

(12) NACH DEM VERTRAG ÜBER DIE INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES
PATENTWESENS (PCT) VERÖFFENTLICHTE INTERNATIONALE ANMELDUNG

(19) Weltorganisation für geistiges Eigentum
Internationales Büro



(43) Internationales Veröffentlichungsdatum
23. Oktober 2003 (23.10.2003)

PCT

(10) Internationale Veröffentlichungsnummer
WO 03/087957 A2

(51) Internationale Patentklassifikation⁷: **G05B 13/00**

SC, SD, SE, SG, SK, SL, TJ, TM, TN, TR, TT, TZ, UA,
UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(21) Internationales Aktenzeichen: PCT/EP03/03424

(84) Bestimmungsstaaten (*regional*): ARIPO-Patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), eurasisches Patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), europäisches Patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI-Patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

(22) Internationales Anmeldedatum:
2. April 2003 (02.04.2003)

(25) Einreichungssprache: Deutsch

(26) Veröffentlichungssprache: Deutsch

Erklärung gemäß Regel 4.17:

— hinsichtlich der Berechtigung des Anmelders, ein Patent zu beantragen und zu erhalten (Regel 4.17 Ziffer ii) für die folgenden Bestimmungsstaaten AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, UZ, VC, VN, YU, ZA, ZM, ZW, ARIPO-Patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), eurasisches Patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), europäisches Patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI-Patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG)

(30) Angaben zur Priorität:
102 16 558.0 15. April 2002 (15.04.2002) DE

(71) Anmelder (für alle Bestimmungsstaaten mit Ausnahme von US): **BAYER AKTIENGESELLSCHAFT** [DE/DE]; 51368 Leverkusen (DE).

(72) Erfinder; und

(75) Erfinder/Anmelder (nur für US): **SCHUPPERT, Andreas** [DE/DE]; Stockberggasse 46, 51515 Kürten (DE). **OHRENBERG, Arne** [DE/DE]; Takustr. 2, 50825 Köln (DE).

(74) Gemeinsamer Vertreter: **BAYER AKTIENGESELLSCHAFT**; 51368 Leverkusen (DE).

Veröffentlicht:

— ohne internationalen Recherchenbericht und erneut zu veröffentlichen nach Erhalt des Berichts

Zur Erklärung der Zweibuchstaben-Codes und der anderen Abkürzungen wird auf die Erklärungen ("Guidance Notes on Codes and Abbreviations") am Anfang jeder regulären Ausgabe der PCT-Gazette verwiesen.

(54) Title: METHOD AND COMPUTER FOR EXPERIMENTAL DESIGN

(54) Bezeichnung: VERFAHREN UND COMPUTER ZUR PLANUNG VON VERSUCHEN

(57) Abstract: The invention relates to a method for experimental design comprising the following steps: A selection of at least one first experiment from an experiment range by a data-driven optimizer, B input of data for the first experiment that has been determined by experimentation, C use of a meta-layer for evaluating the experiment data, whereby the meta-layer contains a neuronal network and/or a hybrid model and/or a rigorous model and/or a data mining method, D input of the experiment data determined by experimentation for the first experiment into the data-driven optimizer, E influencing of the data-driven optimizer by the result of the evaluation, F selection of at least one second experiment from the experiment range by the data-driven optimizer.

(57) Zusammenfassung: Die Erfindung betrifft ein Verfahren zur Planung von Versuchen mit folgenden Schritten: A. Auswahl von zumindest einem ersten Versuch aus einem Versuchsraum durch einen datengetriebenen Optimierer, B. Eingabe von experimentell ermittelten Versuchsdaten des ersten Versuchs, C. Verwendung einer Meta-Schicht für die Auswertung der Versuchsdaten, wobei die Meta-Schicht ein neuronales Netz und/oder ein Hybridmodell und/oder ein rigoroses Modell und/oder eine Data Mining Methode beinhaltet, D. Eingabe der experimentell ermittelten Versuchsdaten des ersten Versuchs in den datengetriebenen Optimierer, E. Beeinflussung des datengetriebenen Optimierers durch das Ergebnis der Auswertung, F. Auswahl von zumindest einem zweiten Versuch aus dem Versuchsraum durch den datengetriebenen Optimierer.



WO 03/087957 A2

Verfahren und Computersystem zur Planung von Versuchen

Die Erfindung betrifft ein Verfahren und ein Computersystem zur Planung von Versuchen sowie ein entsprechendes Computerprogrammprodukt.

Aus dem Stand der Technik ist es bekannt, Versuche mittels statistischer Versuchsplanungsverfahren zu planen. Solche Planungsverfahren werden unter anderem dazu benutzt, mit einer minimalen Anzahl von Versuchen ein empirisches Prozessmodell für den Zusammenhang zwischen den Regel- und Störgrößen in einem Prozess und den resultierenden Produkt- und Prozesseigenschaften zu ermitteln. Eine solche statistische Versuchsplanung lässt sich beispielsweise mit Hilfe des Computerprogramms „STAVEX“ (STATistische Versuchsplanung mit EXPerten-system, Hersteller AICOS Technologie, Schweiz) durchführen. Ein weiteres kommerziell erhältliches Computerprogramm für die Versuchsplanung ist das Programm „Statistica®“, der Fa. StatSoft (Europe) GmbH, Deutschland.

Im Bereich der statistischen Versuchsplanung werden im Stand der Technik verschiedene Versuchsplanungstypen unterschieden. Insbesondere unterscheidet man zwischen der klassischen, vollfaktoriellen Methode und modernen Verfahren nach Taguchi oder Shainin.

Die klassische, vollfaktorielle Methode ist der Ursprung aller statistischen Versuchsplanungsmethoden. Sie basiert auf einem Vergleich aller qualitätsbedingenden Faktoren miteinander in Anlehnung an die Varianzanalyse. Zahlreiche Varianten wurden im Laufe der vergangenen Jahrzehnte erarbeitet und in Forschungs- und Entwicklungslabors validiert.

Das Shainin-DOE (DOE = Design of Experiment) gilt als geeignetes Verfahren für die Prozessoptimierung, weil es sogenannte „starke“ Einflussgrößen isoliert und diese auf Relevanz und Abhängigkeit untersucht.

Das Taguchi-DOE basiert auf vorangegangenen, teilfaktoriellen, orthogonalen Versuchsplänen. Wegen der drastischen Versuchslaufeinsparungen durch Vorauswahl der wichtigsten Einflussgrößen handelt es sich hierbei um ein schnelles und relativ wirtschaftliches Verfahren der Versuchs- und Prozessplanung.

Weitere bekannte statistische Versuchsplanungstypen sind teilfaktorielle Versuchspläne ("fractional factorial"), Plackett-Burmann Versuchspläne, Zentral-zusammengesetzte Pläne („central composite“), Box-Behnken Versuchspläne, D-optimale Pläne („D-optimal Designs“), Mischungspläne, Balancierte Blockpläne, Lateinische Quadrate, Desperado Pläne (vgl. hierzu auch Eberhard Scheffler, Statistische Versuchsplanung and -auswertung; Deutscher Verlag für Grundstoffindustrie, Stuttgart, 1997).

Weitere Versuchsplanungsverfahren sind bekannt aus Hans Bendemer, „Optimale Versuchsplanung“, Reihe Deutsche Taschenbücher (DTB, Band 23, und ISBN 3-87144-278-X) sowie Wilhelm Kleppmann, Taschenbuch Versuchsplanung, „Produkte und Prozesse optimieren“ 2., erweiterte Auflage, ISBN: 3-446-21615-4. Diese Verfahren werden in der Praxis aus Kostengründen oftmals eingesetzt.

