

Wissensentdeckung und Robustheitsanalyse für Simulationsmodelle weltweiter Netze

Sören Bergmann¹, Niclas Feldkamp¹ und Steffen Straßburger¹

¹ Technische Universität Ilmenau, Fachgebiet Informationstechnik in Produktion und Logistik
soeren.bergmann@tu-ilmenau.de

Abstract. Immer komplexere Netzwerke durchdringen verschiedenste Bereiche des täglichen Lebens. Durch die enge Verzahnung in den Netzwerken und die resultierenden Beziehungen der einzelnen Entitäten im Netz entstehen neben den gewünschten positiven Effekten auch systemische Risiken. Ein proaktives Vorgehen, welches es erlaubt, systemische Risiken im Vorhinein zumindest zu erkennen und ggf. zu vermeiden bzw. zumindest die negativen Auswirkungen einzudämmen ist wünschenswert. Ein möglicher Ansatz ist es, das Netzwerk bzgl. Störungen und Umwelteinflüssen möglichst robust zu gestalten. In diesem Beitrag werden die Grundlagen einer auf Data Farming basierenden Methode zur Robustheitsanalyse im Kontext der Produktion vorgestellt. Im Anschluss werden die Herausforderungen diskutiert, die bei der Adaption der Methode auf komplexe weltweite Netze auftreten.

Keywords: network simulation, data farming, data mining, visual analytics, knowledge discovery, complex networks, robustness analysis, robustness optimization.

1 Einleitung

Immer komplexere Netzwerke durchdringen verschiedenste Bereiche des täglichen Lebens. Die Spanne reicht hierbei von Rechnernetzen als Basis vieler moderner Anwendungen über soziale Netzwerke, Energie- oder Verkehrsnetze bis hin zu hoch dynamischen Produktions- und Logistiknetzwerken oder sogar der kombinierten Betrachtung verschiedener Netze in ihrer Wechselwirkung, z. B. die Betrachtung von Logistiknetzwerken in Kombination mit Energie- und Verkehrsnetzen. Durch die enge Verzahnung in den Netzwerken und die resultierenden Beziehungen der einzelnen Entitäten im Netz entstehen neben den gewünschten positiven Effekten auch systemische Risiken, welche mitunter schwer vorauszusehen, abzuschätzen bzw. zu handhaben sind. Ein aktuelles Beispiel hierfür sind die Versorgungsengpässe mit Benzin und Diesel im Sommer 2018 an einigen Tankstellen im Rheinland, die durch den Niedrigpegel des Rheins und die damit einhergehende eingeschränkte Schifffahrt sowie den Brand einer Raffinerie ausgelöst wurden und durch verändertes Kaufverhalten der verunsicherten Autofahrer verstärkt wurden [21].

Ein proaktives Vorgehen, welches es erlaubt, systemische Risiken im Vorhinein zumindest zu erkennen und ggf. zu vermeiden bzw. zumindest die negativen Auswirkungen einzudämmen ist das Gebot der Stunde. Ein Lösungsansatz kann darin bestehen, Netze möglichst robust gegenüber unkontrollierbaren Umwelteinflüssen und Störungen zu gestalten. D. h., eine Netzkonfiguration muss so gewählt werden, dass sie auch unter negativen Umweltbedingungen und beim Auftreten von Störungen im Netz eine möglichst gute Systemperformanz garantiert bzw. negative Effekte mildert bzw. beherrschbar macht. Die große Herausforderung besteht hierbei gerade in komplexen Systemen wie weltweiten Netzen darin, solch eine Konfiguration zu finden.

Einen Ausgangspunkt für eine Methode zur Entwicklung robuster Netzwerkkonfigurationen können die Arbeiten von Feldkamp, Bergmann und Straßburger darstellen, welche für Intralogistik- und Produktionssysteme Untersuchungen zur Entdeckung von sowohl robusten [12] als auch effizienten [14] Konfigurationen veröffentlicht haben.

Ziel dieses Beitrags ist, anzudiskutieren inwiefern diese (oder ähnliche Methoden) für komplexere Netzwerke anwendbar sind und welche Forschungsbedarfe und Herausforderungen hierdurch auftreten. Hierzu werden zunächst im Kapitel 2 die Grundlagen zu den Themen Simulation komplexer Netzwerke, Data Farming, Knowledge Discovery in Simulation Data und zur Robustheitsbewertungsmethode nach Taguchi gelegt. Im Kapitel 3 wird ein für Produktions- und Logistiksysteme entworfenes und bereits veröffentlichtes Konzept zur Robustheitsanalyse vorgestellt. Im darauffolgenden Kapitel 4 wird diskutiert, inwiefern das Konzept auch für komplexe Netzwerke adaptiert werden kann und welche Herausforderungen bzw. Forschungsbedarfe hierbei entstehen. Hierbei sollen Denkansätze skizziert und offene Forschungsfragen abgeleitet werden. Ein Kapitel mit Fazit und Ausblick schließt den Beitrag ab.

2 Grundlagen

2.1 Simulation komplexer Netzwerke

Simulation ist ein etabliertes Werkzeug zur Planung und Steuerung komplexer Systeme. Insbesondere ist Simulation eine Schlüsselkomponente auch hinsichtlich der Herausforderungen im Kontext von Industrie 4.0 [25]. Unter Simulation wird hierbei die „Nachbildung eines Systems mit seinen dynamischen Prozessen in einem experimentierbaren Modell [verstanden], um zu Erkenntnissen zu gelangen, die auf die Wirklichkeit übertragbar sind“ [50].

