



UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE TECNOLOGIA E GEOCIÊNCIAS
DEPARTAMENTO DE ENGENHARIA CIVIL
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA CIVIL

ÁLAMO DITARSO SOUSA PESSOA

**ANÁLISE DE CONFIABILIDADE ESTRUTURAL DE DUTOS CORROÍDOS COM
FOCO NA INFLUÊNCIA DA CORRELAÇÃO ESPACIAL DA CORROSÃO**

Recife/PE

2025

ÁLAMO DITARSO SOUSA PESSOA

**ANÁLISE DE CONFIABILIDADE ESTRUTURAL DE DUTOS CORROÍDOS COM
FOCO NA INFLUÊNCIA DA CORRELAÇÃO ESPACIAL DA CORROSÃO**

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Civil da Universidade Federal de Pernambuco, Centro de Tecnologia e Geociências, como requisito parcial para a obtenção do título de Mestre em Engenharia Civil, Área de Estruturas.

Orientadora: Prof^a. Dra. Silvana Maria Bastos Afonso da Silva – UFPE

Coorientador: Prof^o. Dr. André Teófilo Beck – USP

Recife/PE

2025

.Catalogação de Publicação na Fonte. UFPE - Biblioteca Central

Pessoa, Álamo DiTarso Sousa.

Análise de confiabilidade estrutural de dutos corroídos com foco na influência da correlação espacial da corrosão / Alamo Ditarso Sousa Pessoa. - Recife, 2025.

119f.: il.

Universidade Federal de Pernambuco, Centro de Tecnologia e Geociências, Programa de Pós-Graduação em Engenharia Civil.

Orientação: Silvana Maria Bastos Afonso da Silva.

Coorientação: André Teófilo Beck.

Inclui referências e apêndices.

1. Dutos corroídos; 2. Confiabilidade estrutural; 3. Correlação; 4. Método de Monte Carlo. I. Silva, Silvana Maria Bastos Afonso da. II. Beck, André Teófilo. III. Título.

UFPE-Biblioteca Central

ÁLAMO DITARSO SOUSA PESSOA

**ANÁLISE DE CONFIABILIDADE ESTRUTURAL DE DUTOS CORROÍDOS COM
FOCO NA INFLUÊNCIA DA CORRELAÇÃO ESPACIAL DA CORROSÃO**

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Civil da Universidade Federal de Pernambuco, Centro de Tecnologia e Geociências, como requisito parcial para a obtenção do título de Mestre em Engenharia Civil, Área de Estruturas.

Aprovada em 29/01/2025

Orientadora: Prof^a. Dra. Silvana Maria Bastos Afonso da Silva - UFPE

Coorientador: Prof^o. Dr. André Teófilo Beck - USP

BANCA EXAMINADORA

Participação por videoconferência
Prof. Dr. Paulo Marcelo Vieira Ribeiro (Examinador Interno)
Universidade Federal de Pernambuco

Participação por videoconferência
Prof. Dr. Wellison José de Santana Gomes (Examinador Externo)
Universidade Federal de Santa Catarina

Participação por videoconferência
Prof. Dr. Adriano Dayvson Marques Ferreira (Examinador Externo)
Universidade Federal de Pernambuco

Dedico este trabalho à minha mãe.

AGRADECIMENTOS

Primeiramente, gostaria de expressar agradecimento especial à minha mãe, que sempre me deu suporte incondicional. Seu amor, dedicação e palavras de incentivo foram essenciais para que eu superasse os desafios dessa trajetória.

Em seguida, gostaria de expressar minha profunda gratidão à minha orientadora, professora Silvana Bastos, pela paciência, orientação e confiança depositada em meu trabalho. Sou imensamente grato pela oportunidade. Agradeço também ao professor André Teófilo Beck que, como coorientador, forneceu contribuições valiosíssimas para a devida progressão da pesquisa. Suas dicas e ensinamentos enriqueceram o trabalho de maneira significativa, e sou grato pelo apoio recebido.

Aos meus colegas de laboratório, meu muito obrigado pela companhia e colaboração ao longo de todos esses anos. Em especial, quero agradecer a Paulo Sousa e Savanna Cristina, com quem pude compartilhar momentos de aprendizado, superação e crescimento. Se a experiência foi rica e proveitosa, isto se deu muito em parte devido a vocês.

Ademais, agradeço Bruna Liberal, amiga com quem pude dividir, desde o início, tanto as alegrias como as preocupações inerentes a esta etapa da carreira acadêmica. As constantes palavras de suporte seguramente tornaram o percurso mais leve e menos desafiador.

Gostaria também de expressar minha gratidão às funcionárias da secretaria da pós-graduação, que sempre foram atenciosas e prestativas. Agradeço por todo apoio logístico e pela paciência em lidar com todas as questões administrativas que surgiram ao longo do percurso. Meu sincero agradecimento por todo o suporte prestado.

A todos que de alguma forma contribuíram para a realização deste trabalho, meus sinceros agradecimentos.

RESUMO

Dutos são amplamente utilizados no transporte de *commodities*, especialmente no setor de petróleo e gás, devido à sua segurança, eficiência e economia em longas distâncias. Com uma rede em constante crescimento, espera-se que mantenham sua importância por um longo período, sendo sua operação segura e adequada de grande interesse para as operadoras. O estudo da corrosão se mostra, portanto, de fundamental importância, uma vez que esta é uma das principais causas de acidentes em dutos ao atuar degradando a parede destes, reduzindo sua integridade estrutural e comprometendo sua resistência mecânica. A complexidade no tratamento numérico do fenômeno eletroquímico em si é aumentada pelas dificuldades em modelar as condições ambientais variáveis ao longo do duto – as quais não são distribuídas uniformemente e apresentam correlação espacial. O objetivo desta dissertação foi, portanto, avaliar a influência da correlação entre as taxas lineares de crescimento da profundidade de defeitos em dutos corroídos na probabilidade de falha ao longo do tempo. A probabilidade de falha foi calculada com base no modelo de falha à primeira sobrecarga e a determinação dos coeficientes de correlação entre as taxas de cada defeito foi feita arbitrando-se valores para os parâmetros das funções de autocorrelação: comprimento de correlação e espaçamento entre defeitos; foram avaliadas as diferenças na utilização de duas diferentes funções de autocorrelação: triangular e exponencial quadrática; os resultados foram obtidos para tempos totais de vida útil iguais a 15, 20, 25 e 30 anos e uma quantidade de defeitos variando de 2 a 5 num duto de comprimento arbitrário. Foi possível perceber que há uma influência considerável da correlação entre as taxas de crescimento da profundidade dos defeitos na probabilidade de falha total calculada no sentido de a probabilidade diminuir com um aumento na correlação. Este efeito é notado aumentar conforme cresce a quantidade de defeitos e conforme é considerado um maior tempo de vida útil. A aderência da relação probabilidade de falha vs. correlação a uma curva ajustada, escolhida como aquela com maior R^2 dentre vários tipos de curva, também cresce em função tanto do número de defeitos, como do tempo total.

Palavras-chave: dutos corroídos; confiabilidade estrutural; correlação; método de Monte Carlo

ABSTRACT

Pipelines are widely used in the transportation of commodities, especially in the oil and gas sector, due to their safety, efficiency and economy over long distances. With a constantly growing network, they are expected to maintain their importance for a long period, and their safe and adequate operation is of great interest to operators. The study of corrosion is therefore of fundamental importance, since this is one of the main causes of accidents in pipelines. Corrosion acts by degrading the pipeline wall, reducing its structural integrity and compromising its mechanical resistance. The complexity in the numerical treatment of the electrochemical phenomenon itself is increased by the difficulties in modeling the variable environmental conditions along the pipeline – which are not uniformly distributed and present spatial correlation. The objective of this dissertation was, therefore, to evaluate the influence of the correlation between the linear growth rates of the depth of defects in corroded pipelines on the probability of failure over time. The probability of failure was calculated based on the first passage probability concept and the determination of the correlation coefficients between the rates of each defect was done by arbitrating values for the parameters of the autocorrelation functions: correlation length and spacing between defects; the differences in the use of two different autocorrelation functions, triangular and quadratic exponential, were evaluated; the results were obtained for total lifespans equal to 15, 20, 25 and 30 years and a number of defects ranging from 2 to 5 in an arbitrarily long pipeline. It was possible to notice that there is a considerable influence of the correlation between the linear growth rates of the depth of the defects on the calculated total failure probability in the sense that the probability decreases with an increase in the correlation. This effect is noted to increase as the number of defects increases and as a longer lifespan is considered. The adherence of the probability of failure vs. correlation relationship to a certain fitted curve, chosen as that one with the highest R^2 values among several types of curves, also increases as a function of both the number of defects and the total time.

Keywords: corroded pipelines; structural reliability; correlation; Monte Carlo Method

LISTA DE FIGURAS

Figura 1 – Distribuição conjunta de probabilidades.....	17
Figura 2 – Diferentes comportamentos de dependência entre variáveis aleatórias. .	20
Figura 3 – PDF e CDF de uma distribuição normal com média e desvio-padrão.....	21
Figura 4 – Desenho esquemático de várias realizações de um processo estocástico com parâmetro contínuo.....	28
Figura 5 – Função de autocorrelação triangular (esquerda) e sua respectiva função de densidade espectral de potência.....	37
Figura 6 – Função de autocorrelação exponencial quadrática (esquerda) e sua respectiva função de densidade espectral de potência.....	37
Figura 7 – Exemplo de um processo de pontos.....	38
Figura 8 – Domínios de falha e de segurança dadas as PDFs marginais e conjunta de R e S.....	42
Figura 9 – Equação de estado-limite e os domínios de falha e sobrevivência com base nos valores da função de estado-limite.....	43
Figura 10 – Transformação do espaço original para o espaço isoprobabilístico no FORM.....	46
Figura 11 – Etapas da transformação composta: VAs correlacionadas e em distribuições quaisquer → VAs correlacionadas mas com distribuição normal-padrão → VAs independentes no espaço normal-padrão.....	47
Figura 12 – Ilustração de um problema de confiabilidade dependente do tempo. A ação e resistência, e , respectivamente, variam no tempo.....	51
Figura 13 - As diversas formas de simplificação da geometria real de um defeito para fins de cálculo da pressão interna resistente	55
Figura 14 – Modos de falha em função do comprimento e profundidade do defeito. .	56
Figura 15 – Desenho esquemático de diversas realizações da taxa linear de corrosão ao longo do contínuo , o comprimento longitudinal do duto.....	61
Figura 16 – Distribuição de defeitos igualmente espaçados num duto.....	69
Figura 17 – Máxima diferença relativa (em módulo) para função de autocorrelação triangular (linear).....	71
Figura 18 – Máxima diferença relativa (em módulo) vs. ρ_{Rd} para função de autocorrelação triangular (linear).....	72

Figura 19 – Diferença relativa média (em módulo) para função de autocorrelação triangular (linear).....	73
Figura 20 – Diferença relativa média (em módulo) vs. ρ_{Rd} para função de autocorrelação triangular (linear).....	74
Figura 21 – Máxima diferença relativa (em módulo) para função de autocorrelação exponencial quadrática.....	74
Figura 22 – Diferença relativa média (em módulo) para função de autocorrelação exponencial quadrática.....	75
Figura 23 – Máxima diferença relativa (em módulo) vs. ρ_{Rd} para função de autocorrelação exponencial quadrática.....	76
Figura 24 – Diferença relativa média (em módulo) vs. ρ_{Rd} para função de autocorrelação exponencial quadrática.....	76
Figura 25 – Prob. de falha vs. 'c' vs.'ds' (Rxx: linear, Td=15 anos).....	77
Figura 26 – Prob. de falha vs. 'c' vs.'ds' (Rxx: linear, Td=20 anos).....	78
Figura 27 – Prob. de falha vs. 'c' vs.'ds' (Rxx: linear, Td=25 anos).....	78
Figura 28 – Prob. de falha vs. 'c' vs.'ds' (Rxx: linear, Td=30 anos).....	79
Figura 29 – Prob. falha vs. ρ_{Rd} (linear) para Td=15 anos. T0: 2.88 anos.....	80
Figura 30 – Prob. falha vs. ρ_{Rd} (linear) para Td=20 anos. T0: 2.88 anos.....	80
Figura 31 – Prob. falha vs. ρ_{Rd} (linear) para Td=25 anos. T0: 2.88 anos.....	81
Figura 32 – Prob. falha vs. ρ_{Rd} (linear) para Td=30 anos. T0: 2.88 anos.....	81
Figura 33 – R ² vs. Td para vários números de defeito considerados.....	82
Figura 34 – Prob. de falha vs. 'c' vs.'ds' (Rxx: expquad, Td=15 anos).....	83
Figura 35 – Prob. de falha vs. 'c' vs.'ds' (Rxx: expquad, Td=20 anos).....	84
Figura 36 – Prob. de falha vs. 'c' vs.'ds' (Rxx: expquad, Td=25 anos).....	84
Figura 37 – Prob. de falha vs. 'c' vs.'ds' (Rxx: expquad, Td=30 anos).....	85
Figura 38 – Prob. falha vs. ρ_{Rd} (expquad) para Td=15 anos. T0: 2.88 anos.....	86
Figura 39 – Prob. falha vs. ρ_{Rd} (expquad) para Td=20 anos. T0: 2.88 anos.....	86
Figura 40 – Prob. falha vs. ρ_{Rd} (expquad) para Td=25 anos. T0: 2.88 anos.....	87
Figura 41 – Prob. falha vs. ρ_{Rd} (expquad) para Td=30 anos. T0: 2.88 anos.....	87
Figura 42 – R ² vs. Td para vários números de defeito considerados.....	88
Figura 43 – Histograma com distribuições dos tempos de surgimento dos primeiros 5 defeitos na primeira passada.....	89
Figura 44 – Prob. falha vs. ρ_{Rd} (linear) para Td=50 anos. T0: NHPP.....	90

	3
Figura 45 – Prob. falha vs. ρR_d (linear) para $T_d=75$ anos. T0: NHPP.....	90
Figura 46 – R^2 vs. T_d para vários números de defeito considerados.....	91
Figura 47 – R^2 vs. T_d para vários números de defeito considerados.....	92
Figura 48 – Representação por boxplot dos tempos de surgimento dos defeitos por NHPP para a 1ª execução do algoritmo.....	107
Figura 49 – Representação por histograma dos tempos de surgimento dos defeitos por NHPP para a 1ª execução do algoritmo.....	107
Figura 50 – Representação por boxplot dos tempos de surgimento dos defeitos por NHPP para a 2ª execução do algoritmo.....	108
Figura 51 – Representação por histograma dos tempos de surgimento dos defeitos por NHPP para a 2ª execução do algoritmo.....	108
Figura 52 – Representação por boxplot dos tempos de surgimento dos defeitos por NHPP para a 3ª execução do algoritmo.....	109
Figura 53 – Representação por histograma dos tempos de surgimento dos defeitos por NHPP para a 3ª execução do algoritmo.....	109
Figura 54 – Representação por boxplot dos tempos de surgimento dos defeitos por NHPP para a 4ª execução do algoritmo.....	110
Figura 55 – Representação por histograma dos tempos de surgimento dos defeitos por NHPP para a 4ª execução do algoritmo.....	110
Figura 56 – Representação por boxplot dos tempos de surgimento dos defeitos por NHPP para a 5ª execução do algoritmo.....	111
Figura 57 – Representação por histograma dos tempos de surgimento dos defeitos por NHPP para a 5ª execução do algoritmo.....	111

Sumário

1	INTRODUÇÃO.....	8
1.1	CONSIDERAÇÕES INICIAIS.....	8
1.2	OBJETIVOS.....	13
1.2.1	Objetivo geral.....	13
1.2.2	Objetivos específicos.....	13
1.3	ORGANIZAÇÃO DO TEXTO.....	13
2	REFERENCIAL TEÓRICO.....	14
2.1	VARIÁVEIS ALEATÓRIAS E DISTRIBUIÇÕES DE PROBABILIDADE....	14
2.1.1	Distribuição conjunta de probabilidades.....	17
2.1.1.1	Dependência e independência.....	18
2.1.1.2	Covariância e coeficiente de correlação linear.....	19
2.1.2	Distribuição normal.....	20
2.1.3	Distribuição de Poisson.....	23
2.1.4	Distribuição Exponencial.....	25
2.1.5	Função de uma variável aleatória.....	26
2.2	PROCESSOS ESTOCÁSTICOS.....	27
2.2.1	Funções de distribuição de probabilidade e os momentos de um processo estocástico.....	29
2.2.2	Funções de distribuição de probabilidade e os momentos de dois processos estocásticos.....	31
2.2.3	Estacionariedade.....	32
2.2.4	Estatísticas do contínuo e ergodicidade.....	34
2.2.5	Função de autocorrelação: propriedades e características.....	35
2.2.5.1	Propriedades.....	35
2.2.5.2	Relação com a função de densidade espectral de potência.....	35
2.2.5.3	Expressões usuais.....	37
2.2.6	Processo de pontos de Poisson.....	37
2.3	INCERTEZAS NA ENGENHARIA E O CONCEITO DE CONFIABILIDADE ESTRUTURAL.....	41
2.3.1	O problema básico de confiabilidade e os conceitos de função e equação de estado-limite.....	41

	5
2.3.2	Métodos usuais para cálculo da probabilidade de falha.....43
2.3.2.1	Método de Monte Carlo.....44
2.3.2.2	Principais métodos de transformação e o mapeamento das variáveis aleatórias para o espaço normal-padrão.....45
2.3.2.2.1	<i>Obtenção de distribuições normais equivalentes.....47</i>
2.3.2.2.2	<i>Eliminação da correlação.....49</i>
2.3.3	Confiabilidade no tempo e modelo de falha à primeira sobrecarga...50
2.3.4	Confiabilidade em dutos corroídos.....53
2.3.4.1	Probabilidade de falha para eventos dependentes e independentes no contexto de dutos corroídos.....58
2.3.4.2	Taxa de corrosão como processo estocástico tendo a dimensão longitudinal do duto como parâmetro contínuo.....60
3	METODOLOGIA.....62
3.1	CÁLCULO DA PROBABILIDADE DE FALHA DO DUTO.....62
3.2	OBTENÇÃO DE AMOSTRAS CORRELACIONADAS.....65
3.3	ESTUDO DE CASO.....68
3.4	AJUSTES DE CURVA.....70
4	RESULTADOS.....71
4.1	APROXIMAÇÕES OBTIDAS PARA CORRELAÇÃO DAS AMOSTRAS DAS TAXAS LINEARES DE CRESCIMENTO DA PROFUNDIDADE DOS DEFEITOS.....71
4.1.1	Função de autocorrelação triangular (linear).....71
4.1.2	Função de autocorrelação exponencial quadrática.....74
4.2	PROBABILIDADE DE FALHA DO DUTO.....77
4.2.1	Função de autocorrelação triangular (linear) – tempo determinístico para o surgimento dos defeitos.....77
4.2.2	Função de autocorrelação exponencial quadrática – tempo determinístico para o surgimento dos defeitos.....83
4.2.3	Função de autocorrelação triangular (linear) – tempo aleatório para surgimento dos defeitos e seguindo modelo NHPP.....88
4.2.4	Função de autocorrelação exponencial quadrática – tempo aleatório para surgimento dos defeitos e seguindo modelo NHPP.....91
5	CONCLUSÕES E SUGESTÕES DE TRABALHOS FUTUROS.....94

REFERÊNCIAS.....	6
REFERÊNCIAS.....	96
APÊNDICE A: TABELAS COM FUNÇÕES DE INTERPOLAÇÃO.....	103
APÊNDICE B: GRÁFICOS COM DETALHES DAS REALIZAÇÕES DOS TEMPOS DE SURGIMENTO DOS DEFEITOS POR NHPP.....	107

1 INTRODUÇÃO

Este trabalho tem como fim o estudo da confiabilidade estrutural de dutos corroídos considerando a influência da correlação espacial entre as taxas lineares de crescimento da profundidade de defeitos idealizados – a chamada taxa de corrosão. Os valores para correlação entre estas taxas são obtidos a partir de duas distintas funções de autocorrelação e os resultados são comparados para diferentes tempos totais de análise.

O documento é a dissertação de mestrado do autor, submetida ao Programa de Pós-Graduação em Engenharia Civil da Universidade Federal de Pernambuco como requisito para obtenção do título de Mestre em Engenharia Civil.

1.1 CONSIDERAÇÕES INICIAIS

Dutos são largamente entendidos como um dos meios de transporte de *commodities* mais empregados a nível mundial, principalmente no que se refere ao circuito do petróleo e gás. Sua segurança, eficiência e economia na movimentação destas cargas por longas distâncias são notáveis; dado o tamanho da malha atual, bem como sua expectativa de crescimento, é esperado que mantenham sua importância por tempo considerável. (ZELMATI et al., 2022; WOLDESELLASSE; TEFAMARIAM, 2023b; HUSSAIN et al., 2024; MOTTA; FERREIRA; AFONSO, 2024)

Conforme apresentado por Ahmed e Kabir (2021) citando dados canadenses, em 2017 18% do gás natural do país era consumido por usuários domésticos e 68% pelo setor industrial, com o consumo desta *commodity* sendo esperado crescer em 43% mundialmente até 2040. Hidrocarbonetos como este eram transportados no Canadá por 760.000km de dutos em 2020 – crescimento considerável em comparação aos 580.00km presentes em 2009, conforme dados citados por Bedairi et al. (2012). Estes mesmos autores indicam que 97% do petróleo e gás natural canadenses eram movidos por meio dutoviário naquele ano; Khan et al. (2021) destaca que havia até 2020 117.000km de dutos de grande diâmetro instalados no Canadá. Para os EUA, por sua vez, Dahire et al. (2018) afirmam que, à época, a malha era composta por 305.000 milhas de dutos, distribuídos em mais de 210 sistemas diferentes. A nível mundial, em 2019 o petróleo e o gás correspondiam, somados, a cerca 57% das fontes de energia primária utilizadas no planeta, sendo

33.1% referentes a petróleo e 24.2% a gás. (BRITISH PETROLEUM, 2020; CABRAL, 2023)

Percebe-se, portanto, que, sendo o meio mais empregado para o transporte de fontes energéticas essenciais, a adequada e segura operação de linhas dutoviárias é de imediato interesse das operadoras. Desta forma, especial atenção é dada por estes agentes à prevenção de acidentes que, dada a característica inerentemente perigosa do produto transportado, muitas vezes resultam não somente em devastadores prejuízos ambientais, como contaminação de lençóis-freáticos e terras agricultáveis, e materiais (danos às infraestruturas circundantes, paralisação das operações etc.), mas também possuem um potencial particularmente notável de perda de vidas humanas. Tudo isto não obstante os avanços recentes dos métodos, práticas de segurança, ferramentas preditivas etc.. (MEDEIROS; ALENCAR; DE ALMEIDA, 2017; BIEZMA et al., 2020; KHAN; YARVEISY; ABBASSI, 2021)

As causas para estes acidentes são variadas, sendo a corrosão uma das principais. Dai *et al.* (2017) mostram que nos EUA, de 2010 a 2015, a corrosão foi responsável por 25% das falhas em dutos que transportavam petróleo e 24% daqueles que transportavam gás; para a Europa, os valores são de 24% das falhas acontecendo por corrosão entre 1970 e 2013; na China, até 2015 as falhas por corrosão eram 9%, terceiro maior valor ao se desconsiderar os 50% atribuídos ao furto de combustível. Para o Reino Unido, o relatório da UKOPA (*United Kingdom Onshore Pipeline Operators' Association*) mostra que, entre os anos de 1962 e 2014, a corrosão foi motivo de cerca de 35% das falhas em dutos ao somar os valores percentuais para corrosão externa e para corrosão interna. (HASWELL; MCCONNELL, 2015)

A corrosão enquanto mecanismo de causa, enquanto agente indutor de acidentes, atua degradando a parede do duto durante a operação, reduzindo sua integridade estrutural e fazendo com que a resistência mecânica aos esforços solicitantes diminua progressivamente (MUHLBAUER, 2004; NESSIM, 2011; GOMES; BECK; HAUKAAS, 2013). A consideração de como a corrosão afeta as propriedades de resistência mecânica do duto (ou como pode vir a acarretar vazamentos, por exemplo) ao longo de sua vida útil é componente crucial para a segura operação das malhas, levando a ser rica a literatura acerca da confiabilidade

(aqui entendida como a probabilidade de que não haja falha) de dutos com defeitos de corrosão e como esta se comporta no contínuo do tempo. Trabalhos como o de Li *et al.* (2017), Liu *et al.* (2018) e Mishra, Keshavarzzadeh e Noshadravan (2019) por exemplo, estudam a confiabilidade de dutos como forma de melhor planejar inspeções e ações de manutenção e reparo; Bazán e Beck (2013) utilizaram dos cálculos de probabilidade de falha no tempo como ferramenta para avaliação das diferenças na consideração do crescimento da corrosão ora como simples variável aleatória, ora como processo estocástico; De Leon e Macías (2005), Li *et al.* (2009), Zhang (2019) e Fu *et al.* (2019), por outro lado, analisaram os efeitos advindos da consideração de correlações entre algumas variáveis aleatórias na probabilidade de falha.

De fato, o cerne da problemática em torno da confiabilidade de uma estrutura advém das incertezas atreladas às variáveis sendo consideradas no problema e em como estas afetam o cálculo da previsão da chance de uma determinada condição – neste caso a falha – ser atingida (MELCHERS; BECK, 2018). Naquilo que concerne a corrosão em dutos, incertezas são especialmente pronunciadas devido não somente à dificuldade no estabelecimento de um modelo matemático do fenômeno eletroquímico da corrosão em si, mas também ao fato de que este é um processo que depende de condições ambientais do local em que o duto está instalado. (MURATA, 2011)

Estas condições ambientais em questão (umidade, temperatura, pH do solo, presença de contaminantes, presença de correntes elétricas oriundas de instalações próximas e que possam induzir corrosão etc.) variam a depender de que partes da estrutura estão sendo analisadas e comparadas entre si. Isto faz com que a corrosão não ocorra uniformemente ao longo do comprimento do duto, tendo regiões próximas umas das outras uma tendência a estarem submetidas a condições ambientais similares. Isto leva, por conseguinte, a uma tendência em as formas com que corrosão se desenvolve em pontos próximos uns dos outros também serem similares. A esta tendência se dá o nome de correlação espacial: a medida da relação de dependência linear que propriedades do duto, descritas por variáveis aleatórias, têm entre si a depender da distância entre pontos de referência. (DE LEON; MACÍAS, 2005; LI *et al.*, 2009; PAIK; MELCHERS, 2009; NESSIM, 2011;

MURATA, 2011; MELCHERS; BECK, 2018; BECK, 2019; FU et al., 2019; AMAYA-GÓMEZ et al., 2024)

Dos vários trabalhos que abordam o tema da correlação entre variáveis no problema de confiabilidade estrutural de dutos corroídos, pode-se citar quatro particularmente pertinentes ao estudo desenvolvido no âmbito desta dissertação: De Leon e Macías (2005), Li *et al.* (2009), Zhang (2019) e Fu *et al.* (2019). No primeiro, os autores, ao considerarem o duto corroído como um sistema em série composto por vários segmentos – cada um contendo um defeito –, arbitram valores para os coeficientes de correlação ρ entre as profundidades iniciais do defeito necessárias para o cômputo da pressão de falha (aquela pressão interna a qual o duto – com determinada geometria e características mecânicas – suporta tendo um defeito idealizado – também de geometria determinada – em sua superfície). Com isso, os autores são capazes de calcular as correlações entre os valores das funções de estado-limite e de mostrar que há uma tendência em estas correlações aumentarem conforme aumenta a correlação entre as profundidades iniciais dos defeitos. Assim, demonstram que a partir de valores de correlação de 60% entre as profundidades iniciais dos defeitos a correlação nas funções de estado-limite torna-se maior que 40%, o que pode ser considerado significativo e, portanto, digno de ser levado em consideração. Por fim, chegam à conclusão de que não contemplar a correlação no problema de confiabilidade em dutos leva a soluções conservadoras.

Li *et al.* (2009) estudaram os efeitos da consideração da correlação no problema obtendo valores para as probabilidades de falha – em vez de se valerem apenas dos limites superior e inferior calculados pelas aproximações por limites bimodais como em De Leon e Macías (2005). Os autores procedem, então, com análises da sensibilidade da função de estado-limite, inicialmente avaliando a influência de variações nas médias das variáveis (tido como estudo acerca da sensibilidade às incertezas das variáveis) e, posteriormente, procurando entender como a função de estado-limite se correlacionava, no tempo, com as variáveis do problema; esta é tida como aquela abordagem que também leva em conta a sensibilidade do modelo em si. Concluem não somente corroborando as observações de De Leon e Macías (2005) de que a não-consideração da correlação leva a resultados conservadores, mas também de que aquelas variáveis que mais influem nos valores da função de estado-limite mudam com o tempo. Dentre as

variáveis estudadas, a taxa de corrosão radial apresentou -12.42% da contribuição para a variância da função de estado-limite em $T=12$ anos e -27.43% em $T=16$ anos; a correlação entre a taxa de corrosão radial e os valores da função de estado-limite saltam de -0.34 em $T=12$ anos para -0.76 em $T=16$ anos.

Resultados similares são observados por Zhang *et al.* (2019), onde foram estudados os efeitos da consideração da correlação entre quatro pares de variáveis no problema da confiabilidade de dutos corroídos com a presença de um defeito. Os pares de variáveis estudados foram: diâmetro e espessura da parede do duto; comprimento e profundidade do defeito; taxas radial e axial de corrosão; tensão de escoamento e tensão máxima de tração do material. Ainda nesta referência, quatro modos de falha foram considerados: perfuração da parede do duto, ou pequeno vazamento (*small-leak*); por grande vazamento (*burst*), por ruptura longitudinal após grande vazamento (*rupture*); e a falha composta (quando, num determinado ponto do tempo, o duto apresenta falha simultânea tanto por pequeno vazamento, quanto por grande vazamento). Os autores concluíram que a probabilidade de falha era particularmente sensível à correlação entre as taxas lineares de corrosão radial e axial do defeito; puderam observar, também, que o impacto da correlação diminuía com o tempo.

Finalmente, Fu *et al.* (2019) também estudaram o tópico de correlações no contexto de dutos corroídos. Os autores expõem como a correlação dos valores para diferentes funções de estado-limite consigo mesmas evolui no tempo, demonstrando como a considerável influência desta correlação justifica a interpretação das variáveis aleatórias do problema como processos estocásticos no domínio do tempo e, por conseguinte, fazendo valer a ideia de cálculo da probabilidade de falha como um problema de primeira passagem por barreira (*first passage probability*, na sigla em inglês, ou “falha à primeira sobrecarga”, como também é conhecida), típica do estudo de processos estocásticos. Assim como os outros trabalhos previamente mencionados, aqui também os autores notam uma discrepância considerável nas probabilidades de falha calculadas com e sem a consideração da correlação, tendo a não-consideração implicações conservadoras para os valores obtidos.

1.2 OBJETIVOS

1.2.1 Objetivo geral

Este trabalho tem como objetivo geral investigar de forma detalhada como a confiabilidade de um duto destinado ao transporte de petróleo e gás é influenciado pela correlação espacial das taxas lineares de crescimento da profundidade de defeitos de corrosão idealizados.

1.2.2 Objetivos específicos

Como objetivos específicos, destaca-se:

- Obter amostras de valores aleatórios em distribuições específicas e correlacionados de valores específicos a serem determinados a partir da disposição espacial dos defeitos;
- Avaliar o comportamento das probabilidades de falha calculadas quando em função da correlação entre defeitos adjacentes;
- Comparar as diferenças no conteúdo final das probabilidades calculadas quando da consideração de duas distintas funções de autocorrelação;
- Estudar a aderência dos valores de Probabilidade de Falha *versus* Correlação ρ ao comportamento descrito por uma curva ajustada.

1.3 ORGANIZAÇÃO DO TEXTO

O texto se conforma dividido em seis capítulos:

- Capítulo 1 – Introdução: aspectos gerais da problemática de dutos e corrosão, bem como os desenvolvimentos recentes no tema;
- Capítulo 2 – Referencial Teórico: principais marcos teóricos que embasam a pesquisa desenvolvida: variáveis aleatórias, processos estocásticos, incertezas na engenharia, confiabilidade no tempo etc.;
- Capítulo 3 – Metodologia: métodos utilizados no cálculo das probabilidades de falha, meios com quais amostras correlacionadas foram obtidas, dados empregados no problema etc.;
- Capítulo 4 – Resultados e discussões
- Capítulo 5 – Conclusão: discussão e sugestão de trabalhos futuros

2 REFERENCIAL TEÓRICO

2.1 VARIÁVEIS ALEATÓRIAS E DISTRIBUIÇÕES DE PROBABILIDADE

Dado o espaço amostral Ω , o conjunto de todos os possíveis resultados de um experimento ou fenômeno aleatório, chama-se de evento qualquer subconjunto $\omega_i \in \Omega$. Diz-se, portanto, ser uma variável aleatória aquela função $X(\omega)$ com $\omega \in \Omega$, tal que para cada ponto amostral ω_i , um valor real $x(\omega_i)$ lhe é atribuído (MAGALHÃES, 2011; MAGALHÃES; LIMA, 2015; KULKARNI, 2017). Desta forma, o conjunto $\{X(\omega) \leq x\}$ é um evento para todo e qualquer valor real de x (ELISHAKOFF, 2017; BECK, 2019).

Em outras palavras, uma variável aleatória pode ser entendida como aquela função que mapeia os eventos de um espaço de possibilidades Ω , seu domínio, que pode ser discreto ou contínuo, até a reta dos reais, sua imagem (YATES; GOODMAN, 2005; ANG; TANG, 2007; MONTGOMERY; RUNGER, 2018). A cada valor $x(\omega_i)$ assumido pela variável, dá-se o nome de realização.

Uma variável aleatória é completamente descrita por aquilo que se chama de Função Acumulada de Probabilidade (*Cumulative Distribution Function*, CDF, na sigla em inglês empregada no restante do texto) dada pela probabilidade de que o evento $X \leq x$ ocorra. Matematicamente apresentada na Equação 1 (KULKARNI, 2017)

$$F_X(x) \equiv Pr(X \leq x) = Pr(\{\omega \in \Omega : X(\omega) \leq x\}), -\infty < x < \infty, \quad (1)$$

onde $Pr(\cdot)$ é o operador que indica a probabilidade de um determinado evento acontecer; referido doravante de maneira mais simples por $P(\cdot)$.

A função acumulada de probabilidade é monotonicamente crescente, ou seja, $(F_X(x_1) \leq F_X(x_2), \forall x_1 \leq x_2)$ e limitada superiormente por 1 (i.e. $\lim_{x \rightarrow \infty} F_X = 1$) e inferiormente por 0 (i.e. $\lim_{x \rightarrow -\infty} F_X = 0$).

