

UNIVERSIDADE FEDERAL DE ITAJUBÁ
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

ANÁLISE PARAMÉTRICA USANDO SIMULAÇÕES BASEADAS EM
EMTP: PRINCÍPIOS E POTENCIAIS APLICAÇÕES

LUIS GUILHERME RODRIGUES DE PAULA

Itajubá, junho de 2025

UNIVERSIDADE FEDERAL DE ITAJUBÁ
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

LUIS GUILHERME RODRIGUES DE PAULA

ANÁLISE PARAMÉTRICA USANDO SIMULAÇÕES BASEADAS EM
EMTP: PRINCÍPIOS E POTENCIAIS APLICAÇÕES

Dissertação submetida ao Programa de Pós-Graduação em Engenharia Elétrica como parte dos requisitos para obtenção do título de Mestre em Ciências em Engenharia Elétrica.

Orientador: Prof. Dr. Benedito Donizeti Bonatto

Itajubá, junho de 2025

Agradecimentos

Agradeço aos meus pais, esposa, irmãos, avós e amigos por estarem sempre me apoiando de todas as formas possíveis, de forma que é impossível recompensá-los.

Agradeço a Deus, por tudo.

Agradeço à comunidade acadêmica da UNIFEI. Possuo um imenso sentimento de gratidão à mesma, já que os professores realizam trabalhos de qualidade e os demais funcionários garantem o ambiente acadêmico agradável, o que facilita o estudo.

Agradeço ao meu orientador, professor Dr. Benedito Donizeti Bonatto, por sempre ser atencioso, além de sua alta qualificação para esta tarefa e seu grande conhecimento geral. Agradeço ao Prof. Pedro Paulo Balestrassi pelas ricas contribuições ao trabalho e ao Prof. Antônio Carlos Zambroni de Souza, por ter aceitado ser coorientador em função do afastamento para pós-doutoramento do Prof. Bonatto.

Agradeço também o apoio da CAPES - Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, CNPq – Conselho Nacional de Desenvolvimento Científico e Tecnológico, FAPEMIG - Fundação de Amparo à Pesquisa do Estado de Minas Gerais, FAPESP – Fundação de Apoio à Pesquisa do Estado de São Paulo, CPTEn - Centro Paulista de Estudos da Transição Energética (Projeto Fapesp 2021/11380-5), CePIL – Centro Paulista de Inovação em Iluminação Pública (Projeto Fapesp 2024/01110-9) e UNICAMP – Universidade Estadual de Campinas.

Resumo

Este trabalho apresenta a relevância da análise paramétrica em estudos elétricos utilizando o *Electromagnetic Transients Program* (EMTP), expondo alguns estudos elaborados utilizando o módulo *Parametric Modeling Suite* (PAMSUITE) do software EMTP-RV. É apresentado um exemplo de caso preparado utilizando um conjunto de outras ferramentas computacionais, incluindo o pacote ATP. Ao comparar a abordagem tradicional com a análise paramétrica, observa-se que os resultados produzidos pela análise paramétrica são muito mais detalhados, com os quais é possível tomar decisões mais bem informadas e economicamente viáveis sobre o sistema.

Palavras-chave: EMTP, análise paramétrica, análise de sensibilidade, variabilidade do sistema, incerteza dos parâmetros.

Abstract

This work presents the relevance of parametric analysis in electrical studies using the Electromagnetic Transients Program (EMTP), exposing some studies prepared using the Parametric Modeling Suite (PAMSUITE) module of the EMTP-RV software and presenting an example of a case prepared using a set of other computational tools, including the ATP package. When comparing the traditional approach with parametric analysis, it is observed that the results produced by parametric analysis are much more detailed, with which it is possible to make better informed and economically viable decisions about the system.

Keywords: EMTP, parametric analysis, sensitivity analysis, system variability, parameter uncertainty.

Lista de Figuras

Figura 1 – Função de densidade de probabilidade. Fonte: [13].....	25
Figura 2 – Função de densidade de probabilidade e função de distribuição acumulada. Fonte: [13].....	26
Figura 3 – Formato de uma distribuição normal e probabilidades associadas. Fonte: Autoria Própria.....	28
Figura 4 – Esquema de aplicação básica de métodos de Monte Carlo. Fonte: [16].....	31
Figura 5 – Método Morris: representação de uma rede de 5 níveis ($p = 5$), duas entradas ($k = 2$) e $\Delta = 1/4$. Fonte: [20].	36
Figura 6 – Método Morris: exemplo de trajetória no espaço de entrada quando $k = 3$. Fonte: [25].....	37
Figura 7 – Método Morris: distribuições empíricas obtidas para diferentes valores de Δ , quando $r = 20$ trajetórias: (a) $\Delta = 2/3$. (b) $\Delta = 1/3$	40
Figura 8 – Procedimento de amostragem para fatores de entrada com distribuição normal padrão.....	41
Figura 9 – Gráfico de dispersão Y versus $X_1, \dots, X_4$. Fonte: [25].....	45
Figura 10 – Gráfico de dispersão Y versus $X_1, \dots, X_4$ cortado em fatias. Fonte: [25].	46
Figura 11 – Média dentro de cada fatia do gráfico de dispersão. Fonte: Adaptado de [25]. ...	47
Figura 12 – Classes e formas de onda de sobretensões - ensaios normalizados de tensão suportável. Fonte: [34].	57
Figura 13 – Fluxograma do gerenciamento das simulações. Fonte: Autoria Própria.....	61
Figura 14 – Circuito RLC série. Fonte: Autoria Própria.....	62
Figura 15 – Tipos de amortecimento em circuito RLC série. Fonte: Autoria Própria.....	63

Figura 16 – Tensão versus tempo no capacitor para 1000 simulações. Fonte: Autoria Própria.	65
Figura 17 – Análise de sensibilidade: (a) método de Morris, (b) método de Sobol sem intervalo de confiança. Fonte: Autoria Própria.	66
Figura 18 – Boxplot para comparativo da tensão máxima no capacitor para 1000 simulações. Fonte: Autoria Própria.	67
Figura 19 – Análise de sensibilidade: comparativo dos índices do método Morris para as duas diferentes metodologias. Fonte: Autoria Própria.	68
Figura 20 – Análise de sensibilidade: comparativo dos índices do método Sobol para as duas diferentes metodologias. Fonte: Autoria Própria.	69
Figura 21 – Diagrama unifilar simplificado. Fonte: Autoria Própria.	70
Figura 22 – Curva de duas rampas. Fonte: Adaptado de [34].	71
Figura 23 – Exemplo de curva de duas rampas no EMTP-RV. Fonte: Autoria Própria.	73
Figura 24 – Histograma do módulo da corrente de pico da descarga atmosférica. Fonte: Autoria Própria.	74
Figura 25 – Geometria considerada para a torre. Fonte: Adaptado [43].	76
Figura 26 – Resumo da modelagem adotada para a linha de transmissão. Fonte: Adaptado de [32, 38].	78
Figura 27 – Característica do para-raios. Fonte: Autoria Própria.	81
Figura 28 – Tensão versus tempo. (a) para-raios de entrada de linha, (b) chave seccionadora de <i>by-pass</i> da entrada de linha e (c) para-raios do setor de 230 kV do transformador de potência. Fonte: Autoria Própria.	84
Figura 29 – Tensão versus tempo: curva de maior valor absoluto de tensão de crista. Fonte: Autoria Própria.	85

Figura 30 – (a) probabilidade de o valor máximo absoluto da tensão de crista superar 500 kV e (b) função de distribuição cumulativa do valor máximo absoluto da tensão de crista. Fonte: Autoria Própria.	86
Figura 31 – Análise de sensibilidade: (a) método de Morris, (b) método de Sobol sem intervalo de confiança e (c) método de Sobol com intervalo de confiança. Fonte: Autoria Própria.	88
Figura 32 – Gráfico de dispersão: (a) resistência do pé da torre, (b) corrente de descarga e (c) capacitância do transformador de potência. Fonte: Autoria Própria.	90
Figura 33 – Gráfico de teia de aranha. Fonte: Autoria Própria.	90
Figura 34 – Entrada de dados da linha de transmissão na rotina do EMTP-RV. Fonte: Adaptado Própria.	104
Figura 35 – Entrada de dados no PAMSUITE e definição das saídas de interesse. Fonte: Adaptado Própria.	105
Figura 36 – Planta da subestação. Fonte: Confidencial.	106
Figura 37 – Corte da subestação. Fonte: Confidencial.	107
Figura 38 – Modelagem computacional do sistema em EMTP-RV. Fonte: Autoria Própria.	108

Lista de Tabelas

Tabela 1– Exemplo de valores para uma variável aleatória. Fonte: adaptado de [14].	23
Tabela 2– Exemplo de distribuição de probabilidade. Fonte: adaptado de [14].	23
Tabela 3– Comparação dos métodos de análise de sensibilidade local e global. Fonte: adaptado de [21].	34
Tabela 4– Parâmetros adotados para o circuito RLC. Fonte: Autoria Própria.	64
Tabela 5– Resumo dos principais parâmetros adotados para a linha de transmissão. Fonte: Autoria Própria.	78
Tabela 6– Resumo dos principais parâmetros adotados para os equipamentos da subestação. Fonte: Autoria Própria.	80
Tabela 7– Resumo dos parâmetros variáveis. Fonte: Autoria Própria.	82
Tabela 8– Script em Python: definição das funções para alteração do arquivo “. atp”. Fonte: Autoria Própria.	99
Tabela 9– Script em Python: tratamento de incerteza pelo método de Monte Carlo. atp”. Fonte: Autoria Própria.	100
Tabela 10– Script em Python: análise de sensibilidade pelo cálculo dos índices Morris. Fonte: Autoria Própria.	100
Tabela 11– Script em Python: análise de sensibilidade pelo cálculo dos índices Sobol. Fonte: Autoria Própria.	102
Tabela 12– Script em Python: interpretação dos resultados gerados pelo ATP. Fonte: Adaptado de [45].	102

Lista de Abreviaturas e Siglas

ATP	Programa de transitórios eletromagnéticos <i>Alternative Transients Program</i>
CAA	Condutor de Alumínio com Alma de Aço
CFO	<i>Critical Flashover Voltage</i> (tensão crítica de <i>flashover</i>)
CIGRE	<i>Conseil International des Grands Réseaux Electriques</i> (Conselho Internacional de Grandes Sistemas Elétricos)
EAR	Extra-Alta Resistência
EDF	Empresa francesa <i>Électricité de France</i>
EMTP	<i>Electromagnetic Transients Program</i> (Programa de Transitórios Eletromagnéticos)
GSA	Análise de Sensibilidade Global
IEC	<i>International Electrotechnical Commission</i> (Comissão Eletrotécnica Internacional)
LSA	Análise de Sensibilidade Local
MATLAB	Software de computação numérica da empresa MathWorks
NBI	Nível básico de isolamento ao impulso atmosférico
ONS	Operador Nacional do Sistema Elétrico
OPGW	Cabo para-raios óptico <i>Optical Ground Wire</i>
PAMSUITE	Módulo <i>Parametric Modelling Suite</i> do EMTP-RV
TEM	Transitórios eletromagnéticos
TOV	<i>Temporary overvoltage</i> (sobretensão temporária)
TSO	Operador francês <i>Transmission System Operator</i>
UNIFEI	Universidade Federal de Itajubá

Lista de Símbolos

$\mathbf{B}$	matriz utilizada para o cálculo de $\mathbf{B}^*$ para o método Morris
$\mathbf{B}^*$	matriz que fornece um efeito elementar por entrada para o método Morris
c	velocidade da luz
C	capacitância
C_{DJ}	capacitância do disjuntor
C_{isol}	capacitância da cadeia de isoladores
C_{IP}	capacitância do isolador de pedestal
C_{secc}	capacitância da chave seccionadora
C_{TC}	capacitância do transformador de corrente
C_{TP}	capacitância do transformador de potencial indutivo
C_{TPC}	capacitância do transformador de potencial capacitivo
C_{TR}	capacitância do transformador de força
d_{ml}	distância entre o par de trajetórias m e l
$\mathbf{D}^*$	matriz utilizada para o cálculo de $\mathbf{B}^*$ para o método Morris
e	constante matemática com valor aproximado de 2,7183
E	evento
$E(X)$	média ou valor esperado
EE_i^j	efeito elementar associado ao fator i referente a dois pontos da j -ésima trajetória
$f(x)$	função densidade de probabilidade
$F(x)$	função de probabilidade cumulativa
F_i	distribuição finita de efeitos elementares para o método Morris

G_i	distribuição dos valores absolutos dos efeitos elementares para o método Morris
i	corrente elétrica
I_f	valor de crista da corrente de descarga atmosférica
$J_{k+1,k}$	matriz utilizada para o cálculo de $\mathbf{B}^*$ para o método Morris
k	número de entradas
L	indutância
$L_{\text{aterramento}}$	indutância do condutor de aterramento
N	tamanho da amostra
$N(\mu, \sigma^2)$	distribuição normal
p	nível da rede para o método Morris ou proporção da população
$P(E)$	probabilidade de um evento E
P_{max}	valor máximo de tensão de crista obtido por meio de simulação computacional
$\mathbf{P}^*$	matriz utilizada para o cálculo de $\mathbf{B}^*$ para o método Morris
$\hat{p}$	estimador pontual da proporção da população p
r	número de trajetória para o método Morris
R	resistência
$R_{\text{pe_torre}}$	resistência do pé da torre
S	desvio-padrão da amostra
S_i	efeito de primeira ordem para métodos baseados em variância
$S_{ij \dots k}$	efeito de k -ésima ordem para métodos baseados em variância
S_{Ti}	efeito total de X_i para métodos baseados em variância
Surto_sort	valor sorteado da função densidade de probabilidade utilizado para cálculo da corrente de crista da descarga atmosférica
t	tempo

t_f	tempo de frente da descarga atmosférica
t_h	tempo de meio valor da descarga atmosférica
$U(a, b)$	distribuição uniforme
v	velocidade de propagação da onda
v_c	tensão no capacitor
V	tensão
$V(X)$	variância de X
V_{flash}	tensão mínima para a disrupção do isolamento
$\mathbf{x}^{(i)}$	i -ésimo ponto da trajetória para o método Morris
$\mathbf{x}^*$	valor base a construção de uma trajetória para o método Morris
x_i^*	valor fixado de uma variável aleatória X_i
X	variável aleatória
$\mathbf{X}_{\sim i}$	todos os fatores exceto X_i
$\bar{X}$	média da amostra
Y	admitância
$z_{\alpha/2}$	ponto superior com $100 \cdot \alpha/2$ % da distribuição normal padrão
Z	impedância
ZC_{torre}	impedância de surto da torre
α	nível de significância
Δ	múltiplo predeterminado de $1/(p - 1)$ para o método Morris
λ	parâmetro da função exponencial
μ	média ou valor esperado
μ^*	média de G_i para o método Morris
π	constante matemática com valor aproximado de 3,1415
σ	desvio-padrão

$\Phi(z)$	função de distribuição cumulativa de Z
Ω	espaço amostral ou ohm

Sumário

1	INTRODUÇÃO	17
1.1	Visão Geral	17
1.2	Relevância do tema.....	18
1.3	Objetivos	19
1.4	Estrutura do trabalho.....	19
2	REVISÃO DA LITERATURA	21
2.1	Probabilidade e estatística.....	21
2.1.1	Espaços amostrais e variáveis aleatórias.....	21
2.1.2	Distribuições de probabilidades para variáveis aleatórias discretas	22
2.1.2.1	Distribuições de probabilidades e funções de probabilidade	22
2.1.2.2	Funções de distribuição cumulativa.....	23
2.1.2.3	Média e variância de uma variável aleatória discreta	24
2.1.2.4	Distribuição binomial.....	24
2.1.3	Distribuições de probabilidades para variáveis aleatórias contínuas	25
2.1.3.1	Distribuições de probabilidades e funções de densidades de probabilidade	25
2.1.3.2	Funções de distribuição cumulativa.....	26
2.1.3.3	Média e variância de uma variável aleatória contínua.....	26
2.1.3.4	Distribuição contínua uniforme	27
2.1.3.5	Distribuição normal	27
2.1.3.6	Distribuição exponencial.....	29
2.1.3.7	Distribuição lognormal	29
2.1.4	Dependência estatística e correlação	29
2.1.5	Propagação de incertezas	30
2.1.5.1	Incerteza combinada de medições	30
2.1.5.1.1	Métodos de Monte Carlo.....	30
2.1.5.2	Intervalo de confiança.....	31
2.2	Análise de sensibilidade.....	33
2.2.1	Método Morris.....	35
2.2.1.1	Melhorias propostas na literatura.....	41
2.2.1.2	Conclusão	42
2.2.2	Métodos baseados em variância	43

2.2.2.1	Variância condicional e índice de sensibilidade de primeira ordem	44
2.2.2.2	Modelos não aditivos	47
2.2.2.3	Efeitos de interação e índices de sensibilidade de ordem superior.....	48
2.2.2.4	Efeito total.....	50
2.2.2.5	Computando os índices de sensibilidade	51
2.2.2.6	Diferentes testes para diferentes cenários	53
2.2.2.6.1	Priorização de Fatores.....	53
2.2.2.6.2	Fixação de Fatores	54
2.2.2.6.3	Corte de Variância	55
2.2.2.6.4	Mapeamento de Fatores	55
2.2.2.7	Conclusão	55
2.3	Modelagem computacional em estudos de TEM	56
3	METODOLOGIA	59
4	ESTUDOS DE CASO.....	62
4.1	Estudo de caso 1: Circuito RLC em série	62
4.1.1	Modelagem computacional e parametrização	63
4.1.2	Resultados da simulação	64
4.2	Estudo de caso 2: Sobretensões em subestação devido a descarga atmosférica.....	69
4.2.1	Modelagem computacional e parametrização	70
4.2.1.1	Fonte de frequência fundamental.....	71
4.2.1.2	Descarga atmosférica	71
4.2.1.3	Linha de transmissão, torres metálicas e cadeia de isoladores.....	74
4.2.1.4	Subestação	78
4.2.2	Resultados da simulação	83
5	CONCLUSÃO	92
5.1	Publicações na área	93
5.2	Trabalhos Futuros	93
	REFERÊNCIAS.....	96
	APÊNDICE A – ROTINA DE AUTOMATIZAÇÃO DO PYTHON PARA O ESTUDO DE CASO 1	99
	APÊNDICE B – PARÂMETROS E MODELAGEM DO ESTUDO DE CASO 2 - SOBRETENSÕES DEVIDO A DESCARGA ATMOSFÉRICA	104

1 Introdução

1.1 Visão Geral

Idealmente, em estudos de transitórios eletromagnéticos (TEM), a resposta do sistema é bem definida, ou seja, a estrutura do sistema e as características de seus componentes são perfeitamente definidos. Porém, este cenário está longe da realidade. Por um lado, é muito comum que as características dos componentes não sejam exatamente conhecidas ou até mesmo que sejam totalmente desconhecidas. Neste caso, existe incerteza dos parâmetros do sistema [1]. Por outro lado, o sistema pode apresentar diferentes estruturas, por exemplo, de acordo com o Operador Nacional do Sistema Elétrico (ONS), é necessário considerar as diferentes topologias operacionais de uma subestação em estudos de coordenação de isolamento [2]. Neste caso, há variabilidade da estrutura do sistema [1].

Além disso, a incerteza e a variabilidade do sistema são frequentemente associadas a uma outra questão: a influência de cada parâmetro e/ou configuração nos resultados, ou seja, uma análise de sensibilidade é interessante e pode ser a principal parte do estudo, por exemplo se o objetivo é determinar a melhor estratégia de investimento para reduzir a probabilidade de falta [1].

A incerteza de determinado parâmetro pode ser originada por diversos fatores. Por exemplo, a determinação do parâmetro pode ser dada indiretamente, por meio de medições de campo. Neste caso, pode-se utilizar um processo de identificação dos parâmetros a fim de modelar mais precisamente o componente conforme as medições, ou seja, um processo de casamento entre os dados medidos e modelados, conforme utilizado em [3].

Neste trabalho, o estudo que considera a incerteza dos parâmetros, a variabilidade do sistema, a análise de sensibilidade e/ou identificação de parâmetros é chamado de **análise paramétrica**.

Para lidar com os problemas supracitados, diversas estratégias já foram utilizadas, sendo que na maioria dos estudos de TEM o engenheiro precisa fornecer a solução “a partir do zero”. Desta forma, na maioria dos casos, o engenheiro simplifica o problema a fim de ir de um sistema parcialmente indefinido para um sistema perfeitamente definido [1], ao:

- Negligenciar as imprecisões e os dados desconhecidos ao utilizar valores “típicos” para caracterizar parâmetros desconhecidos;

- Estudar o sistema para um limitado número de configurações e/ou conjunto de parâmetros definidos como críticos, ou seja, cenários de pior caso.

Contudo, os resultados obtidos com essas estratégias, que é chamada neste trabalho de **abordagem tradicional**, podem possuir significantes falhas pois a confiabilidade dos resultados depende muito das suposições feitas pelo engenheiro e do conhecimento do engenheiro na área. Além disso, as estratégias que tradicionalmente consideram o **cenário de pior caso** para lidar com as incertezas nos dados de entrada são ineficientes e nem sempre conservadoras, conforme detalhado em [1].

Em algumas situações, para realizar a análise paramétrica, engenheiros criam *scripts* para um estudo em particular, como em [4, 5, 6]. Estes scripts são poderosos e responsáveis por realizar diversas simulações, considerando diferentes conjuntos de dados e topologias. Em outros casos, como em [7], utiliza-se as próprias capacidades do ATP para a realização de múltiplas simulações.

1.2 Relevância do tema

De acordo com o exposto no subcapítulo anterior, a abordagem tradicional para estudos de TEM pode ser ineficiente.

A Constituição Federal do Brasil em seu Art. 37 estabelece 5 princípios para a administração pública: Legalidade, Impessoalidade, Moralidade, Publicidade e Eficiência [8]. Tendo isso em vista, qualquer ação que mitigue perdas desnecessárias e apresente uma forma de economizar gastos vêm de encontro ao princípio constitucional [9].

A preocupação acerca da incerteza dos parâmetros de entrada parece ser global, já que o tema foi um dos tópicos discutidos pela comissão de estudos C4 - que estuda a performance técnica do sistema elétrico - do CIGRE (do francês *Conseil International des Grands Réseaux Electriques*), no encontro realizado em 2018, em Paris, onde um determinado número de organizações reportou a maneira que lidam com a incerteza dos parâmetros, principalmente em relação aos estudos de harmônicos para novas conexões na rede elétrica. Além disso, foi reportado pela EDF (do francês *Électricité de France*) o uso do módulo PAMSUITE (do inglês *Parametric Modelling Suite*), do software EMTP-RV, para a realização de análise paramétrica [1, 10].

Uma outra preocupação atual é com relação aos modelos matemáticos. Sabe-se que é fundamental que os modelos dos componentes do sistema elétrico sejam representativos

(desempenho aderente ao comportamento real) e que a qualidade das diretrizes do operador do sistema elétrico está intimamente ligada com a qualidade do modelo matemático.

Recentemente, foi emitido pelo ONS um relatório contendo a análise de perturbação ocorrida no Brasil em 15/08/2023 [11], que resultou na separação dos Subsistemas Norte, Nordeste e Acre/Rondônia do restante do Sistema Interligado Nacional brasileiro, onde foi constatado que “a abrupta redução de tensão observada no campo após a perda de uma única linha de transmissão foi consequência do desempenho dos parques eólicos e fotovoltaicos observado em campo ter sido inesperado, muito aquém daquele obtido pelo ONS nos seus estudos, os quais são realizados utilizando-se os modelos matemáticos encaminhados pelos agentes ao ONS”.

Ainda com relação a este relatório, determinou-se a necessidade de validação em campo pelos agentes dos modelos dos parques eólicos e fotovoltaicos. Diante disso, a análise paramétrica pode ser uma ferramenta importante para avaliação dos modelos quanto à identificação dos parâmetros mais importantes e dos negligenciáveis e quanto à calibração do modelo, por exemplo.

1.3 Objetivos

O objetivo principal deste trabalho é ampliar a literatura existente acerca da análise paramétrica aplicada a estudos de TEM. Buscou-se implementar estudos de caso que mostram como este tipo de análise pode beneficiar o engenheiro especialista ao fornecer dados estocásticos mais bem detalhados para tomadas de decisão.

1.4 Estrutura do trabalho

A estrutura da presente dissertação é composta por uma fundamentação teórica, presente no capítulo a seguir, que tem como objetivo o refinamento de informações indispensáveis para o entendimento de conceitos estatísticos e a análise de resultados. Para este capítulo, destacam-se os assuntos relacionados a probabilidades e medidas de estatísticas, método de Monte Carlo e ferramentas para a análise de sensibilidade.

O terceiro capítulo apresenta a metodologia proposta, onde é evidenciada a relação entre a fundamentação teórica e os estudos de caso expostos.

O quarto capítulo contém a aplicação prática do conteúdo presente no segundo capítulo. Destaca-se a realização da análise paramétrica de duas diferentes formas: por meio da criação de um programa computacional utilizando Python para controlar a execução do ATP; e por meio do uso das ferramentas disponíveis no EMTP-RV e PAMSUITE.

Por fim, baseado nos resultados obtidos, é apresentada uma conclusão e recomendações para estudos futuros. As referências utilizadas, que por sua vez foram indispensáveis para a confecção deste trabalho, estão listadas, e em seguida estão os anexos desta dissertação.

2 Revisão da Literatura

Este capítulo é dedicado à revisão dos trabalhos que embasam a dissertação. São revisados conceitos básicos de probabilidade e estatística, transitórios eletromagnéticos e literaturas relacionadas à análise paramétrica.

2.1 Probabilidade e estatística

O campo de estatística lida com a coleta, apresentação, análise e uso dos dados para tomar decisões, resolver problemas e planejar produtos e processos. Devido a muitos aspectos da prática de engenharia envolverem o trabalho com dados, obviamente algum conhecimento de estatística é muito importante para qualquer engenheiro. Especificamente, técnicas estatísticas podem ser uma ajuda poderosa no planejamento de novos produtos e sistemas, melhorando os projetos existentes e planejando, desenvolvendo e melhorando os processos de produção [12].

O termo probabilidade se refere ao estudo da aleatoriedade e da incerteza [13].

Em qualquer situação em que ocorrem diversos resultados, a teoria da probabilidade oferece métodos de quantificação das chances ou possibilidades de ocorrência associadas aos diversos resultados [13].

2.1.1 Espaços amostrais e variáveis aleatórias

Ao medir a corrente em um fio fino de cobre, é conduzido um experimento. Entretanto, em repetições diárias da medida, os resultados poderão diferir levemente, por causa de pequenas variações que não estejam controladas no experimento, como variações na temperatura e oscilações na fonte de corrente. Consequentemente, é dito que este experimento possui um componente aleatório. Em geral não importa quão cuidadosamente um experimento tenha sido planejado e conduzido, a variação está quase sempre presente e sua magnitude pode ser suficientemente grande de tal modo que as conclusões importantes acerca do experimento podem não ser óbvias [12].

Chama-se experimento aleatório aquele que pode fornecer diferentes resultados, muito embora seja repetido toda vez da mesma maneira [12].

O conjunto de todos os resultados possíveis de um experimento aleatório é chamado de espaço amostral do experimento. O espaço amostral é denotado por Ω e é usualmente definido baseado nos objetivos da análise [12].

É útil distinguir entre dois tipos de espaços amostrais: discreto, caso ele consista em um conjunto finito ou infinito contável de resultados, e contínuo, se ele contém um intervalo (tanto finito como infinito) de números reais [12].

Frequentemente é interessante resumir o resultado de um experimento aleatório através de um simples número. A variável que associa um número ao resultado de um experimento aleatório é referida como variável aleatória. A variável aleatória é denotada por uma letra maiúscula, tal como X . Depois de um experimento ser conduzido, o valor medido da variável aleatória é denotado por uma letra minúscula, tal como x [12].

Uma variável aleatória discreta é uma variável aleatória com um faixa finita (ou infinita contável). Uma variável aleatória contínua é uma variável aleatória com um intervalo (tanto finito como infinito) de números reais para sua faixa [12].

A distribuição de probabilidades de uma variável aleatória X é uma descrição das probabilidades associadas com os valores possíveis de X [12].

2.1.2 Distribuições de probabilidades para variáveis aleatórias discretas

2.1.2.1 Distribuições de probabilidades e funções de probabilidade

Para uma variável aleatória discreta, a distribuição é frequentemente especificada por apenas uma lista de valores possíveis, juntamente com a probabilidade de cada um. Em alguns casos, é conveniente expressar a probabilidade em termos de uma fórmula [12].

Para um espaço amostral discreto, a probabilidade de um evento E , denotada por $P(E)$, é igual à soma das probabilidades dos resultados em E [12].

Por exemplo, uma moeda é lançada duas vezes e é observada sua face. O espaço amostral é $\Omega = \{KK, KC, CK, CC\}$. Uma variável aleatória de interesse poderia ser o número de caras. A cada evento simples ou ponto de Ω , associamos um número, que é o valor assumido para a variável aleatória X . Os valores de X são então 0, 1 e 2, conforme Tabela 1:

Tabela 1– Exemplo de valores para uma variável aleatória. Fonte: adaptado de [14].

Evento	KK	KC	CK	CC
X	0	1	1	2

A Tabela 2 mostra a distribuição de probabilidade do exemplo citado:

Tabela 2– Exemplo de distribuição de probabilidade. Fonte: adaptado de [14].

X	0	1	2
$P(X = x_i)$	0,25	0,50	0,25

Para uma variável aleatória discreta X , com valores possíveis $x_1, x_2, \dots, x_n$, a função de probabilidade é uma função tal que [12]:

$$f(x_i) \geq 0 \quad (1)$$

$$\sum_{i=1}^n f(x_i) = 1 \quad (2)$$

$$f(x_i) = P(X = x_i) \quad (3)$$

2.1.2.2 Funções de distribuição cumulativa

Algumas vezes é útil ser capaz de expressar probabilidades cumulativas, tais como $P(X \leq x)$ e que tais probabilidades cumulativas podem ser usadas para encontrar a função probabilidade de uma variável aleatória. Portanto, o uso de probabilidades cumulativas é um método alternativo de descrever a distribuição de probabilidades de uma variável aleatória [12].

De acordo com [12], a função de distribuição cumulativa de uma variável aleatória discreta X , denotada por $F(x)$, pode ser escrita conforme a Eq. (4) e satisfaz as propriedades descritas nas Eq. (5) e (6):

$$F(x) = P(X \leq x) = \sum_{x_i \leq x} f(x_i) \quad (4)$$

$$0 \leq F(x) \leq 1 \quad (5)$$

$$\text{Se } x \leq y, \text{ então } F(x) \leq F(y) \quad (6)$$

2.1.2.3 Média e variância de uma variável aleatória discreta

Dois números são frequentemente usados para resumir uma distribuição de probabilidades para uma variável aleatória X . A média é uma medida do centro ou meio da distribuição de probabilidades e a variância é uma medida da dispersão ou variabilidade na distribuição. Essas duas medidas não identificam unicamente uma distribuição de probabilidades, ou seja, duas distribuições diferentes podem ter a mesma média e a variância [12].

De acordo com [12, 13] a média ou valor esperado de uma variável aleatória discreta X , denotada(o) como μ ou $E(X)$, é calculada(o) conforme Eq. (7). A variância de X , denotada por σ^2 ou $V(X)$, é calculada conforme Eq. (8). O desvio-padrão de X é calculado conforme Eq. (9).

$$\mu = E(X) = \sum_x x \cdot f(x) \quad (7)$$

$$\sigma^2 = V(X) = E[(X - \mu)^2] = \sum_x (x - \mu)^2 \cdot f(x) = \sum_x x^2 \cdot f(x) - \mu^2 \quad (8)$$

$$\sigma = \sqrt{\sigma^2} \quad (9)$$

A média de uma variável aleatória X é uma média ponderada dos valores possíveis de X , com pesos iguais às probabilidades [12].

A variância de uma variável aleatória X é uma medida de dispersão ou espalhamento nos valores possíveis para X [12].

2.1.2.4 Distribuição binomial

Um experimento para o qual as condições abaixo são satisfeitas é denominado experimento binomial [13]:

- O experimento consiste em uma sequência de N experimentos menores denominados tentativas, onde N é estabelecido antes do experimento.
- Cada tentativa pode resultar em um de dois resultados possíveis, chamados de sucesso ou falha.
- As tentativas são independentes, de forma que o resultado de qualquer tentativa particular não influencia o resultado de qualquer outra tentativa.

- A probabilidade de sucesso é constante de uma tentativa para a outra. Denominamos essa probabilidade de p .

Dado um experimento binomial, a variável aleatória binominal a ele associada é definida como quantidade de sucesso nas N tentativas. A função de probabilidade de X é definida conforme Eq. (10) [12]:

$$f(x) = \binom{N}{x} \cdot p^x \cdot (1-p)^{N-x}, x = 0, 1, \dots, N \quad (10)$$

2.1.3 Distribuições de probabilidades para variáveis aleatórias contínuas

2.1.3.1 Distribuições de probabilidades e funções de densidades de probabilidade

Uma função densidade de probabilidade $f(x)$ pode ser usada para descrever a distribuição de probabilidades de uma variável aleatória contínua X [12]. A probabilidade de X ter um determinado valor no intervalo $[a, b]$ é a área contida entre o intervalo abaixo da curva da função de densidade [13], conforme ilustrado na Figura 1:

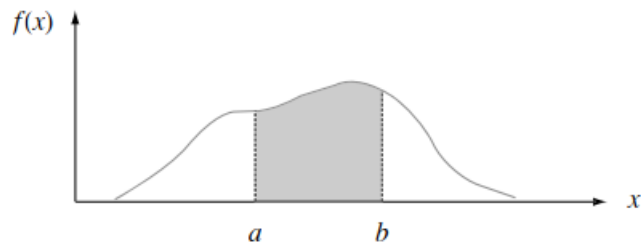


Figura 1 – Função de densidade de probabilidade. Fonte: [13].