Für den Begriff der komplexen Netze liegt keine generell akzeptierte Definition vor. Allgemein kann man aber davon ausgehen, dass komplexe Netzwerke aus der nicht linearen Verbindung einer großen Anzahl von (wiederum ggf. komplexer) Entitäten bestehen. Zudem ist zu konstatieren, dass das Verhalten von komplexen Netzwerken ebenfalls komplex ist und somit das Entwerfen und Steuern solcher Netzwerke keinesfalls trivial ist.

Der Einsatz der Simulation kann sich unter solchen Bedingungen als sehr vorteilhaft erweisen [26]. So gibt es in der Simulationsforschung aber auch in der Praxis eine Vielzahl von Anwendungsgebieten für die Simulation komplexer Netze. Dies reicht von der

Simulation von Supply-Netzwerken [44] über die Betrachtung von Energie- oder Rechnernetzwerken [1, 4] bis hin zu Simulation von sozialen Netzwerken [7, 28, 36]. Hierbei werden verschiedene Simulationsmethoden und Weltansichten eingesetzt. Das Spektrum reicht von systemdynamischer über agentenbasierter bis hin zu diskret ereignisorientierter Simulation.

2.2 Data Farming und Knowledge Discovery in Simulation Data

Die Methode des Data Farmings beschreibt die Verwendung eines Simulationsmodells als Datengenerator. Hierbei wird mit Hilfe von effizientem Experimentdesign und High Performance Computing das Ziel verfolgt, eine möglichst vollständige Abdeckung des Spektrums von Eingangs- und Ergebnisdaten zu erreichen und somit den Informationsgewinn zu verbessern [6, 19, 20, 39]. Analog zu einem Farmer, der sein Feld möglichst effizient kultiviert und versucht, die Qualität und Quantität seiner Ernte zu maximieren, drückt die „Farming“-Metapher aus, dass hierbei der Datenertrag des Simulationsmodells maximiert werden soll [40].

Eine wichtige Kernkomponente der Methode Data Farming ist das Experimentdesign. Hierbei ermöglichen neue Ansätze für die Gestaltung der Simulationsexperimente die umfassende Abbildung möglicher Wertekombinationen von Eingabeparametern bei gleichzeitig vertretbaren Datenmengen [23, 41]. Anfangs wurde die Methode des Data Farming für die militärische Gefechtssimulation entwickelt, wobei ursprünglich auch die inkrementelle Modellerstellung Teil der Methode war. Der Hintergrund hierfür bestand darin, dass sowohl die genutzten Modelle als auch die daran gestellten Fragestellungen eine hohe Komplexität aufweisen, die nicht bzw. nicht gut mit klassischen Simulationsstudien gehandhabt werden können [20]. Abzugrenzen vom Data Farming sind Methoden, welche Experimentdesigns nutzen um spezifizierte Hypothesen zu testen [29, 30].

Aufbauend auf der dargestellten Data-Farming-Methode wurde ein Konzept, siehe Abb. 1., zum Auffinden von versteckten, potenziell nützlichen Wirkzusammenhängen in gefarmten Ergebnisdaten nicht-militärischer Simulationsmodelle, insbesondere im Kontext der Simulation von Produktions- und Logistiksystemen, entwickelt [8, 11]. Hierbei werden die mittels Data Farming erzeugten sehr großen Datenmengen (Big Data) mit Data-Mining-Methoden verarbeitet. Einsetzbar ist hierbei ein breites Spektrum von Methoden wie z. B. Clustering, Entscheidungsbäume oder Frequent Pattern Mining.

Die eigentliche Analyse der generierten Simulationsergebnisdaten sowie deren Beziehungen zu den Eingangsparametern (Faktoren) nutzt zudem interaktive visuelle Analysen. Visualisierung ist im allgemeinen ein wichtiges Werkzeug, wenn eine Interpretation von Daten gefordert ist. Die inkrementelle Kombination von Data Mining und der interaktiven Visualisierung wird auch als Visual Analytics bezeichnet [22, 49].

Übliche bei Simulationsstudien eingesetzte Visualisierungstechniken sind die Animationen, Time-Plots oder Graphen bestimmter Kenngrößen z. B. in einem Konfidenzintervall über replizierte Simulationsläufe [26]. Die bereits angesprochene Forschungsdisziplin Visual Analytics geht über die klassische Visualisierung insoweit hinaus, dass die konsequente Verzahnung von Datenanalyse und -visualisierung angestrebt wird,

deren Verbindungsglied die menschliche Fähigkeit zur Schlussfolgerung darstellt. Gefördert wird dies durch ein hohes Maß an Interaktivität, wie z. B. Filtern und Zoomen [22, 49].

