



Felipe Sawaya Hirschheimer

**Um estudo de caso sobre a implementação de um *APS* de
otimização no gerenciamento e organização da produção**

31/2013

Campinas, 2013



UNIVERSIDADE ESTADUAL DE CAMPINAS
FACULDADE DE ENGENHARIA MECÂNICA

Felipe Sawaya Hirschheimer

**Um estudo de caso sobre a implementação de um *APS*
de otimização no gerenciamento e organização da
produção**

Orientador: Prof. Dr. Antônio Batocchio

Dissertação apresentada para a Faculdade de Engenharia Mecânica para obtenção de título de Mestre em Engenharia Mecânica na área de Materiais e Processos de Fabricação.

ESTE EXEMPLAR CORRESPONDE A VERSÃO FINAL
DA DISSERTAÇÃO DEFENDIDA PELO ALUNO:
FELIPE SAWAYA HIRSCHHEIMER E ORIENTADA
PELO PROF. DR. ANTÔNIO BATOCCHIO

A handwritten signature in blue ink, consisting of several loops and a long horizontal stroke, is positioned above a horizontal line.

Prof. Dr. Antônio Batocchio

H615e Hirschheimer, Felipe Sawaya, 1980-
Um estudo de caso sobre a implementação de um APS de otimização no gerenciamento e organização da produção / Felipe Sawaya Hirschheimer. – Campinas, SP : [s.n.], 2013.

Orientador: Antônio Batocchio.
Dissertação (mestrado) – Universidade Estadual de Campinas, Faculdade de Engenharia Mecânica.

1. Engenharia de Produção. 2. Planejamento da Produção. 3. Pesquisa Operacional. I. Batocchio, Antônio, 1953-. II. Universidade Estadual de Campinas. Faculdade de Engenharia Mecânica. III. Título.

Informações para Biblioteca Digital

Título em inglês: A study case on the implementation of an APS to optimize the management and organization of production

Palavras-chave em inglês:

Production Engineering

Production Planning

Operational Research

Área de concentração: Materiais e Processos de Fabricação

Titulação: Mestre em Engenharia Mecânica

Banca examinadora:

Antônio Batocchio [Orientador]

Oswaldo Luiz Agostinho

Marcus Fabius Henriques de Carvalho

Data de defesa: 20-02-2013

Programa de Pós-Graduação: Engenharia Mecânica



UNIVERSIDADE ESTADUAL DE CAMPINAS
FACULDADE DE ENGENHARIA MECÂNICA
COMISSÃO DE PÓS GRADUAÇÃO EM ENGENHARIA MECÂNICA
DEPARTAMENTO DE ENGENHARIA DE FABRICAÇÃO

Dissertação de Mestrado Acadêmico

**Um estudo de caso sobre a implementação de um APS de
otimização no gerenciamento e organização da produção**

Autor: Felipe Sawaya Hirschheimer

Orientador: Antônio Batocchio

A Banca Examinadora composta pelos membros abaixo aprovou esta Dissertação

A handwritten signature in black ink, appearing to read "Antônio Batocchio", written over a horizontal line.

Prof. Dr. Antônio Batocchio, Presidente
Faculdade de Engenharia Mecânica - UNICAMP

A handwritten signature in black ink, appearing to read "Oswaldo Luiz Agostinho", written over a horizontal line.

Prof. Dr. Oswaldo Luiz Agostinho
Faculdade de Engenharia Mecânica - UNICAMP

A handwritten signature in black ink, appearing to read "Marcius Fabius", written over a horizontal line.

Prof. Dr. Marcius Fabius Henriques de Carvalho
Cenpra/PUC Campinas

Campinas, 20 de Fevereiro de 2013.

Dedicatória

*Aos meus pais Sonia e Mário
e esposa Alena*

Agradecimentos

Em primeiro lugar, gostaria de agradecer ao professor e orientador Antônio Batocchio pela grande oportunidade, paciência e apoio na condução deste trabalho.

Ao Sr. Paulo Reimann, então presidente da Gelita pela oportunidade e coragem em realizar o projeto de implementação do APS de otimização.

À Alcione, então coordenadora de planejamento da Gelita pela confiança, cumplicidade e enorme contribuição para meu conhecimento sobre planejamento e produção de gelatinas.

Ao professor Dr. Marcius Carvalho pelas valiosas contribuições durante a implementação do projeto, pelas referências, conselhos para a construção desta dissertação e gentil aceitação do convite para formação da banca.

Ao professor Dr. Oswaldo Agostinho pelas contribuições para a finalização do texto e por aceitar o convite de participar da banca de qualificação e defesa.

Aos professores Harmud Stadtler da Universidade de Darmstadt e Horst Tempelmeier da Universidade de Cologne, pela grande ajuda dividindo suas belas publicações sobre o assunto aqui discutido.

Aos professores do Departamento de Fabricação da Faculdade de Eng. Mecânica da UNICAMP pelos importantes ensinamentos relacionados a planejamento de produção e operações.

Aos meus pais Sonia e Mário pelo amor, apoio e paciência quando muitas vezes me ausentei dos afazeres familiares.

Finalmente, eu agradeço a minha linda esposa Alena pela revisão final do texto, motivação e enorme apoio durante o longo período de mestrado.

Felipe Sawaya Hischheimer,

Resumo

Nas últimas duas décadas os sistemas de suporte ao planejamento da produção têm sido foco de melhorias por parte dos desenvolvedores de sistemas, empresas e pesquisadores. Os *MRPs*, amplamente utilizados por empresas de diversos segmentos demonstraram, ao longo dos anos, algumas deficiências principalmente no que se refere à gestão de capacidades dos recursos e flexibilidade na programação detalhada da produção. Os Sistemas Puxados, embora tenham sido utilizados como resposta para algumas das deficiências dos *MRPs*, também apresentaram falhas em gerenciar estruturas de materiais complexas e se adaptar a demandas instáveis. Da mesma forma, o *OPT* mesmo apresentando bons métodos para gerenciar o planejamento da produção em ambientes complexos, mostrou-se limitado devido à grande necessidade analítica para gerenciá-lo e dependência de softwares especialistas.

Em função da deficiência destes sistemas, da melhoria no processamento de dados derivados dos avanços da tecnologia computacional e da procura por aumento da eficiência no planejamento das operações das empresas, surgem os *APSs*.

Esta dissertação procura contextualizar os chamados *APSs* dentro da estrutura, processos e sistemas que envolvem o planejamento e a organização da produção. Mais especificamente, o texto explora uma forma específica de implementação dos *APSs*: a utilização de algoritmos de otimização linear.

Um dos grandes objetivos do trabalho é a apresentação de um estudo de caso sobre a implementação de um *APS* que utiliza otimização linear em uma empresa de produção de gelatina. O estudo demonstra como tais sistemas são idealizados e implementados nas empresas e de que forma podem aumentar a performance dos processos de planejamento. Nesse sentido, o texto destaca os modelos matemáticos que envolvem o sistema de otimização linear, bem como a forma com que estes são implementados em ambientes de planejamento da produção.

O estudo de caso também revela as dificuldades e caminhos para implementar os modelos complexos para o suporte da tomada de decisões no dia a dia das empresas. Problemas relacionados a limitações de processamento computacional, bem como a dificuldade de gerenciar os recursos humanos que irão manusear os novos sistemas de planejamento são abordados e demonstram poder corromper a implementação dos *APSs*.

Palavra Chave: Planejamento da Produção, *APS*, Otimização Linear;

Abstract

In the last two decades systems to support production planning have been the improvement focus of system developers, companies and researchers. The MRPs, widely used by various enterprises have shown, over the years, certain shortcomings principally to manage resource capabilities and flexibility in programming detailed production. Although the Pulled Systems have been used as a response to some of the MRP restrictions, they also had failure in managing complex material structures and adapting to demand instability. Likewise, the OPT even with good methods to manage production planning in complex environments, have showed some limitation due to the great analytic need and specialized software requirements.

Due to the deficiency of these systems, the improvement in data processing derived from advances in computer technology and the demand for increasing efficiency in the operational planning, APSs emerged.

This dissertation intends to analyze the so called APS within the structure, processes and systems involved in the planning and organization of production. More specifically, the text exploits a specific form of APSs implementation: the linear optimization algorithms.

One of the major goals of this work is to present a study case on the implementation of an APS in a gelatin production company. The study demonstrates how such systems are devised and implemented in companies, and how they can enhance operational performance. In this sense, the text emphasizes the mathematical models involved the linear optimization systems as well as how they are implemented in production planning environment.

The study case also reveals the difficulties and ways to implement complex models to support decision making in day by day production management. Problems related to computational process limitations and the difficulties of managing the human resources that will handle the new planning systems are also discussed in the text.

Keyword: Production Planning, APS, Linear Optimization

Lista de figuras

Figura 2.1: Estrutura de materiais em um <i>MRP</i>	13
Figura 2.2: <i>MRP II</i> e interface com <i>S&OP</i> e <i>SFC</i>	16
Figura 2.3: Representação do fluxo de informações e matérias em um sistema puxado	19
Figura 2.4: Processo de funcionamento do <i>Kanban</i>	20
Figura 2.5: Ilustração do funcionamento do <i>Heijunka</i>	21
Figura 2.6: Esquema de funcionamento do sistema Tambor, Pulmão, Corda	25
Figura 2.7: Sistema de programação de Tambor, Pulmão, Corda.....	26
Figura 2.8: Configuração da otimização linear em substituição ao funcionamento do <i>ERP</i> ...	30
Figura 2.9: Estrutura de módulos em que os <i>APSs</i> são operados	31
Figura 2.10: Estrutura de funcionamento dos <i>APSs</i>	32
Figura 2.11: Resolução gráfica de otimização linear	37
Figura 2.12. Resolução gráfica do modelo Simplex.....	40
Figura 2.13: Exemplo da resolução gráfica do método <i>Branch and Bound</i>	45
Figura 2.14: Árvore de busca do método <i>Branch and Bound</i>	46
Figura 2.15: Demonstração do plano de corte.....	48
Figura 3.1: Processo de construção de modelagem matemática	58
Figura 3.2: Mapeamento de processo;.....	59
Figura 4.1: Macro mapa de fluxo de valor do processo de planejamento e programação da produção de gelatina.....	69
Figura 4.2: Esquema Ilustrativo do sistema de programação de moagem sem o uso do <i>APS</i> .	71
Figura 4.3: Esquema ilustrativo do processo de programação da produção sem o uso do <i>APS</i>	73
Figura 4.4: Representação real de restrições de <i>mix</i> de produção	75
Figura 4.5: Representação dos processos iniciais de fabricação de gelatina.....	77
Fonte: Arquivos da Gelita.....	77
Figura 4.6: Representação das etapas finais de produção de gelatina: extração, separação e filtração	77
Figura 4.8: Esquema ilustrativo do <i>OBS</i>	82
Figura 4.9: Esquema ilustrativo do <i>GRP</i>	86
Figura 4.10: Representação do problema da capacidade de produção e a inter temporalidade	89

Figura 4.11: Representação gráfica da equação de balanço do sistema de planejamento da produção	90
Figura 4.12: Esquema ilustrativo do sistema de controle da produção	91
Figura 4.13: Protótipo do OBS com 3 propriedades de gelatina e 3 ordens de produção	95
Figura 4.14: Tela de customizações para flexibilização do sistema <i>OBS</i>	103
Figura 4.15: Interface com usuário do <i>OBS</i> para manipulação das propriedades das ordens de moagem	104
Figura 4.16: Interface com o usuário do sistema <i>GRP</i> e suas principais funcionalidades	105
Figura 4.17: Exemplo hipotético de nivelamento de demanda em função de simulação do <i>GRP</i>	113

Lista de tabelas:

Tabela 2.1: Definição das estratégias de produção.....	10
Tabela 2.2: Níveis x atividades do planejamento da produção	11
Tabela 2.3: Tipo de fluxos e suas características.....	12
Tabela 2.4: Dados em um roteiro de produção.....	14
Tabela 2.5: Tempos para abertura de ordem de produção.....	15
Tabela 2.6: Tabela simplex com x_1 e x_2 como variáveis não básicas e x_3 , x_4 e x_5 como variáveis básicas.....	41
Tabela 2.7: Tabela simplex com x_1 e x_5 como variáveis não básicas e x_3 , x_4 e x_2 como variáveis básicas.....	42
Tabela 3.1: Tabela de organização de diagnóstico do processo;	60
Tabela 4.1: Resumo dos problemas no processo de programação de moagem sem o uso do APS	72
Tabela 4.2: Propriedades não contempladas no processo de planejamento de produção de produção sem o uso do <i>APS</i>	74
Tabela 4.3. Resumo dos problemas no processo de planejamento de produção sem o uso do APS	76
Tabela 4.4: Resumo dos problemas no processo de controle de produção sem o uso do <i>APS</i>	80
Tabela 4.5: Representação da exclusão de ordens de moagem e pedido da base de variáveis	102
Tabela 4.6: Interface em tabelas para a identificação dos lotes que fizeram com que a taxa de sucesso não fosse atingida	109

Lista de gráficos

Gráfico 4.1: Planejamento x realização da produção de gelatina em uma planta da Gelita.....	78
Gráfico 4.2: Tempo de processamento em 3 testes com modelo completo	98
Gráfico 4.3: Tempo de processamento de 3 testes com restrição da tolerância de soluções inteiras	100
Gráfico 4.4: Interface do atendimento por planta do programa Taxa de Sucesso.....	107
Gráfico 4.5: Interface do atendimento por perfil de produção do programa Taxa de Sucesso	108
Gráfico 4.6: Entregas perfeitas / estoque antes e depois da implementação do APS	111
Gráfico 4.7: Comparação de tempo médio de processamento da programação de mistura de 70 toneladas de gelatina com o sem o uso do <i>OBS</i>	112
Gráfico 4.8: Comparação do tempo de processamento do planejamento de produção com o sem o uso do <i>GRP</i>	114

Lista de abreviações:

APICS American Production and Inventory Control Society

APO Advanced Planning and Optimizing

APS Advanced Planning Scheduling;

ASCP Advanced Supply Chain Planning

ATO Assembly to Order

BOM Bill of Material

ConWip Constant work in Process

CRP Capacity Requerement Planning

EDD Early Due Date

ERP Enterprise Resource Planning;

ETO Engineer to Order

GRP Gelatine Resouce Planning

MPS Master Production Plan

MTO Make to Order

MTS Make to Stock

MRP Material Resource Planning

OBS Optimized Blending System

ORS Operational Research Society

ORSA Operational Research Society of America

OM Ordem de Moagem

OPT Optimized Production Tecnology

PDCA Plan Do Check and Analyze

RCC Recurso Restritivo Critico

RCCP Rough Cut Capacity Planning

SCF Sistema de Capacidade Finita

SCOOP Scientific Computation of Optimization programs

SFC Shop Floor Control

SKU Stock Keeping Units

S&OP Sales and Operational Plan

SPT short process time

TS Taxa de Sucesso

Sumário

Capítulo 1: Introdução:.....	1
1.1. A evolução da pesquisa operacional.....	2
1.2. A Pesquisa operacional no planejamento e organização da produção.	4
1.3. Definição do problema	5
1.4. Restrições do problema	6
1.5. Objetivos.....	7
1.6. Organização do trabalho.....	7
Capítulo 2: Revisão da bibliografia sobre sistema de produção.....	9
2.1.O <i>MRP II</i>	12
2.2. Sistema puxado.....	18
2.3. <i>OPT</i> - O sistema Tambor Pulmão Corda (<i>Drum Buffer Rope</i>)	22
2.4. <i>APSs</i> (Sistema de Capacidade Finita).....	27
2.5. Otimização linear.....	34
2.5.1. Forma gráfica da otimização linear	36
2.5.2. Método Simplex	38
2.5.3. O método <i>Branch and Bound</i>	44
2.5.4. O método de <i>Cut Plane</i> (plano de corte)	47
2.5.5. O modelo <i>Branch and Cut</i>	49
2.6. <i>APSs</i> de otimização	50
2.6.1. Modelos de dimensionamento de lotes (<i>lot sizing</i>)	51
2.6.2. Modelos de sequenciamento da produção (<i>scheduling</i>).....	54
2.6.3 Outros modelos de otimização aplicados no planejamento e organização da produção .	56
3. O Método para implementação de <i>APSs</i> de Otimização	57
3.1. Diagnóstico do processo.....	58

3.2. Interpretação e modelagem.....	60
3.3 Testes do modelo	62
3.4. Implementação.....	64
Capítulo 4 - Estudo de caso:	67
4.1. Diagnóstico do processo de planejamento de produção	68
4.1.1. Programação de moagem	70
4.1.2. Planejamento de produção.....	72
4.1.3. Controle da produção.....	78
4.2. Modelagem e Interpretação do modelo	80
4.2.1 O Modelo de misturas	81
4.2.2. O Modelo de planejamento da produção.....	85
4.2.3. O modelo de controle da produção.....	90
4.3. Testes do modelo	94
4.3.1. Teste em pequena escala.....	94
4.3.2. Teste em grande escala	95
4.3.2.1. Teste com otimização do saldo em estoque.....	98
4.3.2.2. Teste com alternativas de processamento de programação inteira.....	99
4.4. A Implementação do <i>APS</i> de otimização	102
4.5 Os Ganhos com o uso da otimização linear.....	110
5. Conclusões e considerações.....	115
5.1 Sugestão para estudos futuros.....	117
6. Referências bibliográficas:	118

Capítulo 1: Introdução:

Estudos de algoritmos para a organização e planejamento a produção de bens e serviços são frequentes desde a evolução da produção em série durante e após a segunda guerra mundial (MCCLOSKEY, FORENCE, 1956). Durante todo esse período, os modelos matemáticos e mais especificamente os modelos de otimização têm sido foco de intensos estudos para muitos pesquisadores e empresas (HILLIER, LIEBERMAN, 2010). Este trabalho está inserido neste contexto.

Em uma pequena pesquisa realizada nos bancos de dados das Universidades Estadual de Campina e Universidade de São Paulo, identificamos que existem diversas publicações sobre algoritmos de otimização. Também analisamos inúmeros trabalhos sobre diversos métodos para o gerenciamento e administração da produção e textos relacionados aos chamados *APSs* (*Advanced Planning and Scheduling*) ou sistemas avançados de planejamento e sequenciamento. Contudo, o número de publicações relacionadas ao uso de algoritmos de otimização como suporte a tomada de decisões em ambientes de produção é relativamente restrito.

Ao contrário de muitos trabalhos sobre algoritmos de otimização, esta dissertação não aborda testes empíricos de algoritmos de otimização com o objetivo de aumentar a capacidade processar as variáveis em um tempo mais curto de tempo. De mesma forma, nosso objetivo principal não é realizar uma pesquisa bibliográfica intensa sobre os principais sistemas de planejamento da produção e destacar suas principais virtudes e limitações.

Esta dissertação tenta contextualizar os *APSs* dentro da organização, dos processos e dos sistemas de gestão da produção. O texto propõe um método de implementação de *APSs* de otimização para a melhoria do gerenciamento do planejamento e organização da produção. Na sequência, o trabalho explora um estudo de caso da implementação de um *APS* de otimização em uma empresa de produção de gelatina.

O estudo destaca problemas pertencentes a um ambiente produtivo e de planejamento de produção que estão além da formulação e aplicação de um modelo matemático em um ambiente de planejamento. A dificuldade e os caminhos para a implementação do sistema, os problemas relacionados à adaptação dos recursos humanos e as customizações necessárias para a implementação de um modelo de otimização na prática, são também elementos abordados pelo trabalho.

A evolução da utilização de sistemas de otimização linear no planejamento da produção é um assunto que intriga gestores e pesquisadores da área de administração da produção. Assim,

antes de quaisquer colocações acerca da revisão bibliográfica, sobre a aplicação de modelos matemáticos no planejamento de produção ou do modelo elaborado em si, é importante contextualizar a aplicação de otimização linear em ambientes de produção dentro da evolução da Pesquisa Operacional.

1.1. A evolução da pesquisa operacional

A programação linear faz parte de um campo mais amplo da engenharia e matemática que atrela o uso de equações matemáticas à resolução de problemas que possam ser descritos em equações lógicas, chamado Pesquisa Operacional. A sociedade européia EURO define o termo como: “Uma abordagem científica para a solução de problemas no gerenciamento de sistemas complexos” (ARENALES, 2007 p. 3). A sociedade americana APICS (*American Production and Inventory Control Society*) designa a Pesquisa Operacional como: “teoria e metodologia em matemática, estatística e computação adaptada e aplicada à identificação, formulação, solução, validação implementação e controle dos problemas de tomadas de decisão”. Uma das melhores definições do termo, porém vem de Winston (1995 p.11):

“Abordagem científica para tomada de decisões que procuram determinar como melhor projetar um sistema usualmente sob condições que requerem a alocação ótima de recursos escassos”.

De forma geral e das definições acima colocadas, podemos dizer que Pesquisa Operacional é um enfoque científico sobre a tomada de decisões.

O surgimento desse termo está relacionado a invenção do radar e seu uso em um plano cartesiano na Inglaterra em 1934, no âmbito da segunda guerra mundial. O Ministério Britânico da Aviação, dois anos depois, criou a estação de pesquisa Manor Brawdsei em Suffolk para estudar como a tecnologia do radar poderia ajudar no processo de interceptação de aviões inimigos. O superintendente dessa estação, A. P. Rowe que coordenava estudos sobre o efeito do uso das técnicas de operações advindas do experimento com a interceptação de radar foi o inventor do termo: *Operational Research*. Em seguida, em 1941, foi criado um departamento de Pesquisa Operacional dentro dos serviços militares britânicos que estudava problemas de operações de guerra tais como inspeções de aviões, controle de artilharia aérea e probabilidades de distribuição de equipamentos inimigos (MCCLOSKEY, FORENCE, 1956).

Durante e após a segunda guerra mundial o uso de Pesquisa Operacional se ampliou principalmente nos Estados Unidos e na Inglaterra. Um dos projetos mais ambiciosos sobre o

assunto foi criado em 1947. O SCOOP (*Scientific Computation of Optimization programs*), pertencente às forças militares do Pentágono, tinha o objetivo de apoiar decisões operacionais na força aérea americana e contava com o economista Marshal Wood e com o matemático George Dantzig (MCCLOSKEY, FORENCE, 1956). Foi durante esse projeto que Dantzig, cientista bastante citado neste e em outros trabalhos que falam sobre otimização, formulou o método Simplex para desenvolver problemas de programação linear (HILLIER, LIEBERMAN, 2010). Dantzig, por sua vez se baseou no trabalho do matemático russo Leonid Kantorovich: Métodos matemáticos na organização e no planejamento de produção.

Já o termo programação linear foi sugerido a Dantzig por outro cientista da época, o economista T.C. Koopmans. Ambos, Knatorovich e Koopmans ganharam prêmio Nobel de economia pelas contribuições à teoria de alocação ótima de recursos fundamentada por estudos no campo da Pesquisa Operacional (ARENALES, 2007)

Não muito tempo depois da criação do SCOOP, em 1952 foi fundada a sociedade científica americana de Pesquisa Operacional, ORSA (*Operational Research Society of America*); e em 1953 os Ingleses criaram a ORS (*Operational Research Society*). Essas duas sociedades contribuíram com grande parte dos estudos sobre o tema até o presente momento.

No Brasil a Pesquisa Operacional iniciou-se na década de 1960 com poucos estudos científicos sobre o assunto. O primeiro simpósio sobre Pesquisa Operacional foi realizado em 1968 no ITA em São Jose dos Campos. O SOBRAPO (Sociedade Brasileira de Pesquisa Operacional) foi criado logo em seguida.

Na primeira conferência internacional sobre Pesquisa Operacional realizada em 1957 em Oxford na Inglaterra, foi possível constatar os diferentes focos nos trabalhos dos cientistas ingleses e dos americanos. Os primeiros enfatizavam muito mais trabalhos científicos teóricos enquanto os americanos estavam utilizando os conceitos de Pesquisa Operacional tais como estoque, teoria das filas, programação de máquinas de produção, teoria os jogos, fluxos em redes e otimização linear. (ARENALES, 2007).

Em um estudo feito em 1953 com 160 organizações inglesas, 45 alegaram que possuem um departamento que estudava Pesquisa Operacional ou que apresentavam ao menos uma pessoa dedicada ao tema. Em 1970 foram identificados 766 grupos de Pesquisa Operacional, dos quais 533 estavam no setor industrial ou em serviços financeiros. Foi durante a década de setenta que a Pesquisa Operacional deixou de ser utilizada preponderantemente nas associações sobre o assunto e por cientistas que se aventuram com modelos matemáticos e passou a ser aplicada nas mais diversas áreas de logística e produção da indústria automotiva,

de alimentação, aviação, computação e transporte, telecomunicações, metalurgia, mineração, eletrônica entre outras (ARENALES, 2007).

Os primeiros livros sobre Pesquisa Operacional falavam sobre o tópico de maior uso do tema: a programação e otimização linear, tema de nosso trabalho. *Linear Programming Methods and Application* (Gass, 1958) e *Linear Programming and Extensions* (Dantzig, 1963), exploram os métodos de programação linear e suas aplicações, passando pelo modelo Simplex e pelos vários usos de programação linear nas empresas de produção e distribuição de bens e serviços (ARENALES, 2007).

O termo Pesquisa Operacional está amplamente relacionado à construção de modelos matemáticos que tentam resumir ou representar problemas reais e à alocação de recursos para tomada de decisão. A grande maioria das definições concorda que os modelos matemáticos constituem o cerne da Pesquisa Operacional e de sua ramificação mais popular, a programação linear.

1.2. A Pesquisa operacional no planejamento e organização da produção.

A Pesquisa Operacional e a otimização linear especificamente se relacionam com as decisões no planejamento da produção desde a década de 60 como foi colocado na sessão anterior. Entretanto, seu uso sempre ficou restrito ao suporte do planejamento na tomada de decisão em ambientes simples sendo utilizado em bases de dados independentes do restante das informações das empresas, como é possível se observar em *Linear Programming and Extensions* de 1963, por Datzig.

Recentemente, com o uso mais intensivo da informática nas empresas, a otimização linear vem sendo aplicada como ferramenta de suporte às decisões na produção através dos chamados *APSs*. (*Advanced Planning and Scheduling*). (STADTLER, KILGER, 2005). Os *APSs* são na realidade os nomes atribuídos atualmente aos sistemas de gerenciamento de capacidade finita criados para superar as dificuldades de gestão da produção e cadeia de suprimentos derivados dos *MRPs* (*Material Resouce Planning*) e outros sistemas de suporte a produção (GIROTTI e MESQUITA, 2011).

Günther e Beek (2003) definem os *APSs* como sistemas de suporte a decisão na cadeia de suprimentos nos níveis de planejamento de longo, médio e curto prazo. Usualmente esses sistemas utilizam modelos matemáticos para auxiliar os planejadores a alocar recursos escassos no ambiente de produção e na cadeia de suprimentos.

O uso dos *APSs* é relativamente novo nas empresas e sua aplicação pode abranger desde a alocação de recurso em níveis estratégicos do gerenciamento das cadeias de suprimentos até o sequenciamento do chão de fábrica (STADLER, KILGER, 2005). Sua aplicação vem crescendo e cada vez mais tais sistemas vêm cumprindo o papel importante de suportar a tomada de decisões no planejamento e organização da produção (GIACON, 2010).

Como iremos detalhar adiante, os *APSs* não necessariamente utilizam algoritmos de otimização linear para solucionar problemas de gestão planejamento. O uso de otimização linear é um método específico de funcionamento dos *APSs*. Sua aplicação é um dos focos de desenvolvimento e evolução dos sistemas *APSs* (PEDROSO e CORREA, 1996).

Pode-se dizer que este texto é uma dissertação que aborda a forma com que os *APSs* se contextualizam com os outros elementos pertencentes ao planejamento de produção, bem como descreve seu desenvolvimento em uma empresa de produtos alimentícios.

1.3. Definição do problema

Toda a pesquisa tem início com algum tipo de problema. Para Gil (1995) um problema é algo que dá margem a hesitação ou a perplexidade por ser difícil de concluir ou de resolver. Para a ciência, problema é qualquer questão não solucionada e que é objeto de discussão em qualquer domínio do conhecimento. Este trabalho se encaixa nestas duas definições.

O problema que se pretende destacar neste trabalho consiste em responder às 3 seguintes perguntas:

1. Como os *APSs* e sua forma específica, a otimização linear, se relacionam com a estrutura, o processo e os vários sistemas utilizados para a organização e planejamento de produção nas empresas?
2. Como implementar um sistema *APS* de otimização linear para melhorar a performance do processo de planejamento em um ambiente de produção complexo?
3. Quais as dificuldades e caminhos para se implementar um sistema de otimização para o suporte das tomadas de decisões no planejamento de produção de uma empresa?

Ao longo dos próximos capítulos, o trabalho irá tentar responder estas três perguntas. Uma pesquisa bibliográfica com elementos relacionados à utilização de sistemas de suporte a produção e modelos de otimização nas empresas forma a resposta da primeira pergunta. Para responder a segunda pergunta, detalharemos as fórmulas e métodos matemáticas utilizados na

aplicação de modelos de otimização linear, bem como explorar o método utilizado para a implementação de um *APS* em empresa de produção de gelatina. Por fim, as diversas dificuldades do período de implementação do modelo serão analisadas na tentativa de explicar as complicações de se operacionalizar um *APS* na prática. Isso deve endereçar a terceira pergunta.

1.4. Restrições do problema

Dentre os vários seguimentos de estudo que a Pesquisa Operacional pode nos levar, este trabalho irá se restringir à otimização linear aplicada a processos de organização e planejamento de produção. Por otimização linear pretende-se dizer o segmento da Pesquisa Operacional que estuda os algoritmos exatos que respeitam as condições e hipóteses de linearidade e que exploram os espaços da solução com o objetivo de se conseguir uma resposta ótima do problema (TAHA, 2008).

Embora ao longo da dissertação se explore melhor os conceitos de “organização e planejamento de produção”, por agora é importante assimilar que o termo se restringe ao processo utilizado para alocar recursos de modo a transformar e produzir bens e serviços dentro de um arranjo produtivo (SLACK, 2009).

O texto irá percorrer uma pesquisa bibliográfica que engendra a utilização dos *APSs* nos níveis tático e operacional do planejamento da produção e como estes se relacionam com os outros sistemas, estrutura e processos pertencentes ao planejamento em ambientes produtivos. Nesse sentido, iremos analisar apenas os casos específicos de aplicação de otimização em modelos determinísticos; aqueles em que são conhecidas as variáveis estudadas (TAHA, 2008).

O método para a implementação do *APS* de otimização abordado não tem a intenção de generalizar ou propor um método de implementação de sistemas de suporte para processo de planejamento, mas pode servir como um bom caminho.

As considerações sobre a aplicação de otimização linear em outros ambientes de produção e indústrias serão meras referências às bibliografias ou pequenas constatações e observações da experiência em realizar projetos de consultoria em outras empresas.

De forma sucinta, a dissertação irá se restringir à contextualização do uso dos *APSs* no planejamento da produção e análise da aplicação de um modelo otimização linear, caso específico de implementação do *APS*, em uma empresa de misturas.

1.5. Objetivos

Os objetivos deste trabalho são fundamentalmente as respostas e contextualização das perguntas colocadas na definição do problema. Sendo assim, os objetivos gerais do trabalho são:

1. Contextualizar os *APSs* e sua forma específica, a otimização linear, na estrutura, processos e sistemas utilizados para a organização e planejamento de produção nas empresas;
2. Descrever um método para implementação dos *APSs* de otimização linear para melhorar a performance do processo de planejamento em ambientes de produção complexos;

E o objetivo principal e específico da dissertação é:

3. Analisar um estudo de caso sobre a implementação e operacionalização de um sistema de otimização para o suporte das tomadas de decisões operacionais em uma empresa do segmento alimentício;

Contudo, ao tentar responder a essas perguntas, alguns outros aspectos sobre o tema são indiretamente discutidos. Nesse sentido, ao se contextualizar o uso dos *APSs* dentro dos conceitos de planejamento e organização da produção para o direcionamento da primeira pergunta, critica-se alguns trabalhos acadêmicos e considerações que exploram o tema.

Da mesma forma, ao descrever o processo de construção de um *APS* de otimização, o trabalho impõe um método geral para a construção de modelo de otimização linear para o planejamento de produção nas empresas.

1.6. Organização do trabalho

Este trabalho está estruturado em cinco partes: Após esta breve introdução, no capítulo 2 é realizada uma revisão bibliográfica referente à contextualização das aplicações dos *APSs* nas empresas e sua integração com outros métodos e sistemas dentro dos processos de planejamento e organização da produção. Nesse capítulo os sistemas *MRPs*, *Kanban* e *OPT*

são analisados, bem como os principais modelos de otimização linear aplicados nas empresas. Os modelos de otimização linear, o algoritmo *Simplex* e os métodos que são usualmente utilizados nos modelos de otimização linear de problemas com variáveis inteiras são explorados também no capítulo 2. Nesse sentido, iremos dedicar alguns parágrafos na dissertação sobre método *Branch and Bound* (ramificação e delimitação) e o método *Cut Plane* (plano de corte)

No capítulo 3 descrevemos o método utilizado para a construção e implementação do *APS* de otimização utilizado no estudo de caso.

No capítulo 4 se apresenta o estudo de caso da aplicação de um modelo otimização desenvolvido na empresa Gelita. Nessas sessões são explorados o diagnóstico, o modelo matemático desenvolvido, os testes elaborados e a implementação do modelo na prática.

O capítulo 5, última parte deste trabalho, está reservado para as considerações finais e conclusões desta dissertação.

Capítulo 2: Revisão da bibliografia sobre sistema de produção

O planejamento e a programação da manufatura são atividades das mais complexas para a grande maioria das empresas, abrindo espaço para o desenvolvimento de ferramentas que auxiliem a busca de soluções satisfatórias para estes problemas (HADDAD e CARVALHO, 2003). Slack (2009) define o termo “planejamento e programação da produção” como o processo de alocar recursos escassos para a execução de tarefas em um horizonte de tempo. A otimização linear pode substituir a tomada de decisão intuitiva para problemas complexos por uma abordagem que identifica a melhor alternativa segundo algum critério de análise previamente estabelecido (GAITHER e FRAZIER, 2004). Este capítulo tem o objetivo de contextualizar o uso dos APSs e sua forma específica, a otimização linear, no gerenciamento da produção.

Stadtler e Kilger (2005) afirmam que os APSs são uma resposta a algumas deficiências dos métodos e sistemas de planejamento e programação da produção das empresas. Essa colocação nos conduz ao importante processo de estudar o que leva as empresas e os pesquisadores a realizem esforços para a implementação de APSs e sistema de otimização em ambientes de produção. Para isso é necessário explorar as características fundamentais de outros processos de suporte a tomada de decisão no planejamento e organização da produção que antecederam os APSs.

A tarefa de sintetizar os vários modelos utilizados no planejamento e organização da produção, no entanto, não é simples. Isso porque existem diferentes sistemas computacionais e métodos utilizados para suportar a produção nos diversos tipos e ambientes de produção. Ao mesmo tempo, existem concepções sobre os diferentes sistemas de planejamento que podem ser considerados equivocados se generalizados para todas as suas aplicações. Assim, descreveremos os sistemas de planejamento e programação de forma geral, sem entrar no mérito das variações que podem existir entre os fornecedores de sistemas de suporte e empresas.

Antes de iniciarmos a descrição dos sistemas *MRPII*, *Produção Puxada* e *OPT (Optimized Production Technology)*, é relevante pontuar as diferentes estratégias de manufatura, tipos de produção e atividades relacionadas aos níveis de planejamento. Esses elementos são de grande importância para a contextualização dos sistemas e métodos de planejamento bem como para entender melhor a forma de utilização dos APSs nas empresas.

Slack (2009) define quatro estratégias de manufatura e argumenta que tais estratégias estão relacionadas à natureza do negócio e à dinâmica de produção das empresas.

Estratégias	Descrição
<i>Make to Order (MTO)</i>	Fazer para a ordem. A produção só é realizada após a sinalização real do pedido do cliente
<i>Make to Stock (MTS)</i>	Fazer para o estoque. A produção é realizada para um estoque sem a necessidade de um pedido do cliente.
<i>Assemble to Order (ATO)</i>	Montagem para a ordem. A montagem final é feita em função de um pedido formal, porém o resto da produção realizada para estoque
<i>Engineer to Order (ETO)</i>	Desenvolver para a ordem. O cliente participa no desenho do projeto e a produção o atende

Tabela 2.1: Definição das estratégias de produção

Fonte: Adaptado de Slack (2009)

O *MRPII* que será analisado a seguir pode ser aplicado em qualquer uma das estratégias como algumas ressalvas e restrições. Entretanto, o *Kanban* é usualmente utilizado nos casos em que a produção é feita para um estoque MTS (HOOP, 2008). O *OPT* é um sistema aplicável a todos as estratégias acima principalmente em ambientes complexos (CORREA e GIANESI, 2011).

As atividades referentes aos níveis hierárquicos do planejamento e organização da produção é um aspecto que será abordado durante este e os demais capítulos. A tabela abaixo adaptada de Lustosa (2008) é importante para o estabelecimento das atividades que compõe os diferentes níveis do planejamento de produção.

Níveis	Atividades
Estratégico	Definição do portfólio, Definição da localização; Definição da capacidade para atender a demanda; Nível de serviço;
Tático	Plano mestre de produção; Plano de suprimentos; Alocação de recursos em função da demanda; Avaliação de performance e desempenho da produção geral;
Operacional	Programação da Produção; Gestão dos estoques; Gestão da performance da produção (planejado x real);

Tabela 2.2: Níveis *versus* atividades do planejamento da produção
Fonte: Adaptado de Lustosa (2008)

É comum o uso de sistemas distintos e diferentes arquiteturas de controle para os diferentes níveis de planejamento das organizações.

O tipo de sistema e métodos de planejamento utilizadas nas operações também depende em grande medida do tipo de produção que é realizado (CORREA e GIANASI 2011). A tabela abaixo mostra os diferentes tipos de produção e suas principais características.

Tipo de Produção	Características
Máquina Única	Existe uma única máquina para o processamento das tarefas
Máquina paralela	Em único estágio de produção há duas ou mais máquinas disponíveis que podem executar qualquer tarefa. Cada tarefa pode ser executada em uma máquina
<i>Flow Show</i>	Todas as tarefas possuem o mesmo fluxo de processamento nas máquinas
<i>Job Shop</i>	Cada tarefa tem seu fluxo individual ou rota nas máquinas;
<i>Open Shop</i>	Não há fluxo determinado para cada tarefa;

Tabela 2.3: Tipo de fluxos e suas características

Fonte: adaptado de Slack (2009)

As tarefas são conjuntos de operações inter-relacionadas por restrições de procedência. As restrições de procedência definem as rotas das operações, ou seja, a ordem em que são processadas. Um conjunto de tarefas compõe a produção de uma peça ou de um produto na manufatura dos bens.

