

UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ  
PROGRAMA DE PÓS-GRADUAÇÃO EM COMPUTAÇÃO APLICADA

ROBERTO OLIVEIRA SANTOS

**METODOLOGIA PARA LOCALIZAÇÃO DE ESTAÇÕES  
METEOROLÓGICAS: COMPARAÇÃO ENTRE ABORDAGENS  
EXATA E HEURÍSTICA**

DISSERTAÇÃO

CURITIBA

2013

**ROBERTO OLIVEIRA SANTOS**

**METODOLOGIA PARA LOCALIZAÇÃO DE ESTAÇÕES  
METEOROLÓGICAS: COMPARAÇÃO ENTRE ABORDAGENS  
EXATA E HEURÍSTICA**

Dissertação apresentada ao Programa de Pós-graduação em Computação Aplicada da Universidade Tecnológica Federal do Paraná como requisito parcial para obtenção do grau de Mestre-área de Concentração: Engenharia de Sistemas Computacionais.

Orientador: Prof. Dr. Heitor Silvério Lopes

**CURITIBA**

**2013**

---

Dados Internacionais de Catalogação na Publicação

---

- S237 Santos, Roberto Oliveira  
Metodologia para localização de estações meteorológicas: comparação entre abordagens exata e heurística/ Roberto Oliveira Santos. – 2013.  
96 f. : il. ; 30 cm
- Orientador: Prof. Dr. Heitor Silvério Lopes.  
Dissertação (Mestrado) – Universidade Tecnológica Federal do Paraná. Programa de Pós-graduação em Computação Aplicada. Curitiba, 2013.  
Bibliografia: f. 89-93.
1. Estações meteorológicas - Planejamento. 2. Problemas de localização (Programação). 3. Pesquisa Operacional. 4. Otimização matemática. 5. Computação evolutiva. 6. Métodos de simulação. 7. Computação - Dissertações. I. Lopes, Heitor Silvério Lopes, orient. II Universidade Tecnológica Federal do Paraná. Programa de Pós-graduação em Computação Aplicada. III. Título.

CDD (22. ed.) 004

---

Biblioteca Central da UTFPR, Campus Curitiba

---

**Título da Dissertação**

**“METODOLOGIA PARA LOCALIZAÇÃO DE ESTAÇÕES  
METEREOLÓGICAS: COMPARAÇÃO ENTRE ABORDAGENS EXATA E  
HEURÍSTICA.”**

por

**Roberto Oliveira Santos**

Esta dissertação foi apresentada como requisito parcial à obtenção do grau de MESTRE EM COMPUTAÇÃO APLICADA - Área de Concentração: Sistemas de Informação, pelo PPGCA - Programa de Pós-Graduação em Computação Aplicada - Mestrado Profissional – da Universidade Tecnológica Federal do Paraná - UTFPR - Câmpus Curitiba, às 14:00 horas do dia 21 de março de 2013. O trabalho foi aprovado pela Banca Examinadora, composta pelos professores:

**Programa de Pós-Graduação  
em Computação Aplicada**

---

**Prof. Heitor Silvério Lopes, Dr.**  
presidente - (UTFPR - CT)

---

**Prof. Luiz Fernando Nunes, Dr.**  
(UTFPR - CT)

---

**Profª. Maria Teresinha Arns Steiner, Drª.**  
(UFPR)

---

**Prof. Sergio Scheer, Dr.**  
(UFPR)

**Você** (*Tim Maia*)

Você  
É algo assim  
É tudo pra mim  
É como eu sonhava  
baby

Para minha querida esposa Elaine, com amor!

## **AGRADECIMENTOS**

Tenho comigo, que de tempos em tempos é bom fazer um balanço da vida, comemorando os acertos e aprendendo com os erros e assim vamos ajustando o curso do barco. A conclusão deste mestrado é um destes momentos. Expresso aqui o meu muito obrigado a todas as pessoas que consciente ou inconscientemente ajudaram neste trabalho. Espero que a conclusão deste mestrado não seja somente a obtenção de um título, nem tão pouco a ação de repetir o conhecimento adquirido, mas que eu possa humildemente contribuir para a melhoria das condições de vida do próximo, através do pensamento crítico e da expansão do nosso entendimento da ciência e tecnologia.

Agradeço o professor Dr. Heitor Silvério Lopes, pela oportunidade de trabalharmos juntos e pela orientação sempre serena, competente e produtiva, sem a qual este trabalho não seria concluído. Aos colegas do laboratório de bioinformática pelo companheirismo, animadas conversas e jantares de confraternização. Aos professores do Departamento de Informática, em especial o professor Dr. João Fabro, coordenador do PPGCA, sempre com disposição para nos receber com uma palavra de incentivo e resolver os problemas que surgiram durante este período.

Agradeço a Prof. Dra. Maria Teresinha Arns Steiner, o Prof. Dr. Luiz Nunes e o Prof. Dr. Sérgio Scheer por aceitarem gentilmente participar da banca e pelos inúmeros comentários, sugestões e correções apontadas, que sem dúvida tornaram este trabalho melhor.

Agradeço o Eng. Fabio Sato e o pesquisador PhD. Reinaldo Silveira Bonfim do SIMEPAR, por atuarem informalmente como coorientadores neste trabalho. Considero as nossas conversas sobre este trabalho e sobre pesquisa em geral como uma grande aula sobre metodologia científica.

Agradeço a minha família, que esteve presente e acompanhou todos os passos com paciência e carinho. Agradeço os meus pais, que me ensinaram o valor do conhecimento e não mediram esforços para proporcionar o melhor estudo possível. Agradeço os meus irmãos, fonte inesgotável de apoio e energia positiva.

Agradeço os meus cunhados e sogros pelo apoio e suporte em todos os momentos e os meus sobrinhos queridos, sem os quais a vida seria menos alegre.

Agradeço Elaine, minha querida esposa, pela compreensão nos longos períodos em que estive dedicado aos estudos, por embarcar nesta jornada juntos, realizando também o mestrado, pelo incentivo nos momentos de desânimo, pela confiança nos momentos de dúvida, pelo carinho e amor.

Tudo é uma questão de manter

A mente quieta

A espinha ereta

E o coração tranquilo

**Coração Tranquilo** (*Walter Franco*)



## RESUMO

SANTOS, Roberto. METODOLOGIA PARA LOCALIZAÇÃO DE ESTAÇÕES METEOROLÓGICAS: COMPARAÇÃO ENTRE ABORDAGENS EXATA E HEURÍSTICA. 96 f. Dissertação – Programa de Pós-graduação em Computação Aplicada, Universidade Tecnológica Federal do Paraná. Curitiba, 2013.

Este trabalho apresenta um método para o planejamento da implantação de uma estrutura de rede de monitoramento meteorológico que utiliza o resultado do modelo numérico de previsão do tempo MBAR como parâmetro para a avaliação da qualidade das soluções candidatas. O problema é abordado com o uso dos modelos de problemas de localização discretos: o Problema de Cobertura de Conjuntos (*Set Covering Location Problem*) e o Problema de Cobertura Máxima (*Maximal Covering Location Problem*). O problema de planejamento da estrutura é dividido em duas fases. Na primeira fase, busca-se determinar a quantidade mínima necessária e localização de estações meteorológicas necessárias para atender todos os locais de demanda. Na segunda etapa, busca-se determinar a ordem de instalação das estações meteorológicas, maximizando a área de cobertura a cada nova instalação. O método proposto é detalhado, apresentando as etapas envolvidas, as informações de entrada e saída de cada etapa e as alterações necessárias para avaliação de outros algoritmos. Um conjunto de três cenários foi planejado utilizando os limites políticos do Estado do Paraná e as informações de rodovias, áreas urbanas e altimetria. O primeiro cenário avaliou-se a capacidade das abordagens utilizadas em resolver o Problema de Cobertura de Conjuntos. No segundo cenário avaliou-se a capacidade das abordagens utilizadas em resolver o Problema de Cobertura Máxima e no terceiro cenário propõe-se o Problema de Cobertura Máxima considerando a existência de uma rede de monitoramento anterior no local do experimento. Na resolução do problema foram comparadas abordagens exata e heurística. Na abordagem exata utilizou-se o método *Branch & Bound* para resolução do problema via Programação Linear Inteira Mista. Na abordagem heurística utilizou-se a estratégia evolutiva Evolução Diferencial. O método proposto destaca-se pela flexibilidade na substituição dos métodos utilizados na abordagem, permitindo a avaliação de outras técnicas.

**Palavras-chave:** Estações meteorológicas, Problemas de localização, Pesquisa Operacional, Otimização matemática, Computação evolutiva, Computação

## ABSTRACT

SANTOS, Roberto. METHODOLOGY FOR WEATHER STATION LOCATION: COMPARISON BETWEEN EXACT AND HEURISTIC APPROACHES. 96 f. Dissertação – Programa de Pós-graduação em Computação Aplicada, Universidade Tecnológica Federal do Paraná. Curitiba, 2013.

This paper presents a method for planning the deployment of a network's infrastructure for meteorological monitoring that uses the results of the numerical weather prediction MBAR as parameter for evaluating the quality of candidate solutions. The problem is addressed with the use of models for discrete location problems: the Set Covering Location Problem and Maximal Covering Location Problem. The problem of planning the structure is divided into two phases. In the first phase, we seek to determine the minimum required amount and location of weather stations required to meet all demand. In the second step, we seek to determine the order of installation of weather stations, maximizing the coverage area for each new installation. The proposed method is detailed, presenting the steps involved, the input and output information of each step and the changes necessary to evaluate other algorithms. A set of three scenarios was designed using the political boundaries of the State of Paraná, highways, urban areas and altimetry information. The first scenario evaluates the ability of the approaches used to solve the Set Covering Location Problem. In the second scenario, we evaluated the ability of the approaches used to solve the Maximal Covering Location Problem and the third scenario proposes the Maximum Coverage Location Problem considering the previous existence of a monitoring network. For the solution of these problems it was compared exact and heuristic approaches. The exact approach used the method *Branch & Bound* for solving the problem via Mixed Integer Linear Programming. The heuristic approach used the evolutionary strategy Differential Evolution. The proposed method is distinguished by flexibility for substitution of the methods used, allowing the evaluation of other techniques.

**Keywords:** Weather Stations, Facility Location Problems, Operational Research, Optimization, Monitoring Network, Evolutionary Computation, Computation

## LISTA DE FIGURAS

FIGURA 1	– Sistema de observação convencional .....	21
FIGURA 2	– Sistema de observação automático .....	22
FIGURA 3	– Cobertura da rede global de observações segundo o ECMWF .....	23
FIGURA 4	– A cobertura da rede nacional de estações automáticas segundo o INMET ..	24
FIGURA 5	– Domínio espacial de um modelo de previsão do tempo .....	27
FIGURA 6	– Métodos de interpolação exatos e inexatos. ....	30
FIGURA 7	– Taxonomia dos modelos de localização proposta por Daskin. ....	36
FIGURA 8	– Taxonomia dos modelos de localização discretos proposta por Daskin. ...	38
FIGURA 9	– Modelo esquemático do método proposto. ....	57
FIGURA 10	– Saída da etapa 1 - Lista de locais candidatos e características associadas ..	59
FIGURA 11	– Áreas urbanas em cor amarela e o buffer de cinco quilômetros em cor azul	66
FIGURA 12	– <i>Buffer</i> de 500 metros em torno do eixo das rodovias em cor cinza .....	67
FIGURA 13	– Mapa de altimetria baseado nos dados do projeto SRTM .....	68
FIGURA 14	– Localização das estações do SIMEPAR no Estado do Paraná. ....	71
FIGURA 15	– Tempo de execução em relação à quantidade de candidatos. ....	76
FIGURA 16	– Valor da função objetivo e quantidade de estações em relação à quantidade de candidatos. ....	77
FIGURA 17	– Localização dos candidatos selecionados .....	78
FIGURA 18	– Solução da instância com 1781 candidatos em relação as rodovias e área urbanas .....	79
FIGURA 19	– Tempo de execução dos métodos exato e evolução diferencial para instância com 1781 locais candidatos .....	79
FIGURA 20	– Valor da função objetivo dos métodos exato e evolução diferencial para instância com 1781 locais candidatos .....	81
FIGURA 21	– Valor de <i>fitness</i> normalizado dos métodos exato e evolução diferencial para instância com 1781 locais candidatos .....	81
FIGURA 22	– Localização das estações do SIMEPAR e os respectivos locais candidatos mais próximos .....	82
FIGURA 23	– Plano de implantação de rede real por método exato .....	83
FIGURA 24	– Plano de expansão da rede ideal por método exato - iteração 1 .....	84
FIGURA 25	– Plano de expansão da rede ideal por método exato - iteração 2 .....	84
FIGURA 26	– Plano de expansão da rede ideal por método exato - iteração 3 .....	85
FIGURA 27	– Plano de expansão da rede ideal por método exato - iteração 4 .....	85

## LISTA DE TABELAS

TABELA 1	– Modelos de superfície contínua .....	29
TABELA 2	– Métodos de interpolação .....	29
TABELA 3	– Experimento - Objetivo e características .....	70
TABELA 4	– Experimento - Objetivo e características .....	70
TABELA 5	– Experimento - Objetivo e características .....	72
TABELA 6	– Experimento - Objetivo e características .....	73
TABELA 7	– Experimento - Objetivo e características .....	74
TABELA 8	– Resultados do experimento .....	76
TABELA 9	– Resultados do experimento .....	80
TABELA 10	– Localização das estações existentes .....	95
TABELA 11	– Localização das estações existentes - continuação .....	96

## LISTA DE SIGLAS

CPTEC	Centro de Previsão do Tempo e Estudos Climáticos
DECEA	Departamento de Controle do Espaço Aéreo
DHN	Diretoria de Hidrografia e Navegação
ECMWF	<i>European Centre for Medium-Range Weather Forecasts</i>
GEOS	<i>Geometry Engine Open Source</i>
GOS	<i>Global Observation System</i>
GPRS	<i>General Packet Radio Service</i>
INMET	Instituto Nacional de Meteorologia
INPE	Instituto Nacional de Pesquisas Espaciais
JSON	<i>JavaScript Object Notation</i>
LTM	<i>Long-term monitoring</i>
MBAR	Modelo Brasileiro de Alta Resolução
MDT	Modelo Digital de Terreno
NRMSE	Normalized Root Mean Squared Error
OSGEO	<i>Open Source Geospatial Foundation</i>
PSF	<i>Python Software Foundation</i>
SIG	Sistema de Informação Geográfica
SIMEPAR	Serviço Meteorológico do Paraná
SRTM	Shuttle Radar Topography Mission
TIN	Triangular Irregular Network
WMO	<i>World Meteorological Organization</i>

## SUMÁRIO

<b>1 INTRODUÇÃO</b> .....	<b>16</b>
1.1 JUSTIFICATIVA DO PROJETO .....	17
1.2 OBJETIVOS .....	18
1.2.1 Geral .....	18
1.2.2 Específicos .....	18
1.3 ESTRUTURA DO TRABALHO .....	18
<b>2 FUNDAMENTAÇÃO TEÓRICA</b> .....	<b>19</b>
2.1 SISTEMAS DE OBSERVAÇÃO .....	19
2.1.1 Sistemas de observação convencionais .....	20
2.1.2 Sistemas de observação automáticos .....	20
2.1.3 Seleção de local para instalação .....	21
2.1.4 Situação atual dos sistemas de observação .....	23
2.1.5 Previsão numérica do tempo .....	24
2.2 ANÁLISE ESPACIAL E SISTEMAS DE INFORMAÇÕES GEOGRÁFICAS .....	26
2.2.1 Tipos de dados em análise espacial .....	28
2.3 INTERPOLAÇÃO ESPACIAL .....	28
2.3.1 Modelo de grade .....	28
2.3.2 Métodos de interpolação .....	29
2.3.3 Modelos determinísticos locais .....	30
2.3.4 Sistemas de informações geográficas .....	31
2.4 PROBLEMAS DE LOCALIZAÇÃO .....	32
2.4.1 Problema de Weber .....	32
2.4.2 Modelos de localização .....	33
2.4.3 Componentes de modelos de localização de instalações .....	34
2.4.3.1 Instalações .....	34
2.4.3.2 Localizações candidatas .....	35
2.4.3.3 Usuários .....	35
2.4.4 Taxonomia dos modelos de localização .....	36
2.4.4.1 Modelos analíticos .....	36
2.4.4.2 Modelos contínuos .....	37
2.4.4.3 Modelos de rede .....	37
2.4.4.4 Modelos discretos .....	37
2.4.5 Modelos baseados em cobertura .....	38
2.4.5.1 Cobertura de conjuntos .....	39
2.4.5.2 Cobertura máxima .....	40
2.4.5.3 P-centro .....	41
2.4.6 Modelos baseados em medianas .....	41
2.4.6.1 P-mediana .....	41
2.4.6.2 Custo fixo .....	41
2.4.7 Outros Modelos .....	42
2.5 COMPUTAÇÃO EVOLUCIONÁRIA (CE) .....	42

2.5.1	Evolução Diferencial (ED)	42
2.5.1.1	Estrutura da população	42
2.5.1.2	Processo de Inicialização	44
2.5.1.3	Processo de Mutaç�o	45
2.5.1.4	Processo de Combinaç�o	45
2.5.1.5	Processo de Seleç�o	46
2.5.1.6	Pseudoc�digo	46
2.5.1.7	Par�metros da evoluç�o diferencial	47
2.6	TRABALHOS CORRELATOS	48
2.6.1	Monitoramento de qualidade de �guas subterr�neas	48
2.6.2	Monitoramento do meio ambiente	49
2.6.3	Monitoramento de meteorologia	50
2.6.4	Outros	50
<b>3</b>	<b>METODOLOGIA</b>	<b>51</b>
3.1	O AMBIENTE DE PESQUISA	51
3.2	MATERIAIS	52
3.2.1	JSON	52
3.2.2	Lp-solve	52
3.2.3	PostgreSQL	52
3.2.4	PostGIS	53
3.2.5	Python	53
3.2.6	QGIS	53
3.2.7	Shapefile	53
3.2.8	Shapely	54
3.3	O M�TODOS PROPOSTO	54
3.3.1	Etapa 1 - Definiç�o dos locais candidatos	56
3.3.2	Etapa 2 - Geraç�o das soluç�es candidatas	61
3.3.3	Etapa 3 - Geraç�o das interpolaç�es	62
3.3.4	Etapa 4 - Avaliaç�o das soluç�es candidatas	64
3.3.5	Peso das caracter�sticas dos locais candidatos	65
3.4	PLANEJAMENTO DOS EXPERIMENTOS	65
3.4.1	Definiç�o das restriç�es da modelagem matem�tica	67
3.4.2	Cen�rio 1	68
3.4.2.1	Soluç�o do problema de cobertura de conjuntos por m�todo exato	69
3.4.2.2	Soluç�o problema de cobertura de conjuntos por evoluç�o diferencial	70
3.4.3	Cen�rio 2	71
3.4.3.1	Rede ideal com a exist�ncia pr�via de estaç�es	72
3.4.3.2	Rede real com a exist�ncia pr�via de estaç�es	72
3.4.4	Cen�rio 3	73
3.4.4.1	Plano de crescimento sem a exist�ncia pr�via de estaç�es	73
<b>4</b>	<b>RESULTADOS</b>	<b>75</b>
4.1	CEN�RIO 1	75
4.1.1	Soluç�o do problema de cobertura de conjuntos por m�todo exato	75
4.1.2	Soluç�o problema de cobertura de conjuntos por evoluç�o diferencial	77
4.2	CEN�RIO 2	80
4.2.1	Rede ideal com a exist�ncia pr�via de estaç�es	80
4.3	CEN�RIO 3	82
4.3.1	Plano de crescimento sem a exist�ncia pr�via de estaç�es	82

<b>5 CONCLUSÕES E TRABALHOS FUTUROS .....</b>	<b>86</b>
<b>REFERÊNCIAS .....</b>	<b>89</b>
<b>Apêndice A – MODELO MATEMÁTICO .....</b>	<b>94</b>
<b>Apêndice B – LOCALIZAÇÃO DAS ESTAÇÕES EXISTENTES .....</b>	<b>95</b>



## 1 INTRODUÇÃO

As condições atmosféricas influenciam direta e indiretamente diversas atividades humanas, tais como: geração de energia, agricultura, comércio, indústria e turismo. Também são determinantes em relação às condições de segurança da população, conforme descrito por Sugahara et al. (2010) no caso de desastres naturais como os ocorridos no Vale do Paraíba nos primeiros dias de janeiro de 2010, onde a quantidade de chuva acumulada em cinco dias causou mortes e destruição. Outros eventos extremos mais recentes, como os ocorridos em Santa Catarina em 2010 e no litoral do Paraná e na região serrana do Rio de Janeiro, em 2011, evidenciam o impacto que as condições atmosféricas têm sobre a população. Devido a esta forte relação entre a atividade humana e as condições atmosféricas, é grande o interesse das autoridades e da comunidade científica no processo de coleta de observações e nos processos posteriores que envolvem a previsão das condições atmosféricas.

Segundo Silveira (2010), as observações são geralmente realizadas em pontos previamente conhecidos, que podem ser móveis ou estacionários, como um navio ou a superfície da Terra. As observações são realizadas por diversos tipos de sensores que são reunidos em um equipamento conhecido por estação meteorológica, no caso das observações na superfície da Terra. Ao conjunto destas estações dá-se o nome de rede de estações ou rede de monitoramento. No âmbito mundial, é de responsabilidade da WMO (*World Meteorological Organization* - Organização Meteorológica Mundial) a supervisão do GOS (*Global Observation System* - Sistema Global de Observação), que compreende as redes de monitoramento administradas pelos serviços meteorológicos nacionais. No Brasil, é de responsabilidade do INMET (Instituto Nacional de Meteorologia) a instalação e manutenção da rede de monitoramento que abrange o país. Em nível regional, é de responsabilidade do SIMEPAR (Serviço Meteorológico do Paraná) a rede que abrange o estado do Paraná, a qual conta com estações meteorológicas automáticas e convencionais, além de estações hidrológicas.

Atualmente não existe um método largamente aceito para definição da estrutura de uma rede de monitoramento de estações meteorológicas, em especial no que concerne à localização das estações. Quando da instalação de um novo nó na rede, são observadas características como

segurança do local, facilidade de acesso para manutenção, disponibilidade de espaço físico, meios de comunicação disponíveis, histórico de desastres naturais, entre outros. Apesar da importância destas características não ser questionada, percebe-se a ausência de ferramentas científicas e tecnológicas que permitam mensurar a redundância ou o valor que dados providos por uma nova estação possam agregar aos processos de monitoramento e previsão das condições atmosféricas.

Ainda em relação ao planejamento de uma rede, verifica-se que a quantidade de estações em geral está limitada ao orçamento disponível. Entretanto, os processos que fazem uso dos dados coletados por estas redes almejam estruturas com o máximo possível de nós, para obter uma melhor representação do comportamento das variáveis observadas (tipicamente temperatura, pressão, umidade do ar, entre outras) na região de interesse.

Assim, os responsáveis pelo planejamento de redes de monitoramento deparam-se com o seguinte problema: como implantar redes com custos que respeitem o orçamento disponível e que ao mesmo tempo sejam capazes de coletar dados que representem adequadamente o comportamento das variáveis?

## 1.1 JUSTIFICATIVA DO PROJETO

Segundo Amorim et al. (2011), “O planejamento de redes de estações de monitoramento é um tópico que tem atraído a atenção de planejadores e pesquisadores, visto que as vantagens de um planejamento e gerenciamento adequado destas estações traduzem-se em benefícios econômicos e científicos.”.

No planejamento de uma rede de estações, os dois componentes principais são: o custo econômico da rede (cuja meta é reduzir ao mínimo necessário a quantidade de estações) e a abrangência das observações (cuja meta é expandir ao máximo a quantidade de estações). A área de estudo que compreende este tipo de problema é conhecida por localização de instalações, do inglês *Facility Location*. Segundo Daskin (1995), os problemas de localização de instalações estão relacionados às decisões sobre onde devem ser localizadas as estruturas para atendimento de clientes de forma a otimizar determinados critérios.

O SIMEPAR gerencia uma rede de estações meteorológicas espalhada pelo estado do Paraná e se depara com o problema de definição da localização quando surge a necessidade de instalar novas estações.

Observa-se que o desenvolvimento de um método aplicado ao processo de planejamento da rede de estações que considere a otimização é um avanço a ser alcançado

no serviço de monitoramento meteorológico do SIMEPAR.

## 1.2 OBJETIVOS

### 1.2.1 GERAL

Este trabalho tem por objetivo projetar e implementar um método para planejamento de implantação e expansão de redes de estações meteorológicas, minimizando os custos envolvidos e maximizando a área de cobertura em vistas a melhoria da qualidade dos serviços de monitoramento meteorológico.

### 1.2.2 ESPECÍFICOS

- Apresentar uma sistemática para indicar a localização das estações meteorológicas em uma rede ideal;
- Apresentar uma sistemática para indicar a localização de novas estações meteorológicas em uma rede existente, considerando a capacidade máxima de instalação em um determinado período;
- Expandir a discussão acadêmica através da publicação de artigos em periódicos e congressos;
- Avaliar as abordagens exata e heurística para resolução deste tipo de problema.

## 1.3 ESTRUTURA DO TRABALHO

O presente trabalho está estruturado em cinco capítulos e demais complementos. O segundo capítulo apresenta a fundamentação teórica dos problemas de localização de instalações, os métodos para determinação de soluções exatas ou aproximadas disponíveis atualmente bem como a revisão da literatura, onde são apresentados trabalhos similares publicados nos últimos 10 anos, com breves comentários sobre o ambiente do problema abordado, as técnicas aplicadas e os resultados obtidos. O terceiro capítulo apresenta o método proposto neste trabalho e descreve os cenários dos experimentos realizados. O quarto capítulo apresenta os resultados obtidos nos experimentos na forma de tabelas, gráficos e mapas. O quinto capítulo apresenta as considerações finais, identifica o atendimento aos objetivos propostos e propõe sugestões para trabalhos futuros.

## 2 FUNDAMENTAÇÃO TEÓRICA

### 2.1 SISTEMAS DE OBSERVAÇÃO

Segundo Büyükbas et al. (2006), “Os seres humanos vivem em um ambiente cercado pela atmosfera. Desta forma, todas as mudanças e fenômenos que ocorrem na atmosfera e no ambiente onde se vive afetam direta ou indiretamente os seres humanos. Ser capaz de minimizar os efeitos negativos destes fenômenos e usar os resultados de forma benéfica para os seres humanos é uma das motivações para observar a atmosfera e o meio ambiente.”<sup>1</sup>.

Ainda de acordo com Büyükbas et al. (2006), as seguintes variáveis começaram a ser observadas e seu comportamento futuro é analisado utilizando-se alguns métodos e técnicas bastante simples: vento, temperatura do ar, umidade, precipitação, pressão atmosférica, insolação, visibilidade e nuvens.

Atualmente, existem diversos produtos relacionados à Meteorologia como previsão de curto prazo, modelos de previsão numérica do tempo, pesquisas sobre clima e mudança climática, desenvolvimento sustentável, proteção do meio ambiente e fontes de energia renováveis, entre outros. Todos estes produtos dependem das fontes de dados utilizadas e todas estas têm em comum as observações meteorológicas como sua origem.

