

Farming for Mining – Entscheidungsunterstützung mittels Simulation im Supply Chain Management

Farming for Mining – Decision Support by Simulation in Supply Chain Management

Markus Rabe, Anne Antonia Scheidler, TU Dortmund, Dortmund (Germany),
markus.rabe@tu-dortmund.de, anne-antonia.scheidler@tu-dortmund.de

Abstract: This paper constitutes an approach for combining Simulation and Knowledge Discovery techniques in the field of Supply Chain support and demonstrates relationships among them. We discuss how simulation and Knowledge Discovery will be applied as a part of data farming. In view of the above, we give an insight into first experiment setups using a prototypical implementation. In the experiment, the prototype for the Plant Simulation model is introduced and first results of the Supply Chain simulation scenarios are presented. We demonstrate how to transform the farmed model output into a transactional database which serves as a basis for the Knowledge Discovery process.

1 Einleitung

Die vielfältigen Aufgaben in der Supply Chain (SC) werden heutzutage oftmals mit Unterstützung des Knowledge Discovery in Databases (KDD) gelöst (Gabriel et al. 2009). Zu den Aufgaben zählen insbesondere auch Kernkompetenzen des Supply Chain Managements (SCMs), wie beispielsweise der Decision Support. Typische Fragestellungen des SCMs sind beispielsweise wie eine fristgerechte Lieferung sichergestellt werden kann, welche Transportmittel unter welchen Bedingungen am effizientesten sind oder wo und wann Komponenten in der SC benötigt werden. Betrachtet man diese Fragestellungen des SCMs, so sind insbesondere Fragen nach Zusammenhängen von großer Bedeutung. In diesem Umfeld können Techniken der Wissensentdeckung sinnvoll eingesetzt werden, um die Daten der SC-Systemlandschaft zu untersuchen. Da Techniken der Wissensentdeckung spezifische Anforderungen an die Datenbasis stellen (z. B. Stichprobenumfang, vorliegende Datenformate und eingehende Datenqualität) und auch prozessbegleitende Verifikation und Validierung nicht Teil der Wissensentdeckungsmodelle sind (Düsing 2010), genügt dieser Ansatz den heutigen Anforderungen der Industrie nicht. Um die vorhandenen Verfahrensschwächen aufzulösen, werden in diesem Beitrag Techniken der Wissensentdeckung mit Simulation kombiniert.

2 Supply Chain, Wissensentdeckung und Simulation - aktuelle Problemfelder

Nachfrageschwankungen sind ein hochaktuelles Thema in der SC und zählen zu den größten Herausforderungen im SCM (Capgemini 2011). In Zeiten globaler SCs sind Nachfrageschwankungen selten isolierte Effekte; vielmehr gibt es Korrelationen zwischen unterschiedlichen Effekten. So kann eine erhöhte Nachfrage eines Artikels x mit erhöhten Nachfragen der Artikel y und z korrelieren, obwohl die Artikelgruppen von x , y und z nicht in betriebswirtschaftlichen Beziehungen wie Produktgruppen zueinander stehen. Erhöhte Nachfragen können in der SC zu unterschiedlichen Problemen führen, beispielhaft wären das zunächst erhöhte Transportaufkommen und der darauf folgende Lieferengpass.

Die zugrundeliegenden Daten, in denen Korrelationen zu entdecken sind, sind schwer zu kategorisieren. Dies liegt zum einen an der Vielzahl der beteiligten Systeme und somit an den redundanten Dateninformationen, stark aggregierten Daten sowie heterogenen Datenformaten. Erschwerend wirkt sich zudem aus, dass die betriebswirtschaftliche Sicht auf die Daten und somit auch eine Vielzahl der zugehörigen Systemdokumentation nicht die Datenstrukturen und die möglichen Operationen auf den Datenstrukturen wiedergeben. Die Daten der Basissysteme, wie beispielsweise Datenbanken, dienen jedoch als Eingabedaten für Wissensentdeckungstechniken. Ein Großteil der SC-Daten sind Transaktionsdaten, auch als Bewegungsdaten bezeichnet, die Zustandsänderungen mittels Mengenangaben und Zeitstempeln beschreiben. Diese bilden die Grundlagen für die Wissensentdeckungen, beispielsweise von Wissen über Korrelationen, da die Transaktionen für die Daten- und Managementprozesse von grundlegender Bedeutung sind (Anane et al. 2002).

