

**UNIVERSIDADE ESTADUAL DE CAMPINAS
FACULDADE DE ENGENHARIA MECÂNICA
INSTITUTO DE GEOCIÊNCIAS**

**Aplicação da Metaheurística Busca Dispersa ao
Problema do Ajuste de Histórico**

Autor: **Sergio Henrique Guerra de Sousa**
Orientador: **Prof. Dr. Denis José Schiozer**
Co-orientador: **Dr. Célio Maschio**

FICHA CATALOGRÁFICA ELABORADA PELA
BIBLIOTECA DA ÁREA DE ENGENHARIA E ARQUITETURA - BAE - UNICAMP

So85a Sousa, Sergio Henrique Guerra de
 Aplicação da methaheurística busca dispersa ao
 problema do ajuste de histórico / Sergio Henrique Guerra
 de Sousa. --Campinas, SP: [s.n.], 2007.

 Orientadores: Denis José Schiozer, Célio Maschio
 Dissertação (mestrado) - Universidade Estadual de
 Campinas, Faculdade de Engenharia Mecânica e Instituto
 de Geociências.

 1. Engenharia de reservatório de óleo. 2. Otimização
 matemática. 3. Programação heurística. 4. Simulação
 (Computadores). 5. Programação paralela (Computação).
 6. Engenharia de petróleo. I. Schiozer, Denis José. II.
 Maschio, Célio. III. Universidade Estadual de Campinas.
 Faculdade de Engenharia Mecânica. IV. Instituto de
 Geociências. V. Título.

Título em Inglês: Application of the scatter search methaheuristic to the history
 matching problem

Palavras-chave em Inglês: History matching, Optimization, Scatter search,
 Reservoir simulation

Área de concentração: Reservatórios e Gestão

Titulação: Mestre em Ciências Engenharia de Petróleo

Banca examinadora: Leandro Costa Reis e Sérgio Nascimento Bordalo

Data da defesa: 14/02/2007

Programa de Pós-Graduação: Ciências Engenharia de Petróleo

**UNIVERSIDADE ESTADUAL DE CAMPINAS
FACULDADE DE ENGENHARIA MECÂNICA
INSTITUTO DE GEOCIÊNCIAS**

Aplicação da Metaheurística Busca Dispersa ao Problema do Ajuste de Histórico

Autor: Sergio Henrique Guerra de Sousa
Orientador: Prof. Dr. Denis José Schiozer
Co-orientador: Dr. Célio Maschio

Curso: Ciências e Engenharia de Petróleo

Dissertação de mestrado apresentada à
Subcomissão de Pós-Graduação Interdisciplinar
de Ciências e Engenharia de Petróleo (FEM e
IG), como requisito para a obtenção do título de
Mestre em Ciências e Engenharia de Petróleo.

Campinas, 2007
SP – Brasil

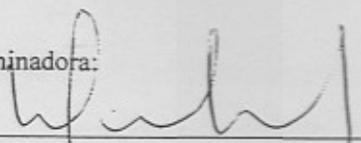
UNIVERSIDADE ESTADUAL DE CAMPINAS
FACULDADE DE ENGENHARIA MECÂNICA
INSTITUTO DE GEOCIÊNCIAS

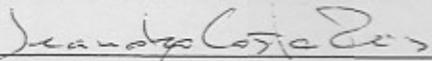
DISSERTAÇÃO DE MESTRADO

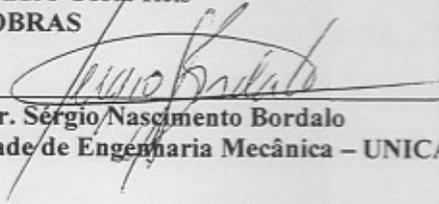
**Aplicação da Metaheurística Busca Dispersa ao
Problema do Ajuste de Histórico**

Autor: Sergio Henrique Guerra de Sousa
Orientador: Prof. Dr. Denis José Schiozer
Co-orientador: Dr. Célio Maschio

Banca Examinadora:


Prof. Dr. Denis José Schiozer, Presidente
Faculdade de Engenharia Mecânica – UNICAMP


Dr. Leandro Costa Reis
PETROBRAS


Prof. Dr. Sérgio Nascimento Bordalo
Faculdade de Engenharia Mecânica – UNICAMP

Campinas, 14 de fevereiro de 2007

DEDICATÓRIA

Dedico este trabalho aos meus pais, que, com seus exemplos de estudo, dedicação e perseverança, abriram o caminho do saber, da academia e da pesquisa científica para mim.

AGRADECIMENTOS

Agradeço a Deus por ter me dado o dom da vida e da razão, sem as quais este trabalho não seria possível.

Agradeço a meus pais, Ivan Sergio Freire de Sousa e Maria de Fátima Guerra de Sousa, pela educação de qualidade da qual sempre usufruí, pelo exemplo de dedicação aos estudos dado desde quando obtiam seus títulos de Ph.D. até o presente e pelo suporte emocional e financeiro nestes anos de UNICAMP desde meu curso de graduação até o mestrado.

Agradeço a meus irmãos, Flavia Maria Guerra de Sousa Aranha Oliveira e Ivan Sergio Freire de Sousa Jr., pelo apoio e incentivo nos meus anos em Campinas.

Agradeço à minha companheira, Sandra Fonsêca Rêgo, pelas noites e finais de semana nas quais me acompanhou para fazer hora extra na UNICAMP, pela abdicação dos feriados, viagens e passeios que fizemos em favor deste trabalho e também pela infinita paciência com meu mau humor nos momentos de maior estresse.

Agradeço ao meu orientador, Prof. Dr. Denis José Schiozer, por despertar meu interesse pela área de Petróleo desde o final de minha graduação, quando fui seu estagiário, por me incentivar a permanecer na área após a formatura, por me dar apoio financeiro neste período todo e, principalmente, pela excelente orientação desde a escolha do tema da pesquisa até as revisões finais em período de festividades. Agradeço também ao meu co-orientador, Dr. Célio Maschio, que, ao longo da elaboração deste trabalho, conseguiu sanar todas as dúvidas às quais o expus, me aconselhou em diversas etapas do processo e, sobretudo, forneceu um exemplo de como ser um pesquisador.

Agradeço também a todos os professores do Departamento de Engenharia de Petróleo da FEM/UNICAMP pelos ensinamentos, pelos projetos nos quais pude dar alguma contribuição e pelas sugestões dadas em suas salas, nos corredores do departamento, ou mesmo na copa em

conversas informais. Aos funcionários e estagiários do DEP, agradeço a infra-estrutura fornecida, o suporte prestado e a convivência antes e durante o período de meu mestrado.

À UNICAMP, eu agradeço a sensação de orgulho que tenho ao defender um trabalho desta natureza em uma instituição tão “jovem” e tão prestigiada quanto é esta Universidade.

Finalmente, agradeço à FAPESP pelo apoio financeiro fornecido durante minha formação anterior na área de otimização, quando cursava disciplinas de mestrado na Faculdade de Engenharia Elétrica da UNICAMP.

“Deixe-me lhe contar o que nós israelenses temos contra Moisés. Ele nos arrastou pelo deserto durante 40 anos para nos trazer ao único lugar do Oriente Médio onde não há petróleo.”

Golda Meir (1898-1978) – Primeira Ministra de Israel

“Petróleo é raramente encontrado onde ele é mais necessário e raramente é o mais necessário onde ele é encontrado.”

L. E. J. Brouwer, Diretor Gerente Senior, Royal Dutch/Shell Oil Co

RESUMO

SOUSA, Sergio Henrique Guerra de, **Aplicação da Metaheurística Busca Dispersa ao Problema do Ajuste de Histórico**. Campinas, 2006, 128pp. Dissertação (Mestrado em Ciências e Engenharia de Petróleo) – Faculdade de Engenharia Mecânica e Instituto de Geociências, Universidade Estadual de Campinas – UNICAMP, 2006.

O problema do ajuste de histórico é uma das tarefas que mais demandam tempo em um estudo de reservatório baseado em simulações de fluxo, porque é um problema inverso onde os resultados (dados de produção) são conhecidos, porém os valores de entrada (a caracterização do reservatório) não são integralmente conhecidos. Adicionalmente, as funções objetivo que medem a qualidade do ajuste costumam ser expressões compostas por uma série de componentes que tornam a topologia do espaço de soluções complexa e repleta de não linearidades. A metodologia adotada neste trabalho foi a modelagem do problema de ajuste de histórico como um problema de otimização combinatória de modo que ele pudesse ser abordado através de processos metaheurísticos. Em particular, a metaheurística Busca Dispersa (Scatter Search) foi acoplada a um algoritmo de Busca Direta baseado no método de Hooke e Jeeves para resolver o problema do ajuste de histórico. Reservatórios sintéticos de solução conhecida foram utilizados para fazer a validação da metodologia e, em seguida, ela foi aplicada a outro reservatório, também sintético, mas com características de reservatórios reais onde a solução do ajuste é desconhecida. São discutidos ao longo do texto o uso da metodologia de forma automática e assistida e também os benefícios do uso da computação distribuída na execução do método. As maiores contribuições deste trabalho em relação à questão do ajuste de histórico são: a introdução de uma nova metodologia versátil para uso automático ou assistido, a discussão de algumas características que dificultam o processo de ajuste e de que forma eles podem ser contornados e também a abordagem do tema do ajuste automático vs. o ajuste assistido ilustrado com exemplos.

Palavras Chave: Ajuste de Histórico, Otimização, Busca Dispersa, Simulação de Reservatórios.

ABSTRACT

SOUSA, Sergio Henrique Guerra de, **Application of the Scatter Search Metaheuristic to the History Matching Problem**. Campinas, 2006, 128pp. Thesis (MS in Petroleum Sciences and Engineering) – Faculdade de Engenharia Mecânica and Instituto de Geociências, Universidade Estadual de Campinas – UNICAMP, 2006.

The history matching problem is one of the most demanding tasks in a reservoir simulation study; because it's an inverse problem where the results (production data) are known but the input data (the reservoir characterization data) are not entirely known. Moreover, the objective function that guide the match is usually made out of a series of components that make the topology of the objective function both complex and full of non-linearities. The methodology adopted in this work was to model the history matching problem as a combinatorial optimization problem in order for it to be solved by metaheuristic processes. In particular, the Scatter Search metaheuristic was coupled with a direct search method based on Hooke and Jeeve's method to solve the history matching problem. Synthetic reservoirs of known solutions were used to validate the methodology and then the methodology was applied to another reservoir, also synthetic, but with characteristics of real reservoirs where the solution is not known in advance. Throughout the text, the mixed use of the methodology on both an assisted and automatic fashion is discussed along with the benefits attained by the use of distributed computing resources. The greatest contributions of this work related to the history matching problem are: the introduction of a new versatile methodology for both automatic and assisted matches, the discussion of some characteristics that burden the entire process and some ways to overcome the difficulties, and also the discussion of some tradeoffs between automatic versus assisted history matching with examples to illustrate the matter.

Keywords: History Matching, Optimization, Scatter Search, Reservoir Simulation.

ÍNDICE

Dedicatória	v
Agradecimentos.....	vi
Resumo.....	ix
Abstract	x
Índice.....	xii
Lista de Figuras	xiv
Lista de Tabelas.....	xvi
Nomenclatura	xviii
Capítulo 1 – Introdução.....	1
1.1 Motivação.....	3
1.2 Objetivo.....	4
1.3 Estrutura da Dissertação.....	5
Capítulo 2 – Fundamentação Teórica.....	7
2.1 Ciclo de Vida de Um Reservatório	7
2.1.1 Exploração.....	7
2.1.2 Avaliação.....	8
2.1.3 Desenvolvimento.....	8
2.1.4 Produção.....	9
2.1.5 Abandono	9
2.2 Simulação Numérica de Reservatórios	9
2.2.1 A Revolução da Computação e a Simulação Numérica.....	11
2.2.2 Ajuste de Histórico.....	13
2.2.3 Os Parâmetros de Ajuste	15
2.2.4 Tipos de Ajuste.....	16
2.2.5 Esquema Geral de um Processo de Ajuste	19
2.3 Otimização Local Versus Otimização Global	21
2.4 Otimização Baseada em Simulações.....	23
2.4.1 Otimização Tradicional vs. Planejamento Estatístico	24
2.4.2 Métodos de Busca Direta	25
2.5 Complexidade de Algoritmos.....	26
2.6 Heurísticas e Metaheurísticas.....	28
2.7 Metaheurísticas em Otimização Combinatorial	30
2.8 A Metaheurística Busca Dispersa	31
2.8.1 O Conjunto de Referência	31
2.8.2 Um Modelo Para a Busca Dispersa.....	33
Capítulo 3 – Revisão Bibliográfica	37
3.1 Otimização	37

3.1.1	Otimização Combinatória	37
3.1.2	Busca Direta	37
3.2	Otimização aplicado ao Ajuste de Histórico	38
3.2.1	Métodos de Ajuste de Histórico Baseados em Gradientes.....	39
3.2.2	Métodos de Busca Direta para o Ajuste de Histórico	40
3.2.3	Métodos híbridos.....	40
Capítulo 4	– Metodologia.....	41
4.1	Modelagem do Problema de Ajuste de Histórico Como um Problema de Otimização Combinatória.....	41
4.2	As Propriedades Incertas	42
4.2.1	Discretização do Domínio.....	42
4.3	A Função Objetivo	43
4.3.1	O Cálculo do Afastamento	44
4.3.2	Normalização dos Afastamentos.....	45
4.4	O Espaço de Soluções	47
4.5	A Implementação da Busca Dispersa.....	48
4.5.1	Geração do Conjunto de Referência Inicial	49
4.5.2	Heurística de Melhoria Local.....	53
4.5.3	Algoritmo de Combinação de Pontos.....	58
4.5.4	Atualização do Conjunto de Referência.....	61
4.5.5	Critérios de Parada	62
4.5.6	As Soluções Encontradas	63
4.6	Validação e Aplicação do Método de Otimização	63
Capítulo 5	– Aplicação.....	67
5.1	Casos de Validação	67
5.1.1	Caso 1A.....	67
5.1.2	Caso 1B	70
5.1.3	Caso 1C	72
5.2	Caso de Aplicação (Caso 2)	75
Capítulo 6	– Resultados e Discussões	79
6.1	Casos de Validação	79
6.1.1	Caso 1A.....	79
6.1.2	Caso 1B	86
6.1.3	Caso 1C	94
6.2	Caso 1C – Ajuste Assistido.....	104
6.3	Caso de Aplicação (Caso 2)	111
6.3.1	Fase 1.....	112
6.3.2	Fase 2.....	113
6.3.3	Fase 3.....	114
6.4	Análise de <i>Speedup</i>	116
Capítulo 7	– Conclusões e Sugestões Para Trabalhos Futuros.....	119
Referências	125

LISTA DE FIGURAS

Figura 2.1 – Atividades típicas em um estudo de simulação – fonte: Mattax <i>et al.</i> , 1990.....	14
Figura 2.2 - Exemplo contendo caso (modelo) base e ajustado.	15
Figura 2.3 - Distribuições de probabilidade de parâmetros do reservatório.	16
Figura 2.4 - Algoritmo básico de um ajuste de histórico 20	20
Figura 2.5 – Gráfico de $x^2 - 4x + 2$ no domínio $0 \leq x \leq 4$	21
Figura 2.6 – Gráfico da função $\text{sen}(x) + x/5$ no domínio $-10 \leq x \leq 10$	22
Figura 2.7 – Otimização Baseada em Simulações na otimização tradicional e em planejamento estatístico.....	24
Figura 2.8 - Classificação dos problemas segundo sua complexidade 27	27
Figura 2.9 - Inserção de vetores no conjunto de referência através de combinações lineares 32	32
Figura 2.10 - Modelo geral do algoritmo da Busca Dispersa..... 34	34
Figura 4.1 - Exemplo de uma propriedade incerta. 42	42
Figura 4.2 - Discretização de uma propriedade incerta..... 43	43
Figura 4.3 - O espaço de soluções é o contradomínio da função objetivo 48	48
Figura 4.4 – Exemplo de um espaço de soluções..... 48	48
Figura 4.5 – Seleção dos valores de formação dos pontos de amostragem. 50	50
Figura 4.6 – Enumeração completa dos valores maior (+) e menor (–) dos parâmetros P_i 51	51
Figura 4.7 – Crescimento das enumerações com o aumento do número de parâmetros de ajuste. 52	52
Figura 4.8 – Espaço de soluções da Tabela 4.1..... 54	54
Figura 4.9 – Combinação linear dos pontos a e b 59	59
Figura 4.10 – Pontos gerados pela combinação linear de a e b 60	60
Figura 5.1 - Mapa de permeabilidade horizontal do Caso 1A 68	68
Figura 5.2 – Comparação das vazões de água do modelo base versus histórico do Caso 1A. 69	69
Figura 5.3 - Mapa de permeabilidade horizontal do Caso 1B..... 71	71
Figura 5.4 – Comparação entre as vazões de água do modelo base e o histórico do Caso 1B 72	72
Figura 5.5 – Mapa de permeabilidade horizontal do Caso 1C..... 73	73
Figura 5.6 – Comparação das vazões de produção de água do modelo base com o histórico (Caso 1C)..... 75	75
Figura 5.7 – Comparação entre os cortes de água do caso base em relação ao histórico (Caso 2). 77	77
Figura 6.1 – Dispersão das soluções encontradas (Caso 1A com 57 níveis). 80	80
Figura 6.2 – Ajuste da vazão de água nos poços produtores do Caso 1A com 57 níveis. 81	81
Figura 6.3 – Custo do algoritmo em número de simulações (Caso 1A com 57 níveis). 82	82
Figura 6.4 – Impacto do número de níveis no número de simulações e no valor da função objetivo (Caso 1A). 83	83
Figura 6.5 – Desempenho da busca dispersa em número de simulações (Caso 1A). 84	84
Figura 6.6 – Evolução da função objetivo por iteração do método (Caso 1A). 85	85
Figura 6.7 – Ajuste das melhores soluções de cada espaço de soluções explorado (Caso 1A). 86	86

Figura 6.8 – Dispersão das soluções encontradas relativo aos valores geradores do histórico.....	88
Figura 6.9 - Ajuste da vazão de água nos poços produtores do Caso 1B com 57 níveis.	88
Figura 6.10 – Custo da metodologia para o Caso 1B de 57 níveis em número de simulações.....	90
Figura 6.11 - Impacto do número de níveis no número de simulações e no valor da função objetivo (Caso 1B)	91
Figura 6.12 – Desempenho da busca dispersa em número de simulações (Caso 1B).....	92
Figura 6.13 - Evolução da função objetivo por iteração do método (Caso 1B).....	93
Figura 6.14 – Ajuste das melhores soluções de cada espaço de soluções explorado (Caso 1B) ...	94
Figura 6.15 – Dispersão das soluções de referência (Caso 1C com 57 níveis).....	96
Figura 6.16 – Ajuste da vazão de água nos poços do Caso 1C com 57 níveis.....	96
Figura 6.17 - Impacto do número de níveis no número de simulações e no valor da função objetivo (Caso 1C).	99
Figura 6.18 – Dispersão das soluções encontradas em relação aos valores geradores do histórico (Caso 1C).....	100
Figura 6.19 – Evolução da função objetivo durante a exploração dos espaços de soluções.....	102
Figura 6.20 – Ajuste da vazão de água dos poços do Caso 1C.....	102
Figura 6.21 – Representação esquemática do caso base do ajuste assistido	105
Figura 6.22 – Espaço de soluções e resultado do ajuste na Fase 1	106
Figura 6.23 – Espaço de soluções e resultado do ajuste na Fase 2	107
Figura 6.24 – Espaço de soluções e resultado do ajuste na Fase 3	108
Figura 6.25 – Ajuste da vazão de água no Caso 1C assistido	109
Figura 6.26 – Ajuste do corte de água dos poços produtores (Caso 2).....	115
Figura 6.27 – Ajuste da vazão de água nos poços injetores (Caso 2).	116
Figura 6.28 – Simulações absolutas versus simulações equivalentes no Caso 1C	118

LISTA DE TABELAS

Tabela 4.1 – Espaço de soluções de exemplo discretizado	49
Tabela 4.2 – Pontos da amostragem reduzida para o espaço da Tabela 4.1.....	52
Tabela 4.3 – Determinação do vetor direção da busca linear.....	55
Tabela 4.4 – Pontos da busca linear até a fronteira.....	56
Tabela 4.5 – Exemplo de um conjunto de pontos explorados e suas distâncias relativas.....	61
Tabela 5.1 - Principais propriedades do reservatório do Caso 1A	68
Tabela 5.2 - Regiões e valores de permeabilidade do histórico do Caso 1A	69
Tabela 5.3 – Quatro diferentes espaços de soluções para o Caso 1A.	70
Tabela 5.4 – Regiões e valores de histórico das permeabilidades do Caso 1B.....	71
Tabela 5.5 – Definição dos mapas de permeabilidade horizontal e vertical (Caso 1C).....	74
Tabela 6.1 – Valores encontrados para os parâmetros de ajuste (Caso 1A com 57 níveis).....	80
Tabela 6.2 – Desempenho da Busca Dispersa (Caso 1A com 57 níveis).....	81
Tabela 6.3 – Soluções encontradas nas diferentes discretizações do espaço de soluções.	84
Tabela 6.4 – Quatro diferentes espaços de soluções para o Caso 1B.....	86
Tabela 6.5 - Valores encontrados para os parâmetros de ajuste (Caso 1B com 57 níveis).....	87
Tabela 6.6 – Custo do ajuste do Caso 1B em número de simulações	89
Tabela 6.7 - Soluções encontradas nas diferentes discretizações do espaço de soluções (Caso 1B)	92
Tabela 6.8 – Quatro diferentes espaços de soluções para o Caso 1C.....	95
Tabela 6.9 - Valores encontrados para os parâmetros de ajuste (Caso 1C com 57 níveis).....	95
Tabela 6.10 – Valores e afastamentos encontrados para as duas melhores soluções de cada espaço.	101
Tabela 6.11 – Espaço de soluções e custo das fases do ajuste assistido	110
Tabela 6.12 – Tabela de cálculo do <i>speedup</i> da Busca Dispersa	117

NOMENCLATURA

Letras Latinas

A	medida do afastamento entre dados simulados e de histórico
D	conjunto de dados de produção/injeção simulados
F_i	função objetivo parcial
FOG	função objetivo global
H	conjunto de dados de histórico de produção/injeção
inf	limite inferior do intervalo de definição de um parâmetro de ajuste
K_x	parâmetro permeabilidade horizontal
K_z	parâmetro permeabilidade vertical
p	parâmetro de ajuste (ou parâmetro incerto)
Q_w	vazão de água (m^3/dia)
S	espaço de soluções para uma dada parametrização e discretização
sup	limite superior do intervalo de definição de um parâmetro de ajuste
t	tempo (dias)
w_i	pesos nas funções objetivo parciais ou global

Sobrescritos

* dado usa interpolação

Subscrito

caso refere-se a um caso qualquer no espaço de soluções

base refere-se ao caso base a partir do qual o espaço de soluções é gerado
obs refere-se a um dado de histórico
sim refere-se a um dado simulado

Abreviaturas

E&P Exploração e Produção

Capítulo 1 – INTRODUÇÃO

A simulação numérica de reservatórios se estabeleceu há anos como uma ferramenta fundamental no gerenciamento de reservatórios (Staggs *et al.*, 1971). Isto porque ela é a melhor ferramenta de análise quantitativa do fluxo multifásico em reservatórios heterogêneos (Mattax *et al.*, 1990). Existem diferentes modelos de simulação para diferentes premissas sobre o reservatório, cabe ao engenheiro de reservatórios escolher aquele que melhor representa a dinâmica do reservatório real. É importante observar que as respostas obtidas do simulador, que incluem os dados de produção e outras variáveis do reservatório como mapas de saturação e pressão, são funções do conjunto de dados de alimentação. Assim, a qualidade dos resultados é diretamente proporcional à qualidade dos dados de entrada.

O objetivo principal do ajuste de histórico é calibrar os dados de entrada de modo que a saída fornecida pelo simulador de reservatórios reflita, da melhor maneira possível, os dados observados. Presume-se que modelos bem ajustados, ou seja, que conseguem refletir o histórico de produção observado no reservatório real, produzam melhores previsões de desempenho do que modelos desajustados. O processo de ajuste de histórico normalmente consiste em minimizar uma função objetivo composta pelas diferenças entre os dados observados e os simulados. Este tipo de função objetivo é não linear e a natureza dos modelos de simulação é tal que normalmente não existe uma única solução para o problema.

O que torna o ajuste de histórico um problema complexo é que se trata de um problema inverso com muitos parâmetros, onde a saída esperada é conhecida (o histórico), mas não se conhece quais dados de entrada levam ao ajuste. Adicionalmente, o problema não é bem-posto, ou seja, não obedece às propriedades definidas pelo matemático francês Hadamard para modelos matemáticos de fenômenos físicos. São elas (Romanov, 1997): (1) a solução existe, (2) a solução é única e (3) a solução depende continuamente dos dados de entrada com alguma topologia

razoável. Pelo menos a segunda propriedade, na maioria das situações, não é satisfeita pelos modelos matemáticos representados pelos simuladores numéricos de reservatórios.

Tradicionalmente o ajuste de histórico tem sido feito manualmente, em um processo laborioso e demorado cujo sucesso é altamente dependente da experiência do engenheiro de reservatórios e também de sua organização e método ao realizar a tarefa. Mais recentemente, diversas iniciativas na direção de automatizar partes ou a íntegra do processo tem sido realizadas.

A maior dificuldade na automação dos processos de ajuste está na dimensão do problema. Existe um número muito grande de variáveis no processo de forma que incluir todas elas no processo gera um problema intratável com a tecnologia atual. Mesmo com o avanço exponencial do poder computacional dos microprocessadores, um futuro onde o problema possa ser resolvido incluindo todas as variáveis não é muito provável porque a sofisticação dos modelos numéricos dos reservatórios aumenta a ponto de consumir a capacidade de processamento adicional disponibilizada.

A denominação de ajuste assistido é utilizada para processos de ajuste de histórico onde parte do processo é manual – normalmente a parte de controle de variáveis incertas e seus limites (chamada de parametrização) e determinação da função-objetivo – e parte do processo é automática – normalmente a otimização ligada a minimização da função-objetivo escolhida e as partes de pré-processamento dos dados, montagem de arquivos e pós processamento dos resultados.

Neste trabalho, o foco é na parte automática voltada para a minimização da função-objetivo. É comum nas metodologias de ajuste a existência de um limite prático no tamanho dos problemas que podem ser resolvidos, quer seja pelo tempo de execução proibitivo, ou pelo grande volume de dados gerados pelas simulações. Porém, esta limitação tende a ser mais crítica nos casos onde a parametrização está incluída no ajuste.

Assim, este trabalho centra suas discussões na proposta de uma metodologia de ajuste assistido onde um método de busca local é guiado pela metodologia da Busca Dispersa na superação de valores ótimos locais. A metodologia proposta enquadra-se na categoria de métodos de busca direta, ou seja, que não faz uso de informações de gradiente.

1.1 MOTIVAÇÃO

A simulação de reservatórios tem um importante papel no gerenciamento de reservatórios. Metodologias que visam diminuir o tempo gasto no ajuste de histórico são sempre relevantes, pois este processo, além de ser realizado muitas vezes ao longo da vida de um campo de petróleo, é geralmente a etapa mais demorada em um estudo de simulação.

Enquanto metodologias de ajuste baseadas em gradientes possuem uma boa taxa de convergência para soluções otimizadas, o uso generalizado de simuladores comerciais dificulta ou inviabiliza a aquisição de informações de gradiente. Por outro lado, os modelos numéricos implementados nos simuladores comerciais possuem uma sofisticação e precisão difíceis de reproduzir em simuladores que possam ser executados de forma acoplada com um processo de otimização. Estas razões sugerem o uso dos simuladores comerciais como caixas pretas em metodologias de ajuste assistido.

Muitas iniciativas de ajuste assistido utilizam metodologias fortemente dependentes de aleatorização¹ como os algoritmos evolutivos. Este tipo de método tem a tendência de só convergir com a realização de um número excessivo de simulações numéricas. Por outro lado, as metodologias baseadas em trajetórias, como a Busca Tabu, costumam fornecer apenas uma solução para o problema do ajuste o que vai contra sua natureza de múltiplas soluções. Sendo uma metodologia híbrida combinando aspectos de algoritmos evolutivos e de algoritmos de trajetória, a Busca Dispersa tem o potencial de incorporar as qualidades das duas abordagens de otimização.

Adicionalmente, poucos trabalhos de ajuste assistido fizeram uso da metodologia da Busca Dispersa. O primeiro trabalho encontrado que mencionava o uso desta metodologia só apareceu no final de 2006 (Cullick *et al.*, 2006), ou seja, durante os últimos meses de elaboração deste trabalho.

Outro fator importante na escolha do tema de pesquisa foi a experiência prévia do autor no uso de métodos de busca direta para resolução do problema do ajuste de histórico. Essa

¹ Aleatorização – refere-se a processos baseados na geração de números pseudo-aleatórios para inserir não determinismo em processos. Este termo é muito usado em estatística e no campo da pesquisa operacional.

experiência consistia no uso de uma ferramenta baseada na metodologia de busca local proposta por Hooke *et al.*, 1961, acoplada com algumas técnicas simples para transformar o procedimento em um de busca global. O acoplamento desta metodologia com a Busca Dispersa pareceu uma maneira muito mais robusta de realizar esta transformação.

Finalmente, tanto a Busca Dispersa quanto a metodologia de busca local (Hooke e Jeeves) podem tirar proveito da computação distribuída para melhorar seu desempenho, e por este motivo, a combinação das duas técnicas pareceu uma boa forma de explorar os benefícios da computação distribuída para o problema.

1.2 OBJETIVO

O principal objetivo desta dissertação é estudar a aplicabilidade da metodologia da Busca Dispersa ao problema do ajuste de histórico. O foco do trabalho é ajustar problemas com vários parâmetros onde o espaço de soluções é não-linear, com a presença de diversos mínimos locais. A metodologia proposta deve ser flexível o suficiente para suportar a realização de ajustes de forma automática ou assistida. Adicionalmente, é importante que a metodologia seja capaz de permitir que um balanço seja feito entre qualidade do ajuste e número total de simulações realizadas. Finalmente, por questões de eficiência, é importante que o algoritmo tire proveito de sistemas distribuídos realizando simulações em paralelo quando possível.

Um objetivo secundário deste trabalho é gerar um programa de computador que implementa a metodologia proposta de forma que esta possa ser melhor descrita e analisada. Pretende-se que a metodologia proposta seja geral o suficiente para ser aplicada em outros problemas que envolvam a simulação numérica de reservatórios. Por este motivo, é importante que um nível de abstração seja inserido na metodologia de forma que a Busca Dispersa enxergue o problema com um conjunto de pontos de um espaço n -dimensional, cada ponto associado a um valor real resultante da avaliação de uma função objetivo, que precisa ser explorado em busca das melhores soluções disponíveis. Outro objetivo do trabalho é escolher um algoritmo de busca direta local que seja robusto (convergência garantida) e eficiente para ser acoplado à Busca Dispersa.

1.3 ESTRUTURA DA DISSERTAÇÃO

Esta dissertação está organizada em oito capítulos. O primeiro capítulo traz uma introdução geral sobre o contexto onde o problema do ajuste de histórico se insere, apresenta a motivação da realização deste trabalho e também apresenta os principais objetivos a serem alcançados.

Uma série de tópicos e conceitos que são importantes para entender as metodologias sobre as quais este trabalho se sustenta encontram-se detalhadas no segundo capítulo juntamente com as escolhas feitas durante a realização do trabalho.

No terceiro capítulo, um panorama é apresentado sobre quais são as origens das técnicas e metodologias usadas na dissertação e também o trabalho realizado é localizado no universo das iniciativas de ajuste de histórico.

O quarto capítulo é dedicado ao detalhamento da metodologia proposta, isto é, são apresentadas as contribuições do autor para o problema do ajuste de histórico e os passos seguidos na direção de demonstrar a aplicabilidade da Busca Dispersa neste contexto são descritas.

No quinto capítulo, os modelos de reservatório utilizados na geração dos resultados e análises são apresentados.

O sexto capítulo contém uma análise dos resultados obtidos com a metodologia proposta para cada um dos modelos apresentados no capítulo cinco.

As conclusões e discussões finais encontram-se no sétimo capítulo.

Capítulo 2 – FUNDAMENTAÇÃO TEÓRICA

Neste capítulo, diversos conceitos, metodologias e técnicas pertinentes ao trabalho desenvolvido serão apresentados e discutidos. Sua integração na composição de uma metodologia específica para a resolução do problema de ajuste de histórico será apresentada no capítulo sobre metodologia.

2.1 CICLO DE VIDA DE UM RESERVATÓRIO

Como todo projeto de exploração de recursos minerais (não renováveis), os reservatórios de petróleo têm um ciclo de vida com fases bem definidas (Jahn *et al.*, 1998). São elas:

1. Exploração.
2. Avaliação.
3. Desenvolvimento.
4. Produção.
5. Abandono.

2.1.1 Exploração

Esta fase é caracterizada por um elevado grau de incertezas sobre a região alvo. Diversas técnicas de sondagem geológica são empregadas na tentativa de caracterizar a bacia sedimentar e quantificar a probabilidade de acumulação de hidrocarbonetos.

Ainda que os estudos geológicos indiquem que a probabilidade de ocorrência de hidrocarbonetos no local seja alta, outras variáveis afetam a decisão de prosseguir ou não com o projeto. Entre elas, destacam-se: dificuldades técnicas de acesso ao potencial reservatório, existência de infra-estrutura para o escoamento da produção, proximidade dos mercados consumidores além do obstáculo formado pelos riscos devido às instabilidades políticas, econômicas e fiscais na locação.

As empresas de Exploração e Produção (E&P) normalmente possuem um conjunto de reservatórios em fase de exploração compondo uma carteira (ou portfólio) de projetos. Os investimentos de exploração são motivados pela necessidade que as empresas do setor têm de manter uma boa razão reserva/produção. Este indicador tem sido usado na indústria para avaliar a saúde financeira destas empresas uma vez que é preciso sempre descobrir novas reservas de modo a compensar a produção declinante dos campos mais maduros. A fase de exploração é finalizada quando um poço exploratório confirma a presença de hidrocarbonetos.

2.1.2 Avaliação

Uma vez que hidrocarbonetos foram encontrados por um poço exploratório, a empresa deve avaliar o potencial econômico da descoberta. Atividades típicas desta fase de vida de um reservatório incluem a perfuração de novos poços para tentar determinar a extensão do campo e a execução de testes de produção para avaliar as condições de produção do reservatório. O objetivo principal desta fase é a redução das incertezas sobre o reservatório de forma que seja possível elaborar um estudo de viabilidade contendo diversas opções técnicas para o desenvolvimento do campo.

2.1.3 Desenvolvimento

O foco desta fase é na confecção e execução de um plano de desenvolvimento para o campo. O plano de desenvolvimento segue as orientações do estudo de viabilidade selecionando as técnicas cujas implementações sejam economicamente viáveis. Os principais objetivos são: minimizar os investimentos e os riscos e maximizar a produção e os lucros.

Os planos de desenvolvimento normalmente incluem diferentes abordagens para o desenvolvimento do campo, um para cada cenário vislumbrado pela equipe de avaliação. Algumas abordagens podem compartilhar certas etapas de desenvolvimento de forma a inserir flexibilidade na adoção de uma ou outra estratégia de produção à medida que as incertezas forem sendo mitigadas pela presença de novos dados.

Esta é para as empresas a fase mais dispendiosa em termos de investimentos, pois todos os recursos presentes no plano de desenvolvimento são adquiridos e/ou fabricados e os contratos de serviços e locação são lavrados.

2.1.4 Produção

A fase de produção tem início quando uma quantidade comercial de hidrocarbonetos começa a fluir do sistema de produção (primeiro óleo). É nesta fase que a empresa começa a recuperar o seu investimento.

Esta fase pode ser subdividida em 3 fases de acordo com o ritmo da produção:

1. Ascensão. Neste período os volumes produzidos aumentam na medida em que novos poços entram em operação.
2. Platô. Nesta etapa novos poços podem entrar em operação, mas alguns poços já se encontram em declínio. A produção mantém-se num mesmo patamar durante este período.
3. Declínio. Nesta etapa, todos os poços apresentam declínio de produção. Esta é a etapa mais longa da fase de produção e persiste enquanto o projeto estiver gerando lucros para a companhia.

Na fase de produção o processo de ajuste de histórico é mais importante devido às incertezas no processo de caracterização e desenvolvimento dos campos os modelos numéricos construídos para representar os reservatórios e gerenciar a produção não reproduzem fielmente o modelo real e , por isso, devem ser calibrados.

2.1.5 Abandono

Esta fase tem início quando o fluxo de caixa do empreendimento torna-se negativo. Nesta fase, a infra-estrutura é removida ou reaproveitada e os poços são selados para evitar vazamentos que possam agredir o meio ambiente.

2.2 SIMULAÇÃO NUMÉRICA DE RESERVATÓRIOS

A simulação numérica é um procedimento fundamental às tarefas de desenvolvimento e gerenciamento de reservatórios porque ela permite estudar o comportamento de reservatórios no nível da produção e injeção individual dos poços. Segundo Fanchi, 2001, “o objetivo primário de um estudo de gerenciamento de reservatórios é determinar as condições ótimas necessárias para que a recuperação de hidrocarbonetos seja maximizada em um campo operado com prudência”.

A prudência neste contexto significa conciliar dois objetivos conflitantes na exploração de recursos naturais não renováveis: (1) maximizar a recuperação e (2) minimizar as despesas (investimentos e custos operacionais).

A obtenção de um modelo numérico de simulação é uma tarefa complexa, envolvendo uma equipe multidisciplinar e aborda diversas fases da vida útil do reservatório. O ponto de partida para a geração do modelo numérico ocorre quando um poço pioneiro apresenta indícios de hidrocarbonetos e um posterior teste de formação indica a viabilidade econômica da reserva encontrada. Com a perfuração de novos poços, mais dados se fazem presentes sendo possível caracterizar a formação através de análises de litologia (tipos das rochas), perfis elétricos, análises de calha e testemunhos, interpretações de dados sísmicos entre outras informações adquiridas no processo. Com estas informações, mapas e seções estruturais são elaborados pela correlação entre as informações de diferentes poços que geram, através de métodos geoestatísticos, os mapas (propriedades da rocha, propriedades do fluido e propriedades da interação rocha-fluido) que alimentam os simuladores numéricos (Thomas *et al.*, 2001).

Todo o processo de caracterização é cercado de incertezas de modo que os resultados de algumas etapas são largamente dependentes de interpretações subjetivas dos dados disponíveis. Tudo isso contribui para um elevado nível de incerteza nos modelos numéricos gerados no processo de caracterização. Neste contexto, a calibração dos modelos numéricos através do processo de ajuste de histórico pode ser visto como uma forma reversa de diminuir as incertezas sobre as características do reservatório.

Existem diversos tipos de simuladores de reservatórios de acordo com as premissas adotadas no modelo numérico subjacente. Entre eles destacam-se, quanto ao tratamento dos fluidos:

- Black-Oil – Neste modelo estão representados apenas dois tipos de hidrocarbonetos, óleo e gás, e a componente água. Outras premissas são: modelo isotérmico e a não alteração da composição destes hidrocarbonetos ao longo da simulação, ficando a solução do gás no óleo e das propriedades do fluido determinadas pela pressão.

- Composicional – No modelo composicional, existem um conjunto de equações para os componentes hidrocarbonetos existentes que podem ser agrupados por similaridade em pseudo-componentes. O modelo também é isotérmico e as propriedades são determinadas por cálculos Termodinâmicos.
- Térmico – Em simuladores térmicos, os gradientes de temperatura dentro do reservatório são importantes e afetam as propriedades, principalmente a mobilidade do óleo em seu interior. Estes simuladores são necessários quando processos térmicos de recuperação avançada são utilizados, como a injeção de vapor no reservatório.

A técnica de discretização mais utilizada nos simuladores comerciais é a de diferenças finitas, onde as equações diferenciais que modelam o escoamento de fluidos em meios porosos são resolvidas através da discretização e solução numérica. Outras técnicas, como elementos finitos, podem ser utilizadas. Entre elas, tem se destacado recentemente a simulação por linhas de fluxo, onde o problema 3D é transformado em uma série de problemas 1D, cada um resolvendo o problema de deslocamento de fluidos através de uma linha de fluxo. A forma e posição das linhas de fluxo são (re)calculadas ao longo da simulação de acordo com o campo de pressão no reservatório. O objetivo deste simulador é ser mais eficiente que aqueles baseados em métodos de elementos finitos, porém, seu desempenho pode ser prejudicado em casos onde os fluidos do reservatório são predominantemente compressíveis de modo que o campo de pressão precisa ser recalculado com maior frequência (Maschio *et al.*, 2002).

A escolha do tipo de simulador a ser usado num estudo de simulação depende das premissas adotadas para o reservatório e também do objetivo do estudo. O simulador de fluxo usado neste trabalho foi o IMEX da CMG, um simulador que utiliza o modelo Black-Oil.

2.2.1 A Revolução da Computação e a Simulação Numérica

Em 1965, o engenheiro eletrônico Gordon E. Moore, co-fundador da *Intel Corporation*, observou que o número de transistores em circuitos integrados dobrava a cada 18 meses aproximadamente (Moore, 1965). Esta taxa de crescimento exponencial, que já foi responsável por diversas revoluções no universo da informática, entre elas, a criação do computador pessoal

nos anos 80, é conhecida hoje como a “Lei de Moore”. O interessante é que ela já era válida em 1965 e permanece válida ainda hoje, ou seja, já se convive com mais de 40 anos de crescimento exponencial no poder de processamento disponível e não há uma previsão de quando esta “lei” deixará de ser válida.

A evolução dos simuladores numéricos de fluxo acompanhou de perto a evolução nos microprocessadores que alimentam a indústria de informática, tanto que, hoje em dia eles são bem aceitos como ferramenta para auxiliar a tomada de decisão na área de E&P. Os problemas associados a estudos baseados em simulação normalmente são devido ao mau uso da ferramenta.

Por exemplo, os simuladores Black-Oil exigem algumas adaptações na transformação do modelo geológico para um modelo de simulação. Uma das mais óbvias é a transferência de escala de um modelo contendo milhões de blocos para outro contendo milhares deles. É claro que há perdas de qualidade durante este processo. Malhas muito refinadas geram soluções mais precisas, porém aumentam bastante o tempo de simulação. Por outro lado, as malhas grosseiras permitem simulações mais rápidas, mas podem fazer com que certos fenômenos que ocorrem no reservatório real não apareçam na simulação, especialmente nas imediações dos poços, onde ocorrem as maiores variações de pressão. O engenheiro de reservatórios deve estar ciente do impacto de cada simplificação na qualidade dos resultados das simulações desempenhadas. Por outro lado, embora os resultados de uma única simulação possam ser discutíveis, a análise comparativa de simulações fornece tendências que se verificam na prática (Mattax *et al.*, 1990).

Uma importante faceta das simulações de reservatórios é que é possível, a partir dos dados de produção/injeção, construir fluxos de caixa detalhados sobre o projeto. Desta forma, diferentes cenários podem ser comparados entre si através de análises econômicas ou simplesmente de produção. Hoje em dia os modelos de simulação estão tão sofisticados que a qualidade dos dados de entrada e a aplicabilidade do tipo de modelo de simulação desenvolvido pelo simulador são mais questionadas que a correção ou qualidade do simulador utilizado. Não obstante, este é um item que deve ser olhado com cuidado, pois os modelos numéricos também podem ser fontes importantes de erros na modelagem de campos de petróleo.

