OPTIMALE LOSGRÖSSEN BEI SERVICEGRADRESTRIKTION, UNSICHERER PRODUKTION UND STOCHASTISCHEM ABSATZ

Hermann Jahnke, Hamburg

1. Einleitung

In ihren Arbeiten beschäftigen sich [WILLIAMS] und [ZIPKIN] mit dem Problem der Bestimmung optimaler Losgrößen bei unsicherer Nachfrage und Produktion für den Fall mehrerer Produkte, die gemeinsam auf (in der einfachsten Situation) einer Anlage intermittierend gefertigt werden.

Dabei kombinieren sie Modelle der Lagerhaltungs- und der Warteschlangentheorie: Fertigungsaufträge vom Umfang der Losgröße werden ausgelöst, sobald der Lagerbestand des betreffenden Produktes eine Meldemenge erreicht. Die Fertigungsaufträge ordnen sich in eine Warteschlange vor der Fertigungsanlage ein, werden gemäß stochastischer Bearbeitungszeiten der Reihe nach bearbeitet und anschließend in ihr Lager transportiert.

Die zu optimierende Kostenfunktion ist an die Zielgrößen der Lagerhaltungstheorie angelehnt, wobei die Rolle der (zufälligen) Lieferzeit von der aus dem Warteschlangenmodell ermittelten Systemzeit (Warte- plus Bearbeitungszeit) der Aufträge bzw. Lose übernommen wird.

Beide Autoren approximieren den Ankunftstrom der Aufträge durch einen Poissonprozeß, was gerechtfertigt erscheint, wenn eine große Anzahl von Produkten betrachtet wird, die eine niedrige Auflagehäufigkeit haben. Diese niedrige Häufigkeit spricht für relativ große Lose, die wiederum zu relativ hohen Lagerbeständen führen. Eine solche Vorgehensweise ist sinnvoll, falls die Rüstkosten hoch, die Lagerkosten niedrig sind.

Die beschriebene Produktionsweise dürfte bei knappen Kapazitäten zu den schon im deterministischen Fall bekannten Problemen der Ablaufplanung (Lossequenzproblem) führen, deren Ergebnis hier die zufällige Wartezeit ist, die vergeht, bis die Abwicklung eines Fertigungsauftrages beginnt. Als Folge dürften sich hohe Fehlmengenkosten einstellen: Tatsächlich deutet WILLIAMS Bericht über simulativ durchgeführte Beispielrechnungen in solche Richtung.

Zudem werden in vielen Fällen die der Sache nach entscheidungsrelevanten und tatsächlich anfallenden Rüstkosten - wenn überhaupt zu ermitteln - vernachlässigbar gering sein, so daß man, unter Beachtung des Kapazitätseffektes von nichtverschwindenden Rüstzeiten, die Lose möglichst klein wählen wird, um die Lagerkosten niedrig zu halten. Die resultierende, eher hohe Auflagehäufigkeit der Sorten wird Konsequenzen für die Wartezeit der Aufträge und als Folge für die Fehlmengen haben. Hohe Fehlmengen(kosten) dürften aber i.d.R. für die Praxis inakzeptabel seien.

Ein Lösungsvorschlag für das Lossequenzproblem im deterministischen Fall findet sich bei [MAGEE, S.310ff]: Wähle die Losgrößen so, daß der Quotient aus Losgröße und Absatzgeschwindigkeit für alle Produkte gleich ist. Die Sorten werden dann in festem Zyklus mit gleicher Häufigkeit aufgelegt.

Stellt ein Unternehmen relativ wenige (Haupt-) Sorten her, die etwa aus den oben dargestellten Gründen eine hohe Auflagehäufigkeit haben, so liegt es auch bei Unsicherheit nahe, sie zu einem solchen Auflagezyklus zusammenzufassen. In diesem wird in vorgegebener, fester Reihenfolge jede Sorte einmal gefertigt.

Ziel dieses im folgenden sehr kurz dargestellten Ansatzes ist es, kostenminimale Losgrößen so zu wählen, daß ein vorgegebener hoher Mindestservicegrad des Lagers eingehalten wird.

2. Herleitung der losgrößenabhängigen Kosten

Diese Kosten sollen zunächst für nur eine Sorte ermittelt werden. Dabei wird angenommen, daß Nachfrage, die auf ein leeres Lager trifft, verlorengeht. Daher ist zunächst eine den Gewinn maximierende Losgröße zu bestimmen. Ferner beschränkt sich die vorliegende Arbeit auf die Betrachtung von einstufigen Produktionsvorgängen, bei denen die Herstellung eines Loses einer Sorte erst vollständig abgeschlossen sein muß, bevor - nach eventueller Umrüstung - ein Los einer anderen Sorte aufgelegt werden kann.

