

GUSTAVO DELA BRUNA NORONHA

**AKMCS E SUAS VARIAÇÕES:  
PROCEDIMENTO DE CONFIABILIDADE  
UTILIZANDO KRIGING E  
SIMULAÇÃO DE MONTE CARLO**

Florianópolis/SC

2018



GUSTAVO DELA BRUNA NORONHA

**AKMCS E SUAS VARIAÇÕES:  
PROCEDIMENTO DE CONFIABILIDADE  
UTILIZANDO KRIGING E  
SIMULAÇÃO DE MONTE CARLO**

Dissertação submetida à Universidade Federal de Santa Catarina como requisito parcial exigido pelo programa de Pós-Graduação em Engenharia Civil - PPGEC, para obtenção do título de MESTRE em Engenharia Civil.

Universidade Federal de Santa Catarina  
Centro Tecnológico  
Programa de Pós-Graduação em Engenharia Civil

Orientador: Rafael Holdorf Lopez

Florianópolis/SC

2018

Florianópolis, 10 de Outubro de 2018

**AKMCS E SUAS VARIAÇÕES:  
PROCEDIMENTO DE CONFIABILIDADE  
UTILIZANDO KRIGING E  
SIMULAÇÃO DE MONTE CARLO**

GUSTAVO DELA BRUNA NORONHA

Dissertação submetida à Universidade Federal de Santa Catarina como requisito parcial exigido pelo programa de Pós-Graduação em Engenharia Civil - PPGEC, para obtenção do título de MESTRE em Engenharia Civil.

---

Eneir Ghisi, Dr. - Coordenador PPGEC

Comissão Examinadora:

---

Prof. Rafael Holdorf Lopez, Dr. - Orientador

---

André Teófilo Beck, Dr. - EESC/USP

---

Leandro Fleck Fadel Miguel, Dr. -  
ECV/UFSC

---

Marcos Souza Lenzi, Dr. - ECV/UFSC

# Resumo

Este estudo tem como objetivo descrever e analisar alguns métodos de Kriging com Monte Carlo disponíveis na literatura, demonstrando assim o estado da arte deste tema em expansão. Este métodos utilizam do metamodelo de Kriging para substituir a função de desempenho dos problemas para resolver a análise de confiabilidade por meio da simulação de Monte Carlo, sendo sua principal intenção reduzir o custo computacional necessário. A pesquisa deste tipo de solução se popularizou com o primeiro método conhecido como AKMCS em 2010 e passou a ser explorado e estudado por uma grande quantidade de profissionais. Com os métodos de Kriging com Monte Carlo é possível realizar análises de confiabilidade com alta precisão e baixo custo computacional, sendo possível utilizar as estatísticas do metamodelo de Kriging para gerar métodos ainda mais eficientes. Com a pesquisa crescente e avanços disponíveis, propõem-se dar informações suficientes para compreender o funcionamento do modelo de Kriging com a simulação de Monte Carlo simples, por importância ou com o uso de Kriging universal e analisar os resultados demonstrados na literatura e para alguns problemas de teste adicionais. Os resultados disponíveis indicam a tendência da pesquisa do Kriging universal ser ampliada, assim como justifica parcialmente a menor quantidade de trabalhos sobre o Kriging com simulação por importância, indicando também as vantagens e desvantagens para cada método descrito.

**Palavras-chaves:** Kriging, AKMCS, confiabilidade, metamodelo, simulação de Monte Carlo.



# Abstract

This study aims to describe and analyze some methods using Kriging with Monte Carlo simulation available in the literature, thus demonstrating the state of the art of this expanding subject. These methods use the Kriging metamodel to replace the performance function to solve reliability analysis through Monte Carlo simulation, and its main intention is to reduce the computational cost required. The research of this type of solution became popular with the first method known as AKMCS in 2010 and started to be explored and studied by a large number of researchers. With Kriging with Monte Carlo methods it is possible to perform reliability analyzes with high precision and low computational cost, being possible to use Kriging metamodel statistics to generate even more efficient methods. With the increasing research and available advances, it is proposed to give sufficient information to understand the Kriging model with the crude Monte Carlo simulation, importance simulation or with the use of universal Kriging and analyze the results demonstrated in the literature and for some additional test problems. The available results demonstrate the tendency of expanding the studies about universal Kriging, as well as partially justify the smaller number of research about Kriging with importance simulation, also indicating the advantages and disadvantages for each method described.

**Keywords:** Kriging, AKMCS, reliability, metamodel, Monte Carlo simulation.





# Lista de ilustrações

|   |    |
|---|----|
| Figura 1 – Modelo estrutural de um edifício. . . . .  | 16 |
| Figura 2 – Mais exemplos de modelos estruturais. . . . .  | 16 |
| Figura 3 – Covariância próxima de zero para variáveis dependentes. . . . .  | 26 |
| Figura 4 – Funções de distribuição de uma variável normal padrão . . . . .  | 27 |
| Figura 5 – Esquema do AKMCS geral . . . . .   | 43 |
| Figura 6 – Comparação entre $\hat{P}_{fI}$ (curva superior em azul), $\hat{P}_{f\pi}$ (curva inferior em vermelho) e $P_{fSMC}$ (linha tracejada, valor de referência). . . . . | 53 |
| Figura 7 – Comparação da PDF ideal(a) e da quase-ótima por meio de $\hat{h}(b)$ . . . . .   | 63 |
| Figura 8 – Solução exata da função $G(\mathbf{x})$ para o problema 1. . . . .   | 68 |
| Figura 9 – Visualização do domínio de falha para o problema 1. . . . .  | 69 |
| Figura 10 – Pontos de suporte usados pelo AKMCS+U no problema 1. . . . .  | 70 |
| Figura 11 – Pontos de suporte mais pontos de amostragem por importância do Meta-IS. . . . .   | 71 |
| Figura 12 – Solução exata da função $G(\mathbf{x})$ para o problema 2. . . . .  | 72 |
| Figura 13 – Visualização do domínio de falha para o problema 2. . . . .   | 72 |
| Figura 14 – Pontos de suporte usados pelo AKMCS+U no problema 2. . . . .  | 73 |
| Figura 15 – Solução exata da função $G(\mathbf{x})$ para o problema 3. . . . .  | 74 |
| Figura 16 – Visualização do domínio de falha para o problema 3. . . . .   | 74 |



# Lista de tabelas

|  |    |
|--|----|
| Tabela 1 – Probabilidade máxima de superar o valor de $e_I$ . . . . .  | 52 |
| Tabela 2 – Comparação número de avaliações da função de desempenho para diferentes números de pontos de suporte adicionados por iteração . . . | 58 |
| Tabela 3 – Resultados para o problema 1 . . . . .  | 69 |
| Tabela 4 – Resultados para o problema 2 . . . . .  | 73 |
| Tabela 5 – Resultados para o problema 3 . . . . .  | 75 |
| Tabela 6 – Resultados para o problema 4 . . . . .  | 76 |



# Lista de símbolos

|                 |  |
|-----------------|--|
| $A, B$          | Conjunto de valores representando um evento. Também foi usado para identificar as equações 2.59 e 2.62.            |
| $\Omega$        | Conjunto com todas as ocorrências do espaço amostral.  |
| $P[A]$          | Probabilidade de ocorrência de um evento $A$ .   |
| $U$             | Operador de união entre conjuntos.   |
| $\cup$          | Operador de intersecção entre conjuntos.   |
| $F_X(x)$        | Função de distribuição acumulada para a variável aleatória $X$ para um valor $x$ .                                 |
| $F_{X,Y}(x, y)$ | Função de distribuição acumulada conjunta das variáveis aleatórias $X$ e $Y$ para um par de valores $(x, y)$ .     |
| $f_X(x)$        | Função de densidade de probabilidade da variável aleatória $X$ para um valor $x$ .                                 |
| $f_{X,Y}(x, y)$ | Função de densidade de probabilidade conjunta das variáveis aleatórias $X$ e $Y$ para um par de valores $(x, y)$ . |
| $E[X]$          | Operador de expectativa para a variável $X$ .  |
| $\mu_X$         | Valor médio para a variável $X$ .  |
| $Var[X]$        | Operador de variância aplicado à variável $X$ .  |
| $Cov[X, Y]$     | Operador de covariância aplicado às variáveis $X$ e $Y$ .  |
| $\sigma_X$      | Valor do desvio-padrão para a variável $X$ .   |
| $\delta_X$      | Valor do coeficiente de variação para a variável $X$ .   |
| $p_X(x)$        | Probabilidade da variável aleatória $X$ ter um valor igual a $x$ .   |
| $p_{X,Y}(x, y)$ | Probabilidade conjunta de um par de variáveis $X$ e $Y$ ser igual ao par de valores $(x, y)$ .                     |
| $p_{X Y}(x y)$  | Probabilidade de massa condicional da variável aleatória $X$ em relação à variável $Y$ .                           |
| $\hat{f}(x)$    | Aproximação para a função $f(x)$ .   |

|                                  |  |
|----------------------------------|--|
| $\hat{\mu}_X$                    | Aproximação para o valor de $\mu_X$ .  |
| $E_f, P_f$                       | Evento falha e probabilidade de ocorrência do evento falha.  |
| $\Omega_f$                       | Conjunto com todas as ocorrências do evento falha no espaço amostral.  |
| $G(\mathbf{x})$                  | Função que define o evento falha por $G(\mathbf{x}) \leq 0$ .  |
| $I_{cond.}[\mathbf{x}]$          | Função indicadora, podendo ser igual a 1 ou 0 caso a condição seja atendida.   |
| $I[x]$                           | Caso particular da função indicadora para o evento falha, se $x \in \Omega_f$ então $I[x] = 1$ , senão $I[x] = 0$ .  |
| $\gamma$                         | Variograma, explicado na sessão 2.5.   |
| $\mathbf{X}, \mathbf{Y}$         | Amostra de ocorrências observadas para a variável aleatória $X$ e sua resposta em $Y$ .  |
| $x^{(i)}, y^{(i)}$               | Índice $i$ para identificação para os pares de variáveis $\mathbf{X}$ e $\mathbf{Y}$ .   |
| $\lambda_i$                      | Pesos usados na aproximação do método de Kriging, o índice $i$ indica seu par $y^{(i)}$ ou $y_i$ .   |
| $\nabla[\cdot]_{(\lambda, M_L)}$ | Operador de gradiente em relação às variáveis $\lambda$ e $M_L$ .  |
| $\Psi$                           | Matriz de correlação para os pontos de suporte do Kriging.   |
| $\psi(\mathbf{x})$               | Vetor de correlação para um ponto $y(\mathbf{x})$ aproximado pelo Kriging.   |
| $\mathbf{1}$                     | Vetor operacional composto apenas de valores unitários com dimensão $n \times 1$ . Necessário para operações vetoriais e matriciais em conjunto a valores escalares. |
| $\mathcal{L}$                    | Operador de estimativa de máxima verossimilhança ou <i>maximum likelihood estimation</i> (MLE).  |
| $\pi(\mathbf{x}_i)$              | Função pi introduzida por Picheny (2009), demonstrada na equação 3.1.  |
| $\Phi(x)$                        | Função de distribuição acumulada de uma variável normal padrão para um valor $x$ .   |
| $U(\mathbf{x})$                  | Função $U$ desenvolvida por Echard, Gayton e Lemaire (2010), descrita na equação 3.5.  |
| $\theta_m$                       | Parâmetro para a função de variograma mais utilizada para o Kriging ordinário.   |

# Sumário

|            |  |           |
|------------|--|-----------|
| <b>1</b>   | <b>INTRODUÇÃO</b>  | <b>15</b> |
| <b>1.1</b> | <b>Objetivos</b>   | <b>17</b> |
| 1.1.1      | Objetivos específicos  | 17        |
| <b>1.2</b> | <b>Organização</b>   | <b>18</b> |
| <b>2</b>   | <b>CONCEITOS NECESSÁRIOS</b>                                     | <b>19</b> |
| <b>2.1</b> | <b>Conceitos de probabilidade</b>                                | <b>19</b> |
| 2.1.1      | Distribuição de probabilidade                                    | 20        |
| 2.1.2      | Valor central e momentos de uma variável aleatória               | 21        |
| 2.1.3      | Probabilidades conjuntas entre variáveis aleatórias              | 23        |
| 2.1.4      | Distribuição normal  | 26        |
| <b>2.2</b> | <b>Probabilidade de falha</b>                                    | <b>26</b> |
| <b>2.3</b> | <b>Simulação de Monte Carlo</b>                                  | <b>28</b> |
| 2.3.1      | Histórico do método  | 28        |
| 2.3.2      | Formulação do SMC  | 29        |
| 2.3.3      | Observações sobre o método                                       | 30        |
| <b>2.4</b> | <b>Kriging</b>   | <b>31</b> |
| <b>2.5</b> | <b>Variograma</b>  | <b>32</b> |
| <b>2.6</b> | <b>Cálculo do Kriging</b>  | <b>33</b> |
| 2.6.1      | Exemplo de resolução do Kriging                                  | 37        |
| <b>3</b>   | <b>INTEGRAÇÃO KRIGING COM MONTE CARLO SIMPLES</b>                | <b>41</b> |
| <b>3.1</b> | <b>Método de AKMCS base - Kriging com aprendizado ativo</b>      | <b>41</b> |
| 3.1.1      | Características do método  | 45        |
| <b>3.2</b> | <b>AKMCS+BiL - adaptação para melhorar a curva limite</b>        | <b>47</b> |
| 3.2.1      | Efeitos da alteração   | 47        |
| <b>3.3</b> | <b>AKMCS+ewd - adaptação com função de aprendizado composta</b>  | <b>48</b> |
| 3.3.1      | Efeitos da alteração   | 49        |
| <b>3.4</b> | <b>Medida de precisão para o AKMCS</b>                           | <b>50</b> |
| 3.4.1      | Análise das medidas propostas                                    | 51        |
| 3.4.2      | Efeitos da alteração   | 53        |
| <b>3.5</b> | <b>DKM - método de Kriging dependente</b>                        | <b>54</b> |
| 3.5.1      | Efeitos da alteração   | 55        |
| <b>3.6</b> | <b>AK-MCSi - método AKMCS melhorado para aplicações práticas</b> | <b>56</b> |
| 3.6.1      | Efeitos da alteração   | 57        |

|            |  |           |
|------------|--|-----------|
| <b>4</b>   | <b>INTEGRAÇÃO DE KRIGING COM MONTE CARLO USANDO AMOSTRAGEM POR IMPORTÂNCIA . . . . .</b> | <b>61</b> |
| <b>4.1</b> | <b>Meta-IS - primeiro modelo de Kriging por importância . . . . .</b>                    | <b>61</b> |
| 4.1.1      | Características do método . . . . .  | 63        |
| <b>4.2</b> | <b>MetaAL-OIS - exemplo do progresso do método . . . . .</b>                             | <b>64</b> |
| 4.2.1      | Análise do trabalho . . . . .  | 65        |
| <b>5</b>   | <b>RESULTADOS . . . . .</b>  | <b>67</b> |
| <b>5.1</b> | <b>Problema com quatro curvas limite . . . . .</b>                                       | <b>68</b> |
| 5.1.1      | Resultados para o problema 1 . . . . .   | 68        |
| <b>5.2</b> | <b>Problema com oscilação - Rastrigin Function . . . . .</b>                             | <b>71</b> |
| 5.2.1      | Resultados para o problema 2 . . . . .   | 71        |
| <b>5.3</b> | <b>Problema com equação de terceiro grau . . . . .</b>                                   | <b>73</b> |
| 5.3.1      | Resultados para o problema 3 . . . . .   | 74        |
| <b>5.4</b> | <b>Problema com grande número de variáveis . . . . .</b>                                 | <b>75</b> |
| 5.4.1      | Resultados para o problema 4 . . . . .   | 75        |
| <b>6</b>   | <b>CONCLUSÃO . . . . .</b>   | <b>77</b> |
| <b>6.1</b> | <b>Trabalhos Futuros . . . . .</b>   | <b>78</b> |
|            | <b>REFERÊNCIAS . . . . .</b>   | <b>81</b> |



# 1 Introdução

Para um projeto de engenharia pode-se definir três fatores de extrema importância, o primeiro é a funcionalidade, o segundo fator é o custo e o terceiro a segurança. A funcionalidade descreve os requisitos do projeto, usualmente na forma de um sistema de equações definido pelas características envolvidas no projeto em si. A relevância do custo é óbvia em um cenário real, a competitividade de um projeto é definida não apenas pelos benefícios que ele gera, mas também pelo custo para obter tais benefícios. A segurança é definida pelo projeto em questão, em casos usuais de engenharia a segurança é definida por limites estruturais ou limites de conforto e segurança. Para este trabalho o termo segurança é usada de maneira ampla desde que possa ser definido por algum sistema de equações, este sistema usualmente é transformado em uma função das variáveis aleatórias do projeto chamada de função de desempenho.

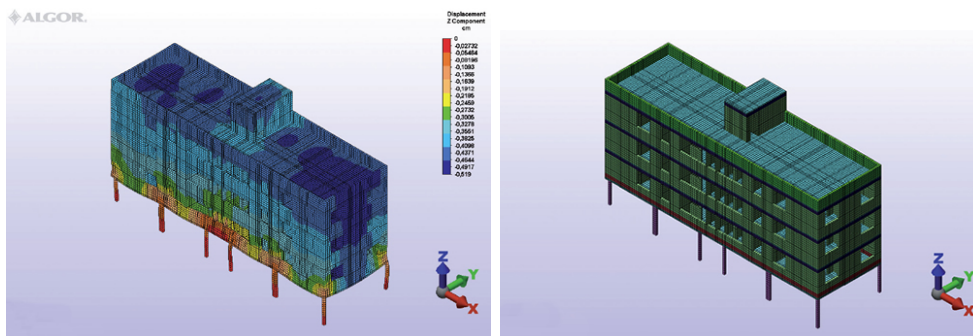
Para definir a segurança de um projeto, deve-se conhecer uma grande quantidade de informações: características dos materiais, esforços atuantes, efeitos de desgaste por exposição ao ambiente, temperaturas usuais, entre outros fatores. É impossível definir todas as características de um determinado projeto a priori, sendo possível separar essas indeterminações como incertezas intrínsecas ou epistêmicas. A incerteza intrínseca faz parte do processo envolvido, alguns exemplos são ventos, chuvas ou até variações em propriedades de um material a ser manufaturado. A incerteza epistêmica por outro lado está relacionada com o conhecimento quanto ao problema, sendo possível reduzir ou eliminar esta incerteza por meio de maior estudo sobre as informações específicas do projeto. Alguns exemplos de incerteza epistêmica são: a carga que um determinado tipo de caminhão terá ou o uso para uma determinada construção, nota-se que a redução ou eliminação desta incerteza exige um conhecimento específico do projeto em si.

Após definir as incertezas pode-se modelar elas por meio de variáveis aleatórias e, junto da descrição da segurança por meio de um sistema de equações, é possível determinar a probabilidade destes requisitos serem ou não violados, chamada de probabilidade de falha ( $P_f$ ). A probabilidade de falha pode ser calculada analiticamente para alguns casos simples, sendo necessário algum método mais avançado para outros casos. Um dos métodos mais usados é a simulação de Monte Carlo (SMC), sendo simuladas execuções das variáveis aleatórias repetidamente até que o valor de  $P_f$  encontrado torne-se estável e confiável. Pode-se afirmar que projetos usuais de engenharia civil devem ter de  $10^5$  realizações do modelo de projeto ou mais para obter um resultado adequado.

Para calcular a segurança de um projeto falta apenas definir o sistema de equações. Para casos práticos de engenharia utilizam-se geralmente modelos numéricos de extenso

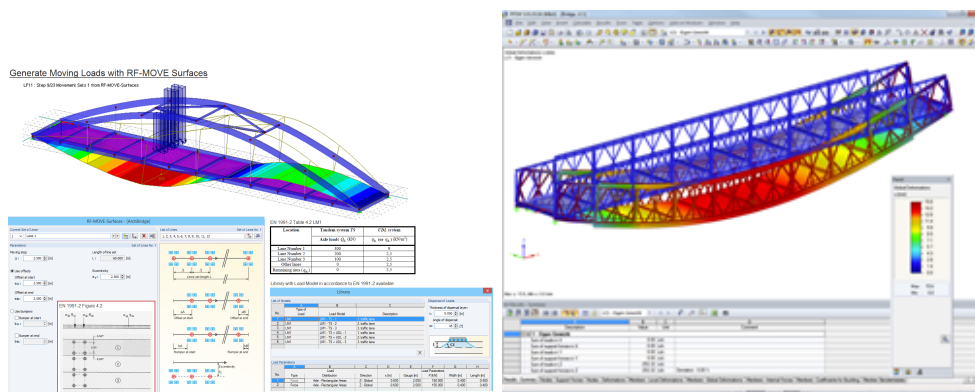
custo computacional. Exemplificando para a área de engenharia civil, um edifício deve ser verificado em cada uma de suas vigas, pilares e fundações, para casos de alvenaria estrutural deve-se considerar adicionalmente todas as paredes e vãos, verificando tanto o estado limite de serviço quanto o estado limite último. Com isso pode-se verificar conceitualmente os motivos para o alto custo computacional de uma estrutura, sendo possível verificar também na Figura 1 e Figura 2 que demonstram visualmente a análise de um projeto.

Figura 1 – Modelo estrutural de um edifício.



Fonte: <http://techne17.pini.com.br/engenharia-civil/172/artigo286838-3.aspx>

Figura 2 – Mais exemplos de modelos estruturais.



Fonte: [www.dlubal.com](http://www.dlubal.com)

Considerando o custo computacional para solucionar o sistemas de equações do modelo numérico de um projeto e a quantidade de vezes que é necessário realizar seu cálculo para obter um bom resultado por meio da simulação de Monte Carlo, pode-se concluir que o uso da SMC é inviável em função do custo envolvido e deve-se utilizar de soluções alternativas. Um método utilizado atualmente envolve o uso de um metamodelo para substituir a função de desempenho original do projeto, sendo demonstrado neste estudo a solução utilizando o metamodelo de Kriging.

O uso do metamodelo implica na aproximação da resposta da função de desempenho com um menor custo computacional. Uma vez que a aproximação será usada juntamente com a simulação de Monte Carlo, basta determinar a quantidade de realizações do modelo numérico que resultam na violação da função de desempenho. Para simplificar, todas as

funções de desempenho utilizadas neste trabalho serão violadas quando seu valor numérico for inferior ou igual a zero ( $G(\mathbf{x}) \leq 0$ ).

Considerando a importância discutida anteriormente de análises de confiabilidade e os elevados custos computacionais envolvidos, assim como da redução de custo por meio de um metamodelo para evitar usar as custosas funções de desempenho, pode-se verificar com facilidade a relevância deste trabalho. Por meio das descrições e análises demonstradas pode-se compreender o funcionamento do Kriging e dos procedimentos de Kriging com Monte Carlo. Este trabalho demonstra as informações necessárias para que um profissional decida qual o método mais adequado para os problemas em questão, assim como descreve em detalhes para identificar as limitações dos métodos descritos. Enquanto para os interesses acadêmicos pode-se ressaltar as classificações definidas e a inclusão de pesquisas recentes e relevantes para o futuro do uso do Kriging com Monte Carlo.

## 1.1 Objetivos

O objetivo principal é demonstrar os avanços obtidos para o método de Kriging com Monte Carlo desde seu desenvolvimento inicial até a atualidade. Além de analisar as vantagens e desvantagens das alterações realizadas neste período disponíveis na literatura.

### 1.1.1 Objetivos específicos

Para alcançar o objetivo principal, os seguintes objetivos foram definidos:

- Descrever de maneira clara o método de Kriging e comentar as fontes para aprofundamento, pois ainda existem poucos trabalhos na literatura que descrevem o método extensivamente.
- Explicar detalhadamente o método AKMCS original buscando disponibilizar ferramentas para comparações e análises, possibilitando então a compreensão dos trabalhos abordados no escopo desta dissertação, assim como para os demais estudos da área.
- Realizar uma seleção de artigos e trabalhos relevantes ao assunto que permitam descrever o estado da arte atual, assim como possibilite compreender os avanços gerados a partir da popularização do método.
- Expor as diferenças nas características do método após as alterações desenvolvidas pelos autores descritos.
- Auxiliar no desenvolvimento de futuras alterações por meio das análises disponíveis.

## 1.2 Organização

O trabalho é iniciado explicando os conceitos necessários para a compreensão das análises realizadas, seguindo para a análise de alguns dos métodos. A análise foi separada em três tipos de combinações de Kriging com Monte Carlo: a integração Kriging com Monte Carlo simples e com amostragem por importância. Após a apresentação dos modelos da literatura, mostram-se alguns resultados para reforçar as análises feitas. Finalizando então com as conclusões e sugestões para trabalhos futuros.

## 2 Conceitos necessários

Para facilitar a compreensão do estudo realizado serão revisados os conceitos dos assuntos utilizados. Também são apresentadas as nomenclaturas utilizadas e a simbologia para os equacionamentos. Os símbolos usados pelos autores analisados serão alterados e unificados na busca de simplificar os equacionamentos sempre que possível, gerando maior clareza e facilitando as análises e comparações.

