Fraunhofer IAIS

Mit Künstlicher Intelligenz und Expertenwissen zu effizienteren Versuchsplänen

KI-gestütztes Design of Experiments in Forschung und Entwicklung

Fraunhofer
IAIS

Fraunhofer-Institut für Intelligente
Analyse- und Informationssysteme IAIS

Mit Künstlicher Intelligenz und Expertenwissen zu effizienteren Versuchsplänen

KI-gestütztes Design of Experiments in Forschung und Entwicklung

Autor*innen

Dorina Weichert
Dr. Gunar Ernis

Herausgeber

Das Fraunhofer IAIS

Als Teil der größten Organisation für anwendungsorientierte Forschung in Europa ist das Fraunhofer-Institut für Intelligente Analyse- und Informationssysteme IAIS mit Sitz in Sankt Augustin bei Bonn eines der führenden Wissenschaftsinstitute auf den Gebieten Künstliche Intelligenz, Maschinelles Lernen und Big Data in Deutschland und Europa. Mit seinen rund 300 Mitarbeitenden unterstützt das Institut Unternehmen bei der Optimierung von Produkten, Dienstleistungen, Prozessen und Strukturen sowie bei der Entwicklung neuer digitaler Geschäftsmodelle. Damit gestaltet das Fraunhofer IAIS die digitale Transformation unserer Arbeits- und Lebenswelt.

www.iais.fraunhofer.de

Lamarr-Institut für Maschinelles Lernen und Künstliche Intelligenz

Das Lamarr-Institut gestaltet eine neue Generation der Künstlichen Intelligenz (KI), die leistungsstark, nachhaltig, vertrauenswürdig und sicher zur Lösung fundamentaler Herausforderungen in Wirtschaft und Gesellschaft beiträgt. Als eines der zentralen KI-Kompetenzzentren Deutschlands steht das Lamarr-Institut für wertebasierte, international wettbewerbsfähige und anwendungsorientierte Spitzenforschung und engagiert sich auf regionaler, nationaler sowie europäischer Ebene in Wissenschaft, Bildung und Technologietransfer.

Getragen wird das Forschungsinstitut von der Technischen Universität Dortmund, der Rheinischen Friedrich-Wilhelms-Universität Bonn sowie den Fraunhofer-Instituten für Intelligente Analyse- und Informationssysteme IAIS in Sankt Augustin und für Materialfluss und Logistik IML in Dortmund. Das Lamarr-Institut wird im Rahmen der KI-Strategie der Bunderegierung durch das Bundesministerium für Bildung und Forschung (BMBF) sowie das Land Nordrhein-Westfalen dauerhaft gefördert.

www.lamarr-institute.org

Inhalt

1. Executive Summary

»Design of Experiments« (DoE), auf Deutsch auch als »statistische Versuchsplanung« bezeichnet, ist ein insbesondere in den Ingenieurswissenschaften etablierter und integraler Bestandteil der Produktentwicklung. Das Ziel besteht in der Regel in der Anwendung unterschiedlicher statistischer Verfahren, um mit möglichst wenigen Experimenten – seien es aufwendige Computersimulationen oder physische Versuche – die optimalen Parametereinstellungen für ein Produkt zu ermitteln: zum Beispiel das zukünftige Design von Fahrzeugen, die Zusammensetzung eines neuen Parfums oder die Suche nach leichter und schneller herstellbaren Polymerverbindungen.

Klassische Verfahren zur Erstellung von Versuchsplänen können in der Regel kein vorhandenes Wissen über die Produkte in Form von bereits durchgeführten Experimenten berücksichtigen. Vielmehr müssen Abhängigkeiten zwischen Eingangs- und Zielgrößen händisch implementiert werden und sind häufig auf einfache (wie z. B. polynomielle) Zusammenhänge beschränkt. Außerdem werden Versuchspläne mit einer vorgegebenen Anzahl an Versuchen erzeugt, sogenannte fixed-size DoE. Diese Versuche sind unabhängig voneinander: Es ist egal, in welcher Reihenfolge sie durchgeführt werden. Dadurch ist es auch innerhalb einer Versuchsreihe nicht möglich, Informationen, die in frühen Experimenten der Reihe gewonnen wurden, in späteren Experimenten der gleichen Reihe zu berücksichtigen.

Während der Arbeit in zahlreichen Projekten in Industrie und Forschung haben wir am Fraunhofer IAIS festgestellt, dass sich die oben skizzierten Nachteile der klassischen Verfahren durch den Einsatz von Methoden der Künstlichen Intelligenz (KI) und des Maschinellen Lernens (ML) adressieren lassen: Sogenannte informierte probabilistische ML-Modelle können mit bereits vorhandenen Daten und dem Expertenwissen der Ingenieur*innen trainiert werden. Sie sind zudem in der Lage, auch komplexe Zusammenhänge zwischen Eingangs- und Zielgrößen automatisch zu erkennen.

Durch die gezielte Kombination von ML-Modellen und Expertenwissen lässt sich die Anzahl der benötigten Versuche um bis zu 50 Prozent reduzieren. Dadurch erreichen Sie eine bessere Ausnutzung Ihrer Ressourcen und führen nur Versuche durch, die auch zielführend für die Optimierung Ihrer Produkte sind. Insgesamt profitieren Sie mithilfe unseres Ansatzes von kürzeren Entwicklungszyklen und einer besseren Nutzung Ihrer Zeit und Ihres Personals.

ML-Modelle können genutzt werden, um einen oder mehrere Versuche mit möglichst optimalen Parametereinstellungen vorzuschlagen. Anschließend wird das ML-Modell mit den hinzugewonnenen Versuchsdaten neu trainiert und der Zyklus aus *»Erstellung eines Versuchsplans – Durchführung der Experimente – Neues Training des ML-Modells«* wird so lange fortgesetzt, bis die optimalen Einstellungen für die Produkte gefunden wurden.

Neuartiger Ansatz: Prozessmodell mit drei Phasen und informierte probabilistische Machine-Learning-Modelle

In diesem Whitepaper beschreiben wir einen neuartigen Ansatz für DoE mit dem Ziel, bei gleichbleibender Qualität der Ergebnisse die Anzahl der benötigten Versuche zu reduzieren. Dabei setzen wir auf ein Prozessmodell mit drei Phasen, das sich in der Praxis bewährt hat.

Als Grundlage für die Modellierung mithilfe von ML verwenden wir sogenannte informierte probabilistische Modelle, da diese Klasse von Modellen besonders gut dazu geeignet ist, Wissen sowohl aus Daten als auch aus vorhandenem Expertenwissen abzuleiten. Expertenwissen kann in sehr verschiedenen Formen auftreten: beispielsweise in Annahmen über die Verteilung von Einflussfaktoren und Zielgrößen, in Monotonie-Annahmen oder in komplexeren Hypothesen wie polynomiellen Zusammenhängen, Symmetrien oder der Periodizität. Die Integration dieses Wissens in die ML-Modelle und damit in die erzeugten Versuchspläne kann dafür sorgen, dass die benötigte Datenmenge für eine Vorhersage der Zielgröße erheblich reduziert werden kann. Außerdem sind sich probabilistische Modelle ihrer eigenen Unsicherheit bewusst: Zu jeder Vorhersage der Zielgröße erhält man eine Abschätzung der eigenen Unsicherheit in Form von Vertrauens- oder Konfidenzintervallen, die wiederum verwendet werden können, um Vorschläge für neue Experimente zu bewerten.

2. Einleitung

Die Sicherung unternehmerischer Existenzen hängt heutzutage maßgeblich davon ab, wie agil Unternehmen sich auf ändernde Kundenanforderungen einstellen. Unternehmen, die schneller auf neue Kundenanforderungen und Umstände reagieren, haben klare Vorteile auf dem Markt. Forschungs- und Entwicklungsabteilungen (F&E), insbesondere in Großunternehmen, sind Garanten für die Sicherung und im besten Fall den kontinuierlichen Ausbau von Marktanteilen. Effiziente F&E-Abteilungen sind das Rückgrat für Produktinnovationen und bilden so die Grundlage für eine langfristig angelegte Strategie zur Existenzsicherung der Unternehmen. Dies zeigt auch ein Blick auf die jährlichen Ausgaben für F&E-Vorhaben: In den letzten zehn Jahren sind die zukunftsgerichteten Investitionen in Deutschland kontinuierlich von 94,2 Mrd. US-Dollar (2,7 Prozent des BIP) auf 132,5 Mrd. US-Dollar (3,2 Prozent des BIP) gestiegen [1, 2].

Ein wesentliches Ziel von F&E-Abteilungen in Deutschland ist die möglichst effiziente Neu- und Weiterentwicklung von Produkten, die einen hohen Beitrag zur Wertschöpfung leisten. Experimentelle Validierung und Optimierung des Produktentwurfs sind insbesondere in frühen Phasen der Planung und Entwicklung wichtig. Es gilt zu vermeiden, dass eine suboptimale Versuchsplanung und -durchführung später zu fehlerhaften Produkten führt. Die frühe Erkennung von Fehlern ist auch deshalb so wichtig, da die Fehlerkosten exponentiell mit den weiteren Stufen der Wertschöpfung wachsen (siehe Infobox: Rule of Ten).

In diesen Experimenten sollten möglichst alle Faktoren (Eingangsparameter) berücksichtigt werden, die einen Einfluss auf die Produkte haben. Eine Vielzahl von Faktoren mit vielen Einstellungsstufen bedingt aber ein exponentielles Wachstum der Anzahl der benötigten Experimente zur Bestimmung der Produkteigenschaften, was als »Curse of Dimensionality« bezeichnet wird [3]. Ein typischer Ansatz, den Raum möglichst vollständig abzutasten und damit alle möglichen Variationen der zu untersuchenden Produkte bewerten zu können, ist das sogenannte »Full Factorial Design«: Hier werden alle Kombinationen der Faktoren für Experimente vorgeschlagen, also jede Einstellung jedes Faktors mit jeder Einstellung jedes anderen Faktors. Aufgrund der vielen kombinatorischen Möglichkeiten sind solche Designs auf wenige Faktoren mit wenigen Einstellungsstufen beschränkt, da sie sehr teuer und zeitaufwendig in der Anwendung sein können. Hier ergibt sich der zentrale Zielkonflikt bei F&E: Die Entwicklungskosten sollen minimiert werden, während die Entwicklungsgeschwindigkeit maximiert werden soll, um die Sicherung der eigenen Marktanteile zu gewährleisten.

Um diesem Zielkonflikt zu begegnen, wird »Design of Experiments« (DoE) verwendet, das auf Deutsch als »statistische

Anwendungsbeispiel: Bestimmung der Dauerfestigkeit von Stählen

Die Abschätzung der Dauerfestigkeit ist ein kostspieliger und manueller Materialcharakterisierungsprozess, bei dem der Stand der Technik ein standardisiertes Experiment- und Analyseverfahren vorsieht [4]. Die Dauerfestigkeit ist eine Materialeigenschaft, die die maximale Belastung beschreibt, die zyklisch auf eine bestimmte Probe ausgeübt werden kann, ohne dass sich ein Defekt/Bruch einstellt. Sie erfordert ein komplexes experimentelles Verfahren, da die Dauerfestigkeit für einen bestimmten Werkstoff nicht direkt beobachtbar ist.

Für eine Probe aus einem bestimmten Material bedeutet dies, dass nur beobachtet werden kann, ob sie bei der aufgebrachten Last bricht, was darauf hindeutet, dass die Last größer als die Dauerfestigkeit der Probe war (ein sogenanntes Versagen), oder ob sie den Vorgang überlebt, was darauf hindeutet, dass die Last kleiner als die Dauerfestigkeit der Probe war (ein sogenannter Durchläufer). Hinzu kommt, dass die Versuche verrauscht sind, d. h., dass Proben aus identischen Materialien bei identischen Belastungen manchmal versagen und manchmal überleben. Daher zielt die Dauerfestigkeitsprüfung darauf ab, eine gültige Statistik der Ausreißer und Ausfälle zu erstellen, um die Parameter der Dauerfestigkeitsverteilung zu schätzen.

