UNIVERSIDADE DE SÃO PAULO
ESCOLA POLITÉCNICA

ARNALDO RABELLO DE AGUIAR VALLIM FILHO

LOCALIZAÇÃO DE CENTROS DE DISTRIBUIÇÃO DE CARGA
CONTRIBUIÇÕES À MODELAGEM MATEMÁTICA

SÃO PAULO
2004

ARNALDO RABELLO DE AGUIAR VALLIM FILHO

LOCALIZAÇÃO DE CENTROS DE DISTRIBUIÇÃO DE CARGA
CONTRIBUIÇÕES À MODELAGEM MATEMÁTICA

Tese apresentada à Escola Politécnica da Universidade de São Paulo como requisito para obtenção do título de Doutor em Engenharia

Área de Concentração:
Engenharia de Transportes

Orientador:
Prof. Dr. Nicolau Dionísio Fares Gualda

SÃO PAULO
2004

A minha mulher Vilma e a meus filhos Renata e Marco Vinicius que acompanharam e apoiaram esta minha longa jornada, para que ao menos o exemplo permaneça. A vocês com muito amor.

AGRADECIMENTOS

São muitos os agradecimentos que seriam devidos a todos que direta ou indiretamente contribuíram para a consecução deste trabalho, pois esta é, na verdade, mais uma etapa dos longos anos de pesquisa e projetos que venho desenvolvendo nas áreas de Pesquisa Operacional e Logística. Vou me ater àqueles que diretamente participaram do período em que estive envolvido no programa de doutoramento. Aos outros que estiveram presentes nos períodos anteriores de minha vida profissional fica aqui o meu muito obrigado pelo estímulo, apoio e oportunidades oferecidas.

Devo agradecer inicialmente ao caro amigo Prof. Dr. Orlando Strambi que me abriu as portas desta universidade mostrando-me os caminhos a percorrer para que eu pudesse chegar ao ponto em que me encontro no momento.

Agradeço ainda à Universidade Presbiteriana Mackenzie na figura do Prof. Dr. Vilmar Pedro Votre, diretor da Faculdade de Computação e Informática pelo apoio, incentivo e importante suporte financeiro proporcionado neste último ano de trabalho demonstrando seu compromisso com a pesquisa científica.

Devo agradecer também aos membros de minha banca de qualificação Prof. Dr. Antonio Galvão Novaes e Prof. Dr. Marco Antonio Brinati pelas críticas, sugestões e comentários precisos que muito auxiliaram na condução e complementação final da pesquisa.

Mas acima de tudo dois agradecimentos são devidos: primeiro ao meu orientador e agora amigo, Prof. Dr. Nicolau Dionísio Fares Gualda pela confiança depositada ao longo de todo esse período e pelas suas orientações sempre objetivas e seguras que tanto enriqueceram o trabalho e o conduziram a um desfecho satisfatório, e também um agradecimento muito especial à minha família que soube suportar com serenidade as dificuldades desta minha longa ausência, apesar da presença, em que estive percorrendo estes árduos, porém estimulantes caminhos da modelagem matemática, conhecidos e valorizados por tão poucos.

A todos vocês o meu mais sincero, muito obrigado.

RESUMO

A localização de instalações está entre as mais importantes decisões logísticas. A questão é tratada, em geral, por técnicas de Pesquisa Operacional, principalmente a programação linear inteira mista, cuja modelagem tem aspectos que podem ser aprimorados. Esta tese apresenta uma proposta metodológica subdividindo o problema em duas fases, visando localizar centros de distribuição de carga (CD's).

A *Fase 1* define as localizações ótimas dos CD's, iniciando com uma agregação de nós de demanda em *clusters* através de uma heurística, complementada pela meta-heurística *simulated annealing* (SA). Em seguida, buscam-se, em cada *cluster,* os "melhores" locais candidatos. Entre estes candidatos definem-se as localizações ótimas, através de um modelo de programação linear inteira mista ou por SA. Foi conduzido um experimento controlado, com solução ótima conhecida, tendo coincidido em vários casos com a solução obtida através da metodologia proposta.

A *Fase 2* define a quantidade "ótima" de CD's, modelando os custos de estoque, armazenagem e vendas perdidas, não considerados na primeira fase. Estes modelos exógenos ao modelo de otimização da primeira fase podem trabalhar sem a restrição de linearidade, trazendo assim, maior realismo a modelagem. Ao final consolida-se o conjunto de custos logísticos (*Fases 1 e 2*) de forma a se identificar a quantidade de CD's que conduz ao mínimo custo total.

A metodologia foi aplicada a um caso real com mais de duzentos pontos de demanda distribuídos sobre os três estados da região sul do país. Os resultados mostraram a aplicabilidade da abordagem proposta.

Palavras Chave: 1. Localização industrial 2. Pesquisa operacional 3. Logística

ABSTRACT

The facility location is one of the most important logistic decisions. The problem is usually handled through the use of operations research techniques, mainly with the use of mixed integer linear programming (MILP), which modeling process can be improved in some of its aspects. This thesis presents a methodological proposal subdividing the problem in two stages, in order to locate distribution centers (DC's).

Stage 1 identifies the optimal DC locations, starting with a heuristics, complemented by the metaheuristic simulated annealing (SA), to cluster the demand nodes. Next, the heuristics searches for "best" DC's candidates in each cluster. Among these "best" candidates are defined the optimal locations, with the use of a MILP model or through the use of SA. In order to validate the methodology, it was conducted a controlled experiment, with a known optimal solution, having matched in many cases with the solutions obtained through the proposed approach.

Stage 2 defines the "optimal" DC number, modeling the costs of inventory, storage and lost sales, which were not considered in the first stage. These models, exogenous to the *Stage 1* optimization model, could represent the costs behavior without the linear restriction, bringing more realism to the modeling process. Finally, the whole set of logistic costs (*Stages 1 and 2*) is combined in order to identify the DC quantity with minimal total cost.

The methodology was applied to a real world problem encompassing more than two hundred demand points spread over the three states of the southern Brazilian region. The results have shown the applicability of the proposed approach.

Keywords: 1. Industrial location. 2. Operations research. 3. Logistics.

LISTA DE FIGURAS
Capítulos 1 a 4

Figura	Título	Pág
1	Escopo da pesquisa	24
4.1	Zoneamento típico	84
4.2	Zonas e suas MPT's	93
4.3	Região com uma MPT central e quatro MPT's radiais	93
4.4a	MPT's vs. Melhores localizações – possibilidade 1	94
4.4b	MPT's vs. Melhores localizações – possibilidade 2	94
4.5	O Conceito de raio máximo de busca	96
4.6	Heurística do ângulo máximo – fatores principais	100
4.7	Nuvem aleatória de pontos para o experimento	109
4.8	Heurística de *clusterização* - Proposta de zoneamento	110
4.9	Zoneamento definido e as MPT's de cada zona.	111
4.10a	Visão geral das MPT's	112
4.10b	MPT's e MPT's Meio, geradas entre MPTs adjacentes	112
4.11a	Candidatos gerados no entorno das MPT's	113
4.11b	Visão geral dos Candidatos	113
4.12	Solução ótima gerada a partir dos candidatos	114
4.13a	Zoneamento por Pontos cardeais	116
4.13b	Zoneamento em Quadrantes	116
4.13c	Zoneamento em faixas verticais	116
4.13d	Zoneamento em faixas horizontais	116
4.13e	Zoneamento Ballou	116
4. 13f	Proposta de zoneamento da Tese	116
4.14	Quadrículas Ilustração com zoneamento em Faixas Horizontais	119
4.15	Quadrículas e *clusters* Ilustração com zoneamento em Pontos Cardeais	119
4.16	Proposta Tese vs. Pontos Cardeais	121
4.17	Proposta Tese vs. Quadrantes	121
4.18	Proposta Tese vs. Faixas Horizontais	122
4.19	Proposta Tese vs. Faixas Verticais	122
4.20	Proposta Tese vs. Proposta Ballou	122
4.21a	*Clusterização* – 4 Zonas	125
4.21b	*Clusterização* – 3 Zonas	125
4.21c	*Clusterização* – 2 Zonas	125
4.21d	*Clusterização* – 1 Zona	125
4.22a	MPT's – 4 Zonas	126
4.22b	MPT's – 3 Zonas	126
4.22c	MPT's – 2 Zonas	126
4.22d	MPT's – Zona única	126
4.23a	Candidatos - 4 Zonas	127
4.23b	Candidatos - 3 Zonas	127
4.23c	Candidatos - 2 Zonas	128
4.23d	Candidatos - Zona única	128
4.24a	Solução - 4 Zonas e 4 CD's	128
4.24b	Solução - 5 Zonas Fixas e 4 CD's	128
4.24c	Solução - 3 Zonas e 3 CD's	129
4.24d	Solução - 5 Zonas Fixas e 3 CD's	129
4.24e	Solução - 2 Zonas e 2 CD's	129
4.24f	Solução - 5 Zonas Fixas e 2 CD's	129
4.24g	Solução - 1 Zona e 1 CD	129
4.24h	Solução - 5 Zonas Fixas e 1 CD	129

LISTA DE FIGURAS
Capítulos 5 e 6

Figura	Título	Pág
5.1	SA:Tela de resultados apresentando ΔE_m^+	144
5.2a	Variação de GH Máximo com o decréscimo de temperatura	146
5.2b	Detalhamento da variação de GH Máximo para temperaturas abaixo de 10,0	146
5.3a	Melhor configuração gerada para o anel 1	148
5.3b	Melhor configuração gerada para o anel 2	148
5.4a	Resultados de índices de monitoramento do processo Temperaturas iniciais (mais altas) – Anel 2	149
5.4b	Índices de monitoramento – Convergência de soluções Temperaturas finais (mais baixas) - Anel 2	150
5.5	Evolução de GH Médio com a redução de temperatura	150
5.6	Convergência da solução "ótima" partindo de uma solução diferente da "ótima"	151
5.7	Distribuição de GH ao longo da simulação – Anel 2	152
5.8	Histograma de GH - Anel 2	152
5.9	Resultado final com "melhor" configuração global Anéis 1 e 2	153
5.10	Convergência da solução "ótima" . Inicia na solução "ótima" e retorna à "ótima"	154
5.11	Localizações "ótimas" de CD´s - SA	155
5.12	Definição de localizações "ótimas" de CD's Processo com meta-heurística SA em 4 Fases	157
5.13	Processo Heurístico em 4 Fases e características do problema	160
5.14	Tela de resultados apresentando ΔE_m^+ na SA de localização	162
5.15a	Variação de Custo Mínimo com o decréscimo de temperatura	164
5.15b	Detalhamento da variação de Custo Mínimo para temperaturas abaixo de 10,0	164
5.16	Melhor configuração gerada	166
5.17a	Resultados de índices de monitoramento do processo Temperaturas iniciais (mais altas)	167
5.17b	Índices de monitoramento – Convergência de soluções Temperaturas finais (mais baixas)	167
5.18	Evolução do Custo Médio com a redução de temperatura	168
5.19	Convergência da solução "ótima" partindo de uma solução diferente da "ótima"	168
5.20	Convergência da solução "ótima" .Inicia na solução "ótima" e retorna à "ótima"	169
5.21	Comparação de resultados – 5 CD's	170
5.22	Comparação de resultados – 4 CD's	171
5.23	Comparação de resultados – 3 CD's	171
5.24	Comparação de resultados – 2 CD's	172
5.25	Comparação de resultados – 1 CD	172
6.1a	Evolução de estoque e estoque médio - Origem	179
6.1b	Evolução de estoque e estoque médio - Destino	179
6.2	Custos fixos de movimentação e Armazenagem - Evolução descontínua	184
6.3	Inconformidade vs. quantidade de CD's	190
6.4	Possível evolução de custos logísticos com a quantidade de CD´s	193

LISTA DE FIGURAS
Capítulo 7

Figura	Título	Pág
7.1	Pontos de Demanda	194
7.2	Anéis desbalanceados	202
7.3	Heurística de *clusterização* - Proposta de zoneamento	204
7.4a	Anel 2: GH Máximo evoluindo com o decréscimo de temperatura	207
7.4b	Anel 2: Evolução de GH Máximo para temperaturas abaixo de 1,2	208
7.5	GH Médio vs. Temperatura	208
7.6	Convergência da solução “ótima” *Clusterização*	209
7.7	*Simulated Annealing* - Zoneamento final	210
7.8	Zoneamento e MPT’s por zona	211
7.9a	MPT’s - Visão geral	212
7.9b	MPT’s e MPT’s Meio	212
7.10	Raios máximos de busca no entorno das MPT’s	214
7.11a	Candidatos gerados no entorno das MPT’s	215
7.11b	Candidatos gerados no entorno das MPT’s Meio	215
7.12	Configuração geral de “melhores” Candidatos	216
7.13a	Custos vs. No. de CD’s	218
7.13b	Distâncias médias aos CD's	218
7.14	Evolução da solução ótima para *p* CD´s	219
7.15a	Solução ótima - 1 CD	220
7.15b	Solução ótima - 2 CD´s	220
7.15c	Solução ótima - 3 CD´s	220
7.15d	Solução ótima - 4 CD´s	220
7.15e	Solução ótima - 5 CD´s	220
7.15f	Solução ótima - 6 CD´s	220
7.15g	Solução ótima - 7 CD´s	220
7.16a	Custo Mínimo evoluindo com o decréscimo de temperatura	223
7.16b	Evolução do custo mínimo temperaturas abaixo de 5,0	223
7.17	Custo médio vs. temperatura	224
7.18	Convergência da solução “ótima” Localização de CD's	224
7.19a	SA: Custos vs. No. de CD’s	225
7.19b	SA: Distâncias aos CD's	225
7.20	SA: solução ótima para *p* CD´s	226
7.21a	Solução “ótima” SA - 1 CD	227
7.21b	Solução “ótima” SA - 2 CD´s	227
7.21c	Solução “ótima” SA - 3 CD´s	227
7.21d	Solução “ótima” SA - 4 CD´s	227
7.21e	Solução “ótima” SA - 5 CD´s	227
7.21f	Solução “ótima” SA- 6 CD´s	227
7.21g	Solução “ótima” SA - 7 CD´s	227
7.22a	MILP vs SA - 1 CD	230
7.22b	MILP vs SA - 2 CD´s	230
7.22c	MILP vs SA - 3 CD´s	230
7.22d	MILP vs SA - 4 CD´s	230
7.22e	MILP vs SA - 5 CD´s	230
7.22f	MILP vs SA - 6 CD´s	230
7.22g	MILP vs SA - 7 CD´s	230

LISTA DE FIGURAS
Capítulo 7 (cont.) e Anexos

Tabela	Título	Pág
7.23a	1 CD - MILP	231
7.23b	1 CD - SA	231
7.24a	2 CD's - MILP	231
7.24b	2 CD's - SA	231
7.25a	3 CD's - MILP	231
7.25b	3 CD's - SA	231
7.26a	4 CD's - MILP	231
7.26b	4 CD's - SA	231
7.27a	5 CD's - MILP	232
7.27b	5 CD's - SA	232
7.28a	6 CD's - MILP	232
7.28b	6 CD's - SA	232
7.29a	7 CD's - MILP	232
7.29b	7 CD's - SA	232
7.30	Nível de estoque no mês i	235
7.31	Volume movimentado no mês i+1	235
7.32	Volume movimentado no mês i	235
7.33	Nível de estoque vs. movimentação	236
7.34	Estoque regulador vs. número de CD'	237
7.35a	Distância de percurso vs. No. de CD's - MILP	239
7.35b	Distância de percurso vs. No. de CD's – SA	239
7.36a	Estoque em trânsito vs. No. de CD's – MILP	240
7.36b	Estoque em trânsito vs. No. de CD's – SA	240
7.37	Estoque de segurança vs. No. de CD's	241
7.38a	Estoque total vs. No. de CD's - MILP	242
7.38b	Estoque total vs. No. de CD's - SA	242
7.39	Configurações de CD´s alocadas a patamares de custo	245
7.40	Variação de custos unitários por patamar de custo	246
7.41	Custos unitários médios por patamar de custo	246
7.42	Custos de armazenagem vs. No. de CD´s	247
7.43a	Modelo de inconformidade - MILP	249
7.43b	Modelo de inconformidade – SA	249
7.44a	Custos de vendas perdidas - MILP	250
7.44b	Custos de vendas perdidas – SA	250
7.45a	Custo logístico global – MILP	253
7.45b	Custo logístico global – SA	253
B.1	Distribuições Normal, Gama, Beta e Rayleigh - Todos os anéis.	281
B.2	Distribuições Normal, Gama, Beta e Rayleigh - Anéis 2 e 3	282
B.3	Distribuições Normal, Gama e Beta - Anel 1	283
B.4	Distribuições Normal, Gama e Beta - Anel 2	284
B.5	Distribuições Normal, Gama e Beta - Anel 3	285

LISTA DE TABELAS

Tabela	Título	Pág
4.1	Seleção de Locais – Atributos	75
4.2	Amostragem descritiva – geração e permutação	108
4.3	Proposta da pesquisa vs. outros critérios - Soluções ótimas Diferença percentual nos custos da operação	123
7.1	Pontos de Demandas	195
7.2	Parâmetros de custos de transferência e distribuição	200
7.3	Fontes de produção	200
7.4	Definição de anéis	201
7.5	Momento de transporte e *clusters* por anel	201
7.6	Momento de transporte e clusters por anel – Nova configuração	203
7.7	Resultados finais - *clusterização*	204
7.8	Resultados finais SA	210
7.9	Valores críticos de Z ao nível 1-β_o. de confiança	213
7.10	Raios Máximos de busca de candidatos no entorno de MPT's	213
7.11	Modelo de otimização: solução ótima para *p* CD´s	218
7.12	*Simulated Annealing*: solução "ótima" para *p* CD´s	225
7.13	MILP vs *Simulated Annealing* solução "ótima" para *p* CD´s	229
7.14a	Volume acima de $D_{máx}$* - Solução via MILP	248
7.14b	Volume acima de $D_{máx}$* - Solução via SA	248
7.15a	Componentes do custo logístico - MILP	252
7.15b	Componentes do custo logístico – SA	252
7.16	Custo logístico global– Soluções MILP e SA	252
A1a	Dados de entrada do modelo MILP	279
A1b	Condições operacionais	279
A1c	Condições de integralidade - variáveis binárias	279
A1d	Número de Variáveis	279

SUMÁRIO

1. INTRODUÇÃO 15

1.1. INTRODUÇÃO AO PROBLEMA 15

1.2. OBJETIVOS DA PESQUISA 22

2. REVISÃO BIBLIOGRÁFICA 25

2.1. RELEVÂNCIA DO PROBLEMA DA LOCALIZAÇÃO 26

2.2. CONTORNOS DO PROBLEMA 28

2.3. MÉTODOS DE SOLUÇÃO 32

2.4. DESCRIÇÃO DOS MODELOS MATEMÁTICOS DE LOCALIZAÇÃO 38

2.5. DADOS DE ENTRADA DOS MODELOS DE LOCALIZAÇÃO 51

3. PROPOSTA DE MODELAGEM 58

3.1. IDÉIAS INTRODUTÓRIAS 58

3.2. SEGMENTAÇÃO DA MODELAGEM 67

4. MODELAGEM DE LOCALIZAÇÃO DE CD's 71

4.1. O USO DE CANDIDATOS 71

4.2. SELEÇÃO QUALITATIVA DE CANDIDATOS 73

4.3. PROCEDIMENTO QUANTITATIVO DE SELEÇÃO DE CANDIDATOS 76

4.3.1. Heurística de "*Clusterização*" 78

4.3.1.1. Homogeneidade das zonas 78

4.3.1.2. Montagem das zonas 81

4.3.1.3. Nível de serviço e número de zonas 85

4.3.1.4. Passos da heurística de *"clusterização"* 87

4.3.2. Seleção dos “Melhores” Candidatos a CD’s 91

4.3.2.1. Heurística do raio máximo 95

4.3.2.2. Heurística do ângulo máximo 99

4.4. MODELO DE OTIMIZAÇÃO PARA LOCALIZAÇÃO DOS CD’s 101

4.5. EXPERIMENTO DE VALIDAÇÃO DAS HEURÍSTICAS 104

4.6. VARIANTES NO USO DAS HEURÍSTICAS 123

5. USO DE META-HEURÍSTICA NO PROBLEMA DA LOCALIZAÇÃO 132

5.1. META-HEURÍSTICA *SIMULATED ANNEALING* 133

5.2. USO DE SA NA MONTAGEM DE ZONAS 139

5.2.1. Descrição do Modelo de Zoneamento 139

5.2.2. Parâmetros da SA e Dados de Entrada - Zoneamento 142

5.2.3. Resultados Experimentais do Zoneamento 147

5.3. USO DE SA NA LOCALIZAÇÃO FINAL DOS CD’s 154

5.3.1. Descrição do Modelo de Localização 155

5.3.2. Parâmetros da SA e Dados de Entrada - Localização 160

5.3.3. Resultados Experimentais da Localização 165

5.4. COMPARAÇÃO DE RESULTADOS 169

6. MODELAGEM DE QUANTIFICAÇÃO DE CD’S 174

6.1. CUSTOS LOGÍSTICOS EXTERNOS AO MODELO DE OTIMIZAÇÃO 174

6.2. MODELAGEM DOS CUSTOS DE ESTOQUE 175

6.3. MODELAGEM DOS CUSTOS DE ARMAZENAGEM
E MOVIMENTAÇÃO DE MATERIAIS 183

6.4. MODELAGEM DOS CUSTOS DE VENDAS PERDIDAS 188

6.5. QUANTIDADE “ÓTIMA” DE CD’s 191

7. APLICAÇÃO PRÁTICA 194

7.1. SITUAÇÃO EM ESTUDO 194

7.2. FASE 1: LOCALIZAÇÃO DE CD's 198

7.2.1. Heurística de *"Clusterização"* Aplicada 198

7.2.2. Meta-Heurística SA na Montagem de Zonas 205

7.2.3. Seleção dos "Melhores" Candidatos 211

7.2.4. Modelo de Otimização Aplicado a Localização dos CD's 215

7.2.5. Meta-Heurística SA para Localização dos CD's 219

7.2.6 MILP e SA: Comparação de Resultados 226

7.3. FASE 2: QUANTIFICAÇÃO DE CD's 232

7.3.1. Modelo de Custos de Estoque 233

7.3.2. Modelo de Custos de Armazenagem e Movimentação de Materiais 243

7.3.3. Modelo de Custos de Vendas Perdidas 247

7.3.4. Determinação da Quantidade "Ótima" de CD's 251

8. CONSIDERAÇÕES FINAIS, CONCLUSÕES E RECOMENDAÇÕES 255

REFERÊNCIAS BIBLIOGRÁFICAS 268

APÊNDICES 278

A: DIMENSÕES DO MODELO MILP 278

B: DISTÂNCIAS A MPT'S – Análise da Hipótese de Normalidade 281

1. INTRODUÇÃO

1.1 INTRODUÇÃO AO PROBLEMA

O problema estudado nesta pesquisa diz respeito à questão da localização de instalações. Esta é uma questão de ampla abrangência e de grande importância em termos estratégicos e táticos para a maior parte das organizações. Em função desta importância é que o tema tem sido amplamente estudado na literatura. Daskin (1995), por exemplo, declara explicitamente que o sucesso ou fracasso de empreendimentos privados ou públicos passa necessariamente pela questão da localização das instalações.

O autor trata da abrangência deste tipo de problema afirmando que a localização afeta praticamente todos os tipos de atividades dos setores público e privado. No caso industrial, por exemplo, Daskin lembra que as indústrias têm a necessidade de decidir sobre as localizações mais indicadas para suas unidades que devem levar em consideração o posicionamento de seus mercados e de seus fornecedores, disponibilidade de infraestrutura, mercado de mão de obra e vários outros fatores que de forma direta ou indireta podem afetar sua produtividade. O mesmo ocorre no caso de varejistas, onde o autor salienta que estes devem identificar os melhores posicionamentos para suas lojas de forma a atender o mercado, mas ao mesmo tempo possibilitar um abastecimento destes estabelecimentos dentro de padrões desejados de produtividade. Outras atividades em que a localização exerce papel central são os atacadistas e distribuidores, que devem decidir sobre os locais de seus centros de distribuição (CD's), que, a exemplo de centros de armazenagem de indústrias, servem de ponto intermediário entre fontes de abastecimento e regiões

de demanda que devem ser atendidas por estes CD's. No setor público o problema se repete, com escolas, garagens de combate a incêndio, hospitais, etc. É assim, uma questão realmente ampla, afetando uma variedade de setores nas mais diversas áreas da atividade econômica.

Olhando-se para a questão sob o ponto de vista de planejamento logístico, verifica-se que dentre as mais importantes decisões logísticas que as organizações devem tomar, pode-se considerar que em termos estratégicos a localização de instalações é a primordial, pois é a partir desta que toda a configuração da rede logística se define. Ballou (1994), considera assim, ao afirmar que a localização de instalações é a base estrutural que formata todo o sistema logístico.

A localização afeta de forma absoluta a produtividade das operações. Neste sentido, Randhawa e West (1995) salientam que a correta localização pode não só gerar significantes melhorias de produtividade, mas ainda possibilitar novos mercados e novos negócios, enquanto que localizações sub-ótimas podem gerar ineficiências em transporte, mão de obra inadequada e gastos adicionais de capital investido em instalações e operações.

Mas a localização afeta também outras questões importantes. Uma destas questões é um tema bastante atual, que é a crescente preocupação que as empresas têm com a satisfação dos clientes. A competição global tem levado as empresas a tratar a satisfação dos clientes como questão fundamental em suas estratégias e esta também está diretamente associada à questão da localização de instalações. Whybark e Yang (1996) comentam esta questão considerando o posicionamento de estoques, que podem estar em plantas industriais ou em centros de distribuição, como fator importante na melhoria do nível de serviço prestado aos clientes.

Localização de instalações é um problema, portanto, que se relaciona a várias questões que estão direta ou indiretamente associadas à produtividade das operações.

No que tange à atividade industrial, a localização assume aspecto de importância fundamental, pois está associada tanto a aspectos de abastecimento das plantas industriais, quanto à distribuição dos produtos acabados, passando ainda pela transferência de estoques entre fábricas e CD's e entre os próprios centros de distribuição.

Em relação à distribuição de produtos, pode-se considerar, na verdade, que toda a eficiência do processo está vinculada à configuração da rede logística. Uma vez definida a rede, todas as outras decisões logísticas serão dependentes desta configuração.

Neste sentido, salienta Coyle (1992), que uma vez que a rede logística esteja definida, a operação deverá ocorrer dentro das restrições que a localização das instalações impõe. Isto mostra que a configuração de uma rede logística é, dentre as decisões logísticas, aquela que vem primeiro e que, assim, determina todas as demais, tanto em termos estratégicos quanto táticos. Mas não bastasse isto, há ainda que se considerar que sua análise não é tarefa nada trivial, sendo multifacetada, pois são vários os fatores que interferem na decisão de localização (Bowersox e Closs, 2001 e Simchi-Levi et al., 2000).

Dada assim, a sua importância e estas características complexas de sua análise, deve-se imaginar que o problema da localização necessita de uma forte base de informações e de ferramentas quantitativas para suportar as decisões a tomar.

O problema da localização, na verdade, é resolvido através de vários métodos, porém, modelos de otimização, baseados em programação linear inteira mista, os

chamados modelos MILP (*mixed-integer linear programming*) são os mais empregados (Ballou, 1994). Estes modelos do tipo MILP correspondem a um problema de programação inteira que apresenta algumas variáveis que devem ser inteiras não negativas e outras que podem ser reais não negativas. Há também o caso particular de o modelo MILP ter variáveis binárias (com valor 0 ou 1), e isto ocorre com freqüência nos modelos aplicados a problemas de localização (Daskin, 1995). Um problema que exija somente variáveis inteiras não negativas é chamado de programação inteira pura (Winston, 1994). Quando o problema requer que as variáveis sejam exclusivamente binárias (0 ou 1) tem-se a programação "0-1" (Winston, 1994) ou programação binária (Erlich, 1982). A programação linear inteira mista (MILP) é assim um tipo de problema de programação inteira mais abrangente, já que admite qualquer destes tipos de variáveis.

Autores como Daskin (1995), Geoffrion e Powers (1995), Crainic e Laporte (1997), Owen e Daskin (1998) fizeram revisões dos principais métodos de otimização utilizados na resolução do problema e mostraram que os modelos matemáticos considerados baseiam-se em critérios de otimização que utilizam fatores de custos associados principalmente ao transporte, estoques e instalações de armazéns, entre outros.

O problema da localização da forma como é entendido neste trabalho pode ser estabelecido dentro dos seguintes termos:

- definir as localizações ótimas das instalações;
- definir a quantidade ótima de instalações;
- definir os fluxos de origem e destino de cada instalação e
- definir o porte de cada instalação.

Os modelos desenvolvidos para a solução deste problema apresentam, em certo sentido, dificuldades metodológicas que ainda não foram vencidas.

Os critérios de decisão tomam por base custos logísticos. Sendo, portanto, eventuais reduções nestes custos que indicam os melhores rumos para a operação em estudo. Note-se, assim, que todas as análises têm como pano de fundo a questão dos custos de uma operação logística. Desta forma, é uma exigência do próprio modelo uma detalhada análise, quantificação e definição de procedimentos tecnicamente bem embasados para se estabelecer estimativas confiáveis destes custos.

Estes custos logísticos incluem pelo menos os componentes de: armazenagem, movimentação de materiais, estoques em armazém, estoques em trânsito, transporte, distribuição, vendas perdidas e administração. As estimativas destes custos se constituem em um dos pilares de toda a análise a ser desenvolvida.

O problema da localização, apesar dos grandes avanços (Geoffrion e Powers, 1995) ainda se apresenta, na verdade, como um campo aberto à pesquisa, pois seja nos aspectos relativos à modelagem propriamente dita, seja em termos de parâmetros que são considerados nos modelos, existem aspectos importantes que ainda podem ser aprimorados (Geoffrion e Powers, 1995; Vidal e GoetschalcKx, 1997).

Possivelmente, o mais importante dos aspectos a considerar seja estudar a questão da geração de soluções viáveis para o problema. Pois estas é que irão se constituir na base para a determinação da solução ótima. Estas soluções viáveis, na verdade, são locais em que as instalações poderiam se situar. São assim, locais candidatos a instalações.

A maior parte dos modelos adota o procedimento de gerar uma solução ótima a partir de um conjunto de soluções viáveis; entretanto, estas soluções viáveis não

são em geral obtidas a partir de critérios objetivos. Autores como Ballou (1994), Bowersox e Closs (2001) e Randhawa e West (1995) apresentam vários fatores qualitativos que influenciam a definição das soluções viáveis, mas a literatura é muita escassa em termos de métodos quantitativos que conduzam a este conjunto de soluções viáveis. Ballou (1994) chega a declarar expressamente que as possíveis soluções para o problema são escolhidas por sua "razoabilidade". Se isto ocorre, a solução ótima do problema mesmo passando pelos mais avançados métodos de otimização, que por sua vez também não são perfeitos, se torna frágil, pois está sujeita à subjetividade da escolha das soluções viáveis.

Na verdade, duas etapas principais podem ser vislumbradas no processo de seleção destas soluções: a agregação dos pontos de demanda e a seleção de locais candidatos a localizações de instalações. Um destes candidatos será a solução ótima. Este é um desafio que se impõe para as pesquisas na área. Desenvolver procedimentos quantitativos que executem estas duas etapas, especialmente para a fase de seleção de candidatos, pois para procedimentos de agregação existem estudos sobre o assunto (Ballou, 1991b; Novaes e Graciolli, 1999, Francis et al., 1999), apesar de não serem especificamente dirigidos para a questão de localização de instalações. Muitas vezes, a finalidade dos processos de agregação é a definição de zonas de distribuição física de produtos, que são utilizadas para a montagem de roteiros de entrega.

Gerar, portanto, de forma adequada, um conjunto de soluções viáveis que se constituam na base para a escolha da solução ótima, é ponto chave do processo de modelagem e da conseqüente determinação da solução ótima.

Outro aspecto complexo a estudar é a questão de compatibilizar uma modelagem que represente adequadamente as relações existentes no processo e que ao

mesmo tempo atenda à necessidade de se estabelecer uma metodologia que conduza o sistema a um ponto de otimização.

Para isto, é preciso uma modelagem que seja sensível às várias compensações (*trade-offs)* de custos que ocorrem numa rede logística e ao tipo de relações matemáticas que se estabelecem entre os componentes de custos e as características da operação. Note-se, por exemplo, que a quantidade de CD's influi diretamente no nível global de estoque da operação, que tende a aumentar de forma não linear com o número de CD's (Fleury e Figueiredo, 2000). Os custos de transporte, por outro lado, tendem a decrescer com o número de CD's (Ballou, 1994). Há, portanto, todo um entrelaçamento de *trade-offs* que por sua complexidade, não se apresenta completamente equacionado na literatura, principalmente por não tratar completamente as evoluções não lineares destes custos. Estes aspectos, portanto, podem ser estudados, e uma vez aprimorados, irão representar um avanço metodológico.

É possível identificar assim, dois aspectos chaves do processo de otimização que podem ser aprimorados:

- a base de soluções viáveis que conduz à solução ótima e
- as relações matemáticas dos modelos de otimização.

O presente estudo irá focar sua atenção em um tipo de instalação que se constitui em ponto comum de decisão em vários tipos de empreendimentos, que é o centro operações logísticas de distribuição de carga ou simplesmente de centros de distribuição (CD's). O trabalho tem, assim, um interesse particular pela localização de CD's.

A pesquisa deverá estar voltada para uma rede logística de operações de transporte em longas distâncias, como é o caso típico de empresas atacadistas brasileiras e/ou

distribuidores que têm várias filiais espalhadas pelo país, podendo também ser aplicado para empresas de operação logística, e até mesmo, indústrias ou redes de varejo que possuem centros de distribuição em diferentes regiões.

Neste estudo se procurarão definir as posições e características dos CD's e os fluxos de carga entre as fontes de produção e os CD's, e entre estes e os pontos de demanda que passam a se constituir em nós de uma rede logística. Isto deve ser alcançado de forma a atender características impostas pelos níveis de serviço esperados para a operação. No caso de CD's, além de determinar sua quantidade e localização, também será estabelecida a carga de produção de cada CD.

É dentro desta linha que se pode especificar cada aspecto dos objetivos da pesquisa, conforme descritos na seção seguinte.

1.2 OBJETIVOS DA PESQUISA

A proposta de pesquisa a ser desenvolvida pode ser colocada em termos de que se entendendo a rede logística como um sistema, o projeto desta rede e o planejamento do seu carregamento serão feitos a partir da definição de seus componentes e estabelecendo-se funções e características de cada componente do sistema. Os componentes aqui são: centros de distribuição (CD's), as ligações entre CD's e origens/destinos de carga, e as magnitudes dos fluxos de carga em cada ligação. O ambiente deste sistema será composto pelas origens e destinos de carga e pelas restrições operacionais e de nível de serviço a serem impostas.

Esta questão, colocada desta forma, pode então ser conceituada como um *modelo de distribuição*, cujas características permitem considerá-lo como um dos tipos do problema da localização de instalações discutido acima.

Formalmente, então, o problema pode ser conceituado conforme segue:

Dado um plano de movimentação de carga, caracterizado pela sua oferta e demanda e pela localização dos centros de produção e centros de consumo deste produto, deve-se projetar a rede logística de distribuição.

A configuração desta rede se caracterizará por:

- *quantidade de CD's na rede logística;*
- *nós da rede em que estarão localizados os CD's;*
- *estoque a ser alocado a cada CD;*
- *fluxo de carga de cada centro de produção para cada CD;*
- *fluxo de carga de cada CD para cada centro de consumo.*

O objetivo central da pesquisa ao tratar este problema é o estudo de dois aspectos chaves da modelagem:

- *aprimorar a base de soluções viáveis que conduz à solução ótima e*
- *viabilizar uma metodologia que possibilite relações matemáticas mais realistas nos processos de otimização.*

Sobre o primeiro tópico, a pesquisa procurará desenvolver um procedimento que trabalhe de forma quantitativa a base de dados dos nós de demanda para estabelecer uma agregação destes nós e um segundo procedimento para seleção de nós, possibilitando assim a geração de um conjunto de soluções viáveis (localizações candidatas a CD's) que irão se constituir na base para a escolha da solução ótima do problema. Esta base de dados deverá, portanto, ser trabalhada de

forma a que possibilite a escolha das “melhores” soluções viáveis (“melhores” locais candidatos a CD’s).

Sobre o segundo tópico a pesquisa deverá reestruturar a modelagem de forma que seja possível se levar em consideração relações matemáticas que eventualmente apresentem outros tipos de comportamento não necessariamente linear, possibilitando assim, representações mais realistas do problema.

A pesquisa procurará, portanto, avançar em termos de modelagem do problema de forma a que melhores soluções sejam obtidas.

Espera-se, desta forma, que o desenvolvimento desta pesquisa possa frutificar em contribuições às metodologias existentes de análise e planejamento de uma rede logística.

A figura 1.1 a seguir dá uma visão geral do escopo do trabalho:

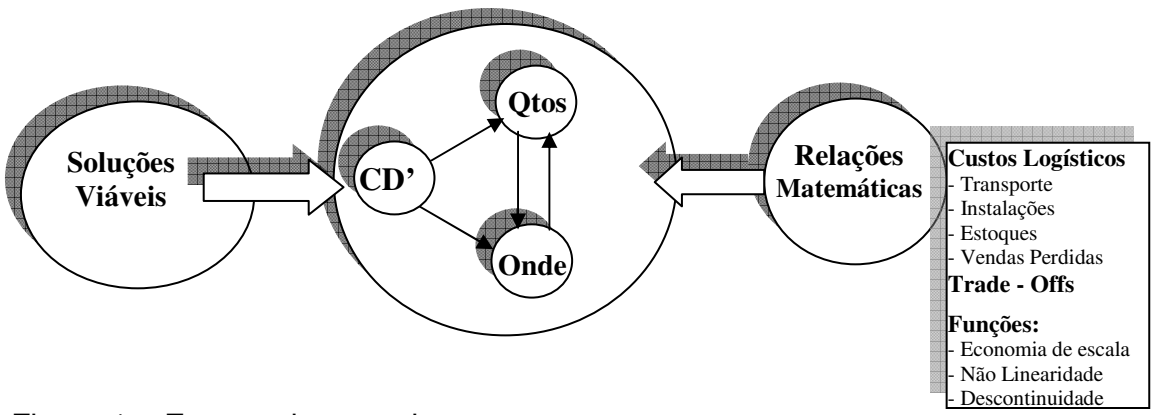

Figura 1 – Escopo da pesquisa

2. REVISÃO BIBLIOGRÁFICA

A análise de localização tem sua importância reconhecida desde meados do século XIX quando foi publicado o trabalho *The isolated state*, onde o economista alemão Joachim von Thünen estabelecia que os principais fatores que determinavam o desenvolvimento econômico eram o preço da terra e o custo do transporte entre os locais de produção e consumo (Bowersox e Closs, 2001).

Em termos, entretanto, de introdução formal da teoria da localização na literatura considera-se que de fato ocorreu (Brandeau e Chiu, 1989; Owen e Daskin, 1998) através de Alfred Weber, em 1909, que analisou como posicionar um depósito de forma a minimizar a soma das distâncias de um conjunto de clientes para esta instalação (Weber, 1929). Weber tratou ainda da questão da mão de obra como fator locacional e do fator aglomerativo, que considerou como sendo a redução de custo que uma indústria obtém ao se localizar nas proximidades de outras empresas do mesmo ramo de atividade.

Walter Isard (1956) tentou unir algumas discussões teóricas, analisando os trabalhos de Weber, Von Thünen e August Lösch, que também havia estudado a questão da localização, tendo como principal foco a organização espacial do sistema econômico. Lösch admitia que a maximização de lucros dirigia a escolha locacional. Isard analisou estas três correntes, gerando um modelo geral de minimização de custos que considerou o custo de transporte como principal fator de escolha de um local. Segundo Isard, ao se escolher um local posicionado a meio caminho entre uma fonte de matéria prima e seu mercado, a empresa na realidade está decidindo sobre a quantidade de insumo a ser consumida com matéria prima e produto. É assim, na verdade, uma questão de substituição de fatores.

2.1 RELEVÂNCIA DO PROBLEMA DA LOCALIZAÇÃO

Apesar da teoria da localização ter tido o seu início em 1909, poucas aplicações foram realmente desenvolvidas até meados da década de 1960. Foi somente nesta época que a questão ganhou interesse e uma série grande de trabalhos começou a ser desenvolvida, continuando a merecer a atenção de pesquisadores até o presente momento (Owen e Daskin, 1998).

A relevância do estudo de localizações é constatada hoje pelo expressivo número de aplicações práticas que são encontradas nas mais diversas áreas de atividade. O trabalho de Brandeau e Chiu (1989) ilustra esta grande aplicabilidade, através de uma pesquisa em que mostram uma vasta relação de aplicações, tanto na área privada quanto pública. Exemplos da área privada são: depósitos e centros de produção; terminais de transporte e centros logísticos de distribuição de carga; plantas industriais; redes de comunicações; centrais elétricas; centrais de táxi. Na área pública tem-se: centrais de veículos de emergência (ambulâncias, bombeiros, etc.); centros de serviços públicos em geral, tais como: centros de saúde, bancos de sangue, centrais de tratamento de lixo, centrais de tratamento de água; localização de bairros residenciais e até mesmo, instalações de defesa.

Outro aspecto a destacar, que como será visto é salientado de forma unânime na literatura, é a importância estratégica da questão da localização de instalações.

Crainic e Laporte (1997) conceituam este aspecto esclarecendo que as decisões estratégicas determinam, na verdade, o desenvolvimento geral das políticas de uma empresa e formatam de modo abrangente as estratégias operacionais do sistema sobre um relativamente longo horizonte de tempo e, dentre estas decisões estratégicas, está a localização de instalações.

Neste sentido, Owen e Daskin (1998), por exemplo, afirmam que a localização de instalações é um aspecto estratégico crítico em uma grande variedade de atividades públicas e privadas e acrescentam ainda, que a decisão sobre uma nova instalação é tipicamente um projeto sensível ao tempo e ao custo. Os elevados custos envolvidos em projetos de instalações fazem, na verdade, com que quase todos os projetos do gênero se caracterizem como investimentos de longo prazo. Espera-se assim, que localizações de instalações que são definidas hoje permaneçam em operação por um longo tempo. É assim, portanto, decisão estratégica vital.

Daskin (1995) em seu texto que trata exclusivamente da localização de instalações ,coloca que decisões pobres neste campo levam a processos de aumentos de custos e, conseqüentemente, a decréscimos de competitividade. O autor considera que apesar de não ser o único fator, necessariamente o sucesso ou fracasso de instalações públicas ou privadas, sempre irá depender da escolha para a localização dessas instalações.

Uma vez que se toma uma decisão de localização, a operação logística irá se desenvolver dentro do conjunto de restrições impostas por aquela localização. É, portanto, uma decisão estratégica que afeta todas as demais que vêm em seguida. Se for uma decisão acertada, todo o sistema logístico será beneficiado, mas se não for uma boa decisão, a penalidade sobre o sistema poderá se estender por um longo período de tempo (Coyle et al., 1992).

Ainda sobre a importância do problema, Ballou (1994) afirma que é esta decisão de localização de instalações que dá forma e estrutura a todo o sistema logístico, e complementa, lembrando que é esta decisão que irá determinar o conjunto de alternativas que estarão disponíveis para a operação do sistema, e que cada

alternativa terá o seu respectivo conjunto de custos associados, podendo, portanto, afetar para cima ou para baixo a performance de todo o sistema logístico.

Ballou ainda afirma que a localização de instalações pode ser considerada como o mais importante problema de planejamento logístico estratégico em um grande número de empresas. É esta decisão que estabelece as condições para a seleção apropriada e a boa administração dos serviços de transporte e dos níveis de estoque.

A importância estratégica da questão é, assim, amplamente reconhecida na literatura.

2.2 CONTORNOS DO PROBLEMA

Os contornos deste problema necessitam de uma definição clara, pois apesar de dito problema de *localização*, as decisões envolvem não somente a determinação de um local, mas também o número e dimensão dessas instalações (Ballou, 1994). Estas instalações, do ponto de vista de modelagem, são na verdade nós de uma rede (Boaventura, 1977; Novaes, 1989), onde ocorrem operações logísticas. Estes nós podem representar plantas industriais, pontos de varejo, depósitos, centros de distribuição ou de abastecimento e centros de serviço em que os produtos e/ou materiais ficarão temporariamente estacionados e/ou estarão sendo processados em seu caminho para o consumidor final.

Esta definição, apesar de abordar os principais aspectos da questão, não mostra, entretanto, todas as facetas do problema, que apresenta muitos aspectos que levam a um nível elevado de complexidade. O próprio Ballou (1994) já apresenta aspectos que envolvem o problema e que demonstram sua não trivialidade. O número de

instalações é um primeiro tópico a ser contemplado (definir um local é muito mais simples do que múltiplos locais). Localização discreta, em que se têm apenas pontos discretos em uma rede para escolher os locais, é uma outra questão a tratar, que é bem diferente quando se considera um espaço contínuo. O grau de agregação dos dados é outro item importante a analisar, pois quando se têm muitos pontos de demanda é operacionalmente inviável trabalhar com os pontos individuais e, neste caso, uma agregação em algum nível será necessária. O horizonte de tempo considerado é outro aspecto importante, pois a natureza dos problemas pode ser estática ou dinâmica. Nos problemas estáticos as localizações são determinadas a partir de dados baseados em um único período bem definido no tempo. Já nos problemas dinâmicos os planos de localização podem cobrir um extenso período de tempo em que as instalações podem inclusive se movimentar de um local para outro. A natureza determinística dos fatores envolvidos no problema é outro ponto importante a tratar. Muitos parâmetros podem ter comportamento estocástico, como: tempos de viagem e demandas, que são fatores importantes no processo de localização.

Como se vê, o problema tem muitos ângulos e aspectos a considerar.

Brandeau e Chiu (1989) procuraram analisar bem a evolução das pesquisas desenvolvidas na área e pelo seu artigo é possível se ter uma visão ampla destas características multifacetadas do problema da localização. Este artigo é na verdade um clássico da literatura, pois os autores procuraram conceituar de forma clara o problema e em seguida desenvolveram uma taxonomia para distinguir os diversos tipos de modelos existentes. Esta taxonomia é baseada em três categorias: o objetivo do problema, que pode ser de otimização ou de não otimização; as variáveis de decisão consideradas no modelo e os parâmetros do sistema, como a

topologia da rede, métricas consideradas, tipos de custos, tipos de demanda, etc. A partir desta taxonomia os autores analisaram e classificaram mais de cinqüenta problemas referenciados na literatura.

Owen e Daskin (1998) também desenvolveram um trabalho abrangente sobre o problema da localização, procurando analisar não só os problemas mais comumente analisados na literatura, mas também contemplando aspectos associados à evolução temporal e às incertezas. No primeiro grupo estão os problemas estáticos e determinísticos e no segundo grupo estão os problemas de localização dinâmica, que procuram capturar os aspectos temporais dos problemas do mundo real. Além disto, os autores analisam ainda a programação estocástica e o planejamento de cenários.

Nos problemas de localização dinâmica procura-se determinar um padrão ótimo de localização ao longo do tempo. Como os comportamentos da demanda e de custos da operação se modificam ao longo do tempo, implementar a solução de um modelo de localização baseado nas condições de uma dada posição do tempo pode se tornar uma solução sub-ótima em algum momento futuro (Ballou, 1994). Note-se assim, que considerando a variação temporal de seus parâmetros, estes problemas não têm uma visão estática da questão, mas por outro lado, ainda adotam uma abordagem determinística, lidando com incertezas no futuro, mas admitindo que se tem conhecimento preciso dos valores atuais dos parâmetros e que se há variação no futuro, esta variação será determinística. É o que se pode chamar de modelos de localização dinâmica determinística (Owen e Daskin, 1998).

Já na programação estocástica considera-se que se tem um conhecimento limitado de parâmetros do modelo. Pode-se ter incerteza sobre parâmetros como: tempos de viagem, custos, localização de demandas e nível de demandas. O objetivo aqui é

determinar localizações que irão ter boa performance (de acordo com um critério definido) sob um número de possíveis valores dos parâmetros. Dois tipos de abordagem são adotados: modelos probabilísticos, que consideram as distribuições de probabilidade das variáveis aleatórias associadas ao problema, e planejamento de cenários, em que são gerados conjuntos de possíveis valores para as variáveis do problema, caracterizando-os como cenários futuros a analisar (Owen e Daskin, 1998).

O foco desta pesquisa está nos problemas estáticos e determinísticos, que representam a grande maioria das aplicações práticas e são classificados em três grandes grupos considerados clássicos da literatura (Tondo, 1992; Daskin, 1995; Crainic e Laporte, 1997; Owen e Daskin, 1998): os problemas das medianas (*median problems*), os problemas de cobertura (*covering problems*) e os problemas do centro (*center problems*).

O primeiro tipo foi introduzido por Hakimi (1964) e pode ser enunciado da seguinte forma: *encontrar a localização de p instalações de forma a minimizar a distância total, ponderada pelas demandas, entre os pontos de demanda e as instalações*. Este é o conhecido problema das p-medianas.

Se o número p de medianas é uma variável, tem-se dois outros tipos de problemas: aquele com instalações sem restrição de capacidade (*uncapacited plant location problem – UPLP*) e aquele com instalações com restrição de capacidade (*capacited plant location problem – CPLP*).

No segundo grupo chamado de problema de cobertura, uma demanda é dita *coberta* quando pode ser atendida dentro de um *tempo especificado* (Owen e Daskin, 1998). Estes problemas são divididos em dois segmentos principais:

um em que um nível de cobertura é requerido (*location set covering problem*) e um outro em que se busca otimização da cobertura (*maximum covering problem*). Nas duas situações o problema faz uso de uma distância especificada de cobertura exógena ao modelo. Já no terceiro grupo, denominado de problema p-centro (Tondo, 1992) também se busca uma cobertura de demandas. Neste caso, porém, o objetivo é localizar um conjunto de p-instalações para garantir a cobertura dos pontos de demanda de forma que a distância total de cobertura seja minimizada. Portanto, ao invés de receber como dado de entrada uma distância especificada, o modelo determina endogenamente a distância mínima de cobertura associada a p-instalações.

Note-se que todos os problemas apresentados até aqui não diferenciam fontes de produção de pontos de demanda.

Quando se considera a diferenciação entre nós de produção e nós de demanda, entra-se no campo de uma outra categoria de modelos. São os chamados modelos de produção - distribuição (*production – distribution models*). Este tipo de problema, na verdade, vai além da localização de instalações, abordando todas as etapas de uma cadeia de suprimento: abastecimento, produção e programação, estoques, centros de distribuição e transporte (Vidal e Goetschalckx, 1997). A análise de localização de instalações considerando-se fontes de produção e pontos de demanda seria uma abordagem parcial do problema de produção – distribuição.

2.3 MÉTODOS DE SOLUÇÃO

Os métodos para solucionar o problema da localização são variados. As abordagens mais encontradas na literatura são a programação linear inteira mista e

procedimentos heurísticos, que na maior parte têm por base um modelo de programação linear. Mas as técnicas vão desde métodos matemáticos exatos que são aplicados a problemas mais simples e que não garantem otimização, passando por modelagem mais avançada como a técnica de simulação e lógica fuzzy, até a programação matemática e heurísticas em que a programação linear inteira mista e métodos heurísticos derivados e/ou complementares são as abordagens mais utilizadas.

Iniciando pelos problemas mais simples, pode-se falar em métodos exatos, que tratam da localização de uma única instalação que deve atender a um conjunto de localidades. Este tipo de método não diferencia pontos de demanda de pontos de produção e ainda considera um espaço contínuo podendo, portanto, localizar a instalação em qualquer ponto do plano. O método do centro de gravidade (Novaes, 1989; Ballou, 1994; Bowersox e Closs, 2001) é o caso de um método exato que procura minimizar a soma das distâncias dos pontos de demanda à instalação. As distâncias são ponderadas pelas demandas (ou custos). Outras abordagens para o problema incluem técnicas gráficas e até mesmo, métodos aproximados (Ballou, 1994).

Um caso mais complexo de método exato é o caso de localizar múltiplas instalações. O chamado método do múltiplo centro de gravidade trata esta questão e o faz de forma análoga ao caso de uma única instalação.

Modelagem por simulação (Law e Kelton, 1982) é uma outra abordagem empregada para solucionar o problema. Uma grande vantagem desta técnica é que ela permite uma descrição apurada do sistema, permitindo considerar um grande número de variáveis, possibilitando assim um tratamento mais realista de seu comportamento, mesmo que este comportamento apresente características complexas, como

flutuações estocásticas e dinâmicas. A desvantagem é que a simulação não garante a otimização do processo e, além disto, a construção de um modelo de um sistema complexo pode ser demorada. Mas, mesmo assim, um modelo que não garanta a otimização, mas tenha uma boa descrição do problema real pode ser mais vantajoso do que um modelo de otimização em que se é obrigado a simplificar certos aspectos da questão devido à complexidade matemática que a sua consideração integral acarretaria para a modelagem e, conseqüentemente, para o método de solução. Em problemas de grandes dimensões, como, por exemplo, com milhares de clientes, dezenas de centros de distribuição e de plantas industriais, modelos de simulação podem tratar a questão sem necessitar de recursos adicionais, como algoritmos especiais, ou máquinas com capacidades excepcionais de processamento. Em outras abordagens de solução isto não seria viável sem métodos e algoritmos especiais e em muitas situações, mesmo assim, não há capacidade de processamento disponível para a solução do problema (Ballou, 1994; Geoffrion e Powers, 1995).

A lógica fuzzy, outra abordagem que é empregada em alguns casos, parte do princípio de que devem ser considerados em certas situações critérios qualitativos e que estes devem ser escritos em termos lingüísticos (baixo, médio, muito baixo, alto, muito alto, etc.) tanto em termos de seus valores quanto de seus pesos relativos. Para isto são então definidas variáveis lingüísticas que permitem tratar estes aspectos qualitativos (Chen, 2001). Esta técnica, entretanto, apesar de procurar ser abrangente em termos de tratamento de variáveis, não é uma abordagem que ganhou muito impulso no tratamento do problema da localização (Crainic e Laporte 1997; Geoffrion e Powers, 1995; Owen e Daskin, 1998).

A programação linear como método de solução do problema é uma técnica muito citada na bibliografia. Uma abordagem deste tipo, baseada no conhecido problema do transporte com transbordo, é apresentada por Novaes (1989), que mostra que um modelo de transbordo pode ser utilizado para localização de uma instalação na presença de outras já em operação. O autor sugere que diferentes localizações podem ser testadas com o modelo, identificando-se aquela que apresenta o menor custo total de transporte.

A técnica, entretanto, que apresenta o maior numero de trabalhos é a programação inteira, na sua forma pura ou mista, (Brandeau e Chiu, 1989; Crainic e Laporte 1997; Geoffrion e Powers, 1995; Owen e Daskin, 1998). Na programação mista os chamados modelos do tipo MILP (*mixed-integer linear programming*) apresentam algumas variáveis que devem ser inteiras não negativas (podendo ser 0-1) e outras que podem ser reais não negativas, conforme visto no capítulo 1.

Os três tipos clássicos de problemas estáticos e determinísticos: p-medianas, cobertura e do centro, por exemplo, que representam a grande maioria das aplicações práticas, estão nessas categorias. No caso das p-medianas, tanto os problemas de medianas com p variável com restrição de capacidade e sem restrição de capacidade (UPLP e CPLP), como os problemas de p fixo, têm métodos de solução baseados em programação inteira.

Esses tipos de modelos apresentam um nível de complexidade no seu algoritmo que eleva de forma exponencial os tempos necessários para se obter uma solução, inviabilizando o uso direto da técnica com os seus métodos de solução clássicos em problemas de grande porte. São os chamados problemas do tipo *NP-completo* (Daskin, 1995). A utilização direta da técnica fica assim restrita a problemas menores, em que o número de variáveis e/ou restrições não é muito elevado. Na

prática, porém, os problemas podem apresentar um número muito grande de variáveis ou de restrições (milhares ou milhões). Assim, métodos alternativos acabam sendo empregados. Estes métodos muitas vezes não descartam os modelos de otimização, mas, na verdade, utilizam o modelo como parte do processo de solução.

Os métodos heurísticos, que tomam como ponto de partida um modelo de programação linear inteira mista são, em geral, os mais utilizados na solução dos problemas de localização (Ballou, 1994; Crainic e Laporte 1997; Daskin, 1995). Essas heurísticas partem de um modelo de otimização e procuram introduzir simplificações e/ou subdivisões no problema, facilitando assim sua solução. As heurísticas podem, ainda, aprimorar a solução com a introdução de novos aspectos não considerados no modelo de otimização, o que torna a modelagem mais próxima do mundo real, apesar de não garantir que a solução encontrada é ótima.

Estes métodos têm sido empregados nos problemas de produção – distribuição, que se constituem em problemas com elevado número de variáveis.

Um representante deste tipo de problema é apresentado por Geoffrion e Graves (1974) em um artigo que é considerado um marco na literatura.

Vidal e GoetschalcKx (1997) que fizeram uma extensa revisão sobre o tema ,apresentam aqueles artigos que consideram como os mais relevantes e partem a discussão justamente deste trabalho de Geoffrion e Graves, que foi o primeiro modelo de programação linear inteira mista realmente importante na área. Mas, na verdade, todos os autores pesquisados que tratam de modelos de programação linear inteira mista envolvendo a distribuição de carga, sem exceção, citam o modelo de Geoffrion e Graves para a solução de problemas de distribuição (Ballou, 1994; Brandeau e Chiu, 1989; Geoffrion e Powers, 1995; Vidal e GoetschalcKx,

1997; Owen e Daskin, 1998; Erenguç et all., 1999). O modelo de Geoffrion e Graves é considerado um modelo de distribuição, pois trata parcialmente o problema da produção – distribuição e até hoje é considerado importante para a solução de problemas de distribuição de produtos.

A solução do modelo foi baseada na técnica de decomposição de J. F. Bender e apresentada por Daskin (1995), que em linhas gerais fixa certas variáveis binárias temporariamente e resolve um problema principal (*master problem*), que pode ser tratado como um conjunto de "problemas do transporte", um para cada produto.

Este método vem sendo hoje substituído por outros métodos (Geofrion e Powers, 1995), como a *Fatorização* (Brown e Olson, 1994), que é um procedimento que isola estruturas especiais no *tableau* de programação linear (Maculan e Pereira, 1980; Winston, 1994). Fazendo uma *partição* da matriz de programação linear (PL) uma parte considerável do *tableau,* acaba sendo representada de forma implícita, sendo uma outra parte explícita. Na verdade, Brown e Olson (1994) fazem uma distinção entre *partição* e *fatorização*. Consideram que métodos de *partição* são baseados em estruturas especiais do próprio problema, enquanto que *fatorização* é baseada em estruturas especiais das *bases* da matriz de PL e, portanto, no *tableau* básico. Um exemplo de *fatorização* que analisam é quando a matriz *base* da PL pode ser particionada em duas: uma grande matriz que passa por atualizações muito pouco freqüentes e uma pequena matriz com atualizações regulares. Com isto, conseguem-se grandes reduções nos tempos de processamento.

Mas, independentemente dos novos métodos, o modelo de Geoffrion e Graves permanece em uso, indicando a sua importância para a solução de localização de instalações no caso de distribuição de produtos.

As estruturas dos modelos citados nesta seção são discutidas agora na seção seguinte.

2.4 DESCRIÇÃO DOS MODELOS MATEMÁTICOS DE LOCALIZAÇÃO

Considerando-se inicialmente os modelos mais simples, pode-se pensar no método exato do centro de gravidade (Novaes, 1989 e Ballou, 1994), que se destina à localização de uma única instalação que irá atender a um conjunto de pontos. O modelo em princípio não diferencia fontes de produção (pontos de origem de carga) das regiões de demanda (destino de carga), mas isto pode ser feito pelo tratamento adequado dos custos de transporte que serão utilizados no processo. O espaço considerado no modelo é um plano contínuo. Na aplicação do método determina-se inicialmente a posição do centro de gravidade de todo o conjunto de pontos (que podem ser pontos de produção ou de demanda de carga). A localização assim definida constitui-se em uma posição inicial para localizar a instalação (equações 2.1a e 2.1b). Em seguida, calcula-se a distância de cada ponto para esta localização inicial e através de iterações sucessivas vai-se determinando novas posições para a instalação considerando-se as distâncias aos pontos de produção e/ou demanda (equações 2.2a e 2.2b).

Passo 1: Coordenadas do Centro de Gravidade

$$X = \frac{\sum_{i=1}^{n} c_i . D_i . X_i}{\sum_{i=1}^{n} c_i . D_i} \qquad (2.1a) \qquad\qquad Y = \frac{\sum_{i=1}^{n} c_i . D_i . Y_i}{\sum_{i=1}^{n} c_i . D_i} \qquad (2.1b)$$

Passo 2: Coordenadas do CD

$$X^k = \frac{\sum_{i=1}^{n} c_i . D_i . X_i \Big/ d_i^{k-1}}{\sum_{i=1}^{n} c_i . D_i \Big/ d_i^{k-1}} \quad (2.2a) \qquad Y^k = \frac{\sum_{i=1}^{n} c_i . D_i . Y_i \Big/ d_i^{k-1}}{\sum_{i=1}^{n} c_i . D_i \Big/ d_i^{k-1}} \quad (2.2b)$$

Nesta formulação tem-se:

X = coordenada X do centro de gravidade

Y = coordenada Y do centro de gravidade

c_i = custo unitário de transporte na região i

D_i = demanda a atender a região i

X_i = coordenada X da região i

Y_i = coordenada Y da região i

X^k = coordenada X do CD na iteração k

Y^k = coordenada Y do CD na iteração k

d_i^{k-1} = distância da região i ao CD na iteração K-1

i= 1, 2, ...n

As iterações são interrompidas quando:

$$|X^k - X^{k-1}| / X^{k-1} \leq \varepsilon \quad e \quad |Y^k - Y^{k-1}| / Y^{k-1} \leq \varepsilon \qquad (2.3)$$

com ε tão pequeno quanto se queira.

Passando-se para os modelos de otimização, não se pode deixar de considerar os modelos baseados em programação linear, que representam uma parte considerável das aplicações práticas (Novaes, 1989; Ballou, 1994; Crainic e Laporte, 1997). Um exemplo que ilustra bem a aplicação desta técnica é o modelo de transbordo analisado por Novaes (1989) e apresentado a seguir.

$$Mín \sum_i \sum_k c_i . F_{ik} . d_{ik} + \sum_k \sum_j c_j . F_{kj} . d_{kj} \tag{2.4}$$

Sujeita a:

$$\sum_K F_{ik} \leq O_i \qquad (i = 1,2,.....N) \tag{2.5}$$

$$\sum_K F_{kj} \geq D_j \qquad (j = 1,2,.....M) \tag{2.6}$$

$$\sum_i F_{ik} = \sum_j F_{kj} \qquad (k = 1,2,.....ND) \tag{2.7}$$

$$F_{ik} \geq 0 \ , \ \forall i,k \tag{2.8}$$

$$F_{kj} \geq 0 \ , \ \forall k,j \tag{2.9}$$

onde: c_i = custo unitário de transporte por peso.km a partir da fonte de produção i;

c_j = custo unitário de transporte por peso.km para o ponto de demanda j;

F_{ik} = fluxo de transporte da fonte de produção i para o centro de distribuição k;

F_{kj} = fluxo de transporte do centro de distribuição k para o ponto de demanda j;

d_{ik} = distância da fonte de produção i ao centro de distribuição k;

d_{kj} = distância do centro de distribuição k ao ponto de demanda j;

O_i = oferta na fonte de produção i;

D_j = demanda no ponto de demanda j;

N = número de fontes de produção;

M = número de pontos de demanda;

ND = número de centros de distribuição.

A função objetivo (2.4) representa os custos globais de transporte das origens para os centros de distribuição (CD) e destes para os pontos de demanda. Em (2.5) e

(2.6) tem-se as restrições de oferta e de demanda. Em (2.7) tem-se as equações de continuidade nos CD's e em (2.8) e (2.9) tem-se as restrições de não negatividade dos fluxos de transporte.

Com este modelo é possível o teste de diferentes posições para um novo CD. Tendo-se ND-1 centros de distribuição, pode-se testar a introdução de um novo CD rodando-se o modelo para cada posição candidata.

Este modelo, portanto, parte de um conjunto de candidatos e define todos os fluxos de transporte que devem ocorrer para minimizar o custo global de transporte. O modelo não permite que se defina a quantidade de CD´s que se deseja, isto é determinado endogenamente pelo próprio modelo. Requer ainda, que sejam conhecidas as condições de oferta e demanda.

Considerando-se agora os modelos de programação inteira, que são os mais freqüentemente empregados, deve-se iniciar pelo modelo utilizado para a solução do problema das p-medianas, que é largamente utilizado.

Este modelo discutido por Daskin (1995) é apresentado abaixo:

$$Mín \sum_i \sum_j D_j . d_{ij} . Y_{ij} \qquad (2.10)$$

Sujeita a:

$$\sum_i Y_{ij} = 1 \qquad (j = 1,2,.....n) \qquad (2.11)$$

$$\sum_i X_i = p \qquad (2.12)$$

$$Y_{ij} - X_i \leq 0 \qquad (i = 1,2,....N \ ; \ j = 1,2,.....n) \qquad (2.13)$$

$$Y_{ij} \in \{ 0,1 \} \ , \ \forall i, j \qquad (2.14)$$

$$X_i \in \{ 0,1 \} \ , \ \forall i \qquad (2.15)$$

onde: D_j = demanda na região representada pelo nó j

d_{ij} = distância ao nó j da instalação localizada em i;

$$Y_{ij} = \begin{cases} 1 \ \textit{se a instalação no nó } i \textit{ atende a região } j \\ 0 \ \textit{em caso contrário} \end{cases}$$

$$X_i = \begin{cases} 1 \ \textit{uma instalação foi ativada no nó } i \\ 0 \ \textit{em caso contrário} \end{cases}$$

A função objetivo (2.10) representa a soma dos momentos de transporte associados a todos os pontos de demanda, que devem ser minimizados.

As restrições (2.11) buscam garantir que cada região seja atendida por uma única instalação. Em (2.12) estabelece-se que o número de instalações deve ser igual a "p". As condições (2.13) determinam que as regiões só podem ser atendidas por uma instalação ativa. Em (2.14) faz-se a ligação entre as instalações e as regiões sendo atendidas. Finalmente, em (2.15), através da variável X_i, define-se se uma instalação é ativa ou não.

Este já é um modelo que pode definir um certo número de instalações em nós de uma rede. Isto é possível graças à programação inteira, que com o uso de variáveis binárias permite que se definam quais os nós que contêm instalações ativas e também quais pontos são atendidos por determinadas instalações.

Este modelo tem aplicação para a definição de uma ou mais instalações (p instalações) que irão atender a um número n de pontos de demanda e/ou produção. O modelo não distingue, entretanto, entre as fontes de produção e os pontos de demanda. Todos os pontos são tratados de forma única. Não há a caracterização de fluxos de abastecimento e fluxos de distribuição.

Outro problema importante citado na literatura (Owen e Daskin, 1998) é o problema da cobertura. Este problema tem dois tipos principais: um primeiro em que se deve atender a uma cobertura requerida e um outro em que se otimiza o nível de cobertura.

O modelo utilizado para a solução do primeiro tipo de problema de cobertura ("*set covering problem*") é apresentado a seguir. Neste caso, o objetivo é minimizar o custo de localização de forma a atender um certo nível de cobertura.

$$Mín \sum_i c_i . X_i \tag{2.16}$$

Sujeita a:

$$\sum_{i \in N_j} X_i \geq 1 \ , \ \forall j \tag{2.17}$$

$$X_i \in \{ 0,1 \} \ , \ \ \forall i \tag{2.18}$$

onde: c_i = custo fixo de localizar uma instalação no nó i;

S = distância de atendimento máxima aceitável (ou tempo);

N_j = conjunto de locais dentro da distância S do nó j (i.e., { Nj = { i | $d_{ij} \leq S$});

$$X_i = \begin{cases} 1 \ uma\ instalação\ foi\ ativada\ no\ nó\ i \\ 0 \ em\ caso\ contrário \end{cases}$$

Neste modelo a função objetivo (2.16) corresponde ao custo da localização das instalações. As restrições (2.17) exigem que todos os nós de demanda tenham ao menos uma instalação situada dentro da distância máxima de atendimento. As restrições restantes (2.18) referem-se às variáveis de decisão que devem ser binárias.

Este modelo permite que se verifique o número necessário de instalações para se atingir um nível desejado de cobertura. Entretanto, em muitas aplicações práticas verifica-se que o orçamento não será suficiente para a implementação de todas as instalações definidas pelo modelo (Owen e Daskin,1998). Surge então o segundo tipo de problema, que agora procura otimizar a cobertura. A otimização é feita no sentido de se utilizar os recursos disponíveis, localizando-se um número fixo de

instalações e buscando-se maximizar a demanda atendida dentro da distância máxima de atendimento. Este problema é conhecido como o problema da maximização da cobertura (*"maximal covering problem"*)

$$Max \sum_{j} D_j . Z_j \tag{2.19}$$

Sujeita a:

$$Z_j \leq \sum_{i \in N_j} X_i \quad , \quad \forall j \tag{2.20}$$

$$\sum_{i} X_i \leq p \tag{2.21}$$

$$X_i \in \{ 0,1 \} \ , \quad \forall i \tag{2.22}$$

$$Z_j \in \{ 0,1 \} \ , \quad \forall j \tag{2.23}$$

onde: D_j = demanda no nó j;

$$Z_j = \begin{cases} 1 \ se\ o\ nó\ j\ é\ coberto \\ 0\ em\ caso\ contrário \end{cases}$$

$$X_i = \begin{cases} 1\ se\ uma\ instalação\ foi\ ativada\ no\ nó\ i \\ 0\ em\ caso\ contrário \end{cases}$$

N_j = conjunto de locais dentro da distância S do nó j (i.e., { Nj = { i | $d_{ij} \leq S$});

Neste modelo a função objetivo (2.19) se refere à demanda coberta e que deve ser maximizada. Em relação às restrições (2.20), note-se que cada nó j só pode ser considerado coberto (Z_j = 1) se existe uma instalação localizada no nó i e que deve estar dentro da distância S do nó j (X_i = 1 para j $\in$ N_j). Se esta instalação não existe, então o lado direito da restrição (2.20) será zero, forçando Z_j a se anular. As restrições (2.20), portanto, determinam que os nós sejam cobertos dentro da

distância máxima aceitável de atendimento. A restrição (2.21) limita o número de instalações, considerando, portanto, limitações de recursos e as restrições (2.22) e (2.23) se referem às variáveis binárias.

O terceiro tipo clássico de modelo de programação linear inteira mista aplicado a problemas de localização é o problema do centro (*"center problem"*), apresentado abaixo:

$$Mín\ D \tag{2.24}$$

Sujeita a:

$$\sum_i X_i = p \tag{2.25}$$

$$\sum_i Y_{ij} = 1\ ,\ \ \forall i \tag{2.26}$$

$$Y_{ij} - X_i \leq 0\ \ ,\ \ \forall\ i, j \tag{2.27}$$

$$D\ \geq\ \sum_i d_{ij}.\ Y_{ij}\ \ ,\ \ \forall\ j \tag{2.28}$$

$$Y_{ij} \in \{\ 0,1\}\ \ ,\ \forall i, j \tag{2.29}$$

$$X_i\ \in \{\ 0,1\ \}\ ,\ \ \forall\ i \tag{2.30}$$

onde:

d_{ij} = distância da instalação i ao nó j

$$Y_{ij} = \begin{cases} 1\ \ se\ a\ instalação\ no\ nó\ \ i\ atende\ a\ região\ \ j \\ 0\ em\ caso\ contrário \end{cases}$$

$$X_i\ = \begin{cases} 1\ uma\ instalação\ \ foi\ ativada\ no\ nó\ i \\ 0\ em\ caso\ contrário \end{cases}$$

Neste caso, a função objetivo (2.24) simplesmente representa a distância máxima entre qualquer nó de demanda e sua instalação mais próxima.