Os tipos de produção definem, em última instância, a complexidade dos modelos matemáticos utilizado nos APSs e, desta forma, elegem os sistemas possíveis ou não de serem implementados em cada ambiente.

Utilizaremos as três tabelas acima ao longo deste capítulo para contextualizar os diversos métodos de planejamento de produção e suas peculiaridades.

Os parágrafos seguintes abordarão o *MRP II*; segundo Correa e Giansi (2011), o sistema usualmente utilizado pelas empresas nos processos de planejamento de processos discretos. O Sistema Puxado que nasceu na indústria automotiva japonesa na década de (TARDIN, 2001), será analisado em seguida. Depois disso iremos caracterizar o sistema Tambor, Pulmão Corda, metodologia de planejamento de produção pertencente ao *OPT* (GOLDRATT, 1990). Após a explorar estes três sistemas, explicamos melhor os APSs.

2.1.O MRP II

Os *ERPs* (*Enterprise Resource Planning*) são os sistemas usualmente utilizados para facilitar o gerenciamento das atividades dentro dos processos de planejamento (HADDAD e

CARVALHO, 2003). Estes sistemas gerenciam todas as operações das empresas e contem, entre outros, os módulos para a gestão financeira e de custos, suprimentos, vendas, distribuição e planejamento da produção. Dentro das várias funções existentes nos módulos dos *ERPs*, o *MRP*, introduzido na década de 1960 por Joe Orlicky e J. Case (CORREA e GIANESI, 2011), tem grande importância para gestão dos recursos e estoques das fábricas.

Usualmente, o *MRP* é um sistema que converte previsões de vendas em um programa de produção e compra de peças identificando quando e em que quantidade se deve processar os materiais (HOOP, 2008). Para isso, o sistema faz uso de uma estrutura de materiais com itens de demanda dependentes e itens de demanda independentes. A diferença entre essas duas classes de itens jaz na forma com eles se relacionam com a demanda de produtos das empresas. Os itens de demanda independentes são aquelas cuja quantidade necessária depende do mercado ou da demanda exógena das empresas. Normalmente, a quantidade desses produtos é determinada por uma previsão de vendas ou pelo histórico de demanda. Já os itens dependentes são aqueles cuja quantidade a ser produzida ou comprada depende diretamente do consumo dos itens de demanda independente (CORREA e GIANESI, 2011).

A figura 2.1 adaptada de Arenales (2007) ilustra uma estrutura hipotética de materiais para três itens finais, referente ao funcionamento do *MRP*:

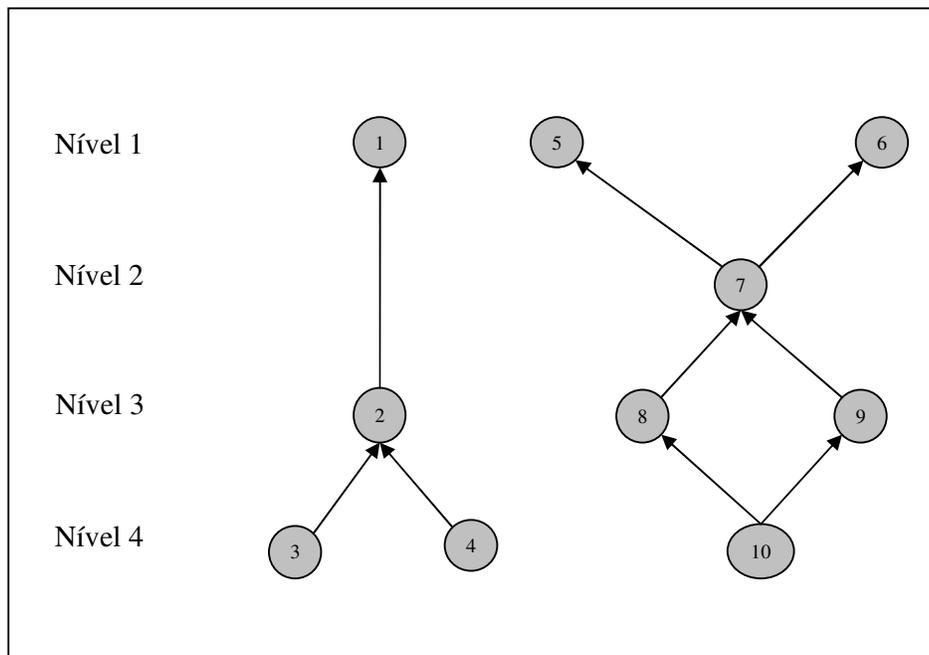


Figura 2.1: Estrutura de materiais em um MRP
Fonte: Adaptado de Arenales, 2007

OS itens 1, 5 e 6 são os itens no *nível 1* e representam os itens finais de demanda independente que são demandados em função das vendas realizadas ou previstas. O item 7 e os itens 2, 8 e 9 pertencem, respectivamente, aos níveis 2 e 3 da estrutura de materiais e são manufacturados. Por último, os itens 3,4 e 10 pertencem ao 4º nível e são comprados pela empresa. Todos os itens com exceção dos itens finais são de demanda dependente na estrutura de produtos.

Segundo Correa e Giansi (2011) os planeadores “importam” o planeamento de vendas que contém as datas e a quantidades de itens que se vendeu ou que se espera vender em um determinado período para as bases de dados do *MRP*. O plano de produção de itens finais é denominado *MPS (Master Production Plan)*. Esses itens de demanda independente estão relacionados aos itens de demanda dependentes, de tal forma que “automaticamente”, ao se planejar o consumo de itens de demanda independentes, a compra ou produção de itens de demanda dependente é realizada.

A relação “automática” entre os itens é dada pela estrutura dos produtos e pelo cadastro dos tempos de processamento e entrega dos componentes (HOOP, 2008). Neste sentido, o sistema *MRP* normalmente estabelece uma sistemática de “data mais tarde”, em que a data de solicitação é sempre feita estabelecendo-se a data mais tarde possível para o atendimento da demanda. Fazendo menção à figura 2.1 e dados os tempos hipotéticos abaixo:

Itens	Níveis	Tempo de entrega/processamento
1	Nível 1	x horas
2	Nível 2	y horas
3	Nível 3	w dias
4	Nível 4	z dias

Tabela 2.4: Dados em um roteiro de produção
Fonte: Adaptado de Arenales, (2007)

Se um cliente demanda o item 1 no período t , então o *MRP* sinaliza, por intermédio de uma ordem de produção, que o item 1, 2, 3 e 4 deve ser feito nos seguintes tempos:

Itens	Níveis	Sinalização para elaboração das ordens de produção
1	Nível 1	T- x horas
2	Nível 2	T - x - y horas
3	Nível 3	T - x - y horas - w dias
4	Nível 4	T - x - y horas - z dias

Tabela 2.5: Tempos para abertura de ordem de produção

Fonte: Adaptado de Arenales (2007)

Hoop (2008) descreve o funcionamento do *MPP* da seguinte forma: após a definição das quantidades de produtos finais e suas respectivas datas de entrega para a preparação do *MPS*, são realizadas quatro etapas:

- **Determinação da quantidade líquida (*netting*):** Nesta etapa é feita a definição da necessidade líquida em função da necessidade bruta, dos estoques planejados, dos estoques existentes e projetados.
- **Definição do tamanho dos lotes (*lot sizing*):** Nesta fase a necessidade líquida de produtos é decomposta em lotes adequados de processamento, formando ordens de produção e ou ordens de compra de produtos finais;
- **Preparação das etapas (*time phasing*):** nesta fase são determinadas as datas de início dos trabalhos em função dos tempos de processamento ou compra (*lead times*) cadastrados no sistema;
- **Explosão da lista de matérias (*BOM* ou *bill of material explosion*):** nessa etapa a lista de materiais é explodida e os tempos de processamento ou compras dos componentes também são determinados em função dos tempo de processamento e compra e da lista de matérias.

Segundo Correa e Gianesi (2011) o *MRP* inicialmente não permitia a visão das restrições de capacidade de produção dos recursos. Os sub módulos, *RCCP* (*Rough Cut Capacity Planning*) e o *CRP* (*Capacity Requirement Planning*), introduzidos na evolução do *MRP* para o *MRP II* possibilitam a visualização e o apontamento da capacidade dos recursos

da fábrica em função do plano de venda em interação com a estrutura de materiais e os tempos de processamento.

O *MRP II* pode ser considerado uma evolução do *MRP* que visava a superação de algumas deficiências importantes do antigo sistema. Ainda assim, embora muito se tenha feito, o sistema ainda é sujeito a falhas como analisaremos adiante.

A figura 2.2 demonstra o funcionamento e os componentes do *MRP II* segundo Correa e Gianesi (2011). O *S&OP* (*Sales and Operational Plan*) será explorado posteriormente nesse trabalho.

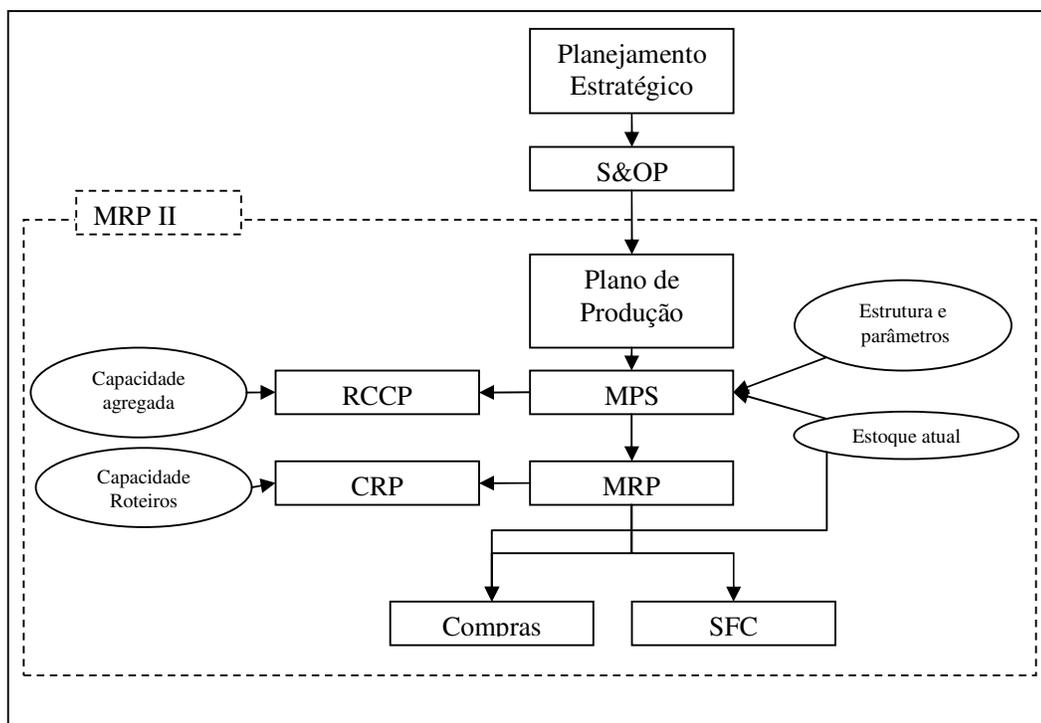


Figura 2.2: MRP II e interface com S&OP e SFC
 Fonte Adaptado de Correa e Gianesi (2001)

O *MRP* gerencia a produção e a gestão de materiais sem a visão de uma capacidade de produção, enquanto o *MRP II* identifica e processa tais limitações apontando em que período e em que quantidade a capacidade de produção foi superada pela demanda (CORREA e GIANESI, 2011). Para isso, o *RCCP* restringe a capacidade da produção ainda em um nível agregado de processamento, enquanto o *CRP* restringe a capacidade depois que o *MRP* já desdobrou o planejamento de vendas em componentes e peças para compras e produção.

Segundo Hoop (2008), o *RCCP* utiliza apenas os recursos críticos relacionados aos itens finais e os componentes derivados de sua explosão. Ou seja, o *RCCP* funciona no nível

do *MPS (Master Planning System)*. Já o *CRP* faz uma interação mais detalhada contemplando a posição dos estoques em processo, roteiro de todas as peças e considera todos os centros de produção.

Correa e Gianesi (2001) argumentam que os *MRPs* podem ser eficientes para o planejamento de itens com estruturas longas e complexas, com vários níveis de componentes e um número grande de itens por nível. As estruturas de materiais em interação com os roteiros de produção e seus tempos auxiliam no processo de gerenciamento dos momentos e quantidades que os produtos devem ser processados.

Haddad e Carvalho (2003) alegam que o *MRP II* não é um sistema suficiente para se gerenciar ou para suportar o gerenciamento da produção. Segundo tais autores, caso exista a sobrecarga de algum recurso em um determinado período, o sistema apenas permite a visualização dessa sobrecarga, sem sugerir uma solução melhor para a alocação dos recursos. Os mesmos alegam que um processo de otimização poderia sugerir a antecipação da produção e compra de alguns itens, bem como a alteração da sequência de produção e ou a postergação de processamento de itens.

Petroni (2002) destaca que um grande problema do *MRP II* é a falta de descrição e exibição do plano detalhado de produção. Segundo o autor, o *MRP II* não trata de forma viável a questão das capacidades dos recursos pois o *CRP* só restringe a produção segundo a capacidade dos recursos depois que o *MRP* foi processado. Ou seja, os planejadores têm que processar o *MRP* novamente se desejarem elaborar um plano factível.

O problema da falta de um plano detalhado de produção está intimamente relacionado aos problemas dos tempos de processamentos (*lead times*) fixos colocados por Hoop (2008). O fato do *MRP* sempre calcular os tempos de entrega e processamento em função do tempo de processamento cadastrado no sistema, sem levar em consideração o nível de trabalho dos recursos, faz com que as datas colocadas para o processamento possam ser diferentes das datas em que realmente as peças serão processadas.

Segundo Giacon (2008), existem ainda duas falhas na forma com que o sistema *MRP* gerencia e planeja a produção. A primeira está relacionada à falta de utilização de roteiros alternativos na reprogramação em função das quebras ou falhas no sistema produtivo. E a segunda, esta na lógica de prioridade baseada nas datas dos pedidos, desconsiderando outros parâmetros como número de trocas de ferramentas (*set ups*), prioridade de clientes e outros parâmetros que possibilitariam a melhor alocação dos recursos.

2.2. Sistema puxado

Segundo Liker (1998), o Sistema Puxado que pertence ao método de gestão de matéria *Just in Time* (no tempo certo), tem o objetivo de planejar e controlar a movimentação e produção de matérias com menor número de recursos e estoque possível, atingindo um nível de serviço desejável. Para isso, o Sistema Puxado faz uso de um sistema de sinalização que funciona de acordo com os níveis de estoque de peças em pontos da produção em que não existe um fluxo de peças (MOLDEN, 1998).

Segundo Womack et. al. (1998), o sistema puxado foi criado pela Toyota Motors, empresa fabricante de carros no período após a Segunda Guerra Mundial. O sistema foi baseado nas gôndolas de produtos dos supermercados ocidentais.

De acordo com Womack et. al. (1998) os sistemas de planejamento puxados são implementados quando não se consegue criar um fluxo determinado de peças nas etapas de processamento. Desta forma, no processo de produção de um produto, quando não é possível realizar um fluxo contínuo, como o processo de suprimento de componentes para uma linha de montagem de carros ou em uma operação de tratamento térmico em que as peças são processadas em lotes, um supermercado é utilizado para sinalizar a necessidade de peças da etapa consecutiva.

De maneira sucinta, o Sistema Puxado exige que exista uma porção de peças nos pontos em que não existe fluxo de peças chamados de “super mercados”. Quando o processo consumidor desse item quiser coletar algumas peças, essas estarão disponíveis nestes estoques controlados de peças. Assim que o supermercado atinge um determinado nível, relacionado ao tempo que o item leva para ser processado, um sinal, normalmente um cartão, é emitido para que o processo anterior desloque ou produza peças, re-suprindo o supermercado. Esse processo pode acontecer várias vezes em uma cadeia de produção. Ou seja, ao longo da cadeia de processamento de uma empresa, podem existir inúmeros supermercados e etapas de sinalização de necessidade de peças (TARDIN, 2001). Adiante se descreve o funcionamento do Kanban especificamente.

Talvez a maior característica do sistema de planejamento puxado seja a eliminação da necessidade de se programar todas as etapas de produção por onde passam as peças (Pace, 2003). O sistema de sinalizações faz com que os operadores da produção consigam definir qual peça e em que quantidade deve ser produzida. A programação do processo de produção é, assim, colocada apenas em um ponto do processo. As outras etapas ou são realizadas em

fluxo ou puxadas segundo o consumo de produtos de acordo como o nível dos supermercado (TAKAHASHI E NAKAMURA, 1998).

A figura 2.3 demonstra como um sistema puxado funciona. As setas que vão das etapas para os estoques representam as sinalizações de necessidade de materiais que são consumidos pelas etapas (operações).

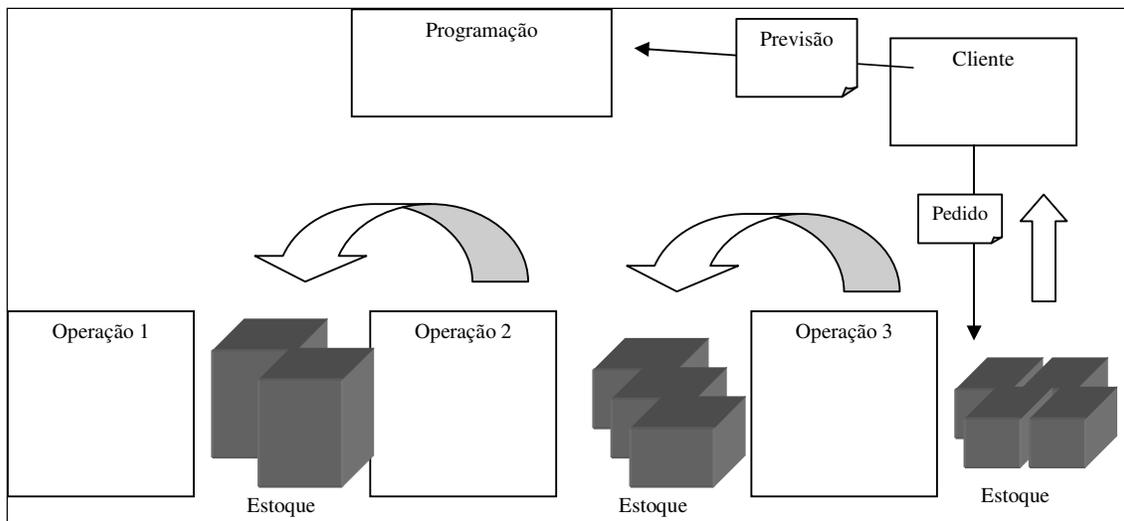


Figura 2.3: Representação do fluxo de informações e matérias em um sistema puxado
Fonte: Adaptado de Tardin (2001)

O modo mais usual de planejamento puxado é o *Kanban*, sistema que utiliza cartões como sinalização de necessidade de produção (LIKER, 1998). Segundo Pace (2003) “Kan” significa cartão e “Ban”, controle. Esse sistema é bastante usual em fábricas de produção discreta.

Os sistemas de produção puxada não se restringem ao Kanban apenas. Existem as variações do sistema Kanban como o Flexible Kanban System (FKS) em que a quantidade de cartões pode variar e o Generalized Kanban System (GKS) explorado por Duri et al (2000). Também existe o sistema *ConWip* (*Constant work in Process*) abordado por Hoop (2008). Tais sistemas não serão abordados aqui.

Tardin (2001) descreve o funcionamento do *kanban* da seguinte forma: Os supermercados de produtos são acompanhados de cartões. Nos dias de hoje, muitas vezes esses cartões são virtuais ou eletrônicos. Quando tais produtos são expedidos, os cartões que os acompanhavam são transportados para um quadro que fica situado na frente de um recurso que transporta ou produz itens. Os cartões são colocados no quadro sob uma régua colorida

em verde, amarelo e vermelho. Os cartões nos quadros sinalizam a falta de produtos no super mercado.

Quando os cartões no quadro atingirem um determinado nível na régua colorida, o operador identifica que é necessário produzir a peça. A figura 2.4 ilustra o funcionamento do Kanban.

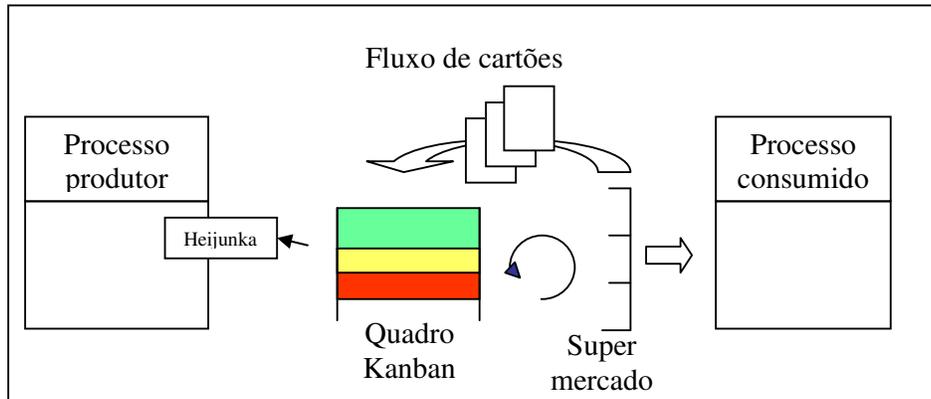


Figura 2.4: Processo de funcionamento do Kanban
Fonte: Adaptado de Shook (2003)

Assim que os cartões nos quadros atingem o nível que sinaliza a necessidade de produzir os itens, eles são colocados em uma régua que contém os horários dos turnos de produção (ou transporte), chamada de *Heijunka* (que significa nivelamento). Estas régua ficam normalmente perto do processo produtor da figura 2.4. Normalmente, o tamanho dos cartões é dimensionado de tal forma que tenha a mesma extensão que o tempo necessário para serem processados na régua de horários *Heijunka*. Assim, à medida que são feitas as sinalizações de produção, os operadores da produção colocam os cartões sobre o *Heijunka* e assim acompanham e planejam a produção “minuto a minuto”. Esse processo possibilita que os gestores da produção tenham uma boa noção dos horários em que as peças serão produzidas. Essa é a forma usual que os gestores do Sistema Puxado utilizam para planejar as capacidades dos recursos e tomar decisões.

O *Heijunka* auxilia os operadores a realizar a priorização dos produtos e a identificarem se o nivelamento da produção está adequado. A figura 2.5 adaptada de Tardin (2001) demonstra a forma usual do *Heijunka*.

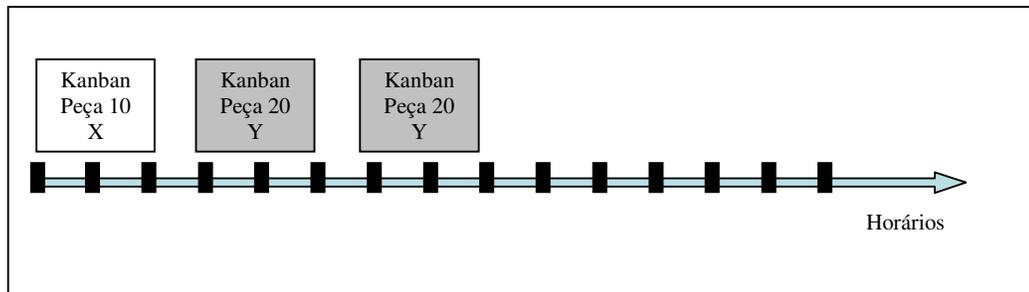


Figura 2.5: Ilustração do funcionamento do Heijunka
Fonte: Adaptado de Tardin (2001 p. 52)

Para criar as condições do funcionamento do sistema Kanban, Molden (1998) define cinco regras para o seu funcionamento:

Regra 1: O Processo seguinte deve retirar peças do supermercado anterior na quantidade adequada e no momento certo.

Isso implica que as peças devem ser acompanhadas de um cartão (autorização) e apenas produzidas quando houver sinalização;

Regra 2: O processo anterior deve produzir produtos para o processo seguinte nas quantidades retiradas por ele.

Ou seja, as quantidades produzidas devem ser iguais aquelas que constam no cartão de sinalização;

Regra 3: Produtos defeituosos não devem passar para o processo seguinte

Uma vez que os estoques no sistema puxado são limitados, deve-se ter certeza que essas peças estejam prontas para serem utilizadas;

Regra 4: O número de cartões deve sempre ser reduzido.

O número de cartões expressa o volume máximo de estoque em processo no sistema. Esse estoque deve sempre ser o menor possível em função do volume e variação da demanda.

Regra 5: Os cartões devem ser utilizados para suportar as pequenas variações de demanda.

Segundo Correa e Giansi (2011), o sistema *Kanban* necessita que a demanda dos itens seja razoavelmente estável, caso contrário os supermercados de peças podem ser “desabastecidos” sem tempo de resposta para a produção.

Tardin (2001) argumenta que o sistema *Kanban* apresenta limitações quando existem muitas peças em diversos níveis nas estruturas dos produtos ou quando os roteiros são muito

complexos com tarefas alternativas. Isso porque os quadros têm tamanho limitado e o sistema *Kanban* é em essência um sistema de planejamento visual. Ao mesmo tempo, a sinalização de produção com cartões não funciona quando existe alternativa de produção em recursos distintos.

Outra limitação do sistema *Kanban* está na dinâmica de priorização do processo de programação. Conforme foi colocado anteriormente, na dinâmica do *Kanban* são os gestores da produção que priorizam a sequência de produção nas máquinas em função dos níveis dos supermercados. Embora o *Heijunka* possibilite uma visão de capacidade e sequenciamento em um curto prazo, muitas vezes tal processo não engloba outras variáveis que influenciam a tomada de decisão como, por exemplo, a priorização dos clientes e a sincronia das peças na fábrica.

Por último, o sistema de cálculo do *Kanban* sugere uma análise de capacidade, porém, esse processo é bastante limitado e normalmente feito em planilhas independentes do sistema *ERP* (CORREA e GIANESI, 2011).

Embora Tardin (2001) e outros autores contraponham o *MPP* e o *Kanban* descritos nas sessões anteriores, esses dois sistemas muitas vezes coexistem nas empresas (PACE, 2003). O sistema *MRP* se adapta melhor a um planejamento de médio prazo enquanto o *Kanban* atende melhor ao processo de programação da produção. Desta forma, o sistema *MRP* e o *Kanban* podem ser usados em diferentes estágios da produção de uma empresa. O primeiro pode ser usado para controlar a compra e aquisição de materiais e o segundo para gerenciar a programação de produção. Ou ainda, o *MRP* pode ser utilizado até uma determinada etapa de produção para depois o processo ser gerenciado por um quadro de *Kanban* e supermercados de reposição (CORREA E GIANESI, 2011).

2.3. OPT - O sistema Tambor Pulmão Corda (*Drum Buffer Rope*)

O sistema Tambor Pulmão Corda consiste no método de planejamento e programação da produção pertencente a metodologia *OPT* (*Optimized Production Technology*); um sistema de administração da produção desenvolvido por pesquisadores israelenses entre os quais Eliyahu Goldratt é principal divulgador (CORREA e GIANESI, 2011). Esse sistema ficou conhecido pelos pesquisadores e gestores da produção como o sistema de produção da Teoria das Restrições, termo que traz o principal alicerce de toda a metodologia que norteia o *OPT*.

A teoria das restrições faz uma distinção entre dois tipos de recursos na fábrica: os recursos gargalos e os recursos não gargalos. Os recursos aqui podem ser entendidos como

qualquer elemento necessário para a produção. Os recursos gargalos são aqueles que ficam ocupados durante todo o tempo de sua disponibilidade em função da demanda de trabalho imposta a ele. Assim, pode-se dizer que um recurso gargalo tem a ocupação de sua disponibilidade igual ou maior do que demanda de trabalho. Já um recurso não gargalo não apresenta ocupação total de sua disponibilidade. (GOLDRATT, 1990).

É importante fazer a diferenciação entre os Recursos Restritivos Críticos (RRC) e os gargalos. Em um processo de produção, sempre há RRCs, mas não necessariamente há gargalos. Isso porque nem sempre a demanda é maior que a capacidade dos recursos, contudo, sempre existe um recurso que tem uma ocupação mais próxima do total de sua disponibilidade em função de sua demanda de trabalho. De qualquer forma, os RRCs e os gargalos cumprem uma função central no sistema da teoria das restrições.

A distinção entre os tipo de recurso é de grande importância para o entendimento do sistema Tambor Pulmão Corda. Os recursos gargalos (ou os RRCs) são os pontos nevrálgicos do planejamento de produção. São eles que definem a cadência e o fluxo da produção nas fábricas. Da mesma forma, os esforços de melhoria para o aumento da capacidade das fábricas devem ser realizados neles. Enquanto isso, os recursos não gargalos têm seu fluxo definido pelos recursos gargalos e não devem ser foco de melhoria para o aumento de capacidade de produção (CORREA e GIANESI, 2011).

O sistema Tambor Pulmão Corda considera que os recursos gargalos ou os RRCs do processo são em essência os elementos que balanceiam o fluxo do sistema como um todo. A utilização dos recursos não gargalos é definida pela estratégia de utilização dos gargalos (GOLDRATT, 1993). Essa abordagem é contraditória à tradicional que dá ênfase no balanceamento do processo pelo uso total da capacidade dos equipamentos.

Os Gargalos não apenas definem o fluxo no sistema Tambor Pulmão Corda como também a capacidade do processo de produção como um todo. Segundo Goldratt (1993) uma hora ganha em um recurso gargalo é uma hora ganha no sistema global. Essa afirmação é um dos mais conhecidos princípios da teoria das restrições e rege os conceitos de definição de lote e estoques no processo.

Tradicionalmente se assume que existe um lote econômico de processamento nas máquinas que é atingido calculando-se o custo unitário de preparação em contraposição ao custo de carregamento do estoque. Contudo quando encaramos que a capacidade de um sistema é definida pelo recurso gargalo então essa afirmação somente é válida para estes (os gargalos). Para todos os outros recursos não gargalos o custo unitário de preparação não seria genericamente aplicável. Assim, segundo Correa e Gianesi (2011 p. 151):

“No OPT, ao contrário do que ocorre na maioria dos sistemas tradicionais, o tamanho do lote de processamento é uma função da situação da fábrica e pode variar de operação para operação. Estes tamanhos de lotes são estabelecidos pela sistemática de cálculo do OPT, que leva em conta os custos de carregar o estoque os custos de preparação, as necessidades de fluxo de determinados itens e os tipos de recursos.”

Os gargalos definem também as estratégias e os tamanhos dos estoques que são dimensionados e colocados em locais tais que consigam proteger os gargalos de flutuações da produção e da demanda. Assim, é comum em um sistema Tambor Pulmão Corda alocar um estoque antes dos RRCs com o objetivo de protegê-los de variações no processo que possam fazer com que fluxo de produtos seja interrompido.

O sistema Tambor Pulmão Corda estabelece que a programação de atividades e a capacidade produtiva devem ser consideradas simultaneamente e não em sequência (GOLDRATT, 1990). Essa peculiaridade sugere uma diferença importante entre este e os sistemas MRPII e Sistema Puxado em que essas duas atividades são feitas separadamente.

Correa e Giansi (2011) descrevem o sistema Tambor Pulmão Corda de acordo com três regras básicas:

Primeira regra: O RRC (que pode ser um gargalo) é carregado de acordo com o total de demanda de trabalho a ser nele processado. A sequência dos trabalhos no RRC é realizada simultaneamente ao carregamento em função das datas de entrega e a prioridade entre as atividades.

Segunda regra: O RRC deve ser protegido contra possíveis variações no processo que podem acontecer em processos anteriores. Este estoque leva o nome de “estoque por tempo de segurança” e é dimensionado segundo a variação dos processos nos diferentes períodos. Podem existir estoques secundários que garantem que peças de linhas paralelas ao RRC fiquem em sincronia com o processo como um todo.

Terceira regra: O RRC deve ser usado para controlar os recursos não gargalos do processo. Os recursos que estão à frente do RRC no roteiro de produção processam os produtos na sequência em que eles foram processados no RRC. Já os recursos anteriores ao RRC, ou que estão paralelos a esse em outras linhas, devem processar os produtos também de acordo com a sequência estabelecida pelo RRC e nas quantidades demandadas pelo pulmão protetor.

As três regras dão origem ao nome do sistema de produção Tambor Pulmão Corda. O RRC representa o tambor que dita a cadência e a capacidade do sistema como um todo. O

Pulmão é representado pelo estoque controlado de peças que protege o recurso crítico. E por fim, a corda representa a sincronia entre o RCC, os recursos não gargalos e os estoques ao longo do processo.

A figura 2.6 representa o esquema Tambor, Pulmão, Corda. Como é possível visualizar, o RCC é protegido por um pulmão que por sua vez apresenta sincronia com os recursos não gargalos como se as peças estivessem amarradas por uma corda inelástica.

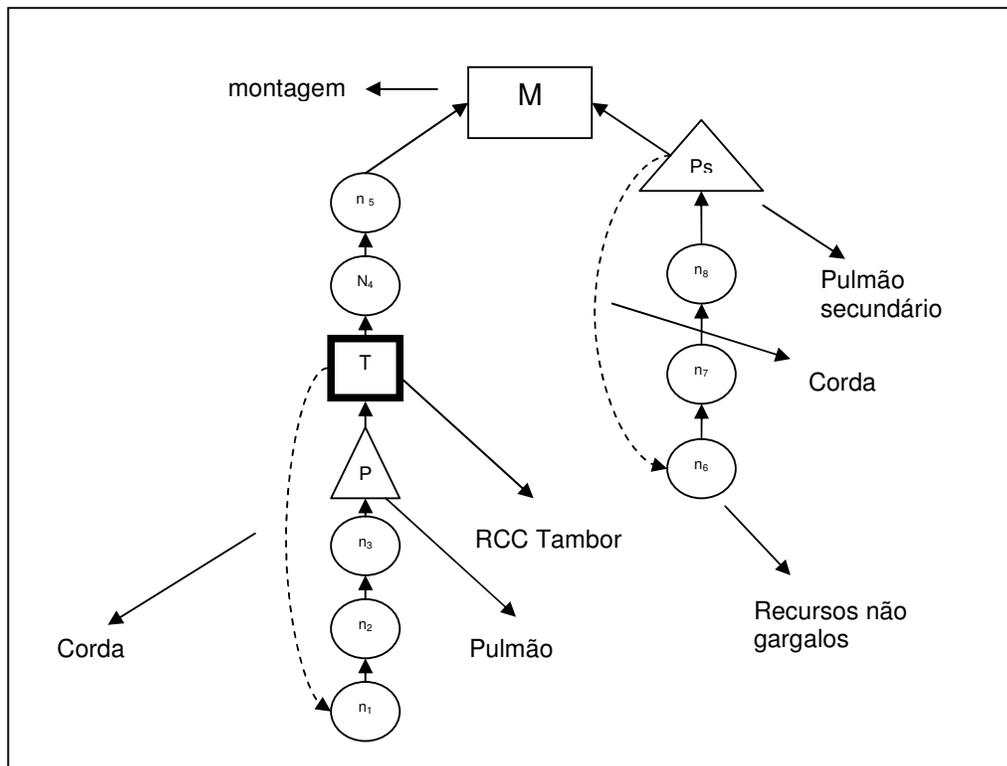


Figura 2.6: Esquema de funcionamento do sistema Tambor, Pulmão, Corda
Fonte: Adaptado de Correa e Giansi (2011 p. 157)

Segundo Correa e Giansi (2011) existem dois métodos de se realizar o planejamento e programação do sistema Tambor, Pulmão, Corda. Estes métodos funcionam em paralelo e de forma integrada.

O módulo *OPT* é um algoritmo de programação finita para frente (*forward scheduling*) que considera as restrições dos recursos, os tempos dos processos e as prioridades para realizar o sequenciamento e definir o tamanho dos lotes de produção para os RCCs. Ao mesmo tempo, o módulo *SERVE* de capacidade infinita para trás (*backward scheduling*) usa como ponto as datas de entrega nos Pulmões da produção que são determinados, em última análise, pelos RCCs.

A figura 2.7 representa o funcionamento integrado dos dois módulos.

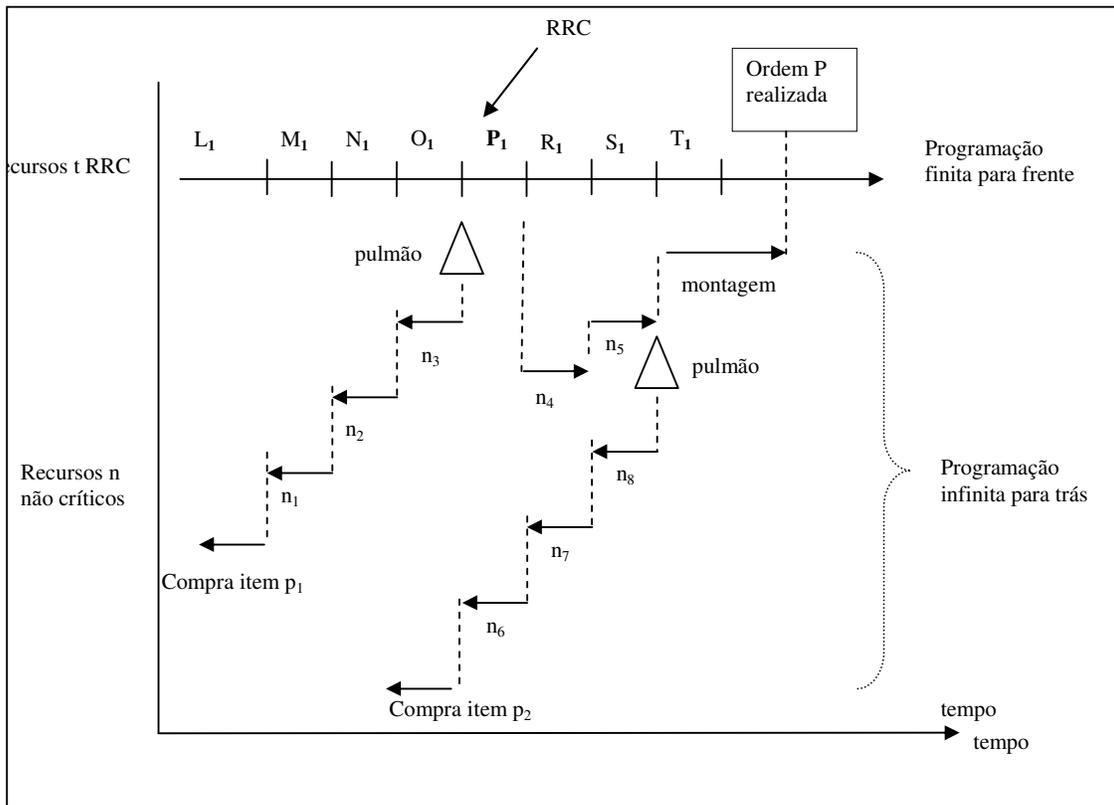


Figura 2.7: Sistema de programação de Tambor, Pulmão, Corda
Fonte: Adaptado de Correa e Giansesi (2011 p. 159)

Desta forma, o sistema gerencia os recursos críticos com um sistema de capacidade finita e os outros recursos não críticos através de um sistema de capacidade finita (CORREA e GIANESI, 2011).