Die beschriebenen Korrelationen in der SC lassen sich mit Techniken des KDDs, einem Prozessmodell zur Wissensentdeckung (Fayyad et al. 1996), aufdecken. Hierbei ist der zentrale Prozessschritt - das Data Mining - so bedeutend, dass heutzutage die Begriffe KDD und Data Mining oftmals synonym verwendet werden (Adriaans und Zantinge 1996). Data Mining deckt Aufgabenstellungen aus den Bereichen der Segmentierung, Klassifizierung, Prognose oder Abhängigkeitsanalyse ab (Petersohn 2005). In der Logistik wurden Problemstellungen aus der Kundensegmentierung, Prognosen im Bereich Controlling, Materialbedarfsplanung und große Bereiche der Finanzdienstleistungen bereits mittels des Einsatzes von Data Mining gelöst (Gabriel et al. 2009). Auch Problemstellungen der SC, wie die bereits erwähnten korrelierten Nachfrageschwankungen, aber auch Fragestellungen nach Gruppierungen von Artikeln oder Transportfahrzeugen, profitieren von der Unterstützung des Data Minings (Rabe und Scheidler 2014). So lässt sich eine korrelierte Nachfrageschwankung mit Regellernern aufdecken, und verschiedene Clusteralgorithmen gruppieren Artikel, Fahrzeuge oder auch Kunden und Lieferanten (Düsing 2010).

Der wesentliche Unterschied von KDD und traditionell-logistischen Verfahren, wie der Clusteranalyse zur Set-Bildung bei Artikeln, liegt in der Mächtigkeit der Verfahren. Während logistische Verfahren im konkreten Beispiel reine Artikeleigenschaften zur Gruppenbildung heranziehen, kann sich das KDD aller vorhandenen Kennzahlen bedienen, die im Zeitalter von Big Data vollautomatisch von den logistischen Systemen protokolliert werden (Mayer-Schönberger und Cukier 2013).

Aus dem entdeckten Wissen des KDDs kann das SCM Handlungsanweisungen, beispielsweise für das SC-Bestellwesen, ableiten. Das Bestellwesen kann im konkreten Fall der Artikelkorrelation eine Clusterung der betroffenen Artikel bilden. Eine Alternative wäre eine Clusterung der artikelproduzierenden Unternehmen, um ein günstiges Transport- und Lagernetzwerk zu entwickeln. Diese Cluster können jedoch im Regelfall nicht ohne Veränderungen in die bestehende SC-Landschaft integriert werden.

Veränderungen können in einer SC aufgrund ihrer emergenten Eigenschaften und inhärenten Wechselwirkungen in der Regel nicht partiell durchgeführt werden. Um die Auswirkungen der unterschiedlichen Wissensintegrationsmöglichkeiten auf Produktion, Lagerhaltung und Distribution zu testen, bedient sich die Logistik der ereignisdiskreten Simulation (Discrete Event Simulation, DES), da es sich bei ihr um eine sehr effiziente Methode zur Handhabung zeitvarianter Systeme handelt (Hellström und Johnsson 2002). Bei dem Aufbau des Simulationsmodells spielt die Frage nach dessen Zweck eine herausragende Rolle (Robinson, 1994). Denn nur unter Berücksichtigung des „Warum?“ erhält man geeignete Simulationsergebnisse, die den Einsatz von Simulation rechtfertigen (Rabe et al. 2008).