De acordo com [12], para uma variável aleatória contínua X , uma função densidade de probabilidade é uma função tal que:

$$f(x) \geq 0 \quad (11)$$

$$\int_{-\infty}^{\infty} f(x) \cdot dx = 1 \quad (12)$$

$$P(a \leq X \leq b) = \int_a^b f(x) \cdot dx, \text{ para qualquer } a \text{ e } b \quad (13)$$

2.1.3.2 Funções de distribuição cumulativa

A função de distribuição cumulativa de uma variável aleatória X é [12]:

$$F(x) = P(X \leq x) = \int_{-\infty}^x f(u) \cdot du, \text{ para } -\infty \leq x \leq \infty \quad (14)$$

Na equação (13), para cada x , $F(x)$ é a área abaixo da curva de densidade à esquerda de x [13]. A Figura 2 exemplifica a obtenção da função de distribuição acumulada:

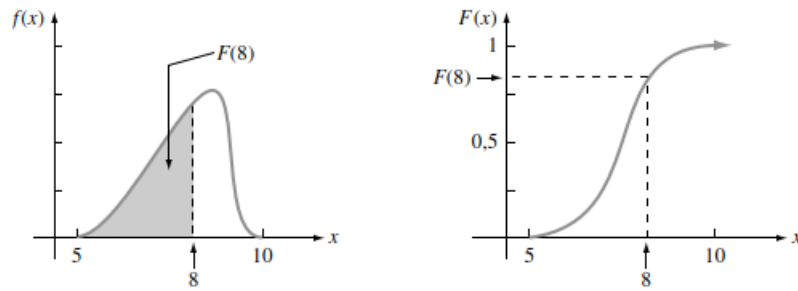


Figura 2 – Função de densidade de probabilidade e função de distribuição acumulada. Fonte: [13].

2.1.3.3 Média e variância de uma variável aleatória contínua

A média e a variância de uma variável aleatória contínua são definidas de modo similar a uma variável aleatória discreta. A integração substitui a soma nas definições de variáveis discretas [12].

Suponha que X seja uma variável aleatória contínua, com uma função densidade de probabilidade $f(x)$. A média ou valor esperado de uma variável aleatória contínua X , denotada(o) como μ ou $E(X)$, é calculada(o) conforme Eq. (15). A variância de X , denotada por σ^2 ou $V(X)$, é calculada conforme Eq. (16). O desvio-padrão de X é calculado conforme Eq. (17) [12].

$$\mu = E(X) = \int_{-\infty}^{\infty} x \cdot f(x) \cdot dx \quad (15)$$

$$\sigma^2 = V(X) = \int_{-\infty}^{\infty} (x - \mu)^2 \cdot f(x) \cdot dx = \int_{-\infty}^{\infty} x^2 \cdot f(x) \cdot dx - \mu^2 \quad (16)$$

$$\sigma = \sqrt{\sigma^2} \quad (17)$$

2.1.3.4 Distribuição contínua uniforme

A distribuição contínua uniforme é denotada por $U(a,b)$. De acordo com [12], uma variável aleatória contínua X , com uma função densidade de probabilidade conforme Eq. (18) tem uma variável aleatória contínua uniforme.

$$f(x) = \frac{1}{(b-a)}, a \leq x \leq b \quad (18)$$

Considerando a Eq. (18) nas Eq. (15) e (16), tem-se as Eq. (19) e (20):

$$\mu = \frac{a+b}{2} \quad (19)$$

$$V(X) = \frac{(b-a)^2}{12} \quad (20)$$

2.1.3.5 Distribuição normal

A distribuição normal, também conhecida como gaussiana e notação $N(\mu;\sigma^2)$, provavelmente é a mais importante distribuição de probabilidade, tanto na teoria quanto na prática da estatística [12].

Muitas populações numéricas possuem distribuições que podem ser ajustadas aproximadamente por uma curva normal apropriada. Inclusive, mesmo quando a distribuição em questão é discreta, a curva normal frequentemente fornece aproximação excelente. Além disso, ainda que as próprias variáveis individuais não sejam normalmente distribuídas, as somas e as médias das variáveis terão uma distribuição aproximadamente normal sob condições adequadas. Essa é a essência do Teorema do Limite Central [13].

A distribuição normal é simétrica em torno da média, o que implica que a média, a mediana e a moda são todas coincidentes. Esta distribuição é assintótica, ou seja, a curva aproxima-se cada vez mais do eixo x, mas nunca o toca efetivamente, conforme ilustrado na Figura 3:

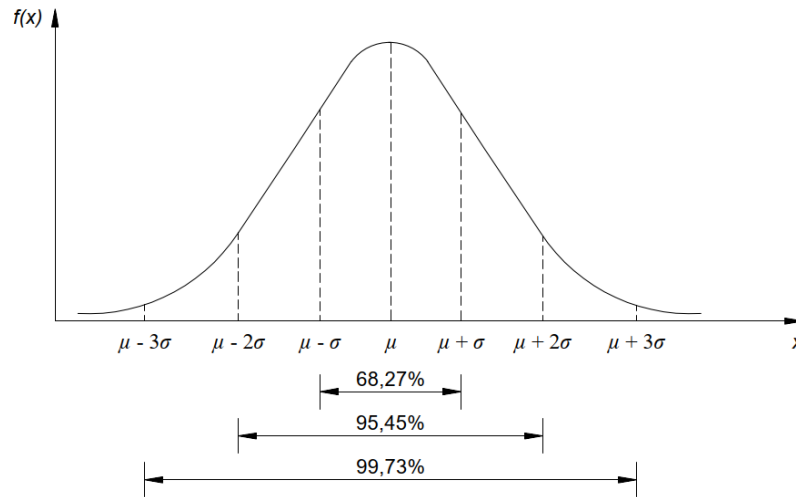


Figura 3 – Formato de uma distribuição normal e probabilidades associadas. Fonte: Autoria Própria.

Se X é uma variável aleatória normal, então a função densidade de probabilidade de X é definida conforme Eq. (21) [14]:

$$f(x) = \frac{1}{\sqrt{2 \cdot \pi} \cdot \sigma} \cdot e^{-\frac{(x-\mu)^2}{2 \cdot \sigma^2}}, -\infty < x < \infty \quad (21)$$

Para calcular $P(a \leq X \leq b)$, quando X é uma variável aleatória normal, deve-se utilizar a Eq. (22):

$$P(a \leq X \leq b) = \int_a^b \frac{1}{\sqrt{2 \cdot \pi} \cdot \sigma} \cdot e^{-\frac{(x-\mu)^2}{2 \cdot \sigma^2}} \cdot dx \quad (22)$$

Nenhuma das técnicas de integração-padrão pode ser usada para calcular a Eq. (22). Em vez disso, quando $\mu = 0$ e $\sigma = 1$, a equação supracitada foi calculada numericamente e tabulada para valores determinados de a a b . Esta tabela também é usada para calcular probabilidades de quaisquer outros valores de μ e σ que estejam em consideração [13].

Uma variável aleatória normal com $\mu = 0$ e $\sigma = 1$ é chamada de variável aleatória normal padrão [13].

2.1.3.6 Distribuição exponencial

Diz-se que X tem uma distribuição exponencial com parâmetro λ ($\lambda > 0$) se a função densidade de probabilidade de X é definida conforme Eq. (23) [13]:

$$f(x; \lambda) = \begin{cases} \lambda \cdot e^{-\lambda \cdot x}, & x \geq 0 \\ 0, & \text{caso contrário} \end{cases} \quad (23)$$

2.1.3.7 Distribuição lognormal

Diz-se que uma variável aleatória não-negativa X tem uma distribuição lognormal se a variável aleatória $\ln(X)$ possui uma distribuição normal. A função densidade de probabilidade resultante de uma variável aleatória lognormal quando $\ln(X)$ tiver distribuição normal com parâmetros μ e σ é definida conforme Eq. (24):

$$f(x; \mu, \sigma) = \begin{cases} \frac{1}{\sigma \cdot x \cdot \sqrt{2 \cdot \pi}} \cdot e^{-[\ln(x) - \mu]^2 / (2 \cdot \sigma^2)}, & x \geq 0 \\ 0, & \text{caso contrário} \end{cases} \quad (24)$$

2.1.4 Dependência estatística e correlação

Duas variáveis aleatórias X e Y são independentes se sua função distribuição de probabilidade conjunta ou sua função densidade de probabilidade for o produto das duas funções distribuição de probabilidade ou funções densidade de probabilidade marginais [13].

Quando duas variáveis aleatórias não são independentes, geralmente é de interesse avaliar quão fortemente está relacionada uma com a outra e a relação entre elas pode ser de vários tipos e, no caso de ser linear, pode-se definir uma medida para essa dependência, a covariância [13, 14]. Caso a relação entre as variáveis aleatórias não for linear, a covariância pode não ser sensível à relação [12].

Há uma outra medida da relação entre duas variáveis aleatórias que é frequentemente mais fácil de interpretar que a covariância, a correlação, que só escalona a covariância através

do desvio-padrão de cada variável. Assim, a correlação é uma grandeza adimensional que pode ser usada para comparar as relações lineares entre pares de variáveis em diferentes unidades de medida [12].

De acordo com [13], se duas variáveis são independentes, então são não-correlacionadas. No entanto, duas variáveis podem ser não-correlacionadas, porém altamente dependentes, pois pode existir uma relação não-linear forte.

2.1.5 Propagação de incertezas

A medição de uma grandeza é dita direta caso apenas um sistema de medição está envolvido, ou seja, o sistema de medição já indica naturalmente o valor do mensurado. Medições são ditas indiretas quando o valor do mensurado é calculado a partir de operações matemáticas efetuadas envolvendo duas ou mais medidas associadas a diferentes características do mensurando [15]. Dessa forma, a incerteza da variável desejada depende das incertezas de cada uma das grandezas de entrada [16].

2.1.5.1 Incerteza combinada de medições

Atualmente existem duas abordagens usadas para modelar a propagação de incertezas. A primeira é o clássico Método da Série de Taylor com termos de ordem superior negligenciados (tornando-o menos “exato”). O segundo é o Método de Monte Carlo, que amostra erros extraídos de distribuições assumidas e simula a execução do teste muitas vezes com um conjunto diferente de erros amostrados a cada vez. Ambos os métodos são informativos, mas com poder de computação rápido e barato, o Método Monte Carlo é geralmente preferido [17].

2.1.5.1.1 Métodos de Monte Carlo

A metodologia básica para a aplicação dos métodos de Monte Carlo para a propagação de incertezas está esquematizada na Figura 4. Para cada parâmetro de entrada são geradas N amostras aleatórias. Essas variáveis randômicas devem respeitar os momentos estatísticos de cada variável (por exemplo, para uma distribuição normal os momentos estatísticos são média e desvio padrão). Para cada conjunto de amostras geradas, o modelo matemático que calcula o

mensurando é avaliado. Finalmente, é feita a análise estatística de todo o conjunto de resultados obtidos, do qual a incerteza do mensurando é definida [16, 17].

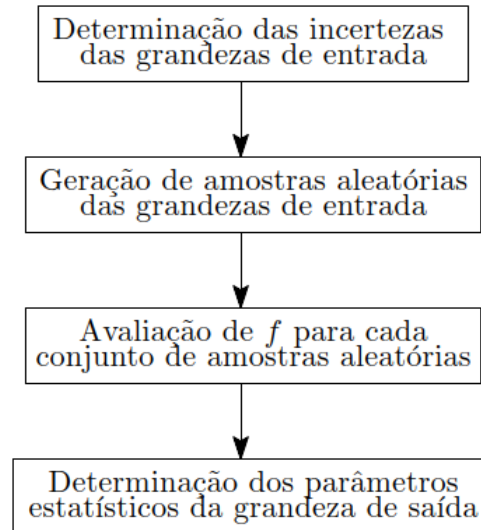


Figura 4 – Esquema de aplicação básica de métodos de Monte Carlo. Fonte: [16].

Outra vantagem da propagação de incertezas por meio de métodos de Monte Carlo é não ser necessário assumir nenhum tipo de distribuição para a variável de saída [16]. Se a simulação é altamente não-linear e/ou se as incertezas das variáveis são relativamente grandes, a distribuição dos resultados de Monte Carlo pode ser assimétrica [17].

2.1.5.2 Intervalo de confiança

Define-se como intervalo de confiança, a faixa de valores dentro da qual espera-se, para um certo nível de confiança desejado, encontrar o valor verdadeiro do mensurando. Por exemplo, para 95,45 % de nível de confiança, o intervalo de confiança correspondente tem 95,45 % de probabilidade de conter o valor verdadeiro do mensurando [16].

Não é possível estar certo de que o intervalo contém o parâmetro verdadeiro desconhecido da população. No entanto, o intervalo de confiança é construído de modo que tenhamos alta confiança de que ele contém o parâmetro desconhecido da população [12].

A maneira para determinar o intervalo de confiança para a média de uma distribuição pode variar dependendo do fato de conhecer ou não a variância da população, bem como o tamanho da amostra aleatória.

Caso o tamanho da amostra N for grande, o teorema do limite central implica que $\bar{X}$ (média da amostra) tenha aproximadamente uma distribuição normal com média μ e variância $\sigma^2/\sqrt{N}$. Entretanto, o desvio-padrão σ é desconhecido. No entanto, quando N é grande, a troca de σ por S (desvio-padrão da amostra) tem pouco efeito na distribuição normal, conseqüentemente o intervalo de confiança com $100 \cdot (1 - \alpha)\%$ de confiança pode ser calculado conforme Eq. (25):

$$\bar{x} - z_{\alpha/2} \cdot \frac{s}{\sqrt{N}} \leq \mu \leq \bar{x} + z_{\alpha/2} \cdot \frac{s}{\sqrt{N}} \quad (25)$$

sendo $z_{\alpha/2}$ o ponto superior com $100 \cdot \alpha/2 \%$ da distribuição normal padrão. Ainda, de acordo com [12], geralmente, N deve ser, no mínimo, igual a 40 para Eq. (25) ser confiável.

Além do intervalo de confiança **para a média** de uma distribuição, frequentemente é necessário construir intervalos de confiança **para a proporção** de uma população. Por exemplo, suponha que uma amostra aleatória de tamanho N tenha sido retirada de uma grande população e que X ($\leq N$) observações nessa amostra pertençam a uma classe de interesse. Então, $\hat{P} = X/N$ é um estimador pontual da proporção da população p que pertence a esta classe [12].

Note que os parâmetros N e p supracitados são os parâmetros de uma distribuição binomial. Além disso, a distribuição amostral de $\hat{P}$ é aproximadamente normal, se p não estiver muito próximo de 0 ou 1 e N for relativamente grande e, neste caso, o intervalo aproximado de confiança de $100 \cdot (1 - \alpha)\%$ para a proporção p da população pode ser calculado conforme Eq. (26) [12]:

$$\hat{p} - z_{\alpha/2} \cdot \sqrt{\frac{\hat{p} \cdot (1 - \hat{p})}{N}} \leq p \leq \hat{p} + z_{\alpha/2} \cdot \sqrt{\frac{\hat{p} \cdot (1 - \hat{p})}{N}} \quad (26)$$

sendo $\hat{p}$ a proporção de observações em uma amostra aleatória de tamanho N que pertença a uma classe de interesse.

A amplitude do intervalo de confiança em conjunto com o nível de confiança desejado pode ser utilizada para determinar o tamanho necessário da amostra.

2.2 Análise de sensibilidade

Neste trabalho a análise de sensibilidade é definida como a maneira pela qual o parâmetro de entrada afeta a saída de um determinado modelo. No entanto, de acordo com [18], existem diferentes definições a depender do campo de estudo.

Conforme [19], a análise de sensibilidade pode ser motivada por alguns fatores, como a necessidade de determinar:

- Quais parâmetros requer maior investigação a fim de detalhar melhor seu modelo, resultando na redução da incerteza na saída de interesse;
- Quais parâmetros são insignificantes e podem ser eliminados do modelo;
- Quais entradas contribuem mais significativamente para a variabilidade da saída;
- Quais parâmetros são mais altamente correlacionados com a saída.

Os métodos de análise de sensibilidade podem ser classificados de acordo com o âmbito, aplicabilidade e características. A classificação mais simples e mais comum são análise de sensibilidade local (LSA) e análise de sensibilidade global (GSA). A LSA é focada nos efeitos de entradas incertas em torno de um ponto (ou caso base), ou seja, é baseado nas derivadas locais de saída do modelo em respeito a variação de um único parâmetro, que indica a rapidez com que a saída aumenta ou diminui localmente em torno do valor do parâmetro de referência. Já a GSA é focada mais nas influências de entradas incertas sobre todo o espaço de entrada [18, 20, 21].

Na GSA, a variância da saída do modelo é avaliada em relação ao quanto os parâmetros variam em todos os seus domínios de incerteza. Isso fornece uma visão mais realista do comportamento do modelo quando usado na prática [22]. A GSA ainda pode ser dividida de acordo com quatro abordagens: regressão; baseada em triagem, baseada em variância, e metamodelo [21]. A Tabela 3 descreve as características dos métodos supracitados.

Tabela 3– Comparação dos métodos de análise de sensibilidade local e global. Fonte: adaptado de [21].

	Método	Subtipo	Características
Local	Local	-	Não explora a incerteza; nenhuma auto verificação; não considera a interação de parâmetros, Implementação simples; fácil entendimento; baixo custo computacional.
Global	Regressão	SRC SRRC <i>t-value</i>	SRC e <i>t-value</i> : adequados para modelos lineares; SRRC: adequado para modelos não-lineares monótono; fácil implementação e entendimento; moderado custo computacional;
	Triagem	Morris	Não explora a incerteza; nenhuma auto verificação; adequado para um maior número de entradas e modelos computacionalmente intensivos; medida qualitativa para classificar fatores.
	Baseado em variância	FAST Sobol	Decompõe a variância da saída do modelo para cada entrada; considera os efeitos principais e de interação; medidas quantitativas; alto custo computacional; Fast não é adequado para distribuições discretas.
	metamodelo	MARS ACOSSO SVM	Adequado para modelos complexos e computacionalmente intensivos; quantifica a variância de saída devido às diferentes entradas; a precisão depende do modelo.

onde:

- SRC: do inglês *standardised regression coefficients*;
- SRRC: do inglês *standardized rank regression coefficient*;
- FAST: do inglês *Fourier amplitude sensitivity test*;
- MARS: do inglês *multivariate adaptive regression splines*;
- ACOSSO: do inglês *adaptive component selection and smoothing operator*;
- SVM: do inglês *support vector machine*.

Diante do exposto na Tabela 3, deve-se utilizar a GSA ao invés da LSA para a análise de sensibilidade aplicada ao estudo de TEM que contempla a incerteza nos parâmetros de

entrada e a interação de parâmetros. Quanto ao método utilizado, é importante que não tenha um alto custo computacional.

Conduzir GSA usando métodos de decomposição de variância em complexos modelos de simulação com muitos fatores de entrada é usualmente inacessível. Uma alternativa é primeiro utilizar o método de triagem a fim de reduzir o número de fatores de entrada e então aplicar o método de decomposição de variância ao modelo reduzido [23].

Os subitens a seguir contém a descrição dos métodos Morris e Sobol, que é utilizado neste trabalho para a análise de sensibilidade.

2.2.1 Método Morris

Modelos matemáticos são, frequentemente, complexos, computacionalmente caros e envolvem vários fatores de entrada. Nestes casos, uma das necessidades é conhecer quais fatores de entrada são significativos para o modelo e, para isso, é essencial a adoção de um método computacionalmente barato, como o método de triagem [20].

O método Morris [24] é um método do tipo triagem utilizado para a GSA onde cada fator de entrada pode assumir um número discreto de valores, chamado níveis. Duas medidas são propostas por Morris para cada fator: μ , que estima o efeito geral do fator na saída, e σ , que estima o conjunto dos efeitos de segunda ordem e de ordem superior em que o fator está envolvido (incluindo curvaturas e efeitos de interação). A medida μ proposta por Morris é obtida ao computar um número, r , de diferentes pontos do espaço de entrada e então tomar a média. O número r de pontos selecionados é chamado de tamanho da amostra do experimento ou número de trajetórias [20].

A descrição do método Morris a seguir foi baseada em [20, 25].

A filosofia do método Morris é determinar quais fatores podem ter efeitos (a) negligenciáveis, (b) linear e aditivo, ou (c) não-linear ou envolvido com interações com outros fatores. O impacto da mudança de um fator de cada vez é avaliado.

Considere que o vetor $\mathbf{X}$, de dimensão k , do modelo de entrada possua componentes X_i , os quais podem assumir os valores do conjunto $\{0, 1/(p-1), 2/(p-1), \dots, 1\}$. A região do experimento, Ω , é uma rede de nível p e dimensão k .

O método sugerido por Morris é baseado em *efeito elementar*. O efeito elementar da i -ésima entrada é definido a seguir. Seja Δ um múltiplo predeterminado de $1/(p-1)$. Para um dado valor $\mathbf{x}$ de $\mathbf{X}$, o efeito elementar da i -ésima entrada é definido conforme Eq. (27):

$$d_i(\mathbf{x}) = \frac{[y(x_1, \dots, x_{i-1}, x_i + \Delta, x_{i+1}, \dots, x_k) - y(\mathbf{x})]}{\Delta} \quad (27)$$

onde $\mathbf{x} = (x_1, x_2, \dots, x_k)$ é qualquer valor selecionado em Ω , de tal modo que o ponto transformado $(\mathbf{x} + \mathbf{e}_i \cdot \Delta)$ ainda está em Ω para cada índice $i = 1, \dots, k$, onde $\mathbf{e}_i$ é um vetor de zeros, exceto no i -ésimo componente, que possui valor unitário.

A distribuição finita de efeitos elementares associados ao i -ésimo fator de entrada é obtida por amostragem aleatória de diferentes de $\mathbf{x}$ de Ω , e é denotada por F_i . O número de elementos de cada F_i é $p^{k-1} \cdot [p - \Delta \cdot (p - 1)]$. Por exemplo, considere $k = 2$, $p = 5$ e $\Delta = 1/4$. Neste caso, tem-se um número total de 20 elementos para cada F_i . A Figura 5 ilustra este exemplo:

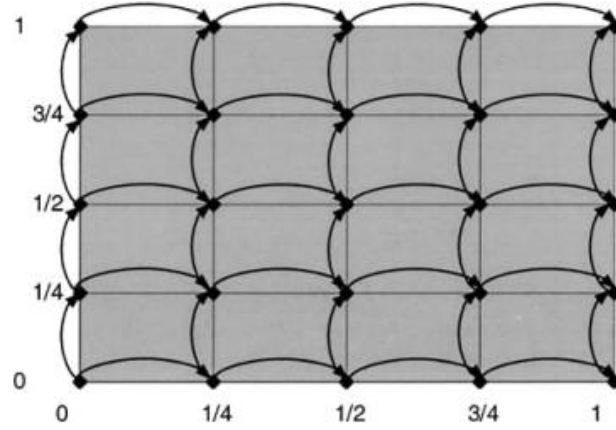


Figura 5 – Método Morris: representação de uma rede de 5 níveis ($p = 5$), duas entradas ($k = 2$) e $\Delta = 1/4$. Fonte: [20].

Cada flecha da figura acima identifica um par de pontos necessário para computar um efeito elementar. As flechas horizontais identificam 20 efeitos elementares relativos a X_1 , enquanto as flechas verticais identificam 20 efeitos elementares relativos a X_2 .

O método Morris se baseia no cálculo da média, μ , e do desvio padrão, σ^2 , de F_i . É necessário determinar o número de trajetórias, r , a fim de estimar as estatísticas da distribuição. Em sua forma mais simples, cada efeito elementar requer a avaliação de y duas vezes. Neste caso, o esforço computacional total necessário para uma amostra aleatória de r valores de cada F_i é $n = 2 \cdot r \cdot k$ execuções. Porém, foi sugerido pelo próprio Morris um *design* mais eficiente (com menor número de execuções). Este *design* é baseado na construção de uma matriz, $\mathbf{B}^*$, de dimensão $k \times (k + 1)$, cujas linhas representam os vetores de entrada $\mathbf{x}_s$, para o qual o

experimento correspondente fornece k efeitos elementares, um para cada fator de entrada, de $k + 1$ execuções.

As trajetórias são geradas da seguinte maneira: primeiramente, é selecionado aleatoriamente um valor base de $\mathbf{x}^*$, para o vetor $\mathbf{X}$. Cada componente x_i de $\mathbf{x}^*$ é amostrado do conjunto $\{0, 1/(p - 1), 2/(p - 1), \dots, 1\}$. O vetor $\mathbf{x}^*$ não faz parte da trajetória, mas é utilizado para gerar todos os pontos da trajetória. O primeiro ponto da trajetória, $\mathbf{x}^{(1)}$, é obtido ao incrementar um ou mais componentes de $\mathbf{x}^*$ pelo valor de Δ . A escolha dos componentes de $\mathbf{x}^*$ a ser incrementado é condicionado por $\mathbf{x}^{(1)}$ estar em Ω . O segundo ponto, $\mathbf{x}^{(2)}$, é gerado de $\mathbf{x}^*$ com a propriedade que ele difere de $\mathbf{x}^{(1)}$ em seu i -ésimo componente, que pode ser incrementado ou decrementado pelo valor de Δ . O índice i é aleatoriamente escolhido no conjunto $\{1, 2, \dots, k\}$. O terceiro ponto, $\mathbf{x}^{(3)}$, é gerado de $\mathbf{x}^*$ com a propriedade que ele difere de $\mathbf{x}^{(2)}$ em apenas um componente, j , para qualquer $j \neq i$, que pode ser incrementado ou decrementado pelo valor de Δ . O processo continua até $\mathbf{x}^{(k+1)}$, que termina a trajetória. Desta forma, tem-se $(k + 1)$ pontos amostrados $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \dots, \mathbf{x}^{(k+1)}$, cujas propriedades chave são: dois pontos consecutivos diferem em apenas um componente; qualquer valor do vetor base $\mathbf{x}^*$ foi selecionado ao menos uma vez para ser acrescentado por Δ . A Figura 6 ilustra um exemplo de trajetória:

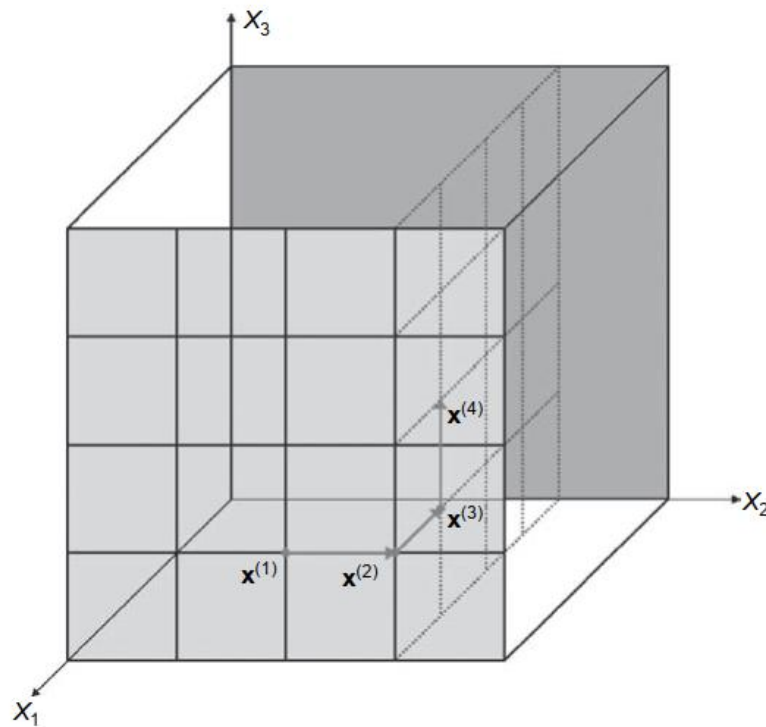


Figura 6 – Método Morris: exemplo de trajetória no espaço de entrada quando $k = 3$. Fonte: [25].

A técnica para gerar trajetórias com as propriedades requeridas é descrita a seguir. Uma trajetória pode ser vista na forma de uma matriz, $\mathbf{B}^*$, com dimensão $(k + 1) \times k$, cujas linhas são os vetores $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \dots, \mathbf{x}^{(k+1)}$. Para construir $\mathbf{B}^*$, o primeiro passo é a seleção de uma matriz $\mathbf{B}$, cuja dimensão é $(k + 1) \times k$, composta de elementos iguais a 0's e 1's e a propriedade chave é que para cada índice de coluna, j , onde $j = 1, \dots, k$, existem duas linhas de $\mathbf{B}$ que diferem somente na j -ésima entrada. Uma escolha conveniente de $\mathbf{B}$ é uma matriz triangular inferior de 1's:

$$\mathbf{B} = \begin{bmatrix} 0 & 0 & 0 & \dots & 0 \\ 1 & 0 & 0 & \dots & 0 \\ 1 & 1 & 0 & \dots & 0 \\ 1 & 1 & 1 & 0 & \dots \\ \dots & \dots & \dots & \dots & \dots \end{bmatrix}$$

A matriz $\mathbf{B}'$, dada pela Eq. (28), é uma potencial candidata para a matriz desejada, no entanto, tem a limitação que o k -ésimo efeito elementar que ela produz não é aleatoriamente selecionado.

$$\mathbf{B}' = \mathbf{J}_{k+1,k} \cdot \mathbf{x}^* + \Delta \cdot \mathbf{B} \quad (28)$$

onde $\mathbf{J}_{k+1,k}$ é uma matriz com dimensão $(k + 1) \times k$ de 1's.

Uma versão aleatória da matriz amostral é dada pela Eq. (29):

$$\mathbf{B}^* = (\mathbf{J}_{k+1,1} \cdot \mathbf{x}^* + (\Delta/2) \cdot [(2 \cdot \mathbf{B} - \mathbf{J}_{k+1,k}) \cdot \mathbf{D}^* + \mathbf{J}_{k+1,k}]) \cdot \mathbf{P}^* \quad (29)$$

onde $\mathbf{D}^*$ é uma matriz diagonal de dimensão k e cada elemento é igual a +1 ou -1, com igual probabilidade, e $\mathbf{P}^*$ é uma matriz de permutação aleatória, com dimensão $k \times k$, que, para cada linha, há um elemento igual a 1 e os demais elementos iguais a 0, e não há duas colunas que possuem os 1's em mesma posição. Lido linha por linha, $\mathbf{P}^*$ fornece a ordem na qual os fatores são movidos; $\mathbf{D}^*$ indica se os fatores aumentarão ou diminuirão seu valor ao longo da trajetória. $\mathbf{B}^*$ fornece um efeito elementar por entrada, que é selecionado aleatoriamente.

A técnica para a montagem da matriz $\mathbf{B}^*$, descrita acima, é repetida um número r de vezes a fim de obter r trajetórias em Ω . Se $\mathbf{x}^{(l)}$ e $\mathbf{x}^{(l+1)}$, com l pertencente ao conjunto $\{1, \dots, k\}$, são dois pontos de j -ésima trajetória e diferem-se em seu i -ésimo componente, o efeito elementar associado ao fator i é dado pela Eq. (30) caso o i -ésimo componente de $\mathbf{x}^{(l)}$ seja aumentado por Δ , ou pela Eq. (31) caso o mesmo seja diminuído por Δ :

$$EE_i^j(\mathbf{x}^{(l)}) = \frac{[y(\mathbf{x}^{(l+1)}) - y(\mathbf{x}^{(l)})]}{\Delta} \quad (30)$$

$$EE_i^j(\mathbf{x}^{(l+1)}) = \frac{[y(\mathbf{x}^{(l)}) - y(\mathbf{x}^{(l+1)})]}{\Delta} \quad (31)$$

Uma vez que r efeitos elementares por entrada estejam disponíveis ($EE_i^j, i = 1, 2, \dots, k, j = 1, 2, \dots, r$), pode-se calcular a média, μ_i , e o desvio padrão, σ_i^2 , da distribuição F_i , conforme Eq. (32) e (33), respectivamente:

$$\mu_i = \frac{1}{r} \cdot \sum_{j=1}^r EE_i^j \quad (32)$$

$$\sigma_i^2 = \frac{1}{r-1} \cdot \sum_{j=1}^r (EE_i^j - \mu)^2 \quad (33)$$

A escolha crítica relacionada à implementação do método é a escolha dos parâmetros p e Δ . A escolha de p está estritamente ligada à escolha de r . Caso seja escolhido um alto valor de p , um alto número de possíveis níveis pode ser explorado, que aparentemente aumenta a precisão da amostragem. Caso esta escolha de p não esteja casada com a escolha de um alto valor de r , o esforço será desperdiçado, já que vários níveis possíveis não serão explorados. Geralmente, quando r é pequeno, é provável que nem todos os níveis possíveis dos fatores de entrada serão explorados no experimento.

Uma escolha conveniente para os parâmetros p e Δ é: p é par; $\Delta = p/[2 \cdot (p-1)]$. Esta escolha tem a vantagem de que a estratégia garante uma amostragem de igual probabilidade de cada F_i . Por exemplo, a parte de cima da Figura 7a (flechas cinza) mostra que, quando $p = 4$, a escolha de $\Delta = p/[2 \cdot (p-1)] = 2/3$ garante que os quatro níveis têm igual probabilidade de ser escolhido. Por outro lado, a Figura 7b ilustra que, para $\Delta = 1/3$, os níveis $1/3$ e $2/3$ são amostrados com maior frequência, já que existem duas flechas apontando para cada um deles. Os dois histogramas abaixo das flechas ilustram as distribuições empíricas obtidas quando geradas $r = 20$ trajetórias.

