

Stoffstromnetzbasierende Optimalplanung des Ressourceneinsatzes industrieller Produktionssysteme

Inauguraldissertation
zur Erlangung des Grades

Doktor der Naturwissenschaften
an der Leuphana Universität Lüneburg

vorgelegt von

Hendrik Lambrecht (Dipl.-Phys.)
geboren am 13. Juni 1977 in Koblenz

Gutachter:

Prof. Dr. Andreas Möller
Prof. Dr. Wolfgang Ruck

Tag der Disputation
22. August 2011

2011

*Für Mathis,
weil er doch schneller war.*

Abstract

Today, many enterprises involved in production are striving for improvements in resource efficiency, for environmental as well as economic reasons. Material flow networks, a method developed in the 1990s, has turned out to be a particularly useful tool with which to support planning processes in the field of industrial energy and material flow management. One drawback to the tool, however, stems from the lack of a method for identifying system configurations with optimal resource usage. To this end, state of the art methods of operations research like mathematical programming could be used. However, their high level of abstractness prevents them from being easily integrated in decision-making processes, particularly in small and medium-sized enterprises.

For this reason, a new approach to optimization of industrial production systems is developed in this thesis that combines the respective advantages of material flow networks and parameter optimization: *material flow network based optimization*. By embedding a material flow network in a parameter optimization problem, the material flow network based optimization problem (MFNOP) is defined. This formal interface between material flow analysis and operations research, however, only results in an expedient planning tool if one succeeds to merge the respective modeling paradigms of material flow analysis and operations research. The MFNOP therefore not only adds automatic system optimization to material flow management, but induces a change in perspective from a traditionally descriptive to *decision-oriented material flow analysis*. In order to map factually or potentially existing action spaces for the operation or design of a production system, one needs to systematically include the corresponding degrees of freedom into the material flow model.

Furthermore, concrete approaches to the solution of the MFNOP are examined. Dependent on the properties of the embedded material flow network there are ways to ensure optimization: One is the algebraic reformulation as a mathematical program and the subsequent application of analytical optimization algorithms; the other is the application of direct search algorithms in a simulation-based framework. An optimization module is conceived that integrates both approaches. By means of numerical experiments the performance of different kinds of direct search algorithms is finally evaluated.

Zusammenfassung

Die Steigerung ihrer Ressourceneffizienz ist heute für viele Unternehmen des produzierenden Gewerbes sowohl aus Gründen des Umweltschutzes als auch zur Steigerung ihrer Wettbewerbsfähigkeit ein wichtiges Anliegen. Als ein in vielfältiger Hinsicht geeignetes Planungsinstrument für das industrielle Energie- und Stoffstrommanagement hat sich in den letzten Jahren die in den 1990er Jahren entwickelte Methode der Stoffstromnetze erwiesen. Ihr Manko besteht hauptsächlich darin, dass bislang keine automatische Suche nach Systemkonfigurationen mit optimalem Ressourcenverbrauch möglich ist. Gängige Optimierungsmethoden des Operations Research wie die mathematische Programmierung sind hierfür prinzipiell geeignet. Aufgrund ihres hohen Abstraktionsniveaus lassen sie sich jedoch nur schwer in betriebliche Planungsprozesse, insbesondere in mittelständischen Unternehmen, einbinden.

Daher wird in dieser Arbeit eine neuartige Methode für die Optimalplanung industrieller Produktionssysteme entwickelt, die die Vorteile der Stoffstromnetze mit denen der Parameteroptimierung verbindet: die *stoffstromnetzbasierter Optimalplanung*. Als formale Schnittstelle zwischen Stoffstromanalyse und Operations Research wird durch die Einbettung eines Stoffstromnetzes in ein Parameteroptimierungsproblem das stoffstromnetzbasierter Optimierungsproblem (SSNOP) definiert. Dabei wird deutlich, dass aus der methodischen Integration nur dann ein sinnvolles Planungswerkzeug resultiert, wenn auch die jeweiligen Modellierungsparadigmen von Stoffstromanalyse und Optimalplanung miteinander verschmolzen werden. Das SSNOP ermöglicht daher nicht nur die angestrebte Erweiterung des Stoffstrommanagements um den Aspekt der automatischen Systemoptimierung. Es induziert darüber hinaus einen Perspektivwechsel von der traditionell eher beschreibenden zu einer *entscheidungsorientierten Stoffstromanalyse*. Nur durch die systematische Berücksichtigung der Freiheitsgrade eines Produktionssystems in Form von vorhandenen oder potenziellen Handlungs- und Gestaltungsspielräumen entstehen Stoffstrommodelle, die gleichzeitig als Optimierungsmodelle einsetzbar sind.

Weiterhin werden in dieser Arbeit konkrete Lösungsansätze für das SSNOP untersucht. In Abhängigkeit von den Eigenschaften des eingebetteten Stoffstromnetzes gibt es zwei Möglichkeiten für die Optimierung: Einerseits die algebraische Reformulierung als mathematisches Programm, welches anschließend mit analytischen Optimierungsalgorithmen gelöst werden kann, andererseits die Anwendung direkter Suchalgorithmen im Rahmen einer simulationsbasierten Optimierung. Weiterhin wird ein Optimierungsmodul für eine stoffstromnetzbasierter Stoffstromanalyse-Software konzipiert, welches beide Lösungsansätze integriert. Mittels numerischer Experimente wird schließlich die Eignung verschiedener direkter Optimierungsalgorithmen zur Lösung des SSNOP untersucht.

Inhaltsverzeichnis

Abstract	i
Zusammenfassung	ii
Inhaltsverzeichnis	iii
1 Einleitung	1
1.1 Ausgangslage und Problemstellung.....	1
1.2 Zielsetzung und Aufbau der Arbeit.....	5
2 Grundlagen der modellbasierten Planung industrieller Produktionssysteme	7
2.1 Modellbasierte Planung im betrieblichen Kontext.....	7
2.1.1 Allgemeine Konzeption der Planung.....	8
2.1.2 Modellbasierte Systemanalyse.....	10
2.2 Das Optimalplanungsparadigma des Operations Research	15
2.2.1 Planung mit Optimierungsmodellen	16
2.2.2 Die mathematische Grundstruktur von Optimierungsmodellen	17
2.2.3 Parameteroptimierung.....	17
2.2.4 Mathematische Programmierung	19
2.2.5 Entscheidungstheoretische Ansätze zur Aggregation mehrerer Extremierungsziele zu einer skalaren Zielfunktion	21
2.2.6 Systemfreiheitsgrade.....	22
2.3 Stoffstromanalysen in der Industrial Ecology.....	23
2.3.1 Das Metabolismusparadigma der Industrial Ecology.....	23
2.3.2 Formale Definition der Stoffstromanalyse.....	26
2.3.3 Stoffstrommanagement als zentrales Handlungsfeld der Industrial Ecology	28
2.4 Stoffstromnetze	29
2.4.1 Strukturanalyse mit Stoffstromnetzen	29
2.4.2 Stoffanalyse und Quantifizierung des Systemzustands.....	31
2.4.3 Prozessmodellierung mit spezifizierten Transitionen	34
2.4.4 Produktionsplanung mit Stoffstromnetzen.....	36
2.5 Zusammenfassung	38

3	Stoffstromnetzbasierter Formulierung von Optimierungsmodellen für den betrieblichen Material- und Energieeinsatz	41
3.1	Das Stoffstromnetz als Kern eines Parameteroptimierungsmodells	42
3.1.1	Das Stoffstromnetz als Optimierungsmodell.....	42
3.1.2	Erweiterungen des Kernmodells um zusätzliche Restriktionen und eine algebraische Zielfunktion zur Bewertung des Systemzustands.....	46
3.1.3	Einbeziehung stoffstromnetzspezifischer Auswertungsmethoden in die Formulierung des Optimierungsmodells	47
3.1.4	Allgemeine Definition des stoffstromnetzbasierten Optimierungsproblems	49
3.2	Modellierung von Handlungs- und Gestaltungsspielräumen	50
3.2.1	Die Knoten des Stoffstromnetzes als isolierte Submodelle.....	51
3.2.2	Freiheitsgrade von Transitionen.....	52
3.2.3	Freiheitsgrade von Stellen	52
3.2.4	Die Elimination von Freiheitsgraden durch Flussrelationen.....	54
3.3	Formulierung ökonomischer und ökologischer Zielgrößen	55
3.3.1	Die betriebliche Stoffbilanz als Bewertungsgrundlage.....	56
3.3.2	Die Präferenzen des Entscheidungsträgers	58
3.3.3	Ökonomische Zielgrößen.....	63
3.3.4	Ökologische Zielgrößen	66
3.3.5	Die Bewertung beliebiger Stoffstromnetze	68
3.3.6	Grenzen der stoffstromnetzbasierter Bewertung betrieblicher Produktion.....	70
3.4	Systematik stoffstromnetzbasierter Optimierungsprobleme.....	70
3.4.1	Ökonomische Zielsetzungen.....	71
3.4.2	Ökologische Zielsetzung: die Minimierung der Umweltbelastung.....	71
3.4.3	Maximierung der Öko-Effizienz	72
3.4.4	Zusammenfassung	73
3.5	Beispielhafte Optimierung eines Verbundproduktionssystems	74
3.5.1	Verschiedene Stoffstromnetze zur Abbildung des Originalsystems.....	75
3.5.2	Optimierungsergebnisse und Interpretation.....	83
3.6	Zusammenfassung	89
4	Lösung des stoffstromnetzbasierter Optimierungsproblems (SSNOP).....	93
4.1	Mathematische und simulationsbasierte Optimierung.....	93
4.1.1	Mathematische Optimierung	93
4.1.2	Simulationsbasierte Optimierung.....	95
4.1.3	Vergleich analytischer und direkter Optimierungsverfahren.....	98

4.2	Algebraische Reformulierung des SSNOP	99
4.2.1	Transformation des Stoffstromnetzes.....	99
4.2.2	Beispielhafte Reformulierung des Verbundproduktionsmodells aus Abschnitt 3.4	103
4.2.3	Transformation stoffstromnetzbezogener Auswertungsmethoden	105
4.3	Konzeption eines Optimierungsmoduls für eine stoffstromnetz-basierte Stoffstromanalyse-Software.....	106
4.3.1	Kombination von simulationsbasierter und mathematischer Optimierung für die Lösung des SSNOPs.....	107
4.3.2	Simulationsbasierte Optimierung.....	108
4.3.3	Mathematische Optimierung	110
4.4	Zusammenfassung	110
5	Leistungsfähigkeit direkter Suchstrategien für die stoffstromnetz-basierte Optimierung	113
5.1	Visualisierung und Leistungsfähigkeitsbewertung numerischer Optimierungsalgorithmen	114
5.1.1	Visualisierung von Optimierungsverläufen	114
5.1.2	Performancekriterien für einzelne Optimierungsläufe	117
5.1.3	Bewertung von Testreihen	118
5.1.4	Software-technische Unterstützung der Testreihenberechnung.....	119
5.2	Beschreibung und Veranschaulichung der Funktionsweise der ausgewählten Optimierungsalgorithmen	121
5.2.1	Complex/Nelder-Mead-Algorithmus (CNM)	122
5.2.2	Evolutionstrategie (ES).....	126
5.2.3	Particle Swarm Optimization (PSO).....	127
5.2.4	Zusammenfassung.....	130
5.3	Performance für verschiedene Testprobleme.....	132
5.3.1	Sphärenmodell	133
5.3.2	Randlösungen – Testproblem nach Bracken und McCormick.....	135
5.3.3	Multimodalität – Testproblem nach Himmelblau	137
5.3.4	Beweglichkeit im Lösungsraum – Korridormodell.....	141
5.4	Abhängigkeit der Konvergenzgeschwindigkeit von der Anzahl der Freiheitsgrade	143
5.5	Einfluss der wichtigsten Algorithmusparameter	145
5.5.1	Abbruchkriterium	145
5.5.2	CNM – Skalierung des Startpolyeders	146
5.5.3	ES – Populationsgröße.....	149
5.5.4	PSO – Einfluss der Trägheit.....	150

5.6	Schlussfolgerungen	152
5.6.1	Quantitative Performancekriterien	152
5.6.2	Qualitative Performancekriterien.....	153
5.6.3	Anwendung der untersuchten Optimierungsalgorithmen auf das SSNOP	155
6	Schlussbetrachtung und Ausblick	157
6.1	Grafik- und simulationsbasierte Optimalplanung industrieller Produktion	157
6.2	Entscheidungsorientierte Stoffstromanalysen	159
Anhang	165
A.1	Mathematische Notation.....	165
A.2	Abkürzungen	167
A.3	LINGO-Code des vollständig transformierten SSNOPs für das Verbundproduktionsbeispiel aus Abschnitt 3.4.....	168
Literatur	171

1 Einleitung

1.1 Ausgangslage und Problemstellung

Die Bedeutung der Ressourceneffizienz für das produzierende Gewerbe

Industrielle Produktion trägt durch den Verbrauch natürlicher Rohstoffe, die Erzeugung nicht nutzbarer Abfälle, sowie Schadstoffemissionen maßgeblich zu vielen Umweltproblemen bei. Weil die Wechselwirkung mit der Ökosphäre im Wesentlichen¹ auf Stoffströmen beruht, setzen moderne Umweltschutzmaßnahmen direkt bei den Stoff- und den damit verbundenen Energieströmen in den Unternehmen an. Dabei wird insbesondere das Ziel verfolgt, dieselbe betriebliche Leistung mit einem geringeren Material- und Energieeinsatz zu erbringen². Für die Steigerung der *Ressourceneffizienz* sprechen aus betrieblicher Sicht jedoch keineswegs nur ökologische Gründe. Der Material- und Energieverbrauch stellt heute im produzierenden Gewerbe mit 47,5% den größten Kostenblock dar (Statistisches Bundesamt, 2010: S.277), welcher zudem in den letzten Jahren kontinuierlich angewachsen ist (IBU, 2010). Die Optimierung ihres Material- und Energieeinsatzes ist heute daher für viele Unternehmen des produzierenden Gewerbes aus ökologischen *und* betriebswirtschaftlichen Gründen ein zentrales Anliegen.

Im einfachsten Fall kann die Ressourceneffizienz einer Produktionsanlage bereits durch eine veränderte *Betriebsführung* gesteigert werden. Häufig sind jedoch *strukturelle Eingriffe* wie das Hinzufügen oder Entfernen technischer Komponenten oder die Schließung von Stoffkreisläufen notwendig. Für die Entwicklung konkreter Maßnahmen zur Steigerung der Ressourceneffizienz muss in jedem Fall der Ausgangszustand eines Produktionssystems hinreichend gut bekannt sein. Bei der Komplexität der meisten Industrieanlagen, die oftmals über viele Jahre hinweg durch sukzessive Anpassungen an wechselnde Rahmenbedingungen gewachsen sind, ist das keineswegs selbstverständlich. Am Anfang eines *betrieblichen Energie- und Stoffstrommanagements* steht daher zunächst die *Stoffstromanalyse*, bei der alle relevanten Stoff- und Energieströme sowie Umwandlungsprozesse und Stoffspeicher eines Produktionssystems erfasst werden.

Während eine unter wirtschaftlichen Gesichtspunkten durchgeführte Stoffstromanalyse sich auf betriebsinterne Prozesse beschränken kann, müssen für die ökologische Bewertung der betriebl-

¹ Ausnahmen sind beispielsweise Lärm, radioaktive Strahlung oder das Problem der Flächenversiegelung (vgl. Braunschweig und Müller-Wenk, 1993: S.32-39).

² Diese Dematerialisierungsstrategie (Bringezu und Moriguchi, 2002:79f) des integrierten Umweltschutzes führt streng genommen nur dann zu einer Umweltentlastung, wenn das Produktvolumen konstant bleibt. Das in den Wirtschaftswissenschaften diskutierte Problem des Rebound-Effektes, d.h. dass aus Produktivitätssteigerungen resultierende Umweltentlastungen durch einen Anstieg der Nachfrage überkompensiert werden, wird hier ausgeklammert.

chen Produktion auch *außerbetriebliche* Produktionsprozesse berücksichtigt werden. Stärker als der *unmittelbare* Rohstoffverbrauch sowie die *direkten* Schadstoffemissionen des Betriebs fallen in der Regel die indirekten Umweltwirkungen durch die Herstellung von Vorprodukten, die Energiebereitstellung sowie die Entsorgung von Produktionsabfällen ins Gewicht. Für eine *vollständige* Bewertung müssen daher im Prinzip alle vom Produktionssystem ausgehenden Stoff- und Energieströme bis an die Grenze der Technosphäre zurückverfolgt werden³. Obwohl an dieser Stelle in der Praxis natürlich Vereinfachungen vorgenommen werden, ist für ökologische Stoffstromanalysen doch in jedem Fall eine überbetriebliche Perspektive notwendig.

Um die bei Stoffstromanalysen anfallenden immensen Datenmengen bewältigen zu können, werden in der Regel Computer eingesetzt. Das Spektrum verwendeter Software-Werkzeuge reicht von einfachen Tabellenkalkulationsprogrammen bis hin zu hoch spezialisierten Anwendungen wie GABI, REGIS oder Umberto®, die insbesondere die grafikbasierte Modellierung von Produktionssystemen ermöglichen und Schnittstellen zu Prozessdatenbanken und ökobilanziellen Wirkungsabschätzungsmethoden bereitstellen (Müller-Beilschmidt, 1997, LfU, 2000, IES, 2010).

Betriebliches Energie- und Stoffstrommanagement mit Stoffstromnetzen

Als ein in vielfältiger Hinsicht geeignetes Planungsinstrument für das industrielle Energie- und Stoffstrommanagement hat sich in den letzten Jahren die in den 1990er Jahren entwickelte Methode der *Stoffstromnetze* erwiesen (Möller, 1993, 1994, 2000, Schmidt, 1997). Stoffstromnetze verwenden für die Abbildung von Produktionssystemen Graphen mit zwei verschiedenen Knotentypen: *Transitionen* repräsentieren die Produktionsprozesse und *Stellen* die Lager eines Produktionssystems. Die Energie- und Stoffströme werden durch gerichtete Kanten zwischen diesen Knoten abgebildet.

Wie andere Stoffstromanalysetechniken eignen sich Stoffstromnetze für die transparente und *konsistente* Erfassung der Stoff- und Energieumsätze industrieller Produktionssysteme. In zwei wesentlichen Punkten unterscheiden sie sich jedoch von den meisten anderen Methoden der Stoffstromanalyse: Zum Einen werden nicht nur Stoffströme, sondern auch *Stoffbestände* abgebildet. So kann die Lagerhaltung in Betrieben und überbetrieblichen Wertschöpfungsketten berücksichtigt werden (Möller und Rolf, 1995: S.38). Zum Anderen lassen sich ganz unterschiedliche Prozessmodelle in die Systemabbildung integrieren. Der *Abstraktionsgrad* bei der Modellierung einzelner Stoffumwandlungsprozesse bewegt sich zwischen dem einer linearen Input-Output-Analyse, wie sie in den Wirtschaftswissenschaften (Kloock, 1969, Leontief, 1986) oder bei Ökobilanzen (Heijungs, 1994, Heijungs und Suh, 2002) vielfach eingesetzt wird, und dem einer detaillierten verfahrenstechnischen Fließbildsimulation beispielsweise mit CHEMCAD oder ASPEN, wie sie

³ Stoffe, die ohne weitere Vor- bzw. Nachbehandlung durch den Menschen zwischen Techno- und Ökosphäre ausgetauscht werden, werden in der Ökobilanzierung (LCA) als Elementarflüsse bezeichnet. Laut ISO 14040 (2006: S.24) ist ein Produktionssystem idealerweise so abzubilden, „dass die Inputs und Outputs an seinen Grenzen Elementarflüsse sind“.

bei der Anlagenentwicklung in der Prozessindustrie üblich ist (Penkuhn, 1997, Fröhling, 2006). Je nach Datenlage und Bedeutung für die untersuchte Fragestellung kann die Funktionsweise einzelner Prozesse daher entweder mit linearen Produktionskoeffizienten (Rezepturen) grob abgeschätzt oder durch komplexe, auch nichtlineare Prozessmodelle wirklichkeitsgetreu simuliert werden.

Die Abbildung der Systemstruktur durch einen Graphen ermöglicht die *Visualisierung* von Systemzuständen durch so genannte Sankey-Diagramme, in denen die Breite der Flussrelationen proportional zur Menge transportierter Stoffe oder Energien dargestellt wird. Komplexe Produktionsanlagen können so übersichtlich dargestellt und Anlagenteile mit besonders hohen Stoffumsätzen – potenzielle Schwachstellen – schnell identifiziert werden. Dies sind wichtige Anhaltspunkte für die Entscheidung, an welchen Stellen eines Produktionssystems Maßnahmen zur Verbesserung der betrieblichen Ressourceneffizienz vorrangig ansetzen sollten (Schmidt, 2008b). Stoffstromnetze eignen sich außerdem für vielfältige *quantitative Auswertungen*. Diese reichen „von einfachen periodenbezogenen Input/Output-Bilanzen über [...] *Produktökobilanzierung* und *Kostenrechnungen* bis hin zu betrieblichen Umweltkennzahlensystemen und Wirkungsabschätzungen von Stoffströmen in der natürlichen Umwelt“ (Möller, 2000: S.2). Sie ermöglichen damit die Berechnung wichtiger betriebswirtschaftlicher und ökologischer Zielgrößen für die Planung des Material- und Energieeinsatzes von Produktionssystemen.

Aufgrund ihrer Anschaulichkeit, der Flexibilität bei der Modellbildung, sowie den vielfältigen Möglichkeiten der visuellen und quantitativen Auswertung, werden Stoffstromnetze heute in vielen Industriezweigen (Hähre, 2000: Kap.5.3, Lang et al., 2003, Geldermann, 2005: Kap.5.3, Schmidt, 2007, Wolff et al., 2009) und insbesondere in der chemischen Industrie (Bode et al., 2007, Thißen, 2010, Viere et al., 2010) als Analyse- und Planungswerkzeug für das betriebliche Energie- und Stoffstrommanagement eingesetzt.

Szenarienanalysen zur „Optimierung“ des Ressourceneinsatzes

Stoffstromnetze können als Beschreibungs- und als Erklärungsmodelle eingesetzt werden. Im ersten Fall wird die Struktur eines Betriebes als Stoffstromnetz abgebildet, sowie alle unbekannt Stoffströme und –bestände erhoben und den Stellen und Kanten des Modells zugeordnet. Auf dieser Grundlage kann der Ist-Zustand übersichtlich dargestellt oder durch entsprechende Kennzahlen bewertet werden. Deutlich weiter reichende Analysemöglichkeiten eröffnen sich hingegen, wenn in den Transitionen des Modells funktionale Zusammenhänge zwischen den jeweiligen Input- und Outputströmen definiert werden. Durch diese so genannte Transitionsspezifikation wird das Stoffstromnetz zum *Erklärungsmodell*, mit dem sich insbesondere bei periodischen Datenerhebungen aufwändige Messungen von Stoffströmen und –beständen durch wesentlich kostengünstigere Modellrechnungen ersetzen lassen. Wichtiger noch im Hinblick auf unternehmerische Planungsprozesse ist die Möglichkeit, mit spezifizierten Stoffstromnetzen *Wirkungsprognosen* zu erstellen. Durch die Abbildung alternativer Systemstrukturen in entsprechend modifizierten Stoffstromnetzen oder die Variation einzelner Modellvariablen können verschiedene *Szenarien* konstruiert und miteinander verglichen werden.

Eine automatische Systemoptimierung mit Stoffstromnetzen ist bislang jedoch nicht möglich. Wenn beim Stoffstrommanagement von Optimierung die Rede ist, so wird darunter üblicherweise die Berechnung einzelner Szenarien mit anschließender Auswahl des besten Szenarios verstanden⁴. Dieser Lösungsansatz stößt jedoch an die Grenzen des praktisch Machbaren, wenn (i) mehrere Verbesserungsmaßnahmen zu evaluieren sind, die sich gegenseitig beeinflussen und deren Wirkung daher nicht isoliert beurteilt werden kann, oder (ii) der Einfluss kontinuierlicher Systemparameter berücksichtigt werden muss. Im ersten Fall wächst die Anzahl zu vergleichender Szenarien *exponentiell* mit der Anzahl untersuchter Maßnahmen, im zweiten Fall ist die vollständige Aufzählung aller Szenarien aufgrund der *unendlichen* Anzahl möglicher Parameterkombinationen ausgeschlossen (Lambrech et al., 2009: S.64).

Optimierungsmethoden des Operations Research

Einen möglichen Ausweg liefern das Modellierungsparadigma und die Analysemethoden des Operations Research (OR): Anstatt mehrere Modelle für verschiedene Szenarien zu entwickeln, werden Planungsprobleme durch ein einziges *Optimierungsmodell* abgebildet, das alle denkbaren Szenarien in Gestalt eines *Lösungsraums* beinhaltet. Dieser Lösungsraum kann dann mit Hilfe spezieller mathematischer Verfahren automatisch nach optimalen Lösungen durchsucht werden.

Am häufigsten werden für *Optimalplanungen* Parameteroptimierungsmodelle eingesetzt, bei denen die verschiedenen Elemente des Lösungsraums durch so genannte Entscheidungsvariablen parametrisiert werden. Die Entscheidungsvariablen bilden die verschiedenen Einflussmöglichkeiten des Entscheidungsträgers auf das geplante System ab. Für die Bewertung verschiedener Lösungen werden die jeweiligen Planungsziele als Funktion dieser Entscheidungsvariablen ausgedrückt. Parameteroptimierungsprobleme werden im OR vorwiegend mit algebraischen Programmiersprachen wie GAMS, AMPL oder LINDO als so genannte *mathematische Programme* formuliert, in denen die Ziele und Restriktionen eines Planungsproblems ausschließlich durch mathematische Gleichungen und Ungleichungen definiert werden. Mathematische Programme können aufgrund ihrer speziellen Struktur mit besonders effizienten Algorithmen gelöst werden. Selbst für sehr umfangreiche Optimierungsmodelle mit vielen tausenden Variablen und Gleichungen können Optima, teilweise sogar exakt, bestimmt werden (Frontline Systems, 2010, Mittelmann, 2010).

Trotz dieser Vorteile beschränkt sich der Einsatz der mathematischen Programmierung in der Prozessindustrie vorwiegend auf Optimalplanungsprojekte in großen Konzernen, bei denen die Modellbildung entweder durch Spezialabteilungen für OR bzw. wissenschaftliches Rechnen oder durch externe Forschungseinrichtungen durchgeführt wird (z.B. Kallrath, 2002, Grossman, 2005, Kannegießer und Günther, 2010, Dutta et al., 2011). In kleineren und mittelständischen Unternehmen (KMU) wird die Optimalplanung dagegen kaum eingesetzt, obwohl gerade in diesen

⁴ Bei der simulationsbasierten Optimierung spricht man in diesem Zusammenhang von einem „ranking-and-selection problem“ (Hong und Nelson, 2009: S.76).

besonders große Optimierungspotenziale vorhanden sind⁵. Was im Hinblick auf die Lösbarkeit von Vorteil ist, erweist sich beim praktischen Einsatz der mathematischen Programmierung nämlich immer wieder als Hemmnis: Der hohe Abstraktionsgrad der mathematischen Notation erschwert einerseits die Modellvalidierung durch die Entscheidungsträger eines Betriebs und andererseits die Vermittlung und damit auch Umsetzung der Optimierungsergebnisse. Für Arbeiter, Techniker, Ingenieure, Werks- oder Produktlinienleiter, die täglich am und mit dem geplanten Produktionssystem arbeiten, sind die als mathematische Programme formulierten Planungsprobleme in der Regel weitgehend unverständlich.

Daher wird im Operations Research bereits seit den 1980er Jahren nach Methoden gesucht, mit denen die Vermittlung zwischen den verschiedenen Akteuren eines Optimalplanungsprozesses verbessert werden kann. Eine Optimalplanung kann nur dann erfolgreich sein, wenn es gelingt, divergierende Problemauffassungen zur Deckung zu bringen und in ein einheitliches Planungsmodell zu integrieren. Jones (1996: S.29) zufolge muss dafür insbesondere „*communication gap*“ zwischen den Entscheidungsträgern, d.h. den Initiatoren und Hauptadressaten einer Optimalplanung, und den mit informationstechnischen und mathematischen Methoden arbeitenden OR-Spezialisten überwunden werden: „The greatest divide, however, usually occurs between those who understand the technical details of mathematics and computer science and those who understand the problem to be solved“. Als Schlüssel hierzu wird der Einsatz *visueller Techniken* für die Formulierung und Auswertung von Optimierungsproblemen gesehen. Im Hinblick auf die Planung industrieller Produktionssysteme haben vor allem das *Visual Interactive Modeling* (Hurrion, 1980, Bell, 1985) und seit der Jahrtausendwende die *simulationsbasierte Optimierung* (Fu et al., 2005) wichtige Beiträge geliefert.

1.2 Zielsetzung und Aufbau der Arbeit

Offensichtlich beinhalten sowohl die stoffstromnetzbasierende Stoffstromanalyse als auch das Optimalplanungsparadigma des Operations Research wichtige Elemente für die Optimierung der Material- und Energieströme in komplexen industriellen Produktionssystemen: Die graphenbasierten Stoffstromnetze bieten einen visuellen, und daher sehr intuitiven Ansatz für Modellierung industrieller Produktionssysteme. Sie verfügen außerdem über Schnittstellen zur Kostenrechnung und der Ökobilanzierung, mit denen leicht ökologische und ökonomische Planungsziele aus dem Stoffstrommodell abgeleitet und quantifiziert werden können. Das Operations Research bietet seinerseits das für die Planung und Steuerung komplexer Systeme besonders geeignete Optimalplanungsparadigma sowie eine Vielfalt möglicher Lösungsansätze und -algorithmen für konkrete Optimierungsprobleme. Daher soll in dieser Arbeit eine Optimalplanungsmethode für den Res-

⁵ Zu diesem Ergebnis kommt eine systematische Auswertung einzelbetrieblicher Beratungen im Rahmen des Impulsprogramms Materialeffizienz der Bundesregierung (Schmidt und Schneider, 2010: S.158f).

sourceneinsatz industrieller Produktionssysteme entwickelt werden, die auf der Verwendung von Stoffstromnetzen für die Systemanalyse beruht.

In Kapitel 2 werden zunächst allgemeine Grundlagen der modellgestützten Planung dargestellt, bevor genauer auf die beiden Planungsansätze eingegangen wird, auf denen die stoffstromnetz-basierte Optimalplanung aufbaut: die Stoffstromanalyse mit Stoffstromnetzen und die Optimalplanung mit Parameteroptimierungsmodellen. Beide Planungsmethoden werden vor dem jeweiligen Hintergrund der Forschungsgebiete dargestellt, in denen sie hauptsächlich angewendet werden. Im Falle der Stoffstromanalyse sind dies die Umweltwissenschaften und hier insbesondere die Industrial Ecology bzw. das Stoffstrommanagement. Im Falle der Optimalplanung das Operations Research.

Darauf aufbauend widmet Kapitel 3 sich der Frage, wie mit Stoffstromnetzen Optimierungsmodelle formuliert werden können. Dabei geht es zunächst um die Definition einer formalen Schnittstelle zwischen Stoffstrommodellen und mathematischen Optimierungsalgorithmen. Es zeigt sich jedoch, dass daraus nur dann ein sinnvolles Analysewerkzeug resultiert, wenn auch die verschiedenen Modellierungsparadigmen von Stoffstromanalyse und Optimalplanung miteinander verschmolzen werden. Das setzt jedoch eine grundlegende Auseinandersetzung mit der Frage voraus, wie Stoffstromanalysen in konkreten Planungs- bzw. Entscheidungssituationen durchgeführt werden sollten. Es wird – sowohl in der Theorie als auch anhand eines fiktiven Fallbeispiels – gezeigt, wie Lösungsräume und Zielgrößen mit Stoffstromnetzen konstruiert und zu sinnvollen Optimierungsmodellen kombiniert werden.

In Kapitel 4 werden mit der mathematischen und der simulationsbasierten Optimierung zwei verschiedene Lösungsansätze für stoffstromnetzbasierende Optimierungsprobleme vorgestellt sowie die jeweiligen Vor- und Nachteile in Abhängigkeit von den Eigenschaften des stoffstromnetz-basierten Optimierungsproblems diskutiert. Auf dieser Grundlage wird das Konzept eines Optimierungsmoduls für eine stoffstromnetzbasierende Stoffstromanalysesoftware skizziert. Kapitel 5 untersucht die Leistungsfähigkeit dreier Optimierungsalgorithmen für die simulationsbasierte Optimierung von Stoffstromnetzen. In Kapitel 6 wird die stoffstromnetzbasierende Optimalplanung schließlich in den jeweiligen Kontext der Industrial Ecology sowie des Operations Research eingeordnet und Schlussfolgerungen im Hinblick auf mögliche Anwendungsfelder, sowie die Grenzen der neu entwickelten Optimalplanungsmethode gezogen.

2 Grundlagen der modellbasierten Planung industrieller Produktionssysteme

Dieses Kapitel befasst sich mit den methodischen Grundlagen der stoffstromnetzbasierter Optimalplanung. In Abschnitt 2.1 wird zunächst das Paradigma der modellbasierten Planung dargestellt. In den folgenden Abschnitten werden das Optimalplanungsparadigma des Operations Research und darauf folgend die Stoffstromanalyse als zentrales Instrument der Industrial Ecology für die Analyse industrieller Metabolismen eingeführt. Abschnitt 2.4 widmet sich der speziellen Stoffstromanalysemethode der Stoffstromnetze, welche als Ausgangspunkt der Methodenentwicklung dieser Arbeit gewählt wurde. Aufbauend auf den in diesem Kapitel dargestellten Grundlagen kann in Abschnitt 2.5 schließlich das Verhältnis von stoffstromnetzbasierter Stoffstromanalyse und Optimalplanung im Hinblick auf die spätere Integration beider Ansätze weiter präzisiert werden.

2.1 Modellbasierte Planung im betrieblichen Kontext

Bei der Planung industrieller Produktionssysteme werden heute überwiegend Computermodelle verwendet. Mit ihnen können planungsrelevante Kennzahlen berechnet, bestimmte Sachverhalte grafisch dargestellt und insbesondere Simulationsrechnungen für geplante Maßnahmen durchgeführt werden. Abgesehen von den vielfältigen Möglichkeiten der *Modellanalyse* trägt jedoch bereits die *Modellbildung* wesentlich zur Strukturierung eines Planungsprozesses bei. Bei der Modellierung des Originalsystems muss zwangsläufig von vielen Details *abstrahiert* werden. Das unterstützt den Planenden dabei, unwichtige Systemeigenschaften auszublenden, die Problemstellung zu präzisieren und konkrete Planungsziele zu formulieren. Außerdem erleichtert die *Komplexitätsreduktion* die Identifikation von Handlungs- bzw. Gestaltungsalternativen. Der Aufbau quantitativer Modelle führt in vielen Fällen außerdem zur Überprüfung verwendeter Datenquellen und reduziert damit Planungsfehler durch fehlerhafte Daten und Informationen.

Beim Neuentwurf eines Produktionssystems besteht offensichtlich keine Alternative zur *modellbasierten Planung*. So hat in der Industrie aufgrund des steigenden Wettbewerbsdrucks und durch die Verfügbarkeit immer leistungsfähigerer Computer die Bedeutung des „virtual prototyping“ beim Produkt- und Anlagendesign in den letzten Jahren stetig zugenommen (Driessen, 2006: S.4). Aber auch für eine Umgestaltung oder veränderte Betriebsführung eines existierenden Produktionssystems bringt die modellgestützte Planung viele Vorteile mit sich: *Produktionsausfälle* durch Umbauarbeiten und Tests werden vermieden und das Risiko *irreversibler Schäden* am Produktionssystem aufgrund unvorhergesehener Effekte einzelner Maßnahmen reduziert. Damit sinkt auch das Risiko für *Fehlinvestitionen*.

In diesem Abschnitt wird zunächst der Begriff der Planung allgemein eingeführt und in den Kontext des unternehmerischen Managements gestellt (§ 2.1.1). Planung beruht auf der gedanklichen Antizipation des Verhaltens geplanter Systeme und ist daher stets auf Modelle derselben angewiesen. In § 2.1.2 werden die in diesem Zusammenhang wesentlichen Aspekte der modellbasierten Systemanalyse eingeführt.

2.1.1 Allgemeine Konzeption der Planung

Planung dient der zielorientierten Gestaltung von Systemen nach dem Willen eines Entscheidungsträgers (Scholl, 2001:8f). Am Anfang jeder Planung steht eine als unbefriedigend empfundene Situation. Ein Planungsproblem besteht zunächst in einer Abweichung zwischen dem tatsächlichen und dem erwünschten Zustand eines Systems. Im Rahmen des betrieblichen Energie- und Stoffstrommanagements wird beispielsweise angestrebt, den Energie- und Materialverbrauch eines Betriebs und die damit verbundenen Umweltbelastungen und Produktionskosten zu reduzieren. Dabei ist in der Regel zunächst unklar, um welchen Faktor und mit welchen Mitteln sich der Ressourcenverbrauch senken lässt. Charakteristisch für den Anfang eines Planungsprozesses ist somit *Unwissen* (Bretzke, 1980: S.33f). Die Aufgabe der Planung besteht in einer solchen Situation darin, die Ausgangssituation eingehend zu analysieren und auf dieser Grundlage geeignete *Maßnahmen* für den Übergang vom Status Quo in einen *besseren* Zustand zu ermitteln. So wird das Planungs- zu einem Entscheidungsproblem, bei dem der Entscheidungsträger zwischen verschiedenen *Optionen* wählen muss. Planung setzt daher zwingend eine *subjektive Wertung* von Zuständen voraus.

Die Phasen des Planungsprozesses

Planung ist ein *Prozess*, der üblicherweise in verschiedene Phasen unterteilt wird (Strohhecker, 1998: S.31, Scholl, 2001: S.12):

- Problemidentifikation und Zielfestlegung
- Informationssuche
- Ermittlung von Handlungsalternativen
- Bewertung der Alternativen
- Auswahl der besten Alternative (Entscheidung)

Empirische Untersuchungen von Entscheidungsprozessen in Unternehmen haben gezeigt, dass diese Phasen nicht streng voneinander getrennt werden können und auch nicht zwangsläufig in der angegebenen Reihenfolge durchlaufen werden. In der Praxis sind unterschiedliche Reihenfolgen sowie Sprünge zwischen den verschiedenen Phasen üblich und im Sinne einer kreativen Lösungsfindung sogar zweckmäßig (Witte et al., 1988). So kann die Bewertung von Handlungsalternativen zu einer erneuten Informationssuche führen, bei der sich beispielsweise neue Handlungsalternativen ergeben. Die verschiedenen Phasen sollten daher eher als Teilaufgaben aufge-

fasst werden, die quasi simultan bearbeitet werden. Ein Planungswerkzeug sollte daher schnelle Übergänge von der einen zur anderen Teilaufgabe ermöglichen.

Planung als Teil des zyklischen Managementprozesses

Der klassische *Managementprozess* umfasst die Aufgaben Planung, Organisation, Führung und Kontrolle (Scholl, 2001: S.10). Der bei einer Planung entwickelte *Plan* enthält konkrete Vorschläge, wie bestimmte Aspekte des ursprünglichen Planungsproblems zu lösen sind. Die erfolgreiche Umsetzung eines Plans ist sowohl eine Frage der Organisation als auch der Führung in einem Unternehmen. Im Anschluss an die Umsetzungsphase kann mit Soll-Ist-Vergleichen im Rahmen des betrieblichen Controllings untersucht werden, inwieweit der entwickelte Plan zur Erreichung des angestrebten Ziels geführt hat. Dabei festgestellte Abweichungen zwischen Plan und Wirklichkeit, sowie neue bzw. präzisere Informationen können zu erneuten Planungsaktivitäten und damit zu einem erneuten Durchgang des Managementzyklus führen.

Reichweiten der Planung

Die Planung ist in Unternehmen hierarchisch organisiert. Auf der obersten Ebene befindet sich die *strategische* Planung, die durch die Entwicklung allgemeiner Ziele und Leitlinien auf die Schaffung und Erhaltung einer wettbewerbsfähigen Produktion zielt (Kiener et al., 2006: S.29f). Bezogen auf das Stoffstrommanagement geht es auf dieser Ebene um sehr allgemeine Fragen wie beispielsweise die geeignete Wahl von Produkten, Produktionsstandorten und -verfahren.

Tabelle 2.1: Reichweiten der verschiedenen Ebenen einer unternehmerischen Planungshierarchie

Planungsebene	Reichweite
strategisch	langfristig (mehr als 5 Jahre)
taktisch	mittelfristig (1-5 Jahre)
operativ	kurzfristig (unter einem Jahr)

Auf der darunter liegenden *taktischen* Planungsebene geht es um die Konkretisierung der vorgegebenen Strategien. Hier wird im Einzelnen festgelegt, welche Arten von Produktionsfaktoren eingesetzt werden können und wie die Produktion organisiert werden kann. Das taktische Stoffstrommanagement beinhaltet Überlegungen zum Umbau bzw. zur Modernisierung vorhandener Prozesse, zur Vermeidung bzw. Verwertung von Abfällen sowie zur Einsparung von Energie (Penkuhn, 1997: S.19). Für eine Verbesserung der Umweltperformance eines Produktionssystems kommen verschiedenen Maßnahmen in Frage. Inputseitig kann nach Möglichkeiten zur Substitution von Primärrohstoffen durch umweltfreundlichere Ersatz- bzw. Sekundärrohstoffe gesucht werden. Verfahrensseitig ist die Erhöhung energetischer bzw. stofflicher Wirkungsgrade, sowie die Vermeidung bzw. Verminderung unerwünschter Emissionen zu überprüfen. Outputseitig können Emissionsminderungstechnologien (End-of-Pipe-Technologien) eingesetzt und Produkti-

onsabfälle der Wiederaufbereitung zugeführt werden (Spengler, 1998: S.45). Maßnahmen auf der taktischen Planungsebene sind in der Regel mit *Investitionen* verbunden.

Auf der untersten Stufe befindet sich die *operative* Planung. Ihr Ziel ist die „optimale Durchführung der Produktion [...] bei vorgegebener technisch-organisatorischer Systemstruktur unter Beachtung der aktuell gültigen rechtlichen, technischen, ökonomischen und ökologischen Rahmenbedingungen“ (Spengler, 1998: S.45). In den Bereich der operativen Produktionsplanung fallen die Festlegung des Produktionsprogramms, die Materialbedarfs- sowie die Ablaufplanung.

Die auf den verschiedenen Planungsebenen gefällten Entscheidungen haben unterschiedliche Tragweiten und sind daher nicht alle gleich wichtig für das Unternehmen. Sie sind – zumindest in größeren Unternehmen – auf verschiedenen Managementebenen angesiedelt. Strategische Entscheidungen fallen im Top-Management, während die taktische und in noch stärkerem Maße die operative Planung produktionsnah erfolgen. Bei der modellbasierten Planung hat dies Auswirkungen auf den Detaillierungsgrad verwendeter Modelle. Während technische Details auf der strategischen Ebene keine Rolle spielen, müssen sie für die darunter liegenden Planungsebenen in zunehmendem Maße in die verwendeten Modelle integriert werden. Penkuhn sieht in einer detaillierten Prozessanalyse eine notwendige Voraussetzung für das operative Stoffstrommanagement: „Um zuverlässige Aussagen über mögliche Potentiale prozessintegrierter Maßnahmen treffen zu können, müssen die Prozessverläufe hinreichend genau bekannt sein und mögliche Einflussgrößen auf den Prozessverlauf systematisch untersucht werden“ (Penkuhn, 1997: S.19).

2.1.2 Modellbasierte Systemanalyse

Bei der Analyse von Produktionssystemen und insbesondere der Planung zukünftiger Eingriffe in deren Steuerung oder Gestaltung werden in der Regel Modelle eingesetzt. In diesem Abschnitt werden die untrennbar miteinander verbundenen Begriffe des Modells und des Systems eingeführt und der Prozess der modellbasierten Systemanalyse näher beleuchtet.

Definition des Modellbegriffs

Definitionen des Modellbegriffs sind so vielfältig wie die Anwendungsgebiete der modellbasierten Systemanalyse. Im Hinblick auf die Steuerung produktionstechnischer Systeme definiert Heidepriem (2000: S.30). „Ein *Modell* M ist die Wiedergabe der in einem bestimmten Zusammenhang wesentlichen Elemente und Relationen eines geplanten oder bestehenden Systems O , das Informationen über dieses System in einer *handhabbaren* Form zur Verfügung stellt.“

Diese Definition beinhaltet alle für diese Arbeit benötigten Aspekte des Modellbegriffs. Modelle können sich entweder im Sinne einer *Abbildung* auf bestehende Systeme, oder im Sinne eines *Vorbilds* auf geplante Systeme beziehen. Modelle sind meistens keine identischen (isomorphen) Kopien, sondern lediglich strukturähnliche (homomorphe) Wiedergaben des Originals. Durch den „Filter“ der Abstraktion werden nur die im Hinblick auf den *Modellzweck* wesentlichen Aspekte der Wirklichkeit berücksichtigt. Neben dem Modellzweck beeinflusst auch das Kriterium der

Handhabbarkeit maßgeblich die Entscheidung, welche Elemente und Relationen des Originals bei der Modellbildung vernachlässigt werden können.

Schließlich beruht Heidepriems Definition des Modellbegriffs auf dem *Systemkonzept*. Sowohl das Original als auch das Modell werden als Systeme mit Elementen und Relationen verstanden und der Modellbildungsprozess als eine Abbildung zwischen beiden Systemen.⁶ Wie beim Modell gibt es auch für den Systembegriff keine einfache, allgemein anerkannte Definition⁷. Daher sollen im Folgenden nur die im Rahmen dieser Untersuchung wesentlichen Grundbegriffe und Eigenschaften von Systemen dargestellt und erläutert werden.

Grundbegriffe des kybernetischen Systemkonzepts

Ein System ist durch seine *Komponenten* und die zwischen ihnen bestehenden *Relationen* bestimmt (vgl. Abbildung 2.1). Die Relationen beschreiben die *Struktur* des Systems. Ein System ist in der Regel Teil übergeordneter Systeme, welche aus seiner Sicht die Systemumwelt bzw. die *Umgebung* bilden. Über die *Systemgrenzen* wirkt die Umgebung auf das System ein; umgekehrt hat auch das System Auswirkungen auf seine Umgebung.

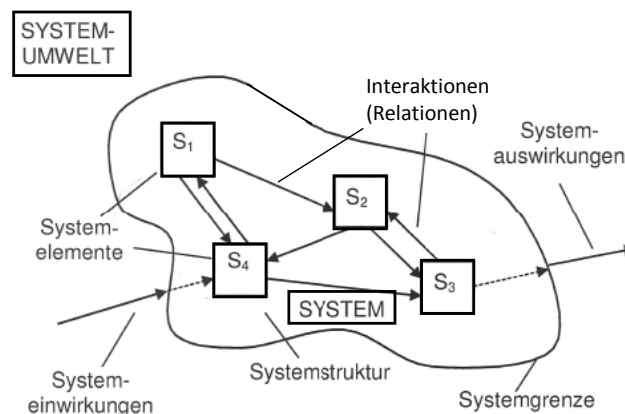


Abbildung 2.1: Bestandteile von Systemen: Komponenten, Relationen, Systemgrenze, Umgebung, Einwirkungen, Auswirkungen (abgeändert aus Bossel, 2004: 36).

Sowohl das System als auch die einzelnen Systemkomponenten besitzen *Attribute*. Da jedes Attribut mindestens zwei verschiedene Werte annehmen kann, repräsentiert es im mathematischen Sinn eine Menge. Je nach erreichbarem Ordnungsgrad innerhalb dieser Menge werden verschie-

⁶ Vgl. hierzu die Definition von Dyckhoff für Modelle in der Produktionstheorie. Ihm zufolge ist Modellbildung „die sinnhafte Abbildung eines oder mehrerer ähnlicher Systeme (Urbilder, Originale) auf ein anderes System (Abbild, Bild)“ (Dyckhoff, 1994: S.23).

⁷ Strohecker (1998: S.105) stellt hierzu fest: „die umfangreiche systemtheoretische Literatur [... wartet], ganz entgegen dem systemtheoretischen Streben nach einer einheitlichen Terminologie, mit fast unüberschaubar vielen, zwar meist ähnlichen, aber eben nicht identischen Definitionen auf“.

dene Attributtypen unterschieden (Tabelle 2.2). Der *Zustand* eines Systems ist durch die Zustände seiner Komponenten festgelegt, also durch eindeutige Werte aller Attribute.

Tabelle 2.2: Klassifikation von Systemattributen nach dem erreichbaren Ordnungsgrad (zusammengestellt nach Niemeyer, 1977: S.40ff).

Typ		Skala		zulässige math. Operationen	Beispiel aus der Stoffstromanalyse
quantitativ	primär	kardinal (metrisch)	Verhältnis	\neq ; $</>$; $+/-$; \times/\div	Stoffströme oder -bestände
			Intervall	\neq ; $</>$; $+/-$	Start- und Endpunkt einer Betrachtungsperiode
qualitativ	sekundär	ordinal (Rangskala)		\neq ; $</>$	Einteilung von Stoffen in Güter, Übel oder Neutra (vgl. Abschnitt 3.3)
	tertiär	nominal		\neq	Materialtyp

Systeme können aus zwei verschiedenen Blickwinkeln betrachtet werden. Aus Sicht der *Systemumgebung* erscheint ein System zunächst als *Black Box* mit einer bestimmten Funktion (Abbildung 2.2). Zur Beschreibung eignet sich das Paradigma des Input-Output-Systems (I/O-System), welches Stimuli aus der Umgebung erfährt (*Input*), und als Reaktion darauf entweder selbst auf die Umgebung einwirkt (*Output*) oder seinen *Zustand* ändert.

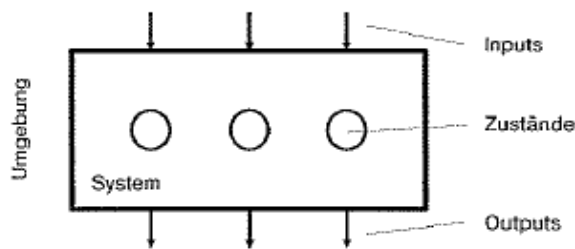


Abbildung 2.2: Black Box: die funktionale Sichtweise von Systemen (Ropohl, 1999: S.76)

Das wesentliche Merkmal von auf dieser Außensicht basierenden I/O-Analysen ist die *Komplexitätsreduktion* durch Ausblendung der inneren Struktur. In diesem Sinne wird Produktion in den Wirtschaftswissenschaften häufig als Black Box beschrieben: Produktionsfaktoren (Inputs) werden verbraucht, um bestimmte Produkte (Output) herzustellen, wobei typischerweise auch Abfälle und Schadstoffemissionen entstehen (Output). Produkte und Abfälle müssen nicht unmittelbar abgesetzt werden, sondern können zunächst in betriebsinternen Lagern verbleiben (Zustandsänderung).

Die komplementäre Sichtweise betont dagegen die *Modularität* von Systemen und die durch die Relationen zwischen den Systemelementen erzeugte *Struktur*. Die Untersuchung dieser Struktur wird als *Systemanalyse* bezeichnet. Das Prinzip der Systemanalyse hat seine Wurzeln im wissenschaftlichen *Reduktionismus*, mit dem in allen wissenschaftlichen Disziplinen „komplexe Phänomene durch Rückführung („Reduktionismus“) auf einfache und einfachste Elemente erklärt und verstanden [wurden. ...] War ein Verständnis der Elemente gewonnen, so wurde ein Verständnis des Ganzen aus dem Zusammenwirken der Elemente abgeleitet“ (Hanssman, 1993: S.9). Die Systemanalyse ist ein iterativer Prozess, der auf der zuvor beschriebenen Black-Box-Sichtweise aufbaut. Im ersten Schritt werden die Grenzen und Außenrelationen eines Systems geklärt. Innerhalb dieser Black Box werden dann relevante Subsysteme herausgearbeitet und immer weiter detailliert. Die Systemanalyse führt durch sukzessive Disaggregationsschritte zu hierarchisch strukturierten Modellen (Abbildung 2.3).

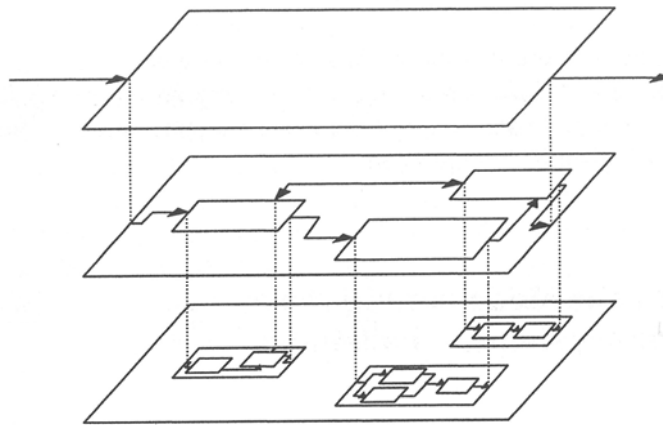


Abbildung 2.3: Hierarchieebenen eines Produktionssystems (nach Penkuhn, 1997: S.28)

Im Prinzip kann eine Systemanalyse so lange fortgeführt werden, bis die Komponenten eines Subsystems unteilbare Elemente sind. Ein solches Idealmodell wird als *White Box* bezeichnet. Aufgrund der begrenzten Genauigkeit verwendeter Analysemethoden, sowie den in der Regel bestehenden zeitlichen oder finanziellen Restriktionen sind in der Praxis so genannte *Grey-Box-Modelle* üblich, in denen Subsysteme, obwohl sie eine innere Struktur besitzen, als Black Box beschrieben werden. Der Grad der Disaggregation richtet sich nach dem Kontext und den Zielen der Systemanalyse.

Modellierung als zweistufiger Prozess – Systemerfassung und Systemabbildung

Niemeyer (1977) untergliedert den Modellbildungsprozess in zwei Stufen (Abbildung 2.4). Bei der *Systemerfassung* entsteht durch Wahrnehmung und geistige Reflektion eines Ausschnitts der Wirklichkeit zunächst das *mentale* Modell. Obwohl es als direktes Ergebnis menschlicher Wahrnehmung entsteht, ist es nicht einfach ein Spiegelbild der Wirklichkeit. Durch vielfältige kreative, vielfach auch unbewusste Leistungen des Gehirns wie Imagination, Assoziation, Extrapolation

usw. kann es weit darüber hinaus gehen. Es ist daher auch kein objektives Bild. Die Systemerfassung wird vielmehr „durch den Willen zur Erkenntnis und Gestaltung gesteuert und weitgehend durch die physische und psychische Kondition des Menschen beeinflusst. Das Resultat [...] sind daher subjektive und zweckorientierte Bilder von Systemen in der menschlichen Vorstellung“ (Niemeyer, 1977: S.38).

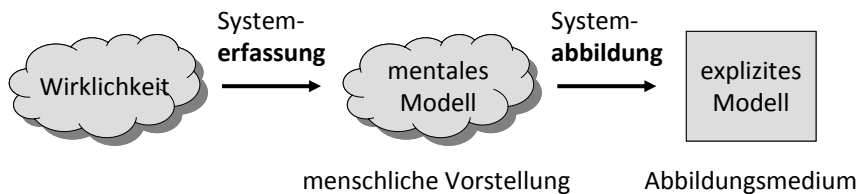


Abbildung 2.4: Zweistufiger Modellbildungsprozess

Bei der anschließenden *Systemabbildung* wird das mentale Modell mittels eines externen Mediums dargestellt und – wenn mehrere Personen an der Modellbildung beteiligt sind – kommuniziert. Sowohl Stoffstromnetze als auch mathematische Programme fallen in diese Kategorie. Sie verwenden für die Abbildung realer Objekten und Zusammenhänge *abstrakte Zeichen*, die nach bestimmten Regeln (Syntax) zum Modell kombiniert werden. Diese Art expliziter Modelle bezeichnet man als *symbolische Modelle*. Die Abbildung der Wirklichkeit beruht auf dem Prinzip, dass „den Zeichen wie auch den Verknüpfungsregeln aus dem jeweils zugrundeliegenden [realen] System entnommene inhaltliche Deutungen gegeben [werden]“ (Meyer, 1996:17). Während syntaktische Regeln meistens strikt definiert sind, da ansonsten bestimmte Methoden der Modellanalyse nicht anwendbar sind, bieten diese semantischen Korrespondenzregeln einen größeren Interpretationsspielraum.

Da bei der Planung Eingriffe in die Wirklichkeit gedanklich *vorweggenommen* werden, ist sie immer auf Modelle angewiesen, um mögliche Konsequenzen des geplanten Handelns zu antizipieren. Die meisten alltäglichen Entscheidungen basieren auf *mentalen* Modellen. Für die Gestaltung oder Steuerung komplexer Systeme sind sie in der Regel jedoch nicht ausreichend. Strohhecker (1998: S.96) nennt dafür mehrere Gründe: Mentale Modelle sind unpräzise. Zusammenhänge zwischen Einflussgrößen werden häufig nur qualitativ erfasst. Sie sind außerdem instabil, da sie sich ständig neuen Erfahrungen anpassen. Da sie nur in der Vorstellung einer einzelnen Person existieren (implizite Modelle), sind sie zudem nicht intersubjektiv überprüfbar. Schließlich sind sie einer systematischen Analyse nur eingeschränkt zugänglich, da das Gehirn ohne Verwendung des relativ langsamen Langzeitgedächtnisses nur eine verhältnismäßig kleine Anzahl von Variablen bzw. Symbolen verarbeiten kann (Miller, 1956). Die Grundlage für komplexe Planungs- und Entscheidungsprozesse bilden daher in der Regel *explizite Modelle*.

Verwendungszweck von Modellen

Modelle werden hinsichtlich ihres Verwendungszwecks üblicherweise in Beschreibungs-, Erklärungs-, Prognose- und Optimierungsmodelle unterteilt (vgl. z.B. Wöhe und Döring, 2000: S.39f).

Beschreibungsmodelle dokumentieren lediglich die Struktur oder den Zustand eines Systems. Ihr Sinn besteht meist darin, komplexe und unübersichtliche Systeme der menschlichen Anschauung zugänglich zu machen. *Erklärungsmodelle* führen den Zustand eines Systems dagegen auf bestimmte Grundannahmen bzw. erklärende Variablen zurück. Sie können daher als *Prognosemodelle* verwendet werden. Man unterscheidet zwischen Wirkungsprognosen, in denen isolierte Systemzustände dargestellt und miteinander verglichen werden, und Entwicklungsprognosen, welche den zeitlichen Verlauf einer Zustandsänderung nachvollziehen. *Optimierungsmodelle* werden mit dem Ziel erstellt, optimale Systemzustände zu ermitteln. Die Übergänge zwischen den verschiedenen Kategorien sind fließend (Abbildung 2.5). Insbesondere beruhen Optimierungsmodelle auf Erklärungsmodellen, da die Wirkung der Veränderung der Entscheidungsvariablen vorhergesagt werden muss. Wirkungsprognosen sind ihrerseits offensichtlich nur auf der Grundlage eines korrekten Beschreibungsmodells möglich.

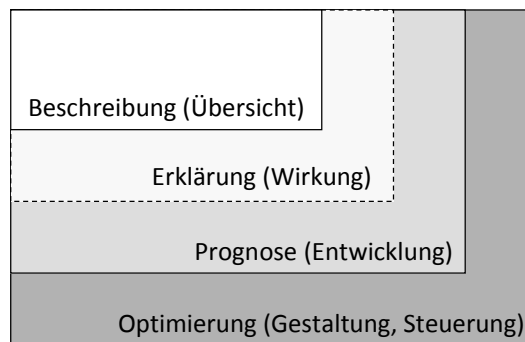


Abbildung 2.5: Klassifikation von Modellen nach dem Verwendungszweck

2.2 Das Optimalplanungsparadigma des Operations Research

Optimalplanung ist eine modellbasierte Planungsmethode, bei der komplexe reale Planungsprobleme in abstrakte, meist quantitative Optimierungsmodelle überführt werden. Die spezielle Struktur solcher Optimierungsmodelle erlaubt die Anwendung mathematischer Verfahren zur *automatischen Suche* nach optimalen Modellzuständen, die Rückschlüsse auf Lösungen für das ursprüngliche Planungsproblem zulassen. Die Anfänge der Optimalplanung reichen in die Zeit des Zweiten Weltkriegs zurück, als britische, amerikanische und sowjetische Militärs erstmals mathematische Modelle für die Planung militärischer Operationen einsetzten. Daher stammt auch die heute noch gebräuchliche Bezeichnung *Operations Research* (OR) für das Forschungsfeld rund um die Anwendung und Methodentwicklung der Optimalplanung⁸. Heute wird die Methoden-

⁸ „Optimalplanung“ ist einer von mehreren erfolglosen Übersetzungsversuchen von „Operations Research“ ins Deutsche (Meyer, 1996: S.2). In dieser Arbeit wird damit jedoch nicht das Forschungsfeld, sondern das Planungsparadigma des Operations Research bezeichnet.

entwicklung des OR überwiegend im zivilen Bereich durch Universitäten oder entsprechende Abteilungen von Großunternehmen vorangetrieben.

2.2.1 Planung mit Optimierungsmodellen

Experimentelle versus modellbasierte Optimierung

Die Optimierung eines Systems S bedeutet im Sinne einer Optimalplanung die Suche nach dem besten Zustand Z^* in der Menge Ω möglicher Systemzustände. Um das System vom Ausgangszustand Z_0 in den optimalen Zustand Z^* zu überführen, müssen diejenigen Systemattribute x_i identifiziert werden, mit denen der Systemzustand gezielt beeinflusst werden kann. Diese *Steuergrößen* bzw. Kontrollparameter müssen frei wählbar sein und den Zustand des Systems eindeutig festlegen: $Z=Z(\mathbf{x})$. Im Prinzip kann die Suche nach dem optimalen Zustand durch Variation der Steuergrößen nach dem Prinzip von Versuch und Irrtum am real existierenden System durchgeführt werden. Derartige *experimentelle Optimierungen* (Schwefel, 1995: S.6) kommen in vielen Fällen jedoch nicht in Frage, entweder da das geplante System noch gar nicht existiert, oder da die mit dem Experimentieren verbundenen Gefahren für das System als zu groß eingeschätzt werden. In diesen Fällen bietet es sich an, mit *Modellen* des Originalsystems zu arbeiten, in denen die Steuergrößen durch *Entscheidungsvariablen* abgebildet werden.

Abbildung von Strukturänderungen des Originalsystems als Zustandsänderungen eines taktischen Optimierungsmodells

Bei der Abbildung eines Planungsproblems als Optimierungsmodell werden verschiedene Handlungsoptionen in Gestalt eines Lösungsraums in das Modell integriert. Handlungsoptionen bestehen im einfachsten Fall in *Zustandsänderungen* des Systems ($S(Z) \rightarrow S(Z')$), die durch bekannte Kontrollparameter gesteuert werden können. Sie können sich jedoch auch auf *strukturelle Veränderungen* des Originalsystems ($S \rightarrow S'$) beziehen. Dieser Unterschied zwischen Zustands- und Strukturänderungen entspricht den verschiedenen Reichweiten der *operativen* und *taktischen* bzw. strategischen Planung. Während bei der operativen Planung das Produktionssystem als gegeben vorausgesetzt und eine optimale Betriebsweise angestrebt wird, bezieht die taktische Planung auch Gestaltungsoptionen mit ein. Entsprechende Optimierungsmodelle sind daher keine bloßen Abbilder der Wirklichkeit, sondern ergänzen diese um (bislang) nicht vorhandene Systemstrukturen. Was in der Realität eine Strukturveränderung ist, wird in taktischen und strategischen Optimierungsmodellen als Zustandsänderung abgebildet. In dieser Konstruktion von Lösungsräumen steckt das kreative Moment einer taktischen Optimalplanung. Die Anwendung mathematischer Optimierungsalgorithmen baut auf diesem Schritt auf und kann natürlich keine besseren als die bei der Modellbildung bereits implizit vorgesehen Lösungen finden.

2.2.2 Die mathematische Grundstruktur von Optimierungsmodellen

Die wesentlichen Bestandteile eines mathematischen Optimierungsmodells sind der Lösungsraum L und die Zielfunktion f . Die Zielfunktion bildet den Lösungsraum auf eine geordnete Menge $(B; <)$ ab: $f: L \rightarrow B$. Die Ordnungsrelation $<$ ist notwendig, um eine eindeutige und vollständige Rangordnung unter allen Elementen von B und damit – durch die Zielfunktion – auch unter allen Elementen von L zu etablieren. $a < b$ bedeutet, dass b in der Rangordnung höher steht als a , gesprochen „ a ist kleiner als b “. Bei der Optimierung wird nun die Lösung $x \in F \subseteq L$ gesucht, für welche die Zielfunktion $f(x)$ den kleinst- bzw. größtmöglichen Wert annimmt. Optimierung ist ein Oberbegriff, der sowohl die Minimierung als auch die Maximierung einer Zielfunktion umfasst.

$$\begin{aligned} \min/\max \quad & f(x) \\ \text{Nb.} \quad & x \in F \end{aligned} \tag{2.1}$$

Das *globale Optimum* ist diejenige Lösung $x^* \in F$, für die gilt

$$f(x^*) < / > f(x) \quad \forall x \in F \tag{2.2}$$

Gilt diese Relation nur für einen Teilbereich von F , so spricht man von einem *lokalen Optimum*. F ist die *zulässige Menge*, und die Elemente von F heißen zulässige Lösungen. Bei $F = L$ wird (2.1) als *unrestringiertes* Optimierungsproblem bezeichnet. Wird dagegen die Menge $F \subset L$ durch entsprechende Nebenbedingungen definiert, ist (2.1) ein beschränktes oder *restringiertes* Optimierungsproblem (vgl. Alt, 2002: S.2). Wenn nichts anderes angegeben ist, wird in dieser Arbeit die Menge der reellen Zahlen als Wertbereich der Zielfunktion aufgefasst: $(B, <) = (R, \leq)$. Außerdem werden ausschließlich so genannte *Parameteroptimierungsmodelle* behandelt (vgl. folgender Abschnitt), deren Lösungen Vektoren des mehrdimensionalen euklidischen Raumes sind ($L = R^n$).

2.2.3 Parameteroptimierung

Entscheidungsvariablen, Koeffizienten und Modellparameter

Das dominierende Modellierungsparadigma für Entscheidungsprobleme ist im Operations Research die Parameteroptimierung. Zur Disposition stehende Steuergrößen des Originalsystems werden durch *Entscheidungsvariablen* x_i beschrieben. Die x_i können als *Indizes* für alle im Rahmen des Planungsvorhabens denkbaren *Alternativen* zur Gestaltung oder Steuerung des Originalsystems verstanden werden⁹.

⁹ Der Vektor der Entscheidungsvariablen *parametrisiert* die Lösungsmenge. Daher die Bezeichnung *Parameteroptimierung*. Sie ist insofern etwas irreführend, als Optimierungsmodelle auch es auch Modellpara-

Daneben enthalten Optimierungsmodelle üblicherweise weitere *quantitative Größen* in Form von Koeffizienten und Parametern. Diese bilden vom Entscheidungsträger nicht beeinflussbare Attribute des Originalsystems ab und bleiben bei der Optimierung konstant. Als *Koeffizienten* werden solche Größen modelliert, die im Rahmen der Problemstellung als unveränderlich angesehen werden können und mit hoher Genauigkeit bekannt sind, beispielsweise technische oder naturwissenschaftliche Konstanten wie Brennwerte, Wärmeleitfähigkeiten, stöchiometrische Koeffizienten etc. Bei einer kurzfristigen Produktionsprogrammplanung werden typischerweise auch Marktpreise als Koeffizienten modelliert, obwohl diese langfristig natürlich Preisschwankungen unterliegen.

Parameter p repräsentieren dagegen Umwelt- oder Systemgrößen, die das Optimierungsergebnis zwar beeinflussen, jedoch nur ungenau bekannt oder veränderlich sind, insbesondere also stochastische Größen. Die Überlegungen dieser Arbeit beschränken sich zwar auf den Fall einer deterministischen Optimierung. Aber auch in diesem Fall ist es sinnvoll, stochastische Größen als Parameter in die Modellierung einzubeziehen, um die Stabilität von Optimierungsergebnissen nachträglich anhand von Sensitivitätsanalysen überprüfen zu können. Als Parameter können aber auch solche Steuergrößen modelliert werden, die aus sachlichen Gründen nicht in die Optimierung einbezogen werden sollen.

Zielgrößen

Zielgrößen sind skalare Maßzahlen zur *Beurteilung des Systemzustands aus Sicht des Entscheidungsträgers*. Bei einem Parameteroptimierungsmodell ergeben sie sich als Funktion der Entscheidungsvariablen und können auch von den Modellparametern abhängen: $\varphi = \varphi(\mathbf{x}; \mathbf{p})$. Typische Zielgrößen bei der Planung industrieller Produktionssysteme sind beispielsweise das Produktionsvolumen, die Produktionskosten oder der Deckungsbeitrag, die produktionsbedingten CO₂-Emissionen oder die eingesetzte Menge (Masse) eines bestimmten Rohstoffes.

Bei der Konstruktion eines Optimierungsmodells können Zielgrößen in unterschiedlicher Weise eingesetzt werden:

- *Extremierungsziele $f(\mathbf{x}; \mathbf{p})$* sollen bei der Optimierung entweder minimiert oder maximiert werden. Für ein wohldefiniertes Optimierungsproblem muss mindestens ein Extremierungsziel angegeben werden, anhand dessen der optimale Zustand identifiziert werden kann.
- *Satisfizierungsziele $\mathbf{g}(\mathbf{x}; \mathbf{p})$* dürfen bei der Optimierung gewisse Grenzen nicht über- oder unterschreiten. Ihr Vorhandensein schränkt die zulässige Menge ein.

meter enthalten können. Diese werden bei der Optimierung jedoch konstant gehalten und nur z.B. für Sensitivitätsanalysen variiert.