Die Durchlaufzeit eines Loses ist die Zeit vom Beginn der Umrüstung der Anlagen auf seine Bearbeitung bis zu seinem geschlossenen Eintreffen im Endproduktlager. Im allgemeinen wird diese Zeit nicht genau vorhergesagt werden können, vielmehr wird sie als stochastische Größe anzusehen sein [TEMPELMEIER, S. 280ff]. Mit der Durchlaufzeit ist die Zeit zwischen zwei Ankünften von Losen der gleichen Sorte im Lager, die Zwischenankunftszeit, zufällig. Im Fall der Betrachtung nur einer Sorte seien $A_1, A_2, A_3, \ldots$ stochastisch unabhängige und identischverteilte Zufallsvariablen, wobei A_i die Zeit zwischen der Ankunft des (i-1)-ten und i-ten Loses der betrachteten Sorte im Lager darstelle. Auf der anderen Seite steht der Absatz der Produkte. In der Sprache der Warteschlangentheorie entspricht er der Bedienung der Kunden: Die im Lager befindlichen Produkteinheiten werden einzeln nacheinander abgesetzt. Steht die i-te Produkteinheit zum Absatz an, so vergehen B_i Zeiteinheiten bis eine Nachfrage nach dieser Einheit auftritt und sie verkauft wird (Zwischenabsatzzeit). $B_1, B_2, B_3, \ldots$, die entsprechenden Zeiten für die Produkteinheiten 1,2,3,..., seien stochastisch unabhängige, identischverteilte Zufallsvariable. Ferner seien die A_i und B_j für alle i und j stochastisch unabhängig, $0<EA_i=1/\lambda<\infty$, $0<EB_i=1/\mu<\infty$ (E sei der Erwartungswertoperator) und die zweiten Momente mögen existieren.

Geht man von stochastisch unabhängigen und identischverteilten Zwischennachfragezeiten aus, werden die Zwischenabsatzzeiten i.a. nicht identischverteilt sein, außer in dem besonders wichtigen Fall eines poissonschen Nachfrageprozesses. Liegt kein solcher Nachfrageprozeß vor, kann man die Annahme identischverteilter Zwischennachfragezeiten als Approximation ansehen, die umso besser ist, je höher der Servicegrad des Lagers ist. Daneben ist zu bedenken, daß das Wissen der Unternehmen sich in vielen Fällen zunächst auf den Absatz und nicht auf die Nachfrage beziehen wird, so daß die Zwischenabsatzzeiten die ursprünglichen, ins Modell einzubeziehenden Größen sind.

Für eine gegebene Losgröße von y Produkteinheiten sind $C_1:=B_1+B_2+\ldots+B_y$, $C_2:=B_{y+1}+\ldots+B_{2y}, \ldots$ die Absatzzeiten der Lose, $A_1, A_2, \ldots$ und $C_1, C_2, \ldots$ also Zwischenankunfts- bzw. Bedienzeiten in einem G/G/1-Warteschlangenmodell, in dem die im Lager befindlichen, auf ihren Absatz wartenden Lose der Größe y die Rolle der Kunden spielen (Eine einfache Version eines solchen Lagerhaltungsmodelles findet sich bei [COHEN, S. 408ff] und [TIJMS, S. 94ff]). Der Warteschlangenlängenprozeß $\{L(t), t\geq 0\}$ ist bekanntlich ein regenerativer stochastischer Prozeß, bei welchem sich also der gleiche wahrscheinlichkeitstheoretische Mechanismus zu wiederholen beginnt, sobald ein ankommendes Los auf ein leeres Lager trifft (Regenerationszeitpunkt). Die erwarteten, mit einer bestimmten gewählten Losgröße verbundenen Kosten und Erlöse sind dann in jedem durch zwei solche Regenerationszeitpunkte gegebenen Regenerationsintervall gleich.

