



ANÁLISE E PROPOSIÇÕES SOBRE O BALANCEAMENTO E USO DE EXCESSO DE CAPACIDADE EM RECURSOS PRODUTIVOS

Fernando Bernardi de Souza

Departamento de Engenharia Mecânica
EESC/USP – NUMA
E-mail: bernardi@sc.usp.br

Sílvio R. I. Pires

Faculdade de Engenharia Mecânica e de Produção
UNIMEP – NUMA
E-mails: sripires@unimep.br, sripires@sc.usp.br

Resumo

A busca por uma alta eficiência dos recursos produtivos sempre foi um preocupação básica da maioria dos gerentes de manufatura, em suas constantes buscas por menores custos de produção. Surgida e baseada nos princípios da produção em massa, nas últimas décadas tal prática tem se mostrado bastante inadequada e se convertido em um forte entrave para uma gestão da produção baseada nas reais dimensões competitivas requeridas pelo mercado atual. Este artigo discute a questão sob a perspectiva do balanceamento e uso de excesso de capacidade em recursos produtivos, apresentando uma abordagem quantitativa com o propósito de ilustrar e enriquecer as análises e proposições realizadas. Os argumentos expostos evidenciam e reforçam a importância de se tratar o balanceamento de capacidade em recursos produtivos de forma integrada, em detrimento à tradicional abordagem voltada para a otimização da eficiência de recursos isolados.

Palavras-chave: *gestão de capacidade, balanceamento de carga, estratégia de manufatura.*

1. Introdução

Durante muito tempo, a busca por uma alta eficiência dos recursos produtivos, como forma de maximizar o rendimento operacional,

tem sido uma constante no dia-a-dia da imensa maioria das empresas de manufatura. Tal prática tem sua origem junto com o surgimento da chamada produção em massa, no início deste século, quando a competição era ainda baseada

essencialmente nos baixos custos produtivos. Com o passar dos anos, o nível de competitividade na manufatura tem crescido muito e novas e crescentes dimensões competitivas têm surgido e se tornado tão importantes (ou mais) do que os tradicionais custos produtivos. Assim, dimensões como a qualidade, a flexibilidade e o desempenho das entregas (englobando a pontualidade e a confiabilidade dos prazos acordados) têm se convertido em importantes elementos de uma estratégia competitiva e de manufatura (SLACK, 1991; PIRES, 1995). Conseqüentemente, um número cada vez maior de publicações têm alertado que altos níveis de utilização dos recursos produtivos geralmente causam impactos negativos em outros importantes indicadores de desempenho atuais, como os tempos de ciclo, os níveis de estoques em processo e nível de confiabilidade no atendimento dos prazos de pedidos. Em termos práticos, uma forma de se reduzir os níveis de utilização, melhorando o desempenho das medidas citadas, sem comprometer o volume de produção (*throughput*) planejado é aumentando a capacidade de produção de alguns recursos, aceitando que estes operem com um certo nível de ociosidade (ver STEELE & PAPKE-SHIELDS, 1993; ATWATER & CHAKRAVORTY, 1994). Estudar como isto ocorre, analisar a questão com mais detalhes e propor algumas formas de alocar capacidade dentro de uma linha de produção são objetivos básicos deste texto. Também, pela própria natureza do tema tratado, o foco deste volta-se para a realidade das empresas de manufatura que produzem de forma intermitente, ou seja, para aquelas em que a questão do balanceamento de capacidades é bem mais crítica.

2. A Questão do Balanceamento de Capacidade dos Recursos Produtivos

É prática comum nas empresas de manufatura a busca constante pelo balanceamento das capacidades produtivas dos seus recursos, ou seja, procura-se alcançar uma situação em que a capacidade de todos os recursos esteja aproxi-

madamente equilibrada entre si e com a demanda de mercado. Tal prática é justificada pela tentativa de se manter todos os recursos continuamente ativados, amortizando os investimentos realizados e procurando estabelecer um fluxo uniforme de materiais pela fábrica. Desta forma, evitar-se-ia que a empresa perdesse vendas por falta de capacidade de seus recursos ou, no caso de haver excesso de capacidade, que a mesma operasse com baixas eficiências. Esta última situação é vista como uma ótima oportunidade de redução de custos a partir da venda deste excesso de capacidade. Deve-se sublinhar, no entanto, que a permanente procura por uma linha de produção balanceada advém de causas mais profundas localizadas principalmente nas práticas convencionais da contabilidade de custos. Este modo de pensar está fundamentado na idéia de que ao se utilizar todos os recursos ao máximo estar-se-ia ganhando e economizando mais dinheiro. Como os sistemas de produção são não balanceados por natureza, uma forma de se conseguir altas eficiências é pela manutenção de altos estoques em processo. Mas como altos estoques também se constituem em uma situação indesejável, a única saída parece apontar para o balanceamento da capacidade dos recursos produtivos (SOUZA, 1997).

Mesmo as empresas que operam com filosofias atuais de administração como o *Just in Time*, cuja premissa básica é trabalhar de acordo com as necessidades do mercado, produzindo sempre em resposta ao “o que”, “quanto”, e “quando” o cliente deseja determinado produto, também lutam por um balanceamento de suas linhas de fabricação. Para tal, essas empresas fazem uso de uma mão-de-obra multifuncional para socorrer os postos de trabalho com falta de capacidade produtiva, equilibrando assim a capacidade desses postos.

A questão natural que surge então diz respeito ao por que tanto tempo e tantos recursos são gastos com insistentes tentativas de balancear a fábrica, mas quando finalmente todos os recursos parecem estar com suas capacidades balanceadas o resultado operacional fica muito

aquém do esperado, a produção não alcança o índice preestabelecido no programa, o *lead time* é maior que o desejado e o número de pedidos não atendidos ou atendidos fora do prazo cresce muito. Geralmente, a culpa pelo fracasso no desempenho da fábrica fica por conta de fatores externos e fora do controle da empresa, como fornecedores “indisciplinados”, mudanças nos prazos e especificações dos pedidos, imprevisibilidade da demanda, entre outros.

Na realidade, boa parte da explicação de tudo isso está, segundo GOLDRATT & COX (1986), na existência de dois fenômenos presentes em todos os ambientes de produção. Estes, apesar de extremamente conhecidos por todos aqueles que trabalham com processos de manufatura, não são devidamente considerados quando agem em conjunto. O primeiro deles refere-se ao fato de que as tarefas de manufatura possuem uma seqüência de operações que são interdependentes, ou seja, uma operação tem de ser realizada antes que outra possa começar. A tal fenômeno se dá o nome de “eventos dependentes”, isto é, um evento subsequente só pode realizar-se depois que outras atividades anteriores a ele ocorram. A seqüência de operações estabelecida no roteiro de produção de um certo item é um típico exemplo de dependência na manufatura. Desta forma, determinadas atividades só podem começar quando certas peças específicas necessárias para aquela operação estejam disponíveis. Estas peças, por sua vez, necessitam ser processadas por um determinado recurso que novamente dependerá do abastecimento de material de um outro recurso e assim por diante. Um outro tipo de dependência também muito importante em ambientes fabris são as dependências decorrentes de interações entre produtos que disputam um mesmo recurso. Situações deste tipo costumam ser tratadas nas questões referentes aos problemas de seqüenciamento de operações.