2.2.2 Ajuste de Histórico

O ajuste de histórico é um processo iterativo aplicado em modelos numéricos de reservatórios onde os parâmetros de entrada são consistentemente alterados com o objetivo de aproximar os resultados simulados com aqueles observados no campo real. A expectativa é que, se um dado modelo é capaz de reproduzir o desempenho passado de um reservatório, as previsões de comportamento fornecidas por ele deverão possuir melhor qualidade que as previsões de modelos desajustados.

Trata-se de um problema inverso onde o número de parâmetros de entrada costuma ser grande, os parâmetros possuem interdependências e o tempo de simulação é fortemente dependente do tamanho da entrada (número de blocos) e do horizonte de tempo escolhido, que normalmente varia de poucos anos a poucas décadas.

Outra característica importante do problema de ajuste de histórico é que não existe uma única solução para o problema, pois diferentes modelos de simulação podem aproximar os dados de histórico com diferentes precisões. Esta é uma característica herdada da modelagem numérica de reservatórios de petróleo, que permitem diversas formulações para um mesmo reservatório real. É importante citar que diferentes modelos usados para reproduzir históricos de produção de forma semelhante resultam em previsões de produção diferentes; por isso, a consideração de múltiplas soluções é importante para verificar incertezas na previsão de desempenho de campos.

O ajuste de histórico é uma das tarefas mais demoradas durante um estudo de simulação. Ela tem início após a etapa de caracterização do reservatório e a geração do(s) modelo(s) de simulação (Figura 2.1). O objetivo primário do ajuste de histórico é garantir que os modelos de simulação gerados produzam previsões satisfatórias sobre o desempenho futuro do reservatório.

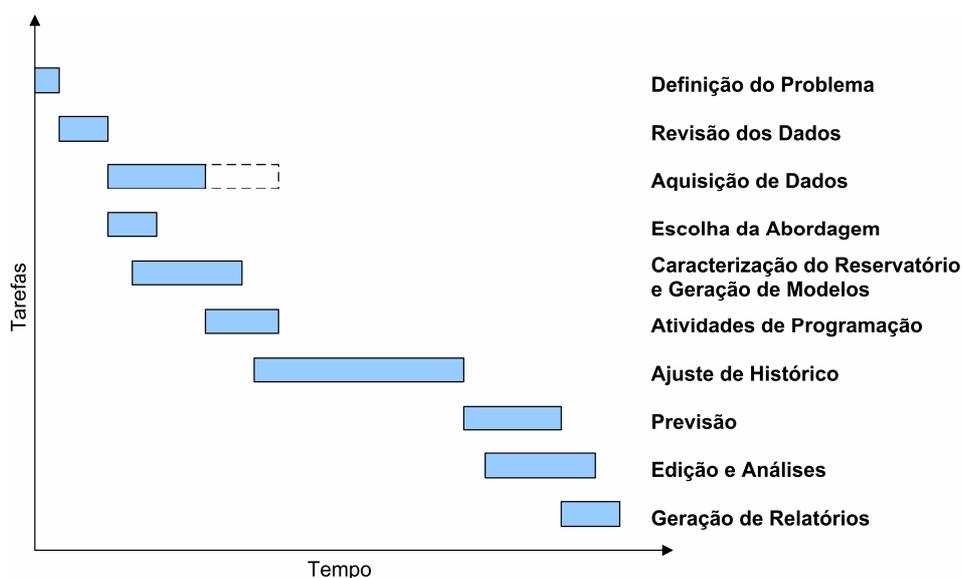


Figura 2.1 – Atividades típicas em um estudo de simulação – fonte: Mattax *et al.*, 1990.

O tempo de estudo tende a ser um dos recursos mais escassos em um projeto de ajuste de histórico, por isso, alguns cuidados devem ser tomados no sentido de economizá-lo:

- O modelo numérico deve ser dimensionado para fazer um balanço entre a qualidade dos resultados obtidos e o tempo total gasto por simulação. Quanto maior o tamanho de um modelo (medido em número de blocos), maior o tempo de simulação, por outro lado, quanto menor for esse modelo, maior será a chance de o modelo não capturar certos efeitos observados no reservatório.
- O número de parâmetros utilizados no ajuste tende a aumentar de forma exponencial o número de simulações requeridas para executar um estudo compreensivo. Fazer uma análise de sensibilidade nos parâmetros de ajuste pode ajudar a eliminar parâmetros pouco influentes e assim diminuir o número de simulações.

O processo de aquisição, interpretação e manipulação dos dados para a geração dos primeiros modelos numéricos é cercado de imprecisões, erros e interpretações subjetivas de modo que precisam ser ajustados para oferecer bons resultados na prática. Uma das formas de testar se o modelo de simulação está coerente com o reservatório real é contrastar os desempenhos do reservatório real e do modelo numérico. Por exemplo, considere o gráfico da Figura 2.2. A curva de pontos vermelhos mostra os dados do histórico de corte de água (*water cut*) para um

determinado poço. A curva sólida azul mostra os dados de corte de água gerados pelo mesmo poço em um modelo numérico desajustado. Finalmente, a curva sólida verde representa os dados de corte de água obtidos a partir de um modelo ajustado.

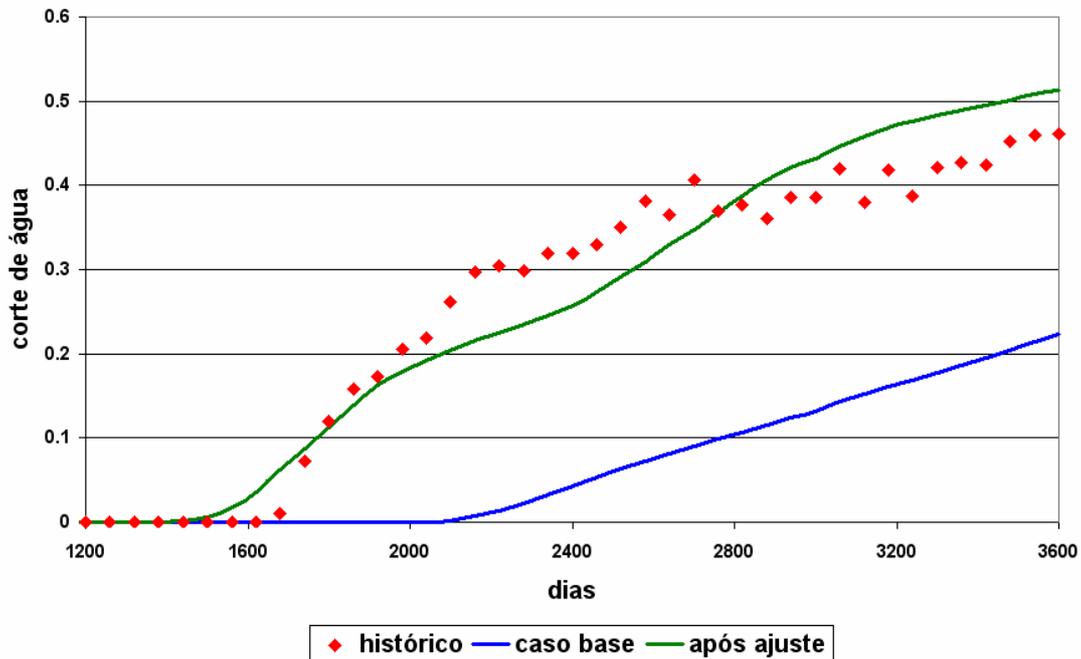


Figura 2.2 - Exemplo contendo caso (modelo) base e ajustado.

2.2.3 Os Parâmetros de Ajuste

As variáveis do processo são os valores dos parâmetros do modelo numérico. O ponto de partida do ajuste é o modelo base, onde todas as propriedades do reservatório estão fixas nos valores mais prováveis; um resultado que vem da etapa de caracterização do reservatório. Uma das formas de medir a qualidade do modelo base é comparando os dados obtidos na simulação com os valores obtidos no campo real. Os valores que esses parâmetros assumem no modelo inicial, chamado de caso base, são originados do processo de caracterização do reservatório, que antecede o ajuste (ver cronograma de atividades da Figura 2.1). Como a caracterização é marcada por um elevado grau de incertezas, insuficiência de dados entre outras dificuldades, os valores dos parâmetros de reservatório nos modelos gerados não devem ser encarados como valores determinísticos, mas sim como uma estimativa para os valores mais prováveis dada a discretização assumida na geração da malha. Assim, cada parâmetro em cada bloco de simulação

pode ser enxergado como uma distribuição de valores em torno do valor mais provável. A Figura 2.3 mostra um exemplo de qual é a distribuição de probabilidades para os valores dos parâmetros porosidade da rocha e permeabilidade horizontal em um dado reservatório.

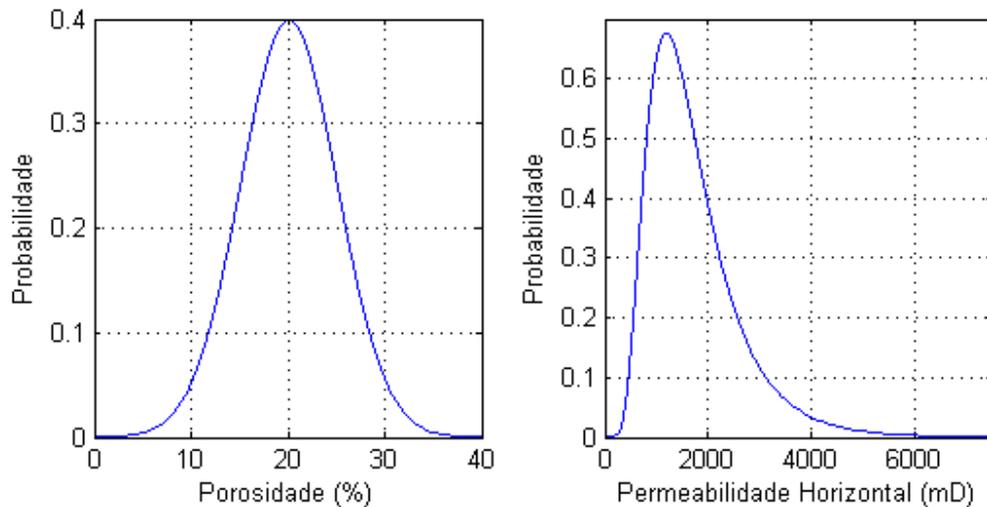


Figura 2.3 - Distribuições de probabilidade de parâmetros do reservatório.

Neste contexto de modelar cada variável do problema como uma variável incerta, com média, desvio padrão e distribuição de probabilidades, o caso base de um ajuste de histórico pode ser encarado como um caso onde as variáveis foram configuradas para suas respectivas médias. Alguns dados típicos considerados durante o processo de ajuste são: vazões de água e gás, corte de água, pressão de fundo dos poços e pressão média do reservatório.

2.2.4 Tipos de Ajuste

Existem diversas abordagens para o ajuste de histórico. Estas podem ser classificadas em três níveis quanto ao grau de automação do processo: (1) ajuste manual, (2) ajuste automático e (3) ajuste assistido. Nas seções a seguir, por simplicidade de exposição, o problema do ajuste será tratado como se fosse aplicado a apenas um modelo de simulação, muito embora a etapa de caracterização possa gerar vários modelos de simulação. Subentende-se que os processos descritos a seguir podem ser aplicados a vários modelos durante a fase de ajuste de histórico de um estudo de simulação.

2.2.4.1 Ajuste de Histórico Manual

O processo manual de ajuste é o mais tradicional e comum na indústria de E&P. Ele consiste em um processo iterativo onde um determinado modelo de simulação é analisado pelo engenheiro de reservatórios, as discrepâncias dos dados simulados e os dados de histórico são avaliados qualitativamente e as decisões sobre quais parâmetros do reservatório serão modificados e de que forma estas modificações se darão são tomadas. O sucesso deste empreendimento é largamente dependente da experiência do engenheiro de reservatórios e de seu nível de organização. Um dos objetivos perseguidos neste caso, além da qualidade dos resultados, é a minimização do número de simulações, pois elas podem ser muito demoradas. As dificuldades associadas a um ajuste manual incluem:

- Grande esforço em tarefas manuais.
- Grande volume de dados que precisam ser analisados em cada iteração do processo.
- A maior parte do tempo dedicado ao ajuste é utilizada na geração de dados intermediários como planilhas e gráficos que permitem avaliar o ajuste.

2.2.4.2 Ajuste de Histórico Automático

Um processo de ajuste totalmente automático consiste em fornecer um ou mais modelos de simulação a um programa de computador juntamente com os dados de histórico que precisam ser ajustados e, ao término da execução, tem-se um conjunto de modelos ajustados como resultado. Pode-se afirmar que este é um objetivo almejado por todos os profissionais envolvidos em atividades de ajuste de histórico, porém, trata-se de uma impossibilidade prática devido ao grande número de possibilidades que se apresentam durante o ajuste.

O maior desafio no desenvolvimento de um processo de ajuste automático é encontrar uma parametrização que leve ao ajuste, ou seja, encontrar um conjunto de parâmetros que, quando modificados dentro de seus devidos intervalos de confiança, levam ao ajuste. Um ajuste verdadeiramente automático teria que, em tese, lidar com todas as variáveis do processo. Uma forma de atacar este problema seria tratar cada parâmetro de reservatório em um bloco como uma variável do problema. Neste contexto, o número de variáveis do problema seria o número de blocos presente na malha multiplicado pelo número de parâmetros de ajuste em cada bloco. É

importante observar que modelos típicos de simulação contêm dezenas ou centenas de milhares de blocos. Este fato isoladamente já estabelece um número proibitivo de variáveis para as técnicas de otimização conhecidas.

Uma forma de reduzir a complexidade destes casos seria agrupar conjuntos de blocos em regiões nas quais as variáveis de otimização estariam definidas, mas não existe uma metodologia de agrupamento que funcione para todos os casos. Adicionalmente, a redução do número de variáveis com esta abordagem ainda é insuficiente para gerar problemas de otimização computacionalmente viáveis.

Em resumo, com a tecnologia atual, não é viável incluir a parametrização no problema de otimização representado pelo ajuste de histórico. Ainda que a tecnologia avance a ponto de viabilizar problemas de otimização não tratáveis hoje, os modelos de simulação não são estanques e também evoluem na direção de aproveitar a capacidade computacional adicional de forma a melhorar a precisão da simulação.

Outra consideração a respeito dos processos de ajuste automáticos é que, se não forem bem desenvolvidos, eles tendem a gerar soluções que não seguem a caracterização geológica dos reservatórios e, em geral, resultam em um esforço computacional excessivo.

2.2.4.3 Ajuste de Histórico Assistido

O ajuste de histórico assistido é uma combinação entre o ajuste de histórico manual e o automático, uma vez que as melhores características de cada metodologia são agrupadas. De um lado a experiência do engenheiro de reservatórios é usada para limitar o número de variáveis mais importantes e escolher os objetivos de cada etapa do ajuste, do outro, técnicas automatizadas são empregadas para avaliar rapidamente a qualidade de cada tentativa de ajuste. A maior parte dos processos de ajuste citados na literatura se enquadram nesta categoria.

Um ponto comum de todos os métodos desta categoria é o uso de funções objetivo que medem a diferença entre os dados simulados e observados de forma a avaliar quantitativamente a qualidade do ajuste (Kabir *et al.*, 2003). A escolha dos termos e pesos que comporão a função objetivo em um ajuste automatizado não é tarefa trivial. É comum que funções objetivo mal projetadas levem o ajuste para regiões inesperadas do espaço de soluções. O correto

estabelecimento da função objetivo é especialmente importante quando um ajuste local está em andamento.

Uma clara vantagem dos procedimentos de ajuste assistido é que o engenheiro de reservatórios consegue utilizar muito mais informações sobre o reservatório em estudo, e sua própria experiência, em suas decisões de ajuste do que qualquer método de otimização poderia. Esta habilidade do engenheiro normalmente se traduz em um número menor de soluções a pesquisar (imposição de restrições no espaço de soluções) e uma maior eficiência na pesquisa do espaço de soluções. O resultado final é que há uma maior chance de um processo assistido gerar resultados melhores que os dos processos automáticos.

Finalmente, uma importante característica da adoção de processos de ajuste assistidos é o aumento na produtividade do engenheiro de reservatórios, na medida em que eles o livram de tarefas repetitivas como a geração de casos de simulação derivados e a avaliação da qualidade do ajuste obtido. É objeto deste trabalho a proposta de um método de ajuste assistido com foco na metodologia de otimização Busca Dispersa.

2.2.5 Esquema Geral de um Processo de Ajuste

Um processo típico de ajuste de histórico consiste em estabelecer um conjunto de parâmetros incertos, ou seja, um conjunto de parâmetros que será alvo de alteração no processo, definir uma função objetivo de forma a avaliar numericamente a qualidade dos diferentes modelos gerados, executar uma análise de sensibilidade para limitar o número de parâmetros que efetivamente irá entrar no ajuste e também obter informações sobre a intensidade e direção das modificações nestes parâmetros e, finalmente, ocorre o processo de ajuste propriamente dito (veja diagrama da Figura 2.4). Neste esquema, supõe-se que a modelagem do reservatório esteja sendo feita de forma adequada, embora em alguns casos seja necessário rever a própria concepção do modelo ou até a formulação empregada.

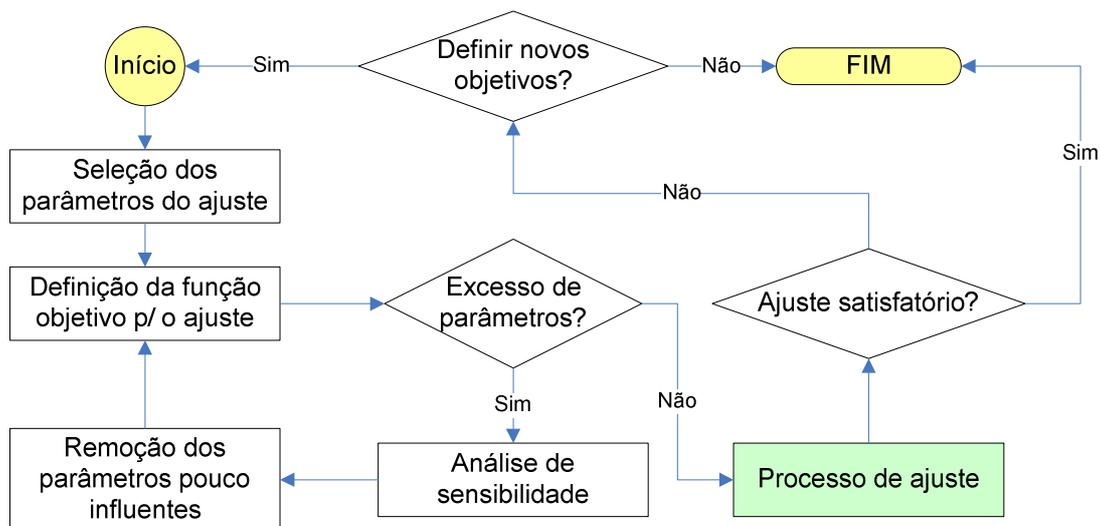


Figura 2.4 - Algoritmo básico de um ajuste de histórico

Um processo de ajuste de histórico pode levar a centenas ou mesmo milhares de simulações numéricas. Processos manuais de ajuste tendem a gerar menos simulações, mas seu sucesso é largamente dependente da experiência do engenheiro (ou equipe) realizando o ajuste. Processos assistidos pelo computador normalmente geram mais simulações que processos manuais, porém possuem diversas vantagens. Entre elas estão:

- A economia de tempo gerada pela automação de procedimentos manuais como a geração de novos modelos modificados ajuda a diminuir o tempo total gasto no processo.
- O uso de computação distribuída pelo algoritmo de ajuste é outro fator que pode reduzir o tempo total do ajuste.
- Após seu início, os algoritmos não exigem a presença do operador durante seu processamento podendo ser realizados fora do horário normal de trabalho.

As vantagens citadas acima não eliminam o custo de ocupação dos recursos computacionais que poderiam estar dedicados a outros processos na empresa. Por este motivo, é fundamental que o número total de simulações desempenhadas pelo processo assistido seja reduzido ao máximo. É isto que torna a análise de sensibilidade uma etapa importante do ajuste de histórico. Na análise de sensibilidade, busca-se encontrar os parâmetros que possuem maior impacto na avaliação da função objetivo escolhida. Parâmetros pouco influentes são candidatos a serem excluídos do processo. Isto é feito fixando seus valores nos valores mais prováveis enquanto o processo de

ajuste encontra os valores dos demais parâmetros que melhoram o valor da função objetivo global. Não realizar uma análise de sensibilidade prévia pode levar o processo assistido a realizar um número muito grande de simulações ou mesmo inviabilizar o ajuste.

2.3 OTIMIZAÇÃO LOCAL VERSUS OTIMIZAÇÃO GLOBAL

No campo da matemática e da pesquisa operacional, a otimização consiste em encontrar um valor mínimo ou máximo para uma dada função objetivo real sistematicamente escolhendo valores reais ou inteiros dentro de um conjunto de valores válidos (o domínio da função). Sem perda de generalidade, é possível descrever qualquer problema de otimização como um de minimização, pois qualquer problema de maximização pode ser transformado em um de minimização apenas invertendo o sinal da função objetivo. Assim, a descrição a seguir leva em conta apenas problemas de minimização.

Suponha que se pretende encontrar o valor de x que leva uma determinada função $f(x)$ a assumir seu valor mínimo dentro de um domínio especificado. O gráfico da Figura 2.5 mostra como se comporta a função $f(x) = x^2 - 4x + 2$ no domínio $0 \leq x \leq 4$.

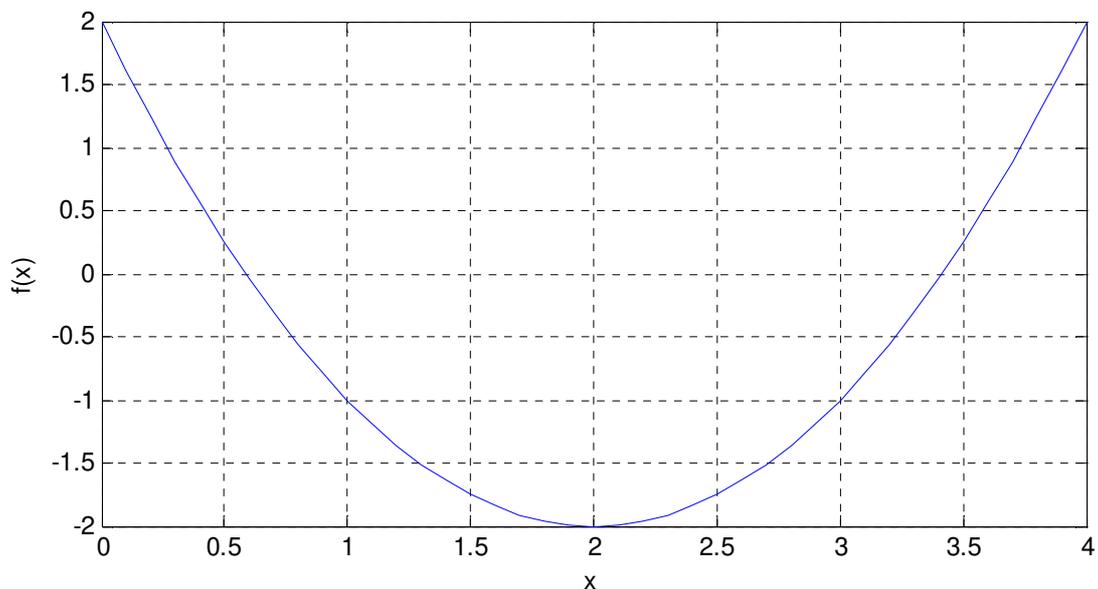


Figura 2.5 – Gráfico de $x^2 - 4x + 2$ no domínio $0 \leq x \leq 4$.

Observe que neste caso o mínimo é único e ocorre em $x = 2$. Como uma função analítica explícita está disponível, este problema de otimização pode ser resolvido analiticamente. Basta encontrar todos os pontos onde a derivada da função se anula, ou seja, resolver a equação $2x - 4 = 0$. Numericamente, no entanto, é necessário definir um ou mais pontos iniciais, dependendo do método adotado, e seguir uma direção de descida para o valor da função objetivo, possivelmente reavaliando esta direção ao longo do caminho, até que não se encontre mais direções de descidas a partir do ponto atual.

Para superfícies com apenas um mínimo local, como a do exemplo da Figura 2.5, encontrar um mínimo local corresponde a encontrar o mínimo global do problema, porém, superfícies mais complexas tipicamente apresentam mais que um mínimo local, logo, o mínimo encontrado pode não ser o mínimo global. Tome, por exemplo, a função $f(x) = \text{sen}(x) + 5/x$ no domínio $-10 \leq x \leq 10$. O gráfico desta função pode ser visto na Figura 2.6. Observe que existem 4 mínimos locais nesta função ($x = -8.055$, $x = -1.772$, $x = 4.511$ e $x = 10$), porém apenas um deles é o mínimo global ($x = -8.055$).

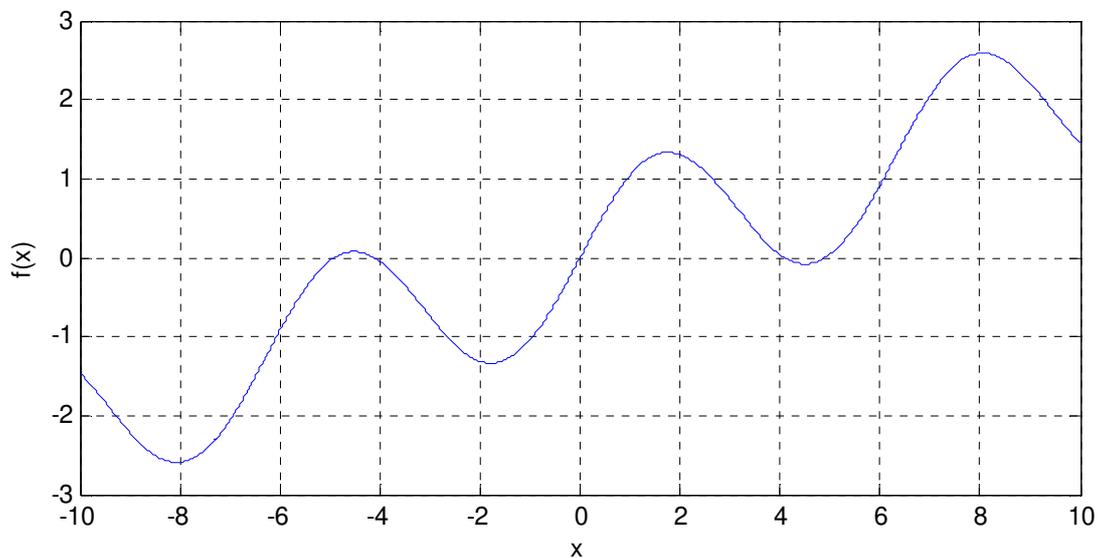


Figura 2.6 – Gráfico da função $\text{sen}(x) + x/5$ no domínio $-10 \leq x \leq 10$.

Algoritmos de busca local são eficientes na busca de valores mínimos locais, mas qual mínimo local será encontrado depende do ponto de partida e da topologia da superfície da função objetivo. Por exemplo, ao usar o Solver do Excel para encontrar os mínimos locais deste problema, usar o ponto de partida $x = -6$, a solução ótima do problema é encontrada

($x = -8.055$), porém iniciando o mesmo algoritmo a partir do ponto $x = 9$, o pior mínimo local é encontrado ($x = 10$). É neste contexto que a busca local é dita míope, pois não consegue “enxergar” além do mínimo local mais próximo. Foi para superar estes problemas que algoritmos de busca global foram criados.

Os algoritmos de busca global não param, necessariamente, ao encontrar um mínimo local, mas usam recursos adicionais para continuar a busca além do ótimo local encontrado. Em sua forma mais simples, um algoritmo de busca global pode simplesmente eleger um conjunto de pontos de partida dentro do domínio da função e realizar uma busca local a partir de cada ponto eleito. Algoritmos mais robustos usam estruturas de memória para tentar obter informações que auxiliem no desenvolvimento de estratégias de intensificação (busca agressiva de valores otimizados) e diversificação (busca de regiões inexploradas do espaço de soluções) ou então usam outras estratégias que incluem a randomização ou mesmo a simulação de processos estocásticos.

2.4 OTIMIZAÇÃO BASEADA EM SIMULAÇÕES

Nos exemplos da Figura 2.5 e Figura 2.6, as funções objetivo, além de simples, poderiam ser resolvidas analiticamente, o que é preferível em problemas de otimização. Acontece que um número considerável de aplicações de otimização não possui uma função objetivo que possa ser expressa analiticamente em função dos dados de entrada, sendo necessário uma etapa adicional de simulação numérica para obter, a partir dos dados de entrada do problema, os parâmetros sendo avaliados pela função objetivo.

Essa junção das tecnologias de simulação com os métodos de otimização é uma característica nova e crescente no meio técnico e científico. Este tipo de problema é conhecido na literatura como problemas de otimização baseados em simulações (*simulation optimization*). O objetivo final da otimização baseada em simulações é identificar quais conjuntos de especificações de modelos, isto é, parâmetros de entrada e/ou premissas estruturais, levam a um desempenho ótimo do modelo de simulação.

2.4.1 Otimização Tradicional vs. Planejamento Estatístico

No universo do planejamento estatístico (ou experimental), esses parâmetros de entrada ou premissas estruturais são chamados *fatores*. As medidas de desempenho do modelo são chamadas *respostas*. Já no campo da otimização, os fatores se tornam *variáveis de decisão* e as respostas são usadas para modelar *funções objetivo* e *restrições* (Figura 2.7). Assim, enquanto que o objetivo no planejamento estatístico é encontrar quais fatores tem maior impacto na resposta, na otimização o objetivo é encontrar um conjunto de valores para as variáveis de decisão que minimizam (ou maximizam) a função objetivo sem violar nenhuma restrição imposta ao modelo (April *et al.*, 2003).

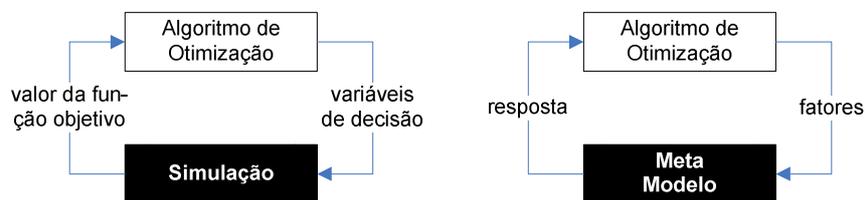


Figura 2.7 – Otimização Baseada em Simulações na otimização tradicional e em planejamento estatístico

O maior problema nos dois mundos é lidar com a explosão combinatorial de soluções possíveis ao se variar os valores assumidos pelas variáveis de decisão (ou fatores). Este grande número de possibilidades gera um espaço de soluções praticamente infinito² de forma que uma enumeração completa de todas as possibilidades é impraticável. É neste contexto que as técnicas de otimização, na forma de metaheurísticas, são utilizadas no processo com o objetivo de explorar de forma inteligente e econômica o espaço de soluções. Uma metaheurística é uma abordagem algorítmica que visa encontrar soluções ótimas aproximadas em problemas de otimização combinatoria. Essencialmente, um algoritmo de alto nível guia uma heurística interna na busca de soluções otimizadas.

No planejamento estatístico, é dado um passo adicional onde a simulação é substituída por um *metamodelo*, que nada mais é que uma expressão algébrica dos fatores de entrada que, quando calculada, aproxima a resposta que seria fornecida por um simulador convencional. O processo de construção do *metamodelo* faz uso de técnicas estatísticas fornecidas pela teoria de planejamento estatístico e implica na execução de uma série pré-determinada de simulações. O

² Se os intervalos especificados pelas variáveis de decisão (ou fatores) não forem discretizados, o espaço de soluções será verdadeiramente infinito e impossível de enumerar.

maior benefício desta técnica é que a avaliação do *metamodelo* tende a ser muito mais rápida que a simulação tradicional possibilitando a exploração de um maior número de soluções pelo algoritmo de otimização. Uma desvantagem é que um grande número de simulações precisa ser feito para obter um *metamodelo* cujas respostas sejam de qualidade. Uma observação importante, é que esta técnica pode ter problemas de precisão para casos complexos com não-linearidades acentuadas.

Neste trabalho, a técnica de planejamento estatístico não foi utilizada, um exemplo de utilização bem sucedida de metamodelos no contexto do ajuste de histórico pode ser visto no trabalho de Risso e Schiozer (2006). O foco da abordagem adotada foi o algoritmo de otimização. Este algoritmo usou como ferramenta de transformação de variáveis de decisão em valores reais para a função objetivo, um simulador de reservatórios *Black-Oil* acoplado a um módulo de cálculo que interpreta a saída do simulador e calcula a função objetivo especificada.

2.4.2 Métodos de Busca Direta

O uso de informações de gradiente tradicionalmente tem sido a força motriz das técnicas de otimização. O próprio conceito de otimalidade é muitas vezes associado à ocorrência de um gradiente nulo. Acontece que existe toda uma classe de problemas de otimização cuja função objetivo não pode ser expressa analiticamente em função dos parâmetros de entrada, ou então os parâmetros de entrada da função objetivo são valores gerados a partir de simulação numérica de modelos reais. Em situações como esta, a obtenção de informações sobre gradientes de primeira ordem ou de ordens superiores é dificultada ou mesmo impraticável. Os métodos de busca direta surgiram como ferramenta de otimização deste tipo de problemas onde os métodos clássicos de otimização falhavam.

Os métodos de busca direta têm suas raízes na década de 50, mas somente na década seguinte o termo busca direta foi sugerido por Hooke e Jeeves (1961) em um artigo publicado no *Journal of the ACM*. Naquela época a comunidade dos teóricos matemáticos envolvidos com otimização não defendiam a idéia por diversas razões, as principais sendo que esses métodos eram desenvolvidos empiricamente com métodos heurísticos, apresentavam convergência lenta e não possuíam estudos matemáticos sérios que comprovassem sua eficácia (Kolda *et al.*, 2003). O fato é que, mesmo naquela época, diversos problemas não tinham solução conhecida pelos

métodos tradicionais de otimização (baseados em informações de gradientes) e os métodos de busca direta eram os únicos que se propunham a resolvê-los.

Nos dias atuais, o avanço da tecnologia da informação permitiu que modelos cada vez mais complexos pudessem ser simulados numericamente o que deu vida nova aos métodos de busca direta que continuam amplamente utilizados em diversas áreas e não mostram sinais de estarem caindo em desuso (Lewis *et al.*, 2000).

O problema de ajuste de histórico é um bom exemplo de um problema de otimização cuja saída (dados de produção dos poços) não pode ser expresso analiticamente em função dos dados de entrada (descrição do reservatório). Por este motivo, um método de otimização de busca direta foi empregado para otimizar soluções fornecidas pela metodologia da Busca Dispersa proposta neste trabalho.

2.5 COMPLEXIDADE DE ALGORITMOS

Problemas de otimização são comuns a todas as áreas de engenharia. Acontece que alguns problemas são mais complicados que outros. Na ciência da computação, um problema é considerado resolvido quando um algoritmo eficiente é conhecido para sua solução. O conceito de eficiência é um tanto específico neste caso, significa que existe um polinômio definido em termos do tamanho da entrada³ que oferece um limite superior para o tempo de execução do algoritmo. É importante observar que o grau deste polinômio deve ser expresso por uma constante. Em outras palavras, um algoritmo é eficiente se ele consegue rodar qualquer instância do problema em tempo polinomial. Um algoritmo só é considerado eficiente pela comunidade de ciência de computação se houver uma prova matemática que o polinômio que determina sua complexidade é mesmo um limitante superior para o tempo de execução.

Acontece que muitos problemas, especialmente aqueles envolvendo matemática discreta, não são bem resolvidos na medida em que não possuem um algoritmo eficiente conhecido. Este tipo de problema consegue ser resolvido com algoritmos, porém, dependendo do tamanho da

³ O tamanho da entrada é uma métrica de quanto espaço em memória é necessário para armazenar os dados do problema. É usualmente medida em bits.

entrada, o caso ou leva um tempo de execução impraticável ou consome mais recursos (memória, por exemplo) do que as máquinas atuais conseguem endereçar.

A dificuldade de obtenção de algoritmos eficientes para determinados problemas motivou a classificação dos problemas em 2 tipos principais: P , que possuem algoritmos polinomiais, e NP , de não polinomiais, ou melhor, que não possuem solução polinomial conhecida. Dentre aqueles problemas do grupo NP , existem alguns problemas especiais que pertencem a um subconjunto dito NP -completo (Figura 2.8).

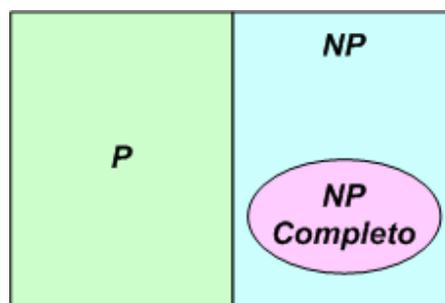


Figura 2.8 - Classificação dos problemas segundo sua complexidade

Os problemas dentro do grupo NP -completo possuem propriedades especiais, entre elas, podem ser citados:

- Os problemas são comprovadamente difíceis.
- Todo problema neste grupo pode ser reduzido (transformado) em qualquer outro a partir de transformações que podem ser realizadas em tempo polinomial.
- A resolução eficiente de qualquer problema deste grupo significa que todos os demais problemas também podem ser resolvidos em tempo polinomial.

Um dos grandes focos do estudo da complexidade de algoritmos é provar se P e NP representam o mesmo conjunto, ou seja, se sempre há uma solução eficiente para todo problema, ou não. Até o momento, persiste a idéia que estes conjuntos são disjuntos. O importante a ser assimilado desta seção é que alguns problemas não conseguem ser resolvidos de forma eficiente, mas isto não quer dizer que não existam algoritmos eficientes que dêem bons resultados para eles. Este é o universo dos algoritmos aproximados, ou das heurísticas (vide Seção 2.6). Para um melhor entendimento das questões relacionadas com a complexidade de algoritmos, é sugerida a leitura da referência Wilf, 1994.

2.6 HEURÍSTICAS E METAHEURÍSTICAS

Como foi citado na Seção 2.5, alguns problemas ainda não possuem algoritmos eficientes conhecidos. Para eles, encontrar a solução ótima tem um custo proibitivo para instâncias de tamanho prático. Para resolver estes problemas, algoritmos aproximados, chamados de heurísticas, foram desenvolvidos para encontrar, senão a solução ótima, soluções boas o suficiente para serem úteis em suas respectivas esferas de atuação. Uma característica dos algoritmos heurísticos é que não há uma prova matemática que a solução encontrada seja ótima, embora algumas heurísticas consigam garantir certa aproximação do ótimo (90% ou mais em relação ao ótimo, por exemplo). O fato é que as heurísticas têm sido utilizadas com sucesso nas aplicações cotidianas da indústria.

Um outro problema relacionado às heurísticas é que elas normalmente assumem que o espaço de soluções do problema é convexo e possui apenas um ponto otimizado que é o ótimo global, mas nem sempre este é o caso. Vários problemas de otimização em matemática discreta, onde uma função opera com valores discretos em suas variáveis independentes, possuem um espaço de solução não linear e não convexo com a existência de diversos ótimos locais. Nestes casos, as heurísticas tendem a ter sua performance prejudicada, pois costumam parar no primeiro ótimo local encontrado.

Várias técnicas têm sido utilizadas para incrementar as heurísticas de forma que elas possam ser adaptadas de métodos de busca local para métodos de busca global. A mais simples delas é simplesmente fornecer vários pontos de partida para a heurística na expectativa que cada rodada convirja para um ótimo local diferente. Neste caso, a resposta seria então a menor (ou maior) solução ótima local encontrada.

Nos últimos 20 anos, emergiu entre os algoritmos aproximados um conjunto de algoritmos denominados *metaheurísticas*. Este termo foi instituído por Fred Glover em seu trabalho sobre a Busca Tabu aplicada a problemas de programação inteira (Glover, 1986). Não existe um consenso sobre a definição do termo, porém as principais propriedades que caracterizam uma *metaheurística* estão resumidas em um trabalho de Blum e Roli (2003):

- Metaheurísticas são estratégias que “guiam” o processo de busca.

- O objetivo é explorar o espaço de soluções de forma eficiente de modo a buscar soluções (quase) ótimas.
- As técnicas que constituem algoritmos metaheurísticos vão de simples procedimentos de busca local até processos complexos de aprendizado.
- Algoritmos metaheurísticos são aproximados e, normalmente, não determinísticos.
- Eles podem incorporar mecanismos para evitar o confinamento do algoritmo a determinadas áreas do espaço de soluções.
- Os conceitos básicos das metaheurísticas permitem uma descrição em nível abstrato.
- As metaheurísticas não são específicas a nenhum problema em particular.
- As metaheurísticas podem fazer uso de conhecimentos específicos ao problema na forma de heurísticas que são controladas pela estratégia de nível superior.
- Atualmente as metaheurísticas mais avançadas usam experiência prévia (incorporada em algum tipo de memória) para guiar a busca.

Alguns exemplos mais conhecidos de metaheurísticas são:

- *Simulated Annealing* (SA) – Um método que usa a analogia entre a forma que um metal se resfria e se solidifica em uma estrutura cristalina de mínima energia e a busca de um mínimo em um sistema mais geral.
- *Algoritmos Genéticos* – Um conjunto de algoritmos populacionais onde cada indivíduo representa uma solução, factível ou infactível, onde, geração a geração busca-se obter indivíduos cada vez mais capazes (minimizados). As principais operações de geração de novos indivíduos envolvem o *crossover* (troca de genes) e mutação (mudança de uma solução para outra).
- *Busca Tabu* – Faz uso intensivo de memória, de curta e longa duração, tanto para escapar de mínimos locais quanto para explorar novas regiões do espaço de soluções. A memória de curta duração usa uma lista de movimentos proibidos (tabu) para tentar escapar de mínimos locais, daí seu nome.
- *Scatter Search ou Busca Dispersa* – Uma evolução da busca tabu, onde o conceito de população foi emprestado dos algoritmos genéticos. O algoritmo tem esse nome devido à forma com que explora o espaço de soluções. Inicia-se com um conjunto

de soluções relativamente espalhado no espaço e novos indivíduos são gerados combinando linearmente dois ou mais indivíduos da geração anterior.

Dois conceitos importantes em metaheurísticas são o de intensificação e diversificação. A intensificação é a tentativa de explorar o conhecimento adquirido pelo algoritmo no sentido de obter soluções melhoradas a partir dos dados existentes, enquanto que a diversificação é usada para direcionar o algoritmo para soluções novas, ou seja, para regiões do espaço de soluções que foram pouco exploradas. Apesar de possuir objetivos opostos, essas duas atividades são complementares no mundo das metaheurísticas. Cada uma das metaheurísticas atuais possui uma ou mais formas de ativar uma ou outra política no decorrer do algoritmo.