Der erwartete Gewinn im ersten Regenerationsintervall für gegebene Losgröße setzt sich aus den Deckungsbeiträgen, den Rüstkosten und den Lagerkosten dieses Intervalls zusammen. Es sei angenommen, daß im Zeitpunkt 0 ein erstes Los im Lager eintrifft, d.h. 0 ist der erste Regenerationszeitpunkt. Die Absatzzeit dieses Loses sei $C_0:=B_1 + B_2 + \ldots + B_y$ (entsprechend für die Herleitung der Gewinngröße $C_i:=B_{(i-1)y+1} + B_{(i-1)y+2} + \ldots + B_{iy}$, i=1,2,...). S sei der zufällige nächste Regenerationszeitpunkt, d.h. der Zeitpunkt, in dem nach dem Ende der ersten Betriebsperiode - Verkauf der letzten gelagerten Produkteinheit - wieder ein Los im Lager eintrifft. Ist

$\{A(t), t \geq 0\}$ der Erneuerungsprozeß zu den Zwischenankunftszeiten der Lose im Lager, so ist A(S) die Anzahl der im ersten Regenerationsintervall im Lager eingetroffenen Lose ohne das im Zeitpunkt 0 eingelagerte und inclusive des den Beginn des zweiten Regenerationsintervalls markierenden Loses.
Die Lagerkosten kann man nun durch eine gedankliche Zweiteilung des Lagerbestandes ermitteln. Sei

$$L^q(t) = \begin{cases} L(t) - 1 & \text{falls } L(t) \geq 1 \\ 0 & \text{sonst,} \end{cases}$$

so ist in der Sprache der Warteschlangentheorie $L^q(t)$ die Anzahl im Zeitpunkt t vor dem Schalter wartender Kunden oder in unserem Zusammenhang die Anzahl auf Lager liegender Lose, ohne dasjenige, das gerade Produkteinheit für Produkteinheit abgesetzt wird.
Unterstellt man nun die Existenz einer Grenzverteilung der Warteschlangenlänge $\{L(t), t \geq 0\}$, so lauten die erwarteten Lagerkosten dieses Teillagers im ersten Regenerationsintervall bei einem Lagerkostensatz von l Geldeinheiten pro Mengen- und Zeiteinheit und bei Losgröße $y \in \mathbb{N}$

$$E\left(\int_0^S y\, l\, L^q(t)\, dt\right) = ES\, [y\, l(EL-\rho)],$$

wobei EL die erwartete Warteschlangenlänge unter der Grenzverteilung, ES die erwartete Länge des Regenerationsintervalls und $\rho = EC_1/EA_1$ die Verkehrsintensität des Warteschlangenmodelles ist. (Es sei $\rho<1$ angenommen, woraus bekanntlich $ES<\infty$ folgt.)
Bis zum Zeitpunkt des Absatzes der letzten Mengeneinheit des letzten Loses im ersten Regenerationsintervall befindet sich jeweils ein Los in "Bedienung", d.h. es wird abgesetzt. Dieses Los bildet den gedanklichen zweiten Teil des Lagers.
Betrachten wir das im Zeitpunkt 0 eintreffende Los. Bis zum Absatz der ersten Produkteinheit vergehen B_1 Zeiteinheiten, bis zu diesem Zeitpunkt ist das ganze Los - y Mengeneinheiten - auf Lager. Anschließend vergehen weitere B_2 Zeiteinheiten bis zum Absatz der nächsten Produkteinheit, während welcher (y-1) Mengeneinheiten auf Lager liegen und so fort, bis nach $B_1 + B_2 + ... + B_y$ Zeiteinheiten das erste Los vollständig abgesetzt ist. Dieser Vorgang wiederholt sich mit dem nächsten Los, bis alle A(S) Lose im ersten Regenerationsintervall abgesetzt sind.

Mit $$Z_i := \sum_{j=1}^{y}(y-j+1)B_{iy+j} \qquad i=0,1,...,A(S)-1$$

erhält man die durch das sich jeweils im Absatz befindende Los verursachten erwarteten Lagerkosten, da A(S) eine Stopzeit für $Z_0, Z_1, ...$ ist, mit der Waldschen Gleichung [HEYMAN/SOBEL, S.171] als

$$E\left(l \sum_{i=0}^{A(S)-1} Z_i\right) = (l/2)ES\,(y\rho+\rho),$$

so daß die gesamten erwarteten Lagerkosten im ersten Regenerationsintervall $l\, ES[y(EL-(\rho/2))+\rho/2]$ sind.