- **Fixierungsziele $\mathbf{h}(\mathbf{x};\mathbf{p})$:** dürfen bei der Optimierung nicht verändert werden. Fixierungsziele eliminieren Freiheitsgrade des Optimierungsproblems und reduzieren die effektive Dimensionalität des Lösungsraums.

Satisfizierungsziele und Fixierungsziele definieren, zusammen mit der Angabe oberer und unterer Schranken für die einzelnen Entscheidungsvariablen, die zulässige Menge F des Parameteroptimierungsproblems (2.3). Jedes einzelne Extremierungsziel eines Planungsproblems kann als Zielfunktion verwendet werden. Da nur skalare Zielfunktionen zu eindeutigen Optimierungsproblemen führen, müssen im Falle mehrerer Extremierungsziele diese in geeigneter Weise zu einer einzigen skalaren Größe zusammengefasst werden.

$$\begin{array}{llll}
 \min/\max & f(\mathbf{x};\mathbf{p}) & & \text{(a)} \\
 & g_i(\mathbf{x};\mathbf{p}) \leq 0 & i = 1 \dots n & \text{(b)} \\
 & h_j(\mathbf{x};\mathbf{p}) = 0 & j = 1 \dots k & \text{(c)} \\
 & \mathbf{l} \leq \mathbf{x} \leq \mathbf{u} & & \text{(d)}
 \end{array} \tag{2.3}$$

2.2.4 Mathematische Programmierung

Mathematische Programme sind Parameteroptimierungsmodelle, in denen die Funktionen f^a , g_i^a und h_j^a aus (2.3) ausschließlich durch *algebraische Ausdrücke* definiert werden, was jeweils durch den hochgestellten Index a zum Ausdruck gebracht wird. In algebraischen Ausdrücken werden Variablen ausschließlich durch Grundrechenoperationen, sowie *elementare* mathematische Funktionen¹⁰ verknüpft. Für die Formulierung algebraischer Optimierungsmodelle werden häufig Anordnungsstrukturen wie Vektoren und Matrizen verwendet. Diese Definition algebraischer Ausdrücke ist in der Mathematik nicht gebräuchlich. Sie wird in dieser Arbeit in Anlehnung an die im OR übliche Terminologie verwendet: Dort werden so genannte *algebraische Modellersprachen*¹¹ für die Formulierung von Optimierungsmodellen mit derartigen mathematischen Ausdrücken verwendet (Fraginière und Gondzio, 2002: S.997). Im Folgenden werden daher die Begriffe „mathematisches Programm“ und „algebraisches Optimierungsmodell“ synonym gebraucht. Die im folgenden Abschnitt diskutierte Transformation des stoffstromnetzbasierten Optimierungsproblems in ein mathematisches Programm wird in diesem Sinne als *algebraische Reformulierung* bezeichnet.

¹⁰ Als *elementare* Funktionen werden hier die bei (Wolfram Research Inc., 2010a) aufgeführten und definierten Funktionen verstanden. Dazu gehören insbesondere das Potenzieren, Wurzelziehen, die Exponential- und Logarithmusfunktion sowie die trigonometrischen Funktionen.

¹¹ Eines der führenden Programmiersprachen im Bereich der mathematischen Programmierung heißt beispielsweise GAMS, was für „General Algebraic Modeling System“ steht (<http://www.gams.com>).

Klassifizierung mathematischer Programme

Je nach den Eigenschaften der Entscheidungsvariablen x , Zielfunktion $f(x)$, sowie der Restriktionen $g_i(x)$ und $h_j(x)$, werden verschiedene Klassen von mathematischen Programmen unterschieden. Tabelle 2.3 führt einige der am häufigsten vorkommenden Problemtypen auf, erhebt jedoch keinen Anspruch auf Vollständigkeit. Die Wahl eines geeigneten Optimierungsverfahrens zur Lösung eines mathematischen Programms richtet sich maßgeblich nach dieser Klassifizierung.

Tabelle 2.3: Spezielle Unterklassen mathematischer Programme.

Eigenschaften	Problemtyp
f, g_i, h_j linear, $x \in R^u$	Lineares Programm (LP)
f, g_i, h_j nicht linear	Nichtlineares Programm (NLP)
$x \in [R^u \times] Z^v$	[gemischt] ganzzahliges Programm, linear [M]ILP oder nicht linear[M]INLP
$n = k = 0$	unrestringierte Optimierung
$x_i = n \cdot s_i, n, s_i \in Z$	diskrete Optimierung

Eine für die Praxis besonders wichtige Unterklasse sind lineare Programme (LP). Beim LP sind die Zielfunktion, sowie alle Restriktionen durch *gewichtete Summen* der Entscheidungsvariablen gegeben. In (2.4) ist die im OR übliche Matrixnotation dargestellt (Dasgupta et al., 2006: S.222). Der transponierte Vektor \mathbf{c}^T enthält die Koeffizienten der Zielfunktion, die Matrix \mathbf{A} die Koeffizienten aller Restriktionen.

$$\begin{aligned}
 \max \quad & \mathbf{c}^T \mathbf{x} \\
 \text{Nb.} \quad & \mathbf{A} \mathbf{x} \leq \mathbf{b} \\
 & \mathbf{x} \geq \mathbf{0}
 \end{aligned}
 \tag{2.4}$$

LPs haben unter allen mathematischen Programmen die günstigsten Lösungseigenschaften. Mit gängigen Solvern lassen sich globale Optima für Probleme mit einer prinzipiell unbeschränkten Anzahl von Entscheidungsvariablen und Restriktionen effizient lösen (Frontline Systems, 2009). Aufgrund dieser günstigen Lösungseigenschaften verwenden auch manche Lösungsstrategien für nichtlineare (NLP) oder diskrete Optimierungsprobleme den Ansatz, diese „schwierigeren“ Probleme durch geeignete LPs zu approximieren und diese ersatzweise zu lösen.

2.2.5 Entscheidungstheoretische Ansätze zur Aggregation mehrerer Extremierungsziele zu einer skalaren Zielfunktion

Gibt es bei einem Planungsproblem mehrere Extremierungsziele $\mathbf{f} = (f_1, \dots, f_n)$, so spricht man von einem *Vektoroptimierungsproblem*. Eindeutige Optima können jedoch nur auf der Grundlage einer skalaren Zielfunktion bestimmt werden. In der Entscheidungstheorie bzw. der multikriteriellen Optimierung werden verschiedene Ansätze verwendet, um vektorielle in skalare Optimierungsprobleme zu überführen (vgl. Neumann und Morlock, 2002: S.135ff, Scholl, 2001: S.46f). Alle im Folgenden dargestellten Methoden setzen voraus, dass der Entscheidungsträger seine Präferenzen bezüglich der Wichtigkeit der verschiedenen Extremierungsziele präzisiert.

Zieldominanz und lexikographische Ordnung

Die einfachste Möglichkeit der Präzisierung ergibt sich im Falle der Zieldominanz, d.h. wenn ein bestimmtes Extremierungsziel f_k alle Übrigen in seiner Wichtigkeit für den Entscheidungsträger dominiert. In diesem Fall reduziert sich das vektorielle auf ein skalares Optimierungsproblem bezüglich f_k .

Auf einem ähnlichen Prinzip basiert der Ansatz der lexikographischen Ordnung. Zunächst müssen dabei alle Extremierungsziele nach ihrer Bedeutung für den Entscheidungsträger in eine strikte Rangordnung gebracht werden. Wie im Falle der Zieldominanz wird zunächst ein skalares Optimierungsproblem bezüglich des wichtigsten Extremierungsziels gelöst. Wenn das Optimum entartet ist, d.h. wenn mehrere Lösungen zu demselben optimalen Ergebnis führen, wird bezüglich der verbleibenden Lösungsmenge ein weiteres skalares Optimierungsproblem mit dem zweitwichtigsten Extremierungsziel als Zielfunktion gelöst. Dieses Vorgehen wird fortgeführt, bis ein eindeutiger optimaler Systemzustand identifiziert werden konnte.

Goal-Programming

Beim so genannten *Goal-Programming* muss der Entscheidungsträger seine Präferenzen in Form von Zielwerten $f_k^\#$ (goals) für alle Extremierungsziele offenlegen. Das Optimum besteht dann in derjenigen Lösung, die der Ideallösung $\mathbf{f}^\#$ am nächsten kommt. Das Optimierungsproblem kann als Minimierung einer L_p -Metrik (2.5) im Ergebnisraum formuliert werden. Die Gewichtungsfaktoren η_k sind optional. Für $p = 1$ ergibt sich die Summe der Abstände (Manhattan-Metrik), für $p = 2$ die aus der Geometrie bekannte euklidische Norm und für $p \rightarrow \infty$ die so genannte Maximum-Metrik, bei der nur die jeweils größte Abweichung einer Ergebniskategorie relevant ist.

$$\psi(l) = \sqrt[p]{\sum_{k \in M} \eta_k |l_k - l_k^\#|^p} \quad (2.5)$$

Zielgewichtung

Im Idealfall kann der Entscheidungsträger quantitative Gewichte g_k angeben, die die Wichtigkeit der verschiedenen Extremierungsziele widerspiegeln. Die Zielfunktion $g(\mathbf{f})$ ergibt sich in diesem Fall durch eine gewichtete Summe aller Extremierungsziele (2.6).

$$g(l) = \sum_{k=1}^{\kappa} g_k \cdot l_k \quad (2.6)$$

2.2.6 Systemfreiheitsgrade

Definition von Systemfreiheitsgraden

Definition 2.1: Die Freiheitsgrade eines Systems S bilden eine Menge von Systemattributen mit den folgenden Eigenschaften:

- a) Der Systemzustand ist eindeutig durch die Spezifizierung aller Freiheitsgrade festgelegt.
- b) Jeder Freiheitsgrad kann unabhängig von den anderen variiert werden.

Aus Definition 2.1 a) folgt unmittelbar: Wird mindestens ein Freiheitsgrad nicht festgelegt, so ist der Zustand des Systems nicht definiert. 2.1 b) impliziert, dass Freiheitsgrade nie als Funktionen anderer Freiheitsgrade dargestellt werden können. Die Anzahl der Freiheitsgrade ist eine eindeutige Systemeigenschaft, die in dieser Arbeit mit $FG(S)$ bezeichnet wird. Die Auswahl der Freiheitsgrade aus allen Systemattributen ist dagegen in der Regel nicht eindeutig. Wenn das System ein Modell ist, werden die Freiheitsgrade auch als *erklärende* bzw. *unabhängige* Variablen und die übrigen Systemattribute als *erklärte* bzw. *abhängige* Variablen bezeichnet.

Freiheitsgrade und die Dimensionalität von Optimierungsmodellen

Bei Optimierungsmodellen wird die Anzahl ihrer Freiheitsgrade auch als *Dimensionalität des Optimierungsproblems* bezeichnet. Liegen bei einem mathematischen Programm ausschließlich Ungleichungsrestriktionen vor, sind die Entscheidungsvariablen x_D die *Freiheitsgrade* des Modells¹². Dieser einfache Zusammenhang zwischen der Anzahl der Entscheidungsvariablen und der Dimensionalität eines Optimierungsproblems gilt nicht mehr, sobald es Gleichungsrestriktionen gibt. Jede Gleichungsrestriktion $h_j(\mathbf{x})=0$ kann als *implizite* Definitionsgleichung für eine der Entscheidungsvariablen in $h_j(\mathbf{x})$ aufgefasst werden und reduziert folglich die Anzahl vorhandener Freiheitsgrade um eins (Schwefel, 1995: S.8).

¹² Die Modellparameter p sind bildlich gesprochen eingefrorene Freiheitsgrade, die bei der Optimierung konstant gehalten werden und daher bei der Beurteilung der Dimensionalität nicht mit berücksichtigt werden.

Algebraische Gleichheitsrestriktion können meistens in *explizite* Definitionsgleichungen der Art $x_j = x_j(x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)$ umgeformt werden¹³. In dieser Form ist die Entscheidungsvariable x_j offensichtlich kein Freiheitsgrad des Systems mehr, sondern eine von den anderen Entscheidungsvariablen $x_{i \neq j}$ abhängige Ergebnisgröße. Da die Gleichungsrestriktion prinzipiell nach einer beliebigen Entscheidungsvariablen aufgelöst werden kann, ist auch klar, dass die Auswahl der Systemfreiheitsgrade nicht eindeutig sein kann. Die Anzahl der Freiheitsgrade des Optimierungsmodells lässt sich berechnen, indem von der Anzahl der Entscheidungsvariablen die Anzahl vorhandener Gleichungsrestriktionen abgezogen wird¹⁴.

2.3 Stoffstromanalysen in der Industrial Ecology

Industrielle Produktion basiert auf dem Verbrauch natürlicher Rohstoffe. Die aus dem Abbau und der Veredelung gewonnen Zwischenprodukte sowie die in weiterführenden Herstellungsprozessen erzeugten Konsumprodukte gelangen nach ihrem Gebrauch bzw. Verbrauch in Form von Abfällen und stofflichen Emissionen wieder zurück in die natürliche Umwelt¹⁵. Ayres (1989) hat für die Beschreibung der materiellen Grundlagen moderner Wirtschaftssysteme den Begriff des *industriellen Metabolismus* geprägt.

2.3.1 Das Metabolismusparadigma der Industrial Ecology

Unter Metabolismus (Stoffwechsel) versteht man in der Biologie „die Gesamtheit der biochemische Vorgänge, die im pflanzlichen, tierischen und menschlichen Organismus oder in Teilen davon ablaufen und dem Aufbau, Umbau und der Erhaltung der Körperfunktion dienen“ (Brockhaus, 1994). Das Metabolismus-Konzept wird somit auf verschiedenen Hierarchiestufen von der Zellbiologie über die Beschreibung einzelner Organe und Organismen bis hin zu ganzen Ökosystemen angewendet (Fischer-Kowalski, 2007: S.90). Wenn man vom biologischen Kontext abstrahiert, ist ein Metabolismus ein *stoffliches I/O-System*, das zur Erfüllung einer spezifischen Funktion Stoffe aus der Umgebung aufnimmt (Input), durch biologische, chemische oder physikalische Vorgänge verändert, oder in andere Stoffe umwandelt (Transformation) und diese teilweise wieder an die Umwelt abgibt (Output) (Abbildung 2.6).

¹³ Eine derartige Umformung ist beispielsweise für alle linearen Gleichheitsrestriktionen möglich.

¹⁴ Die Gleichungen dürfen nicht redundant sein. Sind alle $h_j(x)=0$ linear ist das gleichbedeutend mit linearer Unabhängigkeit der Gleichungsrestriktionen.

¹⁵ „Every substance extracted from the earth’s crust, or harvested from a forest, a fishery or from agriculture [...] soon becomes an actual waste in almost all cases, with a delay of a few weeks to a few years at most“ (Ayres und Ayres, 1996: S.1).

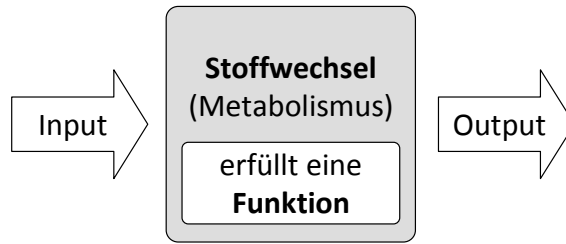


Abbildung 2.6: Stoffwechsel als materielles Input-Output-System

Industrielle Metabolismen

In Analogie zum biologischen Metabolismus versteht Ayres unter dem industriellen Metabolismus “the whole integrated collection of physical processes that convert raw materials and energy, plus labour, into finished products and wastes” (Ayres, 1994). Der Begriff des Metabolismus ist auf einzelne Produktionsanlagen über Produktionsstandorte und Unternehmen (Abbildung 2.7) bis hin zu ganzen Volkswirtschaften anwendbar. Während Stoffströme in wirtschaftswissenschaftlichen Analysen vor allem in ihrer Eigenschaft als Warenströme, d.h. im Hinblick auf ihren gesellschaftlichen Nutzen analysiert werden¹⁶, verschiebt das Metabolismus-Paradigma den Fokus unter einem ökologischen Blickwinkel gezielt von der Wertebene auf die materielle Ebene. Aus der Vorstellung, dass Industrie und Natur – Technosphäre und Ökosphäre – durch Materieströme miteinander verwobene Stoffwechselsysteme sind¹⁷, hat sich in den letzten zwanzig Jahren das umweltwissenschaftliche Forschungsfeld der *Industrial Ecology*¹⁸ entwickelt.

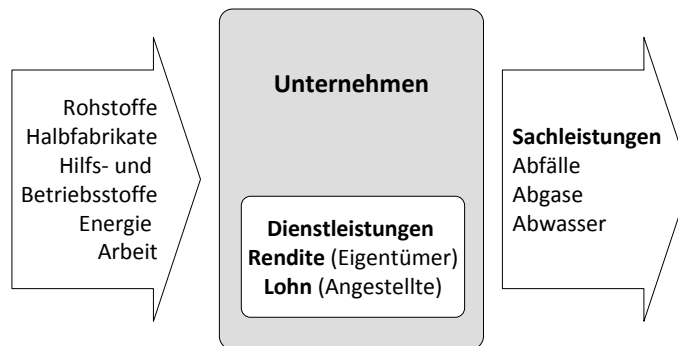


Abbildung 2.7: Das Unternehmen als Metabolismus. Systemfunktionen in fettem Schriftsatz.

¹⁶ Ayres und Kneese (Ayres und Kneese, 1969: S.284) bemerken hierzu: „Almost all of standard economic theory is in reality concerned with services. Material objects are merely the vehicles which carry some of these services, and they are exchanged because of consumer preferences for the services associated with their use or because they can help to add value in the manufacturing process.“

¹⁷ „We may think of both the biosphere and the industrial economy as systems for the transformation of materials“ (Ayres, 1989: S.23)

¹⁸ Die englischsprachige Bezeichnung hat sich auch im deutschen Sprachraum durchgesetzt und wird daher im Folgenden ohne Übersetzung verwendet (Isenmann und von Hauff, 2007).

Das Hauptanliegen der Industrial Ecology besteht in der Entwicklung nachhaltiger Konzepte für die industrielle Produktion und das gesellschaftliche Wirtschaften insgesamt. Dabei werden im Wesentlichen zwei komplementäre Strategien verfolgt: Die Substitution umweltschädlicher und für den Menschen gefährlicher Stoffe in Produkten und Herstellungsprozessen durch harmlose Substanzen wird als *Materialsubstitution* („Detoxification“) bezeichnet. Bei der *Dematerialisierung* wird dagegen ganz allgemein die Reduzierung des Materialverbrauchs techno-ökonomischer Systeme angestrebt (Bringezu und Moriguchi, 2002: S.79f, Jackson, 2002: S.36). Das an den biogeologischen Stoffkreisläufen der Natur orientierte Idealziel der Dematerialisierungsstrategie ist das *Industrial Ecosystem*, bei dem alle Stoffkreisläufe in sich geschlossen sind und allein durch die Nutzung regenerativer Energieformen angetrieben werden. (Frosch und Gallopoulos, 1989: S.94, Lifset und Graedel, 2002: S.5).

Die Außen- und die Innenperspektive von Metabolismen

Für die ökologische Bewertung eines industriellen Metabolismus ist es ausreichend, diesen aus der Außenperspektive durch seine Stoff- und Energiebilanz zu charakterisieren. Darunter versteht man die tabellarische Aufstellung aller Stoffe und Energien, die der Metabolismus innerhalb eines festgelegten Zeitraums mit seiner Umgebung austauscht (Abbildung 2.8).

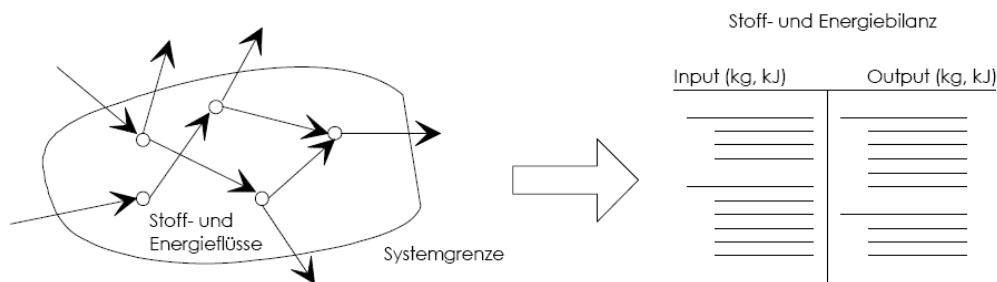


Abbildung 2.8: Beschreibung eines Stoffstromsystems über seine (äußere) Stoff- und Energiebilanz (Möller und Rolf, 1995:35, Abb.1).

Anhaltspunkte für die Verbesserung der Umweltwirkung ergeben sich jedoch nur durch eine weitere Disaggregation. Das wichtigste Analyseinstrument der Industrial Ecology ist daher die *Stoffstromanalyse*. Mit ihr werden industrielle Metabolismen hinsichtlich ihres Stoffaustauschs mit der Umgebung, sowie der Umwandlung und Speicherung von Stoffen in ihrem Inneren untersucht. Durch eine Stoffstromanalyse werden sukzessive relevante Subsysteme eines Metabolismus herausgearbeitet und die zwischen ihnen bestehenden materiellen und energetischen Austauschprozesse analysiert.

Stoffstromanalysen beinhalten drei eng miteinander verbundene Teilaufgaben: Bei der *Stoffanalyse* wird zunächst festgelegt, welche Stoffe und Substanzen im Zusammenhang mit der untersuchten Fragestellung relevant sind. Bei der *Strukturanalyse* werden die Prozesse und Akteure

erfasst, bei denen diese Stoffe umgesetzt werden. Auf Grundlage der Stoff- und Strukturanalyse können Stoffumsätze und Stoffströme schließlich durch Messungen oder Berechnungen *quantifiziert* werden (Enquete-Kommission, 1994: S.557).

2.3.2 Formale Definition der Stoffstromanalyse

In der einschlägigen Literatur wird das Konzept der Stoffstromanalyse immer vergleichsweise vage eingeführt und als weitgehend selbsterklärend angenommen. Brunner und Rechberger (2004: S.3) definieren die Stoffstromanalyse beispielsweise als „vollständige und konsistente Erfassung von Stoffströmen und –beständen in einem gegebenen System“. Für die meisten Zwecke ist das auch tatsächlich ausreichend. Für diese Arbeit, welche sich mit der Stoffstromanalyse aus einer methodologischen Perspektive befasst, soll im Folgenden jedoch der Versuch einer etwas formaleren Definition der Stoffstromanalyse unternommen werden. Dafür wird zunächst der Begriff des Stoffstromsystems definiert.

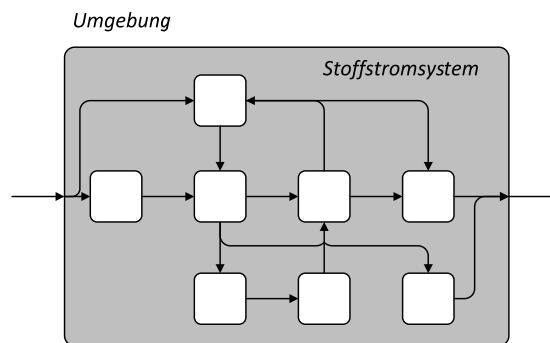


Abbildung 2.9: Veranschaulichung von Definition 2.2. Das Stoffstromsystem (grauer Kasten) besteht aus mehreren Komponenten (weiße Kästen), die über Stoffströme (Pfeile) miteinander verbunden sind.

Definition 2.2: Ein Stoffstromsystem S hat folgende Eigenschaften (Abbildung 2.9):

- S besteht aus einer endlichen Anzahl (räumlich) diskreter Komponenten $e \in (e_1, \dots, e_n) = E$, die untereinander und mit der Umgebung von S Stoffe austauschen.
- Der Stoffaustausch zwischen S und seiner Umgebung ergibt sich als Summe der Stoffströme zwischen den Komponenten und der Umgebung. Dabei gilt: Führt ein Stoffstrom zur Abnahme des Stoffbestands in der Umgebung, wird er als Inputstrom von S bezeichnet. Führt er zu einer Bestandszunahme, handelt es sich um einen Outputstrom.
- Stoffe sind diskrete oder kontinuierliche Objekte. Sie kommen in positiven Quantitäten vor, die sich durch Abzählen oder in geeigneten Einheiten messen lassen: $Q_k \in \mathbb{Z}^+$ oder $Q_k \in \mathbb{R}^+$. Der Index $k \in M$ verweist auf die jeweilige Stoffart. Insgesamt gibt es $\kappa = |M| \in \mathbb{N}$ verschiedene Stoffe im Stoffstromsystem.

- Die Quantität ΔQ_k eines Stoffes, die in einem bestimmten Zeitintervall $\Delta\tau$ zwischen zwei Komponenten e_i und e_j ausgetauscht wird, definiert den *Stoffstrom* $z_k^{(e_i, e_j)} = \Delta Q_k / \Delta\tau$.
- Stoffe können in den Komponenten *gespeichert* oder *transformiert* werden. Transformation bedeutet entweder die *Modifikation* bestimmter Eigenschaften oder die vollständige Umwandlung in Stoffe anderer Art. Im Falle der *Umwandlung* ist die Komponente eine Senke für den umgewandelten Stoffstrom z_k und die Quelle für den neu erzeugten Stoffstrom z_k' .

Unmittelbar darauf aufbauend kann nun die Stoffstromanalyse formal definiert werden:

Definition 2.3: Die Stoffstromanalyse ist eine spezielle Form der modellbasierten Systemanalyse, mit der folgende Ziele verfolgt werden:

- (i) die Abbildung eines Originalsystems als Stoffstromsystem.
- (ii) die Bestimmung aller Stoffströme innerhalb des Systems und an seinen Grenzen innerhalb eines vorgegebenen Zeitintervalls.
- (iii) die Bestimmung aller Stoffbestände im Inneren des Systems, jeweils zu Beginn und am Ende eines Betrachtungszeitraums.

Als Ergebnis jeder Stoffstromanalyse entsteht ein *Stoffstrommodell*. Der Zustand dieses Modells ist durch die Gesamtheit der Stoffbestände und Stoffströme definiert. Der *Periodenbezug* von Stoffstrommodellen führt zu einer Besonderheit. Bestände sind im Modell wie im wirklichen System zeitabhängige Zustandsgrößen. Während Stoffströme aus einer *Aktivität* des realen Systems resultieren, stellen sie im Hinblick auf das Stoffstrommodell ebenfalls Zustandsgrößen dar, die sich durch das *Integral* aller Stoffaustauschprozesse über die Dauer des Betrachtungszeitraums ergeben.

Stoffstromanalyse als Plattform für die Integration multidisziplinärer Ansätze.

Die vielfältigen Ansprüche an moderne Produktionsanlagen wie Wirtschaftlichkeit, Umweltverträglichkeit aber auch eine geringe Störanfälligkeit können nur durch eine *ganzheitliche Planung* erfüllt werden, die das Wissen und die Methoden unterschiedlicher Disziplinen integriert. So müssen neben den technischen und naturwissenschaftlichen Aspekten insbesondere auch betriebswirtschaftliche und ökologische Gesichtspunkte bei der Planung industrieller Produktionssysteme berücksichtigt werden.

Das Prinzip der Stoffstromanalyse eignet sich hierfür besonders gut, da es den Systemanalysemethoden der anderen in diesem Zusammenhang relevanten Disziplinen sehr ähnlich ist. So können die in der betriebswirtschaftlichen Produktionstheorie verwendeten Gozintographen oder Leontief-Modelle, ebenso wie die Fließbildmodelle der Verfahrenstechnik, dem Prinzip nach als Stoffstrommodelle betrachtet werden. Die Unterscheidung liegt im Einzelfall im Detaillierungsgrad, der sich aus den spezifischen Erkenntnisinteressen der verschiedenen Disziplinen ergibt.

2.3.3 Stoffstrommanagement als zentrales Handlungsfeld der Industrial Ecology

Umweltschutz im Sinne der Industrial Ecology bedeutet „die vielfältigen Prozesse, innerhalb derer Industrie und Natur ihre Einheit [fortwährend] erneuern, als solche zu erkennen, zu kontrollieren und zu steuern, sie also planvoll zu gestalten“ (Hofmeister, 1989: S.2). Daher werden Stoffstromanalysen häufig im Kontext eines *Stoffstrommanagements*¹⁹ durchgeführt und stellen in diesem Sinne eine modellbasierte Planungsmethode dar (Abbildung 2.10).

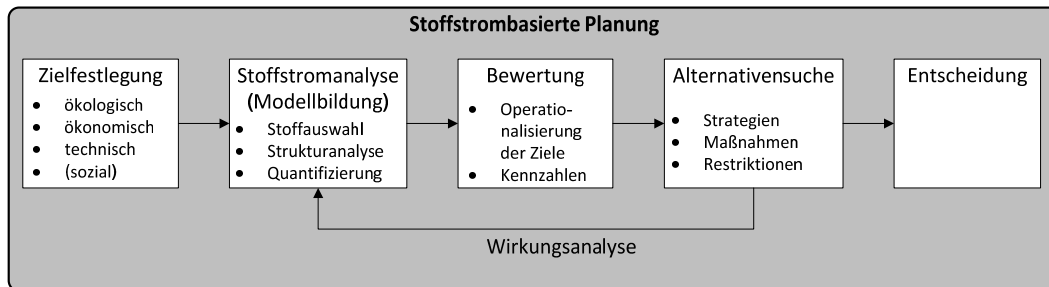


Abbildung 2.10: Stoffstrombasierte Planung (vgl. Enquete-Kommission, 1994: S.556)

Ökobilanzierung und betriebliches Stoffstrommanagement

Für das Stoffstrommanagement können unterschiedliche Bezugspunkte innerhalb eines techno-ökonomischen Systems gewählt werden. So können die Ziele „auf betrieblicher Ebene, in der Kette der an einem Stoffstrom beteiligten Akteure oder auf der staatlichen Ebene entwickelt“ werden (Enquete-Kommission, 1994: S.719). Das *Life Cycle Assessment* (LCA) setzt beispielsweise bei einem bestimmten Produkt an, von dem ausgehend alle mit seiner Herstellung, Nutzung und Entsorgung verbundenen Prozesse der Technosphäre analysiert werden. Der entsprechende Ausschnitt aus der Technosphäre wird als Produktsystem bezeichnet (ISO 14040, 2006: S.21).

Beim *betrieblichen Stoffstrommanagement* steht dagegen das Unternehmen als eigenständig agierende Wirtschaftseinheit im Zentrum der Analyse. Die ökologische Bewertung betrieblicher Produktion erfordert jedoch, genau wie die Produktökobilanzierung, eine überbetriebliche Perspektive. In einer modernen Marktwirtschaft erfolgt die Herstellung von Konsumgütern *arbeits-teilig* in einem komplexen Geflecht vieler einzelner Betriebe, deren Produktion jeweils nur einen kleinen Ausschnitt der gesamten mit der Herstellung verbundenen Prozess- bzw. Wertschöpfungskette umfasst. Die Produktionsaktivitäten in überbetrieblichen Wertschöpfungsketten (Supply Chains) sind abhängig voneinander. Daher entstehen durch eine betriebliche Produktion stets auch indirekte Umweltwirkungen in außerhalb der Betriebsgrenzen liegenden Prozessen zur

¹⁹ Stoffstrommanagement ist das „zielorientierte, ganzheitliche und effiziente Beeinflussen von Stoffströmen in wirtschaftlichen Systemen, insbesondere mit dem Ziel des schonenden Einsatzes von Ressourcen und der reduzierenden Verwendung umweltbelastender Stoffe“ (Brockhaus, 1998).

Herstellung von Vorprodukten und Energie sowie zur Abfallverwertung bzw. –entsorgung, die für eine aussagekräftige ökologische Bewertung jedoch notwendigerweise berücksichtigt werden müssen (Braunschweig und Müller-Wenk, 1993: S.30f).

2.4 Stoffstromnetze

In dieser Arbeit wird von einer speziellen Stoffstromanalysemethode ausgegangen: den in den 1990er Jahren entwickelten *Stoffstromnetzen* (Möller, 1993, 1994, 2000, Schmidt, 1997). Im Folgenden wird gezeigt, wie Produktionssysteme mit Stoffstromnetzen modelliert werden. Abschließend wird auf die weitreichenden Möglichkeiten sowie die Grenzen dieses Ansatzes im Hinblick auf die Planung von Produktionssystemen eingegangen.

2.4.1 Strukturanalyse mit Stoffstromnetzen

Die Abbildung der Struktur von Produktionssystemen mit einem Stellen-Transitions-Netz

Stoffstromnetze verwenden für die Strukturanalyse so genannte *Stellen-Transitions-Netze* (S/T-Netze), wie sie von Petri (1962) zur Beschreibung nebenläufiger Systeme in der theoretischen Informatik eingeführt worden sind (Möller, 2000: S.72ff). Ein S/T-Netz ist ein *gerichteter bipartiter Graph* $N=(P,T,F)$, der aus einer Menge P von Stellen, einer dazu disjunkten Menge T von Transitionen, sowie einer auf diesen Mengen definierten *Flussrelation* $F \subseteq (P \times T) \cup (T \times P)$ besteht.

Für die Modellierung der Struktur eines Produktionssystems mit einem Stoffstromnetz werden Stoffumwandlungsprozesse als Transitionen, Stoffbestände als Stellen und Stoffströme als Flussrelationen abgebildet. Flussrelationen und Transitionen bilden die Aktivität eines Systems ab. Die Stellen repräsentieren dagegen den Zustand des Systems zu einem bestimmten Zeitpunkt, entweder zu Beginn oder am Ende eines betrachteten Zeitraums.

Durch die *Differenzierung* der Netzknoten in Stellen und Transitionen unterscheiden sich Stoffstromnetze von den meisten anderen Stoffstrommodellen, bei denen üblicherweise nur Umwandlungsprozesse und Stoffströme in Flussdiagrammen (flow charts) berücksichtigt werden. Das Stellen-Transitions-Konzept der Stoffstromnetze für die Stoffstromanalyse entspricht dem Prinzip der doppelten Buchführung im betrieblichen Rechnungswesen, bei dem Bestandskonten (Stellen), Erfolgskonten (Transitionen) und Buchungen (Flussrelationen zwischen diesen Konten) eine Rolle spielen. Der Vorteil liegt darin, dass sich auf einer solchen Grundlage aus einem Modell in konsistenter Weise sowohl Zustandskennzahlen (Betriebsbilanz) als auch Aktivitätskennzahlen (Betriebserfolg) ableiten lassen (Möller und Rolf, 1995: S.38ff).





Die vier verschiedenen Stellentypen

Die Syntax der S/T-Netze erfordert, dass sich Transitionen und Stellen beim Aufbau des Stoffstromnetzes strikt abwechseln. Die als Transitionen modellierten Umwandlungsprozesse bezie-

hen ihre Inputs aus Stellen und geben ihre Outputs an anderen Stellen ab. Insgesamt gibt es vier verschiedene Stellentypen (Tabelle 2.4).

Für die Modellierung des Systeminneren werden „normale“ Stellen und Verbindungsstellen verwendet, je nachdem ob Bestandsveränderungen zwischen den Prozessen möglich sind oder nicht. Man kann sich eine normale Stelle wie ein Zwischenlager vorstellen. Die produzierte Quantität eines Zwischenprodukts muss nicht zwangsläufig dessen Verbrauch in den nachfolgenden Prozessen entsprechen. Entsprechend den tatsächlichen Gegebenheiten nimmt der Bestand des Zwischenprodukts in der Stelle entweder zu oder ab, darf normalerweise jedoch nicht negativ werden. Um Prozesskonstellationen abzubilden, in denen keine Zwischenlager vorhanden und die realen Produktionsprozesse direkt miteinander verbunden sind, werden *Verbindungsstellen* verwendet, in denen *keine Bestandsveränderungen* auftreten. Im Gegensatz zu normalen Stellen, die durch ihre Speicherfunktion wie Puffer zwischen den Transitionen wirken, sind Verbindungsstellen *Schnittstellen*, die den Output der vorhergehenden in einen Input der folgende Transition umsetzen.

Tabelle 2.4: Die verschiedenen Stellentypen eines Stoffstromnetzes

	Symbol	Bezeichnung	Funktion
Systeminneres		„normale“ Stelle	Bestandsveränderungen innerhalb des Betrachtungszeitraums sind zulässig.
		Verbindungsstelle	keine Bestandsveränderungen im Betrachtungszeitraum möglich. Daher müssen alle zugehenden Stoffquantitäten innerhalb einer Periode auch wieder abgehen.
Systemgrenze		Inputstelle	repräsentiert die Systemgrenze. Als Quelle eines Systeminputs weist die Inputstelle am Ende einer Periode negative Bestände auf.
		Outputstelle	repräsentiert die Systemgrenze. Als Senke des Systemoutputs weisen Outputstellen zu Periodenende positive Bestände auf.

Input- und Outputstellen repräsentieren die Grenze zwischen einem Stoffstromsystem und seiner Umgebung. Sie werden daher zusammenfassend auch als *Grenzstellen* bezeichnet. Genau genommen stellen sie *außerhalb* des Systems liegende Stoffspeicher dar, die Stoffquellen (Inputstellen) bzw. Stoffsenken (Outputstellen) für das eigentliche System sind. Die Bestandsveränderungen aller Input- und Outputstellen zusammengenommen ergeben die Stoffbilanz des Stoffstromsystems.

Zementherstellung als Beispiel für die Abbildung industrieller Produktion als Stoffstromnetz

Abbildung 2.11 zeigt beispielhaft ein Stoffstromnetz für die industrielle Zementherstellung. Die Produktion wird in zwei Stufen modelliert: Zunächst entsteht in der Drehofenanlage (Transition T1) Klinker, der den Bestand im Klinkerlager (normale Stelle P3) erhöht. Im zweiten Produktionsschritt, der Kugelmühle (T2), wird der Klinker als Rohstoff eingesetzt, um daraus Portland-Zement herzustellen. Wird im Betrachtungszeitraum mehr Klinker hergestellt, als für den nachfolgenden Produktionsschritt benötigt wird, nimmt der Klinkerbestand in P3 zu. P3 stellt somit einen Puffer dar, der die lokalen Prozessmodelle T1 und T2 unabhängig voneinander macht. Alle übrigen Stellen repräsentieren die Entnahme von Stoffen aus der bzw. die Abgabe an die Systemumwelt.

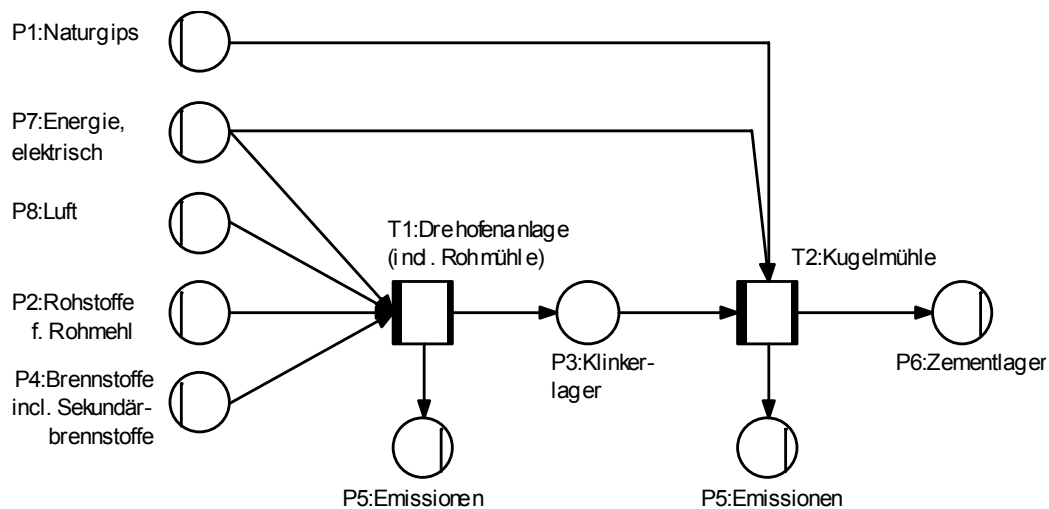


Abbildung 2.11: Beispiel eines Stoffstromnetzes für die Herstellung von Portland-Zement (aus Möller, 2000: S.90)

2.4.2 Stoffanalyse und Quantifizierung des Systemzustands

Alle Stoffströme und -bestände eines Stoffstromsystems werden in Stoffstromnetzen durch quantitative Modellvariablen erfasst, die den einzelnen Netzelementen zugeordnet werden:

- Jeder Stelle $p \in P$ wird ein *Bestandsvektor* zu Beginn $s^{(p)}(\tau)$ und am Ende einer Periode $s^{(p)}(\tau + \Delta\tau)$ zugeordnet.
- Jeder Flussrelation $a \in F$ wird ein *Stoffstromvektor* zugeordnet $\mathbf{z}^{(l,p)}$ bzw. $\mathbf{z}^{(p,l)}$, der das Integral aller Stoffflüsse über den Betrachtungszeitraum $\Delta\tau$ darstellt.

Sankey-Darstellung von Stoffstromvektoren

Abbildung 2.12 veranschaulicht die Zuordnung *vektorieller Stoffströme* zu den Flussrelationen $a_1 = (P1, T1)$ und $a_2 = (T1, P2)$ eines einfachen Stoffstromnetzes: Jeder Flussrelation in Abbildung 2.12 (a) entsprechen *mehrere* Stoffströme in Abbildung 2.12 (b). Die Komponenten des Flussvektors werden als Pfeile dargestellt, deren Breite proportional zur Größe des Stoffstroms ist. Au-

ßerdem wird jeder Stoffart eindeutig eine bestimmte Farbe zugeordnet. Diese Darstellungsweise wird als *Sankey-Diagramm* bezeichnet²⁰. Damit ein sinnvoller Vergleich von Stoffquantitäten möglich ist, werden im Sankey-Diagramm üblicherweise nur Stoffströme mit gleichen physikalischen Grundeinheiten dargestellt.

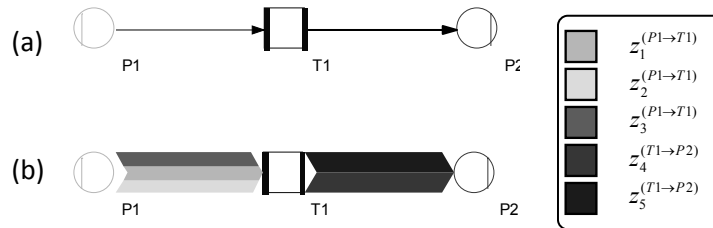


Abbildung 2.12: Vektorielle Darstellung von Stoffströmen in einem Stoffstromnetz

Tabellarische Darstellung von Bestandsvektoren

Abbildung 2.13 zeigt die tabellarische Darstellung der Bestandsvektoren in den Grenzstellen des Stoffstromnetzes aus Abbildung 2.12. Der Betrachtungszeitraum ist das Jahr 2010. Weil keine Anfangsbestände vorhanden sind ($s^{(P1)}(\tau) = s^{(P2)}(\tau) = 0$), werden nur die Endbestände zum Zeitpunkt $\tau + \Delta\tau = 31.12.2010$ ausgewiesen. Entsprechend der in Definition 2.2 festgelegten Vorzeichenkonvention hat der Bestand der Inputstelle $P1$ ein negatives und der der Outputstelle $P2$ ein positives Vorzeichen.

Material	T	Q	01.01.2010	31.12.2010	Unit
1	C			-1 kg	
2	C			-1 kg	
3	C			-1 kg	

(P1)

Material	T	Q	01.01.2010	31.12.2010	Unit
4	C			1.5 kg	
5	C			1.5 kg	

(P2)

Abbildung 2.13: Darstellung der Stoffbestände zu Beginn (1.1.2010) und zum Ende (31.12.2010) des einjährigen Betrachtungszeitraums in den beiden Stellen $P1$ und $P2$ des Stoffstromnetzes aus Abbildung 2.12.

Skalare Darstellung der Stoffströme und –bestände im Stoffstromnetz

Aus Gründen der besseren Nachvollziehbarkeit werden bei der Darstellung der Formulierung stoffstromnetzbasierter Optimierungsprobleme in Kapitel 3 Stoffströme und –bestände als skalare Größen behandelt: Jede Flussrelation des Stoffstromnetzes entspricht in der skalaren Darstellung dem Fluss *genau einer* Stoffart. Entsprechend kann je Stelle genau eine Stoffart gespeichert werden. Das Stoffstromnetz aus Abbildung 2.12 geht dann in die in Abbildung 2.14 dargestellte Form über. Die skalare und die vektorielle Darstellungsweise von Stoffstromnetzen können, bis auf die Nummerierung von Flussrelationen und Stellen, eindeutig ineinander überführt werden.

²⁰ Eine umfassende historische Einordnung sowie methodologische Untersuchung des Einsatzes von Sankey-Diagrammen im Energie- und Stoffstrommanagement findet sich bei Schmidt (2008b, 2008a).

Daher können alle auf Grundlage der skalaren Darstellung gewonnenen Ergebnisse ohne Beschränkung der Allgemeinheit auch auf die Vektordarstellung übertragen werden²¹.

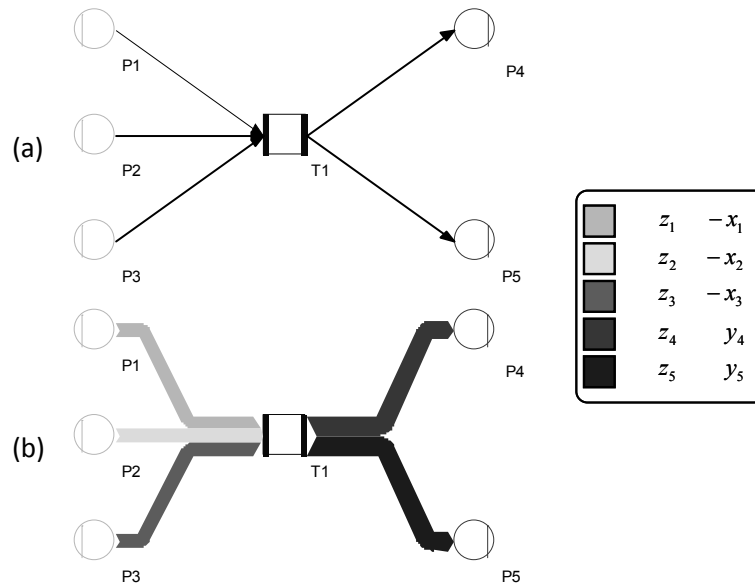


Abbildung 2.14: Skalare Darstellung von Stoffströmen in einem Stoffstromnetz

Verschiedene Notationen für die Stoffströme eines Stoffstromnetzes

In dieser Arbeit werden je nach Kontext verschiedene Notationen für Stoffstromvariablen verwendet (Tabelle 2.5). Bisher wurden Stoffströme durch die Flussrelation indiziert, der sie zugeordnet sind. $z_k^{(t,p)}$ bezeichnet demnach die Quantität des Stoffes k , der von der Transition t zur Stelle p fließt. In dieser Darstellungsweise sind alle Stoffströme positiv: $z_k^{(t,p)} \geq 0$. Die Bipartition des S/T-Netzes bezüglich Stellen und Transitionen erzeugt auch eine Bipartition der Flussrelationen. Daher kann die Indizierung bezüglich der Flussrelationen in eine Indizierung bezüglich der Transitionen transformiert werden, wobei nun explizit zwischen Input- und Outputströmen differenziert wird. Inputströme $x_k^{(t)}$ münden in Transitionen, Outputströme $y_k^{(t)}$ gehen aus diesen hervor. Von der transitionsbezogenen x-y-Notation kann mit (2.7) direkt die Stoffbilanz $\mathbf{z}^{(t)}$ der jeweiligen Transition erzeugt werden.

$$z_k^{(t)} = y_k^{(t)} - x_k^{(t)} \quad (2.7)$$

Die Stoffbilanz $\mathbf{z}^{(t)}$ enthält per definitionem auch negative Komponenten, nämlich die Inputströme der Transition.

²¹ Das ist im Zusammenhang mit der softwaretechnischen Umsetzung wichtig, da in der Stoffstromanalysesoftware Umberto® die Vektorrepräsentation von Stoffstromnetzen verwendet wird.

Tabelle 2.5: Übersicht über die verschiedenen, in dieser Arbeit verwendeten Notationen und Indizierungsweisen für die Stoffströme eines Stoffstromnetzes.

Bezeichnung:	x-y-Notation	z-Notation	
Indizierung bezüglich:	Transition t (Stoffbilanz)		Flussrelation (t,p); (p,t)
Output	$y_k^{(t)}$	$z_k^{(t)}$	$z_k^{(t,p)}$
Input	$x_k^{(t)}$	$-z_k^{(t)}$	$z_k^{(p,t)}$

2.4.3 Prozessmodellierung mit spezifizierten Transitionen

Vom Beschreibungsmodell zum Erklärungsmodell

Sämtliche Stoffströme eines Systems können gemessen und den entsprechenden Flussrelationen zugeordnet werden. Eine solche rein *buchhalterische Verwendungsweise* von Stoffstromnetzen ist sinnvoll, wenn man beispielsweise die Konsistenz von Stoffstromdaten prüfen oder die Stoffumsätze in einem System mit Sankey-Diagrammen visualisieren möchte. Stoffstrommodelle, die in dieser Weise einen spezifischen Systemzustand *beschreiben*, können bei einer veränderten Datenlage jedoch nur dann wiederverwendet werden, wenn erneut alle Daten bestimmt werden.

Stoffstromnetze bieten daher die Möglichkeit, die durch Transitionen abgebildeten Produktionsprozesse durch Formeln bzw. Algorithmen so zu beschreiben, dass aus bekannten Stoffströmen, Stoffbeständen und sonstigen Systemparametern unbekannte Stoffströme und -bestände *berechnet* werden können. Der Aufwand der erstmaligen Datenerhebung wird dadurch nicht zwangsläufig reduziert. Anstatt der Stoffströme müssen nun Parameter der Prozessmodelle bestimmt werden. Dafür erhöht sich jedoch die Wiederverwendbarkeit des Stoffstrommodells. Es wird zu einem *quantitativen Erklärungsmodell*, mit dem der Zustand von Stoffstromsystemen als Funktion einiger erklärender Modellvariablen berechnet werden kann. Kostspielige Messungen bei periodischen Datenerhebungen können so in vielen Fällen durch wesentlich preiswertere Modellrechnungen ersetzt werden (Möller, 2000: S.95).

Verschiedene Arten der Transitionsspezifikation

Durch die Spezifikation werden Transitionen zu lokalen Submodellen eines Stoffstromnetzes, die funktionale Zusammenhänge zwischen den unmittelbar an der Transition anliegenden Stoffströmen und ggf. weiteren Modellparametern festlegen. Es gibt vier verschiedene Arten Transitionen zu spezifizieren, die nachfolgend kurz beschrieben werden sollen²²:

- Lineare Produktions- bzw. Transformationskoeffizienten

²² Für eine detaillierte Darstellung der verschiedenen Möglichkeiten der Transitionsspezifikation sei auf Möller (2000: S.95-105) verwiesen.

- Benutzerdefinierte Funktionen
- Berechnungsalgorithmus, formuliert in einer Skriptsprache
- Subnetze

Bei *linear spezifizierten Transitionen* wird der Transformationsprozess wie in der linearen Aktivitätsanalyse²³ durch unveränderliche *Transformationskoeffizienten* $c_k^{(i)}$ bzw. $a_k^{(i)}$, $b_k^{(i)}$ beschrieben²⁴. Das für alle Stoffströme gleiche Verhältnis zwischen Stoffstrom und Transformationskoeffizient wird in der linearen Aktivitätsanalyse als *Aktivitätsniveau* λ bezeichnet. Diese Form der Transitionsspezifikation ist für Modellentwickler ohne mathematische Vorkenntnisse sehr leicht verständlich. So können beispielsweise die Mengenangaben von *Stücklisten* oder *Rezepturen* unmittelbar als Transformationskoeffizienten verwendet werden. Die Berechnung unbekannter Stoffströme erfolgt mit dem linearen Gleichungssystem (2.8), welches $n+m-1$ Gleichungen für $n+m$ Stoffströme enthält, und somit einfach unterbestimmt ist.

$$\lambda^{(i)} = \frac{z_k^{(i)}}{c_k^{(i)}} = \frac{x_i^{(i)}}{a_i^{(i)}} = \frac{y_j^{(i)}}{b_j^{(i)}} \quad (2.8)$$

$$\forall i \in \{1, \dots, n\}, j \in \{1, \dots, m\}, k \in \{1, \dots, n+m\}$$

Alle Inputs und Outputs einer linear spezifizierten Transition stehen gemäß (2.8) in einem festen Verhältnis zueinander. Diese Eigenschaft eines Produktionsprozesses wird in der Produktionstheorie als *Limitationalität*²⁵ bezeichnet. Variable Kuppelproduktionen oder Prozesse, bei denen Produktionsfaktoren in einem gewissen Maße gegeneinander substituierbar sind, lassen sich mit einer linear spezifizierten Transition daher nicht abbilden.

Das Beispiel der linear spezifizierten Transition verdeutlicht, dass sich der Aufwand der Datenerhebung durch die Spezifikation zunächst nicht reduziert, da die Anzahl von Koeffizienten und anliegenden Stoffströme gleich ist. Ändert sich aber das Prozessniveau, so müssten bei einer un-spezifizierten Transition sämtliche anliegenden Stoffströme neu bestimmt werden, während für

²³ Die lineare Aktivitätsanalyse von Produktionssystemen geht auf die Arbeiten von Koopmans (1951) zurück. Eine übersichtliche Darstellung findet sich z.B. bei Spengler (1998).

²⁴ Die Bezeichnung der Koeffizienten entspricht der z- bzw. x-y-Notation für Stoffströme (vgl. Tabelle 2.5).

²⁵ Ein Prozess heißt *inputlimitational*, wenn alle Inputs eines Prozesses eindeutig vom Output abhängen. Er wird umgekehrt als *outputlimitational* bezeichnet, wenn die Outputmengen eindeutige Funktionen des Inputs sind. Gilt beides, bezeichnet man den Prozess als *limitational* (Dyckhoff, 1998: S.145ff). Das Gegenteil sind *substitutionale* Prozesse, bei denen einzelne Inputs gegeneinander substituiert oder das Verhältnis verschiedener Outputs variiert werden können.

die linear spezifizierte Transition die Messung eines einzigen Stoffstromes genügt, um die restlichen mit dem einfach unterbestimmten Gleichungssystem (2.8) zu berechnen.

Bei der Spezifikation durch *benutzerdefinierte Funktionen* wird vom Modellierer ein Satz von Zuweisungsvorschriften festgelegt, in welche die Input- und Outputströme der Transition wahlweise als Ergebnis- oder Eingangsparameter eingehen. Die Variable auf der linken Seite einer Zuweisung wird dabei durch einen mathematischen Ausdruck einer oder mehrerer Variablen auf der rechten Seite definiert. Dabei gilt das Single-Assignment-Prinzip, wonach eine einmal berechnete Größe nicht durch folgende Zuweisungen überschrieben werden kann (Möller, 2000: S.101f). Außerdem können Transitionen durch einen *Berechnungsalgorithmus* in einer der gängigen Skriptsprachen wie JScript, VBScript oder Python spezifiziert werden (Möller, 2000: S.104). Dies ermöglicht insbesondere die Einbindung externer Softwarekomponenten in die Transitionsberechnung. Schließlich kann eine Transition selbst durch ein Stoffstromnetz auf einer niedrigeren Hierarchiestufe definiert werden. Mit derartigen *Subnetzen* können Produktionssysteme im Sinne einer Systemanalyse schrittweise disaggregiert werden.

2.4.4 Produktionsplanung mit Stoffstromnetzen

Der Stoffbegriff – Von der Stoffstrom- zur Ressourcenflussanalyse

Stoffstromnetze werden vorwiegend für die Analyse von Material- und Energieströmen eingesetzt. Aus methodischer Sicht besteht jedoch kein Grund für eine derartige Einengung des Anwendungsgebietes. „Stoffe“²⁶ im Sinne eines Stoffstromnetzes können alle Arten von Objekten sein, solange sie in geeigneten Einheiten *quantifizierbar* sind. Das Anwendungsgebiet von Stoffstromnetzen bei der Analyse industrieller Produktionssysteme kann daher auch auf immaterielle Produktionsfaktoren wie den Einsatz von Mitarbeitern, finanzielle Ressourcen sowie externe Dienstleistungen ausgeweitet werden. Die Stoffstromanalyse wird dann in einem umfassenderen Sinne zu einer *Ressourcenflussanalyse* (Thißen, 2009: S.155f).

Als theoretisches Fundament für diese inhaltliche Erweiterung des Stoffstromnetzansatzes kann der Vergleich mit dem *abstrakten I/O-Graph* herangezogen werden, der in der Produktionstheorie von Dyckhoff (1998: Kap.3.2.3) für die Abbildung beliebiger Güter- und Dienstleistungsproduktionssysteme vorgeschlagen worden ist. Wie die Stoffstromnetze basiert auch der I/O-Graph auf einem S/T-Netz, dessen Elemente jedoch anders interpretiert werden. Wie im Folgenden gezeigt wird, kann aber jeder I/O-Graph in ein Stoffstromnetz transformiert werden, das ausschließlich linear spezifizierte Transitionen enthält. Dazu müssen seine *Prozessknoten* in linear spezifizierte Transitionen umgewandelt werden, deren Koeffizienten sich aus den Gewichten der Kanten des I/O-Graphen ergeben. Anders als beim Stoffstromnetz, werden den Kanten des I/O-Graphen nämlich nicht die Stoffströme $x_k^{(i)}$ und $y_k^{(i)}$, sondern die Transformationskoeffizienten $a_k^{(i)}$ und $b_k^{(i)}$

²⁶ Zur Verwendung des Stoffbegriffs bei umweltorientierten Stoffstromanalysen vgl. Schmidt (2008c: S.9-14).

zugeordnet. Zuletzt werden die *Objektknoten* des I/O-Graphen in die Stellen eines Stoffstromnetzes überführt. Der abstrakte I/O-Graph kann daher als Subklasse allgemeiner Stoffstromnetze verstanden werden (Abbildung 2.15).

Auf diesem Weg lässt sich weiterhin der Bezug zu anderen Standardmodellen der Produktionsplanung herstellen. So können I/O-Graphen ohne Mehrertragsprozesse in lineare I/O-Modelle (Leontief-Modelle) überführt werden. Gibt es darüber hinaus keine zyklischen Stoffströme innerhalb des Produktionssystems, kann das Leontief-Modell eindeutig in einen Gozintographen überführt werden (vgl. Kloock, 1969: S.89). Abbildung 2.15 fasst die Zusammenhänge zusammen und macht – im Sinne von Teilmengen – deutlich, dass Stoffstromnetze all diese Modellklassen als Sonderfälle beinhalten. Dadurch erschließt sich grundsätzlich ein sehr breites Anwendungsspektrum für Stoffstromnetze im Bereich der betrieblichen Produktionsplanung.

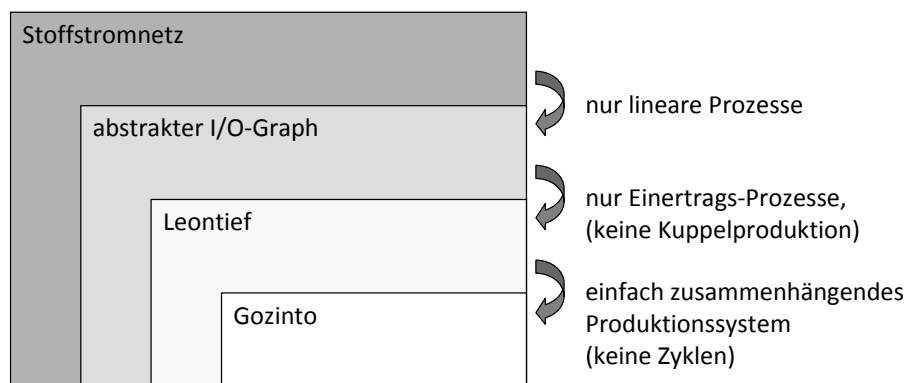


Abbildung 2.15: Verschiedene Modellklassen für die Analyse von Produktionssystemen

Der Periodenbezug von Stoffstromnetzen

Die Anwendbarkeit der Stoffstromnetzmodellierung für die Produktionsplanung wird hingegen durch den Periodenbezug der Stoffstromnetze eingeschränkt. Indem bei der Modellbildung mehrere Perioden betrachtet und sukzessive unter Verwendung der Ergebnisse der jeweils vorhergehenden Periode berechnet werden, können mit Stoffstromnetzen zwar dynamische Modelle erzeugt werden. Ereignisdiskrete Simulationen, wie sie für die Losgrößen-, Ablauf- oder Maschinenbelegungsplanung benötigt werden, können mit Stoffstromnetzen – im Gegensatz zu den ursprünglichen Petrinetzen – jedoch nicht adäquat abgebildet werden. In der Vergangenheit sind die auftragsbezogene Sichtweise der diskreten Simulation und die stoffstromorientierte Perspektive des betrieblichen Umweltschutzes lediglich mit dem Ziel kombiniert worden, die bewährten ökologischen und kostenrechnerischen Bewertungsmethoden der Stoffstromnetze auf die Berechnungsergebnisse eines diskreten Stoffstromsimulators anwenden zu können (Wohlgemuth, 2005:307f).

Daher eignen sich Stoffstromnetze vor allem für die Produktionsplanung in der Prozessindustrie, in der vorwiegend amorphe Arbeitsgüter wie Fluide oder Schüttgüter in kontinuierlichen bzw. (quasi)kontinuierlichen Prozessen verarbeitet werden (Heidepriem, 2000: S.38). Der Zustand

solcher Systeme entspricht zu jedem Zeitpunkt in guter Näherung dem zeitlichen Mittelwert über den gesamten Betrachtungszeitraum. Weniger geeignet sind Stoffstromnetze dagegen für die Planung diskontinuierlicher Einzelstück- oder Chargenprozesse, wie sie häufig in der Fertigungsindustrie bei der Formgebung oder Montage von Werkstücken vorkommen. Für deren Planung müssen die zeitlichen Schwankungen der Stoff-„Ströme“ berücksichtigt werden.

2.5 Zusammenfassung

Zu Beginn dieses Kapitels wurden zunächst die allgemeinen Prinzipien der modellbasierten Planung eingeführt. Demnach beruhen alle rationalen Entscheidungen auf Modellen, mit denen die Konsequenzen verschiedener Handlungsoptionen antizipiert werden können. In alltäglichen Situationen und bei einfachen Planungsproblemen handelt es sich dabei meistens um mentale Modelle. Im Kontext dieser Arbeit, die sich mit der Planung komplexer Produktionssysteme beschäftigt, bedeutet modellbasierte Planung jedoch immer die Planung mit expliziten Modellen wie linearen I/O-Modellen, Stoffstromnetzen oder mathematischen Programmen. Auf der Grundlage des Paradigmas der modellbasierten Planung kann das Verhältnis von Stoffstromanalyse und Optimalplanung im Hinblick auf die Methodenintegration der folgenden Kapitel weiter präzisiert werden.

Optimalplanung ist eine spezielle Form der modellbasierten Planung: Planung mit Optimierungsmodellen. Jedes Optimierungsmodell verfügt über einen Lösungsraum, der verschiedene Handlungsoptionen abbildet. Im Falle der Parameteroptimierung werden alle Lösungen durch Entscheidungsvariablen parametrisiert und durch eine Zielfunktion bewertet. Diese Struktur erlaubt die Anwendung mathematischer Verfahren zur *automatischen* Modelloptimierung, die Rückschlüsse auf Lösungen für das ursprüngliche Planungsproblem zulässt.

In Abschnitt 2.3 wurde die Stoffstromanalyse als zentrales Instrument der Industrial Ecology zur Analyse und Planung industrieller Metabolismen eingeführt. Per Konstruktion ist die Stoffstromanalyse ideal für die Analyse der vernetzten Energie- und Materieströme in Produktionssystemen geeignet. Überdies ermöglicht sie jedoch auch die Integration betriebswirtschaftlicher oder verfahrenstechnischer Analysemethoden und eignen sich daher in besonderer Weise als Werkzeug für eine ganzheitliche Planung moderner Produktionssysteme, die neben ökologischen auch wirtschaftliche und technische Aspekte einbezieht.

In Abschnitt 2.4 wurde schließlich die spezielle Stoffstromanalysemethode der Stoffstromnetze eingeführt, auf der die Methodenentwicklung der folgenden Kapitel aufbaut. Dabei wurde insbesondere gezeigt, wie mit S/T-Netzen die Struktur von Produktionssystemen erfasst und die Funktionsweise einzelner Prozesse durch spezifizierte Transitionen modelliert werden kann. Darüber hinaus wurde die konzeptionelle Kompatibilität der Stoffstromnetze mit verschiedenen Planungsmodellen der betriebswirtschaftlichen Produktionstheorie nachgewiesen und damit gezeigt, dass die stoffstromnetzbasierete Stoffstromanalyse in der Tat ein geeigneter Ansatz für eine integrierte ökonomisch-ökologische Optimalplanung von Produktionssystemen dargestellt. Auf-

grund ihres Periodenbezugs sind die Stoffstromanalyse und die Methode der Stoffstromnetze im Speziellen hingegen nicht für die Lösung ereignisdiskreter Planungsprobleme geeignet. Potenzielle Anwendungsfelder liegen daher eher in der Prozessindustrie.

Sowohl die Stoffstromanalyse als auch die Optimalplanung basieren auf der Verwendung von Modellen. Bei der Abstraktion setzen sie entsprechend ihrem jeweiligen Analyseziel jedoch unterschiedliche Schwerpunkte (Abbildung 2.16). Die auf die Systemoptimierung ausgelegte Optimalplanung richtet ihren Fokus auf die Abbildung von Handlungs- und Gestaltungsoptionen, sowie auf die Bewertung alternativer Systemzustände. Die aus den Umweltwissenschaften hervorgegangene Stoffstromanalyse untersucht Produktionssysteme dagegen vor allem unter dem Gesichtspunkt der zwischen den Systemkomponenten und mit der Systemumgebung ausgetauschten Energie- und Stoffströme. Für eine Kombination beider Ansätze bietet es sich daher an, die Stoffstromanalyse als spezielle Form der Systemanalyse in die Optimalplanung zu integrieren.

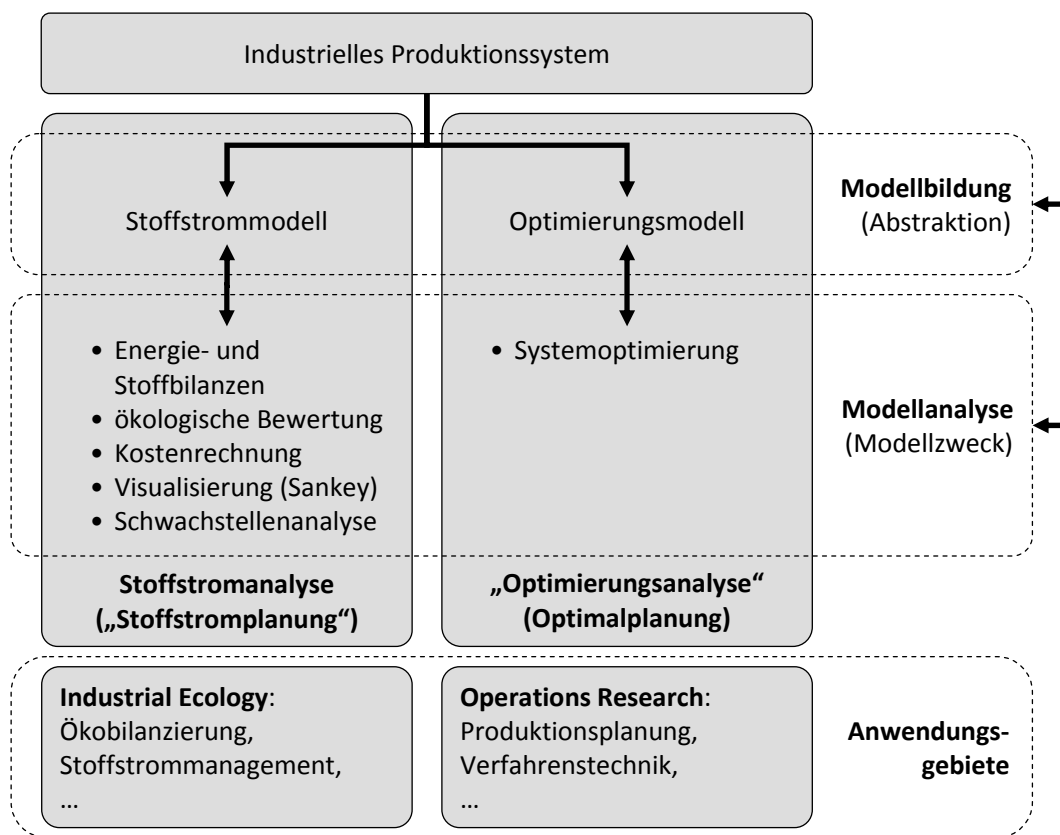


Abbildung 2.16: Die Bestandteile der stoffstrombasierten Optimalplanung

3 Stoffstromnetzbasierende Formulierung von Optimierungsmodellen für den betrieblichen Material- und Energieeinsatz

In Kap. 2 wurden Stoffstromanalyse und Optimalplanung in den jeweiligen Kontext von Industrial Ecology und Operations Research eingebettet. Auf dieser Grundlage wird in diesem Kapitel die Integration beider Ansätze auf der Ebene der *Modellbildung* vollzogen. Die spezielle Stoffstromanalysemethode der Stoffstromnetze wird verwendet, um Optimierungsmodelle für den Energie- und Materialverbrauch industrieller Produktionssysteme zu formulieren. Dabei sind zwei Aspekte zu beachten:

(i) Stoffstromnetze müssen den *formalen Kriterien* mathematischer Optimierungsmodelle genügen, damit entsprechende Optimierungsverfahren auf sie angewendet werden können. In Abschnitt 3.1 wird daher, ausgehend von der Interpretation des Stoffstromnetzes als Erklärungsmodell, dieses schrittweise in die Standardform für Optimierungsmodelle überführt. Dabei zeigt sich, dass Stoffstromnetze in der Regel zu kontinuierlichen Parameteroptimierungsproblemen führen. Dies ist ein erster wichtiger Anhaltspunkt für die Auswahl geeigneter Optimierungsverfahren in Kap. 4. (ii) Für die praktische Einsetzbarkeit der stoffstromnetzbasierenden Optimalplanung muss außerdem untersucht werden, wie reale Planungsprobleme bezüglich der Steuerung und Gestaltung industrieller Produktionssysteme mit Stoffstromnetzen abgebildet werden können. Dies erfordert eine detaillierte Auseinandersetzung damit, wie die modellbildenden Elemente der Stoffstromnetze im Sinne einer Optimalplanung eingesetzt werden können.

