

UNIVERSIDADE FEDERAL DE JUIZ DE FORA
INSTITUTO DE CIÊNCIAS EXATAS / FACULDADE DE ENGENHARIA
PROGRAMA DE PÓS-GRADUAÇÃO EM MODELAGEM
COMPUTACIONAL

Rodrigo Bittencourt Pigozzo

**Análise de sensibilidade de alternâncias cardíacas em modelos celulares de
eletrofisiologia**

Juiz de Fora

2026

Ficha catalográfica elaborada através do Modelo Latex do CDC da UFJF
com os dados fornecidos pelo(a) autor(a)

Pigozzo, Rodrigo Bittencourt.

Análise de sensibilidade de alternâncias cardíacas em modelos celulares de eletrofisiologia / Rodrigo Bittencourt Pigozzo. – 2026.
117 f. : il.

Orientador: Bernardo Martins Rocha

Coorientador: Rodrigo Weber dos Santos

Tese (Doutorado) – Universidade Federal de Juiz de Fora, Instituto de Ciências Exatas / Faculdade de Engenharia. Programa de Pós-Graduação em Modelagem Computacional, 2026.

1. arritmias 2. alternans. 3. análise de sensibilidade. 4. emuladores.
I. M. Rocha, Bernardo, orient. II. Análise de sensibilidade de alternâncias cardíacas em modelos celulares de eletrofisiologia.

Rodrigo Bittencourt Pigozzo

Análise de sensibilidade de alternâncias cardíacas em modelos celulares de eletrofisiologia

Tese apresentada ao Programa de Pós-Graduação em Modelagem Computacional da Universidade Federal de Juiz de Fora, como requisito parcial para obtenção do grau de Doutor em Modelagem Computacional.

Orientador: Bernardo Martins Rocha

Coorientador: Rodrigo Weber dos Santos

Juiz de Fora

2026

Rodrigo Bittencourt Pigozzo

Análise de sensibilidade de alternâncias cardíacas em modelos celulares de eletrofisiologia

Tese apresentada ao Programa de Pós-Graduação em Modelagem Computacional da Universidade Federal de Juiz de Fora como requisito parcial à obtenção do título de Doutor(a) em Modelagem Computacional. Área de concentração: Modelagem Computacional.

Aprovada em 25 de setembro de 2026.

BANCA EXAMINADORA

Prof. Dr. Bernardo Martins Rocha - Orientador

Universidade Federal de Juiz de Fora

Prof. Dr. Grigori Chapiro

Universidade Federal de Juiz de Fora

Prof. Dr. Joventino de Oliveira Campos

Universidade Federal de Juiz de Fora

Prof. Dr. Rafael Sachetto Oliveira

Universidade Federal de São João Del Rei

Prof. Dr. Rafael Alves Bonfim de Queiroz

Universidade Federal de Ouro Preto

Juiz de Fora, 30/09/2026.



Documento assinado eletronicamente por **Bernardo Martins Rocha, Professor(a)**, em 30/09/2026, às 13:08, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Rafael Sachetto Oliveira, Usuário Externo**, em 30/09/2026, às 13:45, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Grigori Chapiro, Professor(a)**, em 30/09/2026, às 14:44, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Joventino de Oliveira Campos, Professor(a)**, em 30/09/2026, às 15:08, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Rafael Alves Bonfim de Queiroz, Usuário Externo**, em 30/09/2026, às 16:41, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).

AGRADECIMENTOS

Primeiramente, gostaria de agradecer a Deus por me dar a vida e a sua graça para continuar trabalhando e me esforçando nos meus estudos todos os dias. Gostaria de agradecer aos meus familiares por todo o apoio neste tempo de vida acadêmica, em especial, gostaria de agradecer à minha esposa por todo o apoio a permanecer neste caminho. Agradeço também a todos os que estudaram comigo, aos que foram colegas de sala e laboratório e de alguma maneira me ajudaram. Agradeço também, principalmente, aos meus orientadores Bernardo e Rodrigo por todo o apoio, conteúdo e oportunidade que me deram neste tempo para realizar meu trabalho. Finalmente, eu gostaria de agradecer a todas as instituições de fomento que contribuem para esta pesquisa: FAPEMIG, CAPES, CNPq e UFJF.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001. Este trabalho também foi financiado pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) sob o número do processo 301120/2025-0.

RESUMO

Arritmias cardíacas são um grave problema de saúde que pode levar à parada cardíaca súbita, à insuficiência cardíaca e, por fim, à morte. O alternans é frequentemente observado em pacientes com arritmias cardíacas, particularmente naquelas que afetam os ventrículos. Trata-se de uma variação, batimento a batimento, na duração ou na amplitude dos potenciais de ação cardíacos, que também pode se refletir em alterações do eletrocardiograma. Compreender os mecanismos que levam à sua ocorrência é, portanto, clinicamente relevante. Nesse esforço, os modelos matemáticos da eletrofisiologia celular são ferramentas fundamentais, por permitirem investigar, de forma controlada, as condições que favorecem o surgimento do alternans. A capacidade de um modelo reproduzir o alternans depende de sua formulação e de seus parâmetros, e ainda são escassos os estudos que identificam, de forma sistemática, quais parâmetros governam a ocorrência desse fenômeno. Determinar esses parâmetros é essencial tanto para compreender os mecanismos do alternans quanto para avaliar a coerência dos próprios modelos. O objetivo principal deste trabalho foi identificar os parâmetros que mais influenciam a ocorrência de alternans, empregando técnicas de análise de sensibilidade, com destaque para a Filtragem de Monte Carlo (MCF), além dos índices de Sobol. Para acelerar essas análises, que exigiriam um número proibitivo de simulações do modelo completo, foram utilizados modelos substitutos, tendo como referência o método de Monte Carlo clássico. Além da Expansão em Caos Polinomial (PCE), foram investigadas duas outras técnicas de modelagem substituta: os Processos Gaussianos (GP) e as Florestas Aleatórias (RF), empregadas tanto na regressão das quantidades de interesse quanto na classificação direta da ocorrência de alternans, com comparação sistemática entre as duas estratégias em termos de precisão e custo computacional. A abordagem foi aplicada a três modelos celulares de complexidade crescente: o modelo simplificado de Mitchell-Schaeffer modificado, o modelo detalhado de ten Tusscher-Panfilov e o modelo de Mahajan-Shiferaw. Os resultados identificaram os parâmetros determinantes para o alternans em cada modelo. Para os modelos biofisicamente detalhados, a condutância da corrente de cálcio do tipo L destacou-se como o parâmetro mais influente, evidenciando a capacidade da abordagem de recuperar, a partir da estrutura interna dos modelos, relações fisiologicamente relevantes, além de identificar interações entre parâmetros não evidentes a priori.

Palavra-chave: arritmias, alternans, índices de Sobol, filtragem de Monte Carlo, expansão em caos polinomial.