Uma variável aleatória X é dita discreta se há um conjunto enumerável, finito ou não, de valores distintos x_1, x_2, \dots, x_n tais que $a_i = P(X = x_i) > 0$ para $i = 1, 2, \dots, n$. Assim sendo, a função descrita na Equação 2, chamada de Função de Probabilidade ou Função de Massa de Probabilidade (*Probability Mass Function*, PMF), descreve a probabilidade de que o evento x_i ocorra:

$$p(x_i) = P(X = x_i) = P(\{\omega \in \Omega : X(\omega) = x_i\}), \text{ para } i = 1, 2, \dots, n, \quad (2)$$

e obedece às seguintes propriedades:

a) $0 \leq p(x_i) \leq 1$, para $i = 1, 2, \dots, n$;

b) $\sum_{x_i \in A} p(x_i) = 1$

Sua CDF, no caso, é uma função degrau com saltos num conjunto discreto A dos números reais onde se encontram os valores x que a variável assume. O tamanho dos saltos em cada x_i é igual a $p(x_i)$ (TAYLOR; KARLIN, 1994; MAGALHÃES, 2011). De fato, conhecida a função de probabilidade, então a função acumulada de probabilidades passa a ser dada pela Equação 3:

$$F_X(x) = \sum_{i \in A_x} p(x_i), \text{ com } A_x = \{i : x_i \leq x\}, \text{ ou } F_X(x) = \sum_{x_i \leq x} p(x_i). \quad (3)$$

Por outro lado, uma variável aleatória X é dita contínua se suas probabilidades não forem associadas a valores específicos x_i , mas a intervalos na linha dos reais. Assim, seu comportamento de probabilidades é descrito por uma CDF F_X caracterizada por ser uma função contínua e suave. Se este é o caso e se existe uma função $f_X(x)$, não-negativa, com $-\infty \leq x \leq \infty$, tal que:

$$P(a \leq X(\omega) \leq b) = \int_a^b f_X(x) dx, \text{ com } -\infty \leq a \leq b \leq \infty, \quad (4)$$

então, $f_X(x)$ é chamada de Função de Densidade de Probabilidade (*Probability Density Function*, doravante referida como PDF) da variável aleatória X . Consequentemente:

$$F_X(x) = \int_{-\infty}^x f_X(u) du, \text{ com } -\infty \leq x \leq \infty, \quad (5)$$

é a CDF de X . (TAYLOR; KARLIN, 1994; ANG; TANG, 2007; KULKARNI, 2017)

Naturalmente, a PDF obedece às seguintes propriedades:

a) $f_X(x) \geq 0, \forall x \in \mathbb{R}$;

b) $\int_{-\infty}^{\infty} f_X(x) dx = 1$

A média (ou valor esperado) de uma variável aleatória é definida como:

$$E[X] = \int_{-\infty}^{\infty} x f_X(x) dx = \mu, \quad (6)$$

ou seja, o centro de gravidade da PDF de X . A média é, também, o momento de primeira ordem da variável aleatória, em que um momento de ordem k qualquer é dado por:

$$E[X^k] = \int_{-\infty}^{\infty} x^k f_X(x) dx = \mu^k. \quad (7)$$

Caso estejamos tratando de uma função $g(X)$ de uma variável aleatória, o valor esperado se torna:

$$E[g(x)] = \int_{-\infty}^{\infty} g(x) f_X(x) dx. \quad (8)$$

A variância, por sua vez, medida da dispersão dos valores em torno da média, é dada por:

$$\text{Var}[X] \equiv E[(X - \mu)^2] = \int_{-\infty}^{\infty} (x - \mu)^2 f_X(x) dx = \sigma^2, \quad (9)$$

em que o valor $\sigma = \sqrt{\text{Var}[X]}$ é chamado desvio-padrão da variável X . Pode ser expressa como:

$$\text{Var}[X] = E[X^2] - E[X]^2. \quad (10)$$

A variância corresponde ao momento de inércia da PDF da variável aleatória (BECK, 2019), e, assim, como o valor esperado, é também um momento desta mesma variável. No caso, a variância é o momento central de segunda ordem da variável X ; de fato, para uma variável aleatória contínua, um momento central de ordem k qualquer é dado por:

$$E[(X - \mu)^k] = \int_{-\infty}^{\infty} (x - \mu)^k f_X(x) dx = m^k, \quad (11)$$

e um momento de ordem k , por sua vez, é dado por:

$$E[X^k] = \int_{-\infty}^{\infty} x^k f_X(x) dx = \mu^k. \quad (12)$$

O caso específico em que $k=0$ resulta, para ambos os momentos, em 1, que é a probabilidade do espaço amostral inteiro; o primeiro momento ($k=1$, ou seja μ^1) é igual à média μ e o primeiro momento central igual a 0.

Tratando-se de amostras com n observações de uma determinada variável aleatória X com média μ e desvio-padrão igual a σ , então a média e variância desta amostra são dadas por:

$$\bar{\mu} = \frac{1}{n} \cdot \sum_{i=1}^n x_i, \quad (13)$$

$$\bar{v} = \frac{1}{(n-1)} \cdot \sum_{i=1}^n (x_i - \bar{x})^2. \quad (14)$$

Ao valor adimensional obtido pela razão entre desvio-padrão e média dá-se o nome de Coeficiente de Variação:

$$CoV = \frac{\bar{\sigma}}{\bar{\mu}}. \quad (15)$$

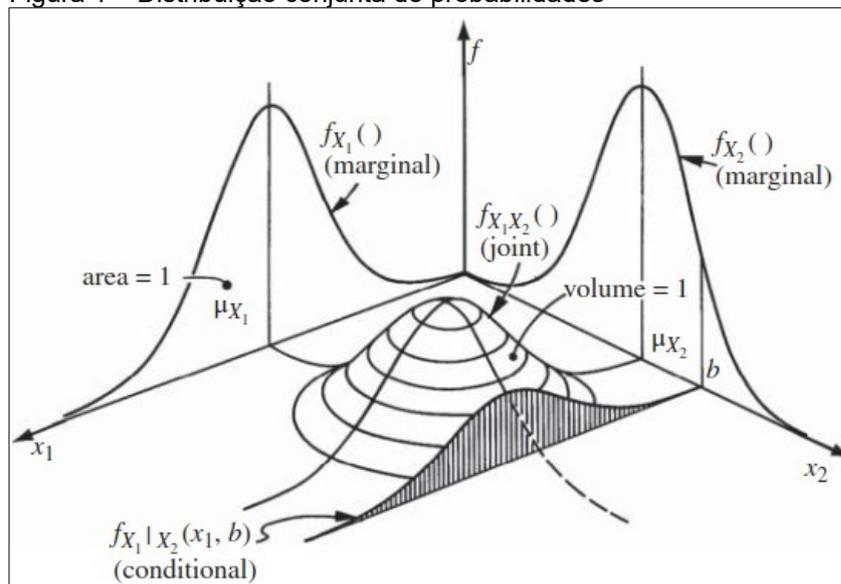
2.1.1 Distribuição conjunta de probabilidades

Se um determinado evento é resultado de duas variáveis aleatórias X e Y , a probabilidade não-nula de que ocorram valores menores que x e y é descrita por uma distribuição cumulativa conjunta de probabilidades (ou distribuição bivariada) de probabilidades. Matematicamente escreve-se a distribuição conjunta como:

$$F_{XY}(x, y) = P([X \leq x \cap Y \leq y]) = P([X \leq x, Y \leq y]), \quad (16)$$

onde as distribuições $F_X(x) = P(X \leq x)$ e $F_Y(y) = P(Y \leq y)$ são chamadas de distribuições marginais de probabilidade, representadas na Figura 1 (ELISHAKOFF, 2017; MELCHERS; BECK, 2018).

Figura 1 – Distribuição conjunta de probabilidades



Fonte: Melchers, R.E.; Beck, A.T. (2018)

Uma vez que $F_{XY}(x, y)$ é uma medida de probabilidade, seu valor está restrito ao intervalo $[0, 1]$ para $-\infty \leq x \leq \infty$ e $-\infty \leq y \leq \infty$, de tal forma que:

$$\begin{aligned} \lim_{\substack{x \rightarrow \infty \\ y \rightarrow \infty}} F_{XY}(x, y) &= F_{XY}(\infty, \infty) = 1 \\ \lim_{\substack{x \rightarrow -\infty \\ y \rightarrow -\infty}} F_{XY}(x, y) &= F_{XY}(-\infty, -\infty) = 0 \end{aligned} \quad (17)$$

Como os eventos $y \leq \infty$ e $x \leq \infty$ são certos, então (IBE, 2014; BECK, 2019):

$$F_{XY}(x, \infty) = P([X \leq x, Y \leq \infty]) = P(X \leq x) \cdot 1 = F_X(x), \text{ e por analogia:} \quad (18)$$

$$F_{XY}(\infty, y) = F_Y(y). \quad (19)$$

Assim, naturalmente:

$$F_{XY}(x, -\infty) = F_{XY}(-\infty, y) = 0 \quad (20)$$

2.1.1.1 Dependência e independência

Duas variáveis são definidas independentes se os eventos $X \leq x$ e $Y \leq y$ são independentes, o que leva a distribuição F_{XY} conjunta a ser expressa por:

$$F_{XY}(x, y) = P([X \leq x, Y \leq y]) = P(X \leq x) \cdot P(Y \leq y) = F_X(x) \cdot F_Y(y), \quad (21)$$

para quaisquer valores de x e y nos reais. Portanto, tem-se que a função de densidade de probabilidade conjunta é dada por:

$$f_{XY}(x, y) = f_X(x) \cdot f_Y(y), \quad (22)$$

Em sendo independentes, as distribuições condicionais de cada uma das variáveis aleatórias X e Y com respeito à outra são iguais às suas distribuições incondicionais (ELISHAKOFF, 2017):

$$\begin{aligned} f_{X|Y}(x|y) &= f_{XY}(x, Y=y) = f_X(x) \\ f_{Y|X}(y|x) &= f_{XY}(X=x, y) = f_Y(y) \end{aligned} \quad (23)$$

Se as variáveis aleatórias X e Y não forem independentes, no entanto, a expressão fica (BECK, 2019):

$$\begin{aligned} f_{X|Y}(x|Y=y) &= \frac{f_{XY}(x, y)}{f_Y(y)}, \text{ desde que } f_Y(y) > 0 \\ f_{Y|X}(y|X=x) &= \frac{f_{XY}(x, y)}{f_X(x)}, \text{ desde que } f_X(x) > 0 \end{aligned} \quad (24)$$

2.1.1.2 Covariância e coeficiente de correlação linear

A covariância é uma medida linear da dependência entre duas variáveis aleatórias X e Y e é definida da seguinte maneira:

$$\text{Cov}[X, Y] \equiv E[(X - \mu_x)(Y - \mu_y)], \quad (25)$$

cujo desenvolvimento leva a (IBE, 2014):

$$\begin{aligned} \text{Cov}[X, Y] &= E[XY - \mu_x Y - \mu_y X + \mu_x \mu_y] \\ " " &= E[XY] - \mu_x E[Y] - \mu_y E[X] + \mu_x \mu_y \\ " " &= E[XY] - \mu_x \mu_y - \mu_y \mu_x + \mu_x \mu_y \\ " " &= E[XY] - \mu_x \mu_y \end{aligned} \quad (26)$$

Se as variáveis são independentes, então $E[XY] = E[X]E[Y] = \mu_x \mu_y$ (MAGALHÃES; LIMA, 2015) e a Equação 26 vai para zero.

A covariância de uma variável aleatória consigo mesma é igual à variância desta variável aleatória:

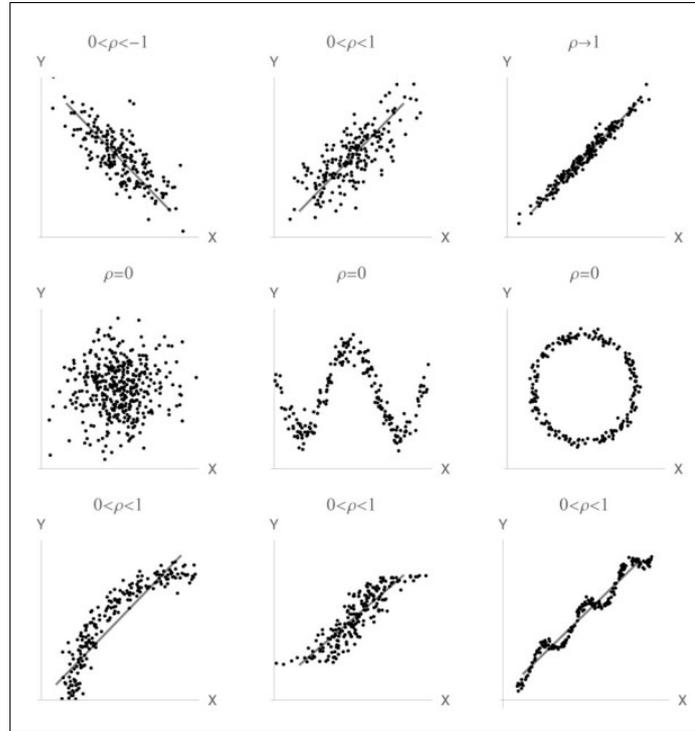
$$\text{Cov}[X, X] = E[(X - \mu_x)^2] = \text{Var}[X], \text{ conforme a Equação 9.} \quad (27)$$

O coeficiente de correlação ρ_{XY} , por sua vez, é uma medida adimensional da covariância entre duas variáveis aleatórias (IBE, 2014):

$$\rho_{XY} = \frac{\text{Cov}[X, Y]}{\sqrt{\text{Var}[X]\text{Var}[Y]}} = \frac{\sigma_{XY}}{\sigma_x \sigma_y}. \quad (28)$$

Apesar de a covariância implicar dependência entre as variáveis X e Y , covariância nula não implica que as variáveis sejam independentes entre si, uma vez que há possibilidade de dependências que não tenham comportamento linear, vide Figura 2 abaixo. (MELCHERS; BECK, 2018; BECK, 2019)

Figura 2 – Diferentes comportamentos de dependência entre variáveis aleatórias



Fonte: Beck, A.T. (2019)

Para um determinado vetor \mathbf{X} de variáveis aleatórias $\mathbf{X} = \{X_1, X_2, \dots, X_n\}$, escreve-se uma matriz de correlação \mathbf{R}_X , quadrada, simétrica e com elementos da diagonal principal iguais a 1, tal como na Equação 29:

$$\mathbf{R}_X = \begin{bmatrix} 1 & \rho_{X_1 X_2} & \cdots & \rho_{X_1 X_n} \\ \rho_{X_2 X_1} & 1 & \cdots & \rho_{X_2 X_n} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{X_n X_1} & \rho_{X_n X_2} & \cdots & 1 \end{bmatrix}. \quad (29)$$

2.1.2 Distribuição normal

Dentre todas as distribuições, tanto contínuas como discretas, talvez a mais importante, conhecida e utilizada seja a distribuição normal (BERGER, 1993; PAPOULIS; PILLAI, 2002; ROUSSAS, 2014). Definida completamente por sua média μ e desvio-padrão σ apenas, tem sua PDF dada por:

$$f_X(x) = \frac{1}{\sigma \cdot \sqrt{2\pi}} \cdot \exp\left(-\frac{1}{2} \cdot \left(\frac{x-\mu}{\sigma}\right)^2\right), \text{ com } -\infty \leq x \leq \infty, \quad (30)$$

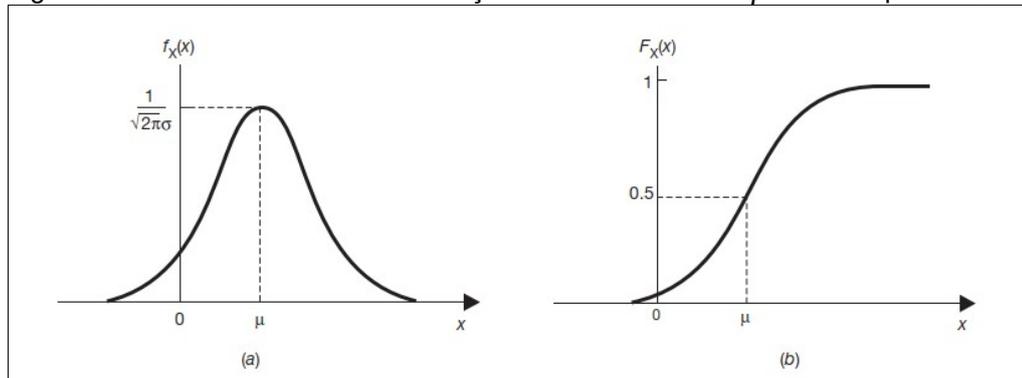
que mostra a importante característica de simetria em torno da média μ , ou seja, $f_X(\mu-x) = f_X(\mu+x)$.

Sua CDF é definida pela integral:

$$F_X(x) = \int_{-\infty}^x \frac{1}{\sigma \cdot \sqrt{2\pi}} \cdot \exp\left(-\frac{1}{2} \cdot \left(\frac{y-\mu}{\sigma}\right)^2\right) dy, \text{ com } -\infty \leq x \leq \infty, \quad (31)$$

que não possui solução analítica fechada, tendo, portanto, seus valores computados numericamente. Seu comportamento pode ser observado na Figura 3, em que se ilustra também a respectiva PDF.

Figura 3 – PDF e CDF de uma distribuição normal com média μ e desvio-padrão σ



Fonte: Hsu (2020)

Uma variável aleatória X com distribuição normal de parâmetros μ e σ pode ser escrita na seguinte notação: $X \sim N(\mu, \sigma)$. O caso particular $X \sim N(0, 1)$ é chamado de distribuição normal padrão.

De fato, se $X \sim N(\mu, \sigma)$, então:

$$Z = \frac{(X - \mu)}{\sigma} \quad (32)$$

é normal-padrão, de forma que afirmações acerca de uma determinada variável aleatória com distribuição normal podem ser obtidas a partir de afirmações sobre uma variável normal-padrão Z equivalente. Esta é a chamada transformação de Hasofer-Lind (1974).

A PDF e a CDF de Z têm, portanto, a seguinte forma:

$$f_Z(z) = \phi(z) = \frac{1}{\sqrt{2\pi}} \cdot \exp\left(-\frac{z^2}{2}\right), \text{ com } -\infty \leq z \leq \infty, \quad (33)$$

$$F_Z(z) = \Phi(z) = \int_{-\infty}^z \phi(y) dy, \text{ com } -\infty \leq z \leq \infty, \quad (34)$$

para os quais há valores tabelados. Nas equações 33 e 34 os termos $\phi(z)$ e $\Phi(z)$ são notações alternativas para a PDF e CDF de uma distribuição normal padrão.

A distribuição normal é notável pelo fato de adequadamente representar o comportamento de vários fenômenos naturais. Berger (1993) e Papoulis e Pillai (2002) relatam que, da Lei de Maxwell, na Física, sob condições apropriadas, os componentes da velocidade de uma molécula de gás serão normalmente distribuídos. Roussas (2014), por sua vez, menciona que a distribuição de pesos e alturas de uma determinada população humana; a renda anual de uma família; as notas de uma turma suficientemente grandes etc. podem, com boa aproximação, serem modeladas conforme uma distribuição normal. Soong (2004) ilustra esse fato com o exemplo de como o consumo de gasolina de carros de uma determinada marca e modelo, todos produzidos sob condições e processos idênticos, difere de unidade para unidade, tendo a aleatoriedade presente em cada componente em particular (advinda de variações na precisão de cada equipamento que produz os componentes, flutuações nos valores das resistências dos materiais, mudanças na qualidade da gasolina utilizada, bem como nos hábitos do motorista etc.) contribuído para as diferenças finais. Não obstante, o consumo tenderia a seguir uma distribuição aproximadamente normal. Ou seja, a aleatoriedade de um determinado fenômeno tende a seguir uma distribuição normal quando este é o resultado da contribuição de diversas pequenas aleatoriedades agindo simultaneamente.

Este é o caso geral de um teorema (Teorema do Limite Central) que afirma que sob condições bastante generalistas, a distribuição a qual se aproxima a média das contribuições de um número qualquer, suficientemente grande, de variáveis aleatórias independentes e identicamente distribuídas (*i.i.d* ou *iid*) é normal. (PAPOULIS; PILLAI, 2002). A expressão formal é a seguinte: seja S_1, S_2, \dots, S_n uma sequência de variáveis *iid*, com médias finitas $E[S_i] = \mu$ e variâncias $Var[S_i] = \sigma^2$, e seja $Z_n = S_1 + S_2 + \dots + S_n$, então o teorema diz que:

$$\lim_{n \rightarrow \infty} Pr\left(\frac{Z_n - n \cdot \mu}{\sigma \cdot \sqrt{n}} \leq x\right) = \Phi(x), \quad (35)$$

ou seja, para um número n suficientemente grande, Z_n possui distribuição aproximadamente normal com média $E[Z]=n\cdot\mu$ e variância $Var[Z]=n\cdot\sigma^2$, i.e. $Z\approx N(n\mu,\sqrt{n}\sigma)$.

2.1.3 Distribuição de Poisson

Nas engenharias, uma importante distribuição discreta é a de Poisson. Ela descreve a probabilidade de k ocorrências (também referidas como “chegadas”), sendo k um número natural, de um determinado evento em um determinado intervalo de interesse de um contínuo (tempo, espaço etc.) dado que este evento tenha uma quantidade de ocorrências λ que se mantém constante ao longo deste intervalo.

Uma distribuição de Poisson representa o número de ocorrências de um evento raro num grande número de tentativas; nisto, ela é o limite de um experimento de Bernoulli quando o número de ensaios independentes tende ao infinito e quando os intervalos tendem a zero (TAYLOR; KARLIN, 1994; DANTAS, 2013; BECK, 2019). A distribuição de Poisson com parâmetro λ toma a forma a seguir:

$$P(X=k)=\frac{e^{-\lambda}\cdot\lambda^k}{k!}, k=0,1,2,\dots \quad (36)$$

O parâmetro λ também pode ser escrito na forma

$$\lambda=r\cdot t, \quad (37)$$

caso em que r se refere, então, à taxa com que as ocorrências acontecem num determinado intervalo de tamanho t do contínuo, também conhecida como taxa de chegadas ou taxa de ocorrências (BECK, 2019). Assim, uma apresentação alternativa se dá como presente a Equação 38:

$$P(k \text{ ocorrências no intervalo } t)=\frac{e^{-r\cdot t}\cdot(r\cdot t)^k}{k!}, k=0,1,2,\dots \quad (38)$$

O exemplo trazido por Yates (2005) ilustra bem esta interpretação: se um evento tem taxa de ocorrência $r=2$ chegadas/segundo, a probabilidade de k chegadas em um intervalo $t=0,25$ segundos é dada pela distribuição de Poisson cujo $\lambda=2\left(\frac{\text{chegadas}}{\text{segundo}}\right)\times 0,25(\text{segundos})=0,5$ chegadas. Ou seja:

$$P(X=k) = \frac{e^{-rt}(rt)^k}{k!} = \frac{e^{-0,5} \cdot (0,5)^k}{k!}, k=0,1,2,\dots$$

Sua CDF é, naturalmente, dada por:

$$F_{X \text{ Poisson}}(x) = P(X \leq x) = \sum_{k=0}^x \frac{e^{-\lambda} \cdot \lambda^k}{k!}, k=0,1,2,\dots \quad (39)$$

Uma variável aleatória com distribuição de Poisson com parâmetro λ tem média e variância iguais a, respectivamente:

$$\begin{aligned} \mu &= \lambda = r \cdot t \\ \sigma^2 &= \lambda = r \cdot t \end{aligned}$$

Uma importante propriedade da distribuição de Poisson concerne à variável aleatória resultante da soma de duas outras variáveis aleatórias X e Y , independentes, distribuídas também conforme Poisson. Sejam X e Y estas variáveis aleatórias independentes, com parâmetros μ e ν , respectivamente, então pela Lei da Probabilidade Total (BECK, 2019), tem-se que a probabilidade de que a soma de X e Y assumam um valor específico n é dada por:

$$P(X+Y=n) = \sum_{k=0}^n P(X=k, Y=n-k) = \sum_{k=0}^n P(X=k) \cdot P(Y=n-k), \quad (40)$$

uma vez que X e Y são independentes.

Assim:

$$P(X+Y=n) = \sum_{k=0}^n \left\{ \frac{\mu^k \cdot e^{-\mu}}{k!} \right\} \cdot \left\{ \frac{\nu^{(n-k)} \cdot e^{-\nu}}{(n-k)!} \right\}, \quad (41)$$

que, multiplicando e dividindo por $n!$, resulta em:

$$P(X+Y=n) = \frac{e^{-(\mu+\nu)}}{n!} \cdot \sum_{k=0}^n \frac{n!}{k! \cdot (n-k)!} \cdot \mu^k \cdot \nu^{(n-k)}, \quad (42)$$

onde o termo:

$$\sum_{k=0}^n \frac{n!}{k! \cdot (n-k)!} \cdot \mu^k \cdot \nu^{(n-k)}, \text{ é a expansão binomial de } (\mu+\nu)^n. \quad (43)$$

Desta forma, a Equação 42 pode ser reescrita como:

$$P(X+Y=n) = \frac{e^{-(\mu+\nu)} \cdot (\mu+\nu)^n}{n!}, \text{ com } n=0,1,2,\dots, \quad (44)$$

que, observando a Equação 36, é justamente uma distribuição de Poisson com parâmetro λ igual a $(\mu+\nu)$.

Ou seja, tem-se que a soma de variáveis independentes distribuídas conforme Poisson é, em si, distribuída conforme Poisson. Ademais, o parâmetro desta variável aleatória é a soma dos parâmetros dos termos individuais que a compõem. (TAYLOR; KARLIN, 1994; BEICHELT et al., 2002)

2.1.4 Distribuição Exponencial

Se a distribuição de Poisson descreve a probabilidade de uma determinada quantidade k de eventos ocorrer em um determinado intervalo, a variável aleatória que descreve os intervalos entre ocorrência destes eventos pode ser modelada pela distribuição Exponencial. Ou seja: tanto o tempo (e o mesmo valor para outros contínuos análogos) até a primeira ocorrência, T_1 , como o tempo *entre* ocorrências, T , seguem a distribuição Exponencial de média e desvio-padrão iguais a $1/r$, em que r tem a mesma interpretação que na distribuição de Poisson: a taxa de ocorrências de um evento dentro de um determinado intervalo.

Assim, seja a variável aleatória T , o tempo até primeira ocorrência, distribuída exponencialmente. Sua PDF e CDF são, representadas matematicamente como:

$$F_T(t) = P(T \leq t) = 1 - e^{-r \cdot t}, \quad (45)$$

$$f_T(t) = \frac{dF_T(t)}{dt} = r \cdot e^{-r \cdot t}. \quad (46)$$

O resultado de $F_T(t)$ pode ser facilmente obtido ao computar-se a probabilidade de o evento $P(T > t)$ ocorrer por meio da distribuição de Poisson. Ora, $P(T > t)$ implica *nenhuma* ocorrência no intervalo de comprimento t , assim, da Equação 38, tem-se:

$$P(T > t) = P(k=0 \text{ ocorrências no intervalo } t) = \frac{e^{-r \cdot t} \cdot (r \cdot t)^0}{0!} = e^{-r \cdot t}, \quad (47)$$

de tal forma que:

$$F_T(t) = P(T \leq t) = 1 - e^{-r \cdot t}, \text{ com } t \geq 0. \quad (48)$$

Uma importante característica da distribuição Exponencial é a chamada falta de memória, sendo a única distribuição contínua a possuir tal propriedade (BEICHELT et al., 2002; KULKARNI, 2017). A falta de memória da distribuição diz respeito a como, caso o contínuo em questão diga respeito ao tempo, por exemplo, e a variável T ao tempo de vida de um determinado item, se este sobreviveu por t

unidades de tempo, então sua vida-útil restante (i.e., $[t, \infty)$) é a mesma que a de um item novo. (TAYLOR; KARLIN, 1994; BEICHELDT et al., 2002)

Assim, se considerarmos $s, t \geq 0$, a probabilidade de que o item sobreviva (isto é, de que seu tempo de vida T) seja maior que $(t+s)$ uma vez que já sobreviveu até o tempo t é dada pela probabilidade condicional $P(T > t+s | T > t)$. Da expressão básica para esta, tem-se:

$$P(T > t+s | T > t) = \frac{P(T > t+s \cap T > t)}{P(T > t)} = \frac{P(T > t+s)}{P(T > t)}, \quad (49)$$

uma vez que $\{T > t+s\} \subset \{T > t\}$.

Portanto:

$$P(T > t+s | T > t) = \frac{e^{-r \cdot (t+s)}}{e^{-r \cdot t}} = e^{-r \cdot s} = P(T > s). \quad (50)$$

Finalmente, seja T_n a variável aleatória que descreve o tempo de chegada da n -ésima ocorrência. Uma vez que $T_n \leq t$ é um evento que acontece se e somente se $N(t) > n$ (o número de chegadas, distribuída conforme Poisson, até o tempo t) então:

$$P(T_n \leq t) = P(N(t) \geq n). \quad (51)$$

De forma que:

$$F_{T_n}(t) = P(N(t) \geq n) = \sum_{i=n}^{\infty} \frac{(r \cdot t)^i}{i!} \cdot e^{-(r \cdot t)}, \text{ com } n=1, 2, \dots \text{ e PDF:} \quad (52)$$

$$f_{T_n}(t) = \frac{(r \cdot t)^{(n-1)}}{(n-1)!} \cdot r \cdot e^{-(r \cdot t)}, \text{ também notado como } f_{T_n}(t) = \frac{r^n \cdot t^{(n-1)}}{(n-1)!} \cdot e^{-(r \cdot t)}. \quad (53)$$

é distribuído segundo uma distribuição Erlang com parâmetros n e r . (BEICHELDT et al., 2002)

2.1.5 Função de uma variável aleatória

Seja $X(\omega)$ uma variável aleatória, então esta é uma função tal que a cada evento ω_i é atribuído um valor real $x(\omega_i)$. A imagem desta função é referida como I_X . Se $g(X)$ é definida $\forall x \in I_X$, então $Y = g(X)$ é uma variável aleatória:

$$Y(\omega) = g(X(\omega)) \quad (54)$$

Com momentos:

$$E[Y] = \int_{-\infty}^{\infty} y f_y(y) dy, \text{ que leva a } E[g(X)] = \int_{-\infty}^{\infty} g(x) f_x(x) dx \quad (55)$$

$$Var[Y] = E[(Y - E[Y])^2] = \int_{-\infty}^{\infty} (g(x) - E[g(X)])^2 f_x(x) dx \quad (56)$$

que são grandezas possíveis de serem determinadas mesmo quando da ausência de informações diretas a respeito da distribuição de probabilidades de Y . (DEVORE, 2016; JOHNSON; MILLER; FREUND, 2017; BECK, 2019)

Para o caso de $X(\omega)$ ser uma variável aleatória discreta, tem-se que $Y(\omega)$ também é discreta e tem momentos iguais a:

$$E[Y] = \sum_{y_i \in A} y_i p(y_i), \text{ que leva a } E[g(X)] = \sum_{y_i \in A} g(x_i) p(x_i) \quad (57)$$

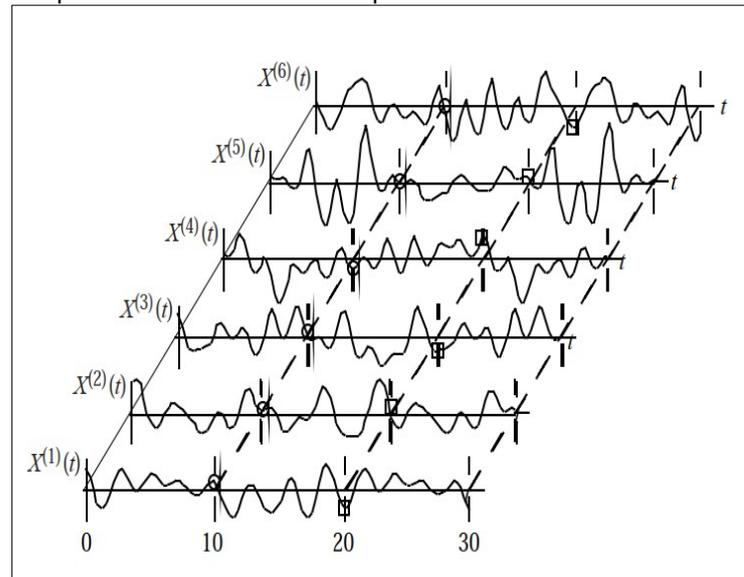
$$Var[Y] = E[Y^2] - E[Y]^2, \text{ que leva a } Var[g(X)] = E[g(X)^2] - E[g(X)]^2 \quad (58)$$

2.2 PROCESSOS ESTOCÁSTICOS

Se variáveis aleatórias são funções que atribuem a cada ponto amostral ω_i de um determinado experimento um valor $x(\omega_i)$, no caso de processos estocásticos, temos uma generalização tal que cada ponto amostral leva a uma função real $x(t, \omega_i)$, com $t \in T$, definida num espaço de probabilidade comum. Assim sendo, o processo estocástico $X(\omega, t)$ pode ser entendido como uma família (ou coleção) de funções com t representando um determinado parâmetro determinístico, geralmente o tempo ou alguma dimensão espacial, mas não limitado a estes exemplos. (KULKARNI, 2017; BECK, 2019).

Um processo estocástico é tal que, para cada realização ω_i fixa, $x(t, \omega_i) = x(t)$ é uma função que depende apenas dos valores particulares que t assume; em um ponto t_i fixo, $X(\omega, t_i) = X(\omega)$ é uma variável aleatória; com ω_i e t_i fixos, $x(\omega_i, t_i) = x$ é um número, (PARZEN, 1999; PAPOULIS; PILLAI, 2002; ELISHAKOFF, 2017). Em geral, no caso de t_i fixo, a variável aleatória $X(\omega)$ é denotada apenas como $X(t_i)$ (a variável aleatória para aquele determinado valor do parâmetro), omitindo-se menção a ω . A Figura 4, que emprega a notação $X^{(\omega_i)}(t)$, mostra como em cada realização o processo é uma função parâmetro, neste caso contínuo, t e como em cada ponto t_i o processo é uma variável aleatória.

Figura 4 – Desenho esquemático de várias realizações ω_i de um processo estocástico com parâmetro contínuo t



Fonte: Lutes, L.D.; Sarkani, S. (2004)

Se T , o conjunto de valores que o parâmetro t assume, é discreto, diz-se que o processo estocástico é de parâmetro discreto e que $X(\omega, t_1), X(\omega, t_2), \dots, X(\omega, t_n)$ é, portanto, uma sequência contável de variáveis aleatórias. Por outro lado, se T é o conjunto dos reais, \mathbb{R} , diz-se que este é um processo estocástico de parâmetro contínuo e que a coleção de variáveis aleatórias definidas para cada ponto t_i é incontável (HOEL et al., 1972). Seja Z o conjunto de valores que $x(t, \omega_i) = x(t)$ pode assumir, chamado doravante de espaço de estado (*state space*), diz-se que este espaço é discreto se os valores que o processo pode assumir são contáveis e contínuo caso contrário. (BEICHELT et al., 2002; PAPOULIS; PILLAI, 2002)

Os exemplos ilustrativos trazidos por Dobrow (2016) são particularmente úteis para a correta compreensão do conceito de processo estocástico. O primeiro exemplo, aquele de um processo estocástico de parâmetro discreto mas de espaço de estado contínuo, é o de uma estação de monitoramento do ar que registra a concentração X_k de poluentes hora a hora. Os valores registrados em uma determinada hora k do dia (o parâmetro discreto) ao longo de, por exemplo, vários dias (realizações) assumem valores contínuos; o tempo em que esses registros acontecem, no entanto, é sempre o mesmo todos os dias, logo, discreto.

O segundo exemplo diz respeito a um processo estocástico de parâmetro contínuo mas de estado de espaço discreto: o número X_t de mensagens recebidas

por uma pessoa desde um momento t até outro. O número de mensagens (estado de espaço) é discreto, mas o momento em que se faz a medição (realização) pertence ao intervalo $[0, \infty)$.

2.2.1 Funções de distribuição de probabilidade e os momentos de um processo estocástico

Tendo sido estabelecido que para um valor fixo do parâmetro t o processo estocástico $X(\omega, t)$ é uma variável aleatória, tem-se que sua distribuição de probabilidades dependem, em geral, de t :

$$F_X(x, t) = P(X(\omega, t) \leq x). \quad (59)$$

Desta maneira, para um valor fixo de t ($t = \{t_1\}$), a CDF da variável aleatória $X(\omega)$ é:

$$F_X(x, t_1) = P(X(\omega, t_1) \leq x), \text{ com PDF:} \quad (60)$$

$$f_x(x, t_1) = \frac{\partial F_X(x, t_1)}{\partial x}, \quad (61)$$

em que o evento em questão consiste em todas aquelas realizações $\omega_1, \omega_2, \dots, \omega_n$ em que os números $x(\omega_1, t_1), x(\omega_2, t_1), \dots, x(\omega_n, t_1)$ não ultrapassam o valor x . Estas CDFs e PDFs são chamadas de funções (cumulativa e de distribuição, respectivamente) de primeira ordem do processo, podendo ser notadas como F_1 e f_1 , ou até $F_{1,x}$ e $f_{1,x}$. A CDF e PDF conjuntas do processo estocástico em dois pontos t , as funções de segunda ordem são (IBE, 2014; ELISHAKOFF, 2017; BECK, 2019):

$$F_X(x_1, x_2; t_1, t_2) = P(\{X(\omega, t_1) \leq x_1, X(\omega, t_2) \leq x_2\}), \text{ com PDF:} \quad (62)$$

$$f_x(x_1, x_2; t_1, t_2) = \frac{\partial^2 F_X(x_1, x_2; t_1, t_2)}{\partial x_1 \partial x_2}, \quad (63)$$

que leva à generalização para um conjunto finito de tamanho n de valores para o parâmetro t , dada por:

$$F_X(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n) = P(\{X(\omega, t_1) \leq x_1, X(\omega, t_2) \leq x_2, \dots, X(\omega, t_n) \leq x_n\}) \quad (64)$$

a chamada CDF de ordem n do processo que o caracteriza completamente (BECK, 2019). A função de distribuição de probabilidades é:

$$f_x(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n) = \frac{\partial^n F_X(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n)}{\partial x_1 \partial x_2 \dots \partial x_n}, \quad (65)$$

podendo, a CDF e a PDF, ser notadas como F_n e f_n ou $F_{n,x}$ e $f_{n,x}$ respectivamente.