Os sistemas de observação podem ser classificados em diversas formas de acordo com o critério utilizado.

Baseado no local onde a observação é realizada, os sistemas se definem em: observações de superfície e observações nas camadas superiores da atmosfera.

Os sistemas de observações de superfície podem ser classificadas em: convencionais e automáticos.

---

<sup>1</sup>All living beings live in an environment surrounded by atmosphere. So, all changes and phenomena occurred in the atmosphere and in the environment lived in affect all livings much, human being as well. To be able to minimize the negative effects of the phenomena occurred and to be able to use the results of those phenomena more beneficial for the human being, it is very important to observe the atmosphere and the environment.

### 2.1.1 SISTEMAS DE OBSERVAÇÃO CONVENCIONAIS

Os sistemas de observação convencionais consistem de um observador humano e alguns instrumentos para medição das variáveis meteorológicas. A Figura 1 apresenta um sistema de observação convencional com instrumentos para leitura das variáveis meteorológicas. Estes sistemas possuem as seguintes características:

- Observam determinadas variáveis como vento, temperatura, umidade relativa, precipitação, nuvens e visibilidade;
- Instrumentos convencionais que dependem de um observador para leitura;
- Subjetividade na leitura;
- Frequência de leitura limitada à quantidade de observadores disponíveis;
- Limitada quantidade de variáveis;
- Instrumentos mecânicos gravando em papel.

### 2.1.2 SISTEMAS DE OBSERVAÇÃO AUTOMÁTICOS

Um sistema de observação meteorológico automático é um conjunto completo composto por:

- Sensores e interfaces para sensores;
- Unidade de coleta de dados;
- Central de controle e unidade de processamento;
- Visor de informações;
- Interface de comunicação;
- Baterias.

A Figura 2 mostra um sistema de observação automático onde os sensores estão concentrados em torno de uma torre utilizada pelo sensor de velocidade e direção do vento e geralmente pelo sistema de comunicação.

Os sistemas de observação automáticos podem ser classificados de acordo com o propósito do sistema.



**Figura 1: Sistema de observação convencional.**

**Fonte: Büyükbas et al. (2006).**

- Sinótico: observa e calcula parâmetros e gera relatórios para análise da previsão do tempo.
- Climatologia: observa e calcula parâmetros e gera relatórios para estudos sobre o clima.
- Agricultura: observa e calcula parâmetros e gera relatórios para atividades agrícolas.
- Aviação: observa e calcula parâmetros para apoiar a segurança de vôo e a aviação.
- Marinha: observa e calcula parâmetros para apoiar a navegação marítima.
- Hidrologia: observa e calcula parâmetros para hidrologia e irrigação.

### 2.1.3 SELEÇÃO DE LOCAL PARA INSTALAÇÃO

Segundo Büyükbas et al. (2006), a determinação do local correto para instalação dos sistemas de observação é o primeiro e mais importante passo para o sucesso de uma rede de sistemas de observação. Estes locais devem ser determinados levando em consideração as



**Figura 2: Sistema de observação automático.**

**Fonte: Büyükbas et al. (2006).**

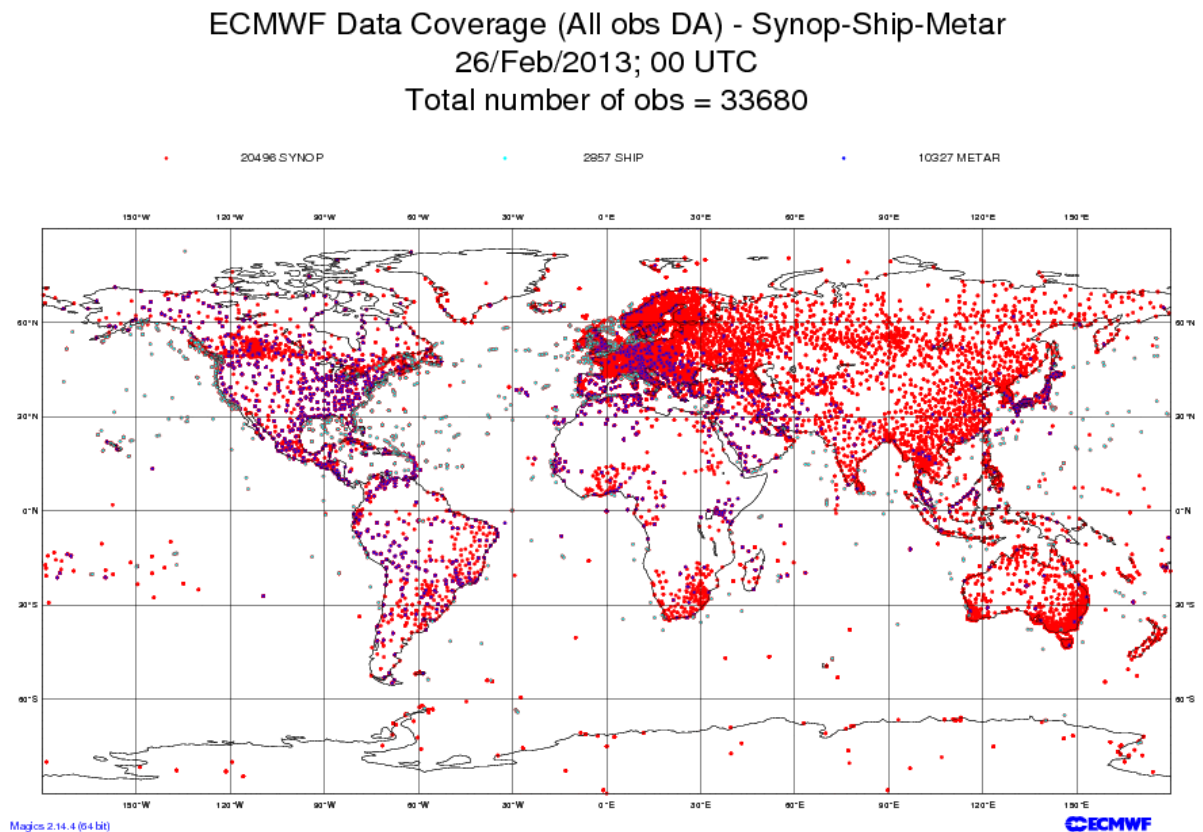
recomendações da WMO. Durante o estudo para determinação dos locais de instalação, os seguintes critérios devem ser considerados:

- Quais as variáveis meteorológicas a serem observadas
- O propósito da obtenção destas variáveis
- O comportamento das variáveis no entorno do local
- O tamanho da área a ser representada pela estação
- A adequação do local às variáveis observadas
- Infraestrutura física e de comunicação disponíveis

### 2.1.4 SITUAÇÃO ATUAL DOS SISTEMAS DE OBSERVAÇÃO

De acordo com Silveira (2010), os atuais sistemas de observação meteorológicos possuem uma história de evolução de 300 anos, desde o advento dos primeiros equipamentos meteorológicos como termômetros (temperatura), barômetros (pressão), higrômetros (umidade) e anemômetros (vento), no século dezoito.

Atualmente, a WMO é o órgão responsável pela integração das diversas redes de observação existentes. Na Figura 3 é possível visualizar a cobertura do GOS segundo o ECMWF (*European Centre for Medium-Range Weather Forecasts* - Centro Europeu de Prognóstico de Tempo de Médio Alcance). Na Figura 4 é possível visualizar a cobertura da rede de estações automáticas do INMET.



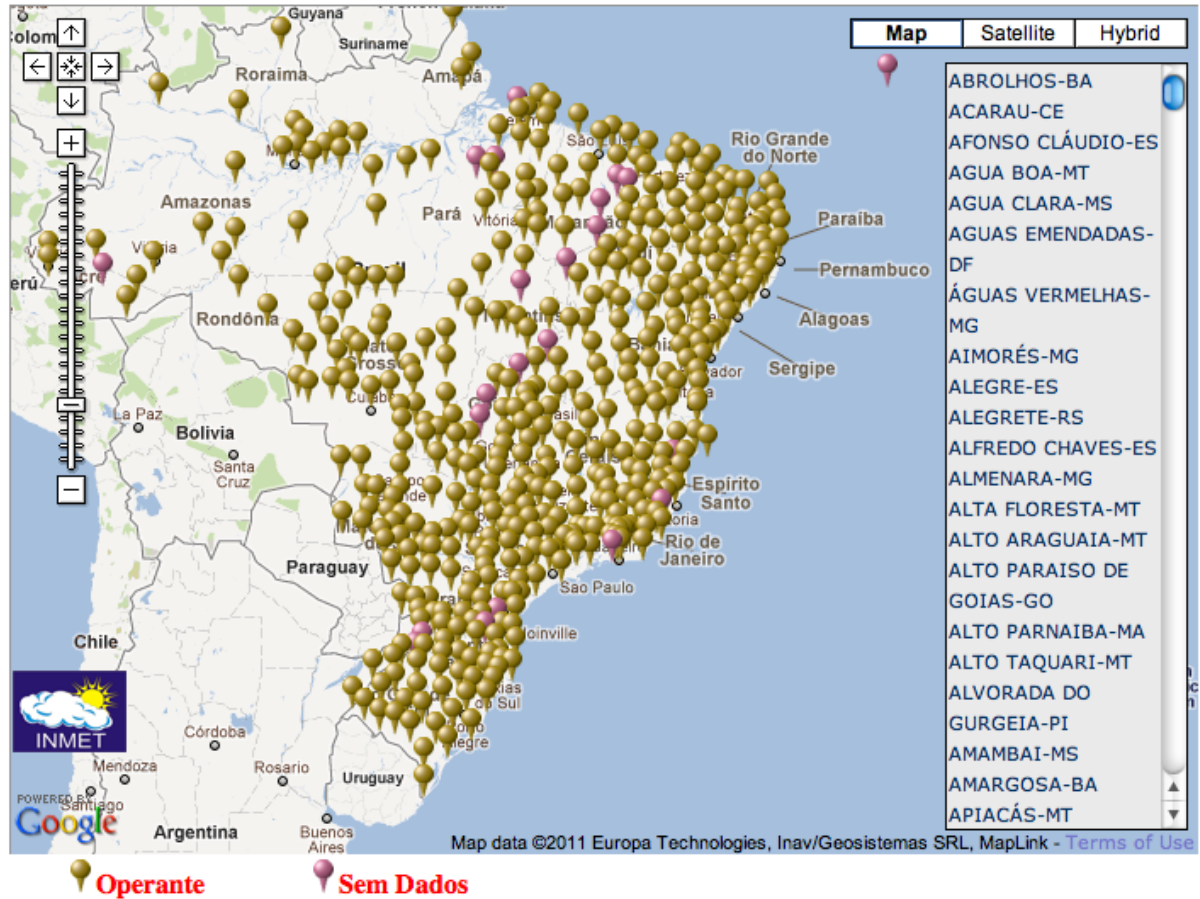
**Figura 3: Cobertura da rede global de observações segundo ECMWF.**

**Fonte: ECMWF (2013)**

O custo de implantação de uma estação meteorológica envolve a aquisição de sensores, proteção dos equipamentos contra roubo e vandalismos, equipamentos de armazenamento e transmissão de dados no caso das estações automáticas e técnico capacitado para leitura dos dados no caso das estações convencionais. Além disto, a transmissão dos dados pode ocorrer



por rede *ethernet*/fibra ótica, *link* de rádio, GPRS (*General Packet Radio Service* - Serviço de Rádio de Pacote Geral) ou satélite.



**Figura 4: A cobertura da rede nacional de estações automáticas do INMET.**

**Fonte: INMET (2013)**

Atualmente, os principais órgãos operacionais de meteorologia do Brasil que mantêm uma rede de observação em nível nacional são: O INMET, o DECEA (Departamento de Controle do Espaço Aéreo) do Comando da Aeronáutica e a DHN (Diretoria de Hidrografia e Navegação) do Comando da Marinha, ambos do Ministério da Defesa, além do INPE (Instituto Nacional de Pesquisas Espaciais) do Ministério da Ciência e Tecnologia.

### 2.1.5 PREVISÃO NUMÉRICA DO TEMPO

A previsão numérica do tempo tem sua origem no século XIX, quando o desenvolvimento da termodinâmica permitiu a formalização de um conjunto de princípios físicos que governam o fluxo da atmosfera. Segundo Lynch (2008), o meteorologista americano Cleveland Abbe reconheceu que “a Meteorologia é essencialmente a aplicação da

hidrodinâmica e termodinâmica na atmosfera”.

Segundo Daley (1991), no início do século XX, o meteorologista e físico norueguês Vilhelm Bjerknes determinou duas condições necessárias para prever os estados futuros da atmosfera, descritas da seguinte forma:

1. O presente estado da atmosfera deve ser caracterizado com o máximo de precisão possível.
2. As leis intrínsecas, pelas quais os subsequentes estados se desenvolvem a partir dos anteriores, devem ser conhecidas.

Estas duas condições definem a previsão do tempo como um problema de valor inicial, porque um estado futuro da atmosfera é assumido como completamente determinado a partir de um estado inicial. Desta forma, Bjerknes desenvolveu uma metodologia para implantação prática do problema de valor inicial, que consiste de três componentes:

- Componente de observação
- Componente de análise
- Componente de prognóstico

O primeiro e o segundo componentes estão relacionados com o estado atual da atmosfera e o terceiro componente está relacionado com a segunda condição. Segundo este modelo, o componente de observação requer uma rede de observação distribuída na atmosfera, de forma a monitorar as variáveis meteorológicas como pressão, temperatura, vento, etc. A distância entre as estações de observação e o tempo entre as observações devem ser suficientemente pequena para resolver adequadamente as escalas espaciais e temporais do problema. Esta condição é a causa principal da necessidade de implantação de densas redes de observação, chegando a distâncias cada vez mais reduzidas.

Richardson (1922) propôs um algoritmo para a elaboração sistemática da previsão numérica do tempo. Por seu trabalho detalhado e visão além da sua época, é considerado o mentor da previsão numérica do tempo. Apesar de não alcançar os resultados esperados do seu trabalho, Richardson indicou quatro áreas críticas que deveriam avançar para que a previsão numérica do tempo pudesse obter resultados práticos. Os avanços indicados foram i) melhor entendimento da dinâmica da atmosfera, ii) algoritmos computacionais para integração

das equações da atmosfera estáveis, iii) observações da atmosfera em períodos regulares e iv) equipamentos mais poderosos para cálculo automático.

Na década de 50, o trabalho de matemáticos e meteorologistas como Jule Gregory Charney, Ragnar Fjortoft, John von Neumann, Charney et al. (1950), utilizando os primeiros computadores ENIAC (Electronic Numerical Integrator and Calculator) localizados na *U.S. Army Aberdeen Proving Ground*, em Maryland, e *Institute of Advanced Studies*, em Princeton, Estados Unidos, produziram as primeiras previsões numéricas bem sucedidas.

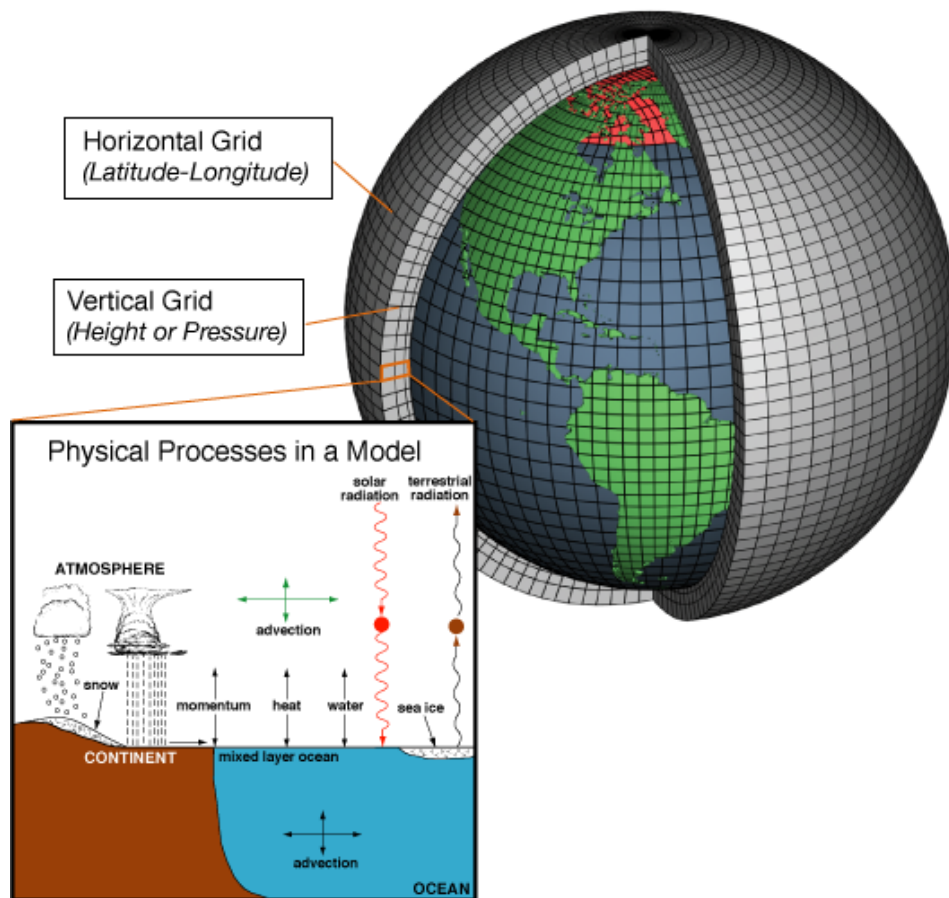
A previsão numérica do tempo utiliza modelos matemáticos que podem ser classificados de acordo com o domínio espacial. Os modelos globais possuem um domínio espacial que cobre toda a superfície terrestre, com um espaçamento horizontal de grade em torno de 100km e com número reduzido de variáveis meteorológicas. Os modelos regionais possuem um domínio espacial reduzido, com um espaçamento horizontal de grade em torno de 5 a 40km e com maior quantidade de variáveis meteorológicas.

A Figura 5 apresenta um modelo esquemático de previsão numérica do tempo onde o domínio espacial tridimensional é dividido em cubos. O modelo executa uma simulação da atmosfera e calcula os valores das variáveis meteorológicas para cada passo de tempo do domínio temporal que pode variar de segundos a minutos. A integração geralmente é realizada para alguns dias (modelos regionais) ou até 15 dias (modelos globais) nos vértices e centro dos cubos.

Atualmente, diversas agências governamentais nacionais e internacionais elaboram previsão numérica do tempo em períodos variando de 5 a 15 dias. No Brasil, o CPTEC (Centro de Previsão do Tempo e Estudos Climáticos) do Instituto Nacional de Pesquisas Espaciais utiliza o modelo regional ETA e sua versão operacional tem espaçamento horizontal de grade de 40km e vertical de 38 camadas (MATOS, 2009). O INMET (Instituto Nacional de Meteorologia) utiliza o modelo MBAR (Modelo Brasileiro de Alta Resolução) com espaçamento horizontal de grade de 10km e vertical de 12 camadas.

## 2.2 ANÁLISE ESPACIAL E SISTEMAS DE INFORMAÇÕES GEOGRÁFICAS

A coleta, análise e divulgação de informações sobre a distribuição geográfica da população humana e animal, recursos naturais e eventos atmosféricos entre outros é uma necessidade constante nas organizações públicas e privadas. A união destas atividades com a Informática levou ao surgimento do Geoprocessamento, um conjunto de conhecimentos matemáticos e computacionais cujo objetivo é o tratamento das informações geográficas.



**Figura 5: Domínio espacial de um modelo de previsão do tempo e as variáveis meteorológicas.**

**Fonte: NOAA (2007).**

Segundo Camara et al. (2001), “Se onde é importante para o seu negócio, então Geoprocessamento é a sua ferramenta de trabalho”. Tem-se nos mapas, um dos produtos mais comuns da atividade de geoprocessamento, sejam impressos ou digitais.

Entretanto, de acordo com Druck et al. (2004), um significativo grupo de fenômenos espaciais possui uma natureza essencialmente dinâmica e as representações estáticas utilizadas em geoprocessamento não os capturam de forma adequada.

Ainda segundo Druck et al. (2004), “Compreender a distribuição espacial de dados oriundos de fenômenos ocorridos no espaço constitui hoje um grande desafio para a elucidação de questões centrais em diversas áreas do conhecimento, seja na saúde, em ambiente, em geologia, em agronomia, entre tantas outras”.

## 2.2.1 TIPOS DE DADOS EM ANÁLISE ESPACIAL

Segundo Druck et al. (2004), a taxonomia mais utilizada para caracterizar os problemas de análise espacial considera três tipos de dados:

- Eventos ou Padrões Pontuais - fenômenos expressos através de ocorrências identificadas como pontos localizados no espaço, denominados processos pontuais. São exemplos: localização de crimes, ocorrências de doenças e localização de espécies vegetais.
- Superfícies Contínuas - estimadas a partir de um conjunto de amostras de campo, que podem estar regularmente ou irregularmente distribuídas. Usualmente, este tipo de dados é resultante de levantamentos naturais, e que incluem mapas geológicos, topográficos, ecológicos, fitogeográficos e pedológicos.
- Áreas com Contagens e Taxas agregadas - tratam-se de dados associados a levantamentos populacionais, como censos e estatísticas de saúde, e que originalmente se referem a indivíduos localizados em pontos específicos do espaço. Por razões de confidencialidade, estes dados são agregados em unidades de análise, usualmente delimitadas por polígonos fechados (setores censitários, zonas de endereçamento postal, municípios).

## 2.3 INTERPOLAÇÃO ESPACIAL

A interpolação espacial é o processo de utilização de pontos com valores conhecidos para estimar valores em outros pontos. Portanto, diz respeito a um conjunto de técnicas que visam a criação de superfícies contínuas a partir de amostras pontuais.

Os modelos de superfície contínua são exemplos de produtos da interpolação espacial. Entre estes modelos pode-se citar o MDT (Modelo Digital de Terreno).

O processo de criação de um modelo cuja representação fosse 100% fiel à realidade demandaria a coleta de um número infinito de pontos. Não sendo possível coletar todos esses pontos, foram desenvolvidas diferentes propostas de representação, sendo as mais populares o TIN, do inglês *Triangular Irregular Network*, e a Grade Regular. A Tabela 1 apresenta um comparativo entre os dois tipos de representação.

### 2.3.1 MODELO DE GRADE

Modelos de grade representam uma superfície a partir de uma matriz na qual cada célula possui um valor quantificador da propriedade representada (por exemplo: altitude,

**Tabela 1: Modelos de superfície contínua**

Grade Regular	TIN
Armazenamento simples	Menos pontos são necessários para que se obtenha a mesma precisão
Compatível com dados de sensoriamento remoto	A resolução se adapta ao modelo
Bons recursos de análise superficial	Armazena a topologia
Processamento lento	Não permite análise de superfície
Pode gerar dados redundantes	Sua estruturação é demorada
Tamanho do pixel uniforme	Não dispõe de algoritmos eficientes para algumas operações.

**Fonte: O Autor (2013).**

temperatura, precipitação, etc).

Nos modelos de superfície que utilizam grade, a interpolação baseia-se no pressuposto de que pontos que estão mais próximos no espaço tendem a apresentar valores de atributos semelhantes. Essa propriedade é conhecida como autocorrelação espacial positiva. A extrapolação refere-se à estimativa de valores em pontos fora da área coberta pelo levantamento.

A maioria dos procedimentos de interpolação assume que existe um gradiente contínuo e regular entre os pontos de amostragem. Esse gradiente pode ser linear ou não linear.

### 2.3.2 MÉTODOS DE INTERPOLAÇÃO

Os métodos de interpolação podem ser classificados em:

- Globais, pois consideram toda a população de pontos;
- Locais, pois consideram uma zona próxima ao ponto que está sendo interpolado.

**Tabela 2: Métodos de interpolação**

Globais	Locais
Análise de superfície de tendência	Polígonos de Thiessen
Modelos de regressão polinomial	Vizinho mais próximo
Análise espectral (séries de Fourier)	Inverso da distância
	Spline
	Krigagem

**Fonte: O Autor (2013).**

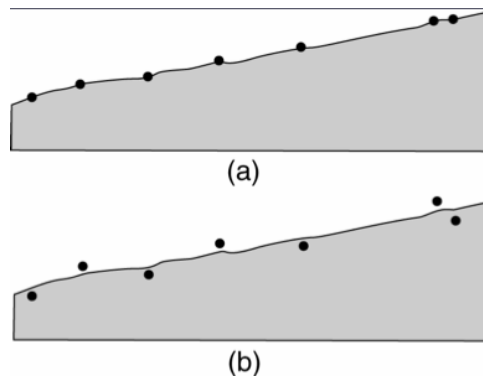
A Tabela 2 apresenta alguns métodos de interpolação espacial.

Os métodos de interpolação também podem ser classificados em:

- Determinísticos, pois não permitem a avaliação de erros associados aos valores previstos;
- Estocásticos, pois permitem a avaliação de erros de previsão com base na estimativa das variâncias.

Ainda pode-se classificar os interpoladores em

- Exatos, quando o método prevê um valor idêntico ao medido no ponto;
- Inexatos, quando o método prevê um valor diferente ao medido no ponto.



**Figura 6: Métodos de interpolação exatos e inexatos.**

A Figura 6a apresenta o resultado de um método de interpolação exato enquanto a 6b apresenta o resultado de um método não exato.

### 2.3.3 MODELOS DETERMINÍSTICOS LOCAIS

Segundo Camara et al. (2001), “uma alternativa simples para gerar uma superfície bidimensional a partir de amostras pontuais é ajustar uma função bidimensional sobre as amostras consideradas, compondo uma superfície cujo valor será proporcional à local intensidade das amostras”.

Este tipo de interpolação pode ser formalizado na seguinte expressão:

$$\hat{z}_i = \frac{\sum_{j=1}^n w_{ij} z_j}{\sum_{j=1}^n w_{ij}}, \quad (1)$$

onde:  $z_i$  é o valor de cota de um ponto  $i$  qualquer da grade,  $z_j$  é a cota de uma amostra  $j$  vizinha do ponto  $i$  da grade e  $w_{ij}$  é um fator de ponderação.

Variações desse esquema são os interpoladores: (a) por vizinho mais próximo; (b) por média simples; (c) por média ponderada. Nestes casos, considera-se uma região em torno do ponto a ser interpolado como contendo os pontos que influenciam na interpolação. A interpolação por vizinho mais próximo é definida pela escolha de apenas uma amostra vizinha para cada ponto de grade. Este interpolador deve ser usado quando se deseja manter os valores de cotas das amostras na grade, sem gerar valores intermediários. A interpolação por média simples considera o valor de cota  $z$  do elemento da grade igual à média aritmética dos valores de cota das amostras vizinhas. Neste caso considera-se que o fator de ponderação  $w_{ij}$  é igual a  $1/n$  para qualquer amostra considerada. Na interpolação por média ponderada o valor de cota de cada elemento da grade é definido pela média ponderada dos valores de cota das amostras vizinhas. A ponderação mais usada na prática é o inverso da distância euclidiana do ponto da grade à amostra considerada, ou seja:

$$w_{ij} = \frac{1}{d_{ij}^k}, \quad (2)$$

onde:  $k$  é o expoente da distância, geralmente igual a 1 ou 2 e  $d_{ij}$  é o valor de distância da amostra  $j$  ao ponto  $i$  da grade, expresso por:

$$d_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \quad (3)$$

#### 2.3.4 SISTEMAS DE INFORMAÇÕES GEOGRÁFICAS

Para Camara et al. (2001), “o termo SIG, Sistema de Informação Geográfica, é aplicado para sistemas que realizam o tratamento computacional de dados geográficos e recuperam informações não apenas com base em suas características alfanuméricas, mas também através de sua localização espacial. Para que isto seja possível, a geometria e os atributos dos dados num GIS devem estar georreferenciados, isto é, localizados na superfície terrestre e representados numa projeção cartográfica”.

