



**UNIVERSIDADE FEDERAL DO OESTE DO PARÁ**  
**PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO E INOVAÇÃO TECNOLÓGICA**  
**PROGRAMA DE PÓS-GRADUAÇÃO EM RECURSOS NATURAIS DA AMAZÔNIA**

**RESOLUÇÃO DE PROBLEMAS DE ESTIMAÇÃO DE PARÂMETROS**  
**UTILIZANDO O ALGORITMO DE OTIMIZAÇÃO TOPOGRÁFICA**

**ROSANE TOLENTINO GUSMÃO MAIA**

**Santarém (PA)**  
**2014**

**ROSANE TOLENTINO GUSMÃO MAIA**

**RESOLUÇÃO DE PROBLEMAS DE ESTIMAÇÃO DE PARÂMETROS  
UTILIZANDO O ALGORITMO DE OTIMIZAÇÃO TOPOGRÁFICA**

**ORIENTADOR:** Prof. WAGNER FIGUEIREDO SACCO, D.SC.

**CO-ORIENTADOR(A):** Profa. ANA CAROLINA RIOS COELHO, D.SC.

Dissertação apresentada como requisito parcial para obtenção do título de Mestre ao Programa de Pós-Graduação em Recursos Naturais da Amazônia da Universidade Federal do Oeste do Pará.

**Áreas de Concentração:** Conservação e Manejo da Biodiversidade na Amazônia.

**Santarém (PA)  
2014**

FICHA CATALOGRÁFICA ELABORADA PELA SEÇÃO TÉCNICA DE AQUISIÇÃO E TRATAMENTO DA INFORMAÇÃO DA DIVISÃO TÉCNICA DE BIBLIOTECA E DOCUMENTAÇÃO-CAMPUS RONDON-UFOPA.

**BIBLIOTECÁRIA RESPONSÁVEL:**

Tolentino, Rosane.  
Resolução de problemas de estimação de parâmetros utilizando o algoritmo de otimização topográfica. Santarém, 2014.  
Nº de Páginas: 54f.

Bibliografia.  
Dissertação (Mestrado). Universidade Federal do Oeste do Pará, 2014.  
Orientador: Prof. Wagner Figueiredo Sacco, D.SC.

1.Otimização. 2. Algoritmos difusos. 3.Análise multivariada.

RESOLUÇÃO DE PROBLEMAS DE ESTIMAÇÃO DE PARÂMETROS  
UTILIZANDO O ALGORITMO DE OTIMIZAÇÃO TOPOGRÁFICA

Esta dissertação foi julgada adequada para a obtenção do Título de Mestre em Recursos Naturais da Amazônia. Área de Concentração: Conservação e Manejo da Biodiversidade na Amazônia. Aprovada em sua forma final pelo Programa de Pós-Graduação *Strictu Sensu* em Recursos Naturais da Amazônia, nível de Mestrado, da Universidade Federal do Oeste do Pará – UFOPA, em \_\_\_ / \_\_\_ / 2014.

Prof. Dr. Luís Reginaldo Ribeiro Rodrigues  
Coordenador do PGRNA da Universidade Federal do Oeste do Pará

Apresentada à Comissão Examinadora Integrada pelos Professores:

---

Prof. Wagner Figueiredo Sacco, D.Sc.

Universidade Federal do Oeste do Pará - UFOPA

---

Prof<sup>ª</sup>. Ana Carolina Rios Coelho, D.Sc.

Universidade Federal do Oeste do Pará - UFOPA

---

Prof. Carlos Eduardo Guerra, D.Sc.

Universidade Federal do Oeste do Pará - UFOPA

---

Prof. Luiz Nélio Henderson Guedes de Oliveira, D.Sc.

Instituto Politécnico da Universidade do Estado do Rio de Janeiro – IPRJ/UERJ

Santarém, março de 2014.

## **AGRADECIMENTOS**

Agradeço aos meus orientadores e a todas as pessoas que, de alguma forma, contribuíram na execução da presente dissertação.

A Ciência não estuda ferramentas, mas  
o que fazemos e o que descobrimos  
com elas.

*Edsger Wybe Dijkstra (1930 - 2002)*

TOLENTINO, Rosane. **Resolução de problemas de estimação de parâmetros utilizando o algoritmo de otimização topográfica**. 2014. 54 f. Dissertação de Mestrado em Recursos Naturais da Amazônia. Área de Concentração: Conservação e Manejo da Biodiversidade na Amazônia. Programa de Pós-Graduação em Recursos Naturais da Amazônia Universidade Federal do Oeste do Pará - UFOPA, Santarém, 2014.

## RESUMO

Neste trabalho, utilizamos um método de otimização a ser aplicado a problemas de otimização de parâmetros das biociências e da engenharia química, muitos dos quais são multimodais. Nesta técnica, primeiramente são geradas soluções iniciais dentro do domínio do problema por meio da sequência *quasi*-aleatória de Sobol com o intuito de explorar o espaço de busca. Em seguida, com a finalidade de encontrar mais de um ótimo, estas soluções foram agrupadas utilizando o algoritmo denominado *Topographical Algorithm* (TA). Finalmente, foram feitas buscas locais através dos métodos determinísticos de otimização Hooke-Jeeves e Rosenbrock com a finalidade de se encontrar os ótimos locais. O método é formado por duas variantes, combinando a inicialização das soluções na primeira fase e dois algoritmos de busca local na terceira. A fim de que este método possa ser avaliado, seus resultados foram comparados com os obtidos por outras metodologias e os resultados alcançados são bastante efetivos.

Palavras-chave: Otimização. Análise multivariada. Métodos de simulação. Programação heurística.

TOLENTINO, Rosane. **Resolução de problemas de estimação de parâmetros utilizando o algoritmo de otimização topográfica.** 2014. 54 f. Dissertação de Mestrado em Recursos Naturais da Amazônia. Área de Concentração: Conservação e Manejo da Biodiversidade na Amazônia. Programa de Pós-Graduação em Recursos Naturais da Amazônia Universidade Federal do Oeste do Pará - UFOPA, Santarém, 2014.

## ABSTRACT

In this work, we solve parameter estimation problems from Biosciences and Chemical Engineering. These problems are highly multimodal, requiring optimization techniques that overcome local optima. In order to do so, we use a clustering optimization technique based on the topographical information on the objective function called Topographical Algorithm (TA). This algorithm consists of three steps: a uniform random sampling of solutions in the search space, the construction of the topograph, and the application of a local optimization algorithm using the topograph minima as starting points. In this work, we use the Sobol quasi-random sequence to perform the first step and the Hooke–Jeeves and Rosenbrock direct search methods for the third step. Both variants of TA are compared against other methods in the literature, and the results obtained are very effective.

Keywords: Optimization. Multivariate analysis. Simulation methods. Heuristic Programming.

# SUMÁRIO

RESUMO.....	iv
ABSTRACT.....	v
LISTA DE TABELAS.....	viii
<b>1 INTRODUÇÃO.....</b>	<b>9</b>
2 O ALGORITMO DE OTIMIZAÇÃO TOPOGRÁFICA (TA).....	11
2.1 GERAÇÃO DAS SOLUÇÕES INICIAIS UTILIZANDO A SEQUÊNCIA DE SOBOL.....	12
2.2 HEURÍSTICA TOPOGRÁFICA (HT).....	18
2.3 OS ALGORITMOS DETERMINÍSTICOS DE BUSCA LOCAL.....	20
2.3.1 Algoritmo Hooke-Jeeves.....	20
2.3.2 Algoritmo Rosenbrock.....	21
<b>3 OS PROBLEMAS DE OTIMIZAÇÃO ABORDADOS.....</b>	<b>24</b>
3.1 FUNÇÕES DE TESTE DA LITERATURA.....	24
3.1.1 Função Easom.....	24
3.1.2 Função Shekel's Foxholes.....	25
3.1.3 Função Branin.....	26
3.1.4 Função Himmelblau.....	27
3.2 PROBLEMAS DE ESTIMAÇÃO DE PARÂMETROS.....	28
3.2.1 Modelo para eliminação enzimática de substratos do fígado- <i>Pig Liver Problem</i> .....	28
3.2.2 Modelo de Reator Catalítico.....	30
3.3.3 Modelo Cinético de Primeira Ordem.....	33
<b>4 RESULTADOS E DISCUSSÕES.....</b>	<b>35</b>
4.1 O AMBIENTE DE TESTES.....	35

4.2	AVALIAÇÃO DOS MÉTODOS DE OTIMIZAÇÃO.....	35
4.2.1	<b>CrITÉrios de avaliaço.....</b>	<b>35</b>
4.2.2	<b>A metodologia experimental.....</b>	<b>36</b>
4.2.3	<b>Resultados do TA aplicado aos problemas abordados.....</b>	<b>36</b>
4.2.4	<b>Comparativo dos resultados consolidados pelos mtodos Sbl-HT-HJ e Sbl-HT-Rsbk com resultados obtidos por outros mtodos.....</b>	<b>38</b>
4.2.4.a	<i>Funço Easom.....</i>	40
4.2.4.b	<i>Funço Shekel's Foxholes.....</i>	41
4.2.4.c	<i>Funço Branin.....</i>	42
4.2.4.d	<i>Funço Himmelblau.....</i>	43
4.2.4.e	<i>Pig Liver Problem (PLP).....</i>	43
3.2.4.f	<i>Modelo de Reator Cataltico.....</i>	45
3.2.4.g	<i>Modelo Cintico de 1ª Ordem.....</i>	45
4	<b>CONCLUSES E TRABALHOS FUTUROS.....</b>	<b>46</b>
	<b>REFERNCIAS.....</b>	<b>47</b>

## LISTA DE TABELAS

<b>Tabela 1</b> - Inteiros ímpares ( $m_i$ ), números direcionais e Sobol [ $\Phi(n)$ ].....	15
<b>Tabela 2</b> - Valores experimentais das taxas de eliminação do fígado.....	28
<b>Tabela 3</b> - Valores dos parâmetros para cada experimento.....	30
<b>Tabela 4</b> – Dados e seus valores estimados.....	32
<b>Tabela 5</b> – Demanda bioquímica de oxigênio ao longo do tempo.....	34
<b>Tabela 6</b> – Resumo dos ótimos globais e $f(x^*)$ das funções e modelos testados.....	34
<b>Tabela 7</b> – Resultados obtidos com busca local pelo Hooke-Jeeves.....	37
<b>Tabela 8</b> – Resultados obtidos com busca local pelo Rosenbrock.....	37
<b>Tabela 9</b> – Resultados de outros métodos aplicados para solução da função Easom.	40
<b>Tabela 10</b> –Resultados da aplicação dos métodos na função <i>Shekel's Foxholes</i> .....	41
<b>Tabela 11</b> - Resultados de outros métodos aplicados para solução da função Branin.....	42
<b>Tabela 12</b> - Resultados de outros métodos aplicados para solução da função Himmelblau.....	43
<b>Tabela 13</b> - Resultados obtidos para o PLP.....	44
<b>Tabela 14</b> - Resultados obtidos para o Modelo de Reator Catalítico.....	45

## 1 INTRODUÇÃO

Na resolução de problemas reais de otimização, particularmente problemas de estimação de parâmetros, deparamo-nos frequentemente com espaços de busca multimodais, isto é, com vários ótimos locais ou mesmo múltiplos ótimos globais (COELHO et al., 2009; COELHO, ALOTTO, 2009; SACCO et al., 2009, por exemplo). Além disso, tem sido uma tendência atual resolver sistemas não-lineares utilizando métodos de otimização (GROSAN; ABRAHAM, 2008; HENDERSON; SACCO, 2008; HIRSCH et al., 2009; HENDERSON et al., 2010; SACCO, HENDERSON, 2011, por exemplo). Neste caso, cada raiz do sistema não-linear será um ótimo global do problema equivalente de otimização, caracterizando assim um problema com vários ótimos globais.

A ideia central de métodos de otimização para a resolução de problemas multimodais é detectar os vários ótimos através de uma ampla varredura do espaço de busca.

Por esta razão, têm sido utilizados, principalmente, algoritmos populacionais com este intuito, como os algoritmos genéticos (HOLLAND, 1975) e evolução diferencial (STORN, PRICE, 1997). Estes algoritmos são associados a mecanismos para a manutenção da diversidade populacional, como por exemplo, em SACCO et al. (2004), GROSAN e ABRAHAM (2008), COELHO e ALOTTO (2009) e YU e SUGANTHAN (2010).

Outro enfoque é o apresentado recentemente por HENDERSON et al. (2010), em que um aperfeiçoamento do algoritmo *Simulated Annealing* (KIRKPATRICK et al., 1983) proposto por CORANA et al. (1987) é associado a uma técnica que os autores chamaram de polarização, que cria uma área de repulsão ao redor de ótimos já detectados através de um engenhoso artifício matemático, levando o algoritmo a vasculhar áreas inexploradas. HIRSCH et al. (2009) utilizam uma técnica filosoficamente semelhante associada ao C-GRASP (HIRSCH et al., 2007).

Uma terceira maneira de atacar problemas de otimização multimodais, que esteve em voga principalmente na década de 80 e início da década de 90, é aquela aplicada por Aimo Törn (ver TÖRN e ZILINSKAS, 1989, por exemplo), em que o espaço de busca é particionado por meio de heurísticas específicas e em cada partição é determinado o ponto com menor valor de função objetivo, que servirá de solução inicial para um algoritmo de busca local (BAZARAA et al., 2006). Assim, em cada partição será encontrado um mínimo local. Este enfoque foi utilizado em diversos métodos, sendo um dos mais bem sucedidos nesta linha o *Topographical Algorithm* (TA) (TÖRN e VIITANEN, 1992, 1994).

O objetivo principal deste trabalho é propor a resolução de problemas de estimação de parâmetros das biociências e da engenharia química partindo do paradigma proposto por TÖRN e VIITANEN (1992, 1994), o TA. Este método é constituído pelos seguintes passos:

1. Soluções iniciais são geradas dentro do domínio do problema por meio da sequência *quasi*-aleatória de SOBOL (1967). Enquanto números pseudo-aleatórios são realizações independentes de uma variável aleatória (ALI, 1994), sequências *quasi*-aleatórias têm por objetivo preencher um hipercubo unitário com pontos distribuídos o mais uniformemente possível (GENTLE, 2005).
2. Estas soluções são agrupadas pela heurística topográfica, que neste trabalho chamaremos de HT (TÖRN e VIITANEN, 1992, 1994).
3. Em seguida, em cada subespaço é acionado um método determinístico de busca local partindo da melhor solução encontrada em cada um deles no passo anterior. Neste trabalho, são utilizados dois métodos determinísticos de busca local, que fazem parte de uma família de métodos ditos de busca-padrão (LEWIS et al., 2000): o Hooke-Jeeves (HOOKE, JEEVES, 1961), que foi aplicado com sucesso associado a um método estocástico por RIOS-COELHO et al. (2010), e o Rosenbrock (ROSENBROCK, 1960), que é uma variante do primeiro método que visa uma maior exploração do espaço de busca.