Nachteilig bei bekannten statistischen Versuchsplanungsverfahren ist, dass die Versuchsplanung und Modellierung ohne Berücksichtigung von zusätzlichem Wissen erfolgt, so dass u.U. keine geeigneten Optima gefunden werden und die Vertrauenswürdigkeit der generierten Ergebnisse und Aussagen fragwürdig ist. Ein weiterer gewichtiger Nachteil von vorbekannten Versuchsplanungsverfahren ist, dass diese bei einer großen Zahl zu berücksichtigender Einflussgrößen zu umfangreich werden. Ferner ist bei bestimmten Systemen z.B. in der Katalyse oder Wirkstoffforschung oftmals die Zielfunktion stark „zerklüftet“ und deshalb schwer mit statistischen Verfahren fassbar.

Aus der WO 00 / 15341 ist ein Verfahren zur Entwicklung von Feststoffkatalysatoren für heterogene katalysierte Reaktionsabläufe bekannt, welches auf einer parallelierten Ausprüfung nach evolutionären Methoden basiert.

Entsprechende evolutionär arbeitende Verfahren sind ferner bekannt aus der WO 00 / 43411, J. chem. Inf. Compute. Sci. 2000, 40, 981-987 „Heterogeneous Catalyst Design Using Stochastic Optimization Algorithms“ sowie aus Applied Catalysis A: General 200 (2000) 63-77 „An evolutionary approach in the combinatorial selection and optimization of catalytic materials“.

Ferner ist aus der US 6,009,379 ein Verfahren zur Steuerung eines Herstellungsprozesses über ein effizientes experimentelles Design bekannt. Dabei werden Versuchs-Punkte gleichmäßig auf einer mehrdimensionalen Kugeloberfläche verteilt, um die einzelnen Herstellungsparameter gleichmäßig zu gewichten.

Die Figur 1 zeigt ein Blockdiagramm eines aus dem Stand der Technik bekannten Systems zur Durchführung von Screening Versuchen, wie es insbesondere in den Bereichen Katalyse, Material- und Wirkstoffforschung eingesetzt wird. Das System beinhaltet eine Substanzbibliothek, das heißt eine sogenannte kombinatorische Library 1 und einen Versuchsaufbau 2 zur Durchführung von High-Throughput-Screening (HTS) oder High-Speed Experimentation (HSE) Versuchen. Solche Screening Versuche werden typischerweise für die Wirkstoffidentifizierung, die Katalysatorforschung (homogen und heterogen), die Materialforschung sowie die Identifizierung von optimalen Reaktionsbedingungen bei chemischen, biochemischen oder biotechnologischen Systemen eingesetzt.

Üblicherweise werden in einem solchen Versuchsaufbau 2 mehrere Versuche parallel durchgeführt. Die Versuchsergebnisse werden in Form einer Datei 3 ausgegeben. Diese Ausgangsdaten oder Teile davon sind zugleich die Eingangsdaten für einen Optimierer 4.

- 4 -

Bei dem Optimierer 4 handelt es sich um einen sogenannten Black-Box-Optimierer, das heißt ein Optimierer, der auf einem datengetriebenen Modell oder auf einem evolutionären Algorithmus basiert. A priori Wissen über Struktur und/oder Wirkungsbeziehungen sind in dem Optimierer 4 nicht vorhanden, vielmehr beschränkt sich dieser auf die Auswertung der Daten als solchen, um eine Auswahl von Versuchen aus der kombinatorischen Library 1 zu treffen.

Typischerweise verwendet der Optimierer 4 die Versuchsdaten 3 bestehend aus Einflussgrößen (Attribute, Faktoren, Strukturmerkmale, Deskriptoren, physikalische Größen, Stoffeigenschaften) und einer Angabe hinsichtlich der Wirkung dieser Größen auf die sogenannten Targets (Zielgrößen), um eine optimale Suchrichtung im Raum der Targets festzulegen.

Die Realisierung eines solchen Black-Box-Optimierers 4 erfolgt beispielsweise über:

- genetische Algorithmen,
- evolutionäre Algorithmen oder Strategien,
- neuronale Netze oder
- sonstige datengetriebene Modellansätze, die auf stochastische oder deterministische oder aus beiden kombinierte Optimierungsstrukturen zurückgreifen.

Ein gemeinsamer Nachteil von solchen aus dem Stand der Technik bekannten Systemen ist, dass a priori Informationen nicht oder nur eingeschränkt in den Black-Box-Optimierer 4 Einfluss finden können und entsprechende Suchstrategien oft langsam oder gegen ungeeignete Suboptima konvergieren. Solche aus dem Stand der Technik bekannte Verfahren sind daher hinsichtlich des Zeit- und Kostenaufwands oft ineffizient. Bei Techniken auf der Basis von evolutionären Algorithmen besteht ferner die Gefahr, dass der Aufwand beim Einsatz des Optimierers zur Erreichung des Optimums höher sein kann als bei einem rationalen oder statistischen Vorgehen.

- 5 -

Der Erfindung liegt daher die Aufgabe zu Grunde ein verbessertes Verfahren zur Planung von Versuchen sowie ein entsprechendes Computersystem und Computerprogrammprodukt zu schaffen.

Die der Erfindung zu Grunde liegende Aufgabe wird mit den Merkmalen der unabhängigen Patentansprüche jeweils gelöst. Bevorzugte Ausführungsformen der Erfindung sind in den abhängigen Patentansprüchen angegeben.

Gegenstand der Erfindung ist ein Verfahren zur Planung von Versuchen zur Erreichung eines Optimierungszieles mit folgenden Schritten:

- A) Auswahl von zumindest einem ersten Versuch aus einem Versuchsraum durch einen datengetriebenen Optimierer in einer Rechneinheit,
- B) Eingabe von experimentell ermittelten Versuchsdaten des ersten Versuchs, in mindestens eine Meta-Schicht in eine Rechneinheit,
- C) Verwendung mindestens einer Meta-Schicht für die Auswertung der Versuchsdaten,
- D) Eingabe der experimentell ermittelten Versuchsdaten des ersten Versuchs in den datengetriebenen Optimierer,
- E) Beeinflussung des datengetriebenen Optimierers durch das Ergebnis der Auswertung in der Meta-Schicht und Prüfung der Zielerreichung,
- F) Auswahl von zumindest einem zweiten Versuch aus dem Versuchsraum durch den datengetriebenen Optimierer,
- G) Wiederholung der Schritte B) bis E) für die Daten des zweiten Versuchs

und

- H) Abbruch der Hexation bei Zielerreichung oder Wiederholung der Schritte A) bis F) für mindestens einen dritten oder weitere Versuche bis zur Zielerreichung.

Das Verfahren wird wiederholt bis das Optimierungsziel erreicht ist oder festgestellt wird, dass das Optimierungsziel möglicherweise nicht erreicht werden kann. Der Abbruch des Verfahrens kann automatisch oder durch den Benutzer erfolgen. Optimierungsziel können das Erreichen bestimmter Bewertungskennzahlen der Versuche sein. Bei den Kennzahlen kann es sich beispielsweise um Ausbeuten Selektivitäten, Raum-Zeit-Ausbeuten, Kosten, physikalische Eigenschaften, Wirkmechanismen, abgeleitete Eigenschaften usw. handeln. Eine Bewertung der Versuche unter Verwendung mehrerer Kennzahlen ist ebenfalls möglich.

Die Erfindung erlaubt die Integration von Wissen zur Beeinflussung des Black-Box-Optimierers mit dem Ziel der Beschleunigung der Konvergenz und/oder der Sicherstellung der Konvergenz zu einem geeigneten Optimum sowie einer Steigerung der Vertrauenswürdigkeit der Ergebnisse. Das Wissen kann dabei als Vorwissen a priori bekannt sein und/oder laufend durch die Auswertung von zuvor durchgeführten Versuchen ergänzt werden.