ABSTRACT

Cardiac arrhythmias are a serious health problem that can lead to sudden cardiac arrest, heart failure, and ultimately death. Alternans is frequently observed in patients with cardiac arrhythmias, particularly those affecting the ventricles. It consists of a beat-to-beat variation in the duration or amplitude of the cardiac action potential, which may also be reflected in changes in the electrocardiogram. Understanding the mechanisms that lead to its occurrence is therefore clinically relevant. In this effort, mathematical models of cellular electrophysiology are fundamental tools, as they allow the controlled investigation of the conditions that favor the onset of alternans. The ability of a model to reproduce alternans depends on its formulation and its parameters, and studies that systematically identify which parameters govern the occurrence of this phenomenon are still scarce. Determining these parameters is essential both to understand the mechanisms of alternans and to assess the consistency of the models themselves. The main objective of this work was to identify the parameters that most influence the occurrence of alternans, employing sensitivity analysis techniques, with emphasis on Monte Carlo Filtering (MCF), in addition to Sobol indices. To accelerate these analyses, which would otherwise require a prohibitive number of full-model simulations, surrogate models (emulators) were used, with the classical Monte Carlo method serving as reference. In addition to Polynomial Chaos Expansion (PCE), two other surrogate modeling techniques were investigated: Gaussian Processes (GP) and Random Forests (RF), employed both in the regression of the quantities of interest and in the direct classification of alternans occurrence, with a systematic comparison between the two strategies in terms of accuracy and computational cost. The approach was applied to three cell models of increasing complexity: the simplified modified Mitchell-Schaeffer model, the detailed ten Tusscher-Panfilov model, and the Mahajan-Shiferaw model. The results identified the determining parameters for alternans in each model. For the biophysically detailed models, the conductance of the L-type calcium current stood out as the most influential parameter, demonstrating the ability of the approach to recover physiologically relevant relationships from the internal structure of the models, as well as to identify interactions between parameters that are not evident a priori.

Key-words: arrhythmias, alternans, Sobol indices, Monte Carlo filtering, polynomial chaos expansion.

LISTA DE ILUSTRAÇÕES

Figura 1 – Novos implantes de marcapasso por milhão de habitantes em 2005, 2009, 2013, 2017 e 2018, em diferentes regiões e países do mundo. Figura adaptada de Mkoko et al. (2020).	16
Figura 2 – Potencial de ação de célula ventricular humana gerado pelo modelo matemático de Tusscher e Panfilov (2006).	17
Figura 3 – O sistema de condução elétrica do coração. O marca-passo fisiológico impulsi- ona a corrente elétrica a partir de um impulso que inicia no nó sinoatrial.	19
Figura 4 – Indicação de ocorrência de Alternans em medições clínicas do trabalho de Zang e Xia (2014) a partir das oscilações na amplitude e duração do AP, a partir também da alternância na forma da onda-T do ECG e a partir da oscilação na quantidade de cálcio (Ca^{++}), em nível de célula.	19
Figura 5 – Potencial de Ação do modelo MMS com uma corrente aplicada em $t = 100$ ms.	27
Figura 6 – Simulação numérica do potencial de ação do modelo TP resultante após a aplicação de uma corrente externa aplicada em $t = 100$ ms.	29
Figura 7 – Simulação numérica do potencial de ação do modelo Mahajan resultante após a aplicação de uma corrente externa aplicada em $t = 100$ ms.	30
Figura 8 – Potencial de Ação típico do modelo MMS e suas fases: (1) despolarização; (2) platô; (3) repolarização e (4) repouso.	31
Figura 9 – Curva de Restituição ($DI \times APD$) do modelo MMS.	32
Figura 10 – APs no modelo MMS e a ocorrência do alternans, para o seguinte conjunto de parâmetros do modelo MMS: $\tau_{in} = 0,1$, $\tau_{out} = 9,0$, $\tau_{open} = 100$, e $\tau_{close} = 120$	34
Figura 11 – A curva de restituição ($BCL \times APD$) do modelo MMS mostra alternans em um BCL de 450ms. Parâmetros do modelo: $\tau_{in} = 0,1$, $\tau_{out} = 4,3$, $\tau_{open} = 50,1$ e $\tau_{close} = 177,5$	34
Figura 12 – Curva de restituição ($BCL \times APD$) do modelo TP e exemplo da ocorrência de alternans. Parâmetros: $G_{CaL} = 2,1 \times 10^{-5}$, $G_{K1} = 4,060$, $G_{Kr} = 0,094$, $G_{Ks} = 0,210$, $G_{Na} = 14,558$, $G_{bCa} = 5,9 \times 10^{-4}$, $G_{bNa} = 4,1 \times 10^{-4}$, $G_{pCa} =$ $0,135$, $G_{pK} = 0,011$, $G_{to} = 0,184$	35
Figura 13 – Exemplo de protocolo de restituição dinâmico.	36
Figura 14 – Resultados do cálculo dos índices de Sobol de primeira ordem e de ordem total para a saída da função Rosenbrock.	41
Figura 15 – Representação esquemática da Filtragem de Monte Carlo.	42
Figura 16 – Representação gráfica do teste de Smirnov para um parâmetro influente X_i . A barra vertical é a estatística $d_{n,\bar{n}}$. Adaptado de Saltelli et al. (2008).	42

Figura 17 – Representação gráfica dos testes de Smirnov das variáveis X_1 e X_2 para o modelo simples $Y = X_1 + X_2$, onde $Y > 1$ é o comportamento alvo a ser alcançado.	43
Figura 18 – Projeção bidimensional da subamostra B para o modelo simples $Y = X_1 + X_2$. O coeficiente de correlação entre X_1 e X_2 na amostra MC é $cc = -0,53$	44
Figura 19 – Resultados da análise MCF para o Exemplo 2: na parte superior os gráficos das CDFs e na parte inferior as correlações dos parâmetros para os casos B e $\bar{B}$	44
Figura 20 – Esquema ilustrativo da construção e utilização de modelo substituto para análise de sensibilidade.	45
Figura 21 – Estrutura de uma árvore de decisão. Cada nó interno aplica uma condição sobre uma variável de entrada, dividindo o conjunto de dados em dois ramos; as folhas correspondem à classe prevista pelo modelo.	49
Figura 22 – Esquema ilustrativo de uma <i>Random Forest</i> : o vetor de entrada $\mathbf{x}$ é avaliado por um conjunto de B árvores de decisão, cada uma treinada sobre uma amostra <i>bootstrap</i> distinta do conjunto de dados. As previsões individuais das árvores são combinadas por voto majoritário (classificação) ou por média (regressão), resultando na predição final $\hat{y}$	51
Figura 23 – Fluxo para o cálculo dos índices de Sobol. Cada quantidade de interesse, APD ou MaxSlope, é aproximada por um emulador (PCE ou GP), a partir do qual os índices de Sobol são obtidos.	53
Figura 24 – Fluxo para a detecção de alternans via Filtragem de Monte Carlo. A partir do MaxSlope, um regressor aproxima Y e aplica o critério $Y > 1$, ou um classificador prevê a classe diretamente. As técnicas indicadas ao lado de cada ramo mostram que PCE atua apenas como regressor, enquanto GP e RF atuam tanto na regressão quanto na classificação.	54
Figura 25 – Análise da qualidade dos emuladores PCE para os modelos MMS e TP em função do tamanho do conjunto de dados N_{PCE} usado em sua construção: em cima, coeficientes R^2 , e em baixo, erro LOO.	60
Figura 26 – SA local OAT dos parâmetros do modelo MMS para o APD.	61
Figura 27 – SA local OAT dos parâmetros do modelo MMS para o MaxSlope.	61
Figura 28 – SA global pelo cálculo dos Índices de Sobol dos parâmetros do modelo MMS para as QoI: APD e MaxSlope.	62
Figura 29 – Resultados do MCF e os valores da estatística de Smirnov para os parâmetros τ_{in} , τ_{out} , τ_{open} e τ_{close} do modelo MMS obtidos através de simulação computacional do modelo no OpenCARP.	63