Assim, temos duas possíveis combinações:

1. Inicialização pelo Sobol, clusterização pela HT e busca local pelo Hooke-Jeeves, doravante denominado **Sbl-HT-HJ**;
2. Inicialização pelo Sobol, clusterização pela HT e busca local pelo Rosenbrock, doravante denominado **Sbl-HT-Rsbk**;

No **capítulo 1**, o método aqui utilizado e seus constituintes são descritos em detalhe. No **capítulo 2**, são descritos os problemas de otimização abordados por este método. No **capítulo 3** são mostrados os resultados obtidos pelo método e é feita uma análise acerca do desempenho de cada uma de suas combinações. Finalmente, no **capítulo 4** são tecidas as conclusões.

## 2 O ALGORITMO DE OTIMIZAÇÃO TOPOGRÁFICA (TA)

Entre as décadas de setenta e noventa, o paradigma de otimização global baseada em clusterização foi estudado por alguns pesquisadores, principalmente na Europa. BECKER e LAGO (1970) foram os pioneiros, seguidos, entre outros, por TÖRN (1973, 1978), TIMMER (1984), TÖRN e VIITANEN (1992, 1994) e ALI e STOREY (1994). Segundo TÖRN e ZILINSKAS (1989), a motivação para explorar métodos de clusterização baseia-se nos seguintes fatores:

- (a) É possível obter uma amostra de pontos em um espaço de busca baseado na concentração de pontos vizinhos mais próximos dos mínimos locais da função objetivo  $f$ ;
- (b) Os pontos da amostra podem ser agrupados em *clusters* contendo os mínimos locais vizinhos mais próximos onde em seguida aplica-se um método de busca local.

Originalmente, o TA é não iterativo e baseia-se na exploração do espaço de busca (ALI, STOREY, 1994). Ele consiste em três passos (TÖRN, VIITANEN, 1994):

1. O espaço de busca é preenchido com  $N$  pontos aleatórios uniformemente distribuídos;
2. Na topografia, que são os grafos com arcos orientados conectando os  $k$  vizinhos mais próximos da amostra, a seta do arco aponta no sentido do ponto com maior valor de função objetivo. O mínimo do grafo serão os pontos cujos vizinhos têm maior valor de função objetivo, i.e., são os pontos onde não há arcos entrando, apenas saindo;
3. Os mínimos do grafo são os pontos de partida para um algoritmo de otimização de busca local. Após executar a busca local utilizando como ponto de partida todos os mínimos obtidos na etapa anterior, a solução do algoritmo será o melhor ponto obtido de todas essas execuções.

Originalmente, TÖRN e VIITANEN (1992,1994) obtiveram soluções iniciais (passo1) preenchendo um hipercubo unitário até que  $N$  pontos, atendendo a condição de um limiar de distância  $\delta$  entre seus vizinhos mais próximos, fossem obtidos. Esses pontos então eram desnormalizados. Os autores ressaltam que qualquer método capaz de preencher inicialmente o espaço de busca o mais uniformemente possível pode ser utilizado. Em um trabalho posterior (TÖRN, VIITANEN, 1996), estes autores utilizaram amostragem quasi-aleatória (GENTLE, 2003; PRESS et al., 2007), bem mais eficiente, em uma versão do TA. Em relação ao número de pontos em seus testes, Törn e Viitanen (1994) trabalharam com  $N=100$  e  $N=200$ .

Neste trabalho, utilizamos a sequência de Sobol, que será exposta na subseção a seguir.

## 2.1 GERAÇÃO DAS SOLUÇÕES INICIAIS UTILIZANDO A SEQUÊNCIA DE SOBOL

*Discrepância e dispersão* são os dois principais conceitos indicativos da extensão da equidistribuição de pontos amostrais. Quanto menores forem a discrepância e a dispersão, mais regularmente estarão distribuídos estes pontos. Sequências de pontos com baixa discrepância e baixa dispersão são usualmente denominadas sequências *quasi-aleatórias* (*quasi-random sequences*).

As sequências *quasi-aleatórias* são úteis em muitas áreas, como em otimização global (NIEDERREITER, 1986) e em aproximações de integrais em altas dimensões (HUA; WANG, 1981). Várias sequências *quasi-aleatórias* tem sido propostas, como as de Faure, Halton, Niederreiter e Sobol (GENTLE, 2005).

Dados  $N$  pontos amostrais dentro de um hipercubo unitário  $d$ -dimensional  $I = [0,1)^d$ , a discrepância indicará se diferentes subdomínios em  $I$  possuem densidades de pontos iguais. Uma definição usual de discrepância é dada por  $\sup_D \left| \frac{A(D)}{N} - \frac{|D|}{|I|} \right|$  (BRATLEY; FOX, 1992), onde  $\sup$  é o supremo de um conjunto de dados,  $|\cdot|$  denota o tamanho de uma região,  $D$  é qualquer subdomínio de  $I$  e  $A(D)$  é o número de pontos dentro de  $D$ .

Dado um conjunto de pontos dentro de  $I$ , a dispersão, por sua vez, indica intuitivamente o tamanho da maior região esférica vazia (sem nenhum ponto) em  $I$ . O tamanho desta região vazia é usualmente expressada pela máxima distância que qualquer ponto tenha em relação à distância do vizinho mais próximo (BRANICKY et al., 2001).

De forma intuitiva, portanto, baixa discrepância e baixa dispersão indicam que os pontos amostrais estão razoavelmente equidistribuídos. Como dissemos acima, sequências de pontos com estas características são ditas sequências *quasi-aleatórias*. Neste trabalho, utilizaremos a sequência de SOBOL (1967), pois os resultados obtidos por GALANTI e JUNG (1997) demonstram a sua superioridade. Estes autores geraram 500 pontos entre 2 e 260 dimensões utilizando as sequências de Faure, Halton e Sobol e observaram que apenas a última gerou distribuições satisfatórias em todas as dimensões.

A sequência de Halton colapsou a partir de 14 dimensões e a de Faure a partir de 25 dimensões.

SOBOL (1967) propôs o método que leva o seu nome para gerar sequências *quasi-aleatórias* com baixas discrepância e dispersão.

A exposição a seguir segue os trabalhos de BRATLEY e FOX (1988) e de GALANTI e JUNG (1997).

A construção de uma sequência multidimensional de Sobol segue um procedimento de quatro passos para cada dimensão. Como estes passos são idênticos para todas as dimensões, o processo será ilustrado para uma única dimensão. Além disso, serão mostrados exemplos para finalizar a compreensão do procedimento. Os passos são os seguintes:

**Passo 1:** Gere um conjunto de inteiros ímpares  $m_i$ , para  $i = 1, 2, \dots, [\log_2 N]$ , que satisfaçam a condição  $0 < m_i < 2^i$ , onde  $N$  é o número de pontos e  $[\log_2 N]$  é o menor inteiro maior que  $\log_2 N$ .  $[\log_2 N]$  é o número máximo de dígitos na expansão de  $N$  na base 2. Isto leva a um processo recursivo descrito a seguir.

Começamos com uma série de inteiros  $h_1, h_2, h_3, \dots$ , onde cada  $h_j$  vale "0" ou "1" e é coeficiente do polinômio de grau  $d$

$$P = x^d + h_1 x^{d-1} + h_2 x^{d-2} + \dots + h_{d-1} x + 1 \quad (1)$$

Nesta discussão, nosso único interesse nesses polinômios é com relação aos seus coeficientes.

Em seguida, geramos o conjunto de  $m_i$  para todos os termos  $i$  utilizando os coeficientes do polinômio e uma relação recursiva para  $i > d$ :

$$m_i = 2h_1 m_{i-1} \oplus 2^2 h_2 m_{i-2} \oplus \dots \oplus 2^{d-1} h_{d-1} m_{i-d+1} \oplus 2^d m_{i-d} \oplus m_{i-d}, \quad (2)$$

onde  $h_1, h_2, \dots, h_{d-1}$  são os coeficientes do polinômio de grau  $d$  e  $\oplus$  é o operador *exclusive-or* (XOR) *bit a bit*:

$$1 \oplus 0 = 0 \oplus 1 \equiv 1$$

$$1 \oplus 1 = 0 \oplus 0 \equiv 0$$

Devido ao fato de a Eq. (3) gerar  $m_i$  apenas para  $i > d$ , os  $d$  primeiros inteiros ímpares ( $m_1, m_2, \dots, m_d$ ) devem ser fornecidos. Felizmente, eles podem ser livremente escolhidos desde que satisfaçam a  $m_i$  ser ímpar e a  $0 < m_i < 2^i$ .

Ilustremos este procedimento utilizando um polinômio de terceiro grau:

$$P = x^3 + x^2 + 1 = 1x^3 + \underbrace{1}_{h_1}x^2 + \underbrace{0}_{h_2}x + 1, \quad (3)$$

com coeficientes  $h_1 = 1$  e  $h_2 = 0$ . A relação de recorrência representada pela Eq. (4) torna-se

$$m_i = 2m_{i-1} \oplus 2^3m_{i-3} \oplus m_{i-3} \quad (4)$$

Escolhendo arbitrariamente os três primeiros  $m_i$ , i.e.,  $m_1 = 1, m_2 = 3$  e  $m_3 = 7$ . Começando com  $i = 4$ , temos da Eq. (6):

$$\begin{aligned} m_4 &= 2m_3 \oplus 8m_1 \oplus m_1 \\ &= 14 \oplus 8 \oplus 1 \\ &= 1110 \oplus 1000 \oplus 0001 \quad \text{expandindo em base 2} \\ &= 0111 \\ &= 7 \quad \text{em base 10.} \end{aligned}$$

Para  $i = 5$ , temos da Eq. (4):

$$\begin{aligned} m_5 &= 2m_4 \oplus 8m_2 \oplus m_2 \\ &= 14 \oplus 24 \oplus 3 \\ &= 01110 \oplus 11000 \oplus 00001 \quad \text{expandindo em base 2} \\ &= 10111 \\ &= 23 \quad \text{em base 10,} \end{aligned}$$

e assim por diante.

Podemos observar que  $m_i$  é um inteiro ímpar satisfazendo a condição  $0 < m_i < 2^i$ . Na Tabela 1, abaixo, mostramos os resultados dos seis primeiros valores de  $m_i$  utilizando o polinômio dado pela Eq. (3).

**Tabela 1** - Inteiros ímpares ( $m_i$ ), números direcionais e Sobol  $[\Phi(n)]$ .

$i$ (ou $n$ )	1	2	3	4	5	6
$m_i$	1	3	7	7	23	17
$m_i$ base 2	1	11	111	111	10111	10001
$v(i)$ base 10	1/2	3/2 <sup>2</sup>	7/2 <sup>3</sup>	7/2 <sup>4</sup>	23/2 <sup>5</sup>	17/2 <sup>5</sup>
$v(i)$ base 2	0,1	0,11	0,111	0,0111	0,10111	0,010001
Sobol $\Phi(n)$	1/2	1/4	3/4	1/8	5/8	3/8

**Passo 2:** Calcule um conjunto de “números direcionais” através da conversão de  $m_i$  para o sistema numérico de base 2. O  $i$ -ésimo número direcional  $v(i)$  para  $i = 1, 2, \dots, [\log_2 N]$  é dado por

$$v(i) = m_i / 2^i \quad (5)$$

Em seguida, convertamos este número para base 2. Como este procedimento é idêntico para todos os números direcionais, ilustramos o mesmo apenas para  $i = 4$ . Partindo da Eq. (7) e utilizando o quarto inteiro ímpar,  $m_{i=4} = 7$  (vide Tabela 1), temos

$$v(4) = \frac{7}{2^4} = \frac{7}{16}$$

Expandindo 7/16 em base 2, teremos

$$\frac{7}{16} = \frac{0}{2^1} + \frac{1}{2^2} + \frac{1}{2^3} + \frac{1}{2^4} + \frac{0}{2^5} + \dots +$$

onde os coeficientes da expansão são 0, 1, 1 e 1. Portanto,  $v(4) = 0,0111$  no sistema numérico de base 2. A Eq. (5) é equivalente a deslocar para a direita a expansão binária de  $m_{i=4} = 7$  em  $i = 4$  dígitos (adicionando “0s” precedentes se necessário). Ou seja, a expansão de 7 na base 2 é “111” e deslocando o ponto decimal binário em quatro dígitos para a esquerda adicionando um 0 precedente, teremos  $v(4) = 0,0111$ . A Tabela 1 exhibe os resultados para os seis primeiros números direcionais.

**Passo 3:** Converta uma sequência consecutiva de inteiros não-negativos ( $n = 0, 1, 2, \dots, N - 1$ ) em suas representações no sistema numérico de base 2. Como a sequência de Sobol utiliza a mesma base 2 para cada dimensão, este passo precisa ser feito apenas uma vez.

**Passo 4:** Calcule o  $n$ -ésimo número de Sobol  $\Phi(n)$  para  $n = 0, 1, \dots, N - 1$  utilizando a seguinte expressão recursiva:

$$\Phi(n + 1) = \Phi(n) \oplus v(c), \quad (6)$$

onde  $\Phi^k(0) = 0$ ,  $v(c)$  é o  $c$ -ésimo número direcional,  $c$  é o *bit* de valor zero mais à direita na expansão de base 2 de  $n$  e  $k$  é a dimensão do problema. O número  $n = 17$ , por exemplo, vale “010001” em representação binária. O *bit* de valor zero mais à direita é o segundo a partir da direita, o que leva a  $c = 2$ .