Die Generierung von zusätzlichem Wissen erfolgt dabei vorzugsweise in Form von „Regeln“, insbesondere über die Struktur-Wirkungsbeziehung mit Data Mining und anderen Verfahren. Diese Regeln können vor, während oder nach einem Optimierungsschritt oder auch fortlaufend in die Versuchsplanung integriert werden, wobei der datengetriebene Optimierer entsprechend beeinflusst wird. Für diese Beeinflussung des datengetriebenen Optimierers ist eine Meta-Schicht vorgesehen.

Durch Einsatz einer solchen Meta-Schicht wird der Black-Box-Optimierer „getuned“. Dabei ist die Meta-Schicht nicht auf eine Methode beschränkt, sondern kann die Kombination verschiedener Methoden beinhalten. Mögliche Methoden sind:

- neuronale Netze,
- Hybridmodell,
- rigorose Modelle,
- Data Mining Methoden, wie zum Beispiel Entscheidungsbaumverfahren, allgemeine Separationsverfahren, Untergruppensuchverfahren, allgemeine Partitionsverfahren, Clusterverfahren, Assoziationsregelgeneratoren und Korrelationsverfahren.

Die Beeinflussung der Arbeitsweise des Optimierers kann dabei unmittelbar erfolgen, indem in die Arbeitsweise des Optimierers eingegriffen wird oder mittelbar, indem die Daten, welche die Basis für die Optimierung bilden, gefiltert werden.

Nach einer bevorzugten Ausführungsform der Erfindung werden Methoden zur Beeinflussung des Optimierers verwendet, durch die der Optimierer und/oder der Optimierungsprozess „getuned“ werden. Zu solchen Methoden gehören beispielsweise Untergruppensuchverfahren oder Korrelationsanalysen oder Attributstatistiken bei Regelgeneratoren.

Nach einer weiteren bevorzugten Ausführungsform der Erfindung sind weitere Meta-Layer vorgesehen, die den jeweils vorhergehenden Meta-Layer verbessern oder in den oder die vorhergehenden Meta-Layer und/oder zusätzlich direkt in den Black-Box-Optimierungsprozess der ersten Ebene eingreifen.

Nach einer weiteren bevorzugten Ausführungsform der Erfindung können in jedem Optimierungsschritt die Eingriffspositionen in den ursprünglichen Optimierungsprozess und die in dem oder den Meta-Layern eingesetzten Methoden oder Kombi-

nationen von Methoden variiert werden. Die Auswahl geeigneter Methoden zur Generierung optimaler Regeln kann dabei automatisch erfolgen.

Nach einer bevorzugten Ausführungsform der Erfindung erfolgt die Beeinflussung des Optimierers durch eine Neubewertung der Versuchsdaten. Beispielsweise können die Versuchsdaten selbst bereits eine Auswertung beinhalten, indem entsprechende Daten, wie zum Beispiel die Ausbeute unmittelbar experimentell ermittelt werden. In diesem Fall kann die Neubewertung durch Filterung der Ausbeute-Daten erfolgen, zum Beispiel indem besonders gute Ausbeuten durch die Datenfilterung stärker gewichtet und besonders schlechte Ausbeuten durch die Datenfilterung schwächer gewichtet werden. Durch diese Art der Datenfilterung lässt sich eine schnellere Konvergenz des Versuchsablaufs erreichen.

Entsprechend kann auch vorgegangen werden, wenn die Versuchsdaten nicht unmittelbar eine experimentell ermittelte Bewertung beinhalten, sondern die Bewertung erst durch dem Experiment nachgeordnete Berechnungen ermittelt wird. In diesem Fall werden nicht experimentell ermittelte Daten, sondern durch Berechnung ermittelte Bewertungen gefiltert oder gewichtet.

Die Art und Weise der Filterung ergibt sich dabei durch Regeln oder andere Zusammenhänge, die aufgrund eines Analyseverfahrens der Versuchsdaten, beispielsweise mittels neuronalen Netzen oder Data Mining Verfahren oder anderen Verfahren aufgefunden wurden.

Nach einer weiteren bevorzugten Ausführungsform erfolgt die Beeinflussung des Optimierers durch eine Verkleinerung, Vergrößerung und/oder Verschiebung des Versuchsraums.

Nach einer weiteren bevorzugten Ausführungsform erfolgt die Filterung durch Vorselektion und/oder Gewichtung der Versuchsdaten. Besonders „schlechte“ Versuchsdaten, das heißt, solche Versuchsdaten, die von, beispielsweise einem Regel-

generator, als ungeeignet erkannt worden sind, werden vorselektiert und aus dem Versuchsraum eliminiert. Ferner können auch ganze Spalten oder Zeilen aus der Versuchsdatenmatrix eliminiert werden, wenn die entsprechenden Parameter von dem Regelgenerator als nicht relevant erkannt worden sind. Hierdurch wird der Versuchsraum verkleinert, was den Gesamtaufwand erheblich verringert.

Eine Gewichtung der Versuchsdaten kann dadurch erfolgen, dass als besonders relevant erkannte Versuchsdaten in der Versuchsdatenmatrix ein oder mehrfach dupliziert werden. Alternativ kann ein Gewichtungskoeffizient eingeführt werden.

Nach einer weiteren bevorzugten Ausführungsform der Erfindung beinhaltet der Black-Box-Optimierer sogenannte Kernmodule oder Kernoperatoren, sowie ein Modul zur Selektion neuer Versuchspunkte. Die Beeinflussung der Arbeitsweise des Optimierers erfolgt dann durch eine Beeinflussung des oder der Kernmodule und/oder des Moduls zur Selektion neuer Versuchspunkte basierend auf Zusammenhängen, die zum Beispiel von einem Regelgenerator erkannt worden sind.

Im weiteren werden bevorzugte Ausführungsformen der Erfindung mit Bezugnahme auf die Zeichnungen näher erläutert. Es zeigen:

- Figur 1 ein Blockdiagramm zur Darstellung eines aus dem Stand der Technik bekannten Versuchsplanungssystems,
- Figur 2 ein Blockdiagramm einer Ausführungsform eines erfindungsgemäßen Versuchsplanungssystems,
- Figur 3 ein Blockdiagramm einer Ausführungsform des erfindungsgemäßen Versuchsplanungssystems mit einer Neubewertung der Versuchsdaten,

- Figur 4 eine Ausführungsform des erfindungsgemäßen Versuchsplanungssystems mit einer Vorselektion und/oder Gewichtung der Versuchsdaten,
- Figur 5 eine Ausführungsform des erfindungsgemäßen Versuchsplanungssystems mit einer Beeinflussung der Selektion neuer Versuchspunkte des Optimierers,
- Figur 6 eine Ausführungsform des erfindungsgemäßen Versuchsplanungssystems mit einer Beeinflussung des oder der Kernmodule des Optimierers.

Das Versuchsplanungssystem der Figur 2 basiert auf einer kombinatorischen Library 5, die sich aufgrund der durch einen Versuchsraum gegebenen Rahmenbedingungen ergibt. Aus dieser kombinatorischen Library 5 werden durch einen Optimierer 6 ein oder mehrere Versuche ausgewählt, die dann in einem Versuchsaufbau 7 beispielsweise mittels eines High-Throughput-Screening oder High-Speed-Experimentation Versuchsverfahren durchgeführt werden. Die entsprechenden Versuchsdaten werden in Form einer Datei 8 ausgegeben.

Für den Optimierer 6 ist in dem Versuchsplanungssystem eine Meta-Schicht 9 vorgesehen. Die Meta-Schicht 9 dient zur Beeinflussung des Optimierers 6 unter Berücksichtigung von a priori bekannten Wissen oder von während der Versuchsdurchführung hinzugewonnenem Wissen. Wissen, zum Beispiel in Form von Regeln oder in Form von trainierten neuronalen Netzen kann dabei fortlaufend durch die Auswertung von Dateien 8 hinzugewonnen werden.