Am Anfang ist es sinnvoll, das KDD auf Echtdateen durchzuführen, um eine konkrete Hypothese für die Simulation zu entwickeln. Jedoch können Echtdateen häufig nicht in ausreichender Menge oder angemessener Qualität zur Verfügung gestellt werden. Gründe hierfür sind mangelnde Datenqualität oder Sicherheitsbestimmungen in Unternehmen (Hildebrand 2011). Zusätzlich erschwert eine komplexe Systemlandschaft heterogener Systeme mit redundanter Datenhaltung eine zweckgebundene Datenextraktion. Unter diesen Bedingungen muss die Ausgangshypothese durch Expertenwissen oder andere Quellen gebildet werden, da ohne adäquate Datenbasis kein Anwendungsfall für das KDD gegeben ist. In beiden Fällen ist die Bildung der Starthypothese nur ein Spezialfall des nachfolgenden Prozesses aus DES und KDD, der im nachfolgenden Kapitel explizit beschrieben wird.

3 Verbindung DES und KDD

Die Veränderungen der SC werden in der Praxis des SCMs als „What-if-Szenarien“ simuliert (Konrad und Stache 2005). Ein „What-if-Szenario“ der Simulation produziert als Simulationsergebnis große Datenmengen. Der Unterschied dieser Datenmenge zu den Echtdateen ist, dass hier die Ausgangssituation (Simulationsaufbau und Parameter) bekannt und kontrollierbar ist. Veränderungen in der Ausgangssituation können gezielt vorgenommen werden, um Veränderungen in der Datenbasis zu produzieren. Dies ist ein Vorgang, der in der Literatur auch im Kontext des Data Farmings bekannt ist (Horne 1997). Der Begriff stammt von der Idee des bewussten „Kultivierens“ von Daten mittels Simulationsszenarien und dem nachfolgenden „Ernten“ der Daten aus Simulationsläufen, oder wie Sanchez (2014) sagt: „They grow data from their models“. Das Ergebnis des Data Farmings ist eine Datenlandschaft, die mittels verschiedener Techniken analysiert werden kann. Horne (2005) nennt als mögliche Techniken Data Mining, Data Analysis und Data Visualization. Diese Techniken sind jedoch nicht ohne Zwischenschritte oder Transformationen auf ein Simulationsergebnis anwendbar. Rabe und Scheidler (2014) haben gezeigt, dass insbesondere die Vorverarbeitung der Daten, die ein essentieller Teil der Wissensentdeckung ist, auch in der Verbindung von Simulation und KDD

von grundlegender Bedeutung ist. Daher wird in diesem Beitrag die Verbindung der Simulation zum KDD, also dem Prozessmodell der Wissensentdeckung (Fayyad et al. 1996), gesetzt. Die erzeugte Datenlandschaft ist durch zwei Dimensionen gekennzeichnet. Zum einen die Dimension Volumen, die sich aus der Anzahl der einzelnen Replikationen ergibt. Zum anderen die Dimension Komplexität, die sich im Kontext von Datenmengen auch als Varietät in der Literatur finden lässt (Che et al. 2013). Die Komplexität der Datenlandschaft wird durch den Aufbau der Simulation bestimmt. Die erzeugte Datenlandschaft kann mittels des KDDs analysiert werden, um zu prüfen, ob durch die Umstrukturierung der SC-Landschaft neue, eventuell ungewollte, Korrelationen entstanden sind (Farming for Mining).

Im konkreten Fall der Clusterung des Materials kann ein Simulationsexperiment die unterschiedlichen Neustrukturierungen von Lagerzonen und Transportmodalitäten simulieren. Auf den so gewonnenen Ergebnisdaten kann mittels der Assoziationsanalyse nach neuen Korrelationen, beispielsweise mit weiteren Artikelgruppen, gesucht werden. Eine Bewertung der neuen Korrelation oder eine Ursachenanalyse ist jedoch durch eine Assoziationsanalyse nicht möglich. Die aus dem KDD gewonnenen Erkenntnisse können über neue Hypothesen in einem iterativen Prozess in die Simulation zurückgeführt werden. Hier führt die Hypothese zu einer Veränderung des Simulationsaufbaus und der Simulationsparameter. Dadurch erfolgt durch mehrere Simulationsläufe aus der ursprünglichen Ist-SC ein Vorschlag für die umgestaltete Soll-SC (Abb. 1).