Figura 30 – Correlações (cc) de pares de parâmetros para os casos de ocorrência de alternans do modelo MMS. A cor indica a densidade local de pontos, das regiões menos densas (roxo) às mais densas (amarelo); as áreas amarelas correspondem, portanto, às combinações de parâmetros mais frequentes entre as amostras que produzem <i>alternans</i>	64
Figura 31 – Convergência da estatística de Smirnov ($d_{n,n}$) da análise MCF em função do tamanho amostral N_s , para cada parâmetro do modelo MMS na quantidade MaxSlope. As curvas mostram a média e as faixas sombreadas o desvio-padrão sobre repetições da subamostragem; as linhas correspondem, de cima para baixo, a τ_{close} (verde), τ_{out} (vermelho), τ_{in} (preto) e τ_{open} (azul).	65
Figura 32 – Esquerda: histograma do APD obtido para um conjunto de amostras através do modelo MMS e do emulador PCE. Direita: gráfico de paridade dos valores do APD obtidos pelo modelo MMS e pelo emulador PCE.	66
Figura 33 – Lado esquerdo está o histograma do emulador PCE do modelo MMS para o MaxSlope com uma predição boa do modelo original. Lado direito está a amostra, por meio dos gráficos de dispersão, dos valores do MaxSlope do modelo MMS original que estão próximos em comparação aos do modelo emulado pelo PCE.	66
Figura 34 – Resultados do MCF e os valores da estatística de Smirnov para os parâmetros τ_{in} , τ_{out} , τ_{open} e τ_{close} do modelo MMS obtido a partir das expressões analíticas.	68
Figura 35 – Projeções bidimensionais das amostras associadas à ocorrência de <i>alternans</i> para cada par de parâmetros do modelo MMS. A cor indica a densidade local de pontos, das regiões menos densas (roxo) às mais densas (amarelo); as áreas amarelas correspondem, portanto, às combinações de parâmetros mais frequentes entre as amostras que produzem <i>alternans</i>	69
Figura 36 – Análise de sensibilidade local OAT para os parâmetros do modelo TP com relação a quantidade APD.	70
Figura 37 – SA local OAT dos parâmetros do modelo TP para o MaxSlope.	71
Figura 38 – SA global pelo cálculo dos Índices de Sobol dos parâmetros do modelo TP para as QoI: APD e MaxSlope.	72
Figura 39 – Resultados do MCF: CDF das amostras que geram alternans (preto) e não geram alternans (vermelho) para os parâmetros do modelo TP, através das quais se faz o cálculo do Teste de Smirnov.	73
Figura 40 – Correlação entre pares de parâmetros do modelo TP para a ocorrência de <i>alternans</i> . A cor indica a densidade local de pontos, das regiões menos densas (roxo) às mais densas (amarelo). Mostram-se apenas dois casos, dado o tamanho do espaço de parâmetros e o fato de que todos apresentaram correlações muito baixas, inferiores a 0,02.	73

Figura 41 – Histograma (esquerda) e gráfico de paridade (direita) dos valores verdadeiros e os preditos pelo emulador PCE do modelo TP para APD.	74
Figura 42 – Histograma (esquerda) e gráfico de paridade (direita) dos valores verdadeiros e os preditos pelo emulador PCE do modelo TP para MaxSlope.	75
Figura 43 – Experimento I: Comparação dos emuladores PCE (a) e GP (b) no conjunto de dados de validação para a quantidade APD: histograma dos valores verdadeiros e estimados pelo emulador (esquerda) e gráfico de paridade (direita).	86
Figura 44 – Experimento I: Histograma das amostras de MaxSlope após pós-processamento, totalizando 2471 amostras válidas.	87
Figura 45 – Experimento I: Comparação dos emuladores PCE (a) e GP (b) no conjunto de dados de validação para a quantidade MaxSlope: histograma dos valores verdadeiros e estimados pelo emulador (esquerda) e gráfico de paridade (direita).	89
Figura 46 – Experimento I: Funções de distribuição acumulada (CDF) obtidas pela técnica MCF para os parâmetros G_{CaL} (a) e G_{K1} (b) do modelo Mahajan, distinguindo os casos de ocorrência e não ocorrência de alternans.	90
Figura 47 – Experimento II: Comparação dos emuladores PCE (a) e GP (b) no conjunto de dados de validação para a quantidade APD: histograma dos valores verdadeiros e estimados pelo emulador (esquerda) e gráfico de paridade (direita).	93
Figura 48 – Experimento II: Comparação dos emuladores PCE (a) e GP (b) no conjunto de dados de validação para a quantidade MaxSlope: histograma dos valores verdadeiros e estimados pelo emulador (esquerda) e gráfico de paridade (direita).	96
Figura 49 – Cinco maiores índices de Sobol de segunda ordem para o MaxSlope no modelo Mahajan-Shiferaw, obtidos com os emuladores PCE e GP para o caso de 3000 amostras.	101
Figura 50 – Gráfico de radar para as quantidades normalizadas: tempo de construção, tempo de inferência da análise MCF e precisão (BA).	108

LISTA DE TABELAS

Tabela 1	– Valores base dos parâmetros do modelo MMS e o intervalo da distribuição uniforme utilizada para a análise de sensibilidade, em que $\mathcal{U}(a, b)$ denota a distribuição uniforme no intervalo $[a, b]$	26
Tabela 2	– Valores base dos parâmetros do modelo TP e o intervalo de uma distribuição uniforme utilizada para análise de sensibilidade.	28
Tabela 3	– Valores base dos parâmetros do modelo Mahajan e o intervalo de uma distribuição uniforme utilizada para análise de sensibilidade.	30
Tabela 4	– Dados sobre a construção dos emuladores PCE para os modelos MMS e TP.	59
Tabela 5	– Cálculo dos Índices de Sobol e Teste de Smirnov para o modelo MMS através dos emuladores PCE.	67
Tabela 6	– Comparativo entre o cálculo dos Índices de Sobol do modelo MMS e do seu emulador PCE para o APD e o MaxSlope.	67
Tabela 7	– Comparativo entre os valores do Teste de Smirnov para o modelo MMS obtidos de forma analítica, numérica (simulado pelo OpenCARP) e pelo emulador PCE.	69
Tabela 8	– Comparativo entre o cálculo dos coeficientes de correlação (cc) do modelo MMS por meio analítico e pelas simulações numéricas do OpenCARP.	70
Tabela 9	– Índices de Sobol ($S1$ e ST) para o modelo TP com resultados obtidos por simulação numérica e pelos emuladores para APD e MaxSlope.	75
Tabela 10	– Estatística de Smirnov calculada pela técnica MCF para os resultados do modelo TP emulado pelo PCE.	75
Tabela 11	– Comparativo entre os Testes de Smirnov calculados, a partir da técnica MCF, para o modelo TP e do seu emulador PCE.	76
Tabela 12	– Comparativo entre o cálculo dos Índices de Sobol do modelo TP e do seu emulador PCE para o APD e o MaxSlope.	77
Tabela 13	– Comparativo entre os Desempenhos Computacionais dos modelos MMS e TP e seus respectivos emuladores PCE.	78
Tabela 14	– Comparativo entre o cálculo dos Índices de Sobol do modelo TP neste trabalho e do modelo TP abordado no trabalho de Corso e Viola (2020), para o APD.	79
Tabela 15	– Experimento I: Índices de Sobol de Primeira Ordem dos parâmetros do modelo Mahajan, custo computacional e métricas de precisão dos emuladores PCE e GP para a quantidade APD, em função do número de amostras de treinamento.	85