Tendo sido estabelecido que para um ponto t fixo o processo estocástico é uma variável aleatória, notada doravante $X(t)$, então, assumindo a existência de $E[X(t)], \forall t \in T$, a tendência (*trend*) ou função de tendência (*trend function*) do processo estocástico é o valor esperado deste processo como função de t (também entendido como desenvolvimento médio do processo) (BEICHELT et al., 2002):

$$\mu_x(t) \equiv E[X(t)], \text{ com } t \in T. \quad (66)$$

Desta forma, a função de tendência, $\mu_x(t)$, descreve o desenvolvimento médio do processo. Se a PDF

$$f_x(t) = \frac{dF_x(t)}{dx}, \forall t \in T, \quad (67)$$

existe, então (BEICHELT et al., 2002):

$$\mu_x(t) = \int_{-\infty}^{\infty} x f_x(x, t) dx. \quad (68)$$

Generalizando o resultado, tem-se que o momento de ordem k da variável $X(t)$ é dado por:

$$E[X(t)^k] \equiv \int_{-\infty}^{\infty} x^k f_x(x, t) dx = \mu_x^k(t). \quad (69)$$

Se o processo estocástico tem $E[X(t)^2] \leq \infty, \forall t \in T$, então ele é chamado de processo de segunda ordem. (IBE, 2014)

Ao momento conjunto das duas variáveis aleatórias $X(t_1)$ e $X(t_2)$ dá-se o nome de função de autocorrelação do processo $X(\omega, t)$, representada por $R_{XX}(t_1, t_2)$, uma medida da similaridade do processo em dois pontos do contínuo t (IBE, 2014; BECK, 2019):

$$E[X(t_1)X(t_2)] \equiv \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 x_2 f_x(x_1, x_2; t_1, t_2) dx_1 dx_2 = R_{XX}(t_1, t_2), \quad (70)$$

que é geralmente uma função de t_1 e t_2 . A função de autocorrelação basicamente define o quanto um determinado processo estocástico é similar a si mesmo quando transladado de $\tau = t_2 - t_1$. (IBE, 2014)

A covariância entre $X(t_1)$ e $X(t_2)$, chamada de *função de autocovariância*, é dada por (ELISHAKOFF, 2017):

$$\begin{aligned} C_{XX}(t_1, t_2) &\equiv E[(X(t_1) - \mu_X(t_1))(X(t_2) - \mu_X(t_2))] \\ " " &= R_{XX}(t_1, t_2) - \mu_X(t_1)\mu_X(t_2) \end{aligned} \quad (71)$$

de forma tal que a variância de $X(t)$ é dada fazendo-se $t_1=t_2=t$ (BECK, 2019).

$$\begin{aligned} \text{Var}[X(t)] &= \sigma_X^2(t) = C_{XX}(t, t) = R_{XX}(t, t) - \mu_X(t)\mu_X(t) \\ " " &= R_{XX}(t, t) - \mu_X(t)^2 \end{aligned} \quad (72)$$

Se $X(t_1)$ e $X(t_2)$ são independentes, então $R_{XX}(t_1, t_2) = \mu_X(t_1)\mu_X(t_2)$ (o produto das médias do processo nos tempos t_1 e t_2), o que leva $C_{XX}(t_1, t_2)$ a ser igual a zero. (IBE, 2014). O mesmo é válido para dois processos distintos $X(\omega, t)$ e $Y(\omega, t)$ em dois pontos t_1 e t_2 do parâmetro contínuo t ; ao mesmo tempo, dois processos são ditos independentes *de maneira genérica* quando o conjunto $X(t_1), X(t_2), \dots, X(t_n)$ é independente do conjunto $Y(s_1), Y(s_2), \dots, Y(s_n)$ para quaisquer valores t_1, t_2, \dots, t_n e s_1, s_2, \dots, s_n . (IBE, 2014; BECK, 2019)

O índice de autocorrelação de um determinado processo $X(\omega, t)$ é definido de maneira similar àquela para variáveis aleatórias:

$$\rho_{XX}(t_1, t_2) = \frac{C_{XX}(t_1, t_2)}{\sigma_X(t_1)\sigma_X(t_2)} \quad (73)$$

As funções $R_{XX}(t_1, t_2)$, $C_{XX}(t_1, t_2)$ e $\rho_{XX}(t_1, t_2)$ caracterizam completamente a estrutura de correlações de um processo estocástico. (BECK, 2019)

2.2.2 Funções de distribuição de probabilidade e os momentos de dois processos estocásticos

Para o caso de dois processos estocásticos $X(t)$ e $Y(t)$, define-se a função de autocorrelação-cruzada como:

$$E[X(t_1)Y(t_2)] \equiv \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x y f_{xy}(x, y; t_1, t_2) dx dy = R_{XY}(t_1, t_2), \quad (74)$$

com função de covariância-cruzada dada por (ELISHAKOFF, 2017; BECK, 2019):

$$\begin{aligned} C_{XY}(t_1, t_2) &\equiv E[(X(t_1) - \mu_X(t_1))(Y(t_2) - \mu_Y(t_2))] \\ " " &= R_{XY}(t_1, t_2) - \mu_X(t_1)\mu_Y(t_2) \end{aligned} \quad (75)$$

Assim, tem-se as matrizes de correlação e de covariância-cruzada C e R , respectivamente:

$$\mathbf{C} = \begin{bmatrix} C_{XX}(t_1, t_2) & C_{XY}(t_1, t_2) \\ C_{YX}(t_1, t_2) & C_{YY}(t_1, t_2) \end{bmatrix} \text{ e } \mathbf{R} = \begin{bmatrix} R_{XX}(t_1, t_2) & R_{XY}(t_1, t_2) \\ R_{YX}(t_1, t_2) & R_{YY}(t_1, t_2) \end{bmatrix}. \quad (76)$$

que são simétricas, ou seja:

$$R_{XY}(t_1, t_2) = R_{YX}(t_2, t_1), \text{ o que é igualmente válido para os elementos de } \mathbf{C} \quad (77)$$

2.2.3 Estacionariedade

Processos estocásticos podem ser classificados de várias maneiras e baseados em diversos critérios; um dos mais usuais diz respeito a como suas características variam ao longo do parâmetro contínuo t : se ele é estacionário ou não. É intuitivo, portanto, concluir que um determinado processo estocástico $X(\omega, t)$ estacionário é aquele cujas características estatísticas não sofrem mudanças no contínuo (PISHRO-NIK, 2014). Esta propriedade sempre se refere a um determinado aspecto específico do processo que se mantém inalterada ao longo dos valores do parâmetro t , havendo vários tipos de estacionariedade a depender de qual propriedade ou característica do processo mantém-se inalterada. (LUTES; SARKANI, 2004)

Mais especificamente, caracteriza-se um processo estocástico como estritamente estacionário (*strictly stationary*) quando suas propriedades estatísticas são invariantes a uma translação no contínuo, ou seja, quando sua CDF é invariante quanto à origem deste contínuo (IBE, 2014; BECK, 2019). Assim, $X(t)$ e $X(t+\varepsilon)$ são idênticos para qualquer ε arbitrário. Formalmente, um processo estocástico é estritamente estacionário quando para todos os $n=1, 2, \dots$ inteiros positivos, para qualquer h e para todas as n -tuplas (t_1, t_2, \dots, t_n) com $t_i \in T$ e $t_i+h \in T$, bem como para todas as n -tuplas (x_1, x_2, \dots, x_n) (BEICHEL et al., 2002; BECK, 2019):

$$F_{n,X}(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n) = F_{n,X}(x_1, x_2, \dots, x_n; t_1+h, t_2+h, \dots, t_n+h). \quad (78)$$

Para a PDF de primeira-ordem do processo estacionário:

$$f_X(x, t) = f_X(x, t+h). \quad (79)$$

Se isto é verdade para todos os valores de h , então também o é para $h=-t$.

Assim:

$$f_X(x, t) = f_X(x, 0), \quad (80)$$

de tal forma que $f_X(x, t)$ é independente de t , o valor absoluto para o parâmetro contínuo, logo:

$$f_X(x,t) = f_X(x), \forall t \in T. \quad (81)$$

Isto leva a média μ_X a também ser independente de t , tornando a Equação 68 constante ao longo de t :

$$\mu_X(t) = \int_{-\infty}^{\infty} x f_X(x) dx = \mu_X \rightarrow \text{constante}, \quad (82)$$

o mesmo sendo válido para a variância em $t_1 = t_2 = t$: $\text{Var}[X(t)] = \sigma_X^2 \rightarrow \text{constante}$. (ELISHAKOFF, 2017)

Se a Equação 78 é verdadeira para qualquer n , também o é para $n=2$, assim, a PDF de segunda ordem de um processo estacionário é tal que satisfaz a condição abaixo:

$$f_{2,X}(x_1, x_2; t_1, t_2) = f_{2,X}(x_1, x_2; t_1 + h, t_2 + h). \quad (83)$$

Procedendo de maneira similar e fazendo $h = -t_1$, tem-se:

$$f_{2,X}(x_1, x_2; t_1, t_2) = f_{2,X}(x_1, x_2; 0, t_2 - t_1) = f_{2,X}(x_1, x_2, \tau), \quad (84)$$

ao chamarmos $t_2 - t_1 = \tau$ (ELISHAKOFF, 2017).

Ou seja, a distribuição conjunta (das VAs $X(t)$ e $X(t+\tau)$) de um processo estritamente estacionário é tal que depende dos valores x_1 e x_2 e de apenas uma grandeza referente ao parâmetro contínuo (i.e. τ), sendo irrelevantes t_1 e t_2 . Nos contextos em que t diz respeito a uma dimensão temporal do problema, esta quantidade τ também é conhecida como tempo de atraso (*lag time* ou *delay time*). (IBE, 2014)

Para a autocorrelação de um processo estritamente estacionário, obtemos que esta também depende apenas de τ (BECK, 2019):

$$\begin{aligned} R_{XX}(t_1, t_2) &= E[X(t_1)X(t_2)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 x_2 f_X(x_1, x_2; t_1, t_2) dx_1 dx_2 \\ " " &= E[X(t)X(t+\tau)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 x_2 f_X(x_1, x_2, \tau) dx_1 dx_2 = R_{XX}(\tau) \end{aligned} \quad (85)$$

↓

$$\therefore R_{XX}(t_1, t_2) = R_{XX}(t, t+\tau) = R_{XX}(\tau)$$

com o mesmo sendo válido para a autocovariância (ELISHAKOFF, 2017; BECK, 2019):

$$\begin{aligned}
C_{XX}(t_1, t_2) &= R_{XX}(t_1, t_2) - \mu_X(t_1)\mu_X(t_2) \\
" " &= R_{XX}(\tau) - \mu_X^2 &= C_{XX}(\tau)
\end{aligned}
\tag{86}$$

↓

$$\therefore C_{XX}(t_1, t_2) = C_{XX}(t, t+\tau) = C_{XX}(\tau)$$

Finalmente, um processo estocástico é dito fracamente estacionário (*weakly stationary*) ou estacionário em sentido amplo (*wide sense stationary*) quando sua média é constante, $\mu_X(t) = \mu_X$, e sua covariância depende apenas de τ , $C_{XX}(t_1, t_2) = C_{XX}(\tau)$ (ELISHAKOFF, 2017), ainda que este não cumpra a condição da Equação 78 (BECK, 2019). Todo processo estocástico estritamente estacionário, portanto, é também fracamente estacionário, mas a recíproca não necessariamente é verdadeira. (IBE, 2014; ELISHAKOFF, 2017; BECK, 2019)

2.2.4 Estatísticas do contínuo e ergodicidade

Todas estas grandezas estatísticas estudadas até aqui, tomadas sob várias realizações ω em um determinado ponto t , são chamadas de estatística de valor esperado ou estatísticas de envelope. (BECK, 2019). Um processo estacionário cujas estatísticas de envelope num determinado ponto t são iguais àquelas tomadas ao longo de todos os valores $t \in T$, as chamadas estatísticas do contínuo, de uma única realização ω_i do processo é chamado de processo ergódico. (IBE, 2014; MELCHERS; BECK, 2018; BECK, 2019)

Formalmente: seja ω_i uma realização de um processo estacionário tal que, como já previamente discutido, $X(\omega_i, t) = x(t)$ é uma função do parâmetro t e que $U(\omega_i) = u$, constante, é a média desta função; a autocorrelação $R_{XX}(\omega_i, \tau) = r(\tau)$ é função de τ . Obviamente, tomadas várias realizações ω_i , nota-se que tanto $U(\omega)$ quanto $R_{XX}(\omega)$ são variáveis aleatórias. Temos que, tomando o limite para t suficiente grande, as estatísticas do contínuo (chamadas de estatísticas temporais, caso t se refira ao tempo) são dadas por (IBE, 2014; MELCHERS; BECK, 2018; BECK, 2019):

$$U(\omega) \equiv \lim_{t \rightarrow \infty} \frac{1}{2t} \int_{-t}^t X(\omega, t) dt, \tag{87}$$

$$R_{XX}(\omega, \tau) \equiv \lim_{t \rightarrow \infty} \frac{1}{2t} \int_{-t}^t X(\omega, t) X(\omega, t+\tau) dt. \tag{88}$$

Chama-se de estritamente ergódico aquele processo cujas estatísticas de envelope podem ser determinadas a partir de uma única realização $X(\omega_i, t) = x(t)$ do processo. (IBE, 2014; BECK, 2019)

2.2.5 Função de autocorrelação: propriedades e características

2.2.5.1 Propriedades

A função de autocorrelação $R_{XX}(t_1, t_2) = R_{XX}(\tau)$ de um processo estocástico tem as seguintes propriedades (IBE, 2014; BECK, 2019):

- A função de autocorrelação é par: $R_{XX}(\tau) = R_{XX}(-\tau)$;
- O máximo valor absoluto de $R_{XX}(\tau)$ ocorre na origem, ou seja, em $\tau = 0$;
- O máximo valor da função de autocorrelação é igual ao segundo momento do processo em t . A quantidade $R_{XX}(0) = E[X(t)^2]$ é usualmente chamada de valor médio quadrático (*mean-square value*);
- Se $R_{XX}(\tau)$ é contínua na origem, consequentemente é contínua em τ ;
- $R_{XX}(\tau) > 0 \quad \forall \tau \in \mathbb{R}^+$;
- $R_{XX}(\tau)$ não pode ter um formato arbitrário;

2.2.5.2 Relação com a função de densidade espectral de potência

Uma das maneiras pelas quais se pode analisar um determinado processo estocástico é através da chamada representação integral de Fourier: a decomposição do processo estocástico em somas infinitas de senos e cossenos, cada uma ocorrendo com amplitude aleatória, mas com frequência ω^1 constante (MELCHERS; BECK, 2018). Assim, o processo $X(t)$ é expresso como:

$$X(t) = \mu + \sum_{k=-K}^K X_k(t), \text{ em que:} \quad (89)$$

$$X_k(t) = V_k \cos(\omega_k t + \Phi_k).$$

Onde:

- μ é a média do processo;
- K é o número de senoides com os quais se discretiza o processo;
- ω_k são as frequências;

1 Não confundir com a notação para uma realização qualquer, presente tanto na seção 2.1 como nesta até então.

- V_k são as amplitudes aleatórias;
- Φ_k , com $k=1, 2, \dots, K$ são os ângulos de fase;
- $X_k(t)$ são funções elementares com média nula e variância igual a $E[V_k^2]/2$.

Todas as amplitudes são aleatórias e os ângulos de fase são independentes, sendo estes últimos uniformemente distribuídos no intervalo $(0, 2\pi)$. (BECK, 2019)

Chama-se função de densidade espectral $S_X(\omega)$ de um determinado processo estocástico estacionário a transformada de Fourier de sua função de autocorrelação $R_{XX}(\tau)$ (ELISHAKOFF, 2017):

$$S_{XX}(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} R_{XX}(\tau) e^{-i\omega\tau} d\tau, \quad (90)$$

e, pela inversão de Fourier, expressa-se $R_{XX}(\tau)$ como:

$$R_{XX}(\tau) = \int_{-\infty}^{\infty} S_{XX}(\omega) e^{i\omega\tau} d\omega, \quad (91)$$

que são os pares de Wiener-Khinchine. (PAPOULIS; PILLAI, 2002; ELISHAKOFF, 2017; BECK, 2019)

Uma outra representação destas expressões é a seguinte (MELCHERS; BECK, 2018; BECK, 2019):

$$S_{XX}(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} R_{XX}(\tau) \cos(\omega\tau) d\tau, \quad (92)$$

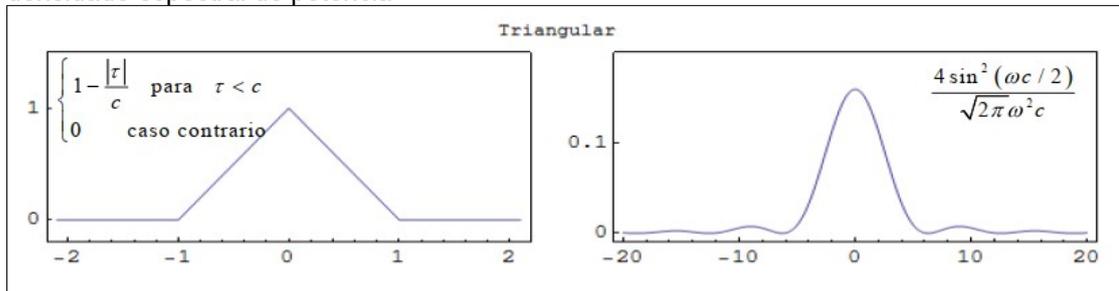
$$R_{XX}(\tau) = \int_{-\infty}^{\infty} S_{XX}(\omega) \cos(\omega\tau) d\omega. \quad (93)$$

Esta parte teórica é particularmente empregada na análise de séries temporais, onde a densidade espectral é uma ferramenta para determinação de mecanismos geradores de uma certa série temporal, e na teoria estatística de comunicações, campo no qual é usual o estudo de sinais os (processos estocásticos) passando por filtros (PARZEN, 1999). Uma outra aplicação desta formulação é o estudo de vibrações aleatórias em sistemas mecânicos elásticos e com comportamentos lineares; sistemas estruturais sensíveis a respostas dinâmicas impostas a partir de carregamentos apenas adequadamente modelados de maneira aleatória, como o vento ou ondas do mar. (MELCHERS; BECK, 2018)

2.2.5.3 Expressões usuais

Alguns pares $R_{XX}(\tau), S_{XX}(\omega)$ são comumente utilizados na modelagem de processos estocásticos. De particular importância para este trabalho são as funções de autocorrelação triangular e exponencial quadrada, ilustradas respectivamente na Figura 5 e na Figura 6 com suas respectivas funções de densidade espectral de potência.

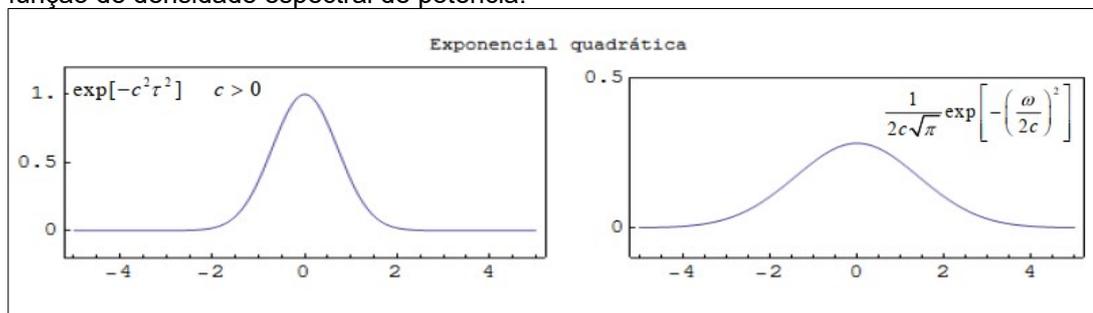
Figura 5 – Função de autocorrelação triangular (esquerda) e sua respectiva função de densidade espectral de potência



Fonte: Beck, A.T. (2019)

Nas equações de ambas as figuras, τ tem a mesma interpretação que nos parágrafos anteriores (i.e., a distância, no contínuo, entre pontos de interesse do processo estocástico), enquanto c é uma espécie de “fator de forma” da curva chamado de comprimento de correlação (BECK, 2019). Estes dois argumentos são suficientes para descrever todo o comportamento das autocorrelações descritas pelas expressões das Figuras 5 e 6.

Figura 6 – Função de autocorrelação exponencial quadrática (esquerda) e sua respectiva função de densidade espectral de potência.



Fonte: Beck, A.T. (2019)

2.2.6 Processo de pontos de Poisson

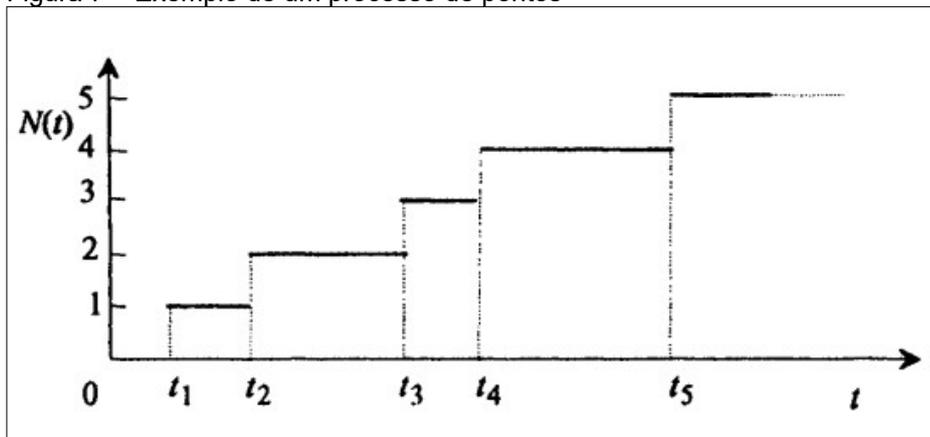
Seja $N(t)$, com $t \geq 0$, o processo estocástico definido pela função degrau de incrementos unitários que descreve o número de ocorrências de um determinado

evento. Esta função é dita um processo de pontos se obedecer às seguintes condições (BEICHELT et al., 2002):

- $N(s) \leq N(t)$ para $s \leq t$;
- Para quaisquer s, t com $s < t$, $N(t) - N(s)$ é o número de eventos que ocorre no intervalo $(s, t]$.

A Figura 7 ilustra uma realização de um processo de pontos:

Figura 7 – Exemplo de um processo de pontos



Fonte: Beichelt, F.; Fatti, L.P. (2002)

Este processo de pontos é dito um processo de pontos de Poisson homogêneo com intensidade r estacionária se (TAYLOR; KARLIN, 1994; BEICHELT et al., 2002):

- $N(0) = 0$;
- Para quaisquer pontos no contínuo $t_0 = 0 < t_1 < t_2 < \dots < t_n$, os incrementos $N(t_1) - N(t_0), N(t_2) - N(t_1), \dots, N(t_n) - N(t_{n-1})$ são variáveis aleatórias independentes e designam o número de ocorrências no intervalo $(t_1, t_2]$;
- Os incrementos do processo no intervalo $[s, t]$, com $s < t$ são distribuídos conforme Poisson com parâmetro $r \cdot (t - s)$:

$$P(N(t) - N(s) = k) = \frac{[r \cdot (t - s)]^k}{k!} \cdot e^{-r \cdot (t - s)}. \quad (94)$$

Ou, de maneira equivalente, fazendo $\tau = t - s$ (BEICHELT et al., 2002):

$$P(N(s + \tau) - N(s) = k) = \frac{(r \cdot \tau)^k}{k!} \cdot e^{-r \cdot \tau}. \quad (95)$$

O termo “estacionária”, utilizado para caracterizar a intensidade do processo de Poisson, vem do fato de que a probabilidade expressa na Equação 95 não

depende de nenhum t ou s específicos, mas sim da distância entre eles (sendo, portanto, irrelevantes seus valores absolutos t e s) e de r , um valor que, neste caso, também independe de t ou s , ou seja, é constante ao longo do contínuo (KULKARNI, 2017). De fato, como já mencionado na seção referente à distribuição Exponencial, r é também justamente o parâmetro da distribuição Exponencial que modela as chegadas entre eventos de um processo de Poisson; esta quantidade ser “estacionária” quer dizer que ela permanece inalterada ao longo do contínuo. (BEICHELT et al., 2002)

Um processo de Poisson é baseado nas seguintes hipóteses (ANG; TANG, 2007):

- a) Um evento pode acontecer aleatoriamente em qualquer ponto do contínuo;
- b) A ocorrência (ou as ocorrências) de um evento num dado intervalo do contínuo são estatisticamente independentes em intervalos que não se sobrepõem;
- c) A probabilidade de ocorrência de um evento num pequeno intervalo Δt é proporcional a este Δt , podendo ser computado por $r \cdot \Delta t$, em que r é a taxa de ocorrência. Ou seja, a grandeza r é a constante de proporcionalidade da probabilidade de que um evento ocorra num intervalo arbitrariamente pequeno. (TAYLOR; KARLIN, 1994).

$$P[N(t+\Delta t) - N(t) = 1] = r \Delta t + o(\Delta t). \quad (96)$$

- d) A probabilidade de duas ou mais ocorrências do evento em Δt é desprezável:

$$P[N(t+\Delta t) - N(t) \geq 2] = o(\Delta t). \quad (97)$$

A Equação 96 pode ser obtida a partir da expansão da Equação 95 em série de Taylor, como em Taylor & Karlin (1994), ou de maneira mais simplificada, conforme presente em Beichelt (2002).

Em ambas as Equações, 96 e 97, o termo $o(\Delta t)$ se refere ao Símbolo de Ordem de Landau (BEICHELT et al., 2002), uma função geral, não-especificada, cujo comportamento é o de ir para zero quanto menor for seu argumento:

$$\lim_{\Delta t \rightarrow \infty} \frac{o(\Delta t)}{\Delta t} = 0. \quad (98)$$

Um processo de Poisson é dito não-homogêneo se o parâmetro r não é constante ao longo do contínuo. Desta forma, sendo $r(t)$ a função que descreve a evolução da intensidade do parâmetro r , a taxa de chegadas, então, tem-se que:

$$\Lambda(\tau) = \int_0^{\tau} r(t) dt, \quad (99)$$

é a quantidade de chegadas num certo intervalo τ . Nota-se que esta grandeza é análoga àquela da Equação 37, agora com o diferencial de que a taxa com a qual essas chegadas acontecem não é constante. De fato, nota-se que, caso $r(t) = r = \text{constante}$, a Equação 99 se torna:

$$\Lambda(\tau) = \int_0^{\tau} r dt = r \cdot \tau = \lambda, \quad (100)$$

e voltamos ao caso homogêneo.

Do Cálculo, a área sob a curva $r(t)$ de t_1 a t_2 é igual à integral da curva até t_2 subtraída da área até t_1 (STEWART; CLEGG; WATSON, 2021). Assim, a probabilidade da Equação 94, é, para o caso não-homogêneo (BEICHELT et al., 2002):

$$P(N(t) - N(s) = k) = \frac{[\Lambda([s, t])]^k}{k!} \cdot e^{-[\Lambda([s, t])]} = \frac{(\Lambda(t) - \Lambda(s))^k}{k!} \cdot e^{-(\Lambda(t) - \Lambda(s))}. \quad (101)$$

O tempo T_1 até a primeira chegada é distribuído conforme:

$$F_{T_1}(t) = 1 - \exp\left(-\int_0^t r(t) dt\right), \text{ com:} \quad (102)$$

$$f_{T_1}(t) = r(t) \exp\left(-\int_0^t r(t) dt\right). \quad (103)$$

que são bastante similares ao caso homogêneo das Equações 45 e 46.

Finalmente, o tempo T_n até a n chegada é distribuído conforme (BEICHELT et al., 2002):

$$F_{T_n}(t) = \sum_{i=n}^{\infty} \frac{[\Lambda(t)]^i}{i!} \cdot e^{-\Lambda(t)}, \text{ com:} \quad (104)$$

$$f_{T_n}(t) = \frac{[\Lambda(t)]^{(n-1)}}{(n-1)!} \cdot r(t) \cdot e^{-\Lambda(t)}, \quad (105)$$

expressões bastante similares àquelas das Equações 52 e 53.

2.3 INCERTEZAS NA ENGENHARIA E O CONCEITO DE CONFIABILIDADE ESTRUTURAL

O comportamento de uma estrutura de engenharia quando submetida a um certo carregamento depende do tipo e da magnitude da carga aplicada, bem como da resistência e rigidez estruturais, das propriedades geométricas de seus componentes, das características dos materiais destes etc.. A consideração da resposta da estrutura ao carregamento aplicado como satisfatória depende do cumprimento de certos requisitos, como o de que a segurança desta estrutura contra colapso, o limite aceitável de danos durante uso, o controle de deformações etc. encontram-se todos dentro de níveis aceitáveis. Esses requisitos são conhecidos como estados limites, e uma violação do estado-limite é uma situação a ser evitada. Uma vez que há incertezas não somente relacionadas aos carregamentos, como também a basicamente todas as outras variáveis levadas em conta durante a fase de projeto, faz-se necessária uma abordagem estatística para a adequada avaliação do cumprimento, por parte da estrutura, a tais estados limites. (ANG; TANG, 2007)

A confiabilidade estrutural é, portanto, o campo da engenharia estrutural que se concentra em garantir a segurança e a durabilidade das estruturas considerando as incertezas das variáveis envolvidas no problema em questão. Seu objetivo principal é a avaliação da probabilidade de que uma estrutura desempenhe sua função pretendida sem falhas (aqui entendidas como violações do estado-limite), especialmente durante o período de tempo entendido como seu ciclo de vida (KAPUR, 2014; MELCHERS; BECK, 2018). Esta ideia pode ser matematicamente expressa como:

$$P_f(t) = 1 - C(t), \quad (106)$$

onde $C(t)$ denota a confiabilidade de uma determinada estrutura num determinado instante do tempo e $P_f(t)$ a probabilidade de falha desta estrutura neste mesmo instante.

2.3.1 O problema básico de confiabilidade e os conceitos de função e equação de estado-limite

O problema básico de confiabilidade é tido como aquele em que se busca determinar a probabilidade de que a solicitação sobre um sistema seja maior que a

capacidade deste sistema em resistir a esta solicitação – i.e., a determinação da probabilidade de falha P_f :

$$P_f = P(R \leq S) = P(R - S \leq 0), \quad (107)$$

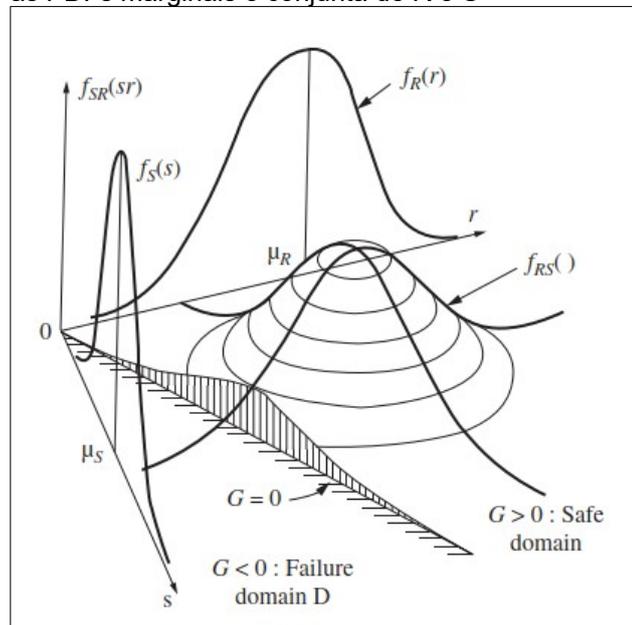
em que a equação $g(R, S) = R - S$, que determina a falha quando $R \leq S$, é chamada de função de estado-limite. Ao caso particular em que é igual a zero, dá-se o nome de equação de estado-limite (BECK, 2019)

Assim:

$$P_f = \int_{\Omega_f} f_{RS}(r, s) dr ds, \quad (108)$$

em que Ω_f é o chamado domínio de falha: o conjunto de todos os pontos amostrais das VAs do problema para os quais a função de estado-limite assume valores iguais ou inferiores a zero; $f_{RS}(r, s)$ é a PDF conjunta destas variáveis. Uma ilustração do domínio de falha e de segurança pode ser encontrada na Figura 8 e Figura 9.

Figura 8 – Domínios de falha e de segurança dadas as PDFs marginais e conjunta de R e S



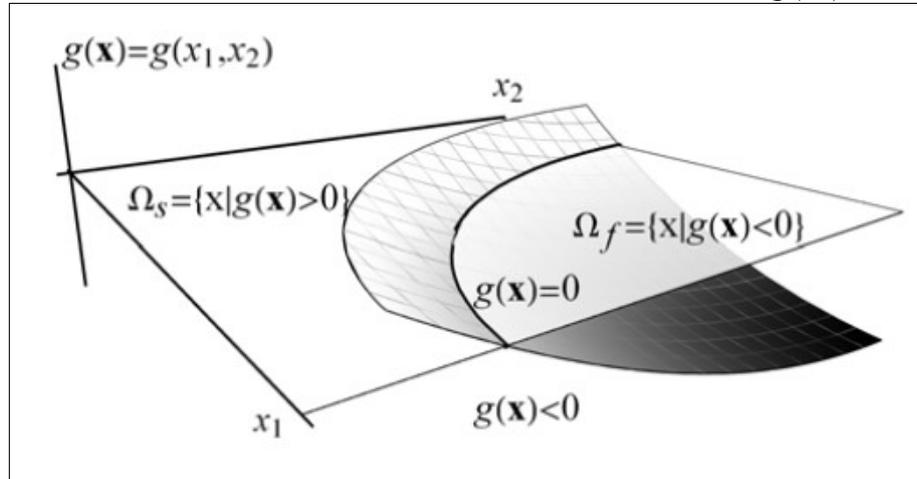
Fonte: Melchers, R.E. e Beck, A.T. (2018)

Mais genericamente, denota-se esta função como $g(\mathbf{X})$, em que $\mathbf{X} = \{X_1, X_2, \dots, X_n\}$ é o vetor de variáveis aleatórias do problema. Portanto, tem-se:

$$P_f = \int_{\Omega_f} f_{\mathbf{X}}(\mathbf{x}) d\mathbf{x}, \quad (109)$$

$$\begin{aligned}\Omega_f &= \{x|g(x)\leq 0\} \\ \Omega_s &= \{x|g(x)>0\}\end{aligned}\quad (110)$$

Figura 9 – Equação de estado-limite $g(\mathbf{X})=0$ e os domínios de falha e sobrevivência com base nos valores da função de estado-limite $g(\mathbf{X})$



Fonte: Beck, A.T. (2019)

A integral na Equação 109 é, em geral, de solução analítica explícita muito difícil – ou mesmo impossível – para problemas de interesse prático real, tanto por causa da complexidade da superfície de f_x – especialmente quando f_x não é multinormal, o que geralmente é o caso –, quanto por causa da quantidade de variáveis envolvidas nos problemas, bem como devido a questões referentes à estimação das características estatísticas conjuntas destas variáveis (ANG; TANG, 2007; PAIK; MELCHERS, 2009; BECK, 2019). Desta forma, é usual o emprego de métodos numéricos para aproximação dos valores de probabilidade de falha da Equação 109.

2.3.2 Métodos usuais para cálculo da probabilidade de falha

Dadas as dificuldades inatas ao cálculo exato das probabilidades de falha expressas pela Equação 109, diversos métodos de estimação de seu valor foram desenvolvidos ao longo do tempo, destacando-se entre estes os métodos de simulação e os de transformação (ou analíticos). (LOPEZ; BECK, 2012)

Os métodos de simulação, dos quais o mais empregado é o Método de Monte Carlo, se valem de amostragem de valores das VAs e posterior cálculo em força-bruta dos valores das funções de estado-limite para cada uma das realizações, tendo como principal desvantagem a alta quantidade de avaliações da função

quando as probabilidades de falha são muito pequenas; a principal vantagem, por outro lado, é a simplicidade de implementação e a generalidade do método. De fato, o Método de Monte Carlo não somente é a referência contra a qual outros métodos são comparados, como recurso quando todos os outros métodos falham. (BECK, 1999; LOPEZ; BECK, 2012)

Os métodos de transformação (*First Order Second Moment* – FOSM, *First Order Reliability Method* – FORM, *Second Order Reliability Method* – SORM, etc.) (BECK, 2019), por sua vez, são assim chamados por não solucionarem a integral na Equação 109 no espaço original das variáveis X , mas sim num mapeamento deste a um espaço de variáveis normais-padrão independentes Y . Neste espaço são obtidos valores importantes para outros métodos – como *Importance Sampling* (BECK, 2019) –, a exemplo do *Most Probable Point* (MPP), o ponto no domínio de falha mais próximo da origem (tomada como a média das Vas X , que se torna a média, igual a zero, das variáveis Y) e o índice de confiabilidade de Hasofer-Lind (1974), β , a distância vetorial da origem ao MPP no espaço de Y . (LOPEZ; BECK, 2012; MELCHERS; BECK, 2018)

2.3.2.1 Método de Monte Carlo

O método de Monte Carlo simples (*Crude Monte Carlo*) (MELCHERS; BECK, 2018; BECK, 2019) baseia-se no conceito de valor esperado de uma função qualquer de uma (ou mais de uma, no caso) variável aleatória, podendo ser entendido como a integração numérica de uma função indicadora $I[X]$ definida como:

$$I[x] = \begin{cases} 1 & \text{se } x \in \Omega_f, \\ 0 & \text{se } x \in \Omega_s \end{cases}, \quad (111)$$

em que Ω_f e Ω_s são os mesmos que na Equação 110.

