ANSÄTZE EINER ALGORITHMISCHEN ANWENDUNG QUANTITATIVER VERFAHREN ZUR EFFIZIENTEN BEDARFSPROGNOSE VON VORPRODUKTEN. ERSTE ERGEBNISSE EINER EMPIRISCHEN UNTERSUCHUNG.

Thomas Cleff

Pforzheim University, Business School, Germany

Kirsten Wüst

Pforzheim University, Business School, Germany

1. DIE BEDARFSPROGNOSE IN DER LAGERHALTUNG EINES DEUTSCHEN MASCHINENBAU-UNTERNEHMENS

Gemäß der betriebswirtschaftlichen Definition gelten Lager als „Knoten im logistischen System, in dem Güter vorübergehend festgehalten und häufig in ihrer Mengenzusammensetzung verändert werden. [Lager] übernehmen sowohl die Funktion von Liefer- und Empfangspunkten, als auch von Auflöse- und Konzentrationspunkten". (Gabler, Wirtschaftslexikon (2004), S. 1843). Die stufenbezogenen Erscheinungsformen sind Eingangslager, Zwischenlager, Erzeugnis- oder Fertigungswarenlager sowie das Versandlager. Die Aufgabe einer jeden Lagerhaltungspolitik besteht darin, in Abhängigkeit von Input- und Output-Strömen, eine kostenoptimale Lagerhaltung herbei zu führen. Während Inputströme sowohl in Ein- und Mehrperiodenmodellen der Lagerhaltung in der Regel durch Bestellmengen und –zeitpunkte selbst gesteuert werden können, werden die Outputströme der Nachfrage nach Fertigprodukten fast immer exogen – also „außerhalb des Systems" - hervorgerufen. Aufgrund dieser Tatsache dürften die theoretischen Lagerhaltungsmodelle mit deterministischer Nachfrage in der Praxis nur in Ausnahmefällen anwendbar sein. Diese Modelle erfüllen nicht die Annahmen zufallsbedingter Lagerabflüsse. Dennoch konstatieren Hillier/Lieberman (2001, S. 987), dass deterministische Modelle hinreichend gute Annäherungen an viele Lagerhaltungssituationen in der Praxis sind: „For example the EOQ [Economic Order Quantity] models have been particularly widely used. These models are sometimes modified to include some type of stochastic demand, such as the stochastic continuous reviewed model does". Im Wesentlichen finden also vor allem die einfachen, intuitiven und „handhabbaren" Modelle der Lagerhaltung in der Praxis Anwendung. Dies gilt insbesondere für die in einer ABC-Analyse als B-Produkte klassifizierten Vorprodukte, denn eine systematische Anwendung komplexer Verfahren für die Vielzahl der eingesetzten Vorprodukte (Teile und Komponenten) eines Unternehmens überfordert nicht nur die Einkaufsabteilungen klein- und mittelständischer Betriebe, sondern auch große und global operierende Unternehmen. Insbesondere Unternehmen aus dem Maschinenbau, in denen die Anzahl der einzusetzenden Vorprodukte überproportional hoch ist, sehen sich diesem Problem gegenüber. Das Beschaffungsmanagement setzt deshalb nicht selten auf „handgestrickte“ Lösungen, in denen vor allem die Erfahrungen und die Intuitionen des Beschaffungsmanagers die Bestellmengen und Bestellzeitpunkte bestimmen, oder auf die „blinde“ Anwendung von vorgefertigten ERP- bzw. Lagerhaltungsprogrammen. Nicht selten werden dabei auch die realisierten und potenziellen Auftragseingänge der Vertriebsabteilung als mögliche Prädiktoren verwendet. Dieses stellt eine unbefriedigende Lösung dar, da Vertriebsabteilungen „Anbahnungen“ tendenziell überschätzt angeben, da sie die eigene Liefersicherheit höher bewerten als die Kosten einer erhöhten Kapitalbindung in Lagerbestände.

Dies hat in der Vergangenheit in vielen Einkaufsabteilungen mittelständischer Unternehmen zu Überlegungen darüber geführt, in wieweit die stochastische Nachfrage nach Vorprodukten – also die Warenabflüsse aus dem Lager - in Hinblick auf die optimale Bestellmenge und den optimalen Bestellzeitpunkt für das Eingangslager prognostiziert werden können, ohne dass hierdurch große personelle und finanzielle Kapazitäten im Unternehmen gebunden werden müssen.

Das vorliegende Papier beschreibt einen standardisierten algorithmischen Ansatz, mit dem der Verbrauch von 110 Vorprodukten eines deutschen Maschinenbauunternehmens für Zeiträume von drei, sechs oder zwölf Monaten mit Hilfe zeitreihenökonometrischer Verfahren prognostiziert werden kann. Im Rahmen dieses Ansatzes werden für jede Vorproduktgruppe die unterschiedlichsten Prognosetechniken systematisch angewendet und danach in Hinblick auf verschiedene Kriterien der Prognosequalität bewertet. Diese systematische Anwendung ermöglicht für einen erheblichen Teil der Vorprodukte eine gute Prognostizierbarkeit.

2. VERWENDETE PROGNOSETECHNIKEN ZUR EDARFSERMITTLUNG

Der Verbrauch der N=110 Vorprodukte liegt auf Monatsbasis für einen Zeitraum von T=77 Monaten vor. Zur Entwicklung und Anpassung der Prognose der Vorproduktzeitreihen $\{z_t\}_{t\in T}$ wählen wir zehn verschiedene ökonometrische Modelle, die zunächst einmal differenziert werden, so dass der Bedarf in Veränderungsraten vorliegt. Dies führt i.d.R. bereits zu stationären Datenreihen – eine Voraussetzung der Anwendung von zeitreihenökonometrischen Verfahren. Um die Stationarität in jedem Fall sicherzustellen, erfolgt eine Überprüfung mit Hilfe des Augmented Dickey Fuller Tests. Liegt keine Stationarität der Zeitreihe vor, so werden sukzessive Differenzen bis zur dritten Ordnung gebildet. Ist die Zeitreihe in der dritten Differenz immer noch nicht stationär, passen wir die Modelle für die erste Differenz an, wobei eine Warnung über die Nichtstationarität ausgegeben wird. Die verwendete Differenz bezeichnen wir mit d, die d-mal differenzierte Zeitreihe mit $\{y_t\}_{t\in T} = \{\Delta^d z_t\}_{t\in T}$. Auf Basis dieser Zeitreihe erfolgt die Anwendung verschiedener ökonometrischer Prognosetechniken, die sich in die beiden Prognoseklassen (1) der Zeitreihenökonometrie (time series method) und (2) der strukturellen Regressionsanalyse (causal method) unterteilen lassen. Die zeitreihenökonometrischen Verfahren beschränken sich auf die Nutzung der Muster und Strukturen der historischen Daten einer Variablen, um dieselbe in die Zukunft zu extrapolieren. Zur Prognose von Lagerabflüssen eines Vorproduktes werden somit lediglich die Informationen über Lagerabflüsse desselben Vorproduktes in der Vergangenheit genutzt. Die Entscheidung, auf welche und wie viele Informationen aus der Vergangenheit zurückgegriffen werden sollte, erfolgt über eine Minimierung des Bayesschen Informationskriteriums BIC (Akaike [1977]). Dieses Kriterium bestimmt die optimale Lag-Struktur der Vorproduktvariablen. Die zeitreihenökonometrischen Prognosetechniken sind in Abbildung 1 als Modelle eins bis sieben beschrieben. Im Anschluss an die Modellbildung wird jeweils ein Portmanteau-Test für die Residuen durchgeführt (Box, Pierce[1970], Ljung, Box[1978]). Die Anzahl der gemeinsam auf Null getesteten Autokorrelationen der Residuen wurde als 20% der verwendeten Zeitpunkte gewählt.