Em (2.25) estabelece-se que o número de instalações deve ser igual a “p”. As restrições (2.26) buscam garantir que cada região seja atendida por uma única instalação e as condições (2.27) determinam que as regiões só podem ser atendidas por uma instalação ativa. Note-se que estas restrições (2.25) a (2.27) são idênticas às restrições (2.11) a (2.13) do problema das p-medianas.

Em (2.28) se estabelece a máxima distância entre qualquer nó de demanda e sua instalação mais próxima. E as restrições (2.29) e (2.30) garantem que as variáveis sejam binárias.

Até aqui foram apresentados modelos que não diferenciam fontes de produção de pontos de demanda. Já quando se considera esta diferenciação entra-se no campo dos modelos de produção – distribuição.

Dentro desta categoria a presente pesquisa procurou focar os modelos de distribuição que procuram definir as localizações de centros de distribuição de carga. Dentre os modelos de distribuição, o modelo desenvolvido por Geoffrion e Graves (1974) tem sido considerado importante (Daskin, 1995) por tratar os principais aspectos do processo de distribuição e por ter utilizado com sucesso um eficiente método de solução. O método de solução empregado utiliza a técnica de decomposição proposta por Bender e discutida por Daskin (1995). Este modelo de Geoffrion e Graves é apresentado abaixo:

$$Mín \sum_{i,j,k,l} c_{ijkl}.x_{ijkl} + \sum_{k}\left[f_k.z_k + v_k.\sum_{i,l} D_{il}.y_{kl} \right] \tag{2.31}$$

Sujeita a:

$$\sum_{k,l} x_{ijkl} \leq S_{ij} \quad , \forall\ i, j \tag{2.32}$$

$$\sum_{j} x_{ijkl} = D_{il} . y_{kl} \quad , \forall \; i,k,l \tag{2.33}$$

$$\sum_{k} y_{kl} = 1 \quad , \quad \forall \; l \tag{2.34}$$

$$\underline{v}_{k} . z_{k} \leq \sum_{i,l} D_{il} . \; y_{kl} \leq \overline{v}_{k} . z_{k} \quad , \quad \forall \; k \tag{2.35}$$

• restrições lineares de configuração em y e/ou z (2.36)

$$y_{kl} \in \{0,1\} \; , \; \forall \; k,l \tag{2.37}$$

$$z_{k} = \in \{0,1\} \; , \; \forall \; k \tag{2.38}$$

Notação utilizada no modelo:

i – índice para produtos;

j = índice para plantas industriais;

k = índice para CD´s;

l = índice para zonas de demandas;

S_{ij} = capacidade de suprimento do produto i na planta j;

D_{il} = demanda do produto i na zona l;

$\underline{v}_{k}$ e $\overline{v}_{k}$ = mínimo e máximo fluxo anual do CD k;

f_k = parcela fixa do custo anual de propriedade e de operação do CD k;

v_k = custo unitário do fluxo através do CD k;

c_{ijkl} = custo unitário de produção e transporte do produto i com origem na planta j passando pelo CD k para atender à zona l;

x_{ijkl} = valor do fluxo do produto i transportado da planta j através do CD k para a zona l (deve-se ter $x_{ijkl} \geq 0$);

y_{kl} = 0, se o CD k não atende à zona l e será 1, se o CD k atende à zona l;

z_k = variável que será 0 se o CD k não for ativado e será 1 se o Cd k for um CD ativo;

Neste modelo a função objetivo (2.31) corresponde a soma dos custos de produção e transporte, mais os custos de armazém(fixos e variáveis).

As restrições (2.32) correspondem às condições de oferta. Em (2.33) estabelece-se que as demandas devem ser atendidas e que os fluxos devem ser nulos quando um CD não atende a uma zona. As condições (2.34) especificam que cada zona deve ser servida por um único CD. Em (2.35) garante-se que o fluxo através de cada CD estará dentro de limites pré-estabelecidos e que há uma relação lógica entre z_k e y_{kl}, isto é, $z_k = 1$ se algum $y_{kl}=1$. As restrições (2.36) correspondem a outros tipos eventuais de condições que possam ocorrer. A única condição é que sejam lineares. As restrições (2.37) e (2.38) garantem que as variáveis sejam binárias.

Deve-se notar que os modelos apresentados até aqui levam em consideração, no processo de otimização, as distâncias entre as instalações e os pontos de atendimento e/ou custos de transporte. Outros custos logísticos, como por exemplo, o custo de manutenção de estoques, em geral não são considerados nos modelos (Wanke, 2001). Em certos modelos estes custos são considerados através de aproximações lineares. O próprio modelo de Geoffrion e Graves permite esta aproximação ao se considerar que os estoque evoluem linearmente com o número de instalações. Esta evolução pode ser incorporada no coeficiente f_k, que representa uma parcela fixa do custo anual de propriedade e de operação do CD k.

Uma outra tentativa para esta incorporação do custo de estoques na função objetivo é apresentada por Nozick e Turnquist (1998 e 2001) que utilizam um modelo de medianas com restrição de capacidade em que se considera um custo de instalação *(fixed charge facility location problem)*. Os autores desenvolvem uma análise de regressão linear (Draper e Smith, 1981) entre os custos de estoques de segurança e o número de centros de distribuição (CD) e incorporam esta relação no coeficiente f_k,

que corresponde a um custo fixo de instalação de um CD. O modelo utilizado pelos autores é apresentado abaixo:

$$Mín \sum_{k} f_k . X_k + \alpha \sum_{kk} \sum_{j} D_j . d_{kj} . Y_{kj} \quad (2.39)$$

Sujeita a:

$$\sum_{k} Y_{kj} = 1 \quad , \quad \forall j \quad (2.40)$$

$$Y_{kj} \leq X_k \, , \quad \forall k, j \quad (2.41)$$

$$Y_{kj} \geq 0 \, , \quad \forall k, j \quad (2.42)$$

$$X_k = 0,1 \quad ; \quad \forall k \quad (2.43)$$

onde: f_k = custo de instalação de um CD no nó k;

D_j = demanda no nó j;

D_{kj} = distância do nó k ao nó j;

α = custo de transporte por unidade de distância por unidade de demanda;

Y_{kj} = fração de demanda no nó j atendida pelo CD localizado no nó k;

$$X_k = \begin{cases} 1 & \textit{uma instalação foi ativada no nó k} \\ 0 & \textit{em caso contrário} \end{cases}$$

Neste modelo a função objetivo (2.39) se refere aos custos de instalação do CD e dos custos de transporte. Em (2.40) garante-se que os CD's atendam à demanda de cada nó j e que esta demanda não é excedida. As restrições (2.41) asseguram que a demanda no nó j não pode ser atendida por um nó k se um CD não foi alocado

àquele nó. Em (2.42) tem-se as restrições de não negatividade para as frações de demanda dos nós j e em (2.43) as restrições da variável binária X_k.

Deve-se notar que este modelo não tem restrição associada à capacidade do CD, por isto é chamado de modelo sem restrição de capacidade *(uncapacitated fixed charge facility location problem)*. Entende-se que esta na verdade foi, dentre as opções de modelos disponíveis, apenas uma escolha de Nozick e Turnquist (1998) que permitisse o desenvolvimento do estudo, pois o foco da pesquisa era a introdução de um custo de estoque na função objetivo e não o tipo de modelo, propriamente dito.

Por uma questão de completar esta revisão bibliográfica é importante registrar também a outra versão deste tipo de modelo, que é aquele que apresenta restrição de capacidade *(capacitated fixed charge facility location problem).* Nesta versão o modelo tem agora um novo conjunto de restrições em que um fator φ_k define a capacidade do CD k (Daskin, 1995). As demais condições são as mesmas do modelo sem restrição de capacidade.

Este modelo com restrição de capacidade é apresentado a seguir, sendo o novo conjunto de restrições representado por (2.47). Todas as demais condições se mantêm.

$$Mín \sum_{k} f_k . X_k + \alpha \sum_{kk} \sum_{j} D_j . d_{kj} . Y_{kj} \tag{2.44}$$

Sujeita a:

$$\sum_{k} Y_{kj} = 1 \quad , \quad \forall j \tag{2.45}$$

$$Y_{kj} \leq X_k \, , \quad \forall k, j \tag{2.46}$$

$$\sum_j D_j.Y_{kj} \leq \varphi_k.X_k\,, \quad \forall\, k \tag{2.47}$$

$$Y_{kj} \geq 0\,, \quad \forall\, k, j \tag{2.48}$$

$$X_k = 0,1 \;\; ; \;\; \forall\, k \tag{2.49}$$

2.5 DADOS DE ENTRADA DOS MODELOS DE LOCALIZAÇÃO

Nesta última parte da revisão bibliográfica será abordada uma questão que é fundamental para a confiabilidade dos resultados obtidos na aplicação dos modelos de programação inteira, que correspondem à grande maioria das aplicações práticas em problemas de localização. Trata-se dos dados que são empregados nestes modelos. Os modelos serão tratados aqui de forma genérica como sendo os do tipo MILP, já que são mais abrangentes que os modelos de programação inteira pura (vide capítulo 1).

A questão dos dados é uma preocupação que está presente em qualquer tipo de aplicação de modelagem matemática, mas no caso dos modelos MILP há uma particularidade que merece atenção adicional.

Os modelos do tipo MILP partem de um conjunto de nós em uma rede logística nos quais as instalações deverão se localizar. Estas instalações deverão atender a um conjunto de pontos de demanda, que por sua vez também correspondem a nós da rede logística.

Dois aspectos fundamentais associados aos nós desta rede estão envolvidos no processo de preparação dos dados de entrada dos modelos MILP. São eles: a

questão da agregação dos nós de demanda em *clusters* e a definição dos nós que podem ser candidatos a instalações. Estes dois pontos são fundamentais, pois determinam todo o resultado do modelo de otimização. É com base em pontos agregados em *clusters,* sendo cada *cluster* representado por um *centróide*, que se determinam as localizações ótimas das instalações. A base para esta escolha das localizações ótimas é um conjunto de nós candidatos pré-definidos. O estabelecimento de critérios para estes dois tópicos se apresenta assim como fundamental para a garantia de boas soluções.

No que diz respeito à agregação dos nós de demanda, Lacerda (1999) lembra que não é incomum ter que se trabalhar com centenas de locais candidatos a instalações e milhares de clientes e nesta linha os autores em geral concordam que devido a esta e a algumas outras razões principais este processo de agregação realmente se faz necessário.

Francis et al. (1999), por exemplo, apresentam três razões que levam ao uso de pontos agregados. O primeiro aspecto a se considerar é o esforço computacional para resolver um problema de localização. A maior parte dos problemas de localização é do tipo *NP-completo* (Daskin,1995), o que leva a se pensar em encontrar formas para reduzir as dimensões do problema. O autor pondera que muito se tem feito através de métodos heurísticos para se aproximar da melhor solução para o problema, mas os problemas com número muito grande de nós de demanda continuam impraticáveis sob o ponto de vista computacional se forem tratados com os nós desagregados. Outra questão a se considerar diz respeito à representação espacial da demanda. A maioria dos problemas de localização destina-se à definição de instalações que serão utilizadas por um razoável período de tempo. Assim, os autores acreditam que dificilmente as instalações irão atender a

uma demanda exatamente igual àquela que está sendo considerada no momento do estudo. Na verdade, uma representação espacial do *padrão de demanda*, mais do que a demanda em si, com todos os seus detalhes, é que caracteriza e determina os resultados do problema real no futuro. O terceiro ponto a se considerar é quando há necessidade de manter os dados confidenciais, o que ocorre em muitos casos. Com os dados agregados esta confidencialidade está preservada.

Já, Ballou (1994) afirma que de um ponto de vista de rede logística, por se tratar de uma questão de planejamento estratégico, não há necessidade de um detalhamento muito grande das informações a ponto de se tratar cada nó de demanda de forma individual, podendo-se realmente agrupá-los em *clusters*.

Existem diferentes formas de agregação de pontos, como através de códigos de endereçamento postal, distritos urbanos, regiões administrativas. Mas, na verdade, algumas normas gerais devem ser seguidas no processo de agregação.

Norman et al. (1999) apresentam uma metodologia geral para agregação de pontos que deve incluir: definição das entidades a agregar; um método para medir a similaridade entre as entidades; um método para agregar as entidades, definindo *clusters*, baseado na medida de similaridade; um método para combinar os membros do *cluster* de forma a criar um modelo agregado. E quando a agregação de pontos é feita em problemas táticos ou operacionais, em algum momento deverá ser providenciada a desagregação da solução encontrada. Assim, nestes casos um método para desagregar a solução do modelo agregado também deve ser definido.

Dentro de uma visão de coleta e distribuição de carga, Valente et al. (1997) citam quatro aspectos a considerar: compacidade, que é a medida de proximidade de um conjunto de pontos de serviço; morfologia, que representa a forma do grupo de pontos, que em geral é definida pelas características físicas da região e/ou a

finalidade do transporte; balanceamento entre os conjuntos agregados do número de pontos e do serviço demandado em cada ponto; homogeneidade entre as regiões de cada conjunto de pontos.

Em um contexto de planejamento de transportes, Baass (1981) definiu um conjunto de critérios que devem ser seguidos na montagem de *clusters* ou zonas, que é o termo empregado no artigo. São eles: buscar o máximo de homogeneidade dentro de cada zona; reter o máximo de interações entre zonas minimizando o volume de viagens intrazonas; limitar o número de viagens por zona; respeitar limites físicos, políticos e históricos quando isto for relevante para os propósitos da análise; evitar formatos indesejáveis para as zonas; agregar somente entidades espacialmente adjacentes; gerar somente zonas contíguas; evitar a formação de ilhas (zonas contidas em outras zonas); variáveis características de cada zona devem ser homogêneas, tais como: população, número de domicílios, área e número de viagens; o delineamento das zonas deve se basear nas zonas do censo. Note-se que alguns destes critérios dizem respeito apenas ao planejamento de transportes, mas a maior parte é adequada também ao planejamento logístico.

Daganzo (1984) desenvolve um trabalho em que seu objetivo não é propriamente definir *clusters*, mas sim determinar as distâncias percorridas em um roteiro em uma zona. Mas, o fato relevante para a presente pesquisa é que a medida que desenvolve experimentos com zonas, adota um padrão de análise que é a zona retangular. Este é um formato que vem se repetindo em todos os artigos analisados, podendo-se dizer que se constitui, assim, em um padrão básico que os autores utilizam em suas análises. Daganzo depois expande o resultado de uma zona retangular para outros formatos. Outro aspecto interessante deste trabalho de Daganzo é o uso de julgamento pessoal para a análise dessas rotas intrazonas. O

autor utilizou um conjunto de indivíduos para desenvolver suas próprias rotas e as comparou posteriormente com procedimentos estruturados. Nesta pesquisa, algo semelhante foi utilizado no capítulo 4, no desenvolvimento de um experimento.

Muyldermans et al. (2002) desenvolveram um estudo para uma operação de pulverização de sal em áreas cobertas por gelo ou neve a partir de um depósito e verificaram que as características relevantes de distritos bem projetados são: a sua habilidade para suportar boas rotas, homogeneidade do esforço de produção, áreas compactas e centralização de depósitos. Segundo os autores, a forma dos distritos deve ser próxima do formato circular ou retangular, que se mostram melhores do que formas alongadas.

Em outro artigo Ballou (1991b) sugere uma heurística para a montagem de *clusters* que é baseada na proximidade geográfica dos pontos de demanda. A heurística parte da idéia de que em princípio cada ponto se constitui em um *cluster*. Em seguida procura-se para cada ponto o *cluster* mais próximo e destes dois gera-se um novo. Este processo se repete até que um número desejado de *clusters* seja atingido. As coordenadas de um *cluster* corresponderão às coordenadas do centro de gravidade entre seus pontos.

Novaes (1989) também estudou a questão de distribuição urbana e através de uma heurística desenvolve a subdivisão de uma área de estudo criando um conjunto de anéis a partir de quadrículas elementares. Mais recentemente Novaes e Graciolli (1999) propõem um novo procedimento para criação de *clusters* destinados a uma operação de distribuição de carga a ser efetuada em uma certa região de estudo. No modelo proposto a região é particionada em setores a partir de um depósito central. Os autores utilizam aqui o mesmo conceito utilizado por Novaes (1989) de quadrículas elementares. Estes setores por sua vez são subdivididos em anéis e

dentro de cada anel uma nova subdivisão é feita gerando os clusters finais denominados pelos autores de *distritos*. As subdivisões vão sendo feitas de forma a buscar uma quase-homogeneidade do esforço de produção. O artigo procura definir estes distritos para uma operação de distribuição de carga com veículos de mesmas características e mede então este esforço de produção através do número de veículos necessários em cada subdivisão.

No que tange ao segundo aspecto importante a ser considerado,ou seja, os locais candidatos a instalações; o que se verifica é que este é um campo em que a literatura é muito escassa em estudos mais aprofundados. E esta escassez se reflete com uma intensidade muito maior quando se procura por estudos envolvendo métodos quantitativos para a definição de candidatos. As poucas referências na área dizem respeito basicamente a propostas de análises qualitativas do assunto com algumas sugestões de um tratamento que poderia ser chamado de *semiquantitativo*. O que se verifica, na verdade, é que a questão vem sendo tratada fundamentalmente sob o ponto de vista qualitativo.

Randhawa e West (1995), por exemplo, tratam da seleção de candidatos dentro desta linha de tratamento qualitativo para a questão. Os autores relacionam um conjunto de fatores que consideram os atributos que os locais candidatos deveriam possuir. Estes fatores, entretanto, em boa parte são diretrizes gerais, o que levam a caracterizar a análise como sendo qualitativa. De forma a agregar alguma objetividade ao processo, seria necessário a atribuição de pesos aos fatores e a partir daí desenvolver um processo de avaliação de diferentes locais, chegando-se a uma lista de candidatos. Poder-se-ia caracterizar, assim, o procedimento como sendo parcialmente quantitativo.

Ballou (1994) também trata do tema e pondera que para indústrias e centros de distribuição de carga, que são os tipos de instalações em que o autor foca sua atenção, a localização é definida em geral a partir de modelos matemáticos discretos de localização que selecionam os melhores locais a partir de uma lista de candidatos pré-selecionados. O autor coloca, entretanto, que estes candidatos são definidos a partir de critérios totalmente subjetivos, dentro daquilo que ele chama de "razoabilidade" (*reasonableness*). Segundo o autor, deve-se considerar que se for "razoável" localizar, por exemplo, uma indústria em um certo local, este poderá então entrar na lista de candidatos.

Outros autores também abordam o assunto, como Bowersox e Closs (2001) e Coyle et al (1992). Estes trabalhos seguem na mesma linha de Randhawa e West (1995), apresentando atributos que os locais deveriam ter. Neste caso, as sugestões dizem respeito, por exemplo, a zoneamento da região que deve permitir instalações industriais e circulação de veículos de carga. Comentam também a questão da infra-estrutura, que deve apresentar um mínimo de condições para a instalação do empreendimento. Discutem ainda a questão da disponibilidade de mão de obra que consideram outro ponto importante para a instalação de empreendimentos industriais. Ficam assim, basicamente no campo qualitativo.

Verifica-se, portanto, que a literatura realmente se restringe a apresentar as características que os locais devem apresentar para se constituírem em candidatos. Uma vez que se encontrem estes locais, a literatura é carente em propor procedimentos matemáticos que estabeleçam objetivamente aqueles locais que se apresentam como os "melhores" candidatos para localização das instalações.

3. PROPOSTA DE MODELAGEM

3.1 IDÉIAS INTRODUTÓRIAS

Foi visto, na revisão bibliográfica, que existem várias metodologias para tratar o problema em estudo. Ballou (1994) relaciona algumas: os métodos exatos estão na primeira categoria comentada e o autor inclui a abordagem do múltiplo centro de gravidade e também a programação linear inteira mista. O primeiro não é um método que conduza a uma solução ótima, enquanto que o segundo sim, tratando-se de um conhecido método de otimização. O autor fala ainda da simulação e de métodos heurísticos. Mas o interesse de análise desta pesquisa está na programação linear inteira mista, tratada como MILP (*"mixed-integer linear programming"*). Na verdade, esta solução através de MILP tem sido tradicionalmente uma das principais abordagens para solução do problema. E, conforme visto na revisão bibliográfica, um modelo clássico que representa bem este tipo de solução é aquele apresentado no trabalho de Geoffrion e Graves (1974), o que segundo alguns autores, apesar de já ter sido desenvolvido há algum tempo, ainda é um dos melhores modelos deste tipo e tem sido bastante empregado em aplicações práticas (ver Vidal e Goetschalcx,1999 e Geoffrion e Graves,1995).

Este modelo de Geoffrion e Graves incorpora grande parte das características encontradas na maioria dos modelos do tipo MILP utilizados para solução do problema. Pode assim ser utilizado como base para se desenvolver uma análise das características principais de um modelo MILP.

Este modelo trabalha com uma rede logística em que é definido um conjunto de nós que são candidatos a locais de instalação de CD´s e a partir desta definição têm-se então, dois tipos de variáveis de decisão:

- *Variáveis do tipo x_{ijkl} , que representam fluxos do produto i, que parte do local de produção j, passa através do CD k e tem como destino a zona de demanda l*
- *Variáveis binárias, do tipo 0 ou 1, que representam abertura de CD's e Zonas Atendidas por CD's (CD é aberto ou não / CD atende zona de demanda ou não)*

Nesta solução clássica a própria estrutura do modelo obriga a que pelo menos duas condições básicas sejam seguidas:

a) as funções do modelo devem ser todas lineares;

b)as variáveis de decisão fazem parte de uma única função objetivo, que é otimizada.

Estas questões estruturais do modelo e suas características levam a certas simplificações na representação do comportamento do sistema que ensejam uma série de considerações. Inicialmente, deve-se considerar que o modelo opera com uma função objetivo que representa neste caso custos que devem ser minimizados. O processo de localizar e quantificar CD's envolve, por outro lado, um conjunto de complexas compensações e evoluções destes custos. E pelo menos uma parte dos custos envolvidos nestas compensações segue evoluções nitidamente não lineares, como é o caso da variação dos custos de estoque, que sofrem um incremento à medida que o número de CD's aumenta. Neste caso, emprega-se com freqüência a chamada *regra da raiz quadrada* que relaciona a evolução do estoque de segurança

de uma operação com a raiz quadrada da proporção do aumento de pontos de estocagem (Ballou, 1997; Fleury et al., 2000). Como se vê, esta é uma relação que apresenta comportamento tipicamente não linear.

Um modelo do tipo MILP, apesar de procurar tratar estas compensações, o faz com custos linearizados, devido à sua própria estrutura.

O modelo, deste modo, conduz a uma perda de realismo e conseqüentemente, a solução ótima assim obtida, pode eventualmente até estar distante da solução ótima do sistema real.

Sobre a segunda condição imposta pelo tipo de metodologia usada, deve-se considerar, inicialmente, que a questão em estudo é, na verdade, a par de outros aspectos, fundamentalmente, um problema que tem duas vertentes principais: a *localização e a quantificação* de CD´s e estes são pontos que normalmente são endereçados no mesmo problema, pois um interfere diretamente no outro. Para ilustrar esta vinculação, pode-se pensar numa situação em que dois CD´s devem ser posicionados numa certa área. Imaginando-se que a demanda é uniformemente distribuída sobre a área de atendimento e que se pretende minimizar as distâncias percorridas para o atendimento da área, então uma provável solução será um CD em cada extremo desta área. Mas, se apenas um CD será localizado, então seu provável posicionamento será no centro da área de atendimento, e neste caso, as duas posições pensadas para o caso de dois CD´s são abandonadas. Nenhuma delas é aproveitada. Percebe-se, assim, uma total vinculação entre localização e quantidade de CD´s.

Como as duas questões – localização e quantidade – devem ser endereçadas no mesmo problema, as metodologias baseadas em MILP tentam encontrar uma solução ótima para o problema procurando seguir esta mesma linha de raciocínio, o

que significa dizer que tratam as duas questões de forma conjunta no mesmo modelo, como é o caso da proposta de Geoffrion e Graves (1974), em que em uma única função objetivo são tratadas todas as variáveis de decisão.

Um questionamento que poderia ser levantado seria o de se pensar se não é este tratamento conjunto que está levando a certas simplificações no modelo? Este é um ponto a ser analisado.

E um outro questionamento é que, independentemente do fato de as duas questões estarem fortemente inter relacionadas, será que isto obriga a que haja necessariamente um tratamento simultâneo das duas?

O que realmente é necessário é que o vínculo entre as soluções seja mantido e não soluções simultâneas.

Ocorre que estes modelos MILP trabalham com um único conjunto de variáveis de decisão, e o algoritmo empregado leva a que sejam tratadas todas de forma simultânea, não restando, portanto, outra alternativa a não ser procurar localizar e quantificar CD's ao mesmo tempo. O tratamento simultâneo, portanto, é função da metodologia utilizada e não das condições do problema, como se poderia pensar à primeira vista. Logo, se fosse possível o uso de uma metodologia que mantivesse o vínculo entre as duas soluções, o tratamento simultâneo não seria necessário e neste caso, então, simplificações que são devidas à metodologia empregada poderiam ser revistas.

A não ser, portanto, pelo fato de se estar trabalhando unicamente com um modelo do tipo MILP, não há razão para se pensar em manter este tipo de tratamento.

Neste ponto convém fazer uma observação sobre os problemas que empregam os modelos do tipo p-medianas. Estes modelos não tratam localização e quantidade de CD's de forma endógena ao modelo. Na verdade, o parâmetro p é justamente a

quantidade de CD's que vai sendo modificada a cada aplicação do modelo. O tratamento, portanto, não é simultâneo. Ocorre, entretanto, que estes modelos não se aplicam ao problema em questão. Este modelo considera que as p-medianas que corresponderão a p CD's devem ser determinadas para atender unicamente a um conjunto de nós de demanda, sem levar em consideração as origens da carga. Tondo (1992) já demonstrou a necessidade de se aplicar algoritmos adicionais ao p-medianas quando se deseja trabalhar um modelo de distribuição, que no estudo em questão se referia à localização de terminais de contêineres.

Portanto, para os modelos de distribuição as soluções simultâneas, referentes à localização e quantificação de CD's, permanecem.

Além destes dois pontos discutidos, existem outras questões que também merecem ser endereçadas.

Os modelos baseados exclusivamente em MILP sofrem de uma perda de flexibilidade, pois a abordagem que normalmente se adota ao se construir o modelo é fazer com que alguns dos coeficientes da função objetivo agreguem mais de um componente de custo logístico. Custos de estoques, armazenagem e transporte, por exemplo, quando são considerados, são geralmente incorporados num único coeficiente (Geoffrion e Graves, 1974; Ballou, 1994; Chopra e Meindl, 2001). Assim, um único parâmetro comanda a evolução de todo um conjunto de custos, que na verdade, evoluem a partir de uma série de variáveis com diferentes tipos de comportamento. Todos os custos são agregados, o que faz com que todos passem a ter comportamento linear e com a mesma taxa de evolução (reta com mesma inclinação). Esta simplificação, além de representar uma perda sensível de realismo, representa ainda uma perda de flexibilidade, pois se fica, desta forma, muito restrito no manuseio deste conjunto de variáveis.

Os custos fixos de implantação normalmente são tratados em outros coeficientes.

Além da perda de flexibilidade, deve-se considerar que fica praticamente impossível considerar todas as características reais num único modelo operando com estas características e com coeficientes que têm comportamento linear.

No que tange a esta não linearidade do modelo, poder-se-ia, inclusive, pensar no desenvolvimento de uma solução através de programação não linear, porém a complexidade matemática seria maior e a idéia desta pesquisa, no que tange a programação matemática, é procurar se manter dentro da programação linear, devido à sua ampla utilização em aplicações práticas e à simplicidade de sua estrutura, que se pretende aproveitar. O objetivo desta pesquisa é manter esta estrutura, porém procurar agregar um maior realismo, de forma a aperfeiçoá-la. E é nesta linha que se procurará desenvolver uma proposta de modelagem.

A idéia, na verdade, é utilizar a favor da pesquisa uma propriedade do problema para procurar desenvolver um aprimoramento desta modelagem.

O que ocorre é que ao se refletir somente sobre a localização de centros de distribuição, desconsiderando-se por um momento a questão da definição da quantidade de CD's, percebe-se que a localização, de um dado número de CD's, depende basicamente de forma direta ou indireta do custo de transporte e do custo do local (custo de uso do solo) onde será instalado o CD (Ballou, 1994; Daskin, 1995; Bowersox e Closs, 2001).

Os demais fatores como estoque e porte do armazém, por exemplo, não devem sofrer influências do posicionamento das instalações.

O ponto a ser considerado, portanto, é que se for fixado o número de CD's, a localização destes CD´s irá depender fundamentalmente do custo de transporte e

do custo do local (implantação e uso do solo). Os demais custos não apresentam relação com a questão.

Isto leva naturalmente a se pensar então que *se não há envolvimento de outros custos na localização de CD´s, por que deveriam estes estar envolvidos na análise?*

Isto na verdade ocorre, porque as duas questões – localização e quantidade - são tratadas de forma conjunta num mesmo modelo do tipo MILP.

Note-se que ao se considerar que existem certos custos que têm envolvimento direto com a quantidade de CD's, mas que não sofrem alteração à medida que a localização dos CD's é modificada, e como localização e quantidade recebem tratamento simultâneo, todos os custos que estão envolvidos com uma ou com outra questão devem necessariamente ser considerados.

Porém, se fosse desenvolvida uma outra abordagem para o problema tratando-o de forma segmentada, seria então possível considerar em cada parte da análise somente os componentes de custos que fossem relevantes às questões abordadas naquele segmento.

Um dos componentes que não tem relação com a posição dos CD's é o estoque. A evolução dos custos de estoques com a quantidade de CD's é um *trade-off* clássico em logística (Ballou,1994, Bowersox e Closs,2001; Fleury et al., 2000). Os custos de estoque aumentam à medida que a *quantidade* de CD's aumenta, independentemente de onde estão posicionados estes CD's. Segundo a regra da raiz quadrada (Ballou,1994; Fleury et al., 2000) os estoques de segurança aumentam na mesma razão da raiz quadrada da proporção do aumento da quantidade de CD's. Há assim, independência do posicionamento dos CD's.

A segmentação do problema não implica em perda de precisão ou em empobrecimento da modelagem, pelo contrário. Com esta linha metodológica, é

possível atender a dois objetivos, que à primeira vista poderiam ser entendidos como contraditórios: o primeiro seria uma tentativa de promover uma facilidade maior para a modelagem e o segundo seria o de ao mesmo tempo torná-la ainda mais realista.

A maior facilidade é obtida pela própria subdivisão do problema, o que naturalmente facilita qualquer análise, devido ao menor número de fatores envolvidos em cada parte. Cada um dos segmentos corresponderá a um submodelo, que em princípio, não deverá ser tão complexo de ser tratado do que um único modelo que esteja abrangendo todos os aspectos da análise. Ao mesmo tempo será possível dar mais realismo aos modelos, pois com a segmentação, cada submodelo poderá passar por um trabalho de maior aprimoramento da modelagem do sistema de forma a torná-la mais próxima das condições reais em estudo. Consegue-se, assim, uma melhoria na qualidade dos modelos.

Na verdade, autores que comentaram estas questões sugerem mesmo a subdivisão do problema. Vidal e Goetschalcx (1997), por exemplo, tratando de um contexto de modelos globais, em que se pode pensar como uma generalização do presente estudo, afirmam textualmente: "*é praticamente impossível considerar todas as características reais num único modelo*" e Geoffrion e Powers (1995), apesar de salientarem toda a evolução dos modelos nas últimas décadas, falam da necessidade de se pensar em *integração de modelos*, além de citar a questão dos *custos não lineares*, onde mostram sua preocupação com as aproximações lineares que têm sido empregadas nos modelos em uso.

Dadas estas considerações, é o caso, então, de se refletir sobre as razões da insistência em se continuar a trabalhar com um único modelo totalmente baseado em programação linear mista para tentar representar um processo que sabidamente

não tem comportamento linear e que, segundo pesquisadores, dada a sua complexidade, deveria ser tratado de forma segmentada.

O modelo baseado em programação linear mista, realmente, tem a vantagem de ser uma abordagem de maior domínio dos técnicos e pesquisadores e de ser ainda uma técnica de otimização. Mas, não há garantias de que esta solução ótima do modelo seja a solução ótima do sistema real, dadas as aproximações que são feitas.

Por que, então, não se propor um aprimoramento desta abordagem, que mantivesse suas características, mas que por outro lado, fosse mais realista?

O fato novo a se levantar então seria efetivamente o de se propor um tratamento separado de localização e quantidade de CD's, procedendo-se assim a uma subdivisão do problema.

A condição para que isto seja feito é que seja mantido o vínculo entre as soluções de cada segmento a ser considerado. Se este vínculo realmente se mantiver, o problema então pode ser subdividido.

No sentido de se manter na linha de um aprimoramento da metodologia, a idéia da segmentação pode ser implementada desenvolvendo-se uma modelagem em que o MILP continue a desempenhar um papel fundamental. Porém, é importante, no sentido deste aperfeiçoamento do modelo que sua função objetivo considere somente os custos que sem perda de realismo possam ser considerados lineares.

O MILP seria mais bem aplicado, assim, à primeira parte do problema, que trataria exclusivamente da localização de CD's e onde os custos, apesar de não serem totalmente lineares, poderiam ser aproximados sem perda de realismo.

Os demais custos seriam tratados de forma externa a este modelo baseado em MILP e seriam incorporados posteriormente, na segunda fase que trataria da quantidade de CD's. Estes outros componentes passam a ter modelos específicos,

que consideram suas particularidades de comportamento, inclusive uma eventual não linearidade. Como o custo logístico é uma composição de custos, é possível se desenvolver um submodelo adequado a cada componente. Cada submodelo poderá ter assim uma dose maior de realismo sem a preocupação de se evitar eventuais complexidades matemáticas que não seriam suportadas por um modelo MILP. Como nesta segunda parte, cada componente de custo varia justamente com a quantidade de CD's, cada submodelo a ser montado deverá então ser uma função que expresse o valor do custo para um dado número de CD's. Se a relação entre custo e quantidade de CD's for não linear, esta função deverá, de forma realista, procurar demonstrar tal relação.

3.2 SEGMENTAÇÃO DA MODELAGEM

A proposta básica de metodologia consiste, portanto, em subdividir o problema em duas fases:

Fase 1. Localização dos CD's;

Fase 2. Quantificação de CD's.

Uma vez que o problema é subdividido, a modelagem fica composta de dois modelos principais:

. modelo de localização de CD's,

tendo por base um modelo MILP e

. modelo de quantificação de CD's,

constituído de submodelos de custos.

Em termos de otimização também haverá uma segmentação. A proposta é de uma primeira otimização referente apenas a uma parcela dos custos logísticos no modelo MILP de localização e uma segunda otimização global considerando um conjunto mais abrangente de custos logísticos no modelo de quantificação.

O modelo MILP irá contemplar apenas os custos que podem ser considerados lineares sem perda de realismo. Já a modelagem para a *Fase 2,* e lembrando as considerações feitas na seção anterior, irá trabalhar com submodelos por componente para aproveitar a característica do custo logístico de ser uma composição de custos. O modelo de custo global será a soma dos modelos individuais.

Em termos de representação de restrições que em geral existem na operação, estas serão consideradas no modelo MILP e também na preparação dos dados para o MILP (seleção de locais candidatos a CD's).

O modelo MILP na *Fase 1* deverá ser aplicado para diferentes valores de p (número de CD's). Para cada valor desse parâmetro p o modelo irá encontrar a localização ótima dos CD´s. Em seguida, na *Fase 2*, será definido o número p_0 de CD's que minimiza o conjunto de custos logísticos globais. Definido o valor "ótimo" p_0, usa-se então a solução ótima de localizações correspondente a p_0.

Neste sentido, a modelagem irá funcionar de forma análoga ao que ocorre quando se deseja construir uma representação gráfica dos vários componentes de custos logísticos evoluindo com o número de CD´s. O que acontece em tal representação é que a curva de custo logístico global corresponde à soma das curvas individuais.

Neste caso, cada ponto da curva do custo de transporte (*Fase 1*) irá corresponder às localizações ótimas para um dado número p de CD´s. As curvas de outros custos que não dependem de localização (*Fase 2*) irão sendo "otimizadas" em relação à

quantidade de CD´s. A soma das curvas, que corresponde ao custo global, irá gerar as localizações ótimas e a quantidade “ótima”.

O problema assim será resolvido através de uma heurística constituída de um processo iterativo conforme definido abaixo:

Passo 1: Inicializar parâmetro. Fazer p = 1;

Passo 2: Modelo de localização

Localizar p CD´s utilizando um modelo MILP;

Passo 3: Modelo de quantificação

Empregar submodelos de custos para cada componente;

Passo 4: Montagem do custo global

Com os resultados do modelo MILP e dos submodelos de custos construir o custo logístico global para p CD´s.

Passo 5: Se p=1, incrementar p de uma unidade e retornar ao passo 2. Em caso contrário, verificar se há aumento ou redução no custo global.

- Se custo aumentar, encerrar o processo.
  Solução foi encontrada. A quantidade de CD´s será igual a p-1.
- Em caso contrário, incrementar p de uma unidade e retornar ao passo 2.

Uma observação final sobre o uso desta heurística diz respeito à decisão sobre os componentes de custo que serão tratados em cada fase. Um dos componentes que merece uma atenção especial é aquele associado ao porte das instalações de armazenagem. A dimensão destas instalações evolui na mesma medida que o estoque. A taxa de variação do nível de estoque direciona a taxa de variação da necessidade de volume útil de armazenagem. A dimensão das instalações é um fator, portanto, que independe da localização dos CD’s. É função exclusiva da

quantidade de CD's. Para um dado número p de CD's a dimensão das instalações está definida e poderia em princípio ser tratada apenas na *Fase 2* da modelagem. O que ocorre, entretanto, é que como o custo unitário do solo pode variar de um local para outro, o custo da área necessária às instalações para um dado número p de CD's poderá também variar entre os locais, apesar de suas dimensões permanecerem estáveis. Este é um fato relevante, pois interfere diretamente no modelo MILP que é tratado na *Fase 1* e que deverá minimizar uma função de custo incluindo aqueles componentes de custos que interferem nas localizações ótimas dos CD's.

Esta questão, portanto, deve ser tratada de forma adequada. O seu tratamento, entretanto, pode apresentar duas variantes:

- a primeira é quando se considera um custo do solo uniforme entre os diferentes nós da rede logística. Neste caso, o custo da área necessária às instalações passa a ser independente do local e pode, portanto, ser integralmente tratado na *Fase 2* da modelagem. Os passos da heurística acima podem ser seguidos exatamente conforme a sua descrição.

- a outra possibilidade é quando este custo não for uniforme. Nesta situação o volume útil necessário de armazenagem para cada quantidade p de CD's deve ser um parâmetro de entrada para a definição deste custo no modelo MILP. Isto significa que se deve trabalhar antes com o submodelo de estoques da *Fase 2* que irá gerar uma necessidade de área de armazém. Com esta necessidade retorna-se à *Fase 1* e se determinam os custos correspondentes em cada local. Estes custos serão dados de entrada para o modelo MILP que é aplicado então para o número p de CD's.

4. MODELAGEM DE LOCALIZAÇÃO DE CD's

4.1. O USO DE CANDIDATOS

A resolução dos problemas de localização através de modelos de programação linear inteira mista requer a determinação prévia de um conjunto de soluções viáveis para o problema. Estas soluções correspondem, na verdade, a um conjunto de locais que serão as possíveis localizações das novas instalações. No caso de um estudo de localização de CD's, este conjunto de locais passa a ser, assim, um conjunto de "candidatos" a CD's. O modelo estabelece uma solução ótima, indicando as melhores localizações dentre aqueles candidatos pré-selecionados. A solução ótima encontrada estará, portanto, totalmente vinculada a estes candidatos, que serão utilizados como dados de entrada do modelo de otimização.

Assim, utilizando o conceito exposto por Francis et al (1999), ao se considerar o problema de interesse como sendo o problema da localização de 1ª ordem, então a seleção de candidatos pode ser vista como o problema da localização de 2ª ordem. Notando-se que são dois problemas diferentes e que devem receber tratamento diferenciado, mas que possuem uma relação de dependência, pois a solução do problema de 1ª ordem é ótima para um dado conjunto de candidatos. Se a seleção de candidatos se altera, a solução ótima também pode se alterar.

Esta é uma situação indesejável do ponto de vista prático, pois ao se resolver um problema, o esperado é que alguma garantia haja de que a solução encontrada é efetivamente a melhor. Não haveria interesse em uma solução que é ótima *apenas* para um conjunto de candidatos. Busca-se a melhor solução dentre todos os

conjuntos de candidatos. Em uma situação real, possivelmente se estaria sujeito até à sensibilidade de um indivíduo.

A definição de candidatos, portanto, é passo fundamental para o sucesso da metodologia.

Essa seleção deveria passar por um procedimento estruturado que garantisse que com o conjunto de candidatos selecionado o modelo iria gerar a efetiva solução ótima do problema. Uma alternativa simplista para esta questão seria definir todos os nós da rede em estudo como candidatos a CD's. Apesar de ser este o tratamento ideal do ponto de vista de garantir a otimização, pois assim o modelo poderia selecionar os CD's dentre todas as alternativas possíveis, isto, entretanto, poderia levar um modelo a assumir proporções imensas, inviabilizando a sua aplicação mesmo com os métodos mais modernos de solução de algoritmos de programação linear inteira mista.

O processo de fatorização (Brown e Olson, 1994), por exemplo, permite resolução de modelos de grandes proporções. Geoffrion e Powers (1995) comentam que com este processo é possível a resolução de problemas com até 7.000.000 de variáveis e 550.000 restrições (os autores não especificam os tipos de restrições). Aparentemente, com este poder computacional poder-se-ia pensar em usar todos os pontos como candidatos e o trabalho de definir candidatos não seria mais necessário. Mas, na verdade, esta etapa da otimização continua sendo importante, pois um modelo mesmo em uma situação comum pode apresentar grandes dimensões. Em problemas urbanos, por exemplo, é possível se ter milhões de pontos de entrega em uma grande cidade (Ballou, 1991b). Mas mesmo com um número de pontos muito menor, as dimensões do modelo podem ser extremamente elevadas.

Uma ilustração deste tipo de situação pode ser encontrada em Novaes (1999), que apresenta uma aplicação para a cidade de São Paulo em que se deve atender a 6.000 pontos de entrega. Para este caso foi feita uma análise da dimensão que teria um modelo como o utilizado nesta pesquisa (ver seção 4.4). Se fossem considerados todos os 6.000 nós da rede como candidatos a CD's, o modelo poderia chegar a 36.006.000 de variáveis binárias, mais 6.000 variáveis reais e 72.024.002 restrições (vide Apêndice A).

Mesmo que seja eliminado um número considerável de candidatos, ainda assim as dimensões seriam elevadas. Supondo-se que dos 6.000 nós apenas 1.000 nós pudessem ser candidatos, o modelo chegaria a 1.000 variáveis reais, 6.001.000 variáveis binárias e 12.009.002 restrições. Para a resolução por fatorização dever-se-ia ter apenas 90 candidatos, eliminando-se, portanto, 5.910 pontos. Assim, o modelo teria 546.182 restrições, além de 540.090 condições para as variáveis binárias. Estes 90 pontos representariam apenas 1,5% do total de pontos. Há, portanto, uma tarefa importante de seleção a ser feita, uma vez que 98,5% dos pontos precisariam ser descartados como locais potenciais à instalação de um CD. Todos estes cálculos são apresentados no apêndice A. Vê-se, assim, que a escolha de candidatos é realmente fundamental em um estudo de localização.

4.2. SELEÇÃO QUALITATIVA DE CANDIDATOS

Uma vez que a solução ótima depende diretamente dos candidatos selecionados, sua seleção deveria se dar através de procedimentos estruturados que garantissem a efetiva solução ótima do problema.

A literatura, porém é muito escassa em procedimentos quantitativos para seleção de candidatos. O tratamento que vem sendo dado ao assunto tem um viés acentuadamente qualitativo. Os textos que são encontrados apresentam em geral determinadas condições que os locais em potencial devem ter, sugerindo, por exemplo, que o zoneamento da região deve permitir instalações industriais e circulação de veículos de carga; que os locais devem ter infra-estrutura adequada; que deve haver disponibilidade de mão de obra; etc (Bowersox e Closs,2001; Coyle et al,1992; Ballou,1994). Porém, uma vez que se tenha um conjunto de locais com estas características, que são tratadas de forma qualitativa, a literatura é carente em sugerir procedimentos matemáticos que definam as melhores localizações destes candidatos.

De acordo com Ballou (1994), no caso de indústrias e CD's a localização é geralmente dominada por fatores econômicos em que modelos discretos de localização selecionam as melhores localizações a partir de uma lista de candidatos pré-selecionados pela sua "razoabilidade" ("*reasonableness*"). O autor também trata um assunto associado, que é a questão da agregação dos pontos de demanda e afirma que de um ponto de vista de rede logística, não é necessário tratar cada ponto individual, podendo-se agrupá-los em *"clusters"*. Em outro artigo (Ballou, 1991) o autor já sugeria um procedimento para agregação dos pontos de demanda, porém não tratou neste artigo da questão da seleção de candidatos.

Randhawa e West (1995) também tratam qualitativamente o assunto, propondo uma lista dos aspectos que consideram como sendo os principais atributos que os locais devem possuir, conforme apresentado na tabela 4.1.

Tabela 4.1 - Seleção de Locais - Atributos

- acesso a mercados;
- acesso a matérias primas e suprimentos;
- disponibilidade de mão de obra, profissionais e executivos;
- disponibilidade de infraestrutura de transporte;
- disponibilidade, qualidade e preço de utilidades e serviços;
- considerar aspectos de legislação e incentivos fiscais;
- considerações ambientais e ecológicas;
- custo, dimensão, zoneamento e topografia das áreas disponíveis;
- aspectos comunitários, incluindo moradia, escolas e custo de vida.

Fonte: Randhawa e West (1995)

Deve-se perceber que a maior parte dos pontos colocados acima corresponde na verdade a diretrizes gerais com um forte caráter qualitativo, que eventualmente podem levar a algum tipo de procedimento que poderia ser considerado parcialmente quantitativo. Isto poderia ser feito, por exemplo, com a atribuição de pesos a cada atributo.

Mas, note-se que mesmo atribuindo-se pesos aos atributos, ainda assim não se teria um procedimento propriamente quantitativo, pois a sua aplicação exigiria que um analista passasse antes por todos os nós atribuindo um grau a cada atributo do nó e em seguida haveria apenas o cômputo da somatória de pesos para cada nó. Há que se considerar que a atribuição de pesos pode ser uma tarefa bastante subjetiva, assim como, a atribuição de graus a cada nó.

Além disso, seria muito difícil uma automatização do processo, que é o que normalmente se visa com um procedimento quantitativo. Segundo Lachtermacher (2002), um dos objetivos da *Management Sciences* é a criação de sistemas computacionais úteis para os usuários não técnicos.

Um procedimento como este se constitui, na verdade, mais em uma forma de classificação dos locais do que de seleção propriamente dita; poderia ser considerado como um processo de triagem. Uma seleção automática conduziria

diretamente aos “melhores” locais e abriria ainda a possibilidade de uma automatização do processo.

Deve-se considerar, assim, que não se garante a solução da questão inicial que foi colocada, referente à quantidade de candidatos, pois em uma situação hipotética inicial, com um número muito elevado de candidatos, após a triagem este número poderia continuar ainda a ser muito elevado.

Observa-se, desta forma, através da literatura, que as análises qualitativas dependentes de julgamento pessoal correspondem, em princípio, ao que se tem de estado da arte em termos de seleção de candidatos. O que se busca agora, com esta pesquisa, é que esta seleção se dê através de um procedimento quantitativo que conduza aos “melhores” candidatos.

4.3. PROCEDIMENTO QUANTITATIVO DE SELEÇÃO DE CANDIDATOS

O que se busca, neste estudo é um procedimento que percorra um “*caminho matemático*” bem definido.

Deve-se desenvolver um método de passagem por todos os nós da rede logística em análise seguindo um critério pré-estabelecido, de forma que se vá selecionando automaticamente aqueles nós que se mostrem mais “adequados” como candidatos a CD’s.

Os melhores candidatos não poderão ser identificados sem que haja um critério estabelecido de seleção e sem um procedimento bem definido que indique a forma de busca. Este processo criterioso de busca entre todos os nós da rede exigiria, então, o desenvolvimento de um procedimento de varredura da área de estudo.

Deve-se notar que se a quantidade de nós não for muito elevada a indicação de candidatos pode representar uma tarefa relativamente simples.

Tendo-se poucos nós, bastaria um certo bom senso e eventualmente alguns cálculos triviais e a questão estaria resolvida.Mas quando se deseja estabelecer a melhor combinação dentre uma quantidade grande de nós, a complexidade da tarefa pode crescer muito, chegando a se tornar um problema não trivial.

Com muitos nós e tendo-se em muitas situações proximidades muito grandes entre os pontos, como estabelecer aqueles que devem integrar o grupo de candidatos? Observa-se que a seleção de nós tem características de um problema de análise combinatória (*n* nós e *p* candidatos), crescendo de forma expressiva a sua complexidade à medida que aumenta o número de nós. Esta, portanto, é uma dificuldade a ser superada.

Uma forma para ultrapassar esta dificuldade, facilitando o trabalho de seleção, seria subdividir a área de estudo. O que significaria, na verdade, reorganizar os nós da rede em *clusters*. O trabalho com um número reduzido de nós é naturalmente mais fácil do que o trabalho com toda a rede de uma só vez. Com a subdivisão, o número de combinações de nós a analisar pode se reduzir muito. Um caminho viável, portanto, é a busca da montagem de *clusters* de nós. E tendo-se estes *clusters* montados, a tarefa passa a ser, então, identificar os “melhores” candidatos no interior dos *clusters*.

A forma escolhida para superar estas dificuldades, portanto, é subdividir o processo em dois estágios principais:

Estágio 1: Subdivisão dos nós da rede em *clusters* e

Estágio 2: Seleção dos “melhores” candidatos de cada *cluster*.

4.3 1. Heurística de "*Clusterização*"

4.3.1.1. Homogeneidade das zonas

Com os nós organizados em *clusters,* o que se está definindo, na verdade, é um zoneamento para a área de estudo, pois cada *cluster* passa a se constituir em uma zona. Cada conjunto de nós passa a ser representado de forma agregada pelo centróide da zona, que corresponde ao seu centro de gravidade.

Uma questão chave para a definição deste zoneamento é a consideração do fato de que se deve procurar manter as zonas dentro de um certo padrão de homogeneidade em relação a algum critério. Sem isto o processo de seleção de candidatos tende a ficar desigual. A seleção poderia ser muito simples em uma zona pequena e bastante complexa em uma zona muito grande.

Convém lembrar também, que do ponto de vista de seleção dos candidatos, em cada zona a situação ideal seria ter todos os nós como candidatos. Assim, uma vez que pode não ser viável ter todos os nós, deve-se procurar selecionar candidatos nas zonas que sejam "representativos" dos demais nós daquela zona.

Representar bem tem o sentido de minimizar divergências. Significa visualizar a zona como uma entidade. Assim, Valente et al (1997) sugerem que entre outros atributos as zonas devem ser compactas e homogêneas. A compacidade medida pela proximidade entre os pontos da zona e a homogeneidade deve ser definida através de volume, área ou algum outro fator. Com isto ao se escolher um nó qualquer, espera-se que os demais estejam bem representados. Mas, deve-se considerar ainda que os nós têm demandas diferentes. Portanto, o que se espera é que as demandas dos nós não selecionados como candidatos estejam bem

representadas pelos candidatos. A diferença de representação entre cada unidade de demanda de um nó e de outro deve ser pequena em relação ao centróide.

Convém salientar aqui as características desses dois atributos. Compacidade significará que o centróide estará “adequadamente próximo” de todas as unidades de demanda da zona. Se todos os nós tivessem o mesmo peso, poder-se-ia pensar em zonas muito grandes que tenderiam a ter os nós dos extremos muito longe do centróide e zonas pequenas, que teriam uma tendência contrária e, neste caso, os nós tenderiam a estar bem representados. No limite chegar-se-ia aos nós individuais.

Já, a homogeneidade significa que a magnitude dos desvios de cada centróide, em relação às demandas dispersas em sua zona, deve ser aproximadamente a mesma em todas as zonas. Tendo-se zonas heterogêneas os nós de cada zona são representados de forma desigual pelos seus respectivos centróides.

Esta homogeneidade de um zoneamento pode ser percebida pela própria sensação visual de semelhança entre as zonas, que é transmitida a um observador. Esta semelhança é sentida em relação a algum critério: forma, quantidade de pontos, área, etc. Para se garantir a homogeneidade é fundamental, portanto, o uso de algum critério, e que nesta pesquisa deverá ser quantitativo.

Deve-se considerar também que este critério de homogeneidade não deve perder de vista a finalidade das zonas, que é a seleção de candidatos a CD’s. Estes candidatos a CD’s serão utilizados como dados de entrada em um modelo de otimização que tem por objetivo a minimização de custos logísticos. De acordo com Francis et al (1999), *“a correta estratégia para agregação espacial deve ser direcionada pela função objetivo do modelo em questão”*. A homogeneidade das zonas deve assim estar relacionada aos fatores que geram tais custos.

Para se estabelecer um critério que atenda a estes objetivos deve-se inicialmente considerar a distância entre os nós, pois um dos objetivos na montagem das zonas deve ser reduzir a dispersão entre os seus nós. Os nós de uma zona, portanto, devem estar próximos entre si já que a distância é um fator de desagregação que eleva custos. Por outro lado, a distância não é o único fator de custo, cada nó tem uma demanda associada que também interfere diretamente nos custos. Como se deseja homogeneidade em relação aos fatores geradores de custos, deve-se então considerar também a demanda do nó, além do fator distância.

Um indicador conhecido e amplamente aceito que considera os dois fatores é o momento de transporte, em geral computado em termos de *ton.km*. Podendo, entretanto, utilizar outras unidades de medida, como, por exemplo, $m^3.km$.

A homogeneidade pode assim ser garantida em termos do momento de transporte, já que este indicador considera de forma conjunta os dois fatores. Este índice pode ser computado ao longo de toda a rede de estudo. E a soma dos momentos de transportes, assim calculados, deverá ser aproximadamente a mesma em cada zona. Com isto, garante-se uma quase homogeneidade das zonas em relação aos fatores geradores de custos.

O processo de cálculo do momento de transporte deve ser único para todos os nós da rede e deve ser estabelecido de forma a que sejam consideradas as posições relativas dos nós e suas demandas. O momento de transporte pode ser calculado em termos de distâncias de cada nó para o seu vizinho mais próximo ou então a partir de todos os nós para um certo referencial fixo. Na primeira situação têm-se distâncias entre os nós e, na segunda, diferenças de distâncias de cada nó para um mesmo referencial. Ambos, portanto, são processos aceitáveis de cálculo do indicador.

4.3.1.2. Montagem das zonas

Os modelos de otimização mais abrangentes (Geoffrion e Graves, 1974) levam em consideração os custos de transporte das regiões de produção aos CD's e destes às regiões de demanda. Assim, a montagem dos *clusters* deve, de alguma forma, também considerar estes dois tipos de fluxos.

O custo dos CD's aos nós de demanda é considerado quando se utiliza o momento de transporte entre os nós da zona, pois serão aproximadamente estas as distâncias que estarão sendo percorridas a partir de um CD que atenda aqueles nós.

Quanto às regiões de origem, deve-se considerar que o modelo de otimização irá gerar uma solução ótima em que os CD's estarão posicionados entre as posições de origem da carga e os nós de demanda, pois este é o tipo de posicionamento que gera o mínimo custo. Assim, este também deve ser o posicionamento dos candidatos a CD's.

Agora, se os candidatos devem ser selecionados no interior das zonas, estas devem ser formatadas e posicionadas no sentido de permitir que tal processo de seleção ocorra. Quando se tem um único ponto de origem da carga, um processo que poderia ser adequado seria o de se procurar estabelecer vetores partindo da região de origem na direção dos nós de demanda e as zonas iriam sendo formadas ao longo desses vetores. Assim, os candidatos escolhidos em cada zona ficarão naturalmente posicionados entre a origem dos fluxos de carga e os nós de demanda. Mas em uma operação com várias origens de carga tal processo fica inviabilizado. Seria necessário estabelecer vários vetores que estariam eventualmente em sentidos opostos. Poderia haver encontros e intersecções de vetores. A montagem das zonas poderia ficar muito complexa.

Para considerar estes fluxos dos nós de produção aos CD's foi desenvolvida a idéia de se determinar um ponto central imaginário na área de estudo que correspondesse à posição ótima de um único CD Central. Este CD Central incorpora todas as tendências das origens da carga e de suas relações com os destinos. Passa a ser o ponto de partida para a formação das zonas, e pode ser considerado como o centro de gravidade (CG) entre os nós de origem e os nós de demanda. Seria como se toda a carga fluísse para o CG e daí para os nós de demanda. O CD Central representa, assim, a dispersão das origens da carga em relação à dispersão dos nós de demanda. O CD Central direciona o ponto de partida das zonas para as tendências predominantes das origens da carga e formata as zonas de acordo com estas tendências. Se as origens estão mais concentradas na região sul, por exemplo, o CG estará deslocado para o sul e todas as zonas serão formatadas a partir deste ponto. O centro do conjunto de zonas fica assim deslocado para a direção sul.

Com este procedimento passa a não importar mais se há uma única origem da carga ou múltiplas origens. As tendências de todas as origens ficam devidamente representadas e consolidadas em um único ponto e a partir daí abre-se como em um leque a formação de zonas num processo de avanço para os nós e de varredura de toda a área de estudo.

Com este ponto central, portanto, garante-se que as origens de carga são adequadamente levadas em consideração no processo de formação das zonas.

Outro aspecto a considerar é o formato das zonas. As zonas devem ter um formato que facilite a escolha de candidatos. Um CD atende a sua região de entorno o que resulta, em geral, em uma forma aproximadamente circular.

Assim, um candidato deveria ter nós no seu entorno que estariam na sua região de atendimento.

Pensando-se assim, em estabelecer um candidato por zona, as zonas não deveriam ter formas muito alongadas, devendo de preferência ser o mais próximo possível do formato de um círculo.

Se a área de estudo é muito extensa, a formação de zonas a partir de um ponto central irá redundar em formatos alongados, ou seja, zonas com muito comprimento e pouca largura. Seriam zonas formatadas no sentido longitudinal.

Novaes (1989) estudou a questão de distribuição urbana e desenvolveu um método heurístico para a subdivisão de uma área de estudo em anéis. Em outro trabalho (Novaes, 1999) o autor propõe a criação de distritos, estabelecendo dentro destes anéis uma subdivisão em zonas. Adotando-se este método consegue-se evitar estas zonas alongadas. No presente caso pode-se, a partir do CD Central, definir anéis em função da distância de cada ponto a este CD. As zonas seriam definidas dentro dos anéis. Todas as zonas, independentemente do anel em que se situam, deveriam seguir o critério de homogeneidade.

A figura 4.1 ilustra o tipo de zoneamento que seria obtido a partir dos critérios propostos. Este padrão de zoneamento está de acordo com o que é previsto pelas teorias clássicas de localização, como a proposta por August Lösch (*in* Isard,1956), em que a formatação de áreas de influência de centros de produção se distribuiriam em um formato aproximadamente concêntrico em torno de um centro principal. Segundo Lösch, a área de cada centro de produção formaria um hexágono dando uma forma geral de colméia para a segmentação da área de estudo. Na verdade, Isard (1956) criticou as hipóteses de Lösch, que assumia uniformidade em relação às distribuições de transporte, de preferências da população e de oferta de matéria

prima na área de estudo. Em seu trabalho Isard propôs modificações, mas a formatação final permanecia em moldes semelhantes aos da figura 4.1.

Como aspecto adicional na definição do zoneamento, deve-se pensar no ponto de corte em cada zona.

A base para este corte é a homogeneidade em termos de momento de transporte, mas outros fatores também precisam ser considerados.

A distância ao próximo nó é um desses fatores. Não seria razoável manter um ponto em uma determinada zona sendo que o seu vizinho mais próximo se encontra em outra zona. Este seria um fator que iria gerar dispersão na zona. Este, portanto, é um aspecto a considerar.

Uma outra consideração a ser feita é que é possível que alguns nós isolados não sejam incorporados a nenhuma zona e acabem sobrando no final do processo. Neste caso, estes pontos podem ser agregados à zona do nó mais próximo.

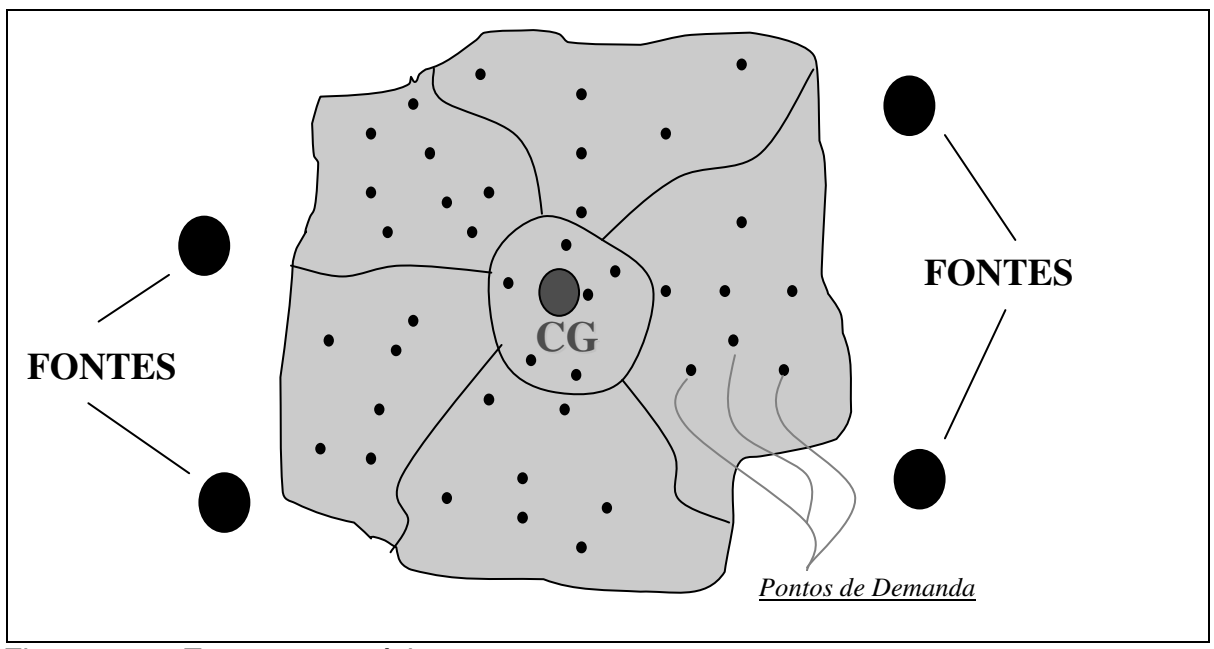

Figura 4.1 – Zoneamento típico

4.3.1.3. Nível de serviço e número de zonas

Além dos aspectos discutidos na seção anterior, a metodologia aqui proposta irá considerar também que o processo de *clusterização* deve ser desenvolvido de forma que um nível de serviço (NS) previamente estabelecido seja atendido e que eventuais restrições operacionais sejam respeitadas.

Na presente metodologia propõe-se que NS seja definido em termos de tempo máximo de atendimento a um ponto de demanda a partir de um CD e como restrição será adotada uma área média desejada por CD. Com isto é possível chegar-se a um número desejado de zonas para a área de estudo.

O processo completo é apresentado abaixo:

Passo 1: Definição de NS

Define-se um NS baseado no Tempo de Viagem (T_α) do CD ao ponto de demanda.

O tempo $T\alpha$ será o tempo máximo admissível de viagem, ao nível de confiança $1-\alpha$.

Assim, para um tempo T a probabilidade de $T \leq T\alpha$ pode ser expressa por:

$$P(T \leq T\alpha) = 1- \alpha \qquad (4.1)$$

Calcula-se a seguir a distância ($D\alpha$) correspondente a $T\alpha$ utilizando-se uma velocidade de cruzeiro (V):

$$D\alpha = T\alpha .V \qquad (4.2)$$

Com $D\alpha$ determina-se a distância média de viagem

Assumindo-se $D \sim N(D_m, S_d)$, tem-se:

$$D\alpha = D_m + S_d. Z\alpha \qquad (4.3)$$

Sendo Cv o coeficiente de variação, tem-se:

$$S_d = Cv.D_m, \qquad (4.4)$$

Logo:

$$D\alpha = D_m + Cv.D_m.\ Z\alpha \qquad (4.5$$

Portanto:

$$D_m = D\alpha\ /\ (1 + Cv.\ Z\alpha) \qquad (4.6)$$

Passo 2: Restrição de Área

Estabelece-se uma Área Média por CD ($A_{média}$) baseada em restrição operacional.

O volume médio de armazenagem por CD ($V_{médio}$) será dado por:

$$V_{médio} = A_{média}.h\ .\eta_A \qquad (4.7)$$

onde : $A_{média}$ = área média do CD ;

h = pé direito ;

η_A = índice de aproveitamento de área construída;

O volume médio útil por CD ($V_{médio}$) é uma parcela do volume disponível para armazenagem, que em geral não é totalmente utilizado (CEL/COPPEAD, 2001). Assim, tem-se:

$$V_{ÚTIL} = V_{médio}.\ \eta_V \qquad (4.8)$$

onde : η_V = índice de aproveitamento do volume de armazenagem.

O volume médio mensal atendido por CD (V_{CD}) será este volume útil multiplicado pelo giro mensal de estoque:

$$V_{CD} = V_{ÚTIL}.G \qquad (4.9)$$

onde: G = giro mensal de estoque

Passo 3: Momento médio de transporte

Determina-se agora o momento médio de transporte por CD (TKM_{CD}) através de:

$$TKM_{CD} = V_{CD}.D_m \qquad (4.10)$$

Passo 4: Número de zonas

O número de CD´s (N_{CD}) é determinado através de:

$$N_{CD} = TKM_{Total} / TKM_{CD} \qquad (4.11)$$

onde: TKM_{tot} = momento de transporte total da área de estudo;

Este momento de transporte total da área de estudo deve ser computado com base na demanda dos pontos a atender e na posição destes pontos.

O número de zonas (N_{zonas}) será igual ao número de CD's

Faz-se então: $N_{zonas} = N_{CD}$ (4.12)

4.3.1.4. Passos da heurística de *"clusterização"*

A partir das considerações feitas até aqui é possível a definição de uma heurística que leve à formação de *clusters* que irão se constituir em zonas.

A idéia básica da proposta é que o processo se inicie com o cômputo das distâncias de todos os nós ao CD central determinando a distância média (μ) ao CD central e o seu desvio padrão (σ). Com isto definem-se em princípio três anéis: um anel de 0 Km até ($\mu - \lambda\,\sigma$) ; um segundo anel de ($\mu - \lambda\,\sigma$) a ($\mu + \lambda\,\sigma$) e um terceiro anel acima de ($\mu + \lambda\,\sigma$)., para $\lambda > 0$. Para anéis adicionais soma-se ao limite do último anel a quantidade $2\lambda\sigma$, até que se atinja o limite da área de estudo.

Este parâmetro λ exerce um papel de coeficiente de ajuste da dimensão dos anéis e do número de zonas às necessidades da área de estudo e do nível de serviço definido.

O número de zonas, na verdade, trabalha em conjunto com a largura dos anéis, redefinindo λ de forma a alcançar um equilíbrio de momento de transporte (MT) médio por zona entre os anéis. Este equilíbrio é buscado de forma a atender também às exigências de nível de serviço, conforme será visto a seguir.

Define-se inicialmente o número de zonas por anel com base no primeiro anel. Este anel deve ter apenas uma zona central para poder abrigar os candidatos no caso de se desejar testar um único CD central. Esta única zona é utilizada também como base central para a formatação das demais zonas. É a partir desta zona que as demais vão se formando no sentido radial.

O número de zonas dos demais anéis é calculado de forma a se manter a mesma média de MT por zona (balanceamento) e de não ultrapassar o número de zonas definido na seção anterior. O número de zonas dos outros anéis é definido então através da divisão do MT total de cada anel pelo MT do primeiro anel. Se este número não se igualar ao valor definido no nível de serviço, deve-se ajustar o anel 1 de forma a atingir o valor desejado. O ajuste é feito alterando-se o valor do parâmetro λ.

Em seguida verifica-se se os anéis estão equilibrados em termos de momento de transporte.

Para se estabelecer quando um anel está desbalanceado, utiliza-se o seguinte critério:

a) Identifica-se o anel com menor média de MT por zona

b) Calcula-se o intervalo de 95% de confiança (IC_{95}) das médias de MT dos demais anéis (Steel e Torrie, 1980)

c) Há distorção em um anel se sua média estiver fora do IC_{95} das médias dos demais anéis.

Se houver anel desbalanceado, deve haver uma redefinição, distribuindo-se os pontos desse anel entre os demais.

Isto é feito com o uso de um critério baseado em distâncias.

c_1) Identifica-se a faixa de distância (distância entre o primeiro e último ponto do anel) em que se encontram 95% dos pontos do anel desbalanceado.

c_2) Divide-se esta distância pelo número total de zonas e agrega-se um múltiplo desta faixa de distância a cada anel, proporcionalmente ao número de zonas de cada um. Com isto o parâmetro λ fica automaticamente ajustado.

c_3) retorna-se ao passo (a).

Havendo redefinição de anéis, redefine-se também o número de zonas. O processo continua até que o equilíbrio *balanceamento de MT - nível de serviço* seja alcançado.

Definidos os anéis varre-se a nuvem de pontos dentro de cada anel, computando-se o momento de transporte entre todos os pontos consecutivos, o que irá possibilitar a montagem de *clusters*.

A descrição completa da heurística é apresentada abaixo:

ESTÁGIO NO.1: *CLUSTERIZAÇÃO* DOS NÓS DE DEMANDA

Passo 01: Computar a distância de todos os nós ao CD central, assim como a média (μ) e o desvio padrão (σ) correspondentes;

Passo 02: Definir os anéis: de 0 Km até $\mu-\lambda\sigma$; $\mu-\lambda\sigma$ a $\mu+\lambda\sigma$ e acima de $\mu+\lambda\sigma$, acrescendo $2\lambda\sigma$ conforme discutido acima.

Passo 03: Conforme discutido acima, definir o número de zonas por anel e verificar se os anéis estão balanceados em termos de MT. Em caso negativo, redistribuir os pontos do anel mais desbalanceado e verificar novamente o balanceamento. Seguir para o próximo passo somente quando o equilíbrio balanceamento de MT - nível de serviço ocorrer.

Passo 04 Em cada anel, ordenar os nós pelo ângulo em relação ao eixo X e varrer a nuvem de nós do anel, computando o MT entre todos os nós consecutivos;

Passo 05: Dentro de cada anel, dividir o MT total pelo número desejado de zonas (este valor será denominado TKM);

Passo 06: Identificar todos os nós "isolados" (vide critério abaixo) e separá-los;

Passo 07: Dentro de cada anel, identificar o maior ângulo entre dois nós consecutivos. A partir do 2º nó deste par, iniciar uma varredura dos pontos do anel no sentido anti-horário e agregar nós a uma zona, até atingir o valor mais próximo de TKM.

Ao iniciar uma nova zona, verificar o nó mais próximo do primeiro nó da zona. Um nó não pode ir para a zona seguinte, se está mais próximo do último nó da zona anterior;

Passo 08: Parar se todos os nós já foram varridos. Os nós "isolados" devem ser agregados à zona do seu nó mais próximo.

A identificação de ponto "isolado" é feita através de:

Passo 1. Calcule a distância de cada nó ao seu nó anterior ($Dist_{ANT}$) e nó posterior ($Dist_{POST}$), na seqüência de varredura para todos os pontos de um anel;

Passo 2. Calcule intervalos com níveis de 90% e 95% confiança para a média de $Dist_{ANT}$ ou de $Dist_{POST}$ (as médias deverão coincidir);

Passo 3: Um nó será considerado “isolado” se tiver uma das $Dist_{ANT}$ e $Dist_{POST}$ na faixa de 90% e outra em 95%, ou então ambas na faixa de 95%. Este nó deve ser separado dos demais para posterior tratamento.