Do conceito de esperança de uma função de uma variável aleatória, tem-se que a probabilidade de falha pode ser estimada a partir da integral da função indicadora sobre todo o domínio Ω de X (DEVORE, 2016; JOHNSON; MILLER; FREUND, 2017; BECK, 2019):

$$p_f = \int_{\Omega} I[x] f_X(x) dx = E[I[x]]. \quad (112)$$

Para toda avaliação da função indicadora $I[X]$ haverá uma avaliação da função de estado-limite $g(X)$ correspondente. (BECK, 2019)

Como mencionado anteriormente, o método de Monte Carlo se vale de amostras no espaço das variáveis de projeto com as quais se avaliará a função de estado-limite e, então, aproximar a integral da Equação 112 acima. Formalmente, escreve-se:

$$E[I[\mathbf{x}]] \approx \tilde{p}_f = \frac{1}{n_s} \sum_{k=1}^{n_s} I[\mathbf{x}_k] = \frac{n_f}{n_s}, \quad (113)$$

com n_f sendo o número de pontos amostrais no domínio de falha Ω_f . (BECK, 2019)

O método de Monte Carlo fornece uma estimativa não-tendenciosa da probabilidade de falha, uma vez que a aproximação \tilde{p}_f tende ao valor exato p_f conforme n_s tende ao infinito. (RUBINSTEIN; KROESE, 2017; MELCHERS; BECK, 2018; BECK, 2019)

Uma expressão para a variância da aproximação \tilde{p}_f como função do tamanho da amostra é dada por (BECK, 2019):

$$\text{Var}[\tilde{p}_f] = \frac{\tilde{p}_f(1-\tilde{p}_f)}{n_s}. \quad (114)$$

Com o coeficiente de variação expresso como:

$$\text{CoV}[\tilde{p}_f] = \frac{\sigma(\tilde{p}_f)}{E[\tilde{p}_f]} = \frac{\sqrt{\text{Var}[\tilde{p}_f]}}{E[\tilde{p}_f]} = \sqrt{\frac{\tilde{p}_f(1-\tilde{p}_f)}{n_s}} \cdot \frac{1}{\tilde{p}_f} = \sqrt{\frac{1-\tilde{p}_f}{n_s \tilde{p}_f}} = \delta_{\tilde{p}_f}, \quad (115)$$

de onde é possível obter que o tamanho n_s da amostra para obter uma estimativa da probabilidade de falha, com um determinado coeficiente de variação $\delta_{\tilde{p}_f}$, tende ao infinito quando \tilde{p}_f tende a zero (Equação 116). (ANG; TANG, 2007; BECK, 2019)

$$\sqrt{\frac{1-\tilde{p}_f}{n_s \tilde{p}_f}} = \delta_{\tilde{p}_f} \rightarrow \delta_{\tilde{p}_f}^2 = \frac{1-\tilde{p}_f}{n_s \tilde{p}_f} \rightarrow n_s = \frac{1-\tilde{p}_f}{\tilde{p}_f \delta_{\tilde{p}_f}^2} \quad \therefore \quad \lim_{\tilde{p}_f \rightarrow 0} n_s = \lim_{\tilde{p}_f \rightarrow 0} \left(\frac{1-\tilde{p}_f}{\tilde{p}_f \delta_{\tilde{p}_f}^2} \right) = \infty. \quad (116)$$

2.3.2.2 Principais métodos de transformação e o mapeamento das variáveis aleatórias para o espaço normal-padrão

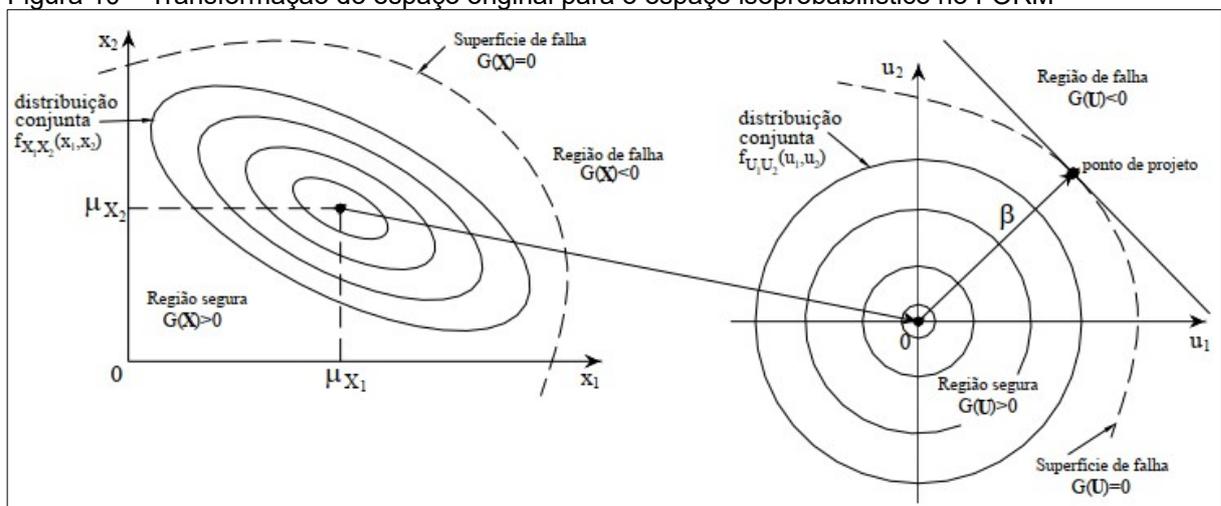
O método FOSM é o mais simples dentre os métodos de transformação citados e apresenta uma abordagem de fácil implementação computacional que foi não somente amplamente usada na análise de confiabilidade, como serviu de base

para o desenvolvimento de métodos mais robustos posteriormente. Este método aproxima a função de estado-limite por uma função linear e usa o primeiro e segundo momentos (média e variância) das variáveis aleatórias que representam os parâmetros incertos (por exemplo, cargas, resistências do material, condições ambientais) do problema. A consideração de média e variância apenas equivale a considerar as variáveis aleatórias todas como tendo distribuição normal, o que é severamente limitante quando se trata de problemas reais práticos. (BECK, 2019)

Os métodos de primeira e segunda ordem FORM e SORM, no entanto, se valem de todas as informações estatísticas das variáveis aleatórias para o cômputo da probabilidade de falha, levando em consideração as distribuições marginais e coeficientes de correlação entre cada um dos pares de variáveis (MELCHERS; BECK, 2018). A aproximação do FORM, assim como o FOSM, advém da linearização da função de estado limite no ponto de projeto; no SORM, a função de estado-limite é aproximada por uma parábola. (MELCHERS; BECK, 2018; BECK, 2019).

O espaço normal-padrão das variáveis Y é chamado de isoprobabilístico pois as curvas de nível (neste contexto também referidas como curvas de equiprobabilidade) são círculos concêntricos centrados na origem. Esta formulação se aproveita da simetria radial, em torno da média, da distribuição conjunta de variáveis normais-padrão independentes, conforme ilustra a Figura 10 para o caso em que há apenas duas variáveis aleatórias. (BECK, 2019)

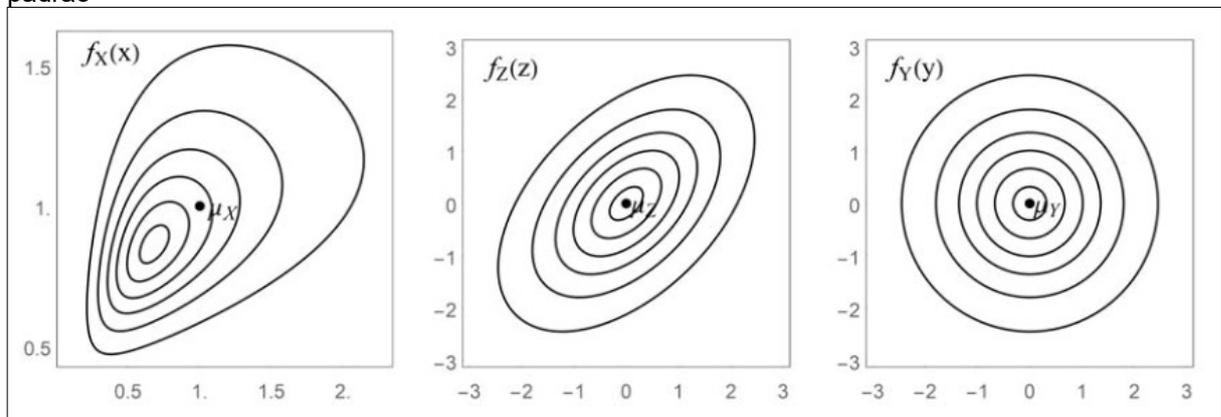
Figura 10 – Transformação do espaço original para o espaço isoprobabilístico no FORM



Fonte: Paliga, C.M. (2008)

Um conceito central nos métodos FORM e SORM é o mapeamento das variáveis aleatórias X do problema de seu espaço original, aquele em que apresentam distribuições quaisquer e onde não são necessariamente independentes (com particular interesse no caso em que a correlação linear entre as variáveis não é necessariamente nula), para o espaço de variáveis com distribuições normais-padrão independentes Y . Dentre os meios para efetuar esta transformação de espaços, destaca-se a Transformação Composta com modelo de Nataf, graficamente representada na Figura 11. (MELCHERS; BECK, 2018; BECK, 2019)

Figura 11 – Etapas da transformação composta: VAs correlacionadas e em distribuições quaisquer → VAs correlacionadas mas com distribuição normal-padrão → VAs independentes no espaço normal-padrão



Fonte: Beck, A.T. (2019)

Esta transformação é tida como composta pois envolve três passos principais (BECK, 2019):

- a) Obtenção de distribuições marginais normais equivalentes, mas possivelmente correlacionadas, a partir das distribuições marginais originais, transformação $X \rightarrow Z$;
- b) Estabelecimento dos coeficientes de correlação equivalentes para as distribuições normais obtidas (modelo de Nataf);
- c) Eliminação da correlação das distribuições normais equivalentes, transformação $Z \rightarrow Y$.

2.3.2.2.1 Obtenção de distribuições normais equivalentes

A obtenção de distribuições normais equivalentes é de solução trivial quando todas as variáveis aleatórias do problema têm distribuições normais e são

independentes entre si, caso no qual basta a aplicação da transformação de Hasofer-Lind (Equação 32) para cada variável.

Em problemas de ordem prática, no entanto, tal cenário não é realista: as variáveis dos problemas via de regra não têm distribuição normal e as correlações entre si não são nulas, o que torna necessário passos adicionais além da obtenção das distribuições normais equivalentes. Para tanto, procede-se com a ideia de aproximar a distribuição original por uma distribuição normal equivalente que preserve seu conteúdo de probabilidades num certo ponto x_i^* . Este é conhecido como o Princípio de Aproximação Normal ou Princípio da Aproximação de Cauda (DITLEVSEN, 1981), representado formalmente por (MELCHERS; BECK, 2018; BECK, 2019):

$$F_{X_i}(x_i^*) = F_{X_i}^{\text{neq}}(x_i^*) = \Phi\left(\frac{x_i^* - \mu_{X_i}^{\text{neq}}}{\sigma_{X_i}^{\text{neq}}}\right) = \Phi(z_i^*), \quad (117)$$

em que $\mu_{X_i}^{\text{neq}}$ e $\sigma_{X_i}^{\text{neq}}$ são a média e desvio-padrão que definem a distribuição normal equivalente.

A rigor, o critério para a determinação destes parâmetros fica em aberto, mas é usual é obtê-los a partir da consideração de que também as funções de densidade sejam iguais no mesmo ponto x_i^* .

$$f_{X_i}(x_i^*) = f_{X_i}^{\text{neq}}(x_i^*) = \frac{1}{\sigma_{X_i}^{\text{neq}} \cdot \sqrt{2\pi}} \cdot \exp\left(-\frac{1}{2} \cdot \left(\frac{x_i^* - \mu_{X_i}^{\text{neq}}}{\sigma_{X_i}^{\text{neq}}}\right)^2\right) = \frac{\phi(z_i^*)}{\sigma_{X_i}^{\text{neq}}}. \quad (118)$$

Desta forma, é possível chegar a expressões para os parâmetros (desvio-padrão e média) da distribuição normal equivalente (BECK, 2019):

$$\sigma_{X_i}^{\text{neq}} = \frac{\phi(z_i^*)}{f_{X_i}(x_i^*)}, \quad (119)$$

$$\mu_{X_i}^{\text{neq}} = x_i^* - z_i^* \cdot \sigma_{X_i}^{\text{neq}}. \quad (120)$$

Se existe dependência linear \mathbf{R}_X entre as variáveis \mathbf{X} no espaço original, então esta informação também deve ser convertida para uma dependência linear \mathbf{R}_Z^2

2 \mathbf{R}_X e \mathbf{R}_Z sendo as matrizes de correlação para variáveis \mathbf{X} e \mathbf{Z} , e cujos elementos são representados por $\rho_{X_{ij}}$ e $\rho_{Z_{ij}}$

entre as variáveis Z no espaço normal-padrão, que fica a cargo da já mencionada Transformação de Nataf. Uma expressão para a relação entre os coeficientes de correlação das variáveis X e Z é dada (BECK, 2019):

$$\rho_{X_{ij}} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} z_i z_j \phi(z_i, z_j, \rho_{Z_{ij}}) dz_i dz_j, \text{ com } i=1,2,\dots,n \text{ e } j=1,2,\dots,n. \quad (121)$$

Na Equação 121 os subscritos i e j se referem aos índices de cada uma das n variáveis aleatórias do problema. Um importante aspecto da conversão dos valores dos coeficientes de correlação entre as variáveis X e Z é que esta é feita de forma iterativa, arbitrando valores candidatos de ρ_Z e avaliando ρ_X até que se atinja ρ_X desejado, apesar de existirem expressões analíticas para certas combinações de distribuições. (MELCHERS; BECK, 2018; BECK, 2019)

2.3.2.2.2 Eliminação da correlação

Para que haja simetria radial num espaço normal-padrão é necessário que as variáveis neste espaço sejam independentes, ou seja, é necessário transformar as variáveis do espaço das variáveis normais-padrão Z para um espaço de variáveis normais-padrão independentes Y . A eliminação da dependência linear expressa pela matriz de correlação R_Z (a qual está atrelada uma matriz de covariâncias C_Z (Equação 28)), se dá pela busca de uma transformação linear (Equação 122) das variáveis tal que as variáveis em Y resultem independentes (BECK, 2019):

$$Y = B^T Z. \quad (122)$$

Ou seja, pretende-se determinar a matriz B tal que a matriz de covariâncias de Y , C_Y , seja igual à matriz identidade I , conforme a Equação 123 (BECK, 2019):

$$\begin{aligned} C_Y &= \text{Cov}[U, U^T] \\ " &= \text{Cov}[B^T Z, Z^T B] \\ " &= B^T \text{Cov}[Z, Z^T] B \\ " &= B^T C_Z B = I \end{aligned} \quad (123)$$

Da Equação 123 é possível demonstrar que C_Z pode tomar a forma de:

$$\begin{aligned} C_Z &= (B^T)^{-1} B^{-1} \\ &= L L^T \end{aligned} \quad , \text{ ao se fazer } L = (B^T)^{-1} \text{ e } L^T = B^{-1}, \quad (124)$$

ou seja: é possível obter B a partir da decomposição de Cholesky da matriz de covariâncias C_Z . Neste caso, L é uma matriz triangular inferior com elementos positivos e reais na diagonal; exige-se que C_Z seja uma matriz positiva-definida. (WATKINS, 2010)

Desta forma, das Equações 122 e 124 obtém-se (BECK, 2019):

$$Y = L^{-1}Z, \quad (125)$$

$$Z = LY, \quad (126)$$

tal que, da Equação 125 é possível obter variáveis normais-padrão independentes Y . A Equação 126, por sua vez, fornece uma relação para a transformação em sentido contrário: do espaço de variáveis Y , independentes e com distribuição normal-padrão, de volta ao espaço de variáveis Z , dependentes e com distribuição normal-padrão.

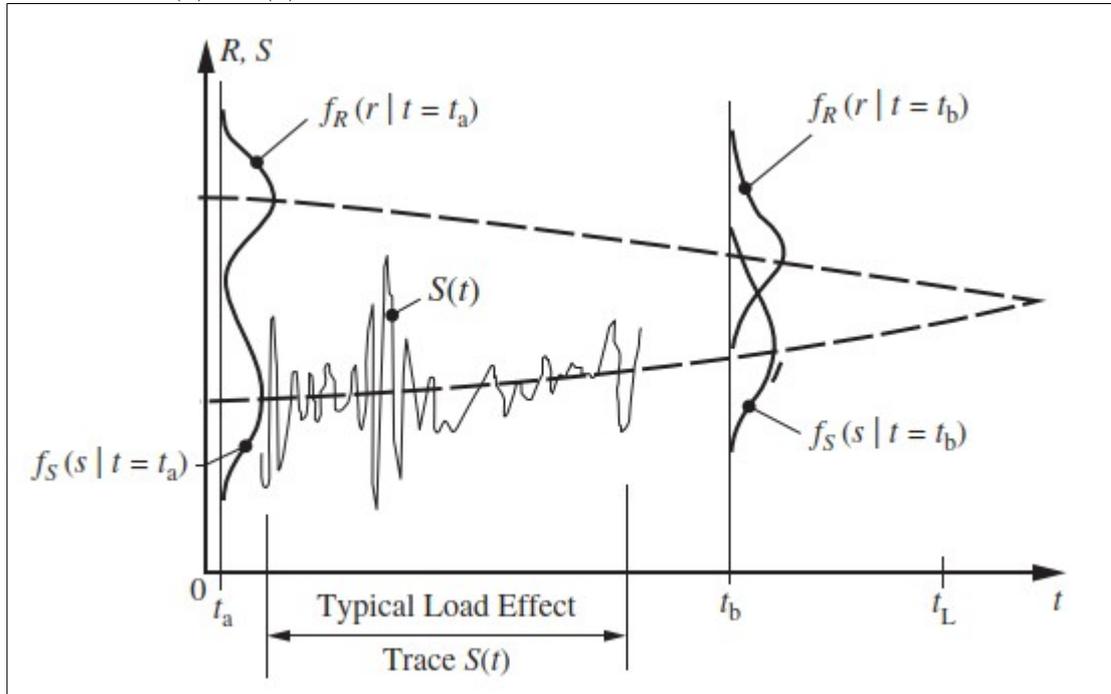
2.3.3 Confiabilidade no tempo e modelo de falha à primeira sobrecarga

Nas situações em que as ações e/ou resistências de uma determinada estrutura ou sistema estrutural não são constantes no tempo, como o caso ilustrado na Figura 12, a probabilidade de que haja falha também não o é. Portanto, tem-se que o problema básico de confiabilidade toma a forma da Equação 127 (MELCHERS; BECK, 2018):

$$p_f(t) = P(R(t) \leq S(t)), \quad (127)$$

de maneira que, se as distribuições $f_R(t)$ e $f_S(t)$ de $R(t)$ e $S(t)$ são conhecidas, a probabilidade $p_f(t)$, chamada probabilidade de falha instantânea, pode ser obtida a partir da Equação 108. (MELCHERS; BECK, 2018)

Figura 12 – Ilustração de um problema de confiabilidade dependente do tempo. A ação e resistência, $S(t)$ e $R(t)$, respectivamente, variam no tempo



Fonte: Melchers, R.E.; Beck, A.T. (2018)

De particular interesse, no entanto, é a probabilidade de que haja falha dentro de um período arbitrário t_D , ou seja, a probabilidade de que o mínimo valor da função de estado limite a qualquer tempo t , com $t \in [0, t_D]$, seja menor que zero. Formalmente (BECK, 2019):

$$p_f(t_D) = P\left(\min_{0 \leq t \leq t_D} g(\mathbf{R}, S, t) \leq 0\right), \quad (128)$$

em que os domínios de falha e sobrevivência podem ser expressos como:

$$\begin{aligned} \Omega_f(t) &= \{\mathbf{r}, s | g(\mathbf{R}, S, t) \leq 0\} \\ \Omega_s(t) &= \{\mathbf{r}, s | g(\mathbf{R}, S, t) > 0\} \end{aligned} \quad (129)$$

Assumindo que a estrutura de interesse, como é o caso de dutos, objeto deste trabalho, falhe à primeira ocorrência de sobrecarga, convém entender a chegada destas sobrecargas como um processo de Poisson, o que faz o tempo entre sobrecargas ser distribuído exponencialmente, conforme discutido na seção 2.1.4. Assim sendo, seja $N^+(r, t_D)$ a notação para a quantidade de sobrecargas no intervalo $[0, t_D]$, e seja $\eta(r, t)$ sua taxa não-constante de chegadas, a confiabilidade da estrutura é dada pela probabilidade de que a quantidade de sobrecargas no

intervalo seja zero. Usando a Equação 38 e 99, escreve-se a confiabilidade no formato do seguinte processo de Poisson não-homogêneo (BECK, 2019):

$$\begin{aligned}
 \mathcal{R}(r, t_D) &= P(N^+(r, t_D)=0) \\
 " " &= \frac{\Lambda(t_D)^0}{0!} \cdot \exp(-\Lambda(t_D)) \\
 " " &= \frac{\left[\int_0^{t_D} \eta(r, t) dt \right]^0}{0!} \cdot \exp\left(-\int_0^{t_D} \eta(r, t) dt\right) \\
 " " &= \exp\left(-\int_0^{t_D} \eta(r, t) dt\right)
 \end{aligned} \tag{130}$$

Em que a probabilidade de falha durante o intervalo $[0, t_D]$ é o complemento da confiabilidade, conforme Equação 106:

$$\begin{aligned}
 p_f(r, t_D) &= 1 - \mathcal{R}(r, t_D) \\
 " " &= 1 - \exp\left(-\int_0^{t_D} \eta(r, t) dt\right),
 \end{aligned} \tag{131}$$

em que está assumido que a probabilidade falha no instante inicial é irrelevante (BECK, 2019).

A quantidade $\eta(r, t)$, a taxa não-constante de chegadas de sobrecarga, é condicional a um determinado valor r da barreira (resistência) e é, em geral, desconhecida. Uma aproximação usual é aquela que a substitui pela taxa de passagens pela barreira de baixo pra cima, $v^+(r, t)$, como na Equação 132, (BECK, 2019)

$$\eta(r, t) \approx v^+(r, t), \tag{132}$$

para a qual há expressões explícitas para certos processos estocásticos. Esta aproximação é apropriada para valores altos da barreira r , e tende à exatidão quando $r \rightarrow \infty$.

A Equação 131 fica, então:

$$p_f(r, t_D) = 1 - \exp\left(-\int_0^{t_D} v^+(r, t) dt\right). \tag{133}$$

Quando a resistência é um vetor de variáveis aleatórias \mathbf{R} , então para cada realização de seus valores, $\mathbf{R}=\mathbf{r}$, haverá uma barreira $r(t)$ distinta, seja ela constante ou variando no tempo. Desta forma, para cada realização há uma probabilidade condicional (a um determinado valor de r) para o qual a expressão da Equação 133 é calculada. Formalmente, isto é expresso tal qual a Equação 134 (BECK, 2019)

$$p_f(t_D|\mathbf{r})=1-\exp\left(-\int_0^{t_D} v^+(\mathbf{r},t)dt\right). \quad (134)$$

Do teorema da probabilidade total, tem-se que a probabilidade de falha incondicional é dada pela integral sobre todos os valores de \mathbf{R} , tal qual expresso na Equação 135.

$$\begin{aligned} p_f(t_D) &= E_{\mathbf{R}}[p_f(t_D|\mathbf{r})] \\ " " &= \int_{\mathbf{R}} p_f(t_D|\mathbf{r})f_{\mathbf{R}}(\mathbf{r})d\mathbf{r}' \end{aligned} \quad (135)$$

que é passível de ser computado pelo Método de Monte Carlo (BECK, 2019). Assim, a probabilidade de falha condicional ao conjunto das realizações \mathbf{r} é dada pela soma da contribuição de cada realização individualmente. Nota-se, ainda, que esta solução é específica para um determinado tempo t para o qual $p_f(t_D|\mathbf{r})$, dentro da integral na Equação 135, é calculado. (BECK, 2019)

2.3.4 Confiabilidade em dutos corroídos

Dutos para transporte de petróleo e gás, sejam eles *on-shore* ou *off-shore*, estão expostos às intempéries dos meios em que se encontram instalados – que são, via de regra, pouco ou muito pouco passíveis de controle pelos operadores –, bem como aos efeitos mecânicos e químicos dos materiais que transportam; sofrem, por exemplo, com ações corrosivas tanto externas, quanto internas (MUHLBAUER, 2004; NORSWORTHY, 2014; TIMASHEV; BUSHINSKAYA, 2016; MAHMOODIAN, 2018). Como nenhuma destas ações é estática no tempo, então a confiabilidade de um duto, assim como acontece para outros tipos de estruturas, muda ao longo do período em que permanece em operação.

As formas pelas quais um duto pode vir a falhar são várias; vão desde aquelas relacionadas a alterações na geometria do duto – flexão longitudinal,

ovalização, flambagem etc. – para além dos limites aceitáveis, até aquelas referentes a, de fato, uma modificação na forma com que o duto opera, como trincas ou perfurações da parede que levem a vazamentos ou, no limite, grande vazamentos/explosões (*burst*) (MAHMOODIAN, 2018; AMAYA-GÓMEZ et al., 2024). No contexto do cálculo da probabilidade de falha no tempo, no entanto, as falhas que envolvem corrosão – um fenômeno que varia significativamente no tempo – se destacam.

A corrosão atua reduzindo a espessura da parede do duto e, assim, reduzindo a resistência mecânica deste aos esforços aos quais é submetido durante sua operação (HOPKINS, 2014). Referentes à ação de redução da seção advinda da corrosão, há dois principais critérios pelos quais se julga um determinado defeito³: profundidade do defeito e pressão interna a qual o duto ainda resiste quando da presença daquele defeito (MUHLBAUER, 2004). Estes dois critérios dão origem a duas das funções de estado-limite mais comuns na literatura a respeito da confiabilidade em dutos com defeitos de corrosão (MUHLBAUER, 2004; TIMASHEV; BUSHINSKAYA, 2016; SU et al., 2023; AMAYA-GÓMEZ et al., 2024):

$$g_1(t) = \gamma \cdot wt - d_{\max}(t), \quad (136)$$

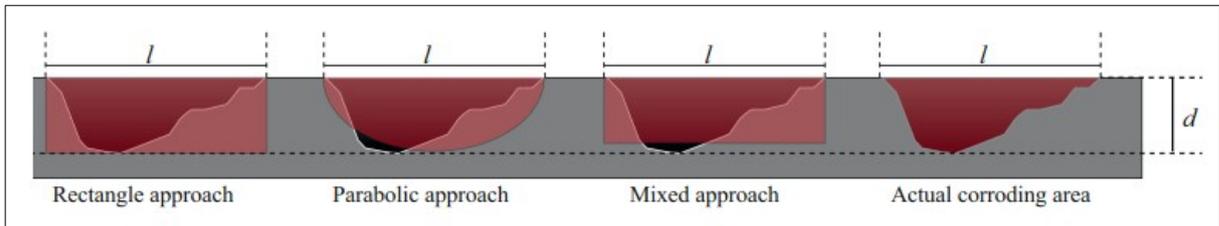
$$g_2(t) = r_b(t) - P_{\text{int}}(t). \quad (137)$$

A expressão em 136 é a função de estado limite que determina o modo de falha por pequeno vazamento (*small leak*) que acusa falha sempre que $d_{\max}(t)$, a profundidade máxima do perfil do defeito no tempo t , for maior que uma determinada fração γ da espessura wt do duto. Os valores para γ podem, naturalmente, variar a depender do problema em questão, mas é o usual da literatura a consideração deste como igual a 0.8 sob a justificativa de ser prática corrente da indústria (CALEYO; GONZÁLEZ; HALLEN, 2002; CALEYO et al., 2009; ZHOU, 2010; ASME, 2012; ZHOU; HONG; ZHANG, 2012; ZHANG; ZHOU, 2013; GONG; ZHOU, 2017, 2018; MISHRA; KESHAVARZZADEH; NOSHADRAVAN, 2019; ZHANG et al., 2019; XIAO et al., 2024a, 2024b). Nota-se, entretanto, que autores como Amaya-Gómez et al. (2024) acompanham normas como a DNV-RP-FP101 (2017) na consideração como igual a 0.85.

3 Aqui entendido como aquele tipo de anomalia indesejada, como a perda de material na seção, que leve, ou possa levar, a uma redução sensível da resistência mecânica do duto ou a um vazamento de qualquer tamanho. (MUHLBAUER, 2004)

A Equação 137, por sua vez, diz respeito à falha causada pelo colapso plástico (*burst*) (TEE; PESINIS, 2017; MISHRA; KESHAVARZZADEH; NOSHADRAVAN, 2019). Este é estabelecido quando a pressão interna de operação do duto, $P_{\text{int}}(t)$, ultrapassa, em um determinado instante, a pressão $r_b(t)$ à qual o mesmo resiste quando da presença de um defeito de determinadas características geométricas. A pressão resistente $r_b(t)$ é usualmente calculada a partir do uso de normas como a ASME B31G e ASME B31G modificada (ASME, 2012), DN-RP-FP101 (DNV GL AS, 2017), bem como de softwares comerciais como o RSTRENG (PRCI, 2020), que aproximam, cada uma a seu modo, o perfil real do defeito por uma simplificação pertinente. O cálculo de $r_b(t)$ por estas normas tem como parâmetros de entrada informações como o comprimento longitudinal e a profundidade máxima do defeito; a espessura da parede e o diâmetro do duto; as tensões últimas do aço etc.. Amaya-Gómez *et al.* (AMAYA-GÓMEZ *et al.*, 2019) fazem uma revisão das diversas formas com as quais é possível obter valores para $r_b(t)$.

Figura 13 - As diversas formas de simplificação da geometria real de um defeito para fins de cálculo da pressão interna resistente $r_b(t)$.



Fonte: Amaya-Gómez *et al.* (2019)

Um terceiro modo de falha, condicional à ocorrência do colapso plástico é a ruptura (*rupture*), indicado na Equação 138:

$$g_3(t) = r_{\text{RP}}(t) - P_{\text{int}}(t) \quad (138)$$

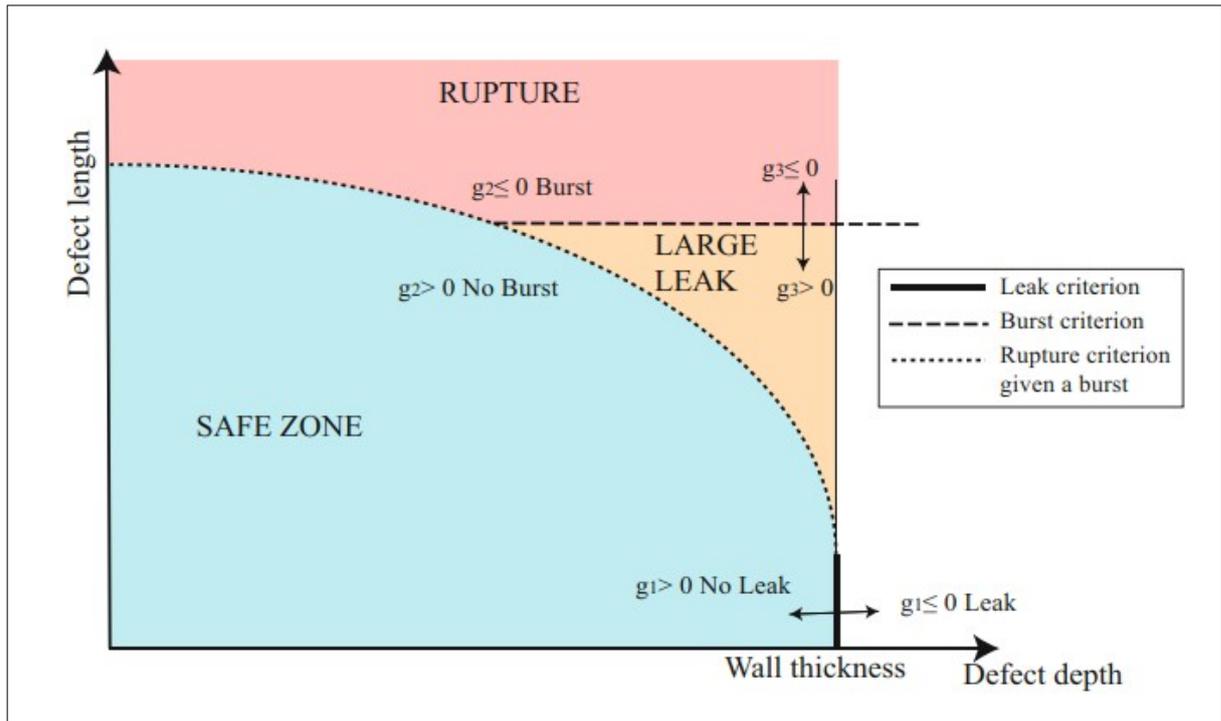
caso em que a falha se dá quando, após o colapso plástico, há a extensão axial instável da seção do defeito (ZHOU, 2010). Na Equação 138 $r_{\text{RP}}(t)$ é a pressão para que haja *rupture*.

Zhou (2010) também afirma que, quando da consideração simultânea destes três modos de falha, há *small leak* quando o evento $\{g_1 \leq 0 \cap g_2 > 0\}$ acontece; *burst* quando $\{g_1 > 0 \cap g_2 \leq 0\}$; *large leak* quando $\{g_1 > 0 \cap g_2 \leq 0 \cap g_3 > 0\}$ e, finalmente, *rupture*

quando $\{g_1 > 0 \cap g_2 \leq 0 \cap g_3 \leq 0\}$. Ou seja, há uma diferenciação do *burst* para caso aconteça (ou não aconteça) *rupture* em seguida.

Uma ilustração da ocorrência de cada um dos modos de falha como função da profundidade e comprimento do defeito se encontra na Figura 14.

Figura 14 – Modos de falha em função do comprimento e profundidade do defeito.



Fonte: Amaya-Gómez *et al.* (2024)

Na Equação 138 a pressão resistente r_{RB} também é função das características mecânicas do material que compõe o duto, bem como da geometria do duto e também do defeito.

Das Equações 136, 137 e 138 nota-se que em todas elas há um parâmetro temporal. Isto se dá devido a todas as funções de estado-limite apresentadas, $g_1(t)$, $g_2(t)$ e $g_3(t)$, deverem sua variação temporal às variações nas dimensões do defeito, seja ao todo (como é o caso da falha por *small leak*) ou em parte (uma vez que a pressão interna atuante pode variar, apesar de geralmente pouco, com o tamanho do defeito). Desta forma, entender como um defeito evolui no tempo – o que equivale a entender como o processo de corrosão afeta suas dimensões – é necessário para a adequada previsão de qual será o estado de um duto dado um intervalo após ter sido instalado.

Assim, destaca-se dois modos possíveis de representação de como as dimensões do defeito variam no tempo: linear e não-linear (seguindo uma lei de potência (*power law*)). O modelo linear, simplificação usual da literatura e ainda bastante empregado, se dá na forma:

$$d(t) = d_0 + Rd \cdot (t - t_0), \quad (139)$$

$$L(t) = L_0 + RL \cdot (t - t_0), \quad (140)$$

em que Rd e RL se referem às taxas lineares de crescimento da profundidade e do comprimento, respectivamente; d_0 e L_0 à profundidade e comprimento no instante de surgimento t_0 do defeito em questão. Ao modelo em que d_0 e Rd são simples variáveis aleatórias, Bazán e Beck (2013) chamam de modelo linear randômico de crescimento da corrosão (*linear random variable corrosion growth model*).

A corrosão, no entanto, é um fenômeno eletroquímico que se autoinibe, de forma tal que a taxa com que avança é maior no início do processo e vai decaindo com o tempo. De acordo com Mahmoodian (2018), apesar de o uso do modelo linear ainda ser tópico de debates na literatura, o modelo não-linear de lei de potência é largamente aceito como válido para representação do crescimento da profundidade de defeitos de corrosão. A formulação matemática é dada pela Equação 141 (BAZÁN; BECK, 2013):

$$d(t) = k(t - t_0)^\alpha, \quad (141)$$

onde k e α são constantes determinadas a partir de ajustes de curva a dados empíricos obtidos em campo e t_0 tem a mesma interpretação que nas Equações 139 e 140 (MAHMOODIAN, 2018). Velázquez *et al.* (2009) e Caleyó *et al.* (2009) obtiveram estimativas para estes valores – e para sua adequada como variáveis aleatórias – a partir de regressões estatísticas que tiveram como base de dados 259 amostras de solo e dutos enterrados coletados num período de três anos no sul do México. Os trabalhos mostraram que há grande variação nos valores de k e α .