O método de gerenciamento da produção do *OPT* pode ser considerado como um sistema de capacidade finita semelhante aos *APSs* comerciais de hoje. As regras que estabelecem suas prioridades são fundamentadas na teoria das restrições e nos conceitos abordados nesta sessão; tal como as regras dos *APSs* são determinadas pelos fornecedores de softwares e pelas empresas que os implementam. Assim, a tentativa de comparação entre *OPT* e os *APSs* não traz contribuições relevantes para a melhor compreensão dos sistemas de suporte a produção. Ressalva feita ao fato de que o *OPT* não utiliza algoritmos de otimização como alguns *APSs* o fazem.

O gerenciamento do sistema Tambor, Pulmão, Corda demanda bons conhecimentos das pessoas que o controlam. Identificar recursos restritivos não é um processo elementar,

bem como indicar parâmetros que auxiliem a definição de estoques estratégicos na produção. Ademais, segundo Correa e Giansesi (2011), é necessário que todos os princípios elencados nesta sessão sejam de domínio dos gestores do sistema, o que não é tão simples como se poderia desejar.

O sistema demanda um software “proprietário” pertencente a uma empresa de consultoria israelita que segundo Correa e Giansesi (2011) não é de fácil utilização. Esse software centraliza as decisões quanto ao planejamento de produção, o que não favorece o comprometimento da força de trabalho com os objetivos do planejamento.

Embora a aplicação íntegra do sistema Tambor Pulmão Corda não seja tão frequente, a teoria das restrições, que norteia os conceitos do *OPT* é amplamente explorada pelos gestores de produção seja qual for sistema de planejamento utilizado. Nos livros textos que auxiliam os gestores a implementar o Sistemas Puxados como Aprendendo a Enxergar (2003) de Shook, J. e Rother, M., ou O kanban na Prática (2003) de Pace, J.H., é comum a definição dos recursos restritivos para a implementação dos supermercados. De mesma forma, o conhecimento dos recursos críticos ajuda na definição das regras do *RCCP* nas implementações dos *MRPs* (Tubino, 2007).

2.4. APSs (Sistema de Capacidade Finita)

Segundo Pedrosa e Corrêa (1996), os Sistemas de Capacidade Finita ou *APSs* foram desenvolvidos em função das limitações do *MRP II* e *Kanban* em gerenciar a capacidade e planos detalhados de produção, da melhoria no processamento computacional das últimas duas décadas e pela busca de competitividade das empresas privadas. Os Sistemas de Capacidade Finita foram introduzidos na manufatura ao mesmo tempo em que o *MRPII* estava sendo implementado nas empresas. Esses sistemas vieram suprir as deficiências que os *MRPs* tinham em gerenciar recursos com capacidade finita e auxiliar as empresas a dimensionar melhor seus lotes e a sequência de produção (TUBINO, 2007).

Os Sistemas de Capacidade Finita, porém, não são utilizados exclusivamente no controle da produção detalhada, embora essa seja uma das suas principais aplicações em ambientes de produção. Os métodos para solucionar problemas de capacidade finita, atualmente chamados de *APSs* são hoje utilizados em vários níveis de planejamento e organização da produção e também em ambientes de transportes e gestão de demanda (STADTLER e KILGER, 2005).

Os *APSs* podem funcionar com algoritmos que não sejam de otimização linear. Tais sistemas podem funcionar com heurística, meta heurística, inteligência artificial ou outros algoritmos específicos (GIACON 2010). A otimização linear é utilizada em alguns *APSs* quando se trata de problemas em que o tempo de processamento não possa inibir o uso do sistema (TEMPELMEIER, 2003). Outrossim, em função do avanço no processamento computacional e também da melhoria dos algoritmos e sistemas de otimização, o desenvolvimento de programas de otimização linear tem sido um dos focos do desenvolvimento dos *APSs*.

Os *APSs* podem ser complementares aos *ERPs* ou podem pertencer a estes em alguns casos (STADTLER, 2005). Atualmente existem softwares *APSs* comerciais que são aplicados em diversas indústrias como *APO (Advanced Planning and Optimizing)* do *ERP* alemão *SAP*. Ou o *ASCP (Advanced Supply Chain Planning)* do *ERP* americano *ORACLE* que possibilita o planejamento otimizado de recursos de produção e transporte. Existem ainda o *I2* e o *Manugistics*, ambos comprados recentemente pela *JDA Software Group*.

Como os *APSs* são relativamente novos nas empresas, existe bastante controvérsia quanto a sua definição ou ainda em como esses sistemas se contextualizam dentro de um ambiente de organização e planejamento de produção (TEMPELMEIER, 2003).

Neste sentido, Stadler e Kilger (2005) definem os *APSs* como ferramentas de gerenciamento de planejamento que têm como objetivo suportar a decisão dos gestores em ambientes complexos. Estes autores assumem que os *APSs* são uma infraestrutura de módulos que funciona em diversos elos e etapas da cadeia de suprimento das empresas como um todo. Ou seja, são sistemas de suporte a operação que funcionam nos níveis estratégico, tático e operacional nas empresas.

Tubino (2007), por sua vez, impõe que os *APSs* são sistemas de gestão da capacidade finita no ambiente de programação detalhada da produção. Segundo o autor:

“Discutidas as regras de sequenciamento mais usuais, cabe agora apresentar o ambiente onde essas regras são normalmente implementadas para gerar a seqüência de ordens de produção conhecidos atualmente como sistemas *APSs*...” (Tubino, 2007 p.119)

Desta forma, Os *APSs* podem, dependendo da configuração e estratégia de produção das empresas, auxiliar a tomada de decisões dos gestores de produção apenas no processo de sequenciamento e planejamento da produção no curto prazo, ficando o processo de gestão de compras e suprimentos a cargo de um *MRP*. Tal aplicação pode ser vista em Kjeldsdotter,

(2005), onde é analisada uma aplicação de um APS apenas no processo de sequenciamento da produção.

Em ambientes de planejamento e organização da produção, os sistemas de capacidade finita podem ter uma aplicação bastante voltada para o sequenciamento da produção no chão das fábricas ou de forma semelhante, nas funções da programação detalhada da produção ou *Shop Floor Control* (TUBINO, 2007).

Segundo Hoop (2008 p.145):

“Uma vez que uma ordem de compra ou produção é emitida, algum controle deve ser feito para se ter certeza que foram realizadas com as quantidades de recursos e na especificação correta (...). Se a tarefa for relacionada a manufatura interna esse controle é feito nas funções do Sistema de Capacidade Finita.”

Neste sentido, o APS possui duas funções: Envio de tarefas (*Job Dispatching*) e Controle de Entradas e Saídas (*input/output control*).

O Envio de tarefas se refere às regras sistêmicas para se estabelecer filas nos recursos de produção que atendam às datas de entregas dos produtos e mantenham as máquinas ocupadas (HOOP, 2008). Existem várias regras de envio de tarefas que se pode estabelecer em um sistema SCF. Algumas delas como a SPT (*short process time*), em que as tarefas são estabelecidas segundo o menor tempo de processamento, minimizam o tempo de atravessamento (*lead time*) de produção e oneram as entrega dos produtos com altos tempos de processamento. Enquanto outros como o EDD (*Early Due Date*) priorizam as datas de entrega enquanto não consideram a ocupação das máquinas.

Segundo Hoop (2008), O controle de entradas e saídas é feito com o objetivo de manter o tempo de processamento (*lead time*) de processamento sob controle. A abordagem identifica se a taxa de liberação de ordens de produção está muito alta ou muito baixa para cada um dos centros e produção e consiste na realização das 4 etapas:

1. Monitorar o estoque em processo em cada um dos centros de produção;
2. Se o estoque em processo em cada um dos centros estiver acima de um determinado nível, então é necessário reduzir a liberação de ordens;
3. Se o estoque em processo em cada um dos centros estiver abaixo de um determinado nível, então é necessário aumentar a liberação de ordens;

4. Se o estoque em processo estiver entre o nível máximo e mínimo estabelecidos, então o taxa de liberação está adequada.

O controle de entradas e saídas é uma sistemática para se identificar o excesso de ordens em função da capacidade que, em conjunto com a função de liberação de tarefas, faz com que se tenha o controle de capacidade em um nível detalhado da operação.

Uma configuração diferente da Kjeldsdotter (2005) e Tubino (2007) dos sistemas APSs nas empresas pode ser encontrada em Haddad e Carvalho (2004), onde é colocada a sugestão de aplicação desse sistema na forma de um algoritmos de otimização linear nos *MRPs*. É como se o *MRPs* pudessem alocar os recursos atentando às restrições de capacidade e demanda, tentando com o objetivo de melhoria da performance da operação. Neste sentido, os APSs atuariam de uma forma mais tática no processo de planejamento da produção sendo usados como uma alternativa aos *MRPs*. A figura 2.8 sugerida por Haddad e Carvalho (2004) ilustra a forma com que os autores enxergam a atuação dos APSs:

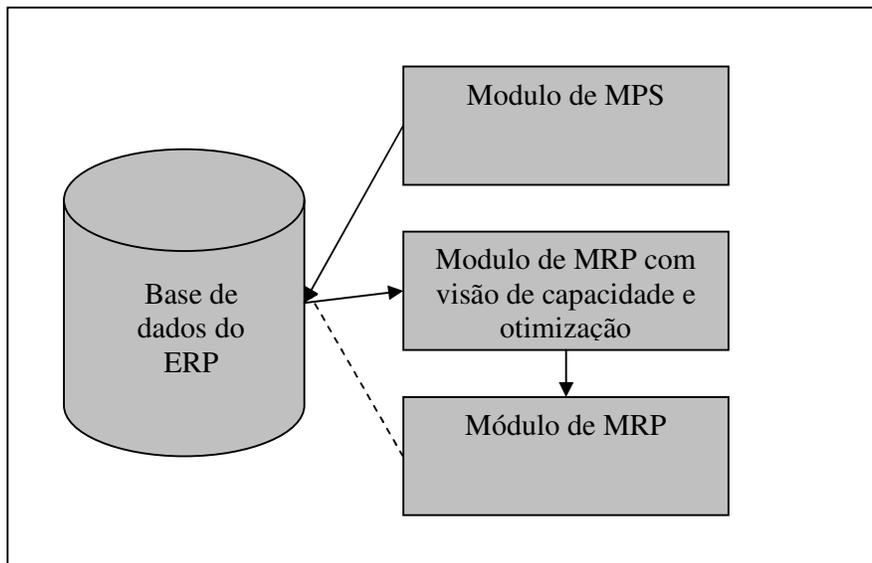


Figura 2.8: Configuração da otimização linear em substituição ao funcionamento do ERP

Fonte: Adaptado de Haddad e Carvalho (2003 p. 23)

A APICS (2007) define os APSs como sendo “qualquer programa computacional que usa lógica ou algoritmos matemáticos avançados para realizar otimização ou simulação em ambientes de capacidade finita”. Ao avaliar as definições do termo, pode-se concluir que não existe uma definição clara para esses sistemas. Eles são determinados por um conjuntos de

módulos integrados na cadeia de suprimentos por uns; e por qualquer aplicação de modelos matemáticos avançados para suportar as decisões em ambientes de capacidade finita por outros.

Além da falta de consenso quanto à definição do que são os APSs e em que nível do planejamento e organização da produção eles podem atuar, também existe na literatura uma discussão acerca de qual seria a melhor estrutura hierárquica para demonstrar o funcionamento dos sistemas de mercado (TEMPLEMEIER, 2005).

Stadler e Kilger (2005) argumentam que embora as aplicações e construções dos APSs sejam diferentes em cada empresa ou para cada provedor da solução, sua arquitetura sempre se encaixa em uma visão comum de duas dimensões: a cadeia de processos de uma empresa e o horizonte de planejamento. A figura 2.9, abaixo, mostra essa arquitetura de módulos.

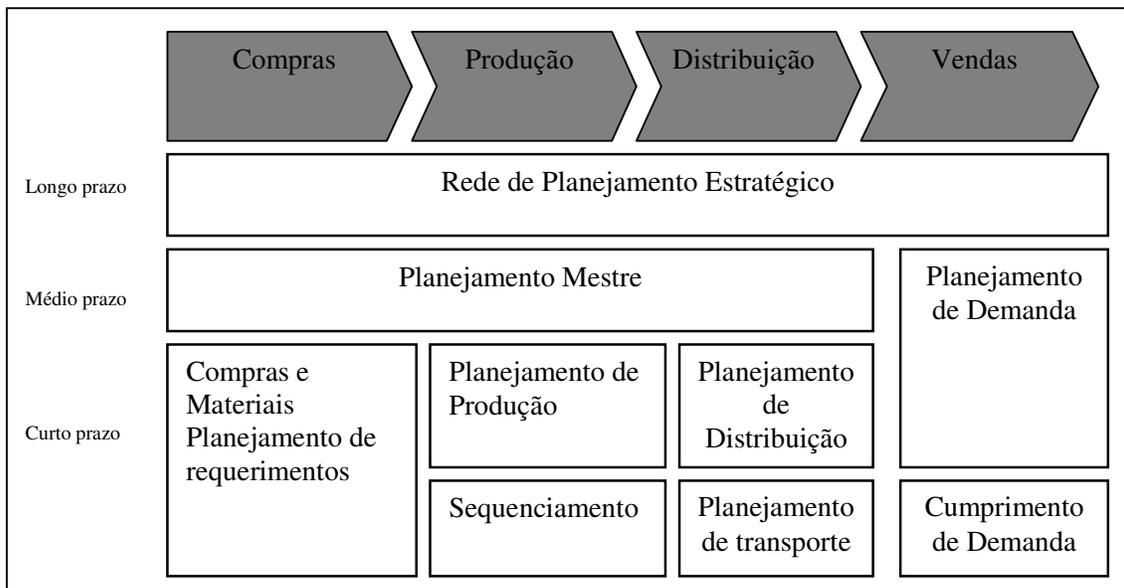


Figura 2.9: Estrutura de módulos em que os APSs são operados
Fonte: Adaptado de Stadler e Kilger (2005)

Os APSs comerciais muitas vezes se utilizam da figura 2.9 citada para demonstrar como esses sistemas funcionam na prática. Contudo, segundo Tempelmeier (2005), citando Drexel et al. (1994), as arquiteturas atuais não atendem as complexidades específicas dos diferentes tipos de manufatura nos níveis baixos do planejamento e programação da produção. Segundo o autor:

“O problema tratado por diferentes módulos de planejamento não diferenciam claramente o tipo de sistema de produção e logística. Entretanto, é sabido que os problemas enfrentados por uma empresa de metais com pequenos lotes de produção em um ambiente de *Job flow* são diferentes dos problemas de planejamento de fabricante que utiliza um *lay out* fixo e um processo de produção orientado para projetos. Essa diferença deve ser refletida no modelo de planejamento usado para suportar o processo de decisão” (TEMPELMEIER 2005, p.4).

Nesse sentido, uma empresa com processo de fluxo de trabalho (*Job flow*) em que os tempos de troca (*set up*) são relevantes para o planejamento de produção, o conhecido problema do dimensionamento dinâmico de lote em vários níveis com visão de capacidade (*multi level capacitated dynamic lot sizing problem*) deve ser tratado. Da mesma forma, em um ambiente de linha de produção com vários estágios em que vários produtos são produzidos com tempos de *set up* relevantes, o problema do sequenciamento de lotes econômicos (*economic lot scheduling problem*) deve ser utilizado (TEMPELMEIER, 2005). Uma solução única de programação da produção não atende as especificidades dos diferentes tipos e ambientes de produção.

Assim, Drexl (1994) propõe uma estrutura teórica diferente para as aplicações dos APSs.

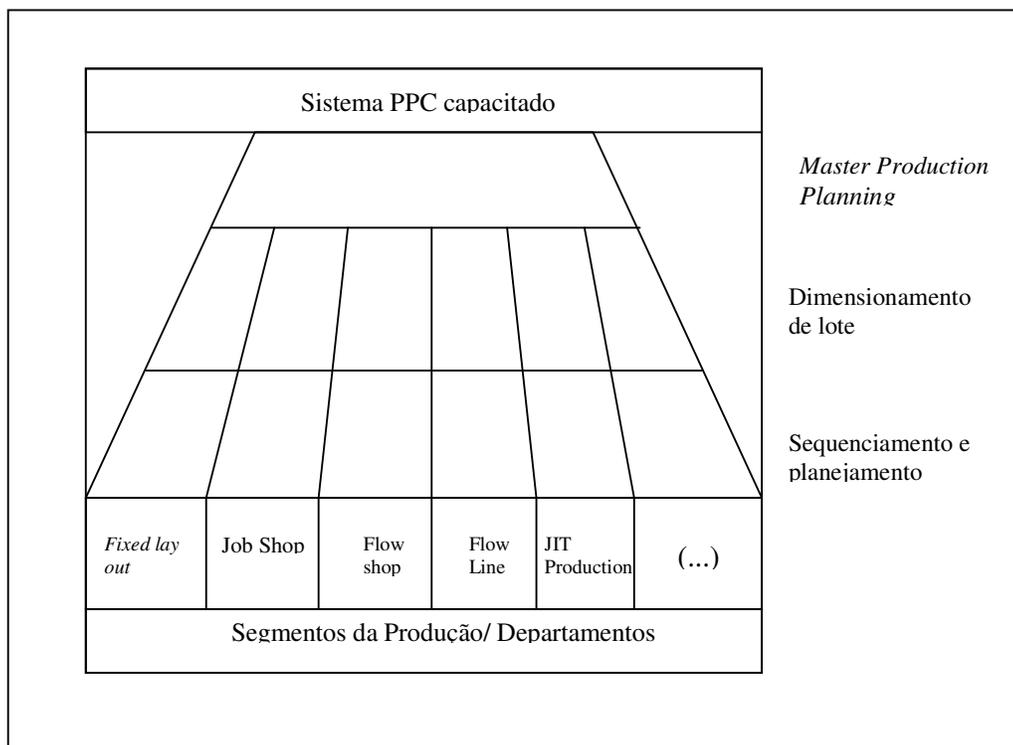


Figura 2.10: Estrutura de funcionamento dos APSs
Fonte: Adaptado de Drexl (1994) citado por Tempelmeier (2005)

A figura 2.10 enfatiza a especialização que os *APSs* devem adotar para atender as demandas dos diferentes tipos de manufatura. Ou seja, os *APSs* têm diferentes abordagens que são determinadas pelo tipo de produção em que serão aplicados. É importante notar que em um nível agregado no médio/longo prazo, os *APSs* não precisam ser específicos para cada tipo de produção.

De uma forma geral, independentemente da melhor estrutura que o traduz ou do seu real significado, os *APSs* apresentam sistemática de gerenciamento que podem ser visualizados em tabelas ou digramas de *Gantt* (GIROTTI e MESQUITA, 2011). Os gráficos de *Gantt* mostram a interdependência entre as atividades ao longo do processo de produção, os tempos de fila e os atrasos dos produtos em função de uma determinada escolha de produção. Essa visualização demonstra um avanço com relação ao *MRP* ou Sistema Puxado. Enquanto o primeiro não considera os tempos de fila dos recursos, o segundo apresenta limitações em estabelecer a inter-relação entre as tarefas ou produtos. Se comparado com o *OPT*, é possível dizer que os *APSs* podem ser bem mais amigáveis; feita a crítica de que o sistema proprietário do *OPT* não é muito amigável.

Normalmente os *APSs* funcionam segundo uma lógica de programação para frente (*forward planning*), como acontece com o *OPT*. Esse método traz maior acuracidade do que os programas de *lead time* fixos dos *MRPs* e apresentam melhor visão das datas de entrega do que as réguas de tempo de *Heijunka* utilizadas para se controlar o Sistema Puxado.

O método de programação para frente em conjunto com a capacidade de realizar as interdependências entre as etapas de produção do *APS* possibilitam que o sistema seja utilizado com uma plataforma de simulação (Tubino, 2008). Os planejadores podem simular o resultado de decisões aumentando as chances de realizar as escolhas certas. Embora seja possível realizar simulações com o sistema *OPT*, este aspecto é um ponto positivo se comparado com o Sistema Puxado ou o *MRP*.

Ao contrário dos *MRPs* e os Sistemas, os *APSs* podem realizar escolhas eficientes ou otimizadas (PEDROSO e CORREA, 1996). Como colocamos em sessões anteriores, os *APSs* não necessariamente utilizam algoritmos de otimização para realizar as decisões no processo de produção. Mesmo que o sistema utilize qualquer metodologia, os *APSs* conseguem seguir regras e parâmetros previamente estabelecidos com o objetivo de alocar de forma eficiente os recursos no planejamento e controle de produção (GIACON, 2010). Essa funcionalidade é particularmente apreciada em ambientes complexos em que existe um número muito grande de variáveis e restrições a serem consideradas no processo de tomada de decisão.

Ao analisar as diversas definições sobre os *APSs* é possível concluir que se trata de um sistema de capacidade finita que inicialmente foi amplamente utilizado na gestão operacional da produção. Com o avanço dos algoritmos e na tecnologia de processamento de dados, os *APSs* passaram a ser aplicados em processos táticos e estratégicos do planejamento da produção.

Embora existam controvérsias quanto à definição do termo, seus níveis de aplicações ou qual seria a melhor estrutura hierárquica para representá-los, os *APSs* têm cumprido um importante papel no planejamento e programação das empresas (STADTLER e KILGER, 2005). Os *APSs*, sejam eles comerciais ou específicos, têm realizado eficientemente seu papel no suporte ao planejamento de produção, desempenhando funções consideradas deficientes nos outros sistemas de gerenciamento. Ademais, em muitos casos os *APSs* utilizam as mesmas informações da produção que os MRPs ou OPT como lista de materiais, roteiros de produção, ordens de entrega, o que facilita sua implementação e uso (PLENERT e KIRCHEMIER, 2000)

Nas sessões seguintes aprofundaremos a discussão sobre os algoritmos de otimização utilizados no estudo de caso sobre a implementação do *APS* do estudo de caso. Esta imersão nas equações e modelos de otimização são necessários para um bom entendimento dos métodos utilizados no modelo de otimização do estudo de caso

2.5. Otimização linear

Existem vários tipos de modelos matemáticos dentro da Pesquisa Operacional (HILLIER e LIEBERMAN, 2010). Este trabalho se concentra nos modelos de otimização lineares apenas. Esta sessão tem o objetivo de explorar os métodos e algoritmos relacionados a otimização linear. Os conceitos que definem o termo são inicialmente introduzidos para em seguida explorarmos os algoritmos usualmente utilizados para resolução dos problemas lineares. Os métodos de redução do tempo de processamento ocupam os últimos parágrafos do capítulo.

Segundo Taha (2008, p. 3):

“Em problemas de otimização, busca-se maximizar ou minimizar quantidades específicas com um certo objetivo, que depende de um número finito de variáveis de entrada. Essas variáveis podem ser independentes uma das outras ou podem ser relacionadas por meio de uma ou mais restrições”.

Um modelo de otimização é um sistema de equações no qual as variáveis, a função objetivo e as restrições são expressas com funções matemáticas relacionadas funcionalmente (HILLIER, LIEBERMAN, 2010).

Assim, pode-se considerar que um modelo de otimização linear é definido por variáveis que estão linearmente relacionadas a uma função objetivo que é limitada por restrições também lineares. As variáveis, respeitando as restrições, mudam de valor até um ponto em que a função objetivo atinge um valor ótimo.

A forma matemática mais comum de se expressar um problema de otimização linear é:

$$\begin{array}{l}
 \text{Otimizar } z = f(c_1x_1, c_2x_2, c_3x_3, \dots, a_nx_n) \\
 \\
 \text{Sujeito a: } \left. \begin{array}{l} g_1(x_1, x_2, x_3, \dots, x_n) \\ g_2(x_1, x_2, x_3, \dots, x_n) \\ (\dots) \\ g_m(x_1, x_2, x_3, \dots, x_n) \end{array} \right\} \begin{array}{l} < \\ > \\ = \end{array} \left\{ \begin{array}{l} b_1 \\ b_2 \\ (\dots) \\ b_n \end{array} \right.
 \end{array}$$

Onde $f(x_1, x_2, x_3, \dots, x_n)$ é a função objetivo do modelo. Os coeficientes c_1, c_2, \dots, c_n são chamados coeficientes do vetor custo C, e desempenham um papel importante na definição dos modelos de otimização lineares. Como se pode presumir, o vetor custo da função objetivo direciona o valor das variáveis na função objetivo. Já $x_1, x_2, x_3, \dots, x_n$ são as variáveis do modelo (ARENALES, 2007, TAHA, 2008).

As variáveis do problema são limitadas pelas funções $g_1(x_1, x_2, x_3, \dots, x_n)$, $g_2(x_1, x_2, x_3, \dots, x_n)$, ..., $g_m(x_1, x_2, x_3, \dots, x_n)$ que descrevem as restrições (ARENALES, 2007, TAHA, 2008).

Quando tratamos de otimização linear, os algoritmos de otimização (apresentados adiante) garantem que o resultado dessas equações atinja a melhor resolução possível dentro das restrições que foram impostas, desde que as condições de linearidade sejam atendidas. Entretanto, é importante destacar que não é incomum que os valores das variáveis levem para um mesmo resultado ótimo. Ou seja, em problemas de otimização linear é comum que existam alguns resultados com diferentes valores para as variáveis que fazem com que a função objetivo tenha o mesmo valor (HILLIER, LIEBERMAN, 2007).

Um problema de otimização é considerado linear quando a função objetivo e as funções de restrições são lineares, ou seja, se:

$$f(x_1, x_2, x_3, \dots, x_n) = c_1x_1 + c_2x_2 + \dots + c_nx_n \text{ e};$$

$$g_i(x_1, x_2, x_3, \dots, x_n) = a_1x_1 + a_2x_2 + \dots + a_nx_n$$

Onde c_n e a_n são constantes conhecidas.

Segundo Arenales (2007) Existem três condições ou hipóteses elementares para que um problema matemático seja considerado um problema linear.

1. Hipótese da aditividade: Essa hipótese pressupõe que se adicionarmos um composto que contenha $0,2kg$ ($0,2 b_1$) de um ingrediente b_1 a outro composto que contenha $0,1kg$ do mesmo ingrediente b_1 ($0,1b_1$), então o produto resultante terá $0,3 kg$ do ingrediente b_1 ($0,3b_1$);
2. Hipótese da proporcionalidade: Se $\frac{1}{2} c$ é a quantidade do elemento c em uma unidade de produto D , então $\frac{1}{2} D$ contem $\frac{1}{4} C$.
3. Hipótese de fracionamento: Essa hipótese pressupõe que os valores inteiros não podem ser fracionados e somente é válida para problemas de otimização linear inteira. Ou seja, Se x_1 tem que ser inteiro, então, $0,99x_1$ é diferente de x_1 .

Dentro dos problemas de otimização linear, existem ainda os modelos de otimização lineares inteiros, que possuem as restrições adicionais de que as variáveis do problema devem ser inteiras. Ou seja:

$$x_1, x_2, x_3, \dots, x_n = (1, 2, 3, \dots, m)$$

Onde m é sempre um número inteiro.

Os problemas de otimização inteira são hoje em dia os problemas mais complexos de se solucionar em função do grande número de interações necessárias para a sua resolução (ARENALISM 2007)

2.5.1. Forma gráfica da otimização linear

O exame da forma gráfica da otimização linear deixa claro o conceito que norteia o funcionamento do algoritmo usado no modelo estudado nesta dissertação. Para facilitar o entendimento, usamos como referência o exemplo da representação gráfica de Taha, (2008). As equações e gráficos aqui descritos tem apenas duas dimensões

Dada a equação de otimização:

$$\text{Otimizar } z = a_1x_1 + a_2x_2,$$

$$\begin{aligned} \text{Sujeito a: } R_1: & b_{11}x_1 + b_{12}x_2 \leq c_1 \\ R_2: & b_{21}x_1 \leq c_2 \\ R_3: & b_{33}x_2 \leq c_3 \end{aligned}$$

As equações expressam o gráfico abaixo:

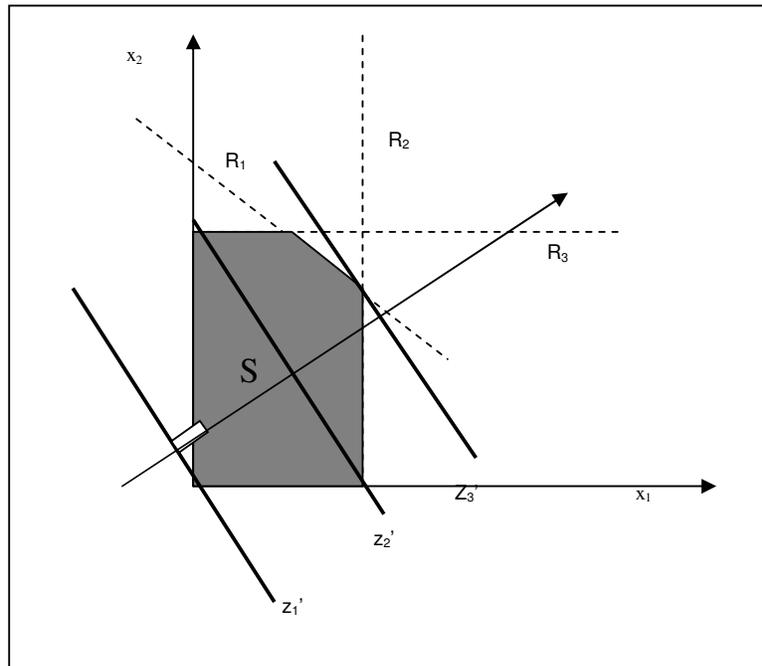


Figura 2.11: Resolução gráfica de otimização linear
Fonte: Adaptado de Taha (2007)

As retas R_1 , R_2 , R_3 representam as equações de restrições do modelo e definem o espaço S (ou área S), região factível, onde todas as respostas do problema estão contidas. (TAHA, 2008). Como este trabalho trata somente de variáveis e restrições positivas, a figura é também restringida pelas restrições de não negatividade representadas pelos quadrantes 2, 3 e 4 do plano cartesiano da figura 2.11. Esse problema possui apenas duas variáveis, o que possibilita que seja representado graficamente em um plano, entretanto muitos modelos possuem muito mais variáveis, o que dificulta a sua visualização gráfica (ARENALES, 2007).

A função objetivo do problema: $z = a_1x_1 + a_2x_2$, é representada no plano por uma curva de nível a qual as retas z_1 , z_2 , z_3 pertencem. A intersecção desta curva de nível com o extremo

da região factível define o ponto de x_1^* , x_2^* que representa solução ótima do problema tal que: $z(x_1^*, x_2^*) \geq z(x_1, x_2)$ para qualquer ponto de S. De forma sucinta, para se obter x_1^* , x_2^* , a reta que representa a função objetivo: $z = a_1x_1 + a_2x_2$, se desloca perpendicularmente e no sentido em que a função z cresce para encontra o extremo do espaço S (ARENALES, 2007; TAHA, 2008).

O ponto ótimo x_1^* , x_2^* é representado na figura por um ponto extremo ou um vértice que representa a intersecção das retas da função objetivo e restrições do problema. De acordo com Hillier e Lieberman (2010), por hipótese, a solução de um problema sempre vai pertencer à equação de suas restrições, pois são essas que impedem que as variáveis que estão diretamente colocadas na função objetivo deixem de apresentar valores como $+\infty$ e $-\infty$. Seguindo esse raciocínio e observando a figura 2, pode-se intuir que se um software pudesse sair de um vértice para outro melhor finitas vezes até encontrar um ponto extremo ótimo, poderíamos solucionar problemas de otimização linear (TAHA, 2008)

O método Simplex, tema da próxima parte do trabalho é um método que possibilita a trajetória entre os vértices da área de factibilidade de forma a conseguir atingir a otimização dos problemas (HILLIER, LIEBERMAN, 2010).

2.5.2. Método Simplex

Uma pessoa com experiência na implementação de soluções relacionadas a modelos matemáticos de otimização poderia concluir ao se deparar com esta sessão sobre o método Simplex, que é indevido o esforço em escrever uma dissertação sobre modelagem matemática quando já não são incomuns os softwares comerciais baseados em algoritmos de otimização linear que realizam o processo sem qualquer trabalho de programação. Entretanto, nem sempre nos deparamos com problemas que são facilmente transformados em modelos matemáticos pré-existentes ou customizáveis, tampouco com softwares com alta flexibilidade de se ajustar a qualquer problema real. Outrossim, mesmo com o uso dos softwares comerciais, a falta de conhecimento nos conceitos dos algoritmos e sua abrangência podem inibir seu uso correto ou até mesmo fazer com que seu uso seja realizado de forma equivocada. Embora não seja necessário entender a eletrônica de uma máquina de calcular para se fazer bom uso dela, nem ser conhecedor da teoria dos números, qual seria a importância de uma máquina de calcular para uma pessoa sem a noção de operações aritméticas (ARENALES, 2007)?

Na pesquisa que foi feita sobre métodos de otimização durante este trabalho, identificou-se que são raras as dissertações de mestrado em engenharia que se aventuram em explicar o modelo simplex. Porém, pensamos ser importante explorar o funcionamento desse método para explicar e detalhar melhor o modelo aqui estudado. Ademais, o conhecimento do funcionamento dos algoritmos e métodos de resolução foi de extrema importância no processo de modelagem do problema de planejamento e programação de produção do estudo de caso, e por isso a decisão de expor os princípios do funcionamento dos algoritmos nesse trabalho.

Em um processo de modelagem matemática e mais ainda na operacionalização de um modelo de otimização para a tomada de decisões reais, o processo de redução do tempo de processamento é fundamental para que o tempo de processamento da tomada de decisão não extrapole o tempo de resposta exigida pelo processo que se tenta resolver. Somente com o conhecimento profundo do funcionamento do algoritmo se conseguiu, no estudo de caso exposto, reduzir o tempo de processamento fazendo com que o tempo de resposta do processo fosse respeitado.

Em Hillier e Lieberman (2010), Taha (2008) e Arenales (2007) o método simplex antes de ser desdobrado é examinado graficamente. O mesmo processo será adotado aqui. Depois de demonstrar graficamente o método simplex, as principais equações e transformações matriciais referentes ao método serão analisadas.

Conforme é demonstrado na figura 3.2, o método Simplex inicia sua interação com os vértices da região factível no ponto $(x_1, x_2) = (0,0)$, ou seja, na origem. O algoritmo então percorre os pontos extremos realizando testes sobre todas as soluções possíveis e escolhendo sempre um vértice que aumente o valor da função objetivo. Esses testes consistem em analisar o valor da função objetivo e avaliar se seu valor é maior do que um valor já analisado anteriormente. Quando não existem mais vértices que aumentem o valor da função o objetivo o algoritmo define a solução ótima.

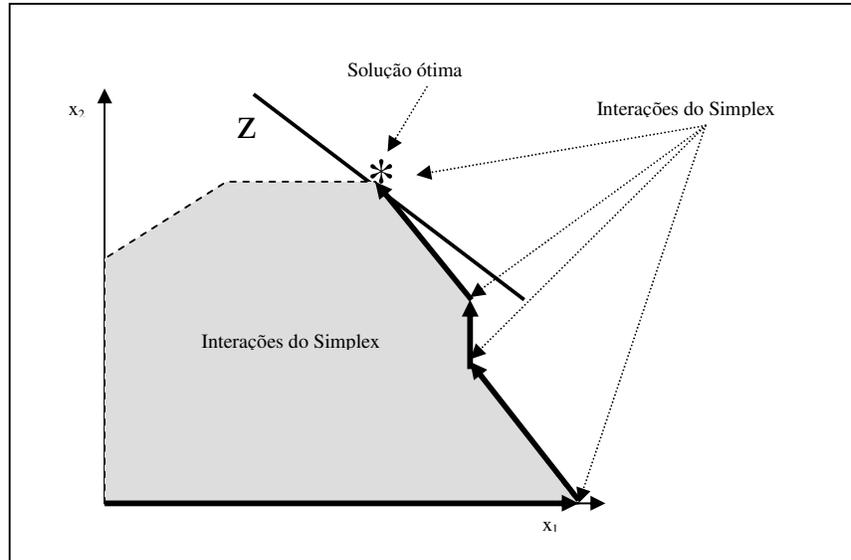


Figura 2.12. Resolução gráfica do modelo Simplex
Fonte: Adaptado de Hillier e Lieberman (2010)

A forma algébrica do método simplex será abordada aqui de forma heterodoxa atrás de matrizes $m \times n$ que serão representadas por tabelas. A abordagem usa como referência Arenales, (2007) e Hillier e Lieberman (2010). A análise do Simplex em tabelas permite a visualização das transformações que ocorrem no algoritmo e se aproxima mais do processo computacional do que as demonstrações das operações matriciais como usualmente o método é apresentado nos livros texto.

Como foi colocado anteriormente, existem muitos sistemas (*softwares*) que usam algoritmos de otimização linear, muitos outros que resolvem problemas de otimização linear e não linear para resolução de problemas e não é objetivo fundamental deste trabalho examinar o método simplex em sua completude.