Utilizando os números direcionais da Tabela 1, ilustraremos o passo 4 para a geração dos seis primeiros números de Sobol baseados no polinômio da Eq. (3). Para  $n = 0$ , a expansão binária de “0” é  $[\dots, a_1(0), a_0(0)] = [\dots, 0, 0]$ . Como o *bit* de valor zero mais à direita é  $c = 1$ , utilizamos o primeiro número direcional,  $v(1) = 0,1$  da Tabela 1. Da Eq. (6) com  $\Phi^k(0) = 0$ , o primeiro número de Sobol vale

$$\begin{aligned} \Phi(1) &= \Phi(0) \oplus v(1) \\ &= 0,0 \oplus 0,1 \quad \text{em binário} \\ &= 0,1 \quad \text{em binário} \\ &= 1/2 \quad \text{em base 10.} \end{aligned}$$

Para  $n = 1$ , a expansão binária é  $[\dots, a_1(1), a_0(1)] = [\dots, 0, 1]$ . Como o *bit* de valor zero mais à direita de “01” é  $c = 2$ , utilizamos o segundo número direcional,  $v(2) = 0,11$  da Tabela 1. Da Eq. (6), o segundo número de Sobol vale

$$\begin{aligned} \Phi(2) &= \Phi(1) \oplus v(2) \\ &= 0,10 \oplus 0,11 \quad \text{em binário} \\ &= 0,01 \quad \text{em binário} \\ &= 1/4 \quad \text{em base 10.} \end{aligned}$$

Para  $n = 2$ , a expansão binária é  $[\dots, a_1(2), a_0(2)] = [\dots, 1, 0]$ . Como o *bit* de valor zero mais à direita de “10” é  $c = 1$ , utilizamos o primeiro número direcional,  $v(1) = 0,1$  da Tabela 1. Da Eq. (6), o terceiro número de Sobol vale

$$\begin{aligned} \Phi(3) &= \Phi(2) \oplus v(1) \\ &= 0,01 \oplus 0,10 \quad \text{em binário} \end{aligned}$$

$$= 0,11 \quad \text{em binário}$$

$$= 3/4 \quad \text{em base 10,}$$

e assim sucessivamente. A última linha da Tabela 1 mostra os seis primeiros números de Sobol (excluindo o 0-ésimo).

Para gerar uma sequência de Sobol  $s$ -dimensional, começando pela primeira dimensão, cada dimensão sucessiva é gerada sequencialmente utilizando-se polinômios distintos. Tipicamente, o procedimento começa com o polinômio de grau mais baixo, movendo-se para graus mais altos com o crescimento de  $s$ . Sumarizando, para cada dimensão:

1. Utilizando o polinômio dado pela Eq. (2) e a relação recursiva dada pela Eq. (4), gere um conjunto de inteiros ímpares  $m_i$  para  $i = 1, 2, \dots, [\log_2 N]$  que satisfaçam a condição  $0 < m_i < 2^i$ , onde  $N$  é o número de pontos e  $[\log_2 N]$  é o menor inteiro maior que  $\log_2 N$ .
2. Utilizando a Eq. (7), calcule o conjunto de números direcionais  $v(i)$  convertendo  $m_i$  em uma fração binária no sistema numérico de base 2 (i.e., deslocando para a direita a expansão binária de  $m_i$  por  $i$  dígitos).
3. Converta uma sequência consecutiva de inteiros não-negativos ( $n = 0, 1, 2, \dots, N - 1$ ) em suas representações no sistema numérico de base 2. Este passo só precisa ser feito uma vez.
4. Começando com  $n = 0$  e movendo consecutivamente para o próximo inteiro de maior valor, localize o *bit* de valor zero mais à direita da expansão binária de  $n$  do passo 3. Rotule este *bit* como o  $c$ -ésimo dígito (contando a partir do dígito da expansão mais à direita). Utilize o  $c$ -ésimo número direcional  $v(c)$  que foi calculado no passo 2 através da Eq. (6) para calcular o  $n$ -ésimo número de Sobol. Repita o procedimento para cada  $n$ .

Repita os passos 1, 2 e 4 para cada uma das  $s$  dimensões, utilizando um polinômio distinto para cada dimensão.

Finalmente, a título de ilustração, relacionamos os primeiros dezesseis termos da sequência de Sobol para três dimensões:  $S_{16}^3 = \{(0,0000;0,0000;0,0000), (0,5000;0,5000;0,5000), (0,7500;0,2500;0,2500), (0,2500;0,7500;0,7500), (0,3750;0,3750;0,6250), (0,8750;0,8750;0,1250), (0,6250;0,1250;0,8750), (0,1250;0,6250;0,3750), (0,1875;0,3125;0,9375), (0,6875;0,8125;0,4375), (0,9375;0,0625;0,6875), (0,4375;0,5625;0,1875),$

$(0,3125;0,1875;0,3125),(0,8125;0,6875;0,8125),(0,5625;0,4375;0,0625),(0,0625,0,9375;0,5625)\}$ .

## 2.2 HEURÍSTICA TOPOGRÁFICA (HT)

No passo 2, a parte principal do método, o grafo é construído através da heurística topográfica. Primeiro, a matriz de distância simétrica  $N \times N$  é obtida calculando a distância euclidiana de cada ponto da amostra em relação aos demais. Em seguida, uma matriz chamada  $kNN$  é construída, contendo o índice de seus  $k$  vizinhos mais próximos (classificados pela distância euclidiana obtida anteriormente). Depois, esta matriz é transformada em um grafo direcionado onde o índice de cada linha da matriz representa o ponto a ser comparado com os demais vizinhos. A composição desta é bastante simples. É formada por sinais positivos (+) ou negativos (-). Os sinais representam a direção dos vetores no grafo, se o  $k$  vizinho mais próximo tem valor de função objetivo maior (sinal positivo – seta chegando no ponto) ou se tem o valor de função objetivo menor (sinal negativo – seta saindo do ponto) em relação ao ponto da linha analisada (ALI, 1994). Um sinal positivo representa a “ponta” do vetor e um sinal negativo representa o “início” do mesmo (TÖRN, VIITANEN, 1994). Finalmente, os pontos que correspondem a setas com sinais somente positivos são os mínimos do TA.

Abaixo, ilustraremos um exemplo adaptado de ALI (1994) de como a heurística topográfica funciona.

Supondo que desejemos minimizar a função

$$f(x, y) = x^2 + y^2,$$

e que seis pontos foram amostrados com os seguintes valores de função calculados:

$$f(P_1) = f(2,5) = 29, f(P_2) = f(1,2) = 5, f(P_3) = f(3,4) = 25, f(P_4) = f(0,1) = 1, f(P_5) = f(5,0) = 25 \text{ e } f(P_6) = f(4,2) = 20.$$

Primeiro, a matriz simétrica  $D$  da distância euclidiana é construída, onde, por exemplo, o elemento  $d_{13}$  corresponde à distância entre  $P_1$  e  $P_3$ :

$$\mathbf{D} = \begin{bmatrix} 0 & 10 & 2 & 20 & 34 & 13 \\ 10 & 0 & 8 & 2 & 20 & 9 \\ 2 & 8 & 0 & 18 & 20 & 5 \\ 20 & 2 & 18 & 0 & 26 & 17 \\ 34 & 20 & 20 & 26 & 0 & 5 \\ 13 & 9 & 5 & 17 & 5 & 0 \end{bmatrix}$$

Em seguida, constrói-se a matriz  $kNN$  contendo os  $k$  pontos vizinhos mais próximos com relação ao ponto representado pelo índice da linha da matriz  $\mathbf{D}$ . Usando  $k=3$ , os vizinhos mais próximos de  $P_1$  (a primeira linha de  $\mathbf{D}$ ) são os pontos com índices 3, 2 e 6 respectivamente. Esses elementos irão compor a primeira linha da matriz  $kNN$ . O processo continua até que a seguinte matriz seja obtida:

$$\mathbf{kNN} = \begin{bmatrix} 3 & 2 & 6 \\ 4 & 3 & 6 \\ 1 & 6 & 2 \\ 2 & 6 & 3 \\ 6 & 2 & 3 \\ 3 & 5 & 2 \end{bmatrix}$$

Esta matriz representa um grafo não direcionado. Computacionalmente ela é obtida classificando-se cada linha de  $\mathbf{D}$  em ordem crescente e tomando como os primeiros  $k$  elementos os seus índices correspondentes na matriz, tomando o cuidado de atribuir aos elementos da diagonal principal um valor alto (ex.,  $10^{+8}$ ), a fim de que eles não sejam selecionados.

Os elementos de  $kNN$  agora receberão um sinal de mais (+) ou menos (-) de acordo com os seus valores funcionais em relação ao valor dos pontos representados pelo índice da linha.

A segunda linha, por exemplo, corresponde a  $P_2$ , cujo valor de função é igual a 5, que é maior que  $f(P_4)=1$  ( $P_4$  é o elemento  $knn_{21}$ ), porém menor que  $f(P_3)=13$  e  $f(P_1)=29$  (elementos  $knn_{22}$  e  $knn_{23}$ , respectivamente). Então  $knn_{21}$  receberá o sinal (-) e os outros dois elementos receberão o sinal (+) originando a matriz a seguinte matriz:

$$\mathbf{kNN} = \begin{bmatrix} -3 & -2 & -6 \\ -4 & +3 & +6 \\ +1 & -6 & -2 \\ +2 & +6 & +3 \\ -6 & -2 & +3 \\ +3 & +5 & -2 \end{bmatrix}$$

Como  $P_4=(0,1)$  é o único ponto que corresponde a uma linha somente com sinais positivos, com o valor de função de todos os  $k$  vizinhos maiores do que o dele. Portanto, dos  $N$  pontos gerados inicialmente ele tem o menor valor de função objetivo e será o ponto de partida para o algoritmo de otimização local. Quando implementamos a heurística topográfica, os sinais podem ser atribuídos no processo de construção de  $kNN$ .

Para o passo 3, Törn e Viitanen (1994) afirmam que quaisquer métodos de otimização local podem ser utilizados. Eles empregaram um algoritmo baseado no gradiente da função, pois os seus testes foram realizados com funções de testes algébricas. Neste trabalho, conforme mencionado, testamos dois algoritmos de busca padrão, que serão descritos na subseção a seguir.

## 2.3 OS ALGORITMOS DETERMINÍSTICOS DE BUSCA LOCAL

### 2.3.1 Algoritmo Hooke-Jeeves

O método Hooke–Jeeves (HJ), (HOOKE; JEEVES, 1961) é um dos métodos de busca direta mais difundidos. Ele faz dois tipos de busca: uma busca exploratória e uma busca padrão. Um sumário do método é dado a seguir para a minimização de uma função  $n$ -dimensional (BAZARAA et al., 2006).

#### **Passo de inicialização**

Sejam  $\mathbf{d}_1, \dots, \mathbf{d}_n$  as direções coordenadas. Escolha um escalar  $\varepsilon > 0$  a ser utilizado para encerrar a execução do algoritmo. Além disso, escolha um passo inicial  $\Delta \geq \varepsilon$  e um fator de aceleração  $\alpha > 0$ . Escolha um ponto inicial  $\mathbf{x}_1$ , faça  $\mathbf{y}_1 = \mathbf{x}_1$  e  $k = j = 1$  e vá para o passo principal.

#### **Passo principal**

1. Se  $f(\mathbf{y}_j + \Delta\mathbf{d}_j) < f(\mathbf{y}_j)$ , a tentativa constituiu um sucesso: faça  $\mathbf{y}_{j+1} = \mathbf{y}_j + \Delta\mathbf{d}_j$  e vá para o passo 2. Se, no entanto,  $f(\mathbf{y}_j + \Delta\mathbf{d}_j) \geq f(\mathbf{y}_j)$ , a tentativa constituiu um fracasso. Neste caso, se  $f(\mathbf{y}_j - \Delta\mathbf{d}_j) < f(\mathbf{y}_j)$ , faça  $\mathbf{y}_{j+1} = \mathbf{y}_j - \Delta\mathbf{d}_j$  e vá para o passo 2; se  $f(\mathbf{y}_j - \Delta\mathbf{d}_j) \geq f(\mathbf{y}_j)$ , faça  $\mathbf{y}_{j+1} = \mathbf{y}_j$ .

2. Se  $j < n$ , substitua  $j$  por  $j + 1$  e repita o passo 1. Senão, vá para o passo 3 se  $f(\mathbf{y}_{k+1}) < f(\mathbf{y}_k)$  e vá para o passo 4 se  $f(\mathbf{y}_{k+1}) \geq f(\mathbf{y}_k)$ .
3. Faça  $\mathbf{x}_{k+1} = \mathbf{y}_{k+1}$  e  $\mathbf{y}_1 = \mathbf{x}_{k+1} + \alpha(\mathbf{x}_{k+1} - \mathbf{x}_k)$ . Substitua  $k$  por  $k + 1$ , faça  $j = 1$  e vá para o passo 1.
4. Se  $\Delta \leq \varepsilon$ , pare:  $\mathbf{x}_k$  é a solução. Senão, substitua  $\Delta$  por  $\Delta/2$ . Faça  $\mathbf{y}_1 = \mathbf{x}_k$ ,  $\mathbf{x}_{k+1} = \mathbf{x}_k$ , substitua  $k$  por  $k + 1$ , faça  $j = 1$  e repita o passo 1.

Os passos 1 e 2 acima descrevem a busca exploratória. No passo 3, há uma aceleração na direção  $\mathbf{x}_{k+1} - \mathbf{x}_k$ . Finalmente, no passo 4 o tamanho do passo  $\Delta$  é reduzido.

### 2.3.2 Algoritmo Rosenbrock

Este algoritmo, introduzido por ROSENBROCK (1960), também é de busca direta, porém foi concebido com o intuito de tentar acelerar a busca pelo ótimo através da construção de direções que utilizam informações sobre a curvatura da função objetivo adquiridas ao longo do processo de otimização (LEWIS et al., 2000). No método Hooke-Jeeves, a busca é realizada em direções ortogonais pré-estabelecidas e fixas durante o processo. Já no método Rosenbrock, quando um novo ponto é atingido ao final de uma iteração, um novo conjunto de vetores ortogonais é construído, conforme explicaremos a seguir.

Sejam  $\mathbf{d}_1, \dots, \mathbf{d}_n$  vetores linearmente independentes com norma unitária e mutuamente ortogonais, ou seja,  $\mathbf{d}_i^t \cdot \mathbf{d}_j = 0$  para  $i \neq j$ . Iniciando do vetor atual  $\mathbf{x}_k$ , a função objetivo  $f$  é minimizada iterativamente ao longo de cada direção, resultando em um ponto  $\mathbf{x}_{k+1}$ . Em particular,  $\mathbf{x}_{k+1} - \mathbf{x}_k = \sum_{j=1}^n \lambda_j \mathbf{d}_j$ , onde  $\lambda_j$  é a distância de deslocamento ao longo de  $\mathbf{d}_j$ . O novo conjunto de direções  $\bar{\mathbf{d}}_1, \dots, \bar{\mathbf{d}}_n$  é formado pelo procedimento de ortogonalização de Gram-Schmidt conforme a seguir (BAZARAA et al., 2006):

$$\mathbf{a}_j = \begin{cases} \mathbf{d}_j, & \text{se } \lambda_j = 0 \\ \sum_{i=j}^n \lambda_i \mathbf{d}_i, & \text{se } \lambda_j \neq 0 \end{cases}$$

$$\mathbf{b}_j = \begin{cases} \mathbf{a}_j, & \text{se } j = 1 \\ \mathbf{a}_j - \sum_{i=1}^{j-1} (\mathbf{a}_j^t \cdot \bar{\mathbf{d}}_i) \bar{\mathbf{d}}_i & \text{se } j \geq 2 \end{cases} \quad (7)$$

$$\bar{\mathbf{d}}_j = \frac{\mathbf{b}_j}{\|\mathbf{b}_j\|}$$

Um sumário do método é dado a seguir para a minimização de uma função  $n$ -dimensional (BAZARAA et al., 2006).