Quanto ao tempo t_0 para surgimento de defeitos, uma abordagem bastante comum na literatura tem sido aquela em que há a consideração de que t_0 é determinístico, constante e igual a 2.88 anos (GOMES; BECK; HAUKAAS, 2013; VALOR *et al.*, 2013; GOMES; BECK, 2014; MISHRA; KESHAVARZZADEH; NOSHADRAVAN, 2019; SOUSA; AFONSO; WILLMERSDORF, 2021; WOLDESELLASSE; TESFAMARIAM, 2023a; XIAO *et al.*, 2024a, 2024b). Este valor

também tem como origem os trabalhos de Velázquez *et al.* (2009) e Caleyó *et al.* (2009) e são referentes a um solo genérico. Devido à falta de trabalhos similares em escopo e profundidade para outras partes do mundo, este tem sido o valor empregado em uma grande parte da literatura a respeito da confiabilidade de dutos corroídos no tempo quando não há maiores considerações a respeito do tempo de surgimento dos defeitos.

Um outro caminho, mais complexo, é a consideração de que os defeitos surgem aleatoriamente no tempo. No caso, um dos mais utilizados (ZHANG; ZHOU, 2014; PESINIS; TEE, 2017; TEE; PESINIS, 2017; SOUSA; AFONSO; WILLMERSDORF, 2021) é aquele primeiro empregado por Kuniewski *et al.* (2009) em que o número de defeitos num dado ponto t do tempo é distribuído conforme um processo de Poisson com taxa $\Lambda(t)$ variando no tempo; um processo de Poisson não-homogêneo, como descrito na seção 2.2.6 .

2.3.4.1 Probabilidade de falha para eventos dependentes e independentes no contexto de dutos corroídos

Para efeito de cálculo de probabilidade de falha, um duto é um tipo de estrutura adequada para consideração como um sistema em série: aquele tipo de sistema em que a falha de um componente implica a falha de todo o sistema, como numa corrente cujo elo se rompe (MUHLBAUER, 2004; GUO *et al.*, 2014). No caso deste trabalho, o sistema tem como seus componentes vários defeitos, de forma que se um defeito entra em estado de falha, então todo o sistema (segmento de duto) falha. A operação segura de um duto, portanto, necessita que todos os defeitos estejam em situação de não-falha. Essa abordagem é adotada por vários trabalhos que tratam da análise de confiabilidade em dutos. (ZHOU, 2010; SAHRAOUI; CHATEAUNEUF, 2016; TEE; PESINIS, 2017; GONG; ZHOU, 2017; LIU *et al.*, 2018; GONG; ZHOU, 2018; FU *et al.*, 2019; CHAKRABORTY; TESHAMARIAM, 2021)

A falha de um sistema em série é representada como a união de todos os n eventos falha:

$$F_{sys} = F_1 \cup F_2 \cup \dots \cup F_n = \cup_{i=1}^n F_i. \quad (142)$$

Se o sistema é dotado de apenas dois componentes – logo, dois possíveis eventos, F_1 e F_2 , que podem levar à falha –, então tem-se a expressão clássica para a probabilidade de união de dois eventos (MAGALHÃES, 2011):

$$P[F_1 \cup F_2] = P[F_1] + P[F_2] - P[F_1 \cap F_2]. \quad (143)$$

Para uma quantidade n qualquer de componentes, a expressão fica:

$$\begin{aligned} P[F_1 \cup F_2 \cup \dots \cup F_n] = & + P[F_1] \\ & + P[F_2] - P[F_2 \cap F_1] \\ & + P[F_3] - P[F_3 \cap F_1] - P[F_3 \cap F_2] + P[F_3 \cap F_2 \cap F_1] \\ & + \dots \end{aligned} \quad (144)$$

Generalizando-se, como (BECK, 2019):

$$P_{f_{sys}} = \sum_{i=1}^n P[F_i] - \sum_{i=2}^n \sum_{\substack{j=1 \\ (j < i)}}^{i-1} P[F_i \cap F_j] + \sum_{i=3}^n \sum_{\substack{j=2 \\ (j < i)}}^{i-1} \sum_{k=1}^{i-2} P[F_i \cap F_j \cap F_k] - \dots, \quad (145)$$

com somatórios alternando de sinal à medida que mais interseções entre eventos são consideradas.

A expressão na Equação 145 pode ser demonstrada igual a (DANTAS, 2013; MELCHERS; BECK, 2018; BECK, 2019):

$$P_{f_{sys}} = 1 - \prod_{i=1}^n (1 - P[F_i]). \quad (146)$$

caso os eventos sejam independentes.

Se não forem independentes, no entanto, as probabilidades de interseção são afetadas, de forma que cálculo da probabilidade de falha, *a partir de informações de probabilidades de falha individuais de cada componente*, de um sistema em série em que os componentes apresentam dependência entre si pode ser computada apenas de forma aproximada, como discutido em Melchers & Beck (2018) e Beck (2019). A simulação direta do sistema pelo Método de Monte Carlo é uma opção, apesar do alto custo computacional envolvido quando as probabilidades do problema são muito baixas, caso haja muitos componentes; ou ambos.

No contexto de dutos, e em particular em trabalhos que lidam com vários defeitos agindo em conjunto, a simples consideração de independência entre defeitos que compõem o sistema ainda é bastante comum na literatura, a exemplo de Tee & Pesinis (2017), Pesinis & Tee (2017) e Sousa (2022). Em ambos os trabalhos os autores trabalharam com a confiabilidade de um sistema com vários

defeitos (mais de uma dezena) agindo simultaneamente. Nesta situação, admiti-los todos independentes entre si é um pressuposto excessivamente conservador.

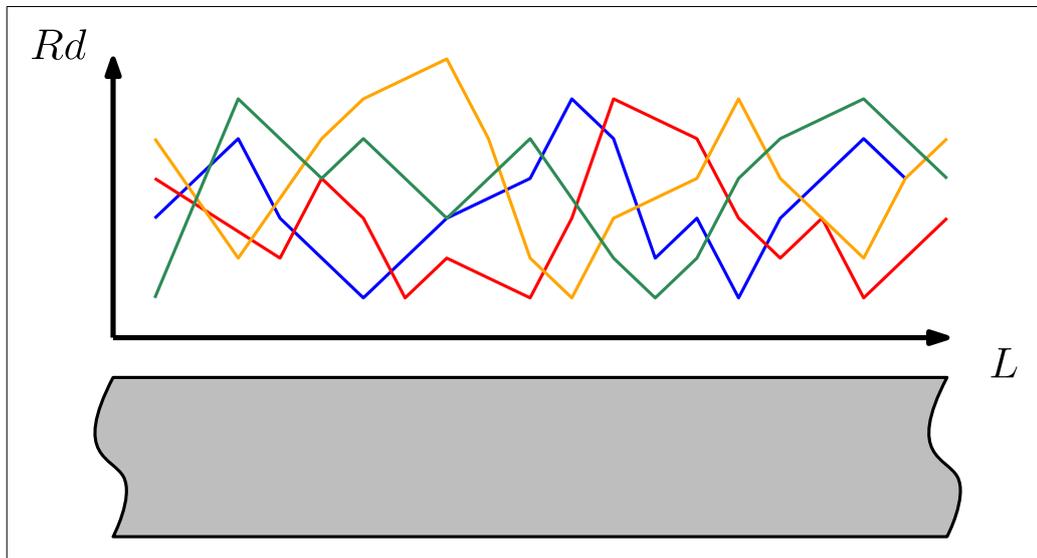
2.3.4.2 Taxa de corrosão como processo estocástico tendo a dimensão longitudinal do duto como parâmetro contínuo

Dutos são estruturas com dimensão longitudinal significativamente maior que as demais (diâmetro e espessura), o que justifica tratá-la como um contínuo ao longo do qual ocorrem fenômenos de interesse. A corrosão, por exemplo, é um destes fenômenos.

As condições ambientais que influem no processo corrosivo, como a composição química do solo, temperatura, umidade e pH, que interagem de maneira complexa com o material do duto, tendem a ser semelhantes em pontos próximos dadas diversas observações. Logo, conclui-se que são espacialmente correlacionadas. (DE LEON; MACÍAS, 2005; LI *et al.*, 2009; AMAYA-GÓMEZ *et al.*, 2024). Assim o sendo, é imediato concluir que o processo de corrosão também apresente correlação ao longo do comprimento do duto, de forma que áreas próximas sujeitas a condições ambientais similares tendem a ter comportamentos de corrosão semelhantes. Essa correlação espacial é esperada diminuir conforme a distância entre os pontos aumenta, refletindo a mudança das condições ambientais dada uma distância suficientemente grande.

De maneira mais formal, pode-se entender a taxa de corrosão R_d , num ponto qualquer do comprimento do duto (Figura 15), como uma variável aleatória que é função de várias outras variáveis aleatórias (umidade, temperatura, pH do ambiente circundante etc.), i.e. $R_d = g(X_1, X_2, \dots, X_n)$, que variam espacialmente e, portanto, que se correlacionam espacialmente consigo mesmas e para com umas às outras, inclusive para com a taxa de corrosão. Desta forma, a taxa de corrosão R_d , como resultado da complexa interação do duto com seu ambiente, termina por também ser dotada de correlação espacial. Assim, justifica-se o tratamento da taxa de corrosão como um processo estocástico ao longo do comprimento longitudinal do duto. (AMAYA-GÓMEZ *et al.*, 2024)

Figura 15 – Desenho esquemático de diversas realizações da taxa linear de corrosão Rd ao longo do contínuo L , o comprimento longitudinal do duto



Fonte: o autor (2024)

3 METODOLOGIA

A problemática proposta neste trabalho consiste no cálculo da probabilidade de falha de um duto corroído quando da presença de vários defeitos e quando as taxas lineares de crescimento da profundidade destes defeitos, Rd , são espacialmente correlacionadas.

A probabilidade de falha é calculada como um problema de falha à primeira sobrecarga e as correlações entre as taxas Rd de cada defeito são obtidas a partir de valores arbitrados para comprimento de correlação c (da função de autocorrelação) e para o espaçamento ds (entre os defeitos). Estes valores são, então, usados como argumento de duas funções de autocorrelação: triangular (linear) e exponencial quadrática.

Resultados são obtidos para uma quantidade de defeitos variando de $N_{\text{def}}=2$ a $N_{\text{def}}=5$ e tempos totais iguais a $T_d=\{12,20,25,30\}$ anos. Ademais, foram estudados o caso em que todos os defeitos têm tempo de surgimento t_0 determinístico e igual a 2.88 anos, bem como o caso em que surgem aleatoriamente no tempo conforme um processo de Poisson não-homogêneo.

3.1 CÁLCULO DA PROBABILIDADE DE FALHA DO DUTO

Neste trabalho, a probabilidade de falha do sistema, $P_{\text{sys, seg}}(t)$, para um segmento de duto, num determinado ponto t do tempo e para uma determinada quantidade N_{def} de defeitos, foi calculada por meio do Método de Monte Carlo com um total de 1×10^6 simulações. De forma a obter parâmetros estatísticos dos resultados, todos os procedimentos (amostragem de variáveis independentes; otimização para obtenção de variáveis correlacionadas; cálculo das taxas de ultrapassagem pela barreira etc.) foram repetidos um total de 5 vezes (doravante referidas como passadas). O relativamente baixo número de repetições justifica-se pela baixa dispersão de todos os resultados obtidos e aqui apresentados. Para a probabilidade de falha, por exemplo, os coeficientes de variação (Equação 15) máximos observados foram de aproximadamente 3%.

O modo de falha estudado foi o *small leak*. Na Equação 147 encontra-se a representação do evento falha para um segmento de duto, numa determinada realização ω num determinado ponto do tempo: a união dos eventos falha de cada

defeito i individualmente, de forma que se um defeito leva à falha, há falha de todo o sistema naquela realização, conforme discutido na seção 2.3.4.1 :

$$F_{\text{sys, seg}}(\omega, t) = \left\{ \bigcup_{i=1}^{N_{\text{def}}} g_{1, i, \text{seg}}(\omega, t) \right\}, \quad (147)$$

onde g_1 se refere à falha especificamente por *small leak*.

A razão entre a probabilidade de falha $P_{\text{sys, seg}}(t)$ para um determinado segmento de duto com comprimento L_{seg} e este mesmo comprimento foi tomada como igual à taxa de ultrapassagens pela barreira $v_{\text{seg}}(t)$ – que caracteriza o processo estocástico quando o parâmetro contínuo é espacial (o comprimento do duto) – para este segmento.

$$v_{\text{seg}}(t) = \frac{P_{\text{sys, seg}}(t)}{L_{\text{seg}}}, \text{ em unidades de } \frac{[\text{prob. falha}]}{[L][T]}, \quad (148)$$

esta quantidade é análoga a r na Equação 100, caso de um processo de Poisson homogêneo, ou seja: a probabilidade de falha $P_{\text{sys, seg}}(t)$ é tomada como constante ao longo do segmento para o qual foi computada.

Neste trabalho, adotou-se apenas um segmento de comprimento $\Delta L_{\text{seg}} = L_{\text{seg}}$ de forma que a taxa de ultrapassagens pela barreira num determinado ponto t do tempo, $v(t)$, pode ser calculada como na Equação 149:

$$v(t) = 1 - \exp\left(-\int_L v_{\text{seg}}(t) \cdot dL_{\text{seg}}\right), \text{ em unidades de } \frac{[\text{prob. falha}]}{[T]}, \quad (149)$$

que, discretizada, se torna:

$$v(t) = 1 - \exp\left(-\sum_{l=1}^{N_{\text{segs}}} v_{\text{seg}, l}(t) \cdot \Delta L_{\text{seg}, l}\right), \quad (150)$$

em que N_{segs} é o número de segmentos e l o índice de cada segmento a ser considerado.

Em nosso caso, como há apenas um segmento com comprimento $\Delta L_{\text{seg}} = L_{\text{seg}}$, faz-se:

$$v(t) = 1 - \exp\left(-v_{\text{seg}}(t) \cdot L_{\text{seg}}\right) = 1 - \exp\left(-P_{\text{sys, seg}}(t)\right), \quad (151)$$

chamando a atenção à similaridade com a Equação 100.

Assim, uma vez de posse da taxa de passagens pela barreira – já considerada para todos os segmentos (em nosso caso apenas um) – num determinado ponto t do tempo, a probabilidade de falha para um tempo total T_d , fica:

$$P_{f_{T_d}} = 1 - \exp\left(-\int_0^{T_d} v(t) \cdot dt\right), \text{ em unidades de } [prob. falha], \quad (152)$$

cuja versão discretizada, mais apropriada para implementação computacional, é expressa pela Equação 153:

$$P_{f_{T_d}} = 1 - \exp\left(-\sum_0^{T_d} v(t) \cdot \Delta t\right). \quad (153)$$

Resultados para $P_{\text{sys, seg}}(t)$, $v(t)$ (e conseqüentemente para o somatório na Equação 153) foram obtidos para pontos t no tempo separados por um intervalo $\Delta t = 0.125 \text{ anos}$, conforme literatura (GOMES; BECK; HAUKAAS, 2013; TEE; PESINIS, 2017; SOUSA; AFONSO; WILLMERSDORF, 2021), até um tempo T_d máximo igual a 100 *anos* e, então, armazenados em arquivos de dados. Os valores de probabilidade de falha da Equação 153 para cada valor de T_d específico (15 *anos*, 20 *anos* etc.) foram, portanto, apurados a partir da recuperação dos dados salvos. Isso foi feito de modo a evitar cálculos redundantes.

A abordagem *Common Random Numbers* (RUBINSTEIN; KROESE, 2017; MELCHERS; BECK, 2018; BECK, 2019) foi adotada – através do controle de *seeds* de geração de números aleatórios – para cada ponto da malha (composta por pares $c-ds$); para cada uma das quantidades de defeito estudada; para cada função de autocorrelação e para cada uma das 5 passadas executadas. Desta forma as simulações de Monte Carlo são executadas com rigorosamente as mesmas amostras de variáveis aleatórias para cada ponto t no intervalo $[0, T_d]$ de um determinado par $c-ds$ da malha, para uma determinada quantidade de defeitos, dada uma determinada função de autocorrelação e em uma determinada passada. Esta medida foi adotada tanto para reduzir o esforço computacional⁴, como para evitar ao máximo que qualquer efeito de dependência entre resultados observados

4 Para um $T_d = 100 \text{ anos}$ com $\Delta t = 0.125 \text{ anos}$, tem-se um total de 801 pontos discretos do tempo nos quais executar o método de Monte Carlo. Amostrando VAs 801 vezes para cada par $c-ds$, para cada quantidade de defeitos etc. é desnecessariamente custoso computacionalmente.

fosse advindo do uso repetido das mesmas amostras de variáveis aleatórias (ou seja, amostras com correlação $\rho=1$ entre si) entre diferentes pontos da malha, quantidade de defeitos etc..

3.2 OBTENÇÃO DE AMOSTRAS CORRELACIONADAS

As amostras das variáveis aleatórias Rd – as taxas lineares de crescimento da profundidade de cada defeito – foram correlacionadas de valores definidos a partir da: função de autocorrelação, comprimento de correlação c e espaçamento ds entre os defeitos. A matriz de autocorrelação R_{XX} , é portanto, função de c e ds , conforme discutido na seção 2.2.5.3 :

$$R_{XX} = \begin{bmatrix} 1 & \rho_{Rd_1 Rd_2} & \cdots & \rho_{Rd_1 Rd_n} \\ \rho_{Rd_2 Rd_1} & 1 & \cdots & \rho_{Rd_2 Rd_n} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{Rd_n Rd_1} & \rho_{Rd_n X_2} & \cdots & 1 \end{bmatrix}. \quad (154)$$

A obtenção de amostras correlacionadas as quais usar no Método de Monte Carlo se deu a partir de um processo iterativo, similar ao que foi discutido na seção 2.3.2.2.1 , e que consistiu em: 1) geração de amostras normais-padrão independentes; 2) obtenção de amostras normais-padrão correlacionadas; 3) conversão destas amostras normais-padrão correlacionadas em amostras correlacionadas e nas distribuições de interesse a partir do uso da CDF inversa; 4) ajuste iterativo da matriz de correlação das amostras normais-padrão até que a correlação desejada se manifestasse nas amostras na distribuição de interesse. Mais detalhadamente:

- Obtenção de amostras normais-padrão independentes de dimensão $N_{MC} \times N_{defs}$, em que N_{MC} é o tamanho arbitrado para o Método de Monte Carlo e N_{defs} o número de defeitos;
- Obtenção de amostras normais-padrão correlacionadas, Z , a partir da transformação expressa na Equação 126, em que a matriz L é a

decomposição de Cholesky de uma matriz de covariâncias C_Z a ser definida de maneira oportuna de forma iterativa;

- Emprego da CDF inversa em Z (Equação 155) para obtenção de amostras correlacionadas nas distribuições de interesse, X . Estas amostras, apesar de agora na distribuição desejada e correlacionadas, estarão correlacionadas de uma matriz $R_{XX \text{ calculada}}$ que pode se distanciar bastante da matriz de correlação alvo $R_{XX \text{ alvo}} = R_{XX}$ computada na Equação 154;

$$F_{X \text{ corr}} = \left\{ F_{X_1}^{-1}(\Phi(Z_1)), F_{X_2}^{-1}(\Phi(Z_2)), \dots, F_{X_n}^{-1}(\Phi(Z_n)) \right\}; \quad (155)$$

Na Equação 155, $F_{X_n}^{-1}$ se refere à CDF inversa da n -ésima variável aleatória que compõe o vetor X e $\Phi(Z_n)$ a CDF de uma distribuição normal-padrão, conforme Equação 34;

- Ajuste iterativo da matriz de correlações R_{ZZ} (que dá origem a C_Z por meio da Equação 28) até que a matriz $R_{XX \text{ calculada}}$ esteja próxima o suficiente – i.e., dentro das tolerâncias estipuladas – da matriz alvo $R_{XX \text{ alvo}}$.

Aqui, o ajuste iterativo da matriz de correlações das amostras normais-padrão, R_{ZZ} , foi efetuado por meio de otimização numérica, com algoritmo baseado em gradiente (*interior-point*⁵), de forma a minimizar a distância entre o vetor contendo os elementos da parte triangular superior (uma vez que a matriz de correlação é simétrica) de $R_{XX \text{ calculada}}$ e de $R_{XX \text{ alvo}}$. A expressão formal da otimização efetuada, é dada pela Equação 156:

$$\begin{aligned} \text{dado: } & \underline{X}_{\text{alvo}} \\ \text{minimize: } & f_{\text{obj}}(\underline{Z}), \\ \text{sujeito a: } & 0 \leq \underline{Z} \leq 1 \end{aligned} \quad (156)$$

não havendo funções com restrições de igualdade ou desigualdade impostas. (ARORA, 2017)

Na Equação 156, \underline{Z} é o vetor de variáveis de projeto (ARORA, 2017) a ser otimizado. Em nosso caso, são os elementos da parte triangular superior da matriz de correlações R_{ZZ} (excetuando-se a diagonal principal, pois seus valores serão sempre iguais a 1). Formalmente expresso Equação 157:

5 Opção padrão da função *fmincon* da linguagem de programação Matlab© (THE MATHWORKS, 2020) utilizada.

$$\underline{Z} = \left[\rho_{Z_{1,2}}, \rho_{Z_{1,3}}, \dots, \rho_{Z_{1,n}}, \rho_{Z_{2,3}}, \rho_{Z_{2,4}}, \dots, \rho_{Z_{2,n}} \right] = \left[\rho_{Z_{i,j}} \right] \quad (157)$$

$$\text{com } \begin{matrix} i=1,2,\dots,N_{\text{def}} \\ j=i+1,\dots,N_{\text{def}} \end{matrix}$$

O mesmo raciocínio se aplica a $\underline{X}_{\text{alvo}}$, parâmetro da função objetivo, e sua relação com $R_{XX \text{ alvo}}$:

$$\begin{aligned} \underline{X}_{\text{alvo}} &= \left[\rho_{X_{\text{alvo } 1,2}}, \rho_{X_{\text{alvo } 1,3}}, \dots, \rho_{X_{\text{alvo } 1,n}}, \rho_{X_{\text{alvo } 2,3}}, \rho_{X_{\text{alvo } 2,4}}, \dots, \rho_{X_{\text{alvo } 2,n}} \right] \\ \text{" " } &= \left[\rho_{Z_{i,j}} \right] \end{aligned} \quad (158)$$

$$\text{com } \begin{matrix} i=1,2,\dots,N_{\text{def}} \\ j=i+1,\dots,N_{\text{def}} \end{matrix}$$

Devido a problemas de convergência numérica observados, três otimizações foram efetuadas separadamente, cada uma com uma função objetivo. Os resultados (amostras e métricas de convergência) foram coletados e foi tomada como amostra a ser utilizada aquela oriunda da otimização que resultou em menor diferença média relativa (em valores absolutos) entre os elementos de $R_{XX \text{ calculada}}$ e $R_{XX \text{ alvo}}$. As três funções objetivo em questão foram:

– Norma do vetor diferença:

$$f_{\text{obj1}}(\underline{Z}) = \left\| \underline{X}_{\text{calc.}} - \underline{X}_{\text{alvo}} \right\| = \sqrt{\left(\rho_{X_{\text{calc.},i,j}} - \rho_{X_{\text{alvo},i,j}} \right)^2} \quad (159)$$

$$\text{com } \begin{matrix} i=1,2,\dots,N_{\text{def}} \\ j=i+1,\dots,N_{\text{def}} \end{matrix}$$

– Média dos absolutos do vetor diferença:

$$f_{\text{obj2}}(\underline{Z}) = \frac{1}{N_{\text{elem}}} \sum_1^{N_{\text{elem}}} \left| \underline{X}_{\text{calc.}} - \underline{X}_{\text{alvo}} \right| \quad \begin{matrix} i=1,2,\dots,N_{\text{def}} \\ j=i+1,\dots,N_{\text{def}} \end{matrix} \quad (160)$$

$$\text{" " } = \frac{1}{N_{\text{elem}}} \sum_1^{N_{\text{elem}}} \left| \left(\rho_{X_{\text{calc.},i,j}} - \rho_{X_{\text{alvo},i,j}} \right) \right| \quad N_{\text{elem}} = N_i * N_j$$

– Raiz do valor quadrático médio (*Root Mean Square*) do vetor diferença:

$$f_{\text{obj3}}(\underline{Z}) = \sqrt{\frac{1}{N_{\text{elem}}} \sum_1^{N_{\text{elem}}} (\underline{X}_{\text{calc.}} - \underline{X}_{\text{alvo}})^2} \quad \begin{array}{l} i=1,2,\dots,N_{\text{def}} \\ j=i+1,\dots,N_{\text{def}} \end{array}, \text{ com } \quad (161)$$

$$" " = \sqrt{\frac{1}{N_{\text{elem}}} \sum_1^{N_{\text{elem}}} (\rho_{X_{\text{calc.},i,j}} - \rho_{X_{\text{alvo},i,j}})^2} \quad N_{\text{elem}} = N_i * N_j$$

Por ser baseado em gradiente, o algoritmo de otimização utilizado necessita de um ponto inicial. Este foi, portanto, tido como os próprios elementos de $\underline{R}_{XX \text{ alvo}}$. Esta decisão foi tomada partindo da ideia de que o ponto ótimo estaria na vizinhança do ponto inicial, conforme Beck (2019), que comenta que a razão r , expressa na Equação 162

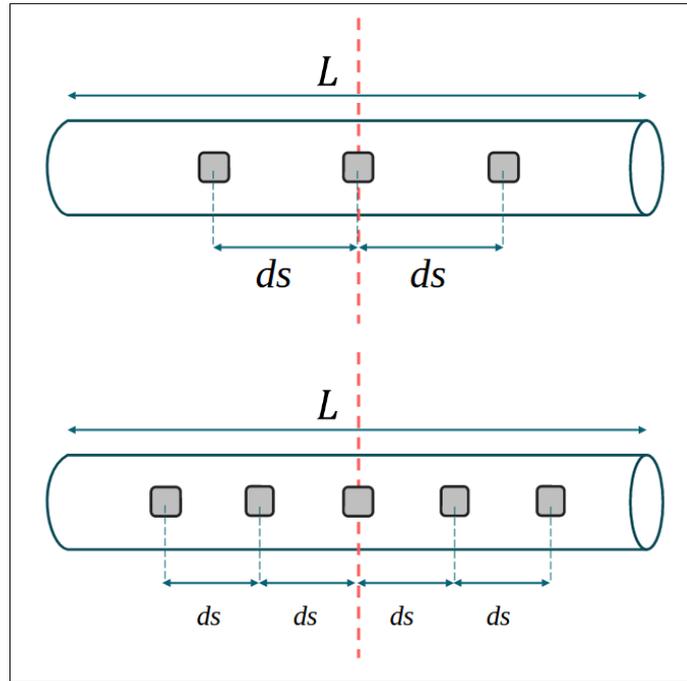
$$r = \frac{\rho_{Z_{i,j}}}{\rho_{X_{i,j}}}, \quad (162)$$

em geral não difere muito da unidade.

3.3 ESTUDO DE CASO

Os valores arbitrados para os espaçamentos entre os defeitos, ds , foram: $ds=[0.1,0.15,0.20,0.25]$; para o comprimento de correlação, c , foram $c=[0.25,0.50,0.75,1]$. A Figura 16 ilustra a distribuição idealizada de defeitos no duto.

Figura 16 – Distribuição de defeitos igualmente espaçados num duto



Fonte: o autor (2024)

As expressões para as funções de autocorrelação (das taxas lineares de crescimento da profundidade dos defeitos) utilizadas são (BECK, 2019):

$$R_{XX \text{ LINEAR}} = \begin{cases} 1 - \frac{|ds|}{c}, & \text{se } |ds| \leq c, \text{ com } c > 0 \\ 0, & \text{caso contrário} \end{cases} \quad (163)$$

$$R_{XX \text{ EXPO. QUAD}} = \exp(-c^2) \cdot ds^2, \text{ com } c > 0 \quad (164)$$

conforme Figura 5 e Figura 6.

As distribuições, valores de média e coeficiente de variação das variáveis aleatórias envolvidas no problema são dados na Tabela 1 abaixo:

Tabela 1 – Variáveis aleatórias do problema

Variável	Média	CoV [%]	Distribuição
Espessura da parede (wt)	8,96 [mm]	1,5 %	Normal
Profundidade inicial do defeito (d_0)	2,64 [mm]	0,31 %	Gama
Taxa de corrosão (R_d)	0,082 [mm/ano]	0,65 %	Gama

Fonte: d_0 e R_d : Bazán e Beck (2013); wt : Tee e Pesinis (2017).

Quanto aos tempos para surgimento dos defeitos, tem-se, como já mencionado ao início do capítulo, o caso em que $t_0 = 2.88 \text{ anos}$ e outro no qual

surtem aleatoriamente no tempo seguindo um processo de Poisson não-homogêneo, conforme discutido ao fim da seção 2.3.4 . Os parâmetros que definem este processo são λ_0 e b e seus valores foram tomados como iguais a 0.0064 e 1, respectivamente, conforme Tee & Pesinis (2017).

3.4 AJUSTES DE CURVA

Superfícies que mostram resultados (diferenças máximas e médias entre a matriz R_{XX} calculada e R_{XX} alvo, bem como probabilidade de falha) em função dos valores de (c, ds) adotados foram condensadas em gráficos de dispersão, 2D, em função da correlação entre dois defeitos adjacentes de forma a se observar um comportamento geral e independente de valores específicos de c e ds . De posse destes gráficos, ajustes de curva foram realizados com o uso de ferramentas automáticas⁶ nativas da linguagem de programação utilizada. Foram obtidos, então, valores de R^2 (coeficiente de determinação, uma medida de *goodness of fit*⁷ de um determinado ajuste de curva (LANCASTER; ŠALKKAUSKAS, 1986)) de várias curvas diferentes ajustadas aos valores de probabilidade de falha referentes a cada quantidade de defeitos. Nas figuras, bem como nas tabelas do Apêndice B, são exibidos os valores do maior R^2 dentre as curvas testadas.

Abaixo, tabela com curvas utilizadas:

Tabela 2 – Modelos de curva para ajuste aos valores dos gráficos de dispersão

Formato	Descrição
$Y = p1*x+p2$	polinômio grau 1 (linear)
$Y = p1*x^2+p2*x+p3$	polinômio grau 2 (quad.)
$Y = p1*x^3+p2*x^2+...+p4$	polinômio grau 3 (cúbico)
$Y = a*\log(x)+b$	função logarítmica
$Y = a*\log(x+b)+c$	função log. custom
$Y = a*x^b+c$	função expoente
$Y = a*\exp(b*x)$	função exponencial
$Y = a*\exp(b*x)+c*\exp(d*x)$	função exponencial tipo 2
$Y = a*(x-b)^n$	função exponencial custom

Fonte: o autor (2024)

6 A *toolbox* específica (*Curve Fitting Toolbox*) para esta tarefa na linguagem de programação científica Matlab© (THE MATHWORKS, 2020).

7 Uma medida da concordância do comportamento dos dados a uma determinada função.

4 RESULTADOS

O estudo se dividiu em duas partes principais: a primeira sendo referente à obtenção das amostras correlacionadas para R_d ; e a segunda aos resultados obtidos para a probabilidade de falha do duto para um tempo total T_d considerado.

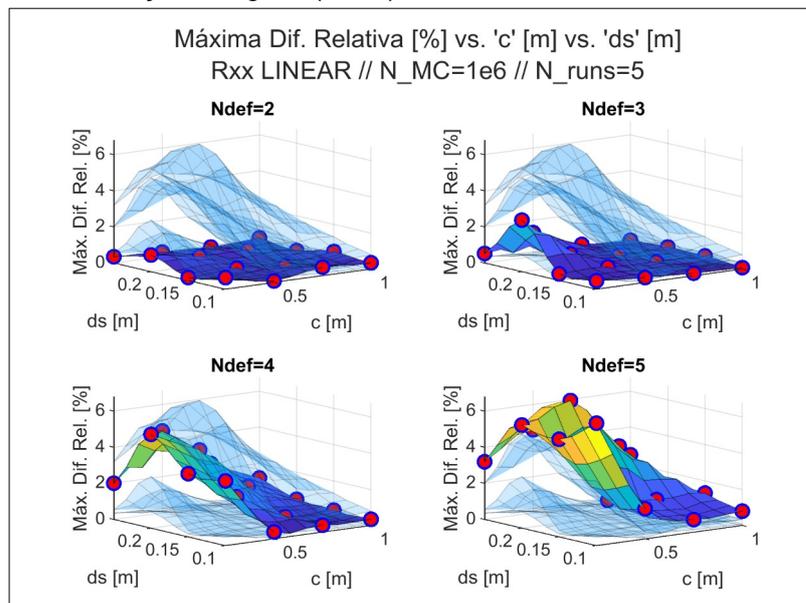
4.1 APROXIMAÇÕES OBTIDAS PARA CORRELAÇÃO DAS AMOSTRAS DAS TAXAS LINEARES DE CRESCIMENTO DA PROFUNDIDADE DOS DEFEITOS

Uma vez que os valores para R_{XX} calculada são obtidos a partir de um processo de otimização numérica, é de se esperar que haja desvios em relação a R_{XX} alvo para os elementos fora da diagonal principal. Esses desvios, de fato, dependem do número de defeitos, do par (c, ds) e da função de autocorrelação da qual se trata. As formas com que isto se dá estão expostas nas Figuras 16 a 19.

4.1.1 Função de autocorrelação triangular (linear)

Para a função de autocorrelação triangular (linear), tem-se um comportamento para a máxima diferença relativa (em módulo) entre R_{XX} calculada e R_{XX} alvo, plotada em função de c e ds conforme exposto na Figura 17:

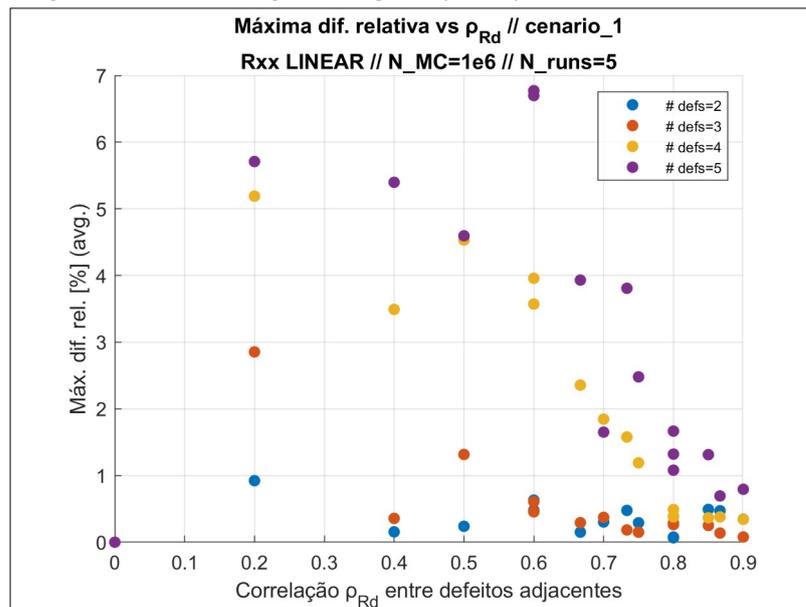
Figura 17 – Máxima diferença relativa (em módulo) para função de autocorrelação triangular (linear)



Fonte: o autor (2024)

Em que se nota um claro comportamento de crescimento das diferenças relativas com o aumento do número de defeitos. Quando em função do coeficiente de correlação ρ_{Rd} entre defeitos adjacentes, nota-se que há uma tendência de queda dos valores de diferença relativa máxima com o aumento da correlação, conforme Figura 18:

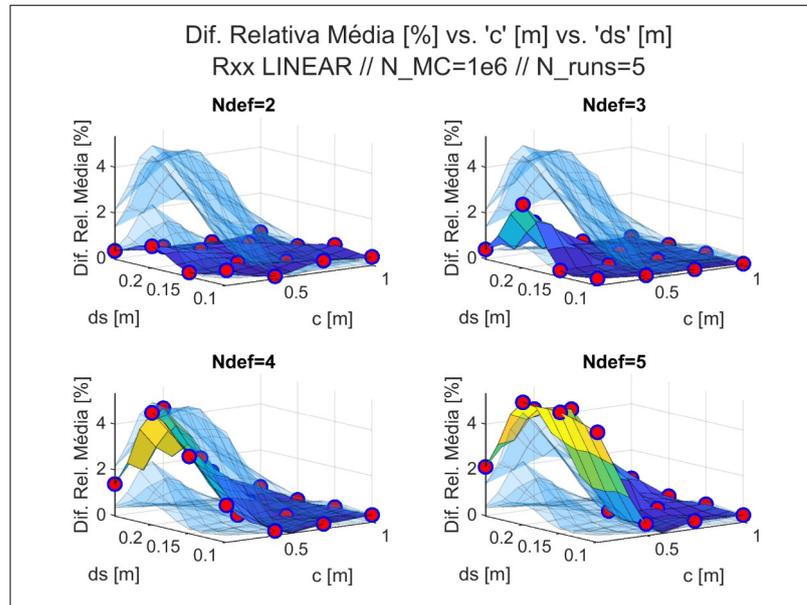
Figura 18 – Máxima diferença relativa (em módulo) vs. ρ_{Rd} para função de autocorrelação triangular (linear)



Fonte: o autor (2024)

O mesmo pode ser observado para os valores médios das diferenças relativas, como na Figura 19.

