Datenbasierte Generierung von situationsbezogenen Entscheidungsregeln für die Kraftwerksführung

Dipl.–Ing. Daniel Schauten, Dipl.–Phys. Timo Slawinski und Prof. Dr. rer. nat. Harro Kiendl

Lehrstuhl für Elektrische Steuerung und Regelung
Fakultät für Elektrotechnik und Informationstechnik
Universität Dortmund
44221 Dortmund
Tel.: 0231 755–4524 Fax: 0231 755–2752
E–Mail: {Schauten, Slawinski, Kiendl}@esr.e-technik.uni-dortmund.de

Zusammenfassung

Das Fuzzy–ROSA–Verfahren stellt ein Rahmenkonzept zur datenbasierten Generierung von Fuzzy–Systemen dar. Zur weiteren Ausgestaltung dieses Rahmens wird das Verfahren anhand des hoch komplexen Anwendungsproblems der datenbasierten Generierung von Entscheidungsregeln für die Kraftwerksführung erprobt. Es werden situationsbezogene Handlungsvorschläge von Prozessexperten hinsichtlich der Einstellung der Werte von drei ausgewählten Strategieparametern eines Kraftwerksführungssystems zur Verbesserung des Prozessverhaltens modelliert. Es zeigt sich, dass die Handlungsvorschläge bezüglich der Einstellung eines dieser drei Parameter mit dem Fuzzy–ROSA–Verfahren in seiner derzeitigen Form sehr gut modellierbar sind. Bei der Modellierung der Handlungsvorschläge für die anderen beiden Strategieparameter stößt das Verfahren jedoch auf Anwendungsgrenzen, die wesentlich damit zusammenhängen, dass die zugehörigen Daten ausgeprägte Widersprüchlichkeiten aufweisen. Von diesem Ergebnis ausgehend, werden neue methodische Ansätze zur weiteren Ausgestaltung des Fuzzy–ROSA–Verfahrens zur Beherrschung auch sehr widersprüchlicher Daten aufgezeigt.

1 Einführung und Aufgabenstellung

Eine zentrale Aufgabe von Energieversorgungsunternehmen besteht in der Kraftwerksführung zur Gewährleistung einer jederzeit sicheren, ausreichenden und kostengünstigen Energieversorgung der angeschlossenen Verbraucher. Diese Aufgabe wird vom Kraftwerksführungssystem übernommen, das die Sollwerte für die Leistungserzeugung aller momentan verfügbaren Kraftwerke online vorgibt. Innerhalb des Kraftwerksführungssystems übernimmt ein *Optimierungsmodul* die Aufgabe, die Sollwerte für die einzelnen Kraftwerke unter Berücksichtigung des prognostizierten Leistungsbedarfs so einzustellen, dass die erforderliche Leistung insgesamt unter Einsatz möglichst geringer Brennstoffkosten erbracht wird und dass eine übermäßige Beanspruchung der Kraftwerke durch Vorgabe möglichst ruhiger Sollwertverläufe vermieden wird. Aufgrund des stochastischen Verhaltens der Verbraucher treten unvermeidliche Differenzen zwischen der Prognose und der tatsächlichen Leistungsnachfrage auf. Zur Vermeidung eines notwendigen Ausgleichs dieser Differenzen durch einen unerwünschten Leistungsaustausch mit den Netzverbundpartnern modifiziert ein schnell aktivierbares Modul zur *Leistungsfrequenzregelung* die Sollwerte der hierfür verfügbaren Kraftwerke. Voraussetzung für die Erfüllbarkeit dieser Regelungsaufgabe ist allerdings, dass die jeweils in Betrieb befindlichen Kraftwerke noch über eine ausreichend große Leistungsreserve verfügen, die zudem ausreichend schnell entbunden werden kann.

Eine besondere Schwierigkeit dieser Prozessführungsaufgabe liegt darin, dass der betrachtete Kraftwerkspark aus unterschiedlichen Kraftwerkstypen besteht, die sich hinsichtlich ihrer spezifischen Erzeugungskosten und dynamischen Eigenschaften stark unterscheiden. Kohlekraftwerke weisen beispielsweise im Vergleich zu Gaskraftwerken geringere spezifische Erzeugungskosten auf. Gaskraftwerke hingegen haben den Vorteil, dass sie schneller auf Sollwertänderungen reagieren können. Deshalb kann die elektrische Energie durch bevorzugten Einsatz der kostengünstigen Kraftwerke um so wirtschaftlicher erzeugt werden, je frühzeitiger und genauer das Optimierungsmodul über zukünftige Leistungsänderungen anhand der Lastprognose informiert wird. Ebenso wirkt es sich kostensenkend aus, wenn die für die Regelung vorgehaltene Leistungsreserve nicht unnötig groß bemessen wird.

Zur Lösung dieser komplexen Kraftwerksführungsaufgabe wurde in Vorprojekten ein neues Verfahren entwickelt, das gegenüber dem Stand der Technik wesentliche Vorzüge aufweist [1, 2, 3, 4]. Ein zentrales Element des Verfahrens besteht darin, dass die für die Regelung vorzuhaltende Reserve nicht anteilmäßig einzelnen Kraftwerken fest zugewiesen wird, sondern statt dessen die Forderung nach einer summarischen Reserve (*Regelband*) erhoben und diese situationsabhängig den beteiligten Kraftwerken zugewiesen wird. Dieses Konzept der summarischen Reserve schafft eine größere Flexibilität, in der ein wesentliches Kosteneinsparpotenzial liegt. Die eigentliche Optimierung, d. h. die Aufteilung der Prognose auf die einzelnen Kraftwerke unter Freihaltung der Reserve, wird durch ein vorausschauendes Verfahren gelöst, das mit Rücksicht auf die erforderliche Onlinefähigkeit suboptimal arbeitet. Darüber hinaus verfügt das Verfahren über eine Reihe von Strategieparametern (wie beispielsweise die Bemessung der vorzuhaltenden summarischen Reserve), die nach dem bisherigen Entwicklungsstand des Verfahrens auf solche Werte, fest oder nach einem Fahrplan, eingestellt werden, die sich im Mittel als günstig erwiesen haben.

Das Ziel dieser Arbeit besteht einerseits darin, auf datenbasiertem Wege mit Hilfe des Fuzzy–ROSA–Verfahrens ein Adaptionsmodul zu entwickeln, das eine situationsabhängige Anpassung der Strategieparameter vornimmt. Dadurch soll eine Verbesserung des Systemverhaltens hinsichtlich der resultierenden Kosten und der Sollwertruhe gegenüber der bisherigen Festeinstellung dieser Werte erzielt werden. Andererseits soll das Anwendungspotenzial des Fuzzy–ROSA–Verfahrens in seinem derzeitigen Entwicklungszustand an diesem hoch komplexen Anwendungsproblem erprobt werden. Insbesondere soll festgestellt werden, bei welchen Problemtypen gute Modellierungsergebnisse zu erwarten sind, bzw. wo sich Anwendungsgrenzen zeigen und ggf. wie diese durch methodische Weiterentwicklungen überwunden werden können.

1.1 Ursprüngliches Modellierungskonzept

Eine wichtige Voraussetzung zur situationsabhängigen Einstellung eines optimalen Parametersatzes ist ein Gütemaß zur Bewertung der Prozessführungsgüte. Wie in [5] gezeigt wird, kann zum Lernen von Adaptionsregeln die Prozessführungsgüte in den Lernprozess des Fuzzy–ROSA–Verfahrens zurückgekoppelt werden. Die prinzipielle Idee besteht darin, dass bezogen auf eine Situation schlechte Parametereinstellungen nicht oder nur abgeschwächt zur Regelgenerierung beitragen und somit im Adaptionsmodul weitestgehend unberücksichtigt bleiben.

Die Bewertung eines Kraftwerksführungssystems ist aufgrund der Vielzahl der Einflussgrößen und der komplexen Abhängigkeiten schwierig. Darüber hinaus müssen auch viele teilweise widersprüchliche Teilgütekriterien berücksichtigt werden. Aus diesen Gründen ist Fuzzy–Logik zur Konzeption eines aussagekräftigen Gütemaßes erfolgreich eingesetzt worden [6]. Darauf aufsetzend wurde die Regelgenerierung mit dem Fuzzy–ROSA–Verfahren in der oben beschriebenen Weise durchgeführt. Allerdings konnte kein Regelsatz gefunden werden, der die Einstellwerte der betrachteten Strategieparameter situationsabhängig so anpasst, dass sich die Prozessführungsgüte nennenswert verbessert.