### Passo de inicialização

Sejam  $\varepsilon > 0$  um escalar que servirá como critério de parada,  $\alpha > 1$  um dado fator de expansão e  $\beta \in (-1,0)$  um dado fator de contração. Escolha  $\mathbf{d}_1, \dots, \mathbf{d}_n$  como as direções coordenadas e sejam  $\bar{\Delta}_1, \dots, \bar{\Delta}_n > 0$  os passos iniciais ao longo dessas direções. Escolha um ponto inicial  $\mathbf{x}_1$ , faça  $\mathbf{y}_1 = \mathbf{x}_1$ ,  $k = j = 1$ ,  $\Delta_j = \bar{\Delta}_j$  e vá para o passo principal.

### Passo principal

1. Se  $f(\mathbf{y}_j + \Delta_j \mathbf{d}_j) < f(\mathbf{y}_j)$ , a  $j$ -ésima tentativa constituiu um sucesso: faça  $\mathbf{y}_{j+1} = \mathbf{y}_j + \Delta_j \mathbf{d}_j$  e vá para o passo 2. Se, no entanto,  $f(\mathbf{y}_j + \Delta_j \mathbf{d}_j) \geq f(\mathbf{y}_j)$ , a tentativa constituiu um fracasso; faça  $\mathbf{y}_{j+1} = \mathbf{y}_j$  e substitua  $\Delta_j$  por  $\beta \Delta_j$ . Se  $j < n$ , substitua  $j$  por  $j + 1$  e repita o passo 1. Senão, se  $j = n$ , vá para o passo 2.
2. Se  $f(\mathbf{y}_{n+1}) < f(\mathbf{y}_1)$ , ou seja, se qualquer uma das  $n$  tentativas do passo 1 foi bem sucedida, faça  $\mathbf{y}_1 = \mathbf{y}_{n+1}$ ,  $j=1$  e repita o passo 1. Agora, considere o caso  $f(\mathbf{y}_{n+1}) = f(\mathbf{y}_1)$ , isto é, quando cada uma das  $n$  tentativas do passo 1 foi um fracasso. Se  $f(\mathbf{y}_{n+1}) < f(\mathbf{x}_k)$ , isto é, ao menos uma tentativa bem sucedida ocorreu no passo 1 durante a iteração  $k$ , vá para o passo 3. Se  $f(\mathbf{y}_{n+1}) = f(\mathbf{x}_k)$ , isto é, nenhuma tentativa foi bem sucedida, pare, ficando  $\mathbf{x}_k$  como uma estimativa da solução ótima se  $|\Delta_j| \leq \varepsilon$  para  $j$ ; senão, faça  $\mathbf{y}_1 = \mathbf{y}_{n+1}$ ,  $j=1$  e vá para o passo 1.
3. Faça  $\mathbf{x}_{k+1} = \mathbf{y}_{n+1}$ . Se  $\|\mathbf{x}_{k+1} - \mathbf{x}_k\| < \varepsilon$ , pare, ficando  $\mathbf{x}_{k+1}$  como uma estimativa da solução ótima. Senão, calcule  $\lambda_1, \dots, \lambda_n$  da relação  $\mathbf{x}_{k+1} - \mathbf{x}_k = \sum_{j=1}^n \lambda_j \mathbf{d}_j$ , forme um novo conjunto de direções utilizando (16) e denote essas direções por  $\mathbf{d}_1, \dots, \mathbf{d}_n$ . Além disso, faça  $\Delta_j = \bar{\Delta}_j$  para cada  $j$ , faça  $\mathbf{y}_1 = \mathbf{x}_{k+1}$ , substitua  $k$  por  $k + 1$ , faça  $j = 1$  e vá para o **passo 1**.

PALMER (1968) observou que o processo de ortogonalização dos vetores de busca por Gram-Schmidt era passível de gerar instabilidades numéricas colapsando o algoritmo. Por isso,

modificou este processo para evitar esta possibilidade, chegando à seguinte expressão para os novos vetores unitários ortogonais, que deve ser utilizada para  $2 \leq j \leq n$ :

$$\bar{\mathbf{d}}_j = \frac{\mathbf{a}_j \|\mathbf{a}_{j-1}\|^2 - \mathbf{a}_{j-1} \|\mathbf{a}_j\|^2}{\|\mathbf{a}_{j-1}\| \|\mathbf{a}_j\| \sqrt{(\|\mathbf{a}_{j-1}\|^2 - \|\mathbf{a}_j\|^2)}} \quad (8)$$

Neste trabalho, utilizamos o algoritmo Rosenbrock com a modificação proposta por PALMER (1968).

### 3 OS PROBLEMAS DE OTIMIZAÇÃO ABORDADOS

Os problemas de otimização abordados neste trabalho utilizando-se o método TA são divididos em duas subseções: funções de teste da literatura, utilizadas para validar a nossa implementação do método, e os problemas de estimação de parâmetros propriamente ditos.

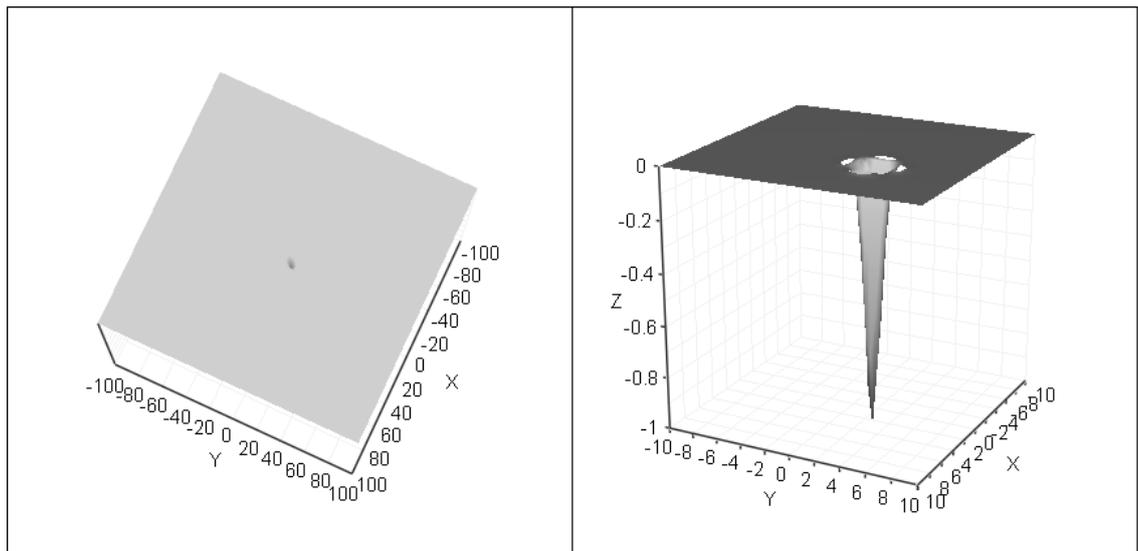
#### 3.1 FUNÇÕES DE TESTE DA LITERATURA

##### 3.1.1 Função Easom

A função Easom (MICHALEWICZ, 1996) é uma função de testes unimodal onde a região do mínimo global possui uma área pequena em relação ao espaço de busca (Figura 1, abaixo), o que faz dela um grande desafio para os algoritmos de otimização. Sua expressão é dada por

$$f(\mathbf{x}) = -\cos(x_1)\cos(x_2)e^{[-(x_1-\pi)^2-(x_2-\pi)^2]}, \quad \mathbf{x} \in [-100,100]^2, \quad (9)$$

onde  $n$  é a dimensão. O mínimo global está no ponto  $\mathbf{x}^* = (\pi, \pi)$  tendo como valor  $f(\mathbf{x}^*) = -1$ .



**Figura 1** - A função Easom.

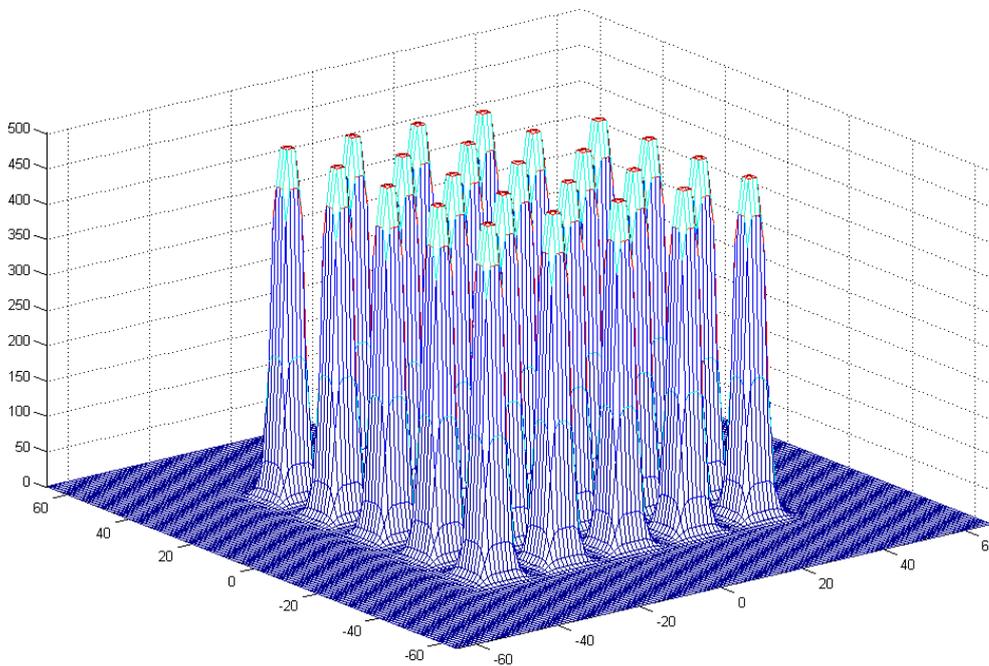
### 3.1.2 Função Shekel's Foxholes

Esta função ficou conhecida através do trabalho de DE JONG (1975). A Shekel's Foxholes (trincheiras de Shekel) é **bidimensional** e possui 25 picos, todos com diferentes alturas, variando entre 476,191 e 499,002. (Figura 2). Esta característica dificulta a atuação dos algoritmos de otimização. O **ótimo global** está localizado no ponto  $(-32, -32)$  correspondendo a  $f(x^*)=499,002$

Esta função é definida por

$$f(x) = 500 - \frac{1}{0,002 + \sum_{i=0}^{24} \frac{1}{1+i+(x_1-a(i))^6+(x_2-b(i))^6}}, \quad x \in [-65,536; 65,536]^2, \quad (10)$$

onde  $a(i) = 16[(i \bmod 5) - 2]$  e  $b(i) = 16[(i/5) - 2]$ .



**Figura 2** - A função *Shekel's Foxholes*.

### 3.1.3 Função Branin

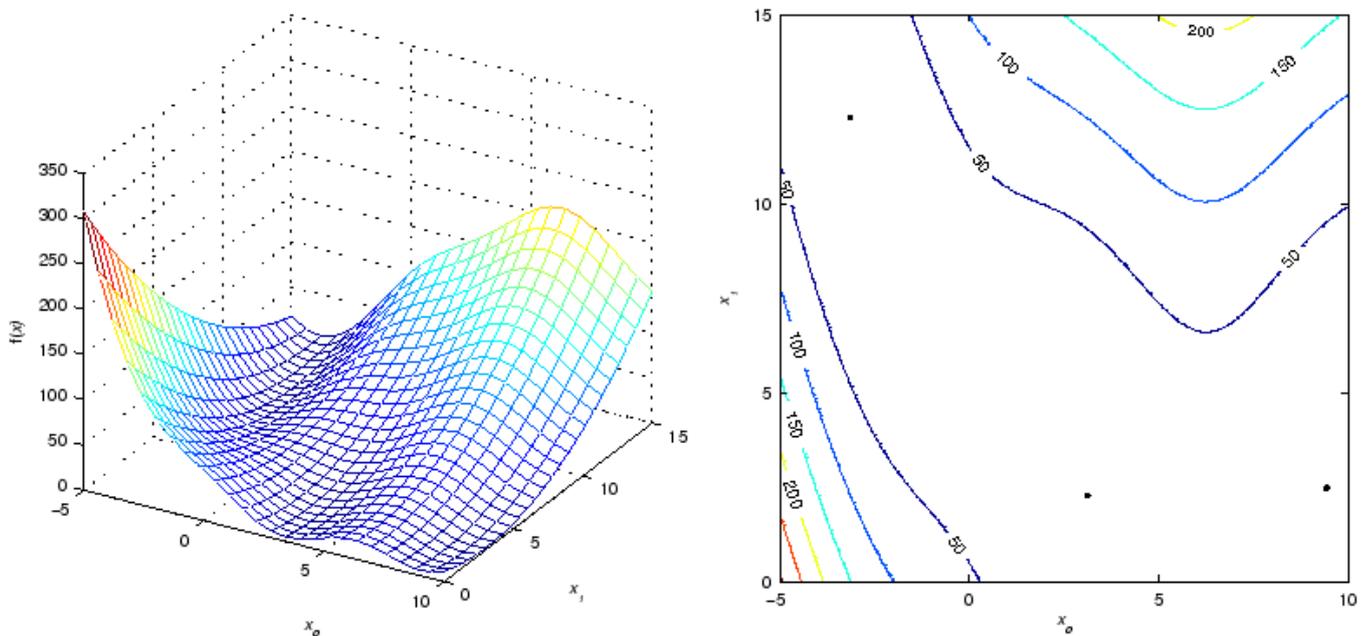
A função Branin (DIXON, SZEGÖ, 1978) possui **três ótimos globais**, sendo dada por

$$f(\mathbf{x}) = a(x_2 - bx_1^2 + cx_1 - d)^2 + g(1 - h) \cos(x_1) + g, \mathbf{x} \in [-5, 15]^2, \quad (11)$$

onde  $a = 1, b = 5/(4\pi^2), c = 5/\pi, d = 6, g = 10, h = 1/(8\pi)$ .