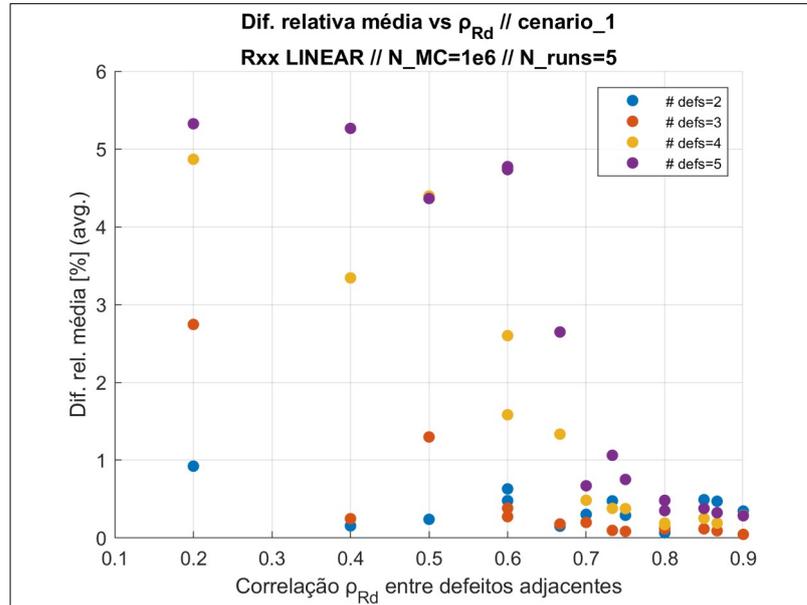
Figura 19 – Diferença relativa média (em módulo) para função de autocorrelação triangular (linear)



Fonte: o autor (2024)

É possível perceber que aqui, assim como para a máxima diferença relativa na Figura 17, os pares (c, ds) em que $c > ds$ tendem a apresentar menores valores de diferença média. Tendo em vista a Equação 163, constata-se que estes são justamente os valores em que há maior coeficiente de correlação entre defeitos adjacentes ρ_{Rd} , e, com isso, o comportamento da diferença relativa média como função desta correlação segue o comportamento da máxima diferença relativa na Figura 18, como é constatado na Figura 20.

Figura 20 – Diferença relativa média (em módulo) vs. ρ_{Rd} para função de autocorrelação triangular (linear)

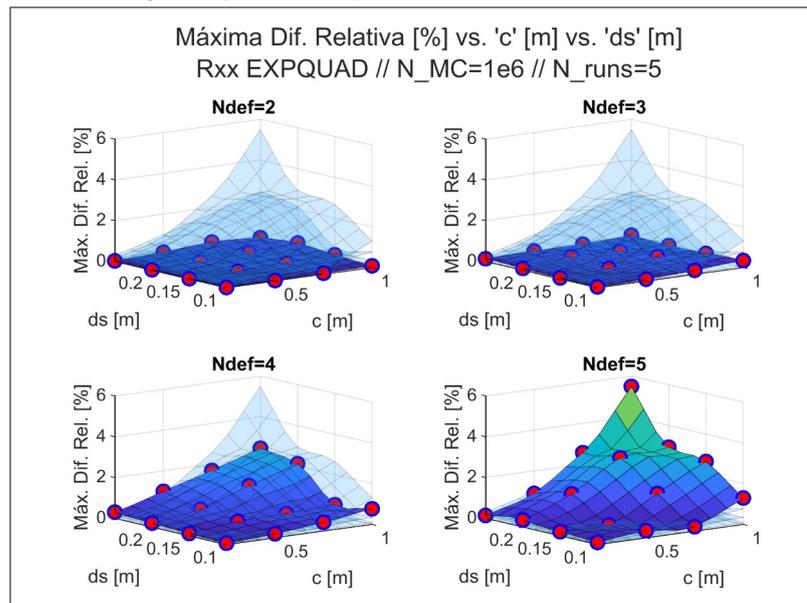


Fonte: o autor (2024)

4.1.2 Função de autocorrelação exponencial quadrática

Quando a função de autocorrelação adotada é a exponencial quadrática, o formato das superfícies para máxima diferença relativa como função dos pares (c, ds) mudam radicalmente, como exposto na Figura 21.

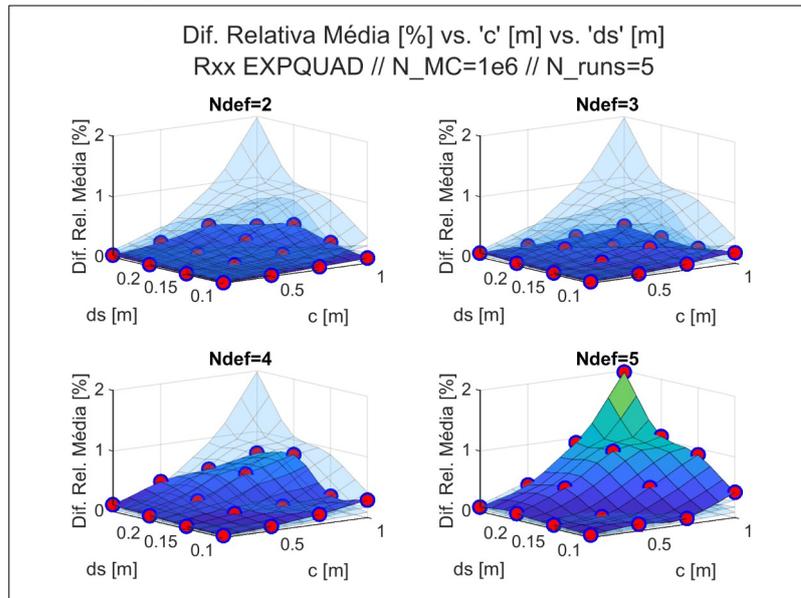
Figura 21 – Máxima diferença relativa (em módulo) para função de autocorrelação exponencial quadrática



Fonte: o autor (2024)

Para a diferença relativa média, a mudança em relação ao observado na Figura 19 para função de autocorrelação triangular (linear) se mantém, como observado na Figura 22.

Figura 22 – Diferença relativa média (em módulo) para função de autocorrelação exponencial quadrática

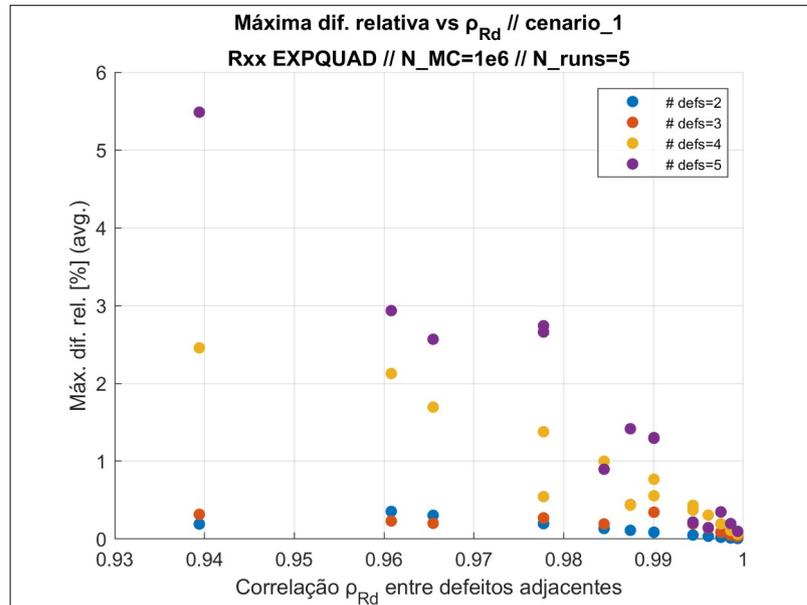


Fonte: o autor (2024)

Através das Figuras 21 e 22 constata-se que, apesar de os valores máximos para máxima diferença relativa serem similares entre a função de autocorrelação triangular e a exponencial quadrática, os valores médios para esta última são bem menores. Isto fica ainda mais explícito quando se observa a relação destes valores com o coeficiente de correlação ρ_{Rd} entre defeitos adjacentes. Ademais, o crescimento dos valores médios das diferenças relativas para a função de autocorrelação exponencial quadrática é, também, muito mais sutil quando comparado ao que se observa para a função de autocorrelação triangular (linear) na 19. Isto induz à conclusão de que, excetuando-se os valores máximos, que são similares entre as duas funções de autocorrelação, o restante das correlações que compõem as matrizes $\mathbf{R}_{XX \text{ calculada}}$ se comporta de maneira muito mais fiel a $\mathbf{R}_{XX \text{ alvo}}$ quando a função de autocorrelação é exponencial quadrática.

Para a máxima diferença relativa, o comportamento é dado na Figura 23

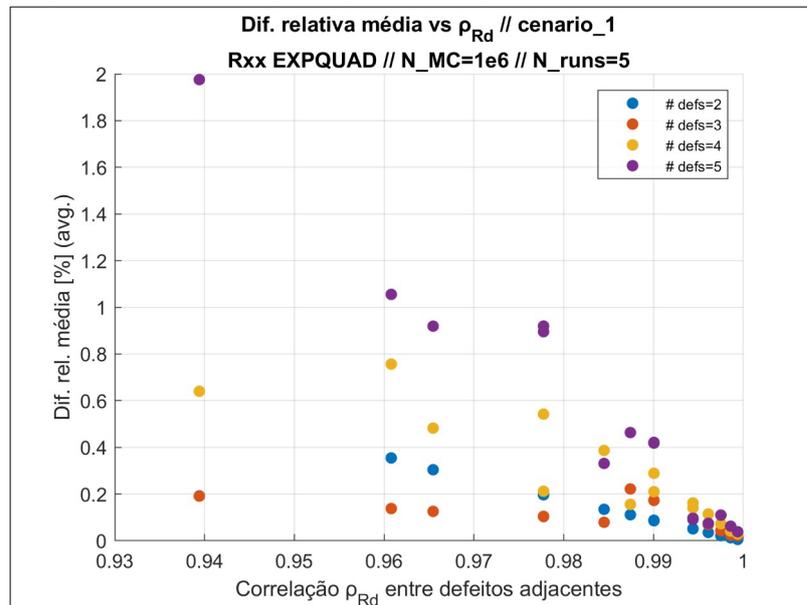
Figura 23 – Máxima diferença relativa (em módulo) vs. ρ_{Rd} para função de autocorrelação exponencial quadrática



Fonte: o autor (2024)

E para a diferença relativa média, a relação com ρ_{Rd} é dada na Figura 24:

Figura 24 – Diferença relativa média (em módulo) vs. ρ_{Rd} para função de autocorrelação exponencial quadrática



Fonte: o autor (2024)

Das Figuras 23 e 24 nota-se ainda mais explícita a tendência de queda das diferenças relativas, tanto máximas quanto médias, com o aumento da correlação ρ_{Rd} entre defeitos adjacentes.

4.2 PROBABILIDADE DE FALHA DO DUTO

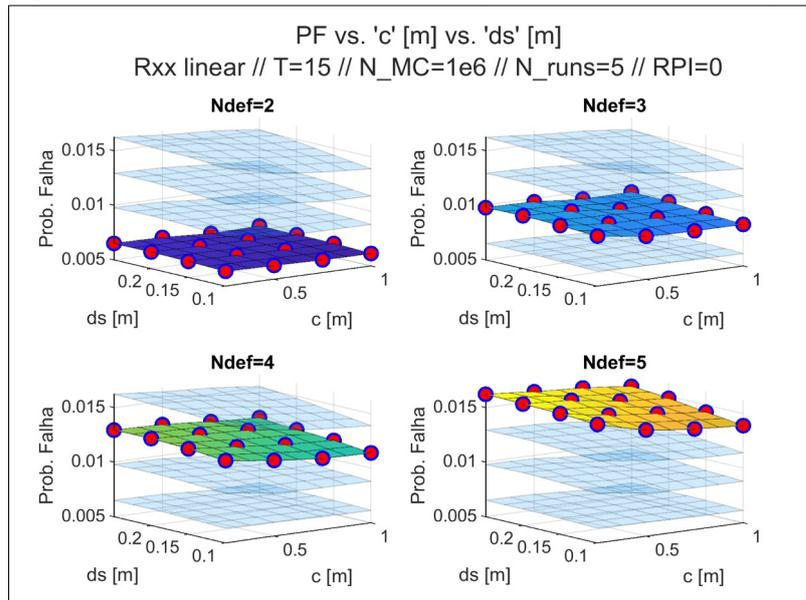
Esta seção apresenta os resultados da probabilidade de falha total Pf_{Td} , calculada conforme a Equação 153, em função do comprimento de correlação c e do espaçamento ds entre os defeitos. Essas quantidades caracterizam completamente o comportamento das correlações entre as taxas de crescimento da profundidade desses defeitos (conforme as Equações 163 e 164) e pretende-se mostrar se e como, por conseguinte, influenciam nos valores da probabilidade de falha total Pf_{Td} obtidos. São expostos os resultados para os casos em que o tempo para surgimento dos defeitos é determinístico e para o caso em que é aleatório segundo modelo NHPP mencionado anteriormente.

4.2.1 Função de autocorrelação triangular (linear) – tempo determinístico para o surgimento dos defeitos

Para o caso em que a função de autocorrelação imposta a Rd é linear e $t_0=2,88$ anos, tem-se o seguinte comportamento descrito nas Figuras 25 a 28:

- Tempo total $T_d=15$ anos:

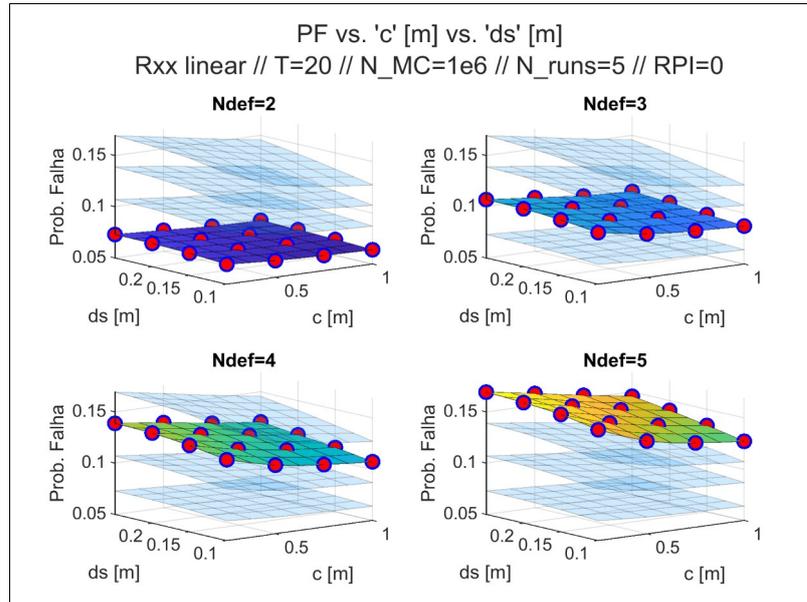
Figura 25 – Prob. de falha vs. 'c' vs.'ds' (Rxx: linear, Td=15 anos)



Fonte: o autor (2024)

- Tempo total $T_d=20$ anos:

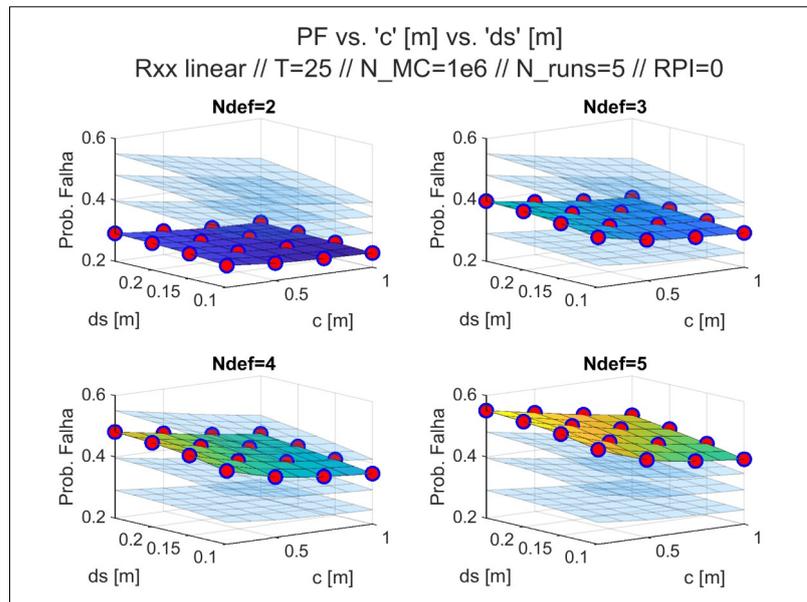
Figura 26 – Prob. de falha vs. 'c' vs.'ds' (Rxx: linear, Td=20 anos)



Fonte: o autor (2024)

– Tempo total $T_d=25$ anos:

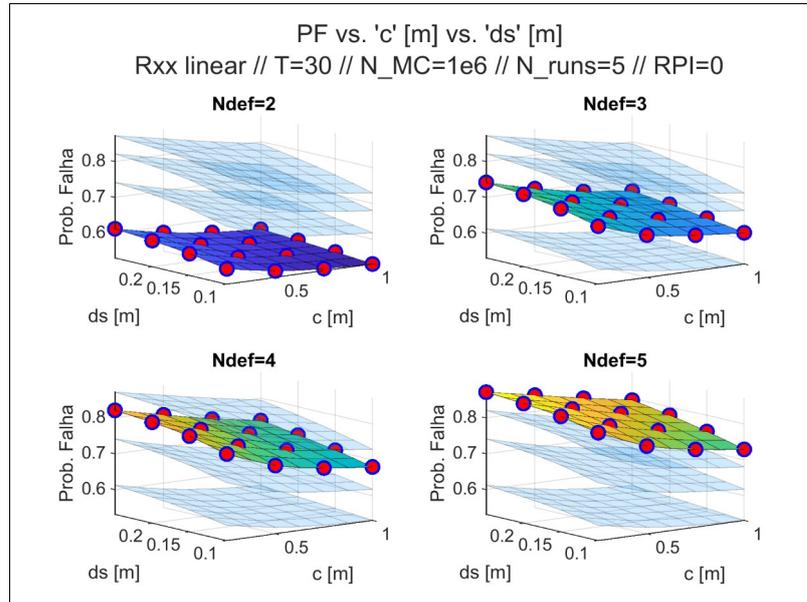
Figura 27 – Prob. de falha vs. 'c' vs.'ds' (Rxx: linear, Td=25 anos)



Fonte: o autor (2024)

– Tempo total $T_d=30$ anos:

Figura 28 – Prob. de falha vs. 'c' vs.'ds' (Rxx: linear, Td=30 anos)



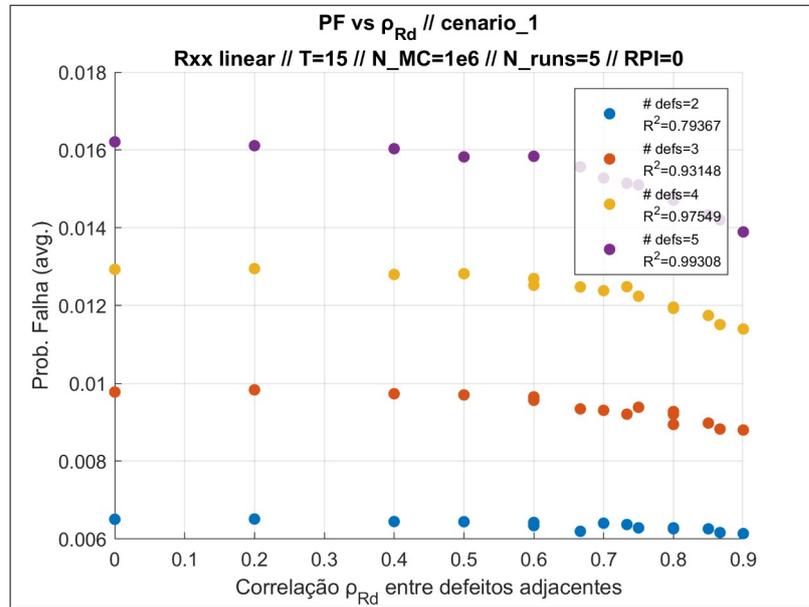
Fonte: o autor (2024)

Das Figuras 25 a 28 é notável o comportamento de mudança suave dos valores da probabilidade de falha *versus* comprimento de correlação c e espaçamento ds . Em todos os casos se tem, como é esperado, um aumento da probabilidade de falha total, Pf_{Td} , com o aumento do número de defeitos. É possível notar, também, que, quanto maiores os valores de tempo total T_d considerados, mais inclinadas se tornam as superfícies que denotam o comportamento da probabilidade de falha total. Aqui, o termo “inclinado” deve ser entendido como naquele sentido em que a probabilidade de falha Pf_{Td} passa a ser maior para aqueles pares (c, ds) em que: 1) ds aumenta e c diminui e 2) c aumenta e ds diminui.

Este comportamento é confirmado pelas Figuras 29 a 32 que mostram a variação da probabilidade de falha calculada em função da correlação ρ_{Rd} entre as taxas lineares de crescimento da profundidade de defeitos adjacentes, que é, naturalmente, função de (c, ds) :

- Tempo total $T_d=15$ anos:

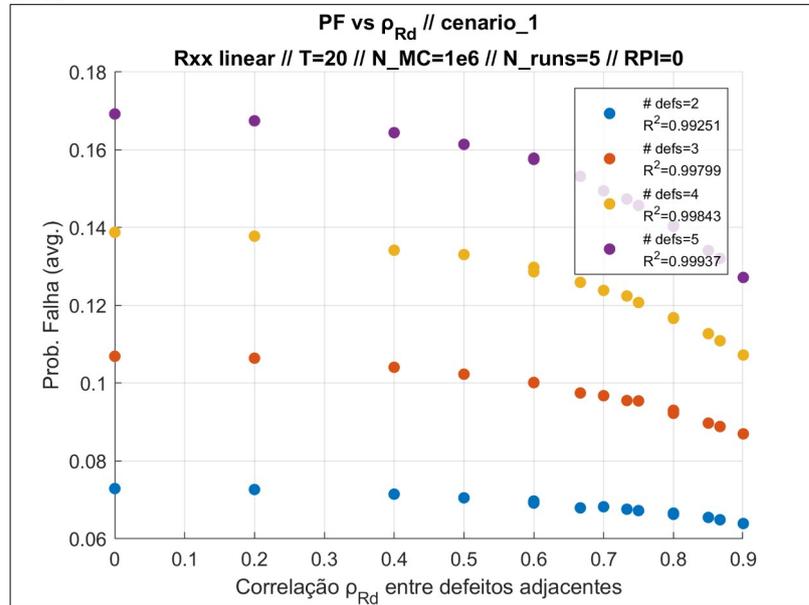
Figura 29 – Prob. falha vs. ρ_{Rd} (linear) para $T_d=15$ anos. T_0 : 2.88 anos



Fonte: o autor (2024)

– Tempo total $T_d=20$ anos:

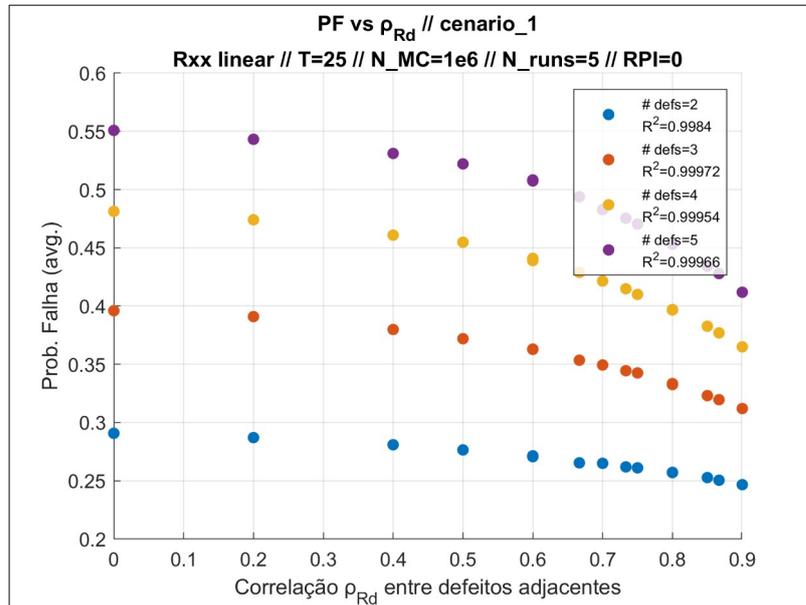
Figura 30 – Prob. falha vs. ρ_{Rd} (linear) para $T_d=20$ anos. T_0 : 2.88 anos



Fonte: o autor (2024)

– Tempo total $T_d=25$ anos:

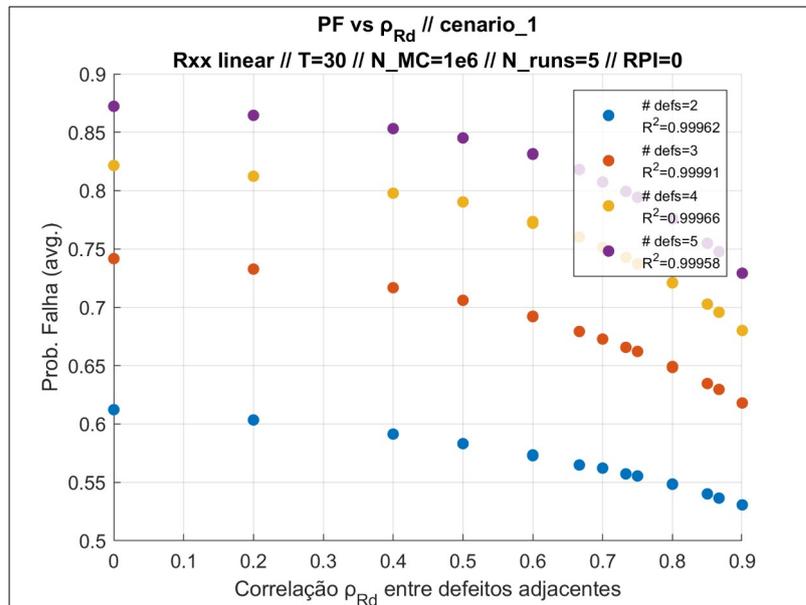
Figura 31 – Prob. falha vs. ρ_{Rd} (linear) para $T_d=25$ anos. T_0 : 2.88 anos



Fonte: o autor (2024)

– Tempo total $T_d=30$ anos:

Figura 32 – Prob. falha vs. ρ_{Rd} (linear) para $T_d=30$ anos. T_0 : 2.88 anos



Fonte: o autor (2024)

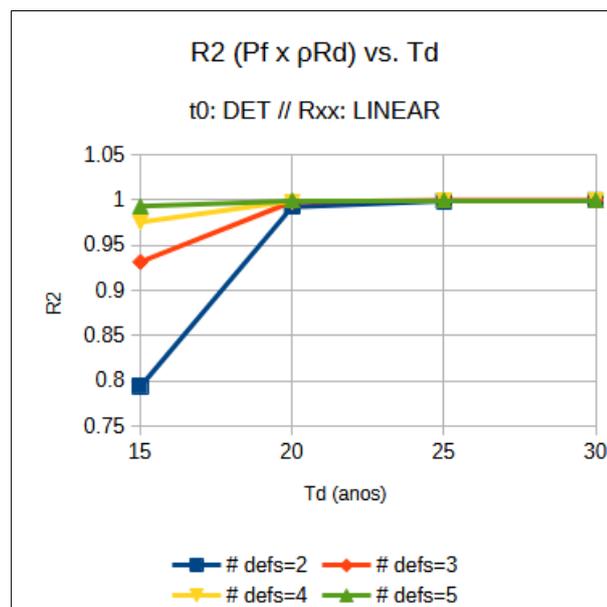
Nas Figuras 29 a 32 encontram-se gráficos de dispersão que condensam as informações das superfícies das Figuras 25 a 28: a probabilidade de falha calculada, Pf_{T_d} , versus a correlação ρ_{RD} definida para cada par (c, ds) que compõe a malha de

cada superfície. O comportamento da probabilidade Pf_{Td} é claro e no sentido de diminuir quando a correlação ρ_{RD} aumenta, corroborando as observações de que a não consideração dos efeitos das dependências entre as variáveis aleatórias do problema leva a resultados conservadores. (DE LEON; MACÍAS, 2005; LI et al., 2009; QIAN et al., 2013b; FU et al., 2019; WANG; XIE; SU, 2024)

Os valores de R^2 nas Figuras 29 a 32 dizem respeito aos melhores ajustes de curva obtidos dentre as funções-modelo expostas na Tabela 2. Tabelas-resumo com as funções ajustadas e os respectivos valores de R^2 para cada quantidade de defeitos, para cada T_d , para cada função de autocorrelação e cada tipo de tempo de surgimento (determinístico ou NHPP) para cada caso são fornecidas nos apêndices. De forma geral, no entanto, é imediato notar que, para todas as quantidades de defeitos estudadas, há um aumento no valor de R^2 para maiores valores de T_d . Isto indica que, à medida que um maior tempo é considerado, a probabilidade de falha Pf_{Td} tende a um comportamento prescrito.

Este fenômeno pode ser melhor observado graficamente na Figura 33. É possível perceber que, exceto para o $T_d=15$ anos e um número de defeitos igual a 2 e 3, todos os outros pontos mostram valores bastante altos para os coeficientes de determinação R^2 .

Figura 33 – R^2 vs. T_d para vários números de defeito considerados



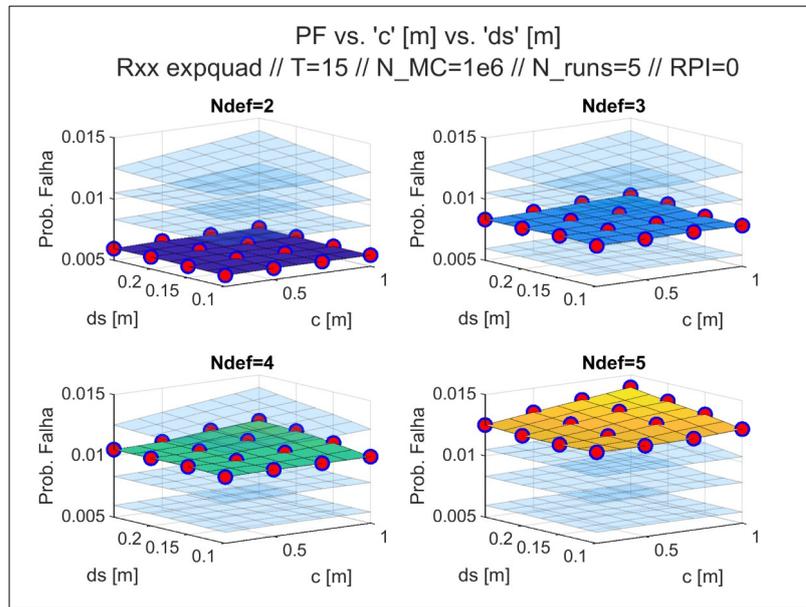
Fonte: o autor (2024)

4.2.2 Função de autocorrelação exponencial quadrática – tempo determinístico para o surgimento dos defeitos

Para o caso em que a função de autocorrelação imposta a Rd é exponencial quadrática e $t_0=2,88\text{ anos}$, tem-se que as superfícies da probabilidade de falha Pf_{Td} versus comprimento de correlação c e espaçamento ds apresentam comportamentos ligeiramente diferentes daqueles apresentados na seção 4.2.1 .

– Tempo total $T_d=15\text{ anos}$:

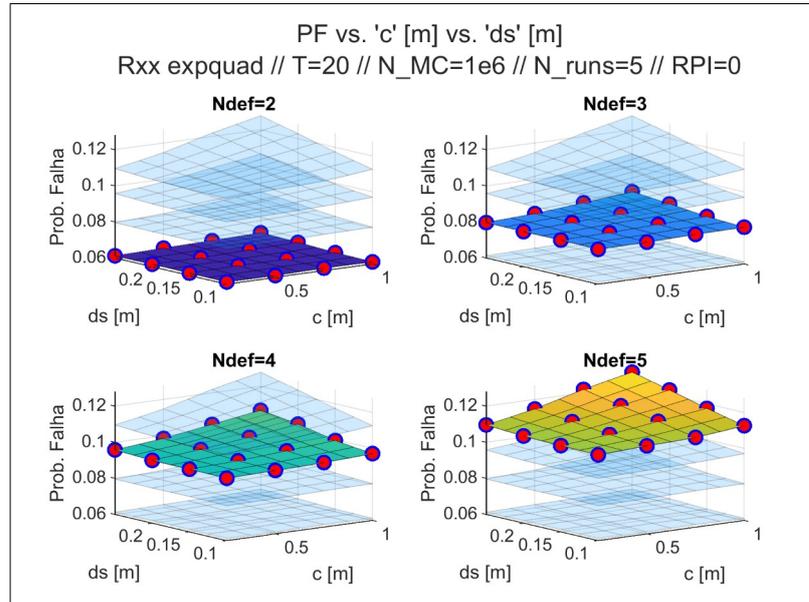
Figura 34 – Prob. de falha vs. 'c' vs.'ds' (Rxx: expquad, Td=15 anos)



Fonte: o autor (2024)

– Tempo total $T_d=20\text{ anos}$:

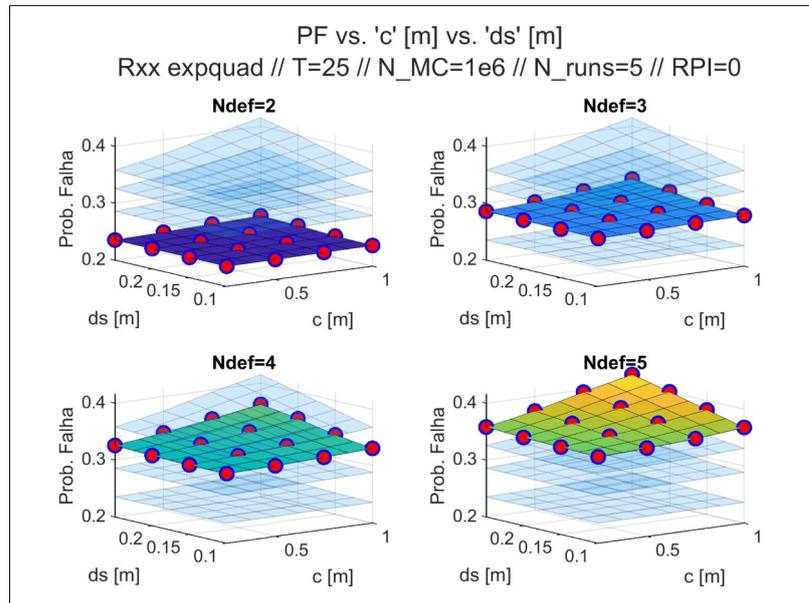
Figura 35 – Prob. de falha vs. 'c' vs.'ds' (Rxx: expquad, Td=20 anos)



Fonte: o autor (2024)

– Tempo total $T_d=25$ anos:

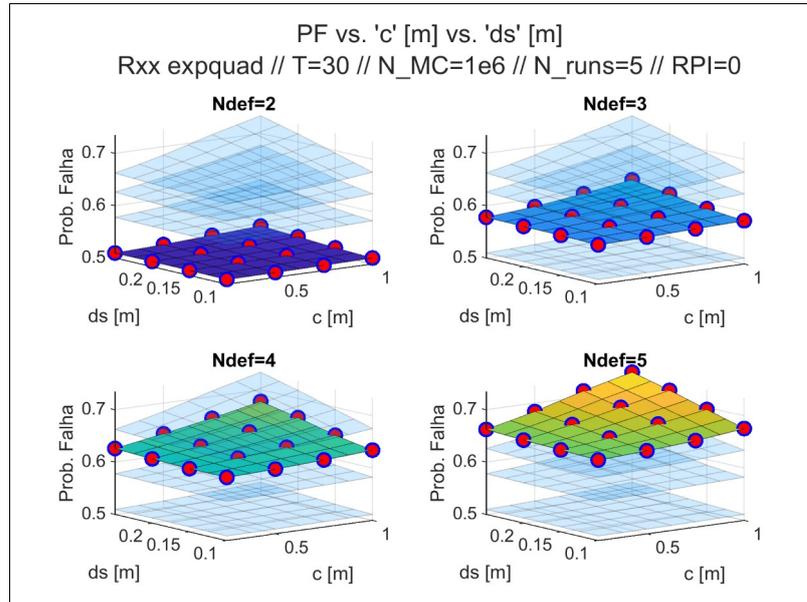
Figura 36 – Prob. de falha vs. 'c' vs.'ds' (Rxx: expquad, Td=25 anos)



Fonte: o autor (2024)

– Tempo total $T_d=30$ anos:

Figura 37 – Prob. de falha vs. 'c' vs.'ds' (Rxx: expquad, Td=30 anos)

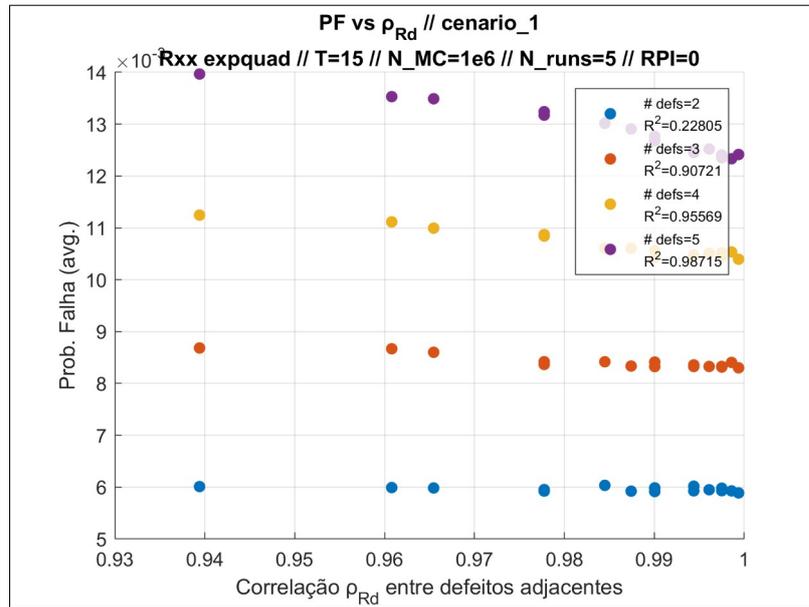


Fonte: o autor (2024)

Novamente há um comportamento conforme o esperado: maior número de defeitos implicando maior probabilidade de falha dado um mesmo tempo total T_d sendo levado em consideração; e, para uma mesma quantidade de defeitos, um aumento da probabilidade de falha com o crescimento do tempo total T_d . É fácil perceber, no entanto, que nas Figuras 34 a 37 as superfícies inclinam-se num sentido diferente que aquele para a função de autocorrelação triangular (linear). Aqui há um aumento da probabilidade de falha quando do aumento simultâneo de c e de ds . Este detalhe, porém, não se reflete nos gráficos da probabilidade de falha Pf_{T_d} versus ρ_{RD} , conforme Figuras 38 a 41.