4.3 2. Seleção dos “Melhores” Candidatos a CD’s

Uma vez que os nós são agregados em *clusters,* estes passam a se constituir em uma nova entidade que foi considerada como sendo uma zona. E da mesma forma que cada nó era inicialmente um candidato em potencial a ser o local de um CD, esta nova entidade “zona” deve agora fornecer pelo menos um candidato a CD. Deve-se então, nesta etapa, identificar em cada zona a “melhor” posição para um candidato. Esta posição deve ser um dos nós da rede logística considerada.

Em princípio, o bem conhecido procedimento do centro de gravidade (COG) apresentado no capítulo 2 nas equações 2.1a a 2.2b, poderia atender a esta necessidade. O que se observa, entretanto, é que as posições assim geradas dificilmente se localizam exatamente em um nó efetivo da rede. Como o COG trabalha com uma área contínua no plano e, portanto, com infinitos pontos, a possibilidade de um único ponto coincidir com um nó da rede passa a ser muito pequena. Além disto, a posição gerada pelo COG pode eventualmente, se situar em área imprópria para a instalação de CD’s, devido à, por exemplo, questões de legislação ou mesmo por se tratar de área muito valorizada. O COG, portanto, por si só não atende integralmente à necessidade.

Por outro lado, desconsiderar o uso do procedimento do COG seria abandonar algo que no mínimo, poderia ser considerado como uma indicação muito clara do local onde deveria se situar um CD para atender a cada zona. Então, se o resultado do COG não coincidir com um nó da rede, é razoável se supor que a melhor localização *na rede* deverá ser algum nó próximo do local proposto. Portanto, se o resultado do COG não pode ser a solução que se busca, pode pelo menos ser considerado como um ponto de partida para um procedimento de busca do "melhor" local na rede.

Este local fornecido pelo COG pode assim ser considerado como a melhor posição do ponto de vista teórico, já que outros fatores podem tornar inviável o seu uso do ponto de vista prático. Este local será então denominado daqui para frente de "melhor posição teórica" (MPT). A figura 4.2 mostra como se apresentariam as MPT's em uma área já segmentada em zonas.

Observe-se que as MPT's formam uma configuração que espelha a própria formação das zonas. Essa configuração é, portanto, um aspecto importante a ser mantido conforme a discussão sobre formato das zonas desenvolvida na seção 4.3.1.2. Candidatos selecionados no entorno das MPT's irão também espelhar essa formação.

O "melhor" local na rede deverá estar localizado no entorno da MPT, mas não necessariamente o nó mais próximo da MPT será o "melhor" local, pois a solução é função das posições e volumes dos pontos de demanda e também das outras MPT's. Dependendo da posição relativa das MPT's e dos nós no entorno da MPT sob análise até mesmo um outro nó um pouco mais distante poderá se constituir na "melhor" solução.

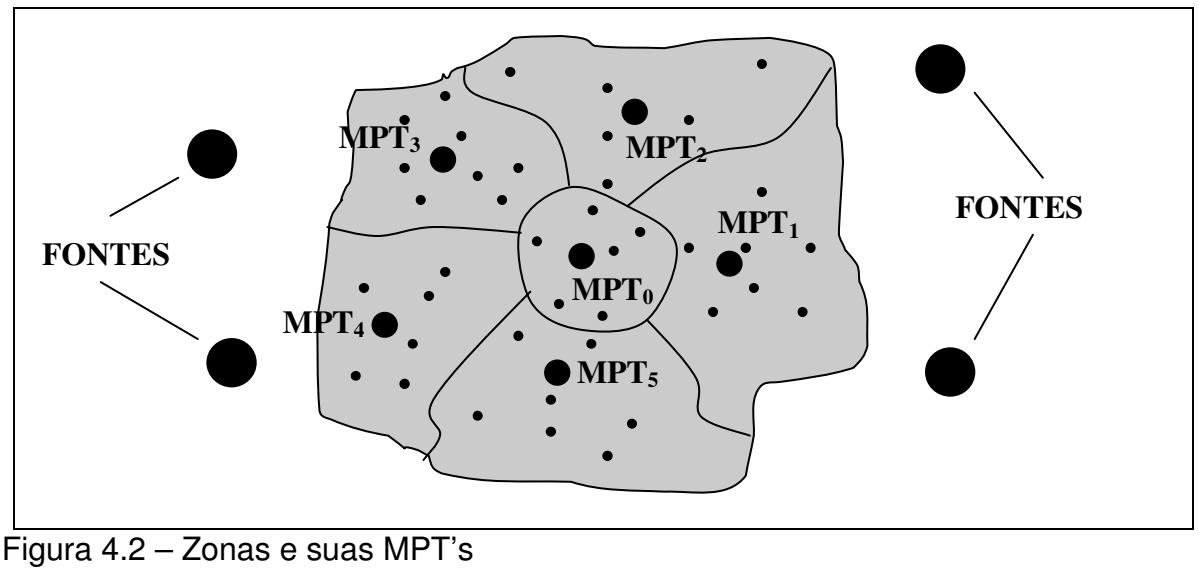

Figura 4.2 – Zonas e suas MPT's

O que deve ser feito então é o desenvolvimento de um processo de busca na rede dos nós considerados "mais próximos" das MPT's geradas para cada zona. Todos esses nós "mais próximos" deverão ser considerados como candidatos a CD´s.

Por outro lado, a proximidade entre as MPT's assim definidas e as localizações ótimas de CD's irá depender do número de CD's que se deseja instalar na região. Supondo-se uma região com cinco zonas, conforme figura 4.3, então se o número desejado de CD's for de três CD's, possivelmente nem todas as melhores localizações estarão no entorno de MPT's.

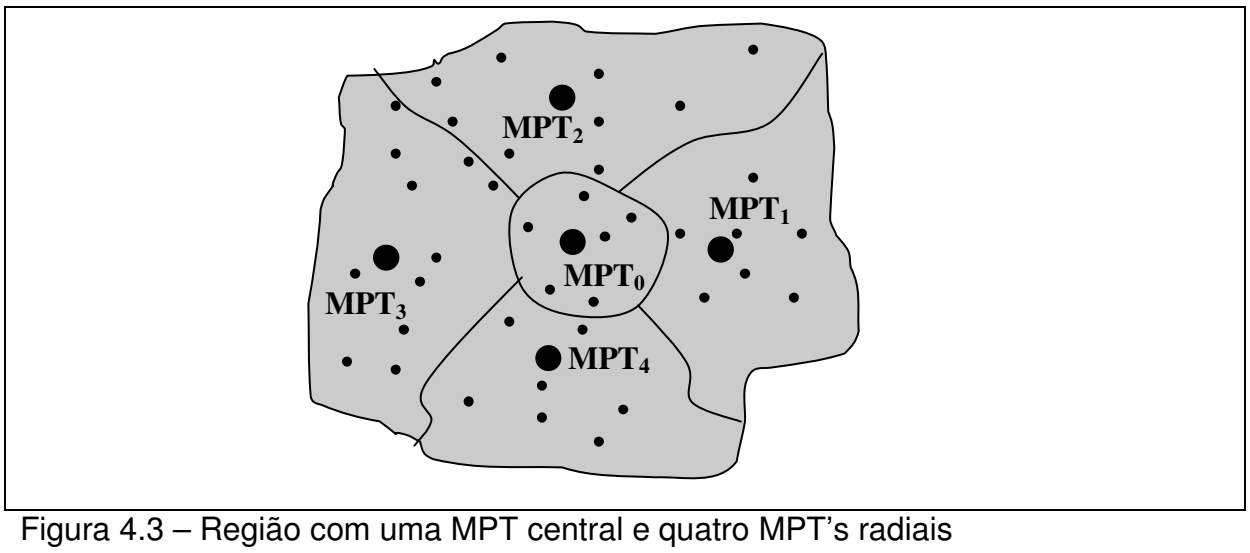

Figura 4.3 – Região com uma MPT central e quatro MPT's radiais

As figuras 4.4a e 4.4b apresentam duas possíveis configurações para estas localizações.

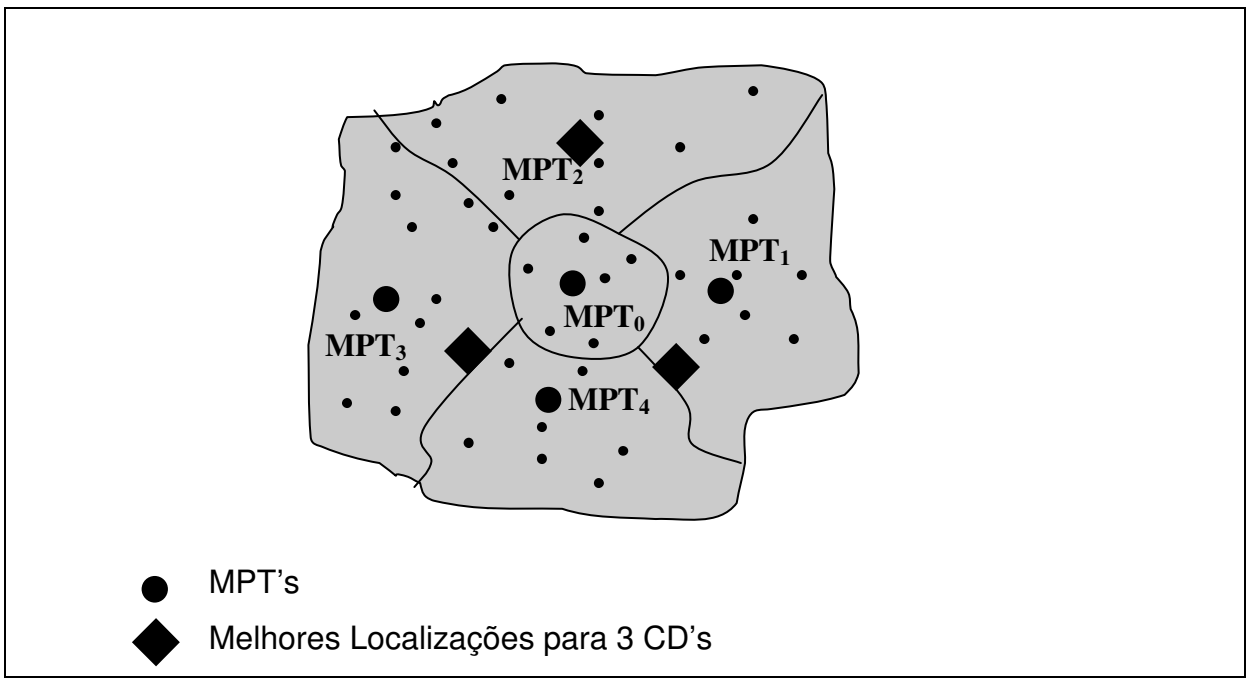

Figura 4.4a –MPT's vs. Melhores localizações – possibilidade 1

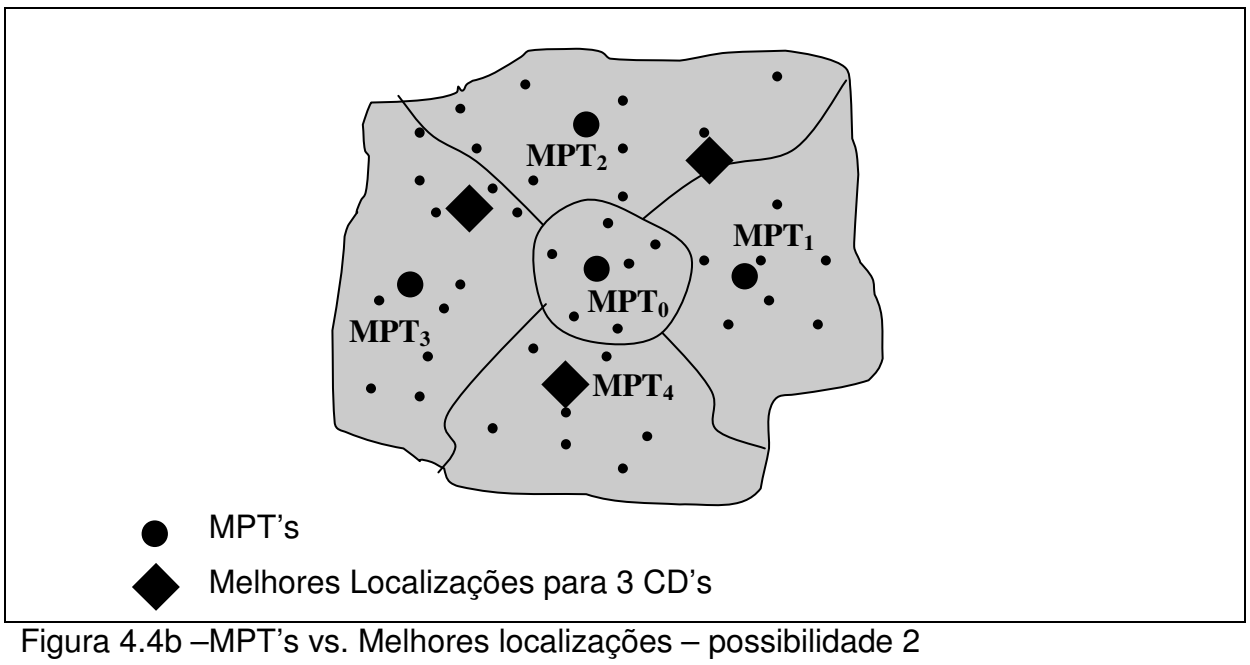

Figura 4.4b –MPT's vs. Melhores localizações – possibilidade 2

Na primeira possibilidade somente a MPT2 coincide com uma localização e as outras duas estão entre as MPT's 3 e 4 e entre as MPT's 4 e 1. Na segunda possibilidade mostra-se que se a MPT 4 coincidir com uma localização de CD então os outros dois CD's poderão se situar entre as MPT's 1 e 2 e MPT's 2 e 3.

Esta questão se resolve estabelecendo-se pontos intermediários entre as MPT's para localizar candidatos a CD's. Assim, estes pontos intermediários passam a se constituir também em pontos em que no seu entorno, se desenvolverá uma procura de nós na rede. Estes pontos intermediários se denominarão MPT's Meio. Uma MPT Meio é determinada como a média das coordenadas de duas MPT's adjacentes ponderadas pelas respectivas demandas. Assim, o seu posicionamento tenderá a se deslocar para a região que mais necessita de CD´s.

Quanto à questão da adjacência de MPT's, utiliza-se aqui a teoria dos grafos (Boaventura, 1977; Novaes, 1989). Neste caso, pode-se definir um grafo G(N,A) constituído de N nós correspondentes às MPT's e A arcos unindo pares de MPT's que estariam em zonas fronteiriças. Utilizando o conceito de adjacência de nós de um grafo, duas MPT's ligadas por um arco seriam consideradas adjacentes. As zonas dessas MPT's serão então também consideradas adjacentes.

4.3.2.1. Heurística do raio máximo

Esta heurística tem por objetivo encontrar as MPT's de cada zona e as respectivas MPT's Meio e proceder a uma busca em seu entorno para identificar os nós que serão considerados os "melhores" candidatos a CD's. Uma vez encontrada a MPT, a heurística estabelece um raio de busca no seu entorno, chamado R_{max},, dentro do qual a procura pelos candidatos na rede acontece.

Assim: $R_{max} = \beta D_{ij}$ (4.13)

onde: β = coeficiente redutor do raio ($0 \leq \beta \leq 0,5$);

D_{ij} = distância entre posições teóricas i, j adjacentes.

Com $\beta \leq 0,5$ garante-se que um nó selecionado para a MPT i sempre estará mais próximo desta do que de qualquer outra MPT.

A adjacência de MPT's é definida conforme discutido na seção anterior. As zonas fronteiriças adjacentes são em princípio todas as zonas que circundam uma dada zona. Assim, sendo i uma zona interna a um anel, as zonas adjacentes à zona i serão as zonas i-1 e i+1. Agora, se i = número de zonas do anel, então a teórica adjacente i+1 é substituída pela zona 1. As zonas de um dado anel terão ainda zonas adjacentes nos anéis superior e inferior, quando estes existirem.

Os nós encontrados dentro de R_{max} serão os "melhores" candidatos que se procurava. O procedimento é visualizado através da figura 4.5.

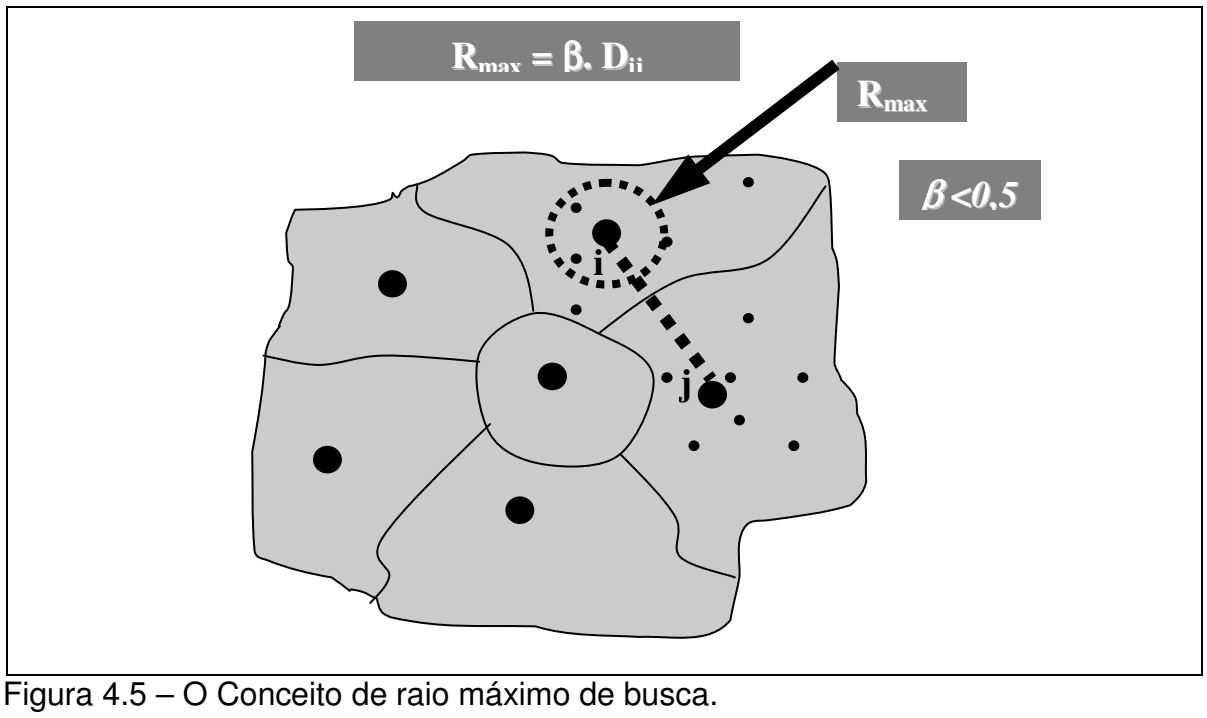

Figura 4.5 – O Conceito de raio máximo de busca.

A definição de R_{max}, é feita a partir de um procedimento que toma por base a distribuição de probabilidades da distância (D^z) dos nós de demanda de cada zona às suas respectivas MPT's.

Para o desenvolvimento do procedimento teve-se o cuidado de verificar o comportamento desta distribuição de probabilidades (vide Apêndice B). Isto foi feito a partir de um conjunto de dados reais em que se tinham aproximadamente duzentos e trinta pontos de demanda. Estes pontos estavam distribuídos por praticamente toda a região sul do Brasil. Estes dados foram, inclusive, utilizados posteriormente na aplicação prática apresentada no capítulo 7. Nestes dados foi aplicada a heurística de *clusterização,* gerando um zoneamento e em seguida calculadas as MPT's de cada zona. Foram então construídas distribuições empíricas para diferentes conjuntos de zonas, inclusive para a área total de estudo. Os diferentes testes mostraram que dentre as distribuições analisadas – Gama, Erlang, Normal, Beta e Rayleigh – a que apresentou os melhores resultados foi a Distribuição Normal.

Assumindo-se, portanto, que D^z segue uma distribuição Normal com média μ_z e desvio padrão σ_z , $D^z \sim N(\mu_z,\sigma_z)$, R_{max} será definido com base no limite inferior dessas distâncias para um certo nível de confiança desejado (Steel e Torrie, 1980).

A formalização deste princípio pode ser feita definindo-se inicialmente uma distância limite $D^z_{\beta o}$, como sendo:

$$D^z_{\beta o} = d^z - Z_{\beta o}.s^z \qquad (4.14)$$

onde:

d^z = média de distâncias dos nós da zona z à sua MPT, que será uma estimativa de μ_z

$Z_{\beta o}$.= valor crítico da N(0,1) ao nível 1-β_o de confiança

s^z = desvio padrão das distâncias dos nós da zona z à sua MPT, que será uma estimativa de σ_z

Assim, R_{max} pode então ser definido por:

$$R_{max} = \begin{cases} 0, & se\ D^z_{\beta_o} < 0 \\ D^z_{\beta_o}, & se\ 0 \le D^z_{\beta_o} \le 0{,}5.D_{ij} \\ 0{,}5.D_{ij}, & se\ D^z_{\beta_o} > 0{,}5.D_{ij} \end{cases} \qquad (4.15)$$

Com isto tem-se o coeficiente β dentro dos limites estabelecidos em sua definição ($0 \le \beta \le 0{,}5$).

O detalhamento dos passos desta heurística é apresentado a seguir:

ESTÁGIO 2A: HEURÍSTICA DO RAIO MÁXIMO

Passo 1: Fazer i=1;

Passo 2: Selecionar a zona i e seguir para o passo 3;

Passo 3: Encontrar para a zona i sua MPT – "melhor" posição teórica utilizando o procedimento do centro de gravidade;

Passo 4: Computar as distâncias de todos os nós da zona i para sua MPT e determinar o raio máximo da zona i (R_{max});

Passo 5: Identificar na zona i os nós dentro da área circular definida por R_{max}, tendo como centro a MPT da zona. Estes serão os "melhores" candidatos a CD's;

Passo 6: Fazer i=i+1. Parar se i=número de zonas. Caso contrário, retornar ao passo 2.

Em seguida a heurística é aplicada as MPT's Meio.

4.3.2.2. Heurística do ângulo máximo

Se nenhum nó da rede é encontrado dentro do raio máximo de uma dada MPT ou MPT Meio, um novo raio (R_{max1}) é definido para este caso. Isto é feito diminuindo-se o nível de confiança desejado.

Assim, tem-se:

$$R_{max1} = \mu_D - z_{\alpha 1}.\sigma_D \qquad (4.16)$$

sendo $\alpha_1 > \alpha$

A heurística do raio máximo é então novamente aplicada. Porém, como o raio foi incrementado, os nós encontrados dentro da nova faixa não podem agora ter um ângulo (θ) muito aberto em relação ao eixo que une duas MPT's adjacentes, de forma a não distorcer a configuração de MPT´s. Note-se que manter esta configuração é um aspecto importante para se propiciar a geração de candidatos que realmente espelhem as localizações potenciais de CD´s, conforme discutido na introdução da seção 4.3.2.

Então é necessário agora desenvolver uma nova verificação para identificar os nós que estão dentro de R_{max1} e que não apresentem um ângulo muito aberto com aquele eixo (figura 4.6). Isto é verificado em termos da área do setor circular formado pelo segmento que une o ponto em questão e a MPT e o próprio ângulo θ. Assim, dentre os pontos dentro de R_{max1} escolhe-se aquele que apresentar o setor circular com menor área.

Estágio 2b: HEURÍSTICA DO ÂNGULO MÁXIMO

Passo 1: Faça i=1;

Passo 2: Selecione a zona i e siga para o passo 3;

Passo 3: Computar o novo raio máximo da zona i (R_{max1});

Passo 3: Compute o ângulo θ pata todos os nós da zona i dentro de R_{max1};

Passo 4: Usando o ângulo θ e a distância do nó à MPT determine a área do setor circular de todos os nós selecionados no passo 3. Separe o nó que apresentar o setor circular com menor área;

Passo 5: Faça i=i+1. Pare se i=número de zonas, em caso contrário, retorne ao passo 2.

Os nós selecionados nos estágios 2a e 2b serão os "melhores" candidatos que estavam sendo procurados.

Estas heurísticas serão avaliadas na seção 4.5, através de um experimento de validação.

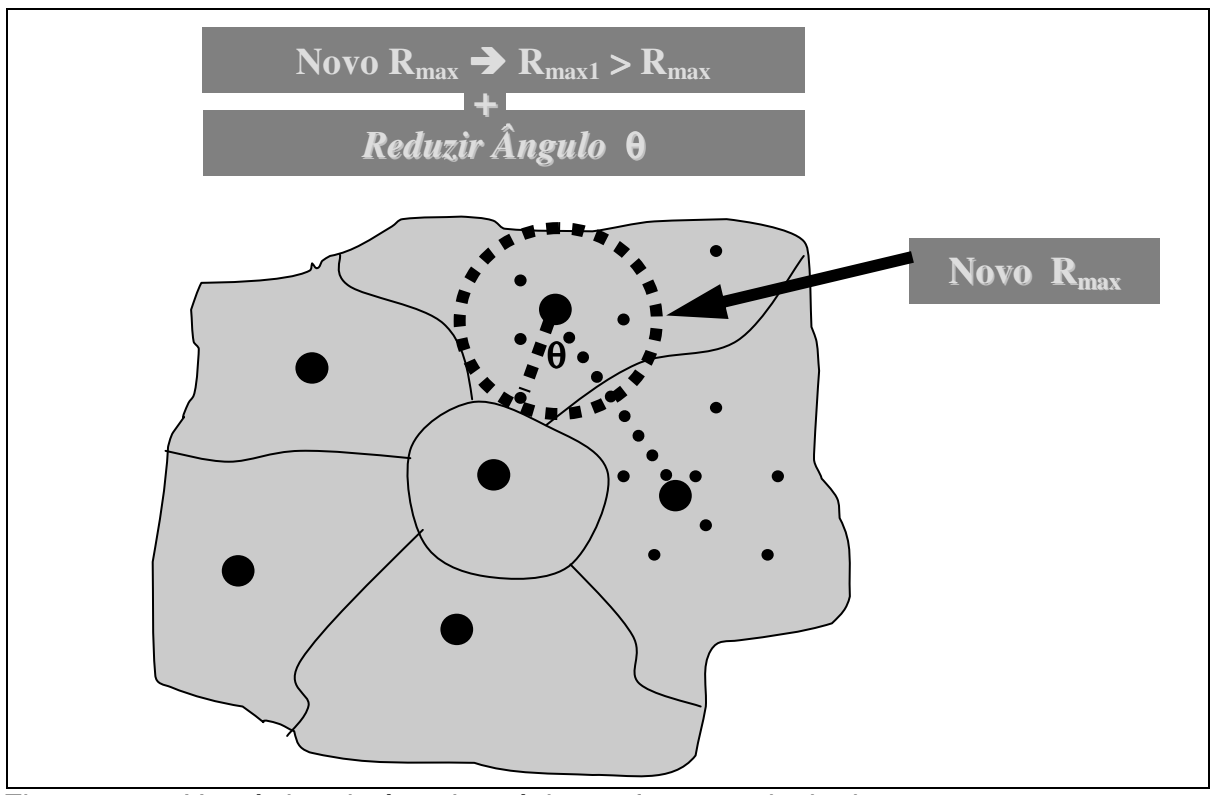

Figura 4.6 – Heurística do ângulo máximo – fatores principais

4.4. MODELO DE OTIMIZAÇÃO PARA LOCALIZAÇÃO DOS CD's

A construção do modelo de otimização foi baseada na idéia de p-medianas (Daskin, 1995; Crainic e Laporte, 1997), em alguns aspectos apresentados por outros autores, como Geoffrion e Graves (1974) e nos modelos clássicos de transporte com transbordo (Novaes, 1978 e 1989).

O modelo aqui proposto procura utilizar aspectos importantes de cada abordagem e acrescenta outras características com o intuito de procurar adequá-lo às necessidades da própria metodologia.

Os índices utilizados no modelo são descritos abaixo:

i = índice para as fontes de produção; i=1,2,.....I; I=número de origens.

j = índice para zonas de demanda; j=1,2,.....J; J=número de zonas.

k = índice para CD's; k=1,2,.....K; K=número de candidatos.

Considerando-se que o modelo será utilizado exclusivamente para localizar CD's e não para determinar a sua quantidade ótima, a sua função objetivo foi construída de forma a representar somente os custos envolvidos neste processo, quais sejam: custos de transporte referentes à transferência de carga entre as fontes de produção e os CD's, e custos de distribuição dos CD's para as regiões de demanda, bem como os custos de localização que evoluem de acordo com a necessidade de volumes a movimentar, representando a necessidade de área de armazenagem.

O modelo completo é apresentado na seqüência.

MODELO DE PROGRAMAÇÃO LINEAR INTEIRA MISTA

$$Mín \sum_{i,k} c_{ik}.x_{ik} + \sum_{k,j} c_{kj}\, D_j.y_{kj} + \sum_{k} f_k.Y_k \qquad (4.17)$$

Sujeita a:

$$\sum_{k} y_{kj} = 1 \qquad (j = 1,2,.....J)$$. zona é atendida por único CD (4.18)

$$\sum_{k} Y_k = p$$. deve-se ter p CD's (4.19)

$$y_{kj} - Y_k \leq 0 \qquad (k = 1,2,....K \;;\; j = 1,2,.....J)$$. zona é atendida por CD ativo (4.20)

$$\sum_{j} y_{kj} - Y_k \geq 0 \qquad (k = 1,2,.....K)$$. CD ativo atende ao menos uma zona (4.21)

$$\sum_{k} x_{ik} \leq O_i \qquad (i = 1,2,.....I)$$. restrições de oferta (4.22)

$$\sum_{i} x_{ik} = \sum_{j} D_j.y_{kj} \qquad (k = 1,2.....K)$$. atendimento da demanda (4.23)

$$y_{kj} = 0,1 \qquad (k = 1,2,....K \;;\; j = 1,2,....J)$$ 0 = CD não atende zona (4.24)
1 = CD atende zona

$$Y_k = 0,1 \qquad (k = 1,2,....K)$$ 0 = CD fechado (4.25)
1 = CD aberto

Notação - Função Objetivo

A função objetivo (equação 4.17) é composta de três termos que são analisados abaixo:

1º termo:

c_{ik} = custo unitário de transferência entre a região i de produção e o CD k;

x_{ik} = fluxo do produto entre a região i de produção e o CD k (deve-se ter $x_{ik} \geq 0$)

2º termo:

c_{kj} = custo unitário de distribuição entre o CD k e a zona j de demanda;

D_j = demanda da zona j;

y_{kj} = variável binária de decisão $y_{kj} = 0$ ...CD k não atende zona j;

$y_{kj} = 1$ ...CD k atende zona j;

3º termo:

f_k = custo fixo de localização do CD k;

Y_k = variável binária de decisão $Y_k = 0$ ...CD k é ativo;

$Y_k = 1$ ...CD k não é ativo.

Notação - Restrições:

O_i= oferta de produto na região de produção i;

As restrições (4.18) estabelecem que somente um y_{kj} para um dado valor de j (uma dada zona) pode ser igual a 1 (y_{kj} =1 implica que CD K atende à zona j). Define-se assim que cada zona é atendida por um único CD, o que não impede que cada CD atenda mais de uma zona. Geoffrion e Graves (1974) utilizam estas restrições e Daskin (1995) também as utiliza.

A restrição (4.19) é característica do modelo p-medianas, pois é esta que define que se deve ter p CD's. Esta restrição é importante, pois na *Fase 2* da metodologia serão estudados os custos logísticos globais para diferentes quantidades de CD´s.

Em (4.20) estabelece-se que cada zona é atendida por um CD ativo. Note-se que y_{kj} não pode ser 1 (zona j é atendida pelo CD k) se Y_k for zero (CD k não está ativo).

As restrições (4.21) foram introduzidas nesta proposta de modelo para garantir que se o CD k for ativo (Y_k=1) não ficará sem atender alguma zona, caso em que ter-se-

ia $\Sigma y_{kj}=0$. Se CD K é ativo deve atender uma ou mais zonas e, portanto, deve-se ter $\Sigma y_{kj}\geq 0$.

Note-se que sem esta restrição o modelo poderia decidir que um determinado CD não iria atender a nenhuma zona, ficando assim desativado. A *Fase 2* da metodologia, que está voltada para o cômputo do custo logístico global, seria assim prejudicada, pois ficar-se-ia sem a informação de uma parcela de custos para este valor p de CD´s. Com esta restrição obriga-se o modelo a fornecer uma solução para todas as quantidades de CD´s desejadas.

Em (4.22) e (4.23) tem-se condições típicas dos modelos de transporte. As restrições de oferta são representadas em (4.22), enquanto que as equações de continuidade nos CD's, em conjunto com o atendimento das demandas, são garantidas pelas restrições (4.23). E finalmente, as restrições 4.24 e 4.25 correspondem às condições das variáveis binárias (0,1) y_{kj} e Y_k.

Este modelo deve ser aplicado sucessivamente para diferentes valores de p. Para cada valor de p o modelo irá encontrar a localização ótima dos p CD´s. Em seguida, passa-se para a *Fase 2* da modelagem, em que se procura definir a quantidade "ótima" (p_0) de CD's que minimiza os custos logísticos globais. Definido o valor ótimo p_0, usa-se então a solução ótima de localizações (*Fase 1*) correspondente a p_0.

4.5. EXPERIMENTO DE VALIDAÇÃO DAS HEURÍSTICAS

O experimento desenvolvido para validar a metodologia proposta para a *Fase 1* foi baseado em um conjunto de pontos gerados aleatoriamente no plano. Novaes

(1989) mostrou que a distância entre pontos próximos pode ser representada pela distribuição de Rayleigh. Isto é feito imaginando-se um círculo de raio r (r>0) com o ponto P em seu centro. O autor procura estabelecer a função de probabilidade acumulada da distância d de um ponto P ao seu ponto mais próximo, *F(d)=P(d≤r)*, demonstrando que se a probabilidade de que existam n outros pontos (n>0) no círculo de raio r for dada por uma distribuição de Poisson, então F(d) será uma distribuição de Rayleigh, conforme abaixo:

$$F(d) = 1 - e^{-\lambda \pi d^2} \qquad (4.26)$$

onde: d= distância ao ponto mais próximo;

λ = densidade de pontos/km^2.

A partir desta distribuição foi então gerada uma nuvem de pontos que possibilitou uma experimentação com controle apurado.

Para a geração dos pontos foi definido um sistema de eixos cartesianos e sorteado um par de coordenadas para ser utilizado como semente. A partir desta semente foi gerado um raio *d* correspondente ao círculo de distâncias em que o ponto mais próximo à semente deveria se posicionar. Isto foi feito com o uso da distribuição de Rayleigh. Utilizando-se em seguida uma distribuição uniforme com faixa de variação entre 0° e 360° foi gerado o ângulo em que este ponto deveria se localizar no círculo, definindo-se assim sua posição final. Repetindo-se este processo a partir de cada ponto gerado foram definidas as posições de todos os pontos necessários ao experimento.

Para a geração destas distâncias na distribuição de Rayleigh e ângulos na distribuição uniforme foi utilizada Simulação de Monte Carlo, seguindo-se a metodologia proposta por Saliby (1989), que sugere o uso de amostragem descritiva. De forma análoga foram geradas demandas para os pontos a partir de uma distribuição normal.

Na amostragem descritiva que foi utilizada os valores são sorteados ao longo de toda a gama de variação da distribuição. São chamados valores descritivos justamente por permitir descrever toda a distribuição, evitando-se uma possível tendenciosidade devida a uma eventual concentração do sorteio de valores aleatórios em um determinado segmento da faixa de variação da distribuição. Em seguida a aleatoriedade do processo é garantida através de um sorteio da seqüência de ocorrência destes valores na corrida de simulação. Este sorteio da seqüência é feito permutando-se aleatoriamente a ordenação dos valores descritivos, conforme os passos abaixo:

i) Definir um vetor de índices $P(i) = i$, $i=1,2,...n$, que define a seqüência de uso dos n valores descritivos gerados;

ii) Gerar um número aleatório n_0 entre i e n;

iii) Permutar o valor de $P(n_0)$ com $P(i)$;

iv) Fazer $i=i+1$. Se $i>n$, encerrar o processo, caso contrário retornar ao passo ii.

Ao término do procedimento o vetor $P(i) = i$, $i=1,2,...n$ contém a seqüência aleatória de utilização dos valores descritivos na simulação.

Com isto garante-se total aleatoriedade, além de que efetivamente toda a gama de variação da distribuição estará representada. A tabela 4.2 mostra este processo. Na primeira parte da tabela é apresentado um conjunto de valores descritivos conforme sua ordem de geração. Na segunda parte da tabela, correspondente à permutação, é apresentada a seqüência em que os valores descritivos devem ser utilizados na simulação. Assim, por exemplo, o primeiro valor a ser utilizado na simulação deve ser o décimo segundo número gerado (0,3833), enquanto que o primeiro valor gerado (0,0167) deve ser o vigésimo terceiro a ser utilizado na simulação.

O resultado final de todo este processo foi a geração de uma nuvem de pontos aleatórios que está apresentada na figura 4.7.

Considerando-se assim estes pontos como nós de demanda em uma rede logística, foram aplicados todos os estágios da heurística proposta para a *Fase 1*, quais sejam:

Estágio 1: Heurística de *Clusterização*;

e

Estágio 2: Seleção dos “melhores” Candidatos

Estágio 2a: Heurística do Raio Máximo

Estágio 2b: Heurística do Ângulo Máximo

O processo tem início então pela Heurística de *Clusterização*. O ponto de partida desta heurística é a determinação de um CD central, em função do qual se definem anéis e posteriormente os *clusters*. A heurística estabelece que para a definição do posicionamento deste CD seja empregado o método do centro de gravidade discutido no capítulo 2.

Tabela 4.2 – Amostragem descritiva – geração e permutação

GERAÇÃO		PERMUTAÇÃO		
Ordem de Geração	**Valor Descritivo (entre 0 e 1)**	**Seqüência de uso na Simulação**	**Ordem de Geração**	**Valor Descritivo**
1º	0,0167	1º	12º	0,3833
2º	0,0500	2º	28º	0,9167
3º	0,0833	3º	7º	0,2167
4º	0,1167	4º	11º	0,3500
5º	0,1500	5º	15º	0,4833
6º	0,1833	6º	5º	0,1500
7º	0,2167	7º	13º	0,4167
8º	0,2500	8º	14º	0,4500
9º	0,2833	9º	9º	0,2833
10º	0,3167	10º	2º	0,0500
11º	0,3500	11º	10º	0,3167
12º	0,3833	12º	19º	0,6167
13º	0,4167	13º	24º	0,7833
14º	0,4500	14º	18º	0,5833
15º	0,4833	15º	22º	0,7167
16º	0,5167	16º	21º	0,6833
17º	0,5500	17º	20º	0,6500
18º	0,5833	18º	25º	0,8167
19º	0,6167	19º	29º	0,9500
20º	0,6500	20º	23º	0,7500
21º	0,6833	21º	6º	0,1833
22º	0,7167	22º	16º	0,5167
23º	0,7500	23º	1º	0,0167
24º	0,7833	24º	26º	0,8500
25º	0,8167	25º	27º	0,8833
26º	0,8500	26º	8º	0,2500
27º	0,8833	27º	30º	0,9833
28º	0,9167	28º	4º	0,1167
29º	0,9500	29º	3º	0,0833
30º	0,9833	30º	17º	0,5500

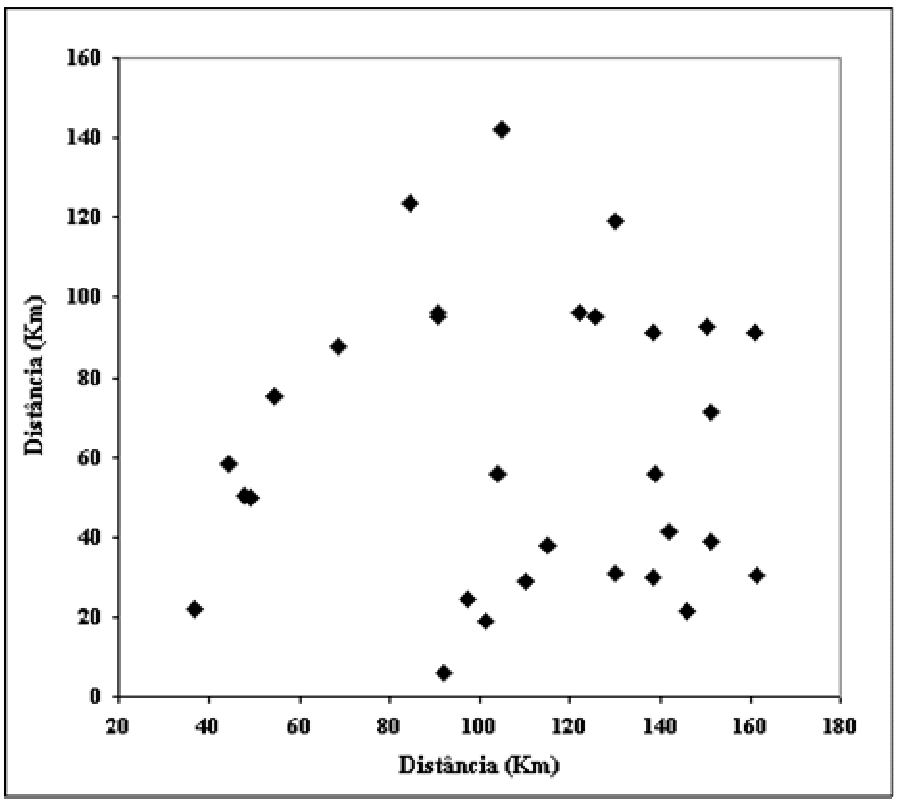

Figura 4.7 - Nuvem aleatória de pontos para o experimento

A partir deste ponto central foram definidos anéis e setores atendendo aos requisitos de homogeneidade entre os anéis, conforme estabelecido na heurística, chegando-se a uma configuração final de uma única zona no primeiro anel e quatro zonas em um segundo anel (figura 4.8). O objetivo do experimento neste estágio foi o de avaliar a validade da heurística para um dado número de zonas. Não foi necessário, assim, utilizar o conceito de nível de serviço para definir esse número de zonas. Mesmo porque este conceito de nível de serviço foi detalhadamente trabalhado na aplicação prática desenvolvida na pesquisa, como será visto no capítulo 7.

Tendo-se assim um zoneamento definido, seguiu-se para o estágio No. 2 da heurística, correspondente à seleção dos “melhores” candidatos. Este estágio começa pela heurística do Raio Máximo, onde inicialmente são definidas as MPT’s

de cada zona, conforme apresentado na figura 4.9. As figuras 4.10a e 4.10b já apresentam uma visão exclusiva das MPT's e das MPT's Meio.

Pelas figuras pode-se verificar o formato circular característico da configuração de MPT's.

Finalizando-se a heurística do raio máximo e complementando-a com a heurística do ângulo máximo, identifica-se no entorno das MPT's e das MPT's Meio um conjunto de candidatos a CD's (figuras 4.11a e 4.11b).

É possível notar que ao final do processo passa-se a ter o conjunto de candidatos em um formato que circunda a área de estudo, percorrendo praticamente todo o seu entorno.

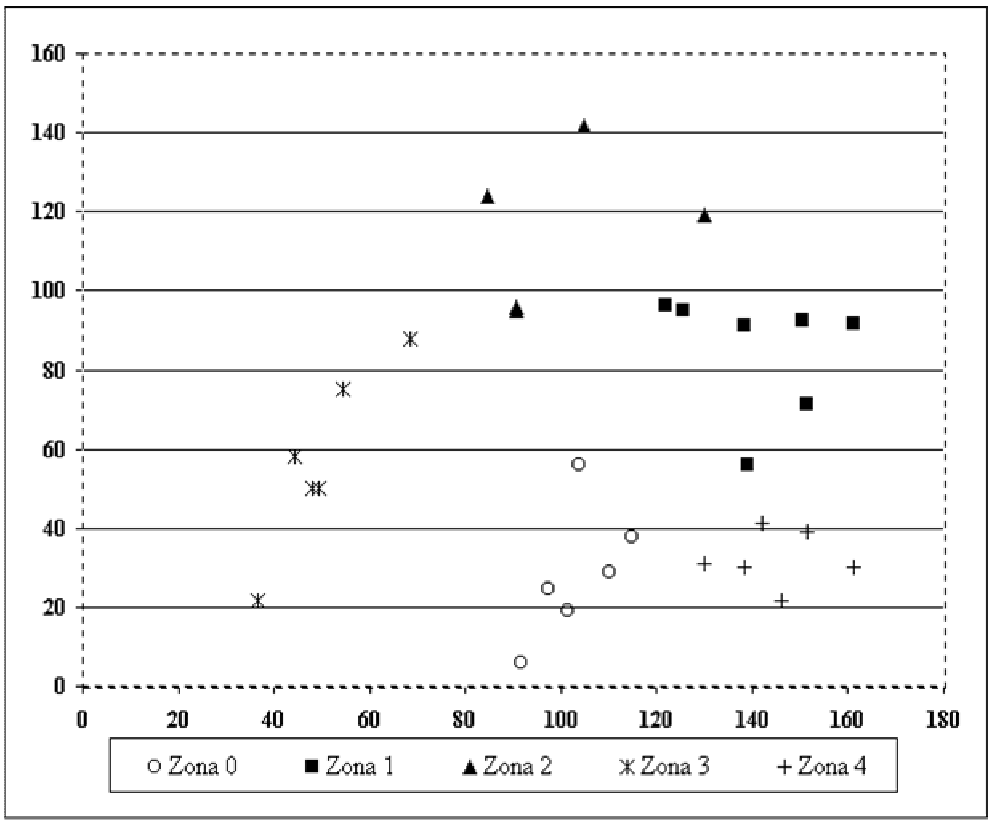

Figura 4.8 Heurística de *clusterização* - Proposta de zoneamento

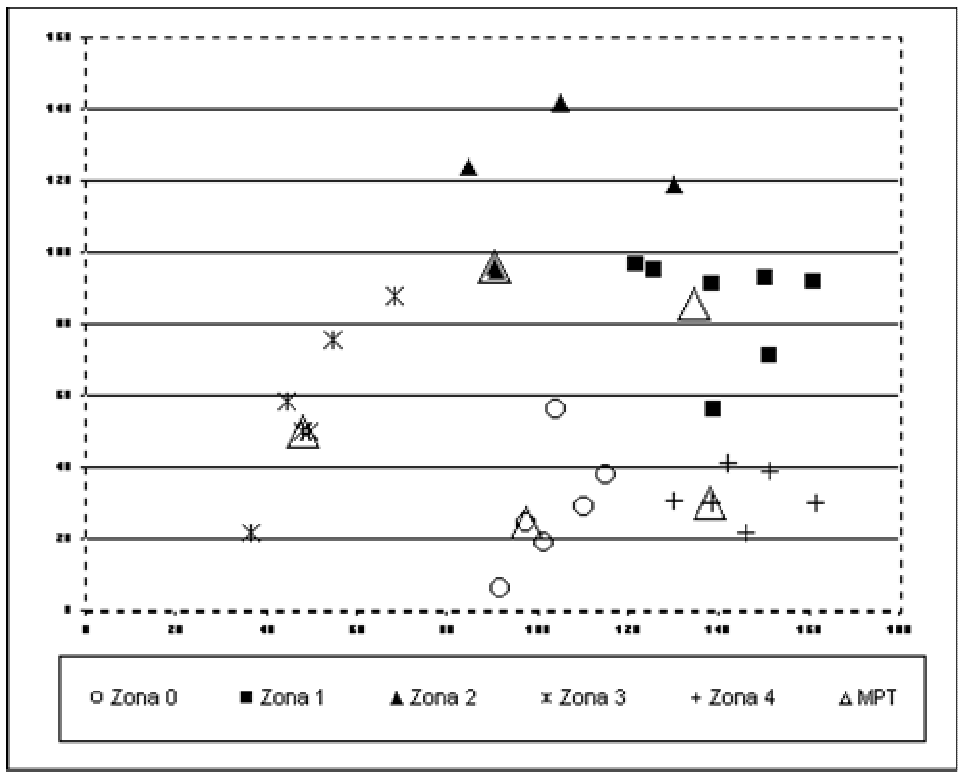

Figura 4.9 Zoneamento definido e as MPT's de cada zona.

Os candidatos foram utilizados como dados de entrada para o modelo de otimização, que teve necessidade também dos custos de transporte para alimentar os coeficientes de sua função objetivo. Estes custos foram extraídos de uma aplicação prática (Vallim, 1997) que apresentava custos para transferência de produtos entre fontes de origem e CD's e de distribuição dos CD's para os pontos de demanda. Com isto garante-se que os dados de entrada mantêm a proporcionalidade real de uma situação prática. A figura 4.12 apresenta a solução ótima obtida para cinco CD's, onde se pode verificar que as posições ótimas circundam a área de estudo, localizando-se de forma espaçada em uma formação que leva a uma abrangência efetiva de praticamente todo o seu entorno.

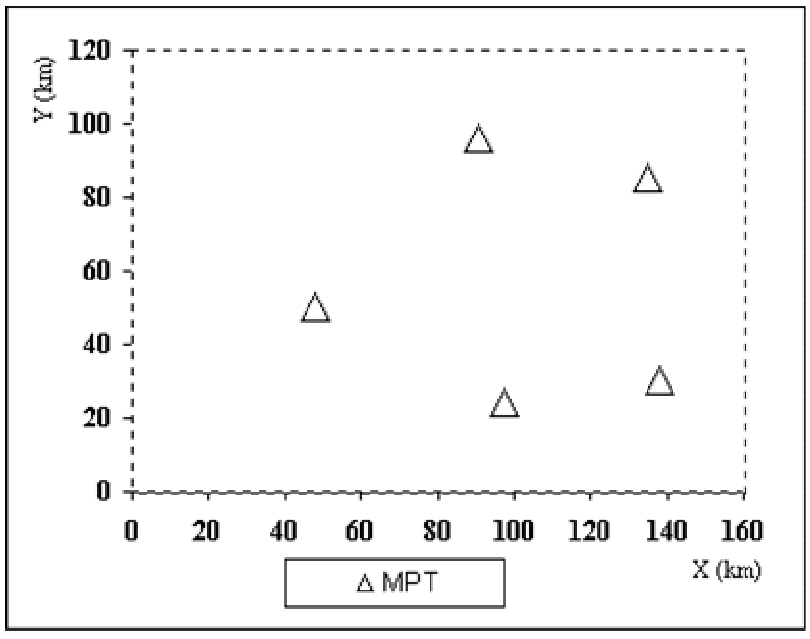

Figura 4.10a Visão geral das MPT's

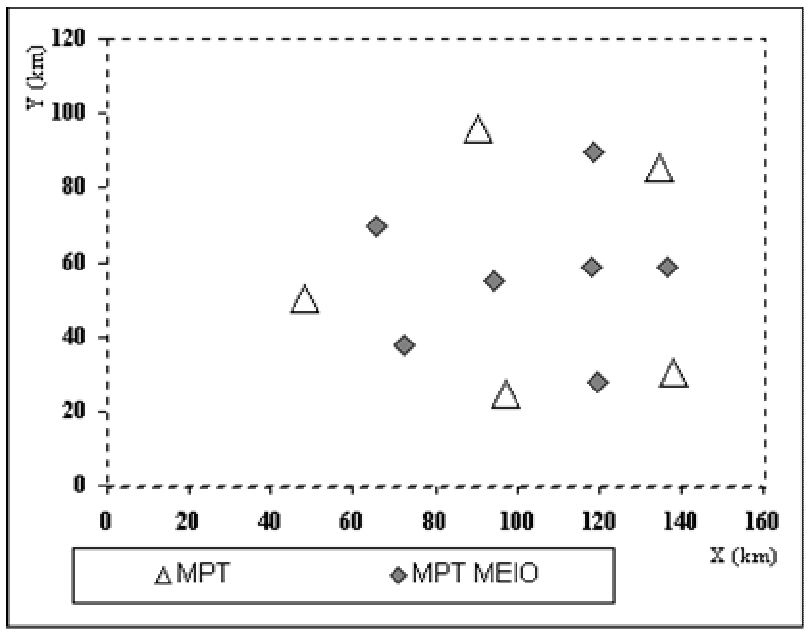

Figura 4.10b MPT's e MPT's Meio, geradas entre MPTs adjacentes

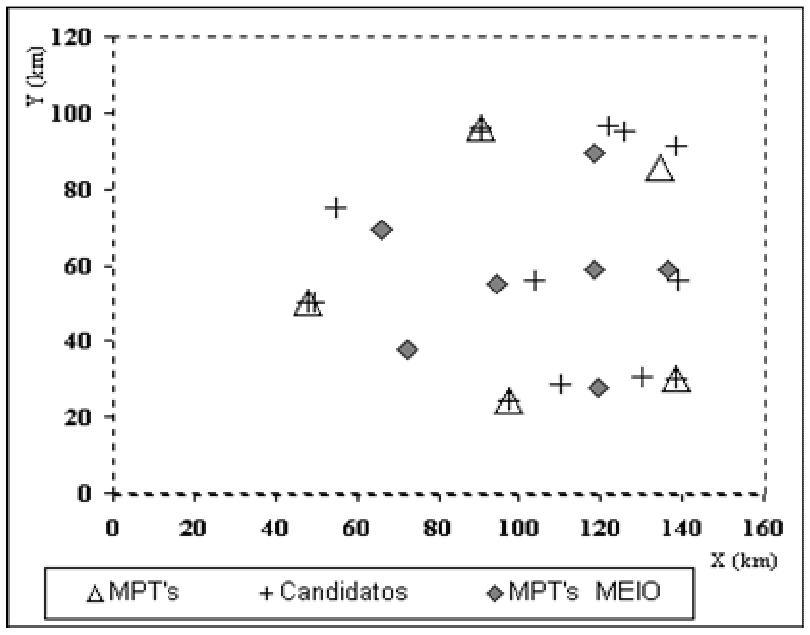

Figura 4.11a Candidatos gerados no entorno das MPT's

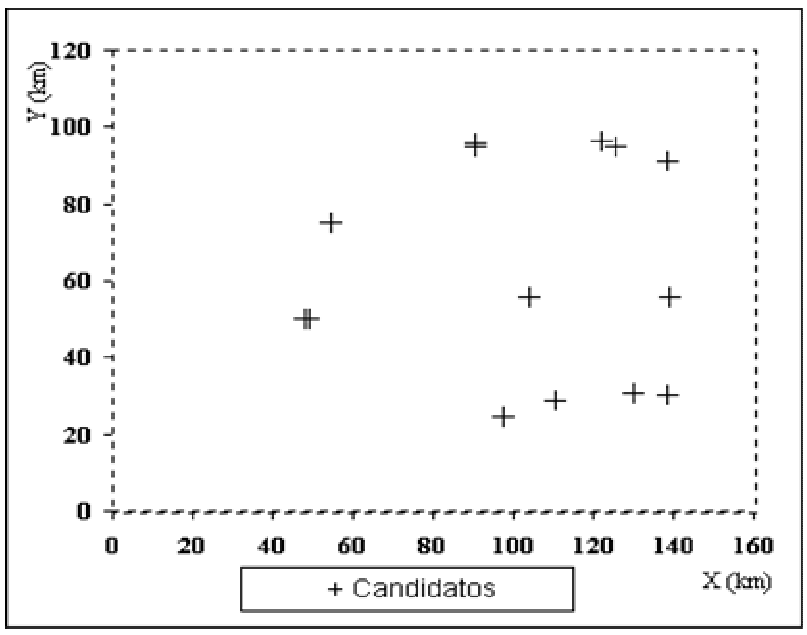

Figura 4.11b Visão geral dos Candidatos

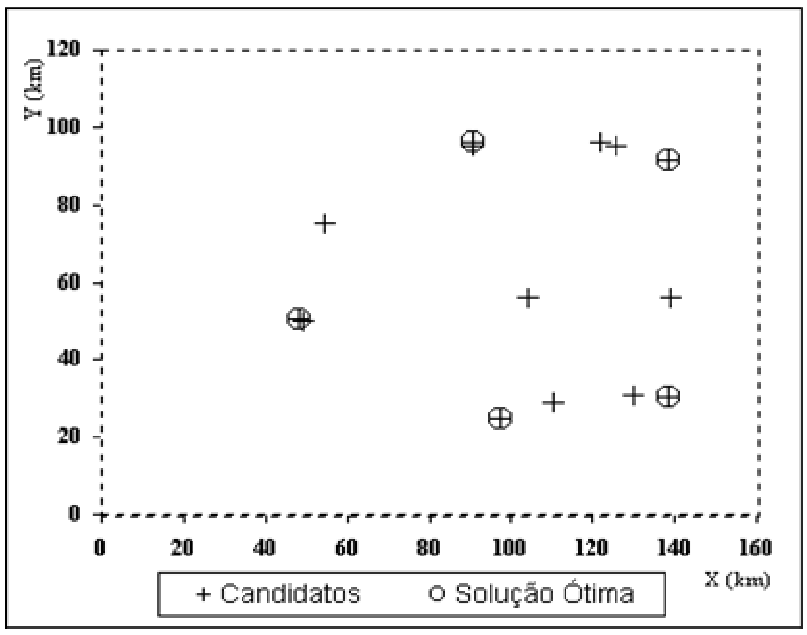

Figura 4.12 - Solução ótima gerada a partir dos candidatos

Os resultados obtidos precisavam ser confrontados com referenciais estabelecidos de forma a validá-los. Para isto procurou-se compará-los a procedimentos qualitativos clássicos. Quatro critérios foram empregados para desenvolver uma agregação geográfica dos nós de demanda: pontos cardeais (norte, sul, leste e oeste), quatro quadrantes (sudeste, nordeste, sudoeste e noroeste), faixas verticais e faixas horizontais. Os resultados foram ainda comparados com um quinto critério proposto por Ballou (1991b) em um artigo em que o autor, apesar de não tratar especificamente da questão dos candidatos a CD's, sugere um procedimento para agregação de pontos de demanda.

O primeiro critério, zoneamento segundo os pontos cardeais, foi definido a partir do centro de gravidade da área estabelecendo-se duas retas a 45° em relação aos eixos. Os quadrantes seguiram procedimento semelhante, definindo-se duas retas perpendiculares aos eixos. Para as faixas verticais identificou-se a amplitude total da área entre as coordenadas horizontais e com este dado e, mais, estabelecendo-se

um número de faixas, definiu-se a amplitude para as faixas. De forma análoga, as faixas horizontais foram construídas a partir da identificação da amplitude total da área entre as coordenadas verticais e definindo-se faixas de igual amplitude entre os pontos extremos.

A heurística proposta por Ballou (1991b) para a montagem de *clusters* está baseada no conceito de proximidade geográfica dos pontos de demanda.

A idéia proposta por Ballou é que se desenvolva um processo de busca do *cluster* mais próximo de cada ponto e que se agregue o ponto a este *cluster*.

A descrição completa da heurística é apresentada abaixo:

HEURÍSTICA "BALLOU"

Passo 1: Inicie com todos os pontos e suas coordenadas. Cada ponto constituirá um *cluster*;

Passo 2: Compute as distâncias para todos os pares de pontos;

Passo 3: Identifique o par de pontos com menor distância. Combine este par em um novo *cluster*. O número de *clusters* é reduzido em 1;

Passo 4: As coordenadas para este *cluster* serão aquelas do centro de gravidade entre os pontos;

Passo 5: Repita este procedimento de agregação até que o número de *clusters* desejado seja atingido.

Os zoneamentos resultantes destes cinco critérios e mais o zoneamento proposto nesta pesquisa são apresentados nas figuras 4.13a a 4.13f.

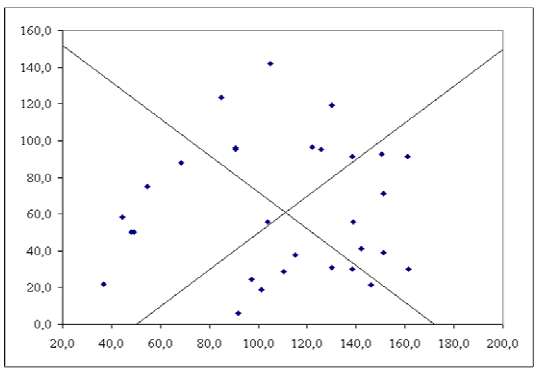

Figura 4.13a Zoneamento por Pontos cardeais

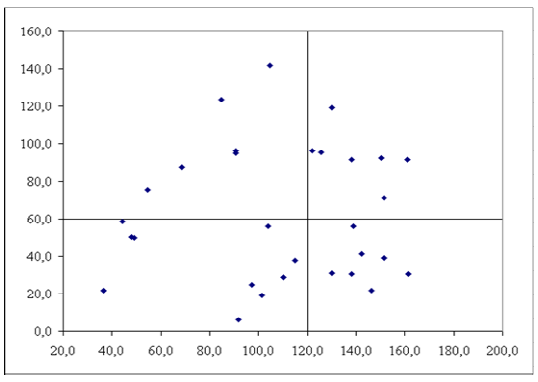

Figura 4. 13b Zoneamento em Quadrantes

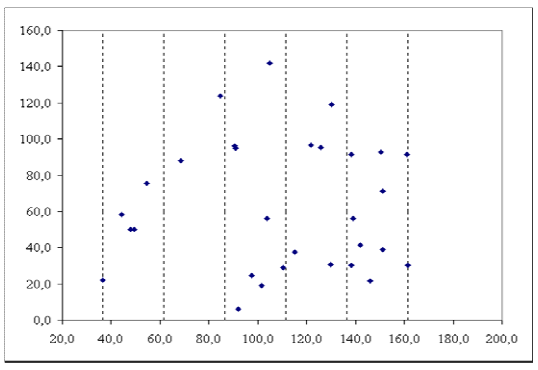

Figura 4.13c Zoneamento em faixas verticais

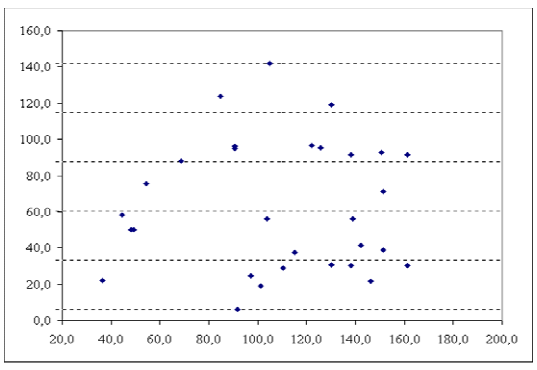

Figura 4.13d Zoneamento em faixas horizontais

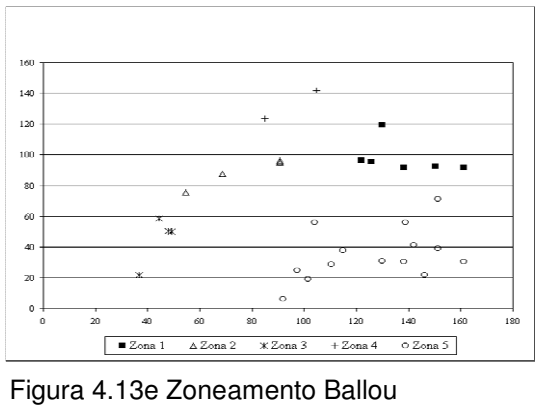

Figura 4.13e Zoneamento Ballou

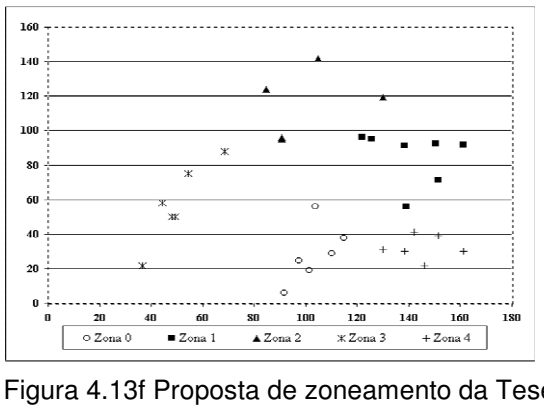

Figura 4.13f Proposta de zoneamento da Tese

Na seqüência foram selecionados candidatos para cada um dos cinco tipos zoneamentos. Esta seleção foi feita de forma qualitativa, através de uma observação da região e procurando-se selecionar candidatos ao longo de toda a área de estudo, a exemplo de Daganzo (1984). Este é um procedimento subjetivo, que pode depender da visão de cada observador. De forma a procurar minimizar esta subjetividade, mas ao mesmo tempo manter o procedimento em bases qualitativas, foi desenvolvido um processo de varredura da área de estudo que procurou representar aquilo que um observador dentro do *senso comum* faria para selecionar de forma qualitativa candidatos a CD's.

A idéia foi criar uma seleção qualitativa de candidatos, mas ao mesmo tempo tentar tornar o procedimento isento de tendenciosidades que poderiam surgir em função de características específicas de cada observador.

Procurou-se assim criar um cenário de isenção, estabelecendo uma seqüência bem definida que simulasse com fidedignidade o processo executado por um observador representando um padrão de *senso comum* que estivesse selecionando visualmente candidatos a CD's. Com isto estabelece-se um padrão claro e bem definido que pode ser reproduzido.

Os candidatos foram selecionados por zona. Para se estabelecer a quantidade a selecionar em cada zona, determinou-se o percentual de pontos de cada zona em relação ao total de pontos da área de estudo. Assim, para um dado número total de candidatos, define-se de forma proporcional quantos candidatos deve-se ter em cada zona. A quantidade total de candidatos foi definida como sendo o mesmo número de candidatos que foram selecionados na heurística proposta nesta pesquisa. Desta forma as comparações são feitas em igualdade de condições.

Os candidatos são selecionados de forma a serem representativos de um grupo de outros pontos da zona. Assim, para um dado número n_i de candidatos a selecionar na zona i, define-se n_i grupos de pontos próximos entre si. Tem-se, portanto, n_i *clusters* de pontos. Para cada *cluster* escolhe-se visualmente um ponto como candidato.

Internamente às zonas deve-se ainda estabelecer a proporção de pontos da zona em cada *cluster*. Assim, se a zona i tem m_i pontos cada *cluster* terá em média m_i / n_i pontos. No processo de varredura deve-se procurar manter esta média em cada *cluster*.

A montagem de *clusters* requer um processo de varredura de cada zona em que os pontos vão sendo agregados a um *cluster* à medida que vão sendo “varridos”. A varredura, entretanto, deve ser feita de forma racional seguindo o trajeto que o “olho” de um observador independente seguiria.

Para proporcionar precisão a este processo e possibilidade de reprodução a área de estudo foi subdividida em quadrículas (figura 4.14) e a varredura feita quadrícula a quadrícula garantiu total abrangência da área.

Outro aspecto a considerar é a direção e sentido de varredura. Em cada zona as quadrículas foram varridas em camadas formadas por avanços no sentido horizontal e vertical, alternadamente, de maneira a formar como que uma *onda* que avança por toda a área da zona. Esta *onda* pode se iniciar em qualquer quadrícula extrema da zona (cantos sudeste, nordeste, noroeste ou sudoeste), dependendo do seu formato. Supõe-se que este deve ser o trajeto que o “olho” de um observador faria para pesquisar a área e selecionar candidatos.

A título de ilustração apresenta-se na figura 4.15 o zoneamento em pontos cardeais, com a área subdividida em quadrículas e os *clusters* definidos dentro de cada zona.

Este processo foi desenvolvido para cada um dos cinco critérios de zoneamento que foram utilizados como referenciais de comparação, chegando-se, em cada caso, à definição do conjunto de candidatos a CD's.

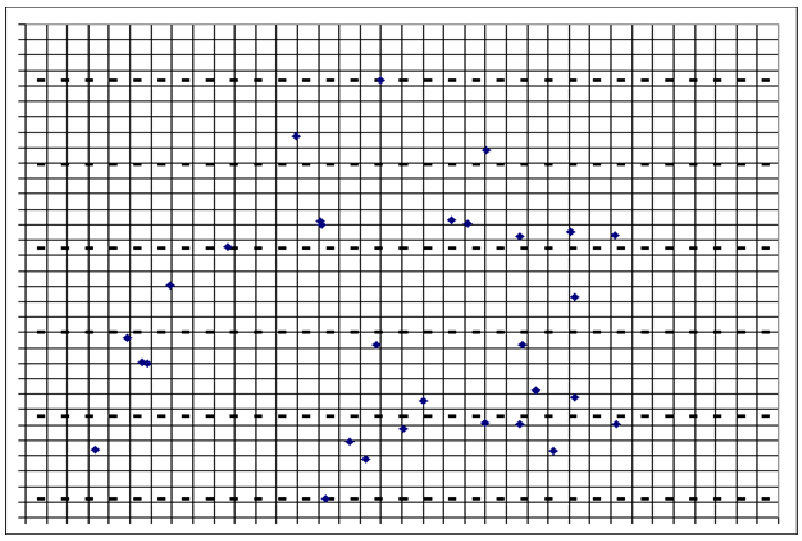

Figura 4.14 - Quadrículas
Ilustração com zoneamento em Faixas Horizontais

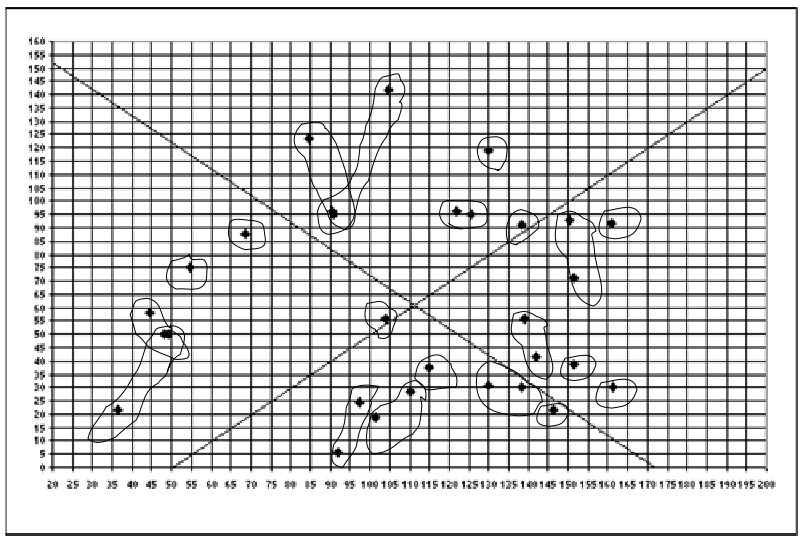

Figura 4.15 - Quadrículas e *clusters*
Ilustração com zoneamento em Pontos Cardeais

Os dados destes candidatos foram inseridos no modelo de otimização e as soluções ótimas foram comparadas com os resultados obtidos para a proposta da pesquisa. As figuras 4.16 a 4.20 apresentam os resultados, comparando-os com a solução ótima obtida anteriormente através da metodologia proposta nesta pesquisa.

Através destas figuras percebe-se que a distribuição dos pontos correspondentes à solução ótima ao longo da área de estudo é sempre mais abrangente na solução obtida com a proposta desta tese, quando comparada com as soluções obtidas através dos outros critérios. As soluções encontradas com os outros não se apresentaram bem distribuídas por toda a área de estudo. Em determinadas situações os CD's ficaram deslocados para um determinado lado e em outras situações um CD ficou deslocado situando-se em um local com baixa incidência de pontos de demanda, como no caso do zoneamento Ballou. Estes tipos de situações não ocorreram com a solução ótima obtida com a proposta da pesquisa.

As figuras 4.16 a 4.20 apresentam uma comparação visual das soluções para os casos de quatro e cinco CD's. Entretanto, testes foram executados com o número de CD's variando de um a cinco, para todos os critérios de zoneamento. A comparação entre estes resultados e a proposta da pesquisa é apresentada na tabela 4.3. Naquela tabela mostra-se a diferença percentual entre o custo total da operação na solução ótima em cada critério e a proposta da tese, para cada número de CD's testado.

As configurações de CD's obtidas nas soluções ótimas da proposta da pesquisa apresentaram custos menores ou iguais aos fornecidos pelos outros critérios em todas as situações testadas.

Com estes resultados tem-se uma indicação da eficácia da metodologia proposta, uma vez que em todos os casos testados a proposta da pesquisa fornece soluções tão boas ou melhores que aquelas obtidas através de critérios tradicionalmente utilizados neste tipo de problema.