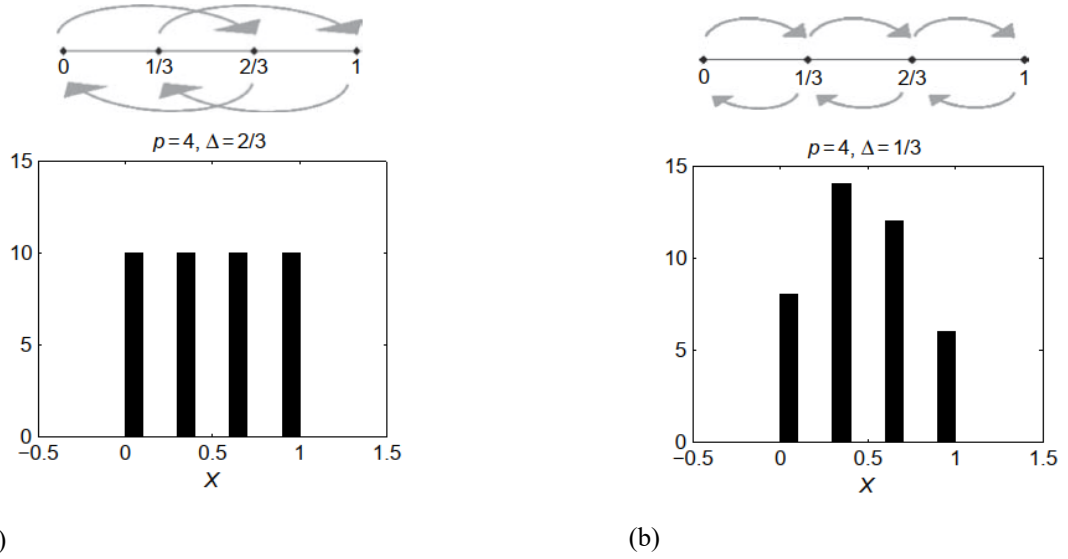


Figura 7 – Método Morris: distribuições empíricas obtidas para diferentes valores de Δ , quando $r = 20$ trajetórias: (a) $\Delta = 2/3$. (b) $\Delta = 1/3$.

Fonte: Adaptado de [25].

É ilustrado em [25] um caso no qual um número ímpar de níveis é considerado ($p = 5$). Nesta condição, é impossível alcançar igual probabilidade para os efeitos elementares, independentemente do valor atribuído para Δ . Além disso, em alguns casos, por exemplo $p = 5$ e $\Delta = 3/4$, existem efeitos elementares impossíveis de serem amostrados.

Se um fator de entrada segue uma distribuição uniforme, os níveis são obtidos simplesmente dividindo em partes iguais o intervalo em que cada fator varia. Se um fator segue distribuições não uniformes, então os valores de entrada não são amostrados diretamente; em vez disso, a amostragem é realizada no espaço dos quantis das distribuições, que é um hipercubo k -dimensional (cada quantil varia em $[0, 1]$). A Figura 8 mostra o caso em que um fator X segue uma distribuição padrão normal e o espaço de quantis é investigado através de seis quantis ($q_1, q_2, \dots, q_6$). Utilizando o inverso da função de distribuição cumulativa normal (curva cinza), obtêm-se níveis correspondentes para X ($L_1, L_2, \dots, L_6$).

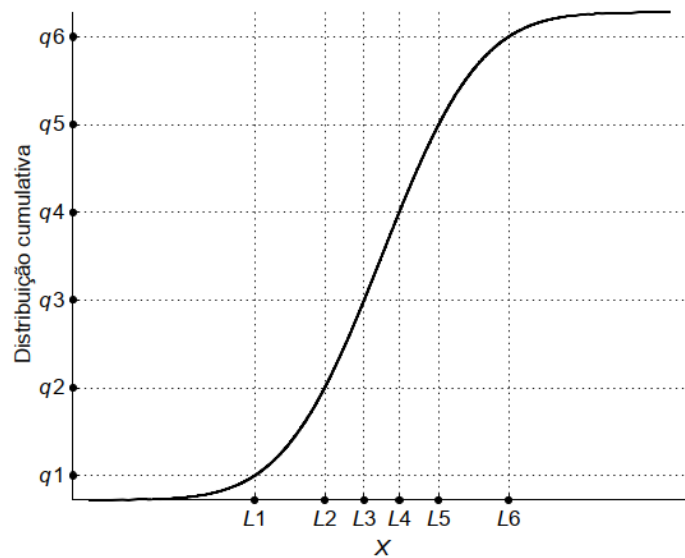


Figura 8 – Procedimento de amostragem para fatores de entrada com distribuição normal padrão.

Fonte: Adaptado de [25].

Se a distribuição a ser amostrada tiver suporte infinito, os quantis a serem utilizados podem ser escolhidos, por exemplo, cortando as caudas da distribuição.

Neste ponto, vale a pena considerar brevemente o papel de Δ . Se e somente se Δ for escolhido igual para todos os fatores de entrada, o seu papel na definição do efeito elementar para efeitos de sensibilidade torna-se irrelevante, pois Δ^{-1} é simplesmente uma constante que multiplica cada efeito elementar, o que não afeta os resultados da análise de sensibilidade.

2.2.1.1 Melhorias propostas na literatura

Algumas melhorias foram propostas em [26] e serão abordadas neste subitem.

Foi proposto o uso de μ^* , que é definido como a estimativa da média de G_i - distribuição dos valores absolutos dos efeitos elementares.

O uso de μ^* é conveniente pois resolve o problema de falha na identificação do fator com considerável influência no modelo, no qual μ é vulnerável. O tipo de erro supracitado pode ocorrer quando a distribuição F_i contém tanto valores positivos quanto valores negativos, isto é, quando o modelo é não monotônico ou tem efeitos interativos. Nestes casos, alguns efeitos podem cancelar outros ao comutar μ , produzindo baixo valor da média para um fator de entrada importante. A fim de evitar este erro, Morris [24] recomendou considerar os valores de μ e σ simultaneamente, já que um fator com efeitos elementares de diferentes sinais teria um baixo valor de μ , mas um considerável valor de σ . A representação gráfica no plano (μ, σ) permite

uma melhor interpretação dos resultados ao considerar, ao mesmo tempo, as duas medidas de sensibilidade. A abordagem de Morris é, em geral, válida, mas pode tornar-se problemática no caso de grandes modelos com várias saídas de interesse. Por outro lado, μ^* é uma medida prática e concisa para se utilizar, especialmente quando existem diversas variáveis de saída [25, 26]. Calcula-se μ_i^* conforme Eq. (34):

$$\mu_i^* = \frac{1}{r} \cdot \sum_{j=1}^r |EE_i^j| \quad (34)$$

Em geral, como a estimativa de μ é obtida sem custo computacional extra, é recomendado calcular os três índices μ , σ , μ^* , a fim de extrair a máxima quantidade de informação do modelo. Por exemplo, a comparação entre μ e μ^* fornece informações sobre os sinais dos efeitos que o fator tem na saída. Se μ e μ^* são altos, implica não apenas que o fator tem um grande efeito sobre a saída, mas também que o sinal desse efeito é sempre o mesmo. Se, pelo contrário, μ for baixo enquanto μ^* for alto, significa que o fator examinado tem efeitos de sinais diferentes dependendo do ponto do espaço em que o efeito é calculado.

Outra melhoria aplicada ao método Morris sugerida em [26] está relacionada com processo de amostragem. A ideia é selecionar r trajetórias de tal forma a maximizar o distanciamento entre elas no espaço de entrada. O procedimento se inicia ao gerar um alto número de trajetórias, por exemplo $M \sim 500 - 1000$, e então selecionar r trajetórias com maior espaçamento, onde o conceito de espaçamento é baseado na definição da distância, d_{ml} , entre o par de trajetórias m e l . Em outras palavras, d_{ml} é a soma das distâncias geométricas entre todos os pares de pontos de duas trajetórias sob análise. As melhores r trajetórias de M são selecionadas ao maximizar as distâncias d_{ml} entre elas. Esta estratégia otimiza a varredura do espaço de entrada e é simples de implementar.

2.2.1.2 Conclusão

O método apresentado é simples, fácil de implementar e de interpretar e sua principal vantagem é o baixo custo computacional, pois o número total de simulações é uma função linear que depende do número de fatores de entrada examinados [20].

Sugere-se utilizar o método quando o problema é selecionar alguns fatores de entrada importantes entre um grande número contido no modelo [20].

O método Morris, como todos os outros métodos do tipo triagem, fornece medida qualitativa, ou seja, capaz de ranquear os fatores de entrada em ordem de importância, mas não quantificar o quão um determinado fator é mais importante que outro [20].

2.2.2 Métodos baseados em variância

Os subcapítulos e considerações a seguir referentes aos métodos baseados em variância foram baseados em [25].

A maioria dos modelos vive a sua vida operacional como “determinística”. Cada vez que são “interrogados”, são alimentados com um conjunto determinístico de valores para as variáveis de entrada e a saída é investigada para possíveis inferências. A análise de sensibilidade para estes modelos geralmente implicará alterar uma entrada de cada vez para testar o seu efeito na saída. No entanto, neste subcapítulo é mostrado uma modelagem diferente, na qual considera-se explorar o modelo sobre diferentes combinações de valores para as entradas com incerteza.

Recomenda-se usar a variância como uma medida sumária de incerteza sempre que a aplicação permitir, a fim de explorar as propriedades estatísticas de variância descritas neste capítulo para investigar como os fatores contribuem para a variância.

Características interessantes de métodos baseados em variância são:

- independência do modelo: a medida de sensibilidade é livre de modelo;
- capacidade de capturar a influência de toda a gama de variação de cada fator de entrada;
- apreciação dos efeitos de interação entre fatores de entrada;
- capacidade para lidar com grupos de fatores de entrada: fatores incertos podem pertencer a diferentes tipos lógicos, e pode ser desejável decompor a incerteza de acordo com estes tipos.

O inconveniente de métodos baseados em variância é o alto custo computacional, como é discutido a seguir. Inclusive, esta é a razão pela qual muitas pesquisas recentes visam encontrar um algoritmo numérico eficiente para seu cálculo, sendo que medidas de sensibilidade baseadas na decomposição da variância do resultado do modelo são relativamente recentes na literatura.

2.2.2.1 Variância condicional e índice de sensibilidade de primeira ordem

Considere o modelo genérico da Eq. (35):

$$Y = f(X_1, X_2, \dots, X_k) \quad (35)$$

Cada X possui uma faixa de variação ou incerteza não nula e deseja-se determinar o que aconteceria com a incerteza de Y caso fosse possível fixar um fator. Imagine que seja fixado o fator X_i em um valor específico x_i^* . Seja $V_{X_{\sim i}}(Y|X_i = x_i^*)$ a variância resultante de Y , tomada sobre $X_{\sim i}$ (todos os fatores exceto X_i). Isto é chamado de variância condicional, pois é condicional a X_i ser fixado em x_i^* . Pode-se imaginar que, tendo congelado uma fonte potencial de variação (X_i), a variância resultante $V_{X_{\sim i}}(Y|X_i = x_i^*)$ será menor que a variância total ou incondicional correspondente $V(Y)$. É possível, portanto, conceber a utilização de $V_{X_{\sim i}}(Y|X_i = x_i^*)$ como medida da importância relativa de X_i , raciocinando que quanto menor $V_{X_{\sim i}}(Y|X_i = x_i^*)$, maior será a influência de X_i . Há dois problemas com esta abordagem. Primeiro, torna a medida de sensibilidade dependente da posição do ponto x_i^* para cada fator de entrada, o que é impraticável. Em segundo lugar, pode-se projetar um modelo que, para fatores específicos X_i e ponto fixo x_i^* , produza $V_{X_{\sim i}}(Y|X_i = x_i^*) > V(Y)$. Se, em vez disso, tomarmos a média desta medida sobre todos os pontos possíveis x_i^* , a dependência de x_i^* desaparecerá. Isso é escrito como $E_{X_i}(V_{X_{\sim i}}(Y|X_i))$, que é sempre menor ou igual a $V(Y)$ e possui a relação descrita na Eq. (36):

$$E_{X_i}(V_{X_{\sim i}}(Y|X_i)) + V_{X_i}(E_{X_{\sim i}}(Y|X_i)) = V(Y) \quad (36)$$

Portanto, um pequeno valor de $E_{X_i}(V_{X_{\sim i}}(Y|X_i))$, ou um alto valor de $V_{X_i}(E_{X_{\sim i}}(Y|X_i))$, implica que X_i é um fator importante.

Note que, de acordo com a Eq. (36), $V_{X_i}(E_{X_{\sim i}}(Y|X_i)) \leq V(Y)$. A variância condicional $V_{X_i}(E_{X_{\sim i}}(Y|X_i))$ é chamada de efeito de primeira ordem de X_i em Y e a medida dada pela Eq. (37) é conhecida como índice de sensibilidade de primeira ordem X_i em Y . S_i é um número sempre entre 0 e 1.

$$S_i = \frac{V_{X_i}(E_{X \sim i}(Y|X_i))}{V(Y)} \quad (37)$$

Valores altos de S_i indica que a variável é importante.

O índice de primeira ordem representa a contribuição do efeito principal de cada fator de entrada para a variância do produto.

Um outro caminho para enxergar a variância condicional é exemplificado a seguir. Para isso, considere a Figura 9:

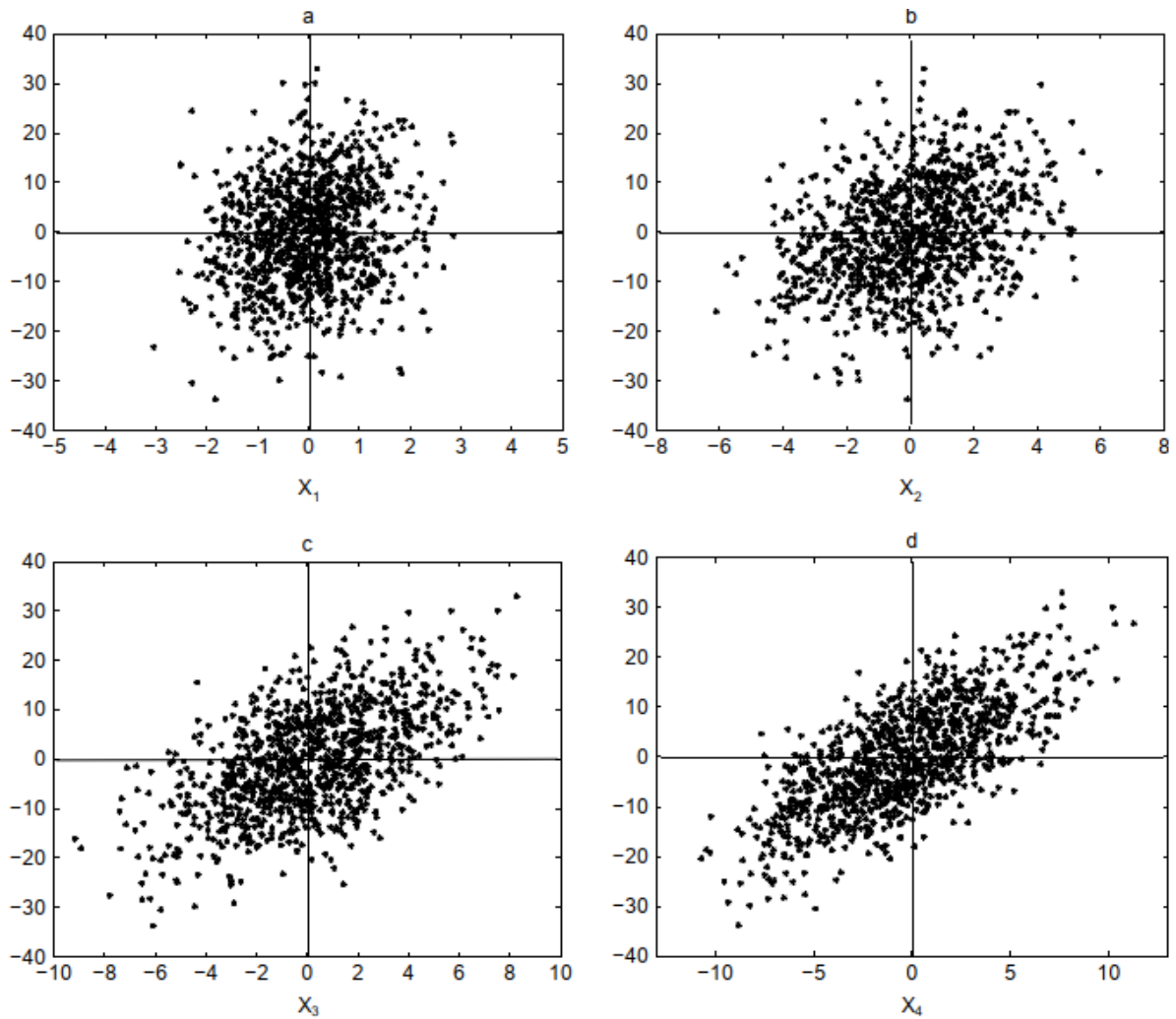


Figura 9 – Gráfico de dispersão Y versus $X_1, \dots, X_4$. Fonte: [25].

O gráfico de dispersão da Figura 9 revela que a ordem de influência dos fatores de entrada na saída Y segue a relação a seguir:

$$X_4 > X_3 > X_2 > X_1 \quad (38)$$

Sendo que a saída Y é mais sensível ao X_4 do que X_3 , por exemplo, devido ao padrão ou formato do gráfico. Pode-se dizer que se tem um padrão quando a distribuição dos pontos Y sobre a abscissa, ou seja, sobre o fator X_i não é uniforme. Em outras palavras, deve-se analisar se caso o eixo X_i for cortado em fatias, serão observadas diferenças na distribuição dos pontos Y nas fatias (ver Figura 10), e se o valor médio de Y em cada fatia varia entre as fatias (ver Figura 11).

Na Figura 10, que é igual à Figura 9 com a adição de fatias, e na Figura 11 pode-se ver que o fator X_4 é o mais influente, conforme relação da Eq. (38), de acordo com o quanto o valor médio de Y varia de uma fatia para outra. Nota-se também que o fator X_3 é influente, mas ligeiramente menos influente que o fator X_4 .

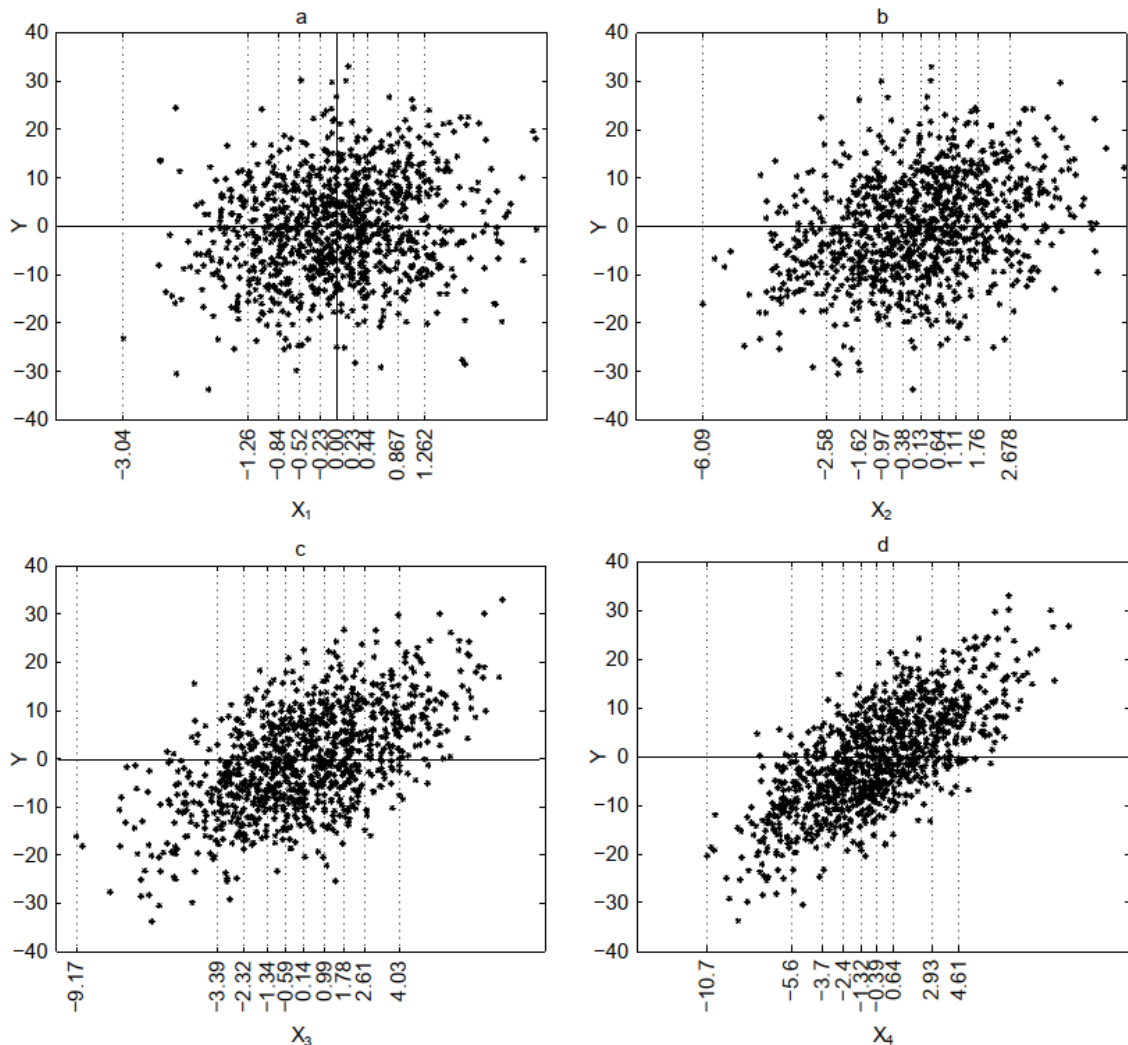


Figura 10 – Gráfico de dispersão Y versus $X_1, \dots, X_4$ cortado em fatias. Fonte: [25].

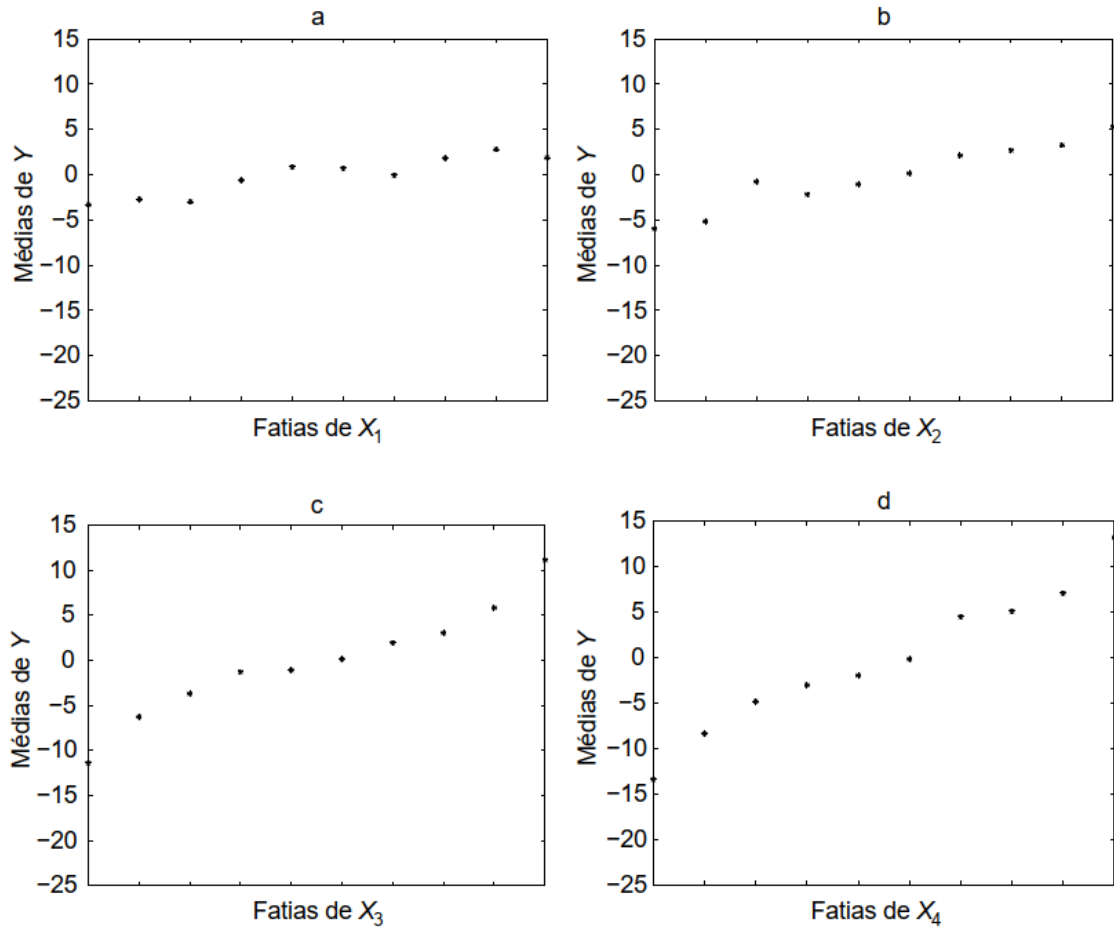


Figura 11 – Média dentro de cada fatia do gráfico de dispersão. Fonte: Adaptado de [25].

Tomando o limite para fatias muito finas, encontra-se $V_{X_i}(E_{X_{\sim i}}(Y|X_i))$. Note, de fato, que o valor esperado de Y em uma fatia muito fina corresponde a manter X_i fixo enquanto calcula-se a média de todos menos X_i , que é exatamente $E_{X_{\sim i}}(Y|X_i)$. O operador de variância também é facilmente compreendido.

2.2.2.2 Modelos não aditivos

A definição de modelos não aditivos é: a soma dos índices de primeira ordem é menor que um, ou seja, $\sum_{i=1}^k S_i < 1$. A aditividade de um modelo depende das características de seus fatores de entrada, de modo que é suficiente, por exemplo, transformar uma constante do modelo em um fator para mudar o modelo de aditivo para não aditivo, embora o modelo permaneça inalterado na forma.

2.2.2.3 Efeitos de interação e índices de sensibilidade de ordem superior

As interações entre os fatores de entrada podem implicar, por exemplo, que valores extremos do produto Y estão exclusivamente associados a combinações particulares de dados do modelo, de uma forma que não é descrita pelos efeitos de primeira ordem S_i . As interações representam características importantes dos modelos e são mais difíceis de detectar do que os efeitos de primeira ordem.

O matemático Sobol provou que, dada uma função f quadrado integrável sobre Ω^k , os termos da Eq. (39) podem ser calculados usando as esperanças condicionais do resultado Y , desde que a decomposição da função f , descrita pela Eq. (39), atenda certas condições.

Ainda, cabe ressaltar que na expansão considerada por Sobol, cada termo é uma função apenas dos fatores em seu índice, por exemplo $f_i = f_i(X_i)$, $f_{ij} = f_{ij}(X_i, X_j)$, etc.

$$f = f_0 + \sum_i f_i + \sum_i \sum_{j>i} f_{ij} + \dots + f_{12\dots k} \quad (39)$$

Em particular, tem-se, Eq. (40), (41) e (42):

$$f_0 = E(Y) \quad (40)$$

$$f_i = E(Y|X_i) - E(Y) \quad (41)$$

$$f_{ij} = E(Y|X_i, X_j) - f_i - f_j - E(Y) \quad (42)$$

Considere, por exemplo, dois fatores em vez de um, para a variância condicional, conforme representado abaixo:

$$\frac{V(E(Y|X_i, X_j))}{V(Y)} \quad (43)$$

Com $i \neq j$. Os índices dos operadores E e V foram eliminados. Na verdade, eles não são necessários caso seja considerada a convenção de que o argumento que condiciona o operador interno, X_i, X_j neste caso, é também o conjunto sobre o qual se aplica o operador externo, ou seja, a variância é assumida sobre X_i, X_j (pode-se escrever V_{X_i, X_j}). Por padrão, o operador interno, a média E , deve ser assumido sobre todos menos X_i, X_j .

Uma característica útil da decomposição descrita pelas Eq. (41) e (42) é que, para os dois fatores genéricos, X_i, X_j , tem-se Eq. (44):

$$V(E(Y|X_i, X_j)) = V_i + V_j + V_{ij} \quad (44)$$

onde:

$$V_i = V(E(Y|X_i)) \quad (45)$$

$$V_j = V(E(Y|X_j)) \quad (46)$$

O termo V_{ij} é o termo de interação entre os fatores X_i, X_j . Ele captura aquela parte da resposta de Y a X_i, X_j que não pode ser escrita como uma superposição de efeitos separadamente devido a X_i e X_j . O termo V_{ij} é o efeito conjunto de X_i e X_j menos os efeitos de primeira ordem para os mesmos fatores. V_{ij} é conhecido como efeito de segunda ordem. Fórmulas análogas podem ser escritas para termos de ordem superior, permitindo ao analista quantificar as interações de ordem superior.

Ao considerar a expansão descrita pela Eq. (39), tem-se a seguinte decomposição de variância para fatores independentes um do outro:

$$V(Y) = \sum_i V_i + \sum_i \sum_{j>i} V_{ij} + \dots + V_{12\dots k} \quad (47)$$

Ao dividir a Eq. (47) em ambos os lados por $V(Y)$, tem-se a Eq. (48):

$$\sum_i S_i + \sum_i \sum_{j>i} S_{ij} + \dots + S_{12\dots k} = 1 \quad (48)$$

Em aplicações práticas, o modelo objeto da análise será um programa de computador, e a única forma de verificar se existe ou não uma interação será estimá-la numericamente. O problema é que o desenvolvimento em série da Eq. (48) tem até 2^{k-1} termos. Para $k = 3$, isso dá apenas 7 termos, $S_1, S_2, S_3, S_{12}, S_{23}, S_{13}, S_{123}$. Para $k = 10$, resulta em 1023 termos, o que é muito, na prática. No entanto, pode-se utilizar o termo de *efeito total*, que é descrito a seguir a fim de ajudar a resolver este problema.

2.2.2.4 Efeito total

Os efeitos totais são uma consequência direta da abordagem de decomposição de variância e do procedimento de estimativa de Sobol, embora tenham sido explicitamente introduzidos e tornados computacionalmente acessíveis por outros investigadores.

O índice de efeito total contabiliza a contribuição total para a variação do produto devido ao fator X_i , ou seja, o seu efeito de primeira ordem mais todos os efeitos de ordem superior devidos às interações.

Para um modelo de três fatores, por exemplo, o efeito total de X_1 é a soma de todos os termos da Eq. (48) onde o fator X_1 é considerado:

$$S_{T1} = S_1 + S_{12} + S_{13} + S_{123} \quad (49)$$

Os índices totais são úteis na análise de sensibilidade, pois fornecem informações sobre as características não aditivas do modelo. Como mencionado, para um modelo puramente aditivo $\sum_{i=1}^k S_i = 1$, enquanto para um determinado fator X_j uma diferença significativa entre S_{Tj} e S_j sinaliza uma interação importante envolvendo esse fator. Os índices totais poderiam ser calculados em princípio calculando-se todos os termos na decomposição da Eq. (48). No entanto, existem até $2^k - 1$ destes.

Outra forma de encontrar o índice total é decompor a variância do produto $V(Y)$ em termos de efeito principal e residual conforme descrito pela Eq. (36), mas condicionando desta vez em relação a todos os fatores exceto um, ou seja, $X_{\sim i}$, conforme Eq. (50):

$$V(Y) = V(E(Y|X_{\sim i})) + E(V(Y|X_{\sim i})) \quad (50)$$

A medida $V(Y) - V(E(Y|X_{\sim i})) = (V(Y|X_{\sim i}))$ é a variância restante de Y que restaria, em média, se fosse possível determinar os valores verdadeiros de $X_{\sim i}$. A média é calculada sobre todas as combinações possíveis de $X_{\sim i}$, uma vez que $X_{\sim i}$ são fatores incertos e seus “valores verdadeiros” são desconhecidos. Dividindo por $V(Y)$, obtém-se o índice de efeito total para X_i :

$$S_{Ti} = \frac{E[V(Y|X_{\sim i})]}{V(Y)} = 1 - \frac{V[E(Y|X_{\sim i})]}{V(Y)} \quad (51)$$

Existem técnicas que permitem estimar índices totais com o mesmo custo dos índices de primeira ordem, contornando assim o problema da dimensionalidade. Costuma-se calcular o conjunto de todos os S_i mais o conjunto de todos os S_{Ti} para obter uma descrição bastante boa das sensibilidades do modelo a um custo razoável.

De acordo com [25], foi argumentado em uma série de trabalhos ([20] e suas referências) que uma caracterização boa, sintética, embora não exaustiva, do padrão de sensibilidade para um modelo com k fatores é dada pelo conjunto total de termos de primeira ordem mais os efeitos totais. Por exemplo, para um sistema com 10 fatores são necessários 20 termos em vez de 1023.

2.2.2.5 Computando os índices de sensibilidade

Nesta seção é descrito o procedimento numérico baseado em Monte Carlo para calcular o conjunto completo de índices de primeira ordem e de efeito total para um modelo de k fatores.

O método descrito a seguir é atribuível a [27] e representa uma extensão da abordagem original fornecida por [28] e [29].