Os **ótimos globais** estão localizados em  $\mathbf{x}^* = (-\pi; 12, 275), (\pi; 2, 275), (3\pi; 2, 475)$  e  $f(\mathbf{x}^*) = 5/(4\pi)$ .

A Figura 3 mostra o gráfico e as curvas de nível dessa função. Os pontos negros correspondem aos ótimos globais.



**Figura 3** - Função Branin.

### 3.1.4 Função Himmelblau

Esta função possui **quatro ótimos globais** dispostos irregularmente no espaço de busca (MAHFOUD, 1995). Sua expressão é dada por:

$$f(\mathbf{x}) = (x_1^2 + x_2 - 11)^2 + (x_1 + x_2^2 - 7)^2, \mathbf{x} \in [-6,6]^2. \quad (12)$$

Os **mínimos** desta função estão localizados em  $\mathbf{x}^* = (3,000; 2,000), (3,584; -1,848), (-3,779; -3,283), (-2,805; 3,131)$  e  $f(\mathbf{x}^*) = 0$ . A Figura 4 mostra que o ótimo global correspondente a  $\mathbf{x}^* = (-3,779; -3,283)$  está isolado em relação aos demais, sendo por isso de determinação mais difícil.

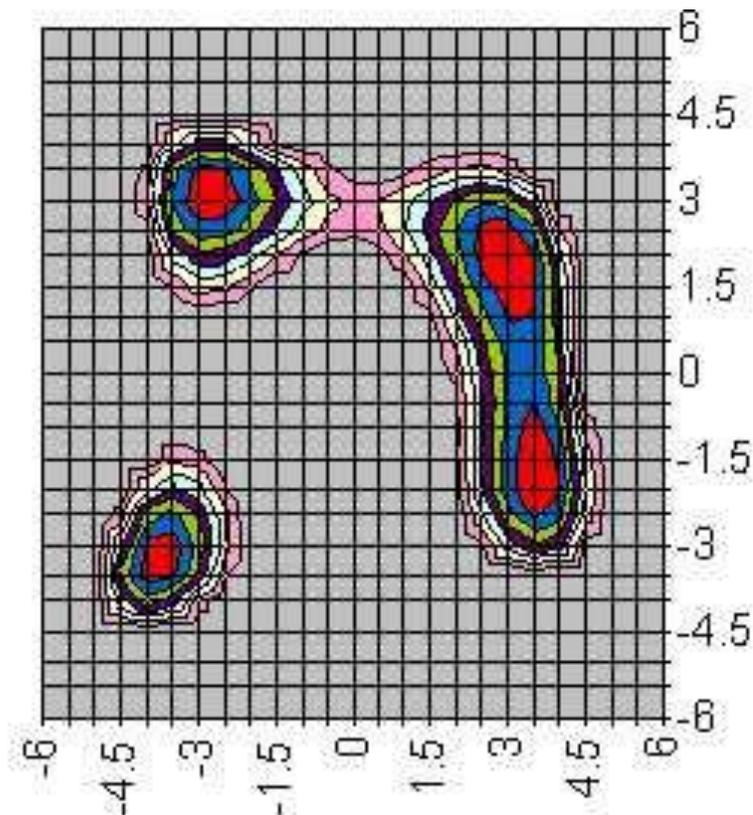


Figura 4 - Curvas de nível da função Himmelblau.

### 3.2 PROBLEMAS DE ESTIMAÇÃO DE PARÂMETROS

Serão apresentados a seguir três problemas estimação de parâmetros de modelos não-lineares, cujas soluções já foram obtidas por outros métodos. O primeiro conhecido como *Pig Liver Problem* (PLP) é um problema de estimação de parâmetros para a obtenção da taxa média de eliminação do substrato no fígado de um porco, que é anatomicamente semelhante ao fígado humano. O segundo um problema utilizado para modelar a *hidrogenação catalítica* na fase de gás do fenol em um reator catalizador pseudodiferencial. O terceiro denominado de *Modelo Cinético de Primeira Ordem* descreve uma reação irreversível de primeira ordem em um tanque de mistura com batelada.

#### 3.2.1 Modelo para eliminação enzimática de substratos do fígado (*Pig Liver Problem*)

ALI et al. (1997) utilizaram este problema como teste para um conjunto de métodos estocásticos, tendo constatado que ele é um problema de otimização bastante desafiador, possuindo **vários ótimos locais** que podem funcionar como atratores para os algoritmos de otimização, em detrimento do ótimo global que deseja-se encontrar (multimodalidade), além de ser altamente não-linear e não-diferenciável (ALI, 2007). Também foi abordado por RIOS-COELHO (2009) visando a testar métodos híbridos de otimização. RIOS-COELHO (2009) constatou inclusive que o PLP **possui dezenas de ótimos locais**.

A Tabela 2 mostra os resultados experimentais das taxas de eliminação  $V_{ij}$  do  $j$ -ésimo experimento realizado no  $i$ -ésimo fígado realizado por ROBINSON et al. (1983).

**Tabela 2** - Valores experimentais das taxas de eliminação do fígado.

$i \downarrow$	$j \rightarrow$	1	2	3	4	5
1	$V_{ij}$	0,09	0,23	0,23	0,33	0,38
2	$V_{ij}$	0,05	0,11	0,17	0,24	0,35
3	$V_{ij}$	0,26	0,36	0,55	0,57	-
4	$V_{ij}$	0,15	0,21	0,36	0,41	0,41
5	$V_{ij}$	0,16	0,33	0,67	0,70	0,74

Fonte: ROBINSON et al.,1983.

Empregamos o modelo introduzido por ALI (1994), utilizando doze parâmetros, dois deles de interesse e dois auxiliares. Os dois parâmetros de interesse são a constante de

Michaelis  $k_m$  para a interação substrato-enzima e o coeficiente de variação  $\epsilon^2$  para as propriedades de capilaridades assumidas para constituir o fígado (ROBINSON et al., 1983). Os dois parâmetros auxiliares, para cada fígado  $i$ , são o desvio-padrão  $\sigma_i$  de  $\log V_{ij}$  e a taxa máxima de eliminação  $V_{max}$  de todo o fígado.

A função logarítmica de verossimilhança a ser maximizada pode ser escrita como

$$f(k_m, \epsilon^2, \sigma_1, \dots, \sigma_2, V_{max_1}, \dots, V_{max_5}) = -\sum_{i=1}^5 [n_i \log \sigma_i + R_i^2 / 2\sigma_i^2], \quad (13)$$

onde  $n_i$  é o número de experimentos realizados no  $i$ -ésimo fígado. Os dados são analisados usando uma técnica de estatística bayesiana com uma distribuição de erro  $\log V_{ij}$  considerada normal (ROBINSON et al., 1983). Portanto,

$$R_i^2(V_{max}, K_{max}, \epsilon^2) = -\sum_{j=1}^{n_i} (\log V_{ij} - \log \hat{V}_{ij})^2, \quad (14)$$

onde  $\hat{V}_{ij}$  é o valor do modelo para o  $j$ -ésimo experimento no  $i$ -ésimo fígado, que pode ser obtido resolvendo numericamente a seguinte equação não-linear:

$$\frac{V_{max_i}}{\hat{V}_{ij}} \left\{ (1 - \epsilon)^2 \left\{ \frac{V_{max_i} / 2F_i K_{max}}{1 + \frac{\hat{V}_{ij} / 2F_i K_{max}}{\exp[(V_{max_i} - \hat{V}_{ij}) / F_i K_{max}] - 1}} \right\} \right\} = \frac{K_{max}}{\hat{c}_{ij}} + 1, \quad (15)$$

onde  $\hat{c}_{ij}$  é a média logarítmica de  $c_{ij}$  e  $\bar{c}_{ij}$ , sendo dado por

$$\hat{c}_{ij} = (c_{ij} - \bar{c}_{ij}) / \log(c_{ij} / \bar{c}_{ij}). \quad (16)$$

Os valores das variáveis  $F_i$ ,  $c_{ij}$  e  $\bar{c}_{ij}$  para o  $j$ -ésimo experimento no  $i$ -ésimo fígado são dados na Tabela 2.

**Tabela 3** - Valores dos parâmetros para cada experimento.

$i \downarrow$	$F_i \downarrow$	$j \rightarrow$	1	2	3	4	5
1	1,03	$c_{ij}$	0,14	0,43	0,46	1,69	3,69
		$\bar{c}_{ij}$	0,04	0,20	0,23	1,30	3,20
2	0,94	$c_{ij}$	0,08	0,15	0,28	0,43	1,65
		$\bar{c}_{ij}$	0,02	0,04	0,09	0,17	0,95
3	1,14	$c_{ij}$	0,23	0,36	0,80	1,47	4,12
		$\bar{c}_{ij}$	0,08	0,13	0,48	1,14	3,78
4	1,22	$c_{ij}$	0,16	0,35	2,00	18,4	24,4
		$\bar{c}_{ij}$	0,04	0,13	1,53	18,0	23,6
5	0,96	$c_{ij}$	0,36	0,63	2,63	26,5	-
		$\bar{c}_{ij}$	0,09	0,26	1,92	26,0	-

O espaço de busca utilizado foi o mesmo de Ali et al. (1997):  $\mathbf{x} \in [0,03; 1]^{12}$ . O **ótimo global** está localizado em  $\mathbf{x}^* = (0,2247; 0,1689; 0,0358; 0,0896; 0,0519; 0,0060; 0,0312; 0,3986; 0,4302; 0,6051; 0,4580; 0,7376)$  correspondendo a  $f(\mathbf{x}^*) \approx \mathbf{59,8403}$  (RIOS-COELHO, 2009).

### 3.2.2 Modelo de Reator Catalítico

Este problema de estimação de parâmetros foi introduzido por ROD e HANČIL (1980) para modelar a hidrogenação catalítica na fase de gás do fenol em um reator catalizador pseudodiferencial.

O problema anterior foi resolvido utilizando mínimos quadrados, assumindo que as variáveis não estavam sujeitas a erros de medição. Para este problema consideramos os erros de medição para todas as variáveis. E, ao invés de simplesmente resolver por mínimos quadrados, é necessário utilizar uma aproximação por máxima verossimilhança (TJOA e BIEGLER, 1992). A função objetivo para este modelo tem a seguinte forma (GAU e STADTHERR, 2002):

$$\min_{\theta, \tilde{x}_i} \sum_{i=1}^m \sum_{j=1}^n \frac{(\tilde{x}_{ij} - x_{ij})^2}{\sigma_j^2} \quad (17)$$

$$\text{sujeito a } \mathbf{f}(\theta, \mathbf{x}_i) = \mathbf{0}, \quad \mathbf{i} = \mathbf{1}, \dots, \mathbf{m}. \quad (18)$$

Nas equações acima,  $x_i = (x_{i1}, \dots, x_{in})^T$  representa a medição das variáveis dos experimentos de  $i = 1, \dots, m$  e  $\tilde{x}_i = (\tilde{x}_{i1}, \dots, \tilde{x}_{in})^T$  são os valores desconhecidos correntes e  $\sigma_j$  é o desvio padrão associado à variável  $j$  (GAU e STADTHERR, 2002). Contudo, ao aumentarmos a dimensão do problema de otimização, a aproximação por máxima verossimilhança envolve não somente o parâmetro  $\theta$ , mas também o valor teórico de  $x_i$ .

O modelo é descrito pela equação:

$$x_3 = \frac{\theta_1 \theta_2^2 \theta_3 x_1 x_2^2}{(1 + \theta_1 x_1 + \theta_2 x_2)^3}, \quad (19)$$

As variáveis  $x_1$ ,  $x_2$  e  $x_3$  representam a pressão parcial do fenol e do hidrogênio e a taxa de reação inicial.  $\theta_1$ ,  $\theta_2$  e  $\theta_3$  são os parâmetros a serem estimados. O desvio padrão é especificado em 0,0075 para  $x_1$ , 0,0075 para  $x_2$  e 2,5 para  $x_3$  (TJOA e BIEGLER, 1992). A tabela 4 mostra os dados e seus valores estimados.

**Tabela 4** – Dados e seus valores estimados.

<b>X<sub>1</sub></b>		<b>X<sub>2</sub></b>		<b>X<sub>3</sub></b>	
<b>Dado</b>	<b>Valor Estimado</b>	<b>Dado</b>	<b>Valor Estimado</b>	<b>Dado</b>	<b>Valor Estimado</b>
0,015	0,018	0,235	0,236	6,25	2,54
0,030	0,031	0,220	0,220	4,90	3,11
0,045	0,045	0,205	0,205	2,90	3,21
0,100	0,100	0,150	0,150	1,75	1,94
0,180	0,180	0,070	0,070	0,30	0,35
0,015	0,023	0,485	0,486	12,30	8,88
0,030	0,034	0,470	0,471	14,00	10,83
0,045	0,038	0,455	0,453	5,00	10,77
0,045	0,047	0,455	0,456	14,20	11,82
0,100	0,100	0,400	0,400	10,81	10,73
0,143	0,143	0,357	0,357	7,81	8,12
0,167	0,167	0,333	0,333	6,41	6,72
0,250	0,250	0,250	0,250	3,90	3,06
0,333	0,333	0,167	0,167	3,60	1,09
0,030	0,023	0,720	0,720	13,00	14,74
0,045	0,044	0,705	0,705	20,00	20,72
0,100	0,100	0,650	0,649	19,81	22,45
0,180	0,180	0,570	0,570	15,10	15,88
0,240	0,241	0,510	0,510	8,90	11,10
0,300	0,300	0,450	0,450	7,50	7,51
0,360	0,360	0,390	0,389	2,00	4,85
0,026	0,015	0,974	0,974	13,00	14,56
0,050	0,048	0,950	0,950	30,00	30,67
0,100	0,100	0,900	0,901	37,50	34,89
0,250	0,248	0,750	0,752	25,00	20,52
0,150	0,150	0,850	0,850	31,50	30,98
0,333	0,334	0,667	0,666	10,00	13,34
0,500	0,500	0,500	0,500	4,00	5,26

Os **parâmetros ótimos** deste problema com 59 variáveis são  $\theta_1 = 7,39696 \text{ atm}^{-1}$ ,  $\theta_2 = 0,63782 \text{ atm}^{-1}$  e  $\theta_3 = 1769,71 \text{ mol/kg h}$  para um valor de  $f(x^*) = 30,3072$  e para as variáveis  $x_i$ , utilizamos a mesma região de busca  $[x_{1i} - 3\sigma_1, x_{1i} + 3\sigma_1]$  e  $[x_{2i} - 3\sigma_2, x_{2i} + 3\sigma_2]$ , para  $i = 1, \dots, 28$ . (GAU AND STADTHER, 2002).