Para Silva (2003), “Os Sistemas de Informações Geo-referenciadas ou Sistemas de Informações Geográficas (SIGs) são usualmente aceitos como sendo uma tecnologia que possui o ferramental necessário para realizar análises com dados espaciais e, portanto, oferece, ao ser implementada, alternativas para o entendimento da ocupação e utilização do meio ambiente,



compondo o chamado universo da Geotecnologia, ao lado do Processamento Digital de Imagens (PDI) e da Geoestatística. A tecnologia GIS está para as análises geográficas, assim como o microscópio, o telescópio e os computadores estão para outras ciências (Geologia, Astronomia, Geofísica, Administração, entre outras)”.

## 2.4 PROBLEMAS DE LOCALIZAÇÃO

Problemas de localização podem ser entendidos e modelados de diferentes maneiras de acordo com a experiência e objetivos do observador. De acordo com Eiselt e Marianov (2011), um matemático provavelmente definiria um problema de localização da seguinte forma: “dado um espaço métrico e um conjunto de pontos conhecidos, determine o número adicional de pontos de forma a otimizar uma função de distância entre os novos pontos e os existentes”. Já um geógrafo poderia descrever o problema da seguinte maneira: “dadas algumas regiões onde estão localizados consumidores conhecidos, a tarefa é determinar a localização de uma certa quantidade de supermercados de forma a atender estes consumidores”. Por sua vez, cientistas da computação podem querer determinar “o número mínimo de formas geométricas idênticas necessárias para cobrir uma área contínua e a posição dos seus respectivos centróides”.

Assim, percebe-se que problemas desta natureza podem ser foco de estudo de diferentes áreas do conhecimento. Entretanto, ainda segundo Eiselt e Marianov (2011), existem elementos básicos que são comuns a estas diferentes abordagens e são:

- um espaço, no qual uma unidade de medida de distância é definida;
- um conjunto de pontos fornecidos, identificados como clientes ou demandas;
- um conjunto de localizações candidatas a hospedar novos pontos, identificadas como instalações.

### 2.4.1 PROBLEMA DE WEBER

No decorrer da história e de forma mais intensa a partir do século XVII vários estudiosos trabalharam em problemas de localização bastante similares entre si. Estes trabalhos geraram farta literatura sobre o tema e estes problemas assumiram diferentes nomes como problema de Fermat, problema generalizado de Fermat, problema de Fermat-Torricelli, problema de Steiner, problema generalizado de Steiner, problema de Steiner-Weber, problema de uma mediana, problema da mediana central, problema minisum, entre outros (DASKIN, 1995).

O primeiro modelo formal de um problema de localização foi desenvolvido pelo economista alemão Alfred Weber. O seu trabalho Teoria da localização de indústrias (Uber den Standort der Industrie), de 1909, foi pioneiro na área de modelagem de problemas de localização industrial e pode ser enunciado da seguinte forma:

“Deseja-se encontrar o ponto  $(x^*, y^*)$  que minimiza a soma das distâncias com pesos a partir deste ponto até  $n$  pontos fixos com coordenadas  $(a_i, b_i)$ . Os pesos associados aos pontos fixos são identificados por  $w_i$ ”.

O problema pode ser formulado como:

$$\min_{x,y} \left\{ W(x,y) = \sum_{i=1}^n w_i d_i(x,y) \right\} \quad (4)$$

onde  $d_i(x,y) = \sqrt{(x-a_i)^2 + (y-b_i)^2}$  é a distância Euclidiana entre os pontos  $(x,y)$  e  $(a_i, b_i)$ .

#### 2.4.2 MODELOS DE LOCALIZAÇÃO

Os problemas de localização de instalações estão relacionados às decisões sobre onde devem ser posicionadas instalações para o atendimento de clientes de forma a otimizar determinados critérios tais como: redução de custos, aumento da demanda atendida, balanceamento de carga de atendimento, redução do tempo de atendimento, redução da distância entre cliente e a instalação, etc. Dependendo do tipo de bem ou serviço requerido, as instalações podem trabalhar de forma cooperativa ou competitiva (DASKIN, 1995).

A resolução de problemas de localização implica necessariamente a modelagem da realidade. A transformação implícita de aspectos qualitativos que compõem esta realidade e a representação das relações existentes entre os diferentes elementos eleva ainda mais a complexidade inerente e pode comprometer a obtenção de uma solução satisfatória (AMORIM, 2008).

Nos últimos anos tem-se assistido uma considerável disseminação da pesquisa relacionada à definição e tipificação de problemas de localização bem como dos respectivos modelos para sua abordagem e resolução. A literatura disponível sobre o tema revela grande diversidade, quer na definição e tipificação de problemas de localização, quer nas respectivas propostas de resolução, (DASKIN, 1995) e (EISELT; MARIANOV, 2011).

### 2.4.3 COMPONENTES DE MODELOS DE LOCALIZAÇÃO DE INSTALAÇÕES

Segundo Corrêa (2003), existem três componentes que podem ser considerados básicos para a análise de modelos de locação de facilidades:

- Instalações: objetos que devem ter a sua localização identificada para prover um serviço ou bem.
- Localizações candidatas: o conjunto de pontos candidatos para a localização das instalações.
- Usuários: usuários das instalações que demandam certos serviços.

#### 2.4.3.1 INSTALAÇÕES

O termo “instalações” denota uma grande variedade de objetos para os quais é necessário determinar a localização espacial a fim de otimizar sua interação com outros objetos (SCAPARRA; SCUTELLA, 2001). Como exemplo de instalação podem ser utilizados armazéns, usinas, subestações, escolas, hospitais e muitas outras estruturas comerciais, industriais ou públicas para os quais é necessário determinar a localização com o objetivo de otimizar sua interação com os outros elementos.

Além da localização da instalação é importante analisar também três outras características: número, tipo e custo associado.

Dependendo do modelo, o número de novas instalações pode ser um dado de entrada ou pode ser uma variável do problema.

A característica “tipo” envolve capacidade, serviço e considerações de infraestrutura. No caso mais simples, o problema de localização requer a localização de instalações idênticas com respeito à capacidade de atendimento e tipo de serviço suprido. Entretanto, em muitas aplicações é necessário locar simultaneamente instalações com diferenças entre elas. Um exemplo típico é dado por um sistema de distribuição onde armazéns e fábricas devem ser localizadas a fim de produzir e distribuir bens de forma eficiente. As fábricas e armazéns podem ter capacidade de fornecimento diferente dependendo da região em que serão localizadas. Considerando o tipo de serviço provido, os modelos de locação podem ser classificados em *single-service* e *multi-service*, baseado na informação de que as instalações podem prover somente um tipo ou vários tipos de serviços ou bens. Outra questão é o fato da locação admitir instalações que podem suprir uma demanda infinita em oposição a problemas que procuram

locar instalações com capacidade limitada de produção ou de suprimentos. Considerando este fator, os problemas podem ser classificados como não capacitado ou capacitado. A capacidade pode depender da região onde a instalação será localizada.

O custo associado com a instalação a ser locada é normalmente referente ao custo fixo de uma nova instalação e ao custo de manutenção da instalação em funcionamento. O custo fixo depende da localização da instalação, já o custo de manutenção depende não só da localização da instalação, mas também da distância da instalação aos clientes ou da dificuldade em atendê-los devido a problemas de terreno, clima, etc.

#### 2.4.3.2 LOCALIZAÇÕES CANDIDATAS

O segundo componente essencial para modelos de localização de instalações é a localização física onde as instalações podem ser instaladas, isto é, o conjunto de localizações candidatas a hospedar uma instalação.

#### 2.4.3.3 USUÁRIOS

O objetivo dos problemas de locação de instalações é justamente locar o recurso para melhor atender a demanda dos usuários. O nome usuário normalmente é utilizado para identificar as pessoas que demandam um certo serviço ou bem. No caso de localização de instalações o nome usuários também é utilizado para identificar qualquer objeto que precisa interagir com alguma instalação a ser locada. Por exemplo, os terminais remotos são considerados usuários de uma rede de comunicações.

Com respeito aos usuários em uma análise de locação de instalações é importante conhecer sua distribuição, sua demanda e seu perfil. Cada uma destas características deve ser analisada para definir o melhor método para atendê-los.

Considerando a distribuição dos usuários, deve ser analisado se os mesmos estão distribuídos uniformemente sobre um certo espaço geográfico ou se eles estão mais concentrados em determinados pontos do espaço ou ainda em certos vértices de uma rede.

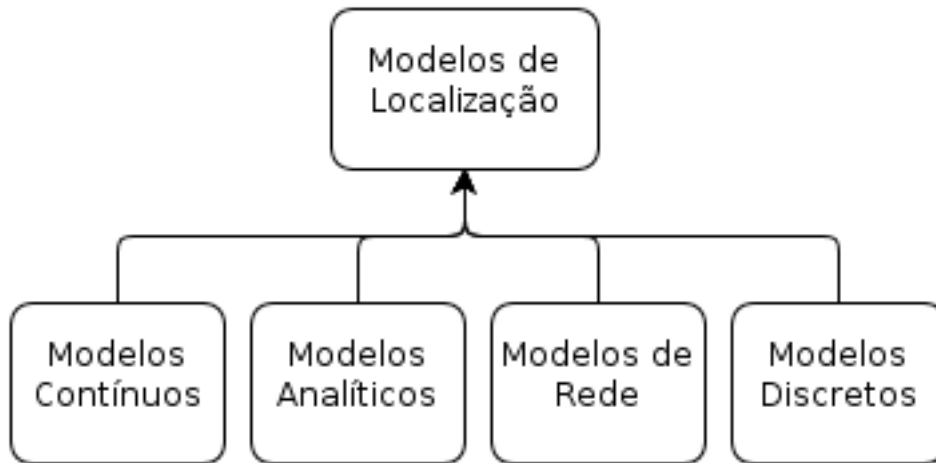
Com relação à demanda, para cada usuário é assumido um peso que reflete o total de serviços ou produtos que ele necessita ou demanda. Em certas situações, os usuários são agrupados para facilitar o trabalho de análise. O agrupamento pode ser feito por área administrativa, como bairros ou municípios, ou por área geográfica, como áreas de um quilômetro quadrado. Neste caso, cada agrupamento é considerado um usuário individual onde a demanda pode ser a soma das demandas de cada usuário real, a média das demandas ou outro

método de cálculo que estiver adaptado à situação em estudo.

Quanto ao perfil, em algumas aplicações os usuários são livres para escolher qual instalação vai servi-los. É bastante comum utilizar a distância entre a instalação e o usuário como fator importante na decisão sobre qual instalação vai atendê-lo.

#### 2.4.4 TAXONOMIA DOS MODELOS DE LOCALIZAÇÃO

Daskin (2008) apresenta uma taxonomia dos modelos modernos de localização baseado no espaço no qual o modelo é construído. Os modelos são classificados inicialmente em analíticos, contínuos, de rede e discretos, conforme a Figura 7.



**Figura 7: Taxonomia dos modelos de localização proposta por Daskin.**

**Fonte: Daskin (2008).**

##### 2.4.4.1 MODELOS ANALÍTICOS

Os modelos analíticos são modelos de localização simples. Estes modelos tipicamente assumem que a demanda está distribuída de alguma forma sobre o espaço e que as instalações podem ser localizadas em qualquer lugar dentro deste espaço. Estes modelos são resolvidos normalmente com o uso de cálculo ou outras técnicas simples. Apesar das características muito específicas destes modelos limitarem sua aplicação na realidade, as idéias derivadas podem ser aplicadas em diversos contextos (DASKIN, 2008).

#### 2.4.4.2 MODELOS CONTÍNUOS

Os modelos contínuos assumem que a demanda surge em pontos específicos do espaço. O problema de Weber (WEBER, 1929) é um exemplo típico desta classe de modelos, em que as demandas ocorrem em  $n$  pontos distribuídos no espaço. A localização de uma demanda  $i$  é dada por  $(x_i, y_i)$  para  $i = 1, 2, \dots, n$  e a intensidade é definida por  $h_i$ . O problema de Weber consiste em determinar a localização de uma instalação  $(X, Y)$  a fim de minimizar a soma das distâncias ponderadas desde os pontos de demandas até a instalação. A formulação matemática deste problema pode ser apresentada da seguinte forma: *Minimizar*  $\sum_{i=1}^n h_i \sqrt{(x_i - X)^2 + (y_i - Y)^2}$ . Este modelo é resolvido através da aplicação de procedimentos numéricos iterativos com o algoritmo de Weiszfeld (WEISZFELD, 1936). Drezner et al. (2004) analisa esta classe de modelos, o algoritmo de Weiszfeld e uma variedade de melhorias para garantir a convergência da solução. Plastria (2004) apresenta uma introdução a esta classe de modelos.

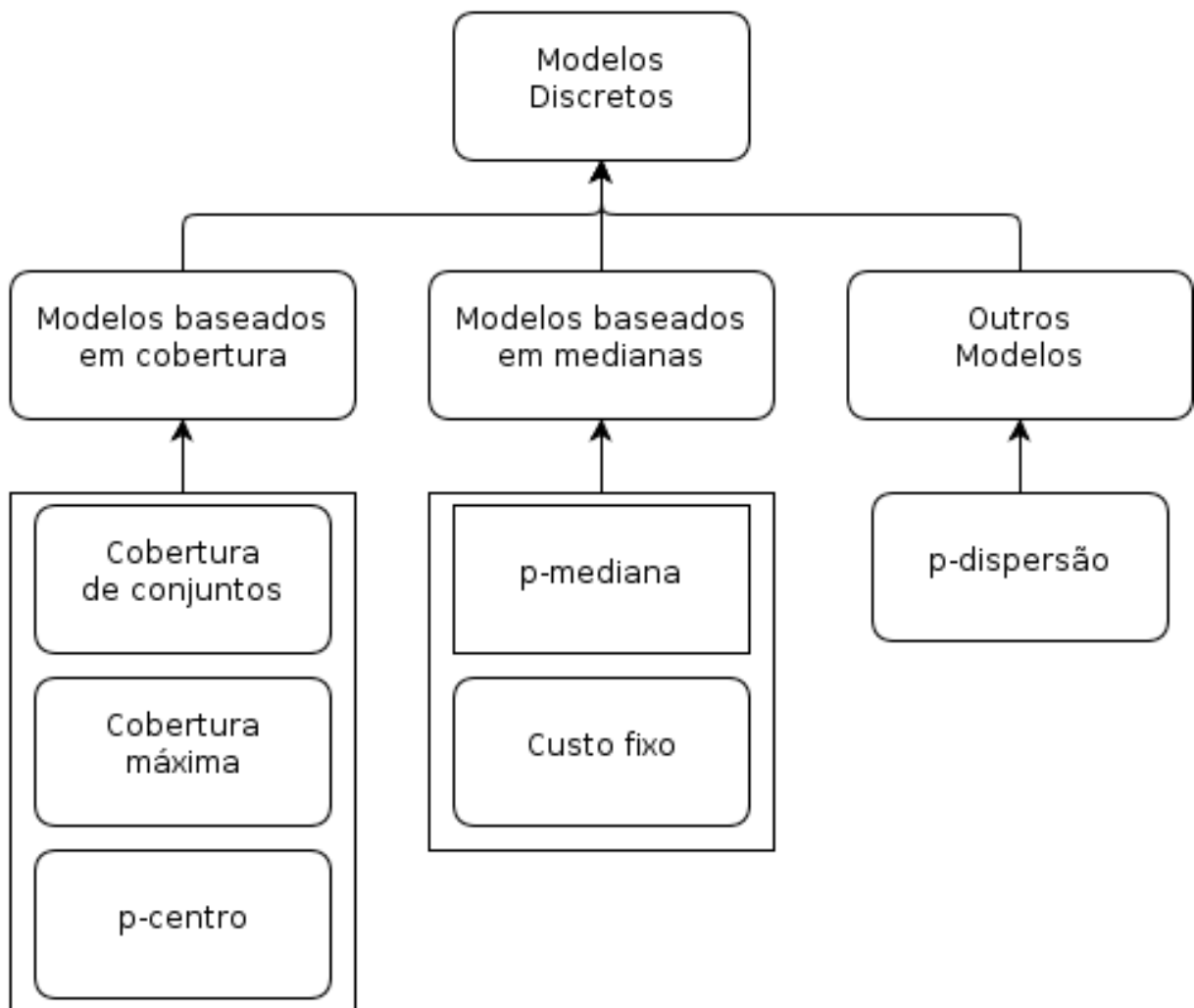
#### 2.4.4.3 MODELOS DE REDE

Os modelos de rede assumem que as demandas e instalações podem ser localizadas somente sobre uma rede composta por nós e arestas. Frequentemente as demandas surgem somente nos nós da rede, enquanto as instalações podem ser localizadas tanto nos nós quanto nas arestas da rede. O foco de boa parte da literatura é encontrar algoritmos com tempo de execução polinomial, geralmente para problemas em estruturas específicas como árvores (DASKIN, 2008), (TAMIR, 1996).

#### 2.4.4.4 MODELOS DISCRETOS

Os modelos discretos de localização surgem como a quarta classe de modelos. Nesta classe, os modelos caracterizam-se por terem ou não uma métrica de distância definida. A distância ou custo entre um par de nós pode ser arbitrária ou seguir determinada técnica (distância Euclidiana, Manhattan, rede grandes círculos de distância). As demandas geralmente surgem em nós e as instalações podem estar localizadas em um conjunto limitado e conhecido de lugares.

A Figura 8 apresenta a taxonomia para os modelos da classe dos modelos discretos.



**Figura 8: Taxonomia dos modelos de localização discretos proposta por Daskin.**

**Fonte: Daskin (2008).**

#### 2.4.5 MODELOS BASEADOS EM COBERTURA

Em alguns problemas de localização, uma distância máxima é conhecida *a priori*, e é chamada de distância de cobertura. Modelos baseados em cobertura utilizam esta informação para definir se uma demanda está coberta ou servida por determinada instalação ou conjunto de instalações. Nestes modelos, a distância entre uma demanda e a instalação mais próxima não faz parte da função objetivo mas é incluída na modelagem matemática como uma restrição. Portanto, a função objetivo envolve geralmente encontrar a quantidade mínima de instalações necessárias para cobrir todos os pontos de demanda ou maximizar a área de cobertura dado que o número de instalações é fixo.

### 2.4.5.1 COBERTURA DE CONJUNTOS

Modelos baseados em cobertura de conjunto assumem que existe uma distância crítica ou tempo limite dentro do qual um ponto de demanda ou cliente pode ser considerado coberto por uma instalação, por exemplo, o raio de cobertura de um sensor de temperatura pode ser considerado 50km. Estes modelos apresentam como resultado a quantidade de instalações necessárias para a cobertura total. Toregas et al. (1971) foi o pioneiro neste modelo com um trabalho sobre a localização de serviços de emergência com restrições de distância e tempo cuja solução utilizou programação linear (CURRENT et al., 2004).

O modelo de cobertura de conjunto pode ser expresso através dos seguintes conjuntos e parâmetros de entrada:

$I$  = conjunto de nós de demanda com índice  $i$

$J$  = conjunto de nós candidatos para hospedar uma instalação com índice  $j$

$d_{ij}$  = distância entre o nó de demanda  $i$  e o nó candidato a instalação  $j$

$D_c$  = distância de cobertura

$N_i = \{j | d_{ij} \leq D_c\}$  conjunto dos nós candidatos que podem cobrir o nó de demanda  $i$

e a variável de decisão

$$x_j = \begin{cases} 1 & \text{se instala no nó } j \\ 0 & \text{se não instala no nó } j \end{cases} \quad (5)$$

Pode-se assim, definir o problema de cobertura de conjunto da seguinte forma:

Minimizar

$$\sum_{j \in J} x_j \quad (10)$$



Sujeito a

$$\sum_{j \in N_i} x_j \geq 1, \forall i \in I \quad (11)$$

$$x_j \in 0, 1, \forall j \in J \quad (12)$$

$$(13)$$

A função objetivo 10 visa minimizar a quantidade de nós selecionados para instalação, pois conforme a equação 5,  $x_j = 1$  indica que o nó hospeda uma instalação, enquanto  $x_j = 0$  indica o contrário. A restrição 11 garante que cada nó de demanda é coberto por ao menos uma instalação. A restrição 12 define a natureza da variável  $x_j$  como binária (sim ou não). A função objetivo pode ser generalizada para incluir características específicas de cada nó candidato com o uso de pesos associados a cada nó. Neste caso, o objetivo torna-se minimizar o custo de instalação dos nós selecionados ao invés da quantidade de instalações. Garey e Johnson (1979) demonstraram que nos dois casos o problema é *NP – hard*.

#### 2.4.5.2 COBERTURA MÁXIMA

O problema de cobertura máxima é similar ao problema de cobertura de conjuntos com a diferença que a quantidade de instalações é conhecida e definida como entrada do modelo. Neste modelo, o objetivo é atender a maior área de cobertura possível com o número pré-definido de instalações. Observe-se aqui que o termo “maior” não está necessariamente vinculado a uma área euclidiana, mas a uma medida de demanda por parte do usuário.

Maximizar

$$\sum_{i \in I} h_i * Z_i \quad (19)$$

Sujeito a

$$\sum_{j \in J} X_j = p \quad (20)$$

$$Z_i - \sum_{j \in N_i} X_j \leq 0, \forall i \in I \quad (21)$$

$$X_j \in 0, 1, \forall j \in J \quad (22)$$

$$Z_i \in 0, 1, \forall i \in I \quad (23)$$

O modelo de cobertura máxima pode ser formulado utilizando as expressões (19, 20, 21, 22 e 23).

A função objetivo (19) maximiza o número de pontos de demanda atendidos, pois conforme a expressão 23,  $Z_j = 1$  indica que o nó esta coberto por uma instalação, enquanto  $Z_j = 0$  indica o contrário. A restrição (20) determina o número de instalações que serão instaladas. A restrição 21 associa os locais candidados com os pontos de demanda enquanto as restrições (22 e 23) definem o modelo com um problema de programação inteira binária.

#### 2.4.5.3 P-CENTRO

O modelo P-centro do inglês (*p-center*) tem por objetivo determinar a menor distância de cobertura possível de forma que todos os pontos de demanda possam ser cobertos.

#### 2.4.6 MODELOS BASEADOS EM MEDIANAS

Enquanto os modelos de cobertura geralmente tratam distância como um problema binário, os modelos de medianas trabalham com as distâncias reais (entende-se neste trabalho a distância real como a distância mensurada na unidade de medida selecionada para o modelo).

##### 2.4.6.1 P-MEDIANA

O modelo da p-mediana define a localização de  $p$  instalações de forma a minimizar a distância ponderada entre os pontos de demanda e a instalação mais próxima.

##### 2.4.6.2 CUSTO FIXO

O modelo de custo fixo é similar ao modelo de p-mediana considerando somente que o custo de instalação é igual para todas as localizações candidatas.

### 2.4.7 OUTROS MODELOS

O modelo p-dispersão maximiza a distância mínima entre duas instalações, evitando desta forma que duas ou mais instalações fiquem próximas demais dentro dos limites definidos no modelo.

## 2.5 COMPUTAÇÃO EVOLUCIONÁRIA (CE)

Segundo (WEINERT, 2010), “A CE abrange uma série de métodos computacionais inspirados na teoria da evolução das espécies, mais especificamente no conceito de seleção natural. Estes métodos manipulam soluções candidatas, também chamadas de indivíduos. Um conjunto de indivíduos constitui uma população. Estes indivíduos são apresentados a um ambiente, no caso o problema, onde os mais aptos sobreviverão. A evolução dos indivíduos, e conseqüente sobrevivência e adaptação dos mesmos, é realizada por meio de um processo iterativo de exploração do espaço de soluções através de métodos de seleção e operadores genéticos específicos. Cada processo iterativo constitui uma geração. Fazendo um paralelo com a evolução das espécies, uma nova geração acontece a cada nova reprodução, por meio dos descendentes. A aptidão de um indivíduo é mensurada por uma função de adequabilidade, conhecida na computação evolucionária como função de *fitness*”.

### 2.5.1 EVOLUÇÃO DIFERENCIAL (ED)

A evolução diferencial (*differential evolution*) é um método de otimização com uma estratégia evolucionária, que utiliza populações e é baseado no algoritmo *Genetic Annealing* desenvolvido por Kenneth Price e publicado em outubro de 1994 na revista *Dr. Dobb's Journal* (PRICE, 1994). A aplicação da *Genetic Annealing* ao problema de 5 dimensões de Chebyshev, proposta por Rainer Storn, levou ao desenvolvimento da evolução diferencial. Price et al. (2005) e Feoktistov (2006) apresentam ampla introdução ao tema, incluindo aplicações e implementações para análise.

#### 2.5.1.1 ESTRUTURA DA POPULAÇÃO

Uma implementação da evolução diferencial pode ser composta por três populações de vetores:  $P_x$ ,  $P_v$  e  $P_u$ , todas contendo  $N$  vetores de parâmetros do tipo real com dimensão  $D$ , onde:

- $P_x$  - população de vetores atual;

- $P_v$  - população de vetores produzidos no processo de mutação;
- $P_u$  - população de vetores produzidos no processo de combinação entre  $P_x$  e  $P_v$ .

A população atual,  $P_x$ , é composta por vetores,  $\mathbf{x}_{i,g}$ , cuja origem é a solução inicial do problema ou o resultado de comparação com outros vetores e pode ser representada da seguinte forma:

Seja

$$P_{x,g} = (\mathbf{x}_{i,g}) \quad (26)$$

$$\mathbf{x}_{i,g} = (x_{j,i,g}) \quad (27)$$

Onde

$$g = 0, 1, \dots, g_{max}$$

$$i = 0, 1, \dots, N - 1$$

$$j = 0, 1, \dots, D - 1$$

Nas expressões 26 e 27, o índice  $g$  indica a geração à qual o vetor pertence e  $g_{max}$  indica o número máximo de gerações (normalmente utilizado como critério de parada para o método). Além disto, cada vetor recebe um índice,  $i$ , dentro da população  $g$ , que varia de 0 a  $N - 1$ . Parâmetros dentro dos vetores são referenciados por meio do índice  $j$ , que varia de 0 a  $D - 1$ .