1. Gera-se uma matriz de dimensão $N \times 2 \cdot k$ e define-se duas matrizes A e B , cada uma contendo metade da amostra, onde k é o número de fatores de entrada e N é chamado de amostra base. Para dar ordem de grandeza, N pode variar de algumas centenas a alguns milhares. Sobol recomenda o uso de sequências de números quase aleatórios.

$$A = \begin{bmatrix} x_1^{(1)} & x_2^{(1)} & \dots & x_i^{(1)} & \dots & x_k^{(1)} \\ x_1^{(2)} & x_2^{(2)} & \dots & x_i^{(2)} & \dots & x_k^{(2)} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ x_1^{(N-1)} & x_2^{(N-1)} & \dots & x_i^{(N-1)} & \dots & x_k^{(N-1)} \\ x_1^{(N)} & x_2^{(N)} & \dots & x_i^{(N)} & \dots & x_k^{(N)} \end{bmatrix} \quad (52)$$

$$B = \begin{bmatrix} x_{k+1}^{(1)} & x_{k+2}^{(1)} & \dots & x_{k+i}^{(1)} & \dots & x_{2 \cdot k}^{(1)} \\ x_{k+1}^{(2)} & x_{k+2}^{(2)} & \dots & x_{k+i}^{(2)} & \dots & x_{2 \cdot k}^{(2)} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ x_{k+1}^{(N-1)} & x_{k+2}^{(N-1)} & \dots & x_{k+i}^{(N-1)} & \dots & x_{2 \cdot k}^{(N-1)} \\ x_{k+1}^{(N)} & x_{k+2}^{(N)} & \dots & x_{k+i}^{(N)} & \dots & x_{2 \cdot k}^{(N)} \end{bmatrix} \quad (53)$$

2. Define-se uma matriz C_i , formada pelas colunas de B , exceto a i -ésima que é tomada de A :

$$C_i = \begin{bmatrix} x_{k+1}^{(1)} & x_{k+2}^{(1)} & \dots & x_i^{(1)} & \dots & x_{2 \cdot k}^{(1)} \\ x_{k+1}^{(2)} & x_{k+2}^{(2)} & \dots & x_i^{(2)} & \dots & x_{2 \cdot k}^{(2)} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ x_{k+1}^{(N-1)} & x_{k+2}^{(N-1)} & \dots & x_i^{(N-1)} & \dots & x_{2 \cdot k}^{(N-1)} \\ x_{k+1}^{(N)} & x_{k+2}^{(N)} & \dots & x_i^{(N)} & \dots & x_{2 \cdot k}^{(N)} \end{bmatrix} \quad (54)$$

3. Calcula-se a saída do modelo para todos os valores de entrada nas matrizes de amostra A , B e C_i obtendo três vetores de saídas do modelo de dimensão $N \times 1$, tais que:

$$y_A = f(A), y_B = f(B), y_{C_i} = f(C_i) \quad (55)$$

Pode-se antecipar que os vetores das Eq. (55) são tudo o que é preciso para calcular os índices de primeiro efeito e de efeito total S_i e S_{Ti} , para um determinado fator X_i . Como existem k fatores, o custo desta abordagem é de $N + N$ execuções do modelo para as matrizes A , B mais k vezes N para estimar k vezes o vetor de saída correspondente à matriz C_i . O custo total é, portanto, $N \cdot (k + 2)$.

4. O índice de primeira ordem pode ser calculado de acordo com a Eq. (56):

$$S_i = \frac{y_A \cdot y_{C_i} - f_0^2}{y_A \cdot y_A - f_0^2} \quad (56)$$

onde:

$$f_0^2 = \left(\frac{1}{N} \sum_{j=1}^N y_A^{(j)} \right)^2 \quad (57)$$

5. Similarmente, calcula-se os índices de efeito total de acordo com a Eq. (58):

$$S_{Ti} = 1 - \frac{y_B \cdot y_{C_i} - f_0^2}{y_A \cdot y_A - f_0^2} \quad (58)$$

A explicação do motivo das equações acima funcionarem pode ser encontrada em [25], bem como sugestões de melhorias e estimativa do erro provável – que depende de N .

Outros métodos para o cálculo dos índices são discutidos em [30].

2.2.2.6 Diferentes testes para diferentes cenários

É comum encontrar na literatura casos em que diferentes testes de sensibilidade são aplicados ao mesmo problema de forma não estruturada. Esta prática pode produzir uma variedade de resultados – por exemplo em termos de classificação dos fatores por ordem de importância – sem qualquer orientação sobre em que se deve acreditar ou privilegiar. Sugere-se, em vez disso, uma consideração cuidadosa de (a) o produto de interesse e (b) o conceito de “importância”, tal como se aplica ao problema em questão. Isto permitiria, em geral, a identificação do cenário mais adequado para um determinado problema e, por sua vez, o teste de sensibilidade a ser aplicado. Uma lista de cenários possíveis [20] é fornecida nos subitens a seguir.

Ainda cabe ressaltar que a utilidade das medidas de sensibilidade baseadas na variância deriva da sua ampla gama de aplicações. Dos quatro cenários descritos a seguir, os três primeiros são suscetíveis de análise baseada na variância.

2.2.2.6.1 Priorização de Fatores

Priorização de Fatores é usada para identificar um fator (ou grupo de fatores) que, quando fixado em seu valor real, leva à maior redução na variância do resultado, $V(Y)$. Por outras palavras, o fator identificado (ou grupo de fatores) é aquele que representa a maior parte da variância do produto. Portanto, esta configuração nos permite:

- Detectar e classificar aqueles fatores que precisam ser mais bem medidos para reduzir a variância da saída;
- Detectar os fatores que têm melhor chance de serem estimados em um subsequente processo de estimação numérica ou experimental.

Este último ponto é particularmente interessante porque o analista pode identificar os fatores a serem estimados antes de qualquer estimativa ser feita ou medições.

Um bom candidato ao título de “fator mais influente” seria aquele fator que, em média, uma vez fixado, causaria a maior redução na variância. “Em média”, neste caso, significa que

devemos calcular a média da fixação do fator sobre a distribuição do próprio fator. É fácil ver que neste cenário $E_{X_i}(V_{X_{\sim i}}(Y|X_i))$ é a medida a ser usada. Quanto menor $E_{X_i}(V_{X_{\sim i}}(Y|X_i))$ e, portanto, quanto maior $V_{X_i}(E_{X_{\sim i}}(Y|X_i))$, mais provável é que o fator X_i seja o fator que se deve medir primeiro para reduzir ao máximo a variância. Assim, vincula-se o cenário de priorização de fatores a uma medida, o índice de sensibilidade de primeira ordem S_i . Isto é uma aposta, pois não se conhece a posição do verdadeiro valor de um fator sobre sua distribuição, ou seja, um determinado fator ainda poderia superar a estimativa baseada na análise de sensibilidade e reduzir a variância além do imaginado, ou reduzir a variância usando um fator diferente daquele que foi identificado por meio da análise de sensibilidade.

2.2.2.6.2 Fixação de Fatores

Fixação de Fatores é usada para identificar fatores no modelo que, deixados livres para variar em sua faixa de incerteza, não contribuem significativamente para a variância do resultado. Os fatores identificados podem então ser fixados em qualquer valor dentro de sua faixa de variação sem afetar a variação da saída. Esta análise pode ser realizada em grupos de fatores, especialmente para modelos grandes, para identificar subconjuntos de fatores não influentes. Às vezes, os fatores são configurados para representar estruturas alternativas para os componentes do modelo (por exemplo, simples versus complexo) e muitas vezes podem ser alcançadas simplificações significativas do modelo quando esses fatores são considerados não influentes.

Das discussões nos subcapítulos anteriores ficará claro que $S_i = 0$ é uma condição necessária, mas insuficiente para fixar o fator X_i , pois este fator pode estar envolvido em interações com outros fatores de modo que, embora o seu termo de primeira ordem seja zero, pode haver termos de ordem superior diferentes de zero.

A condição $S_{Ti} = 0$ é necessária e suficiente para que X_i seja um fator não influente. Se $S_{Ti} \cong 0$, então X_i pode ser fixado em qualquer valor dentro da sua faixa de incerteza sem afetar sensivelmente o valor da variância de saída $V(Y)$. O erro de aproximação cometido quando esta simplificação do modelo é realizada depende do valor de S_{Ti} (ver [31]). Os índices totais são adequados para a configuração de fixação de fator.

2.2.2.6.3 Corte de Variância

Corte de Variância é usada para reduzir a variação de saída abaixo de uma determinada tolerância. Isto pode ser desejável na análise de risco ou de confiabilidade, onde o analista está interessado em certificar-se, por exemplo, de que a incerteza da confiabilidade de um determinado componente do sistema está abaixo de uma determinada tolerância. Neste cenário, o analista pretende garantir que a incerteza seja reduzida a um determinado valor, atuando sobre o menor número possível de fatores.

2.2.2.6.4 Mapeamento de Fatores

Mapeamento de Fatores é usada para estudar quais valores dos fatores de entrada levam à realização do modelo em um determinado intervalo do espaço de saída. Por exemplo, pode-se querer destacar as realizações do modelo que ficam acima do percentil 95 porque correspondem a condições de risco numa planta industrial ou a uma perda financeira considerável. Neste cenário, investiga-se qual combinação de fatores leva às realizações em análise.

2.2.2.7 Conclusão

Os métodos baseados em variância são quantitativos, ou seja, são capazes de quantificar o quão um determinado fator é mais importante que outro. Neste tipo de método, pode ser fornecida, por exemplo, a porcentagem exata da variância de saída devido a cada fator de entrada, ou grupo de fatores. No entanto, os métodos quantitativos são computacionalmente mais caros, quando comparados aos métodos qualitativos, e inviável quando há vários fatores de entrada em sua análise ou o modelo é complexo [20].

Com relação aos índices de sensibilidade, recorda-se as seguintes propriedades:

- Qualquer que seja a força das interações no modelo, S_i indica quanto se poderia reduzir, em média, a variância do produto se X_i pudesse ser fixado; portanto, é uma medida do efeito principal;
- Quaisquer que sejam as interações no modelo, $S_{i_1, i_2, \dots, i_s}^c$ indica o quanto a variância poderia ser reduzida, em média, se fosse possível fixar $X_{i_1}, X_{i_2}, \dots, X_{i_s}$. Lembremos que c denota o efeito conjunto;

- Por definição, S_{Ti} é maior que S_i , ou igual a S_i no caso de X_i não estar envolvido em nenhuma interação com outros fatores de entrada. A diferença $S_{Ti} - S_i$ é uma medida de quanto X_i está envolvido nas interações com qualquer outro fator de entrada;
- $S_{Ti} = 0$ implica que X_i não é influente e pode ser fixado em qualquer lugar de sua distribuição sem afetar a variância do produto;

2.3 Modelagem computacional em estudos de TEM

Uma representação precisa de um componente elétrico é essencial para uma análise confiável do TEM. A simulação computacional de fenômenos transitórios pode exigir uma representação de componentes de rede válida somente para uma determinada faixa de frequência do fenômeno [32]. Por exemplo, em estudos de sobretensão devido à manobra de equipamentos de pátio da subestação, a modelagem do transformador de potencial capacitivo pode ser negligenciada. Já para estudos de sobretensão devido à descarga atmosférica, este equipamento é modelado como um capacitor de acordo com [33].

Em alguns casos, mesmo se a versão de banda larga estiver disponível, ela pode exibir ineficiência computacional ou exigir dados muito complexos [32].

A modelagem de componentes elétricos que leva em consideração a dependência de frequência dos parâmetros pode ser atualmente obtida por meio de modelos matemáticos que são precisos o suficiente para uma faixa específica de frequências. Cada faixa de frequências geralmente corresponde a algum fenômeno transitório específico [32]. A Figura 12 ilustra a classificação do fenômeno pela norma IEC (do inglês *International Electrotechnical Commission*) [34] de acordo com suas características.

Classe	Baixa frequência		Transitório		
	Contínua	Temporária	Frente lenta	Frente rápida	Frente muito rápida
Formas de onda das tensões ou sobretensões					
Faixas de formas de onda das tensões ou sobretensões	$f = 50 \text{ Hz}$ or 60 Hz $T_t \geq 3\,600 \text{ s}$	$10 \text{ Hz} < f < 500 \text{ Hz}$ $0,03 \text{ s} \leq T_t \leq 3\,600 \text{ s}$	$20 \mu\text{s} < T_p \leq 5\,000 \mu\text{s}$ $T_2 \leq 20 \text{ ms}$	$0,1 \mu\text{s} < T_1 \leq 20 \mu\text{s}$ $T_2 \leq 300 \mu\text{s}$	$3 \text{ ns} < T_f \leq 100 \text{ ns}$ $0,3 \text{ MHz} < f_1 < 100 \text{ MHz}$ $30 \text{ kHz} < f_2 < 300 \text{ kHz}$
Formas de onda normalizada das tensões					1)
Ensaio normalizado de tensão suportável	1)	Ensaio de frequência fundamental de curta duração	Ensaio de impulso de manobra	Ensaio de impulso atmosférico	1)
1) A ser especificado pela norma do respectivo equipamento.					

Figura 12 – Classes e formas de onda de sobretensões - ensaios normalizados de tensão suportável. Fonte: [34].

Em [33] pode-se encontrar uma visão geral sobre as várias origens dos TEM e suas faixas de frequência mais comuns. Por exemplo, é descrito em [33] que a faixa de frequência dos fenômenos gerados devido a descarga atmosférica pode variar de 10 kHz até 3 MHz. Neste caso, [33] classifica o fenômeno como sendo do tipo “frente rápida”.

A simulação de um fenômeno transitório implica não somente na seleção de modelos, mas na seleção da área do sistema que deve ser representada. Algumas regras a serem consideradas na simulação de TEM ao selecionar modelos e a área do sistema podem ser resumidas a seguir [35]:

- Selecione a zona do sistema levando em consideração a faixa de frequência dos transitórios; quanto maiores as frequências, menor a zona modelada.

- Minimize a parte do sistema a ser representada. Um número maior de componentes não significa necessariamente maior precisão, pois pode haver uma probabilidade maior de modelagem insuficiente ou errada. Além disso, uma representação muito detalhada de um sistema geralmente exigirá um tempo de simulação maior.
- Implemente uma representação adequada das perdas. Como seu efeito nas tensões máximas e frequências de oscilação é limitado, elas não desempenham um papel crítico em muitos casos. Existem, no entanto, alguns casos (por exemplo, ferorressonância ou comutação de banco de capacitores) para os quais as perdas são críticas para definir a magnitude das sobretensões.
- Considere uma representação idealizada de alguns componentes se o sistema a ser simulado for muito complexo. Tal representação facilitará a edição do arquivo de dados e simplificará a análise dos resultados da simulação.
- Realize um estudo de sensibilidade se um ou vários parâmetros não puderem ser determinados com precisão. Os resultados derivados desse estudo mostrarão quais parâmetros são preocupantes.

A determinação ou estimação dos parâmetros do TEM é, provavelmente, a tarefa mais difícil e demorada de muitos estudos [32]. Já com relação aos parâmetros da simulação, deve-se definir o tempo total da simulação e o passo de integração. É importante verificar, para cada tipo de fenômeno estudado, se o passo de integração escolhido está adequado, em função da faixa de frequências envolvidas no fenômeno em análise e das características dos componentes modelados [2].

As versões existentes do software EMTP não fornecem ao usuário uma estimativa ótima do tamanho do passo de integração a ser utilizado para um determinado caso de simulação. Cabe ao usuário escolher e ajustá-lo adequadamente. A mesma situação acontece para o tempo máximo de simulação [36]. Em [36] há uma discussão a respeito destes parâmetros, bem como um método proposto para estimá-los.

Propõe-se em [34] utilizar o valor do passo de integração de um décimo do período correspondente à maior frequência prevista no estudo e ressalta-se que o valor do passo de integração tem que ser menor que o tempo de propagação dos elementos da rede.

3 Metodologia

Este capítulo é dedicado para a apresentação da metodologia empregada neste trabalho, onde é relacionada a fundamentação teórica e os estudos de caso propostos utilizando análise paramétrica.

Para possibilitar a criação de um estudo de caso, inicialmente, é necessário conhecer a teoria clássica acerca de transitórios eletromagnéticos, como: modelagem de circuitos elétricos por meio de resistências, indutâncias e capacitâncias; leis de Kirchhoff de tensão e corrente; análise no domínio do tempo; equações diferenciais, métodos de integração numérica etc.

Além dos conceitos acerca de transitórios eletromagnéticos, é fundamental entender como estes conceitos são implementados em programas computacionais. Desta forma, recomenda-se a leitura de [37], onde é descrito um método computacional genérico para solucionar sistemas mono ou polifásicos submetidos a transitórios eletromagnéticos.

De maneira resumida, em uma simulação de transitórios eletromagnéticos, os equipamentos são modelados de acordo com suas respectivas equações, que podem ou não envolver equações diferenciais. Além disso, a representação de determinado equipamento pode variar em função do tipo de fenômeno a ser estudado.

Após a modelagem dos equipamentos, aplica-se as leis de Kirchhoff a fim de construir as equações da rede elétrica como um todo. Então, é necessário escolher a melhor forma de solucioná-las. Em [37], é proposta uma sequência de “fotografias” do transitório eletromagnético em intervalos discretos e utiliza-se o método de integração trapezoidal para integração das equações diferenciais acerca de indutâncias e capacitâncias concentradas, por exemplo.

Ao solucionar as equações da rede elétrica, consegue-se verificar o seu comportamento ao longo do tempo, sendo que o tempo total de observação depende do fenômeno a ser estudado.

Uma vez compreendido os conceitos supracitados, é necessário definir quais tópicos, acerca de análise paramétrica, serão explorados. Ressaltam-se os seguintes tópicos:

- Simulação com incertezas de parâmetros de entrada utilizando o método de Monte Carlo;
- Análise de sensibilidade (a fim de avaliar como a incerteza nos parâmetros de entrada influência no resultado), onde é possível, por exemplo:
 - Detectar fatores que precisam ser mais bem medidos para reduzir a variância de saída;

- Identificar fatores que, deixados livres para variar em sua faixa de incerteza, não contribuem significativamente para a variância de saída. Então, estes fatores podem ser fixados em qualquer valor dentro de sua faixa de variação);
- Realizar corte de variância por meio da intervenção no menor número possível de fatores.
- Estudos de otimização (a fim de determinar os parâmetros de entrada para atingir um resultado pré-determinado);
- Calibração de modelo (a fim de ajustar os parâmetros de entrada para aproximar o resultado do modelo ao resultado esperado – aquele obtido por meio de medições de campo, por exemplo).

Posteriormente à definição do sistema elétrico sob estudo e o(s) tópico(s) acerca de análise paramétrica a ser(em) explorado(s), é necessário definir os parâmetros de entrada com base no valor da grandeza e verificar se há incerteza acerca deste valor. Caso positivo, é necessário atribuir uma distribuição de probabilidade para o parâmetro em questão e, no caso desta distribuição ser desconhecida, é comum utilizar a distribuição uniforme para a caracterização.

Além disso, faz-se necessário a utilização de ferramentas computacionais. Neste trabalho, considerou-se a utilização de dois diferentes métodos: por meio do uso EMTP-RV em conjunto com PAMSUITE para a análise paramétrica; por meio de um *script* em Python para controlar a execução do ATP. Basicamente, descreve-se a segunda metodologia da seguinte forma: cria-se o circuito no ATPDraw e o executa, de forma a criar automaticamente o arquivo “.atp” deste circuito. Neste arquivo criado, localiza-se, manualmente, os parâmetros a serem variados. Após isso, insere-se no Excel os conjuntos de parâmetros, aleatoriamente gerados, a serem considerados no estudo. Por fim, elabora-se um conjunto de *scripts* de forma a conseguir:

- Interpretar os parâmetros de entrada presentes no Excel;
- Automatizar e controlar a execução do ATP considerando este conjunto de parâmetros;
- Coletar e armazenar os resultados de interesse.

O fluxograma apresentado na Figura 13 detalha o gerenciamento das simulações:

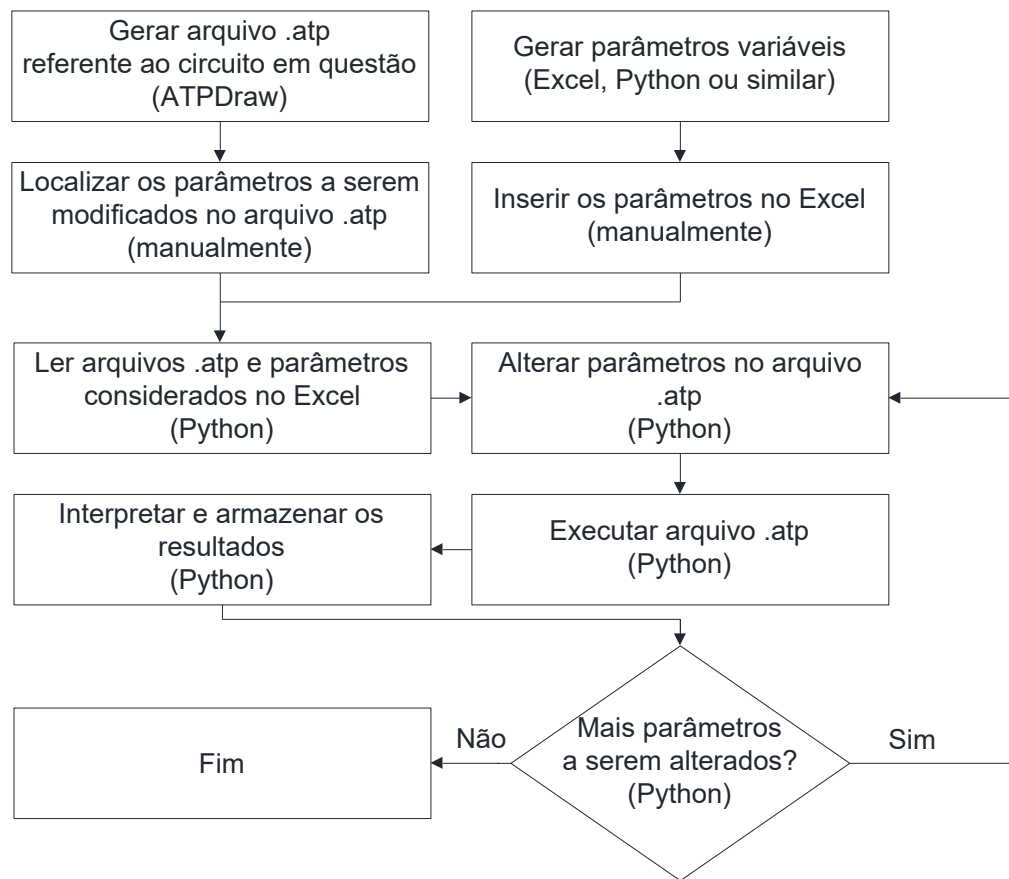


Figura 13 – Fluxograma do gerenciamento das simulações. Fonte: Autoria Própria.

Ainda, cabe ressaltar que se disponibilizou os *scripts* utilizados no Apêndice A – Rotina de automatização do Python para o estudo de caso 1.

Por fim, pode ser necessário o conhecimento em estatística para a interpretação do resultado de saída, a depender do estudo de caso considerado.

4 Estudos de Caso

Este capítulo é dedicado para a apresentação de estudos de caso referentes à aplicação de análise paramétrica em estudos de TEM. Nos estudos apresentados a seguir, faz-se uso do EMTP-RV em conjunto com PAMSUITE para a análise paramétrica. Para o Estudo de caso 1, para controlar a execução do ATP utilizou-se, adicionalmente, um *script* em Python.

4.1 Estudo de caso 1: Circuito RLC em série

É proposto, neste subcapítulo, a análise do comportamento da tensão em um capacitor (v_c) devido à aplicação de um degrau de tensão (V), no instante 0 s, em um circuito RLC série inicialmente desenergizado, conforme Figura 14:

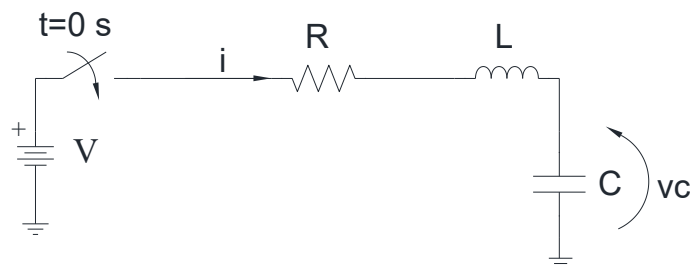


Figura 14 – Circuito RLC série. Fonte: Autoria Própria.

onde R , L e C são, respectivamente, os valores de resistência, indutância e capacitância.

O objetivo do estudo proposto é analisar a influência dos parâmetros V , R , L e C na tensão v_c .

Sabe-se que, pela lei das tensões de Kirchhoff, pode-se escrever a Eq. (59):

$$L \cdot \frac{di}{dt} + R \cdot i + v_c = V \quad (59)$$

Ainda, a solução $v_c(t)$ pode ter amortecimento:

- Forte: caso a condição (60) seja satisfeita;

- Fraco: caso a condição (61) seja satisfeita;
- Crítico: caso a condição (62) seja satisfeita.

$$\frac{R}{2 \cdot L} > \frac{1}{\sqrt{L \cdot C}} \quad (60)$$

$$\frac{R}{2 \cdot L} < \frac{1}{\sqrt{L \cdot C}} \quad (61)$$

$$\frac{R}{2 \cdot L} = \frac{1}{\sqrt{L \cdot C}} \quad (62)$$

A Figura 15 ilustra um exemplo de comportamento de $v_c(t)$ em função do tipo de amortecimento.

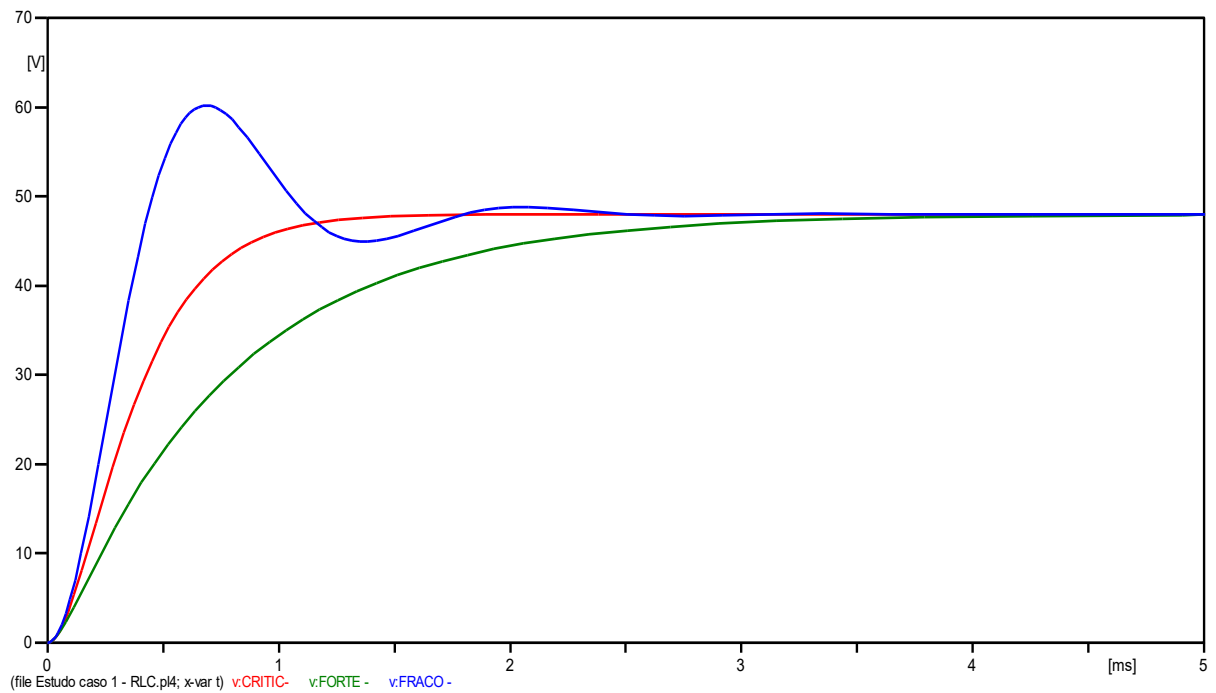


Figura 15 – Tipos de amortecimento em circuito RLC série. Fonte: Autoria Própria.

4.1.1 Modelagem computacional e parametrização

Propõe-se a realização da análise paramétrica por meio do estudo de propagação de incerteza pelo método de Monte Carlo e cálculo dos índices de sensibilidade por meio dos métodos Morris e Sobol.

Os valores considerados para o circuito RLC constam na Tabela 4. Estes valores foram arbitrariamente escolhidos, isto é, sem evidência experimental, com a única premissa de possibilitar a ocorrência dos três tipos de amortecimento supracitados.

Tabela 4– Parâmetros adotados para o circuito RLC. Fonte: Autoria Própria.

Parâmetro	Modelagem adotada
V	U(40;60) V
R	U(600;1400) Ω
L	U(60;140) mH
C	U(250;550) nF

Utilizou-se duas diferentes metodologias para a realização dos estudos supracitados. Na primeira, fez-se uso do conjunto EMTP-RV e PAMSUITE. Já na segunda, elaborou-se *scripts* em Python, responsáveis por controlar a execução do ATP, bem como calcular os índices de sensibilidade. Por fim, os resultados obtidos por meio de ambas as metodologias foram comparados e constam na próxima seção.

4.1.2 Resultados da simulação

Ao considerar, para o circuito representado pela Figura 15, 1000 conjuntos de parâmetros de entrada, gerados aleatoriamente e conforme a Tabela 4, obteve-se, para a tensão $v_c(t)$, os comportamentos ilustrados na Figura 16:

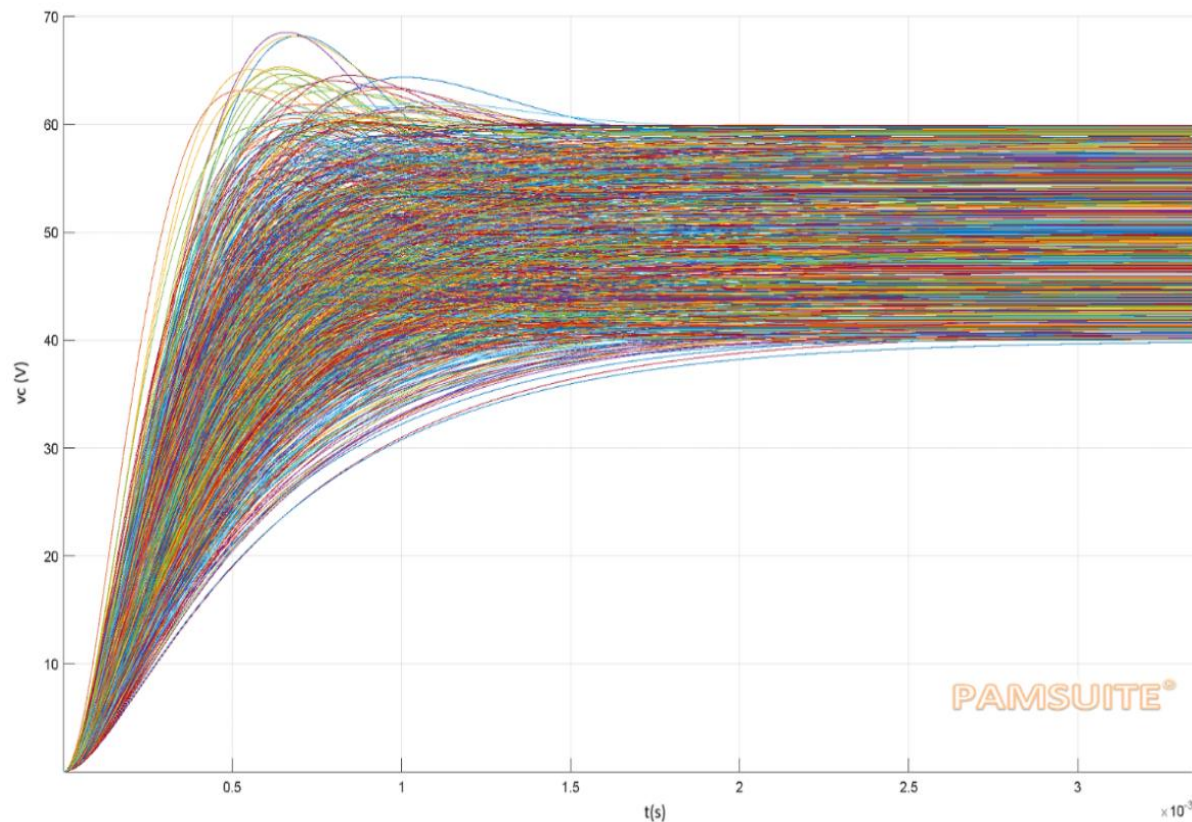
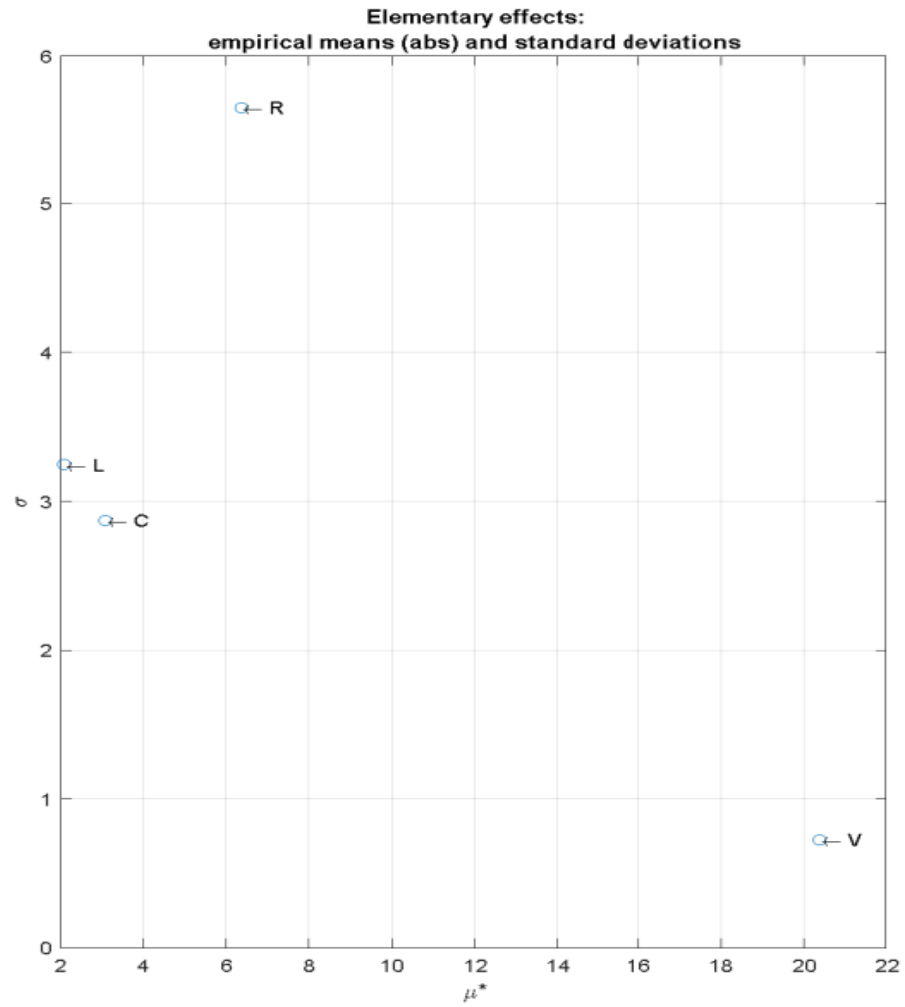
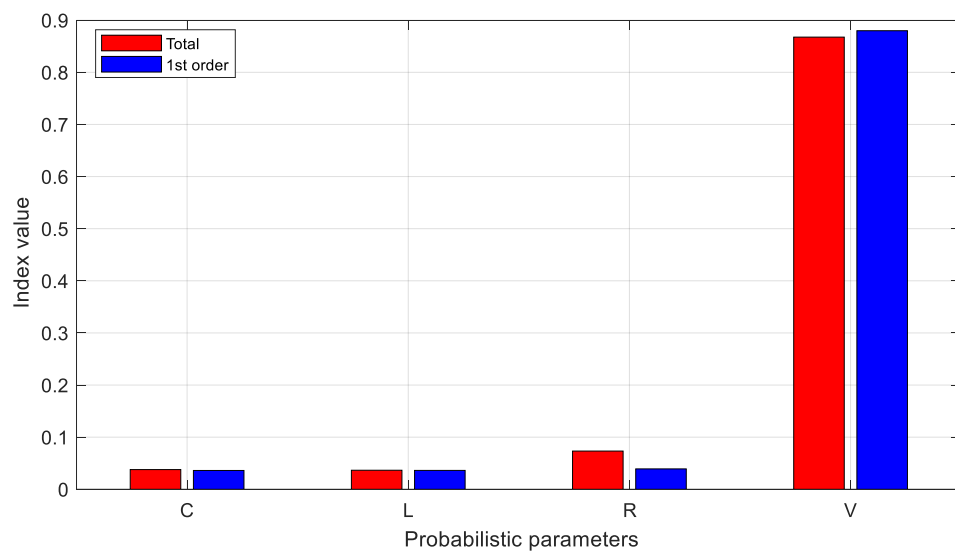


Figura 16 – Tensão versus tempo no capacitor para 1000 simulações. Fonte: Autoria Própria.