Diese Schwierigkeit ist möglicherweise darauf zurückzuführen, dass der Grad der oben angesprochenen Suboptimalität beim vorliegenden Optimierungsalgorithmus, wie unten näher erläutert, situationsabhängig ist. Dies kann dazu führen, dass selbst bei fest eingestellten Strategieparametern in ähnlichen Prozesssituationen stark unterschiedliche Gütewerte auftreten. Ein solches widersprüchliches Verhalten spiegelt sich in mehrdeutigen Lerndaten wider. In wieweit diese Mehrdeutigkeiten/Widersprüchlichkeite die Regelgenerierung behindern bzw. zu einer unerwünschten Kompromissbildung führen kann, wird in Abschnitt 5 diskutiert.

Aufgrund der erkannten Unzulänglichkeiten wurde die vom Kraftwerksführungssystem zu lösende Optimierungsaufgabe grundlegend neu strukturiert. Als Ergebnis wurde Sie auf das mathematische lösbare Problem zurückgeführt, zwei quadratische Zielfunktionen, die die Kosten und die Sollwertunruhe beschreiben, unter Einhaltung expliziter linearer Ungleichheitsbedingungen simultan zu optimieren. Dieses Problem wurde für eine Reihe von Betriebssituationen durch offline Optimierung gelöst. Dabei wurde aufgedeckt, dass der Grad der Suboptimalität des originären Optimierungsalgorithmusses situationsabhängig ist. Gelänge es, das neu formulierte Optimierungsproblem in Echtzeit zu lösen, läge damit eine neue Kraftwerksführungsstrategie vor, die bereits ohne situationsabhängige Anpassung der Strategieparameter dem bisherigen überlegen wäre. Die Lösung dieses nunmehr präzise formulierten Problems ist nicht Gegenstand dieses Beitrags und wird anderweitig verfolgt.

1.2 Modifiziertes Modellierungskonzept

Um die oben formulierten Ziele, ausgehend vom derzeitigen Kraftwerksführungssystem, dennoch zu erreichen, wurden die beschriebene Schwierigkeit durch eine Modifikation des Modellierungskonzepts umgangen. Ausgangspunkt hierfür war die Überlegung, dass die Prozessexperten vor Ort aufgrund langjähriger Erfahrung die Prozessführungsgüte in unterschiedlichen Prozesssituationen intuitiv bewerten und daraufhin Vorschläge darüber machen können, wie die Werte bestimmter Strategieparameter zur Verbesserung der Prozessführungsgüte zu verändern sind. Dies betrifft insbesondere drei Strategieparameter:

Anpassung der Basisprognose: Verbesserung der Genauigkeit der langfristigen Lastprognose durch eine geeignete Offsetverschiebung der Basisprognose.

Größe des Regelbandes: Vorgabe der mindestens vorzuhaltenden summarischen Leistungsreserve zur Ermöglichung des Eingriffs der *Leistungsfrequenzregelung*. Durch diese Vorgabe wird der Spielraum für die kostenoptimale Lastaufteilung vorgegeben.

Zusatzreserve: Vergrößerung der verfügbaren Leistungsreserve durch Vorverlegung bzw. Zurückverlegung von An- und Abfahrtzeiten bestimmter Kraftwerke.

Das modifizierte Modellierungskonzept besteht darin, ohne den Umweg über eine Prozessführungsgüte direkt Handlungsvorschläge der Prozessexperten hinsichtlich einer situationsabhängigen Adaption dieser drei Strategieparameter zu modellieren.

Hierzu wurden insgesamt vier Prozessexperten Informationen über die jeweiligen Prozesssituationen in geeigneter Form präsentiert. Daraufhin haben Sie in Ansehung dieser Prozesssituationen Handlungsvorschläge zur Modifikation der Einstellwerte der obigen drei Strategieparameter abgegeben (Abschnitt 2). Ausgehend von der so gewonnenen Datenbasis wurden im Rahmen einer Datenvorverarbeitung die abgegebenen Expertenvorschläge miteinander verglichen. Des Weiteren wurde anhand der Prozessdaten nach Prozessgrößen gesucht, die mit den Vorschlägen der Prozessexperten gut korrelieren und dementsprechend für die Modellierung der Vorschläge der Prozessexperten als relevant anzusehen sind, die aber untereinander nicht redundant sind (Abschnitt 3). Aufsetzend auf diesen Prozessgrößen wurde das Fuzzy–ROSA–Verfahren in seiner derzeitigen Form zur datenbasierten Modellierung der Handlungsvorschläge der Prozessexperten eingesetzt (Abschnitt 4). Dabei wurden bezüglich der *Anpassung der Basisprognose* gute Modellierungsergebnisse erzielt, bei der Modellierung der Vorschläge hinsichtlich der *Größe des Regelbandes* und der *Zusatzreserve* zeigten sich hingegen Anwendungsgrenzen des Modellierungsverfahrens. Zu ihrer Überwindung wurden neue methodische Konzepte zur Datenvorverarbeitung und der eigentlichen Modellierung entwickelt.

2 Akquisition und Aufbereitung der Daten

Zur datenbasierten Modellierung hoch dimensionaler Systeme ist bei der Datenakquisition durch Expertenbefragung zu berücksichtigen, dass der Prozess durch eine große Zahl von Prozessgrößen gekennzeichnet ist, die ein Experte nicht simultan erfassen kann. Deshalb muss eine sinnvolle Auswahl von möglichst wenigen Prozessgrößen, die die Prozesssituationen dennoch angemessen beschreiben, getroffen werden. Zudem müssen diese Größen so präsentiert werden, dass alle Situationen als Grundlage für den Entscheidungsprozess überschaubar und bewertbar sind. Zur Verdeutlichung dieser Problematik sei angemerkt, dass beim vorliegenden Kraftwerksführungssystem eine Datenbasis zur Verfügung steht, die weit über 100 im Minutentakt gemessene Prozessgrößen enthält. Diese Größen sind

Soll- und Istwerte der Kraftwerke, innere Systemgrößen des Kraftwerksführungssystems, Trendgrößen bezüglich zeitlich zurückliegender Werte sowie zukünftiger Prognosewerte und kalendarische Größen wie Tageszeit und Tagestyp.

Zur Informationskondensation wurden zunächst artverwandte Prozessgrößen, wie solche, die zu ähnlichen Kraftwerkstypen gehören, zu einer einzigen Größe zusammengefasst. Ferner wurden Daten durch Mittelwertbildung über Stundenintervalle konzentriert, und es wurde für Trendgrößen ein maximaler Zeithorizont von zwei Stunden festgelegt. Auf diese Weise wurde eine reduzierte Datenbasis, die sich nur noch auf 47 Prozessgrößen bezieht, erzeugt. Diese wurde in einer grafischen Benutzeroberfläche, die sich an vertraute Flussdiagramme des vorhandenen Prozessleitsystems anlehnt, präsentiert. Insgesamt wurden 882 Betriebssituationen in dieser Form aufbereitet und vier voneinander unabhängig agierenden Prozessexperten zur Abgabe ihrer Handlungsvorschläge präsentiert.

3 Datenvorverarbeitung

Im Rahmen der Datenvorverarbeitung wurden zunächst die abgegebenen Handlungsvorschläge zur Bewertung ihrer Aussagekraft miteinander verglichen. Anschließend wurden diejenigen Prozessgrößen bzw. davon abgeleitete Größen ermittelt, die mit den Handlungsvorschlägen der Experten gut korrelieren und daher aussichtsreiche Kandidaten für relevante Einflussgrößen zur Modellierung der Handlungsvorschläge der Experten sind.

3.1 Vergleich der Handlungsvorschläge

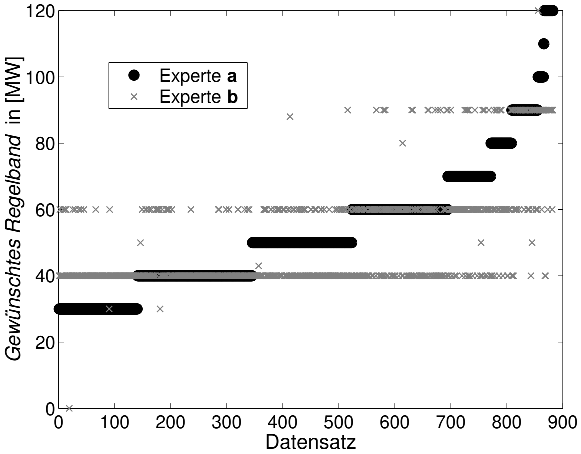
Jeder der vier Prozessexperten **a**, **b**, **c** und **d** machte für insgesamt 882, 882, 135 bzw. 404 unterschiedliche Situationen Vorschläge hinsichtlich der drei Eingriffsmöglichkeiten *Anpassung der Basisprognose*, *Größe des Regelbandes* und *Zusatzreserve*. Um diese Handlungsvorschläge der einzelnen Experten miteinander zu vergleichen, wurden sie für alle möglichen Expertenpaare einer Korrelationsanalyse unterzogen (Tabelle 1).