Tabela 16 – Experimento I: Índices de Sobol de Primeira Ordem e Totais dos parâmetros do modelo Mahajan, custo computacional e métricas de precisão dos emuladores PCE e GP para a quantidade MaxSlope, em função do número de amostras de treinamento.	88
Tabela 17 – Experimento I: Estatística de Smirnov (KS) obtida pelo método de filtragem de Monte Carlo com emuladores PCE e GP para a quantidade MaxSlope, em função do número de amostras de treinamento.	89
Tabela 18 – Experimento I: Estatística de Smirnov (KS) obtida pelo método de filtragem de Monte Carlo com o classificador RF para a quantidade MaxSlope, em função do número de amostras de treinamento.	91
Tabela 19 – Experimento II: Índices de Sobol de Primeira Ordem dos parâmetros do modelo Mahajan, custo computacional e métricas de precisão dos emuladores PCE e GP para a quantidade APD, em função do número de amostras de treinamento.	92
Tabela 20 – Experimento II: Índices de Sobol de Primeira Ordem e Totais dos parâmetros do modelo Mahajan, custo computacional e métricas de precisão dos emuladores PCE e GP para a quantidade MaxSlope, em função do número de amostras de treinamento.	95
Tabela 21 – Experimento II: Estatística de Smirnov (KS) obtida pelo método de filtragem de Monte Carlo utilizando emuladores PCE e GP, em função do número de amostras de treinamento.	97
Tabela 22 – Experimento II: Estatística de Smirnov (KS) obtida pelo método de Filtragem de Monte Carlo com o classificador RF para a quantidade MaxSlope, em função do número de amostras de treinamento.	97
Tabela 23 – Experimento III: Índices de Sobol de Primeira Ordem dos parâmetros do modelo Mahajan, custo computacional e métricas de precisão dos emuladores PCE e GP para a quantidade APD, em função do número de amostras de treinamento.	98
Tabela 24 – Experimento III: Índices de Sobol de Primeira Ordem e Totais dos parâmetros do modelo Mahajan, custo computacional e métricas de precisão dos emuladores PCE e GP para a quantidade MaxSlope, em função do número de amostras de treinamento.	99
Tabela 25 – Experimento III: Estatística de Smirnov (KS) obtida pelo método de Filtragem de Monte Carlo utilizando emuladores PCE e GP, em função do número de amostras de treinamento.	100
Tabela 26 – Experimento III – Estatística de Smirnov (KS) obtida pelo método de Filtragem de Monte Carlo com o classificador RF para a quantidade MaxSlope, em função do número de amostras de treinamento.	101

Tabela 27 – Comparativo dos coeficientes de determinação R^2 obtidos com 3000 amostras de treinamento para os emuladores PCE e GP das quantidades APD e MaxSlope do modelo Mahajan em função do número de parâmetros.	102
Tabela 28 – Comparativo da acurácia balanceada (BA) para os emuladores PCE, GP e o classificador RF no Experimento III (13 parâmetros) em função do número de amostras de treinamento – modelo Mahajan.	103
Tabela 29 – Comparativo dos tempos de construção (em segundos) dos emuladores PCE, GP e do classificador RF em função do número de amostras de treinamento.	104
Tabela 30 – Comparativo dos tempos de inferência MCF (em segundos) dos emuladores PCE, GP e do classificador RF em função do número de amostras de treinamento	104
Tabela 31 – Acurácia balanceada (BA) para os cinco métodos no Experimento III (13 parâmetros), em função do número de amostras de treinamento – quantidade MaxSlope, modelo Mahajan.	105
Tabela 32 – Coeficiente de determinação R^2 para os regressores PCE, GP e RF da quantidade MaxSlope no Experimento III (13 parâmetros) em função do número de amostras de treinamento.	106
Tabela 33 – Tempos de construção e inferência MCF (em segundos) para os cinco métodos no Experimento III (13 parâmetros), em função do número de amostras de treinamento – quantidade MaxSlope, modelo Mahajan.	106

LISTA DE ABREVIATURAS E SIGLAS

AP	<i>Potencial de Ação</i>
APD	<i>Duração do Potencial de Ação</i>
AUC	<i>Área sob a Curva ROC</i>
BCL	<i>Comprimento do Ciclo Básico</i>
CDF	<i>Função de Distribuição Acumulada</i>
DI	<i>Intervalo Diastólico</i>
ECG	<i>Eletrocardiograma</i>
EDO	<i>Equação Diferencial Ordinária</i>
GP	<i>Processos Gaussianos</i>
MaxSlope	<i>Máxima Inclinação da Curva de Restituição</i>
MC	<i>Monte Carlo</i>
MCF	<i>Filtragem de Monte Carlo</i>
MMS	<i>Mitchell-Schaeffer Modificado</i>
MS	<i>Mitchell-Schaeffer</i>
MSC	<i>Morte Súbita Cardíaca</i>
OAT	<i>Análise de Sensibilidade Um-de-Cada-Vez</i>
PCE	<i>Expansão em Caos Polinomial</i>
QoI	<i>Quantidade de Interesse</i>
RF	<i>Florestas Aleatórias</i>
SA	<i>Análise de Sensibilidade</i>
TP	<i>Ten Tusscher-Panfilov</i>

SUMÁRIO

1	INTRODUÇÃO	16
1.1	Motivação	16
1.2	Potencial de ação e o sistema de condução elétrica do coração	17
1.3	Alternans elétrico	18
1.4	Análise de sensibilidade de modelos cardíacos	20
1.5	Objetivos	22
1.6	Organização do texto	22
2	MODELOS DA ELETROFISIOLOGIA CARDÍACA	24
2.1	Revisão da literatura	24
2.2	Modelo de Mitchell e Schaeffer (MS)	25
2.3	Modelo de Mitchell e Schaeffer Modificado (MMS)	25
2.4	Modelo de Tusscher-Panfilov (TP)	26
2.5	Modelo de Mahajan-Shiferaw	28
2.6	Conceitos de eletrofisiologia cardíaca	29
2.6.1	Duração do potencial de ação	31
2.6.2	Curva de restituição	32
2.6.3	Alternans elétrico	33
2.6.4	Protocolo de restituição	35
3	ANÁLISE DE SENSIBILIDADE E MODELOS SUBSTITUTOS	37
3.1	Análise de Sensibilidade	37
3.1.1	Análise de sensibilidade local	38
3.1.2	Análise de sensibilidade global baseada nos índices de Sobol	38
3.1.3	Exemplo índices de Sobol	40
3.2	Filtragem de Monte Carlo (MCF)	40
3.2.1	Exemplo MCF 1	42
3.2.2	Exemplo MCF 2	43
3.3	Modelos Substitutos e aprendizado de máquina	44
3.3.1	Expansão em Caos Polinomial (PCE)	46
3.3.2	Processos Gaussianos (GP)	47
3.3.3	Random Forests (RF)	48
4	MÉTODOS	52
4.1	Quantidades de interesse	52
4.2	Índices de Sobol	53
4.3	Filtragem de Monte-Carlo	53
4.4	Métricas	54
4.4.1	Métricas de precisão na regressão	54
4.4.2	Métricas de precisão na classificação	55

4.4.3	Implementação Computacional	56
5	RESULTADOS PARA OS MODELOS MMS E TP	57
5.1	Descrição dos experimentos	57
5.1.1	Escolha dos parâmetros	58
5.1.2	Amostragem	58
5.1.3	Construção dos emuladores PCE	58
5.2	Modelo MMS	59
5.2.1	Análise de sensibilidade local	60
5.2.2	Índices de Sobol	60
5.2.3	Filtragem de Monte Carlo	62
5.2.4	Resultados com emuladores PCE do modelo MMS	65
5.2.5	Verificação e comparativo das análises com a fórmula analítica do modelo MMS	67
5.3	Modelo TP	70
5.3.1	Índices de Sobol	71
5.3.2	Filtragem de Monte Carlo	72
5.3.3	Resultados com emuladores PCE	72
5.3.4	Comparativo das Análises do Modelo TP	76
5.4	Desempenho computacional	76
5.5	Discussões	78
5.5.1	Verificação dos experimentos para o modelo MMS	78
5.5.2	Comparação com a literatura	78
5.5.3	Sobre os modelos e seus parâmetros	79
5.5.4	Comparativo das técnicas de análise de sensibilidade	80
6	ANÁLISE DE SENSIBILIDADE DO MODELO MAHAJAN-SHIFERAW	83
6.1	Descrição dos experimentos	83
6.1.1	Experimento I - 8 parâmetros	84
6.1.1.1	Análise para APD	84
6.1.1.2	Análise para MaxSlope com PCE e GP	85
6.1.1.3	Análise para MaxSlope com Random Forest	90
6.1.2	Experimento II - 11 parâmetros	91
6.1.2.1	Análise para APD	91
6.1.2.2	Análise para MaxSlope com PCE e GP	94
6.1.2.3	Análise para MaxSlope com RF	94
6.1.3	Experimento III - 13 parâmetros	96
6.1.3.1	Análise para APD	97
6.1.3.2	Análise para MaxSlope com PCE e GP	98
6.1.3.3	Índices de Sobol de segunda ordem	100
6.2	Comparativo entre os experimentos	102