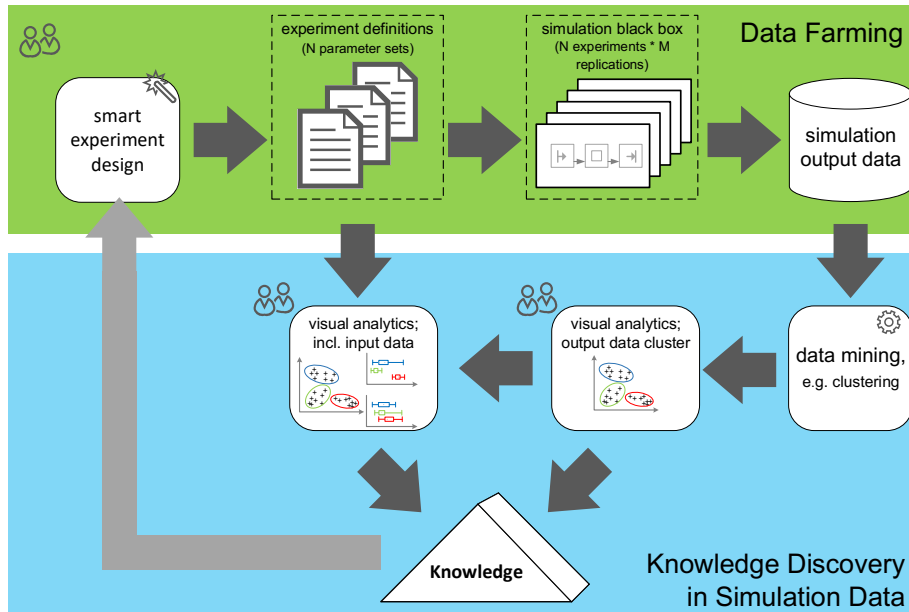


Abb. 1. Der Knowledge Discovery in Simulations (KDS) Prozess [8]

Dieses Konzept wurde in den Ansatz zur Wissensentdeckung in Simulationsdaten integriert [9, 10] und in Industriefallstudien erprobt. Abb. 1 zeigt das schematische Vorgehensmodell für die Wissensentdeckung in Simulationsdaten. Anwendungen in der Praxis, u. a. im Bergbau [11] oder auch im Bereich Automotive [13, 46] zeigen das enorme Potential aber auch offene Herausforderungen der Methode.

Verwandte

2.3 Die Taguchi-Methode zur Robustheitsbewertung

Genichi Taguchi [47, 48] entwickelte eine häufig im Qualitätsmanagement eingesetzte Methode, um Entscheidungsalternativen nicht nur anhand ihrer Ergebniswerte, sondern auch anhand der Streuung der Ergebnisse bei Auftreten von Störeinflüssen (engl. Noise) zu bewerten. Anders ausgedrückt ist die beste System- oder Prozesskonfiguration möglicherweise nicht immer diejenige mit dem höchsten Einzelwert bzw. besten Mittelwert einer untersuchten Zielgröße, sondern diejenige Konfiguration, die unter Einfluss von verschiedenen Störeinflüssen stabile Werte für die Zielgrößen aufweisen. Zu trennen sind bei derartigen Betrachtungen verschiedene Arten von Faktoren, die die Ergebnisse (f) eines Systems beeinflussen. Hierbei wird unterschieden zwischen Kon-

trollfaktoren (b), welche auch im Realsystem steuerbar sind und Störfaktoren (z), welche nicht bzw. kaum beeinflussbar sind (siehe Abb. 2). Eine Einstellung für alle Kontrollfaktoren wird im Folgenden als Systemkonfiguration, eine Belegung der Störfaktoren wird als Störkonfiguration bezeichnet.

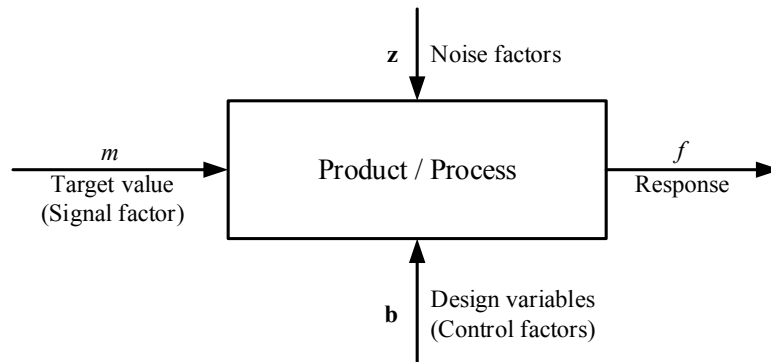


Abb. 2. Block diagram of a product/process [32].

Taguchi nutzt Verfahren der statistischen Versuchsplanung und entwickelte entsprechende Bewertungsfunktionen (so genannte Verlustfunktionen), welche den Qualitätsverluste auf Basis der Abweichung von einem gewünschten Wert berechnen [47]. Diese Sichtweise ist eine Möglichkeit der Varianzreduzierung, die im weiteren Sinne als Datenkompressionsmaßnahme angesehen werden kann, um den Einfluss von Störgrößen auf die Zuverlässigkeit des zugrundeliegenden Systems zu bewerten [2]. Taguchis Arbeiten haben einen starken Einfluss auf den Bereich der Robustheitsanalyse im Qualitätsmanagement und werden bis heute weiter erforscht und unter Statistikern intensiv diskutiert [3, 31]. In der Praxis findet die Methode bis heute Anwendung in verschiedensten Bereichen von den klassischen Ingenieurwissenschaften bis hin zur Biotechnologie [24, 35, 43].

Tab. 1. Verschiedene Verlustfunktionen [34]

Typ der (Qualitäts-) Verlustfunktion	Formel
Nominal das Beste (def. Zielwert)	$\bar{L} = k[\sigma^2 + (\bar{y} - \tau)^2]$
Je kleiner je besser	$\bar{L} = k[\bar{y}^2 + \sigma^2]$
Je größer je besser	$\bar{L} = k \left[\sum (1/y^2) \right] / n$