O simplex em tabelas funciona conforme se demonstra abaixo. Seja um problema de otimização:

$$\begin{aligned} & \text{Otimizar } f = a_1x_1 + a_2x_2, \\ & \text{Sujeito a: } r_1: b_{1,1}x_1 + b_{1,2}x_2 \leq c_1 \\ & \quad \quad r_2: b_{2,1}x_1 + b_{2,2}x_2 \leq c_2 \\ & \quad \quad r_3: b_{3,1}x_1 + b_{3,2}x_2 \leq c_3 \end{aligned}$$

Colocando o problema na forma padrão: com a criação de variáveis de folga para cada desigualdade:

$$\begin{aligned}
 & \text{Minimizar } z = a_1x_1 + a_2x_2, \\
 & \text{Sujeito a: } r_1: b_{1,1}x_1 + b_{1,2}x_2 + x_3 + 0 + 0 = c_1 \\
 & \quad \quad \quad r_2: b_{2,1}x_1 + b_{2,2}x_2 + 0 + x_4 + 0 = c_2 \\
 & \quad \quad \quad r_3: b_{3,1}x_1 + b_{3,2}x_2 + 0 + 0 + x_5 = c_3
 \end{aligned}$$

Em que as variáveis básicas do problema (aquelas que são criadas para se estabelecer a igualdade das restrições) são definidas por.

$$\begin{aligned}
 r_1: x_3 &= c_1 - b_{1,1}x_1 - b_{1,2}x_2 \\
 r_2: x_4 &= c_2 - b_{2,1}x_1 - b_{2,2}x_2 \\
 r_3: x_5 &= c_3 - b_{3,1}x_1 - b_{3,2}x_2
 \end{aligned}$$

A representação deste problema em uma tabela a qual chamamos de tabela Simplex pode ser expressa conforme a tabela 3.1.

Variáveis	x_1	x_2	x_3	x_4	x_5	
Coefficiente da função objetivo	a_1	a_2	0	0	0	F
Coefficiente das restrições	$b_{1,1}$	$b_{1,2}$	1	0	0	c_1
Coefficiente das restrições	$b_{2,1}$	$b_{2,2}$	0	1	0	c_2
Coefficiente das restrições	$b_{3,1}$	$b_{3,2}$	0	0	1	c_3

Tabela 2.6: Tabela simplex com x_1 e x_2 como variáveis não básicas e x_3 , x_4 e x_5 como variáveis básicas

Fonte: Baseado em Arenales (2007)

É importante identificar que existe uma matriz identidade na tabela que somente expressa os coeficientes das variáveis básicas, fazendo com que, ao usar cálculos matriciais básicos, o problema seja resolvido.

Na origem da variação das variáveis, $(x_1, x_2) = (0, 0)$, o valor da função objetivo é 0 e todas as restrições do problema são atendidas. Contudo, essa solução pode não minimizar a função objetivo. Por exemplo, caso o coeficiente de x_1 na função objetivo, a_1 , seja negativo, para qualquer valor de x_1 o valor da função objetivo é menor que 0 e desta forma, atinge-se uma resolução melhor que $(x_1, x_2) = (0, 0)$ (ARENALES, 2007)

Como existem valores não negativos que minimizam a função objetivo, o algoritmo deve procurar por uma resposta melhor do que $(x_1, x_2) = (0, 0)$. Conforme foi concluído anteriormente, a solução ótima sempre pertence a um vértice da região factível o que torna possível concluir que é possível procurar por valores melhores para a função objetivo em que as variáveis básicas x_3 ou x_4 , ou x_5 , ou ainda as variáveis não básicas x_1 , x_2 , são iguais a 0.

O processo adequado para saber qual das variáveis temos que estabelecer como 0, é identificar quais delas limitam o decréscimo da função objetivo e então estabelecer seu valor máximo. Na forma padrão de minimização em que a função objetivo é expressa, o processo consiste em procurar pela variável não básica com o coeficiente negativo e então identificar qual a restrição que limita seu crescimento, ou seja, faz com que a função objetivo atinja valores menores.

Ao encontrar a variável cujo coeficiente seja negativo, a função da restrição que contém tal variável deve ser igualada a zero. Assim, estamos percorrendo os vértices de nossa zona de soluções factíveis. Em outras palavras, segundo a definição das variáveis básicas e supondo que o coeficiente a_2 a variável x_2 fosse negativo, deve-se estabelecer que $x_n = c_n/b_{n,m}$, tal que x_n é variável não básica e c_n e $b_{n,m}$, respectivamente, o elemento independente e o coeficiente da variável x_2 na equação que restringe x_n . É possível identificar que quando $b_{n,m}$ -é negativo, então o problema não tem solução limitada, pois nada impede que o valor de x_2 cresça fazendo com que o valor da função objetivo seja $-\infty$. Esse procedimento deve ser feito para todo $b_{n,m} \geq 0$.

Agora, ao invés de termos a solução $x_1 = 0, x_2=0$, temos que x_1 e $x_n = 0$, sendo x_n uma variável básica. Para se estabelecer as condições em que as variáveis básicas formam uma matriz identidade, deve-se “pivotar” a tabela anterior e expressar a função objetivo conforme as variáveis que são diferentes de 0. Segundo Arenales (2007), “pivotar” significa fazer operações de multiplicação dos vetores e adicionar o resultado a outro vetor. Essa operação também é chamada de eliminação de Gauss, e não altera o resultado do sistema.

Ao se reproduzir a tabela anterior e supondo que a variável básica x_5 foi anulada (igualada a 0), conforme as alterações descritas, temos:

Variáveis	x_1	x_2	x_3	x_4	x_5	D
Coeficiente da função objetivo	a_1	0	0	0	a_5	F
Coeficiente das restrições	$b'_{1,1}$	0	1	0	$b_{1,5}$	d_1
Coeficiente das restrições	$b'_{2,1}$	0	0	1	$b_{2,5}$	d_2
Coeficiente das restrições	$b'_{3,2}$	1	0	0	$b_{3,5}$	d_3

Tabela 2.7: Tabela Simplex com x_1 e x_5 como variáveis não básicas e x_3, x_4 e x_2 como variáveis básicas

Fonte: Adaptado de Arenales (2007)

Segundo Arenales (2007), esse processo faz com que a variável x_5 saia da base e a variável x_2 entre na base. É como se estivéssemos redefinindo as variáveis básicas e não

básicas, estabelecendo que as variáveis que se anulam (se igualam a zero), são as novas variáveis não básicas e as demais, as básicas, cujo o coeficiente na função objetivo é 0.

Com o “pivoteamento”, uma matriz identidade agora é formada pelos coeficientes da das funções de x_3 , x_4 e x_2 e a função objetivo foi escrita novamente com os coeficientes de x_1 e x_5 e uma nova solução pôde ser encontrada. Esse processo poderia acontecer várias vezes até que uma solução que minimize o valor da função objetivo seja atingida.

O procedimento de se variar o valor de apenas uma das variáveis básicas (que explicam a função objetivo) que foi descrito anteriormente é chamado de *estratégia simplex*. No caso acima, x_2 variou enquanto o valor de x_1 se manteve constante. Da mesma forma, o processo que se inicia com a tabela 2.1 e termina na tabela 2.2, ou seja, o processo que parte de uma solução básica factível para outra solução básica factível é chamada de *interação do método simplex* e acontece variáveis vezes quando se utiliza o método simplex para se chegar no resultado ótimo de um problema (ARENALES, 2007). De forma análoga à exposta no processo gráfico do simplex, a forma algébrica de resolução identifica a região de soluções factíveis ao estabelecer as restrições que limitam o crescimento da função objetivo pela variação das variáveis básicas, e varia as outras variáveis básicas percorrendo os vértices da área de soluções factíveis até se alcançar um valor ótimo.

Como se pode intuir, quanto maior o número de variáveis, maior o número de interações e como consequência lógica, maior o tempo de interações. Métodos derivados do simplex podem fazer com que o tempo de processamento seja menor ao se eliminar soluções óbvias. Alguns destes algoritmos serão avaliados neste trabalho.

Segundo Arenales (2007) as interações do método simplex, que foram expostas nos parágrafos anteriores, podem ser descritas de maneira genérica, quando expressas de forma padrão, pelos seguintes passos:

1. Determine a tabela simplex $m \times n$ (m é o número de equações e n é o número de variáveis básicas e não básicas do problema). É importante frisar que a função objetivo deve estar escrita em termos das variáveis não básicas, ou seja, os coeficientes das variáveis básicas são nulos.
2. Inicie as interações no ponto em que as variáveis não básicas são iguais a 0.
3. Identifique o menor coeficiente a_n das variáveis não básicas da função objetivo. Se $a_n \geq 0$, a solução na interação atual é ótima;
4. Determine $cn/a_{jn} = x_n$, onde c_n é o valor independente e a_{jn} é o coeficiente de x_n na função que o restringe. Se a_{jn} for negativo então não existe solução finita para o

problema. Esse processo faz com que uma restrição básica seja zerada (essa variável entrará na base). Segundo os termos usados por Arenales, (2007), a variável x_n entra na base.

5. Atualize a tabela simplex com o valor fixo de x_n , escrevendo a função objetivo em termos das novas variáveis não básicas e faça o “pivoteamento” da matriz até chegar na matriz identidade com as variáveis.
6. Faça interação = interação + 1 e retorne para o ponto 3.

Os seis passos são de forma simplificada e reduzida o mecanismo de funcionamento do método simplex, principio fundamental deste trabalho e da maioria dos sistemas que tentam resolver problemas de otimização linear.

2.5.3. O método *Branch and Bound*

Os problemas de otimização linear inteira são complexos e normalmente estão associados a tempos de processamento elevados. Isso porque os algoritmos precisam testar as várias soluções inteiras dentro da zona de respostas factíveis para chegar em uma resolução ótima. Tais problemas podem ser solucionados por uma série de métodos e algoritmos sendo a mais usual delas o método *Branch and Bounds*, desenvolvido por A. Lang e g. Doing em 1960 (HILLIER, LIEBERMAN, 2010).

De acordo com Hillier e Lieberman (2010, p. 489)

“o conceito básico subjacente à técnica de ramificação e avaliação progressiva (*branch and bounds*) é dividir para conquistar. Já que o grande problema original é muito difícil de ser resolvido diretamente, ele é dividido em subconjuntos cada vez menores até esses subproblemas poderem ser vencidos”

De forma sintética, o algoritmo *Branch and Bounds* soluciona o problema de forma relaxada, isso é, sem as restrições inteiras. Se o resultado da interação apresentar uma ou mais variáveis não inteiras, então o algoritmo identifica uma variável e divide o problema (*branch*) em dois subproblemas, fixando o valor desta variável no seu limite inferior e no seu limite superior. A variável escolhida para ser avaliada quanto aos seus limites é denominada variável de ramificação. Por limite superior e inferior estamos nos referindo, respectivamente, aos valores inteiros logo abaixo e logo acima da variável que apresentou valores não inteiro na

resolução do problema relaxado (TAHA, 2008). Essa divisão pode ser visualizada na figura 2.13.

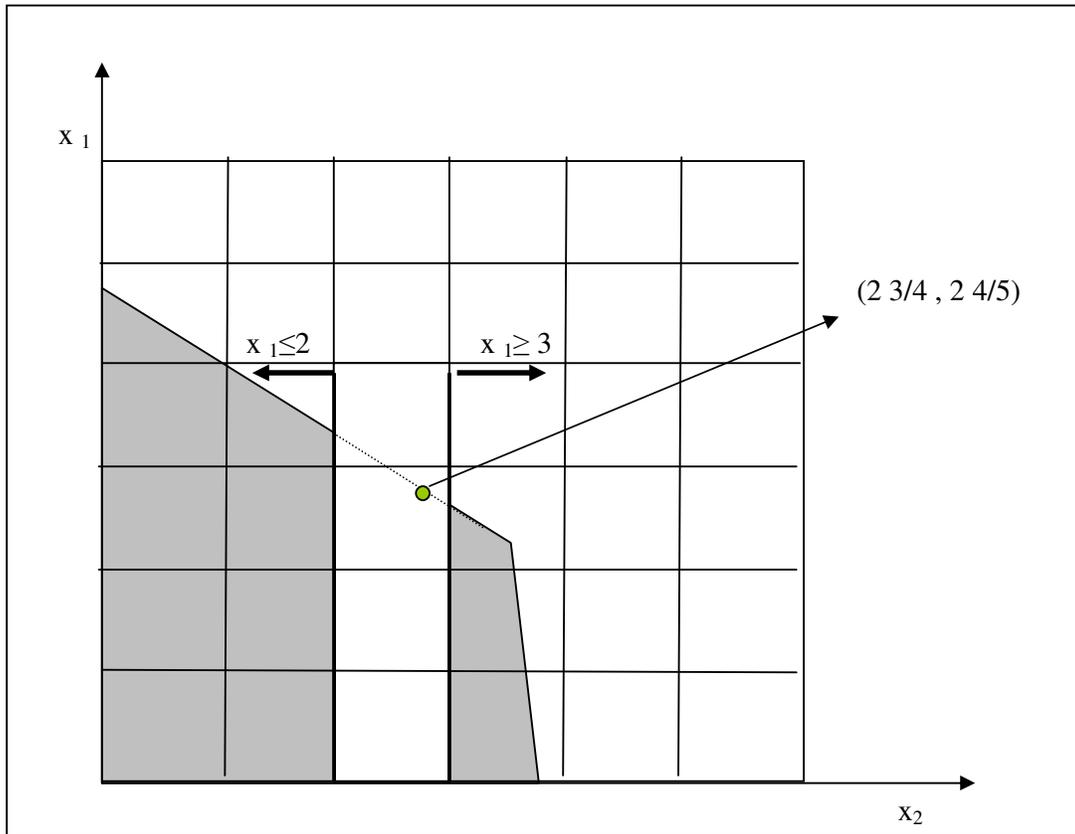


Figura 2.13: Exemplo da resolução gráfica do método *Branch and Bound*
Fonte: Adaptado de Taha (2008) e Arenales (2007)

Em seguida, ambos os problemas (para a o limite superior e o limite inferior da variável escolhida) são examinados e resolvidos pelo método simplex (ou dual simplex). Se a resolução dos subproblemas não apresentar respostas inteiras para as outras variáveis envolvidas no problema, então o problema é novamente dividido novamente criando outras ramificações.

O melhor resultado da resolução dos subproblemas é escolhido como resultado incumbente, o que significa dizer que é o melhor resultado até o momento (TAHA, 2008). Ele é comparado com os resultados das outras ramificações. Se as outras ramificações não apresentarem resultados que possam ser melhores que o resultado incumbente, ou apresentarem resultados não válidos, então essas ramificações são eliminadas. Esse processo dá origem ao termo em inglês *Bound* (delimitação). As várias divisões do problema e as eliminações de

$$\begin{aligned}
 & \text{Minimizar } z = a_1x_1 + a_2x_2, \\
 & \text{Sujeito a: } r_1: b_{1,1}x_1 + b_{1,2}x_2 + x_3 + 0 + 0 = c_1 \\
 & \quad r_2: b_{2,1}x_1 + b_{2,2}x_2 + 0 + x_4 + 0 = c_2 \\
 & \quad r_3: b_{3,1}x_1 + b_{3,2}x_2 + 0 + 0 + x_5 = c_3
 \end{aligned}$$

E que a solução para o problema sem as restrições inteiras seja $(x_1^*, x_2^*) = (z, w)$ onde z e w não são inteiros,

Então o método iria criar 2 subproblemas:

$x_1 = h$, onde h é o maior número inteiro mais próximo de z ;

$x_1 = y$, onde y é o menor número inteiro mais próximo de z ;

O problema seria solucionado para $x_1 = h$ e $x_1 = y$ e criaria outros dois 4 subproblemas. 2 subproblemas para cada valor de x_2 . Os subproblemas seriam eliminados se a ramificação fosse piores que a solução encontrada, até que se encontrasse o melhor resultado e todas as variáveis tivessem valores inteiros.

O método *Branch and Bound* foi amplamente utilizado no estudo de caso de otimização linear na indústria de alimentos, juntamente com o método de realizar cortes que será descrito na próxima sessão.

2.5.4. O método de *Cut Plane* (plano de corte)

Como se pode intuir da descrição acima, a resolução de modelos com programação inteira podem tomar grande tempo de processamento mesmo com o uso do método *Branch and Bounds* e para isso alguns outros métodos complementares são necessários para aumentar a performance de processamento dos modelos. O método *cut* foi criado ainda na década de 50 por Ralph. E. Gomory para a aplicação de problemas com variáveis inteiras (WINSTON, 1995).

Um corte (*cut*) é uma restrição adicional àquelas impostas pelo problema, automaticamente gerada no modelo. Os cortes são gerados dentro do espaço de soluções para eliminar soluções não inteiras dos subproblemas (TAHA, 2008).

De acordo com Arenales (2007 p.251):

“Algoritmos de planos de corte buscam obter uma aproximação da envoltória convexa da região factível de um problema de programação inteira que contemple um ponto extremo correspondente a uma solução ótima. Essa aproximação é obtida por meio de cortes ou desigualdades válidas (...).”

De forma sucinta, o método de plano de corte funciona adicionando-se ao modelo de variáveis inteiras restrições ou desigualdades chamadas de plano de corte, que satisfaçam todas as soluções inteiras da zona de factibilidade e violem (não contemplem) a melhor solução sem a imposição das variáveis inteiras (problema no módulo relaxado). O problema é solucionado com a adição da restrição e se o valor obtido for pior que a solução atual inteira, a restrição imposta é como consequência eliminada (TAHA, 2007).

O processo faz com que algumas soluções sejam cortadas melhorando a performance do algoritmo e aproximando a procura por soluções inteiras da solução ótima com o modelo relaxado, descartando subgrupos de soluções que apresentam valores piores que o atual.

A figura abaixo apresenta de forma gráfica o método de cortes.

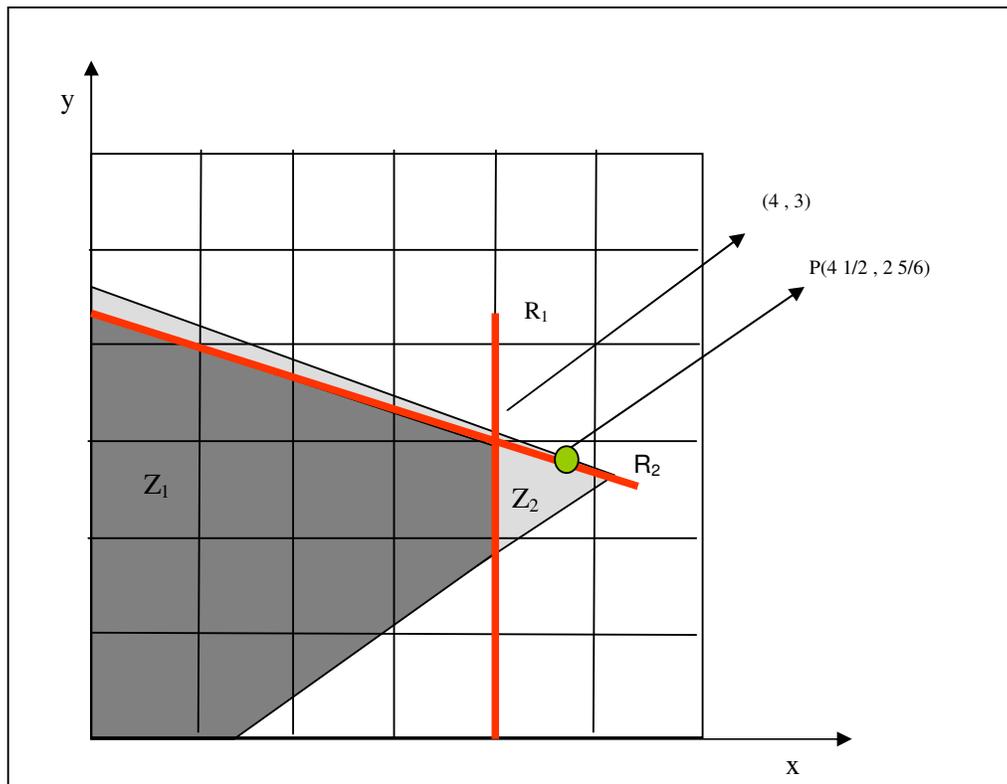


Figura 2.15: Demonstração do plano de corte
Fonte: Adaptado de Hillier e Lieberman (2010)

As retas R_1 e R_2 (em vermelho) são planos de corte que segregam o ponto P, solução ótima do modelo relaxado, e as soluções da área z_1 , que contempla as soluções viáveis inteiras. A área z_2 é a área de soluções ótimas do modelo relaxado. Se as soluções pertencentes aos planos de corte não forem maiores que a solução atual inteira, elas são cortadas.

Taha (2007) afirma que os cortes adicionados não eliminam nenhuma das soluções inteiras viáveis da área de solução possíveis original, mas devem passar no mínimo por um ponto inteiro viável ou inviável. “Esses são os requisitos básicos de qualquer corte” (TAHA, 2007 p.168).

Como iremos descreveremos na sessão sobre testes do modelo implementado na Gelita, existem diversas formas de se executar cortes no espaço de soluções possíveis. Esta sessão teve o objetivo de descrever a forma geral de funcionamento de planos de cortes para melhoria do esforço de processamento.

2.5.5. O modelo *Branch and Cut*

Em algoritmos de alta performance, o método *Branch and Bounds* é conjugado ao método *Cut Plane* para se obter tempos de processamento menores (FRONTLINE PLATFORM, 2006).

De acordo com Arenales (2007 p. 256):

“Esse método combina a estratégia dos métodos de *Branch and Bounds* e de plano de cortes com o objetivo de reduzir o número de nós da árvore de *Branch and Bounds*. Em cada nó da árvore adicionam-se desigualdades válidas de modo a obter um limite superior mais apertado no nó.”

Os cortes podem ser aplicados em cada um dos nós da árvore de busca do método *Branch and Bounds*. Os problemas resultantes depois dos cortes apresentam normalmente um número menor de variáveis não inteiras, o que faz com que o número de nós que devem ser analisados seja reduzido.

Incorporando o plano de cortes ao método *Branch and Bounds* em cada um dos nós da árvore temos:

1. O subproblema é resolvido utilizando-se o método *Simplex* sem a imposição das restrições inteiras.
2. Se a solução encontrada apresentar somente variáveis inteiras, então o subproblema é comparado com o melhor valor atual (incumbente);

3. Restrições são adicionadas ao modelo com o objetivo de reduzir a quantidade de variáveis não inteiras e como consequência diminuir o número de interações;
4. Se uma ou mais de uma variável não contiver valores inteiros, então o método escolhe uma variável e ramifica (*branch*) o problema em dois (com o limite inferior e com o limite superior);
5. O processo de Delimitação (*bounding*) é realizado eliminando as ramificações (conjunto de subproblemas) que apresentem solução necessariamente piores do incumbente;
6. Se existem ainda nós a serem analisados, volta-se para o passo 1, caso contrário a resposta incumbente é apresentada como a melhor solução;

O método *Branch and Cut* foi o método utilizado no estudo de caso presente no capítulo 4. Como iremos analisar, algumas outras formas de cortes foram incorporadas ao algoritmo com o objetivo de diminuir o tempo de processamento do sistema.

Uma vez detalhados os modelos utilizados no estudo de caso, é importante identificar como estes modelos são aplicados no planejamento e organização da produção. Na sessão seguinte iremos analisar os modelos de otimização mais utilizados nos APSs de otimização para o planejamento das empresas.

2.6. APSs de otimização

Os APSs são normalmente aplicados nas empresas para suporte a tomada de decisões em ambientes complexos. Soluções ótimas derivadas de algoritmos de otimização e outros modelos matemáticos têm ajudado empresas a tomarem decisões e se destacarem em mercados de alta concorrência (STADTLER, KILGER, 2005).

Os modelos de otimização linear relacionados a organização e planejamento da produção podem contemplar as mais diversas variáveis que compõe esse complexo ambiente. De acordo com Taha (2008 p.20).

“Há inúmeras aplicações de Programação linear em controles de produção e de estoque, indo desde a simples alocação de capacidade de máquinas para atender a demanda até o caos mais complexo da utilização de equipes para “atenuar” o efeito da mudança imprevisível na demanda para determinada projeção de planejamento...”

Problemas na área de produção em geral são encontrados nos três níveis hierárquicos dos processos das organizações. Problemas no nível estratégico estão normalmente

relacionados a escolhas de arranjos de máquinas, na determinação da capacidade de produção para atender à demanda agregada ou ainda na melhor escolha para a localização das fábricas (ARENALES, 2007). Tal aplicação não será investigada neste trabalho.

A designação da quantidade de recursos: mão de obra, horas extras e subcontratação, bem como a escolha de quando e quanto se deve produzir em cada um dos períodos do horizonte são as variáveis usualmente analisadas nas aplicações de otimização linear no nível tático e operacional. De forma geral as capacidades dos recursos e os limites da demanda são restrições importantes para os modelos bem como os custos de produção, o custo de preparação (*set up*), e estoque elementos presentes no vetor custo da função objetivo.

Embora não exista nenhuma taxonomia que determine os tipos de problema de otimização linear, podemos segregar as aplicações em poucos tipos de problemas existentes na literatura e exemplos práticos. Nos níveis operacional e tático se utiliza a otimização linear para definir a sequência de operações nas máquinas (*scheduling*), as quantidades e os instantes em que os produtos devem ser produzidos em um nível mais detalhado de tempo (*lot sizing*) (ARENALES, 2007).

As aplicações de otimização linear no planejamento e organização da produção guardam uma relação íntima com o tipo de manufatura imposta no processo (TEMPELMEIER, 2005). Assim, uma empresa que produz bens de produção e que como consequência frequentemente apresenta um processo em ilhas de trabalho (*Job Shop*), faz o uso de modelos de sequenciamento (*scheduling*) e da programação dinâmica para a identificação das peças que serão alocadas nos recursos. Da mesma forma, uma empresa de bens de consumo onde normalmente se encontra um processo de produção em fluxo como células em fluxo (*flow cells*) ou fluxos em linha (*flow lines*), os modelos de definição de tamanho dos lotes (*lot sizing*) são provavelmente mais abordados. Abaixo descrevemos os 2 tipos mais utilizados no planejamento e organização da produção segundo os modelos encontrados em Taha (2008) Arenales (2007).

2.6.1. Modelos de dimensionamento de lotes (*lot sizing*)

Os problemas de dimensionamento de lotes são bem conhecidos na literatura e são abordados nos principais textos sobre Pesquisa Operacional e suas aplicações (ARENALES, 2007). Nesta sessão, iremos discutir os modelos determinísticos, de horizontes finitos e descritos em períodos em que a demanda em cada um dos períodos é dinâmica.

Os modelos de dimensionamento se dividem em poucos tipos. Existem os modelos que tratam apenas um de período ou de múltiplos períodos, de um produto ou de multiprodutos (que podem se desdobrar em diversos níveis de produtos), que apresentam capacidade restritiva ou irrestrita de produção/ armazenagem e os modelos que contemplam ou não custos de preparação (*set up cost*) (ARENALES, 2007). Os problemas de dimensionamento de lote podem apresentar várias combinações com os elementos citados.

Além desses, ainda existem os chamados modelos de dimensionamento e programação de lotes (*lot sizing and scheduling*). Isso porque os problemas de dimensionamento de lote normalmente tratam de períodos grandes, tipicamente uma semana, sem considerar a priorização dos lotes nos períodos. Desta forma, segundo Arenales (2007 p. 213):

“Uma tendência recente da literatura consiste em integrar as decisões de dimensionamento e programação de lotes em cada período”.

Os modelos de dimensionamento de lote são aplicáveis em ambientes em que a escolha do tamanho dos lotes interfere na produtividade e no nível de serviço da programação da produção. Como tais modelos não definem os recursos em que as tarefas serão realizadas, o modelo não é geralmente aplicável em um ambiente de ilhas de processo (*Job shop*). Pode-se dizer que o modelo é usualmente aplicado em ambientes em fluxo linear (*flow lines*) ou em células e em ambientes de fluxo de trabalho (*flow shop*) (TEMPELMEIER , 2005).

A seguir descrevemos os modelos matemáticos em que existem múltiplos períodos, múltiplos itens, restrição de capacidade e tempos de preparação conforme Arenales (2007, p. 206 - 211) Sendo assim, temos:

Nomenclatura:

D_{it} = demanda do item i no período t ;

B_i = tempo de produzir uma unidade do item i ;

C_t = capacidade de produção em horas de uma máquina ou instalação no período t ;

Sp_i = tempo de preparação de máquina para processar o item i ;

S_i = custo de preparação do item i ;

h_i = custo unitário do estoque do item i ;

I_{i0} = estoque inicial do item i ;

n = número de itens finais;

T = número de períodos do horizonte de planejamento;

x_{it} = quantidade de itens i produzidos no período t (tamanho do lote);

I_{it} = estoque do item i no fim do período t

$Y_{it} = 1$ se o item i é produzido no período t , e 0 caso contrário;

$W_{it} = 1$ se o estado de preparação para o produto i continua do período $t-1$ ao período t ; 0 caso contrário;

$Q_{ik} = 1$ se não existir preparação do item em t que utiliza o recurso k , dado que a preparação de um item específico aconteceu em $t-1$ e é mantida até $t+1$;

Modelo de lot sizing: múltiplos períodos, múltiplos itens, restrição de capacidade e tempos de preparação.

$$\text{Min } \sum_n \sum_t (s_i y_{it} + h_i I_{it})$$

$$R1: I_{it} = I_{i,t+1} + x_{it} - d_{it}$$

$$R2: \sum (s_i y_{it} + b_{ki} x_{it}) \leq C_k t$$

$$R3: \sum w_{it} \leq 1$$

$$R4: w_{it} \leq y_{i,t-1} + w_{i,t-1}$$

$$R5: w_{it+1} + W_{it} \leq 1 + q_{tk}$$

$$R6: y_{it} + q_{tk} \leq 1$$

$$R7: x_{it} \leq M_{it} (y_{it} + w_{it})$$

$$R8: \min\{(\min c_k t - sp_i/b_{ki}), \sum d_{it}\}$$

$$R9: q_{tk} \geq 0$$

$$R10: W_{it} = \text{binomial};$$

tal que o valor de W_{it} é igual a 1 se o estado de preparação do produto i continua do período $t-1$ ao período t ;

A função objetivo tenta minimizar o custo total de preparação e estoque. A restrição $R1$ refere-se ao balanceamento do estoque dos itens i nos períodos t . A restrição $R2$ expressa que para cada período t e para cada produto i , o tempo total de preparação e produção é limitado pela capacidade disponível ck_t .

As restrições $R3$ e $R4$, $R5$, $R6$ e $R7$ apresentam grande complexidade. A restrição $R3$ garante que para cada recurso k , no máximo, uma preparação é preservada de um período para outro. $R4$ assegura que uma preparação é preservada no período t ($w = 1$) somente se o item i foi preparado em $t-1$ ($Y_{i,t-1}$) ou a preparação foi preservada do período $t-2$ ao período $t-1$ ($w_{i,t-1}$). $R5$ impõe que uma preparação pode se manter por dois períodos consecutivos t e $t+1$,

isto é, $q_{ik} = 1$ se $w_{it} = 1$ e $w_{i, t+1} = 1$; e a restrição $R6$ assegura que isso pode acontecer se não houver preparação no mesmo período. A restrição $R7$ garante que a produção do item i no período t pode ocorrer se o item i estiver preparado no período t ($y_{it} = 1$) ou se a preparação em $t-1$ foi preservada.

A produção máxima M_{it} do item i no período t é dada por $R8$, e as restrições $R9$ e $R10$ definem os tipos de variáveis.

2.6.2. Modelos de sequenciamento da produção (*scheduling*)

Tubino (2007) argumenta que o processo de gerenciamento das atividades no chão de fábrica, o que envolve a designação de tarefas (*jobs*) nas máquinas e sua programação (*scheduling*) pode trazer grande complexidade para os gestores em função do número de variáveis e restrições existentes no sistema.

Os problemas de *scheduling* estão normalmente associados a ambientes de produção do tipo *Job shop* ou *flow shop* em que é necessário realizar o sequenciamento das tarefas nas máquinas mantendo-se os atrasos sob controle ao mesmo tempo em que a utilização das máquinas é otimizada. Os modelos são aplicáveis também em ambientes de uma máquina ou máquinas paralelas (idênticas, uniformes e não relacionadas) (ARENALES, 2007).

Lustosa (2008) argumenta que as principais medidas de desempenho para se avaliar um processo de *scheduling* são:

- *Makespan*: tempo do processamento total das ordens, medido pelo intervalo de tempo entre a abertura da primeira ordem de produção até a conclusão da última ordem de produção de um item.
- Atraso total: Medido pelo tempo em que um item está atrasado;
- Número de tarefas atrasadas: número de tarefas que extrapolaram o tempo em que deveriam ser entregues;

Os modelos de sequenciamento (*scheduling*) devem estabelecer uma boa troca entre tais medidas de desempenho. A seguir definimos o modelo de *Job Shop*.

Segundo Arenales, 2007 p. 222)

“Um *job shop* clássico é um ambiente de produção com n tarefas e m máquinas em que cada tarefa é processada nas m máquinas com um roteiro pré-estabelecido.”

O modelo definido a seguir é baseado em Arenales (2007 p. 222 e 223) e expressa um aplicação de otimização linear em problemas de *scheduling* em um ambiente de *Job shop*. Com poucas modificações nas restrições do problema, o mesmo modelo pode ser aplicado em um ambiente de *shop flow*;

Nomenclatura:

C_i = instante de término do processamento da tarefa i ;

$X_{ij} = 1$ se a tarefa i precede imediatamente a tarefa j , 0 caso contrário;

P_{ik} = tempo de processamento da tarefa i na máquina k ;

S_{ijk} = tempo de preparação da máquina k para processar a tarefa j imediatamente após a tarefa i ;

D_i = data de entrega da tarefa i ;

M = número grande

C_{ik} = Instante de termino do processamento da tarefa i no máquina k ;

$X_{ijk} = 1$ se a tarefa i precede a tarefa j na máquina k , 0 caso contrário;

Modelo geral de sequenciamento (*scheduling*) em ambientes *Job Shop*:

$$\text{Min } \sum i(m)$$

$$R1: C_{i,i(1)} \geq P_{i,i(1)};$$

$$R2: C_{i,i(k+1)} \geq c_{i(k)} + P_{i,i(k+1)};$$

$$R3: C_{jk} \geq C_{ik} + P_{jk} - M_{(1-x_{ijk})}$$

$$R4: C_{ik} \geq C_{jk} + P_{ik} - M_{x_{ijk}}$$

A função objetivo do modelo expressa a minimização do tempo de fluxo total das tarefas. *R1* garante que a primeira operação de cada tarefa i é completa após o respectivo tempo de processamento. *R2* impõe que operação $k+1$ da tarefa i é concluída depois do término da operação k e do tempo e processamento da operação $k+1$. As restrições *R3* e *R4* são “disjuntivas” e indicam que na máquina k , a tarefa i precede a tarefa j , ou a tarefa j precede a tarefa i . Se $x_{ijk} = 1$ então de *R3* e *R4* tem-se que $C_{jk} \geq C_{ij} + P_{jk}$ e $C_{ij} - C_{ik} \geq P_{ik} - M$. Isso é, a restrição *R4* é desativada. De forma análoga se $x_{ijk} = 0$, então $c_{jk} - c_{ik} \geq P_{jk} - M$ e $C_{jk} \geq C_{ij} + P_{jk}$. e então *R3* é desativada. A restrição *R4* estabelece o tipo das variáveis.

2.6.3 Outros modelos de otimização aplicados no planejamento e organização da produção

A literatura é bastante rica quanto à aplicação de otimização linear no ambiente produtivo. Willians (1999) apresenta modelos relacionados à alocação de mão de obra nas linhas de produção. Taha (2008) expõe as aplicações de otimização linear em mistura e refinamento com diferentes preços e restrições de produção para os insumos. Arenales (2007), por sua vez, aborda problemas de balanceamento de linhas de montagem com restrições de procedência entre as tarefas e limites de cargas em cada uma dos recursos. Ainda dentro das aplicações de produção, o autor também expõe problemas de programação de projetos que podem ser similares aos problemas de sequenciamento (*scheduling*) se não são incorporados nos problemas aspectos como recursos renováveis que são limitados ao longo do tempo (como capital disponível ou matérias primas).

Existem diversos problemas que podem ser aplicados na produção, porém o problema de sequenciamento (*scheduling*) e definição do tamanho do lote (*lot sizing*) são os mais tradicionais problemas de alocação de recursos nos ambientes de planejamento e organização da produção (ARENALES, 2007). Ademais, pela sua complexidade ao modelar tais problemas se estabelece uma base de conhecimento que pode ajudar a resolução dos problemas menos incidentes.

A descrição dos modelos de mistura de componentes para a produção de bens não discretos é realizado no quarto capítulo, o estudo de caso.

3. O Método para implementação de APSs de Otimização

Os modelos matemáticos cumprem uma função importante de descrever os problemas e eventos reais. Vários modelos físicos, químicos, financeiros, possuem leis e regras que podem ser descritas usando modelos matemáticos (TAHA, 2008). Para se capturar a dinâmica dos modelos reais, um processo de mapeamento deve ser feito com o objetivo de avaliar quais são os elementos de decisão, as restrições e principalmente as restrições do problema.

Como em qualquer outro processo de modelagem, simplificações devem ou podem ser feitas para representar o modelo real. Um modelo matemático deve ser suficientemente detalhado para capturar os elementos essenciais do problema, mas ainda possível de ser simulado ou resolvido usando-se resumos que não expressam necessariamente a realidade (ARENALES, 2007). A validação dos modelos matemáticos é feita com base no grau de representatividade ou de aderência ao do modelo ou situação real (TAHA, 2008).

Entretanto, a construção de modelos matemáticos que imitam ou representam a realidade não é fáceis de ser realizada. A modelagem matemática de um problema complexo usando-se variáveis livres e restrições numéricas já é por si só um desafio que pode inibir o progresso da tarefa. Dependendo da complexidade e da extensão do problema não se consegue construir com facilidade um modelo tal que represente com alta aderência as dimensões do problema real e ainda seja exequível ou que apresente uma resolução adequada dentro de um curto espaço de tempo (ARENALES, 2007)

A modelagem é o pré-requisito dos testes funcionais e de performance. Os testes visam aproximar o modelo da realidade e o expor aos diversos problemas que podem surgir em sua operacionalização (LANGLEY e MOAN, 2011). Nesse sentido, o modelo deve ser submetido às mesmas condições que irão influenciar o modelo posto em funcionamento. Esforços em customização e mudanças no modelo para o aperfeiçoamento do tempo de processamento serão eventualmente necessários

Ao executar os testes é possível analisar os resultados e identificar se o modelo realmente conseguiu representar com adequada aderência a realidade do problema. O processo de leitura e análise dos resultados de um modelo matemático não são nada simples quando se trata problemas complexos. É nessa etapa do processo de construção de um modelo matemático que, preponderantemente, segundo os resultados deste estudo, o modelo deve ser revisado. Isso pode acontecer em função da modelagem que não demonstrou aderência o suficiente para representar a realidade ou em função do alto tempo de processamento.