Der derzeitige Stand der Technik für Dauerfestigkeitsprüfungen ist genormt. Die Norm legt mehrere Analysemethoden und ein iteratives experimentelles Verfahren, die sogenannte Treppenstufenmethode, fest. Bei der Treppenstufenmethode werden die während des Versuchs zu verwendenden Laststufen vor Versuchsdurchführung festgelegt. Versagt nun eine Probe auf einer Laststufe, wird die Last um eine Stufe verringert; ist sie ein Durchläufer, wird die Last um eine Stufe erhöht. Besonders kritisch für den Anwender ist die vorige Festlegung der Laststufen sowie der initialen Last – sie beeinflussen nicht nur die erzielte Genauigkeit der Schätzung der Dauerfestigkeit, sondern auch die generelle Auswertbarkeit der gemachten Versuche.

Bei Verwendung unserer KI-gestützten DoE-Methode gehen als Parameter zum einen die chemische Zusammensetzung der Stähle und zum anderen die Ergebnisse aus den Lastversuchen (aufgebrachte Last, Durchläufer ja/nein) ein. Hierdurch können wir für jeden Versuch eine Last empfehlen, die nicht nur in jedem Fall auswertbar ist, sondern auch die Genauigkeit der Schätzung der Dauerfestigkeit so schnell wie möglich erhöht. Mit diesem Ansatz lassen sich im Vergleich zum Treppenstufenverfahren bis zu 30 Prozent der benötigten Versuche einsparen.

Versuchsplanung« bezeichnet wird. DoE bezeichnet eine Sammlung unterschiedlichster (statistischer) Verfahren zur Planung, Durchführung, Analyse und Interpretation von kontrollierten Versuchen, um die Faktoren zu bewerten, die das Ergebnis der Experimente beeinflussen. Konkrete Aufgaben von DoE-Verfahren sind beispielsweise die Bestimmung des minimal erforderlichen Versuchsumfangs zur Einhaltung von Genauigkeitsvorgaben, die optimale Auswahl von Versuchen oder Methoden zum Umgang mit Störgrößen.

Der grundsätzliche Ansatz, statistische Verfahren zur Versuchsplanung einzusetzen, reicht bis zum Anfang des 20. Jahrhunderts zurück. Zum Beispiel beruhen Verfahren wie Randomisierung, Block-Randomisierung oder Wiederholungen bei der Versuchsplanung sowie die Varianzanalyse bei der Versuchsauswertung auf den Arbeiten von R. A. Fisher [5].

Heutige Ansätze profitieren sehr stark von rasanten Fortschritten bei der Entwicklung von Verfahren des Maschinellen Lernens (ML) oder der Künstlichen Intelligenz (KI). Aus geschäftlicher Sicht sind die Anforderungen an die statistische Versuchsplanung klar: Mit DoE geplante Versuche sollen statistisch valide den Einfluss verschiedener Faktoren auf vorgegebene Zielgrößen bestimmen und dabei möglichst wenig kosten.

Rule of Ten

Die 10er-Regel (Rule of Ten) besagt, dass sich die Kosten eines unentdeckten Fehlers um den Faktor 10 für jede weitere Stufe der Wertschöpfungskette erhöhen. Das bedeutet, je früher Fehler erkannt und vermieden werden, desto günstiger ist dies für die Organisation [6, 7, 8]. Insbesondere externe Kosten, wie etwa Nacharbeit, Garantieleistungen, Produkthaftungskosten oder Rücktritt vom Kaufvertrag sind weitaus umfangreicher als interne Kosten wie etwa Fehlerverhütungskosten oder Prüfkosten.

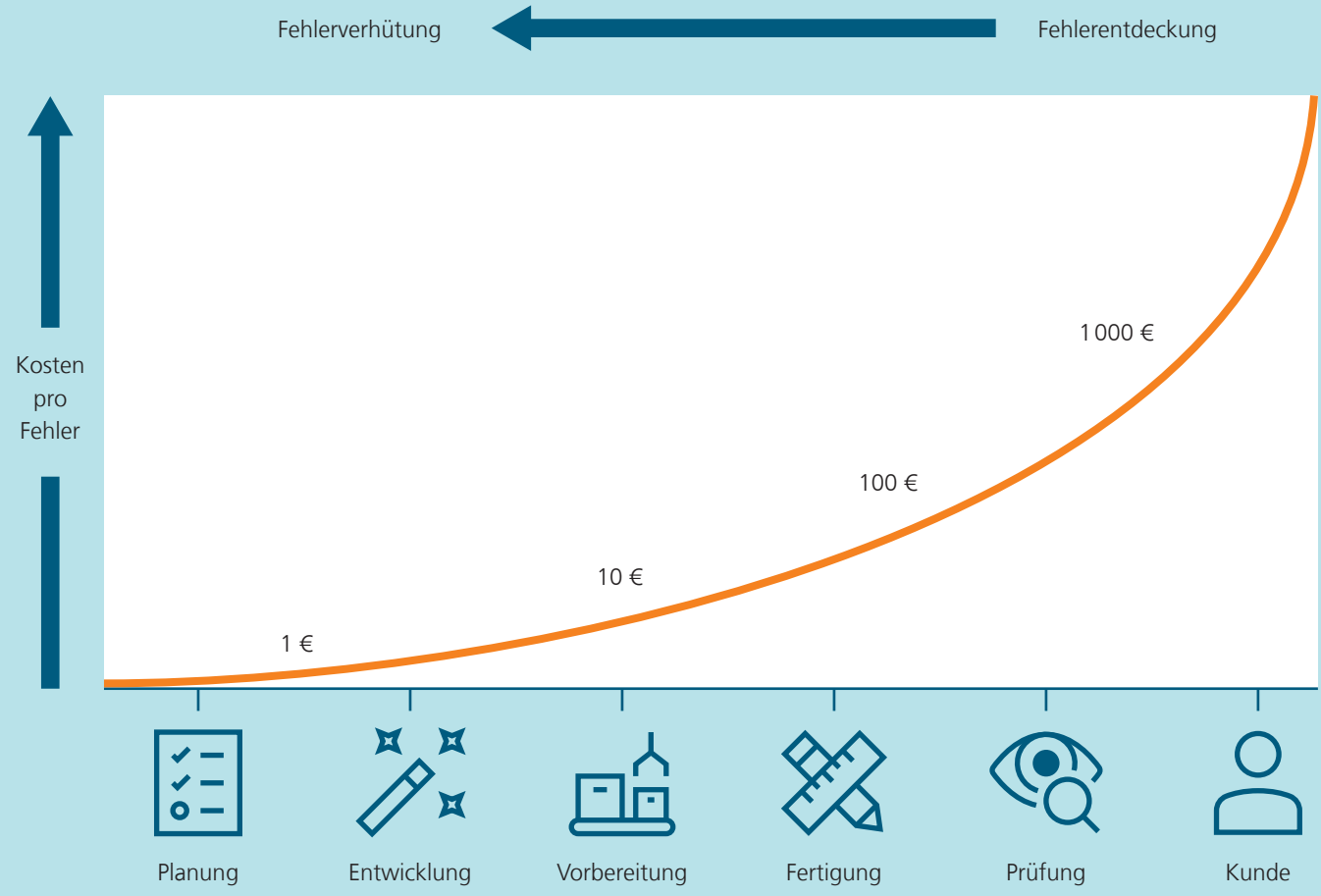


Abbildung 1: Rule of Ten. Grafik nach [8].

Wird dies durch eine Digitalisierung des Produktenwicklungsprozesses unterstützt, kann der Entwicklungszyklus für neue Produkte dramatisch verkürzt werden [9].

Der gezielte Einsatz von KI-Methoden in der Produktentwicklung und insbesondere für die Versuchsplanung wird in Zukunft immer bedeutsamer werden:

- In der **Automobilindustrie** wird KI für die Entwicklung neuer Modelle genutzt. Das Ziel ist es, die Markteinführungszeit für die Modellreihen der nächsten Generation zu verkürzen und Kompatibilität mit den neuesten Vorschriften und Compliance-Anforderungen der Branche sicherzustellen. Außerdem wird die KI auch eingesetzt, um die Lebensdauer bestehender Modelle zu verlängern [10, 11].
- In der **Kosmetikindustrie** wird KI verwendet, um neue Kombinationen von Parfum-Formulierungen zu erzeugen, die spezifischen Designzielen entsprechen. Das KI-System soll außerdem zukünftig in der Ausbildung von Parfümeuren eingesetzt werden [12].
- In den **Materialwissenschaften** werden KI-Modelle verwendet, um neue Kandidaten für neue Polymerverbindungen schneller und flexibler herstellen zu können [13].

Mit unserem Ansatz setzen wir auf ein dreiphasiges Prozessmodell, in dem wir einerseits vorhandenes Expertenwissen und andererseits die Möglichkeit, ML-Modelle aus Daten lernen zu lassen, kombinieren.

Phase 1: Definition der Zielgrößen und möglicher Einflussfaktoren

Phase 2: Mithilfe von Versuchsplänen fester Größe (sogenannten »fixed-size DoEs«) wird ermittelt, ob die unterschiedlichen Faktoren Einfluss auf die Zielgröße haben und wie groß die Einflüsse gegebenenfalls sind

Phase 3: Basierend auf ML-Methoden werden sequenziell Versuchspläne erstellt, die auf den jeweils aktuellen Messergebnissen beruhen. Ziele hierbei können die Verfeinerung des ML-Modells oder das Finden eines globalen Optimums der Zielgrößen sein

3. Design of Experiments: Grundlegendes

Dieses Kapitel beleuchtet relevante Hintergründe moderner DoE-Methoden. So wird erklärt, weshalb eine Verwendung von Simulationen physische Experimente nicht ersetzen kann, wie mit den verschiedenen Arten der auftretenden Unsicherheit umgegangen wird, welche statistische Methodik zugrunde liegt und welchen weiteren Herausforderungen begegnet werden muss.

3.1. Die Relevanz physischer Experimente

Unsere zentrale Frage ist, wie man die Arbeit im Bereich der F&E effizienter gestalten kann. Ein möglicher Weg ist die Verwendung von Simulationen der Produkte mit dem Ziel, diese möglichst frühzeitig während der Entwicklung optimal zu konzipieren und auszulegen. Dies ist für komplexe Entwicklungsaufgaben aber eine Herausforderung, denn trotz der kontinuierlich wachsenden Leistungsfähigkeit von Computern ist die Entwicklung und die Anwendung von Simulationen sehr aufwendig und teuer.

Mit steigenden Anforderungen an die Genauigkeit der Simulation steigen auch die Anforderungen an die zugrunde liegenden Modelle, was die Kosten bei der Entwicklung erhöht. Dazu kommen noch Rechenzeit und -kosten bei der Anwendung der Simulationen. Ein anderer möglicher Weg, F&E effizienter zu gestalten, ist die Verwendung von gezieltem DoE. Hier besteht der Vorteil, dass sich DoE-Verfahren auf Simulationen und auf physische Versuche anwenden lassen und damit während der Entwicklung in beiden Bereichen Zeit und Kosten (beispielsweise in Form von Rechenleistung oder Material) eingespart werden können.

3.2. Der Umgang mit Unsicherheiten

Sowohl bei Simulationen, aber auch bei physisch durchgeführten Experimenten, stellt sich die Frage nach dem Umgang mit inhärenten Unsicherheiten. In der praktischen Anwendung gibt es vielfältige Ursachen für Unsicherheiten, die sich im Wesentlichen in zwei Kategorien einordnen lassen: die aleatorische (statistische) und die epistemische (systematische) Unsicherheit. Die aleatorische Unsicherheit ergibt sich aus zufälligen Schwankungen der physikalischen Eigenschaften der Produkte wie Streuungen in Materialparametern oder fertigungsbedingter Varianz. Die epistemische Unsicherheit spiegelt die verschiedenen Formen unzureichenden Wissens bei der mathematischen Modellierung wie z. B. die unvollständige oder nur beschränkt gültige Information bei der Herleitung der Modelle wider.