Strukturelle Prognosetechniken gehen hingegen von der Annahme aus, dass für die Entwicklung einer zu prognostizierenden Variablen kausal andere erklärende Effekte bzw. Variablen verantwortlich sind. Diese erklärenden Variablen müssen die Eigenschaft besitzen, den Entwicklungen des Lagerabflusses eines Vorproduktes zeitlich ausreichend vorauszulaufen, um den Lagerabfluss zum Zeitpunkt t bereits zu einem früheren Zeitpunkt t-x erkennbar zu machen. Als erklärende Variable kommen eine Reihe von Prädiktor-Indizes in Frage, die in der Praxis häufig auch als Frühindikatoren zur konjunkturellen Entwicklung verwendet werden. Während in den USA gerne der Purchasing Manager Index (PMI) verwendet wird, gelten für die Entwicklungen auf dem deutschen Markt vor allem der Einkaufmanager Index des Bundesverbandes Materialwirtschaft, Einkauf und Logistik e.V. (BME) und der Ifo-Geschäftsklimaindex als gute Frühindikatoren.

Abbildung 1: Verwendete Prognosetechniken

	Modell	**Beschreibung**
Referenz-modell	1	Als „**naives Vergleichs- oder Referenzmodell**“ für die folgenden Verfahren wird ein Random Walk $y_t = y_{t-1} + \varepsilon_t$ verwendet. Für die Prognose ergibt sich so beim Random Walk, dass $\hat{y}_t = y_{t-1}$ gilt. Der prognostizierte Wert entspricht also jeweils dem der letzten Beobachtung.
Zeitreihenökonometrische Verfahren	2	**MA(q)-Modell**: Es wird ein reiner Moving Average Prozess q-ter Ordnung $y_t = \varepsilon_t - \beta_1 \varepsilon_{t-1} - ... - \beta_q \varepsilon_{t-q}$ angepasst.
	3	**AR(p)-Modell**: Es wird ein reiner autoregressiver Prozess p-ter Ordnung $y_t = \alpha_1 y_{t-1} + ... + \alpha_p y_{t-p} + \varepsilon_t$ angepasst.
	4	Die optimale Ordnung eines ARMA(p,q)-Modells $y_t = \alpha_1 y_{t-1} + ... + \alpha_p y_{t-p} = \varepsilon_t - \beta_1 \varepsilon_{t-1} - ... - \beta_q \varepsilon_{t-q}$ wird in Anlehnung an das Hannan-Rissanen-Verfahren (Hannan, Rissanen [1982]) bestimmt. Zunächst wird ein AR(p)-Modell unterstellt. Es wird die in Modell 3 als optimal bestimmte Ordnung p zur Anpassung des Prozesses ausgewählt. Für den so angepassten Prozess werden daraufhin die Residuen, sowie die Lags der Residuen ermittelt. Im Anschluss erfolgt eine Regression der Lags der Residuen und der als optimal ermittelten autoregressiven Terme auf die abhängige Vorproduktvariable. Die optimal zu verwendenden Lag-Strukturen der Residuen werden über das Bayessche Informationskriterium (BIC) ermittelt. Da der ARMA-Prozess für die im voraus d-mal differenzierte Zeitreihe angepasst wird, liegt ein **ARIMA(p,d,q)**-Modell für die Zeitreihe $\{z_t\}$ vor.

	Modell	Beschreibung
	5	Die Werte der Zeitreihe des betrachteten Vorproduktes werden zunächst logarithmiert, d.h. die Zeitreihe $\{\ln(z_i)\}_{i\in I}$ betrachtet. Für die so transformierte Zeitreihe wird – wie bei Modell 4 - ein optimales ARIMA(p,d.q)-Modell in Anlehnung an das Hannan-Rissanen-Verfahren bestimmt. Die Ordnung für die Differenz d entspricht wiederum den Erfordernissen aus dem Augmented Dickey Fuller Test. Für die differenzierte logarithmierte Zeitreihe wird zunächst ein optimaler AR(p)-Prozess angepasst, für den die Residuen und die Lags der Residuen berechnet werden. Die potentiellen AR- und MA-Terme werden auf die Zeitreihe regressiert. Es werden die Ordnungen gewählt, für die das Bayessche Informationskriterium minimal ist.
	6	Desweiteren wird die Zeitreihe des Bedarfs des untersuchten Vorproduktes über 3 Monate geglättet. Hierbei wird der letzte beobachtete Wert zur Hälfte, der vorletzte und der drittletzte Wert je zu einem Viertel gewichtet, so dass $z_{t,neu} = 0{,}5 \cdot z_t + 0{,}25 \cdot z_{t-1} + 0{,}25 \cdot z_{t-2}$ gilt. An die so modifizierte Zeitreihe wird erneut ein ARIMA(p,d,q)-Modell nach dem Hannan-Rissanen-Verfahren angepasst.
	7	Im siebten Modell wird bei der Anpassung der Zeitreihe eine Niveaukomponente berücksichtigt, d.h. für den **ARIMA(p,d,q)**-Prozesses wird ein **konstanter Term** zugelassen, so dass $y_t = \mu + \alpha_1 y_{t-1} + ... + \alpha_p y_{t-p} = \varepsilon_t - \beta_1 \varepsilon_{t-1} - ... - \beta_q \varepsilon_{t-q}$ gilt.
Strukturelle Regression	8	Hierbei werden zunächst für jede einzelne unabhängige Variable die optimalen Lag-Längen bei einer multiplen Regression auf die Vorproduktvariable ermittelt, indem sukzessive unterschiedliche Lag-Längen und Kombinationen von unterschiedlichen Lag-Längen als Regressoren ausgetestet werden und mit Hilfe des Bayesschen Informationskriteriums (BIC) die optimale Lag-Struktur bestimmt wird. Für die so ermittelten optimalen Lag-Strukturen aller unabhängigen Variablen wird im zweiten Schritt eine schrittweise Rückwärtsregression auf die abhängige Vorproduktvariable ausgeführt. In jedem Durchlauf wird also von den Variablen, für die der p-Wert größer als 0,05 ist, diejenige Variable mit dem höchsten p-Wert ausgeschlossen. Es wird solange fortgefahren, bis nur noch signifikante (p≤0,05) Regressoren im Modell verbleiben.
	9	Dieses Modell entspricht dem Modell 8, bei der anstelle der Rückwärtsregression eine Vorwärtsregression mit schrittweisem Einschluss der Variablen mit dem jeweils kleinsten p-Wert kleiner gleich 0,05 durchgeführt wird. Verliert eine bereits im Modell befindliche unabhängige Variable im fortlaufenden Iterationsprozess ihre Signifikanz, wird sie aus dem Modell ausgeschlossen, wenn der p-Wert größer als 0,1 wird. Die potentiellen unabhängigen Variablen sind diejenigen, die auch für Modell 8 bestimmt wurden.
Mischverfahren	10	Schließlich werden die Modelle 4 und 8 kombiniert. Für Vorproduktvariable $y_t = z_t - z_{t-1}$ wird eine Regression auf die ersten Differenzen der prognostizierten Werte aus dem ARIMA(p,d,q)-Prozess in Modell 4 $(f_{4,t} - f_{4,t-1})$ und dem strukturellen Modell 8 $(f_{8,t} - f_{8,t-1})$ vorgenommen. Die Prognose für y_t ergibt sich jeweils, indem zu dem bekannten Wert y_{t-1} der Schätzwert der Regressionsfunktion hinzugerechnet wird, d.h. $y_t = y_{t-1} + \hat{\beta}_0 + \hat{\beta}_1(f_{4,t} - f_{4,t-1}) + \hat{\beta}_2(f_{8,t} - f_{8,t-1})$.