O segundo fenômeno está relacionado com a imprevisibilidade de certos eventos. Eventos incertos irão sempre ocorrer em sistemas como os de produção, de tal forma que qualquer

operação da fábrica estará sujeita às denominadas “flutuações estatísticas”. “Flutuações estatísticas” resultam de diversos acontecimentos que não podem ser previstos com exatidão, como pontualidade e qualidade de materiais comprados, quebra de máquinas, disponibilidade de ferramentas, absenteísmo, ausência de material para processar, capacidade do equipamento ou variações no comportamento do operador. Como conseqüência, o tempo de execução de uma mesma operação pode variar aleatoriamente ou de acordo com alguma recorrência estatística e, desta forma, os *lead times* das tarefas serão na verdade valores médios ou esperados dos tempos de processamento. Com o objetivo de minimizar tais incertezas, as empresas lançam mão de diversas ferramentas provenientes principalmente do movimento da Qualidade Total, como o Controle Estatístico de Processo em conjunto com certos diagramas de causa-efeito (diagrama “6 M” ou “Espinha de Peixe”), Manutenção Produtiva Total, padronização de métodos de trabalho, entre outras. Contudo, a menos que se tenha uma linha totalmente robotizada e com confiabilidade de 100%, sempre existirão eventos imprevistos. Pode-se afirmar que as empresas simplesmente não possuem recursos suficientes para controlar todos os componentes aleatórios de todas as suas operações. Ainda que para cada recurso, visto isoladamente, a soma de suas “flutuações estatísticas” tenda a zero, quando examinada a partir de uma cadeia seqüencial de operações (eventos dependentes), cada variação acima da média em qualquer elemento do sistema tende a não ser absorvida pelo sistema como um todo, enquanto que as flutuações abaixo da média são propagadas ao longo da cadeia. Assim, eventuais adiantamentos de determinadas tarefas não serão compensadas pelos atrasos em outras e o desvio médio em toda a seqüência de operações será negativo.

Pode-se ilustrar essa situação supondo-se uma linha de produção composta por dois recursos (A e B) sendo o recurso B alimentado por peças provenientes de A. Logo, B é

dependente de A. Considera-se também que as condições de processamento de cada um dos recursos sejam caracterizadas por uma distribuição uniforme 10 ± 4 peças por dia. A pergunta é se no final de 10 dias a linha será capaz de produzir 100 peças (10 peças/hora x 10 dias), uma vez que cada recurso tem capacidade para, em média, produzir 10 peças por dia. A Tabela 1 ilustra uma possível situação resultante de uma simulação dirigida das condições apresentadas. Os valores, em número de peças, apresentados em cada coluna da tabela, representa a capacidade de produção do recurso naquele dia, e não o quanto esse recurso efetivamente produziu. Os valores contidos na tabela foram escolhidos pelos autores. Exemplos como esse podem ser encontrados em HOPP & SPEARMAN (1996), UMBLE & SRIKANTH (1990) ou CORRÊA & GIANESI (1993).

Nota-se pela Tabela 1 que, ainda que cada recurso tenha tido capacidade de produzir 100 peças em 10 dias, a linha como um todo produziu apenas 80 peças. Isto porque nos 5 primeiros dias o recurso B estava operando mais rápido que A, só que, por falta de peças, sua produção caiu ao nível do recurso mais lento A. Já nos últimos 5 dias, o recurso A estava mais rápido, porém esta rapidez não pôde ser completamente aproveitada pela linha devido agora à baixa capacidade apresentada pelo recurso B.

GOLDRATT & COX (1986) argumentam que com o intuito de alcançar o balanceamento desejado, as empresas procuram, tanto quanto possível, eliminar capacidades consideradas em excesso em determinados locais e, assim, evitarem que qualquer recurso fique parado. No entanto, segundo os autores, quanto mais próximo se chega de uma fábrica balanceada mais perto se está da falência, pois, em termos da meta, tal atitude não resultará num aumento do ganho (ou seja, das vendas) enquanto que o inventário e a despesa operacional também não diminuirão. Esses autores citam ainda que devido à combinação dos dois fenômenos anteriormente citados, pode-se provar matemati-

camente que quando a “capacidade produtiva é reduzida exatamente até a demanda de mercado, nem mais nem menos, o ganho cai e o inventário aumenta até o teto”. E, como o inventário aumenta, o custo de manter o inventário, que é despesa operacional, sobe.

Todavia, os efeitos comentados acima não acontecem, na realidade, exatamente como foi demonstrado na simulação. É possível que a administração, ao perceber que seus níveis de inventário estão aumentando e que a produção esperada não está sendo alcançada, tome medidas que acabem levando ao desbalanceamento da linha, como uso de horas extras, por exemplo.

Muitos dirigentes industriais costumam gastar boa parte do tempo procurando equilibrar a capacidade de seus recursos produtivos. No entanto, quando a empresa finalmente consegue alcançar um estágio de balanceamento que poderia se chamar de “desejável”, a fábrica começa a atrasar seus pedidos à medida que aumenta os níveis de estoque em processo. “As conseqüências de uma fábrica perfeitamente balanceada, com nenhum excesso de capacidade, não são bem entendidas. Em uma fábrica de manufatura, excesso ou capacidade não utilizada num dado recurso é geralmente visto como custo em excesso. Em sistemas tradicionais de custeio, este excesso de custo é freqüentemente o foco de projetos de redução de custos. Em uma vã tentativa de minimizar os custos em cada processo/recurso, muitos gerentes de produção gastam um tempo considerável tentando balancear as capacidades dos recursos em suas fábricas. Mas quando o excesso de capacidade é eliminado, a capacidade de reação dos vários recursos desaparecem e o inevitável acontece. A fábrica começa a ficar atrasada com relação ao plano de produção enquanto que inventários de produto em processo aumentam e o ganho fica prejudicado. A culpa pelo não cumprimento dos prazos é atribuída a fatores incontroláveis. Enquanto isso, os gerentes têm que recorrer ao uso de horas extras, ou outros meios disponíveis de aumentar a capacidade e desta forma atender

Tabela 1 – Produção real de uma linha composta por dois recursos com tempos de processamento variáveis (em número de peças).

Dia	Recurso A	Recurso B	Produção da linha
1	6	8	6
2	6	12	6
3	6	11	6
4	6	9	6
5	6	10	6
6	14	6	6
7	14	14	14
8	14	10	10
9	14	7	7
10	14	13	13
Total	100	100	80

o plano de produção. Ironicamente, os gerentes estarão pagando um prêmio extra pela capacidade que tanto trabalharam para colocar em ordem” (UMBLE & SRIKANTH, 1990, pp. 63).