Die Rüstkosten des ersten Regenerationsintervalls ergeben sich aus der Anzahl in diesem Zeitraum aufgelegter Lose A(S) multipliziert mit den pro Rüstvorgang anfallenden $R \geq 0$ Geldeinheiten. Da $EA(S) = \lambda ES$ ist (A(S) ist Stopzeit für A_1, A_2,... und $S=A_1+ \ldots +A_{A(S)}$), folgt für die erwarteten Rüstkosten $E(R \cdot A(S)) = R\lambda ES$.
Die Anzahl im ersten Regenerationsintervall abgesetzter Lose A(S) multipliziert mit der Losgröße y ergibt die Absatzmenge, für deren Erwartungswert gilt (Beachte: $EC_1=yEB_1$!) $E(yA(S)) = y\lambda ES = \mu\rho ES$, so daß der erwartete gesamte Deckungsbeitrag des ersten Regenerationsintervalls bei einem Deckungsbeitrag von d Geldeinheiten pro Produkteinheit $d\mu\rho ES$ Geldeinheiten beträgt.
Sowohl der Erwartungswert von C_1, als auch derjenige von A_1, ferner ES hängen von der gewählten Losgröße ab. Da insbesondere die erwartete Länge des ersten Regenerationsintervalles die Wahl der optimalen Losgröße nicht beeinflussen soll, ist das geeignete Kriterium der erwartete Gewinn im ersten Intervall, dividiert durch ES, also

$$G(y) = d\mu - (1-\rho)d\mu - yl(EL-(\rho/2)) - (l\rho/2) - \lambda R.$$

Die ersten beiden Summanden haben dabei folgende Interpretation: Absatz kann nur solange stattfinden, bis das Lager leer ist, d.h. bis zum Augenblick des letzten Absatzvorgangs. Die erwartete Länge dieser Betriebsperiode ist (A(S) ist Stopzeit für C_0, C_1,...)

$$E(\sum_{i=0}^{A(S)-1} C_i) = \rho ES.$$

ρ ist also der erwartete Anteil des Regenerationsintervalles, während dessen das Lager nicht leer ist. Entsprechend ist $(1-\rho)ES$ die mittlere Zeit, die zwischen dem Absatz der letzten Produkteinheit und dem Eintreffen eines neuen Loses im leeren Lager (Regenerationszeitpunkt) vergeht, also die erwartete Leerzeit des Lagers, während der auftretende Nachfrage verloren geht.
Der erwartete Absatz ist demnach gleich dem Produkt aus mittlerem Absatz pro Zeiteinheit, falls das Lager nicht leer wird (μ) und der mittleren Zeit, während der das Lager nicht leer ist (ρES). Entsprechend ist μES der erwartete Absatz im ersten Regenerationsintervall, falls das Lager zu keinem Zeitpunkt leer wäre, $\mu(1-\rho)ES$ entsprechend der durch Leerstand des Lagers im Mittel entgangene Absatz. Der zweite Term im Gewinnausdruck oben stellt die pro Zeiteinheit erwartete Fehlmenge, gewichtet mit dem Deckungsbeitrag, also die erwarteten Fehlmengenkosten pro Zeiteinheit dar.
An dieser Stelle ergibt sich eine Interpretation der Verkehrsintensität ρ. Nach dem eben Gesagten ist $\rho = (\rho\mu ES)/(\mu ES)$, also gleich dem Verhältnis von erwartetem Absatz zur insgesamt auftretenden erwarteten Nachfrage im Regenerationsintervall, ρ ist also ein Servicegrad des Lagers [NEUMANN, S. 273].
Äquivalent zur Maximierung des pro Zeiteinheit erwarteten Gewinnes ist die Minimierung der Summe der pro Zeiteinheit erwarteten Lager-, Rüst- und Fehlmengenkosten:

$$\lambda R + ylEL - yl(\rho/2) +(l/2)\rho + d\mu(1-\rho) \tag{2.1}$$

Wir hatten die Existenz der Grenzverteilung der Warteschlangenlänge unterstellt, die wir benötigten, um in (2.1) ihren Mittelwert EL für gegebene Losgröße y ausrechnen zu können. Diese Existenz ist gesichert, falls $\rho<1$ ist (und S eine nichtarithmetische Verteilung hat, wovon im weiteren ausgegangen werden soll) [HEYMAN/SOBEL, S. 183]. Im allgemeinen läßt sich für ihren Erwartungswert jedoch kein geschlossener Ausdruck angeben, mit dem sich sinnvoll arbeiten ließe. Er soll daher im folgenden approximiert werden, wozu unterstellt wird, daß ein hoher Mindestservicegrad vorgegeben ist. Man wird aber den Servicegrad auch nicht zu hoch wählen: für $\rho \geq 1$ wächst bekanntlich mit einer sich 1 annähernden Wahrscheinlichkeit der Lagerbestand über jede gegebene Grenze hinaus, das Lager "explodiert" [SCHASSBERGER, S. 72]. Auf Dauer ökonomisch sinnvoll sind also nur Servicegrade kleiner als, aber nahe bei 1.