A fase de implementação é o último passo e se refere ao uso constante e contínuo do modelo para fundamentar as decisões da empresa no processo de planejamento e controle da produção. Se os testes foram realizados com eficiência e as modificações devidamente efetuadas e aprovadas, a operacionalização é certa. Se o problema for complexo e as decisões que se devem realizar importantes para o negócio, este processo é uma boa receita para um projeto de grande agregação de valor para a empresa.

Arenales (2007) sugere uma figura que ilustra o processo de modelagem matemática:

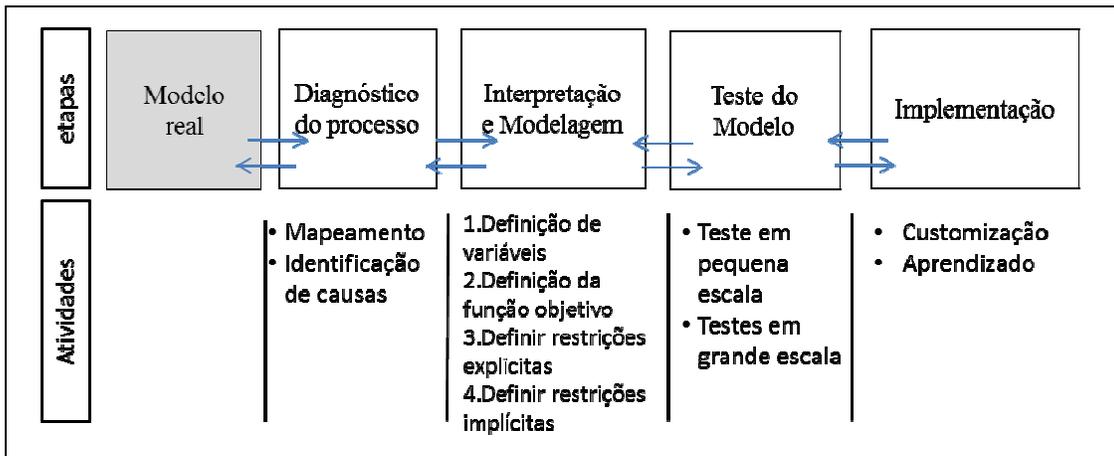


Figura 3.1: Processo de construção de modelagem matemática

Fonte: Adaptado de Arenales (2007)

A figura 3.1 irá servir como mapa para a descrição do estudo de caso; ela demonstra o método utilizado para construir o modelo de otimização. Como não se poderia deixar de pensar, o modelo de otimização linear que se aplicou na Gelita percorreu os 6 passos da figuras 3.1 apresentando conclusões importantes dentro do objetivo e do problema que se pretende destacar neste trabalho.

3.1. Diagnóstico do processo

Chamamos de diagnóstico a atividade de identificação e avaliação de um processo (LANGLEY e MOAN, 2011). O processo de diagnóstico aqui destacado refere-se a avaliação do funcionamento do planejamento de produção como um todo. Isso é, aos elementos relacionados ao sistema, estrutura e processo em si foco do estudo.

Existem várias metodologias de diagnóstico de processos sendo um dos mais conhecidos na gestão de melhorias, o mapeamento de fluxo (SHOOK, 2003). Normalmente o

mapeamento serve para descrever um processo de forma simples. Na construção de sistema de planejamento esta atividade é importante para se conhecer e aprender como o processo funciona atualmente e quais são os principais problemas que devem ser saneados (Shook, 2003). Ao mesmo tempo, o mapeamento também serve para identificar os requisitos que o sistema deve possuir para que sirva de suporte a tomada de decisão no planejamento. Segundo Deming (1982) o primeiro passo para ganhar controle e conseguir melhorar um processo é conhecer e entender como ele funciona.

Normalmente o mapeamento descreve o encadeamento das atividades que acontecem em uma cadeia de processos (Shook, 2003). São usadas caixas com a descrição das atividades e setas que determinam as entradas e saídas das informações ou produtos que são processados nas atividades (Slack, 2008). A figura 3.2 descreve um mapeamento de processo.

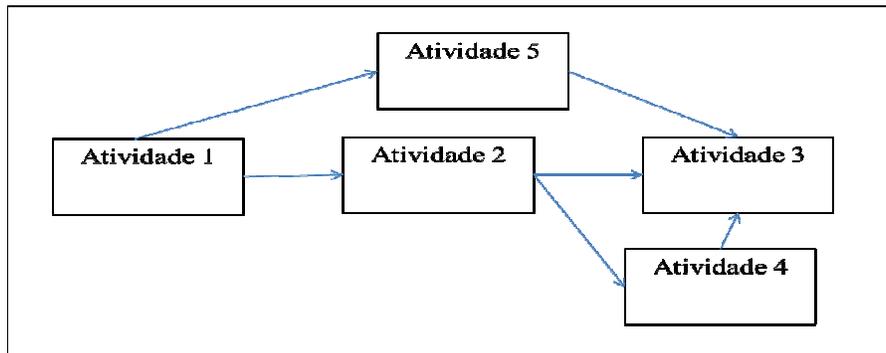


Figura 3.2: Mapeamento de processo;
Fonte: Adaptado de Slack (2008)

Com o mapeamento consegue-se entender quais sistemas e eventuais problemas impedem que seja realizado com maior eficiência ou eficácia (Slack, 2008). Assim, o segundo passo na realização do diagnóstico é a avaliação dos problemas que acontecem durante o processo de planejamento.

Os problemas não necessariamente indicam quais são os requisitos ou funções que um sistema ou processo deve realizar, mas sim, quais são os efeitos das possíveis causas existentes no processo atual. É preciso coletar evidências ou estudar sobre quais são causas dos problemas atuais que possam apontar para um processo ou sistema mais eficiente (Falconi, 2004).

Deming (1986) e Falconi (2004) apontam a importância da análise e estudo dos problemas atuais para melhoria dos processos. O estudo e a análise de dados como os tempos das etapas de planejamento, o grau de assertividade do planejamento frente ao que realmente

foi realizado ou os limites e restrições do processo de planejamento são elementos fundamentais para a construção de um sistema *APS* de suporte ao planejamento de produção. Tais estudos ajudam a identificação das causas dos problemas e conseqüentemente quais elementos o novo processo e sistema deve realizar para aumentar a performance de planejamento de produção. A tabela 3.1 adaptada de Falconi (2004) ilustra a organização dos problemas, dá importância destes problemas para o processo e os efeitos que eles causam e quais são suas causas fundamentais.

Problemas	Grau de importância do Problema	Efeitos	Causa
Descrição dos principais problemas existentes no processo	Identificação quantitativa do grau de importância dos problemas para o processo	Identificação dos efeitos qualitativos ou quantitativos causados pelos problemas	Causas ou potenciais causas/ evidências dos problemas

Tabela 3.1: Tabela de organização de diagnóstico do processo;
Fonte: Adaptado de Falconi (2004);

O preenchimento da tabela 3.1 é realizado com as informações do mapeamento e a análise dos dados levantados para se evidenciar as causas dos problemas. O conhecimento das pessoas que conhecem e gerenciam o processo é um elemento fundamental para que o diagnóstico do processo, e como consequência a identificação correta das causas aconteça adequadamente (LANGLEY e MOAN, 2011).

Para a construção de um *APS* de otimização, o foco de atuação está na eliminação ou minimização dos problemas relacionados ao sistema e metodologia de planejamento de produção.

3.2. Interpretação e modelagem

Esta sessão explora o método de modelagem de um *APS* de otimização para o planejamento e organização da produção. Mais especificamente, descreveremos aqui o método de modelagem de um problema de otimização linear que compõe o método de funcionamento da maioria dos *APSs* de otimização. Taha (2008) sugere 4 passos para transformar um problema real em um modelo de otimização linear. Estes 4 pontos

Passo 1: Defina as variáveis ou incógnitas do problema.

As variáveis do problema são as incógnitas cujas respostas o modelo de otimização deve definir. Este primeiro passo demanda um alto conhecimento do processo de planejamento e de quais são as variáveis que os gestores definem em sua rotina de tomada de decisões.

Passo 2: Defina a grandeza que se deseja otimizar ou minimizar e a expresse matematicamente em uma função.

Todo o modelo de planejamento de produção tem um objetivo claro. Na grande maioria das vezes os sistemas de planejamento têm o objetivo de aumentar a performance dos recursos de produção e atender de melhor forma possível os clientes do processo (SLACK, 2009). Assim, de forma generalizada, os sistemas de planejamento tentam de forma direta ou indireta diminuir o uso de recursos produtivos para atender um número maior de clientes com produtos que desejam (Falconi, 2004).

Passo 3: Identifique todas as exigências, restrições e limitações do problema real e as expresse matematicamente em função lineares.

Os recursos produtivos são normalmente elementos finitos. As fábricas e equipamentos apresentam capacidade produtiva restrita à disponibilidade com que trabalham. Da mesma forma, normalmente os clientes aceitam receber seus pedidos apenas dentro de uma janela de entrega restrita e na quantidade solicitada. O terceiro passo exige que se descreva de forma matemática em funções lineares as restrições físicas e exigências do problema.

Passo 4. Expresse as condições implícitas do problema.

Tais condições não são estipuladas explicitamente pelo problema, mas são evidentes a partir da situação física real que está sendo observada. Geralmente estas condições envolvem requisitos de não negatividade, valores que devem ser considerados como inteiros ou regras de prioridade que não são explícitas diretamente pelo modelo. Um cliente que deve ser preferido a outro deve estabelecer um valor de prioridade maior do que um cliente que tem menos importante. Da mesma forma, se os gestores preferem a utilização de um recurso produtivo a outro em função de risco não mensurado, o recurso com menor prioridade deve

fazer com que a função objetivo atinja valores menores do que quando utilizam o recurso com menor risco.

3.3 Testes do modelo

Os testes ajudam a estabelecer se um sistema ou processo é adequado ou causa impacto positivo no processo antes de ser implementado e operacionalizado no dia a dia pelas pessoas. Assim segundo (Langley e Moan, 2011 pg 47).

“Testar é uma maneira de experimentar temporariamente, na prática, uma mudança e aprender sobre seu potencial impacto. (...). Os testes devem ser elaborados de tal maneira que o investimento em tempo, dinheiro e risco seja o menor possível, ao mesmo tempo em que se aprenda o suficiente para passar à implementação completa.”

Em testes de modelos matemáticos, o primeiro passo é identificar se as restrições e as variáveis expressam realmente o problema real (Taha, 2008). Ao contrário de outros testes que tentam identificar se uma mudança no processo resultou em uma melhoria, os testes de modelos matemáticos têm o valor de identificar a aderência ao modelo real. Desta forma, estamos identificando a relação do modelo matemático com o mundo real, ou testando o grau de convicção que se pode ter no modelo matemático (LANGLEY e MOAN, 2011).

Segundo Langley e Moan (2011), para que os testes sejam bem executados e controlados, duas fases devem ser realizadas:

Testes em pequena escala

Manter os testes em pequena escala no primeiro momento possibilita a identificação da aderência do modelo de forma mais simples e rápida. Não é necessária uma análise sofisticada para identificar que o resultado da interação do modelo tem uma relação forte com o modelo real. Assim, o número menor de variáveis ajuda na redução de tempo e complexidade para a avaliação do modelo.

No caso de modelos matemáticos de otimização linear, os testes de pequena escala não contemplam os eventuais problemas que podem acontecer no tempo de processamento do sistema. Se por um lado isso agiliza o processo de testes, por outro pode resultar em uma perda de tempo, dado que segundo Kjellsdotter (2001), os sistemas de otimização linear que possuem muitas variáveis e restrições apresentam elevados tempos de processamento.

Embora os testes de pequena escala tentem estabelecer a relação entre o modelo e o problema real, é importante argumentar que por se tratar de um teste com variáveis reduzidas, ele pode não trazer todas as variáveis e complexidade do modelo real. A avaliação do teste deve ser realizada com cautela e as condições de comparação com o modelo real devem ser reduzidas.

Testes em grande escala

À medida que os testes em pequena escala ajudam no aprendizado e possibilitam ajustes no modelo, são incluídos elementos diferentes que auxiliam a identificar possíveis falhas no modelo (LANGLEY e MOAN, 2011). Em testes de grande escala de modelo de otimização, elementos importantes do modelo real são incorporados nos testes para o aumento da aderência ao problema real.

Conforme Hillier e Liederman (2010), o número de variáveis em um problema de programação inteira aumenta exponencialmente as interações para a resolução dos problemas. Os testes em grande escala dos modelos de otimização trazem o problema de tempo de processamento à tona. Neste sentido, os testes ajudam não só a medir a aderência do modelo o problema real, como também a encontrar um equilíbrio entre o número de variáveis do modelo real contempladas no sistema e o tempo de processamento.

Os testes de grande escala dos modelos de otimização têm dois objetivos distintos. O primeiro é identificar se o modelo é aderente ao problema real apesar de suas simplificações e generalizações. Neste sentido, o teste tem grande similaridade com os testes de pequena escala. Entretanto, agora as condições de comparação são maiores e toda a complexidade do modelo real deve ser avaliada e comparada com o modelo matemático. O segundo é a adequação do número de variáveis ou complexidade do modelo com o tempo de processamento. Quando tratamos de modelos matemáticos grandes e com variáveis inteiras este tema toma grandes proporções. Entre as várias formas de se reduzir o tempo de processamento de interação de um modelo de otimização com variáveis inteiras, duas são de grande importância.

A primeira é o afrouxamento das tolerâncias das variáveis inteiras e de otimização. Pode-se optar por não obter um valor exatamente inteiro para as variáveis inteiras do modelo, mas um valor bem próximo e com isso diminuir o tempo de processamento. Da mesma forma, pode-se escolher uma distância ou uma porcentagem entre a resposta final e a resposta ótima do modelo sem variáveis inteiras. Assim que a interação atinge uma resposta que atende a

todas as restrições e esteja dentro da distância estabelecida entre a resposta atual e a resposta ótima, o valor atual da função objetivo é aceito o modelo interrompe a interação.

A segunda se refere à tentativa de diminuição do número de variáveis inteiras do modelo por intermédio de restrições não inteiras e condições de otimização. Assim, pode-se estabelecer que um variável pode ser menor ou igual a um valor inteiro e que a função objetivo atinge maiores valores quanto maior é o valor desta variável. Esta colocação faz com que o valor da variável tenda para um valor inteiro, mas não existe garantia de que isso aconteça.

3.4. Implementação

Embora possa se pensar que após os testes os sistemas de otimização simplesmente funcionam, isso não é uma verdade absoluta. Podemos dividir a fase de implementação de um APS em duas grandes fases que podem ou não acontecer ao mesmo tempo.

Adaptação do sistema APS (Customização)

Conforme afirma Tonini (2006), a implementação de softwares como o APS necessita de uma fase de customização para adaptar o sistema aos processos e tarefas que as empresas executam na prática. Oliveira (2011) e Correa e Giansi (2011), colaboram com essa colocação afirmando que uma das grandes dificuldades na implementação de softwares é o nível de especificidades dos processos e tarefas das empresas que demandam customizações e adaptações.

Embora existam os APSs de mercado como citamos no capítulo 2, estes softwares podem ser bastante genéricos e, dependendo da indústria ou setor, sua aplicabilidade pode ser relativamente baixa. Assim, depois da fase de testes do sistema de otimização, é necessário realizar pequenas customizações no sistema para que se adapte ao modelo real.

A realização de adaptações ou customização dos APSs não significa que os testes foram feitos de forma ineficiente ou de forma errada. Elas são coerentes com o uso no dia a dia das pessoas de um novo sistema. Demais disso, os processos de planejamento e organização da produção são constantemente modificados em função do dinamismo dos mercados, da melhoria dos processos e mudanças na dinâmica das empresas (Slack, 2006). Tais modificações impõem modificações nos sistemas de suporte.

Os *APSs* são utilizados por pessoas. Estas pessoas têm que ter facilidade para manuseá-los e gerenciá-los, caso contrário a implementação do sistema pode ser arruinada. Relatórios analíticos, sistemas que facilitem a execução e manipulação de dados para a tomada de decisão, são importantes no processo de implementação e para tornar o manuseio e análise das informações o mais fácil e prático possível. Como impõe Langley e Moan (2011), a implementação de uma mudança ou de um novo sistema pode trazer o total comprometimento ou a hostilidade declarada pelas pessoas que gerenciam o processo. Assim, a adaptação do sistema no processo de implementação cumpre um papel fundamental de adaptar o sistema ao processo a as pessoas que gerenciam o planejamento e organização da produção.

Aprendizado sobre Otimização no Planejamento

Da mesma forma que existe a fase de adaptação ou customização dos sistemas aos métodos que as pessoas gerenciam o processo de planejamento e organização da produção, deve existir a fase de adaptação das pessoas aos conceitos de otimização do sistema. A mudança de um processo de gerenciamento qualquer para um sistema de otimização não é sutil e tampouco simples. A leitura das respostas de um sistema de otimização demanda uma capacidade analítica e muitas vezes holística que necessariamente era exercitada antes do processo de implementação do sistema.

O êxito da implementação de um *APS* depende da interação produtiva dos gestores de planejamento e o sistema. Quanto maior o conhecimento profundo dos gestores do processo de planejamento quanto aos fundamentos dos sistemas *APSs*, maior é a produtividade de todos os recursos necessários para a realização da produção (TEMPELMEIER, 2005).

O processo de analisar os resultados das interações dos *APSs* por si faz com que os recursos humanos adquiram conhecimento sobre o processo de planejamento com uso de otimização, porém é necessário também o conhecimento de conceitos fundamentais de otimização para se conseguir usar modelos para apoio de tomada de decisões. Como afirma Tempelmeier (2005 p. 8):

“Para um planejador bem informado que está ciente dos tipos de problemas que são resolvidos e quais não o são, o uso do *APS* pode fazer com que a produtividade de suas ações aumente (...). Porém, para um planejador não muito bem informado, a frustração da falta de entendimento do funcionamento do sistema pode fazer com que os antigos métodos baseados na “regra do dedão.”

O processo de capacitação em sistemas e conceitos de planejamento também é consideração de grande importância para Tempelmeier (2005 p.8):

“ Lições bem definidas sobre competências em planejamento devem ser consideradas como um pré-requisito para o sucesso de uma implementação do *APS*”.

Assim, pode-se concluir que o êxito da implementação de um *APS* de otimização está fundamentalmente estabelecido no conhecimento analítico dos gestores com relação ao sistema e na adaptação do sistema ao processo e as pessoas que os gerenciam.

Capítulo 4 - Estudo de caso:

Este capítulo expõe o estudo de caso da aplicação de um *APS* de otimização em uma empresa que produz bens alimentícios e farmacêuticos. Nas sessões que se seguem são expostos o diagnóstico do problema, a modelagem, os testes e a implementação do *APS* utilizado na Gelita.

O modelo explorado consiste na mistura de dois ou mais produtos componentes com certas propriedades para se atingir outros produtos finais com propriedade derivadas da média ponderada dos seus produtos componentes. Este processo de misturas é conhecido dentro da literatura e está exposto nas obras de Williams (1999), em um exemplo de Taha (2008) e na clássica exposição do problema da ração apresentado por Arenales (2007) e outros autores.

O gerenciamento de processos de misturas pode ser bastante complexo. Isso porque os gestores destes tipos de processos enfrentam os mesmos problemas existentes em um processo de produção discreta: atraso de fornecedores, variação nos processos de produção e variação de demanda. No entanto, nos processos de mistura, são várias as formas de se obter um mesmo produto usando-se insumos com propriedades diferentes em proporções diversas.

O caminho para atingir níveis de controle e planejamento eficientes em empresas que misturam insumos, contudo, não é simples. O uso dos *MRPs* devido à inexistência de lista de matérias, roteiros de produção e fidelidade de componentes, não pode ser cogitado. Por sua vez, os modelos baseados no sistema puxado são inócuos para gerenciar processos contínuos pela inexistência de unidades de estoque ou *SKUs* (*stock keeping Units*) específicos.

Desta forma, as empresas que gerenciam processos de mistura, muitas vezes estabelecem processos empíricos manuais em que se comparam especificações da demanda futura com o estoque de insumos nos armazéns e tonéis, usando-se longas planilhas de dados para fazer o planejamento e a programação das operações. No entanto, tal comparação não é tão simples quanto se possa desejar.

A manipulação empírica e manual desses estoques demanda operações espúrias dos produtos na tentativa de criar famílias de características semelhantes de produtos causando erros de análise no momento de se fazer o planejamento. Demais disso, um grande questionamento desse processo gira em torno do uso ótimo dos estoques. Isso porque se pode misturar os produtos de diversas formas para se produzir os produtos demandados pelos clientes. Estas misturas feitas manualmente podem comprometer o processo de planejamento e programação e onerar a qualidade e o prazo das entregas aos clientes.

Métodos mais eficientes tanto de controle e planejamento de produção quanto de empenho de misturas podem ser elaborados com o uso de otimização linear (ARENALLES, 2007). A modelagem de um sistema de equações linearmente dependentes e relacionadas com o fim de atingir um objetivo específico faz com que as empresas que se utilizam de misturas consigam atingir um nível de controle e acurácia de programação melhor.

Na próxima sessão desse capítulo iremos explorar os conceitos que norteiam a diagnóstico e implementação do APS de otimização. Usamos aqui o método explorado no capítulo 3.

4.1. Diagnóstico do processo de planejamento de produção

A *Gelita South América* é a divisão da América Latina da Gelita AG, empresa alemã líder mundial em fabricação e venda de gelatina animal. A Gelita oferta gelatina para os seguimentos alimentício e farmacêutico do mundo inteiro, possuindo mais de 2500 especificações de produtos distintos. As principais aplicações de gelatina são as matérias primas de iogurtes, gelatinas em pó comestíveis e cápsulas de remédios.

A cadeia de produção da Gelita é composta por quatro elementos. A empresa compra matéria prima de curtumes de couro bovino e suíno localizados em diversas regiões do Brasil. Esse material é enviado para as três plantas de produção de gelatina. Nas plantas o couro bovino é transformado em gelatina na forma de grãos depois de passar por um longo período de produção. Após a transformação do couro em gelatina, o produto semiacabado é enviada para uma unidade de moagem e mistura na forma de lotes de produção. Neste estágio a gelatina é armazenada em grandes sacos de aproximadamente uma tonelada chamados de *big bags* (grandes sacos).

Estes *big bags* são utilizados pelos planejadores para compor os pedidos dos clientes. Os planejadores literalmente realizam uma seleção dos lotes de produção que serão misturados dentro dos moinhos para atender as especificações de propriedades dos clientes. Iremos descrever melhor o que são as propriedades e as especificações das gelatinas adiante. Por hora é importantes definir as propriedades como as características que a gelatina apresenta em seu estágio final tal como viscosidade, PH, odor, cor entre outros. As especificações são os intervalos (ou *ranges*) de propriedades requisitadas pelos clientes.

Finalmente os pedidos são analisados em laboratório para identificar se suas propriedades estão dentro daquela que foram solicitadas pelos clientes, e transportados para os clientes. Esse processo é descrito detalhadamente ao longo do capítulo.

Antes de prosseguirmos vale a pena explorar uma figura que demonstra o mapa de fluxo de valor dos processos de planejamento da empresa. As setas largas representam o fluxo de matérias entre os fornecedores de matéria prima, as plantas produtoras de gelatina, a unidade de moagem e os clientes de dentro e fora da América Latina. As setas mais finas representam o fluxo de informações.

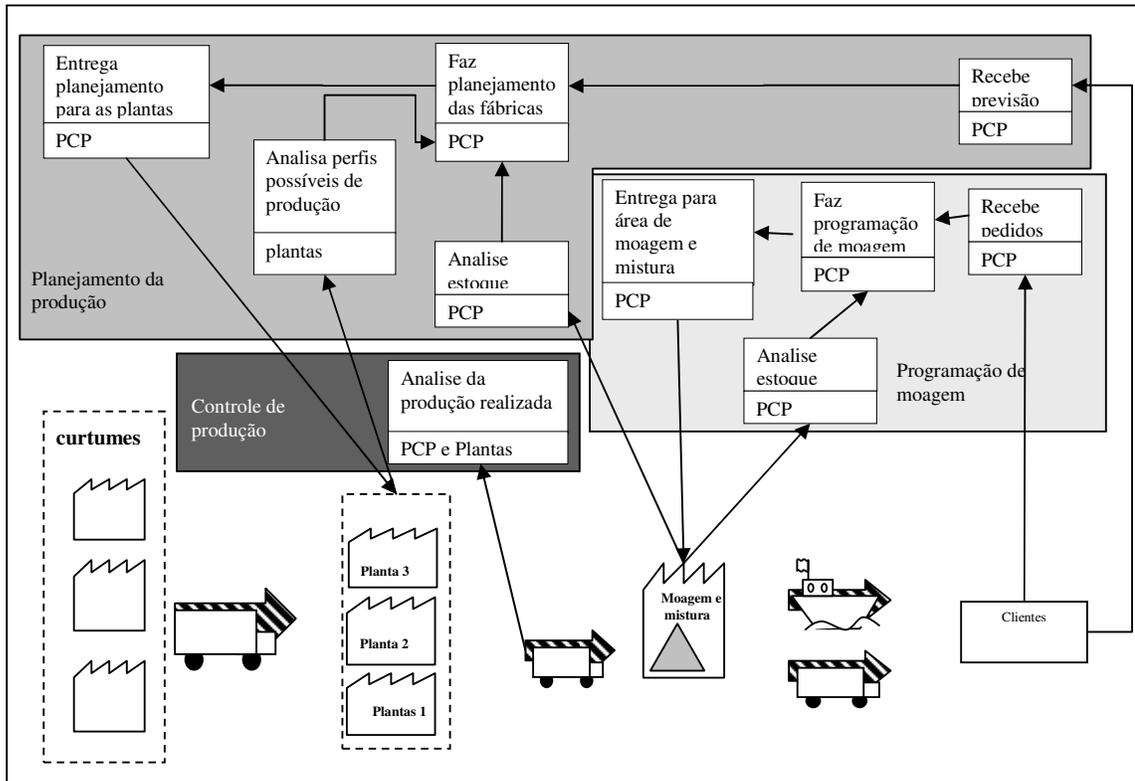


Figura 4.1: Macro mapa de fluxo de valor do processo de planejamento e programação da produção de gelatina

Fonte: Elaborado pelo autor

A figura 4.1 ilustra os três ciclos do macro processo de planejamento e controle de produção de gelatina:

O ciclo de programação de moagem (4.1.1) que se inicia com o recebimento dos pedidos e é finalizado com a emissão das Ordens de Moagem.

O ciclo de planejamento da produção (4.1.2.) que começa com o recebimento das previsões dos pedidos e termina com a emissão da sinalização de produção;

O processo de controle de produção (4.1.3) que analisa qual foi a acuracidade da produção de gelatina das plantas diante do que fora especificado.

Nas próximas sessões aprofundaremos a descrição do sistema de planejamento da Gelita antes do uso do *APS* de otimização.

4.1.1. Programação de moagem

Nossa descrição começa de traz para frente. A programação de mistura é a última etapa do planejamento de produção das gelatinas antes de serem transportadas para os clientes. Para atender às especificações dos pedidos de gelatina dos clientes, os planejadores da Gelita programam a sequência de moagem de gelatinas nos moinhos. Para isso as especificações dos pedidos dos clientes que devem ser entregues em um curto prazo são comparadas com propriedades do estoque de gelatina em processo (chamados de lote de produção) para se determinar quais os lotes devem ser misturados para atingir as especificações dos pedidos. Ou seja, nesta etapa são selecionados os lotes de produção que devem ser misturados para se atingir as especificações dos pedidos dos clientes.

As propriedades são as características intrínsecas das gelatinas tais como PH, viscosidade, cor, odor e entre outras. As gelatinas apresentam varias dezenas de propriedades, assim como os lotes de produção apresentam propriedades diferentes uns dos outros. Entende-se por especificações um conjunto de intervalos de propriedades os quais representam as necessidades dos clientes.

De uma forma simples, os produtos que são entregues para os clientes não podem conter propriedades fora dos intervalos estabelecidos nas especificações.

Uma vez estabelecidos quais são os lotes de produção que devem ser misturados para se atender aos pedidos, é emitida uma Ordem de Moagem (OM) com os lotes de produção que devem ser colocados nos moinhos. As OM são então sequenciadas em função de suas datas de expedição e enviadas para a área de moagem que se responsabiliza por coletar os lotes de produção (ou *big bags*) no armazém e dispensar seu conteúdo nos moinhos.

O Objetivo primordial do processo de programação de moagem é atender o maior número possível de pedidos dentro das especificações dos clientes.

A figura 4.2 ilustra a concepção geral do processo de programação de moagem antes do uso do *APS*.

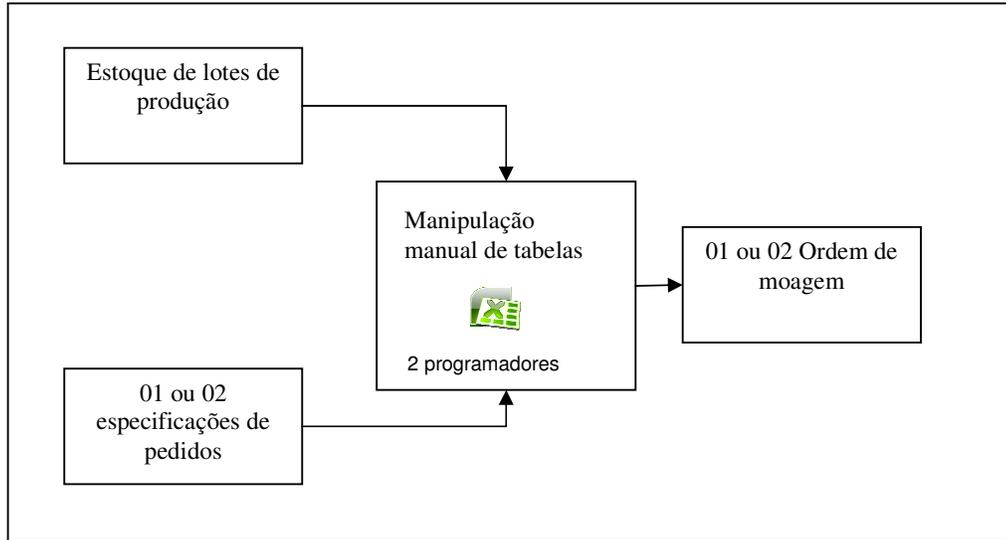


Figura 4.2: Esquema Ilustrativo do sistema de programação de moagem sem o uso do APS

Fonte: Elaborado pelo autor

Realizar as montagens manualmente, porém, dificilmente faz com que as montagens sejam feitas com o uso ótimo do estoque. Isso se deve ao fato de haver centenas de milhares de variáveis a serem analisadas na mistura dos lotes de produção para conseguir o maior número possível de ordens de moagem. Demais disso, antes da implementação dos APSs, os pedidos eram montados de maneira individual, ou seja, se empenhava um pedido e depois que esse estivesse no sistema como uma OM para gerar o produto final, analisava-se a sequência de pedidos e se fazia outra análise de mistura. Essa rotina fazia com que alguns lotes de produção pudessem ser utilizados erroneamente, causando o mau uso do estoque. O número de pedidos atendidos pode ser maior, utilizando-se o mesmo montante de estoque, caso se analisasse todos os pedidos em certo horizonte de tempo em conjunto.

Outro desconforto importante do processo anterior ao uso do APS era o tempo de processamento de se empenhar manualmente as OMs. Em uma coleta de dados sobre o tempo de planejamento das montagens das ordens de moagem, diagnosticamos que em média demorava-se entre 30 e 120 minutos para se empenhar 20 toneladas. Este tempo é relativamente longo se comparado com os tempos de empenho atingidos com a implementação do APS. Esse tempo de atravessamento, além de fazer com que o custo administrativo para calcular as misturas seja elevado, também aumenta a probabilidade de que houvesse quebras de produção devido a atrasos de entrega de ordens de produção.

Para organizarmos os desconfortos e causas da sistemática de programação de moagem sem o uso do APS fazemos um pequeno resumo em forma de tabela.

Processo	Problemas	Efeitos	Causa
Programação de moagem	Não otimização do uso do estoque ao realizar as misturas	- Falta de emissão das ordens de produção ocasionando atrasos; - Atrasos para os fornecedores; Altos estoques;	- Falta de sistema de otimização de misturas;
	Longos tempos para se programar as ordens de moagem	- Altos custos administrativos para realizar a programação de moagem;	- Falta de sistema para facilitar o processo misturas;
		- Risco de atrasos de pedidos pela demora na elaboração das ordens de moagem;	- Demora na aprovação das ordens de moagem com exceção;
			- Falta de recursos para realizar o processo;

Tabela 4.1: Resumo dos problemas no processo de programação de moagem sem o uso do APS

Fonte: Relatório de diagnóstico, Gelita

4.1.2. Planejamento de produção

Os lotes de produtos a que se fez menção nos parágrafos anteriores são os produtos elaborados pelas unidades produtivas que tratam o couro bovino até o ponto em que se transforma em gelatina. O planejamento das propriedades mais adequadas para a composição de tais lotes é feito pelos planejadores da produção que comparam os pedidos em um médio prazo com o perfil de produção de cada planta e com o estoque de lotes de produção existente no armazém. Depois desta comparação, os planejadores sinalizam para as plantas quais tipos de gelatinas deve-se produzir.

De forma sucinta, o planejamento da produção tem o objetivo de especificar o tipo de gelatina que deve ser produzido para atender aos pedidos em um médio prazo. A gelatina que esta sendo especificada no planejamento da produção irá compor o estoque de lotes de produção que serão misturados e entregues para os clientes.

Antes da implementação do APS, as informações sobre os estoques de lotes de produção eram manualmente agrupados bem como os pedidos e a previsão dos pedidos dos

clientes em um médio prazo (aproximadamente 2 meses). A análise era realizada apenas para três variáveis, bloom, viscosidade e transmitância. O planejador comparava os pedidos e a previsão de pedidos com o estoque de lotes de produção. Os pedidos que não se conseguia atender com o estoque de lotes de produção já existentes, eram direcionados para as plantas realizarem a produção nas próximas semanas. O planejador se comunicava com os gestores da planta sobre os tipos de produção que as plantas conseguiam produzir, o que nomeamos de “perfil das plantas”. O processo de planejamento de produção era realizado conforme é mostrado na figura 4.3.

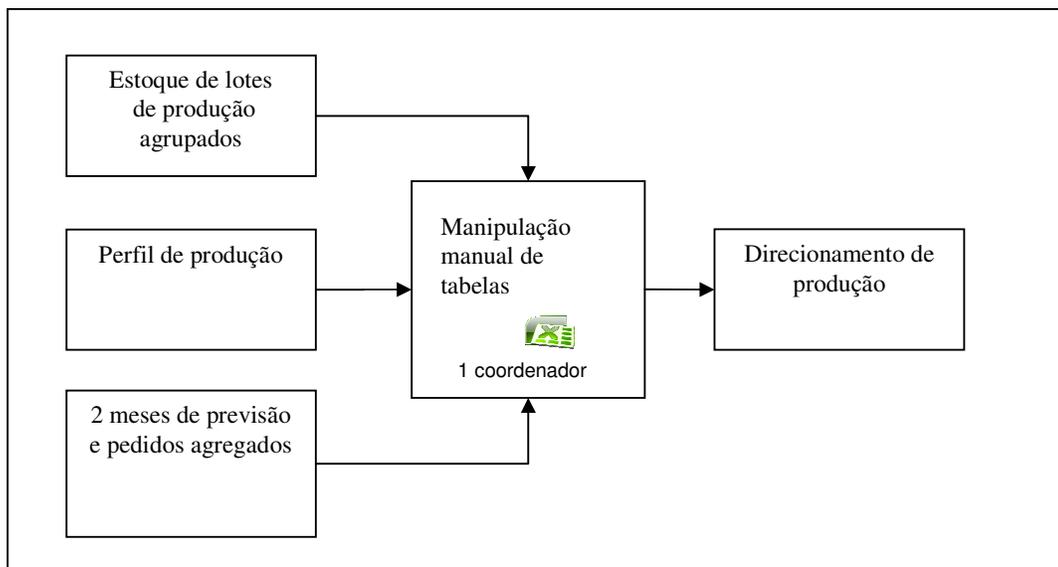


Figura 4.3: Esquema ilustrativo do processo de programação da produção sem o uso do APS

Fonte: Elabora pelo autor

As misturas de gelatinas entre os lotes de produção existentes no estoque e os lotes de produção que eram solicitados para as fábricas eram feitos manualmente e apresentavam baixa assertividade. Isso deve-se à dificuldade de simular ou otimizar misturas manualmente usando-se planilhas de dados. Os lotes de produção existentes no estoque eram agrupados segundo suas principais propriedades e comparados com os pedidos. O planejador então simulava em longas planilhas de dados as misturas entre os lotes de produção e os perfis de produção das plantas. Não eram raros os casos em que este planejamento não era acontecia na prática e os pedidos dos clientes não eram atendidos, ou eram atendidos de uma foram diferentes daquela planejada.

As especificações dos clientes possuem, contudo, dezenas de propriedades as quais devem ser consideradas quando se está planejando a produção. A falta de definição dessas propriedades no momento da produção causa o risco de não se conseguir obter os lotes de produção adequados para a montagem dos pedidos. A tabela 4.2 demonstra algumas propriedades existentes nas especificações dos clientes que não eram contempladas no processo de planejamento das fabricas.

Variáveis contempladas no processo de planejamento de produção	Variáveis existentes em mais de 70% dos pedidos dos clientes
Bloom	Bloom
Viscosidade	Viscosidade
Transmitância	Transmitância
X	% Batch
X	GI (partículas sólidas)
X	Cinzas
X	PH
X	PI
X	Odor
X	Umidade
X	Restrições de plantas
X	Rastreamento de lotes

Tabela 4.2: Propriedades não contempladas no processo de planejamento de produção de produção sem o uso do APS

Fonte: Elaborado pelo autor

O tempo para realizar o planejamento descrito era demasiadamente longo. Segundo um mapeamento de fluxo de valor que fora feito no departamento de planejamento e controle de produção antes do APS, o processo de planejamento demorava de 2 a 3 dias para ser realizado. O tempo de processamento e a complexidade das análises fazia com que o planejamento fosse feito com uma frequência muito pequena. Dado o grau de incerteza da produção em atender completamente ao planejamento (que iremos analisar adiante), essa é uma frequência muito baixa. Como alega Falconi (2004), um processo deve ser controlado e revisto com frequência adequada para ser bem gerenciado. Caso o controle não aconteça, desvios podem acontecer e impactando indicadores importantes do negócio.