2.7 METAHEURÍSTICAS EM OTIMIZAÇÃO COMBINATORIAL

Uma otimização é dita do tipo combinatorial quando suas variáveis de decisão têm domínios discretos. Em termos matemáticos, uma *Otimização Combinatória* (OC) pode ser definida como (Blum *et al.*, 2003):

Definição 1 – Seja $P = (S, f)$ um problema de OC onde se tem:

um conjunto de variáveis de decisão $X = \{x_1, \dots, x_n\}$

um conjunto de domínios finitos para as variáveis D_1, \dots, D_n

um conjunto de restrições sobre as variáveis

uma função objetivo que será minimizada $f : D_1 \times \dots \times D_n \rightarrow \mathfrak{R}^+$

onde: $S = \{s = \{(x_1, v_1), \dots, (x_n, v_n)\} \mid v_i \in D_i, s \text{ satisfaz as restrições sobre as variáveis} \}$

Assim, S é chamado de espaço de soluções, visto que cada elemento do conjunto é um candidato à solução do problema. Para resolver o problema de OC, deve-se encontrar $s^* \in S$ de tal forma que $f(s^*) \leq f(s) \forall s \in S$. Neste caso, s^* é uma solução ótima global de $P = (S, f)$ e $S^* \subseteq S$ é o conjunto de soluções ótimas globais.

Os algoritmos que buscam resolver os problemas de OC dividem-se essencialmente em dois grupos: (1) completos e (2) aproximados. Os algoritmos completos dão a garantia de

obtenção da solução ótima em um tempo “razoável” desde que o espaço de soluções seja finito, porém seu desempenho é fortemente dependente da composição e complexidade do espaço de soluções. São exemplos de problemas de OC para os quais se conhece algoritmos completos eficientes: fluxos de custo mínimo em redes de transmissão, determinação de uma árvore espalhada mínima em grafos entre outros (Wolsey, 1998). Em diversos outros casos, estes algoritmos possuem tempo de execução exponencial e só são úteis para casos de pequeno porte. Por sua vez, os algoritmos aproximados sacrificam a garantia de obtenção da solução ótima global em benefício de encontrar boas soluções em um tempo razoável.

2.8 A METAHEURÍSTICA BUSCA DISPERSA

Nesta seção será apresentada uma versão geral, independente do problema sendo resolvido, da metaheurística *Busca Dispersa*. Esta seção foi baseada nos trabalhos Glover *et al.*, 2000 e Glover, 1998. Mais adiante, a aplicação da Busca Dispersa ao problema de ajuste de histórico será explicada.

2.8.1 O Conjunto de Referência

A Busca Dispersa foi inicialmente concebida como mais um mecanismo dentro do ferramental da Busca Tabu, mas nunca foi devidamente explorada pelos pesquisadores deste método. Foi somente após ter sido encarado como um método à parte que seus méritos foram devidamente reconhecidos. Diferentemente da Busca Tabu, este método trabalha sempre com uma população de soluções, aos moldes dos métodos evolutivos como a classe dos Algoritmos Genéticos. Essa população é conhecida na Busca Dispersa como *conjunto de referência*. Cada elemento deste conjunto é denominado *vetor solução*, ou simplesmente *vetor*, na medida em que uma solução pode ser representada como um vetor de valores para as variáveis de decisão do problema. O vetor solução na Busca Dispersa é equivalente ao indivíduo dos algoritmos genéticos.

A cada iteração, o conjunto de referência é atualizado com novos vetores gerados a partir daqueles previamente existentes no conjunto. As operações de “reprodução” neste contexto não seguem as existentes nos Algoritmos Genéticos onde há troca de “material genético” (operação

de *crossover*) segundo parâmetros aleatórios. Na Busca Dispersa, a geração de novos vetores se dá através da combinação linear (convexa ou não convexa) de dois ou mais vetores do conjunto de referência. Os novos vetores são então comparados, e os melhores, julgamento baseado em critérios de qualidade e diversidade, são incorporados ao conjunto de referência.

A Figura 2.9 mostra a evolução de um conjunto de referência em um espaço de soluções bidimensional. A Figura 2.9a mostra o conjunto de referência inicial contendo os vetores A , B e C . Em seguida, vários novos vetores são gerados pela combinação linear convexa e não convexa do par de vetores (A,B) . Destes, apenas o vetor 1 é selecionado para inserção no conjunto de referência. Nos demais diagramas, os vetores 2, 3, 4 e 5 são inseridos de forma análoga com a combinação dos pares de vetores (B,C) , $(1,C)$, $(1,2)$ e $(3,A)$, respectivamente. Uma observação importante é que a Busca Dispersa não adiciona vetores ao conjunto de referência em sua forma original, mas aplica uma heurística de melhoria local de modo a inserir versões refinadas dos vetores neste conjunto.

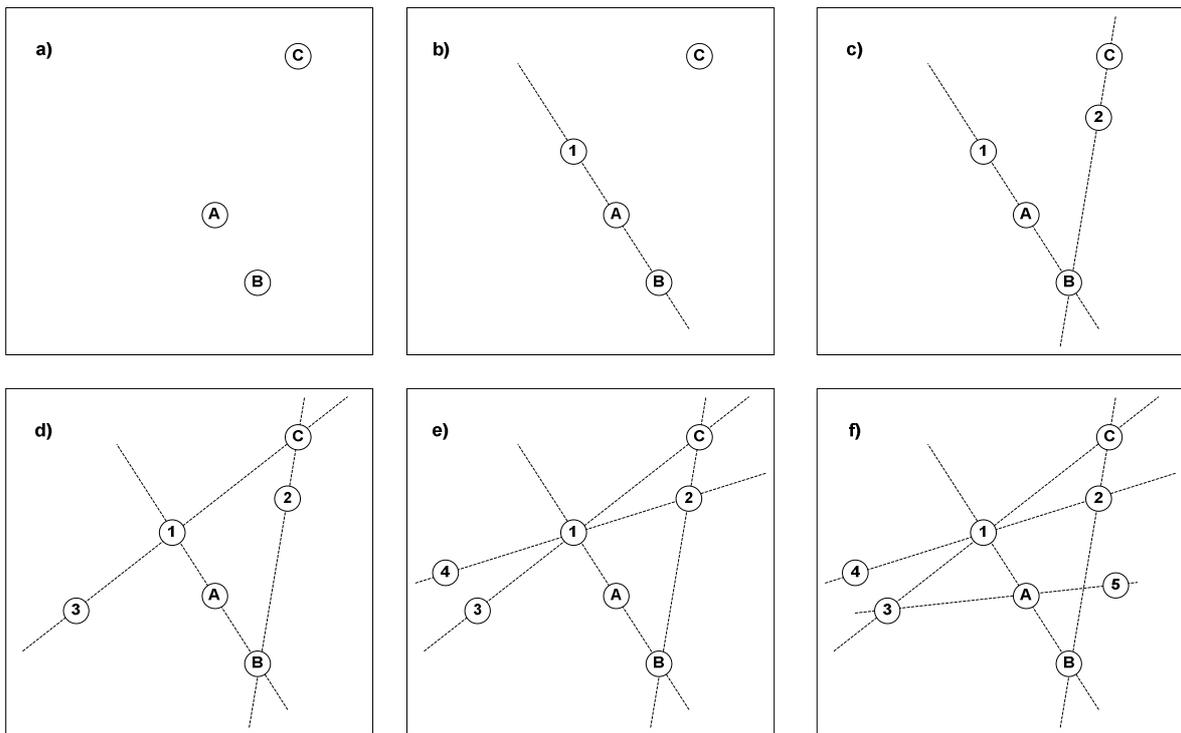


Figura 2.9 - Inserção de vetores no conjunto de referência através de combinações lineares

Ao contrário do que o exemplo da Figura 2.9 possa sugerir, o conjunto de referência não cresce sem limites. Enquanto que as populações dos Algoritmos Genéticos são tipicamente da

ordem de 100 ou mais indivíduos, na Busca Dispersa o tamanho do conjunto de referência é de 20 vetores ou menos. O objetivo de manter o número de vetores baixo é para que o número de subconjuntos de dois ou mais elementos se mantenha baixo, reduzindo as opções de combinação de vetores na criação de novos vetores candidatos a entrar no conjunto de referência.

2.8.2 Um Modelo Para a Busca Dispersa

Neste momento é possível delinear um modelo geral para as implementações da Busca Dispersa (ver fluxograma da Figura 2.10). Essencialmente um conjunto de referência inicial é gerado e, a partir daí, alguns ciclos de geração de novos vetores com a correspondente atualização do conjunto de referência são executados até que o conjunto de referência não mais se modifique. Quando isto ocorre, há a possibilidade de geração de um novo conjunto de referência reiniciando todo o processo. Para esta apresentação introdutória, é considerado como único critério de parada o número de iterações do algoritmo.

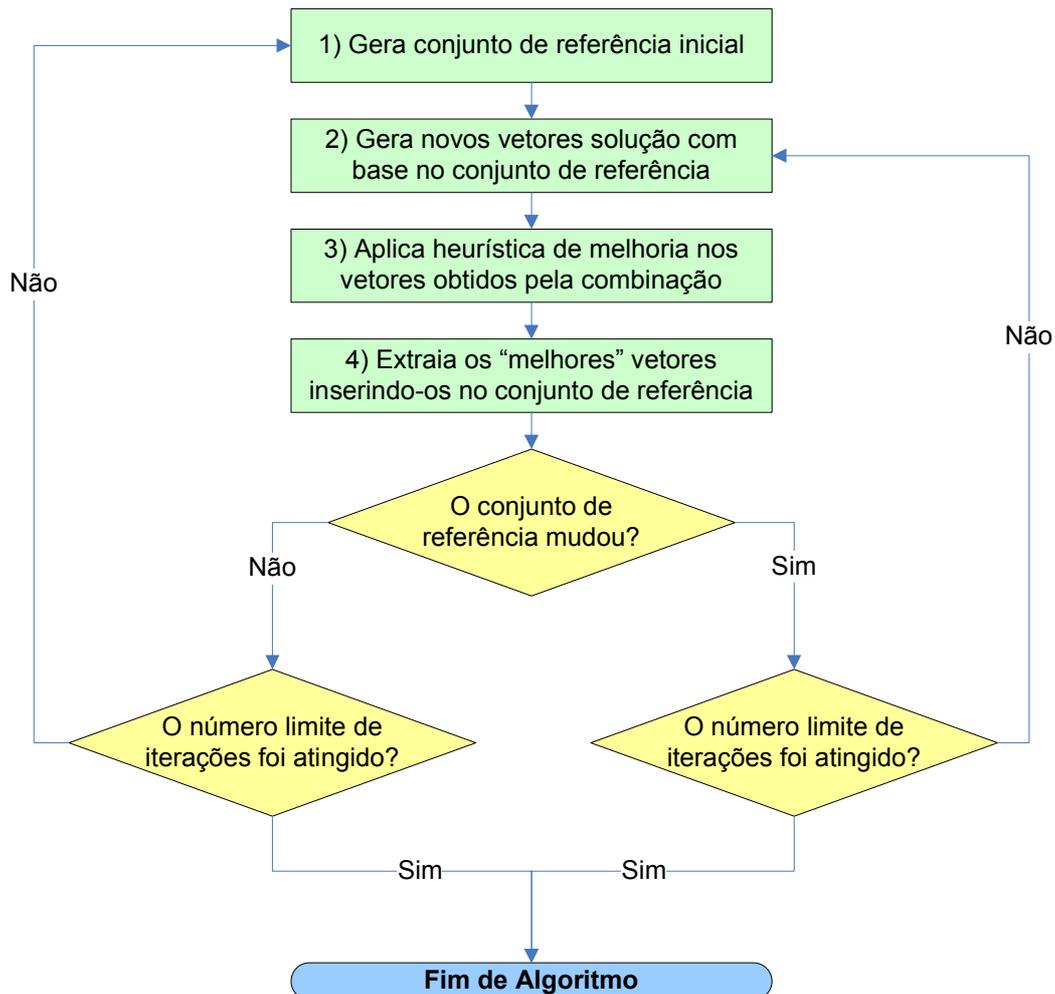


Figura 2.10 - Modelo geral do algoritmo da Busca Dispersa

Segue um detalhamento dos quatro principais passos do algoritmo geral da Figura 2.10.

- **Geração do conjunto de referência inicial** – A diversidade deve guiar a criação do conjunto de referência inicial. Cada vetor deve passar por um processo de melhoramento local (heurística de melhoria) antes de ser incorporado ao conjunto de referência. Isto faz com que cada elemento no conjunto de referência seja um ótimo local. Em caso de reinício do algoritmo, as melhores soluções da execução anterior devem ser automaticamente incorporadas no conjunto para manter um conjunto de soluções de elite para compor a resposta da busca.
- **Geração dos novos vetores solução** – O algoritmo de geração de novas soluções deve gerar vetores tanto dentro quanto fora da região convexa delimitada pelos vetores que originaram a geração. É possível que os vetores gerados pela

combinação linear dos vetores do conjunto de referência estejam fora do espaço de soluções do problema. Em casos como esse algum algoritmo de factibilização deve ser empregado ou então estes vetores devem ser ignorados.

- **Heurística de melhoria** – Um procedimento de melhoria deve ser capaz de, a partir de um vetor solução, obter uma solução melhorada que seja um ótimo local, ou seja, que nenhum vetor “vizinho” tenha melhor valor para a função objetivo que guia a otimização. Esta mesma heurística é aplicada nos pontos que farão parte do conjunto de referência inicial.
- **Atualização do conjunto de referência** – Como o conjunto de referência tem um número limitado de vetores solução, em algum ponto do algoritmo será necessário julgar qual entre dois vetores é o “melhor”. O valor da função objetivo é um critério de qualidade, mas manter a diversidade do conjunto de referência também deve ser um objetivo do algoritmo.

Capítulo 3 – REVISÃO BIBLIOGRÁFICA

Este trabalho utiliza novas técnicas de otimização na resolução do problema do ajuste de histórico. Por este motivo, esta revisão bibliográfica divide-se em duas partes: uma contendo referências a trabalhos desenvolvidos na área de otimização e a outra contendo um breve resumo dos esforços de otimização empregados no ajuste de histórico.

3.1 OTIMIZAÇÃO

3.1.1 Otimização Combinatória

Blum e Roli (2003), fizeram uma extensa pesquisa sobre como as metaheurísticas são usadas para resolver problemas de otimização combinatória. Os autores começam definindo o termo metaheurística, classificam as metaheurísticas de acordo com diversos critérios (por exemplo: inspiração ou não em processos naturais, uso de populações de soluções ou apenas uma solução simples, uso de uma função objetivo estática ou dinâmica etc), descrevem os principais métodos metaheurísticos baseados em trajetória (solução única) e populacionais, e, finalmente, analisam as diferentes estratégias de balanceamento das características de intensificação e diversificação das metaheurísticas.

3.1.2 Busca Direta

Talvez o trabalho mais famoso na área de busca direta tenha sido o de Hooke e Jeeves (1961), pois foi nele que o termo busca direta (*direct search*) foi cunhado. A inovação deste trabalho foi formalizar uma metodologia de otimização que não era baseada em informações de gradiente. Seu objetivo inicial era resolver uma classe de problemas para a qual as informações de gradiente eram difíceis ou impossíveis de obter. As aplicações iniciais do método foram: problemas de ajuste de curvas, resolução de equações de integração, maximização e minimização

de funções com e sem restrições nas variáveis independentes e a resolução de sistemas de equações. Hoje, quase 50 anos depois, os métodos de busca direta continuam influenciando o cenário da otimização de sistemas complexos.

Um interessante relatório interno da NASA (Lewis *et al.*, 2000) discorre sobre os motivos pelos quais os métodos de busca direta ainda hoje não foram substituídos por métodos “mais modernos”. Entre as razões, destacam-se: os métodos diretos funcionam bem na prática, têm obtido sucesso onde outros métodos falharam, têm sido usados como primeira abordagem em um processo de otimização devido à facilidade de implementação e, finalmente, os métodos do tipo Quasi-Newton não são aplicáveis a todos os problemas de otimização. Adicionalmente o relatório resume as principais características dos métodos diretos, os classifica de acordo com suas abordagens e analisa suas propriedades de convergência.

No trabalho de Kolda, Lewis et al. (2003), os métodos de busca direta são analisados na ótica da análise matemática. O trabalho faz um breve histórico da aplicação dos métodos de busca direta na área de otimização e então analisa teoricamente diversas abordagens para esta metodologia. Os principais métodos têm sua convergência e desempenho assintótico estudados e algumas limitações de aplicabilidade são determinadas. Um importante resultado da leitura deste trabalho foi a escolha para uso na tese de um método de busca direta baseado na metodologia sugerida por Hooke e Jeeves (1961) em comparação com outros métodos de busca direta como a busca da bússola (*compass search*) e o simplex. A razão principal desta escolha é a robustez do método que converge qualquer que seja a topologia do espaço de soluções.

3.2 OTIMIZAÇÃO APLICADO AO AJUSTE DE HISTÓRICO

Como já foi citado na Seção 2.4.2, os algoritmos de otimização podem ser divididos em dois tipos de acordo com o uso ou não de informações de gradiente: métodos baseados em gradiente e métodos de busca direta. Os métodos de otimização baseados em gradiente são os mais tradicionais uma vez que antecedem os métodos de busca direta e possuem mais estudos matemáticos de convergência e aplicabilidade. Por outro lado, os métodos de busca direta têm sido aplicados em problemas que não conseguem ser atacados pelos métodos tradicionais pela

ausência de um modelo analítico para o problema ou mesmo pela não “diferenciabilidade” da função objetivo da otimização.

Na área de ajuste de histórico, os métodos de gradiente têm utilizado as informações de gradiente que são obtidas diretamente dos modelos de diferenças finitas utilizados pelos simuladores de fluxo modernos. Entre suas principais qualidades está a maior velocidade de convergência em relação aos métodos de busca direta. Porém, nem sempre os simuladores numéricos permitem a obtenção de informações de gradiente, pois usualmente são programas comerciais de código fechado que não suportam o acoplamento de rotinas de otimização em sua metodologia. Adicionalmente os métodos de busca direta tendem a ser mais fáceis de implementar e menos sujeitos a singularidades numéricas. As seções a seguir mostram os principais trabalhos de otimização baseados em gradiente e de busca direta pesquisados durante a fase de revisão bibliográfica deste trabalho.

3.2.1 Métodos de Ajuste de Histórico Baseados em Gradientes

Arenas *et al.*, 2001, associaram o método de otimização Levenberg-Marquardt, baseado em gradientes, com a técnica de inserir pontos de referência (*pilot points*) cujas permeabilidades são ajustadas via dados de sísmica 4D com o restante das permeabilidades ajustadas via krigagem.

Normalmente os métodos baseados em gradientes conseguem apenas encontrar ótimos locais, porém, Gomez *et al.*, 2001 aplicaram uma técnica chamada tunelamento (*tunneling*) para encontrar uma série de soluções melhoradas a partir de uma solução mínima inicial tornando a metodologia proposta em um método de busca global. Posteriormente, Mantica *et al.*, 2002, trabalharam em outro método de busca global ao analisar criticamente iniciativas de otimização baseadas em *simulated annealing*. O resultado foi chamado de otimização caótica (*chaotic optimization*) e consiste em usar um método que explora o espaço de soluções usando uma analogia com um sistema de partículas energizadas cujas “cargas” são função de seus valores de função objetivo, a seqüência em que foram encontradas e alguns outros parâmetros determinados pelo usuário. O resultado final é um conjunto diversificado de soluções de partida que serão posteriormente otimizadas localmente por um método de gradientes tradicional.

3.2.2 Métodos de Busca Direta para o Ajuste de Histórico

O uso de metaheurísticas para assistir o processo de ajuste de histórico é bem difundido. Talvez o tipo de metaheurística mais utilizada seja da classe dos algoritmos evolutivos, mais especificamente Algoritmos Genéticos. Schulze-Riegert *et al.*, 2002 usou um Algoritmo Genético que fazia uso de informações de geoestatística para reduzir o espaço de soluções e também acelerar a convergência do método.

Uma técnica de ajuste semi-automático foi desenvolvida em Li *et al.*, 2006, onde um metamodelo construído a partir de um planejamento estatístico é iterativamente melhorado pela adição de simulações em sua construção de modo a melhor capturar as não linearidades do problema. Em seguida, a superfície de resposta gerada é explorada globalmente em busca dos mínimos relevantes ao problema de ajuste. A inovação deste trabalho está na forma em que a precisão da superfície de resposta é incrementalmente melhorada antes que o procedimento de ajuste tenha início.

Em Cullick *et al.*, 2006, uma metodologia híbrida baseada nas buscas tabu e dispersa foi aplicada ao problema do ajuste de histórico. A contribuição deste trabalho foi comparar o desempenho do método usando unicamente a simulação numérica de reservatórios no estilo otimização baseada em simulações (Seção 2.4), substituindo o simulador numérico por um metamodelo implementado em uma rede neural e também usando o metamodelo para gerar pontos de partida para o otimizador baseado em simulações. Outro aspecto interessante deste trabalho foi o treinamento da rede neural com simulações fornecidas por um planejamento aproximadamente ortogonal (*nearly orthogonal array design*, NOA).

3.2.3 Métodos híbridos

Alguns pesquisadores preferem unir o melhor dos dois mundos, gradientes e busca direta, na tentativa de aliar a robustez dos métodos de busca direta com a velocidade de convergência dos métodos baseados em gradiente. Este foi o caminho adotado por Leitão (1997) em sua dissertação de mestrado. Neste trabalho, o autor inicia o processo de ajuste com um método de busca direta e, a partir do momento onde a busca direta começa a ter convergência lenta, o processo é revertido para um baseado em gradientes.

Capítulo 4 – METODOLOGIA

O principal objetivo deste trabalho foi resolver o problema do ajuste de histórico usando a metaheurística da Busca Dispersa (ou *Scatter Search*), de forma que o uso de estratégias de intensificação e diversificação pudesse balancear as necessidades de encontrar soluções de qualidade e de explorar bem o espaço de soluções. Um dos passos adotados na metodologia proposta consiste em modelar o problema de ajuste de histórico como um problema de otimização combinatória de forma que possa ser resolvido pela Busca Dispersa. Resumidamente, os principais passos da metodologia empregada seguem:

1. Modelagem do problema de ajuste de histórico como um problema de otimização combinatória.
2. Desenvolvimento de um algoritmo para a Busca Dispersa aplicado ao problema de ajuste de histórico.
3. Validação do algoritmo proposto com base em casos onde a solução é conhecida.
4. Aplicação em um caso onde a solução não é conhecida.

4.1 MODELAGEM DO PROBLEMA DE AJUSTE DE HISTÓRICO COMO UM PROBLEMA DE OTIMIZAÇÃO COMBINATÓRIA

A metodologia aplicada neste trabalho consistiu em modelar o problema de ajuste de histórico de uma forma que fosse possível resolvê-lo, de maneira ótima ou quase ótima, com o auxílio da *metaheurística* Busca Dispersa. O primeiro passo nesta direção é propor uma forma de comparar o desempenho de modelos de simulação distintos.

4.2 AS PROPRIEDADES INCERTAS

Após alguns estudos comparativos entre o desempenho do reservatório simulado e o real, o responsável pelo ajuste formula algumas hipóteses sobre quais propriedades do reservatório devem ser modificadas (por exemplo, porosidade, permeabilidade, compressibilidade da rocha entre outras), qual deve ser a amplitude destas modificações ($\pm 10\%$, $\pm 20\%$ etc.) e também qual será o escopo dessas mudanças (região específica ou o reservatório todo). Dá-se o nome *propriedade incerta* a cada conjunto de propriedade, amplitude da modificação e escopo.

No exemplo da Figura 4.1, a propriedade incerta é a porosidade da rocha. Seu intervalo de validade está entre 14% e 34% e o escopo é o conjunto de blocos selecionados (em azul). Definido desta forma, esta propriedade pode assumir qualquer valor real entre 14% e 34% gerando um número infinito de possibilidades. Casos contínuos como estes tendem a usar avaliações de gradiente em relação à função objetivo, onde se determina uma *direção de decida* e o tamanho do *passo* a tomar nesta direção.

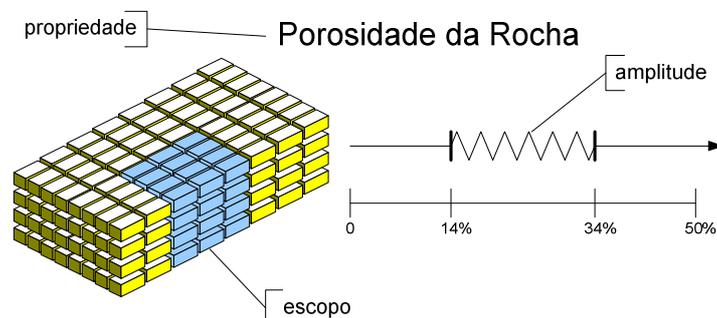


Figura 4.1 - Exemplo de uma propriedade incerta.

4.2.1 Discretização do Domínio

Como o objetivo final deste trabalho é aplicar uma metaheurística de otimização combinatória para resolver o problema de ajuste, e como a otimização combinatória se aplica a domínios finitos, é necessário substituir o intervalo contínuo que define cada propriedade incerta por um conjunto finito de valores discretos que buscará representar o intervalo original, sem perda significativa de qualidade das soluções. Uma forma de fazer isso é dividir o intervalo contínuo em certo número de subintervalos e considerar os valores dos limites de cada um deles como os valores discretos que formam o novo intervalo discretizado.

Aplicando este conceito na propriedade incerta definida na Figura 4.1, fixando o número de subintervalos em 4, obtém-se a discretização exibida na Figura 4.2, onde o intervalo $[14\%,34\%]$ é substituído pelo conjunto discreto $\{v \mid v \in \{14\%,19\%,24\%,29\%,34\%\}\}$. Quanto maior for o número de intervalos na geração da discretização, melhor será a representatividade da propriedade incerta em relação ao seu intervalo original, mas o número de simulações exigidas para mapear todas as possibilidades desta propriedade cresce de forma linear com o número de intervalos.

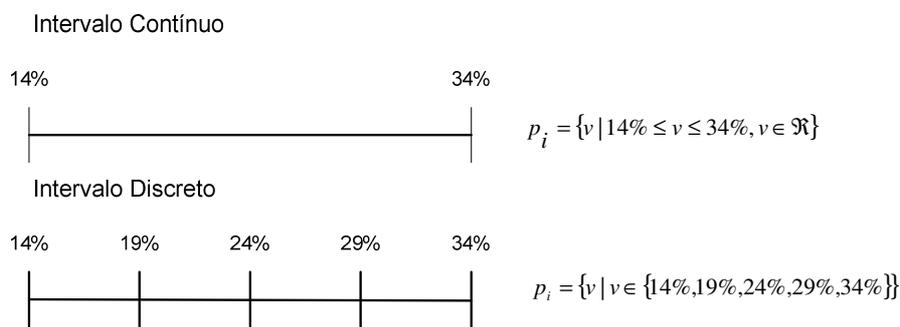


Figura 4.2 - Discretização de uma propriedade incerta.

O ajuste de histórico normalmente apresenta um número de parâmetros elevado, o que aumenta a complexidade do problema. Desta forma, é comum fazer uma análise prévia de sensibilidade da variação dos valores dos parâmetros de ajuste em relação à função objetivo (Seção 4.3), eliminando aqueles cujas perturbações de valor dentro de sua faixa de variação pouco influem na avaliação da função objetivo. Embora a análise de sensibilidade seja uma tarefa importante para todo processo de ajuste, o foco deste trabalho é na exploração eficaz do espaço de soluções gerado pela combinação dos parâmetros de ajuste selecionados, ou seja, esta exploração ocorre após a etapa de análise de sensibilidade. Por este motivo, o processo de análise de sensibilidade não está descrito neste trabalho.

4.3 A FUNÇÃO OBJETIVO

Como o objetivo principal do ajuste de histórico é deslocar as curvas simuladas em direção às curvas do histórico, e fazer isto de forma eficiente, é preciso criar um índice que possa ser usado para determinar numericamente se duas curvas estão próximas ou não. Neste trabalho o afastamento foi usado para este fim.

4.3.1 O Cálculo do Afastamento

Uma boa expressão para o afastamento dos pontos da curva de produção em relação ao histórico precisa garantir as seguintes propriedades:

- O sinal do afastamento deve indicar se a curva de valores simulados está majoritariamente acima (valores positivos) ou abaixo (valores negativos) da curva de histórico.
- Quando a curva simulada está ora acima, ora abaixo da curva do histórico, as distâncias calculadas em cada região não devem se cancelar mutuamente, mas sim serem somadas.

Com estes dois objetivos em mente definimos:

Definição 2 – Distância entre Curvas. Seja $H_p(t) = \{h_p(t_1), h_p(t_2), \dots, h_p(t_n)\}$ o conjunto de valores observados para um dado parâmetro de produção ou injeção p nos tempos t_1 a t_n (histórico de produção ou de injeção) e seja também $D_p(t) = \{d_p(t_1), d_p(t_2), \dots, d_p(t_n)\}$ o conjunto de valores simulados para o mesmo parâmetro p nos mesmos tempos t_1 a t_n (valores simulados). Define-se a distância entre estes conjuntos de pontos como:

$$\text{distância} = \sum_{i=1}^n (h_p(t_i) - d_p(t_i))$$

Observe que, segundo esta definição, valores negativos de distância indicam que a curva simulada está majoritariamente abaixo da curva do histórico e valores positivos indicam que os pontos simulados estão majoritariamente acima da curva do histórico. Observe também que um valor nulo para a distância pode tanto significar que a curva simulada está ajustada em relação ao histórico ou que a curva simulada esteve uniformemente distribuída acima e abaixo da curva de histórico, pois termos positivos e negativos se anulam no somatório.

A definição matemática do conceito de afastamento usado neste trabalho segue:

Definição 3 – Afastamento. Sejam $H_p(t) = \{h_p(t_1), h_p(t_2), \dots, h_p(t_n)\}$ e $D_p(t) = \{d_p(t_1), d_p(t_2), \dots, d_p(t_n)\}$ os conjuntos de pontos do histórico e dos dados simulados

conforme explicado na Definição 3. O afastamento da curva simulada do parâmetro p em relação à sua curva de histórico é calculado numericamente como:

$$A(p) = \underbrace{\frac{\sum_{i=1}^n (h_p(t_i) - d_p(t_i))}{\sum_{i=1}^n (h_p(t_i) - d_p(t_i))}}_{\text{sinal}} \times \underbrace{\sum_{i=1}^n (h_p(t_i) - d_p(t_i))^2}_{\text{módulo}}$$

ou seja, o módulo do afastamento é definido como a somatória do quadrado das diferenças e o sinal é dado pela distância calculada.

Esta definição de afastamento exige que o vetor do histórico possua o mesmo número de elementos que o vetor de dados simulados e também que cada par de dados $(h(p, t_i), d(p, t_i))$ tenha o mesmo referencial de tempo. Embora seja uma prática comum solicitar ao simulador gerar a saída dos dados simulados nos mesmos tempos em que os dados de histórico estão disponíveis, pode acontecer de estes vetores não estarem sincronizados ou mesmo não possuírem o mesmo número de elementos. Processos automatizados precisam levar em conta esta possibilidade de modo que se define aqui um *afastamento interpolado* onde o vetor de histórico funciona como referência enquanto que os respectivos dados simulados são interpolados para obter a correspondência desejada. Assim, substitui-se na Definição 3 cada par $(h_p(t_i), d_p(t_i))$ pelo par $(h_p(t_i), d_p^*(t_i))$, onde $d_p^*(t_i)$ é um ponto interpolado da série de dados simulados. Assim, a nova expressão do afastamento interpolado é:

$$A^*(p) = A^* = \frac{\sum_{i=1}^n (h_p(t_i) - d_p^*(t_i))}{\left| \sum_{i=1}^n (h_p(t_i) - d_p^*(t_i)) \right|} \times \sqrt{\sum_{i=1}^n (h_p(t_i) - d_p^*(t_i))^2} \quad (4.1)$$

4.3.2 Normalização dos Afastamentos

Pode-se ficar tentado a usar o próprio valor do afastamento como função objetivo a ser minimizada pelo método de ajuste, porém pode ser visto na Equação 4.1 que o afastamento é função de apenas um parâmetro de produção ou injeção. Muitas vezes ocorre que ao ajustar um determinado parâmetro de produção, vazão de água, por exemplo, a pressão de fundo de poço ou

a pressão média do reservatório podem ficar desajustadas levando a um desempenho ruim do modelo para a previsão de comportamento futuro. É por este motivo que a função objetivo global do ajuste normalmente envolve mais que um cálculo de afastamento.

O problema desta abordagem é a necessidade de compatibilizar as unidades que expressam esses afastamentos, cujos valores calculados podem diferir por várias ordens de grandeza fazendo com que um ou mais termos dominem o valor da função objetivo. Uma solução para este problema foi sugerida por Maschio e Schiozer (2005), onde se propõe o uso dos afastamentos obtidos no caso base para normalizar cada um dos afastamentos obtidos nos casos derivados. Nesta abordagem, o ajuste é obtido minimizando uma *função objetivo global* (FOG) que é dada por:

$$FOG = \frac{\sum_{i=1}^n w_i F_i(p_i, t)}{\sum_{i=1}^n w_i}, w_i \in \mathfrak{R}^+ \quad (4.2)$$

ou seja, a *FOG* é uma média ponderada das funções objetivo parciais. Observa-se que as funções objetivo parciais são dependentes de seus respectivos parâmetros de produção/injeção e do horizonte de simulação na forma de uma variável de tempo. Além do fato de os pesos w_i serem não negativos, pois a *FOG* terá que ser minimizada, não existe uma restrição sobre os valores que os pesos devem assumir, mas é comum que cada peso esteja no intervalo $[0;1]$ e que a somatória dos pesos seja 1.

As funções objetivo parciais, por sua vez, são definidas como segue:

$$F(p) = \frac{A_{\text{caso}}^*(p)}{A_{\text{base}}^*(p)} \quad (4.3)$$

onde os afastamentos do caso base e do caso derivado são calculados de acordo com a Equação 4.1. Observe que, no caso particular onde o caso sendo avaliado é o próprio caso base, o valor da função objetivo parcial relativa ao parâmetro incerto p é sempre 1. Observe também que quando o caso derivado possui um afastamento menor que o caso base (possui um ajuste melhor que o base) $F(p) < 1$. Analogamente, se o caso derivado apresenta um afastamento maior que o caso

base, $F(p) > 1$. Estas constatações são independentes do parâmetro de produção/injeção sendo avaliado de modo que as funções objetivo parciais são compatibilizadas apesar de possuírem diferentes unidades e ordens de grandeza.

4.4 O ESPAÇO DE SOLUÇÕES

Considerando que se dispõe de n propriedades incertas devidamente discretizadas, $p_i \mid 1 \leq i \leq n, i \in \mathfrak{X}$, cada uma formada por $m_i \in \mathfrak{X}^*$ valores reais, $v_{ij} \mid 1 \leq i \leq n, 1 \leq j \leq m_i \text{ e } i, j, n, m_i \in \mathfrak{X}$, pode-se imaginar um espaço de soluções onde cada *ponto* neste espaço representa uma combinação de valores fixados para cada propriedade incerta. Assim, cada ponto deste espaço é um caso derivado do caso base, onde as propriedades incertas assumem valores discretos dentro do intervalo de confiança de cada propriedade. No caso onde os valores do caso base para as propriedades incertas pertencem aos intervalos destas propriedades e estes valores não foram suprimidos pela discretização dos intervalos, o caso base faz parte do espaço de soluções.

Definição 4 – Espaço de Soluções. Seja $P = \{p_1, p_2, \dots, p_n \mid p_i \in \mathfrak{X}\} \subset \mathfrak{X}^n$ um dado conjunto de propriedades incertas conforme definição anterior. Considere que cada conjunto p_i seja formado por m_i valores reais v_{ij} , ou seja: $p_i = \{v_{ij} \mid 1 \leq j \leq m_i \text{ onde } j, m_i \in \mathfrak{X} \text{ e } v_{ij} \in \mathfrak{X}\}$. Seja agora o conjunto de soluções $S = \left\{ s_k = \{v_{1x_{k_1}}, v_{2x_{k_2}}, \dots, v_{nx_{k_n}}\} \text{ tal que } n = |P| \text{ e } 1 \leq k \leq \prod_{i=1}^n m_i \right\}$, ou seja, cada s_k é uma possível combinação entre os valores possíveis de cada propriedade incerta. Define-se o *espaço de soluções* do domínio P o conjunto S formado por todas as combinações de valores fixados para cada propriedade incerta em P . Por sua vez, cada elemento de S (s_k) pode ser avaliado pela *FOG* (Equação 4.2) mapeando o espaço de soluções em valores reais (vide diagrama da Figura 4.3).

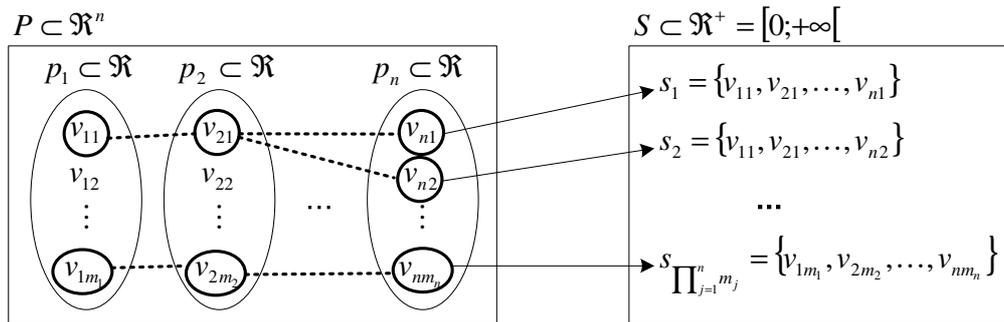


Figura 4.3 - O espaço de soluções é o contradomínio da função objetivo

Com estas definições, já é possível expressar o problema de ajuste em termos da definição de um conjunto de propriedades incertas discretizadas e uma *FOG*. Este espaço de soluções é combinatorial e, portanto, capaz de ser investigado com métodos metaheurísticos.

4.5 A IMPLEMENTAÇÃO DA BUSCA DISPERSA

O formato geral da Busca Dispersa foi resumido na Seção 2.8. Nesta seção, descreve-se como foram desenvolvidas as principais funções deste algoritmo que serão demonstradas no espaço de soluções da Figura 4.4.

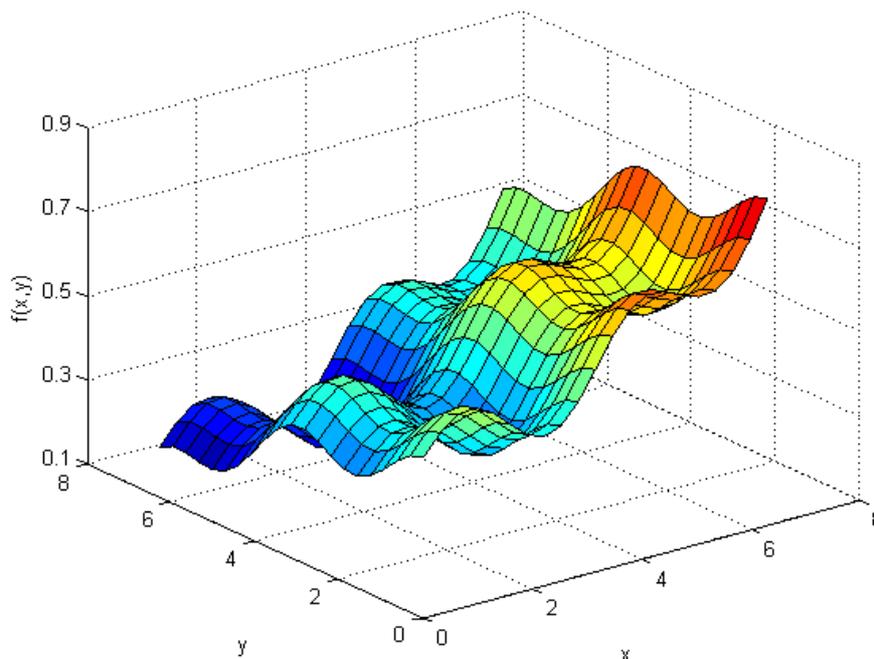


Figura 4.4 – Exemplo de um espaço de soluções

Este espaço de soluções foi gerado analiticamente a partir da função objetivo expressa na Equação 4.4 onde os parâmetros de ajuste x e y variam no intervalo $[0: 2\pi]$.

$$f(x, y) = \frac{8 \cdot \sin(2 \cdot x) + 5 \cos(2 \cdot y) + x \cdot y + x^2 - \frac{2}{3} \cdot y^2 - 240}{100} \quad (4.4)$$

Como a metodologia proposta explora espaços de soluções discretizados, a Tabela 4.1 mostra uma possível discretização para este espaço. Os valores em vermelho mostram os mínimos locais neste espaço enquanto que o valor vermelho em negrito ($x = 6.283, y = 2.513$) mostra o valor mínimo global.

Tabela 4.1 – Espaço de soluções de exemplo discretizado

		x										
		0.000	0.628	1.257	1.885	2.513	3.142	3.770	4.398	5.027	5.655	6.283
y	0.000	0.483	0.563	0.545	0.471	0.470	0.581	0.701	0.723	0.688	0.726	0.877
	0.628	0.445	0.525	0.508	0.434	0.432	0.544	0.664	0.686	0.651	0.689	0.840
	1.257	0.382	0.462	0.444	0.370	0.369	0.480	0.600	0.622	0.587	0.625	0.776
	1.885	0.368	0.448	0.431	0.357	0.356	0.467	0.587	0.609	0.574	0.612	0.763
	2.513	0.406	0.486	0.469	0.394	0.393	0.505	0.624	0.646	0.612	0.650	0.801
	3.142	0.417	0.497	0.480	0.405	0.404	0.515	0.635	0.657	0.622	0.660	0.812
	3.770	0.353	0.433	0.416	0.342	0.340	0.452	0.571	0.594	0.559	0.597	0.748
	4.398	0.263	0.343	0.326	0.252	0.250	0.362	0.481	0.504	0.469	0.507	0.658
	5.027	0.224	0.304	0.286	0.212	0.211	0.322	0.442	0.464	0.429	0.467	0.618
	5.655	0.235	0.315	0.298	0.223	0.222	0.334	0.453	0.475	0.440	0.479	0.630
	6.283	0.219	0.299	0.282	0.208	0.206	0.318	0.438	0.460	0.425	0.463	0.614

4.5.1 Geração do Conjunto de Referência Inicial

Como foi descrito na Seção 2.8.1, o conjunto de referência deve conter apenas soluções que passaram por um processo de otimização local, ou seja, que são ótimos locais. Porém, garantir uma boa amostragem inicial do espaço de soluções, otimizando todos os pontos da amostra onera bastante a inicialização do algoritmo (Sousa *et al.*, 2006). Por este motivo, a restrição de otimalidade foi relaxada na construção do conjunto de referência inicial de forma que os quatro melhores pontos da amostra inicial são inseridos sem otimização no conjunto de referência. O algoritmo base de geração do conjunto de referência inicial segue:

1. Gere uma amostragem inicial contendo pontos internos do espaço mais o ponto central do espaço de soluções.
2. Avalie o valor da função objetivo para todos os pontos obtidos no Passo 1.