Die Abschnitte 3.2 und 3.3 behandeln daher die Frage, wie mit Stoffstromnetzen *sinnvolle* Optimierungsmodelle für Produktionssysteme formuliert werden. „Sinnvoll“ bedeutet, dass stärker auf die semantische Ebene der Modellbildung eingegangen wird. Abschnitt 3.2 behandelt die Frage, wie die für jede Optimierung wesentlichen *Spielräume* für die Gestaltung oder den Betrieb eines Produktionssystems mit Stoffstromnetzen abgebildet werden. In Abschnitt 3.4 wird gezeigt, wie mit Stoffstromnetzen berechnete ökonomische und ökologische *Zielgrößen* zu sinnvollen Optimierungsproblemen kombiniert werden. In Abschnitt 3.5 wird die stoffstromnetzbasierende Formulierung von Optimierungsmodellen am Beispiel eines fiktiven Verbundproduktionssystems aus der Prozessindustrie illustriert.

3.1 Das Stoffstromnetz als Kern eines Parameter- optimierungsmodells

In diesem Abschnitt wird gezeigt, wie Stoffstromnetze in Optimierungsmodelle eingebettet werden können. Die Überführung eines Stoffstromnetzes in ein stoffstromnetzbasierendes Optimierungsmodell kann mehrere Stufen umfassen. Zunächst kann prinzipiell jedes Stoffstromnetz als Optimierungsmodell *interpretiert* werden (§ 3.1.1). Das ist eine wichtige Grundvoraussetzung für die Verwendung von Stoffstromnetzen für die mathematische Optimierung. In der Regel sind Ergänzungen erforderlich, um ein Stoffstromnetz in ein sinnvolles Optimierungsmodell zu überführen. Die einfachste Erweiterung eines als Kernmodell verstandenen Stoffstromnetzes besteht in der Einführung von Ober- und Untergrenzen bezüglich einzelner Modellvariablen. Außerdem können algebraische Ausdrücke mehrerer Modellvariablen verwendet werden, um eine Zielfunktion oder zusätzliche Restriktionen zu formulieren (§ 3.1.2). Auch speziell für das Stoffstrommanagement entwickelte stoffstromnetzbasierende Auswertungsmethoden können in die Modellbildung einbezogen werden (§ 3.1.3). All diese modellbildenden Elemente werden schließlich zur allgemeinen Definition des stoffstromnetzbasierten Optimierungsproblems (SSNOP) zusammengefasst (§ 3.1.4). Damit wird die formale Grundlage für die Anwendung mathematischer Optimierungsalgorithmen auf Stoffstromnetze in den späteren Kapiteln gelegt.

3.1.1 Das Stoffstromnetz als Optimierungsmodell

Das Stoffstromnetz als quantitatives Erklärungsmodell

Stoffstromnetze werden überwiegend als Erklärungsmodelle eingesetzt (vgl. Abschnitt 2.4.3). Die Modellierung einzelner Teilprozesse eines Produktionssystems durch spezifizierte Transitionen ermöglicht die Berechnung unbekannter Stoffströme z_k und Stoffbestände s_j aus bekannten Systemattributen. Diese werden bei der Modellierung mit Stoffstromnetzen entweder als manuell vorgegebene Stoffströme und -bestände oder als Netz- und Transitionsparameter p^N bzw. p^T modelliert. Der Zustand eines Stoffstromnetzes wird durch den Vektor $Z = (\mathbf{z}, \mathbf{s}, \mathbf{p}^T, \mathbf{p}^N) \in R^u$ vollständig beschrieben.

Anstatt die Modellvariablen des Stoffstromnetzes nach semantischen Kriterien in Stoffströme, Stoffbestände und sonstige Modellparameter zu unterteilen, kann man sie, im Sinne eines Erklärungsmodells, in *erklärende Variablen* e_i und *abhängige Variablen* bzw. Ergebnisgrößen r_j unterteilen²⁷.

²⁷ Diese Zuordnung ist im Allgemeinen nicht eindeutig. Bei einer linear spezifizierten Transition werden beispielsweise durch die Vorgabe eines beliebigen anliegenden Stoffstromes alle weiteren Stoffströme eindeutig festgelegt. Jeder der anliegenden Stoffströme kann a priori also erklärende Variable oder Ergebnisgröße sein. Welcher Stoffstrom *sinnvollerweise* als erklärende Variable aufzufassen ist, kann auf der

In seiner Eigenschaft als Erklärungsmodell wird das Stoffstromnetz im Folgenden als Vektorfunktion $\mathbf{r}^S(\mathbf{e})$ dargestellt²⁸. \mathbf{r}^S wird durch die *Struktur* des Stellen-Transitions-Netzes sowie die *Spezifikation* einzelner Transitionen definiert. Die Auswertung von \mathbf{r}^S erfolgt durch die Datenausgleichsrechnung. Der Modellzustand wird in dieser Perspektive durch den Vektor $Z = (\mathbf{e}, \mathbf{r}^S) \in R^u$ beschrieben und durch den Vektor der erklärenden Variablen eindeutig festgelegt: $Z = Z(\mathbf{e})$.

Kombiniert man semantische und funktionelle Gesichtspunkte, so ergibt sich das in Abbildung 3.1 abgebildete 2x4-Raster mit acht verschiedenen Kategorien, denen die Variablen eines stoffstromnetzbasierten Erklärungsmodells zugeordnet werden können. De facto reduziert sich die Anzahl möglicher Kategorien um zwei, weil sich die Werte von Netz- und Transitionsparametern bei der Berechnung eines Stoffstromnetzes nicht ändern und sie daher keine Ergebnisgrößen sein können.

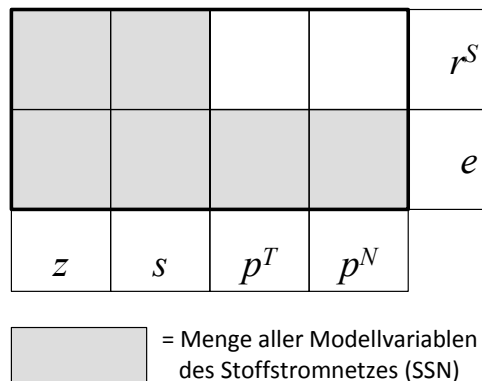


Abbildung 3.1: Das Stoffstromnetz als Erklärungsmodell. Jede Modellvariable kann einer der sechs grau dargestellten Kategorien des abgebildeten 2x4-Rasters zugeordnet werden. Die Anzahl der Variablen, die sich in den jeweiligen Kategorien befinden, ist, anders als es die Flächendarstellung suggeriert, in der Regel nicht gleich verteilt. Im Normalfall gibt es deutlich mehr Flüsse und Bestände als Modellparameter und deutlich mehr Ergebnisgrößen als erklärende Variablen.

Interpretation quantitativer Erklärungsmodelle als Parameteroptimierungsmodelle

Als nächstes wird gezeigt, wie ein beliebiges quantitatives Erklärungsmodell $\mathbf{r}(\mathbf{e})$ mit $e_i, r_j \in R$ als Optimierungsmodell *reinterpretiert* werden kann. Der erklärende Vektor e indiziert alle denkbaren Zustände des Erklärungsmodells $\mathbf{r}(\mathbf{e})$. Es bietet sich daher an, eine beliebige Teilmenge der e_i als Entscheidungsvektor \mathbf{x} zu *interpretieren*. Die übrigen Komponenten werden dann als Parameter p des Optimierungsproblems betrachtet, d.h. bei der Suche nach optimalen Lösungen kon-

abstrakten Modellebene nicht entschieden werden, sondern ergibt sich allein aus dem zugrunde liegenden realen Sachverhalt.

²⁸ Im Folgenden steht $\mathbf{r}(\mathbf{e})$ für ein beliebiges Erklärungsmodell. Der hochgestellt Index S verweist dagegen auf ein Stoffstromnetz in seiner Eigenschaft als Erklärungsmodell.

stant gehalten²⁹. Durch entsprechende Umsortierung der Komponenten des erklärenden Vektors kann dieser in die Form $\mathbf{e} = (x_1, \dots, x_n, p_1, \dots, p_m)$ gebracht werden, bei der zuerst die n Entscheidungsvariablen und danach die m Parameter des als Optimierungsmodell interpretierten Erklärungsmodells stehen. Als Zielgröße der Optimierung kommt jede beliebige Komponente des Ergebnisvektors $r_k(\mathbf{e}) = r_k(\mathbf{x}; \mathbf{p})$ in Frage.

Auf diese Weise kann *jedes* quantitative Erklärungsmodell als Optimierungsmodell interpretiert werden. Der Übergang besteht lediglich in einem *Perspektivwechsel*: Anstatt mit dem Modell $\mathbf{r}(\mathbf{e})$ im Sinne einer Wirkungsprognose den Systemzustand Z_0 zu berechnen, der sich für einen bestimmten erklärenden Vektor e_0 ergibt, wird bei der Optimierung nach dem Entscheidungsvektor \mathbf{x}^* gesucht, bei dem sich das Modell im optimalen Zustand Z^* befindet, der durch ein Maximum bzw. Minimum der Zielfunktion $r_k(\mathbf{x}^*; \mathbf{p})$ gekennzeichnet ist.

Modellinterne Restriktionen und der Modellgültigkeitsbereich F^M

Bis hierhin wurde angenommen, dass der Definitionsbereich des Erklärungsmodells r dem gesamten Wertebereich von x entspricht und sich daher das unrestringierte Optimierungsproblem (3.1) ergibt.

$$\begin{aligned} \min/\max \quad & r_k(\mathbf{x}; \mathbf{p}) \\ & \mathbf{x} \in L = R^n \end{aligned} \tag{3.1}$$

Diese Annahme ist im Allgemeinen jedoch falsch. Die *formale Struktur* und die *Semantik* des Erklärungsmodells können dazu führen, dass nicht alle Entscheidungsvektoren x zu einem *gültigen* oder *sinnvollen* Modellergebnis führen. Das gilt insbesondere, wenn es sich bei dem Erklärungsmodell um ein komplexes Computermodell handelt. Außerhalb des *Modellgültigkeitsbereichs*³⁰ $F^M \subset L$ kann $r(\mathbf{x}; \mathbf{p})$ unter Umständen nicht einmal ausgewertet werden, z.B. weil Auswertungsalgorithmen ergebnislos abgebrochen werden oder zu fehlerhaften Ergebnissen führen. Die Reinterpretation eines Erklärungsmodells als Optimierungsmodell führt daher im Allgemeinen zum restringierten Optimierungsproblem (3.2):

²⁹ Wie bei der Auswahl der erklärenden Variablen selbst, gibt es auf der formalen Modellebene keinerlei Anhaltspunkte dafür, welche der erklärenden Variablen sinnvollerweise als Entscheidungsvariable und welche als Parameter der Optimierung zu betrachten ist. Wiederum kann dies nur anhand des zugrunde liegenden *realen Problems* entschieden werden. Als Entscheidungsvariablen werden typischerweise nur Modellgrößen in Frage kommen, die *beeinflussbare* Attribute des Realsystems repräsentieren.

³⁰ Die Bezeichnung stammt aus der Arbeit von Peters et al. (1999: S.127f) zur modellbasierten Designoptimierung von Mikrosystemen. Im mathematischen Sinne definiert der Modellgültigkeitsbereich den Definitionsbereich der Funktion $r(e)$.

$$\begin{aligned} \min/\max \quad & r_k(\mathbf{x}; \mathbf{p}) \\ \text{Nb.} \quad & \mathbf{x} \in F^M \subseteq R^n \end{aligned} \tag{3.2}$$

Beispiel für eine modellinterne Restriktion eines Stoffstromnetzes

Wie eine solche Einschränkung der Lösungsmenge durch modellinterne Restriktionen aussehen kann, wird hier am Beispiel des in Abbildung 3.2 dargestellten Stoffstroms $x_k^{(t)}$ gezeigt. Es wird angenommen, dass $x_k^{(t)}$ einen materiellen Stoffstrom repräsentiert.

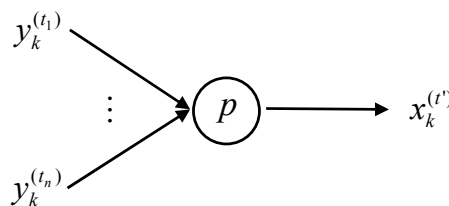


Abbildung 3.2: Lokale Modellstruktur zur Ableitung der modellinternen Restriktionen bezüglich des Stoffstroms $x_k^{(t)}$

Gemäß einer semantischen Regel von Stoffstromnetzen können Masse- und Energieströme nicht negativ werden. Daher gilt: $x_k^{(t)} \geq 0$. Gleichermaßen unzulässig sind negative Bestände an Stellen. Daraus und aus der lokalen Modellstruktur folgt eine Obergrenze: Der Output $x_k^{(t)}$ der Stelle p kann den Anfangsbestand zuzüglich der im Modellierungszeitraum erfolgten Zuflüsse von anderen Transitionen nicht übersteigen. Zusammenfassend ergibt sich Restriktion (3.3), die den ursprünglichen Definitionsbereich $x_k^{(t)} \in R$ deutlich einschränkt.

$$0 \leq x_k^{(t)} \leq \sum_{t \rightarrow p} y_k^t + s_k^p(\tau_0) \tag{3.3}$$

Das Stoffstromnetz als Parameteroptimierungsmodell

Gleichung (3.2) kann unmittelbar auf *Stoffstromnetze* übertragen werden, sofern diese Erklärungsmodelle sind:

$$\begin{aligned} \min/\max \quad & r_k^S(\mathbf{x}; \mathbf{p}) \\ \text{Nb.} \quad & \mathbf{x} \in F^S \subseteq R^n \end{aligned} \tag{3.4}$$

Da alle Modellvariablen eines Stoffstromnetzes quantitative Variablen sind, gehört (3.4) zur Klasse der *Parameteroptimierungsprobleme*. Stoffstromnetze verwenden bislang ausschließlich reelle

Variablen zur Quantifizierung von Stoffströmen, -beständen und sonstigen Modellparametern, unabhängig davon, ob die realen Systemattribute diskret oder kontinuierlich sind. Aus formaler Sicht ist das wegen $Z \subset R$ problemlos möglich. Im Hinblick auf die Klassifizierung des Optimierungsproblems (3.4) lässt sich daher nicht pauschal festlegen, ob es sich um ein kontinuierliches, ganzzahliges oder gemischt-ganzzahliges Parameteroptimierungsproblem handelt.

3.1.2 Erweiterungen des Kernmodells um zusätzliche Restriktionen und eine algebraische Zielfunktion zur Bewertung des Systemzustands

Ober- und Untergrenzen bezüglich einzelner Modellvariablen

Die modellinternen Restriktionen des Stoffstromnetzes $\mathbf{r}^S(\mathbf{e})$ betreffen die *Struktur* eines Produktionssystems, *technische Eigenschaften* seiner Teilprozesse sowie naturwissenschaftliche Grundprinzipien der Stoffumwandlung wie die Erhaltung von Masse und Energie. Nicht jede beliebige Kombination von Stoffströmen und -beständen ist daher mit dem Modell $\mathbf{r}^S(\mathbf{e})$ vereinbar (Lambrech et al., 2009: S.66-71).

Darüber hinaus können sich bei der Planung industrieller Produktionssysteme weitere Einschränkungen ergeben. Typische Beispiele sind die begrenzte Verfügbarkeit von Produktionsfaktoren, gesetzliche Emissionsgrenzwerte, vertraglich festgelegte Mindestliefermengen oder begrenzte Kapazitäten von Anlagen. Derartige Restriktionen können zusätzlich zum Kernmodell (3.4) als Ober- und Untergrenzen bezüglich einzelner Modellvariablen formuliert werden:

$$\begin{aligned} \mathbf{l} &\leq (\mathbf{r}^S, \mathbf{x}) \leq \mathbf{u}, \\ \mathbf{r}^S &\in R^m, \mathbf{x} \in R^n, \mathbf{l}, \mathbf{u} \in R^{n+m} \cup \{-\infty, +\infty\} \end{aligned} \tag{3.5}^{31}$$

Verwendung algebraischer Ausdrücke mehrerer Modellvariablen des Stoffstromnetzes zur Formulierung einer Zielgröße und zusätzlicher Restriktionen

Für die Optimalplanung eines Produktionssystems benötigt man Zielgrößen, die den Zustand Z des gesamten Systems hinreichend gut repräsentieren. Einzelne Stoffströme bzw. -bestände, wie sie in (3.4) als Zielfunktion verwendet werden, sind dafür nicht immer geeignet. Für praktische Anwendungen kann das Erklärungsmodell $\mathbf{r}^S(\mathbf{e})$ daher um die Definition geeigneter Zielgrößen $\varphi(Z)$ erweitert werden. Im einfachsten Fall werden solche Zielgrößen durch algebraische Ausdrücke der Modellgrößen $(\mathbf{r}^S, \mathbf{e})$ des Stoffstromnetzes definiert: $\varphi^a(\mathbf{r}^S, \mathbf{e})$.

Algebraische Zielgrößen können einerseits als Zielfunktion f des Optimierungsproblems oder zur Formulierung weiterer Restriktionen g und h verwendet werden. Berücksichtigt man diese Erweiterung des Kernmodells in der formalen Definition des Optimierungsmodells, so wird aus dem als

³¹ Die Ungleichungen für Vektoren in (3.5) sind komponentenweise zu verstehen.

Optimierungsmodell interpretierten Stoffstromnetz (3.4) das *stoffstromnetzbasierende Optimierungsmodell* (3.6):

$$\begin{aligned}
 \min/\max \quad & f^a(\mathbf{r}^S, \mathbf{x}; \mathbf{p}) \\
 \text{Nb.} \quad & \mathbf{x} \in F^S \subseteq R^n \\
 & g^a(\mathbf{r}^S, \mathbf{x}) \leq 0 \\
 & h^a(\mathbf{r}^S, \mathbf{x}) = 0 \\
 & l \leq (\mathbf{r}^S, \mathbf{x}) \leq u
 \end{aligned}
 \tag{3.6}$$

Mit (3.6) wird das Stoffstromnetz $\mathbf{r}^S(\mathbf{x}; \mathbf{p})$ in ein algebraisches Optimierungsmodell eingebettet.

3.1.3 Einbeziehung stoffstromnetzspezifischer Auswertungsmethoden in die Formulierung des Optimierungsmodells

Der Vorteil algebraischer Ausdrücke ist ihre Flexibilität bei der Zielgrößendefinition. Die Möglichkeiten zur Kombination von Stoffstrom- und Stoffbestandsvariablen sind praktisch unbegrenzt. Bei Systemen mit vielen Teilprozessen und Stoffströmen können die entsprechenden algebraischen Ausdrücke jedoch sehr umfangreich werden. Häufig müssen ineinander verschachtelte Terme verwendet werden. Das vermindert die Transparenz des Modells und erhöht damit seine Fehleranfälligkeit bei der Modellbildung.

Für die stoffstromnetzbasierende Zielgrößendefinition gibt es noch eine weitere Möglichkeit. Im Kontext des betrieblichen Stoffstrommanagements sind in der Vergangenheit verschiedene *stoffstromnetzbezogene Auswertungsmethoden* entwickelt worden, mit denen die ökologische Wirkung bzw. ökonomische Performance von Produktionssystemen quantifiziert werden können (vgl. Abbildung 3.3). So können *Energie- und Stoffbilanzen* für das gesamte Modell oder ausgewählte Teilbereiche erstellt werden. Die *Leistungsverrechnung* ermöglicht den Bezug solcher Bilanzen auf einzelne Referenzströme innerhalb des Systems. Dies ist eine notwendige Voraussetzung sowohl für die *produktbezogene Ökobilanzierung* als auch eine stoffstrombasierte *Kostenträgerrechnung*. Die erzeugten Energie- und Stoffbilanzen können anschließend monetär oder hinsichtlich der mit ihnen verbundenen ökologischen Wirkungen bewertet werden.

Gerade die ökologische Bewertung erfordert typischerweise die Einbeziehung umfangreicher Datenbanken und weiterer Modelle. Zum Einen um Umweltwirkungen durch Prozesse zu berücksichtigen, die nicht Teil des modellierten Produktionssystems sind, mit diesem jedoch indirekt über die Bereitstellung von Gütern oder die Entsorgung entstandener Abfälle zusammenhängen. Zum Anderen um die tatsächliche Auswirkung der in der Sachbilanz ausgewiesenen materiellen Stoffströme auf die Umwelt abzuschätzen.

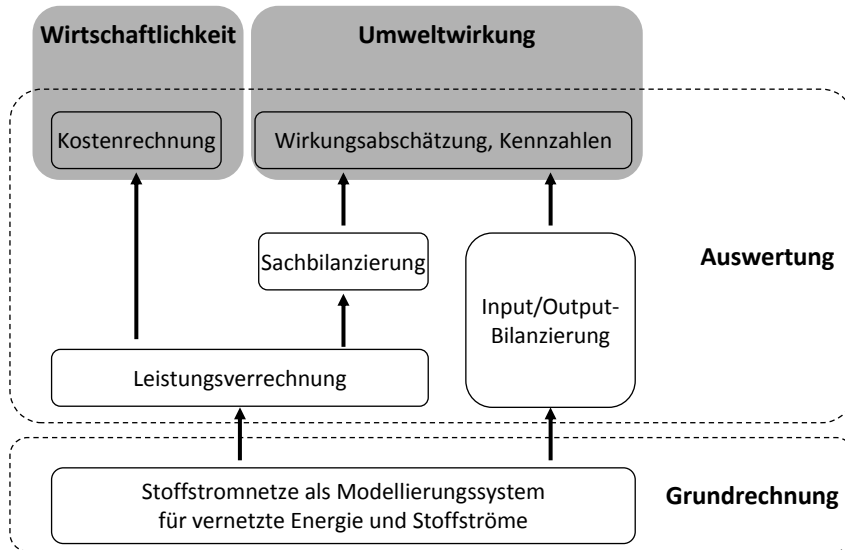


Abbildung 3.3: Stoffstromnetze als Grundrechnung zur Erfassung der Energie- und Stoffströme eines Produktionssystems. Aussagen zur Wirtschaftlichkeit oder ökologischen Wirkung des Systems lassen sich erst durch weiter gehende Auswertungen machen (modifiziert nach Möller, 2000: S.44, Abb.7).

Die Ergebnisse quantitativer Modellauswertungen, wie sie in Abbildung 3.3 dargestellt sind, werden im Folgenden durch die Vektorfunktion $a(Z)$ repräsentiert. Hinter der einfachen Darstellung von $a(Z)$ verbergen sich komplexe Berechnungsalgorithmen und Datenbanken, die nicht durch einfache algebraische Ausdrücke $f^a(Z)$ ersetzt werden können. $a(Z)$ bzw. einzelne Komponenten $a_i(Z)$ können vielmehr als weiterer Bestandteil stoffstromnetzbasierter Optimierungsprobleme in die Formulierung algebraischer Ausdrücke für Zielfunktion und Nebenbedingung integriert werden und stellen in diesem Sinn ein wertvolles Hilfsmittel für die Konstruktion aussagekräftiger Zielgrößen für die ökologische und ökonomische Optimierung des Material- bzw. Energieverbrauchs industrieller Produktionssysteme dar.

3.1.4 Allgemeine Definition des stoffstromnetzbasierten Optimierungsproblems

Durch die Integration der Ergebnisse stoffstromnetzspezifischer Auswertungsmethoden $a(Z)$ in die Form (3.6) erhält man schließlich das stoffstromnetzbasierende Optimierungsproblem (SSNOP) in seiner allgemeinsten Form (3.7).

$$\begin{aligned}
 \text{min/max } & f^a(\mathbf{a}, \mathbf{r}^S, \mathbf{x}; \mathbf{p}) && \text{(a)} \\
 \text{Nb. } & \mathbf{x} \in F^S \subseteq R^n && \text{(b)} \\
 & g^a(\mathbf{a}, \mathbf{r}^S, \mathbf{x}) \leq 0 && \text{(c)} \\
 & h^a(\mathbf{a}, \mathbf{r}^S, \mathbf{x}) = 0 && \text{(c)} \\
 & \mathbf{l} \leq (\mathbf{a}, \mathbf{r}^S, \mathbf{x}) \leq \mathbf{u} && \text{(d)}
 \end{aligned}
 \tag{3.7}$$

Alle Variablen und Komponenten der Vektorfunktionen in (3.7) sind reelle Zahlen. Das SSNOP ist daher ein *kontinuierliches Parameteroptimierungsproblem*. Hinter den wie gewöhnliche Funktionen mehrerer Variablen aussehenden Zielgrößen $\varphi=(f, \mathbf{g}, \mathbf{h})$ (3.7,a und c) verbergen sich tatsächlich auf mehreren Ebenen ineinander verschachtelte Funktionen (Abbildung 3.4). Daher kann die Auswertung stoffstromnetzbasierter Zielgrößen bis zu drei aufeinander aufbauende Berechnungsschritte involvieren: die Berechnung unbekannter Stoff- und Energieströme des Stoffstromnetzes mit der Datenausgleichsrechnung, die Anwendung spezifischer Auswertungsmethoden und schließlich die Auswertung der algebraischen Zielgröße $\varphi^a(\mathbf{a}, \mathbf{r}^S, \mathbf{x}; \mathbf{p})$.

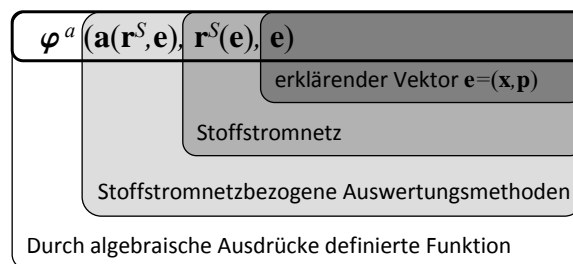


Abbildung 3.4: Die verschiedenen Modellbildungsebenen zur Formulierung einer stoffstromnetzbasierten Zielfunktion.

Das SSNOP (3.7) ist die für die Anwendung von Optimierungsverfahren auf Stoffstromnetze benötigte *formale Struktur*. Es bildet die *Schnittstelle* zwischen Stoffstromanalyse und Optimalplanung, indem es das Stoffstromnetz r^S , eine Zielfunktion (3.7,a) sowie verschiedenen Arten von Restriktionen (3.7,b-d) zur Definition eines Optimierungsproblems kombiniert (Abbildung 3.5).

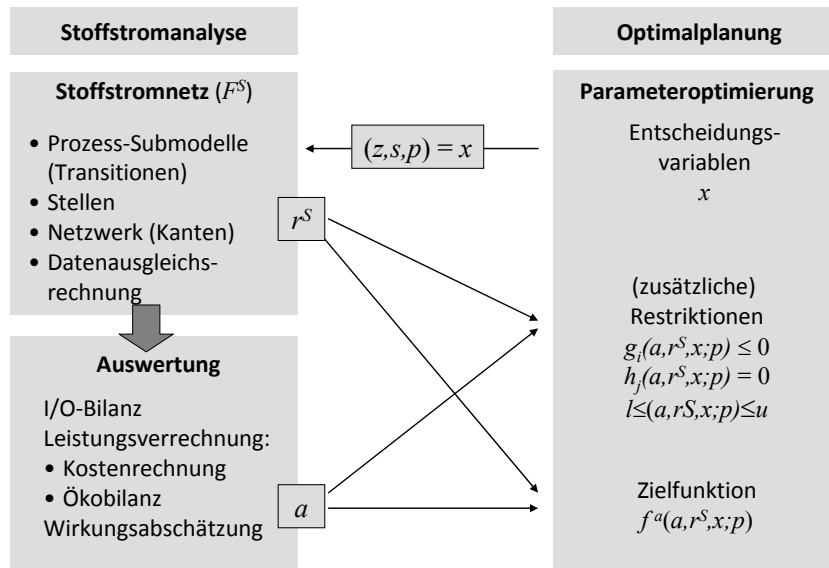


Abbildung 3.5: Ergänzung des Stoffstromnetzes zum stoffstromnetzbasierten Parameteroptimierungsmodell (verändert aus Schmidt et al., 2007: S.275). Entscheidungsvariablen x sind im Stoffstromnetz vorhandene Flüsse (z), Bestände (s) bzw. Netz- oder Transitionsparameter (p). Der Vektor r^S enthält die quantitativen Ergebnisse der Stoffstromnetzberechnung. Die Ergebnisse darauf aufbauender Auswertungsmethoden werden im Vektor a zusammengefasst.

3.2 Modellierung von Handlungs- und Gestaltungsspielräumen

Mit einer stoffstromnetzbasierenden Optimalplanung wird das Ziel einer Systemoptimierung verfolgt, weshalb es im Gegensatz zu herkömmlichen Stoffstromanalysen nicht ausreichend ist, das analysierte Produktionssystem ausschließlich im Status Quo abzubilden. Vielmehr müssen in das Modell alternative Systemzustände integriert werden, aus denen durch den Einsatz entsprechender Algorithmen der optimale Zustand ausgewählt werden kann. Daher befasst sich dieser Abschnitt mit der Frage, wie mit den Modellierungselementen von Stoffstromnetzen systematisch *Systemfreiheitsgrade*, d.h. Spielräume zur Verbesserung der Steuerung und der Gestaltung industrieller Produktionssysteme, abgebildet werden können.

Wie aus den Überlegungen des vorhergehenden Abschnitts hervorgeht, können Systemfreiheitsgrade bei einem stoffstromnetzbasierenden Optimierungsmodell ausschließlich auf der Ebene des Stoffstromnetzes modelliert werden (vgl. Abbildung 3.4). Alle darauf aufbauenden Schritte der Modellbildung wie die Anwendung stoffstromnetzspezifischer Auswertungsmethoden $a(Z)$ oder die Formulierung algebraischer Zielgrößen $\varphi(a, Z)$ dienen entweder der *Bewertung* bereits eindeutig festgelegter Systemzustände oder einer noch stärkeren Einschränkung der Lösungsmenge ($F \subseteq F^S$) durch zusätzliche algebraischer Restriktionen (3.7,c und d).

Dem Modellgültigkeitsbereich F^S eines Stoffstromnetzes kommt daher eine zentrale Bedeutung bei der Optimalplanung zu. Er gibt den Rahmen vor, innerhalb dessen mit dem stoffstromnetzbasierenden Optimierungsmodell überhaupt optimale Lösungen gefunden werden können. Nur wenn

Handlungs- und Gestaltungsoptionen für die Verbesserung eines Produktionssystems bei der Modellbildung hinreichend berücksichtigt werden, kann die Anwendung mathematischer Optimierungsalgorithmen zu befriedigenden Lösungen für die Praxis führen. In diesem Abschnitt wird daher untersucht, welchen Einfluss die einzelnen modellbildenden Elemente sowie die Struktur eines Stoffstromnetzes auf den Modellgültigkeitsbereich F^S haben.

Der Modellgültigkeitsbereich eines Stoffstromnetzes entsteht durch die Kombination der lokalen Lösungsräume seiner Netzelemente. Daher wird zunächst untersucht, wie viele Freiheitsgrade einzelne Transitionen (§ 3.2.2) und Stellen (§3.2.3) besitzen. Dabei zeigt sich, dass die Freiheitsgrade von der Art der Transitionsspezifikation bzw. vom Stellentyp abhängen. Darauf aufbauend wird eine Formel abgeleitet werden, die die Freiheitsgrade der Netzelemente zu denen des Gesamtmodells kombiniert (§3.2.4).

3.2.1 Die Knoten des Stoffstromnetzes als isolierte Submodelle

Um die Anzahl der Freiheitsgrade einzelner Netzknoten zu bestimmen, müssen diese zunächst gedanklich aus der Netzstruktur herausgelöst werden (Abbildung 3.6). Sowohl Transitionen als auch Stellen werden als Submodelle des Stoffstromnetzes betrachtet, die jeweils n eingehende Stoffströme in spezifischer Weise mit m von ihnen ausgehenden Stoffströmen verknüpfen.

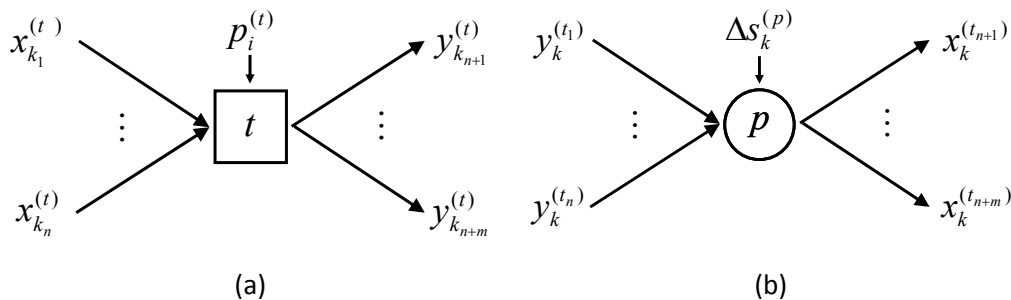


Abbildung 3.6: Transitionen (t) und Stellen (p) als lokale quantitative Erklärungsmodelle. $x_k^{(t)}$ und $y_k^{(t)}$ bezeichnen jeweils Input- bzw. Outputströme des Stoffes k bezüglich der Transition t . $p_i^{(t)}$ sind die Transitionparameter von t und $\Delta S_k^{(p)}$ die Bestandsveränderung des Stoffes k in der Stelle p . Alle diese quantitativen Modellgrößen können Freiheitsgrade repräsentieren.

Da sich die Bipartition der Stoffströmen in Inputs $x(t)$ und Outputs $y(t)$ auf die Transitionen des Netzes bezieht, verhalten sich Transitionen und Stellen hinsichtlich dieser Modellvariablen spiegelverkehrt:

- Die Inputs einer Transition sind in der hier verwendeten skalaren Darstellungsform n skalare Ströme der Stoffe k_1 bis k_n , die aus verschiedenen Stellen hervorgehen. Outputs sind m Flüsse der Stoffe k_{n+1} bis k_{n+m} .

- Die an einer Stelle eingehenden Stoffströme sind dagegen Outputs der gleichen Stoffart k verschiedener Transitionen t_1 bis t_n und die ausgehenden Ströme sind gleichartige Inputs der Transitionen t_{n+1} bis t_{n+m} .

Außer den anliegenden Stoffströmen verfügen Stellen über Stoffbestände bzw. Bestandsveränderung $\Delta s_k^{(p)}$, wodurch sie zu einem lokalen Modell mit $n+m+1$ Variablen werden. Für die Spezifikation von Transitionen können lokale Transitionsparameter $p^{(t)}$ eingeführt werden, weshalb diese über $n+m+p$ Variablen verfügen (wobei $p=|p^{(t)}|$).

3.2.2 Freiheitsgrade von Transitionen

Hinsichtlich der Anzahl der Freiheitsgrade können bei Transitionen zwei Extremfälle unterschieden werden: Bei einer *unspezifizierten Transition* können alle anliegenden Stoffströme frei gewählt werden. Sie besitzt daher genauso viele Freiheitsgrade, wie Stoffströme mit ihr verbunden sind $FG(t)=n+m$. Das andere Extrem ist die *linear spezifizierte Transition*. Sie kann durch ein *einfach unterbestimmtes lineares Gleichungssystem* repräsentiert werden, zu dessen eindeutiger Lösung die Festlegung eines beliebigen anliegenden Stoffstroms genügt. Die Transition hat daher genau einen Freiheitsgrad: $FG(t)=1$. Für durch *benutzerdefinierte Funktionen* oder *Skripte* spezifizierte Transitionen gilt allgemein $0 \leq FG(t) \leq n+m+p$, wobei p die Anzahl der Transitionsparameter ist. Die genaue Anzahl der Freiheitsgrade hängt von der konkreten Spezifikation ab.

Für durch *Subnetze* spezifizierte Transitionen gelten dieselben Regeln zur Bestimmung ihrer Freiheitsgrade, wie für das Stoffstromnetz insgesamt. Daher muss in diesem Fall zum Abzählen der Freiheitsgrade *rekursiv* vorgegangen werden, wobei auf der jeweils untersten Hierarchiestufe zu beginnen ist, d.h. in dem Subnetz, das selbst keine Subnetze mehr enthält. Für dieses können dann die Freiheitsgrade bestimmt werden und für das Abzählen auf der nächst höheren Hierarchiestufe verwendet werden usw.

3.2.3 Freiheitsgrade von Stellen

Normale Stellen und Grenzstellen

Transitionen stehen für Prozesse, in denen Stoffe umgewandelt werden. Sie sind daher Senken für bestimmte Stoffarten und Quellen für andere. Im Gegensatz dazu sind an Stellen die Quantitäten aller Stoffarten *Erhaltungsgrößen*. Daher gilt stets die *Kontinuitätsgleichung* (3.8):

$$s_k^{(p)}(\tau_1) = s_k^{(p)}(\tau_0) + \sum_{t \rightarrow p} y_k^{(t)} \cdot \Delta\tau - \sum_{p \rightarrow t} x_k^{(t)} \cdot \Delta\tau \quad \forall p \in P, k \in M \quad (3.8)$$

Darin sind τ_0 und τ_1 respektive der Start- und Endzeitpunkt der Betrachtungsperiode $\Delta\tau = \tau_1 - \tau_0$, P die Menge aller Stellen und M die Menge aller Stoffarten. Gleichung (3.8) liest sich wie folgt: Der Bestand des Stoffes k in Stelle p am Ende einer Betrachtungsperiode (linke Seite) entspricht dem

Anfangsbestand zuzüglich eingehender und abzüglich abgehender Stoffflüsse (rechte Seite). Anfangs- und Endbestand können zur Bestandsveränderung $\Delta S_k^{(p)}$ zusammengefasst werden:

$$\Delta S_k^{(p)} = \left(\sum_{t \rightarrow p} y_k^{(t)} - \sum_{p \rightarrow t} x_k^{(t)} \right) \cdot \Delta \tau \quad (3.9)$$

Durch diese Erhaltungseigenschaft werden Stellen – genau wie Transitionen – zu *lokalen Erklärungsmodellen*. (3.9) liefert eine Gleichung für $n+m+1$ Unbekannte. Jede normale Stelle hat daher, genau wie eine unspezifizierte Transition, $FG(p) = n+m$ Freiheitsgrade. Das gilt auch für *Input-* und *Outputstellen* mit der Besonderheit, dass diese über keine Eingangs- bzw. Ausgangsknoten verfügen. Daher gilt $FG(p_{In}) = m$ und $FG(p_{Out}) = n$.

Verbindungsstellen

Die *Verbindungsstelle* unterscheidet sich von der „normalen“ Stelle dadurch, dass sie keine Bestände und daher auch keine Bestandsveränderungen haben kann ($\Delta S_k^{(p)} \equiv 0$). Die Kontinuitätsgleichung (3.8) nimmt daher die besonders einfache Form (3.10) an: Summe der Inputs gleich Summe der Outputs.

$$\sum_{t \rightarrow p} y_k^{(t)} = \sum_{p \rightarrow t} x_k^{(t)} \quad (3.10)$$

Weil an der Verbindungsstelle keine Bestandsveränderungen zulässig sind, hat sie einen Freiheitsgrad weniger als die normale Stelle: $FG(p_V) = n+m-1$. Der obere Teil von Tabelle 3.1 fasst die bis hierhin dargestellten Freiheitsgrade von normalen, Grenz- und Verbindungsstellen zusammen. Im unteren Teil sind spezielle *Konstellationen* von Verbindungsstellen gesondert aufgeführt, die bei der Modellierung industrieller Produktionssysteme häufiger auftreten.

Bei der *Verteilerstelle* wird ein eingehender Stoffstrom auf verschiedene Prozesse aufgeteilt. Der Verteiler repräsentiert damit die Entscheidung bezüglich verschiedener Verwendungsoptionen. Bei der *Sammlerstelle* werden umgekehrt gleichartige Stoffströme unterschiedlicher Herkunft gebündelt. Weil das Mischungsverhältnis frei wählbar ist, verfügt die Sammlerstelle über n Freiheitsgrade. Sammlerstellen repräsentieren Entscheidungen zwischen vollständig substituierbaren Produktionsfaktoren. Beispielsweise zwischen elektrischer Energie aus Kohle, Kern- oder Windkraft. Eine besondere Rolle spielt die *direkte Kopplung*. Durch sie wird der Output eines Prozesses mit dem Input eines anderen gleichgesetzt.

Tabelle 3.1: Schematische Darstellung der Stellentypen in verschiedenen Netzstrukturen. Jeweils unter Angabe der algebraischen Ersatzformel und der Anzahl der Freiheitsgrade. $t \rightarrow p$ bezeichnet die Menge der Eingangs- und $p \rightarrow t$ die der Ausgangsknoten.

Funktion	lokale Netzstruktur	Ersatzformel	FG(p)
normale Stelle		(3.8) bzw. (3.9)	$n+m$
Input		$\Delta s_k^{(p)} = -\sum_{p \rightarrow t} x_k^{(t)} \cdot \Delta \tau$	m
Output		$\Delta s_k^{(p)} = \sum_{t \rightarrow p} y_k^{(t)} \cdot \Delta \tau$	n
Verbindung		(3.10)	$n+m-1$
Verteiler		$y_k^{(t)} = \sum_{p \rightarrow t} x_k^{(t)}$	m
Sammler		$\sum_{t \rightarrow p} y_k^{(t)} = x_k^{(t)}$	n
direkte Kopplung		$y_k^{(t_1)} = x_k^{(t_2)}$	1

3.2.4 Die Elimination von Freiheitsgraden durch Flussrelationen

In den beiden vorangegangenen Abschnitten wurde gezeigt, wie die Anzahl der Freiheitsgrade verschiedener Stellen- und Transitionstypen bestimmt werden kann. Dabei wurden die an den Netzknoten ein- und ausgehenden Flussrelationen genau wie der Stoffbestand einer Stelle und

die ggf. vorhandenen Transitionsparameter als quantitative Attribute des Netzknotens betrachtet. Diese Vorgehensweise führt zu einer *Doppelzählung*, sobald zwei Elemente durch dieselbe Flussrelation verbunden sind.

Um die Freiheitsgrade des gesamten Stoffstromnetzes $FG(N)$ zu bestimmen, muss daher von der Summe der Freiheitsgrade der isoliert betrachteten Transitionen und Stellen die Anzahl der Flussrelationen subtrahiert werden. Mit dieser einfachen Abzählregel kann die Dimensionalität eines stoffstromnetzbasierten Optimierungsproblems bestimmt werden.

$$FG(N = \{P, T, F\}) = \sum_{t \in T} FG(t) + \sum_{p \in P} FG(p) - |F| \quad (3.11)$$

P , T und F bezeichnen in Formel (3.11) jeweils die Mengen aller Stellen, Transitionen und Flussrelationen des Stoffstromnetzes N . Die nachträgliche Korrektur von Doppelzählungen durch Subtraktion der Anzahl vorhandener Flussrelationen $|F|$ ist nur gültig, wenn ausschließlich skalare Stoffströme verwendet werden. Bei Vektorflüssen muss jede Flussrelation mit der Anzahl der von Null verschiedenen Komponenten des ihr zugeordneten Flussvektors gewichtet werden.

3.3 Formulierung ökonomischer und ökologischer Zielgrößen

Dieser Abschnitt befasst sich aus einer *semantischen* Perspektive mit der stoffstromnetzbasierten Konstruktion *sinnvoller* Zielgrößen für die ökonomische und ökologische Optimierung von Produktionssystemen. Um die Darstellung so einfach wie möglich zu halten, wird die Frage der Systembewertung zunächst aus einer betrieblichen Perspektive angegangen und die betriebliche Stoffbilanz als Ausgangspunkt gewählt (§3.3.1). Die Zielgrößenformulierung muss nicht nur den objektiven Zustand eines Systems berücksichtigen, sondern auch die subjektiven Werturteile des Entscheidungsträgers. Daher wird in §3.3.2 gezeigt, wie die Präferenzen des Unternehmers in zwei Stufen – zunächst beim Übergang von der Stoffstrom- zur Ergebnisebene und anschließend bei der Aggregation der Produktionsergebnisse zu skalaren Kennzahlen – explizit in die Zielgrößenkonstruktion einfließen. In den beiden darauffolgenden Abschnitten werden auf dieser Grundlage konkrete stoffstromnetzbasierende ökonomische und ökologische Zielgrößen für die Optimierung industrieller Produktionssysteme hergeleitet. Für die ökologische Bewertung wird dabei insbesondere auf Methoden der Ökobilanzierung zurückgegriffen (§3.3.4). Die für betriebliche Stoffbilanzen entwickelten Bewertungsmethoden werden in §3.3.5 schließlich auf den allgemeinen Fall von Stoffstromnetzen übertragen, die betriebsinterne Bestandsveränderungen berücksichtigen oder deren Grenzen nicht mit denen des Betriebs übereinstimmen. Abschließend werden die Grenzen der stoffstromnetzbasierenden Bewertung von Produktionssystemen aufgezeigt.

3.3.1 Die betriebliche Stoffbilanz als Bewertungsgrundlage

Definition der betrieblichen Stoffbilanz

Der Ausgangspunkt für die Bewertung der ökologischen und ökonomischen Performance eines Betriebes ist seine Stoffbilanz. Die betriebliche Stoffbilanz z^S beschreibt den gesamten Material- und Energieumsatz eines Unternehmens in einem gegebenen Zeitraum. Um sie mit einem Stoffstromnetz zu ermitteln, muss eine Bilanzgrenze um alle innerbetrieblichen Transitionen gezogen, und alle sie schneidenden Stoffströme summiert werden. Abbildung 3.7 zeigt beispielhaft ein aus vier Stoffumwandlungsprozessen (T3-T6) bestehendes Produktionssystem. Alle Stoffströme der betrieblichen Stoffbilanz sind fett gedruckt.

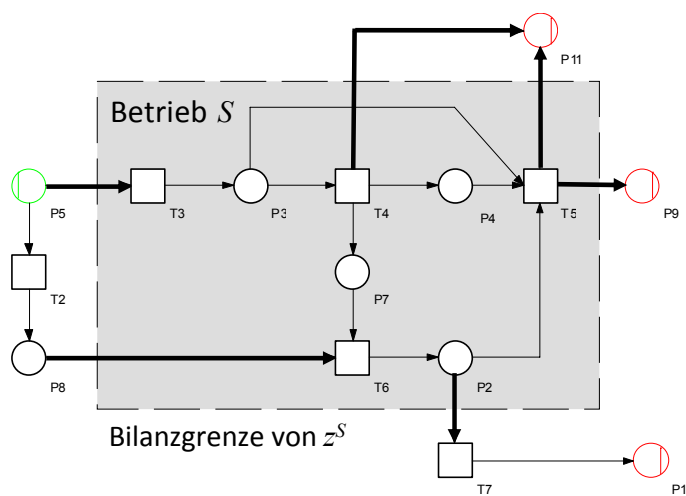


Abbildung 3.7: Veranschaulichung der betrieblichen Stoffbilanz. Die Transitionen T3-T6 bilden innerbetriebliche Stoffumwandlungsprozesse ab. Die fett gedruckten Flussrelationen bilden die betriebliche Stoffbilanz.

Das Grundmodell: exakte Abbildung einer Produktion ohne interne Bestandsveränderungen

Um die Darstellung der stoffstromnetzbasierter Zielgrößendefinition so einfach wie möglich zu halten, wird im Folgenden zunächst von Stoffstromnetzen ausgegangen, die (i) die betriebliche Produktion *exakt* abbilden und dabei (ii) *keine Nettoveränderung innerbetrieblicher Stoffbestände* im Betrachtungszeitraum berücksichtigen. Mit diesem *Grundmodell* lassen sich bereits die meisten praktisch relevanten Fälle abbilden. In § 3.3.5 wird dann die Verallgemeinerung auf beliebige Stoffstromnetze vorgenommen.

Keine Nettoveränderung betriebsinterner Stoffbestände

Die Vernachlässigung interner Bestandsveränderung bedeutet formal, dass alle innerbetrieblichen Transitionen des Stoffstromnetzes durch *Verbindungsstellen* gekoppelt sind (P2-P4 in Abbildung 3.7). Wenn man Verbindungsstellen als striktes Verbot von Bestandsveränderungen zwischen den durch sie verbundenen Prozessen versteht, könnten mit derartigen Stoffstromnetzen nur *kontinuierliche Produktionssysteme* mit perfekt aufeinander abgestimmten Prozessschrit-

ten abgebildet werden. Man muss jedoch berücksichtigen, dass die Stoffumsätze eines Produktionssystems mit Stoffstromnetzen immer bezüglich eines definierten Zeitintervalls abgebildet werden. Eine Verbindungsstelle besagt daher nicht, dass zu keinem Zeitpunkt Bestandsveränderungen auftreten können, sondern dass das *Integral* aller Bestandsveränderungen über den gewählten Betrachtungszeitraum verschwindet. In diesem Sinne können auch viele *Batchprozesse* oder *diskrete Fertigungsprozesse* mit kleineren Zwischenlagern mit dem oben definierten Grundmodell abgebildet werden.

Ohne interne Bestandsveränderungen entspricht die betriebliche Stoffbilanz z^S der *Nettoerzeugung*³² von Stoffen durch das Produktionssystem. Alle darüber hinaus erzeugten Zwischenprodukte werden innerhalb des Betrachtungszeitraums auch wieder verbraucht und können daher weder die natürliche Umwelt beeinflussen noch zu einer dauerhaften Wertschöpfung des Systems beitragen. Die ökologische und ökonomische Bewertung der Produktionsaktivität hängt daher ausschließlich von den Stoffströmen der betrieblichen Stoffbilanz ab.

Exakte Abbildung der betrieblichen Produktion durch das Stoffstromnetz

Bei einer exakten Abbildung eines Betriebs durch ein Stoffstromnetz entsprechen die Modellgrenzen den Betriebsgrenzen. Alle Grenzstellen befinden sich gerade außerhalb des Betriebes, entweder in der Ökosphäre oder im Verfügungsbereich anderer Wirtschaftssubjekte. Die entsprechenden Stoffströme zwischen Betrieb und diesen Grenzstellen werden als *Elementarströme* und *Warenströme* bezeichnet (Abbildung 3.8).

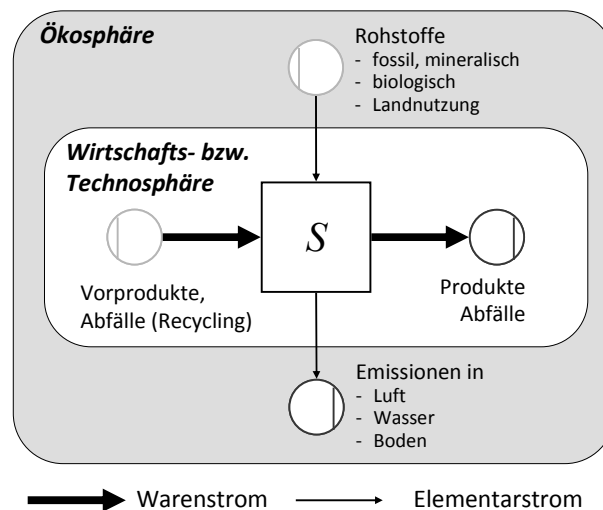


Abbildung 3.8: Warenströme und Elementarströme in der betrieblichen Stoffbilanz.

Die *direkten Umweltwirkungen* der betrieblichen Produktion gehen von den Elementarströmen aus. Dabei handelt es sich um die Nutzung natürlicher Rohstoffe sowie die Emission von Schad-

³² Komponenten mit einem negativen Vorzeichen stellen einen *Nettoverbrauch* dar.

stoffen in die Umwelt. Warenströme ergeben sich dagegen aus *wirtschaftlichen Transaktionen* des Betriebs mit anderen Wirtschaftssubjekten wie Kunden, Zulieferern, Energieversorgern oder Abfallentsorgern³³.

Zweistufigkeit der stoffstromnetzbasieren Zielgrößendefinition

Die stoffstromnetzbasierete Zielgrößendefinition erfolgt *zweistufig*: Zunächst wird mit dem Stoffstromnetz die Stoffbilanz als Funktion einiger vorgegebener Stoffströme, -bestände und Parameter berechnet: $\mathbf{z}^S = \mathbf{z}^S(\mathbf{z}, \mathbf{s}, \mathbf{p})$. Darauf aufbauend werden dann ökonomische und ökologische Zielgrößen definiert: $\varphi = \varphi(\mathbf{z}^S)$. Interne Stoffströme gehen in der Regel nicht *explizit* in die Definition von Zielgrößen ein. Um beispielsweise die Umweltwirkung der Einführung eines innerbetrieblichen Recyclings zu bewerten, muss nicht die Umstrukturierung der innerbetrieblichen Stoffströme selbst betrachtet werden, sondern deren Wirkung auf die betriebliche Stoffbilanz. So reduziert das Recycling das zu entsorgende Abfallaufkommen und verringert darüber hinaus den Primärrohstoffbedarf.

3.3.2 Die Präferenzen des Entscheidungsträgers

Vom materiellen Standpunkt der Stoffstromanalyse aus betrachtet ist Produktion die *Kombination* bzw. *Transformation* von Stoffen, wobei bestimmte Ausgangsstoffe (*Inputs*) verbraucht und Produkte (*Outputs*) erzeugt werden. Für die Optimalplanung eines Produktionssystems ist es jedoch nicht ausschlaggebend, ob Stoffe entstehen oder vernichtet werden, sondern wie diese Vorgänge vom Entscheidungsträger *bewertet* werden.

3.3.2.1 Übergang von der Stoffstrombilanz zur Ergebnisbilanz

Differenzierung von Stoffen in Güter, Übel und Neutra

Diese Bewertung hängt wiederum davon ab, wie der Entscheidungsträger das Vorhandensein der entsprechenden Stoffe beurteilt. Dyckhoff (1991: S.291) unterscheidet in diesem Zusammenhang drei *Erwünschtheits-Kategorien*, denen sämtliche bei einer Produktionsaktivität involvierten Objekte zugeordnet werden können:

- *Güter* sind erwünschte Stoffe, die als Produktionsfaktoren eingesetzt oder an Kunden verkauft werden können. Sie besitzen aus Sicht des Produzenten einen positiven Gebrauchs- oder Tauschwert.
- *Übel* sind unerwünschte Stoffe, weil sie für Mensch und Natur schädlich sind oder als störend (z.B. übel riechend) empfunden werden.

³³ d.h. Abfälle werden als „Waren“ betrachtet, solange sie Gegenstand von Transaktionen innerhalb der Wirtschaftssphäre sind. Erst bei ihrer Endlagerung auf einer Mülldeponie werden sie zu Elementarströmen.

- *Neutra* sind Stoffe, von denen weder ein Nutzen für den Produzenten noch eine Gefahr für die Umwelt ausgeht, weshalb der Produzent ihrem Vorhandensein gegenüber indifferent ist.

Diese Klassifizierung kombiniert ökonomische und ökologische Gesichtspunkte. Verkürzt gesagt ist ein *Gut* für Produktion oder Konsum nützlich, ein *Übel* für Mensch oder Umwelt schädlich. In der Klasse der „*Neutra*“ verbleiben lediglich die wert- und harmlosen Stoffe. Die Klassifizierung erfolgt – der in den meisten Unternehmen üblichen Prioritätensetzung entsprechend – zweistufig (Tabelle 3.2). Zunächst werden unter ökonomischen Gesichtspunkten alle nützlichen Stoffe als Güter identifiziert. Bei rationalem Verhalten können nur nutzlose Stoffe, für die keine weitere Verwendung innerhalb der Wirtschaftssphäre besteht, in die natürliche Umwelt gelangen. Sie werden je nachdem, ob sie umweltschädlich oder harmlos sind, den Übeln oder den *Neutra* zugeordnet.

Tabelle 3.2: Zweistufige Klassifizierung aller Stoffe in Güter, Übel und *Neutra*

Bewertung		
ökonomisch	ökologisch	insgesamt
nützlich	-	Güter
nutzlos	schädlich	Übel
	harmlos	<i>Neutra</i>

Ein „ökologisches Übel“ kann im Sinne von Opportunitätskosten auch zu einem „ökonomischen Übel“ werden. Wird ein Produzent durch entsprechende Umweltgesetze z.B. verpflichtet, einen bestimmten Produktionsabfall zu entsorgen, muss er Ressourcen dafür aufwenden, die dann nicht mehr für die eigentliche Produktion zur Verfügung stehen. Alternativ kann er auch einen Dritten mit der Entsorgung beauftragen. Die Produktionsabfälle tauchen dann als Ware in der betrieblichen Stoffbilanz auf, für die jedoch kein Geld eingenommen, sondern ausgegeben wird.

Erträge und Aufwendungen der Produktion

Aufgrund der Einteilung der Stoffe in Güter, Übel und *Neutra* können nun auch die Stoffströme der betrieblichen Stoffbilanz danach beurteilt werden, ob sie den Nutzen des Entscheidungsträgers steigern oder mindern. So sind die Erzeugung von Gütern und die Vernichtung von Übeln erstrebenswerte Ergebnisse der Produktion. Sie werden als *Erträge* bezeichnet. Umgekehrt stellen der Verbrauch von Gütern und die Erzeugung von Übeln *Aufwendungen* dar. Das Verhältnis zwischen den Kategorien Input/Output und Aufwand/Ertrag wird in Abbildung 3.9 am Beispiel einer Anlage zur thermischen Nachverbrennung illustriert. Diese verbrennt umweltschädliche leichtflüchtige organische Verbindungen (Übel) bei Bedarf unter Zugabe von Erdgas (Gut) zu weniger schädlichen Abgasen (Übeln) und erzeugt dabei thermische Energie (Gut).

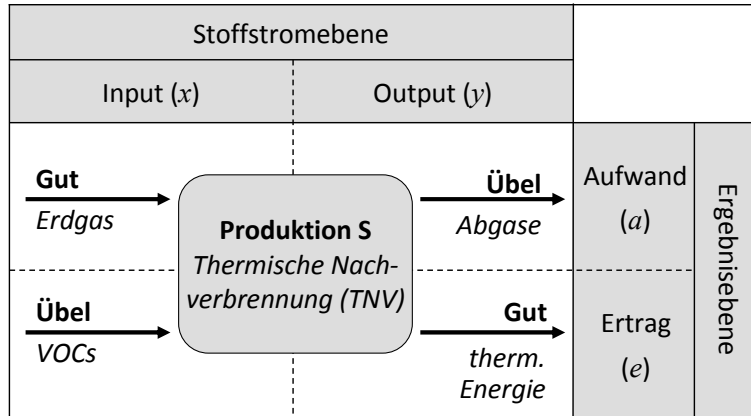


Abbildung 3.9: Stoffstrom- und Ergebnisebene zur Beschreibung der mit dem Produktionssystem S verbundenen Stoffströme. Als Beispiel dient eine Einheit zur thermischen Nachverbrennung (TNV) leichtflüchtiger organischer Verbindungen (VOC) wie sie in der chemischen Industrie und im Druckgewerbe vielfach eingesetzt wird.

Die Ergebnisfunktion zur Transformation der Stoffstrom- in die Ergebnisdarstellung

Formal wird der Übergang von der objektiven Stoffstromebene auf die bewertete Ergebnisebene durch die *Ergebnisfunktion* $P(\cdot)$ beschrieben. Diese ordnet jedem Stoffstromvektor \mathbf{z} einen Ergebnisvektor \mathbf{l} zu: $\mathbf{l} = P(\mathbf{z})$ (Dyckhoff, 1994: S.62ff). Die Ergebnisfunktion ist durch die komponentenweise Vorzeichentransformation (3.12) gegeben, welche die Präferenzen des Entscheidungsträgers bezüglich aller in der Produktion involvierten Stoffe abbildet.

$$l_k = \delta_k \cdot z_k$$

wobei $l, z \in R^k$ und $\delta_k = \begin{cases} +1 & k \text{ ist ein Gut (erwünscht)} \\ 0 & k \text{ ist Neutrum (unbeachtet)} \\ -1 & k \text{ ist Übel (unerwünscht)} \end{cases}$ (3.12)

Die Ergebnisbilanz $\mathbf{l}^S = P(\mathbf{z}^S)$ ist somit bis auf die Vorzeichen ihrer Komponenten mit der Stoffbilanz identisch. Insbesondere werden die Produktionsergebnisse nach wie vor in physikalischen Einheiten gemessen. Dyckhoff (1998: S.124) spricht daher in Abgrenzung zum monetären Aufwands- bzw. Ertragsbegriff des externen Rechnungswesens von *mengenmäßigen* bzw. *realen* Erträgen und Aufwendungen der Produktion.

Identifikation effizienter Systemzustände durch die Dominanzrelation

Durch den Übergang von der Stoffstrom- zur Ergebnisebene kann mit der nachfolgend definierten *Dominanzrelation* ein einfaches Kriterium für den Vergleich zweier Systemzustände hinsichtlich ihrer Güte aus Sicht des Entscheidungsträgers formuliert werden.

Definition 3.1: Gilt im komponentenweisen Vergleich der Ergebnisvektoren für zwei Zustände Z_1 und Z_2 des Produktionssystems

$$l_k^{Z_1} \geq l_k^{Z_2} \quad \forall k \in M \tag{3.13}$$

so sagt man, dass der Zustand Z_1 den Zustand Z_2 **dominiert** (Dyckhoff, 1998: S.136). Weil in Z_1 alle Erträge mindestens gleich groß und alle Aufwendungen mindestens ebenso gering sind wie in Z_2 , wird der Zustand Z_1 vom Entscheidungsträger gegenüber Z_2 vorgezogen.

Für die Optimierung von Produktionssystemen ist die Dominanzrelation insofern unzureichend, als sie nur eine *Halbordnung* im Ergebnisraum definiert, d.h. nur für Zustandspaare definiert ist, bei denen der komponentenweise Vergleich der Ergebnisvektoren einheitlich ausfällt. Diese Problematik wird in Abbildung 3.10 für ein einfaches Produktionssystem mit nur zwei Stoffströmen veranschaulicht, von denen der erste einen Ertrag (e_1) und der zweite einen Aufwand (a_2) darstellt. Durch den komponentenweisen Vergleich wird der Produktionsergebnisraum ausgehend vom Ausgangszustand Z_0 in drei verschiedene Bereiche aufgeteilt. Alle Produktionsergebnisse, die im Bereich B liegen, dominieren den Ausgangszustand Z_0 . So auch der Zustand Z_1 , der auf dem Rand von B liegt. Z_1 hat den gleichen Aufwand wie Z_0 , liefert jedoch einen höheren Ertrag. Umgekehrt dominiert Z_0 alle Zustände, die im Bereich A liegen und damit insbesondere den Zustand Z_2 . Dagegen kann für alle Zustände, die im Bereich C liegen, auf der Basis des komponentenweisen Vergleichs grundsätzlich nicht entschieden werden, ob sie dem Zustand Z_0 vorzuziehen sind oder nicht. Z_3 weist bspw. einen geringeren Aufwand als Z_0 auf ($l_2^{(Z_3)} > l_2^{(Z_0)}$). Dafür ist jedoch auch der Ertrag geringer ($l_1^{(Z_3)} < l_1^{(Z_0)}$). Die Ergebnisse des komponentenweisen Vergleichs sind somit widersprüchlich. Hinsichtlich des Ertrags schneidet Z_0 besser ab, hinsichtlich des Aufwands jedoch Z_3 .

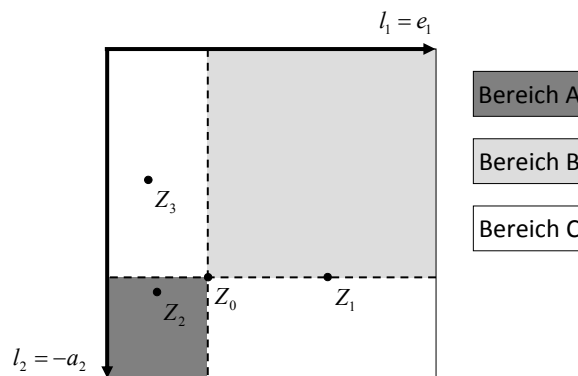


Abbildung 3.10: Vergleich verschiedener Ergebnisse eines Produktionssystems S mit zwei Stoffströmen, von denen einer einen Ertrag (e_1) und einer einen Aufwand (a_2) darstellt.

Mit der Dominanzrelation kann nun ein Kriterium für *effiziente Zustände* des Produktionssystems formuliert werden.

Definition 3.2: Alle Zustände Z^{eff} eines Produktionssystems, deren Ergebnis von keinem anderen zulässigen Zustand $Z \in F$ dominiert wird, heißen *effizient*.

Die Eigenschaft der Effizienz definiert innerhalb des Modellgültigkeitsbereichs F^S die so genannte *Pareto-Menge*. Für alle effizienten Zustände gilt: Die Verbesserung eines Ergebnisses kann – unter Berücksichtigung der Restriktionen, die F^S definieren – nur auf Kosten der Verschlechterung eines anderen Ergebnisses erreicht werden. Weil die einzelnen Erträge und Aufwendungen der Produktion untereinander jedoch nicht vergleichbar sind, kann grundsätzlich nicht beurteilt werden, welcher unter den effizienten Zuständen der beste ist. Effizienz ist ein notwendiges, aber kein hinreichendes Kriterium für die Bestimmung eines eindeutigen Optimums $Z^* \in F^S$.

Das Minimal- und das Maximalprinzip als fundamentale Vektoroptimierungsprobleme der Ergebnisebene

Betriebliche Produktion unterliegt – wie jedes auf Zwecke hin gerichtete menschliche Handeln – dem *allgemeinen Vernunftsprinzip*, wonach es erstrebenswert ist, ein gegebenes Ziel mit möglichst geringem Mitteleinsatz zu erreichen (Wöhe und Döring, 2000: S.1). Dieses *Rationalprinzip* kann auf die Ebene realer Produktionsergebnisse übertragen werden, indem der Zweck mit den Erträgen e^S und die Mittel mit den Aufwendungen a^S der Produktion identifiziert werden. Dann ergeben sich die beiden in Tabelle 3.3 dargestellten Optimierungsprobleme, die den beiden Ausprägungen des Rationalprinzips als Minimal- und als Maximalprinzip entsprechen.

In der Praxis ist vor allem das *Minimalprinzip* von Bedeutung. In vielen Fällen werden Produktionssysteme geplant, die nur einen Ausschnitt aus einem größeren Produktionssystem bzw. aus der Gesamtheit wirtschaftlicher Aktivitäten eines Unternehmens darstellen. Sie sind daher typischerweise in eine unternehmensinterne Planungshierarchie eingebunden, durch die das Produktionsvolumen (der Ertrag) auf der taktischen bzw. strategischen Planungsebene, bspw. durch vertragliche Lieferverpflichtungen, bereits festgelegt worden ist.

Tabelle 3.3: Das Rationalprinzip übertragen auf die Ergebnisdarstellung der Produktion. Aufwand und Ertrag sind Vektoren, weshalb die Minimierung und Maximierung komponentenweise zu verstehen ist.

Ausprägung	Formale Schreibweise	Wortlaut
Minimalprinzip	$\min \mathbf{a}$ Nb. $\mathbf{e} = \text{const.}$	minimiere den Aufwand bei gegebenem Ertrag
Maximalprinzip	$\max \mathbf{e}$ Nb. $\mathbf{a} = \text{const.}$	maximiere den Ertrag bei gegebenem Aufwand

Das Ziel der Optimalplanung besteht dann darin, eine Systemkonfiguration zu ermitteln, mit der diese Vorgabe aufwandsminimal hergestellt werden kann. Die umgekehrte Zielsetzung, nämlich eine festgelegte Rohstoffmenge ertragsmaximierend zu verarbeiten, kommt in der betrieblichen Praxis dagegen so gut wie nie vor (vgl. Müller-Christ, 2001: S.536).

3.3.2.2 Aggregation des Produktionsergebnisses zu einer skalaren Zielfunktion

Das auf die Ergebnisebene übertragene Minimal- und Maximalprinzip (Tabelle 3.3) führt zu *Vektoroptimierungsproblemen*. Eindeutige Optima können jedoch nur auf der Grundlage einer skalaren Zielfunktion bestimmt werden. Daher müssen die Produktionsergebnisse mit einer der in § 2.2.4 dargestellten entscheidungstheoretischen Methoden zu einer skalaren Zielfunktion aggregiert werden. Das setzt jedoch voraus, dass der Entscheidungsträger seine Präferenzen bezüglich der einzelnen Produktionsergebnisse weiter präzisiert.