Die Meta-Schicht 9 ergänzt also den datengetriebenen Optimierer 6 durch zusätzliches Wissen, um eine Konvergenz der Versuchsreihe zu beschleunigen. Durch die Meta-Schicht 9 lässt sich also die Konvergenzgeschwindigkeit eines Black-Box

Optimierungsverfahrens, welches in dem Optimierer 6 realisiert ist, durch Integration von Vorwissen und/oder Regelstrukturen verbessern.

Diese Integration kann auf verschiedene Arten und Weisen erfolgen, wie zum Beispiel durch:

- A Informationsgestützte Zusatzselektion der Testensembles, d.h. Einschränkung der zu prüfenden kombinatorischen Library durch die mit Datamining gefundenen Regeln und kein Eingriff in den Optimierer
- B gezielte Gewichtung der Optimierungsschritte in Richtung von als optimal identifizierten Librarybereichen, d. h. Eingriff in das Suchverfahren des Optimierers
- C Tuning der Selektionsregeln der Black-Box-Optimierungsverfahren, d. h. direkter Eingriff in das Bewertungsverfahren des Optimierers oder Modifikation der Bewertungsgrößen vor Eingang in den Optimierer

Die Eingriffsformen A, B und C können grundsätzlich auch in Kombination erfolgen, d.h. es können bei einem Optimierungsschritt auch Eingriffe bei A und B, B und C, A und C oder A und B und C erfolgen. Von Optimierungsschritt zu Optimierungsschritt können sich die Eingriffspositionen und Eingriffskombinationen sowie die im Metalayer verwendeten Methoden ändern. Die Eingriffe können auch aus nachgeordneten Metalayers erfolgen.

Bei der Optimierung durch statistische Versuchsplanung ist das Vorgehen ähnlich zu der Verwendung eines Black-Box-Optimierers, das heißt auch hier wird in dem Optimierungsprozess über die Meta-Schicht in einer oder mehreren der oben beschriebenen Formen eingegriffen. Zum Beispiel erfolgt die Integration von Vorwissen, indem bei der Wahl der Einflussgrößen deren Gültigkeitsbereich und/oder zusätzliche Beschränkungen des Gültigkeitsbereichs bei der Kombination von Einflussgrößen einfließen.

Für die sequentielle statistische Versuchsplanung können durch den Einsatz von Data Mining Methoden oder anderen oben beschriebenen Methoden weitere Informationen über Einflussgrößen einfließen und in die Versuchsplanung integriert werden, das heißt nach dem ersten, zweiten ..., n-ten Durchlauf wird jeweils aufgrund der zusätzlichen Informationen der Versuchsraum verändert.

Die Veränderung erfolgt durch

- a) Hinzunahme oder Entfernung von Einflussgrößen
- b) Veränderung der Gültigkeitsbereiche der einzelnen oder kombinierten Einflussgrößen
- c) Kombination aus a) und b).

Von besonderem Vorteil ist dabei, dass „klassische“ aus dem Stand der Technik an sich bekannte Versuchsplanungsverfahren für einen Black-Box - oder einen statistischen Optimierer weiter benutzt werden können; diese Versuchsplanungsverfahren werden durch die vorliegende Erfindung verbessert, indem durch die Berücksichtigung von Vorwissen oder während des Versuchsablaufs hinzu gewonnenem Wissen das Verfahren schneller oder überhaupt konvergiert.

Insbesondere wird die Konvergenzgeschwindigkeit durch das erfindungsgemäße „Tuning“ beim Optimieren der Versuchsplanung für Katalysatoren, Wirkstoffe oder Materialien oder Reaktionsbedingungen deutlich erhöht. Ein weiterer Vorteil ist, dass sich die Anzahl der Versuche bei gleicher Ergebniserwartung reduzieren lässt, womit ein geringerer Material- und Zeitaufwand und eine bessere Ausnutzung der Anlagen erzielbar ist.

Von besonderem Vorteil ist ferner, dass durch die Integration des Vorwissens kein Verlust von Forschungsinvestitionen beim Einsatz von HSE- oder HTS-Technologien oder bei einem kombinatorischen Vorgehen gegeben ist.

Die Figur 3 zeigt eine Ausführungsform des Versuchsplanungssystems, bei der eine Neubewertung der Versuche erfolgt.

In dem Versuchsaufbau 7 werden ein oder mehrere zuvor aus der kombinatorischen Library 5 (vgl. Figur 2) ausgewählte Versuche durchgeführt. Die entsprechenden Versuchsdaten werden in Form der Datei 8 ausgegeben. Die Versuchsdaten können dabei selbst bereits eine Bewertung beinhalten, wenn entsprechende Daten unmittelbar experimentell erfasst werden können. Ein Beispiel hierfür ist die experimentelle Bestimmung der Ausbeute. Die Ausbeute ist gleichzeitig eine Bewertung der durchgeführten Versuche.

In anderen Fällen kann es erforderlich sein, zusätzlich eine Bewertung der Versuchsdaten in einem Bewertungsmodul 10 vorzunehmen. Beispielsweise beinhaltet das Bewertungsmodul 10 eine Berechnungsvorschrift für die Berechnung einer Bewertung basierend auf einen oder mehreren der Versuchsdaten.

Die Datei 8 und gegebenenfalls das Ergebnis des Bewertungsmoduls 10 werden in die Meta-Schicht 9 eingegeben. Die Meta-Schicht 9 beinhaltet ein Modul 11 zur Realisierung eines Data Mining (DM) Algorithmus, eines neuronalen Netzes oder eine Hybridverfahrens oder eines anderen geeigneten Datenanalyseverfahrens.

Durch die Anwendung eines solchen Verfahrens werden Regeln generiert, das heißt zusätzliche Informationen und Erkenntnisse zum Verständnis des in den Versuchen betrachteten chemischen Systems. Das Modul 11 hat also die Funktion eines Regelgenerators. Entsprechende Regeln und Nebenbedingungen werden in dem Modul 12 der Meta-Schicht 9 formuliert.

Basierend auf diesen Regeln und Nebenbedingungen erfolgt dann gegebenenfalls in dem Modul 13 eine Neubewertung des oder der Versuche. Dies kann so erfolgen, dass ein Neuberwertung eines Versuchs nur dann vorgenommen wird, sofern ein vorgegebener Schwellwert überschritten wird. Alternativ kann auch der Benutzer

eingreifen, um die Neubewertung zu aktivieren oder zu deaktivieren. Die Neubewertung kann darin bestehen, dass als „schlecht“ erkannte Versuche eine verschlechterte Bewertung erhalten und als „gut“ erkannte Versuche eine verbesserte Bewertung.

Basierend auf der Datei 8, die gegebenenfalls neu bewertete Versuchsdaten beinhaltet, erstellt dann der Black-Box-Optimierer 6 einen weiteren Versuchsplan 18. Die entsprechenden Versuche werden dann wiederum in dem Versuchsaufbau durchgeführt und so weiter.

Die Figur 4 zeigt eine alternative Ausführungsform, bei der die Filterung nicht durch eine Neubewertung der Versuchsdaten, sondern durch eine Vorselektion und/oder Gewichtung erfolgt. Das Versuchsplanungssystem der Figur 4 ist dabei prinzipiell gleich aufgebaut, wie das der Figur 3, wobei an Stelle des Moduls 13 ein Modul 15 für die Vorselektion und/oder Gewichtung verwendet wird.

Die Versuchsdaten werden also nicht neu oder anders bewertet, sondern es können mittels des Moduls 15 auf der Grundlage der ermittelten Regeln beispielsweise Versuche ausgeblendet oder stärker oder schwächer gewichtet werden. Dadurch findet eine Vorselektion statt, ohne dass die Bewertung der Versuche selbst geändert wird.

Die Figur 5 zeigt eine weitere Ausführungsform des erfindungsgemäßen Versuchsplanungssystems. Die Ausführungsform der Figur 5 unterscheidet sich von den Ausführungsformen der Figur 3 und der Figur 4 dadurch, dass unmittelbar in den Optimierer 6 eingegriffen wird.