Vor allem für letzteren werden sowohl für den Gesamtindex als auch für einzelne Teilindizes branchenspezifische Einzelergebnisse publiziert.[54] Schätzt die Abnehmerindustrie eines Unternehmens die Geschäftslage in den nächsten sechs Monaten rückläufig ein, wird auch der Absatz des Zulieferunternehmens in der nächsten Zeit zurückgehen. Als Prädiktoren für den Abfluss aus einem Einkaufslager eignen sich theoretisch also die elf möglichen Teilindizes der jeweiligen Abnehmerindustrien. In den strukturellen Prognosemodellen 8 bis 9 in Abbildung 1 wird unterstellt, dass die Bedarfszeitreihe des betrachteten Vorprodukts linear von den elf Konjunkturindikatoren abhängt. Die zehnte Prognosetechnik versucht die Ergebnisse des zeitreihenökonometrischen Modells vier mit dem strukturellen Prognosemodell 8 miteinander zu verbinden.

3. VERWENDETE KENNZIFFERN ZUR BEWERTUNG DER PROGNOSE-GÜTE

Bei der Beurteilung der Güte eines Prognosemodells muss berücksichtigt werden, dass das Modell nicht nur die zur Erstellung des Modells verwendeten Daten, sondern auch Daten, die bei der Erstellung des Modells noch nicht vorlagen, möglichst gut wiedergibt. Um diese „Generalisierbarkeit“ einer Prognose sicherstellen zu können, wird in der Praxis häufig auf die Cross-Validation-Methode zurückgegriffen, bei der der Datensatz in zwei Teile geteilt wird. Auf Basis des so genannten Trainingsdatensatzes wird das Prognosemodell erstellt, um es am aktuellen Rand der Zeitreihe – dem sogenannten Testdatensatz - auf seine Prognosegüte hin zu überprüfen.

Bei den vorliegenden Bedarfsdaten wurde für die Länge des zu prognostizierenden Zeitraumes ein Testdatensatz am aktuellen Rand der Zeitreihe abgeschnitten. So konnten bei einer Kurzzeitprognose über drei Monate ein Trainingsdatensatz mit der Länge der gesamten Zeitreihe abzüglich der letzten drei Monate verwendet werden. Der dreimonatige Testdatensatz am aktuellen Rand dient zur Überprüfung der Prognosegüte. Die längsten Bedarfsprognosen erstreckten sich über 12 Monate, so dass der Trainingsdatensatz in diesem extremen Fall immerhin noch (77-12=) 65 Beobachtungen enthielt.

Im Folgenden wurden mehrere verschiedene Maße berechnet und sowohl auf den Trainings- als auch auf den Testdatensatz angewendet, um die Güte der Prognose zu beurteilen:

1. Der „Root Mean Squared Error“ (RMSE) ist das meist verwendete Maß für den Erfolg einer Prognose. Hierbei werden die mittleren quadratischen Abweichungen der Prognosewerte von den wahren Werten gemittelt (MSE). Der RMSE entsteht durch Wurzelbildung aus dem MSE, d.h. $RMSE = \sqrt{\frac{1}{N}\sum_{t=1}^{N}\left(y_t - \hat{y}_t\right)^2}$.

[54] Die Teilindizes sind (1) die derzeitige Geschäftslage, (2) die inländische Produktionstätigkeit, (3) der Bestand an Fertigwaren, (4) die Nachfragesituation, (5) die Inlandsverkaufspreise, (6) der Auftragsbestand im Vergleich zum Vormonat, (7) die Auslandsaufträge, (8) das Exportgeschäft, (9) die Entwicklung der Beschäftigtenzahl, (10) die Entwicklung der Verkaufspreise für die nächsten 3 Monate sowie (11) die Geschäftslage für die nächsten 6 Monate.

2. Zur Normierung des Prognosefehlers ist das Prognosemaß auf eine Referenzgröße zu beziehen. Der Theil'sche Ungleichheitskoeffizient (Theil's U) bezieht die Wurzel aus dem mittleren quadratischen Fehler (RMSE) bei dem verwendeten Prognosemodell auf den RMSE der naiven Prognose, d.h.
$$U = \frac{\sqrt{\frac{1}{N}\sum_{t=1}^{N}\left(y_t - \hat{y}_t\right)^2}}{\sqrt{\frac{1}{N}\sum_{t=1}^{N}\left(y_t - \hat{y}_{Naive\ \mathrm{Pr}ognose,t}\right)^2}}.$$
Bei einer perfekten Prognose nimmt der MSE und damit auch Theil's U den Wert Null an. Ist die Prognose nur so gut wie eine naive Prognose einer Fortschreibung der jeweils letzten Beobachtung in die Zukunft, so hat Theil's U den Wert Eins. Nimmt Theil's U Werte größer (kleiner) als Eins an, so ist das gewählte Prognoseverfahren schlechter (besser) als eine naive Fortschreibung der Vergangenheitswerte in die Zukunft. Zugegebenermaßen würde wohl kaum ein Beschaffungsmanagement ausschließlich Vergangenheitswerte fortschreiben. Anstelle dessen wären die Werte einer „unternehmensüblichen" Prognose als Referenzreihe angemessen gewesen. Diese lagen allerdings nicht vor, so dass an dieser Stelle mit der üblichen Referenzreihe der naiven Fortschreibung von Vergangenheitswerten kalkuliert wird.
3. Die mittlere prozentuale Abweichung mittelt die prozentualen Abweichungen der Prognosen bezogen auf die wahren Daten, d.h. $MPA = \frac{1}{N}\sum_{t=1}^{N}\frac{\left|y_t - \hat{y}_t\right|}{y_t}$.
4. Für die praktische Beschaffungsprognose ist ein weiteres Maß von besonderer Bedeutung. Es interessiert nämlich häufig nur, ob für den gesamten Prognosezeitraum die erforderlichen Bestellmengen in ihrer Gesamtsumme gut prognostiziert und nicht zu weit über- oder unterschätzt werden. Abweichungen nach oben und unten können sich innerhalb des Prognosezeitraums tendenziell ausgleichen, sobald der Fehlerterm nicht autokorreliert ist. Dementsprechend wird bei der durchschnittlichen gesamten prozentualen Abweichung im Prognosezeitraum (MGPA) die Summe der Prognosefehler auf die Summe der wahren Bedarfswerte bezogen, d.h. $MGPA = \frac{\sum_{t=1}^{N}(\hat{y}_t - y_t)}{\sum_{t=1}^{N} y_t}$.