Aleatorische und epistemische Unsicherheit

Aleatorische Unsicherheit: Diese Art von Unsicherheit ist auch als statistische Unsicherheit bekannt. Sie beschreibt den Umstand, dass gewisse Prozesse statistischer Natur und dementsprechend nicht vollständig reproduzierbar sind. Wird zum Beispiel eine Messgröße durch voneinander unabhängigen Messungen erfasst, stellen die Resultate der einzelnen Messungen Werte einer Zufallsvariable mit einer zugehörigen Verteilung dar. Die Schätzung für die Messgröße wird typischerweise durch den empirischen Mittelwert und die empirische Standardabweichung angegeben.

Epistemische Unsicherheit: Diese Art von Unsicherheit ist auch als systematische Unsicherheit bekannt. Sie ist darauf zurückzuführen, dass es bei der Bewertung von Experimenten Dinge gibt, die man im Prinzip wissen könnte, in der Praxis aber nicht weiß. Dies ist häufig darauf zurückzuführen, dass die zur Bewertung verwendeten Modelle bestimmte Effekte vernachlässigen oder dass bestimmte Daten bei der Bewertung absichtlich nicht berücksichtig wurden.

Die Bedeutung dieser Begriffe kann einfach anhand eines Würfels verständlich gemacht werden. Wenn 4 Würfe eines Würfels die Zahlen 2, 4, 3, und 5 ergeben haben, verbleibt die Frage, wie der Würfel aussieht. Auf der einen Seite kann angenommen werden, dass der Würfel (wie die meisten Würfel) 6 Seiten hat, jedoch kann man sich dessen nicht sicher sein. Diese epistemische Unsicherheit (= die Unsicherheit, wie der Würfel aussieht) kann durch weitere Versuche reduziert werden, sodass die Anzahl der Seiten und ihre Wurfwahrscheinlichkeiten bestimmt werden. Ist die epistemische Unsicherheit minimiert, verbleibt lediglich die aleatorische Unsicherheit: die Unsicherheit, welchen Wert der Würfel bei einem Wurf annehmen wird.

Mathematisch werden Unsicherheiten typischerweise in Form von Wahrscheinlichkeitsverteilungen charakterisiert. Aus dieser Perspektive bedeutet epistemische Unsicherheit, dass nicht sicher ist, wie die relevante Wahrscheinlichkeitsverteilung aussieht, und aleatorische Unsicherheit bedeutet, dass nicht sicher ist, wie eine aus einer Wahrscheinlichkeitsverteilung gezogene Zufallsstichprobe aussehen wird.

Mithilfe der Unsicherheiten lassen sich sogenannte Vertrauens- oder Konfidenzintervalle der mit Unsicherheiten behafteten Parameter konstruieren. Diese Konfidenzintervalle geben Bereiche an, die die Parameter mit einer bestimmten Wahrscheinlichkeit einschließen. Typische Werte für diese Überdeckungswahrscheinlichkeiten sind 95 Prozent oder 99 Prozent und Ereignisse heißen signifikant (mit einem Signifikanzniveau von 5 Prozent oder 1 Prozent), wenn deren Parameterwerte außerhalb dieser so konstruierten Intervalle liegen.

Konkret auf die Arten von Unsicherheit bezogen, bedeutet die Anwendung von DoE-Verfahren, eine ausreichende Anzahl von Experimenten vorzuschlagen, um die epistemische Unsicherheit so stark zu verringern, dass die aleatorische Unsicherheit bestimmt werden kann, da diese mithilfe von Experimenten gut quantifiziert werden kann. Beide Unsicherheitskategorien sollten immer mit einbezogen werden, um final etwa die Frage beantworten zu können, ob sich das Ergebnis zufällig verändert hat oder weil ein Faktor variiert wurde.

3.3. Die statistische Methodik des DoE-Prozesses

Die Erfahrung aus zahlreichen Projekten hat uns gezeigt, dass DoE nicht als Anwendung statistischer Verfahren, sondern als methodische Herangehensweise verstanden werden muss, um nachhaltig zu spürbaren wirtschaftlichen Erfolgen zu führen – wie etwa der Reduktion der Zeit bis zur Markteinführung (Time-to-Market).

Am Fraunhofer IAIS und in unseren Kundenprojekten betrachten wir den DoE-Prozess als einen iterativen Prozess (siehe Abbildung 2):

- Aufstellen einer Arbeitshypothese
- Design des Versuchsplans
- Durchführen der Experimente
- Analyse der Ergebnisse

Je nach Aufgabenstellung sind mehrere Iterationen nötig, um mithilfe von DoE ein Produkt zu entwickeln: Der Prozess reicht von der Bestimmung der relevanten Faktoren bis zu einer optimalen Einstellung dieser Faktoren.

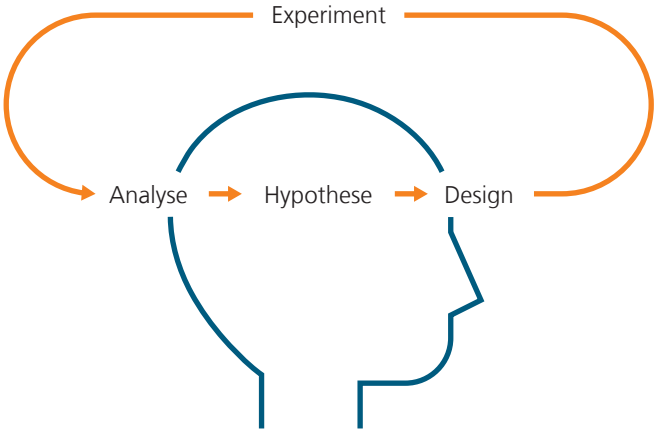


Abbildung 2: Schematische Darstellung des iterativen DoE-Prozesses: Aufstellen einer Arbeitshypothese, Design des Versuchsplans, Durchführen der Experimente und Analyse der Ergebnisse. Grafik nach [14].

Bei der Aufstellung passender Arbeitshypothesen verfolgen wir am Fraunhofer IAIS die Strategie, intensiv auf das Wissen der jeweiligen F&E-Expert*innen zurückzugreifen, mit denen wir zusammenarbeiten. Ein spezielles Augenmerk liegt dabei auf möglichen Vorannahmen (Bias), die die Expert*innen unter Umständen mitbringen und die nur schwer geprüft werden können. In diesem Fall unterstützen wir unsere Partner*innen dabei, passendere Arbeitshypothesen zu formulieren, aus denen sich ableiten lässt, welche Daten zur Unterstützung oder Ablehnung der Hypothese benötigt werden.

Damit lässt sich mithilfe einer DoE-Methodik ein Versuchsplan erstellen, der darauf zugeschnitten ist, die Hypothese zu überprüfen. Ein bekanntes Beispiel einer solchen Hypothese und dem zugehörigen Experiment ist das »Lady-Tasting-Tea-Experiment« von Ronald Fisher [5].

3.4. Herausforderungen bei der Durchführung und Analyse von DoEs

Die nächsten Schritte nach der Erstellung des Versuchsplans im DoE-Prozess sind die Durchführung und die Analyse der Experimente. Hierbei gibt es einige Herausforderungen zu beachten:

- **Reproduzierbarkeit:** Aufgrund der statistischen Natur vieler Prozesse spielt die aleatorische Unsicherheit eine Rolle. Experimente sind in der Realität nie vollständig reproduzierbar, aber man kann mit genügend Daten die Unsicherheit häufig ausreichend genau abschätzen. Trotzdem sollte stets bedacht werden, dass die aleatorische Unsicherheit nicht größer sein darf als die beobachtbaren Effekte. Hierbei können insbesondere nicht kontrollierbare Faktoren eine Rolle spielen.
- **Messung/Kontrolle der Faktoren:** Wie bei der Reproduzierbarkeit spielt auch hier die aleatorische Unsicherheit eine Rolle. So ist es möglich, dass Faktoren, die nicht kontrolliert werden können, Einfluss auf die Messergebnisse haben, zum Beispiel die Raumtemperatur eines Labors, in dem ein chemischer Versuch durchgeführt wird. In diesem Fall ist es wichtig, diese Faktoren zu dokumentieren, um ihren Einfluss prüfen zu können. Weiterhin unterliegen die einstellbaren Faktoren wie z. B. eine Ofentemperatur ähnlichen statistischen Schwankungen wie die Messergebnisse; sie haben aufgrund der statistischen Natur der Prozesse nicht jedes Mal den exakt gleichen Wert, sondern sind um einen Mittelwert verteilt. Idealerweise sind die Schwankungen wesentlich kleiner als die verwendeten Einstellungen des Faktors, der im Experiment untersucht werden soll.
- **Dokumentation der Experimente/Ergebnisse:** Was auf den ersten Blick trivial klingen mag, nämlich die Dokumentation von Experimenten und deren Ergebnissen, führt in der Praxis immer wieder zu Problemen. Häufig werden Ergebnisse von Experimenten analog erfasst und müssen für die weitere Verarbeitung aufwendig digitalisiert werden.

Bei der anschließenden Auswertung der Daten finden typischerweise Visualisierungsmethoden, klassische statistische Verfahren (wie etwa die Varianzanalyse oder statistische Tests) und moderne ML-Verfahren ihre Anwendung. Insbesondere die Anwendung von ML-Verfahren ist sehr gut für die Analyse geeignet, wenn wenige Annahmen über die Daten getroffen werden können. Die Ergebnisse der Analyse können wiederum verwendet werden, um neue Hypothesen abzuleiten.

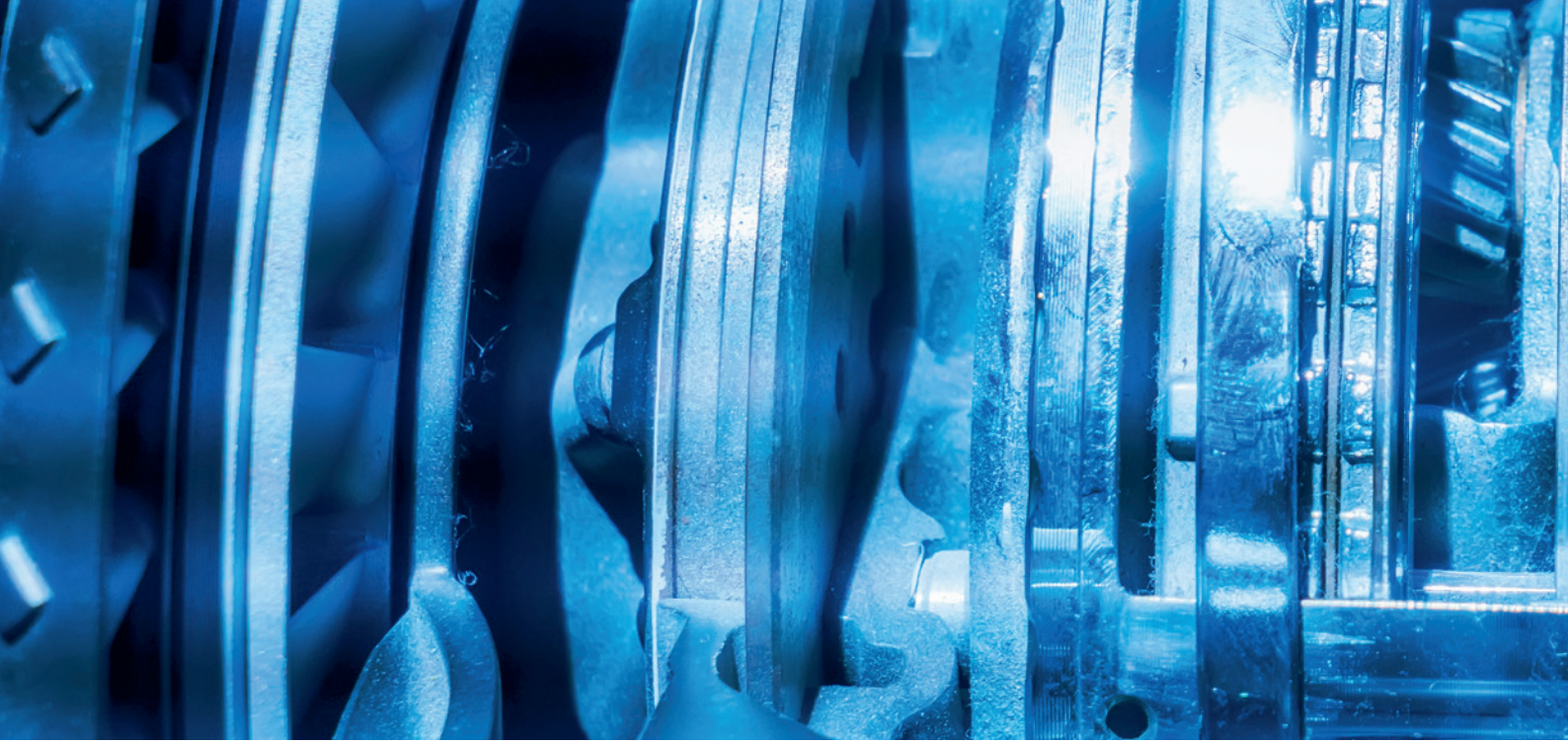
Lady Tasting Tea

Beim Lady-Tasting-Tea-Experiment handelt es sich um ein randomisiertes Experiment mit folgender Versuchsanordnung:

Lady Bristol behauptet unterscheiden zu können, ob bei einer Tasse Tee zuerst die Milch und dann der Tee eingeschenkt wurde (milk-in-first) oder ob zuerst der Tee und dann die Milch eingeschenkt wurde (tea-in-first).