3. Selecione os 5 melhores pontos para comporem o conjunto de referência inicial.

O objetivo da amostragem inicial é definir, com um número reduzido de simulações, regiões promissoras do espaço de soluções para iniciar o método. Cada parâmetro de ajuste, juntamente com seu intervalo de variação, pode ser encarado como gerador de um eixo no espaço de soluções. Como estes parâmetros são caracterizados por uma distribuição de probabilidades com média e desvio padrão (vide Figura 2.3), os valores extremos do intervalo não são recomendados para a geração de pontos de amostragem porque possuem menor probabilidade que aqueles localizados mais próximos da média. Os valores que fazem parte das soluções de amostragem são aqueles localizados a meio caminho do centro do intervalo, um do lado esquerdo, outro do lado direito (Figura 4.5). Para casos onde a distribuição não é simétrica e uniforme, pode-se fazer uma transformação de variáveis para considerar a probabilidade de ocorrência associada a cada atributo. Pode-se também trabalhar com a divisão do espaço de soluções de forma irregular para levar em conta esse fator.

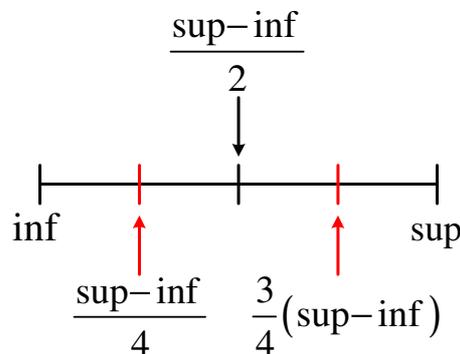


Figura 4.5 – Seleção dos valores de formação dos pontos de amostragem.

Como a metodologia proposta assume valores discretos no intervalo dos parâmetros de ajuste, pode ser necessário que os valores menor e maior, identificados adiante como ‘-’ e ‘+’, sejam fixados nos valores discretos mais próximos. Tendo definido um valor ‘-’ e ‘+’ para cada parâmetro de ajuste, é possível montar um conjunto de soluções de amostragem do espaço variando combinações destes valores em cada parâmetro. A Figura 4.6 mostra um exemplo de enumeração completa de amostragem baseado nos valores ‘-’ e ‘+’ de um problema com quatro variáveis P_i .

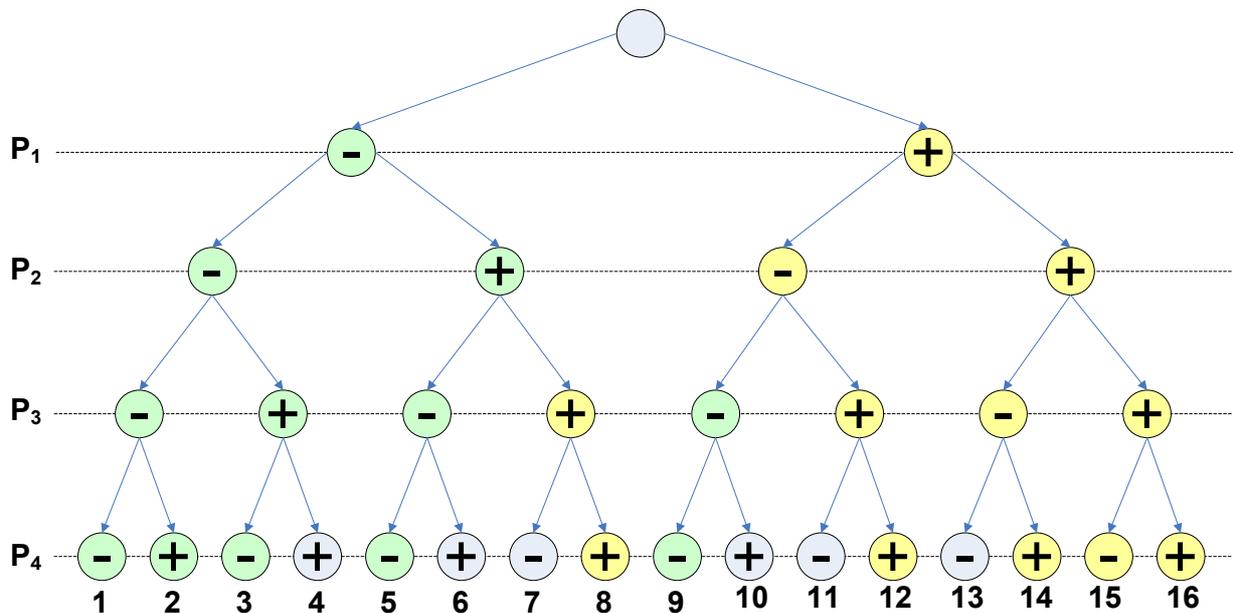


Figura 4.6 – Enumeração completa dos valores maior (+) e menor (-) dos parâmetros P_i .

O problema da enumeração completa é que ela cresce exponencialmente com o número de parâmetros: 2^n , onde n é o número de parâmetros de ajuste. Por este motivo, um subconjunto reduzido da enumeração completa foi gerado para compor a amostragem. Inicia-se com duas soluções iniciais: uma com todos os parâmetros configurados para o valor menor (caminho 1 na Figura 4.6) e outra com todos os parâmetros configurados para o valor maior (caminho 16 na Figura 4.6). Em seguida, ambos os vetores geram soluções modificadas onde apenas um valor menor (ou maior) é substituído por outro valor maior (ou menor). Por exemplo, trocar o valor de P_2 no caminho 1 de '-' para '+' gera o caminho 5, trocar o valor de P_3 no caminho 16 gera o caminho 14 e assim por diante. Os caminhos exibidos em verde na Figura 4.6 (1, 2, 3 e 5) são caminhos gerados a partir do caminho 1, os caminhos exibidos em amarelo (16, 15, 14 e 12) são caminhos gerados a partir do caminho 16. Já os caminhos de cor mista amarela e verde são caminhos que podem ser gerados tanto pela modificação do caminho 1 quanto do caminho 16.

A Figura 4.7 mostra um gráfico do crescimento do número de caminhos (soluções) de amostragem gerado pelas enumerações completa e reduzida. Observe que até 3 parâmetros de ajuste, não há diferença entre uma e outra enumeração, porém, a partir de 4 parâmetros, a enumeração completa apresenta crescimento exponencial enquanto que a enumeração reduzida provoca apenas um aumento linear. De fato, a expressão que define o número de caminhos gerados pela amostragem reduzida é $2n + 2$.

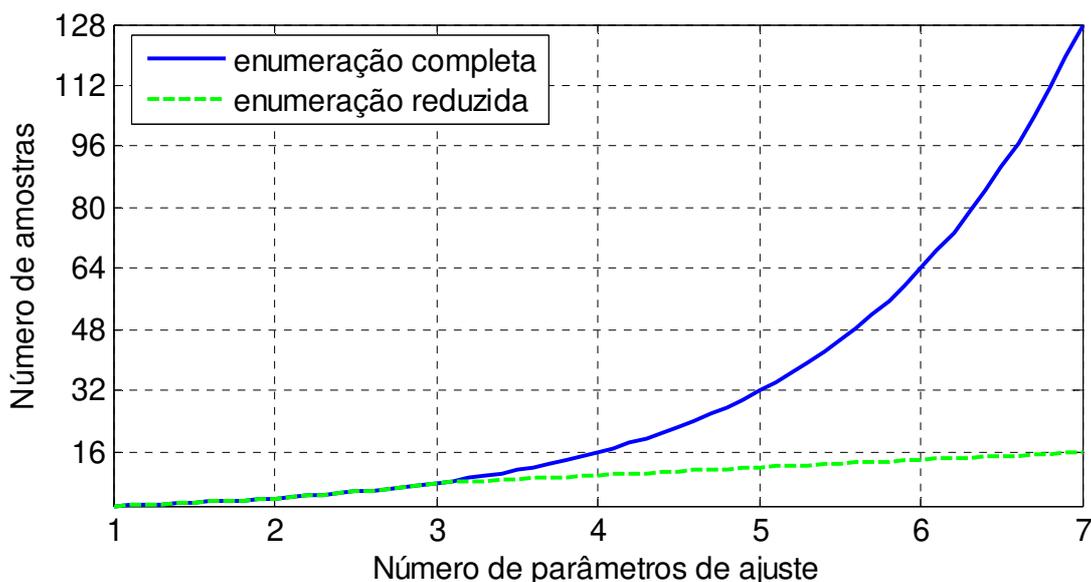


Figura 4.7 – Crescimento das enumerações com o aumento do número de parâmetros de ajuste.

Para concluir a descrição do algoritmo de amostragem reduzido proposto na metodologia, o ponto central do espaço de soluções, que é obtido configurando todos os parâmetros de ajuste para o valor central de seus intervalos de variação, é inserido neste conjunto.

Como exemplo do algoritmo de amostragem reduzido, considere o espaço de soluções definido na Tabela 4.1. Os valores menor e maior tanto para x quanto para y seriam 1.571 e 4.712, respectivamente. Acontece que estes valores não estão presentes na discretização assumida, de forma que os valores mais próximos precisam ser usados. Porém neste caso, como o valor está exatamente à mesma distância em relação àqueles mais próximos, há mais de uma escolha para a definição de ‘-’ e ‘+’. O critério de desempate (arbitrário) usado é que os menores valores têm prioridade. Logo, os valores discretizados escolhidos seriam 1.257 e 4.398 para ambos os parâmetros de ajuste (eles têm intervalos iguais). Para concluir o exemplo da amostragem, a Tabela 4.2 traz o conjunto de pontos que faria parte da amostragem inicial gerada pela metodologia proposta.

Tabela 4.2 – Pontos da amostragem reduzida para o espaço da Tabela 4.1

x	y	Valor FO	Observação
1.257	1.257	0.444	ponto menor
4.398	4.398	0.504	ponto maior
1.257	4.398	0.326	ponto misto
4.398	1.257	0.622	ponto misto
3.142	3.142	0.515	ponto central

Após a geração dos pontos de amostragem, todos eles são simulados, avaliados com relação à função objetivo e ordenados de forma crescente. Em seguida, as melhores simulações são inseridas sem otimização no conjunto de referência inicial. O número de soluções inseridas no conjunto de referências é configurado pelo parâmetro **N_REFSET_INICIAL**. Este parâmetro foi configurado para 5 neste trabalho. Valores muito baixos podem levar o algoritmo a falhar se o espaço de soluções possui poucos mínimos, porque a cada iteração duas soluções do conjunto de referência são inseridas na lista tabu e não podem ser usadas para a geração de novas soluções tentativas. Um valor entre 5 e o dobro do número máximo de iterações é adequado para este parâmetro.

4.5.2 Heurística de Melhoria Local

Uma das principais suposições da Busca Dispersa é que haja um algoritmo capaz de transformar uma solução tentativa em uma solução otimizada. Isto é necessário porque a principal estrutura da Busca Dispersa, o conjunto de referência, aceita apenas a inserção de soluções otimizadas⁴.

Três algoritmos de busca direta foram testados para a Busca Dispersa: o Método da Bússola (Kolda *et al.*, 2003), uma variação discretizada do Método de Newton (Dennis Jr. *et al.*, 1974) e a busca proposta por Hooke e Jeeves (Hooke *et al.*, 1961). Destes, o método que se mostrou mais robusto foi o de Hooke e Jeeves. O Método de Newton não funcionou bem no espaço de soluções discretizado utilizado, especialmente quando a solução corrente está próxima de um mínimo local, e o Método da Bússola levou a um número excessivo de simulações. Assim, foi adotada neste trabalho uma versão especializada do Método de Hooke e Jeeves com busca diagonal e simulações feitas em paralelo, como será descrito a seguir.

Considere novamente o espaço de soluções apresentado na Tabela 4.1. Uma visualização gráfica deste espaço de soluções aparece na Figura 4.8. As regiões representadas pelas cores quentes (laranja, amarelo e vermelho) são regiões onde o valor da função objetivo é alto enquanto que as regiões de cores frias (verde, ciano e azul) são regiões onde o valor da função objetivo é baixo. Embora a região apresentada na figura seja contínua, o espaço de soluções da Tabela 4.1 é

⁴ Na verdade, a metodologia proposta elimina esta restrição durante a inicialização do método conforme descrito na Seção 4.5.1.

discreto, e seus pontos aparecem na Figura 4.8 como os pontos onde a grade vertical cruza com a grade horizontal. Os pontos marcados com 'X' (vermelhos) são pontos de mínimo local, ou seja, aqueles cujos vizinhos (leste, oeste, norte e sul) tem maiores valores para a função objetivo.

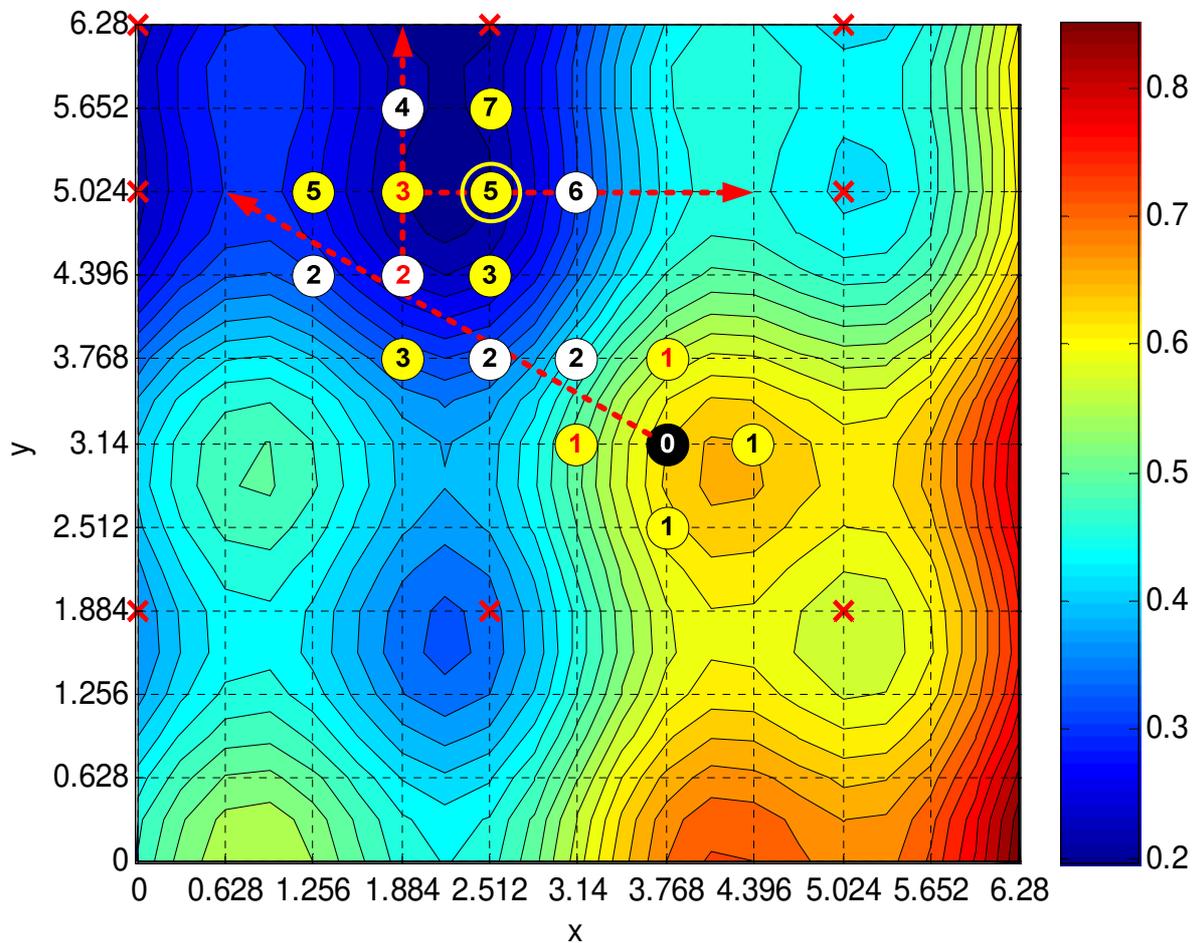


Figura 4.8 – Espaço de soluções da Tabela 4.1

O método de busca local utilizado consiste em intercalar seqüências de buscas exploratórias, onde a direção de descida é estabelecida, e buscas lineares, onde a direção de descida é explorada na busca de soluções melhoradas no espaço. Uma observação importante sobre o método Hooke e Jeeves é que ele sempre se inicia e termina com uma busca exploratória. A cada busca exploratória, um gradiente é aproximado pelos valores das funções objetivo dos pontos estudados. Quando uma busca exploratória não consegue achar pontos melhores que o atual ao visitar todos os pontos vizinhos, um mínimo local foi encontrado e o algoritmo de busca local chega ao fim.

Os pontos numerados na Figura 4.8 mostram o caminho percorrido pelo algoritmo de busca local, sendo que o Ponto 0 (zero) é o ponto de partida e o Ponto 5 (cinco) com um anel amarelo é o ponto de chegada.

O algoritmo se inicia com uma avaliação da vizinhança do Ponto 0 para determinar a direção de descida do método. A vizinhança é obtida variando-se as coordenadas do ponto central da busca dispersa para cima e para baixo, uma coordenada por vez. Em um espaço 2D, como o da Figura 4.8, isto corresponde a avaliar os pontos vizinhos a leste, oeste, norte e sul. Observe que avaliar também os vizinhos nas direções nordeste, noroeste, sudeste e sudoeste, implica em visitar uma vizinhança muito maior no caso geral. De fato, esta diferença consiste em avaliar $2n$ ou 2^n pontos do espaço, onde n é o número de dimensões (parâmetros de ajuste) do problema. Claramente é mais viável a avaliação de uma vizinhança que cresce linearmente com o número de dimensões do que outra que cresce de forma exponencial.

Voltando ao exemplo da Figura 4.8, a vizinhança do Ponto 0 corresponde aos pontos marcados com o número 1. Destes, apenas os marcados com o número 1 em vermelho fornecem valores melhorados para a função objetivo. Assim, a direção de descida é seguindo uma diagonal, indicada na figura com uma seta vermelha partindo do Ponto 0, onde os valores de x decrescem e os de y crescem. Observe que num espaço discreto, muitas vezes não é possível andar exatamente sobre a diagonal como mostra a busca linear partindo do Ponto 0. Neste caso, a trajetória apresenta ziguezague em torno da diagonal.

A Tabela 4.3 mostra o ponto de origem da busca exploratória inicial, $[x=3.770; y=3.142]$ cujo valor da função objetivo é 0.635. Os dois pontos vizinhos que melhoraram a função objetivo foram: $[x=3.142; y=3.142]$ e $[x=3.770; y=3.770]$.

Tabela 4.3 – Determinação do vetor direção da busca linear

Descrição	x	y	Valor FO	Δx	Δy	$ \Delta FO $
Origem da busca exploratória	3.770	3.142	0.635	-	-	-
Ponto de melhoria 1	3.142	3.142	0.515	-0.628	0.000	0.120
Ponto de melhoria 2	3.770	3.770	0.571	0.000	0.628	0.064
Vetor direção da busca linear				-0.628	0.334	-

O vetor direção da busca linear é formado pela composição das diferenças parciais normalizadas pela maior diferença obtida na função objetivo. Assim, o Ponto de melhoria 1

contribui com um Δx de -0.628 enquanto que o Ponto de melhoria 2 contribui com 53% de 0.628 na direção y, ou seja, com um Δy de 0.334.

Os pontos da busca linear são então gerados aplicando (somando) o vetor direção ao ponto de origem da busca exploratória até que o valor da função objetivo do novo ponto gerado seja maior que o do último ponto calculado ou que a fronteira do espaço de soluções seja alcançada. É importante notar que os pontos da busca linear provavelmente não farão parte do espaço de soluções discreto original. Assim, é necessário arredondar as coordenadas obtidas de forma que estas representem o ponto discreto mais próximo.

A Tabela 4.4 mostra todos os pontos que seriam alvos da busca linear até a fronteira. O Ponto 0 é o ponto de origem da busca exploratória enquanto que os demais são obtidos pela soma sucessiva do vetor direção obtido na Tabela 4.3. Os pontos $[x; y]$ nem sempre fazem parte do espaço de soluções original, por isso os pontos $[x'; y']$ foram obtidos arredondando as respectivas coordenadas para os valores discretos mais próximos.

Tabela 4.4 – Pontos da busca linear até a fronteira

ponto	x	y	x'	y'	Valor FO
0	3.770	3.142	3.770	3.142	0.635
1	3.142	3.476	3.142	3.770	0.452
2	2.514	3.809	2.513	3.770	0.340
3	1.886	4.143	1.885	4.398	0.252
4	1.258	4.477	1.257	4.398	0.326
5	0.630	4.811	0.628	5.027	0.304
6	0.002	5.144	0.000	5.027	0.224

Conforme foi explicado acima, a busca linear persiste até que um valor pior (maior) de função objetivo seja atingido. Isto ocorre no exemplo da Tabela 4.4 ao avaliar o Ponto 4. O resultado da busca linear neste caso é o Ponto 3. O ponto de resultado da busca linear é então submetido a uma nova busca exploratória e o ciclo de buscas exploratórias e lineares continua até que a busca exploratória não encontre nenhuma nova direção de descida, ou seja, o ponto central da busca linear é menor que todos os seus vizinhos.

Assim, a busca local mostrada esquematicamente na Figura 4.8 se inicia com a busca exploratória em torno do Ponto 0 e se encerra com uma busca exploratória em torno do Ponto 5 com círculo branco.

4.5.2.1 Uso de Computação Distribuída

Como pôde ser visto no pequeno exemplo da Figura 4.8, o número de simulações pode crescer bastante se o espaço de soluções for complexo o suficiente para exigir um número grande de reavaliações da direção de descida (buscas exploratórias). Por outro lado, quando o espaço de soluções é relativamente bem comportado, a busca linear leva a diversas avaliações da função objetivo até encontrar um ponto onde ela não melhora.

Quando se tem apenas um processador disponível, não há muito que fazer senão executar as simulações de fluxo sequencialmente avaliando o valor da função objetivo para cada caso. Porém, se existe a disponibilidade de um ambiente de computação distribuída, é possível desempenhar parte das tarefas de simulação em paralelo. Este é o caso, por exemplo, quando o usuário da metodologia tem acesso a um cluster de computadores.

Paralelismo na Busca Exploratória

Na busca exploratória, é sempre necessário avaliar todos os pontos vizinhos antes de determinar a direção de descida adotada pela busca linear. Se existem m máquinas à disposição, e $2n$ pontos a serem avaliados, onde n é o número de parâmetros de ajuste, será possível reduzir o tempo total de simulação consideravelmente.

A análise a seguir supõe que as m máquinas sejam iguais em relação à performance e estejam ociosas e que todos os processos de simulação levem um mesmo tempo t para ser executado em qualquer uma delas. Rodar as $2n$ simulações levaria um tempo $2nt$. Por outro lado, se as simulações forem realizadas em grupos de m simulações, uma para cada máquina disponível, o tempo total de simulação pode cair por um fator m . Este fator de decréscimo do tempo de simulação é conhecido como *speedup* e é computado pela razão entre o tempo serial pelo tempo paralelo de execução (Equação 4.5).

$$S = \frac{T_s}{T_p} = \frac{2nt}{2nt/m} = m \quad (4.5)$$

Idealmente o *speedup* é igual ao número de processadores, mas uma série de fatores como tempo de comunicação, diferenças de desempenho nas máquinas, divisão dos recursos

computacionais com outros processos e etc., fazem com que este valor dificilmente seja alcançado.

A busca exploratória implementada na Busca Dispersa proposta faz uso de um programa de computação distribuída que distribui os processos para as máquinas disponíveis. Os resultados do *speedup* obtido por esta implementação serão apresentados no capítulo de resultados.

Paralelismo na Busca Linear

Como não é possível saber, a priori, quantas simulações serão realizadas pela busca linear, a paralelização desta etapa da busca local é menos óbvia que a realizada na busca exploratória. Uma vez definida a direção de descida após a realização de uma busca exploratória, já é possível gerar todos os pontos potenciais da busca linear do ponto de partida até um ponto de fronteira do espaço de soluções (ver exemplo da Tabela 4.4).

A idéia da paralelização realizada pela Busca Dispersa proposta é sempre manter certo número de pontos do conjunto potencial da busca linear pré-avaliados (simulados) de forma a otimizar o tempo total de execução do algoritmo. Se um ponto não simulado for atingido pela busca, um novo conjunto de pontos é pré-avaliado, sempre sob demanda e em pacotes.

Para não penalizar o tempo de execução total do método nos casos em que apenas uma ou nenhuma simulação melhora o valor corrente da função objetivo, não é recomendado pré-simular mais pontos do que existam máquinas disponíveis para realizar as simulações. A recomendação desta metodologia é adotar um número entre $m/2$ e m como tamanho do pacote de simulações prévias, onde m é o número de máquinas disponíveis. Em particular, no desenvolvimento deste trabalho havia 7 máquinas disponíveis e o tamanho do pacote ficou configurado para 4, porque a rede não era de uso exclusivo.

4.5.3 Algoritmo de Combinação de Pontos

A cada nova iteração do algoritmo da Busca Dispersa, um par de pontos do conjunto de referência é combinado linearmente para gerar novas soluções candidatas a entrar neste grupo de soluções. Considere o espaço de soluções esquemático da Figura 4.9. Dois pontos deste espaço

definem uma reta. Esta reta pode ser expressa como uma equação parametrizada por um valor real t como mostra a Equação 4.6.

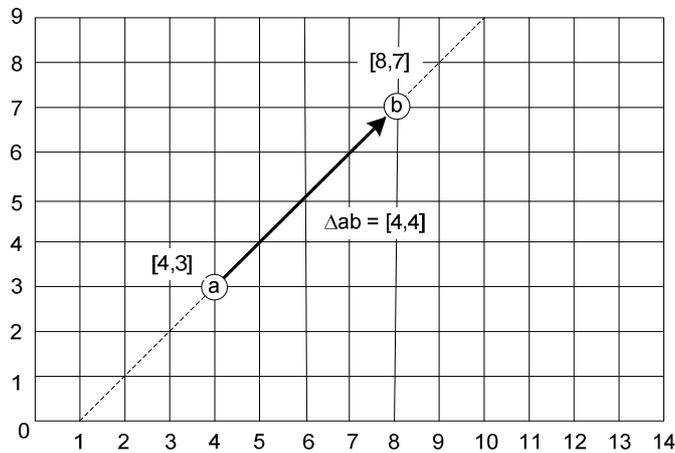


Figura 4.9 – Combinação linear dos pontos a e b

Observe que quando $t = 0$, $[x,y] = [x_1,y_1]$ e quando $t = 1$, $[x,y] = [x_2,y_2]$. Isto implica dizer que quando se usa t entre 0 e 1, faz-se uma combinação linear convexa dos pontos, conseqüentemente, valores de t menores que zero ou maiores que 1 geram uma combinação linear não convexa.

$$[x, y] = [x_1, y_1] + t \times [(x_2 - x_1), (y_2 - y_1)] \quad (4.6)$$

O algoritmo da Busca Dispersa gera pontos tanto dentro da região convexa quanto fora dela de modo que o conjunto de referência não fique preso dentro de sua própria região convexa. O algoritmo de geração das soluções encontra o intervalo de validade do parâmetro t (o ponto gerado não pode estar fora do espaço de soluções) e então gera um total de quatro pontos igualmente espaçados neste intervalo. Por exemplo, no espaço de soluções da Figura 4.9, a equação da reta definida pelos pontos a e b é: $[x, y] = [4,3] + t \times [4,4]$. O intervalo de validade de t neste caso é $[-0.75, 1.5]$, pois $[1,0] = [4,3] - 0.75 \times [4,4]$ e $[10,9] = [4,3] + 1.5 \times [4,4]$. Os quatro pontos que seriam gerados pelo método de combinação linear de pontos neste caso seriam $[2.8, 1.8]$, $[4.6, 3.6]$, $[6.4, 5.4]$ e $[8.2, 7.2]$. Porém, como o espaço de soluções é discreto e não contínuo, as coordenadas reais precisam ser convertidas para coordenadas inteiras, ou seja: $[3, 2]$, $[5, 4]$, $[6, 5]$ e $[8, 7]$. Observe na Figura 4.10 que o Ponto b coincide que o último ponto gerado na combinação e que, apesar de os valores de t que geraram os quatro pontos estarem igualmente

espaçados, o procedimento de arredondamento das coordenadas dos pontos para dentro do domínio do espaço de soluções removeu esta propriedade dos pontos discretos.

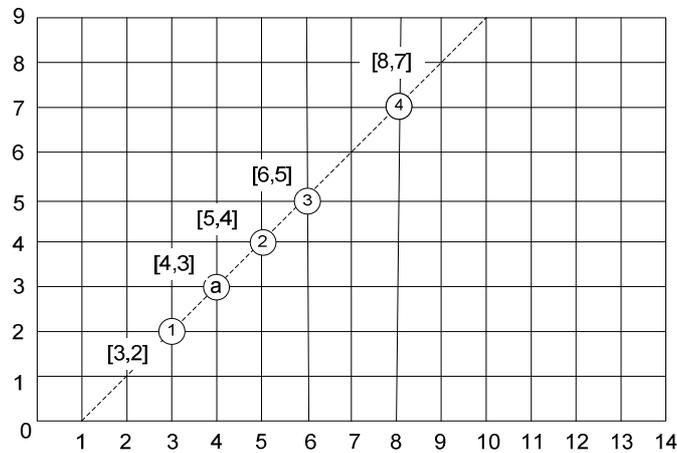


Figura 4.10 – Pontos gerados pela combinação linear de a e b

4.5.3.1 Critério de Seleção de Pontos Para Combinação

Uma nova combinação linear de pontos do conjunto de referência ocorre a cada iteração (exceto na primeira⁵), logo, um critério para a escolha de qual par de pontos será utilizado na geração de novos pontos se faz necessário.

Uma estratégia de intensificação é usada na escolha do primeiro ponto do par a ser combinado. É escolhido sempre o ponto não tabu com o menor valor absoluto para a função objetivo. Se nenhum ponto nestas condições existir, iterações artificiais são contabilizadas pela lista tabu até que uma solução não tabu exista no conjunto de referência. Tome, por exemplo, o conjunto de pontos da Tabela 4.5, sendo que a solução 3 é tabu. Neste caso, o ponto não tabu com o menor valor absoluto para o valor da função objetivo é o Ponto 5 ($x = 2525$, $y = 1735$ e valor da função objetivo = 0.255).

⁵ Com o relaxamento da otimalidade das soluções criadas durante o algoritmo de geração do conjunto de referência inicial, o processo de geração de soluções novas a partir de soluções de referência foi removido da iteração inicial onde, em seu lugar, ocorre a pré-seleção da melhor solução de referência para ser otimizado pela busca local.

Tabela 4.5 – Exemplo de um conjunto de pontos explorados e suas distâncias relativas

	1	2	3	4	5	
	x = 1788 y = 1428	x = 4312 y = 6539	x = 3543 y = 5488	x = 2159 y = 1791	x = 2525 y = 1735	
1	x = 1788 y = 1428	0	5700	4423	519	798
2	x = 4312 y = 6539	5700	0	1303	5213	5126
3	x = 3543 y = 5488	4423	1303	0	3947	3888
4	x = 2159 y = 1791	519	5213	3947	0	370
5	x = 2525 y = 1735	798	5126	3888	370	0
	Valor FO	0.539	0.748	0.293	0.799	0.255

Solução Tabu

Já o segundo ponto a ser escolhido segue uma estratégia de diversificação. Nesta estratégia, é selecionado como ponto secundário aquele que seja não-tabu e mais distante de todos os pontos visitados pelo método. A distância aqui é definida como a somatória das distâncias entre o ponto em questão e os demais pontos explorados. Cada distância individual é computada como a raiz da somatória das diferenças ao quadrado dos índices que compõem os pontos. Tomando novamente a Tabela 4.5 como exemplo e supondo que a tabela contém todos os pontos simulados até o momento, o ponto mais distante da solução 5 é a solução 2, que não é tabu. Se não houver pontos não tabu, iterações artificiais são inseridas na lista tabu até que pelo menos uma solução tabu vire não tabu e, portanto, candidata a fazer parte do par de pontos a serem combinados.

A idéia de combinar um ponto de intensificação e outro de diversificação é explorar o espaço de soluções entre uma boa solução conhecida e outra solução representativa de uma região pouco explorada no espaço de soluções. Para evitar ciclagem no método escolhendo um mesmo par de pontos para combinação no futuro, ambos os pontos são inseridos na lista tabu com durações tabu diferentes. O número de iterações que uma dada solução permanece na lista tabu é um parâmetro da Busca Dispersa que é denominado de **DURACAO_TABU**. Se a **DURACAO_TABU = n**, a solução de intensificação é inserida na lista tabu por n iterações, enquanto que a solução de diversificação é inserida na lista tabu por $n + 2$ iterações. Neste trabalho, o parâmetro **DURACAO_TABU** foi configurado para 4 iterações, metade do valor do número limite de iterações do método, que foi configurado para 8.

4.5.4 Atualização do Conjunto de Referência

Pela proposta original da Busca Dispersa, todos os pontos gerados pelo algoritmo de combinação linear (Seção 4.5.3) deveriam ser otimizados e inseridos no conjunto de referência, porém Sousa *et al.*, 2006, apontaram que este procedimento aumenta significativamente o número total de simulações realizadas pela Busca Dispersa de forma que foi sugerido que apenas

a melhor solução gerada fosse submetida para otimização pelo algoritmo da busca linear. O resultado desta otimização é então inserido no conjunto de referência.

4.5.5 Critérios de Parada

Dois critérios de parada são propostos para esta metodologia, um que busca limitar o total de recursos (tempo e poder computacional) utilizado pelo método e outro que detecta situações de ausência de melhoria.

Seguindo a proposta original da Busca Dispersa (Glover, 1998), um dos critérios de parada é quando o algoritmo atinge o número máximo de iterações permitidas pelo método (parâmetro **MAX_ITERACOES**). Nos testes realizados, um valor adequado para a metodologia proposta é de 8 iterações. Outra forma de limitar o uso de recursos computacionais é impor um limite no número total de simulações realizadas, mas este limite não foi usado nesta metodologia por dois motivos:

1. É difícil prever quantas simulações serão exigidas pela Busca Dispersa, pois isto é função tanto de fatores conhecidos como o número de parâmetros de ajuste e o número de níveis discretos utilizados nos parâmetros quanto de fatores desconhecidos como a topologia do espaço de soluções.
2. O custo de utilização do simulador de fluxo é fixo independente do número de simulações realizadas durante o período de vigência de sua licença de uso.

Em situações diferentes, um limite tanto no número de simulações quanto no tempo de execução do algoritmo podem ser adicionados se necessário.

O segundo critério de parada é interromper o algoritmo quando a qualidade do conjunto de referência não evolui com o tempo. Como o objetivo da Busca Dispersa é encontrar não apenas uma, mas diversas soluções possíveis para o problema do ajuste de histórico, a evolução do algoritmo é medida em termos da melhoria do valor da função objetivo das melhores soluções do conjunto de referência. Um outro parâmetro da Busca Dispersa é configurado para informar ao método quantas soluções são consideradas na medição de evolução da qualidade do conjunto de referência: o parâmetro **SOLUCOES_REFSET**. Este parâmetro em associação a outro parâmetro da Busca Dispersa, o parâmetro **MAX_ITERACOES_SEM_MELHORIA**, define o segundo critério de parada. Assim, se as soluções monitoradas do conjunto de referência não

melhorarem durante um número pré-especificado de iterações, a Busca Dispersa é interrompida. Nos estudos realizados durante este trabalho, o parâmetro **SOLUCOES_REFSET** foi configurado para 5 enquanto que o parâmetro **MAX_ITERACOES_SEM_MELHORIA** foi configurado para 2.

4.5.6 As Soluções Encontradas

No decorrer da Busca Dispersa, diversas buscas locais foram realizadas, cada uma delas encontrando, potencialmente, soluções ótimas locais distintas. Estas soluções ótimas locais estão agrupadas no conjunto de referência. Assim, pode-se afirmar que o objetivo final da Busca Dispersa é obter um conjunto de soluções ótimas locais que sejam uma boa amostragem das soluções possíveis para o problema.

Não obstante, a metodologia não pretende, encontrar todos os ótimos locais do espaço de soluções sendo explorado, pois já foi estabelecido anteriormente que o número de simulações para garantir este objetivo seria inviável. Por este motivo, é tarefa do usuário da metodologia interpretar se o conjunto de soluções encontrado é suficiente para os objetivos fixados durante a parametrização do problema de ajuste. Por exemplo, se em n iterações da Busca Dispersa, apenas uma solução ótima local foi encontrada, é bastante provável que este seja o ótimo global deste espaço de soluções. Por outro lado, se em n iterações foram encontradas n soluções ótimas locais, pode-se concluir que o espaço de soluções é bastante complexo e mais iterações seriam necessárias para melhor explorá-lo.

4.6 VALIDAÇÃO E APLICAÇÃO DO MÉTODO DE OTIMIZAÇÃO

A metodologia proposta inclui uma etapa de validação do método de otimização que usa modelos de reservatório cuja parametrização e solução ótima de ajuste são conhecidas de modo que seja possível avaliar a convergência do método proposto. Os modelos de validação são 3 reservatórios sintéticos com 1, 2 e 4 arranjos *five-spot* (5, 8 e 13 poços, respectivamente), que apresentam um nível de complexidade incremental para a metodologia na medida em que o número de variáveis do problema de ajuste aumenta (4, 8 e 16 parâmetros de ajuste, respectivamente). Outro aspecto analisado pela exploração dos espaços de soluções representados

pelos casos de validação é o crescimento do custo de execução da metodologia quando o número de variáveis do problema aumenta.

Nos casos de validação, a função objetivo utilizada sempre foi a vazão de água nos poços produtores conforme a Equação 4.7:

$$f(S_i) = \frac{\sum_{p=p_1}^{p_n} w_p \times \overbrace{\frac{\sum_{t=1}^n (Q_{w_{sim_i}}(t) - Q_{w_{obs_i}}(t))^2}{\sum_{t=1}^n (Q_{w_{sim_i}}(t) - Q_{w_{obs_{base}}}(t))^2}}^{\text{afastamento normalizado}}}{\underbrace{\sum_{p=p_1}^{p_n} w_p}_{\text{média ponderada dos afastamentos}}} \quad (4.7)$$

onde:

S_i é uma dada solução dentro do conjunto de soluções.

p é uma variável que representa os poços produtores do caso.

w_p é um peso associado ao afastamento de um dado poço.

t representa os tempos onde valores de histórico são conhecidos.

$Q_{w_{sim_i}}(t)$ é o valor simulado da vazão de água no tempo t .

$Q_{w_{obs_i}}(t)$ é o valor registrado (histórico) da vazão de água no tempo t .

$Q_{w_{sim_{base}}}(t)$ é o valor simulado da vazão de água no tempo t no caso base (origem do ajuste).

Observe que na Equação 4.7 os afastamentos são normalizados com aqueles encontrados no caso base. Desta forma, uma simples inspeção no valor da função objetivo indica se houve ou não melhoria em relação ao caso base. Valores abaixo de 1.0 indicam que houve melhoria, valores acima deste valor indicam piora em relação ao caso base. Finalmente, a Equação 4.7 permite associar um peso w_p ao afastamento normalizado de cada poço de forma a privilegiar ou penalizar o ajuste de um ou mais poços na busca do ajuste.

Em uma segunda etapa, o Caso 1C é resolvido de forma assistida de forma usar informações sobre a topologia do reservatório para restringir o número de soluções possíveis para o problema com o objetivo de melhorar o resultado obtido num processo automático. Em

seguida, um outro ajuste assistido é realizado em um reservatório mais complexo. Este reservatório não possui uma solução para o ajuste conhecida, de fato, nem uma parametrização para o problema é conhecida de antemão. Enquanto que os casos de validação são resolvidos de forma automática, nestes casos um processo de ajuste assistido é utilizado mostrando a aplicabilidade da Busca Dispersa também neste contexto de ajuste.

Na primeira abordagem (validação), os procedimentos de ajuste ocorrem de forma automática visto que a parametrização e os valores que ajustam o modelo são previamente conhecidos para validação do método de otimização. Em contraste, na segunda abordagem (aplicação), os parâmetros de ajuste bem como seus intervalos de variação estão sendo estudados concomitantemente com o ajuste de histórico, o que leva a um procedimento de ajuste assistido.

Capítulo 5 – APLICAÇÃO

Este capítulo descreve os modelos de simulação utilizados nas etapas de validação e aplicação da metodologia.

5.1 CASOS DE VALIDAÇÃO

Os modelos de validação são reservatórios sintéticos desenvolvidos com arranjos *five-spot* de poços, isto é, conjuntos de produção compostos de um poço injetor central e quatro poços produtores periféricos. Existe um total de três casos de validação: 1A, 1B e 1C. O Caso 1A possui um único arranjo *five-spot*, o Caso 1B tem dois arranjos *five-spot* enquanto que o Caso 1C possui quatro arranjos *five-spot*. Esses reservatórios são caracterizados por serem homogêneos em todas as suas propriedades exceto pelas permeabilidades horizontais e verticais como será visto a seguir.

5.1.1 Caso 1A

O Caso 1A é um reservatório sintético discretizado com $23 \times 23 \times 5$ blocos. Cada bloco do reservatório tem 50 m de lado e 5 m de altura. Todas as propriedades são constantes nos blocos exceto as permeabilidades horizontais e verticais, que são constantes apenas por regiões. A Figura 5.1 mostra o mapa de permeabilidade horizontal para uma camada. Não há diferenças de propriedades de uma camada para a outra.

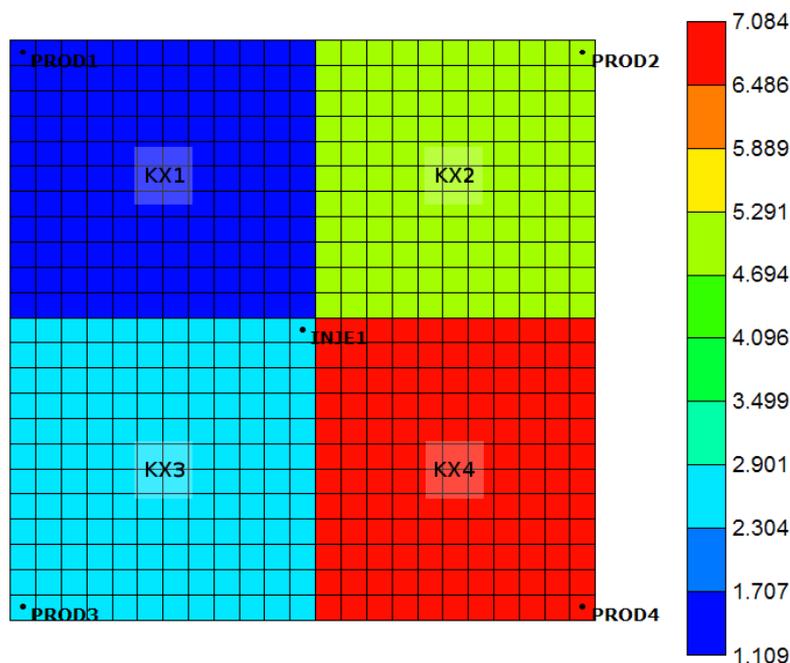


Figura 5.1 - Mapa de permeabilidade horizontal do Caso 1A

A Tabela 5.1 traz um resumo das principais propriedades do reservatório do Caso 1A. O histórico deste caso foi gerado após a divisão do reservatório em quatro regiões de permeabilidades distintas, uma região por área de drenagem entre cada par injetor produtor. A definição destas regiões e as permeabilidades fixadas estão detalhadas na Tabela 5.2.