O processo de agrupamento manual das informações no planejamento de produção dificultava a definição dos pedidos seriam atendidos e quais não seriam. Isso porque as fábricas têm limitações em produzir certos tipos de gelatina em um determinado período. O resultado lógico disso era o atraso de pedidos que poderiam ser postergados ou produzidos com antecedência pelos gestores de produção e da demanda.

A colocação acima é mais fácil de ser compreendida com o uso de uma figura que expressa uma situação real de incapacidade de produção em função do “mix” de produtos. Na figura 4.7 existe, em um horizonte curto de tempo, o agrupamento de pedidos com valores altos de duas propriedades: *bloom* e viscosidade (duas principais propriedades das gelatinas). A capacidade de produção desse tipo de gelatina em um curto espaço de tempo é limitada e desta forma, sem a existência de um estoque estratégico desse tipo de gelatina, é impossível se atender a uma sequência de pedidos com altos valores de bloom e viscosidade como o exposto. Como são muitas as variáveis que se deve analisar para concluir que a tal sequência é impossível de ser atendida, muitas vezes um agrupamento desses era mantido até o momento de entrega da gelatina, causando atrasos.

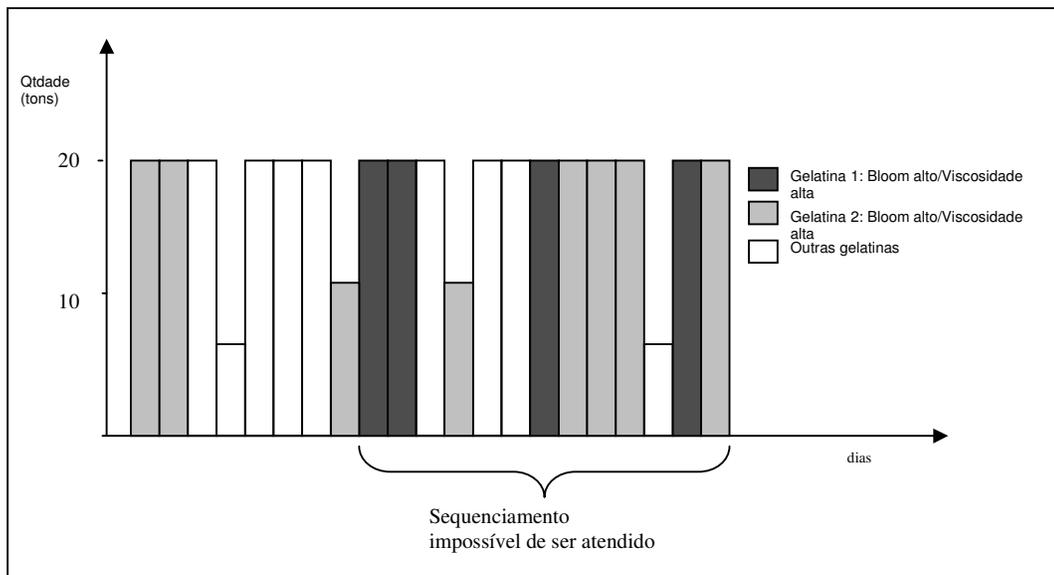


Figura 4.4: Representação real de restrições de *mix* de produção
Fonte: Elaborado pelo autor

Entretanto, o maior desconforto no modo com que se planejava a produção sem o uso de algoritmos de otimização estava na dificuldade de otimizar custos da produção. Existem diferentes custos de produção para diferentes perfis de gelatina. Por exemplo, é mais barato confeccionar gelatina com valores de transmitância baixos do que produzir gelatina com

valores de transmitância mais elevados. De mesma forma, é mais barato produzir gelatina com valores de *Bloom* (principal propriedade das gelatinas) menores do que com valores de *Bloom* maiores. Em um sistema manual de planejamento não se consegue, com tempo e recursos limitados, estabelecer qual é escolha de produção de menor custo. Abaixo resumimos os principais desconfortos e causas no processo de planejamento da produção sem o uso de algoritmos de otimização.

Processo	Problemas	Efeitos	Causa
Planejamento de produção	Falta do uso das propriedades importantes	Possível definição inadequada da gelatina para atender aos pedidos;	Falta de sistema para facilitar a simulação do planejamento;
			Falta de tempo recursos para realizar o processo
	Longo tempo de processamento do processo	Falta de frequência adequada para planejamento e consequente falta de controle da produção	Falta de sistema para facilitar o processo de planejamento.
			Falta de recursos para realizar o processo;
	Falta de sistemática para sinalizar a falta de capacidade para atendimento dos pedidos	Aceitação de pedidos que não poderão ser produzidos causando atrasos na entrega	- Falta de sistema de simulação de misturas;
	Dificuldade na análise de misturas entre gelatinas	Falta de acuracidade do planejamento em função do processo de programação de misturas;	Falta de sistema para facilitar o processo de planejamento;
		Possível definição de produção com maior custo;	
	Dificuldade na análise de custos de produção	Possível definição da produção com maiores custos	Falta de sistema de otimização e simulação de custos em diferente cenários;
Dificuldade na definição dos custos de produção;			

Tabela 4.3. Resumo dos problemas no processo de planejamento de produção sem o uso do APS

Fonte: Relatório de Diagnóstico, Gelita

Depois de planejada a produção, inicia-se a fabricação da gelatina. O processo de produção começa com o recebimento e corte da matéria prima. O couro bovino é recebido dos curtumes e em seguida cortado em tiras para início do processo de tratamento. O primeiro

processo por que passa o material é a pré-lavagem, onde são retirados fisicamente os materiais que não fazem parte da composição da gelatina (pedaços de carne, pêlos, pedaço de casco, entre outros).

Após essa etapa o couro é enviado para um tanque, denominado “calado”, onde recebe um agressivo tratamento químico preparatório para a retirada do extrato de gelatina do couro. A figura 4.2 ilustra esses processos.

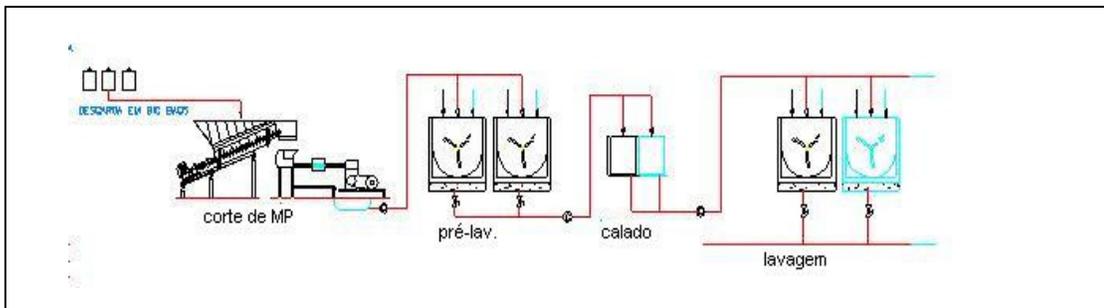


Figura 4.5: Representação dos processos iniciais de fabricação de gelatina
Fonte: Arquivos da Gelita

Depois desse intensivo tratamento, o couro vai para uma segunda lavagem onde é neutralizado e limpo para depois seguir para o processo de retirada do colágeno. O colágeno é extraído por um processo de hidrólise em que uma mistura de gelatina e água é aquecida para separar o material protéico de outras matérias do couro.

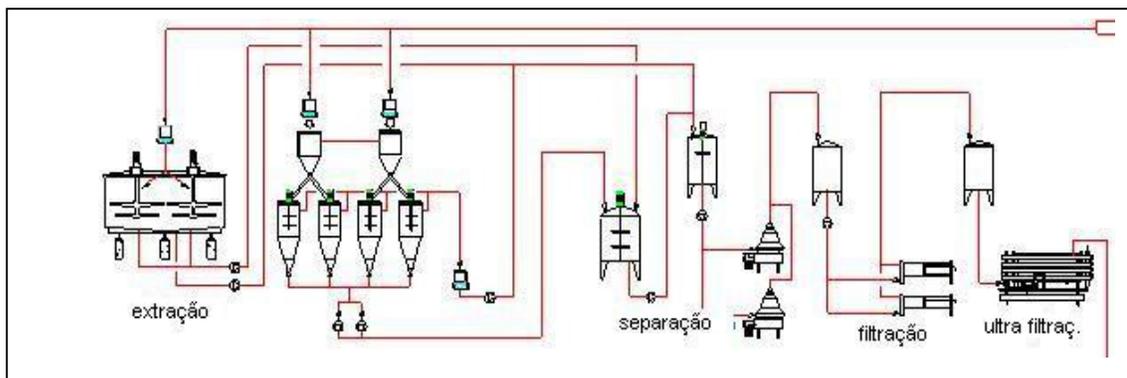


Figura 4.6: Representação das etapas finais de produção de gelatina: extração, separação e filtração
Fonte: Arquivos da Gelita

A gelatina, agora em uma forma líquida, passa por um processo de separação, filtração e deionização, para depois sofrer um processo de evaporação e tomar a forma de um caldo

concentrado de colágeno. A figura 4.3 ilustra essas etapas. Um processo de auto-concentração é então seguido por um processo de gelificação para, por fim, o produto passar por um processo de secagem onde finalmente tomará uma forma sólida.

Neste ponto, aproximadamente 15% do peso total de couro que fora limpo no início do processo permanece na produção. O último processo é a etapa de moagem e homogeneização, onde a gelatina toma a forma de grãos.

Um dos fatores mais importantes no processo de produção de gelatina pelas fábricas é a incerteza na determinação das propriedades do produto. As propriedades planejadas nem sempre são iguais as propriedades obtidas na produção. Neste sentido, o processo de produção de gelatina descrito acima é um processo estocástico do ponto de vista dos modelos matemáticos.

Essa incerteza pode ser visualizada no gráfico 4.1 que demonstra a realização e o planejamento da produção em uma amostra de 7 meses de produção. Os triângulos vermelhos mostram a quantidade e o valor de uma propriedade da gelatina que foi planejada. As barras do histograma demonstram o perfil de gelatina que foi realmente produzido.

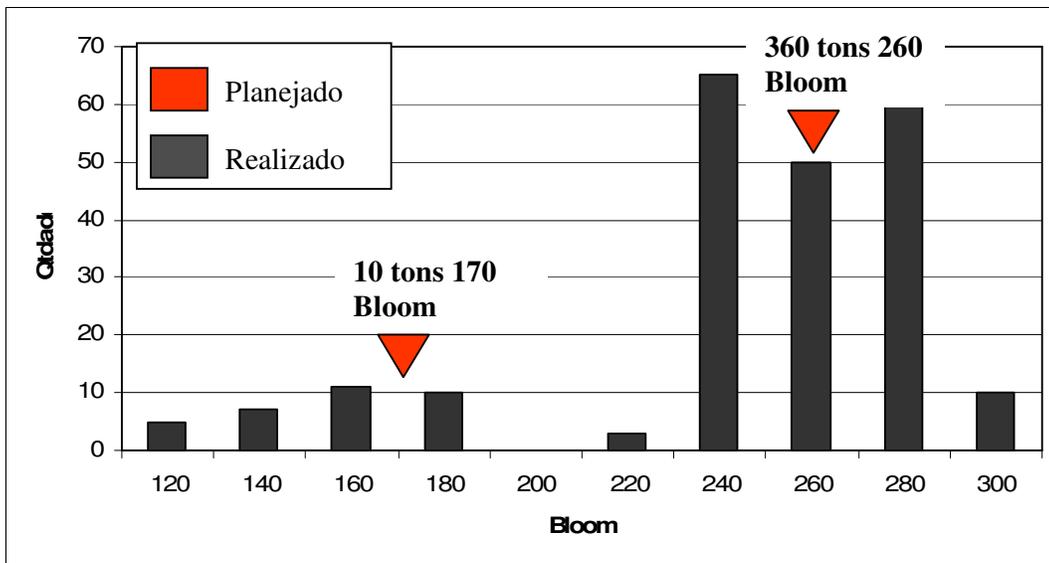


Gráfico 4.1: Planejamento x realização da produção de gelatina em uma planta da Gelita

Fonte: Relatório de diagnóstico, Gelita

4.1.3. Controle da produção

O processo de avaliação da aderência da produção dos lotes de produção em função do que fora definido pelo planejamento não acontecia com tanta frequência. Para realizar esse controle, os gestores do planejamento teriam que tentar misturar os diferentes lotes de produção e identificar se, na média, eles atendem o que fora solicitado. A figura 4.8 ilustra o processo.

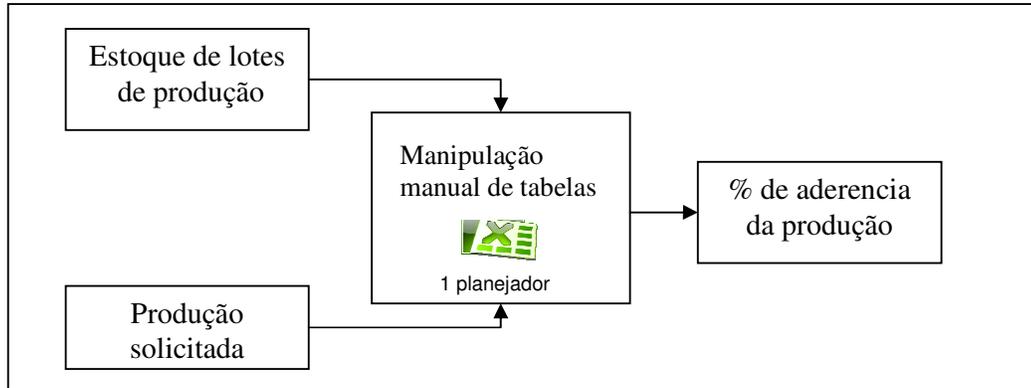


Figura 4.7: Esquema ilustrativo do processo de controle de produção da Gelita sem o uso do APS

Fonte: Elaborado pelo autor

Porém a complexidade em medir a eficácia das plantas em atender ao planejamento, o que acontecia em função do número muito elevado de propriedades a serem controladas e a impossibilidade de simular as misturas inibia que esse processo fosse realizado.

A falta do controle do planejamento dificulta o estabelecimento da real causa da falta de atendimento dos pedidos. Além disso, não se consegue saber porque a acuracidade da produção em realizar o planejamento está baixa ou alta. Durante o processo produtivo, usam-se vários parâmetros de produção tal como temperatura da água, concentração de insumo, quantidade de matéria prima, entre outros. Se não se mede a acuracidade da produção de forma sistêmica, a definição de qual o processo de produção mais adequado se torna mais difícil.

Como fizemos com os processos de programação de moagem abaixo consta o resumo dos problemas do controle da produção

Processo	Problemas	Efeitos	Causa
Controle da produção	Falta de sistemática para realizar o controle da produção	Falta de controle da acuracidade da produção	Falta de sistema para facilitar as misturas;
			Falta de metodologia para apurar a acuracidade da produção;
	Longo tempo de processamento do processo	- Desuso do controle de produção	Falta de sistema para facilitar as misturas;
			Processo pouco importante;

Tabela 4.4: Resumo dos problemas no processo de controle de produção sem o uso do APS

Fonte: Relatório de Diagnóstico Gelita

4.2. Modelagem e Interpretação do modelo

Arenales (2007) argumenta que o planejamento de processos de mistura sem o uso de otimização é uma tarefa complexa. O diagnóstico do planejamento e programação reitera tal afirmação. Como se pode intuir do que fora descrito nas sessões anteriores, o uso do método empírico não mostra ser suficiente para uma alta eficiência no planejamento e organização da produção de misturas.

Nesta sessão dissertaremos sobre o processo de aplicação do modelo de otimização linear para a superação dos problemas descritos. Ou seja, vamos para a segunda etapa de nosso processo de implementação do modelo de otimização linear.

Segundo Hillier e Lieberman (2010) os modelos de otimização linear somente são aplicáveis em problemas complexos com um número muito grande de variáveis e restrições que devem ser analisadas antes da tomada de decisão. Do que foi descrito nas sessões dedicadas ao diagnóstico, pode-se concluir que os problemas dos processos de planejamento e organização da produção da Gelita são complexos e necessitam de sistema para o suporte das decisões.

O objetivo final de um modelo de otimização é superar as dificuldades existentes na forma usual com que as decisões são tomadas, dado um problema real (TAHA, 2008). Desta forma, de acordo com os pressupostos acima e colocadas as causas dos problemas, o uso da otimização linear pareceu ser viável nos problemas de misturas encontrados no modelo real do planejamento e organização da produção.

Os quatro passos mencionados no capítulo 3 serão passados aqui para descrever o modelo de otimização linear realizado para embasar o planejamento e a organização da

produção na Gelita. Os modelos serão apresentados para os três ciclos do processo de planejamento descritos na sessão 4.2.

4.2.1 O Modelo de misturas

Os principais problemas do processo de programação antes da implementação do *APS* eram a não otimização do uso do estoque e o tempo de processamento da programação. Para solucionar tais problemas criou-se um *APS* de otimização na tentativa de se obter um número maior de pedidos com o estoque disponível dentro de um tempo menor.

Esse sistema confronta o estoque disponível de gelatina (lotes de produção) com os pedidos que se pretendia atender em um horizonte curto de planejamento para simular as misturas das ordens de moagem. Entretanto, para otimizar a alocação da gelatina nos pedidos corretos é necessário processar um número maior de ordens de moagem. Isso porque com um número maior de ordens de produção é possível atender um número maior de OMs é possível se misturar um número maior de pedidos. Ao se montar um pedido de cada vez é possível que se use lotes de produção que possibilitem a montagem de um número maior de Oms. Mesmo que essas ordens de moagem não sejam enviadas para o departamento de moagem, o estoque de lotes de produção está empenhado para atendê-las em um futuro próximo e não seriam utilizados em outros pedidos. Ou seja, é preciso aumentar o lote de processamento de 01 ou 02 OMs por processamento para aproximadamente 50 OMs, ou o equivalente a uma semana de moagem;

Imaginávamos que o uso de um algoritmo de otimização linear por si só diminuiria o tempo de processamento das OMs reduzindo os custos administrativos e também os riscos de parada dos moinhos por falta ou atraso nas ordens.

Com isso deu-se início à montagem de um *APS* para realizar a otimização da produção que foi intitulado *OBS (Optimized Blending System)* ou Sistema de Otimização de Misturas. A concepção geral do sistema é otimizar as misturas aumentando o horizonte de pedidos atendidos por programação e alocar de forma ótima os lotes de produção nas respectivas ordens de moagem. A figura 4.9 expressa a concepção geral do modelo produzido.

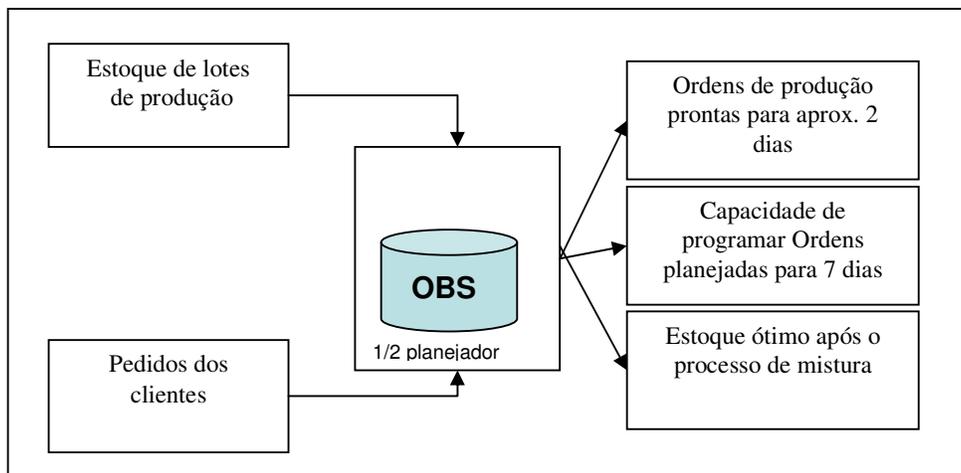


Figura 4.8: Esquema ilustrativo do OBS
Fonte: Elaborado pelo autor

Os passos sugeridos por Taha (2008) serão percorridos para descrever a modelagem do problema de programação de moagem.

Passo 1: Definição das variáveis do problema: As variáveis do modelo são as quantidades de cada lote de produção empenhadas em cada uma das ordens de moagem. Ou seja, cada lote de produção tem o potencial hipotético de ser usado em cada uma das OMs com o fim de configurar uma mistura cuja média ponderada atinja as propriedades requisitadas pelos clientes. Ou seja, variáveis do modelo são:

$$x_{ij}$$

tal que x é a alocação do Lote de Produção i na Ordem de Moagem j . Ou seja, o modelo define uma matriz de variáveis X i por j de variáveis

Passo 2: Definição da função objetivo: Segundo Arenales (2007) os vetores da função objetivo devem ser estabelecidos com o fim de se obter a maior rentabilidade dentro dos objetivos da empresa. Em nosso caso, a expectativa é atender o maior número de pedidos possível dada a quantidade de lotes de produção disponível. Portanto, a função objetivo estabelece um valor a ser maximizado que cresce proporcionalmente ao número de pedidos atendidos. Como muitas vezes o estoque de lotes de produção não é o suficiente para atender todos os pedidos, estabeleceu-se que a função objetivo atinge maiores valores com o atendimento dos pedidos com maior prioridade. Ou seja, o vetor custos da função objetivo

tem relação direta com a prioridade dos pedidos. Também estabelecemos os coeficientes que maximizam a qualidade da gelatina no estoque para os próximos períodos. Ou seja, a função é objeto é:

$$\text{Maximizar } \sum b_j x_{i,j} - \sum r_i x_{i,j}$$

tal que b_j é um coeficiente do vetor de custos B *l por j* que estabelece a prioridade dos pedidos j que devem ser atendidos;

e r_i é um é uma coeficiente do vetor de custos R *vetor l por i* que estabelece pesos para o estoque ruim de gelatina i .

Muitas vezes em função da má qualidade das matérias primas ou qualquer outro fator inesperado de produção, lotes de produção são produzidos com alguma propriedade não solicitada pelos clientes. Um exemplo dessas propriedades não requeridas são as gelatinas com odor muito alto ou com altos valores de GI (impurezas). A não utilização desses lotes de produção piora a qualidade do estoque, o que aumenta o custo financeiro de manutenção dos estoques e dificulta o trabalho de realizar a programação. Desta forma, estabelecemos que o estoque de baixa qualidade deveria ser diluído nas ordens de moagem sem restrições para essas propriedades, respeitando-se, evidentemente os limites das especificações dos clientes.

Passo 3: Definição das restrições: Uma parte das restrições do modelo é determinada pelas especificações dos clientes. Desta forma, os modelos de programação linear da Gelita possuem equações que determinam a média das propriedades resultantes da mistura dos lotes de produtos utilizados em um pedido. Essa média mantém-se entre os valores máximos e mínimos solicitados pelas especificações dos clientes, pois os valores limites das especificações são determinados como restrições do modelo. Ou seja, as restrições do modelo são dadas por:

$$\sum g_i x_{i,j} < qgmax_j$$

$$\sum g_i x_{i,j} > qgmin_j;$$

Tal que g = propriedade do lote de produção i ;

q = quantidade de lotes de produção alocadas na ordem de produção j ;

$gmax_j$ = valor máximo da propriedade g na ordem de produção j ;

$gmin_j$ = valor mínimo da propriedade g na ordem de produção j ;

Ainda existem as restrições de utilização. Isso porque em alguns casos os clientes não aceitavam que a gelatina não possuísse lotes rastreados, ou lotes que eram provenientes de uma certa planta. Por exemplo, a gelatina destinada para clientes judeus não podia ter proveniência de plantas que processavam gelatina de couro suíno. Tais imposições nos remetem a outros tipos de variáveis.

$$\sum h_i x_{i,j} = 0.$$

Tal que h indica uma propriedade do lote de produção i não aceita pelo cliente;

Por último, existem as restrições de utilização. Em função de não se conseguir uma homogeneização perfeita dos lotes de produção no momento da mistura nos moinhos de gelatina, os planejadores evitam utilizar lotes de produção com propriedades muito distantes daquelas requeridas pelas ordens de produção. Desta forma, o modelo demanda outras restrições:

$$\sum c x_{i,j}$$

Tal que c = binário cujo valor positivo é expresso por:

$$c(g_i) : g_i - (g_{max_j} - g_{min_j})/2 \leq d$$

onde d = distancia estipulada pelos planejadores;

Passo 4: Estabelecer condições implícitas: As condições implícitas do modelo de programação de moagem são de grande importância para o funcionamento do sistema. Suas condições estabelecem que não se pode obter resultados negativos de estoque no final do processo de moagem:

$$e_i - x_{i,j} \geq 0$$

Onde e_i = estoque total de lotes de produção j ;

Também estabelecem que a quantidade de gelatina empenhada nas ordens de produção devem ser sempre menores que quantidades solicitadas pelas mesmas:

$$\sum x_{i,j} \leq l_j;$$

Onde l é a quantidade de gelatina especificada pela ordem de produção j ;

E por último que os lotes de produção devem ser utilizados por completo nas ordens de produção. Isso se deve a uma grande dificuldade de se armazenar os *Big Bags* em que são armazenados os lotes de produção depois que uma fração destes foi utilizada. Ou seja, é necessário que se utilize as unidades de armazenamento como um todo. Desta forma:

$Q_{i,j} X_{i,j} =$ onde $q_{i,j}$ é uma variável binomial pertencente a matriz Q *i por j*

O modelo de misturas dessa forma pode ser expresso pela seguinte equação de otimização linear:

Maximizar $bqx_{i,j}$

Sujeito a:

$$\sum g_i x_{i,j} < qgmax_j;$$

$$\sum g_i x_{i,j} > qgmin_j;$$

$$\sum h_i x_{i,j} = 0;$$

$$\sum cx_{i,j}; \text{ tal } c \text{ é binomial}$$

$$E_i - x_{i,j} > 0;$$

$$\sum x_{i,j} \leq l_j;$$

$$qx_{i,j}, \text{ com } q = \text{variável binomial};$$

4.2.2. O Modelo de planejamento da produção

O modelo empírico de planejamento de produção causava grandes problemas no processo de planejamento de produção. Um deles era o fato de não se contemplar um número maior de propriedades no planejamento. Isso porque o planejamento era realizado manualmente, o que dificulta a análise de muitas variáveis. Pensávamos que esse problema seria ser superado com a utilização de um algoritmo que facilitasse o processo de análise de variáveis. O mesmo raciocínio foi feito com relação ao desconforto relacionado à dificuldade de analisar misturas e o longo tempo de processamento do planejamento. Ou seja, imaginava-se que tais problemas seriam superados com a construção de um modelo matemático de otimização que facilitasse a construção de cenários e contemplasse as misturas no processo bem como foi descrito no processo de otimização de misturas.

De forma semelhante, um sistema que simula o atendimento dos pedidos com o estoque de lotes de produção e o perfil de produção das plantas, poderia ser capaz de sinalizar

quais pedidos não conseguiriam ser atendidos e como consequência teriam que ser postergados ou antecipados. Ou seja, a problemática das restrições de mix de produção, também poderia ser superada com a simulação da produção com as previsões dos pedidos ou mesmo com os pedidos firmes, mas com um grau de antecipação bem maior.

Para superar o desconforto da dificuldade de analisar os custos de produção durante o processo de planejamento, pensamos em atribuir os custos reais dos diferentes tipos de produção que a planta pode realizar e inserir a minimização desses custos na função objetivo. Desta forma criamos um sistema de planejamento de produção intitulado GRP do Inglês *Gelatine Resource Planning*, ou Planejamento de Recursos de Gelatina. O nome faz referencia ao *MRP* descrito no capítulo 2.

A concepção geral do GRP é planejar de forma ótima os estoques de produção que deveriam ser produzidos em função do estoque de lotes de produção existentes, perfil de produção das plantas e da previsão dos pedidos ou pedidos colocados no sistema. A figura 4.9 expressa o esquema do planejamento de produção. A novidade esta no aspecto intertemporal do problema que será exposto mais adiante.

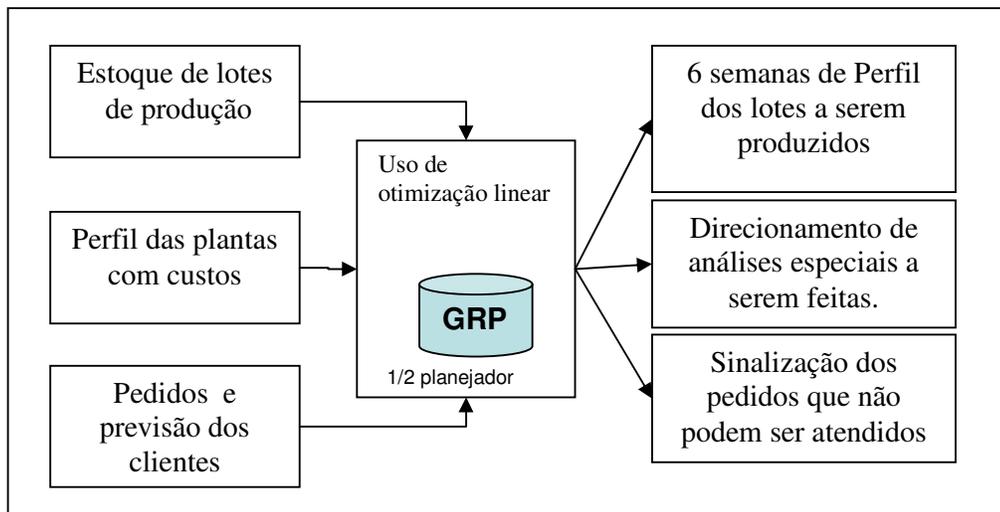


Figura 4.9: Esquema ilustrativo do GRP
Fonte: Elaborado pelo autor

Percorreremos os quatro passos de Taha (2008) para expor o modelo matemático do GRP:

Passo 1. Definição das variáveis. A maior diferença entre o modelo de programação linear aplicado no ponto de mistura daquele que fora estabelecido para auxiliar o planejamento da

produção está na presença dos perfis de gelatina produzidos pelas plantas e o problema da intertemporalidade. Na Gelita cada planta de fabricação produz perfis específicos de gelatina em cada um dos períodos de produção. As variáveis são determinadas pela alocação de cada um desses perfis de produção nas ordens de produção em cada período. Desta forma, a função objetivo é dada por;

$$x_{i,j,t}$$

Tal que x é uma variável pertencente à matriz X que define a alocação do perfil de produção ou lote de produção i no pedido j e no período t .

Passo 2: Definir a função objetivo: O objetivo é, tal como é feito no modelo de otimização de mistura, montar o maior número possível de pedidos com o menor uso de estoque possível dada uma certa prioridade de pedidos. Os pedidos mais próximos da data atual são os prioritários. Porém aqui adicionamos os custos ao perfil de produção das plantas, o que resulta na seguinte função objetivo:

$$\text{Maximizar: } \sum b x_{i,j,t} - \sum k x_{i,j,t}$$

Onde b representa coeficientes que pertencem ao vetor custo B $I \times i$ que estabelecem valores para a priorização dos pedidos e;

k representa o coeficiente que pertencem ao vetor custo K $I \times i$ que estabelecem valores para os custos de produção

A cada perfil de produção é atribuído um custo de produção distinto que é considerado na equação da função objetivo. Desta forma, a resolução do modelo que determina o perfil de produção que melhor atende aos pedidos do horizonte de planejamento considera o custo dos perfis de produção em sua escolha. A produção sugerida, então, é estabelecida pelo arranjo dos perfis de produção que melhor atende a demanda e tem o menor custo de produção.

Passo 3, Estabelecer as restrições: Tal como na otimização de misturas, aqui temos as **restrições** das propriedades das especificações dos clientes;

$$\sum g_i x_{i,j,t} \leq q g_{max,j,t}$$

$$\sum g_i x_{i,j,t} \geq q g_{min,j,t}$$

Tal que g_i = propriedade do lote de produção i ;

q = quantidade de lotes de produção alocadas no pedido j no período t ;
 $gmax_j$ = valor máximo da propriedade g no pedido j no período t
 $gmin_j$ = valor mínimo da propriedade g na ordem de produção j no período t

Da mesma forma, os ranges das propriedades dos perfis de produção são restritos:

$$\sum c x_{i,j} \leq 1$$

Tal que c = variável binário cujo valor positivo é expresso por:

$$C(g_i): g_i - (gmax_j - gmin_j)/2 \geq d$$

onde d = distância estipulada pelos planejadores;

Passo 4: Estabelecer condições implícitas:

Ao contrário do estoque de lotes de produção que tem como limite de uso a quantidade presente nos armazéns, os perfis de produção têm como limite de uso os limites de produção nos períodos. A capacidade de produção é dada, portanto de forma intertemporal. Cada período de produção tem um limite de produção que é determinado pela capacidade de produção das plantas. Ou seja:

$$\sum x_{l,j,t} \leq h_t$$

Onde h é capacidade de produção da planta no período t ;

Muitas vezes a planta apresenta capacidade restrita para um perfil de produção o que traz outra restrição:

$$x_{l,j,t} \leq f_t$$

onde $f \leq h$ e determina a capacidade de se produzir um determinado perfil de produção;

Entretanto, embora existam restrições para perfis de produção nos períodos, também existe a possibilidade de estocar gelatina de um período para o outro. Ou seja, o problema real nos diz que quando a demanda de gelatina é maior do que a capacidade de produção para um período, é possível produzir tal gelatina em um período anterior e armazená-la para que depois seja consumida. Essa lógica foi elaborada com referência em Carvalho e Haddad

(2003). A figura 4.10 ilustra as condições de capacidade e demanda em multi períodos para um perfil de produção;

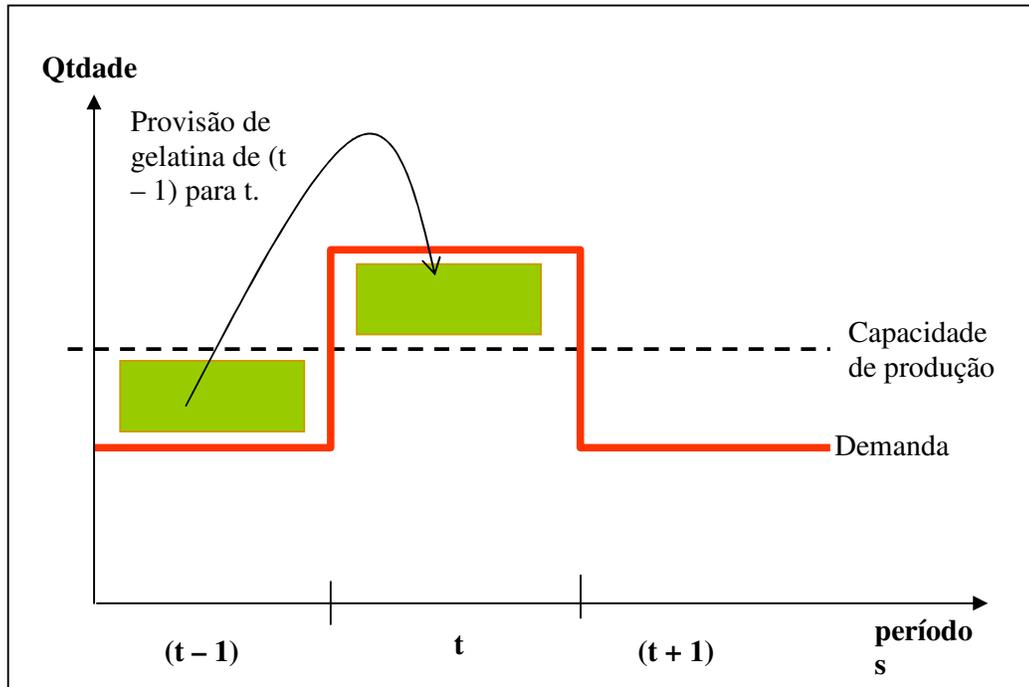


Figura 4.10: Representação do problema da capacidade de produção e a intertemporalidade

Fonte: Adaptado de Carvalho e Haddad, 2003

Na figura a demanda está abaixo da capacidade no período no período $t-1$. Entretanto, no período t acontece o inverso, a demanda para o perfil da produção é maior que a capacidade de produção. O modelo matemático deve ser capaz de provisionar gelatina do período $t-1$ para t e atender o máximo de pedidos possível.

Desta forma, temos que estabelecer um modelo que contemple o provisionamento de estoque entre os períodos, ou de uma forma mais simples, a transação de gelatina entre os períodos discriminados no horizonte de produção. Em um período t para um perfil de produção j da produção temos que fazer com que exista a possibilidade de transacionar estoques para haja um balanceamento entre demanda e produção. Chamamos essas equações de equações de balanço intertemporal (CARVALHO e HADDAD, 2003). De forma esquemática temos:

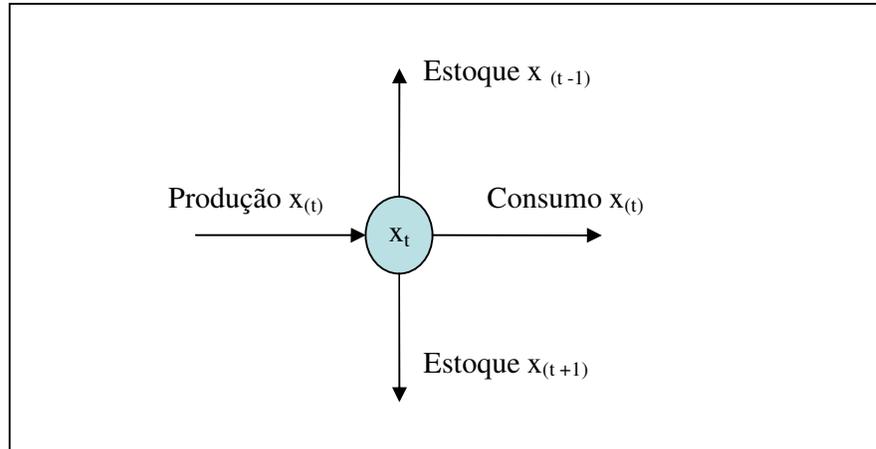


Figura 4.11: Representação gráfica da equação de balanço do sistema de planejamento da produção
Fonte: Adaptado de Carvalho e Haddad (2003)

Tal imposição estabelece a seguinte restrição:

$$p_{i,(k-1)} + y_{i,j,(k-1)} \leq x_{i,j,k} + y_{i,j,k}$$

onde p é quantidade produzida no período do perfil de produção i no período t , e $y_{j,t}$ é a quantidade de gelatina do perfil de produção i que será estocada no período t para o período $t+1$.