### 2.1 Conceitos de probabilidade

De acordo com Montgomery, Runger, Hubele (2003) e Silva (1999), a medida de uma determinada variável é chamada de experimento. Frequentemente existirão fatores que influenciam esta medida e que não são possíveis controlar, sendo estes fatores chamados de componentes aleatórios. Pode-se chamar de variável aleatória o resultado de um experimento não previsível, sendo divididas em variáveis aleatórias discretas e contínuas (REDDY, 2011).

Segundo Degroot e Schervish (2012) uma variável aleatória é uma função com valores reais definida em um espaço amostral  $\Omega$ . A partir do espaço amostral pode-se denominar um subconjunto Evento, que se refere a um resultado cuja veracidade só pode ser determinada após a realização do experimento (CLARKE; DISNEY, 1979). Para facilitar a compreensão será determinado um evento  $A$ , sua probabilidade de ocorrência é representada por  $P[A]$ , sendo esta medida sujeita aos seguintes axiomas (PAPOULIS; PILLAI, 2002):

1.  $P[A] \geq 0$ .  
Nenhum evento pode ter probabilidade negativa.
2.  $P[\Omega] = 1$ .  
O espaço amostral é certo, ou seja, a probabilidade é 100%.
3.  $P[A \cup B] = P[A] + P[B]$ ,  
se  $A$  e  $B$  forem eventos mutuamente exclusivos ( $P[A \cap B] = 0$ ).

De acordo com Beck (2014), para um evento qualquer  $A$  pode-se definir sua probabilidade de duas formas, a definição clássica de probabilidade define que se o evento  $A$  possui  $N_A$  eventos favoráveis em um total de  $N$  eventos possíveis, pode-se calcular a probabilidade de ocorrência do evento  $A$  por:

$$P[A] = \frac{N_A}{N}. \quad (2.1)$$

Para a definição frequentista a probabilidade de um evento é calculada à posteriori, baseando-se em um grande número de observações do experimento. Caso o evento  $A$  seja observado  $n_A$  vezes em  $n$  realizações do experimento, a probabilidade de ocorrência de  $A$  é dada por:

$$P[A] = \lim_{n \rightarrow \infty} \frac{n_A}{n}. \quad (2.2)$$

No entanto, esta definição é limitada por não ser possível realizar infinitos experimentos.

### 2.1.1 Distribuição de probabilidade

A partir da definição da probabilidade de um evento  $A$  é possível definir a função de distribuição acumulada ou *cumulative distribution function (CDF)* (KRISHNAIAH; RAO, 1988), usando  $A = \{x \in \Omega / X \leq x\}$  onde  $X$  é a variável aleatória em questão, resultando em:

$$F_X(x) = P[A] = P[\{X \leq x\}], \quad (2.3)$$

onde  $F_X(x)$  é a CDF da variável  $X$ . Em outras palavras, o número  $F_X(x)$  corresponde à probabilidade da variável  $X$  assumir qualquer valor menor que  $x$ .

A função de distribuição acumulada de probabilidade possui as seguintes propriedades (ANG; TANG, 2007):

1.  $F_X(-\infty) = 0$  e  $F_X(\infty) = 1$ .

É impossível o valor de  $X$  ser menor que  $-\infty$ , além de ele sempre ser menor que  $\infty$ .

2.  $F_X(\cdot)$  é monotonicamente crescente, ou seja,  $F_X(x_1) \leq F_X(x_2) \forall x_1 < x_2$ .

Em outras palavras  $P[\{X \leq x_1\}] \leq P[\{X \leq x_2\}]$ ,

pois  $P[\{X \leq x_2\}] = P[\{X \leq x_1\}] + P[\{x_1 \leq X < x_2\}]$ .

3.  $F_X(\cdot)$  é contínua pela direita:

$$\lim_{\Delta_X \rightarrow 0^+} F_X(x + \Delta_X) = F_X(x+) = F_X(x). \quad (2.4)$$

Também pode-se escrever como  $F_X(x+) = P[\{X \leq x+\}] = P[\{X \leq x\}]$ ,

pois  $X \leq x+ \rightarrow X < x+$  e  $X < x+ \rightarrow X \leq x$ .

4. Para  $x_1 < x_2$  temos:

$$P[\{x_1 < X \leq x_2\}] = F_X(x_2) - F_X(x_1). \quad (2.5)$$

Porque  $P[\{x_1 < X \leq x_2\}] = P[\{X \leq x_2\}] - P[\{X \leq x_1\}]$ .

5. Se  $F_X(x)$  é descontínua em  $x_1$ , então:

$$P[\{X = x_1\}] = F_X(x_1) - F_X(x_1-), \quad (2.6)$$

que é igual ao salto de  $F_X(x)$  em  $x_1$ .

Pois  $X \leq x- \rightarrow X < x$ , portanto  $P[\{X \leq x_1\}] - P[\{X < x_1\}] = P[\{X = x_1\}]$ .

6. Se  $F_X(x)$  é contínua, não possui saltos, e portanto:

$$P[\{X = x\}] = 0 \quad \forall x \in \Omega. \quad (2.7)$$

Isso pode ser confirmado a partir da 5ª propriedade.

Com a CDF é possível descrever uma variável aleatória e identificá-la em variável discreta, contínua ou mista. Se  $F_X(x)$  for contínua em todo o domínio então a variável é contínua, caso possua incrementos de  $F_X$  apenas nas descontinuidades ela é discreta, caso contrário é uma variável aleatória mista.

Outra função de grande importância é aquela que representa a probabilidade de ocorrência de um determinado ponto  $x$  ( $P[\{X = x\}]$ ), sendo esta chamada de função de massa de probabilidade ou *probability mass function (PMF)* (ROSS, 2014). Pode-se chegar em  $p_X(x) = P[\{X = x\}]$  utilizando a 5ª propriedade mostrada, onde  $p_X$  é a PMF.

A PMF existe apenas quando há descontinuidade em  $F_X$ , pois para variáveis contínuas  $P[\{X = x\}] = 0$ . Pode-se definir então uma função de densidade de probabilidade ou *probability density function (PDF)* como sendo a derivada em  $x$  da CDF:

$$f_X(x) = \frac{dF_X(x)}{dx}, \quad (2.8)$$

sendo  $f_X(x)$  a PDF para um ponto  $x$  qualquer. As propriedades da PDF podem ser verificadas em Ushakov (1999), juntamente com outras propriedades para as funções de distribuição anteriores.

### 2.1.2 Valor central e momentos de uma variável aleatória

Muitas vezes as funções de probabilidade de uma variável não é especificada, sendo usados outras medidas para descreve-la. Os valores mais usados para descrever de forma aproximada uma variável aleatória sem utilizar as funções de probabilidade são seus valores centrais e momentos (ANG; TANG, 2007).

O valor central mais utilizado para uma variável aleatória é o valor esperado, sendo ele uma média ponderada que leva em conta as diferentes probabilidades associadas a cada valor  $x \in \Omega$ . Esta média ponderada define o valor esperado ( $E[X]$ ) ou valor médio ( $\mu$ ) da variável aleatória. Define-se este valor por:

$$E[X] = \sum_{x_i} x_i p_X(x_i) \quad \forall x_i \in \Omega \quad (2.9)$$

para variáveis discretas ou:

$$E[X] = \int_{-\infty}^{+\infty} x f_X(x) dx \quad (2.10)$$

para variáveis contínuas. Observando  $E[X]$  é possível generalizá-lo como um operador linear  $E[\cdot]$ , sendo verificada sua linearidade pela formulação apresentada.

Existem outros valores centrais que não serão utilizados neste estudo como valor modal e mediana.

Além do valor central, utiliza-se o momento de segunda ordem para indicar a dispersão de uma variável, sendo a variância o mais usado. Pode-se calcular a variância por meio da equação:

$$Var[X] = \sum_{x_i} (x_i - \mu_X)^2 p_X(x_i) \quad \forall x_i \in \Omega \quad (2.11)$$

para variáveis discretas ou:

$$Var[X] = \int_{-\infty}^{+\infty} (x - \mu_X)^2 f_X(x) dx \quad (2.12)$$

para contínuas.

Para estimar o valor central ou variância usando uma amostra de tamanho  $n$  pode-se calcular por:

$$\begin{aligned} \hat{\mu}_X &= \frac{1}{n} \sum_{i=1}^n x_i \\ \hat{Var}[X] &= \frac{1}{n-1} \sum_{i=1}^n (x_i - \hat{\mu}_X)^2, \end{aligned} \quad (2.13)$$

sendo usado no denominador da variância  $(n-1)$  para gerar um resultado não tendencioso em relação ao tamanho da amostra (DEVORE, 2011).

Pode-se expandir a Eq.2.12 usando o operador linear  $E[\cdot]$  apresentado nas equações 2.9 e 2.10, resultando em:

$$\begin{aligned} Var[X] &= E[(X - \mu_X)^2] = E[(X - E[X])^2] \\ Var[X] &= E[X^2 - 2XE[X] + E[X]^2] = E[X^2] - 2E[XE[X]] + E[E[X]^2] \\ Var[X] &= E[X^2] - 2E[X]E[X] + E[X]^2 = E[X^2] - E[X]^2 \\ Var[X] &= E[X^2] - \mu_X^2. \end{aligned} \quad (2.14)$$

De acordo com Urbano (2010), a medida mais usada para indicar a dispersão de uma variável é a raiz quadrada da variância, chamada de desvio-padrão ( $\sigma_X$ ). Por este motivo, a variância também pode ser escrita como o quadrado do desvio-padrão, sendo muitas vezes denotada como  $\sigma^2$  ao invés de  $Var$ .

$$\sigma_X = \sqrt{Var[X]} = \sqrt{E[X^2] - \mu_X^2}. \quad (2.15)$$

Outro meio conveniente para indicação da dispersão é o coeficiente de variação ( $\delta_X$ ), sendo descrito pela equação:

$$\delta_X = \frac{\sigma_X}{\mu_X}, \quad (2.16)$$

apesar da média  $\mu_X$  poder ser negativa, o resultado de  $\delta_X$  é considerado positivo por convenção (NOWAK; COLLINS, 2000). A vantagem de utilizar o coeficiente de variação é



reduzir o impacto da escala associada a variável, entretanto não é possível definir um  $\delta_X$  quando a média é nula.

O coeficiente de variação também pode ser encontrado na literatura de outras formas como *c.v.* (BUSSAB; MORETTIN, 2012),  $V_X$  (NOWAK; COLLINS, 2000),  $CV_X$  (SUDRET, 2007) e até mesmo como *relative standard deviation* ou RSD (MOHAMMADI; EGHTEHAD; MOHAMMADI, 2018). Para este trabalho será utilizado o símbolo  $\delta_X$  apresentado na equação 2.16.

### 2.1.3 Probabilidades conjuntas entre variáveis aleatórias

Para qualquer par de variáveis aleatórias  $X$  e  $Y$ , existem probabilidades associadas para as ocorrências dos valores  $x$  e  $y$ . Pode-se definir a função de distribuição acumulada conjunta ou *joint distribution function* (CDF conjunta) (KINGSBURY, 2005) a partir da equação 2.3 para um caso de duas variáveis por:

$$F_{X,Y}(x, y) = P[\{X \leq x, Y \leq y\}]. \quad (2.17)$$

Similarmente, pode-se definir uma PMF conjunta e uma PDF conjunta por:

$$\begin{aligned} p_{X,Y}(x, y) &= P[\{X = x, Y = y\}] \\ f_{X,Y}(x, y) &= \frac{\partial^2 F_{X,Y}(x, y)}{\partial x \partial y}. \end{aligned} \quad (2.18)$$

Quando uma variável possui alguma relação com outra, pode-se dizer que existe uma distribuição condicional. Segundo Joyce (2014), para um par discreto de variáveis  $X$  e  $Y$  pode-se escrever:

$$p_{X|Y}(x|y) = P[\{X = x|Y = y\}] = \frac{p_{X,Y}(x, y)}{p_Y(y)} \quad (2.19)$$

se  $p_Y(y) \neq 0$  onde  $p_{X|Y}(x|y)$  indica a probabilidade de massa condicional de  $X$  (PMF condicional), similarmente:

$$p_{Y|X}(y|x) = P[\{Y = y|X = x\}] = \frac{p_{X,Y}(x, y)}{p_X(x)} \quad (2.20)$$

quando  $p_X(x) \neq 0$ . O símbolo  $|$  indica a condição especificada, por exemplo  $P[\{Y = y|X = x\}]$  pode ser lido como a probabilidade de  $Y = y$  caso  $X = x$  ocorra. Semelhante ao descrito para a PMF condicional, pode-se escrever a PDF condicional:

$$\begin{aligned} f_{X|Y}(x|y) &= \lim_{\Delta_X \rightarrow 0} \lim_{\Delta_Y \rightarrow 0} P[\{x \leq X \leq x + \Delta_X | y \leq Y \leq y + \Delta_Y\}] \\ f_{X|Y}(x|y) &= \frac{f_{X,Y}(x, y)}{f_Y(y)}. \end{aligned} \quad (2.21)$$

Pode-se ainda equacionar uma distribuição condicional de  $X$ , dado que  $Y = y$  por:

$$F_{X|Y}(x|y) = P[\{X \leq x|Y = y\}], \quad (2.22)$$

sendo  $F_{X|Y}(x|y)$  o CDF condicional de  $X$  em relação a  $Y$ .

A informação da dependência entre variáveis aleatórias pode ser reestruturada isolando as probabilidades conjuntas para se assemelhar as equações 2.18 da seguinte forma:

$$\begin{aligned} p_{X,Y}(x, y) &= P[\{X = x\}]P[\{Y = y|X = x\}] = P[\{Y = y\}]P[\{X = x|Y = y\}] \\ f_{X,Y}(x, y) &= f_X(x)f_{Y|X}(y|x). \end{aligned} \quad (2.23)$$

Com isso pode-se desenvolver as propriedades para as distribuições conjuntas de maneira simplificada, por exemplo:

$$\begin{aligned} \int_{-\infty}^{+\infty} f_{X,Y}(x, y)dy &= \int_{-\infty}^{+\infty} f_X(x)f_{Y|X}(y|x)dy \\ \int_{-\infty}^{+\infty} f_{X,Y}(x, y)dy &= F_X(x) \int_{-\infty}^{+\infty} f_{Y|X}(y|x)dy \\ \int_{-\infty}^{+\infty} f_{X,Y}(x, y)dy &= f_X(x). \end{aligned} \quad (2.24)$$

Quando não existe dependência entre as variáveis  $X$  e  $Y$  as equações anteriores podem ser reescritas por:

$$\begin{aligned} F_{X,Y}(x, y) &= P[\{X \leq x, Y \leq y\}] = P[\{X \leq x\}]P[\{Y \leq y\}] = F_X(x)F_Y(y) \\ p_{X,Y}(x, y) &= P[\{X = x\}]P[\{Y = y\}] \\ f_{X,Y}(x, y) &= f_X(x)f_Y(y). \end{aligned} \quad (2.25)$$

A dependência ou independência entre variáveis pode ser importante, como observado na equação 2.25, portanto é necessário desenvolver formas de medir a relação entre variáveis aleatórias. Para medir a relação linear entre duas variáveis aleatórias é possível utilizar a covariância (GATTI, 2005):

$$Cov[X, Y] = E[(X - \mu_X)(Y - \mu_Y)]. \quad (2.26)$$

Desenvolvendo o equacionamento do operador  $E[\cdot]$ :

$$\begin{aligned}
Cov[X, Y] &= E[(X - \mu_X)(Y - \mu_Y)] \\
Cov[X, Y] &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x - \mu_X)(y - \mu_Y) f_{X,Y}(x, y) dx dy \\
Cov[X, Y] &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (xy - x\mu_Y - y\mu_X + \mu_X\mu_Y) f_{X,Y}(x, y) dx dy \\
Cov[X, Y] &= \begin{cases} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} xy f_{X,Y}(x, y) dx dy \\ -\mu_Y \int_{-\infty}^{+\infty} x \int_{-\infty}^{+\infty} f_{X,Y}(x, y) dy dx \\ -\mu_X \int_{-\infty}^{+\infty} y \int_{-\infty}^{+\infty} f_{X,Y}(x, y) dx dy \\ +\mu_X\mu_Y \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f_{X,Y}(x, y) dx dy \end{cases} \quad (2.27) \\
Cov[X, Y] &= \begin{cases} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} xy f_{X,Y}(x, y) dx dy \\ -\mu_Y \int_{-\infty}^{+\infty} x f_X(x) dx \\ -\mu_X \int_{-\infty}^{+\infty} y f_Y(y) dy \\ +\mu_X\mu_Y \end{cases} \\
Cov[X, Y] &= E[XY] - \mu_Y\mu_X - \mu_X\mu_Y + \mu_X\mu_Y \\
Cov[X, Y] &= E[XY] - \mu_X\mu_Y.
\end{aligned}$$

Além disso, a covariância de uma variável com ela mesma resulta em:

$$\begin{aligned}
Cov[X, X] &= E[X^2] - \mu_X\mu_X \\
Cov[X, X] &= E[X^2] - \mu_X^2 \\
Cov[X, X] &= Var[X].
\end{aligned} \quad (2.28)$$

Considerando a equação 2.26, é possível obter um valor qualquer dentre os números reais, tornando difícil analisar o resultado da covariância. Para usar uma medida adimensional de covariância usa-se o coeficiente de correlação:

$$Cor[X, Y] = \frac{Cov[X, Y]}{\sigma_X\sigma_Y}, \quad (2.29)$$

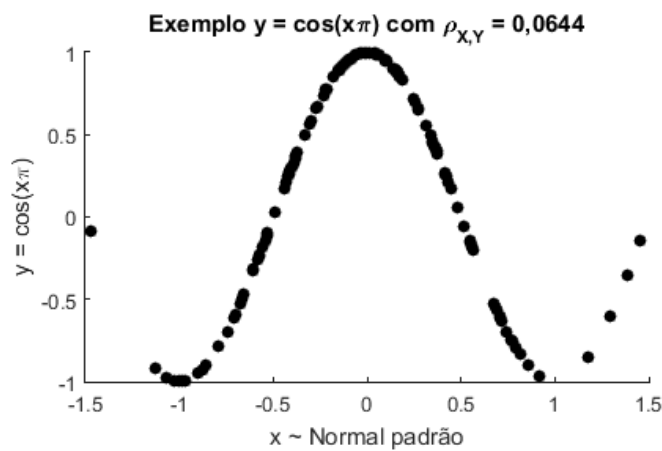
sendo  $\sigma$  o desvio-padrão. Supondo uma relação linear entre as variáveis da forma  $y = nx$ , resulta em:

$$\begin{aligned}
Cor[X, Y] &= \frac{Cov[X, Y]}{\sigma_X\sigma_Y} \\
Cor[X, Y] &= \frac{E[XnX] - \mu_Xn\mu_X}{\sigma_X|n|\sigma_X} \\
Cor[X, Y] &= \frac{nE[X^2] - n\mu_X^2}{|n|\sigma_X^2} \\
Cor[X, Y] &= \frac{n}{|n|} \frac{E[X^2] - \mu_X^2}{\sigma_X^2} \\
Cor[X, Y] &= \frac{n}{|n|} \frac{\sigma_X^2}{\sigma_X^2} \\
Cor[X, Y] &= \frac{n}{|n|} 1.
\end{aligned} \quad (2.30)$$

Sendo assim, caso  $Y$  tenha uma relação linear com  $X$  o resultado será próximo de  $-1$  ou  $1$ , enquanto valores próximos de  $0$  indicam independência linear, podendo existir dependência não-lineares. A demonstração do limite superior e inferior da correlação pode ser encontrada em Larson (1982).

Importante ressaltar que a covariância, por indicar relação linear, pode não identificar outras relações de dependência. A figura 3 demonstra  $Y = \cos(X\pi)$ , sendo  $X$  uma variável normal padrão, apesar de serem variáveis dependentes, o coeficiente de correlação resulta em  $0,06$  aproximadamente.

Figura 3 – Covariância próxima de zero para variáveis dependentes.



#### 2.1.4 Distribuição normal

A distribuição normal, também chamada de distribuição Gaussiana, é uma das distribuições mais utilizadas na literatura atualmente (PINHEIRO et al., 2012). Sua PDF é dada por:

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu)^2/2\sigma^2}, \quad (2.31)$$

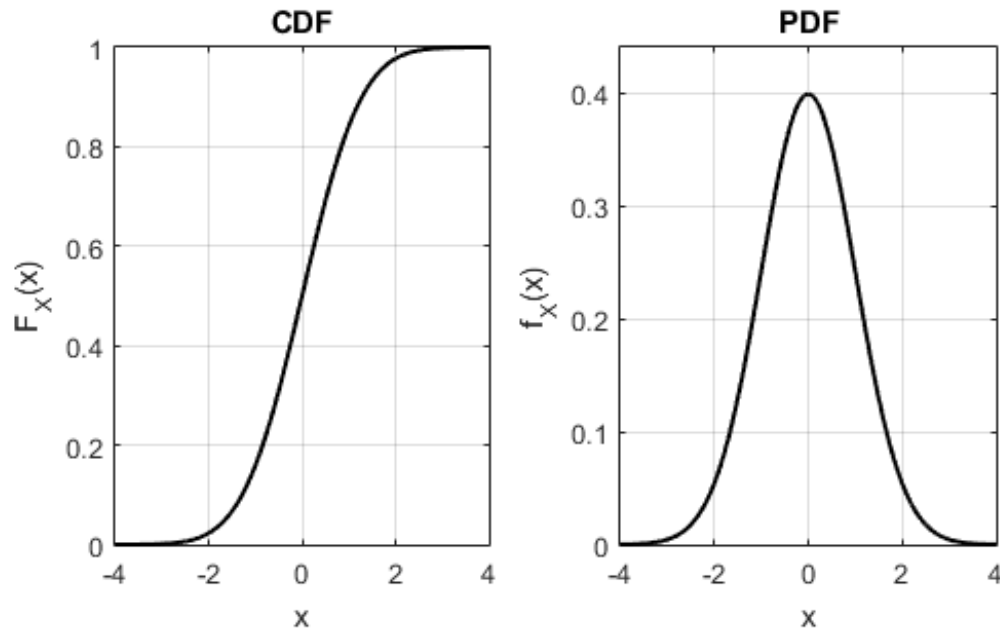
sendo  $\mu$  e  $\sigma$  os parâmetros que definem a média e desvio-padrão da distribuição em questão. Na figura 4 pode-se observar os gráficos da CDF e da PDF respectivamente.

Quando  $\mu = 0$  e  $\sigma = 1$  temos um caso particular de distribuição chamado de normal padrão. Apesar de sua importância, as características particulares desta distribuição não serão abordadas neste estudo, para mais detalhes pode-se consultar os trabalhos citados anteriormente.

## 2.2 Probabilidade de falha

A falha ocorre quando um sistema ou componente não desempenha sua função de acordo com o estipulado (BIROLINI, 2017), logo a probabilidade de falha de um sistema

Figura 4 – Funções de distribuição de uma variável normal padrão



é definida pela a probabilidade de não desempenhar sua função para dadas condições em um intervalo de tempo (EBELING, 1997).

Sendo possível determinar um evento falha ( $E_f$ ) em função das variáveis ( $\mathbf{X}$ ), é possível descrever a probabilidade de falha ( $P_f$ ) por:

$$P_f = P[E_f] = P[\{\mathbf{X} \in E_f\}], \quad (2.32)$$

observando que  $\mathbf{X}$  pode ser uma ou mais variáveis (KAPUR; LAMBERSON, 1977).

Usualmente define-se o evento de falha pelos limites de falha, ou seja as equações que separam a região de falha do resto do domínio  $\Omega$  por

$$E_f = \{G(\mathbf{x}) \leq 0 / \forall \mathbf{x} \in \Omega\}, \quad (2.33)$$

onde  $G(\mathbf{x})$  é a função que define o limite de falha. Existe a possibilidade da região de falha ser definida por mais de uma função, quando isto ocorre basta descreve-la por:

$$E_f = \{G_j(\mathbf{x}) \leq 0 / \forall \mathbf{x} \in \Omega, j \in \{1; \dots, m\}\}, \quad (2.34)$$

sendo  $m$  o número de equações necessárias para delimitar a região de falha. Pode-se também denominar de domínio de falha ( $\Omega_f$ ) o conjunto determinado pelo evento falha.

Segundo Beck (2014), pode-se calcular a probabilidade de falha por meio da integração:

$$P_f = \int_{\Omega_f} f_{\mathbf{X}}(\mathbf{x}) d\mathbf{x} \quad (2.35)$$

ou ainda pela integração:

$$P_f = \int_{\Omega} I[\mathbf{x}] f_{\mathbf{X}}(\mathbf{x}) d\mathbf{x}, \quad (2.36)$$

onde  $I[\mathbf{x}] = 1$  se  $\mathbf{x} \in \Omega_f$  ou  $I[\mathbf{x}] = 0$  caso  $\mathbf{x} \notin \Omega_f$ .

## 2.3 Simulação de Monte Carlo

Para solucionar problemas de confiabilidade simples é possível utilizar os métodos analíticos, entretanto diversas complicações podem existir para tornar a solução analítica inviável. Portanto são utilizados métodos de calcular aproximadamente a probabilidade de falha. O método usado neste trabalho será a simulação de Monte Carlo (*Monte Carlo Simulation*).

O SMC é um método de amostragem, cuja vantagem é a utilização direta de experimentos para obter soluções matemáticas ou informações probabilísticas onde o sistema de equações que definem os limites de falha não pode ser facilmente solucionado (CHOI; GRANDHI; CANFIELD, 2007).