UMBLE & SRIKANTH (1990) ressaltam também que as operações em qualquer ambiente de manufatura jamais estarão realmente balanceadas. Os dois fatores que mais contribuem para o inerente desbalanceamento das capacidades dos recursos são (1) a maneira como se dá o incremento de capacidade e (2) a presença de flutuações estatísticas e eventos dependentes que forcem os gerentes a desbalancearem suas fábricas. O primeiro aspecto refere-se ao fato que, muitas vezes, para se conseguir um perfeito balanceamento das capacidades produtivas, um particular recurso deverá trabalhar com 3,45 operadores enquanto que se deve ter 4,78 de uma outra determinada máquina. Para evitar o excesso de capacidade, muito provavelmente a empresa optará por 3 operadores e 4 máquinas daquele tipo e lançará mão de horas extras para aumentar a capacidade destes centros produtivos. O segundo fator diz respeito aos efeitos indesejáveis da presença simultânea de flutuações estatísticas e eventos dependentes sobre os estoques de produtos em processamento, *lead times*, despesas operacio-

nais e ganho. Desta forma, ainda segundo esses autores, sem uma apropriada quantidade de reserva de capacidade, o ganho, inventário, despesa operacional e o *lead time* são negativamente afetados, e a empresa será incapaz de operar eficientemente. Finalmente, os gerentes não terão alternativa a não ser dirigir fábricas desbalanceadas.

3. Uma Abordagem Quantitativa do Impacto da Variabilidade e do Grau de Utilização dos Recursos no Desempenho das Linhas de Produção

Muito do que existe em termos de teoria ou métodos de estudo de sistemas de produção se baseia em considerações que inexistem em ambientes produtivos reais. Uma dessas suposições mais utilizada, e muitas vezes ainda não corretamente entendida, é a *variabilidade*. A variabilidade está presente em todo sistema de manufatura e se manifesta de diversas maneiras. As principais fontes de variabilidade em ambientes de manufatura podem ser classificados em:

- *Variabilidade natural*: é a variabilidade inerente aos tempos de processamento naturais, excluindo quebras, *setups* ou outras

Tabela 2 – Relações matemáticas importantes para a análise em questão (baseado em HOPP & SPEARMAN, 1996).

Situação	Natural	Interruptiva	Não-interruptiva
Exemplo	Confiabilidade da máquina	Falhas randômicas	<i>Setups</i> e Retrabalho
Parâmetro	t_0, c_0^2	m_f, m_r	N_s, t_s
t_e	t_0	$T_0/A, A = m_f / (m_f + m_r)$	$t_0 + t_s / N_s$
s_e^2	$t_0^2 c_0^2$	$s^2_0 / A^2 + 2 t_0 m_r (1-A) / A$	$s^2_0 + s^2_0 / N_s + (N_s - 1) t_s^2 / N_s^2$
c_e^2	c_0^2	$C^2_0 + 2(1-A)A m_r / t_0$	s_e^2 / t_e^2

influências externas. Estão inclusas nesta categoria diversas fontes de variação, geralmente relacionadas ao operador, e que por isso são menos marcantes em processos automatizados. No entanto, mesmo nos processos mais rigidamente controlados, há sempre alguma fonte de variabilidade natural, como aquelas relacionadas a pequenas variações na composição dos materiais, por exemplo.

- *Variabilidade proveniente de quebras*: em muitos sistemas de produção as quebras não programadas são as maiores fontes de variabilidade nos tempos efetivos de processo. Duas medidas são freqüentemente utilizadas para avaliação do impacto que as quebras têm sobre a variabilidade de um recurso fabril: o tempo médio entre falhas (MTTF – *Mean Time To Failure*) e o tempo médio para reparos (MTTR – *Mean Time To Repair*). Pode-se provar que paradas freqüentes, porém curtas, é preferível em termos de variabilidade, às paradas infreqüentes, porém longas. Isso vem a justificar, por exemplo, um uso mais efetivo das manutenções preventivas.
- *Variabilidade proveniente de setups*: ao contrário do que ocorre com as quebras das máquinas, os *setups*, assim como as manutenções preventivas, são formas de interrupções do processo produtivo que ocorrem após terminada uma determinada tarefa. Por serem de natureza não-interruptiva, os *setups* devem

ter um tratamento diferenciado para cálculo de variabilidade. De qualquer forma, como é de se esperar, quanto menos peças são produzidas entre cada *setup* mais variável é o processo.

- *Variabilidade proveniente de retrabalhos*: problemas de qualidade, como o retrabalho, são outra fonte importante de variabilidade em sistemas de manufatura. Como o retrabalho possui também natureza não-interruptiva, deve ter tratamento semelhante ao dado à variabilidade de *setups*. Portanto, retrabalho tem efeitos análogos àqueles do *setup*, ambos roubam capacidade e contribuem grandemente para a variabilidade dos tempos efetivos de processo.

A Tabela 2 resume as fórmulas matemáticas para cálculo dos tempos médios de processo, variância e coeficiente de variação quadrático para cada uma das situações expostas anteriormente. Para tal, a seguinte notação, baseada em HOPP & SPEARMAN (1996) foi usada:

t_e = tempo médio efetivo de processo
 s_e^2 = variância do tempo médio efetivo de processo

$c_e^2 = s_e^2 / t_e^2$ = coeficiente de variação quadrático do tempo médio efetivo de processo

t_0 = tempo médio natural de processo
 s_0^2 = variância do tempo médio natural de processo

$c_0^2 = s_0^2 / t_0^2$ = coeficiente de variação quadrático do tempo médio natural de processo

m_f = tempo médio entre falhas

m_r = tempo médio para reparo
 A = disponibilidade da máquina
 N_s = número de peças produzidas entre cada *setup*
 t_s = tempo médio de duração de um *setup*

As discussões anteriores focalizaram somente a variabilidade dos tempos de processo em estações produtivas individuais. No entanto, a variabilidade de uma estação pode afetar o comportamento de outras estações em uma linha por outro tipo de variabilidade denominada de variabilidade de fluxo, no qual fluxo refere-se à transferência de peças de uma estação para outra. Assim, se uma estação precedente possui um tempo de processo altamente variável, a estação alimentada por ela também será altamente variável. Desta forma, para se analisar o efeito da variabilidade sobre uma linha deve-se caracterizar a variabilidade em fluxos.

Para se iniciar o estudo de fluxos deve-se definir o parâmetro de chegada de peças em um recurso. As saídas destas estações serão as chegadas em outros recursos. Logo, uma vez que se descreve a variabilidade das chegadas em um recurso e determina-se como este afeta a variabilidade das saídas destes mesmos recursos (e portanto as chegadas em outros recursos) pode-se caracterizar a variabilidade da linha inteira. As discussões que se seguem foram desenvolvidas a partir de HOPP & SPEARMAN (1996). Sejam:

t_a = tempo médio entre chegadas
 $r_a = 1 / t_a$, taxa média de chegadas
 s_a^2 = variância do tempo entre chegadas
 $c_a^2 = s_a^2 / t_a^2$, = coeficiente de variação quadrático

Da mesma forma definem-se os parâmetros t_d , r_d , s_d^2 , c_d^2 para a saída de peças de um recurso. Pode-se concluir então que:

$$t_a(i+1) = t_d(i) \text{ e}$$

e portanto,

$c_a^2(i+1) = c_d^2(i)$, onde i é a posição de um recurso na linha, ou seja, a saída de um recurso é igual à chegada do recurso que o sucede, conforme ilustra a Figura 1.

Para se caracterizar a variabilidade na saídas de peças de um recurso deve-se reconhecer que essa é resultante de dois fatores:

- (1) variabilidade nas chegadas ao recurso;
- (2) variabilidade nos tempos de processo.