6.2.1	Qualidade dos emuladores	102
6.2.2	Desempenho computacional	103
6.2.3	Consistência da análise de sensibilidade	103
6.3	Comparativo entre regressores e classificadores	104
6.3.1	Qualidade de classificação	105
6.3.2	Qualidade de regressão	106
6.3.3	Desempenho computacional	106
6.3.4	Discussão integrada dos resultados	107
6.4	Discussão e posicionamento na literatura	108
7	CONCLUSÕES	110
7.1	Publicações	111
7.2	Limitações e Trabalhos Futuros	112
	REFERÊNCIAS	113

1 INTRODUÇÃO

Neste capítulo apresenta-se uma breve introdução com a motivação do estudo realizado, alguns conceitos básicos sobre fisiologia cardíaca necessários para estudar o fenômeno de interesse deste trabalho e a organização do texto.

1.1 Motivação

De acordo com dados da Sociedade Brasileira de Arritmias Cardíacas (SOBRAC), as doenças cardíacas afetam cerca de 20 milhões de pessoas no Brasil e é a causa de mais de 320 mil mortes súbitas (MSC) por ano no país. Um coração saudável apresenta entre 50 a 100 batimentos por minuto a depender de fatores como sexo, idade e prática desportiva, no entanto, os estímulos elétricos que dão origem ao batimento podem ter várias alterações de ritmo. A Figura 1 ilustra um pouco a preocupação com os casos de doenças cardiovasculares e arritmias aumentando no mundo, pois mostra o aumento do número de casos de implante de marca-passo por milhão de habitantes em algumas regiões e países do mundo, mostrado por Mkoko et al. (2020).

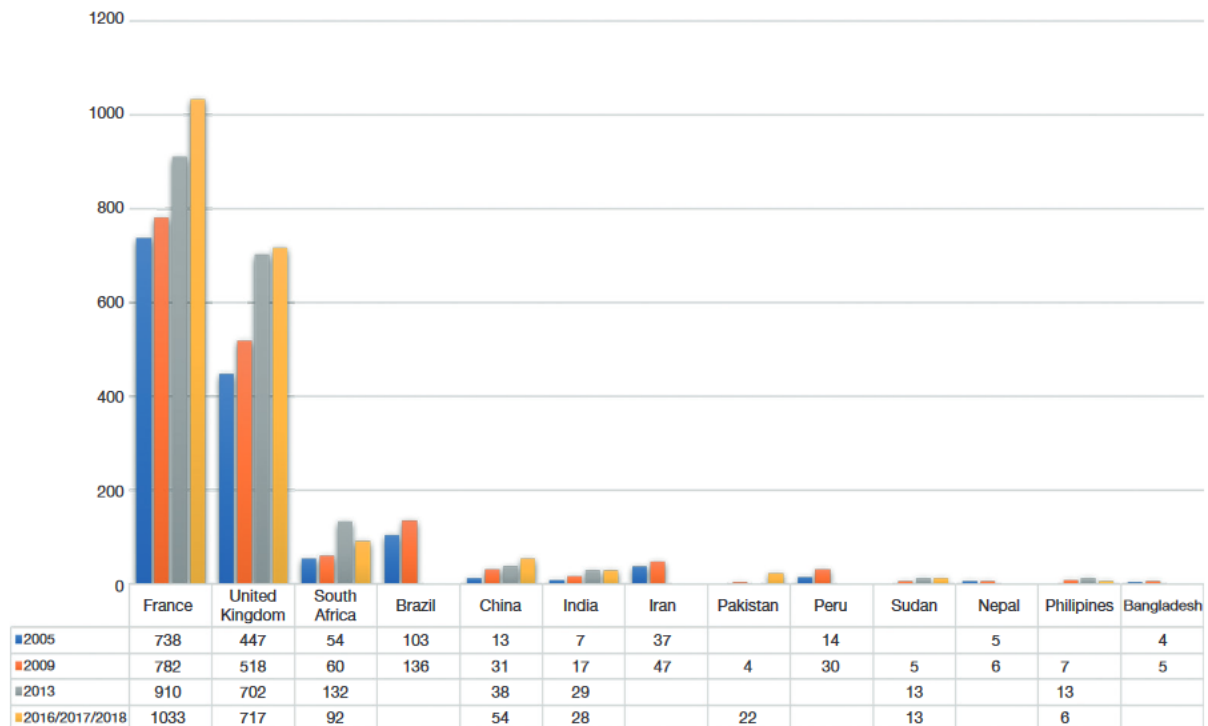


Figura 1 – Novos implantes de marcapasso por milhão de habitantes em 2005, 2009, 2013, 2017 e 2018, em diferentes regiões e países do mundo. Figura adaptada de Mkoko et al. (2020).

Arritmias cardíacas são sequências de batimentos cardíacos irregulares, que afetam o ritmo do coração, fazendo com que ele bata muito rápido, lento ou irregularmente, levando a sérios problemas de saúde, incluindo parada cardíaca súbita, insuficiência cardíaca e

morte. A morte súbita cardíaca (MSC) e a arritmia representam um importante problema de saúde pública mundial, sendo responsáveis por uma porcentagem alta de todas as mortes do mundo. Alterações no sistema de condução elétrica do coração relacionadas à idade aumentam a probabilidade de aparecimento de certas arritmias. Apesar dos avanços recentes na prevenção da MSC, sua incidência na população em geral permaneceu inaceitavelmente alta, como se observa no modelo de Shen e Zipes (2014).

As arritmias nem sempre são percebidas diretamente pelo paciente, que muitas vezes nota apenas suas consequências, como fraqueza, tontura ou desmaio. O diagnóstico pode ser feito por meio do eletrocardiograma (ECG) ou de outras ferramentas clínicas, como o estudo eletrofisiológico. O ECG é o principal exame para detectar arritmias e determinar suas causas You et al. (2021).

1.2 Potencial de ação e o sistema de condução elétrica do coração

O potencial de ação (AP, do inglês *action potential*) no coração é um fenômeno eletrofisiológico essencial que ocorre nas células cardíacas, particularmente nos cardiomiócitos (células musculares cardíacas). O AP consiste em uma rápida variação do potencial transmembrânico, desencadeada pela troca de íons (principalmente sódio, potássio e cálcio) através de canais iônicos específicos localizados na membrana celular. A Figura 2 apresenta o potencial transmembrânico em função do tempo e a trajetória de um potencial de ação típico de uma célula ventricular humana, gerado pelo modelo matemático proposto por Tusscher e Panfilov (2006).