	Anpassung der Basisprognose				*Größe des Regelbandes*				*Zusatzreserve*			
	a	**b**	**c**	**d**	**a**	**b**	**c**	**d**	**a**	**b**	**c**	**d**
a	1	0.954	0.926	0.966	1	0.816	0.796	0.905	1	0.653	0.487	0.310
b		1	0.930	0.990		1	0.836	0.789		1	0.325	0.348
c			1	0.963			1	0.755			1	0.500
d				1				1				1

Tabelle 1: Korrelationskoeffizienten für die Vorschläge aller möglichen Expertenpaare

Es zeigt sich, dass die Vorschläge aller Experten hinsichtlich der *Anpassung der Basisprognose* sehr ähnlich und damit sehr aussagekräftig sind (Korrelationen größer als 0.9). Die Vorschläge hinsichtlich der *Größe des Regelbandes* und vor allem hinsichtlich der *Zusatzreserve* zeigen demgegenüber deutlich geringere Korrelationen (Ausnahme: Die Vorschläge der Experten **a** und **d** hinsichtlich der *Größe des Regelbandes* korrelieren mit dem Koeffizienten 0.905).

Abbildung 1 verdeutlicht die Unterschiede der Handlungsvorschläge der Experten **a** und **b** bezüglich der *Größe des Regelbandes* und der *Zusatzreserve*. Über der Abszisse sind für alle 882 von beiden Experten bearbeiteten Prozesssituationen die Werte der zugehörigen Expertenvorschläge aufgetragen. Im Interesse einer möglichst großen Übersichtlichkeit wurden die Kraftwerkssituationen nicht chronologisch, sondern nach zunehmenden Vorschlagswerten des Experten **a** sortiert. Diese Darstellung verdeutlicht die großen Unterschiede zwischen den Vorschlägen beider Experten, allerdings zeigen die Tendenzen der Vorschläge noch gewisse Ähnlichkeiten.



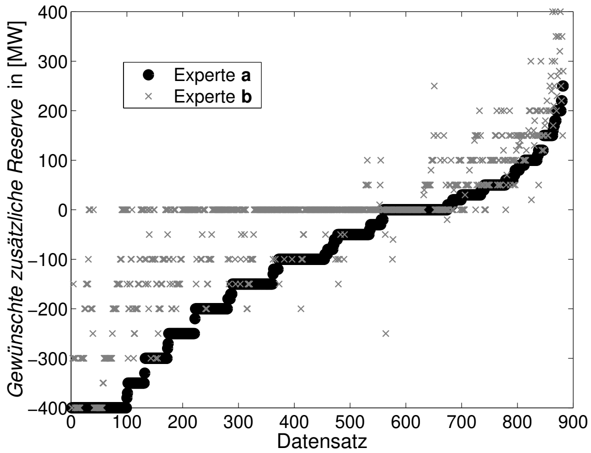


Abbildung 1: Expertenvorschläge zu *Größe des Regelbandes* und *Zusatzreserve* (links bzw. rechts)

3.2 Auswahl relevanter Einflussgrößen

Die Verwendung sämtlicher 47 Prozessgrößen, die im Rahmen der Datenvorverarbeitung ausgewählt bzw. gebildet worden sind, als Grundlage für die datenbasierte Modellierung mit dem Fuzzy–ROSA–Verfahren, führt zu einem großen Suchraum. Zwar ist in [7, 8] gezeigt worden, dass auch für solche komplexe Anwendungen das Fuzzy–ROSA–Verfahren eingesetzt werden kann. Dennoch ist es sinnvoll, überflüssige Prozessgrößen von Anfang an zu eliminieren und damit einen unnötig großen Suchraum zu vermeiden. Hieraus ergeben sich mehrere Vorteile:

- Eine Verkleinerung des Suchraumes auf den tatsächlich relevanten Teil erhöht die Wahrscheinlichkeit dafür, innerhalb einer vorgegebenen Rechenzeit einen möglichst guten Regelsatz zu finden.
- Das Entfernen irrelevanter Größen verbessert häufig die Genauigkeit des resultierenden Fuzzy–Modells.

Für die Auswahl der Prozessgrößen wurde das in [9] vorgestellte Verfahren zur datenbasierten Komplexitätsreduktion eingesetzt. Seine erste Stufe besteht darin, für jede der ursprünglich 47 Prozessgrößen anhand der aufgezeichneten Daten zu untersuchen, ob sie in dem Sinne relevant sind, dass die Expertenvorschläge von ihnen abhängen. In der zweiten

Stufe werden sämtliche Paare von Prozessgrößen daraufhin untersucht, ob sie in dem Sinne redundant sind, dass eine Größe von der anderen abhängig ist. Abweichend von [9] wurde für die Untersuchung zweier Größen auf Abhängigkeit nicht durchgängig das dort entwickelte Eindeutigkeitsmaß, das im Gegensatz zur Korrelation nicht nur lineare, sondern auch nichtlineare Abhängigkeiten entdeckt, verwendet. Statt dessen wurde zunächst mit der Korrelation gearbeitet, die einfacher berechenbar ist und zudem auch keine Vorgabe einer Granularisierung der untersuchten Größen voraussetzt. Diese Modifikation führt zu einer Zeitersparnis, aber zu keinem Informationsverlust, denn nichtlineare Abhängigkeiten werden zwar von der Korrelation nicht erkannt, aber von dem nachgeschalteten Eindeutigkeitsmaß entdeckt. Für die vorgeschaltete Analyse, basierend auf der Korrelation, wurden Größen als relevant angesehen, deren Korrelationskoeffizient bezüglich der Ausgangsgröße größer als 0.5 ist. Ferner wurden zwei Größen als redundant angesehen, deren zugehöriger Korrelationskoeffizient größer als 0.9 ist. Bei dieser Voranalyse wurden beispielsweise 12 redundante Prozessgrößenpaare gefunden, die jeweils Soll- und dazugehörige Istwerte einer Kraftwerkseinheit bzw. stündlich aufeinander folgende Leistungswerte sind.

Bei der nachgeschalteten Anwendung des Verfahrens zur datenbasierten Komplexitätsreduktion wurden durch Variation der Akzeptanzschwellen unterschiedlich umfangreiche Sätze von Prozessgrößen generiert und für die eigentliche Modellierung verwendet.

Da a priori nicht bekannt ist, ob die Experten bei der Abgabe ihrer Vorschläge nur die Werte der Einzelgrößen oder auch die daraus abgeleiteten sekundären Größen berücksichtigen, wurde von Hand nach abgeleiteten Größen gesucht, die möglicherweise stärker als Einzelgrößen mit der Ausgangsgröße korrelieren. Dabei wurde gefunden, dass sich durch Differenzbildung bestimmter Prozessgrößen, die nur schwach mit der Ausgangsgröße korrelieren, neue Größen bilden lassen, die stark mit der Ausgangsgröße korrelieren. Die so gefundenen Größen wurden den oben ermittelten Größen hinzugefügt.

4 Modellierung der Expertenvorschläge

Zur Modellierung der Expertenvorschläge wurde für jede der drei genannten Strategieparameter mit Hilfe des Fuzzy–ROSA–Verfahrens ein separates Teilmodell generiert. Dabei wurden konkurrierend zueinander unterschiedliche Mengen von Prozessgrößen als Modelleingangsgrößen verwendet und die resultierenden Modellgüten miteinander verglichen. Nach einer kurzen Einführung in das Fuzzy–ROSA–Verfahren (Abschnitt 4.1) wird die eingeschlagene Vorgehensweise zur Modellierung der Expertenvorschläge am Beispiel des Strategieparameters *Anpassung der Basisprognose* detailliert beschrieben (Abschnitt 4.2). Die Modellierungsergebnisse für die *Größe des Regelbandes* und die *Zusatzreserve* werden in den Abschnitten 4.3 und 4.4 beschrieben. In diesen Abschnitten wird aufgezeigt, welche Anwendungsgrenzen des Fuzzy-ROSA-Verfahrens in seinem derzeitigem Entwicklungsstand sich bei der Behandlung des vorliegenden Anwendungsproblems zeigen. In Abschnitt 5 werden neue methodische Ansätze skizziert, die auf die Überwindung dieser Anwendungsgrenzen abzielen.