Fisher startet daraufhin mit der Hypothese, dass Lady Bristol nicht in der Lage ist, zu unterscheiden, was zuerst eingeschenkt wurde. Er schlägt als Versuchsplan vor, dass Lady Bristol in zufälliger Reihenfolge acht Tassen Tee serviert werden (vier zuerst mit Tee und vier zuerst mit Milch gefüllt), also acht Experimente durchzuführen. Die Hypothese wird nur dann verworfen, wenn Lady Bristol es schafft, alle acht Tassen korrekt zu klassifizieren: Nur in diesem Fall ist die Wahrscheinlichkeit, zufällig richtig zu liegen, kleiner als 5 Prozent und damit geringer als das übliche Signifikanzniveau von $p = 0.05$.

Das Lady-Tasting-Tea-Experiment ist damit ein Musterbeispiel eines gelungenen DoE: Es veranschaulicht sowohl die korrekte Aufstellung einer Hypothese als auch, wie es möglich ist, einen Versuchsplan zu erstellen, der diese mit so wenig Experimenten wie möglich überprüft.



4. Formalisierung und Operationalisierung des Prozesses

Für die praktische Anwendung haben wir am Fraunhofer IAIS auf Basis unserer Erfahrungen aus Industrie- und Forschungsprojekten den DoE-Prozess in drei Phasen unterteilt (siehe Abbildung 3). Die ersten beiden Phasen beschreiben das bekannte Vorgehen, welches in der Industrie üblich ist. Phase 3 erweitert das Vorgehen des klassischen DoE um einen sequenziellen Ansatz, der mithilfe von ML-Modellen die Versuchsanzahl um bis zu 50 Prozent reduzieren kann. Die Phasen bezeichnen die grundsätzliche Art, mit der Versuchspläne erzeugt werden. Unser Fokus liegt dabei auf der Verbesserung und Weiterentwicklung unserer Methoden für die Erstellung von sequenziellen Versuchsplänen. Das Ziel besteht darin, mithilfe von KI-Methoden und gezielten Versuchen neue Daten zu erzeugen, die die Modelle verbessern, oder möglichst schnell einen optimalen Wert der Zielgröße zu erreichen.

In Phase 1 eines DoEs werden zunächst die Zielgrößen (Responsevariablen) und die möglichen (Einfluss-)Faktoren definiert. Sie ist die Grundlage für die Erstellung eines erfolgreichen Entwicklungsprojektes und sollte daher mit großer Sorgfalt geschehen. In Phase 1 ist der Einfluss des Expertenwissens am größten, da Wissen über die Zusammenhänge der Faktoren und Zielgrößen direkt eingebracht werden kann.

In Phase 2 werden mit fixed-size DoEs, also DoEs, bei denen eine feste Anzahl von Versuchen vor der Durchführung der Versuche festgelegt wird, erste Informationen über den Prozess gesammelt. In dieser Screening-Phase wird typischerweise zwischen DoEs unterschieden, mit denen man feststellen kann, ob ein Faktor einen Einfluss auf die Zielgröße hat, und DoEs, um darauf aufbauend die Größe dieses Einflusses festzustellen. Wird der Einfluss genügend genau bestimmt, können erste Modelle des Prozesses erstellt werden.

Diese Schritte lassen sich in der Praxis häufig nicht klar voneinander abgrenzen. Es ist aber wichtig, festzuhalten, dass es sich um unterschiedliche Fragestellungen handelt.

Nach der Erstellung eines initialen Datensatzes findet der Übergang zu Phase 3 statt, in der mit sequenziellen Versuchsplänen gearbeitet wird. Hierbei werden nach und nach Versuche zu den existierenden Versuchsplänen hinzugefügt, um ein möglichst genaues Modell der Zielgröße zu erzeugen oder optimale Kombinationen von Faktoren für eine Zielgröße zu finden. Dabei ist zu beachten, dass diese Ziele durchaus konkurrieren können – um zum Beispiel ein gutes Optimum zu finden, muss nicht der gesamte Eingangsraum bekannt sein.

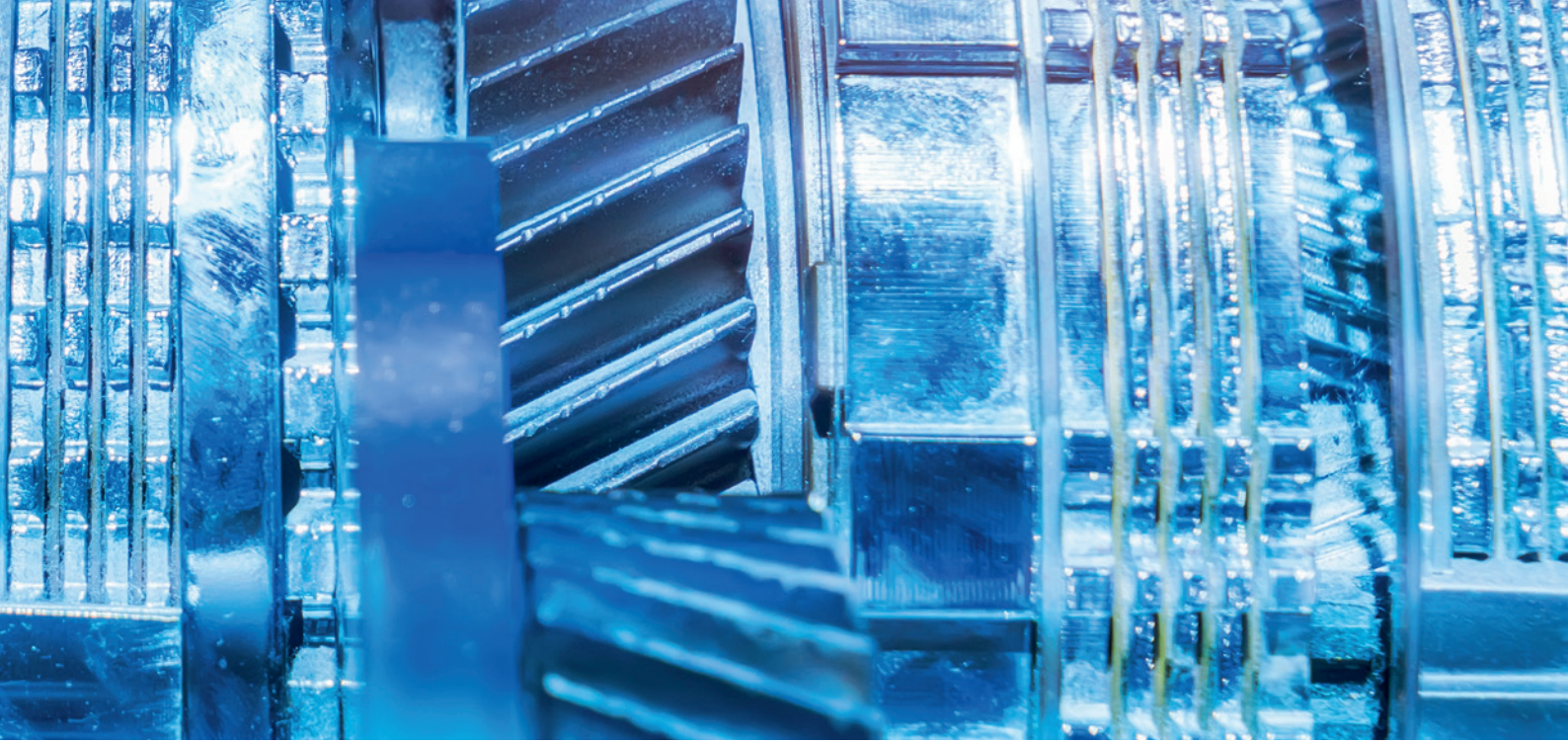


Abbildung 3: Schematische Darstellung des dreiphasigen DoE-Prozessmodells. Der Einfluss des Expertenwissens ist in der zweiten Phase am größten und wird in Richtung Phase 3 nach und nach geringer, während der Einfluss der Daten (Versuchsergebnisse) steigt.

4.1. Der Einfluss von Expertenwissen und Daten

Für den Erfolg eines DoEs haben zwei Dinge einen sehr großen Einfluss: das vorhandene Wissen von Expert*innen und die erzeugten Daten. Zu Beginn des DoE-Prozesses ist der Einfluss des Wissens besonders groß, während der Einfluss der Daten im Verlauf des Prozesses stetig zunimmt.

Der Begriff des Expertenwissens bezieht sich sowohl auf die Domänenexpertise der Anwender*innen allgemein als auch auf die Fähigkeit, statistisch valide und sinnvolle Versuchshypothesen zu formulieren. Besonders wichtig ist hierbei, alle notwendigen Faktoren bei der Hypothesenbildung zu berücksichtigen – es besteht immer das Risiko, wichtige Alternativhypothesen zu vernachlässigen und dies aufgrund des experimentellen Designs nicht zu bemerken. Ein sehr einfaches Beispiel für diesen Fall ist die Annahme eines monotonen Verlaufs der Zielgrößen, die bei manchen Designs verwendet wird. Hierbei werden Experimente allein an den Minima und Maxima der Faktoren durchgeführt, um festzustellen, ob diese einen großen Einfluss haben. Dies vernachlässigt aber die Möglichkeit einer starken Änderung der Zielgrößen zwischen den Extrema

der Faktoren, was dazu führen kann, dass diese Faktoren nicht weiter berücksichtigt werden.

Wenn wenig oder nur unsicheres Expertenwissen vorhanden ist, lohnt es sich, mehr Experimente durchzuführen, um verschiedene Hypothesen zu prüfen. In extremer Ausprägung führt dies zu sogenannten raumfüllenden Designs, deren Hauptziel es ist, möglichst viel Information zu sammeln. Bei ausreichendem Expertenwissen ist es einfacher, Hypothesen, wie ein monotones Prozessverhalten, aufzustellen und anschließend zu prüfen. Je mehr gesichertes Wissen vorhanden ist, desto genauer können valide Hypothesen formuliert werden. Die Kombination von Expertenwissen und (Experiment-)Daten ergibt jedoch erst die gewünschte Prozessbeschreibung – wobei die Anzahl der notwendigen Experimente sich mit steigendem Expertenwissen verringert.

Expertenwissen kann auf verschiedene Arten in Hypothesen formuliert werden: zum Beispiel als einfache statistische Hypothese (»die Zielgröße ist normalverteilt«), als komplexere Hypothese unter der Annahme von Monotonie (»wenn Faktor A sich vergrößert, steigt auch die Zielgröße«), als polynomielles Modell (»es besteht ein linearer/quadratischer/… Zusammenhang zwischen Faktor und Zielgröße«) und als informiertes Machine-Learning-Modell.