Dominiert ein bestimmter Ertrag e_i alle anderen Produktionsergebnisse in seiner Bedeutung für den Entscheidungsträger, so bietet es sich an, dem Prinzip der Zieldominanz entsprechend e_i als Zielfunktion zu wählen. Ein etwas differenzierteres Ergebnis kann unter Umständen mit dem Ansatz der lexikographischen Ordnung erzielt werden, weil Nebenerträge bei der Optimierung zumindest nicht prinzipiell ausgeblendet werden. Die Anwendung der lexikographischen Ordnung ist insbesondere bei der Minimierung produktionsbedingter Umweltbelastungen denkbar. Im Kontext der aktuellen Umweltdebatte rangiert die Klimaproblematik in der öffentlichen Wahrnehmung deutlich vor anderen Umweltproblemen. Daher kann es sinnvoll sein, ein Produktionssystem vorrangig im Hinblick auf die Treibhausgasemissionen („Carbon Footprint“) zu optimieren und erst bei verbleibenden Freiheitsgraden weitere Wirkungskategorien wie die Versauerung, Eutrophierung oder Ökotoxizität in die Optimierung einzubeziehen.

Im Idealfall kann der Beitrag aller Produktionsergebnisse zu einer skalaren Zielgröße quantifiziert und das Prinzip der Zielgewichtung (2.6, S.22) angewendet werden. Die Zielfunktion $g(\mathbf{l})$ ergibt sich als Skalarprodukt aus einem *Gewichtvektor* \mathbf{g} mit dem Ergebnisvektor \mathbf{l} . Damit $g(\mathbf{l})$ wohldefiniert ist, müssen sowohl der Gewichtvektor also auch der Ergebnisvektor die $(\kappa \times l)$ -Vektoren sein. Außerdem müssen die Komponenten des Gewichtungsvektors *spezifische Größen* sein, deren Dimension sich als Quotient aus der Maßeinheit der Zielfunktion und der des jeweiligen Produktionsergebnisses l_k ergibt (3.14).

$$[g_k] = \frac{[g(l)]}{[l_k]} \quad (3.14)$$

3.3.3 Ökonomische Zielgrößen

Geeignete Maße für die ökonomische Performance eines Produktionssystems

Ökonomische Zielgrößen müssen den *Nutzen* quantifizieren, der sich aus einer Produktionsaktivität ergibt. Im Sinne des Gebrauchswertes kann dieser Nutzen bereits an den realen Erträgen eines Produktionssystems festgemacht werden. In der Regel wird er jedoch anhand einer *monetären Bewertung* $m(\cdot)$ der Ergebnisbilanz \mathbf{I}^S quantifiziert. Da der Nutzen betrieblicher Produktion in einem marktwirtschaftlichen Umfeld primär in der Steigerung des betrieblichen Geldvermögens

besteht, liefert diese offensichtlich ein sinnvolles Maß für die ökonomische Performance eines Produktionssystems.

Auf der methodischen Ebene entspricht die monetäre Bewertung einer Aggregation des Produktionsergebnisses durch eine Zielgewichtung (3.15). Die Gewichte p_k sind durch die *spezifischen Preise* gegeben, die für den Warentausch mit anderen Wirtschaftssubjekten gelten. Außer Marktpreisen im eigentlichen Sinn können die p_k auch andere *preisähnliche* Bestandteile wie *Entsorgungsgebühren*, *Abgaben* oder *Steuern* beinhalten. Daher können sie sich auch auf die Elementarströme der betrieblichen Stoffbilanz beziehen. Beispielsweise in Form von Preisen für CO₂-Zertifikate oder Grundwasserentnahmegebühren („Wasserpfennig“).

$$m(\mathbf{l}^S) = \sum_{k=1}^K p_k \cdot l_k = \sum_i p_i e_i - \sum_j p_j a_j = E(\mathbf{e}) - K(\mathbf{a}) = G(\mathbf{l}^S) \quad (3.15)$$

Umsatzerlöse und Produktionskosten

Aufgrund der Vorzeichenkonvention für die Ergebnisbilanz zerfällt die gewichtete Summe (3.15) in eine Differenz zweier Summenterme. Die positiven Erträge können zu den *stoffstrombasierten Umsatzerlösen* $E = m(\mathbf{e})$ zusammengefasst werden und die negativen Aufwendungen zu den *stoffstrombezogenen Produktionskosten* $K = m(\mathbf{a})$. Bei materieller Güterproduktion werden die Umsatzerlöse *vollständig* durch entsprechende Stoffströme erfasst. Die stoffstrombasierten Kosten decken dagegen nur Material- und Energiekosten ab. Diese stellen zwar den größten Kostenblock im produzierenden Gewerbe dar (Statistisches Bundesamt, 2010: S.277), andere wichtige Kostenfaktoren wie Mieten, Investitionen oder Personal haben jedoch keine unmittelbaren Entsprechungen auf der Stoffstromebene.

Betriebserfolg und Deckungsbeitrag

In der Betriebswirtschaftslehre werden unterschiedliche Indikatoren für die betriebliche Wertschöpfung verwendet. Die Differenz zwischen dem monetär bewerteten Ertrag einer Produktion und dem monetär bewerteten Einsatz an Produktionsfaktoren wird als *Betriebserfolg* bezeichnet (Wöhe und Döring, 2000: S.46ff.; 982ff.). Werden in der Differenz nur die *variablen* Produktionskosten betrachtet, spricht man dagegen vom *Deckungsbeitrag*. Dieser quantifiziert die Geldmenge, die aus der Produktion zur Deckung der fixen Kostenbestandteile durch Potenzialfaktoren wie Grundstücke, Fabrikgebäude, Maschinen oder auch Arbeitskräfte zur Verfügung steht.

Ob $G(\mathbf{l}^S)$ eher dem Betriebserfolg oder dem Deckungsbeitrag entspricht, hängt wesentlich von der Modellierung des Stoffstromnetzes ab. Bei der Abbildung von Produktionssystemen als Stoff-

stromnetz ist es *üblich*, nur die für die laufende Produktion nötigen Aufwendungen³⁴ als Stoffströme zu modellieren. Potenzialfaktoren werden auf der Stoffstromebene dagegen nicht oder nur teilweise berücksichtigt. $G(I^S)$ wird hier daher als *stoffstrombezogener Deckungsbeitrag* bezeichnet. Das heißt jedoch nicht, dass sich der Betriebserfolg mit Stoffstromnetzen nicht quantifizieren ließe. Einerseits ist es durchaus möglich, zumindest die materiellen Potenzialfaktoren, als Stoffströme in das Modell zu integrieren. Andererseits werden fixe Kostenbestandteile an den Transitionen bei der stoffstromnetzbasierten Kostenrechnung (vgl. § 3.1.3) standardmäßig berücksichtigt.

Produktivität und Wirtschaftlichkeit

Die *Produktivität* ρ misst das Verhältnis von realem Ertrag zum realen Aufwand. Beides sind jedoch vektorielle Größen, die Objekte unterschiedlicher Art zusammenfassen: verschiedene im Produktionsprozess eingesetzte Roh-, Hilfs- und Betriebsstoffe, Arbeitseinsatz des Personals, Schadstoffemissionen, hergestellte Produkte usw. Dieses Problem wird regelmäßig dadurch umgangen, dass *Faktorproduktivitäten* ρ^f betrachtet werden. Insbesondere bei Prozessen mit einem einzigen Ertrag bzw. einem Hauptertrag wird dieser zu einzelnen Faktoraufwendungen ins Verhältnis gesetzt (3.16). So können z.B. Arbeitsproduktivitäten, Maschinenproduktivitäten oder eben auch Rohstoffproduktivitäten berechnet werden (Kiener et al., 2006: S.18f.).

$$\rho = \frac{e_i}{a_j} \quad (3.16)$$

Eine andere Methode, um das Einheitenproblem bei der Quantifizierung der Gesamtproduktivität einer Produktion zu umgehen, besteht im Übergang von den realen zu monetär bewerteten Größen. Das Verhältnis (3.17) von monetär bewertetem Ertrag (Erlös) zu monetär bewertetem Aufwand (Kosten) wird dann nicht mehr als Produktivität, sondern als *Effizienz* bzw. *Wirtschaftlichkeit* bezeichnet (Wöhe und Döring, 2000: S.48). In diesem Sinne werden auch bei der Quantifizierung von *Material-* und *Energieeffizienz* monetäre Größen verwendet. Es wird dann jeweils der Ertrag ins Verhältnis zum gesamten Material- bzw. Energieaufwand gesetzt.

$$\rho = \frac{E(e)}{K(a)} \quad (3.17)$$

³⁴ Diese werden in der Produktionstheorie als Verbrauchsfaktoren bezeichnet. Dieser Begriff beschränkt sich jedoch auf den Güterverbrauch, während die Aufwendungen auch die Erzeugung von Übeln mit einschließen (vgl. §3.3.2).

3.3.4 Ökologische Zielgrößen

Für die *ökologische Bewertung* $u(\cdot)$ eines Produktionssystems eignen sich grundsätzlich die Methoden der Ökobilanzierung (ISO 14040, 2006). Die Ökobilanzierung basiert einerseits auf der Betrachtung ganzer Produkt/*lebenswege* „von der Wiege bis zur Bahre“. Ein solcher Lebensweg beginnt mit der Rohstoffgewinnung, führt über diverse Produktionsstufen, die Nutzung bis hin zum Recycling bzw. der endgültigen Entsorgung aller bei Herstellung und Nutzung angefallenen Abfälle. Dieser Aspekt kommt besser in der englischen Bezeichnung *Life Cycle Assessment (LCA)* zum Ausdruck. Andererseits werden Umweltbelastungen bei der Ökobilanzierung *medienübergreifend* erfasst und bewertet. Durch diese – im doppelten Sinne – ganzheitlichen Analyseprinzipien wird gewährleistet, dass Verlagerungen von Umweltbelastungen zwischen Lebenszyklusphasen oder Umweltmedien nicht fälschlicherweise als ökologische Verbesserungen erscheinen. Für die Bestimmung ökologischer Gewichte für die Stoff- bzw. Ergebnisbilanz ergeben sich daraus zwei wesentliche Konsequenzen, die im Folgenden näher dargestellt werden.

Lebenszyklusanalyse: von der Kern- zur Komplementärbilanz

Die betriebliche Stoffbilanz erfasst nur die *direkte Umweltwirkung* („*gate-to-gate*“) eines Betriebes. Der überwiegende Teil der Stoffströme einer betrieblichen Stoffbilanz sind jedoch Warenströme, die den Betrieb mit anderen Wirtschaftssubjekten wie Zulieferbetrieben, Energieversorgern, Entsorgungsunternehmen oder Kunden verbinden. Diese haben zwar keinen direkten Einfluss auf die Umwelt, wohl aber die mit ihnen verbundenen vor- und nachgelagerten Prozesse innerhalb der Wirtschaftssphäre, so z.B. die Erzeugung der für die Produktion benötigten Energie oder die Entsorgung von Produktionsabfällen.

Um diese *indirekten* Umweltwirkungen im Sinne einer Ökobilanz zu erfassen, muss die betriebliche Stoffbilanz als *funktionelle Einheit* betrachtet werden und durch eine *Sachbilanzierung* die mit den vor- und nachgelagerten Produktions- bzw. Stoffumwandlungsprozessen verbundenen Elementarflüsse quantifiziert werden. Erst durch die Erweiterung der betrieblichen *Kernbilanz* z^S um eine *Komplementärbilanz* z^K (Abbildung 3.11) ist es möglich, aussagekräftige Kennzahlen für die produktionsbedingte Umweltwirkung zu berechnen (Braunschweig und Müller-Wenk, 1993: S.54). Werden in der Komplementärbilanz nur die Aufwendungen des Produktionssystems berücksichtigt, spricht man von einer *cradle-to-gate*-Bilanzierung: $u = u(\mathbf{a})$. Werden dagegen auch die Nutzung und Entsorgung der hergestellten Produkte in die Analyse einbezogen, liegt eine *cradle-to-grave*-Bilanzierung, d.h. eine Lebenszyklusanalyse im eigentlichen Sinne vor: $u = u(\mathbf{a}) + u(\mathbf{e})$.

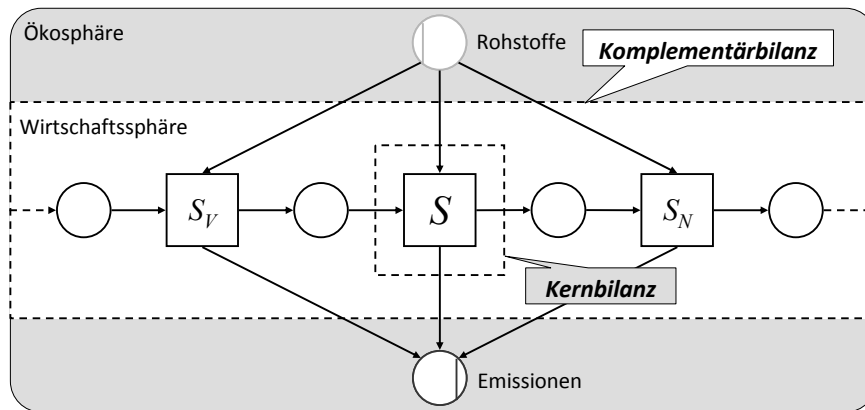


Abbildung 3.11: Erweiterung der betrieblichen Kernbilanz zur Komplementärbilanz, die außer dem Betrieb selbst (S) auch die für die betriebliche Produktion notwendigen vorgelagerten (S_V) und nachgelagerten (S_N) Prozesse innerhalb der Wirtschaftssphäre berücksichtigt.

Die Umweltwirkung als Vektor verschiedener Wirkungsindikatoren

Aufbauend auf der Komplementärbilanz kann eine *Wirkungsabschätzung* durchgeführt werden, bei der ausgehend von den erfassten produktionsbedingten Elementarströmen die Wirkung der Produktion in verschiedenen Umweltproblembereichen quantifiziert wird. Daher ist die ökologische Bewertung $\mathbf{u}(\cdot)$ im Gegensatz zur monetären Bewertung eine *vektorielle* Funktion, deren Komponenten die produktionsbedingten Umweltbelastungen in den verschiedenen Wirkungskategorien abbilden. Die Wirkungsabschätzung erfolgt zweistufig. Bei der so genannten *Klassifizierung* werden zunächst alle Elementarströme den verschiedenen *Wirkungskategorien* zugeordnet: Klimawandel, Versauerung, Ressourcenverbrauch, Eutrophierung, Human- bzw. Ökotoxizität, Naturraumbeanspruchung etc. Bei der anschließenden *Charakterisierung* wird die Wirkung verschiedener Substanzen in derselben Wirkungskategorie anhand einheitlicher Indikatoren quantifiziert und damit vergleichbar gemacht (Tabelle 3.4).

Tabelle 3.4: Beispiele für Wirkungskategorien, -indikatoren und die entsprechenden Maßeinheiten aus Klöpffer und Grahl (2009: S.254ff)

Wirkungskategorie	Wirkungsindikator	Maßeinheit
Klimawandel	Treibhauspotenzial (global warming potential GWP)	kg CO ₂ -Äquivalente
Ozonabbau	Ozone depletion potential ODP	kg CFC11-Äquivalente
Sommersmog	Smogbildungspotenzial (photochemical ozone creation potential POCP)	kg Ethen-Äquivalente
-	Kumulierter Energie-Aufwand (KEA)	J

Ein Indikator, der in der Praxis gerne für die ökologische Bewertung herangezogen wird, ist der *kumulierte Energie-Aufwand* (KEA) (Klöpffer und Grahl, 2009: S.78). Zu seiner Berechnung werden die Energieinhalte aller Primärenergieträger der Komplementärbilanz summiert. Der KEA kann nicht unmittelbar einer bestimmten Wirkungskategorie zugeordnet werden, liefert jedoch einen Anhaltspunkt für die Energieintensität der Produktion.

3.3.5 Die Bewertung beliebiger Stoffstromnetze

3.3.5.1 Produktion mit internen Bestandsveränderungen

Spielen bei einer stoffstrombasierten Optimalplanung innerbetriebliche Bestandsveränderungen eine wichtige Rolle, müssen "normale" Stellen zur Abbildung der entsprechenden Lager verwendet werden. Dies hat zur Folge, dass die betriebliche Stoffbilanz die Nettoproduktion eines Betriebes nicht mehr vollständig erfasst. Nachfolgend wird zunächst gezeigt, welche Auswirkungen dies jeweils für die ökonomische und ökologische Bewertung betrieblicher Produktion hat. Da die Bewertung interner Bestandsveränderungen stets mit gewissen Schwierigkeiten verbunden ist, wird im Anschluss gezeigt, wie diese durch entsprechende Modifikationen des Modells umgangen werden kann.

Auswirkung auf die ökologische und die ökonomische Bewertung

Aus *ökologischer Sicht* sind betriebsinterne Bestände irrelevant, da Umweltwirkungen nur von Bestandsveränderungen in der Ökosphäre ausgehen können. Grundsätzlich wäre es jedoch denkbar, das *Umweltrisiko* zu bewerten, das im Schadensfall beispielsweise von einem Tank für giftige Substanzen ausgeht. In der vorliegenden Arbeit wird jedoch durchweg von einer deterministischen Modellierung ausgegangen.

Aus *ökonomischer Sicht*, d.h. im Sinne der betrieblichen Wertschöpfung, müssen interne Bestandsveränderungen dagegen durchaus berücksichtigt werden. Dafür gibt es zwei Möglichkeiten. Einerseits können interne Bestandsveränderungen genauso wie die Warenströme der betrieblichen Stoffbilanz behandelt werden. Bei einer Bestandszunahme wird dann angenommen, dass dieser Bestand potenziell verkauft werden kann bzw. – im Falle von Abfällen – kostenpflichtig entsorgt werden muss und entsprechend monetär bewertet. Formal wird die Zunahme interner Bestände also wie ein Outputstrom und die Abnahme wie ein Inputstrom behandelt. Bei Zwischenprodukten für die keine außerbetriebliche Nachfrage und daher auch kein Marktpreis existiert, bietet sich andererseits die Bewertung zu Herstellungskosten an, wie sie für das externe Rechnungswesen im Handelsgesetzbuch festgeschrieben ist (Wöhe und Döring, 2000: S.990).

Aus *ökonomischer Sicht* führt bereits das bloße Vorhandensein von Lagerbeständen, abgesehen von den ggf. notwendigen Aufwendungen der Herstellung des gelagerten Stoffes, zu so genannten *Kapitalbindungskosten*. Damit sind die Opportunitätskosten gemeint, die daraus resultieren, dass das für den Einkauf bzw. für die Herstellung des gelagerten Stoffes aufgewendete Kapital

nicht anderweitig am Kapitalmarkt eingesetzt werden kann, um damit Rendite zu erzielen. Die entgangene Rendite entspricht den Kapitalbindungskosten.

Einfluss des Betrachtungszeitraums auf die Bewertung interner Bestandsveränderungen

Streng genommen stellen interne Bestände einen *zukünftigen Nutzen* dar, der erst entsteht sobald sie verkauft oder in der betrieblichen Produktion eingesetzt werden. Insofern sollte die Modellierung von Bestandsveränderungen in einem einperiodigen Optimierungsmodell vermieden werden. Um Bestandsveränderungen des Realsystems zu berücksichtigen, könnte das Modell um weitere Perioden erweitert werden, in denen die Bestände explizit entweder verkauft, d.h. zu Outputströmen der betrieblichen Stoffbilanz werden, oder in der Produktion eingesetzt werden und damit zu einer Reduktion der Inputströme führen. Es wird dann nicht für die einzelne Periode, wohl aber für den gesamten Betrachtungszeitraum gefordert, dass keine Nettobestandsveränderungen stattfinden. Bei einer solchen Modellierweise erübrigt sich die explizite Bewertung von Beständen, weil Bestandsveränderungen einer Periode sich auf die betriebliche Stoffbilanz anderer Perioden auswirken und damit implizit bewertet werden. Eine Alternative zum mehrperiodigen Modell stellt die Ausdehnung des Betrachtungszeitraums des einperiodigen Modells dar, die so zu wählen ist, dass wiederum keine Nettoveränderungen interner Bestände auftreten.

3.3.5.2 Das Auseinanderfallen von Modell- und Betriebsgrenzen

In Abschnitt 3.3 wurde bislang davon ausgegangen, dass das zu bewertende Stoffstromnetz die betriebliche Produktion exakt abbildet. In der Praxis gibt es zwei Arten von Abweichungen gegenüber diesem Idealfall: Einerseits kann es Stoffstromnetze geben, die nur Ausschnitte der gesamten betrieblichen Produktion abbilden. Dieser Fall tritt z.B. auf, wenn in größeren Unternehmen einzelne Produktlinien, Anlagen oder Standorte optimiert werden sollen. Die Modellgrenzen spiegeln in solchen Fällen typischerweise den Zuständigkeitsbereich eines bestimmten Entscheidungsträgers wieder. Sie verlaufen *innerhalb* der Betriebsgrenzen, weshalb es in solchen Modellen auch innerbetriebliche Grenzstellen gibt. Wie bei der Modellierung interner Bestände müssen die Bestände in diesen innerbetrieblichen Grenzstellen entweder zu Herstellungskosten oder beispielsweise mit unternehmensinternen Verrechnungspreisen bewertet werden.

Andererseits kann es Stoffstromnetze geben, die auch außerbetriebliche Produktionsprozesse abbilden. Wie in §3.3.4 gezeigt wurde, kann das z.B. für die Aufstellung einer Komplementärbilanz für die ökologische Bewertung notwendig sein. Dann verlaufen die Modellgrenzen *außerhalb* der Betriebsgrenzen. In diesem Fall kann die betriebliche Stoffbilanz nicht mehr einfach an den Grenzstellen des Modells abgelesen werden, sondern sie muss explizit anhand der Stoffströme bestimmt werden, welche die im Inneren des Stoffstromnetzes verlaufenden Betriebsgrenzen schneiden. Schließlich können auch beide genannten Abweichungen gleichzeitig auftreten.

3.3.6 Grenzen der stoffstrombasierten Bewertung betrieblicher Produktion

In Abschnitt 3.3 wurde dargestellt, wie Produktionssysteme basierend auf ihrer Abbildung als Stoffstromnetz ökologisch und ökonomisch bewertet werden. Nicht alle ökonomischen und ökologischen Auswirkungen industrieller Produktion können auf der Grundlage einer Stoffstromanalyse nachvollzogen werden. Geht man nur von den Stoffströmen aus, so werden Personalkosten, Mieten, Abschreibungen für Maschinen etc. in der Regel nicht berücksichtigt. Für das betriebliche Stoffstrommanagement wurde in der Vergangenheit jedoch bereits ein stoffstromnetzbasierendes Kostenrechnungssystem entwickelt, mit dem solche fixen Kostenbestandteile an Transitionen berücksichtigt werden können (Möller, 2000). Für die ökologische Bewertung von Produktionssystemen ergeben sich aus dem Stoffstrombezug keine wesentlichen Einschränkungen. Umweltprobleme wie Lärmbelastungen, radioaktive Strahlung oder Flächenversiegelung, die nicht unmittelbar durch Stoffströme verursacht werden, spielen in den meisten Produktionsprozessen nur eine untergeordnete Rolle.

3.4 Systematik stoffstromnetzbasierter Optimierungsprobleme

Die im vorigen Abschnitt eingeführten ökonomischen und ökologischen Zielgrößen können auf unterschiedliche Weise zu stoffstrombasierten Optimierungsproblemen kombiniert werden. Sie können als *Extremierungsziele* in die Zielfunktion integriert werden oder als *Satisfizierungsziele* zusätzliche Restriktionen definieren. Im ersten Fall wird die Minimierung oder Maximierung der Zielgröße angestrebt, im zweiten deren Beschränkung durch explizite Unter- oder Obergrenzen.

Während die Anzahl der Satisfizierungsziele bei einem Optimierungsproblem grundsätzlich unbeschränkt ist, müssen mehrere Extremierungsziele für eine mathematische Optimierung zu einer einzigen skalaren Zielfunktion aggregiert werden. Dafür kommen dieselben entscheidungstheoretischen Methoden (§ 2.2.5) infrage, die in den vorangegangenen Abschnitten zur Herleitung skalarer ökologischer und ökonomischer Zielgrößen aus der vektoriellen betrieblichen Stoff- bzw. Ergebnisbilanz verwendet wurden.

Theoretisch können stoffstromnetzbasierende Optimierungen gleichermaßen unter betriebswirtschaftlichen wie ökologischen Gesichtspunkten durchgeführt werden. Auf der methodischen Ebene besteht keinerlei Unterschied zwischen beispielsweise der Maximierung des Deckungsbeitrags eines Betriebs oder der Minimierung seiner Treibhausgasemissionen. In der Praxis besteht jedoch eine ausgeprägte *Asymmetrie* zwischen dem ökonomischen und dem ökologischen Aspekt. Während betriebliche Produktion, die zu finanziellen Verlusten führt, in einem marktwirtschaftlichen System auf Dauer nicht bestehen kann, wirken sich Umweltbelastungen nicht unmittelbar nachteilig auf das Unternehmen selbst aus. Bei der stoffstrombasierten Optimalplanung im betrieblichen Kontext haben daher stets die für das Unternehmen existenziellen betriebswirtschaftlichen Aspekte Vorrang vor den ökologischen Gesichtspunkten.

3.4.1 Ökonomische Zielsetzungen

Eine zentrale Zielsetzung der operativen Produktionsplanung ist die Maximierung des *Deckungsbeitrags* (Kiener et al., 2006: S.17). In den Deckungsbeitrag fließen sowohl die Umsatzerlöse $E(e)$ als auch die Produktionskosten $K(a)$ ein (3.15). Durch das negative Vorzeichen der Produktionskosten werden das Minimal- und Maximalprinzip aus § 3.3.2 zu einem einheitlichen Maximierungsproblem kombiniert, bei dem *Absolutänderungen* beider Beiträge gleich gewichtet werden. Bestehen bei der Planung bereits verbindliche Vorgaben bezüglich des Produktionsprogramms, geht man bei der Optimalplanung typischerweise zur *Minimierung der Produktionskosten* über. Untergrenzen für einzelne Systemerträge müssen dann durch entsprechende Restriktionen bei der Formulierung des Optimierungsproblems berücksichtigt werden.

Andererseits kann es auch von Interesse sein, die *Effizienz der Produktion zu maximieren*. Eine Steigerung der Effizienz schafft Spielräume für Preissenkungen und führt damit zu Wettbewerbsvorteilen. Die Effizienz ergibt sich gemäß (3.17) als Quotient aus den Umsatzerlösen und den Produktionskosten. Wie bei der Deckungsbeitragsmaximierung werden bei der Effizienzmaximierung Minimal- und Maximalprinzip kombiniert. Die Minimierung der Produktionskosten wird in diesem Fall dadurch in ein Maximierungsproblem überführt, dass deren Kehrwert in die Zielfunktion eingeht. Das führt zwangsläufig auf nichtlineare Optimierungsprobleme. Durch die Verhältnissbildung werden bei der Effizienz *relative Änderungen* in den Erlösen und den Kosten gleich stark gewichtet.

Obwohl normalerweise die Erzielung von Umsatzerlösen im Vordergrund steht, kann es kurzfristig aus strategischen Gründen für ein Unternehmen vorteilhaft sein, reale Produkterträge zu maximieren, beispielsweise um seine Marktanteile in dem betreffenden Produktsegment auszubauen. Die sich unmittelbar aus dem Maximalprinzip ergebende *Ertragsmaximierung* stellt als Ziel betrieblicher Optimalplanung in der Praxis jedoch eine Ausnahme dar.

3.4.2 Ökologische Zielsetzung: die Minimierung der Umweltbelastung

Aus ökologischer Sicht ist das primäre Ziel einer Optimalplanung die *Minimierung der produktionsbedingten Umweltbelastung*. Dieses Minimierungsproblem ist im „cradle-to-gate“-Fall formal analog zur ökonomischen Kostenminimierung. Der einzige Unterschied besteht darin, dass die Aufwendungen nicht monetär, sondern mit ihrer jeweiligen Umweltwirkung bewertet werden. Hierbei ergibt sich jedoch die Schwierigkeit, dass die ökologische Bewertung durch eine ökobilanzielle Wirkungsabschätzung eine *vektorielle Funktion* ist, deren Komponenten die verschiedenen Wirkungsindikatoren sind. Um zu einer *skalaren Zielfunktion* für eine ökologische Optimierung zu gelangen, sind in der Praxis zwei Wege üblich.

In der ISO-Normierung der Ökobilanzierung ist als optionaler Bestandteil der Wirkungsabschätzung die so genannte *Gewichtung* vorgesehen. Bei diesem Schritt werden verschiedene Wirkungsindikatoren im Sinne einer Zielgewichtung miteinander verrechnet (3.18), wobei in der Norm ausdrücklich darauf hingewiesen wird, dass die numerischen Gewichtungsfaktoren γ_k auf

Werturteilen beruhen und daher nicht objektivierbar sind (ISO 14044, 2006: §4.4.3.4.1). Daher wird der Anwendungsbereich der Gewichtung vorsorglich stark eingegrenzt: „Die Gewichtung [...] darf nicht in Ökobilanz-Studien angewendet werden, die für die Verwendung in zur Veröffentlichung vorgesehenen vergleichenden Aussagen bestimmt sind“ (ISO 14044, 2006: §4.4.5).

$$u(\mathbf{l}) = \sum_k \gamma_k \cdot u_k(\mathbf{l}) \quad (3.18)$$

Eine weitere Möglichkeit besteht darin, nur die für den Entscheidungsträger wichtigste Wirkungskategorie zu betrachten: $u(\mathbf{l}) = u_k(\mathbf{l})$. Das genau ist der Ansatz des so genannten **Carbon Footprints**, einer Ökobilanz mit dem Treibhauspotenzial als einzigem Wirkungsindikator. Die Einengung der Wirkungsanalyse spiegelt die Dominanz der Klimaproblematik in der Umweltdebatte der letzten Jahre und die gegenwärtige gesellschaftliche Prioritätensetzung im Umweltschutz wider. Obwohl dies in Fachkreisen sehr kritisch gesehen wird, zeigt die Erfahrung, dass der Carbon Footprint nicht zuletzt aufgrund dieser Vereinfachung heute bereits auf eine wesentlich größere Akzeptanz in Öffentlichkeit und Wirtschaft stößt, als das bei Ökobilanzen in der Vergangenheit der Fall war (Weidema et al., 2008). Die weniger wichtigen Wirkungskategorien müssen nicht zwangsläufig ausgeblendet werden, sondern können als Satisfizierungsziele berücksichtigt werden.

Aus unternehmerischer Sicht müssen Untergrenzen für die ökonomische Performance festgelegt werden, weil sonst auch negative Deckungsbeiträge oder die Einstellung der Produktion als Lösungen möglich sind. Hier zeigt sich wieder die Asymmetrie betriebswirtschaftlicher und ökologischer Aspekte bei der Produktionsplanung: Während eine ausschließlich ökonomische Optimierung denkbar ist, führt eine ausschließlich ökologische Optimierung unter praktischen Gesichtspunkten nicht zu sinnvollen Ergebnissen.

3.4.3 Maximierung der Öko-Effizienz

In den Umweltwissenschaften wurde in den 1990er Jahren in Analogie zur ökonomischen Effizienz (3.17) die **Ökoeffizienz** (Schaltegger und Sturm, 1995: S.20) bzw. **ökologische Produktivität** (Braunschweig und Müller-Wenk, 1993: S.17) als Verhältnis betrieblicher Wertschöpfung und produktionsbedingter Umweltbelastung definiert. Es handelt sich also offensichtlich um eine gemischt ökonomisch-ökologische Zielgröße. Mit den in diesem und dem vorhergehenden Abschnitt definierten stoffstrombasierten Kennzahlen kann die Öko-Effizienz folgendermaßen formalisiert werden:

$$\rho^{\text{öko}} = \frac{G(e)}{u(a) + u(e)} \quad (3.19)$$

Formal entspricht eine Maximierung der Öko-Effizienz der ökonomischen Effizienzmaximierung. Während eine zum Deckungsbeitrag analoge Summenbildung ökologischer und ökonomischer Zielgrößen aufgrund unterschiedlicher Dimensionalität nicht möglich ist, kann ein Verhältnis problemlos gebildet und als Zielgröße eines Maximierungsproblems verwendet werden. Relative Änderungen in der Umweltbelastung und bei der ökonomischen Performance werden gleich gewichtet.

3.4.4 Zusammenfassung

Tabelle 3.5 fasst die Ergebnisse dieses Abschnittes abschließend zusammen. Die verschiedenen Typen ökonomischer und ökologischer Optimierungsprobleme können demnach auf vier unter formalen Gesichtspunkten ähnliche Grundprobleme zurückgeführt werden: die Aufwandsminimierung und die Ertragsmaximierung als unmittelbare Entsprechung zum Minimal- bzw. Maximalprinzip, außerdem die Produktivitätsmaximierung und die Erfolgsmaximierung als Problemklassen, die sich aus der additiven bzw. multiplikativen Kombination von Minimal- und Maximalprinzip ergeben.

Tabelle 3.5: Klassifizierung der wichtigsten stoffstrombasierten Optimierungsproblemtypen nach der Art der verwendeten Zielfunktion. $m(.)$ steht für die monetäre und $u(.)$ für eine skalare ökologische Bewertung der in der betrieblichen Ergebnisbilanz aufgeführten Erträge (e) und Aufwendungen (a).

Typ	Ausprägung			
	ökologisch		ökonomisch	
Aufwandsminimierung	Umweltwirkung	$u(\mathbf{a})^{35}$	Kosten	$m(\mathbf{a})$
Ertragsmaximierung	-	-	Ertrag (real)	e_k
			Erlös	$m(\mathbf{e})$
Produktivitätsmaximierung	Produktivität	$\frac{e_k}{a_1}$	Produktivität	$\frac{e_k}{a_1}$
	Ökoeffizienz	$\frac{e_k}{u(\mathbf{a})}; \frac{m(\mathbf{e})}{m(\mathbf{a})}$ ³⁵	Wirtschaftlichkeit (Effizienz)	$\frac{m(\mathbf{e})}{m(\mathbf{a})}$
Erfolgsmaximierung	-	-	Betriebserfolg bzw. Deckungsbeitrag	$m(\mathbf{e})-m(\mathbf{a})$

³⁵ Entspricht einer „cradle-to-gate“-Bilanzierung. Bei „cradle-to-grave“ wird der ökologische Aufwand durch die Größe $u(\mathbf{a}) + u(\mathbf{e})$ quantifiziert.

Die Asymmetrie ökologischer und ökonomischer Aspekte hinsichtlich ihrer Bedeutung für die betriebliche Planung wird in Tabelle 3.5 dadurch deutlich, dass es keine ökologische Ausprägung der Ertrags- und der Erfolgsmaximierung gibt. Bei den vorrangig ökologisch motivierten Optimierungsproblemen ist zudem die Integration ökonomischer Zielgrößen zwingend notwendig. Im Fall der Öko-Effizienz-Maximierung ist das offensichtlich, weil diese Zielgröße bereits ökonomische Zielgrößen beinhaltet. Aber auch die Minimierung der Umweltwirkung kann in der Praxis nur sinnvoll unter ökonomischen Nebenbedingungen wie Untergrenzen für Erträge oder Umsatzerlöse formuliert werden.

3.5 Beispielhafte Optimierung eines Verbundproduktionssystems

Zur Illustration der theoretischen Überlegungen dieses Kapitels, wird in diesem Abschnitt ein fiktives *Verbundproduktionssystem* optimiert. Von einer Verbundproduktion oder verbundenen Produktion spricht man, wenn ein Produktionssystem mehrere Erträge hervorbringt und dafür zumindest teilweise dieselben Produktionsfaktoren eingesetzt werden. Im Gegensatz zur Kuppelproduktion kann das quantitative Verhältnis der Erträge weitgehend frei variiert werden (vgl. hierzu Möller, 2000: S.142).

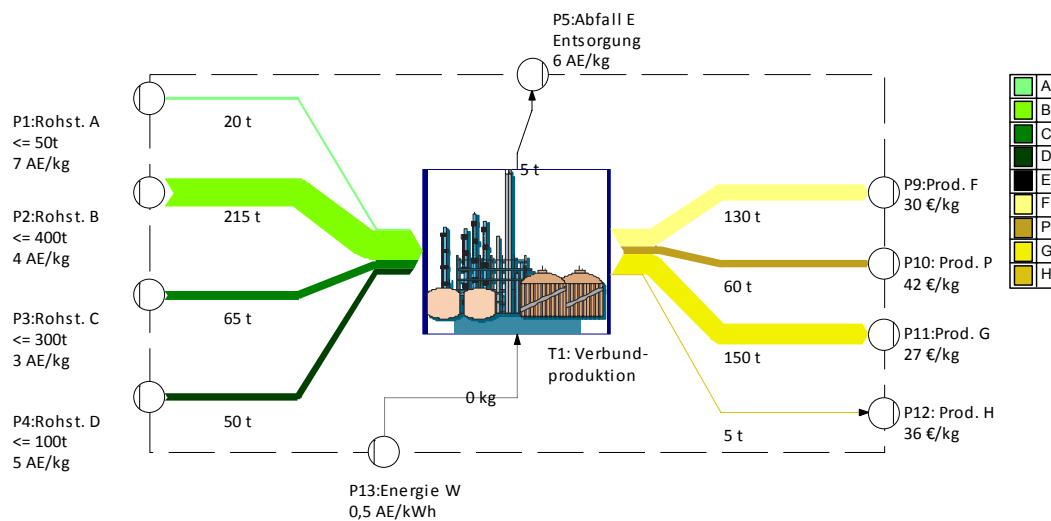


Abbildung 3.12: Fiktives Verbundproduktionssystem zur Illustration der stoffstromnetz-basierten Optimierung. Für alle Erträge sind an den Systemgrenzen Marktpreise [€/kg] angegeben. Aufwendungen werden in Aufwandseinheiten [AE/kg] quantifiziert. Für die Ausgangsstoffe sind außerdem Verfügbarkeitsgrenzen angegeben. Die Jahresproduktion im Status Quo wird als Sankey-Diagramm³⁶ visualisiert.

³⁶ Da der Energieverbrauch in der Grundeinheit kJ quantifiziert wird, wird er im Massen-Sankey-Diagramm der anderen Stoffströme nicht dargestellt (grauer Pfeil mit Masse 0 kg).

Abbildung 3.12 zeigt in der Außenperspektive die Anlage, mit der aus den Vorprodukten A , B , C und D die Produkte F , G , H und P hergestellt werden. Dabei wird die Energiemenge W verbraucht und es entstehen Schadstoffemissionen in Höhe von E . Derartige *Verbundproduktionssysteme* sind typisch für die Prozessindustrie. An Produktionsstandorten z.B. der chemischen Industrie befinden sich häufig mehrere Verfahrenszüge, die bestimmte Ressourcen (z.B. Erdgas, Druckluft, Elektrizität, Arbeitskräfte) oder Anlagen (z.B. Abwasserreinigung) gemeinsam nutzen.

Für die Optimierung des Verbundproduktionssystems werden im Folgenden verschiedene stoffstromnetzbasierende Optimierungsmodelle verwendet. In Abschnitt 3.5.1 werden mehrere Stoffstromnetze eingeführt, die stets dasselbe Originalsystem repräsentieren, dabei jedoch in unterschiedlichem Maße Details seiner internen Struktur und seiner Teilprozesse erfassen. Sie erzeugen daher auch verschiedene Lösungsräume und liefern, selbst bei identischer Zielfunktion, verschiedene Optimierungsergebnisse. Mit diesen Modellen werden in Abschnitt 3.5.2 verschiedene Optimierungsfragestellungen untersucht: eine klassische Gewinnoptimierung, eine Aufwandsminimierung, die sowohl im ökologischen Sinn als Minimierung schädlicher Umweltwirkungen als auch im ökonomischen Sinn als Kostenminimierung interpretiert werden kann, sowie die Maximierung der Effizienz der Produktion.

3.5.1 Verschiedene Stoffstromnetze zur Abbildung des Originalsystems

Im Folgenden werden verschiedene Stoffstromnetze des Verbundproduktionssystems formuliert. Die Reihenfolge ist so gewählt, dass sie einer fortschreitenden Verfeinerung der Systemanalyse entsprechen. Zunächst wird das Produktionssystem vollständig als Black Box behandelt und durch eine *unspezifizierte Transition* beschrieben. Im nächsten Modell wird diese Transition linear spezifiziert. Diese beiden Modelle dienen vor allem zur Demonstration wie sich diese beiden Transitionstypen bei Optimierungen verhalten. Realistischer sind die beiden darauf folgenden Modelle, bei denen das Verbundsystem als zweistufige Produktion mit drei Teilprozessen dargestellt wird.

3.5.1.1 Das Black Box-Modell: Einstufige Modellierung der Produktion mit einer un spezifizierten Transition

In Abschnitt 3.2.2 wurde gezeigt, dass bei *unspezifizierten Transitionen* alle anliegenden Stoffströme unabhängig voneinander variiert werden können und daher Freiheitsgrade darstellen. Aus formaler Sicht spricht zunächst nichts gegen die Verwendung derartig flexibler Modellelemente für die Formulierung eines Optimierungsmodells. Anhand eines Gedankenexperiments soll hier jedoch verdeutlicht werden, dass sich un spezifizierten Transitionen nicht für die Formulierung *sinnvoller* Optimierungsmodelle industrieller Produktionssysteme eignen.

Es soll angenommen werden, dass die Verbundproduktion, wie in Abbildung 3.12 dargestellt, durch eine einzige un spezifizierte Transition modelliert werde. Die anliegenden Stoffströme können nun in Aufwendungen und Erträge eingeteilt werden. Mit einer un spezifizierten Transition ist es nun ohne Weiteres vereinbar, die Aufwendungen auf Null zu setzen und die Erträge beliebig zu steigern. Das widerspricht offensichtlich jeglicher Erfahrung und wurde daher vom Begründer der

produktionstheoretischen Aktivitätsanalyse (Koopmans, 1951: S.33ff) durch das Postulat der „Impossibility of the Land of Cockaigne“ ausgeschlossen. Dyckhoff (1994: S.75) hat darauf hingewiesen, dass die „Unmöglichkeit eines Schlaraffenlandes“ für den Fall einer reinen Güterproduktion aus den Naturgesetzen der Masse- und Energieerhaltung folgt.

Bei aller Einfachheit liefert dieses Gedankenexperiment bereits einen wichtigen Hinweis auf die Rolle von Transitionen in stoffstromnetzbasierten Optimierungsmodellen: Transitionen wirken durch ihre Spezifikation wie *Restriktionen*, die den Raum aller *denkbaren* Input-Output-Kombinationen auf die technisch *machbaren* Lösungen einschränken. Unspezifizierte Transitionen sind aus semantischen Gründen für die Formulierung von Optimierungsmodellen für industrielle Produktionssysteme unbrauchbar und werden daher im Weiteren nicht mehr verwendet.

3.5.1.2 Das 1D-Modell: Einstufige Modellierung der Produktion mit einer linear spezifizierten Transition

Wie im vorigen Abschnitt wird die Verbundproduktion als *einstufiges* Produktionssystem mit einer Transition abgebildet, die nun jedoch *linear spezifiziert* ist (vgl. Abbildung 3.13). Wie in § 3.2.2 gezeigt wurde, steht für die Optimierung dieser linear spezifizierten Transition nur ein einziger Freiheitsgrad zur Verfügung. Daher die Bezeichnung 1D-Modell.

Transition Specifications T1 - Input/Output Relations						
Input / Output		Allocation Rules		Cost Center Costs		Cost Drivers
Var	Place	Material	Coeffi...	B. Unit	DQ	
X00	P1	▲ A		20 kg	●	
X01	P2	▲ B		215 kg	●	
X02	P3	▲ C		65 kg	●	
X03	P4	▲ D		50 kg	●	
X04	P13	▲ W		8,950 EUR	●	
Constraints						
Var	Place	Material	Coeffi...	B. Unit	DQ	
Y00	P5	▲ E		5 kg	●	
Y01	P9	▲ F		130 kg	●	
Y02	P11	▲ G		150 kg	●	
Y03	P12	▲ H		5 kg	●	
Y04	P10	▲ P		60 kg	●	

Abbildung 3.13: Spezifikation der Transition, die die gesamte Verbundproduktion abbildet. Screenshot aus der stoffstromnetzbasierter Stoffstromanalysesoftware Umberto®.

Für eine betriebliche Stoffstromanalyse wäre ein derartiges Modell von geringem Nutzen, weil es die innere Struktur des Produktionssystems nicht transparent macht. Hier geht es jedoch nicht um ein glaubhaftes Beispiel für eine betriebliche Stoffstromanalyse, sondern darum, wie sich die Darstellung von Subsystemen durch lineare Transitionen auf deren Optimierung auswirkt. Insofern mag man sich vorstellen, dass das Verbundsystem aus Abbildung 3.12 nur einen Ausschnitt in einer umfassenderen Stoffstromanalyse darstellt, beispielsweise bei einer Ökobilanz, die auch vor- und nachgelagerte Prozesse berücksichtigt. Bei der Analyse derart umfangreicher Produktionssysteme wäre die stark vereinfachte Darstellung des Verbundsystems mittels einer einzigen linear spezifizierten Transition durchaus denkbar.

3.5.1.3 Das 3D-Modell: Mehrstufige Modellierung der Produktion mit alternativen Verwendungsmöglichkeiten für Zwischenprodukte

Nun wird das Modell der Verbundproduktion weiter disaggregiert. Es wird berücksichtigt, dass die Produktion in *zwei Stufen* erfolgt und aus drei Teilprozessen besteht (Abbildung 3.14). In der ersten Stufe werden die Rohstoffe A-D in den Prozessen *T1* und *T2* zu den Produkten *F-H* verarbeitet. Sie können als Endprodukte verkauft oder in der zweiten Stufe (*T3*) zum Produkt *P* weiterverarbeitet werden.

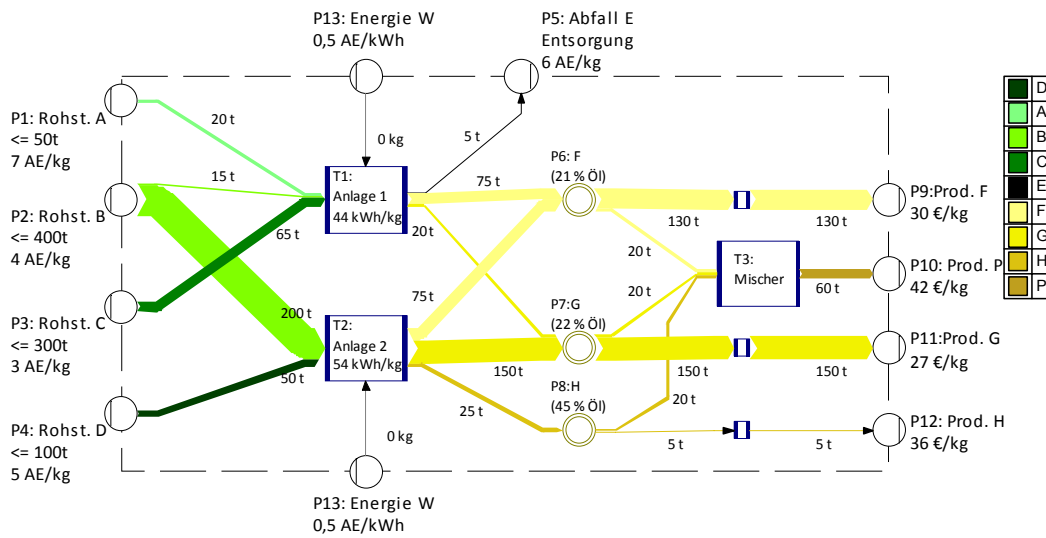


Abbildung 3.14: Verbundproduktion modelliert als zweistufige Technik. Die Stoffströme des Ausgangszustands sind als Sankey-Diagramm dargestellt.

Zusammenhang zwischen der Stoffbilanz des Stoffstromnetzes und denen der Transitionen

Die Teilprozesse werden wie im vorigen Abschnitt als linear spezifizierte Transitionen modelliert. Sie sind ausschließlich durch Verbindungsstellen gekoppelt, d.h. Bestandsveränderungen im Systeminneren werden nicht berücksichtigt. Daher kann die Stoffbilanz z des gesamten Systems als Summe der Transitionsbilanzen $z^{(t)}$ der Teilprozesse dargestellt werden (3.20).

$$z = \sum_{t \in T} z^{(t)} = \sum_{t \in T} \lambda^{(t)} \cdot c^{(t)} \tag{3.20}$$

Die zweite Gleichung in (3.20) gilt, weil alle Transitionen linear spezifiziert sind. Für das in Abbildung 3.14 dargestellte Stoffstromnetz im Ist-Zustand ergibt sich damit die Vektorgleichung (3.21): Die erste Spalte gibt die physikalischen Einheiten und die Stoffarten der Vektorkomponenten an. Die Gesamtstoffbilanz entspricht der des 1D-Modells aus dem vorigen Abschnitt. Die Stoffbilanzen der einzelnen Transitionen können mit dem Stoffstromnetz berechnet werden.

$$\begin{array}{c} [z] \\ \begin{pmatrix} \text{kg } A \\ \text{kg } B \\ \text{kg } C \\ \text{kg } D \\ \text{kg } E \\ \text{kg } F \\ \text{kg } G \\ \text{kg } H \\ \text{kg } P \end{pmatrix} \end{array}
 \begin{array}{c} z^{(1)} \\ \begin{pmatrix} -20 \\ -15 \\ -65 \\ 0 \\ 5 \\ 75 \\ 20 \\ 0 \\ 0 \end{pmatrix} \end{array}
 +
 \begin{array}{c} z^{(2)} \\ \begin{pmatrix} 0 \\ -200 \\ 0 \\ -50 \\ 0 \\ 75 \\ 150 \\ 25 \\ 0 \end{pmatrix} \end{array}
 +
 \begin{array}{c} z^{(3)} \\ \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ -20 \\ -20 \\ -20 \\ 60 \end{pmatrix} \end{array}
 =
 \begin{array}{c} z \\ \begin{pmatrix} -20 \\ -215 \\ -65 \\ -50 \\ 5 \\ 130 \\ 150 \\ 5 \\ 60 \end{pmatrix} \end{array}
 \quad (3.21)$$

Semantik der Freiheitsgrade von Stellen

In Abbildung 3.15 sind die Freiheitsgrade der einzelnen Netzelemente angegeben (vgl. Abschnitt 3.2). Alle Transitionen sind linear spezifiziert und haben deshalb genau einen Freiheitsgrad (+1). Es gibt jedoch Stellen, die über mehrere Freiheitsgrade verfügen. Am Beispiel dieser Stellen kann nun die *Semantik* der formalen Regeln aus Tabelle 3.1 verdeutlicht werden.

- Die *Inputstelle* P2 hat zwei Freiheitsgrade. Formal ergibt sich das daraus, dass die Ersatzformel $\Delta s_B^{(P2)} = -(x_B^{(T1)} + x_B^{(T2)})$ aus einer Gleichung und drei Variablen besteht. Die beiden Freiheitsgrade können nun aber auch im Kontext des Stoffstromsystems *interpretiert* werden. An der Stelle P2 müssen zwei Entscheidungen getroffen werden, nämlich (i) in welcher Quantität Rohstoff B in der Produktion eingesetzt wird und (ii) zu welchen Anteilen er auf die beiden Prozesse T1 und T2 verteilt wird. Parametrisiert werden die Freiheitsgrade durch eine beliebige Auswahl zweier Modellvariablen aus der Menge $(\Delta s_B^{(P2)}, x_B^{(T1)}, x_B^{(T2)})$.
- An der *Verteilerstelle* P8 liegt eine sehr ähnliche Entscheidungssituation vor. Das Produkt H kann hier wahlweise zum Verkauf (T6) oder für die interne Produktion von P in T3 verwendet werden. Der einzige Unterschied zu P2 besteht darin, dass H kein Systeminput ist, sondern innerhalb des Systems in T2 hergestellt wird.
- Die Entscheidungssituationen in P6 und P7 sind identisch. Es handelt sich um kombinierte *Verteiler-* und *Sammlerstellen*. Wie an P8 besteht hier die Wahl zwischen Verkauf und interner Verwendung der Produkte F und G in T3. Im Gegensatz zu P8 ist jedoch nicht festgelegt, woher die zu verteilenden Quantitäten der Produkte F und G stammen. Beide können sowohl in T1 als auch in T2 hergestellt werden.

Freiheitsgrade von Stellen repräsentieren somit *Entscheidungsspielräume* hinsichtlich der *Verteilung variabler Stoffströme auf die gegebene Struktur* eines Stoffstromsystems.

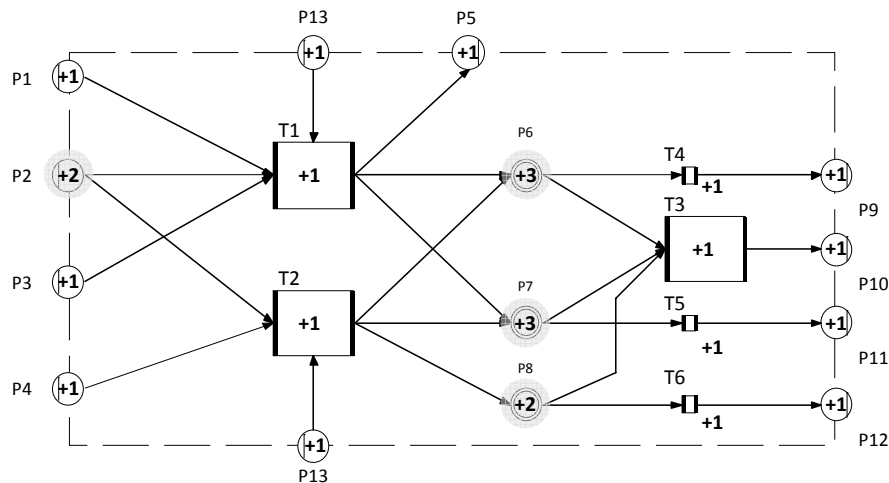


Abbildung 3.15: Freiheitsgrade der einzelnen Netzelemente

Interdependenz lokaler Entscheidungen über das Netz

Die mit den Stellen $P2$, $P6$, $P7$ und $P8$ verbundenen Entscheidungen über die Verteilung von Stoffströmen sind jedoch nicht unabhängig voneinander. Wird in $P2$ beispielsweise festgelegt, wie viel des Rohstoffs B zur Produktion in $T2$ aufgewendet wird, legt diese Entscheidung das Aktivitätsniveau von $T2$ und damit die Ausbringungsmenge $y_H^{(T2,P8)}$ fest. In $P8$ verbleibt damit nur noch ein Freiheitsgrad, nämlich die Entscheidung zwischen Weiterverarbeitung ($T3$) und Verkauf ($T6$). Das ist ein Beispiel für das in Abschnitt 3.2.4 allgemein abgeleitete Prinzip, wonach zur Bestimmung der Freiheitsgrade des Gesamtsystems die Summe der Freiheitsgrade der Netzknoten um die doppelt gezählten Flussrelationen bereinigt werden muss. Im vorliegenden Fall ergibt sich:

$$FG(N) = \sum_{t \in T} FG(t) + \sum_{p \in P} FG(p) - |F| = 6 + 20 - 23 = 3 \quad (3.22)$$

Von insgesamt 23 Stoffflüssen können somit nur drei als Entscheidungsvariablen gewählt werden. Die übrigen Stoffflüsse ergeben sich durch die Modellberechnung. Drei Freiheitsgrade ergeben immer noch ein recht überschaubares Optimierungsmodell, bieten jedoch bereits einen größeren Entscheidungsspielraum als die Abbildung des Systems mit einer einzigen linear spezifizierten Transition. Eine mögliche Parametrisierung der Freiheitsgrade des Gesamtsystems ist:

$$x_0 = \begin{pmatrix} x_B^{(T1)} \\ x_D^{(T2)} \\ x_F^{(T4)} \end{pmatrix} = \begin{pmatrix} 15 \text{ t} \\ 50 \text{ t} \\ 130 \text{ t} \end{pmatrix} \quad (3.23)$$

Die Zahlenwerte für die Komponenten des Vektors auf der rechten Seite von (3.23) entsprechen dem Ist-Zustand des Stoffstromsystems.

3.5.1.4 Das 7D-Modell: Mehrstufige Modellierung der Produktion mit alternativen Verwendungsmöglichkeiten für Zwischenprodukte und substitutionalen Teilprozessen

Für die Formulierung des 7D-Modells wird angenommen, dass in Wirklichkeit (i) $T1$ ein *flexibler Kuppelproduktionsprozess* ist und (ii) an $T2$ und $T3$ die Möglichkeit zur *Substitution* von Produktionsfaktoren besteht.

Modellierung der Transition T1 als flexible Kuppelproduktion

Da $T1$ eine flexible Kuppelproduktion ist, kann das Verhältnis, in dem die Produkte F und G aus $T1$ hervorgehen, variiert werden, d.h. $T1$ ist nicht mehr outputlimitational. Aus technischen Gründen muss der Anteil des Produktes F an der Gesamtproduktion jedoch mindestens 20% betragen. Dieser Sachverhalt wird durch die Restriktion (3.24) ausgedrückt.

$$y_F^{(T1)} \geq 0,2 (y_F^{(T1)} + y_G^{(T1)}) \quad (3.24)$$

Tabelle 3.6 enthält die Spezifikation von $T1$ mittels benutzerdefinierter Funktionen in Umberto®. Die Transitionsspezifikation hat folgende Eigenschaften:

- sie ist *linear*: die Stoffströme tauchen in allen Definitionsgleichungen ausnahmslos in der ersten Potenz auf und es gibt keine Produktterme mehrerer Stoffströme.
- sie ist *inputlimitational*: In den Zeilen 2-7 werden die eingehenden Stoffströme proportional zum Masseumsatz berechnet, sie stehen folglich unabhängig von diesem immer im gleichen Verhältnis zueinander.
- das Verhältnis der beiden Produkte F und G kann frei gewählt werden. Die Zeilen 15-17 definieren diese Form der *verbundenen Produktion*.

Tabelle 3.6: Modellierung der Transition T1 als flexible Kuppelproduktion. Spezifikation erfolgt als Satz benutzerdefinierter Funktionen in Umberto®. Die Spalte „Notation“ enthält die Zuordnung der Stoffstrombezeichner der Transitionsspezifikation in Umberto® und der in dieser Arbeit üblichen Notation (vgl. § 2.4.2).

Spezifikation		Notationen	
<pre> 1 ;starres Verhältnis benötigter Inputs 2 X00=0.2*INP 3 X01=0.15*INP 4 X02=0.65*INP 5 INP=X00/0.2 6 INP=X01/0.15 7 INP=X02/0.65 8 ;5% Abfall 9 Y00=0.05*INP 10 INP=Y00/0.05 11 ;flexible Kuppelproduktion für Y01 und Y02 12 Y02=INP-Y00-Y01 13 Y01=INP-Y00-Y02 14 Y00=INP-Y01-Y02 15 16 ; Arbeit und Energie als Input-Material W: 17 X03=44*INP </pre>		X00	$x_A^{(T1)}$
		X01	$x_B^{(T1)}$
		X02	$x_C^{(T1)}$
		X03	$x_W^{(T1)}$
		Y00	$y_E^{(T1)}$
		Y01	$y_F^{(T1)}$
		Y02	$y_G^{(T1)}$

Modellierung der Transitionen T2 und T3 als Prozesse mit substituierbaren Inputs

Die Spezifikation von T2 als inputvariablem Prozess geht aus Tabelle 3.7 hervor. In den Zeilen 4-6 wird das Verhältnis der beiden Rohstoffe B und D zunächst so modelliert, als seien diese *vollkommene Substitute*. Für die Optimierung wird jedoch die Einschränkung gemacht, dass aus prozesstechnischen Gründen mindestens 20% der eingesetzten Rohstoffmasse aus dem Stoff B bestehen muss. Wie im Falle der Modellierung von $T1$ wird dieser Sachverhalt als zusätzliche algebraische Restriktion (3.25) formuliert werden.

$$x_B^{(T2)} \geq 0,2 (x_B^{(T2)} + x_D^{(T2)}) \tag{3.25}$$

Andererseits ist $T2$ outputlimitational, d.h. bei bekanntem Input können die Quantitäten der verschiedenen Outputs eindeutig berechnet werden. Die entsprechenden Zuweisungen befinden sich in den Zeilen 12-18. Die Modellierung erfolgt spiegelbildlich zur limitationalen Modellierung des Inputs in $T1$. Die Spezifikation von T2 ist linear.

Der Prozess T3 ist ein reiner Mischprozess. Die Ausgangsstoffe F , G und H sind dementsprechend vollkommene Substitute. Die Einsatzverhältnisse führen zu einem entsprechenden Mischungsverhältnis. Hier wird nun angenommen, dass die Stoffe verschiedene Ölgehalte haben: $\rho_F=21\%$,

$\rho_G=22\%$ und $\rho_H=45\%$. An das durch die Mischung erzeugte Produkt P wird nun die Qualitätsanforderung gestellt, dass es ein Ölgehalt von mindestens 27% aufweisen soll: $\rho_P \geq 27\%$. Daraus ergibt sich die Restriktion (3.26).

$$0,27y_P^{(T3)} \leq 0,21x_F^{(T3)} + 0,22x_G^{(T3)} + 0,45x_H^{(T3)} \quad (3.26)$$

Tabelle 3.7: Modellierung der Transition T2 als Prozess mit zwei substituierbaren Inputs. Die Spalte „Notation“ enthält die Zuordnung der Stoffstrombezeichner der Transitionsspezifikation in Umberto® und der in dieser Arbeit üblichen Notation (vgl. § 2.4.2).

Spezifikation	Notationen	
<pre> 1 / Gesamtumsatz (INP) als Fkt. der Inputs 2 INP=X01+X00 3 X00=INP-X01 4 X01=INP-X00 5 / X00 und X01 weitgehend substituierbar. 6 / Einschränkung: X00 >= 0.2*INP, 7 / Ungleichungsrestriktion hier nicht formuliert 8 9 / Output als Fkt. des Gesamtumsatzes (INP) 10 Y00=0.3*INP 11 Y01=0.6*INP 12 Y02=0.1*INP 13 / Umkehrung 14 INP=Y00/0.3 15 INP=Y01/0.6 16 INP=Y02/0.1 17 18 / Arbeit und Energie als Input-Material W: 19 X02=54*INP </pre>	X00	$x_B^{(T2)}$
	X01	$x_D^{(T2)}$
	X02	$x_W^{(T2)}$
	Y00	$y_F^{(T2)}$
	Y01	$y_G^{(T2)}$
	Y02	$y_H^{(T2)}$

Einfluss der variablen Prozessmodellierung auf die Freiheitsgrade des Stoffstromnetzes

Zur Untersuchung der Frage, wie diese Änderungen in der Abbildung der Elementarprozesse den Lösungsraum des Optimierungsmodells beeinflussen, geht man wiederum von Abbildung 3.15 des vorigen Abschnitts aus. Die Beiträge von Stellen und Flussrelationen bleiben unverändert. Die Transitionen T1 und T2 haben jeweils einen Freiheitsgrad mehr: Bei T1 kann das Outputverhältnis von F und G gewählt werden, bei T2 das Verhältnis der Inputs B und D. T3 erhält sogar zwei zusätzliche Freiheitsgrade, weil das Mischungsverhältnis aller drei Inputs frei wählbar ist. Die Anzahl der Freiheitsgrade steigt entsprechend der veränderten Freiheitsgrade der Transitionen von 3 auf 7. Zur Parametrisierung der Freiheitsgrade müssen sieben unabhängige Modellvariablen ausgewählt werden. Hier wird folgender Entscheidungsvektor gewählt:

$$x_0 = \begin{pmatrix} x_B^{(T1)} \\ x_B^{(T2)} \\ x_D \\ y_F^{(T1)} \\ y_F \\ y_G \\ x_H^{(T3)} \end{pmatrix} = \begin{pmatrix} 15 \text{ t} \\ 200 \text{ t} \\ 50 \text{ t} \\ 75 \text{ t} \\ 130 \text{ t} \\ 150 \text{ t} \\ 20 \text{ t} \end{pmatrix} \quad (3.27)$$

3.5.2 Optimierungsergebnisse und Interpretation

Bei der Optimierung des Verbundproduktionssystems werden im Folgenden drei verschiedene Varianten untersucht:

- Gewinnmaximierung
- Aufwandsminimierung
- Effizienzmaximierung

Alle genannten Zielgrößen hängen ausschließlich von der Stoffbilanz des Gesamtsystems ab. Sie sind unabhängig von der inneren Struktur des zur Systemabbildung verwendeten Stoffstromnetzes und können hier allgemein für alle in diesem Abschnitt verwendeten Stoffstromnetze formuliert werden.

3.5.2.1 Gewinnmaximierung

Die Optimierung des mit der Verbundproduktion erwirtschafteten Gewinns ist eine klassische Problemstellung des operativen Produktionsmanagements. Zur Berechnung des Gewinns werden alle Stoffströme an den Systemgrenzen monetär bewertet. Reale Erträge führen zu Umsatzerlösen und reale Aufwendungen verursachen Produktionskosten. Die Zuordnung der Stoffströme zu Erträgen und Aufwendungen des Systems geschieht durch die in § 3.3.2 eingeführte Vorzeichen-Transformation δ . Die monetäre Bewertung erfolgt durch die Marktpreise p_k der jeweiligen Stoffart. Zusammenfassend ergibt sich Gleichung (3.28) zur Berechnung des Gewinns des Stoffstromnetzes. Darin haben alle Stoffströme z die Einheit [kg/Zeiteinheit] bzw. [kWh/Zeiteinheit], spezifische Preise p die Einheit bzw. [€/kWh]. Kosten, Erlöse und Gewinn werden folglich in [€/Zeiteinheit] angegeben.

$$\text{Gewinn } (G_0) = \sum_k z_{k,0} \cdot \delta_k \cdot p_k$$

$$\text{wobei gilt: } \delta_k = \begin{cases} +1 & \text{Gut} \\ 0 & \text{für } k \text{ ist ein Neutrum} \\ -1 & \text{Übel} \end{cases} \quad (3.28)$$

Damit ergibt sich in der Ausgangssituation folgender Gewinn:

$$\begin{aligned}
 G_0 &= \sum_{l=\{F,G,H,P\}} z_{l,0} \cdot p_l - \sum_{k=\{A,B,C,D,W,E\}} z_{k,0} p_k \\
 &= \text{Erlös } (E_0) - \text{Kosten } (K_0) \\
 &= 10.650 \text{ k€}/a - 10.425 \text{ k€}/a = 225 \text{ k€}/a
 \end{aligned}
 \tag{3.29}$$

Gewinnmaximierung im 1D-Modell

Die Gewinnoptimierung des 1D-Modells kann leicht explizit durchgeführt werden. Da das Stoffstromnetz aus einer einzigen linear spezifizierten Transition besteht, wird sein Zustand durch das Aktivitätsniveau $\lambda^{(T1)} = \lambda$ dieser Transition bereits vollständig beschrieben. Im Ausgangszustand beträgt es:

$$\lambda_0 = \frac{z_{k,0}^{(T1)}}{c_k^{(T1)}} = 10^3
 \tag{3.30}$$

Da alle Stoffströme der Transition und damit auch die Stoffbilanz des gesamten Systems linear vom Aktivitätsniveau abhängen, kann auch der mit der Produktion verbundene Gewinn als Produkt des Aktivitätsniveaus mit einer von der Transitionspezifikation abhängigen Konstante $G^{(T1)}$ ausgedrückt werden:

$$G(\lambda) = \lambda \cdot \sum_k c_k^{(T1)} \delta_k p_k = \lambda \cdot G^{(T1)} \quad \text{mit} \quad G^{(T1)} = 225 \text{ €/a}
 \tag{3.31}$$

Da $G^{(T1)}$ positiv ist, muss zur Ergebnismaximierung in diesem einfachen Modell das Aktivitätsniveau λ soweit wie möglich gesteigert werden. Dabei ist das Aktivitätsniveau nach oben durch die Verfügbarkeitsrestriktionen der Rohstoffe A-D (Abbildung 3.12) beschränkt. Bezeichnet man die rohstoffspezifischen oberen Schranken mit u_k , dann ergeben sich folgende vier Restriktionen:

$$x_k = \lambda \cdot a_k \leq u_k \quad \text{mit} \quad k = \{A, B, C, D\}
 \tag{3.32}$$

Das optimale Aktivitätsniveau entspricht der kleinsten Schranke bezüglich λ (3.33). Weiter kann die Produktion nicht gesteigert werden, ohne mindestens eine der vier Verfügbarkeitsrestriktionen (3.32) zu verletzen. Der berechnete optimale Gewinn in Höhe von ca. 419 k€/a entspricht einer Gewinnsteigerung von 86% gegenüber dem Ist-Zustand.

$$\lambda^* = \min\left(\frac{u_k}{a_k}\right) = \frac{u_B}{a_b} = \frac{400 \text{ t}}{215 \text{ kg}} = 1860 \quad (3.33)$$

$$\Rightarrow G^* = \lambda^* G^{(T1)} \approx 419 \text{ k€}/a$$

Gewinnmaximierung im 3D-Modell

Die drei Freiheitsgrade des 3D-Modells werden durch den in §3.5.1.3 angegebenen Stoffstromvektor x_0 parametrisiert, der hier als Entscheidungsvektor eines dreidimensionalen linearen Optimierungsproblems interpretiert wird. Die Optimierung ergibt die Lösung:

$$x^* = \begin{pmatrix} 37,5 \text{ t} \\ 90,6 \text{ t} \\ 270,1 \text{ t} \end{pmatrix}, G^* = 820 \text{ k€}/a \quad (3.34)$$

Die Modellierung der Verbundproduktion als zweistufige Technik und die damit verbundene Erschließung von Freiheitsgraden führen zu einer Lösung, mit der eine Gewinnsteigerung um 265% gegenüber dem Ausgangszustand einhergeht. Dieses Ergebnis liegt damit knapp um einen Faktor 2 über dem des 1D-Modells. Die Entscheidungsspielräume an den Verteiler- und Sammlerstellen führen dazu, dass das Stoffstromnetz insgesamt nicht mehr limitational ist, obwohl es ausschließlich aus limitationalen Transitionen aufgebaut ist. Daher besteht bei der Optimierung die Möglichkeit, nicht nur das Niveau der Produktion anzupassen, sondern auch das Einsatzverhältnis der Produktionsfaktoren. Dass diese zusätzlichen Freiheitsgrade für die Optimierung des Modells notwendig sind, erkennt man daran, dass der optimale Entscheidungsvektor x^* linear unabhängig vom Ausgangsvektor x_0 ist³⁷.