4.1 Das Fuzzy–ROSA–Verfahren

Die Grundidee des Fuzzy–ROSA–Verfahrens [10, 11, 12, 13, 14, 7] besteht darin, einen Raum von Hypothesen zu spezifizieren und darin enthaltene Hypothesen anhand der Daten daraufhin zu testen, ob sie statistisch relevant sind und damit einen signifikanten Teilaspekt des zu modellierenden Systems beschreiben. Bei bestandenem Test [13, 15, 16] wird die Hypothese zur Regel erhoben und mit einem Relevanzmaß versehen, das angibt, wie gut die Regel durch die Daten gestützt wird. Alle so gefundenen Regeln werden zur Regelbasis zusammengefasst, die durch nachgeschaltete Reduktionsmechanismen noch verkleinert werden kann. Da die so generierten Regeln im Sinne des Regeltestverfahrens relevant und damit lokal vernünftig sind, erhält man auf diese Weise transparente und interpretierbare Regelbasen. Ein wesentliches Kennzeichen des Fuzzy–ROSA–Verfahrens ist zudem, dass damit auch generalisierende Regeln gefunden werden können. Diese Regeln können auch Bereiche des Eingangsdatenraumes, in denen keine Lerndaten liegen, abdecken. Diese Generalisierungsfähigkeit des Verfahrens erlaubt somit auch die Einsatzmöglichkeit bei unbekannten Situationen. Besonders vorteilhaft erweist sich zudem die Möglichkeit einer Inspektion und Bearbeitung der generierten Regeln. Dadurch können die Experten nachträglich jede einzelne Regel akzeptieren oder korrigieren, sowie weitere Regeln hinzufügen.

Der Generierungsprozess des Fuzzy–ROSA–Verfahrens ist in vier Hauptschritte unterteilt. Für jeden Schritt sind unterschiedliche Strategien verfügbar, um die Modellierung an unterschiedliche Anforderungen anpassen zu können. Im ersten Schritt der *Projektdefinition* werden die Zugehörigkeitsfunktionen für die vorgesehenen Eingangsgrößen des Modells und der Ausgangsgröße festgelegt. Dies kann wissensbasiert, datenbasiert oder heuristisch erfolgen [17]. Ferner wird in diesem Schritt der Hypothesenraum spezifiziert. Dies erfolgt durch Einstellung der maximalen Kombinationstiefe, die die Maximalzahl der in einer Regelprämisse auftretenden linguistischen Variablen quantifiziert. Für den anschließenden Schritt der *Regelgenerierung* kann abhängig von der Größe des Hypothesenraumes eine komplette Suche, eine evolutionäre Suche [18, 19] oder eine Kombination aus beiden Suchstrategien gewählt werden. Nach Generierung des Regelsatzes kann die Anzahl der Regeln durch Offline–Verfahren zur *Regelreduktion* verringert werden [20]. Dabei können verschiedene Anforderungen, wie zum Beispiel eine vollständige Überdeckung aller Eingangssituationen, eine gleichmäßige Ausschöpfung der Daten, ein möglichst kleiner Modellierungsfehler oder eine möglichst kleine Anzahl von Regeln berücksichtigt werden. Im abschließenden Schritt der *Regelsatzanalyse und -optimierung* kann das Modellierungsergebnis analysiert und das Eingangs-/Ausgangsverhalten des erhaltenen Fuzzy–Modells durch Anpassung noch freier Parameter nachoptimiert werden. Hierzu kann auch das Verfahren zur optimierenden Konfliktreduktion, eine Kombination aus Reduktion der Regelanzahl und Eingangs-/Ausgangsoptimierung, angewendet werden [21, 22].

4.2 Modellierung *Anpassung der Basisprognose*

Zur datenbasierten Modellierung der Expertenvorschläge zur *Anpassung der Basisprognose* mit Hilfe des Fuzzy–ROSA–Verfahrens werden zum Vergleich unterschiedliche Sätze von Prozessgrößen als Eingangsgrößen verwendet und die so generierten Fuzzy–Module miteinander verglichen. Die Zugehörigkeitsfunktionen für jede Eingangsvariable wird anhand einer Analyse der Häufigkeitsverteilung, die durch Projektion der Lerndaten auf die zur betreffenden Eingangsvariablen gehörigen Koordinatenachse entsteht, bestimmt. Für die eigentliche Regelgenerierung werden alle im Fuzzy–ROSA–Verfahren implementierten Hypothesentests und Regelbewertungsverfahren eingesetzt. Die Regelgenerierung wird auf den Lerndaten vorgenommen. Zur Bewertung der Modellgüte wird das Modell anhand der Validierungsdaten getestet. Dabei wird die vorhandene Datenbasis im Verhältnis 1 : 1 zufällig in Lern- und Validierungsdaten aufgeteilt. Als Maß für die Modellgüte wird der Korrelationskoeffizient zwischen tatsächlichem Expertenvorschlag und Modellausgang verwendet. Zur Einordnung der resultierenden Fehler wird der mittlere absolute Fehler **a** (Mittelwert) bzw. **b** (Mittelwert) angegeben, der beschreibt, wie stark die einzelnen Handlungsvorschläge im Mittel vom Mittelwert dieser Handlungsvorschläge abweichen.

Was die Wahl der unterschiedlichen Regeltest- und -bewertungsverfahren für die Regelgenerierung angeht, so ergaben sich mit dem mittelwertbasierten Regeltest [15, 16] stets die geringsten Modellfehler. Die mit diesem Test erzielten Ergebnisse zur Modellierung der Vorschläge der Experten **a** bzw. **b** sind in Tabelle 2 beispielhaft für mehrere Modelle angegeben, die für unterschiedliche Wahl der Eingangsgrößen erhalten wurden.

Modell	**Validierungsdaten**		**Lerndaten**	
	MAF	Korrelation	MAF	Korrelation
a(Mittelwert)	90.3 MW	-	90.0 MW	-
a(47)	10.1 MW	0.991	11.2 MW	0.994
a(29)	8.8 MW	0.994	8.6 MW	0.994
a(4)	6.20 MW	0.997	6.09 MW	0.995
a(2)	5.56 MW	0.998	5.41 MW	0.999
b(Mittelwert)	90.3 MW	-	90.0 MW	-
b(47)	9.52 MW	0.990	9.84 MW	0.989
b(29)	6.48 MW	0.992	6.41 MW	0.991
b(4)	4.60 MW	0.994	4.43 MW	0.993
b(2)	2.98 MW	0.997	2.95 MW	0.999

Tabelle 2: Modellierungsergebnisse zur *Anpassung der Basisprognose* für unterschiedlich ausgewählte Modelleingangsgrößen

Die mit **a**(47) und **b**(47) gekennzeichneten Modelle verwenden alle ursprünglichen 47 Prozessgrößen, während sich die Modelle **a**(29) und **b**(29) auf den Satz von Eingangsgrößen beziehen, der sich durch Reduktion gemäß Abschnitt 3.2 ergibt. Dabei wurden für alle diese Modelle nur Regeln zugelassen, deren Prämissen jeweils nur maximal zwei der Eingangsgrößen berücksichtigen. Die Modelle **a**(4) und **b**(4) beziehen sich jeweils nur noch auf Eingangsgrößen, die durch datenbasierte Komplexitätsreduktion nach [9] aus den obigen 29 Eingangsgrößen ausgewählt wurden. In diesem Fall erlaubte die vergleichsweise geringere Anzahl von Eingangsgrößen die Berücksichtigung von jeweils maximal vier Eingangsgrößen in den Regelprämissen.

Die Modelle **a**(2) und **b**(2) schließlich stützen sich auf nur noch zwei Eingangsgrößen. Diese wurden so gefunden, dass der obige Satz von vier Eingangsgrößen durch Hinzufügung abgeleiteter Größen in Form von Differenzen aus einigen der 47 ursprünglichen Größen zunächst vergrößert und anschließend durch Anwendung des Verfahrens zur datenbasierten Komplexitätsreduktion reduziert wurde. Obwohl die resultierenden Modelle sich nur noch auf zwei Eingangsgrößen stützen weisen sie den geringsten Modellfehler auf. Dies unterstreicht die überragende Bedeutung, die der Auswahl bzw. Bildung relevanter Eingangsgrößen zukommt. Die Qualität der Modelle **a**(2) und **b**(2), die mit nur noch 24 bzw. 34 Regeln eine deutlich kleinere Regelbasis als die Modelle aufweisen, die bei Verwendung von mehr als zwei Eingangsgrößen erhalten wurden, wird durch Abbildung 2 unterstrichen. Dargestellt sind analog zu Abbildung 1 für die nach steigenden Werten der Expertenvorschläge sortierten Datensätzen die Ausgangsgrößenwerte des Modells und die zugehörigen Expertenvorschläge. Ebenso wie Tabelle 2 zeigt auch dieses Bild, dass das Modell **b**(2) die Vorschläge des Experten **b** noch besser als das Modell **a**(2) die Vorschläge des Experten **a** wiedergibt.