### 2.3.1 Histórico do método

A simulação de Monte Carlo foi desenvolvida em maior parte por Stanislaw Ulam por volta de 1946. Com a criação do primeiro computador eletrônico (ENIAC) e para solucionar problemas termonucleares uma reunião foi feita em 1946 onde Enrico Fermi, John von Neumann, Stan Ulam, Nicholas Metropolis e outros estavam presentes. Foi nesta reunião que iniciou a criação do método de Monte Carlo (METROPOLIS, 1987).

O nome do método foi reforçado com uma carta de John von Neumann em 1947 para Robert Richtmyer “Stan Ulam, John von Neumann, and the Monte Carlo Method”, sendo o nome dado por Nicholas Metropolis em homenagem ao tio de Stan que pegava empréstimos com parentes porque precisava ir apostar em Monte Carlo (METROPOLIS, 1987).

Apesar da carta de Neumann, o método muitas vezes é creditado apenas ao Nicholas Metropolis e Stanislaw Ulam em função da publicação (METROPOLIS; ULAM, 1949). Além disso o método também foi usado anterior a data publicada por Enrico Fermi desde 1933.

Apesar das fundações teóricas do método serem conhecidas há muito tempo, a simulação de variáveis aleatórias manualmente é um processo exaustivo (SOBOL, 1994). Um dos problemas que Fermi resolveu utilizando o Monte Carlo resultou no ganho de um premio Nobel em 1938 e a criação do Los Alamos em 1943.

Foi a criação de Los Alamos que possibilitou a reunião dos pesquisadores citados anteriormente, entretanto Fermi não havia publicado seus métodos (ANDERSON, 1986). Durante os cálculos do ENIAC, Fermi desenvolveu sua máquina de Monte Carlo analógica, o FERMIAC, que está em exibição no museu de ciência de Bradbury em Los Alamos, Novo México.

### 2.3.2 Formulação do SMC

A simulação de Monte Carlo consiste na solução da integração descrita na equação 2.36, entretanto ao invés de solucionar a integração numericamente, utilizam-se pontos aleatórios no domínio de maneira a economizar tempo computacional (LEMAIRE, 2009). A técnica usada para solucionar integrações por Monte Carlo baseia-se no fato da teoria de probabilidade clássica ter embasamento matemático para o comportamento de médias e tendência de estabilizar seu valor com o aumento do tamanho da amostra (DITLEVSEN; MADSEN, 2007).

Para determinar o valor de  $P_f$  pode-se trabalhar a partir da equação 2.36 repetida aqui:

$$P_f = \int_{\Omega} I[\mathbf{x}]f_{\mathbf{x}}(\mathbf{x})dx \quad (2.36 \text{ Repetida})$$

$$P_f = E(I[\mathbf{x}]), \quad (2.37)$$

pois como definido na equação 2.10, a expressão indica a média do valor  $I[\mathbf{x}]$  em todo o domínio  $\Omega$  (SUDRET, 2007). Logo é possível descrever a probabilidade de falha com base numa amostra de tamanho finito por:

$$E[I[\mathbf{x}]] \approx \frac{1}{n} \sum_{i=1}^n I[\mathbf{x}_i] = \frac{n_f}{n} = \hat{P}_f, \quad (2.38)$$

onde  $\hat{P}_f$  indica uma estimativa de  $P_f$ ,  $n$  é o número total de pontos amostrais e  $n_f$  é o número de pontos amostrais dentro do domínio de falha.

Utilizando os conceitos da teoria clássica de probabilidade por meio da equação 2.2, pode-se verificar que:

$$P_f = \lim_{n \rightarrow \infty} \frac{n_f}{n} = \lim_{n \rightarrow \infty} \hat{P}_f. \quad (2.39)$$

Considerando que é impossível utilizar todos os pontos em um domínio contínuo, o resultado de  $\hat{P}_f$  não apenas é aproximado, como também possui uma variação inerente à simulação. De acordo com Lemaire (2009), pode-se calcular a variância da estimativa por:

$$Var [\hat{P}_f] = Var \left[ \frac{1}{n} \sum_{i=1}^n I[\mathbf{x}_i] \right] = \frac{1}{n^2} \sum_{i=1}^n Var [I[\mathbf{x}_i]], \quad (2.40)$$

pois  $I[\mathbf{x}_i]$  é determinado pelo resultado aleatoriamente independente das realizações dos limites de falha. Pode-se usar a equação 2.14 para chegar em:

$$Var [\hat{P}_f] = \frac{1}{n^2} n Var [I[\mathbf{x}_i]] = \frac{1}{n^2} (E [I[\mathbf{x}_i]^2] - E [I[\mathbf{x}_i]]^2), \quad (2.41)$$

pois cada realização  $i$  possui as mesmas características de distribuição.

Pode-se mostrar que a equação 2.41 estima a variância de maneira tendenciosa, e que para obter uma estimativa não tendenciosa basta multiplicar o resultado por  $n/(n-1)$

(PAPOULIS, 1991). Chega-se então na equação:

$$\widehat{Var} [\hat{P}_f] \approx \frac{1}{n-1} \left[ \frac{1}{n} \sum_{i=1}^n I[\mathbf{x}_i] - \left( \frac{1}{n} \sum_{i=1}^n I[\mathbf{x}_i] \right)^2 \right]. \quad (2.42)$$

Após estimar a variância, é possível determinar um intervalo de confiança para o resultado encontrado, sendo essencial para qualquer estimador providenciar uma avaliação da precisão do valor estimado. O intervalo de confiança define uma forma precisa de descrever a incerteza do SMC (KROESE; TAIMRE; BOTEV, 2011).

Apesar de termos o valor da variância e da média, é necessário determinar a distribuição que a estimativa de falha possui. Considerando que a estimativa da probabilidade de falha é uma média da soma de várias realizações de  $I[\mathbf{x}]$ , e como  $I[\mathbf{x}]$  pode ser considerada uma variável aleatória discreta com dois resultados possíveis (0 ou 1), então pode-se afirmar que  $\hat{P}_f$  será uma distribuição normal quando  $n \rightarrow \infty$  (MELCHERS; BECK, 2018).

Sendo assim é possível determinar o intervalo de confiança por:

$$P[\hat{P}_f - k\hat{\sigma} < P_f < \hat{P}_f + k\hat{\sigma}] = C, \quad (2.43)$$

onde  $\hat{\sigma}$  é  $\sqrt{\widehat{Var}}$ ,  $P[\cdot]$  indica a probabilidade de ocorrência de  $[\cdot]$  e  $k$  é uma constante que define ou é definida pelo valor de confiança  $C$ . Um intervalo com 95% de confiança resulta em  $k = 1,96$ . Pode-se verificar este resultado, assim como utilizar outros valores, a partir das tabelas de distribuição para variáveis normais padrão.

Utilizando um intervalo proporcional a probabilidade de falha pode-se utilizar o coeficiente de variância  $\delta$  da equação 2.16:

$$\delta_{P_f} \equiv \frac{\sqrt{\widehat{Var} [\hat{P}_f]}}{E [\hat{P}_f]} \approx \frac{1}{\sqrt{nP_f}}, \quad (2.44)$$

sendo possível definir a precisão da estimativa com base na probabilidade de falha esperada e número de pontos amostrais utilizados. Para um  $P_f$  com uma ordem de  $10^{-p}$ , obter um  $\delta_{P_f} \leq 10\%$  requer aproximadamente  $10^{(p+2)}$  (BECK, 2014).

### 2.3.3 Observações sobre o método

De acordo com Melchers e Beck (2018), apenas é vantajoso utilizar o SMC quando o tamanho da amostra é inferior ao número de pontos de integração para solucionar uma integração numérica. Também deve-se levar em conta a complexidade do cálculo do limite de falha por meio das equações  $G_j(\mathbf{x})$  e considerar o número de variáveis aleatórias envolvidas na simulação.

Segundo Sudret (2007) o método descrito pode, teoricamente, ser aplicado em qualquer problema independente da complexidade. Entretanto o SMC é inviável quando



o custo computacional para cada execução individual não seja desprezível e exigir uma quantidade de execuções elevada.

Para problemas de confiabilidade estrutural típicos a probabilidade de falha possui uma ordem de  $10^{-3}$  a  $10^{-6}$ , portanto para ter um coeficiente de variação adequado é necessário de  $10^5$  a  $10^8$  (BECK, 2014).

Para aumentar a quantidade de problemas que a simulação de Monte Carlo pode solucionar, variações do método foram desenvolvidas, seja para reduzir o número de pontos necessários ou o custo computacional de cada ponto amostral. Por existirem outras formas de Monte Carlo, o procedimento descrito anteriormente foi chamado de Monte Carlo simples, direto, bruto ou cru (*crude Monte Carlo* em inglês).

Neste trabalho demonstra-se a redução do custo computacional para o cálculo de cada ponto amostral por meio de metamodelagem. Com o uso de um metamodelo é possível simplificar o cálculo do limite de falha transformando-o um modelo aproximado de menor custo computacional. Com esta alteração é possível utilizar o “Monte Carlo simples” em problemas de alta complexidade, desde que o metamodelo seja suficiente para representar corretamente o limite de falha.

## 2.4 Kriging

O metamodelo que será usado é o Kriging, sendo também chamado de krigagem. O Kriging se origina dos estudos de Wold em 1938, de Kolmogorov em 1939 e 1941 e de Wiener em 1949 (OLEA, 1999). Porém o método só foi popularizado em 1963 por meio dos trabalhos de Georges Matheron e seus colegas, dando o nome de Kriging por se basearem nos estudos de regressão e mineração de Daniel G. Krige realizados em 1951 (OLIVER; WEBSTER, 2015).

O Kriging é um método de regressão com forte embasamento estatístico utilizando os dados amostrais e as relações espaciais entre estes dados (CHAUVET; GALLI, 1982; VENKATRAM, 1988). Uma das vantagens deste embasamento estatístico é a análise da variância ser específica para o método, alguns métodos gerais para estimar a variância podem ser encontrados nos textos de Ryan (2009), onde pode-se encontrar também diversos modelos de regressão distintos.

O Kriging considera que os pontos desconhecidos são variáveis aleatórias com uma determinada tendência e um valor de ruído. A tendência é o valor esperado para um ponto qualquer no espaço amostral e o ruído pode ser representado por um desvio-padrão esperado para este mesmo ponto (CRESSIE, 1993; FAZIO; ROISENBERG, 2013).

$$\hat{y} = t + r. \quad (2.45)$$

Existem diferentes formas do modelo Kriging, entretanto para descrever a “krigagem ordinária” (*Ordinary Kriging, OK*) muitas vezes utiliza-se apenas “Kriging” por ser o tipo mais comum. Pode-se definir a diferença entre os tipos de Kriging simplificados por (CHIÈS; DELFINER, 2012):

- Kriging Simples (*Simple Kriging, SK*): a tendência média é conhecida, sendo ela constante ou nula;
- Kriging/Kriging Ordinário (*OK*): possui uma tendência média constante porém desconhecida;
- Kriging Global ou Universal (*Universal Kriging, UK*): a tendência é uma função das coordenadas, geralmente desconhecida com coeficientes ajustados durante a otimização do modelo.

Outro tipo relevante de Kriging é a cokrigagem ou *coKriging*. Este tipo é interessante por ser utilizado para problemas de mais de uma variável de interesse, particularmente usado quando há certeza de influência por erros de medição (CHIÈS; DELFINER, 2012). Apesar de sua importância, a diferença principal é a dimensão de resposta, inclusive existindo coKriging simples, ordinário e universal.

## 2.5 Variograma

Para que seja possível gerar aproximações a partir de um conjunto de dados pode-se utilizar mais do que apenas as características individuais e as relações entre pares isolados de variáveis aleatórias. Portanto analisam-se as funções de densidade espectral (*spectral density function*) (BOX; JENKINS, 1976).

Generalizando, a função espectral de densidade é um meio de identificar o comportamento de uma determinada variável ao longo de uma dimensão, usualmente buscando periodicidades. É com esta caracterização de períodos que a aproximação de dados desconhecidos pode ser aprimorada.

Buscando na literatura, encontra-se formas específicas dessa função espectral, mas existem algumas contradições entre autores de áreas distintas. Chatfield indica as diferenças entre os termos mais utilizados (CHATFIELD, 2003), por exemplo para indicar a validade do termo “periodograma” na formulação demonstrada por ele em relação ao termo tecnicamente correto “espectrograma” descrito por Anderson (1971).

Neste trabalho será utilizada a definição de variograma usada em geo estatística. Segundo Isaaks e Srivastava (1989) a variação entre duas realizações de uma função aleatória  $Y(x)$  é caracterizada pelo dobro da função do variograma ( $2\gamma(y_i, y_j)$ ), sendo o

variograma descrito pela equação:

$$\gamma(y_i, y_j) = \frac{E[(Y(x_i) - Y(x_j))^2]}{2} \quad (2.46)$$

Por definição, para uma posição  $x_i$  fixa, todos os pontos de mesma distância  $h$  devem obter o mesmo resultado no variograma, portanto pode-se escrever:

$$\gamma(x, h) = \frac{E[(Y(x) - Y(x + \vec{v}))^2]}{2} \quad (2.47)$$

sendo  $\vec{v}$  um vetor de módulo  $h$ . Além disso, caso o variograma da função aleatória  $Y(x)$  não varie com a posição  $x$ , pode-se escrever o variograma apenas em função da distância ( $\gamma(h)$ ) (JOURNEL; HUIJBREGTS, 1978).

Para determinar o variograma a partir de uma amostra, usa-se a definição de “nuvem de variograma” (*variogram cloud*) (WACKERNAGEL, 2003). Esta nuvem é gerada pelo gráfico definido por:

$$\begin{aligned} \text{abcissas} &\rightarrow \text{dist}[x_i, x_j] = h \\ \text{ordenadas} &\rightarrow \gamma_{ij}^* = \frac{(y_i - y_j)^2}{2}, \end{aligned} \quad (2.48)$$

sendo  $y_i$  a medida da variável aleatória na posição indicada por  $x_i$ ,  $\gamma_{ij}^*$  o valor do variograma experimental do par  $i, j$  e  $\text{dist}[x_i, x_j]$  a distancia entre as posições  $x_i$  e  $x_j$ .

Existindo uma nuvem experimental, pode-se estimar o variograma como sendo a função que melhor descreve o comportamento da “nuvem de variograma”.

Importante ressaltar que o Kriging considera que o variograma de fato não varia apenas com a distância  $h$  e respeita uma função contínua. Pode-se observar também que o valor de  $\gamma(0)$  é nulo, pois a diferença esperada para a equação 2.47 é nula quando  $\vec{v}$  tem módulo zero.

## 2.6 Cálculo do Kriging

Para o desenvolvimento dos cálculos, utilizou-se na maior parte o trabalho de Oliver e Webster (2015) e reestruturou-se o equacionamento para manter uma mesma notação em todo o texto.

Supondo uma determinada função aleatória  $G(x) = y$  e usando uma amostra de pontos observados  $\mathbf{X} = \{\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \dots, \mathbf{x}^{(n)}\}^T$  para gerar a resposta  $\mathbf{Y} = \{y^{(1)}, y^{(2)}, \dots, y^{(n)}\}^T$ . Busca-se aproximar a resposta  $y^{(0)}$  para um determinado ponto  $\mathbf{x}^{(0)}$ . O conjunto  $\mathbf{X}$  é formado por pontos denominados de pontos de suporte, além disso a função  $y$  pode depender de mais de uma variável, sendo estas variáveis comprimidas em um vetor único  $\mathbf{x}^{(i)} = x_1^{(i)}, x_2^{(i)}, \dots, x_k^{(i)}$ . Destacando que  $\mathbf{X}$  é uma matriz  $n \times k$  e  $\mathbf{Y}$  é um vetor  $n \times 1$  onde

$y^{(i)} = y(\mathbf{x}^{(i)})$ , porém a notação matricial  $\mathbf{Y}$  será mantida,  $n$  representa o número de pontos de suporte e  $k$  a dimensão de  $\mathbf{x}$ . Será usada a notação  $y_i$  no lugar de  $y^{(i)}$  com o objetivo de facilitar a compreensão e visualização de todo o equacionamento pois  $y^{(i)}$  possui apenas uma dimensão dimensão.

Para aproximar um novo ponto  $x^{(0)}$  que resulta em  $y_0$  com o Kriging teremos uma combinação dos pontos de suporte por meio de uma soma de pesos:

$$\hat{y}_0 = \sum_{i=1}^n \lambda_i y_i, \quad (2.49)$$

sendo  $\hat{y}_0$  o símbolo usado para o valor aproximado de  $y_0$ .

Para tornar o Kriging um aproximador não tendencioso em relação a média, além da equação 2.49, tem-se que:

$$\sum_{i=1}^n \lambda_i = 1 \quad (2.50)$$

e portanto:

$$\begin{aligned} E[\hat{y}_0 - y_0] &= 0, \\ \mu(\hat{y}_0) &= y_0. \end{aligned} \quad (2.51)$$

A variância de  $\hat{y}_0$  pode ser calculada por:

$$\sigma^2(\hat{y}_0) = E[(\hat{y}_0 - y_0)^2] = 2 \sum_{i=1}^n \lambda_i \gamma(y_i, y_0) - \sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j \gamma(y_i, y_j), \quad (2.52)$$

sendo  $\gamma$  o semi-variograma. Para compreender melhor como chega-se no resultado da equação 2.52 observa-se que:

$$\begin{aligned} E[(\hat{y}_0 - y_0)^2] &= E[\hat{y}_0^2 - 2y_0\hat{y}_0 + y_0^2] \\ \sigma^2(\hat{y}_0) &= E[A + B + y_0^2]. \end{aligned} \quad (2.53)$$

Usando a separação em  $A$  e  $B$  para facilitar a visualização de cada parte:

$$\begin{aligned} A &= \hat{y}_0^2 = \left( \sum_{i=1}^n \lambda_i y_i \right)^2 \\ A &= \left( \sum_{i=1}^n \lambda_i y_i \right) \left( \sum_{j=1}^n \lambda_j y_j \right) \\ A &= \sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j (y_i y_j), \end{aligned} \quad (2.54)$$

e como:

$$y_i y_j = \frac{1}{2} [y_i^2 + y_j^2 - (y_i - y_j)^2], \quad (2.55)$$

então escreve-se:

$$\begin{aligned}
 A &= \sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j \frac{1}{2} [y_i^2 + y_j^2 - (y_i - y_j)^2] \\
 A &= \frac{1}{2} \sum_{j=1}^n \sum_{i=1}^n \lambda_i \lambda_j y_i^2 + \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j y_j^2 - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j (y_i - y_j)^2 \\
 A &= \frac{1}{2} \sum_{j=1}^n \lambda_j \sum_{i=1}^n \lambda_i y_i^2 + \frac{1}{2} \sum_{i=1}^n \lambda_i \sum_{j=1}^n \lambda_j y_j^2 - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j (y_i - y_j)^2,
 \end{aligned} \tag{2.56}$$

sabendo que:

$$\sum_{j=1}^n \lambda_j \sum_{i=1}^n \lambda_i y_i^2 = \sum_{i=1}^n \lambda_i \sum_{j=1}^n \lambda_j y_j^2 \tag{2.57}$$

e como definido na equação 2.50:

$$\sum_{i=1}^n \lambda_i = 1, \tag{2.58}$$

a equação 2.56 pode ser resumida em:

$$A = \sum_{i=1}^n \lambda_i y_i^2 - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j (y_i - y_j)^2. \tag{2.59}$$

Enquanto  $B$ :

$$\begin{aligned}
 B &= -2y_0 \hat{y}_0 = -2y_0 \left( \sum_{i=1}^n \lambda_i y_i \right) \\
 B &= -2 \sum_{i=1}^n \lambda_i (y_0 y_i),
 \end{aligned} \tag{2.60}$$

de maneira similar a equação 2.55 tem-se que:

$$y_i y_0 = \frac{1}{2} (y_i^2 + y_0^2 - (y_i - y_0)^2), \tag{2.61}$$

portanto:

$$\begin{aligned}
 B &= -\frac{2}{2} \sum_{i=1}^n \lambda_i (y_i^2 + y_0^2 - (y_i - y_0)^2) \\
 B &= -\sum_{i=1}^n \lambda_i y_i^2 - \sum_{i=1}^n \lambda_i y_0^2 + \sum_{i=1}^n \lambda_i (y_i - y_0)^2 \\
 B &= -\sum_{i=1}^n \lambda_i y_i^2 - y_0^2 + \sum_{i=1}^n \lambda_i (y_i - y_0)^2.
 \end{aligned} \tag{2.62}$$

Aplicando as equações 2.59 e 2.62 na equação 2.53, chega-se em:

$$\begin{aligned}
\sigma^2(\hat{y}_0) &= E \begin{bmatrix} A \\ +B \\ +y_0^2 \end{bmatrix} = E \begin{bmatrix} \sum_{i=1}^n \lambda_i y_i^2 - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j (y_i - y_j)^2 \\ - \sum_{i=1}^n \lambda_i y_i^2 - y_0^2 + \sum_{i=1}^n \lambda_i (y_i - y_0)^2 \\ +y_0^2 \end{bmatrix} \\
\sigma^2(\hat{y}_0) &= E \left[ \sum_{i=1}^n \lambda_i (y_i - y_0)^2 - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j (y_i - y_j)^2 \right] \quad (2.63) \\
\sigma^2(\hat{y}_0) &= E \left[ \sum_{i=1}^n \lambda_i \frac{2}{2} (y_i - y_0)^2 \right] - E \left[ \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j (y_i - y_j)^2 \right] \\
\sigma^2(\hat{y}_0) &= 2 \sum_{i=1}^n \lambda_i E \left[ \frac{(y_i - y_0)^2}{2} \right] - \sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j E \left[ \frac{(y_i - y_j)^2}{2} \right].
\end{aligned}$$

Pode-se definir o variograma  $\gamma$  entre os pontos  $y_i$  e  $y_j$  pela equação 2.46, tornando fácil visualizar que a equação 2.63 resultará na equação 2.52, repetida parcialmente aqui:

$$\sigma^2(\hat{y}_0) = 2 \sum_{i=1}^n \lambda_i \gamma(y_i, y_0) - \sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j \gamma(y_i, y_j). \quad (2.52 \text{ Repetida})$$

Para definir os pesos então, basta minimizar o valor aproximado do desvio-padrão de  $\hat{y}_0$  considerando a restrição de igualdade da equação 2.50. Para isso, utiliza-se do método do multiplicador de Lagrange:

$$\nabla \left[ \sigma^2(\hat{y}_0) + M_L \left( 1 - \sum_{i=1}^n \lambda_i \right) \right]_{(\lambda, M_L)} = 0, \quad (2.64)$$

sendo  $M_L$  o multiplicador de Lagrange e  $\nabla[\cdot]_{(\lambda, M_L)}$  o operador de gradiente em relação a cada variável  $\lambda_i$  e ao  $M_L$ . O multiplicador de Lagrange torna o problema de otimização restrungido por  $\sum \lambda_i = 1$  em um problema não restrungido. Para melhor compreensão do multiplicador de Lagrange recomenda-se os trabalhos de Dano (1975), Bertsekas (1999), Arora (2004), Griva, Nash e Sofer (2009), Luenberger e Ye (2016).

Substituindo o desvio da equação 2.52 na equação 2.64 obtêm-se:

$$\nabla \left[ 2 \sum_{i=1}^n \lambda_i \gamma(y_i, y_0) - \sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j \gamma(y_i, y_j) + M_L - M_L \sum_{i=1}^n \lambda_i \right]_{(\lambda, M_L)} = 0 \quad (2.65)$$

e resolvendo o gradiente:

$$\begin{aligned}
\frac{\partial}{\partial \lambda_i} &\rightarrow \gamma(y_i, y_0) - \sum_{j=1}^n \lambda_j \gamma(y_i, y_j) - \frac{M_L}{2} = 0 \\
\frac{\partial}{\partial M_L} &\rightarrow 1 - \sum_{i=1}^n \lambda_i = 0.
\end{aligned} \quad (2.66)$$

Reorganizando as equações em formato matricial, pode-se chegar no sistema:

$$\begin{bmatrix} \gamma(y_1, y_1) & \gamma(y_1, y_2) & \dots & \gamma(y_1, y_n) & 1 \\ \gamma(y_2, y_1) & \gamma(y_2, y_2) & \dots & \gamma(y_2, y_n) & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \gamma(y_n, y_1) & \gamma(y_n, y_2) & \dots & \gamma(y_n, y_n) & 1 \\ 1 & 1 & \dots & 1 & 0 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \\ \frac{M_L}{2} \end{bmatrix} = \begin{bmatrix} \gamma(y_1, y_0) \\ \gamma(y_2, y_0) \\ \vdots \\ \gamma(y_n, y_0) \\ 1 \end{bmatrix}. \quad (2.67)$$

Observando este sistema pode-se perceber que, definindo a função do semi-variograma, os resultados para cada  $\lambda_i$  podem ser calculados, e assim calcular os resultados de  $\hat{y}_0$  e  $\sigma(\hat{y}_0)$ . Portanto tudo que precisa ser feito é definir a função  $\gamma(\cdot, \cdot)$ .