Uma vez inicializada, a evolução diferencial aplica o processo de mutação em vetores selecionados aleatoriamente, na população  $P_x$ , para produzir uma população intermediária,  $P_{v,g}$ , com vetores,  $\mathbf{v}_{i,g}$ , da forma:

Seja

$$P_{v,g} = (\mathbf{v}_{i,g}) \quad (30)$$

$$\mathbf{v}_{i,g} = (v_{j,i,g}) \quad (31)$$

Onde

$$\begin{aligned} g &= 0, 1, \dots, g_{max} \\ i &= 0, 1, \dots, N - 1 \\ j &= 0, 1, \dots, D - 1 \end{aligned}$$

Cada vetor na população atual,  $P_x$ , é então re combinado com um vetor da população  $P_v$ , para produzir uma população candidata,  $P_u$ , da forma:

Seja

$$P_{u,g} = (\mathbf{u}_{i,g}) \quad (34)$$

$$\mathbf{u}_{i,g} = (u_{j,i,g}) \quad (35)$$

Onde

$$\begin{aligned} g &= 0, 1, \dots, g_{max} \\ i &= 0, 1, \dots, N - 1 \\ j &= 0, 1, \dots, D - 1 \end{aligned}$$

### 2.5.1.2 PROCESSO DE INICIALIZAÇÃO

Antes que a população possa ser inicializada, ambos os limites superior e inferior devem ser especificados para cada parâmetro. Estes valores podem ser armazenados em dois vetores de inicialização de dimensão  $D$ ,  $b_L$  e  $b_U$ , onde os índices  $L$  e  $U$  indicam os limites inferior e superior respectivamente. Uma vez que os limites de inicialização são especificados, um gerador de números aleatórios associa a cada parâmetro de todos os vetores um valor dentro da faixa definida anteriormente. Por exemplo, o valor inicial ( $g = 0$ ) do  $j$  –ésimo parâmetro do  $i$  –ésimo vetor é:

$$x_{j,i,0} = rand_j(0, 1) * (b_{j,U} - b_{j,L}) + b_{j,L}. \quad (36)$$

O gerador de números aleatórios,  $rand_j(0, 1)$ , retorna um número aleatório uniformemente distribuído na faixa  $[0, 1)$ , ou seja,  $0 \leq rand_j(0, 1) < 1$ . O índice  $j$ , indica

que um novo valor aleatório é gerado para cada parâmetro. Mesmo que um parâmetro seja inteiro ou discreto, ele deve ser inicializado com um valor real, já que a evolução diferencial manipula todos os parâmetros como valores de ponto flutuante independente do seu tipo.

### 2.5.1.3 PROCESSO DE MUTAÇÃO

Uma vez inicializada a população, a evolução diferencial aplica o processo de mutação e recombinação para produzir a população  $P_u$ , vetores candidatos. Em particular, a evolução diferencial adiciona uma subtração vetorial a um terceiro vetor aleatoriamente selecionado. A expressão 37 apresenta como acontece a combinação entre três vetores diferentes entre si,  $\mathbf{x}_{r0,g}$ ,  $\mathbf{x}_{r1,g}$ ,  $\mathbf{x}_{r2,g}$ , escolhidos aleatoriamente na população  $P_x$ , para criar um novo vetor,  $\mathbf{v}_{i,g}$ , na população  $P_v$ :

$$\mathbf{v}_{i,g} = \mathbf{x}_{r0,g} + F * (\mathbf{x}_{r1,g} - \mathbf{x}_{r2,g}) \quad (37)$$

O fator de escala,  $F \in (0, 1+)$  é um número real positivo que controla a taxa na qual a população evolui. Enquanto não existe limite superior para o valor de  $F$ , valores efetivos raramente são maiores que 1,0.

### 2.5.1.4 PROCESSO DE COMBINAÇÃO

Para complementar a estratégia de busca baseada em mutação, a evolução diferencial aplica o processo de combinação uniforme. A combinação produz vetores candidatos, na população  $P_u$ , a partir de parâmetros que são copiados de dois diferentes vetores. Em particular, a evolução diferencial combina um vetor da população atual,  $P_x$ , com um vetor produzido no processo de mutação, na população  $P_v$ , conforme a expressão 38:

$$\mathbf{u}_{i,g} = u_{i,j,g} = \begin{cases} v_{j,i,g} & \text{if } (rand_j(0,1) \leq Cr \text{ or } j = j_{rand}) \\ x_{j,i,g} & \text{caso contrário} \end{cases} \quad (38)$$

A probabilidade de combinação,  $Cr \in [0, 1]$ , é um valor definido pelo usuário que controla a quantidade de parâmetros que são copiados no processo de mutação. Para determinar qual fonte de dados contribui para um dado parâmetro, a combinação uniforme compara o valor de  $Cr$  com a saída de um gerador de números aleatórios uniforme,  $rand_j(0,1)$ . Se o número aleatório é menor ou igual a  $Cr$ , o parâmetro do candidato é herdado do vetor da população,  $P_v$ , produto do processo de mutação,  $\mathbf{v}_{i,g}$ ; caso contrário, o parâmetro é copiado do vetor da

população  $P_x$ ,  $\mathbf{x}_{i,g}$ . Além disto, o parâmetro do vetor candidato com índice  $j_{rand}$  escolhido aleatoriamente é tomado do vetor da população  $P_v$ , para garantir que o vetor candidato não seja uma duplicação de  $\mathbf{x}_{i,g}$ .

### 2.5.1.5 PROCESSO DE SELEÇÃO

Se o vetor candidato,  $\mathbf{u}_{i,g}$  tem o valor de função objetivo menor ou igual ao melhor vetor conhecido,  $\mathbf{x}_{i,g}$ , então  $\mathbf{u}_{i,g}$  assume como o melhor vetor conhecido na próxima geração, caso contrário o melhor vetor conhecido continua sendo o mesmo, ao menos para a próxima geração. Uma vez que uma nova população é definida, o processo de mutação, combinação e seleção é repetido até que um valor satisfatório da função objetivo seja encontrado, ou um critério de parada seja satisfeito, por exemplo, número máximo de iterações.

### 2.5.1.6 PSEUDOCÓDIGO

O algoritmo 2 apresenta o pseudocódigo do método Evolução Diferencial.

---

#### Algoritmo 1: Pseudocódigo do método Evolução Diferencial

---

```

1 Parâmetros:  $N, D, Cr, F$ 
2 início
3   Gerar a população inicial de tamanho  $N$  chamada  $X$ 
4   enquanto critério de parada não é atendido faça
5     Gerar uma população chamada  $U$ 
6     para cada indivíduo  $x_i$  da população  $X$  faça
7       Selecionar três diferentes indivíduos da população  $X$ :  $x_{r0}, x_{r1}, x_{r2}$ ,
       todos diferentes entre si e do indivíduo  $x_i$ .
8       para cada elemento  $j$  do vetor que representa o indivíduo  $x_i$  faça
9         Calcula um valor aleatório  $rand$ 
10        se  $rand \leq Cr$  então
11           $u_{ji} = x_{jr0} + F * (x_{xr1} - x_{jr2})$ 
12        senão
13           $u_{ji} = x_{ji}$ 
14        para cada indivíduo  $x_i$  da população  $X$  faça
15          se  $f(u_i) \leq f(x_i)$  então
16             $x_i = u_i$ 

```

---

### 2.5.1.7 PARÂMETROS DA EVOLUÇÃO DIFERENCIAL

A evolução diferencial tem como característica o conjunto reduzido de parâmetros necessários para execução do método. Estes parâmetros são:

- $F$  - fator de valor constante e real;
- $Cr$  - probabilidade de cruzamento;
- $N$  - tamanho da população.
- $D$  - tamanho do vetor que representa um indivíduo da população

Segundo Storn (1996), as regras a seguir podem ser utilizadas como apoio para definição dos parâmetros.

- a população inicial deve ser gerada o mais próximo possível da superfície da função objetivo;
- deve-se considerar a probabilidade de cruzamento menor do que 0,3 por exemplo. Caso não ocorra convergência, utilizar um  $Cr \in [0, 8; 1]$ ;
- para muitas aplicações,  $N = 10 * D$  é uma boa escolha.  $F$  é, normalmente, escolhido no intervalo  $[0, 8; 1]$ ;
- quanto maior for o tamanho da população escolhida, menor é o valor de  $F$ ;
- tem-se um bom sinal de convergência quando os parâmetros do melhor membro da população varia muito de geração para geração, especialmente durante o início do processo de minimização. Mesmo se o seu valor da função objetivo decrescer lentamente;
- não há necessariamente um desempenho inapropriado, quando o valor da função objetivo do melhor membro da população apresentar platôs durante o processo de minimização. Entretanto, isto indica que a minimização levará um tempo maior até encontrar o mínimo global (ou próximo dele) ou que o aumento do tamanho da população poderá ser benéfico para a convergência;
- o valor da função objetivo do melhor membro da população não pode cair de forma brusca. Caso isto aconteça, a otimização está em um mínimo local;



## 2.6 TRABALHOS CORRELATOS

Segundo Amorim (2008), o planejamento e otimização de redes não é um assunto novo no panorama científico e existem vários estudos cujo objetivo é encontrar métodos que conduzam a uma estrutura de rede otimizada para monitoramento de algum fenômeno natural ou produzido pelo ser humano.

Na revisão da literatura relativa ao tema, agrupou-se as publicações de acordo com o objetivo da rede de monitoramento a ser otimizada.

### 2.6.1 MONITORAMENTO DE QUALIDADE DE ÁGUAS SUBTERRÂNEAS

Entre as publicações pesquisadas, parecem ser as redes de monitoramento da qualidade de águas subterrâneas as que até o momento despertaram maior atenção das pesquisas nesta área.

Li et al. (2004) desenvolveram um modelo híbrido de Algoritmos Genéticos e Otimização por Colônia de Formigas (ACO - *Ant Colony Optimization*) para monitoramento de longo prazo de água subterrânea a fim de maximizar a relação custo/benefício das amostras e o intervalo entre coletas.

Li e Hilton (2005a, 2005b, 2006a, 2006b) estenderam o modelo de Li et al. (2004) para minimizar o número de poços utilizados para coleta das amostras incluindo restrições de estimativa de erros e controle de qualidade dos dados. Este experimento alcançou resultados atingindo o ótimo global ou próximo a este.

Yeh et al. (2006) demonstram como a integração entre a análise geo-estatística multivariada e os algoritmos genéticos pode derivar numa metodologia eficaz para a definição da rede ótima de monitoração da qualidade de águas subterrâneas. O método desenvolvido minimiza a variância de estimação do fator espacial e, na opinião dos autores, proporciona informação suficiente, mas não redundante, de forma a perceber completamente o fenômeno espacial.

Dhar e Datta (2009) apresentam uma metodologia para projeto de redes de monitoramento de qualidade de água subterrânea utilizando uma formulação mista de programação linear e inteira. A metodologia proposta utilizou a técnica de Krigagem ordinária para estimação das concentrações de contaminação.

## 2.6.2 MONITORAMENTO DO MEIO AMBIENTE

As redes de monitoramento do meio ambiente também despertam bastante interesse entre os artigos identificados. Os parâmetros para avaliação das redes neste grupo dependem diretamente do dado a ser observado.

Ruiz-Cardenas et al. (2012) apresenta um problema de otimização de redes de monitoramento de ozônio na costa oeste do Estados Unidos. O método aplicado neste trabalho utilizou cadeias de Markov em conjunto com Recozimento Simulado. Segundo os autores, os resultados apresentados foram bastante satisfatórios considerando a área de extensão do trabalho.

Inoue et al. (2005) apresentam o problema de otimização de redes de monitoramento da contaminação do solo. O objetivo deste trabalho foi maximizar a probabilidade de detecção e cobrir as áreas onde produtos perigosos são transportados. Os autores aplicaram algoritmos genéticos em conjunto com o método estatístico *Latin Hypercube Sampling*.

Jimenez et al. (2005) apresentam um trabalho de otimização de rede de monitoramento de temperatura em um lago da Colômbia. O método aplicado envolve as técnicas de modelagem numérica, geo-estatística e computação evolucionária. Os autores dividem o processo em quatro etapas: (i) modelagem numérica para definição das variáveis, (ii) divisão da área de estudo em regiões, (iii) interpolação utilizando o método de Krigagem, (iv) otimização das redes avaliadas.

Martinez et al. (2010) apresentam o problema de localização de estações de monitoramento do nível de corpos d'água na superfície incluindo lagos e rios. O trabalho desenvolvido propõe a otimização da rede minimizando o número de pontos de coleta sem prejuízo da informação. Os autores aplicaram busca Tabu separadamente em lagos e rios e compararam os resultados com os dados observados. O algoritmo proposto foi integrado a um SIG criando uma ferramenta computacionalmente eficiente para otimização de redes de monitoramento.

Ruiz-Cardenas et al. (2010) apresentam um trabalho de otimização de rede de monitoramento de qualidade de água em rios que compara os resultados de três diferentes técnicas: Algoritmo Genético Híbrido (algoritmo genético padrão com um operador de busca local), o Recozimento Simulado e um Algoritmo Genético padrão. Nos resultados apresentados pelo autor, o algoritmo genético híbrido apresentou melhor resultado em todos os conjuntos de teste.

Icaga (2005) apresenta um trabalho de otimização de uma rede de monitoramento de poluentes na Turquia. Para otimização foram utilizados Algoritmos Genéticos e a melhor rede

de monitoramento apresentada conseguiu diminuir a quantidade de estações de 33 para 14.

### 2.6.3 MONITORAMENTO DE METEOROLOGIA

Entre as redes de monitoramento de meteorologia destacam-se as redes de radares meteorológicos, responsáveis pela previsão de curto prazo (menos de três horas) nos serviços meteorológicos.

Minciardi et al. (2003) modelam o problema de localização de uma rede de radares meteorológicos utilizando o problema de máxima cobertura com pesos visando a elaboração de uma metodologia a ser utilizada por um sistema de suporte a decisão.

Jaimes et al. (2009) modelam o problema de localização de uma torre de estação meteorológica utilizando otimização multi-critério com restrições.

Amorim et al. (2011) modelam o problema de localização de estações meteorológicas climatológicas em Portugal utilizando técnicas de interpolação com geo-estatística (Krigagem) para avaliação das soluções candidatas e os algoritmos Recozimento Simulado e *Greedy* para otimização da rede. A proposta dos autores envolve dois passos: (1) inicializar uma rede baseada na rede já existente e reduzir o número de estações que compõem a rede baseado na redundância da informação. (2) a partir da rede reduzida, iniciar o processo de otimização visando preencher as regiões onde a incerteza associada ao processos geo-estatísticos são maiores.

### 2.6.4 OUTROS

Cameron (2004) apresenta o problema de otimização de rede de monitoramento de longo prazo LTM (*Long-term monitoring* - Monitoramento de longo prazo) alterando o foco de adição de informações e dados para enfatizar a identificação e minimização da redundância estatística de forma a reduzir os custos a longo prazo. Um sistema ótimo é definido como sendo o com menor perda de informação e maior ganho em termos de economia de custos.

Ratle (2001) apresenta considerações sobre o impacto de avaliar as soluções geradas pelos algoritmos evolucionários através de métodos numéricos computacionalmente intensivos e apresenta o método de interpolação por Krigagem como uma alternativa viável para mensurar a qualidade dos soluções avaliadas por algoritmos evolucionários.

Jaramillo et al. (2002) apresenta uma comparação entre algoritmos genéticos e heurísticas tradicionais para resolução de problemas de localização de instalações.

### 3 METODOLOGIA

Considerando os objetivos deste trabalho, apresentados no capítulo 1, e a fundamentação teórica relacionada ao tema, apresentada no capítulo 2, o capítulo atual analisa o problema de planejamento da instalação de estações meteorológicas. Ao conjunto das estações meteorológicas instaladas, chama-se de rede de estações meteorológicas, pois, apesar do fato de que as estações são operacionalmente independentes entre si, considera-se que as mesmas trabalham de forma conjunta para coletar dados que representem a atmosfera da melhor forma possível.

#### 3.1 O AMBIENTE DE PESQUISA

Para a realização dos experimentos do modelo proposto e avaliação da viabilidade da sua implantação, utilizou-se neste trabalho como ambiente de pesquisa o Sistema Meteorológico do Paraná - SIMEPAR.

O SIMEPAR está situado em Curitiba, no Centro Politécnico da Universidade Federal do Paraná. Foi instituído em 1993, por meio do Decreto Estadual nº 1752, sendo mantido por sete anos pelo convênio firmado pela COPEL – Companhia Paranaense de Energia e o IAPAR – Instituto Agrônomo do Paraná, com a interveniência da Universidade Federal do Paraná. Devido ao processo de reestruturação do setor energético e da necessidade de autonomia para buscar novos negócios, no ano 2.000, por meio do Decreto Estadual nº 2047 as atividades do SIMEPAR foram transferidas para o Instituto Tecnológico vinculado a Seti - Secretária de Ciência, Tecnologia e Ensino Superior do Estado do Paraná.

Este instituto tem por finalidade fornecer ao público em geral informações de natureza meteorológica, hidrológica e ambiental, bem como conceber, desenvolver e executar projetos de pesquisa científica e tecnológica, formação e capacitação profissional.

No que se referem às previsões, estas são realizadas com base em telemetria e sensoriamento remoto, modelos numéricos e da análise crítica de meteorologistas. Para

apoiar a interpretação destes profissionais especialistas na climatologia da região, que resulta no produto final de previsão, existe um conjunto integrado de sistemas de monitoramento, composto por estações meteorológicas de superfície, sistemas de detecção de raios, radares e satélites, que transmite dados de forma automática ao centro computacional. Os dados são integrados, armazenados e processados por programas computacionais gerando as informações. Complementando sua concepção funcional, o SIMEPAR conta com suporte computacional para o desenvolvimento de aplicativos utilizados em monitoramento e previsão (SIMEPAR, 2013).

## 3.2 MATERIAIS

As próximas subseções apresentam as ferramentas computacionais utilizadas para a execução dos experimentos.

### 3.2.1 JSON

O JSON, do inglês *JavaScript Object Notation* e traduzido como Notação de Objetos JavaScript, é um formato de texto para trocas de dados, independe da linguagem de programação, pois usa convenções que são familiares as linguagens C, C++, C#, Java, JavaScript, Perl, Python e outras. JSON se constitui de uma estrutura de coleção de pares nome/valor e uma lista ordenada de valores (JSON, 2013).

### 3.2.2 LP-SOLVE

O `lp_solve` é um programa para resolução de problemas de Programação Linear Inteira Mista e está baseado nos métodos *Simplex* e *Branch & Bound*. Através do método *Branch & Bound*, é possível lidar com variáveis inteiras, semi-variáveis contínuas e conjuntos especiais ordenados. Basicamente trata-se de uma biblioteca com um conjunto de rotinas, chamado de API que pode ser utilizado por diferentes linguagens de programação, como: C, VB, .NET, Delphi, Java entre outras (BERKELAAR et al., 2004).

### 3.2.3 POSTGRESQL

O PostgreSQL é um sistema de banco de dados objeto-relacional de código aberto. Este banco de dados pode ser executado na maioria dos sistemas operacionais, incluindo Linux (AIX, BSD, HP-UX, SGI IRIX, Mac OS X, Solaris, Tru64) e Windows. Tem suporte completo para chaves estrangeiras, associações, visões, gatilhos e store procedures (em múltiplas

linguagens). Inclui suporte para a maioria dos tipos definidos no padrão SQL:2008, incluindo INTEGER, NUMERIC, BOOLEAN, CHAR, VARCHAR, DATE, INTERVAL e TIMESTAMP. Também possui suporte para armazenamento no formato binário de grandes objetos como imagens, sons ou vídeos. Possui interface nativa para programação C/C++, Java, .Net, Python, Ruby, Tcl, ODBC entre outros e uma documentação completa (POSTGRESQL, 2013).

#### 3.2.4 POSTGIS

O PostGIS é uma extensão espacial para o banco de dados PostgreSQL, que adiciona suporte para objetos geográficos e permite o uso de operadores espaciais nas consultas e outros comandos (POSTGIS, 2013).

#### 3.2.5 PYTHON

Python é uma linguagem de programação que utiliza o paradigma de Orientação a Objetos. Esta linguagem é utilizada em uma variedade de aplicações e tem como características: sintaxe clara, orientação a objetos intuitiva, extensas bibliotecas padrão e módulos para diversas tarefas. Várias distribuições Linux vem com o Python instalado. A PSF, do inglês *Python Software Foundation*, traduzido para Fundação de Software Python, é uma organização sem fins lucrativos que detém os direitos de propriedade intelectual da linguagem de programação Python (PYTHON, 2013).

#### 3.2.6 QGIS

O Quantum GIS é um software livre de SIG, desenvolvido em 2002 pela OSGEO, do inglês *Open Source Geospatial Foundation*, traduzido para Fundação de Código Fonte Aberto Geoespacial. As funções básicas do QGIS incluem a geração de mapas, visualização, criação e edição de dados espaciais, estruturação de dados em SIG. Este software suporta os formatos de dados vetoriais (como *Shapefile*), dados raster (como modelos digitais de elevação do terreno), e de banco de dados, como o PostGIS (QGIS, 2013).

#### 3.2.7 SHAPEFILE

O *shapefile* é um formato de arquivo com extensão shp de dados espaciais desenvolvido pela ESRI, do inglês *Environmental Systems Research Institute*, com uma especificação aberta para interoperabilidade por dados entre os software desenvolvidos pela ESRI como também

de outros fornecedores. Um *shapefile* suporta características de pontos, linhas e polígonos, os polígonos são usados para representar áreas e as linhas e pontos podem representar rios e lagos.

Cada item do *shapefile* pode ter atributos que os descrevem como a temperatura ou a profundidade, sendo que cada atributo tem um registro um para um com o item da forma associada (ESRI, 1998)

### 3.2.8 SHAPELY

Shapely é um conjunto de módulos da linguagem de programação Python utilizado para manipulação e análise de objetos geométricos planos, que possui interface com a biblioteca GEOS, do inglês *Geometry Engine Open Source*, permitindo a execução de operações espaciais. Os tipos fundamentais de objetos geométricos são implementadas como pontos, curvas e superfícies (SHAPELY, 2013).

## 3.3 O MÉTODO PROPOSTO

Neste trabalho, propõem-se um método para elaboração de um plano de implantação de redes de estações meteorológicas. O método é composto por etapas que são apresentadas com os respectivos parâmetros de entrada e saída. As etapas são independentes entre si, permitindo a substituição do conteúdo de uma etapa por outro similar, desde que, os parâmetros de entrada e saída sejam mantidos.

O planejamento da instalação de uma rede de estações meteorológicas pode ser dividido em duas fases:

- Fase 1: determinação da quantidade mínima e localização das estações necessárias para a cobertura de todos os locais de demanda;
- Fase 2: determinação da ordem de instalação das estações visando maximizar a área de cobertura da rede a cada nova instalação.

Na primeira fase, o profissional responsável, chamado planejador, procura determinar a quantidade mínima de estações necessária para atender a região de interesse. Esta região de interesse, no escopo deste trabalho, é o Estado do Paraná. A determinação da quantidade mínima de estações envolve a análise:

- dos locais onde podem ser instaladas as estações, chamados de locais candidatos;

- das características logísticas, como vias de acesso ao local de instalação e a proximidade de áreas urbanas;
- das características meteorológicas, como locais cuja coleta de dados represente com fidelidade as condições da atmosfera;
- das características estratégicas, como prevenção de desastres naturais, em que torna-se importante coletar dados meteorológicos de locais como áreas de risco indicadas pela Defesa Civil.

Inicialmente, em uma abordagem na qual considera-se todos os locais candidatos com o mesmo valor ou peso associado, a função objetivo pode ser representada pela equação 39.

$$\text{Min} \quad \sum_{i=0}^n x_i \quad (39)$$

onde  $n$  é a quantidade de locais candidatos e  $x_i$  representa os locais candidatos e assume valor “1”, caso receba uma estação e valor “0” caso contrário. Na equação 39 todos os locais candidatos tem o mesmo valor ou peso.

Entretanto, ao considerar as características logísticas, meteorológicas e estratégicas como valores ou pesos  $h$  que podem ser diferentes para cada local candidato  $h_i$ , substitui-se o objetivo da fase. Ao invés da minimização da quantidade de estações meteorológicas conforme equação 39, busca-se minimizar uma função objetivo que considera também os pesos  $h_i$ , conforme equação 40

$$\text{Min} \quad \sum_{i=0}^n h_i * x_i \quad (40)$$

onde  $n$  é a quantidade de locais candidatos,  $h_i$  é o valor ou peso associado ao local candidato  $i$  e  $x_i$  representa os locais candidatos e assume valor “1”, caso receba uma estação e valor “0” caso contrário.

Na segunda fase, o planejador determina a ordem de instalação das estações, visando maximizar a área de cobertura da rede a cada nova instalação. Como na fase 1, deve-se considerar as características logísticas, meteorológicas e estratégicas dos locais candidatos.

Assim, o objetivo da fase 2 é maximizar a área cobertura da rede de estações considerando uma quantidade limitada de equipamentos disponíveis.



Maximizar

$$\sum_{i=0}^n h_i * z_i \quad (43)$$

Sujeito a

$$\sum_{i=0}^n x_i = p \quad (44)$$

As equações 43 e 44 apresentam a formulação matemática deste objetivo, onde:

- $z_i$  pode assumir o valor 1, caso o local candidato esteja coberto por alguma estação, e o valor 0, caso contrário,
- $x_i$  pode assumir o valor 1, caso o local candidato receba uma estação, e o valor 0, caso contrário;
- $p$  é a quantidade de estações que a empresa/instituição tem capacidade de instalar.

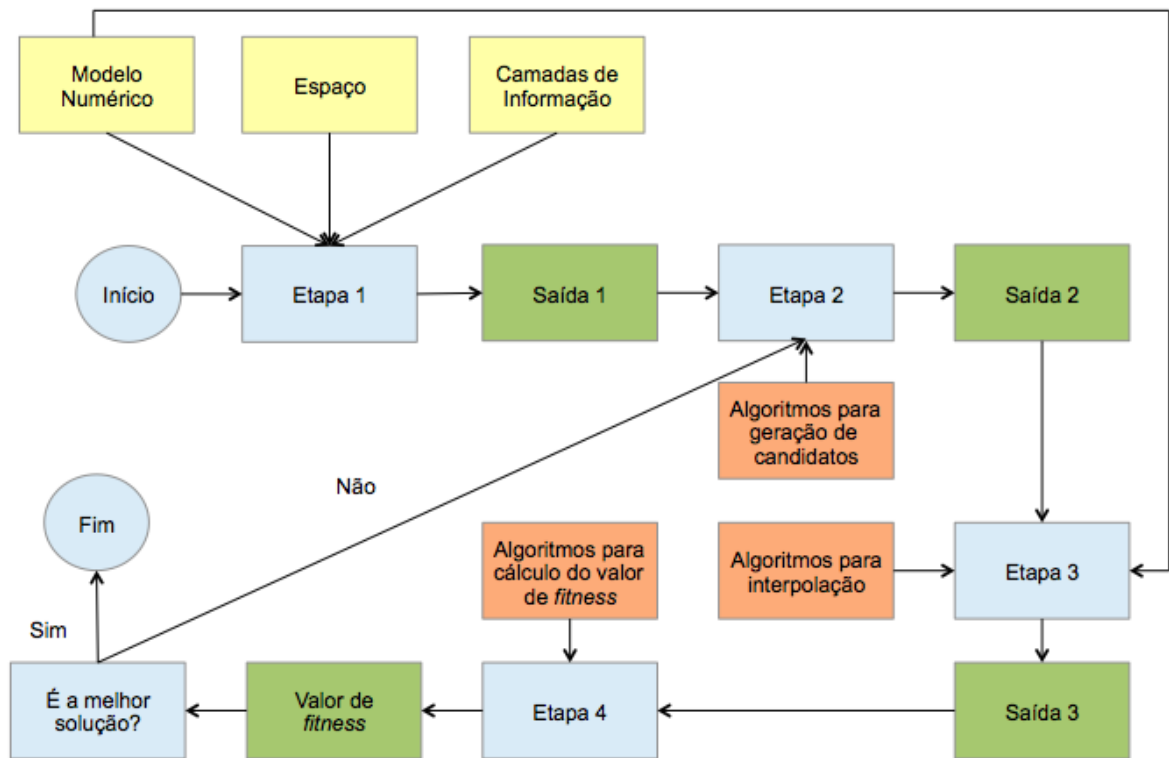
A Figura 9 apresenta o diagrama do fluxo de informações do método proposto. O método é composto por quatro etapas que possuem entradas e saídas explicitamente definidas de forma a manter o baixo acoplamento entre as etapas e permitir a substituição do conteúdo de cada por outro similar, sem alteração no método.