Im Beschaffungsmanagement geht es aber nicht nur darum, eine Punktschätzung mit einem möglichst geringen Prognosefehler für den zukünftigen Bedarf von Vorprodukten zu ermitteln. Vielmehr ist ebenfalls von Interesse, ob die möglichen Prognosefehler symmetrisch und ohne große Ausreißer um die Punktschätzer verteilt sind. Gerade Ausreißer können nämlich dazu führen, dass ein künftiger Bedarf massiv über- oder unterschätzt wird, was letztlich möglicherweise zu hoher Kapitalbindung oder zu Produktionsausfällen führen kann.

Deshalb wurden weiterhin die Schiefe und die Kurtosis der Prognosefehler für den Testdatensatz ermittelt. Der Test von Schiefe und Kurtosis wurde nach D'Agostino, Balanger, and D'Agostino, Jr. (1990) mit dem empirischem Adjustment von Royston (1991c) durchgeführt. Eine rechtsschiefe Verteilung der Prognosefehler $e_t = y_t - \hat{y}_t$, d.h. das verstärkte Auftreten von negativen Prognosefehlern, signalisiert eine überdurchschnittlich große Überschätzung einzelner Werte, eine linksschiefe Verteilung eine überdurchschnittlich große Unterschätzung. Bei einem p-Wert kleiner oder gleich 0,05 und einer Schiefe größer als Null wird die Verteilung als rechtsschief, bei einer negativen Schiefe als linksschief angenommen. Bei einem insignifikanten p-Wert größer als 0,05 liegen keine ungleichmäßig großen Über- oder Unterschätzungen vor und wir bezeichnen die Verteilung als „nicht schief". Ergibt sich für die Kurtosis ein Wert größer als drei bei einem p-Wert kleiner oder gleich 0,05, so darf angenommen werden, dass die Verteilung „zu starken Ausreißern neigt", bei einer Kurtosis kleiner als 3 und/oder einem insignifikanten p-Wert größer als 0,05 „neigt die Verteilung nicht zu starken Ausreißern".

4. ERGEBNIS DER PROGNOSEFÄHIGKEIT VON VORPRODUKTEN

Die N=110 Datenreihen wurden mit den oben beschriebenen zehn Prognosetechniken kurz-, mittel- und langfristigen Prognosen unterzogen. Bei den kurzfristigen Prognosen wurde ein Prognosezeitraum von 3 Monaten, bei den mittelfristigen von 6 Monaten und bei den langfristigen Prognosen ein Prognosezeitraum von 12 Monaten gewählt, deren Prognosegüte im weiteren Verlauf getrennt voneinander untersucht wurde.

Bei der Anpassung der Modelle wurde jeweils das Modell als optimal gewählt, das den geringsten Theil'schen Ungleichheitskoeffizienten während der Trainingsphase hatte und dessen Residuen als „Weißes Rauschen" (White Noise) angesehen werden können. Letztere Bedingung wird mit Hilfe des Portmanteau-Tests überprüft. Ist der Portmanteau-Test für das automatisch gewählte Modell negativ, so wird dasjenige Modell mit dem niedrigsten Theil's U in der Trainingsphase verwendet, für das der Portmanteau-Test ergibt, dass die Residuen unabhängig und standardnormalverteilt sind. Die oben beschriebenen Gütemaße der Prognose in Bezug auf den Trainings- und den Testdatensatz ermöglichen danach weitere Aussagen hinsichtlich der Qualität des gewählten Prognosemodells.

Insgesamt fiel auf, dass die angepassten Modelle im Trainingsdatensatz stärker zur überdurchschnittlich starken Überschätzung als zur überdurchschnittlich starken Unterschätzung der tatsächlichen Werte neigen. Eine starke Überschätzung der wahren Werte tritt je nach Länge des Prognosezeitraumes in 39,8% bis 44,4% der Fälle auf. Verbunden mit der Tatsache, dass mehr als 50% der Zeitreihen zu Ausreißern neigen (s. Kurtosis in Tabelle 1), ist dies ein erstes Indiz für die Tatsache, dass eine erhebliche Anzahl von Vorproduktreihen durch sporadische Nachfrageeinbrüche und sogar durch „Nullnachfragen" gekennzeichnet sind. Sporadische Ausreißer nach oben kommen hingegen selten vor. Nur 2,7% bis 4,6% der Prognosen weisen eine starke Unterschätzung auf (vgl. Tabelle 1).

Tabelle 1: Schiefe und Kurtosis der Prognose mit bester Prognosegüte im Trainingsdatensatz

Schiefe	Kurzfristige Prognose (3-Monatsprognose)		Mittelfristige Prognose (6-Monatsprognose)		Langfristige Prognose (12-Monatsprognose)	
	Anzahl	Prozent	Anzahl	Prozent	Anzahl	Prozent
rechtsschief	3	2,7	4	3,7	5	4,6
Nicht schief	62	56,3	61	56,5	55	51
linksschief	45	40,9	43	39,8	48	44,4
Gesamt	110	100,0	108	100,0	108	100,0
Kurtosis	**Kurzfristige Prognose (3-Monatsprognose)**		**Mittelfristige Prognose (6-Monatsprognose)**		**Langfristige Prognose (12-Monatsprognose)**	
	Anzahl	Prozent	Anzahl	Prozent	Anzahl	Prozent
Neigt nicht zu Ausreißern	53	48,2	54	50,0	53	49,1
Neigt zu Ausreißern	57	51,8	54	50,0	55	50,9
Gesamt	110	100,0	108	100,0	108	100,0

4.1. Kurzfristige Prognose

Tabelle 2 zeigt die Verteilung der als optimal bestimmten Modelle der 110 Vorprodukte bei der kurzfristigen 3-Monats-Prognose.

Tabelle 2: Modelle mit bester Prognosegüte aus Theil's U im Trainingsdatensatz (3-Monatsprognose)

Prognosetyp	Prognose-modell	Gewählt[1)]	Schlechter als Referenzmodell		Besser als Referenzmodell	
			Anzahl	Prozent	Anzahl	Prozent
Zeitreihenökonometrische Verfahren	Modell 2	21,8%	5	20,8%	19	79,2%
	Modell 3	5,5%	4	66,7%	2	33,3%
	Modell 4	45,5%	18	39,1%	28	60,9%
	Modell 5	1,8%	1	50,0%	1	50,0%
	Modell 6	3,6%	2	50,0%	2	50,0%
	Modell 7	19,1%	7	33,3%	14	66,6%
Strukturelle Regression	Modell 8	1,8%	1	50,0%	1	50,0%
	Modell 9	0,9%	0	0,0%	0	0,0%
Mischverfahren	Modell 10	4,5%	4	80,0%	1	20,0%
	Gesamt		42	38,2%	68	61,8%

[1)] Aufgrund von Mehrfachnennungen ergeben sich eine Gesamtsumme von über 110 Zeitreihen und eine prozentuale Summe von mehr als 100%.