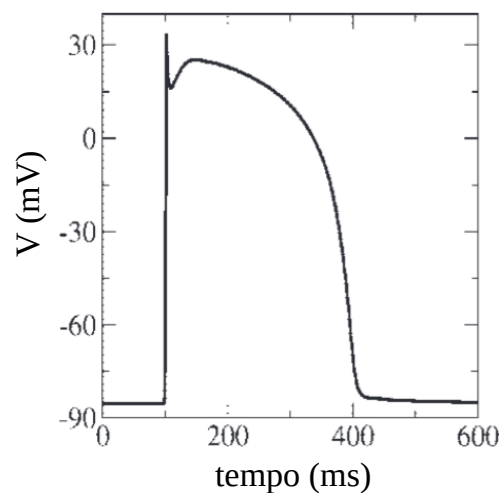


Figura 2 – Potencial de ação de célula ventricular humana gerado pelo modelo matemático de Tusscher e Panfilov (2006).

A propagação desse impulso elétrico pelo tecido cardíaco depende da interconexão funcional entre as células, mediada pelos discos intercalares, estruturas especializadas que contêm junções comunicantes (ou *gap junctions*). Essas junções permitem a transferência direta de íons e outros pequenos sinais elétricos entre as células adjacentes, assegurando a

continuidade do AP de uma célula para outra. Dessa forma, o impulso elétrico gerado em uma célula rapidamente se propaga pelas células vizinhas, criando um efeito de sincronização que é fundamental para a contração coordenada.

Esse processo garante que o AP iniciado no nó sinoatrial, que atua como o marca-passo do coração, se propague eficientemente através do sistema de condução e atinja os ventrículos. A sincronização das contrações resultantes nos átrios e ventrículos é crucial para o controle adequado da frequência e do ritmo cardíaco, permitindo o bombeamento eficiente do sangue para todo o organismo.

O pulso elétrico iniciado em cada batimento cardíaco tem a sua origem no marca-passo natural do coração, chamado nó sinoatrial, localizado no alto do átrio direito, conforme ilustrado na Figura 3. A frequência cardíaca é determinada pela frequência com que o marca-passo fisiológico gera este impulso.

A Figura 3 mostra o nó sinoatrial, que origina o pulso elétrico que resulta em contração e é transmitido aos átrios direito e esquerdo, fazendo com que estes se contraiam. Quando o impulso elétrico alcança o nó atrioventricular, ele sofre um pequeno retardo em sua propagação. O impulso então percorre o feixe de His, onde se divide entre o ramo direito para o ventrículo direito e o ramo esquerdo para o ventrículo esquerdo. O impulso propaga-se então pelos ventrículos, provocando a contração e, assim, bombeando o sangue do ventrículo direito para os pulmões e do ventrículo esquerdo para o restante do corpo.

1.3 Alternans elétrico

Alternans elétrico é um fenômeno no qual há uma variação, batimento-a-batimento, na duração ou amplitude dos potenciais de ação dos miócitos cardíacos. Alternans também pode se manifestar como uma alternância, batimento-a-batimento, da onda-T no ECG, como discutido por You et al. (2021), e pode ser identificado como um precursor de arritmias ventriculares e MSC, apresentado por Rosenbaum et al. (1994). Na Figura 4 pode-se observar exemplos deste comportamento alternante que pode se manifestar no ECG, tanto na amplitude quanto na duração do potencial de ação e, também, no transiente de cálcio intracelular.

É importante ressaltar que a Figura 4 apresenta dados experimentais e medições clínicas, e nos indica que, de fato, o alternans elétrico pode ser um precursor de arritmias perigosas, como fibrilação ventricular, mostrado nos trabalhos de Rosenbaum et al. (1994), Pastore et al. (1999) e Wilson e Rosenbaum (2007). Não desconsiderando o fato de que mudanças nas propriedades do nível celular também podem refletir em variações na amplitude ou no formato da onda-T do ECG, como colocado por You et al. (2021). Portanto, identificar e monitorar alteração elétrica em pacientes com distúrbios cardíacos é de grande importância pois pode influenciar na prevenção de arritmias cardíacas.

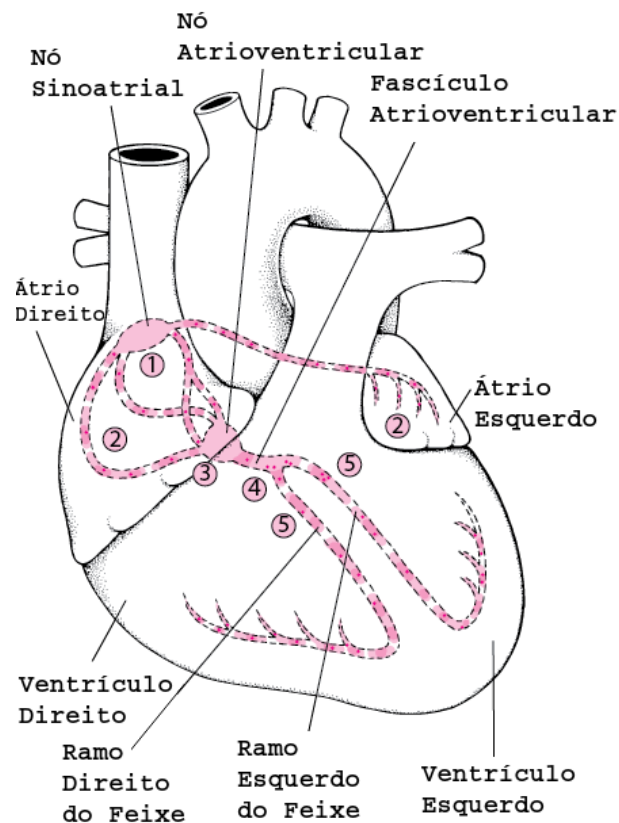


Figura 3 – O sistema de condução elétrica do coração. O marca-passo fisiológico impulsiona a corrente elétrica a partir de um impulso que inicia no nó sinoatrial.



Figura 4 – Indicação de ocorrência de Alternans em medições clínicas do trabalho de Zang e Xia (2014) a partir das oscilações na amplitude e duração do AP, a partir também da alternância na forma da onda-T do ECG e a partir da oscilação na quantidade de cálcio (Ca^{++}), em nível de célula.

Do ponto de vista matemático e computacional, diversos estudos apresentados por Krogh-Madsen e Christini (2007), Elshrif e Cherry (2014), Eastman et al. (2016), Jeong, Marcellinus e Lim (2022) analisaram a ocorrência do fenômeno alternans por meio de simulações computacionais de modelos celulares do potencial de ação. Estudos computacionais baseados em simulações da atividade elétrica do tecido cardíaco também já foram realizados visando elucidar e compreender melhor a ocorrência de alternans, como nos modelos de Tusscher e Panfilov (2006), Walmsley et al. (2015).

Entretanto, é importante ressaltar que muitos modelos matemáticos para eletrofisiologia de miócitos cardíacos não conseguem reproduzir esse fenômeno, enquanto outros o fazem, como por exemplo os modelos apresentados por Elshrif e Cherry (2014) e Eastman et al. (2016). Nesse contexto, faltam estudos na literatura que indiquem quais dos parâmetros desses modelos específicos mais contribuem para a ocorrência do alternans. No trabalho realizado por Jeong, Marcellinus e Lim (2022), uma abordagem baseada em modelagem populacional foi usada para descobrir qual canal iônico teve o maior impacto na ocorrência de alternans. Eles descobriram que as variações na corrente de platô de cálcio (Ca^{++}) do modelo investigado foram altamente influentes. Entretanto, a abordagem utilizada carece de um procedimento mais sistemático baseado em técnicas de análise de sensibilidade global.