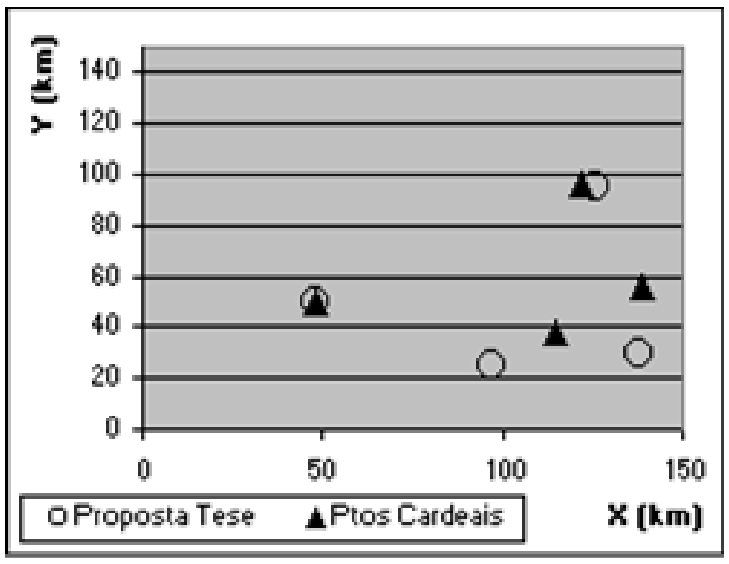

Figura 4.16 - Proposta Tese vs. Pontos Cardeais

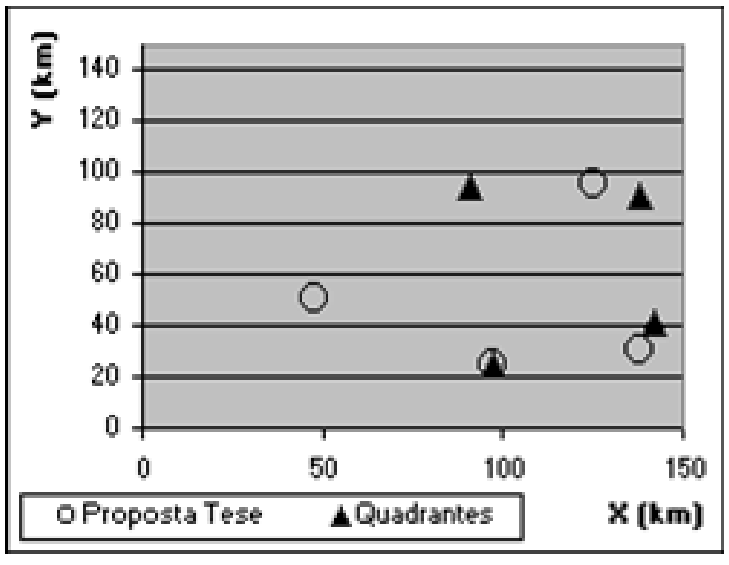

Figura 4.17 - Proposta Tese vs. Quadrantes

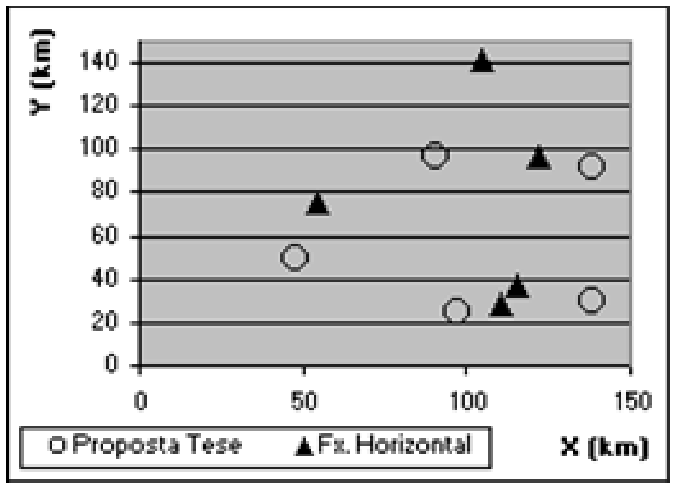

Figura 4.18 - Proposta Tese vs. Faixas Horizontais

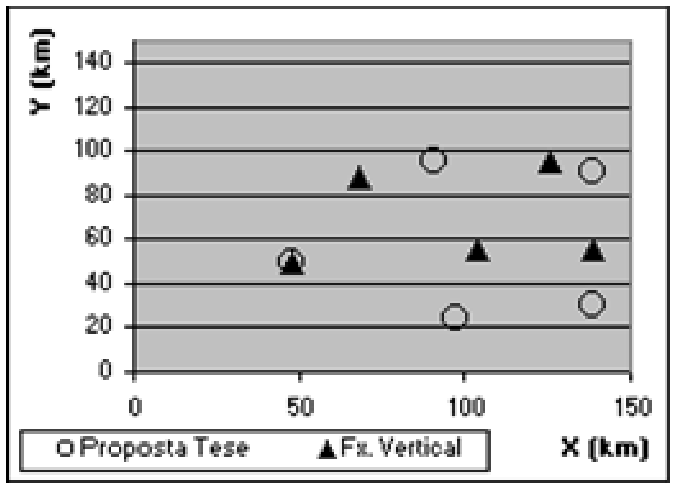

Figura 4.19 - Proposta Tese vs. Faixas Verticais

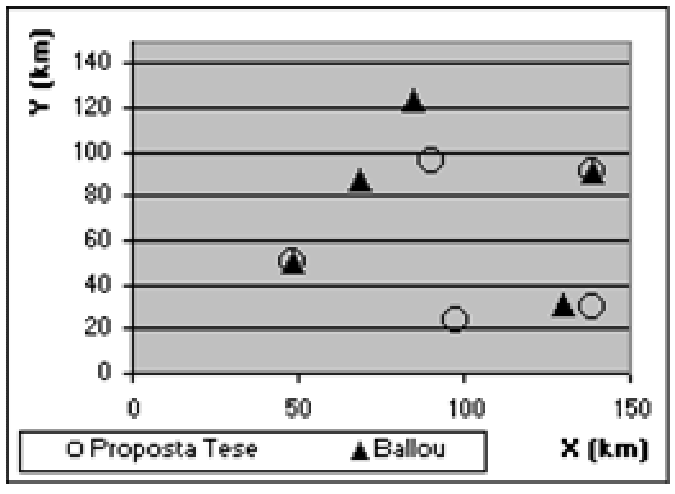

Figura 4.20 - Proposta Tese vs. Proposta Ballou

Tabela 4.3 – Proposta da pesquisa vs. outros critérios - Soluções ótimas
Diferença percentual nos custos da operação

No. de CD's	**ΔCusto**				
	Pontos Cardeais	Quadrantes	Faixas Horizontais	Faixas Verticais	Ballou
5	----	----	+ 34,52%	+ 22,78%	+ 5,62%
4	+ 3,88%	+ 8,92%	+ 27,52%	+ 13,15%	+ 1,23%
3	+ 2,41%	+ 12,26%	+ 18,93%	+ 8,40%	+ 0,74%
2	+ 0,75%	+ 3,06%	+ 3,25%	+ 0,19%	+ 0,04%
1	0,00%	0,00%	+ 6,33%	0,00%	+ 3,48%

Como observação final deve-se salientar que de forma a se proceder a uma validação final da metodologia proposta foi construído um modelo MILP desagregado para solucionar o problema. Neste caso, não houve a aplicação, portanto, do processo de *clusterização*. Neste modelo cada ponto de demanda foi considerando de forma individual. Isto foi feito para que se pudesse identificar qual seria a efetiva solução ótima para o problema. Em alguns casos a solução ótima encontrada para este modelo foi a mesma obtida pelo modelo MILP que recebeu como dados de entrada os resultados da heurística proposta nesta pesquisa. Em outros casos foi muito próxima. Uma análise completa destes resultados é apresentada no capítulo 5, seção 5.4.

4.6 VARIANTES NO USO DAS HEURÍSTICAS

Ã medida que se trabalha com a geração de zonas, uma idéia que surge naturalmente é a de se associar o número de zonas ao número de CD's que se deseja localizar. Adotando-se este procedimento, para se localizar n CD's desenvolver-se-ia um processo de agregação dos pontos de demanda em n zonas.

Com um número de zonas igual ao número de CD's é possível se definir o número de candidatos necessário apenas para o dado número de CD's que se tem interesse e não para quantidades maiores do que aquela em análise.

Nesta linha de raciocínio, considerando-se inicialmente uma primeira situação em que n zonas são geradas, tem-se n MPT's e um conjunto de q_1 candidatos gerados no entorno dessas n MPT's ($q_1 \geq n$). Caso se queira localizar p CD's ($p<n$) e sejam geradas p zonas, ter-se-á agora p MPT's. No entorno destas p MPT's será gerado um conjunto de q_2 candidatos ($q_2 \geq p$). Como se tem $p<n$, há uma tendência de se ter $q_2<q_1$ o que gera modelos MILP menores.

Para se testar esta idéia foi desenvolvido um experimento para se identificar a localização de um, dois três e quatro CD's para o mesmo conjunto de pontos utilizados no experimento descrito na seção 4.5.

A estes pontos foi então aplicada a heurística, porém agora para diferentes números de zonas. Lembrando que a heurística desenvolvida tem dois estágios, a seqüência seguida foi:

Estágio 1: Heurística de "Clusterização";

e

Estágio 2: Seleção dos "melhores" Candidatos

Estágio 2a: Heurística do Raio Máximo

Estágio 2b: Heurística do Ângulo Máximo

Iniciando-se pela Heurística de *Clusterização* partiu-se do mesmo CD central já definido na seção 4.5 para cuja definição foi empregado o método do centro de gravidade. A partir deste ponto central foram definidos anéis e setores conforme estabelecidos na heurística, chegando-se a uma configuração final para as zonas.

As figuras 4.21a a 4.21d apresentam os resultados da heurística para quatro, três, duas e uma única zona.

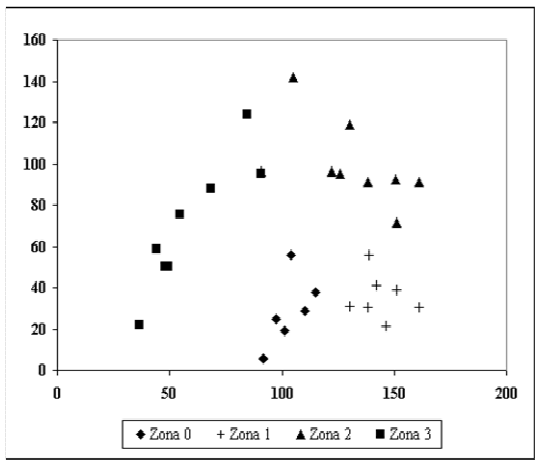

Figura 4.21a - *Clusterização* – 4 Zonas

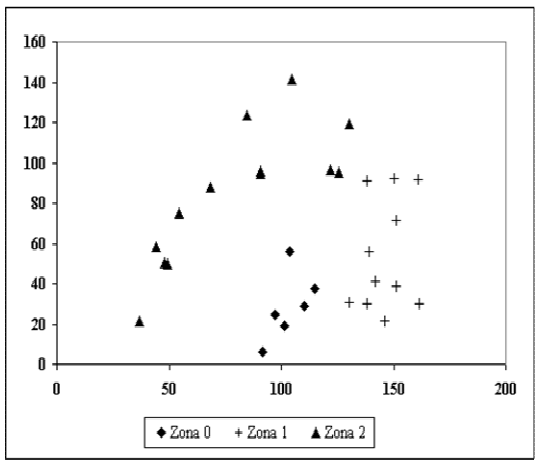

Figura 4.21b - *Clusterização* – 3 Zonas

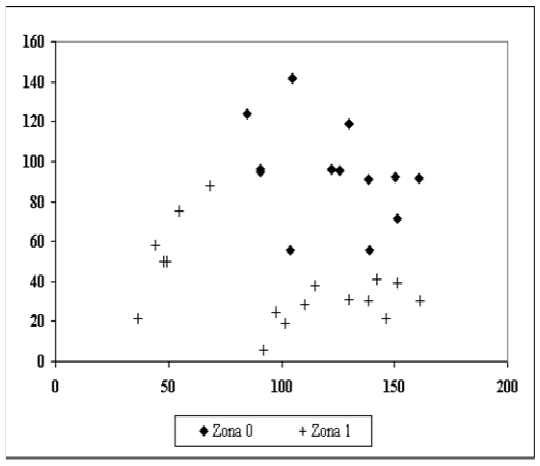

Figura 4.21a - *Clusterização* – 2 Zonas

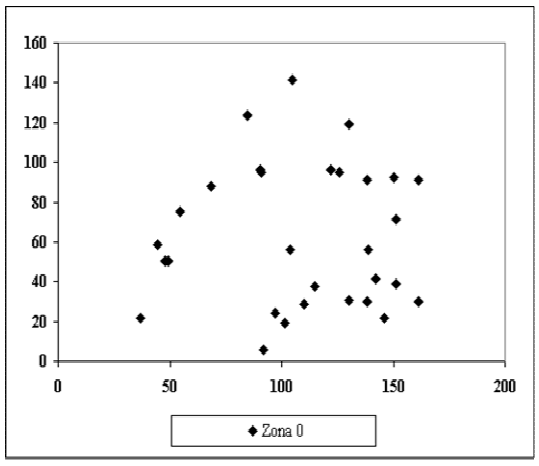

Figura 4.21b - *Clusterização* – 1 Zona

Será então a partir destes quatro zoneamentos definidos que se aplicará o estágio 2 da heurística, de forma a identificar os “melhores” locais candidatos a CD’s. Como se tem uma variação de uma a quatro zonas, será possível localizar de um a quatro CD’s.

A partir do zoneamento definido seguiu-se então para o estágio No. 2 da heurística.

Este estágio tem seu início com a heurística do Raio Máximo. Nesta heurística são identificadas inicialmente as MPT's de cada zona. O resultado desta etapa pode ser observado nas figuras 4.22a a 4.22d, que apresentam as MPT's para cada zoneamento desenvolvido.

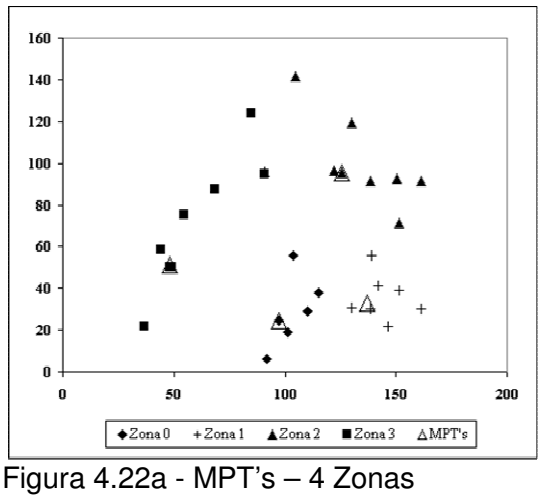

Figura 4.22a - MPT's – 4 Zonas

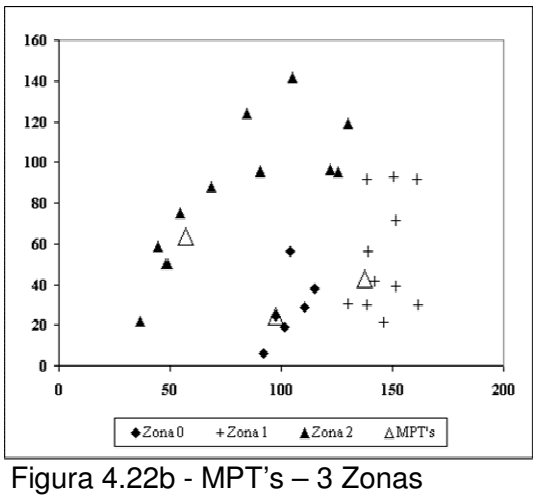

Figura 4.22b - MPT's – 3 Zonas

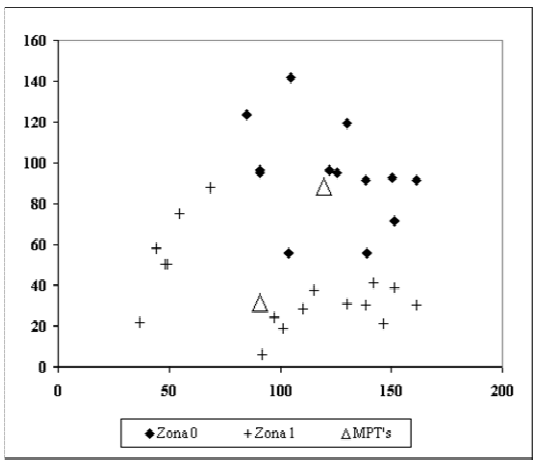

Figura 4.22c - MPT's – 2 Zonas

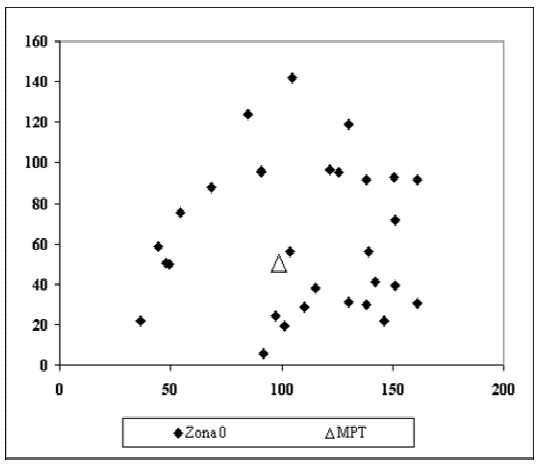

Figura 4.22d - MPT's – Zona única

Note-se que neste caso não foi desenvolvido o procedimento para se gerar as MPT's Meio. Desta vez isto não foi necessário, porque como já se tem uma definição de um número de zonas igual ao número de CD's a localizar, as MPT's Meio perdem sua função. A razão de uma MPT Meio é justamente compensar a diferença entre número de zonas e número de CD's.

Na seqüência o processo deve ser complementado com a aplicação da heurística do Ângulo Máximo.

Será a partir desta heurística que serão identificados os candidatos finais que deverão se posicionar no entorno das MPT's. Este conjunto de candidatos a CD's pode ser visualizado através das representações gráficas apresentadas nas figuras 4.23a a 4.23d.

Note-se que ao final do processo os candidatos se posicionam de forma a circundar a área de estudo. Aqui, porém, não se localizam em todo o seu entorno, como ocorreu no experimento anterior (seção 4.5). Isto não ocorreu, pois no experimento anterior havia um maior número de MPT's cujo objetivo é justamente permitir uma efetiva distribuição das posições teóricas ao redor de toda a área de estudo.

Os candidatos a CD's foram utilizados como dados de entrada para o mesmo modelo de otimização utilizado na seção 4.5, tendo-se obtido soluções ótimas para cada quantidade de CD's testada.

Os resultados deste novo experimento foram então confrontados com os resultados obtidos no experimento anterior, em que o número de zonas era fixo. A título de uma comparação inicial visual são apresentadas as figuras 4.24a a 4.24h.

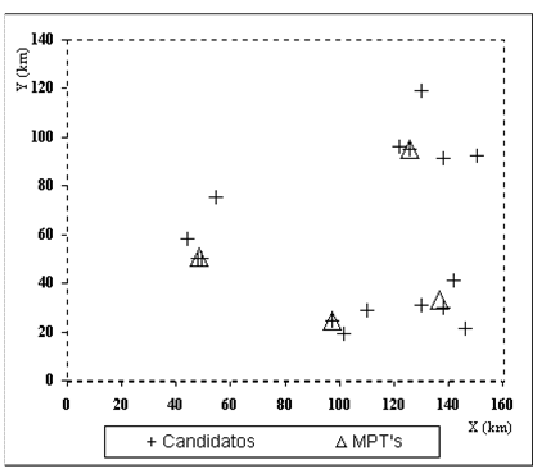

Figura 4.23a - Candidatos - 4 Zonas

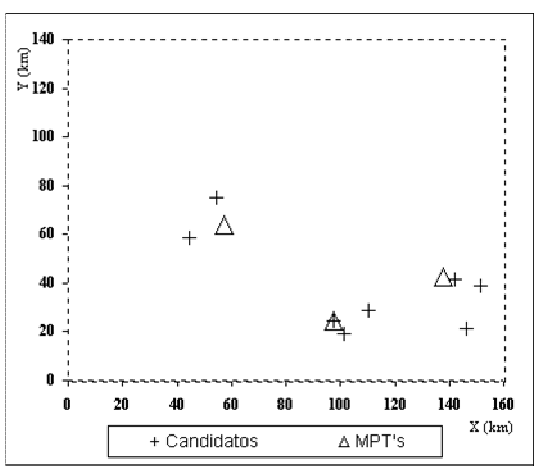

Figura 4.23b - Candidatos - 3 Zonas

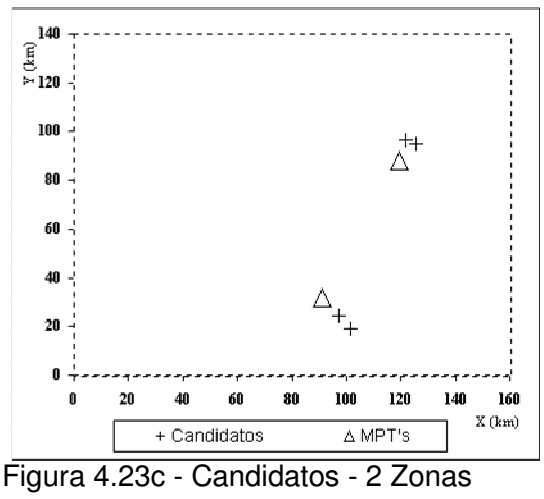

Figura 4.23c - Candidatos - 2 Zonas

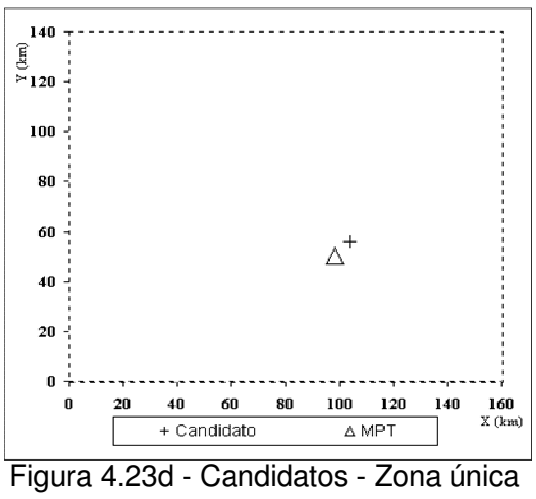

Figura 4.23d - Candidatos - Zona única

Visualmente verifica-se grande diferença nos casos de três e dois CD's. E em termos de custos da operação as zonas variáveis não apresentaram bons resultados. Em todas as situações os custos da solução ótima foram mais elevados do que a solução com zonas fixas, exceto no caso de zona única em que o resultado foi o mesmo.

O que ocorre é que, com zonas fixas gerando-se MPT's e MPT's-Meio, o processo proporciona um conjunto de candidatos circundando toda a região de estudo. Assim, estes candidatos acabam fornecendo melhores resultados que aqueles gerados por um conjunto de zonas variáveis.

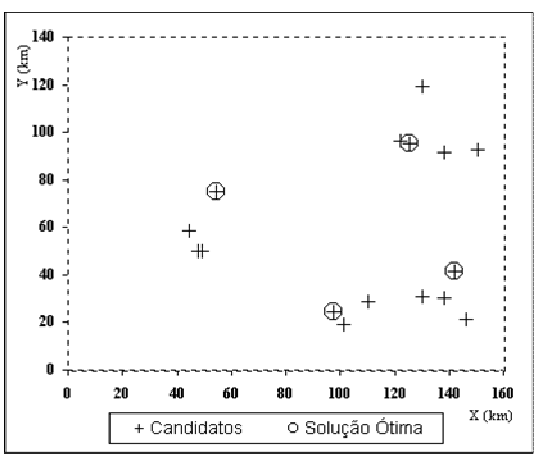

Figura 4.24a Solução - 4 Zonas e 4 CD's

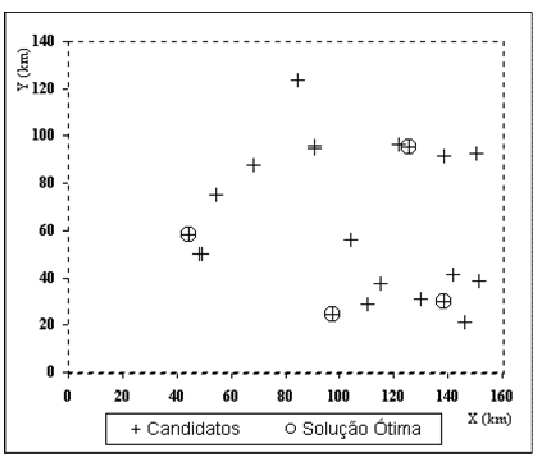

Figura 4.24b Solução - 5 Zonas Fixas e 4 CD's

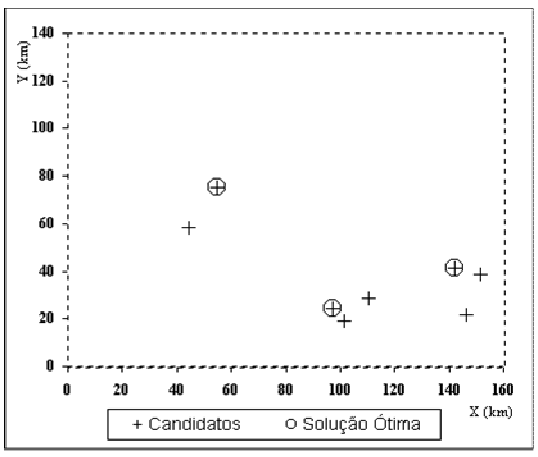

Figura 4.24c Solução - 3 Zonas e 3 CD's

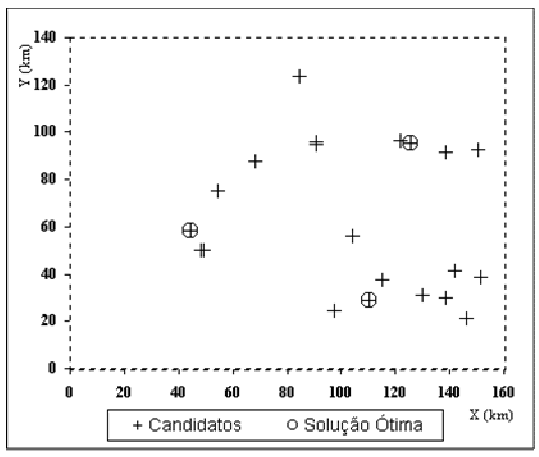

Figura 4.24d Solução - 5 Zonas Fixas e 3 CD's

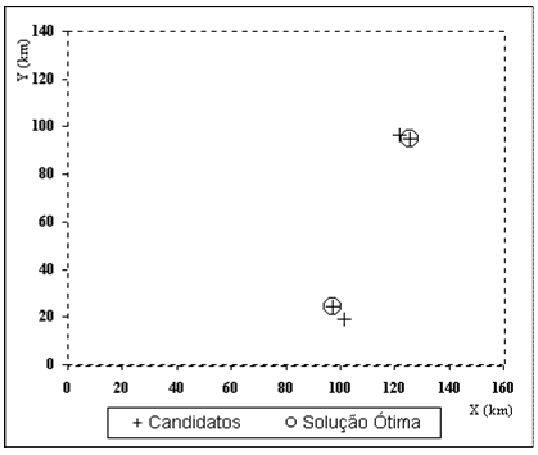

Figura 4.24e Solução - 2 Zonas e 2 CD's

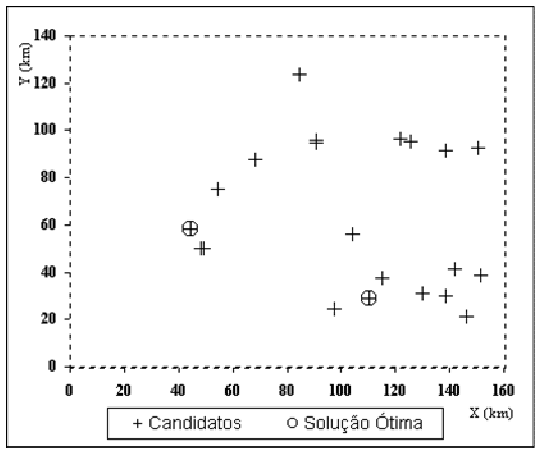

Figura 4.24f Solução - 5 Zonas Fixas e 2 CD's

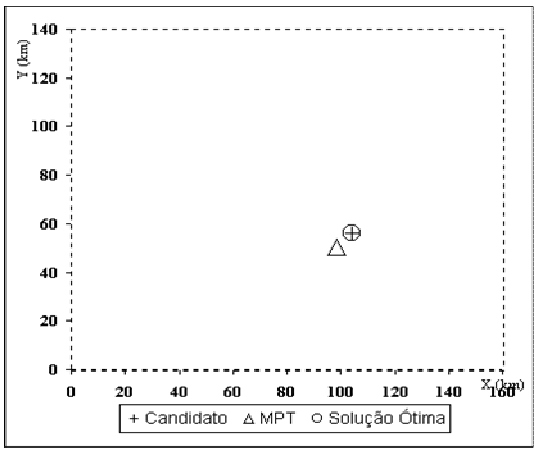

Figura 4.24g Solução - 1 Zona e 1 CD

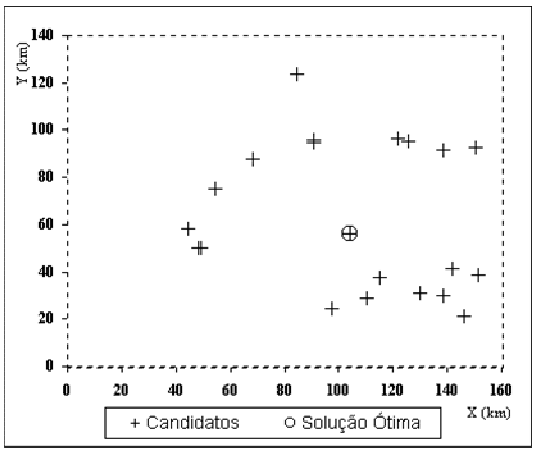

Figura 4.24h Solução - 5 Zonas Fixas e 1 CD

Já com zonas variáveis a configuração final irá depender do ponto de início do processo de zoneamento. Deve-se lembrar que para se gerar as zonas adota-se um determinado ponto de demanda para iniciar o processo de montagem dos *clusters*. Tendo-se n pontos de demanda pode-se ter n diferentes pontos de início do processo. Assim, dependendo de onde se inicia o processo, cada zona poderá ter um conjunto diferente de pontos e, conseqüentemente, uma MPT diferente.

Através das figuras pode-se visualizar claramente as diferenças entre os candidatos escolhidos como solução ótima através do processo de zonas variáveis e do de zonas fixas.

Note-se, entretanto, que as soluções ótimas de 4 CD's são muito próximas em zonas variáveis e fixas. O mesmo não ocorre com 3 CD's e 2 CD's, em que as soluções ótimas são bem diferentes.

Isto ocorre devido à proporção entre um eventual aumento do momento de transporte e a magnitude deste indicador em cada zona.

Note-se que quando se tem um número menor de zonas (duas ou três) tem-se um maior número de pontos em cada zona. Com muitos pontos a soma das distâncias percorridas tende a ser mais alta e, assim, se há uma distância muito grande entre um par de pontos esta distância não terá um peso tão significativo no aumento do momento de transporte, quanto o que teria quando o número de zonas é maior e o número de pontos em cada zona é menor.

Assim, o princípio de se iniciar a geração de zonas no segundo ponto do par que apresentar a maior distância continua a ser um critério válido, mas a sua eficácia é maior quando o número de zonas cresce, que é o que ocorre com zonas fixas, pois neste caso pontos distantes realmente não devem pertencer à mesma zona, o que proporcionalmente elevaria muito mais o momento de transporte da zona. Já com

um número pequeno de zonas a elevação do momento de transporte continua a ocorrer, mas proporcionalmente pode ser bem menor, reduzindo a eficácia do critério.

Verifica-se, assim, que no caso de zonas fixas o critério adotado para escolha do ponto de início da geração das zonas tem uma eficácia maior, pois neste caso o número de zonas não deverá ser muito pequeno.

Um procedimento que poderia ser adotado para minimizar este efeito seria o de iniciar o processo de zoneamento em diferentes posições ao longo de um giro de 360°, gerando-se, assim, diferentes conjuntos de zonas e suas respectivas MPT's. Ter-se-ia, desta forma, MPT's ao longo de todo o giro de 360°.

Este processo poderia inclusive ser utilizado para zonas fixas ou variáveis, pois em ambos os casos se caracterizaria como uma alternativa para buscar melhores soluções.

Um processo assim, entretanto, é bastante trabalhoso e pode se tornar muito lento se todos os pontos forem testados como pontos de início. Uma alternativa para esta questão será de certa fora tratada no próximo capítulo, com o uso de uma meta-heurística.

5. USO DE META-HEURÍSTICA NO PROBLEMA DA LOCALIZAÇÃO

A questão colocada ao final do capítulo anterior - variar o ponto de início de geração das zonas – tem como uma alternativa viável de solução o uso de uma meta-heurística para procurar melhorar a solução inicial. Esta proposta está na mesma linha de alguns trabalhos recentes, como Zhou et al. (2002), que desenvolve com o apoio de algoritmos genéticos um processo de geração de zonas de atendimento de centros de distribuição. O modelo busca o balanceamento de custo de transporte e nível de serviço e segundo os autores os resultados foram satisfatórios.

Novaes et al. (2000) desenvolve um outro trabalho nessa linha, procurando montar distritos em um sistema de distribuição física que busca maximizar a homogeneidade entre distritos. Esta homogeneidade é definida em termos de esforço de distribuição, que por sua vez é baseado no número de veículos necessários em cada zona. Os autores desenvolvem uma heurística para geração dos distritos e a partir desta solução inicial aplicam um algoritmo genético associado a um método de otimização direta, o chamado método do gradiente (Novaes, 1978), para maximizar a homogeneidade entre os distritos.

Um outro trabalho que também trata do assunto é o de D'Amico et al. (2002). Este artigo apresenta o problema de redesenhar jurisdições policiais na cidade de Buffalo, New York. O trabalho é desenvolvido com a meta-heurística *simulated annealing,* que a partir de uma dada solução inicial busca novas soluções que minimizem a disparidade entre as cargas de trabalho dos policiais.

Nos três casos, portanto, faz-se uso de uma meta-heurística que busca a otimização de uma função objetivo que representa algum tipo de medida de homogeneização das zonas geradas.

É neste sentido que se procurará caminhar nas próximas seções. O processo de modelagem das zonas deverá avançar passando a incorporar o uso de uma meta-heurística para melhorar a solução inicial encontrada.

5.1 META-HEURÍSTICA *SIMULATED ANNEALING*

A opção deste trabalho foi pelo uso de *simulated annealing,* pois analisando a literatura verificou-se que esta técnica tem apresentado bons resultados em dois tipos de problemas relacionados aos assuntos desta pesquisa. O primeiro é o chamado *graph partioning problem* e para o qual vários autores têm utilizado *simulated annealing* (Johnson et al., 1989; Pirlot, 1996; D'Amico, 2002). Este problema é na verdade semelhante à questão do zoneamento aqui estudada. O segundo problema refere-se à alocação de pontos a um *hub*, o chamado *p-hub problem* para o qual Abdinnour-Helm (2001) desenvolveu uma aplicação de *simulated annealing* e afirma que esta é uma abordagem indicada devido aos bons resultados que apresenta em problemas de alocação em geral. Este é semelhante à questão final de definição dos CD's e de suas zonas de atendimento.

A *simulated annealing* faz parte de uma categoria de técnicas consideradas como *meta-heurísticas* ou *heurísticas gerais* (Pirlot, 1996). Esta denominação é dada, pois estes métodos se constituem, na verdade, em *"princípios gerais de busca organizados em estratégias gerais",* não podendo assim, ser considerados como algoritmos, mas sim realmente como métodos ou então como meta-algoritmos (Pirlot, 1996).

Segundo Viana (1998) a técnica de *simulated annealing,* em particular, deriva de uma analogia com um processo físico cujo termo pode ser traduzido por "têmpera

simulada". Aqui neste trabalho a sigla SA será adotada para representá-la. A têmpera é o processo de aquecer certos materiais a altas temperaturas e em seguida ir reduzindo gradualmente esta temperatura até que o material atinja um equilíbrio térmico, tornando-o consistente e rígido. SA, assim, corresponderia segundo Viana, a uma simulação algorítmica do processo físico de têmpera.

O que ocorre na têmpera é que a altas temperaturas as partículas do material estão sob alta energia e se arranjam de forma totalmente aleatória. À medida que lentamente a temperatura vai baixando, começa a ocorrer uma organização das partículas no sentido de padrões regulares, até que estas se estabilizem em uma posição que dá ao material propriedades de maior resistência. (Johnson et al.,1989; Viana,1998; Kado, 1998).

A abordagem de SA é aplicada a problemas de otimização combinatória em que uma função objetivo é equivalente ao nível de energia do sistema físico.

Cada configuração das partículas do material pode ser comparada a uma solução viável de um problema de otimização. A melhor configuração das partículas seria aquela de energia mínima, correspondendo ao valor mínimo da função objetivo em um problema de otimização (Kado,1998).

A redução de temperatura deve ser lenta, pois para se atingir o equilíbrio térmico a uma dada temperatura é preciso que esta permaneça estável por um período suficientemente longo. E o mesmo se aplica aos problemas de otimização. Assim, um conjunto grande de "perturbações" deve ser imposto ao sistema a cada nível de temperatura.

Neste ponto são necessárias definições de alguns termos que serão adotados ao longo do texto.

- Instância de um problema refere-se a um conjunto de valores especificados aos parâmetros de entrada em um determinado instante (Viana, 1998). Assim, uma rede com três CD's e quatro pontos de demanda e mais os custos de transporte em cada aresta que ligue estes vértices seria uma instância de um problema.
- Configuração equivale a um dos arranjos possíveis dos componentes do sistema em estudo. Os fluxos definidos entre CD's e destinos, por exemplo.

  A cada configuração corresponde um valor da função objetivo, assim, uma configuração é equivalente ao estado de um sistema no sentido estocástico em que, segundo Kleinrock (1975), o estado de uma variável X(t) corresponde ao valor de X no tempo t. Kleinrock, inclusive, faz uma associação com o movimento de uma partícula no espaço. A cada posição da partícula corresponderia um estado do sistema.
- Perturbação é um procedimento que transforma uma configuração E_i em uma configuração E_{i+1}. Na prática, E_{i+1} é uma configuração "ligeiramente" modificada de E_i (Viana, 1998). De modo similar, a uma pequena mudança de posicionamento de um átomo de um material (Kado, 1998).
- Vizinhança é o conjunto de configurações que podem ser geradas a partir de uma dada configuração.

No algoritmo da SA parte-se de uma solução inicial e entra-se em um processo iterativo de introdução de pequenas perturbações na configuração do sistema. Feita uma perturbação na configuração do sistema, computa-se o valor da função objetivo. No caso de minimização, se o valor for menor, aceita-se a nova configuração. Esta solução melhor é chamada de solução superior. Se a função objetivo aumentar de valor, gerando uma solução inferior, a nova configuração

também poderá ser aceita, mas com uma probabilidade que é determinada pela diferença entre os dois valores da função objetivo e pelo parâmetro denominado "temperatura" do sistema.

Um ponto crucial do método é o processo de geração de novas configurações. A SA exige um processo sistemático para se escolher a nova configuração a cada iteração. Dentro desta sistematização, o processo deve incluir uma forma aleatória de geração da nova configuração e, além disto, a perturbação deve se constituir em uma pequena alteração, de preferência interferindo em um único elemento do sistema (um "átomo" na analogia com o processo físico). Com isto, o comportamento da SA pode ser modelado como uma sucessão de cadeias de Markov, uma para cada valor de temperatura (Colorni et al., 1997), em que cada mudança de configuração corresponde a um estado do sistema.

Ao final do processo as soluções devem convergir para um único valor, a exemplo do que ocorre em uma cadeia de Markov, em que a longo prazo as probabilidades de mudança de um estado para outro convergem para uma probabilidade limite (Ross, 1972; Kleinrock, 1975; Novaes, 1975).

A probabilidade de aceitação de uma mudança é chamada de "critério de Metropolis" (Viana, 1998) e é calculada a partir do "Fator de Boltzman (B)", que é dado pela função:

$$B = \mathbf{exp(-\Delta E/T)} \qquad (5.1)$$

onde: T = temperatura atual

$$\Delta E = v(s) - v(s_o) \qquad (5.2)$$

sendo :

$v(s)$ = valor da função objetivo da configuração em teste

$v(s_o)$ = valor da função objetivo da configuração atual

Em SA o valor de B é comparado a um número n_o entre 0 e 1 gerado aleatoriamente. Se $B>n_o$ então a nova configuração será aceita.

Inicialmente a temperatura T é colocada em um valor bem alto. Assim, a mudança de s para s_o ocorre com freqüência. Conforme T vai diminuindo, B também diminui e as aceitações de soluções inferiores passam a ser mais raras. Quando T está muito baixa e por muitas iterações não ocorrem mudanças, considera-se que a solução está "congelada" e o processo é encerrado.

A implementação deste processo envolve a definição de alguns parâmetros:

- Temperatura inicial (T_o);
- Número máximo de iterações em cada temperatura (p);
- Número máximo de sucessos em cada temperatura (L);
- Número total de iterações (m);
- Método de resfriamento;

A definição desses parâmetros é ponto chave para o sucesso do método. E neste sentido alguns procedimentos precisam ser estabelecidos de forma a se determinar o valor de cada um.

No que diz respeito à temperatura inicial (T_o), esta pode ser obtida através de experimentos, como mostra D'Amico et al. (2002), em que o valor da função objetivo permaneceu aproximadamente constante em temperaturas acima de 0,1. Abaixo desta temperatura o valor da função objetivo decrescia até que por volta de 0,017 voltava a crescer. Os autores concluíram, então, que estas deveriam ser as temperaturas iniciais e finais.

Uma forma, entretanto, mais direta é proposta por Johnson, et al. (*in* Viana,1998) em que T_o é definida através de:

$$T_o = -\Delta E_m^+/\ln(\chi_o) \qquad (5.3)$$

onde: ΔE_m^+ = media aritmética de incrementos na função objetivo;

χ_o = valor empírico em torno de 0,8.

O número máximo (p) de iterações em cada temperatura pode ser entendido como sendo o comprimento da cadeia de Markov e para este parâmetro uma alternativa de sucesso é apresentada por Viana (1998), que mostra que p deve ser proporcional ao número (n) de variáveis no sistema e que bons resultados são obtidos fazendo-se p = 100n. O uso de L=0,1p traz também bons resultados.

Quanto ao método de resfriamento, autores sugerem que o esquema de decréscimo seja logarítmico e que, assim, quase com certeza o processo irá convergir para uma solução ótima (Johnson et al.,1989). Neste caso, a equação proposta seria:

$$T_i = C/(1+\log_i) \qquad (5.4)$$

onde: Ti = temperatura na iteração i

C = um valor constante.

Este processo, entretanto, leva um longo tempo e os resultados práticos mostram que se aplicando um resfriamento geométrico os resultados são suficientemente bons (Johnson, 1989; Viana, 1998). O método de resfriamento geométrico seguiria a seguinte equação:

$$T_i = Rc^i.T_o \qquad (5.5)$$

Deve-se notar que com o uso deste método, uma vez definidas as temperaturas inicial e final e a taxa de resfriamento, é possível estimar o número de iterações através de:

$$m = \log(T_m/T_o) / \log Rc \qquad (5.6)$$

Nos experimentos desenvolvidos nesta pesquisa foi constatado que estas equações constituem um bom referencial para iniciar a definição dos parâmetros, mas que pequenos ajustes são necessários, pois em alguns casos os parâmetros são superestimados aumentando desnecessariamente o tempo de processamento. Na verdade, esta necessidade de ajustes não ocorreu somente nesta pesquisa, tendo sido relatada a mesma questão nos experimentos desenvolvidos por Johnson et al. (1991).

Em relação à configuração inicial do sistema, a SA, apesar de ser considerada uma meta-heurística (Johnson et al., 1989; Viana, 1998; Pirlot, 1996), não requer qualquer preparação anterior, podendo em princípio iniciar de qualquer ponto. Johnson et al. (1989) ressaltam, entretanto, que iniciar o processo a partir de uma "boa" solução pode ser nas palavras dos autores de "significância crucial" para obtenção de uma solução de melhor qualidade ou redução dos tempos de processamento, ou ambos. Daí uma razão forte para no caso da presente pesquisa se ter uma heurística de *clusterização* que gere inicialmente esta "boa" solução.

5.2 USO DE SA NA MONTAGEM DE ZONAS

5.2 1. Descrição do Modelo de Zoneamento

Como já colocado anteriormente, o processo de SA exige uma sistemática absolutamente definida para se escolher uma outra solução a cada iteração. Para a montagem de zonas este processo na verdade se constitui em um rearranjo de zonas e é baseado no conceito de solução vizinha.

No presente caso, uma solução vizinha de uma solução em particular é definida como sendo aquela mesma configuração de zonas, exceto por um único ponto que é re-alocado para uma zona adjacente àquela à qual pertence.

O conceito de zona adjacente é estabelecido a partir do próprio processo de montagem de zonas (vide capítulo 4). Lembrando que a montagem das zonas dentro de um determinado anel inicia-se com uma ordenação crescente, segundo o ângulo com o eixo horizontal, de todos os pontos de demanda daquele anel. As zonas vão sendo montadas varrendo-se todos estes pontos no sentido anti-horário. As zonas adjacentes correspondem, assim, à zona anterior e posterior neste processo de varredura.

Esta definição de vizinho tem como resultado que a cada iteração um ponto sorteado na fronteira entre duas zonas adjacentes estará mudando de uma zona para outra se a nova solução for aceita.

O processo de aceitação de uma outra solução tem por base uma função objetivo que representa o *grau de homogeneidade* (GH) das zonas. Se este GH aumentar, a nova solução é aceita; se o GH diminuir, a nova solução será aceita desde que P(Boltzman) - a probabilidade de Boltzman – seja maior ou igual a um número aleatório gerado entre 0 e 1.

O grau de homogeneidade é definido através de:

$$GH = \begin{cases} \left(1 - \dfrac{s_{TKM}}{TKM_m}\right).100, & se \;\; s_{TKM} \leq TKM_m \\ 0, & em\ caso\ contrário \end{cases} \tag{5.7}$$

onde:

s_{TKM} = desvio padrão dos momentos de transporte das zonas

TKM_m = média do momento de transporte das zonas

O modelo para a aplicação deste processo tem uma estrutura montada em fases, conforme apresentadas a seguir:

Fase 1. Lê dados de entrada

Identificação de pontos e respectivos anéis

Identificação de zonas e respectivos pontos

Identificação de parâmetros de controle da SA

Fase 2: Preparação do processo:

a) Se todos anéis já foram selecionados encerrar a SA.

Em caso contrário, selecionar o próximo anel

b) Montar seqüência de varredura de pontos dentro do anel

Fase 3: Desenvolve processo:

a) Se total de Iterações atingiu limite, retornar para Fase 2

b) Faz rearranjo de zonas

b1) Identificar os pontos limítrofes de cada zona

b2) Sortear um destes pontos

b21) Se ponto estiver no início da seqüência de varredura da zona, transferir o ponto para zona anterior

b22) Se ponto estiver no final da seqüência de varredura da zona, transferir o ponto para zona seguinte

c) Verificar se nova configuração atende restrições

c1) se atender, seguir para próximo passo

c2) se não atender, retornar para passo (b)

d) Recalcular GH

c1) Se GH aumentou ou se manteve, aceitar a nova solução

c2) Se GH se reduziu, aceitar a nova solução se P(Boltzman)>n_o, sendo n_o um número aleatório gerado entre 0 e 1

e) Verificar se iterações na temperatura ou número de sucessos permitidos na temperatura atingiram limite

e1) se um deles chegou ao limite permitido, reduzir temperatura e retornar ao passo (a)

e2) se nenhum chegou ao limite permitido, retornar ao passo (b)

Para a implementação deste modelo foi desenvolvido um programa de computador que contempla todo o desenvolvimento do processo, cujos parâmetros, operação e resultados são descritos nas seções seguintes.

5.2.2 Parâmetros da SA e Dados de Entrada - Zoneamento

A experimentação com SA foi feita com os mesmos dados do experimento desenvolvido na validação das heurísticas (seção 4.5)

O sistema desenvolvido recebeu informações referentes às zonas atribuídas a cada ponto, às coordenadas e às demandas dos pontos. Além disto, são fornecidos os ângulos de cada ponto com o eixo horizontal, de forma a possibilitar o processo de varredura descrito acima.

Como restrição estabeleceu-se um mínimo de pontos que deveria ser mantido em cada zona, evitando-se assim a ocorrência de zonas muito pequenas. Foram feitos testes com limites mínimos de três, quatro e cinco pontos, de forma a se verificar a influência deste parâmetro nos resultados.

O número de zonas seguiu o padrão já definido para atender aos requisitos de homogeneidade entre os anéis, porte dos CD's e nível de serviço desejado. Este

padrão já havia sido estabelecido em um total de quatro zonas no anel 2 e uma única zona no anel 1.

Em termo de parâmetros da SA, foram definidos aqueles discutidos na seção 5.1, quais sejam:.

Temperatura inicial (T_o)

Número máximo de iterações em cada temperatura (p)

Número máximo de sucessos em cada temperatura (L);

Número total de iterações (m);

Método de resfriamento.

A determinação destes parâmetros baseou-se nas diretrizes também discutidas na mesma seção 5.1.

Os padrões adotados para cada um dos parâmetros são apresentados a seguir.

a) Temperatura inicial (T_o)

Esta foi definida a partir da equação 5.3, utilizando-se valor empírico de 0,8 para χ_o, conforme sugerido por Viana (1998).

Assim, obtém-se a equação final :

$$T_o = 4{,}48.\ \Delta E_m^+ \qquad (5.8)$$

O valor de ΔE_m^+ é obtido através de resultados experimentais, conforme indicado por Viana (1998) e Johnson (1989). A figura 5.1 apresenta uma tela de saída do software que foi desenvolvido para a SA, mostrando o cálculo final de ΔE_m^+ no experimento desenvolvido.

Com $\Delta E_m^+ = 16{,}74$ tem-se $T_o = 74{,}99$. Foi adotado então $T_o = 75{,}0$.

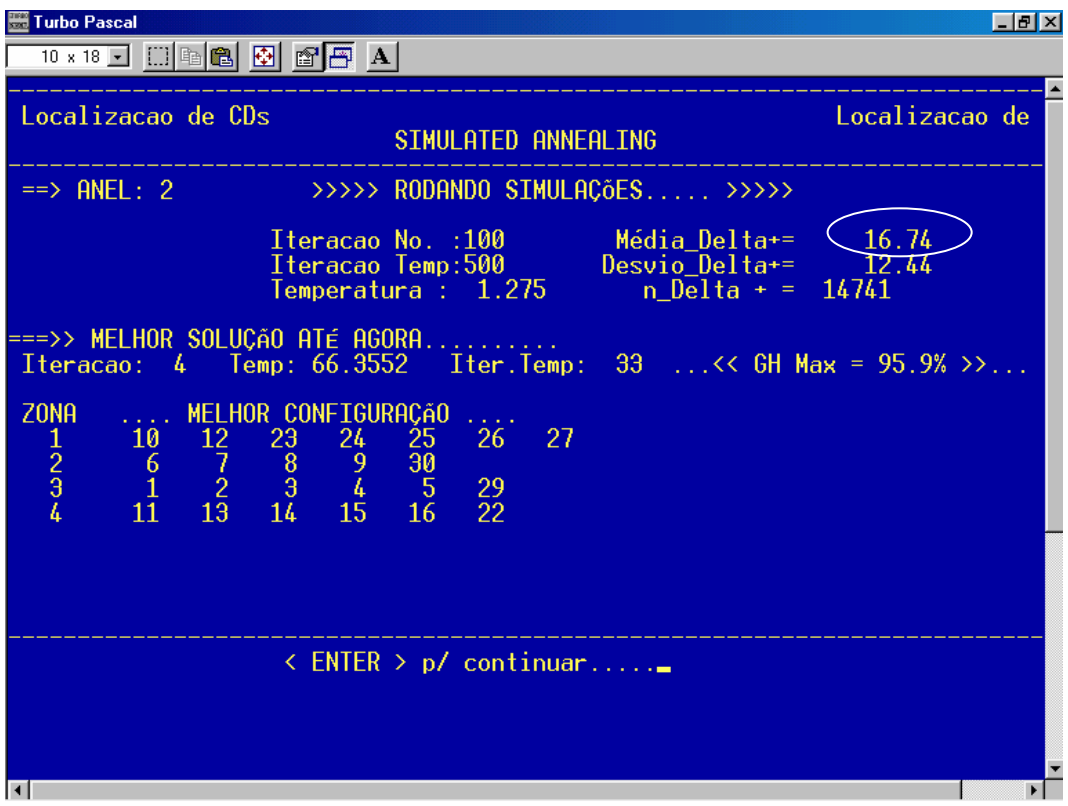

Figura 5.1 - SA:Tela de resultados apresentando ΔE_m^+

b) Número máximo de iterações em cada temperatura (p)

Considerando-se um número definido de quatro zonas no anel 2, que é o maior dos dois anéis, tem-se, para uma dada configuração, oito pontos limítrofes. É com base nestes pontos que se define o número de variáveis envolvidas no processo e a própria vizinhança de uma dada configuração de zonas.

Conforme a diretriz apresentada na seção 5.1, ter-se-ia então p = 100.8 = 800 iterações a cada temperatura (comprimento da Cadeia de Markov).

Os resultados experimentais mostraram que este número estava superdimensionado. Na verdade, com um número menor de iterações as soluções já estavam convergindo. Foi adotado então o número de 500 iterações a cada temperatura, que se mostrou plenamente satisfatório.

c) Número máximo de sucessos em cada temperatura (L)

Por decorrência da definição de p=500 iterações, e adotando-se L igual a 10% de p (seção 5.1), tem-se: L=50 iterações.

d) Método de resfriamento;

Conforme as observações de Johnson(1989) e Viana(1998), que sugerem o uso de uma taxa geométrica de resfriamento variando entre 0,90 e 0,99, foram desenvolvidos alguns testes para verificar a mais adequada para o presente estudo, chegando-se ao valor de Rc =0,96 como sendo aquele que apresentou o melhor grau de convergência da solução dentro de um número de iterações compatível com os valores definidos para o estudo.

d) Número total de iterações (m);

Resultados experimentais mostraram que o valor máximo da função objetivo definida (GH - grau de homogeneidade) apresentava grande oscilação até uma temperatura em torno de 10,0, quando então iniciava-se uma tendência de estabilização. A convergência final da solução, porém, efetivamente ocorreu para temperaturas abaixo de 3,0.

Isto pode ser observado pela figura 5.2a, que apresenta a variação de GH máximo à medida que a temperatura vai decrescendo. Com este gráfico, entretanto, não foi possível uma visualização completa do processo, pois, em temperaturas mais baixas, a diferença absoluta entre duas temperaturas seguidas é muito pequena, o que leva a um agrupamento dos pontos nestas faixas de temperatura devido ao efeito da escala. É importante lembrar que conforme a taxa de resfriamento adotada (Rc=0,96), na iteração i+1 (i=1,2,...m) a temperatura corresponderá a 96% da temperatura na iteração i (T_{i+1}= 0,96.T_i; i=1,2, m). Quando as temperaturas são

baixas, portanto, a diferença absoluta entre duas temperaturas seguidas passa a ser muito pequena.

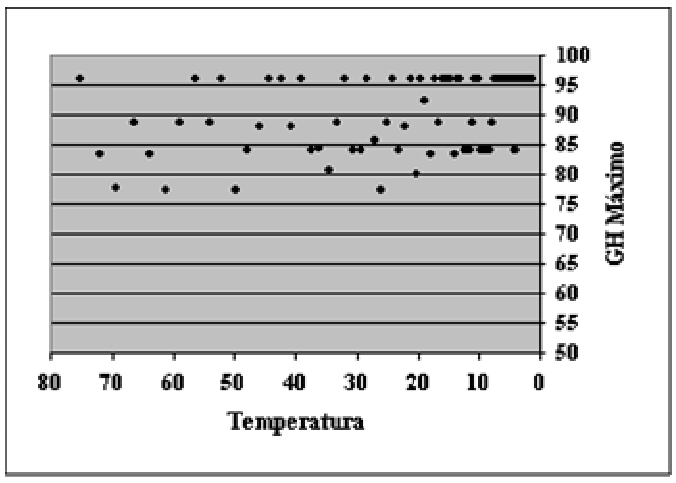

Figura 5.2a - Variação de GH Máximo com o decréscimo de temperatura

Um detalhamento da fase final deste processo para permitir uma melhor visualização é apresentado então na figura 5.2b, em que se percebe com clareza a convergência da solução quando são apresentadas apenas as temperaturas abaixo de 10,0.

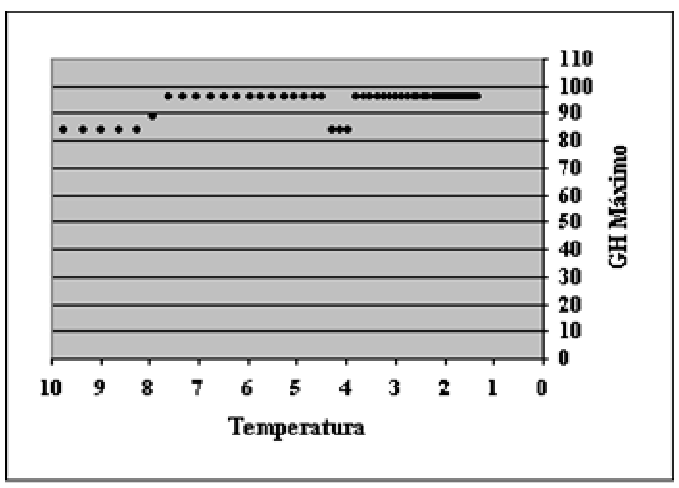

Figura 5.2b - Detalhamento da variação de GH Máximo para temperaturas abaixo de 10,0

Assim, dentro de uma abordagem mais conservadora, foi adotado um valor mínimo de temperatura (T_m) de cerca de 1,0, o que garantiria com segurança a convergência da solução. Como o modelo estava sendo executado em tempos de processamento bastante baixos, a atitude conservadora não gerou lentidão no processo.

Com o valor de T_m definido e utilizando-se a equação 5.6 foi possível determinar o número m máximo de iterações:

$$m = \log(1{,}0 / 75{,}0) / \log(0{,}96) \rightarrow \mathbf{m = 105{,}76}$$

Foi adotado então o número máximo de 100 iterações, resultando um valor final de temperatura um pouco superior a 1,0. Obteve-se $T_m = 1{,}31799$.

5.2.3. Resultados Experimentais do Zoneamento

O processo foi todo desenvolvido com o apoio de um programa de computador construído especificamente para esta pesquisa. Este programa apresenta ao término da simulação os resultados finais de cada anel com a melhor configuração gerada (figuras 5.3a e 5.3b).

Na seqüência o sistema apresenta índices que monitoram a evolução do processo a cada temperatura.

Os índices calculados a cada temperatura i (i=1,2,....m) são:

. GH Médio

. Desvio Padrão de GH

. Coeficiente de Variação de GH

. Taxa de Soluções Inferiores

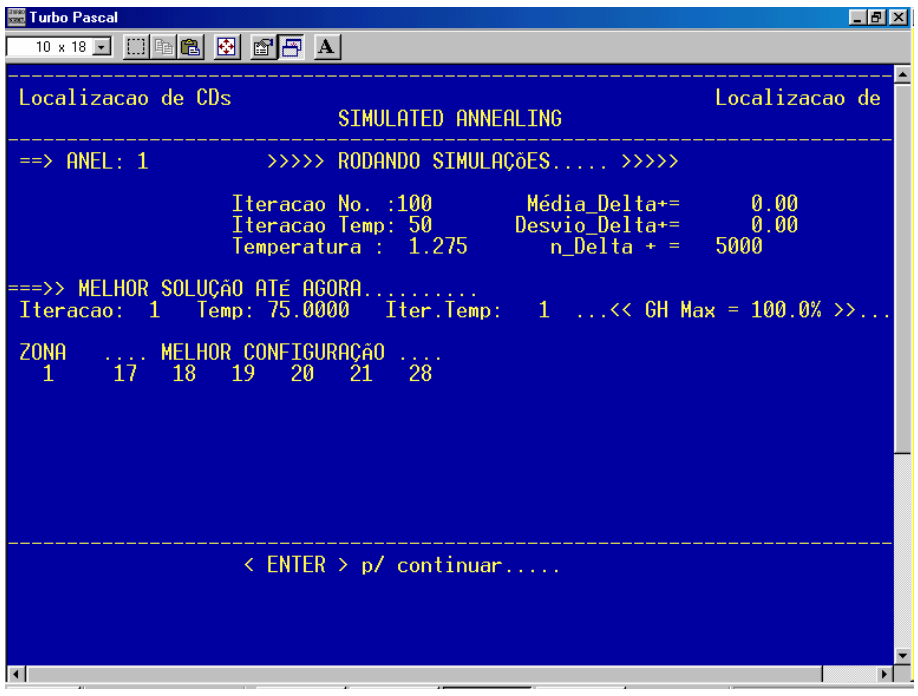

Figura 5.3a - Melhor configuração gerada para o anel 1

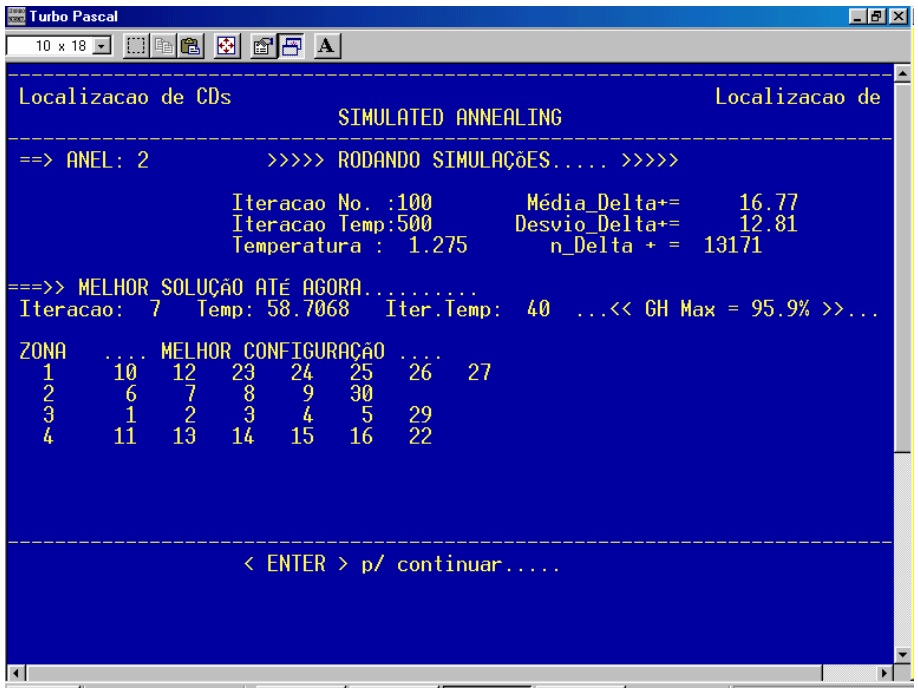

Figura 5.3b - Melhor configuração gerada para o anel 2

. Taxa de Aceitação de Soluções Inferiores

. GH Máximo obtido na temperatura i.

O software apresenta o valor da Temperatura, o Número de iterações naquela temperatura e os resultados de todos estes índices.

A figura 5.4a apresenta a saída do sistema com o resultado dos índices nas temperaturas iniciais (mais altas) e a figura 5.4b mostra os mesmos índices nas temperaturas finais (mais baixas), quando as soluções convergem para um mesmo GH Máximo.

Gráficos do processo de evolução do GH Médio e do GH Máximo a cada temperatura são apresentados nas figuras 5.5 e 5.6. Percebe-se por estes diagramas a eficácia do processo de SA. Na figura 5.5 o GH Médio cresce consistentemente de forma aproximadamente exponencial ao longo de todo o processo e o GH Máximo de cada temperatura converge de forma definitiva para um valor acima de todos os outros quando a temperatura cai para valores bem baixos.

Turbo Pascal

Temp	Iterac	GH	Desv	CV	Tx Inf	TxAcInf	GHmax
75.0000	54	66.8	12.7	19.0	57.4	87.1	86.6
72.0000	51	61.9	9.9	16.1	56.9	96.6	82.4
69.1200	54	72.4	6.7	9.3	51.9	85.7	84.0
66.3552	56	69.2	7.1	10.3	55.4	80.6	84.0
63.7010	51	63.6	7.6	11.9	51.0	96.2	81.9
61.1530	53	69.5	7.1	10.3	54.7	89.7	83.5
58.7068	55	69.6	10.5	15.1	54.5	83.3	95.9
56.3586	55	68.9	9.6	13.9	52.7	82.8	88.6
54.1042	55	66.1	10.9	16.5	49.1	81.5	95.9
51.9400	58	71.9	12.0	16.7	53.4	74.2	92.3
49.8624	58	71.6	11.9	16.6	56.9	75.8	95.9
47.8679	58	66.0	8.8	13.4	56.9	75.8	85.7
45.9532	50	67.2	12.7	18.9	50.0	100.0	84.0
44.1151	53	66.0	9.5	14.4	54.7	89.7	85.7
42.3505	53	56.1	11.3	20.1	54.7	89.7	80.3
40.6565	53	68.3	7.6	11.1	56.6	90.0	84.0
39.0302	50	66.4	11.5	17.3	50.0	100.0	88.0
37.4690	52	65.9	9.1	13.8	59.6	93.5	77.2
35.9703	56	63.7	10.9	17.1	55.4	80.6	84.4
34.5314	55	70.0	6.9	9.8	52.7	82.8	83.5

....< ENTER > p/ continuar...

Figura 5.4a - Resultados de índices de monitoramento do processo
Temperaturas iniciais (mais altas) – Anel 2

Turbo Pascal

Temp	Iterac	GH	Desv	CV	Tx Inf	TxAcInf	GHmax
2.8626	352	92.2	3.6	3.9	92.9	7.6	95.9
2.7481	304	91.2	4.1	4.5	92.4	9.6	95.9
2.6382	285	91.5	5.2	5.7	91.6	10.0	95.9
2.5326	427	89.7	4.5	5.0	93.7	5.8	95.9
2.4313	276	87.9	8.0	9.1	91.7	10.7	95.9
2.3341	263	84.7	9.2	10.9	90.1	10.1	95.9
2.2407	500	93.6	2.4	2.6	96.6	3.7	95.9
2.1511	500	91.8	3.6	4.0	96.2	4.0	95.9
2.0650	500	93.4	2.5	2.7	97.4	3.1	95.9
1.9824	500	93.7	2.3	2.4	96.6	3.5	95.9
1.9031	500	92.8	3.2	3.4	96.6	3.5	95.9
1.8270	500	92.9	2.6	2.8	97.2	3.1	95.9
1.7539	500	93.6	2.5	2.7	98.0	2.2	95.9
1.6838	500	93.3	2.7	2.9	98.4	1.2	95.9
1.6164	500	93.6	2.2	2.4	97.0	3.1	95.9
1.5518	500	92.9	3.1	3.4	97.8	2.5	95.9
1.4897	500	93.1	2.8	3.0	98.2	1.6	95.9
1.4301	500	92.2	3.2	3.5	97.8	3.1	95.9
1.3729	500	93.2	2.7	2.9	97.4	2.7	95.9
1.3180	500	93.4	2.6	2.8	98.2	1.4	95.9

....< ENTER > p/ continuar...

Figura 5.4b - Índices de monitoramento – Convergência de soluções
Temperaturas finais (mais baixas) - Anel 2

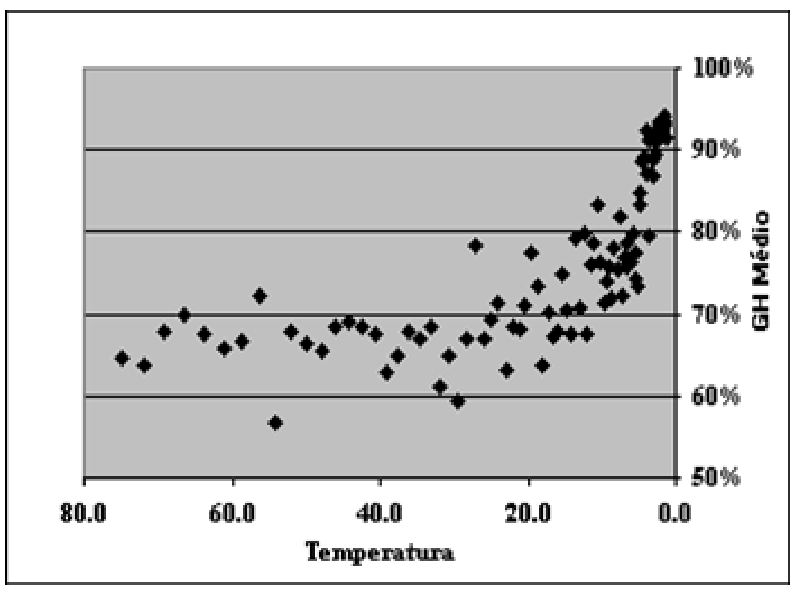

Figura 5.5 – Evolução de GH Médio com a redução de temperatura.

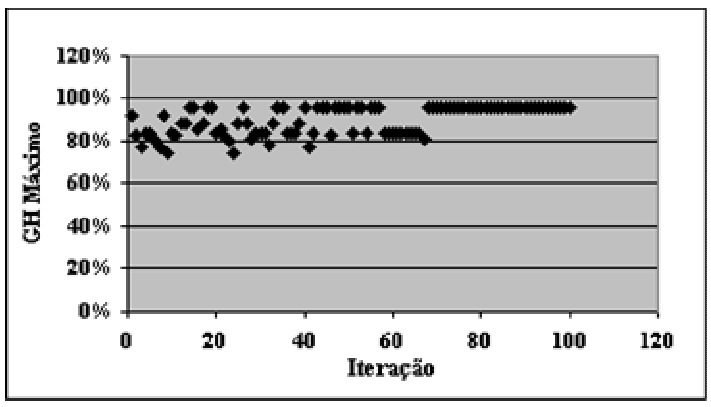

Figura 5.6 – Convergência da solução “ótima” .
Partindo de uma solução diferente da “ótima”.

A figura 5.7 mostra a distribuição de freqüências de GH ao longo de todo o processo de simulação e em seguida é apresentado seu histograma (figura 5.8). A montagem das freqüências foi feita com classes de amplitude de 10%.

A última classe, entretanto, foi subdividida para permitir um maior detalhamento dos GH’s mais elevados. O formato do histograma retrata o processo de simulação da SA que se desenvolveu de forma bem eficiente, abrangendo uma ampla gama de soluções e com um comportamento que lembra a distribuição normal, já que apresenta uma simetria bem razoável. A distribuição atinge uma gama variada de soluções, desde soluções bastante ruins, com GH’s abaixo de 50%, até soluções muito boas, em que os GH’s ultrapassam 95%.

A figura 5.9 apresenta a melhor configuração global, aquela que gerou um Grau de Homogeneidade de 96,1%. Este melhor resultado do experimento corresponde à mesma solução que foi encontrada diretamente pela Heurística de *Clusterização* apresentada no capítulo 4. A heurística proposta nesta pesquisa, portanto, se mostrou bem eficaz, gerando uma solução que não foi possível de ser melhorada através da meta-heurística.