Am häufigsten wird das Modell 4 (ARIMA(p,d,q) ohne Niveau) in 45,5% der Fälle gewählt, gefolgt von Modell 2 (reiner MA-Prozess) mit 21,8% der Fälle. In vier Fällen führten Modell 3 und Modell 4 zu demselben Modell (reines AR(p)-Modell), in einem Fall Modell 8 und Modell 9. Es wird im Folgenden wieder von 110 Zeitreihen ausgegangen, die identischen Modelle werden bei der jeweils kleineren Modellziffer (d.h. bei Modell 3 und Modell 8) aufgeführt.

Ob sich die Ergebnisse in Bezug auf die Anwendung des jeweiligen Prognosemodells auf den Testdatensatz ähneln, zeigt Tabelle 2. Bei 61,8% der Modelle ist Theil's U auch in der Prognosephase kleiner als Eins, d.h. die Prognose ist besser als die naive Ergebnisfortschreibung aus der jeweiligen Vorperiode. Hier schneidet das reine Ma(q)-Modell 2 mit 79,2% am besten ab, gefolgt von dem ARIMA(p,d,q)-Modell mit Niveau (Modell 7) mit 66,7% sowie dem ARIMA(p,d,q)-Modell ohne Niveau (Modell 4) mit 60,9%.

Insgesamt sind diese Ergebnisse eher ernüchternd, denn immerhin wäre zu erwarten gewesen, dass sich bei weit mehr als 61,8% der Zeitreihen auch für den Testdatensatz bessere Prognosen ergeben, als durch die naive Zeitreihenfortschreibung. Es ist daher nicht verwunderlich, dass sich nur bei 38,2% der Zeitreihen eine mittlere prozentuale Abweichung von weniger als 20% ergibt. Bei rund 32 Prozent liegt diese Abweichung zwischen 20 und 40 Prozent und bei ca. 30 Prozent der Zeitreihen sogar über 40

Prozent. Bezieht man den Fehler nun nicht auf die durchschnittliche monatliche Abweichung, sondern auf die durchschnittliche Abweichung, die im gesamten Prognosezeitraum von drei Monaten durch Aggregation der prognostizierten und tatsächlichen Monatswerte auf Quartalswerte auftritt, sind die Ergebnisse nur leicht verbessert: Die durchschnittliche gesamte prozentuale Abweichung im Prognosezeitraum (MGPA) liegt bei der kurzfristigen Prognose zunächst bei rund 54 Prozent der prognostizierten Zeitreihen bei absolut weniger als 20%, bei insgesamt 25 Prozent liegt die absolute MGPA zwischen 20% und 40%. Bei noch immer rund 21 Prozent der Zeitreihen ist die absolute MGPA größer als 40% (davon für 10 Prozent größer als 60 Prozent!).

Bei welchen Verläufen der Vorproduktzeitreihe häufen sich hohe Gesamtabweichungen im Prognosezeitraum? Kann man bereits an der Trainingsphase erkennen, ob es zu einer starken Über- oder Unterschätzung im Prognosezeitraum kommen wird? Dazu haben wir die Verläufe der Zeitreihen mit hohen Abweichungen analysiert und festgestellt, dass diese besonders häufig dann auftraten, wenn im Trainingszeitraum viele Zeitpunkte existierten, an denen kein Bedarf verzeichnet wurde. In einer gesonderten Analyse haben wir daher alle Zeitreihen ausgeschlossen, bei denen in der Trainingsphase nach dem anfänglichen Auftreten von positiven Werten acht (entspricht 12,5% der Trainingszeitpunkte) oder mehr Werte kleiner als zehn zu verzeichnen waren. Hierbei handelt es sich um Vorprodukte, die entweder

- in der weit zurückliegenden Vergangenheit zu den Vorprodukten gehört haben und in der jüngeren Vergangenheit nur noch als Ersatzteile dienen oder
- um Vorprodukte, deren Nachfrage sporadisch im großen Mengen auftritt und für die in darauffolgenden Zeiträumen Nullnachfragen bestehen.

Für diese Vorprodukte eignen sich die quantitativen Prognoseverfahren wenig. Es sollte deshalb auf andere Verfahren zurückgegriffen werden. Für die verbliebenen 74 Zeitreihen ist die Verteilung der durchschnittlichen gesamten prozentualen Abweichungen in Tabelle 3 angegeben. Für rund 61 Prozent der prognostizierten Reihen ist der Absolutbetrag des MGPA kleiner als 20 Prozent, für 27 Prozent der Zeitreihen liegt er zwischen 20 Prozent und 40 Prozent und nur bei 12 Prozent ist der MGPA größer als 40 Prozent.

Tabelle 3: MGPA im Prognosezeitraum bei Vorprodukten mit weniger als acht Werten kleiner zehn in der Trainingsphase (3 Monate)

Abweichung	Anzahl	Prozent	Abweichung bis höchstens 20 Prozent	Abweichung bis höchstens 40 Prozent
[-0,6;-0,4)	1	1,4		
[-0,4;-0,2)	18	24,3		
[-0,2;0)	30	40,5	60,8	87,8
[0;0,2)	15	20,3		
[0,2;0,4)	2	2,7		
[0,4;∞)	8	10,8		
Gesamt	74	100,0		

Auch wenn sich durch die Einschränkung der Bedarfsreihen die Prognosequalität verbessert, kann man sich in der praktischen Anwendung wohl kaum mit Abweichungen von mehr als 20 Prozent zufrieden geben. Zusätzlich wurde deshalb für alle Zeitreihen analysiert, ob die in der Trainingsphase ermittelte Schiefe und Kurtosis der Zeitreihe bereits Indikatoren für die Abweichung der Prognose im Prognosezeitraum sind.

Tabelle 4 zeigt den Zusammenhang zwischen der klassierten Fehlprognose - gemessen als mittlerer gesamter prozentualer Abweichung in der Prognosephase - und der Schiefe in der Trainingsphase sowie den Zusammenhang zwischen der klassierten Fehlprognose und der Kurtosis.