A relativa contribuição destes dois fatores depende da utilização dos recursos (fração de tempo ocupado produzindo num longo prazo). Formalmente, pode-se definir a utilização como sendo:

$u = r_a t_e / m$, onde m é o número de máquinas ou recursos.

Quando o valor de u está muito perto de um, o recurso virtualmente nunca fica esperando por peças para produzir. Portanto, nessa situação os tempos entre saídas do recurso serão essencialmente idênticos aos tempos de processamento, ou

$$c_d^2 = c_e^2$$

Da mesma forma, quando u tende a zero o tempo de processo corresponderá a uma pequena parcela do tempo entre saídas e portanto,

$$c_d^2 = c_a^2$$

Para todos os valores de u pode-se interpolar os resultados anteriores na seguinte fórmula:

$$c_d^2 = u^2 c_e^2 + (1 - u^2) c_a^2,$$

a qual rende para situações com m máquinas ou recursos,

$$c_d^2 = 1 + (1 - u^2)(c_a^2 - 1) + u^2(c_e^2 - 1)/\sqrt{m}.$$

Os resultados acima para variabilidade de tempos de processo e fluxo são essenciais para caracterizar os efeitos da variabilidade em uma linha de produção. Com a finalidade, no entanto, de se avaliar o impacto destes tipos de variabilidade nos três parâmetros básicos de desempenho de uma linha, *WIP* (*Work in Process* – Estoque em Processo), *CT* (*Cycle Time* – Tempo de Ciclo) e *TH* (*Throughput* – Taxa de Produção), deve-se fazer uso da Teoria das Filas.

Primeiramente, lembra-se que os tempos de processamento (incluindo quebras, *setups*, etc.) na maioria das indústrias representam uma pequena

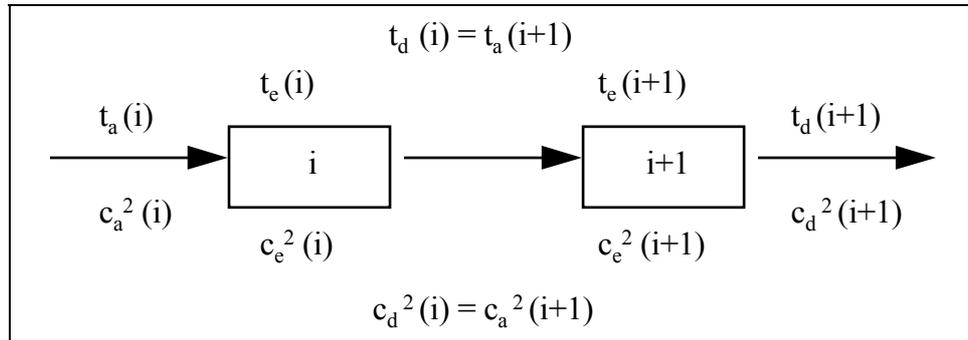


Figura 1 – Ilustração dos parâmetros utilizados nas equações matemáticas desenvolvidas (baseada em HOPP & SPEARMAN, 1996).

parcela do tempo de ciclo total de produção. A maioria do tempo é gasto esperando (na forma de filas) pela disponibilidade de vários recursos. Portanto, torna-se fundamental conhecer as causas desses tempos de espera na forma de filas.

Lembrando que o tempo médio total gasto na estação, CT, e o tempo médio gasto em fila, CT_q , são aditivos, portanto:

$$CT = Ct_q + t_e \quad (I)$$

Aplicando agora a lei de Little ($TH=WIP/CT$) e sabendo-se que o *throughput* é a própria taxa de chegada, r_a , conclui-se:

$$WIP = r_a \times CT \quad (II)$$

Aplicando agora a lei de Little somente para a fila obtém-se:

$$WIP_q = r_a \times Ct_q \quad (III)$$

Um dos mais simples sistemas de filas para se analisar é a M/M/1. Ainda que bastante simples e, portanto, não permitindo uma representação acurada da maioria dos recursos de manufatura, ela possibilita valiosos *insights* para boa parte dos sistemas reais. Esta é caracterizada por possuir tempos entre chegadas com distribuição exponencial ou de Markov, um único recurso com tempo de processamento também exponencial, uma disciplina de atendimento tipo FCFS (*First Come First Service* – primeiro a chegar, primeiro a ser

servido) e espaço ilimitado para as tarefas esperando na fila.

Pode-se demonstrar que o nível de estoque em uma fila M/M/1 é:

$$WIP(M/M/1) = u / 1-u \quad (IV)$$

e que aplicando-se a lei de Little e lembrando que $u = r_a / r_e$ tem-se:

$$CT(M/M/1) = WIP(M/M/1) / r_a = t_e / 1-u \quad (V)$$

Da equação I pode-se concluir que o tempo médio em fila é:

$$Ct_q(M/M/1) = CT(M/M/1) - t_e = t_e u / 1-u \quad (VI)$$

E da equação III:

$$WIP_q(M/M/1) = r_a \times Ct_q(M/M/1) = u^2 / 1-u \quad (VII)$$

As equações IV, V, VI e VII possibilitam concluir alguns pontos interessantes. Observa-se que o tempo médio de fila, e portanto também o nível médio de estoque em processo, aumentam de maneira não linear, segundo o nível de utilização. É claro que CT e WIP também aumentam quando t_e aumenta (ou seja, quanto mais lenta a máquina maiores serão os valores de CT e WIP), porém os níveis de estoque e o tempo de ciclo aumentam de maneira muito mais rápida quando o nível de utilização da máquina

tende a 1 ou 100%. O que se pode concluir até aqui é que, em sistemas M/M/1 o principal fator causador do tempo de fila é grau de utilização, u , da máquina. Em outras palavras, isso significa que conseguindo um índice de utilização de 100% da máquina, o nível de estoque e o tempo de fila serão infinitos.

Como já relatado, sistemas M/M/1 possuem pouca representatividade em termos reais. No entanto, pode-se estimar valores de tempo de ciclo para sistemas mais gerais (com distribuições probabilísticas completamente gerais) os quais, ao contrário dos sistemas M/M/1, são razoavelmente acurados em típicos sistemas de manufatura.

Desta forma, pode-se obter, por aproximação, para sistemas G/G/1 o seguinte valor para o tempo médio de fila:

$$Ct_q(G/G/1) = ((c_a^2 + c_e^2) / 2) \times CT_q(M/M/1)$$

ou

$$Ct_q(G/G/1) = ((c_a^2 + c_e^2) / 2) t_e u / (1-u) \quad \text{(VIII)}$$

Esta aproximação permite concluir alguns pontos interessantes. Primeiro, ela é exata para a fila M/M/1 (basta substituir os valores de c_a e c_e por 1, uma vez que os coeficientes de variabilidade valem 1 em distribuições exponenciais). Depois, ela separa o tempo médio de fila em três termos, ou seja, um termo adimensional da *variabilidade*, $(c_a^2 + c_e^2) / 2$, um termo de *utilização*, $u / (1-u)$, e um termo de *capacidade*, t_e . Nota-se então que o tempo de processamento (t_e) é responsável, apenas, por uma pequena parcela do tempo médio de fila.