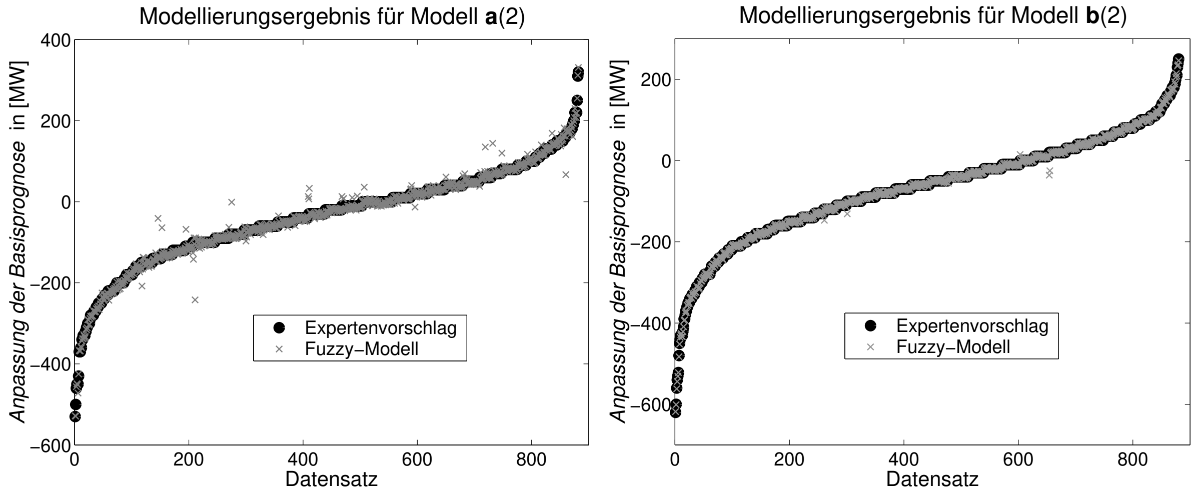


Abbildung 2: Modellierungsergebnisse zur *Anpassung der Basisprognose* für die Experten **a** und **b** (links bzw. rechts)

Da sich die Vorschläge beider Experten auf den gleichen Sachverhalt beziehen, ist der Schluss naheliegend, dass das Modell **a**(2) im Vergleich zum Modell **b**(2) deshalb einen größeren Modellierungsfehler aufweist, weil die Vorschläge des Experten **a** größere Zufallsanteile als die des Experten **b** enthalten.

Aus dieser Beobachtung lässt sich die Feststellung ableiten, dass der Modellfehler allein zur Bewertung eines Modells nicht ausreicht. Vielmehr interessiert neben einer guten Übereinstimmung zwischen Modellausgang und Lern– bzw. Validierungsdaten auch ein ausreichend glattes Modellverhalten: Kleine Änderungen der Eingangssituationen sollen nicht zu unplausibel großen Änderungen des Ausgangsgrößenwerts führen. Wenn ein im Hinblick auf den Modellfehler inkorrekt erscheinendes Modell in diesem Sinne glatter ist, als es die Lerndaten sind, hat man einen Hinweis darauf, dass das Modell tatsächlich korrekter als die Lerndaten ist.

4.3 Modellierung *Größe des Regelbandes*

Analog zur obigen Vorgehensweise wird auch die Modellierung der Handlungsvorschläge zur *Größe des Regelbandes* für stufenweise reduzierte Anzahlen von verwendeten Prozessgrößen durchgeführt. Wegen der geringeren Korrelationskoeffizienten (Abschnitt 3.1) wurden jedoch hier die Daten im Verhältnis 2:1 (statt 1:1) zufällig in Lern- und Validierungsdaten aufgeteilt.

Modell	**Validierungsdaten**		**Lerndaten**	
	MAF	Korrelation	MAF	Korrelation
a(Mittelwert)	15.1 MW	-	15.2 MW	-
a(47)	10.4 MW	0.724	10.1 MW	0.774
a(29)	7.05 MW	0.884	7.13 MW	0.877
a(7)	10.1 MW	0.78	8.11 MW	0.89
b(Mittelwert)	13.0 MW	-	12.9 MW	-
b(47)	10.53 MW	0.691	11.2 MW	0.714
b(29)	7.42 MW	0.861	6.09 MW	0.854
b(7)	4.9 MW	0.94	2.15 MW	0.98

Tabelle 3: Modellierungsergebnisse zu *Größe des Regelbandes* für unterschiedlich gewählte Modellierungsgrößen

Analog zu Tabelle 2 sind die erzielten Modellierungsergebnisse bezüglich der Handlungsvorschläge der Experten **a** und **b** in Tabelle 3 zusammengefasst. Die Modelle **a**(47) und **b**(47) verwenden wieder alle ursprünglichen 47 Prozessgrößen, die Modelle **a**(29) bzw. **b**(29) alle nach der Reduzierung gemäß Abschnitt 3.2 verbliebenen Prozessgrößen. Für beide Fälle wurden wiederum nur Regeln mit Berücksichtigung von maximal zwei Eingangsgrößen zugelassen. Ausgehend von den 29 Größen wurden mittels datenbasierter Komplexitätsreduktion jeweils die 7 relevantesten selektiert. Dabei wurden für beide Experten dieselben Einflussgrößen gefunden. Diese geringe Anzahl von Größen erlaubte es, in den Regelprämissen für die Regeln der Modelle **a**(7) und **b**(7) maximal vier Eingangsgrößen zuzulassen. Die generierten Regelsätze wurden unter Verwendung der Lerndaten mit Hilfe der optimierenden Konfliktreduktion [21] optimiert. Dabei ergibt sich für den Experten **b** eine Verringerung des mittleren betraglichen Modellfehlers (MAF) um 25% auf den Lerndaten, sowie eine Verbesserung um 20% auf den Validierungsdaten (Modell **b**(7)). Für Experte **a** ergibt sich für Modell **a**(7) im Vergleich zu Modell **a**(29) eine verminderte Modellgenauigkeit, die auch mit der optimierenden Konfliktreduktion nicht nennenswert verbessert werden konnte. Dies deutet darauf hin, dass die Reduktion auf 7 Prozessgrößen zu weit gehend ist. Abbildung 3 visualisiert die Qualität der erhaltenen Modelle **a**(7) und **b**(7).

Insgesamt zeigt sich, dass die Handlungsvorschläge beider Experten akzeptabel modelliert werden. Obwohl für den Experten **b** wesentlich geringere Modellfehler erzielt werden als für den Experten **a**, zeigt Abbildung 3 für beide Experten eine ähnlich große Streuung der Modellvorschläge. Hier ist jedoch zu berücksichtigen, dass die Vorschläge des Experten **b** den geringeren Wertebereich von 40 MW bis 90 MW im Wesentlichen durch nur drei diskrete Werte abdecken. Demgegenüber verteilen sich die Vorschläge des Experten **a** in dem größeren Bereich von 30 MW bis 120 MW auf 10 Werte, d. h. sie sind differenzierter.

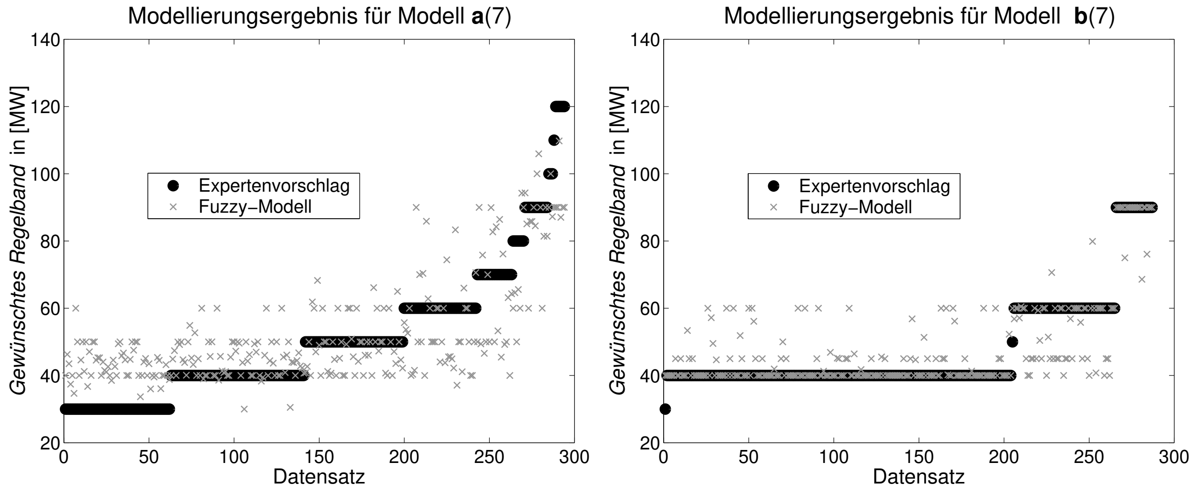


Abbildung 3: Modellierungsergebnisse zu *Größe des Regelbandes* für die Experten **a** und **b** (links bzw. rechts)