– Tempo total $T_d=15$ anos:

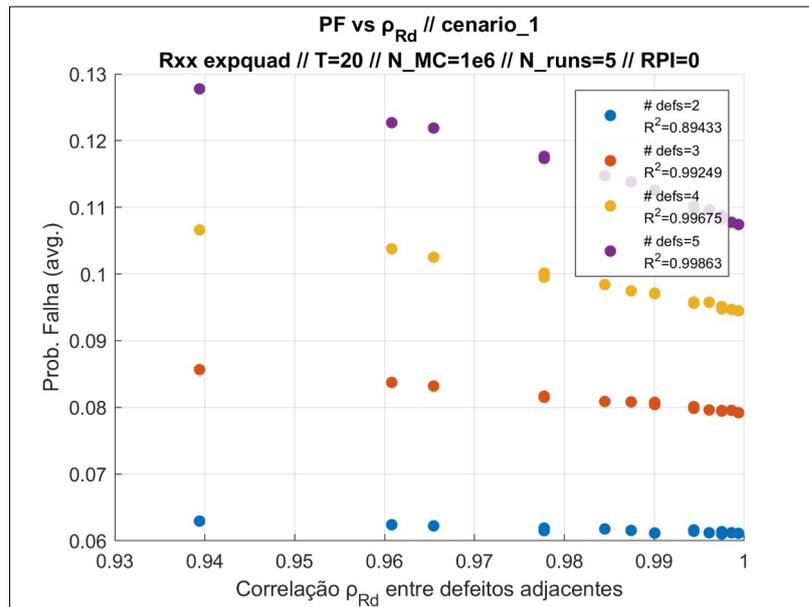
Figura 38 – Prob. falha vs. ρ_{Rd} (expquad) para $T_d=15$ anos. T_0 : 2.88 anos



Fonte: o autor (2024)

– Tempo total $T_d=20$ anos:

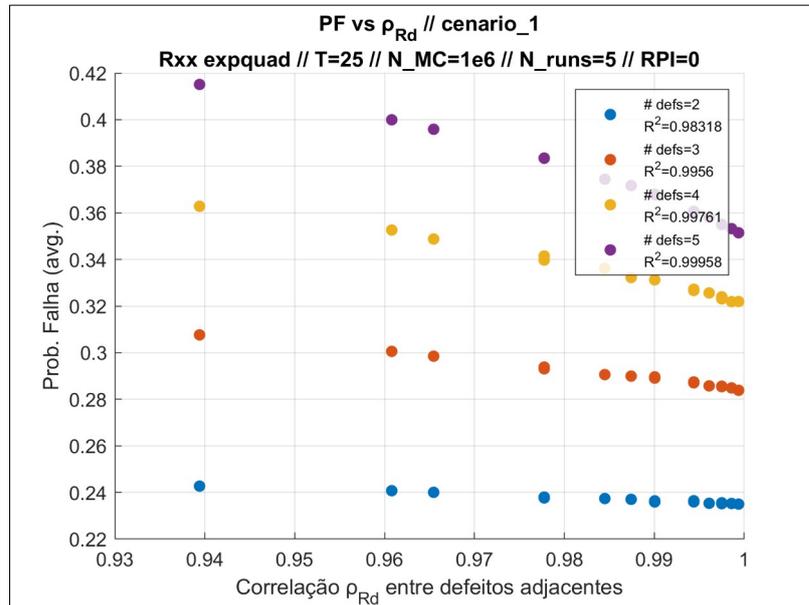
Figura 39 – Prob. falha vs. ρ_{Rd} (expquad) para $T_d=20$ anos. T_0 : 2.88 anos



Fonte: o autor (2024)

– Tempo total $T_d=25$ anos:

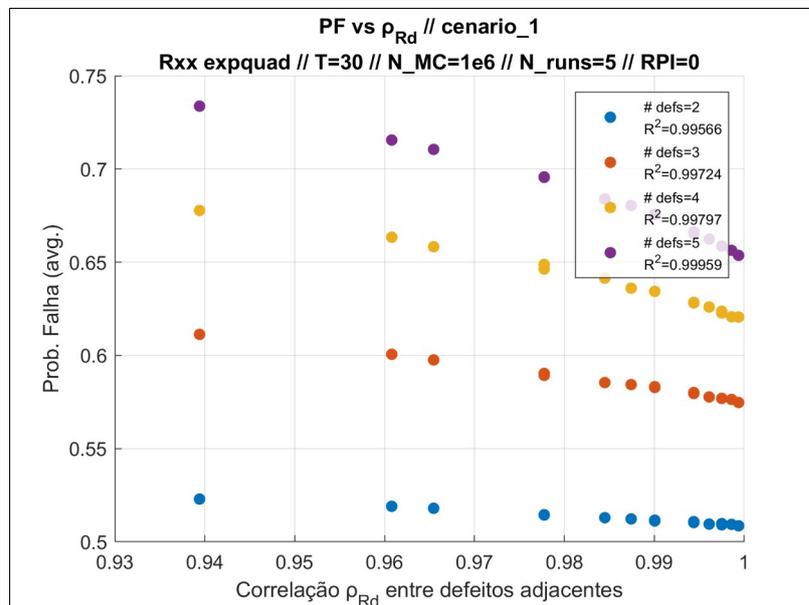
Figura 40 – Prob. falha vs. ρ_{Rd} (expquad) para $T_d=25$ anos. T_0 : 2.88 anos



Fonte: o autor (2024)

– Tempo total $T_d=30$ anos:

Figura 41 – Prob. falha vs. ρ_{Rd} (expquad) para $T_d=30$ anos. T_0 : 2.88 anos



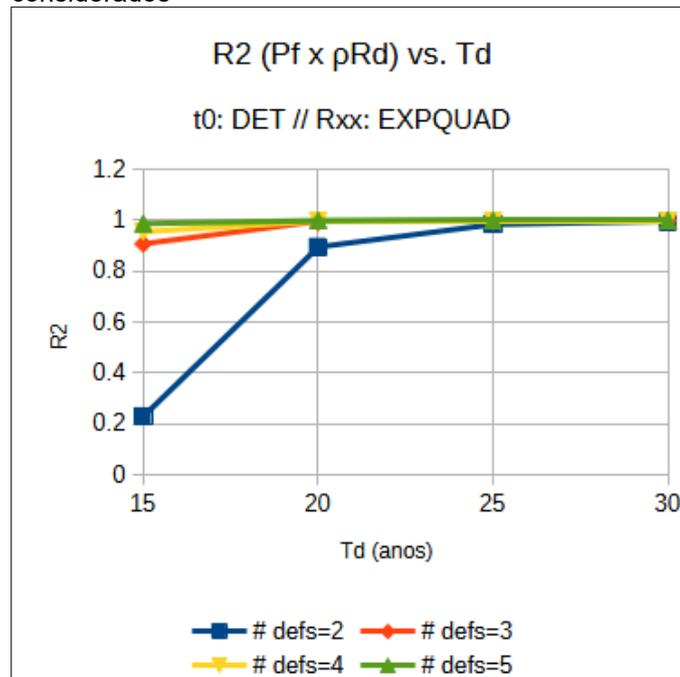
Fonte: o autor (2024)

Das Figuras 38 a 41 é possível perceber que há bastante semelhança entre os comportamentos quando do emprego da função de autocorrelação triangular (linear) em comparação com a função exponencial quadrática – mesmo

considerando que, neste último caso, a faixa de valores para a correlação ρ_{RD} é muito mais estreita e se concentra em valores sempre acima de 0.9.

O gráfico de R^2 vs. T_d , como consequência, também apresenta uma evolução de formato similar (Figura 42), com os coeficientes de determinação sendo, em geral, bastante altos e crescendo conforme aumenta T_d . Apesar disso, a observação feita para o caso em que o número de defeitos é igual a 2 e $T_d=15\text{anos}$ é, aqui, ainda mais válida, uma vez que o R^2 deste ponto é baixíssimo, implicando que o comportamento da probabilidade Pf_{T_d} não é bem determinado a partir da função obtida pelo ajuste de curva.

Figura 42 – R^2 vs. T_d para vários números de defeito considerados



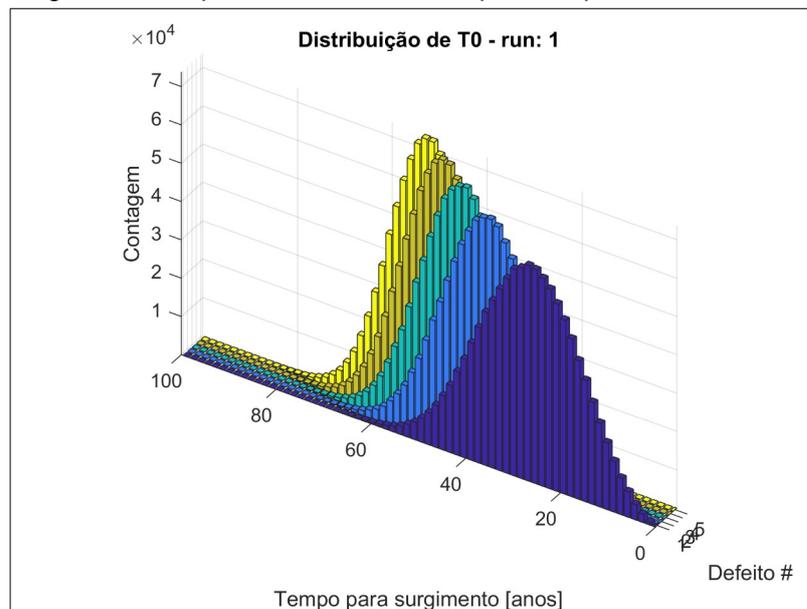
Fonte: o autor (2024)

4.2.3 Função de autocorrelação triangular (linear) – tempo aleatório para surgimento dos defeitos e segundo modelo NHPP

Para o caso em que a função de autocorrelação imposta a R_d é linear e t_0 segue modelo NHPP notou-se um comportamento diferente daquele anteriormente observado para a relação entre Pf_{T_d} e ρ_{RD} até $T_d=30\text{anos}$: aqui não houve nenhuma aderência satisfatória à curva ajustada mesmo para este tempo de vida útil considerado. Atribui-se este comportamento ao surgimento mais tardio dos defeitos

quando da consideração destes como seguindo o modelo NHPP com os parâmetros utilizados: o primeiro defeito surge, em média, aproximadamente aos 28 anos; o segundo defeito, aproximadamente aos 37 anos, conforme o caso para a primeira passada ilustrado na Figura 43. Mais detalhes das realizações dos tempos de surgimento por NHPP são encontrados nos apêndices.

Figura 43 – Histograma com distribuições dos tempos de surgimento dos primeiros 5 defeitos na primeira passada

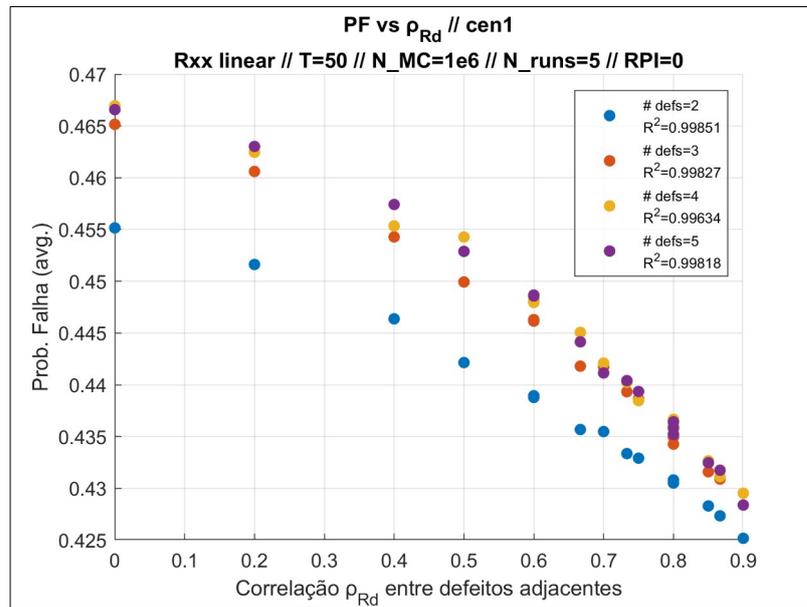


Fonte: o autor (2024)

Como forma de verificar o resultado obtido, foram computados os valores para o caso em que T_d é igual a 50, 75 e 100 anos. As respostas da probabilidade de falha Pf_{T_d} para 50 e 75 anos são expostas, a fim de ilustração, nas Figuras 44 e 45:

- Tempo total $T_d=50$ anos:

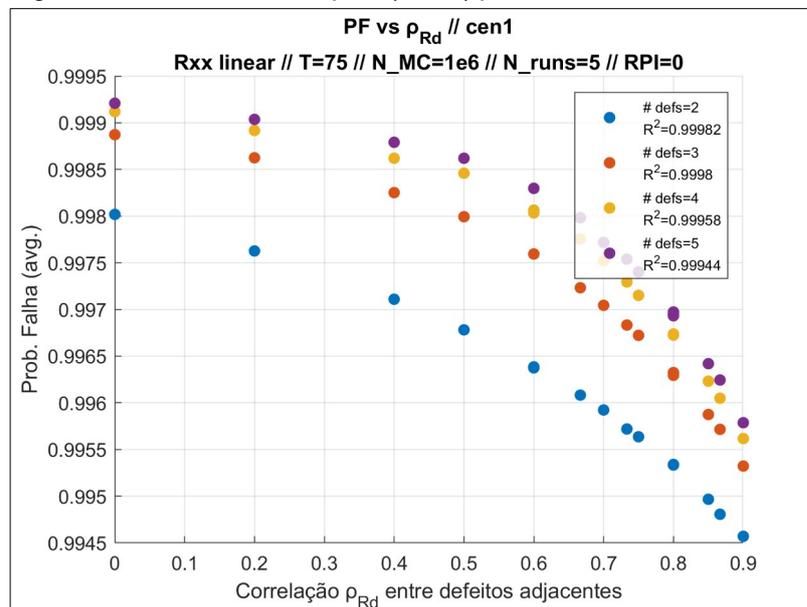
Figura 44 – Prob. falha vs. ρ_{Rd} (linear) para $T_d=50$ anos. T0: NHPP



Fonte: o autor (2024)

– Tempo total $T_d=75$ anos:

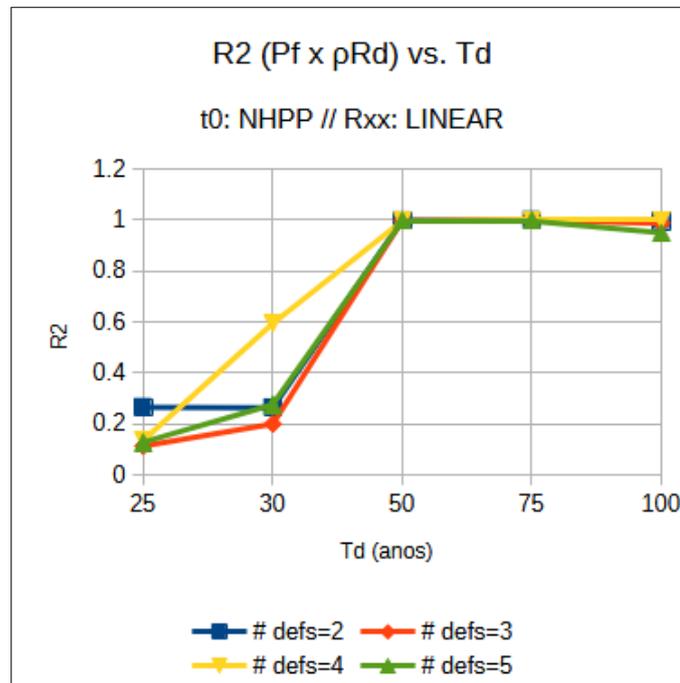
Figura 45 – Prob. falha vs. ρ_{Rd} (linear) para $T_d=75$ anos. T0: NHPP



Fonte: o autor (2024)

Sendo o gráfico de R^2 vs. T_d , exposto na Figura 46, tal que:

Figura 46 – R^2 vs. T_d para vários números de defeito considerados



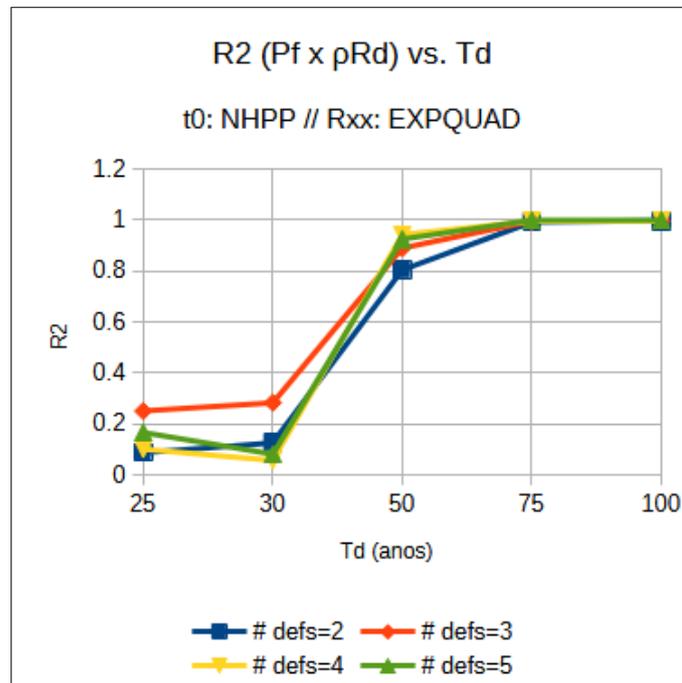
Fonte: o autor (2024)

De onde se nota o claro salto do coeficiente de determinação R^2 para maiores valores de T_d . Ou seja: quando da consideração de defeitos surgindo aleatoriamente no tempo, a tendência de convergência para um comportamento definido por uma função (mensurado em termos de R^2) se dá bem mais tarde do que quando os defeitos surgem todos a um mesmo tempo determinístico. A presença de aleatoriedade oriunda da consideração de um tempo para surgimento t_0 não-determinístico não mostra, portanto, grande influência dado um tempo T_d suficientemente grande.

4.2.4 Função de autocorrelação exponencial quadrática – tempo aleatório para surgimento dos defeitos e seguindo modelo NHPP

Para o caso em que a função de autocorrelação imposta a R_d é exponencial quadrática e t_0 segue o modelo NHPP, é observado um comportamento bastante similar ao anterior: altos valores de R^2 ocorrem apenas para altos valores de T_d , conforme Figura 47.

Figura 47 – R^2 vs. T_d para vários números de defeito considerados



Fonte: o autor (2024)

A tendência se mantém e os valores obtidos para Pf_{T_d} tendem, para todas as quantidades de defeito estudadas, a convergir a um comportamento previsível conforme se aumenta o tempo total de análise. A principal diferença entre o caso em que $t_0=2.88\text{anos}$ e aquele em que t_0 segue um modelo NHPP se dá, no exemplo estudado, apenas em qual o tempo necessário para que R^2 atinja valores altos.

Este fenômeno poderia ser mitigado (e os efeitos da aleatoriedade *em si* melhor avaliados) com o ajuste dos parâmetros λ_0 e b – que definem o surgimento dos defeitos pelo modelo NHPP, como mencionado na seção 3.3 – de forma que ao menos o primeiro defeito surgisse, em média, em $t_0=2.88$ anos. No sentido oposto, o tempo determinístico t_0 poderia ser alterado e tornado igual à média dos valores observados para o primeiro defeito que surge pelo modelo NHPP quando $\lambda_0=0.0064$ e $b=1$, que é de aproximadamente 28 anos. Desta forma, os resultados da consideração ora determinística, ora aleatória do tempo t_0 para surgimento dos defeitos poderiam ser melhor comparados.

De forma geral, nota-se que, em todos os gráficos de probabilidade de falha Pf_{T_d} versus o coeficiente de correlação ρ_{RD} entre dois defeitos adjacentes (Figuras 29 a 32; 38 a 41; e 44 e 45), a tendência de declínio de Pf_{T_d} com o aumento de ρ_{RD}

acelera – tanto para uma maior quantidade de defeitos dado um mesmo tempo T_d , quanto para a mesma quantidade de defeitos mas entre tempos T_d crescentes. Esta constatação corrobora as observações de Qian *et al.* (2013a, 2013b)

Ademais, o formato em si da curva segue aquele indicado por De Leon & Macías (2005) que obtiveram resultados semelhantes – mas onde esta relação foi dada como função do coeficiente de confiabilidade β em vez da probabilidade de falha diretamente. Curvas semelhantes podem ser encontradas no trabalho de Li *et al.* (2009).

5 CONCLUSÕES E SUGESTÕES DE TRABALHOS FUTUROS

Com base nos resultados obtidos, é possível concluir que:

- As aproximações obtidas para as correlações impostas às amostras foram satisfatórias. Em particular, as amostras correlacionadas pela função exponencial quadrática apresentaram menores valores médios para diferença relativa (entre a matriz obtida e a desejada) quando em comparação com as amostras correlacionadas pela função triangular. Isto indica que houve, para este caso, uma melhor aproximação no geral, apesar de os valores máximos serem similares;
- A tendência de aumento das diferenças relativas (máximas e médias) em relação à matriz de correlação alvo com o aumento do número de defeitos se deve, em grande parte, a dificuldades do algoritmo de otimização na busca pelo ponto ótimo;
- Há um impacto perceptível na probabilidade de falha total calculada para um tempo de vida T_d quando da consideração da correlação entre taxas lineares de crescimento da profundidade, Rd , no problema;
- Tomar as taxas Rd como independentes (i.e. correlação nula) leva a resultados conservadores, corroborando com a literatura a respeito do tema;
- A consideração de que os defeitos surgem aleatoriamente no tempo seguindo um modelo de Poisson não-homogêneo (NHPP com parâmetros $\lambda_0=0.0064$ e $b=1$) faz com que os efeitos da correlação ρ entre as taxas Rd apareçam somente bem mais tarde no ciclo de operação do duto. A tendência de “quanto maior o tempo total T_d considerado, maior a aderência da probabilidade de falha Pf_{Td} a um comportamento determinado por uma função (mensurado por R^2)”, no entanto, é mantida.

Como trabalhos futuros no tema sugere-se:

- Obtenção de resultados quando do ajuste dos parâmetros λ_0 e b (que definem o modelo NHPP) de forma fazer com que o primeiro defeito surja, em média, em 2.88 anos;

- Implementação do caso em que a função de estado-limite considerada seja *burst*, bem como o caso conjunto em que *small-leak*, *burst* e *rupture* atuem concorrendo entre si;
- Investigação dos efeitos da correlação espacial entre outras variáveis que não a taxa linear de crescimento da profundidade dos defeitos Rd , a exemplo da taxa RL (crescimento do comprimento do defeito);
- Observação do comportamento e comparação dos resultados da otimização numérica para obtenção de amostras correlacionadas quando do uso de outros algoritmos como o *Sequential Quadratic Programming* (SQP) e Simplex;
- Estudos a respeito da otimização de custos operacionais e de cronogramas de manutenção quando da consideração da correlação espacial entre variáveis do problema;
- Averiguação dos impactos da implementação de outros métodos para consideração da correlação entre variáveis aleatórias, como a teoria de cópula e/ou aproximação por limites de Ditlevsen para interseção entre dois eventos;
- Avaliação das diferenças observadas nos resultados quando do cálculo da probabilidade de falha do sistema por métodos diferentes do Método de Monte Carlo simples (*FORM*, *SORM*, *Importance Sampling*, *Selective Monte Carlo* etc..)

Referências

- AHMED, S. K.; KABIR, G. An Integrated Approach for Failure Analysis of Natural Gas Transmission Pipeline. **CivilEng**, v. 2, n. 1, p. 87–119, 1 fev. 2021.
- AMAYA-GÓMEZ, R. et al. Reliability assessments of corroded pipelines based on internal pressure – A review. **Engineering Failure Analysis**, v. 98, p. 190–214, abr. 2019.
- AMAYA-GÓMEZ, R. et al. **Corrosion and Reliability Assessment of Inspected Pipelines**. Cham: Springer International Publishing, 2024.
- ANG, A. H.-S.; TANG, W. H. **Probability concepts in engineering: emphasis on applications in civil & environmental engineering**. 2nd ed ed. Hoboken: J. Wiley & sons, 2007.
- ARORA, J. S. **Introduction to optimum design**. 4th edition ed. Amsterdam Boston: Elsevier/Academic Press, 2017.
- ASME. **Manual for Determining the Remaining Strength of Corroded Pipelines - Supplement to ASME B31 Code for Pressure Piping**. New York, NY, 2012.
- BAZÁN, F. A. V.; BECK, A. T. Stochastic process corrosion growth models for pipeline reliability. **Corrosion Science**, v. 74, p. 50–58, set. 2013.
- BECK, A. T. **Um método para análise de confiabilidade para elementos finitos**. Dissertação (Mestrado em Engenharia Mecânica)—Florianópolis: Universidade Federal de Santa Catarina, 1999.
- BECK, A. T. **Confiabilidade estrutural**. [s.l.] Elsevier, 2019.
- BEDAIRI, B. et al. Failure prediction for Crack-in-Corrosion defects in natural gas transmission pipelines. **International Journal of Pressure Vessels and Piping**, v. 96–97, p. 90–99, ago. 2012.
- BEICHELT, F. et al. **Stochastic processes and their applications**. London: Taylor & Francis, 2002.
- BERGER, M. A. **An Introduction to Probability and Stochastic Processes**. New York, NY: Springer New York, 1993.
- BIEZMA, M. V. et al. Most fatal oil & gas pipeline accidents through history: A lessons learned approach. **Engineering Failure Analysis**, v. 110, p. 104446, mar. 2020.
- BRITISH PETROLEUM. **Statistical Review of World Energy 2020**. [s.l: s.n.].
- CABRAL, M. A. F. S. **Uma metodologia para avaliação de integridade de dutos corroídos baseada em banco de dados**. Dissertação (Mestrado em Engenharia Civil)—Recife, PE: Universidade Federal de Pernambuco, 2023.

CALEYO, F. et al. Probability distribution of pitting corrosion depth and rate in underground pipelines: A Monte Carlo study. **Corrosion Science**, v. 51, n. 9, p. 1925–1934, set. 2009.

CALEYO, F.; GONZÁLEZ, J. L.; HALLEN, J. M. A study on the reliability assessment methodology for pipelines with active corrosion defects. **International Journal of Pressure Vessels and Piping**, v. 79, n. 1, p. 77–86, jan. 2002.

CHAKRABORTY, S.; TESHAMARIAM, S. Subset simulation based approach for space-time-dependent system reliability analysis of corroding pipelines. **Structural Safety**, v. 90, p. 102073, maio 2021.

DAHIRE, S. et al. Bayesian Network inference for probabilistic strength estimation of aging pipeline systems. **International Journal of Pressure Vessels and Piping**, v. 162, p. 30–39, maio 2018.

DAI, L. et al. Analysis and Comparison of Long-Distance Pipeline Failures. **Journal of Petroleum Engineering**, v. 2017, p. 1–7, 25 maio 2017.

DANTAS, C. A. B. **Probabilidade: um curso introdutório**. 3. ed. rev., 1. reimpr ed. São Paulo: Edusp, 2013.

DE LEON, D.; MACÍAS, O. F. Effect of spatial correlation on the failure probability of pipelines under corrosion. **International Journal of Pressure Vessels and Piping**, v. 82, n. 2, p. 123–128, fev. 2005.

DEVORE, J. L. **Probability and statistics for engineering and the sciences**. Ninth edition ed. Boston, MA: Cengage Learning, 2016.

DITLEVSEN, O. Principle of Normal Tail Approximation. **Journal of the Engineering Mechanics Division**, v. 107, n. 6, p. 1191–1208, dez. 1981.

DNV GL AS. **Recommended Practice DNVGL-RP-F101 - Corroded Pipelines**. , 2017.

DOBROW, R. P. **Introduction to stochastic processes with R**. Hoboken, New Jersey: Wiley, 2016.

ELISHAKOFF, I. **Probabilistic methods in the theory of structures: strength of materials, random vibrations, and random buckling**. Third edition ed. New Jersey London Singapore Beijing Shanghai Hong Kong Taipei Chennai Tokyo: World Scientific, 2017.

FU, G. et al. Reliability analysis of corrosion affected underground steel pipes considering multiple failure modes and their stochastic correlations. **Tunnelling and Underground Space Technology**, v. 87, p. 56–63, maio 2019.

GOMES, W. J. S.; BECK, A. T. Optimal inspection and design of onshore pipelines under external corrosion process. **Structural Safety**, v. 47, p. 48–58, mar. 2014.

GOMES, W. J. S.; BECK, A. T.; HAUKAAS, T. Optimal inspection planning for onshore pipelines subject to external corrosion. **Reliability Engineering & System Safety**, v. 118, p. 18–27, out. 2013.

GONG, C.; ZHOU, W. First-order reliability method-based system reliability analyses of corroding pipelines considering multiple defects and failure modes. **Structure and Infrastructure Engineering**, v. 13, n. 11, p. 1451–1461, 2 nov. 2017.

GONG, C.; ZHOU, W. Importance sampling-based system reliability analysis of corroding pipelines considering multiple failure modes. **Reliability Engineering & System Safety**, v. 169, p. 199–208, jan. 2018.

GUO, B. et al. **Offshore pipelines: design, installation, and maintenance**. 2nd edition ed. Amsterdam: Elsevier, GPP, 2014.

HASOFER, A. M.; LIND, N. C. Exact and Invariant Second-Moment Code Format. **Journal of the Engineering Mechanics Division**, v. 100, n. 1, p. 111–121, fev. 1974.

HASWELL, J. V.; MCCONNELL, R. A. **UKOPA Pipeline Product Loss Incidents and Faults Report (1962-2014)**. [s.l.] United Kingdom Onshore Pipeline Operator's Association, dez. 2015.

HOEL, P. G. et al. **Introduction to stochastic processes**. Boston: Houghton Mifflin, 1972.

HOPKINS, P. Assessing the significance of corrosion in onshore oil and gas pipelines. Em: **Underground Pipeline Corrosion**. [s.l.] Elsevier, 2014. p. 62–84.

HSU, H. P. **Schaum's outlines: probability, random variables, and random processes**. Fourth edition ed. New York: McGraw-Hill Education, 2020.

HUSSAIN, M. et al. Energy pipeline degradation condition assessment using predictive analytics – challenges, issues, and future directions. **Journal of Pipeline Science and Engineering**, v. 4, n. 3, p. 100178, set. 2024.

IBE, O. C. **Fundamentals of applied probability and random processes**. 2nd edition ed. Amsterdam ; Boston: Elsevier/AP, 2014.

JOHNSON, R. A.; MILLER, I.; FREUND, J. E. **Miller & Freund's probability and statistics for engineers**. Ninth edition ed. Boston: Pearson, 2017.

KAPUR, K. C. **Reliability engineering**. Hoboken, N.J: J. Wiley & Sons, 2014.

KHAN, F.; YARVEISY, R.; ABBASSI, R. Risk-based pipeline integrity management: A road map for the resilient pipelines. **Journal of Pipeline Science and Engineering**, v. 1, n. 1, p. 74–87, mar. 2021.

KULKARNI, V. G. **Modeling and analysis of stochastic systems**. Third edition ed. Boca Raton, FL: CRC Press, Taylor & Francis Group, 2017.

KUNIEWSKI, S. P.; VAN DER WEIDE, J. A. M.; VAN NOORTWIJK, J. M. Sampling inspection for the evaluation of time-dependent reliability of deteriorating systems under imperfect defect detection. **Reliability Engineering & System Safety**, v. 94, n. 9, p. 1480–1490, set. 2009.

LANCASTER, P.; ŠALKKAUSKAS, K. **Curve and surface fitting: an introduction**. London ; Orlando: Academic Press, 1986.

LI, S.-X. et al. A method of probabilistic analysis for steel pipeline with correlated corrosion defects. **Corrosion Science**, v. 51, n. 12, p. 3050–3056, dez. 2009.

LI, X. et al. Optimal maintenance strategy for corroded subsea pipelines. **Journal of Loss Prevention in the Process Industries**, v. 49, p. 145–154, set. 2017.

LIU, X. et al. Multi-level optimization of maintenance plan for natural gas pipeline systems subject to external corrosion. **Journal of Natural Gas Science and Engineering**, v. 50, p. 64–73, fev. 2018.

LOPEZ, R. H.; BECK, A. T. Reliability-based design optimization strategies based on FORM: a review. **Journal of the Brazilian Society of Mechanical Sciences and Engineering**, v. 34, n. 4, p. 506–514, dez. 2012.

LUTES, L. D.; SARKANI, S. (EDS.). **Random vibrations: analysis of structural and mechanical systems**. Amsterdam Boston: Elsevier, 2004.

MAGALHÃES, M. N. **Probabilidade e Variáveis Aleatórias**. São Paulo (SP): Edusp, 2011.

MAGALHÃES, M. N.; LIMA, A. C. P. DE. **Noções de Probabilidade e Estatística**. 7ª Edição ed. São Paulo: Edusp, 2015.

MAHMOODIAN, M. **Reliability and maintainability of in-service pipelines**. Cambridge, MA: Gulf Professional Publishing, an imprint of Elsevier, 2018.

MEDEIROS, C. P.; ALENCAR, M. H.; DE ALMEIDA, A. T. Multidimensional risk evaluation of natural gas pipelines based on a multicriteria decision model using visualization tools and statistical tests for global sensitivity analysis. **Reliability Engineering & System Safety**, v. 165, p. 268–276, set. 2017.

MELCHERS, R. E.; BECK, A. T. **Structural reliability analysis and prediction**. Third edition ed. Hoboken, NJ: Wiley, 2018.

MISHRA, M.; KESHAVARZZADEH, V.; NOSHADRAVAN, A. Reliability-based lifecycle management for corroding pipelines. **Structural Safety**, v. 76, p. 1–14, jan. 2019.

MONTGOMERY, D. C.; RUNGER, G. C. **Applied statistics and probability for engineers**. Seventh edition ed. Hoboken, NJ: Wiley, 2018.

