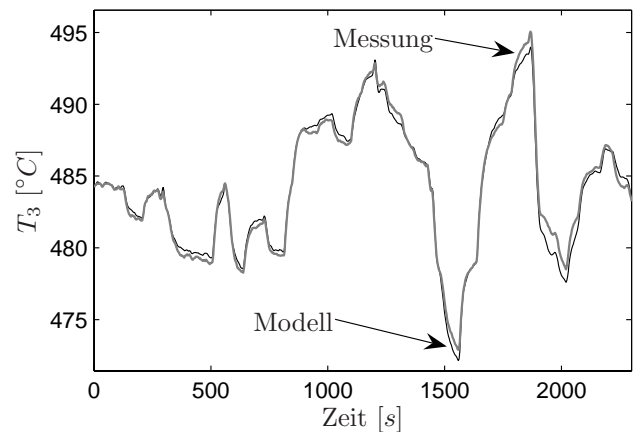


Bild 8: Vergleich zwischen simulierter und gemessener Abgastemperatur bei einem Validierungsversuch

Um die Anwendung des Motormodells unter realen Bedingungen zu testen, wurde eine starke Tag/Nacht-Schwankung der Ansauglufttemperatur T_1 bei verkürzter Periodendauer mittels Prüfstandslüftungssystem aufgebracht. Dieser Test stellt eine Sensitivitätsanalyse des erstellten Modells bezüglich der Umgebungsbedingungen dar. Die Umgebungstemperatur verändert die

Dichte des Gemischs und beeinflusst damit das Turboladersystem Verdichter-Turbine maßgeblich. Der Einfluss der Umgebungstemperatur T_1 ist - bezogen auf die Dynamik des identifizierten Modells - quasistatisch. Er wird deshalb parametrisch über die Dichtekorrektur in c_{Gas} berücksichtigt. Zu quantitativen Analyse wurde ein entsprechender Validierungsversuch mit variierender Temperatur T_1 durchgeführt. In Abb.9 wird die Abgas-temperaturmessung mit zwei Simulationsmodellen verglichen: Einerseits mit einem Modell mit antrainierten Stationärpunkten und andererseits mit einem, dessen Parameter herkömmlich optimiert wurden. Hier ist deutlich zu sehen, wie das Konzept der Stationärpunkte die Genauigkeit verbessert und den systematischen Bias beseitigt, der beim Modell ohne Stationärpunkte bis zu $2^\circ C$ beträgt.

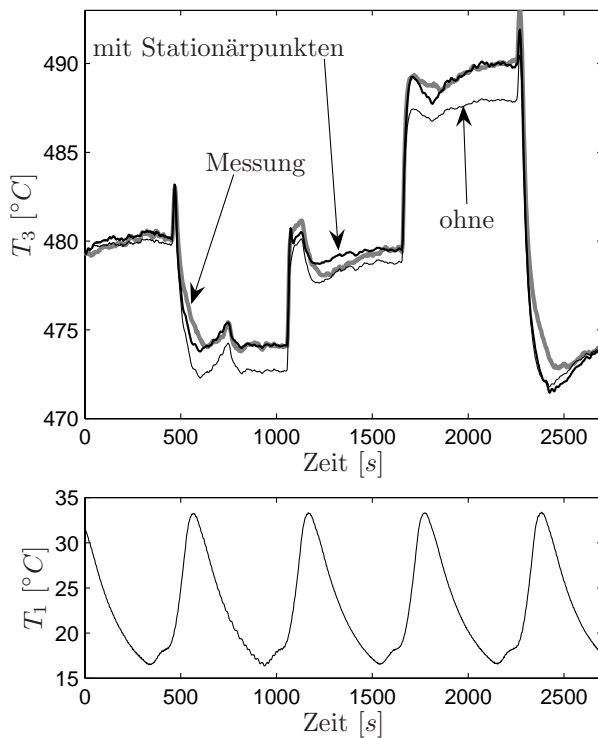


Bild 9: Modellvalidierung bei schwankender Ansauglufttemperatur

6 Schlussfolgerungen und Ausblick

In diesem Artikel wurden zwei neue Konzepte für die dynamische Prozessidentifikation mit lokalen Neuro-Fuzzy Netzwerken beschrieben. Einerseits wurde die Total Least Squares Methode vorgestellt, welche zur biasfreien Bestimmung der lokalen Modellparameter angewendet werden kann und andererseits wurde gezeigt, wie die Einhaltung stationärer Verstärkungen durch das Modell erzwungen werden kann. Anhand eines Praxisbeispiels wurde die Anwendbarkeit beider Konzepte erläutert.

Aktuelle Entwicklungen beinhalten eine automatisierte Totzeitanalyse bei Neuro-Fuzzy Modellidentifikation sowie die Berechnung von Konfidenzintervallen für dynamische Modelle mit verrauschten Ein- und Ausgangsdaten.