Tabela 5.1 - Principais propriedades do reservatório do Caso 1A

Propriedade	Valor	Unidade
Porosidade	30%	-
Compressibilidade da Rocha	2.07E-09	1/kPa
Permeabilidade na direção i	(mapa)	mD
Permeabilidade na direção j	(igual à Kx)	mD
Permeabilidade na direção k	(mapa)	mD
Densidade do óleo	786.58	kg/m ³
Densidade do gás	0.96	kg/m ³
Densidade da água	1038	kg/m ³
Compressibilidade do óleo	1.23094E-06	1/kPa
Compressibilidade da água	2.06E-09	1/kPa
Pressão de referência	100	kPa
Fator volume de formação da água	1.041	m ³ /m ³

Tabela 5.2 - Regiões e valores de permeabilidade do histórico do Caso 1A

Região	Blocos			Kx (mD)	Kz (mD)
	i	j	k		
Kx ₁	1 a 12	1 a 11	1 a 5	1109	222
Kx ₂	13 a 23	1 a 11	1 a 5	5270	1054
Kx ₃	1 a 12	12 a 23	1 a 5	2530	506
Kx ₄	13 a 23	12 a 23	1 a 5	7084	1416

Com esta configuração, o modelo de reservatório foi simulado e um histórico de vazão de água foi gerado. Com o objetivo de inserir incertezas nos dados de histórico, cada valor simulado foi perturbado com um valor aleatório entre -10% e +10%. A função objetivo utilizada neste caso é aquela descrita pela Equação 4.7.

Para gerar o modelo base deste caso, as permeabilidades horizontais foram modificadas em todas as quatro regiões para o valor 4500 mD. As permeabilidades verticais permaneceram as mesmas do caso gerador do histórico. A comparação entre as vazões simuladas de água e aquelas presentes no histórico é mostrada na Figura 5.2. Observe que os poços mais desajustados no modelo base são os poços PROD1 e PROD4.

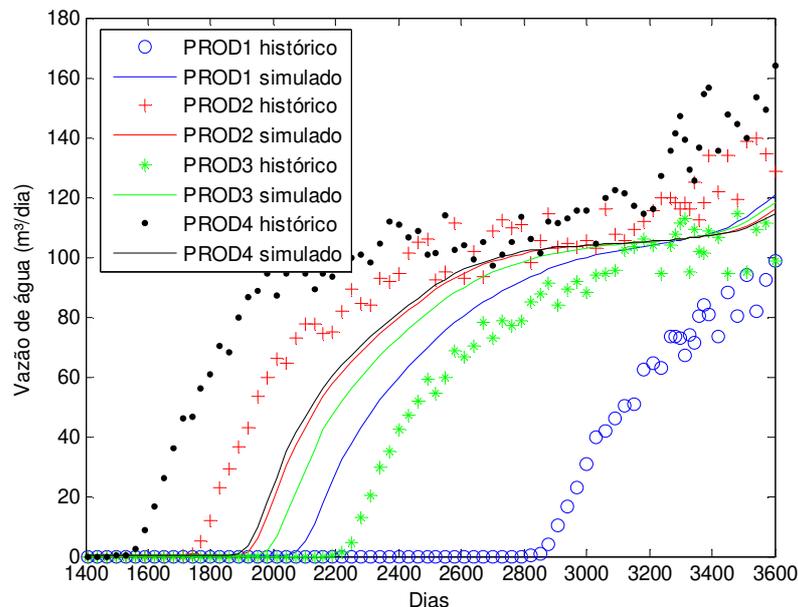


Figura 5.2 – Comparação das vazões de água do modelo base versus histórico do Caso 1A.

A metodologia da Busca Dispersa foi utilizada para resolver 4 diferentes instâncias do problema, cada qual com um número diferente de níveis usados na discretização dos parâmetros de ajuste (ver Tabela 5.3).

Tabela 5.3 – Quatro diferentes espaços de soluções para o Caso 1A.

Parâmetro de Ajuste	Valor Mínimo (mD)	Valor Máximo (mD)	Δ (mD) 57 níveis	Δ (mD) 29 níveis	Δ (mD) 15 níveis	Δ (mD) 8 níveis
KX1	1000	8000	125	250	500	1000
KX2	1000	8000	125	250	500	1000
KX3	1000	8000	125	250	500	1000
KX4	1000	8000	125	250	500	1000

Estes diferentes níveis discretos (8, 15, 29 e 57) foram definidos de forma que as soluções presentes em um modelo menos refinado estivessem também presentes no modelo mais refinado. Esta característica está evidenciada na Tabela 5.3 pelo fato de os tamanhos dos intervalos escolhidos em cada caso (colunas dos Δ) são múltiplos do intervalo usado no caso mais refinado. Desta forma, esperava-se que a qualidade das soluções obtidas no espaço mais refinado (com 57 níveis discretos) fossem superiores às obtidas nos espaços de soluções mais grosseiros (29, 15 e 8 níveis discretos).

5.1.2 Caso 1B

Este caso é composto de dois arranjos *five-spot* dispostos lado a lado como mostra a Figura 5.3. O reservatório possui uma malha de 46×23×5 onde cada bloco possui 50 m nos lados e 5 m de altura.

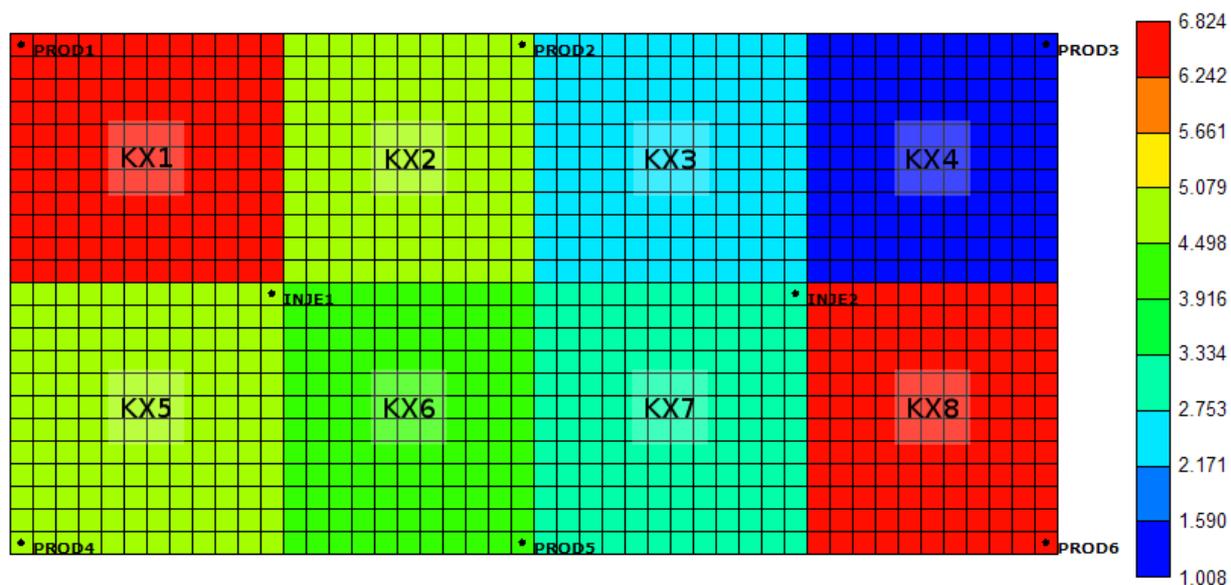


Figura 5.3 - Mapa de permeabilidade horizontal do Caso 1B

Novamente, o reservatório é homogêneo exceto pelas permeabilidades horizontais e verticais. Estes variam por região conforme especificado na Tabela 5.4. As demais propriedades deste reservatório são as mesmas do Caso 1A que se encontram expressos na Tabela 5.1.

Tabela 5.4 – Regiões e valores de histórico das permeabilidades do Caso 1B

Região	Blocos			Kx (mD)	Kz (mD)
	i	j	k		
KX ₁	1 a 12	1 a 11	1 a 5	6257	1251
KX ₂	13 a 23	1 a 11	1 a 5	4663	933
KX ₃	24 a 36	1 a 11	1 a 5	2256	451
KX ₄	36 a 46	1 a 11	1 a 5	1008	202
KX ₅	1 a 12	12 a 23	1 a 5	4803	961
KX ₆	13 a 23	12 a 23	1 a 5	4119	824
KX ₇	24 a 36	12 a 23	1 a 5	3277	655
KX ₈	36 a 46	12 a 23	1 a 5	6824	1365

Este modelo foi simulado para gerar o histórico de vazões de água dos poços produtores e, assim como no Caso 1A, os dados de histórico foram perturbados por um ruído aleatório entre -10% e +10%. Novamente, para gerar o modelo base, todas as permeabilidades horizontais da Tabela 5.4 foram mudadas para 4500 mD e as permeabilidades verticais foram mantidas constantes.

A Figura 5.4 mostra o quanto a homogeneização da permeabilidade horizontal afetou o ajuste da vazão de água dos poços produtores. O poço menos afetado foi o PROD4 e o mais afetado foi o PROD3.

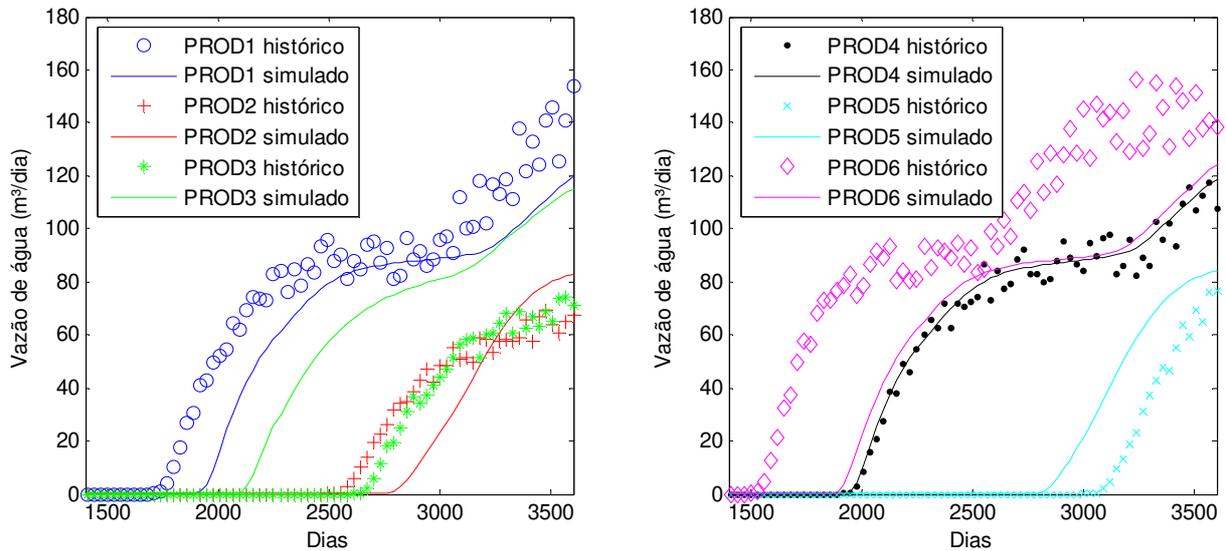


Figura 5.4 – Comparação entre as vazões de água do modelo base e o histórico do Caso 1B

5.1.3 Caso1C

Assim como o Caso 1B é a duplicação do Caso 1A, o 1C é a duplicação do Caso 1B. Assim, o Caso 1B contém um total de 4 arranjos *five-spot* com nove poços produtores e 4 injetores. O reservatório possui uma malha $46 \times 46 \times 5$ onde cada bloco possui 50 m de lado e 5 m de altura. Também neste caso, o reservatório é essencialmente homogêneo com as propriedades já descritas na Tabela 5.1. As únicas propriedades descritas através de mapas são as permeabilidades horizontal e vertical. A Figura 5.5 mostra a localização dos poços produtores e injetores no reservatório juntamente com o mapa de permeabilidade horizontal, que é constante apenas nas regiões de influência dos injetores com seus respectivos produtores.

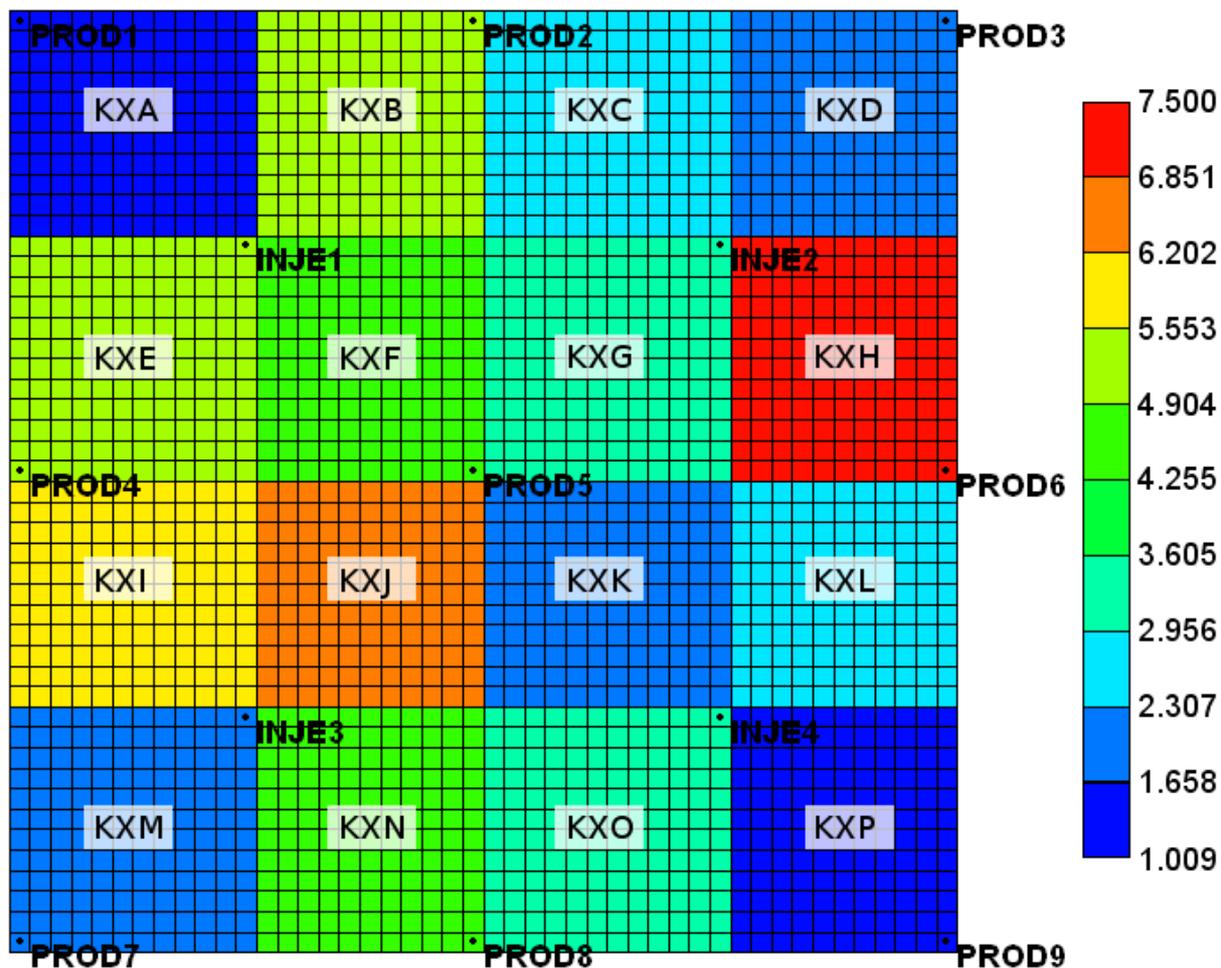


Figura 5.5 – Mapa de permeabilidade horizontal do Caso 1C

A Tabela 5.5 traz a definição do mapa das permeabilidades horizontal e vertical do modelo utilizado para gerar o histórico de produção. Assim como nos demais casos de validação, o histórico de produção obtido da simulação foi perturbado com a adição de um ruído aleatório de -10% a +10%.

Tabela 5.5 – Definição dos mapas de permeabilidade horizontal e vertical (Caso 1C)

Região	Blocos			Kx (mD)	Kz (mD)
	i	j	k		
KX _A	1 a 12	1 a 11	1 a 5	1009	202
KX _B	13 a 23	1 a 11	1 a 5	4945	989
KX _C	24 a 36	1 a 11	1 a 5	1253	471
KX _D	36 a 46	1 a 11	1 a 5	1900	380
KX _E	1 a 12	12 a 23	1 a 5	5095	1019
KX _F	13 a 23	12 a 23	1 a 5	4359	872
KX _G	24 a 36	12 a 23	1 a 5	3452	690
KX _H	36 a 46	12 a 23	1 a 5	7500	1500
KX _I	1 a 12	24 a 34	1 a 5	5800	1160
KX _J	13 a 23	24 a 34	1 a 5	6226	1245
KX _K	24 a 36	24 a 34	1 a 5	2219	444
KX _L	36 a 46	24 a 34	1 a 5	2500	500
KX _M	1 a 12	35 a 46	1 a 5	2100	420
KX _N	13 a 23	35 a 46	1 a 5	4595	919
KX _O	24 a 36	35 a 46	1 a 5	3128	626
KX _P	36 a 46	35 a 46	1 a 5	1105	221

Adicionalmente, o modelo base foi obtido do modelo de geração do histórico através da homogeneização da permeabilidade horizontal para 4500 mD em todas as regiões definidas na Tabela 5.5. O impacto desta mudança no ajuste pode ser visto nos gráficos da Figura 5.6.

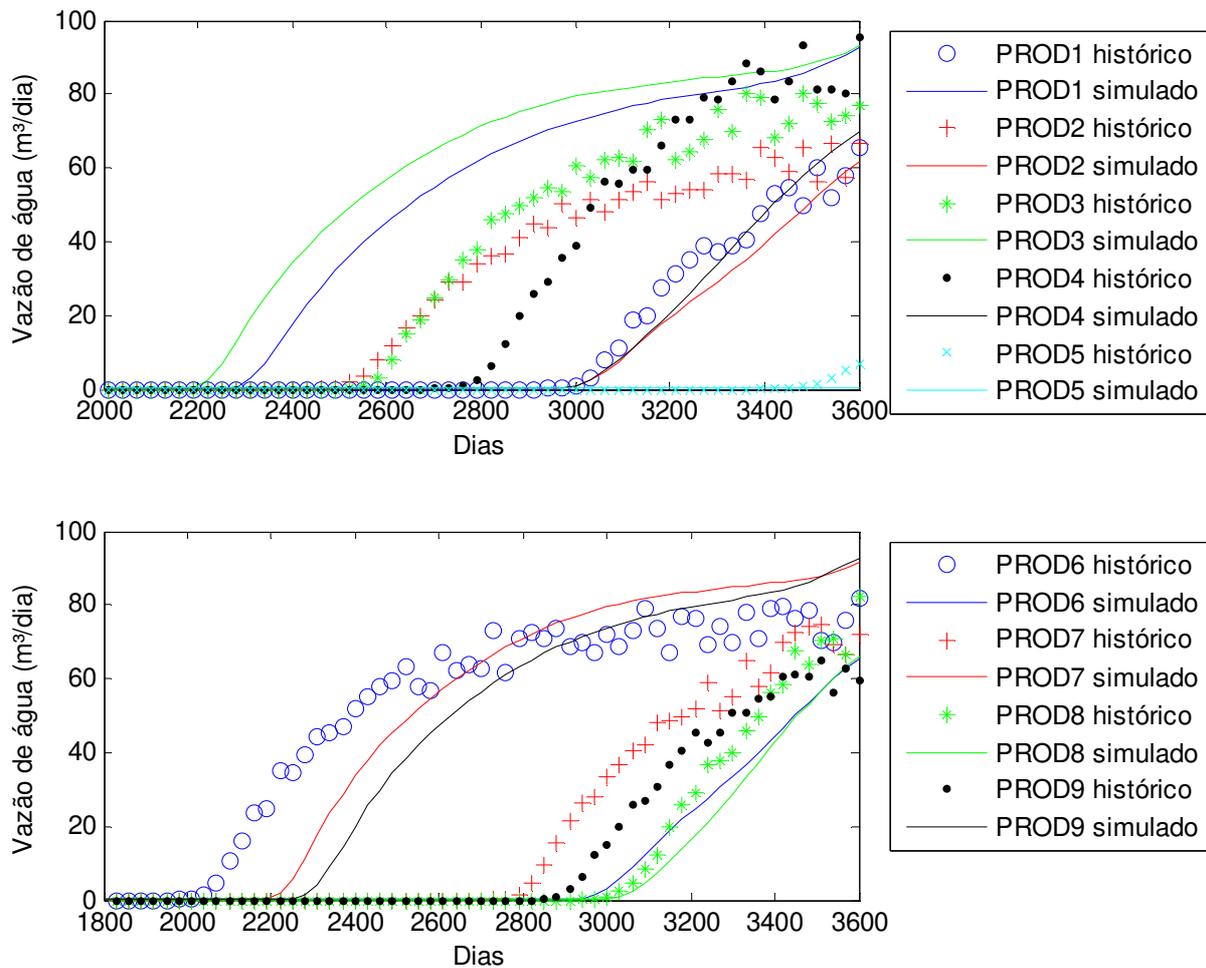


Figura 5.6 – Comparação das vazões de produção de água do modelo base com o histórico (Caso 1C)

5.2 CASO DE APLICAÇÃO (CASO 2)

O Caso 2 é um caso sintético com características de um caso real; o modelo reproduz um sistema turbidítico depositado em águas profundas. O histórico foi gerado a partir de um modelo bastante refinado enquanto o modelo de simulação (base) foi obtido a partir de uma reinterpretação dos dados originais com algumas simplificações e alterações de propriedades na geração de um modelo bem mais grosseiro. O objetivo destes procedimentos foi inserir grandes incertezas na localização dos canais de permeabilidade e nos valores das propriedades. Este caso foi gerado na Petrobras e utilizado no trabalho Silva *et al.*, 2005. Ele possui dois tipos principais de rocha: uma de maior qualidade que forma os canais do sistema e outra de menor qualidade que compõe a areia de fundo que preenche a maior parte do reservatório. No modelo gerador do

histórico, os poços estão todos localizados dentro dos canais de permeabilidade de modo que estes desempenham um importante papel na produção do campo. Algumas dificuldades associadas ao caso base do ajuste incluem: alguns poços produtores e injetores se localizam fora da região de canais após a reinterpretação dos dados sísmicos, os canais bem delimitados do caso gerador de histórico encontram-se difusos e diversos poços injetores estão com problemas de injetividade. Maiores detalhes sobre o caso podem ser obtidos diretamente na fonte citada (referido como Caso 3 naquele trabalho).

A Figura 5.7 mostra como está o comportamento do corte de água nos poços produtores no caso base em relação aos respectivos históricos de produção. As curvas são apresentadas duas a duas, identificadas por cor, onde a curva formada por símbolos é a curva de histórico e a curva contínua representa o comportamento do respectivo poço no caso base.

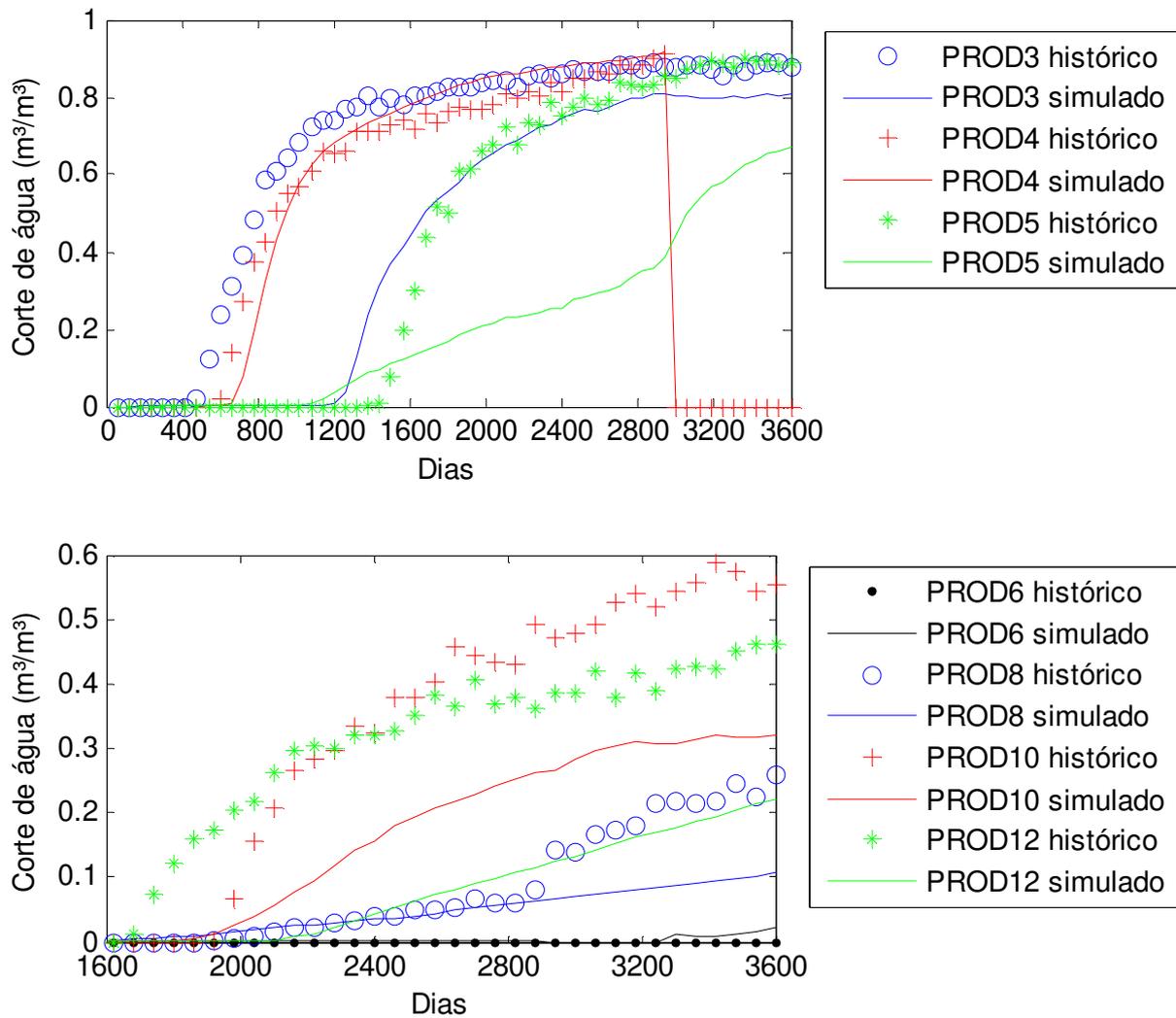


Figura 5.7 – Comparação entre os cortes de água do caso base em relação ao histórico (Caso 2).

Observe que o único poço ajustado no caso base é o PROD4. O poço PROD6 não possui produção de água no período de simulação (3600 dias) por isto ele pode ser excluído do ajuste.

Capítulo 6 – RESULTADOS E DISCUSSÕES

Este capítulo está dividido em duas partes: a apresentação dos resultados obtidos através dos casos de validação (Casos 1A, 1B e 1C) e a dos resultados obtidos através do modelo de aplicação (Caso 2).

6.1 CASOS DE VALIDAÇÃO

Diversas análises foram feitas sobre a Busca Dispersa a partir da realização do ajuste de histórico dos casos de validação. A primeira análise foi relacionada à qualidade do ajuste obtido que pode ser feita tanto de forma gráfica, com a inspeção das curvas simuladas versus as curvas do histórico, quanto de forma quantitativa, através da proximidade dos parâmetros de ajuste em relação aos valores usados na geração dos dados de histórico. Outras análises interessantes comparam o desempenho do algoritmo frente a mudanças na formulação do problema como, por exemplo:

- Aumento no número de parâmetros de ajuste.
- Aumento no número de níveis discretos utilizados na definição dos parâmetros de ajuste.
- Impacto do uso da computação distribuída em partes do processo.

6.1.1 Caso 1A

Dos resultados da investigação dos diferentes espaços de solução deste caso, apenas os resultados do caso mais refinado (57 níveis) serão apresentados na íntegra. Os demais (29, 15 e 8 níveis) serão discutidos relativamente ao caso mais refinado.

6.1.1.1 Caso 1A – 57 Níveis

A execução da Busca Dispersa encontrou apenas dois ótimos locais neste espaço de soluções. A Tabela 6.1 mostra os valores utilizados nos parâmetros de ajuste para gerar os dados de histórico (valores alvo na tabela) e os valores ótimos locais encontrados.

Tabela 6.1 – Valores encontrados para os parâmetros de ajuste (Caso 1A com 57 níveis).

Parâmetros de ajuste	Valores alvo (histórico)	Melhor solução	2ª melhor solução
<i>KX1 (mD)</i>	1109	1125	1125
<i>KX2 (mD)</i>	5270	5250	5375
<i>KX3 (mD)</i>	2530	2500	2625
<i>KX4 (mD)</i>	7084	7000	7250

Outra visualização para o conjunto de soluções encontradas relativa aos valores alvo encontra-se na Figura 6.1. Observe que as duas soluções situam os parâmetros de ajuste muito próximo dos respectivos valores alvo.

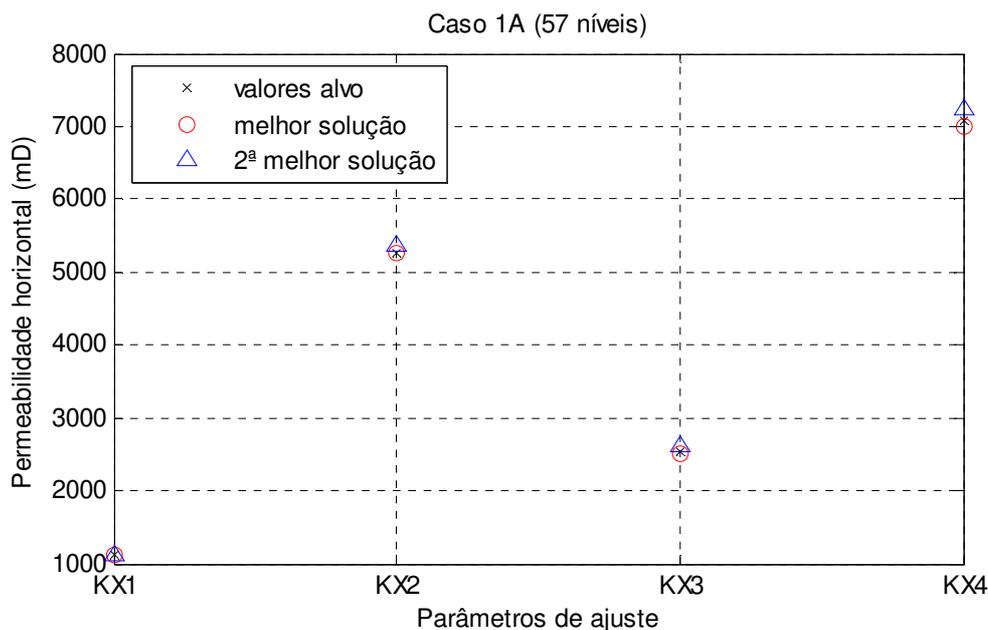


Figura 6.1 – Dispersão das soluções encontradas (Caso 1A com 57 níveis).

As maiores dispersões são percebidas nos parâmetros KX4 e KX3, aqueles parâmetros cujos valores alvo trazem permeabilidades mais elevadas. Contudo, como pode ser visto na Figura 6.2, ambas as soluções fornecem ajustes de boa qualidade para todos os poços produtores.

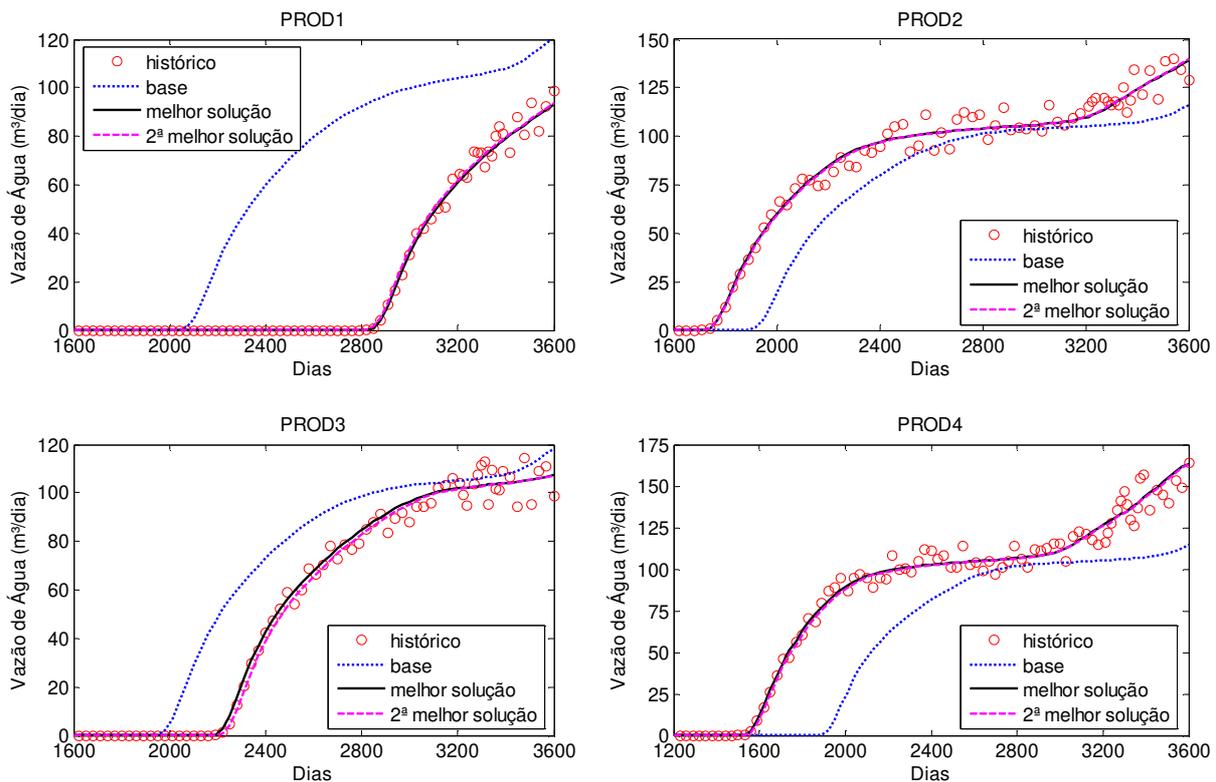


Figura 6.2 – Ajuste da vazão de água nos poços produtores do Caso 1A com 57 níveis.

Com relação ao desempenho da metodologia neste caso, a melhor solução para o problema foi encontrada após 174 das 487 simulações realizadas (Tabela 6.2). Foram gastas 11 simulações na inicialização do método e, em seguida foram gastas uma média de 119 simulações por iteração da Busca Dispersa.

Tabela 6.2 – Desempenho da Busca Dispersa (Caso 1A com 57 níveis).

Iteração	Nº Simulações	Simulações Acumuladas	Valor FO	Melhor FO
0	11	11	0.6403	0.6403
1	163	174	0.0369	0.0369
2	94	268	0.0375	0.0369
3	144	412	0.0369	0.0369
4	75	487	0.0369	0.0369

A Figura 6.3a mostra como foram consumidas as simulações em cada iteração dada pelo método. Em sua inicialização (iteração 0), foram gastas 11 simulações na amostragem inicial para a composição do primeiro conjunto de referência. Na primeira iteração, todas as simulações são devidas ao algoritmo de busca direta, que já parte do melhor ponto obtido durante a amostragem inicial sem gerar soluções combinadas a partir das soluções de referência. Nesta iteração, há um

balanço entre as simulações gastas durante as buscas exploratórias e as lineares. A partir da segunda iteração, todo início de iteração é caracterizado pela realização de quatro simulações dedicadas a gerar um novo ponto de partida para o método de busca direta. Há, durante as iterações, um equilíbrio no número de simulações gastas em buscas lineares e exploratórias.

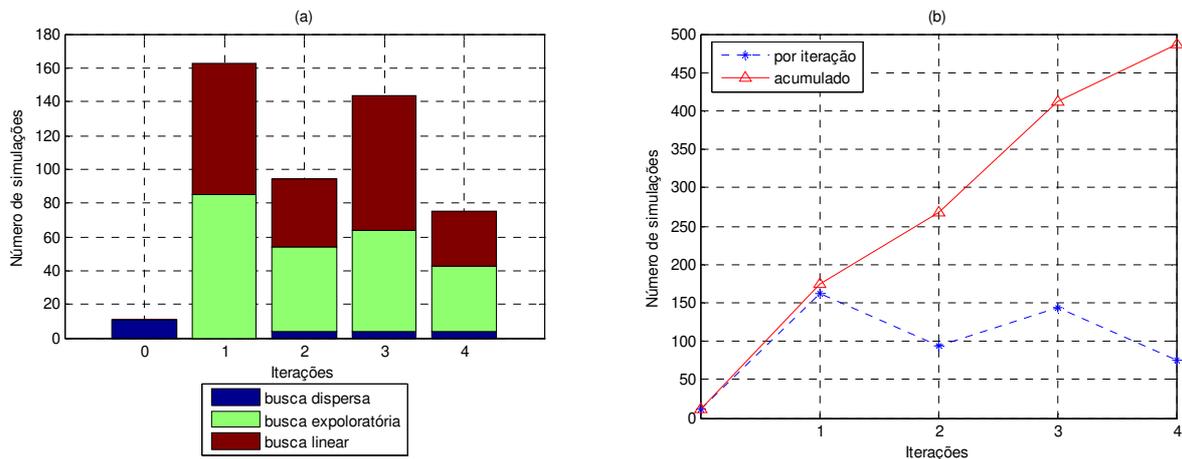


Figura 6.3 – Custo do algoritmo em número de simulações (Caso 1A com 57 níveis).

A Figura 6.3b mostra como o número de simulações cresce com a realização de novas iterações e a também a contribuição de cada iteração neste número. A curva acumulada (vermelha) mostra que o crescimento do número de simulações é aproximadamente linear com o número de iterações. A média neste caso ficou em 119 simulações por iteração.

6.1.1.2 Caso 1A – Análise Comparativa

A Figura 6.4 mostra como a mudança do número de níveis discretos do espaço de soluções impacta o número total de simulações realizadas pelo método e também o valor final encontrado pela função objetivo. Observe que em todas as discretizações a melhoria em relação ao caso base ficou acima de 92% (os valores otimizados da função objetivo ficaram abaixo de 0.08). Porém, quanto maior a discretização do espaço de soluções, melhor é o ajuste medido pela função objetivo.

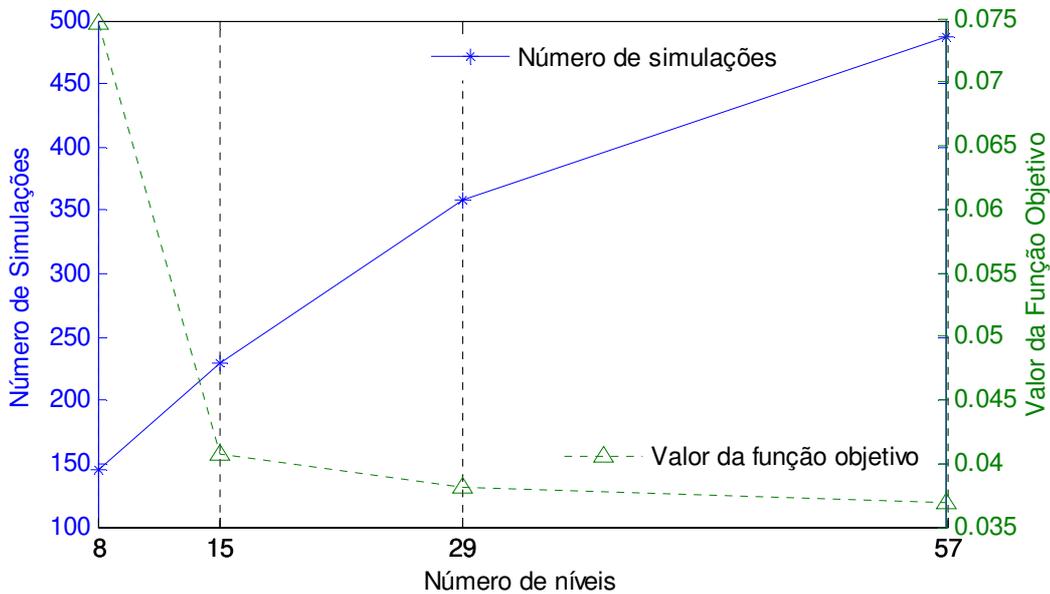


Figura 6.4 – Impacto do número de níveis no número de simulações e no valor da função objetivo (Caso 1A).

Outro aspecto interessante evidenciado pelo gráfico da Figura 6.4 é que à medida que o número de níveis discretos aumenta linearmente, o número de simulações exigido pela Busca Dispersa parece crescer apenas de forma logarítmica. De fato, nos casos analisados, enquanto que a taxa de crescimento do número de níveis aumenta de uma discretização para outra (88%, 93% e 97%), a taxa de crescimento do número de simulações diminui (59%, 55% e 36%). Isto mostra certa escalabilidade do método proposto em relação à discretização assumida. Por outro lado, excetuando-se o ganho no valor da função objetivo ao refinar a discretização de 8 para 15 níveis, não há grandes vantagens em optar por discretizações acima de 15 níveis para esta instância de ajuste.

A Figura 6.5 mostra uma outra visão dos dados de desempenho relativo ao número de simulações realizadas mostrando, não apenas o número final de simulações realizadas, mas como este número foi sendo obtido ao longo das iterações. A Figura 6.5a mostra o desempenho acumulado da Busca Dispersa de acordo com a evolução das iterações enquanto que a Figura 6.5b mostra a variação no custo de cada iteração em termos do número de simulações requerido.

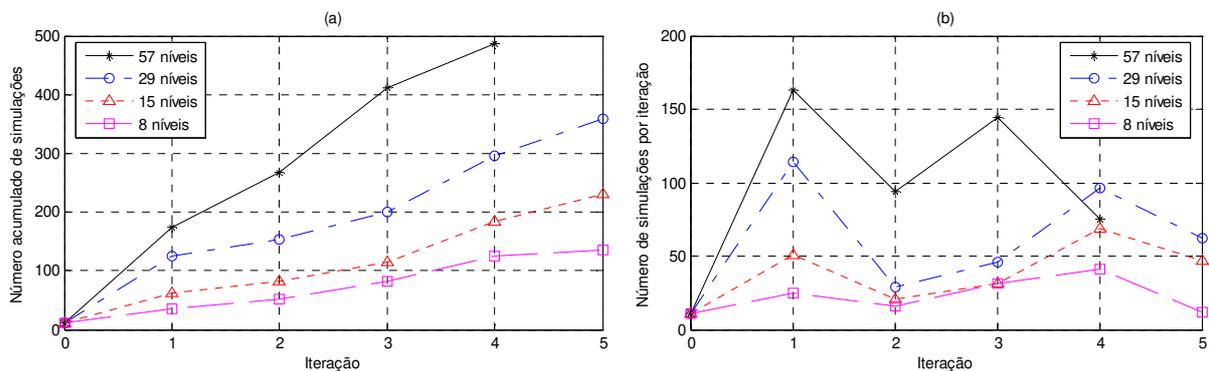


Figura 6.5 – Desempenho da busca dispersa em número de simulações (Caso 1A).