Com a função  $\gamma(\cdot, \cdot)$  e os valores de  $\mathbf{Y}$  definidos, observa-se que os valores de  $\lambda_i$  dependerão de  $y_0 = G(x^{(0)})$ . Sendo  $\mathbf{x}^{(0)}$  a variável de entrada desconhecida pode-se usar a notação simplificada  $\lambda_i(\mathbf{x})$  e  $\hat{y}(x) = \sum_{i=1}^n y_i \lambda_i(\mathbf{x})$ .

### 2.6.1 Exemplo de resolução do Kriging

O método de Kriging mais comum na literatura para uso conjunto com a simulação de Monte Carlo é o Kriging ordinário, sendo que, majoritariamente, apenas uma mesma função de correlação é usada. Portanto usa-se esta seção para exemplificar o cálculo completo do Kriging, sendo desenvolvido o equacionamento por meio da função de correlação descrita na equação 2.68.

Supondo uma determinada função  $G(\mathbf{x}) = y$  e usando os pontos de suporte  $\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}^T$  para gerar a resposta  $\mathbf{Y} = \{y_1, y_2, \dots, y_n\}^T$ .

Considera-se então que o conjunto  $\mathbf{Y}$  possui uma tendência fixa igual a  $\mathbf{1}\hat{\mu}$ , sendo  $\mathbf{1}$  um vetor  $n \times 1$  de valores 1, e que as variáveis do conjunto se correlacionam por meio da expressão:

$$\gamma(y^{(i)}, y^{(j)}) = \exp\left(-\sum_{m=1}^k \theta_m |x_m^{(i)} - x_m^{(j)}|^{p_m}\right), \quad (2.68)$$

sendo  $\theta_m$  e  $p_m$  parâmetros.

Com isso pode-se montar uma matriz  $n \times n$  de correlação para os pontos de suporte e a matriz de covariância:

$$\mathbf{\Psi} = \begin{bmatrix} \gamma(y_1, y_1) & \dots & \gamma(y_1, y_n) \\ \vdots & \ddots & \vdots \\ \gamma(y_n, y_1) & \dots & \gamma(y_n, y_n) \end{bmatrix}, \quad (2.69)$$

$$Cov(\mathbf{Y}, \mathbf{Y}) = \sigma^2 \mathbf{\Psi}. \quad (2.70)$$

A escolha da função que define a correlação da “variável”  $y$  mostrada na equação 2.68 representa o que se espera da função a ser aproximada. Para o caso geral assume-se que será contínua e suave.

A correlação escolhida depende das distâncias entre cada coordenada dos pontos de  $\mathbf{x}^{(i)}$  e  $\mathbf{x}^{(j)}$ , assim como os parâmetros  $\theta_m$  e  $p_m$  dependem da coordenada de  $x$  em questão. Apesar dos parâmetros não terem significado direto, os valores de  $\theta_m$  indicam os pesos das coordenadas de  $x_m$  quando  $p_m = 2$ . Quanto maior for o valor de  $\theta_m$  maior é a relevância da coordenada  $m$  (FORRESTER; SÓBESTER; KEANE, 2008). Essa afirmação nem sempre é válida caso  $\mathbf{x}$  não tenha a mesma distribuição ou proporção entre suas dimensões.

Escolher os valores de  $p_m$  e  $\theta_m$  pode ser um processo trabalhoso, entretanto usando a estimativa de máxima verossimilhança é possível reduzir a dispersão do aproximador com métodos de otimização para determinar os melhores valores para esses termos.

$$\mathcal{L}(\mathbf{Y}|\mu, \sigma) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left[-\sum_{i=1}^n \frac{(y_i - \mu)^2}{2\sigma^2}\right] \quad (2.71)$$

A função 2.71 pode ser expressa também da seguinte forma:

$$\mathcal{L}(\mathbf{Y}|\mu, \sigma) = \frac{1}{(2\pi\sigma^2)^{n/2} |\Psi|^{1/2}} \exp\left[-\frac{(\mathbf{Y} - \mathbf{1}\mu)^T \Psi^{-1} (\mathbf{Y} - \mathbf{1}\mu)}{2\sigma^2}\right] \quad (2.72)$$

Aplica-se um logaritmo para simplificar as equações, resultando em:

$$\ln(\mathcal{L}) = -\frac{n}{2} \ln(2\pi) - \frac{n}{2} \ln(\sigma^2) - \frac{1}{2} \ln|\Psi| - \frac{(\mathbf{Y} - \mathbf{1}\mu)^T \Psi^{-1} (\mathbf{Y} - \mathbf{1}\mu)}{2\sigma^2} \quad (2.73)$$

Usando as derivadas para encontrar o valor máximo para  $\mathcal{L}$  teremos:

$$\hat{\mu} = \frac{\mathbf{1}^T \Psi^{-1} \mathbf{Y}}{\mathbf{1}^T \Psi^{-1} \mathbf{1}} \quad (2.74)$$

$$\hat{\sigma}^2 = \frac{(\mathbf{Y} - \mathbf{1}\hat{\mu})^T \Psi^{-1} (\mathbf{Y} - \mathbf{1}\hat{\mu})}{n} \quad (2.75)$$

Com isso verifica-se que  $\hat{\mu}$  e  $\hat{\sigma}$  são números definidos pela amostra por meio dos termos  $n$  e pelo vetor  $\mathbf{Y}$ , assim como pela matriz  $\Psi$  que possui a função escolhida para correlação com os parâmetros a serem otimizados. Por fim otimiza-se a equação 2.73, sendo usados apenas os termos que variam com os parâmetros  $\theta_m$  e  $p_m$ :

$$\ln(\mathcal{L}) \approx -\frac{n}{2} \ln(\sigma^2) - \frac{1}{2} \ln|\Psi| \quad (2.76)$$

Importante notar que a equação 2.76 tem baixo custo computacional quando a amostra não é grande, caso contrário  $\Psi$  possui  $n \times n$  termos, sendo o mesmo invertido para o cálculo do  $\sigma^2$ , podendo ter um custo computacional significativo.

Após a otimização dos termos  $p_m$  e  $\theta_m$ , está preparado o modelo de Kriging, sendo este modelo usado para aproximar os pontos desconhecidos. Considerando que o ponto desconhecido seja  $\mathbf{x}$  aproxima-se a resposta de  $y$  por meio da seguinte equação:

$$\hat{y} = \hat{\mu} + \psi(\mathbf{x})^T \Psi^{-1} (\mathbf{Y} - \mathbf{1}\hat{\mu}) \quad (2.77)$$



Sendo  $\psi$  o vetor de correlação para o ponto  $y$  e cada elemento do conjunto de controle  $\mathbf{Y}$ .

$$\psi(\mathbf{x}) = \begin{pmatrix} \gamma(y_1, y) \\ \vdots \\ \gamma(y_n, y) \end{pmatrix} = \begin{pmatrix} \gamma(\mathbf{x}^{(1)}, \mathbf{x}) \\ \vdots \\ \gamma(\mathbf{x}^{(n)}, \mathbf{x}) \end{pmatrix} \quad (2.78)$$

Percebe-se que o vetor  $\psi$  não possui nenhum valor indeterminado, pois os termos da função de correlação foram determinados com a otimização da equação 2.76. Após gerar o modelo de Kriging a aproximação dos pontos possui um custo fixo relativamente baixo utilizando um vetor constante  $\Psi^{-1}(\mathbf{Y} - \mathbf{1}\hat{\mu})$  e o valor de  $\hat{\mu}$ .



## 3 Integração Kriging com Monte Carlo simples

O uso do Kriging para simulações estocásticas teve início com os estudos de Michell e Morris (1992). Mas a integração do Kriging com Monte Carlo só foi popularizada após os trabalhos de Echard, Gayton e Lemaire (2010) com o primeiro modelo do AKMCS disponível na literatura e Dubourg, Deheeger e Sudret (2011) usando técnicas de amostragem por importância em simulação de Monte Carlo com Kriging. Apesar do nome AKMCS ser o mais utilizado, é possível encontrar outros termos similares como *Active Learning Kriging* (ALK).

Neste capítulo serão analisadas algumas formas de Kriging utilizando a simulação de Monte Carlo simples. As análises serão do modelo inicial de AKMCS, seguido de duas variações de função de aprendizado (AKMCS+BiL e AKMCS+ewd). Continuando então uma análise de medidas de erro para o AKMCS, sendo possível usar estas medidas como critério de parada para os modelos discutidos anterior e posteriormente. Na sequência é analisado um procedimento com maior embasamento ao cálculo estatístico de Kriging, denominando-se de método de Kriging dependente (DKM). E finalmente mostra-se uma versão de Kriging mais focada ao uso prático para problemas de baixa probabilidade e funções de alto custo usando técnicas de redução de custo computacional (AK-MCSi), tanto por reduzir o uso de memória quanto processamento.

### 3.1 Método de AKMCS base - Kriging com aprendizado ativo

Echard, Gayton e Lemaire (2010) desenvolveram o primeiro modelo que popularizou o Kriging para uso com Monte Carlo na literatura, como afirmado anteriormente. Para o desenvolvimento inicial foram usadas duas formas distintas de função de aprendizado, chamadas de “EFF” (*Expected Feasibility Function*) e “U”, sendo usados nas tabelas de resultados deste e outros autores como “AKMCS+EFF” e “AKMCS+U”. Porém pode-se encontrar em comparações o uso do termo “AKMCS” referindo-se ao “AKMCS+U”, porque a função U é mais adequada que a função EFF (ECHARD; GAYTON; LEMAIRE, 2010), portanto a função padrão é definida desde o início. A utilização da função U tornou-se o padrão não apenas para o AKMCS, mas para qualquer método de Kriging com Monte Carlo, sendo a função de aprendizado utilizada ou a função a ser superada na maioria dos casos.

Sendo o primeiro método de Kriging com aprendizado programado integrado à simulação de Monte Carlo, os outros métodos que se originaram dele serão explicados de

maneira simplificada ressaltando apenas as diferenças.

Pode-se descrever a sequência do AKMCS de maneira resumida por: inicia-se o AKMCS gerando os pontos amostrais da simulação de Monte Carlo e sorteando alguns pontos de suporte iniciais para montar o primeiro modelo de Kriging; usa-se o Kriging para calcular todos os pontos amostrais de acordo com o modelo aproximado; os pontos são avaliados de acordo com uma função de aprendizado e, caso não seja atendido o critério de parada, o melhor ponto é escolhido para ser adicionado ao conjunto de controle do Kriging, volta-se então para a etapa do cálculo do modelo de Kriging; caso o critério de parada seja atendido verifica-se o coeficiente de variação da probabilidade de falha da simulação de Monte Carlo (checa-se se a quantidade de pontos amostrais é suficiente para representar a probabilidade de falha encontrada), caso o coeficiente de variação esteja elevado, aumenta-se o tamanho da amostra e o programa é reiniciado, caso ele esteja adequado o programa é finalizado e os resultados são mostrados.

Para descrever detalhadamente o procedimento estudado será utilizada a Figura 5 possibilitando visualizar o funcionamento do programa pelo esquema simplificado. Para uma descrição mais detalhada, deve-se descrever cada etapa do procedimento individualmente, porém será necessário detalhar os termos usados:

Pontos de suporte: é o nome dado para o conjunto de pontos que calculados pela função de desempenho que controla os parâmetros otimizados do Kriging;

$n_{Init}$ : é o número de pontos iniciais para o Kriging;

$n_{Final}$ : é o número de pontos de suporte do modelo de Kriging após o último ciclo, também é o número de avaliações da função de desempenho;

$\hat{P}_f$ : é o valor para a probabilidade de falha encontrado pelo procedimento, portanto o número de pontos no domínio de falha estimado pelo Kriging sobre o número de pontos amostrais do Monte Carlo.;

SampleNum: o número total de pontos amostrais da simulação de Monte Carlo;

Função de desempenho: A função que descreve o evento “falha”, sendo considerada de alto custo computacional, justificando o uso de um metamodelo para a aproximação. O termo “falha” é atribuído ao evento indesejado. Também pode-se encontrar na literatura o termo inglês *performance function*, função custo (*cost function*) ou função real.

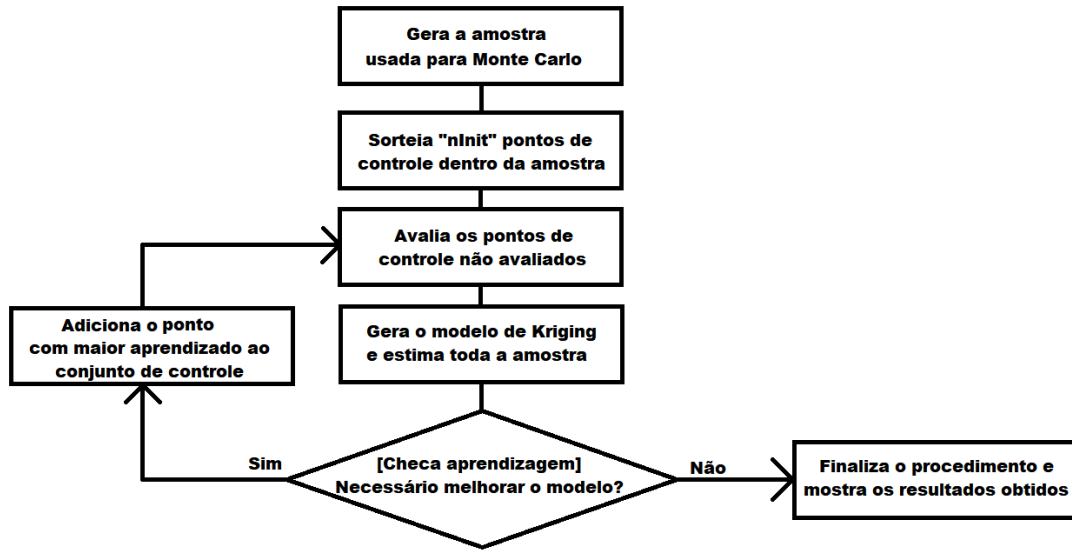
Passo 1 Gerar a amostra usada para o Monte Carlo.

Nesta etapa cada ponto amostral é gerado segundo as funções de densidade de probabilidade, sendo esta a mesma amostra usada para a simulação de Monte Carlo simples.

Passo 2 Sortear  $n_{Init}$  pontos de suporte iniciais.

Nesta etapa são escolhidos os pontos de suporte iniciais. Neste caso usam-se pontos aleatórios dentro da amostra de Monte Carlo, dessa forma aproveita-se melhor o

Figura 5 – Esquema do AKMCS geral



erro nulo dos pontos de suporte.

Passo 3 Avaliar a função de desempenho nos pontos de suporte ainda não avaliados.

A função de desempenho deve ser avaliadas nos pontos de suporte. Considerando que este passo é executado mais de uma vez, pontos já avaliados não precisam ser recalculados.

Passo 4 Gera o modelo de Kriging e aproxima-se todos os pontos da amostra.

Primeiramente os pontos de suporte são utilizados para gerar os sistemas de equações do Kriging, então os parâmetros do sistema são otimizados e depois os pontos da amostra são avaliados pelo modelo resultante.

Passo 5 [Tomada de decisão] Checagem de aprendizagem.

Avalia-se a função de aprendizagem escolhida e compara-se o melhor candidato com o valor limite de aprendizagem (critério de parada). Caso o critério não seja atendido deve-se ir para o Passo “A”, caso contrário prossegue-se ao Passo 6.

Para o AKMCS original observou-se a necessidade de forçar o programa a ir ao Passo “A” caso  $\hat{P}_f = 0$ , porque existe a possibilidade de atender ao critério de parada antes de encontrar o primeiro ponto de falha.

Passo A Adiciona novo ponto de suporte.

O passo A é o passo de aprendizado, sendo escolhido o melhor ponto da amostra para melhorar o modelo de Kriging. Após esta etapa, volta-se para o passo 3.

Passo 6 Finaliza o procedimento e mostra os resultados.

Por fim, o programa verifica o número de pontos que violaram a restrição e mostra

os valores obtidos para  $P_f$ .

Deve-se comparar o tamanho da amostra com a probabilidade de falha e decidir se o resultado é adequado com base no coeficiente de variação, apresentado na equação 2.44.

Ainda no Passo 6 pode-se escrever a probabilidade de falha segundo a equação 2.38 ou usando a função  $\pi(\mathbf{x}_i)$  introduzida por Picheny (2009):

$$\pi(\mathbf{x}_i) = P[\hat{G}(\mathbf{x}_i) \leq 0] = \Phi\left(\frac{-\mu_{\hat{G}}(\mathbf{x}_i)}{\sigma_{\hat{G}}(\mathbf{x}_i)}\right) \quad (3.1)$$

$$\hat{P}_f \approx \frac{1}{n} \sum_{i=1}^n \mathbf{I}_{\pi \geq 0,5}(\mathbf{x}_i), \quad (3.2)$$

sendo  $\mathbf{I}_{\pi \geq 0,5}(\mathbf{x}_i) = 1$  quando  $\hat{G}(\mathbf{x}_i) < 0$ , e  $\mathbf{I}_{\pi \geq 0,5}(\mathbf{x}_i) = 0$  caso  $\hat{G}(\mathbf{x}_i) > 0$ . A função identificadora  $I$  também é representada na literatura por  $\mathbb{1}$ . Enquanto  $\Phi$  indica a função de distribuição acumulada de uma variável normal padrão.

A função  $\pi(\mathbf{x}_i)$  representa a probabilidade da variável  $\mathbf{x}_i$  superar o limite de falha  $G(\mathbf{x}_i) \leq 0$  definido aproximadamente pelo método de Kriging  $\hat{G}(\mathbf{x}_i)$ .

Para o passo 5, deve-se definir a função de aprendizado e o critério de parada a ser utilizado. Pelo trabalho de Echard, Gayton e Lemaire (2010) são utilizadas duas funções de aprendizado. A primeira é o EFF estudado inicialmente no trabalho de Bichon et al. (2008), sua formulação é dada por:

$$\begin{aligned} EFF(\mathbf{x}) = & \hat{G}(\mathbf{x}) \left[ 2\Phi\left(\frac{-\hat{G}(\mathbf{x})}{\sigma_{\hat{G}}(\mathbf{x})}\right) - \Phi\left(\frac{-\varepsilon - \hat{G}(\mathbf{x})}{\sigma_{\hat{G}}(\mathbf{x})}\right) - \Phi\left(\frac{\varepsilon - \hat{G}(\mathbf{x})}{\sigma_{\hat{G}}(\mathbf{x})}\right) \right] \\ & - \sigma_{\hat{G}}(\mathbf{x}) \left[ 2\Phi\left(\frac{-\hat{G}(\mathbf{x})}{\sigma_{\hat{G}}(\mathbf{x})}\right) - \Phi\left(\frac{-\varepsilon - \hat{G}(\mathbf{x})}{\sigma_{\hat{G}}(\mathbf{x})}\right) - \Phi\left(\frac{\varepsilon - \hat{G}(\mathbf{x})}{\sigma_{\hat{G}}(\mathbf{x})}\right) \right] \\ & + \left[ \Phi\left(\frac{\varepsilon - \hat{G}(\mathbf{x})}{\sigma_{\hat{G}}(\mathbf{x})}\right) - \Phi\left(\frac{-\varepsilon - \hat{G}(\mathbf{x})}{\sigma_{\hat{G}}(\mathbf{x})}\right) \right], \end{aligned} \quad (3.3)$$

onde  $\sigma_{\hat{G}}$  é o desvio-padrão estimado pelo Kriging e  $\varepsilon$  é um parâmetro, sendo usado  $\varepsilon = 2\sigma_{\hat{G}}^2$ . Para esta função de aprendizado usa-se o critério de parada:

$$\max(EFF(\mathbf{x})) \leq 0,001. \quad (3.4)$$

Esta função é mais indicada para casos onde uma análise balanceada na região do limite de falha e uma busca global é desejada (BICHON et al., 2008). Segundo Echard, Gayton e Lemaire (2010), a função EFF não garante um bom funcionamento para o AKMCS, desenvolvendo e sugerindo então o uso da função U:

$$U(\mathbf{x}) = \frac{|\hat{G}(\mathbf{x})|}{\sigma_{\hat{G}}(\mathbf{x})}. \quad (3.5)$$

Esta função indica o número de desvios-padrão que a estimativa da função de desempenho para o ponto  $\mathbf{x}$  está do limite de falha. O limite de aprendizado sugerido para esta função é  $U \geq 2$ , portanto o menor valor indica o melhor candidato.

A função  $U$  foi baseada nos trabalhos de Cox e John (1992) onde usava-se o Kriging para otimização, buscando o ponto de menor valor possível dentro do domínio. Portanto o limite de aprendizado busca explorar todo o domínio e refinar a região de interesse extensivamente, possuindo precisão excessiva para o uso em casos usuais de confiabilidade.

### 3.1.1 Características do método

As características do método são inicialmente dependentes do tipo de Kriging e das escolhas advindas dele, para este e a maioria dos casos usou-se o Kriging ordinário com a função de correlação para a formulação do Kriging como sendo:

$$\gamma(y^{(i)}, y^{(j)}) = \exp\left(-\sum_{m=1}^k \theta_m |x_m^{(i)} - x_m^{(j)}|^2\right), \quad (3.6)$$

similar ao indicado na equação 2.68 com  $p_m = 2$ . Como esta correlação é considerada padrão para o Kriging, será omitida das próximas descrições.

As propriedades do Kriging ordinário com a função de correlação exponencial quadrática podem ser encontradas em literaturas como Journel e Huijbregts (1978), sendo elas:

1. Possui o desvio padrão minimizado por meio de uma etapa de otimização.
2. O intervalo de uso do estimador gerado não é restringido pelo intervalo do conjunto de controle.
3. O estimador resulta em valores de peso maiores para pontos de suporte próximos do ponto estimado.
4. O Kriging ordinário evita agrupamentos de pontos com alto peso, levando em consideração os pesos próximos.
5. Para os pontos de suporte, o valor estimado é exato.
6. Incapacidade de lidar com pontos de suporte repetidos.
7. Independência da origem do espaço amostral, levando em consideração apenas as distâncias relativas.

Além disso, pela formulação indicada anteriormente, é possível verificar que o Kriging gera curvas contínuas e relativamente suaves, sendo considerado por alguns como requisito para utilizar o Kriging (OLEA, 1999).

Apesar de não ser uma característica intrínseca ao Kriging, em problemas com grandes oscilações é necessário um maior número de pontos de suporte para aproximar corretamente todo o domínio. Observa-se entretanto que para problemas de confiabilidade onde existem apenas algumas regiões de interesse, a má representação das oscilações distantes do limite de falha não afetam o resultado final.

As outras características advém da distribuição inicial de pontos e da função de aprendizado usada.

A distribuição inicial usada foi uma seleção aleatória de pontos dentro da amostra de Monte Carlo, portanto há uma tendência maior de selecionar pontos mais próximos dos valores médios, pois há uma maior quantidade de pontos nesta região. Sendo assim essa distribuição gera uma aproximação da região central adequada mesmo para poucos pontos iniciais, porém a aproximação é inicialmente inadequada para a região de falha, uma vez que a probabilidade de existir algum ponto na região de falha é proporcional à probabilidade de falha do problema em questão.

Apesar de representar a região de falha de maneira insuficiente na etapa inicial, pode-se ter um bom resultado final com o uso do aprendizado subsequente caso a função usada para selecionar os novos pontos de suporte seja adequada. Sabendo que Echard, Gayton e Lemaire (2010) indicam que a função EFF não desempenha bem para o AKMCS, será apenas comentado que o seu uso gera uma aproximação melhor de todo o domínio. Por selecionar pontos de suporte nas regiões de interesse e fora delas a função EFF obtém resultados bons, mas possui um maior número de avaliações da função de desempenho.

Quanto a função de aprendizado  $U$ , sua formulação indica que os novos pontos de suporte são determinados de acordo com a incerteza individual de cada ponto. Esta incerteza está relacionada diretamente com a probabilidade de ser ou não definido corretamente como um ponto de falha, podendo ser descrito também pela função  $\pi$  da equação 3.3.

Considerando isso, pode-se afirmar que a função  $U$  busca o ponto que gera o maior efeito na aproximação do Kriging individualmente. Para verificar sua convergência para o resultado exato pode-se observar que cada ponto adicionado no conjunto de controle do Kriging gera, no mínimo, uma melhor aproximação para o ponto de maior valor  $U$ , portanto qualquer imprecisão tenderá a zero com o aumento do número de pontos de suporte.

Quanto ao critério de parada, no trabalho de Echard, Gayton e Lemaire (2010) não há uma análise detalhada, apenas comenta-se que  $U > 2$  seria um limite adequado para a parada da etapa de aprendizado. Foram testados alguns problemas e valores, sendo difícil definir a priori um valor verdadeiramente adequado para o limite de aprendizado. Na maioria dos casos 2 é um valor conservador, porém existem problemas no qual  $U > 2$  não é suficiente para explorar todo o domínio da amostra gerando uma aproximação da



região de falha insatisfatória. No trabalho de Noronha e Lopez (2016) há uma análise do valor limite de aprendizado, indicando os efeitos da alteração do valor limite para  $U$ .

## 3.2 AKMCS+BiL - adaptação para melhorar a curva limite

Observando o trabalho de Peijuan et al. (2017) verifica-se uma tentativa lógica de melhorar o procedimento do AKMCS+U. Esta versão é similar ao AKMCS original, porém na etapa de adicionar o ponto de melhor valor de aprendizado existem alterações. O artigo em questão não nomeia seu método, sendo escolhido o termo “AKMCS+BiL” (*AKMCS+Best in Line*) para referir-se a ele.