Já, para os parâmetros a serem estimados, adotamos uma região de busca em um intervalo mais amplo do que a utilizada por estes autores,  $\theta_1, \theta_2 \in [0, 10]$  e  $\theta_3 \in [1000, 2000]$ .

### 3.2.3 Modelo Cinético de Primeira Ordem

Uma reação de primeira ordem é aquela na qual a velocidade é diretamente proporcional à concentração do reagente. São consideradas irreversíveis aquelas nas quais pelo menos um dos reagentes é totalmente consumido ao final da reação.

Neste problema temos um modelo altamente não linear que descreve uma reação irreversível de primeira ordem ( $A \rightarrow B$ ) em um tanque de mistura com batelada. O conteúdo do reator é misturado completamente. Todos os elementos são expostos ao tratamento por um tempo igual à permanência do substrato no reator. Este reator é usado em aplicações experimentais de laboratório (o frasco no qual se processa o teste da Demanda Bioquímica do Oxigênio-DBO é um exemplo), ou nos processos com pequeno volume a ser tratado.

Assim, dada uma reação da forma  $A \rightarrow B$ , pode-se definir a variação da concentração de B com o passar do tempo pela seguinte equação:

$$y = \theta_1 \left(1 - e^{-\frac{\theta_2 x}{\theta_1}}\right), \quad (20)$$

onde, em sua formulação original (BATES e WATTS, 1988 *apud* SCHWAAB et al, 2008),  $y$  é a variável dependente e corresponde à concentração de B,  $x$  é a variável independente e corresponde ao tempo de reação e  $\theta_1$  e  $\theta_2$  são os parâmetros do modelo que serão ajustados.

O modelo definido na Equação (20) foi ajustado aos dados experimentais apresentados na Tabela 5, sendo minimizada a função de mínimos quadrados. Os dados apresentados referem-se a medidas da demanda bioquímica de oxigênio em  $mg/L$  a variável dependente, em função do tempo de amostragem em *dias*, a variável independente. (SCHWAAB et al, 2008).

**Tabela 5** – Demanda bioquímica de oxigênio ao longo do tempo.

$x(\text{dias})$	$y(\text{mg} / \text{L})$
1,0	8,3
2,0	10,3
3,0	19,0
4,0	16,0
5,0	15,6
7,0	19,8

Fonte: (BATES e WATTS, 1988 *apud* SCHWAAB et al, 2008).

O resultado obtido foi um **valor mínimo de 25,99 da função objetivo**, com os parâmetros,  $\theta 1$  e  $\theta 2$  iguais a **19,14 e 0,5311**, respectivamente

Finalmente, na tabela 6 fazemos um resumo dos ótimos globais dos problemas aqui abordados.

**Tabela 6** – Resumo dos ótimos globais e  $f(x^*)$  das funções e modelos testados.

Funções	Ótimo(s) Global(is) ( $x^*$ )	$f(x^*)$
Easom	$(\pi, \pi)$	-1
Shekel's Foxholes	$(-32, -32)$	499,002
Branim	$(-\pi; 12,275), (\pi; 2,275),$ $(3\pi; 2,475)$	$5/(4\pi)$
Himmelblau	$(3,000; 2,000), (3,584; -1,848),$ $(-3,779; -3,283), (-2,805; 3,131)$	$f(x^*) = 0.$
Pig Liver Problem (PLP)	$(0,2247; 0,1689; 0,0358;$ $0,0896; 0,0519; 0,0060; 0,0312;$ $0,3986; 0,4302; 0,6051; 0,4580;$ $0,7376)$	$\approx 59,8403$
Modelo de Reator Catalítico	$\theta_1 = 7,39696 \text{ atm}^{-1}, \theta_2 =$ $0,63782 \text{ atm}^{-1}$ e $\theta_3 = 1769,71$ $\text{mol/kg h}$	30,3072
Modelo Cinético de 1ª Ordem	$\theta_1 = 19,14$ e $\theta_2 = 0,5311$	25,99

## 4 RESULTADOS E DISCUSSÕES

### 4.1 O AMBIENTE DE TESTES

Os métodos de otimização introduzidos neste trabalho foram todos implementados em linguagem de programação C++ (RILEY, 2005). Três funções utilizadas em alguns destes métodos são adaptações de programas-fonte disponibilizados por pesquisadores para utilização para fins científicos, conforme explicaremos a seguir.

O código gerador dos pontos iniciais pela sequência de Sobol é uma função adaptada a partir do código disponibilizado por JOE e KUO (2010), que foi originalmente apresentado por JOE e KUO (2003). O método de Sobol gera pontos no hipercubo  $n$ -dimensional unitário. Portanto, cada componente de um ponto terá valor entre 0 e 1. Como é necessário "desnormalizar" o  $i$ -ésimo componente, utilizamos a seguinte equação:

$$x_i = l_{inf_i} + s_i(l_{sup_i} - l_{inf_i}), \quad (21)$$

onde  $s_i$  é o valor normalizado gerado pela sequência *quasi*-aleatória de Sobol.

Finalmente, a função que executa o método Rosenbrock com a modificação proposta por PALMER (1968) foi adaptada de um código fonte em C disponibilizado por Vanden Berghen em sua tese doutorado (VANDEN BERGHEN, 2004).

Todos os nossos experimentos foram realizados em um *notebook* HP Envy 15 com processador Intel Core  $i7^{\text{TM}}$  64 *bits*, com 1,6 GHz e 6 Gb de RAM rodando no Windows 7 Professional versão 64 *bits*. Os códigos foram compilados com o GNU gcc adaptado para Windows (MinGW) utilizando o compilador Code::Blocks versão 10.5.

### 4.2 AVALIAÇÃO DOS MÉTODOS DE OTIMIZAÇÃO

#### 4.2.1 Critérios de avaliação

Utilizamos como principal critério de avaliação, o número de avaliações da função objetivo necessário para obter o melhor valor em cada execução bem sucedida com os nossos métodos que foi comparado com **número médio** de avaliações da função-objetivo necessárias para obter o melhor pelos outros métodos.

#### 4.2.2 A metodologia experimental

Os métodos foram aplicados às funções de teste Shekel's Foxholes, Branin, Himmelblau e Eason e a três problemas consagrados na literatura: *Pig Liver Problem* (PLP), Modelo de Reator Catalítico, Problema de Cinética de Primeira Ordem, já descritos anteriormente.

Realizamos para cada método aplicado a cada problema **cem** execuções independentes.

Os tamanhos das populações adotados para todos os métodos foram  $N=100$  e  $N=200$ . Este critério foi utilizado porque supõe-se que **quanto maior a dimensão do espaço de busca, maior a população necessária para obter sucesso na busca pelo ótimo global**. Uma população fixa para problemas de várias dimensões poderia levar a poucos indivíduos para um problema de alta dimensionalidade ou a muitos indivíduos para um problema de baixa dimensionalidade, levando a um esforço computacional desnecessário.

Os parâmetros adotados no **algoritmo Hooke-Jeeves** foram  $\alpha = 0,8$ ;  $\Delta_{\text{inicial}} = 10^{-3}$ ;  $\varepsilon = 10^{-6}$  e **número máximo de iterações igual a 5000**. O algoritmo é encerrado quando  $\Delta \leq \varepsilon$  ou quando o número de iterações atingir o máximo, o que ocorrer primeiramente.

Para o **algoritmo Rosenbrock**, os parâmetros adotados foram  $\alpha = 2,0$ ;  $\beta = -0,5$ ;  $\varepsilon = 10^{-6}$  e **número máximo de iterações igual a 5000**. O algoritmo também é encerrado quando  $\Delta \leq \varepsilon$  ou quando o número de iterações atingir o máximo, o que ocorrer primeiramente.

Nossos algoritmos são concebidos para problemas de **minimização**. No caso do **PLP**, que é um problema de **maximização**, basta **multiplicar a sua função objetivo por  $-1$** , resolvendo-o então como um problema de minimização.

#### 4.2.3 Resultados do TA aplicado aos problemas abordados

Nesta subseção, serão mostrados os resultados obtidos pelo TA aplicado aos problemas descritos. Exibiremos, primeiramente, os resultados obtidos por cada variante do novo método em cada problema e, em seguida, os resultados consolidados comparando-os com os da literatura.

As tabelas 7 e 8 mostram os resultados obtidos pelo método para as funções de teste da literatura, com busca local pelos métodos Hooke-Jeeves e o Rosenbrock, respectivamente.

**Tabela 7** – Resultados obtidos com busca local pelo Hooke-Jeeves.

<b>MÉTODO Sbl-HT-HJ</b>								
<b>Número de Avaliações de Função Objetivo</b>								
<b>N</b>	<b>K</b>	<b>Easom</b>	<b>Shekel's Foxholes</b>	<b>Branin</b>	<b>Himmelblau</b>	<b>PLP</b>	<b>Reator Catalítico</b>	<b>Cinética de 1ª Ordem</b>
100	5	365	362	699	784	2130	16242	702
	10	365	362	699	1.443	2130	6651	702
	15	365	246	699	1.618	2130	6651	462
	20	365	246	699	1.112	2130	6651	462
	25	365	246	699	1.334	2130	6651	462
200	5	465	462	799	884	2230	16342	802
	10	465	462	1.031	1.543	2230	16342	562
	15	465	462	1.020	1.212	2230	6751	562
	20	465	346	790	1.212	2230	6751	562
	25	465	346	790	1.434	2230	6751	562

**Tabela 8** – Resultados obtidos com busca local pelo Rosenbrock.

<b>MÉTODO Sbl-HT-Rsbk</b>								
<b>Número de Avaliações de Função Objetivo</b>								
<b>N</b>	<b>K</b>	<b>Easom</b>	<b>Shekel's Foxholes</b>	<b>Branin</b>	<b>Himmelblau</b>	<b>PLP</b>	<b>Reator Catalítico</b>	<b>Cinética de 1ª Ordem</b>
100	5	3.190	1.920	1.086	2.066	6006	468444	1240
	10	2.716	926	1.086	1.764	6006	493224	1240
	15	2.350	<b>474*</b>	1.086	1.240	6006	493224	980
	20	2.278	<b>474*</b>	1.086	1.240	6006	493224	980
	25	1.990	<b>474*</b>	1.086	1.018	6006	493224	980
200	5	3.290	2.020	1.186	2.166	6106	468544	1340
	10	2.816	1.026	2.612	1.864	6106	468544	1080
	15	2.450	2.114	1.384	1.340	6106	493324	1080
	20	2.378	2.038	1.160	1.340	6106	493324	1080
	25	2.090	1.802	1.144	1.746	6106	493324	1080

**\*Falhou!**

A função Easom é unimodal, constituída por um platô com uma "agulha" isolada, onde está o mínimo. Esta função é desafiadora uma vez que a região onde  $x^*$  se encontra é muito pequena e o algoritmo tende a não explorar esta região levando a um valor de função objetivo  $f(x^*)$  diferente de zero, que é o mínimo global da função.

Ambas variantes tiveram desempenho bastante satisfatório para este problema. Para esta função em particular, a variante com busca local pelo Hooke-Jeeves funciona melhor que com o Rosenbrock. Foram necessárias 365 contra 1990 avaliações de função objetivo para chegar em  $f(x^*)=0$ , respectivamente.

O resultado pode ser justificado pela peculiaridade da função Easom, em que o ótimo está na em um local de difícil acesso. Como os eixos do método Rosenbrock vão sendo rotacionados ao longo do processo de busca, isto acaba desfavorecendo este método em detrimento ao Hooke-Jeeves, onde as direções de busca são fixas.

Para a função *Shekel's Foxholes* o método **SBL-HT-HJ** teve um desempenho melhor ainda, sendo necessárias apenas 246 avaliações de função objetivo com  $N=100$  e  $K=25$ . Para estes mesmos padrões o método **Sbl-HT-Rsbk** apresentou falhas. Quando o  $N$  foi aumentado para 200 e com  $K=25$ , o método de busca Hooke-Jeeves teve o seu número de avaliações cerca de 80% menor do que o método Rosenbrock, com redução significativa do esforço computacional.

Em relação às funções Branin e Himmelblau e aos modelos de estimação de parâmetros PLP, Modelo Reator Catalítico e Modelo Cinética de Primeira Ordem o método **Sbl-HT-HJ** também mostrou resultados superiores aos do método **Sbl-HT-Rsbk**, com uma única exceção na função Himmelblau para o  $N=100$  e  $K=25$ , onde o método **Sbl-HT-Rsbk** conseguiu chegar ao valor mínimo global da função com um número de avaliações de função objetivo inferior ao método **Sbl-HT-HJ** (1.018 e 1.334 respectivamente). As maiores diferenças nos resultados foram para o Modelo Reator Catalítico, descrito em 3.2.1.

#### **4.2.4 Comparativo dos resultados consolidados pelos métodos Sbl-HT-HJ e Sbl-HT-Rsbk com resultados obtidos por outros métodos.**

Aqui, são comparados os resultados obtidos pelos novos métodos (Tabelas 5 e 6) com resultados obtidos por SACCO et al. (2006) utilizando o **Algoritmo Genético**, o *Simulated Annealing* e o **Algoritmo de Colisão de Partículas** (PCA, SACCO et al., 2006), com resultados obtidos por YU e SUGANTHAN (2010) utilizando algoritmos genéticos associados a duas técnicas de manutenção de diversidade populacional, o *Clearing* (CLR, PÉTROWSKI, 1996), o *Ensemble of Niching Algorithms* (ENA, YU e SUGANTHAN,

2010), com resultados obtidos pelo **GLOBAL** (CSENDES et al., 2008), pelo **C-GRASP** (HIRSCH et al., 2010), pelo **HJPCA** (RIOS-COELHO et al., 2010), que são algoritmos que representam o "estado da arte" em otimização global e com quatro métodos híbridos testados por RIOS-COELHO (2011) descritos a seguir:

a) **MT-FCM-HJ** que propõe a inicialização pelo Mersenne Twister, gerador de números pseudo-aleatórios introduzido por MATSUMOTO e NISHIMURA (1998), que possui 623 dimensões de equidistribuição, o que significa que sequências com até 623 números de comprimento são equiprováveis sendo extremamente rápido, evitando as operações de multiplicação e divisão e beneficiando-se da arquitetura dos computadores atuais (MATSUMOTO, NISHIMURA, 1998 *apud* RIOS-COELHO, 2011); a clusterização é feita pelo *Fuzzy Clustering Means* ou *Fuzzy c-Means* (FCM) algoritmo de clusterização nebulosa, apresentado com rigor definitivo por BEZDEK (1981 *apud* RIOS-COELHO, 2011), que reconhece nuvens esféricas de pontos em um espaço  $n$ -dimensional. Cada *cluster* é representado pelo seu centróide, também chamado de protótipo por ser representativo de todas as amostras atribuídas ao *cluster*. O FCM trabalha com um número fixo de classes pré-determinado, daí o nome *c-means*, onde  $c$  é o número de classes. Este algoritmo é definido por HÖPPNER *et al.* (1999) como um algoritmo de clusterização simples, robusto e de fácil implementação e a busca intra-cluster pelo Hooke-Jeeves.

b) **Sobol-FCM-HJ** que propõe a inicialização por Sobol, descrito em 2.1; a clusterização pelo FCM e a busca intra-cluster pelo Hooke-Jeeves, descrito em 2.3.1.

c) **MT-FCM-Rosenbrock** com inicialização pelo Mersenne Twister; clusterização pelo FCM; e buscas intra-cluster pelo Rosenbrock, descrito em 2.3.2.

c) **Sobol-FCM-Rosenbrock** com inicialização por Sobol; clusterização pelo FCM e busca intra-cluster pelo Rosenbrock.