In Tab. 1 sind drei der häufiger angewendeten Verlustfunktionen aufgeführt. Hierbei ist \bar{L} der durchschnittliche Verlust für eine gegebene Systemkonfiguration über alle Störkonfiguration hinweg, k bezeichnet eine feste Konstante, die als *Qualitätsverlustkoeffizient* bezeichnet wird. Dieser Wert wird geschätzt und dient der Umrechnung auf einen Geldwert. \bar{y} und σ^2 stellen den Mittelwert und die Varianz des betrachteten Ergebniswerts für eine Systemkonfiguration dar (y stellt die Ausgabe eines einzelnen je-

des Experiments dar). Welche Verlustfunktion genutzt wird, hängt von der Charakteristik der betrachteten Ergebnisdaten ab. Die nominell beste Verlustfunktion zielt darauf ab, die Variabilität um einen gewünschten Zielwert zu verringern und damit Ausgangswerte oberhalb und unterhalb dieses Ziels zu sanktionieren, beispielsweise die erforderliche Ausgangsspannung einer elektrischen Schaltung. Die Funktion "Je kleiner je besser" zielt darauf ab, eine bestimmte Leistung, beispielsweise Kosten, Stress oder Energieverbrauch, zu minimieren. Schließlich wird die „je größer je besser“ Verlustfunktion verwendet, um einen Ausgabewert wie Zuverlässigkeit, Stärke oder Effizienz zu maximieren. Praxisbeispiele aus der Fertigungstechnik minimieren den Strahlungsverlust aus einem Mikrowellenherd bzw. maximieren die Haftfestigkeit eines Schweißpunkts [2, 34].

Eine umfassende tiefergehende Betrachtung zum Thema Taguchi-Methode und andere robuste Konstruktionskonzepte ist in Park et al. [32] zu finden.

Da Robustheitsentwurf und -analyse ein wiederkehrendes Thema in der Simulationmethodikforschung ist, wurde Taguchis Methode bereits von realen Experimenten auf Simulationsexperimente übertragen. Insbesondere Anwendungen im Bereich des Meta Modelling und/oder der Gefechtsfeldsimulation wurden in der Vergangenheit publiziert [5, 18, 37, 38].

In unserem im Folgenden skizzierten Ansatz wird die Robustheitsanalyse basierend auf den Verlustfunktionen von Taguchi mit einem Versuchsaufbau im großen Maßstab sowie mit visuell unterstützten Methoden zur Erkennung von Wissen für Fertigungssimulationen kombiniert.

3 Vorarbeit – Konzept zur Robustheitsverbesserung von Produktionssystemen

Der Nutzen der Kombination von Data Farming und Taguchis Robustheitsbewertungsmethode wurde im Kontext von Produktions- und Logistiksystemen bereits erfolgreich in [12, 42] angewendet. In diesen Publikationen bedeutet Robustheit des betrachteten Produktionssystems, dass sich dessen Leistung bzgl. einer Menge von definierten Zielgrößen, z.B. Ausbringungsmengen, Auslastungen von Stationen oder Puffern, möglichst stabil gegenüber Schwankungen im Produktmix als Störeinfluss verhält. Der genutzte Ansatz basiert hierbei auf der Nutzung der im Kapitel 2.2 vorgestellten Methode des Data Farming sowie der Methode von Taguchi (siehe Kapitel 2.3).

Der erste Schritt in der Methodik ist das Bestimmen der zu betrachtenden Stell- und Störgrößen (im Folgendem auch als Entscheidungsfaktoren bzw. Störfaktoren des Experimentdesigns bezeichnet), deren Art (kontinuierlich, diskret, kategorisch oder komplex) sowie deren Wertebereiche bzw. Ausprägungen. Stellgrößen sind hierbei auch in der Realität durch den Betreiber einstellbare Parameter. Störgrößen hingegen sind nicht oder kaum im realen System einstellbare, das System jedoch beeinflussende Umweltbedingungen.

Im folgendem zweiten Schritt werden zwei Experimentpläne, einer für die Stellgrößen und einer für die Störgrößen des Systems, erzeugt und anschließend gekreuzt. Dieses Vorgehen erlaubt eine von jeglichem Bias befreite Robustheitsanalyse.

Für die Entscheidungsfaktoren können die in der Data-Farming-Forschung gebräuchlichen Methoden zum Experimentdesign verwendet werden, beispielsweise der nearly orthogonal latin hypercube (NOLH), diese Methode gewährleistet ein sehr gutes Experimentdesign und ist hierbei deutlich effizienter als ein Standard- n^k -Design [51]. Eine typischerweise in Produktions- und Logistiksystemen betrachtete Störgröße ist der komplexe Faktor Produktmix. Der Produktmix stellt die Verteilung der verschiedenen Produkte bzw. Produktvariationen im Produktionsprogramm dar. Ein Experimentdesign für solch einen Produktmix ist deutlich schwieriger zu realisieren. Obwohl die Anzahl der Experimente für den Produktmix nicht exponentiell wächst, da die einzelnen Produktanteile nicht unabhängig voneinander sind (steigt der Anteil eines Produktes im Mix muss der Anteil anderer Produkte im gleichen Umfang sinken), wächst sie dennoch mit der zunehmenden Anzahl von Produkten im Mix sowie der gewählten Auflösung, d. h. der Schrittweite von 0% bis 100% [27]. Zum Beispiel würden bei einem vollfaktoriellen Design für 5 Produkte und einer angenommenen Schrittweite von 5% $(5 + (100/5))! / ((100/5)! (5 - 1)!) = 265.650$ Experimente nötig sein. Daher ist eine vollständige faktorielle Abdeckung des Produktmixes gerade bei steigender Zahl von Produkten offensichtlich nicht möglich.

Um die Anzahl der Experimente zu reduzieren, besteht ein einfacher Ansatz darin, eine Data-Farming-Entwurfsmethode wie NOLH zu verwenden und die Summe jeder Zeile auf eins zu normalisieren. Solch ein Vorgehen muss jedoch mit Vorsicht betrachtet werden, da die gewünschten Eigenschaften eines guten experimentellen Designs negativ beeinflusst werden können, z. B. kann sowohl die Ausgewogenheit der Faktorzerte nicht mehr gewährleistet werden. Zudem wird auch die Korrelationsarmut zwischen den Eingangsfaktoren (Orthogonalität) nicht mehr garantiert.