Para a análise de sensibilidade, gerou-se os índices Morris e Sobol (ver Figura 17), a fim de verificar quais variáveis possuem maior importância na tensão máxima de saída no capacitor.



(a)



(b)

Figura 17 – Análise de sensibilidade: (a) método de Morris, (b) método de Sobol sem intervalo de confiança.

Fonte: Autoria Própria.

Nota-se, por meio da Figura 16, que os casos simulados se diferenciam quanto aos amortecimentos para o circuito RLC. Já por meio da Figura 17, verifica-se que, dada as características dos parâmetros de entrada da Tabela 4, o fator de entrada mais influente para a tensão máxima de saída no capacitor é a tensão da fonte, V . Ainda, existem interações entre os parâmetros que influenciam no valor desta tensão.

Além disso, seria interessante, para os índices Sobol apresentados na Figura 17, a inclusão de intervalos de confiança, para uma análise mais precisa. Nota-se também que, para o índice Sobol relativo ao fator de entrada V , obteve-se $S_{T1} < S_1$, o que não é possível por definição e, provavelmente, ao considerar um número maior de amostras resolveria este problema.

Os resultados e gráficos apresentados anteriormente foram construídos utilizando o conjunto EMTP-RV e PAMSUITE.

Após a simulação pelo método de Monte Carlo, cujos resultados constam na Figura 16, exportou-se para o Excel os parâmetros de entrada aleatoriamente gerados pelo PAMSUITE, a fim de utilizá-los para a reprodução das simulações, no entanto, agora utilizando o conjunto ATP e Python. Após isso, comparou-se os resultados de tensão máxima no capacitor, para as 1000 simulações, entre as duas diferentes abordagens e, conforme esperado e exposto na Figura 18, os resultados são aparentemente idênticos:

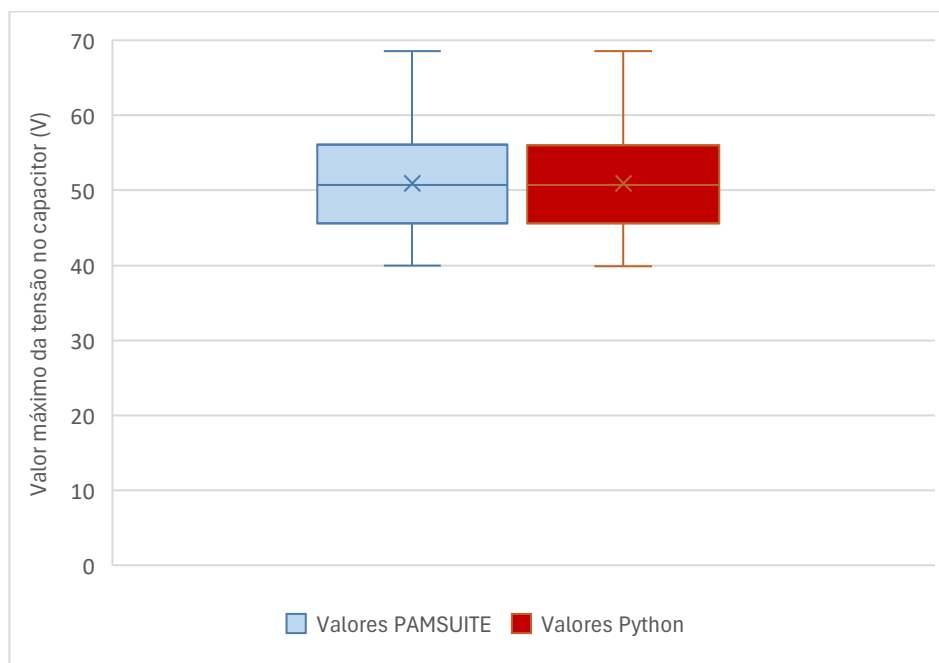


Figura 18 – Boxplot para comparativo da tensão máxima no capacitor para 1000 simulações. Fonte: Autoria Própria.

Com relação à análise de sensibilidade, não foi possível utilizar os mesmos parâmetros de entrada para as duas diferentes metodologias, visto que o PAMSUITE não fornece os dados de entrada utilizados para os cálculos dos índices Morris e Sobol, nem é possível determinar-se estes dados, visto que é gerado aleatoriamente no PAMSUITE por meio de uma rotina externa.

Portanto, para o cálculo dos índices Morris por meio do *script* em Python, gerou-se aleatoriamente e através do próprio *script*, os parâmetros de entrada conforme detalhado na seção de apêndices. Já para o cálculo dos índices Sobol, considerou-se a criação dos parâmetros de entrada pelo método quase Monte Carlo. Para esta criação, utilizou-se o módulo *Statistical functions* disponível para Python. A Figura 19 e a Figura 20 ilustram uma comparação visual dos resultados obtidos por meio das duas diferentes metodologias, considerando os conjuntos EMTP-RV e PAMSUITE e Python e ATP.

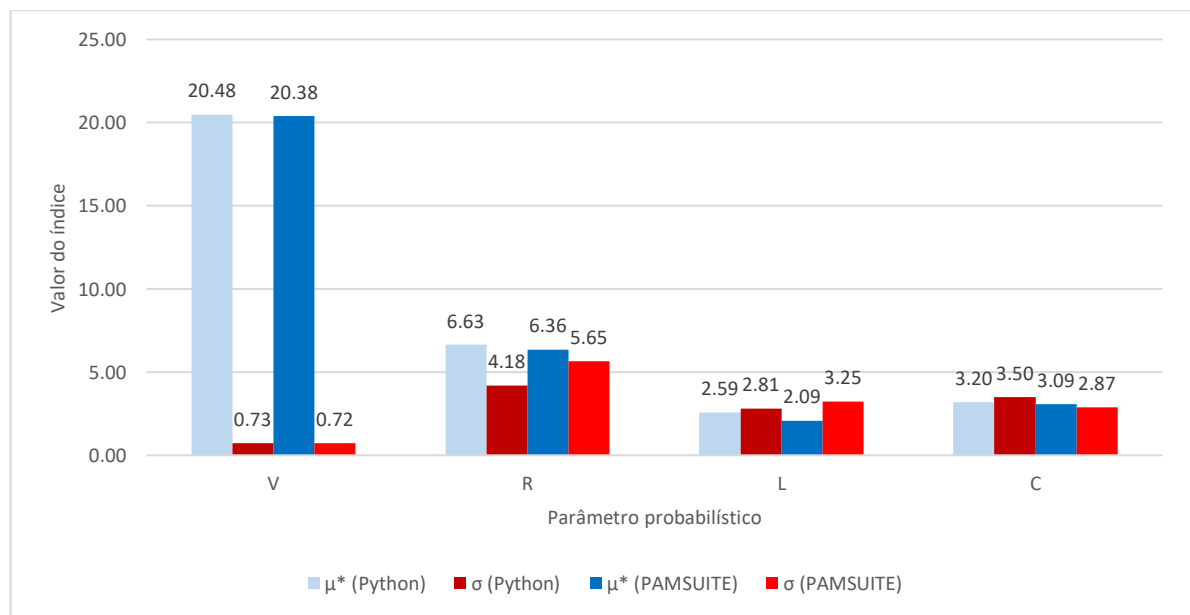


Figura 19 – Análise de sensibilidade: comparativo dos índices do método Morris para as duas diferentes metodologias. Fonte: Autoria Própria.

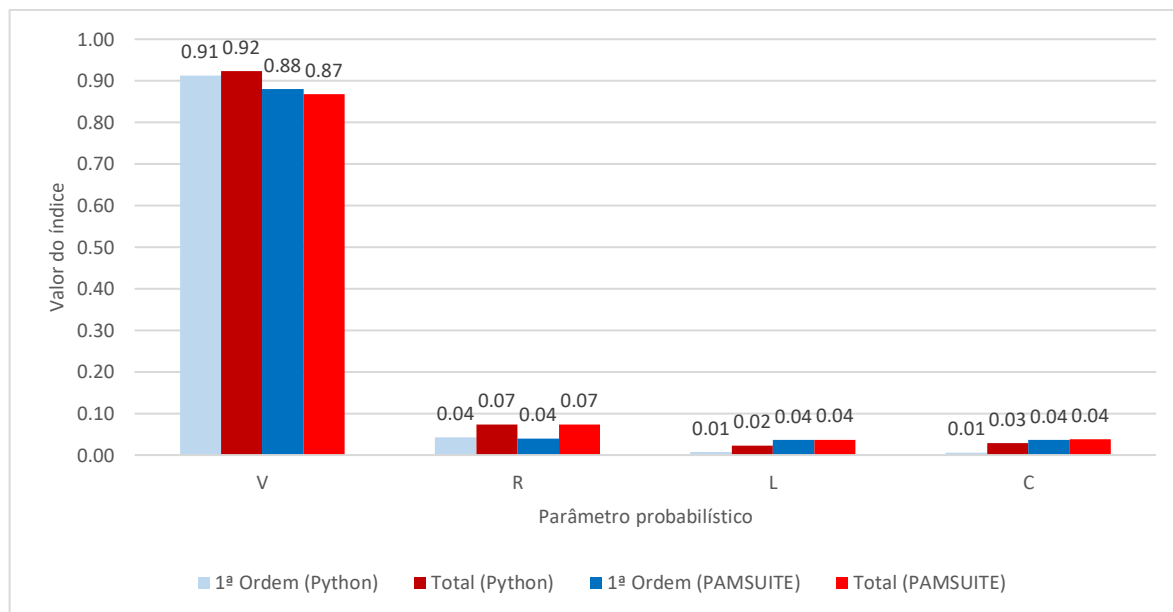


Figura 20 – Análise de sensibilidade: comparativo dos índices do método Sobol para as duas diferentes metodologias. Fonte: Autoria Própria.

Pode-se observar, por meio da Figura 19 e Figura 20, que os resultados obtidos pelas duas diferentes metodologias são aproximadamente iguais, conforme esperado. No entanto, cabe destacar que o ideal seria utilizar ferramentas estatísticas para validar esta aparente igualdade.

4.2 Estudo de caso 2: Sobretensões em subestação devido a descarga atmosférica

Sobretensões no sistema elétrico devido a descarga atmosférica podem ocorrer quando há a descarga direta ou indireta. Descargas diretas são aquelas que incidem sobre os componentes do sistema elétrico, enquanto as descargas indiretas caracterizam-se pela incidência nas proximidades dos sistemas elétricos.

Quanto às descargas diretas em uma linha de transmissão, as sobretensões podem ser causadas devido à descarga direta em um condutor de fase ou devido à descarga direta em um condutor de blindagem, e, neste último caso, se há ruptura do isolamento, caracteriza-se a ocorrência de *back-flashover*.

Neste subcapítulo, propõe-se estudar as sobretensões no setor de 230 kV de uma subestação, cujo arranjo físico consta no Apêndice B – Parâmetros e modelagem do estudo de caso 2 - sobretensões devido a descarga atmosférica. Ainda, propõe-se que estas sobretensões são causadas devido à ocorrência de *back-flashover*. Para uma fácil compreensão do problema

proposto, elaborou-se o diagrama unifilar simplificado da Figura 21, que representa o sistema de estudo:

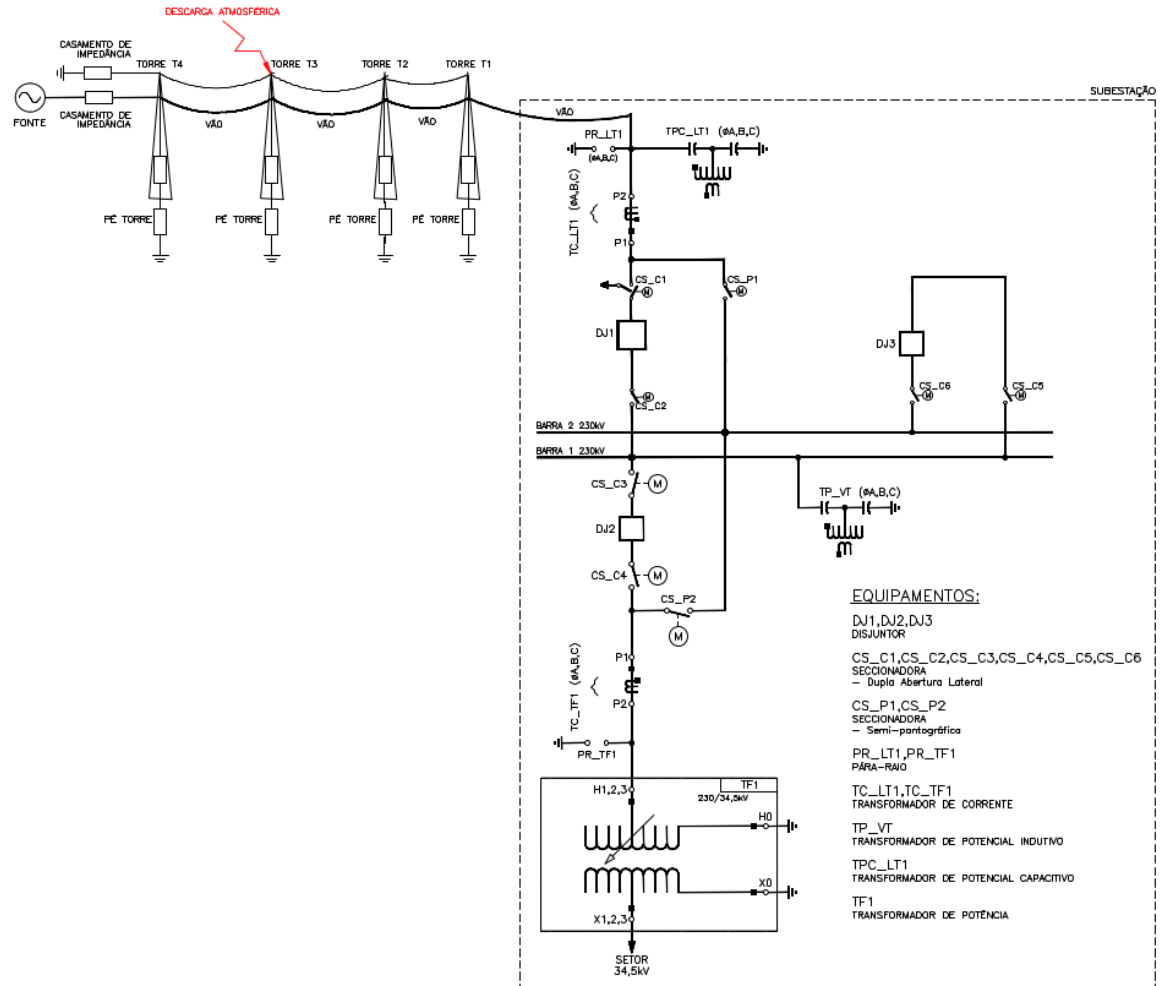


Figura 21 – Diagrama unifilar simplificado. Fonte: Autoria Própria.

4.2.1 Modelagem computacional e parametrização

A seguir, são descritos os critérios e considerações adotadas para a modelagem computacional do sistema sob estudo. Diretrizes para a modelagem deste tipo de estudo podem ser encontradas em [32, 34, 33, 2, 38].

Considerou-se a utilização do EMTP-RV - para o estudo do TEM – em conjunto com o módulo PAMSUITE – para a análise paramétrica.

4.2.1.1 Fonte de frequência fundamental

A tensão de frequência de fundamental em cada isolador adiciona ou subtrai continuamente das tensões criadas no isolador devido à descarga atmosférica, alterando assim continuamente a amplitude da corrente de elétrica necessária para causar o *flashover* [39].

Modelou-se a fonte de frequência fundamental utilizando-se uma fonte de tensão e um resistor de adaptação em série, cujo valor equivale à matriz de impedância de surto da linha de transmissão, conforme sugerido em [34]. Para a fonte de tensão, considerou-se uma fonte trifásica senoidal de tensão 230 kV.

4.2.1.2 Descarga atmosférica

A simplificação geralmente aceita na engenharia de subestações é a adoção de um surto de tensão que represente a onda de corrente propagável para o interior da subestação, devido ao surto atmosférico que atinge a linha de transmissão. No entanto, fisicamente, o surto atmosférico é caracterizado por um impulso de corrente. Desta forma, a sua representação mais precisa é modelar o impulso de corrente no programa de simulação, de acordo com as características de forma e amplitude [2].

Propõe-se em [34] modelar o impulso de corrente de duas formas: por meio da curva CIGRE ou por meio da curva de duas rampas. Escolheu-se, neste estudo, a curva de duas rampas, que embora seja menos precisa, possui o resultado mais conservador e é mais simples de modelar. A Figura 22 ilustra a curva de duas rampas, onde I_f , t_f e t_h são: o valor de crista da corrente de descarga, o tempo de frente e o tempo de meio valor, respectivamente.

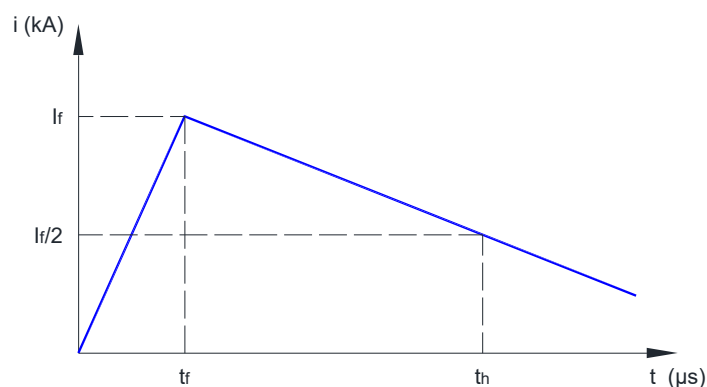


Figura 22 – Curva de duas rampas. Fonte: Adaptado de [34].

Com relação à distribuição de probabilidade, adotou-se um dos métodos expostos em [34], em que a corrente de pico é uma variável aleatória. No método adotado, a distribuição de probabilidade $f(i)$ é considerada log-normal e é dada pela Eq. (63):

$$f(i) = \frac{1}{\sqrt{2 \cdot \pi} \cdot \sigma \cdot i} \cdot e^{-\frac{1}{2} \left(\frac{\ln(i/\mu)}{\sigma} \right)^2} \quad (63)$$

onde μ e σ são definidos pelo usuário e dado em kA. Este método tem origem em um estudo detalhado em [40], onde são sugeridos os valores de 33,3 kA e 0,605 kA para μ e σ , respectivamente, para estudos de *back-flashover*. Estes valores têm como base 408 observações ao redor do mundo de primeira descarga do tipo polaridade negativa. É válido ressaltar que a grande maioria das descargas atmosféricas são do tipo polaridade negativa [40].

Considerou-se, no estudo presente neste subcapítulo, a descarga de polaridade negativa e com os parâmetros descritos anteriormente. No entanto, a fim de estudar apenas os casos em que há a ocorrência do *back-flashover* e simplificar a fórmula da distribuição de probabilidade da Eq. (63), realizou-se o procedimento a seguir:

- Gerou-se, por meio do *software* estatístico Minitab, 408 valores respeitando uma distribuição log-normal e com os parâmetros descritos anteriormente;
- Dos valores gerados, selecionou-se aqueles maiores que 110 kA. Utilizou-se estes dados como valores de entrada da ferramenta *Goodness of Fit Test*, do Minitab, a fim de determinar a distribuição de probabilidade que melhor se aproxima do conjunto de dados selecionados;
- De acordo com o resultado da ferramenta supracitada, foi possível aproximar os dados selecionados da distribuição log-normal para uma distribuição exponencial, cuja equação aproximada é apresentada na Eq. (64)

$$f(i) = \frac{1}{47} \cdot e^{-\frac{(i-110)}{47}} \quad (64)$$

- Para a Eq. (64), considerou-se que $110 \text{ kA} \leq i \leq 200 \text{ kA}$, de tal forma que nunca será possível obter um valor de corrente de pico maior que 200 kA, pois no estudo realizado em [40], nenhum valor medido de corrente de descarga ultrapassou 200 kA.

Ainda, cabe ressaltar que a Eq. (64) foi utilizada para fins de simplificação da Eq. (63), mas o ideal seria utilizar diretamente a Eq. (63).

Com relação a t_h , considerou-se o valor de 77,5 μ s, que é proposto em [34] para estudos de sobretensões devido a descarga atmosférica, enquanto que para t_f considerou-se a Eq. (65), que considera o valor médio de t_f em função da corrente de pico, para valores maiores que 20 kA [34]:

$$t_f = 0,154 \cdot I_f^{0,624} \quad (65)$$

A Figura 23 ilustra um exemplo da curva de duas rampas modelada com as características supracitadas.

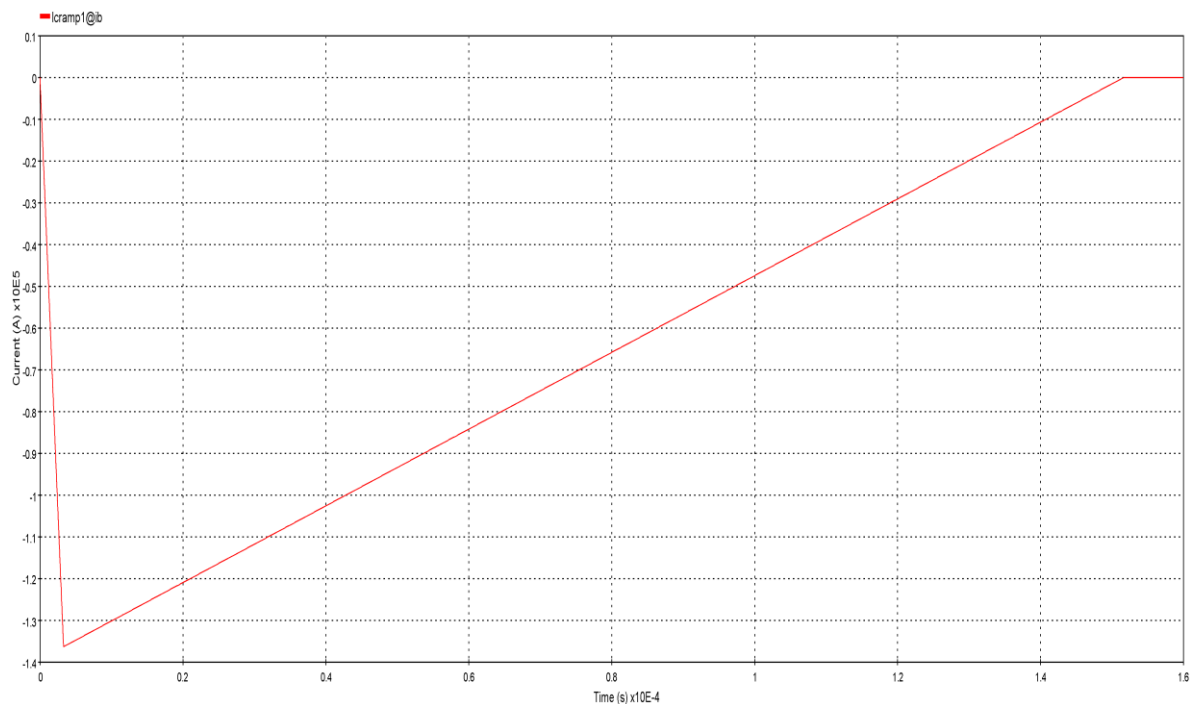


Figura 23 – Exemplo de curva de duas rampas no EMTP-RV. Fonte: Autoria Própria.

Após as simulações realizadas no PAMSUITE, exportou-se os dados de entrada e saída utilizados na simulação para Excel. Após isso, gerou-se os dados da Figura 24, que ilustra o histograma do módulo da corrente de pico da descarga atmosférica.

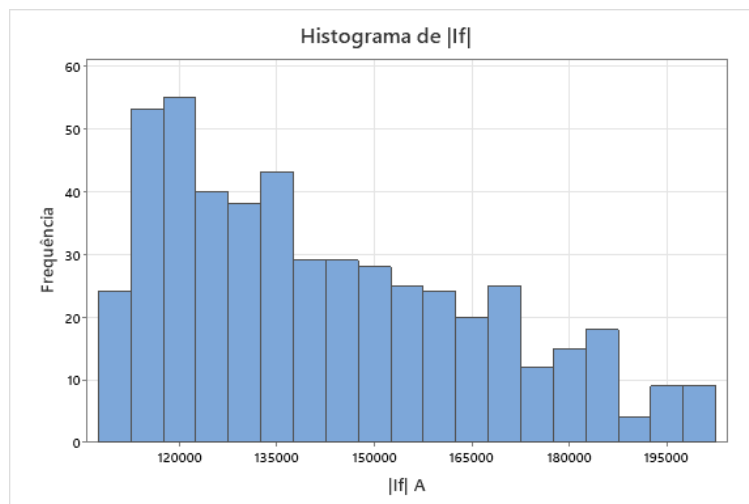


Figura 24 – Histograma do módulo da corrente de pico da descarga atmosférica. Fonte: Autoria Própria.

4.2.1.3 Linha de transmissão, torres metálicas e cadeia de isoladores

O cálculo de sobretensões causadas por descarga atmosférica requer um modelo mais detalhado, no qual torres, impedâncias dos pés das torres, isoladores, condutores de fase e de blindagem são representados. Como a descarga atmosférica é um fenômeno transitório de frente rápida, um modelo multifásico com parâmetros distribuídos, incluindo assimetria do condutor e efeito corona, é necessário para a representação de cada vão [32]. No entanto, no presente estudo desconsiderou-se o efeito corona, uma vez que este efeito tende a reduzir o declive da onda de impacto, e é uma aproximação conservativa negligenciá-lo [34].

De acordo com [38], o valor do pico de tensão em uma torre atingida por uma descarga atmosférica pode ser influenciado pelas reflexões das torres adjacentes. O número suficiente de torres adjacentes a serem consideradas para um estudo preciso pode ser determinado selecionando um número de vãos de linha tal que o tempo de propagação entre a torre atingida e a torre mais distante é maior que a metade de t_f .

Expõe-se em [34] as representações comuns para modelagem de torres. Dentre elas, adotou-se, no presente estudo, a representação da torre por meio de uma linha de propagação sem perdas com parâmetros constantes. A impedância de surto da torre depende dos detalhes da estrutura [38]. Considerou-se que a impedância de surto da torre pode variar de 100 a 300 Ω , e a velocidade de propagação igual à velocidade da luz, conforme valores típicos expostos em [38].

Com relação ao pé da torre, fenômenos físicos envolvidos quando a corrente está circulando através de um eletrodo de aterramento são bastante complexos, exceto para

configurações idealistas simplificadas. Portanto, é bastante difícil representar com precisão ao mesmo tempo a propagação e a ionização. O domínio de validade dos modelos é difícil de avaliar [34]. Adotou-se um dos modelos expostos em [34], considerando apenas uma resistência.

Para o processo de *flashover*, considerou-se uma chave ideal, normalmente aberta e que fecha após a condição de arco ser atingida. Se o passo de integração adotado na simulação não for muito pequeno, este modelo é adequado [34]. Quanto à condição de ocorrência do arco, alguns modelos de suportabilidade a impulso atmosférico são propostos em [32] e é constatado que nenhuma abordagem única pode ser recomendada para todas as aplicações. Ainda, a suposição de que o *flashover* ocorrerá assim que a tensão atinge um determinado valor não é recomendada para o este tipo de estudo [34].

Entretanto, adotou-se, para fins de simplificação, que a condição de arco é obtida quando a tensão entre os terminais da cadeia de isoladores atinge um determinado valor. Considerou-se que este valor é aleatoriamente obtido por meio da distribuição normal de média igual ao valor do CFO (do inglês *critical flashover voltage*) e desvio padrão σ igual a 3% do CFO. Ainda, considerou-se o valor de NBI (nível básico de isolamento ao impulso atmosférico) normalizado [41] de 850 kV e, tendo este valor como a tensão suportável estatística conforme definição em [41], bem como sua probabilidade de falha igual a 10%, pode-se relacionar o NBI com o CFO da seguinte forma Eq. (66) [32]:

$$NBI = CFO \cdot \left(1 - 1,28 \cdot \frac{\sigma}{CFO}\right) \quad (66)$$

Além disso, a distribuição gaussiana é ilimitada à direita e à esquerda; ou seja, é definida entre $+\infty$ e $-\infty$. Um limite de $-\infty$ indica que existe uma probabilidade de flashover para uma tensão igual a zero, o que é fisicamente impossível [32]. Portanto, considerou-se que a faixa de tensão de ruptura considera quatro desvios-padrão.

Uma vez modelada a chave ideal conforme características descritas anteriormente, pode-se considerar um capacitor em paralelo com esta chave, de tal forma a representar os efeitos de acoplamento entre o condutor e a estrutura [38]. Valores típicos de capacitância para isoladores podem ser encontrados em catálogos, bem como em publicações como [38, 42].

Para os condutores da linha de transmissão, utilizou-se a rotina *Line/Cable Data*, do EMTP-RV para a determinação dos parâmetros, usando a geometria da torre e dados dos condutores. Considerou-se o modelo de parâmetros constantes, calculados na frequência de 500

kHz e efeito pelicular, como são geralmente calculados [38]. Considerou-se a resistividade elétrica do solo igual a $1000 \Omega \cdot m$.

Dispõe-se, na nota técnica [43], parâmetros elétricos, dentre outras informações para o planejamento, de configurações típicas de linhas de transmissão aéreas, em corrente alternada, nos níveis de tensão nominal de 69, 138, 230, 345, 440, 500 e 765 kV. Escolheu-se, arbitrariamente, uma destas configurações típicas para a tensão nominal de 230 kV, que é ilustrada na Figura 25.