Des Weiteren fällt auf, dass die generierten Fuzzy–Modelle häufig auch Werte empfehlen, die zwischen den Expertenvorschlägen liegen, d. h., dass hierfür keine Lerndaten vorhanden sind. Verantwortlich für diesen Effekt ist der zur Regelgenerierung verwendete mittelwertbasierte Regeltest [23, 24]. Dieser Test wurde zur Kompromissbildung im Sinne eines minimalen mittleren quadratischen Modellierungsfehlers bei stark streuenden Ausgangswerten für ähnliche Eingangssituationen konzipiert. Wegen der hier vorliegenden teilweise sehr ausgeprägten Widersprüchlichkeit führt dieser Test jedoch vielfach insofern zu inadäquaten Ergebnissen, als dass die resultierenden Empfehlungen der Regeln sich stark von den in den Daten enthaltenen Werten entfernen.

Wird die Regelgenerierung unter Verwendung der konfidenten Trefferquote [15, 16] als Regeltest wiederholt, der in einer mehrdeutigen Situationen denjenigen Expertenvorschlag herausstellt, der am häufigsten in den Lerndaten vertreten ist, so zeigen sich ebenfalls Anwendungsgrenzen des Fuzzy–ROSA–Verfahrens in seiner derzeitigen Form bei Vorliegen stark widersprüchlicher Daten. Aufgrund der starken Auffächerung der Ausgangsgrößenwerte bei ähnlichen Eingangssituationen stehen häufig nicht mehr ausreichend viele Daten zur statistischen Absicherung der Regeln zur Verfügung, so dass zur Modellierung für diese Situationen u. U. keine Regeln gefunden werden.

4.4 Modellierung *Zusatzreserve*

Zur Modellierung der Handlungsvorschläge zur *Zusatzeserve* wurde zunächst, analog wie oben für die *Anpassung der Basisprognose* beschrieben, vorgegangen. Tabelle 4 zeigt die erhaltenen Ergebnisse.

Die aufgelisteten Modelltypen beziehen sich mit einer Ausnahme auf dieselben Prozessgrößen. Abweichend davon werden bei Modell **a**(21) als Eingangsgrößen nicht die o. a. zusätzlich generierten Größen, nämlich die Soll–Istwert–Differenzen und der Leistungsgradient, verwendet. Diese sekundären Größen wurden erst wieder bei den Modellen

Modell	**Validierungsdaten**		**Lerndaten**	
	MAF	Korrelation	MAF	Korrelation
a(Mittelwert)	134.7 MW	-	134.8 MW	-
a(47)	119.4 MW	0.463	113,8 MW	0.502
a(21)	89.1 MW	0.725	87.9 MW	0.762
a(29)	59.3 MW	0.863	55.3 MW	0.894
a(10)	57.1 MW	0.872	52.9 MW	0.903
b(Mittelwert)	89.3 MW	-	89.4 MW	-
b(47)	81.8 MW	0.581	77.2 MW	0.637
b(29)	58.2 MW	0.775	52.0 MW	0.849
b(10)	32.7 MW	0.868	31.3 MW	0.895

Tabelle 4: Modellierungsergebnisse zu *Zusatzreserve* für unterschiedliche gewählte Modellierungsgrößen

a(29) und **a**(10) verwende. Die dadurch bewirkte deutliche Verringerung des Modellfehlers belegt die Nützlichkeit der Generierung sekundärer Größen aus den primär gegebenen Größen. Ausgehend von den 29 Prozessgrößen wurden durch die datenbasierte Komplexitätsreduktion 10 Größen selektiert. Dabei ergaben sich hier im Gegensatz zu den oben beschriebenen Modellierungen für die Experten unterschiedliche Prozessgrößen und auch deutlich geringere Unterschiede im Relevanzmaß. Die zugehörigen Modelle **a**(10) und **b**(10) wurden unter Verwendung des mittelwertbasierten Regeltests mit einer maximalen Kombinationstiefe von vier generiert.

Die Modelle **a**(10) und **b**(10) weisen den geringsten Modellfehler auf. Gegenüber den Fehlern (**a**(Mittelwert) bzw. **b**(Mittelwert)) ist der Modellfehler auf 50% bzw. 30% reduziert. Die Genauigkeit der Modelle wird allerdings für beide Experten auch hier durch starke Mehrdeutigkeiten bzw. Widersprüchlichkeiten in den Daten beeinträchtigt. Mögliche Ursachen dafür, dass die Experten für vergleichbare Kraftwerkssituationen stark unterschiedliche Handlungsvorschläge machen, sind, dass

- die Strategieparameter *Größe des Regelbandes* und *Zusatzreserve* nicht voneinander unabhängig sind und sich gegenseitig bedingen,
- in die Handlungsvorschläge Expertenwissen und in den Daten nicht enthaltene Fakten eingingen.

Als ein ad hoc Ansatz zur besseren Beherrschung der hier vorliegenden stark widersprüchlichen Daten wurden für die Modellierung nur diejenigen Prozesssituationen berücksichtigt, bei denen die Reserve–Empfehlungen beider Experten um weniger als 60 MW voneinander abweichen, die Experten sich also weitgehend „einig“ waren. Zu jeder dieser Prozesssituationen wurde der Mittelwert der beiden Expertenvorschläge gebildet und als Ausgangsgröße für ein Mittelwertmodell **ab** gewählt. Auf Grundlage von 7 selektierten Einflussgrößen wurde die Regelsuche auch hier mit Kombinationstiefe vier mit dem mittelwertbasierten Regeltest durchgeführt. Das generierte Modell enthält ca. 400 Regeln und liefert einen mittleren betraglichen Fehler von 41 MW auf Trainings- und 44 MW auf Validierungsdaten.

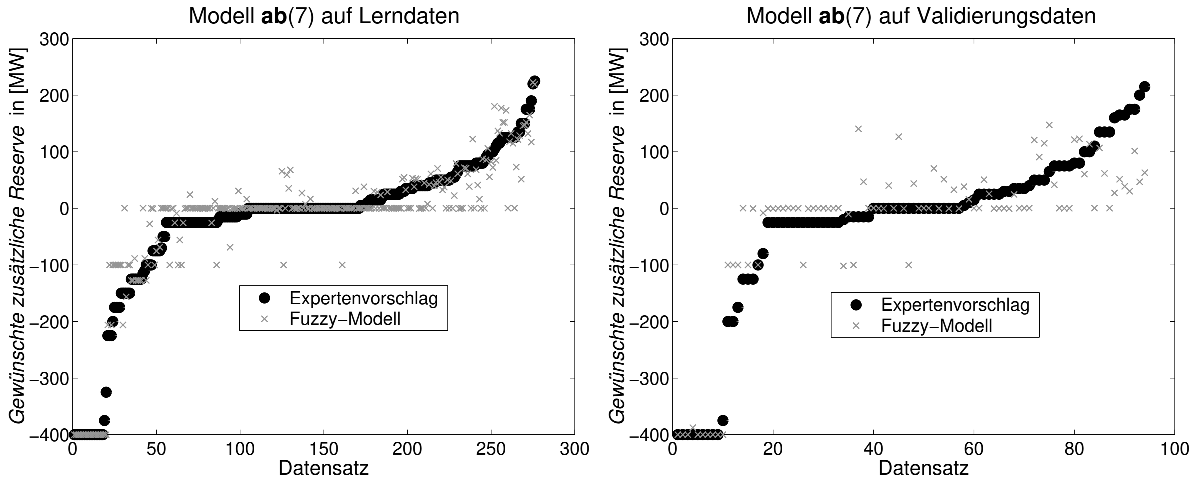


Abbildung 4: Modellierungsergebnisse zur *Zusatzreserve* für das Mittelwertmodell **ab**(7) auf Lern– und Validierungsdaten (links bzw. rechts)

Eine Regelsatzoptimierung mit der optimierenden Konfliktreduktion auf den Lerndaten reduzierte die Regelanzahl auf 65. Dabei wurde der Modellfehler auf den Lerndaten drastisch auf 26 MW verringert und eine Korrelation von 0,96 zwischen Lerndaten und Ausgangswerten des Fuzzy–Modells erreicht. Bezüglich der Validierungsdaten liefert der so optimierte Regelsatz lediglich eine geringfügige Verringerung des Modellfehlers auf 41 MW, bei einer Korrelation von 0,91. Dieser Unterschied der Verbesserungen für Lern- und Validierungsdaten deutet entweder auf Overfitting oder auf eine ungenügende Abdeckung des Eingangsraumes durch die Lerndaten mit der Folge einer nicht ausreichend guten Generalisierungsfähigkeit des generierten Regelsatzes hin.