Diese Situation "starken Verkehrs" (bzw. von heavy traffic) ist aber gerade diejenige, in der sich der stochastische Prozeß $\{L(t), t \geq 0\}$ durch eine geeignete Brownsche Bewegung approximieren läßt [HEYMAN/SOBEL, S. 478]. Die Brownsche Bewegung hat wieder eine Grenzverteilung, deren Erwartungswert man als Näherungswert für die im Mittel zu erwartende Anzahl gelagerter Lose im eigentlichen G/G/1-Modell (EL) bei hohem Servicegrad nutzen kann:

$$EL \approx (1/[2\rho(1-\rho)(EA_1)^2])[(\mathrm{Var}\ C_1/\rho)+\rho^2 \mathrm{Var}\ A_1)] \qquad (2.2)$$

[HEYMAN/SOBEL, S. 482], wobei Var für den Varianzoperator steht.

Dabei ist für gegebene Losgröße $y \in \mathbb{N}$ $EC_1 = y\ EB_1 = y/\mu < \infty$ und $\mathrm{Var}\ C_1 = y\ \mathrm{Var}\ B_1 < \infty$. Nicht so leicht lassen sich Erwartungswert und Varianz von A_1 (Zwischenankunftszeit der Lose) herleiten, da sie natürlich von den jeweiligen betrieblichen Gegebenheiten abhängen.

Um zu konkreten Aussagen zu gelangen, sollen im weiteren spezielle Annahmen über den Produktionsvorgang getroffen werden. Es sei betont, daß hier auch andere Annahmen möglich sind, die sich auf ähnliche Weise behandeln lassen.

I.d.R. wird man für ein Produkt aus der Herstellungspalette eines Betriebes eine minimale, technisch determinierte Bearbeitungszeit (ggf. genommen über alle Produktionsstufen) angeben können, von der jedoch i.a. die Durchlaufzeit nach oben abweicht.

Die Abweichung zwischen der Summe der technisch determinierten Bearbeitungszeiten und der Durchlaufzeit eines Loses wird i.a. von der Losgröße abhängen. Üblicherweise wird man die Verteilung der Abweichung nicht kennen, so daß diese Abhängigkeit im Modell nur über die Momente ihrer Verteilung erfaßt werden kann. Sei S_y die zufällig Abweichungszeit für ein Los der Größe y. Sie sei für je zwei Lose stochastisch unabhängig.

Im weiteren sei davon ausgegangen, daß es für die einzelnen Produkteinheiten eine zufällige Abweichungszeit (etwa Zeit für Qualitätsprüfung und Nacharbeitung) von X_i Zeiteinheiten für die i-te Produkteinheit gibt. Dabei sei für alle i $X_i \geq 0$ fast sicher und $0 < EX_1 = EX_i = a < \infty$. Ferner sei $b := \mathrm{Var}\ X_1 \in (0,\infty)$. Damit gilt $ES_y = ay$.

Die Abweichungszeiten der einzelnen Einheiten eines Loses werden nicht unabhängig voneinander auftreten: Die Abnutzung des Werkzeugs, Maschinendefekte, allgemein mangelnde Faktorqualität wirken nicht nur auf einzelne, sondern systematisch auf Gruppen von oder gar alle Produkteinheiten. Daher ist nicht mit einer mit wachsender Losgröße sinkenden Variabilität der Summe aus technisch determinierter Bearbeitungs- und Abweichungszeit zu rechnen. Vielmehr soll im weiteren angenommen werden, daß der Variationskoeffizient dieser Summe gleich bleibt, d.h. es soll für jede Losgröße $\mathrm{Var}\ A_1 = by^2$ gelten.