A única alteração está na seleção dos novos pontos de suporte, sendo verificados pontos fora do conjunto amostral. A função de aprendizado utilizada para a classificação ainda é a função  $U$ .

Para escolha do ponto de melhor valor de aprendizado observa-se o sinal estimado pelo Kriging  $\hat{G}(\mathbf{x})$  e procura-se o ponto da amostra de Monte Carlo mais próximo onde o sinal estimado seja contrário. Existindo estes dois pontos, geram-se novos pontos temporários entre eles seguindo a equação:

$$\mathbf{x}_t = \mathbf{x}_b(1 - t) + \mathbf{x}_s(t), \quad (3.7)$$

sendo  $t$  um valor entre 0 e 1, o índice  $b$  identifica o melhor ponto segundo a função  $U$  e o índice  $s$  indica o ponto de sinal oposto. Sendo para o trabalho citado usados 1.000 pontos dispostos entre os dois.

Importante ressaltar que, por usar as distâncias entre pontos, deve-se normalizar as variáveis aleatórias para eliminar os efeitos que as diferentes ordens de grandeza na amostra poderiam gerar.

### 3.2.1 Efeitos da alteração

A alteração feita obteve resultados melhores que o original, pois usa de pontos fora da amostra buscando estimar com precisão apenas o limite da região de falha. Portanto, com a definição mais exata de  $\hat{G} = 0$ , definir se está ou não dentro do limite de falha torna-se mais preciso, mesmo que o valor absoluto tenha uma aproximação pior.

Assim como a versão original do AKMCS, não existe garantia da convergência dentro do critério de parada determinado ( $U > 2$ ), poderia ser necessário o uso de uma maior quantidade de pontos de suporte. Outro possível efeito da alteração é a aglomeração de pontos no limite de falha, sendo indesejado o uso excessivo em uma pequena área do domínio enquanto o mesmo não tenha sido apropriadamente explorado.

Por fim, o uso do melhor ponto entre o candidato e o ponto mais próximo dele com sinal oposto gera uma incapacidade de encontrar modos de falha após a distribuição inicial de pontos. Caso exista uma única região aproximada com  $\hat{G} < 0$ , independente de onde os candidatos sejam escolhidos, o ponto a ser adicionado ao conjunto de controle será próximo do limite dessa região.

### 3.3 AKMCS+ewd - adaptação com função de aprendizado composta

O trabalho de Xiao, Zuo e Guo (2018) também foca-se em alterar a função de aprendizado buscando uma melhor forma de distribuir os novos pontos de suporte. Alterando a função  $U$  pela função de aprendizado “ $C$ ” que segue a seguinte equação:

$$C(\mathbf{x}_j) = e_p(\mathbf{x}_j)w(\mathbf{x}_j)[\mathbf{d}_{Emin}(\mathbf{x}_j, \mathbf{x}_{(s)})]^\alpha, \quad (3.8)$$

para melhor compreensão desta equação cada termo será explicado separadamente. Novamente o procedimento não foi nomeado, por tanto será usado AKMCS+ewd porque a função de aprendizado  $C$  utiliza três termos, “ $e \cdot w \cdot d$ ”.

O termo  $e_p(\mathbf{x})$  se relaciona com a medida de erro por validação cruzada k-fold. Sendo  $\pi(\mathbf{x}_j)$  o valor da função  $\pi$  atual e  $\pi_{-i}(\mathbf{x}_j)$  a função  $\pi$  originada pelo modelo de Kriging sem o ponto de suporte  $i$ , tem-se que:

$$e_p(\mathbf{x}_j) = \sqrt{\frac{1}{K} \sum_{i=1}^K (\pi(\mathbf{x}_j) - \pi_{-i}(\mathbf{x}_j))^2}. \quad (3.9)$$

O termo  $w(\mathbf{x})$  origina-se do trabalho de Dubourg, Sudret e Bourinet (2011) e utiliza a função de densidade de probabilidade de  $\mathbf{x}$  para comparar com um valor limite determinado, onde atribue-se valor 1 ou 0 para o  $w$ , ou seja, é um critério de eliminação.

$$w(\mathbf{x}_j) = \mathbb{1}[f_x(\mathbf{x}_j) > \rho_0^k], \quad (3.10)$$

onde  $\rho$  é o valor definido na iteração  $k$  do procedimento. Sendo que  $\rho$  é calculado a partir da probabilidade de falha estimada pelo Kriging na iteração anterior e sua fórmula completa pode ser verificada em Xiao, Zuo e Guo (2018).

Finalmente, o termo  $[\mathbf{d}_{Emin}(\mathbf{x}_j, \mathbf{x}_{(s)})]^\alpha$  é a distância mínima entre o ponto  $\mathbf{x}_j$  e os pontos do conjunto de controle  $\mathbf{x}_{(s)}$ . Além disso, o expoente  $\alpha$  utilizado variou entre 0, 2, 0, 3, 0, 4 ou 0, 5 ao longo dos problemas resolvidos por Xiao, Zuo e Guo (2018), buscando aumentar ou reduzir a relevância da distância.

Assim como no AKMCS+BiL, deve-se normalizar as variáveis aleatórias para que os termos calculados não sejam afetados pelas diferenças de ordem de grandeza. Tanto

o termo  $\mathbf{d}$  envolvendo distância, quanto o termo  $w$  que envolve a análise da distribuição aleatória das variáveis perdem sua eficiência caso a normalização das variáveis não seja realizada.

Quanto ao critério de parada, usa-se a verificação cruzada k-fold ( $k = 5$  ou  $10$ ) em relação a probabilidade de falha encontrada até chegar em um erro limite segundo a equação:

$$\frac{\Delta \hat{P}_f^{max}}{\hat{P}_f} \leq \varepsilon_f, \quad (3.11)$$

sendo  $\Delta \hat{P}_f^{max} = \max(|\hat{P}_f - \hat{P}_f^{-i}|)$  onde  $\hat{P}_f^{-i}$  é a probabilidade encontrada com o ponto de suporte  $i$  fora do modelo de Kriging e  $\varepsilon_f = 0,01$  ou  $0,02$ . Além disso, para qualquer iteração onde  $\hat{P}_f = 0$ , usa-se  $\frac{\Delta \hat{P}_f^{max}}{\hat{P}_f} = 1$ .

### 3.3.1 Efeitos da alteração

O uso da função de aprendizado  $C = e \cdot w \cdot d$  trouxe melhorias em relação ao AKMCS+U, entretanto não é uma melhora constante. Para alguns casos houve melhora no número de chamadas da função de desempenho, porém o erro da estimativa da probabilidade de falha aumentou, em outros o erro foi reduzido com o aumento das chamadas da função de desempenho, e apenas para alguns casos o AKMCS+ewd foi melhor que o AKMCS+U em ambos os critérios.

Pode-se verificar com a formulação da função de aprendizado que existem características interessantes e que poderiam ser reaproveitadas para qualquer função de aprendizado. A ideia do termo  $w$ , por exemplo, é limitar os novos pontos de suporte a uma região de interesse, enquanto o uso de um termo adicional relacionado a distância gera uma menor aglomeração dos pontos de suporte.

No trabalho de Xiao, Zuo e Guo (2018) há uma análise quanto a medida de peso para o termo da distância  $\alpha$  no qual avalia-se a faixa de valores mais eficazes. Essa análise é feita para problemas com uma maior ou menor quantidade de variáveis aleatórias na função de desempenho, levando a sugestão de usar pesos menores para problemas de maiores dimensões.

Reforça-se que o problema de menor probabilidade de falha feito por Xiao, Zuo e Guo (2018) possui  $P_f \approx 0,5\%$  onde o AKMCS+ewd possui o pior desempenho, enquanto os outros três problemas possuem  $P_f > 2,8\%$ . Este detalhe pode estar relacionado também com o critério de parada, pois usando “5-fold” ou “10-fold” a medida de precisão tende a ter menores variações com o aumento do conjunto de controle. Chega-se então a uma parada prematura quando o número de pontos de suporte necessários para ter uma boa aproximação aumenta.

### 3.4 Medida de precisão para o AKMCS

Apesar da maior parte das pesquisas se focar na função de aprendizado, fato determinante no desempenho do AKMCS, os pesquisadores Jian et al. (2017) buscaram uma forma mais adequada de estimar a precisão do procedimento.

Primeiramente, o texto traz uma alteração importante para o cálculo da probabilidade de falha, ao invés de usar a estimativa da função de desempenho pelo Kriging para definir os pontos da região de falha, usa-se o valor médio e desvio estimados para determinar a probabilidade de cada ponto estar no domínio de falha:

$$\begin{aligned}\hat{P}_f &\approx E[\hat{P}_f] \approx \frac{1}{n} \sum_{i=1}^n \frac{\mu_{\hat{G}}(\mathbf{x}_i)}{\sigma_{\hat{G}}} \\ \hat{P}_f &\approx \frac{1}{n} \sum_{i=1}^n \pi(\mathbf{x}_i),\end{aligned}\tag{3.12}$$

sendo  $\pi$  a função detalhada na equação 3.1.

O trabalho de Jian et al. (2017) também apresenta de maneira detalhada as medidas de erro por meio do embasamento estatístico da formulação do modelo de Kriging. Resumidamente pode-se escrever duas das medidas propostas segundo a equação:

$$\begin{aligned}e_{\hat{G}} &= \int |\mathbf{I}_{G \leq 0}(\mathbf{x}) - \mathbf{I}_{\hat{G} \leq 0}(\mathbf{x})| f(\mathbf{x}) d\mathbf{x} \\ e_I &= \int \mathbf{I}_{U \leq U_{ext}} f(\mathbf{x}) d\mathbf{x},\end{aligned}\tag{3.13}$$

sendo  $U$  a função de aprendizado descrita por Echard, Gayton e Lemaire (2010),  $U_{ext}$  é um valor extremo para  $U$  onde sugere-se usar  $U_{ext} = 2$  e  $f(\mathbf{x})$  é a PDF conjunta de  $\mathbf{x}$ . A função  $e_{\hat{G}}$  pode ser escrita também por:

$$e_{\hat{G}} = \frac{1}{n} \sum_{i=1}^n \Phi(-U(\mathbf{x}_i)),\tag{3.14}$$

observando que a função  $\Phi(-U(\mathbf{x}_i))$  é a probabilidade de definir incorretamente o sinal de  $\hat{G}(\mathbf{x}_i)$ , pois  $\Phi$  representa a função de distribuição acumulada para uma variável normal padrão.

Para a medida  $e_I$  entretanto, pode-se escrever como:

$$e_I = \frac{1}{n} \sum_{i=1}^n \mathbf{I}_{U \leq U_{ext}},\tag{3.15}$$

ou seja, qualquer ponto no qual  $U(\mathbf{x}_i) \leq U_{ext}$  é contabilizado para a medida de erro máximo total.

Além das duas medidas descritas, Jian et al. (2017) propõem uma nova medida mais adequada para o cálculo de probabilidade de falha usado pelo mesmo:

$$\begin{aligned} e_{\hat{P}_f}^2 &= E \left[ \left( \int \mathbf{I}_{G \leq 0}(\mathbf{x}) f(\mathbf{x}) d\mathbf{x} - \int \pi(\mathbf{x}) f(\mathbf{x}) d\mathbf{x} \right)^2 \right] \\ e_{\hat{P}_f}^2 &\approx E \left[ \left( \int_{U \leq U_{ext}} \mathbf{I}_{G \leq 0}(\mathbf{x}) f(\mathbf{x}) d\mathbf{x} - \int_{U \leq U_{ext}} \pi(\mathbf{x}) f(\mathbf{x}) d\mathbf{x} \right)^2 \right] \\ e_{\hat{P}_f}^2 &\approx E \left[ \frac{1}{n^2} \sum_{i=1}^n (\mathbf{I}_{G \leq 0}(\mathbf{x}_i) - \pi(\mathbf{x}_i))^2 \right], \end{aligned} \quad (3.16)$$

podendo ser descrito como soma do  $n$  pontos da amostra de Monte Carlo por:

$$e_{\hat{P}_f}^2 \approx \frac{1}{n^2} \left\{ \begin{aligned} &\sum_{U(\mathbf{x}_i) \leq U_{ext}} \Phi(U(\mathbf{x}_i)) \Phi(-U(\mathbf{x}_i)) \\ &+ \sum_{U(\mathbf{x}_i) \leq U_{ext}} \sum_{U(\mathbf{x}_j) \leq U_{ext}} [P[\hat{G}(\mathbf{x}_i) \leq 0, \hat{G}(\mathbf{x}_j) \leq 0] - \pi(\mathbf{x}_i)\pi(\mathbf{x}_j)] \end{aligned} \right\}. \quad (3.17)$$

### 3.4.1 Análise das medidas propostas

Inicialmente será analisada a medida de erro  $e_I$  por ser a mais simples das três. O erro  $e_I$  é a soma de todos os pontos com uma certeza estimada de estar dentro ou fora do domínio de falha igual a  $\Phi(U_{ext})$ . Utilizando  $U_{ext} = 2$  temos que  $\Phi(U_{ext}) \approx 97,7\%$ .

Supondo  $n_e$  pontos com  $U \leq U_{ext}$  e fazendo uma leve análise pode-se concluir que o valor de  $e_I$  considera um limite de  $n_e$  pontos errados, sendo qualquer caso de  $n_e + 1$  ou mais erros desconsiderado. A probabilidade de ocorrência de um caso com ao menos  $n_e + 1$  erros é igual a combinação de  $n_e + 1$  pontos errados e  $n - (n_e + 1)$  corretos. Para ser possível calcular, pode-se considerar o caso de maior probabilidade onde existem  $n_e$  pontos onde  $U = 0$  e  $n - n_e$  pontos com  $U > U_{ext}$ , onde todos os pontos não exatos possuem  $U \approx U_{ext}$ :

$$P[n_{erros} \geq n_e + 1] = 0,50^m \sum_{i=1}^{n_e+1} \frac{(n - n_e - n_c)! 0,978^{n-n_e-i-n_c} 0,002^i}{(n - n_e - n_c - i)! (i)!}, \quad (3.18)$$

sendo  $n$  o número de pontos amostrais,  $n_e$  o número de erros estimados por  $U \leq U_{ext}$  e  $n_c$  o número de pontos “corretos” ou aproximadamente exatos (usualmente os pontos de suporte e próximos deles). Com isso é possível definir a probabilidade máxima da medida de erro  $e_I$  estar inadequada.

Importante ressaltar que este valor é um valor extremo, pois é quase impossível que todos os pontos se agrupem em  $U = 0$  e  $U \approx U_{ext}$ , porém uma análise combinatória real seria computacionalmente impossível para qualquer amostra maior que 200 pontos. Para demonstrar alguns valores pode-se observar a tabela 1, sendo possível reparar que para uma situação com mais de 1000 pontos a estimativa de erro  $e_I$  não só não será superada, quanto é extremamente conservadora. Não são mostrados valores maiores que  $n = 190$  pois o fatorial de 200 gera números na ordem de  $10^{374}$  gerando erros de precisão computacional consideráveis.

Tabela 1 – Probabilidade máxima de superar o valor de  $e_I$ 

| n   | $n_e$ | $n_c$ | Probabilidade |
|-----|-------|-------|---------------|
| 10  | 1     | 1     | 0,685%        |
| 100 | 10    | 10    | 0,003%        |
| 150 | 10    | 10    | 0,0016%       |
| 190 | 10    | 10    | 0,0009%       |

Outro detalhe é que a medida  $e_I$  foi desenvolvida por Sun et al. (2017) para ser usada quando a probabilidade de falha é definida pela equação 3.1 e não pela soma da função  $\pi$  como na equação 3.12; com isso pode-se concluir que a probabilidade do valor real superar a medida  $e_I$  é ainda menor.

Para a segunda medida, será analisado o  $e_{\hat{G}}$ . Esta medida de precisão calcula o erro esperado de forma similar a soma da probabilidade de falha da equação 3.12, sendo cada ponto responsável por  $1/n$  do  $\hat{P}_f$  e tendo uma chance de ser estimado errado de  $\Phi(-U)$ . O  $e_{\hat{G}}$ , ao contrario da primeira medida, não é um valor extremo, seu calculo gera um resultado esperado de erro. Por ser um valor médio de erro, Jian et al. (2017) utiliza um multiplicador, 2 ou 3, para uma comparar o  $e_{\hat{G}}$  com o  $e_I$  de maneira justa.

As duas medidas de precisão discutidas anteriormente desconsideram efeitos de compensação no erro, caso pontos estimados dentro e fora do domínio de falha estejam errados, o erro total é a diferença das duas quantidades. Além disso, ambas as medidas foram desenvolvidas para uma probabilidade de falha definida pela equação 3.2, e não pela equação 3.12. Por este motivo desenvolve-se o equacionamento de  $e_{\hat{P}_f}^2$  (JIAN et al., 2017).

Por ser usado em um equacionamento que considera o valor esperado de  $\hat{P}_f$  ao invés de usar o valor médio de  $\hat{G}((x))$  na função identificadora  $\mathbf{I}$ , os valores calculados para  $e_{\hat{P}_f}^2$  são menores.

Analisando o equacionamento de  $e_{\hat{P}_f}^2$  e comparando-o com uma medida mais simples, sendo ela a diferença das duas formas de calcular a probabilidade de falha (equações 3.2 identificada por  $\hat{P}_{fI}$  e equação 3.12 por  $\hat{P}_{f\pi}$ ):

$$|\hat{P}_{fI} - \hat{P}_{f\pi}| = \frac{1}{n} \left| \sum_{i=1}^n \mathbf{I}_{\pi \geq 0,5}(\mathbf{x}_i) - \sum_{i=1}^n \pi(\mathbf{x}_i) \right| \quad (3.19)$$

pode-se ver claramente que a diferença é a posição do operador de módulo. No caso de  $e_{\hat{P}_f}^2$  aplica-se a soma das diferenças absolutas, resultando em uma estimativa de erro maior. Este maior valor ocorre por anular a compensação de erro comentada anteriormente e somar erros opostos.

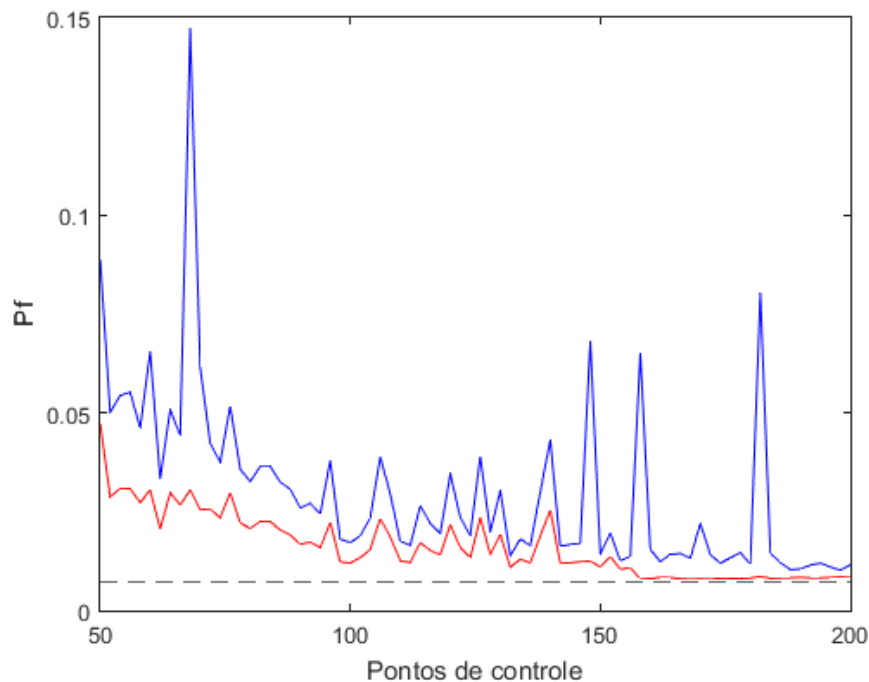
### 3.4.2 Efeitos da alteração

Apesar de não haver grandes alterações no resultado final, a alteração do cálculo da probabilidade de falha gera um efeito nos resultados intermediários e principalmente nos resultados iniciais. Com o uso de pesos parciais de falha para cada ponto da amostra os valores estimados para  $\hat{P}_f$  ao longo das iterações do procedimento possuem uma estabilidade maior.

Esta estabilidade ocorre porque os pontos com maior imprecisão geram um efeito parcial, enquanto os pontos mais precisos são considerados exatos, sendo suas falhas contabilizadas por um peso total próximo de 1 ou 0. Além disso, de maneira geral, não ocorrem estimativas onde  $\hat{P}_f = 0$  nas etapas iniciais, pois existirão pontos com imprecisão que irão contabilizar algum valor de falha.

Outra característica é que o cálculo por meio da soma da função  $\pi$  gera resultados mais próximos do valor de  $P_f$  real com uma menor quantidade de avaliações da função de desempenho. Importante ressaltar que os valores de  $\hat{P}_f$  convergem para o mesmo resultado com o aumento do conjunto de controle, sendo aproximadamente iguais para a maioria dos critérios de falha propostos na literatura.

Figura 6 – Comparação entre  $\hat{P}_{fI}$  (curva superior em azul),  $\hat{P}_{f\pi}$  (curva inferior em vermelho) e  $P_{fSMC}$  (linha tracejada, valor de referência).



### 3.5 DKM - método de Kriging dependente

O trabalho de Zhu e Du (2016) busca utilizar as características do Kriging para gerar uma função de aprendizado mais adequada para encontrar o valor da probabilidade de falha. Sendo assim, o equacionamento necessário para gerar a função de aprendizado final é detalhado. Por não usar a função U, o critério de parada também foi alterado, buscando ainda se adequar melhor com o uso do Kriging para estimar a probabilidade de falha.

Este é um dos poucos métodos nomeados por seu autor. Ele escolhe o nome DKM (*Dependent Kriging Method*) por ser um procedimento desenvolvido exclusivamente para utilização com Kriging. Segundo Zhu e Du (2016), os outros métodos da literatura podem ser nomeados de “IKM” (*Independent Kriging Methods*) por desconsiderarem que o Kriging segue um processo gaussiano, sendo desconsiderada a dependência de suas realizações e simplificando a precisão final do método.

Primeiramente são demonstrados os cálculos estatísticos resultantes do uso de Kriging para o cálculo da probabilidade de falha, chegando nos seguintes resultados (ZHU; DU, 2016):

$$\begin{aligned}\hat{P}_f &= \frac{1}{n} \sum_{i=1}^n \mathbf{I}_{\hat{G} \leq 0}(\mathbf{x}_i) \\ E[\hat{P}_f] &= \frac{1}{n} \sum_{i=1}^n \pi_i \\ \text{Var}[\hat{P}_f] &= \frac{1}{n^2} \left[ \sum_{i=1}^n \sum_{j=1}^n (\pi_{ij} - \pi_i \pi_j) \right],\end{aligned}\tag{3.20}$$

sendo:

$$\begin{aligned}\pi(\mathbf{x}_i) &= P[\hat{G}(\mathbf{x}_i) \leq 0] = \pi_i \\ \pi_{ij} &= P[\hat{G}(\mathbf{x}_i) \leq 0, \hat{G}(\mathbf{x}_j) \leq 0] \\ \pi_{ii} &= \pi_i \\ \text{Var}[\hat{P}_f] &= \frac{1}{n^2} \text{Var} \left[ \sum_{i=1}^n \mathbf{I}_{\hat{G} \leq 0}(\mathbf{x}_i) \right] \\ \text{Var}[\hat{P}_f] &= \frac{1}{n^2} \left[ \sum_{i=1}^n \sum_{j=1}^n \text{Cov}[\mathbf{I}_{\hat{G} \leq 0}(\mathbf{x}_i), \mathbf{I}_{\hat{G} \leq 0}(\mathbf{x}_j)] \right]\end{aligned}\tag{3.21}$$

$$\text{Cov}[\mathbf{I}_{\hat{G} \leq 0}(\mathbf{x}_i), \mathbf{I}_{\hat{G} \leq 0}(\mathbf{x}_j)] = \pi_{ij} - \pi_i \pi_j.$$

Com isso define-se que a função de aprendizado mais adequada seria:

$$c_i = \sum_{j=1}^n (\pi_{ij} - \pi_i \pi_j),\tag{3.22}$$

sendo  $c_i$  o valor de aprendizado para o ponto  $\mathbf{x}_i$ . Portanto é escolhido o ponto de maior contribuição a partir de uma estatística envolvendo todos os pontos amostrais.



Para o critério de parada, escolhe-se um percentual de confiança para o resultado final igual a  $1 - 2\alpha$ , resultando num intervalo de resultados válidos de  $E[\hat{P}_f] \pm \Phi^{-1}(\alpha)\sigma_{\hat{P}_f}$ . O erro relativo permitido pode ser computado por:

$$\frac{|E[\hat{P}_f] \pm \Phi^{-1}(\alpha)\sigma_{\hat{P}_f} - E[\hat{P}_f]|}{E[\hat{P}_f]} = \left| \frac{\Phi^{-1}(\alpha)\sigma_{\hat{P}_f}}{E[\hat{P}_f]} \right| = \frac{|\Phi^{-1}(\alpha)| \sigma_{\hat{P}_f}}{E[\hat{P}_f]}. \quad (3.23)$$

Aceitando um erro relativo  $\varepsilon$  pequeno, pode-se definir o critério de parada como:

$$\frac{\sigma_{\hat{P}_f}}{E[\hat{P}_f]} \leq \left| \frac{\varepsilon}{\Phi^{-1}(\alpha)} \right|. \quad (3.24)$$

### 3.5.1 Efeitos da alteração

A alteração da função aprendizado aplicada gerou resultados que, no geral, necessitam de menos de metade das chamadas da função de desempenho em relação ao procedimento original do AKMCS+U. Os resultados mostrados por Zhu e Du (2016) indicam que o uso de  $\pi_{ij} - \pi_i\pi_j$  gera uma estimativa eficiente por levar em consideração a influencia de todos os pontos amostrais para a definição do melhor candidato.