1.4 Análise de sensibilidade de modelos cardíacos

A análise de sensibilidade (SA, do inglês *sensitivity analysis*) avalia como a variabilidade de algum parâmetro de entrada influencia a variabilidade de quantidades de saída de modelos, identificando parâmetros-chave que afetam o seu comportamento. As técnicas de SA podem ser divididas em local e global. A análise de sensibilidade local examina essa relação ao redor de um ponto específico no espaço de entrada, geralmente calculando derivadas parciais para determinar a sensibilidade. Ela é eficiente, mas pode ser limitada a regiões restritas do espaço de parâmetros. Em contraste, a análise de sensibilidade global considera o impacto das incertezas nas entradas em toda a sua faixa de variação, proporcionando uma visão mais abrangente. Pode-se citar métodos globais, como os índices de Sobol, propostos por Saltelli et al. (2008), que capturam interações entre os parâmetros, sendo mais indicados para sistemas complexos e não lineares. Pode-se dizer que os índices de Sobol oferecem uma análise quantitativa detalhada da contribuição dos parâmetros para a variância da saída. Entretanto, os índices de Sobol não classificam os parâmetros com base em seu impacto em comportamentos ou categorias específicas do sistema, o que pode ser de interesse para estudar determinados fenômenos ou respostas de modelos.

A técnica de Filtragem de Monte Carlo (MCF, do inglês *Monte Carlo filtering*), proposta por Saltelli et al. (2008), foca na classificação de resultados, identificando os

parâmetros que determinam a ocorrência de um comportamento de interesse. O MCF baseia-se na divisão dos resultados em categorias (como, por exemplo, sucesso ou falha) e, em seguida, na avaliação da influência de cada parâmetro de entrada na probabilidade de ocorrência de cada categoria.

Diversas técnicas de análise de sensibilidade têm sido aplicadas no contexto da eletrofisiologia cardíaca a nível celular, como nos trabalhos de Britton et al. (2013), Geneser, Kirby e Sachse (2006), Groenendaal et al. (2015), Niederer et al. (2009), Novaes et al. (2022), Pathmanathan, Cordeiro e Gray (2019) e em nível de simulações de tecido e órgão, como em Barone, Fenton e Veneziani (2017), Corso e Viola (2020), Pagani e Manzoni (2021), Sánchez et al. (2018), Campos et al. (2020). Entretanto, destaca-se que poucos trabalhos abordaram até então o comportamento de alternans através de análise de sensibilidade, sendo abordado apenas no trabalho de Jeong, Marcellinus e Lim (2022). Por outro lado, também observa-se que poucos trabalhos utilizaram a técnica MCF no contexto da eletrofisiologia cardíaca, sendo o trabalho de Pathmanathan, Cordeiro e Gray (2019) um desses exemplos, o qual utilizou o MCF para analisar a ocorrência de determinados formatos (*spike and dome*) no potencial de ação.

Além disso, vale ressaltar que em estudos de quantificação de incertezas e de análise de sensibilidade, nos quais o número de execuções do modelo é muito alto e pode tornar o estudo computacionalmente caro, é comum o uso de emuladores ou modelos substitutos como alternativas para acelerar essas análises. Em particular, os emuladores baseados na Expansão em Caos Polinomial (PCE, do inglês *polynomial chaos expansion*) têm sido utilizados com sucesso neste contexto, como se pode ver novamente nos trabalhos de Campos et al. (2023), Geneser, Kirby e Sachse (2006), Novaes et al. (2022), Campos et al. (2020).

Além do PCE, outras técnicas de modelagem substituta têm sido empregadas na aceleração de análises de sensibilidade e na quantificação de incertezas. Os Processos Gaussianos (GP, do inglês *Gaussian processes*), também conhecidos como *Kriging*, constroem uma distribuição de probabilidade sobre funções e fornecem, além da predição, uma estimativa de sua incerteza (RASMUSSEN; WILLIAMS, 2005). As Florestas Aleatórias (RF, do inglês *random forests*), por sua vez, baseiam-se em conjuntos de árvores de decisão e destacam-se pela robustez e pela capacidade de lidar com relações não lineares entre os parâmetros (BREIMAN, 2001; BREIMAN et al., 1984). Diferentemente do PCE, essas duas técnicas podem ser empregadas não apenas na regressão de quantidades contínuas, mas também diretamente na classificação de comportamentos do modelo, o que as torna particularmente adequadas à Filtragem de Monte Carlo.

1.5 Objetivos

O objetivo principal deste trabalho é aplicar técnicas de análise de sensibilidade a conjuntos de parâmetros de diferentes modelos da eletrofisiologia cardíaca celular, acelerando essas análises por meio de modelos substitutos (emuladores), para identificar quais parâmetros têm maior influência na ocorrência, ou não, de alternans da duração do potencial de ação. Para isso, são investigadas três técnicas de modelagem substituta (a Expansão em Caos Polinomial, PCE; os Processos Gaussianos, GP; e as Florestas Aleatórias, RF), empregadas tanto na regressão das quantidades de interesse quanto na classificação direta da ocorrência de alternans, e seus desempenhos são comparados de forma sistemática.

Como objetivos específicos, este trabalho se propõe a:

- analisar três modelos celulares de complexidade crescente, sendo um deles simplificado (fenomenológico) e dois biofisicamente detalhados;
- aplicar técnicas clássicas de análise de sensibilidade, como a local do tipo *One-At-a-Time* e os índices de Sobol a quantidades relacionadas a alternans;
- empregar a Filtragem de Monte Carlo como técnica de análise de sensibilidade regional e à identificação dos parâmetros que determinam a ocorrência de alternans;
- acelerar essas análises com emuladores de regressão baseados em Expansão em Caos Polinomial, Processos Gaussianos e Florestas Aleatórias, tanto para os índices de Sobol quanto para a Filtragem de Monte Carlo;
- empregar os Processos Gaussianos e as Florestas Aleatórias também como classificadores, guiando a Filtragem de Monte Carlo pela predição direta da ocorrência de alternans;
- comparar as estratégias de regressão e de classificação em precisão e custo computacional.

1.6 Organização do texto

O restante deste documento está organizado da seguinte forma. No Capítulo 2, apresentam-se os três modelos matemáticos para a geração do potencial de ação em miócitos cardíacos avaliados neste trabalho, além de outros conceitos fundamentais da eletrofisiologia cardíaca. No Capítulo 3, apresentam-se de forma geral as técnicas de análise de sensibilidade (análise local, índices de Sobol e Filtragem de Monte Carlo) e os modelos substitutos empregados (PCE, GP e RF). No Capítulo 4, descrevem-se os métodos empregados: as quantidades de interesse adotadas (APD e MaxSlope), as técnicas de análise de sensibilidade (a análise local do tipo *One-At-a-Time*, os índices de Sobol e a Filtragem de Monte Carlo), os modelos substitutos utilizados (PCE, GP e RF) e as

métricas de avaliação. Os dois capítulos seguintes apresentam os resultados, organizados em duas etapas. O Capítulo 5 aplica a abordagem baseada em emuladores PCE aos modelos MMS e TP, avaliando a viabilidade da metodologia. O Capítulo 6 estende a abordagem ao modelo Mahajan-Shiferaw e às técnicas GP e RF, tanto na regressão quanto na classificação, e compara sistematicamente as diferentes estratégias. Por fim, no Capítulo 7, apresentam-se as conclusões e as possibilidades de trabalhos futuros.

2 MODELOS DA ELETROFISIOLOGIA CARDÍACA

Este capítulo apresenta os modelos da eletrofisiologia cardíaca no nível celular que serão utilizados para estudos de análise de sensibilidade para ocorrência de alternans neste trabalho.