#### 4.2.4.a Função Easom

A tabela 9 mostra os resultados obtidos pelos métodos para a função Easom.

**Tabela 9** – Resultados de outros métodos aplicados para solução da função Easom.

Método	Sucesso	Média	Mediana	D.P.	Máximo	Mínimo
GLOBAL	100/100	1.604	-	-	-	-
C-GRASP	100/100	5.093	-	-	-	-
HJPCA	100/100	1084,8	326	1884,5	11.819	181
MT-FCM-HJ	99/100	2.658,5	1.526	3.448,4	22.601	346
Sobol-FCM-HJ	100/100	2.751,2	2.749	5,8	2.780	2.748
MT-FCM-Rosenbrock	100/100	3.080,9	2.267	2.687,6	12.900	239
Sobol-FCM-Rosenbrock	99/100	4.596,6	3.197	2.512,9	13.547	1.907

Comparando os resultados dos métodos introduzidos neste trabalho com os da tabela 9, percebe-se que a Easom, função unimodal, cujo região do mínimo global possui uma área pequena em relação ao espaço de busca, favorece o método **SOBOL-HT-HJ** que teve o menor número de avaliações de função objetivo (246). Por outro lado, favorece também os algoritmos que buscam apenas um ótimo global, como é o caso dos outros métodos da literatura comparados GLOBAL, C-GRASP e HJPCA. Os métodos que optaram pelo algoritmo determinístico de busca local Rosenbrock, tiveram resultados piores dos que utilizaram o Hooke-Jeeves, fato que pode ser explicado pela particularidade da localização do ótimo da função Easom, em um local de difícil acesso, dificultando a busca feita pelo algoritmo Rosenbrock, que realiza mudanças das direções de busca, sendo mais exploratório que o Hooke-Jeeves, cuja busca se dá em uma só direção.

A clusterização pela heurística topográfica contendo os mínimos locais vizinhos mais próximos, aliado ao método gerador de soluções iniciais SOBOL e ao método de busca local Hooke-Jeeves, mostrou resultados muito superiores aos dos demais métodos comparados no teste da função Easom.

#### 4.2.4.b Função *Shekel's Foxholes*

A Função *Shekel's Foxholes*, descrita em 3.2.1, tem sido pouco utilizada na literatura de otimização global. A Tabela 10 mostra alguns resultados obtidos na tentativa de se chegar ao seu ótimo global.

**Tabela 10** – Resultados da aplicação dos métodos na função *Shekel's Foxholes*.

Método	Sucesso	Média	Mediana	D.P.	Máximo	Mínimo
GA	45/100	9.109	-	-	-	-
SA	100/100	16.971	-	-	-	-
PCA	100/100	7.360	-	-	-	-
CLR	0,28	-	-	-	-	-
ENA	0,996	-	-	-	-	-
MT-FCM-HJ	92/100	6.411,1	4.659	4.950,2	19.973	90
Sobol-FCM-HJ	87/100	6.669,6	6.381	5.171,4	20.742	1.854
MT-FCM-Rsbk	92/100	3.277,0	1.975	3.638,1	16.776	90
Sobol-FCM-Rsbk	100/100	392,0	392	0,0	392	392

O **GA**, que é populacional, não tem bom desempenho para esta função. Isto deve-se ao fato de esta função possuir 24 ótimos locais com valor próximo ao do ótimo global e o algoritmo genético ter propensão a levar toda a população a ficar retida em um ótimo local em domínios multimodais (GOLDBERG, 1989).

O **SA** e o **PCA**, são algoritmos que seguem o paradigma de Metropolis (METROPOLIS et al., 1953), em que na busca pelo ótimo, uma solução pior que a melhor no momento pode ser aceita com uma certa probabilidade, o que evita a convergência para ótimos locais. Isto explica o ótimo desempenho destes métodos na função *Shekel's Foxholes*.

O **CLR** obteve desempenho modesto para um método especialmente projetado para funções multimodais.

A explicação para o desempenho apenas razoável dos métodos **MT-FCM-HJ**, **Sobol-FCM-HJ** e **MT-FCM-Rsbk** vem do fato de a função ter 25 ótimos locais de valores próximos. Em algumas execuções, o procedimento de busca local provavelmente falhou pelo fato de o ótimo global não estar em nenhum dos clusters obtidos pelo método FCM ou estar em uma região inexplorada de algum cluster. O **Sobol-FCM-Rsbk**, por sua vez, chegou ao ótimo global em 100/100 execuções. Isto se deve a duas razões: em primeiro lugar, o método de Sobol preenche o espaço de busca de maneira mais homogênea que o Mersenne-Twister;

em segundo, o algoritmo Rosenbrock, pelas mudanças das direções de busca é mais exploratório que o Hooke-Jeeves.

Porém, o método de busca Hooke-Jeeves, quando associado ao SOBOL e à HT obteve os melhores resultados em relação à todas as demais funções de teste. O método **Sbl-HT-HJ** alcançou o ótimo global em apenas 246 avaliações de função objetivo para  $N=100$  e  $K=15, 20$  e  $25$ . Esta performance foi superior à do método **Sbl-HT-Rsbk**, que conseguiu atingir o ótimo global com 926 avaliações de função objetivo para  $N=100$  e  $K=10$ , apesar do método de busca Rosenbrock ser considerado mais exploratório.

#### 4.2.4.c Função Branin

Como vimos na seção 3.1.3, a função Branin possui três ótimos globais. Na Tabela 11 exibimos os resultados obtidos para esta função.

**Tabela 11** - Resultados de outros métodos aplicados para solução da função Branin.

Método	Sucesso	Média	Mediana	D.P.	Máximo	Mínimo
MT-FCM-HJ	98/100	4.054,4	4062	1.458,0	7.104	1.470
Sobol-FCM-HJ	100/100	3.874,0	3.886	82,8	4.027	3.504
MT-FCM-Rsbk	100/100	2.184,4	2.203	880,7	4.558	988
Sobol-FCM-Rsbk	100/100	3.036,0	3.036	0,0	3.036	3.036

Os algoritmos testados são concebidos para determinar todos os ótimos globais de um domínio multimodal. Neste caso, consideramos que a execução foi bem sucedida apenas se todos os ótimos foram determinados. Os números de avaliações de função objetivo exibidos são aqueles necessários para a determinação de todos os ótimos.

Observe que todos os outros métodos mostrados na Tabela 11 necessitaram em média de menos esforço computacional para encontrarem os três ótimos globais do que o combinação MT-FCM-HJ mais uma vez não obteve 100% de sucesso. A combinação MT-FCM-Rsbk foi a mais bem sucedida, provavelmente por ter distribuído pontos próximos aos três ótimos globais. O Sobol-FCM-Rsbk alia a inicialização quasi-aleatória à maior capacidade exploratória do Rosenbrock e atinge os três ótimos ainda na fase determinística. Comparando os resultados obtidos através dos métodos **Sbl-HT-HJ** e **Sbl-HT-Rsbk**, observa-se que mais uma vez esses resultados foram melhores, indicando uma eficiência

maior dos métodos Sobol e da heurística topográfica na geração de populações iniciais e na clusterização desses indivíduos gerados.

#### 4.2.4.d Função Himmelblau

Conforme vimos na seção 3.1.4, a função Himmelblau possui quatro ótimos globais. Apesar de ser uma função clássica, tem sido pouco utilizada na literatura recente. Somente dispomos de resultados obtidos por YU e SUGANTHAN (2010) conforme descrito na seção 3.2.4. Na Tabela 12 são apresentados os resultados para esta função.

**Tabela 12** - Resultados de outros métodos aplicados para solução da função Himmelblau.

Método	Sucesso	Média	Mediana	D.P.	Máximo	Mínimo
CLR	0,6	-	-	-	-	-
ENA	1,0	-	-	-	-	-
MT-FCM-HJ	100/100	3.905,5	3.794	723,0	6.333	1.953
Sobol-FCM-HJ	100/100	3.369,7	3.536	414,7	4.003	2.621
MT-FCM-Rsbk	100/100	2.237,6	2.254	428,8	3.244	1.206
Sobol-FCM-Rsbk	100/100	2.172,0	2.172	0,0	2.172	2.172

Mais uma vez o CLR obtém desempenho insatisfatório para um método para problemas multimodais. Dadas as características da função, mais uma vez o algoritmo Rosenbrock, devido às mudanças de direções de busca, saiu-se melhor que o Hooke-Jeeves. O esquema de inicialização do Mersenne-Twister associado ao OBL saiu-se ligeiramente pior que os dois outros esquemas. O Sobol-FCM-Rsbk mais uma vez atingiu todos os ótimos ainda na fase determinística. Os métodos **Sbl-HT-HJ** e **Sbl-HT-Rsbk** mais uma vez superaram os resultados dos demais métodos para a função Himmelblau, com uma média no número de avaliações ligeiramente superior no método **Sbl-HT-HJ**.

#### 4.2.4.e Pig Liver Problem (PLP)

Como vimos, o PLP, descrito em 3.2.1, é um problema para estimação de parâmetros em um modelo para eliminação enzimática de substratos do fígado e possui dezenas de ótimos locais e apenas um ótimo global,  $f(\mathbf{x}^*) \approx 59,8403$  (RIOS-COELHO, 2009).

Na literatura de otimização global, há resultados para este problema por Ali et al. (1997) utilizando o TMSL (ALI, 1994), por ALI et al. (2002) aplicando variações do *Simulated Annealing* denominadas DSA, ASA e LSA (ALI et al. 2002) e, mais recentemente, por ALI (2007) utilizando variações do *Controlled Random Search* (PRICE, 1978) denominadas CRS2, CRS2C1 e CRS2C2 e o algoritmo evolução diferencial (STORN; PRICE, 1997) e uma variação denominada DEC (ALI, 2007).

A metodologia experimental varia. ALI et al. (1997) informam simplesmente o melhor resultado obtido e o número de avaliações necessário para chegar nele naquela rodada. ALI et al. (2002) realizam cinco execuções de cada algoritmo e informam o número de sucessos e o número médio de avaliações necessários para chegar ao ótimo global. ALI (2007) utiliza uma metodologia mais próxima à deste trabalho, realizando cem execuções de cada algoritmo e informando o número de sucessos e o número médio de avaliações necessário para chegar ao ótimo global. RIOS-COELHO (2009) aplicou o PLP a uma série de algoritmos, dentre eles o HJPCA, que obteve os melhores resultados. (Tabela 13).

**Tabela 13** - Resultados obtidos para o PLP.

<b>Método</b>	<b>Sucesso</b>	<b>Média</b>	<b>Mediana</b>	<b>D.P.</b>	<b>Máximo</b>	<b>Mínimo</b>
TMSL	1/1	-	-	-	-	2.011
DSA	3/5	53.378	-	-	-	-
ASA	3/5	82.765	-	-	-	-
LSA	2/5	187.326	-	-	-	-
CRS2	55/100	55.294	-	-	-	-
CRS2C1	76/100	49.761	-	-	-	-
CRS2C2	48/100	44.283	-	-	-	-
DE	100/100	74.280	-	-	-	-
DEC	100/100	68.294	-	-	-	-
HJ-PCA	100/100	3.939,5	-	3.655,5	27.927	1.999

O resultado do TMSL é o correspondente à melhor execução. ALI *et al.* (1997) não utilizaram uma metodologia adequada, em que deveria ter sido informado ao menos o valor médio de avaliações para um certo número de execuções. Dos outros algoritmos aplicados por ALI (1997) ao PLP, apenas o DE e sua variante DEC são bem sucedidos, encontrando o máximo global em 100% das execuções. Porém, o esforço computacional requerido é uma ordem de grandeza mais alto que o dos métodos introduzidos neste trabalho. Dentre os nossos métodos, os inicializados pelo Mersenne Twister não obtiveram 100% de aproveitamento. O

método Rosenbrock, mesmo sendo mais exploratório que o Hooke-Jeeves, consegue um desempenho inferior a este quando associado ao SBOL e ao TA. Os métodos propostos neste trabalho foram extremamente bem sucedidos neste problema, com o método **Sbl-TA-HJ** superando inclusive o HJPCA, que possuía os melhores resultados até o momento.

### 3.2.4.f *Modelo de Reator Catalítico*

No caso deste problema, temos na Tabela 14 resultados obtidos pelo Algoritmo Evolução Diferencial Canônico (DE) e por uma versão com inicialização populacional por Sobol (Sobol-DE). Estes resultados foram obtidos recentemente (RIOS-COELHO *et al*, 2014).

**Tabela 14** - Resultados obtidos para o Modelo de Reator Catalítico.

<b>Método</b>	<b>População</b>	<b>Sucesso</b>	<b>Média</b>	<b>Máximo</b>	<b>Mínimo</b>
DE	100	7/100	7.886.874	9.725.345	3.867.260
	500	100/100	579.446	631.204	509.125
Sobol-DE	100	2/100	9.314.050	9.797.103	8.830.998
	500	100/100	584.685	636.740	505.999

Para a população de cem indivíduos, ambas as versões do DE falham na maior parte das execuções. Além disso, mesmo quando bem sucedidas, necessitam de um esforço computacional de uma ordem de grandeza entre  $10^2$  e  $10^3$  da requerida pelas duas versões do TA.

### 3.2.4.g *Modelo Cinético de 1ª Ordem*

Este é um problema clássico da literatura da área de otimização de parâmetros. Recentemente, SCHWAAB *et al.* (2008) atacaram este problema utilizando o PSO. Os autores relatam apenas que foram utilizadas 40 partículas em 1.000 iterações, o que equivale a 40.000 avaliações da função objetivo. O TA necessitou entre 1% e 3% do esforço computacional do PSO, dependendo do tamanho da população e do parâmetro  $K$  utilizado pelo primeiro método.