Danksagung

Dieses Projekt wurde von AVL List GmbH und GE-Jenbacher finanziert und ermöglicht.

Literatur

- [1] P. Balle, D. Juricic, A. Rakar, and S. Ernst. Identification of Nonlinear Processes and Model Based Fault Isolation using Local Linear Models. *Proceedings of the 1997 American Control Conference*, 1:47–51, 1997.
- [2] T. A. Johansen, R. Shorten, and R. Murray-Smith. On the Interpretation and Identification of Dynamic Takagi-Sugeno Fuzzy Models. *IEEE Transactions on Fuzzy Systems*, 8(3):297–313, 2000.
- [3] O. Nelles. *Nonlinear System Identification*. Springer Verlag, 1st edition, 2002.
- [4] D. Kukulj and E. Levi. Identification of Complex Systems Based on Neural and Takagi-Sugeno Fuzzy Model. *IEEE Transactions on Systems, Man and Cybernetics, Part B*, 34:272 – 282, Feb. 2004.
- [5] S. Jakubek and N. Keuth. Optimierte Neuro-Fuzzy-Modelle für Auslegungsprozesse und Simulation im Automotive-Bereich. *Automatisierungstechnik - at*, 53:425–433, September 2005.
- [6] D. E. Gustafson and W. C. Kessel. Fuzzy Clustering With the Fuzzy Covariance Matrix. In *Proceedings of the IEEE CDC*, pages 761–766. IEEE, 1979.
- [7] O. Nelles, A. Fink, R. Babuska, and M. Setnes. Comparison of Two Construction Algorithms for Takagi-Sugeno Fuzzy Models. *International Journal of Applied Mathematics and Computer Science*, 10:835–855, 2000.
- [8] C. Heij and W. Scherrer. Consistency Of System Identification By Global Total Least Squares. *Automatica*, 35:993–1008, 1999.
- [9] Roorda B. Algorithms for Global Total Least Squares Modelling of Finite Multivariable Time Series. *Automatica*, 31(3):391 – 404, 1995.
- [10] I. Markovskiy, J.C. Willems, S. Van Huffel, B. De Moor, and R. Pintelon. Application of Structured Total Least Squares for System Identification. In *43rd IEEE Conference on Decision and Control*, volume 4, pages 3382–3387. IEEE, 2004.
- [11] P. P. N. de Groen. An Introduction to Total Least Squares. *Nieuw Archief Voor Wetkunde*, 14:237, 1996. <http://www.citebase.org/cgi-bin/citations?id=oai:arXiv.org:math/9805076>.
- [12] R. Zimmerschied, M. Weber, and Isermann R. Stationäre und dynamische Motorvermessung zur Auslegung von Steuerkenfeldern - Eine kurze Übersicht. *Automatisierungstechnik - at*, 53:87–94, Februar 2005.
- [13] Norman R. Draper and Harry Smith. *Applied Regression Analysis*. Wiley, 3rd edition, 1998.

Manuskripteingang: 23. Mai 2006.

Dr. Stefan Jakubek ist Assistent am Institut für Mechanik und Mechatronik der Technischen Universität Wien. Hauptarbeitsgebiete: Fehlerdiagnose, Systemidentifikation und Modellierung mit Hilfe von Neuronalen Netzen.

Adresse: Technische Universität-Wien, Institut für für Mechanik und Mechatronik, Gußhausstr. 27-29/E325 A5, A 1040 Wien
E-Mail: jakubek@impa.tuwien.ac.at

DI Christoph Hametner ist Assistent am Institut für Mechanik und Mechatronik der Technischen Universität Wien. Hauptarbeitsgebiete: dynamische Prozessidentifikation mit neuronalen Netzen

Adresse: Technische Universität-Wien, Institut für für Mechanik und Mechatronik, Gußhausstr. 27-29/E325 A5, A 1040 Wien
E-Mail: hametner@impa.tuwien.ac.at

Dr. Nikolaus Keuth ist Entwicklungsingenieur bei AVL List GmbH im Bereich Messtechnik und Optimierungstechnologien. Hauptarbeitsgebiete: Statische und dynamische Modellbildung, Optimierung mit Hilfe Neuronaler Netze sowie statistische Versuchsplanung.

Adresse: AVL List GmbH, Hans-List Platz 1, A-8020 Graz
E-Mail: Nikolaus.Keuth@avl.com

DI Andreas Voigt ist Assistent am Institut für Mechanik und Mechatronik der Technischen Universität Wien. Hauptarbeitsgebiete: modellbasierte Fehlerdiagnose und Regelung von Großgasmotoren

Adresse: Technische Universität-Wien, Institut für für Mechanik und Mechatronik, Gußhausstr. 27-29/E325 A5, A 1040 Wien
E-Mail: voigt@impa.tuwien.ac.at