Informierte ML-Modelle sind besonders flexibel, was die Integration unterschiedlichen Wissens betrifft, sodass auch komplexe Hypothesen einfach in das Modell integriert werden können [15]. Im Gegensatz zu klassischen ML-Modellen wie einem neuronalen Netz benötigen informierte ML-Modelle durch das formulierte Wissen eine geringere Datenmenge; sie lernen nicht allein aus Daten, um eine Vorhersage zu treffen, sondern nutzen auch das in Hypothesen formulierte Expertenwissen. Somit kann schon mit einem einfachen informierten ML-Modell die benötigte Datenmenge für eine verlässliche Vorhersage einer Zielgröße erheblich reduziert werden.

In probabilistischer Form sind informierte ML-Modelle außerdem in der Lage, epistemische und aleatorische Unsicherheit abzubilden. Werden nun Daten aus einem DoE mit dem Modell kombiniert (durch Modelltraining oder -update), ist es nicht nur möglich, das Prozessverhalten vorherzusagen, sondern man kann zusätzlich eine Aussage darüber machen, wie sicher man sich dieser Vorhersage ist und wie viel Unsicherheit auf epistemischer Unsicherheit beruht.

Mit steigender Datenmenge sinkt die epistemische Unsicherheit, sodass immer verlässlichere Vorhersagen durch das Modell getroffen werden können. Zusätzlich sinkt der Einfluss des Expertenwissens und es ist sogar möglich, Falschannahmen aufzudecken. Diese treten insbesondere bei hochdimensionalen Problemen (Probleme mit einer großen Anzahl von Faktoren) auf, für die die Formulierung von Hypothesen besonders schwierig ist. Probabilistische informierte ML-Modelle sind somit die ideale Ergänzung in den Phasen 1 und 2 des DoE-Prozesses: Sie vereinen bei hoher Flexibilität Expertenwissen mit den gewonnenen Daten in einem gemeinsamen mathematisch-präzisen Ausdruck (Modell). Trotzdem sollte hierbei immer die Angemessenheit der jeweiligen Formulierungsmethode hinterfragt werden: Einfache Hypothesen benötigen kein komplexes Modell, komplexe Hypothesen können oft auch mathematisch nicht einfach ausgedrückt werden.

Von Daten zum probabilistischen informierten Machine-Learning-Modell

Die datenbasierte Modellierung hat eine lange Tradition und hat sich besonders in den letzten Jahren immer weiterentwickelt. Die Methode der kleinsten Quadrate (»Least Squares«), ermöglicht es seit Beginn des 19. Jahrhunderts, lineare Modelle zu erstellen, wenn die korrekte Modellfunktion (zum Beispiel ein Polynom) bekannt ist. In unserem Beispiel (Bild b) wurde durch die Daten aus Bild a eine lineare Funktion gefittet. Ein (nichtlineares) ML-Modell (Bild c) besitzt hingegen mehr Flexibilität und passt sich den Daten an: Der große Unterschied zum Modell in b ist, dass das Modell (zum Beispiel der Grad des Polynoms) durch die Daten bestimmt ist – mit steigender Modellkomplexität werden die Modelle immer stärker durch sogenannte Hyperparameter gesteuert, die lediglich einen indirekten Einfluss auf das Modellierungsergebnis haben. Informierte ML-Modelle (Bild d) dagegen erlauben eine direkte Einflussnahme – hier wurde zum Beispiel angenommen, dass außerhalb der Daten ein fester Wert vorhergesagt werden soll. Somit kann die Flexibilität von ML-Modellen mit bekanntem Wissen verbunden werden.

Probabilistische informierte ML-Modelle (Bild e) verfolgen einen ähnlichen, aber von der Herkunft gegenteiligen Ansatz: Hier entsteht nicht ein Modell, welches am besten zu den Daten passt und das nur einen einzelnen Wert vorhersagt, sondern es wird eine Menge möglicher Modelle definiert und dann aufgrund der Daten immer weiter eingegrenzt. Ist zum Beispiel die Hypothese, dass der Wert der Zielgröße immer auf einen festen Wert zurückfällt, sagt das Modell dort nicht nur einen Wert vorher, sondern eine Verteilung der möglichen Werte aufgrund der bestehenden Datenlage. Die Unsicherheit dieser Verteilung ist aus der epistemischen und der aleatorischen Unsicherheit zusammengesetzt, die das Modell annimmt. In Bereichen mit vielen Daten ist die epistemische Unsicherheit gering; sie wächst in datenarmen Bereichen.

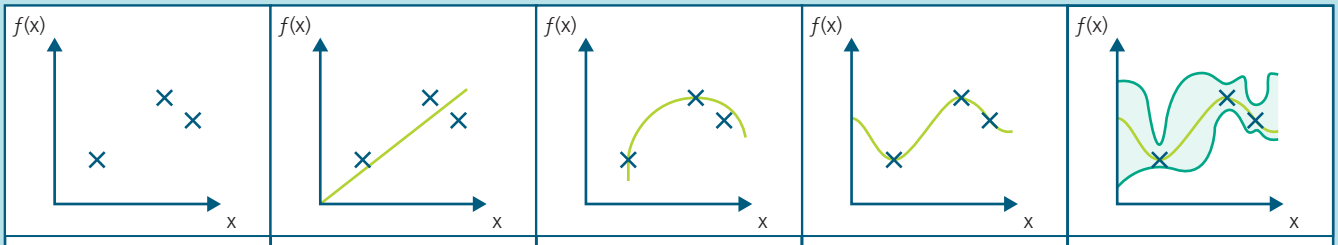


Abbildung 4: Vergleich von verschiedenen Modelltypen

4.2. Phase 1: Definition der Einflussfaktoren und Zielgrößen

Die Grundlage des DoE-Prozesses bildet Phase 1: Die Definition der Einflussfaktoren und Zielgrößen. Die Definition der Zielgrößen bildet das Projektziel ab: Hier sollte eindeutig geklärt werden, welche Variablen warum erfasst werden sollen. Für eine Neuentwicklung im Lebensmittelbereich können dies verschiedene Geschmacksaspekte, Brennwerte oder die Konsistenz eines Produktes sein; im Bereich der Elektronik eine bestimmte Leitfähigkeit oder aber (wie auch in vielen anderen Bereichen) die Fertigungskosten eines Produktes.

Auf Seiten der Einflussfaktoren gilt es zunächst zu bestimmen, ob alle relevanten Faktoren kontrollierbar sind. Falls erwartet wird, dass nicht kontrollierbare Faktoren Einfluss haben, müssen diese dokumentiert werden, da sonst die Gefahr besteht, dass die Ergebnisse der Experimente unbrauchbar sind. Da nur sehr wenige DoE-Verfahren die Ergänzung von kontrollierbaren Faktoren während/nach ihrer Durchführung zulassen, ohne die Anzahl der Versuche stark zu steigern, sollte die Festlegung mit großer Sorgfalt geschehen.

Zu guter Letzt ist es hilfreich, bereits in dieser Phase vorhandenes Expertenwissen wie Hypothesen und Randbedingungen strukturiert zu dokumentieren. Sie können sowohl bei der Auswahl des DoE-Verfahrens als auch bei der späteren Evaluation des DoE-Prozesses sehr hilfreich sein.

4.3. Phase 2: Fixed-size DoE

Generell ist der DoE-Prozess iterativ: Zunächst werden Hypothesen formuliert, dann wird ein Versuchsplan zur speziellen Prüfung der Hypothese entworfen, anschließend erfolgt die Experimentdurchführung und abschließend werden die experimentellen Ergebnisse analysiert, woraufhin eine neue Hypothese formuliert werden kann.

Doch auch wenn dieser Prozess iterativ ist, kann man bei einem genaueren Blick zwischen fixed-size DoEs (Versuchsplänen fester Größe) und sequenziellen DoEs, je nach Phase des DoE-Prozesses, unterscheiden.

Bei fixed-size DoEs, die Teil der Phase 2 des DoE-Prozesses sind, ist die Anzahl der Versuche in einem Versuchsplan vor der Durchführung vom Anwender festgelegt. Hierbei hängt die Anzahl der Versuche von der Komplexität der zu prüfenden Hypothese und dem zu erreichenden Signifikanzniveau des Ergebnisses ab. Vereinfacht gesprochen steigt die Anzahl der notwendigen Versuche sowohl mit der Komplexität der Hypothese als auch mit dem geforderten Signifikanzniveau. Grundsätzlich ist diese Art von DoE sehr vorteilhaft, wenn sehr konkrete und einfache Hypothesen geprüft werden. Gleichzeitig sind ihre Ergebnisse ein idealer Ausgangspunkt für das Training von Vorhersagemodellen, zum Beispiel probabilistischen informierten ML-Modellen, die für sequenzielle experimentelle Designs benötigt werden. Es ist daher von Vorteil, fixed-size DoEs in den Anfangsphasen des DoE-Prozesses zu verwenden, um grundlegendes Wissen über die Zielgrößen aufzubauen und in Daten auszudrücken.

Wie jedes DoE beruhen fixed-size DoEs auf Hypothesen, die mithilfe der im DoE gewonnenen Daten geprüft oder ergänzt werden. Um die Hypothesen zu formulieren, sollte eine geeignete Methode verwendet werden: Für einfache Zusammenhänge (wie die Monotonie) sind einfache Basis-Methoden sinnvoll, für komplexe Zusammenhänge können

zum Beispiel probabilistische informierte ML-Modelle verwendet werden. Hierbei ist insbesondere darauf zu achten, dass die verwendete Methode nur korrekte Hypothesen wiedergibt: Es darf zum Beispiel keine Monotonie impliziert werden, wenn nicht feststeht, dass diese Hypothese wahr ist, da sonst die Aussage der experimentellen Ergebnisse nicht zuverlässig ist.

4.3.1. Fixed-size DoE: Methodenüberblick

Um ein fixed-size DoE zu erstellen, sind drei verschiedene Methodengruppen anwendbar: klassische Ingenieursmethoden, Methoden aus dem Bereich des Designs von Computer-Experimenten und Designs, die auf probabilistischen informierten ML-Modellen beruhen.

Ein ausführlicher Vergleich klassischer Ingenieurs-DoEs mit DoEs aus dem Bereich von Computer-Experimenten findet in Abbildung 5 auf der nachfolgenden Seite statt. Während DoEs aus dem Ingenieurbereich häufig sehr vereinfachende, grundlegende Annahmen prüfen und daher die notwendige Anzahl von Experimenten so stark wie möglich minimieren, verzichten DoEs aus dem Design von Computer-Experimenten so stark wie möglich auf Annahmen und benötigen daher mehr Versuche.

KI-gestützte DoEs erweitern die Möglichkeiten von fixed-size DoEs: Mit ihnen können mit einer geringen Anzahl von Experimenten komplexe Annahmen auf robuste Art überprüft werden. So wird Domänenwissen zunächst in eine Modellannahme überführt und anschließend werden mithilfe von informationstheoretischen Ansätzen die Experimente bestimmt, die diese Modellannahme optimal prüfen. So können zum Beispiel die Experimente gewählt werden, die die epistemische Unsicherheit maximal reduzieren. Im Gegensatz zu Ingenieursdesigns sind auf probabilistischen, informierten ML-Modellen beruhende DoEs flexibler in der Formulierung von Hypothesen, da abstraktere Eigenschaften, wie Symmetrien und Glattheit, und nicht zum Beispiel der Grad eines Polynoms (wie bei Response-Surface-Methoden) definiert werden müssen. Außerdem sind sie, wie DoEs aus dem Bereich der Computer-Experimente, robust gegenüber Fehldefinitionen und ermöglichen das Hinzufügen und Entfernen von Datenpunkten während der Versuchsdurchführung. Ihr besonderer Nutzen konnte bereits in verschiedenen Studien gezeigt werden, ist jedoch stark vom Anwendungsfall abhängig [16].

Vergleich verschiedener fixed-size DoE-Methoden

Ziel von fixed-size DoEs ist die Prüfung von grundlegenden Hypothesen und damit die Generierung von Basiswissen. Hierbei kann, neben KI-gestützten Designs, zwischen traditionellen Ingenieursmethoden und Methoden aus dem Design von Computer-Experimenten unterschieden werden.