Gewinnmaximierung im 7D-Modell

Die Optimierung des 7D-Modells liefert schließlich die Lösung (3.35). Die Berücksichtigung der Substituierbarkeit von Inputs und Outputs an den Transitionen des Stoffstromnetzes führt zu einer weiteren deutlichen Verbesserung des Optimierungsergebnisses. Der Gewinn kann unter Berücksichtigung aller Freiheitsgrade des Systems fast um den Faktor 10 gesteigert werden.

³⁷ d.h. es gibt keinen skalaren Faktor α , für den $x^* = \alpha x_0$ ist.

$$x^* = \begin{pmatrix} 37,5 \text{ t} \\ 162,5 \text{ t} \\ 100,0 \text{ t} \\ 237,5 \text{ t} \\ 376,2 \text{ t} \\ 111,0 \text{ t} \\ 46,3 \text{ t} \end{pmatrix}, G^* = 2220 \text{ k€}/a \quad (3.35)$$

Zusammenfassung der Ergebnisse der Gewinnmaximierung mit den verschiedenen Modellen

Zusammenfassend wird deutlich, dass die Gewinnmaximierung zu umso besseren Ergebnissen führt, je mehr Freiheitsgrade des Originalsystems bei der Abbildung als Stoffstromnetz berücksichtigt werden. Aus Tabelle 3.8 geht außerdem hervor, dass die Maximierung des Gewinns mit einer deutlichen Aufwandssteigerung einhergeht und außerdem auch zu einer Steigerung der Effizienz führt.

Tabelle 3.8: Ergebnisse der Gewinnmaximierung. Die Reihen mit den fett gedruckten Zahlen enthalten die Zielgröße der Optimierung, die in der letzten Spalte als Balkendiagramm dargestellt ist. Die Spalten entsprechen den verschiedenen Stoffstromnetzen aus Abschnitt 3.5.1.

	Referenz	1D	3D	7D	Gewinn
Erlös [€]	10.650,00	19.813,95	21.520,31	23.220,00	
Aufwand [AE]	10.425,00	19.395,35	20.700,00	21.000,00	
Gewinn [€]	225	419	820	2.220	
Effizienz [€/AE]	1,0216	1,0216	1,0396	1,1057	
Gewinnanstieg (relativ) [%]	0	86	265	887	

3.5.2.2 Aufwandsminimierung

Die Aufwandsminimierung spielt sowohl bei der ökonomischen als auch bei der ökologischen Optimierung der Verbundproduktion eine wichtige Rolle. Je nach verwendeter Bewertungsfunktion für das Produktionsergebnis resultiert eine Kostenminimierung oder die Minimierung schädlicher Umweltwirkungen des Produktionssystems. Um die Einstellung der Produktion als mathematische Lösung des Problems auszuschließen, werden wahlweise zwei unterschiedliche Restriktionen betrachtet: (i) der *Ertrag* darf sich nicht verringern, (ii) der *Erlös* darf nicht sinken. Während im ersten Fall der Outputvektor des Systems unverändert bleiben muss, kann er im zweiten Fall verändert werden unter der Voraussetzung, dass der Erlös gleich bleibt bzw. nicht sinkt (gegen eine Steigerung ist nichts einzuwenden).

Nebenbedingung: konstanter oder steigender Systemertrag

Bei der Aufwandsminimierung unter der Nebenbedingung, dass kein realer Ertrag sinken darf, ist eine Aufwandsreduktion nur im Falle des 7D-Modells möglich. Aufgrund der Limitationalität des 1D-Modells ist eine Reduzierung des Aufwands bei konstantem Ertrag offensichtlich unmöglich. Doch selbst die zusätzlichen Freiheitsgrade durch alternative Verwendungsmöglichkeiten für interne Zwischenprodukte beim 3D-Modell sind nicht ausreichend, den Aufwand unter der Nebenbedingung $e \geq e_0$ weiter zu senken.

Tabelle 3.9: Ergebnisse der Aufwandsminimierung bei konstantem Ertragsvektor. Die Reihen mit den fett gedruckten Zahlen enthalten die Zielgröße der Optimierung, deren relative Veränderung bezüglich des Status Quo in der letzten Spalte als Balkendiagramm dargestellt ist. Die Spalten entsprechen den verschiedenen Stoffstromnetzen aus Abschnitt 3.5.1.

	Referenz	1D	3D	7D	relativer Aufwandsrückgang
Erlös [€]	10.650,00	10.650,00	10.650,00	10.650,00	
Aufwand [AE]	10.425,00	10.425,00	10.425,00	10.140,68	
Gewinn [€]	225,00	225,00	225,00	509,32	
Effizienz [€/AE]	1,0216	1,0216	1,0216	1,0502	
Aufwandsrückgang [%]	0,0%	0,0%	0,0%	-2,7%	

Nebenbedingung: konstante oder steigende Umsatzerlöse

Aus Tabelle 3.10 geht hervor, dass bei der Aufwandsminimierung deutlich bessere Ergebnisse erzielt werden, wenn nicht bezüglich des Ertragsvektors, sondern nur für die entsprechenden Umsatzerlöse $m(e)$ eine Untergrenze festgelegt wird.

Tabelle 3.10: Ergebnisse der Aufwandsminimierung bei konstantem Erlös. Die Reihen mit den fett gedruckten Zahlen enthalten die Zielgröße der Optimierung, die in der letzten Spalte als Balkendiagramm dargestellt ist. Die Spalten entsprechen den verschiedenen Stoffstromnetzen aus Abschnitt 3.5.1.

	Referenz	1D	3D	7D	relativer Aufwandsrückgang
Erlös [€]	10.650,00	10.650,00	10.650,00	10.650,00	
Aufwand [AE]	10.425,00	10.425,00	10.134,46	9.487,07	
Gewinn [€]	225,00	225,00	515,54	1.162,93	
Effizienz [€/AE]	1,0216	1,0216	1,0509	1,1226	
Aufwandsrückgang [%]	0,0	0,0	-2,8	-9,0	

Das Ergebnis wird nachvollziehbar, wenn man die optimalen Ertragsvektoren für die verschiedenen Modelle betrachtet (Tabelle 3.11). Da beim 1D-Modell das Verhältnis der Erträge durch die Transitionskoeffizienten fixiert ist, kann der Umsatzerlös nur konstant bleiben, wenn der gesamte Ertragsvektor unverändert bleibt. Daher entsprechen sich die ersten beiden Spalten in Tabelle 3.11. Sowohl beim 3D als auch beim 7D-Modell kann jedoch der Aufwand deutlich gesenkt werden, indem die Möglichkeit der Veränderung des Produktionsprogramms ausgenutzt wird. Während im Optimum des 3D-Modells hauptsächlich F produziert wird, wird der größte Teil des Umsatzes im Optimum des 7D-Modells mit dem Verkauf von P erzielt.

Tabelle 3.11: Ertragsvektoren der Verbundproduktion. Die erste Spalte enthält als Referenz den Ertrag des Ist-Zustandes. Alle weiteren Spalten enthalten die optimalen Erträge der jeweiligen Optimierungsmodelle 1D, 3D und 7D für die Aufwandsminimierung bei konstanten Umsatzerlösen.

	Referenz	1D	3D	7D
e_P	60,0	60,0	34,3	140,8
e_F	130,0	130,0	210,4	91,8
e_G	150,0	150,0	107,2	73,4
e_H	5,0	5,0	0,0	0,0

3.5.2.3 Effizienzmaximierung

Auch hinsichtlich der Effizienzmaximierung zeigt sich eine zunehmende Verbesserung der Optimierungsergebnisse mit steigender Anzahl von Freiheitsgraden des Optimierungsmodells.

Tabelle 3.12: Ergebnisse der Effizienzmaximierung. Die Reihen mit den fett gedruckten Zahlen enthalten die Zielgröße der Optimierung, die in der letzten Spalte als Balkendiagramm dargestellt ist. Die Spalten entsprechen den verschiedenen Stoffstromnetzen aus Abschnitt 3.5.1.

	Referenz	1D	3D	7D	relative Effizienzsteigerung
Erlös [€]	10.650,00	10.650,00	6.975,00	13.920,00	
Aufwand [AE]	10.425,00	10.425,00	6.562,50	12.400,00	
Gewinn [€]	225,00	225,00	412,50	1.520,00	
Effizienz [€/AE]	1,02	1,02	1,06	1,12	
Effizienzsteigerung [%]	0,0	0,0	4,0	9,9	

Wiederum ist das 1D-Modell für diese Optimierungsfragestellung weitgehend untauglich. Aufgrund der Limitationalität der linear spezifizierten Transition ist das Verhältnis von bewertetem Ertrag zu bewertetem Aufwand unabhängig vom Aktivitätsniveau stets dasselbe. Die Effizienz des Prozesses wird bereits bei der Transitionsspezifikation festgelegt und kann in einer Optimierung

nicht verbessert werden. Mit anderen Worten: die Zielgröße Effizienz hängt nicht von der einzigen vorhandenen Entscheidungsvariable ab.

Die Ergebnisse für das 3D und das 7D Modell sind dagegen interessant (vgl. Tabelle 3.12). Während die Höhe der Effizienzsteigerung erwartungsgemäß mit der Anzahl der Freiheitsgrade des jeweiligen Optimierungsmodells zunimmt, liefern beide Modelle ansonsten recht gegenläufige Ergebnisse. Im Optimum des 3D-Modells wird der Aufwand deutlich reduziert. Dennoch können sowohl der Gewinn als auch die Effizienz der Produktion gesteigert werden. Im 7D-Modell nehmen Erlös und Aufwand dagegen zu. Dieses erste Indiz für strukturell sehr unterschiedliche Lösungen wird durch eine vergleichende Darstellung der optimalen Systemzustände in Abbildung 3.16 bestätigt. Bei der optimalen Lösung des 3D-Modells (a) werden die Teilprozesse T2 und T3 stillgelegt und ausschließlich mit T1 produziert. Beim 7D-Modell (b) kehrt sich die Situation vollständig in ihr Gegenteil und T1 wird als einziger Prozess stillgelegt.

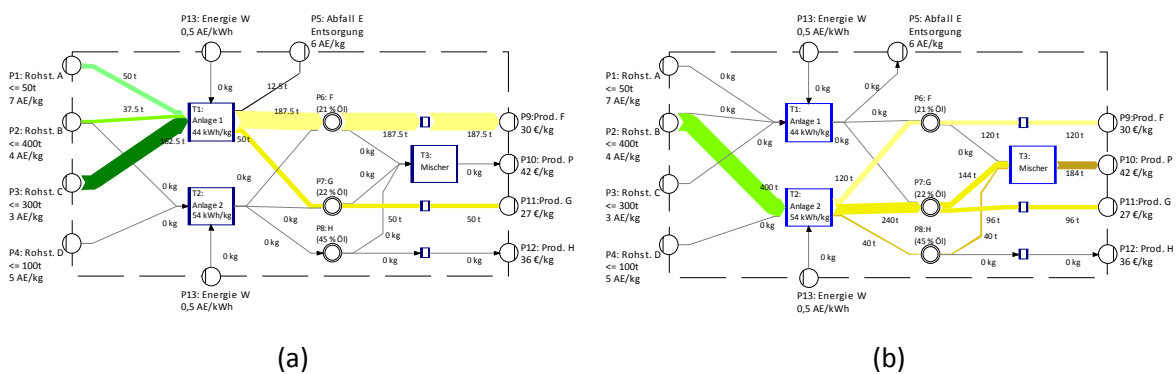


Abbildung 3.16: Sankey-Diagramme der effizienzmaximalen Systemzustände für (a) das 3D- und (b) das 7D-Modell für die stoffstromnetzbasierende Optimierung.

Im Vergleich mit der Gewinnmaximierung in §3.5.2.1 fällt schließlich auf, dass Gewinn und Effizienzsteigerung zwar korrelieren, für das Verbundproduktionssystem jedoch zu unterschiedlichen optimalen Zuständen führen.

3.6 Zusammenfassung

In diesem Kapitel wurde untersucht, wie die Stoffstromanalyse mit Stoffstromnetzen als Modellieretechnik im Rahmen einer Optimalplanung eingesetzt werden kann. In Abschnitt 3.1 wurde mit dem stoffstromnetzbasierten Optimierungsproblem (SSNOP) zunächst eine formale Schnittstelle zwischen Stoffstromnetzen und mathematischen Optimierungsverfahren des OR entwickelt. Dabei wurde die zentrale Bedeutung des Modellgültigkeitsbereichs eines Stoffstromnetzes aufgezeigt. Er steckt den Bereich ab, innerhalb dessen bei der stoffstromnetzbasierten Optimierung nach optimalen Lösungen gesucht werden kann. Nur wenn *Handlungs- und Gestaltungsoptionen* für die Verbesserung eines Produktionssystems bei der Modellbildung hinreichend berücksichtigt

werden, kann die Anwendung mathematischer Optimierungsalgorithmen zu befriedigenden Lösungen für die Praxis führen.

In Abschnitt 3.2 wurde daher – zunächst auf einer formalen Ebene – untersucht, wie Freiheitsgrade bei der Systemabbildung mit Stoffstromnetzen berücksichtigt werden können. Es wurde gezeigt, welche Art von Freiheitsgraden mit den verschiedenen Stellen- und Transitionstypen abgebildet werden können, und wie diese durch die Flussrelationen zum Modellgültigkeitsbereich des gesamten Stoffstromnetzes kombinieren. So konnte gezeigt werden, dass das Stellen-Transitions-Konzept der Stoffstromnetze prinzipiell die notwendige Flexibilität bietet, um die für die Planung industrieller Produktionssysteme notwendigen Steuerungs- bzw. Gestaltungsspielräume abzubilden. Der darauf folgende Abschnitt befasste sich mit der stoffstromnetzbasierter Konstruktion *sinnvoller* Zielgrößen für die ökonomische und ökologische Optimierung von Produktionssystemen. Die Überlegungen dieses Abschnitts sind für das stoffstromnetzbasierter Stoffstrommanagement nicht vollkommen neu, wurden hier jedoch speziell unter dem Blickwinkel der Optimalplanung dargestellt.

Ein zentrales Ergebnis dieses Kapitels ist dagegen die Tatsache, dass sich stoffstrom(netz)basierter Optimierungsprobleme im Hinblick auf die üblichen Gründe für das betriebliche Stoffstrommanagements weitgehend systematisieren lassen. In Abschnitt 3.4 wurde zunächst zwischen betriebswirtschaftlichen und vorwiegend ökologisch motivierten Optimierungsproblemen unterschieden. Klassische betriebswirtschaftliche Optimierungsprobleme sind die Kostenminimierung oder die Gewinn-, Deckungsbeitrags- oder Effizienzmaximierung. Hierbei können ökologische Randbedingungen einfließen. Bei den vorwiegend ökologisch motivierten Optimierungsproblemen geht es typischerweise um die Minimierung der produktionsbedingten Umweltbelastungen, beispielsweise also die Minimierung des Carbon Footprints eines Unternehmens. Ohne die Integration ökonomischer Nebenbedingungen wäre hier stets die Stilllegung eines Betriebs die optimale, für praktische Fragestellungen jedoch kaum befriedigende Lösung. Eine dritte Kategorie von Optimierungsproblemen ergibt sich durch die Verwendung der Ökoeffizienz als Zielfunktion. Hier werden ökologische und ökonomische Belange auf der Ebene der Zielfunktion kombiniert.

Anhand der Optimierung eines fiktiven Verbundproduktionssystems mit unterschiedlichen Optimierungsmodellen wurde in Abschnitt 3.5 die *semantische* Ebene der stoffstromnetzbasierter Optimalplanung genauer beleuchtet. So wurde deutlich, dass substitutionale Prozesse sowie Mehrertragsprozesse, bei denen das Verhältnis der Erträge zumindest in gewissen Grenzen beeinflussbar ist, wichtige Freiheitsgrade für die Systemoptimierung erschließen. Dasselbe konnte für innersystemische Verzweigungen von Stoffströmen gezeigt werden. So kann an Sammlerstellen zwischen verschiedenen Bezugsquellen eines Produktionsfaktors gewählt werden, beispielsweise zwischen elektrischem Strom aus verschiedenen Kraftwerken. Umgekehrt repräsentieren Verteilerstellen verschiedene Optionen für die Verwendung eines Zwischenprodukts, insbesondere z.B. die Entscheidung zwischen Verkauf und interner Weiterverarbeitung

Die verschiedenen stoffstromnetzbasierter Optimierungsmodelle für das Verbundproduktionssystem unterschieden sich einerseits im Hinblick auf das für die Systemabbildung verwendete

Stoffstromnetz, andererseits durch die Art der Zielfunktion. Obwohl die quantitativen Ergebnisse streng genommen ausschließlich für das Fallbeispiel gelten, vermitteln sie einen durchaus verallgemeinerbaren Eindruck vom Einfluss der Modellbildung auf die Güte gefundener Optima. Unabhängig von der jeweiligen Zielfunktion zeigte sich, dass umso bessere Optimierungsergebnisse erzielt werden konnten, je mehr Freiheitsgrade bei der Abbildung des Originalsystems berücksichtigt werden.

4 Lösung des stoffstromnetzbasierenen Optimierungsproblems (SSNOP)

Dieses Kapitel befasst sich mit der Lösung des im vorigen Kapitel entwickelten stoffstromnetzbasierenen Optimierungsproblems (SSNOP). Grundsätzlich kommen zwei Lösungsansätze in Frage: die *simulationsbasierte Optimierung* (SBO) und die *mathematische Optimierung*. Im Folgenden wird untersucht, welcher Ansatz bei welchen Eigenschaften eines SSNOPs vorzuziehen ist bzw. überhaupt für die Lösung in Frage kommt. Darauf aufbauend wird die softwaretechnische Umsetzung der Lösung stoffstromnetzbasierenen Optimierungsprobleme erörtert. Dabei wird der Ansatz verfolgt, eine in der von Möller (2000) beschriebenen Weise auf dem Stoffstromnetzkonzept basierende Stoffstromanalysesoftware³⁸ um ein *Optimierungsmodul* zu erweitern, das die notwendige Funktionalität für die Ergänzung des Stoffstrommodells zum SSNOP und für die automatische Modelloptimierung zur Verfügung stellt.

4.1 Mathematische und simulationsbasierte Optimierung

4.1.1 Mathematische Optimierung

Der im Operations Research am häufigsten verwendete Ansatz zur Lösung von Parameteroptimierungsproblemen ist die mathematische Programmierung. Entscheidungs- bzw. Optimierungsprobleme werden dabei zunächst als *mathematische Programme* formuliert. Weil mathematische Programme die Struktur von Optimierungsproblemen direkt in der Sprache der Mathematik abbilden, können für die Optimierung besonders leistungsfähige Optimierungsverfahren bzw. Solver eingesetzt werden, die schnell und mit hoher Sicherheit globale Optima finden. Je nach Problemeigenschaften können sogar exakte Lösungen gefunden werden. Im Hinblick auf die Lösung des SSNOPs ist jedoch zu bedenken, dass dieses zunächst in ein mathematisches Programm *transformiert* werden muss (vgl. Abschnitt 4.1.3).

Analytische Optimierungsverfahren

Die Optimierung mathematischer Programme wird in dieser Arbeit als *mathematische Optimierung* bezeichnet. Für die Optimierung mathematischer Programme können *analytische* Optimie-

³⁸ Ein Beispiel ist die Stoffstromanalysesoftware Umberto (ifu Hamburg, 2011).

rungsverfahren³⁹ eingesetzt werden. Diese setzen auf der Ebene der Problemformulierung an, d.h. sie nutzen für die Lösung eines Optimierungsproblems seine mathematische Struktur aus. Es gibt Verfahren mit und solche ohne garantierten Erfolg (Syrjakow, 2003: S.20).

Das klassische Beispiel für ein *exaktes* analytisches Verfahren mit garantiertem Erfolg ist der Simplexalgorithmus für lineare Programme. Er basiert auf einer geometrischen Interpretation der Problemstruktur. Bei linearen Programmen hat der Lösungsraum die Gestalt eines Polyeders. Weil die Zielfunktion linear ist, liegt das globale Optimum immer in einer Ecke, ggf. auch entlang einer Kante dieses Polyeders. Indem sich der Algorithmus entlang der Kanten des Polyeders bewegt, kann er das globale Optimum in einer endlichen Anzahl von Iterationen exakt bestimmen.

Analytische Verfahren *ohne garantierten Erfolg* beruhen häufig auf lokalen Suchstrategien, d.h. sie versuchen die Problemlösung ausgehend von einem Startwert iterativ zu verbessern. Beim Durchsuchen des Lösungsraumes verwenden sie analytische Zusatzinformationen, insbesondere Ableitungen der Zielfunktion, um geeignete Suchrichtungen und Schrittweiten zu bestimmen. Die wichtigsten Vertreter solcher ableitungsbasierter Verfahren sind die *Methode des steilsten Abstiegs* (Gradientenverfahren) und *Newtonverfahren* (Luenberger und Ye, 2008: S.234 u. 246ff).

Solver

Für die Lösung mathematischer Programme werden üblicherweise so genannte *Solver* verwendet. Manche Solver implementieren einen einzelnen Optimierungsalgorithmus. Häufig kombinieren sie jedoch mehrere Optimierungsverfahren, um entweder verschiedenen Typen von Optimierungsproblemen abzudecken oder besser mit schwierigen Sonderfällen umgehen zu können⁴⁰. Solver sind in den meisten Fällen eigenständige Softwarekomponenten, die in Kombination mit verschiedenen Entwicklungsumgebungen für mathematische Programme (z.B. GAMS, AMPL oder LPL) eingesetzt werden können. Mittelmann (2010) bietet einen umfassenden Überblick über die aktuell verfügbaren Solver. Für die Lösung linearer Programme stehen nicht weniger als 40 verschiedene Solver zur Verfügung.

Anwendbarkeit auf das SSNOP

Der Vorteil der mathematischen Programmierung liegt darin, dass heute eine Vielzahl sehr leistungsfähiger Optimierungsalgorithmen und Solver für die Lösung mathematischer Programme zur Verfügung steht. Um diese jedoch auf das SSNOP anwenden zu können, muss dieses zuvor vollständig in ein mathematisches Programm transformiert wurde (Abbildung 4.1). Dazu müssen die durch das Stoffstromnetz definierte Funktion $r^S(x;p)$ sowie die durch stoffstromnetzbasierete

³⁹ Analytische Optimierungsverfahren werden in Abgrenzung zu den so genannten direkten Optimierungsverfahren (vgl. §0) auch als *indirekte* Optimierungsverfahren bezeichnet (vgl. Schwefel, 1995: S.13ff, Syrjakow, 2003: S.11).

⁴⁰ Der besonders vielseitig einsetzbare Solver KNITRO kombiniert beispielsweise drei verschiedene Optimierungsstrategien (Ziena Optimization Inc., 2009: S.2f).

Auswertungsmethoden gewonnenen Kennzahlen $a(r^S, x; p)$ vollständig durch *algebraische* Ausdrücke *reformuliert* werden.

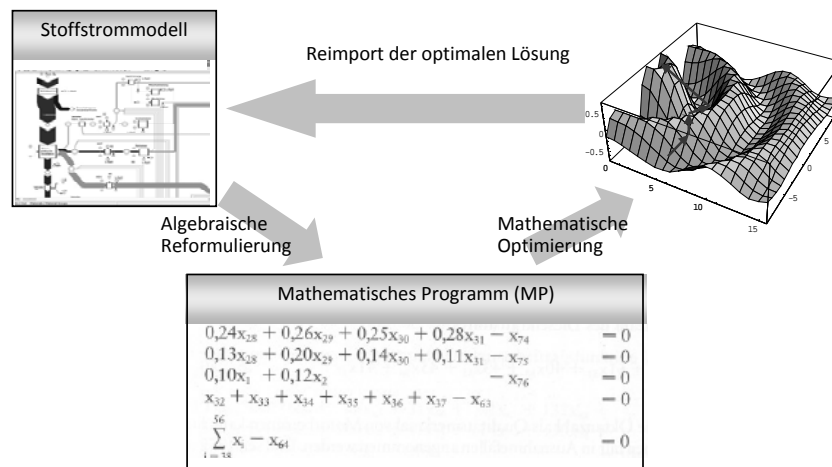


Abbildung 4.1: Prinzip der algebraischen Reformulierung und mathematischen Optimierung eines SSNOP

Das sequentielle Schema aus stoffstromnetzbasierter Problemspezifikation, algebraischer Reformulierung und mathematischer Optimierung entspricht dem Ansatz des *Visual Interactive Modeling* (Bell, 1985, Bell et al., 1999). Die Stoffstromnetze fungieren in diesem Fall als intuitive grafische Benutzerschnittstelle für die Formulierung mathematischer Programme für die Produktionsplanung.

4.1.2 Simulationsbasierte Optimierung

Die simulationsbasierte Optimierung ist ein Optimierungsansatz, mit dem optimale Zustände von Simulationsmodellen bestimmt werden können. Sie unterscheidet sich von der mathematischen Programmierung grundsätzlich dadurch, dass das Optimierungsproblem nicht ausschließlich durch algebraische Ausdrücke und mathematische Gleichungen und Ungleichungen spezifiziert wird, sondern – zumindest teilweise – durch ein eigenständiges Simulationsmodell⁴¹. Im englischen Sprachraum sind auch die Bezeichnungen *Simulation Optimization* (Fu et al., 2005, OptTek Systems Inc., 2010) und *Optimization via Simulation (OvS)* (Hong und Nelson, 2009) gebräuchlich.

Der Lösungsansatz

Für die SBO wird das Simulationsmodell in einen numerischen Optimierungsprozess eingebettet. Das iterative Lösungsprinzip ist in Abbildung 4.2 dargestellt:

⁴¹ Driessen (2006: S.3) definiert simulationsbasierte Optimierung entsprechend als „*optimization involving computer simulations as function evaluations*“.

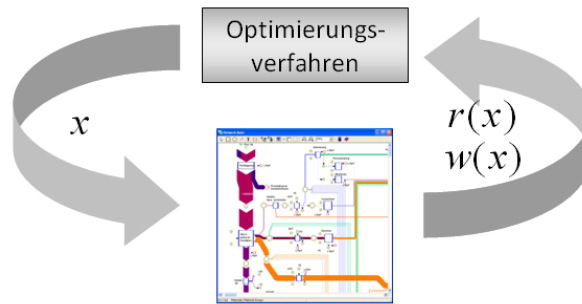


Abbildung 4.2: Prinzip der simulationsbasierten Optimierung am Beispiel eines eingebetteten Stoffstromnetzes

Ein Optimierungsverfahren gibt in der k -ten Iteration eine Lösung $\mathbf{x}^{(k)}$ vor. Die erklärenden Variablen des Simulationsmodells werden mit den entsprechenden Werten initialisiert und anschließend die abhängigen Ergebnisgrößen $\mathbf{r}(\mathbf{x}^{(k)})$ berechnet. Mit dieser *Modellrückgabe* wertet das Optimierungsverfahren die Zielfunktion und Restriktionen aus und ermittelt den nächsten Lösungsvorschlag $\mathbf{x}^{(k+1)}$. Auf diese Weise wird das Optimum schrittweise *approximiert*. Weil die innere Struktur des Simulationsmodells für den Optimierungsprozess belanglos ist, wird die simulationsbasierte Optimierung auch als *Black-Box-Optimierung* bezeichnet (Driessen, 2006: S.4, Syrjakow, 2003: S.13).

Formale Definition des simulationsbasierten Optimierungsproblems (SimOP)

Das simulationsbasierte Optimierungsproblem (SimOP) kann formal wie in (4.1) definiert werden. Es unterscheidet sich vom mathematischen Programm ausschließlich dadurch, dass die Zielgrößen $\varphi=(f, \mathbf{g}, \mathbf{h})$ explizit von den Berechnungsergebnissen $\mathbf{r}(\mathbf{x})$ eines *Simulationsmodells* abhängen.

$$\begin{aligned}
 & \min/\max \quad f^a(\mathbf{x}, \mathbf{r}) \\
 & \text{Nb.} \quad g_i^a(\mathbf{x}, \mathbf{r}) \leq 0 \quad i=1 \dots n \\
 & \quad \quad h_j^a(\mathbf{x}, \mathbf{r}) = 0 \quad j=1 \dots k \\
 & \quad \quad \mathbf{l} \leq \mathbf{x} \leq \mathbf{u}
 \end{aligned} \tag{4.1}$$

Anwendungsfelder

In der Praxis wird die SBO üblicherweise dann eingesetzt, wenn (i) bereits ein Simulationsmodell des geplanten Systems vorhanden ist und eine erneute Modellierung als mathematisches Programm mit einem zu großen Aufwand verbunden wäre oder (ii) das Originalsystem Eigenschaften aufweist, die sich nur durch eine *Simulation im engeren Sinne* abbilden lassen. Während im ersten Fall die mathematische Programmierung als Alternative zumindest theoretisch möglich ist, ist die simulationsbasierte Optimierung im zweiten Fall die einzige Option.

Simulationsmodelle (im engeren Sinne) ahmen das Verhalten dynamischer Systeme nach, d.h. der zeitliche Ablauf systeminterner Prozesse und Ereignisse wird im Computer nachvollzogen und die Veränderung wichtiger Systemeigenschaften durch entsprechende Modellvariablen simuliert. Wichtige modellbildende Elemente sind neben mathematischen Ausdrücken die Verwendung von *Rekursionen*, *Zählschleifen* und *Wenn-dann-Abfragen*, also zentraler Elemente aller imperativen Programmiersprachen. Zustandsänderungen des modellierten Systems ergeben sich *endogen* als Ergebnis einer Kausalkette. Diese Form der Abbildung ermöglicht auch die vergleichsweise einfache Integration *stochastischer Effekte* (Niemeyer, 1972: S.12).

Einen besonderen Stellenwert hat die simulationsbasierte Optimierung für die Planung von Produktionsabläufen erlangt. Typische Beispiele sind Maschinenbelegungspläne, die Auftragsterminierung oder die Organisation von Transportsystemen. Dafür müssen häufig komplex ineinander verschachtelte Lagerhaltungs-, Warteschlangen- und Reihenfolgenplanungsprobleme gelöst werden (Meyer, 1996: S.140). Häufig sind dabei Unsicherheiten wie Maschinenausfälle oder Lieferchwankungen zu berücksichtigen. In der betriebswirtschaftlichen Literatur zur simulationsbasierten Optimierung wird Simulation daher mit einer *stochastischen, ereignisdiskreten Simulation* gleichgesetzt (Fu, 2002: S.194, Hong und Nelson, 2009).

Doch auch im Methodenportfolio der Ingenieurs- und Naturwissenschaften hat die simulationsbasierte Optimierung ihren festen Platz gefunden. Hier geht es in der Regel um die kontinuierliche Simulation von *Differentialgleichungen* mit komplizierten Randbedingungen (Kolda et al., 2003: S.391). In der Automobil- oder Luftfahrtindustrie muss beispielsweise bei der Entwicklung neuer Fahrzeuge deren Luftwiderstand bestimmt und ihre Form entsprechend optimiert werden. Das ist nur möglich, indem der Luftwiderstand durch entsprechende Simulationsmodelle bestimmt wird (Driessen, 2006: S.4f)

Direkte Optimierungsverfahren zur Lösung des SimOP

Das Simulationsmodell stellt bei der Optimierung eine Black Box dar, die lediglich die Berechenbarkeit von $r(\mathbf{x})$ gewährleistet. Deshalb kommen analytische bzw. ableitungsbasierte Verfahren für die SBO nicht in Frage. Besser geeignet sind *direkte Suchstrategien*, die auf einer *punktuellen* Auswertung von Zielfunktion und Restriktionen beruhen. Hooke und Jeeves (1961: S.212) definieren direkte Suchstrategien als *“sequential examination of trial solutions involving comparison of each trial solution with the “best” obtained up to that time together with a strategy for determining (as a function of earlier results) what the next trial solution will be”*. Der bekannteste Algorithmus dieser Art ist zweifellos der Simplex-Algorithmus von Nelder und Mead (1965), der auch heute noch in vielen Anwendungen eingesetzt wird⁴². Die meisten direkten Optimierungsalgorithmen sind *heuristischer* Natur. Im Gegensatz zu den analytischen Optimierungsverfahren beruhen sie nicht auf speziellen mathematischen Eigenschaften eines bestimmten Optimierungsprob-

⁴² Der Nelder-Mead-Algorithmus ist beispielsweise im Computeralgebra-Paket Mathematica als optionale Methode zur nichtlinearen numerischen Optimierung implementiert (Wolfram Research Inc., 2010b).

lemtyps, sondern verwenden plausible, in der Regel in numerischen Experimenten überprüfte Ansätze, um eine optimale Lösung schrittweise anzunähern.

Anwendbarkeit auf das SSNOP

Durch die formale Definition des SimOP (4.1) wird offensichtlich, dass die SBO unmittelbar für die Lösung des SSNOP verwendet werden kann. Wenn die Berechnungsergebnisse des Stoffstromnetzes (\mathbf{r}^S) sowie der darauf aufbauenden Auswertungsmethoden (\mathbf{a}) zur Modellrückgabe \mathbf{r} zusammengefasst werden, geht das SSNOP (3.7, S.49) – bis auf die explizite Einbeziehung des Modellgültigkeitsbereichs in die Nebenbedingungen – direkt in das simulationsbasierte Optimierungsproblem (4.1) über.

4.1.3 Vergleich analytischer und direkter Optimierungsverfahren

Historische Entwicklung

Viele direkte Suchverfahren wurden, wie der Nelder-Mead-Algorithmus, in den Anfängen der computergestützten numerischen Optimierung in den 1950er und 60er Jahren entwickelt. Sie hatten gegenüber analytischen Verfahren den Vorteil, dass die rechenintensive und für damalige Computer anspruchsvolle Berechnung von Gradienten umgangen werden konnte. Nachdem die rasante Entwicklung der Informatik diesem Argument jedoch weitgehend die Grundlage entzogen hatte, verloren direkte Suchstrategien ab den 1970er Jahren im Bereich der mathematischen Optimierung stark an Bedeutung und wurden weitgehend durch die mathematische Programmierung verdrängt. Es wurde nun gemeinhin kritisiert, dass sich die Konvergenz direkter Suchverfahren nicht beweisen ließe und sie in der Regel auch langsamer als analytische Verfahren konvergierten. An dieser Einschätzung hat sich bis heute nicht viel geändert, obwohl es für viele direkte Suchalgorithmen durchaus möglich ist, rigorose Konvergenzbeweise aufzustellen (Kolda et al., 2003: S.387f).

Vorteile direkter Verfahren

In manchen Fällen sind direkte Suchstrategien den analytischen Optimierungsverfahren jedoch grundsätzlich überlegen. Weil sie nicht auf der Berechnung von Zielfunktionsgradienten beruhen, im engeren Sinne auch nicht auf einer numerischen Annäherung beispielsweise durch finite Differenzen, können direkte Optimierungsverfahren auch auf Optimierungsprobleme mit *nicht-stetigen* oder *stochastischen Zielfunktionen* angewendet werden. Trosset hat darauf hingewiesen, dass direkte Suchstrategien sich sogar für die Lösung von Optimierungsproblemen mit einer *ordinalen* Zielfunktion eignen. Er definiert direkte Suchverfahren daher als „*any algorithm that depends on the objective function only through the ranks of a countable set of function values.*“ Trosset (2000: S.5)

4.2 Algebraische Reformulierung des SSNOP

Die algebraische Reformulierung eines stoffstromnetzbasiereten Stoffstrommodells umfasst zwei Stufen⁴³. Zunächst muss die durch das Stoffstromnetz gegebene Funktion $r^S(x)$ durch algebraische Ausdrücke definiert werden. Sofern das Modell durch stoffstromnetzbasierete Auswertungsmethoden $a(r^S, x)$ berechnete Kennzahlen beinhaltet, müssen auch für diese algebraische Funktionsdefinitionen gefunden werden. In diesem Abschnitt werden Transformationsvorschriften für die modellbildenden Elemente von Stoffstromnetzen, sowie für stoffstromnetzspezifische Auswertungsmethoden hergeleitet, mit denen die automatische algebraische Reformulierung geeigneter SSNOPs als mathematische Programme möglich wird. Dieser Ansatz erlaubt zudem die Definition eindeutiger Kriterien für die algebraische Reformulierbarkeit eines SSNOP.

4.2.1 Transformation des Stoffstromnetzes

Für die algebraische Reformulierung von Stoffstromnetzen muss untersucht werden, wie Transitionen, Stellen und Flussrelationen transformiert werden können. Eine besondere Bedeutung kommt den Transitionen und Verbindungsstellen zu, weil diese die Informationen zur Berechnung unbekannter Daten, d.h. die eigentliche Funktionsspezifikation enthalten⁴⁴.

Die Flussrelationen eines Stoffstromnetzes repräsentieren *Stoffströme* innerhalb des Systems. In dieser Eigenschaft werden sie bei der algebraischen Reformulierung in numerische Entscheidungsvariablen überführt. Darüber hinaus erzeugen sie die *Grundstruktur* des Stoffstromnetzes, indem sie verschiedene Stellen und Transitionen in wohl definierter Weise miteinander verbinden. Diese Eigenschaft wird bei der algebraischen Reformulierung nur indirekt abgebildet: Bei der Reformulierung einzelner Netzknoten tauchen genau dann dieselben Flussvariablen auf, wenn diese im Stoffstromnetz durch eine Flussrelation verbunden sind.

4.2.1.1 Stellen und implizite Restriktionen bezüglich einzelner Modellvariablen

Die algebraische Reformulierung von Stellen erfolgt anhand der in Tabelle 3.1 (S.54) aufgeführten *algebraischen Ersatzformeln*. Wie in § 3.2.3 gezeigt wurde, spiegeln diese Formeln das an allen Stellen des Stoffstromnetzes geltende Prinzip der Massen- bzw. Energieerhaltung wieder. Die Stoffbestände $s_k^{(p)}$ von normalen Stellen und Grenzstellen müssen außerdem durch entsprechende Entscheidungsvariablen berücksichtigt werden.

Implizite Restriktionen, die sich aufgrund der Semantik von Stoffstromnetzen ergeben, können ohne Weiteres algebraisch reformuliert werden. Im Wesentlichen geht es hier um die Nichtnegativitätsbedingung für Stoffströme und Stoffbestände in normalen Stellen. Bei der algebraischen

⁴³ vgl. Abbildung 3.5 (S.50) zum hierarchischen Aufbau stoffstromnetzbasierter Stoffstrommodelle.

⁴⁴ vgl. hierzu das Flussdiagramm des Berechnungsalgorithmus für Stoffstromnetze bei Möller (2000: Abb.28, S.108).

Reformulierung müssen die entsprechenden Modellvariablen explizit mit der Untergrenze 0 versehen werden:

$$\begin{aligned} z_k^{(t,p)}, z_k^{(p,t)} > 0 & \quad \forall k \in M, (t,p), (t,p) \in F \\ s_k^{(p)} > 0 & \quad \forall k \in M, p \in P \end{aligned} \tag{4.2}$$

4.2.1.2 Transitionen

Bei der Transformation von Transitionen muss die Transitionsspezifikation algebraisch reformuliert werden. Daher wird im Folgenden zwischen den verschiedenen Spezifikationstypen unterschieden.

Linear spezifizierte Transitionen

Linear spezifizierte Transitionen sind formal äquivalent zu einem einfach unterbestimmten linearen Gleichungssystem (vgl. Abschnitt 2.3). Sie können daher problemlos gemäß Formel (2.8, S.35) in einen Satz von Gleichungsrestriktionen transformiert werden.

Benutzerdefinierte Funktionen

Die algebraische Reformulierung einer mittels *benutzerdefinierter Funktionen* spezifizierten Transition ist ebenfalls möglich. Die Funktionsdefinitionen können in Gleichungsrestriktionen überführt werden, indem der Zuweisungsoperator als Gleichheitszeichen interpretiert wird. Neben elementaren Funktionen und Grundrechenoperationen umfasst die Syntax benutzerdefinierter Funktionen auch *logische Ausdrücke* wie AND(), OR() oder IF(), mit der z.B. Fallunterscheidungen bei der Transitionsberechnung abgebildet werden können. Die algebraische Reformulierung solcher Ausdrücke erfordert entweder die Verwendung *nichtstetiger Funktionen* wie der Heaviside'schen Stufenfunktion⁴⁵ $\theta(x)$ oder die Einführung *binärer Entscheidungsvariablen*.

Die Funktion $\mathbf{IF}(x_1, x_2, x_3)$ hat beispielsweise drei reellwertige Argumente, von denen das erste das Entscheidungskriterium und die beiden anderen mögliche Rückgabewerte sind. Bei $x_1 \neq 0$ (wahr) liefert die Funktion den Wert von x_2 , bei $x_1=0$ (falsch) den Wert von x_3 zurück. Die algebraische Reformulierung lautet

$$\begin{aligned} \mathbf{IF}(x_1, x_2, x_3) &= f(x_1) \cdot x_2 + (1-f(x_1)) \cdot x_3; \quad x_i \in R, i=1,2,3 \\ \text{mit } f(x) &= 1 - \theta(x) \cdot \theta(-x) \end{aligned} \tag{4.3}$$

⁴⁵ Die Heaviside'sche Stufenfunktion hat den Wert 0 für alle negativen reellen Zahlen und den Wert 1 für alle positiven Zahlen einschließlich der 0 (Wolfram Research Inc., 2010c).

Wenn es der Kontext der Transitionsspezifikation erlaubt, kann x_i bei der Reformulierung als binäre Variable interpretiert werden. Dann vereinfacht sich (4.3) zu

$$\mathbf{IF}(x_1, x_2, x_3) = x_1 \cdot x_2 + (1 - x_1) \cdot x_3 \quad x_i \in R, i=2,3; x_1 \in \{0,1\} \quad (4.4)$$

Die algebraische Reformulierung durch benutzerdefinierte Funktionen spezifizierter Transitionen mit logischen Ausdrücken führt daher i.A. entweder auf gemischt-ganzzahlige mathematische Programme oder auf kontinuierliche Optimierungsprobleme, die Unstetigkeiten bezüglich der Zielfunktion oder einzelner Restriktionen enthalten.

Algorithmische Prozessspezifikation durch Skripte

Die algebraische Reformulierung algorithmisch spezifizierter Transitionen führt auf ein grundsätzliches Problem. Die Berechnungsalgorithmen werden in *imperativen* Skript- bzw. Programmiersprachen formuliert⁴⁶. Sie bestehen somit aus einer Reihe von Anweisungen, die streng *sequentiell* abgearbeitet werden⁴⁷. Die schrittweise Berechnungsvorschrift kann durch *Kontrollstrukturen* wie Schleifen weiter strukturiert oder Anweisungsblöcke zu *Prozeduren* oder Funktionen zusammengefasst werden. Modellvariablen sind in der Regel dynamisch, d.h. ihr Wert kann sich im Verlauf der Transitionsberechnung verändern.

Durch diese Eigenschaften unterscheiden sich algorithmisch spezifizierte Transitionen fundamental von *deklarativen* mathematischen Programmen, bei denen „die Lösung eines Problems auf einer hohen Abstraktionsebene spezifiziert und nicht auf einer niedrigen Maschinenebene ‚ausprogrammiert‘ wird“ (Scholl, 2006: S.8)⁴⁸. Die Gleichungen und Ungleichungen mathematischer Programme können als ein System logischer Aussagen aufgefasst werden, die alle simultan erfüllt sein müssen. Die *Reihenfolge* dieser Aussagen ist unerheblich. Außerdem sind alle Modellvariablen *statisch*, d.h. sie stehen für einen unveränderlichen Wert oder Wertebereich.

Die Inkompatibilität zwischen dem imperativen und dem deklarativen Programmierparadigma ist somit die *Ursache* für die Unmöglichkeit, algorithmisch spezifizierte Transitionen in entsprechende Fragmente eines mathematischen Programms zu überführen. Ein dynamisches Berechnungsverfahren kann nicht ohne Weiteres in eine statische deklarative Problemspezifikation überführt werden. Es ist im Einzelfall zwar denkbar, dass für eine algorithmische Spezifikation eine deklarative

⁴⁶ z.B. JScript, VBScript oder Python (Möller, 2000: S.104)

⁴⁷ Ein *Algorithmus* ist „eine endliche Beschreibung eines schrittweisen Problemlösungsverfahrens zur Ermittlung gesuchter Datenobjekte (ihrer Werte) aus gegebenen Werten von Datenobjekten, in dem jeder Schritt aus einer Anzahl ausführbarer, eindeutiger Aktionen und einer Angabe über den nächsten Schritt besteht“ (Pomberger und Dobler, 2008: S.33).

⁴⁸ Das Zitat von Scholl bezieht sich allgemein auf deklarative Programme und schließt den Spezialfall des mathematischen Programms daher natürlich ein. Zum Unterschied zwischen imperativen und deklarativen Programmiersprachen vergleiche auch Henning und Vogelsang (2007: S.27f).

tive Ersatzformulierung existiert. Diese kann aber nur von einem mathematisch versierten Modellierer gefunden werden, der außerdem noch über eine hinreichend gute Kenntnis des modellierten Prozesses verfügt. Im Allgemeinen ist es jedoch unmöglich, imperative in deklarative Programme zu überführen, ohne dabei weitere Vereinfachungen des Modells vorzunehmen.

Einbindung externer Softwarekomponenten zur Transitionsberechnung

Transitionen, zu deren Berechnung auf externe Softwarekomponenten zurückgegriffen wird, basieren wie algorithmisch spezifizierte Transitionen auf Skripten, können im Hinblick auf die algebraische Formulierung jedoch als eigene Kategorie gesehen werden. Während ein Berechnungsalgorithmus grundsätzlich nachvollziehbar ist, stellt eine externe Softwarekomponente offensichtlich eine Black Box dar, die im Regelfall nicht algebraisch reformuliert werden kann.

4.2.1.3 Algebraische Reformulierbarkeit von Stoffstromnetzen

Um ein Stoffstromnetz vollständig in ein mathematisches Programm zu überführen, müssen *alle* darin enthaltenen Stellen und Transitionen einzeln algebraisch reformuliert werden. Da alle Stellentypen algebraisch reformulierbar sind, hängt es allein von den Eigenschaften der Transitionen ab, ob das Stoffstromnetz in seiner Gesamtheit reformulierbar ist (Tabelle 4.1). Die algebraische Reformulierbarkeit eines Stoffstromnetzes kann demnach anhand eines Ausschlusskriteriums festgelegt werden: Ein Stoffstromnetz ist genau dann algebraisch reformulierbar, wenn es keine durch Berechnungsalgorithmen oder externe Softwarekomponenten spezifizierte Transitionen enthält.

Tabelle 4.1: Reformulierbarkeit verschiedener Stoffstromnetztypen. Diese werden differenziert nach der Art der verwendeten Transitionen. In der letzten Spalte wird für die reformulierbaren SSNOP angegeben, zu welchem Typ mathematischen Programm sie führen.

Spezifikationstyp	reformulierbar	Problemtyp
linear	ja	LP
benutzerdefinierte Funktionen		
– nur lineare Ausdrücke und Funktionen		LP
– mind. ein nichtlinearer Ausdruck		NLP
– logische Funktionen		[M]ILP, [M]INLP
Skripte (Algorithmen)	nein	-
Einbindung ext. Anwendung		

4.2.2 Beispielhafte Reformulierung des Verbundproduktionsmodells aus Abschnitt 3.4

Zur Veranschaulichung der vorangegangenen Überlegungen wird nun gezeigt, wie das 7D-Modell der Verbundproduktion aus Abschnitt 3.4 algebraisch reformuliert werden kann. An dieser Stelle wird exemplarisch auf die Transformation einer Transition und einer Stelle eingegangen. Der in der algebraischen Modellierungssprache LINGO abgefasste Code des vollständig transformierten Modells findet sich in Anhang A.3.

Aus Abbildung 4.3 geht hervor, wie die Flussrelationen des Stoffstromnetzes in Entscheidungsvariablen des mathematischen Programms überführt werden. Die Konstruktion der Variablennamen entspricht der in dieser Arbeit verwendeten Notation. Je nachdem, ob es sich um Inputs oder Outputs einer Transition handelt, beginnen die Variablennamen entweder mit x oder mit y . Danach folgen die Bezeichnung der Stoffart und schließlich die Nummer der betreffenden Transition. $xW1$ bedeutet demnach einen Input des Stoffes W in $T1$. Wenn der Stoffstrombezeichner keine Transitionsnummer beinhaltet, handelt es sich um Input- bzw. Outputströme des Gesamtnetzes.

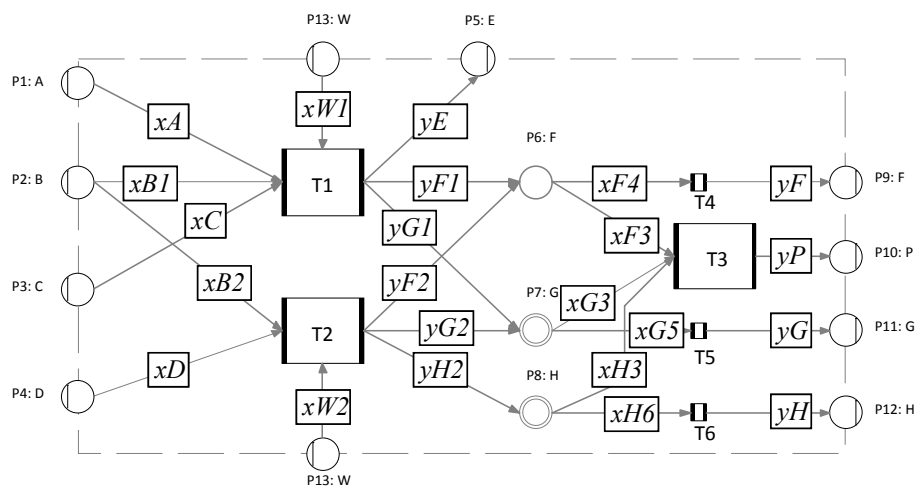
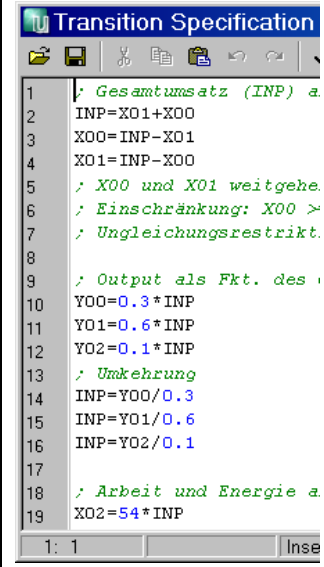


Abbildung 4.3: Zuordnung der Flussrelationen des Stoffstromnetzes der Verbundproduktion (7D-Modell) zu den Entscheidungsvariablen des reformulierten algebraischen Modells.

Tabelle 4.2 enthält die algebraische Reformulierung der Transitionsspezifikation T2 des 7D-Modells der chemischen Verbundproduktion aus Abschnitt 3.4 (vgl. Tabelle 3.7, S.82). Die Zeilen 2-4 der Transitionsspezifikation enthalten verschiedene Umformungen derselben Gleichung. Weil es sich jedoch nicht um mathematische Gleichungen, sondern um Zuweisungsoperationen

handelt, sind die Zeilen im Sinne der Transitionsspezifikation nicht äquivalent⁴⁹. Bei der algebraischen Reformulierung können sie jedoch zu einer einzigen Restriktion [Umsatz_T2] zusammengefasst werden. Die Zeilen 10-12 werden jeweils in die Restriktionen [yF_T2],[yG_T2] und [yH_T2] transformiert. Die Zeilen 14-16 enthalten wiederum Umformungen der Zeilen 10-12 und müssen daher nicht reformuliert werden.

Tabelle 4.2: Algebraische Reformulierung der Transitionsspezifikation von T2 des Stoffstromnetzes (7D-Modell) der chemischen Verbundproduktion. Die Spalten enthalten (v.l.n.r.) die Transitionsspezifikation in der Syntax der benutzerdefinierten Funktionen von Umberto®, die algebraische Reformulierung in LINGO sowie die Zuordnung zwischen den verschiedenen Notationen von Umberto®, LINGO und den in dieser Arbeit üblichen Stoffstrombezeichnern.

Benutzerdefinierte Funktionen	LINGO – Code (algebraische Modellersprache)	Notationen		
 <pre> 1 / Gesamtumsatz (INP) a 2 INP=X01+X00 3 X00=INP-X01 4 X01=INP-X00 5 / X00 und X01 weitgehe. 6 / Einschränkung: X00 > 7 / Ungleichungsrestrikt. 8 9 / Output als Fkt. des 10 Y00=0.3*INP 11 Y01=0.6*INP 12 Y02=0.1*INP 13 / Umkehrung 14 INP=Y00/0.3 15 INP=Y01/0.6 16 INP=Y02/0.1 17 18 / Arbeit und Energie a 19 X02=54*INP </pre>	<pre> !*** Transition T2 ***; [Umsatz_T2] x2=xd+xb2; [yF_T2] yf2=0.3*x2; [yG_T2] yg2=0.6*x2; [yH_T2] yh2=0.1*x2; [xW_T2] xw2 = 54*x2; [xB_T2] xb2 >= 0.2*x2; </pre>	X00	x_{B2}	$x_B^{(T2)}$
		X01	x_D	$x_D^{(T2)}$
		X02	x_{W2}	$x_W^{(T2)}$
		Y00	y_{F2}	$y_F^{(T2)}$
		Y01	y_{G2}	$y_G^{(T2)}$
		Y02	y_{H2}	$y_H^{(T2)}$
		INP	x_2	-

Der zum gesamten Masseumsatz proportionale Energieaufwand des Prozesses (Zeile 19) wird in die Restriktion [xW_T2] überführt. Außerdem enthält der LINGO-Code auch die Ungleichheitsrestriktion [xB_T2], die entsprechend Formel (3.25) festlegt, dass bezogen auf die gesamte in T2 umgesetzte Produktmasse mindestens 20% des Stoffes B eingesetzt werden müssen.

⁴⁹ Solche im mathematischen Sinne redundanten Umformungen sind bei der Transitionsspezifikation mit benutzerdefinierten Funktionen durchaus sinnvoll, weil sie verschiedene Rechenrichtungen für die Auswertung der Transition ermöglichen: in Zeile 2 wird beispielsweise der gesamte Masseumsatz von T2 (INP) als Funktion der beiden Inputströme X00 und X01 berechnet. Ist nun der Stoffstrom X00 unbekannt, so kann er mit Zeile 3 berechnet werden, vorausgesetzt, INP konnte zuvor beispielsweise mit einer der Zuweisungen in den Zeilen 14-16 bestimmt werden.

Als Beispiel für die algebraische Reformulierung einer Stelle, soll hier das Beispiel der kombinierten Sammler- und Verteilerstelle P7 dienen (4.3). Sie ergibt sich unmittelbar aus der algebraischen Ersatzformel (3.10, S.53) für Verbindungsstellen⁵⁰.

$$[G_P7] \quad y_{g1} + y_{g2} = y_g + x_{g3}; \quad (4.5)$$

Dass in den reformulierten Gleichungsrestriktionen [yG_T2] und [G_P7] jeweils die Stoffstromvariable y_{G2} auftaucht, entspricht der Tatsache, dass beide Knoten im Stoffstromnetz durch eine Flussrelation miteinander verbunden sind.

4.2.3 Transformation stoffstromnetzbezogener Auswertungsmethoden

Stoffstromnetzbasierende Auswertungsmethoden ermöglichen, ausgehend von einem vollständig spezifizierten Stoffstromnetz, die Berechnung verschiedener Kennzahlen zur Charakterisierung des Systemzustands. Sofern das Stoffstromnetz selbst algebraisch reformulierbar ist, stellt auch die Transformation der meisten mit stoffstromnetzbasierenden Auswertungsmethoden erzeugten Kennzahlen kein grundsätzliches Problem dar.

Input-Output-Bilanzen

Input-Outputbilanzen weisen alle Stoffströme aus, die eine bestimmte Bilanzgrenze schneiden. Das kann die Modellgrenze sein oder eine beliebige im Inneren des Modells verlaufende Bilanzgrenze. Für die Aufstellung einer Input-Output-Bilanz müssen (i) Stoffströme ausgewählt und ggf. (ii) Stoffströme gleicher Art summiert werden, beispielsweise der Wasserverbrauch oder die CO₂-Emissionen in verschiedenen Prozessen (Transitionen) der betrieblichen Produktion. Die Einzelposten einer Input-Output-Bilanz sind daher entweder Stoffstromvariablen des Modells oder werden durch einfache Summenbildung aus diesen abgeleitet. Die algebraische Reformulierbarkeit ist offensichtlich möglich.

Zustandsbezogene Kennzahlen

Darüber hinaus bieten Entwicklungsumgebungen für stoffstromnetzbasierende Stoffstrommodelle auch die Möglichkeit, Modellparameter und –variablen zu beliebigen Kennzahlen für den modellierten Systemzustand zu aggregieren⁵¹. Dafür werden typischerweise algebraische Ausdrücke, insbesondere Summen und Produkte, verwendet, die sich direkt in mathematische Programme integrieren lassen.

⁵⁰ Auf der rechten Seite von (4.3) steht y_g anstatt x_{G5} , weil bei der algebraischen Reformulierung des Stoffstromnetzes die Transitionsspezifikation von T5 nicht transformiert wurde (vgl. Anhang A.3).

⁵¹ Bei der Stoffstromanalysesoftware Umberto® übernimmt diese Funktion der so genannte Kennzahleneditor.

Leistungsverrechnung

Die Leistungsverrechnung ermöglicht eine verursachungsgerechte Zuordnung von Aufwendungen und Erträgen eines als Stoffstromnetz abgebildeten Produktionssystems. Die Leistungsverrechnung ist damit sowohl der Kern einer stoffstrombasierten *Kostenträgerrechnung* als auch das Bindeglied zwischen der periodenbezogenen Stoffstromanalyse eines Betriebs und den *Produktökobilanzen* seiner verschiedenen Erzeugnisse. Für die Leistungsverrechnung muss das vollständig berechnete Stoffstromnetz in einen linearen Aufwands-Ertragsgraphen überführt werden. Anschließend wird durch Inversion der entsprechenden Adjazenzmatrix die Zuordnung der Systemaufwendungen zu den verschiedenen Systemerträgen (Referenzflüssen) vorgenommen. Diese Berechnungsschritte der Leistungsverrechnung sind nicht ohne Weiteres algebraisch reformulierbar. Sie sind jedoch unverzichtbar, wenn auf einzelne Referenzflüsse bezogene Kennzahlen optimiert werden sollen.

Es stellt sich jedoch die Frage, ob die Optimierung referenzflussbezogener Kennzahlen überhaupt sinnvoll ist. Bei der Leistungsverrechnung wird das ursprüngliche Produktionssystem künstlich⁵² in Teilsysteme zerlegt, die den verschiedenen Referenzflüssen (Kostenträgern) des Systems zugeordnet werden. Bei der Optimierung einer referenzflussbezogenen Zielgröße besteht daher immer die Gefahr, dass das ausgewählte Teilsystem zu Ungunsten des Gesamtsystems optimiert wird. Insbesondere können Synergien, die sich bei der realen Herstellung verschiedener Produkte ergeben, bei der Optimierung nicht ausgenutzt werden. Wird eine einzelproduktbezogene Optimierung dennoch gewünscht, ist ein zweistufiges Verfahren denkbar. Zunächst wird durch die Leistungsverrechnung das Teilsystem separiert und anschließend nur dieses Teilsystem algebraisch reformuliert. Damit kann die Transformation der Leistungsverrechnungsalgorithmen in einer problemadäquaten Weise umgangen werden.

4.3 Konzeption eines Optimierungsmoduls für eine stoffstromnetzbasierete Stoffstromanalysesoftware

Aufbauend auf den vorangegangenen Überlegungen zur simulationsbasierten und mathematischen Optimierung des SSNOPs, wird in diesem Abschnitt das Konzept eines Optimierungsmoduls für eine stoffstromnetzbasierete Stoffstromanalysesoftware skizziert. Ein solches Optimierungsmodul muss drei verschiedene Aufgabenbereiche abdecken, die nachfolgend in der Reihenfolge aufgezählt sind, in der sie bei einer stoffstromnetzbasiereten Optimalplanung notwendig werden.

(M) *Modellbildung*: Ergänzung des Stoffstrommodells zum Optimierungsmodell (SSNOP)

⁵² Das wird insbesondere im Falle von Kuppelproduktionsprozessen deutlich. Die Zwangsläufigkeit der Herstellung mehrerer Kuppelprodukte wird durch Allokationsregeln künstlich aufgehoben und nur der Teil des Prozessaufwands berücksichtigt, der aufgrund einer subjektiven Einschätzung dem Teil des Prozessoutputs zugeordnet wird, der in die Produktion des Referenzflusses eingeht.

(L) Lösung des SSNOP

(A) Analyse der Optimierungsergebnisse

Den Überlegungen von Abschnitt 3.1 entsprechend besteht die Ergänzung des Stoffstrommodells zum SSNOP aus folgenden Schritten:

(M1) Entscheidungsvariablen aus der Menge der erklärenden Modellvariablen auswählen

(M2) die Zielfunktion $f(\mathbf{x}, \mathbf{r})$ konstruieren

(M3) Ober- und Untergrenzen für einzelne Modellvariablen festlegen (optional)

(M4) Algebraische Restriktionen $g_i(\mathbf{x}, \mathbf{r})$ und $h_j(\mathbf{x}, \mathbf{r})$ formulieren (optional)

Die Anforderungen, die sich hinsichtlich der Lösung (L) des SSNOPs ergeben, werden in den nachfolgenden Abschnitten ausführlicher diskutiert. Der letzte Schritt einer stoffstromnetzbasieren Optimalplanung besteht in der Analyse (A) der Optimierungsergebnisse. Dazu zählen die numerische Ausgabe gefundener Lösungen, der Datenexport, Sensitivitätsanalysen oder die Visualisierungen von Optimierungsergebnissen. Darauf wird im Rahmen dieser Arbeit im Detail nicht weiter eingegangen.

4.3.1 Kombination von simulationsbasierter und mathematischer Optimierung für die Lösung des SSNOPs

Für die Lösung des SSNOP kommen, wie die Überlegungen der vorigen Abschnitte gezeigt haben, je nach den Eigenschaften des Stoffstromnetzes sowohl die mathematische Optimierung (MO) als auch die simulationsbasierte Optimierung (SBO) in Frage. Die SBO ist der MO insofern überlegen, als sie die Optimierung beliebiger SSNOPs ermöglicht. In den Fällen, in denen die algebraische Reformulierung des SSNOPs möglich und die MO daher anwendbar ist, führt sie in der Regel dagegen schneller zu besseren Lösungen. Im Idealfall sollten daher beide Ansätze im Optimierungsmodul integriert werden.

Das war bei der Entwicklung des prototypischen Optimierungsmoduls im Forschungsprojekt KOMSA aus Zeitgründen nicht möglich, weshalb aufgrund der breiteren Anwendbarkeit zunächst der simulationsbasierte Lösungsansatz implementiert wurde. In Kap. 5 werden direkte Optimierungsalgorithmen für die Optimierung des stoffstromnetzbasieren Optimierungsproblems vorgestellt und hinsichtlich ihrer Leistungsfähigkeit getestet.

Das in dieser Arbeit entworfene Optimierungsmodul soll dagegen beide Ansätze unterstützen. Weil sich MO und SBO nur bei der Lösung (L) des SSNOP unterscheiden, wird bei der Entwicklung einer Softwarearchitektur für das Optimierungsmodell folgendermaßen vorgegangen: In § 4.3.2 wird zunächst ein Optimierungsmodul für die SBO entworfen (Lambrecht und Schmidt, 2010: S.612). In § 0 wird gezeigt, wie dessen Architektur modifiziert werden muss, um zusätzlich die mathematische Optimierung zu ermöglichen.

4.3.2 Simulationsbasierte Optimierung

Für die simulationsbasierte Optimierung des SSNOP muss das Optimierungsmodul sowohl die Initialisierung als auch die Kontrolle des iterativen Optimierungsprozesses ermöglichen:

- (L1) *Initialisierung*: Auswahl eines geeigneten Optimierungsalgorithmus sowie die Einstellung algorithmusspezifischer Parameter.
- (L2) *Kontrolle*: Auswertung der Zielfunktion, Prüfung von Restriktionen sowie das Protokollieren relevanter Informationen zum Optimierungsverlauf.

Um seine Funktion zu erfüllen, benötigt das Optimierungsmodell insgesamt drei Schnittstellen (Abbildung 4.4). Während des Optimierungsverlaufs (L2) vermittelt das Modul *automatisch* zwischen dem Stoffstrommodell und dem verwendeten Optimierungsalgorithmus. Dafür benötigt es einerseits eine bidirektionale Schnittstelle zum Stoffstrommodell, um Modellberechnungen zu initiieren und die Ergebnisse auszulesen. Andererseits ist auch eine bidirektionale Schnittstelle zum Optimierungsalgorithmus notwendig, um die Ergebnisse der Auswertungen von Zielfunktion und Restriktionen zu übermitteln und neue Lösungsvorschläge vom Algorithmus zu empfangen.

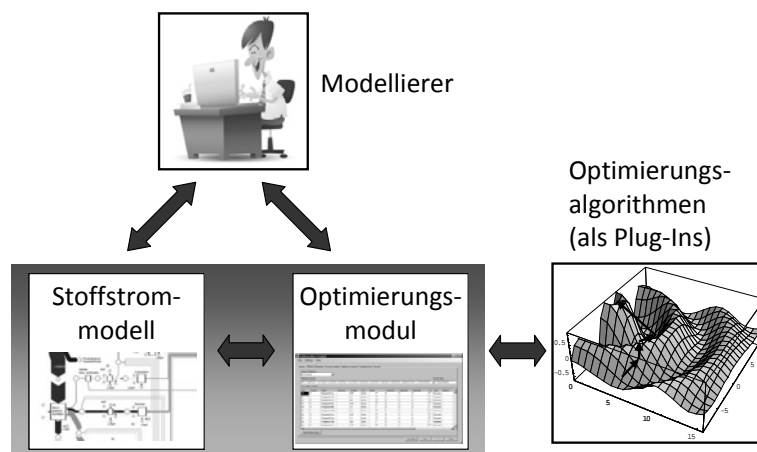


Abbildung 4.4: Schnittstellenarchitektur eines Optimierungsmoduls für eine stoffstromnetzbasierter Stoffstromanalyse-Software. Das SSNOP umfasst das Stoffstrommodell und die mit dem Optimierungsmodul durchgeführten zusätzlichen Spezifikationen (M1-M4). Es wird durch den grauen Kasten um diese beiden Komponenten dargestellt.

Damit das Optimierungsmodul als Werkzeug zur Modellbildung (M), für Optimierungsexperimente (L1) und die Ergebnisinterpretation (A) eingesetzt werden kann, benötigt es eine ebenfalls bidirektionale Benutzerschnittstelle: Die Ergänzung des Stoffstrommodells zum SSNOP (M) sowie die Initialisierung des Optimierungsprozesses (L1) erfordern Eingaben des Anwenders. Im Anschluss an die numerische Optimierung müssen dann protokollierte Daten zum Optimierungsverlauf und insbesondere die Optimierungsergebnisse dem Benutzer zur Verfügung gestellt werden.

Im Forschungsprojekt KOMSA wurde nach diesem Vorbild ein prototypisches Optimierungsmodul als Plug-In für die Stoffstromanalyse-Software Umberto® entwickelt⁵³.

Konkretisierung des allgemeinen SSNOP im Hinblick auf die softwaretechnische Umsetzung

Für die softwaretechnische Umsetzung der SBO von Stoffstrommodellen ist es zweckmäßig, die allgemeine Formulierung des SSNOP (3.7, S.49) in zweierlei Hinsicht zu konkretisieren (4.6)⁵⁴. Zum Einen wird in (4.6) zwischen *expliziten Restriktionen* (c, d) und *impliziten Restriktionen* (b, e) unterschieden. Explizite Restriktionen haben im Hinblick auf die SBO den Vorteil, dass sie ohne die zeitaufwändige Berechnung des Stoffstromnetzes ausgewertet werden können. Sie werden daher vor der Modellberechnung geprüft und können letztere somit überflüssig machen. Besteht bei der Modellierung die Wahl, so ist stets die explizite einer impliziten Formulierung von Restriktionen vorzuziehen.

$$\begin{array}{llll}
 \text{min/max} & f(x, r(x)) & x \in R^u, r \in R^{v+w} & \text{(a)} \\
 \text{Nb.} & g_i(x, r(x)) \begin{cases} \geq \\ = \\ \leq \end{cases} 0 & i = 1 \dots m & \text{(b)} \\
 & h_j(x) \begin{cases} \geq \\ = \\ \leq \end{cases} 0 & j = 1 \dots n & \text{(c) (4.6)} \\
 & l_i^x \leq x_i \leq u_i^x & i = 1 \dots k & \text{(d)} \\
 & l_i^r \leq r_i(x) \leq u_i^r & i = 1 \dots l & \text{(e)} \\
 & w(x) = \{ \} & & \text{(f)}
 \end{array}$$

Zum Anderen beinhaltet die *Textrückgabe* $w(x)$ des Modells (f) Warnungsmeldungen, die bei der Berechnung des Stoffstromnetzes auftreten und unter Umständen darauf hinweisen, dass die vom Optimierungsalgorithmus vorgegebene Lösung $\mathbf{x}^{(k)}$ nicht im *Modellgültigkeitsbereich* liegt⁵⁵. Das muss bei der Konzeption eines Optimierungsmoduls und bei der Auswahl geeigneter Optimierungsstrategien berücksichtigt werden: Bei entsprechenden Warnungsmeldungen kann die numerische Modellrückgabe *unvollständig* oder *falsch* sein. Dann wird die Zielfunktion unter

⁵³ Die Details der softwaretechnischen Umsetzung sind bei Zimmermann et al. (2009) dokumentiert.