Um resultado interessante deste estudo de caso é que o método encontrou sempre duas soluções otimizadas em todas as discretizações, ou seja, neste caso a redução do número de níveis discretos não parece ter afetado o número de mínimos locais do problema. Os ótimos locais encontrados estão resumidos na Tabela 6.3. Na primeira parte da tabela os valores otimizados da permeabilidade horizontal são mostrados. Na segunda parte da tabela, são mostradas as diferenças entre os valores encontrados e os valores alvo. Já a última linha mostra o valor da função objetivo encontrada em cada caso.

Tabela 6.3 – Soluções encontradas nas diferentes discretizações do espaço de soluções.

Parâmetro	Alvo (mD)	8a (mD)	8b (mD)	15a (mD)	15b (mD)	29a (mD)	29b (mD)	57a (mD)	57b (mD)
KX1	1109	2000	1000	1000	1500	1000	1250	1125	1125
KX2	5270	6000	5000	5000	6000	5250	5500	5250	5375
KX3	2530	3000	2000	2500	3000	2500	2750	2500	2625
KX4	7084	8000	6000	7000	8000	7000	7500	7000	7250
Parâmetro	Alvo (mD)	Diferenças (Alvo – Solução)							
		8a (mD)	8b (mD)	15a (mD)	15b (mD)	29a (mD)	29b (mD)	57a (mD)	57b (mD)
KX1	1109	891	-109	-109	391	-109	141	16	16
KX2	5270	730	-270	-270	730	-20	230	-20	105
KX3	2530	470	-530	-30	470	-30	220	-30	95
KX4	7084	916	-1084	-84	916	-84	416	-84	166
Valor da FO	-	0.0747	-0.0793	-0.0407	0.0544	0.0381	0.0401	-0.0369	-0.0375

A Figura 6.6 mostra como a Busca Dispersa oscilou entre os dois mínimos locais encontrados em cada um dos espaços de soluções explorados. Observe que a Busca Dispersa fez apenas quatro iterações porque as duas soluções encontradas foram obtidas já nas duas primeiras iterações, assim, após duas iterações sem alteração no conjunto de referência, o algoritmo foi interrompido por falta de convergência. Nos demais espaços de solução, houve modificação do conjunto de referência somente na 3ª iteração, daí o método persistir na exploração durante 5 iterações.

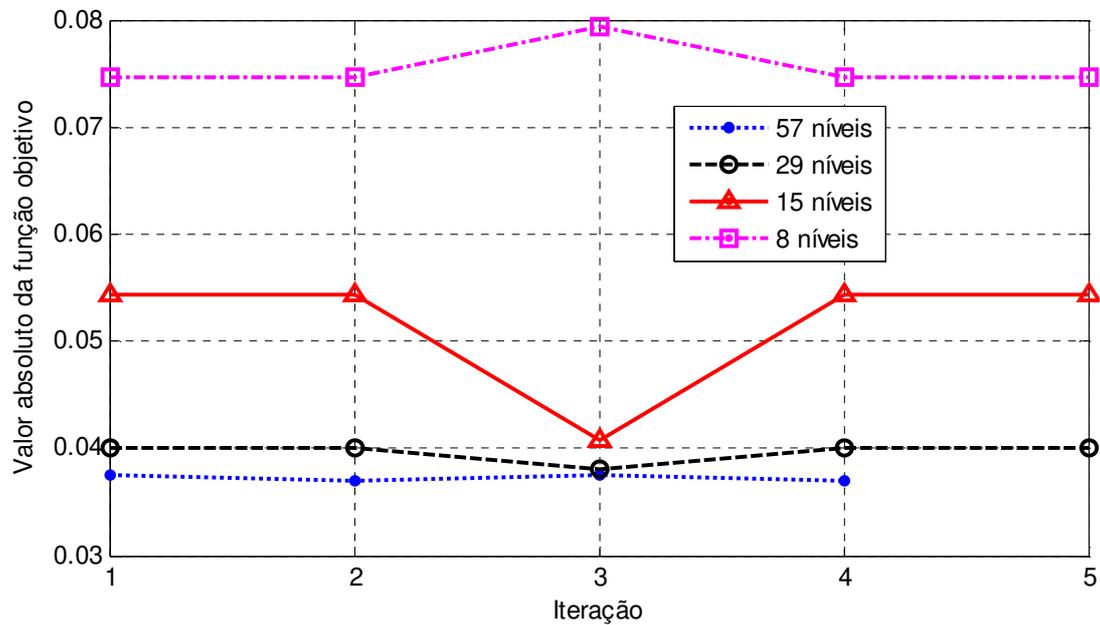


Figura 6.6 – Evolução da função objetivo por iteração do método (Caso 1A).

Uma visão qualitativa dos ajustes da vazão de água obtidos para os diferentes espaços de soluções explorados pode ser visto na Figura 6.7. A curva formada pelos círculos vermelhos representa o histórico de vazão de água. A curva contínua azul mostra o desempenho do caso base. A curva contínua preta mostra o ajuste do caso refinado de 57 níveis, a pontilhada em ciano representa o ajuste do caso de 29 níveis, a pontilhada magenta mostra o ajuste do caso de 15 níveis e, finalmente, a curva pontilhada verde mostra o ajuste caso de 8 níveis.

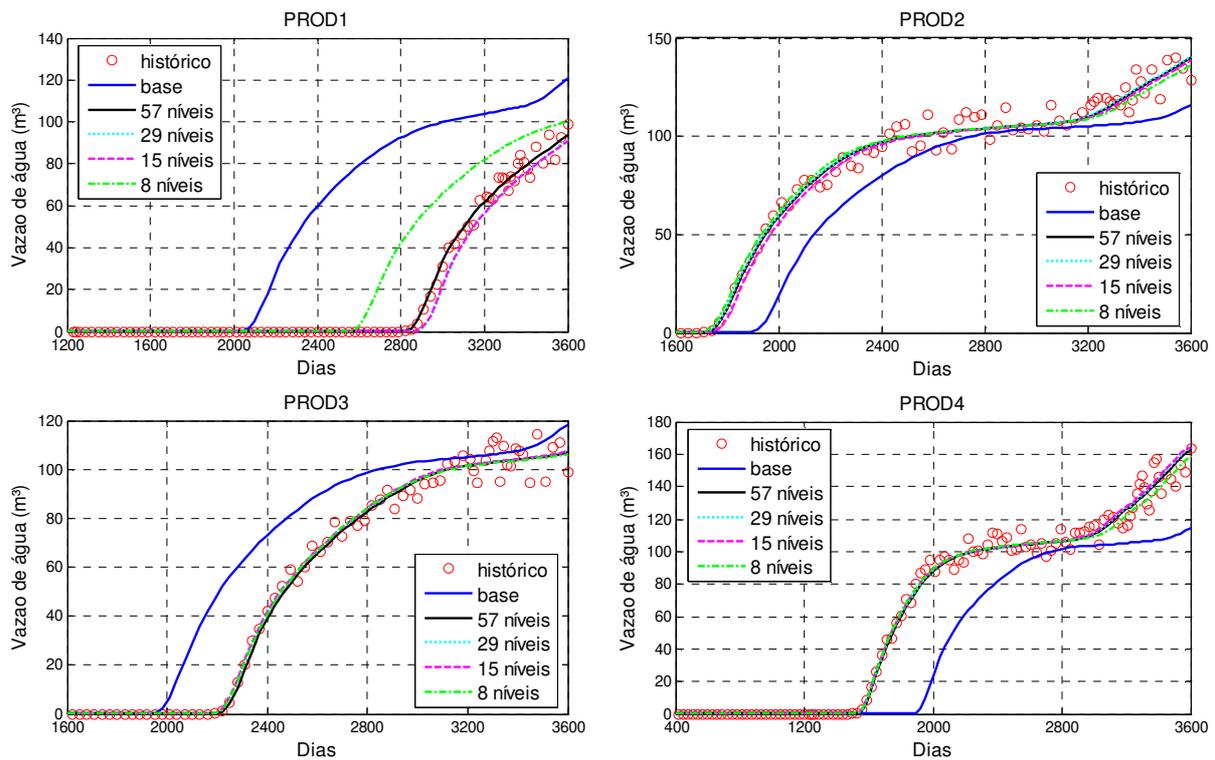


Figura 6.7 – Ajuste das melhores soluções de cada espaço de soluções explorado (Caso 1A).

Observe que todos os poços estão bem ajustados em todos os espaços de soluções investigados, exceto o poço PROD1, cujo espaço de soluções com 8 níveis de discretização não permitiu um ajuste satisfatório com a irrupção de água ocorrendo com cerca de 200 dias de antecedência.

6.1.2 Caso 1B

Desta vez, os parâmetros de ajuste são 8, um para cada área de drenagem entre os pares injetor/produtor dos dois arranjos *five-spot*. Este problema foi otimizado usando quatro espaços de solução distintos cujas discretizações são mostradas na Tabela 6.4.

Tabela 6.4 – Quatro diferentes espaços de soluções para o Caso 1B.

Parâmetro de Ajuste	Valor Mínimo (mD)	Valor Máximo (mD)	Δ (mD) 57 níveis	Δ (mD) 29 níveis	Δ (mD) 15 níveis	Δ (mD) 8 níveis
Kx_i ($i = 1$ a 8)	1000	8000	125	250	500	1000

Seguindo o padrão de apresentação dos resultados, uma descrição detalhada dos resultados obtidos com o espaço de soluções mais refinado é mostrada. Em seguida, os resultados dos demais espaços de soluções são contrastados com este.

6.1.2.1 Caso 1B – 57 níveis

Os valores encontrados para os parâmetros de ajuste estão expressos na Tabela 6.5. Observe que as reduções no valor da função objetivo variaram entre 76% e 83% nas soluções de referência.

Tabela 6.5 - Valores encontrados para os parâmetros de ajuste (Caso 1B com 57 níveis)

Parâmetro de Ajuste	Alvo	Soluções de Referência					
		1ª	2ª	3ª	4ª	5ª	6ª
KX1	6257	5750	5375	6125	5000	4875	4500
KX2	4663	4750	2875	3375	2250	1875	1500
KX3	2256	2875	6125	6375	5875	5500	5375
KX4	1008	1125	1000	1125	1250	1125	1125
KX5	4803	4500	4125	4500	3875	3750	3500
KX6	4119	3625	3375	3750	3125	3000	2625
KX7	3277	3750	4500	4500	4500	4500	4750
KX8	6824	7125	8000	8000	8000	8000	8000
Valor FO	-	-0.1736	-0.1953	0.1970	-0.2105	-0.2175	0.2391

Outra visualização para estes valores é o diagrama de dispersão das soluções encontradas relativo aos valores usados na geração do histórico (Figura 6.8). Os valores alvo estão indicados pelo xis vermelho, enquanto que a melhor solução encontrada é marcada pela cruz negra. Se este diagrama de dispersão for comparado com o do Caso 1A (Figura 6.1), percebe-se que há uma maior dispersão das soluções e isto se deve ao fato de o número de mínimos encontrado ser maior: 6 mínimos locais em 8 iterações versus apenas 2 mínimos locais no primeiro caso. Isto ocorre porque um aumento no número de parâmetros de ajuste provoca um aumento correspondente no número de combinações que levam a um ajuste satisfatório.

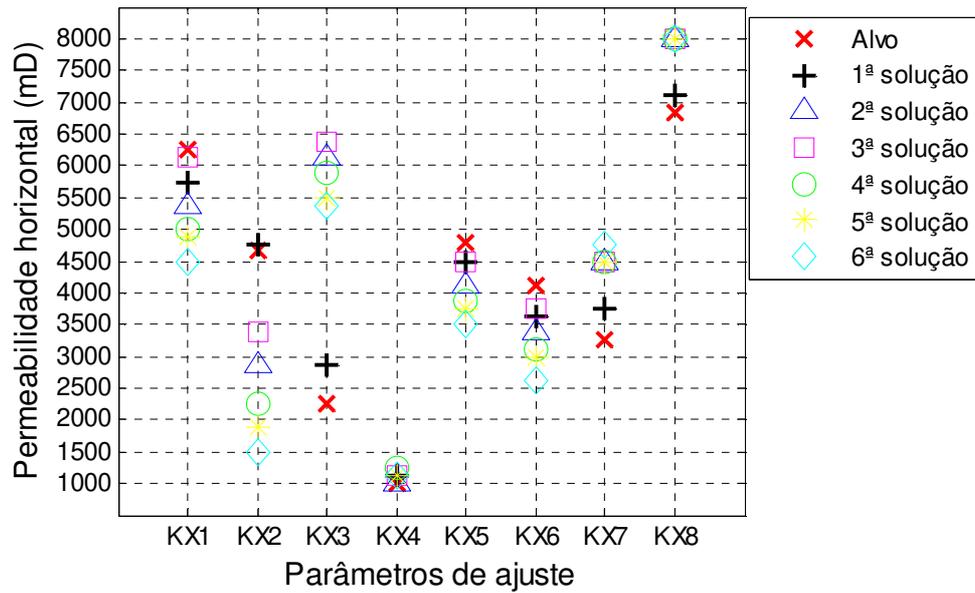


Figura 6.8 – Dispersão das soluções encontradas relativo aos valores geradores do histórico.

Apesar de a dispersão estar mais evidente neste caso (1B) do que no anterior (1A), observando o ajuste qualitativo das soluções de referência (Figura 6.9), percebe-se que todas as soluções encontradas conseguem ajustar bem a vazão de água. Isso mostra que quando o número de parâmetros aumenta, aumenta também o número de soluções viáveis para o problema.

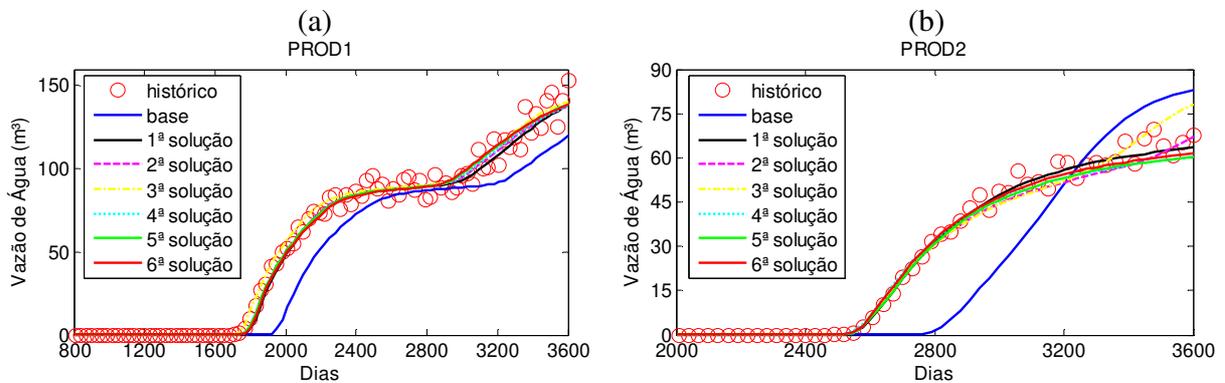


Figura 6.9 - Ajuste da vazão de água nos poços produtores do Caso 1B com 57 níveis.

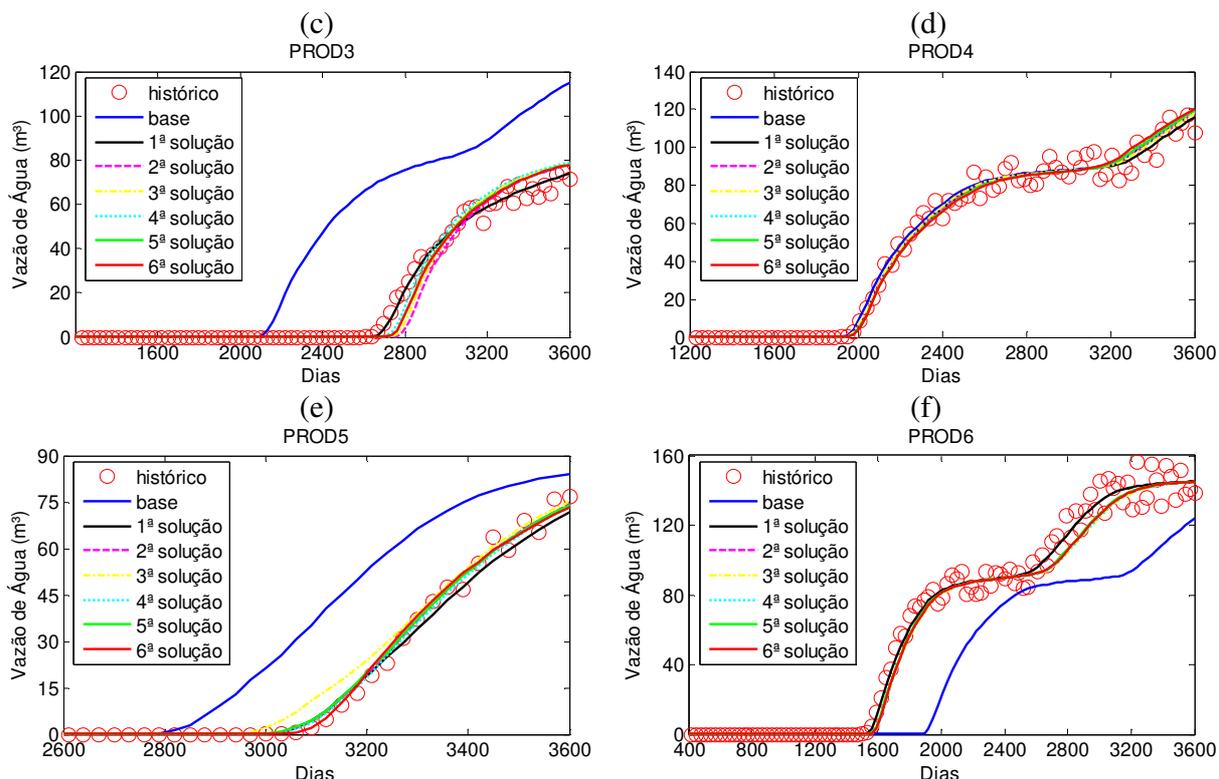


Figura 6.9- Ajuste da vazão de água nos poços produtores do Caso 1B com 57 níveis (continuação).

A Tabela 6.6 traz um resumo do custo de cada iteração da Busca Dispersa, medido em número de simulações, juntamente com o custo acumulado do processo. A inicialização do algoritmo gastou 19 simulações e, a partir daí, em média, foram gastas 282 simulações por iteração somando um total de 1994 simulações. Observe que a melhor solução do método foi obtida logo na 1ª iteração, ou seja, em um processo de ajuste supervisionado, seria possível interromper a execução da busca dispersa após 444 simulações se o engenheiro detectasse que a solução encontrada já era suficientemente boa para o objetivo do ajuste.

Tabela 6.6 – Custo do ajuste do Caso 1B em número de simulações

Iteração	Nº Simulações	Simulações Acumuladas	Valor FO	Melhor FO
0	19	19	1.0000	1.0000
1	425	444	0.1736	0.1736
2	266	710	0.1953	0.1736
3	323	1033	0.2175	0.1736
4	253	1286	0.1970	0.1736
5	288	1574	0.2105	0.1736
6	335	1909	0.1953	0.1736
7	85	1994	0.2391	0.1736

Na Figura 6.10a observa-se que a maior parte das simulações é gasta na etapa de busca exploratória do algoritmo de busca local. Este resultado é um pouco diferente do observado no Caso 1A onde havia certo balanço entre as os dois tipos de busca (linear e exploratória). Isto mostra que os espaços de soluções com maior número de parâmetros de ajuste, ou seja, com um maior número de eixos, tendem a ser mais complexos necessitando um maior número de avaliações da direção de descida seguida pela busca linear. A predominância da busca exploratória como principal fonte de simulações no método também explica o crescimento linear no número acumulado de simulações na Figura 6.10b. Observando a curva azul da Figura 6.10b, é possível perceber que a primeira iteração é a mais custosa do processo. A partir da segunda iteração, algumas soluções visitadas pelo processo de otimização local já foram simuladas de forma que o custo cai para um patamar mais baixo. Finalmente, a sétima e última iteração começa sua busca a partir de uma solução próxima a um mínimo local previamente visitado e por este motivo seu custo foi reduzido em relação às demais iterações.

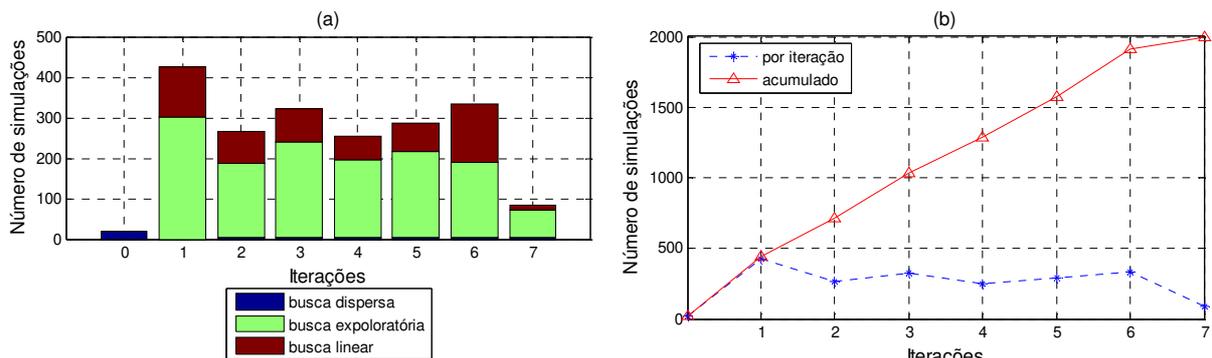


Figura 6.10 – Custo da metodologia para o Caso 1B de 57 níveis em número de simulações

6.1.2.2 Caso 1B – Análise Comparativa

Nesta seção são comparadas as soluções obtidas para o ajuste do Caso 1B considerando a exploração dos 4 diferentes espaços de solução descritos na Tabela 6.4. A Figura 6.11 mostra como se comportou a Busca Dispersa ao explorar os quatro diferentes espaços de soluções definidos. Observe que as mesmas tendências observadas para o Caso 1A (Figura 6.4) se repetem neste caso: (1) o aumento linear no número de níveis discretos dos parâmetros de ajuste não provoca um aumento linear no número total de simulações dadas pelo método, mas um crescimento que se ajusta mais com um crescimento logarítmico; (2) quanto mais refinado for o

espaço de soluções, melhor o desempenho do método em termos do melhor valor encontrado para a função objetivo.

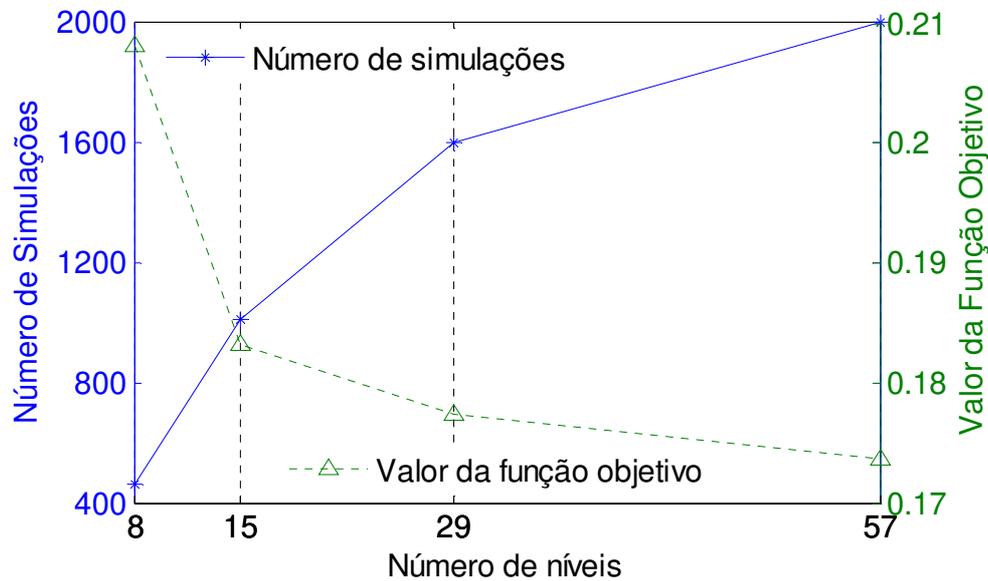


Figura 6.11 - Impacto do número de níveis no número de simulações e no valor da função objetivo (Caso 1B)

Observe no gráfico da Figura 6.11 que as melhores soluções obtidas em todos os espaços de soluções representam uma melhoria na função objetivo de 79% a 83% em relação ao caso base. Enquanto que no Caso 1A (Figura 6.4) não se recomendava usar uma discretização maior que 15 níveis para aquele caso, aqui a escolha é mais difícil, pois os mínimos encontrados com 15, 29 e 57 não estão mais relativamente tão próximos. A escolha neste caso depende da quantidade de recursos à disposição do estudo.

A Figura 6.12 mostra como cada iteração contribuiu para o custo total do método em termos do número de iterações. Cada curva representa o desempenho da Busca Dispersa ao explorar um dos quatro espaços de solução explorados. Em ambos os gráficos da Figura 6.12, pode ser observado que a Busca Dispersa atingiu o número limite de iterações apenas no espaço de soluções formado por 29 níveis discretos. O menor número de iterações foi realizado no caso de 15 níveis discretos. Em todos os casos, o crescimento do número de simulações com as iterações foi aproximadamente linear (Figura 6.12a). Pode ser visto na Figura 6.12b que o número médio de simulações por iteração do método é tanto maior quanto maior é o número de níveis discretos utilizados na confecção do espaço de soluções. Observe também que à medida

que o algoritmo se aproxima das últimas iterações, o número de simulações por iteração cai mostrando que o caminho percorrido pelo algoritmo já possui diversas soluções simuladas.

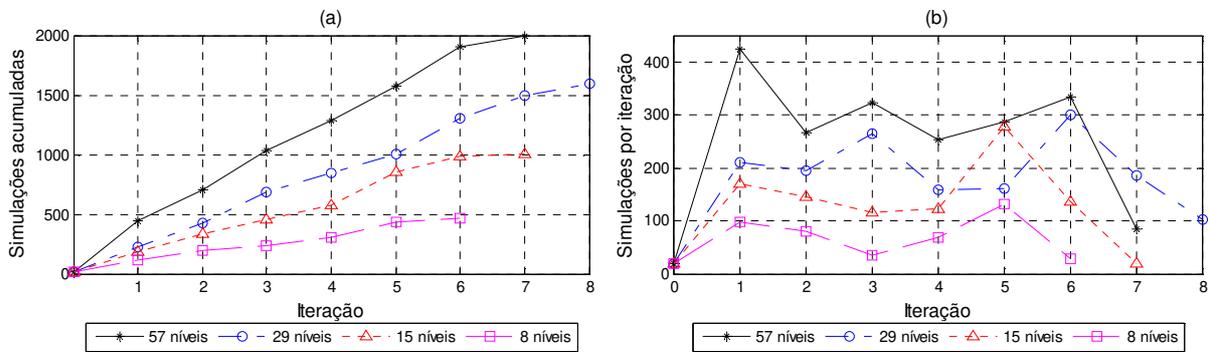


Figura 6.12 – Desempenho da busca dispersa em número de simulações (Caso 1B)

A Tabela 6.7 mostra as duas melhores soluções encontradas nos quatro espaços de soluções analisados. A primeira parte da tabela mostra os valores obtidos para a permeabilidade horizontal dos oito parâmetros de ajuste em contraste com os valores alvo, que geraram o histórico de produção. A segunda parte da tabela mostra o quanto os valores encontrados se distanciam do valor alvo. Finalmente, a última linha mostra o valor da função objetivo das soluções apresentadas.

Tabela 6.7 - Soluções encontradas nas diferentes discretizações do espaço de soluções (Caso 1B)

Parâmetro	Alvo (mD)	8a (mD)	8b (mD)	15a (mD)	15b (mD)	29a (mD)	29b (mD)	57a (mD)	57b (mD)
KX1	6257	7000	6000	7000	6000	5750	5750	5750	5375
KX2	4663	6000	3000	5500	5000	5000	5000	4750	2875
KX3	2256	4000	5000	4000	3500	3500	3750	2875	6125
KX4	1008	1000	1000	1500	1000	1250	1250	1125	1000
KX5	4803	5000	4000	5000	4500	4500	4500	4500	4125
KX6	4119	4000	3000	4000	3500	3500	3500	3625	3375
KX7	3277	4000	2000	4000	4000	4000	4250	3750	4500
KX8	6824	8000	8000	8000	7500	7500	8000	7125	8000
Parâmetro	Alvo (mD)	Diferenças (Solução - Alvo)							
		8a (mD)	8b (mD)	15a (mD)	15b (mD)	29a (mD)	29b (mD)	57a (mD)	57b (mD)
KX1	6257	743	-257	743	-257	-507	-507	-507	-882
KX2	4663	1337	-1663	837	337	337	337	87	-1788
KX3	2256	1744	2744	1744	1244	1244	1494	619	3869
KX4	1008	-8	-8	492	-8	242	242	117	-8
KX5	4803	197	-803	197	-303	-303	-303	-303	-678
KX6	4119	-119	-1119	-119	-619	-619	-619	-494	-744
KX7	3277	723	-1277	723	723	723	973	473	1223
KX8	6824	1176	1176	1176	676	676	1176	301	1176
Valor da FO	-	-0.2082	-0.2774	0.1832	-0.1834	-0.1774	-0.1805	-0.1736	-0.3905

A Figura 6.13 mostra como as soluções que compõem o conjunto de referências nos quatro espaços de soluções foram localizadas ao longo das iterações. Um dado interessante desta figura é que em todos os casos, exceto o de 15 níveis, a melhor solução foi encontrada na primeira iteração. A melhor solução do caso de 15 níveis foi encontrada na terceira iteração. Como foi visto na Seção 6.1.2.1, o caso de 57 níveis teve 6 soluções distintas encontradas em 7 iterações da Busca Dispersa. Já no caso de 29 níveis foram encontradas 6 soluções distintas em 8 iterações, no de 15 níveis, 5 soluções em 7 iterações e no de 8, apenas 3 soluções distintas em 6 iterações.

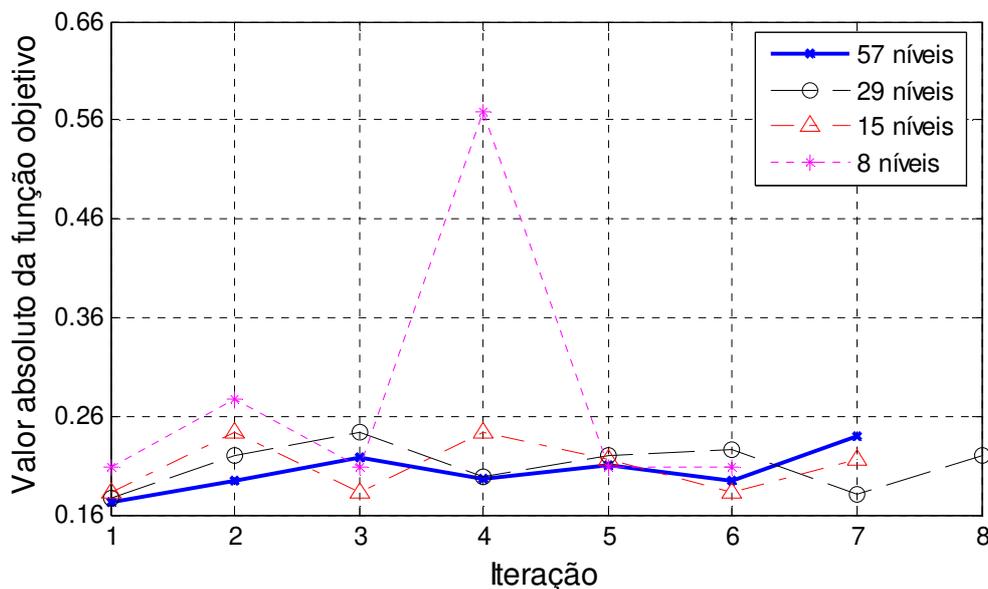


Figura 6.13 - Evolução da função objetivo por iteração do método (Caso 1B)

A Figura 6.14 mostra o melhor ajuste obtido para a vazão de água em cada um dos espaços de soluções explorados. Os poços PROD1, PROD4, PROD5 e PROD6 tiveram um bom ajuste independente da discretização utilizada. O poço PROD2 ficou bem ajustado apenas com os espaços mais refinados (29 e 57 níveis). Com a discretização caindo para 15 níveis, o poço PROD2 fica desajustado a partir de 3300 dias (de um total de 3650 de histórico); com 8 níveis, o tempo de irrupção de água é antecipado e o ajuste não é satisfatório. Já o poço PROD3 também teve um bom ajuste com espaços de 57 e 29 níveis discretos, mas os espaços menos refinados (8 e 15 níveis) conseguiram gerar um ajuste satisfatório, porém, de qualidade inferior aos espaços mais refinados.

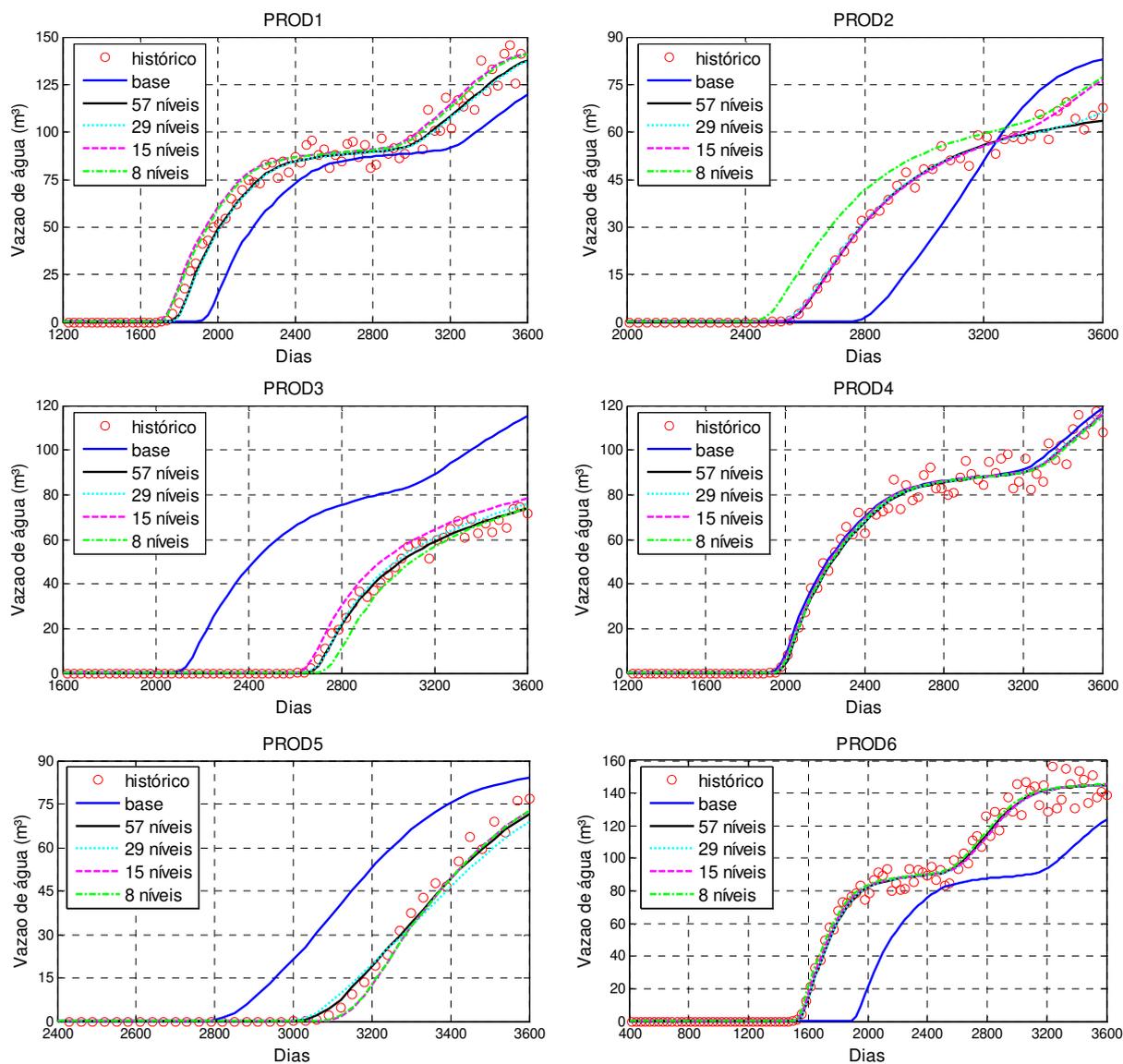


Figura 6.14 – Ajuste das melhores soluções de cada espaço de soluções explorado (Caso 1B)

6.1.3 Caso 1C

A descrição dos resultados obtidos neste caso seguirá o padrão definido nas seções anteriores, com os resultados do espaço de soluções refinado sendo descritos em detalhes enquanto que os dos demais espaços de soluções serão descritos relativamente a ele.

A Busca Dispersa explorou quatro espaços de soluções diferentes do Caso 1C, cujas discretizações encontram-se definidas na Tabela 6.8.

Tabela 6.8 – Quatro diferentes espaços de soluções para o Caso 1C.

Parâmetro de Ajuste	Valor Mínimo (mD)	Valor Máximo (mD)	Δ (mD) 57 níveis	Δ (mD) 29 níveis	Δ (mD) 15 níveis	Δ (mD) 8 níveis
Kx_i (i = A a P)	1000	8000	125	250	500	1000

Observe que os espaços de soluções foram dimensionados de forma que um caso mais refinado sempre contém todas as soluções dos casos menos refinados. Desta maneira, é possível avaliar o impacto da discretização sobre o número e qualidade dos mínimos.

6.1.3.1 Caso 1C – 57 Níveis

A Busca Dispersa realizou um total de 8 iterações nas quais 7 soluções otimizadas distintas foram encontradas (vide Tabela 6.9). Pode ser observado que, em relação ao afastamento do caso base, as soluções encontradas conseguiram reduzir o afastamento de 96% a 98%.

Tabela 6.9 - Valores encontrados para os parâmetros de ajuste (Caso 1C com 57 níveis).

Parâmetros de Ajuste	Alvo	Soluções de Referência							
		1ª	2ª	3ª	4ª	5ª	6ª	7ª	8ª
KXA	1009	1000	1375	1500	1375	1875	2000	2000	1875
KXB	4945	6250	6250	7000	6625	6875	7000	7375	7250
KXC	1253	4125	5125	5375	5000	5125	4625	5000	4750
KXD	1900	2500	2625	2750	2750	2625	2625	2625	2625
KXE	5095	5250	5250	5125	5250	5875	5750	6000	6125
KXF	4359	6875	6750	6875	6875	7125	7250	7250	7375
KXG	3452	6250	6750	6500	6875	6125	6000	5625	6375
KXH	7500	8000	8000	8000	8000	8000	8000	8000	8000
KXI	5800	6500	6750	6500	6500	6625	6750	6625	6500
KXJ	6226	6625	7125	6625	6500	6875	6750	6750	7000
KXK	2219	6750	6125	7000	7375	7000	7125	7250	7250
KXL	2500	3625	2625	3625	4000	4250	4250	5000	4750
KXM	2100	2250	2500	2250	2250	2500	2375	2500	2625
KXN	4595	4750	4875	4750	4750	4875	4875	4875	4875
KXO	3128	4500	4250	4500	4625	4625	4625	4750	4750
KXP	1105	1750	1625	1750	1875	1875	1875	2000	2000
Valor FO		0.0205	0.0237	0.0239	0.0246	0.0348	0.0339	0.0410	0.0423

A Figura 6.15 mostra um diagrama de dispersão dos parâmetros de ajuste em relação aos valores geradores do histórico (alvo). Observe que as soluções aproximaram bem certos parâmetros (KXI, KXJ, KXM e KXN), porém outros valores alvo parecem ter sido ignorados pela Busca Dispersa (KXB, KXC, KXF, KXG e KXK). Outra observação interessante é que houve uma tendência geral do processo de superestimar a permeabilidade alvo. Acredita-se que a ausência de um ajuste combinado de pressão no campo tenha sido responsável por esta tendência.

Caso o impacto disso fosse observado na qualidade do ajuste de pressão, isso poderia ser corrigido com um passo adicional no processo de ajuste, mas esse não é o objetivo da etapa de validação.

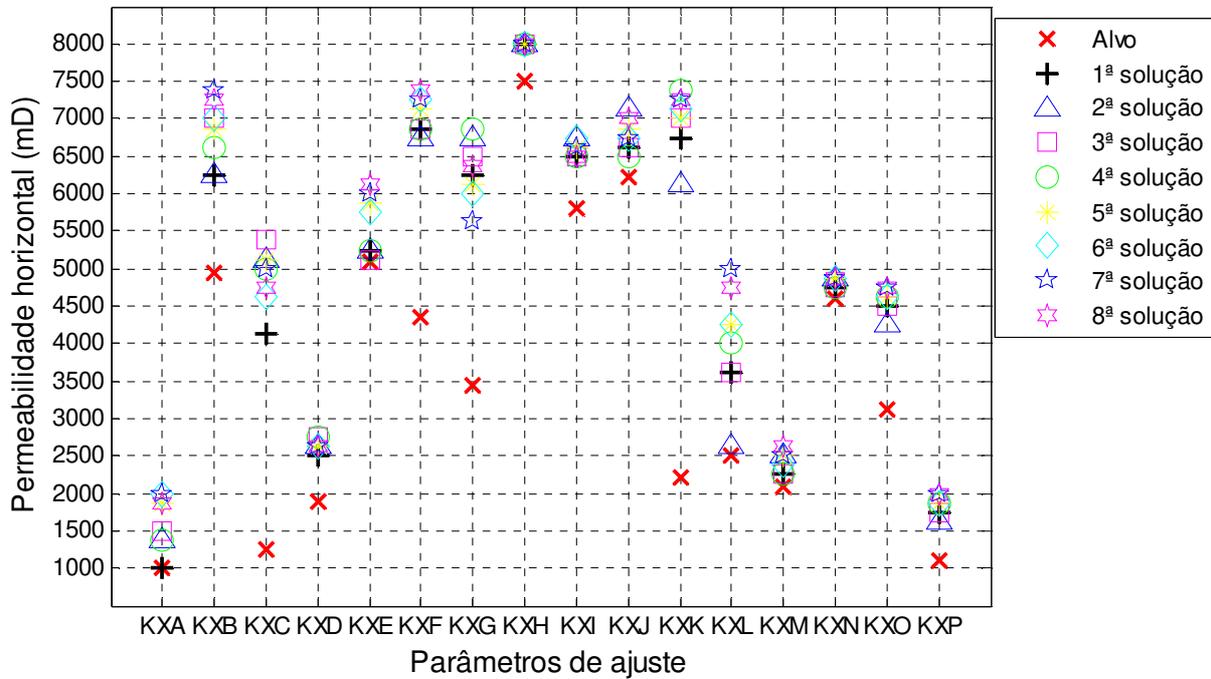


Figura 6.15 – Dispersão das soluções de referência (Caso 1C com 57 níveis)

Estes resultados parecem sugerir que a metodologia não convergiu na direção do ajuste dos poços, mas observe nos gráficos da Figura 6.16a à Figura 6.16e que o ajuste da vazão de água dos poços está, em geral, bem razoável. Os poços menos ajustados são o PROD6 (Figura 6.16c), seguido do poço PROD1 (Figura 6.16), onde apenas a 2ª melhor solução encontrada conseguiu um ajuste satisfatório.