Pode-se inferir ainda o tempo médio de fila para sistemas gerais constituídos por m máquinas (G/G/m) a seguinte equação:

$$Ct_q(G/G/m) = ((c_a^2 + c_e^2) / 2) (u^{(2(m+1) - 1)} / m(1-u)) t_e \quad \text{(IX)}$$

Novamente, observa-se da equação anterior que o tempo médio de fila aumenta de maneira não linear com o nível de utilização, mesmo para sistemas mais gerais e com m máquinas.

4. Análise dos Resultados

Nesta etapa serão analisadas três questões importantes referentes ao problema de alocação de cargas entre recursos, que são os impactos resultantes, o balanceamento *versus* o desbalanceamento e a alocação de capacidades produtivas de maneira desigual entre recursos.

4.1 Impactos Resultantes

A análise apresentada anteriormente permite concluir alguns pontos interessantes. O primeiro deles é que, da equação IX, o tempo de ciclo torna-se infinito quando o índice de utilização dos recursos é um. A consequência disso é que, para ser estável, um recurso deve ter uma taxa de processamento sempre maior que a taxa de chegada. Uma maneira de interpretar essa conclusão é que uma vez que todo sistema possui uma certa randomicidade, pode-se sempre esperar uma possível seqüência de eventos que provoque uma falta de material no gargalo, ao menos por um curto intervalo de tempo. A única maneira de impedir que o gargalo pare por falta de material para processar é começar com uma quantidade infinita de material na frente dele. Portanto, para uma taxa de chegada (r_a) igual à taxa de processamento (r_e), deve haver uma quantidade de WIP em fila, ou pela lei de Little, um tempo de ciclo também infinito.

HOPP & SPEARMAN (1996) explicam como isso pode acontecer em situações reais, nas quais geralmente os gerentes fábri liberam material para fábrica segundo a máxima capacidade disponível e nem por isso possuem ciclos de tempo infinitos. Esses autores chamam isso de “ciclo vicioso da hora extra”, o qual se comporta da seguinte maneira:

- (1) a capacidade da planta é computada considerando-se diversos fatores de indisponibilidade do recurso (quebras, *setups*, retrabalho, paradas para almoço, etc.);
- (2) o programa mestre de produção é estabelecido segundo a capacidade efetiva da fábrica. As taxas de liberação são agora essencialmente a mesma da capacidade;

- (3) mais cedo ou mais tarde, devido às randomicidades na chegada das ordens, tempos de processos, ou ambos, o processo gargalo fica sem material para trabalhar;
- (4) há agora um déficit de trabalho que nunca mais poderá ser recuperado e o nível de WIP aumenta;
- (5) desde que o sistema esteja em sua capacidade máxima, a taxa de produção (*throughput*) permanecerá relativamente constante. Assim, pela lei Little, conclui-se que o sistema terá um aumento proporcional no tempo de ciclo;
- (6) as ordens de produção começam a ficar atrasadas;
- (7) os clientes começam a reclamar;
- (8) depois que o WIP e os tempos de ciclo estiverem relativamente altos, e os clientes tiverem reclamado, os gerentes decidem tomar uma atitude;
- (9) hora extra (ou subcontratação ou rejeição de novas ordens) é autorizada;
- (10) a capacidade efetiva é agora significativamente maior que a taxa de liberação;
- (11) o nível de WIP decresce, os tempos de ciclo diminuem e o serviço ao cliente melhora. Os gerentes prometem nunca mais fazer isso de novo;
- (12) volta ao passo 1!

A discussão anterior faz concluir que, em um regime estável, toda planta liberará trabalho numa taxa média que é estritamente menor que a sua capacidade média. Uma outra maneira de interpretar esta mesma conclusão é que não se pode exigir a capacidade máxima efetiva de todos os recursos de uma planta fabril, ainda que seja justamente isso que os medidores de desempenho tradicionalmente exigem ao se buscar maximização de índices como utilização e eficiência. Deve-se lembrar ainda que qualquer erro ao se estimar a capacidade para valores acima de sua capacidade (que é o que normalmente ocorre até como forma de forçar os recursos a maximizarem seus índices de desempenho) terá impactos ainda maiores em termos de aumento dos ciclos de tempo e do estoque em processo.

Uma outra contribuição interessante que as equações anteriormente desenvolvidas podem dar refere-se à propagação da variabilidade. Ao observar-se a equação (VIII)

$$Ct_q(G/G/1) = ((c_a^2 + c_e^2) / 2) t_e u / 1-u$$

pode-se notar que os tempos de fila aumentam proporcionalmente com o aumento tanto da variabilidade da chegada, c_a^2 , quanto da variabilidade do processo, c_e^2 . Ao dividir o termo variabilidade em duas componentes, a equação anterior mostra que a posição da variabilidade traz conseqüências ao desempenho da linha produtiva. Especificamente, ela implica que processos altamente variáveis no início de uma linha impactam o tempo de espera de todos os processos posteriores a eles, enquanto que processos altamente variáveis no fim de uma linha causarão filas somente lá. Em outras palavras, a variabilidade no início de roteiros de fabricação provocam um impacto maior nos níveis de WIP e tempos de ciclo quando comparados a variabilidades semelhantes presentes em processos situados mais no final do mesmo roteiro. A conclusão é que maiores esforços deveriam ser focalizados para recursos localizados no início da linha onde o impacto da variabilidade é maior. Novamente, pode-se notar deficiências nos medidores de desempenho comumente adotados, os quais não refletem as diferenças posicionais dos recursos.

4.2 Balanceamento *versus* Desbalanceamento

As equações desenvolvidas para cálculo dos tempos de ciclo de cada recurso, o qual é predominado pelo tempo de fila das tarefas, e a lei de Little, permitem concluir alguns pontos interessantes quanto ao uso de linhas balanceadas.

Em primeiro lugar, ao se projetar uma linha balanceada espera-se extrair da mesma um *throughput* equivalente a capacidade efetiva média de seus recursos (a qual se pressupõe, ou se busca, ser bastante próxima entre cada recurso), pela maximização do uso destes

recursos. Aliás, é justamente isso que os indicadores de desempenho (como eficiência, utilização, produtividade, etc.) geralmente utilizados remetem a fazer. No entanto, como já discutido anteriormente, a alta taxa de utilização dos recursos implica em altos tempos de fila e portanto, em altos tempos de ciclo. Pela lei de Little, isto é o mesmo que afirmar que a linha está operando com altos níveis de estoque.

De fato, como também já comentado, devido às inerentes variabilidades dos ambientes produtivos, é simplesmente impossível extrair-se 100% de capacidade de uma linha de produção. Tal possibilidade só seria factível se pudesse manter uma quantidade infinita de material antes do recurso mais lento, ou gargalo, para compensar toda e qualquer vulnerabilidade implicitamente presente nos processos fabris (GOLDRATT, 1991). Entretanto, se fosse possível finalmente encontrar uma configuração na qual todos os recursos tivessem a mesma capacidade produtiva, a única saída para explorar a capacidade de toda a linha ao máximo seria manter infinitas quantidades de estoque antes de cada elemento produtivo.