Der Optimierer beinhaltet in dieser Ausführungsform ein oder mehrere Kernmodule 16, das heißt sogenannte Kernoperatoren. Ferner beinhaltet der Optimierer 6 ein Modul 17 zur Selektion neuer Versuchspunkte. Die Arbeitsweise des Moduls 17 wird durch die von dem Modul 12 formulierten Regeln und Nebenbedingungen beein-

flusst, das heißt es werden beispielsweise von dem Modul 17 selektierte neue Versuchspunkte verworfen, so dass eine Rückkopplung von dem Modul 17 zu dem Kernmodul 16 erfolgt, um entsprechende weitere Versuchspunkte als Ersatz für die verworfenen Versuchspunkte zu selektieren.

Nach Ablauf der eigentlichen Optimierung in dem oder den Kernmodulen 16, werden also durch das Modul 17 neue Versuche bzw. Versuchspunkte zur Optimierung der Zielgrößen des betrachteten Systems vorgeschlagen. Bei diesem System kann es sich beispielsweise um ein chemisches, biotechnologisches, biologisches, enzymatisches System handeln.

Auf der Grundlage der über den Regelgenerator, das heißt die Meta-Schicht 9 erarbeiteten Regeln werden Versuche, die den generierten Regeln widersprechen, aussortiert und gegebenenfalls durch neue Versuche über den Optimierer, das heißt, das oder die Kernmodule 16, ergänzt.

Das Aussortieren kann dabei streng, das heißt vollständig, oder weich, mit einer gewissen Gewichtung erfolgen. Auch die neu geplanten Versuche müssen dann wiederum durch das Modul 17. Auf diese Weise wird erreicht, dass Informationen, die von dem oder den Kernmodulen 16 nicht berücksichtigt werden oder berücksichtigt werden können, nachträglich in die Versuchsplanung einfließen.

Alternativ kann ein separates Modul 18 dem Optimierer 6 nachgeschaltet sein, um die Nachselektion der von dem Modul 17 selektierten neuen Versuchspunkte vorzunehmen. Dies entspricht einem Test in dem Modul 18, ob die neuen Versuchspunkte, die von dem Modul 17 vorgeschlagen worden sind, regelkonform sind. Wenn bei diesem Test Versuchspunkte eliminiert werden, ist wiederum ein Rückfluss erforderlich, um entsprechende alternative neue Versuchspunkte zu planen.

Die Ausführungsform des Versuchsplanungssystems der Figur 6 entspricht dem Versuchsplanungssystem der Figur 5 mit dem Unterschied, dass nicht die Arbeits-

weise des Moduls 17 beeinflusst wird oder eine Nachselektion in dem Modul 18 stattfindet, sondern dass unmittelbar die Arbeitsweise des oder der Kernmodule 16 des Optimierers 6 beeinflusst wird.

Beispiele für Kernoperatoren von neuronalen Netzen sind die Art und Anzahl von Einflussgrößen sowie die Gewichtung einzelner Datenpunkte.

Beispiele für Kernoperatoren von evolutionären Algorithmen am Beispiel des genetischen Algorithmus sind der Selektionsoperator (Auswahl einer neuen Reihe von Versuchen), der Mutationsoperator und der Cross-Over Operator.

Die Regeln und Informationen, die durch den Regelgenerator erzeugt werden, werden in dem Algorithmus des Optimierers selbst bei der Ausführung berücksichtigt.

Für Optimierer, die mit Neuronalen Netzen gekoppelt sind, bedeutet dies, dass durch die Regeln der Versuchsraum eingeschränkt oder die Datensätze in besonderer Weise gewichtet werden.

Bei evolutionärer Algorithmus-Optimierern werden die zusätzlichen Informationen in einem oder mehreren Kernoperatoren berücksichtigt. Dies bedeutet, dass z.B. bestimmte Kreuzungen, Selektionen oder Mutationen verboten oder bevorzugt durchgeführt werden.

Für beide Optimierertypen hat dies zur Folge, dass bei einer vollständigen Automatisierung des Workflows über Schnittstellen in die entsprechenden Programmteile des Optimierers eingegriffen wird oder die Informationen über manuelle oder programmgesteuerte Änderungen von Optimierungsparametern in den Optimierer einfließen.

Die Ausführungsformen der Figuren 3 bis 6 können miteinander kombiniert werden, das heißt es können mehrere Regelgeneratoren, das heißt Meta-Schichten 9 unabhängig voneinander in den Optimierungsablauf eingebunden sein. Diese Regeln

können mit verschiedenen, unabhängig voneinander eingesetzten Verfahren und Methoden erzeugt und in dem Modul 12 zusammengeführt werden.

Die Berücksichtigung der Regeln, die von dem oder den Regelgeneratoren der Meta-Schichten 9 formuliert worden sind, erfolgt entweder automatisch über definierte Schnittstellen und unter Einhaltung vorgegebener Schwellwerte oder durch manuelle Formulierung von Regeln für den Bereich des Optimierers, in den der Regelgenerator eingreift.

Bezugszeichenliste

kombinatorische Library	1
Versuchsaufbau	2
Datei	3
Optimierer	4
Kombinatorische Library	5
Optimierer	6
Versuchsaufbau	7
Datei	8
Meta-Schicht	9
Bewertungsmodul	10
Modul	11
Modul	12
Modul	13
Versuchsplan	14
Modul	15
Kernmodul	16
Modul	17
Modul	18

Patentansprüche

1. Verfahren zur Planung von Versuchen zur Erreichung eines Optimierungszieles mit folgenden Schritten:

- A) Auswahl von zumindest einem ersten Versuch aus einem Versuchsraum durch einen datengetriebenen Optimierer in einer Rechereinheit,
- B) Eingabe von experimentell ermittelten Versuchsdaten des ersten Versuchs, in mindestens eine Meta-Schicht in eine Rechereinheit,
- C) Verwendung mindestens einer Meta-Schicht für die Auswertung der Versuchsdaten,
- D) Eingabe der experimentell ermittelten Versuchsdaten des ersten Versuchs in den datengetriebenen Optimierer,
- E) Beeinflussung des datengetriebenen Optimierers durch das Ergebnis der Auswertung in der Meta-Schicht und Prüfung der Zielerreichung,
- F) Auswahl von zumindest einem zweiten Versuch aus dem Versuchsraum durch den datengetriebenen Optimierer,
- G) Wiederholung der Schritte B) bis E) für die Daten des zweiten Versuchs

und

- 20 -

- H) Abbruch der Hexation bei Zielerreichung oder Wiederholung der Schritte A) bis F) für mindestens einen dritten oder weitere Versuche bis zur Zielerreichung.
2. Verfahren nach Anspruch 1, wobei der Versuchsraum vor der Auswahl von dem zumindest einem zweiten Versuch durch den Optimierer und/oder durch die Meta-Schicht verändert wird, insbesondere beschränkt, verschoben oder vergrößert wird.
 3. Verfahren nach Anspruch 1 oder 2, wobei die Meta-Schicht ein neuronales Netz und/oder ein Hybridmodell und/oder ein rigoroses Modell und/oder eine oder mehrere Data Mining Methoden beinhaltet.
 4. Verfahren nach Anspruch 1, 2 oder 3, wobei zur Ermittlung der experimentellen Versuchsdaten Versuche aus dem Bereich der Wirkstoffforschung, Materialforschung, Katalysatorforschung, Biotechnologie und/oder Optimierung von Reaktionsbedingungen durchgeführt werden.
 5. Verfahren nach einem der vorhergehenden Ansprüche 1 bis 4, wobei die Beeinflussung des datengetriebenen Optimierers durch Filterung der Versuchsdaten basierend auf der Auswertung erfolgt.
 6. Verfahren nach Anspruch 5, wobei die Filterung durch eine Neubewertung der Versuchsdaten erfolgt.
 7. Verfahren nach Anspruch 5, wobei die Filterung durch Gewichtung und/oder Vorselektion der Versuchsdaten erfolgt.
 8. Verfahren nach Anspruch 7, wobei die Gewichtung durch einen Gewichtungsparemeter oder durch ein- oder mehrfache Duplizierung der Versuchsdaten erfolgt.