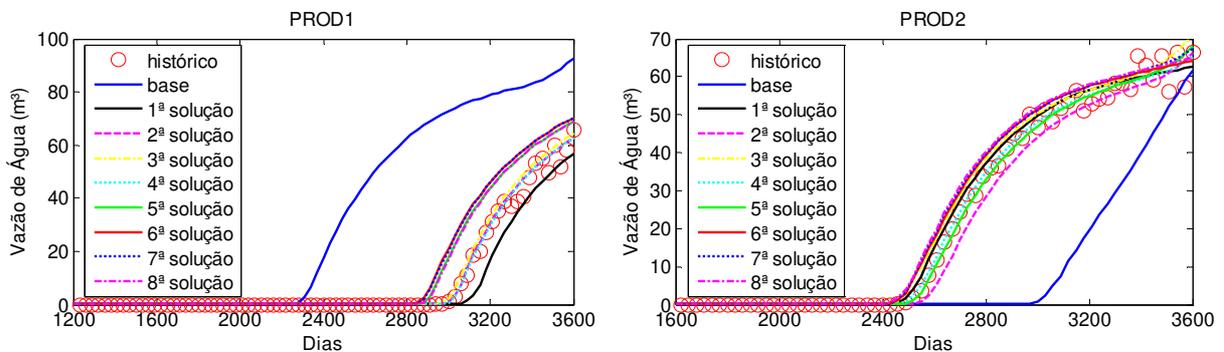


Figura 6.16 – Ajuste da vazão de água nos poços do Caso 1C com 57 níveis.

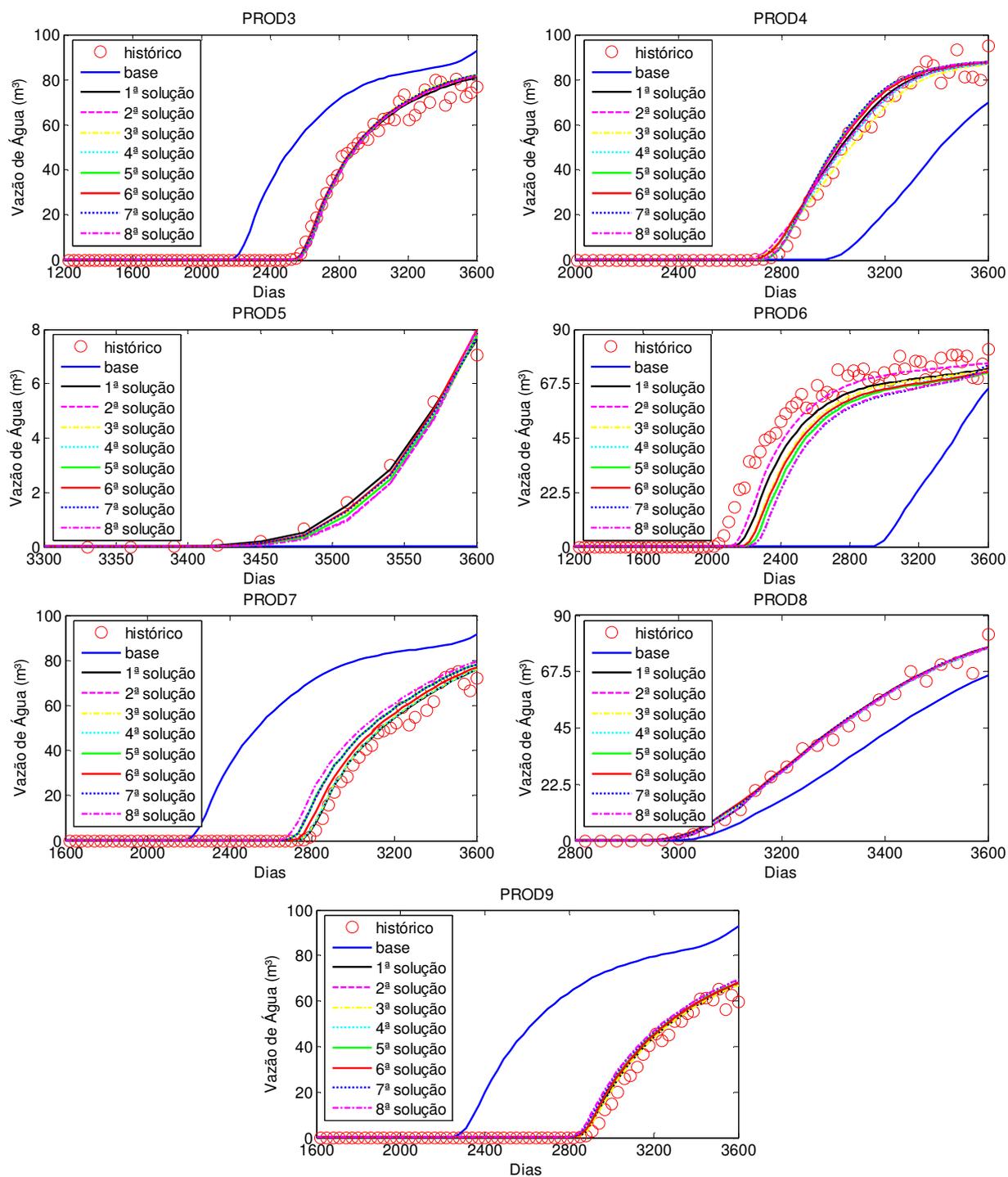


Figura 6.16 – Ajuste da vazão de água nos poços do Caso 1C com 57 níveis (continuação).

Os problemas encontrados neste ajuste devem-se aos seguintes fatores:

1. O ponto final de convergência de cada Busca Direta realizada durante as iterações da Busca Dispersa depende apenas do ponto de partida fornecido

pela lógica do método e da topologia do espaço de soluções. Sendo o espaço de soluções fixo durante uma determinada execução da Busca Dispersa, é responsabilidade do método fornecer pontos de partida suficientemente diversificados de forma a explorar o espaço de soluções de forma eficiente. Esta tarefa se torna mais difícil em problemas com um grande número de variáveis como este.

2. Esta metodologia de ajuste de histórico está baseada em uma função objetivo construída a partir das diferenças ao quadrado dos dados simulados em relação aos observados. A composição da função objetivo global com 9 poços possui um comportamento altamente não linear o que gera espaços de solução complicados com diversos mínimos locais. Este fato pôde ser comprovado com a localização de 8 mínimos locais distintos encontrados a partir de 8 pontos de partida também distintos. Neste contexto, se algum ponto de partida mais próximo da solução alvo tivesse sido gerado durante as iterações do método, a solução alvo teria sido encontrada.
3. Como a topologia do espaço de soluções é definida pela função objetivo, é importante que ela consiga expressar adequadamente quais termos do ajuste são mais importantes que outros. A função objetivo deste caso é simplesmente uma média aritmética dos afastamentos da vazão de água dos 9 poços envolvidos no ajuste. Assim, pequenos desajustes em um ou outro poço não servem como alerta para o método mudar radicalmente a direção do ajuste. Em casos como este, a definição de um bom ponto de partida é essencial para o sucesso do método.
4. Finalmente, quando o caso base está muito desajustado em relação aos demais pontos do espaço de soluções, quase todas as direções tomadas pelo método constituem direções de descida e a convergência do método na direção do mínimo global fica dificultada. Em outras palavras, quando todas as soluções ótimas locais se parecem quando comparadas sob a ótica da função objetivo, fica mais difícil encontrar a melhor delas.

Seguindo o padrão desenvolvido na apresentação dos resultados dos casos 1A e 1B, a seção a seguir apresenta uma análise comparativa das soluções obtidas durante a exploração deste e dos

demaís espaços de soluções referentes ao Caso 1C. Porém, como o ajuste automático não ficou muito preciso em alguns poços, este caso terá uma seção adicional mostrando os resultados de um procedimento de ajuste assistido que conseguiu obter um resultado melhorado para o caso com o benefício adicional de uma grande economia no número de simulações realizadas.

6.1.3.2 Caso 1C – Análise Comparativa

Observe no gráfico da Figura 6.17 que a qualidade da melhor solução obtida pelo método tem uma grande dependência do número de níveis discretos escolhido. Seguindo uma tendência observada nas análises anteriores (Figura 6.4 e Figura 6.11) há uma grande vantagem em termos de função objetivo ao aumentar a discretização de 8 para 15 níveis (redução de 37% no valor da função objetivo). O que não se observou anteriormente foi o grande benefício de aumentar a discretização de 29 para 57 níveis (49% de redução no valor da função objetivo).

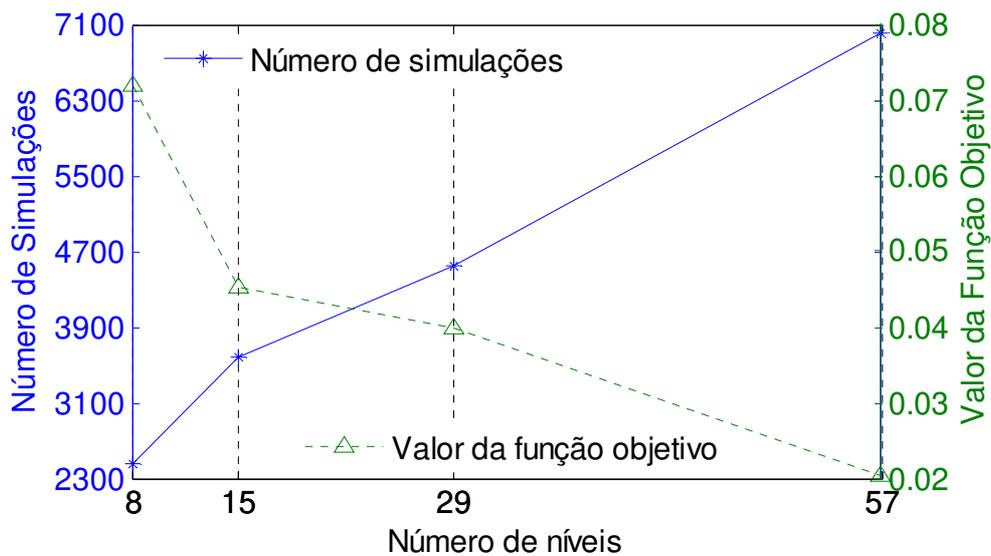


Figura 6.17 - Impacto do número de níveis no número de simulações e no valor da função objetivo (Caso 1C).

O que também é diferente deste caso para os demais é a tendência do crescimento do número de iterações ao refinar a discretização. Nos casos anteriores o aumento era aproximadamente logarítmico, mas neste caso o padrão foi aproximadamente linear. Isto se deveu ao fato de todos os espaços de soluções terem sido explorados com o número máximo de iterações, ou seja, o número de soluções ótimas locais cresceu tanto com um maior número de

parâmetros de ajuste que o critério de parada que observa a falta de evolução da qualidade do conjunto de referência (Seção 4.5.5) não foi acionado.

A Figura 6.18 mostra como as melhores soluções obtidas a partir dos quatro espaços de soluções explorados se localizam relativamente aos valores alvo. Novamente, pode ser observada a tendência geral de superestimação das permeabilidades. Este problema provavelmente pode ser resolvido com a inclusão das pressões de fundo de poço na função objetivo, porém o programa desenvolvido ainda não suportava este dado de histórico e não houve tempo hábil para sua inclusão neste trabalho.

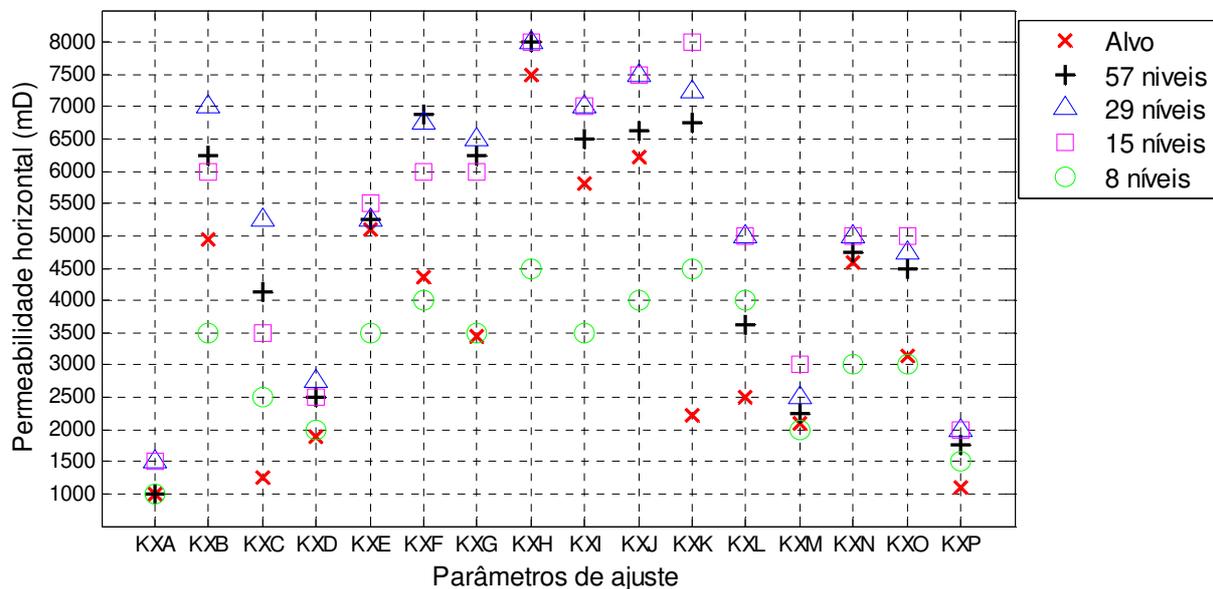


Figura 6.18 – Dispersão das soluções encontradas em relação aos valores geradores do histórico (Caso 1C).

A Tabela 6.10 mostra as duas melhores soluções encontradas nos quatro espaços de soluções explorados. A primeira parte da tabela mostra os valores dos parâmetros de ajuste na solução alvo e nas soluções encontradas enquanto que a segunda parte da tabela mostra a distância entre a solução encontrada e a solução alvo. Valores positivos indicam que a solução encontrada está acima do valor alvo e os negativos indicam que ela está abaixo da solução alvo.

Note que, exceto pela melhor solução do Caso 1C de 8 níveis, as soluções encontradas estão basicamente acima dos valores alvo como já foi percebido no diagrama de dispersão do Caso 1C de 57 níveis (Figura 6.15) e também no diagrama de dispersão contendo as melhores soluções obtidas na exploração dos quatro espaços de soluções (Figura 6.18).

Tabela 6.10 – Valores e afastamentos encontrados para as duas melhores soluções de cada espaço.

Parâmetro	Alvo (mD)	8a (mD)	8b (mD)	15a (mD)	15b (mD)	29a (mD)	29b (mD)	57a (mD)	57b (mD)
KXA	1009	1000	2000	1500	1500	1500	1500	1000	1375
KXB	4945	3500	7000	6000	6500	7000	6500	6250	6250
KXC	1253	2500	5000	3500	5500	5250	5000	4125	5125
KXD	1900	2000	3000	2500	2500	2750	2500	2500	2625
KXE	5095	3500	6000	5500	5000	5250	5250	5250	5250
KXF	4359	4000	7000	6000	6500	6750	6750	6875	6750
KXG	3452	3500	6000	6000	6000	6500	6000	6250	6750
KXH	7500	4500	8000	8000	8000	8000	8000	8000	8000
KXI	5800	3500	7000	7000	7000	7000	7000	6500	6750
KXJ	6226	4000	6000	7500	7500	7500	7000	6625	7125
KXK	2219	4500	7000	8000	7500	7250	7750	6750	6125
KXL	2500	4000	8000	5000	5000	5000	5000	3625	2625
KXM	2100	2000	3000	3000	3000	2500	2750	2250	2500
KXN	4595	3000	5000	5000	5000	5000	5000	4750	4875
KXO	3128	3000	5000	5000	5000	4750	5000	4500	4250
KXP	1105	1500	2000	2000	2000	2000	2250	1750	1625
Parâmetro	Alvo (mD)	Diferenças (Solução-Alvo)							
		8a (mD)	8b (mD)	15a (mD)	15b (mD)	29a (mD)	29b (mD)	57a (mD)	57b (mD)
KXA	1009	-9	991	491	491	491	491	-9	366
KXB	4945	-1445	2055	1055	1555	2055	1555	1305	1305
KXC	1253	1247	3747	2247	4247	3997	3747	2872	3872
KXD	1900	100	1100	600	600	850	600	600	725
KXE	5095	-1595	905	405	-95	155	155	155	155
KXF	4359	-359	2641	1641	2141	2391	2391	2516	2391
KXG	3452	48	2548	2548	2548	3048	2548	2798	3298
KXH	7500	-3000	500	500	500	500	500	500	500
KXI	5800	-2300	1200	1200	1200	1200	1200	700	950
KXJ	6226	-2226	-226	1274	1274	1274	774	399	899
KXK	2219	2281	4781	5781	5281	5031	5531	4531	3906
KXL	2500	1500	5500	2500	2500	2500	2500	1125	125
KXM	2100	-100	900	900	900	400	650	150	400
KXN	4595	-1595	405	405	405	405	405	155	280
KXO	3128	-128	1872	1872	1872	1622	1872	1372	1122
KXP	1105	395	895	895	895	895	1145	645	520
Valor da FO	-	0.0719	0.0787	0.0453	0.0582	-0.0400	0.0445	0.0205	0.0237

A Figura 6.19 mostra a seqüência em que as soluções de referência foram sendo encontradas pela Busca Dispersa nos diferentes espaços de soluções. Observe que as soluções do espaço menos refinado (8 níveis) estão quase sempre piores que os demais espaços enquanto que as soluções encontradas no espaço mais refinado (57 níveis) estão todas melhores que as das demais soluções. Já as soluções nos espaços de 15 e 29 níveis estão muito parecidas com uma pequena vantagem para o espaço de soluções de 29 níveis.

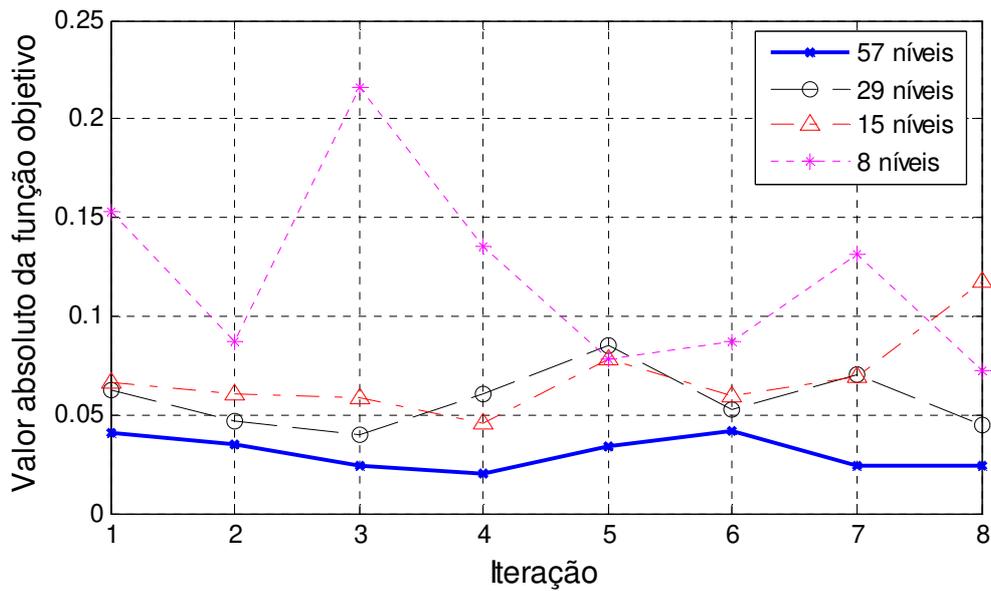


Figura 6.19 – Evolução da função objetivo durante a exploração dos espaços de soluções.

Uma observação interessante sobre o gráfico da Figura 6.19 é que a melhor solução é usualmente obtida na primeira metade do método e, depois de obtido este limite inferior, o algoritmo não consegue ultrapassá-lo dentro do limite de iterações imposto (8 iterações).

Os gráficos da Figura 6.20a à Figura 6.20e mostram o melhor ajuste obtido pela Busca Dispersa nos quatro espaços de soluções explorados. Exceto pelo poço PROD1, a Busca Dispersa foi mais precisa no ajuste ao explorar o espaço de soluções mais refinado (57 níveis).

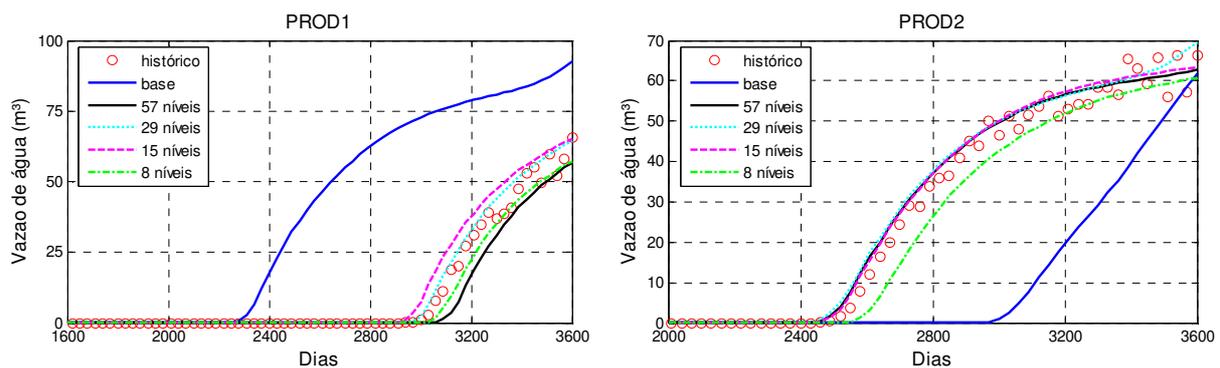


Figura 6.20 – Ajuste da vazão de água dos poços do Caso 1C.

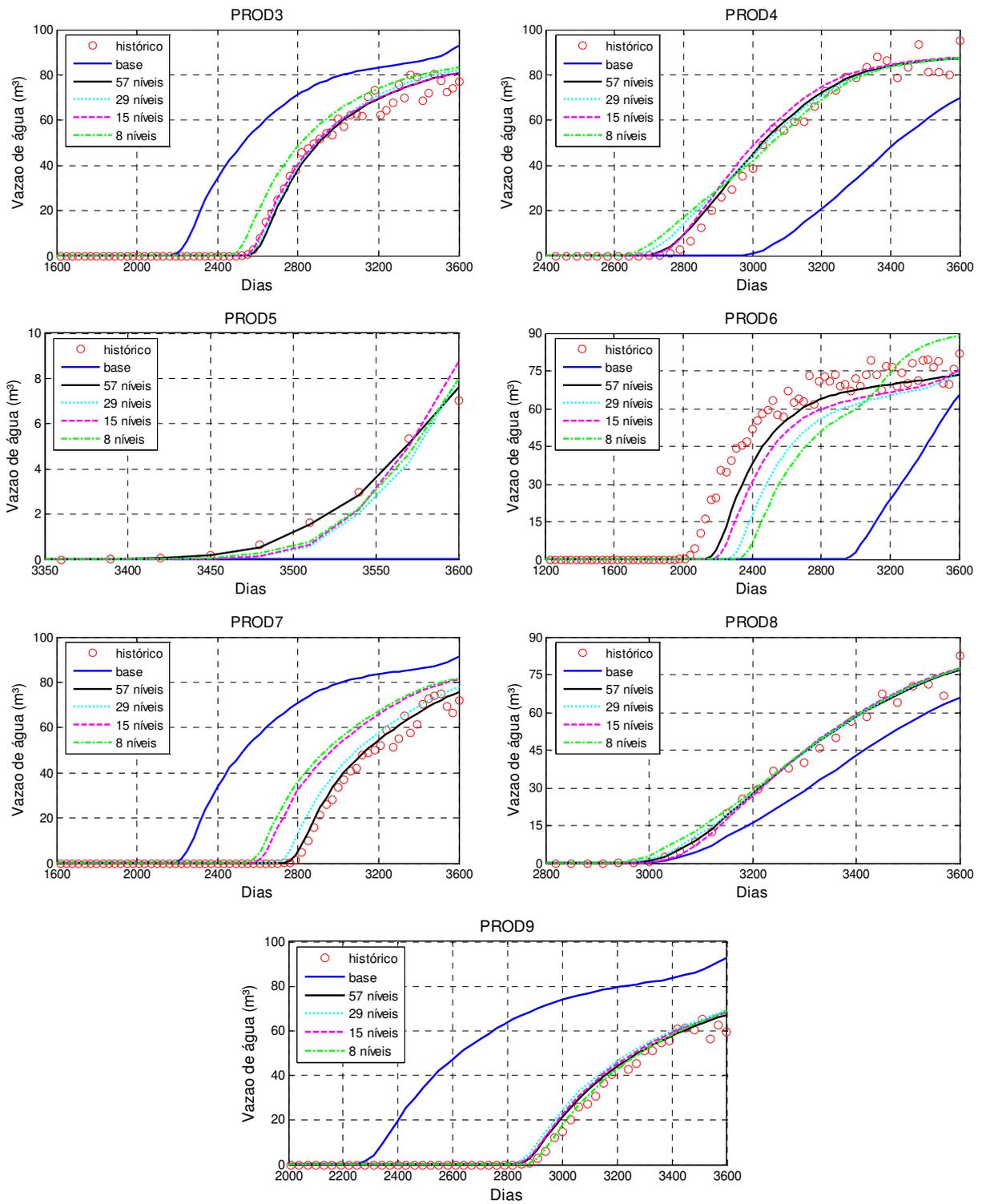


Figura 6.20 – Ajuste da vazão de água dos poços do Caso 1C (continuação)

O que pôde ser observado em todos os casos analisados até o momento (1A, 1B e 1C) é que há uma tendência geral de melhoria da solução encontrada à medida que mais simulações são

realizadas, ou seja, que o espaço de soluções é mais explorado. Isto pode ser observado na evolução do valor da função objetivo ao longo das iterações da Busca Dispersa (Figura 6.6, Figura 6.13 e Figura 6.19). Este comportamento é esperado uma vez que não há, a princípio, uma forma de saber a priori qual é a região do espaço de soluções que conterá as melhores soluções para o problema. Assim, quanto melhor explorado for o espaço de soluções, maiores serão as chances de descoberta de soluções de melhor qualidade. Em resumo, existe uma relação direta entre o número de simulações realizadas e a qualidade das soluções encontradas.

6.2 CASO 1C – AJUSTE ASSISTIDO

Como já foi citado anteriormente, os resultados obtidos no procedimento automático para o caso 1C não ficaram tão bons quanto os obtidos nos casos com menos parâmetros de ajuste (casos 1A e 1B, com 4 e 8 variáveis, respectivamente). Isso está de acordo com o observado na prática e é um dos motivos pelos quais não se usa procedimentos totalmente automáticos quando o número de atributos cresce. Como o espaço de soluções cresce muito com o aumento do número de atributos, ou aumenta muito o número de simulações ou a probabilidade de se encontrar as melhores soluções diminui.

Outra análise importante dos resultados obtidos para o ajuste automático do Caso 1C é que em 8 iterações da Busca Dispersa, 7 soluções ótimas locais foram encontradas. Este resultado sugere que talvez o número de iterações realizadas não foi o suficiente para explorar adequadamente o espaço de soluções. Neste momento há que se considerar realizar mais iterações da metodologia ou então mudar a estratégia de otimização para um procedimento que faz uso de informações adicionais do problema que não podem ser capturadas pela função objetivo, como a relação existente entre determinados parâmetros de ajuste e a geometria do reservatório permitindo a partição do problema global em subproblemas de menor porte, por exemplo.

Como o número de simulações realizado de forma automática para o Caso 1C já foi bastante elevado, optou-se por adotar um procedimento de ajuste assistido usando a metodologia da Busca Dispersa para dividir o espaço de soluções global em espaços menores e mais

adequados para investigação automática. A idéia é usar o conhecimento das características do reservatório de modo a limitar o espectro das buscas realizadas pela metodologia proposta.

Ao adotar uma técnica de ajuste assistido, as suposições do engenheiro que faz o ajuste necessariamente eliminam algumas possibilidades de exploração do espaço de soluções. Assim, se as suposições excluïrem boas soluções do espaço de soluções, o ajuste assistido fica comprometido. Foi isto que ocorreu em duas tentativas de ajuste assistido deste caso que não estão descritas neste trabalho. A descrição da tentativa de ajuste assistido que deu certo segue.

A Figura 6.21 mostra de forma esquemática o reservatório do Caso 1C com ênfase nos parâmetros de ajuste e no esquema de produção. Os poços produtores são numerados da esquerda para a direita e de cima para baixo (PROD1 a PROD9) e os injetores seguem a mesma convenção (INJE1 a INJE4). Para uma visão mais detalhada do reservatório, refira-se à Figura 5.5.

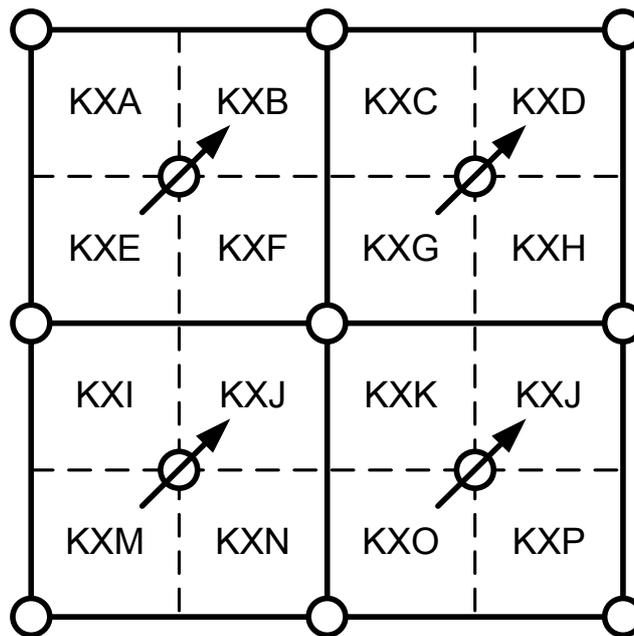


Figura 6.21 – Representação esquemática do caso base do ajuste assistido

Observe que os únicos poços que tem uma área de drenagem dependente diretamente de apenas um poço injetor são os das extremidades (regiões de drenagem KXA, KXD, KXM e KXP). Assim, a primeira fase deste ajuste consistiu em determinar o valor de permeabilidade horizontal destas regiões como mostra o esquema esquerdo da Figura 6.22. As vazões de água dos poços produtores marcados em vermelho estão presentes na função objetivo, os demais não.

Uma observação é que as demais regiões de permeabilidade em cada arranjo *five-spot* foram agrupadas em uma única região (somando mais quatro parâmetros de ajuste) de forma que uma permeabilidade equivalente seja computada pela Busca Dispersa de forma a melhor ajustar os poços extremos. A ausência desta contrapartida, quando se deixou o restante das regiões em cada *five-spot* no caso base foi a razão de falha de uma das tentativas anteriores de ajuste assistido.

O espaço de soluções utilizado nesta primeira fase do ajuste assistido foi composto pelos 8 parâmetros de ajuste da Figura 6.22 com uma discretização de 29 níveis no intervalo de permeabilidades de 1000 mD a 8000 mD. O resultado deste ajuste aparece esquematizado no diagrama direito da Figura 6.22, com as regiões marcadas em amarelo fixas até o final do ajuste global.

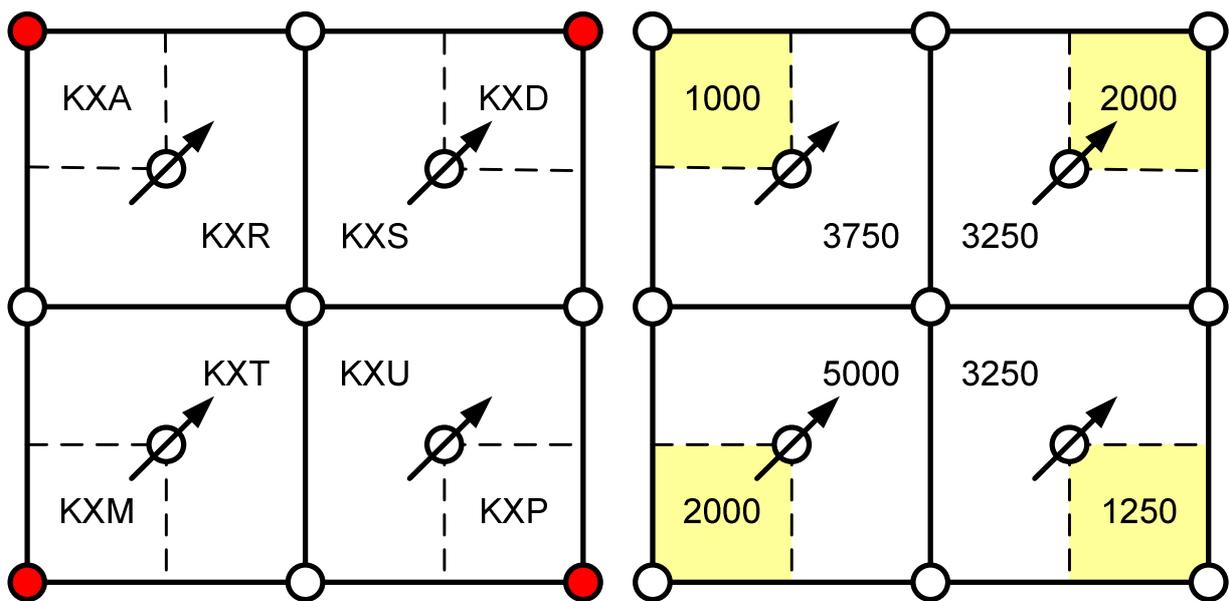


Figura 6.22 – Espaço de soluções e resultado do ajuste na Fase 1

Como o poço central (PROD5) possui um histórico de produção de água pouco significativo, se iniciando aproximadamente a partir de 3400 dias num horizonte de 3650 dias e atingindo um limite de produção inferior a 8 m³/dia, optou-se por ajustar na Fase 2 as regiões de drenagem dos poços externos (PROD2, PROD4, PROD6 e PROD8) como mostra o diagrama esquerdo da Figura 6.23. As áreas de drenagem do poço central (PROD5) permaneceram fixas nos valores encontrados ao final da Fase 1.

Por sua vez, a função objetivo manteve os poços incluídos na Fase 1 e adicionou as vazões de água dos novos poços envolvidos no processo. Os poços cujo histórico de produção de água foram incluídos na função objetivo aparecem em vermelho no diagrama esquerdo da Figura 6.23. O resultado desta etapa do ajuste aparece no diagrama direito da Figura 6.23, e as regiões marcadas em amarelo estão fixadas nos valores apresentados até o fim do ajuste global. O espaço de soluções novamente foi composto dos parâmetros de ajuste variando no intervalo de 1000 mD a 8000 mD com 29 níveis discretos.

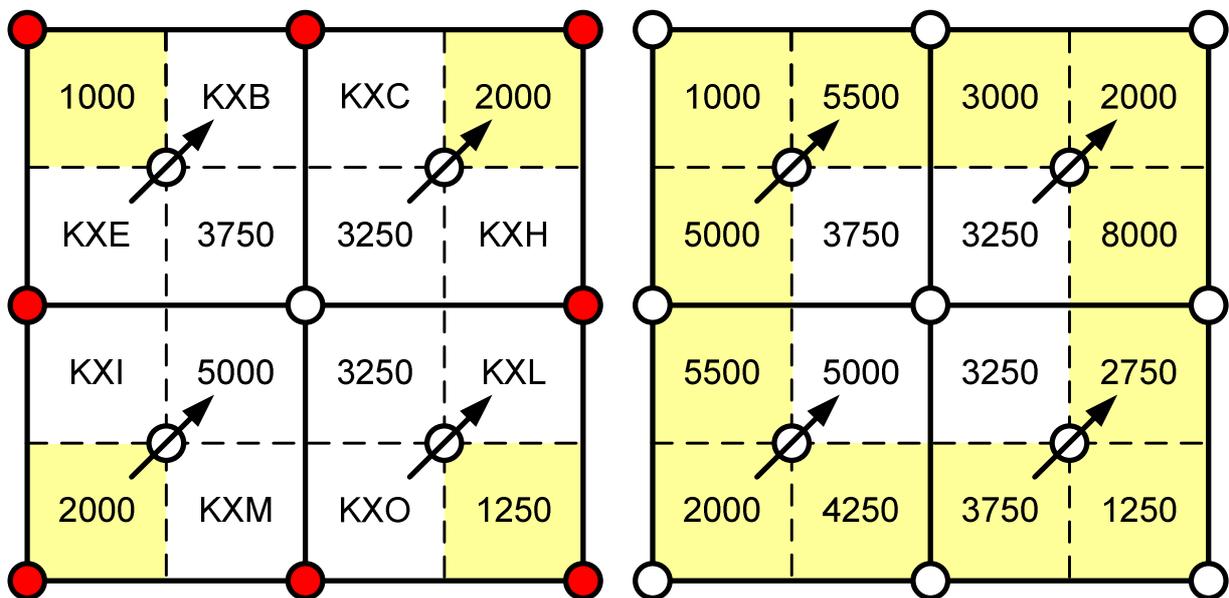


Figura 6.23 – Espaço de soluções e resultado do ajuste na Fase 2

Na última fase do ajuste, restavam fixar as 4 regiões remanescentes que afetam a produção de água do poço central (PROD5), assim essas regiões foram convertidas novamente em parâmetros de ajuste como mostra o diagrama esquerdo da Figura 6.24.

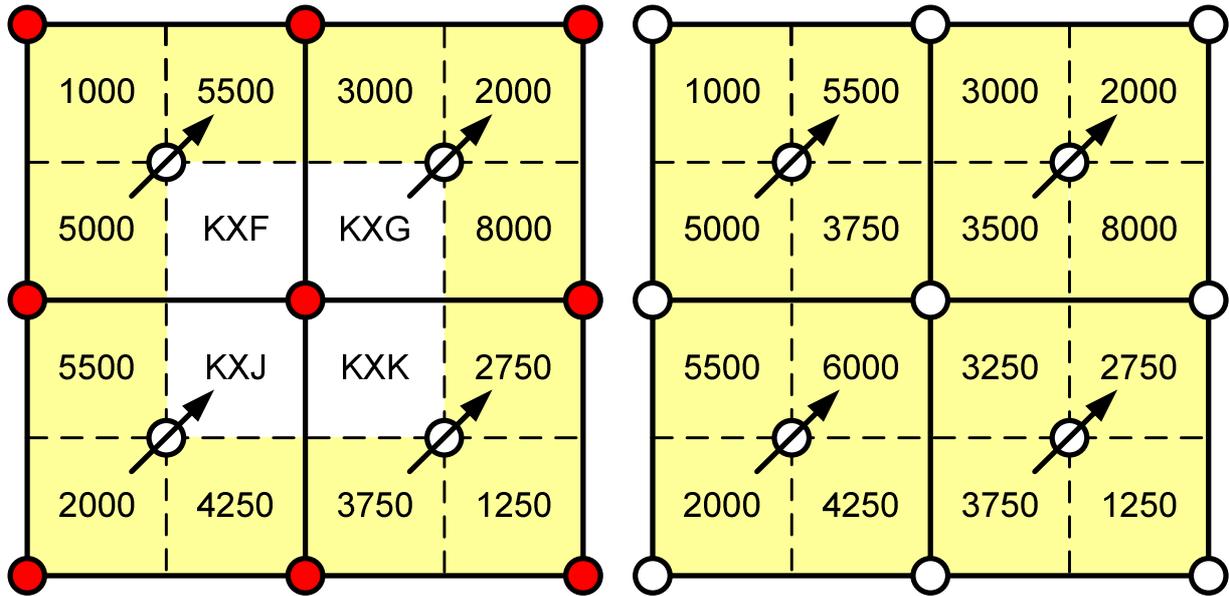


Figura 6.24 – Espaço de soluções e resultado do ajuste na Fase 3

Observe que, nesta fase, o histórico de vazão de água de todos os poços foi incluído na função objetivo. O resultado final desta última etapa aparece no diagrama direito da Figura 6.24.

O ajuste da vazão de águas nos poços pode ser visto nos gráficos da Figura 6.25. Observe que os poços que integraram a função objetivo na Fase 1 (PROD1, PROD3, PROD7 e PROD9) encontram-se ajustados já ao final da Fase 1 e não se desajustam ao final das demais fases. Fato semelhante ocorre com os poços alvos da otimização durante a Fase 2 (PROD2, PROD4, PROD6 e PROD8). Estes ficam ajustados após o final da Fase 2 e não se desajustam na fase seguinte. Finalmente, o poço central (PROD5) só se ajusta ao final da última fase de otimização. Observe, porém, que este poço é pouco importante para o ajuste global do reservatório uma vez que seu histórico é muito curto e as vazões envolvidas são muito menores que as dos demais poços.