3. Das Modell für die losweise Fertigung mehrerer Sorten auf einer Anlage

Im Fall mehrerer (Haupt-) Sorten sollen alle herzustellenden Sorten zu einem Zyklus zusammengestellt werden, in dem jede Sorte genau einmal aufgelegt wird (einfache Reihenfolge).
Die weiteren Überlegungen beziehen sich auf den einfachsten Fall n=2. Die Ergebnisse lassen sich aber bei einfachen Reihenfolgen ohne weiteres verallgemeinern. Die denkbaren einfachen Reihenfolgen sind (y_1,y_2) oder (y_2,y_1), wobei y_i die Losgröße der Sorte i bezeichne, und o.E. sei (y_1,y_2) die bei Berücksichtigung von reihenfolgeabhängigen Rüstzeiten günstigere. Der Absatz der beiden Sorten sei unabhängig voneinander, so daß wir, falls Produkteinheiten auf Lager sind, stochastisch unabhängige Zwischenabsatzeiten B_{11}, B_{12}, B_{13},... und B_{21}, B_{22}, B_{23},..., bzw. auf die Lose bezogen von C_{11}, C_{12}, C_{13},... und C_{21}, C_{22}, C_{23},... vorliegen haben.
Es sei wieder $EB_{il}=1/\mu_i$ bzw. $EC_{il}=y_i/\mu_i$, i=1,2. Auf der Produktionsseite stellt sich die Frage nach der Zeit zwischen der Ankunft zweier Lose der gleichen Sorte im Lager.
Unter den Annahmen des zweiten Abschnittes gilt folgende Überlegung: Nehmen wir an, ein Los der Sorte 2 sei gerade im Lager angekommen. Nach einer eventuell verschwindenden Ruhezeit von τ_2 Zeiteinheiten wird die Anlage in $r_1>0$ Zeiteinheiten auf die Bearbeitung der Sorte 1 umgerüstet, diese nimmt bei einer technisch determinierten Produktionsgeschwindigkeit von $v_1>0$ pro Mengeneinheit $1/v_1$ Zeiteinheiten in Anspruch und wird durch die zufällige Abweichungszeit S_1 verlängert. Nachdem die Sorte 1 die Produktion verläßt, fallen wieder Ruhezeiten von τ_1 und Rüstzeiten von $r_2>0$ Zeiteinheiten an, bevor die Bearbeitung von Sorte 2 beginnt, die nach y_2/v_2+S_2 Zeiteinheiten beendet wird. D.h. bis zur nächsten Ankunft eines Loses der Sorte 2 vergehen

$$y_1/v_1 + S_1 + y_2/v_2 + S_2 + z \tag{3.1}$$

oder im Mittel $1/\lambda(y_1,y_2) = \vartheta_1 y_1+\vartheta_2 y_2+z$ Zeiteinheiten, wobei $\vartheta_i=1/v_i+a_i$, i=1,2 und $z=\tau_1+r_1+\tau_2+r_2$ sei. Wie in Abschnitt 2 sei ferner $ES_i=a_i y_i$, i=1,2, für $a_1>0$ und $a_2>0$ und darüber hinaus S_1 und S_2 stochastisch unabhängig. Durch (3.1) werden die dem ersten Abschnitt entsprechenden Zwischenankunftszeiten (von Losen der Sorte 2 im Lager) A_{21}, A_{22}, A_{23},... als identischverteilte, stochastisch unabhängie Zufallvariable gegeben, Produktion und Absatz dieser Sorte können also durch ein G/G/1-Warteschlangenmodell wie in Abschnitt 2 beschreiben werden.
Etwas anders ist jedoch die Situation für die erste Sorte. Zwar ergibt sich für sie aus (3.1) eine Folge von identischverteilten, stochastisch unabhängien Zufallsvariablen A_{12}, A_{13}, A_{14},... (mit dem obigen Erwartungswert). Jedoch unterscheidet sich hiervon die Zeit A_{11} bis zur Ankunft des ersten Loses dieser Sorte im Lager: Sie beträgt nur $A_{11}=y_1/v_1+S_1+r_1(+\tau_1)$ Zeiteinheiten. A_{11} ist zwar von A_{12}, A_{13},... stochastisch unabhängig, hat jedoch i.a. nicht die gleiche Verteilung. Nun ist aber die Grenzverteilung des stochastischen Prozesses $\{L_1(t), t\geq 0\}$ in dieser Situation von A_{11} unabhängig, solange diese nur eine nichtnegative Zufallsvariable und ihr Erwartungswert endlich ist. Andererseits erfaßt (2.1) nur die Kosten ab dem frühesten Zeitpunkt, zu dem ein Los im Lager eintrifft, d.h. das Kostenkriterium ist unabhängig von A_{11} und wir können daher das tatsächliche A_{11} durch eine wie A_{12} verteilte Zufallsvariable ersetzen. Dann sind A_{11}, A_{12}, A_{13},... stochastisch unabhängige und identischverteilte Zufallsvariable und wir erhalten wie für Sorte 2 eine Beschreibung durch ein G/G/1-Modell.
Die Verkehrsintensität der beiden Sorten

$$\rho_j(y_1,y_2)=y_j / \mu_j(\vartheta_1 y_1+\vartheta_2 y_2+z) \qquad j=1,2 \tag{3.2}$$

läßt sich wieder als Servicegrad des Absatzlagers für das jeweilige Produkt j interpretieren. Der erwartete

Lagerbestand beider Produkte unter der stationären Verteilung, gemessen in der Anzahl auf Lager befindlicher Lose EL_1 (Produkt 1) und EL_2 (Produkt 2), soll wieder durch (2.2) approximiert werden. Dazu sei angenommen, daß für beide Servicegrade eine knapp unter 1 liegende, untere Grenze gegeben sei.