## 4 CONCLUSÕES E TRABALHOS FUTUROS

Neste trabalho, mostramos a efetividade do algoritmo de otimização topográfica aplicado a problemas práticos de estimação de parâmetros com alta multimodalidade e grau de dificuldade nas áreas de biociências e engenharia química. Os resultados aqui obtidos corroboram as conclusões de SACCO et al. (2014), que no primeiro artigo científico da literatura em que o TA foi aplicado a um problema prático de otimização (no caso, da engenharia nuclear), afirmaram que trata-se de um algoritmo que obtém resultados competitivos com um esforço computacional no mínimo uma ordem de grandeza abaixo do requerido por modernas metaheurísticas, como o algoritmo de enxame de partículas e o método evolução diferencial.

Em termos de trabalhos futuros, pretendemos utilizar métodos de busca direta mais modernos no terceiro passo do algoritmo, conforme CONN et al. (2009) .

## REFERÊNCIAS

- ALI, M. M. Some Modified Stochastic Global Optimization Algorithms with Applications. Tese de Doutorado, Department of Mathematical Sciences, Loughborough University of Technology, Loughborough, UK, 1994.
- ALI, M. M. Synthesis of the  $\beta$ -distribution as an aid to stochastic global optimization. *Computational Statistics & Data Analysis*, v. 52, p. 133 – 149, 2007.
- BATES, D. M., WATTS, D. G. *Nonlinear Regression Analysis and its Applications*. New York: Wiley 1988.
- BAZARAA, M. S.; SHERALI, H. D.; SHETTY, C. M. *Nonlinear Programming: theory and algorithms*. Hoboken, NJ: Wiley Interscience, 2006.
- BECKER, R. W., LAGO, G. V. A global optimization algorithm. In: *Proceedings of the 8<sup>th</sup> Allerton Conference on Circuits and System Theory*, Monticello, IL, USA, p. 3-12, 1970.
- BEZDEK, J.C. *Pattern Recognition with Fuzzy Objective Function Algorithms*. New York: Plenum Press, 1981.
- BRANICKY, M. S.; LaVALLE, S. M.; OLSON, K.; YANG, L. Quasi-randomized path planning. In: *Proceedings of the 2001 IEEE International Conference on Robotics and Automation*, 2001, p. 1481-1487.
- BRATLEY, P.; FOX, B. Algorithm 659: Implementing Sobol's quasirandom sequence generator. *ACM Transactions on Mathematical Software*, v. 14, p. 88-100, 1988.
- BRATLEY, P.; FOX, B. Implementation and tests of low discrepancy sequences. *ACM Transactions on Modeling and Computer Simulation*, v. 2, p. 195-213, 1992.
- COELHO, L. S.; ALOTTO, P. Electromagnetic optimization based on an improved diversity-guided differential evolution approach and adaptive mutation factor. *COMPEL: The International Journal for Computation and Mathematics in Electrical and Electronic Engineering*, v. 28, p. 1112-1120, 2009.
- COELHO, L. S.; SOUZA, R. C. T.; MARIANI, V. C. Improved differential evolution approach based on cultural algorithm and diversity measure applied to solve economic load dispatch problems. *Mathematics and Computers in Simulation*, v. 79, p. 3136-3147, 2009.
- CONN, A. R., SCHEINBERG, K., VICENTE, L. N. *Introduction to Derivative-Free Optimization*. Philadelphia: SIAM, 2009.
- CORANA, A.; MARCHESI, M.; MARTINI, C; RIDELLA, S. Minimizing multimodal functions of continuous variables with the “simulated annealing” algorithm. *ACM Transactions on Mathematical Software*, v. 13, p. 262, 1987.
- CSENDES, T.; PÁL, L.; SENDÍN, J. O. H.; BANGA, J. R. The global optimization method revisited. *Optimization Letters*, v. 2, p. 445-454, 2008.

DE JONG, K. A. An analysis of the behavior of a class of genetic adaptive systems. Tese de doutorado, University of Michigan, Ann Arbor, MI, 1975.

DIXON, L. C. W.; SZEGÖ, G. P. Towards Global Optimization 2. Amsterdam: North-Holland, 1978.

GALANTI, S., JUNG, A. Low-discrepancy sequences. *J. Deriv.*, v. 3, p. 63–83, 1997.

GAU, C. Y., STADTHER, M. A. Deterministic Global Optimization for Errors-in-Variables Parameter Estimation. *AIChE J.*, v. 48, p. 1192–1197, 2002.

GENTLE, J. E. Random Number Generation and Monte Carlo Methods. New York: Springer, 2005.

GROSAN, C.; ABRAHAM, A. A New Approach for Solving Nonlinear Equations Systems. *IEEE Transactions on Systems, Man and Cybernetics-Part A: Systems and Humans*, v. 38, p. 698-714, 2008.

HENDERSON N.; SACCO, W. F. Prediction of double retrograde vaporization by hybrid global-local optimization using fuzzy clustering means. *Chemical Product and Process Modeling*, v. 3, p. 47, 2008.

HENDERSON N.; SACCO, W. F.; PLATT, G. M. Finding more than one root of nonlinear equations via a polarization technique: An application to double retrograde vaporization. *Chemical Engineering Research and Design*, v. 88, p. 551-561, 2010.

HIRSCH, M. J.; PARDALOS, P. M.; RESENDE, M. G. C. Solving systems of nonlinear equations with continuous GRASP. *Nonlinear Analysis: Real World Applications*, v. 10, p. 2000-2006, 2009.

HIRSCH, M. J.; PARDALOS, P. M.; RESENDE, M. G. C. Speeding up continuous GRASP. *European Journal of Operational Research*, v. 205, p. 507-521, 2010.

HOLLAND, J. H. Adaptation In Natural and Artificial Systems. Ann Arbor: University of Michigan Press, 1975.

HOOKE, R.; JEEVES, T. A. Direct search solution of numerical and statistical problems. *Journal of the Association for Computing Machinery*, v. 8, p. 212–229, 1961.

HÖPPNER, F.; KLAWONN, F.; KRUSE, R.; RUNKLER, T. Fuzzy Cluster Analysis. Chichester, England: John Wiley & Sons, Ltd, 1999.

HUA, L. K.; WANG, Y. Applications of Number Theory to Numerical Analysis. Berlin: Springer, 1981.

JOE, S.; KUO, F. Y. Remark on Algorithm 659: Implementing Sobol's quasirandom sequence generator. *ACM Transactions on Mathematical Software*, v. 29, p. 49–57, 2003.

JOE, S.; KUO, F. Y. Sobol sequence generator. Disponível em: <http://web.maths.unsw.edu.au/~fkuo/sobol>> Acesso em: jan. 2014.

- KIRKPATRICK, S.; GELATT, C. D.; VECCHI, M. P. Optimization by simulated annealing. *Science*, v. 220, p. 671–680, 1983.
- LEWIS, M. L.; TORCZON, V; TROSSET, M. W. Direct search methods: then and now. *Journal of Computational and Applied Mathematics*, v. 124, p. 191-207, 2000.
- MAHFOUD, S. W. Niching Methods for Genetic Algorithms. Tese de doutorado, Illinois Genetic Algorithm Laboratory, University of Illinois at Urbana- Champaign, Urbana-Champaign, IL, USA, 1995.
- MATSUMOTO, M. Mersenne Twister Home Page. Disponível em: <<http://www.math.sci.hiroshima-u.ac.jp/~m-mat/MT/emt.html>>. Acesso em: set. 2010.
- MATSUMOTO, M.; NISHIMURA, T. Mersenne twister: a 623-dimensionally equidistributed uniform pseudo-random number generator. *ACM Transactions on Modeling and Computer Simulation*, v. 8, p. 3–30, 1998.
- METROPOLIS, N.; ROSENBLUTH, A. W., ROSENBLUTH, M. N.; TELLER, A. H.; TELLER, E. Equations of State Calculations by Fast Computing Machines, *Journal Chem. Phys.*, v.21, p.1087-1092, 1953.
- MICHALEWICZ, Z. *Genetic Algorithms + Data Structures = Evolution Programs*. Berlin: Springer Verlag, 1996.
- NIEDERREITER, H. Quasi-Monte Carlo methods for global optimization. In: *Proceedings of the 4th Pannonian Symposium on Mathematical Statistics: 1986*, p. 251-267.
- PALMER, J. R. An improved procedure for orthogonalising the search vectors in Rosenbrock's and Swann's direct search optimisation methods. *The Computer Journal*, v. 12, p. 69–71, 1968.
- PÉTROWSKI, A. A Clearing Procedure as a Niching Method for Genetic Algorithms. In: *Proceedings of the International Conference on Evolutionary Computation 1996 (ICEC '96)*, p. 798-803, 1996.
- PRESS, W. H., TEUKOLSKY, S. A., VETTERLING, W. T., FLANNERY, B. P. *Numerical Recipes – The Art of Scientific Computing*. Cambridge, UK: Cambridge University Press, 2007.
- RILEY, R. *C++*. London: Hodder & Stoughton Educational, 2005.
- RIOS-COELHO, A. C. Estimaco de parmetros em modelos para eliminao enzimtica de substratos do fgado: um estudo via otimizao global. Dissertao de mestrado, Programa de Ps-Graduao em Modelagem Computacional, Instituto Politcnico da Universidade do Estado do Rio de Janeiro, Nova Friburgo, RJ, 2009.
- \_\_\_\_\_ Otimizao multimodal atravs de novas tcnicas baseadas em clusterizao nebulosa. Tese de doutorado, Programa de Ps-Graduao em Modelagem Computacional, Instituto Politcnico da Universidade do Estado do Rio de Janeiro, Nova Friburgo, RJ, 2011.

RIOS-COELHO, A. C.; SACCO, W. F.; HENDERSON, N. A Metropolis algorithm combined with the Hooke-Jeeves local search method applied to global. *Applied Mathematics and Computation*, v. 217, p. 843–853, 2010.

\_\_\_\_\_ On initial populations of differential evolution for practical optimization problems. *Submetido ao periódico internacional indexado Swarm and Evolutionary Computation*, 2014.

ROD, V., HANČIL, V. Iterative Estimation of Model Parameters when Measurements of All Variables are Subject to Error. *Comput. Chem. Eng.*, v. 4, p. 33-38, 1980.

ROSENBROCK, H. H. An Automatic Method for Finding the Greatest or Least Value of a Function. *Computer Journal*, v. 3, p. 175-184, 1960.

ROBINSON, P. J.; PETTITT, A. N.; ZORNING, J.; BASS, L. A Bayesian Analysis of Capillary Heterogeneity in the Intact Pig Liver. *Biometrics*, v. 39, p. 61–69, 1983.

SACCO, W. F.; MACHADO, M. D.; PEREIRA, C. M. N. A.; SCHIRRU, R. The fuzzy clearing approach for a niching genetic algorithm applied to a nuclear reactor core design optimization problem. *Annals of Nuclear Energy*, v. 31, p. 55-69, 2004.

SACCO, W. F.; DE OLIVEIRA, C. R. E.; PEREIRA, C. M. N. A. Two stochastic optimization algorithms applied to nuclear reactor core design. *Progress in Nuclear Energy*, v. 48, p. 525-539, 2006.

SACCO, W. F.; HENDERSON, N.; RIOS-COELHO, A. C.; ALI, M. M.; PEREIRA, C. M. N. A. Differential evolution algorithms applied to nuclear reactor core design. *Annals of Nuclear Energy*, v. 36, p. 1093-1099, 2009.

SACCO, W. F.; HENDERSON, N. Finding All Solutions of Nonlinear Systems Using a Hybrid Metaheuristic with Fuzzy Clustering Means. *Applied Soft Computing*, v. 11, p. 5424-5432, 2011.

SACCO, W. F.; HENDERSON, N.; RIOS-COELHO, A. C. Topographical global optimization applied to nuclear reactor core design: Some preliminary results. *Annals of Nuclear Energy*, v. 65, p. 177-163, 2014.

SCHWAAB, M., BISCAIA JR., E. C., MONTEIRO, J. L., PINTO, J. C. Nonlinear parameter estimation through particle swarm optimization. *Chemical Engineering Science*, v. 63, p. 1542-1552, 2008.

SOBOL, I. M. On the distribution of points in a cube and the approximate evaluation of integrals. *USSR Computational Mathematics and Mathematical Physics*, v. 7, p. 86–112, 1967.

STORN, R., PRICE, K. Differential Evolution - A simple and efficient heuristic for global optimization over continuous spaces. *Journal of Global Optimization*, v. 11, p. 341–359, 1997.

TIMMER, G.T. Global Optimization – A Stochastic Approach, Ph.D. Thesis, Erasmus University, Rotterdam, The Netherlands, 1984.

TJOA, I. B., BIEGLER, L. T. Reduced Successive Quadratic Programming Strategy for Errors-in-Variables Estimation. *Comput. Chem. Eng.*, v. 16, p. 523–533, 1992.

TÖRN, A. Global optimization as a combination of global and local search. In: *Proceedings of Computer Simulation Versus Analytical Solutions for Business and Economic Models*, Gothenburg, Sweden, p. 191-206, 1973.

\_\_\_\_\_ A search clustering approach to global optimization. In: L.C.W. Dixon and G.P. Szegö (Eds.), *Towards Global Optimization 2*, North-Holland Publishing Company, p. 49-62, 1978.

TÖRN, A.; ZILINSKAS, A. *Global Optimization - Lecture Notes in Computer Science*. Berlin: Springer-Verlag, 1989.

TÖRN, A.; VIITANEN, S. Topographical Global Optimization. In: C.A. Floudas and P.M. Pardalos (Eds.), *Recent Advances in Global Optimization*, Princeton University Press, Princeton, NJ, p. 384-398, 1992.

\_\_\_\_\_ Topographical Global Optimization Using Pre-Sampled Points. *J. Glob. Optim.* v. 5, p. 267-276, 1994.

\_\_\_\_\_ Iterative Topographical Global Optimization. In: C.A. Floudas and P.M. Pardalos (Eds.), *State of the Art in Global Optimization*, Kluwer Academic Publishers, the Netherlands, p. 353-363, 1996.

VANDEN BERGHEN, F. Constrained, non-linear, derivative-free, parallel optimization of continuous, high computing load, noisy objective functions. Tese de doutorado, Faculté de Sciences Appliquées, Université Libre de Bruxelles, Bruxelles, Belgique, 2004.

YU, E. L.; SUGANTHAN, P. N. Ensemble of niching algorithms. *Information Sciences*, v. 180, p. 2815-2833, 2010.