Supondo-se que a empresa esteja conformada com o argumento que diz ser impossível extrair 100% da capacidade da linha e estipule como meta obter um *throughput* equivalente a 95% da capacidade individual de cada recurso. Ela poderá abaixar bastante a quantidade de estoque em processo mas, como ainda terá de proteger a capacidade de cada um dos recursos individuais (já que a linha é supostamente balanceada, se ao menos um destes recursos perder mais que 5% de sua capacidade efetiva por falta de material a linha inteira será incapaz de gerar o *throughput* esperado), deverá garantir um bom nível de estoque antes de cada recurso.

Uma outra forma de se compreender este mesmo problema é que um nível de utilização de 95% de cada recurso deve, segundo as fórmulas apresentadas na Tabela 2, gerar um valor ainda relativamente alto do tempo de ciclo para cada recurso da linha e por isso, também, um alto tempo de ciclo para linha como um todo (que

nada mais é que a soma dos tempos de ciclo dos recursos individuais).

Por outro lado, supondo-se que seja adicionada capacidade em alguns recursos da linha de maneira a desbalanceá-la, alguém poderia justificar que tal alternativa necessitaria de investimentos e que ainda aumentaria (via depreciação) as despesas da empresa. No entanto, deve-se lembrar que os recursos não são igualmente caros e que provavelmente alguns são bem menos custosos do que outros, o que permitiria, com pouco investimento, alavancar a capacidade de alguns recursos. Uma vez que esses recursos passarão a ter um excesso de capacidade em relação ao(s) recurso(s) gargalo(s), seus níveis de utilização serão menores, o que refletirá numa redução dos seus respectivos tempos de fila e logo, do ciclo de tempo total (se ainda levar-se em consideração a parcela, mesmo que pequena, do tempo de ciclo referente ao próprio tempo de processamento, esta redução será ainda maior). A conclusão é que, para atingir-se o mesmo valor do *throughput* desejado no caso da linha balanceada, necessita-se agora de uma quantidade menor de estoque em processo (WIP).

A justificativa pela adoção de uma linha balanceada não deve ser baseada apenas em critérios financeiros, ao se comparar, por exemplo, custos de investimento em capacidade adicional *versus* economia no investimento de capital na forma de inventário de material em processo. Tal visão deixa de considerar diversos outros aspectos de natureza menos tangível, mas nem por isso menos importantes, tais como:

- (1) menores tempos de ciclo permitem que as empresas que operam por encomenda ofereçam menores *lead times* de seus produtos para os clientes;
- (2) menores *lead times* ofertados possibilitam que a empresa opere com previsões de demanda em horizonte mais curtos e, portanto, com menos risco de erros de previsões;
- (3) níveis reduzidos de estoque em processo permitem que a empresa possa identificar e resolver problemas de qualidade com muito mais eficiência;

(4) empresas que operam com menores níveis de estoque em processo são mais flexíveis a mudanças de volume, *mix* de produção e alterações de engenharia.

Fora esses motivos, há ainda um outro, muitas vezes negligenciado, quando se discute questões relativas à distribuição de cargas entre recursos produtivos. Do ponto de vista logístico, é muito mais fácil programar e controlar sistemas produtivos que possuem um ou muitos poucos gargalos que sistemas de produção com alto grau de balanceamento. Se fosse possível, ainda que esses poucos gargalos sejam relativamente estáveis, as atividades de gerenciamento ficariam ainda mais facilitadas. Nesses casos, bastaria apenas gerar um programa que maximizasse a utilização do(s) recurso(s) gargalo(s) e sincronizasse a produção dos demais recursos em função deste programa.

Deve-se ressaltar, entretanto, que toda discussão anterior é válida somente para linhas de produção em que os recursos não sejam conectados um com outro de alguma maneira física (como numa esteira, por exemplo), ou seja, existe mais flexibilidade para acúmulo de material entre os recursos e, portanto, mais independência entre eles. Para os casos em que os recursos estejam conectados fisicamente um ao outro, a parada de um recurso implica na parada da linha inteira. Nesta situação a velocidade da linha deve ser estritamente menor que a capacidade das estações de trabalho para se evitar ao máximo possíveis paradas por falta de material.

As discussões anteriores se apóiam em justificativas para se aumentar a capacidade de alguns recursos mantendo a capacidade de alguns outros constantes. Isto aumentaria o valor médio da capacidade da linha como um todo. Existe, no entanto, uma outra linha de pesquisa que busca encontrar formas distintas de se alocar cargas de trabalho ou capacidades produtivas entre recursos que otimizem alguma variável em estudo (*throughput* ou estoque em processo, por exemplo) mantendo-se o valor médio da capacidade da linha inteira. Em outras, para um mesmo

nível de investimento, qual a melhor maneira de se distribuir capacidades produtivas?

4.3 Alocação de Capacidade Produtiva de Maneira Desigual entre Recursos

FRY & RUSSELL (1993), em um estudo comparativo sobre diferentes formas de alocação de cargas entre os recursos, citam seis tipos de critérios de distribuição de cargas de trabalho entre recursos de uma fábrica fictícia (*job-shop* híbrido) que compartilha características tanto de um sistema *flow-shop* quanto de um sistema *job-shop*. Os seis critérios testados foram: balanceado, em forma de tigela (*bowl*), tigela invertida ou “V”, linear, em passo (*step*) e segmentado, conforme ilustra a Figura 2. Os autores avaliaram cada um dos critérios citados segundo três formatos de alocação de carga (mais acentuado, moderado e pouco acentuado), três níveis de excesso de capacidade (10, 15 e 20%) e dois níveis de variabilidade (alto e baixo).

Segundo esses autores, a primeira estratégia de alocação de cargas é igualar ou balancear as capacidades em todas as estações de trabalho, a qual é defendida pela maioria da literatura de balanceamento de linha. Sob esta estratégia cada máquina é utilizada por uma igual porcentagem de tempo.

A segunda estratégia é conhecida como *bowl* devido à forma como as capacidades são distribuídas entre os recursos. O primeiro trabalho de pesquisa publicado que fez referência ao fenômeno *bowl* foi o trabalho de HILLIER & BOLING (1966). Na pesquisa, os autores verificaram que o número de estações de trabalho e a capacidade de estocagem entre as estações são fatores de projeto críticos que influenciam fortemente a eficiência do sistema. Segundo HILLIER & BOLING (1966), em sistemas constituídos por duas estações de trabalho a taxa de produção é maximizada quando se consegue balancear perfeitamente a linha de produção. Entretanto, a alocação desigual de cargas em uma forma apropriada rende uma taxa de produção mais alta em um