### 3.3.1 ETAPA 1 - DEFINIÇÃO DOS LOCAIS CANDIDATOS

A primeira etapa do método é responsável por definir a grade regular de pontos que serão utilizados como os locais candidatos à instalação de uma estação meteorológica e extrair as características deste locais.

Esta etapa possui três parâmetros de entrada:

- Modelo Numérico - é responsável por definir o espaçamento horizontal da grade regular de pontos
- Espaço - é responsável por definir a região onde a grade regular pontos esta contida
- Camadas de Informação - são responsáveis por fornecer as características para cada elemento da grade regular de pontos



**Figura 9: Modelo esquemático do método proposto para o problema de planejamento da implantação de uma rede de estações.**

Fonte: O Autor (2013).

O Algoritmo 2 apresenta o fluxo de ações da etapa 1.

Conforme o Algoritmo 2, a etapa 1 inicia com a leitura dos parâmetros de entrada. O Modelo Numérico de Previsão do Tempo é utilizado para definir o espaçamento horizontal da grade de pontos que será utilizada na simulação. O Espaço é utilizado para delimitar a área da simulação onde a grade de pontos é construída. Com estes parâmetros a etapa 1 cria a lista de pontos. Após a criação da lista de pontos, o processo extrai para cada ponto, o dado armazenado nas Camadas de Informação Geográfica. Este processo faz uso de rotinas de geoprocessamento para ler as Camada de Informação de formato *raster* e vetorial. O resultado da etapa 1 é a lista dos pontos que têm associado a cada elemento da lista, as informações extraídas das Camadas de Informação Geográfica, conforme apresentado na Figura 10. Os pontos desta lista também são chamados locais candidatos. A saída da etapa 1 é utilizada como entrada da etapa 2.

Os parâmetros de entrada são as informações iniciais necessárias para a determinação do plano de implantação da rede de estações.

Na implantação de uma rede de estações meteorológicas, em que um dos objetivos é caracterizar a atmosfera, existe o desafio relacionado à avaliação da contribuição da inclusão de

---

**Algoritmo 2:** Pseudocódigo da etapa 1
 

---

**Data:** Parâmetros de entrada

**Result:** Grade regular de pontos com lista de características associadas a cada elemento.

```

1 início
2   Ler o parâmetro de entrada Modelo Numérico
3   Gerar uma lista dos pontos a partir da grade utilizada no Modelo Numérico
   de Previsão do Tempo e do Espaço
4   Inicializar a lista com as Camadas de Informação Geográfica
5   enquanto existir Camadas de Informação Geográfica na lista faça
6     Inicializar a lista de pontos;
7     enquanto existir pontos na lista faça
8       Extrair o dado da Camada de Informação Geográfica para o ponto;
9       Excluir o ponto da lista;
10    fim
11    Excluir a Camada de Informação Geográfica da lista;
12  fim
13 fim

```

---

uma nova estação a uma rede existente (AMORIM, 2008). Diversas abordagens de avaliação foram apresentadas no capítulo anterior. Para satisfazer o problema de avaliação, o método proposto utiliza o resultado do Modelo Numérico de Previsão do Tempo, que é um produto de um processo de cálculo das variáveis meteorológicas para cada vértice de uma grade horizontal regular e portanto, fornece um dado sobre uma variável meteorológica em qualquer ponto do espaço de abrangência do Modelo. O objetivo do uso do Modelo Numérico é permitir a consulta do valor de uma variável meteorológica em um local candidato a instalação. Desta maneira, torna-se possível estimar o valor de uma variável meteorológica em um local candidato durante a simulação das opções de redes de estações. Com esta informação, é proposta uma forma de avaliação da rede de estações nas seções seguintes.

O Modelo Numérico de Previsão do Tempo utiliza uma grade horizontal regular. Para cada vértice desta grade são calculados os valores das variáveis meteorológicas. A abordagem aplicada no método proposto, restringe os locais candidatos aos vértices da grade utilizada pelo Modelo Numérico. Assim, transforma-se o problema de planejamento da implantação de uma rede de estações de um espaço contínuo, onde a estação pode ser instalada em qualquer local, para um espaço discreto, onde as estações podem ser instaladas somente nos vértices desta grade. De acordo com Daskin (2008), a opção por discretizar o espaço disponível para os locais candidatos, coloca este problema na classe dos modelos de localização discretos. Desta

```

{
  "0": {
    "altclasse": 5,
    "altitude": 879.0,
    "cod": 0,
    "estrada": 0,
    "lat": -25.3,
    "lon": -49.5,
    "municipio": 0,
    "municipio_buffer": 1
  },
  "1": {
    "altclasse": 3,
    "altitude": 1026.0,
    "cod": 1,
    "estrada": 0,
    "lat": -25.3,
    "lon": -49.4,
    "municipio": 0,
    "municipio_buffer": 1
  },
  ...
  '9": {
    "altclasse": 1,
    "altitude": 886.0,
    "cod": 9,
    "estrada": 0,
    "lat": -25.4,
    "lon": -49.1,
    "municipio": 0,
    "municipio_buffer": 1
  }
}

```

**Figura 10: Saída da etapa 1 - Lista de locais candidatos e características associadas.**  
**Fonte: O Autor (2013).**

maneira, pode-se abordar o problema utilizando modelos de cobertura de conjunto, cobertura máxima ou p-medianas (DREZNER; HAMACHER, 2004).

O Espaço é a região que o método considera como disponível para a instalação das estações. Considera-se que as demandas também estão discretizadas nos mesmos vértices da grade de pontos do Modelo Numérico. Assim, os pontos de demanda coincidem com os locais candidatos a instalação. Como o método proposto utiliza o Modelo Numérico de Previsão do Tempo como referência para a realidade da atmosfera, é necessário que a grade de pontos da simulação utilize a mesma grade de pontos do modelo. Desta forma, a grade utilizada nas simulações é regular e coincidente com a grade de pontos do Modelo Numérico. Neste método, a demanda no espaço é definida como distribuída uniformemente na superfície e é discretizada em uma grade regular onde os vértices são coincidentes com os vértices dos locais candidatos.

As Camadas de Informação Geográfica são responsáveis por caracterizar os locais

candidatos à instalação. O local de instalação de uma estação meteorológica é definido a partir de critérios baseados nas características do local e seus arredores. O método proposto permite a extração destas características das Camadas de Informação Geográfica. Por exemplo, dado um local candidato  $x_i$ , localizado na coordenada  $(lon, lat)$  e uma Camada de Informação Geográfica referente ao relevo do terreno, o método propõe um processo de extração do valor da altitude na coordenada  $(lon, lat)$ . Desta forma, é possível associar a característica de altitude ao local candidato. Esta operação pode ser realizada para todos os vértices da grade e todas as camadas de informação geográfica. O número de camadas é variável de acordo com a disponibilidade da informação e o método pode ser aplicado com nenhuma, uma ou mais camadas. As camadas de informação são compostas por dados georreferenciados que podem ser contínuos ou discretos no espaço. No caso dos dados georreferenciados contínuos, é necessário discretizar a informação para a grade utilizada no espaço do problema. Estas camadas também podem ser classificadas de acordo com o formato de armazenamento em: *raster* ou vetorial. As camadas vetoriais podem ser classificadas em: pontos, linhas e polígonos.

Cada fonte de dado vetorial representa uma característica que pode ser associada aos locais candidatos. Cada classe de uma característica deve possuir um coeficiente associado. Este coeficiente deve refletir o peso que cada característica deve ter no processo de geração de soluções candidatas e é discutido em mais detalhes na subseção 3.3.5.

As características podem ser agrupadas em características logísticas como:

- proximidade de área urbanas;
- proximidade de vias de acesso;

ou como características físicas do ambiente:

- orografia;
- uso do solo;
- tipo de solo;
- vegetação;

ou características estratégicas como:

- proximidade de áreas de agricultura,

- proximidade de áreas de usinas hidroelétricas e eólicas,
- proximidade de áreas de risco de desastres naturais.

Desta forma, é possível criar um conjunto de características com valores distintos para cada local candidato. Do ponto de vista do monitoramento dos fenômenos meteorológicos, deseja-se implantar uma rede de estações possa cobrir todas os locais candidatos, ou caso não seja possível, cobrir a maior quantidade possível de locais candidatos que representem as diferentes condições da atmosfera (SILVEIRA, 2010). Deste ponto de vista, dada uma limitação de somente 2 estações para instalação, deve-se procurar locais candidatos que possam representar condições distintas da atmosfera ao invés de condições similares.

As características logísticas envolvem a facilidade de deslocamento e acesso aos locais candidatos. Estradas ou áreas urbanas próximas indicam que um local pode ser mais adequado que outro sem acesso por estrada ou distante da área urbana mais próxima.

As características físicas envolvem as condições do espaço selecionado para instalação da rede e são geralmente condições com pequena possibilidade de alteração, como o relevo de um terreno ou o tipo de solo.

As características estratégicas envolvem políticas públicas, como monitoramento de risco de geada no campo, monitoramento da quantidade de precipitação próximo às barragens com influência direta na operação de usinas hidroelétricas, monitoramento do regime de ventos de forma a indicar áreas mais adequadas para instalação de parques eólicos ou monitoramento de variáveis relacionadas a desastres naturais como precipitação e vento.

### 3.3.2 ETAPA 2 - GERAÇÃO DAS SOLUÇÕES CANDIDATAS

A etapa 2 tem como parâmetro de entrada o resultado da etapa anterior e é responsável por gerar soluções candidatas baseado nas informações disponíveis, como as características disponíveis do locais candidatos e o peso que cada característica deve assumir na geração das soluções.

Este processo faz uso do algoritmos do grupo I, que são algoritmos geradores de soluções candidatas. Dentro deste grupo é possível dispor de qualquer método que resolva o problema de planejamento na instalação de redes de estações meteorológicas. Este trabalho aborda o uso do método *Branch & Bound* e a estratégia evolutiva Evolução Diferencial, mas é possível expandir o método para fazer uso de Algoritmos Genéticos, inteligência de enxame, heurísticas tradicionais ao tema de localização de instalações como p-medianas, entre outros.

A possibilidade de trocar o algoritmo gerador de soluções é um ponto positivo no método proposto, pois permite a avaliação de diferentes técnicas sem a necessidade de reescrever todo o programa.

As implementações dos algoritmos devem ser adaptadas para fazer uso das características disponíveis nas camadas de informação geográfica.

O método proposto prioriza o baixo acoplamento entre estes algoritmos e as demais etapas, para que uma alteração em um algoritmo não implique em alterações das demais etapas. Com isto, permite-se também explorar o método proposto com diversos algoritmos, bastando que as entradas e saídas entre as etapas sejam respeitadas.

O resultado do etapa 2 é uma ou mais soluções para o problema de definição da estrutura da rede de estações. O produto é formado por uma lista dos locais candidatos com valores binários associados, indicando se o local hospeda ou não uma estação. O método utiliza o valor “0” para quando o local não hospeda uma instalação e “1” para o caso contrário.

### 3.3.3 ETAPA 3 - GERAÇÃO DAS INTERPOLAÇÕES

Esta etapa é responsável por extrair valores do Modelo Numérico para o locais candidatos indicados no resultado da etapa anterior para instalação de estações meteorológicas. Desta forma, o método proposto simula o valor de uma variável meteorológica que a estação coletaria caso fosse instalada no local candidato. Após a extração destes valores para cada local candidato indicado para instalação de estações, a etapa utiliza um algoritmo de interpolação espacial para calcular o valor da variável nos demais locais candidatos que não hospedam uma estação meteorológica.

Conforme discutido anteriormente, o objetivo do uso do modelo numérico de previsão do tempo no modelo proposto é poder utilizá-lo como referência na avaliação da qualidade das soluções candidatas. Para isto, o método propõe a geração de uma grade horizontal regular igual à do modelo numérico, cujos os valores nos vértices são produzidos pelo algoritmo de interpolação.

Os algoritmos de interpolação espacial recebem o resultado da etapa anterior, onde estão as indicações dos locais candidatos que recebem uma estação, extraem do modelo numérico os valores das variáveis meteorológicas em uso e associam ao local candidato e aplicam uma técnica de interpolação espacial nos dados, associando aos locais candidatos que não tem estação um valor oriundo da interpolação. Desta forma, o produto deste grupo de algoritmos é uma grade regular coincidente com a grade do modelo numérico, na qual os

vértices selecionados para hospedar uma estação possuem valor da variável meteorológica com origem no modelo numérico e nos vértices restante o valor tem origem na interpolação.

Neste trabalho, utiliza-se o algoritmo de interpolação espacial de Barnes, também conhecido como análise de Barnes. É uma técnica proposta por Barnes (1964) para gerar uma superfície estimada para variáveis meteorológicas. Esta técnica utiliza uma rodada inicial para gerar uma superfície suavizada e de forma opcional, rodadas subsequentes para melhorar a superfície ajustando os valores da interpolação para valores mais próximos dos observados, ou no caso deste trabalho, dos valores do modelo numérico.

Os refinamentos subsequentes calculam uma superfície de erro utilizando a diferença entre a estimativa da rodada anterior e o valor da observação. A superfície de erro calculada é adicionada à superfície anterior para refinar a estimativa.

O Algoritmo 3 apresenta o pseudo código do algoritmo de interpolação de Barnes.

---

**Algoritmo 3:** Pseudocódigo do método de interpolação espacial de Barnes

---

- 1 Para a primeira rodada, o valor estimado  $E_g$  em cada vértice da grade  $g_{x,y}$  é:
  - 2 
$$E_g = \frac{\sum(w_i * o_i)}{\sum w_i}$$
  - 3 onde:
  - 4  $o_i$  é o valor da variável meteorológica no modelo numérico para o  $i$ -ésimo local candidato com estação
  - 5  $w_i$  é o valor do peso para o  $i$ -ésimo local candidato com estação definido como
  - 6 
$$w_i = \exp \frac{-d_i^2}{L^2 C}$$
  - 7 onde:
  - 8  $d_i$  é a distancia entre o vértice analisado e o  $i$ -ésimo local candidato com estação
  - 9  $L$  é a escala de comprimento, que é determinada pelo espaçamento da observação e a escala natural do fenômeno que está sendo analisado.
  - 10  $C$  é o fator de convergência, que controla quanto refinamento é realizado a cada rodada. Na primeira rodada o fator de convergência é 1. Nas rodadas seguintes o valor variando entre 0.2 e 0.3 é efetivo
  - 11 Durante as rodadas de refinamento, a estimativa para cada vértice é recalculada usando:
  - 12 
$$E'_g = E_g + \frac{\sum(w_i * (o_i - E_i))}{\sum w_i}$$
  - 13 onde:
  - 14  $E_i$  é o valor estimado no vértice que contém o  $i$ -ésimo local candidato com estação.
- 

O método de interpolação espacial de Barnes é utilizado neste trabalho por ser um método geralmente implementado nas ferramentas de visualização meteorológica (MAGICS++, 2013).

O resultado desta etapa é uma grade regular coincidente com a grade do modelo



numérico, com valores do modelo nos vértices selecionados para instalação de estação e valores interpolados nos vértices restantes.

Neste etapa, também é possível substituir o algoritmo de interpolação utilizado sem impacto nas outras etapas. Desta forma, é possível avaliar diferentes técnicas de interpolação sem reescrever o programa completamente, bastando que os parâmetros de entrada e saída sejam respeitados.

### 3.3.4 ETAPA 4 - AVALIAÇÃO DAS SOLUÇÕES CANDIDATAS

A etapa 4 é responsável pelo cálculo do valor de *fitness* das soluções candidatas. Esta etapa utiliza o resultado da etapa anterior em conjunto com os algoritmos para cálculo do valor de *fitness* das soluções.

O método proposto utiliza medidas de nível de qualidade da solução baseadas na comparação entre as grades do modelo numérico e do resultado da etapa 3. O método proposto utiliza inicialmente a raiz quadrada do erro quadrático médio normalizado, do inglês *Normalized Root Mean Squared Error* - NRMSE. Esta abordagem penaliza soluções candidatas cuja interpolação gera uma superfície diferente da superfície do modelo numérico.

A NRMSE é o algoritmo disponível para cálculo do valor de *fitness* na prova de conceito e a sua expressão matemática é apresentada na equação 45.

$$NRMSE = \sqrt{\frac{\sum_{i=1}^N (o_i - c_i)^2}{\max(o) - \min(o)}} \quad (45)$$

onde:  $o$  são os valores lidos do modelo numérico,  $c$  são os valores calculados pelo algoritmo de interpolação e  $N$  é a quantidade de locais candidatos existentes no espaço do problema.

O NRMSE é apresentado em porcentagem e valores próximos do zero indicam que os valores calculados pelo algoritmo de interpolação estão próximos dos valores do modelo numérico.

Após a etapa 4, o método compara o valor de *fitness* da solução candidata com o valor de outras soluções e seleciona a rede com maior valor de *fitness*, em casos de maximização (cobertura máxima), e com menor valor, em casos de minimização (cobertura de conjunto). Este processo é iterativo e retorna à etapa 2 gerando novas soluções até que um critério de parada seja alcançado.

### 3.3.5 PESO DAS CARACTERÍSTICAS DOS LOCAIS CANDIDATOS

O método proposto permite definir um peso para cada Camada de Informação Geográfica. Assim, seja  $n_i \in [0,1]$ , o peso da Camada de Informação Geográfica, onde  $i = 0, 1, \dots, n$  indica a  $i$ -ésima camada e  $a_{ji}$  o valor no local do candidato  $i$  na Camada de Informação  $j$ , onde  $a_{ji} \in [0,1]$ . O peso ponderado de cada local candidato,  $p_i$ , é dado pela equação 48.

Seja

$$p_i = \sum_{j=0}^n n_j * a_{ji} \quad (48)$$

Onde

$$0 \leq p_i \leq 1 \quad (49)$$

Este peso é utilizado para definir a influência de determinada característica na simulação da rede de estações. Determinados experimentalmente os pesos das características físicas, não há necessidade de alteração do peso antes que ocorra uma mudança ambiental, como a construção de novas estradas, crescimento de centros urbanos, entre outros. Já os pesos das características estratégicas podem ser ajustados com mais frequência de acordo com os objetivos da simulação. Redes de estações com foco em monitoramento de desastres naturais devem associar um peso proporcionalmente maior à proximidade de área de risco do que de área de agricultura, por exemplo. A determinação dos pesos neste trabalho foi realizado de forma experimental e validado com apoio de profissionais da área de meteorologia.

### 3.4 PLANEJAMENTO DOS EXPERIMENTOS

Os experimentos descritos nesta seção envolvem cenários comuns ao planejamento de implantação de uma rede de estações meteorológicas, tais como:

- Cenário 1, plano de implantação de uma rede ideal, desconsiderando a existência de rede de estações previamente instalada;
- Cenário 2, plano de implantação de uma rede real, considerando a existência de uma rede de estações previamente instalada;

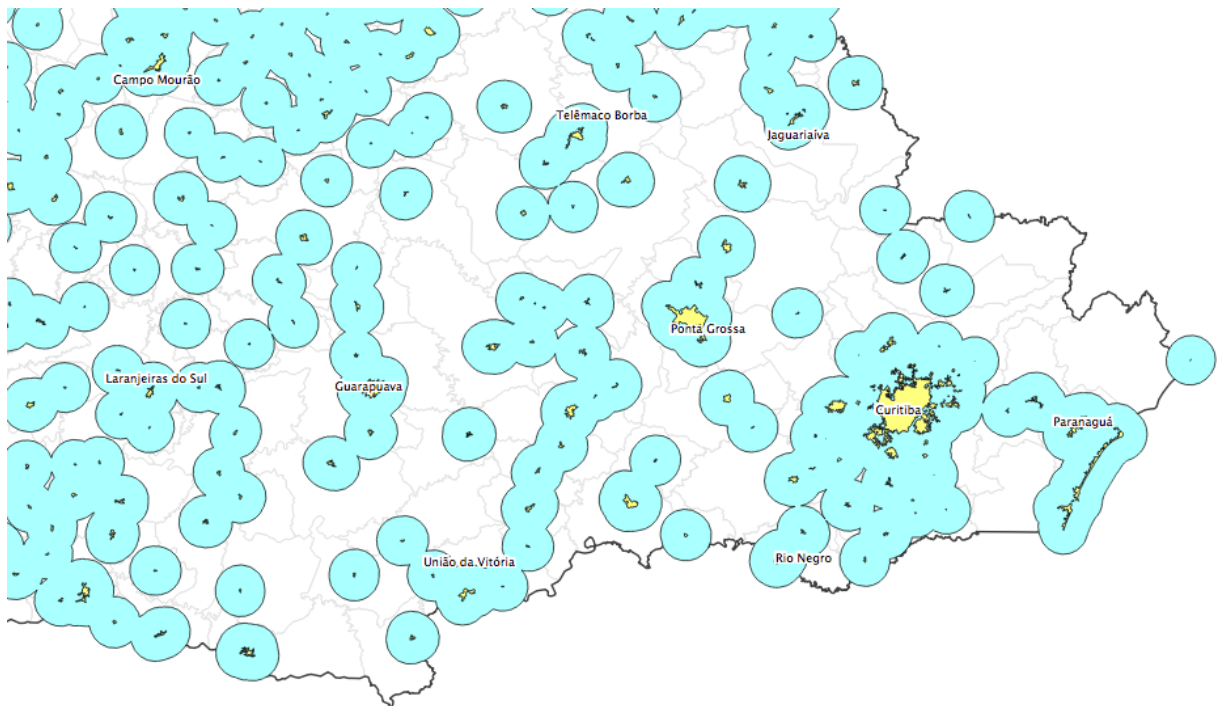
- Cenário 3, plano de expansão de uma rede ideal, por etapas, considerando uma capacidade limitada de instalação de estações por ano.

Os parâmetros de entrada utilizados nos experimentos são:

- Modelo Numérico de Previsão do Tempo: modelo MBAR com variável de temperatura média coletada no dia 06/08/2013 (a data do modelo foi escolhida aleatoriamente)
- Espaço: limites políticos do Estado do Paraná
- Camadas de Informação Geográfica: áreas urbanas, rodovias e altitude.

A partir das camadas de informação fornecidas, utilizou-se operações de geoprocessamento para criar um *buffer* em torno do eixo das rodovias com 500 metros para cada lado do eixo. O mesmo foi feito em relação as áreas urbanas, criando um *buffer* de cinco quilômetros.

A Figura 11 apresenta o mapa das áreas urbanas utilizadas em cor amarela e o *buffer* de cinco quilômetros em cor azul.



**Figura 11: Áreas urbanas em cor amarela e o buffer de cinco quilômetros em cor azul.**  
**Fonte: O Autor (2013).**

Nos experimentos realizados, a área urbana e o *buffer* são considerados uma única camada de informação geográfica. O peso desta camada de informação foi definido

experimentalmente como 1, e os valores atribuídos aos locais candidatos foram definidos da seguinte forma: 0,7 quando o local candidato está situado dentro da área urbana, 0,3 quando o local candidato está situado dentro da área do *buffer* da área urbana e 0 quando não estiver em nenhuma das situações anteriores.

A segunda camada de informação utilizada nos experimentos é a de rodovias. É uma camada vetorial, formada pelos eixos principais das rodovias internas e que cruzam o Estado do Paraná. Uma operação de *buffer* também foi realizada nesta camada, utilizando um raio de 500 metros. A Figura 12 mostra a camada após a operação de *buffer*.



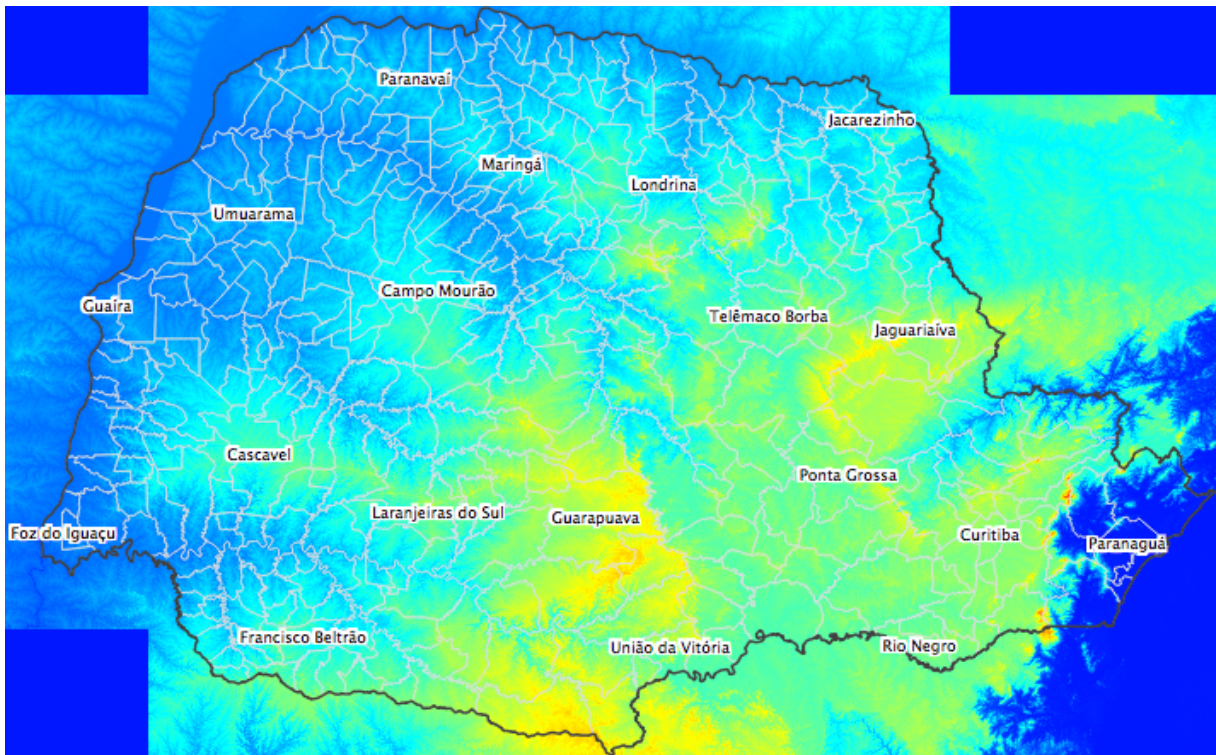
**Figura 12: Buffer de 500 metros em torno do eixo das rodovias em cor cinza.**  
**Fonte: O Autor (2013).**

A terceira camada de informação utilizada nos experimentos é a de altimetria. É uma camada *raster* cuja origem é o projeto SRTM (SRTM, 2013).