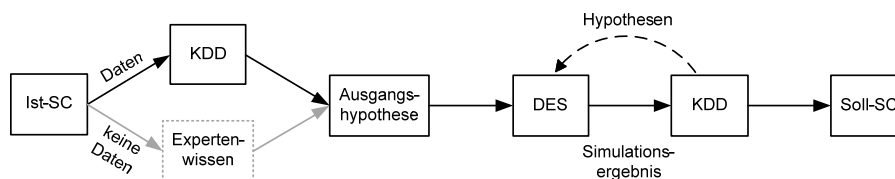


Abbildung 1: KDD und Simulation

4 Das Simulationsmodell für Transaktionsdaten

Das vorliegende Simulationsmodell basiert auf der Arbeit von Arndt (2014) und ist am Fachgebiet „IT in Produktion und Logistik“ entwickelt worden. Das Modell ist in Plant Simulation implementiert worden. Plant Simulation wird von Siemens Industry Software GmbH & Co. KG vertrieben. Die Simulationssoftware, die Einsatz im Umfeld der digitalen Fabrik findet, gestattet dem Benutzer Prozesse aus den unterschiedlichen Bereichen zu modellieren und zu simulieren. Das Leistungsspektrum umfasst die Abbildung der Materialflussteuerung, Visualisierungen in 2D und 3D, Bausteine und Klassenbibliotheken sowie Analysetools (Bangsow 2010).

Abbildung 2 zeigt einen einfachen Ausschnitt des Plant Simulation SC-Prototyps mit Lieferanten und Kunden. Hierbei liegt der fachliche Fokus auf der unternehmensübergreifenden Modellierung einer globalen SC. Das Simulationsmodell dient zu Generation von Transaktionsdaten einer globalen Automobilwertschöpfungskette. Die Eingabegrößen des Modells sind auf der linken Seite dargestellt und umfassen Wareneingang, Lieferdauer und Lieferabweichung. Die Ausgabegrößen sind Lieferdatum (LD), Soll-Wareneingangsdatum (Soll-WE). Ist-Wareneingangsdatum

(Ist-WE) sowie die Anzahl der Wareneingänge, die momentan auf einen fixen Wert eingestellt sind. Die Simulation startet mit dem Warenausgang eines Lieferanten, der die bewegten Elemente (BE) in erster Annäherung mit normalverteilterm Abstand erzeugt. Ein erzeugtes BE wird an die nachfolgenden Stationen verteilt. Die Ankunft eines BEs an einer nachfolgenden Station symbolisiert das Eingangsdatum einer Lieferung. In Stationen können BEs eine bestimmte Zeit warten, was die Lieferdauer einer Sendung nachstellt. Die Lieferdauer ist in das Modell als diskrete, empirische Verteilung eingebunden. Zuletzt werden Sendungen nach diversen Bearbeitungszeiten, Zwischenlagern und Lieferungen zur Senke transportiert. Die Senke stellt den tatsächlichen Wareneingang beim Original Equipment Manufacturer (OEM) dar.

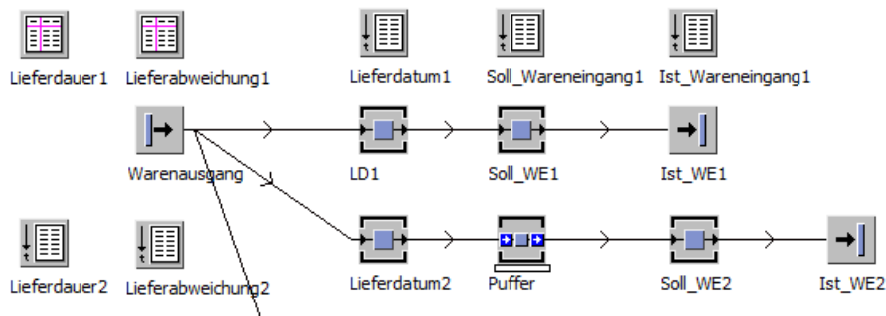


Abbildung 2: Ausschnitt des SC-Modells (Plant Simulation)