⁵⁴ (4.6) orientiert sich weitgehend an der Definition des simulationsbasierten Optimierungsproblems bei Driessen (2006: S.9).

⁵⁵ Restriktion (4.6 f) entspricht damit der Restriktion (3.7 b) des allgemeinen SSNOP.

Umständen nicht oder nur fehlerhaft berechnet und darf nicht in den Optimierungsprozess einbezogen werden. Gleiches gilt für die impliziten Restriktionen. Für die Optimierung kommen daher nur Algorithmen in Frage, die außerhalb des zulässigen Bereiches nicht auf Zielfunktionsauswertungen angewiesen sind (vgl. Peters, 2001: S.40).

4.3.3 Mathematische Optimierung

Die mathematische Optimierung kann problemlos in die im vorigen Abschnitt entwickelte Schnittstellen-Architektur integriert werden. Dazu muss die notwendige Funktionalität für die algebraische Reformulierung sowie den Datenaustausch mit einer Software für die mathematische Programmierung ergänzt werden. Der Unterschied zur SBO besteht in der Ausgestaltung des Lösungsschritts (L):

- (L'1) Algebraische Reformulierung des SSNOP (Abschnitt 4.1.3)
- (L'2) Auswahl eines geeigneten Solvers
- (L'3) Export des mathematischen Programms und externe Optimierung
- (L'4) Reimport der Optimierungsergebnisse

Für die Auswahl eines Solvers gibt es – abhängig von der verwendeten Optimierungssoftware – zwei mögliche Varianten. Entweder wird der Solver den Eigenschaften des mathematischen Programms entsprechend (LP, NLP, ILP etc.) automatisch gewählt⁵⁶. Bei den meisten Entwicklungsumgebungen für mathematische Programme legt jedoch der Modellierer fest, welcher Solver für die Lösung des Optimierungsproblems verwendet wird⁵⁷. Weil die stoffstromnetzbasierter Optimalplanung grundsätzlich den Anspruch erhebt, den Anwender mit möglichst wenigen Details bezüglich des Optimierungsprozesses zu konfrontieren, sollte bei der Verwendung einer Optimierungssoftware ohne automatische Problemklassifizierung diese in (L'2) vorgenommen werden.

4.4 Zusammenfassung

Gegenstand von Kapitel 4 war die Entwicklung von Lösungsstrategien für das SSNOP, sowie deren Überführung in ein Softwarekonzept für eine stoffstromnetzbasierter Optimalplanung. In Abschnitt 4.1 wurden mit der mathematischen und der simulationsbasierten Optimierung zwei mögliche Lösungsansätze für das SSNOP vorgestellt. Während die mathematische Optimierung eine Problemformulierung als mathematisches Programm und damit eine entsprechende Transformation des SSNOPs voraussetzt, erfordern die in der simulationsbasierten Optimierung typischerweise eingesetzten direkten Suchstrategien lediglich die Berechenbarkeit von Zielfunktio-

⁵⁶ z.B. bei der Software LINGO (LINDO, 2004), die für die algebraische Reformulierung des Verbundproduktionsbeispiels in § 4.2.2 verwendet wurde.

⁵⁷ z.B. bei GAMS (Brooke et al., 2010).

nen und Restriktionen eines Parameteroptimierungsproblems. Daher können Stoffstromnetze *unverändert* in den simulationsbasierten Optimierungsprozess integriert werden. In der Regel gilt jedoch: je mehr analytische Informationen über ein Problem ein Optimierungsverfahren ausnutzen kann, desto genauer und schneller kann das Optimum bestimmt werden. Genau hier liegen die Stärken der in der mathematischen Programmierung eingesetzten analytischen Verfahren, die bei entsprechend günstigen Eigenschaften des Optimierungsproblems in der Lage sind, globale Optima exakt zu bestimmen. Hinsichtlich der Leistungsfähigkeit einsetzbarer Optimierungsalgorithmen ist die mathematische Programmierung in der Regel der simulationsbasierten Optimierung vorzuziehen.

Daher wurde im darauffolgenden Abschnitt 4.2 der Frage nachgegangen, ob und wie SSNOPs in mathematische Programme überführt werden können. Dazu wurden konkrete Transformationsvorschriften für die modellbildenden Elemente der Stoffstromnetze – Stellen, Transitionen und Flussrelationen – entwickelt. Es zeigte sich, dass eine algebraische Reformulierung aller Arten von Stellen, sowie der linear und durch benutzerdefinierte Funktionen spezifizierten Transitionen möglich ist. Hinsichtlich der algebraischen Reformulierung stoffstromnetzbezogener Auswertungsmethoden konnte weiterhin gezeigt werden, dass sowohl Input-Output-Bilanzen als auch stoffstromnetzbasierete Kennzahlen ohne Weiteres transformierbar sind. Lediglich die Leistungsverrechnung mit ihren Algorithmen für die Erzeugung linearer Aufwands-Ertragsgraphen sowie der anschließenden Matrixinversion stellt eine Schwierigkeit dar. In diesem Zusammenhang wurde jedoch die Überlegung angestellt, dass eine Leistungsverrechnung und die damit verbundene Ausblendung der Wechselwirkung verschiedener Produktlinien eines Betriebes Freiheitsgrade reduziert, und damit im Rahmen einer Optimalplanung nicht unbedingt sinnvoll ist. Des Weiteren konnte gezeigt werden, dass, falls sie dennoch gewünscht ist, die Reihenfolge von Leistungsverrechnung und Optimierung vertauscht, und damit die Notwendigkeit einer algebraischen Reformulierung umgangen werden kann.

Weil die Vorteile algorithmisch spezifizierter Transitionen hinsichtlich der Modellbildung nicht unmittelbar gegen diejenigen deklarativer Spezifikationsformen hinsichtlich der Lösbarkeit aufgewogen werden können, wurde in Abschnitt 4.3 ein Optimierungsmodul konzipiert, welches sowohl die simulationsbasierte Optimierung, als auch die algebraische Reformulierung von Stoffstromnetzen in Verbindung mit einer anschließenden mathematische Optimierung ermöglicht.

5 Leistungsfähigkeit direkter Suchstrategien für die stoffstromnetzbasierende Optimierung

Das im vorigen Kapitel entworfene Optimierungsmodul zur Lösung stoffstromnetzbasierter Optimierungsprobleme (SSNOP) gibt einen konzeptionellen Rahmen vor, in den *verschiedene* Lösungsalgorithmen integriert werden können. Die meisten Algorithmen können zudem über spezifische Parameter wie Populationsgrößen, Expansionskoeffizienten etc. in ihrem Lösungsverhalten beeinflusst werden. Bei der stoffstromnetzbasierenden Optimierung stellt sich daher grundsätzlich die Frage, (i) welcher Algorithmus (ii) mit welcher Parametrisierung für die Lösung eines konkreten SSNOPs am besten geeignet ist.

Vor diesem Hintergrund werden in diesem Kapitel unterschiedliche Optimierungsalgorithmen auf ihre Leistungsfähigkeit bei der Lösung von SSNOPs hin getestet⁵⁸. Dafür wird das im Forschungsprojekt KOMSA entwickelte Softwarepaket verwendet: die Stoffstromanalyse-Software Umberto[®] zusammen mit einem eigens dafür entwickelten Optimierungsmodul, dem so genannten Optimierungsleitstand⁵⁹, welcher für die simulationsbasierte Optimierung drei verschiedene direkte Suchstrategien als Plug-Ins zur Verfügung stellt (Lambrecht und Zimmermann, 2008: S.315-317).

Die besondere Schwierigkeit bei der Beurteilung von Optimierungsverfahren liegt darin, dass ihre Leistungsfähigkeit nie absolut beurteilt werden kann, sondern immer nur bezüglich eines bestimmten Optimierungsproblems bzw. Problemtyps. Daher hat sich in der bis ins 18. Jh. zurückreichenden Geschichte der Optimierungstheorie auch nie eine *universelle Lösungsstrategie* durchsetzen können (Schwefel, 1995: S.165). Folglich werden in diesem Kapitel keine Algorithmen an sich bewertet, sondern *numerische Optimierungen*, d.h. die Anwendung *bestimmter* Algorithmen mit *definierten* Parametereinstellungen auf *konkrete* Optimierungsprobleme. Durch die Bewertung *vieler* Optimierungsläufe für *verschiedene* Testprobleme werden in einer experimentellen Vorgehensweise und mit statistischen Gütekriterien Rückschlüsse auf die Algorithmen selbst gezogen (Rardin und Uzsoy, 2001).

⁵⁸ Dieses Kapitel ist eine grundlegend überarbeitete Fassung des Beitrags von Lambrecht (2009).

⁵⁹ Der KOMSA-Optimierungsleitstand entspricht, bis auf die Möglichkeit zur Reformulierung und mathematischen Optimierung von SSNOPs, dem in Abschnitt 4.3 skizzierten Optimierungsmodul.

5.1 Visualisierung und Leistungsfähigkeitsbewertung numerischer Optimierungsalgorithmen

Dieser Abschnitt befasst sich zunächst mit geeigneten Darstellungsformen für die Visualisierung exemplarischer Optimierungsverläufe. Danach werden die Optimierungsdauer und die Approximationsgenauigkeit als quantitative Performancekriterien für einzelne Optimierungen eingeführt. Darauf aufbauend können statistische Kennzahlen entwickelt werden, die zur Bewertung von Testreihen mit vielen Optimierungsläufen und damit mittelbar zur Leistungsbewertung der Algorithmen selbst dienen. Abschließend wird auf die softwaretechnische Unterstützung der automatischen Testreihenberechnung eingegangen.

5.1.1 Visualisierung von Optimierungsverläufen

Während einer numerischen Optimierung wird das Stoffstrommodell wiederholt vom Leitstand aus aufgerufen, um die Zielfunktion für bestimmte Testlösungen $\mathbf{x}_i^{(k)}$ zu berechnen: $f_i^{(k)} = f(\mathbf{x}_i^{(k)})$. Der Index k in dieser Notation zählt die Iterationen und i nummeriert die verschiedenen Modellaufrufe innerhalb einer Iteration. Im Leitstand werden jedoch nicht alle berechneten Daten ausgegeben. Sowohl in der tabellarischen als auch der graphischen Ausgabe (Abbildung 5.1) erscheinen iterationsweise nur die besten bis dahin gefundenen Lösungen ($\mathbf{x}_B^{(k)}, f_B^{(k)}$).

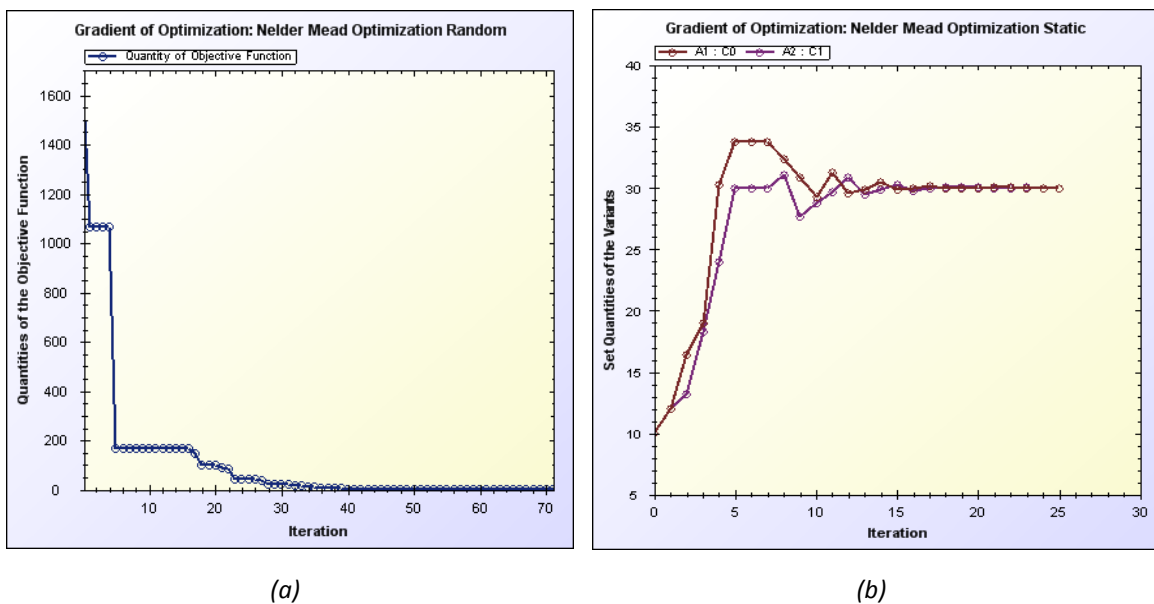


Abbildung 5.1: Visualisierung eines Optimierungslaufs im Leitstand: Gegen die Iteration werden (a) der beste bisher gefundene Zielfunktionswert bzw. (b) die zugehörigen Werte der Entscheidungsvariablen abgetragen.

Optimierungstrajektorie

Um die Bewegung eines Optimierungsalgorithmus im Lösungsraum besser zu veranschaulichen, wird in diesem Kapitel an mehreren Stellen die *Optimierungstrajektorie* dargestellt (Abbildung 5.2b). Dazu werden die Entscheidungsvariablen in kartesischen Koordinaten gegeneinander abgetragen. Zur besseren Orientierung werden Restriktionen durch eine dicke Umrandung des zulässigen Bereichs und die Zielfunktion durch Niveaulinien ebenfalls dargestellt.

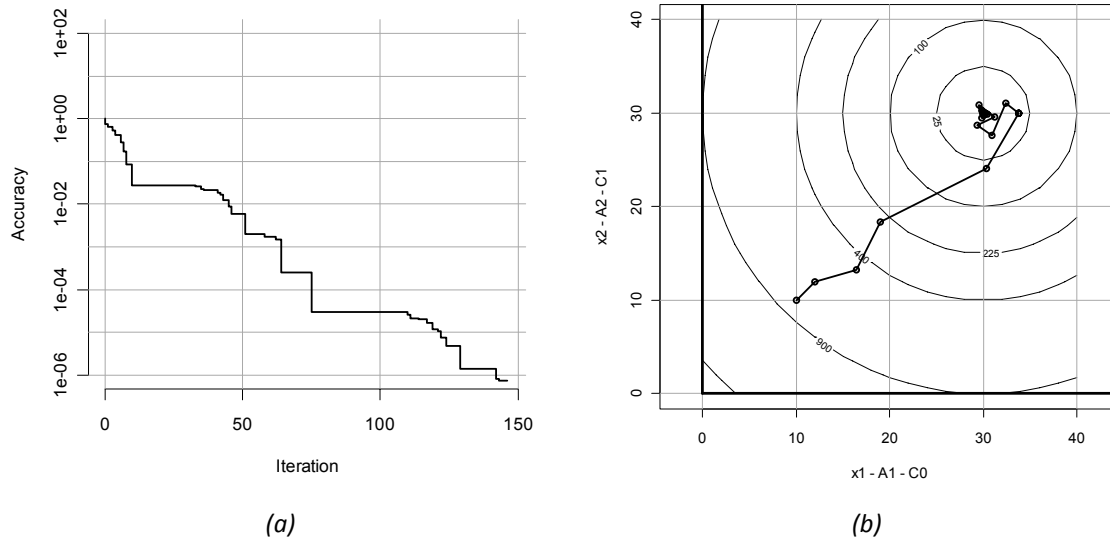


Abbildung 5.2: Anwendung der in diesem Kapitel üblicherweise verwendeten Darstellung auf die Optimierung aus Abbildung 5.1. (a) Logarithmus des Optimumsabstands gegen Iteration. (b) Werte der Entscheidungsvariablen in kartesischem Koordinatensystem.

Normierung der Zielfunktionswerte – der relative Optimumsabstand

In diesem Kapitel werden überwiegend Optimierungen an Testproblemen mit *bekanntem* Optimum durchgeführt. Um auch Optimierungen von Optimierungsproblemen mit stark unterschiedlichen Wertebereichen der Zielfunktion besser miteinander vergleichen zu können, wird für die Visualisierung von Optimierungsverläufen in der Regel nicht die Zielfunktion selbst, sondern eine daraus abgeleitete normierte Größe, der so genannte *relative Optimumsabstand* z verwendet (5.1). Es handelt sich dabei um den Abstand einer Lösung $f_i^{(k)}$ vom optimalen Zielfunktionswert f^* . z ist eine relative Größe, da der Abstand auf die Entfernung des Startwertes $f^{(0)}$ vom Optimum bezogen wird.

$$z_i^{(k)} = \frac{|f_i^{(k)} - f^*|}{|f^{(0)} - f^*|} \quad (5.1)$$

Gleichung (5.1) bildet ein beliebiges Optimierungsproblem für f auf ein *Minimierungsproblem* für z ab. Der relative Optimumsabstand z variiert im Verlauf einer Optimierung im Wesentlichen

zwischen den Werten $z^{(0)} = 1$ und $z^* = 0$. Abstände $z_i^{(k)} > 1$ können jedoch auftreten, wenn sich der Algorithmus zunächst vom Optimum entfernt, oder die Zielfunktion im Lösungsraum nicht monoton verläuft. Da $z_i^{(k)}$ während einer Optimierung häufig um mehrere Größenordnungen variiert, bietet sich die Verwendung einer *logarithmische Skala* an. So kann das Konvergenzverhalten der Algorithmen auch nahe des Optimums gut dargestellt werden (Abbildung 5.2a).

Best-Vektor-Darstellung versus Berücksichtigung aller Modellauswertungsergebnisse

Für die detaillierte Analyse einzelner Optimierungsläufe ist die im Leitstand implementierte Ausgabe der Bestlösungen nicht ausreichend. Kann in der Iteration k keine Verbesserung der Zielfunktion erreicht werden, so ist $f_B^{(k)} = f_B^{(k-1)}$. Geschieht dies mehrere Male in Folge, so entsteht eines jener *Plateaus*, die immer wieder im Verlauf einer $f_B^{(k)}$ -Kurve in Abbildung 5.2(a) zu sehen sind. Die scheinbare Stagnation der Optimierung kann nur dann aufgelöst werden, wenn alle vom Algorithmus bestimmten Suchlösungen in die Darstellung übernommen werden⁶⁰. Der Wechsel zur Darstellung aller sukzessiv berechneten Versuchslösungen $f_i^{(k)}$ (rote Kurve in Abbildung 5.3) macht deutlich, dass im Verlaufe eines Plateaus eine stetige Konvergenz der Suchlösungen stattfindet, die schließlich auch die Voraussetzung für eine weitere Verbesserung der Bestlösung darstellt.

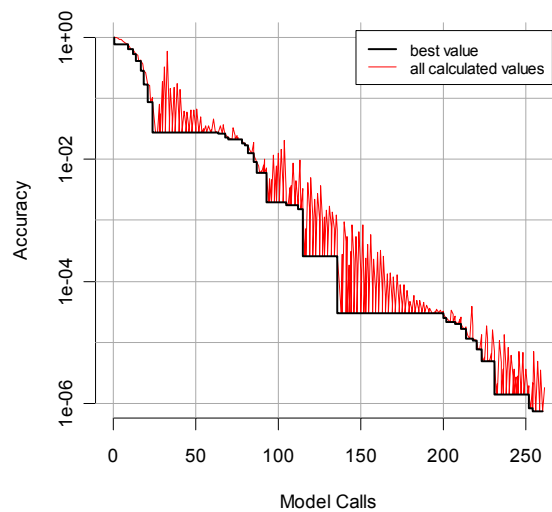


Abbildung 5.3: Optimumsabstand gegen Modellaufruf für die Optimierung aus Abbildung 5.2. Die rote Kurve weist alle Zielfunktionswertungen (Modellaufrufe) aus. Die schwarze Kurve nur die besten bis zum aktuellen Modellaufruf gefundenen Zielfunktionswert.

⁶⁰ Diese können vom Optimierungsleitstand in eine CSV-Datei exportiert werden.

5.1.2 Performancekriterien für einzelne Optimierungsläufe

Bei der Beurteilung der Leistungsfähigkeit heuristischer Optimierungsverfahren stehen zwei Kriterien im Vordergrund: eine gute Optimierung sollte (i) *schnell* zu einem Ergebnis kommen, welches (ii) *nah* beim tatsächlichen Optimum liegt (Rardin und Uzsoy, 2001: S.261). Diese beiden Gütekriterien werden im Folgenden durch die *Optimierungsdauer* (efficiency) und die *Approximationsgenauigkeit* (accuracy) operationalisiert.

Optimierungsdauer (Konvergenzgeschwindigkeit)

Die Optimierungsdauer wird typischerweise durch die *Anzahl benötigter Modellaufrufe* gemessen. Sie ist vom Erfolg der Optimierung, d.h. der erzielten Approximationsgenauigkeit, unabhängig. Sofern die Optimierung nicht manuell abgebrochen wird, gibt die Optimierungsdauer an, wie viele Berechnungen bis zur Erfüllung des im Algorithmus implementierten Abbruchkriteriums durchgeführt wurden. Gelegentlich wird die Optimierungsdauer auch anhand der Anzahl benötigter *Iterationen* angegeben. Dies ist aber nur sinnvoll, solange verschiedene Läufe desselben Verfahrens verglichen werden, da verschiedene Verfahren in der Anzahl je Iteration benötigter Modellaufrufe – teilweise erheblich – voneinander abweichen. Beispielsweise benötigt CNM, solange er auf keine Restriktionen trifft, maximal zwei Modellaufrufe je Iteration, während ES das Modell für alle Nachkommen einer Population neu berechnet, was in der Regel ein Vielfaches an Modellaufrufen erfordert.

Approximationsgenauigkeit

Die Approximationsgenauigkeit misst, wie nah die bei einer numerischen Optimierung gefundene Lösung bei dem tatsächlichen Optimum eines Optimierungsproblems liegt. Sie kann daher nur quantifiziert werden, wenn das Optimum bekannt ist. Es ist daher naheliegend, für die Definition der Approximationsgenauigkeit vom im vorigen Abschnitt definierten relativen Optimumsabstand $z_i^{(k)}$ (5.1) auszugehen. Als *Approximationsgenauigkeit* wird der relative Optimumsabstand z_B der besten Lösung bezeichnet, die im Verlauf einer numerischen Optimierung ermittelt wurde.

Der Optimumsabstand und damit die Approximationsgenauigkeit können prinzipiell auch im Lösungsraum definiert werden. Die Abstände von Zielfunktionswerten in (5.1) müssen dann durch entsprechende Entfernungen von Lösungsvektoren ersetzt werden (5.2). Für die Berechnung der Entfernung können geeignete Metriken, wie z.B. die euklidische Norm, verwendet werden.

$$z_i^{(k)} = \frac{|\mathbf{x}_i^{(k)} - \mathbf{x}^*|}{|\mathbf{x}^{(0)} - \mathbf{x}^*|} \quad (5.2)$$

5.1.3 Bewertung von Testreihen

Einflussfaktoren auf die Güte numerischer Optimierungen

Wie im vorigen Abschnitt gezeigt wurde, kann die Güte von Optimierungsalgorithmen nur indirekt über die Bewertung numerischer Optimierungsläufe verglichen werden. Außer vom Optimierungsproblem selbst, hängt die Güte einer numerischen Optimierung von folgenden Faktoren ab:

- von der vorgegebenen *Startlösung* $x^{(0)}$
- von der *Initialisierung* (im Falle eines pseudo-zufallsbasierten Verfahrens)
- von der *Parametrisierung* des Algorithmus
- vom gewählten *Abbruchkriterium*

Statistische Vorgehensweise

Daher können sich mehrere Optimierungsläufe hinsichtlich Optimierungsdauer und Approximationsgenauigkeiten sogar dann unterscheiden, wenn es sich um denselben Algorithmus handelt, der immer wieder auf dasselbe Optimierungsproblem angewendet wird. Um zwei Algorithmen hinsichtlich ihrer Eignung zur Optimierung eines bestimmten Problems miteinander zu vergleichen, muss daher ein *statistischer* Ansatz gewählt werden (Peters et al., 1999: S.132, Rardin und Uzsoy, 2001, Syrjakow, 2003: S.101). Nicht einzelne Optimierungen müssen bewertet und miteinander verglichen werden, sondern *repräsentative Stichproben* von Optimierungsläufen, die im Folgenden als *Testreihen* bezeichnet werden. Die einzelnen Optimierungsläufe einer Testreihe unterscheiden sich demnach (i) in der *Startlösung* und, falls ein pseudo-zufallsbasiertes Verfahren verwendet wird, (ii) in dessen *Initialisierung*. Repräsentativität der Stichprobe bedeutet in diesem Kontext also, dass der Bereich zulässiger Startlösungen und Initialisierungen hinreichend gut abgedeckt wird.

Anders verhält es sich mit den Parametern des Algorithmus und dem Abbruchkriterium. Die numerischen Optimierungsexperimente dienen im Falle dieser Einflussfaktoren dazu, *Regeln* zu entwickeln, mit denen entschieden werden kann, bei welcher Art des Optimierungsproblems welche Parametrisierung gewählt werden sollte (vgl. Abschnitt 5.5). Daher soll der Einfluss dieser beiden Parameter gerade nicht statistisch betrachtet werden.

Aggregation der Gütekriterien: Optimierungsdauer und Treffhäufigkeit

Für die vergleichende Beurteilung von Testreihen müssen die ursprünglich für einzelne Optimierungen definierten Gütekriterien Optimierungsdauer und Approximationsgenauigkeit zu aussagekräftigen neuen Maßzahlen aggregiert werden. Im Falle der *Optimierungsdauer* ist es offensichtlich sinnvoll, das *arithmetische Mittel* der Optimierungsdauern aller Optimierungen einer Testreihe als Kennzahl zu verwenden. Anhand einer einfachen Varianzanalyse kann abgeschätzt

werden, ob die Optimierungsdauern verschiedener Testreihen signifikant voneinander abweichen.

Im Falle der *Approximationsgenauigkeit* führt die Mittelwertbildung nicht zu einer aussagekräftigen Größe. Hier wird ein anderer Ansatz verwendet. Man gibt eine Untergrenze für die Approximationsgenauigkeit vor. Dadurch wird ein Intervall definiert, in dessen Mitte sich das tatsächliche Optimum befindet. Als erfolgreich werden dann alle Optimierungen betrachtet, deren beste Lösungen in dieses Intervall fallen. Der Anteil erfolgreicher Optimierungen in einer Testreihe wird als *Treffhäufigkeit* bezeichnet. Bei Optimierungsproblemen mit mehreren Optima werden durch dieses Vorgehen mehrere Intervalle um das globale und die lokalen Optima definiert. Man kann dann zwischen der lokalen und der globalen Treffhäufigkeit differenzieren, je nachdem, ob man den Anteil der Optimierungen meint, deren Lösung in ein beliebiges Zielintervall fällt, oder den Teil, deren Lösung in das Zielintervall um das globale Optimum fällt.

5.1.4 Softwaretechnische Unterstützung der Testreihenberechnung

Um die Testreihenberechnung mit dem Optimierungsleitstand effektiv durchführen zu können, wurde dieser um ein Modul erweitert, welches die notwendigen Methoden zur Verfügung stellt. Die Einstellungen dazu werden im Konfigurationsfenster Abbildung 5.4 vorgenommen. Dieses gliedert sich in drei Teile. Unter „Sample Properties“ werden die für die Erzeugung von Stichproben mit verschiedenen Startvektoren und Initialisierungen des Algorithmus notwendigen Angaben vorgenommen (Tabelle 5.1).

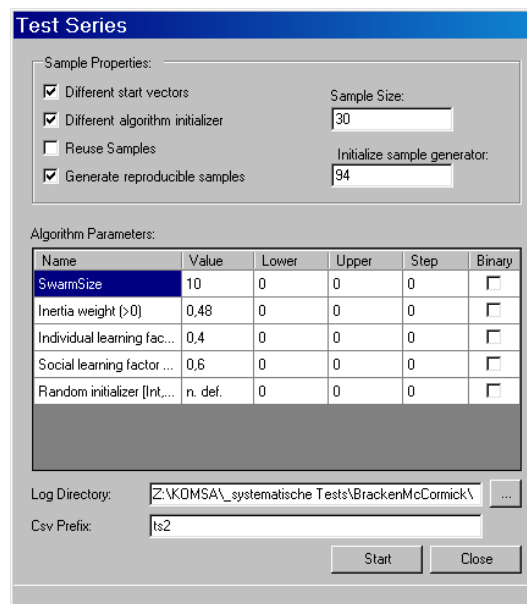


Abbildung 5.4: Screenshot des Testreihen-Fensters

Für systematische Tests des Einflusses der Parameter des Algorithmus auf die Optimierung kann in der Liste „Algorithm Parameters“ festgelegt werden, in welchen Grenzen (Lower, Upper) diese bei der Berechnung der Testreihe schrittweise (Step) variiert werden.

Das Ergebnis einer Testreihen-Berechnung sind (i) CSV-Dateien für die einzelnen Optimierungsläufe, sowie (ii) eine Übersichtsdatei mit den wichtigsten Daten zu den Einzelläufen: Startvektor $x^{(0)}$, beste gefundene Lösung x_B bzw. f_B , sowie die jeweils benötigte Anzahl Iterationen und Modellaufrufe. Aus diesen Daten können Treffhäufigkeiten und Optimierungsdauern berechnet werden. Besonders wichtig hinsichtlich der Reproduzierbarkeit einzelner Optimierungsläufe ist die Protokollierung der zufällig erzeugten Initialisierung r_A der pseudo-zufallsbasierten Suchalgorithmen. Im unteren Teil des Fensters wird das Verzeichnis gewählt, in welches die CSV-Dateien mit den Ergebnissen abgelegt werden. Außerdem eine alphanumerische Kombination, mit der die Namen aller Ergebnisdateien beginnen sollen (CSV Prefix).

Tabelle 5.1: Angaben zur Stichprobenerzeugung im Testreihen-Fenster

Feldname	Typ	Zweck
Sample Size	integer	Der Stichprobenumfang gibt an, aus wie vielen Optimierungen eine Stichprobe besteht.
Different start vectors	bool	true = die einzelnen Optimierungen einer Stichprobe haben verschiedene Startvektoren. Diese werden durch eine Monte-Carlo-Suche im zulässigen Bereich ermittelt. false = alle Stichprobenelemente haben den gleichen Startvektor.
Different algorithm initializer	bool	true = einzelne Optimierungen unterscheiden sich in der Initialisierung des Algorithmus. false = alle Initialisierungen des Algorithmus sind gleich.
Reuse sample	bool	true = die gleiche Stichprobe wird für alle im unteren Bereich des Testreihenfensters spezifizierten Kombinationen der spezifischen Parameter des Algorithmus verwendet.
Generate reproducible samples	bool	true = die Stichprobenerzeugung (Zufallsgenerator für Startvektoren und Initialisierungen des Algorithmus) wird selbst initialisiert. So können gleiche Stichproben in verschiedenen Testreihen verwendet werden.
Initialize sample generator	integer	Startwert für den Zufallsgenerator der Stichprobenerzeugung.

Der Umfang einer Testreihe, d.h. die Anzahl benötigter Optimierungsläufe, ergibt sich aus der Multiplikation des Stichprobenumfangs mit der Anzahl der in der Liste definierten Parameterkombinationen des Algorithmus. Wird in Abbildung 5.4 zusätzlich zum Stichprobenumfang von 30 beispielsweise der Parameter „Population size“ mit einer Schrittweite von 4 zwischen der Untergrenze 6 und der Obergrenze 14 variiert, müssen insgesamt $3 \cdot 30 = 90$ Optimierungen durchgeführt werden.

Obwohl das Testreihenmodul des Optimierungsleitstands primär für numerische Tests mit den implementierten Algorithmen konzipiert wurde, kann es auch als Metastrategie der Optimierung eingesetzt werden.

5.2 Beschreibung und Veranschaulichung der Funktionsweise der ausgewählten Optimierungsalgorithmen

Die in diesem Kapitel untersuchten Optimierungsverfahren sind Heuristiken aus verschiedenen „Familien“ direkter Suchstrategien (Tabelle 5.2). CNM ist ein Hybrid aus zwei sehr ähnlichen Hillclimbing-Strategien, dem bereits zitierten Nelder-Mead-Algorithmus (vgl. § 4.1.2) und dem durch eine Erweiterung⁶¹ desselben für restringierte Optimierungsprobleme entstandenen Complex-Algorithmus (Box, 1965). Hillclimbing-Verfahren sind *lokale* Suchstrategien, welche – vergleichbar einem Bergsteiger im Nebel – von ihrem jeweiligen Standpunkt aus ihre unmittelbare Umgebung nach besseren Lösungen absuchen. Ist die Suche erfolgreich, wird das Verfahren von der neuen Lösung aus wiederholt. Ansonsten wird es abgebrochen und davon ausgegangen, dass es sich bereits am „Gipfel“, d.h. bei der optimalen Lösung, befindet.

Tabelle 5.2: Übersicht der getesteten direkten Optimierungsverfahren

Algorithmus		Verfahrenstyp
Name	Abkürzung	
Complex/Nelder-Mead	CNM	Hillclimbing-Verfahren
Evolutionsstrategie	ES	Evolutionäres Verfahren
Particle Swarm Optimization	PSO	Schwarmintelligentes Verfahren

⁶¹ COMPLEX ist eine Abkürzung für „constrained simplex“ (Schwefel, 1995: S.61f).

Die Evolutionsstrategie (Rechenberg, 1973) ist eine globale Suchstrategie, die auf den evolutionären Prinzipien der Rekombination und Mutation von Testlösungen beruht. Die Particle Swarm Optimization (Eberhart und Kennedy, 1995) gehört zur Klasse der schwarm-intelligenten Optimierungsalgorithmen. Wie bei den meisten Verfahren dieser Art ist ihre Suchstrategie einem natürlichen Vorbild nachempfunden. So basiert das Lösungsprinzip auf einer Population von Testlösungen, die sich wie die Individuen eines Vogelschwarms auf ihrer Suche nach Futter verhalten.

Alle drei Verfahren haben gemeinsam, dass sie auf *Suchmengen* M_k zulässiger Lösungen basieren, welche *iterativ* verbessert werden. Entsprechend der zugrunde liegenden Metaphern werden diese Suchmengen als Simplex (CNM), Population (ES) oder Schwarm (PSO) bezeichnet. Die Verfahren unterscheiden sich jedoch in der Art, wie neue Lösungen S_{k+1} in der Umgebung von M_k konstruiert werden. In diesem Abschnitt werden die *Konstruktionsprinzipien* aller untersuchten Optimierungsalgorithmen erklärt⁶² und anhand typischer Optimierungsverläufe für ein einfaches zweidimensionales Optimierungsproblem veranschaulicht.

Auf dieser Grundlage können erste Rückschlüsse über *spezifische Stärken und Schwächen* der verschiedenen Suchstrategien gezogen werden, die in den darauf folgenden Abschnitten anhand numerischer Optimierungsexperimente weiter untersucht werden. Die Visualisierung der Optimierungsverläufe im Lösungsraum liefert das grundlegende Verständnis, um das Verhalten der Algorithmen für realistische, d.h. insbesondere auch höherdimensionale, Optimierungsprobleme besser zu verstehen.

5.2.1 Complex/Nelder-Mead-Algorithmus (CNM)

Nelder-Mead- und Complexalgorithmus sind Hillclimbing-Strategien, die nicht von einer einzigen Lösung, sondern von einer Suchmenge ausgehen. Beim ursprünglichen Nelder-Mead-Verfahren wird für die Lösung eines n -dimensionalen Optimierungsproblems ein als Simplex bezeichneter Polyeder mit $n+1$ Ecken schrittweise im Lösungsraum verlagert. Dafür wird die Zielfunktion jeweils in allen Ecken ausgewertet und insbesondere die Ecke \mathbf{x}_s mit dem schlechtesten Zielfunktionswert ermittelt. Das Fortbewegungsprinzip besteht nun darin, \mathbf{x}_s am Schwerpunkt der übrigen Ecken zu spiegeln, um so eine neue Testlösung in der Umgebung des Simplex zu erzeugen. Je nach Erfolg dieser Operation kann der Simplex auch verformt, d.h. expandiert oder kontrahiert, werden. Der Complex-Algorithmus weicht vom Nelder-Mead-Verfahren im Wesentlichen in zwei Punkten ab: er verwendet einen Polyeder mit mehr als $n+1$ Ecken und führt bei jeder Iteration eine Expansion durch. Beide Maßnahmen sollen verhindern, dass der Polyeder in der Nähe einer Ungleichungsrestriktion in einen Subraum mit reduzierter Dimensionalität kollabiert (Schwefel, 1995: S.62).

⁶² Zu den Details der softwaretechnischen Implementierung siehe Zimmermann et al. (2009: S.82-96).

Aufbau des Startpolyeders

Beim ursprünglichen Nelder-Mead-Verfahren werden für die Lösung eines n -dimensionalen Optimierungsproblem stets Polyeder mit $n+1$ Ecken verwendet. Beim Complex-Algorithmus werden Polyeder mit $k > n+1$ Ecken verwendet. Damit soll verhindert werden, dass der Polyeder in der Nähe von Restriktionen in einen Subraum kollabiert (Box, 1965: S.44). Das bedeutet für $n = 2$ anschaulich, dass alle drei Punkte des Simplex auf einer Geraden zu liegen kommen. In der Folge kann der Algorithmus nur noch entlang dieser Geraden suchen und wird das Optimum daher im Allgemeinen verfehlen. Schwefel (1995: S.62) schlägt eine Polyedergröße von $k = 2n$ vor. Beim CNM-Algorithmus kann die Polyedergröße frei gewählt werden. Als Standard wird bei niederdimensionalen Problemen ($<5D$) ein Polyeder mit $k = 2n$ und darüber ein Simplex mit $k = n+1$ Ecken verwendet. Damit die Optimierungsverläufe in diesem Abschnitt besser nachvollziehbar sind, wurden für alle entsprechenden Abbildungen Simplizes, in diesem Fall also einfache Dreiecke, verwendet.

Für die Erzeugung des Startpolyeders M_0 ausgehend von der vorgegebenen Startlösung $\mathbf{x}^{(0)}$ stehen bei CNM drei unterschiedliche Varianten zur Verfügung. Bei der *deterministischen* Erzeugung des Startpolyeders (CNMd) werden jeweils einzelne Komponenten der Startlösung $\mathbf{x}^{(0)}$ mit dem Faktor 1,2 multipliziert. Diese Variante liegt der Erzeugung des roten Startsimplex M_0 links unten in Abbildung 5.5(b) zugrunde. Sie entspricht dem ursprünglich von Nelder und Mead vorgeschlagenen Verfahren (vgl. Schwefel, 1995: S.58).

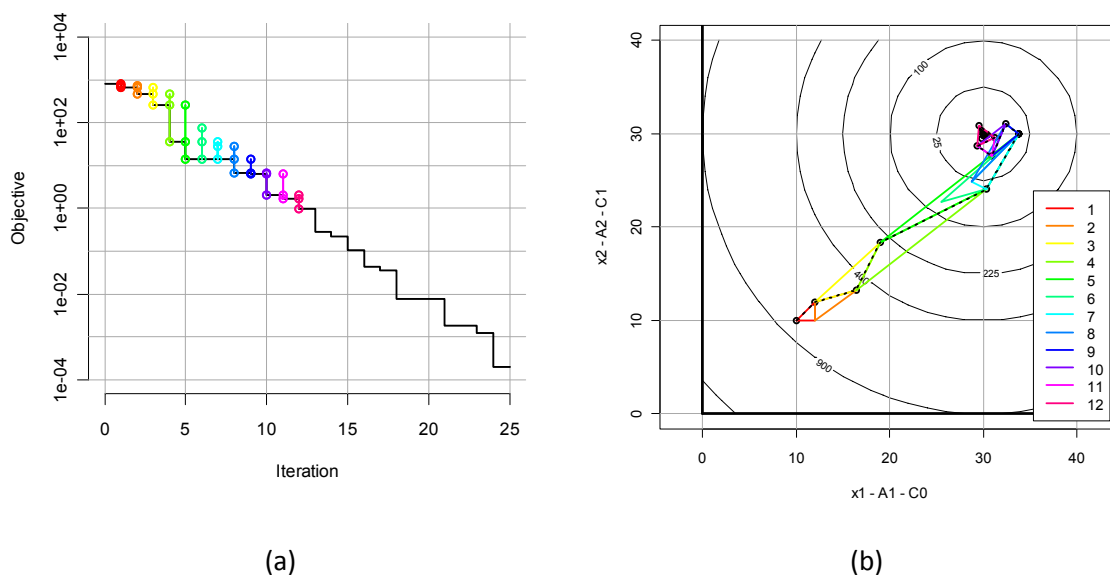


Abbildung 5.5: Complex/Nelder-Mead (CNM). Visualisierung einer exemplarischen Optimierung. (a) Optimumsabstand: die Werte der Polyederecken sind iterationsweise in einheitlichen Farben dargestellt und durch eine vertikale Linie miteinander verbunden. Die schwarze Kurve stellt den besten Zielfunktionswert je Iteration dar. (b) Lösungsraum: die Ecken der Simplizes für jede Iteration sind durch Kanten miteinander verbunden. So kann die Simplex-Bewegung im Lösungsraum nachvollzogen werden.

Außerdem wurde eine *pseudo-zufallsbasierte* Erzeugungsmethode implementiert (CNMr). Sie entspricht dem ursprünglich von Box (1965: S.43) vorgeschlagenen Verfahren. Dabei werden die $\mu-1$ vom Startvektor verschiedenen Ecken des Startpolyeders zufällig in einem durch die jeweiligen Ober- und Untergrenzen (l_i^x, u_i^x) der einzelnen Entscheidungsvariablen definierten Bereich gewählt.⁶³

$$[x_j^{(0)}]_i = l_i^x + r \cdot (u_i^x - l_i^x) \quad \text{mit } i = 1 \dots n; j = 1 \dots \mu-1 \quad (5.3)$$

und r gleichverteilt in $(0,1)$

Eine zweite Variante von CNMr stellt eine Kreuzung der von Nelder und Mead bzw. Box vorgeschlagenen Methoden dar, bei der (5.3) nur auf eine einzelne Komponente von $\mathbf{x}^{(0)}$ angewendet wird, während alle übrigen unverändert bleiben. Weiterhin wurde ein Skalierungsfaktor implementiert, welcher den Suchbereich für den deterministischen Aufbau um eine lokale Umgebung um $\mathbf{x}^{(0)}$ einschränkt.

Weiterer Optimierungsverlauf

In den nächsten vier Iterationen (2 bis 5) wird der Polyeder stets reflektiert und expandiert. So *bewegt* er sich in guter Näherung entlang des Gradienten der Zielfunktion (etwa senkrecht zu den Niveaulinien) und *streckt sich* in dieser zunächst erfolgreichen Richtung. In der sechsten Iteration schlägt die Reflexion jedoch fehl, d.h. der erzielte Zielfunktionswert ist schlechter als der Ausgangspunkt und damit der schlechteste Punkt der vorhergehenden Iteration. In dieser Situation findet eine *Kontraktion* statt: der Ausgangspunkt wird in Richtung des Schwerpunktes der restlichen Ecken verschoben. In Iteration 7 erfolgt aus denselben Gründen eine erneute Kontraktion.

Obwohl die Zielfunktion nicht verbessert werden kann, verändert der Simplex in den beiden Iterationen 6 und 7 seine Form: in einer Art Bremsvorgang wird er nun in derselben Richtung *gestaucht*, in welcher er vorher expandierte. Er wird effektiv *kleiner* (konvergiert), ohne jedoch zu einer global besseren Lösung zu finden. Wird in einem Graphen nur der beste gefundene Zielfunktionswert $f_B^{(k)}$ dargestellt, führt dieses Verhalten zu einem jener *Plateaus*, die bereits im Zusammenhang mit Abbildung 5.3 diskutiert wurden. Am Zusammenziehen des Polyeders können zwei unterschiedliche Konvergenzkriterien für das Verfahren festgemacht werden, die in der Praxis beide vorkommen: (a) die Varianz der Zielfunktionswerte an den Polyederecken, (b) seine Ausdehnung bzw. sein Volumen im Lösungsraum. Für beide kann ein Mindestwert festgelegt werden, bei dessen Unterschreiten der Algorithmus abbricht.

Erst bei Iteration 8 kann erneut reflektiert (nicht jedoch expandiert) und damit eine Verbesserung des bis dahin gefundenen besten Vektors erreicht werden. Es folgen die Iterationen 9 bis 12 mit den Operationen Kontraktion, Reflexion, Kontraktion und Reflexion. Der Simplex der letzten dar-

⁶³ Dieses Verfahren entspricht im Prinzip einer Monte-Carlo-Suche.

gestellten Iteration entspricht in der Größe etwa dem Ausgangssimplex. Das Optimum liegt in seinem Inneren, so dass im Weiteren vorwiegend Kontraktionen notwendig sind, um dieses weiter anzunähern.

Unterteilung der Optimierung in zwei Phasen

Der in Abbildung 5.5 dargestellte Optimierungslauf zeigt mehrere charakteristische Eigenschaften von CNM, die im Folgenden näher diskutiert werden sollen. Während der Optimierung können grob zwei *unterschiedliche Phasen* unterschieden werden: Zu Anfang ist die Ausdehnung des Simplex klein im Vergleich zu seiner Entfernung vom Optimum. Durch Reflexion und Expansion ist das Verfahren in dieser Phase sehr gut in der Lage, den Simplex effektiv in Richtung des Optimums zu bewegen. So kann in nur fünf Iterationen der Zielfunktionswert auf 1/100 seines Ausgangswertes reduziert werden ($f^* = 0$). Durch die Verlagerung und Vergrößerung des Simplex in dieser ersten Phase sind seine Ausdehnung und seine Entfernung vom Optimum in Iteration 5 in derselben Größenordnung. Ab hier tritt das Verfahren nun in die zweite Phase ein, bei der eine weitere Verbesserung nur durch alternierende Kontraktionen und Reflexionen erzielt werden kann. In dieser Phase ist das Verfahren bei der Annäherung des Gradienten jedoch deutlich ineffizienter, als die meisten ableitungsbasierten Methoden. Dies führt zu der aus der Literatur bekannten vergleichsweise langsamen Konvergenz in der Nähe des Optimums (vgl. z.B. Alt, 2002: S.36).

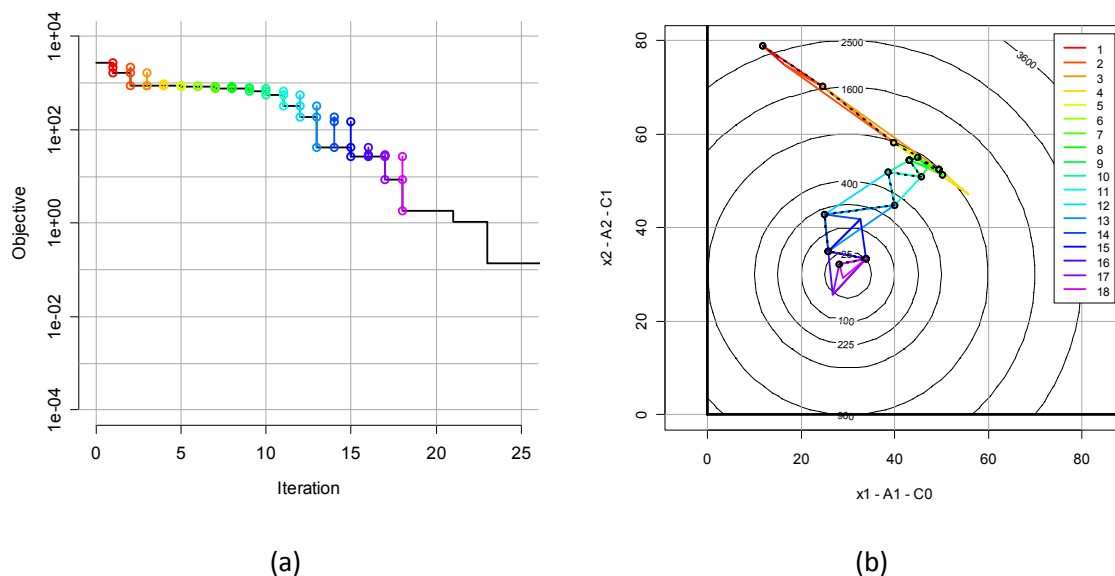


Abbildung 5.6: CNM-Optimierungslauf mit zwei orthogonalen Bewegungsrichtungen des Simplex bei der Näherung des Optimums.

Neuausrichtung des Polyeders im Lösungsraum

Die Plateaus der $f_B^{(k)}$ -Kurve einer CNM-Optimierung treten vor allem bei *Neuausrichtungen* des Polyeders im Lösungsraum auf. Dass lässt sich besonders anschaulich am Beispiel der in Abbildung 5.6 dargestellten Optimierung mit dem CNMr-Algorithmus zeigen. Der zufällig erzeugte Startsimplex hat eine längliche Form und liegt etwa in einem 45°-Winkel zum Gradienten. Diese Ausgangsorientierung dominiert zunächst die Auswahl der Suchrichtung, und es wird das Optimum ermittelt, welches entlang dieser Richtung liegt. Dort kontrahiert der Polyeder über etwa sechs Iterationen hinweg, um sich schließlich fast orthogonal zur ursprünglichen Richtung und in guter Näherung des lokalen Gradienten der Zielfunktion weiterzubewegen. Die mit der Neuausrichtung verbundene Kontraktionsphase, in der die Ecken des Polyeders schrittweise auf dessen Schwerpunkt hin zusammengezogen werden, kann insbesondere bei höherdimensionalen Optimierungsproblemen zu hohen Optimierungsdauern führen.

5.2.2 Evolutionsstrategie (ES)

Die Startpopulation M_0 wird bei der ES wie bei der Monte-Carlo-Variante von CNMr erzeugt. Die iterative Entwicklung der Suchmenge verläuft im Falle der evolutionären Strategie (ES) ganz anders als bei CNM. Während das Prinzip der Reflexion des CNM-Algorithmus noch als eine Art heuristische Annäherung des Zielfunktionsgradienten verstanden werden kann, basiert die Erzeugung neuer Lösungsvektoren bei der ES auf den vollständig zufallsbasierten Prozessen der Rekombination und Mutation.

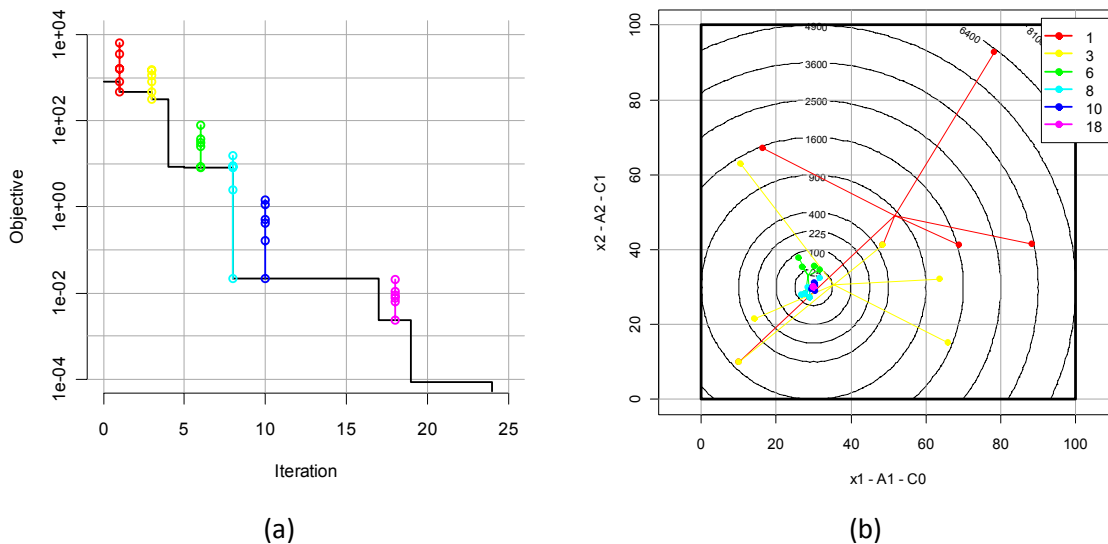


Abbildung 5.7: Evolutions-Strategie (ES). Visualisierung einer exemplarischen Optimierung. (a) Optimums-abstand: die Werte der Individuen einer Population sind jeweils in derselben Farbe dargestellt und durch eine Linie miteinander verbunden. Die schwarze Kurve stellt den besten Zielfunktionswert je Iteration dar (b) Lösungsraum: alle Individuen der Population einer bestimmten Generation (Iteration) sind durch Kanten miteinander verbunden.

Erzeugung neuer Lösungen durch Rekombination und Mutation

Bei der *Rekombination* werden zwei Lösungen in M_k zufällig als „Eltern“ gewählt. Je besser der Zielfunktionswert einer Lösung, desto höher die Wahrscheinlichkeit, als Elternteil ausgewählt zu werden. Je zwei Elternlösungen erzeugen zwei neue Lösungen (Kinder), die zufällig auf der Geraden verteilt werden, welche die beiden Elternlösungen miteinander verbindet. Bei der anschließenden *Mutation* werden die neu erzeugten Lösungsvektoren noch einmal zufällig um ihre in der Rekombination ermittelte Position gestreut. Die Streuweite hängt von der Streuung der Populationsmitglieder in M_k ab.

Auswahl der neuen Suchmenge durch Selektion

Die in jeder Iteration durchgeführte Selektion, bei der jeweils nur die besten Individuen einer Generation beibehalten werden, führt zu einer *zunehmenden Konzentration* der Population um das Optimum. Während die Startpopulation in Abbildung 5.7(b) sich noch nahezu über den gesamten Lösungsraum erstreckt, liegt sie nach acht Iterationen bereits innerhalb der 25er-Niveaulinie. Der Schwerpunkt jeder Population ist in Abbildung 5.7(b) als Schnittpunkt der Linien dargestellt, welche die einzelnen Individuen sternförmig miteinander verbinden. Der evolutionäre Konzentrationsprozess kann, wie im hier dargestellten Beispiel, zu einer effektiven Verlagerung des Schwerpunktes der Population führen. Eine dem CNM vergleichbare *zielgerichtete* Verlagerung der Suchmenge findet hingegen nicht statt.

Wie beim CNM können auch bei ES Plateaus in der $f_B^{(k)}$ -Kurve entstehen, beispielsweise zwischen den Iterationen 8-17 in Abbildung 5.7(a). Da die ES auf der zufälligen Erzeugung neuer Lösungen basiert, ist natürlich nicht garantiert, dass unter den Nachkommen einer Population eine Lösung ist, die besser als das beste Individuum der Elterngeneration bewertet wird. In einem solchen Fall gilt dann $f_B^{(k+1)} = f_B^{(k)}$ und es entsteht ein Plateau. Im Extremfall kann sogar die gesamte Population zwischen zwei Iterationen unverändert bleiben ($M_{k+1} = M_k$), wenn alle Nachkommen schlechtere Zielfunktionswerte haben als die Eltern. Das Zufallsmoment bei Rekombination und Mutation sorgt in diesem Fall jedoch dafür, dass der Algorithmus nicht in eine Endlosschleife läuft: Aus ein und derselben Population können verschiedene Nachkommen erzeugt werden. Im Allgemeinen verbleiben Individuen über mehrere Iterationen bzw. Generationen hinweg innerhalb der Population.

5.2.3 Particle Swarm Optimization (PSO)

Während bei CNM und ES Lösungen über mehrere Iterationen hinweg in der Suchmenge verbleiben können, gibt es beim Particle Swarm Algorithmus prinzipiell keine derartigen Überschneidungen: $M_{k+1} \cap M_k = \{ \}$. Bei der PSO werden die Elemente der Suchmenge als Individuen bzw. *Teilchen* eines *Schwarms* verstanden, die sich auf ihrer Suche nach der optimalen Lösung auf stetigen Bahnen durch den Lösungsraum bewegen. Deshalb weicht auch die Darstellung in Abbildung 5.8 von den entsprechenden Abbildungen für CNM und ES ab. Anstatt die Suchmengen für verschiedene Iterationen als zusammenhängende Gebilde aufzufassen und durch farbliche

Differenzierung zu visualisieren, werden nun die Positionen der einzelnen Schwarmteilchen über mehrere Iterationen hinweg miteinander verbunden.

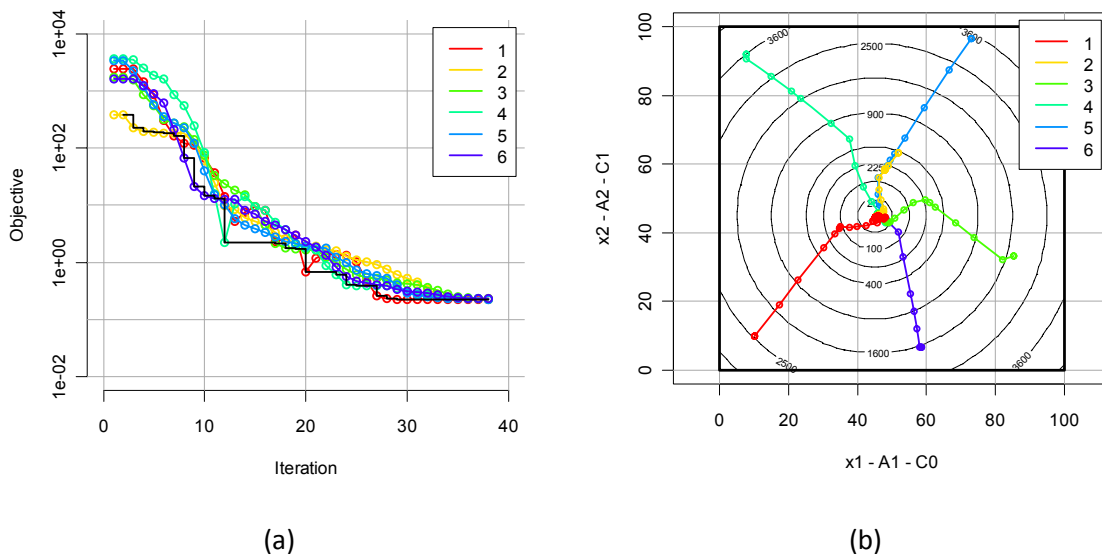


Abbildung 5.8: Particle Swarm Optimization (PSO). Visualisierung der exemplarischen Optimierung mit einem aus sechs Teilchen bestehenden Schwarm. (a) Optimumsabstand: alle Ergebnisse eines Teilchens sind über mehrere Iterationen hinweg in derselben Farbe dargestellt und durch eine Linie miteinander verbunden. Die schwarze Kurve stellt den besten Zielfunktionswert dar, der von dem gesamten Schwarm bis zur jeweiligen Iteration gefunden wurde. (b) Im Lösungsraum werden die Flugbahnen der verschiedenen Teilchen dargestellt.

Schrittweise Berechnung der Flugbahnen der Schwarmteilchen

In jeder Iteration werden für alle Teilchen neue Positionen durch die Vektorsumme $\mathbf{x}_i^{(k+1)} = \mathbf{x}_i^{(k)} + \delta\mathbf{x}_i^{(k)}$ berechnet. Bei der Berechnung des Flugbahnstücks $\delta\mathbf{x}_i^{(k)}$ orientieren sich die Teilchen einerseits an ihren eigenen bis dahin gefundenen besten Ergebnissen $\mathbf{x}_{B,i}^{(k)}$, andererseits am besten Ergebnis des gesamten Schwarms $\mathbf{x}_B^{(k)}$.

$$\delta\mathbf{x}_i^{(k)} = iW \cdot \delta\mathbf{x}_i^{(k-1)} + ILF \cdot (\mathbf{x}_{B,i}^{(k)} - \mathbf{x}_i^{(k)}) + SLF \cdot (\mathbf{x}_B^{(k)} - \mathbf{x}_i^{(k)}) \quad (5.4)$$

In Gleichung (5.4) tauchen mehrere Algorithmusparameter auf, welche vom Anwender eingestellt werden können: iW (inertia weight), ILF (individual learning factor) und SLF (social learning factor). Durch die Orientierung am Bestvektor $\mathbf{x}_B^{(k)}$ der gesamten Population hängt die mittlere Fluggeschwindigkeit des Schwarms von der Entfernung der anderen Schwarmmitglieder ab: je weiter verstreut der Schwarm, desto schneller fliegen die Teilchen.

Richtungsänderungen können nicht vollkommen abrupt erfolgen, da der *Trägheitsfaktor* iW in die Berechnung der Flugbahn eingeht. Ohne „äußere Kräfte“ durch bessere Suchergebnisse in der zurückliegenden Flugbahn oder durch andere Teilchen des Schwarms, fliegt ein Teilchen mit gleichbleibender Geschwindigkeit auf einer Geraden durch den Lösungsraum. Dieses Trägheits-

prinzip soll eine vorzeitige Konvergenz des Schwarms verhindern und dafür sorgen, dass auch Bereiche des Lösungsraums durchsucht werden, die zunächst zu einer Verschlechterung der Suchergebnisse eines Teilchens führen. Die Wirkung der Trägheit wird bei einer Vergrößerung des „Zieleinlaufs“ der Optimierung aus Abbildung 5.8 (b) deutlich (Abbildung 5.9). Offensichtlich laufen die Teilchen nicht so direkt aufeinander zu, wie es in Abbildung 5.8 (b) erscheint. Insbesondere Teilchen 1 folgt aufgrund des Trägheitsterms beim Annähern des Optimums einer Serpentinbahn.

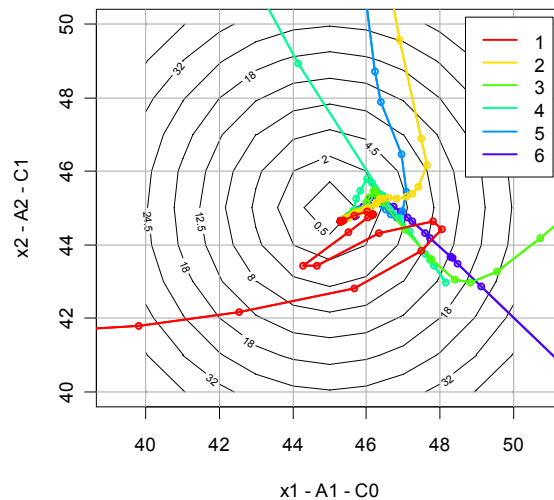


Abbildung 5.9: Vergrößerung des „Zieleinlaufs“ der Schwarmteilchen für die Optimierung aus Abbildung 5.8

Weiterhin zeigt die Vergrößerung, dass die Teilchen deutlich versetzt zum Optimum $\mathbf{x}^* = (45, 45)$ konvergieren. Dieses Ergebnis verweist auf ein grundsätzliches Problem des Algorithmus. Beim Zusammenziehen des Schwarms bremsen die Teilchen ab, d.h. der Beitrag der Trägheit zur Bewegung wird immer kleiner. Abgesehen von der Trägheit ist in dem Verfahren jedoch kein Prinzip angelegt, welches es dem Schwarm ermöglichen würde, Ergebnisse außerhalb der eigenen Suchhistorie zu „sehen“ und sich darauf zuzubewegen. Wird die Trägheit zu klein, konvergieren die Teilchen auf das beste Ergebnis des Schwarms, welches nicht zwangsläufig dem globalen Optimum des Optimierungsproblems entsprechen muss.

5.2.4 Zusammenfassung

Die in diesem Abschnitt gewonnenen Einsichten in die Funktionsweise der drei unterschiedlichen Algorithmen werden im Folgenden kurz zusammengefasst. Dabei wird insbesondere auf die spezifischen Vor- und Nachteile der verschiedenen Verfahren eingegangen.

Complex/Nelder-Mead (CNM)

Vorteil: Durch das Grundprinzip der Reflexion kann sich die Suchmenge sehr effektiv innerhalb des Lösungsraums *bewegen*. So können auch solche Optima gut gefunden werden, die weit vom Startsimplex entfernt liegen (Abbildung 5.5). Das Verfahren kann ebenfalls gut mit *Restriktionen* umgehen, indem der Polyeder sich Restriktionen annähert, neu ausrichtet und sich dann entlang dieser bewegt (Abbildung 5.10). Mit der *Expansionsregel* verfügt das Verfahren über eine sehr effektive Möglichkeit, den Polyeder, d.h. die Suchmenge, zu vergrößern. Dies fördert zum Einen seine Beweglichkeit (s.o.) und wirkt zum Anderen einer vorzeitigen Konvergenz des Polyeders entgegen.

Nachteil: CNM ist ein lokales Optimierungsverfahren. Bei Optimierungsproblemen mit multimodaler Zielfunktion müssen daher grundsätzlich mehrere Optimierungsläufe vorgenommen werden.

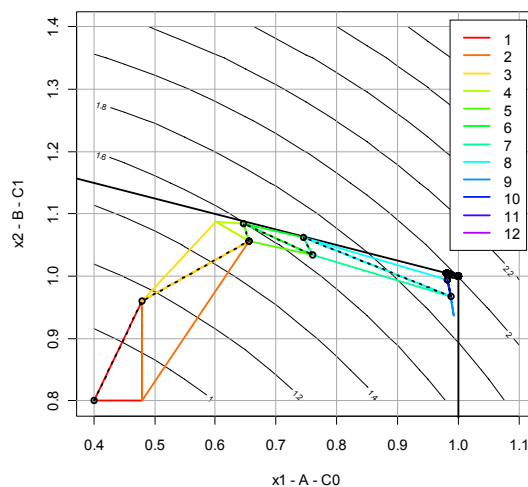


Abbildung 5.10: Verhalten von CNM beim Treffen auf eine Restriktion. Dargestellt ist die Polyeder-Bewegung im Lösungsraum. Die Trajektorie der Bestvektoren $x_B^{(k)}$ ist schwarz gestrichelt dargestellt. Sie beschreibt eine Zickzacklinie entlang der von links oben nach rechts unten verlaufenden Restriktion.

Evolutionäre Strategie (ES)

Vorteil: Die ES ist ein globales Optimierungsverfahren (Syrjakow, 2003: S.21), das prinzipiell robust ist gegenüber Optimierungsproblemen mit ungünstigen analytischen Eigenschaften wie Unstetigkeiten von Zielfunktion und Restriktionen oder sogar einem nicht zusammenhängenden zulässigen Bereich.

Nachteil: Die Suchpopulation ist weniger beweglich im Lösungsraum als z.B. der Polyeder von CNM. Durch die Rekombination alleine können nur Lösungen erzeugt werden, die in unmittelbarer Nähe⁶⁴ der aktuellen Population liegen. Allein das Mutationsprinzip kann zu einer effektiven Expansion bzw. Verschiebung der Population im Lösungsraum führen. Entsprechende „Schritt-richtungen“ sind jedoch zufällig, und Schrittweiten werden umso kleiner, je stärker die Population konvergiert. Diese Eigenschaften der ES können zu Schwierigkeiten führen, wenn sich die Startpopulation M_0 weit entfernt vom Optimum befindet (vgl. Korridormodell § 5.3.4).

Particle Swarm Optimization (PSO)

Vorteil: Das Paradigma der sich unabhängig voneinander bewegenden Teilchen erlaubt grundsätzlich ein sehr effektives Durchsuchen des Lösungsraums. Individuelle Teilchen können sich unabhängig voneinander bewegen und so gleichzeitig unterschiedliche Bereiche des Lösungsraums durchsuchen. Der Trägheitsterm bei der Berechnung neuer Flugrichtungen und -geschwindigkeiten ermöglicht insbesondere auch Suchschritte, mit denen zunächst eine Verschlechterung der Zielfunktion einhergeht.

Nachteil: Ähnlich wie bei ES gibt es kein Prinzip, welches auf Expansion der Suchpopulation angelegt ist. Die „soziale Intelligenz“ der Teilchen beschränkt sich auf die Kenntnis der Position des erfolgreichsten Teilchens des Schwarms. Das macht es – trotz Trägheitsterm – prinzipiell schwierig, gute Lösungen zu finden, die außerhalb der näheren Umgebung der Startpopulation liegen. Weiterhin zeigt sich, dass der Algorithmus sehr empfindlich auf die Einstellungen der Parameter für Trägheit, individuelles und soziales Lernen reagiert. Ungünstige Einstellungen können zu chaotischem Verhalten des Schwarms und einer ausbleibenden Konvergenz führen (Abbildung 5.11). Dies erschwert den Einsatz dieses Algorithmus für unerfahrene Anwender.

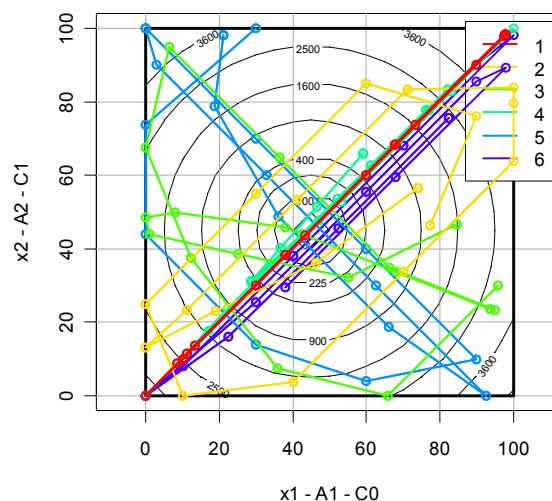


Abbildung 5.11: Chaotisches Verhalten des Schwarms bei einer Optimierung mit PSO.