Die Implementierung von Einschränkungen, wie sie bei Produktmischen vorliegen, in space filling Designs wie z. B. NOHL ist eine enorm herausfordernde Aufgabe und ein laufendes Forschungsthema [15, 33]. Für den Zweck erster Machbarkeitsstudien soll jedoch der erwähnte Ansatz ausreichen.

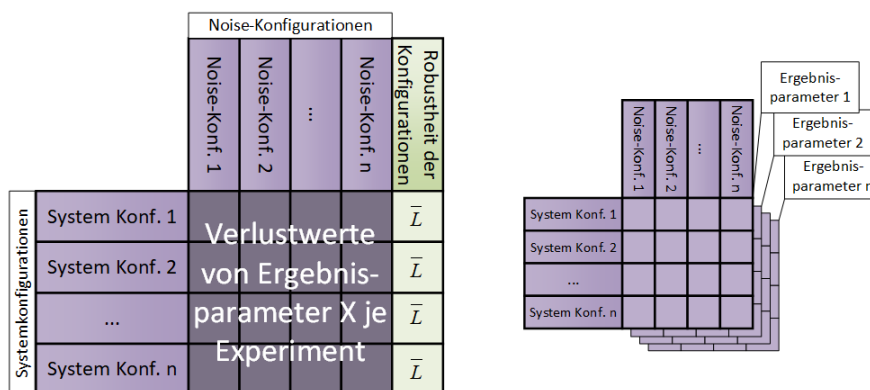


Abb. 3. Matrix des gekreuzten Experimentdesigns als Basis der Robustheitsanalyse [12].

In Schritt drei werden die erzeugten Experimentdesigns gekreuzt (crossed design), was zu einem endgültigen Experimentdesign mit jeder Kombination von Stell- und

Störkonfigurationen führt. Das Ergebnis ist mitunter ein sehr großer Experimentplan. Um eine realisierbare Laufzeit zu gewährleisten, müssen die Experimente parallel ausgeführt werden und/oder Hochleistungsrechner verwendet werden.

Der Vorteil des Ansatzes besteht darin, dass die Ergebnisdaten nach Durchführung der Experimente in matrixähnlichen Tabellen zusammengestellt werden können. Die Matrizen zeigen, wie sich jede Systemkonfiguration für jede Störfaktorkonfiguration bzw. Produktmischung verhält (siehe Abb. 3). Jede Zelle enthält den entsprechend der gewählten Verlustfunktion berechneten Robustheitswert eines ausgewählten Ergebnisparameters x . Wenn Replikationen durchgeführt wurden, wird jede Zelle mit dem durchschnittlichen Verlust aller Replikationen eines Experiments gefüllt. Für jede Zeile kann der durchschnittliche Verlust ermittelt werden, der die Robustheit jeder Systemkonfiguration darstellt. In komplexen Fertigungssystemen sind Betrachtungen der Performance aber auch der hier fokussierten Robustheit meist mehrkriteriell. Dies führt zwangsweise meist zur Betrachtung mehrerer Ergebnisparameter. Somit ergeben sich entsprechend mehrdimensionale Matrizen, siehe Abb. 3 rechts. Im Mittelpunkt des Interesses steht das Finden von Konfigurationen, die für viele, im besten Fall alle Ausgabeparametern robustes Verhalten aufweisen.

Vereinfacht könnte man die Robustheitswerte mehrerer Ergebnisparameter gewichten und eine Nutzwertanalyse durchführen. Somit wären die Robustheitswerte aller Ausgabeparameter zu einer einzigen Zahl zusammengefasst. Dieser Ansatz ist meist jedoch nicht zu empfehlen, da insbesondere das objektive Einstellen der Gewichtungsfaktoren oft nicht möglich ist. Daher wird für eine eingehendere Analyse, analog zu [11], die interaktive visuell geführte Analyse, unterstützt durch geeignetes Data Mining, empfohlen. Dieser Ansatz kann helfen, diejenigen Konfigurationen zu finden, die den günstigsten Kompromiss zwischen den verschiedenen Ausgabeparametern bieten. Hierzu wurde ein zweistufiger Prozess entwickelt, in dem zunächst die Systemkonfigurationen mithilfe von Clustering-Algorithmen in Klassen ähnlicher Robustheitsgruppen zusammengefasst werden. Diese Algorithmen kennzeichnen die Simulationsexperimente entsprechend ihrer Klassenzugehörigkeit. Experimente derselben Klasse weisen somit bzgl. den Robustheitsmaßen ähnliche Werte auf. Im zweiten Schritt können Verfahren des überwachten Lernens eingesetzt werden. Hierbei werden mathematische Modelle auf Basis der Entscheidungsfaktoren und der Klassifizierung der resultierenden Experimente trainiert. Diese Modelle geben zum einen Auskunft darüber, welche Entscheidungsfaktoren überhaupt Einfluss auf die Klassenzuordnung haben und zum anderen welche konkreten Ausprägungen der Entscheidungsfaktoren zu einer bestimmten Klassenzuordnung beitragen. Sind diese Abhängigkeiten bekannt, kann man Rückschlüsse darauf ziehen, wie das System robust gestaltet werden kann. Eine Fallstudie hierzu ist [11] zu entnehmen.