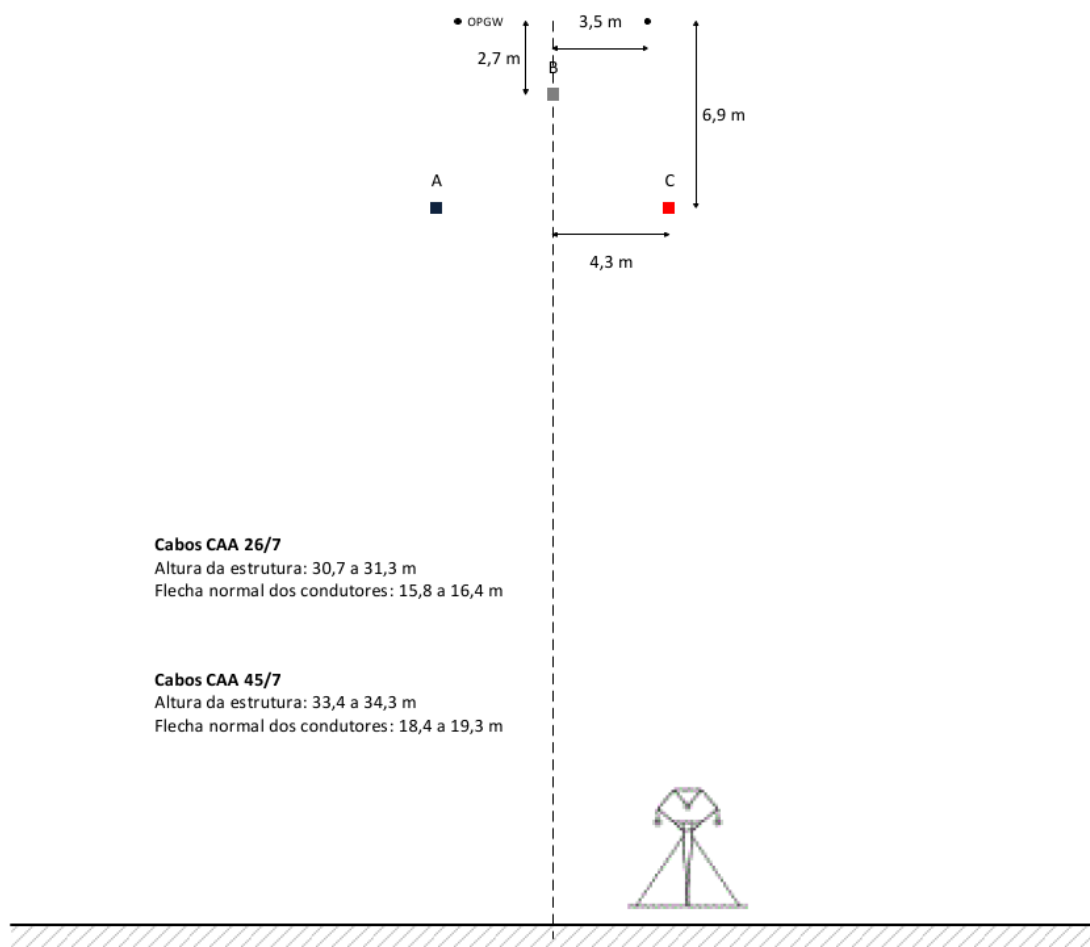


Figura 25 – Geometria considerada para a torre. Fonte: Adaptado [43].

onde OPGW (do inglês *optical ground wire*) refere-se ao cabo para-raios que possui fibra óptica e CAA refere-se aos condutores de alumínio com alma de aço. Os valores 26/7 e 45/7 referem-se à formação do cabo CAA.

Quanto às características dos cabos, os dois cabos superiores na Figura 25 são os cabos para-raios, enquanto os outros três cabos são os condutores de fase. No estudo realizado por [43], considerou-se, para os cabos para-raios, o cabo OPGW 13,3 mm e o cabo convencional tipo EAR (Extra-Alta Resistência) 3/8", classe A. Esses cabos são comumente utilizados na prática [43]. Já para os três cabos de fase, considerou-se, no presente estudo, um feixe por fase e condutor *bluejay*. As características dos cabos para-raios e do condutor *bluejay*, como resistência em corrente contínua e raio externo, podem ser encontradas em [39, 43].

Para os vãos representados na Figura 21, considerou-se, arbitrariamente, o valor de 100 m para o primeiro vão, que conecta a subestação à primeira torre de transmissão. Para os demais vãos, adotou-se 450 m. Este último valor tem como base valores preliminares expostos em [43], que relacionam o vão com a tensão nominal do sistema.

O Apêndice B – Parâmetros e modelagem do estudo de caso 2 - sobretensões devido a descarga atmosférica contém um exemplo de entrada de dados para a rotina *Line/Cable Data*.

A fim de resumir as características principais da modelagem adotadas, considerar a Figura 26 e Tabela 5, onde:

- [Z]: Matriz de impedância série;
- [Y]: Matriz de admitância *shunt*;
- [v]: Matriz de velocidade de propagação da onda;
- Vflash: Tensão mínima para a ruptura do isolamento;
- C_isol: Capacitância da cadeia de isoladores;
- ZC_torre: Impedância de surto da torre;
- c: Velocidade da luz;
- R_pe_torre: Resistência do pé da torre.

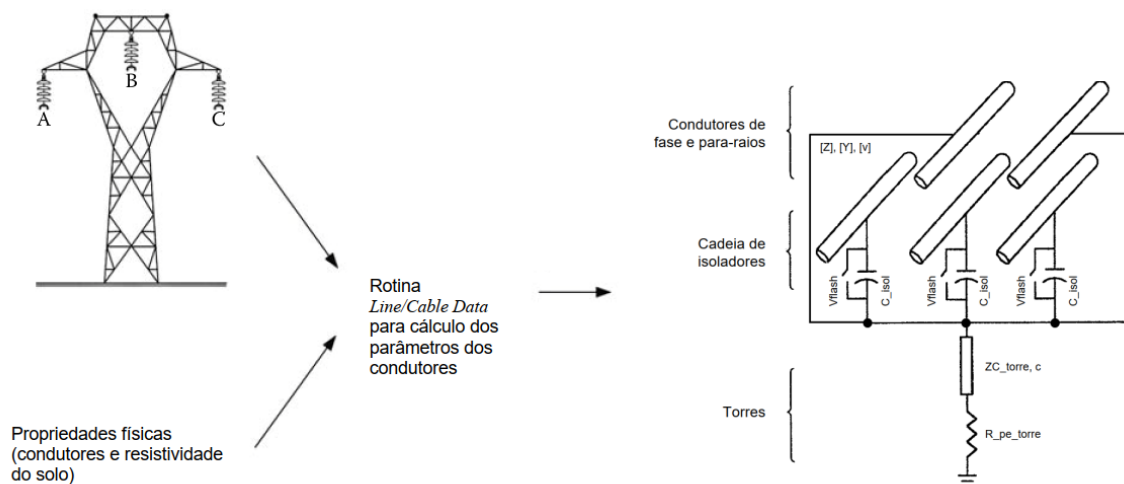


Figura 26 – Resumo da modelagem adotada para a linha de transmissão. Fonte: Adaptado de [32, 38].

Tabela 5– Resumo dos principais parâmetros adotados para a linha de transmissão. Fonte: Autoria Própria.

Parâmetro	Modelagem adotada
$[Z]$, $[Y]$, $[v]$	Parâmetros constantes, conforme cálculo da rotina <i>Line/Cable Data</i> .
Vflash	N(885;25,5 ²) kV
C_isol	U(3;10) pF
ZC_torre	U(100;300) Ω
R_pe_torre	U(10;50) Ω

4.2.1.4 Subestação

O barramento e os condutores entre os pontos de descontinuidade dentro da subestação e as conexões entre os equipamentos da subestação são representados por seções de linha. Essas seções de linha são modeladas por seções de parâmetros distribuídos não transpostos com impedâncias de surto, se forem maiores que 3 m. Caso contrário, uma indutância de $1 \mu\text{H/m}$ é usada. Os parâmetros da linha podem ser calculados usando uma rotina [38].

Diante do exposto, modelou-se os barramentos e os condutores de maneira semelhante àquela considerada para linha de transmissão, utilizando a geometria da subestação – de acordo com os projetos de planta e corte – e a rotina *Line/Cable Data* do EMTP-RV. Para os barramentos e condutores, adotou-se o uso do condutor *rail*, cujas características podem ser encontradas em [43].

Quanto aos equipamentos do pátio da subestação:

- Pode-se modelar o transformador por meio de sua capacitância para terra, caso seja de interesse apenas as tensões do transformador para a terra [34]. A modelagem dos transformadores geralmente é dividida em casos em que a transferência de surto de um enrolamento para outro não é de interesse e, por outro lado, onde esta transferência deve ser levada em consideração [33].
- Disjuntores e chaves seccionadoras são representados por suas capacitâncias entre contatos e entre contatos e terra [34]. Se disjuntores e chaves seccionadoras tem mais que um suporte, apropriadas capacitâncias devem ser adicionadas ao modelo. O estado (aberto ou fechado) destes equipamentos deve ser considerado e pode ser representado por meio de chaves ideais [38].
- Capacitância para terra de todos os isoladores deve ser representada [38].
- Os para-raios de óxido metálico podem ser representados por meio de resistor não linear e indutância de $1 \mu\text{H/m}$ [34].

Diante do exposto anteriormente, considerou-se a representação do transformador de corrente, transformador de potencial indutivo e capacitivo, transformador de potência e isoladores de pedestal por meio de suas respectivas capacitâncias para terra, C_{TC} , C_{TP} , C_{TPC} , C_{TR} e C_{IP} , respectivamente. Para o transformador, a transferência de surto de um enrolamento para outro não foi considerada. Quanto aos disjuntores e chaves seccionadoras, considerou-se chaves ideais para representar seus respectivos estados, bem como as capacitâncias entre contatos e entre contatos e terra dos disjuntores, C_{DJ} , e das chaves seccionadoras, C_{secc} .

Para a modelagem da conexão entre o ponto de aterramento do equipamento de pátio e o sistema de aterramento, considerou-se uma indutância de $1 \mu\text{H/m}$. No entanto, considerou-se que o comprimento deste condutor de aterramento é variável, de tal forma que a indutância modelada deste condutor, $L_{aterramento}$, seja variável.

Com relação aos valores considerados na modelagem, geralmente pouca ou nenhuma informação é dada pelos fabricantes sobre as capacitâncias de sequência positiva e sequência zero dos equipamentos [44]. No entanto, valores típicos podem ser encontrados em algumas referências como em [34, 44]. Estes valores podem ser disponibilizados, por exemplo, por meio de um range de valores, para contemplar variações entre as características do equipamento, como tensão nominal, ou por meio de valores extremos baseados em um conjunto de valores adquiridos de diferentes fabricantes, como em [38].

A Tabela 6 expõe os dados utilizados para os equipamentos da subestação, sendo que, para C_TC, C_TP, C_TPC, C_IP e C_secc, adotou-se valores típicos presentes em [38]. Para C_TR e C_DJ, considerou-se os valores disponibilizados em [34] e [44], respectivamente.

Tabela 6– Resumo dos principais parâmetros adotados para os equipamentos da subestação. Fonte: Autoria Própria.

Parâmetro	Modelagem adotada
C_TC	U(250;680) pF
C_TP	U(500;550) pF
C_TPC	U(5000;8000) pF
C_TR	U(2500;3000) pF
C_IP	U(80;120) pF
C_secc	U(50;100) pF
C_DJ	U(50;100) pF

Em Apêndice B – Parâmetros e modelagem do estudo de caso 2 - sobretensões devido a descarga atmosférica, é disponibilizada a interface com o PAMSUITE, onde os dados variáveis da Tabela 5 e Tabela 6 são definidos, bem como as saídas de interesse do estudo.

A Figura 27 ilustra a característica considerada para o resistor não-linear, considerado para a modelagem do para-raios. Considerou-se o para-raios modelo PEXLIM da fabricante ABB. Os dados podem ser encontrados nos catálogos do fabricante.

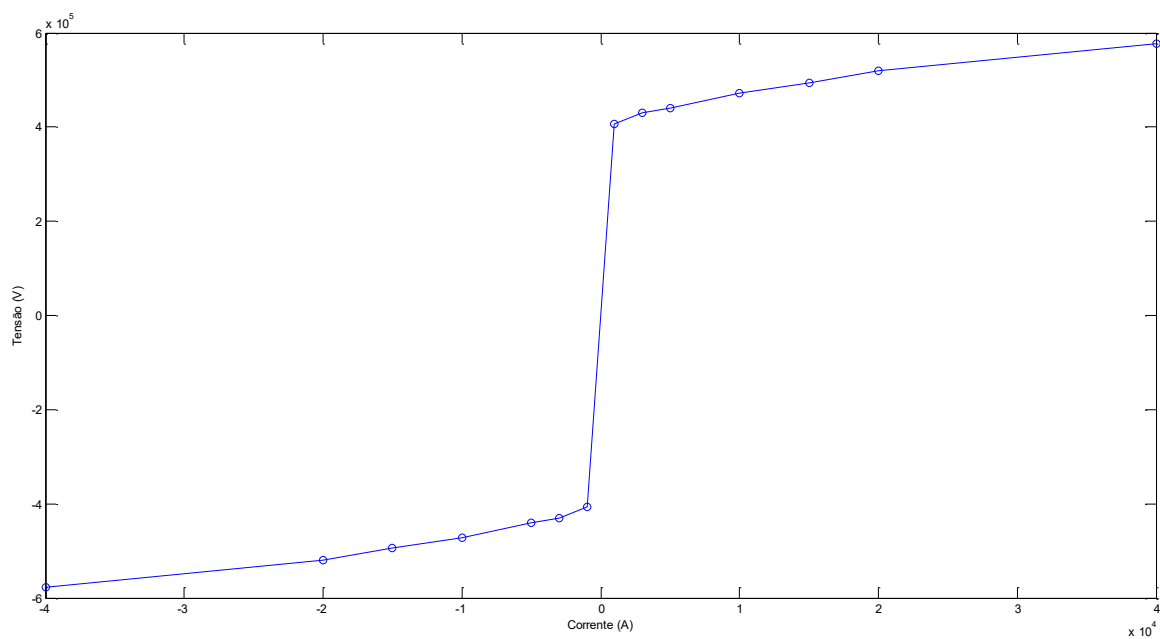


Figura 27 – Característica do para-raios. Fonte: Autoria Própria.

Por fim, na Tabela 7 consta, de maneira resumida, a modelagem de todas as variáveis aleatórias consideradas neste estudo:

Tabela 7– Resumo dos parâmetros variáveis. Fonte: Autoria Própria.

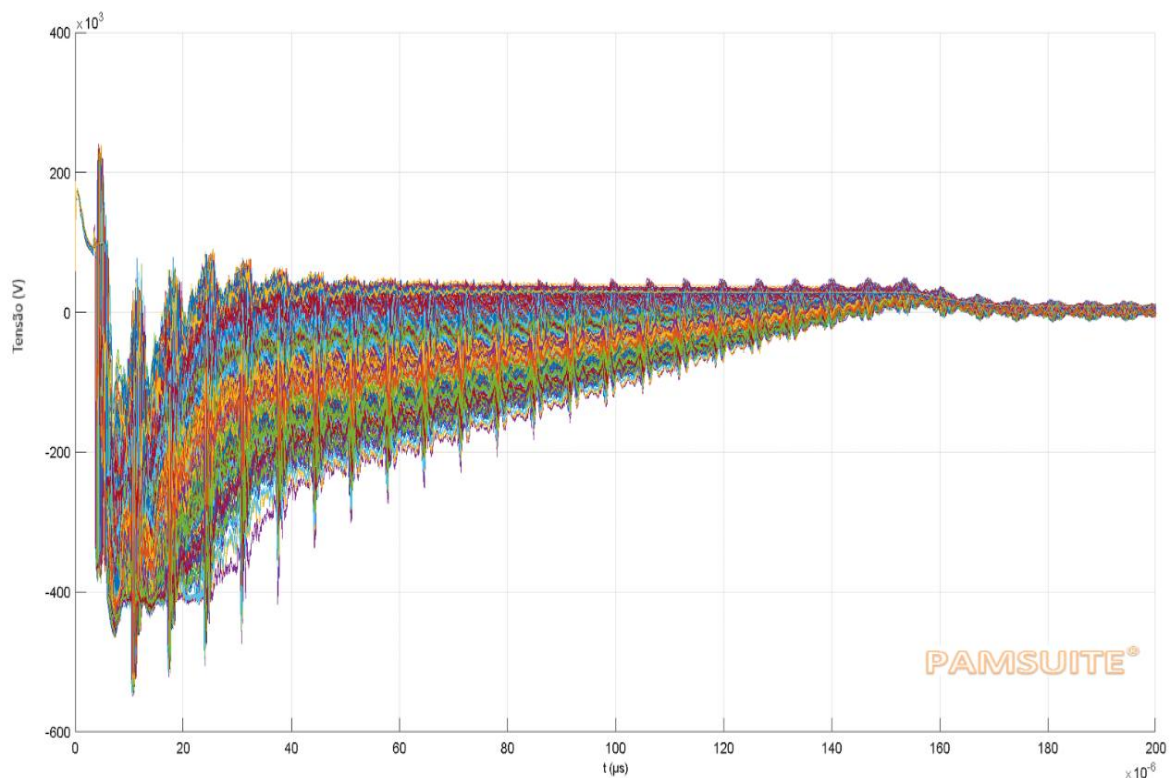
Variável/ Equipamento	Símbolo	Modelagem adotada	Valor
Tensão mínima para ruptura da cadeia de isoladores	Vflash	Chave ideal	N(885;25,5 ²) kV
Cadeia de isoladores	C_isol	Capacitor	U(3;10) pF
Estrutura da torre	ZC_torre	Parâmetros constantes (Bergeron)	U(100;300) Ω
Pé da torre	R_pe_torre	Resistor	U(10;50) Ω
Transformador de corrente	C_TC	Capacitor	U(250;680) pF
Transformador de potencial indutivo	C_TP	Capacitor	U(500;550) pF
Transformador de potencial capacitivo	C_TPC	Capacitor	U(5000;8000) pF
Transformador de potência	C_TR	Capacitor	U(2500;3000) pF
Isolador de pedestal	C_IP	Capacitor	U(80;120) pF
Chave seccionadora	C_secc	Capacitor	U(50;100) pF
Disjuntor	C_DJ	Capacitor	U(50;100) pF
Condutores menores que 3 m	L_aterramento	Indutor	U(0,7;1,0) μ·H/m
Descarga atmosférica	Surto_sort (símbolo representa o valor da função densidade de probabilidade da corrente de crista)	Fonte de corrente	$f(i) = \frac{1}{47} \cdot e^{-\frac{(i-110)}{47}}$, 110 kA ≤ i ≤ 200 kA (valor de crista da corrente de descarga)

4.2.2 Resultados da simulação

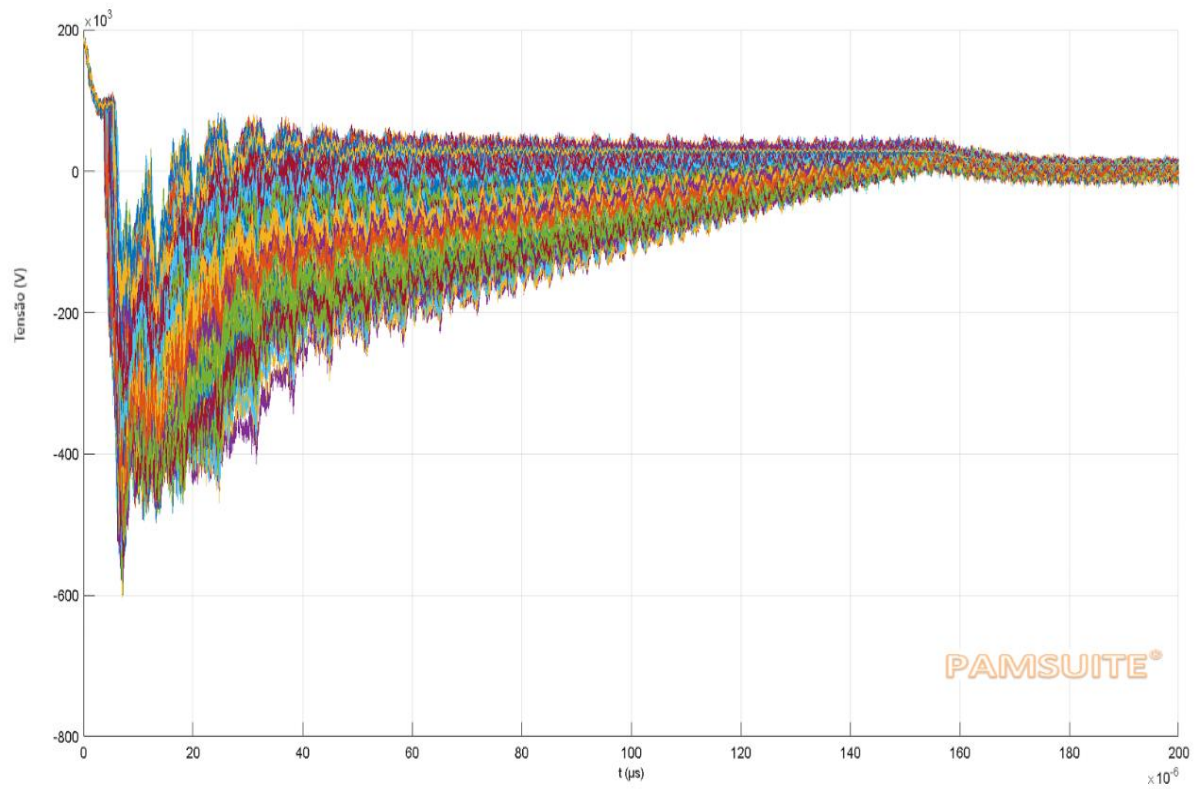
Este subcapítulo contém os detalhes dos resultados obtidos para o estudo de caso proposto.

Conforme citado no subcapítulo anterior, utilizou-se EMTP-RV e PAMSUITE para a elaboração do estudo. Detalhes desta implementação computacional constam no Apêndice B – Parâmetros e modelagem do estudo de caso 2 - sobretensões devido a descarga atmosférica.

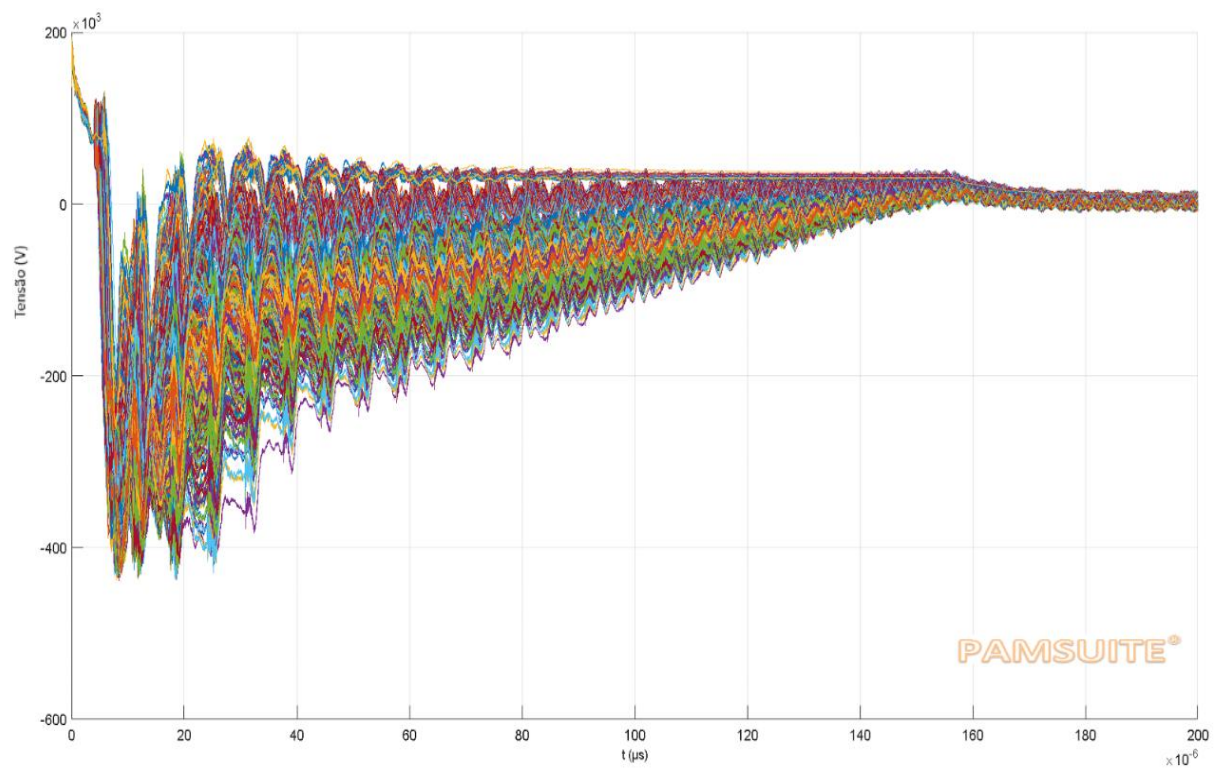
Considerou-se a utilização do método de Monte Carlo para o estudo de propagação de incerteza, utilizando 500 observações. **Escolheu-se, arbitrariamente, o estudo de sobretensão na fase B de três diferentes pontos da subestação: para-raios de entrada de linha, chave seccionadora de *by-pass* da entrada de linha e para-raios do transformador, pontos indicados, respectivamente, por P1, P2 e P3 na modelagem do EMTP-RV.** A Figura 28 ilustra as tensões obtidas nestes pontos, enquanto a Figura 29 ilustra o caso contendo o maior valor de crista de tensão entre os valores coletados.



(a)



(b)



(c)

Figura 28 – Tensão versus tempo. (a) para-raios de entrada de linha, (b) chave seccionadora de *by-pass* da entrada de linha e (c) para-raios do setor de 230 kV do transformador de potência. Fonte: Autoria Própria.

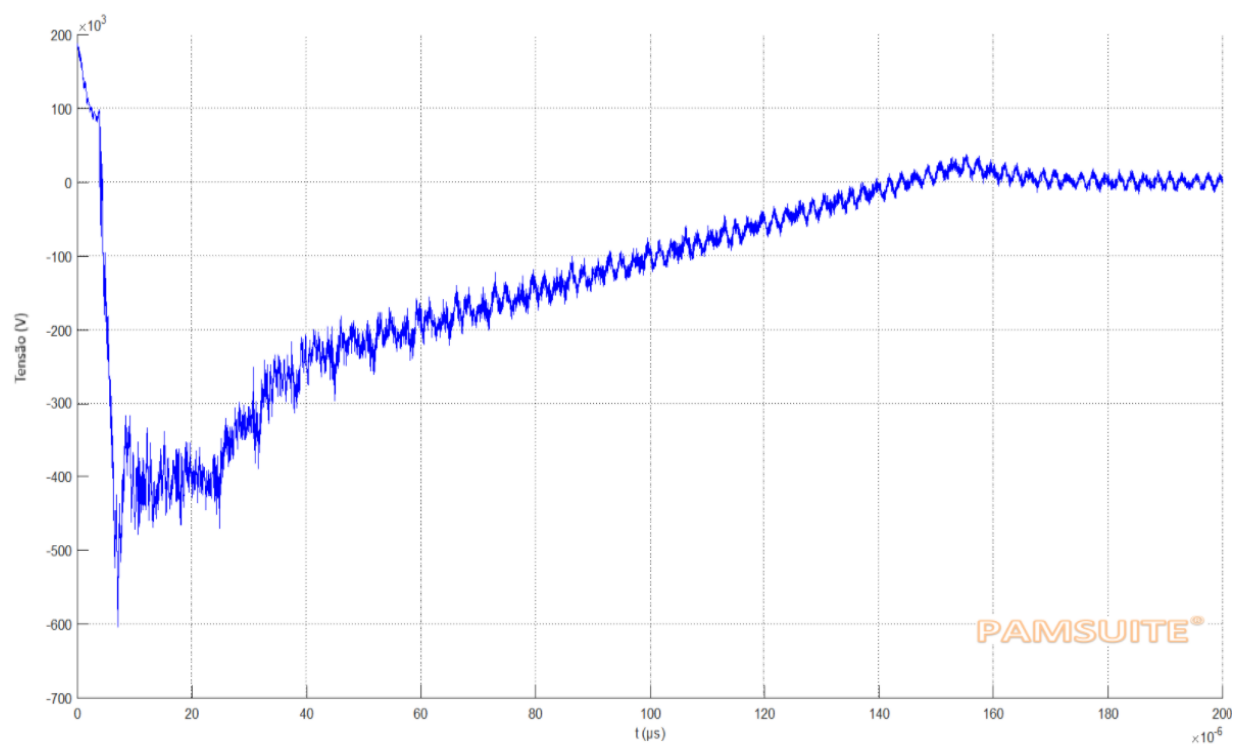
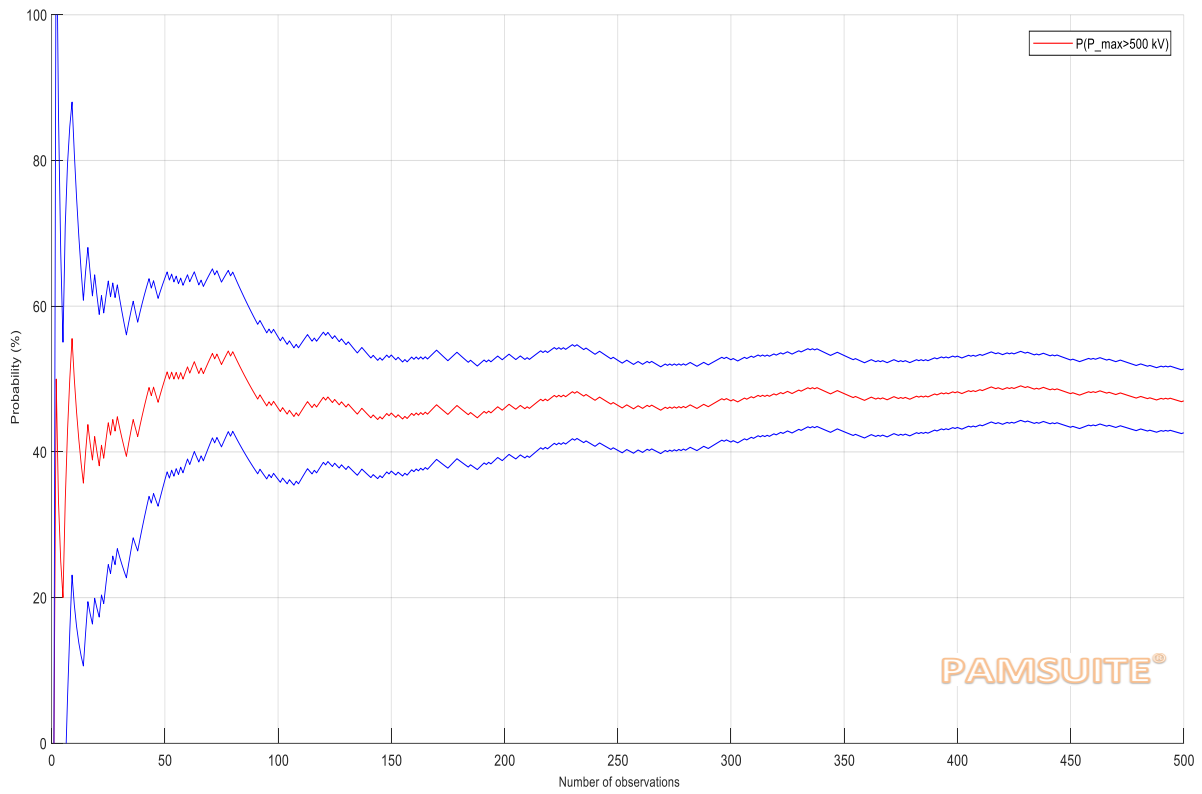
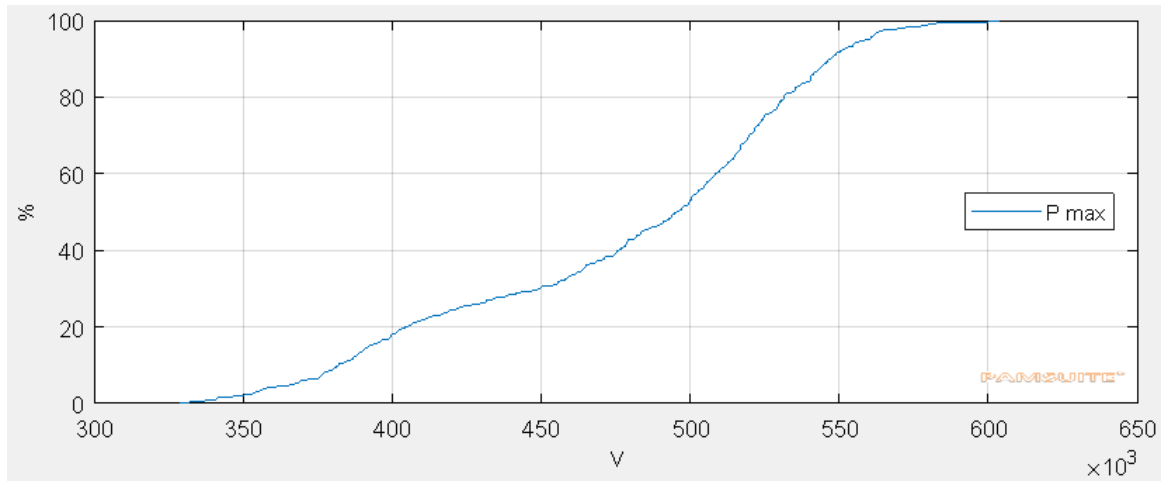


Figura 29 – Tensão versus tempo: curva de maior valor absoluto de tensão de crista. Fonte: Autoria Própria.

Adicionalmente, optou-se por coletar o valor máximo absoluto, $P_{\max}$, para cada uma das 500 observações. A Figura 30 ilustra algumas características de $P_{\max}$.