Ingenieursmethoden

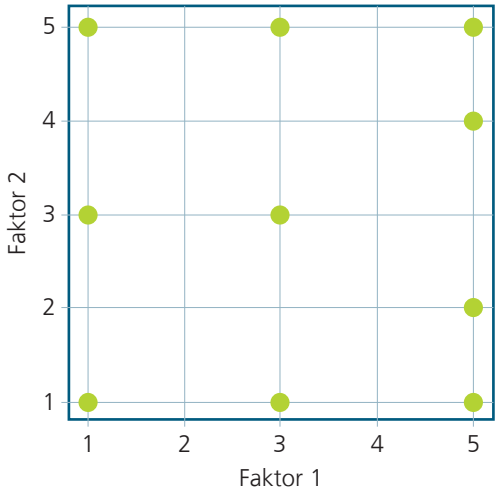
Traditionellerweise werden verschiedene Designs aufeinander aufgebaut. Während man durch ein Screening-Design, wie dem *Plackett-Burman-Design* [17] durch Betrachtung von Maximal-/Minimalwerten herausfinden kann, ob ein Faktor relevant ist, kann man anschließend durch *Fractional-Factorial* [18] und *Response-Surface-Designs* [14] den Einfluss dieser Faktoren klären.

Wichtig ist hierbei, die Versuchsannahmen zu beachten: Diese können in monotonem Verhalten, dem Vorrang der Einzelfaktoren vor ihren Interaktionen oder einem Polynomansatz bestehen. Weiterhin spielen sogenannte *Confounding Factors* eine große Rolle: Es kann sein, dass sich ein Einfluss nicht mehr eindeutig einem Faktor oder einer Interaktion zuordnen lässt.

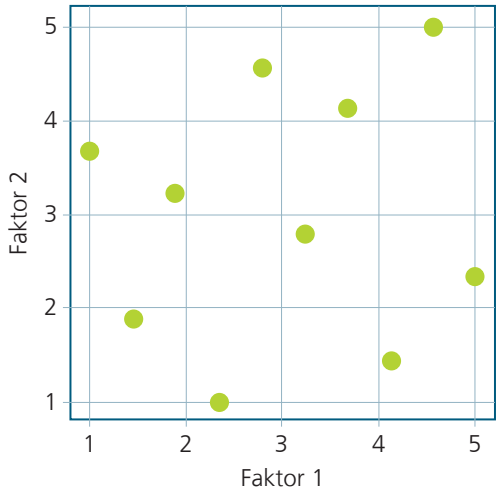
Design von Computer-Experimenten

Methoden aus dem Design von Computer-Experimenten beruhen weniger auf Prozessannahmen und sind somit weniger empfindlich gegen Planungsfehler, benötigen jedoch eine größere Datenmenge. Dies macht sie zwar kostenintensiver, aber robuster.

Aus diesem Bereich sind *Latin-Hypercube-Designs* [19] recht bekannt, während sich die Nutzung *quasirandomer Sequenzen* [20] bislang auf spezielle Bereiche beschränkt. Auch reine *randome Designs* sind möglich, wenn auch nur bei einer sehr großen Anzahl von Faktoren empfehlenswert. Ein weiterer Vorteil dieser Designs ist das einfache Hinzufügen und Entfernen von Punkten und Faktoren zum Versuchsplan – so ist eine besonders große Flexibilität für den Anwender gegeben.



a) D-optimales Response-Surface-Design mit 10 Punkten für ein quadratisches Modell mit 5 Diskretisierungen.



b) Latin-Hypercube-Design mit 10 Punkten für kontinuierliche Faktoren.

Abbildung 5: Vergleich von verschiedenen statischen DoEs: Während das Response-Surface-Design eine vorige Diskretisierung des Raums, eine Entscheidung für eine Modellklasse und ein Optimalitätskriterium benötigt, werden beim Latin-Hypercube-Design die Punkte gleichmäßig im Versuchsraum verteilt.

4.4. Phase 3: Sequenzielles DoE

Bei einer guten Kombination von fixed-size und sequenziellem Design of Experiments ist eine starke Verminderung der Anzahl der notwendigen Versuche zur Findung eines validen Modells oder eines Optimums nachweisbar. In der Fachliteratur konnte gezeigt werden, dass je nach Problem nur ein Viertel der möglichen Versuche für ein Modell gleicher Prädiktionsqualität benötigt wird [16]. In Kundenprojekten mit komplexen Problemstellungen konnten wir bereits Verringerungen der Versuchsanzahl von 20 bis 50 Prozent erreichen. Die individuellen Größen der verschiedenen Versuchsdesigns hängt hierbei vom speziellen Problem ab: Wenn erwartet wird, dass das Prozessverhalten sehr komplex ist, aber nicht durch Experten verlässlich beschrieben werden kann, ist ein größeres fixed-size Design empfehlenswert; ein einfaches oder genauer beschriebenes Prozessverhalten ermöglicht, früher sequenzielle Designs zu verwenden. Der Hauptanspruch an das fixed-size DoE in diesem Fall ist, im Modell Basiswissen über das Verhalten der Zielgrößen zu kodieren.

Grundsätzlich lässt sich im Rahmen des iterativen DoE-Prozesses jedes fixed-size DoE nutzen, um damit interessante Subbereiche zu explorieren und somit sequenzielle DoEs zu erzeugen, die die zweite Phase des DoE-Prozesses darstellen. Interessante Datenbereiche können z. B. Parameterkombinationen sein, die eine besondere Anforderung an die Zielgröße erfüllen. Dies kann zum Beispiel ein besonders hoher Wert sein (das beste Produkt), ein gegenüber kleinen Änderungen der Faktoren unempfindlicher Wert (das robusteste Produkt) oder die Erfüllung eines bestimmten Musters (das gewünschte Produkt).

Klassische Ansätze hierbei sind das manuelle Bestimmen von interessanten Subräumen und die Untersuchung mit einem DoE der Response-Surface-Methode, bei dem bestimmte (typischerweise polynomielle) Beziehungen zwischen Faktoren und Zielgröße angenommen werden. Durch die Raumeingrenzung bedeutet dies eine *manuelle* lokale Verfeinerung des Modells.

Am Fraunhofer IAIS haben wir immer wieder die Beobachtung gemacht, dass die Bestimmung lokaler Einflüsse und das Finden von Optima auch mithilfe von informierten probabilistischen ML-Modellen automatisiert realisiert werden kann, was deutlich effizienter ist als die manuelle Suche nach Unterräumen. Der Ablauf eines KI-basierten sequenziellen DoEs ist in Abbildung 6 dargestellt: Zunächst wird das Modell mit den bestehenden Daten und dem existierenden Expertenwissen trainiert. Anschließend werden mithilfe der Modellvorhersage über die Verteilung des Funktionswertes ein oder mehrere Versuche empfohlen. Nach Durchführung dieser Versuche im Labor wird das Modell mit den entstandenen Daten neu trainiert. Dieser Zyklus wird fortgesetzt, bis die gewünschte Zielsetzung erreicht ist.

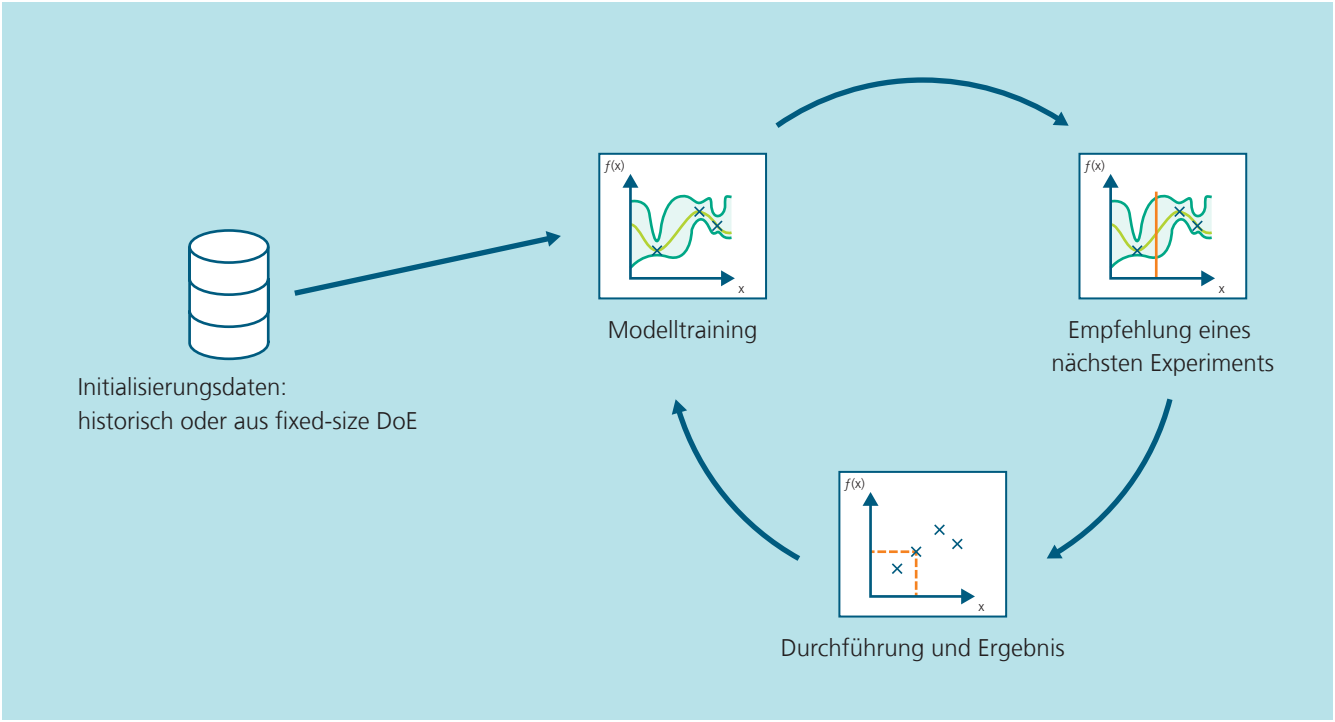


Abbildung 6: Ablauf des KI-basierten sequenziellen DoEs.

Der Einsatz von KI-basierten Methoden bietet viele Vorteile: Zunächst funktionieren diese Methoden auch ohne eine manuelle Eingrenzung des Versuchsraums, da sie in der Lage sind, die entsprechenden Bereiche automatisiert zu finden. Dieser Prozess ist nicht nur automatisiert, sondern auch global: Anstatt eines abgegrenzten Bereichs werden sequenziell alle sinnvoll möglichen Bereiche untersucht. Die Automatisierung macht zusätzlich das DoE weniger vom individuellen Versuchsingenieur abhängig und häufig auch besser reproduzierbar. Weiterhin arbeiten KI-basierte Methoden häufig dateneffizienter als die iterative Anwendung von fixed-size DoEs, da sie mit jedem einzelnen Versuch weiter lernen.

4.4.1. KI-basiertes sequenzielles DoE: Methodenüberblick

Die Basis für ein KI-basiertes sequenzielles DoE bildet ein informiertes probabilistisches ML-Modell, das grundlegendes Wissen über den Prozess abbildet. Dieses Modell dient dazu, erstes Wissen über die Zielgröße abzubilden. Hierbei kann das Wissen unterschiedlich repräsentiert sein: Ein Teil besteht aus Daten, wie historischen Daten oder Daten aus einem gezielten fixed-size DoE, ein anderer Teil aus Expertenwissen und/oder Hypothesen.