Unterstellt man $\mu_j<\vartheta_j^{-1}$ (D.h. der erwartete Absatz pro Zeiteinheit ist kleiner als die bei Vernachlässigung von Rüst- und Ruhezeiten z erwartete Produktionsmenge der Sorte j pro Zeiteinheit. Diese Annahme korrespondiert mit der im Rahmen der Herleitung der klassichen Losgrößenformel üblichen Prämisse, daß die Produktionsgeschwindigkeit größer als die Absatzgeschwindigkeit sei.), so ergeben sich aus der Forderung $\rho_j(y_1,y_2)<1$, $j=1,2$, obere Schranken für zulässige Losgrößen:

$$y_1 < (z\mu_1+\mu_1\vartheta_2 y_2) \;/\; (1-\mu_1\vartheta_1) \tag{3.3}$$

$$y_1 > ((1-\mu_2\vartheta_2)y_2 \,/\, \mu_2\vartheta_1) - z/\vartheta_1 \tag{3.4}$$

Wegen $z>0$ (Wir unterstellen ja $r_i>0$, $i=1,2$!), haben die beiden zugehörigen Begrenzungsgeraden einen Schnittpunkt im positiven Quadranten, falls $\mu_1\vartheta_1+\mu_2\vartheta_2<1$ ist.

Nun ist μ_j annähernd die pro Zeiteinheit vom Produkt j ($j=1,2$) abgesetzte Menge und der obige Ausdruck die zu ihrer Herstellung (incl. Prüfung und eventueller Nachbearbeitung) notwendige Zeit, wobei Rüst- und Ruhezeiten vernachlässigt werden. Stände also oben "$\geq$" an Stelle der "$<$"-Relation, so wären die geforderten hohen Servicegrade eventuell nicht aufrecht zu erhalten, da ja tatsächlich Rüstzeiten anfallen. Daher sei im weiteren $\mu_1\vartheta_1+\mu_2\vartheta_2<1$ unterstellt, woraus sofort wegen $\mu_j\vartheta_j>0$ $\mu_j\vartheta_j<1$, $j=1,2$, folgt. Untere Schranken für y_1 und y_2 ergeben sich aus der jeweiligen Einhaltung eines Mindestservicegrades $0<\rho_j^-<1$, $j=1,2$:

$$y_1 \geq (\rho_1^-\mu_1 z + \rho_1^-\mu_1\vartheta_2 y_2) \;/\; (1-\rho_1^-\mu_1\vartheta_1) \tag{3.5}$$

$$y_1 \leq -z/\vartheta_1 + ((1-\rho_2^-\mu_2\vartheta_2 y_2) \,/\, \rho_2^-\mu_2\vartheta_1) \tag{3.6}$$

Wie oben folgt wieder die Existenz eines Schnittpunktes der zugehörigen Begrenzungsgraden im positiven Quadranten aus $\mu_1\vartheta_1+\mu_2\vartheta_2<1$.

Für beide Sorten ergibt sich (unter Annahmen analog Abschnitt 2) für die Varianz der Zwischenankunftszeiten wegen der stochastischen Unabhängigkeit der jeweiligen Abweichungszeiten $\mathrm{Var}\, A_{j1}=b_1y_1^2+b_2y_2^2$, $j=1,2$ mit $b_1>0$, $b_2>0$. Setzt man die gewonnenen Ausdrücke in die Approximation (2.2) für EL_1 und EL_2 ein, diese wiederum in den Kostenausdruck für die beiden Sorten nach (2.1) und summiert über die Sorten, so erhält man die unter den Nebenbedingungen (3.3) bis (3.6) zu minimierende Kostenfunktion. Fordert man - was allerdings nicht notwendig ist und hier nur der einfacheren Darstellung dient - einen gleichen Servicegrad für beide Sorten, also mit $\eta:=(y_1,y_2)$: $\rho_1(\eta)=\rho_2(\eta)=\rho(\eta)$, so lautet sie