4 Herausforderung beim Transfer des Konzeptes auf komplexe Netzwerke

Der Transfer der in Kapitel 3 vorgestellten Methoden von Produktionssystemen auf nahezu beliebige weltweite/ komplexe Netze erscheint möglich und sehr sinnvoll. Die

Absicherung der Robustheit ist hierbei ein proaktives Vorgehen, das erlaubt, systemische Risiken im Vorhinein zu senken bzw. deren negativen Auswirkungen zu dämpfen. Die ungelöste Herausforderung für komplexe weltweite Netze besteht hierbei nun gerade darin, eine Methode zu entwickeln, die erlaubt, mit adäquaten Aufwand solche robuste Konfiguration zu finden. Beim Transfer der Methode zur Robustheitsanalyse im Kontext von Produktions- und Logistiksystemen auf komplexe weltweite Netze treten spezielle Herausforderungen, Forschungs- und Entwicklungsbedarfe auf.

Im ersten Schritt sind diese Forschungsbedarfe zu definieren. Eine erste Auswahl relevanter Forschungsfragen ist im Folgenden aufgeführt:

1. Wie kann eine Vorgehensweise zur Identifizierung relevanter Stellgrößen, externer Einflussgrößen und Ergebniswerte in großen Netzwerken gestaltet werden?
2. Wie können identifizierte Stellgrößen als Faktoren in Experimentdesigns abgebildet werden? Welche Eigenschaften und Besonderheiten sind zu beachten?
3. Welche Experimentdesignmethoden aus dem Kontext Data Farming können angewendet werden?
4. Wie können die komplexen Eigenschaften des Netzwerkes, z. B. die Netztopologie, als Faktoren in einer Data-Farming-Studie abgebildet werden?
5. Wie können externe Einflüsse als Störgrößen im Sinne der Robustheitsanalysen mittels Data Farming geeignet abgebildet werden?
6. Welche Algorithmen bzw. Methoden können, neben der Taguchi Methode, zur Bewertung der Robustheit von komplexen Netzen genutzt werden? Wie sind die Methoden zu parametrieren?
7. Welche Data-Mining-Methoden und Visualisierungstechniken können eingesetzt werden?
8. Was kann als Validierungsmodell eines komplexen Netzwerkes genutzt werden?
9. Wie können auch bei komplexen Netzwerken und großen Experimentplänen akzeptable Laufzeiten garantiert werden?

Zur Bearbeitung der Forschungsfragen ist eine vielfältige Kombination von Methoden aus verschiedenen Forschungsdisziplinen möglich. Denkbar ist beispielsweise zur effizienteren Ausführung der Modelle der Einsatz von Methoden der verteilten Simulation. Das Ausführen von großen Experimentplänen im Kontext der verteilten Simulation ist hierbei die zu lösende Aufgabe [45]. Weiterhin sind Methoden der Künstlichen-Intelligenz-Forschung für einige der Forschungsfragen zu überprüfen. Beispielsweise ist das Nutzen generativer Neuronaler Netze zur Erstellung von guten Experimentplänen denkbar. Der Einsatz von so genannten Generative Adversarial Networks zur Optimierung der Robustheit könnte ein weiterer möglicher Forschungsansatz sein [16]. Auch die bereits angesprochene Forschungsdisziplin des Visual Analytics und des Data Mining sind einzubeziehen und auf die Analyse von Simulationsdaten aus komplexen Netzwerken auszuweiten [17, 22].

5 Fazit und Ausblick

Die Beherrschung komplexer Netzwerke, wie Rechner- und Energienetze oder auch Supply-Netzwerke wird im täglichen Lebens immer wichtiger. Eine proaktive Verbesserung der Netzwerke durch eine robuste Konfiguration kann die Beherrschbarkeit solcher, mitunter kritischer Systeme, deutlich steigern.

Die Adaption der im Beitrag vorgestellten Methode zur Robustheitsanalyse von Produktionssystemen auf komplexe Netzwerke ist hierbei ein möglicher Schritt. Die hierbei offenen Forschungsbedarfe wurden ausschnittsweise vorgestellt.

Literaturverzeichnis

1. Acha E (2004) *FACTS. Modelling and simulation in power networks*. Wiley, Chichester, Hoboken, NJ
2. Ben-Gal I (2005) On the Use of Data Compression Measures to Analyze Robust Designs. *IEEE Trans. Rel.* 54(3): 381–388. doi: 10.1109/TR.2005.853280
3. Box G (1988) Signal-to-Noise Ratios, Performance Criteria, and Transformations. *Technometrics* 30(1): 1. doi: 10.2307/1270311
4. Breslau L, Estrin D, Fall K et al. (2000) Advances in network simulation. *Computer* 33(5): 59–67. doi: 10.1109/2.841785
5. Dellino G, Kleijnen JPC, Meloni C (2009) Robust Simulation-Optimization using Metamodels. In: Rossetti MD, Hill RR, Johansson B et al. (eds) *Proceedings of the 2009 Winter Simulation Conference (WSC 2009)*. (WSC 2009) : Austin, Texas : 13-16 December 2009. IEEE Inc, Piscataway, N.J., pp 540–550
6. Elmegreen BG, Sanchez SM, Szalay AS (2014) The Future of Computerized Decision Making. In: Tolk A, Diallo SD, Ryzhov IO et al. (eds) *Proceedings of the 2014 Winter Simulation Conference*. IEEE Inc, Piscataway, N.J., pp 943–949
7. Epstein JM (2006) *Generative social science. Studies in agent-based computational modeling*. Princeton studies in complexity. Princeton University Press, Princeton
8. Feldkamp N, Bergmann S, Strassburger S (2015) Knowledge Discovery in Manufacturing Simulations. In: Taylor SJE, Mustafee N, Son Y-J (eds) *Proceedings of the 3rd ACM SIGSIM Conference on Principles of Advanced Discrete Simulation*. ACM, New York, NY, USA, pp 3–12
9. Feldkamp N, Bergmann S, Strassburger S (2015) Visual Analytics of Manufacturing Simulation Data. In: Yilmaz L, Chan WKV, Moon I et al. (eds) *Proceedings of the 2015 Winter Simulation Conference*. IEEE Inc, Piscataway, N.J., pp 779–790
10. Feldkamp N, Bergmann S, Strassburger S (2016) Innovative Analyse- und Visualisierungsmethoden für Simulationsdaten. In: Nissen V, Stelzer D, Straßburger S et al. (eds) *Multikonferenz Wirtschaftsinformatik (MKWI) 2016*. Technische Universität Ilmenau 09. - 11. März 2016, vol 3. TU Ilmenau Universitätsbibliothek, Ilmenau, pp 1737–1748