Technisch gesehen wird nach Eintreffen des BEs in der Senke die BE vernichtet und der Liefervorgang ist abgeschlossen. Die Dokumentation der Simulationsausgabedaten erfolgt über Zeitleisten in Tabellenform die aus Werten (1 für Ankunft, 0 für Ausgang) und Zeitstempeln bestehen. Die Zeitstempel sind, wie auch der Zeitanteil von Soll-WE und Ist-WE, vom Datentyp „DateTime“. In der Tabelle wird für jedes ankommende BE in den entsprechenden Stationen ein Eintrag erstellt und das Ergebnis einer Replikation ist für jede relevante Station eine Tabelle von Ausgabedaten. Tabelle 1 zeigt einen Ausschnitt einer aggregierten Sicht aus Warenausgang, Soll-WE und Ist-WE eines Lieferpfades. Die Tabelle wurde um eine alphanumerische ID angereichert. Der alphanumerische Teil der ID stellt die Replikation dar, den numerische Teil einer BE-Aktion in der jeweiligen Replikation.

Tabelle 1 zeigt vier Transaktionsdatensätze (A1 bis A4) und ihre Zeitstempel. Diese Tabelle kann, mit leichter Modifikation der ID zu Surrogatschlüsseln, in eine relationale Transaktionsdatenbank überführt werden. Die Transaktionsdatenbank dient als Quelle für KDD-Verfahren. Je nach eingesetztem Verfahren ist zu prüfen, ob die Daten im Anschluss an die Simulation für die Weiterverarbeitung durch das KDD in andere Darstellungsformen transformiert werden müssen. Alternativ besteht die Möglichkeit, die benötigten Datenformate für das KDD direkt in der Simulation festzulegen. Beide Vorgehensweisen haben Vor- und Nachteile, die wir abschließend an dem konkreten Praxisbeispiel des vorliegenden Simulationsmodells diskutieren wollen.

Tabelle 1: Transaktionsdaten Plant Simulation (Auszug)

ID	Warenausgang		Soll-WE		Ist-WE	
	Zeit	Wert	Zeit	Wert	Zeit	Wert
A1	12.02.2015 00:00	1	12.02.2015 00:00	1	13.02.2015 00:00	1
A1	12.02.2015 00:00	0	13.02.2015 00:00	0	13.02.2015 00:00	0
A2	26.03.2015 16:42	1	27.03.2015 16:00	1	27.03.2015 16:44	1
A2	27.03.2015 16:42	0	27.03.2015 16:42	0	27.03.2015 16:44	0
A3	25.04.2015 15:56	1	25.04.2015 15:56	1	02.05.2015 15:56	1
A3	25.04.2015 15:56	0	02.05.2015 15:56	0	02.05.2015 16:26	0
A4	14.05.2015 23:25	1	17.05.2015 23:25	1	18.05.2015 23:25	1
A4	17.05.2015 23:25	0	18.05.2015 23:55	0	18.05.2015 23:25	0

Als Beispiel soll die Erweiterung der Transaktionen im Simulationsmodell um Mengenattribute diskutiert werden. Tabelle 2 zeigt vier Transaktionen mit den Mengenangaben der Lieferkette.

Tabelle 2: Alternative A - Transaktionsdaten mit Mengenattributen (Auszug)

ID	Warenausgang		Soll-WE		Ist-WE	
	Zeit	Menge	Zeit	Menge	Zeit	Menge
A1	12.02.2015 00:00	100	12.02.2015 00:00	100	13.02.2015 00:00	100
A2	12.02.2015 00:00	105	13.02.2015 00:00	105	13.02.2015 00:00	105
A3	26.03.2015 16:42	150	27.03.2015 16:00	160	27.03.2015 16:44	150
A4	27.03.2015 16:42	2000	27.03.2015 16:42	2000	27.03.2015 16:44	2000

Eine Vielzahl von KDD-Verfahren kann die numerischen Mengenangaben nicht verarbeiten. Alternative A besteht in einer Modifikation des Simulationsmodells. Die Mengenangaben können in der Simulation durch Container (z.B. M1 – M4) in geeigneter Abstraktionsstufe ersetzt werden (Tab. 3).