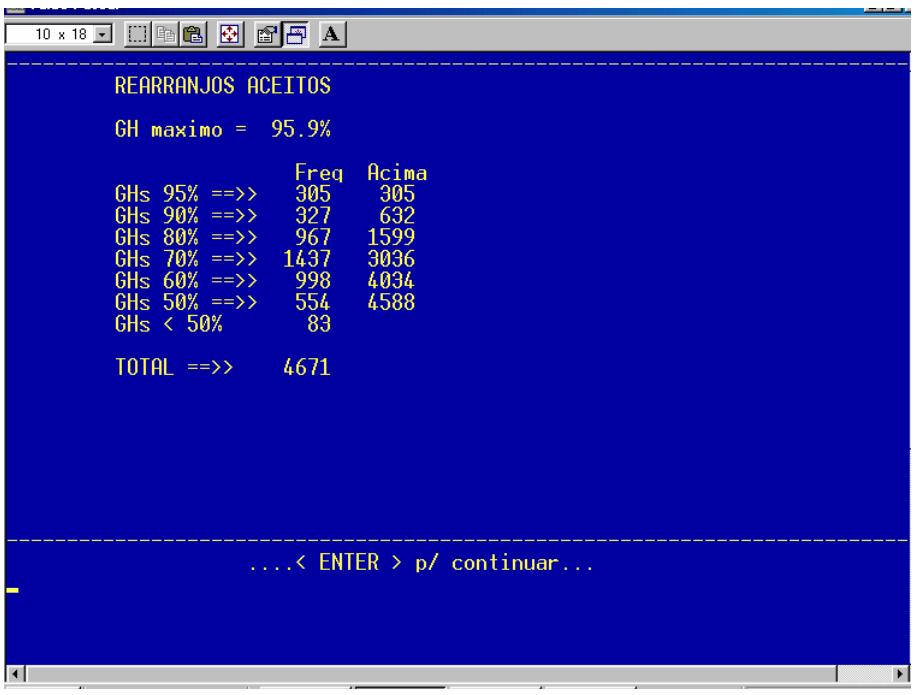

Figura 5.7 - Distribuição de GH ao longo da simulação – Anel 2

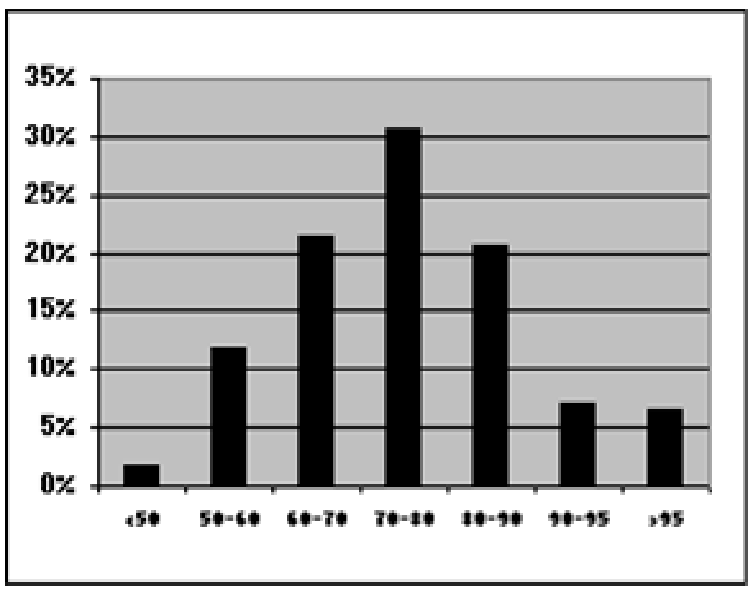

Figura 5.8 – Histograma de GH - Anel 2

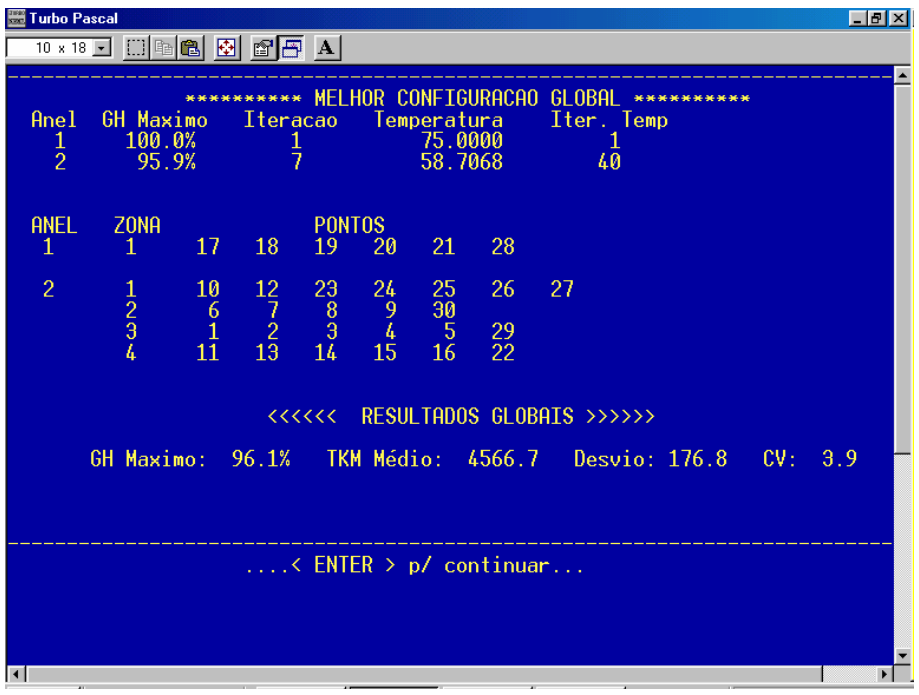

Figura 5.9 – Resultado final com “melhor” configuração global Anéis 1 e 2

Sobre as soluções geradas pela SA deve ser salientado que estas convergiram invariavelmente para a “melhor” solução. Esta “melhor” solução, entretanto, variou em função da restrição imposta sobre o número mínimo de pontos por zona. Foram feitos testes simulando um mínimo de 3, 4 e 5 pontos e um grau de 96,1% foi obtido com mínimos de 4 e 5 pontos, mas não foi possível atingir esta solução para o mínimo de 3 pontos. Análise de sensibilidade destes parâmetros é importante, portanto, que seja desenvolvida.

Outro aspecto importante refere-se ao fato de que para diferentes configurações iniciais o processo sempre convergiu para a “melhor” solução com GH=96,1%. A figura 5.6 mostra a convergência quando a configuração inicial é distinta da “melhor” configuração; já na figura 5.10 utilizou-se como configuração inicial a própria solução “ótima” e verifica-se que o processo leva ao final a um retorno a esta mesma solução.

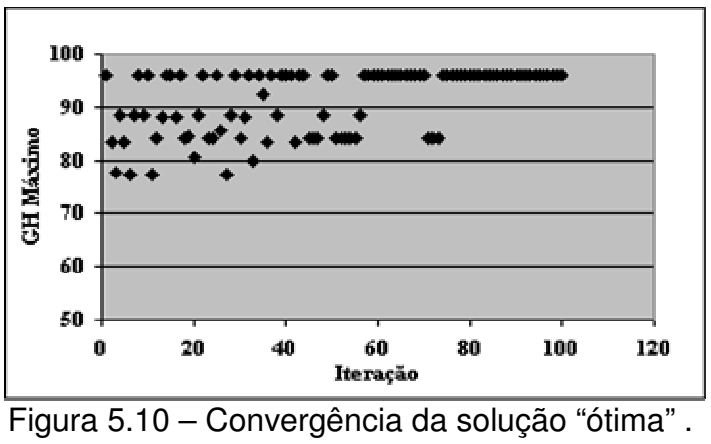

Figura 5.10 – Convergência da solução “ótima” .
Inicia na solução “ótima” e retorna à “ótima”

5.3. USO DE SA NA LOCALIZAÇÃO FINAL DOS CD’s

Os bons resultados obtidos com a aplicação das heurísticas e da meta-heurística utilizadas nesta pesquisa sugeriram a possibilidade de se partir diretamente para a solução final do problema de localização, prescindindo-se assim do modelo MILP. Buscar-se-ia identificar os melhores locais para um dado número de CD’s, nos moldes de um problema de p-medianas, através do desenvolvimento de um novo processo heurístico que conduzisse a esta solução.

Os resultados que se tem até agora com o processo heurístico já desenvolvido se sintetizam em um conjunto de candidatos que se situam no entorno de uma MPT em cada uma das zonas definidas. Esses resultados correspondem à etapa imediatamente anterior ao uso do modelo MILP. A nova heurística deve partir então destes candidatos e desenvolver um processo que conduza a configuração “ótima” de locais a instalar os CD´s. Pretende-se, portanto, propor um processo heurístico que possa substituir o modelo de otimização apresentado no capítulo 4.

Como até aqui se vem trabalhando com SA, a opção foi por se manter esta mesma técnica como base para este processo.

O problema agora, porém, tem características mais complexas do que o zoneamento como pode ser visualizado pela figura 5.11.

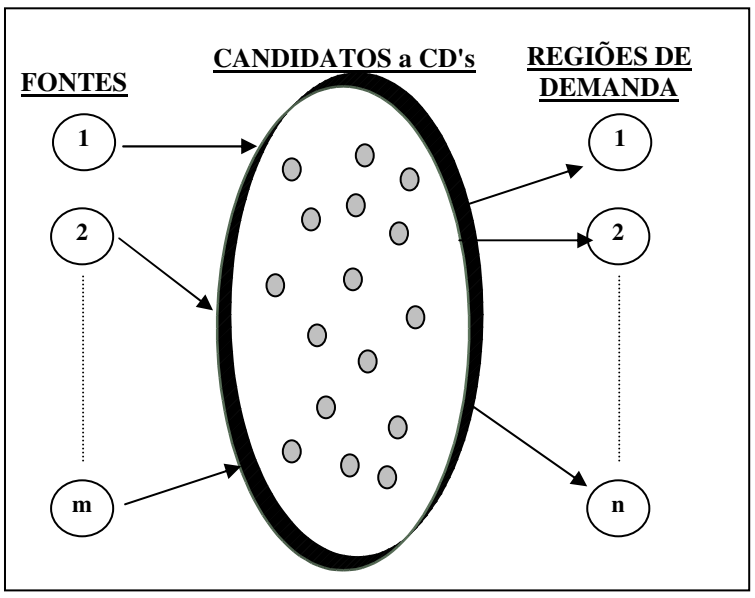

Figura 5.11 – Localizações “ótimas” de CD´s - SA

A descrição da metodologia com a qual pretende-se tratar esta questão mais complexa é apresentada nas próximas seções.

5.3.1. Descrição do Modelo de Localização

Como já colocado anteriormente, o processo de SA exige uma sistemática claramente definida para se escolher uma outra solução a cada iteração, envolvendo a definição de uma função objetivo e de um rearranjo da configuração do sistema.

No que tange à função objetivo, foi definida uma função de custo nos mesmos moldes da equação 4.17, que foi utilizada no modelo MILP. A função Z definida engloba custos de transporte e de localização conforme abaixo:

$$Mín \sum_{i,k} c_{ik}.x_{ik} + \sum_{k,j} c_{kj} D_j.y_{kj} + \sum_{k} f_k.Y_k \qquad (5.9)$$

onde:

c_{ik} = custo unitário de transferência entre a região i de produção e o CD k;

x_{ik} = fluxo do produto entre a região i de produção e o CD k;

c_{kj} = custo unitário de distribuição entre o CD k e a zona j de demanda;

D_j = demanda da zona j;

y_{kj} = variável binária de alocação $y_{kj} = 0$ ...CD k não atende zona j;

$y_{kj} = 1$ ...CD k atende zona j;

f_k = custo fixo de localização do CD k;

Y_k = variável binária de decisão $Y_k = 0$ ...CD k é ativo;

$Y_k = 1$ ...CD k não é ativo.

Em relação ao rearranjo, o processo se constitui em uma troca de alocações de CD's a regiões de demanda e é baseado no conceito de solução vizinha.

No presente caso, uma solução vizinha de uma solução em particular é definida como sendo aquela mesma configuração de alocações de CD´s a zonas de demanda, exceto por uma única zona cujo atendimento é re-alocado para um outro CD.

A vizinhança incorpora ainda uma troca de local candidato a CD. Assim, um vizinho será também um novo conjunto de CD's em que um único CD é trocado por um outro selecionado aleatoriamente no conjunto de locais candidatos.

Dada a maior complexidade deste problema, foi necessário o desenvolvimento de um processo misto que envolve a SA, mas que também incorpora em sua estrutura um modelo de programação linear, conforme será visto a seguir.

Para o desenvolvimento deste processo foi feita uma subdivisão em quatro fases (figura 5.12):

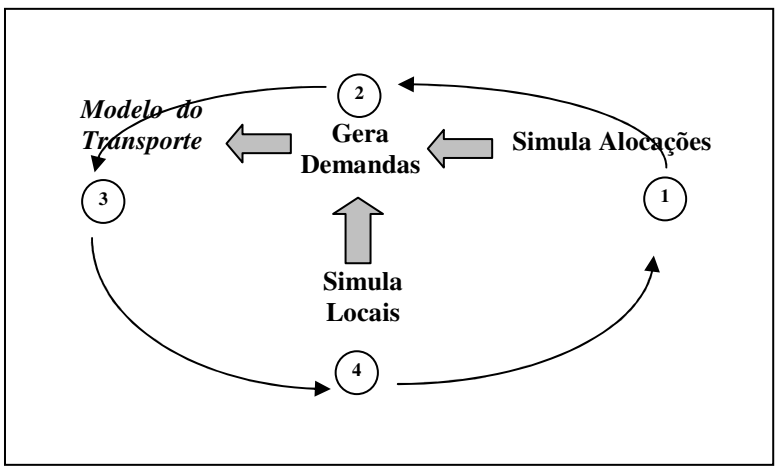

Figura 5.12 – Definição de localizações "ótimas" de CD's
Processo com meta-heurística SA em 4 Fases

Processo - Fase 1: Simulação de alocações de CD's a Regiões de Demanda

. Gera-se aleatoriamente uma troca de alocação.

. Uma região de demanda sorteada passa a ser atendida por um novo CD sorteado.

Processo - Fase 2: Geração de Demandas por CD

. Com as alocações definidas tem-se as Demandas de cada CD

Processo - Fase 3: Resolução de um "Problema do Transporte"

. Nesta fase deve-se determinar os fluxos das Fontes de Produção para os CD's

. Tem-se um conjunto de Fontes com capacidades definidas e um

conjunto de CD's com demandas agora conhecidas.

. Conhece-se também os custos unitários de transporte.

. Deseja-se buscar os fluxos que minimizam o custo global do transporte

. Tem-se assim definido um *Problema do Transporte* (Novaes, 1978).

. Roda-se então um algoritmo de Programação Linear com o *modelo de transporte* para otimizar esses fluxos.

Processo - Fase 4: Simulação de novos Candidatos

. Gera-se aleatoriamente uma troca de CD

.. Simula-se por sorteio um novo local candidato e retorna-se para a Fase 1.

A implementação destas quatro fases foi feita dentro do processo de simulação e busca da SA, conforme descrição na seqüência.

Processo de SA em 4 Fases

a) Leitura de dados de entrada e inicialização de parâmetros

Identificação de pontos de demanda e respectivas zonas

Identificação de candidatos por zona

Identificação de fontes de produção

Inicialização de parâmetros operacionais e de custos

Inicialização de parâmetros de controle da SA

b) Se total de Iterações atingiu limite, encerrar processo.

Em caso contrário, seguir para o próximo passo.

c) Gerar número aleatório n_o entre 0 e 1 para definir tipo de rearranjo a ser feito

c_1) Se n_o<=0,5, fazer rearranjo de CD´s

c_{11}) Da lista de candidatos sortear um novo ponto candidato P1

c_{12}) Sortear um ponto candidato da lista de CD´s em teste (P2)

c_{13}) Trocar o candidato P2 pelo candidato P1

c_2) Se n_o>0,5, fazer rearranjo de alocações de zonas a CD´s

c_{21}) Sortear uma zona

c_{22}) Trocar a alocação desta zona para um novo CD sorteado

d) Verificar se atende restrições

d_1) se atender, seguir para o próximo passo

d_2) se não atender, retornar para o passo (c)

e) Gerar demandas por CD

f) Resolver o *problema do transporte* entre fontes e CD´s

g) Recalcular a função objetivo de custo (equação 5.9)

g_1) Se o custo se reduziu ou se manteve, aceitar a nova solução

g_2) Se o custo aumentou, aceitar a nova solução se P(Boltzman)>n_o, sendo n_o um novo número aleatório gerado entre 0 e 1.

h) Verificar se iterações na temperatura ou número de sucessos da temperatura atingiram limite

h_1) se um deles atingiu, reduzir temperatura e retornar para passo (b)

h_2) se nenhum atingiu, retornar para passo (c)

Uma visão global da metodologia de solução envolvendo as quatro fases do processo, associado às características do problema é apresentada na figura 5.13.

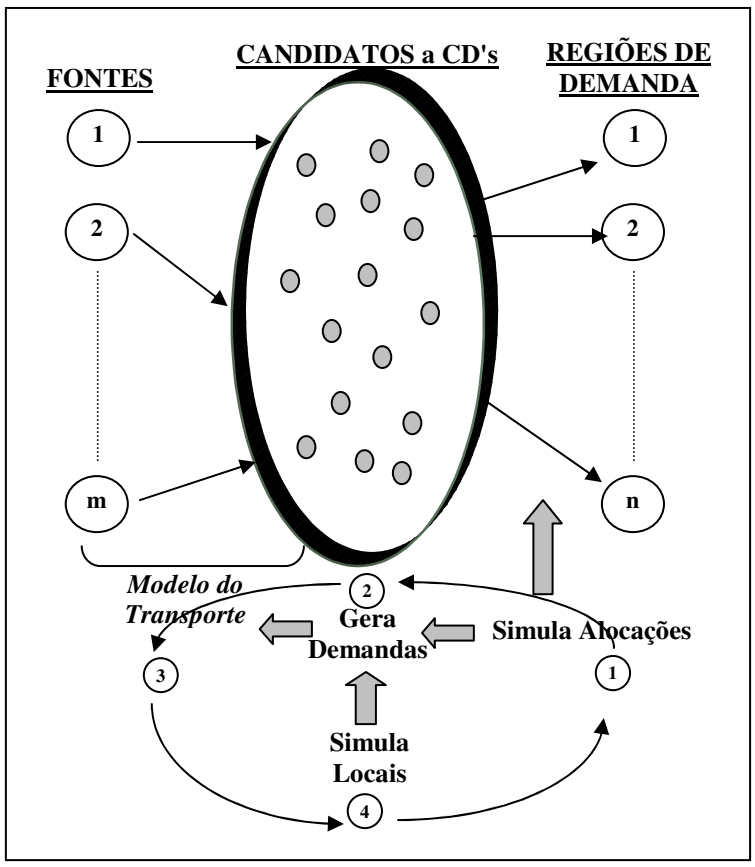

Figura 5.13 – Processo Heurístico em 4 Fases e características do problema

A implementação de todo este processo foi feita com o apoio de um programa de computador desenvolvido especificamente para esta pesquisa. Os parâmetros utilizados, aspectos operacionais e resultados obtidos são descritos nas seções seguintes.

5.3.2. Parâmetros da SA e Dados de Entrada - Localização

A experimentação com SA foi feita com os mesmos dados do experimento desenvolvido na validação das heurísticas (seção 4.5)

O sistema desenvolvido recebeu informações referentes às zonas atribuídas a cada ponto, às coordenadas e às demandas dos pontos. Além disto, são fornecidos também os candidatos e as fontes de produção.

Como restrição estabeleceu-se um máximo de candidatos em cada zona para evitar tendenciosidade. Foram feitos testes com um máximo de um e dois candidatos por zona.

Em termo de parâmetros da SA, foram definidos aqueles discutidos na seção 5.1, que são:.

Temperatura inicial (T_o)

Número máximo de iterações em cada temperatura (p)

Número máximo de sucessos em cada temperatura (L);

Número total de iterações (m);

Método de resfriamento;

A determinação destes parâmetros baseou-se nas diretrizes já tratadas na mesma seção 5.1, conforme já havia sido feito no uso da SA para a geração do zoneamento.

As definições desses parâmetros são apresentadas a seguir.

a) Temperatura inicial (T_o)

Esta foi definida a partir da equação 5.8. A figura 5.14 apresenta uma tela de saída do software desenvolvido para a AS, mostrando o cálculo final de ΔE_m^+ no experimento desenvolvido.

Com $\Delta E_m^+ = 679{,}44$ tem-se $T_o = 3043{,}89$. Conforme sugerido por Viana (1998), entretanto, passou-se a adotar uma divisão de ΔE por 100 antes de ser utilizado na probabilidade de Boltzman. Com isto reduz-se a escala e, conseqüentemente, a

própria temperatura, que fica também dividida por 100. O valor de To calculado passa a ser então T_o = 30,4389. Decidiu-se então por adotar T_o = 30,0.

b) Número máximo de iterações em cada temperatura (p)

Considerando-se que no experimento o número de candidatos girou em torno de 20 e que o número de zonas era igual a 5, tem-se um total de 25 variáveis no problema.

Conforme a diretriz apresentada na seção 5.1 ter-se-ia então p = 100.25 = 2500 iterações a cada temperatura.

Os resultados experimentais mostraram, entretanto, que este número estava além do necessário para que as soluções convergissem.

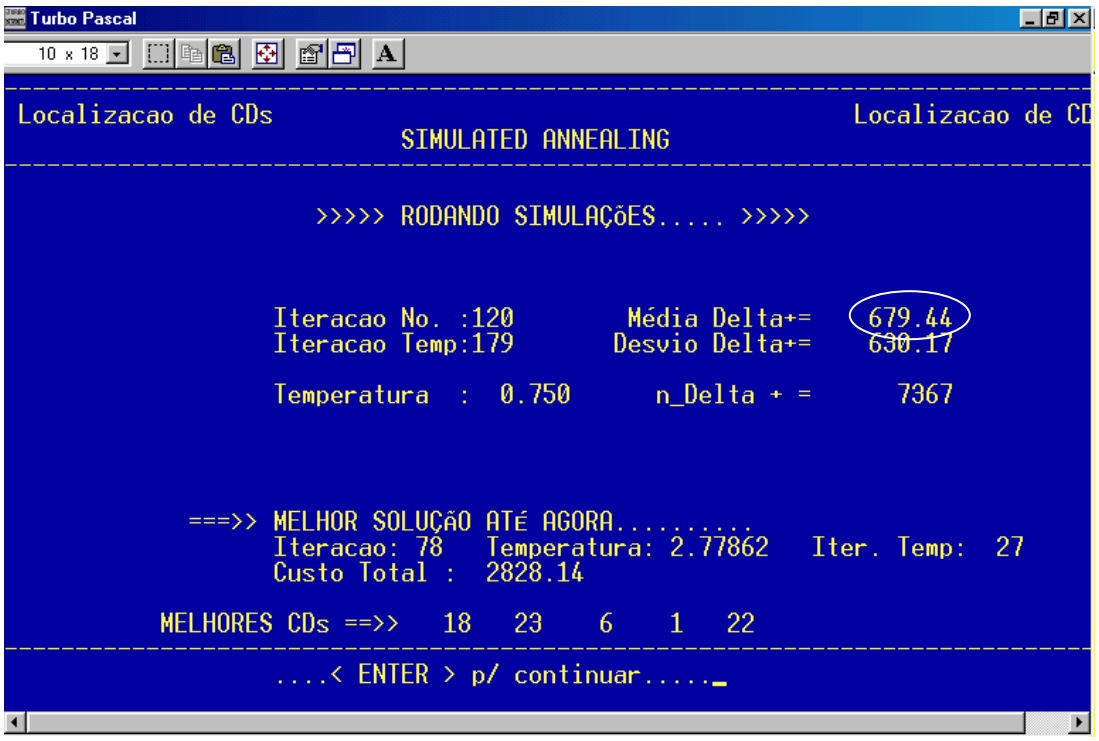

Figura 5.14 - Tela de resultados apresentando ΔE_m^+ na SA de localização

Na verdade, com um número bem menor de iterações as soluções já estavam convergindo. Foi adotado então o número de 500 iterações a cada temperatura, que se mostrou satisfatório.

c) Número máximo de sucessos em cada temperatura (L);

Por decorrência da definição de p=500 iterações, e adotando-se L igual a 10% de p (seção 5.1), tem-se: L=50 iterações.

d) Método de resfriamento;

Seguindo novamente as observações de Johnson(1989) e Viana(1998), que sugerem o uso de uma taxa geométrica de resfriamento variando entre 0,90 e 0,99, foram desenvolvidos testes para identificar a mais adequada neste caso. Chegou-se ao valor de Rc =0,97 como sendo aquele que apresentava o melhor grau de convergência da solução dentro de um número de iterações compatível com os valores definidos para o estudo.

d) Número total de iterações (m);

Resultados experimentais mostraram que o valor máximo da função objetivo definida apresentava grande oscilação até uma temperatura em torno de 8,0, quando então iniciava-se uma tendência de estabilização. A convergência final da solução, porém, efetivamente ocorre para temperaturas abaixo de 2,0.

Isto pode ser observado pela figura 5.15a, que apresenta a variação de Custo Mínimo à medida que a temperatura vai decrescendo. Este gráfico, entretanto, não permite uma boa visualização do processo, devido ao efeito da escala, a exemplo

do que ocorreu na SA de zoneamento; para temperaturas muito baixas a diferença absoluta entre duas temperaturas seguidas torna-se muito pequena.

Um detalhamento da fase final deste processo, propiciando uma visualização adequada é apresentado na figura 5.15b, em que se nota com mais clareza a convergência da solução. Naquele gráfico são apresentadas apenas temperaturas abaixo de 10,0.

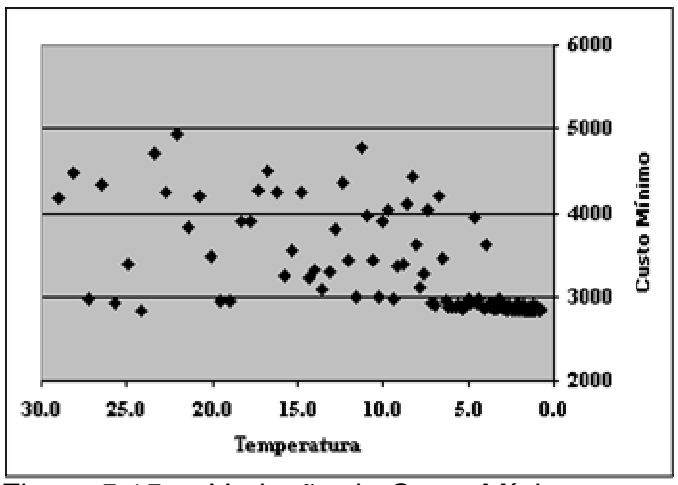

Figura 5.15a-: Variação de Custo Mínimo com o decréscimo de temperatura

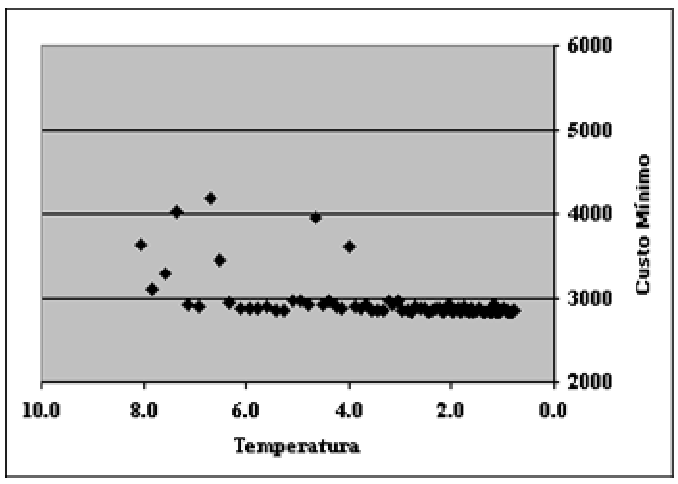

Figura 5.15b-: Detalhamento da variação de Custo Mínimo para temperaturas abaixo de 10,0

Assim, dentro de uma abordagem mais conservadora, foi adotado em princípio um valor mínimo de temperatura (T_m) de 1,0, o que garantiria com segurança a convergência da solução. Como o modelo estava sendo executado em tempos de processamento baixos, a atitude conservadora não gerou lentidão no processo.

Com o valor de T_m definido e utilizando-se a equação 5.6, foi possível determinar o número m máximo de iterações:

$$m = \log(1{,}0 / 30{,}0) / \log(0{,}97) \rightarrow \mathbf{m = 111{,}66}$$

Para garantira a convergência do processo, foi adotado então o número máximo de 120 iterações, resultando um valor final de temperatura um pouco inferior a 1,0. Este valor ficou em $T_m = 0{,}7997$.

5.3.3. Resultados Experimentais da Localização

O programa de computador desenvolvido para a simulação apresenta os resultados finais da melhor configuração gerada (figura 5.16). Na seqüência o sistema apresenta índices que monitoram a evolução do processo a cada temperatura.

Os índices calculados a cada temperatura i (i=1,2,....m) são:

. Custo Médio

. Desvio Padrão do Custo

. Coeficiente de Variação do Custo

. Taxa de Soluções Inferiores

. Taxa de Aceitação de Soluções Inferiores

. Custo Mínimo obtido na temperatura i.

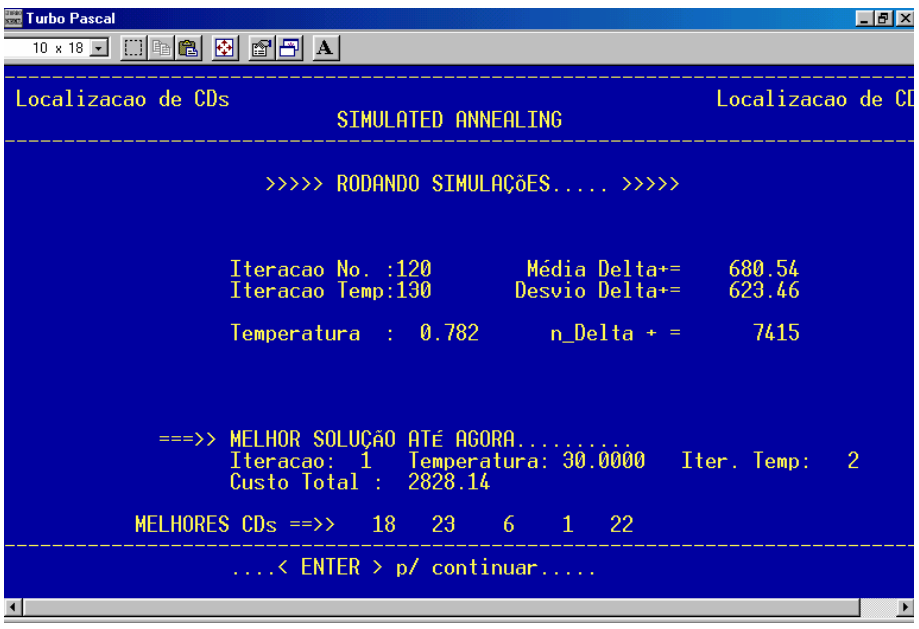

Figura 5.16 - Melhor configuração gerada

O sistema apresenta o valor da Temperatura, o Número de iterações naquela temperatura e os resultados de todos os índices.

A figura 5.17a apresenta a saída do sistema com o resultado dos índices nas temperaturas iniciais (mais altas) e a figura 5.17b mostra os mesmos índices nas temperaturas finais (mais baixas), quando as soluções convergem.

Gráficos do processo de evolução do Custo Médio e do Custo Mínimo a cada temperatura são apresentados nas figuras 5.18 e 5.19.

Percebe-se por estes diagramas a eficácia do processo de SA. Na figura 5.18 o custo médio se reduz consistentemente ao longo de todo o processo e o custo mínimo de cada temperatura converge de forma definitiva para um valor abaixo de todos os outros quando a temperatura cai para valores bem baixos.

A figura 5.16 apresentou a melhor configuração global para cinco CD's, aquela que gerou o custo mínimo. Este melhor resultado do experimento corresponde à mesma solução encontrada diretamente pelo modelo MILP, apresentado no capítulo 4.

Turbo Pascal

Temp	Iterac	CT	Desv	CV	Tx Inf	TxAcInf	CTmin
30.00	54	5125.5	860.5	16.8	57.4	87.1	2828.14
29.10	54	5051.6	661.0	13.1	57.4	87.1	3098.96
28.22	53	5459.8	574.5	10.5	60.4	90.6	4161.54
27.38	52	5284.4	895.3	16.9	55.8	93.1	3591.40
26.55	53	5290.2	622.6	11.8	52.8	89.3	4121.13
25.76	56	4683.4	745.4	15.9	55.4	80.6	3477.87
24.98	57	4962.9	498.4	10.0	56.1	78.1	4261.24
24.23	54	5012.1	806.0	16.1	50.0	85.2	2853.08
23.51	57	5784.6	954.6	16.5	63.2	80.6	3740.05
22.80	52	5292.6	855.6	16.2	46.2	91.7	3654.09
22.12	54	5271.8	948.5	18.0	51.9	85.7	3506.98
21.45	52	5079.0	886.2	17.4	59.6	93.5	3423.95
20.81	53	5116.0	795.2	15.5	50.9	88.9	2955.16
20.19	51	5733.0	839.8	14.6	52.9	96.3	4311.80
19.58	60	5063.8	768.8	15.2	58.3	71.4	3371.68
18.99	59	4622.4	856.1	18.5	59.3	74.3	2849.06
18.42	53	5350.3	813.3	15.2	52.8	89.3	2929.01
17.87	56	5619.5	692.5	12.3	55.4	80.6	3835.15
17.33	53	5669.0	469.1	8.3	47.2	88.0	4629.54
16.81	54	5357.1	730.9	13.6	55.6	86.7	3703.33

....< ENTER > p/ continuar.....

Figura 5.17a - Resultados de índices de monitoramento do processo
Temperaturas iniciais (mais altas)

Turbo Pascal

1.426	148	2992.8	137.4	4.6	81.1	18.3	2828.14
1.383	125	2968.8	85.6	2.9	77.6	22.7	2847.22
1.342	155	2950.3	49.7	1.7	84.5	19.8	2849.04
1.302	117	2958.3	72.1	2.4	78.6	27.2	2860.90
1.262	105	3037.9	88.4	2.9	79.0	33.7	2884.89
1.225	110	2960.2	67.8	2.3	77.3	29.4	2847.22
1.188	133	2952.4	50.9	1.7	82.7	24.5	2847.22
1.152	97	2951.0	65.1	2.2	74.2	34.7	2849.04
1.118	130	3010.1	95.3	3.2	79.2	22.3	2847.22
1.084	119	2946.5	65.3	2.2	79.8	27.4	2829.97
1.052	132	2944.4	94.6	3.2	81.1	23.4	2828.14
1.020	161	2917.7	68.9	2.4	84.5	18.4	2828.14
0.989	152	2937.2	64.0	2.2	84.9	20.9	2829.97
0.960	148	3001.8	142.4	4.7	81.1	18.3	2847.22
0.931	134	2959.6	74.7	2.5	81.3	22.9	2853.08
0.903	154	2983.3	75.1	2.5	83.8	19.4	2829.97
0.876	127	2997.6	114.5	3.8	80.3	24.5	2849.06
0.850	132	2961.9	74.2	2.5	79.5	21.9	2860.92
0.824	153	2973.3	71.3	2.4	83.7	19.5	2829.97
0.799	130	2929.4	56.8	1.9	83.1	25.9	2829.97

....< ENTER > p/ continuar.....

Figura 5.17b - Índices de monitoramento – Convergência de soluções
Temperaturas finais (mais baixas)

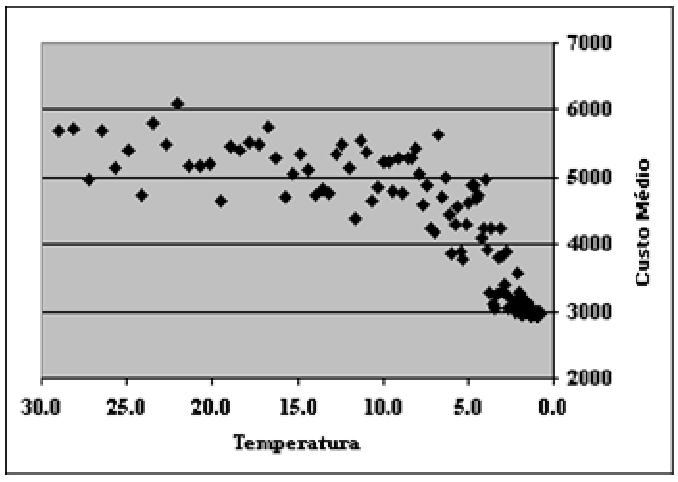

Figura 5.18 – Evolução do Custo Médio com a redução de temperatura.

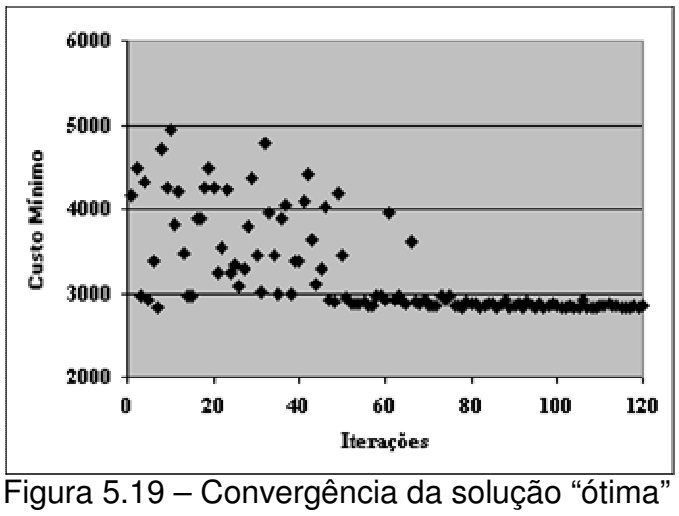

Figura 5.19 – Convergência da solução “ótima” . Partindo de uma solução diferente da “ótima”.

A heurística proposta nesta pesquisa, portanto, se mostrou eficaz gerando uma solução tão boa quanto a de um modelo de otimização para o problema estudado.

Sobre as soluções geradas pela SA ao longo do processo, deve ser salientado que estas convergiram invariavelmente para a “melhor” solução.

Outro aspecto importante se refere ao fato de que, para diferentes configurações iniciais, o processo sempre convergiu para a “melhor” solução. A figura 5.19

mostrou a convergência quando a configuração inicial é distinta da “melhor” configuração; já na figura 5.20 utilizou-se como configuração inicial a própria solução “ótima” e verifica-se que o processo leva ao final a um retorno a esta mesma solução.

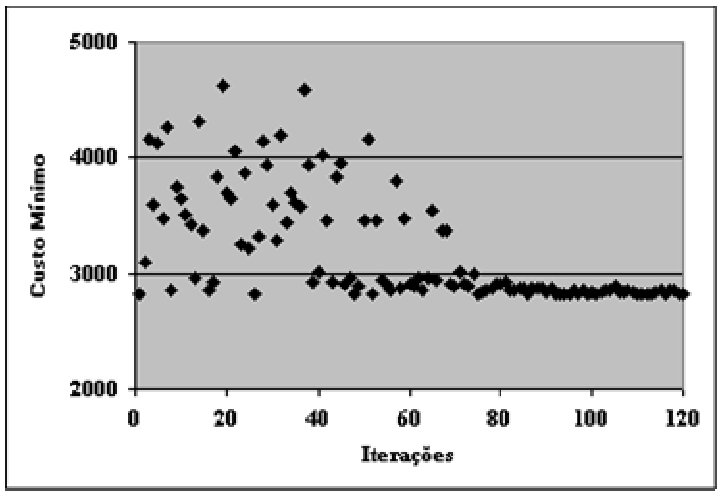

Figura 5.20 – Convergência da solução “ótima” .
Inicia na solução “ótima” e retorna à “ótima”

5.4. COMPARAÇÃO DE RESULTADOS

Após toda esta extensa experimentação com o uso de diferentes métodos, convém agora que se faça um balanço dos resultados obtidos nos dois processos desenvolvidos: o experimento com o modelo MILP e a aplicação da meta-heurística SA.

No caso do experimento com o modelo MILP, já foi observado na seção 4.5 que havia sido construído um modelo MILP desagregado para solucionar o problema. Este modelo não passou pelo processo de *clusterização,* tendo-se considerado cada ponto de demanda de forma individual. Com isto consegue-se identificar a

efetiva solução ótima do problema, tendo-se assim um padrão seguro de comparação.

Tem-se, portanto, três resultados distintos a partir de três processos de experimentação desenvolvidos:

. resultados do modelo MILP agregado;

. resultados da *simulated annealing*

. resultados do modelo MILP desagregado *(a solução ótima)*.

Os resultados deste último modelo estarão sendo denominados aqui de solução ótima, já que como neste caso os pontos de demanda não foram agregados, esta pode ser considerada como a efetiva solução ótima do problema.

Os resultados dos três processos podem ser visualizados através das figuras 5.21 a 5.25.

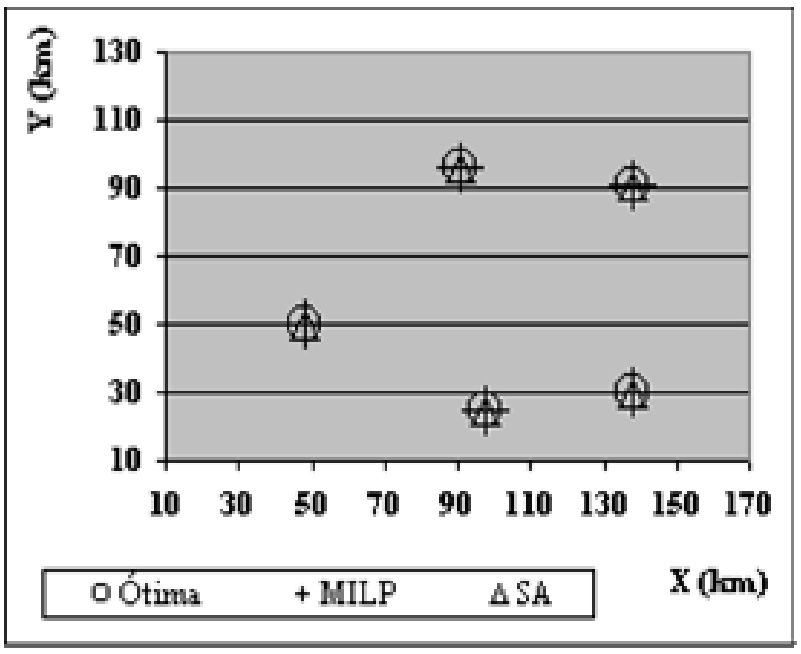

Figura 5.21 – Comparação de resultados – 5 CD's

Em alguns casos a solução ótima encontrada (modelo desagregado) foi a mesma obtida pelo modelo MILP agregado e pela SA. Isto ocorreu nos casos de cinco CD´s

e de um único CD.. Em outros casos as soluções foram muito próximas, sendo praticamente coincidentes, como se observa nas configurações de quatro CD´s e três CD´s.

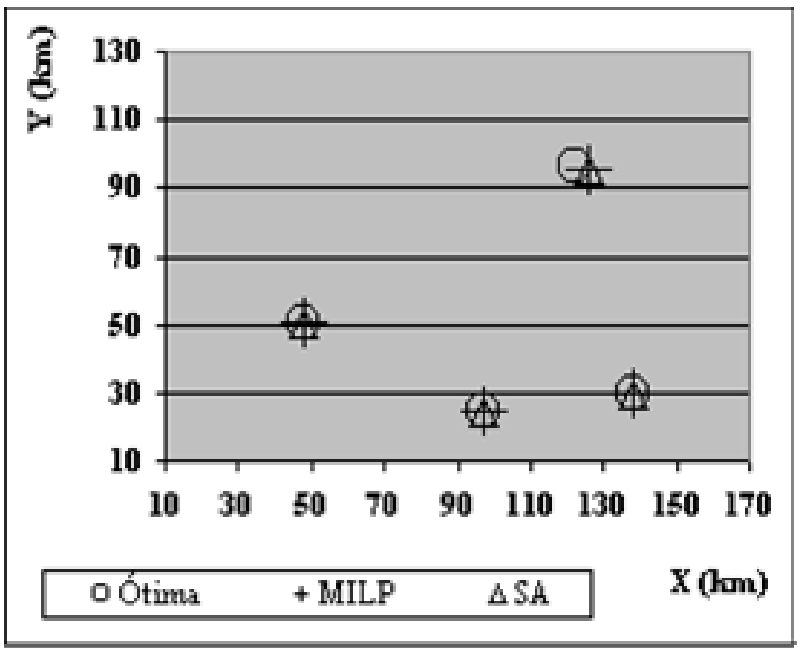

Figura 5.22 – Comparação de resultados – 4 CD's

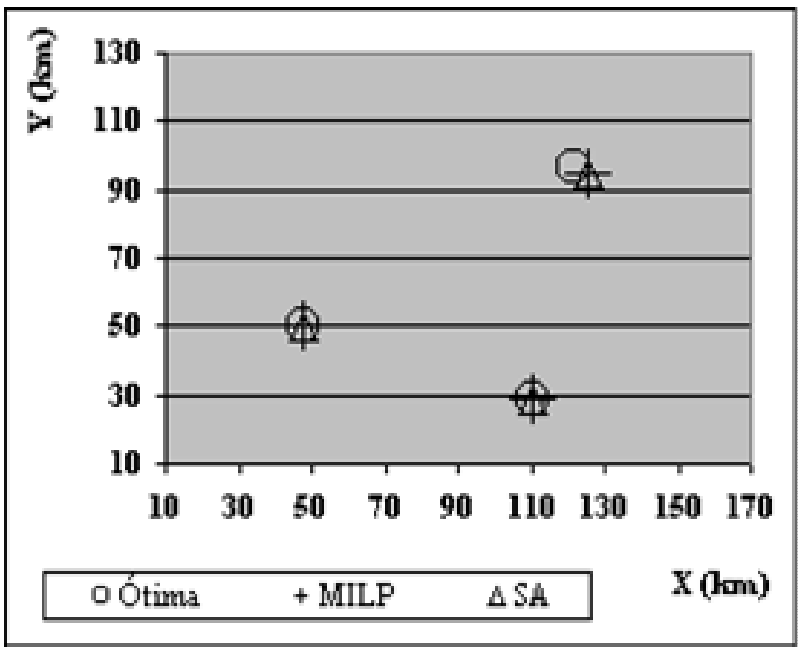

Figura 5.23 – Comparação de resultados – 3 CD's

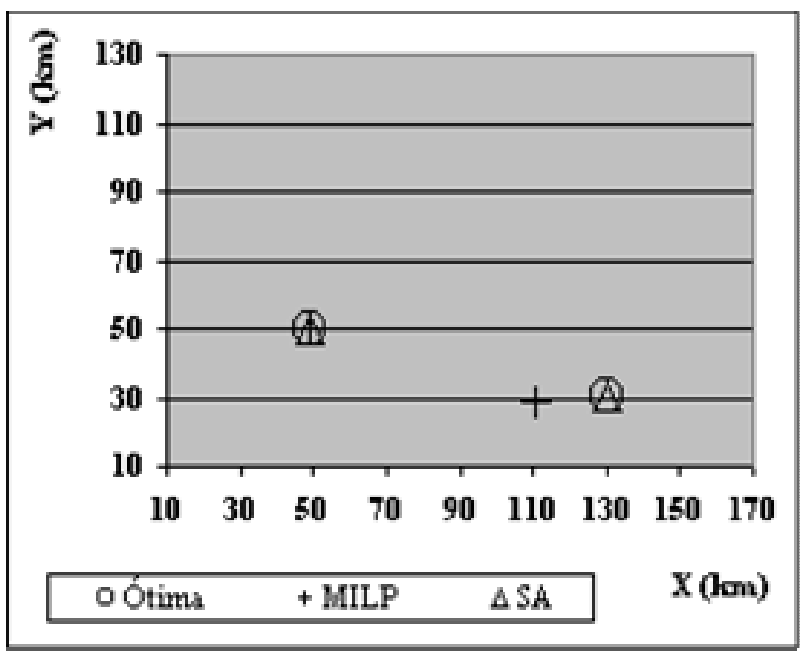

Figura 5.24 – Comparação de resultados – 2 CD's

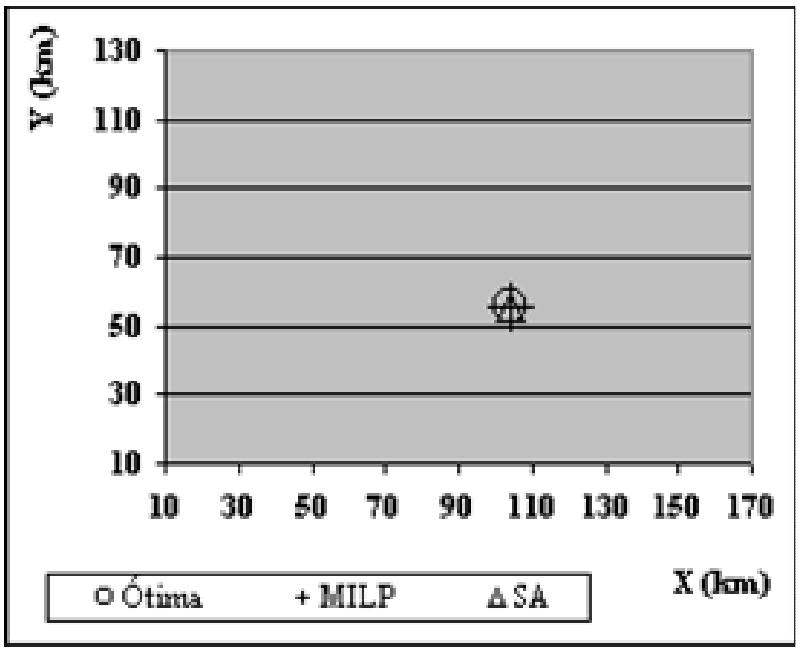

Figura 5.25 – Comparação de resultados – 1 CD

Nestas duas situações as soluções pelo modelo MILP e pela SA foram as mesmas, não coincidindo com a solução ótima em apenas um único ponto. Mas, mesmo neste caso, a diferença foi muito pequena. O único caso em que houve uma ligeira

diferença entre as soluções foi para a configuração de dois CD's. Aqui a solução pelo modelo MILP agregado teve um dos pontos ligeiramente deslocado para oeste, enquanto a solução pela SA coincidiu exatamente com a solução ótima.

Este conjunto de resultados mostra que a metodologia proposta está conduzindo o problema a uma solução muito próxima da solução ótima, senão coincidente. Seja pelo uso de um modelo MILP, seja através da meta-heurística SA.

6. MODELAGEM DE QUANTIFICAÇÃO DE CD's

Este capítulo apresenta a *Fase 2* da metodologia proposta na pesquisa. Nesta fase pretende-se encontrar a quantidade "ótima" de CD's, adotando-se uma abordagem que considera outros custos logísticos além daqueles já contemplados na *Fase 1*. Estes novos custos serão tratados de forma externa ao modelo adotado na *Fase 1*, seja o modelo MILP, seja a meta - heurística de localização. Esta é uma fase, portanto, que recebe os resultados da *Fase 1* e agrega novos elementos para estabelecer a solução final.

Na *Fase 1* foram considerados na modelagem os custos de transporte e de localização dos CD's. Nesta nova etapa a análise é ampliada, passando-se a incluir outros custos logísticos, podendo-se chegar inclusive, se desejado (a metodologia assim o permite) à totalidade destes custos. O objetivo da análise será assim a determinação da quantidade de CD's que conduz à minimização de um conjunto de custos logísticos considerados.

Uma vez estabelecida a quantidade "ótima" de CD's, estes deverão se situar em locais que foram definidos na *Fase 1.*

6.1. CUSTOS LOGÍSTICOS EXTERNOS AO MODELO DE OTIMIZAÇÃO

Na abordagem aqui adotada supõe-se que os custos de transporte e de localização já estarão determinados na *Fase 1* e de forma externa aos modelos lá utilizados são computados agora os demais custos logísticos. Estes custos, uma vez computados,

são então somados aos custos de transporte e localização já otimizados, tendo-se assim, um Custo Logístico Global.

A quantidade de CD's que apresentar o menor valor para este custo global será considerada a "melhor" quantidade.

Esses novos custos poderão nesta *Fase 2* receber um tratamento mais realista do que aquele que seria dispensado caso fossem tratados na *Fase 1*. Custos, por exemplo, que têm comportamento não linear, podem ter este tipo de tratamento, como é o caso de custos de estoques e de sua evolução com o número de CD´s.

Um outro aspecto importante do método é que agora passa a ser possível a consideração de variações em custos unitários. Ao se variar a quantidade de CD´s, por exemplo, os custos fixos de armazenagem podem permanecer estáveis ou sofrer alterações. Os dois tipos de situações podem ser adequadamente tratados em uma modelagem como a aqui proposta.

Nas próximas seções apresenta-se o detalhamento da modelagem para os três tipos de custos aqui considerados: estoques, armazenagem e movimentação de materiais, e vendas perdidas.

6.2. MODELAGEM DOS CUSTOS DE ESTOQUE

Para efeito deste trabalho, estão sendo considerados como custos de estoque aqueles componentes que estão associados ao valor do material em estoque. Não estão incluídos aqui, portanto, quaisquer custos relacionados à estrutura física e administrativa necessária para manter o material armazenado. Estes serão tratados na próxima seção.

Dentro desta visão o componente principal destes custos de estoques é o chamado custo de oportunidade do capital que está investido em estoque (Ballou, 1992). Este é um custo que é computado através de formulações já bem definidas em matemática financeira, conforme será visto na equação 6.1 (Adler 1978; Fabrycky,1998).

A idéia deste custo de oportunidade é que quando um capital encontra-se imobilizado (sem aplicação financeira ou na produção), esta quantia só será recebida numa data futura e perde-se, assim, a oportunidade do investimento. Representa no fundo, o "valor do tempo" durante o qual a quantia não está disponível para aplicação.

A equação (6.1) representa o cálculo deste capital investido em estoque para uma taxa ***i*** de juros:

$$C_{fn} = V_e \,.\, [\,(1 + i)^{tp} - 1\,] \qquad (6.1)$$

onde: C_{fn} = custo financeiro de estoque após t_p períodos.

V_e = valor investido em estoque

i = taxa de juros por unidade de tempo

t_p= tempo médio de permanência de cada unidade em estoque

Considerando-se que t_p representa o tempo médio de permanência de cada unidade em estoque e substituindo-se em 6.1 V_e pelo custo unitário do material (V_u), obtém-se o custo unitário de estoque (CE_u), que representa o custo que irá incidir sobre cada unidade de demanda.

$$CE_u = V_u \,.\, [\,(1 + i)^{tp} - 1\,] \qquad (6.2)$$

O parâmetro t_p pode ser interpretado também como sendo a chamada cobertura de estoque (Ballou,1992; Chopra e Mendl, 2001).

Para uma dada taxa de demanda, a cobertura de estoque é o tempo em que aquele estoque será consumido. Não ocorrendo novas entradas de material em armazém durante um período de cobertura, significa que ao final desse período o estoque estará zerado.

Assim, cobertura de estoque na verdade é o estoque expresso em unidades de tempo de demanda.

Este parâmetro t_p é calculado através de.

$$t_p = C_R + t_t + C_s \qquad (6.3)$$

sendo:

C_R = cobertura do estoque regulador;

t_t = tempo em trânsito;

C_s = cobertura do estoque de segurança;

Estes três parâmetros são detalhados mais abaixo. Antes disto, entretanto, convém salientar que a esta parcela principal do custo financeiro deve-se acrescentar outros custos gerados pelo estoque, tais como: seguros; obsolescência de produtos; perdas por faltas e/ou avarias; outras taxas.

O custo de estoque final será a soma de todos estes itens:

$$CE_u = V_u \,.\, [\,(1 + i)^{tp} - 1\,] + O_{est}.V_u \qquad (6.4)$$

onde: O_{est} = taxa de outros custos de estoques, englobando seguros, obsolescência, perdas, etc.

V_u = valor unitário do material em estoque

A formulação de submodelos que representam as coberturas dos diferentes tipos de estoque é apresentada a seguir.

a) Cobertura de estoque regulador

Estoque regulador ou cíclico é aquele que é utilizado para regular a oferta e demanda. É um estoque que permanece armazenado nos CD´s por um certo período até que seja embarcado para os pontos de demanda. Tem a função assim, de um “pulmão” para ir atendendo à demanda. (Chopra e Mendl, 2001).

Este estoque é composto de uma parcela para atender a necessidade imediata de embarque e uma outra que antecipa necessidades. Será analisada inicialmente a parcela correspondente às necessidades de embarque, que será denominada aqui de estoque sincronizado. Nos caso de sistema *just in time* esta é a única parcela que existe (Blumenfeld et al.,1985).

Considere-se que V_E seja o lote produzido nas fontes de produção e que será embarcado para os CD´s. Este lote V_E representa então o estoque necessário no momento do embarque. O lote V_E corresponderá ao estoque regulador máximo entre embarques. Conforme Blumenfeld et al. (1985) demonstram, considerando-se uma taxa de demanda diária *d*, para uma taxa de produção constante e igual à taxa de demanda, a relação V_E/d representa a cobertura máxima de estoque. Este princípio de equivalência de taxas de produção e de demanda será adotado nesta modelagem.

Assim, neste período entre embarques, o estoque médio E_M na origem (produção) será a metade de V_E, ou seja,

$$E_M = V_E / 2 \qquad (6.5)$$

Como se tem estoque também no destino (consumo), o estoque médio no destino será também $V_E/2$. O estoque médio total será então, $V_E/2 + V_E/2$, ou seja, será V_E (figuras 6.1a e 6.1b).

A cobertura total de estoque regulador sincronizado (C_{RS}) será então:

$$C_{RS} = V_E / d \qquad (6.6)$$

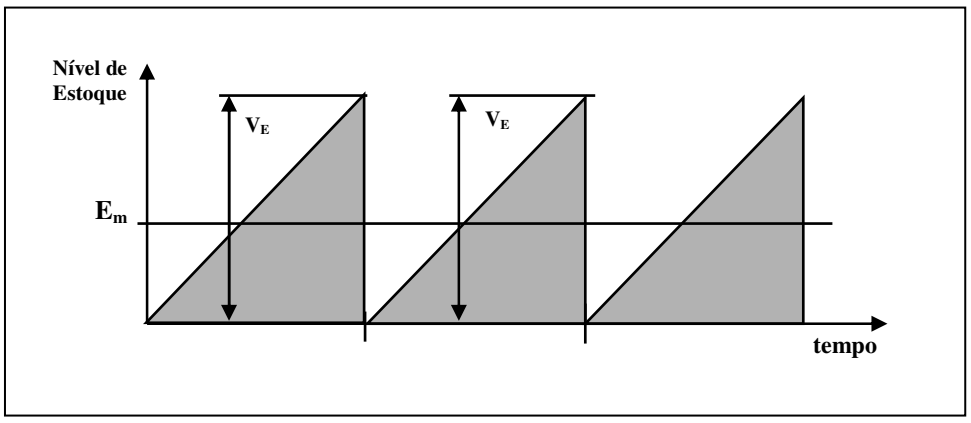

Figura 6.1a – Evolução de estoque e estoque médio - Origem

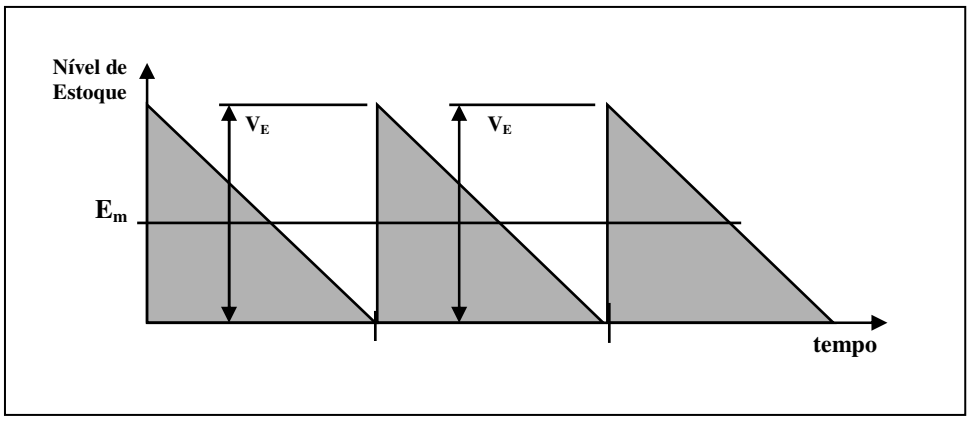

Figura 6.1b – Evolução de estoque e estoque médio - Destino

Note-se que C_{RS} é também uma medida do nível de serviço, pois representa o período (T_E) entre embarques. Assim, para um nível de serviço definido de atendimento, a cada T_E dias tem-se um lote de embarque $V_E = d.T_E$. Logo, tem-se $C_{RS} = T_E$.

Portanto, C_{RS} também pode ser visto como uma função do nível de serviço definido.

Uma outra parcela do estoque regulador é o estoque adquirido por antecipação para fazer frente à demanda futura ou até mesmo para aproveitar oportunidades de

mercado. Segundo Ballou (1994) este estoque pode ser estimado através de uma análise de regressão (Draper e Smith, 1981) entre os níveis de estoque da empresa (E) e o volume movimentado (V), com uma equação do tipo:

$$E = a.V^{b} \qquad (6.7)$$

em que a e b são parâmetros a determinar pela regressão.

Assim, a estimativa final do estoque regulador (C_R) será:

$$C_R = V_E / d + a.V^{b} \qquad (6.8)$$

Se o volume V considerado for diário, será então a própria demanda d.

b)Estoque em trânsito

Este estoque corresponde ao tempo em trânsito (t_t) do material, ou seja, o tempo que o estoque permanece em viagem. Isto corresponde, na verdade, a uma necessidade de cobertura adicional de estoque devida ao trânsito do material.

Esta cobertura quando expressa em horas será:

$$t_t = D_{CD} / V_m \quad \textit{(em horas)} \qquad (6.9)$$

onde; D_{CD}= distância média percorrida em km (transferência e distribuição)

V_m = velocidade de viagem (km/h)

Estes parâmetros de distância e velocidade devem ser ponderados com os respectivos volumes movimentados.

Agora, definindo-se J=jornada diária de trabalho (em horas/dia) tem-se:

$$t_t = D_{CD} / (V_m.J) \quad \textit{(em dias)} \qquad (6.10)$$

c) Estoque de segurança

Estoque de segurança é uma quantidade adicional que se mantém em armazém para proteção contra variações de demanda e de prazos de entrega (Correa et al., 2001)

No caso de uma única instalação o estoque de segurança é determinado por (Chopra e Mendl,2001):

$$Es = Z_{\alpha}.s_d.(LT)^{1/2} \qquad (6.11)$$

onde:

Z_{α}= valor crítico da distribuição normal reduzida ao nível 1- α de confiança

s_d=desvio padrão da demanda

LT = lead time (medido na mesma periodicidade da demanda)

Este nível *1-α* corresponde ao nível de serviço oferecido em termos de disponibilidade do material em estoque. Com um nível *1-α*, espera-se que somente α% da demanda não será atendida por falta do material (Ballou, 1992).

A cobertura correspondente será:

$$C_s = E_s/d = Z_{\alpha}.s_d.(LT)^{1/2}/d \qquad (6.12)$$

Sabe-se, entretanto, que à medida que a quantidade de CD's aumenta, o estoque de segurança sofre uma variação. Esta variação é no sentido crescente e na mesma razão da raiz quadrada da proporção deste aumento (Ballou,1992; Fleury et al.,2000).

Assim, se o número de CD´s passa de p_1 CD´s para p_2 ($p_2>p_1$) então a relação 6.13 pode estimar a evolução do estoque de segurança:

$$Es_2 = \sqrt{\frac{p_2}{p_1}}.Es_1 \qquad (6.13)$$

onde:

Es_i = estoque de segurança com p_i CD´s , (i=1,2)

Considerando-se como base um único CD, então para *p* CD´s ($p \geq 1$) a variação de estoque de segurança será de $\sqrt{p}$.

Assim, o equacionamento final da cobertura de segurança, para um número *p* de CD´s será:

$$C_s=(Z_\alpha.s_d.(LT)^{1/2}/d). \sqrt{p} \qquad (6.14)$$

A partir de 6.6, 6.8 e 6.12 tem-se então:

$$t_p = \frac{V_E}{d} + a.V^b + \frac{D_{CD}}{V_m.J} + \frac{Z\alpha.s_d\sqrt{LT}}{d}\sqrt{p} \qquad (6.15)$$

Lembrando que t_p representa o tempo que cada unidade de material demandado deverá permanecer em estoque e voltando-se agora à equação 6.2, pode-se completar o modelo definindo-se o custo de estoques que irá incidir sobre cada unidade de demanda.

Com base nas equações 6.2 e 6.4 e utilizando-se a formulação estabelecida em 6.15 obtém-se o modelo para o *custo unitário de estoque (CE_u)* para um nível de serviço de estoque igual a *1-α* e para uma quantidade *p* de CD´s:

$$CE_U = V_u.\left\{\left[(1+i)^{\left(\frac{V_E}{d} + aV^b + \frac{D_{CD}}{Vm.J} + \frac{Z\alpha.s_d\sqrt{LT}}{d}\sqrt{p}\right)} - 1\right] + O_{est}\right\} \qquad (6.16)$$

Para se obter o custo global de estoques (C_{est}), basta multiplicar o custo unitário acima pela demanda global adotada (D).

O modelo final de custo de estoques será então:

$$C_{est} = D.V_u.\left\{\left[(1+i)^{\left(\frac{V_E}{d} + a.V^b + \frac{D_{CD}}{V_m.J} + \frac{Z\alpha \cdot s_d\sqrt{LT}}{d}\sqrt{p}\right)} - 1\right] + O_{est}\right\} \quad (6.17)$$

6.3. MODELAGEM DOS CUSTOS DE ARMAZENAGEM E MOVIMENTAÇÃO DE MATERIAIS

Nesta seção serão tratados os custos de operação dos CD's no que se refere à estrutura necessária para desenvolver os processos de armazenagem e movimentação dos materiais. Serão considerados aqui os custos fixos que correspondem a equipamentos, mão de obra e ao próprio imóvel utilizado na armazenagem e movimentação dos materiais. Esta necessidade de área, entretanto, pode ser tratada também na *Fase 1* da metodologia, que pode levar em conta no modelo o custo da área necessária de armazém em cada local para operar a demanda atendida pelos CD´s. Isto, se não houver homogeneidade de custos do solo na área de estudo. Havendo homogeneidade, todo o conjunto de custos de armazenagem e movimentação de materiais deve preferencialmente ser tratado nesta *Fase 2,* de modo a formar um bloco único.

Será considerado nesta modelagem que estes custos são na verdade fixos dentro de certos níveis de produção, admitindo-se, portanto, que estarão sujeitos a um processo de evolução não contínua. O modelo considera que esta evolução se dará em patamares, à medida que o estoque mantido em CD evolui. A hipótese do modelo é de que se necessita de uma certa estrutura operacional fixa para atender

a uma faixa de estoque a ser movimentado. Dentro desta faixa o custo se mantém, mas ultrapassado um certo limite, há necessidade de estrutura adicional e o custo, portanto, cresce. Estes custos, assim, são fixos dentro de cada patamar. Desta forma, o que ocorre é um processo de evolução descontínua dos custos. Será admitido que a evolução em patamares está associada a módulos de área de armazenagem. Cada módulo correspondente a uma área padrão com sua respectiva estrutura (equipamentos, mão de obra, etc). Assim, atingido o limite superior do nível de estoque de um dado patamar de custo, este evolui verticalmente para o patamar superior correspondente a um acréscimo do módulo de área (figura 6.2).

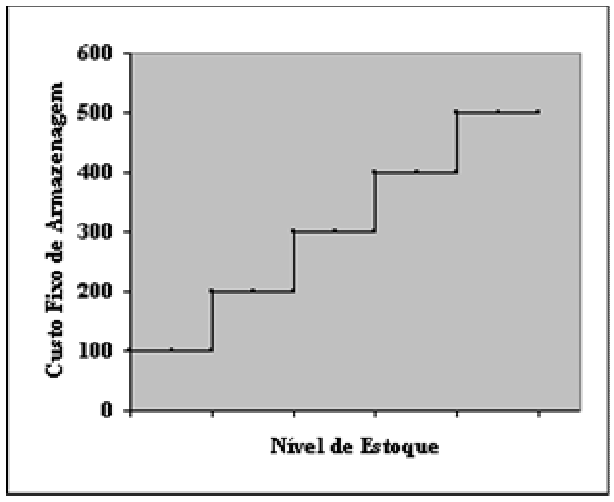

Figura 6.2 – Custos fixos de movimentação e Armazenagem - Evolução descontínua

No mundo real um módulo de área pode apresentar, na verdade, pequenas variações em sua estrutura, à medida que o volume a armazenar varia entre os limites inferior e superior do patamar. Mão de obra, por exemplo, pode ter uma variação quase contínua e mesmo alguns tipos de equipamentos, como, paletes, carrinhos e outros. Assim, na prática, nem todos os custos se mantém

absolutamente fixos dentro de cada patamar. Uma parcela desses custos pode apresentar variações.

Considerando-se este fato, a metodologia aqui adotada assume que os patamares irão se comportar na verdade mais como um referencial de custo do que como um custo fixo propriamente dito. Assim, a proposta do modelo é que para cada patamar se estabeleça um custo padrão *unitário* fixo e não um custo *total* fixo. Este custo padrão será aplicado a qualquer volume situado entre seus limites. O que fica fixo, portanto, dentro de cada patamar, é o custo unitário. Para representar este padrão será adotado o custo médio do patamar. Assim, dentro de um patamar sempre se terá o mesmo custo médio unitário, porém o custo total apresentará pequenas variações à medida que o volume varia dentro do patamar. Ao se passar para um patamar superior o custo padrão aumenta e aí se sente o efeito do patamar. Tem-se assim, dois efeitos: o primeiro, do patamar, que define um *nível* de custo, e um segundo dentro do patamar, que é função da variação do próprio volume dentro dos limites definidos.

Note-se que com esta proposta é possível a consideração de *economias de escala* devidas a eventuais mudanças tecnológicas entre patamares. Este é um aspecto que traz para a modelagem maior realismo, pois no que tange a equipamentos é um fato concreto a opção por mudanças de tecnologia, algumas vezes significativas, à medida que crescem as dimensões dos armazéns.

Por outro lado, se não houver esta mudança tecnológica o modelo também se adeqüa da mesma forma. O aspecto *economia de escala* é considerado no modelo pelo valor do custo médio unitário em cada patamar. Este custo é que define a questão. A modelagem, portanto, está aberta a duas situações: mudança de patamar sem economia de escala e mudança de patamar com economia de escala.

Definindo-se, agora, uma faixa de estoque situada entre dois níveis L_j e L_{j+1} para os quais o custo total de armazenagem e movimentação não se "altera", é possível desenvolver um equacionamento para este custo.

Os custos unitários de um patamar correspondentes aos diferentes níveis de estoque são representados por uma função de potência e o custo médio é aquele correspondente ao volume médio mantido em armazém, conforme a equação:

$$CM_{Lj\,;\,Lj+1} = CF_{Lj\,;\,Lj+1}.\ VM^{-1}_{Lj\,;\,Lj+1} \qquad (6.18)$$

onde: $CM_{Lj\,;\,Lj+1}$ = custo fixo médio de armazenagem e movimentação no patamar de estoque L_j a L_{j+1}

$CF_{Lj\,;\,Lj+1}$ = custo fixo de armazenagem e movimentação no patamar de estoque L_j a L_{j+1}

$VM_{Lj\,;\,Lj+1}$ = $(L_j + L_{j+1})/2$ = volume médio do patamar

O custo fixo ($CF_{Lj\,;\,Lj+1}$) de armazenagem e movimentação de materiais deve ser definido em cada patamar L_j a L_{j+1}. Sua composição é apresentada por autores como Ballou (1992) e Moura (1997) os quais mostram que no que tange a imóveis e equipamentos, seus componentes principais são depreciação e custo de capital investido e no que se refere a mão de obra, devem ser considerados os custos de pessoal administrativo e pessoal operacional.