⁶⁴ Per Konstruktion innerhalb des mehrdimensionalen Quaders, der die Suchmenge gerade umhüllt.

5.3 Performance für verschiedene Testprobleme

Nachdem im vorigen Abschnitt die wesentlichen Eigenschaften der verschiedenen Suchalgorithmen beschrieben und an *beispielhaften* Optimierungsläufen dargestellt wurden, befasst sich der vorliegende Abschnitt mit der *systematischen* Untersuchung der Leistungsfähigkeit der Algorithmen. Dabei wird auf die in Abschnitt 5.1 eingeführten Kennzahlen und Methoden zurückgegriffen.

Eigenschaften der ausgewählten Testprobleme

Für die Optimierungsexperimente wurden vier Standardprobleme für die Beurteilung der Leistungsfähigkeit numerischer Optimierungsverfahren ausgewählt (vgl. z.B. Schwefel, 1995, Peters, 2001, Alt, 2002), die jeweils bestimmte Eigenschaften „realer“ stoffstromnetzbasierter Optimierungsprobleme abbilden. In der Reihenfolge der in diesem Abschnitt diskutierten Testfälle sind dies:

- *Nichtlinearität* der Zielfunktion: Bei SSNOPs ist dies die Folge entweder (i) nichtlinearer Transitionsspezifikationen oder (ii) der Optimierung nichtlinearer Zielgrößen, beispielsweise einer Effizienz, bei deren Definition Quotienten von Stoffströmen auftauchen.
- *Randlösungen*: Bei realen Produktionssystemen kann davon ausgegangen werden, dass Optima häufig Randlösungen sind, d.h. dass im Optimum Ungleichungsrestriktionen bezüglich bestimmter Systemparameter bindend werden.
- *Multimodalität*: Die Nichtlinearität einer Zielfunktion geht häufig mit dem Vorhandensein *mehrerer* lokaler Optima einher. Wie das Testproblem nach Bracken und McCormick zeigt (§ 5.3.2), kann Multimodalität auch bei einer an sich unimodalen Zielfunktion durch das Vorhandensein entsprechender Restriktionen entstehen.
- *Quasi-Reduktion der effektiven Dimensionalität*: Dicht beieinander liegende Ober- und Untergrenzen für einzelne Variablen können dazu führen, dass die Ausdehnung des zulässigen Bereichs in verschiedenen Richtungen des Lösungsraums sehr unterschiedlich ausfällt.

Bewertung der Treffhäufigkeit in den nachfolgenden Tests

Die Treffhäufigkeit der verschiedenen Algorithmen wird, wie in § 0 dargestellt, bewertet. Zur Beurteilung der Treffhäufigkeit wird als Obergrenze für die Approximationsgenauigkeit $z_B < 10^{-3}$ festgelegt. Eine Optimierung muss den Abstand des anfänglichen Zielfunktionswertes $f(\mathbf{x}^{(0)})$ vom Optimalwert somit um mehr als das 1000fache reduzieren, um als erfolgreich zu gelten.

5.3.1 Sphärenmodell

Dieses *einfache quadratische Optimierungsproblem* (Tabelle 5.3, Abbildung 5.12) verwendet Schwefel, um die Konvergenzeigenschaften der von ihm untersuchten numerischen Optimierungsverfahren zu testen.

Tabelle 5.3: Sphärenmodell: Definition und wesentliche Eigenschaften

Ziel	$\min f(x) = \sum_{i=1}^n (x_i - b)^2$
Restriktionen	explizite Ober und Untergrenzen $0 \leq x_i \leq 100, \quad i = 1, \dots, n$
Optima	ein globales (M1) Minimum: $\mathbf{x}^* = (b, \dots, b)^T, f^* = 0$
Quelle	(Schwefel, 1995: S.179-204)
wesentliche Eigenschaften	<ul style="list-style-type: none"> • nichtlineare Zielfunktion (quadratisch) • ein globales Optimum • ohne weiteres auf beliebig viele Dimensionen erweiterbar.

Es wurden zwei Anpassungen vorgenommen. Das ursprünglich unrestringierte Problem wird um explizite Ober- und Untergrenzen für alle Entscheidungsvariablen ergänzt. Die Untergrenzen ergeben sich aus der Logik der Stoffstromanalyse-Software, da alle Entscheidungsvariablen als Stoffflüsse modelliert werden und daher positiv sein müssen. Um dennoch eine *innere Lösung* zu bekommen, wird der Offsetparameter b eingeführt. In hinreichender Entfernung der Suchpopulation von den Ober- und Untergrenzen, insbesondere also bei Konvergenz gegen das innere Optimum, ist das resultierende Optimierungsproblem dem bei Schwefel verwendeten äquivalent. Die mit dem Sphärenmodell gewonnenen Ergebnisse können daher unmittelbar mit denen von Schwefel verglichen werden.

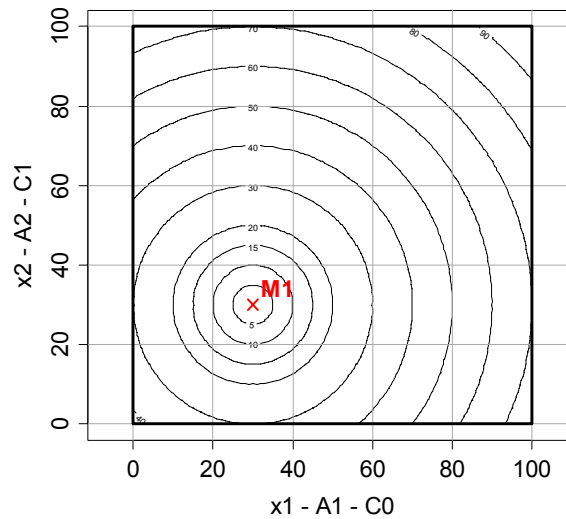


Abbildung 5.12: Lösungsraum des Sphärenmodells mit Isoquanten der Zielfunktion

CNM schneidet sowohl hinsichtlich Treffhäufigkeit (100%) als auch Optimierungsdauer deutlich besser ab als ES und PSO (Abbildung 5.13). Hinsichtlich beider Leistungskriterien ergibt sich für die Lösung des Sphärenmodells damit eine klare Rangfolge in der Eignung der untersuchten Algorithmen: CNM > ES > PSO.

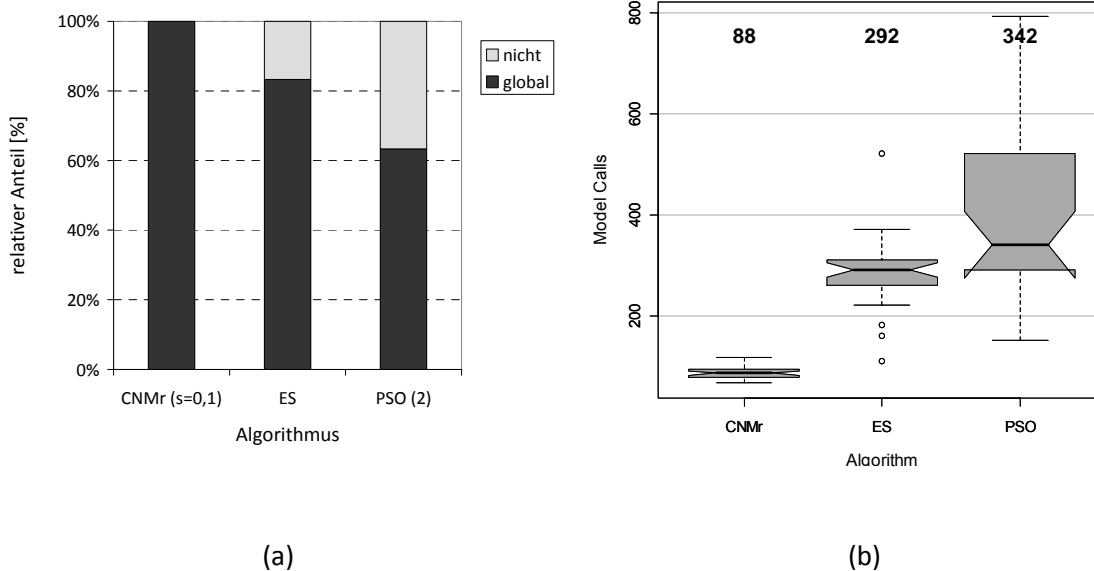


Abbildung 5.13: Performance für Sphärenmodell. (a) Treffhäufigkeit (b) Optimierungsdauer: Lage des Median im Diagramm angegeben.

5.3.2 Randlösungen – Testproblem nach Bracken und McCormick

Tabelle 5.4: Bracken-McCormick-Testproblem: Definition und wesentliche Eigenschaften

Ziel	$\max f(x) = x_1^2 + x_2^2$		
Restriktionen	Ober und Untergrenzen einzelner Entscheidungsvariablen: $0 \leq x_1 \leq 1, 0 \leq x_2 \leq 2$ gemischte Restriktion: $h_1(x) = 5 - x_1 - 4x_2 \geq 0$		
Optima	ein globales (M1) und ein lokales Maximum:		
	Optimum	x_1	x_2
	M1	1	1
	M2	0	1,25
Quelle	(Schwefel, 1995: S.348, Peters, 2001: S.118)		
wesentliche Eigenschaften	<ul style="list-style-type: none"> • nichtlineare Zielfunktion (quadratisch) • ein globales und ein lokales Optimum • beide Optima sind Randlösungen, bei denen jeweils zwei Restriktionen bindend sind. • 2D, nicht ohne weiteres erweiterbar. 		

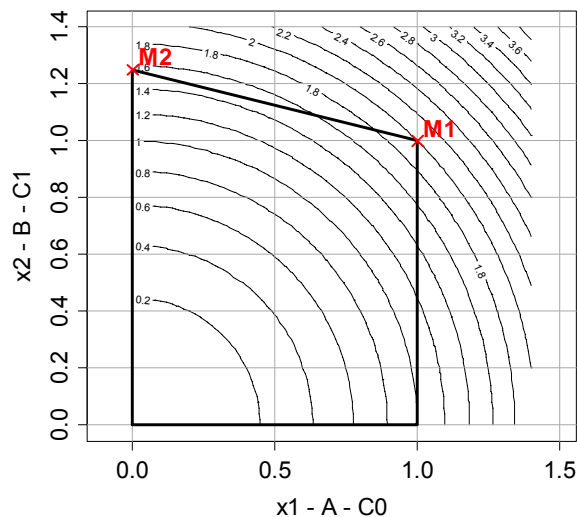


Abbildung 5.14: Testproblem nach Bracken und McCormick.

Hinsichtlich der beiden Leistungskriterien Treffhäufigkeit und Optimierungsdauer ergibt sich beim Vergleich der Algorithmen bei der Lösung des Bracken-McCormick-Testproblems (Tabelle 5.4, Abbildung 5.14) ein ähnliches Bild wie beim Sphärenmodell: CNM > ES > PSO.

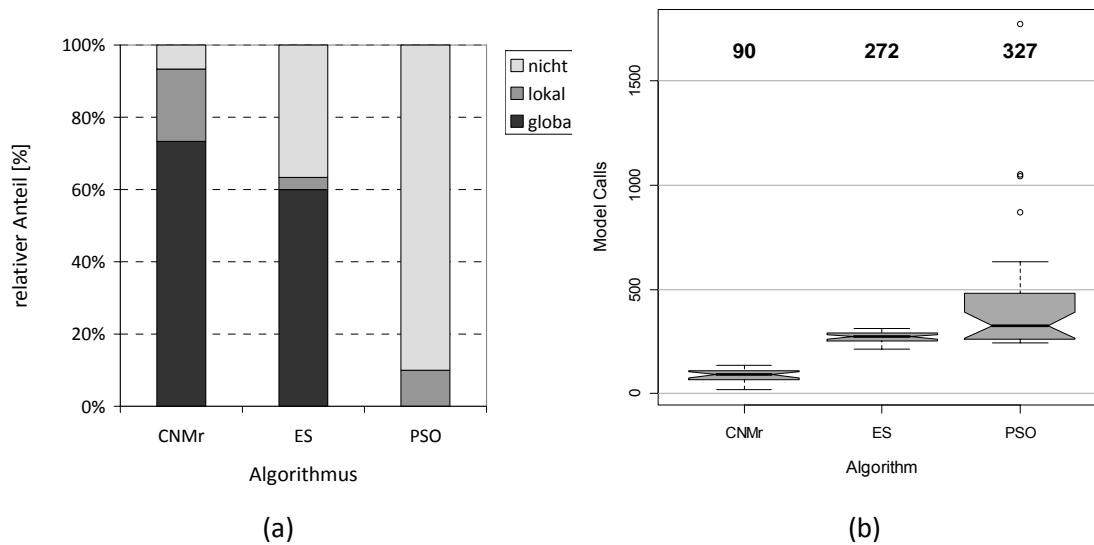


Abbildung 5.15: Performance für Bracken-McCormick-Testproblem. (a) Trefferhäufigkeit (b) Optimierungsdauer: Lage des Median im Diagramm angegeben.

Allerdings kann CNM bei diesem Testproblem keine 100%ige Treffhäufigkeit mehr erzielen. Der Grund hierfür sind zwei lokale Minima auf dem Rand des zulässigen Bereichs bei $\mathbf{x} = (1,0)$ und $\mathbf{x} \cong (1.18, 0.29)$. An diesen beiden Stellen ist der Gradient der Zielfunktion senkrecht zur Restriktion. Die entsprechend in der unmittelbaren Umgebung dieser lokalen Extrema nur schwach ausgeprägte Komponente des Gradienten parallel zur Restriktion ist für den am Rande des zulässigen Bereichs auf kleinem Raum zusammengezogenen Polyeder offenbar zu schwach, um ihn in Richtung eines der beiden gesuchten Maxima zu „ziehen“.

Das Ergebnis von ES ist mit einer globalen Konvergenz von 60% noch zufriedenstellend. Der Einschätzung liegt folgende Überlegung zugrunde: Die Treffhäufigkeit kann als Trefferwahrscheinlichkeit p für einen beliebigen Optimierungslauf interpretiert werden. Gemäß der Bernoulli-Verteilung für binäre Entscheidungsvariablen kann die Wahrscheinlichkeit der globalen Konvergenz des Verfahrens durch *Mehrfachstart* mit unterschiedlichen Startpopulationen sehr effektiv gesteigert werden. Bei m Optimierungen steigt die Wahrscheinlichkeit, das globale Optimum zu finden, auf $1 - (1 - p)^m$. Mit einer Treffhäufigkeit von $p = 60\%$ ergibt das bei fünf Optimierungen bereits eine Trefferwahrscheinlichkeit von 98,9%. Insofern sind Treffhäufigkeiten $< 100\%$ per se kein schlechtes Ergebnis. Allerdings führen Mehrfachstarts natürlich zu einer multiplikativen Verlängerung der Optimierungsdauer. Die Treffhäufigkeit kann somit nur auf Kosten des Optimierungsaufwands gesteigert werden.

PSO hat mit dem Testfall offensichtlich ein prinzipielles Problem (Abbildung 5.15a). Nur 10% aller Optimierungen konvergieren gegen eine optimale Lösung, und diese zudem gegen das lokale Optimum. Ein genauere Blick auf die Ergebnisse der einzelnen Optimierungsläufe zeigt jedoch, dass es sich weniger um ein Problem der Treffhäufigkeit, sondern mehr um eines der Approxima-

tionsgenauigkeit handelt. Tatsächlich konvergiert das Verfahren in den meisten Fällen gegen das lokale oder globale Optimum. Dies jedoch stets mit einer schlechteren Genauigkeit als der für diese Untersuchung festgelegten Untergrenze von 10^{-3} . Deshalb werden die Optimierungen nicht als Treffer gewertet. Die schlechte Treffhäufigkeit ist somit das Ergebnis der in § 5.2.3 beschriebenen Problematik der vorzeitigen Konvergenz des Schwarms in der Nähe der optimalen Lösung, da die Trägheit der einzelnen Schwarmteilchen zu klein wird.

5.3.3 Multimodalität – Testproblem nach Himmelblau

Tabelle 5.5: Modifiziertes Himmelblau-Testproblem: Definition und wesentliche Eigenschaften

Ziel	$\min f(x_1, x_2) = f_H(x_1 - a, x_2 - a) + (x_1 + x_2)$ <p>$f_H(\mathbf{x}) = (x_1^2 + x_2 - 11)^2 + (x_1 + x_2^2 - 7)^2$ ist die ursprüngliche Himmelblau-Funktion. $a = 5$ ist ein Offsetparameter, mit dem bei diesem modifizierten Testproblem alle vier Optima in den positiven Quadranten verschoben werden. Durch den Driftterm $(x_1 + x_2)$ wird die Entartung der Minima der Himmelblau-Funktion aufgehoben und ein eindeutiges globales Minimum (M1) definiert.</p>																							
Restriktionen	explizite Ober und Untergrenzen: $0 \leq x_1, x_2 \leq 10$																							
Optima	ein globales (M1) und drei lokale Minima: <table border="1" style="margin-left: 20px;"> <thead> <tr> <th>Optimum</th> <th>x1</th> <th>x2</th> <th>f(x)</th> </tr> </thead> <tbody> <tr> <td>M1</td> <td>1,2085</td> <td>1,7017</td> <td>2,9238</td> </tr> <tr> <td>M2</td> <td>2,1798</td> <td>8,1190</td> <td>10,3125</td> </tr> <tr> <td>M3</td> <td>7,9935</td> <td>6,9739</td> <td>14,9838</td> </tr> <tr> <td>M4</td> <td>8,5769</td> <td>3,1203</td> <td>11,7166</td> </tr> </tbody> </table>				Optimum	x1	x2	f(x)	M1	1,2085	1,7017	2,9238	M2	2,1798	8,1190	10,3125	M3	7,9935	6,9739	14,9838	M4	8,5769	3,1203	11,7166
Optimum	x1	x2	f(x)																					
M1	1,2085	1,7017	2,9238																					
M2	2,1798	8,1190	10,3125																					
M3	7,9935	6,9739	14,9838																					
M4	8,5769	3,1203	11,7166																					
Quelle	(Alt, 2002: S.12)																							
wesentliche Eigenschaften	<ul style="list-style-type: none"> • nichtlineare Zielfunktion (Polynom 4. Grades) • multimodal (mehrere Optima) • alle Optima liegen im Inneren der konvexen Lösungsmenge • 2D, nicht ohne weiteres erweiterbar. 																							

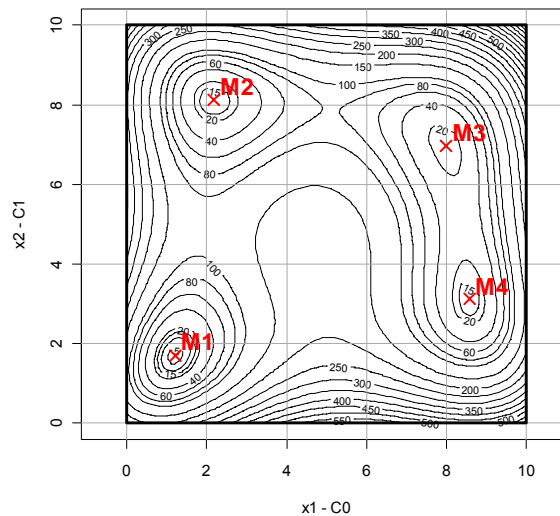


Abbildung 5.16: Modifiziertes Testproblem nach Himmelblau

Abbildung 5.17 stellt die Performance der getesteten Algorithmen in der gleichen Weise dar wie in den vorigen Abschnitten. Zur weitergehenden Differenzierung zwischen den verschiedenen lokalen Optima beim Himmelblau-Testproblem werden die Ergebnisse zusätzlich tabellarisch aufgeführt (Tabelle 5.6).

Tabelle 5.6: Treffhäufigkeit der verschiedenen Algorithmen für das Himmelblau-Testproblem. Die Spalte „lokal“ weist die Summe der Optimierungsläufe aus, die gegen ein lokales Minimum konvergiert sind.

	M1	M2	M3	M4	lokal	lokal und global
CNMr	30,0	30,0	30,0	10,0	70,0	100,0
ES	0,0	6,7	20,0	36,7	63,3	63,3
PSO	6,7	0,0	0,0	10,0	10,0	16,7

Sowohl ES als auch PSO schneiden bei der globalen und der lokalen Treffhäufigkeit deutlich schlechter ab als CNMr. Das Ergebnis steht im Widerspruch beispielsweise zu den numerischen Resultaten von Peters et al. (1999), bei denen die ES grundsätzlich besser abschneidet als der Complex- bzw. Simplexalgorithmus. Im Gegenzug fällt jedoch auf, dass ES bei der zitierten Untersuchung für alle Testprobleme mindestens um das 100fache längere Optimierungsdauern aufweist als Complex und Simplexalgorithmus. Daher muss angenommen werden, dass wesentlich größere Populationen verwendet und dadurch eine höhere Approximationsgenauigkeit erreicht wurde. Bei den Testreihen dieses Abschnitts umfasste die Populationsgröße für PSO und ES lediglich 10 Individuen. Dieser Wert ist im Sinne der Approximationsgenauigkeit sicher nicht optimal. Die Ergebnisse ließen sich sehr wahrscheinlich verbessern, wenn größere Suchmengen verwendet würden. Das würde jedoch auch zu längeren Optimierungsdauern führen (vgl. §5.5.3).

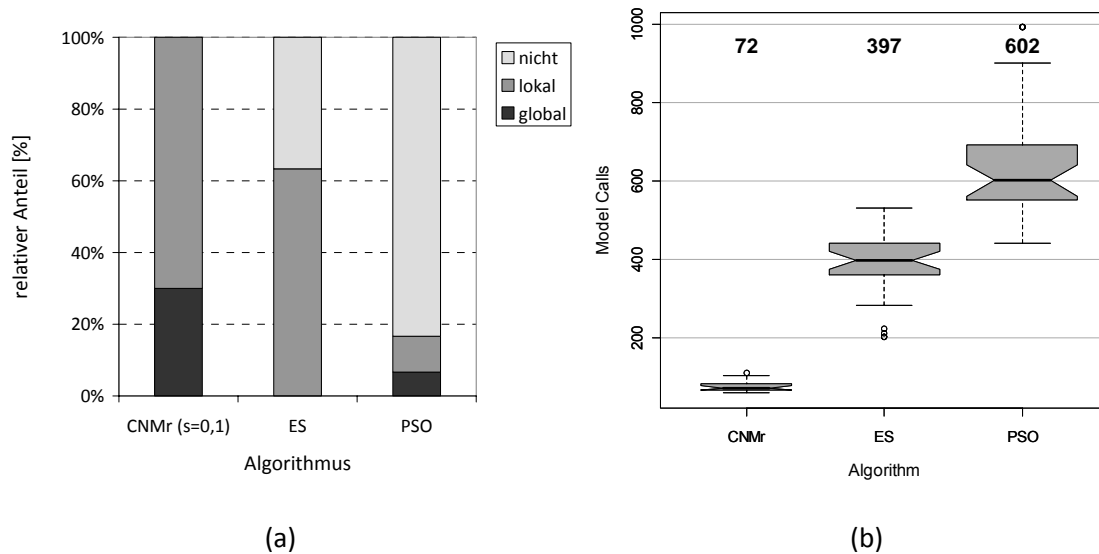


Abbildung 5.17: Performance für Himmelblau-Testproblem. (a) Trefferhäufigkeit (b) Optimierungsdauer: Lage des Median im Diagramm angegeben.

Einfluss des Startvektors auf das Optimierungsergebnis

Anhand der Testreihen für das Himmelblau-Testproblem können einige interessante Eigenschaften der verschiedenen in diesem Kapitel untersuchten Optimierungsverfahren aufgezeigt werden. Wie in Abschnitt 5.2 ausgeführt, gehen alle Optimierungsverfahren von einer *anfänglichen Suchmenge* M_0 aus. Diese beinhaltet neben dem vorgegebenen Startvektor $\mathbf{x}^{(0)}$ weitere Testlösungen, die bei allen Verfahren zufällig im zulässigen Bereich ausgewählt werden. Genau genommen ist dem Suchprinzip des jeweiligen Algorithmus eine *Monte-Carlo-Suche* vorgeschaltet⁶⁵. Dies führt dazu, dass der Startvektor das Optimierungsergebnis weit weniger beeinflusst als man zunächst meinen könnte. Abbildung 5.18 zeigt am Beispiel mehrerer CNMr-Optimierungen, dass von ein und demselben Startvektor, je nach Initialisierung des Zufallsgenerators für den Aufbau der Startsuchmenge, sehr unterschiedliche Optimierungspfade ausgehen können. In zwei Fällen spielt der Startvektor dennoch eine ausgezeichnete Rolle: (i) bei Verwendung des Skalierungsfaktors s der CNM-Implementierung von KOMSA (vgl. § 5.5.2), welcher die Monte-Carlo-Suche auf eine beliebig kleine Umgebung des Startvektors beschränkt. (ii) Wenn Restriktionen dazu führen, dass die Monte-Carlo-Suche effektiv auf einen kleinen Bereich um den Startvektor beschränkt wird (vgl. Korridormodell § 5.3.4).

⁶⁵ Die einzige Ausnahme ist der vollständig deterministische CNMs-Algorithmus.

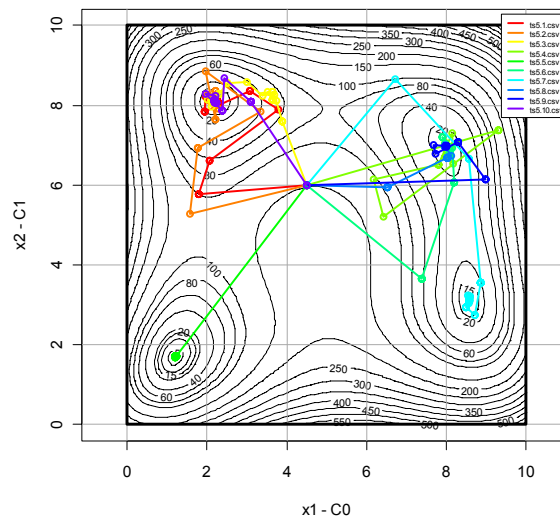


Abbildung 5.18: Zehn CNMr-Läufe ($s=1$) mit gleichem Startpunkt aber unterschiedlichen Initialisierungen des Zufallsgenerators für den Aufbau des Startpolyeders.

Vergleich der Optimierungstrajektorien von ES und CNM

Der Vergleich verschiedener Optimierungstrajektorien $\mathbf{x}_B^{(k)}$ von CNM und ES in Abbildung 5.19 illustriert gut die grundsätzlich unterschiedlichen Optimierungsprinzipien. Während CNM stets gegen ein bestimmtes Minimum konvergiert, springt ES mehrfach zwischen den Einzugsbereichen der Optima M1, M2 und M4 hin und her. Da die Startpopulation anfangs über den gesamten zulässigen Bereich verteilt wird, kann das Verfahren quasi simultan gegen mehrere Optima konvergieren. Hieran zeigt sich, dass ES prinzipiell eine globale Optimierungsstrategie ist.

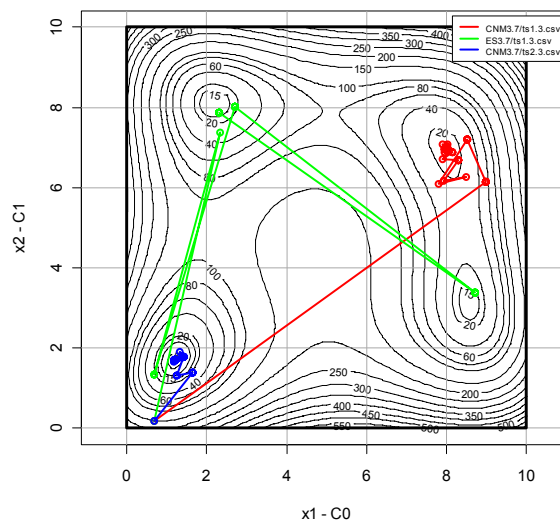


Abbildung 5.19: CNMr und ES im Vergleich. Drei verschiedene Optimierungsläufe, die jeweils von derselben Startlösung ausgehen. CNMr mit Skalierungsfaktor $s = 1$: (rot) ES (grün) CNMr mit $s = 0,1$ (blau).

5.3.4 Beweglichkeit im Lösungsraum – Korridormodell

Das Korridormodell (Tabelle 5.7, Abbildung 5.20) führt zum deutlichsten Auseinanderfallen der Performanceergebnisse von CNM auf der einen und ES und PSO auf der anderen Seite. Während CNM in der Lage ist, das Problem mit 100%iger Treffhäufigkeit zu lösen, stoßen ES und PSO offensichtlich auf grundsätzliche Schwierigkeiten.

Tabelle 5.7: Korridormodell: Definition und wesentliche Eigenschaften

Ziel	$\max f(\mathbf{x}) = x_2 + (1 + \gamma) \cdot x_1$
Restriktionen	<p>Ober und Untergrenzen einzelner Entscheidungsvariablen:</p> $0 \leq x_1, x_2 \leq 5$ <p>gemischte Restriktionen:</p> $h_1(\mathbf{x}) = x_1 + x_2 - 3 \geq 0$ $h_2(\mathbf{x}) = (3 + \varepsilon) - x_1 - x_2 \geq 0$
Optima	<p>ein globales Maximum:</p> $\mathbf{x}^* = (3 + \varepsilon, 0)^T, f(\mathbf{x}^*) = (1 + \gamma) \cdot (3 + \varepsilon) \quad \text{für } \gamma > 1.$
Quelle	Eigenkonstruktion
wesentliche Eigenschaften	<ul style="list-style-type: none"> • lineares Programm (LP) mit eindeutigem globalem Optimum • Lokaler Aufbau des Startpolyeders um den Startvektor $\mathbf{x}^{(0)}$ kann für alle Verfahren durch $\varepsilon \rightarrow 0$ erzwungen werden.

Die Lösungsraumgeometrie des Korridormodells wird wesentlich von den beiden gemischten Restriktionen h_1 und h_2 geprägt. Diese verlaufen im Abstand ε parallel zueinander. Für die Abbildung 5.20 wurde $\varepsilon = 1$ relativ groß gewählt. Der zulässige Bereich ist zwar länglich, aber eindeutig eine zweidimensionale Fläche. Für $\varepsilon \rightarrow 0$ konvergiert er jedoch gegen die „untere“ Restriktion $h_1(\mathbf{x}) = x_1 + x_2 - 3$ und die Dimensionalität des Problems wird praktisch um eins reduziert.

Dies erzwingt einen lokalen Aufbau der Startmenge M_0 um den Startvektor $\mathbf{x}^{(0)}$. Durch die Wahl von $\mathbf{x}^{(0)}$ kann außerdem erreicht werden, dass sich die Startmenge, verglichen mit ihrer Ausdehnung, weit entfernt vom Optimum befindet ($|\mathbf{x}^{(0)} - \mathbf{x}^*| \gg \varepsilon$). Mit diesem Testproblem lässt sich daher gezielt die Beweglichkeit der Suchmengen im Lösungsraum überprüfen. Abbildung 5.21(a) bestätigt in drastischer Weise die aus einer Analyse der Funktionsweise der verschiedenen Suchstrategien in Abschnitt 5.2 gezogene Schlussfolgerung: weder ES noch PSO verfügen über effektive Mechanismen, um sich bei ihrer Suche im Lösungsraum weit über die durch M_0 vorgegebene

Umgebung hinaus zu bewegen. Hierin liegt die Ursache für die sehr niedrigen Treffhäufigkeiten (<10%) beider Verfahren.

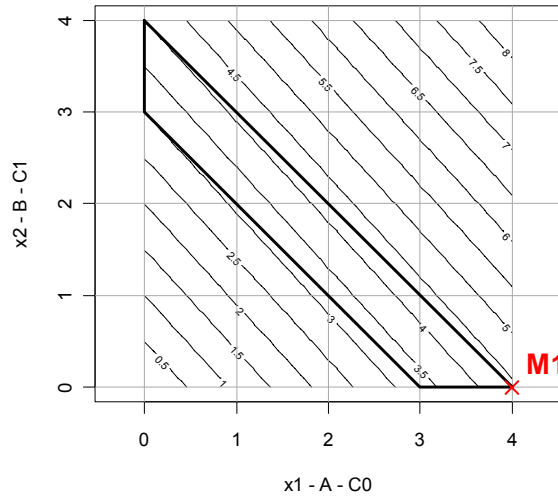


Abbildung 5.20: Das Korridormodell.

Wiederum kann CNM das Testproblem am schnellsten lösen. Allerdings konvergiert in diesem Fall ausnahmsweise PSO besser als ES. Dieses Ergebnis ist angesichts der winzigen Treffhäufigkeit jedoch kaum von Bedeutung.

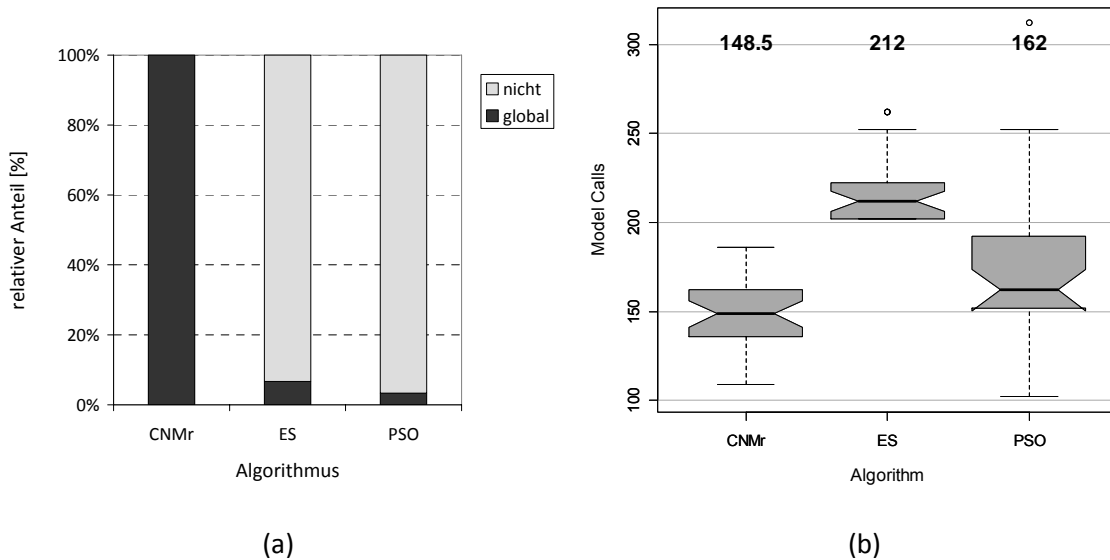


Abbildung 5.21: Performance für Korridormodell. (a) Treffhäufigkeit (b) Optimierungsdauer: Lage des Median im Diagramm angegeben.

Im Vergleich mit den anderen Testproblemen fällt auf, dass CNM im Mittel gut 50% länger braucht, um die optimale Lösung zu finden. Offensichtlich stellt das Korridormodell für alle Algo-

rithmen eine Herausforderung dar, obwohl es als lineares Programm im Sinne der üblichen Klassifizierung von Optimierungsmodellen im Operations Research zu den am besten lösbarsten Problemen gehört.

5.4 Abhängigkeit der Konvergenzgeschwindigkeit von der Anzahl der Freiheitsgrade

Die Abhängigkeit der Optimierungsdauer von der Dimensionalität des Optimierungsproblems wird wie bei Schwefel (1995: S.179-204) anhand des Sphärenmodells aus § 5.3.1 ermittelt. Das Sphärenmodell ist der einfachste Fall eines quadratischen Optimierungsproblems und kann mit beliebig vielen unabhängigen Entscheidungsvariablen definiert werden. Die aus den numerischen Tests mit diesem Modell gewonnenen Ergebnisse entsprechen der so genannten Q-Eigenschaft, die für manche analytischen Optimierungsverfahren auch theoretisch hergeleitet werden kann. Während die Q-Eigenschaft die Anzahl der Iterationen angibt, die für die exakte Bestimmung des Optimums notwendig sind, muss im Fall der numerischen Optimierung natürlich eine Genauigkeitsanforderung vorgegeben werden. Der besseren Vergleichbarkeit wegen wird hier, wie bei Schwefel (a.a.O), festgelegt, dass die Anzahl der Modellaufrufe ermittelt wird, bis das Optimum auf weniger als 1% des anfänglichen Abstands vom Optimum angenähert wurde ($z_B \leq 10^{-2}$).

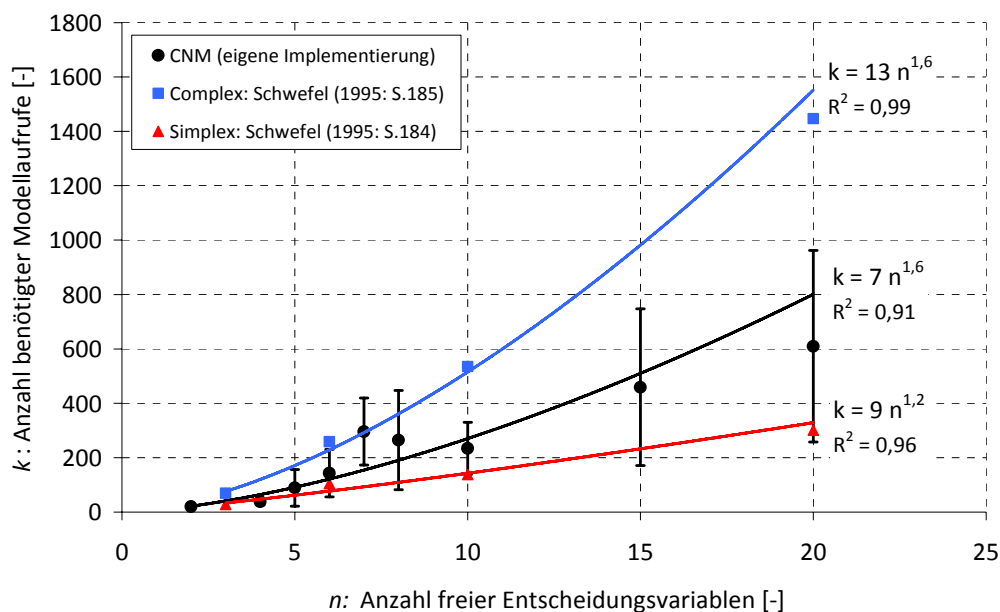


Abbildung 5.22: CNM, Optimierungsdauer in Abhängigkeit von der Dimension des Sphärenmodells. Balken geben die Streubereiche aus den numerischen Tests an. Außerdem abgebildet: Testergebnisse von Schwefel (1995:184f) für den Complex- (blau) und den Nelder-Mead-Algorithmus (rot).

Abbildung 5.22 zeigt die numerischen Testergebnisse für CNM zusammen mit den von Schwefel (1995: S.184f) jeweils für die Simplex- und Complex-Strategie ermittelten Daten. Im Vergleich schneidet CNM insgesamt besser ab als der Complex-Algorithmus, benötigt im Mittel jedoch durchweg mehr Modellauswertungen als das Simplex-Verfahren. Für den untersuchten Bereich zeigen CNM und Complex-Algorithmus dasselbe asymptotische Anwachsen der Optimierungsdauer mit der Dimensionalität des Optimierungsproblems: $O(n^{1,6})$.

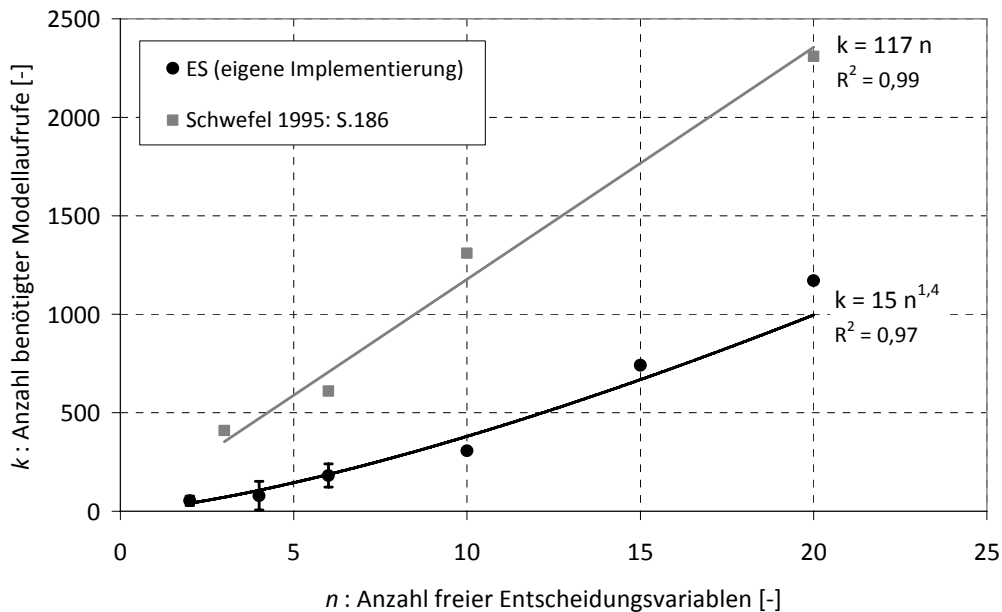


Abbildung 5.23: ES, Optimierungsdauer in Abhängigkeit von Dimension des Optimierungsmodells.

Die in KOMSA implementierte Evolutionsstrategie (ES) findet die Problemlösung durchweg schneller als die in den numerischen Tests von Schwefel (1995: S.185) verwendete Implementierung (Abbildung 5.23). Aus diesem Ergebnis kann jedoch nicht unmittelbar der Schluss gezogen werden, dass die KOMSA-Implementierung grundsätzlich effizienter ist. Die Optimierungsdauer wird bei der ES maßgeblich durch die Populationsgröße beeinflusst (vgl. § 5.5.3). Für die hier durchgeführten Testreihen wurden die Populationsgrößen an die Dimensionalität des Problems angepasst. Die Populationen bestanden jeweils aus $1,5 n$ Individuen. D.h. im höchstdimensionalen der untersuchten Testprobleme mussten daher je Iteration 30 neue Testlösungen erzeugt werden. Bei Schwefel (a.a.O) wurden für alle Problemgrößen gleichbleibend 100 Individuen verwendet. Entscheidend ist, wie im Falle der Hillclimbing-Strategien, der Vergleich des asymptotischen Verhaltens der Optimierungsdauer mit der Problemgröße. Und hier sieht man, dass die Optimierungsdauern beider Verfahren mit $O(n)$ anwachsen. In diesem Sinne erweisen sich also beide Implementierungen als gleich schnell.

5.5 Einfluss der wichtigsten Algorithmusparameter

Jeder Algorithmus verfügt über *spezifische Parameter*, mit denen Einfluss auf den Verlauf konkreter Optimierungen genommen werden kann. Diese Parameter werden über einen entsprechenden Dialog im Leitstand eingestellt. Eine systematische und statistisch fundierte Untersuchung der Wirkung aller Parameter würde den Rahmen dieser Arbeit sprengen und soll an dieser Stelle daher nicht vorgenommen werden. Anstatt dessen wird die Wirkungsweise der wichtigsten Parameter exemplarisch dargestellt und diskutiert. Als „wichtig“ werden dabei solche Parameter betrachtet, die in den vielen Optimierungsexperimenten dieses Kapitels eine deutliche und eindeutige Wirkung auf die jeweiligen Optimierungsverläufe gezeigt haben.

5.5.1 Abbruchkriterium

Einen starken Einfluss auf das Optimierungsergebnis hat das Abbruchkriterium, das im Optimierungsleitstand unabhängig vom jeweils verwendeten Suchverfahren festgelegt wird. Bei direkten Suchverfahren gibt es zwei gebräuchliche Varianten. Beim Complex-Algorithmus wird es beispielsweise als hinreichendes Kriterium für die Konvergenz des Verfahrens interpretiert, wenn die Standardabweichung (5.5) in aufeinanderfolgenden Iterationen gefundener Bestlösungen $f_B^{(k)}$ einen vorgegebenen Schwellenwert S^{\min} unterschreitet: $S_f^{(k)} \leq S^{\min}$. Neben dem Schwellenwert muss im Leitstand festgelegt werden, wie viele Iterationen in die Berechnung der Standardabweichung eingehen sollen.

$$S_f^{(k)} = \sqrt{\frac{1}{n_I - 1} \cdot \sum_{\kappa=k-n_I}^k (f_B^{(\kappa)} - \bar{f}_B)^2} \quad \text{mit } \bar{f}_B = \frac{1}{n_I} \sum_{\kappa=k-n_I}^k f_B^{(\kappa)} \quad (5.5)$$

Die Analyse numerischer Optimierungen in diesem Kapitel zeigt jedoch, dass in diesem Sinne auch Situationen als Konvergenz *erscheinen* können, bei denen ein Algorithmus zwar über mehrere Iterationen keine Verbesserung der Optimallösung $f_B^{(k)}$ erzielt, die Suchmenge jedoch keineswegs gegen einen einheitlichen Wert konvergiert, beispielsweise bei der Neuausrichtung eines CNM-Polyeders an einer Restriktion. Daher bietet sich ein anderes Konvergenzkriterium an, das sich auf die Standardabweichung der Lösungen der verschiedenen Elemente der Suchmenge M_k bezieht.

$$S_x^{(k)} = \sqrt{\frac{1}{|M| - 1} \cdot \sum_{i=1}^{|M|} (f_i^{(k)} - \bar{f}_i^{(k)})^2} \quad \text{mit } \bar{f}_i^{(k)} = \frac{1}{|M|} \sum_{i=1}^{|M|} f_i^{(k)} \quad (5.6)$$

Dies bedeutet anschaulich, dass eine Untergrenze für die Streubreite der Suchmenge angegeben wird, bei deren Unterschreitung das Optimierungsverfahren abbricht. Dieses Abbruchkriterium

wurde von Nelder und Mead (1965) für ihr Simplex-Verfahren und in der Implementierung der Evolutionsstrategie nach Schwefel (1995: S.146) verwendet.

5.5.2 CNM – Skalierung des Startpolyeders

Optimierungen mit CNM zeigen keine ausgeprägte Abhängigkeit von der Einstellung der spezifischen Parameter. Bereits Box (1965: S.45) berichtet, dass die Optimierungsgüte nicht signifikant von der Wahl des Reflexionsfaktors α abhängt. In numerischen Tests wurde dieser zwischen 1 und 1,5 variiert. Die Voreinstellungen der KOMSA-Implementierung für Reflexions-, Kontraktions- und Expansionsfaktoren $(\alpha; \beta; \gamma) = (1,1; 0,5; 2)$ führen in allen Testproblemen – im Rahmen der statistischen Schwankungsweite – zu guten Optimierungsergebnissen. Der Parameter mit dem größten Einfluss auf den Optimierungsverlauf ist der Skalierungsfaktor s für den Startpolyeder. Mit diesem Skalierungsfaktor kann der Aufbau des Startpolyeders gemäß (5.7) auf eine lokale Umgebung um den Startvektor $\mathbf{x}_1^{(0)}$ festgelegt werden.

$$[x_j^{(0)}]_i = [x_1^{(0)}]_i + r \cdot s \cdot (u_i^x - l_i^x) \quad \text{mit } i = 1 \dots n; j = 2 \dots \mu \quad (5.7)$$

$s \in (0,1)$ und r gleichverteilt in $(-1,1)$

In Abbildung 5.24 sind zehn Optimierungsläufe für das Himmelblau-Testproblem aus § 5.3.3 mit $s = 1$ abgebildet⁶⁶.

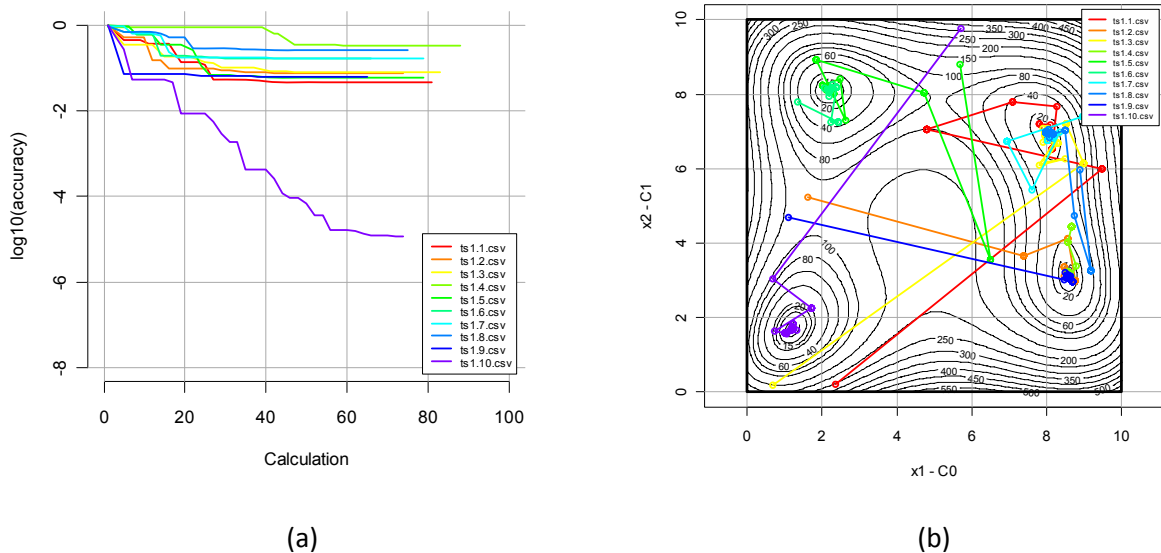


Abbildung 5.24: Zehn CNM-Läufe mit Standardeinstellungen für das Himmelblau-Testproblem.

⁶⁶ das entspricht der Voreinstellung im Leitstand.

Es fällt auf, dass die Lage der Startlösung keinen maßgeblichen Einfluss darauf hat, gegen welches Minimum der Algorithmus konvergiert: der Startpunkt des einzigen Laufes, welcher gegen das globale Optimum M1 konvergiert ($ts1.10$ = lila Kurve), liegt deutlich näher am lokalen Minimum M2 links oben. Andererseits liegen die Startpunkte der Läufe $ts1.1$ und $ts1.3$ im Einzugsbereich von M1, konvergieren jedoch gegen das deutlich entfernter liegende Minimum M3.

Dieses zunächst überraschende Verhalten lässt sich mit dem Aufbau des Startpolyeders leicht erklären. In Abbildung 5.25(a) ist für den erfolgreichen Lauf $ts1.10$ die Bewegung des Polyeders durch den Lösungsraum dargestellt. Die zufällig erzeugten Ecken des Startpolyeders, d.h. alle bis auf den vorgegebenen Startvektor, werden innerhalb des gesamten zulässigen Bereichs ausgewählt. Die Ausdehnung des so konstruierten Startpolyeders ist daher groß verglichen mit der räumlichen Varianz der Zielfunktion. Mit anderen Worten, der Startpolyeder (rot = erste Iteration) überdeckt ein Gebiet mit sehr unterschiedlichen Richtungen des Gradienten. Anschaulich gesprochen: die Höhenlinien verlaufen in seinem Inneren nicht parallel. Daher ergibt sich auch keine eindeutige Suchrichtung. Das heuristische Prinzip, nach dem CNM iterativ neue Polyeder erzeugt, führt in dieser Situation zu einer Konvergenz gegen das Minimum, welches der Ecke des Startpolyeders mit dem besten Zielfunktionswert am nächsten liegt, in diesem Fall also das globale Minimum M1.

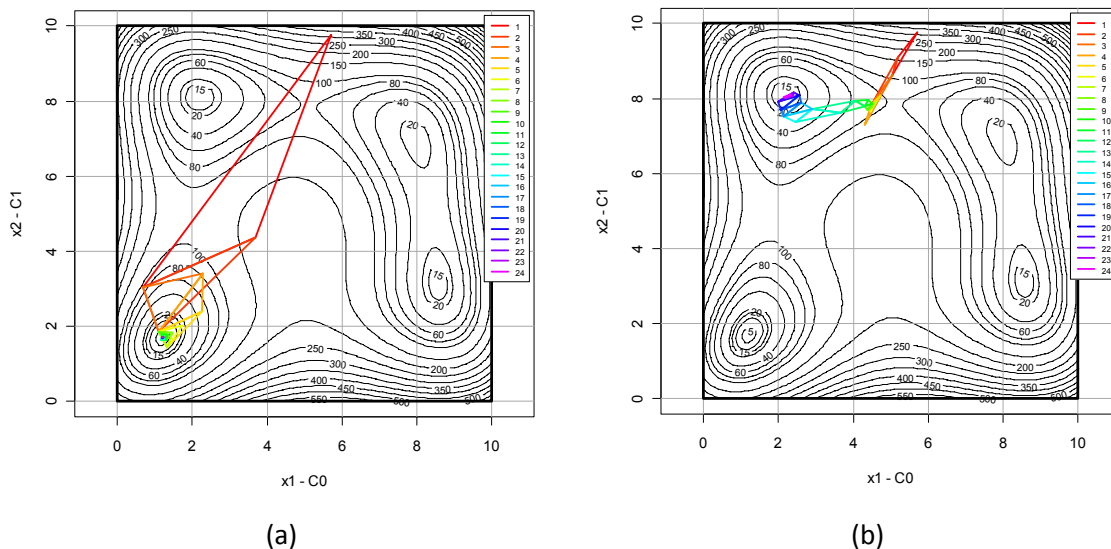


Abbildung 5.25: Zwei verschiedene CNMr-Läufe mit unterschiedlichem Skalierungsfaktor für den Aufbau des Startpolyeders: (a) $s = 1$ (b) $s = 0,1$.

Ganz anders verläuft dagegen ein Optimierungslauf, der vom selben Startpunkt ausgeht, bei dem der Skalierungsfaktor um 90% auf $s = 0,1$ reduziert wurde. Die räumliche Ausdehnung des Startpolyeders in Abbildung 5.25(b) ist nun relativ klein bezogen auf die räumliche Variabilität der Zielfunktion. Durch sukzessive Reflexions- und Expansionsschritte findet die für CNM typische Bewegung des Polyeders durch den Lösungsraum statt, die in drei Etappen jeweils in guter Nähe-

zung dem lokalen Gradienten der Zielfunktion folgt. Das Verfahren konvergiert jetzt – so wie man es von einem lokalen Verfahren erwartet – gegen das lokale Optimum M2.

Abbildung 5.26 zeigt in einer direkten Gegenüberstellung, wie sich die Verkleinerung des Skalierungsfaktors auf die zehn Optimierungsläufe aus Abbildung 5.24 auswirkt: die Reichweite der Optimierungen sinkt, und die CNMr-Läufe konvergieren nun immer gegen das nächstgelegene Optimum.

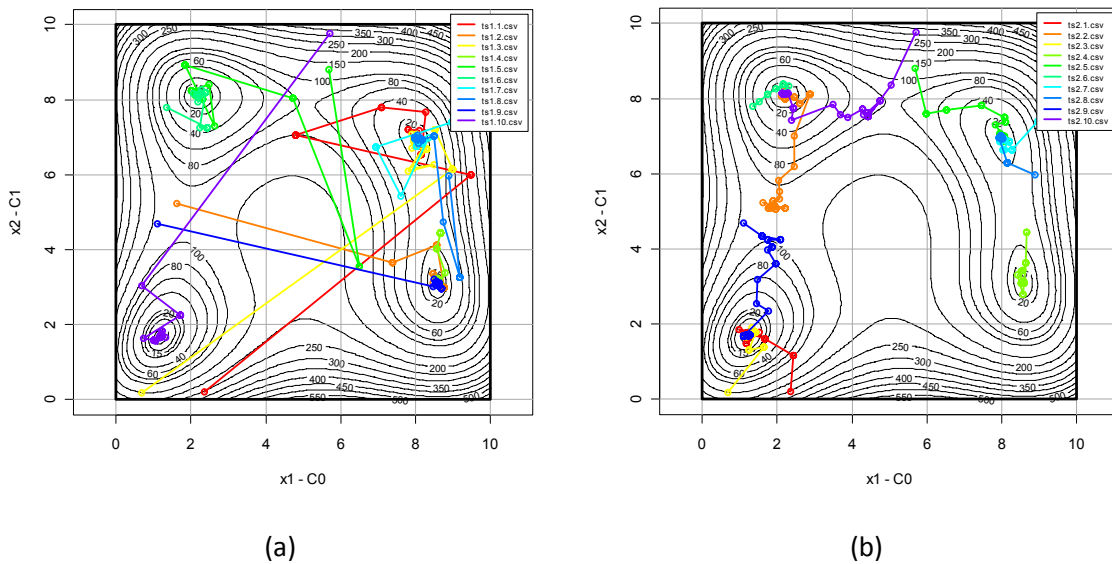


Abbildung 5.26: Wirkung des Skalierungsfaktors auf das Konvergenzverhalten von CNMr. (a) $s=1$ (b) $s=0,1$.

Die vorangegangene Analyse führt zu folgenden Schlussfolgerungen: Mit dem Skalierungsfaktor kann der Einfluss der Startlösung auf das Optimierungsergebnis geregelt werden: je kleiner s , desto stärker wird das Optimierungsergebnis vom vorgegebenen Startvektor beeinflusst. Man könnte auch sagen, desto „lokaler“ das Verfahren. Ist eine gute Startlösung bekannt, ist es daher sinnvoll, einen kleinen Skalierungsfaktor s zu wählen. Große s -Werte führen andererseits zu einer größeren Reichweite des Verfahrens. So kann es auch weiter entfernt liegende Optima finden. Welches Optimum genau gefunden wird, hängt jedoch weniger von dessen Güte als vom Zufall beim Aufbau des Startpolyeders ab.

Der Einfluss des Skalierungsfaktors hängt auch wesentlich von der *Topologie der Zielfunktion* ab. Im linearen Fall ist der Gradient im gesamten Lösungsbereich konstant. Dann ist die Wahl des Skalierungsfaktors von untergeordneter Bedeutung. Wird der Startpolyeder lokal ($s \ll 1$) aufgebaut, werden ggf. einige Iterationen benötigt, um sich in Richtung des Optimums zu bewegen. Wird er global ($s=1$) aufgebaut, müssen dagegen einige Kontraktionen stattfinden. Es ist a priori nicht klar, welches Verfahren schneller zum Ergebnis führt. Schwieriger wird es im Falle nichtlinearer und insbesondere multimodaler Zielfunktionen, oder, wenn die Eigenschaften der Zielfunktion wie bei einem SSNOP weitgehend unbekannt sind. Nachfolgend werden für einen solchen

Fall zwei Optimierungsstrategien vorgeschlagen, die sich darin unterscheiden, wie die Güte der Startlösung eingeschätzt wird:

- (i) Die Startlösung wird als gut eingeschätzt: In diesem Fall geht man von einem kleinen Skalierungsfaktor $0,01 \leq s \leq 0,1$ aus und führt wenige Optimierungen durch (etwa zwischen drei und zehn). Führen alle zu vergleichbaren Ergebnissen, wurde mit hoher Wahrscheinlichkeit das globale Optimum bestimmt.
- (ii) Die Güte der Startlösung ist unklar: Mehrfachstart des Verfahrens. Für $s \rightarrow 1$, muss die Startlösung selbst nicht unbedingt zufällig variiert werden. Wird dagegen $s \ll 1$ gewählt, sollte die Startlösung zufällig variiert werden, damit der gesamte Lösungsbereich hinreichend gut abgedeckt wird.

5.5.3 ES – Populationsgröße

Im Falle der evolutionären Strategie (ES) gibt es zwei Parameter, mit denen der Optimierungsverlauf maßgeblich beeinflusst werden kann.

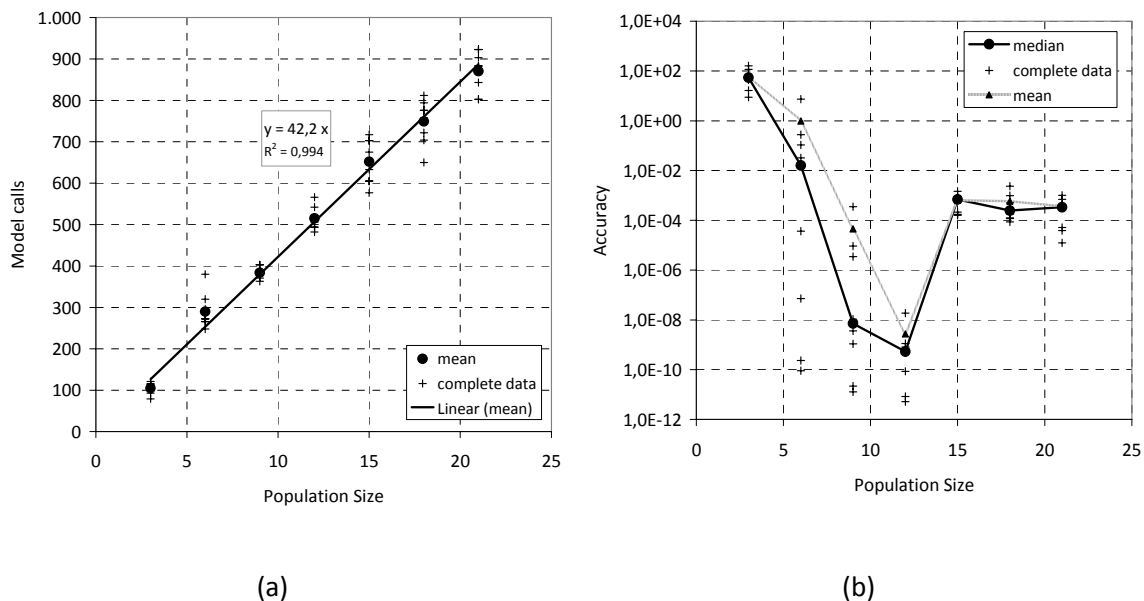


Abbildung 5.27: Einfluss der Populationsgröße bei ES auf (a) die Optimierungsdauer und (b) die Approximationsgenauigkeit für das 2-dimensionale Sphärenmodell (§ 5.3.1).

Abbildung 5.27 zeigt den Einfluss der *Populationsgröße* auf

- (a) die Optimierungsdauer: diese steigt in guter Näherung linear mit der Populationsgröße. Dies bedeutet, dass die Anzahl benötigter Iterationen im untersuchten Bereich unabhängig von der Populationsgröße ist.

- (b) die Approximationsgenauigkeit. Ausgehend von einer minimalen Population von 3 Individuen (entsprechend einem Simplex) kann die Approximationsgenauigkeit zunächst drastisch verbessert werden. Bei 12 Individuen wird ein Optimum erreicht, bevor sich die Approximationsgenauigkeit offensichtlich unabhängig von der Populationsgröße bei ungefähr 10^{-3} einpendelt.

Das Ergebnis zum Einfluss der Populationsgröße auf die Approximationsgenauigkeit ist überraschend. Es stellt sich die Frage, warum sich die Approximationsgenauigkeit zunächst drastisch bis auf etwa 10^{-9} verbessert, dann jedoch auf einem deutlich schlechteren Niveau bei etwa 10^{-3} einpendelt. Offensichtlich handelt es sich bei diesem Ergebnis um ein mit dem implementierten Abbruchkriterium verbundenes Artefakt (§ 5.5.1). Größere Populationen verlangsamen die Konvergenz der ES. Das auf die Bestlösungen aufeinanderfolgender Iterationen bezogene Abbruchkriterium der KOMSA-Implementierung führt daher bei größeren Populationen auch häufiger zu einem vorzeitigen Abbruch. Es ist dieser Effekt, welcher die Approximationsgenauigkeit ab einer Populationsgröße von $n = 15$ wieder stark ansteigen lässt.

Daher muss das Abbruchkriterium zukünftig so geändert werden, dass es sich auf die Varianz der Zielfunktionswerte aller Individuen bezieht. Dann ist zu erwarten, dass die Approximationsgenauigkeit sich monoton mit der Populationsgröße verbessert. Eine solche Änderung hätte jedoch auch zur Folge, dass die Optimierungsdauer überproportional mit der Populationsgröße wächst. In jedem Fall muss der Anwender der ES bei der Wahl der Populationsgröße zwischen Optimierungsdauer und Approximationsgenauigkeit abwägen. Genauere Optimierungsergebnisse gehen stets auf Kosten der Schnelligkeit der ES.

5.5.4 PSO – Einfluss der Trägheit

Abbildung 5.28(a) veranschaulicht die Wirkung der *Trägheit* auf das Konvergenzverhalten der einzelnen Teilchen für das Sphärenmodell. Der monoton gegen das Optimum $f^* = 0$ fallende Trend wird von Oszillationen überlagert, die dadurch entstehen, dass die Teilchen trägheitsbedingt immer wieder über das Ziel hinausschießen.

Abbildung 5.29(a) macht deutlich, welche Gefahr ein zu großer Trägheitsterm birgt. Es handelt sich um eine Optimierung, die bis auf die Wahl des Trägheitsfaktors IW mit derjenigen aus Abbildung 5.28 identisch ist. Da IW von 0,5 auf 0,8 erhöht wurde, bleibt die Konvergenz der Teilchen nun jedoch aus. In elliptischen Bahnen, deren Ausdehnung mit der Größe des zulässigen Bereichs vergleichbar ist, bewegen sie sich um das Optimum wie Himmelskörper um das Zentrum eines Gravitationsfeldes.

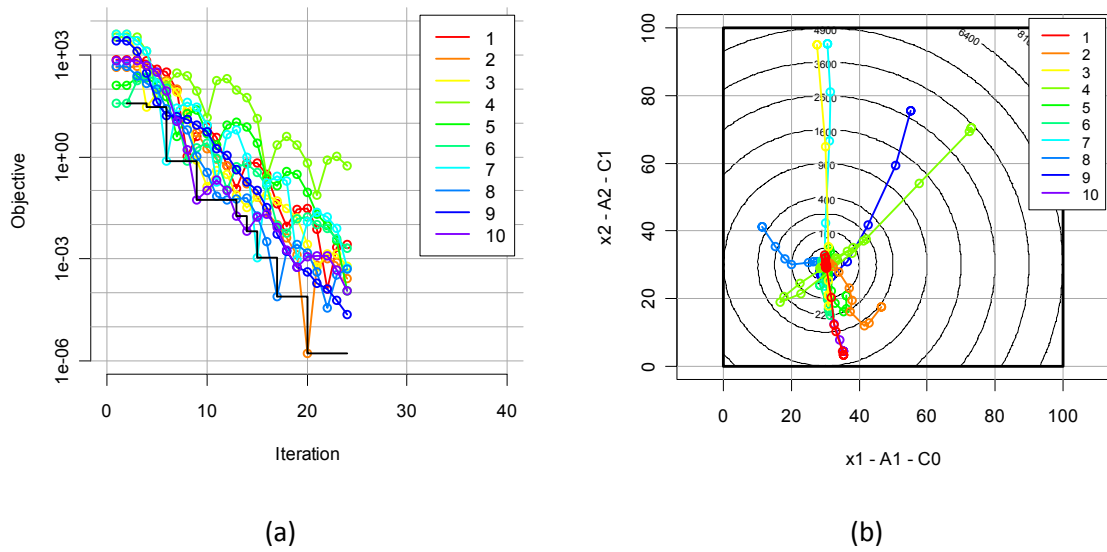


Abbildung 5.28: PSO (a) Zielfunktion: alle Zielfunktionswerte eines Teilchens (über mehrere Iterationen hinweg) sind in derselben Farbe dargestellt und durch eine Linie miteinander verbunden. Die schwarze Kurve stellt den besten Zielfunktionswert je Iteration dar. (b) Lösungsraum: Flugbahnen der verschiedenen Individuen durch Linien veranschaulicht.

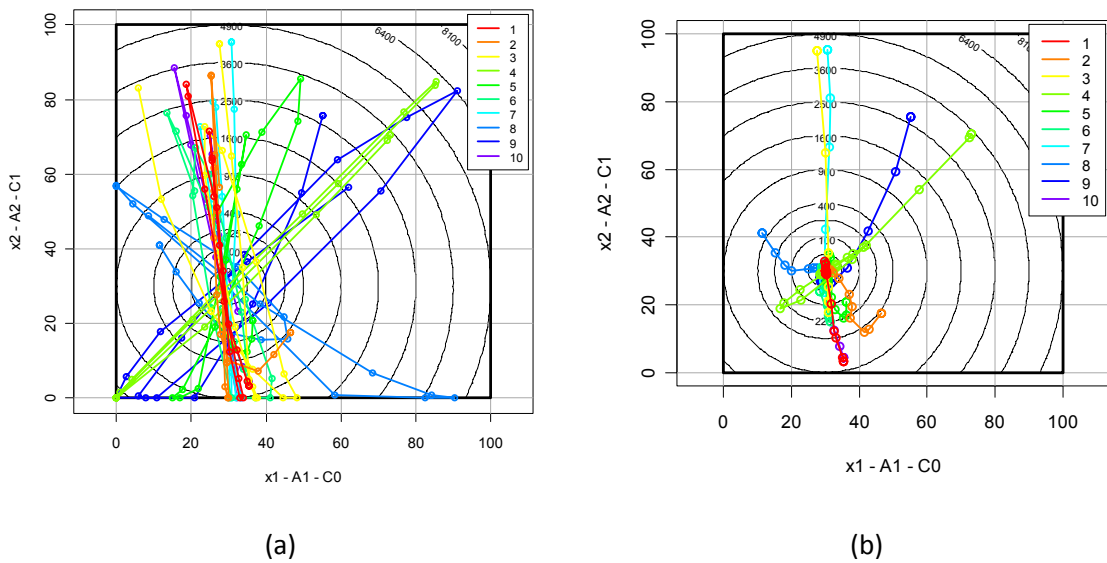


Abbildung 5.29: Verhältnis des Trägheitsterms (iW) zu den Faktoren für individuelles (ILF) und soziales Lernen (SLF). Wertebelegung: (a) $iW=0,8$, $ILF=SLF=0,5$ (b) $iW=ILF=SLF=0,5$.