11. Feldkamp N, Bergmann S, Strassburger S et al. (2016) Knowledge Discovery in Simulation Data: A Case Study of a Gold Mining Facility. In: Roeder TMK, Frazier PI, Szechtman R et al. (eds) Proceedings of the 2016 Winter Simulation Conference. IEEE Inc, Piscataway, N.J., pp 1607–1618
12. Feldkamp N, Bergmann S, Strassburger S et al. (2017) Knowledge Discovery and Robustness Analysis in Manufacturing Simulations. In: Chan V, D’Ambrogio A, Zacharewicz G et al. (eds) Proceedings of the 2017 Winter Simulation Conference. IEEE Inc
13. Feldkamp N, Bergmann S, Strassburger S et al. (2017) Knowledge Discovery in Simulation Data – a Case Study for a Backhoe Assembly Line. In: Chan V, D’Ambrogio A, Zacharewicz G et al. (eds) Proceedings of the 2017 Winter Simulation Conference. IEEE Inc, pp 4456–4458
14. Feldkamp N, Bergmann S, Strassburger S et al. (2018) Combining Data Farming And Data Envelopment Analysis For Measuring Productive Efficiency In Manufacturing Simulations. In: Rabe M, Juan AA, Mustafee N et al. (eds) Proceedings of the 2018 Winter Simulation Conference. IEEE inc., Piscataway, N.J.
15. Golchi S, Loepky JL (2016) Space Filling Designs for Constrained Domains. In: Aggarwal M, George OE (eds) International Conference on Design of Experiments, pp 1–28
16. Goodfellow I, Bengio Y, Courville A (2016) Deep learning. MIT Press, Cambridge, Massachusetts, London, England
17. Han J, Kamber M, Pei J (2012) Data Mining Concepts and Techniques, 3. ed. The Morgan Kaufmann series in data management systems. Elsevier/Morgan Kaufmann, Amsterdam
18. Horne G, Åkesson B, Meyer T et al. (2014) Data farming in support of NATO. Final Report of Task Group MSG-088. STO technical report, TR-MSG-088. North Atlantic Treaty Organisation, Neuilly-sur-Seine Cedex
19. Horne GE, Meyer TE (2005) Data Farming: Discovering Surprise. In: Kuhl ME, Steiger NM, Armstrong FB et al. (eds) Proceedings of the 2005 Winter Simulation Conference. IEEE Inc, Piscataway, N.J., pp 1082–1087
20. Horne GE, Meyer T (2010) Data farming and defense applications. In: Armstrong R, McNamara J, Pinelli TE (eds) MODSIM World Conference and Expo. Langley Research Center, Hampton, VA, pp 74–82
21. Janis Beenen (2018) Niedrigwasser lässt einzelne Tankstellen leerlaufen. Süddeutsche Zeitung 2018(7.11.2018)
22. Keim DA, Mansmann F, Schneidewind J et al. (2008) Visual Analytics: Scope and Challenges. In: Simoff S, Boehlen MH, Mazeika A (eds) Visual Data Mining: Theory, Techniques and Tools for Visual Analytics, 1st edn. Springer, Berlin, Heidelberg
23. Kleijnen JPC, Sanchez SM, Lucas TW et al. (2005) State-of-the-Art Review: A User’s Guide to the Brave New World of Designing Simulation Experiments. INFORMS Journal on Computing 17(3): 263–289
24. Konduk BA, Ucisik AH (1999) Determination of Primary Parameters relevant to the Adequacy of Haemodialysis through Taguchi Method. In: Blanchard S,