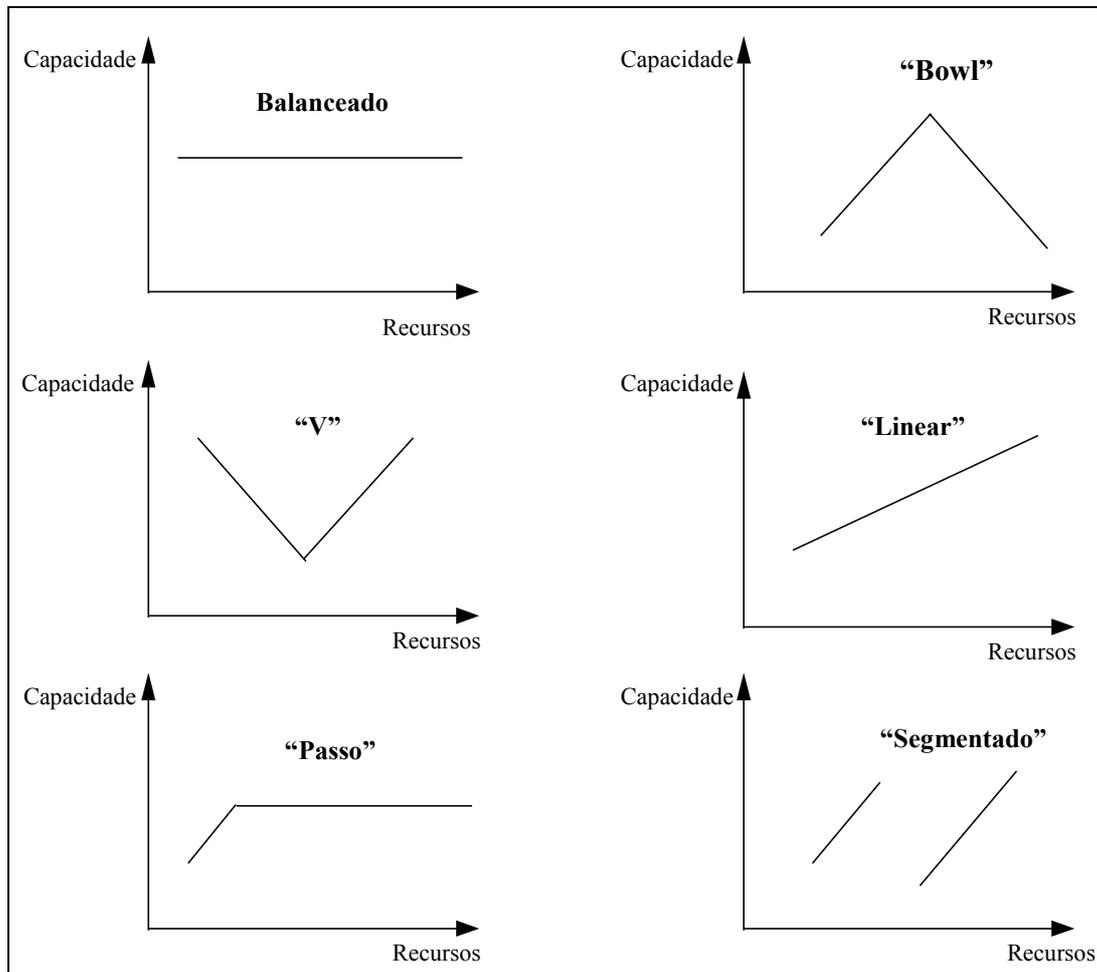


Figura 2 – Critérios de alocação de capacidades produtivas (FRY & RUSSELL, 1993).

sistema com três estações, com ou sem estoques, do que em uma linha de produção perfeitamente balanceada. Foi verificado que a taxa de produção para sistemas com três estações é maximizada alocando-se menos tempo de processamento médio na estação intermediária em comparação com as estações localizadas no extremo da linha. A isto os autores chamam de fenômeno *bowl*, no qual a taxa de produção é maximizada designando-se menores tempos de processamento médio às estações intermediárias em relação às estações localizadas nas duas pontas da linha. Tal constatação foi verificada para sistemas de linhas de produção com três ou mais estações (ver também HILLIER &

BOLIMG, 1979; MISHRA, A. *et al.*, 1985; THOMPSON & BURFORD, 1988; SO, 1989). Nota-se que a Figura 2 ilustra gráficos que relacionam capacidades (e não tempos de processamento) produtivas com recursos de manufatura.

Uma terceira estratégia, citada por FRY & RUSSELL (1993), é chamada de *bowl invertido* (ou em “V”), na qual os recursos situados no centro da linha recebem a menor quantidade de capacidade, enquanto as capacidades dos recursos posteriores e anteriores a ele têm suas capacidades aumentadas linearmente.

A quarta estratégia para alocação de capacidade é denominada de *linear* e representa o caso

em que o recurso com menor capacidade é o primeiro da linha. Os recursos subseqüentes a ele passam a ter um incremento linear de suas respectivas capacidades, de maneira que o último departamento da linha receberá a maior capacidade de produção. FRY & RUSSELL (1993) citam os trabalhos de DAVIS (1966) e EL-RAYAH (1979) como os pioneiros nas pesquisas deste tipo de critério de alocação de cargas de trabalho. Outros defensores desta mesma estratégia são GOLDRATT & FOX (1989), que argumentam que a máquina mais lenta deveria ser localizada na frente do processo.

A quinta estratégia de alocação de capacidade incluída em FRY & RUSSELL (1993) é a *segmentada*, sugerida por GOLDRATT (1981). A crença aqui é que as flutuações (variabilidades) nos processos tendem a se acumular na direção do fluxo de materiais numa linha de produção. Deste modo, um recurso é afetado pelas flutuações do recurso precedente e afeta o recurso subseqüente. A sugestão de GOLDRATT (1981) é que o processo deveria ser segmentado em dois ou mais mini-processos por pulmões de capacidade cuidadosamente localizados onde cada segmento se comporta segundo um projeto de alocação de capacidades linear. Por exemplo, numa linha de sete recursos, poder-se-ia ter o quarto recurso com a menor capacidade dentre eles. Os recursos cinco, seis e sete teriam capacidades linearmente crescentes em relação ao recurso quatro e respectivamente idênticas aos do recurso um, dois e três. Desta maneira, os recursos que recebessem uma quantidade acumulada maior de flutuação ou variabilidade teriam mais capacidade para compensá-la.

A sexta e última estratégia é a alocação em *passo* ou *step*. Neste caso a primeira estação tem a menor capacidade e as demais estações que a sucedem têm capacidades levemente maiores que a primeira, porém iguais entre si. Segundo FRY & RUSSELL (1993), WIGHT (1970) sugeriu este tipo de critério como forma de facilitar o controle de entrada e saída de tarefas pela linha, localizando a operação gargalo no início da mesma. Essa estratégia seria então uma

combinação entre balanceamento e aumento linear de capacidades de produção, com o gargalo localizado no início do processo de fabricação.

Os resultados obtidos por FRY & RUSSELL (1993) mostram que a apropriada alocação de capacidades é dependente sobre o grau de flutuação dos processos e do tamanho da folga na capacidade. Em fábricas com baixo nível de excesso de capacidade, desbalancear intencionalmente a capacidade resultaria em desempenhos nunca piores que a alocação balanceada de cargas e em situações de alta variabilidade o desbalanceamento é sempre superior ao balanceamento. Por outro lado, quando a folga na capacidade é alta, 20% no caso, a estratégia de balanceamento de capacidades deveria ser escolhida. Nas situações em que linhas desbalanceadas superaram as linhas balanceadas, os autores verificaram melhores resultados para as estratégias “linear” e em “V”.