(a)



(b)

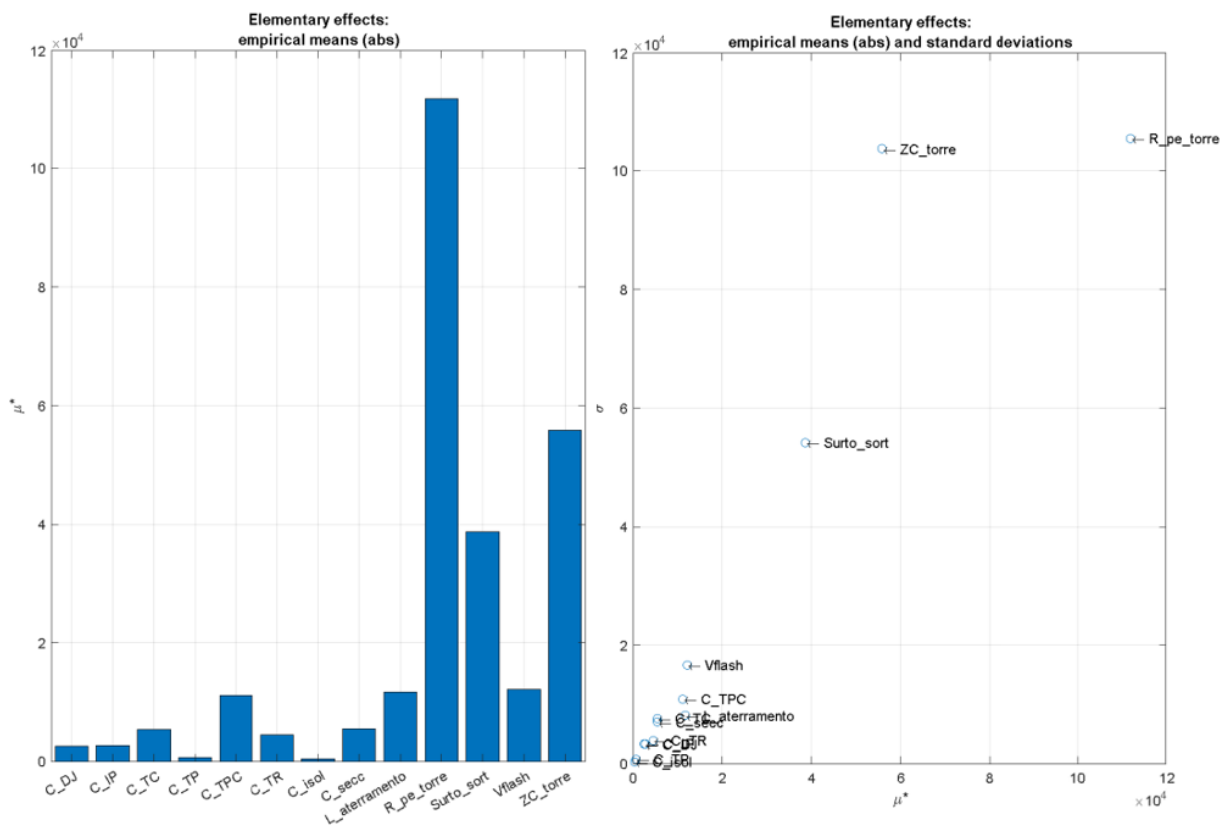
Figura 30 – (a) probabilidade de o valor máximo absoluto da tensão de crista superar 500 kV e (b) função de distribuição cumulativa do valor máximo absoluto da tensão de crista. Fonte: Autoria Própria.

Ao ampliar a Figura 30(a), é possível ver que o valor de $\hat{p}$ é igual 47%. Considerando o cálculo do intervalo de confiança conforme Eq. (26), bem como $\alpha = 5\%$, vem que o intervalo de confiança é: $\hat{p} - 1,96 \cdot \sqrt{\frac{0,47 \cdot (1-0,47)}{500}} \leq p \leq \hat{p} + 1,96 \cdot \sqrt{\frac{0,47 \cdot (1-0,47)}{500}}$, que resulta em uma

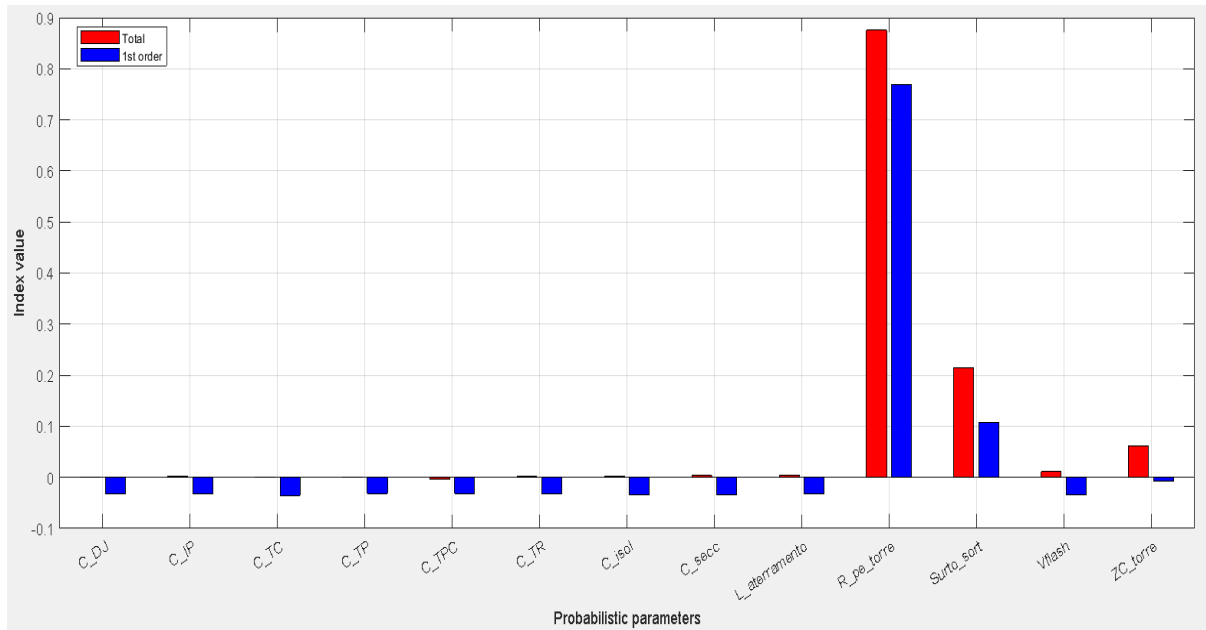
amplitude igual a $2 \cdot 1,96 \cdot \sqrt{\frac{0,47 \cdot (1-0,47)}{500}}$, que por sua vez equivale a cerca de 18,5% de $\hat{p}$. Note que é possível diminuir ainda mais esta amplitude, ao considerar um valor maior para N .

A Figura 30(b) mostra que a maior sobretensão obtida (cerca de 600 kV) ainda possui uma certa margem em relação ao NBI de 850 kV para todos os percentis.

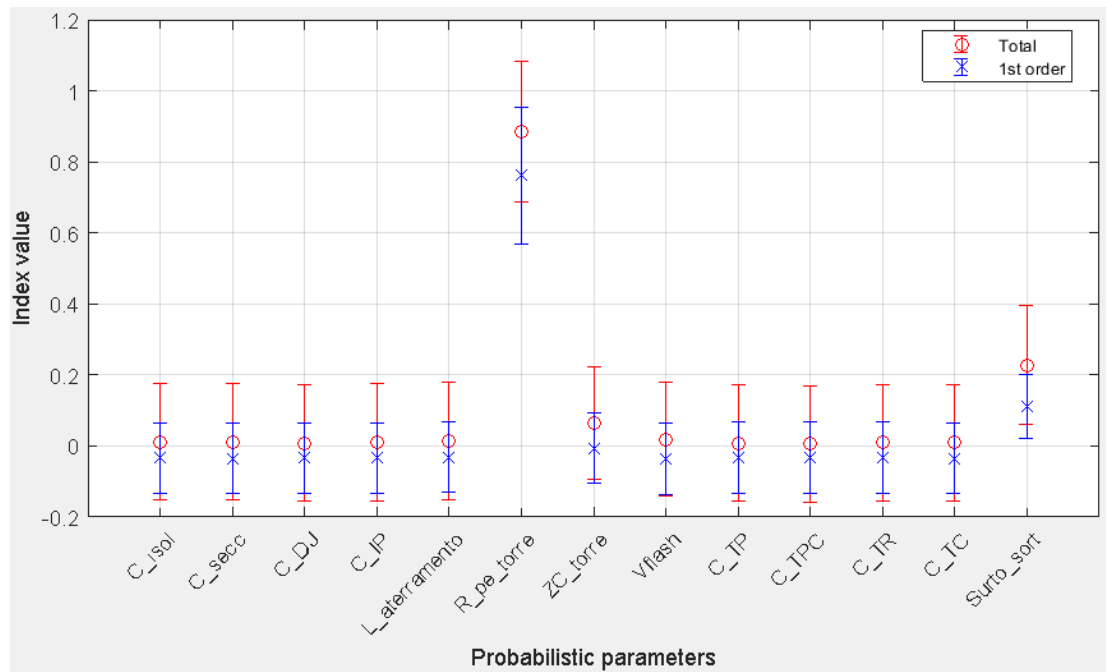
Além disso, realizou-se um estudo de análise de sensibilidade, de forma a conhecer quais variáveis possuem maior influência no valor de $P_{\max}$. A Figura 31 ilustra os resultados deste estudo.



(a)



(b)



(c)

Figura 31 – Análise de sensibilidade: (a) método de Morris, (b) método de Sobol sem intervalo de confiança e (c) método de Sobol com intervalo de confiança. Fonte: Autoria Própria.

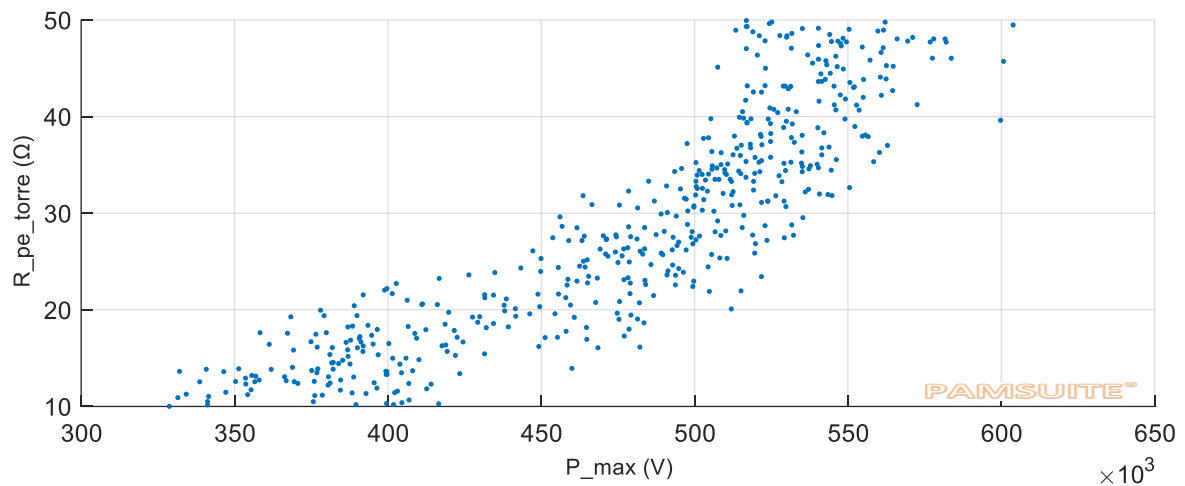
onde Surto_sort corresponde a um valor sorteado da função densidade de probabilidade, que é utilizado para o cálculo de I_f , conforme Eq. (65).

Cabe ressaltar que, na Figura 31, apesar do método Sobol ter apresentado alguns resultados negativos para os índices de primeira ordem e total, há um intervalo de confiança

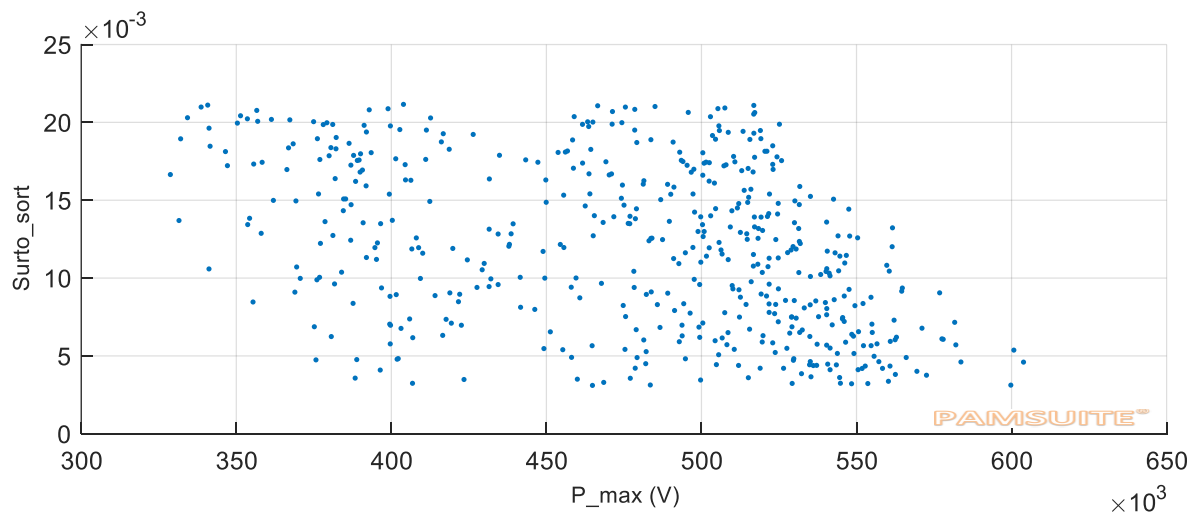
para estes valores, que inclui o valor zero e valores positivos. Ainda, de acordo com os resultados apresentados, nota-se que a resistência do pé da torre, R_{pe_torre} , é um parâmetro importante nesta modelagem para a determinação de P_{max} .

Além disso, pode-se concluir, por meio dos valores determinísticos da Figura 31(b), que o parâmetro R_{pe_torre} é mais importante que o parâmetro $Surto_sort$, que por sua vez é mais importante que o parâmetro C_{TR} , para o cálculo de P_{max} . No entanto, de acordo com a Figura 31(c), os intervalos considerados para $Surto_sort$ e C_{TR} se sobrepõem. Então, para que esta conclusão seja feita com o mesmo nível de confiança da Figura 31(c), é necessário considerar um número maior para a amostra base.

A Figura 32 ilustra os gráficos de dispersão, que reforçam a afirmação anterior.



(a)



(b)

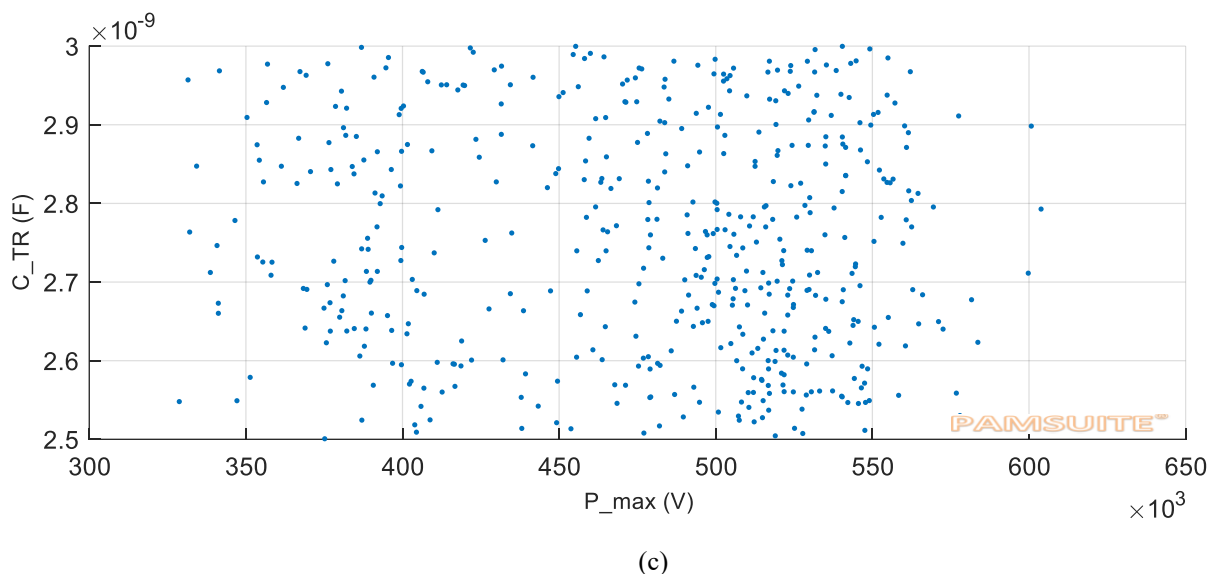


Figura 32 – Gráfico de dispersão: (a) resistência do pé da torre, (b) corrente de descarga e (c) capacitância do transformador de potência. Fonte: Autoria Própria.

Uma outra ferramenta interessante disponível no PAMSUITE para a análise de sensibilidade é o gráfico de teia de aranha. Neste tipo de gráfico, os valores dos parâmetros são normalizados para uma melhor leitura. A Figura 33 ilustra um gráfico de teia de aranha, onde os conjuntos de valores dos parâmetros que resultam em $P_{\max}$ maior que 530 kV estão destacados em vermelho. Os demais conjuntos de valores são representados por linhas azuis.

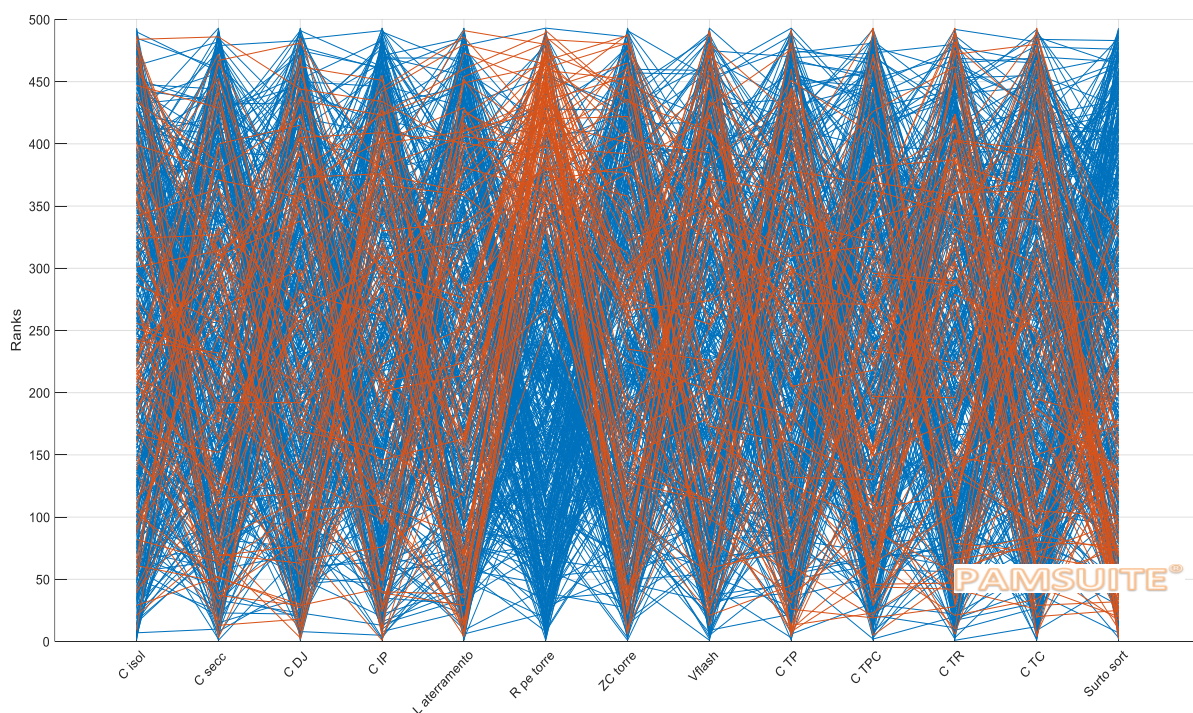


Figura 33 – Gráfico de teia de aranha. Fonte: Autoria Própria.

Nota-se, por meio da Figura 33, que o parâmetro R_{pe_torre} é predominante para resultar na condição supracitada.

Por fim, pode-se concluir que os resultados obtidos por meio da análise paramétrica são bastante detalhados e podem ser utilizados, por exemplo, como base para uma melhor tomada de decisão. Por exemplo, suponha que por algum motivo fosse necessário reduzir os níveis de sobretensões obtidos nas simulações deste sistema sob estudo. Neste caso, com a abordagem tradicional, o engenheiro poderia, eventualmente, concentrar os esforços em determinar precisamente os parâmetros de capacitância de todos os equipamentos de pátio e da indutância dos condutores menores que 3 m que atendam o objetivo, o que poderia levar a um grande gasto de recursos e não chegar ao resultado desejado. No entanto, de acordo com a análise de sensibilidade realizada, fica claro que o ideal seria diminuir a resistência do pé da torre. Similarmente, caso o objetivo fosse diminuir a variância da saída, o ideal seria uma melhor medição da resistência do pé da torre, considerando o tipo do solo, geometria, processo de ionização etc.

5 Conclusão

Neste estudo, foram aplicados os conceitos de estatística e utilizou-se ferramentas computacionais para elaboração de análise paramétrica destinada a estudos de transitórios eletromagnéticos.

Nota-se o grande ganho da análise paramétrica quando comparada à abordagem tradicional, pois é possível produzir resultados extremamente detalhados, que podem ser utilizados como base para tomar decisões de maneira mais rápida, confiante e precisa.

A utilização do software PAMSUITE em conjunto com o EMTP-RV mostrou-se muito poderosa e a aplicação não se limita aos exemplos considerados neste trabalho. Alternativamente, foi proposta uma abordagem que leva em conta a utilização de *scripts* em Python e o pacote ATP para a realização da análise paramétrica. Claramente, o objetivo não é comparar as duas diferentes abordagens, visto que os *scripts* desenvolvidos para este trabalho são específicos para o estudo de caso proposto e poderiam ser melhorados.

Conforme descrito no capítulo referente à metodologia empregada, ressalta-se os seguintes tópicos relativos à análise paramétrica em estudos de transitórios eletromagnéticos:

- Simulação com incertezas de parâmetros de entrada utilizando o método de Monte Carlo;
- Análise de sensibilidade;
- Estudos de otimização;
- Calibração de modelos.

Embora nem todos os tópicos supracitados foram explorados nos estudos de caso apresentados neste trabalho, os princípios e potenciais aplicações do uso de análises paramétricas em simulações computacionais de transitórios eletromagnéticos baseadas no EMTP foram apresentados e discutidos nesta dissertação.

A identificação das variáveis aleatórias de maior impacto nos resultados permite priorizar medições ou melhorias nos dados de entrada, bem como nos modelos utilizados. Portanto, a análise paramétrica é ferramenta importante para engenheiros e tomadores de decisão.

5.1 Publicações na área

Este trabalho resultou em uma publicação e apresentação em congresso internacional (conferência IEEE URUCON 2024, patrocinada pelo Conselho IEEE Cono Sur e pela Seção IEEE Uruguai), conforme referência a seguir:

- L. G. R. Paula, B. D. Bonatto and P. P. Balestrassi, "*Parametric Analysis Applied to Electrical Studies Using Electromagnetic Transients Programs*" IEEE URUCON, 2024.

Além disso, o seguinte trabalho foi submetido à conferência INDUSCON 2025, patrocinada pelo IEEE e pela IAS (do inglês *Industry Applications Society*):

- L. G. R. Paula and B. D. Bonatto, "*Parametric Analysis Using EMTP-Based Simulations: Principles and Potential Applications*" INDUSCON, 2025.

5.2 Trabalhos Futuros

Recomenda-se, para trabalhos futuros:

- Melhorar os *scripts* propostos ou utilizar outra ferramenta, como o EMTP-RV API, a fim de automatizar ainda mais a execução, visto que, nos *scripts* propostos:
 - A identificação dos parâmetros a serem modificados é feita de forma manual no arquivo “.atp” e, posteriormente, inseridos como entrada nos *scripts*;
 - É necessário que todos os valores de um determinado parâmetro de entrada possuam uma mesma quantidade de caracteres, de modo a não alterar a estrutura do arquivo “.atp” quando o valor deste parâmetro for modificado. Portanto, pode ser necessário incluir no *script* um tratamento dos dados de entrada, por exemplo, por meio da adoção de um certo número de casas decimais;
 - Não há uma interface gráfica com o usuário, o que resulta em uma dificuldade para identificação dos campos a serem alterados para cada estudo de caso;

- Não há implementada uma forma do *script* sinalizar eventuais erros na simulação realizada pelo ATP.
- Aplicar, no *script* realizado para o cálculo dos índices Morris, a melhoria citada no capítulo de revisão da literatura, ao maximizar as distâncias d_{ml} , a fim de otimizar a varredura do espaço de entrada. Além disso, para os índices Sobol, poderia ser incluído um critério de convergência com base no cálculo do erro provável dos índices, em vez de definir preliminarmente o tamanho da amostra.
- Ainda com relação aos *scripts*, é válida a utilização de semente aleatória a fim de garantir a reprodutibilidade.
- Realizar alterações no estudo de caso apresentado neste trabalho, referente a descarga atmosférica. Pode-se:
 - Alterar a topologia da rede, ao considerar diferentes cenários operativos da subestação;
 - Melhorar o modelo, por exemplo:
 - incluir o efeito corona e a variação da resistência de aterramento devido à ionização do solo;
 - considerar modelos mais complexos para a estrutura da torre, bem como um modelo mais detalhado para o arco elétrico e o instante de sua ocorrência;
 - Modificar mais parâmetros do modelo, como a frequência adotada na rotina para o cálculo dos parâmetros dos condutores e o ponto de incidência da descarga atmosférica.
- Considerar uma melhor representação das funções densidade de probabilidade ao incluir medições reais ou dados provenientes de literatura, quando aplicável.
- Explorar outros tópicos acerca de estatística, tais como a inclusão de:
 - diferenciação entre fatores controláveis e não controláveis;
 - aplicação de inferência estatística para experimentos comparativos, considerando indicador *p-value* como base para as conclusões;
 - outras ferramentas, como gráfico de Pareto, gráfico de interações de segunda ordem e a utilização dos recursos do aplicativo *Crystal Ball*, Minitab etc.
 - explorar os diferentes métodos de amostragem de Monte Carlo.

- Aplicar análise paramétrica para outros modelos de TEM, como estudos de ferorressonância, religamento automático, tensão de restabelecimento transitória, modelagem de parques aerogeradores e fotovoltaicos (conforme problema exposto em [11]) etc.

Referências

- [1] M. Martinez-Duro, "Going parametric in EMT studies: EDF methods and tools for input data uncertainties, sensitivity analysis and parameter identification," *IPST*, 2019.
- [2] "ONS: Diretrizes para a Elaboração de Projetos Básicos para Empreendimentos de Transmissão: Estudos Elétricos, Especificação das Instalações, de Equipamentos e de Linhas de Transmissão," 2013.
- [3] M. Fratila, A. El-Akoum, M. Martinez-Duró, P. Poujade and M. Flores, "Stability studies with parameter uncertainties," *IPST*, 2019.
- [4] A. Holdyk and J. Holboell, "Implementation of the parametric variation," *IPST*, 2013.
- [5] J. C. Martins, *Avaliação dos impactos das variações sistêmicas nas análises de viabilidade do religamento monopolar através do ATP via Python*, Juiz de Fora: Universidade Federal de Juiz de Fora, 2024.
- [6] L. Paula, B. Bonatto and P. Balestrassi, "Parametric Analysis Applied to Electrical Studies Using Electromagnetic Transients Programs," *IEEE URUCON*, 2024.
- [7] J. A. Martinez-Velasco and F. Castro-Aranda, "Parametric Analysis of the Lightning Performance of Overhead Transmission Lines Using an Electromagnetic Transients Program," *IPST*, 2003.
- [8] República Federativa do Brasil: Constituição Federal, 2019.
- [9] I. Oliveira, *Análise Econômica da Inserção de Painéis Fotovoltaicos no Campus da UNIFEI - Itajubá com Simulação de Monte Carlo*, Itajubá: Universidade Federal de Itajubá, 2019.
- [10] Z. Emin, H. Motoyama, A. B. I. Papič and A. Halley, "General Report for Study Committee C4 (Power System Technical Performance), CIGRE Paris Session," Paris, 2018.
- [11] "ONS: Análise da perturbação do dia 15/08/2023 às 08h30min," 2023.
- [12] D. C. Montgomery and G. C. Runger, *Estatística Aplicada e Probabilidade para Engenheiros*, 4ª ed., LTC, 2009.
- [13] J. L. Devore, *Probabilidade e Estatística para Engenharia e Ciências*, 6ª ed., São Paulo: Cengage Learning, 2006.
- [14] M. Pereira and P. Pereira, *Notas de Aula de Estatística Aplicada à Engenharia*.
- [15] A. Jr. and A. Sousa, *Fundamentos de Metrologia Científica e Industrial*, 1ª ed., São Paulo: Manole, 2008.
- [16] L. Bonomo, *Análise de Incertezas e de Sensibilidade Global para Educação de Impedância Acústica*, Florianópolis: Universidade Federal de Santa Catarina, 2021.
- [17] H. W. Coleman and W. G. Steele, *Experimentation, Validation, and Uncertainty Analysis for Engineers*, 4ª ed., Hoboken: John Wiley e Sons, 2009.
- [18] X. Song, J. Zhang, C. Zhan, Y. Xuan, M. Ye and C. Xu, "Global sensitivity analysis in hydrological modeling: Review of concepts, methods, theoretical framework, and applications," *Journal of Hydrology* 523, p. 739–757, 14 Fevereiro 2015.
- [19] D. M. Hamby, "A Review of Techniques for Parameter Sensitivity Analysis of Environmental Models," *Environmental Monitoring and Assessment*, pp. 135-154, 1994.

- [20] A. Saltelli, S. Tarantola, F. Campolongo and M. Ratto, *Sensitivity Analysis in Practice – a Guide to Assessing Scientific Models*, Ispra: John Wiley & Sons, 2004.
- [21] W. Tian, "A Review of Sensitivity Analysis Methods in Building Energy Analysis," *Renewable and Sustainable Energy Reviews* 20, p. 411–419, 11 Janeiro 2013.
- [22] F. R. Marin and Q. d. J. v. Lier, *Modelagem de Crescimento de Culturas Agrícolas - Análise de Sensibilidade - Notas de Aula*.
- [23] D. Garcia, I. ArosteGUI and R. Pillezo, "Robust combination of the Morris and Sobol methods in complex multidimensional model," *Environmental Modelling and Software* 122, 24 Setembro 2019.
- [24] M. D. Morris, "Factorial Sampling Plans for Preliminary Computational Experiments," *Technometrics*, pp. 161-174, 2 Maio 1991.
- [25] A. Saltelli and M. Ratto, *Global Sensitivity Analysis: The Primer*, Ispra: John Wiley & Sons, 2008.
- [26] F. Campolongo, J. Cariboni and A. Saltelli, "An effective screening design for sensitivity analysis of large models," *Environmental Modelling & Software* 22, pp. 1509-1518, 15 Outubro 2006.
- [27] A. Saltelli, "Making best use of model evaluations to compute sensitivity indices," *Computer Physics Communications* 145, pp. 280-297, 2002.
- [28] I. M. Sobol, "Sensitivity estimates for nonlinear mathematical models," *Matematicheskoe Modelirovanie* 2, p. 112–118, 1990 in Russian, translated in English in 1993.
- [29] T. Homma and A. Saltelli, "Importance measures in global sensitivity analysis of nonlinear models," *Reliability Engineering and System Safety* 52, p. 1–17, 5 Janeiro 1996.
- [30] A. Saltelli, P. Annoni, I. Azzini, F. Campolongo, M. Ratto and S. Tarantola, "Variance based sensitivity analysis of model output. Design and estimator," *Computer Physics Communications* 181, p. 259–270, Setembro 2009.
- [31] I. Sobol, S. Tarantola, D. Gatelli, S. Kucherenko and W. Mauntz, "Estimating the approximation error when fixing unessential factors in global sensitivity analysis," *Reliability Engineering & System Safety* 92, vol. 92, p. 957–960, 28 Agosto 2006.
- [32] J. A. Martinez-Velasco, *Power System Transients: Parameter Determination*, CRC Press, 2009.
- [33] "CIGRE WG 33.02: Guidelines for representation of network elements when calculating transients," 1999.
- [34] "IEC TR 60071-4: Insulation co-ordination – Part 4: Computational guide to insulation co-ordination and modelling of electrical networks," 2004.
- [35] J. A. Martinez, "Parameter Determination for Power Systems Transients," *IEEE Power Engineering Society General Meeting*, 2007.
- [36] J. C. G. d. Siqueira, B. D. Bonatto, J. R. Martí, J. A. Hollman and H. W. Dommel, "A discussion about optimum time step size and maximum simulation time in EMTP-based programs," *Electrical Power and Energy Systems* 72, p. 24–32, 26 Março 2015.
- [37] H. W. Dommel, "Digital Computer Solution of Electromagnetic Transients in Single- and Multiphase Networks," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-88, pp. 388-399, Abril 1969.
- [38] "Modeling Guidelines for fast front transients," *IEEE Transactions on Power Delivery*, Vol 11, No 1, Janeiro 1996.

- [39] Transmission Line Reference Book 345 kV and Above, 2^a ed., Palo Alto: Electric Power Research Institute, 1982.
- [40] "CIGRE WG 33.01: Guide to procedures for estimating the lightning performance of transmission lines," 1991.
- [41] *ABNT NBR 6939: Coordenação do isolamento - Procedimento*, 2018.
- [42] N. Tonmitr, K. Tonmitr and E. Kaneko, "The Effect of Controlling Stray and Disc Capacitance of Ceramic String Insulator in the Case of Clean and Contaminated Conditions," *Procedia Computer Science* 86, vol. 86, pp. 333-336, 2016.
- [43] *EPE: EPE-DEE-NT-012/2018 - Definição de Parâmetros Iniciais Para o Planejamento de Linhas de Transmissão Aéreas*.
- [44] *IEEE Std C37.011-2019: IEEE Guide for the Application of Transient Recovery Voltage for AC High-Voltage Circuit Breakers with Rated Maximum Voltage above 1000 V*, 2019.
- [45] L. M. N. d. Mattos, 2020. [Online]. Available: <https://github.com/ldemattos/readPL4>. [Accessed 2024].

Apêndice A – Rotina de automatização do Python para o estudo de caso 1

Descreve-se neste apêndice as rotinas implementadas em Python para a realização de análise paramétrica no Estudo de Caso 1.

A Tabela 8 contém as funções do Python responsáveis pela alteração automática do arquivo “.atp” do Estudo de Caso 1. Estas funções são utilizadas nos *scripts* da Tabela 9, Tabela 10 e Tabela 11, responsáveis pelo estudo de tratamento de incerteza pelo método de Monte Carlo (Tabela 9) e pela análise de sensibilidade por meio do cálculo dos índices Morris (Tabela 10) e Sobol (Tabela 11).