Segundo Fabrycky (1998) e Adler (1978), uma das formas de se calcular a depreciação seria através da equação (6.19):

$$Dp_{Lj} = [\,(1 - i_r).\ E_{Lj\,;\,Lj+1}\,] / V_u \qquad (6.19)$$

onde:

Dp_{Lj} = depreciação mensal no patamar L_j , L_{j+1}

i_r = índice residual de instalações / equipamentos

E_{Lj} = necessidade de investimento em equipamentos

para atender até o nível L_{j+1} de estoque.

V_u = vida útil das instalações / equipamentos (meses)

Utilizando-se a equação (6.1), agora para o valor investido em equipamentos e fazendo-se t_p=1mês, obtém-se a equação (6.20) a ser aplicada no cálculo do custo de capital empregado:

$$Cf_{Lj} = E_{Lj} . [(1 + i_m) - 1] \quad (6.20)$$

onde: Cf_{Lj} = custo de capital empregado em instalações / equipamentos para atender até o nível L_{j+1} de estoque.

i_m = taxa mensal de juros.

O custo fixo de armazenagem e movimentação final no patamar de estoque L_j , L_{j+1} será representado pela equação 6.21.

$$CF_{Lj\,;\,Lj+1} = Dp_{Lj} + Cf_{Lj} + C_{MO-Lj} \quad (6.21)$$

onde: C_{MO-Lj} = outros custos de armazenagem referentes à pessoal e despesas de administração e de operações para atender até o nível L_{j+1} de estoque.

Tendo-se $CF_{Lj;Lj+1}$ pode-se finalmente, determinar o custo de armazenagem e movimentação de materiais para um dado CD. Aplicando-se um custo unitário por dia de armazenagem ao estoque médio que será mantido em armazém, medido em dias de consumo (C_R + C_S) e multiplicando-se o resultado pela demanda esperada (D), tem-se o custo final de armazenagem e movimentação de materiais (C_{Arm}) do CD:

A equação abaixo apresenta este resultado:

$$C_{Arm} = [(Dp_{Lj} + Cf_{Lj} + C_{MO-Lj}). VM^{-1}_{Lj\,;\,Lj+1} / D_u]. (C_R + C_S).D. \quad (6.22)$$

onde: D_u = dias úteis de operação por mês

D= demanda do CD

Note-se que uma parcela da equação (6.22) representa o custo unitário por dia de armazenagem (C_{MU}). Este corresponde à divisão do custo fixo pelo número de dias úteis de operação, conforme 6.23:

$$C_{MU} = (Dp_{Lj} + Cf_{Lj} + C_{MO-Lj}) / D_u \qquad (6.23)$$

Este é um importante parâmetro para representar as variações de custo do processo e que é aqui captado pelo modelo.

Assim, o custo final pode ser representado por 6.24:

$$C_{Arm} = C_{MU}. VM^{-1}_{Lj\,;\,Lj+1} . (C_R + C_S).D \quad . \qquad (6.24)$$

6.4. MODELAGEM DOS CUSTOS DE VENDAS PERDIDAS

Para efeito desta metodologia está sendo considerado que os custos de vendas perdidas correspondem aos ganhos não auferidos por receitas que não se concretizaram devido a deficiências logísticas. Estas perdas em geral estão associadas a questões do tipo: cortes em pedidos por faltas em estoque; cancelamentos de pedidos por um atendimento insatisfatório; devoluções devido a erros de pedidos; atrasos em entregas devido a falhas de distribuição.

Entende-se aqui por deficiências logísticas aquelas operações que não estão em conformidade com o nível de serviço proposto. Tem-se, assim, um índice de inconformidade (I_{IC}). Supõe-se que será dentro desta parcela da operação que as perdas de vendas poderão ocorrer. Sobre esta parcela então se tem um índice de vendas perdidas (I_{VP}). O índice líquido de vendas perdidas, portanto, será o resultado da multiplicação destes dois índices. Note-se que é importante esta consideração em dois segmentos, pois o nível de inconformidade tem maior

fundamentação nos dados da própria operação logística. Já o índice de vendas perdidas depende em parte de fontes externas à operação logística, como as áreas de marketing e vendas, que podem fornecer estimativas deste índice associadas a danos à imagem da empresa e conseqüentes perdas de receita, quando o nível de serviço não é atendido.

Sobre os custos que efetivamente ocorrem nestes casos, considera-se que quando uma receita não se concretiza, o que a empresa realmente perde é a margem de comercialização (*mark up*) que teria sobre aquela venda não realizada. Entendendo-se aqui que margem de comercialização é a diferença entre o preço de venda do produto e seu custo de produção ou custo de aquisição no caso de um intermediário que apenas revende produtos. Assim, o custo unitário de vendas perdidas (V_{pu}) pode ser expresso através da equação 6.25:

$$V_{pu} = I_{VP} \,.\, I_{IC} \,.\, C_{UN} \,.\, M_C \qquad (6.25)$$

onde: I_{VP} = índice de vendas perdidas

I_{IC} = índice de inconformidade

C_{UN} = custo unitário do produto

M_C = margem de comercialização

Com esta expressão está-se associando a cada unidade vendida um custo adicional correspondente às unidades não vendidas por inconformidade com o nível de serviço. Cada unidade do produto passa a carregar, assim, um custo extra (uma penalidade) pelas deficiências logísticas da operação. À medida que o nível de serviço logístico cresce, este custo de vendas perdidas tende a diminuir, melhorando assim a margem líquida do produto.

Sobre o nível de inconformidade associado a diferentes quantidades de CD's espera-se que seu comportamento possa ser modelado por uma curva como a apresentada na figura 6.3.

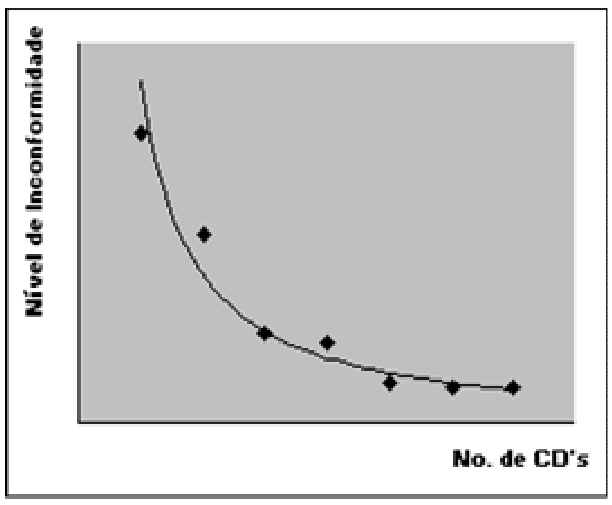

Figura 6.3 – Inconformidade vs. quantidade de CD's

A hipótese que fundamenta este "modelo de inconformidade" é a de que conforme a quantidade de CD's aumenta, estes vão se distribuindo sobre a área de estudo e, conseqüentemente, vão se aproximando dos clientes. Com isto, o processo de distribuição vai se tornando mais ágil devido à maior proximidade. O grau de atendimento ao nível de serviço proposto, portanto, deve sofrer um incremento, o que significa que o nível de inconformidade sofre uma redução gradual à medida que o número de CD's cresce. Isto deve ocorrer de forma mais intensa no início da curva, quando o aumento relativo da quantidade de CD's é maior. Já quando se dispõe de uma rede de distribuição com um número maior de instalações, o incremento de mais uma unidade não afeta significativamente o nível de serviço, logo o decréscimo do nível de inconformidade sofre uma desaceleração.

O equacionamento deste modelo de inconformidade pode ser expresso por:

$$I_{IC} = a_1.p^{-a_2} \qquad (6.26)$$

onde: p = número de CD's;

a_1 = coeficiente a ser estimado

a_2 = expoente a ser estimado

Com este índice pode-se agora determinar o volume total que não está em conformidade (VI) com o nível de serviço através de:

$$VI = I_{IC} \cdot D \tag{6.27}$$

onde: D = demanda esperada (em geral, mensal)

A estimativa final do custo de vendas perdidas (V_{pd}) é obtida associando-se as equações 6.25, 6.26 e 6.27, resultando em:

$$V_{pd} = I_{VP} \cdot a_1 . p^{-a_2} \cdot D \cdot C_{UN} \cdot M_C \tag{6.28}$$

6.5. QUANTIDADE "ÓTIMA" DE CD's

Com a definição dos modelos dos componentes de custos logísticos é possível a determinação do número *p* de CD's que minimiza os custos logísticos globais. Nas seções anteriores foram definidos modelos para os custos de estoque, armazenagem e movimentação de materiais e vendas perdidas. Estes modelos devem ser tratados agora em conjunto com a modelagem da *Fase 1* da metodologia que aborda os custos de transporte e localização.

Todos esses modelos na verdade passam se constituir em submodelos de um modelo de custo logístico global composto de quatro componentes:

- Transporte (*Fase 1*)
- Estoque (*Fase 2*)
- Armazenagem e movimentação (*Fases 1 e 2*)
- Vendas perdidas (*Fase 2*)

A montagem deste custo global é feita consolidando-se os componentes de custo para cada número p de CD's analisado.

Vista de uma forma mais ampla, a questão é, na verdade, solucionada por uma heurística constituída de um processo iterativo conforme abaixo:

Passo 1: Inicializar parâmetro. Fazer p = 1;

Passo 2: Submodelos de custos

Empregar submodelos para cada componente de custo considerando-se uma configuração de p CD's (*Fase 2*);

Passo 3: Montagem do custo global

Com os resultados da *Fase 1* e dos submodelos de custos construir o custo logístico global para p CD´s.

Passo 4: Se p=1, incrementar p de uma unidade e retornar ao passo 2. Em caso contrário, avaliar a variação de custo global ($\Delta C_{p,p+1}$) entre p e p+1 CD´s.

- Se ($\Delta C_{p,p+1}$)>0 parar. A solução foi encontrada.
  Quantidade "ótima" de CD´s = p.
- Em caso contrário, incrementar p de uma unidade e retornar ao passo 2.

Definido o valor "ótimo" p, os CD´s deverão se localizar nos locais candidatos já definidos para p CD*´s na Fase 1.*

A figura 6.4 mostra uma representação gráfica de uma possível evolução deste processo.

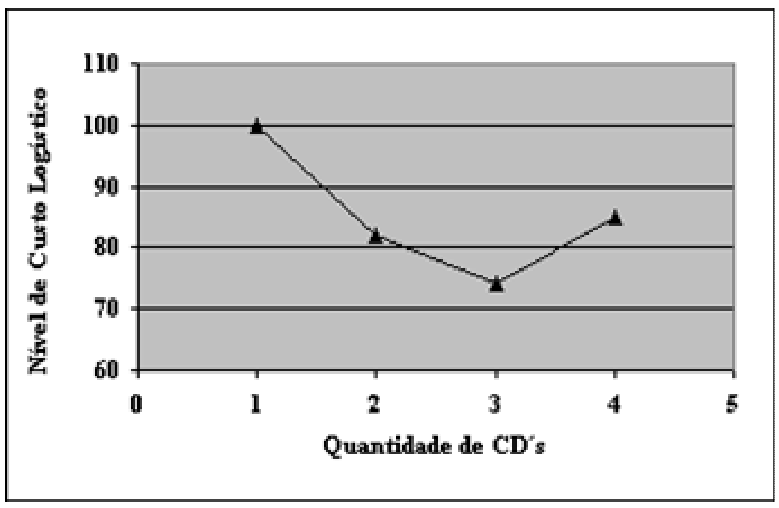

Figura 6.4 -– Possível evolução de custos logísticos com a quantidade de CD´s

Uma alternativa para esta heurística seria computar o custo logístico global para todas as quantidades de CD's que se deseje analisar. A quantidade "ótima" de CD's seria aquela que fornecesse o menor custo logístico global.

7. APLICAÇÃO PRÁTICA

7.1. SITUAÇÃO EM ESTUDO

A aplicação da metodologia proposta foi baseada em um caso real de uma cadeia de lojas de departamento com duzentos e vinte e nove estabelecimentos situados na região sul do Brasil (Vallim, 1997).

Estas lojas estão distribuídas pelos estados do Rio Grande do Sul, Santa Catarina e Paraná (figura 7.1), movimentando um total de 45.651 m^3 por mês e com uma frota de noventa e oito veículos. Na operação analisada a frota roda uma média de 811.574 quilômetros mensais, com 1.088 viagens por mês. O conjunto de pontos com suas coordenadas (em Km) e demandas mensais (em m^3) é apresentado na tabela 7.1. Para a definição das coordenadas foi adotada uma origem do sistema cartesiano localizada no canto sudoeste de um mapa da região sul do país.

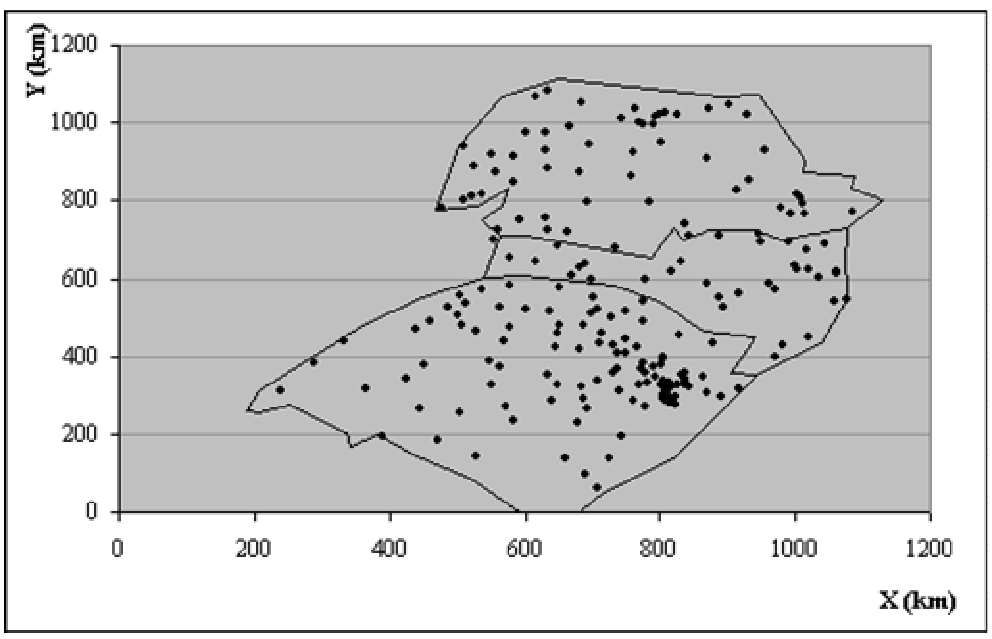

Figura 7.1 – Pontos de Demanda

Tabela 7.1 - Pontos de Demandas (parte 1)

Seq	Loja	X (km)	Y (km)	Volume (m^3)
1	1	790,0	372,5	439,0
2	33	790,0	372,5	222,6
3	2	802,5	377,5	762,0
4	117	802,5	380,0	499,2
5	222	802,5	380,0	273,2
6	31	777,5	360,0	116,5
7	6	775,0	382,5	414,7
8	8	772,5	367,5	126,2
9	17	805,0	392,5	93,8
10	12	792,5	345,0	93,5
11	3	805,0	400,0	106,2
12	42	805,0	335,0	207,2
13	19	782,5	332,5	227,0
14	43	812,5	332,5	95,5
15	106	802,5	327,5	226,1
16	26	832,5	352,5	294,2
17	13	837,5	355,0	336,0
18	15	770,0	325,0	398,8
19	21	737,5	370,0	127,9
20	11	767,5	422,5	111,1
21	5	747,5	410,0	116,8
22	37	825,0	327,5	358,3
23	36	817,5	322,5	401,8
24	38	837,5	340,0	177,5
25	34	810,0	317,5	739,4
26	94	810,0	317,5	202,8
27	200	810,0	317,5	531,1
28	35	732,5	360,0	101,3
29	18	837,5	332,5	225,3
30	54	807,5	312,5	1.077,4
31	20	737,5	407,5	217,6
32	170	837,5	327,5	328,2
33	61	810,0	307,5	404,3
34	28	842,5	322,5	353,4
35	175	805,0	300,0	353,3
36	16	862,5	345,0	168,5
37	30	740,0	312,5	256,3
38	224	750,0	442,5	89,9
39	24	732,5	430,0	197,1
40	29	805,0	292,5	643,3
41	60	815,0	292,5	352,9
42	201	815,0	292,5	216,0
43	55	822,5	295,0	317,2
44	202	822,5	295,0	254,2
45	203	822,5	295,0	384,9
46	66	807,5	287,5	643,1
47	74	807,5	287,5	342,0
48	91	807,5	287,5	402,7
49	105	807,5	287,5	248,7
50	148	807,5	287,5	458,4
51	182	807,5	287,5	240,9
52	4	827,5	452,5	470,4
53	204	970,0	397,5	192,0
54	184	982,5	430,0	459,9
55	161	1.020,0	450,0	235,6
56	149	1.075,0	545,0	331,9
57	220	1.075,0	545,0	79,9
58	155	1.057,5	542,5	165,1
59	10	877,5	432,5	219,5
60	186	1.062,5	615,0	12,2
61	123	1.060,0	620,0	37,3
62	152	1.035,0	602,5	34,8
63	154	970,0	570,0	87,7
64	114	1.020,0	625,0	42,6
65	194	1.020,0	625,0	37,3
66	208	1.002,5	625,0	16,5
67	101	1.045,0	687,5	182,4
68	103	1.045,0	687,5	192,5
69	223	1.045,0	687,5	384,7
70	205	1.000,0	632,5	23,7
71	119	960,0	587,5	165,8
72	127	1.017,5	672,5	282,7
73	196	1.085,0	770,0	16,8
74	195	917,5	562,5	114,5
75	96	892,5	527,5	384,7
76	126	990,0	695,0	301,0
77	138	1.015,0	767,5	28,8
78	218	887,5	550,0	138,9
79	107	1.012,5	792,5	23,4
80	108	1.012,5	792,5	22,0

Tabela 7.1 - Pontos de Demandas (parte 2)

Seq	Loja	X (km)	Y (km)	Volume (m^3)
81	124	1.012,5	792,5	11,3
82	131	1.012,5	792,5	13,2
83	139	1.012,5	792,5	16,6
84	147	1.012,5	792,5	33,5
85	151	1.012,5	792,5	10,6
86	168	995,0	767,5	12,4
87	193	1.007,5	805,0	8,5
88	211	950,0	692,5	171,7
89	163	1.002,5	815,0	13,9
90	181	980,0	782,5	14,0
91	133	947,5	712,5	178,4
92	177	870,0	587,5	156,4
93	169	955,0	930,0	13,7
94	197	932,5	855,0	12,2
95	134	887,5	710,0	219,2
96	136	915,0	825,0	19,8
97	189	930,0	1.022,5	29,7
98	190	902,5	1.045,0	12,8
99	164	842,5	707,5	339,5
100	142	870,0	910,0	19,0
101	97	830,0	645,0	177,2
102	212	837,5	740,0	162,9
103	176	872,5	1.037,5	13,9
104	99	817,5	620,0	194,0
105	132	825,0	1.022,5	26,7
106	180	825,0	1.022,5	16,1
107	199	807,5	1.027,5	29,4
108	166	802,5	950,0	2,9
109	183	800,0	1.022,5	23,2
110	150	792,5	1.015,0	24,3
111	141	790,0	997,5	22,5
112	125	785,0	795,0	293,9
113	187	775,0	995,0	2,4
114	162	770,0	1.000,0	12,2
115	159	765,0	1.035,0	15,2
116	73	760,0	922,5	16,2
117	115	777,5	600,0	230,8
118	25	757,5	865,0	25,8
119	145	742,5	1.012,5	35,0
120	7	775,0	540,0	62,2

Seq	Loja	X (km)	Y (km)	Volume (m^3)
121	9	775,0	490,0	185,0
122	156	685,0	1.052,5	18,9
123	143	695,0	945,0	20,3
124	185	735,0	677,5	243,0
125	174	667,5	990,0	11,8
126	217	635,0	1.085,0	8,9
127	165	680,0	875,0	1,9
128	100	692,5	795,0	137,9
129	171	617,5	1.067,5	16,0
130	206	630,0	975,0	7,7
131	14	750,0	515,0	108,6
132	157	630,0	927,5	10,6
133	219	632,5	885,0	3,6
134	140	600,0	977,5	25,1
135	116	662,5	720,0	215,6
136	104	690,0	637,5	181,0
137	22	582,5	912,5	46,4
138	210	700,0	600,0	103,3
139	207	630,0	757,5	89,1
140	27	680,0	627,5	51,1
141	128	582,5	847,5	114,5
142	130	550,0	920,0	36,9
143	111	632,5	727,5	237,9
144	137	650,0	685,0	52,3
145	118	557,5	872,5	61,8
146	64	727,5	500,0	113,3
147	63	702,5	550,0	426,9
148	178	510,0	940,0	52,3
149	89	670,0	610,0	243,3
150	109	525,0	890,0	57,1
151	110	592,5	752,5	93,5
152	39	707,5	522,5	299,8
153	153	537,5	817,5	50,5
154	135	522,5	810,0	45,0
155	129	617,5	645,0	72,4
156	225	700,0	512,5	85,0
157	173	560,0	727,5	79,0
158	167	510,0	802,5	29,7
159	179	652,5	577,5	108,5
160	172	555,0	700,0	180,6

Tabela 7.1 - Pontos de Demandas (parte 3)

Seq	Loja	X (km)	Y (km)	Volume (m^3)
161	93	577,5	655,0	88,1
162	209	477,5	782,5	54,6
163	59	712,5	460,0	157,9
164	72	687,5	480,0	436,2
165	113	577,5	580,0	155,5
166	221	637,5	517,5	156,7
167	75	537,5	570,0	146,4
168	82	652,5	480,0	191,4
169	112	600,0	520,0	123,2
170	23	710,0	432,5	87,3
171	40	562,5	525,0	100,6
172	70	505,0	555,0	102,8
173	229	647,5	460,0	71,0
174	47	512,5	537,5	121,1
175	56	487,5	527,5	225,2
176	160	577,5	475,0	98,3
177	45	500,0	507,5	86,8
178	144	680,0	417,5	130,3
179	46	507,5	480,0	305,3
180	83	645,0	425,0	130,0
181	51	460,0	490,0	82,8
182	53	527,5	465,0	283,1
183	65	570,0	440,0	346,3
184	52	440,0	467,5	283,1
185	76	332,5	437,5	242,0
186	84	547,5	390,0	142,3
187	57	452,5	380,0	200,6
188	86	290,0	382,5	252,6
189	216	562,5	375,0	70,5
190	227	425,0	340,0	93,6
191	95	237,5	312,5	348,5
192	188	237,5	312,5	228,5
193	121	365,0	315,0	346,1
194	228	632,5	350,0	139,5
195	62	550,0	325,0	454,9
196	88	550,0	325,0	407,5
197	120	550,0	325,0	504,1
198	90	445,0	265,0	196,2
199	77	650,0	325,0	242,9
200	92	505,0	257,5	320,1
201	192	390,00	192,50	403,9
202	32	707,5	335,0	281,8
203	80	572,5	270,0	159,9
204	71	685,0	320,0	318,2
205	48	640,0	285,0	279,1
206	214	472,5	182,5	140,0
207	79	582,5	235,0	158,6
208	49	687,5	290,0	347,1
209	98	527,5	145,0	362,9
210	122	527,5	145,0	72,6
211	226	692,5	267,5	181,3
212	85	677,5	230,0	158,3
213	78	660,0	137,5	185,4
214	67	690,0	95,0	581,4
215	87	690,0	95,0	504,7
216	215	760,0	287,5	213,2
217	68	725,0	137,5	198,6
218	41	742,5	195,0	400,2
219	81	707,5	60,0	617,8
220	44	777,5	272,5	495,9
221	69	815,0	282,5	584,8
222	198	815,0	282,5	284,0
223	58	822,5	277,5	386,0
224	102	822,5	277,5	375,5
225	158	815,0	282,5	399,0
226	191	822,5	277,5	1.375,0
227	50	870,0	305,0	298,0
228	146	890,0	295,0	272,4
229	213	917,5	315,0	181,6

7.2 FASE 1: LOCALIZAÇÃO DE CD'S

O processo de análise se inicia pela *Fase 1,* que se apóia em um processo heurístico em dois estágios (ver seção 4.3):

Estágio 1: Heurística de *Clusterização*;

Estágio 2: Heurística de Seleção dos "melhores" Candidatos

Estágio 2a: Heurística do Raio Máximo

Estágio 2b: Heurística do Ângulo Máximo

7.2.1 Heurística de *"Clusterização"* Aplicada

Conforme visto na seção 4.3, a heurística de *Clusterização* parte de um número pré-definido de clusters que é função do nível de serviço desejado e de uma restrição operacional relativa à área dos CD's. O desenvolvimento deste processo requer ainda a adoção de alguns parâmetros e o uso das equações apresentadas na seção 4.3.

O passo 1 do processo requer a definição do tempo máximo admissível do CD a um ponto qualquer de demanda para um nível de confiança desejado. Foi adotado um tempo máximo de quatro horas a um nível de confiança (1- α) de 95%. Assim, tem-se $T_{0,05}$ = 4 horas e adotando-se uma velocidade de cruzeiro de 60 km/h obtém-se a distância correspondente a $D_{0,05}$=240km.

A partir de uma varredura dos pontos da região de estudo e calculando-se distâncias entre pontos, foi obtida uma estimativa de 1,032 para o coeficiente de variação (CV) das distâncias. Com o nível de confiança de 95% tem-se $Z_{0,05}$=1,645 e, agora, utilizando-se a equação 4.6, obtém-se a distância média (D_m):

$$D_m = 240 / (1 + 1{,}032.\ 1{,}645) = 89{,}0 \text{ km}$$

No passo 2 deve-se determinar o volume por CD. Por questões de viabilidade operacional da empresa estudada, foi estabelecida uma área média desejada ($A_{média}$) de 2.000m^2. Os atuais armazéns da empresa têm pé direito médio de 7,45m. Foi adotado então um pé direito (h) de 7,5m. O aproveitamento de área construída (η_A) na atual operação é de 55,6%. Adotou-se então aproveitamento de 55% e com o uso da equação 4.7 obtém-se o volume médio ($V_{médio}$) por CD:

$$V_{médio} = 2000.7{,}5\ .0{,}55 = 8.250\text{m}^3$$

O volume útil de armazenagem é obtido aplicando-se a $V_{médio}$ um índice de aproveitamento. Não havia dados disponíveis sobre este coeficiente. Foi adotado então o valor de 76%, que é a média do setor varejista, conforme pesquisa efetuada pelo Centro de Estudos Logísticos (CEL/COPPEAD, 2001). Com isto, tem-se:

$$V_{ÚTIL} = 8250.\ 0{,}76 = 6.270\text{m}^3$$

O volume mensal atendido por CD (V_{CD}) será este valor multiplicado pelo giro médio mensal. O giro atual da empresa é de 1,05 e foi mantido.

Desta forma, tem-se: $V_{CD} = 6270.\ 1{,}05 = 6.584\text{m3}$

No passo 3 determina-se o momento de transporte por CD, a partir do volume atendido e da distância média aos pontos de demanda. Assim, tem-se:

$$TKM_{CD} = 6584.\ 89{,}0 = 585.869 \text{ m}^3\text{.Km}$$

O número de CD's é obtido através da equação 4.11, em que se divide o momento de transporte total da área de estudo (TKM_{Total}) pelo momento de transporte por CD (TKM_{CD}). A partir da varredura dos pontos da área de estudo obteve-se TKM_{Total}=3.790.481 m^3.Km. Assim, tem-se:

$$N_{CD} = 3.790.481 / 585.869 = 6{,}47$$

Adota-se então o número de sete CD's e mantém-se este mesmo valor para o número de zonas (N_{zonas} = 7). A heurística de *clusterização* deverá buscar, assim, a formação de sete clusters.

Seguindo-se com a heurística foi determinado um CD central, empregando-se o método do centro de gravidade. Os dados para a aplicação do método foram extraídos da própria operação (Vallim, 1997). Os custos de transferência e distribuição são apresentados abaixo (tabela 7.2)

Tabela 7.2 – Parâmetros de custos de transferência e distribuição

Tipo de Operação	Tipo de Veículo	Capacidade Bruta (m^3)	Índice de Aproveitamento %	Capacidade Líquida (m^3)	Km no trimestre	R$/ no Trimestre	R$/km	**R$/$m^3$.km**
Transferência	Carreta	101	90%	90,9	2.782.610	3.069.100,00	1,10	**0,0121**
Distribuição	Leve	44	75%	33,0	374.572	496.557,00	1,33	**0,0402**

As fontes de produção estavam situadas nas regiões das cidades de Porto Alegre e Curitiba e no interior do Rio Grande do Sul, conforme tabela 7.3.

Tabela 7.3 – Fontes de produção

Fonte de Produção	**X (km)**	**Y (km)**	**Oferta (m3)**
Região de Porto Alegre	807,5	287,5	20.543
Região de Curitiba	1.012,5	792,5	15.978
Interior RGS	687,5	480,0	9.130
		Total. :	45.651

Com estes dados e aplicando-se o método do centro de gravidade aos pontos de demanda obteve-se:

. Coordenada X do CD central = 791,6

. Coordenada Y do CD central = 373,0

A partir deste ponto central foram definidos anéis, conforme estabelecidos na heurística.

Para a definição dos anéis determina-se inicialmente a distância média (μ) e o desvio padrão (σ) dos pontos de demanda ao CD central. Estes parâmetros

resultaram em: μ = 291,5 km e σ = 197,5 km. Com isto tem-se a definição de três anéis conforme mostrado na tabela 7.4:

Tabela 7.4 – Definição de anéis

1º anel:	até	$\mu - \sigma$ ➔	até 94,0 km
2º anel:	de	$\mu - \sigma$ ➔	de 94,0 km
	a	$\mu + \sigma$ ➔	a 489,0 km
3º anel:	após	$\mu + \sigma$ ➔	após 489,0 km

Passa-se a seguir à montagem dos *clusters* finais, que serão setores dentro dos anéis. Deve-se, primeiramente, determinar o número de clusters por anel. O anel 1 que se encontra no entorno do CD central terá apenas uma zona.

Deve-se manter uma só zona para servir de base para a montagem das zonas dos outros anéis. A partir do momento de transporte desta zona inicial define-se a quantidade de zonas dos demais anéis. Esta quantidade de zonas deve atender ao número máximo estabelecido em função de nível de serviço e restrição de área. O número de clusters em cada anel será o resultado da divisão do momento de transporte do anel pelo momento de transporte do anel 1. Os momentos de transporte dos três anéis foram computados e o número de clusters foi definido, conforme apresentados na tabela 7.5.

Tabela 7.5 – Momento de transporte e *clusters* por anel

ANEL	m^3.km / Anel	m^3.km / Cluster	No. de Zonas			m^3.km % Cluster	m^3.km % Anel
			Definido / Calculado	No. Inteiro	Valor Final		
1	437.895	437.895	**1,0**	1,0	1	13,6%	13,6%
2	2.701.204	450.200	6,2	6,0	6	14,0%	84,2%
3	70.660	70.660	0,2	0,0	1	2,2%	2,2%

Conforme a seção 4.3.1.4, a identificação de um anel desbalanceado é feita identificando-se o anel com menor média de momento de transporte por *cluster* e

que aqui é o anel 3 (figura 7.2). Este anel se encontra fora do intervalo de confiança de 95% da média dos demais anéis, conforme demonstram os cálculos abaixo:

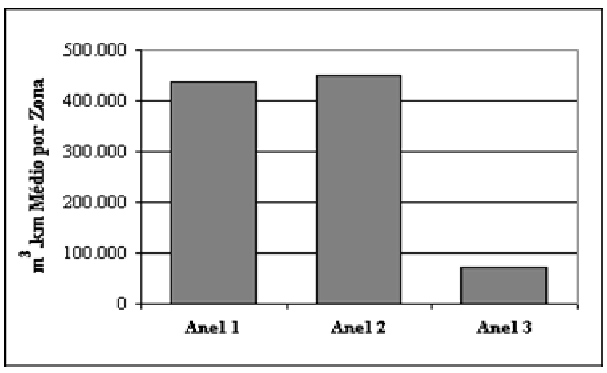

Figura 7.2 – Anéis desbalanceados

A partir dos momentos de transporte dos anéis tem-se:

Média do momento de transporte = 444.048 m^3.km

Desvio padrão = 8.701 m^3.km

O intervalo ao nível 1-α de confiança é obtido por (Steel e Torrie, 1980):

$$IC = \mu_{MT} \pm Z_{\alpha/2} \,.\, \sigma_{MT} / (n)^{1/2} \qquad (7.1)$$

onde: μ_{MT} = média do momento de transporte;

$Z_{\alpha/2}$. = valor crítico da distribuição normal reduzida ao nível 1- $\alpha/2$

σ_{MT} = desvio padrão do momento de transporte

n = número de observações (anéis considerados).

Para um intervalo ao nível de 95% de confiança, tem-se: $Z_{0,025}$ = 1,96. Logo, os limites do intervalo são:

Limite Inferior = 444.048 - 1,96. 8.701/$(2)^{1/2}$ = 431.990 m^3.km

Limite Superior = 444.048 + 1,96. 8.701/$(2)^{1/2}$ = 456.107 m^3.km

Portanto, os anéis se encontram desbalanceados e o processo precisa ser revisto.

Aplicando-se agora o critério definido na seção 4.3.1.4, os anéis 1 e 2 deverão ser reconstruídos, incorporando o anel 3. Para isto, identifica-se a faixa de distância em que se encontram 95% dos pontos do anel desbalanceado. Divide-se esta distância pelo número total de zonas e agrega-se um múltiplo desta faixa de distância a cada anel proporcionalmente ao número de zonas de cada um.

. faixa de 95% das Distâncias: 198,1km

. distância a acrescer ao Anel 1: 24,8km

. distância a acrescer ao Anel 2: 173,3km

Com estes novos anéis reaplica-se a heurística.

Os dados obtidos para esta nova configuração são apresentados na tabela 7.6.

Tabela 7.6 – Momento de transporte e clusters por anel – Nova configuração

ANEL	m^3.km / Anel	m^3.km / Cluster	No. de Zonas			m^3.km % Cluster	m^3.km % Anel
			Definido/ Calculado	No. Inteiro	Valor Final		
1	552.666	552.666	**1,0**	1,0	1	15,1%	15,1%
2	3.095.992	515.999	5,6	6,0	6	14,1%	84,9%

Como se observa agora, os anéis estão balanceados em termos de momento de transporte por *cluster* e pode-se então prosseguir na heurística.

Dentro dos anéis foram construídos setores conforme os procedimentos da heurística, chegando-se a uma configuração final para os *clusters* (figura 7.3). Foram estabelecidos sete clusters, passando cada um agora a se constituir em uma zona de demanda. Os resultados finais do processo são apresentados na tabela 7.7.

Note-se que o grau de homogeneidade obtido foi de 92,2 %. Este conjunto de clusters, entretanto, não se constitui ainda no zoneamento final, pois o resultado da heurística é submetido em seguida à meta-heurística *simulated annealing*. .na tentativa de melhorar este grau de homogeneidade do zoneamento.

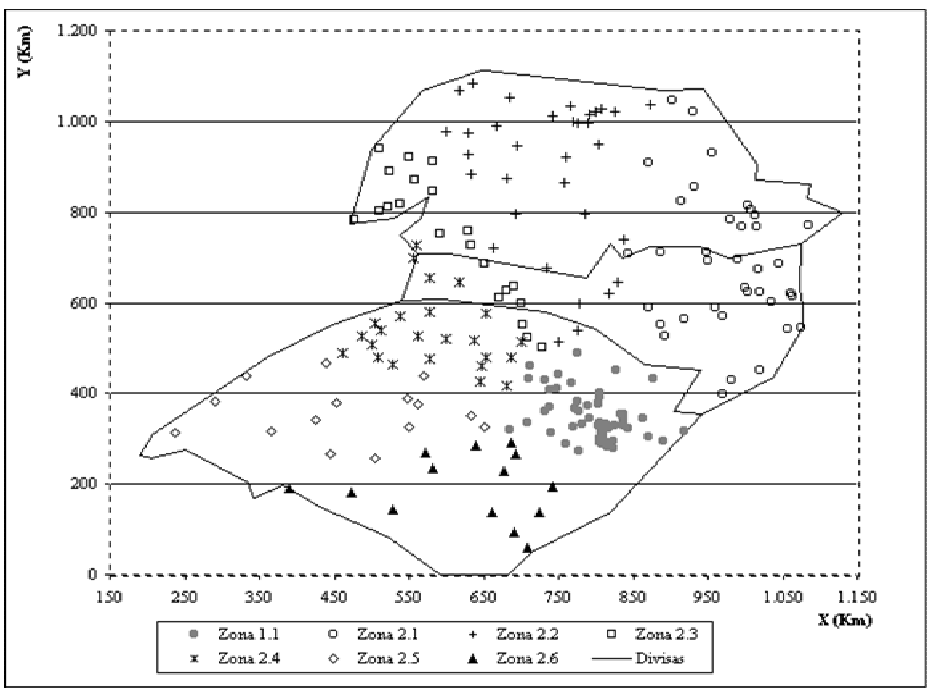

Figura 7.3 - Heurística de *clusterização* - Proposta de zoneamento

Tabela 7.7 – Resultados finais - *clusterização*

Anel	Zona	m^3. km
1	1.1	552.666
2	2.1	511.335
2	2.2	501.166
2	2.3	522.700
2	2.4	453.866
2	2.5	502.520
2	2.6	429.808
	Total	3.474.060
	Média	496.294
	Desvio	38.516
	GH	92,2%

7.2.2 Meta-Heurística SA na Montagem de Zonas

A SA recebeu inicialmente informações referentes ao zoneamento gerado e às coordenadas, demandas e ângulos de cada ponto com o eixo horizontal. O início do processo se dá com a definição dos parâmetros da SA que, conforme já discutido na seção 5.1, são os seguintes:

Temperatura inicial (T_o)

Número máximo de iterações em cada temperatura (p)

Número máximo de sucessos em cada temperatura (L);

Número total de iterações (m);

Método de resfriamento.

Os parâmetros foram analisados para a aplicação da SA ao anel 2, já que o anel 1 constitui-se de uma única zona. A discussão e definição desses parâmetros são apresentadas a seguir.

a) Temperatura inicial (T_o)

Esta foi definida a partir da equação 5.8. O valor de ΔE_m^+ é obtido através de resultados experimentais, conforme indicado por Viana (1998) e Johnson (1989). O cálculo final de ΔE_m^+ resultou em $\Delta E_m^+ = 0{,}85$.

Portanto, tem-se $T_o = 3{,}8$. Foi adotado então $T_o = 4{,}0$.

b) Número máximo de iterações em cada temperatura (p)

Considerando-se o caso do anel 2, que é o maior anel, tem-se um número definido de seis zonas, com um total de 160 pontos. Mantendo-se um mínimo de dez pontos

por zona, restam 100 pontos a cada iteração que podem ser submetidos a uma alteração, de forma a se proceder a um rearranjo do sistema.

Conforme a diretriz apresentada na seção 5.1, ter-se-ia então p = 100.100 ou p=10.000 iterações a cada temperatura.

Os resultados experimentais mostraram que este número estava bem dimensionado.

c) Número máximo de sucessos em cada temperatura (L)

Por decorrência da definição de p=10.000 iterações, e adotando-se L igual a 10% de p (seção 5.1), tem-se: L=1.000 iterações.

d) Método de resfriamento

Conforme as observações de Johnson(1989) e Viana(1998), que sugerem o uso de uma taxa geométrica de resfriamento variando entre 0,90 e 0,99, foram desenvolvidos alguns testes para verificar a mais adequada para o presente estudo, chegando-se ao valor de Rc =0,98 como sendo aquele que apresentou o melhor grau de convergência da solução, dentro de um número de iterações adequado aos demais parâmetros definidos para o estudo.

d) Número total de iterações (m)

Resultados experimentais mostraram que o valor máximo da função objetivo definida (GH - grau de homogeneidade) apresentava oscilação até uma temperatura em torno de 1,2. Neste ponto iniciava-se uma tendência de estabilização a um patamar de GH em torno de 60%. Em seguida, quando a temperatura atinge algo próximo de 0,5, há uma nova subida com GH se aproximando de 80%. A convergência final da solução efetivamente ocorre para temperaturas abaixo de 0,3.

Isto é observado na figura 7.4a, que representa o anel 2. No gráfico se destaca o ponto de início do processo, que parte da solução gerada pela heurística de *clusterização*. A figura apresenta a variação de GH máximo à medida que a temperatura vai decrescendo. Em temperaturas mais baixas a diferença absoluta entre duas temperaturas seguidas é muito pequena, o que leva a um agrupamento dos pontos nestas faixas de temperatura. Assim, um detalhamento da fase final do processo é apresentado na figura 7.4b, que mostra com mais clareza a convergência da solução.

Adotando-se uma abordagem conservadora, foi definido um valor mínimo de temperatura (T_m) de cerca de 0,25, o que garantiria com segurança a convergência da solução.

Com o valor de T_m definido e utilizando-se a equação 5.6, determina-se o número <u>m</u> máximo de iterações:

$$m = \log(0{,}25 / 4{,}0) / \log(0{,}98) \rightarrow \mathbf{m = 137{,}2}$$

Foi adotado então o número máximo de 140 iterações, resultando um valor final de temperatura de 0,2413.

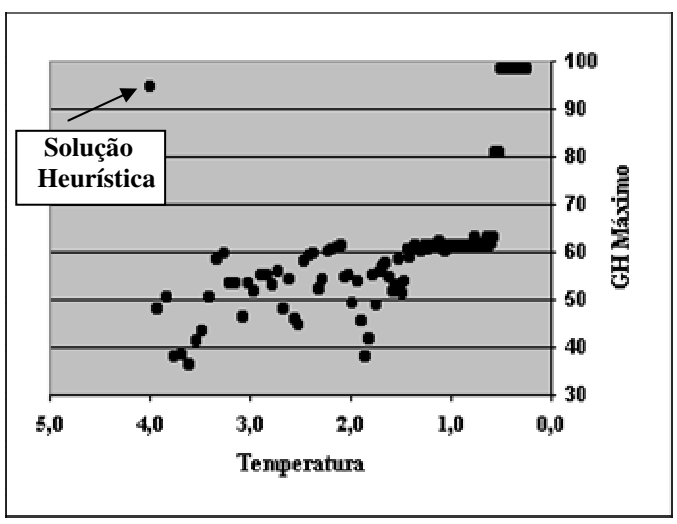

Figura 7.4a -: Anel 2: GH Máximo evoluindo com o decréscimo de temperatura

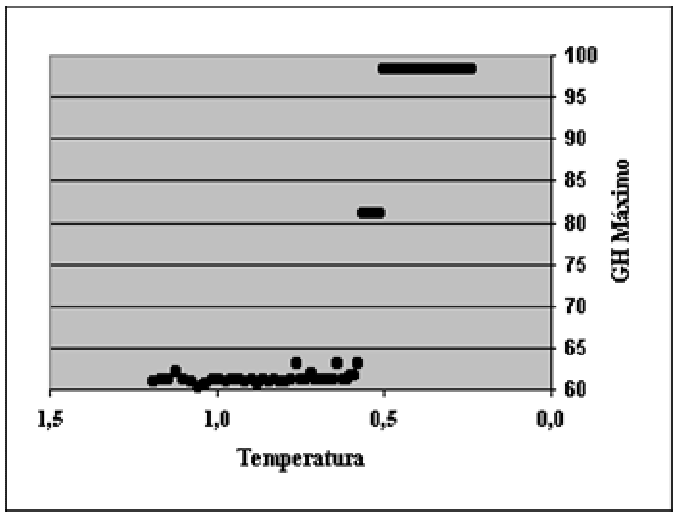

Figura 7.4b -: Anel 2: Evolução de GH Máximo
para temperaturas abaixo de 1,2

Definidos esses parâmetros, a meta-heurística foi rodada em um computador com processador AMD-K6 de 450mhz e com 156mb de memória RAM. O tempo de processamento ficou em torno de 36 minutos. O gráfico do processo de evolução do GH Médio para o anel 2 a cada temperatura é apresentado na figura 7.5. Percebe-se por esta figura e pela figura 7.6 a eficácia do processo de SA.

Na figura 7.5 o GH Médio cresce consistentemente ao longo de todo o processo e

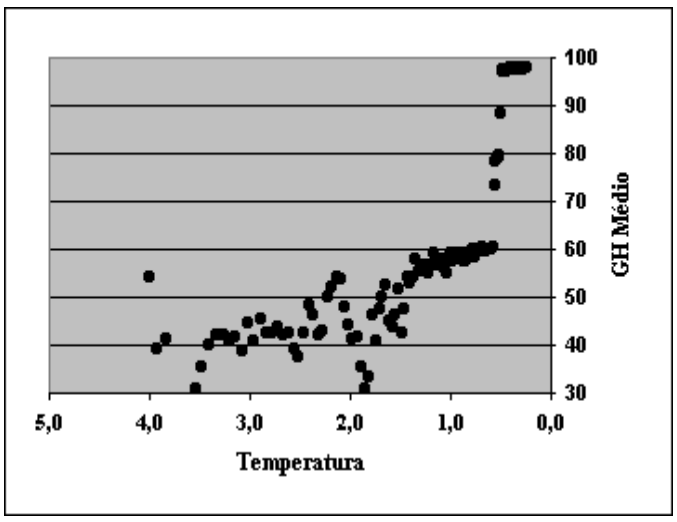

Figura 7.5 – GH Médio vs. Temperatura

na figura 7.6 visualiza-se o GH Máximo de cada temperatura, que converge de forma definitiva para um valor acima de todos os outros quando o processo caminha para seu final.

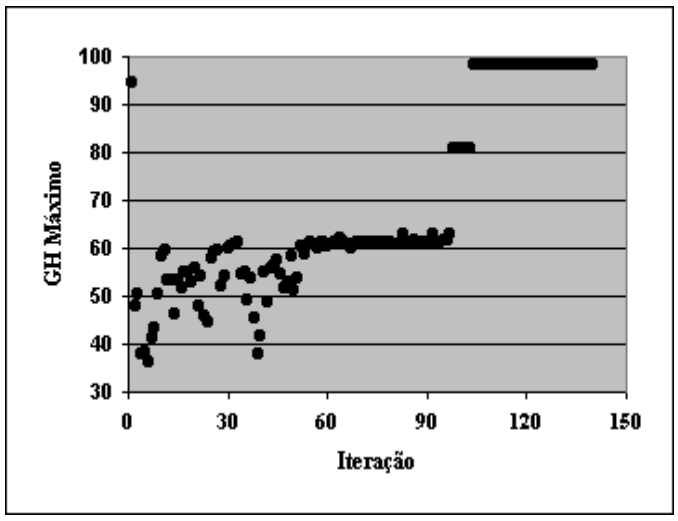

Figura 7.6 – Convergência da solução "ótima"
Clusterização

O processo foi aplicado dezenas de vezes ao conjunto de dados. As soluções geradas pela SA convergiram invariavelmente para a "melhor" solução, o que pode ser considerado um resultado bem satisfatório, já que segundo Viana (1998), se admite que em cerca de 10% dos caso não haja convergência.

A solução "ótima" gerada é apresentada na tabela 7.8. Pode-se observar que como resultado final a meta-heurística conseguiu uma evolução no GH global da área de estudo, apesar de pequena (+1,2%), alterando a alocação de *clusters* de seis pontos de demanda (2,6% dos pontos). Os resultados demonstram a complementaridade dos métodos, mas representam também uma indicação da eficácia da heurística de *clusterização,* que por si só já gera resultados bem próximos da solução "ótima".

A figura 7.7 mostra o zoneamento final gerado pela SA.

Tabela 7.8 – Resultados finais SA

Anel	Zona	m^3	m^3. km
1	1.1	22.561	552.666
2	2.1	5.282	508.555
2	2.2	2.020	499.331
2	2.3	2.656	494.183
2	2.4	3.195	431.418
2	2.5	4.865	488.770
2	2.6	5.072	483.166
	Total	45.651	3.458.089
		Média	494.013
		Desvio	33.259
		GH	93,3%

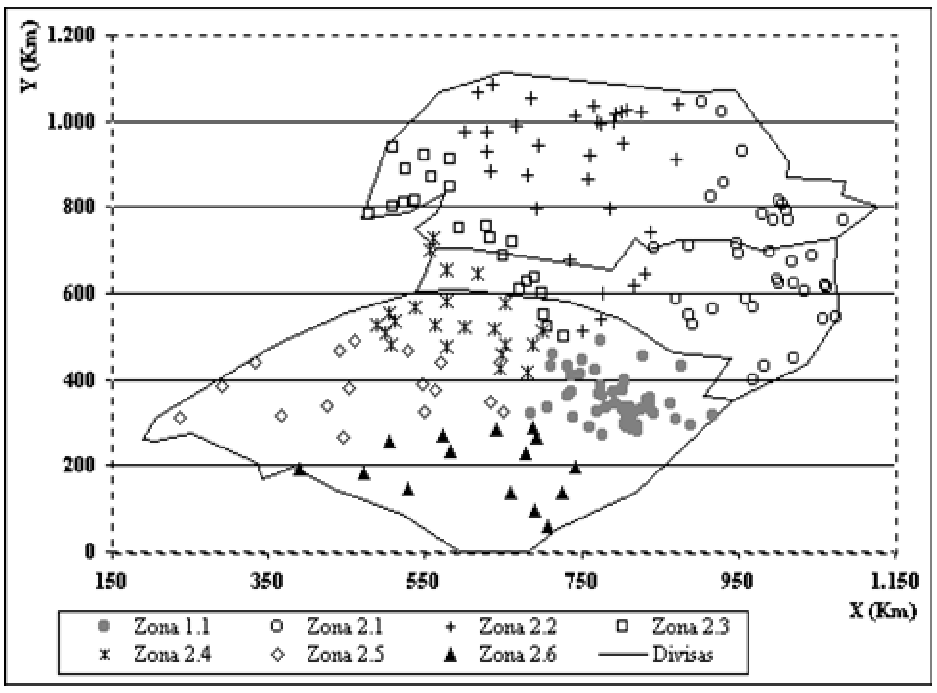

Figura 7.7 - *Simulated Annealing* - Zoneamento final

7.2.3 Seleção dos “Melhores” Candidatos

A partir do zoneamento definido seguiu-se para o estágio No. 2 da heurística correspondente à seleção dos “melhores” candidatos. Este estágio começa pela heurística do Raio Máximo, que necessita inicialmente da definição das MPT´s em cada zona. As MPT´s são determinadas através do método do centro de gravidade apresentado no capítulo 2 (equações 2.1a a 2.2b). A aplicação deste método aos dados do problema resultou no conjunto de MPT´s apresentado na figura 7.8.

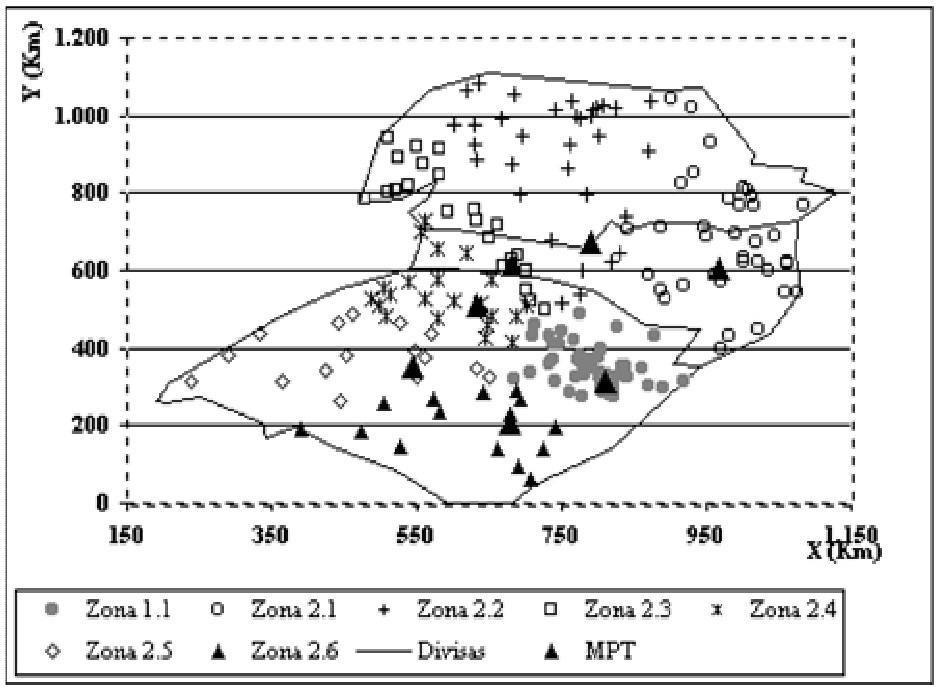

Figura 7.8 - Zoneamento e MPT’s por zona.

Em seguida são determinadas as MPT’s Meio, que são posições intermediárias entre as MPT’s adjacentes. As figuras 7.9a e 7.9b apresentam uma visão geral das

MPT's e MPT's Meio, podendo-se verificar a sua formação característica, aproximadamente circular.

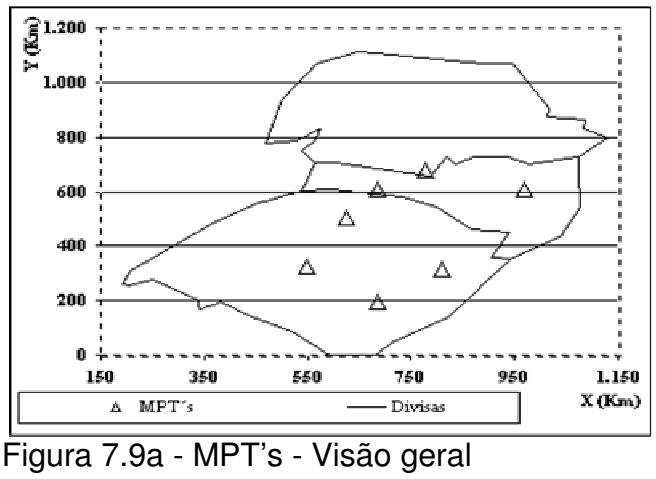

Figura 7.9a - MPT's - Visão geral

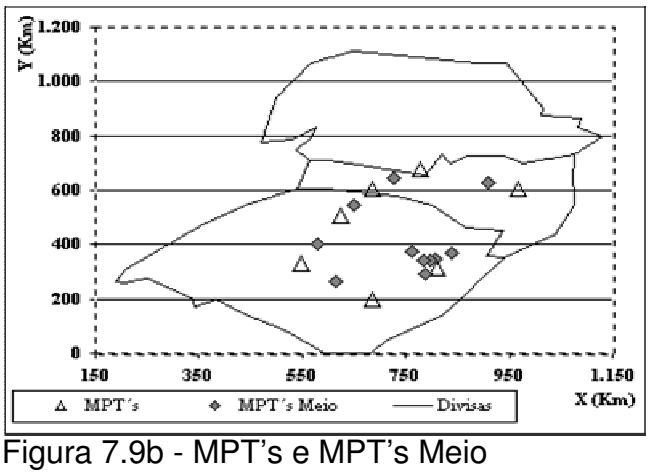

Figura 7.9b - MPT's e MPT's Meio

Na seqüência identifica-se, no entorno das MPT's e das MPT's Meio, o conjunto dos "melhores" candidatos a CD's Isto é feito com o estabelecimento de um raio de busca (o raio máximo) de pontos na rede no entorno de cada MPT e MPT Meio. Conforme descrito no capítulo 4, este raio é baseado na distribuição de probabilidades das distâncias dos pontos de demanda às MPT´s. A partir da distribuição estabelece-se um limite inferior de distância a um nível 1- β_o de confiança para cada posição teórica. Neste caso, foram adotados os níveis de 95%

e 90% de confiança (tabela 7.9). Um ponto encontrado no primeiro nível já se constitui em um candidato a CD. Se o ponto estiver no segundo nível deverá passar pela heurística do ângulo máximo. Convém lembrar, que conforme os procedimentos de cálculo definidos no capítulo 4, se um limite for negativo assume-se limite zero e que seu valor máximo será a metade da distância entre as MPT's adjacentes.

Adotando-se esses critérios foram calculados os raios máximos para os dois níveis de confiança adotados. Os resultados para o caso das MPT's são apresentados na tabela 7.10.

Tabela 7.9 – Valores críticos de Z ao nível 1-β_o. de confiança

Nível de Confiança ➔	95%	90%
Valor de β_o. ➔	β_o. = 0,05	β_o. = 0,10
Valor de Z ➔	1,6449	1,2816

Fonte: Principles and procedures of statistics. Steel, R.G.D. e Torrie, J.H. -1980

Tabela 7.10 – Raios Máximos de busca de candidatos no entorno de MPT's

MPT	Distância Média a MPT (km)	Desvio Padrão (km)	Raio 1 95% Confiança (Km)	Raio 2 90% Confiança (Km)
1.1	57,7	44,6	0,0*	0,5
2.1	153,1	90,1	4,9	37,7
2.2	257,8	114,6	69,3	101,5**
2.3	183,4	120,5	0,0*	29,0
2.4	103,9	53,2	16,5	35,8
2.5	130,7	99,1	0,0*	3,6
2.6	119,6	65,7	11,6	35,5

** Limite inferior era negativo

** Metade da distância entre as MPT's adjacentes

Uma visualização das áreas de busca ao redor de cada MPT é apresentada na figura 7.10.

A busca naquelas áreas gera um conjunto final de "melhores" candidatos no entorno de todas as posições teóricas. As figuras 7.11a e 7.11b apresentam este conjunto para as MPT´s e MPT´s Meio. Ao final, na figura 7.12, tem-se o conjunto completo

de “melhores” candidatos, que acabam circundando praticamente toda a região de estudo.

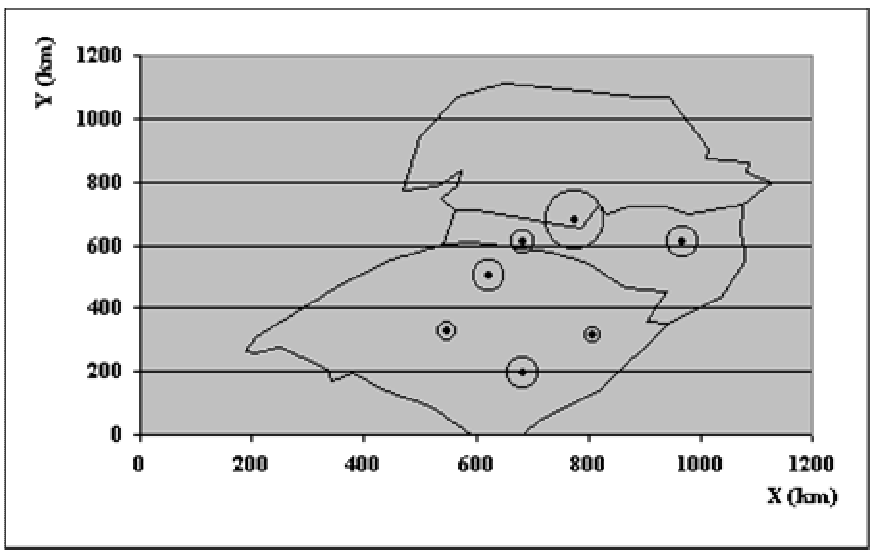

Figura 7.10 – Raios máximos de busca no entorno das MPT’s

É importante salientar a relevância de se ter esta configuração de “melhores” candidatos com este formato quase circular. Este tipo de formação abre a possibilidade de se ter como soluções viáveis para as localizações de CD’s uma multiplicidade de conformações com quantidades distintas de CD’s e que poderão se adaptar a diferentes posições relativas entre si, apresentando assim um efeito de um aparente “giro” desses CD’s ao longo de toda a região de estudo.

Soluções para diferentes quantidades de CD’s podem ser testadas e quaisquer que sejam as posições ótimas relativas desses CD’s, haverá uma grande possibilidade de se ter posições candidatas para atendê-las. E note-se ainda, que isto é conseguido sem a necessidade de geração de um grande número de candidatos, o que tornaria inviável do ponto de vista de tempo de processamento a utilização do modelo de otimização que será usado na próxima etapa.

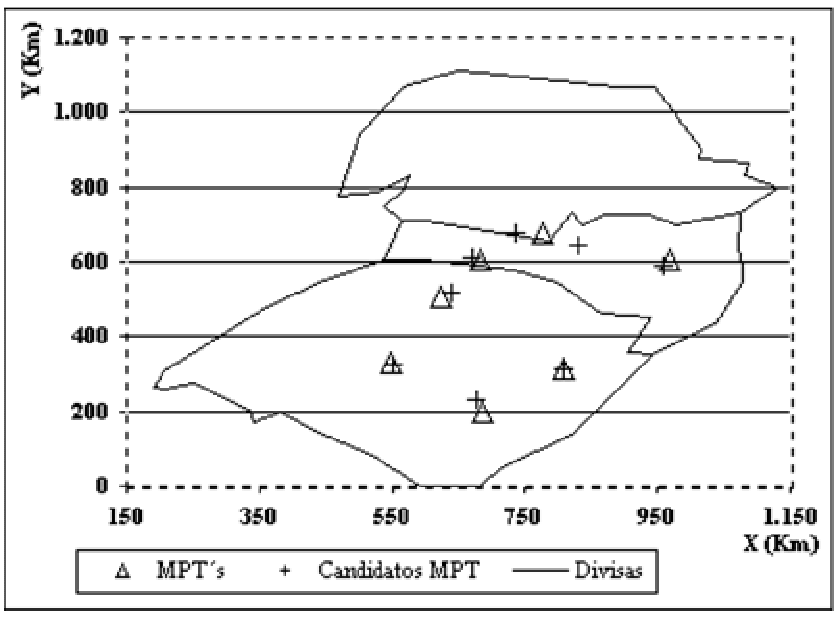

Figura 7.11a - Candidatos gerados no entorno das MPT's

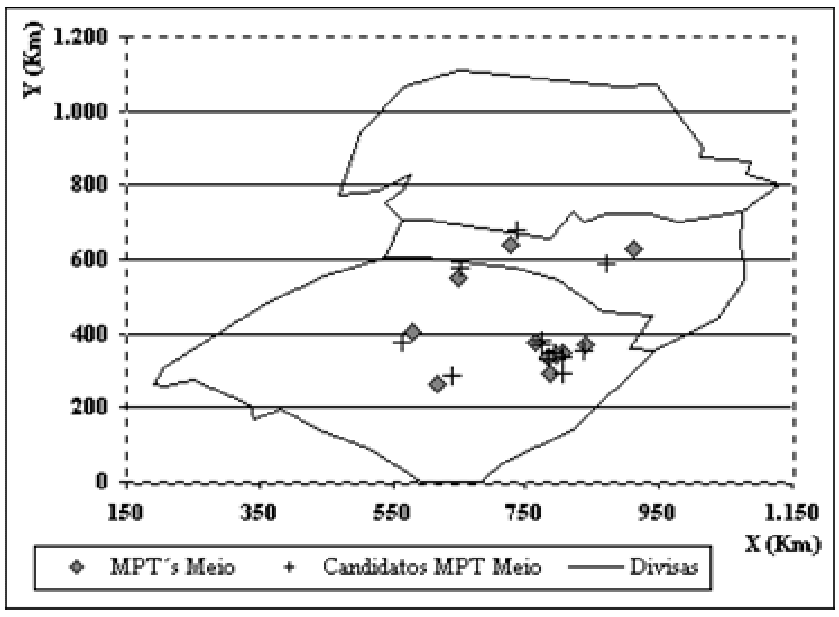

Figura 7.11b - Candidatos gerados no entorno das MPT's Meio

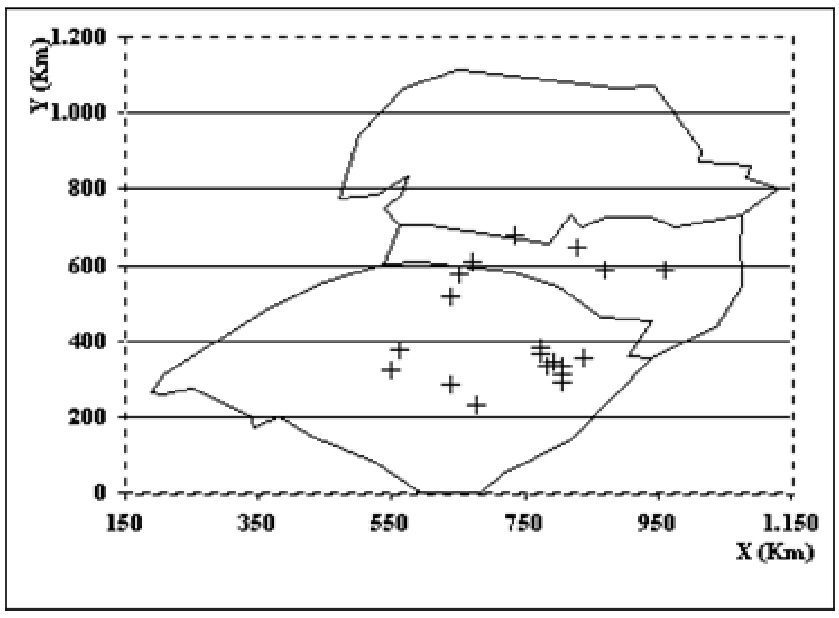

Figura 7.12 - Configuração geral de “melhores” Candidatos

7.2.4. Modelo de Otimização Aplicado a Localização dos CD’s

Nesta etapa final da *Fase 1* da metodologia é aplicado um modelo de otimização ao problema, de forma a se definir as posições ótimas para um número *p* de CD´s. O modelo empregado na análise é aquele apresentado na seção 4.4 (equações 4.17 a 4.25).

A composição desse modelo e de seus dados de entrada, procura no fundo, representar alguns aspectos que são básicos em uma operação de distribuição de carga e que se coadunam com o enfoque desta fase, que é estabelecer a localização dos CD’s.

Os pontos essenciais que os dados de entrada procuram representar são:

. Posições das fontes de produção

Isto é feito através dos custos de transporte para a transferência de mercadorias destas origens de carga para cada posição candidata a CD;

. Posições dos “melhores” candidatos a CD´s

Essas posições dos candidatos são espelhadas pelos custos de transferência citados acima e pelos custos de distribuição de carga de cada local candidato para as zonas definidas para a região de estudo.

. Posições das zonas de demanda

A posição de cada zona é representada pelo seu centro de gravidade (CG) e os custos de distribuição são estabelecidos a partir dos candidatos a CD's para cada CG.

Pode-se considerar, portanto, que os candidatos a CD´s se constituem nos dados básicos de entrada para o modelo de otimização, uma vez que são as posições desses candidatos que definem toda a composição de custos de transferência e de distribuição de carga.

Os custos unitários de transferência e de distribuição de carga utilizados no modelo foram extraídos da própria operação retratada neste estudo de caso (Vallim, 1997). São ainda utilizados como dados de entrada as ofertas de cada fonte de produção e as demandas de carga dos pontos a atender na operação de distribuição, conforme já apresentados na tabela 7.1. Uma outra questão a ser tratada diz respeito aos custos do solo, que afetam diretamente os custos fixos de instalação. Conforme já discutido na seção 3.2, considerando-se este custo uniforme ao longo da região o custo da área necessária às instalações passa a ser independente do local e pode ser integralmente tratado na *Fase 2* da modelagem. Este foi o critério adotado nesta aplicação. Assim, a decisão dos melhores locais se restringe exclusivamente aos custos de transporte.

O modelo foi rodado para o parâmetro *p* variando entre um e sete CD's, que corresponde ao número de zonas, já que cada zona é atendida por um CD.

Os resultados da solução ótima da função de custos que devia ser minimizada no modelo são apresentados na tabela 7.11, em função das quantidades de CD´s analisadas. A figura 7.13a permite uma visualização deste processo. Pela figura

verifica-se uma redução contínua do custo de distribuição, ao passo que o custo de transferência inicia um processo de aumento a partir de três CD's. Este comportamento é esperado, pois as transferências de carga aumentam à medida que o número de CD's cresce. Já a distribuição deve apresentar redução de custos, pois com mais CD's há uma maior proximidade dos clientes. As distâncias, portanto, ficam menores, como pode ser visto pela figura 7.13b. Em termos de custos totais verifica-se uma acentuada queda á medida que o número de CD´s cresce.

Tabela 7.11 – Modelo de otimização: solução ótima para *p* CD´s

No. *p* de CD´s	Custo de Transferência (R$)	Custo de Distribuição (R$)	**Custo Global de Transporte (R$)**
P=1	129.885,73	353.355,31	483.241,04
P=2	84.162,84	272.548,55	356.711,40
P=3	85.514,20	221.430,91	306.945,11
P=4	92.087,54	195.461,22	287.548,75
P=5	95.509,71	172.531,00	268.040,72
P=6	98.967,51	167.859,10	266.826,61
P=7	100.813,09	165.022,27	265.835,36

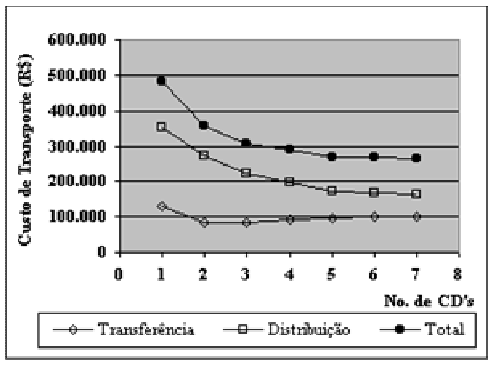

Figura 7.13a – Custos vs. No. de CD's

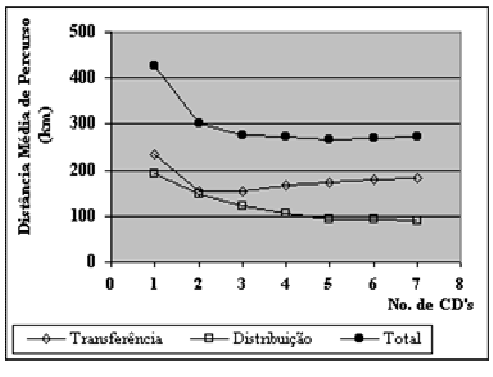

Figura 7.13b – Distâncias médias aos CD's

A construção de um modelo de regressão (figura 7.14) mostra o tipo de curva que pode ser empregada para estimar a evolução desses custos. Os resultados

mostram ainda o alto grau de aderência do fenômeno ao modelo proposto (R^2=0,9524).

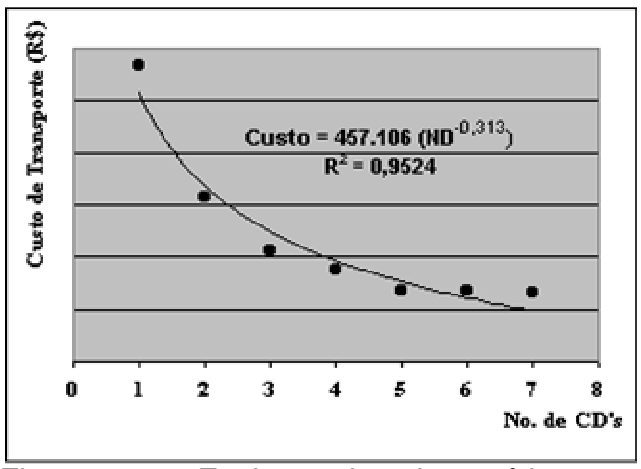

Figura 7.14 – Evolução da solução ótima para *p* CD´s

As figuras 7.15a a 7.15g apresentam as soluções ótimas encontradas para as configurações do sistema variando entre um e sete CD's. As figuras mostram que, à medida que esta quantidade de CD's cresce, as posições ótimas vão se distribuindo na área de estudo de forma a abranger praticamente todo o seu entorno. Com estas figuras visualiza-se todas as variantes das soluções.

7.2.5. Meta-Heurística SA para Localização dos CD's

A meta-heurística SA tem início com a definição de seus parâmetros e critérios conforme discutido na seção 5.1, sendo que os aspectos a definir são: temperatura inicial, número máximo de iterações em cada temperatura, número máximo de sucessos em cada temperatura, método de resfriamento, número total de iterações.