Pode-se chegar a conclusão de que uma função de aprendizado que considera a influencia dos outros pontos é mais adequada até mesmo na formulação do Kriging. Cada novo ponto pode ser descrito por uma soma ponderada dos valores dos pontos de suporte. Como essa ponderação é baseada na correlação utilizada, usar a estimativa da covariância para definir os novos pontos de suporte é teoricamente a melhor forma de classificar os candidatos.

Apesar das vantagens indicadas, a formulação da função de aprendizado gera um custo computacional elevado por necessitar de todas as probabilidades conjuntas para o cálculo do termo  $\pi_{ij}$ . Para alguns casos onde a função real é extremamente custosa, esse aumento de custo para reduzir a quantidade final de pontos de suporte pode ser útil.

Usar todos os  $n$  pontos da amostra geram um gasto computacional desnecessário, mesmo que este gasto seja menor que a função de desempenho, utiliza-se então um critério de pré-seleção. De acordo com Zhu e Du (2016), sugere-se uma pré-seleção baseada na contribuição de  $\hat{P}_f$  e  $Var[\hat{P}_f]$  para resultados adequados. Outras funções de aprendizados para definir a pré-seleção poderiam ser utilizadas, como pro exemplo a função  $U$ . Outras variações de AKMCS poderiam também fazer uma “pós-seleção” por meio da função  $c_i$  nos candidatos mais promissores e obter possíveis melhorias.

### 3.6 AK-MCSi - método AKMCS melhorado para aplicações práticas

Segundo Lelièvre et al. (2018) o AK-MCSi é uma versão melhorada do AKMCS, usando o “i” de *improvement*. São descritas no seu trabalho três melhorias ao procedimento original. Para facilitar, as melhorias serão tratadas separadamente.

A primeira alteração é o uso de uma simulação de Monte Carlo sequencial. Por usar uma quantidade muito grande de pontos ( $10^{(N+2)}$  para problema com  $P_f \approx 10^{-N}$ ), a memória exigida para guardar todas as informações de cada ponto é elevada. Por este motivo utiliza-se uma sequencia de simulações de Monte Carlo com uma menor quantidade de pontos.

Cada grupo de amostras da sequencia de simulações utiliza  $n_p$  pontos amostrais, então é utilizado o procedimento AK-MCSi para chegar na probabilidade de falha de cada agrupamento. Após o primeiro grupo, todo o grupo subsequente utiliza o conjunto de controle para o Kriging do grupo anterior. A sequencia continua sendo executada até que o coeficiente de variação esteja adequado segundo:

$$\delta_{\hat{P}_f} = \sqrt{\frac{1 - \hat{P}_f}{N_p n_p \hat{P}_f}}, \quad (3.25)$$

sendo  $N_p$  o número de grupos até então,  $n_p$  é o número de pontos por grupo e  $\delta_{\hat{P}_f}$  é o coeficiente de variação. Este valor é comparado com um valor estipulado  $\delta_{lim}$ . É usado  $\hat{P}_f$  na equação 3.25 por ser usado o modelo de Kriging, mas esta alteração também pode ser aplicada diretamente no Monte Carlo simples.

A segunda alteração é o uso de “enriquecimento múltiplo” para o conjunto de controle. Acrescentar mais de um ponto de suporte por iteração é proposto para reduzir o custo computacional total do método, não apenas por reduzir o número de vezes que é calculado o modelo de Kriging, mas também por possibilitar o uso de multi-processamento para definir os valores da função de desempenho por meio de paralelização.

Para adicionar novos pontos de suporte de maneira adequada, usa-se uma técnica similar ao “K-mean clustering” disponível nos trabalhos de Macqueen (1967) e Mackay (2003). Resumidamente gera-se um número  $K$  de centroides, agrupam-se os pontos amostrais em  $K$  grupos segundo o diagrama de Voronoi, avalia-se o ponto médio de cada agrupamento para torná-lo o novo centroide  $K$  e repete-se o procedimento até que o resultado de cada centroide permaneça suficientemente constante.

A etapa de definir o novo valor do centroide por meio da média não é a mais adequada para uso no AK-MCSi (LELIÈVRE et al., 2018), portanto uma média ponderada

por meio do valor da função de aprendizado  $U$  é utilizada:

$$c_k = \frac{\sum_{i=1}^{n_k} \left(\frac{1}{U_i}\right)^2 \mathbf{x}_i}{\sum_{i=1}^{n_k} \left(\frac{1}{U_i}\right)^2}, \quad (3.26)$$

sendo  $n_k$  o número de pontos no agrupamento  $k$  e  $c_k$  o novo valor do centroide. Esta etapa é nomeada de “K-w-means clustering” por usar uma média ponderada (*weighted mean*). E ao final da etapa de agrupamento o ponto amostral mais próximo do centroide de cada grupo é escolhido para ser adicionado ao conjunto de controle do Kriging.

A terceira e última alteração é o critério de parada, sendo ele definido pela probabilidade dos pontos amostrais estarem corretamente estimados. O equacionamento exato seria:

$$P[Ce] = P[Ce_1 \cap Ce_2 \cap \dots \cap Ce_{n_p}], \quad (3.27)$$

sendo  $Ce_i$  o evento “o ponto amostral  $i$  está certo”,  $Ce$  é o evento no qual todos os pontos amostrais da simulação em questão estão corretos e  $n_p$  é o número de pontos amostrais de cada simulação de Monte Carlo. Como o valor exato seria difícil de determinar pela dependência entre pontos que o Kriging gera, simplifica-se a equação considerando que os pontos possuem independência por:

$$P[Ce] = \prod_{i=1}^{n_p} P[Ce_i]. \quad (3.28)$$

Assim é possível comparar com um valor de probabilidade limite  $P[Ce] \leq P_{lim}$ .

O AK-MCSi consiste então de uma sequência de simulações de Monte Carlo usando o Kriging. Cada simulação deve passar por uma etapa de aprendizado até ser aprovada pelo critério de parada  $P[Ce] \leq P_{lim}$ . Por fim calcula-se a equação 3.25 para decidir se é necessário aumentar o número de pontos amostrais total com o uso de mais uma simulação, o processo é repetido até que o  $\delta_{\hat{p}_f}$  esteja dentro do limite definido.

### 3.6.1 Efeitos da alteração

Inicialmente deve-se reforçar que o objetivo das alterações propostas por Lelièvre et al. (2018) é reduzir o custo computacional, tanto por processamento quanto por uso de memória, e este objetivo pode ser alcançado mesmo com o aumento do número de avaliações da função de desempenho por causa da possibilidade de paralelização.

Para reforçar esta afirmação, a tabela 2 apresenta a comparação de número de avaliações da função de desempenho com a alteração do número de pontos de suporte adicionados por iteração do programa disponível no trabalho de Lelièvre et al. (2018). Nela é possível verificar que há uma redução do tempo computacional apesar do aumento das avaliações da função de desempenho.

Tabela 2 – Comparação número de avaliações da função de desempenho para diferentes números de pontos de suporte adicionados por iteração

| Metodo  | $N_{add}$ por it. | aval.função de desemp | Tempo CPU[s] | N° iterações |
|---------|-------------------|-----------------------|--------------|--------------|
| AKMCS+U | 1                 | 339                   | 3006         | 320          |
| AK-MCSi | 1                 | 394                   | 902          | 381          |
| AK-MCSi | 2                 | 400                   | 486          | 197          |
| AK-MCSi | 5                 | 430                   | 246          | 89           |
| AK-MCSi | 7                 | 454                   | 224          | 69           |
| AK-MCSi | 10                | 490                   | 177          | 54           |
| AK-MCSi | 15                | 515                   | 149          | 40           |
| AK-MCSi | 20                | 520                   | 128          | 32           |
| AK-MCSi | 25                | 570                   | 143          | 29           |
| AK-MCSi | 30                | 620                   | 161          | 27           |
| AK-MCSi | 35                | 685                   | 163          | 26           |
| AK-MCSi | 40                | 700                   | 162          | 24           |
| AK-MCSi | 45                | 740                   | 160          | 23           |
| AK-MCSi | 50                | 820                   | 351          | 23           |

Fonte: Adaptado de Lelièvre et al. (2018).

Enfatiza-se que os tempos computacionais encontrados na tabela 2 são majoritariamente em função do tempo usado para solucionar os termos do Kriging. O problema em questão é a função de Rastrigin modificada:

$$G(x_1, x_2) = d - [(x_1^2 - 5\cos(2\pi x_1)) + (x_2^2 - 5\cos(2\pi x_2))], \quad (3.29)$$

onde  $d = 15$ . O tempo computacional da função de desempenho para este caso é desprezível em relação a avaliação do modelo de Kriging, ainda mais em relação a todo o procedimento de AKMCS ou AK-MCSi. A redução neste caso ocorre por ter um menor número de iterações.

Para a redução do uso de memória utiliza-se a segmentação da amostra total, pois com a divisão em diversas simulações é possível chegar no limite definido para o coeficiente de variação do Monte Carlo sem que a memória seja um problema. Por não usar apenas o Monte Carlo, os efeitos dessa segmentação devem ser verificados no modelo de Kriging. Na etapa de aprendizado das primeiras simulações existe uma menor quantidade de pontos para serem classificados pela função de aprendizado, portanto toda a aprendizagem pode piorar com a redução da amostra. Este efeito não foi analisado no trabalho de Lelièvre et al. (2018), porém é improvável que haja grandes efeitos no resultado da probabilidade de falha final encontrada, mas ele pode ser um dos fatores que gera aumento do número de chamadas da função de desempenho.

Outro efeito da segmentação da amostra de Monte Carlo é a redução do custo computacional em função da redução da ordem de grandeza das matrizes usadas. Portanto

a segmentação atinge o objetivo definido inicialmente com êxito. Porém uma quantidade muito pequena de pontos em cada simulação resulta em um aumento no custo computacional. O trabalho de Lelièvre et al. (2018) não entra em detalhes nos motivos para isto, apenas comenta sobre e mostra um gráfico de número de pontos amostrais por simulação contra tempo de execução do programa.

Quanto a esse aumento de custo por redução excessiva da amostra de cada simulação, verifica-se que ocorre em função da redução de pontos disponíveis para as etapas de aprendizado. Pode-se concluir que o aumento do número de chamadas da função de desempenho deve ser o fator principal no tempo computacional, mas para problemas de testes da literatura o fator determinante é o aumento do número de termos para a etapa de otimização do modelo de Kriging. Observa-se que o tempo computacional da etapa de otimização depende da quantidade de parâmetros a ser otimizado, usualmente essa quantidade é igual ao número de pontos no conjunto de controle.

Por fim, deve-se observar que o tempo computacional usado para a avaliação de função de desempenho não é dependente apenas do número de avaliações feitas, pois a intenção de usar a técnica do “K-mean clustering” é aproveitar da paralelização. Portanto o número ideal para “ $N_{add}$ ” seria o maior número de processos disponíveis para paralelizar a função de desempenho ou um múltiplo desse número, verificando que o número de avaliações por iteração é igual ao  $N_{add}$ . Para um computador pessoal o número estaria geralmente entre 3 e 8, enquanto uma distribuição em rede possibilitaria valores maiores. Caso o tempo de avaliação da função de desempenho seja aproximadamente igual paralelizado ou não, o uso de  $N_{add}$  resulta em um tempo computacional inferior, chegando ao limite máximo de  $Tempo_{Paralelizado} = Tempo_{NaoParalelizado} / N_{ProcessosParalelos}$  para uma mesma quantidade de avaliações da função de desempenho.



## 4 Integração de Kriging com Monte Carlo usando amostragem por importância

Verificando a literatura, pôde-se observar uma forma de Kriging com Monte Carlo distinta do AKMCS ganhando popularidade no meio científico. Com o uso do AKMCS o procedimento torna-se suscetível a erros por meio de amostragem insuficiente ou de baixa qualidade, fazendo o resultado possuir erros grandes enquanto os critérios de falha indicam o contrário. Este fato é verificado por usar valores de aproximação para o cálculo do erro, sendo impossível garantir o controle do erro (RAZAALY; CONGEDO, 2018).

Neste capítulo serão apresentados dois métodos de amostragem por importância, sendo o primeiro considerado a forma original do “Meta-IS” (*Metamodeling - Importance Sampling*), seguido do MetaAL-OIS (*Metamodel-based combining Active Learning and quasi-Optimal Importance Sampling*). Sendo o Meta-IS considerado a forma original, as características da integração com Monte Carlo por importância serão descritas posteriormente junto do método.

Apesar de existirem alguns outros modelos sobre o assunto como os trabalhos de Gaspar, Teixeira e Soares (2017), Zhao e Yue (2015), suas diferenças estão na forma de definir a amostragem inicial. Logicamente, a amostragem inicial gera um impacto considerável na resposta final, mas deve-se escolher a amostragem de acordo com o problema em questão. Caso exista mais de uma região de falha a maioria das variações de amostragem inicial perdem eficiência da forma que estão desenvolvidas. Ainda assim, pode-se usar os conceitos desenvolvidos pelos autores citados para gerar uma amostragem inicial diferente e possivelmente mais eficiente.

Para reforçar esta afirmação recomenda-se observar a amostragem indicada por Xiong e Tan (2017). É utilizada a amostragem RBIS (*Radial-Based Importance Sampling*) para gerar uma busca inicial de pontos de suporte em uma região de falha provável, usando uma etapa inicial de aprendizado para gerar o conjunto de controle inicial e depois seguindo com o modelo de “AKMCS+” desejado.

### 4.1 Meta-IS - primeiro modelo de Kriging por importância

O trabalho de Dubourg, Sudret e Deheeger (2011) desenvolve uma das primeiras formas do uso de Kriging com a amostragem por importância, onde são descritos detalhadamente os conceitos para utilizar o Kriging para selecionar a distribuição da amostra da simulação de Monte Carlo.

Pode-se calcular a probabilidade de falha por:

$$\begin{aligned}
 P_f &= \int_{\Omega} \mathbf{I}[\mathbf{x}] f_{\mathbf{x}}(\mathbf{x}) = E[\mathbf{I}(\mathbf{x})] \\
 P_{f \text{ IS}} &= \int_{\Omega} \mathbf{I}[\mathbf{x}] \frac{f_{\mathbf{x}}(\mathbf{x})}{h(\mathbf{x})} h(\mathbf{x}) d\mathbf{x} = E_h[\mathbf{I}(\mathbf{x}) \frac{f_{\mathbf{x}}(\mathbf{x})}{h(\mathbf{x})}] \\
 \hat{P}_{f \text{ IS}} &\approx \frac{1}{n} \sum_{i=1}^n \mathbf{I}(\mathbf{x}_i) \frac{f_{\mathbf{x}}(\mathbf{x}_i)}{h(\mathbf{x}_i)}.
 \end{aligned} \tag{4.1}$$

Sendo a amostra de  $[\mathbf{x}_1, \dots, \mathbf{x}_n]$  definida por meio da densidade de probabilidade  $h$ , obtendo variância:

$$\text{Var}[\hat{P}_{f \text{ IS}}] = \frac{1}{n-1} \left[ \frac{1}{n} \sum_{i=1}^n \left( \mathbf{I}(\mathbf{x}_i) \frac{f_{\mathbf{x}}(\mathbf{x}_i)}{h(\mathbf{x}_i)} \right)^2 - \hat{P}_{f \text{ IS}} \right]. \tag{4.2}$$

Portanto, após a determinação de  $h(\mathbf{x})$  é possível definir a probabilidade de falha e variância da amostra independentemente do metamodelo usado para a escolha da PDF, portanto independente do conjunto de controle usado.

Para definir a densidade de probabilidade usa-se:

$$\hat{h}(\mathbf{x}) \approx \pi(\mathbf{x}) f_{\mathbf{x}}(\mathbf{x}), \tag{4.3}$$

observando que a função  $\pi(\mathbf{x})$  tende a 1 caso  $\mathbf{x}$  esteja na região de falha e tende a 0 caso contrário. Considerando que a função  $h$  deve gerar pontos próximos do limite de falha, o termo  $f_{\mathbf{x}}(\mathbf{x})$  faz com os pontos de menor contribuição para  $P_f$  continuem com menor probabilidade de ocorrência segundo a função  $\hat{h}(\mathbf{x})$ .

Para definir a função  $\pi$  adequadamente com uma menor quantidade de pontos amostrais, utiliza-se então o Kriging. O método usado para a etapa de aprendizagem do Kriging é semelhante ao AK-MCSi descrito na seção 3.6 por usar o “K-mean clustering”. Por utilizar um método de clustering similar, as alterações realizadas por Lelièvre et al. (2018) poderiam ser aplicadas nesta etapa do Meta-IS e talvez seu desempenho fosse melhorado.

Para simplificar, o Meta-IS segue o seguinte esquema:

Passo 1 Gera-se uma amostra inicial de, por exemplo,  $10^4$  pontos.

Passo 2 Reduz-se esta amostra por meio do método de K-mean, podendo ser alterado pelo K-w-mean do AK-MCSi nas iterações seguintes, pois o “w” necessita do Kriging para ser calculado e a primeira iteração não possui pontos de suporte.

Passo 3 Avalia-se então os K pontos pela função de desempenho e acrescenta-os ao conjunto de controle do Kriging.

Passo 4 Soluciona-se o modelo de Kriging e retorna ao Passo 1 até o critério de parada ser atendido.



Passo 5 Após o aprendizado do modelo de Kriging finalizar, utiliza-se ele para calcular a função  $\hat{h}(x)$ , permitindo realizar a amostragem por importância.

Passo 6 Realiza-se a amostragem com a função de densidade de probabilidade  $\hat{h}$  e avalia-se os pontos da amostra segundo a função de desempenho  $g(x)$ .

Passo 7 Continua-se gerando pontos amostrais no Passo 6 até que o critério de parada seja atingido.

O critério de parada usado no Passo 4 é:

$$\hat{\alpha} = \frac{1}{m} \sum_{i=1}^m \frac{\mathbf{I}(\mathbf{x}_i)}{P[\hat{G}_{\mathbf{X} \setminus \mathbf{x}_i}(\mathbf{x}_i) \leq 0]}, \quad (4.4)$$

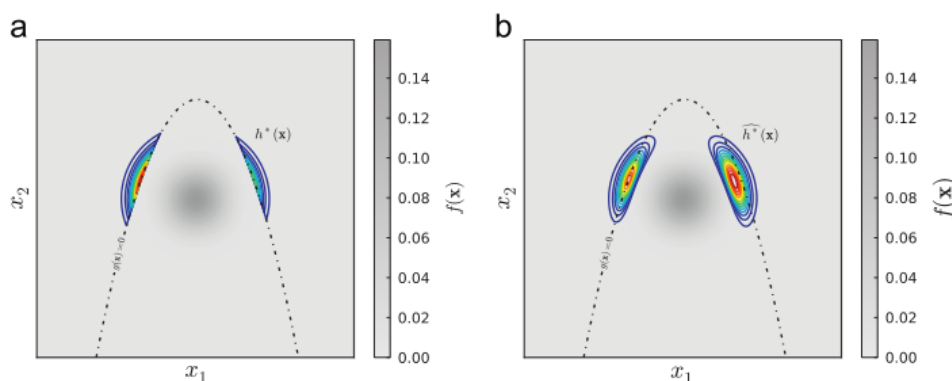
onde  $\mathbf{X} \setminus \mathbf{x}_i$  indica o uso do Kriging sem o ponto de suporte  $\mathbf{x}_i$ . Sendo  $\hat{\alpha}$  aceitável quando possui um valor próximo de 1. Segundo Dubourg, Sudret e Deheeger (2011) um valor entre 0,1 e 10 é suficiente, entretanto é necessário um número mínimo de pontos de suporte, sendo sugerido  $m \geq 30$ .

Enquanto para o critério de parada do Passo 7 usa-se um coeficiente de variação máximo baseado na variância definida na equação 4.2.

#### 4.1.1 Características do método

O Meta-IS em si possui muitas características advindas da simulação de Monte Carlo com amostragem por importância, pois o Kriging é usado “apenas” para definir a função  $h$  que define a PDF da amostra. Logicamente, o Kriging afeta o resultado final, entretanto o gráfico disponível no trabalho de Dubourg, Sudret e Deheeger (2011) demonstra a eficiência do método desenvolvido para descrever a PDF ideal.

Figura 7 – Comparação da PDF ideal(a) e da quase-ótima por meio de  $\hat{h}(b)$ .



Fonte: (DUBOURG; DEHEEGER; SUDRET, 2011)

A maior vantagem do Meta-IS sobre o AKMCS é a independência do resultado final em relação ao conjunto de controle, pois a função  $G(\mathbf{x})$  original é utilizada para avaliar a

função  $\mathbf{I}[\mathbf{x}]$ . Além disso, por possuir variância definida por meio da equação 4.2, pode-se determinar o intervalo de confiança adequado por meio do número de pontos amostrais.

Apesar desta vantagem, a maior desvantagem é o uso de um maior número de avaliações da função de desempenho para obter o resultado final. Com o uso de uma distribuição  $\hat{h}$  aproximada por meio do Kriging, a quantidade de avaliações da função de desempenho torna-se a soma do número de pontos de suporte para definir a função  $\pi$  adequadamente e a quantidade de pontos da amostra por importância. Esta desvantagem será mostrada com maior detalhamento por meio do trabalho do MetaAL-OIS a seguir.

## 4.2 MetaAL-OIS - exemplo do progresso do método

O trabalho de Razaaly e Congedo (2018) é apresentado por um motivo particular: o procedimento desenvolvido é separado em dois métodos distintos, sendo eles o MetaAL-OIS baseado no metamodelo e o MetaAL-OIS independente do modelo de Kriging, em outras palavras, o “tipo AKMCS” e o “tipo Meta-IS”. Sendo interessante mostrar a comparação direta entre Kriging com Monte Carlo simples ou por importância. Além disso é interessante o uso do trabalho de Razaaly e Congedo (2018) por ser recente e dispor de grandes referências para o tema específico de Kriging com Monte Carlo por importância.

Para compreender melhor o procedimento sugere-se ler o texto original, mas simplificadaamente pode-se definir o MetaAL-OIS pela fase de aprendizado do metamodelo e a fase da simulação com amostragem por importância. A fase de aprendizado se assemelha ao AKMCS, sendo inclusive usado este termo no pseudo-código descrito.

Para descrever o método, deve-se informar que existe um novo conjunto utilizado, sendo ele o “conjunto de base de amostragem”. Estes pontos são usados para uma etapa de amostragem por meio do método de MCMC Metropolis-Hastings modificado descrito em Razaaly e Congedo (2018). A fase de aprendizado pode ser resumida por:

- Passo 1 Os pontos de suporte iniciais são selecionados por meio do hipercubo latino (LHS) e seus resultados computados por meio da função de desempenho.
- Passo 2 Gera-se uma grande amostra independente (semelhante a simulação de Monte Carlo simples) e usa-se a função  $U$  para selecionar novos pontos de suporte até  $U \leq 2$ . Esta etapa é basicamente como o AKMCS (ECHARD; GAYTON; LEMAIRE, 2010), entretanto adiciona-se um critério de distância mínima entre os pontos de suporte e o novo ponto a ser adicionado.
- Passo 3 Retorna-se ao Passo 2 até que exista um número mínimo de pontos amostrais na região de falha. Caso o número mínimo seja atingido prossegue-se para o Passo 4. Define-se então o valor mínimo de pontos de falha como  $K$ , sendo ele a quantidade de pontos da base de amostragem, inicialmente 1.

Passo 4 Usa-se o método de K-mean (com o  $K$  do Passo 3) para encontrar o centroide do conjunto dos pontos amostrais que estão na região de falha. Estes centroides serão usados como ponto base de amostragem. Além disso, cria-se um valor auxiliar  $i = 1$ .

Passo 5 Usa-se o método modificado de MCMC Metropolis-Hastings para gerar uma amostra de  $N_{MCMC}$  pontos. Utiliza-se o ponto de base de amostragem da posição  $i$  e a função de densidade alvo:

$$\hat{h}(\mathbf{x}) = \frac{\hat{\mathbf{I}}(\mathbf{x})f_{\mathbf{x}}(\mathbf{x})}{\hat{P}_f}. \quad (4.5)$$

Passo 6 Se existir na amostra gerada por meio do MCMC um ponto que respeite a distância mínima em relação aos pontos de suporte e possua valor de  $U$  inferior a 2, o ponto é adicionado ao conjunto de controle do Kriging.

Passo 7 Aumenta-se o valor de  $i$  em 1, que então é comparado com o valor de  $K$  do Passo 3. Caso  $i$  seja menor, volta-se ao Passo 5.

Passo 8 Caso o número de pontos de suporte tenha sido alterado desde a última execução do Passo 4, retorna-se ao Passo 2. Caso contrário finaliza-se a etapa de aprendizado do MetaAL-OIS.