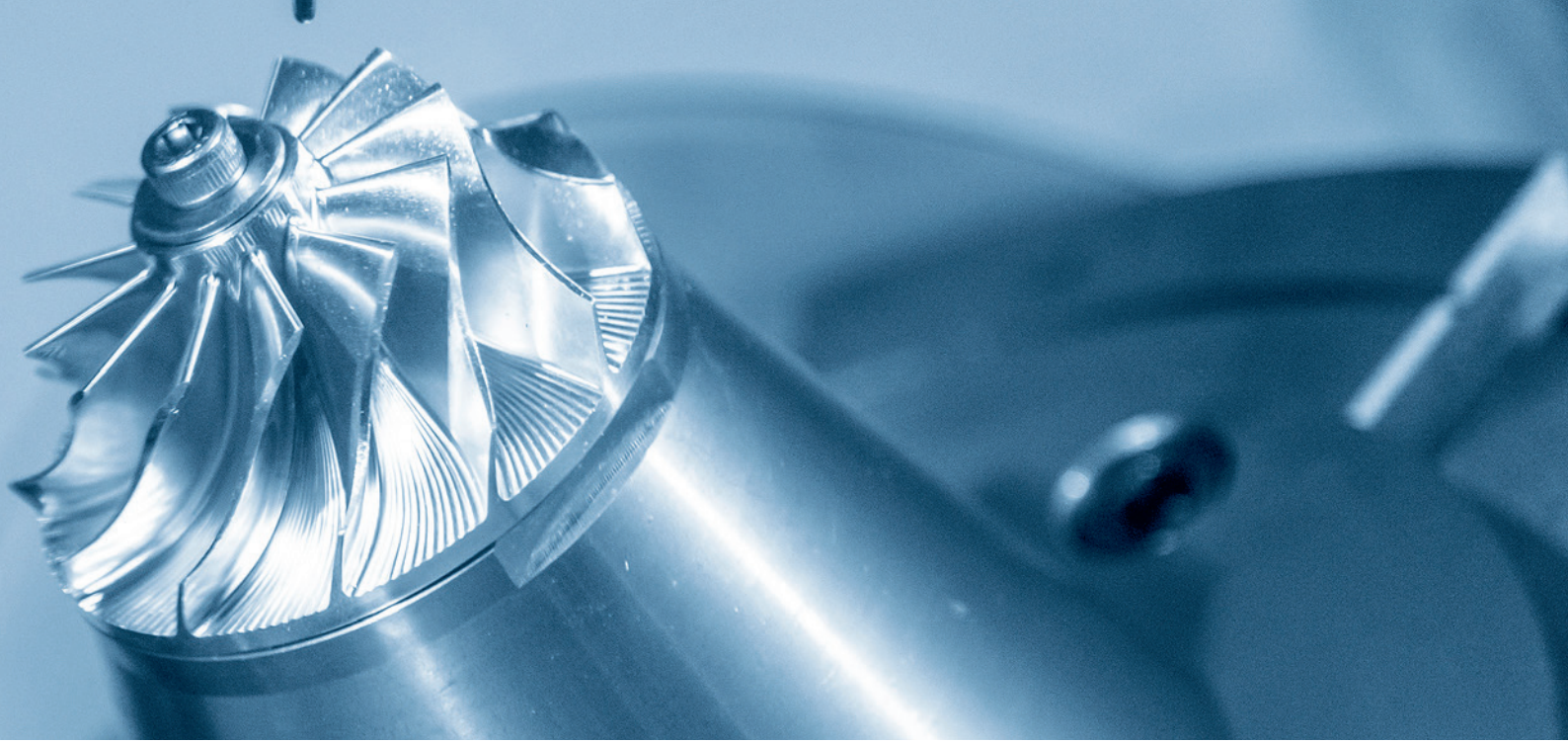
Im nächsten Schritt wird dieses Modell sequenziell verbessert, indem basierend auf dem Modell neue Experimente vorgeschlagen und durchgeführt werden. Anschließend wird das in diesen Experimenten gewonnene Wissen (im einfachsten Fall erzeugte Daten) wieder dem Modell zugeführt und damit neu trainiert. Bei mehrfacher Durchführung dieses Prozesses (Modellupdate, Vorschlag eines Experimentes, Durchführung) kann man somit von einem kontinuierlich lernenden Modell sprechen.

Beim KI-basierten sequenziellen DoE wird zwischen zwei Zielsetzungen unterschieden:

1. Die lokale Verfeinerung des Modells, auch als »Active Learning« bezeichnet.
2. Die Suche nach einem Optimum, auch »Bayessche Optimierung« genannt.

Im Fall von Active Learning wird sequenziell durch das Hinzufügen von experimentellen Daten die epistemische Unsicherheit des Modells verringert: Das Hauptziel der Methode ist ein adäquates Modell des Prozesses. Hierdurch konzentriert sich die Methode auf Bereiche mit großen Fluktuationen der Zielgrößen, da hier die meiste Information über den Prozess vorliegt. Somit ist es möglich, mit einer so geringen Versuchsanzahl wie möglich ein Vorhersagemodell für den Prozess zu erschaffen. In der Anwendung kann dieses Modell dann für neue Kombinationen der Eingangsfaktoren Vorhersagen über den Wert der Zielgröße und die Unsicherheit der gemachten Vorhersage treffen, wodurch weitere Experimente entfallen.

Bayessche Optimierung hingegen konzentriert sich auf optimale Bereiche der Zielgröße. Hierbei muss das Verfahren zwischen der Untersuchung von unbekannten Bereichen (Exploration) und der bekannter, vorteilhafter Bereiche (Exploitation) ausgleichen. Dies führt dazu, dass Bayessche Optimierung andere Versuche vorschlägt, als Active Learning es tun würde. Generell kann es Bereiche geben, in denen eine große Variabilität des Prozesses vorliegt (also für Active Learning besonders relevante Bereiche), in dem die Werte der Zielgröße aber weit weg von anderen möglichen Optima liegen. Abbildung 7 verdeutlicht dies für einen einfachen möglichen Fall. Während für Active Learning der Bereich mit vielen Schwankungen der Zielgröße besonders relevant ist,



konzentriert sich Bayessche Optimierung auf den Ort des Minimums der unbekannten Zielgröße (eine Maximierung ist mathematisch durch Umkehrung des Vorzeichens auf demselben Weg zu erreichen). Das Ergebnis Bayesscher Optimierung ist also im Idealfall ein einzelner Punkt: die Kombination der Inputfaktoren, die zum optimalen Funktionswert führt.

Trotz der offensichtlichen Unterschiede ist der Wechsel von Active Learning zu Bayesscher Optimierung für den Anwender einfach möglich, da beide Methoden auf derselben Idee und auch demselben Modell beruhen. Diese Idee ermöglicht es zusätzlich, während des Vorgehens weitere Faktoren einfach hinzuzufügen oder wegzulassen, sodass für Endanwender*innen eine größtmögliche Flexibilität in Bezug auf wechselnde Anforderungen besteht [21].

Anwendungsstudien bestätigen, was wir in unserer Forschung und in Projekten mit Kunden immer wieder festgestellt haben: KI-basiertes Vorgehen ist besonders dateneffizient. Damit können durch Einsparung von Experimenten Ressourcen, Zeit und Kosten gespart werden.

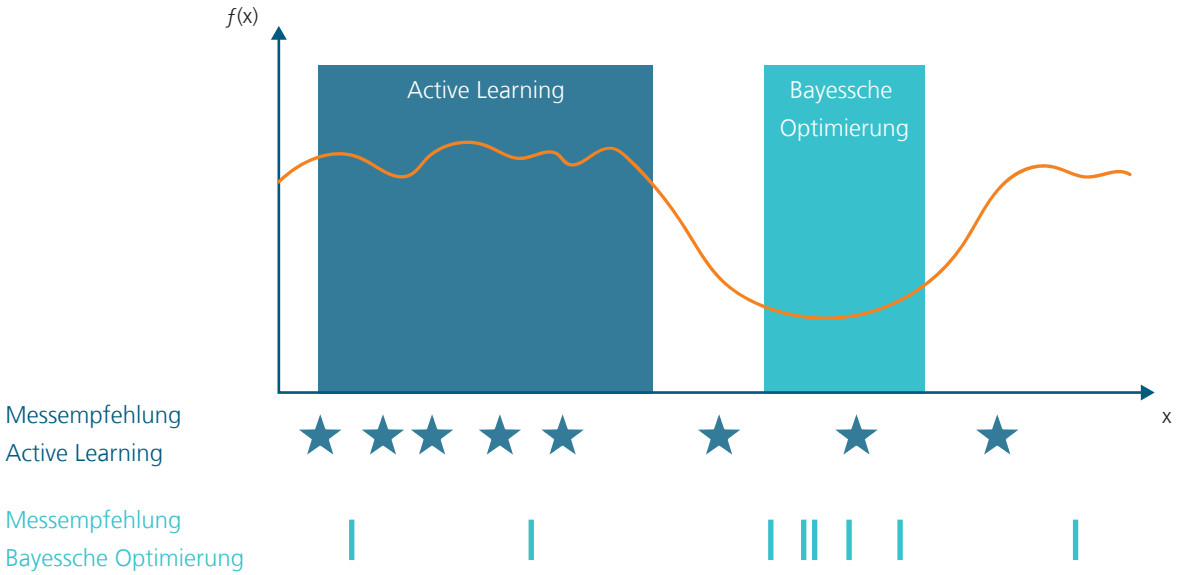


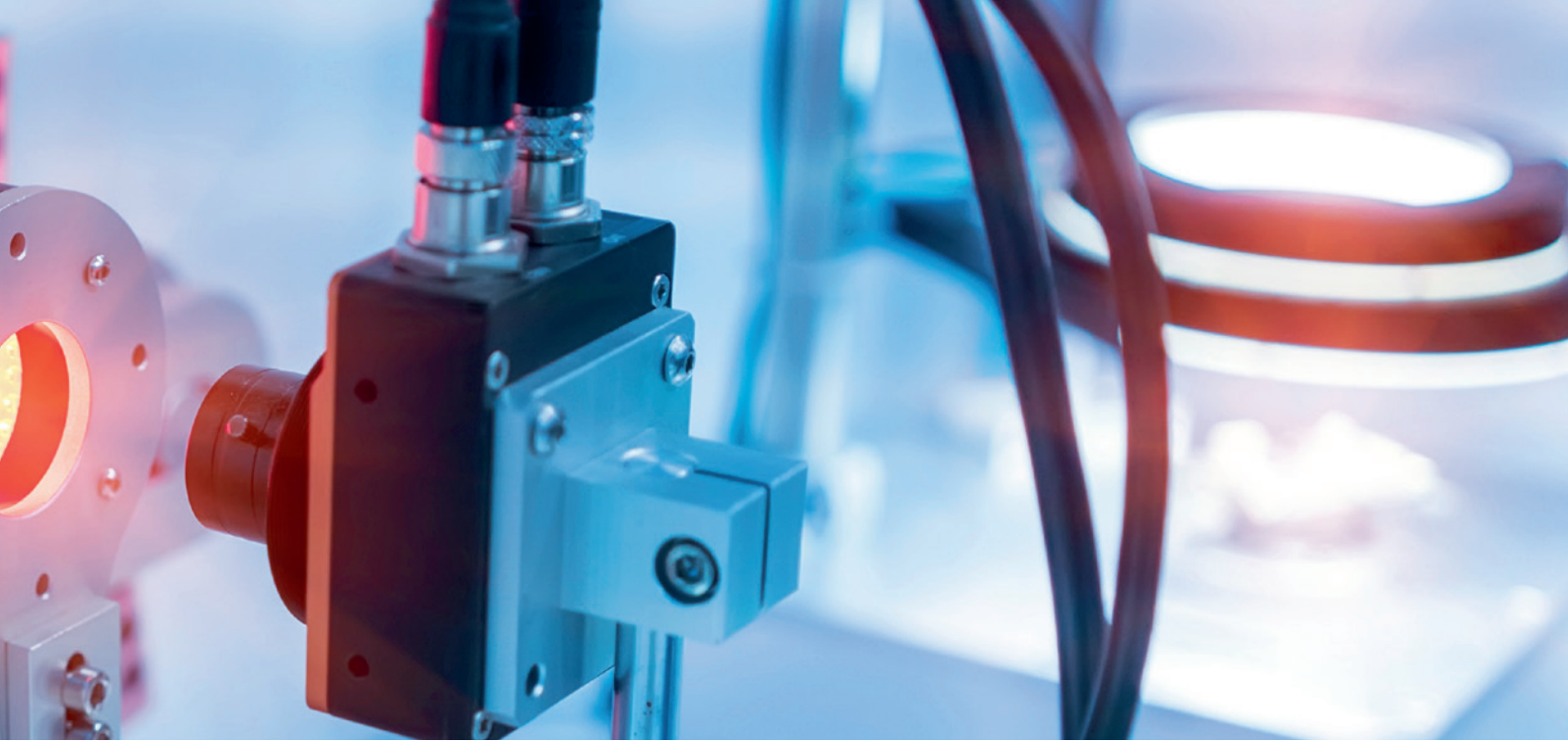
Abbildung 7: Darstellung der Fokusbereiche für Active Learning und Bayessche Optimierung. Messempfehlungen für Active Learning sind als Sternchen, für Bayessche Optimierung als vertikale Linien markiert. Während sich Active Learning sequenziell auf Bereiche konzentriert, in denen die Zielgröße viele Schwankungen aufweist, fokussiert sich Bayessche Optimierung auf die Optima der Zielgrößen.