A Figura 13 mostra um mapa com a camada de informação geográfica do SRTM. Esta camada fornece a informação de altitude com uma resolução vertical de 90 metros.

### 3.4.1 DEFINIÇÃO DAS RESTRIÇÕES DA MODELAGEM MATEMÁTICA

Tanto no problema de cobertura de conjuntos quanto no modelo de cobertura máxima, define-se um conjunto  $N_j$  dos pontos de demanda que podem ser atendidos pelo local candidato  $x_j$ . Ser atendido, no contexto destes problemas, significa também ser coberto e depende do raio



**Figura 13: Mapa de altimetria baseado nos dados do projeto SRTM. Fonte: O Autor (2013).**

de cobertura utilizado no problema. Nos experimentos deste trabalho utilizou-se um raio de cobertura de 50 quilômetros, definido experimentalmente, com base na representatividade de um ponto de coleta de dados da variável temperatura média. Ou seja, utiliza-se a hipótese que um dado coletado de temperatura pode representar uma área de até 50 quilômetros em torno do ponto de coleta.

Entretanto, locais com um razoável desnível dentro deste raio podem fazer com que o dado coletado não represente a região como um todo. Um exemplo desta situação é a diferença de altitude entre o município de São José dos Pinhais e Morretes, cujas sedes estão distantes aproximadamente 40 quilômetros, mas possuem uma diferença de altitude de aproximadamente 700 metros. Para evitar que os algoritmos do grupo I, geradores de soluções candidatas, desconsiderem estas situações, adiciona-se uma segunda restrição para cada local candidato, cuja altitude é representada por  $Z_j$ , indicando quais os locais candidatos do conjunto  $N_j$ , cuja altitude  $z_{ij}$ , atende a condição  $Z_j - \delta_z \leq z_{ij} \leq Z_j + \delta_z$ , onde  $\delta_z$  é a distância vertical. Utilizou-se nos experimentos a distância vertical de 100 metros, definida experimentalmente.

### 3.4.2 CENÁRIO 1

O planejamento da implantação de uma rede ideal não considera a existência de uma rede previamente existente. Portanto, não existem estações instaladas que precisam ser

respeitadas na modelagem matemática. Além disto, não existe um limite máximo na quantidade de instalações.

O objetivo neste cenário é determinar a quantidade mínima de estações necessária para atender todos os pontos de demanda e a localização de cada estação.

O modelo matemático utilizado neste cenário é o modelo de cobertura de conjuntos e é definido na forma de um modelo de Programação Linear Inteira ou também conhecido por Programação Inteira.

Minimizar

$$\sum_{i=1}^m x_i p_i \quad (56)$$

sujeito a

$$p_i = 1 - \sum_{j=1}^n \frac{l_j c_{ij}}{n} \quad (57)$$

$$l_j \in [0, 1] \quad (58)$$

$$c_{ij} \in [0, 1] \quad (59)$$

$$\sum_{j \in N_i} x_j \geq 1, \forall i \in I \quad (60)$$

$$x_j \in 0, 1, \forall j \in J \quad (61)$$

#### 3.4.2.1 SOLUÇÃO DO PROBLEMA DE COBERTURA DE CONJUNTOS POR MÉTODO EXATO

Neste experimento avalia-se a capacidade computacional disponível para encontrar a solução ótima do problema de localização de estações. Utiliza-se o modelo de cobertura de conjuntos, que é do tipo *NP – Hard* (GAREY; JOHNSON, 1979). Ao determinar os limites desta capacidade, pretende-se analisar a necessidade ou não do uso de métodos não exatos para obtenção de soluções próximas da ótima.

Para execução deste experimento, são utilizadas quantidades crescentes de locais candidatos (25, 50, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1100, 1200, 1300, 1400, 1500, 1600, 1700 e 1781) e mede-se o tempo médio necessário para encontrar a solução ótima, o valor de *fitness* calculado e a quantidade de estações utilizada.

Utiliza-se o programa *lp\_solve* (BERKELAAR et al., 2004) para resolução do modelo. O algoritmo 4, disponível no apêndice, apresenta a modelagem matemática no formato do programa *lp\_solve*.

A Tabela 3 apresenta as características deste experimento.

**Tabela 3: Experimento - Objetivo e características**

Característica	Descrição
Objetivo	Avaliar a capacidade computacional disponível para encontrar a solução ótima do problema de localização de estações
Modelo aplicado	Cobertura de conjunto
Objetivo	Minimização
Método de resolução	<i>Branch &amp; Bound</i> com programa <i>lp_solve</i>
Número de estações	Indeterminado
Resultado do experimento	Locais candidatos selecionados para instalação, valor da função objetivo e quantidade de estações

Fonte: O Autor (2013).

### 3.4.2.2 SOLUÇÃO PROBLEMA DE COBERTURA DE CONJUNTOS POR EVOLUÇÃO DIFERENCIAL

Neste experimento avalia-se a capacidade da estratégia evolutiva Evolução Diferencial para encontrar a solução ótima do problema de localização de estações utilizando o modelo de cobertura de conjuntos.

Para execução deste experimento, são utilizadas quantidades crescentes de locais candidatos e mede-se o tempo médio necessário para encontrar a solução ótima, o valor de *fitness* calculado e a quantidade de estações.

A Tabela 4 apresenta as características deste experimento.

**Tabela 4: Experimento - Objetivo e características**

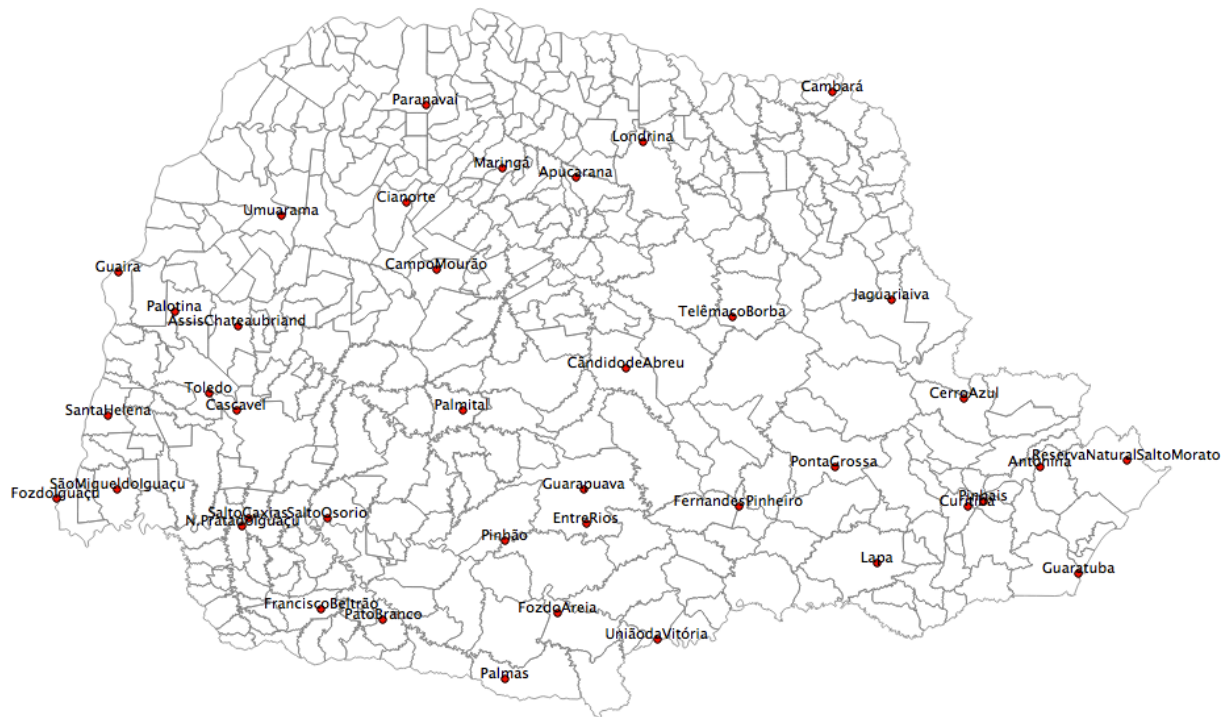
Característica	Descrição
Objetivo	Avaliar a capacidade da estratégia evolutiva Evolução Diferencial para encontrar a solução ótima do problema de localização de estações utilizando o modelo de cobertura de conjuntos
Modelo aplicado	Cobertura de conjunto
Tipo de otimização	Minimização
Método de resolução	Evolução Diferencial
Número de estações	Indefinido
Resultado do experimento	Locais candidatos selecionados para instalação e valor da função objetivo

Fonte: O Autor (2013).

### 3.4.3 CENÁRIO 2

Neste experimento considera-se a existência de estações instaladas e deseja-se determinar a estrutura da rede ideal com cobertura sobre todos os locais de demanda. As estações do SIMEPAR foram consideradas como as estações já existentes. A Tabela 10, disponível no apêndice, apresenta a lista das 40 estações.

A Figura 14 apresenta um mapa com a localização das estações do SIMEPAR no Estado do Paraná.



**Figura 14: Localização das estações do SIMEPAR no Estado do Paraná. Fonte: O Autor.**

O modelo matemático utilizado neste cenário é similar ao primeiro experimento do cenário 1, mas possui uma restrição adicional, a equação 75, que força a seleção das estações do SIMEPAR como parte da solução.

Minimizar

$$\sum_{i=1}^m x_i p_i \quad (69)$$



sujeito a

$$p_i = 1 - \sum_{j=1}^n \frac{l_j c_{ij}}{n} \quad (70)$$

$$l_j \in [0, 1] \quad (71)$$

$$c_{ij} \in [0, 1] \quad (72)$$

$$\sum_{j \in N_i} x_j \geq 1, \forall i \in I \quad (73)$$

$$x_j \in 0, 1, \forall j \in J \quad (74)$$

$$\sum_{w=0}^{40} x_w = 40 \quad (75)$$

As Tabelas 5 e 6 apresentam as características dos experimentos deste cenário.

### 3.4.3.1 REDE IDEAL COM A EXISTÊNCIA PRÉVIA DE ESTAÇÕES

Este experimento envolve o planejamento da implantação e expansão de uma rede de estações. Considera-se neste experimento a existência de limite na capacidade de instalações de estações, por exemplo, 20 estações por ano.

**Tabela 5: Experimento - Objetivo e características**

Característica	Descrição
Objetivo	Determinar os locais de instalação e o quantidade mínima de estações necessárias para cobrir todos os locais de demanda considerando as localização das estação existentes.
Modelo aplicado	Cobertura de conjunto
Tipo de otimização	Minimização
Número de locais candidatos	1781
Número de estações	Fixo, a ser determinado durante execução dos experimentos
Número de estações existentes	40
Resultado do experimento	Lista de locais candidatos selecionados para instalação das novas estações

Fonte: O Autor (2013).

### 3.4.3.2 REDE REAL COM A EXISTÊNCIA PRÉVIA DE ESTAÇÕES

Este experimento envolve o planejamento da implantação e expansão de uma rede de estações. Considera-se neste experimento a existência de limite na capacidade de instalações de estações, por exemplo, 20 estações por ano.

**Tabela 6: Experimento - Objetivo e características**

Característica	Descrição
Objetivo	Determinar os locais de instalação e o quantidade mínima de estações necessárias para cobrir todos os locais de demanda considerando as localização das estação existentes.
Modelo aplicado	Cobertura de conjunto
Tipo de otimização	Minimização
Método de resolução	Evolução Diferencial
Número de locais candidatos	1781
Número de estações	Fixo, a ser determinado durante execução dos experimentos
Número de estações existentes	40
Resultado do experimento	Lista de locais candidatos selecionados para instalação das novas estações

Fonte: O Autor (2013).

### 3.4.4 CENÁRIO 3

No cenário 3, deseja-se implantar uma rede de estações por etapas. A capacidade de instalar estações no período de um ano é limitada a 20, um valor definido *a priori*. Neste ambiente, é interessante planejar a instalação de forma que a primeira etapa cubra a maior área possível, respeitando as restrições impostas. Da mesma forma, nas etapas seguintes, é desejável maximizar a área de cobertura de cada novo conjunto de instalações. Este tipo de problema se caracteriza como um problema de cobertura máxima.

#### 3.4.4.1 PLANO DE CRESCIMENTO SEM A EXISTÊNCIA PRÉVIA DE ESTAÇÕES

Neste experimento não existem estações instaladas previamente e deseja-se determinar a sequência de instalação de um determinado número de estações de forma a maximizar a quantidade de demanda coberta a cada instalação. A Tabela 7 apresenta as características deste cenário.

**Tabela 7: Experimento - Objetivo e características**

Característica	Descrição
Objetivo	Determinar o local de instalação das estações conforme um plano de expansão para 5 anos. Considera-se que a cada ano são instaladas 20 novas estações. Deseja-se obter a melhor sequência de instalações de forma a maximizar a cobertura da demanda existente.
Modelo aplicado	Cobertura de conjunto
Tipo de otimização	Maximização
Número de locais candidatos	1781
Número de estações	10 novas estações por ano
Número de estações existentes	40
Localização das estações existentes	Tabela 10
Número de estações	20 estações por ano
Camadas de informação	Altimetria. Proximidade de estradas ou vias de acesso. Proximidade de área urbanas.
Resultado do experimento	Lista de locais candidatos selecionados para instalação das novas estações em grupos de 20 ordenados por ano de instalação

Fonte: O Autor (2013).

## 4 RESULTADOS

Neste capítulo são apresentados os resultados obtidos nos experimentos planejados no capítulo 3. Devido a limitação de tempo na realização deste trabalho, optou-se por não implementar as etapas 3 e 4 do método proposto.

### 4.1 CENÁRIO 1

#### 4.1.1 SOLUÇÃO DO PROBLEMA DE COBERTURA DE CONJUNTOS POR MÉTODO EXATO

Este experimento prevê a busca da solução do problema de localização de estações com o modelo do problema de cobertura de conjuntos, utilizando a técnica de resolução de problema de programação inteira via método *Branch & Bound* para a solução exata.

Na Tabela 8 são mostrados os resultados do experimento para os diferentes tamanhos de instância. Observa-se um crescimento no tempo de execução à medida que a quantidade de candidatos é incrementada. Verifica-se também que o crescimento do valor da função objetivo não é proporcional ao crescimento do tempo de execução.

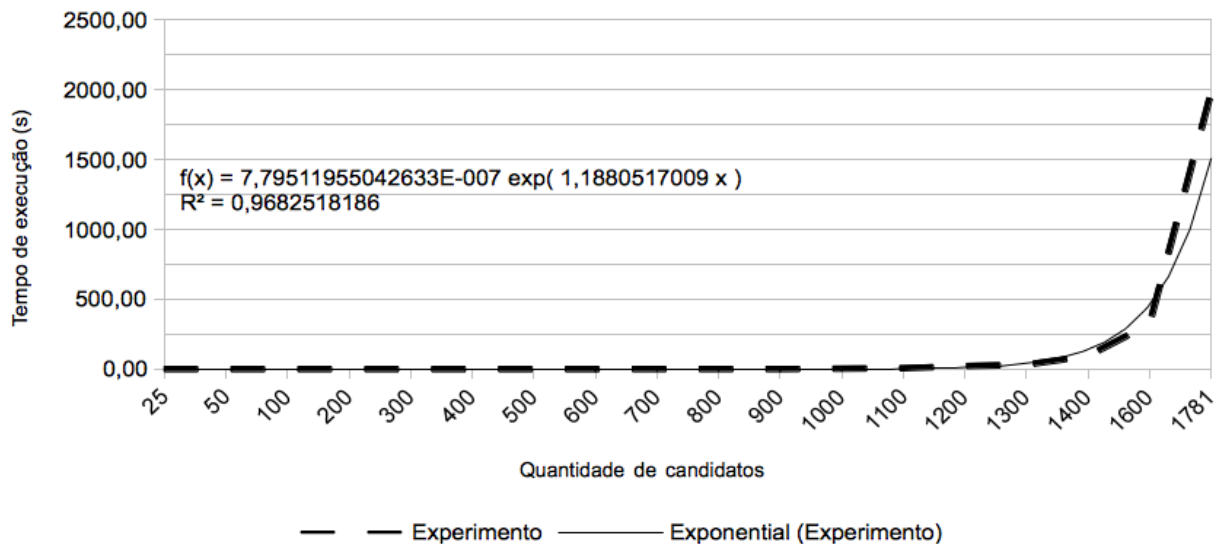
A Figura 15 apresenta o comportamento do tempo médio de execução em relação à quantidade de candidatos. É possível identificar um comportamento similar ao de uma função exponencial. Este comportamento não é totalmente inesperado, visto que a revisão da literatura indica que modelo de cobertura de conjuntos é da classe de problemas *NP-Hard*. Esta característica do modelo de cobertura de conjunto reforça a necessidade do uso de heurísticas, como descrito no método proposto. A Figura 16 apresenta a quantidade de estações e valor da função objetivo em relação à quantidade de candidatos.

A Figura 17 apresenta o espaço do experimentos, o Estado do Paraná, com a solução exata para 1781 locais candidatos com raio de cobertura de 50km e distância vertical de 100m. Estão representados os locais candidatos e os locais selecionados para instalação. Observa-se que a solução ótima mantém uma cobertura sobre todos os pontos de demanda e que

**Tabela 8: Resultados do experimento**

Quantidade de candidatos	Tempo de execução (min)	Valor da função objetivo	Quantidade de estações
25	0,00	0,00	2
50	0,00	1,05	5
100	0,00	2,90	10
200	0,00	5,95	17
300	0,00	8,65	23
400	0,00	11,45	32
500	0,00	12,35	34
600	0,00	14,40	39
700	0,05	17,00	46
800	0,07	19,05	50
900	0,12	19,95	54
1000	3,43	22,50	59
1100	6,57	23,90	64
1200	22,32	24,60	68
1300	31,92	30,00	79
1400	94,27	31,55	86
1500	115,28	33,30	92
1600	320,88	35,00	97
1700	767,13	35,80	99
1781	1.982,30	36,85	102

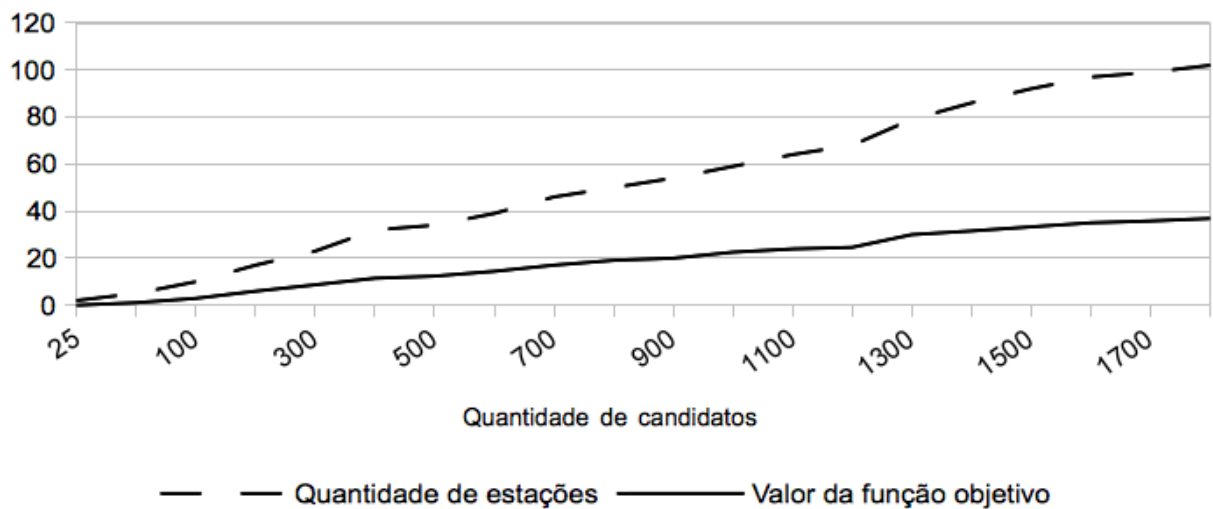
Fonte: O Autor (2013).



**Figura 15: Tempo de execução em relação à quantidade de candidatos. Problema da cobertura de conjuntos com solução por método exato.**

Fonte: O Autor (2013).

visualmente os locais de instalação estão dispostos de forma a minimizar a cobertura duplicada de pontos de demanda.



**Figura 16:** Valor da função objetivo e quantidade de estações em relação à quantidade de candidatos. Problema da cobertura de conjuntos com solução por método exato.

Fonte: O Autor (2013).

Desta maneira, a opção por utilizar o modelo de cobertura de conjuntos parece ser adequada haja vista a solução apresentada. Entretanto, conforme discutido anteriormente, o uso deste modelo pode tornar necessário a aplicação de métodos não exatos. Os próximos experimentos tratam desta situação, onde o método exato é comparado com heurísticas.

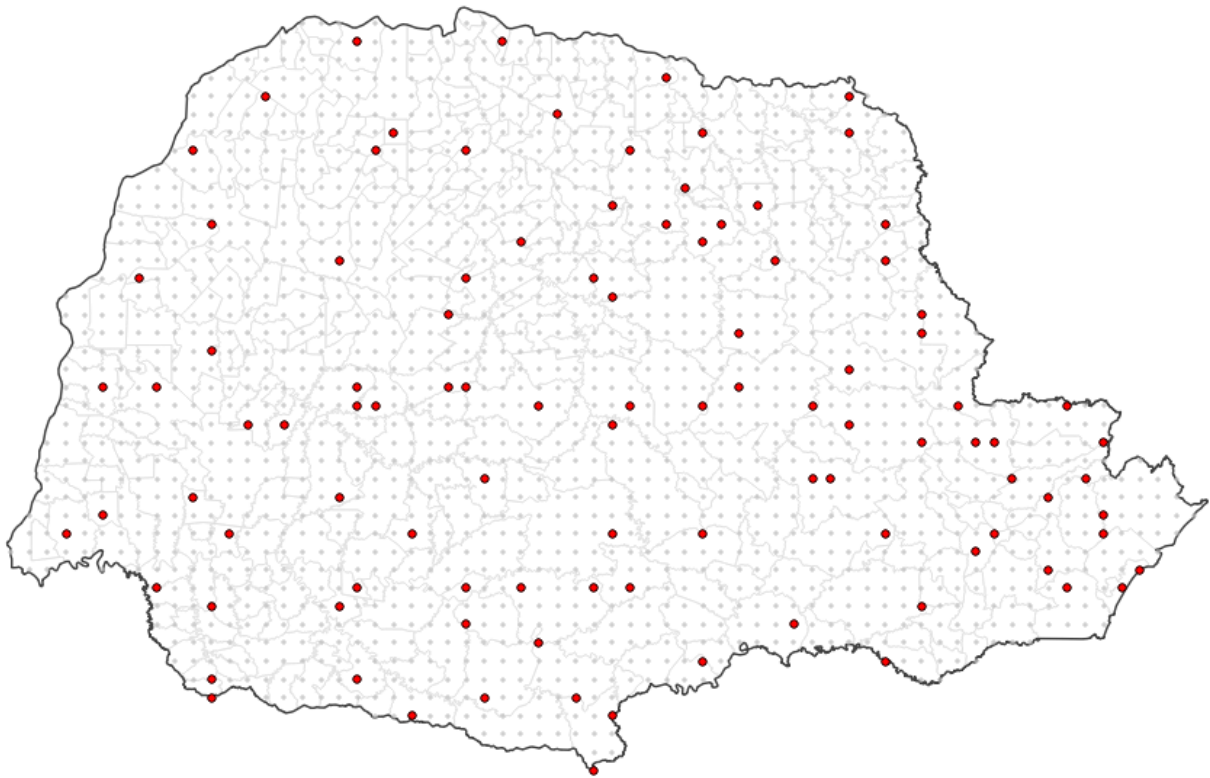
A importância do experimento está em determinar o número mínimo de estações necessário para cobertura de todos os pontos de demanda considerando pesos diferenciados entre eles e camadas de informação estratégicas que possam influenciar o resultado.

A Figura 18 apresenta o resultado do primeiro experimento do cenário 1 para a instância com 1781 candidatos em conjunto com as camadas de informação de rodovias e áreas urbanas. É possível observar que a modelagem matemática utilizada foi adequada, pois poucos locais de instalação foram selecionados fora da proximidade das áreas urbanas e rodovias.

#### 4.1.2 SOLUÇÃO PROBLEMA DE COBERTURA DE CONJUNTOS POR EVOLUÇÃO DIFERENCIAL

Neste experimento utilizou-se a estratégia evolucionária Evolução Diferencial para resolver o problema de cobertura de conjuntos. As instâncias utilizadas neste experimento são as mesmas do primeiro experimento do cenário 1, pois o objetivo é avaliar a capacidade da evolução diferencial em encontrar a solução ótima ou uma solução quase ótima.

Para cada tamanho de instância, que significa a quantidade de locais candidatos a instalação de uma estação, executou-se a simulação 10 vezes.



**Figura 17: Localização dos candidatos selecionados para instância com 1781 candidatos.**

**Fonte: O Autor (2013).**

Na Tabela 9 estão disponíveis os resultados do experimento para as diferentes quantidades de candidatos.

Em comparação com o tempo de execução da solução exata, a evolução diferencial obteve tempos de execução superior até a instância com 1.600 candidatos. Para as instâncias com 1700 e 1781 candidatos, a evolução diferencial obteve tempos significativamente inferiores. A Figura 19 apresenta o tempo de execução dos dois métodos.

Neste experimento, a evolução diferencial foi capaz de encontrar a solução ótima para as instâncias com 25, 50, 100, 200, 300 e 400 candidatos. A partir da instância com 500 candidatos, a evolução diferencial apresenta soluções aproximadas do ótimo conhecido, se afastando à medida que a quantidade de candidatos aumenta. Este comportamento é esperado em heurísticas devido ao aumento exponencial no espaço de busca. Entretanto, observa-se na Figura 20 que este afastamento aparentemente assume um comportamento linear, permitindo avaliar o quão longe o resultado está da solução exata. No caso deste experimento, verifica-se que o resultado da evolução diferencial chega ser quase 25% maior que a solução exata.