Para cada conjunto de parâmetros de entrada simulado pelo ATP, é gerado um arquivo “.pl4”, que contém os resultados da respectiva simulação. No entanto, os dados são binários e, para convertê-los, adaptou-se a rotina proposta em [45], conforme exposto na Tabela 12.

Tabela 8– Script em Python: definição das funções para alteração do arquivo “. atp”. Fonte: Autoria Própria.

1	<code>import fileinput</code>
2	<code>import sys</code>
3	<code>import os</code>
4	<code>import pandas as pd</code>
5	<code>import numpy as np</code>
6	<code>from lib_readPL4_py3 import readPL4</code>
7	<code># Função de alteração dos valores no arquivo ".atp"</code>
8	<code>def replacement(file, previousw, nextw):</code>
9	<code> for line in fileinput.input(file, inplace=1):</code>
10	<code> line = line.replace(previousw, nextw)</code>
11	<code> sys.stdout.write(line)</code>
12	<code># Função execução do arquivo ".atp"</code>
13	<code>def executa (V_executa,R_executa,L_executa,C_executa):</code>
14	<code> # Condição inicial dos parâmetros de entrada (dados do ".atp")</code>
15	<code> File = r'C:\ATPDraw\ATP\Estudo1.atp'</code>
16	<code> V_0 = '11XX0003 0 48.89'</code>
17	<code> R_0 = ' XX0001XX0002 1234'</code>
18	<code> L_0 = ' XX0002VC 567'</code>
19	<code> C_0 = ' VC .891'</code>
20	<code> vc_max = []</code>
21	<code> for i in range(len(V_executa)):</code>
22	<code> replacement(File,V_0,'11XX0003 0 '+V_executa[i])</code>
23	<code> replacement(File,R_0,' XX0001XX0002 '+R_executa[i])</code>
24	<code> replacement(File,L_0,' XX0002VC '+L_executa[i])</code>
25	<code> replacement(File,C_0,' VC '+C_executa[i])</code>
26	<code> V_0='11XX0003 0 '+V_executa[i]</code>
27	<code> R_0=' XX0001XX0002 '+R_executa[i]</code>
28	<code> L_0=' XX0002VC '+L_executa[i]</code>
29	<code> C_0=' VC '+C_executa[i]</code>
30	<code> #Executa .ATP</code>
31	<code> os.system(r'C:\ATP\atpmingw\tpbig.exe disk C:\ATPDraw\ATP\Estudo1.atp s -r"')</code>
32	<code> #Conversão ".pl4" e aquisição da saída de interesse</code>
33	<code> pl4 = readPL4(r'C:\ATPDraw\ATP\Estudo1.pl4')</code>
34	<code> vc_max.append((float(np.max(np.absolute(pl4[1][:,1]))))) # 1 é posição da saída de interesse, vc, no arquivo ".pl4"</code>

35	#Voltando o arquivo ".atp" nas condições iniciais
36	replacement(File,V_0,'11XX0003 0 48.89')
37	replacement(File,R_0,' XX0001XX0002 1234')
38	replacement(File,L_0,' XX0002VC 567')
39	replacement(File,C_0,' VC .891')
40	return vc_max

Tabela 9– Script em Python: tratamento de incerteza pelo método de Monte Carlo. atp”. Fonte: Autoria Própria.

1	import fileinput
2	import sys
3	import os
4	import pandas as pd
5	import numpy as np
6	from lib_readPL4_py3 import readPL4
7	from Estudo1_Monte_Carlo import executa
8	from Estudo1_Monte_Carlo import replacement
9	# Leitura do excel com os parametros de entrada para as simulações
10	df_parametros = pd.read_excel ('Dados Estudo 1.xlsx ' , sheet_name='entrada' , dtype=str)
11	# Atribuindo as variáveis de entrada
12	V=pd.Series.to_numpy(df_parametros.loc [: , 'V'])
13	R=pd.Series.to_numpy(df_parametros.loc [: , 'R'])
14	L=pd.Series.to_numpy(df_parametros.loc [: , 'L'])
15	C=pd.Series.to_numpy(df_parametros.loc [: , 'C'])
16	# Simulando os casos
17	vc_max=executa(V,R,L,C)
18	df=pd.DataFrame(vc_max)
19	df.to_excel(r'resultado.xlsx')

Tabela 10– Script em Python: análise de sensibilidade pelo cálculo dos índices Morris. Fonte: Autoria Própria.

1	import fileinput
2	import sys
3	import os
4	import random
5	import statistics
6	import pandas as pd
7	import numpy as np
8	from lib_readPL4_py3 import readPL4
9	from Estudo1_Monte_Carlo import executa
10	from Estudo1_Monte_Carlo import replacement
11	# Entrada de dados
12	# Parâmetros TEM:
13	V=[40,60] # U(40;60) V
14	R=[600,1400] # U(600;1400) Ω
15	L=[60,140] # U(60;140) mH
16	C=[.25,.55] # U(250;550) nF
17	k=4 # Número de parâmetros variáveis (V, R, L e C)
18	# A seguir, válido apenas para distribuição uniforme
19	Intervalo_entrada=np.full((k+1,k),np.array([[V[1]-V[0],R[1]-R[0],L[1]-L[0],C[1]-C[0]]]))
20	Minimo_entrada=np.full((k+1,k),np.array([[V[0],R[0],L[0],C[0]]]))
21	# Parâmetros método Morris
22	p=4 # Nível da rede
23	r=10 # Número de trajetórias
24	Delta=p/(2*(p-1)) # Passo
25	# Construção das trajetórias: parâmetros fixos
26	J1=np.ones((k+1,1)) # Jk+1,1
27	J2=np.ones((k+1,k)) # Jk+1,k
28	B=np.tri((k+1),k,-1) # B
29	conjunto_aux=np.linspace(0,1,p) # Conjunto base para determinar x*
30	conjunto=[]
31	for i in range (int(p/2)): # Escolhendo os elementos do conjunto base tal que x(1) ainda esteja em Ω
32	conjunto.append(conjunto_aux[i])
33	# Construção das trajetórias: parâmetros variáveis

```

34 def Morris (V,R,L,C,k,p,Delta,J1,J2,B,conjunto):
35     # Determinando x*, chamado no script de "x_"
36     x_=[]
37     for i in range (k):
38         x_.append(random.choice(conjunto))
39     x=np.array([x_])
40     # Determinando D*, chamado no script de "D_"
41     D_=[]
42     D_aux=[]
43     for i in range (k):
44         D_aux.append(random.choice([1,-1]))
45     D=np.diag(D_aux)
46     # Determinando P*, chamado no script de "P_".
47     # P* indica o primeiro fator a variar. Após isso, D* indica se este fator
    irá aumentar ou diminuir
48     P=np.zeros((k,k))
49     permut=[]
50     for i in range(k):
51         permut.append(i)
52     permut=np.random.choice(permut,size=k,replace=False) # Criando a permutação aleatória
53     for i in range (k):
54         P_[permut[i],i]=1
55     # Determinando B* chamado no script de "B_"
56     B_=((J1@x_)+((Delta/2)*(((2*B-J2)@D_)+J2)))@P_
57     # Determinando valores de entrada no ".atp"
58     Entrada_ATP=((B_*Intervalo_entrada)+Minimo_entrada)
59     # Compatibilizando o formato da entrada de dados no ".atp"
60     Entrada_ATP[:,0]=Entrada_ATP[:,0].round(1)
61     Entrada_ATP[:,1]=Entrada_ATP[:,1].round()
62     Entrada_ATP[:,2]=Entrada_ATP[:,2].round()
63     Entrada_ATP[:,3]=Entrada_ATP[:,3].round(2)
64     V=list(map(str, Entrada_ATP[:,0]))
65     R=list(map(str,list(map(int, Entrada_ATP[:,1]))))
66     L=list(map(str,list(map(int, Entrada_ATP[:,2]))))
67     C=list(map(str, Entrada_ATP[:,3]))
68     for i in range (k+1):
69         V[i]='{:0<5}'.format(V[i])
70         R[i]='{:0>4}'.format(R[i])
71         L[i]='{:0>3}'.format(L[i])
72         C[i]='{:0<4}'.format(C[i])
73     return V,R,L,C,D_aux,permut
74 # Cálculo dos índices Morris
75 EE_V=[]
76 EE_R=[]
77 EE_L=[]
78 EE_C=[]
79 EE_Vabs=[]
80 EE_Rabs=[]
81 EE_Labs=[]
82 EE_Cabs=[]
83 for i in range (r):
84     V_Morris,R_Morris,L_Morris,C_Morris,D_aux,permut=Morris
    (V,R,L,C,k,p,Delta,J1,J2,B,conjunto)
85     vc_maximo=executa(V_Morris,R_Morris,L_Morris,C_Morris)
86     EE=[]
87     for j in range (k):
88         if D_aux[permut[j]]==1:
89             EE.insert(j,(vc_maximo[permut[j]+1]-vc_maximo[permut[j]])/Delta)
90         else:
91             EE.insert(j,(vc_maximo[permut[j]]-vc_maximo[permut[j]+1])/Delta)
92     EE_V.append(EE[0])
93     EE_R.append(EE[1])
94     EE_L.append(EE[2])
95     EE_C.append(EE[3])
96     EE_Vabs.append(abs(EE[0]))
97     EE_Rabs.append(abs(EE[1]))
98     EE_Labs.append(abs(EE[2]))
99     EE_Cabs.append(abs(EE[3]))
100 V_media=statistics.mean(EE_Vabs)
101 V_desvio=statistics.stdev(EE_V)

```

102	R_media=statistics.mean(EE_Rabs)
103	R_desvio=statistics.stdev(EE_R)
104	L_media=statistics.mean(EE_Labs)
105	L_desvio=statistics.stdev(EE_L)
106	C_media=statistics.mean(EE_Cabs)
107	C_desvio=statistics.stdev(EE_C)

Tabela 11– Script em Python: análise de sensibilidade pelo cálculo dos índices Sobol. Fonte: Autoria Própria.

1	import fileinput
2	import sys
3	import os
4	import pandas as pd
5	import numpy as np
6	from lib_readPL4_py3 import readPL4
7	from Estudo1_Monte_Carlo import executa
8	from Estudo1_Monte_Carlo import replacement
9	import statistics
10	# Leitura do excel com os parametros de entrada para as simulações
11	df_parametros = pd.read_excel ('Dados Estudo 1.xlsx ' , sheet_name='Sobol' , dtype=str)
12	if ((len(df_parametros))%2) != 0:
13	print('Entrar com um valor par para o conjunto de parâmetros de entrada')
14	exit()
15	k=int(len(df_parametros.columns))
16	N=int(len(df_parametros)/2)
17	# Atribuindo as variáveis de entrada
18	V=pd.Series.to_numpy(df_parametros.loc [: , 'V'])
19	R=pd.Series.to_numpy(df_parametros.loc [: , 'R'])
20	L=pd.Series.to_numpy(df_parametros.loc [: , 'L'])
21	C=pd.Series.to_numpy(df_parametros.loc [: , 'C'])
22	# Construção das matrizes
23	V_A=V[:N]
24	R_A=R[:N]
25	L_A=L[:N]
26	C_A=C[:N]
27	V_B=V[N:]
28	R_B=R[N:]
29	L_B=L[N:]
30	C_B=C[N:]
31	ya=np.array(executa(V_A,R_A,L_A,C_A))
32	yb=np.array(executa(V_B,R_B,L_B,C_B))
33	fo2=np.square(np.mean(ya))
34	ya2_fo2=((ya@(np.transpose(ya)))/N)-fo2
35	yc_V=np.array(executa(V_A,R_B,L_B,C_B))
36	yc_R=np.array(executa(V_B,R_A,L_B,C_B))
37	yc_L=np.array(executa(V_B,R_B,L_A,C_B))
38	yc_C=np.array(executa(V_B,R_B,L_B,C_A))
39	# Cálculo índices de Sobol
40	S1_V=((ya@(np.transpose(yc_V)))/N)-fo2)/(ya2_fo2)
41	ST_V=1-(((yb@(np.transpose(yc_V)))/N)-fo2)/(ya2_fo2)
42	S1_R=((ya@(np.transpose(yc_R)))/N)-fo2)/(ya2_fo2)
43	ST_R=1-(((yb@(np.transpose(yc_R)))/N)-fo2)/(ya2_fo2)
44	S1_L=((ya@(np.transpose(yc_L)))/N)-fo2)/(ya2_fo2)
45	ST_L=1-(((yb@(np.transpose(yc_L)))/N)-fo2)/(ya2_fo2)
46	S1_C=((ya@(np.transpose(yc_C)))/N)-fo2)/(ya2_fo2)
47	ST_C=1-(((yb@(np.transpose(yc_C)))/N)-fo2)/(ya2_fo2)

Tabela 12– Script em Python: interpretação dos resultados gerados pelo ATP. Fonte: Adaptado de [45].

1	# Read PISA's binary PL4
2	def readPL4(pl4file):
3	import mmap
4	import struct
5	import pandas as pd
6	import numpy as np
7	miscData = {

8	'deltat':0.0,
9	'nvar':0,
10	'pl4size':0,
11	'steps':0,
12	'tmax':0.0
13	}
14	# open binary file for reading
15	with open(pl4file, 'rb') as f:
16	pl4 = mmap.mmap(f.fileno(), 0, access=mmap.ACCESS_READ)
17	# read DELTAT
18	miscData['deltat'] = struct.unpack('<f', pl4[40:44])[0]
19	# read number of vars
20	miscData['nvar'] = struct.unpack('<L', pl4[48:52])[0] // 2
21	# read PL4 disk size
22	miscData['pl4size'] = struct.unpack('<L', pl4[56:60])[0]-1
23	# compute the number of simulation miscData['steps'] from the PL4's file size
24	miscData['steps'] = (miscData['pl4size'] - 5*16 - miscData['nvar']*16) // \
25	((miscData['nvar']+1)*4)
26	miscData['tmax'] = (miscData['steps']-1)*miscData['deltat']
27	# generate pandas dataframe to store the PL4's header
28	dfHEAD = pd.DataFrame(columns=['TYPE', 'FROM', 'TO'])
29	for i in range(0,miscData['nvar']):
30	pos = 5*16 + i*16
31	h = struct.unpack('3x1c6s6s',pl4[pos:pos+16])
32	xx=pd.DataFrame({'TYPE': int(h[0]),\
33	'FROM': h[1],\
34	'TO': h[2]}, index=[i])
35	dfHEAD = pd.concat([dfHEAD,xx], ignore_index=True)
36	# Correct 'TO' and 'FROM' columns types
37	dfHEAD['FROM'] = dfHEAD['FROM'].str.decode('utf-8')
38	dfHEAD['TO'] = dfHEAD['TO'].str.decode('utf-8')
39	# Check for unexpected rows of zeroes
40	# See https://github.com/ldemattos/readPL4/issues/2
41	expsize = (5 + miscData['nvar'])*16 + miscData['steps']*(miscData['nvar']+1)*4
42	nullbytes = 0
43	if miscData['pl4size'] > expsize:
44	nullbytes = miscData['pl4size']-expsize
45	# read and store actual data, map it to a numpy read only array
46	data = np.memmap(f,dtype=np.float32,mode='r',shape=(miscData['steps'],miscData['nvar']+1),offset=(5 + miscData['nvar'])*16 + nullbytes)
47	return dfHEAD,data,miscData

Apêndice B – Parâmetros e modelagem do estudo de caso 2 - sobretensões devido a descarga atmosférica

Descreve-se neste apêndice o arranjo físico da subestação considerada no Estudo de Caso 2, bem como os detalhes da modelagem computacional.

A Figura 34 ilustra a entrada de dados na rotina do EMTP-RV para cálculo dos parâmetros da linha de transmissão, para as seções cujo vão é 450 m.

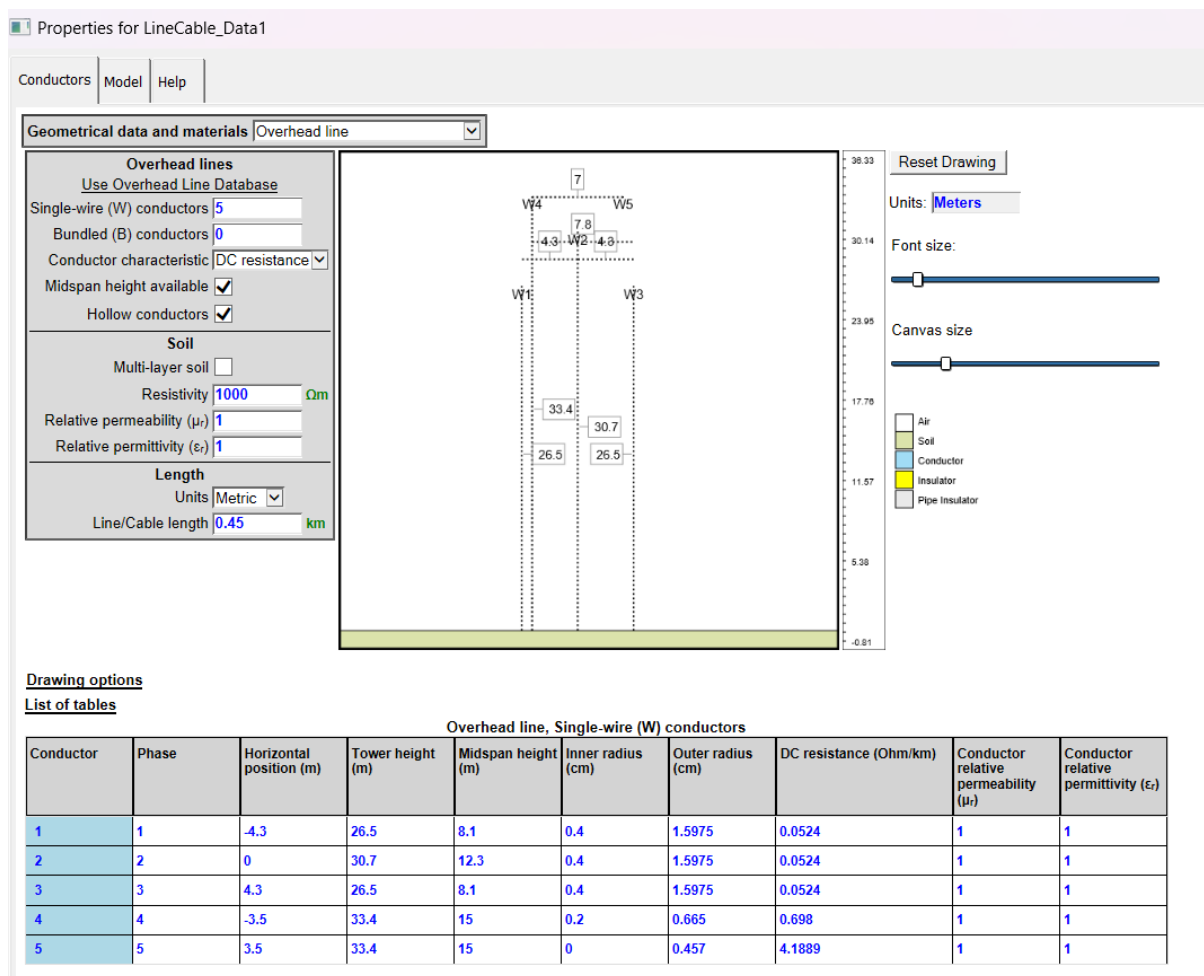


Figura 34 – Entrada de dados da linha de transmissão na rotina do EMTP-RV. Fonte: Adaptado Própria.

A Figura 35 ilustra a entrada de dados no PAMSUITE, bem como a definição das saídas de interesse, para a simulação de Monte Carlo.

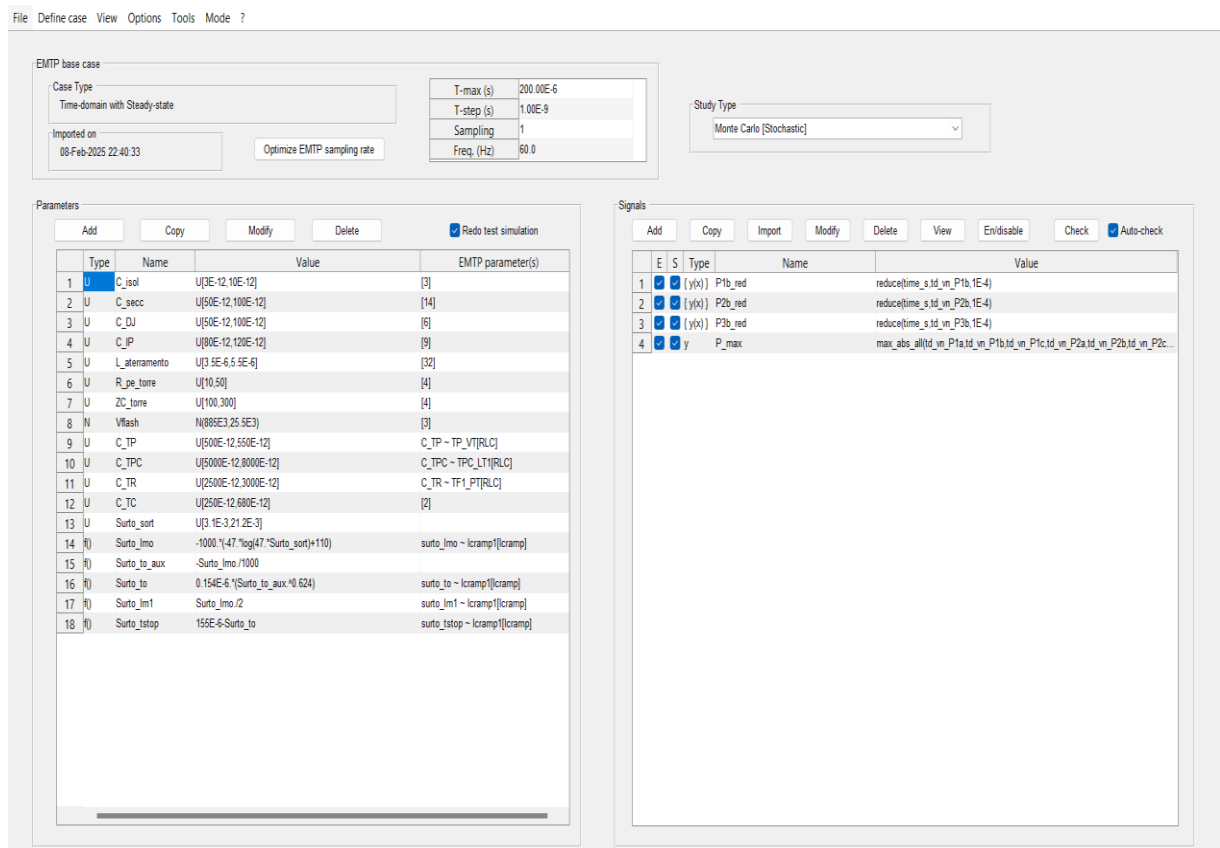
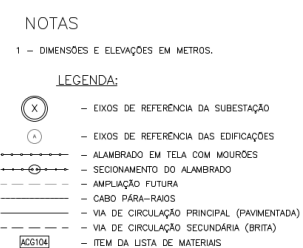
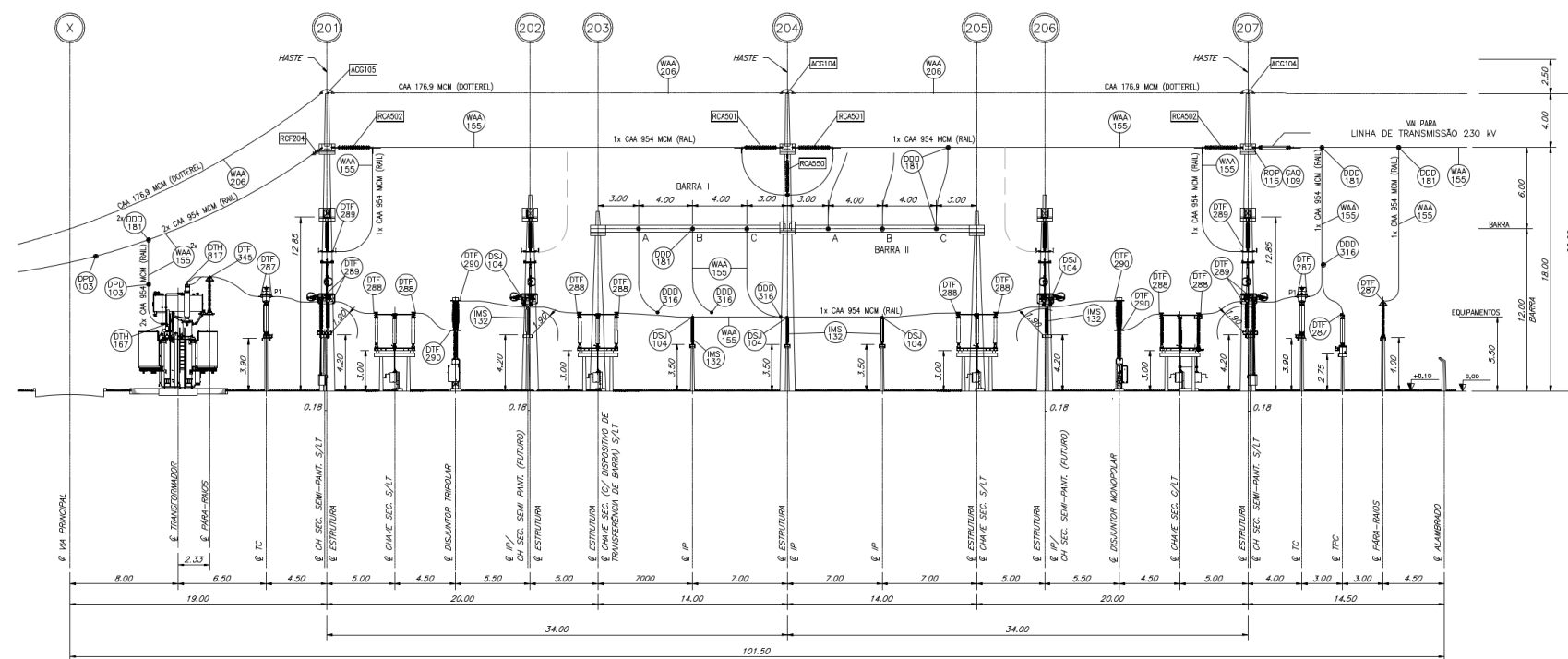


Figura 35 – Entrada de dados no PAMSUITE e definição das saídas de interesse. Fonte: Adaptado Própria.

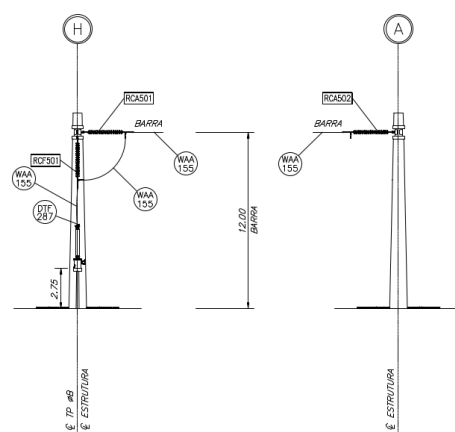
A Figura 36 ilustra o projeto da planta da subestação, enquanto a Figura 37 ilustra os detalhes do corte do setor 230 kV da subestação. Já a Figura 38 mostra toda a modelagem computacional do sistema, realizada em EMTP-RV.



106



CORTE A-A
ESC. 1:200



CORTE C-C
ESC. 1:200

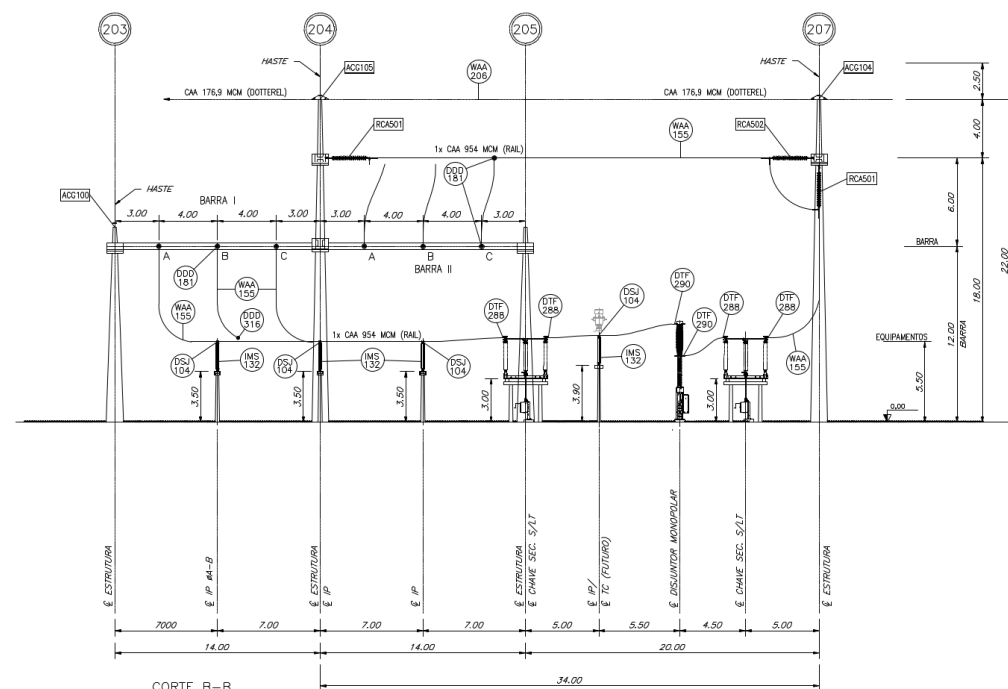
CORTE D-D
ESC. 1:200

NOTAS

1 - DIMENSÕES E ELEVÇÕES EM METROS.

LEGENDA:

- (X) - EIXOS DE REFERÊNCIA DA SUBESTAÇÃO
- (DD 181) - ITEM DA LISTA DE MATERIAIS
- CABO PARA-RAIOS
- ACC104 - ITEM DA LISTA DE MATERIAIS



CORTE B-B
ESC. 1:200

Figura 37 – Corte da subestação. Fonte: Confidencial.

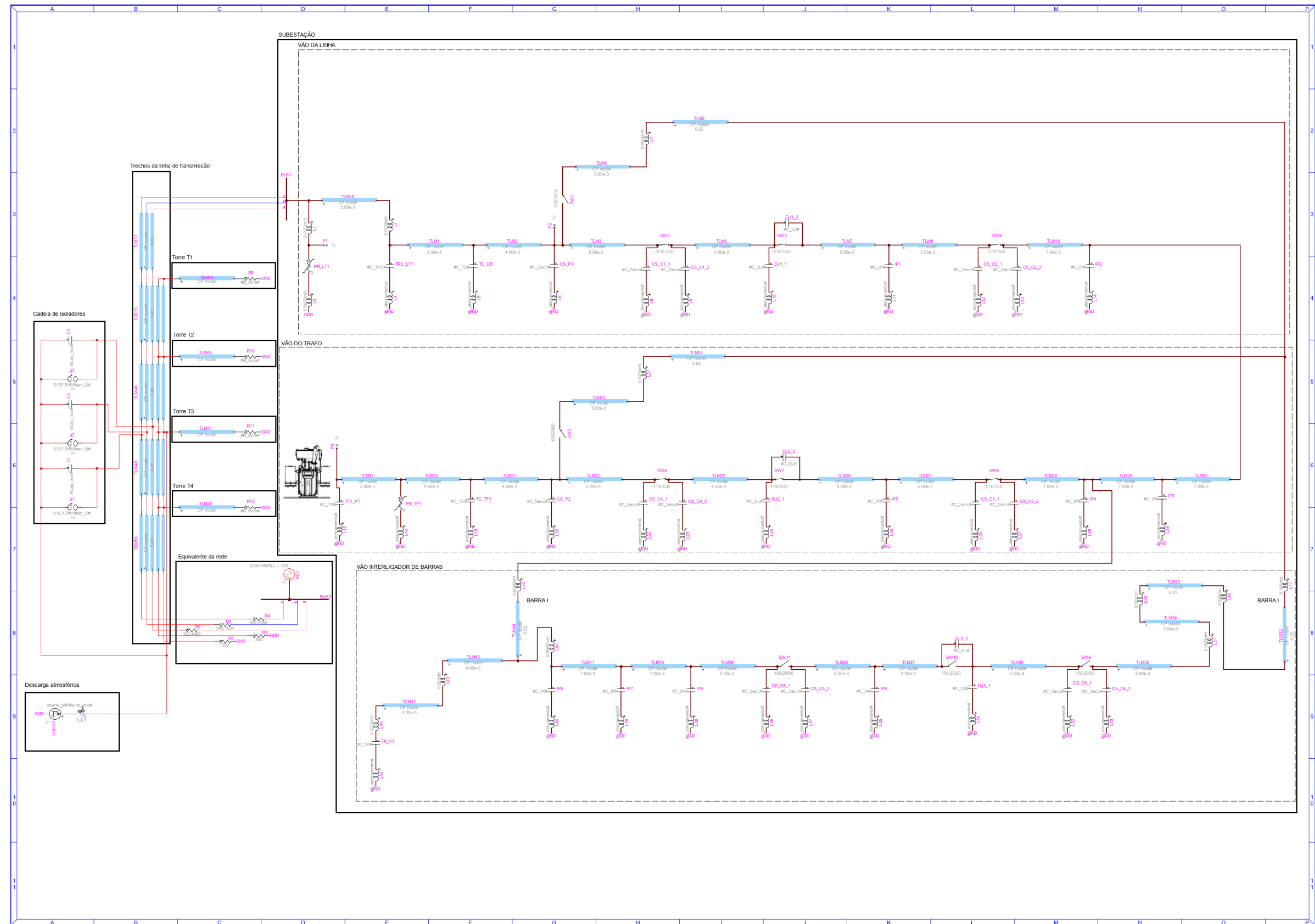


Figura 38 – Modelagem computacional do sistema em EMTP-RV. Fonte: Autoria Própria.