5. Mögliche Anwendungen

Die Anwendung des DoE-Prozesses ist in allen Bereichen empfehlenswert, in denen Experimente teuer sind, wobei dies sowohl die Kosten für Material- und Personaleinsatz als auch die zeitliche Komponente (zum Beispiel Alterungsexperimente) betreffen kann. Das Anwendungsfeld, insbesondere für KI-basierte sequenzielle DoEs, ist daher sehr divers. Unsere Projekte sowie dokumentierte Anwendungen finden sich in Chemie und Pharmazie [22], in klassischen Ingenieursfeldern wie der Konstruktion [23], Simulation [24] und Robotik [25], sowie im Hyperparameter-Tuning für Machine-Learning-Algorithmen [26].

So divers die Anwendungsfelder sind, so divers sind auch die spezifischen Herausforderungen an das Versuchsdesign: Sowohl Einflussfaktoren als auch die Zielgrößen können numerisch, diskret oder kategoriell sein; sie können schwach verrauscht (z. B. große Simulationen) oder stark verrauscht (z. B. menschliches Feedback) sein. Weiterhin gilt es, spezielle Strukturen und Randbedingungen abzubilden: So können in der Konstruktion nur bestimmte Bauteilgeometrien zusammengefügt werden oder in der chemischen Formulierung Kombinationen von Faktoren nicht mehr verarbeitbar sein.

Wenn solche Anforderungen bekannt sind, gilt es, diese als Hypothese in allen Phasen des DoE-Prozesses zu berücksichtigen. Hier haben sich informierte probabilistische ML-Modelle nicht nur in der zweiten, sequenziellen Versuchsphase, sondern auch für fixed-size Designs bewährt [16], sodass die Anzahl der Experimente um rund 20 bis 50 Prozent reduziert werden kann.



6. Umsetzung mit dem Fraunhofer IAIS

Das Fraunhofer IAIS ist ein verlässlicher Partner bei der Umsetzung von DoE-Projekten. Wir begleiten Sie durch alle Schritte des DoE-Prozesses und unterstützen Sie im Umgang mit allen Herausforderungen Ihres Projekts.

Unsere Expert*innen helfen bei der Definition des konkreten Projektziels, insbesondere beim Aufstellen einer geeigneten Zielfunktion und bei der vollständigen Bestimmung der Einflussfaktoren und der Zielgrößen.Wir unterstützen Ihre Expert*innen bei der Übersetzung Ihres Domänenwissens in passende Randbedingungen für die Faktoren und sorgen damit für eine solide Basis, damit Ihr geplantes Projekt Ihre fachlichen und unternehmerischen Ziele erreicht.

Für die weiteren Schritte des DoE-Prozesses bieten wir Ihnen sowohl fachliche, technische als auch Anwendungsschulungen an. Darüber hinaus bieten wir bei komplexeren Problemen auch individuell abgestimmte Beratungen an, von denen sowohl Entwickler*innen als auch Anwender*innen profitieren.

7. Fazit und Ausblick

Design of Experiments ist eine Kernkompetenz von F&E-Abteilungen und beeinflusst deren Arbeitsmethoden stark. Dabei handelt es sich einerseits um ein etabliertes Forschungsfeld und andererseits bieten rasante Entwicklungen im Bereich KI große Chancen, Produktentwicklungszyklen durch gezielte Experimente zu verkürzen und dadurch Kosten zu reduzieren. Der in diesem Whitepaper beschriebene KI-basierte DoE-Prozess mit einer Kombination aus fixed-size und sequenziellen Versuchsplänen erlaubt es, das in Ihrem Unternehmen vorhandene Expertenwissen optimal mit den Möglichkeiten der datengetriebenen Modellierung zu verbinden.

Herausfordernde Aufgaben, wie die Optimierung von Produkteigenschaften oder der gezielte Aufbau einer zukunftssicheren Datenbasis, sind mithilfe der von uns verwendeten Methoden effizient lösbar. Dabei bieten die von uns eingesetzten Verfahren eine hohe Flexibilität in der Anwendung: Sowohl das Hinzufügen als auch das Entfernen von Faktoren in Experimenten ist möglich, ebenso der Wechsel zwischen Bayesscher Optimierung und Active Learning. Gezielt eingesetzt lassen sich so gleichzeitig Entwicklungskosten sparen und Daten erzeugen, mit denen zukünftige Fragestellungen in deutlich höherer Geschwindigkeit beantwortet und Entwicklungen beschleunigt werden können.

Wir sind gespannt, in welchen Kontexten Sie DoE-Verfahren einsetzen und laden Sie herzlich ein, mit uns ins Gespräch zu kommen.

Literaturverzeichnis

[1] OECD Data, Gross domestic spending on R&D (Chart), 26 September 2022. [Online]. Available: https://data.oecd.org/chart/6F0M.

[2] OECD Data, Gross domestic spending on R&D, 26 September 2022. [Online]. Available: https://data.oecd.org/rd/gross-domestic-spending-on-r-d.htm.

[3] R. Bellman, Adaptive Control Processes: A Guided Tour, Princeton University Press, 1961.

[4] DIN 50100:2016-12, Schwingfestigkeitsversuch - Durchführung und Auswertung von zyklischen Versuchen mit konstanter Lastamplitude für metallische Werkstoffproben und Bauteile, 2016.

[5] R. A. Fisher, The Design of Experiments, Edinburgh: Oliver & Boyd, 1935.

[6] Wikipedia Foundation, Konstruktionsfehler, 26 September 2022. [Online]. Available: https://de.wikipedia.org/wiki/Konstruktionsfehler#Auswirkungen_und_Vermeidung.

[7] R. Dillerup und R. Stoi, Unternehmensführung. 4. Auflage, Vahlen, 2013.

[8] Six Sigma Black Belt, Fehlerkosten 10er Regel Zehnerregel (Rule of ten), 26 September 2022. [Online]. Available: https://www.sixsigmablackbelt.de/fehlerkosten-10er-regel-zehnerregel-rule-of-ten.

[9] pwc, Digital Product Development 2025, 26 September 2022. [Online]. Available: https://www.pwc.de/de/digitale-transformation/pwc-studie-digital-product-development-2025.pdf.

[10] ThinkML, Use of AI in Manufacturing Industry, 26 September 2022. [Online]. Available: https://thinkml.ai/use-of-ai-in-manufacturing-industry/.

[11] Forbes, 10 Ways AI Is Improving New Product Development, 26 September 2022. [Online]. Available: https://www.forbes.com/sites/louiscolumbus/2020/07/09/10-ways-ai-is-improving-new-product-development/.

[12] IBM Research, Using AI to Create New Fragrances, 26 September 2022. [Online]. Available: https://www.ibm.com/blogs/research/2018/10/ai-fragrances/.

[13] IBM Research, How to use AI to discover new drugs and materials with limited data, 26 September 2022. [Online]. Available: https://research.ibm.com/blog/ai-discovery-with-limited-data.

[14] G. E. P. Box und N. R. Draper, Response surfaces, mixtures, and ridge analyses, John Wiley & Sons, 2007.

[15] L. von Rueden, S. Mayer, K. Beckh, K. Georgiev, S. Giesselbach, R. Heese, B. Kirsch, J. Pfrommer, A. Pick, R. Ramamurthy, M. Walczak, J. Garcke, C. Bauckhage und J. Schuecker, Informed Machine Learning – A Taxonomy and Survey if Integrating Prior Knowledge into Learning Systems, *IEEE Transactions on Knowledge and Data Engineering,* 2021.

[16] A. Krause, A. P. Singh und G. Carlos, Near-Optimal Placements in Gaussian Processes: Theory, Efficient Algorithms and Empirical Studies, *Journal of Machine Learning Research 9,* pp. 235-284, 2008.

[17] R. Plackett und J. P. Burman, The Design of Optimum Multifactorial Experiments, *Biometrika,* pp. 305-325, 1946.

[18] American Society for Quality Control, *ASCQ Glossary and Tables for Statistical Quality Control,* Milwaukee, 1983.

[19] M. D. Mckay, R. J. Beckman und W. J. Conover, A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code, *Technometrics,* pp. 55-61, 2000.

[20] H. Niederreiter, Random Number Generation and Quasi-Monte Carlo Methods, Society for Industrial and Applied Mathematics, 1992.

[21] B. Solnik, D. Golovin, G. Kochanski, J. E. Karro, S. Moitra und D. Sculley, Bayesian Optimization for a Better Dessert, in *Proceedings of the 2017 NIPS Workshop on Bayesian Optimization*, Long Beach, USA, 2017.

[22] B. J. Shields, J. Stevens, J. Li, M. Parasram, F. Damani, J. I. Martinez Alvarado, J. M. Janey, R. P. Adams und A. G. Doyle, Bayesian reaction optimization as a tool for chemical synthesis, *Nature 590,* pp. 89-96, 2021.

[23] A. Deshwal, S. Belakaria und J. R. Doppa, Bayesian Optimization over Hybrid Spaces, in *Proceedings of the 38th International Conference on Machine Learning 139*, 2021.

[24] H. M. Sheikh und P. S. Marcus, Bayesian Optimization For Multi-Objective Mixed-Variable Problems, ArXiv, https://arxiv.org/abs/2201.12767, 2022.

[25] A. Cully, J. CLune, D. Tarapore und J.-B. Mouret, Robots that can adapt like animals, *Nature 521,* pp. 503-507, 2015.

[26] J. Snoek, H. Larochelle und R. P. Adams, Practical Bayesian optimization of machine learning algorithms, in *Proceedings of the 25th International Conference on Neural Information Processing Systems - Volume 2*, 2012.

Impressum

Herausgeber
Fraunhofer-Institut für Intelligente Analyse- und Informationssysteme IAIS
Schloss Birlinghoven 1
53757 Sankt Augustin
www.iais.fraunhofer.de

Autor*innen
Dorina Weichert
Dr. Gunar Ernis

Redaktion
Inga Daase
Silke Loh
Evelyn Stolberg

Grafik
Angelina Lindenbeck
Pascal Ochel

Layout
Pascal Ochel
Achim Kapusta

Bildnachweise
R. Gino Santa Maria - stock.adobe.com (Titelseite)
dizfoto1973 - stock.adobe.com (Seiten 6–7)
Have a nice day - stock.adobe.com (Seiten 8–9)
nevodka.com - stock.adobe.com (Seite 11)
Brian Jackson - stock.adobe.com (Seiten 12–13)
wywenka - stock.adobe.com (Seite 14)
Julia Romanenko - stock.adobe.com (Seite 15)
xiaoliangge - stock.adobe.com (Seiten 16–17)
luchschenF - stock.adobe.com (Seiten 18–19)
Sergey Ryzhov - stock.adobe.com (Seiten 20–21)
Pixel_B - stock.adobe.com (Seiten 24–25)
xiaoliangge - stock.adobe.com (Seiten 26–27)
279photo - stock.adobe.com (Seiten 28–29)

Kontakt

Dr. Gunar Ernis
Tel. +49 2241 14-2689
design-of-experiments@iais.fraunhofer.de

Fraunhofer-Institut für Intelligente
Analyse- und Informationssysteme IAIS
Schloss Birlinghoven 1
53757 Sankt Augustin
www.iais.fraunhofer.de