Tabelle 4: Kreuztabelle von Schiefemaß und Kurtosis in der Trainingsphase und MGPA in der Prognosephase (3 Monate) - weniger als 8 Werte kleiner 10

Schiefe[1)]	MGPA < -0,2	MGPA ϵ [-0,2;0,2)	MGPA ≥ 0,2	Gesamt
Schief	10 (5,6)	11 (13,4)	1 (3,0)	22
Nicht schief	9 (13,4)	34 (31,6)	9 (7,0)	52
Gesamt	19	45	10	74
Kurtosis[1)]	**MGPA < -0,2**	**MGPA ϵ [-0,2;0,2)**	**MGPA ≥ 0,2**	**Gesamt**
Neigt nicht zu Ausreißern	5 (9,8)	26 (23,1)	7 (5,1)	38
Neigt zu Ausreißern	14 (9,2)	19 (21,9)	3 (4,9)	36
Gesamt	19	45	10	71

[1)] In Klammern die jeweils erwartete Häufigkeit bei Unabhängigkeit

Betrachtet man die Schiefe (p=0,03) und die Kurtosis (p=0,03) der Prognosefehler im Trainingszeitraum ergibt sich für beide Maße ein Zusammenhang zu der MGPA im Prognosezeitraum. Schiefe – und hierbei vor allem rechtsschiefe - Verteilungen neigen zu starken Unterschätzungen in der Prognosephase. Dies erklärt sich durch zum Teil starke Ausreißer nach oben, die sich aufgrund der glättenden Eigenschaften der Modelle in einem kurzen Prognosezeitraum von nur drei Monaten besonders bemerkbar machen. Verteilungen, die zu Ausreißern neigen, neigen nämlich signifikant häufig zu starken Unterschätzungen (vgl. Kurtosis in Tabelle 4). Nichtschiefe Verteilungen neigen hingegen überdurchschnittlich häufig zu guten Prognosen mit einem betragsmäßigen MGPA von unter 20%, aber leider auch häufiger zu Überschätzungen. Allerdings weisen insgesamt 50% der schiefen Verteilungen einen betragsmäßigen MGPA von über 20% auf, während es bei den nichtschiefen Verteilungen nur rund 35% sind. Bei kurzfristigen Prognosen sollte bei schiefen Verteilungen der Fehlerterme deshalb mit Verzerrungen gerechnet werden. Bei mittelfristigen und langfristigen Prognosen sind derartige Effekte nicht zu beobachten, da sich mögliche Ausreißer über die Länge des Prognosezeitraumes „ausnivellieren".

4.2. Mittelfristige Prognose

Bei der mittelfristigen Prognose ergibt sich für das Modell mit dem geringsten Theil'schen Ungleichheitskoeffizienten während der Trainingsphase, für das gleichzeitig der Portmanteau-Test ein positives Ergebnis ergibt, die in Tabelle 5 gezeigte Verteilung. Bei zwei Zeitreihen erweist sich das Referenzmodell als beste Anpassung. Diese Zeitreihen werden aus der Analyse herausgenommen. Im Folgenden beziehen wir uns daher auf eine Grundgesamtheit von 108 Zeitreihen.

Tabelle 5: Modelle mit bester Prognosegüte aus dem Trainingsdatensatz angewendet auf den Testdatensatz (6-Monatsprognose)

Prognosetyp	Prognose-modell	Gewählt[1)]	Schlechter als Referenzmodell		Besser als Referenzmodell	
			Anzahl	Prozent	Anzahl	Prozent
Zeitreihenökonometrische Verfahren	Modell 2	29,6%	12	37,5%	20	62,5%
	Modell 3	3,7%	2	50,0%	2	50,0%
	Modell 4	40,7%	18	43,9%	23	56,1%
	Modell 5	3,7%	2	50,0%	2	50,0%
	Modell 6	5,6%	1	16,7%	5	83,3%
	Modell 7	13,0%	5	35,7%	9	64,3%
Strukturelle Regression	Modell 8	2,8%	2	66,7%	1	33,3%
	Modell 9	1,9%	0	,0%	0	,0%
Mischverfahren	Modell 10	4,6%	0	,0%	4	100,0%
	Gesamt		42	38,9%	66	61,1%

[1)] Aufgrund von Mehrfachnennungen ergeben sich eine Gesamtsumme von über 110 Zeitreihen und eine prozentuale Summe von mehr als 100%.

Am häufigsten wird wieder das Modell 4 (40,7%) gefolgt von Modell 2 (29,6%) zur Klassifikation herangezogen. Tabelle 5 zeigt außerdem die Güte der Prognose in der Testphase. Nur bei 61,1% der Modelle ist Theil's U auch in der Prognosephase kleiner als Eins, d.h. die Prognose ist besser als die naive Wertfortschreibung. 56,1% der

ARIMA(p,d,q)-Modelle ohne Niveau (Modell 4), sowie 64,3% der ARIMA(p,d,q)-Modelle mit Niveau (Modell 7) liegen mit ihrem Theil's U der Prognosephase unter Eins. Bei den reinen MA(q)-Modellen (Modell 2) sind es bei der mittelfristigen Prognose nur 62,5%.

Wir schließen wiederum die Zeitreihen aus, bei denen im Trainingszeitraum acht oder mehr Werte kleiner als zehn auftreten. Es ergibt sich die in Tabelle 6 gezeigte Verteilung der mittleren gesamten prozentualen Abweichungen. Bei 66% der prognostizierten Reihen ist die MGPA absolut gesehen kleiner als 20%, für 18% der Zeitreihen liegt sie zwischen 20% und 40% und bei 16% der Zeitreihen ist sie größer als 40% (davon für nur 3% größer als 60%).

Tabelle 6: MGPA im Prognosezeitraum bei Vorprodukten mit weniger als acht Werten kleiner zehn in der Trainingsphase (6 Monate)

Abweichung	Anzahl	Prozent	Abweichung bis höchstens 20 Prozent	Abweichung bis höchstens 40 Prozent
[-0,6;-0,4)	2	2,6		
[-0,4;-0,2)	3	3,9		
[-0,2;0)	22	28,9	65,7	84,1
[0;0,2)	28	36,8		
[0,2;0,4)	11	14,5		
[0,4;0,6)	10	13,1		
Gesamt	76	100,0		

Schiefe (p=0,256) und Kurtosis (p=0,054) der Prognosefehler im Trainingszeitraum liefern hier keinen signifikanten Zusammenhang zu der durchschnittlichen gesamten prozentualen Abweichung. Ausreißer führen also nicht überdurchschnittlich häufig zu starken Fehlprognosen. Dies könnte auch ein Grund dafür sein, weshalb der Anteil der Prognosen mit einem betragsmäßigen MGPA kleiner als 20% um fünf Prozentpunkte höher liegt als bei Kurzzeitprognosen. Hieraus auf die Qualität der Anpassung zu schließen ist insofern fraglich, als dass der Anteil der Prognosen mit erheblichen Fehlern von mehr als 40% bei den mittelfristigen Prognosen um vier Prozentpunkte höher als bei den kurzfristigen Prognosen liegt.

4.3. Langfristige Prognose

Bei der langfristigen Prognose erweisen sich das ARIMA(p,d,q)-Modell (38%) sowie das reine MA(q)-Modell (25,9%) als die Prognosemodelle mit der besten Güte im Trainingsdatensatz. Zwei Zeitreihen konnten nicht angepasst werden, da sich für die Residuen kein Weißes Rauschen ergab. Wir beziehen uns daher wieder auf eine Grundgesamtheit von 108 Zeitreihen.

Tabelle 7 gibt die Güte der Prognose in seiner Anwendung auf den Testdatensatz wieder. Auch bei der langfristigen Prognose schneiden das ARIMA(p,d,q)-Modell (Modell 4) mit 59,5% der Prognosemaße kleiner Eins sowie das ARIMA(p,d,q)-Modell mit Niveau (Modell 7) mit 60,0% gut ab. Das reine MA(q)-Modell (Modell 2) liefert zu 57,1% eine bessere Prognosegüte als das naive Modell.