Nota-se aqui que a pesquisa de FRY & RUSSELL (1993) procurou estudar a melhor forma de alocar capacidades produtivas entre recursos de uma fábrica e não, a partir de uma configuração já existente, acrescentar capacidade em alguns recursos como forma de se alcançar o mesmo *throughput* com um menor tempo de ciclo e menor estoque em processo. O objetivo do trabalho desses autores foi enfatizar que mais importante que o uso eficiente dos recursos é seu uso efetivo. Aceitar que sob certas condições é mais aconselhável desbalancear as cargas entre os recursos implica em aceitar também que certos recursos serão menos utilizados que outros e que os medidores de desempenho a serem adotados nestes casos deverão refletir esta nova condição. Antes de se propor uma discussão mais aprofundada do trabalho de FRY & RUSSELL (1993), a idéia aqui é mostrar que a alta eficiência de uma linha de manufatura pode estar, e muitas vezes está de fato, relacionada com baixas eficiências de seus recursos individuais, e que tais níveis de “ineficiências” podem ser apropriadamente alocados. Isto é, deve-se não só aceitar certos níveis de ociosida-

de dos recursos fabris como também questionar se o balanceamento de capacidades produtivas é a melhor maneira de se alocar cargas de trabalho aos recursos.

5. Conclusões

As discussões realizadas neste artigo procuram mostrar que existe, de fato, uma defasagem entre as reais necessidades competitivas das empresas de manufatura e suas práticas usuais, geralmente substanciadas por seus indicadores de desempenho adotados. Para a grande parte das empresas de manufatura, integrar a gestão produtiva às reais necessidades do mercado, implica em entender (e aceitar) que existe uma inter-relação muito estreita entre seus recursos produtivos e desses com as tarefas a serem realizadas. Adicionalmente, deve-se considerar que essa interdependência vem na maioria das vezes recheada de um elevado grau de variabilidade dos processos e de dependência de eventos externos imprevisíveis (cancelamen-

tos de pedidos, por exemplo). Tratar então cada recurso como se esse operasse de maneira isolada de todos os fatores que o controlam, principalmente de forma não conectada com as necessidades do mercado é, no mínimo, um erro gerencial grave. No entanto, é justamente isto que acontece quando exige-se de cada recurso de uma empresa de manufatura 100% de eficiência. Fica claro então que para as empresas industriais obterem uma alta flexibilidade, um bom desempenho no atendimento de prazos e baixos *lead times*, é necessário, muitas vezes, aceitar um decréscimo nas cargas de trabalho da fábrica e menores índices de eficiência de seus recursos.

Finalmente, cabe destacar que novas pesquisas vêm sendo desenvolvidas pelos autores com vistas a validação das conclusões aqui obtidas. Tais pesquisas estão engajadas a estudar, principalmente via simulação, os efeitos resultantes das variabilidades dos processos produtivos e das diversas formas de se alocar capacidades produtivas aos recursos fabris.

Referências Bibliográficas

- ATWATER, J.B. & CHAKRAVORTY, S.S.:** "Does protective capacity assist managers in competing along time-based dimensions?" *Production & Inventory Management Journal*, third quarter, p.53-59, 1994.
- CORRÊA, H.L. & GIANESI, I.G.N.:** *Just In time, MRP II e OPT: um enfoque estratégico*. São Paulo, Atlas, 1993.
- DAVIS, L.E.:** "Pacing effects on manned assembly lines". *International Journal of Production Research*, v.4, n.3, p.171-184, 1966.
- EL-RAYAH, T.E.:** "The efficiency of balanced and unbalanced production lines". *International Journal of Production Research*, v.17, n.1, p.61-75, 1979.
- FRY, T.D. & RUSSELL, G.R.:** "Capacity allocation strategies in a hypothetical job-shop". *International Journal of Production Research*, v.31, n.5, p.1097-1115, 1993.
- GOLDRATT, E.M.:** "The unbalanced plant". *APICS 24th Annual International Conference Proceedings*, 1981.
- GOLDRATT, E.M.:** *A síndrome do palheiro: garimpando informação num oceano de dados*. São Paulo, IMAM, 1991.
- GOLDRATT, E.M. & COX, J.:** *A meta: um processo de aprimoramento contínuo*. São Paulo, Educator, 1986.
- GOLDRATT, E.M. & FOX, R.E.:** *A corrida*. São Paulo, IMAM, 1989.
- HILLIER, F.S. & BOLING, R.W.:** "The effect of some design factors on the efficiency of production lines with variable operations times". *The Journal Of Industrial Engineering*, v.17, n.12, p.651-658, 1966.
- HILLIER, F.S. & BOLING, R.W.:** "On the optimal allocation of work in symmetrically unbalanced production line systems with variable operation times". *Management Science*, v.25, n.8, p.721-728, 1979.
- HOPP, W.J. & SPEARMAN, M.L.:** *Factory Physics: Foundations of Manufacturing Management*. Boston, Massachusetts. Irwin / McGraw-Hill, 1996.

- MISHRA, A.; ACHARYA, D.; RAO, N.P. & SASTRY, G.P.:** “Composite stage effects in unbalancing of series production systems”. *International Journal of Production Research*, v.23, n.1, p.1-20, 1985.
- PIRES, S.:** *Gestão Estratégica da Produção*. Piracicaba, Editora Unimep, 1995.
- SLACK, N.:** *The manufacturing advantage*. Londres, Mercury Books, 1991.
- SO, K.C.:** “On the efficiency of unbalancing production lines”. *International Journal of Production Research*, v.27, n.4., p.717-729, 1989.
- SOUZA, F.B.:** “Fundamentos de Teoria das Restrições e uma aplicação em uma Metodologia de Integração de Empresa”. São Carlos. Dissertação (Mestrado) – Escola de Engenharia de São Carlos, Universidade de São Paulo, 1997.
- STEELE, D.C. & PARKE-SHIELDS, K.:** “Capacity slack: strategic alternative to lead time”. *Production and Inventory Management Journal*, fourth quarter, p.1-5, 1993.
- THOMPSON, JR., W.W. & BURFORD, R.L.:** “Some observations on the bowl phenomenon”. *International Journal of Production Research*, v.26, n.8, p.1367-73, 1988.
- UMBLE, M.M. & SRIKANTH, M.L.:** *Synchronous Manufacturing: principles for world class excellence*. Cincinnati, South-Western, 1990.
- WIGHT, O.W.:** “Input/output control: A real handle on lead time”. *Production and Inventory Management*, v.11, n.3, p.9-31, 1970.

ANALYSIS AND PROPOSITIONS ABOUT THE BALANCING AND USE OF OVER CAPACITY IN PRODUCTIVE RESOURCES

Abstract

The search for high efficiency of productive resources has been a basic concern of most manufacturing managers for a long time, in their constant attempts to obtain lower production costs. This managerial practice emerged with the mass production principles, and during recent decades has been shown to be inadequate and has converted itself into a great obstacle for manufacturing management based on the real competitive dimensions required by the current marketplace. This paper deals with this subject from the perspective of the balancing and use of over capacity, and presents a quantitative approach to facilitate the analysis and propositions conducted. The final remarks highlight the importance of dealing with productive resources capacity in an integrated manner, instead of the traditional practice focused on the optimization of isolated resources.

Key words: capacity management, capacity balancing, manufacturing strategy.