O modelo geral dessa forma é expresso por:

$$\text{Maximizar: } \sum b x_{i,j,t} - \sum k x_{i,j,t}$$

$$\text{Sujeito a: } \sum g_i x_{i,j,t} \leq qgmax_j$$

$$\sum g_i x_{i,j,t} \geq qgmin_j;$$

$$\sum c x_{i,j,t} \leq l_j;$$

$$\sum x_{i,j,k,t} \leq h_b;$$

$$X_{i,j,k,t} \leq f_k;$$

$$p_{i,(t-1)} + y_{i,j,(t-1)} \leq x_{i,j,t} + y_{i,j,t}$$

4.2.3. O modelo de controle da produção

Desenvolver um sistema de controle de produção com o uso de algoritmos de otimização é relativamente simples. A verificação do atendimento do plano de produção,

agora poderia ser feita com um sistema semelhante *OBS*. Os perfis solicitados de gelatina cumprem o mesmo papel dos pedidos no *OBS*. Tenta-se calcular as médias das propriedades dentro dos limites especificados. Agora, porém, os limites são estabelecidos pelos perfis de gelatina especificados pelos planejadores da produção. Intitulamos este sistema de “Taxa de Sucesso”.

O sistema tem como saída uma proporção do que se conseguiu produzir dentro do planejado. Como são conhecidas as propriedades solicitadas pelos planejadores e dados os lotes de produção, o sistema consegue estabelecer as causas do não atendimento do perfil solicitado conforme figura 4.12.

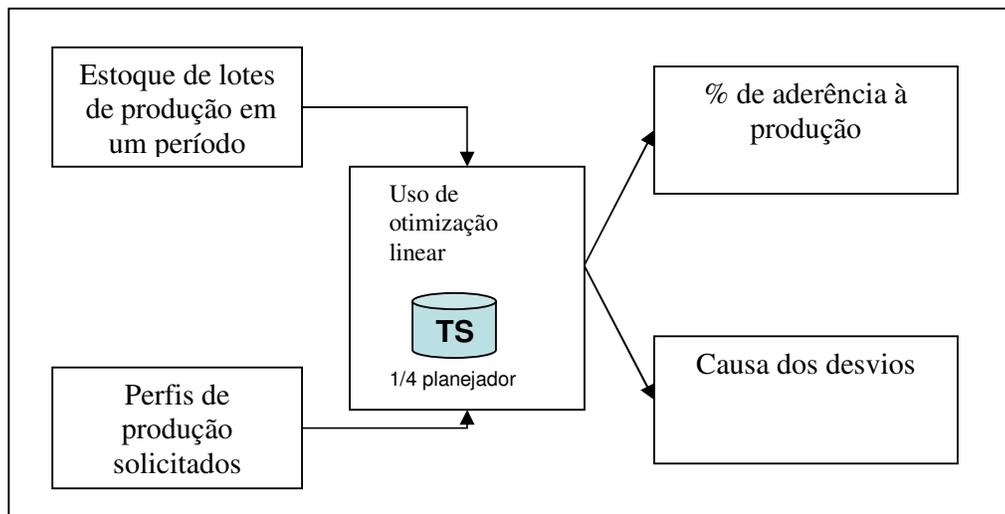


Figura 4.12: Esquema ilustrativo do sistema de controle da produção

Fonte: Elaborado pelo autor

Os passos 1 a 4 para a construção do modelo são semelhantes aos mesmos passos para o modelo do *OBS*. Com poucas exceções.

Passo 1: Definição das variáveis do problema: As variáveis do modelo são as quantidades de cada lote de produção empenhadas em cada uma dos perfis de produção solicitados;

$$x_{i,j}$$

tal que x é a alocação do lote de produção i perfil de produção j . Ou seja, o modelo define uma matriz de variáveis X i por j variáveis.

Passo 2: Definição da função objetivo: No caso do controle da produção, o vetor custo é igual a 1 para todos os perfis de produção e não existe benefício em não utilizar sempre os melhores lotes para atender os clientes; Desta forma temos:

$$\text{maximizar } \sum x_{i,j}$$

Ou seja, a função objetivo tenta maximizar a quantidade montada de perfis solicitados sem pesos para qualquer perfil;

Passo 3: Definição das restrições: Da mesma forma que no *OBS*:

$$\sum g_i x_{i,j} < qgmax_j$$

$$\sum g_i x_{i,j} > qgmin_j;$$

Tal que g_i = propriedade do lote de produção i ;

q = quantidade de lotes de produção alocados nos perfil de produção solicitado j ;

$gmax_j$ = valor máximo da propriedade g no perfil de produção solicitado j ;

$gmin_j$ – valor mínimo da propriedade g no perfil de produção solicitado j ;

Existem as restrições de ranges de uso de gelatinas em que:

$$\sum cx_{i,j} \geq li;$$

Tal que li é a quantidade em estoque dos lotes de produção e ;

c é uma variável binário cujo valor positivo é expresso por:

$$g_i - (gmax_j - gmin_j)/2 < t$$

onde t = distancia estipulada pelos planejadores;

A única restrição existente no modelo matemático de controle de produção que não existe no modelo do *OBS* é o uso de gelatina exclusivo de um determinado período para conseguir misturar um determinado perfil de produção. Isso se deve ao fato de que a produção deve ser direcionada para atingir um determinado perfil de produção em determinado período. Não se pode contar com o fato de que eventualmente o perfil de produção de uma determinada semana seja atingido usando-se o perfil de produção de outra semana. Isso seria sorte.

Desta forma, temos a restrição:

$$a_{i,j}x_{i,j} \geq li$$

Tal que a é um coeficiente binomial pertencente a matriz A que respeita a seguinte equação condicional $a = 1$ se o lote de produção i for direcionado para o perfil de produção j ; e li é a quantidade em estoque dos lotes de produção.

Muitas vezes essa restrição era ignorada para se verificar qual foi o atendimento ao planejamento de produção misturando-se todos os tipos lotes de produção. Isto é, contando com o fator sorte.

Passo 4: Estabelecer condições implícitas: As restrições implícitas são as que dizem respeito aos limites do estoque que pode ser utilizado para montar os perfis de produção solicitados;

$$e_j - x_{i,j} \geq 0$$

Onde e_j é estoque total dos lotes de produção j ;

Também se estabelece que a quantidade de gelatina empenhada nos perfis de produção deve ser sempre menores que quantidade solicitada pelos programadores:

$$\sum x_{i,j} \leq l_j;$$

Onde l_j é a quantidade de gelatina especificada para os perfis de produção solicitados j ;

O modelo de controle de produção pode ser expresso pela seguinte equação de otimização linear:

Maximizar $x_{i,j}$

Sujeito a:

$$\sum g_i x_{i,j} \leq qgmax_j;$$

$$\sum g_i x_{i,j} \geq qgmin_j;$$

$$\sum c_{x_{i,j}} = 1; \text{ com } c \text{ binomial}$$

$$e_i - x_{i,j} \geq 0;$$

$$qx_{i,j} \leq l_j \text{ com } q \text{ binomial};$$

4.3. Testes do modelo

A conclusão do modelo matemático possibilitou a realização de testes com um protótipo em pequena escala e testes em grande escala com o número de variáveis compatíveis com a realidade operacional da Gelita.

Nesta sessão iremos explorar tais simulações e análises do *OBS* até conseguir chegar ao modelo que fosse compatível com a dinâmica de operacionalização da gestão de produção da empresa. Como o *OBS*, ao contrário de muitos sistemas *APSs*, suporta um processo de tomadas de decisão diária no nível operacional do planejamento da produção da empresa, sua funcionalidade em termos de facilidade de uso, flexibilidade e rapidez é requisito chave para seu bom funcionamento. Os outros sistemas, seja pela menor frequência de uso, seja por se tratar de modelos de otimização não inteiros, não tiveram tanta necessidade de uma fase complexa de testes.

4.3.1. Teste em pequena escala

O primeiro protótipo do sistema *OBS* simulava a mistura de 3 ordens de moagem utilizando especificações de clientes com 3 propriedades apenas e um conjunto de lotes de produção compatíveis com as propriedades das ordens de produção, ou seja, o conjunto de ordens de moagem era atendido com os lotes de produção existentes.

Para simplificar o teste definimos que as unidades de armazenamento seriam iguais a 1 quando na realidade a gelatina nos estoques estão armazenados em *big bags* cujo peso se aproxima do peso de 1000 kg e deve ser utilizado integralmente. Ao mesmo tempo, não foi imposto qualquer vetor custo para a otimização da qualidade do estoque.

O modelo matemático testado é estabelecido segundo o seguinte modelo matemático:

Maximizar $b_{x_i,j}$

Restrito a:

$$E_j - x_j \geq 0$$

$$\sum x_i \leq Q_i;$$

$$\sum p_j x_j \leq p_i \text{ Max};$$

$$\sum p_j x_j \geq p_i \text{ Max};$$

Tal que p é usado para três propriedades: Bloom, Viscosidade e Transmitância;

As notações são aquelas utilizadas na determinação dos modelos matemáticos nas sessões anteriores. A figura 4.14 mostra o teste realizado com o Solver® da Excel Microsoft

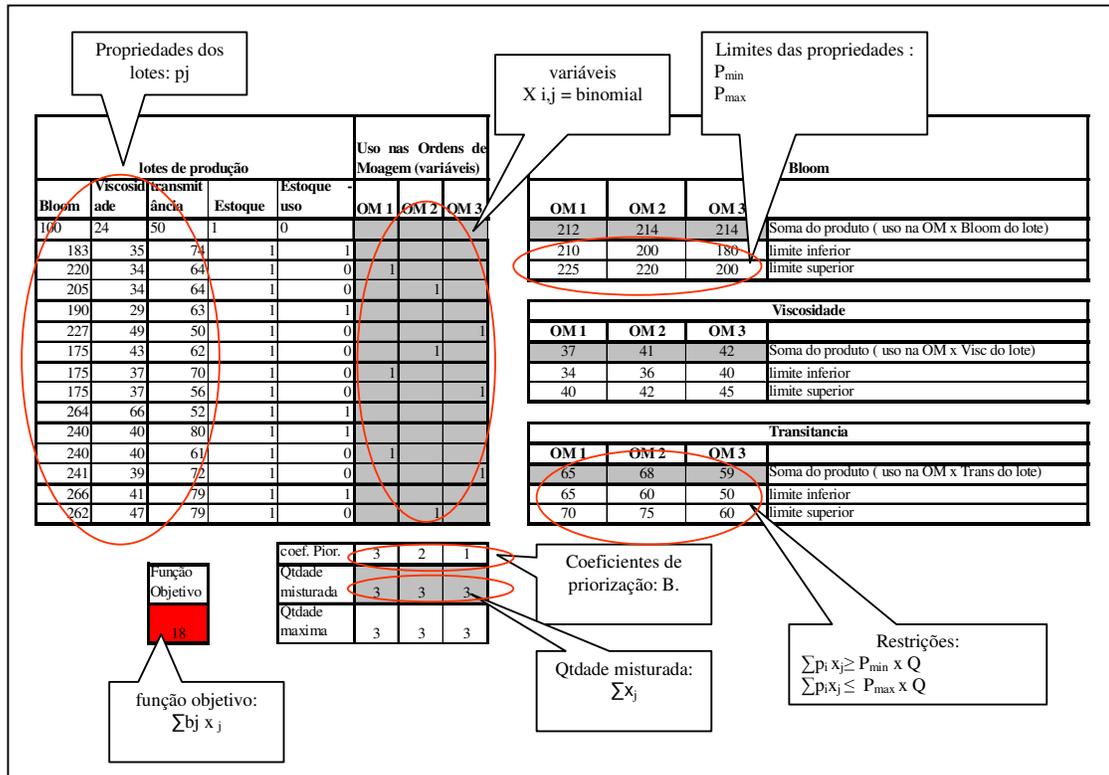


Figura 4.13: Protótipo do OBS com 3 propriedades de gelatina e 3 ordens de produção
Fonte: Arquivos da Gelita

O modelo realizou as misturas em menos de 01 segundo, o que possibilitou a realização dos testes em maior escala. Neste caso, o teste apenas possibilitou estabelecer confiança nas restrições, variáveis e função objetivo do modelo matemático. O aumento do número de OMs, lotes de produção disponíveis e principalmente a adição de todas as propriedades do modelo real poderiam aumentar em milhares de vezes o número de variáveis e restrições, o que aumentaria o tempo de processamento. Isso era importante ser testado.

4.3.2. Teste em grande escala

Os testes em grande escala, apesar de importantes, não eram conceitualmente complexos de serem realizados. Eles apenas representavam o modelo em pequena escala com a adição de um número maior de lotes de produção, ordens de moagem, propriedades e restrições que deveriam ser atendidas. Nesse sentido, foram incorporadas as quantidades reais

dos lotes de produção e o restante da função objetivo para a otimização da qualidade da gelatina dos estoques. Também adicionamos quantidades reais de pedidos presentes em um horizonte de 5 a 7 dias e todas as propriedades dos pedidos dos clientes. O modelo matemático testado foi aquele definido na sessão 4.4.1

$$\text{Maximizar } \sum b_j x_j - k_i x_i$$

Sujeito a:

$$E_i - e_i x_i \geq 0;$$

$$\sum x_{i,j} \leq l_j;$$

$$\sum g_i x_{i,j} \leq q_{gmaxj};$$

$$\sum g_i x_{i,j} \geq q_{gminj};$$

$$\sum h_i x_{i,j} = 0;$$

$$\sum c_i x_{i,j} = 1;$$

Com q sendo aplicado para 19 propriedades: Bloom, Viscosidade, Transmitância 450, % de Batch, transmitancia650, GI, odor, PI, PH, Perda, Sulfato, Umidade, Cinza, Visco 30 Rox, Tipo G, SPM, alto GI, Alto Odor, E/F. O modelo possuía 60.000 variáveis, 60.650 restrições (entre restrições binomiais e outras).

Os problemas nos testes foram avaliados e analisados, o que impediu que problemas acontecessem na operacionalização do sistema. Descreveremos os principais problemas referentes aos testes nos próximos parágrafos.

Problemas no vetor custo

Arenales (2007) afirma que as funções com múltiplos objetivos devem estabelecer uma troca adequada entre os objetos que se pretende otimizar.

A modificação da função objetivo (adição da condição de otimização do estoque) causou um impacto na resolução do problema. Identificamos que algumas ordens de produção não eram montadas, apesar de existirem lotes de produção que se misturados atingiam a especificação requerida. O problema estava no valor dos coeficientes que se estabeleceu para o vetor custo da função objetivo. O teste foi realizado com a seguinte função objetivo:

$$\sum b_j l_j x_j - \sum k_i l_j x_i$$

O vetor custo da função objetivo estava relacionado ao coeficiente b_j das prioridades das ordens de produção e os coeficiente negativo k_i do uso de gelatina de má qualidade, o que otimiza a qualidade do estoque restante. Entretanto, dependendo da grandeza de tais coeficientes, a função objetivo apresenta maior valor utilizando lotes de produção com baixa qualidade do que realmente completar os pedidos com as quantidades de produto solicitados pelos clientes.

Se $b_j < k_i$ então é preferível utilizar um lote com baixa qualidade em uma ordem de moagem do que um utilizar qualquer outro lote de produção que atenda a quantidade total da ordem de moagem. Desta forma estabelecemos que:

$b_j l_j x_j \gg \gg k_i l_j x_i$ ou seja $b_j \gg \gg k_i$ para qualquer alocação de lotes de produção no ordens de moagem

Desta forma, o modelo sempre prefere atender aos pedidos do que otimizar a qualidade do estoque o que faz com que todas as ordens possíveis sejam atendidas.

Problemas de tempo de processamento

Os 3 primeiros testes com dados reais foram processados em um tempo maior que 30 minutos e com grande variação de tempo entre eles, conforme é possível visualizar no gráfico 4.2. Esse tempo é demasiadamente longo para o processo operacional de programação de mistura. Isso se deve ao fato de ser necessário um número grande de simulações para se programar todas as ordens de moagem de um dia.

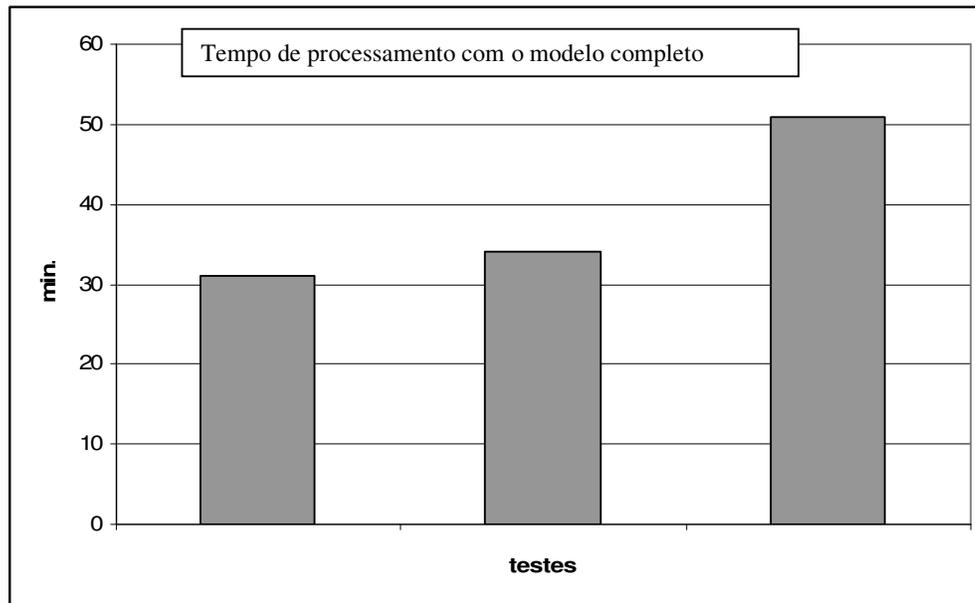


Gráfico 4.2: Tempo de processamento em 3 testes com modelo completo
Fonte: Arquivos Gelita

O problema com a função objetivo foi fácil de ser constatado e eliminado. Entretanto, para solucionar o problema do tempo e processamento referente ao modelo de otimização com variáveis inteiras foram necessários outros testes.

4.3.2.1. Teste com otimização do saldo em estoque

O tempo de processamento estava tornando a implementação do APS inviável. Desta forma, tentou-se retirar a programação inteira do modelo e minimizar a diferença entre os lotes de produção que foram alocados nas ordens de produção e o resíduo do estoque. Tentou-se assim fazer com que a função objetivo atingisse maiores valores com o uso de lotes inteiros transformando a função objetivo em:

$$\sum b_j l_j x_j - \sum k_i x_i - \sum (l_j - x_j)$$

O resultado do teste mostrou que embora muitos lotes fossem usados de maneira inteira, o modelo usava alguns poucos lotes de produção de forma fracionada, o que era incompatível com a operacionalização do gerenciamento do estoque que impossibilitava a

armazenagem dos *big bags* fracionados. O teste não deu certo e então tivemos que utilizar alternativas para reduzir o tempo de processamento;

4.3.2.2. Teste com alternativas de processamento de programação inteira

O modelo de programação inteira era necessário para resolução do OBS e, portanto, teríamos que encontrar alguma forma de reduzir o tempo de processamento sem a eliminação da produção inteira do modelo. O *software* que utilizamos, *Express Solver Engine da Frontline System* apresenta uma série de opções de processamento para solução de problemas de otimização linear. Entre os parâmetros utilizados para reduzir o tempo de processamento dois apresentaram grande êxito:

Tolerância de soluções inteiras

Como foi explorado no capítulo 2, o método *Branch and Cut* de resolução de problemas com restrições inteiras primeiro soluciona o problema de forma relaxada, isso é, ignorando as restrições inteiras para depois testar as diversas soluções inteiras existentes dentro da zona de respostas factíveis. O modelo matemático utilizado pelo Express Solver Engine® guarda o valor da função objetivo da resolução do modelo relaxado e o compara com os valores da resolução do problema com as soluções inteiras. Ao mesmo tempo, o sistema disponibiliza uma caixa de texto onde se pode estabelecer qual é a diferença percentual aceitável entre a resolução do modelo relaxada e resolução do modelo com as restrições inteiras usando a seguinte equação:

$$Vfo \text{ com restrições inteiras} - Vfo \text{ sem restrições inteiras} / Vfo \text{ sem restrições inteiras}$$

Onde *Vfo* é o valor da função objetivo para o problema.

Ao percorrer a árvore do método *Branch and cut*, uma vez que o modelo encontre uma solução inteira dentro da porcentagem estabelecida pelo usuário, apresenta-se a resolução e o sistema pára de realizar outros testes com soluções inteiras. Desta forma estabelecemos que se uma solução com uma diferença de até 5% da resolução do modelo relaxada fosse encontrada, o sistema poderia deixar de realizar outras interações.

Os tempos de resolução caíram drasticamente. Foram feitos mais de 20 testes com conjuntos de estoques e ordens de produção distintas e com aproximadamente o mesmo

número de variáveis: entre 25 e 35 ordens de produção e 650 e 780 lotes linhas de estoque o que equivale a 21300 variáveis binomiais. Nesse conjunto de testes se obteve uma média de tempo de 6,85 minutos e um desvio padrão de 2,1 minutos. Ou seja, não somente o tempo de processamento diminuiu como o desvio também decresceu significativamente. Ademais, utilizou-se esse parâmetro nos 3 testes de grande escala que foram realizados. O resultado pode ser visto no gráfico 4.3.

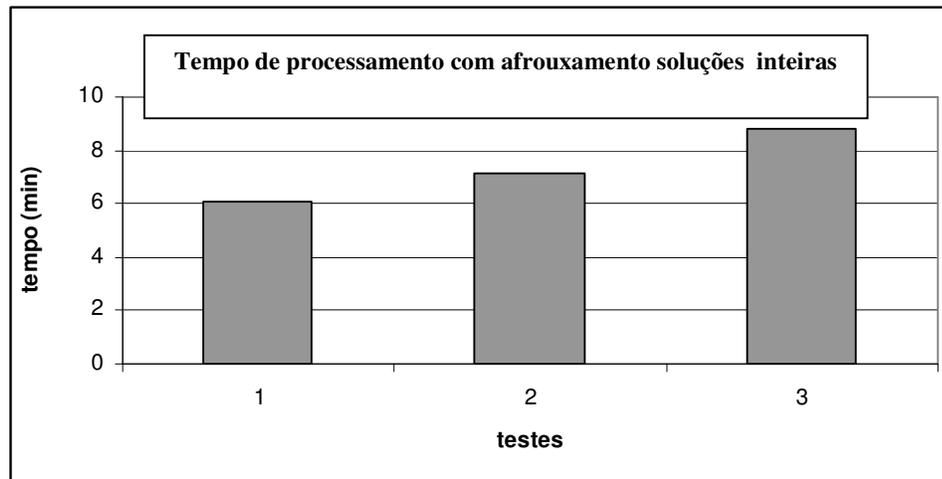


Gráfico 4.3: Tempo de processamento de 3 testes com restrição da tolerância de soluções inteiras
Fonte: Arquivos da Gelita

Embora se soubesse que a resposta obtida com essa alteração não seria a melhor, como estávamos lidando de um programa que serve para controlar a programação de moagem de uma fábrica inteira, tínhamos que conjugar performance e esforço de processamento. Ou em outras palavras, estabelecer uma boa troca (*trade off*) entre o quão perto da melhor resposta se consegue chegar em um razoável tempo de processamento. Contudo, os tempos ainda eram altos e outras soluções para diminuir o tempo de processamento ainda eram necessárias.

Restrição de adição de nós

O método *branch and cut* estabelece cortes de respostas (nódulo) na área de respostas factíveis que obviamente não são melhores que a resposta atual O Express Solver Engine® além de realizar os cortes tradicionais, possibilita a definição dos valores incrementais da função objetivo que se deseja analisar através de um caixa de texto. Ou seja, o sistema

possibilita ajustar o quanto o valor da função objetivo referente ao próximo nóculo deve ser maior do que a melhor resposta atual para que seja analisado. Desta feita, se um nóculo que atenda as restrições inteiras estabelecer um valor para a função objetivo cujo valor percentual é menor do que aquele na caixa de texto, então o nóculo e o sub segmento de respostas relacionadas a ele na árvore de resolução do método de *branch and cut* não será analisado.

Estabelecemos que o sistema somente analisaria nós que apresentassem um valor incremental maior que 1% que a atual melhor resposta. Tal parâmetro possibilitou uma redução significativa do tempo de processamento. Mais de 20 testes foram feitos com conjuntos semelhantes àqueles que foram utilizados para analisar a tolerância das soluções inteiras e os tempos nunca ultrapassaram os 6 minutos.

Redução do número de variáveis inteiras

Embora tenhamos reduzido o tempo de processamento significativamente, ainda existia um desconforto com o tempo que o sistema demorava para realizar algumas simulações das misturas quando se processava um conjunto grande de lotes de produção e/ou ordens de moagem.

A última solução que se aplicou para a redução do tempo de processamento não está relacionada às opções de processamento do Express Solver Engine; mas sim na própria dinâmica de execução da programação de moagem. Embora os programadores processassem o OBS com a presença de várias ordens de produção, a liberação dessas, era feita de forma individual. Ou seja, analisava-se cautelosamente os lotes de produção empenhados nas ordens de moagem pelo sistema antes de enviá-las para o armazém onde os lotes seriam segregados e colocados nos moinhos. Uma vez que uma ordem de produção já tivesse sido analisada e aprovada, os lotes de produção empenhados nessas ordens aprovadas poderiam deixar de ser variáveis fazendo com que o número de variáveis e restrições no modelo fosse reduzido.

A tabela 4.5 mostra a matriz de variáveis binomiais $x_{i,j}$ com i lotes de produção e j ordens de moagem do sistema *OBS*. Como a análise e aprovação da alocação dos lotes de produção nas ordens de moagem, tais linhas e colunas podem deixar de ser variáveis fazendo com que o problema fique menor e, como consequência, reduzindo o tempo de processamento.

Ordens de moagem excluídas

Lotes de produção excluídos

ordens de moagem

	OM 1	OM 2	OM 3	OM 4	OM 5	OM 6	OM 7	OM 8	OM 9	OM 10	OM 11	OM 12	OM 13	OM 14	OM 15	OM 16	OM 17	OM 18
lote 1																		
lote 2																		
lote 3																		
lote 4																		
lote 5																		
lote 6																		
lote 7																		
lote 8																		
lote 9																		
lote 10																		
lote 11																		
lote 12																		
lote 13																		
lote 14																		
lote 15																		
lote 16																		
lote 17																		
lote 18																		
lote 19																		
lote 20																		
lote 21																		

Tabela 4.5: Representação da exclusão de ordens de moagem e pedido da base de variáveis

Fonte: elaborado pelo autor

Após estas 3 soluções o tempo de processamento do *OBS* ficou adequado ao tempo requerido para a operacionalização da programação de ordens de moagem e pôde ser implementado. A performance da busca de melhor solução foi reduzida e uma boa condição de otimização versus tempo de processamento foi atingida.

4.4. A Implementação do APS de otimização

Após os testes, logo se tentou utilizar os modelos de *OBS*, *GRP* e *TS* para suportar as decisões nas operações da Gelita e alguns problemas relacionados à falta de automatização e customização surgiram.

O processo de programação de misturas na Gelita é bastante dinâmico. Continuamente são liberados resultados laboratoriais dos lotes de produção que podem ser utilizados nas misturas. Ao mesmo tempo, o sequenciamento de programação pode sofrer alterações em função de alterações nas prioridades dos pedidos ou ainda mudança na programação de embarques dos produtos finais.

Tal dinamismo estabelece a necessidade de uma flexibilidade muito grande no manuseio das alocações dos lotes de produção. Desta forma, criou-se um programa que possibilitava que os programadores pudessem, com grande facilidade, retirar e colocar lotes de produção das ordens de moagem recalculando a média da mistura automaticamente.

A figura 4.14 apresenta a estrutura de interface com o usuário em que as manipulação das alocações dos lotes de produção nas OMs eram feitas. Com uso dessa tela, os programadores conseguiam alocar lotes de produção nas OMs, retirar lotes de produção das OMs e segregar categorias de gelatina para atender um determinado pedido específico. A especificação do processo foi atendida.

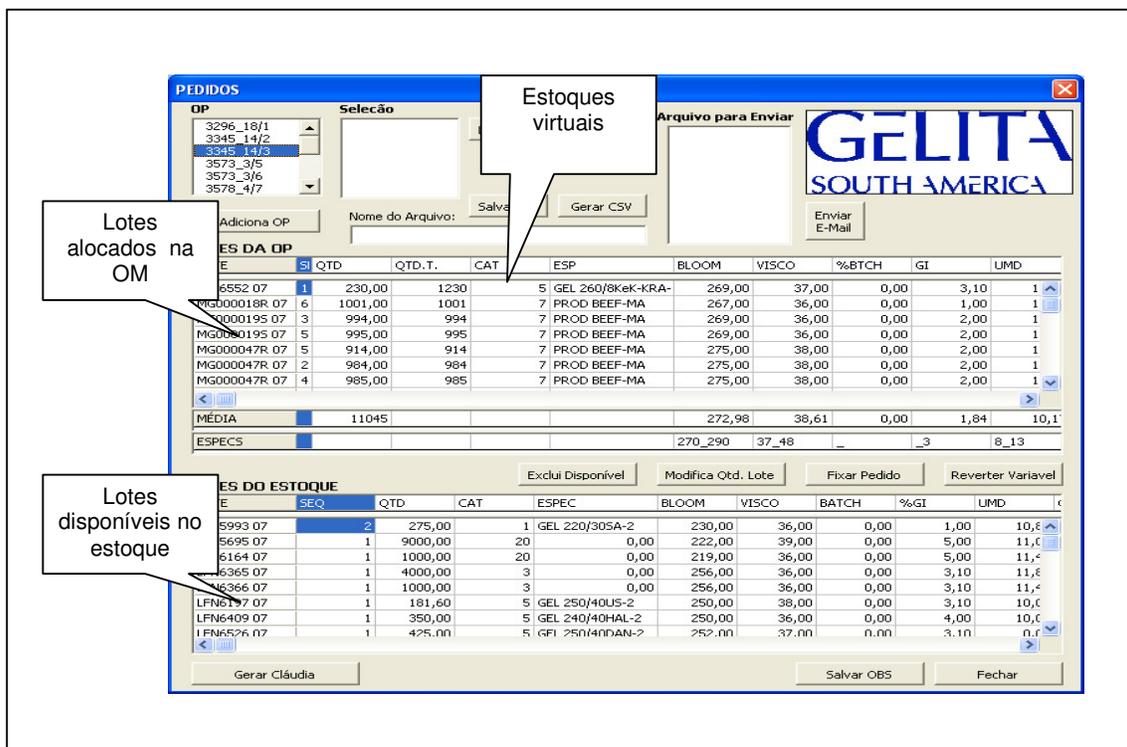


Figura 4.14: Tela de customizações para flexibilização do sistema OBS
Fonte: OBS, Gelita

Alem das customizações mencionadas acima, os programadores e planejadores solicitaram alta flexibilidade para o manuseio das propriedades das ordens. Esse requerimento era necessário porque muitas vezes se simulavam especificações mais apertadas do que o eram realmente. Isso acontecia em função de certa variação no processo de mistura que eventualmente fazia com que a média ponderada das misturas dos lotes de produção não fosse atingida quando se media as propriedades dos lotes dos produtos finais.

Demais disso, como é argumentado por Hillier e Lieberman (2010), os modelos de otimização linear sempre irão utilizar o limite dos recursos escassos de um problema e, desta forma, se o estoque de lotes de produção estiver com gelatina com Bloom baixo, por exemplo, todas as ordens de produção vão ser misturadas com um valor muito baixo ou até mesmo com o valor do limite mínimo de Bloom das especificações. Porém isso é um grande risco em função da variação natural que pode acontecer entre as propriedades calculadas e as propriedades medidas nos lotes dos produtos finais.

A figura 4.15 mostra a estrutura da interface com o usuário da manipulação das propriedades dos pedidos.

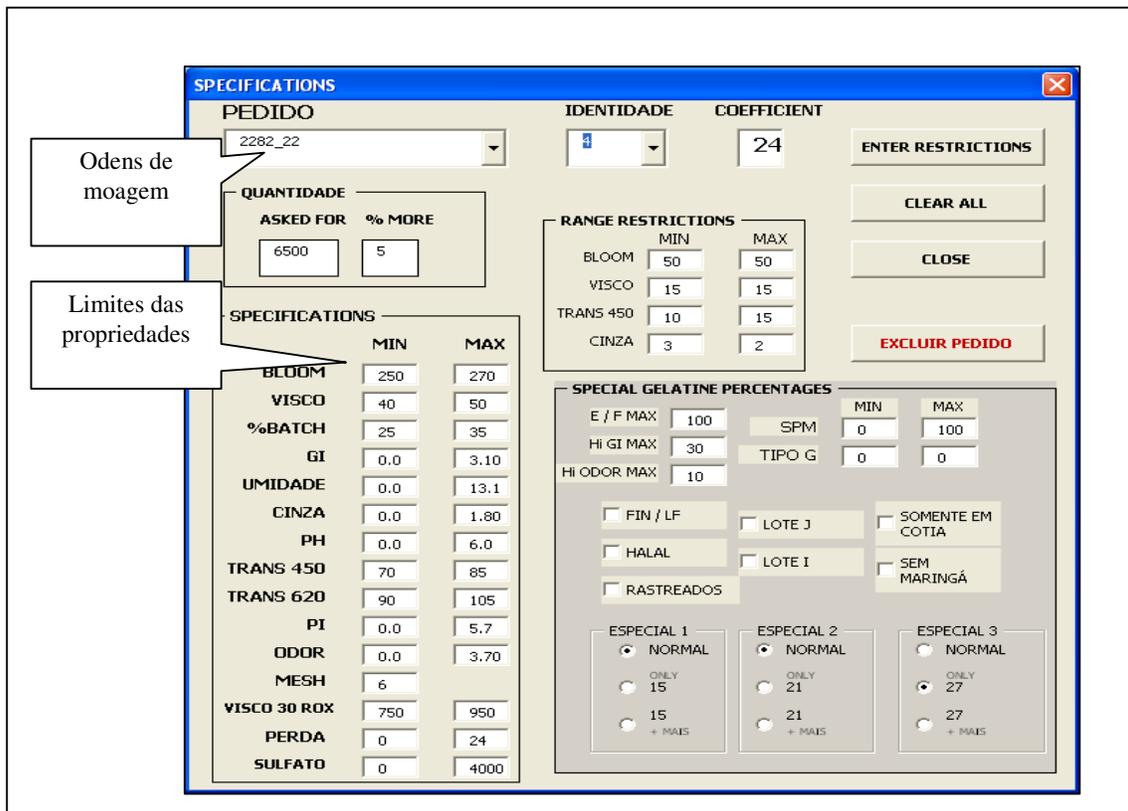


Figura 4.15: Interface com usuário do OBS para manipulação das propriedades das ordens de moagem

Fonte: OBS, Gelita

O sistema GRP também demandou algumas customizações. A capacidade das plantas no modelo real poderia ser alterada em função de manutenções nas instalações ou testes de produção de novas gelatinas, ao mesmo tempo em que os perfis de produção que as plantas

eram capazes de produzir poderiam se alterados em diferentes períodos em função da qualidade das matérias prima e outros fatos.

Desta forma, embora a capacidade de produção e os perfis de gelatina pudessem ser alterados no modelo, era necessário que se criasse uma interface simples de simular diferentes cenários em que se conseguisse visualizar com muita clareza quais os pedidos seriam atendidos e quais os pedidos deixariam de ser atendidos. O formulário de simulação do *GRP* pode ser visualizado na figura 4.16.

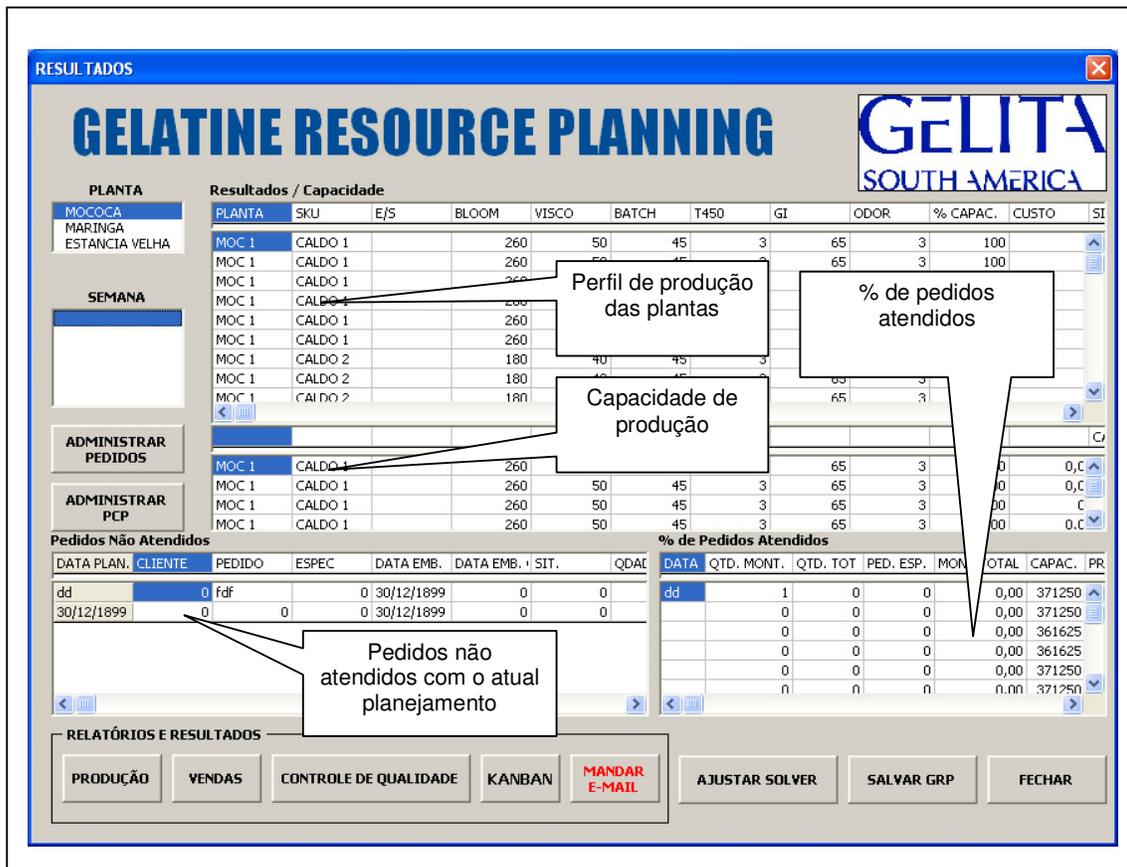


Figura 4.16: Interface com o usuário do sistema GRP e suas principais funcionalidades;
Fonte GRP, Gelita

Embora não se tenha realizado um projeto de implementação de *S&OP* (*Sales and operational Planning*), pode-se afirmar que a *GRP* ajudou muito na sincronia entre a demanda e a oferta de bens na Gelita. Garwood (2002 p.1) diz que o *S&OP* é uma estratégia utilizada para suprir a lacuna entre as vendas e marketing e a produção. Ou seja, que esse processo ajuda no balanceamento entre a cadeia de suprimentos e a demanda de bens da empresa. Correa e Gianesi (2001) argumentam que o *S&OP* auxilia no fortalecimento do

nível de serviço das empresas intensificando a ligação entre o planejamento estratégico e o plano operacional. Esses autores afirmam que o *S&OP* serve para evidenciar e demonstrar as flutuações no mercado e como a operação deve responder a isso.