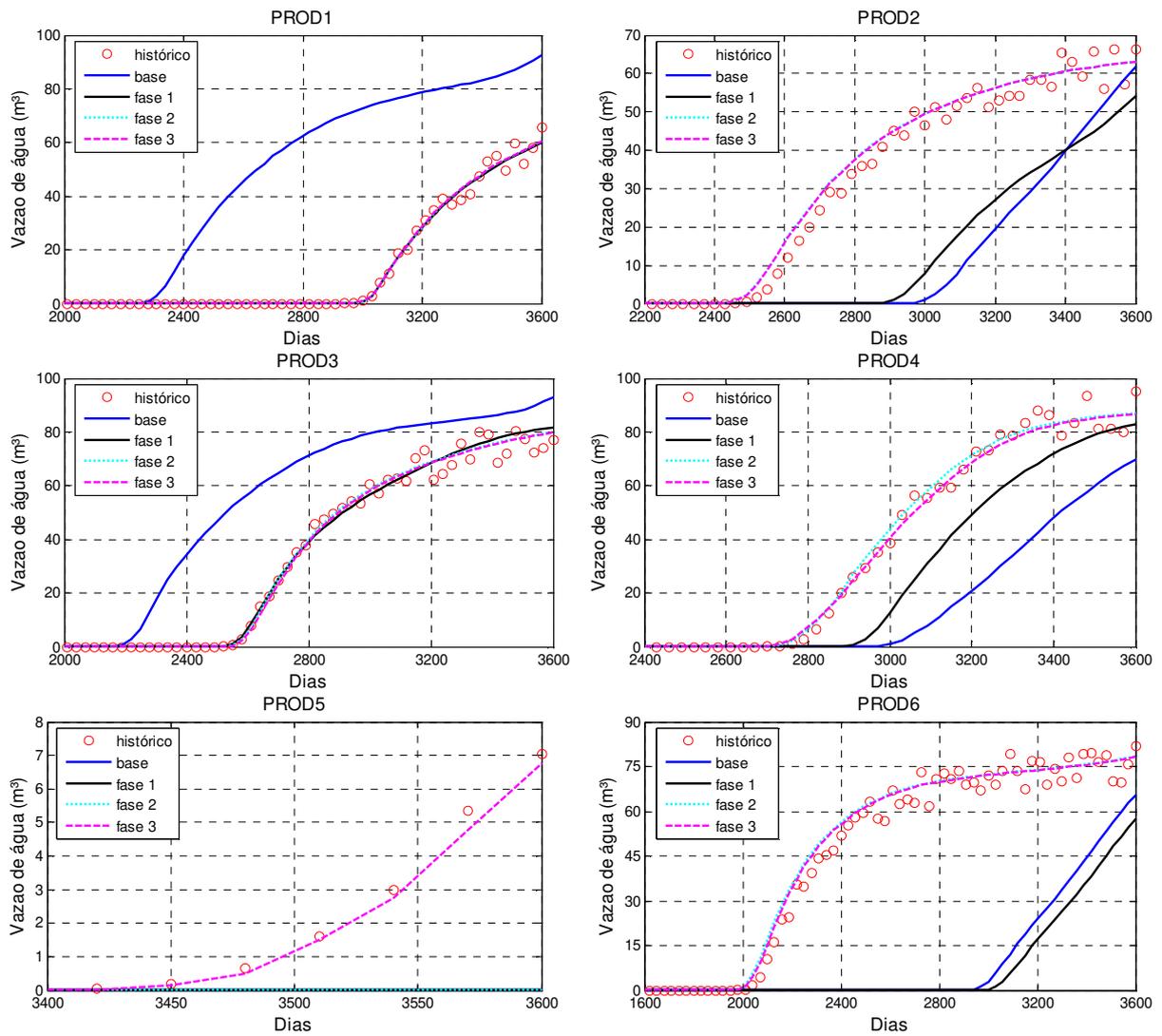


Figura 6.25 – Ajuste da vazão de água no Caso 1C assistido

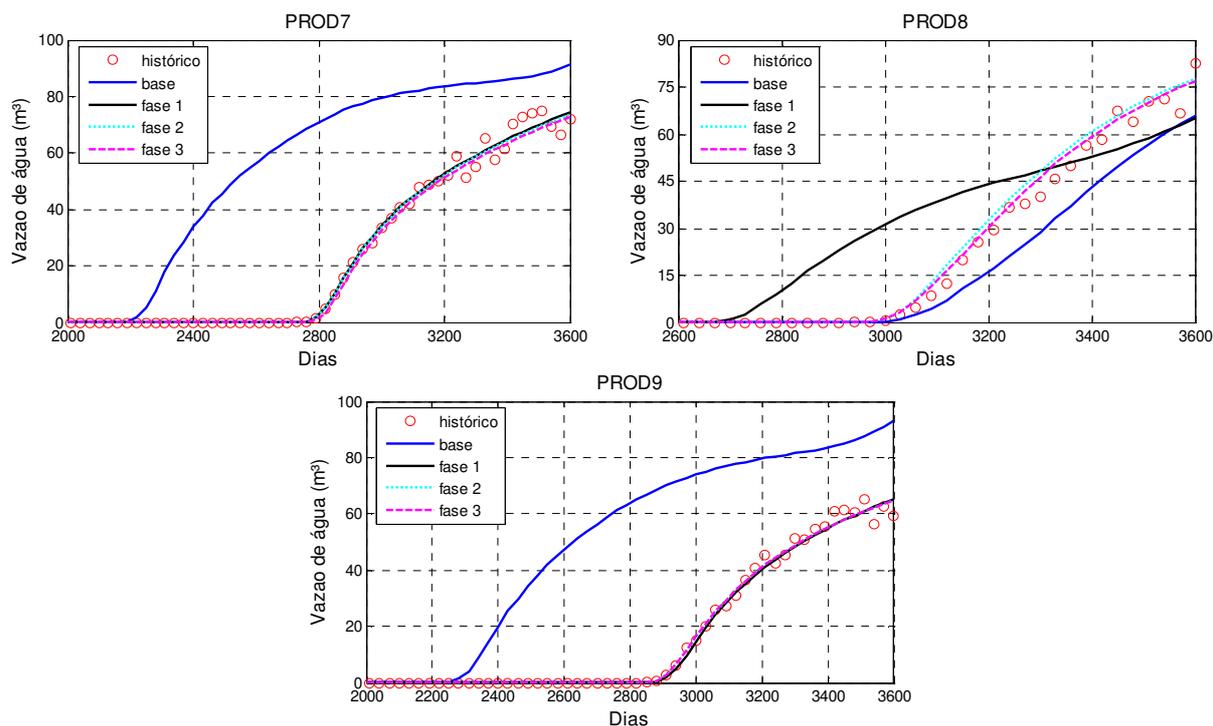


Figura 6.25 – Ajuste da vazão de água no Caso 1C assistido (continuação)

Ao todo, este processo envolveu 3 fases de ajuste, as duas primeiras com 8 parâmetros de ajuste com discretização de 29 níveis enquanto que a última fase contou com apenas 4 parâmetros de ajuste e 29 níveis. O resumo do espaço de soluções utilizado nas fases do ajuste assistido juntamente com seus respectivos custos encontram-se na Tabela 6.11.

Tabela 6.11 – Espaço de soluções e custo das fases do ajuste assistido

Fase	Parâmetros	Níveis	Número de Simulações
1	8	29	892
2	8	29	1632
3	4	29	211
Total			2735

Observe que, potencialmente, há uma grande economia a ser feita ao usar informações que não podem ser fornecidas ao método automático. No ajuste completamente automático, envolvendo todos os 16 parâmetros de ajuste, o caso com 29 níveis executou 4571 simulações enquanto que o de 57 níveis executou 7004 simulações, e nenhum deles conseguiu chegar à qualidade de ajuste do uso assistido da metodologia. É importante lembrar que houve duas outras tentativas de ajuste assistido que não deram certo e suas iterações não foram contabilizadas. Por

este motivo, a experiência do engenheiro de reservatórios é um fator chave nos ganhos de produtividade do ajuste assistido.

Pode-se ressaltar ainda que o procedimento assistido poderia ter sido utilizado também nos Casos 1A e 1B, mas o ganho potencial da parte assistida diminui à medida que o número de atributos diminui. Dessa forma, pode-se afirmar que para casos mais simples, com um menor número de variáveis, deve-se adotar a tendência de um processo mais automático e para casos mais complexos, com um maior número de atributos, a parte assistida passa a ser importante para viabilizar melhores soluções. O mesmo tipo de raciocínio poderia ser feito com relação ao tempo computacional de uma simulação; com simulações mais rápidas, pode-se usar processos mais automáticos que demandam um maior número de avaliações da função-objetivo; para caso mais complexos, com simulações mais demoradas, a parte assistida passa a ter uma maior importância.

A extrapolação do raciocínio anterior poderia levar à conclusão que para casos reais com muitos atributos incertos e cujo tempo computacional é, em geral, grande o processo deveria ser conduzido de forma manual, como se observa em muitos casos da literatura e como praticado muitas vezes pela Indústria. Entretanto, a proposta do ajuste assistido deve levar em conta que o processo pode ser dividido em partes menores possíveis de serem automatizadas como é o caso da aplicação do método no Caso 2, a seguir.

6.3 CASO DE APLICAÇÃO (CASO 2)

Este caso está descrito na Seção 5.2. Considerando que a dinâmica do reservatório original é baseada em canais de permeabilidade, todas as tentativas de ajuste neste caso são baseadas em modificações na permeabilidade horizontal. Uma diferença adicional deste caso com relação aos demais é que há problemas de injetividade no campo, onde alguns poços injetores não conseguem honrar o histórico de injeção de água.

Inicialmente, um ajuste global das permeabilidades foi tentado usando a Equação 6.1. Como o reservatório consiste basicamente de canais de alta permeabilidade (os canais) e regiões de baixa permeabilidade, a idéia foi usar o coeficiente A como uma forma de melhorar a qualidade da areia de fundo sem alterar significativamente a permeabilidade dos canais (A variou

entre 0 mD e 500 mD). O coeficiente B foi inserido para essencialmente aumentar a permeabilidade nos canais, seus valores variaram entre 0.3 e 2.5. O coeficiente A agia como um fator de homogeneização do reservatório enquanto que o fator B agia como fator de aumento da heterogeneidade. O problema desta abordagem é que a melhor solução encontrada praticamente não alterou o desempenho do caso base, por este motivo o ajuste global foi abandonado neste caso.

$$Kx_{\text{bloco}}^* = A + B \cdot Kx_{\text{bloco}} \quad (6.1)$$

Uma nova metodologia de ajuste foi tentada onde uma série de transformações locais na permeabilidade horizontal foi realizada em fases, com a qualidade dos ajustes sendo avaliados localmente. O resultado foi um ajuste composto de três fases, onde cada etapa é iniciada com a definição de: um conjunto de parâmetros de ajuste, os intervalos de variação para eles e as respectivas discretizações dos intervalos. Os resultados parciais obtidos em toda iteração da Busca Dispersa eram avaliados de forma que, se os objetivos locais eram atingidos, a solução parcial era aceita e o método era interrompido economizando as simulações restantes.

6.3.1 Fase 1

Os objetivos desta primeira fase de ajuste foram melhorar a injetividade dos poços INJ1, INJ2, INJ7 e INJ9 e antecipar a chegada da água no poço PROD3 (ver a curva azul do poço PROD3 na Figura 6.26). A parametrização nesta fase consistiu em:

1. Definir um canal de permeabilidade entre o poço injetor INJ1 e o produtor PROD3 (parâmetro de ajuste A).
2. Definir uma barreira de permeabilidade à esquerda do poço injetor INJ1 (parâmetro de ajuste B).
3. Definir um canal de permeabilidade entre o poço injetor INJ9 e o produtor PROD8 (parâmetro de ajuste C).
4. Definir um canal de permeabilidade entre o poço injetor INJ7 e o produtor PROD8 (parâmetro de ajuste D).
5. Definir um canal de permeabilidade entre o poço injetor INJ2 e poço produtor PROD4 (parâmetro de ajuste E).

Assim, foram definidos cinco parâmetros de ajuste sendo que os parâmetros de ajuste dos canais de permeabilidade foram definidos como intervalos de 1000 mD a 8000 mD com 15 níveis discretos enquanto que o parâmetro de ajuste da barreira de permeabilidade foi definido como um intervalo de 0 mD a 500 mD com 11 níveis discretos.

Uma solução viável foi encontrada após uma iteração da Busca Dispersa que foi finalizada com 35 simulações. Após esta fase, os poços injetores INJ1 e INJ9 passaram a honrar completamente a injeção de água do histórico enquanto que os injetores INJ2 e INJ7 mostraram melhorias significativas em suas curvas de injeção (ver curva preta na Figura 6.27). Adicionalmente, o poço produtor PROD3 apresentou um bom ajuste (curva preta na Figura 6.26).

6.3.2 Fase 2

Os objetivos nesta fase foram antecipar a irrupção de água no poço produtor PROD12 e reajustar a produção de água no poço PROD8, que ficou muito desajustada após a Fase 1. Desta vez, quatro parâmetros de ajuste com 15 níveis discretos foram definidos:

1. Definição de um canal de permeabilidade entre o poço produtor PROD8 e o poço injetor INJ9 (parâmetro de ajuste A).
2. Definição de um canal de permeabilidade entre o poço produtor PROD12 e o poço injetor INJ9 (parâmetro de ajuste B).
3. Definição de um canal de permeabilidade ligando o poço produtor PROD12 e o poço injetor INJ11 (parâmetro de ajuste C).
4. Definição da permeabilidade na região em torno do poço injetor INJ9 para melhorar a injetividade (parâmetro de ajuste D).

Um resultado satisfatório foi obtido após duas iterações da Busca Dispersa ao custo de 48 simulações. Como resultado desta etapa do processo, o ajuste do poço produtor PROD12 apresentou uma grande melhoria em relação à etapa anterior (curva verde da Figura 6.26), uma melhoria de menor intensidade foi constatada no poço produtor PROD8 (curva verde da Figura 6.26), os poços injetores ajustados INJ1, INJ9 e INJ11 mantiveram seus respectivos ajustes enquanto que houve pequenas melhorias na injetividade dos poços injetores INJ2 e INJ7 (curvas verdes da Figura 6.27).

6.3.3 Fase 3

O principal objetivo desta fase foi melhorar o ajuste do poço produtor PROD5. O histórico de produção do poço mostra que a água chega ao poço na forma de uma única frente de água, porém as curvas simuladas indicam que pelo menos duas frentes distintas chegam ao poço em tempos distintos. Por sua vez, os mapas de permeabilidade do modelo (Silva *et al.*, 2005) mostram que o poço PROD5 está sob influência pelo menos dos poços injetores INJ7 e INJ2 e, possivelmente, INJ1. Adicionalmente, o gráfico do ajuste do poço PROD5 mostra que não há água suficiente na região deste poço. Por estes motivos, esta fase se concentrou em ajustar a influência dos três poços injetores citados no poço PROD5. A parametrização adotada foi:

1. Definir um canal de permeabilidade do poço injetor INJ1 até o poço produtor PROD5 (parâmetro de ajuste A).
2. Definir um canal de permeabilidade do poço injetor INJ1 até o poço produtor PROD3 (parâmetro de ajuste B).
3. Definir uma região de alta permeabilidade em torno do poço produtor PROD5 para facilitar a chegada de água (parâmetro de ajuste C).
4. Definir uma região de alta permeabilidade em torno do poço injetor INJ1 para melhorar a injetividade deste poço (parâmetro de ajuste D).

Todos os parâmetros de ajuste acima foram definidos com intervalos de 1000 mD a 8000 mD com 15 níveis discretos. Uma solução satisfatória foi obtida já durante a amostragem inicial do algoritmo ao custo de 19 simulações. Nesta etapa, o poço produtor PROD5 obteve uma significativa melhoria em relação às demais etapas (ver curva magenta da Figura 6.26). Os efeitos colaterais desta etapa foram um leve desajuste no poço produtor PROD3 e no poço injetor INJ7.

Como o objetivo principal deste caso era ilustrar o procedimento de ajuste assistido para um caso real e mostrar as dificuldades associadas a este tipo de ajuste, e não fazer um ajuste completo do caso, este trabalho limitou-se a executar apenas estas três etapas iniciais do processo de ajuste assistido. A qualidade do ajuste parcial obtido nas etapas pode ser observada com detalhes na Figura 6.26 e na Figura 6.27. O processo poderia continuar de acordo com os objetivos específicos traçados para este caso.

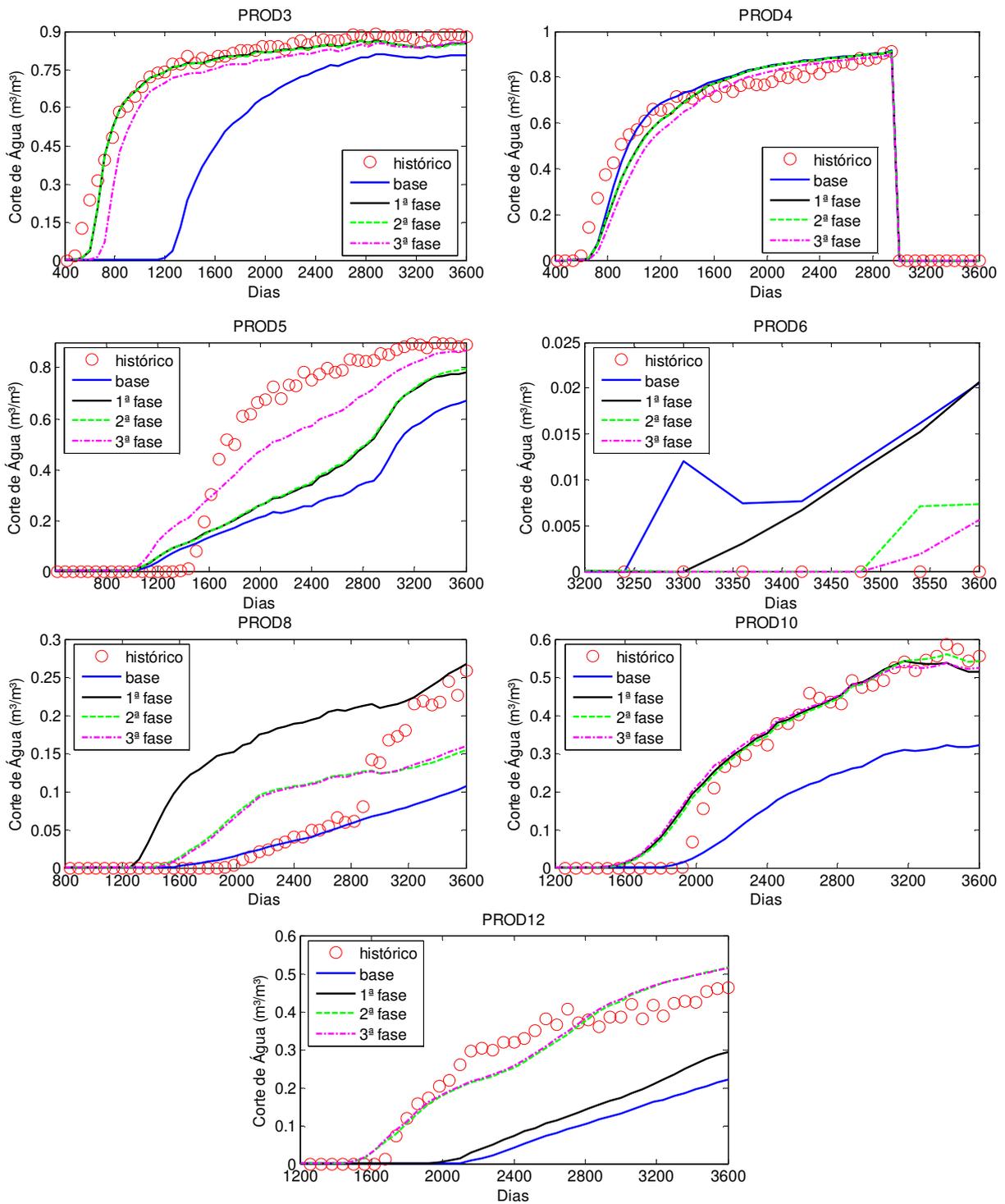


Figura 6.26 – Ajuste do corte de água dos poços produtores (Caso 2).

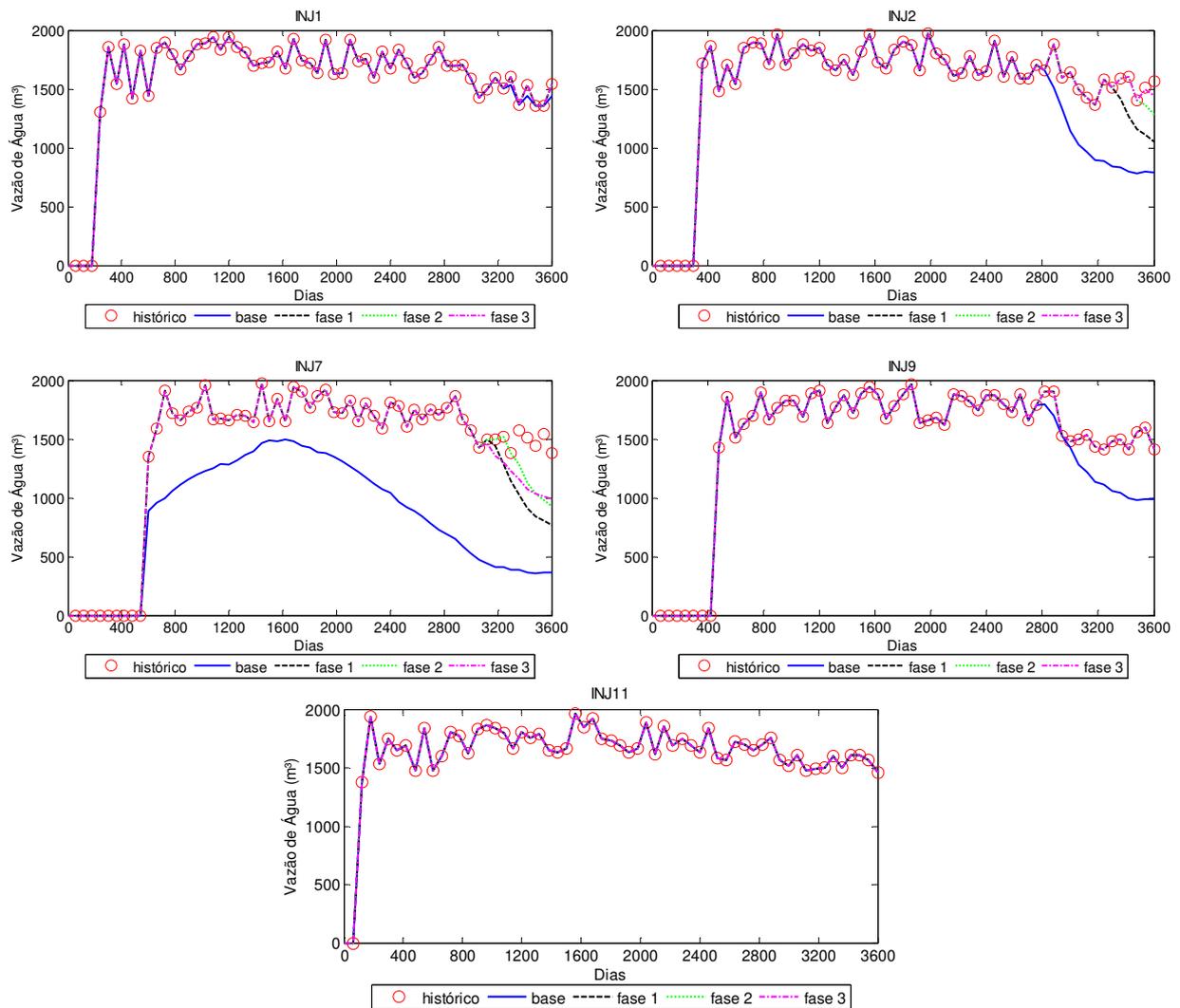


Figura 6.27 – Ajuste da vazão de água nos poços injetores (Caso 2).

6.4 ANÁLISE DE *SPEEDUP*

Conforme foi citado na Seção 4.5.2.1, a implementação da Busca Dispersa fez uso da computação distribuída em diversas etapas do processo: na geração do conjunto de referência inicial, durante a avaliação da vizinhança da busca exploratória, durante a busca linear e também durante a combinação de soluções de referência na geração de um novo ponto de partida para a busca local.

A análise do impacto do uso da computação distribuída não foi medida de forma ampla durante este trabalho, pois ela exigiria a computação do tempo serial do algoritmo, ou seja, a

avaliação do tempo de execução do algoritmo não realizando as simulações em paralelo. Como muitas simulações foram realizadas ao longo do processo, optou-se por analisar uma parte da execução do algoritmo da Busca Dispersa na exploração do espaço de soluções de 57 níveis do Caso 1B. Em particular, o trecho analisado foi a inicialização do algoritmo seguido da primeira iteração, incluindo um total de 444 simulações.

Para fazer as medidas de tempo, um programa adicional foi criado onde a lógica da Busca Dispersa foi substituída por um simples escalonador de simulações capaz de simular e analisar a função objetivo em série ou em paralelo. Com esta ferramenta, duas agendas de simulação foram alimentadas ao programa: uma contendo os pontos visitados pelo algoritmo fornecidos em série ou em paralelo da mesma forma que foram tratados na Busca Dispersa e a outra contendo todos os pontos visitados especificados em série. Os resultados da análise estão resumidos na Tabela 6.12.

Tabela 6.12 – Tabela de cálculo do *speedup* da Busca Dispersa

	Tempo Serial	Tempo Paralelo
Início	18:07:57	20:36:42
Fim	20:27:31	21:12:08
Duração	02:19:34	00:35:26
	<i>Speedup</i>	3.94

Observe que o tempo total de execução da parte inicial do algoritmo levou pouco mais que duas horas quando executado de forma serial enquanto que a versão paralela levou apenas 35 minutos para executar. O *speedup* calculado neste caso foi aproximadamente 4. Como o algoritmo foi usado em uma rede com sete computadores, o *speedup* máximo esperado era de 7. Algumas razões contribuíram para a observação de um *speedup* mais baixo. Entre eles, destacam-se:

- Enquanto as simulações de fluxo eram feitas em paralelo, a transformação dos arquivos de saída do simulador em planilhas que pudessem ser lidas pelo programa de cálculo do valor da função objetivo era feita em série na máquina na qual o processo de otimização estava executando.
- Apesar de as simulações estarem sendo feitas em paralelo, o disco no qual os resultados das simulações eram armazenados se localizava no nó central da rede

de forma que havia uma concorrência no acesso ao recurso de leitura e armazenamento dos dados a serem processados.

Observe que, apesar de ser relativamente baixo, o *speedup* medido mostra que o custo medido simplesmente em número de simulações realizadas pode ser injusto quando a Busca Dispersa é comparada com métodos intrinsecamente seqüenciais, uma vez que o tempo de execução total é reduzido para $\frac{1}{4}$ do tempo serial. Esta observação permite uma reinterpretação das tabelas de desempenho do algoritmo. Dividindo-se o número total de simulações realizadas pelo *speedup* obtido pelo uso da computação distribuída, é possível definir o conceito de número equivalente de simulações. Este valor equivalente de simulações poderia ser usado na comparação do desempenho da Busca Dispersa com outros algoritmos onde a computação distribuída não pode ou simplesmente não foi utilizada, pois de nada adianta possuir um grande poder computacional em uma rede se a otimização se dá em apenas uma máquina.

Como exemplo dos benefícios da utilização da computação distribuída no Caso 1C, a relação entre o número total de simulações e o número equivalente de simulações é mostrado na Figura 6.28. Observe que existe uma grande vantagem em usar a computação distribuída em casos como este onde as 7004 simulações absolutas realizadas no caso mais refinado (57 níveis discretos) podem ser contabilizadas como 1751 simulações equivalentes quando comparadas com algoritmos seriais.

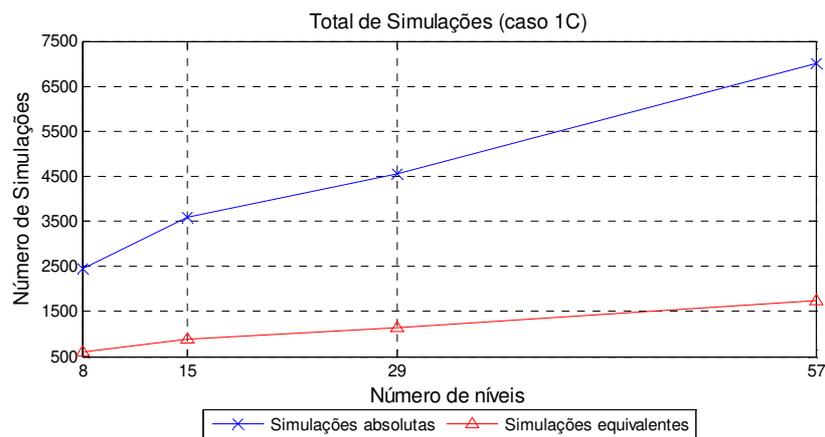


Figura 6.28 – Simulações absolutas versus simulações equivalentes no Caso 1C

Capítulo 7 – CONCLUSÕES E SUGESTÕES PARA TRABALHOS FUTUROS

Diversas são as causas que tornam o problema do ajuste de histórico tão complexo. Entre as principais citam-se: (1) o grande número de parâmetros incertos, (2) o problema não pode ser expresso analiticamente em função dos parâmetros de caracterização do reservatório exigindo uma simulação numérica para obter os dados de produção e (3) as funções objetivo que medem o ajuste costumam medir o afastamento através de funções baseadas em diferenças quadradas que geram topologias complexas para o espaço de soluções. Estas dificuldades puderam ser comprovadas durante a análise dos casos de validação (Seção 6.1) que apresentaram diversos mínimos locais durante a execução do método e, em alguns casos, a exploração destes espaços de solução gerou uma solução ótima local diferente ao final de cada iteração da Busca Dispersa.

A presença de diversos mínimos locais nos problemas estudados confirma uma afirmação comum na literatura que é a possibilidade de existência de múltiplas soluções viáveis para o ajuste. Encontrar diversas soluções para o ajuste, caso existam, é importante, pois permite uma melhor avaliação da complexidade do problema. Múltiplas soluções sugerem maior nível de incertezas quanto ao modelo e podem motivar uma análise mais detalhada das incertezas existentes com o objetivo de, ou reduzi-las através da aquisição de novas informações, ou quantificá-las para tomar decisões melhores sobre a exploração do campo. Neste contexto, a suposição que diversas soluções, armazenadas no conjunto de referência (Seção 2.8.1), são encontradas pela Busca Dispersa faz desta metodologia um bom candidato a método de otimização para o problema.

Para espaços de soluções muito complexos e com diversos mínimos locais, uma discretização menos refinada pode levar a soluções que melhor representam o objetivo final do ajuste como pôde ser observado na comparação dos ajustes obtidos no Caso 1C para os espaços de 8 e 57 níveis (Seção 6.2). Naquela comparação, os valores encontrados para as

permeabilidades horizontais ficaram bem mais próximos dos valores alvo apesar de a função objetivo não refletir isso.

A metodologia apresentada mostrou-se útil tanto em procedimentos automatizados de ajuste (Casos 1A, 1B e 1C) quanto em procedimentos de ajuste assistido (Casos 1C e 2). Para que processos automatizados de ajuste funcionem, é fundamental que a parametrização do problema esteja correta e que uma discretização apropriada seja selecionada, pois espaços de soluções muito complexos apresentam um grande número de mínimos locais que dificultam a convergência destes métodos para os mínimos globais. Por outro lado, dificuldades de convergência de métodos automatizados podem ser contornadas com um procedimento de ajuste assistido como foi feito no Caso 1C (Seção 6.2).

O ajuste assistido por si só não é garantia de convergência, é preciso que as premissas assumidas pelo engenheiro de reservatórios não contrastem com a dinâmica do reservatório. Em casos como este, o ajuste fica comprometido ou impossibilitado como aconteceu na primeira tentativa de ajuste assistido do Caso 1C. Uma característica importante do ajuste assistido é que ele limita as iniciativas de busca a uma região específica do espaço de soluções da parametrização adotada. Se por um lado isto permite reduzir o número de simulações na busca do ajuste, ela também limita as chances de encontrar soluções de ajuste não consideradas por quem está realizando o ajuste. Esta limitação não costuma ocorrer nos algoritmos automatizados que tendem a melhor explorar o espaço de soluções.

Um aprendizado que se extrai das experiências obtidas durante a realização deste trabalho é que um bom algoritmo de otimização para o ajuste de histórico deve saber dosar estratégias de intensificação com estratégias de diversificação. Passar tempo demais na intensificação leva a soluções muito parecidas entre si enquanto que o oposto leva a um conjunto de soluções de baixa qualidade. A metodologia proposta apresentou um bom balanço entre as estratégias de intensificação e diversificação. Este fato ficou evidenciado pelo aumento da dispersão das soluções de referência na medida em que o espaço de soluções ficou mais complexo com o aumento do número de parâmetros de ajuste.

Com relação ao custo da execução da Busca Dispersa medida em número total de simulações realizadas, o método apresentou certa demora em convergir para soluções ótimas

locais em espaços de soluções mais complexos, pois nestes casos a busca linear não consegue progredir de forma tão rápida levando a freqüentes reavaliações da direção de decida pela técnica da busca exploratória. Porém, mesmo nos casos mais complexos, o aumento do número de simulações com o aumento da discretização do espaço de soluções apresentou um crescimento aproximadamente linear, o que mostra uma boa escalabilidade da metodologia.

Um dos procedimentos adotados pela metodologia proposta que permitiu um melhor sentimento do progresso da otimização foi a normalização dos afastamentos encontrados em relação ao caso base. Com ela, é possível não só compatibilizar medidas de diferentes ordens de grandeza, mas avaliar a qualquer momento o grau de otimização já conseguido pela metodologia. Isto é especialmente importante em procedimentos de ajuste assistido onde o algoritmo pode ser interrompido quando uma determinada expectativa de otimização for alcançada.

É importante citar que nenhum método que pretende aumentar a produtividade do engenheiro de reservatórios tira sua responsabilidade de usar seu senso crítico e intuição sobre o problema. Assim, quanto maior for o entendimento do engenheiro sobre os detalhes da metodologia em uso, mais útil será a ferramenta, mais fácil será interpretar os resultados obtidos e mais formas de fazer bom uso dela serão descobertas. Isto é especialmente verdade em processos de ajuste assistidos, onde a melhor parametrização é mais uma incógnita do problema.

De forma resumida, as principais contribuições deste trabalho para o problema do ajuste foram:

1. O uso de parâmetros de ajuste discretos para tornar o espaço de soluções discreto e, portanto, mais adequado para ser resolvido por processos metaheurísticos.
2. A adaptação da metodologia da Busca Dispersa para o problema de ajuste, mostrando que seus diversos parâmetros de configuração (número de iterações, critérios de parada) e seu potencial de fornecer múltiplas soluções se mostrou adequada para explorar o espaço de soluções típico de problemas de ajuste de histórico.
3. O uso da metodologia proposta tanto de forma automática quanto assistida, mostrando que cada tipo de ajuste tem sua aplicabilidade e apresenta prós e contras que devem ser avaliados pelo engenheiro de reservatórios.

4. A demonstração, na análise dos resultados obtidos, que existe uma relação direta entre o número de simulações realizadas e a qualidade das soluções encontradas. Esta propriedade, que pode ser percebida de maneira relativamente intuitiva, não havia sido demonstrada em trabalhos anteriores.

Como sugestão de trabalhos futuros com a Busca Dispersa aplicada ao problema do ajuste de histórico, recomenda-se:

1. Estudar heurísticas adicionais na avaliação da direção de descida na busca exploratória pelo uso de informações de buscas exploratórias prévias. Se uma mesma direção foi seguida em um eixo do problema em n avaliações anteriores, será que a direção contrária precisa ser avaliada nesta busca? Estudos deste tipo podem diminuir o custo da busca exploratória nesta metodologia.
2. No decorrer deste trabalho o espaço de soluções foi sempre considerado combinatorial, ou seja, os eixos foram discretizados de forma que o tamanho do passo dado na otimização local era sempre fixo e igual à distância entre um e outro nível discreto. Um trabalho interessante seria trabalhar em espaços de soluções contínuos com o passo dado variando de acordo com um gradiente estimado na busca de soluções otimizadas. A expectativa é que se consiga dar passos mais largos quando a heurística local está distante de um mínimo local, porém, é preciso tomar cuidado para não tornar a convergência muito lenta quando perto de mínimos locais.
3. A informação do número de passos dado pela última busca linear também pode ser usada ao compor heurísticas para diminuir o número de simulações realizado pela busca exploratória. Quanto maior for o número de passos dado pela busca linear, mais distante esta busca exploratória se encontra em relação à busca exploratória anterior, e menor será o efeito ou validade das heurísticas adotadas.
4. Informações sobre o comprimento da busca linear também podem ser usadas para determinar desacelerações na convergência do método de busca local. Ao detectar situações como esta, a técnica de busca local poderá adotar um algoritmo diferente na tentativa de acelerar a busca nesta etapa final. Em

outras palavras, métodos híbridos podem ser estudados para aumentar a velocidade de convergência da busca local.

5. Outra abordagem para reduzir o custo do método medido em número de simulações de fluxo realizadas é usar metamodelos durante o processo de avaliação do espaço de soluções. A sugestão é que tanto soluções apontadas como sendo de alta qualidade quanto de baixa qualidade sejam confirmados por simulações de fluxo e que, de preferência, estas confirmações sejam usadas para melhorar a qualidade das respostas fornecidas pelo metamodelo.
6. O uso de metamodelos também pode viabilizar a eliminação das relaxações ao algoritmo base da Busca Dispersa que foram adotadas nesta metodologia para viabilizar seu uso no contexto do ajuste de histórico. Em particular, todas as soluções da amostragem inicial (Seção 4.5.1) poderiam ser otimizadas assim como as soluções tentativas geradas pelo algoritmo de combinação de soluções de referência (Seção 4.5.3).
7. Com o uso da metodologia proposta, é possível estudar um conjunto de recomendações para procedimentos de ajuste dado um conjunto de características do reservatório ou da topologia do espaço de soluções gerado. Estes resultados poderiam ser usados no desenvolvimento de sistemas especialistas para o problema do ajuste de histórico.
8. Outra linha de pesquisa futura poderia ser a comparação da metodologia proposta com outros métodos de otimização. Os parâmetros de comparação poderiam ser número de simulações, qualidade da(s) solução(ões) final(is) obtida(s), aplicabilidade de computação distribuída e, finalmente, tempo total de execução.

REFERÊNCIAS

- April, J., Glover, F., Kelly, J. P. e Laguna, M. Practical Introduction to Simulation Optimization. In: Winter Simulation Conference, New Orleans, Louisiana, USA, 2003, pp.71-78.
- Arenas, E., Kruijsdijk, C. V. e Oldenziel, T. Semi-Automatic History Matching Using the Pilot Point Method Including Time-Lapse Seismic Data. In: SPE Annual Technical Conference and Exhibition, New Orleans, Louisiana, USA. *Proceedings*. SPE, 2001, pp.1-11.
- Blum, C. e Roli, A. Metaheuristics in Combinatorial Optimization: Overview and Conceptual Comparison. *ACM Computing Surveys*, v.35, (3), pp.268-308, 2003.
- Cullick, A. S., Johnson, D. e Shi, G. Improved and More-Rapid History Matching With a Nonlinear Proxy and Global Optimization. In: SPE Annual Technical Conference and Exhibition, San Antonio, Texas, U.S.A. *Proceedings*. SPE, 2006,
- Dennis Jr., J. E. e More, J. J. *Quasi-Newton Methods, Motivation and Theory*. Cornell University. 1974
- Fanchi, J. R. *Principles of applied reservoir simulation*: Gulf Professional Publishing, 2001, 355 p.
- Glover, F. Future Paths for Integer Programming and Links to Artificial Intelligence. *Computers and Operations Research*, v.13, (5), pp.533-549, 1986.

- Glover, F. *A Template for Scatter Search and Path Relinking*. In: J.-K. Hao, E. Lutton, *et al* (Ed.). *Selected Papers from the Third European Conference on Artificial Evolution*. London, UK: Springer-Verlag, 1998. A Template for Scatter Search and Path Relinking, p.3-54
- Glover, F., Laguna, M. e Martí, R. Fundamentals of Scatter Search and Path Relinking. *Control and Cybernetics*, v.29, (3), pp.653-684, 2000.
- Gomez, S., Gosselin, O. e Barker, J. W. Gradient-Based History Matching With a Global Optimization Method. *SPE Journal*, pp.200-208, 2001.
- Hooke, R. e Jeeves, T. A. 'Direct Search' Solution of Numerical and Statistical Problems. *Journal of the ACM*, v.8, (2), pp.212-229, 1961.
- Jahn, F., Cook, M. e Graham, M. *Hydrocarbon Exploration and Production*: Elsevier, 1998, 396 p. (Developments in Petroleum Science)
- Kabir, C. S., Chien, M. C. H. e Landa, J. L. Experiences With Automated History Matching. In: SPE Reservoir Simulation Symposium, Houston, Texas, U.S.A., 2003,
- Kolda, T. G., Lewis, R. M. e Torczon, V. Optimization by Direct Search: New Perspectives on Some Classical and Modern Methods. *SIAM Review*, v.45, (33), pp.385-482, 2003.
- Leitão, H. C. *Ajuste de Histórico Automatizado Através de Otimização Multivariada e Paralelização Externa*. Campinas: Faculdade de Engenharia Mecânica, Universidade Estadual de Campinas, 1997, 165 p. Dissertação (Mestrado).
- Lewis, R. M., Torczon, V. e Trosset, M. W. *Direct Search Methods: Then and Now*. Institute for Computer Applications in Science and Engineering. Hampton, Virginia, U.S.A.: May 2000. 2000. (ICASE Report No. 2000-26)

- Li, B. e Friedmann, F. *Semiautomatic Multiple Resolution Design for History Matching*. *SPE Annual Technical Conference and Exhibition*. San Antonio, Texas, U.S.A.: SPE 2006.
- Mantica, S., Cominelli, A. e Mantica, G. Combining Global and Local Optimization Techniques for Automatic History Matching Production and Seismic Data. *SPE Journal*, pp.123-130, 2002.
- Maschio, C., Ligerio, E. L. e Schiozer, D. J. *Estudo Comparativo da Simulação de Reservatórios de Petróleo por Linhas de Fluxo e Pelo Método Tradicional*. IX Congresso Brasileiro de Engenharia e Ciências Térmicas. Caxambu, MG, Brazil 2002.
- Maschio, C. e Schiozer, D. J. Assisted History Matching Using Streamline Simulation. *Petroleum Science and Technology*, v.23, (7-8/2005), pp.761-774, 2005.
- Mattax, C. C. e Dalton, R. L. *Reservoir Simulation*: Society of Petroleum Engineers, v.13, 1990, 184 p. (SPE Monograph Series)
- Moore, G. E. *Cramming more components onto integrated circuits*. *Electronics*. 38 1965.
- Risso, V. F. e Schiozer, D. J. *Ajuste de Modelos Numéricos de Campos de Petróleo Combinando Dados de Produção com Sísmica 4D*. Rio Oil & Gas Expo and Conference. Rio de Janeiro, RJ - Brasil 2006.
- Romanov, V. G. *Ill-posed Problems of Mathematical Physics and Analysis*: American Mathematical Society, 1997, 290 p. (Translations of Mathematical Monographs)
- Schulze-Riegert, R. W., Axmann, J. K., Haase, O., Rian, D. T. e You, Y.-L. Evolutionary Algorithms Applied to History Matching of Complex Reservoirs. *SPE Reservoir Evaluation & Engineering*, pp.163-173, 2002.

Silva, F. P. T., Rodrigues, J. R. P., Paraizo, P. L. B., Romeu, R. K., Peres, A. M. M., Oliveira, R. M., Pinto, L. B. e Maschio, C. *Novel Ways of Parameterizing the History Matching Problem. SPE Latin American and Caribbean Petroleum Engineering Conference*. Rio de Janeiro, Brazil 2005.

Sousa, S. H. G. D., Maschio, C. e Schiozer, D. J. Solving the History Matching Problem With the Scatter Search Metaheuristic. In: Iberian Latin American Congress On Computational Methods In Engineering, Belém-PA, Brazil, 2006, pp.18.

Staggs, H. M. e Herbeck, E. F. Reservoir Simulation Models - An Engineering Overview. *Journal of Petroleum Technology*, v.23, (December), pp.1428-1436, 1971.

Thomas, J. E., Triggia, A. A., Correia, C. A., Filho, C. V., Xavier, J. A. D., Machado, J. C. V., Filho, J. E. D. S., Paula, J. L. D., Rossi, N. C. M. D., Pitombo, N. E. S., Gouvêa, P. C. V. D. M., Carvalho, R. S. e Barragan, R. V. *Fundamentos de Engenharia de Petróleo*. Rio de Janeiro: Editora Interciência, 2001, 271 p.

Wilf, H. S. *Algorithms and Complexity*. Philadelphia, PA, U.S.A., 1994, 139 p.

Wolsey, L. A. *Integer Programming*. New York, NY: John Wiley & Sons, Inc., 1998, 264 p. (Wiley-Interscience Series in Discrete Mathematics and Optimization)