Tabelle 7: Modelle mit bester Prognosegüte aus dem Trainingsdatensatz angewendet auf den Testdatensatz (12-Monatsprognose)

Prognosetyp	Prognose-modell	Gewählt[1)]	Schlechter als Referenzmodell		Besser als Referenzmodell	
			Anzahl	Prozent	Anzahl	Prozent
Zeitreihenökonometrische Verfahren	Modell 2	25,9%	12	42,9%	16	57,1%
	Modell 3	5,6%	3	50,0%	3	50,0%
	Modell 4	38,0%	15	40,5%	22	59,5%
	Modell 5	4,6%	2	40,0%	3	60,0%
	Modell 6	9,3%	2	20,0%	8	80,0%
	Modell 7	13,9%	6	40,0%	9	60,0%
Strukturelle Regression	Modell 8	4,6%	3	60,0%	2	40,0%
	Modell 9	1,9%	0	0,0%	1	100,0%
Mischverfahren	Modell 10	0,9%	0	0,0%	1	100,0%
	Gesamt		43	39,81%	65	60,19%

[1)] Aufgrund von Mehrfachnennungen ergeben sich eine Gesamtsumme von über 110 Zeitreihen und eine prozentuale Summe von mehr als 100%.

Bezieht man sich nur auf die 78 Zeitreihen, bei denen weniger als acht Werte im Trainingszeitraum kleiner als zehn sind, ist die Verteilung der durchschnittlichen gesamten prozentualen Abweichungen in Tabelle 8 gegeben. Hier haben 60% der Abweichungen einen Absolutbetrag kleiner 20%, bei weiteren 31% liegt der Absolutbetrag der Abweichungen zwischen 20% und 40%, nur 9% der Abweichungen sind absolut größer als 40% (davon sind 4% der absoluten Abweichungen größer als 60%).

Tabelle 8: MGPA im Prognosezeitraum bei Vorprodukten mit weniger als acht Werten kleiner zehn in der Trainingsphase (12 Monate)

Abweichung	Anzahl	Prozent	Abweichung bis höchstens 20 Prozent	Abweichung bis höchstens 40 Prozent
[-1;-0,4)	2	2,6		
[-0,4;-0,2)	4	5,1		
[-0,2;0)	26	33,3	60,2	90,9
[0;0,2)	21	26,9		
[0,2;0,4)	20	25,6		
[0,4;0,8)	5	6,4		
Gesamt	78	100,0		

Die Qualität der Prognosen entspricht somit derjenigen der kurzfristigen Prognosen: Zwar erreichen langfristige Prognosen nicht so häufig wie mittelfristige Prognosen einen betragsmäßigen MPGA von unter 20%, allerdings weisen sie dafür seltener einen MGPA-Prognosefehler von über 40% auf. Auch die überdurchschnittlich häufige Unterschätzung bei schiefen Verteilungen und die überdurchschnittlich häufigen Überschätzungen bei nicht-schiefen Verteilungen ist analog zur Struktur kurzfristiger Prognosen. Der p-Wert der Schiefe beträgt p=0,04, wobei der Anteil der Zellen mit einer erwarteten Häufigkeit kleiner als fünf mit 33% grenzwertig hoch liegt und die Signifikanz der Ergebnisse verzerrt. Für die Kurtosis ergibt sich keine signifikante Abhängigkeit des MGPA von der Ausreisseranfälligkeit der Zeitreihen (p=0,160).

Tabelle 9: Kreuztabelle von Schiefemaß Skew in der Trainingsphase und MGPA in der Prognosephase (12 Monate) – weniger als 8 Werte kleiner als zehn

Schiefe[1)]	MGPA < -0,2	MGPA ϵ [-0,2;0,2)	MGPA ≥ 0,2	Gesamt
Schief	5 (2,5)	20 (19,3)	7 (10,3)	32
Nicht schief	1 (3,5)	27 (27,7)	18 (14,7)	46
Gesamt	6	47	25	78

1) In Klammern die jeweils erwartete Häufigkeit bei Unabhängigkeit

5. FAZIT

Zufällig schwankende Nachfragen nach Vorprodukten (Teile und Komponenten) machen die Verwendung von stochastischen Modellen der Lagerhaltung notwendig. In dem vorgestellten algorithmischen Ansatz wird aufgrund von Gütekriterien automatisch ein optimales Prognosemodell ermittelt, das für die Prognose des Bedarfs verwendet wird. Für alle gewählten Prognosezeiträume erwies sich das ARMA Modell der d-differenzierten Zeitreihe als bestes Prognosemodell, gefolgt von einfachen Moving Average und ARIMA Modellen (vgl. hierzu Tabelle 10). Die Bedeutung autoregressiver Verfahren nimmt aber mit der Länge des Prognosezeitraumes ab. Strukturelle Ansätze erweisen sich fast nie als beste Prognosemodelle. Auch wenn ihre Qualität mit der Länge des Prognosezeitraumes zunimmt, ergeben selbst bei der 12-Monatsprognose nur in rund sieben Prozent der Zeitreihen für strukturelle Modelle die besten Prognosen.

Dabei waren die Ergebnisse der Prognose ohne weitere Vorauswahlkriterien zunächst nicht befriedigend, bei 20% der Vorprodukte ergab sich für den kurzfristigen Prognosezeitraum noch eine durchschnittliche gesamte prozentuale Abweichung von mehr als 40%. Verantwortlich hierfür sind vor allem die Zeitreihen sporadisch nachgefragter Vorprodukte bzw. die Zeitreihen mit auslaufender und unregelmäßiger Nachfrage. Die Güte der Prognose verbesserte sich, je seltener Zeiträume mit fehlender Nachfrage auftraten. Die „glättenden“ Eigenschaften der Prognoseverfahren sind nur schwer in der Lage, unregelmäßige Monatsnachfragen mit dem Wert Null zu prognostizieren. Bei den wichtigen und aktuellen Vorprodukten kommen derartige „Nullnachfragen“ kaum vor, so dass dieses Problem vor allem bei den Vorprodukten auftritt, die aufgrund einer möglicherweise veralteten Technologie nur noch als Ersatzteile nachgefragt werden. Je mehr sich eine kontinuierlich hohe Nachfrage eines Vorproduktes in der weit zurückliegenden Vergangenheit durch eine unstetige und geringe Nachfrage in der näheren Vergangenheit ersetzt, umso geringer erweist sich die Qualität der Prognoseergebnisse. Bei Einschränkung der Zeitreihen auf diejenigen, bei denen weniger als acht Werte kleiner als zehn ausfielen, fiel der Anteil der MGPA größer als 40% bei der kurzfristigen Prognose von 20 Prozent auf zwölf Prozent, bei mittelfristigen Prognosen auf 16 Prozent und bei langfristigen Prognosen auf neun Prozent.