9. Verfahren nach einem der vorhergehenden Ansprüche 1 bis 8, wobei der Optimierer mindestens ein Kernmodul und ein Modul zur Selektion neuer Versuchspunkte aufweist.
10. Verfahren nach Anspruch 9, wobei die Beeinflussung des datengetriebenen Optimierers durch Beeinflussung des Moduls zur Selektion neuer Versuchspunkte erfolgt.
11. Verfahren nach Anspruch 10, wobei die Beeinflussung des Moduls zur Selektion neuer Versuchspunkte bei Überschreiten eines Schwellwerts und/oder durch Benutzervorgaben erfolgt.
12. Verfahren nach Anspruch 9, wobei die Beeinflussung des datengetriebenen Optimierers durch Beeinflussung des Kernmoduls erfolgt.
13. Verfahren nach Anspruch 12, wobei die Beeinflussung des Kernmoduls bei Überschreiten eines Schwellwerts und/oder durch Benutzervorgaben erfolgt.
14. Computerprogrammprodukt zur Durchführung eines Verfahrens nach einem der vorhergehenden Ansprüche 1 bis 13.
15. Computersystem zur Planung von Versuchen mit:
 - A. einem datengetriebenen Optimierer (6) zur Auswahl von zumindest einem ersten Versuch aus einem Versuchsraum (1),
 - B. einer Meta-Schicht (9) für den datengetriebenen Optimierer für die Auswertung von für den ersten Versuch experimentell ermittelten Versuchsdaten, wobei die Meta-Schicht ein neuronales Netz und/oder ein Hybridmodell und/oder ein rigoroses Modell und/oder Data

Mining Methoden beinhaltet, und wobei die Meta-Schicht für die Beeinflussung des datengetriebenen Optimierers basierend auf dem Ergebnis der Auswertung ausgebildet ist.

16. Computersystem nach Anspruch 15, mit Mitteln zur Filterung (13; 15) der Versuchsdaten basierend auf der Auswertung.
17. Computersystem nach Anspruch 16, wobei die Mittel zur Filterung zur Neubewertung der Versuchsdaten ausgebildet sind.
18. Computersystem nach Anspruch 16, wobei die Mittel zur Filterung zur Gewichtung und/oder Vorselektion der Versuchsdaten ausgebildet sind.
19. Computersystem nach einem der vorhergehenden Ansprüche 15 bis 18, wobei der Optimierer zumindest ein Kernmodul (16) und ein Modul (17) zur Selektion neuer Versuchspunkte aufweist.
20. Computersystem nach Anspruch 19, wobei die Meta-Schicht zur Beeinflussung des Moduls zur Selektion neuer Versuchspunkte ausgebildet ist.
21. Computersystem nach Anspruch 19, wobei die Meta-Schicht zur Beeinflussung der Kernmodule ausgebildet ist.