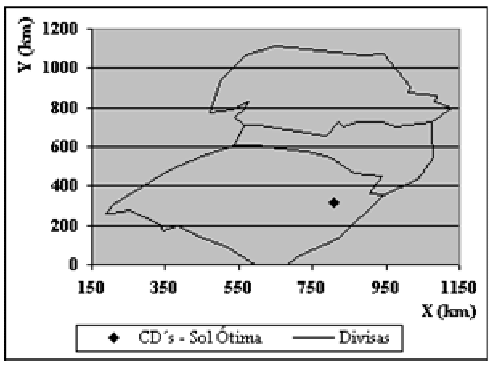

Figura 7.15a – Solução ótima - 1 CD

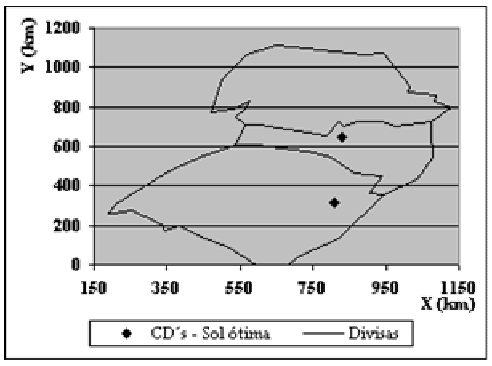

Figura 7.15b – Solução ótima - 2 CD´s

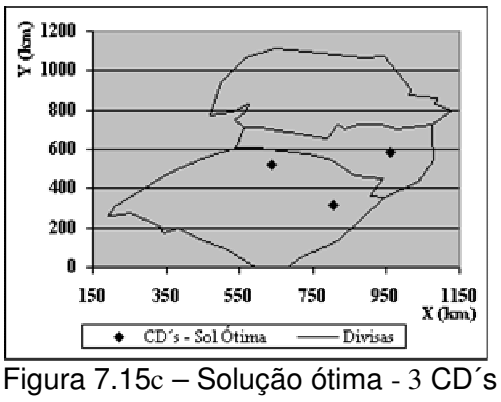

Figura 7.15c – Solução ótima - 3 CD´s

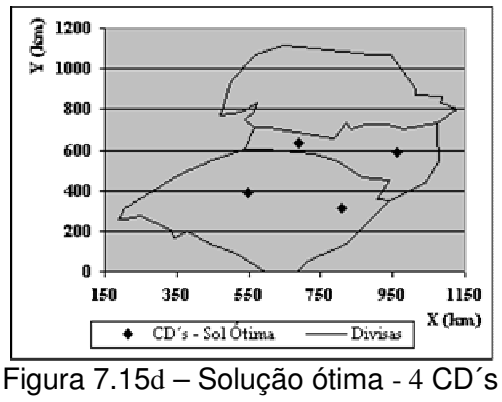

Figura 7.15d – Solução ótima - 4 CD´s

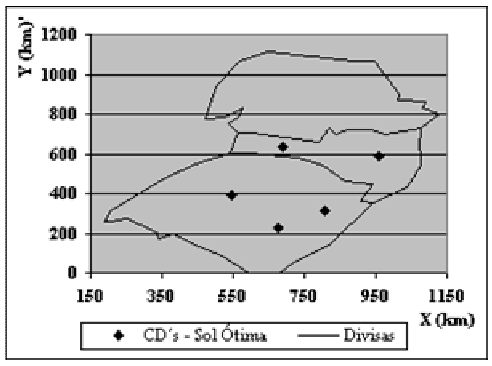

Figura 7.15e – Solução ótima - 5 CD´s

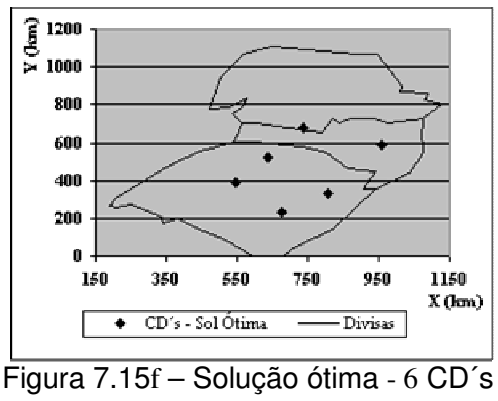

Figura 7.15f – Solução ótima - 6 CD´s

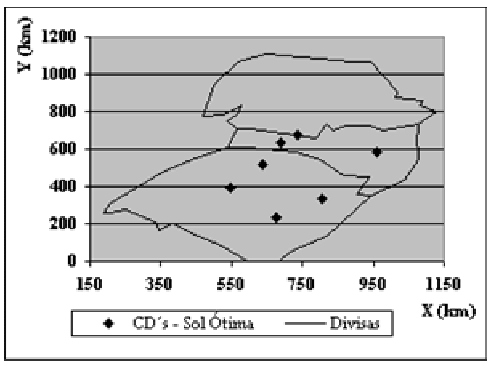

Figura 7.15g – Solução ótima - 7 CD´s

Estas questões são analisadas e definidas na seqüência.

a) Temperatura inicial (T_o)

Esta foi definida a partir da equação 5.8. O valor de ΔE_m^+ é obtido através de resultados experimentais, conforme indicado por Viana (1998) e Johnson (1989). Seu valor foi dividido por 10.000 para reduzir o número de caracteres das variáveis internas do programa de computador empregado na experimentação. O cálculo final de ΔE_m^+ resultou assim em $\Delta E_m^+ = 5{,}9$. Portanto, tem-se $T_o = 26{,}5$. Foi adotado então $T_o = 25{,}0$.

b) Número máximo de iterações em cada temperatura (p)

Considerando-se que se tinha um número de 20 candidatos e que o número de zonas era igual a 7, tem-se um total de 27 variáveis no problema.

Conforme a seção 5.1 ter-se-ia então p = 100.27 = 2.700 iterações a cada temperatura.

Pelos resultados experimentais, entretanto, verificou-se que este número estava além do necessário para que as soluções convergissem.

Na verdade, com cerca de 1.000 iterações as soluções já estavam convergindo. Após alguns testes foi adotado o valor de 800 iterações a cada temperatura, que se mostrou suficiente para gerar a convergência das soluções.

c) Número máximo de sucessos em cada temperatura (L)

Por decorrência da definição de p=800 iterações, e adotando-se L igual a 10% de p (seção 5.1), tem-se: L=80 iterações.

d) Método de resfriamento

Seguindo o padrão de Johnson(1989) e Viana(1998), que sugerem o uso de uma taxa geométrica de resfriamento variando entre 0,90 e 0,99, foram desenvolvidos alguns testes para verificar a mais adequada. Os resultados mostraram que o valor de Rc =0,975 era o mais adequado, pois foi aquele que apresentou o maior grau de convergência da solução dentro de um número de iterações operacionalmente viável.

e) Número total de iterações (m)

Os testes mostraram que o valor mínimo da função objetivo, que representa o custo da operação, apresentava oscilação até uma temperatura em torno de 1,5. Neste ponto iniciava-se uma tendência de estabilização, e a convergência final da solução efetivamente ocorria para temperaturas em torno 0,5. Isto é observado na figura 7.16a, onde é possível acompanhar a variação do custo mínimo obtido a cada nível de temperatura à medida que a temperatura decresce. Um detalhamento da fase final do processo é apresentado na figura 7.16b, que mostra a convergência final da solução.

Foi definido, assim, um valor mínimo de temperatura (T_m) de cerca de 0,5, o que garantiria a convergência da solução.

Com o valor de T_m definido, e utilizando-se a equação 5.6, determina-se o número m máximo de iterações:

$$m = \log(0{,}5 / 25{,}0) / \log(0{,}975) \rightarrow \mathbf{m = 154{,}5}$$

Foi adotado então o número máximo de 150 iterações, resultando um valor final de temperatura de 0,574.

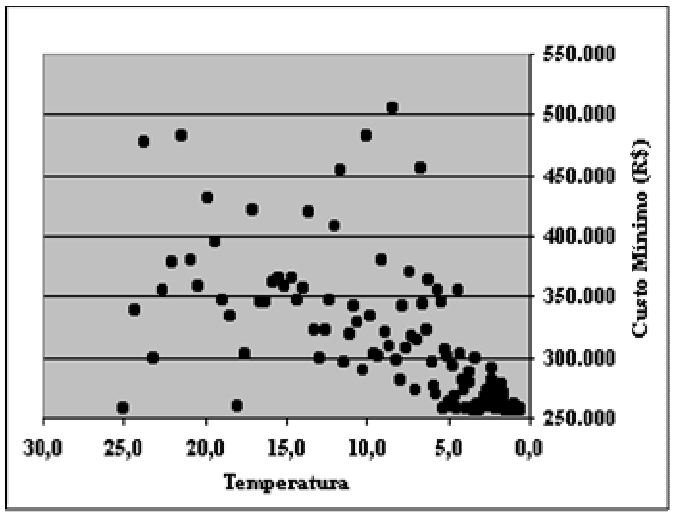

Figura 7.16a - Custo Mínimo evoluindo com o decréscimo de temperatura

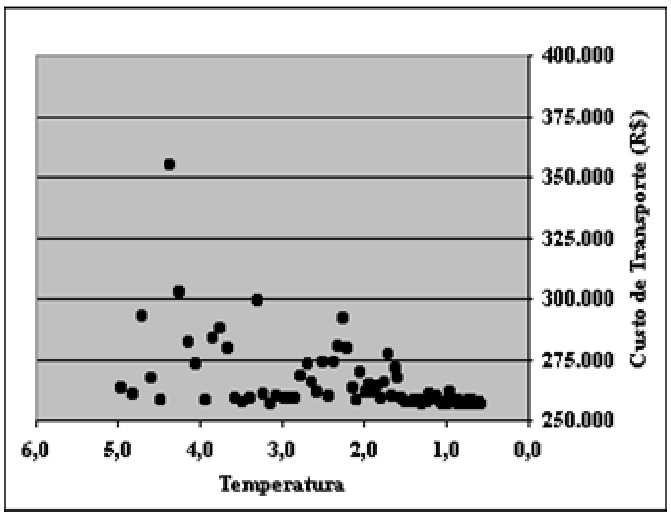

Figura 7.16b - Evolução do custo mínimo temperaturas abaixo de 5,0

Definidos esses parâmetros, a meta-heurística foi rodada em um computador com processador AMD-K6 de 450mhz e com 156mb de memória RAM. O tempo de processamento ficou entre 12 e 14 minutos. O programa desenvolvido para a simulação monitora a evolução do processo a cada temperatura através de alguns índices, conforme discutido na seção 5.3. Um dos índices mais importantes é o custo

médio da operação a cada temperatura, pois este fornece uma indicação da tendência do processo em relação aos custos. O gráfico da evolução do custo médio a cada nível de temperatura é apresentado na figura 7.17. Aquele gráfico é uma indicação da eficácia do processo. Verifica-se a queda consistente dos custos ao longo de todo o processo e de forma mais acentuada a partir de uma temperatura em torno de 5,0. Na figura 7.18 é possível ainda, visualizar-se o custo mínimo obtido pela SA a cada iteração. Verifica-se que ao final do processo de simulação este custo converge de forma definitiva para um valor abaixo de todos os outros. Este processo de convergência ocorreu em todos os experimentos feitos.

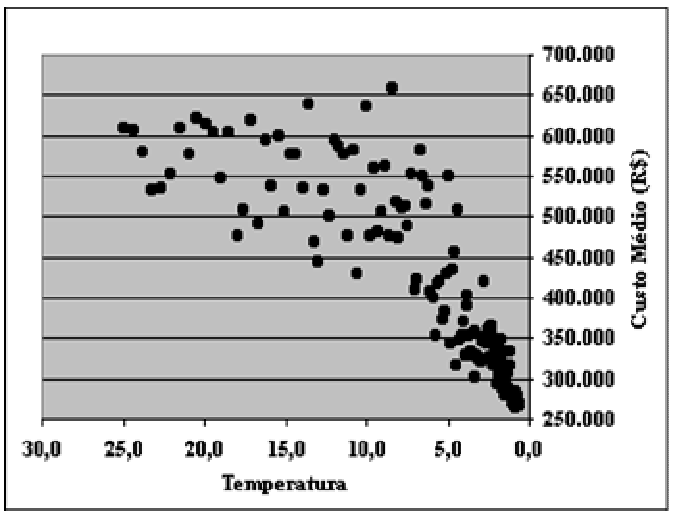

Figura 7.17 - Custo médio vs. temperatura

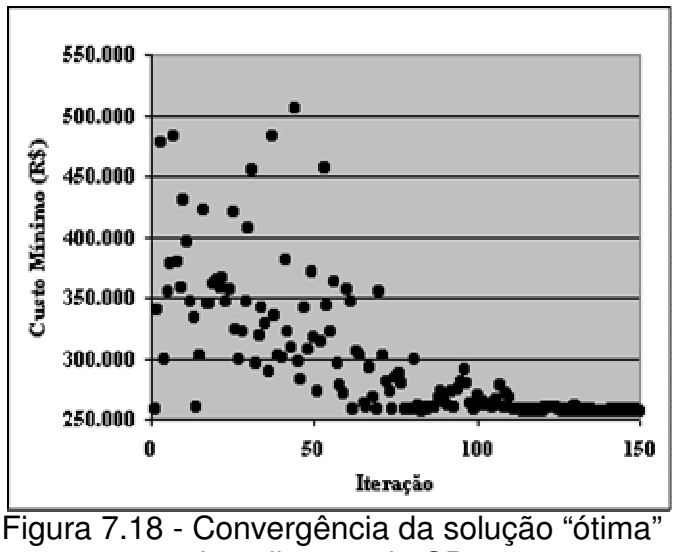

Figura 7.18 - Convergência da solução “ótima”
Localização de CD's

A exemplo da experimentação feita com o modelo MILP, aqui também o modelo foi rodado para o parâmetro *p* variando entre um e sete CD´s. A tabela 7.12 apresenta os resultados da solução ótima (função objetivo de custos) para as diferentes quantidades de CD´s analisadas. Na figura 7.19a visualiza-se este processo. A figura mostra uma redução contínua do custo de distribuição, enquanto o custo de transferência cai no início, mas após dois CD's começa a crescer. Este comportamento esperado foi semelhante ao encontrado no modelo MILP. As distâncias médias de percurso têm o mesmo tipo de comportamento, como pode ser visto pela figura 7.19b. No que tange aos custos totais, verifica-se uma queda acentuada á medida que o número de CD´s aumenta (figura 7.20).

Tabela 7.12 – *Simulated Annealing*: solução "ótima" para *p* CD´s

No. *p* de CD´s	Custo de Transferência (R$)	Custo de Distribuição (R$)	**Custo Global de Transporte (R$)**
P=1	130.070,37	347.768,33	477.839,70
P=2	82.530,02	272.194,75	354.726,77
P=3	83.541,14	221.136,60	304.680,74
P=4	87.849,20	196.116,83	283.970,03
P=5	92.922,89	174.061,25	266.989,14
P=6	93.817,56	164.576,87	258.400,43
P=7	94.056,45	162.426,46	256.489,91

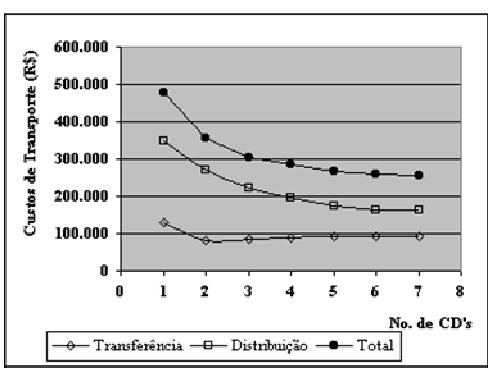

Figura 7.19a – SA: Custos vs. No. de CD's

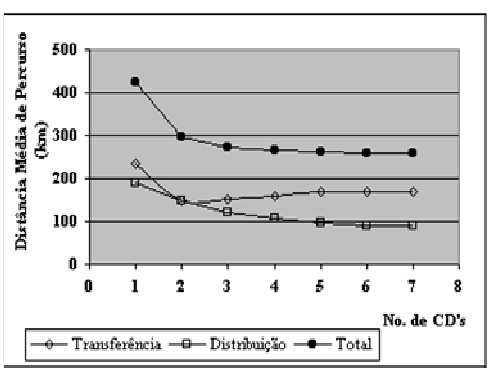

Figura 7.19b – SA: Distâncias aos CD's

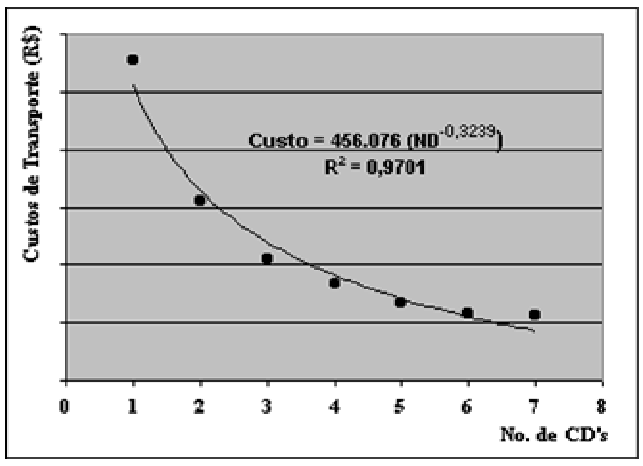

Figura 7.20 – SA: solução ótima para *p* CD´s

Aqui também foi construído um modelo de regressão (figura 7.20) que apresenta o tipo de curva que pode ser empregada para se estimar a evolução desses custos. Os resultados mostram o alto grau de aderência do processo ao modelo proposto (R^2=0,9701). Este é um modelo, portanto, que poderia ser empregado em uma necessidade de estimativas de custos para diferentes números de CD´s.

As figuras 7.21a a 7.21g apresentam as soluções ótimas encontradas para as configurações do sistema variando entre um e sete CD's. As figuras mostram o mesmo tipo de comportamento encontrado com a solução através do MILP. À medida que o número de CD's cresce, as posições ótimas vão se distribuindo na área de estudo de forma a envolver praticamente todo o seu entorno.

7.2.6. MILP e SA: Comparação de Resultados

Os resultados obtidos com o modelo MILP e a meta-heurística *simulated annealing* foram comparados de forma a permitir uma análise da aplicação dos dois métodos.

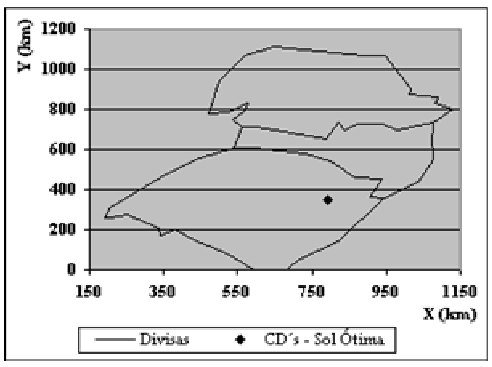

Figura 7.21a – Solução “ótima” SA - 1 CD

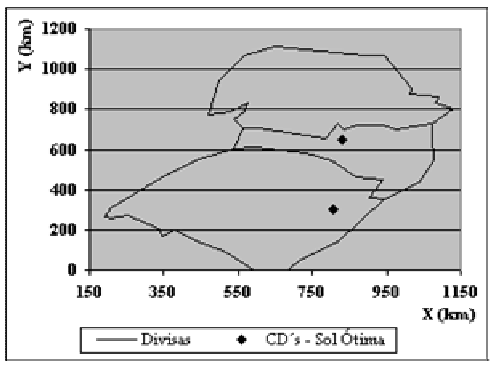

Figura 7.21b – Solução “ótima” SA - 2 CD´s

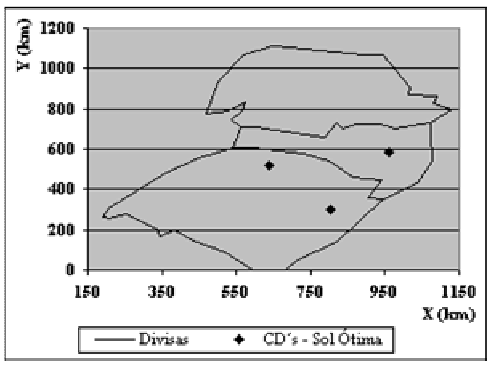

Figura 7.21c – Solução “ótima” SA - 3 CD´s

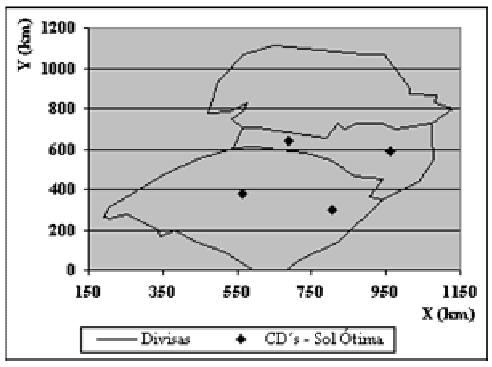

Figura 7.21d – Solução “ótima” SA - 4 CD´s

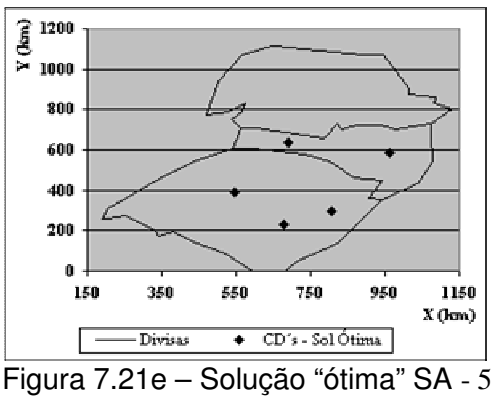

Figura 7.21e – Solução “ótima” SA - 5 CD´s

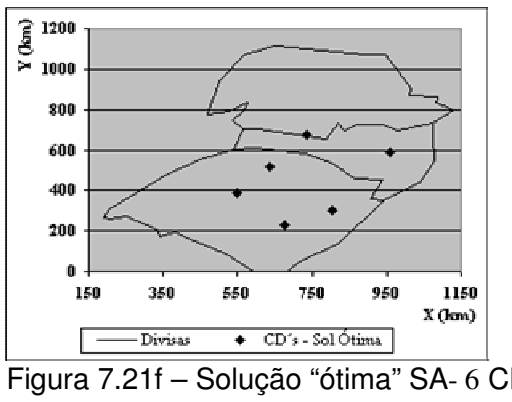

Figura 7.21f – Solução “ótima” SA- 6 CD´s

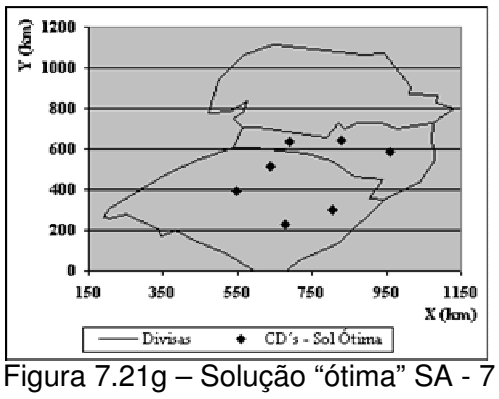

Figura 7.21g – Solução “ótima” SA - 7 CD´s

As figuras 7.22a a 7.22g mostram as soluções dos dois métodos, sobrepostas nos mesmos gráficos.

Verifica-se que houve coincidência de soluções em vários casos e que, quando isto não ocorreu, os candidatos escolhidos em cada caso foram muito próximos. A única exceção foi a configuração com sete CD's, em que, dentre os sete candidatos selecionados, houve uma diferença mais significativa na escolha de um deles. Quanto aos outros seis candidatos, as posições ou foram bem próximas ou coincidentes. Esta diferença é justificada pelo fato de que o modelo MILP trata as zonas de demanda de forma agregada, concentrando toda a carga em seu centro de gravidade. Já a meta-heurística SA considera os pontos de demanda de forma individual.

É possível observar que na solução gerada pela SA os CD´s ficam distribuídos na área de estudo de uma forma mais harmoniosa, enquanto que no MILP há um direcionamento mais a oeste da área. Isto é conseqüência da própria posição dos centros de gravidade, que acaba se refletindo na qualidade da solução.

A tabela 7.13 apresenta uma comparação dos custos finais obtidos nos dois métodos. Em todas as situações houve uma pequena vantagem para a SA. Há uma diferença maior no caso de seis CD's, pois apesar de ter havido coincidência dos candidatos selecionados em cinco CD's, não houve coincidência na zona 1.1, que é aquela que apresenta o maior volume (vide tabela 7.8). Já a maior diferença de custo ocorreu justamente para o caso de sete CD´s, mostrando que a solução via SA pode ser mais próxima do menor custo real do que aquela gerada pelo modelo MILP.

Tabela 7.13 – MILP vs *Simulated Annealing*
solução "ótima" para *p* CD´s

No. *p* de CD´s	Custo de Transporte (R$)		Δ % SA / MILP
	MILP (R$)	*Simulated Annealing* (R$)	
P=1	483.241,04	477.839,70	-1,1%
P=2	356.711,40	354.726,77	-0,6%
P=3	306.945,11	304.680,74	-0,7%
P=4	287.548,75	283.970,03	-1,2%
P=5	268.040,72	266.989,14	-0,4%
P=6	266.826,61	258.400,43	-3,2%
P=7	265.835,36	256.489,91	-3,5%

Na seqüência apresentam-se nas figuras 7.23a a 7.29b os centros de gravidade das sete zonas definidas na seção 7.2.2, conectados aos CD's que deverão atendê-las conforme estabelecido na solução do modelo MILP e da meta-heurística SA.. Com isto visualizam-se todas as variantes das soluções.

Quanto a esta alocação de zonas aos CD´s verifica-se que a proximidade entre os dois métodos também foi grande. A exceção aqui é o caso de quatro CD's, em que as diferenças entre a modelagem agregada (MILP) e a desagregada (SA) aparecem novamente. A diferença de custo entre os dois métodos foi a terceira mais elevada (1,2%), apesar de não ter se situado no patamar mais alto que seria acima de 3%.

A meta-heurística proposta, portanto, mostrou-se eficaz gerando soluções ligeiramente melhores que a de um modelo de otimização tradicionalmente empregado para este tipo de problema.

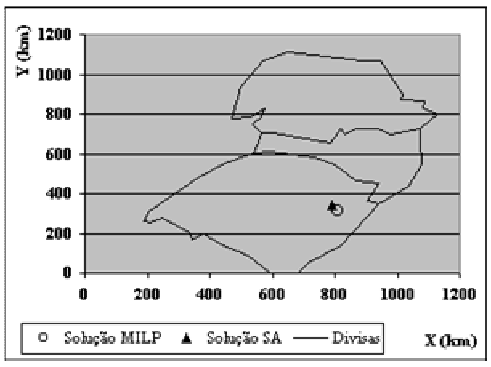

Figura 7.22a – MILP vs SA - 1 CD

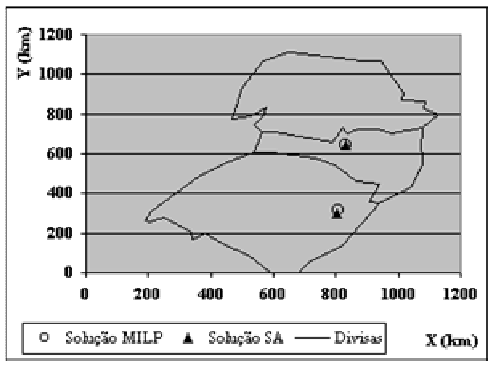

Figura 7.22b – MILP vs SA - 2 CD´s

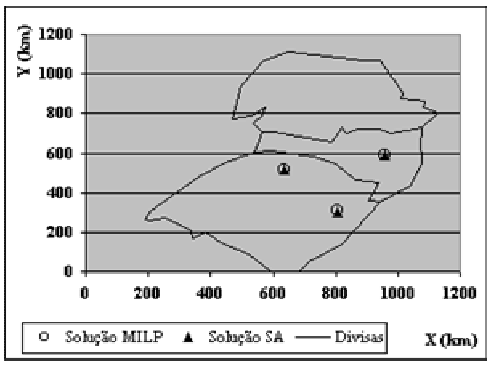

Figura 7.22c – MILP vs SA - 3 CD´s

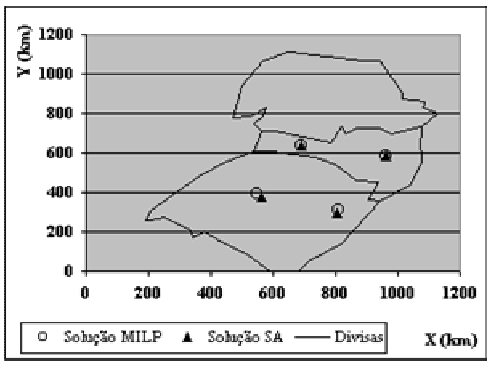

Figura 7.22d – MILP vs SA - 4 CD´s

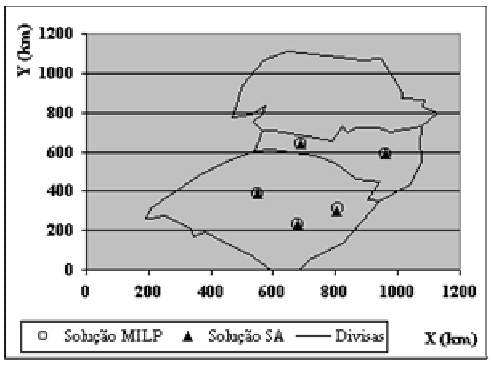

Figura 7.22e – MILP vs SA - 5 CD´s

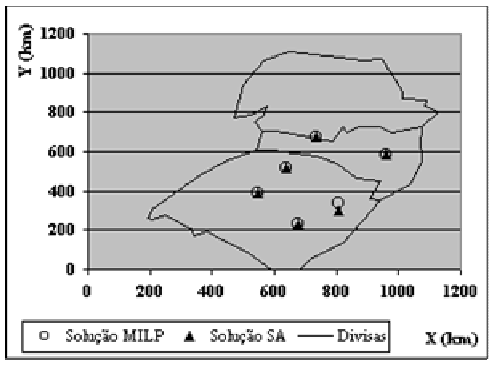

Figura 7.22f – MILP vs SA - 6 CD´s

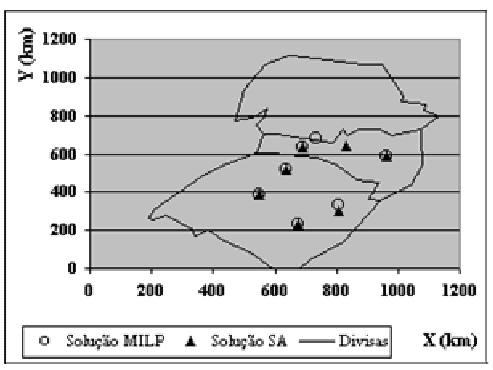

Figura 7.22g – MILP vs SA - 7 CD´s

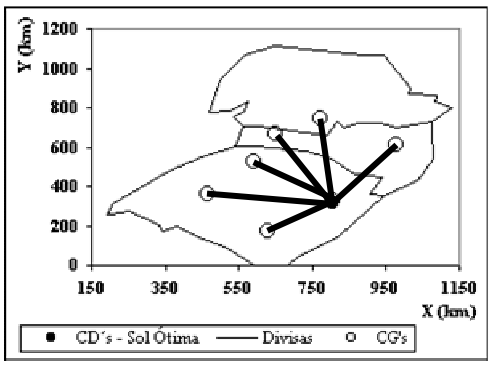

Figura 7.23a – 1 CD - MILP

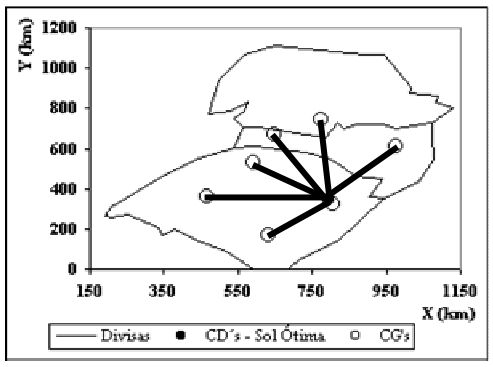

Figura 7.23b – 1 CD - SA

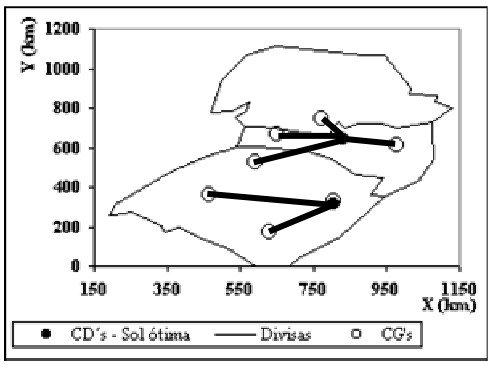

Figura 7.24a – 2 CD's - MILP

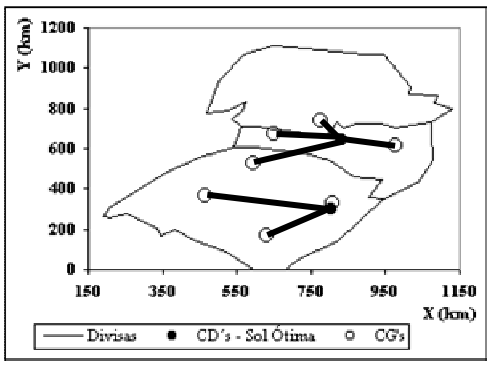

Figura 7.24b – 2 CD's - SA

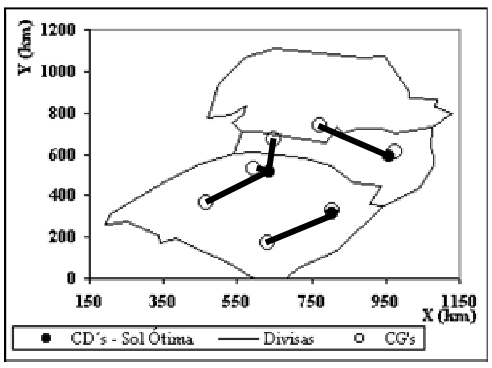

Figura 7.25a – 3 CD's - MILP

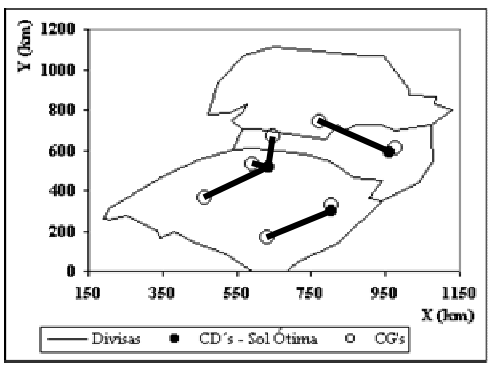

Figura 7.25b – 3 CD's - SA

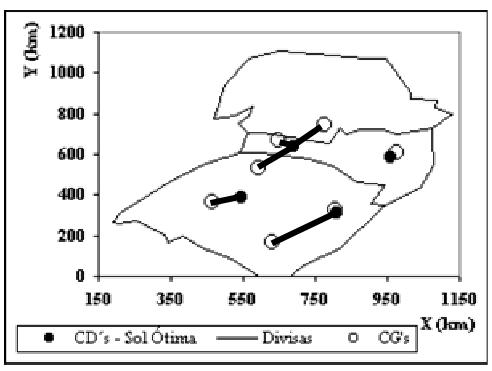

Figura 7.26a – 4 CD's - MILP

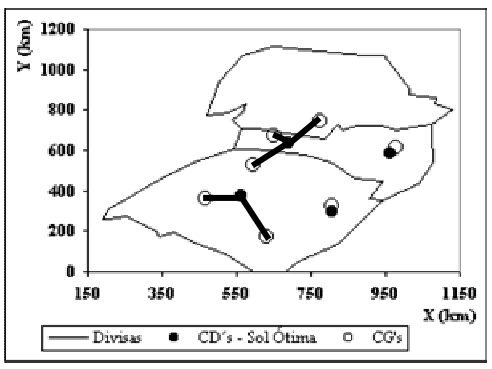

Figura 7.26b – 4 CD's - SA

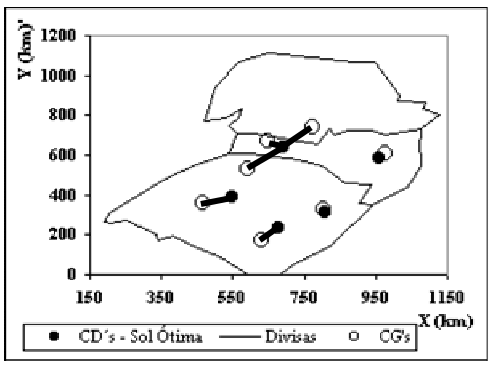

Figura 7.27a – 5 CD's - MILP

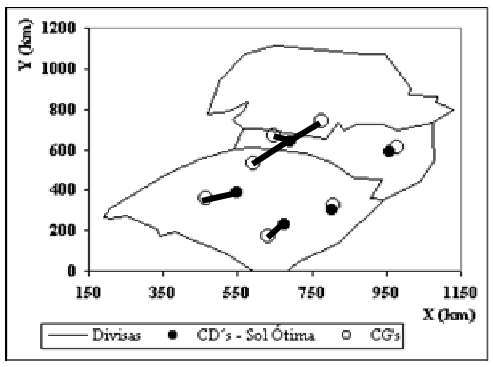

Figura 7.27b – 5 CD's - SA

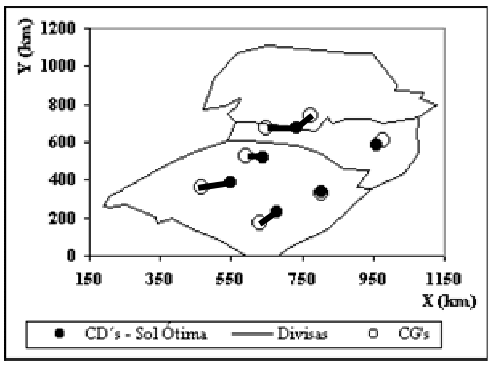

Figura 7.28a – 6 CD's - MILP

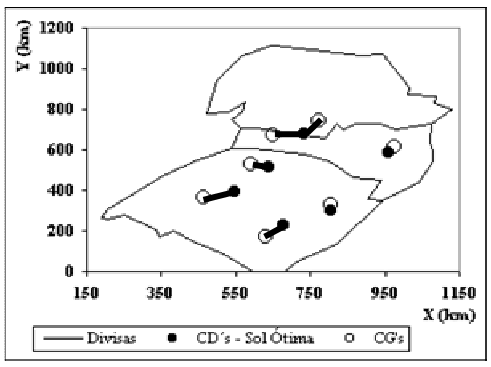

Figura 7.28b – 6 CD's - SA

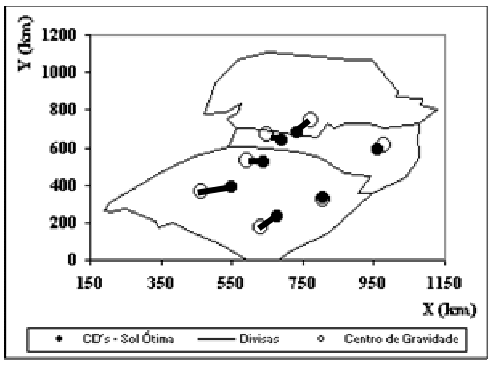

Figura 7.29a – 7 CD's - MILP

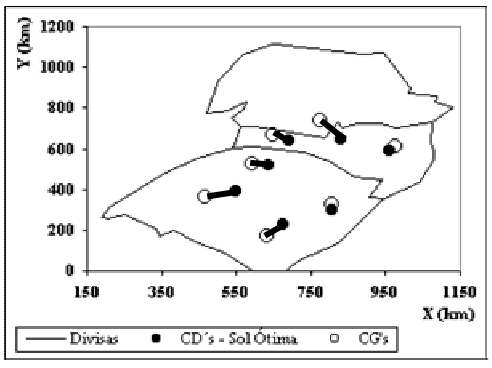

Figura 7.29b – 7 CD's - SA

7.3. FASE 2: QUANTIFICAÇÃO DE CD'S

Nesta *Fase 2* da metodologia busca-se identificar a quantidade "ótima" de CD's.

Para isto serão computados os custos de estoque, armazenagem e movimentação de materiais e vendas perdidas. Estes custos serão estimados para diferentes quantidades de CD's e, ao final, serão consolidados com o custo de transporte já computado na *Fase 1*. Tendo-se assim o custo logístico global, será possível identificar a quantidade de CD's que leva ao menor custo total. Esta será considerada a quantidade "ótima" de CD's. Uma vez estabelecido o número "ótimo" de CD's, estes deverão se situar nos locais definidos na *Fase 1.*

7.3.1. Modelo de Custos de Estoque

O componente principal dos custos de estoques é o chamado custo de oportunidade do capital que está investido em estoque (equação 6.4), que deve ser aplicado a diferentes tipos de coberturas de estoque, conforme visto na seção 6.2.

A determinação dessas múltiplas coberturas de estoque é apresentada a seguir.

a) Cobertura de estoque regulador

Estoque regulador ou cíclico é aquele que é utilizado para regular a oferta e a demanda. É um estoque que permanece armazenado nos CD's por um certo período até que seja embarcado para os pontos de demanda.

Conforme definido este estoque é composto de uma parcela para atender à necessidade imediata de embarque e uma outra que antecipa necessidades. Será calculada inicialmente a parte do estoque regulador correspondente à necessidade de embarque, ou seja, aquela parte que está sendo chamada aqui de estoque sincronizado.

Pelos dados da operação da empresa, tem-se que o volume médio embarcado (V_E) é de 44,02 m^3. Sabe-se ainda que o número médio de dias de operação da empresa

é de 23,79 dias/mês e que a demanda mensal é de 45.651 m3. Portanto, para o caso de um CD, tem-se:

$$d = 45.651/23,79 = 1.918,9\ m^3/dia$$

Através de 6.6 tem-se a cobertura total de estoque sincronizado para um CD que será:

$$C_{RS} = 44,02\ /1.918,9 = 0,02\ dias$$

A outra parcela do estoque regulador é o estoque adquirido por antecipação para fazer frente à demanda futura, ou até mesmo para aproveitar oportunidades de mercado. Seguindo o padrão definido na seção 6.2, o estoque de antecipação foi determinado através de uma análise de regressão (Draper e Smith,1981) entre os níveis de estoque da empresa e o volume movimentado. Esses estoques são na verdade utilizados para os consumos dos próximos períodos. As figuras 7.30 e 7.31 mostram esta relação. Os volumes movimentados no mês i+1 seguem padrão semelhante aos estoques do mês anterior (mês i). Já ao se construir uma curva para a movimentação do mês i (figura 7.32) não se identifica semelhança com o comportamento do estoque do mesmo mês.

Assim, conforme apresentado na figura 7.33, foi desenvolvida uma análise de regressão entre o nível de estoque da empresa ao final do mês i (E_i) e o volume movimentado no mês i+1 (V_{i+1}), resultando na seguinte equação:

$$E_i = 0,0021 V_{i+1}^{0,8339} \qquad (7.1)$$

Apesar da semelhança de comportamento entre as curvas de flutuação mensal dessas variáveis, o coeficiente de determinação para este modelo não foi muito elevado (R^2=0,6531), mas segundo Ballou (1994) deve-se usar a melhor curva que se puder obter.

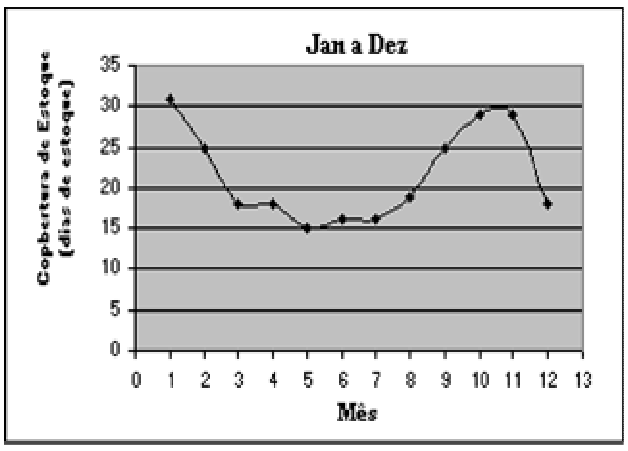

Figura 7.30 – Nível de estoque no mês i

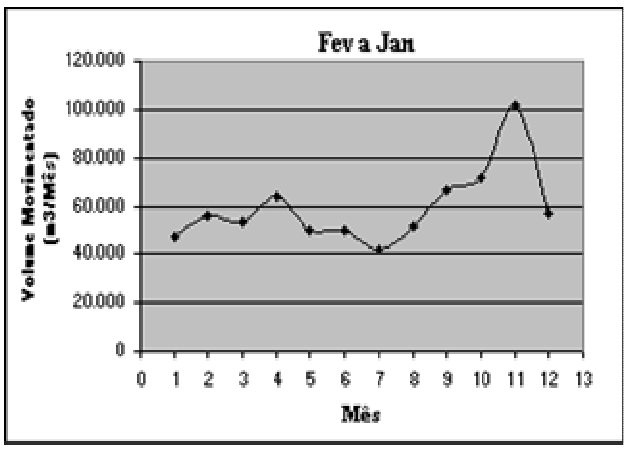

Figura 7.31 – Volume movimentado no mês i+1

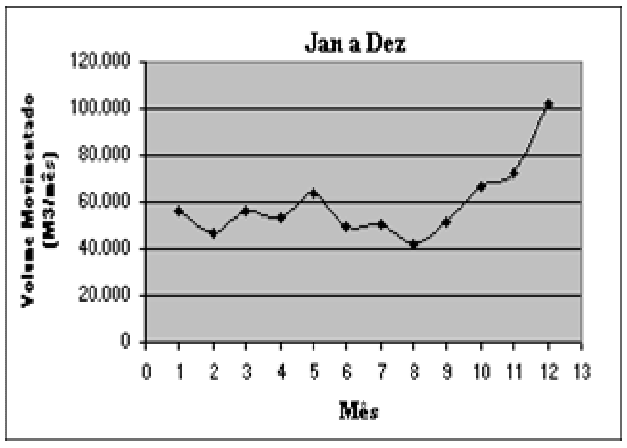

Figura 7.32 – Volume movimentado no mês i

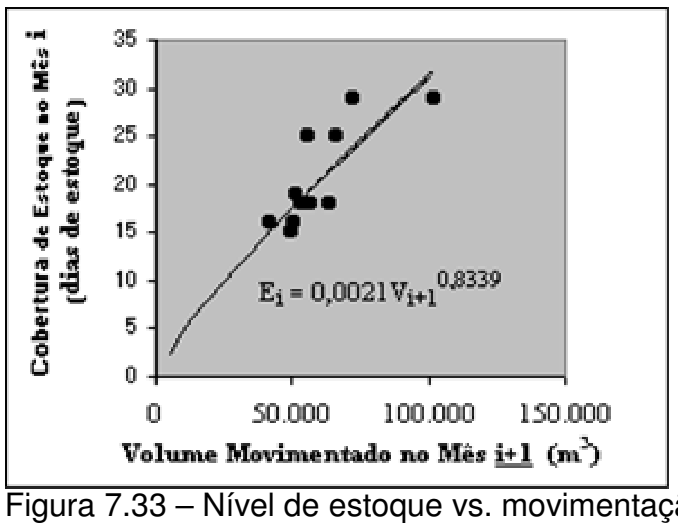

Figura 7.33 – Nível de estoque vs. movimentação

E neste caso, em particular, as características da operação indicam que não será possível a obtenção de uma curva melhor que esta, pois falta nesta operação uma política mais bem definida para a reposição de estoque. Não é adotado nenhum dos sistemas clássicos de gestão de estoques, como o sistema de *quantidade fixa / período variável* ou de *quantidade variável / período fixo* (Chopra e Meindl, 2001) ou qualquer outro, nem tampouco são dimensionados os lotes de reposição a partir de algum critério com embasamento científico, como, por exemplo, o lote econômico de compra (Ballou, 1994). Os lotes de reposição vão sendo definidos a partir de necessidades momentâneas e de oportunidades de mercado em relação a custos de aquisição.

Assim, para o caso de um CD considerando-se a demanda média mensal de 45.651m^3 e aplicando-se 7.1, tem-se que a cobertura média do estoque de antecipação (C_{RA}) será:

$$C_{RA} = 0{,}0021.(45.651)^{0{,}8339} = 16{,}13 \text{ dias}$$

Desenvolvendo-se estes mesmos cálculos para um número maior de CD's até um máximo de sete instalações e transformando-se os níveis de estoque em custos através da equação 6.4, em conjunto com 6.8, chega-se a uma curva da evolução do custo de estoque regulador com o número de CD's (figura 7.34).

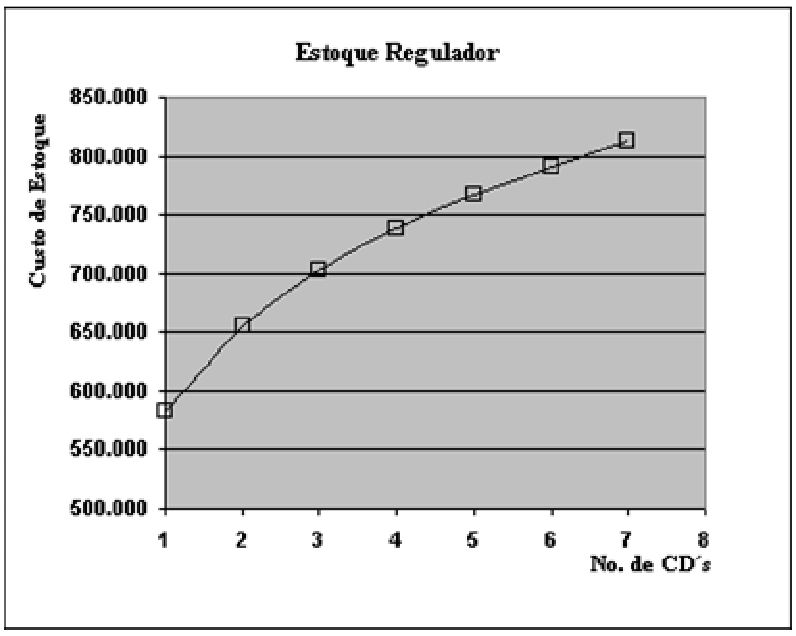

Figura 7.34 –Estoque regulador vs. número de CD's

b)Estoque em trânsito

Este estoque corresponde ao tempo que o estoque permanece em viagem. Isto se reflete, na verdade, em uma necessidade de cobertura adicional de estoque devido ao trânsito do material.

No caso deste tipo de estoque há uma influência da distância média de percurso que irá definir se o material ficará um tempo maior ou menor em trânsito. Conseqüentemente, os resultados aqui serão afetados pela técnica empregada na solução da *Fase 1*: MILP ou SA. Pelas figuras 7.13b e 7.19b tem-se a variação da distância média de percurso em função do número de CD's para as soluções do modelo de otimização MILP e para a meta-heurística SA. Para estas duas situações

foram construídos modelos de regressão que relacionam as variáveis (figuras 7.35a e 7.35b). Com esses modelos e utilizando-se a equação 6.10 e posteriormente 6.4, tem-se instrumentos para estimar o tempo em trânsito da carga de acordo com a quantidade de CD's. Para esta aplicação prática foi adotada uma velocidade de cruzeiro de 60 km/h e uma jornada de trabalho de oito horas diárias, que são compatíveis com a operação.

Aplicando-se esses modelos às diferentes quantidades de CD's testadas obtém-se uma curva da evolução do estoque em trânsito (figuras 7.36a e 7.36b). Como os modelos são muito parecidos para as duas situações, as curvas resultantes também se apresentam semelhantes.

c) Estoque de segurança

Estoque de segurança é a quantidade adicional mantida em armazém para proteção contra variações de demanda e de prazos de entrega (*lead time*).

Para alimentar a formulação vista na seção 6.2, que estima as coberturas deste tipo de estoque, foi estimado a partir dos dados da operação em estudo um desvio padrão de 444,6m^3 para a demanda diária da empresa. Adotando-se um nível de serviço de 95%, tem-se o valor crítico da distribuição normal reduzida igual a 1,645.

O *lead time* médio da operação está em torno de dois dias. Com estes parâmetros e utilizando-se a equação 6.12 pode-se determinar o estoque de segurança para o caso de um CD:

$$C_s = 1{,}645 \,.\; 444{,}6.2^{1/2} / 1.989{,}9 = 0{,}54 \text{ dia}$$

Utilizando-se a seguir a regra da raiz quadrada, determina-se o nível de estoque de segurança para outras quantidades de CD's (equação 6.14). Utilizando-se depois a equação 6.4 para esses diferentes números de CD´s tem-se o custo correspondente.

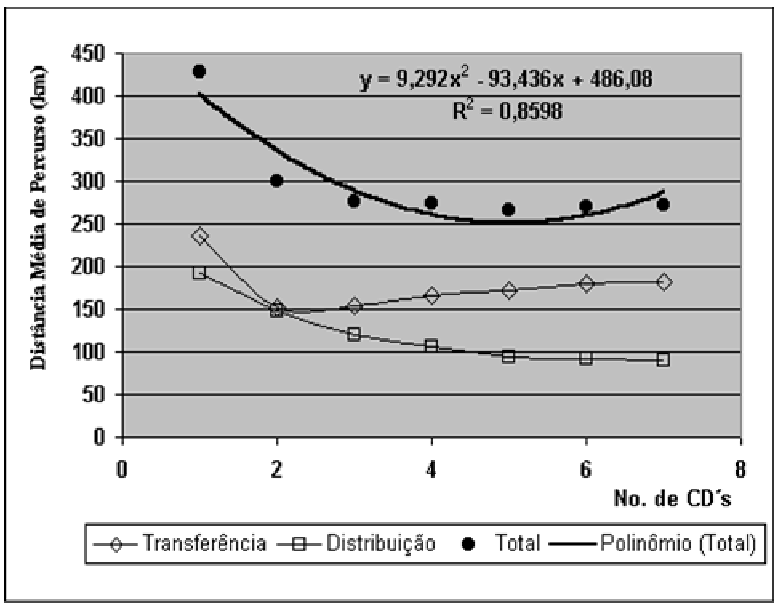

Figura 7.35a – Distância de percurso vs. No. de CD's - MILP

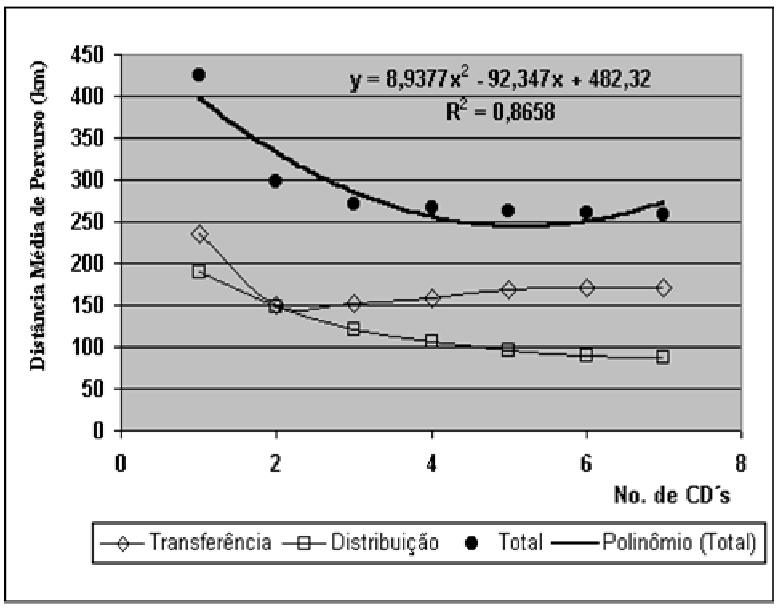

Figura 7.35b – Distância de percurso vs. No. de CD's – SA

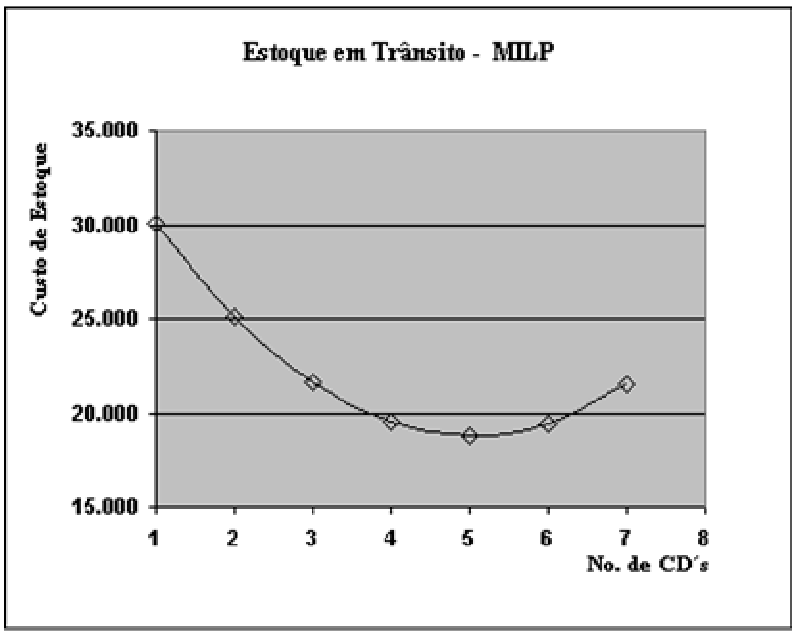

Figura 7.36a – Estoque em trânsito vs. No. de CD's – MILP

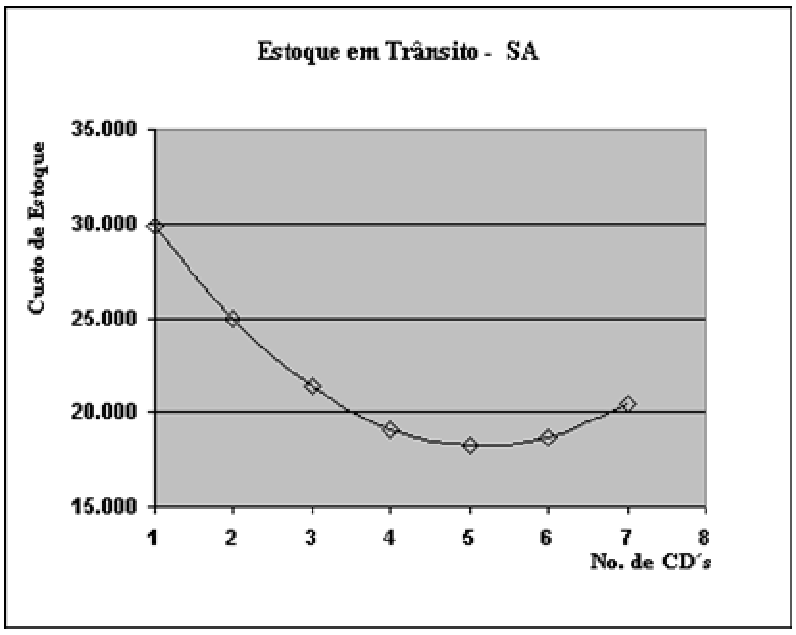

Figura 7.36b – Estoque em trânsito vs. No. de CD's – SA

O resultado desta análise é apresentado na figura 7.37, onde se pode visualizar o comportamento do estoque de segurança à medida que cresce a quantidade de CD's.

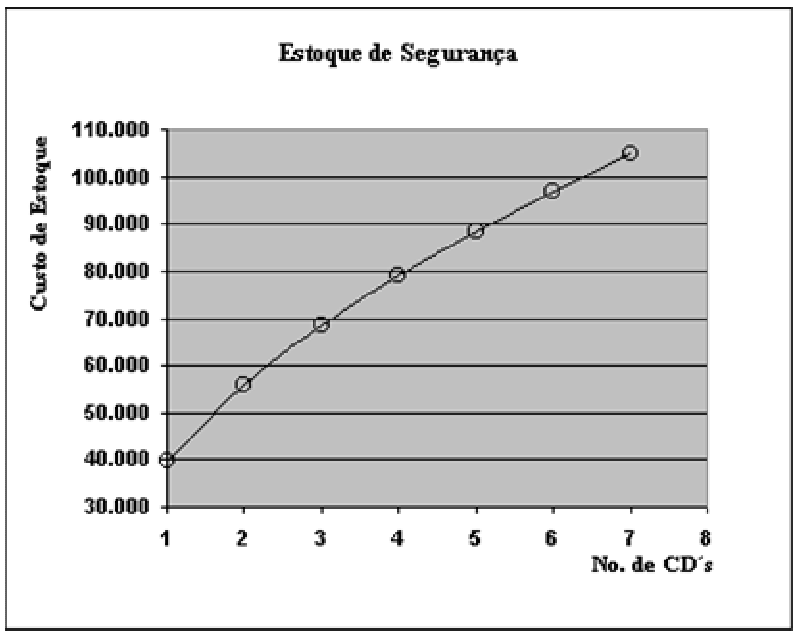

Figura 7.37 – Estoque de segurança vs. No. de CD's

d) Estoque total

Considerando-se agora a partir de estatísticas da operação uma taxa de 0,02% de outros custos de estoque e utilizando-se a equação 6.15, tem-se o custo total de estoque.

Aplicando-se a equação para diferentes valores de p tem-se uma estimativa da evolução desses custos, conforme apresentado nas figuras 7.38a e 7.38b.

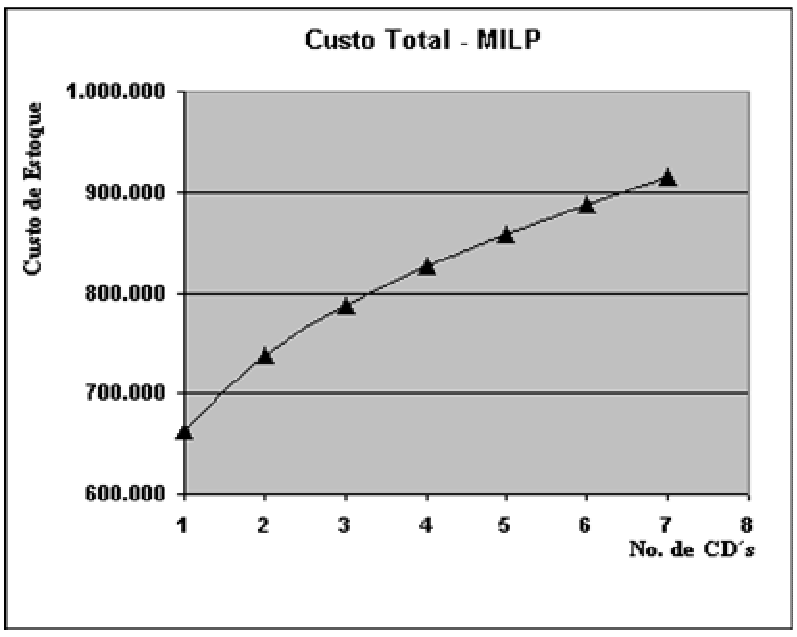

Figura 7.38a – Estoque total vs. No. de CD's - MILP

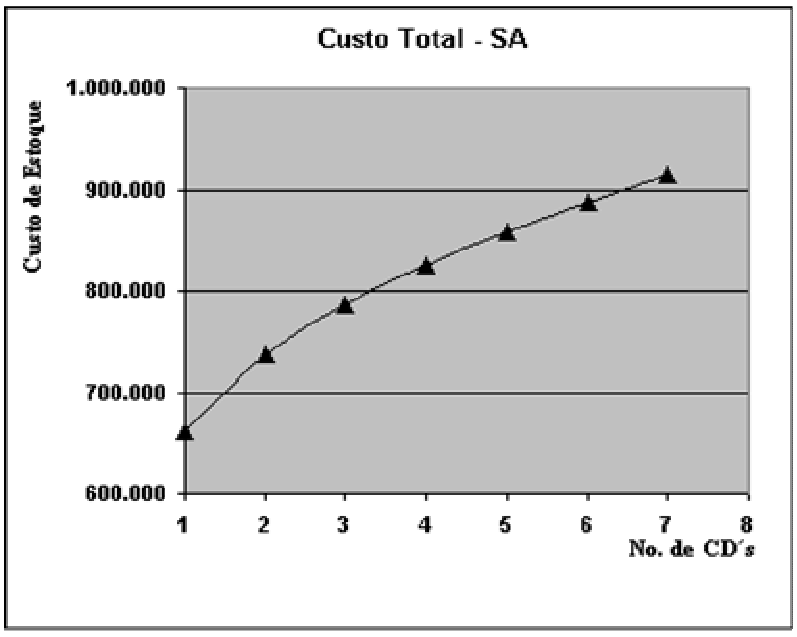

Figura 7.38b – Estoque total vs. No. de CD's - SA

7.3.2. Modelo de Custos de Armazenagem e Movimentação de Materiais

Conforme a metodologia apresentada na seção 6.3, para cada configuração de CD analisada (um a sete CD´s) deve-se estabelecer o patamar de custo fixo de armazenagem e movimentação correspondente. É necessário, portanto, conhecer o estoque médio previsto por CD ($C_R + C_S$) e enquadrá-lo em um patamar.

Os patamares foram definidos a partir de uma área e estrutura conhecida na operação em estudo. Para a operação de um armazém com 12.000m^2 de área útil de armazenagem tem-se conhecimento de todos os custos fixos para sua operação. O módulo de área considerado foi de 1.000m^2, o que significa que a área evolui em blocos iniciando em 1.000m^2 e terminado em 12.000m^2. Considerou-se este limite superior, pois a operação em estudo ocorria com um único armazém. Sendo esta, portanto, a dimensão máxima necessária. Este módulo de área está compatível com a área construída considerada no nível de serviço da operação, que foi de 2.000m^2. Considerando que o aproveitamento de área nesta operação era de 55%, o módulo corresponde aproximadamente aos 2.000m^2.

O pé direito adotado foi de 7,5m, que é muito próximo do atualmente utilizado no armazém em uso (7,45m). E foi ainda utilizado um coeficiente de aproveitamento da área útil de armazém de 76%, que é um parâmetro levantado em estudo realizado em empresas brasileiras distribuídas em várias regiões do país (CEL/COPPEAD, 2001).

A partir destes dados foram construídos onze patamares (medidos em m^3) de volume útil efetivo de armazenagem. Os patamares variaram de 5.700m^3 a 68.400m^3 cada, com uma amplitude de 5.700m^3. O patamar 1 é o de nível mais alto

de custo, que corresponde ao maior volume, e o patamar 11 o de nível mais baixo, associado à menor faixa de volume.

Os níveis esperados de estoque (C_R + C_S) em armazém para cada configuração de CD's foram transformados em volumes necessários de armazém utilizando-se a demanda média diária (medida em m^3) e o índice de aproveitamento de volume útil de 76%. Considerando-se assim, esses níveis esperados de estoque em CD's, procedeu-se ao enquadramento de cada configuração no seu respectivo patamar de custo.

Foram considerados, no cômputo dos custos de armazenagem e movimentação de materiais de um módulo de área, quatro componentes: imóveis, equipamentos, salários e despesas gerais de armazém. A partir deste custo do módulo tem-se o custo de todos os patamares. Os valores foram levantados na própria operação.

A figura 7.39 apresenta as sete configurações analisadas, com a respectiva alocação a seu patamar de custo. Naquela figura o maior volume se refere à configuração com um único CD e o menor ao caso de sete CD's. Note-se que as configurações com cinco, seis e sete CD´s situam-se todas no mesmo patamar de custo, que corresponde a um volume variando 5.700 m^3 e 11.400m^3. Este aumento de volume sem aumento de custo representa aqui um ganho que pôde ser captado pelo modelo.

O uso da equação 6.24 leva agora aos custos finais. Para isto é necessário o cômputo dos custos unitários médios por patamar, que por sua vez utilizam o número médio de dias úteis de operação. A média para esta operação foi de 23,79 dias por mês.

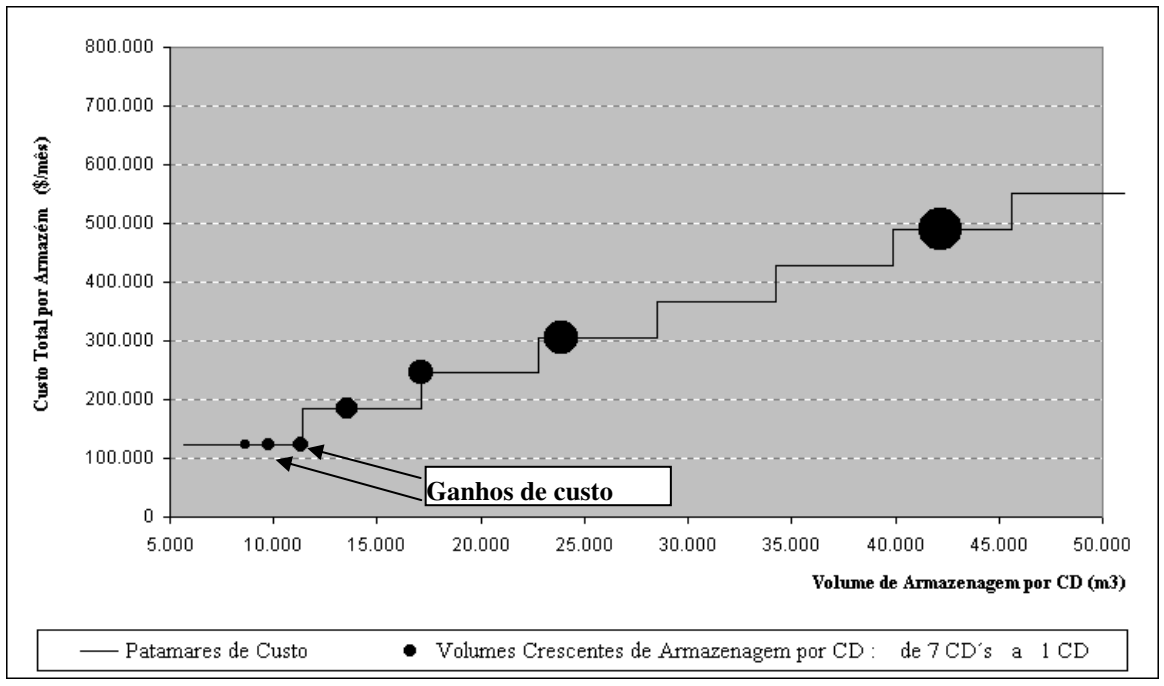

Figura 7.39 - Configurações de CD´s alocadas a patamares de custo

Uma representação da variação obtida para os custos unitários em cada patamar é apresentada na figura 7.40 e ao centro vê-se o custo médio de cada patamar. Uma curva somente dos custos médios é apresentada na figura 7.41, onde se visualiza a queda de custos à medida que o volume por CD aumenta, lembrando que este aumento de volume por CD ocorre com a redução do numero de CD´s.

Aplicando-se a equação 6.24 tem-se o custo final correspondente a cada quantidade de CD´s analisada (figura 7.42).

Nota-se que nos último três pontos da curva da figura 7.42 a taxa de incremento de custo se reduziu. Estes são os pontos que estavam situados no mesmo patamar de custo e que, portanto, não tiveram aumento no custo médio unitário por dia de armazenagem. O aumento ocorre apenas pelo efeito do crescimento do volume a armazenar. Os demais pontos sofreram influência dos dois efeitos: custo unitário médio e volume em armazém.