Tabelle 10: Zusammenfassung der Ergebnisse

Prognose-zeitraum	Betrag MGPA			MGPA <(-20%)	Anteil besser als Referenzmodell	Beste drei Modelle	Anteil strukt. Modelle
	<20%	<20%	>40%				
Kurzfristig	61%	27%	12%	26%	62%	4 (46%) 2 (22%) 7 (19%)	1,8%
Mittelfristig	66%	18%	16%	7%	61%	4 (41%) 2 (30%) 7 (13%)	4,7%
Langfristig	60%	31%	9%	8%	61%	4 (38%) 2 (26 %) 7 (14%)	6,5%

Vergleicht man die unterschiedlichen Prognosezeiträume, so sind die Anteile gut prognostizierter Werte (MGPA kleiner 20%) mit 66% bei der mittelfristigen Prognose (kurzfristig 61%, langfristig 60%) am höchsten, allerdings liegen auch die schlecht prognostizierten Werte (MGPA größer 40%) bei der mittelfristigen Prognose mit 16% am höchsten (kurzfristig – 12%, langfristig – 9%). Nimmt man für das praktische Beschaffungsmanagement einen betragsmäßigen MGPA kleiner als 20% als akzeptablen Grenzwert an, so lassen sich mit Hilfe der hier beschriebenen Verfahren immerhin mehr als 60% aller Vorproduktgruppen zufriedenstellend prognostizieren. Die Prognosen der restlichen 40% der Zeitreihen sollten nur mit äußerster Vorsicht interpretiert werden. Hier bedarf es einer weiteren und tiefergehenden Analyse für die Gründe derartig großer Abweichungen. Erste Untersuchungen dazu deuten an, dass es sich hierbei nicht selten um sehr heterogene Vorproduktgruppen handelt, bei der sich hinter der Vorproduktnummer bestimmte Aggregate von Vorprodukten verbergen. So steckt hinter der Bezeichnung „Kondensatoren“ strenggenommen nicht ein bestimmtes, sondern eine Vielzahl von unterschiedlichen Typen und Technologiegenerationen von Kondensatoren. Einzelne Kondensator-Typen können aus völlig unterschiedlichen Gründen innerhalb dieser Vorproduktgruppe stark schwanken und unerwartete – und damit unvorhersehbare - Verläufe der Zeitreihe herbeiführen. Die Beschränkung der Prognose auf homogene Vorproduktgruppen erhöht die Verlässlichkeit der Prognosen signifikant.

Es fällt auf, dass mit zunehmendem Prognosezeitraum die Werte eher über- als unterschätzt werden. Sind bei der kurzfristigen Prognose noch 26% der Abweichungen 20% oder mehr zu niedrig, sind es bei der mittelfristigen Prognose nur noch 7% und bei der langfristigen Prognose 8%, während Abweichungen über 20% nach oben bei der kurzfristigen Prognose 11%, bei der mittelfristigen 13% und bei der langfristigen Prognose 30% ausmachen.

Insgesamt sollte das hier beschriebene Verfahren der algorithmischen Anwendung unterschiedlicher quantitativer Prognosetechniken nur dann angewendet werden, wenn folgende Tatbestände erfüllt sind:

- Prognosen sollten nur für disaggregierte – also für homogene Vorproduktnummern – angewendet werden. Bei der Verwendung von Vorproduktaggregaten sollte sichergestellt werden, dass sich die einzelnen Vorprodukttypen nicht allzu heterogen in der Zeitreihe entwickeln.
- Der MGPA der Zeitreihe sollte in der Vergangenheit den Wert von 20% nicht überschreiten. Dies sollte kontinuierlich überprüft werden.
- Zeitreihen mit sporadischen Nullnachfragen oder sporadischen Ausreißern nach oben (z.B. bei Ersatzbeschaffung von auslaufenden Vorprodukttypen) ergeben häufig nur zufällig gute Prognosen. Die Nachfrage dieser Zeitreihen sollte mit anderen Verfahren prognostiziert werden.
- Schiefe Verteilungen sollten nur mit Vorsicht angewendet werden. Die Gefahr einer deutlichen Unterschätzung der Nachfrage ist – insbesondere bei kurzfristigen Prognosen - groß.

- Bei Beachtung dieser Voraussetzungen, dürfte die algorithmische – und daher einfach durch den Computer zu ermittelnde – Vorgehensweise, die praktische Aufgabe der Prognose von Lagerabflüssen für einen erheblichen Teil von Verbrauchmaterialien vereinfachen.

REFERENCES:

1. Akaike, H. (1977): On entropy maximization principle, Applications of Statistics, 27-41, Amsterdam.
2. D'Agostino, R., Balanger, A., D'Agostino, R. Jr. (1990): A Suggestion for Using Powerful and Informative Tests on Normality, American Statistician, 44, 316-321.
3. Box, G.E.P.; Pierce, D.A.(1970): Distribution of the Autocorrelations in Autoregressive Moving Average Time Series Models, Journal of American Statistic Association, 65, 1509-1526.
4. Gabler, Wirtschaftslexikon (2004), 16. Aufl., Bd. 3, Wiesbaden.
5. Hannan, E.J., Rissanen, J. (1982): Recursive estimation of mixed autoregressive-moving average order, Biometrika, 69, 81-94.
6. Hillier, F. S., Lieberman, G.J. (2001): Introduction to Operations Research, 7th Ed., Boston u.a..
7. Ljung, G.M.; Box, G.E.P.(1978): On a Measure of Lack of Fit in Time Series Models; Biometrika 65, Nr. 2, 297-303.
8. Royston, P. (1991): Estimating Departure from Normality, Statistics in Medicine, 10, 1283-1293.

Abstract

Zufällig schwankende Nachfragen nach Vorprodukten bzw. Teilen und Komponenten machen die Verwendung von stochastischen Modellen der Lagerhaltung notwendig. Das vorliegende Papier beschreibt einen standardisierten algorithmischen Ansatz, mit dem der Verbrauch von Vorprodukten für die Zeiträume von drei, sechs oder zwölf Monaten mit Hilfe zeitreihenökonometrischer Verfahren prognostiziert werden kann. Im Rahmen dieses Ansatzes werden für jede Vorproduktgruppe die unterschiedlichsten quantitativen Prognosetechniken angewendet. Zu den Techniken zählen unter anderem AR-, MA-, ARMA-, ARIMA- und strukturelle Regressionsmodelle. Durch algorithmisches Vorgehen wird aufgrund von Gütekriterien (z. B. die Prognosefähigkeit in einem Testdatensatz) ein optimales Prognosemodell ermittelt, das für die Prognose des Bedarfs verwendet wird. Für alle gewählten Prognosezeiträume erwies sich das ARMA-Modell der d-differenzierten Zeitreihe als bestes Prognosemodell, gefolgt von einfachen Moving Average und ARIMA-Modellen. Die Bedeutung autoregressiver Verfahren nimmt aber mit der Länge des Prognosezeitraumes ab. Strukturelle Ansätze erweisen sich allerdings fast nie als beste Prognosemodelle, auch wenn deren Bedeutung mit der Länge des Prognosezeitraumes zunimmt. Der algorithmische Ansatz ermöglicht für einen erheblichen Teil (rund 60 Prozent) der Vorprodukte eine gute Prognosequalität. Die Güte der Prognose verbesserte sich, je seltener Zeiträume mit fehlender Nachfrage auftreten. Bei Beachtung ausgearbeiteter Voraussetzungen, dürfte diese algorithmische – und daher einfach durch den Computer zu ermittelnde – Vorgehensweise, die praktische Aufgabe der Prognose von Lagerabflüssen für einen erheblichen Teil von Vorprodukten bzw. Teilen und Komponenten vereinfachen.

Key words*: Inventory Management; Forecasting; Material Requirement Planning; Time Series*