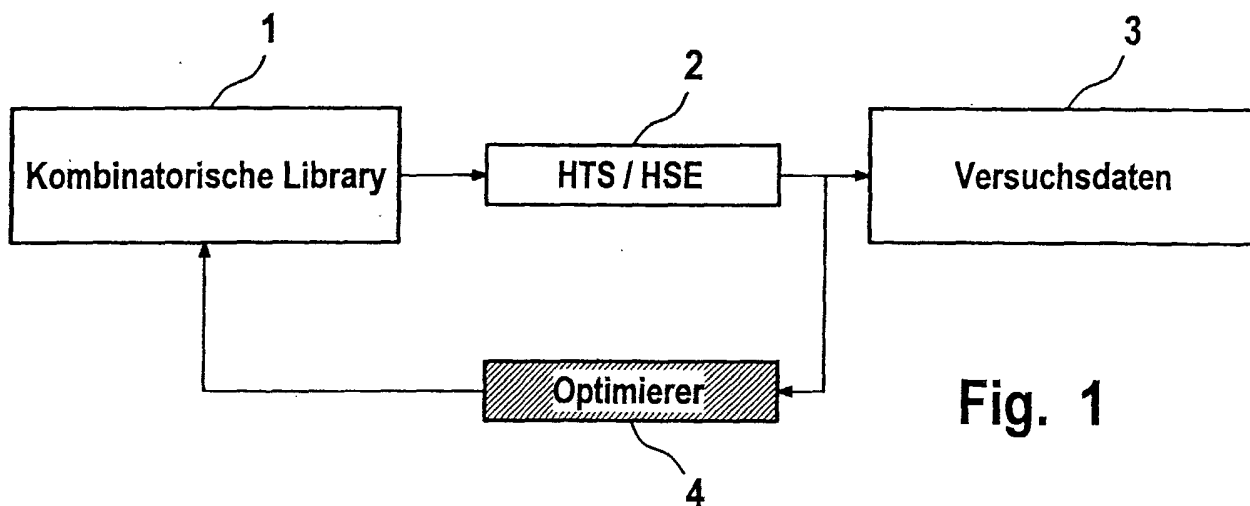


Fig. 1

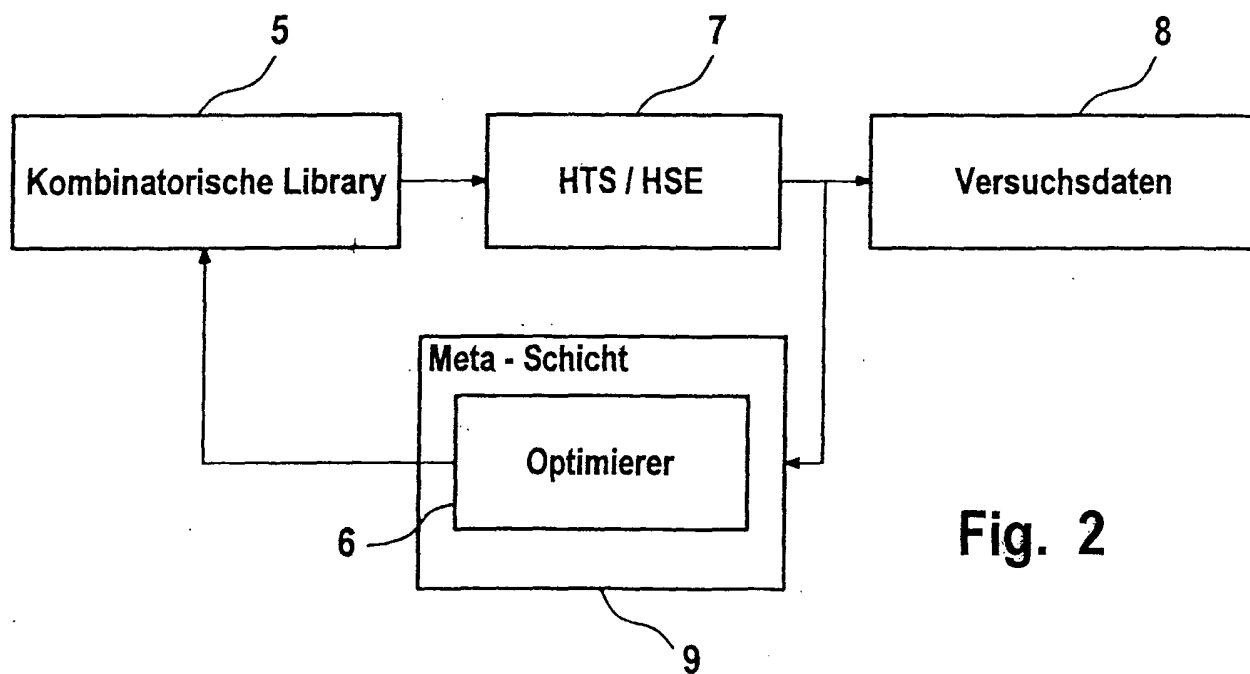


Fig. 2

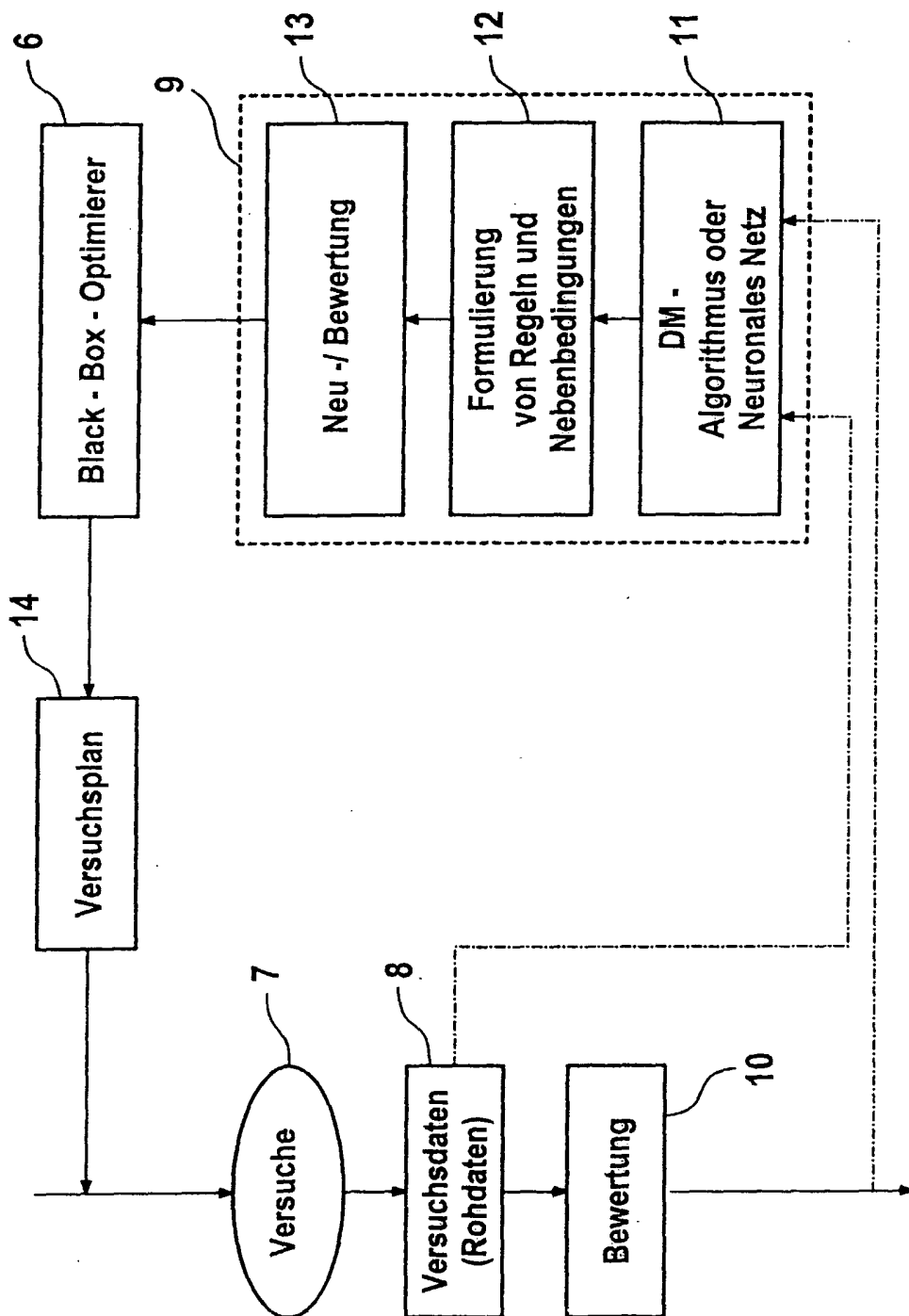


Fig. 3