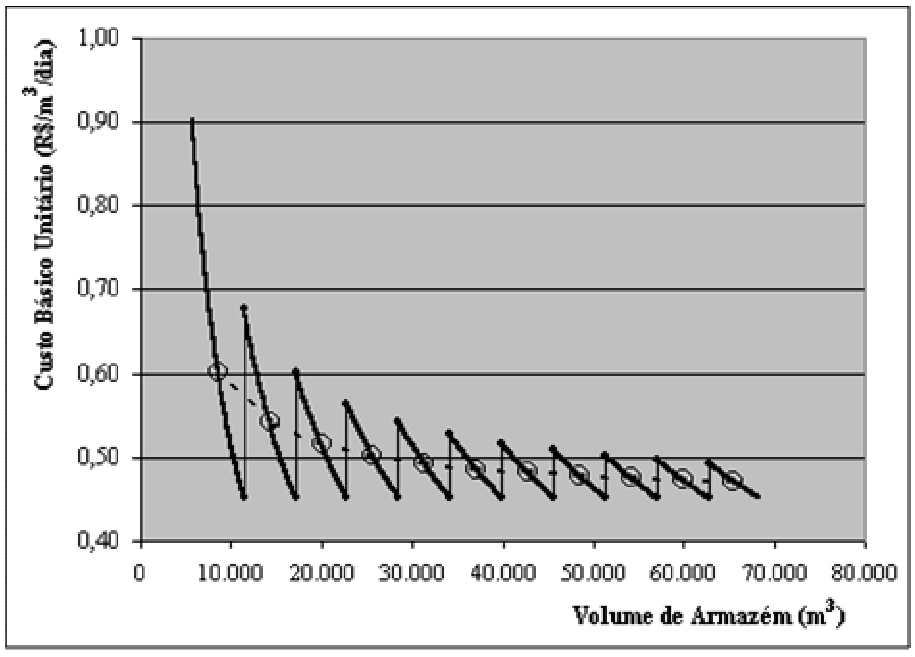

Figura 7.40 – Variação de custos unitários por patamar de custo

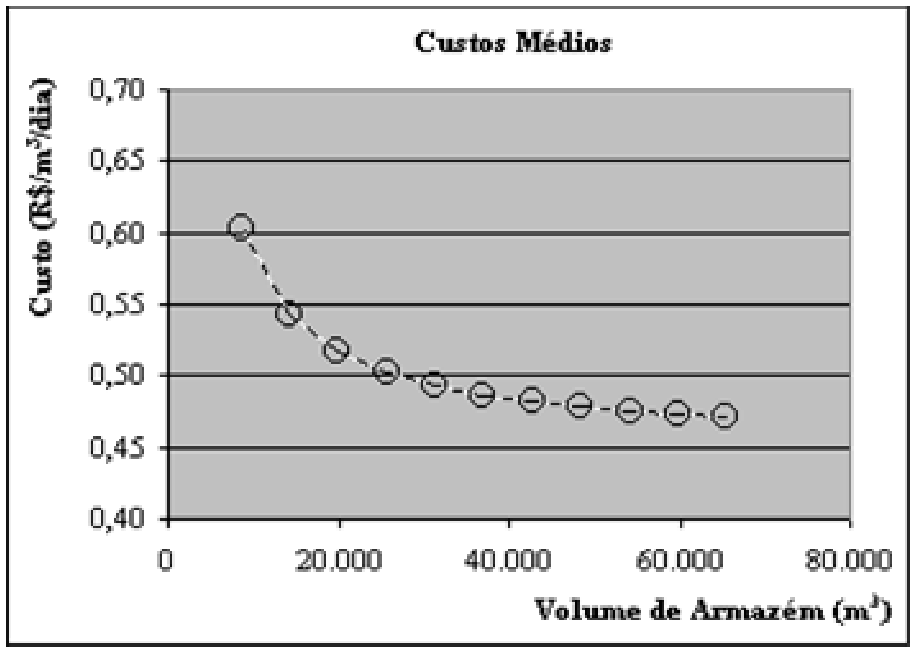

Figura 7.41 - Custos unitários médios por patamar de custo

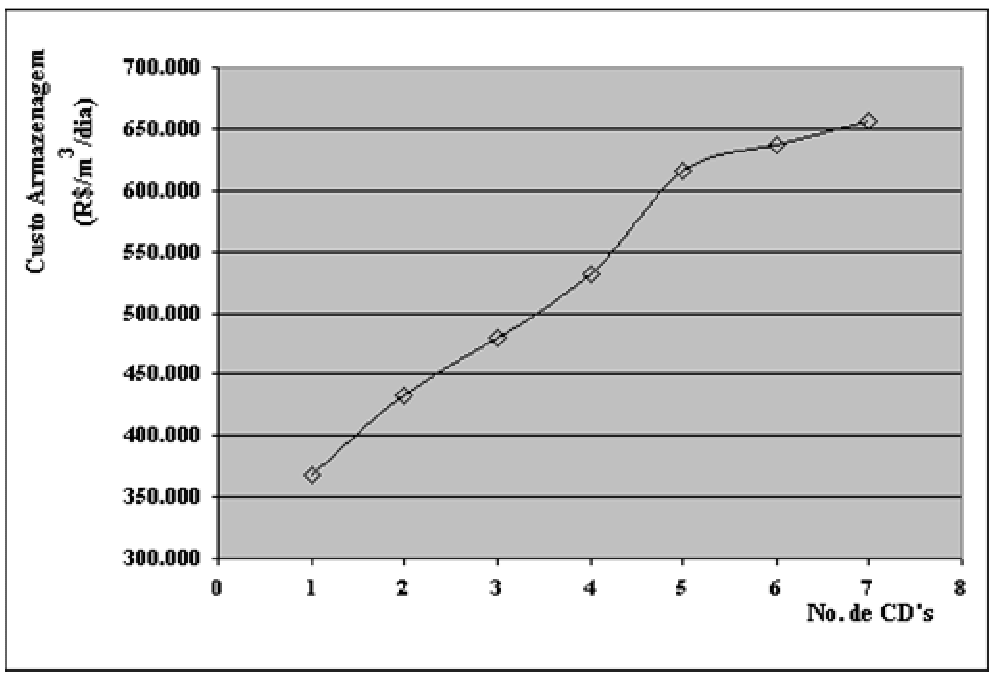

Figura 7.42 - Custos de armazenagem vs. No. de CD´s

7.3.3. Modelo de Custos de Vendas Perdidas

Conforme visto na seção 6.4, os custos de vendas perdidas correspondem aos ganhos não auferidos por receitas que não se concretizaram devido à não conformidade com o nível de serviço estabelecido.

Deve-se, portanto, inicialmente identificar o nível de inconformidade dos pontos de demanda da área de estudo. Este dado em conjunto com outros parâmetros (índice de vendas perdidas, custo do produto e margem de comercialização) irão possibilitar avaliar os custos de vendas perdidas.

O nível de serviço definido estabelece que a distância de um ponto de demanda ao CD que o atende deve ser de no máximo 240 km (seção 7.2.1). Pode-se, então,

avaliar para cada quantidade de CD's analisada a incidência de pontos com distâncias superiores a este valor e somar os volumes desses pontos. Com isto tem-se o volume de inconformidade.

Estes resultados vão apresentar variações entre a solução obtida com o modelo MILP e a solução da meta-heurística SA, pois o posicionamento dos CD's apresentou pequenas diferenças nas duas situações. Foram feitas assim duas análises, conforme apresentado nas tabelas 7.14a e 7.14b.

Tabela 7.14a - Volume acima de Dmáx*
Solução via MILP

No. de CD's	Volume (m^3)	% da Demanda
7	2.084	4,6%
6	2.134	4,7%
5	2.393	5,2%
4	4.992	10,9%
3	6.522	14,3%
2	12.400	27,2%
1	17.975	39,4%

* Dmáx = 240 km do CD de atendimento

Tabela 7.14b - Volume acima de Dmáx*
Solução via SA

No. de CD's	Volume (m^3)	% da Demanda
7	2.089	4,6%
6	2.134	4,7%
5	2.393	5,2%
4	4.881	10,7%
3	5.435	11,9%
2	11.314	24,8%
1	17.367	38,0%

* Dmáx = 240 km do CD de atendimento

A partir destes dados é possível agora estimar os parâmetros do modelo de inconformidade, conforme definido na seção 6.4. As figuras 7.43a e 7.43b

apresentam as curvas e os modelos obtidos para as soluções via MILP e via SA, podendo-se constatar o alto grau de aderência dos dados aos modelos propostos.

Em relação aos índices de vendas perdidas, não se dispunha de dados confiáveis, sabendo-se apenas que deveriam ser valores em torno de 10% a 15% do volume de inconformidade. Para tratar a questão foram então propostos três cenários: perdas de 5%, 10% e 15%. Em cada um destes cenários foi estimado o custo de vendas perdidas.

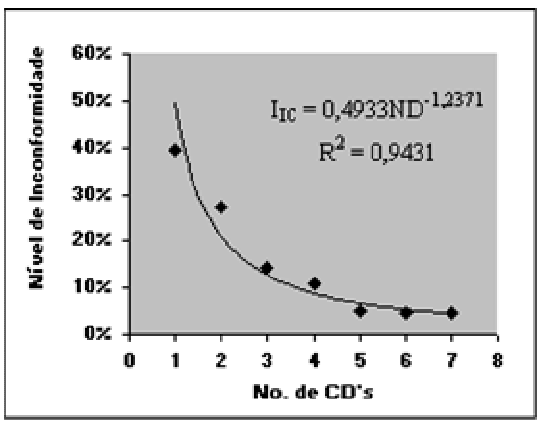

Figura 7.43a – Modelo de inconformidade - MILP

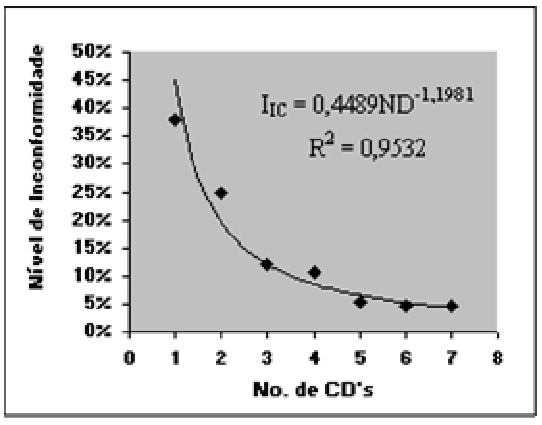

Figura 7.43b – Modelo de inconformidade – SA

Os outros parâmetros que são necessários para a avaliação de vendas perdidas são: o custo unitário do produto (C_{UN}), que neste caso era de R$943,91, e a margem de comercialização (M_C), que foi identificada como sendo de 14,5%.

Com estes parâmetros e os modelos de inconformidade definidos pode-se aplicar a equação 6.28, obtendo-se as estimativas de custos de vendas perdidas para as diferentes quantidades de CD's testadas. Uma visualização dos resultados dessas estimativas é apresentada nas figuras 7.44a e 7.44b para as soluções via MILP e via SA.

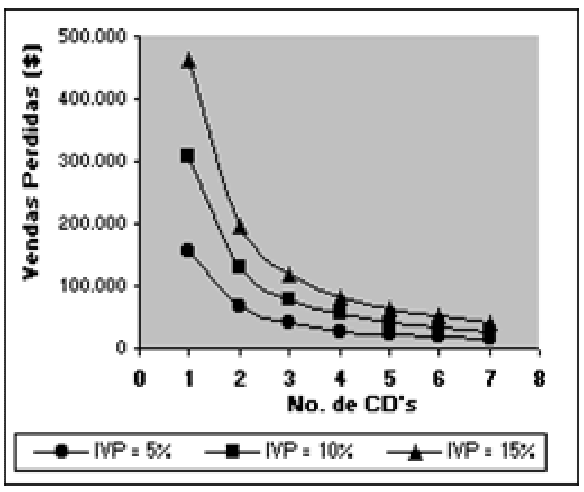

Figura 7.44a – Custos de vendas perdidas - MILP

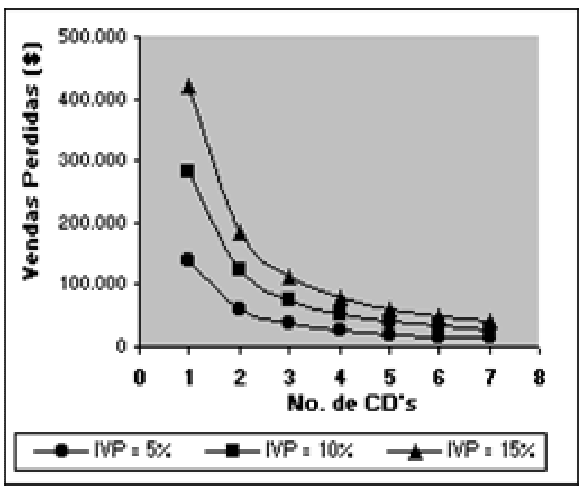

Figura 7.44b – Custos de vendas perdidas – SA

A maior quantidade de CD's favorece o atendimento do nível de serviço, reduzindo o nível de inconformidade. Com isto os custos de vendas perdidas se reduzem de forma acentuada, conforme mostram as curvas.

7.3.4. Determinação da Quantidade “Ótima” de CD’s

Com os custos de estoque, armazenagem e movimentação de materiais e vendas perdidas definidos é possível a determinação do número *p* de CD’s que minimiza os custos logísticos globais.

O problema é resolvido consolidando-se os quatro componentes de custos considerados nesta aplicação prática:

- Transporte (*Fase 1*)
- Estoque (*Fase 2*)
- Armazenagem e movimentação (*Fase 2*)
- Vendas perdidas (*Fase 2*)

Isto deve ser feito para todas as configurações de CD’s analisadas (p=1,2,...7) considerando-se as soluções obtidas através do modelo MILP e da meta-heurística SA.

Foi aqui aplicada a heurística proposta na seção 6.5, identificando-se o valor de p que minimiza o custo logístico total. Este valor ficou entre 2 e 3 CD’s, dependendo do nível de vendas perdidas considerado. Entretanto, como foram simuladas configurações para até sete CD’s, foram consolidados os custos para todas estas configurações testadas.

Os componentes finais de custos logísticos são apresentados nas tabelas 7.15a e 7.15b. E os resultados finais do custo logístico global são apresentados na tabela 7.16.

Tabela 7.15a – Componentes do custo logístico - MILP

No. de CD's	Transporte (R$/mês)	Estoque (R$/mês)	Armazenagem (R$/mês)	Vendas Perdidas I_{VP} = 5% (R$/mês)	I_{VP} = 10% (R$/mês)	I_{VP} = 15% (R$/mês)
1	483.241,04	662.215,61	367.294,14	154.109,42	308.218,83	462.328,25
2	356.711,40	737.920,56	433.353,62	65.377,00	130.753,99	196.130,99
3	306.945,11	787.441,90	480.039,90	39.589,72	79.179,45	118.769,17
4	287.548,75	826.000,28	531.639,55	27.734,53	55.469,05	83.203,58
5	268.040,72	858.770,09	615.930,71	21.044,25	42.088,49	63.132,74
6	266.826,61	888.185,89	637.546,68	16.794,93	33.589,87	50.384,80
7	265.835,36	915.602,58	656.581,45	13.879,01	27.758,01	41.637,02

Tabela 7.15b – Componentes do custo logístico – SA

No. de CD's	Transporte (R$/mês)	Estoque (R$/mês)	Armazenagem (R$/mês)	Vendas Perdidas I_{VP} = 5% (R$/mês)	I_{VP} = 10% (R$/mês)	I_{VP} = 15% (R$/mês)
1	477.839,70	661.986,22	367.294,14	140.238,63	280.477,26	420.715,89
2	354.726,77	737.692,75	433.353,62	61.122,85	122.245,71	183.368,56
3	304.680,74	787.161,94	480.039,90	37.603,54	75.207,07	112.810,61
4	283.970,03	825.614,20	531.639,55	26.640,33	53.280,66	79.920,99
5	266.989,14	858.223,80	615.930,71	20.390,68	40.781,36	61.172,04
6	258.400,43	887.425,20	637.546,68	16.389,46	32.778,92	49.168,38
7	256.489,91	914.573,21	656.581,45	13.625,60	27.251,20	40.876,81

Tabela 7.16 – Custo logístico global– Soluções MILP e SA

No. de CD's	MILP I_{VP} = 5% (R$/mês)	MILP I_{VP} = 10% (R$/mês)	MILP I_{VP} = 15% (R$/mês)	SA I_{VP} = 5% (R$/mês)	SA I_{VP} = 10% (R$/mês)	SA I_{VP} = 15% (R$/mês)
1	1.666.860,21	1.820.969,63	1.975.079,04	1.647.358,70	1.787.597,33	1.927.835,96
2	1.593.362,58	1.658.739,57	1.724.116,57	1.586.896,00	1.648.018,86	1.709.141,71
3	1.614.016,62	1.653.606,35	1.693.196,07	1.609.486,11	1.647.089,64	1.684.693,18
4	1.672.923,11	1.700.657,64	1.728.392,16	1.667.864,11	1.694.504,44	1.721.144,77
5	1.763.785,76	1.784.830,01	1.805.874,26	1.761.534,33	1.781.925,01	1.802.315,69
6	1.809.354,12	1.826.149,05	1.842.943,98	1.799.761,78	1.816.151,24	1.832.540,70
7	1.851.898,40	1.865.777,40	1.879.656,41	1.841.270,18	1.854.895,78	1.868.521,38

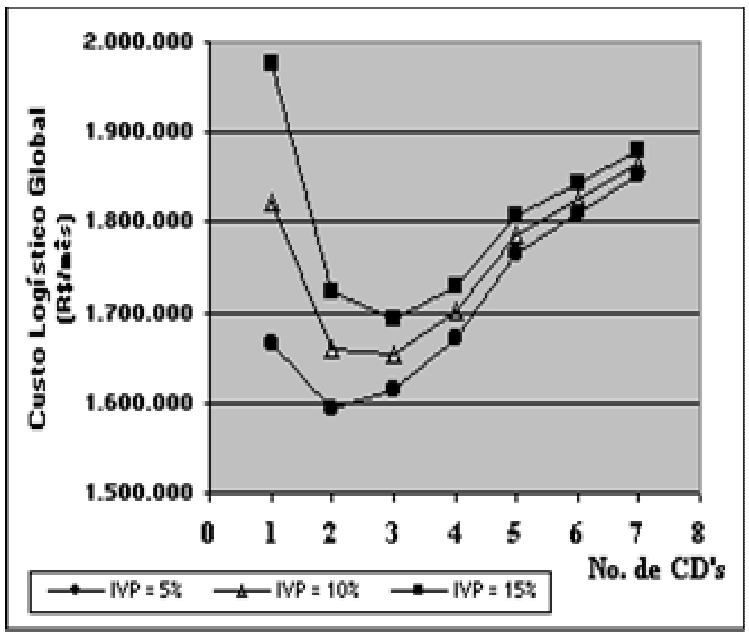

Figura 7.45a – Custo logístico global – MILP

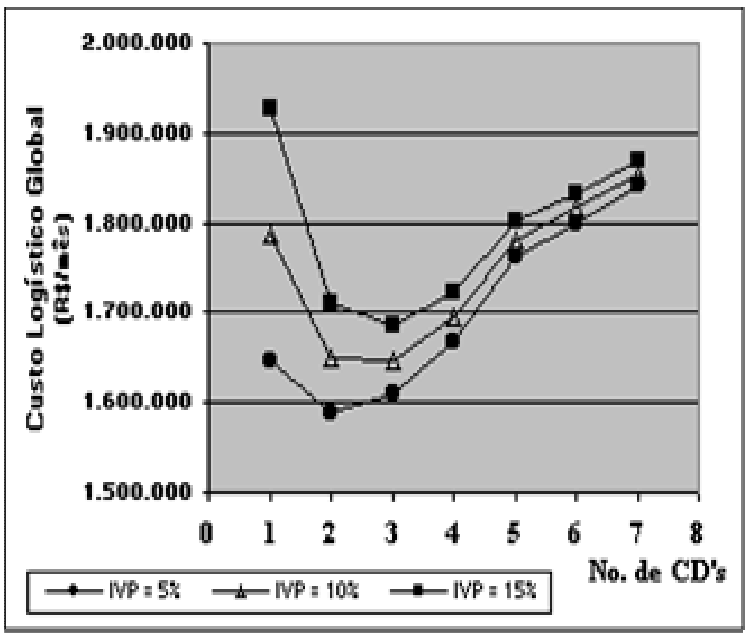

Figura 7.45b – Custo logístico global – SA

Definido o valor “ótimo” p, os CD´s deverão se localizar nos locais candidatos já definidos para p CD*´s na Fase 1.*

8. CONSIDERAÇÕES FINAIS, CONCLUSÕES E RECOMENDAÇÕES

Conforme discutido nos capítulos iniciais, o problema da localização é um dos mais importantes em termos de decisões estratégicas logísticas. O caso dos centros de distribuição objeto desta pesquisa não é exceção. Todos os resultados obtidos no trabalho mostraram, que uma vez definida a quantidade e posicionamento dos CD's, toda o desempenho da operação fica efetivamente condicionado a esta decisão. Os níveis de estoque serão diretamente influenciados, os custos de armazéns serão maiores ou menores dependendo dos locais e quantidades de CD's definidos, as vendas perdidas poderão crescer ou diminuir em face da maior ou menor proximidade entre CD's e clientes, e os custos de transporte, incluindo transferências e distribuição, também sofrerão o impacto direto desta decisão. Os principais custos logísticos, portanto, recebem a influência direta da localização de CD's. É uma decisão, assim, que deve passar por um profundo processo de análise, pois pode, inclusive, determinar a viabilidade econômico-financeira de uma operação.

Este foi o assunto tratado neste trabalho: a localização de centros de distribuição.

Vários aspectos foram estudados ao longo desta pesquisa, sempre com o objetivo de se procurar propor algum tipo de contribuição no sentido de que a modelagem desenvolvida se aproximasse mais da realidade que tenta representar, gerando soluções que realmente se ajustem às situações práticas do mundo real.

Neste sentido, poder-se-ia destacar o conjunto de tópicos relacionados a seguir, como sendo os principais pontos abordados no decorrer da pesquisa:

- Subdivisão do problema da localização

  Considerando-se neste caso dois aspectos mais importantes:

  - a conceituação e fundamentação do princípio das duas *Fases* e

- o desenvolvimento de metodologia contemplando todo o processo.

A formulação da metodologia, em suas *Fases 1 e 2*, exigiu o tratamento e desenvolvimento de uma série de aspectos que merecem ser salientados:

- *Fase 1*:
  - Heurística de *clusterização;*
  - Heurística de seleção de candidatos;
  - Uso de meta-heurística SA em *clusterização;*
  - Uso de meta-heurística SA na localização final de CD's;
  - Uso de *modelo do transporte* em conjunto com meta-heurística SA;
- *Fase 2*:
  - Modelagem do custo de estoque
    - . considerando comportamento não linear;
    - . considerando quantidade e localização de CD's;
  - Modelagem do custo de armazenagem e movimentação
    - . considerando comportamento não linear;
    - . considerando economias de escala;
    - . considerando quantidade e localização de CD's;
  - Modelagem do custo de vendas perdidas
    - . considerando comportamento não linear;
    - . considerando quantidade e localização de CD's;
  - Nível de serviço definindo contornos para a modelagem.

Incluindo a subdivisão do problema, tem-se assim neste conjunto, pelo menos dez aspectos que surgem como resultados principais desta pesquisa.

Uma discussão desses vários tópicos é conduzida a seguir.

Dentre os aspectos estudados, há um que definiu um norte para a linha da pesquisa aqui desenvolvida. Este primeiro grande aspecto a ser tratado nestas considerações finais, diz respeito à subdivisão do problema da localização.

Conforme salientado nos capítulos 1 e 2, o problema da localização vem sendo tratado na maior parte dos casos, de forma a considerar simultaneamente a questão da quantidade e da localização das instalações. E isto se dá de forma mais intensa quando a solução se desenvolve através de modelos de otimização do tipo MILP.

Com a proposta de subdivisão tem-se a *Fase 1,* que trata da localização propriamente dita, para diferentes quantidades de CD's, e a *Fase 2,* que aborda a quantificação de CD´s nas localizações "ótimas" definidas na *Fase 1*.

A idéia da subdivisão em si não é necessariamente uma inovação, mas não se identifica na literatura uma metodologia abrangente e mais aprofundada para desenvolver esta idéia. Além disto, quando se trata de solução do problema via MILP, o que geralmente se encontra é realmente solução em uma única fase. Aqui, portanto, pode se concluir que foi oferecida uma contribuição ao assunto.

A pesquisa procurou inicialmente estruturar adequadamente a idéia de subdivisão, definindo aspectos a tratar em cada fase e estabelecendo os passos a percorrer nessas fases. Além disso, foi desenvolvida uma metodologia completa abrangendo do início ao fim todo o problema da localização, considerando seus vários aspectos com um razoável nível de detalhamento. Os vários aspectos foram discutidos e fundamentados, bem como, definidos procedimentos para sua execução. Estes procedimentos foram testados e, pelos resultados, poder-se-ia concluir, inclusive pela viabilidade de sua implementação em um sistema informatizado. Entre a forma de apresentação das propostas desta pesquisa e a especificação de um sistema para a sua automatização já se tem uma parte razoável do caminho percorrido.

Pode-se concluir ainda que a subdivisão propiciou um melhor aproveitamento das características favoráveis ao uso de modelos MILP e criou caminhos para suprir suas deficiências. Na *Fase 1* procura-se aproveitar o que há de melhor nos modelos de otimização do tipo MILP, propondo, além disso, mecanismos para torná-los ainda mais eficazes e eficientes. Enquanto que, na *Fase 2,* busca-se propiciar uma considerável flexibilidade em termos dos modelos a empregar na representação dos componentes de custo logístico lá tratados, dando margem para que diferentes tipos de modelos com variados graus de sofisticação possam ser empregados na solução. A metodologia proposta suporta com relativa facilidade os diferentes níveis e tipos de modelagem que se deseje empregar na solução dos componentes de custo contemplados naquela *Fase 2.*

Sobre a *Fase 1,* a metodologia procurou suprir uma lacuna detectada na literatura no que diz respeito à seleção de locais candidatos a CD's. Estes se, constituem em peça chave para a garantia de obtenção de uma solução ótima de um modelo de otimização do tipo MILP que realmente represente com realismo a situação prática em estudo. A literatura tem tratado de forma basicamente qualitativa o assunto dando margem a que as decisões fiquem em muitos casos a cargo de julgamento pessoal. Os locais candidatos a CD's se constituem em soluções viáveis para o modelo de otimização. Se estas soluções viáveis não são bem escolhidas, a solução ótima estará definitivamente comprometida.

A pesquisa propõe uma heurística para definir os "melhores" locais candidatos a CD's. Para se chegar a esse conjunto de "melhores" candidatos foi necessário, na verdade, o desenvolvimento de três heurísticas. Uma primeira que faz a *clusterização* dos pontos de demanda, preparando a área de estudo para o processo de seleção. Em seguida, duas heurísticas para a seleção final: raio máximo e ângulo

máximo. Tem-se, assim, um processo totalmente quantitativo para a definição de candidatos a CD's. Vale salientar que neste processo estão sendo aproveitados instrumentos que em geral são utilizados para outros fins, com o objetivo aqui de direcionar a escolha dos candidatos.

Uma inovação que a heurística de *clusterização* utiliza é o fato de trabalhar com o conceito de homogeneidade em termos de momento de transporte, para estabelecer zonas na área de estudo. Os modelos MILP normalmente usam este conceito, considerado em termos de custos, como uma parcela da função a ser minimizada. Aqui se usa o conceito na busca de uma definição de áreas homogêneas em relação ao objetivo da otimização.

Outro aspecto a destacar é idéia de uso das MPT's para criar uma área de busca aos "melhores" candidatos a CD's. Aqui se trabalha com o tradicional método do centro de gravidade, porém voltado para uma nova finalidade. O centro de gravidade é uma posição natural para um CD em uma zona. Por que então não usar esta posição como ponto de partida para um processo de busca das "melhores" localizações em cada zona? Com esta inovação é possível agora situar em uma rede logística as "melhores" posições candidatas a CD's, e não no espaço euclidiano considerado no método do centro de gravidade.

Uma parte importante da pesquisa em termos de aplicação de técnicas avançadas foi o trabalho com a meta-heurística *simulated annealing*. O uso de uma meta-heurística propicia de uma forma geral um aperfeiçoamento das soluções obtidas com uma heurística e dá um maior grau de confiabilidade aos resultados, e isto é o que se procurou fazer nesta pesquisa. Os resultados obtidos com a SA mostraram a sua efetividade para o trabalho em conjunto com as heurísticas propostas.

A SA foi utilizada inicialmente como um complemento à heurística de *clusterização,* propiciando uma melhoria do processo como um todo. Posteriormente foi empregada também para gerar as localizações finais dos CD's (localizações "ótimas"), como uma alternativa ao uso do modelo MILP. O uso da SA foi adotado aqui também, pois inicialmente se observou que repetidamente as configurações geradas pela heurística de seleção dos "melhores" candidatos apresentavam um formato tipicamente circular, envolvendo praticamente toda a área de estudo. Posteriormente, ao se comparar essas configurações com os resultados da solução ótima do modelo MILP, percebeu-se que uma vez definidos os "melhores" candidatos, não se estava muito distante da solução ótima. Isto deu ensejo e motivação para que se partisse diretamente para a solução "ótima" do problema, sem o uso do modelo MILP, propondo um novo processo para atingir tal solução. Neste ponto surgiu a SA novamente, como uma alternativa de ferramenta eficaz para buscar tal solução "ótima".

Sobre esta aplicação da SA, um aspecto que merece ser comentado é que para identificar as localizações "ótimas" dos CD's foi introduzido no fluxo interno da SA uma aplicação clássica de programação linear, que é o modelo de transporte (Novaes, 1978). A SA, através de simulações, procura os melhores fluxos para o processo de distribuição, definindo assim as necessidades de volume em cada CD. O modelo de transporte é então utilizado para otimizar os fluxos de transferência de carga das fontes de produção ao conjunto de CD's em teste. De forma intermitente, a SA vai gerando novos conjuntos de CD's a testar O modelo de transporte desta forma resolve sua parte do problema e retorna sua solução ao fluxo normal da SA. Esta, por sua vez, propõe uma alternativa para a outra parte da questão. Este resultado da SA volta a ser dado de entrada para o modelo de transporte e o

processo continua assim de forma iterativa, até que o número máximo de iterações seja alcançado.

Tem-se, portanto, um trabalho conjunto de um modelo de transporte aninhado dentro do fluxo de uma SA. Isto é algo que deve ser salientado, pois em princípio está-se propondo algo novo aqui em termos de uso das técnicas. Esta proposta ao ser testada funcionou da forma esperada nas aplicações desenvolvidas.

Na *Fase 2* da metodologia, o que se buscou ao longo de toda a pesquisa foi trazer à modelagem uma maior dose de realismo. Este foi seu objetivo principal. Como visto nos capítulos iniciais, as soluções obtidas unicamente através um modelo MILP trazem em si a condição de trabalhar com modelos lineares, o que na maior parte dos casos não ocorre no mundo real. Assim, esta proposta de tratar cada componente de custo com um submodelo próprio, de forma externa ao modelo de otimização, permite flexibilidade suficiente para gerar modelos com considerável proximidade das situações reais.

O primeiro submodelo desenvolvido foi o de custos de estoques. Neste caso, foram considerados todos os seus componentes, incluindo os estoques reguladores, estoques de segurança e estoques em trânsito. Os modelos gerados consideraram de forma ampla as teorias existentes sobre o comportamento desses estoques, agregando em um único modelo diferentes proposições que tratam cada um dos componentes. A inclusão no modelo dessas abordagens distintas se deu de forma natural, sem a preocupação de desenvolver aproximações para comportamentos lineares. Além disto, a modelagem levou em consideração a variação dos níveis de estoque com a quantidade de pontos de estocagem, tanto no que tange a estoques em armazém quanto a estoque em trânsito. Obteve-se, assim, um modelo abrangente para tratar a questão.

Sobre o submodelo de custos de armazenagem alguns avanços significativos podem ser registrados. Uma primeira consideração que não é comum se encontrar neste tipo de modelagem é o fato de que o modelo assume que os custos de armazenagem são fixos dentro de certos limites de volumes a armazenar. Ultrapassado o limite, o custo fixo cresce segundo módulos de área. O modelo proposto trabalha assim com um processo de evolução dos custos fixos segundo patamares. Esta constância do custo, entretanto, não é admitida como sendo absoluta dentro de um patamar. O modelo trabalha com a possibilidade de ocorrência de dois efeitos em um patamar: o custo médio e o custo total. O primeiro foi sempre mantido fixo em cada patamar, mas o custo total cresce dentro do patamar com a evolução do volume a armazenar. Isto traz maior realismo à modelagem, já que na prática, dada uma estrutura de armazém nem todos os componentes de custo permanecem efetivamente constantes à medida que o volume em armazém se modifica. O modelo deu suporte ainda à possibilidade de ocorrência de economias de escala. Aumentos significativos de volume que geram mudanças de patamares de custos poderiam dar ensejo ao uso de novas tecnologias, com maior produtividade e gerando economias de escala. Esta mudança no padrão de custo é perfeitamente suportada pelo modelo.

Um último aspecto a salientar é o fato de que este submodelo também trata a variação da quantidade de CD's. Trabalhando em conjunto com o modelo de estoques, aloca os custos de armazenagem, em diferentes patamares em função da área necessária para acomodar os estoques gerados pelo submodelo de estoques. Foi assim construído um modelo abrangente que considera aspectos variados da problemática de armazenagem.

O último submodelo contemplado na *Fase 2* da metodologia foi o modelo de custos de vendas perdidas. Alguns avanços também podem ser registrados neste caso.

As vendas perdidas são quase sempre citadas na literatura de logística, mas nem sempre são apresentadas modelagens para o processo. Este trabalho procurou propor uma ferramenta concreta para a questão, construindo funções matemáticas que prevêem o comportamento da variável à medida que o número de CD's evolui, e simultaneamente associando os resultados a decisões logísticas anteriores.

Isto foi alcançado relacionando as taxas de vendas perdidas ao grau de não cumprimento dos níveis de serviço estabelecidos, denominados aqui de níveis de inconformidade. Define-se assim um modelo de inconformidade cujos parâmetros são estimados com os resultados da localização ótima dos CD's. Consegue-se, desta forma, criar uma relação matemática entre decisões de localização e perdas de receitas devido à inconformidade. Integra-se em um mesmo modelo decisões de localização e o nível de serviço estabelecido, que por sua vez irão gerar estimativas de taxas de vendas perdidas. O modelo resultante aborda assim aspectos relevantes para a questão e faz isto de forma realista, uma vez que se utiliza dos tipos de equações matemáticas que melhor aderem ao comportamento do processo, independentemente de serem equações lineares ou não.

Esta consideração do nível de serviço é algo a se destacar na pesquisa, pois nem sempre é um aspecto tratado adequadamente nas modelagens de processos logísticos. Deve-se notar que o nível de serviço surge em dois pontos da modelagem: inicialmente como base para a heurística de *clusterização,* fazendo parte do processo de definição do número de *clusters,* e mais tarde, volta a ser considerado no modelo de vendas perdidas.

Deve-se mencionar aqui também o experimento desenvolvido para validação das heurísticas propostas na *Fase 1* da metodologia. A montagem do próprio experimento já procurou de certa forma resgatar alguns aspectos um tanto quanto "esquecidos" na literatura, que é a simulação descritiva proposta por Saliby (1989). Este tipo de técnica de simulação garante que toda a gama de variação da variável simulada seja contemplada, o que nem sempre ocorre na simulação aleatória clássica. Uma vez montado o experimento destaca-se o fato de que a solução ótima de um modelo MILP desagregado que considerou cada ponto de demanda de forma individual coincidiu com a solução obtida pelo modelo MILP agregado (tratando os pontos em zonas) que foi alimentado com os resultados das heurísticas propostas. E ainda a solução "ótima" gerada pela SA também coincidiu com as outras duas. Tem-se a segurança, portanto, de que as heurísticas conduziram o problema à sua efetiva solução ótima. O que já é uma indicação de que a metodologia aponta para uma direção correta.

Saliente-se ainda que o experimento propiciou a comparação dos resultados das heurísticas propostas com várias outras formas de agregação dos pontos de demanda, obtendo-se de forma consistente os melhores resultados através das heurísticas.

Finalmente, neste balanço final cumpre mencionar a aplicação prática desenvolvida que propiciou o uso de toda a metodologia incluindo as *Fases* 1 e 2.

Os dados contemplavam uma vasta área de estudo, com mais de duzentos pontos de demanda, assumindo proporções, portanto, de um problema real considerável, que necessita de ferramentas eficientes de modelagem para uma tomada de decisão segura.

Com essas características do problema a metodologia se mostrou adequada para gerar as decisões de localização. Particularmente as heurísticas da *Fase 1* se mostraram eficientes e eficazes gerando um pequeno número de “melhores” candidatos, apesar da grande quantidade de pontos na rede logística que poderiam se tornar candidatos a CD’s. Outro fato importante da aplicação prática foi a proximidade entre as soluções das heurísticas e as soluções melhoradas pela meta-heurística. Isto foi uma indicação da eficácia das heurísticas propostas.

Em termos de proximidade, deve-se lembrar também das soluções ótimas de localização obtidas através do modelo MILP e através da SA. Neste caso, além das soluções quase coincidentes, constatou-se até uma certa vantagem para os resultados através da SA, mostrando que o tratamento desagregado da SA pode ser uma vantagem a considerar, apesar de ser uma técnica que não garante a otimização.

No que tange aos resultados obtidos com os modelos de custos da *Fase 2,* deve-se mencionar que se apresentaram com a coerência esperada, gerando comportamentos absolutamente justificáveis em todos os aspectos, dando desta forma, uma indicação do direcionamento correto das hipóteses consideradas nos modelos e do equacionamento proposto.

Esta pesquisa não teve a intenção de ser exaustiva. Assim, algumas reflexões que foram se consolidando ao longo do trabalho podem ficar agora como recomendações para trabalhos futuros.

Um primeiro aspecto a ser tratado seria a questão do uso da *simulated annealing*. Esta meta-heurística foi utilizada na pesquisa e apresentou bons resultados, mas seria interessante do ponto de vista de comparação de resultados avaliar o comportamento de outras técnicas. Assim, a aplicação de outros tipos de meta-

heurística dentro do mesmo contexto em que a *simulated annealing* foi empregada nesta pesquisa seria um campo para novas pesquisas.

A modelagem do comportamento dos custos de transporte é outro ponto que também pode ser trabalhado. Há campo ainda para aprimoramento desses modelos. Um tratamento semelhante ao proposto para os custos de armazenagem com evolução descontínua é uma alternativa de avanço das pesquisas com o modelo, já que desconto por quantidade é uma prática encontrada no mercado. Desenvolver este tratamento de forma realista em conjunto com um modelo MILP é um desafio a ser vencido.

Outro aspecto a considerar é o fato de que esta pesquisa está fundamentalmente tratando de um produto único, se bem que em muitos aspectos poderia ser aproveitada para múltiplos produtos. De qualquer modo, tratar de forma explícita o caso de múltiplos produtos seria um avanço a ser pensado para novas pesquisas.

O tratamento dinâmico das demandas é algo também não contemplado de forma explícita neste trabalho. Na verdade, foi estabelecido desde o princípio que a pesquisa se ocuparia exclusivamente de modelos estáticos, podendo, entretanto, ser utilizada dinamicamente com o uso de cenários. Considerar, entretanto, esta evolução temporal da demanda de forma direta nos modelos é algo que pode trazer novos e estimulantes resultados, especialmente para os modelos de custos utilizados na segunda fase da metodologia.

Há, portanto, campo aberto para novas pesquisas.

Como palavras finais espera-se que com este trabalho tenha-se conseguido demonstrar a possibilidade de se trazer avanços à modelagem do problema da localização com a introdução de novos procedimentos e com tratamentos mais realistas de determinados comportamentos. Espera-se que com isto tenha-se feito

alguma contribuição ao estado da arte no assunto e, assim, com a seminação deste campo, eventualmente algo possa vir a germinar em algum *local* no futuro.

REFERÊNCIAS BIBLIOGRÁFICAS

Aarts, E.H. L. e Van Laarhoven, P. J. M. (1985). “Statistical cooling: a general approach to combinatorial optimization problems“. *Philips Journal of Research. Vol. 40, No. 4, p.193-226.*

Abdinnour-Helm, S. (2001). “Using simulated annealing to solve the p=hub median problem“; *International Journal of Phisical Distribution & Logistics Management, Vol.31, Iss.3; p-203-214.*

Adler, H.A. (1978). “Avaliação econômica dos projetos de transporte: metodologia e exemplos“; Livros Técnicos e Científicos Editora. Rio de Janeiro.

Ahuja, R. K., Magnanti, T. L., Orlin, J. B. (1993). “Network Flows“. *Prentice Hall., New Jersey*

Baass, K. (1981). “Design of zonal systems for aggregate transportation planning models“; *Transportation Research Record, n.807, p.1-6*.

Ballou, R.H. (1991a). “The accuracy in estimating truck class rates for logistical planning“; *Transportation Research A, vol.25 A, n.6, p.327-337*.

_____ (1991b). “Measuring transport costing error in customer aggregation for facility location “; *Transportation Journal, Spring, p.49-59*.

_____ (1992). “Business logistics management“; *Prentice Hall International Editions,* N.J.

_____ (1997). “Business logistics – Importance e some research opportunities“; *Gestão e Produção, V 4, No.2*.

Barros, L. (1997). “A global view of industrial logistics“; *Gestão e Produção, V 4, No.2*.

Bender, P. S. (1997). “How to design an optimal worldwide supply chain“; *Supply Chain Management Review, Spring 1997 – Vol.1, N.1.*

Blanchard, B.S. (1992). “Logistics engineering and management“. *Prentice Hall International Inc.* 556 p.

Blumenfeld, D.E., Burns, L.D., Diltz, J.D., Daganzo, C.F. (1985). “Analyzing trade-offs between transportation, inventory and production costs on freight networks“; *Transportation Research B, 19 B, No.5, p. 361-380.*

Boaventura Netto, P. (1977). "Teoria dos grafos" Coordenação dos Programas de Pós Graduação de Engenharia - Universidade Federal do Rio de Janeiro – Comissão de Publicações. Rio de Janeiro, RJ. 284 p.

Bowersox, D.J.; Closs, D.J. (1997). “Brazilian logistics – a time for transition“; *Gestão e Produção, V 4, No.2*.

Bowersox, D.J.; Closs, D.J.(2001). "Logística empresarial: o processo de integração da cadeia de suprimento" Editora Atlas, 1ª Edição. São Paulo, SP. 594 p.

Bradley, G. H., Brown, G. G., Graves, G. W. (1977). “Design and implementation of large-scale primal transshipment algorithms“. *Management Science, vol. 24 , n.1, p. 1 -34.*

Brandeau, M. L., Chiu, S.S. (1989). “An overview of representative problems in location research“; *Management Science, vol. 35 , n.6, p.645-674.*

Brown, G. G., Graves, G. W., Honczarenko, M. D. (1987). “Design and operation of a multicommodity production/distribution system using primal goal decomposition“. *Management Science, vol. 33 , n.11, p. 1469 -1480.*

Brown, G. G., Olson, M. P. (1994). “Dynamic factorization in large-scale optimization“; *Mathematical Programming, n.64, p.17 - 51.*

Camm, J.D.; Chorman, T. E.;. Dill, F. A; Evans, J. R.; Sweeney, D. J.; Wegryn, G. W. (1997). “Blending OR/MS, Judgment and GIS: Restructuring P&G’s Supply Chain“ *Interfaces 27: Jan-Fev , p.128-142*.

CEL/COPPEAD (2001) –. “Armazenagem em grandes empresas brasileiras“ *Centro de Estudos Logísticos - Universidade Federal do Rio de Janeiro*.

Chen,CX.T. (2001). “A fuzzy approach to select the location of the distribution center“; *Fuzzy Sets and Systems, – 118, p. 65-73.*

Chopra, S, Meindl,P. (2001) "Supply chain management: strategy, planning and operation" Prentice Hall, Upper Sadle River, NJ.

Christopher, M. (1997) "Logística e gerenciamento da cadeia de suprimentos" Editora Pioneira, São Paulo, SP. 240 p

Cohen,M. A., Lee, H.L. (1988). “Strategic analysis of integrated production-distribution systems: models and methods“; *Operations Research, – Vol. 36, No. 2 , Mar-Apr, p. 216-228.*

Cohen,M. A., Moon, S. (1991). “An integrated plant loading model with economies of

scale and scope"; *European Journal of Operational Research, No.502 ,p. 266 - 279.*

Colorni,A., Dorigo, M., Maffioli, F., Maniezzo, V., Righini, G., Trubian, M. (1997). "Heuristics from nature for hard combinatorial optimization problems"; *International Transactions in Operational Research, No.3.1 ,p. 1 - 38.*

Correa, H.L.; Gianesi, I.G.N.; Caon, M. (2001). "Planejamento, programação e controle da produção". Editora Atlas. São Paulo. 452 p.

Coyle, J.J.; Bardi, E.J.; Langley Jr., C.J. (1992). "The management of business logistics". *West Publishing Co.*

Crainic, T.G. e Laporte, G. (1997). "Planning models for freight transportation". European Journal of *Operational Research 97, p. 409-438.*

Daganzo, C.F. (1984). "The length of tours in zones of different shapes"; *Transportation Research B, 18 B, No.2, p. 135-145.*

D'Amico, S. J.; Shoou-Jiun, W.; Batta, R.; Rump, C. M. (2002). "A simulated annealing approach to police district design". *Computers and Operations Research. Vol 29, p. 667-684.*

Daskin, M. S. (1995). "Network and discrete location". *John Wiley & Sons Inc.* N.Y.

Dornier, P.P., Ernst, R., Fender, M., Kouvelis, P. (1998). "Global operations and logistics"; *John Wiley & Sons Inc.* N.Y.

Draper, N., Smith, H. (1981). "Applied regression analysis"; *John Wiley & Sons Inc.* N.Y., 2nd edition. 709 p.

Erenguç, S. S., Simpson, N. C., Vakhari, A. J. (1999). “Integrated production/distribution planning in supply chains: An invited review“; *European Journal of Operational Research, n.115, p.219 - 236.*

Erlenkotter, D. (1978). “A dual-based procedure for uncapacited facility location“. *Operations Research, vol. 26 , n.6, p. 992 - 1009.*

Erlich,P.J. (1982). “Pesquisa operacional“; Editora Atlas , 4a Edição. São Paulo-SP. 279 p.

Fabrycky, W.J.; Thuesen,G. J.; Verma,D. (1998). “Economic decision analysis“; Prentice Hall. 3rd Edition. New Jersey.

Fleury, P.F.;Wanke, P; Figueiredo, K.F.(2000) "Logística empresarial: a perspectiva brasileira" Editora Atlas, 1ª Edição. São Paulo.372 p.

Francis, R.L.; Rushton, G.; Rayco, M.B. (1999). "A synthesis of aggregation methods for multi-facility location problems: strategies for containing error," *Geographical Analysis. p. 67-87,31.*

Geoffrion,A.M., Graves, G.W. (1974). “Multicommodity distribution system design by Benders decomposition“; *Management Science, – Vol. 20, No. 5, 822-844.*

Geoffrion,A.M., Powers, R.F. (1995). “Twenty years of strategic distribution system design: an evolutionary perspective“; *Interfaces, – Vol. 25, N. 5, p.103-127.*

Gualda, N. D. F., (1995) “Terminais de transporte: contribuições ao planejamento e ao dimensionamento operacional”, *Tese de Livre Docência, EPUSP. São Paulo.*

Hajek, B. (1988). “Cooling schedules for optimal annealing“; *Mathematics of Operations Research, Vol. 13, No. 2; p.311-329.*

Hakimi, S.L. (1964). “Optimum locations of switching centers and the absolute centers and medians of s graph“; *Operations Research, 12, p. 450-459.*

Hall, R.W. (1987a). “Direct versus terminal freight routing on a network with concave costs“; *Transportation Research, 21 B, No.4, p. 287-298.*

Harmon, Roy L. (1994). “Reinventando a distribuição“. *Editora Campus.,* Rio de Janeiro, RJ. 419 p.

Isard, W. (1956). “Location and space-economy”. The M.I.T. Press. Cambridge, MA. 350 p.

Johnson, D.S. Aragon, C.R. McGeoch, L.A. Catherine, S. (1989). “Optimization by simulated annealing: an experimental evaluation; part I, graph partioning“; *Operations Research, Nov/Dec; 37, 6, p.865-892.*

_____ (1991). “Opimization by simulated annealing: an experimental evaluation; part II, graph coloring and number partitioning“; *Operations Research, May/Jun; 39, 3, p.378-406.*

Kado, T. M. N. (1998) “Análise do método de SA (simulated annealing) aplicado à roteirização de veículos coletores de resíduos sólidos urbanos domiciliares”, *Dissertação de Mestrado. Escola de Engenharia de São Carlos da USP. São Carlos.*

Lacerda, L. (1999). “Considerações sobre o estudo de localização de instalações“.

Artigos do Centro de Estudos em Logística – COPPEAD - Universidade Federal do Rio de Janeiro, RJ. Disponível em: <http://www.cel.coppead.ufrj.br/fs-public.htm.>. Acesso em 15 de junho de 2001.

Lachtermacher, G. (2002). “Pesquisa operacional na tomada de decisões: modelagem em Excel”. Editora Campus. Rio de Janeiro, RJ. 322 p.

Law, A.M., Kelton, W.D. (1982). “Simulation modeling and analysis”. Mc Graw Hill Book Co. New York, NY. 400 p.

Maculan Filho, N., Pereira, M.V.F. (1980). “Programação linear”. Editora Atlas. São Paulo, SP. 182 p.

Mehrotra, A.; Johnson, E. L. Nemhauser, G. L. (1998). “An optimization based heuristic for political districting“. *Management Science. Vol.44, No.8; p.1100-1114.*

Mentzer, J.T.; Kahn, K.B. (1995). “A framework of logistics research“. *Journal of Business Logistics . Vol.16, No.1.*

Moura, R. A. (1997). “Armazenagem e distribuição física“. *IMAN – Instituto de Movimentação e Armazenagem de Materiais, São Paulo.*

Muyldermans, L.; Cattrysse, D.; Van Oudheusden, D.; Lotan, T. (2002). “Districting for salt spread operations. *European Journal of Operations Research*. *Vol. 139, p. 521-532.*

Norman, S.K.; Rogers, D.F.; Levy, M.S. (1999). “Error bound comparisons for aggregation/desaggregation techniques applied to the transportation problem“. *Computers &Operations research. Vol.26, p. 1003-1014.*

Novaes, A.G. (1978). “Métodos de otimização“. *Editora Edgard Blücher Ltda.,* São Paulo, SP. 463 p.

_____ (1989).“Sistemas logísticos“. Livraria Pioneira Editora, São Paulo. 372 p.

Novaes, A.G.; Graciolli, O. D. (1999). “Designing multi-vehicle delivery tours in a grid-cell format“. *European Journal of Operations Research*. *Vol. 119, p. 613-614.*

Novaes, A.G.; Cursi, J. E. S.; Graciolli, O. D. (2000). “A continuous approach to the designing of physical distribution systems“. *Computers and Operations Research*. *Vol. 27, p. 877-893.*

Nozick, L.K., Turnquist, M.A. (1998). “Integrating inventory impacts into a fixed-charge model for locating distribution centers“. *Transportation Research E, Vol. 34, No.3, 173-186.*

Owen, S.H. e Daskin, M. S. (1998). “Strategic facility location: a review“. *European Journal of Operational Research. No. 111, 423-447.*

Pirlot, M. (1996). “General local search nethods“. *European Journal of Operational Research.* No. 92, p.493-511.

Randhawa, Sabah U. ; West, Thomas M. (1995). “An integrated approach to facility location problems”. Computers and Engineering. *Vol 29, No 1-1, p. 261-265.*

Ross, S.M. (1972). “Introduction to probability models”. Academic Press. New York. 267 p.

Saliby, E. (1989). “Repensando a simulação: a amostragem descritiva”. Editora Atlas. São Paulo. 182 p.

Shapiro,J. F., Singhal, V.M., Wagner,S. N. (1993). “Optimizing the value chain“; *Interfaces, Vol. 23:2, March-April, p.102-117.*

Simchi-Levi, D.; Kaminski, P.; Simchi-Levi, E. (2000) "Designing and managing the supply chain: concepts, strategies and case studies" Mc Graw Hill, Boston, MA.

Spiegel, M. (1994) "Estatística" Makron Books, São Paulo. 643 p.

Steel, R.G.D. e Torrie, J.H.. (1980) "Principles and procedures of statistics" Mc Graw Hill, New York, NY. 633 p.

Tondo, C. M., (1992) “Um modelo matemático para a localização estratégica de terminais de contêineres no interior: aplicação ao estado de São Paulo”, *Tese de Doutorado EPUSP. São Paulo.*

Valente, A M.; Passaglia, E. ;Novaes, A.G. (1997). “Gerenciamento de transporte e frotas“. *Editora Pioneira.,* São Paulo, SP. 215 p.

Vallim F^{o} , A. R. A. (1997). “Projeto de otimização da distribuição“; Via- Rethys Consultoria – Relatório técnico. São Paulo. 144 p.

Viana G., V. R. (1998). “Meta-heurísticas e programação paralela em otimização combinatória“; Universidade Federal do Ceará - UFC Edições. Fortaleza. 248 p.

Vidal, C.J., GoetschalcKx, M. (1997). “Strategic production- distribution models: a critical review with emphasis on global supply chain models“; *European Journal of Operational Research, – Vol. 98, p. 1-18.*

Wanke, P. (2001). “Aspectos fundamentais do problema de localização de instalações em redes logísticas “; *Artigos do Centro de Estudos em Logística* – COPPEAD - Universidade Federal do Rio de Janeiro, RJ. Disponível em: <http://www.cel.coppead.ufrj.br/fs-public.htm.>. Acesso em 10 de janeiro de 2002.

Winston, W.L. (1994). “Operations research: applications and algorithms“; Duxbury Press. Belmont, CA. 3th Edition. 1318 p.

Weber, A. (1929). “Theory of the location of industries“; The University of Chicago Press. Chicago. 7th Edition. 256 p.

Whybark, D.C., Yang, S.. (1996). “Positioning inventory in distribution systems“; *International Journal of Production Economics, – Vol.45, p. 271-278.*

Zhou, G.; Hokey, M., Gen M. (2002). “The balanced allocation of customers to multiple distribution centers in the supply chain network: a genetic algorithm approach”. *Computers and Industrial Engineering*. *V. 43, p. 251-261*.

APÊNDICE A: DIMENSÕES DO MODELO MILP

Neste apêndice são determinadas as dimensões em termos de número de variáveis e número de restrições para o modelo abaixo, à medida que se altera a quantidade de locais candidatos a CD's.

$$Mín \sum_{i,k} c_{ik}.x_{ik} + \sum_{k,j} c_{kj}\, D_j.y_{kj} + \sum_{k} f_k.Y_k \qquad \text{(A1)}$$

Sujeita a:

$$\sum_{k} y_{kj} = 1 \qquad (j = 1,2,.....J)$$. zona é atendida por único CD (A2)

$$\sum_{k} Y_k = p$$. deve-se ter p CD's (A3)

$$y_{kj} - Y_k \leq 0 \qquad (k = 1,2,....K \;;\; j = 1,2,.....J)$$. zona é atendida por CD ativo (A4)

$$\sum_{j} y_{kj} - Y_k \geq 0 \qquad (k = 1,2,.....K)$$. CD ativo atende ao menos uma zona (A5)

$$\sum_{k} x_{ik} \leq O_i \qquad (i = 1,2,.....I)$$. restrições de oferta (A6)

$$\sum_{i} x_{ik} = \sum_{j} D_j.y_{kj} \qquad (k = 1,2.....K)$$. atendimento da demanda (A7)

$$y_{kj} = 0,1 \qquad (k = 1,2,....K \;;\; j = 1,2,....J)$$ 0 = CD não atende zona (A8)
1 = CD atende zona

$$Y_k = 0,1 \qquad (k = 1,2,....K)$$ 0 = CD fechado (A9)
1 = CD aberto

A notação empregada no modelo é descrita a seguir:

. Índices

i = índice para as fontes de produção; i=1,2,.....I; I=número de origens.

j = índice para zonas de demanda; j=1,2,.....J; J=número de zonas.

k = índice para CD's; k=1,2,.....K; K=número de candidatos.

. Função Objetivo

A função objetivo (equação A1) é composta dos três termos analisados abaixo:

1º termo:

c_{ik} = custo unitário de transferência entre a região i de produção e o CD k;

x_{ik} = fluxo do produto entre a região i de produção e o CD k (deve-se ter $x_{ik} \geq 0$)

2º termo:

c_{kj} = custo unitário de distribuição entre o CD k e a zona j de demanda;

D_j = demanda da zona j;

y_{kj} = variável binária de decisão $y_{kj} = 0$ ...CD k não atende zona j;

$y_{kj} = 1$ ...CD k atende zona j;

3º termo:

f_k = custo fixo de localização do CD k;

Y_k = variável binária de decisão $Y_k = 0$ ...CD k é ativo;

$Y_k = 1$ ...CD k não é ativo.

. Restrições:

O_i= oferta de produto na região de produção i;

. **Dimensões do Modelo**

Variando-se o número de candidatos a CD's tem-se os resultados abaixo:

Tabela A1a – Dados de entrada do modelo MILP

No. de Candidatos a CD's:	K =	**6000**	**1000**	**90**
No. de Fontes:	I =	1	1	1
No. de Pontos de Demanda:	J =	6000	6000	6000

Tabela A1b - Condições operacionais

Restrições	No. de Restrições			
A2	J restrições =	6.000	6.000	6.000
A3	1 restrição =	1	1	1
A4	K.J restrições =	36.000.000	6.000.000	540.000
A5	K restrições =	6.000	1.000	90
A6	I restrições =	1	1	1
A7	K restrições =	6.000	1.000	90
	Total Operacionais	**36.018.002**	**6.008.002**	**546.182**

Tabela A1c - Condições de integralidade - variáveis binárias

Restrições	No. de Restrições			
A8	K.J restrições =	36.000.000	6.000.000	540.000
A9	K restrições =	6.000	1.000	90
	Total Binárias	**36.006.000**	**6.001.000**	**540.090**
Total Geral : Operacionais + Binárias		**72.024.002**	**12.009.002**	**1.086.272**

Tabela A1d – Número de Variáveis

Variável	No. de Variáveis			
x_{ik}	I.K variáveis =	6.000	1.000	90
y_{kj}	K.J variáveis =	36.000.000	6.000.000	540.000
Y_k	K variáveis =	6.000	1.000	90
	Total de Variáveis	**36.012.000**	**6.002.000**	**540.180**

APÊNDICE B: DISTÂNCIAS A MPT's – Análise da Hipótese de Normalidade

Este apêndice apresenta o comportamento da distribuição de probabilidades das distâncias de pontos de demanda as suas respectivas MPT's.

A análise foi desenvolvida a partir de um conjunto de dados reais em que se tinham duzentos e vinte e nove pontos de demanda distribuídos por praticamente toda a região sul do Brasil. Estes dados foram utilizados posteriormente na aplicação prática apresentada no capítulo 7.

Foi aplicada a heurística de *clusterização* aos dados gerando um zoneamento e em seguida calculadas as MPT's de cada zona. Foram então construídas distribuições empíricas para diferentes conjuntos de zonas, inclusive para a área total de estudo. Os diferentes testes mostraram que dentre as distribuições analisadas – Gama, Erlang, Normal, Beta e Rayleigh – a que apresentou os melhores resultados na maior parte dos casos foi a Distribuição Normal, conforme pode ser observado pelas figuras B.1a a B.5c apresentadas a seguir.

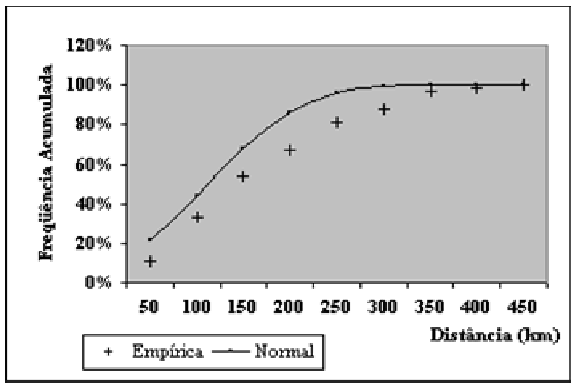

Figura B.1a Distribuição Normal- Todos os anéis

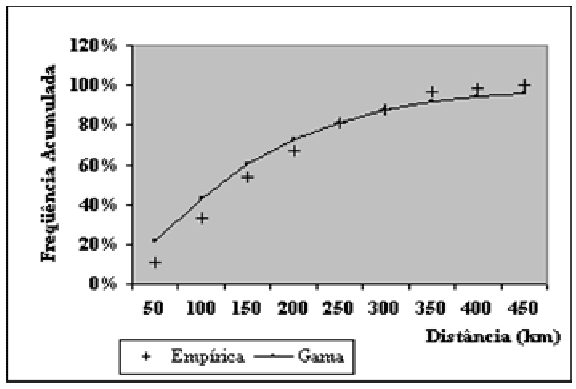

Figura B.1b-Distribuição Gama- Todos os anéis

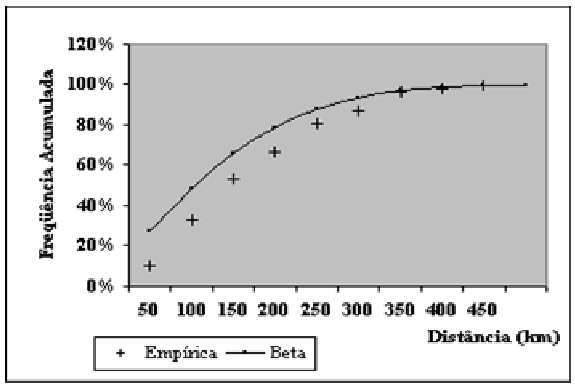

Figura B.1c Distribuição Beta- Todos os anéis

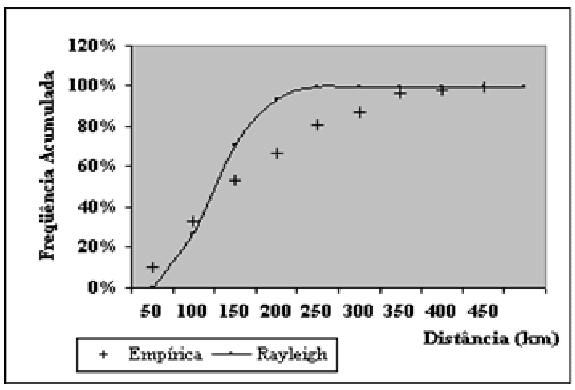

Figura B.1d-Distribuição Rayleigh- Todos os anéis

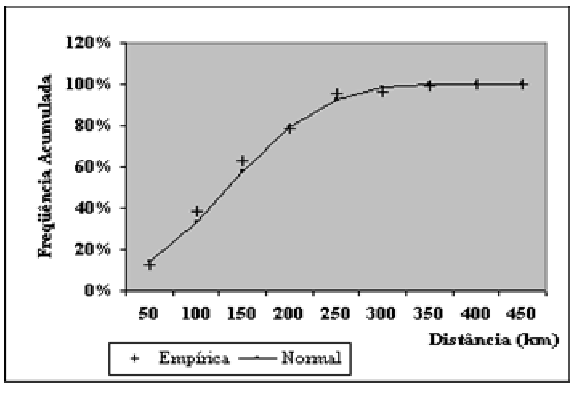

Figura B.2a Distribuição Normal- Anéis 2 e 3

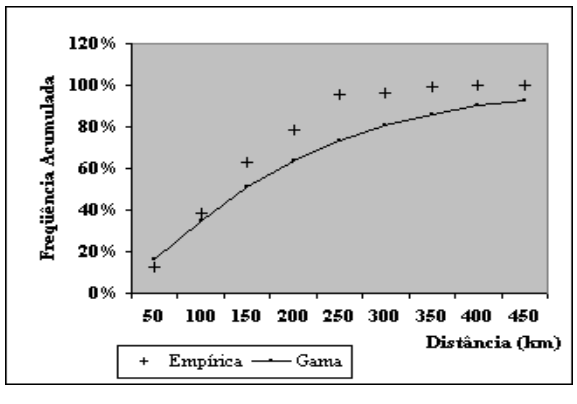

Figura B.2b-Distribuição Gama- Anéis 2 e 3

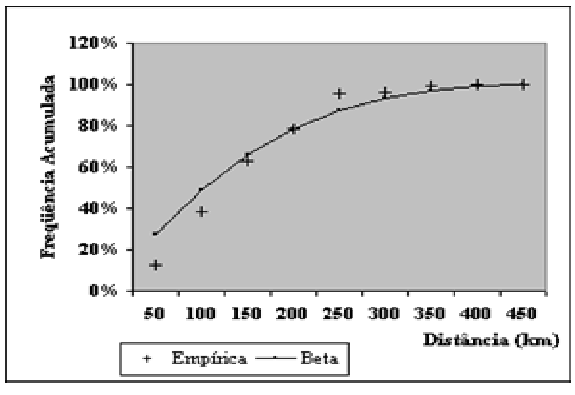

Figura B.2c Distribuição Beta- Anéis 2 e 3

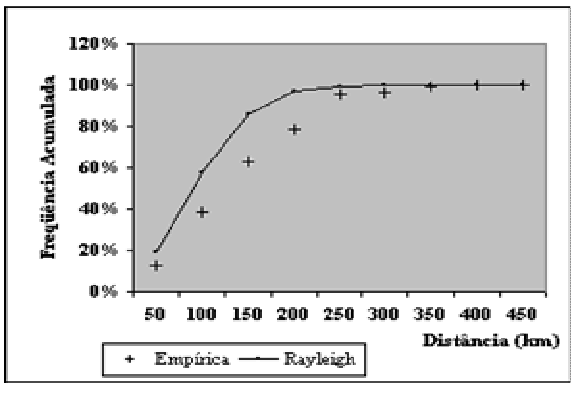

Figura B.2d-Distribuição Rayleigh- Anéis 2 e 3

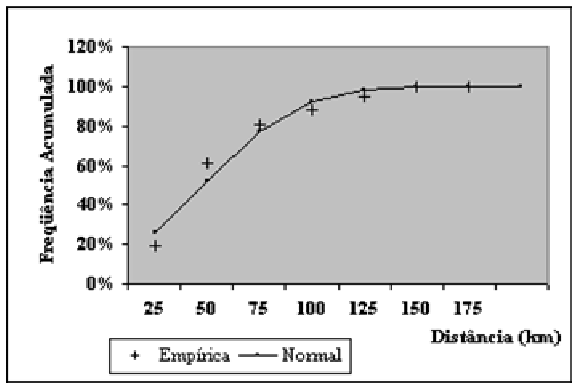

Figura B.3a Distribuição Normal - Anel 1

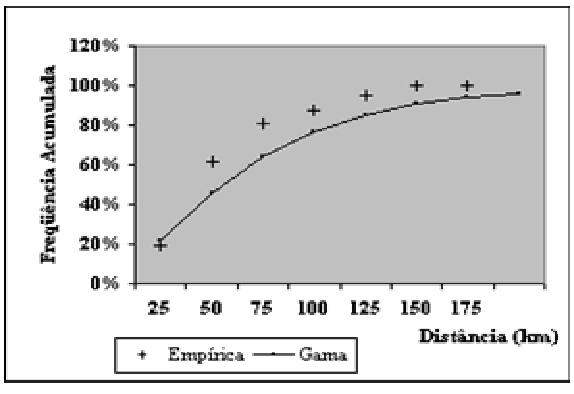

Figura B.3b Distribuição Gama - Anel 1

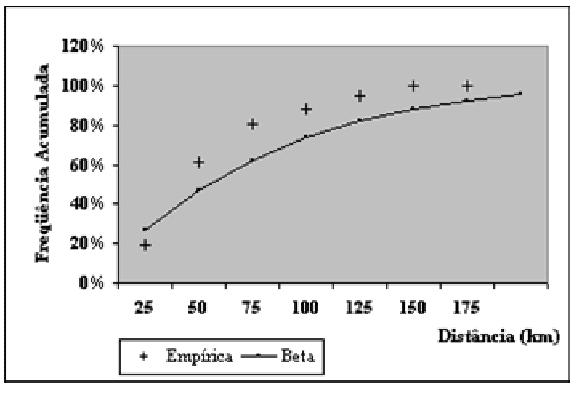

Figura B.3c-Distribuição Beta- Anel 1

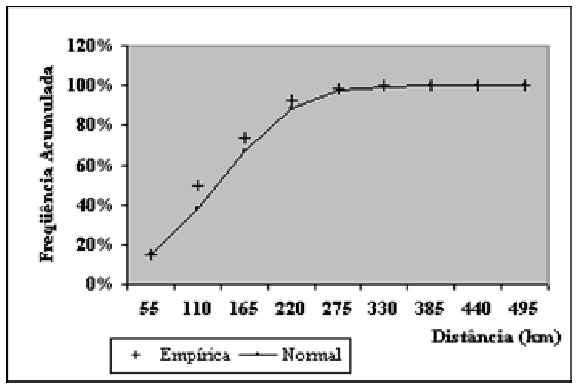

Figura B.4a Distribuição Normal - Anel 2

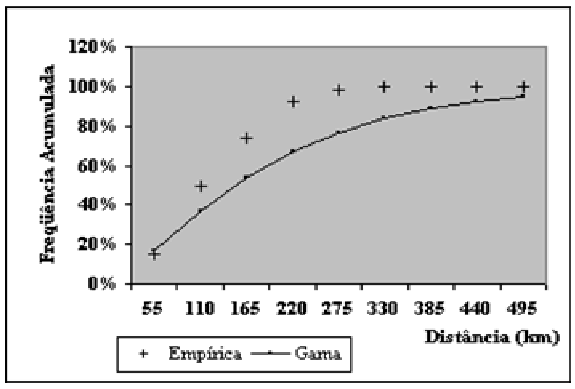

Figura B.4b Distribuição Gama - Anel 2

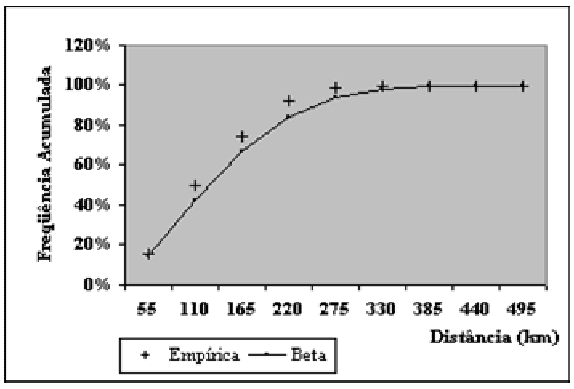

Figura B.4c-Distribuição Beta- Anel 2

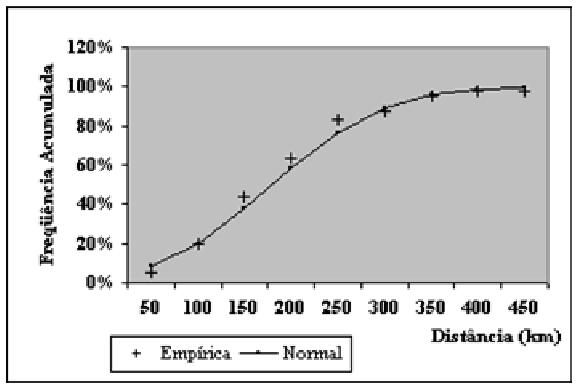

Figura B.5a Distribuição Normal - Anel 3

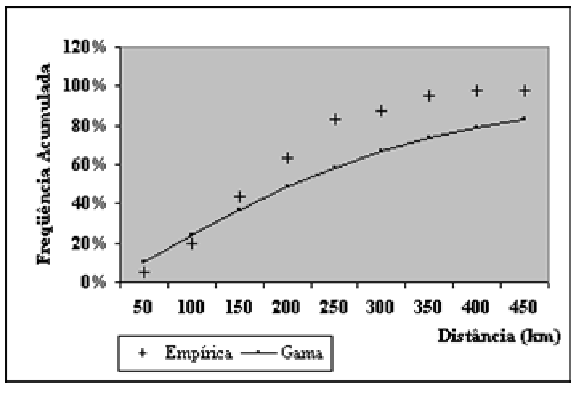

Figura B.5b Distribuição Gama - Anel 3

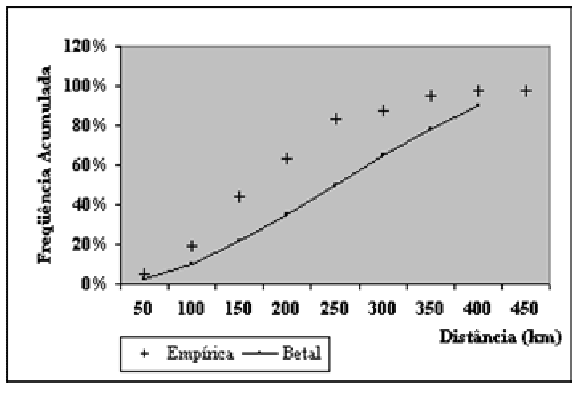

Figura B.5c-Distribuição Beta- Anel 3