Die aus Abbildung 4 ersichtliche Modellierungsgüte kann nur als akzeptabel eingestuft werden. Es zeigt sich, dass der aus den gemittelten Handlungsvorschlägen generierte Datensatz noch nennenswerte Mehrdeutigkeiten bzw. Widersprüchlichkeiten aufweist, so dass für einige Situationen keine statistisch relevanten Regeln gefunden werden konnten. In diesen Situationen wird der Handlungsvorschlag 0 MW ausgegeben. Damit wird nochmals die Notwendigkeit zur vereinheitlichten Handhabung mehrdeutiger Daten im Sinne des in Abschnitt 4.3 skizzierten Konzepts unterstrichen.

5 Schlussfolgerungen und Perspektiven für methodische Weiterentwicklungen

Diese Arbeit verfolgte das Ziel, das Fuzzy–ROSA–Verfahren zur datenbasierten Regelgenerierung am Beispiel des komplexen Anwendungsproblems der situationsbezogenen Generierung von Entscheidungsregeln für die Kraftwerksführung einzusetzen. Dabei sollten die Leistungsfähigkeit des Verfahrens erprobt, Anwendungsgrenzen aufgezeigt und ggf. Ansätze für methodische Weiterentwicklungen zur Überwindung der Anwendungsgrenzen entwickelt werden.

Als erstes Ergebnis wurde gefunden, dass es wegen vorher nicht bekannter Eigenarten des Kraftwerksführungssystems sinnvoll ist, das ursprüngliche Konzept zur Generierung der Entscheidungsregeln zu modifizieren: Statt es auf eine Prozessbewertung mit Hilfe eines Gütemaßes zu stützen, erschien der Weg gangbarer, die Tatsache auszunutzen, dass Prozessexperten aufgrund langjähriger Erfahrungen in der Lage sind, situationsabhängige Vorschläge zur Adaption von Strategieparametern zu machen und so erhobene Daten dann mit dem Fuzzy–ROSA–Verfahren zu modellieren.

Dabei wurde festgestellt, dass man in den Fällen, wo die zu modellierenden Daten keine ausgeprägten Mehrdeutigkeiten/Widersprüchlichkeiten aufweisen, mit Standardeinstellungen des Fuzzy–ROSA–Verfahrens, d. h. ohne aufwendige interaktive Eingriffe, sehr gute Modellierungsergebnisse erhält.

Zugleich zeigte sich aber, dass die Auswahl relevanter Eingangsgrößen für die Modellierung von entscheidender Bedeutung ist: Den Prozessexperten wurde als Entscheidungsgrundlage eine Vielzahl von Prozessgrößen präsentiert. Es war nicht vorab bekannt und deshalb aus den Daten zu schließen, welche dieser Größen die Experten tatsächlich nutzen, und darüber hinaus, ob sie auf die Momentanwerte dieser Größen, auf ihren bisherigen Zeitverlauf oder ggf. auch auf neue Größen reagieren, die durch Kombination der ursprünglichen Größen abgeleitet worden sind. Zur Auswahl relevanter und nicht redundanter Größen wurde hier das in [9] vorgeschlagene Verfahren eingesetzt. In seinem derzeitigen Entwicklungsstand arbeitet es allerdings so, dass sich die Relevanzanalyse jeweils nur auf Einzelgrößen bezieht. In den hier behandelten Beispielen hat sich aber gezeigt, dass es wünschenswert ist, auch das Zusammenspiel mehrerer Größen zu berücksichtigen. Es erscheint daher lohnend, existierende Verfahren wie [9] mit folgender Zielrichtung zu erweitern: Aus eine gegebenen Menge von Prozessgrößen ist eine möglichst nicht zu umfangreiche Teilmenge von Prozessgrößen auszusondern, die insgesamt für die intendierte Modellierung eine hohe Relevanz besitzt. Ein solches Verfahren ließe sich organisch dahingehend erweitern, dass aus den primär gegebenen Prozessgrößen durch Verknüpfungen, wie Summen-, Differenz-, Produkt- und Quotientenbildungen, neue Größen gebildet und dem primär gegebenen Satz hinzugefügt werden. So ließen sich Sätze von relevanten Prozessgrößen finden, die neben den ursprünglichen Prozessgrößen auch daraus abgeleitete Größen enthalten.

Weiterhin wurde die Notwendigkeit erkannt, über ein Maß zu verfügen, mit dem die Glattheit mehrdimensionaler Wirkungszusammenhänge bewertet werden kann. Damit kann man vor Beginn der eigentlichen Modellierung die Daten bewerten und ggf. selektiv glätten. Ebenso kann man damit die resultierenden Modelle bewerten: So könnte von zwei konkurrierenden Modellen dasjenige vorzuziehen sein, das zwar einen etwas größeren Modellfehler, dafür aber ein glatteres Verhalten zeigt. Denn wenn man die Modellausgangsgröße wie hier für situationsabhängige Prozesseingriffe nutzen will, ist ein Modell unerwünscht, das zu stark variierenden, schnell aufeinander folgenden Prozesseingriffen führt. Noch zwingender ist die Forderung nach einem glatten Modellverhalten, wenn das Modell zur Prozessbewertung als Grundlage für eine nachgeschaltete Optimierung dient. Insbesondere kann damit vermieden werden, dass die Optimierung in die falsche Richtung läuft. Ein erster Vorschlag zur Entwicklung eines Glattheitsmaßes wird in [25] vorgestellt.

Bei der Modellierung der Expertenvorschläge hinsichtlich der Strategieparameter *Größe des Regelbandes* und *Zusatzreserve* zeigten sich Anwendungsgrenzen des Fuzzy–ROSA–Verfahrens in seinem derzeitigen Entwicklungsstand, weil die Daten in diesen Fällen ausgeprägte Widersprüchlichkeiten/Mehrdeutigkeiten aufweisen. Wenn diese auf Störungen zurückzuführen sind, kann man sie vor der eigentlichen Modellierung geeignet glätten oder bei der Modellierung mit dem Fuzzy–ROSA–Verfahren den eigens für diese Fälle entwickelten mittelwertbasierten Test einsetzen. Auf beiden Wegen werden aber die in den Daten vorhandenen Widersprüchlichkeiten/Mehrdeutigkeiten durch eine mehr oder minder transparente Kompromissbildung aufgelöst. Dies ist inadäquat, wenn in der betreffenden Anwendung unterschiedliche Entscheidungen bei gleichen Eingangssituationen korrekt, Kompromisswerte aber nicht akzeptabel sind.

Zur besseren Beherrschung von Mehrdeutigkeiten/Widersprüchlichkeiten in den Daten wurde das Konzept der impliziten Modellierung entwickelt [26]. Damit wird in einem ersten Modellierungsschritt ein sogenanntes implizites Modell erzeugt, das die Verteilung der Lerndaten im Eingangs–Ausgangsraum beschreibt. Hierbei ist entscheidend, dass die Aufstellung des impliziten Modells durch etwaige Mehrdeutigkeiten/Widersprüchlichkeiten in den Daten nicht tangiert wird, da der Zusammenhang, den das implizite Modell zu beschreiben hat, stets eindeutig ist. In den Daten vorhandene Widersprüchlichkeiten/Mehrdeutigkeiten behindern also die Aufstellung des impliziten Modells nicht, sondern werden von diesem erfasst. Hierauf aufsetzend werden die Mehrdeutigkeiten/Widersprüchlichkeiten in einem zweiten Modellierungsschritt durch Wahl einer geeigneten Defuzzifizierungseinrichtung situationsgerecht aufgelöst. Dabei ist entscheidend, dass diese Wahl einer geeigneten Defuzzifizierung erst nach erfolgter Erstellung des impliziten Modells erfolgen muss. Dies ermöglicht auch das Ausprobieren von Varianten, ohne dass eine neue Modellierung erforderlich wäre. Schließlich hat sich hier gezeigt, dass die mit herkömmlichen Defuzzifizierungsverfahren erzielbaren Ergebnisse in bestimmten Situationen nicht mit der Entscheidung von Prozessexperten übereinstimmen. Dies gab den Anstoß zur Weiterentwicklung des Defuzzifizierungsverfahrens auf der Basis des Inferenzfilters in dem Sinne, dass situationsabhängig automatisch entschieden wird, ob eine Auflösung von Mehrdeutigkeiten/Widersprüchlichkeiten eher im Sinne eines Kompromisses oder einer Entscheidung als angemessener erscheint [26].