A análise do comportamento da curva do valor de *fitness* normalizado, apresentado na

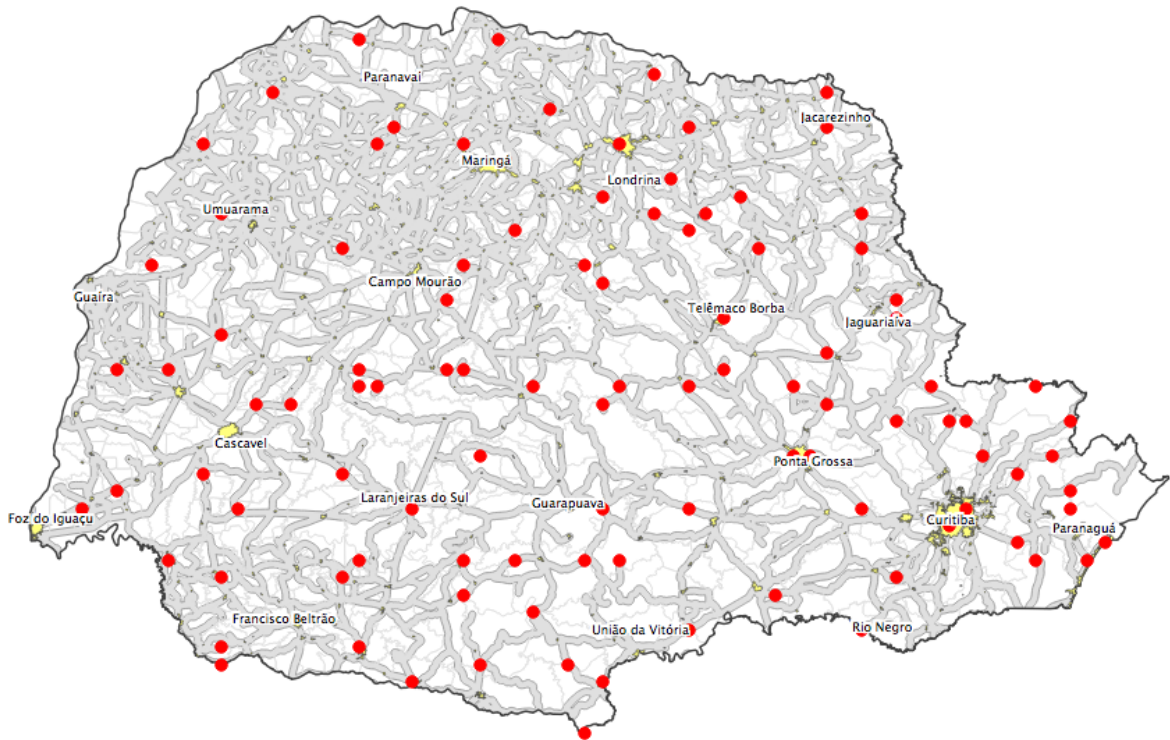


Figura 18: Localização da solução para 1781 candidatos em relação as rodovias e área urbanas.

Fonte: O Autor (2013).

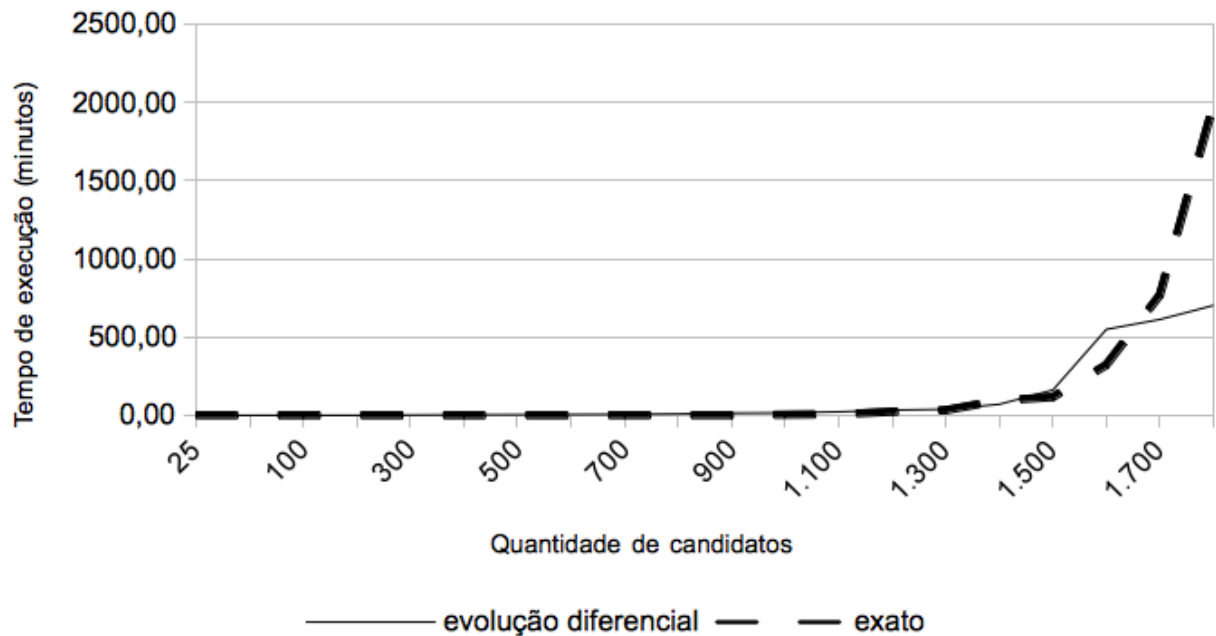


Figura 19: Tempo de execução dos métodos exato e evolução diferencial para instância com 1781 locais candidatos.

Fonte: O Autor (2013).



**Tabela 9: Resultados do experimento**

CAND	POP	ITER	FITN	ROD	TEMP
25	120	6.000	0,00 ± 0,00	11,25 ± 2,28	0,00 ± 0,00
50	120	6.000	1,05 ± 0,00	31,50 ± 0,50	4,50 ± 0,00
100	120	6.000	3,40 ± 0,00	637,00 ± 641,15	10,00 ± 0,00
200	120	6.000	6,54 ± 0,15	3.635,00 ± 1.760,92	23,00 ± 0,00
300	240	12.000	9,17 ± 0,02	1.334,50 ± 272,51	102,75 ± 0,43
400	240	12.000	12,11 ± 0,22	3.091,50 ± 1.101,44	186,75 ± 0,43
500	240	12.000	14,32 ± 0,24	3.310,00 ± 1.236,23	242,00 ± 0,51
600	240	12.000	15,40 ± 0,20	6.586,75 ± 2.719,69	297,75 ± 0,43
700	240	12.000	18,52 ± 0,32	8.876,50 ± 1.460,50	361,50 ± 0,50
800	240	18.000	21,60 ± 0,10	13.626,50 ± 101,50	521,00 ± 0,50
900	240	21.000	23,75 ± 0,35	20.035,50 ± 439,50	829,50 ± 0,50
1000	240	60.000	27,31 ± 0,78	45.585,50 ± 1.255,18	967,00 ± 0,43
1100	240	120.000	29,05 ± 1,15	72.789,00 ± 1.157,20	1.345,50 ± 0,80
1200	240	120.000	30,15 ± 0,69	105.637,00 ± 2.378,00	1.978,40 ± 0,93
1300	240	180.000	35,90 ± 0,80	131.230,00 ± 1.763,00	2.312,50 ± 1,50
1400	240	180.000	39,75 ± 0,57	157.676,00 ± 2.389,75	4.217,75 ± 2,30
1500	240	240.000	42,95 ± 0,37	168.233,50 ± 1.856,00	9.675,50 ± 4,10
1600	240	480.000	51,00 ± 0,91	453.264,00 ± 1.562,80	32.944,00 ± 6,80
1700	240	480.000	50,10 ± 1,02	476.281,00 ± 2.756,10	36.737,00 ± 9,30
1781	240	480.000	50,50 ± 0,87	478.015,15 ± 1.389,00	19.237,00 ± 11,67

CAND: quantidade de candidatos

POP: tamanho da população

ITER: número máximo de iterações

FITN: valor médio de *fitness*

ROD: número médio, em segundos, da iteração na qual a melhor solução é descoberta

TEMP: tempo médio, em segundos, da execução do algoritmo de evolução diferencial

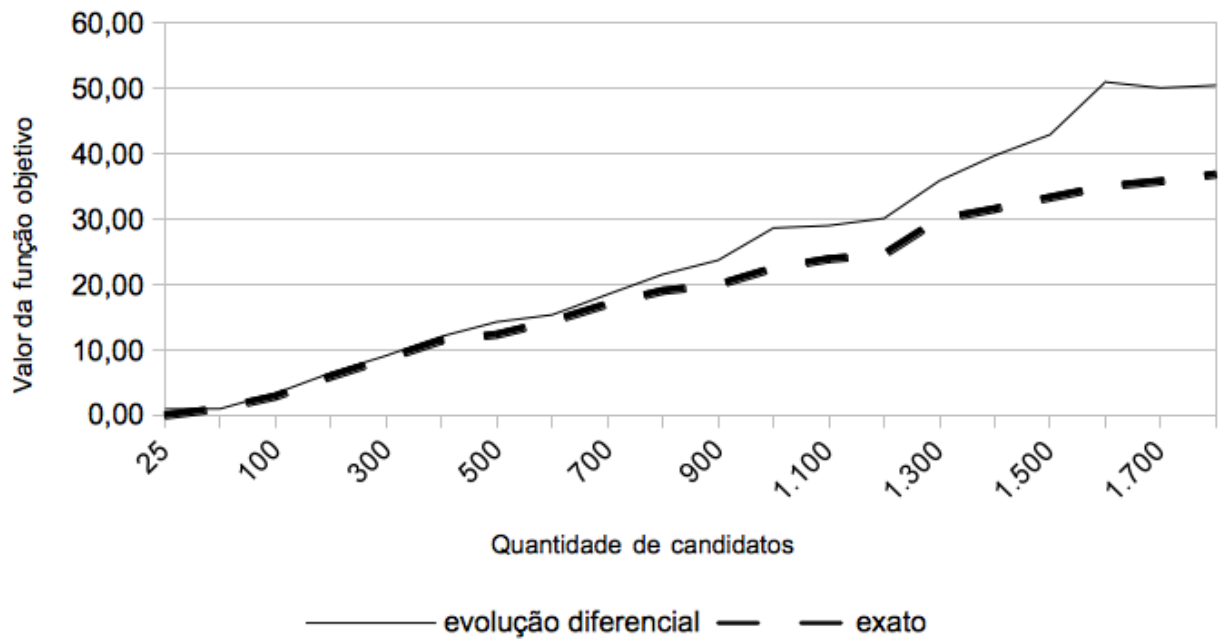
Fonte: O Autor (2013).

Figura 21, indica que ainda há possibilidade de melhora no valor de *fitness* apresentado. Como a diferença de tempo de execução entre os métodos nesta instância é da ordem de horas, é possível que para esta instância e outras maiores seja adequado aumentar ainda mais o número máximo de iterações utilizado como um dos critérios de parada.

## 4.2 CENÁRIO 2

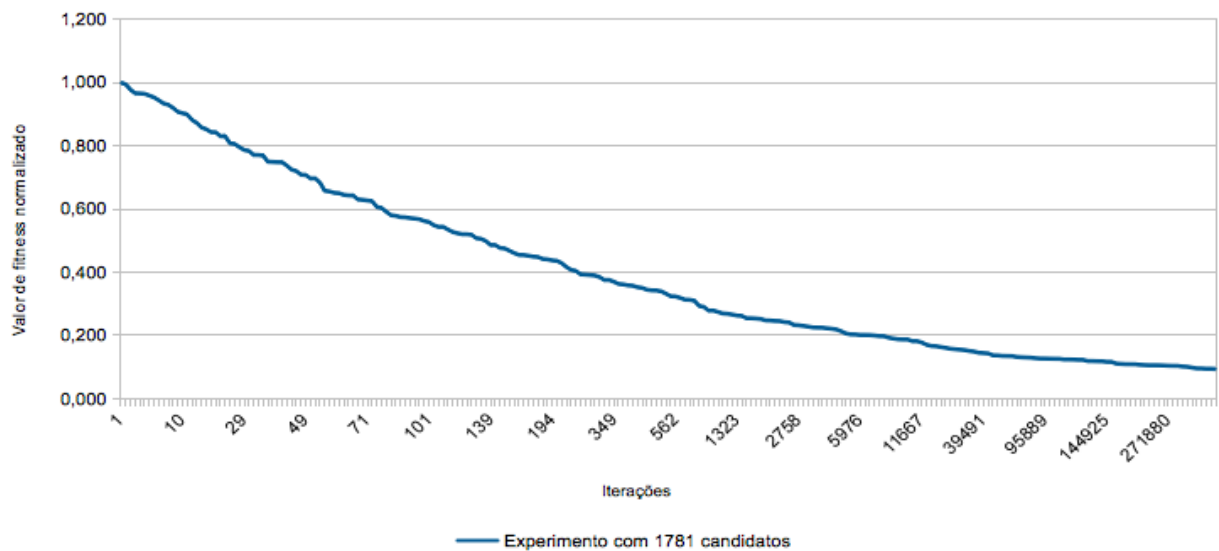
### 4.2.1 REDE IDEAL COM A EXISTÊNCIA PRÉVIA DE ESTAÇÕES

Este experimento produz um plano de implantação, composto por uma quantidade de estações necessárias para cobrir todos os pontos de demanda no Estado do Paraná, bem como a respectiva localização de cada ponto. O experimento faz uso da programação inteira para modelar matematicamente o problema, que se encaixa nos grupos dos problemas de cobertura



**Figura 20:** Valor da função objetivo dos métodos exato e evolução diferencial para instância com 1781 locais candidatos.

Fonte: O Autor (2013).



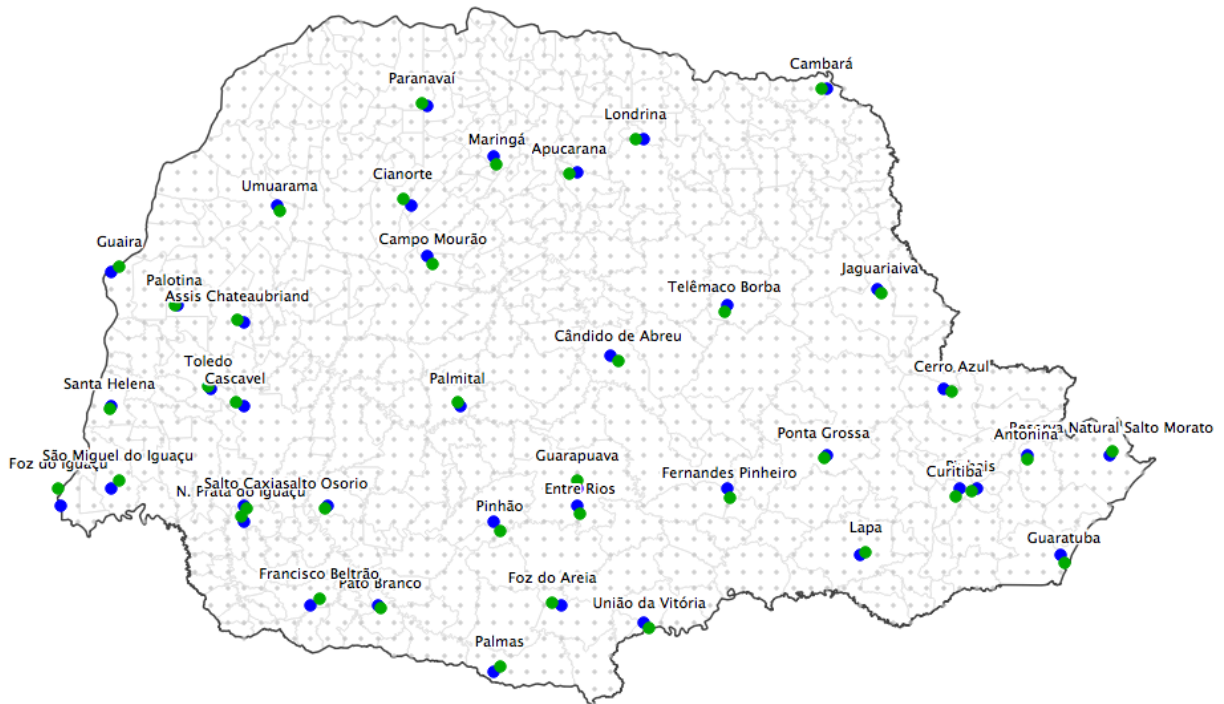
**Figura 21:** Valor de *fitness* normalizado dos métodos exato e evolução diferencial para instância com 1781 locais candidatos.

Fonte: O Autor (2013).

de conjunto.

O termo “rede real”, utilizado no título deste experimento, significa que se considera na modelagem do problema a existência de estações previamente instaladas, sob a forma de restrições, que exige que as estações façam parte da solução do problema. A localização

das estações existentes não coincide com a localização dos locais candidatos utilizados nos experimentos. Portanto, substitui-se cada uma das estações existentes pelo local candidato mais próximo. A Figura 22 apresenta as estações existentes, na forma de pontos verdes, e os respectivos locais candidatos mais próximos, na forma de pontos azuis. Utilizou-se para esta operação, a linguagem de programação *Python* (PYTHON, 2013) e as funcionalidades de geoprocessamento da biblioteca *Shapely* (SHAPELY, 2013).



**Figura 22:** Localização das estações do SIMEPAR e os respectivos locais candidatos mais próximos.

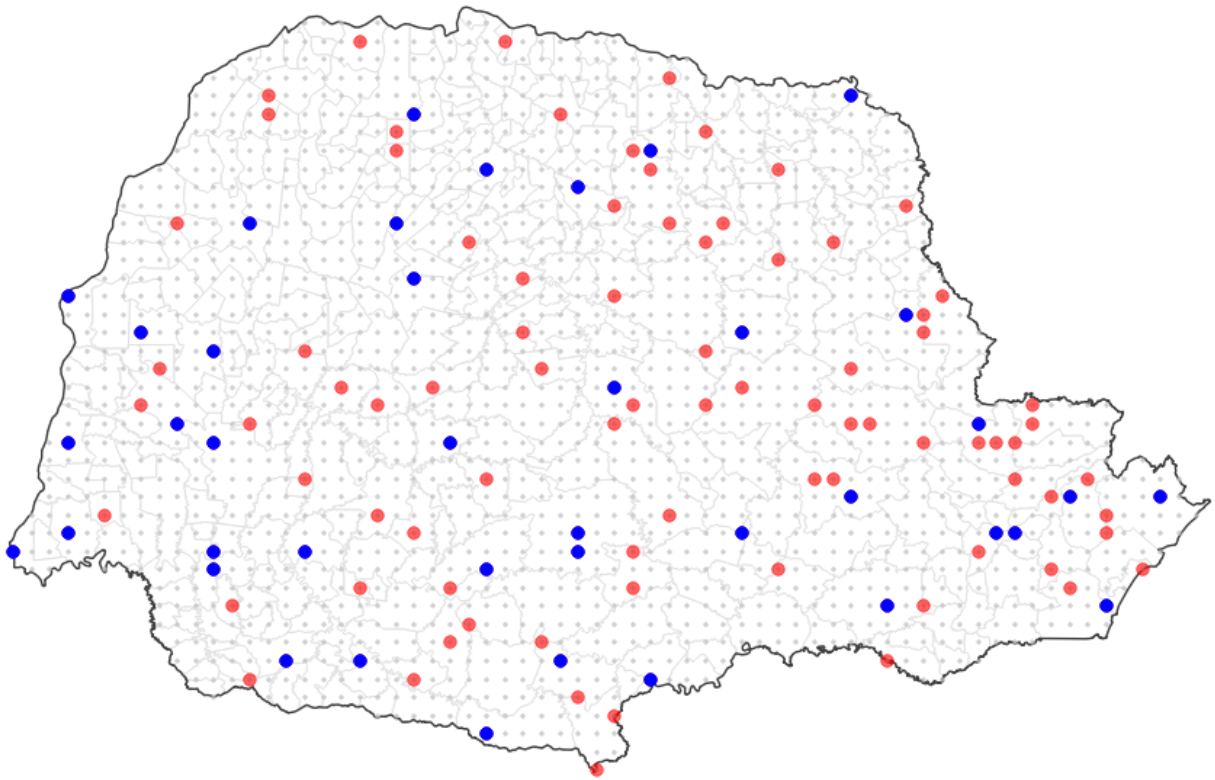
Fonte: O Autor (2013).

A Figura 23 apresenta o resultado do plano de implantação de uma rede ideal por método exato. Os pontos em cor azul, 40 ao total, representam as estações existentes em dezembro de 2012 e os pontos em vermelho, 62 ao total, representam as estações necessárias para garantir a cobertura de todos os pontos de demanda, representados por pequenos pontos em cor cinza e que somam 1781, no Estado do Paraná.

### 4.3 CENÁRIO 3

#### 4.3.1 PLANO DE CRESCIMENTO SEM A EXISTÊNCIA PRÉVIA DE ESTAÇÕES

Neste experimento é produzido um plano de expansão de uma rede ideal e já se conhece, *a priori*, quais os locais onde serão instaladas as novas estações, resultado do primeiro

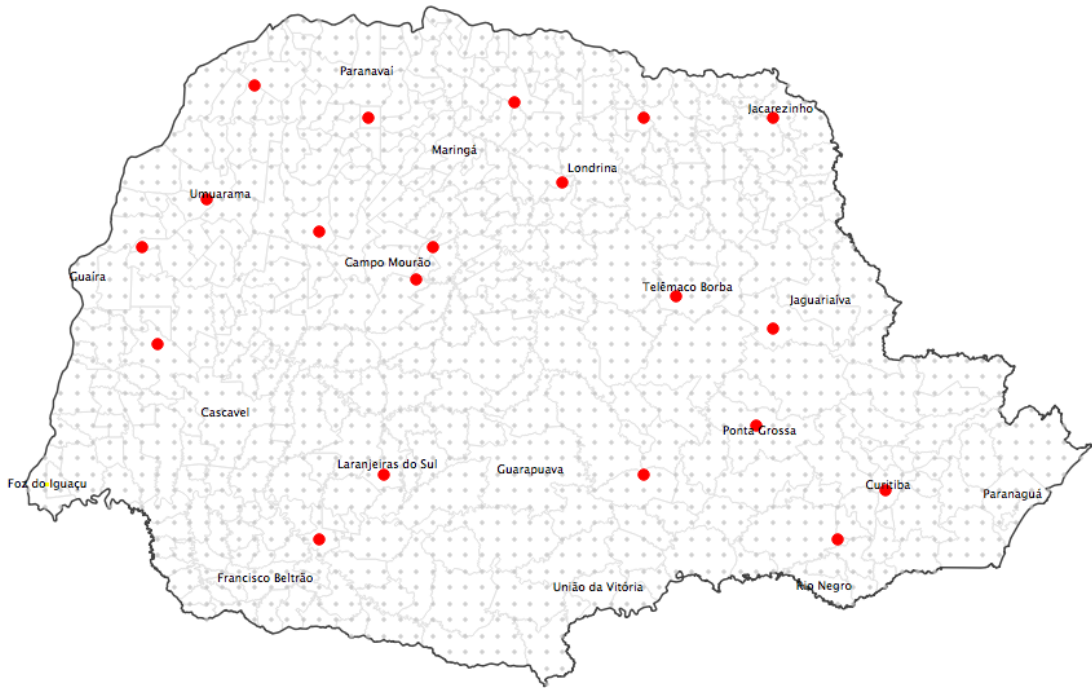


**Figura 23: Plano de implantação de rede real por método exato.**

**Fonte: O Autor (2013).**

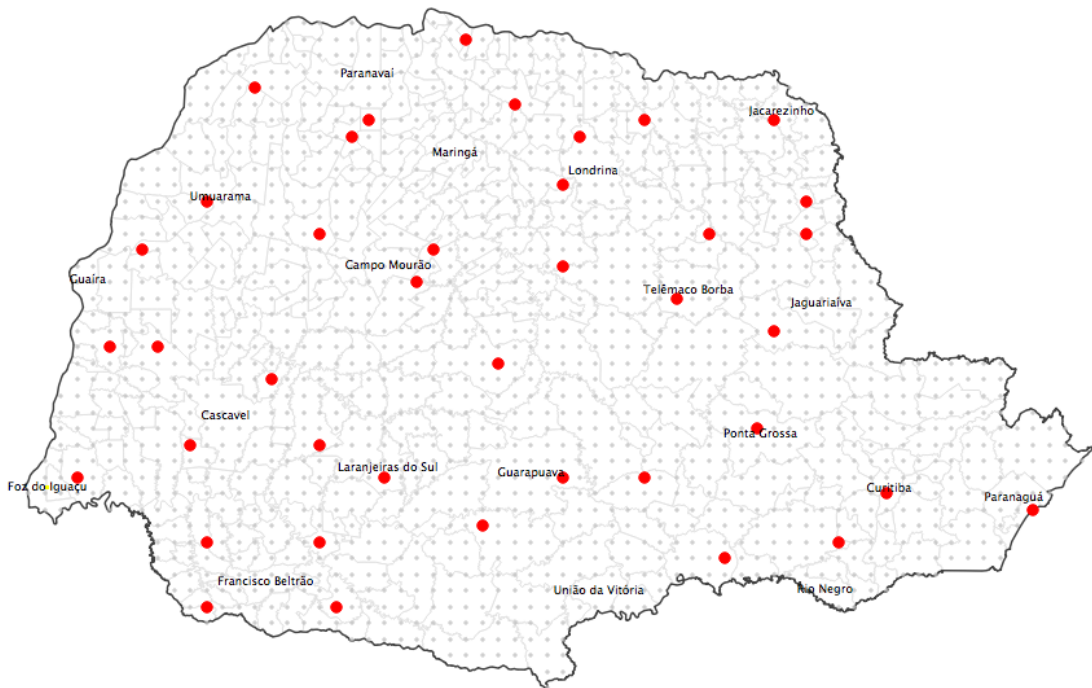
experimento do cenário 1, conforme apresenta a Figura 17. Portanto, deseja-se definir qual a melhor sequência de instalação das estações, visando maximizar o valor da função objetivo e em consequência, aumentar a representatividade do conjunto de estações selecionadas a cada iteração. Definiu-se também, que a cada iteração devem ser selecionadas 20 novas estações para serem instaladas, exceto a quinta rodada, em que são 22 novas estações.

As Figuras 24, 25, 26 e 27 apresentam os resultados para as iterações 1, 2, 3 e 4 respectivamente. Observa-se com o avanço das iterações um aumento na densidade de estações no espaço do experimento. Ainda, na Figura 27, apresenta-se em pontos pretos, as estações restantes, 22 ao total, que são instaladas na quinta iteração. Verifica-se também, que o aumento na quantidade de estações na região norte do litoral paranaense e na zona da Mata Atlântica ocorre somente nas últimas iterações, saindo de 2 estações na iteração 2 para 11 estações ao final da iteração 5.