Nas comparações feitas no trabalho de Razaaly e Congedo (2018), para os resultados do MetaAL-OIS baseado no metamodelo é utilizado o modelo final da etapa de aprendizado do MetaAL-OIS para avaliar uma amostra de Monte Carlo simples. É possível descrever esta etapa como uma alteração ao AKMCS por meio da adição de um critério de parada adicional.

Por fim, usa-se o modelo de Kriging para gerar a amostragem por importância de maneira semelhante ao Meta-IS, porém a função  $\hat{h}$  é definida pela equação 4.5 do Passo 5.

### 4.2.1 Análise do trabalho

Neste caso é inadequado avaliar os efeitos da “alteração”, pois o método é completamente diferente quanto ao uso do Kriging. O trabalho de Razaaly e Congedo (2018) traz um problema em comum com o trabalho de Dubourg, Deheeger e Sudret (2011), sendo ele:

$$G(x_1, x_2) = \min \left\{ \begin{array}{l} c - 1 - x_2 + \exp(-x_1^2/10) + (x_1/5)^4 \\ \frac{c^2}{2} - x_1x_2 \end{array} \right\}, \quad (4.6)$$

e para  $c$  foram usados os valores 3, 4 e 5, sendo  $P_f$  igual a  $3,35 \times 10^{-3}$ ,  $8,68 \times 10^{-5}$  e  $9,48 \times 10^{-7}$ . A quantidade de avaliações da função de desempenho do MetaAL-OIS é no geral menor que do Meta-IS e a maior diferença ocorre com  $c = 5$ , onde o Meta-IS usa 40 pontos de aprendizado e 2900 de amostragem por importância, enquanto o MetaAL-OIS

usa mais pontos de aprendizado (84) e uma menor amostra (496), obtendo um total de avaliações bem inferior.

Pode-se creditar essas diferenças ao modo que o Kriging é utilizado, quando o Meta-IS foi desenvolvido haviam poucos estudos sobre o assunto e seu uso era limitado. Apesar disso, a base do Kriging com Monte Carlo por importância se manteve: estima-se o domínio de falha por meio do Kriging para gerar uma função de densidade  $h$ , por meio dessa função define-se a simulação por importância.

Apesar da dependência do conjunto de controle, utilizar o Kriging com Monte Carlo simples ou por importância geram resultados suficientemente próximos, tornando as avaliações da função de desempenho adicionais da amostragem por importância excessivas.

O modelo de Kriging no término da etapa de aprendizagem pode ser usado de maneira similar ao AKMCS, sendo usado o modelo de Kriging em uma amostra de Monte Carlo simples. Considerando o MetaAL-OIS como uma modificação para o AKMCS pode-se avaliar a inclusão da distância mínima na função  $U$  e a alteração do critério de parada com mais uma função de aprendizado secundária.

A inclusão da distância mínima gera uma eliminação de pontos de suporte “desnecessários”, pois os pontos próximos geram pouco efeito no resultado final do modelo. Entretanto “pouco efeito” para o limite da região de falha pode ser suficiente para alterar seu resultado, sendo que a diferença entre selecionar um ponto de suporte próximo do limite de falha ou exatamente no limite de falha foi demonstrada no trabalho de Peijuan et al. (2017) com o AKMCS+BiL. Entretanto a alteração do resultado por meio dessa distância mínima é pequena o bastante para não afetar a PDF  $h$  e seu uso na amostragem por importância.

O uso de uma amostra adicional resulta em um maior número de avaliações da função de desempenho sem grandes efeitos no valor final, entretanto a variância pode ser calculada de maneira simples. Observa-se também que, caso o critério de parada do AKMCS e similares fosse modificado para incluir uma quantidade de pontos de suporte similar ao número de avaliações da função de desempenho, as versões de AKMCS provavelmente teriam um resultado melhor que o Meta-IS e que o MetaAL-OIS independente do Kriging.

## 5 Resultados

Com o objetivo de expandir os resultados disponíveis e poder reforçar as conclusões das análises mostradas anteriormente, buscou-se na literatura alguns problemas de *benchmark* para confiabilidade.

Para reduzir a variação dos resultados será gerada apenas uma amostra de pontos para a simulação de Monte Carlo no caso do AKMCS e suas variações, sendo resolvidos primeiramente a simulação de Monte Carlo simples para gerar um valor de referência e então serão executadas 10 vezes as versões de Kriging com Monte Carlo estudadas. Quanto ao número de execuções foi utilizado apenas 1 execução para o Meta-IS em função do tempo que restava para finalizar o trabalho.

Os métodos usados foram seguindo os códigos e pseudocódigos dos métodos originais (AKMCS+U (ECHARD; GAYTON; LEMAIRE, 2010), AKCSM+BiL (PEIJUAN et al., 2017), AKMCS+ewd (XIAO; ZUO; GUO, 2018) e Meta-IS (DUBOURG; SUDRET; BOURINET, 2011)), sendo usado 10 pontos de suporte iniciais. Sendo que para o AKMCS+ewd foi utilizado  $\alpha = 0,8$ .

Além disso, o Meta-IS (DUBOURG; SUDRET; BOURINET, 2011) foi alterado um pouco para aproveitar os algoritmos já utilizados, sendo usado para o aprendizado inicial o AKMCS+U, sendo mantido o critério de parada  $\hat{\alpha}$  limite do Meta-IS original.

Por fim, será utilizada também uma versão de Kriging com Monte Carlo simples sem aprendizado para demonstrar a necessidade da adição adequada de pontos de suporte. Mesmo sendo apenas para comparação geral, tentou-se reduzir a variação e melhorar os resultados do método por meio de uma melhor distribuição no domínio. O procedimento Kriging+MCS pode ser descrito por:

Passo 1 Gera-se a amostra de Monte Carlo.

Passo 2 Para gerar uma distribuição com menor variância e buscando explorar todo o domínio utiliza-se o hipercubo latino (*lhs*). É criada uma distribuição por hipercubo com valores entre 0 e 1 inicialmente, então expande-se para todo o domínio da amostra por:

$$lhs \cdot (max - min) + min, \quad (5.1)$$

sendo *max* o valor máximo de cada dimensão e *min* o valor mínimo.

Passo 3 Otimiza-se os parâmetros do modelo de Kriging por meio dos pontos de suporte gerados.

Passo 4 Avalia-se toda a amostra com o Kriging e determina-se aproximadamente o valor da probabilidade de falha.

Para a demonstração dos valores, será apresentado o valor médio e desvio da mesma forma que os resultados disponíveis em Sun et al. (2017), considerando que o uso da melhor execução pode levar a demonstrações injustas na literatura pela aleatoriedade envolvida nos métodos.

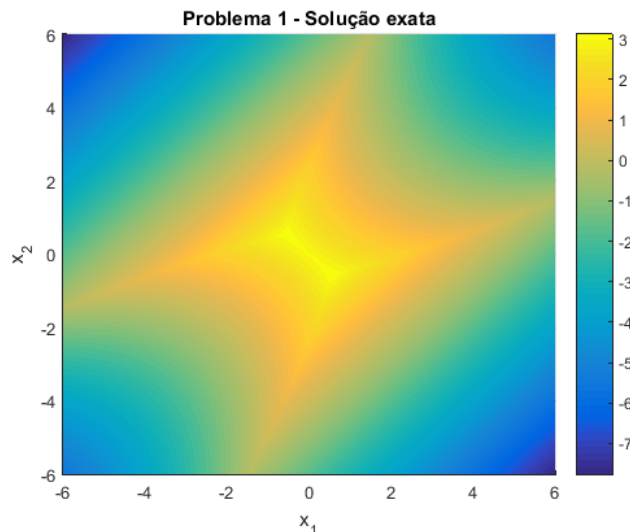
## 5.1 Problema com quatro curvas limite

O primeiro problema pode ser encontrado nos trabalhos de Echard, Gayton e Lemaire (2010), Cadini, Santos e Zio (2014), Sun et al. (2017), Gomes (2018), sendo um sistema de quatro equações onde  $x_1$  e  $x_2$  são variáveis normais padrão independentes.

$$G(\mathbf{x}) = \min \begin{cases} 3 + 0,1(x_1 - x_2)^2 + \frac{x_1+x_2}{\sqrt{2}} \\ 3 + 0,1(x_1 - x_2)^2 - \frac{x_1+x_2}{\sqrt{2}} \\ x_1 - x_2 + \frac{6}{\sqrt{2}} \\ x_2 - x_1 + \frac{6}{\sqrt{2}} \end{cases} \quad (5.2)$$

Para melhor ilustrar este problema mostra-se a figura 8 onde estão representados os valores de  $G(\mathbf{x})$  e a figura 9 demonstrando o domínio de falha em vermelho.

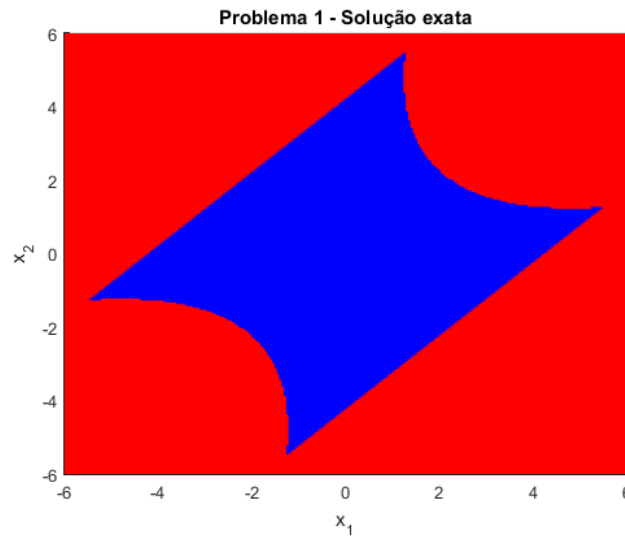
Figura 8 – Solução exata da função  $G(\mathbf{x})$  para o problema 1.



### 5.1.1 Resultados para o problema 1

Primeiramente, para este e os outros resultados será usado o mesmo formato de tabela, a primeira coluna indica o método utilizado, a segunda o número de avaliações da função de desempenho, ressaltando que para o Meta-IS usa-se “número de pontos de

Figura 9 – Visualização do domínio de falha para o problema 1.



suporte do Kriging”+“pontos de amostra por importância”. A terceira coluna contém o valor de  $P_f$  encontrado. A quarta mostra o valor de coeficiente de variação ( $\delta_{P_f}$ ) em porcentagem para 10 execuções do método, sendo o SMC e o Meta-IS exceções: para eles são usadas as equações 2.44 e 4.2. Por fim, a última coluna indica o erro em relação ao SMC usado como referência, sendo o valor usado em porcentagem para facilitar a visualização.

Tabela 3 – Resultados para o problema 1

| Metodo      | $N_{call}$      | $\hat{P}_f (\times 10^{-3})$ | $\delta_{P_f} (\%)$ | $\Delta_{P_f} (\%)$ |
|-------------|-----------------|------------------------------|---------------------|---------------------|
| SMC         | $1 \times 10^6$ | 4,43                         | 1,50                | -                   |
| Kriging+SMC | 250             | 3,60                         | 688,3               | 18,7                |
| AKMCS+U     | 145,6           | 4,44                         | 0,35                | 0,1                 |
| AKMCS+BiL   | 153,2           | 4,43                         | 0,00                | 0,0                 |
| AKMCS+ewd   | 139,1           | 4,42                         | 0,10                | 0,0                 |
| Meta-IS     | 66+268          | 4,40                         | 1,01                | 0,2                 |

Por meio da tabela 3 pode-se perceber que o Meta-IS não foi capaz de gerar um bom resultado, sendo tanto seu erro ( $\Delta_{P_f}$ ) como seu desvio maiores que os outros métodos. Porém, como o cálculo do erro usa a amostra de Monte Carlo simples como referência, os baseados no AKMCS tendem a obter erros próximos de zero. Para uma comparação justa é necessário levar a variância do SMC em consideração.

Como o coeficiente de variação do SMC foi calculado pela equação 2.44, todos os métodos que utilizam o Monte Carlo simples herdam essa variância por usarem a mesma amostragem. Considerando isso, a variação do Meta-IS é menor que a do SMC, mas a quantidade de avaliações da função de desempenho ainda é superior. É possível definir a priori qual o valor de variação que o Meta-IS terá, mas deve-se lembrar que o tamanho

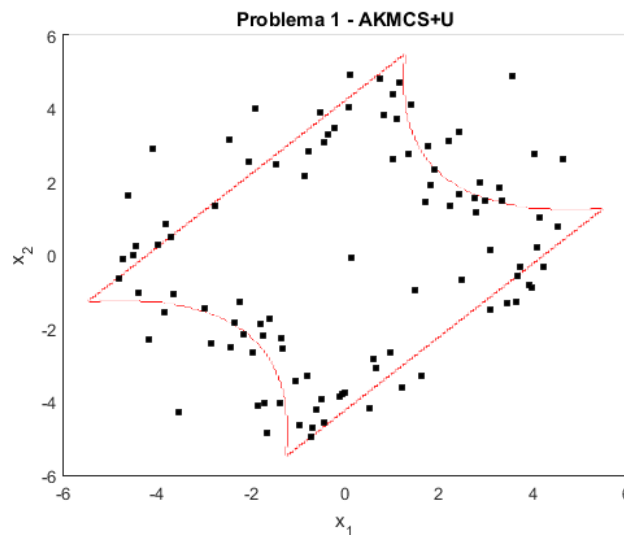
da amostra final aumenta quanto menor a variação escolhida. Para todos os exemplos escolheu-se usar uma variância proporcional ao do SMC de referência.

Por outro lado, para melhorar o coeficiente de variação dos métodos de Monte Carlo simples, basta utilizar uma amostra maior. Esta amostra maior não apenas serve para ampliar o denominador na equação 2.44, como também proporciona um domínio de pontos maior para as funções de aprendizado explorarem.

Com isso pode-se concluir que, enquanto o problema for tempo computacional os métodos de Kriging com Monte Carlo simples são recomendados, mas para os casos onde a memória é um problema os métodos de Kriging com amostragem por importância passam a ser mais interessantes. Entretanto, para solucionar o problema de memória é possível utilizar também o AK-MCSi.

É possível verificar na figura 10 os pontos de suporte para o AKMCS+U, entretanto não serão mostrados os gráficos para a função aproximada  $\hat{G}(\mathbf{x})$  por serem praticamente iguais ao da função de desempenho da figura 8. Além disso, os gráficos para os outros métodos baseados no AKMCS não são mostrados por serem muito semelhantes visualmente.

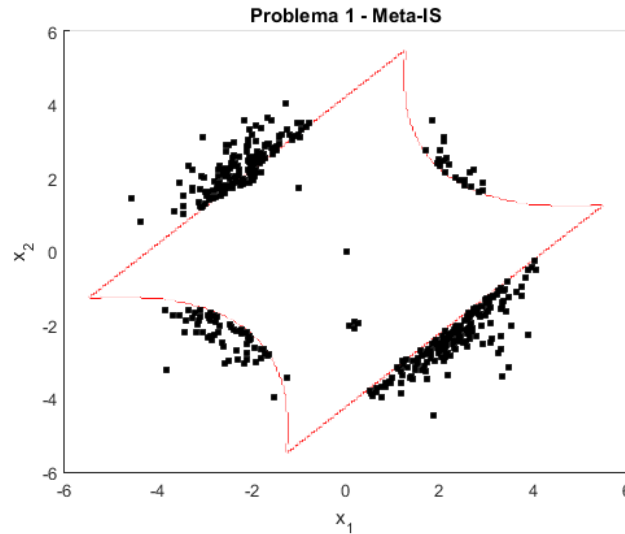
Figura 10 – Pontos de suporte usados pelo AKMCS+U no problema 1.



Na figura 11 pode-se observar que a amostragem dos pontos por importância nas regiões de maior probabilidade de ocorrência para cada sub-região. Portanto pode-se concluir que a inclusão do Kriging para o Monte Carlo por importância gera bons resultados e cumpre seu objetivo como um aproximador da função de densidade de probabilidade ideal ( $\hat{h}(\mathbf{x})$ ).



Figura 11 – Pontos de suporte mais pontos de amostragem por importância do Meta-IS.



## 5.2 Problema com oscilação - Rastrigin Function

O segundo problema é conhecido na literatura pelo nome *modified Rastrigin function* e pode ser encontrado nos trabalhos de Echard, Gayton e Lemaire (2010), Sun et al. (2017), Razaaly e Congedo (2018), Lilièvre et al. (2018), Xiao, Zuo e Guo (2018). O objetivo de usar esta equação é para verificar o desempenho na solução de um problema não convexo e com diversas regiões de falha. As variáveis possuem distribuição normal padrão e são independentes.

$$G(x) = 10 - (x_1^2 - 5\cos(2\pi x_1)) - (x_2^2 - 5\cos(2\pi x_2)). \quad (5.3)$$

Sua representação gráfica pode ser verificada na figura 12 e seu domínio de falha na figura 13. Com estes gráficos observa-se que os limites de falha são extensos e desconexos, prevendo desde o começo a maior dificuldade para aproximar estes resultados que oscilam em torno do valor  $G(\mathbf{x}) = 0$ .

### 5.2.1 Resultados para o problema 2

Neste problema pode-se observar os efeitos das oscilações geradas pelas funções cosseno afetando a qualidade e quantidade de pontos de suporte necessários (Tabela 4), mesmo assim os métodos são capazes de gerar resultados adequados por meio de uma maior quantidade de avaliações da função de desempenho em relação a funções mais simples. Para este caso pode-se verificar também uma possível vantagem do Meta-IS, pois o valor limite de 500 avaliações da função de desempenho observado no AKMCS+ewd foi utilizado para evitar problemas na execução dos métodos.

Para um número maior que 400 pontos de suporte, a otimização e utilização do modelo de Kriging começa a ficar lenta, sendo que para números maiores que 500 ocorreram

Figura 12 – Solução exata da função  $G(\mathbf{x})$  para o problema 2.

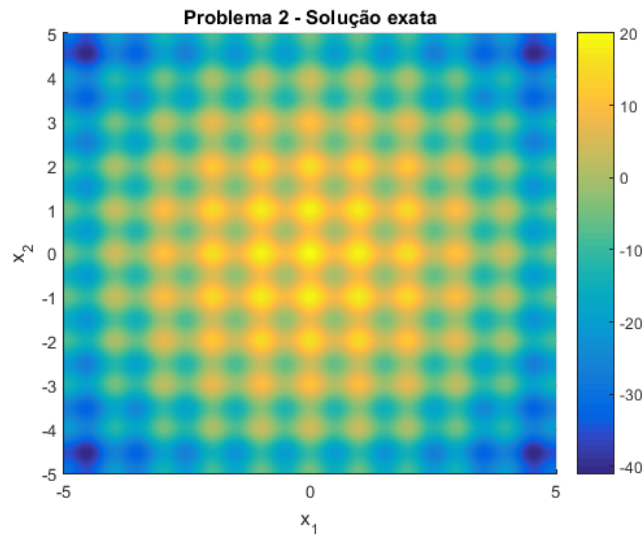
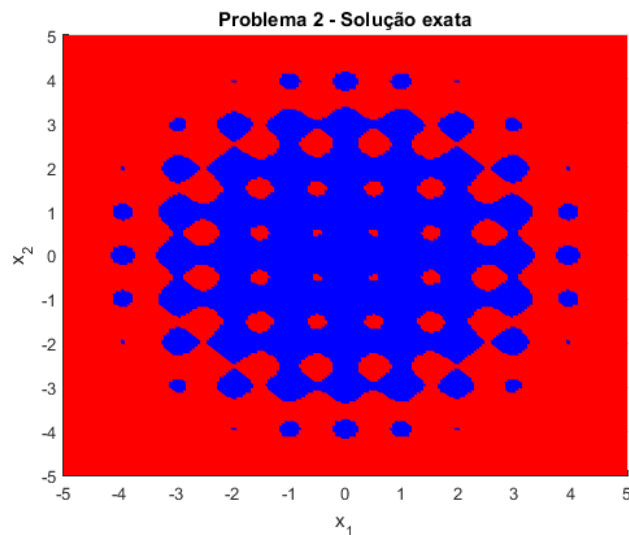


Figura 13 – Visualização do domínio de falha para o problema 2.



problemas de congelamento do processo e fechamento automático. Foi utilizado o Matlab em um computador com 16GB de RAM e processador Intel i5-6400. Observou-se também que parte do problema pode ser em função da temperatura do processador, pois foi possível solucionar um modelo de Kriging com 750 pontos de suporte quando utilizado o comando direto, mas como os métodos aumentam o conjunto de controle gradualmente utilizou-se o limite de 500 para evitar problemas.

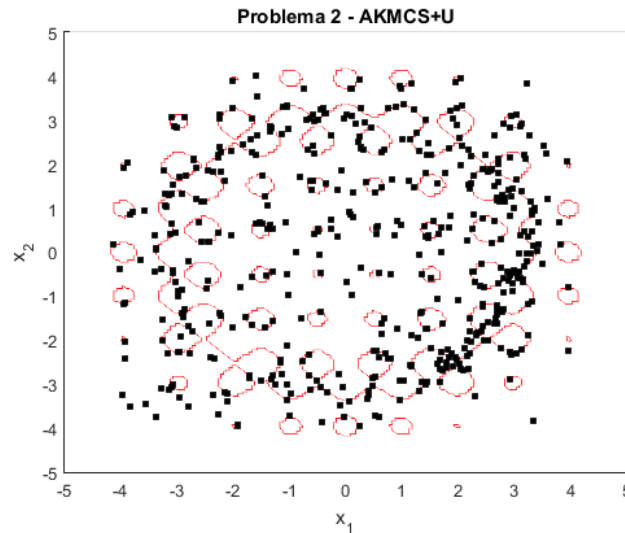
A distribuição dos pontos de suporte para os AKMCS+U pode ser verificada na figura 14, sendo que os outros métodos não são representados por terem poucas diferenças visuais. Com este gráfico pode-se verificar a tendência dos pontos se aglomerarem em valores próximos do limite de falha, apesar da necessidade de uma grande quantidade de pontos para representar corretamente cada trecho do limite representado no gráfico pelas linhas vermelhas.

Tabela 4 – Resultados para o problema 2

| Metodo                 | $N_{call}$      | $\hat{P}_f(\times 10^{-2})$ | $\delta_{Pf}(\%)$ | $\Delta_{Pf}(\%)$ |
|------------------------|-----------------|-----------------------------|-------------------|-------------------|
| SMC                    | $1 \times 10^6$ | 7,34                        | 0,37              | -                 |
| Kriging+SMC            | 250             | 1,47                        | 691,4             | 80,0              |
| Kriging+SMC            | 500             | 2,23                        | 508,7             | 69,6              |
| AKMCS+U                | 419,7           | 7,34                        | 0,60              | 0,1               |
| AKMCS+BiL              | 422,5           | 7,32                        | 0,52              | 0,3               |
| AKMCS+ewd <sup>1</sup> | 483,0           | 7,46                        | 0,88              | 1,7               |
| Meta-IS                | 298+922         | 7,24                        | 0,31              | 1,4               |

1- Foram desconsiderados dois casos no qual  $\hat{P}_f$  estava com erro elevado, pois para estes casos o critério de parada foi número máximo de pontos de suporte.

Figura 14 – Pontos de suporte usados pelo AKMCS+U no problema 2.



### 5.3 Problema com equação de terceiro grau

Este problema pode ser encontrado em Wang e Grandhi (1996), Kaymaz (2005), Peijuan et al. (2017), consiste uma equação polinomial de terceiro grau utilizando as variáveis aleatórias normais  $x_1$  de média 10 e desvio-padrão 5 e  $x_2$  de média 9,9 e desvio 5.

$$G(x) = x_1^3 + x_1^2 x_2 + x_2^3 - 18. \quad (5.4)$$

Observa-se visualmente na figura 15 e na figura 16 os valores da função e o domínio de falha. Para este problema a média das variáveis é 10, porém como o domínio de falha existe apenas em uma região deslocou-se o gráfico para melhor representar o limite de falha.

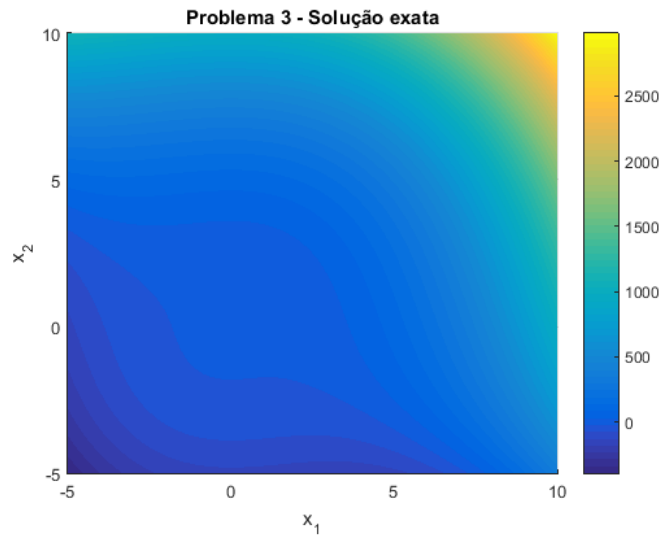
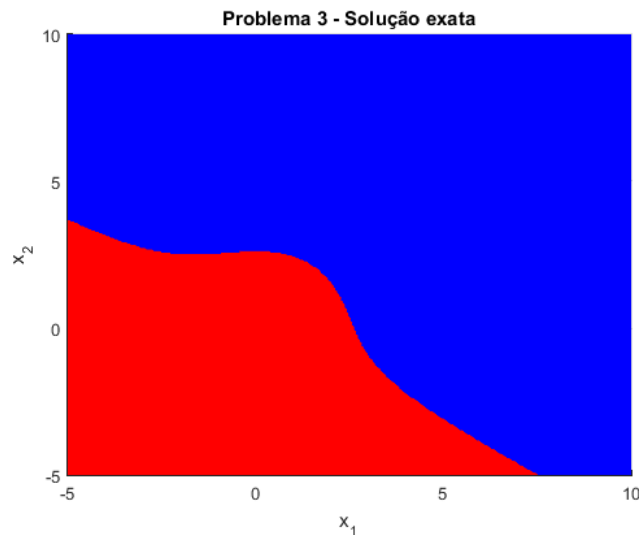
Figura 15 – Solução exata da função  $G(\mathbf{x})$  para o problema 3.