$$\begin{aligned} K(\eta) = {} & (\alpha_1+\alpha_2)\Phi(\eta) + (\beta_1+\beta_2)/\Phi(\eta) + (l_1+l_2)\rho(\eta)/2 + \\ & + ((\gamma_1+\gamma_2+\gamma_3+\gamma_4)\rho^4(\eta))/\Phi(\eta)\xi(\eta) - (\delta_1+\delta_2)\rho^2(\eta)/\xi(\eta) \end{aligned} \tag{3.7}$$

mit $\Phi(\eta):=1-\rho(\eta)$, $\xi(\eta):=1-\mu_1\vartheta_1\rho(\eta)-\mu_2\vartheta_2\rho(\eta)$, $\alpha_i:=\mu_i d_i$, $\beta_i:=(l_i\mu_i^2\ \mathrm{Var}\ B_{1i})/2$, $\gamma_i:=(l_i b_i\mu_i^3 z)/2$,

$\delta_i:=(l_i\mu_i z)/2$ für $i=1,2$ und $\gamma_3:=(l_1 b_2\mu_1\mu_2^2 z)/2$, $\gamma_4:=(l_2 b_1\mu_1^2\mu_2 z)/2$. Dabei wurden wegen der in der Einleitung erwähnten Probleme verschwindende Rüstkosten unterstellt.
Da die Anzahl ganzzahliger Losgrößen, die (3.3) bis (3.6) erfüllen - falls es überhaupt solche gibt - , endlich ist, folgt die Existenz eines (3.7) minimierenden η^*, so daß man auf den üblichen Wegen durch Ausnutzen der Eigenschaften von K eine optimale Lösung finden kann.

4. Diskussion

Die feste Sortenreihenfolge im obigen Ansatz erlaubt es, den Kapazitätseffekt der Rüstzeiten zu berücksichtigen, Losgrößen also sinnvoll auch bei zu vernachlässigenden Rüstkosten zu ermitteln, ferner unter Einhaltung von Servicegradrestriktionen diese kostenminimal zu wählen: Die feste Produktionsfolge der Sorten macht das Fertigungs- und Absatzsystem berechenbar. Allerdings führt sie im Vergleich mit dem in der Einleitung dargestellten Meldemengensystem auch zu einer gewissen Starrheit, so daß man vermuten könnte, sie führte auf im Mittel höhere Lagerbestände.
Nun sind die interessierenden Größen eines Systems mit Meldemenge m_i und festen Losgrößen y_i (i=1,2) ungleich schwerer zu berechnen: Man benötigt die Verteilung der Systemzeit (siehe Einleitung), die man i.d.R. nicht kennt. Die Argumente für die von WILLIAMS und ZIPKIN benutzten Approximationen gelten aber im Falle relativ weniger Sorten mit hoher Auflagehäufigkeit nicht ohne weiteres. Ein exakter Vergleich beider Produktionsweisen scheint daher nicht möglich.
Aber immerhin kann man die aus dem Ansatz mit fester Sortenreihenfolge errechenbaren mittleren Lagerbestände mit solchen vergleichen, die für entsprechende Meldemengensysteme durch Simulationsuntersuchungen ermittelt werden.
Zwar stehen umfangreiche Untersuchungen hierzu noch aus, erste Ergebnisse deuten jedoch in eine andere Richtung als die geäußerte Vermutung: Um gleichhohe Servicegrade zu garantieren wie im System mit fester Sortenreihenfolge, scheinen hohe Meldemengen nötig zu sein, so daß die mittleren Lagerbestände des Meldemengensystemes die Vergleichsbestände bei fester Reihenfolge unter Umständen sogar übersteigen.

Literatur

Cohen, J.W. — The Single Server Queue, Amsterdam, London 1969

Heyman, D.P., Sobel, M.J. — Stochastic Models in Operations Research, Vol. I, New York, St. Louis 1982

Magee, J.F. — Production Planning and Inventory Control, New York, London 1958

Neumann, K. — Operationsresearchverfahren, Bd. II, München, Wien 1977

Tempelmeier, H. — Materiallogistik, Berlin, Heidelberg 1988

Tijms, H.C. — Stochastic Modelling and Analysis, Chichester, New York 1986

Williams, T.M. — Special products an uncertainty in production/inventory systems, Europ. J. Opl. R. 15(1984), 46-54

Zipkin, P.H. — Models for Design and Control of Stochastic, Multi-Item Batch Production Systems, OR, Vol. 34, 1986, 91-104