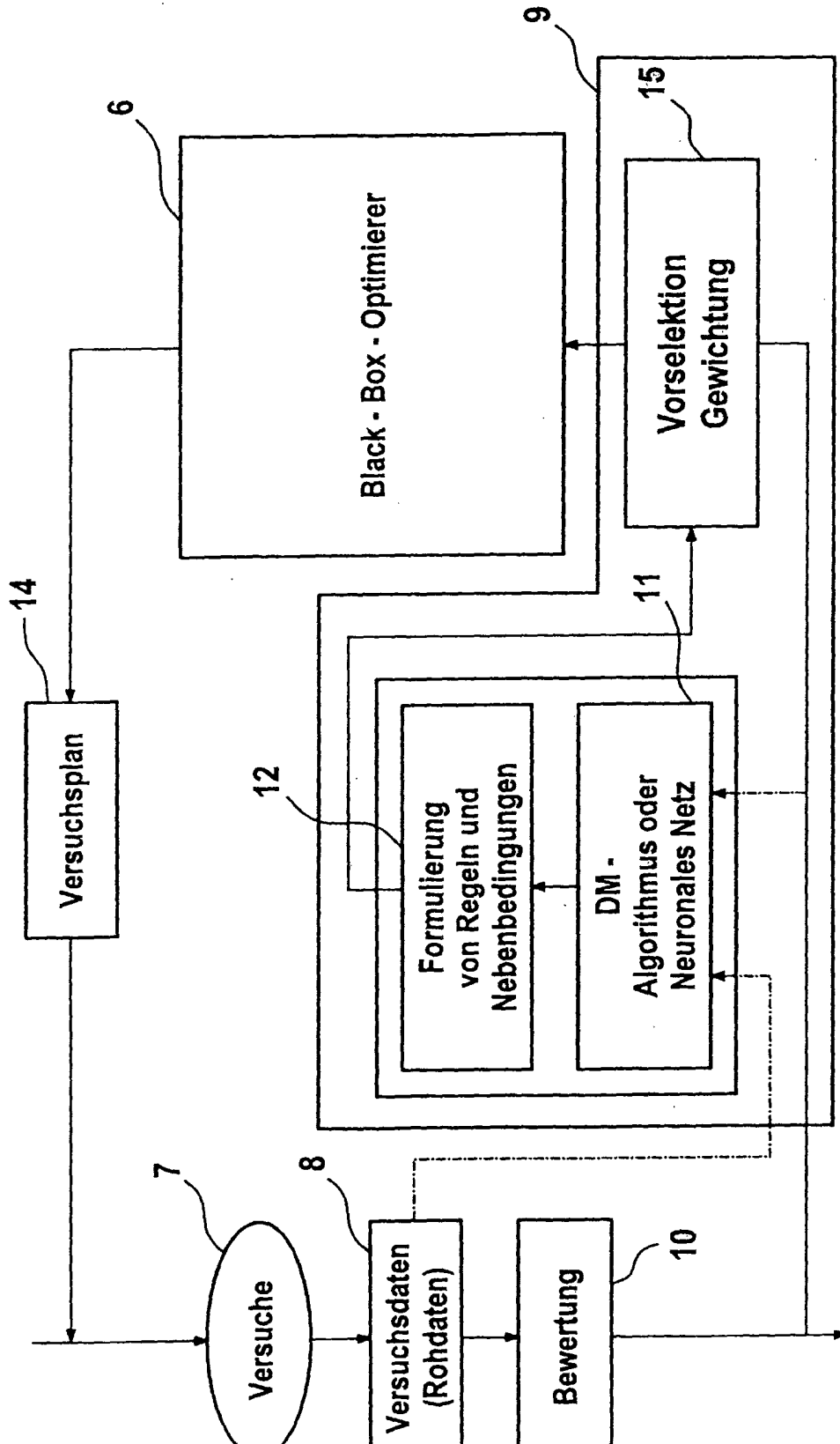


Fig. 4

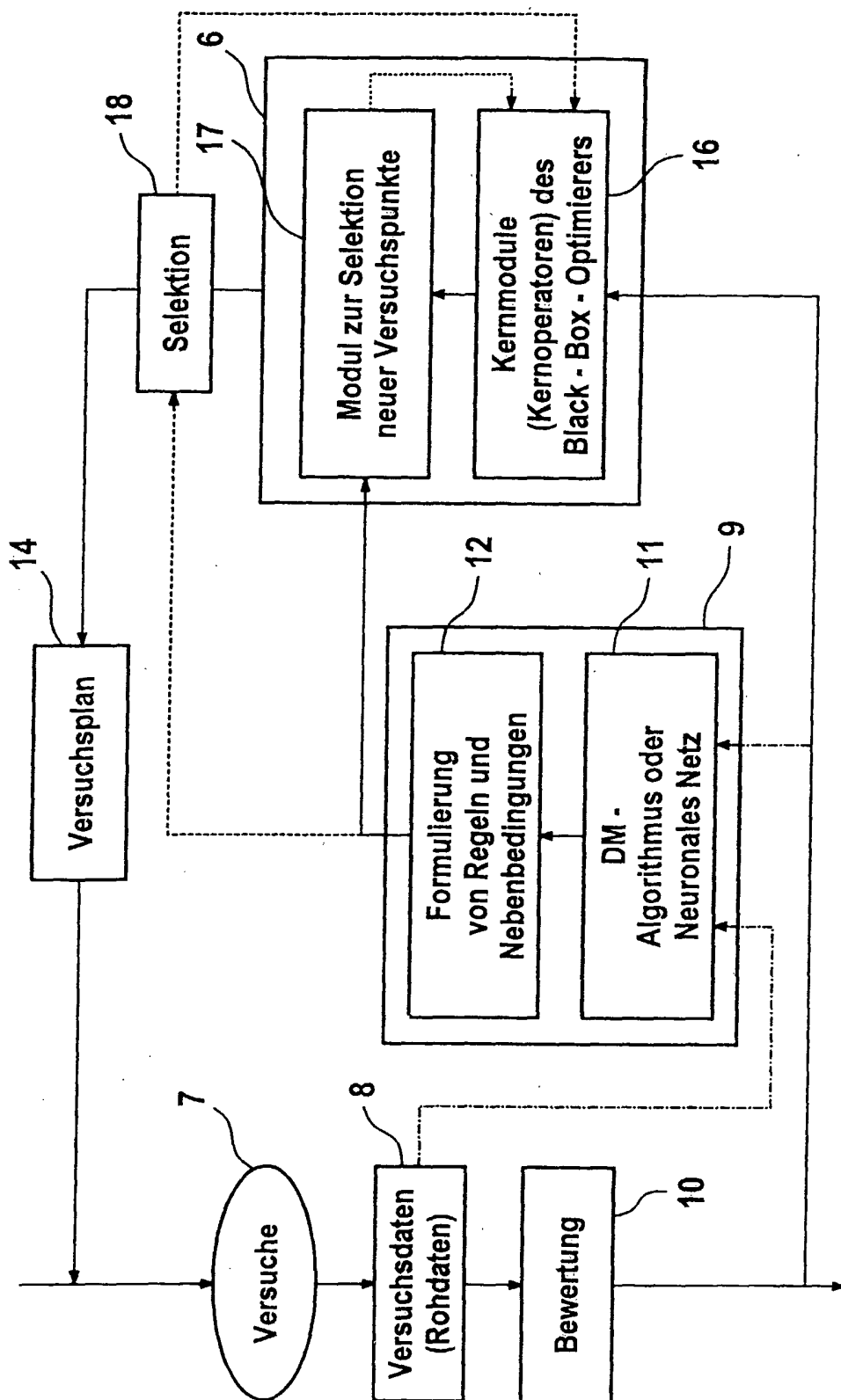


Fig. 5

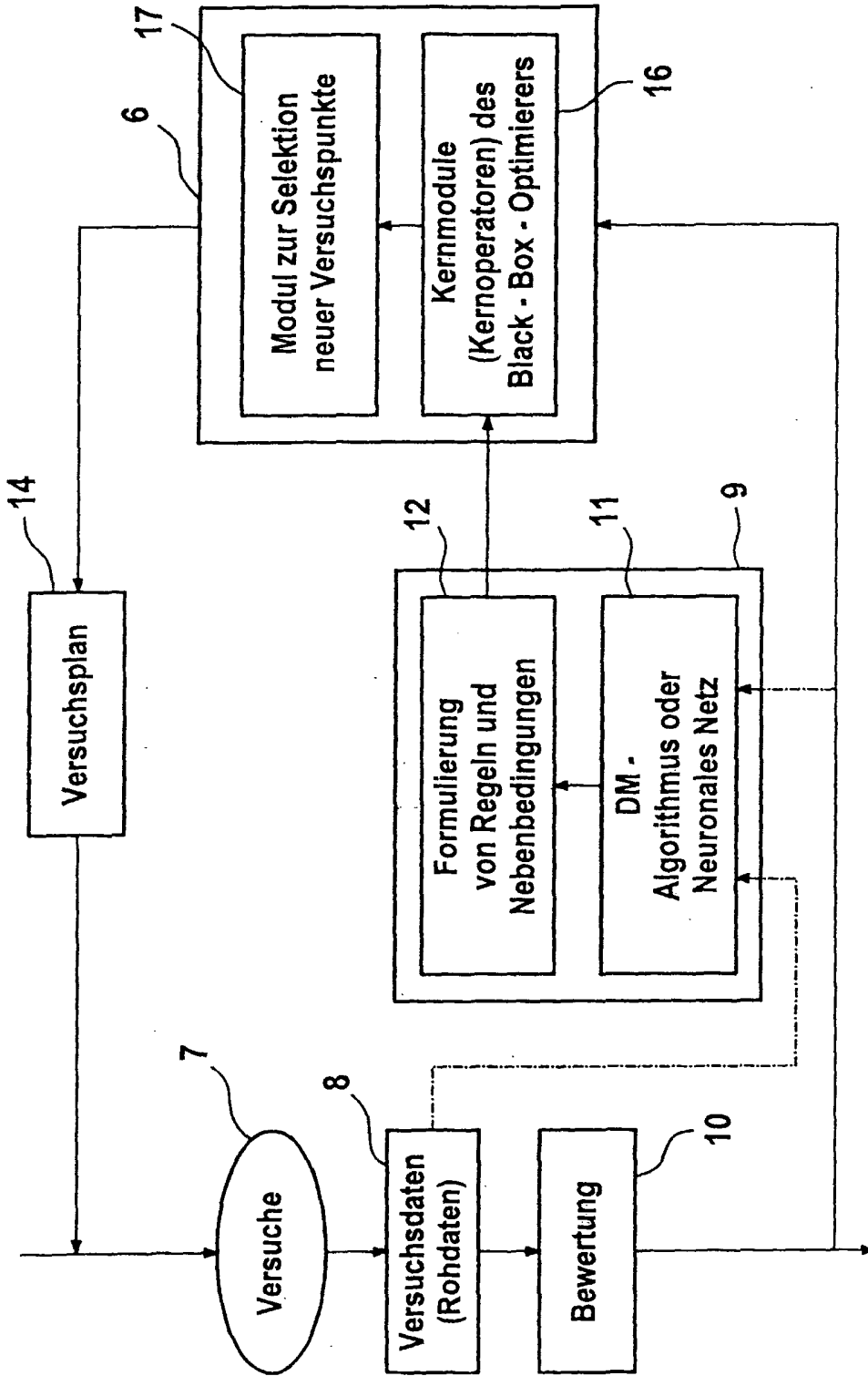


Fig. 6