- MOTTA, R. D. S.; FERREIRA, A. D. M.; AFONSO, S. M. B. Probabilistic assessment of complex corrosion in pipelines considering River-Bottom Profile information. **Engineering Failure Analysis**, v. 165, p. 108801, nov. 2024.
- MUHLBAUER, W. K. **Pipeline risk management manual: ideas, techniques, and resources**. 3rd ed ed. Burlington (Mass.): Elsevier / Gulf Professional Publishing, 2004.
- MURATA, T. INTRODUCTORY NOTES ON CLIMATE CHANGE, LIFE-CYCLE DESIGN, AND CORROSION OF STEEL. Em: **Uhlig's Corrosion Handbook**. 3. ed. [s.l.] Electrochemical Society, 2011.
- NESSIM, M. Estimating the Risk of Pipeline Failure Due to Corrosion. Em: REVIE, R. W. (Ed.). **Uhlig's Corrosion Handbook**. 1. ed. [s.l.] Wiley, 2011. p. 75–82.
- NORSWORTHY, R. Understanding corrosion in underground pipelines: basic principles. Em: **Underground Pipeline Corrosion**. [s.l.] Elsevier, 2014. p. 3–34.
- PAIK, J. K.; MELCHERS, R. E. (EDS.). **Condition assessment of aged structures**. Boca Raton, Fla.: CRC Press, 2009.
- PALIGA, C. M. **Análise probabilística de vigas de concreto armado recuperadas à flexão, através do método de Monte Carlo utilizando um modelo de elementos finitos**. Tese (Doutorado em Engenharia Civil)—Porto Alegre, RS: Universidade Federal do Rio Grande do Sul, 2008.
- PAPOULIS, A.; PILLAI, S. U. **Probability, random variables, and stochastic processes**. Fourth edition ed. Boston New Delhi: McGraw-Hill, 2002.
- PARZEN, E. **Stochastic processes: An unabridged, corrected republication of the work first published by Holden-Day, Oakland, Calif., 1962**. Philadelphia, Pa: Society for Industrial and Applied Mathematics (SIAM, 3600 Market Street, Floor 6, Philadelphia, PA 19104), 1999.
- PESINIS, K.; TEE, K. F. Statistical model and structural reliability analysis for onshore gas transmission pipelines. **Engineering Failure Analysis**, v. 82, p. 1–15, dez. 2017.
- PISHRO-NIK, H. **Introduction to probability, statistics, and random processes**. Blue Bell, PA: Kappa Research, LLC, 2014.
- PRCI. **PRCI RSTRENG**. PRCI/Technical Toolboxes, , 2020.
- QIAN, G. et al. Probabilistic leak-before-break analysis with correlated input parameters. **Nuclear Engineering and Design**, v. 254, p. 266–271, jan. 2013a.
- QIAN, G. et al. Effect of correlated input parameters on the failure probability of pipelines with corrosion defects by using FITNET FFS procedure. **International Journal of Pressure Vessels and Piping**, v. 105–106, p. 19–27, maio 2013b.
- ROUSSAS, G. G. **Introduction to probability**. 2nd edition ed. Amsterdam Boston: Elsevier, 2014.

RUBINSTEIN, R. Y.; KROESE, D. P. **Simulation and the Monte Carlo method**. Third edition ed. Hoboken, New Jersey: Wiley, 2017.

SAHRAOUI, Y.; CHATEAUNEUF, A. The effects of spatial variability of the aggressiveness of soil on system reliability of corroding underground pipelines. **International Journal of Pressure Vessels and Piping**, v. 146, p. 188–197, out. 2016.

SOONG, T. T. **Fundamentals of probability and statistics for engineers**. Hoboken, NJ: John Wiley & Sons, 2004.

SOUSA, P. F. S. **PLANEJAMENTO DE MANUTENÇÕES EM DUTOS CORROÍDOS USANDO METAMODELOS**. Tese (Doutorado em Engenharia Civil)—Recife: Universidade Federal de Pernambuco, 2022.

SOUSA, P. F. S.; AFONSO, S. M. B.; WILLMERSDORF, R. B. Reliability-based preventive maintenance planning for corroded pipelines using a RBF surrogate model. **Journal of the Brazilian Society of Mechanical Sciences and Engineering**, v. 43, n. 12, p. 521, dez. 2021.

STEWART, J.; CLEGG, D.; WATSON, S. **Calculus: early transcendentals**. Ninth edition ed. Boston, MA: Cengage, 2021.

SU, H. et al. (EDS.). **Advanced intelligent pipeline management technology**. Singapore: Springer, 2023.

TAYLOR, H. M.; KARLIN, S. **An introduction to stochastic modeling**. Rev. ed ed. Boston: Academic Press, 1994.

TEE, K. F.; PESINIS, K. Reliability prediction for corroding natural gas pipelines. **Tunnelling and Underground Space Technology**, v. 65, p. 91–105, maio 2017.

THE MATHWORKS, INC. **MATLAB**. The Mathworks, INC., , 2020.

TIMASHEV, S.; BUSHINSKAYA, A. **Diagnostics and Reliability of Pipeline Systems**. Cham: Springer International Publishing, 2016. v. 30

VALOR, A. et al. Reliability assessment of buried pipelines based on different corrosion rate models. **Corrosion Science**, v. 66, p. 78–87, jan. 2013.

VELÁZQUEZ, J. C. et al. Predictive Model for Pitting Corrosion in Buried Oil and Gas Pipelines. **CORROSION**, v. 65, n. 5, p. 332–342, maio 2009.

WANG, Y.; XIE, M.; SU, C. Multi-objective maintenance strategy for corroded pipelines considering the correlation of different failure modes. **Reliability Engineering & System Safety**, v. 243, p. 109894, mar. 2024.

WATKINS, D. S. **Fundamentals of matrix computations**. 3. ed ed. Hoboken, NJ: Wiley, 2010.

WOLDESELLASSE, H.; TESFAMARIAM, S. Failure assessment of oil and gas transmission pipelines using an integrated Bayesian belief network and GIS model.

International Journal of Pressure Vessels and Piping, v. 205, p. 104984, out. 2023a.

WOLDESELLASSE, H.; TEFAMARIAM, S. Risk analysis of onshore oil and gas pipelines: Literature review and bibliometric analysis. **Journal of Infrastructure Intelligence and Resilience**, v. 2, n. 4, p. 100052, nov. 2023b.

XIAO, R. et al. Time varying reliability analysis of corroded gas pipelines using copula and importance sampling. **Ocean Engineering**, v. 306, p. 118086, ago. 2024a.

XIAO, R. et al. Rapid failure risk analysis of corroded gas pipelines using machine learning. **Ocean Engineering**, v. 313, p. 119433, dez. 2024b.

YATES, R. D.; GOODMAN, D. J. **Probability and stochastic processes: a friendly introduction for electrical and computer engineers**. 2nd. ed. Hoboken, NJ: John Wiley & Sons, 2005.

ZELMATI, D. et al. A probabilistic approach to estimate the remaining life and reliability of corroded pipelines. **Journal of Natural Gas Science and Engineering**, v. 99, p. 104387, mar. 2022.

ZHANG, P. et al. Failure probability of corroded pipeline considering the correlation of random variables. **Engineering Failure Analysis**, v. 99, p. 34–45, maio 2019.

ZHANG, S.; ZHOU, W. System reliability of corroding pipelines considering stochastic process-based models for defect growth and internal pressure. **International Journal of Pressure Vessels and Piping**, v. 111–112, p. 120–130, nov. 2013.

ZHANG, S.; ZHOU, W. Cost-based optimal maintenance decisions for corroding natural gas pipelines based on stochastic degradation models. **Engineering Structures**, v. 74, p. 74–85, set. 2014.

ZHOU, W. System reliability of corroding pipelines. **International Journal of Pressure Vessels and Piping**, v. 87, n. 10, p. 587–595, out. 2010.

ZHOU, W.; HONG, H. P.; ZHANG, S. Impact of dependent stochastic defect growth on system reliability of corroding pipelines. **International Journal of Pressure Vessels and Piping**, v. 96–97, p. 68–77, ago. 2012.

APÊNDICE A: TABELAS COM FUNÇÕES DE INTERPOLAÇÃO

- $t_0 = 2.88$ anos e autocorrelação = LINEAR

– $t_D = 15$ anos

N_defs	Eq_PF_vs_rhoRd	R2
2	$Y = (-0.00041345 \cdot x^3) + (5.5675e-05 \cdot x^2) + (-9.7643e-05 \cdot x) + (0.0065139)$	0.79367
3	$Y = (-0.0015241 \cdot x^3) + (-0.00017049 \cdot x^2) + (0.00028116 \cdot x) + (0.0097821)$	0.93000
4	$Y = (-0.005615 \cdot x^3) + (0.0042261 \cdot x^2) + (-0.0010097 \cdot x) + (0.012954)$	0.97549
5	$Y = (-0.0085038 \cdot x^3) + (0.006706 \cdot x^2) + (-0.0017477 \cdot x) + (0.01622)$	0.99308

– $t_D = 20$ anos

N_defs	Eq_PF_vs_rhoRd	R2
2	$Y = -0.004027 \cdot \exp(1.6491 \cdot x) + (0.077022 \cdot \exp(0.068664 \cdot x))$	0.99251
3	$Y = -0.0028263 \cdot \exp(2.5321 \cdot x) + (0.10985 \cdot \exp(0.046372 \cdot x))$	0.99799
4	$Y = -0.0020237 \cdot \exp(3.2332 \cdot x) + (0.14076 \cdot \exp(0.027482 \cdot x))$	0.99843
5	$Y = (-0.091413 \cdot x^3) + (0.048144 \cdot x^2) + (-0.016084 \cdot x) + (0.16925)$	0.99937

– $t_D = 25$ anos

N_defs	Eq_PF_vs_rhoRd	R2
2	$Y = -0.015194 \cdot \exp(1.6731 \cdot x) + (0.3061 \cdot \exp(0.034743 \cdot x))$	0.99840
3	$Y = -0.013222 \cdot \exp(2.3178 \cdot x) + (0.40953 \cdot \exp(0.024318 \cdot x))$	0.99972
4	$Y = (-0.18425 \cdot x^3) + (0.075163 \cdot x^2) + (-0.047535 \cdot x) + (0.48138)$	0.99954
5	$Y = (-0.27435 \cdot x^3) + (0.1416 \cdot x^2) + (-0.059482 \cdot x) + (0.55094)$	0.99966

– $t_D = 30$ anos

N_defs	Eq_PF_vs_rhoRd	R2
2	$Y = -0.011693 \cdot \exp(2.1442 \cdot x) + (0.62415 \cdot \exp(-0.022623 \cdot x))$	0.99962
3	$Y = -0.013889 \cdot \exp(2.5491 \cdot x) + (0.75566 \cdot \exp(0.00029476 \cdot x))$	0.99991
4	$Y = (-0.26138 \cdot x^3) + (0.13872 \cdot x^2) + (-0.069695 \cdot x) + (0.82182)$	0.99966
5	$Y = (-0.35591 \cdot x^3) + (0.23735 \cdot x^2) + (-0.083418 \cdot x) + (0.87279)$	0.99958

- $t_0 = 2.88$ anos e autocorrelação = EXPONENCIAL QUADRÁTICA

– $t_D = 15$ anos

N_defs	Eq_PF_vs_rhoRd	R2
2	$Y = (-1.1809 * x^3) + (3.435 * x^2) + (-3.3307 * x) + (1.0826)$	0.22805
3	$Y = (7.7264 * x^3) + (-22.4944 * x^2) + (21.8173 * x) + (-7.0409)$	0.90721
4	$Y = (8.8312 * x^3) + (-25.8128 * x^2) + (25.1277 * x) + (-8.1357)$	0.95569
5	$Y = (-5.4195 * x^3) + (15.4628 * x^2) + (-14.7234 * x) + (4.6923)$	0.98715

– $t_D = 20$ anos

N_defs	Eq_PF_vs_rhoRd	R2
2	$Y = (7.4649 * x^3) + (-21.7141 * x^2) + (21.0179 * x) + (-6.7075)$	0.89433
3	$Y = (3.1903 * x^3) + (-9.4872 * x^2) + (9.2907 * x) + (-2.9146)$	0.99249
4	$Y = (14.2939 * x^3) + (-42.9553 * x^2) + (42.7695 * x) + (-14.0137)$	0.99675
5	$Y = (-33.7853 * x^3) + (95.1055 * x^2) + (-89.4521 * x) + (28.2389)$	0.99863

– $t_D = 25$ anos

N_defs	Eq_PF_vs_rhoRd	R2
2	$Y = (23.1874 * x^3) + (-67.9001 * x^2) + (66.1279 * x) + (-21.1801)$	0.98318
3	$Y = (-20.9851 * x^3) + (60.026 * x^2) + (-57.585 * x) + (18.8283)$	0.99560
4	$Y = (-38.0975 * x^3) + (105.9094 * x^2) + (-98.5841 * x) + (31.0933)$	0.99761
5	$Y = (-124.3928 * x^3) + (351.9171 * x^2) + (-332.554 * x) + (105.3802)$	0.99958

– $t_D = 30$ anos

N_defs	Eq_PF_vs_rhoRd	R2
2	$Y = (20.9984 * x^3) + (-62.0818 * x^2) + (60.9129 * x) + (-19.3207)$	0.99566
3	$Y = (-30.7555 * x^3) + (87.4106 * x^2) + (-83.3349 * x) + (27.255)$	0.99724
4	$Y = (-116.6361 * x^3) + (331.9465 * x^2) + (-315.6095 * x) + (100.9186)$	0.99797
5	$Y = (-205.5531 * x^3) + (584.0419 * x^2) + (-553.9864 * x) + (176.1501)$	0.99959

- $t_0 = NHPP$ e autocorrelação = LINEAR

– $t_D = 25$ anos

N_defs	Eq_PF_vs_rhoRd	R2
2	$Y = (0.0002429x^3) + (-6.018e-05x^2) + (-9.5267e-05x) + (0.0020757)$	0.26497
3	$Y = (-0.00036728x^3) + (0.000444x^2) + (-8.2027e-05x) + (0.0020548)$	0.11245
4	$Y = (0.0008562x^3) + (-0.0012216x^2) + (0.00043255x) + (0.0020792)$	0.13386
5	$Y = (-0.00054974x^3) + (0.00078047x^2) + (-0.00034907x) + (0.002122)$	0.12741

– $t_D = 30$ anos

N_defs	Eq_PF_vs_rhoRd	R2
2	$Y = (4.9375e-05x^3) + (0.00070671x^2) + (-0.00061487x) + (0.01072)$	0.26352
3	$Y = (-0.0032499x^3) + (0.0044939x^2) + (-0.0016311x) + (0.0108)$	0.19941
4	$Y = (0.0023541x^3) + (-0.0037068x^2) + (0.0012775x) + (0.010839)$	0.59427
5	$Y = (-0.00033053x^3) + (0.00024441x^2) + (-0.00020011x) + (0.010796)$	0.27533

– $t_D = 50$ anos

N_defs	Eq_PF_vs_rhoRd	R2
2	$Y = (0.0019896x^3) + (-0.022072x^2) + (-0.014589x) + (0.45524)$	0.99851
3	$Y = (-0.014316x^3) + (-0.0084682x^2) + (-0.021746x) + (0.46522)$	0.99827
4	$Y = (-0.014307x^3) + (-0.014295x^2) + (-0.017723x) + (0.46683)$	0.99634
5	$Y = (-0.00035412x^3) + (-0.037264x^2) + (-0.0081822x) + (0.46647)$	0.99818

– $t_D = 75$ anos

N_defs	Eq_PF_vs_rhoRd	R2
2	$Y = -0.00015544 \exp(3.0219x) + (0.99818 \exp(-0.0014005x))$	0.99982
3	$Y = (-0.0057526x^3) + (0.0026961x^2) + (-0.0016843x) + (0.99888)$	0.99980
4	$Y = (-0.0081889x^3) + (0.0052465x^2) + (-0.0019613x) + (0.99913)$	0.99958
5	$Y = (-0.0094018x^3) + (0.0066263x^2) + (-0.0021332x) + (0.99922)$	0.99944

– $t_D = 100$ anos

N_defs	Eq_PF_vs_rhoRd	R2
2	$Y = (-8.2389e-07x^3) + (6.3478e-07x^2) + (-2.0997e-07x) + (1)$	0.99661
3	$Y = (-8.6969e-07x^3) + (8.1419e-07x^2) + (-2.0811e-07x) + (1)$	0.98692
4	$Y = -9.8504e-12 \exp(10.4487x) + (1 \exp(-1.0795e-08x))$	0.99958
5	$Y = (-7.0952e-07x^3) + (7.4396e-07x^2) + (-1.9547e-07x) + (1)$	0.94980

- $t_0 = NHPP$ e autocorrelação = EXPONENCIAL QUADRÁTICA

– $t_D = 25$ anos

N_defs	Eq_PF_vs_rhoRd	R2
2	$Y = (-1.3849x^3) + (4.0476x^2) + (-3.9413x) + (1.2808)$	0.08648
3	$Y = (0.3659x^3) + (-0.99732x^2) + (0.9011x) + (-0.26758)$	0.24974
4	$Y = (-0.78787x^3) + (2.3383x^2) + (-2.3122x) + (0.76383)$	0.09993
5	$Y = (-0.46456x^3) + (1.3989x^2) + (-1.4022x) + (0.46993)$	0.16551

– $t_D = 30$ anos

N_defs	Eq_PF_vs_rhoRd	R2
2	$Y = (-7.3148x^3) + (21.2825x^2) + (-20.6334x) + (6.6763)$	0.12570
3	$Y = (5.341x^3) + (-15.3756x^2) + (14.7457x) + (-4.7003)$	0.28261
4	$Y = (0.18546x^3) + (-0.48386x^2) + (0.41392x) + (-0.10487)$	0.05674
5	$Y = (3.6327x^3) + (-10.5324x^2) + (10.176x) + (-3.2657)$	0.07998

– $t_D = 50$ anos

N_defs	Eq_PF_vs_rhoRd	R2
2	$Y = (35.3984x^3) + (-102.6114x^2) + (99.0521x) + (-31.4192)$	0.80425
3	$Y = (24.0685x^3) + (-69.129x^2) + (66.0592x) + (-20.5771)$	0.88972
4	$Y = (88.1616x^3) + (-256.105x^2) + (247.8279x) + (-79.4624)$	0.94350
5	$Y = (30.9996x^3) + (-90.3923x^2) + (87.7367x) + (-27.9224)$	0.92529

– $t_D = 75$ anos

N_defs	Eq_PF_vs_rhoRd	R2
2	$Y = (1.4169x^3) + (-4.1775x^2) + (4.0931x) + (-0.33885)$	0.99204
3	$Y = (-0.4421x^3) + (1.2577x^2) + (-1.2105x) + (1.3888)$	0.99427
4	$Y = (0.22051x^3) + (-0.75202x^2) + (0.81271x) + (0.71285)$	0.99605
5	$Y = (0.14797x^3) + (-0.60335x^2) + (0.72455x) + (0.72489)$	0.99884

– $t_D = 100$ anos

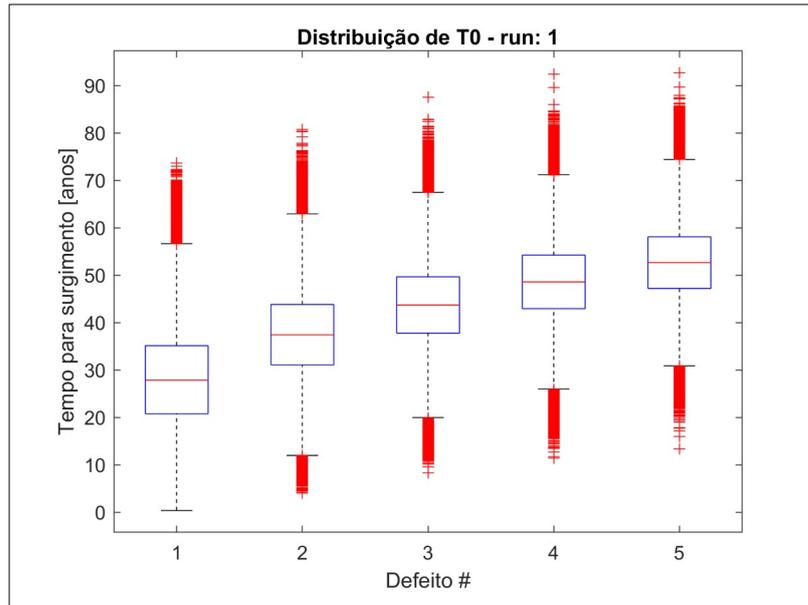
N_defs	Eq_PF_vs_rhoRd	R2
2	$Y = (0.00015112x^3) + (-0.00045778x^2) + (0.00045919x) + (0.99985)$	0.99725
3	$Y = (-0.00022214x^3) + (0.00062374x^2) + (-0.00058613x) + (1.0002)$	0.99750
4	$Y = (-0.00056496x^3) + (0.001596x^2) + (-0.0015049x) + (1.0005)$	0.99819
5	$Y = (-0.00088857x^3) + (0.0025164x^2) + (-0.0023774x) + (1.0007)$	0.99959

APÊNDICE B: GRÁFICOS COM DETALHES DAS REALIZAÇÕES DOS TEMPOS DE SURGIMENTO DOS DEFEITOS POR NHPP

- Run #1:

Boxplot:

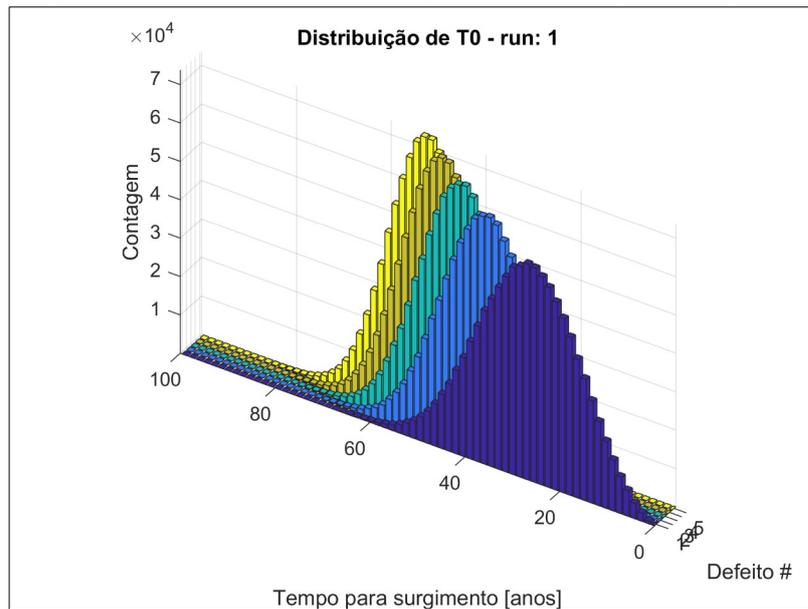
Figura 48 – Representação por boxplot dos tempos de surgimento dos defeitos por NHPP para a 1ª execução do algoritmo



Fonte: o autor (2024)

Histogramas:

Figura 49 – Representação por histograma dos tempos de surgimento dos defeitos por NHPP para a 1ª execução do algoritmo

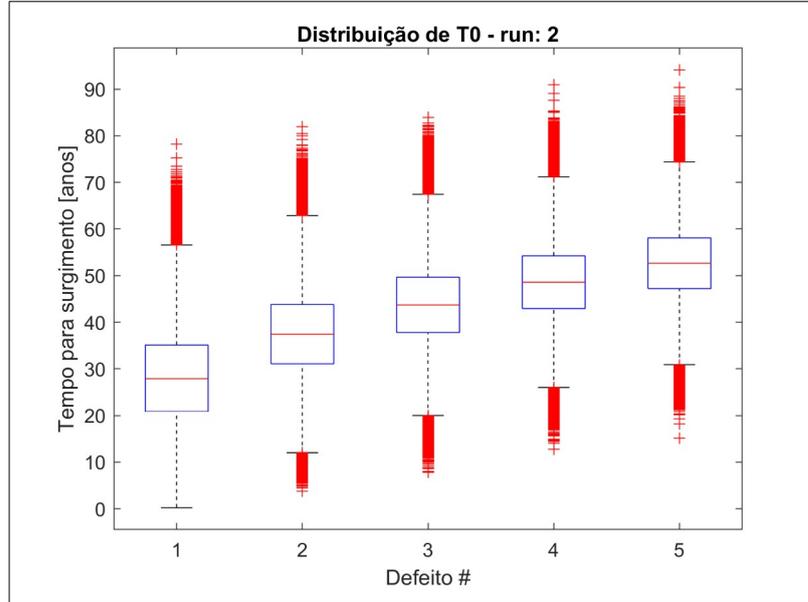


Fonte: o autor (2024)

● Run #2:

Boxplot:

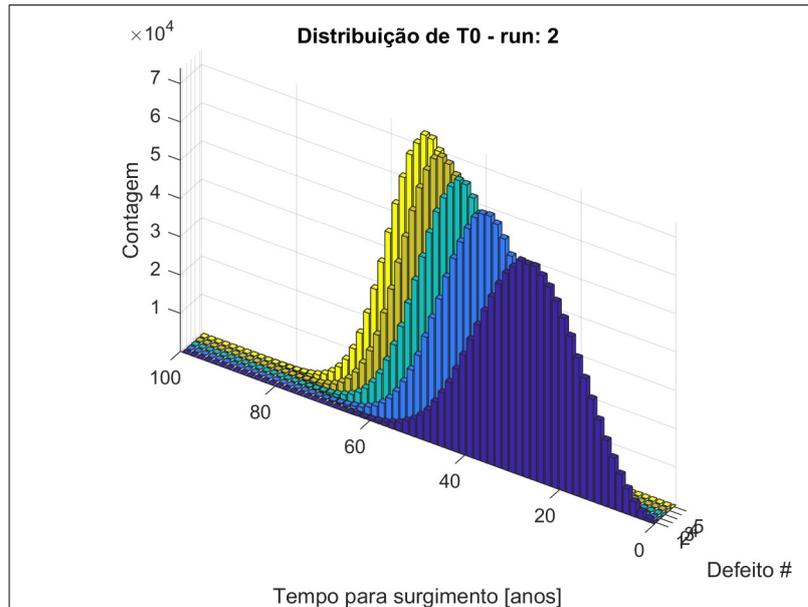
Figura 50 – Representação por boxplot dos tempos de surgimento dos defeitos por NHPP para a 2ª execução do algoritmo



Fonte: o autor (2024)

Histogramas:

Figura 51 – Representação por histograma dos tempos de surgimento dos defeitos por NHPP para a 2ª execução do algoritmo

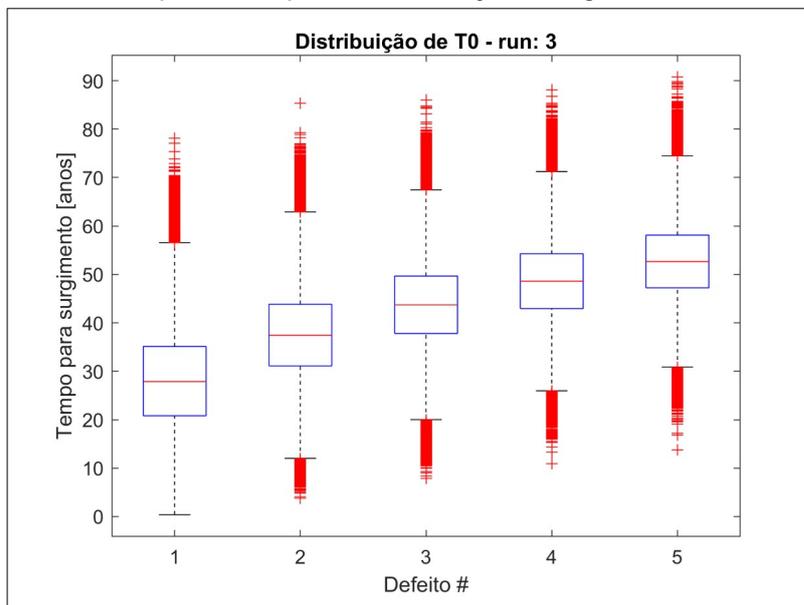


Fonte: o autor (2024)

- Run #3:

Boxplot:

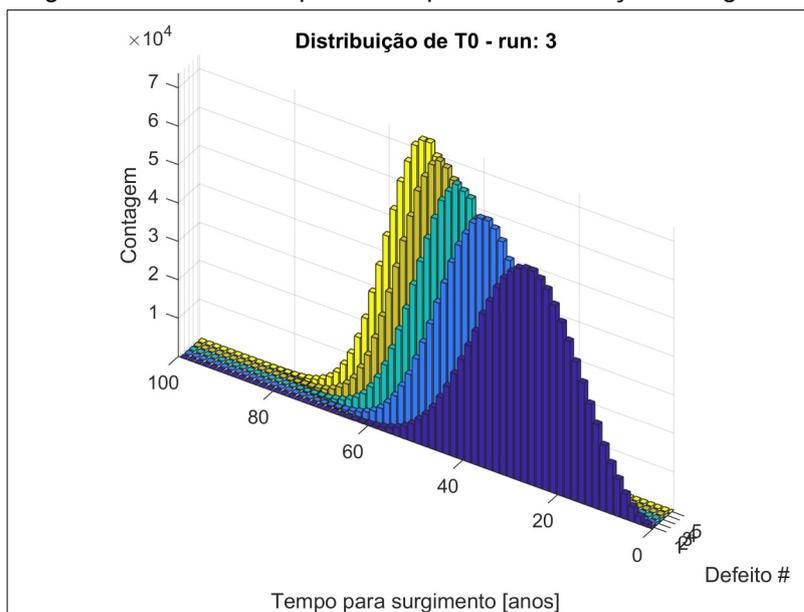
Figura 52 – Representação por boxplot dos tempos de surgimento dos defeitos por NHPP para a 3ª execução do algoritmo



Fonte: o autor (2024)

Histogramas:

Figura 53 – Representação por histograma dos tempos de surgimento dos defeitos por NHPP para a 3ª execução do algoritmo

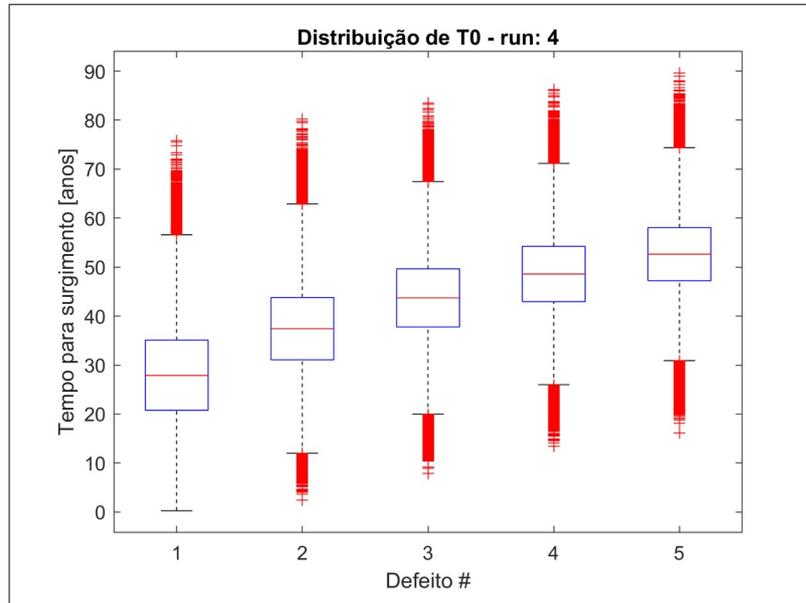


Fonte: o autor (2024)

- Run #4:

Boxplot:

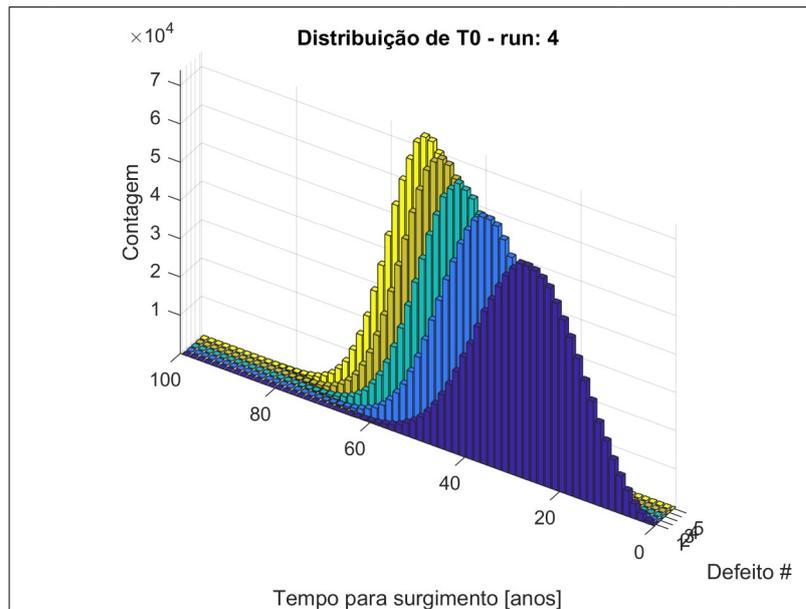
Figura 54 – Representação por boxplot dos tempos de surgimento dos defeitos por NHPP para a 4ª execução do algoritmo



Fonte: o autor (2024)

Histogramas:

Figura 55 – Representação por histograma dos tempos de surgimento dos defeitos por NHPP para a 4ª execução do algoritmo

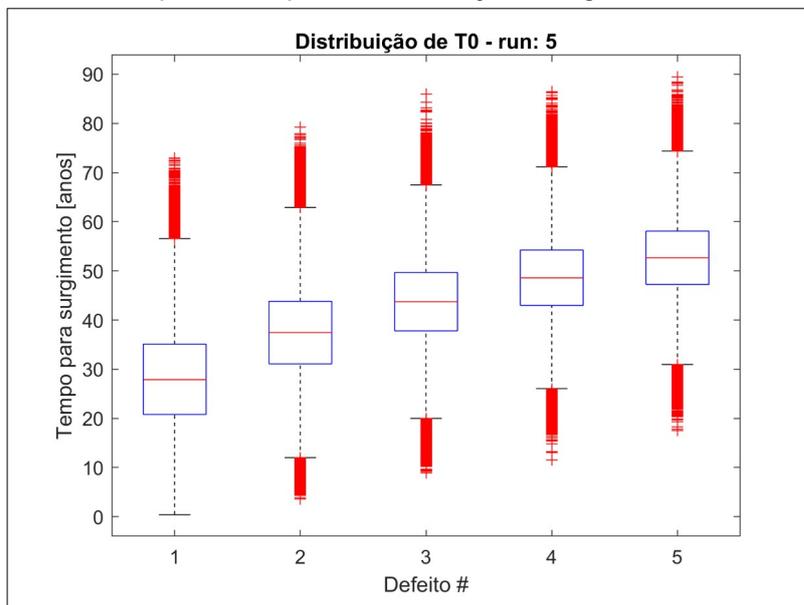


Fonte: o autor (2024)

● Run #5:

Boxplot:

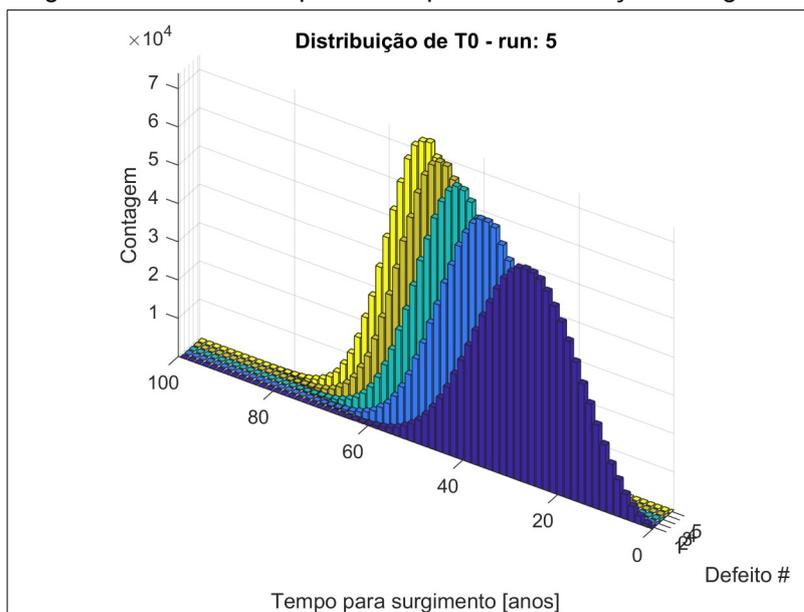
Figura 56 – Representação por boxplot dos tempos de surgimento dos defeitos por NHPP para a 5ª execução do algoritmo



Fonte: o autor (2024)

Histogramas:

Figura 57 – Representação por histograma dos tempos de surgimento dos defeitos por NHPP para a 5ª execução do algoritmo



Fonte: o autor (2024)