5.6 Schlussfolgerungen

Mit den in diesem Kapitel durchgeführten numerischen Optimierungen wurden der Complex/Nelder-Mead-Algorithmus (CNM), die Evolutionsstrategie (ES) und die Particle Swarm Optimization (PSO) hinsichtlich ihrer Leistungsfähigkeit untereinander und mit Ergebnissen aus der Literatur zur experimentellen Evaluierung von Optimierungsheuristiken verglichen.

Tabelle 5.8: Quantitative und qualitative Kriterien für die vergleichende Beurteilung der in diesem Kapitel untersuchten direkten Optimierungsverfahren

Quantitativ	<ul style="list-style-type: none"> • Optimierungsdauer • Treffhäufigkeit
Qualitativ	<ul style="list-style-type: none"> • Lokale versus globale Verfahren • Zufallsabhängigkeit • Abhängigkeit von der Start-Suchmenge und Beweglichkeit im Lösungsraum • Umgang mit Restriktionen • Abbruchkriterium • Robustheit bezüglich der Algorithmusparameter

In der vorliegenden Analyse schneidet CNM durchweg besser ab als die anderen beiden Verfahren. Diese Beurteilung beruht einerseits auf der *quantitativen Auswertung* von Testreihen und andererseits auf der Analyse *qualitativer Eigenschaften* der jeweiligen Algorithmen an entsprechenden Fallbeispielen (Tabelle 5.8). Die Ergebnisse werden nachfolgend zusammengefasst.

5.6.1 Quantitative Performancekriterien

Optimierungsdauer

CNM benötigt bei allen untersuchten Testproblemen weniger Modellaufrufe als ES und PSO. Offensichtlich erfolgt die Konstruktion neuer und erfolgreicher Suchlösungen bei CNM effizienter als bei den stärker zufallsbasierten Verfahren ES und PSO: zum Einen benötigt CNM eine kleinere Anfangspopulation⁶⁷ (Startpolyeder), zum Anderen sind je Iteration typischerweise deutlich weniger Modellaufrufe notwendig. Zeitkritisch kann bei CNM die Neuorientierung des Polyeders im Lösungsraum sein, die in der Regel mit mehreren sukzessiven Kontraktionen des Polyeders verbunden ist. Dies gilt insbesondere in höheren Dimensionen, d.h. bei Polyedern mit vielen Ecken.

⁶⁷ Die Startpopulationen von ES und PSO sind für alle Testreihen mehr als doppelt so groß wie der Startpolyeder von CNM.

Im schlechtesten Fall können diese Neuorientierungen auch zu einem vorzeitigen Abbruch führen, was jedoch weniger das Kriterium der Optimierungsdauer als das der Treffhäufigkeit berührt.

Treffhäufigkeit

Auch bei diesem Performancekriterium schneidet CNM für alle Testprobleme aus Abschnitt 5.3 am besten ab. Insbesondere bei der Himmelblau-Testfunktion mit mehreren Optima im Inneren des zulässigen Bereichs ist dies durchaus überraschend. Hier wäre zu erwarten gewesen, dass ein prinzipiell global angelegtes Optimierungsverfahren wie die ES bei der globalen Treffhäufigkeit besser abschneidet. Dass in den vorliegenden Testreihen andere Performanceergebnisse erzielt wurden, mag damit zusammenhängen, dass die Suchpopulation der ES mit 10 Individuen vergleichsweise klein gewählt war. In § 5.5.3 wurde gezeigt, dass die mittlere Approximationsgenauigkeit – und damit letztlich auch die Treffhäufigkeit – sensibel von der Wahl der Populationsgröße abhängt und unter Umständen deutlich gesteigert werden kann, wenn die Suchpopulation vergrößert wird. Andererseits kann eine höhere Treffhäufigkeit von ES nur zum Preis einer höheren Optimierungsdauer erzielt werden, die bereits für 10 Individuen relativ hoch ist.

5.6.2 Qualitative Performancekriterien

Lokale versus globale Suche

Syrjakow (2003: S.18-21) ordnet das Complex- und das Nelder-Mead-Verfahren – und damit indirekt auch CNM als Synthese aus beiden – den lokalen Optimierungsverfahren zu, während die fundamental zufallsbasierten Verfahren wie Evolutionsstrategien und die reine Monte-Carlo-Suche als globale Suchstrategien charakterisiert werden. Dass eine solche Zuordnung nicht immer vollkommen trennscharf ist, zeigt die Diskussion des Skalierungsfaktors für den CNM-Algorithmus (§ 5.5.2): ob sich das Verfahren „eher“ lokal oder global verhält, hängt stark von der Größe des Startpolyeders ab. Diese kann über den Skalierungsfaktor gesteuert werden. Für $s \ll 1$ konvergiert CNM ausgehend von der Startlösung zum nächstgelegenen Optimum, für $s \rightarrow 1$ zu einem beliebigen, nicht zwangsläufig jedoch dem globalen, Optimum. Dieses für ein lokales Verfahren im Prinzip atypische Verhalten resultiert aus der Tatsache, dass der Startpolyeder M_0 entsprechend der von Box (1965) für den Complex-Algorithmus vorgeschlagenen Methode mit einer „globalen“ Monte-Carlo-Suche im gesamten zulässigen Bereich aufgebaut wird. Mit dem Skalierungsfaktor kann diese Suche im Falle der KOMSA-Implementierung jedoch lokal beschränkt werden, weshalb das Verfahren für kleine Skalierungsfaktoren dann auch tatsächlich ein lokales Verhalten zeigt. Ein Skalierungsfaktor für ES und PSO wurde nicht implementiert und ist, wie der nachfolgende Abschnitt zeigt, bei diesen Verfahren grundsätzlich auch nicht sinnvoll. Die Suchstrategien sowohl von ES als auch PSO erzwingen – zumindest in einem frühen Stadium – keine eindeutige Konvergenz aller Testlösungen auf dasselbe Optimum. Darin liegt das spezifisch „globale“ dieser Verfahren. Gelingt – insbesondere bei entsprechenden Restriktionen – nur der lokale Aufbau der Startpopulation, funktionieren jedoch auch diese Verfahren eher lokal bzw. haben überhaupt Schwierigkeiten gegen ein weit entferntes Optimum zu konvergieren.

Zufallsabhängigkeit

Bei CNMr, ES und PSO wird die Anfangssuchmenge M_0 durch eine Monte-Carlo-Suche, d.h. zufällig, aufgebaut. Insofern sind all diese Algorithmen zufallsbasiert. Im folgenden Optimierungsverlauf unterscheiden sich die Verfahren jedoch in der Zufallsabhängigkeit: während CNM ausgehend von M_0 deterministisch weiter voranschreitet, greifen ES und PSO bei jeder Iteration auf weitere stochastische Größen zurück. Der einzige vollständig deterministische Algorithmus ist CNMs. Hier erfolgt der Aufbau des Startpolyeders entsprechend den von Nelder und Mead vorgeschlagenen Regeln (Schwefel, 1995: S.58).

Beweglichkeit - Abhängigkeit von der anfänglichen Suchmenge

Reflexion und Expansion führen bei CNM dazu, dass sich die Polyeder späterer Iterationen deutlich von dem Startpolyeder entfernen können. Dabei wird der lokale Gradient in vielen Fällen recht gut angenähert. Die Lösungen von ES und PSO sind stärker von der Lage und Streuung der $\mathbf{x}_i^{(0)}$ abhängig. Bei ES sind es Rekombination und Mutation, bei PSO die Trägheit der einzelnen Individuen, die es den Algorithmen grundsätzlich erlauben, über den von M_0 abgedeckten Bereich der zulässigen Menge hinaus zu kommen. Dennoch werden Optima generell besser gefunden, wenn sie sich innerhalb oder in der Nähe des von M_0 abgedeckten Bereichs befinden. So kann auch das deutlich bessere Abschneiden von CNMr beim Korridormodell erklärt werden: die Restriktionen führen dazu, dass auch die Startsuchmengen von PSO und ES effektiv lokal aufgebaut werden müssen (alle anderen Werte sind unzulässig). Da das Optimum, verglichen mit der Streuung der Startpopulation, weit entfernt und in einer durch die Orientierung des Korridors ausgezeichneten Richtung liegt, gelingt es den Verfahren nur in wenigen Fällen, es zu finden.

Diese mangelnde Beweglichkeit von ES und PSO führt somit dazu, dass die Optimierungsperformance (Schnelligkeit wie Treffhäufigkeit) stärker als bei CNM von der Startpopulation, d.h. den Ergebnissen der anfänglichen Monte-Carlo-Suche, abhängig ist.

Umgang mit Restriktionen

Alle Verfahren gehen weitgehend gleich mit Restriktionen um: unzulässige Versuchslösungen werden schrittweise in den zulässigen Bereich „zurückgezogen“. Dennoch kann CNM effektiver mit Beschränkungen umgehen als PSO und ES. Der Grund ist die zuvor diskutierte hohe Beweglichkeit. Die Expansion ermöglicht es der Suchmenge (Polyeder), sich entlang der Restriktion neu zu orientieren. Eine ähnlich effektive Neuorientierung und insbesondere Expansion der Suchmenge ist weder bei ES noch bei PSO möglich. Dadurch kann es bei diesen Verfahren verstärkt zu vorzeitiger Konvergenz an einer Restriktion kommen.

Abbruchkriterium

Einen deutlichen Einfluss auf das Optimierungsergebnis hat das Abbruchkriterium (§ 5.5.1). Hier muss grundsätzlich zwischen Treffhäufigkeit und Optimierungsdauer abgewogen werden. *Vorzeitige Konvergenz* kann dadurch vermieden werden, dass die Variabilitätsschwelle gesenkt wird.

Andererseits kann, wenn die Variabilitätsschwelle sich auf den Verlauf der $f_B^{(k)}$ -Kurve bezieht, die Anzahl der Iterationen erhöht werden, die in die Berechnung der Standardabweichung eingeht. Beide Maßnahmen gehen jedoch zu Lasten der Optimierungsdauer, die dadurch steigt.

Robustheit bezüglich der Wahl der Algorithmusparameter

Auch in dieser Eigenschaft erweist sich CNM als den beiden anderen Verfahren überlegen. Die Performance des Algorithmus hängt nicht kritisch von der Wahl irgendeines der Parameter ab. Die Voreinstellungen sind für eine breite Spanne möglicher Optimierungsprobleme geeignet. Diese Robustheit ist für den praktischen Einsatz vorteilhaft: erfahrene Anwender können über die Einstellung der Parameter einen gewissen durchaus nützlichen Einfluss auf den Optimierungsverlauf nehmen. Unerfahrene Anwender können sie jedoch weitgehend ignorieren und müssen nicht befürchten, dass sie dadurch prinzipiell schlechtere Ergebnisse erzielen.

ES ist ebenfalls ziemlich robust bezüglich der Wahl der Algorithmusparameter. Lediglich die Populationsgröße hat einen deutlichen Effekt auf die Leistungsfähigkeit des Algorithmus. Außerdem kann über die Mutation in einem gewissen Maß die Beweglichkeit der Suchmenge gesteuert werden. Je höher der „Mutation Scope“, desto besser werden Bereiche in die Suche einbezogen, welche die Startpopulation nicht abdeckt. Andererseits führt ein hoher Wert der „Mutation Scope“ auch zu einer Verlangsamung der Konvergenz, d.h. zu längeren Optimierungsdauern.

PSO ist hinsichtlich der Wahl der Algorithmusparameter am problematischsten. Das Verhältnis der sozialen und individuellen Lernfaktoren (ILF, SLF) zum Trägheitsterm (iW) entscheidet maßgeblich darüber, wie schnell das Verfahren konvergiert. Kritisch ist auch die absolute Wahl dieser Faktoren: da die Suchpopulation zu Beginn über den gesamten zulässigen Bereich verstreut ist, können die standardmäßig vorgegebenen Werte $(iW, ILF, SLF) = (0,8 | 2 | 1)$ zu sehr großen Schrittwerten und damit zu einem chaotischen Verhalten der Teilchen und ausbleibender Konvergenz der Suchmenge führen. Zu kleine Werte können jedoch zu einer sehr starken Verlangsamung der Konvergenz führen. Die geeignete Wahl dieser Parameter scheint überdies vom Optimierungsproblem selbst abzuhängen und erfordert einiges Fingerspitzengefühl. Neben den ernüchternden Performanceergebnissen ist dies ein wesentlicher Grund, weshalb der Algorithmus – zumindest in seiner jetzigen Implementierung – für die Optimierung von SBOs weniger geeignet scheint.

5.6.3 Anwendung der untersuchten Optimierungsalgorithmen auf das SSNOP

Zusammenfassend spricht einiges dafür, bei der Lösung eines SSNOPs vor allem auf den CNM-Algorithmus zurückzugreifen. Ob die vollständig deterministische Variante CNMs oder diejenige mit einem zufallsbasierten Aufbau des Startpolyeders (CNMr) verwendet wird, ist von untergeordneter Bedeutung. Da bei CNMs die Ecken des Startpolyeders durch komponentenweise Variation des Startvektors aufgebaut werden, verlaufen seine Kanten zu einem großen Teil parallel zu den Koordinatenachsen des Lösungsraums. Ob diese Vorzugsrichtung einen Vor- oder Nachteil für die Optimierung mit sich bringt, lässt sich allgemein jedoch nicht entscheiden, da dies von der Struktur des jeweiligen Optimierungsproblems abhängt.

Da Optimierungsheuristiken grundsätzlich mit dem Nachteil behaftet sind, dass sie die Bestimmung des globalen Optimums nicht garantieren können, ist es in jedem Fall sinnvoll, bei der Optimierung auf mehrere Algorithmen zurückzugreifen. Je nachdem, wie viel Zeit für die Lösung eines SSNOPs investiert werden kann, kann folgendermaßen vorgegangen werden: zunächst wird eine einzelne Optimierung mit CNM ausgehend von einem Startpunkt durchgeführt, welcher idealerweise bereits eine gute Schätzung des Optimums darstellt. Um die Zuverlässigkeit des Ergebnisses zu überprüfen, können Testreihen (Mehrfachstarts) mit unterschiedlichen Startvektoren und/oder Startpopulationen vorgenommen werden. Gemäß der Bernoulli'schen Wahrscheinlichkeitsverteilung kann so die Gesamtwahrscheinlichkeit zur Ermittlung des globalen Optimums – gegenüber einem einzelnen Optimierungslauf – deutlich vergrößert werden. Schließlich können zur Absicherung des gefundenen Ergebnisses ggfs. zusätzliche Einzeloptimierungen bzw. Testreihen mit ES und PSO durchgeführt werden.

6 Schlussbetrachtung und Ausblick

Die Stoffstromanalyse der Stoffstromnetze eignet sich insbesondere wegen ihrer anschaulichen Art der Systemabbildung für die Analyse und Planung des Material- und Energieeinsatzes industrieller Produktionssysteme. Ihr Manko besteht hauptsächlich darin, dass bislang keine automatische Modelloptimierung möglich ist, obwohl beim betrieblichen Energie- und Stoffstrommanagement üblicherweise Optimierungsfragestellungen, wie die nach minimalen Schadstoffemissionen, Produktionskosten, oder ganz allgemein des minimalen Ressourcenverbrauchs, auftreten. Gängige Optimierungsmethoden des Operations Research wie die mathematische Programmierung wären hierfür geeignet, lassen sich aufgrund ihres hohen Abstraktionsniveaus jedoch nur schwer im betrieblichen Umfeld, insbesondere bei kleineren und mittelständischen Unternehmen vermitteln.

Vor diesem Hintergrund wurde in dieser Arbeit eine neuartige Methode für die Optimalplanung industrieller Produktionssysteme entwickelt, welche die Vorteile der Stoffstromnetze mit denen der Parameteroptimierung verbindet. In diesem Kapitel wird die stoffstromnetzbasierende Optimalplanung nun in den Kontext der Forschungsfelder gestellt, aus denen sie durch die Synthese von Stoffstromanalyse und Optimalplanung hervorgegangen ist. Dabei wird insbesondere den Fragen nachgegangen, welchen Beitrag die stoffstromnetzbasierende Optimalplanung für die aktuelle Forschung in diesen Bereichen leisten kann, und welcher zukünftige Forschungsbedarf noch besteht. Aus der Perspektive des Operations Research stellt die stoffstromnetzbasierende Optimalplanung eine *grafik- und simulationsbasierte* Methode für die Optimalplanung industrieller Produktionssysteme dar (Abschnitt 6.1). Aus Sicht der Industrial Ecology steht dagegen der Entscheidungsbezug von Stoffstromanalysen im Vordergrund (Abschnitt 6.2).

6.1 Grafik- und simulationsbasierte Optimalplanung industrieller Produktion

Die stoffstromnetzbasierende Optimalplanung wurde in dieser Arbeit als Alternative zum im OR dominierenden Optimalplanungsparadigma der mathematischen Programmierung entwickelt. Im Hinblick auf die Intuitivität der Modellbildung und die Vermittelbarkeit der Modelle im betrieblichen Kontext ist sie den deklarativen algebraischen Modelliersprachen der mathematischen Programmierung insbesondere wegen ihrer graphenbasierten Abbildung von Produktionssystemen, sowie dem Rückgriff auf Simulationsmethoden bei der Prozessspezifikation grundsätzlich überlegen. Vorteilhaft für die Systembewertung sind außerdem die vorhandenen Schnittstellen zu kostenrechnerischen und ökobilanziellen Methoden.

Die mathematische Programmierung hat ihren Vorteil hingegen bei der Bestimmung optimaler Lösungen. Gängige Solver können komplexe Modelle, selbst solche mit ungünstigen analytischen

Eigenschaften wie Nichtlinearitäten, Unstetigkeiten oder diskreten Modellvariablen, mit einer prinzipiell unbeschränkten Anzahl von Entscheidungsvariablen und Restriktionen zuverlässig lösen. Die simulationsbasierte Optimierung eines stoffstromnetzbasiereten Optimierungsproblems (SSNOP) erfordert dagegen den Einsatz direkter Suchstrategien, die in der Regel langsamer und weniger robust als die analytische Verfahren sind (vgl. Abschnitt 4.1). Unter dem Gesichtspunkt der Modelloptimierung sollte daher immer der mathematischen Programmierung der Vorzug gegeben werden. Eine Abwägung zwischen beiden Optimalplanungsmethoden muss jedoch zugleich die Aspekte der *Modellbildung* (Kap. 3) und der *Modelloptimierung* (Kap. 4 und 5) berücksichtigen.

Die Simulation elementarer Prozesse durch algorithmisch spezifizierte Transitionen

In diesem Zusammenhang stellen insbesondere die in Kapitel 4 hergeleiteten *Transformationsregeln für die algebraische Reformulierung* von SSNOPs ein wichtiges Ergebnis dar. Zum Einen liefern sie ein eindeutiges Kriterium für die Entscheidung zwischen mathematischer und simulationsbasierter Optimierung eines SSNOPs: eine mathematische Optimierung ist genau dann möglich, wenn das SSNOP nur lineare und durch benutzerdefinierte Funktionen spezifizierte Transitionen enthält. Zum Anderen könnte auf ihrer Grundlage zukünftig eine Art *Compiler* entwickelt werden, welcher das am Problemverständnis des Menschen orientierte SSNOP in ein für analytische Optimierungsalgorithmen geeignetes mathematisches Programm übersetzt, unter der genannten Voraussetzung allerdings, dass keine algorithmischen oder durch externe Softwarekomponenten spezifizierten Transitionen verwendet werden.

Nicht immer ist es jedoch möglich, bei der Modellbildung völlig auf derartige Spezifikationsformen zu verzichten. Insbesondere bei der Modellierung von Produktionssystemen der Prozessindustrie spielen sie eine wichtige Rolle. Die Funktionsweise einzelner Verfahrensstufen, wie Verbrennungsprozesse oder chemische Reaktoren, kann vielfach nur unter Rückgriff auf Berechnungsalgorithmen *simuliert* werden. In anderen Fällen ist eine *deklarative* Spezifikation mit mathematischen Funktionen und Gleichungen zwar theoretisch möglich, die algorithmische Problemformulierung für den Anwender jedoch *anschaulicher*. Weil mit der Entwicklung der stoffstromnetzbasiereten Optimalplanung ausdrücklich das Ziel verfolgt wird, ein intuitives Werkzeug für einen breiten Anwenderkreis zu entwickeln, fällt das Kriterium der Anschaulichkeit durchaus ins Gewicht, selbst wenn im Hinblick auf die Lösbarkeit eines Problems die deklarative Spezifikation vorzuziehen wäre (vgl. hierzu auch Niemeyer, 1972: S.15, Fu, 2002: S.211).

Berührungspunkte zum „Visual Interactive Modeling“

Die stoffstromnetzbasierete Optimalplanung knüpft an die Bemühungen innerhalb des OR an, intuitivere Werkzeuge als die in der mathematischen Programmierung verwendeten algebraischen Modelliersprachen für die Optimalplanung von Produktionssystemen zu entwickeln. Die meisten Entwicklungen stammen hier aus den Bereichen des „Visual Interactive Modeling“ (VIM) und der simulationsbasierten Optimierung. Tatsächlich kann die simulationsbasierte Optimierung als ein Ansatz des VIM angesehen werden, dessen Ursprünge auf Hurriions (1980) Entwicklung

eines „visual interactive simulation models“ für die Reihenfolgen- und Ablaufplanung in der Produktion zurückgehen (Bell et al., 1999: S.155). Vor allem aufgrund der sich rasant verbessernden PC-Rechenleistung konnte sie seit dem Ende der 1990er Jahre einen anderen Zweig verdrängen, welcher die Forschung im VIM zuvor dominiert hatte: die Verwendung visueller Modellertechniken, um den Zugang zur mathematischen Programmierung für Anwender in der Praxis zu erleichtern (Jones, 1994) und damit Planungsprozesse optimal zu unterstützen: „A Visual Interactive Model combines meaningful pictures and easy interactions to stimulate creativity and insight; promoting a process of ‘generate and test’, it facilitates a rapid cycle of learning“ (Belton und Elder, 1994).

Das in Abschnitt 4.3 konzipierte Optimierungsmodul integriert nun *beide* VIM-Ansätze: Stoffstromnetze können als benutzerfreundliches „Frontend“ zur Formulierung mathematischer Programme verwendet, oder in eine simulationsbasierte Optimierung integriert werden. Die Verwendung als Benutzerschnittstelle für die mathematische Programmierung ist vor allem deshalb interessant, weil Stoffstromnetze eine gut etablierte Methode des betrieblichen Stoffstrommanagements mit weit reichenden Methoden der ökologischen und ökonomischen Systembewertung sind. Das haben sie den „ad hoc“-Entwicklungen des VIM voraus, die sich wie z.B. das grafikbasierte Planungswerkzeug für die Prozessindustrie von Draman et al. (2002) nach einer Entwicklungs- und Erprobungsphase häufig nicht langfristig in der Praxis durchsetzen konnten. Durch die Kombination der mathematischen und der simulationsbasierten Optimierung entsteht zudem ein besonders flexibles Planungswerkzeug, das sich nicht an den Anforderungen eines bestimmten Lösungsansatzes, sondern an den Bedürfnissen der Planenden orientiert.

6.2 Entscheidungsorientierte Stoffstromanalysen

Sowohl die theoretischen Überlegungen als auch die exemplarische Anwendung der stoffstromnetzbasierter Optimalplanung machen deutlich, dass die methodische Integration von Stoffstromanalyse und Optimalplanung keine „Einbahnstraße“ ist. Als Ergebnis erhält man nicht nur die angestrebte Erweiterung des Stoffstrommanagements um den Aspekt der automatischen Systemoptimierung. Die Einbettung der Stoffstromanalyse in die Optimalplanung hat auch Auswirkungen auf die Art der Stoffstromanalyse. Sie erzwingt den Perspektivwechsel von der traditionell eher beschreibenden hin zu einer *entscheidungsorientierten* Stoffstromanalyse. Nur durch die systematische Berücksichtigung von Freiheitsgraden in Form von vorhandenen oder potenziellen Handlungs- und Gestaltungsspielräumen entstehen Stoffstrommodelle, die gleichzeitig als Optimierungsmodelle einsetzbar sind.

Die stoffstromnetzbasierter Optimalplanung wurde in dieser Arbeit vorwiegend aus der Perspektive des betrieblichen Stoffstrommanagements heraus entwickelt. Das Stoffstrommanagement im allgemeinen Sinn der Industrial Ecology kann sich jedoch auf beliebige Teilsysteme des industriellen Metabolismus beziehen. So werden beim Life Cycle Assessment (LCA) über Betriebsgrenzen hinweg alle mit der Herstellung, Nutzung und Entsorgung eines Produktes verbundenen Prozesse der Technosphäre untersucht. Auch die Untersuchung von Supply Chains oder Produktionsnetz-

werken erfordert eine überbetriebliche Sichtweise. Hier soll daher der Frage nachgegangen werden, welche Implikationen das aus einer betrieblichen Perspektive entwickelte Paradigma der entscheidungsorientierten Stoffstromanalyse auf die Analyse beliebiger industrieller Metabolismen hat.

Attributive versus handlungskonsequenzbezogene Ökobilanzierung

Stoffstromanalysen sind in Gestalt der Sachbilanzierung (LCI) der quantitative Kern jeder Ökobilanz (LCA). Je nachdem, unter welchen Gesichtspunkten die Sachbilanzierung durchgeführt wird, unterscheidet man zwischen *attributiven* und *handlungskonsequenzbezogenen* Ökobilanzen⁶⁸.

Das Ziel der in der Praxis überwiegend eingesetzten attributiven Sachbilanzierung ist die systematische Analyse aller durch ein Produkt oder einen Produktionsprozess hervorgerufenen umweltrelevanten Stoffströme. Sie basiert auf der Vorstellung, dass jedem Produkt oder jeder Dienstleistung, formalisiert in der so genannten funktionellen Einheit, eine bestimmte Umweltwirkung *eindeutig* zugeordnet werden kann (Frischknecht, 2005: S.12f). Genau dieses leisten die üblicherweise eingesetzten linearen Input/Output-Modelle, in denen die einzelnen Prozesse eines Produktsystems in einer den linear spezifizierten Transitionen sehr ähnlichen Weise abgebildet werden⁶⁹ (Heijungs, 1994, Heijungs und Suh, 2002). Um die Eindeutigkeit der Abbildung zu gewährleisten, ist eine strikt *limitationale* Abbildung des gesamten Produktsystems notwendig. Andernfalls könnten ein und derselben funktionellen Einheit verschiedene Umweltwirkungsvektoren zugeordnet werden. Genau auf diesem Prinzip baut jedoch eine ökologische Optimierung auf: die Senkung von Umweltbelastungen bei gleichbleibendem Nutzen (funktionelle Einheit). Die *beschreibende Stoffstromanalyse* der attributiven Ökobilanz blendet die dafür notwendigen Freiheitsgrade eines Produktsystems gezielt aus und führt daher zu Stoffstrommodellen, die für eine automatische Systemoptimierung per definitionem völlig ungeeignet sind. Um das Prinzip einer ökologischen Optimierung mit dem der Ökobilanzierung zu verbinden, muss daher der „Attributionsgedanke“ fallen gelassen und Substitutionsmöglichkeiten innerhalb des Produktionssystems systematisch in die Analyse integriert werden.

Den dafür notwendigen konzeptionellen Rahmen liefert die entscheidungsorientierte Stoffstromanalyse. Sie weist in dieser Hinsicht deutliche Parallelen zur handlungskonsequenzbezogenen Ökobilanzierung auf, welche die Auswirkungen von Eingriffen in ein gegebenes Produktsystem,

⁶⁸ Im Englischen haben sich für die Unterscheidung in den letzten Jahren die Begriffe „attributional“ und „consequential“ weitgehend durchgesetzt (Finnveden et al., 2009: S.3), obwohl nach wie vor auch viele Synonyme wie „descriptive“ oder „retrospective“ für die attributional LCA und „change-oriented“ oder „market-based“ für die consequential LCA verwendet werden. Die deutsche Übersetzung wurde von Walk (2009: S.119) übernommen.

⁶⁹ Der einzige Unterschied besteht darin, dass linear spezifizierte Transitionen im Gegensatz zu den in linearen I/O-Modellen verwendeten Leontief-Transformationen auch Prozesse mit mehreren Outputs darstellen können, in diesem Sinne also sogar allgemeiner sind (vgl. §2.4.4).

beispielsweise hinsichtlich des Produktdesigns, auf die umweltrelevanten Stoffströme analysiert (Finnveden et al., 2009: S.3, Tillman, 2000). Bei handlungskonsequenzbezogenen Ökobilanzen wird üblicherweise jedoch eine volkswirtschaftliche Perspektive eingenommen und der Fokus auf Freiheitsgrade gerichtet, die sich durch marktwirtschaftliche Interaktionen der Unternehmen einer Supply Chain über Gütermärkte ergeben. Auf der methodischen Ebene führt diese Sichtweise zu einem Übergang von den linearen I/O-Modellen zu allgemeinen Gleichgewichtsmodellen, mit denen das Zusammenspiel von Angebot und Nachfrage beispielsweise durch entsprechende Preiselastizitäten abgebildet werden kann.

Die stoffstromnetzbasierter Optimalplanung nimmt dagegen eine betriebswirtschaftliche Perspektive ein. Im Hinblick auf die Bewertung produktionsbedingter Umweltwirkungen knüpft sie mit der in § 3.3.4 vorgenommenen Unterscheidung zwischen einer Kern- und einer Komplementärbilanz an die *betriebliche Ökobilanzierung* an. Ihr Fokus liegt auf den innerbetrieblichen Freiheitsgraden, die sich nicht aus marktwirtschaftlichen Prinzipien ergeben, sondern aus den Entscheidungen der Unternehmensführung, zu deren Unterstützung sie konzipiert ist. Der Fokus auf die Kernbilanz resultiert zum Einen aus der Tatsache, dass nur dieser Teil der Sachbilanz auch im unmittelbaren Einflussbereich der Unternehmensführung liegt. Zum Andern liegen die für eine operative oder taktische Optimalplanung notwendigen Daten meistens nur für innerbetriebliche Prozesse in hinreichender Qualität und Detaillierung vor. Die der betrieblichen Produktion vor- und nachgelagerte Prozesse können dagegen mit den Methoden der attributiven Sachbilanzierung abgebildet werden, die beispielsweise auch den Prozessdaten vieler Ökobilanzdatenbanken wie ecoinvent zugrunde liegen (Frischknecht et al., 2005).

Betriebliches Stoffstrommanagement

In das betriebliche Stoffstrommanagement, aus dessen Perspektive die stoffstromnetzbasierter Optimalplanung in dieser Arbeit entwickelt wurde, fügt sich die entscheidungsorientierte Stoffstromanalyse problemlos ein. Vorhandene Systemfreiheitsgrade, beispielsweise in Gestalt von Prozessparametern oder durch die Verzweigung von Stoffströmen innerhalb des Betriebs oder Querverbindungen zwischen unterschiedlichen Produktlinien, werden hier – meist weitgehend unreflektiert – in das Stoffstrommodell integriert. Zwar werden auch für betriebliche Stoffstromanalysen linear spezifizierte Transitionen als einfache Näherungen des Verhaltens realer Transformationsprozesse verwendet. Das führt in der Regel jedoch nicht zu einer der attributiven Ökobilanzierung vergleichbaren systematischen Vernachlässigung von Freiheitsgraden. Optimierbare Stoffstrommodelle ergeben sich in vielen Fällen daher auf einem ganz natürlichen Weg, worin auch die ursprüngliche Motivation für die in dieser Arbeit vorgenommene Methodenentwicklung liegt.

Dennoch können die Ergebnisse aus Kapitel 3 auch im Falle betrieblicher Stoffstromanalysen eine wertvolle Leitlinie für die Modellbildung darstellen. Einzelne Produktionsschritte sind demnach immer unter dem Gesichtspunkt abzubilden, ob es neben den tatsächlich vorhandenen auch potenziell denkbare Substitutionsmöglichkeiten zwischen den Inputs und Outputs einzelner Transitionen gibt. Außerdem wird durch die entscheidungsorientierte Stoffstromanalyse die Bedeu-

tung der zwischen den Produktionsschritten gelegenen „Stellen“ gegenüber einer bloß beschreibenden Sichtweise aufgewertet. So verbindet sich mit einer Stelle grundsätzlich die Frage, ob alternative Verwendungs- oder Beschaffungsmöglichkeiten für bestimmte Produkte bzw. Produktionsfaktoren bestehen. Es muss jedoch der fundamental gestaltende Charakter von Planungsprozessen berücksichtigt werden. Insbesondere für taktische Planungen ist die einfache Abbildung eines Produktionssystems meistens nicht ausreichend. Vielmehr müssen alternative Strukturen, wie sie sich durch veränderte Stoffstromführungen oder den Einsatz neuer Techniken ergeben, in das stoffstromnetzbasierende Planungsmodell integriert werden.

Geeignete Transitionsspezifikationstypen für die entscheidungsorientierte Stoffstromanalyse

Die entscheidungsorientierte Stoffstromanalyse ist demnach ein spezielles Paradigma der Systemanalyse, welches neben den für jede Stoffstromanalyse üblichen Stoffumwandlungsprozessen und Stoffströmen explizit auch prozess- und strukturbezogene Systemfreiheitsgrade einbezieht. In Kapitel 3 konnte gezeigt werden, dass das Stellen-Transitions-Konzept der Stoffstromnetze dafür gute Voraussetzungen liefert. Diese müssen bei der Modellierung jedoch auch genutzt werden. Der Fokus zukünftiger Forschung muss daher von den formalen Möglichkeiten der Stoffstromnetze auf die mentalen Modelle ihrer Nutzer verschoben werden. Um den Entscheidungsbezug besser in der Stoffstromanalyse zu verankern, muss insbesondere über Alternativen zur linear spezifizierten Transition nachgedacht werden. Während dieser Transitionstyp für die Belange beschreibender Ökobilanzen oder Stoffstromanalysen bestens geeignet ist, führt die limitationale Abbildung von Transformationsprozessen zur systematischen Ausblendung von Freiheitsgraden. Daher sollten, wo dies im Hinblick auf die realen Prozesse sinnvoll und aufgrund der Datenlage möglich ist, wirklichkeitsnähere Formen der Prozessmodellierung gewählt werden, bei denen Input- und Outputströme zumindest in gewissen Grenzen frei variiert werden können.

Die Transitionsspezifikation durch benutzerdefinierte Funktionen oder Skripte liefert im Prinzip eine für die flexible Prozessmodellierung ausreichende Grundlage. Für eine bessere Unterstützung der entscheidungsorientierten Stoffstromanalyse wäre jedoch über entsprechend angepasste Spezifikationsmodelle nachzudenken, die vom Modellierer z.B. gezielt Substitutionsmöglichkeiten zwischen Produktionsfaktoren oder prozessbedingte Restriktionen abfragen. Im einfachsten Fall wäre es beispielsweise denkbar, die Angabe von Transformationskoeffizienten bei linear spezifizierten Transitionen durch die Definition von *Wertebereichen* zu ersetzen. Speziell im Hinblick auf die Systemoptimierung in der Prozessindustrie wäre außerdem eine engere Anbindung an die verfahrenstechnische Prozessmodellierung sinnvoll. So könnten für die Modellierung von Unit Operations thermodynamisch fundierte Modelle aus dem Bereich der Fließbildsimulation übernommen werden. Dafür müssten neben den bisher in Stoffstromnetzen erfassten quantitativen jedoch auch *qualitative* Informationen zu den Stoffströmen, wie beispielsweise Druck, Temperatur oder Konzentrationen, integriert werden.

Anhang

A.1 Mathematische Notation

Allgemeines

R	Menge der reellen Zahlen. $R^+ = \{x \in R x \geq 0\}$
Z	Menge der ganzen Zahlen
N	Menge der natürlichen Zahlen: $N = \{1, \dots, +\infty\}$
x	Einzelne Variable (Skalar)
\mathbf{v}	Vektor: $\mathbf{v} = (x_1, \dots, x_n)$

Optimierung

L	Zulässige Menge. Für eine kontinuierliche Parameteroptimierung mit m Entscheidungsvariablen gilt $L \subset R^m$.
\mathbf{x}	Entscheidungsvektor (Lösung). Enthält als Komponenten die einzelnen Entscheidungsvariablen: $\mathbf{x} = (x_1, \dots, x_m)$.
$f(\mathbf{x})$	Zielfunktion
$r(\mathbf{x})$	Numerische Rückgabe des Stoffstrommodells, insbesondere Flüsse und Bestände des Stoffstromnetzes
$w(\mathbf{x})$	Textrückgabe des Stoffstrommodells. Insbesondere Warnmeldungen eines Stoffstrommodells, die bei der Berechnung auftreten. Da die Reihenfolge keine Rolle spielt, wird w als Menge und nicht als Vektor dargestellt. $w = \{ \}$ bedeutet, dass keine Warnmeldungen aufgetreten sind.
\mathbf{x}^*	Optimale Lösung
$\mathbf{x}^{(0)}$	Startvektor
$\mathbf{x}_i^{(k)}$	In der k -ten Iteration neu berechnet Testlösungen ($i = 1, \dots, \mu$).
S_k	in der k -ten Iteration neu erzeugte Entscheidungsvektoren
M_k	Suchmenge in der k -ten Iteration. Wird je nach Verfahren auch bezeichnet als Simplex bzw. Polyeder (CNM), Population (ES), Schwarm (PSO).
$f_i^{(k)} = f(\mathbf{x}_i^{(k)})$	Zur Vereinfachung können alle Indizes vom Entscheidungsvektor auf die entsprechende Zielfunktion übertragen werden.

$\mathbf{x}_B^{(k)}$	Beste Lösung der k -ten Iteration k : $\mathbf{x}_B^{(k)} = \min_i \{\mathbf{x}_i^{(k)}\}$
\mathbf{x}_B	Beste Lösung des gesamten Optimierungslaufs: $\mathbf{x}_B = \min_{i,k} \{\mathbf{x}_i^{(k)}\}$
S	Stichprobe
$n = S $	Stichprobenumfang
r_A	Initialisierung des zufallsbasierten Algorithmus ($r_A \in N$)
$\mathbf{p}(A)$	Algorithmisspezifische Parameter. Für CNM beispielsweise: $\mathbf{p}(A) = \{\alpha, \beta, \gamma\}$
r_T	Initialisierung des Stichprobengenerators ($r_T \in N$)

Stoffstromanalyse und Stoffstromnetze

E	Menge der Komponenten eines Stoffstromsystems
σ	Anzahl der Komponenten eines Stoffstromsystems: $\sigma = E $
M	Menge der verschiedenen Stoffarten eines Stoffstromsystems
κ	Anzahl verschiedener Stoffarten: $\kappa = M $
z	Stoffstrom
p	Modellparameter
$p^{(N)}$	Netzparameter
$p^{(T)}$	Transitionsparameter
s	Stoffbestand in einer Stelle
T	Menge der Transitionen
P	Menge der Stellen
F	Menge der Flussrelationen
N	Stoffstromnetz: $N = \{P, T, F\}$

A.2 Abkürzungen

CNM	Complex/Nelder-Mead-Verfahren
ES	Evolutions-Strategie
LCA	Life Cycle Assessment = Ökobilanzierung
LCI	Life Cycle Inventory = Sachbilanzierung
LP	Lineares Programm
MILP	Gemischt ganzzahliges lineares Programm (mixed integer linear programm)
MINLP	Gemischt ganzzahliges nichtlineares Programm (mixed integer nonlinear programm)
MFCA	Material Flow Cost Accounting
MO	Mathematisch-analytische Optimierung
MP	Mathematische Programmierung
NLP	Nichtlineares Programm
OR	Operations Research
PSO	Particle Swarm Optimization
SimOP	Simulationsbasiertes Optimierungsproblem
SBO	Simulationsbasierte Optimierung
SSA	Stoffstromanalyse
SSN	Stoffstromnetz
SSNOP	Stoffstromnetzbasierendes Optimierungsproblem
S/T-Netz	Stellen-Transitions-Netz
SSM	Stoffstrommanagement
POP	Parameteroptimierung(-sproblem)

A.3 LINGO-Code des vollständig transformierten SSNOPs für das Verbundproduktionsbeispiel aus Abschnitt 3.4

Nachfolgend findet sich das in der algebraischen Modellersprache LINGO abgefasste mathematische Programm, das aus der algebraischen Reformulierung des Verbundproduktionsbeispiels (7D-Modell) aus Abschnitt 3.4 resultiert. Weil es sich um eine manuelle Transformation handelt, wurden die im Sinne der mathematischen Programmierung redundanten Spezifikationen der Transitionen $T4-T6$ nicht reformuliert.

Die Transitionen $T4-T6$ werden in dem Stoffstromnetz nur deshalb benötigt, weil die Verteilerstellen $P6-P8$ nicht direkt mit den Outputstellen $P9, P11$ und $P12$ verbunden werden können. Weil in $T4-T6$ keine Stoffumwandlung stattfindet, sind die jeweiligen Input- und Outputströme identisch. Bei einer automatischen algebraischen Reformulierung führt dies, beispielsweise im Falle von $T4$, zur Gleichung $x_{F4} = y_F$. Damit wird dem mathematischen Programm ein zusätzlicher Freiheitsgrad (x_{F4}) hinzugefügt, der durch die der Transitionsspezifikation entsprechenden Gleichungsrestriktion jedoch gleich wieder vernichtet wird. Bei der manuellen Transformation wurde daher so getan, als würden die Stoffflüsse y_F, y_G und y_H jeweils direkt aus den Stellen $P6-P8$ hervorgehen.

```
MODEL:
TITLE Verbund-Produktion 7D;

!*****;
!* Zielfunktion *;
!*****;

! Definition der Hilfsgrößen 'ertrag' und 'aufwand' für Zielfunktionsdefinition;
! Koeffizienten enthalten spezifische Preise aller Stoffe;
ertrag = 42*yp + 30*yf + 27*yg + 36*yh;
aufwand = 7*xa + 4*xb + 3*xc + 5*xd + 0.5*xw1 + 0.5*xw2 + 6*ye;

! Verschiedene Zielfunktionsvarianten
! *** 1) Effizienzmaximierung (NLP);
!Max = ertrag/aufwand;
! *** 2) Aufwands- bzw. Kostenminimierung (LP);
Min = aufwand;
! *** 3) Gewinnmaximierung (LP);
!Max = ertrag - aufwand;

!*****;
!* Restriktionen ` ` *;
!*****;
!*** Rohstoffverfügbarkeit;
[Ober_xA] xa<=50;
[Ober_xB] xb<=400;
[Ober_xC] xc<=300;
[Ober_xD] xd<=100;

! *** Nachfrage;
```

```
![Unter_yf] yf>=130;
![Unter_yg] yg>=150;
![Unter_yh] yh>=5;
![Unter_yp] yp>=60;
[ertrag_con] ertrag >= 10650;

!*** Stelle P2 ***;
[xB_P2] xb=xb1 + xb2;

!*** Transition T1 ***;
[xA_T1] xa=0.2*x1;
[xB_T1] xb1=0.15*x1;
[xC_T1] xc=0.65*x1;
[Output_T1] x1=yf1+yg1+ye;
[yE_T1] ye=0.05*x1;
[xW_T1] xw1 = 44*x1;
[yF_T1] yf1>=0.2*x1;

!*** Transition T2 ***;
[Input_T2] x2=xd+xb2;
[yF_T2] yf2=0.3*x2;
[yG_T2] yg2=0.6*x2;
[yH_T2] yh2=0.1*x2;
[xW_T2] xw2 = 54*x2;
[xB_T2] xb2 >= 0.2*x2;

!*** Stelle P6 ***;
[F_P6] yf1+yf2 = yf+xf3;

!*** Stelle P7 ***;
[G_P7] yg1+yg2 = yg+xg3;

!*** Stelle P8 ***;
[H_P8] yh2=yh+xh3;

!*** Transition T3 ***;
[P_T3] yp=xf3+xg3+xh3;
[P_Oelgehalt_T3] 0.27*yp <= (0.21*xf3+0.22*xg3+0.45*xh3);
```


Literatur

- Alt, W. (2002): *Nichtlineare Optimierung*. Vieweg, Braunschweig, Wiesbaden.
- Ayres, R. U. (1989): Industrial Metabolism. In: Ausubel, J. H. und H. E. Sladovich (Hrsg.): *Technology and Environment*. National Academy Press, Washington, D.C.
- Ayres, R. U. (1994): Industrial Metabolism: Theory and Policy. In: Ayres, R. U. und U. E. Simonis (Hrsg.): *Industrial Metabolism: Restructuring for Sustainable Development*. United Nations University Press, Tokyo.
- Ayres, R. U. und L. W. Ayres (1996): *Industrial ecology: towards closing the materials cycle*. Edward Elgar, Cheltenham.
- Ayres, R. U. und A. V. Kneese (1969): Production, Consumption and Externalities. *American Economic Review*, 59: 282-297.
- Bell, P. C. (1985): Visual Interactive Modeling as an OR Technique. *Interfaces*, 15(4): 26-33.
- Bell, P. C., C. K. Anderson, D. S. Staples und M. D. Elder (1999): Decision-makers' perceptions of the value and impact of visual interactive modelling. *Omega, Int. J. Mgmt Sci.*, 27: 155-165.
- Belton, V. und M. D. Elder (1994): Decision support systems: Learning from visual interactive modeling. *Decision Support Systems*, 12: 355-364.
- Bode, A., E. Uerdingen und T. Wisniewski (2007): Industrial Process Systems Engineering: Identification and Realization of Economic Potentials. *European Congress of Chemical Engineering*, Copenhagen.
- Bossel, H. (2004): *Systeme Dynamik Simulation. Modellbildung, Analyse und Simulation komplexer Systeme*. Books on Demand GmbH, Norderstedt.
- Box, M. J. (1965): A new method of constrained optimization and a comparison with other methods. *Comp. J.*, 8(1): 42-52.
- Braunschweig, A. und R. Müller-Wenk (1993): *Ökobilanzen für Unternehmen: eine Wegleitung für die Praxis*. Verlag Paul Haupt, Bern u.a.
- Bretzke, W.-R. (1980): *Der Problembezug von Entscheidungsmodellen*. Mohr, Tübingen.
- Bringezu, S. und Y. Moriguchi (2002): Material flow analysis. In: Ayres, R. U. und L. W. Ayres (Hrsg.): *A Handbook of Industrial Ecology*. Edward Elgar, Northampton.
- Brockhaus (1994): *Brockhaus-Enzyklopädie: in 24 Bd. - 19., völlig neubearb. Aufl.* F.A. Brockhaus GmbH, Mannheim.
- Brockhaus (1998): *Brockhaus-Enzyklopädie: 21. Aufl.* F.A. Brockhaus GmbH, Mannheim.
- Brooke, Kendrick, Meeraus und Raman (2010): *GAMS - A User's Guide*. GAMS Development Corporation, Washington, D.C.
- Brunner, P. H. und H. Rechberger (2004): *Practical Handbook of Material Flow Analysis*. CRC Press, Boca Raton et al.

- Dasgupta, S., C. Papadimitriou und U. Vazirani (2006): *Algorithms*. McGraw-Hill.
- Draman, M., I. K. Altinel, N. Bajgoric, A. T. Ünal und B. Birgören (2002): A clone-based graphical modeler and mathematical model generator for optimal production planning in process industries. *European Journal of Operational Research*, 137(3): 483-496.
- Driessen, L. T. (2006): *Simulation-based Optimization for Product and Process Design*. Dissertation, Universiteit van Tilburg.
- Dutta, G., N. Gupta und R. Fourer (2011): An optimization-based decision support system for strategic planning in a process industry: the case of aluminium company in India. *Journal of the Operational Research Society*, 62: 616-626.
- Dyckhoff, H. (1991): Berücksichtigung des Umweltschutzes in der betriebswirtschaftlichen Produktionstheorie. In: Ordelheide, D., B. Rudolph und E. Büselmann (Hrsg.): *Betriebswirtschaftslehre und ökonomische Theorie*. Verlag C.E. Poeschel, Stuttgart: 275-309.
- Dyckhoff, H. (1994): *Betriebliche Produktion. Theoretische Grundlagen einer umweltorientierten Produktionswirtschaft*. Springer, Berlin et al.
- Dyckhoff, H. (1998): *Grundzüge der Produktionswirtschaft. Einführung in die Theorie betrieblicher Wertschöpfung*. Springer, Berlin et al.
- Eberhart, R. und J. Kennedy (1995): A New Optimizer Using Particle Swarm Theory. *Sixth International Symposium on Micro Machine and Human Science*: 39-43.
- Enquete-Kommission (1994): Enquête-Kommission "Schutz des Menschen und der Umwelt" des Deutschen Bundestages: *Die Industriegesellschaft gestalten - Perspektiven für einen nachhaltigen Umgang mit Stoff- und Materialströmen*. Economica Verlag, Bonn.
- Finnveden, G., M. Z. Hauschild, T. Ekvall, J. Guinée, R. Heijungs, S. Hellweg, A. Koehler, D. Pennington und S. Suh (2009): Recent developments in Life Cycle Assessment. *Journal of Environmental Management*, 91: 1-21.
- Fischer-Kowalski, M. (2007): Multidisziplinäre Wurzeln der Industrial Ecology. In: Isenmann, R. und M. Von Hauff (Hrsg.): *Industrial Ecology: Mit Ökologie zukunftsorientiert wirtschaften*. Elsevier, München: 89-100.
- Fraginière, E. und J. Gondzio (2002): Optimization modeling languages. In: Pardalos, P. M., G. Mauricio und C. Resende (Hrsg.): *Handbook of Applied Optimization*. Oxford University Press, New York: 993-1007.
- Frischknecht, R. (2005): *Methoden der Umweltbewertung technischer Systeme. Teil 1: Ökobilanzen (Life cycle assessment, LCA)*. ETH Zürich, Studiengang Umweltwissenschaften.
- Frischknecht, R., N. Jungbluth, H.-J. Althaus, G. Doka, R. Dones, T. Heck, S. Hellweg, R. Hirschler, T. Nemecek, G. Rebitzer und M. Spielmann (2005): The ecoinvent Database: Overview and Methodological Framework. *Int. J. LCA*, 10(1): 3-9.
- Fröhling, M. (2006): *Zur taktisch-operativen Planung stoffstrombasierter Produktionssysteme - dargestellt an Beispielen in der stoffumwandelnden Industrie*. Deutscher Universitäts-Verlag, Wiesbaden.
- Frontline Systems (2009): *Linear and Quadratic Programming Problems*. Frontline Systems, <http://www.solver.com/xlsengines.htm#Linear%20and%20Quadratic%20Programming%20Problems>, 4.2.2010.

- Frontline Systems (2010): *Optional Plugin Solver Engines*. Frontline Systems, <http://www.solver.com/xlsengines.htm>, 6.4.2011.
- Frosch, R. A. und N. E. Gallopoulos (1989): Strategies for Manufacturing. *Scientific American*, 261(September Special Issue): 94-102.
- Fu, M. C. (2002): Optimization for Simulation: Theory vs. Practice. *INFORMS Journal on Computing*, 14(3): 192-215.
- Fu, M. C., F. W. Glover und J. April (2005): Simulation Optimization: A review, new developments, and applications. *37th Winter Simulation Conference*, Orlando, USA.
- Geldermann, J. (2005): *Mehrzielentscheidungen in der industriellen Produktion*. Habilitationsschrift, Fakultät für Wirtschaftswissenschaften, Universität Karlsruhe (TH), Karlsruhe.
- Grossman, I. (2005): Enterprise-wide Optimization: A New Frontier in Process Systems Engineering. *AIChE Journal*, 57(7): 1846-1857.
- Hähre, S. (2000): *Stoffstrommanagement in der Metallindustrie. Zur Gestaltung techno-ökonomischer Bewertung industrieller Recyclingkonzepte*. VDI Verlag, Düsseldorf.
- Hanssman, F. (1993): *Einführung in die Systemforschung: Methodik der modellgestützten Entscheidungsvorbereitung*. Oldenbourg, München.
- Heidepriem, J. (2000): *Prozessinformatik 1*. Oldenbourg Industrieverlag, München.
- Heijungs, R. (1994): A generic method for the identification of options for cleaner products. *Ecological Economics*, 10: 69-81.
- Heijungs, R. und S. Suh (2002): *The Computational Structure of Life Cycle Assessment*. Kluwer, Dordrecht.
- Henning, P. A. und H. Vogelsang (Hrsg.) (2007): *Taschenbuch Programmiersprachen*. Carl Hanser Verlag, München.
- Hofmeister, S. (1989): *Stoff- und Energiebilanzen. Zur Eignung des physikalischen Bilanz-Prinzips als Konzeption der Umweltplanung*. Technische Universität Berlin, Berlin.
- Hong, L. J. und B. L. Nelson (2009): A brief introduction to optimization via simulation. *41th Winter Simulation Conference*, Austin, TX.
- Hooke, R. und T. A. Jeeves (1961): "Direct Search" Solution of Numerical and Statistical Problems. *J. ACM*, 8(2): 212-229.
- Hurrion, R. (1980): An interactive visual simulation system for industrial management. *European Journal of Operational Research*, 5(4): 86-93.
- IBU (2010): *Betrieblicher Umweltschutz in Baden-Württemberg. Eine Informationsplattform des Wirtschaftsministeriums Baden-Württemberg*. Informationszentrum für betrieblichen Umweltschutz (IBU). Wirtschaftsministerium Baden-Württemberg, <http://www.umweltschutz-bw.de/>, 27.3.2010.
- IES, Institute for Environment and Sustainability, DG Joint Research Centre (2010): *LCA Tools, Services and Data*. <http://lca.jrc.ec.europa.eu/lcainfohub/toolList.vm>, 7.12.2010.
- ifu Hamburg (2011): *Umberto - Software für ihr Stoffstrommanagement*. <http://www.ifu.com/de/products/umberto/index.htm>, 25.1.2011.

- Isenmann, R. und M. von Hauff (2007): Einführung in die Industrial Ecology. In: Isenmann, R. und M. Von Hauff (Hrsg.): *Industrial Ecology: Mit Ökologie zukunftsorientiert wirtschaften*. Elsevier, München: 1-16.
- ISO 14040 (2006): *Umweltmanagement – Ökobilanz – Grundsätze und Rahmenbedingungen*. DIN Deutsches Institut für Normung e.V., Berlin.
- ISO 14044 (2006): *Umweltmanagement – Ökobilanz – Anforderungen und Anleitungen*. DIN Deutsches Institut für Normung e.V., Berlin.
- Jackson, T. (2002): Industrial ecology and cleaner production. In: Ayres, R. U. und L. W. Ayres (Hrsg.): *A Handbook of Industrial Ecology*. Edward Elgar, Northampton: 36-43.
- Jones, C. V. (1994): Visualization and Optimization. *ORSA Journal on Computing*, 6(3): 221-257.
- Jones, C. V. (1996): *Visualization and Optimization*. Kluwer Academic Publishers, Boston, Dordrecht, London.
- Kallrath, J. (2002): Combined strategic and operational planning - an MILP success story in chemical industry. *OR Spectrum*, 24: 315-341.
- Kannegiesser, M. und H.-O. Günther (2010): An integrated optimization model for managing the global value chain of a chemical commodities manufacturer. *J Oper Res Soc*, 62(4): 711-721.
- Kiener, S., N. Maier-Scheubeck, R. Obermaier und M. Weiß (2006): *Produktions-Management: Grundlagen der Produktionsplanung und -steuerung*. Oldenbourg Verlag, München Wien.
- Kloock, J. (1969): *Betriebswirtschaftliche Input-Output-Modelle*. Betriebswirtschaftlicher Verlag Dr. Th. Gabler, Wiesbaden.
- Klöpper, W. und B. Grahl (2009): *Ökobilanz (LCA). Ein Leitfaden für Ausbildung und Beruf*. WILEY-VCH, Weinheim.
- Kolda, T. G., R. M. Lewis und V. Torczon (2003): Optimization by Direct Search: New Perspectives on Some Classical and Modern Methods. *SIAM Review*, 45(3): 385-482.
- Koopmans, T. C. (1951): Analysis of Production as an Efficient Combination of Activities. In: Koopmans, T. C. (Hrsg.): *Activity analysis of production and allocation. Proceedings of a conference*. Wiley, New York.
- Lambrecht, H. (2009): Leistungsfähigkeit der implementierten Algorithmen. In: Schmidt, M., H. Lambrecht und A. Möller (Hrsg.): *Stoffstrombasierte Optimierung. Wissenschaftliche und methodische Grundlagen sowie softwaretechnische Umsetzung*. MV-Verlag, Münster.
- Lambrecht, H. und M. Schmidt (2010): Material Flow Networks as a Means of Optimizing Production Systems. *Chemical Engineering & Technology*, 33(4): 610-617.
- Lambrecht, H., M. Schmidt und A. Möller (2009): Stoffstromnetze aus Sicht der Optimierungstheorie. In: Heyde, F. und C. Tammer (Hrsg.): *Methoden der Mehrkriteriellen Entscheidungstheorie. Tagungsband des Workshops der GOR-Arbeitsgruppen „Entscheidungstheorie und –praxis“ und „OR im Umweltschutz“*. Shaker Verlag, Aachen.
- Lambrecht, H. und M. Zimmermann (2008): Combination of Optimization Methods and Material Flow Analysis for Improvement of Operational Material Use (KOMSA): Concept and its Implementation. In: Möller, A., B. Page und M. Schreiber (Hrsg.): *Environmental Informatics and Industrial Ecology*. Shaker Verlag, Aachen.

- Lang, C., M. Keil und D. Heubach (2003): *Stoffstrommanagement für klein- und mittelständische Unternehmen der Mineralwasserbranche*. Konsortium des Forschungsprojekts INTUS - BMBF-Fkz: 01 RU 0102.
- Leontief, W. (1986): *Input-Output Economics*. Oxford University Press, New York, Oxford.
- LfU (2000): *Betriebliche Energie- und Stoffstrommanagementsysteme. Methoden, Praxiserfahrungen, Software - eine Marktanalyse*. Landesanstalt für Umweltschutz Baden-Württemberg, Karlsruhe.
- Lifset, R. und T. E. Graedel (2002): Industrial Ecology: goals and definitions. In: Ayres, R. U. und L. W. Ayres (Hrsg.): *A Handbook of Industrial Ecology*. Edward Elgar, Northampton.
- LINDO (2004): *LINGO. The modeling language and optimizer*. LINDO Systems Inc., Chicago.
- Luenberger, D. G. und Y. Ye (2008): *Linear and Nonlinear Programming*. Springer.
- Meyer, M. (1996): *Operations Research, Systemforschung*. UTB, Stuttgart.
- Miller, G. A. (1956): The Magical Number Seven, Plus or Minus Two: Some Limits on our Capacity for Processing Information. *The Psychological Review*, 63(2): 81-97.
- Mittelman, H. D. (2010): *Decision Tree for Optimization Software*.
<http://plato.asu.edu/guide.html>, 27.5.2010.
- Möller, A. (1993): Datenerfassung für das Öko-Controlling: Der Petri-Netz-Ansatz. In: Arndt, H.-K. (Hrsg.): *Umweltinformationssysteme für Unternehmen*. Berlin.
- Möller, A. (1994): Stoffstromnetze. In: Hilty, L. M., A. Jaeschke, B. Page und A. Schwabl (Hrsg.): *Informatik für den Umweltschutz, Band II: Anwendungen für Unternehmen und Ausbildung*. Metropolis-Verlag, Marburg: 223-2302.
- Möller, A. (2000): *Grundlagen stoffstrombasierter Betrieblicher Umweltinformationssysteme*. Dissertation, Universität Hamburg, Fachbereich Informatik.
- Möller, A. und A. Rolf (1995): Methodische Ansätze zur Erstellung von Stoffstromanalysen unter besonderer Berücksichtigung von Petri-Netzen. In: Schmidt, M. und A. Schorb (Hrsg.): *Stoffstromanalysen in Ökobilanzen und Öko-Audits*. Springer, Berlin u.a.: 33-58.
- Müller-Beilschmidt, P. (1997): Software zur Unterstützung der Ökobilanzierung - ein Überblick. In: Schmidt, M. und A. Häuslein (Hrsg.): *Ökobilanzierung mit Computerunterstützung. Produktbilanzen und betriebliche Bilanzen mit dem Programm Umberto*. Springer, Berlin: 3-10.
- Müller-Christ, G. (2001): *Umweltmanagement. Umweltschutz und nachhaltige Entwicklung*. Vahlen, München.
- Nelder, J. A. und R. A. Mead (1965): A Simplex Method for Function Minimization. *Comp. J.*, 7(4): 308-313.
- Neumann, K. und M. Morlock (2002): *Operations Research*. Hanser, München.
- Niemeyer, G. (1972): *Die Simulation von Systemabläufen mit Hilfe von FORTRAN IV. GPSS auf FORTRAN-Basis*. Walter de Gruyter, Berlin. New York.
- Niemeyer, G. (1977): *Kybernetische System- und Modelltheorie. Systems dynamics*. Vahlen, München.

- OptTek Systems Inc. (2010): *Simulation Optimization*.
<http://www.optquest.com/About/simulationoptimization.html>, 21.10.2010.
- Penkuhn, T. (1997): *Umweltintegriertes Stoffstrommanagement in der Prozeßindustrie - dargestellt am Beispiel der operativen Produktionsplanung der Ammoniaksynthese*. Peter Lang, Frankfurt am Main.
- Peters, D. (2001): *Unterstützung des Entwurfsprozesses von Mikrosystemen mit direkten Optimierungsverfahren*. Doktorarbeit, Universität Bremen.
- Peters, D., M. Gorges-Schleuter, W. Jakob und S. Parodat (1999): Modellbasierte Optimierung mit mathematischen Methoden - ein Vergleich von klassischen und evolutionären Verfahren. 7. GMM-Workshop "Methoden und Werkzeuge zum Entwurf von Mikrosystemen".
- Petri, C. A. (1962): *Kommunikation mit Automaten*. Schriftenreihe des Instituts für Instrumentelle Mathematik, Bonn.
- Pomberger, G. und H. Dobler (2008): *Algorithmen und Datenstrukturen. Eine systematische Einführung in die Programmierung*. Pearson Studium, München.
- Rardin, R. L. und R. Uzsoy (2001): Experimental Evaluation of Heuristic Optimization Algorithms: A Tutorial. *Journal of Heuristics*, 7: 261-304.
- Rechenberg, I. (1973): *Evolutionsstrategie - Optimierung technischer Systeme nach Prinzipien der biologischen Evolution*. Fommann-Holzboog, Stuttgart.
- Ropohl, G. (1999): *Allgemeine Technologie*. Hanser, München.
- Schaltegger, S. und A. Sturm (1995): *Öko-Effizienz durch Öko-Controlling: Zur praktischen Umsetzung von EMAS und ISO 14'001*. Schäffer-Poeschel Verlag, Stuttgart.
- Schmidt, A. (2007): *Stoffstrombilanzierungsmodelle für thermische Abfallbehandlungsanlagen zur Ermittlung und Bewertung von optimalen Betriebszuständen*. Diplomarbeit, Hochschule Bremen.
- Schmidt, M. (1997): Stoffstromnetze zwischen produktbezogener und betrieblicher Ökobilanzierung. In: Schmidt, M. und A. Häuslein (Hrsg.): *Ökobilanzierung mit Computerunterstützung. Produktbilanzen und betriebliche Bilanzen mit dem Programm Umberto*. Springer, Berlin: 11-24.
- Schmidt, M. (2008a): The Sankey Diagram in Energy and Material Flow Management. Part I: History. *Journal of Industrial Ecology*, 12(1): 82-94.
- Schmidt, M. (2008b): The Sankey Diagram in Energy and Material Flow Management. Part II: Methodology and Current Applications. *Journal of Industrial Ecology*, 12(2): 173-185.
- Schmidt, M. (2008c): *Umweltorientierte Stoffstromanalysen unter Berücksichtigung produktions-theoretischer Ansätze*. Doktorarbeit, Leuphana Universität Lüneburg, Lüneburg.
- Schmidt, M., H. Lambrecht und A. Möller (2007): Optimisation Approaches in Material Flow Models of Manufacturing Systems. In: Hryniewicz, O., J. Studzinski und M. Romaniuk (Hrsg.): *Environmental Informatics and System Research*. Shaker Verlag, Aachen: 271-278.
- Schmidt, M. und M. Schneider (2010): Kosteneinsparungen durch Ressourceneffizienz in produzierenden Unternehmen. *uwf*, 18(2010): 153-164.
- Scholl, A. (2001): *Robuste Planung und Optimierung: Grundlagen, Konzepte und Methoden, Experimentelle Untersuchungen*. Physica-Verl., Heidelberg.

- Scholl, M. H. (2006): Funktionale vs. imperative Programmierung. *Folien zur Vorlesung Deklarative Programmierung*. Fachbereich Informatik und Informationswissenschaft, Konstanz <http://www.inf.uni-konstanz.de/dbis/teaching/ss06/fp/fp.pdf>.
- Schwefel, H.-P. (1995): *Evolution and Optimum Seeking*. Wiley-Interscience, New York.
- Spengler, T. (1998): *Industrielles Stoffstrommanagement. Betriebswirtschaftliche Planung und Steuerung von Stoff- und Energieströmen in Produktionsunternehmen*. Erich Schmidt Verlag, Berlin.
- Statistisches Bundesamt (2010): *Produzierendes Gewerbe. Kostenstruktur der Unternehmen des Verarbeitenden Gewerbes sowie des Bergbaus und der Gewinnung von Steinen und Erden. 2008*. Fachserie 4 Reihe 4.3 Statistisches Bundesamt, Wiesbaden.
- Strohhecker, J. (1998): *System- und objektorientierte Simulation betriebswirtschaftlicher Entscheidungen*. Duncker & Humblot, Berlin.
- Syrjakow, M. (2003): *Web- und Komponenten-Technologien in der Modellierung und Simulation*. Habilitation, Fakultät für Informatik, Universität Karlsruhe.
- Thißen, N. (2009): Mehrstufiger Batchprozess in der chemischen Industrie. In: Schmidt, M., H. Lambrecht und A. Möller (Hrsg.): *Stoffstrombasierte Optimierung. Wissenschaftliche und methodische Grundlagen sowie deren softwaretechnische Umsetzung*. Monsenstein und Vannerdat, Münster.
- Thißen, N. (2010): Mass and Energy Flow Analysis Supports Process Improvement. *Chemical Engineering & Technology*, 33(4): 573-581.
- Tillman, A.-M. (2000): Significance of decision-making for LCA methodology. *Environ. Impact Assessment Rev.*, 20(1): 113-123.
- Trosset, M. C. (2000): *On the Use of Direct Search Methods for Stochastic Optimization*. Tech. Rep. 00-20, Department of Computational and Applied Mathematics, Rice University, Houston, TX.
- Viere, T., H. Brüner und J. Hedemann (2010): Verbund-Simulation - Strategic planning and optimization of integrated production networks. *Chemical Engineering & Technology*, 33(4): 582-588.
- Walk, W. (2009): Softwareunterstützung zur Erteilung von Verwertungs-Gutschriften in „attributiven“ Ökobilanzen. In: Feifel, S., W. Walk, S. Wursthorn und L. Schebeck (Hrsg.): *Ökobilanzierung 2009: Ansätze und Weiterentwicklungen zur Operationalisierung von Nachhaltigkeit. Tagungsband der fünften Ökobilanz-Werkstatt*. Universitätsverlag Karlsruhe, Karlsruhe: 117-124.
- Weidema, B. P., M. Thrane, P. Christensen, J. Schmidt und S. Løkke (2008): Carbon Footprint. A Catalyst for Life Cycle Assessment? *Journal of Industrial Ecology*, 12(1): 3-6.
- Witte, E., R. Bronner, P. R. Wossidlo und O. Grün (1988): Innovative Entscheidungsprozesse. In: Witte, E., J. Hauschildt und O. Grün (Hrsg.): Mohr (Paul Sibeck), Tübingen: 200-321.
- Wöhe, G. und U. Döring (2000): *Einführung in die allgemeine Betriebswirtschaftslehre*. Vahlen, München.
- Wohlgemuth, V. (2005): *Komponentenbasierte Unterstützung von Methoden der Modellbildung und Simulation im Einsatzkontext des betrieblichen Umweltschutzes: Konzeption und prototypische Entwicklung eines Stoffstromsimulators zur Integration einer stoffstrombasier-*

ten Perspektive in die auftragsbezogene Simulationssicht. Dissertation, Fachbereich Informatik, Universität Hamburg, Hamburg.

Wolff, S., H. Albers, M. Wittmaier, S. Langer, H. Lambrecht und T. Brinkmann (2009): Resource-Manager: A material flow model for recycling and waste management. In: Kungolos, A., K. Aravossis, A. Karagiannidis und P. Samaras (Hrsg.): *2nd International Conference on Environmental Management, Engineering, Planning and Economics (CEMEPE) & SECOTOX Conference.* GRAFIMA Verlag, Thessaloniki.

Wolfram Research Inc. (2010a): *Elementary Functions.*
<http://functions.wolfram.com/ElementaryFunctions/>, 27.1.2011.

Wolfram Research Inc. (2010b): *Global Optimization.*
<http://www.wolfram.com/products/applications/globalopt/>, 7.6.2010.

Wolfram Research Inc. (2010c): *Heaviside Theta.*
<http://functions.wolfram.com/GeneralizedFunctions/HeavisideTheta/02/>, 1.2.2011.

Ziena Optimization Inc. (2009): *KNITRO 6.0. User's Manual.*

Zimmermann, M., T. Viere und I. Meinshausen (2009): Prototypische Umsetzung des Optimierungsmoduls. In: Schmidt, M., H. Lambrecht und A. Möller (Hrsg.): *Stoffstrombasierte Optimierung. Wissenschaftliche und methodische Grundlagen sowie softwaretechnische Umsetzung.* MV-Verlag, Münster.