- Eckstein E, Fouke J (eds) Proceedings of the first Joint BMES/EMBS Conference, p 627
25. Krückhans B, Meier H (2013) Industrie 4.0 – Handlungsfelder der Digitalen Fabrik zur Optimierung der Ressourceneffizienz in der Produktion. In: Dangelmaier W, Laroque C, Klaas A (eds) Proceeding der 15. ASIM Fachtagung Simulation in Produktion und Logistik 2013. Entscheidungsunterstützung von der Planung bis zur Steuerung. HNI-Verlagsschriftenreihe, Paderborn, pp 31–40
 26. Law AM (2014) Simulation Modeling and Analysis, 5th edn. McGraw-Hill Series in Industrial Engineering and Management Science. McGraw Hill Book Co, New York, N.Y.
 27. Ledi T, Spagon P, del Castillo E et al. (2013) e-Handbook of Statistical Methods. 5. Process Improvement. <http://www.itl.nist.gov/div898/handbook/pri/pri.htm>
 28. Lorig F, Timm IJ (2014) How to Model the "Human Factor" for Agent-based Simulation in Social Media Analysis? Work in Progress Paper. In: Proceedings of the 2014 Symposium on Agent Directed Simulation. SCS, San Diego, CA, USA
 29. Lorig F, Becker CA, Timm IJ (2017) Formal Specification of Hypotheses for Assisting Computer Simulation Studies. In: Proceedings of the Symposium on Theory of Modeling & Simulation. Society for Computer Simulation International, San Diego, CA, USA
 30. Lorig F, Leberherz DS, Berndt JO et al. (2017) HYPothesis-driven Experiment Design in Computer Simulation Studies. In: Chan V, D'Ambrogio A, Zacharewicz G et al. (eds) Proceedings of the 2017 Winter Simulation Conference. IEEE Inc
 31. Nair VN, Abraham B, MacKay J et al. (1992) Taguchi's Parameter Design. A Panel Discussion. *Technometrics* 34(2): 127. doi: 10.2307/1269231
 32. Park G-J, Lee T-H, Lee KH et al. (2006) Robust Design. An Overview. *AIAA Journal* 44(1): 181–191. doi: 10.2514/1.13639
 33. Petelet M, Looss B, Asserin O et al. (2010) Latin hypercube sampling with inequality constraints. *AStA Adv Stat Anal* 94(4): 325–339. doi: 10.1007/s10182-010-0144-z
 34. Phadke MS (1989) Quality engineering using robust design. Prentice Hall, Englewood Cliffs, N.J.
 35. Rao RS, Kumar CG, Prakasham RS et al. (2008) The Taguchi Methodology as a Statistical Tool for Biotechnological Applications: A Critical Appraisal. *Biotechnol J* 3(4): 510–523. doi: 10.1002/biot.200700201
 36. Rodermund SC, Lorig F, Berndt JO et al. An Agent Architecture for Simulating Communication Dynamics in Social Media. In: Multiagent System Technologies
 37. Sanchez SM (2000) Robust design: seeking the best of all possible worlds. In: Joines JA, Barton RR, Kang K et al. (eds) Proceedings of the 2000 Winter Simulation Conference. IEEE Inc, Piscataway, N.J., pp 69–76

38. Sanchez SM (1994) A Robust Design Tutorial. In: Tew JD, Manivannan S, Sadowski DA et al. (eds) Proceedings of the 1994 Winter Simulation Conference, pp 106–113
39. Sanchez SM (2007) Work Smarter, Not Harder: Guidelines for Designing Simulation Experiments. In: Henderson SG, Biller B, Hsieh M-H et al. (eds) Proceedings of the 2007 Winter Simulation Conference. December 9 - 12, 2007, Washington, DC, U.S.A. IEEE, Piscataway, N.J., pp 84–94
40. Sanchez SM (2014) Simulation Experiments: Better Data, Not Just Big Data. In: Tolk A, Diallo SD, Ryzhov IO et al. (eds) Proceedings of the 2014 Winter Simulation Conference. IEEE Inc, Piscataway, N.J., pp 805–816
41. Sanchez SM, Wan H (2009) Better than a petaflop: The power of efficient experimental design. In: Rossetti MD, Hill RR, Johansson B et al. (eds) Proceedings of the 2009 Winter Simulation Conference (WSC 2009). (WSC 2009) : Austin, Texas : 13-16 December 2009. IEEE Inc, Piscataway, N.J., pp 60–74
42. Schulze T, Feldkamp N, Bergmann S et al. (2018) Data Farming und simulationsbasierte Robustheitsanalyse für Fertigungssysteme. In: Deatcu C, Schramm T, Zobel K (eds) Tagungsband ASIM 2018 – 24. Symposium Simulationstechnik, pp 243–251
43. Song J, Dong F, Zhao J et al. (2017) Optimal Design of Permanent Magnet Linear Synchronous Motors based on Taguchi Method. IET Electric Power Applications 11(1): 41–48. doi: 10.1049/iet-epa.2016.0164
44. Stefanovic D, Stefanovic N, Radenkovic B (2009) Supply network modelling and simulation methodology. Simulation Modelling Practice and Theory 17(4): 743–766. doi: 10.1016/j.simpat.2009.01.001
45. Strassburger S, Schulze T, Fujimoto R (2009) Future trends in distributed simulation and distributed virtual environments. In: Alexopoulos C, Goldsman D, Wilson JR (eds) Advancing the frontiers of simulation: a Festschrift in honor of George Samuel Fishman. Springer, Heidelberg, pp 231–261
46. Straßburger S, Bergmann S, Feldkamp N et al. (2018) Data Farming Research Project with Audi and VW. In: Siemens AG (ed) 2018 Plant Simulation Worldwide User Conference
47. Taguchi G (1988) System of Experimental Design, 3rd edn. Unipub, White Plains, New York
48. Taguchi G (1995) Quality engineering (Taguchi methods) for the development of electronic circuit technology. IEEE Trans. Rel. 44(2): 225–229. doi: 10.1109/24.387375
49. Thomas JJ, Cook KA (2005) Illuminating the Path. Research and Development Agenda for Visual Analytics, 1st edn. IEEE Computer Society, Los Alamitos, California
50. Verein Deutscher Ingenieure (2014) Simulation of Systems in Materials Handling, Logistics and Production - Fundamentals VDI 3633-1(VDI 3633-1)
51. Vieira H, Sanchez SM, Kienitz KH et al. (2011) Improved efficient, nearly orthogonal, nearly balanced mixed designs. In: Jain S, Creasey R, Himmelspach J et al. (eds) Proceedings of the 2011 Winter Simulation Conference. IEEE Inc, Piscataway, NJ, pp 3600–3611