Tabelle 3: Alternative B - Transaktionsdaten mit Containern (Auszug)

ID	Warenausgang		Soll-WE		Ist-WE	
	Zeit	Menge	Zeit	Menge	Zeit	Menge
A1	12.02.2015 00:00	M1	12.02.2015 00:00	M1	13.02.2015 00:00	M1
A2	12.02.2015 00:00	M1	13.02.2015 00:00	M1	13.02.2015 00:00	M1
A3	26.03.2015 16:42	M1	27.03.2015 16:00	M2	27.03.2015 16:44	M1
A4	27.03.2015 16:42	M4	27.03.2015 16:42	M4	27.03.2015 16:44	M4

Die Alternative A besteht in einer nachträglichen Diskretisierung der numerischen Attribute in eine benutzerspezifische Anzahl von Containern mit geeigneter Zuordnungsfunktion (z.B. nach Häufigkeitsverteilung der Werte). Alternative B erspart mögliche Transformationsschritte vor den KDD-Verfahren, was sich positiv auf die Gesamtdauer des Verfahrens auswirkt. Allerdings ist die erzeugte Datenlandschaft spezifisch auf das Verfahren angepasst und kleine Verfahrensänderungen erfordern mitunter größere Anpassungen im Simulationsmodell. Alternative A ist im Durchschnitt zeitaufwendiger als Alternative B, da die nachträglichen Transformationen zusätzliche Zeit in der Prozesskette benötigen. Der Vorteil bei Alternative A besteht jedoch in der Flexibilität, denn bei Veränderungen im KDD-Bereich müssen nur die Transformationsschritte adaptiert werden. Dies ist zumeist schneller und kostengünstiger zu bewerkstelligen als Anpassungen im Simulationsmodell.

5 Zusammenfassung und Ausblick

Der vorliegende Beitrag diskutiert die Grundlagen der Entscheidungsunterstützung mittels Simulation im Umfeld des SCMs. Hierzu wurden die aktuellen wissenschaftlichen Grundlagen und Problemstellungen beleuchtet und die beteiligten Themengebiete miteinander verknüpft. Es wurde gezeigt, wie Daten generell mittels eines Data-Farming-Simulationsmodells in Plant Simulation erzeugt werden können. Im Anschluss wurde demonstriert, wie diese Daten aus dem Simulationsmodell (Simulationsergebnisse) exportiert werden können. Die damit verbundenen grundlegenden fachlichen Überlegungen wurden an zwei Alternativen für den Import in die KDD-Umgebung diskutiert. Als Ergebnis der Untersuchungen, wurde festgestellt, dass bei dem Experimentaufbau (vgl. Abb. 1) eine Entscheidung zwischen Prozessgeschwindigkeit (Alternative B) und Prozessflexibilität (Alternative A) zu treffen ist. Beide Varianten zeigen, wie die prinzipielle Überführung der Simulationsergebnisse in das KDD durchzuführen ist. Ebenso wird dargelegt, welche Konfigurationsmöglichkeiten dem Prozessmodellierer zur Verfügung stehen.

Die kommenden Forschungsarbeiten umfassen zum einen weitere Untersuchungen an der Schnittstelle von Simulation und KDD im SC-Kontext, um die automatische Datenverarbeitung zu unterstützen. Hierzu wird in den nächsten Schritten der Import der KDD-Ergebnisse (Hypothesen) in das Simulationsmodell diskutiert. Gleichzeitig

wird das Simulationsmodell kontinuierlich erweitert, um die Komplexität der resultierenden Datenbasis zu erhöhen und in Folge den Informationsgehalt der Simulationsergebnisse zu steigern. Die Erweiterungen der nächsten Phase umfassen Mengenattribute und kurzfristige Änderungen innerhalb der Bestellprozesse sowie die Generierung zusätzlicher Transaktionsattribute.