Figura 16 – Visualização do domínio de falha para o problema 3.



### 5.3.1 Resultados para o problema 3

Neste caso pode-se observar a solução de um problema simples onde todos os métodos obtiveram um bom desempenho. Neste caso a quantidade mínima de 30 pontos de suporte para o modelo de Kriging do Meta-IS seria desnecessária, pois com uma quantidade inferior de pontos é possível descrever o problema.

Observa-se que este é um problema bem mais simples, pois até o Kriging sem aprendizado foi capaz de gerar uma média adequada. Apesar da média adequada, o coeficiente de variação obtido entre as 10 execuções permite verificar que o Kriging sem aprendizado não é confiável, pois a dispersão dos resultados entre uma execução e outra é muito grande.

Com este caso é possível verificar a necessidade de utilizar informações estatísticas

Tabela 5 – Resultados para o problema 3

| Metodo      | $N_{call}$      | $\hat{P}_f(\times 10^{-3})$ | $\delta_{Pf}(\%)$ | $\Delta_{Pf}(\%)$ |
|-------------|-----------------|-----------------------------|-------------------|-------------------|
| SMC         | $2 \times 10^5$ | 5,87                        | 2,92              | 0                 |
| Kriging+SMC | 250             | 5,86                        | 72,34             | 0,2               |
| AKMCS+U     | 31,0            | 5,87                        | 0,00              | 0,0               |
| AKMCS+BiL   | 23,8            | 5,87                        | 0,00              | 0,0               |
| AKMCS+ewd   | 26,0            | 5,87                        | 0,00              | 0,0               |
| Meta-IS     | 30+100          | 5,87                        | 2,48              | 0,0               |

como média e desvio para comparações justas, porque mesmo o Kriging sem aprendizado foi capaz de gerar um erro aproximadamente nulo apenas usando a média como referencia.

## 5.4 Problema com grande número de variáveis

O quarto problema encontra-se em Kaymaz (2005), Echard, Gayton e Lemaire (2010), Dubourg, Deheeger e Sudret (2011), Gomes (2018). Este problema consiste em função de performance analítica onde o número  $n$  de variáveis pode ser alterado sem modificar significativamente o nível da probabilidade de falha (ECHARDE; GAYTON; LEMAIRE, 2010), sendo adotado o valor  $n = 40$ .

$$G(x, n) = (n + 3\sigma\sqrt{n}) - \sum_{i=1}^n x_i. \quad (5.5)$$

As variáveis possuem distribuição lognormal com média unitária e desvio-padrão  $\sigma$ , sendo  $\sigma = 0,2$  neste caso.

Considerando a quantidade de variáveis envolvidas não é possível demonstrar graficamente os resultados.

### 5.4.1 Resultados para o problema 4

Neste problema é possível verificar na tabela 6 que o AKMCS+BiL não foi o responsável pelo melhor resultado entre os AKMCS's, possivelmente com o aumento do número de dimensões houve uma dificuldade de determinar a localização do limite de falha por meio da seleção de pontos de sinais opostos, porém por causa do número elevado de dimensões do problema é difícil avaliar qual o real motivo.

Além disso, a função de aprendizado usada no AKMCS+ewd demonstrou o melhor desempenho. O motivo mais provável é o uso de uma função mais complexa e específica com o objetivo de melhorar o aprendizado, resultando numa vantagem neste caso com um maior número de dimensões no espaço amostral. A princípio pode-se considerar que o

Tabela 6 – Resultados para o problema 4

| Metodo      | $N_{call}$      | $\hat{P}_f(\times 10^{-3})$ | $\delta_{Pf}(\%)$ | $\Delta_{Pf}(\%)$ |
|-------------|-----------------|-----------------------------|-------------------|-------------------|
| SMC         | $3 \times 10^5$ | 1,81                        | 4,29              | 0                 |
| Kriging+SMC | 250             | 0,12                        | 3887              | 93,4              |
| AKMCS+U     | 118,1           | 1,81                        | 0,00              | 0,0               |
| AKMCS+BiL   | 129,6           | 1,81                        | 0,00              | 0,0               |
| AKMCS+ewd   | 115,5           | 1,81                        | 0,00              | 0,0               |
| Meta-IS     | 80+468          | 1,82                        | 3,92              | 0,3               |

AKMCS+ewd deve ser melhor que os outros métodos testados para problemas com grande número de variáveis.

## 6 Conclusão

Este trabalho demonstrou alguns dos diversos modelos de Kriging com Monte Carlo existentes apresentados na literatura. Outros métodos foram estudados e não incluídos por menor relevância para a compreensão do estado da arte no tema. Observou-se que a maioria dos autores usa formas distintas para a função de aprendizado, sendo necessário estudar cada uma delas separadamente para sua análise. A existência de uma classificação de funções de aprendizado em funções focadas em exploração ou refinamento poderia ser útil, entretanto para uma função ser eficiente deve ser capaz de desempenhar bem as duas características.

Espera-se que a descrição do modelo de Kriging deste trabalho seja suficiente para futuras análises, sendo demonstrado não apenas a função de correlação mais utilizada, como também o equacionamento sem definir seu uso a priori. Além disso, com a inclusão de fontes relevantes no assunto buscou-se facilitar o aprofundamento dos conhecimentos envolvidos. Sugere-se ainda que para introduções mais detalhadas ao modelo de Kriging busque-se literaturas na área de geoestatística e mineração, pois são as áreas na qual o modelo de Kriging foi introduzido.

Quanto ao uso do Kriging para descrever a função de amostragem na simulação de Monte Carlo, demonstrou-se o procedimento original do AKMCS juntamente com sua caracterização e análise garantindo ao leitor a possibilidade de compreendê-lo para uso em aplicações ou estudos futuros.

Por meio da exposição dos trabalhos escolhidos, verifica-se que o uso de Kriging está se consolidando para solucionar problemas de confiabilidade por meio da integração com Monte Carlo. Métodos que não aproveitam de todo o tratamento estatístico que o Kriging possibilita ainda são maioria na literatura, entretanto pode-se mostrar principalmente por Zhu e Du (2016) que estudos para melhor utilização das características do Kriging podem ser encontrados.

Além do uso da estatística, existe na literatura a inclusão de métodos híbridos e tentativas de utilização de modelos de Kriging mais complexos, sendo o maior dos exemplos atuais disto o PC-Kriging de Schobi, Sudret e Marelli ???. Espera-se que métodos ainda mais complexos por meio de equacionamentos estatísticos detalhados e procedimento mais robustos sejam estudados, possivelmente garantindo a convergência com uma pequena quantidade de avaliações da função de desempenho. A princípio não se pode garantir a existência de convergência exista sem a caracterização do problema, ou seja para problemas *black box*), entretanto a abordagem para soluções de engenharia possibilitaria a garantia de convergência para um grupo de problemas com características definidas.

Conclui-se também que, apesar da maior complexidade do Meta-IS e outros métodos de Kriging e Monte Carlo por meio de amostragem por importância, sua vantagem é um tanto quanto limitada. A possibilidade de fazer uma análise estatística detalhada onde existe independência em relação aos pontos de suporte pode ser necessária. No entanto, para o resultado final é mais indicado o uso do AKMCS. Os resultados da literatura ainda indicam que o Meta-IS possui um menor desvio-padrão, porém com uma maior quantidade de avaliações da função de desempenho, sendo difícil de comprovar que o AKMCS com a mesma quantidade de avaliações resultaria num desvio semelhante.

## 6.1 Trabalhos Futuros

Para pesquisas futuras ainda existe uma grande disponibilidade de assuntos quanto ao uso de Kriging com Monte Carlo. Para iniciar, pode-se analisar com um maior detalhamento as formas mais adequadas de seleção dos pontos de suporte iniciais, pois como comentado anteriormente, a maioria dos autores sugere de 4 a 20 sem definir maiores detalhes.

Outra possibilidade é a inclusão de outras funções de aprendizado no uso do Meta-IS, porém assim como o MetaAL-OIS, já existem alguns trabalhos nesta área. Ainda assim, seria possível fazer uma análise das funções de aprendizado para verificar se as funções que funcionam para o AKMCS funcionam para o Meta-IS ou vice-versa.

Pode-se pesquisar também as vantagens do uso de funções de aprendizado de forma sequencial, utilizando funções que enfatizam a exploração de todo o domínio inicialmente e para as etapas finais funções que buscam o maior refino do limite de falha. Para funções de exploração sugere-se testes com distância entre pontos ou a inclusão delas como no AKMCS+ewd, enquanto para refino usaria-se funções como a dp AKMCS+BiL. Observando que casos com apenas uma região de falha provavelmente não se beneficiariam com uma etapa exploratória de maneira significativa, podendo inclusive ser um uso ineficiente das avaliações da função de desempenho.

Além disso observou-se uma pequena quantidade de estudos nos quais a função de correlação do modelo de Kriging é alterada. É possível estudar novas correlações buscando características que a função de correlação exponencial usada atualmente não é capaz de representar satisfatoriamente. Apesar de observar na literatura que a correlação usada é suficiente para aproximar a região de falha em uma extensa variedade de problemas, pode-se pesquisar funções de correlação que obtêm um resultado mais eficiente utilizando um menor número de avaliações da função de desempenho.

Por fim, pode-se buscar dar um tratamento estatístico mais detalhado ao procedimento e aplicá-lo na função de aprendizado como apresentado por Zhu e Du (2016), sendo possível fazer um tratamento similar para o Meta-IS. Espera-se que o Meta-IS utilize um



---

tratamento similar ao Zhu e Du (2016) por usar o Kriging para descrever a função  $\pi$ .



# Referências

- ANDERSON, H. L. Metropolis, monte carlo, and the maniac. *Los Alamos Science*, v. 14, n. 96, p. 108–1986, 1986. Citado na página 28.
- ANDERSON, T. W. *The Statistical Analysis of Time Series*. [S.l.]: Wiley, 1971. Citado na página 32.
- ANG, A. H.-S.; TANG, W. H. Probability concepts in engineering: emphasis on applications in civil & environmental engineering. Wiley New York, v. 1, 2007. Citado 2 vezes nas páginas 20 e 21.
- ARORA, J. S. *Introduction to Optimum Design*. [S.l.]: Elsevier, 2004. Citado na página 36.
- BECK, A. T. *Curso de Confiabilidade Estrutural*. São Paulo: Escola de engenharia de São Carlos, 2014. Citado 4 vezes nas páginas 19, 27, 30 e 31.
- BERTSEKAS, D. P. *Nonlinear Programming*. [S.l.]: Athena Scientific, 1999. Citado na página 36.
- BICHON, B. J. et al. *Efficient global reliability analysis for nonlinear implicit performance functions*. 2008. 2459–2468 p. Citado na página 44.
- BIROLINI, A. *Reliability engineering: theory and practice*. [S.l.]: Springer, 2017. Citado na página 26.
- BOX, G. E. P.; JENKINS, G. M. *Time Series Analysis forecasting and control*. [S.l.]: Holden-Day, 1976. Citado na página 32.
- BUSSAB, W. de O.; MORETTIN, P. A. *Estatística Básica*. [S.l.]: Saraiva, 2012. Citado na página 23.
- CADINI, F.; SANTOS, F.; ZIO, E. An improved adaptive kriging-based importance technique for sampling multiple failure regions of low probability. *Reliability Engineering & System Safety*, Elsevier, v. 131, p. 109–117, 2014. Citado na página 68.
- CHATFIELD, C. *The analysis of time series: an introduction*. [S.l.]: Chapman & Hall/CRC, 2003. Citado na página 32.
- CHAUVET, P.; GALLI, A. *Universal Kriging*. Nancy, França: Centre de géostatistique et morphologie mathématique, 1982. Citado na página 31.
- CHIÈS, J.-P.; DELFINER, P. *Geostatistics: Modeling Spatial Uncertainty*. [S.l.]: John Wiley & Sons, 2012. Citado na página 32.
- CHOI, S.; GRANDHI, R.; CANFIELD, R. *Reliability-based Structural Design*. EUA: Springer Science & Business Media, 2007. Citado na página 28.
- CLARKE, A. B.; DISNEY, R. L. *Probabilidade e processos estocásticos*. [S.l.]: LTC, 1979. Citado na página 19.

- COX, D. D.; JOHN, S. *A statistical method for global optimization*. 1992. 1241–1246 p. Citado na página 45.
- CRESSIE, N. A. *Statistics for spatial data: revised edition*. New York, EUA: Wiley-Interscience, 1993. Citado na página 31.
- DANO, S. *Nonlinear and Dynamic Programming*. [S.l.]: Springer-Verlag, 1975. Citado na página 36.
- DEGROOT, M. H.; SCHERVISH, M. J. *Probability and statistics*. [S.l.]: Pearson Education, 2012. Citado na página 19.
- DEVORE, J. L. *Probability and statistics for engineering and the sciences*. [S.l.]: Cengage learning, 2011. Citado na página 22.
- DITLEVSEN, O.; MADSEN, H. O. *Structural Reliability Methods*. [S.l.]: John Wiley & Sons, 2007. Citado na página 29.
- DUBOURG, V.; DEHEEGER, W.; SUDRET, B. Metamodel-based importance sampling for structural reliability analysis. *Probabilistic Engineering Mechanics*, EUA, 2011. Citado 5 vezes nas páginas 41, 61, 63, 65 e 75.
- DUBOURG, V.; SUDRET, B.; BOURINET, J.-M. Reliability-based design optimization using kriging surrogates and subset simulation. *Structural and Multidisciplinary Optimization*, v. 44, n. 5, p. 673–690, Nov 2011. ISSN 1615-1488. Disponível em: <<https://doi.org/10.1007/s00158-011-0653-8>>. Citado 2 vezes nas páginas 48 e 67.
- EBELING, C. E. *An introduction to reliability and maintainability engineering*. [S.l.]: McGraw-Hill, 1997. Citado na página 27.
- ECHARD, B.; GAYTON, N.; LEMAIRE, M. *AK-MCS: An Active Learning Reliability Method Combining Kriging and Monte Carlo Simulation*. França: Clermont Université, Institut Français de Mécanique Avancée, 2010. Citado 10 vezes nas páginas 12, 41, 44, 46, 50, 64, 67, 68, 71 e 75.
- FAZIO, V. S.; ROISENBERG, M. *Interpolação espacial: uma comparação analítica entre krigagem e redes RBF*. Florianópolis, SC: Dissertação de mestrado de ciências da computação da universidade federal de Santa Catarina, 2013. Citado na página 31.
- FORRESTER, A.; SÓBESTER, A.; KEANE, A. *Engineering Design via Surrogate Modelling a Practical Guide*. UK: Wiley, 2008. Citado na página 38.
- GASPAR, B.; TEIXEIRA, A.; SOARES, C. G. Adaptive surrogate model with active refinement combining kriging and a trust region method. *Reliability Engineering & System Safety*, Elsevier, v. 165, p. 277–291, 2017. Citado na página 61.
- GATTI, P. L. *Probability Theory and mathematical statistics for engineers*. [S.l.]: Spon Press, 2005. Citado na página 24.
- GOMES, W. J. *Structural reliability analysis using artificial neural networks and bootstrap techniques*. Florianópolis, SC: ICVRAM-ISUMA uncertainties conference, 2018. Citado 2 vezes nas páginas 68 e 75.

GRIVA, I.; NASH, S. G.; SOFER, A. *Linear and Nonlinear Optimization*. [S.l.]: Society for Industrial and Applied Mathematics, 2009. Citado na página 36.

ISAAKS, E. H.; SRIVASTAVA, R. M. *An introduction to applied geostatistics*. [S.l.]: Oxford University Press, 1989. Citado na página 32.

JIAN, W. et al. Two accuracy measures of the kriging model for structural reliability analysis. *Reliability Engineering & System Safety*, Elsevier, v. 167, p. 494–505, 2017. Citado 3 vezes nas páginas 50, 51 e 52.

JOURNEL, A.; HUIJBREGTS, C. *Mining Geostatistics*. New Jersey: The Blackburn Press, 1978. Citado 2 vezes nas páginas 33 e 45.

JOYCE, D. *Conditional Distributions*. [S.l.]: Clark University, 2014. Citado na página 23.

KAPUR, K. C.; LAMBERSON, L. R. *Reliability in engineering design*. [S.l.]: John Wiley & Sons, 1977. Citado na página 27.

KAYMAZ, I. *Application of kriging method to structural reliability problems*. [S.l.]: Elsevier, 2005. 133–151 p. Citado 2 vezes nas páginas 73 e 75.

KINGSBURY, N. *Joint and Conditional cdfs and pdfs*. Connexions, 2005. Disponível em: <<http://cnx.org/content/m10986/2.8/>>. Citado na página 23.

KRISHNAIAH, P.; RAO, C. *Handbook of Statistics*. [S.l.]: Elsevier, 1988. Citado na página 20.

KROESE, D. P.; TAIMRE, T.; BOTEV, Z. I. *Handbook of Monte Carlo Methods*. [S.l.]: John Wiley & Sons, 2011. Citado na página 30.

LARSON, H. J. *Introduction to probability theory and statistical inference*. [S.l.]: John Wiley & Sons, 1982. Citado na página 26.

LELIÈVRE, N. et al. *AK-MCSi: A Kriging-based method to deal with small failure probabilities and time-consuming models*. 2018. 1 - 11 p. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0167473017301558>>. Citado 6 vezes nas páginas 56, 57, 58, 59, 62 e 71.

LEMAIRE, M. *Structural Reliability*. [S.l.]: John Wiley & Sons, 2009. Citado na página 29.

LUENBERGER, D. G.; YE, Y. *Linear and Nonlinear Programming*. [S.l.]: Springer, 2016. Citado na página 36.

MACKAY, D. J. C. *Information Theory, Inference and Learning Algorithms*. [S.l.]: University of Cambridge, 2003. Citado na página 56.

MACQUEEN, J. Some methods for classification and analysis of multivariate observations. In: OAKLAND, CA, USA. *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*. [S.l.], 1967. v. 1, n. 14, p. 281–297. Citado na página 56.

MELCHERS, R. E.; BECK, A. T. *Structural Reliability Analysis and Prediction*. [S.l.]: John Wiley & Sons, 2018. Citado na página 30.

- METROPOLIS, N. *The beginning of the Monte Carlo method*. [S.l.]: Los Alamos Science, 1987. Citado na página 28.
- METROPOLIS, N.; ULAM, S. The monte carlo method. *Journal of the American statistical association*, Taylor & Francis, v. 44, n. 247, p. 335–341, 1949. Citado na página 28.
- MICHELL, T. J.; MORRIS, M. D. *The spatial correlation function approach to response surface estimation*. EUA: Proceedings of Winter Simulation Conference, Ed. 24, 1992. Citado na página 41.
- MOHAMMADI, M.; EGHTEHAD, M.; MOHAMMADI, H. Stochastic analysis of pull-in instability of geometrically nonlinear size-dependent fgm micro beams with random material properties. *Composite Structures*, Elsevier, 2018. Citado na página 23.
- MONTGOMERY, D. C.; RUNGER, G. C.; HUBELE, N. F. *Estatística aplicada e probabilidade para engenheiros*. Rio de Janeiro: LTC, 2003. Citado na página 19.
- NORONHA, G. D. B.; LOPEZ, R. H. *Modelo de krigagem para confiabilidade*. Florianópolis, SC: Trabalho de conclusão de curso de engenharia civil da universidade federal de Santa Catarina, 2016. Citado na página 47.
- NOWAK, A. S.; COLLINS, K. R. *Reliability of structures*. [S.l.]: McGraw-Hill, 2000. Citado 2 vezes nas páginas 22 e 23.
- OLEA, R. A. *Geostatistics for engineers and earth scientists*. Kansas, EUA: Kluwer Academic Publishers, 1999. Citado 2 vezes nas páginas 31 e 45.
- OLIVER, M. A.; WEBSTER, R. *Basic steps in geostatistics: the variogram and kriging*. New York, EUA: Springer, 2015. Citado 2 vezes nas páginas 31 e 33.
- PAPOULIS, A. *Probability, Random variables, and stochastic process*. [S.l.]: McGraw-Hill, 1991. Citado na página 30.
- PAPOULIS, A.; PILLAI, S. U. *Probability, random variables, and stochastic processes*. EUA: McGraw-Hill, 2002. Citado na página 19.
- PEIJUAN, Z. et al. A new active learning method based on the learning function  $u$  of the ak-mcs reliability analysis method. *Engineering Structures*, 2017. Citado 4 vezes nas páginas 47, 66, 67 e 73.
- PICHENY, V. *Improving accuracy and compensating for uncertainty in surrogate modeling*. Tese de pós doutorado — Universidade da Florida, EUA, 2009. Citado 2 vezes nas páginas 12 e 44.
- PINHEIRO, J. I. D. et al. *Probabilidade e estatística: quantificando a incerteza*. [S.l.]: Elsevier Brasil, 2012. Citado na página 26.
- RAZAALY, N.; CONGEDO, P. M. *Novel algorithm using Active Metamodel Learning and Importance Sampling: application to multiple failure regions of low probability*. [S.l.]: Elsevier, 2018. 92–114 p. Citado 4 vezes nas páginas 61, 64, 65 e 71.
- REDDY, T. A. *Applied data analysis and modeling for energy engineers and scientists*. New York: Springer, 2011. Citado na página 19.

- ROSS, S. M. *Introduction to Probability Models*. [S.l.]: Elsevier, 2014. Citado na página 21.
- RYAN, T. P. *Modern Regression Methods*. [S.l.]: John Wiley & Sons, 2009. Citado na página 31.
- SCHOBI, R. et al. *Combining Polynomial Chaos Expansion and Kriging*. ETH Zurich, Switzerland: Orange Labs research, 2014. Nenhuma citação no texto.
- SILVA, P. A. L. *Probabilidades & estatística*. Rio de Janeiro: Reichmann & Affonso, 1999. Citado na página 19.
- SOBOL, I. M. *A primer for the Monte Carlo Method*. [S.l.]: CRC Pres, 1994. Citado na página 28.
- SUDRET, B. *Uncertainty propagation and sensitivity analysis in mechanical models Contributions to structural reliability and stochastic spectral methods*. França: Universidade Blaise Pascal, 2007. Citado 3 vezes nas páginas 23, 29 e 30.
- SUN, Z. et al. Lif: A new kriging based learning function and its application to structural reliability analysis. *Reliability Engineering& System Safety*, 2017. Citado 3 vezes nas páginas 52, 68 e 71.
- URBANO, J. ao. *Estatística - uma nova abordagem*. [S.l.]: Ciência Moderna Ltda, 2010. Citado na página 22.
- USHAKOV, N. G. *Selected Topics in Characteristic Functions*. [S.l.]: VSP, 1999. Citado na página 21.
- VENKATRAM, A. On the use of kriging in the spatial analysis of acid precipitation data. *Atmospheric Environment (1967)*, Elsevier, v. 22, n. 9, p. 1963–1975, 1988. Citado na página 31.
- WACKERNAGEL, H. *Multivariate Geostatistics: An Introduction with Applications*. [S.l.]: Springer-Verlag, 2003. Citado na página 33.
- WANG, L.; GRANDHI, R. V. *Safety index calculation using intervening variables for structural reliability analysis*. [S.l.]: Elsevier, 1996. 1139–1148 p. Citado na página 73.
- XIAO, N.-C.; ZUO, M. J.; GUO, W. Efficient reliability analysis based on adaptive sequential sampling design and cross-validation. *Applied Mathematical Modelling*, v. 58, p. 404 – 420, 2018. ISSN 0307-904X. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0307904X18300829>>. Citado 4 vezes nas páginas 48, 49, 67 e 71.
- XIONG, B.; TAN, H. *A robust and efficient structural reliability method combining radial-based importance sampling and Kriging*. [S.l.]: Springer, 2017. 1–11 p. Citado na página 61.
- ZHAO, H. et al. An efficient reliability method combining adaptive importance sampling and kriging metamodel. *Applied Mathematical Modelling*, Elsevier, v. 39, n. 7, p. 1853–1866, 2015. Citado na página 61.

ZHU, Z.; DU, X. Reliability analysis with monte carlo simulation and dependent kriging predictions. *Journal of Mechanical Design*, v. 138, 2016. ISSN 0307-904X. Disponível em: <<http://dx.doi.org/10.1115/1.4034219>>. Citado 5 vezes nas páginas 54, 55, 77, 78 e 79.