Como iremos demonstrar adiante, a visualização da simulação simples do atendimento dos pedidos em função da capacidade de produção e perfil de produção das plantas permitiu a sincronização entre os pedidos e a programação de vendas diminuindo a incidência de pedidos que não poderiam ser atendidos em função da capacidade de mix de produção das plantas. De certa forma, o *GRP* se tornou uma ferramenta fundamental para estabelecer um fluxo eficiente de informações entre o planejamento de produção e o gerenciamento de demanda fortalecendo o nível de serviço da Gelita e o planejamento integrado de operações e vendas como impõe o *S&OP*.

As customizações do sistema de Taxa de Sucesso se restringiram a mostrar de forma gráfica a taxa de atendimento ao planejamento das plantas e o porquê do não atendimento. Segundo Falconi (2004), a análise das causas dos problemas precede a realização de ações para eliminar ou minimizar o problema nos processos. Desta forma, confeccionamos a interface do gráfico 4.4, que representa a taxa de sucesso no atendimento ao planejamento de produção das plantas.

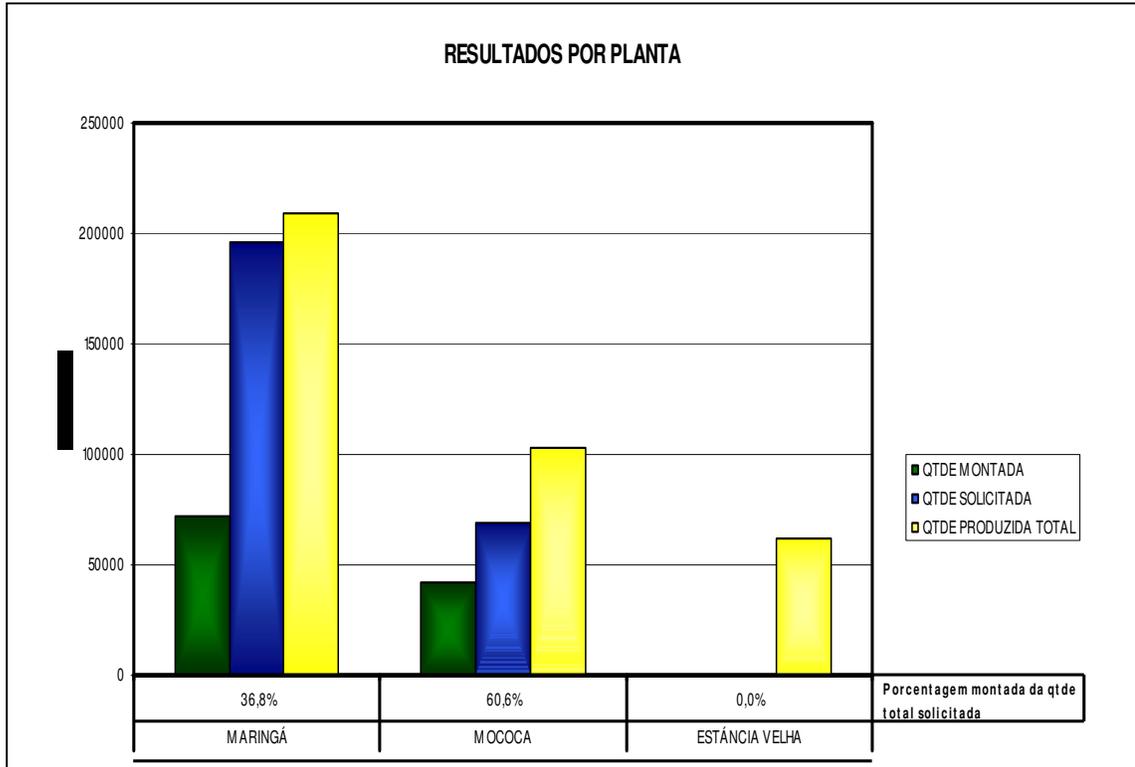


Gráfico 4.4: Interface do atendimento por planta do programa Taxa de Sucesso
Fonte: Taxa de Sucesso, Gelita

Também se elaborou uma interface gráfica para que fosse possível visualizar dentro da programação para as plantas como se comportou o atendimento de cada um dos perfis de produção durante um determinado período.

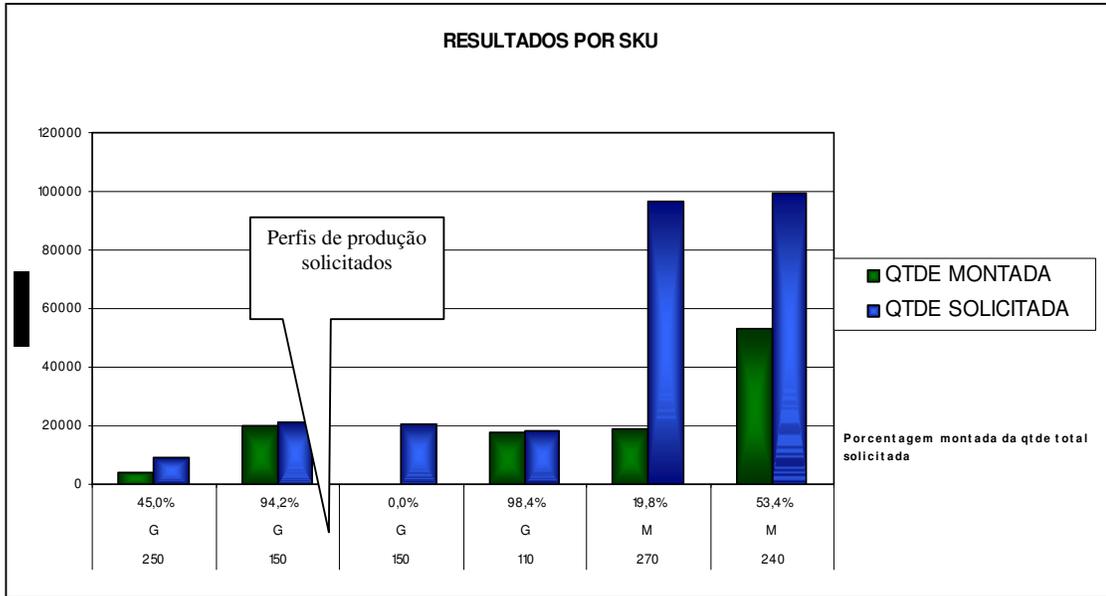


Gráfico 4.5: Interface do atendimento por perfil de produção do programa Taxa de Sucesso

Fonte: Taxa de Sucesso, Gelita

Por último, uma interface em tabelas para se identificar quais lotes de produção causaram a falta de aderência entre a produção e o planejamento foi desenvolvida. Tal interface permitiria aos gestores a análise das propriedades das gelatinas produzidas e sua comparação com o que fora planejado. O TS auxiliou a realização dos ciclos de melhoria baseados no PDCA (*Plan Do Check and Analyze*) de Falconi (2004). Essa interface pode ser visualizada na tabela 4.6.

MOCOCA - LOTES NAO USADOS												
Produção	M. Prima R/S ou I	Sigla	Lote	Peso Kg	Data Fabricação	Bloom g	Visco cP	T 450nm %	T 620nm %	Odor 6,67%	GI	Cinzas %
Standard	R	GJ	6119	1904	23/7/2007	133	29	64	90	3	2	0,7
Standard	R	GJ	6120	1970	23/7/2007	128	32	62	88	3	2	0,9
Standard	R	GV	6121	1520	23/7/2007	185	39	64	87	3,3	2	1,1
Standard	R	GV	6122	700	23/7/2007	204	39	59	83	3,7	2	1,6
Standard	S	GV	6123	724	23/7/2007	164	38	71	92	3,7	2	0,9
Standard	R	GJ	6124	1710	23/7/2007	202	46	64	89	4	2	1,7
Standard	R	GJ	6125	520	23/7/2007	226	49	64	89	3,7	2	1,2
Standard	R	GV	6126	480	23/7/2007	248	45	65	89	4	2	0,9
Standard	S	GV	6127	1710	23/7/2007	273	46	72	92	3	2	0,4
Standard	S	GV	6128	800	24/7/2007	254	46	67	89	3,7	2	0,6
Standard	S	GV	6135	865	24/7/2007	181	33	68	91	4	1	0,8
Standard	R	GV	6141	1933	25/7/2007	89	20	55	87	4	2	1,5
Standard	R	GV	6142	1700	25/7/2007	82	21	55	88	4	2	1,4
Standard	R	GV	6143	1790	25/7/2007	117	31	54	81	3,3	2	1
Standard	R	GV	6144	1160	25/7/2007	132	31	54	81	3,3	1	0,7
Standard	R	GV	6145	760	25/7/2007	143	32	58	83	3,7	1	0,9
Standard	R	GV	6146	1020	25/7/2007	157	29	50	85	3,7	1	1,1
Standard	R	GV	6147	1560	25/7/2007	156	27	52	85	3,3	1	0,5
Standard	S	GV	6148	1690	26/7/2007	130	26	58	83	3	2	0,2
Standard	S	GV	6149	825	26/7/2007	155	25	60	82	3,7	1	0,3
Standard	S	GG	6150	550	26/7/2007	154	22	59	81	4	2	0,6
Standard	S	GV	6151	1680	26/7/2007	153	29	51	81	3,3	2	1
Standard	S	GV	6152	375	26/7/2007	118	33	50	81	3,3	2	1,3
Standard	S	GV	6155	1920	26/7/2007	161	29	66	90	3	2	0,8
Standard	S	GV	6157	1950	26/7/2007	151	26	67	91	3,7	2	1
Standard	S	GV	6158	820	26/7/2007	154	28	63	89	3,7	2	1,4

fora de todas as specs

dentro mais do que uma espec

Lotes de produção para atender um perfil de produção

Tabela 4.6: Interface em tabelas para a identificação dos lotes que fizeram com que a taxa de sucesso não fosse atingida
Fonte: Taxa de sucesso, Gelita

Ao contrário do que se fazia anteriormente, os sistemas de otimização implementados fazem otimizações considerando um cenário abrangente e complexo de lotes de produção, pedidos e perfil de produção. As alocações dos pedidos com o uso dos otimizadores não atendem os objetivos pontuais e unidimensionais de outrora o que pode causar falta de confiança ou descrédito nos sistemas.

O processo de mudança na forma de gerenciamento do planejamento da produção não foi simples. O grau de confiança que as pessoas tinham nas informações processadas pelo APS não era nada alta. Isso porque o processo de analisar as respostas de um algoritmos de otimização que processa centenas de milhares de variáveis não é rápido nem tampouco fácil. O sistema demandava uma perfeita sincronia entre a análise do processamento do APS e a manipulação das OMs para que realmente causasse impacto na performance do planejamento.

Foram realizadas mais de 40 horas de treinamento para os programadores e planejadores da produção para que utilizassem os sistemas com destreza e com o objetivo de auxiliá-los nos processos operacionais. Como resultado dos treinamentos, identificou-se que as habilidades manuais de seleção e identificação de lotes compatíveis com as ordens de produção deixaram de ser importantes e foram substituídas por habilidades analíticas de

compreensão das causas das alocações que os sistemas de otimização realizaram. O perfil das pessoas que gerenciavam as misturas e o planejamento da produção foi completamente alterado. A destreza em manipular dados em tabelas foi substituída pela necessidade de realizar análises complexas de otimização envolvendo grande número de clientes, ordens de produção e custos. Embora esta mudança de perfil pareça ser facilmente gerenciável, o descompasso entre a forma com que o sistema opera e o antigo método de manipulação dos dados para se realizar o planejamento causou, em um primeiro momento, dificuldade das pessoas em operar o sistema e pontuais danos na qualidade e custo dos produtos produzidos.

4.5 Os Ganhos com o uso da otimização linear

De uma forma geral, o planejamento e o controle de produção da Gelita deixam de ser feitos pelo agrupamento de lotes com propriedades semelhantes usando a experiência e o empirismo e passou a ser feito com o uso de algoritmos de otimização que garantem uma escolha muito próxima da escolha ótima dos recursos.

A implementação do *APS* ajudou a empresa a aumentar as entregas dentro do prazo acordado com o cliente. Antes da implementação do *APS* seja pela falta de antecipação das restrições de produção e mistura, seja pela própria eficiência do processo de misturas manuais, não eram raros índices de entregas no prazo abaixo de 80% em um determinado período. Ou seja, apenas 80% dos clientes, ou menos, recebiam seus pedidos dentro do prazo estabelecido.

Depois da implementação do *APS* este índice melhorou. A Gelita passou a conseguir entregar constantemente até de 95% a 98% dos seus pedidos dentro do prazo. Evidentemente este índice pode variar em função de acontecimentos no processo de produção, mas o gráfico 4.6 bem como o histórico de índice de “entregas perfeitas” (nome utilizado para entregas dentro do prazo estabelecido) indicam que o *APS* realmente mudou os índices de entrega da empresa de patamar.

Outro efeito da implementação do *APS* foi a redução da necessidade de altos estoques de lotes de produção para o atendimento dos pedidos dos clientes. Em função do uso de otimização, o estoque de lotes de produção de gelatina poderia ser menor do que o estoque necessário para se planejar as misturas manualmente.

Assim, o uso do *OBS* permitiu o atendimento de um número relativamente maior de pedidos com o uso de uma certa quantidade de estoque se comparamos com o método anterior. O melhor uso dos estoques possibilita o fortalecimento do nível de serviço de entrega

de gelatina com um nível mais restrito de estoque de lotes de produção. O gráfico 4.6 demonstra este ganho substancial no atendimento de pedidos ao ilustrar as entregas perfeitas, divididos pelo estoque de lotes de produção disponível. Esse é um indicador relativo e somente tem a função de comparar desempenho de entrega relativo ao nível de estoque entre diferentes períodos. É importante mencionar que a operacionalização do sistema se iniciou em fevereiro de 2006 e que os meses de dezembro e fevereiro representam bem o patamar anterior de performance.

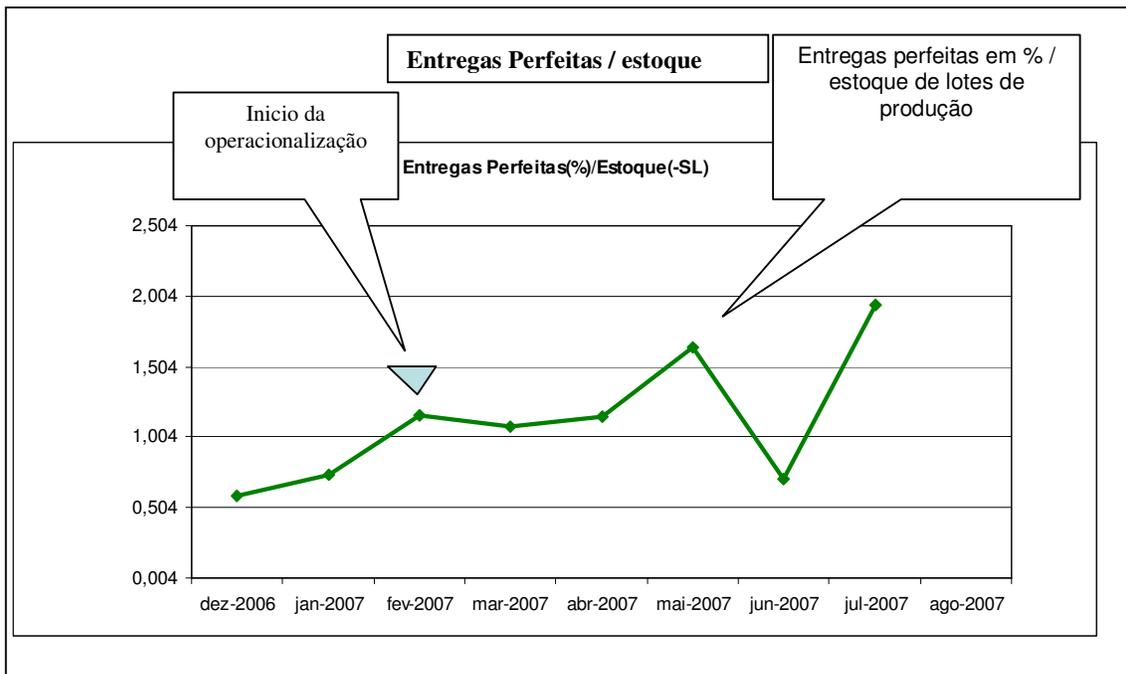


Gráfico 4.6: Entregas perfeitas / estoque antes e depois da implementação do APS
Fonte: Arquivos da Gelita

O uso do *OBS* também trouxe ganhos no processo de suporte da operação. O tempo necessário para realizar a programação e moagem com o uso do modelo de otimização é cerca de 4 vezes menor do o mesmo processo utilizando-se do empirismo. De um tempo de aproximadamente 8 horas para a confecção de 70 toneladas de gelatina, passamos para um tempo de aproximadamente 2 horas. Essa redução permite não só a redução de custos administrativos no setor de planejamento e controle de produção, como também possibilita o envio de ordens de produção com uma maior antecedência para o setor de moagem, o que reduz do risco de paradas inesperadas na produção. O gráfico 4.7 mostra os ganhos no tempo de processamento das ordens de moagem.

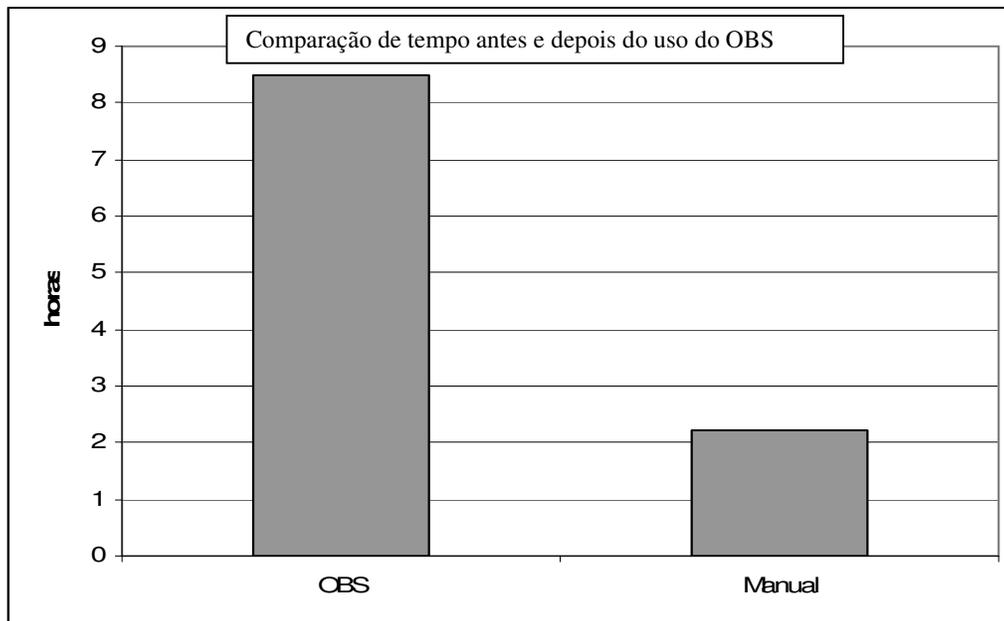


Gráfico 4.7: Comparação de tempo médio de processamento da programação de mistura de 70 toneladas de gelatina com o sem o uso do *OBS*
Fonte: Arquivos da Gelita

A inserção de valores de custos nos perfis de produção no GRP para o planejamento de produção possibilitou a determinação da escolha dos perfis de produção com o menor custo. Ao contrário do processo de planejamento manual, o modelo agora contempla o número desejado de propriedades existentes nas especificações dos clientes. Esse fato faz com que o planejamento seja mais adequado aos perfis dos pedidos sinalizando para a produção quais são exatamente as propriedades necessárias para o atendimento dos pedidos.

Outrossim, o maior ganho proveniente do uso do *GRP* está no balanceamento das vendas em função da capacidade de produção de mix de produto e capacidade nominal de produção das plantas. Isso possibilitou a antecipação do conhecimento de que os pedidos ou a previsão dos pedidos não poderiam ser atendidos, fazendo com que os gestores de demanda entrassem em contato com os clientes e negociando outros prazos de entrega. A medição desses ganhos pode ser observada no gráfico 4.6 que mostra nível de serviços em relação ao nível de estoque.

A figura 4.18 demonstra ilustrativamente como o nivelamento da demanda acontece. Em breves palavras, as simulações no *GRP* possibilitava a visualização dos pedidos que não podem ser atendidos, o que serve como sinalização para os gestores de demanda negociarem novas datas com os clientes.

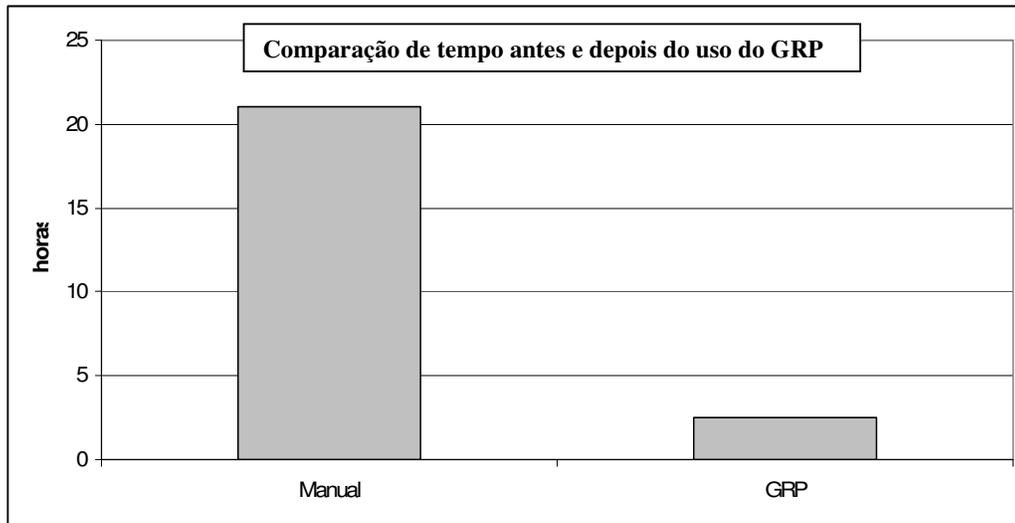


Gráfico 4.8: Comparação do tempo de processamento do planejamento de produção com o sem o uso do *GRP*

Fonte: Arquivos Gelita

Por último, o sistema de controle de produção, Taxa de Sucesso, também trouxe benefícios significativos para a cadeia de produção da Gelita. A comparação entre os perfis de produção planejados e os lotes produzidos pelas plantas faz com que se tenha um melhor controle da qualidade de produção, o que traz embasamento para fazer mudanças nos processos de produção, e auxilia a determinação das causas de eventuais atrasos de pedidos.

Ao analisar a aplicação do *APS* de otimização, podemos concluir que trouxeram valiosos benefícios para a Gelita. Grande parte destes benéficos são derivados da alta complexidade existente na programação e planejamento em um ambiente de misturas e a dificuldade de gerenciar esses processos sem o uso de sistemas de suporte. Desta forma, as aplicações dos modelos de otimização na Gelita cumpriram com o papel da aplicação de Pesquisa Operacional das empresas: suportar a tomada de decisão em ambientes complexos de serem gerenciados (WINSTON, 1995).

5. Conclusões e considerações

Segundo esse estudo, o uso de algoritmos matemáticos na forma dos *APSs* para o suporte de decisões nos processos de planejamento e organização da produção tem crescido significativamente nos últimos anos (STADTLER e KILGER, 2005). Isso se deve, entre outros motivos, ao aumento da velocidade de processamento computacional derivados do avanço tecnológico das últimas duas décadas, as ineficiências dos antigos sistemas de suporte a produção e a procura pela melhoria da eficiência operacional das empresas (PEDROSO e CORREA, 1996).

Esta dissertação teve o objetivo geral de contextualizar o uso dos *APSs* na estrutura, processos e métodos de apoio ao planejamento e organização da produção. No capítulo 2 argumentamos que, do ponto de vista evolutivo dos sistemas de suporte a produção, os *MRPs* demonstraram ser ineficientes com relação a gestão de capacidades finitas e na flexibilidade em realizar a programação detalhadas da produção (CORREA e GIANESI, 2011; HOOP, 2008). Por outro lado, os sistemas puxados embora tenham conseguido, até certo ponto, suprir algumas deficiências do *MRPs*, ainda deixaram algumas lacunas com relação gestão de estruturas complexas e em menor grau, na avaliação de capacidade dos recursos (CORREA e GIANESI, 2008; TARDIN, 2001). De mesma forma, apesar de o *OPT*, apresentar uma boa sistemática para gerenciar a produção, demonstrou ser dependente de um sistema proprietário e um pouco complexo de ser gerenciado. Das deficiências desses sistemas surgem os *APSs*.

Segundo a pesquisa bibliográfica deste trabalho, parece não haver consenso com relação a abrangência e definição do termo *APS*, sendo a definição genérica da *APICs* “qualquer programa computacional que usa lógica ou algoritmos matemáticos avançados para realizar otimização ou simulação em ambientes de capacidade finita” aparentemente a mais adequada em função das diferentes visões sobre o tema.

Nossa pesquisa apontou que os *APSs* eram preponderantemente utilizados na gestão operacional do sequenciamento de produção na década de 90. Porém, com a evolução dos algoritmos e da tecnologia de processamento computacional, os *APSs* passaram a ser utilizados em processos mais táticos e estratégicos do planejamento de produção. Assim, os *APSs* podem funcionar como suporte para os sistemas *MRPs* na determinação da capacidade de produção (HADDAD e CARVALHO, 2003), como métodos de programação detalhada da produção (TUBINO, 2007), ou ainda, desempenhar um papel mais abrangente nos níveis estratégico como alocação ótima de recursos das plantas em um médio prazo como argumentam Stadtler e Kilger (2005).

Os *APSs* atualmente no mercado como o *APO* do SAP e o *ASCP* da ORACLE têm sido cada vez mais demandados pelas empresas para o controle e o suporte no planejamento e organização da produção (STADTLER, 2005). Contudo, a aplicação de soluções genéricas de mercado em função das peculiaridades dos diversos tipo e ambientes de produção, principalmente no que se refere a programação detalhada da produção, tem sido limitada. Assim, diferentes tipos de processos de produção demandam diferentes modelos de *APSs* ou otimização linear (TEMPELMEIER 2005)..

Os problemas mais conhecidos e incidentes são aqueles que suportam a definição do tamanho dos lotes (*lotes sizing*) característicos dos processos de *Process Flow* e *cell flow* e o seqüenciamento de produção (*scheduling*) freqüentemente vistos em processos de *Job Shop* (ARENALES, 2007).

O trabalho também procurou responder a pergunta de como implementar um *APS* de otimização linear no planejamento e programação em ambientes complexos de produção. No capítulo 2 iniciamos a respostas dessa questão apresentando o método Simplex e o *branch and Cut* para resolução dos problemas de programação inteira que foram utilizados no modelos implementados na Gelita. No capítulo 3 continuamos a responder a pergunta ao descrever quatro etapas do método para a construção e implementação de um *APS* de otimização. O processo se inicia com o diagnóstico e é seguido de um processo de modelagem, teste e finalmente implementação do sistema.

Contudo, o objetivo principal do trabalho era a descrição e análise da aplicação de um *APS* de otimização em uma empresa de produção de bens alimentícios e ao descrever este caso, expor as dificuldade e caminhos para a implementação de um *APS* na prática. No capítulo 4 descrevemos as quatro etapas do processo implementação do *APS*. Os vários percalços relacionados à necessidade de customização para ganhos de flexibilidade e adaptação do sistema ao processo de planejamento como bem anteciparam Tonini (2011) e Oliveira (2011) foram avaliados. Também foram expostos os problemas relacionados ao treinamento dos recursos humanos que gerenciam os sistemas de otimização citados por Tempelmeier (2005). Um perfil analítico e o conhecimento dos conceitos de otimização se mostraram necessários para a operacionalização dos modelos implementados na Gelita.

O uso do OBS para a programação de misturas e do GRP para o planejamento de produção possibilitaram a otimização do uso dos estoques, a escolha ótima dos perfis de produção e a redução dos tempos administrativos para a realização do planejamento e programação. Ademais, a facilidade na simulação de misturas ótimas possibilitou a melhor visualização das restrições de mix de produção das plantas, causando maior performance na

gestão dos pedidos em um médio prazo. A implementação do *APS* na Gelita provou causar grande impacto em indicadores de grande importância para o resultado econômico da companhia.

Embora o caminho trilhado para a melhoria do processo de planejamento e organização da produção exposto no estudo de caso não tenha sido tão retilíneo quanto gostaríamos, Os *APSs* de otimização parecem ser um solução eficiente para ambientes complexos de planejamento de produção. Acreditamos que tais sistemas possam também ter grande valor para empresas caracterizadas por outros processos de produção.

A aplicação de modelos de otimização no forma dos *APSs* é um processo pouco explorado pela literatura acadêmica brasileira. Esperamos que a dissertação tenha apresentado elementos importantes sobre o assunto e que possa agregar valor nos estudos sobre o tema que sucedem o sucedem.

5.1 Sugestões para estudos futuros

Este trabalho explora a implementação e uso de um *APS* no gerenciamento tático e operacional da produção. Estudos sobre o uso e implementação de *APSs* na tomada de decisões em níveis estratégicos das empresas são também aparentemente escassos na academia brasileira.

Tais estudos aumentam o escopo do presente trabalho ao explorar aspectos como os níveis de recursos para o atingimento das metas orçamentarias, capacidade necessária dos ativos e localização das fábricas e centros de distribuição das empresas. Ainda, estes trabalhos incrementam a abrangência da presente dissertação, contemplando elementos não pertencentes a uma empresa específica, mas sim as cadeias de suprimentos como um todo.

Por último, seria de grande contribuição para a academia e para as empresas brasileiras a proposição de um método adequado e eficiente para a implementação dos *APSs* nos respectivos níveis de decisão e etapas de operação das empresas. Nosso trabalho teve a intenção apenas de descrever o método utilizado no estudo de caso de sucesso.

6. Referências bibliográficas:

ARENALES, Marcos. Pesquisa Operacional. Rio de Janeiro; Elsevier, 2007;

BICHENO, JOHN. *The Leantoolbox*. Buckinham. PICSIE Books, 2000.

CORRÊA, H, L. e CORRÊA C. A. Administração de produção e operações: manufatura e serviços: uma abordagem estratégica.. São Paulo; Atlas 2004.

CORRÊA, H, L.; GIANESI, I. G. *Just in Time, MRPII, e OPT*, Uma Enfoque Estratégico. São Paulo, Editora Atlas, 2007.

DEMIN, W. EDWARDS. *Out of the Crisis*. MIT Press, 1986.

Duckworth, W. E.; *A Guide to Operational Research*, Halsted Press, 3 ed., 1977.

Duri, C., Fren, Y., Di Masclo M. *Comparisson between three pull control Policies: Kanban, base stock, Generalized Kanban*, Annals of Operational Research, 2000 v.93, p. 41-69.

ENTRUP, M.L. *Advanced Planning ad Scheduling in Fresh Food Industries*. 1 ed, Germany, Physica - Verlag Heidenberg, 2005.

FALCONI, C.,V. Gerenciamento da Rotina do Trabalho do Dia a Dia. 8 ed. Minas Gerais: Editora INDG, 2004.

FRONTILINE PLATFORM, *Premium Solver - User Guide*, 2006.

FRONTILINE USER GUIDE, *Premium Solver Platnform, Xpress Solver User Guide*, 2006.

GAITHER, N. FRAZIER, G. A administração da Produção e Operações. São Paulo; Pioneira Thompson Learning, 2004. 8ª Ed.

GARTNER GLOBAL SURVEY. Disponível em: www.gartner.com/it/page.jsp?id=1357514. Acesso em 10/02/2012.

GARWOOD, D. Sales and Operational Planning. R.D.Garwood Articles, USA, 2002. Disponível em: www.rdgarwood.com/archive/hot14. Acesso em 05/01/2009.

GAVIRA, M., O. Simulação Computacional como Ferramenta de Aquisição de Conhecimento. 2003. Dissertação de Mestrado. Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos.

GIL, A. C. Métodos e Técnicas de Pesquisa Social. 4 ed. São Paulo, Atrás, 1995.

GIROTTI, J. L.; MESQUITA, A. M. Uma análise Comparativa das Abordagens MRP e APS pra Programação Detalhada da Produção. Artigo. SIMPOE, 2011.

GIACON, E., Implementação de Sistema de Programação Detalhada: Levantamento das Práticas de Programação da Produção na Indústria, 2010. Dissertação de Mestrado - Escola Politécnica, Universidade de São Paulo, São Paulo.

GOLDRATT, E. What is this thing called Theory of Constraints and how should it be implemented? Great Barrington, MA: North River Press, 1990.

GOLDRATT, E. A Meta: Um processo de aprimoramento contínuo. São Paulo: Educator, 1993.

HADDAD, R. B. B; CARVALHO, M. F. H, ROCHA, B. R., Integração Entre ERP e Programação Matemática, Um Estudo de Caso na Indústria de Autopeças. Campinas, SP; CEMPRA -Centro de Pesquisa Renato Archer, 2005.

HILLIER, S. F, LIEBERMAN, J. G. Introdução a Pesquisa Operacional, 8ª Edição, Porto Alegre, AMGH, 2010.

HOOP, W. J. Factory Physics. New York, McGraw- Hill/Erwin, 2008;

HUANG, M.; WANG, D. *A Simulation and Comparative Study of CONWIP, Kanban and MRP Production Control Systems in a Cold Rolling Plant; Production Planning and Control*, 1998, V. 9, n.8, p.803 – 812.

LANGLEY, J. G; MOAN, D. R; NOLAN, M. K; NOLAN, W. T, N, L. C; PROVOST, P. Lloyd. – Modelo de Melhoria Uma abordagem prática para melhorar o desempenho organizacional. Campinas, Mercado das Letras, 2011.

Liker, J. *Becoming Lean*. New York: Productivity Press, 1998.

LIKER, JEFFREY K. – *The Toyota Way*. New York: McGraw-Hill, 2004.

- LUSTOSA et al. Planejamento e Controle da Produção. Rio de Janeiro, Elsevier, 2008.
- KJELLSDOTTER, L., *User Experiences of Advanced Planning and Scheduling (APS Systems in Production Scheduling)*; Sweden, University of Technology Gothenburg, 2001.
- MCCLSOKEY, J. F; FLORENCE, N. *Operational Reseach for Management*. John Hopkins Press, 1956.
- MONDEN, Y., *Toyota Production System*, 3 Ed. Tokyo, 1998.
- OLIVEIRA, A., S. *Sistemas de Informação Gerenciais em Empresas Multinacionais: Um estudo de caso da Implementação Global do ERP e BI*. 2011. Dissertação de Mestrado. Faculdade de Economia, Administração e Contabilidade. Universidade de São Paulo, São Paulo.
- ORACLE ASCP USER GUIDE – *Advanced Supply Chain Planning, Implementation and User Guide, Volume 1 and 2*, 2001.
- PACE, H. João; *O Kanban na Prática*. Rio de Janeiro, Qualimark, 2003.
- PEDROSO, M. C; CORRÊA, H. L. *Sistemas de Programação com Capacidade Finita. Uma decisão Estratégica?* ERA. Revista de Administração de Empresas. Vol. 36, n4, p60-73, 1996.
- PETRONI, A. *A Critical Factor of MRP Implementation in Small and Medium Size Firms*. *Intenational Journal of Operation & Production Management*, v22, n3, p. 324, 348, 2002.
- PLENERT, G.; KIRCHMIER, B. *Finite Capacity Scheduling: selection and implementation*. New York: John Wiley and Sons, 2000.
- SAP LIBRARY, disponível em: [help,sap.com](http://help.sap.com) (Acessado em fev/2012).
- SHOOCK. J. *Aprendendo a Enxergar*. São Paulo: Lean Institute Brasil, 2003.
- SLACK. N; CHANBERS, S. *A Administração da Produção*. 3 ed. São Paulo, Editora Atlas, 2009.
- STADTLER, H. *Supply Chain Management and Advanced Planning – basics, overview and challenges*. *European Journal of Operational Research*. Vol 163, Issue 3 p. 575-588, 2005.

STADTLER, A; KILGER, C, *Supply Chain Management and Advanced Planning Concepts, Models, Softwares and Case Studies*, 3 ed. Berlin: Springer, 2005.

TAKAHASHI, K.; NAKAMURA, N., *Reactive Logistics in a JIT Enviroment, Production Planning and Control*, 2000.

TAHA, A. Handy – Pesquisa Operacional, São Paulo, Pearson Prentice, Hall, 2008.

TARDIN, G. G. - O kanban e o Nivelamento de Produção, 2001. Tese de Mestrado, Faculdade de Engenharia Mecânica, Unicamp, Campinas.

TAYLOR, D. A nova Concorrência. Editora Pearson, 2005.

TEMPELMER, H. *Sypply Chain with Advanced Planning Systems, Working Paper, University of Colagne*, 2000.

TONINI, A., C. A Contribuição do Seis Sigma para a melhoria dos processos de Software, 2006. Dissertação de Mestrado. Eng. de Produção, Escola Politécnica, Universidade de São Paulo, São Paulo.

TUBINO, D. F. Planejamento e Controle da Produção : Teoria e Prática. São Paulo, Ed. Atlas, 2007.

WILLIANS, H.P., *Model Building in Mathematical Programming*. 4 ed. John & Son ltd, 1999.

WINSTON W.L.. *Introduction to Mathematics Programming, Aplication and Algorithms*. Duxbury Press Belton, California, 1995.

Womack, J., Jones, D. A Mentalidade Enxuta nas Empresas. Rio de Janeiro: Editora Campus, 1998.

Womack, J., Jones, D., Ross, D. A Máquina que Mudou o Mundo. 2 ed. Rio de Janeiro: Editora Campus, 1992.