Literatur

- Anane, R.; Younas, M.; Tsai, C.-F.; Chao, K.-M.: Agent-based transactional framework for the supply chain. In: International Conference on Machine Learning and Cybernetics, Peking (China), 2002-11-04, 2002, S. 1956–1961.
- Arndt, V.: Ereignisdiskrete Simulation einer Supply Chain zur Generierung von Transaktionsdaten. Dortmund, Technische Universität Dortmund, Fachgebiet IT in Produktion und Logistik, Masterarbeit, 2014. Online verfügbar unter <http://www.itpl.mb.tu-dortmund.de/rabe/?page=theses&lang=en>
- Adriaans, P.; Zantinge, D.: Data mining. Harlow (England), Reading (MA): Addison-Wesley 1996.
- Bangsow, S.: Manufacturing simulation with Plant Simulation and SimTalk. Berlin: Springer 2010.
- Capgemini: Nachfrageschwankungen sind größte Herausforderung im Supply Chain Management. Pressemitteilung, online verfügbar unter <http://www.de.capgemini.com/news/herausforderung-supply-chain-management>, geprüft am 11.02.2015.
- Che, D.; Safran, M.; Peng, Z.: From Big Data to Big Data Mining: Challenges, issues, and opportunities. In: Hutchison, D.; Kanade, T.; Kittler, J.; Kleinberg, J.M.; Mattern, F.; Mitchell, J.C. et al. (Hrsg.): Database Systems for Advanced Applications. Berlin, Heidelberg: Springer 2013, S. 1–15.
- Düsing, Roland (2010): Knowledge Discovery in Databases. In: Chameni, P.; Gluchowski, P. (Hrsg.): Analytische Informationssysteme. Business-Intelligence-Technologien und -Anwendungen, 4. Aufl. Berlin [u.a.]: Springer, S. 282–295.
- Fayyad, U.M.; Piatetsky-Shapiro, G.; Smyth, P.: From Data Mining to Knowledge Discovery in Databases. American Association for Artificial Intelligence 17 (1996) 3, S. 37–54.
- Gabriel, R.; Gluchowski, P.; Pastwa, A.: Data warehouse & data mining. Herdecke, Witten: W3L 2009.
- Hellström, D.; Johnsson, M.: Using discrete event simulation in supply chain planning. In: Solem, O. (Hrsg.): Proceedings of the 14th Annual Conference for Nordic Researchers in Logistics, Trondheim (Norway), June 13-14, 2002, S. 13–28.
- Hildebrand, K. (Hg.): Daten- und Informationsqualität. Wiesbaden: Vieweg + Teubner 2011.
- Horne, G.E.; Brandstein, A.G., 1997: Data Farming: A meta-technique for research in the 21st Century. Naval War College Newport, RI, November 1997.
- Horne, G.E.; Meyer, T.E.: Data Farming: Discovering surprise. In: Kuhl, M.E.; Steiger, N.M.; Armstrong, F.B.; Joines, J.A. (Hrsg.): Proceedings of the 2005 Winter Simulation Conference, Piscataway: IEEE Press 2005, S. 1082–1087.
- Konrad, G.; Stache, U.: Theorie, Anwendbarkeit und strategische Potenziale des Supply Chain Management. Wiesbaden: Deutscher Universitätsverlag/GWV Fachverlage 2005.

- Mayer-Schönberger, V.; Cukier, K.: Big Data. Die Revolution, die unser Leben verändern wird. München: REDLINE 2013.
- Petersohn, H.: Data Mining. Verfahren, Prozesse, Anwendungsarchitektur. München, Wien: Oldenbourg 2005.
- Rabe, M.; Wenzel, S.; Spieckermann, S.: Verifikation und Validierung für die Simulation in Produktion und Logistik. Berlin, Heidelberg: Springer 2008.
- Rabe, M.; Scheidler, A.A.: An approach for increasing the level of accuracy in Supply Chain Simulation by using patterns on input data. In: Tolk, A.; Diallo, S.Y.; Ryzhov, I.O.; Yilmaz, L.; Buckley, S.; Miller, J.A. (Hrsg.): Proceedings of the 2014 Winter Simulation Conference, Piscataway: IEEE Press 2014, S. 1897-1906.
- Robinson, S.: Successful simulation. A practical approach to simulation projects. London, New York: McGraw-Hill 1994.
- Sanchez, S.M.: Simulation experiments: Better data, not just big data. In: Tolk, A.; Diallo, S.Y.; Ryzhov, I.O.; Yilmaz, L.; Buckley, S.; Miller, J.A. (Hrsg.): Proceedings of the 2014 Winter Simulation Conference, Piscataway: IEEE Press 2014, S. 805–816.