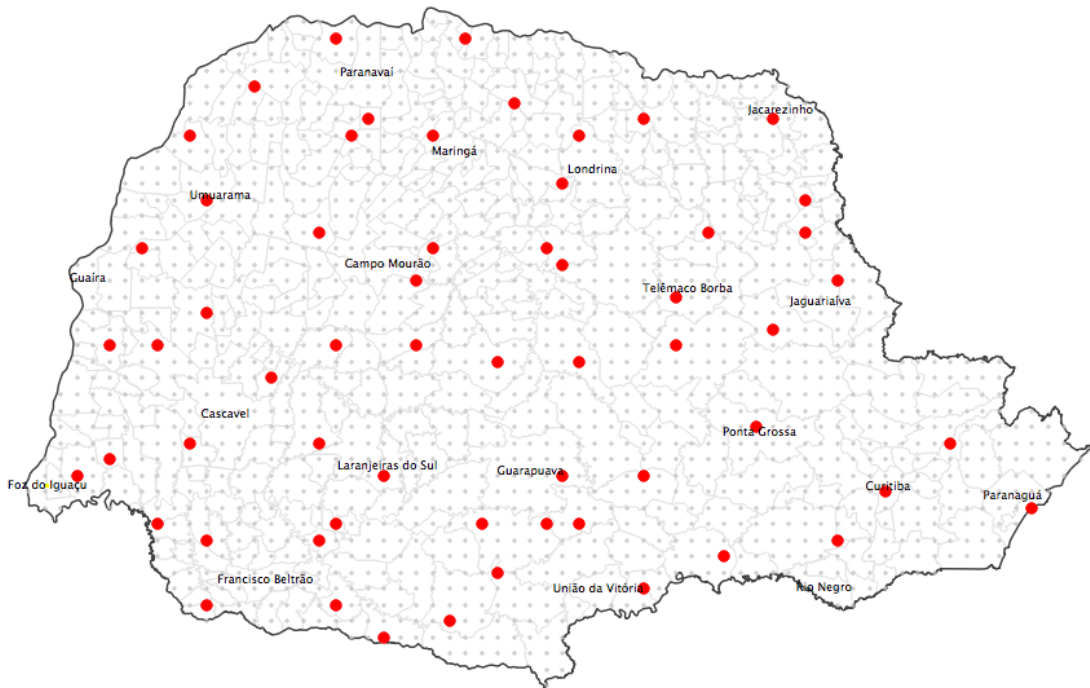
**Figura 24: Plano de expansão da rede ideal por método exato - iteração 1.**

**Fonte: O Autor (2013).**



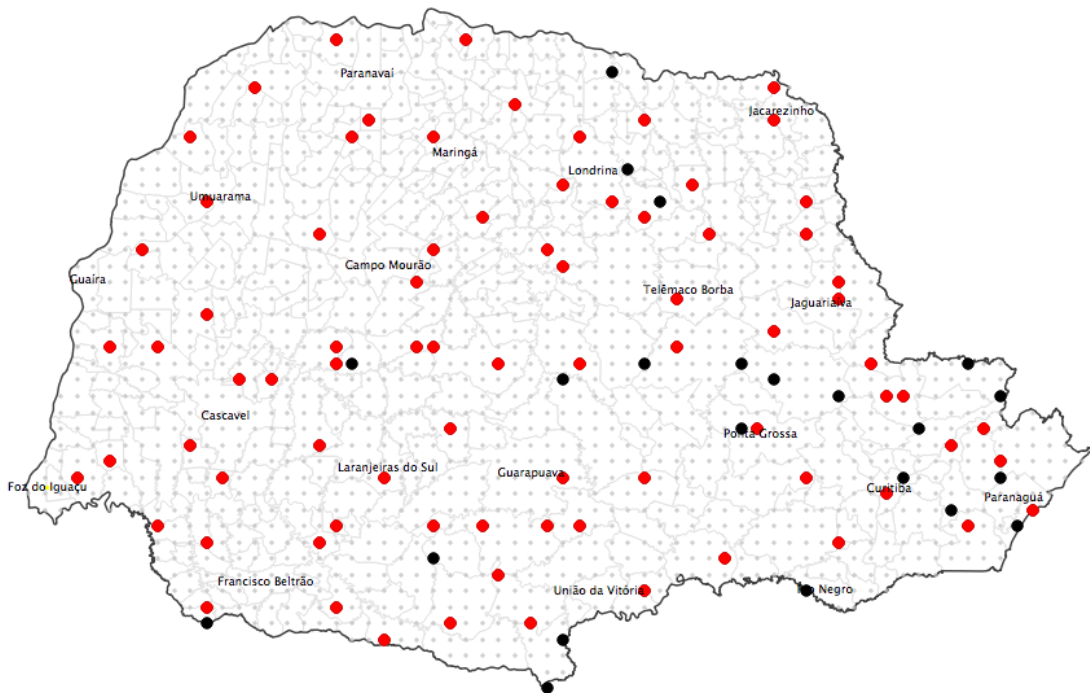
**Figura 25: Plano de expansão da rede ideal por método exato - iteração 2.**

**Fonte: O Autor (2013).**



**Figura 26: Plano de expansão da rede ideal por método exato - iteração 3.**

**Fonte: O Autor (2013).**



**Figura 27: Plano de expansão da rede ideal por método exato - iteração 4.**

**Fonte: O Autor (2013).**

## 5 CONCLUSÕES E TRABALHOS FUTUROS

A implantação ou expansão de redes de estações meteorológicas é um tema que tem atraído a atenção da administração pública. Eventos de desastres naturais ocorridos nos últimos anos têm destacado a necessidade de entender os eventos atmosféricos como forma de antecipar os impactos causados por chuvas e ventos, entre outros fenômenos.

Neste cenário, o uso do modelo de previsão do tempo é um importante aliado na simulação de condições da atmosfera no futuro próximo (de um a quinze dias). Apesar disto, tanto a previsão numérica do tempo quanto outros sistemas de previsão dependem fundamentalmente de observações coletadas para caracterizar o estado atual ou futuro da atmosfera.

Os atuais sistemas de observação utilizam estações meteorológicas, compostas por diversos sensores, para coletar os dados. O investimento no Brasil, no final de 2012, em uma estação, envolve valores aproximados de R\$ 30.000,00. Além do custo de aquisição ou montagem de uma estação, também devem ser considerados os custos com transmissão dos dados via telefonia celular ou outro meio. Assim, os administradores públicos responsáveis pela gestão da rede de estações meteorológicas precisam encontrar um equilíbrio entre a cobertura dos pontos de demanda no território e atender as limitações de custos e capacidade de instalação.

As questões apresentadas no parágrafo anterior ressaltam a importância de utilizar um método científico para planejamento da implantação de uma rede de estações. Este problema, de planejamento da instalação das estações, se encaixa em dois modelos discretos de localização: cobertura de conjunto e cobertura máximo. Este modelos são do tipo *NP-Hard*, o que implica no uso de heurísticas para grandes instâncias.

Este trabalho apresentou um método para o problema de localização de estações meteorológicas. O método apresenta como pontos positivos: a independência entre módulos que permitiu a avaliação com uso de dois diferentes algoritmos e os conceitos de camadas de informação para extração das características dos locais candidatos.

Os mapas apresentados nas Figuras 15, 16, 17, 18, 22 e 23 mostram que o método

proposto, em conjunto com o uso de modelos matemáticos, permitiu a descoberta de soluções adequadas e que respeitaram as restrições impostas. Em especial, verificou-se que a imposição de uma restrição referente à distância vertical entre estações próximas permitiu a distribuição das estações de forma que regiões com grande diferença de altitude tenham uma quantidade maior de estações para refletir as diferentes condições da atmosfera. Verificou-se também que poucas estações foram definidas em locais que não estão próximos de áreas urbanas ou de vias de acesso, garantindo que a solução atenda as restrições logísticas do problema.

Foram utilizadas duas técnicas para algoritmos geradores de soluções candidatas. A primeira, exata, mostrou-se viável para áreas menores onde a quantidade de opções para instalação não ultrapassou o valor de 1200. À medida que a quantidade de locais candidatos aumentou, observou-se que o tempo médio para encontrar a solução exata cresceu com um comportamento similar ao exponencial. Para estes casos, avaliou-se o algoritmo evolucionário Evolução Diferencial. O algoritmo foi aplicado nas mesmas instâncias utilizadas com o método exato, sendo capaz de encontrar a solução ótima para instâncias com até 400 locais candidatos. Em instâncias maiores, o método encontrou soluções aproximadas, que se afastaram da solução ótima à medida que a quantidade de locais candidatos aumentou, chegando a erro de aproximadamente 24% no experimento com 1781 candidatos.

Uma análise do comportamento da função de *fitness* para este experimento indica que há espaço para melhoria na qualidade da solução, envolvendo uma análise mais completa dos parâmetros do algoritmo, visto que na execução dos experimentos foram utilizados parâmetros definidos experimentalmente. Entretanto, o algoritmo se apresenta como uma opção que pode ser explorada para o problema de otimização combinatória devido à facilidade de implementação, pequena quantidade de parâmetros necessários para execução e apresentação de resultados adequados.

O método, de forma geral, apresenta-se como uma opção viável para o planejamento de redes de estações meteorológicas, pois permite ao planejador inserir as informações disponíveis na forma de camadas de informações e utilizar diferentes algoritmos para avaliar os resultados. A resolução espacial das soluções apresentadas ficaram restritas somente à resolução das camadas fornecidas.

O resultado do primeiro experimento do cenário 1, sobre a quantidade mínima necessária de estações, indica que existe espaço para expansão na rede atual do SIMEPAR, saindo da situação atual, com 40 estações, para a quantidade indicada no experimento, 102 estações. Esta quantidade é adequada aos planos da instituição para os próximos anos. O resultado do primeiro experimento do cenário 2 apresentou uma rede ideal em que as



estações existentes foram consideradas e o experimento do cenário 3 apresentou uma estratégia para expansão da rede considerando uma capacidade anual para instalação de 20 estações meteorológicas.

O desenvolvimento desta dissertação foi acompanhada pela área de Desenvolvimento e Integração do SIMEPAR e foi considerado que em próximos planejamentos de instalação de estações, a método ora estudado será avaliado com subsídio do novo trabalho.

O método proposto pode auxiliar bastante os planejadores em estudos futuros procurando sempre obter melhores soluções em menos tempo.

Para tais estudos poderia ser testada a utilização de algoritmos gulosos para a fase de determinação da quantidade mínima necessária de estações para cobrir todo o espaço em análise e o algoritmo de p-mediana de Teitz & Bart para a fase de planejamento da expansão das redes.

Outra melhoria poderia ser o uso de raios de cobertura variáveis de acordo com a variabilidade das observações. Regiões com pouca variabilidade podem utilizar um raio de cobertura maior enquanto áreas com maior variabilidade podem utilizar raios menores, implicando em uma quantidade maior de estações na região.

O método proposto poderia ser ajustado para outras aplicações como localização de sensores de descargas elétricas e radares meteorológicos, pois são instalações que podem beneficiar-se do uso do modelo numérico do tempo na avaliação da qualidade das soluções.

O método proposto também poderia ser ajustado para usar dados de sensoriamento remoto, como satélites meteorológicos, na avaliação da qualidade das soluções em vez do modelo numérico de previsão do tempo.

## REFERÊNCIAS

- AMORIM, A. M. T. **Un contributo para definição de uma rede otimizada de monitorização da temperatura em Portugal Continental**. Dissertação (Mestrado) — Instituto Superior Técnico, Lisboa, Portugal, Maio 2008. Dissertação para obtenção do Grau de Mestre em Sistemas de Informação Geográfica.
- AMORIM, A. M. T. et al. Optimizing the location of weather monitoring stations using estimation uncertainty. **International Journal of Climatology**, v. 3, n. 6, p. 941–952, 2011.
- BARNES, S. L. A technique for maximizing details in numerical weather map analysis. **Journal of Applied Meteorology**, v. 3, p. 396–409, 1964.
- BERKELAAR, M.; EIKLAND, K.; NOTEBAERT, P. **Open source (Mixed-Integer) Linear Programming system**. 2004.
- BÜYÜKBAS, E. et al. Instruments and Observing Methods Report. **Training Material on Automated Weather Observing Systems**. Genebra: WMO, 2006. 284 p.
- CAMARA, G.; DAVIS, C.; MONTEIRO, A. M. V. **Introdução a Ciência da Geoinformação**. São José dos Campos: INPE, 2001. 345 p.
- CAMERON, K. Better optimization of long-term monitoring networks. **Bioremediation Journal**, v. 8, p. 89–107, 2004.
- CHARNEY, J. G.; FJORTOFT, R.; von Neumann, J. Numeric integration of the vorticity equation. **Tellus**, v. 2, p. 237–254, 1950.
- CORRÊA, G. C. **Otimização da Locação de Subestações para o Planejamento da Expansão da Rede Elétrica de Distribuição**. Dissertação (Mestrado) — Programa de Pós-Graduação em Engenharia Elétrica e Informática Industrial, Universidade Tecnológica Federal do Paraná, Curitiba, PR, 2003.
- CURRENT, J.; DASKIN, M.; SCHILLING, D. Discrete network location models. In: DREZNER, Z.; HAMACHER, H. W. (Ed.). **Facility Location Applications and Theory**. Berlin: Springer-Verlag, 2004. p. 81–118.
- DALEY, R. **Atmospheric data analysis**. Cambridge: Cambridge University Press, 1991. 472 p.
- DASKIN, M. S. **Network and Discrete Location: Models, Algorithms and Applications**. New York: Wiley, 1995. 520 p.
- DASKIN, M. S. What you should know about location modeling. **Naval Research Logistics**, v. 55, p. 284–294, 2008.
- DHAR, A.; DATTA, B. Global optimal design of ground water monitoring network using embedded kriging. **Ground Water**, v. 47, n. 6, p. 806–815, 2009.

DREZNER, Z.; HAMACHER, H. W. (Ed.). **Facility Location Applications and Theory**. Berlin: Springer-Verlag, 2004. 330 p. (Facility Location Applications and Theory).

DREZNER, Z. et al. The weber problem. In: DREZNER, Z.; HAMACHER, H. W. (Ed.). **Facility Location Applications and Theory**. Berlin: Springer-Verlag, 2004. p. 1–36.

DRUCK, S. et al. (Ed.). **Análise Espacial de Dados Geográficos**. Planaltina: EMBRAPA, 2004. 208 p.

ECMWF, E. C. f. M.-R. W. F. **ECMWF Data Coverage (All obs DA) - Synop-Ship-Metar**. 2013. Disponível em: <<http://www.ecmwf.int/products/forecasts/d/charts/monitoring/coverage/dcover>>.

EISELT, H. A.; MARIANOV, V. (Ed.). **Foundations of Location Analysis. International Series in Operations Research & Management Science**. New York, Dordrecht, Heidelberg, London: Springer, 2011.

ESRI. **ESRI Shapefile Technical Description**. 1998. Disponível em: <<http://www.esri.com/library/whitepapers/pdfs/shapefile.pdf>>.

FEOKTISTOV, V. **Differential Evolution: In Search of Solutions (Springer Optimization and Its Applications)**. Secaucus, NJ, USA: Springer-Verlag New York, Inc., 2006. ISBN 0387368957.

GAREY, M. R.; JOHNSON, D. S. **Computers and Intractability: A guide to the Theory of NP-Completeness**. New York: W. H. Freeman and Co., 1979.

ICAGA, Y. Genetic algorithm usage in water quality monitoring networks optimization in Gediz (Turkey) river basin. **Environmental Monitoring and Assessment**, v. 108, p. 261–277, 2005.

INMET, I. N. d. M. **Rede de Estações**. 2013. Disponível em: <<http://www.inmet.gov.br/portal/index.php?r=estacoes/mapaEstacoes>>.

INOUE, K.; TANAKA, T.; MASAKI, I. Application of a genetic algorithm to the design of groundwater monitoring well networks under conditions of uncertainty. In: **Proceedings of the Fifteenth International Offshore and Polar Engineering Conference**, p. 615–624, 2005.

JAIMES, A. et al. Selecting the best location for a meteorological tower: A case study of multi-objective constraint optimization. In: **Martine Ceberio (ed.), Abstracts of the Second Workshop on Constraint Programming and Decision Making CoProD'09**, p. 56–60, 2009.

JARAMILLO, J. H.; BHADURY, J.; BATA, R. On the use of genetic algorithms to solve location problems. **Computers & Operations Research**, v. 29, p. 761–779, 2002.

JIMENEZ, N. et al. A methodology for the design of quasi-optimal monitoring networks for lakes and reservoirs. **Journal of Hydroinformatics**, v. 7, n. 2, p. 105–116, 2005.

JSON. **JSON - JavaScript Object Notation**. 2013. Disponível em: <<http://www.json.org/json-pt.html>>.

LI, Y.; HILTON, A. B. C. Reducing spatial sampling in long-term groundwater monitoring networks using ant colony optimization. **International Journal of Computational Intelligence Research**, v. 1, n. 1, p. 19–28, 2005a.

- LI, Y.; HILTON, A. B. C. Analysis of the primal and dual problem for long-term groundwater monitoring spatial optimization. **In: Proceedings of World Water and Environmental Resources Congress**, Anchorage, Alaska, p. 1–11, 2005b.
- LI, Y.; HILTON, A. B. C. An algorithm for groundwater long-term monitoring spatial optimization by analogy to ant colony optimization for TSP. **In: Proceedings of World Environmental and Water Resources Congress**, Omaha, Nebraska, p. 1–6, 2006a.
- LI, Y.; HILTON, A. B. C. Bayesian statistics-based procedure for the groundwater long-term monitoring temporal optimization problem. **In: Proceedings of World Environmental and Water Resources Congress**, Omaha, Nebraska, p. 1–5, 2006b.
- LI, Y.; HILTON, A. B. C.; TONG, L. Development of ant colony optimization for long-term groundwater monitoring. **In: Proceedings of World Water and Environmental Resource Congress 2004**, p. 1–10, 2004.
- LYNCH, P. The origins of computer weather prediction and climate modeling. **Journal of Computational Physics**, v. 227, p. 3431–3444, 2008.
- MAGICS++. **ECMWF's Meteorological plotting software**. 2013. Disponível em: <<https://software.ecmwf.int/wiki/pages/viewpage.action?pageId=14975084>>.
- MARTINEZ, S. I.; MERWADE, V.; MAIDMENT, D. Linking GIS, hydraulic modeling, and tabu search for optimizing a water level-monitoring network in South Florida. **Journal of Water Resources Planning and Management**, v. 136, n. 2, p. 167–176, 2010.
- MATOS, P. P. O. **Impacto da utilização de dados de temperatura da superfície do mar de alta resolução espacial em um modelo de previsão numérica do tempo**. Dissertação (Mestrado) — Instituto Nacional de Pesquisas Espaciais, 2009.
- MINCIARDI, R.; SACILE, R.; SICCARDI, F. Optimal planning of a weather radar network. **Journal of Atmospheric and Oceanic Technology**, v. 20, p. 1251–1263, 2003.
- NOAA. **Climate Model**. Novembro 2007. [http://celebrating200years.noaa.gov/breakthroughs/climate\\_model](http://celebrating200years.noaa.gov/breakthroughs/climate_model).
- PLASTRIA, F. Continuous covering location problems. In: DREZNER, Z.; HAMACHER, H. W. (Ed.). **Facility Location Applications and Theory**. Berlin: Springer-Verlag, 2004. p. 37–80.
- POSTGIS. **About PostGIS**. 2013. Disponível em: <<http://www.postgis.org>>.
- POSTGRESQL. **About PostgreSQL**. 2013. Disponível em: <<http://www.postgresql.org/about>>.
- PRICE, K. Genetic annealing. **Dr. Dobb's**, october 1994. Disponível em: <<http://www.drdoobbs.com/database/genetic-annealing/184409333>>.
- PRICE, K.; STORN, R. M.; LAMPINEN, J. A. **Differential Evolution: A Practical Approach to Global Optimization (Natural Computing Series)**. Secaucus, NJ, USA: Springer-Verlag New York, Inc., 2005. ISBN 3540209506.

- PYTHON. **Linguagem de programação Python**. 2013. Disponível em: <<http://www.python.org>>.
- QGIS. **Quantum GIS - About QGIS**. 2013. Disponível em: <<http://www.qgis.org/en/about-qgis.html>>.
- RATLE, A. Kriging as a surrogate fitness landscape in evolutionary optimization. **Artificial Intelligence for Engineering Design, Analysis and Manufacturing**, v. 15, p. 37–49, 2001.
- RICHARDSON, L. F. **Weather Prediction by Numerical Process**. Cambridge: Cambridge University Press, 1922. 250 p.
- RUIZ-CARDENAS, R.; FERREIRA, M. A. R.; SCHMIDT, A. M. Stochastic search algorithms for optimal design of monitoring networks. **Environmetrics**, v. 21, p. 102–112, 2010.
- RUIZ-CARDENAS, R.; FERREIRA, M. A. R.; SCHMIDT, A. M. Evolutionary Markov chain Monte Carlo algorithms for optimal monitoring network designs. **Statistical Methodology**, v. 9, p. 185–194, 2012.
- SCAPARRA, M. P.; SCUTELLA, M. G. **Facilities, Locations, Customers: Building Blocks of Location Models. A Survey**. 2001.
- SHAPELY. **Python package for manipulation and analysis of planar geometric objects**. 2013. Disponível em: <<http://pypi.python.org/pypi/Shapely>>.
- SILVA, A. de B. **Sistemas de Informações Geo-referenciadas: conceitos e fundamentos**. Campinas, SP: Editora UNICAMP, 2003.
- SILVEIRA, R. **Introducción a el tópico de red de observaciones**. [S.l.], 2010.
- SIMEPAR. **SIMEPAR - Sistema Meteorológico do Paraná**. 2013. Disponível em: <<http://www.simepar.br>>.
- SRTM. **Shuttle Radar Topography Mission**. 2013. Disponível em: <<http://www2.jpl.nasa.gov/srtm/>>.
- STORN, R. On the usage of differential evolution for function optimization. **Biennial Conference of the North American Fuzzy Information Processing Society (NAFIPS1996)**, Berkeley, p. 519 – 523, 1996.
- SUGAHARA, S.; SILVEIRA, R. B.; ROCHA, R. P. Estimativa da probabilidade do evento extremo de precipitação de janeiro de 2000 no Vale do Paraíba, baseada na distribuição generalizada de Pareto. **Revista Brasileira de Geofísica**, v. 28, n. 2, p. 193–208, 2010.
- TAMIR, A. An  $o(pn^2)$  algorithm for p-median and related problems on tree graphs. **Operations Research Letters**, v. 19, p. 59–64, 1996.
- TOREGAS, C. et al. The location of emergency service facilities. **Operations Research**, v. 19, p. 1363–1373, 1971.
- WEBER, A. **Theory of the Location of Industries**. Chicago: University of Chicago Press, 1929.

WEINERT, W. R. **Computação Evolucionária para Indução de Regras de Autômatos Celulares Multidimensionais**. Tese (Doutorado) — Programa de Pós-Graduação em Engenharia Elétrica e Informática Industrial, Universidade Tecnológica Federal do Paraná, Curitiba, PR, 2010.

WEISZFELD, E. Sur le point pour lequel la somme des distances de  $n$  points donnes est minimum. **Tohoku Math Journal**, v. 43, p. 355–386, 1936.

YEH, M.-S.; LIN, Y.-P.; CHANG, L.-C. Designing an optimal multivariate geostatistical groundwater quality monitoring network using factorial kriging and genetic algorithms. **Environmental Geology**, p. 101–121, 2006.

## APÊNDICE A – MODELO MATEMÁTICO

---

### Algoritmo 4: Modelagem matemática do problema da cobertura de conjunto

---

1 MIN:  $0.85 x_0 + 0.85 x_1 + 0.35 x_2 + 0.85 x_3 + 0.85 \dots + 0.85 x_{23} + 0.85 x_{24}$ ;  
 2  $x_0 + x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 + x_8 + x_9 + \dots + x_{20} + x_{21} + x_{22} > 1$ ;  
 3  $x_0 + x_2 + x_3 + x_4 + x_6 + x_7 + x_8 + x_9 + x_{10} + x_{11} + \dots + x_{21} + x_{22} > 1$ ;  
 4  $x_0 + x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 + x_8 + x_9 + \dots + x_{22} + x_{23} > 1$ ;  
 5  $x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 + x_8 + x_{11} + x_{12} + \dots + x_{23} > 1$ ;  
 6  $x_0 + x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 + x_8 + x_9 + \dots + x_{23} + x_{24} > 1$ ;  
 7  $x_0 + x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 + x_8 + x_9 + \dots + x_{23} + x_{24} > 1$ ;  
 8  $x_0 + x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 + x_8 + x_9 + \dots + x_{23} + x_{24} > 1$ ;  
 9  $x_0 + x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 + x_8 + x_9 + \dots + x_{23} + x_{24} > 1$ ;  
 10  $x_0 + x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 + x_8 + x_9 + \dots + x_{23} + x_{24} > 1$ ;  
 11  $x_0 + x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 + x_8 + x_9 + \dots + x_{23} + x_{24} > 1$ ;  
 12  $x_0 + x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 + x_8 + x_9 + \dots + x_{22} + x_{23} > 1$ ;  
 13  $x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 + x_8 + x_{11} + x_{12} + \dots + x_{22} + x_{23} > 1$ ;  
 14  $x_0 + x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 + x_8 + x_9 + \dots + x_{23} + x_{24} > 1$ ;  
 15  $x_0 + x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 + x_8 + x_9 + \dots + x_{23} + x_{24} > 1$ ;  
 16  $x_0 + x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 + x_8 + x_9 + \dots + x_{23} + x_{24} > 1$ ;  
 17  $x_0 + x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 + x_8 + x_9 + \dots + x_{23} + x_{24} > 1$ ;  
 18  $x_0 + x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 + x_8 + x_9 + \dots + x_{23} + x_{24} > 1$ ;  
 19 ...  
 20  $x_2 + x_3 + x_4 + x_6 + x_7 + x_8 + x_9 + x_{10} + x_{11} + x_{12} + \dots + x_{23} + x_{24} > 1$ ;  
 21  $x_2 + x_3 + x_4 + x_6 + x_7 + x_8 + x_9 + x_{10} + x_{11} + x_{12} + \dots + x_{23} + x_{24} > 1$ ;  
 22 BIN  $x_0, x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}, x_{11}, \dots, x_{22}, x_{23}, x_{24}$ ;

---

O algoritmo 4 apresenta a modelagem matemática do problema de cobertura de conjuntos no formato do programa *lp\_solve*.

## APÊNDICE B – LOCALIZAÇÃO DAS ESTAÇÕES EXISTENTES

A Tabela 10 apresenta uma lista das estações existentes e que foram utilizadas nos experimentos envolvendo planejamento de uma rede real.

**Tabela 10: Localização das estações existentes**

Nome	Longitude	Latitude
Antonina	-48.80000	-25.22000
Apucarana	-51.54470	-23.50750
Assis Chateaubriand	-53.54110	-24.38860
Cambará	-50.03330	-23.00000
Campo Mourão	-52.36660	-24.05000
Cândido de Abreu	-51.25000	-24.63330
Cascavel	-53.55000	-24.88330
Cerro Azul	-49.25000	-24.81660
Cianorte	-52.54760	-23.65775
Curitiba	-49.23033	-25.44817
Entre Rios	-51.48330	-25.55000
Fernandes Pinheiro	-50.58330	-25.45000
Foz do Areia	-51.65000	-26.08330
Foz do Iguaçu	-54.61670	-25.40000
Francisco Beltrão	-53.05080	-26.05930
Guaira	-54.25000	-24.06670
Guarapuava	-51.50000	-25.35000
Guaratuba	-48.57610	-25.84520
Jaguariaiva	-49.67700	-24.22690
Lapa	-49.76660	-25.78330
Londrina	-51.15000	-23.30000
Maringá	-51.98330	-23.45000
N.Prata do Iguaçu	-53.51660	-25.56660
Palmas	-51.96670	-26.46670
Palmital	-52.21660	-24.88330
Palotina	-53.91660	-24.30000
Paranavaí	-52.43330	-23.08330
Pato Branco	-52.68330	-26.11660



**Tabela 11: Localização das estações existentes - continuação**

Nome	Longitude	Latitude
Pato Branco	-52.68330	-26.11660
Pinhais	-49.13330	-25.41670
Pinhão	-51.96250	-25.64944
Ponta Grossa	-50.01660	-25.21660
Reserva Natural Salto Morato	-48.28810	-25.17790
Salto Caxias	-53.48330	-25.51660
Salto Osorio	-53.01670	-25.51670
Santa Helena	-54.31030	-24.91680
São Miguel do Iguaçu	-54.25460	-25.35280
Telêmaco Borba	-50.61660	-24.33330
Toledo	-53.71670	-24.78330