Literatur

[1] Albrecht, J. ; Albers, K. ; Stelzner, P.: Sequentielle vorausschauende Vorsteuerung eines Kraftwerksparks durch ein Führungssystem. In: *at – Automatisierungstechnik* 44 (1996), Nr. 8, S. 381–390

[2] Albrecht, J.: *Vorausschauende optimale Steuer- und Regelstrategien zur Verbesserung der Kraftwerksführung*. Düsseldorf : VDI Verlag, 1997 (Fortschritt–Berichte VDI, Reihe 8, Nr. 616)

[3] Albrecht, J. ; Kiendl, H. ; Michalske, A. ; Albers, K. ; Stelzner, P.: *Verfahren zum Regeln der Leistung eines Kraftwerksparks*, 1995. – Patent, DE 195 10 342

[4] Kiendl, H. ; Albrecht, J.: *Verfahren zur sequentiellen Vorsteuerung eines Prozesses*, 1997. – Patent, DE 195 10 343

[5] Schwane, U.: *Datenbasierte Generierung von Adaptionsregeln und Anwendung zur Erhöhung der Bahngenauigkeit eines Industrieroboters.* Fortschritt–Berichte VDI, Reihe 8, Nr. 748. Düsseldorf : VDI Verlag, 1999

[6] Praczyk, J.: *Entwurf von Fuzzy–Gütemaßen zur Prozeßbewertung.* Düsseldorf : VDI Verlag, 1999 (Fortschritt–Berichte VDI, Reihe 8, Nr. 796)

[7] Slawinski, T. ; Krone, A. ; Krause, P. ; Kiendl, H.: The Fuzzy–ROSA Method: A Stastically Motivated Fuzzy Approach for Data–Based Generation of Small Interpretable Rule Bases in High–Dimensional Search Spaces. In: Last, M. (Hrsg.) ; Kandel, A. (Hrsg.) ; Bunke, H. (Hrsg.): *Data Mining and Computational Intelligence.* Physica-Verlag Heidelberg, 2001, S. 141–166

[8] Slawinski, T. ; Jessen, H. ; Praczyk, J. ; Krause, P. ; Krone, A. ; Kiendl, H.: Einsatz der datenbasierten Fuzzy–Modellierungen für komplexe Anwendungen. In: *Computational Intelligence im indutriellen Einsatz, Tagung Baden Baden, 2000.* Düsseldorf : VDI–Verlag, 2000 (VDI–Berichte, Nr. 1526), S. 119–124

[9] Praczyk, J. ; Kiendl, H. ; Slawinski, T.: Finding Relevant Process Characteristics with a Method for Data–Based Complexity Reduction. In: Reusch, B. (Hrsg.): *Computational Intelligence: Theory and Applications (Proceedings of 6th Fuzzy–Days).* Dortmund : Springer Verlag, 1999, S. 548–555

[10] Kiendl, H. ; Krabs, M.: Ein Verfahren zur Generierung regelbasierter Modelle für dynamische Systeme. In: *at – Automatisierungstechnik* 37 (1989), Nr. 11, S. 423–430

[11] Krabs, M.: *Das ROSA–Verfahren zur Modellierung dynamischer Systeme durch Regeln mit statistischer Relevanzbewertung.* Düsseldorf : VDI Verlag, 1994 (Fortschritt–Berichte VDI, Reihe 8, Nr. 404)

[12] Kiendl, H.: *Fuzzy–Control methodenorientiert.* München : Oldenbourg Verlag, 1997

[13] Krone, A.: *Datenbasierte Generierung von relevanten Fuzzy–Regeln zur Modellierung von Prozesszusammenhängen und Bedienstrategien.* Düsseldorf : VDI Verlag, 1999 (Fortschritt–Berichte VDI, Reihe 10, Nr. 615)

[14] Krone, A. ; Krause, P. ; Slawinski, T. ; Knicker, R.: WINROSA 2.0 and DORA for Windows 6.2. In: *Symposium on System Identification (SYSID 2000), Santa Barbara, USA* Bd. CD–ROM, 2000, S. 54

[15] Jessen, H.: Test and Rating Strategies for Automatic Fuzzy Rule Generation and Application to Load Prediction. In: Mohammadian, M. (Hrsg.): *New Frontiers in Computational Intelligence and its Applications.* IOS Press, 2000, S. 11–21

[16] Jessen, H.: *Test- und Bewertungsverfahren zur regelbasierten Modellierung und Anwendung in der Lastprognose.* Düsseldorf : VDI Verlag, 2000 (Fortschritt–Berichte VDI, Reihe 8, Nr. 836)

[17] Krone, A. ; Slawinski, T.: Data–Based Extraction of Unidimensional Fuzzy Sets for Fuzzy Rule Generation. In: *Proceedings of the Seventh IEEE International Conference on Fuzzy Systems (FUZZ–IEEE '98), Anchorage, USA, 1998* Bd. 2. Piscataway, NJ : IEEE Press, 1998, S. 1032–1037

[18] Krone, A. ; Kiendl, H.: An Evolutionary Concept for Generating Relevant Fuzzy Rules from Data. In: *International Journal of Knowledge–based Intelligent Engineering Systems* 1 (1997), Nr. 4, S. 207–213

[19] Slawinski, T. ; Krone, A. ; Hammel, U. ; Wiesmann, D. ; Krause, P.: A Hybrid Evolutionary Search Concept for Data–based Generation of Relevant Fuzzy Rules in High Dimensional Spaces. In: *Proceedings of IEEE International Conference on Fuzzy Systems (FUZZ–IEEE '99) Seoul, Korea, 1999* Bd. 3. Piscataway, NJ : IEEE Press, 1999, S. 1432–1437

[20] Krone, A.: Advanced Rule Reduction Concepts for Optimizing Efficiency of Knowledge Extraction. In: *Proceedings of the Fourth European Congress on Intelligent Techniques and Soft Computing (EUFIT '96)* Bd. 2. Aachen : Verlag Mainz, 1996, S. 919–923

[21] Krone, A. ; Krause, P. ; Slawinski, T.: A New Rule Reduction Method for Finding Interpretable and Small Rule Bases in High Dimensional Search Spaces. In: *Proceedings of the Ninth IEEE International Conference on Fuzzy Systems, (FUZZ–IEEE '00), San Antonio, USA, 2000* Bd. 2. Piscataway, NJ : IEEE Press, 2000, S. 696–699

[22] Krause, P. ; Krone, A. ; Slawinski, T.: Fuzzy System Identification by Generating and Evolutionary Optimizing Fuzzy Rule Bases Consisting of Relevant Fuzzy Rules. In: *Reihe Computational Intelligence*, CI–84/00, Sonderforschungsbereich 531, Universität Dortmund, 2000

[23] Jessen, H. ; Slawinski, T: Mittelwertbasierter Regeltest und Bewertung für das Fuzzy-ROSA-Verfahren und Anwendung zur Lastprognose. In: *Tagungsband des 8. Workshops Fuzzy Control* VDI/VDE GMA–FA 5.22, Forschungsbericht der Fakultät für Elektrotechnik, Nr. 0298, Universität Dortmund, 1998, S. 67–81

[24] Jessen, H.: A Mean–Value Based Test- and Rating Strategy for Automatic Fuzzy Rule Generation and Application to Load Prediction. In: *International Conference on Computational Intelligence for Modelling Control*. Wien, Östereich : IOS Press, 1999

[25] Kiendl, H.: Implizite Modellierung, inkrementeller Relevanzindex und Rauigkeitsmaß: neue Strategieelemente für die datenbasierte Modellierung. In: *Proceedings 10. Workshops Fuzzy Control des GMA–FA 5.22*. Universtät Dortmund : Wissenschaftliche Berichte Forschungszentrum Karlsruhe (FZKA 6509), 2000, S. 1–14

[26] Kiendl, H. ; Krause, P.: Implicit Modeling to Cope with Ambigous Data. In: *Proceedings of the IASTED International Conference Modeling, Identification and Control (MIC 2001)* Bd. 1, 2001, S. 291–296