2.1 Revisão da literatura

O progresso no tratamento de doenças cardíacas, especialmente fibrilação ventricular ou arritmia, requer uma compreensão do comportamento elétrico de células cardíacas. Do ponto de vista da modelagem matemática, as primeiras extensões das equações de Hodgkin–Huxley para células cardíacas foram introduzidas para a fibra de Purkinje no trabalho de Noble (1962). Posteriormente, Beeler e Reuter (1977) introduziram o primeiro modelo para a dinâmica de um miócito ventricular (fibra miocárdica), também baseado na formulação de Hodgkin e Huxley.

Neste capítulo, apresenta-se os modelos matemáticos para a geração do AP em miócitos cardíacos que podem descrever alternans elétricos em termos do APD. Existe uma vasta literatura sobre modelos de células cardíacas para geração de potenciais de ação baseados em sistemas de equações diferenciais ordinárias (EDOs) com diferentes graus de complexidade. Os modelos variam desde o modelo clássico de Hodgkin e Huxley (1952) e o modelo de Noble (1962) até modelos ventriculares humanos mais recentes como o desenvolvido por Tomek et al. (2019) e conhecido como ToR-ORD.

Em geral, pode-se distinguir os modelos entre os qualitativos e os biofísicos detalhados. Um exemplo do primeiro caso é o modelo Mitchell e Schaeffer (2003) (MS) que consiste em duas EDOs e é muito semelhante ao modelo clássico de FitzHugh (1955), Nagumo, Arimoto e Yoshizawa (1962). Uma característica particularmente interessante do modelo MS, devido à sua simplicidade matemática, é a possibilidade de analisar sem recorrer a simulações numéricas. O modelo permite a representação de um mapa unidimensional para análise do comportamento de alternans e permite a classificação do espaço de parâmetros para o diferenciar entre cenários, com e sem alternans. Esta abordagem de classificação foi introduzida no trabalho inicial de Mitchell e Schaeffer (2003) (modelo MS original). Entre os modelos biofísicamente detalhados e quantitativos de células cardíacas, um dos modelos mais utilizados é o proposto por Tusscher e Panfilov (2006) (TP) para células ventriculares humanas. Além destes, o modelo proposto por Mahajan e Shiferaw (2008) para células ventriculares de coelho, igualmente biofísicamente detalhado e que é capaz de representar a ocorrência de alternans, também foi explorado neste trabalho. Estes três modelos, portanto, serão utilizados neste trabalho e por isso serão detalhados a seguir.

2.2 Modelo de Mitchell e Schaeffer (MS)

O modelo de Mitchell e Schaeffer (2003) original contém duas variáveis, o potencial transmembrânico $v(t)$ e uma variável de recuperação $h(t)$, e estas satisfazem equações diferenciais ordinárias. O modelo pode ser obtido após redimensionamento apropriado, como um caso especial do modelo de Karma (1994). O potencial transmembrânico, que é adimensional, é governado pela seguinte EDO:

$$\frac{dv}{dt} = J_{\text{in}}(v, h) + J_{\text{out}}(v) + J_{\text{stim}}(t), \quad (2.1)$$

onde as três correntes J_{in} , J_{out} , e J_{stim} tem as seguintes descrições:

- (i) J_{in} é a corrente interna, uma combinação de todas as correntes que aumentam o potencial através da membrana (principalmente sódio e cálcio), e é dada por

$$J_{\text{in}}(v, h) = \frac{hC(v)}{\tau_{\text{in}}}, \quad (2.2)$$

onde a função cúbica $C(v)$ é dada por

$$C(v) = v^2(1 - v), \quad (2.3)$$

e descreve a dependência da tensão da corrente de entrada. A força da corrente J_{in} é especificada pela constante de tempo τ_{in} ;

- (ii) J_{out} é a corrente de saída, uma combinação das correntes que diminuem o potencial da membrana (principalmente a de potássio), e é dada por

$$J_{\text{out}} = -\frac{v}{\tau_{\text{out}}}; \quad (2.4)$$

- (iii) J_{stim} é a corrente de estímulo, é uma corrente externa aplicada.

A variável de gating h , é adimensional e varia entre 0 e 1, satisfaz a EDO abaixo

$$\frac{dh}{dt} = \begin{cases} \frac{1-h}{\tau_{\text{open}}} & \text{se } v < v_{\text{gate}}, \\ \frac{-h}{\tau_{\text{close}}} & \text{se } v > v_{\text{gate}}, \end{cases} \quad (2.5)$$

onde τ_{open} e τ_{close} são as constantes de tempo com as quais o gating h abre e fecha, e v_{gate} é o valor da tensão fixado para o modelo. Em resumo, τ_{in} , τ_{out} , τ_{open} e τ_{close} são parâmetros do modelo.

2.3 Modelo de Mitchell e Schaeffer Modificado (MMS)

O modelo MMS consiste em uma modificação do modelo Mitchell e Schaeffer (2003) original proposta por Corrado e Niederer (2016), alterando o modelo MS original

para se evitar o comportamento oscilatório, também conhecido como comportamento de marca-passo, uma vez que ocorre a geração de vários potenciais de ação seguidos de forma espontânea. Foi verificado por Corrado e Niederer (2016) que o modelo original de MS apresenta o comportamento de marca-passo quando o espaço dos parâmetros é varrido, como é o caso em estudos de ajuste de parâmetros ou análises de sensibilidade. Essa alteração do MMS para remoção do comportamento do marca-passo feito por Corrado e Niederer (2016) é crucial no contexto de análise de sensibilidade, que faz uma busca exploratória no espaço de parâmetros, e precisa evitar esse comportamento para identificar corretamente a ocorrência ou não de alternans.

Este modelo também consiste de duas variáveis: o potencial transmembrânico $v(t)$ e a variável de recuperação denotada por $h(t)$. O seguinte sistema de EDOs descreve o modelo:

$$\frac{dv}{dt} = h \frac{(v + a)(v + a - \lambda)(1 - v)}{\tau_{in}} - \frac{v}{\tau_{out}} + J_{stim} \quad (2.6)$$

$$\frac{dh}{dt} = \begin{cases} \frac{1-h}{\tau_{open}} & v < v_{gate}, \\ -\frac{h}{\tau_{close}} & v > v_{gate}, \end{cases} \quad (2.7)$$

onde a , λ , τ_{in} , τ_{out} , τ_{open} e τ_{close} são parâmetros do modelo. Em particular, considera-se o modelo com $a = 0$, $\lambda = v_{gate}$ e $v_{gate} = 0,13$. Pode-se perceber que a diferença entre os modelos MS e MMS consiste no valor de λ na primeira EDO (2.6).

A Tabela 1 apresenta os valores base dos parâmetros do modelo MMS utilizados neste trabalho, bem como intervalos onde o comportamento do modelo Mitchell e Schaeffer (2003), Corrado e Niederer (2016) foi avaliado.

Tabela 1 – Valores base dos parâmetros do modelo MMS e o intervalo da distribuição uniforme utilizada para a análise de sensibilidade, em que $\mathcal{U}(a, b)$ denota a distribuição uniforme no intervalo $[a, b]$.

Parâmetros	Valores base	Distribuição
τ_{in}	0.3	$U(0,1; 0,5)$
τ_{out}	6.0	$U(2; 10)$
τ_{open}	120.0	$U(40; 200)$
τ_{close}	150.0	$U(50; 250)$

A Figura 5 mostra um potencial de ação do modelo MMS, gerado após a aplicação de um estímulo externo no tempo $t = 100$ ms, considerando os valores base dos parâmetros do modelo mostrados na Tabela 1.

2.4 Modelo de Tusscher-Panfilov (TP)

Em geral, a membrana celular é modelada como um capacitor conectado em paralelo com resistências variáveis e baterias representando diferentes correntes iônicas. O

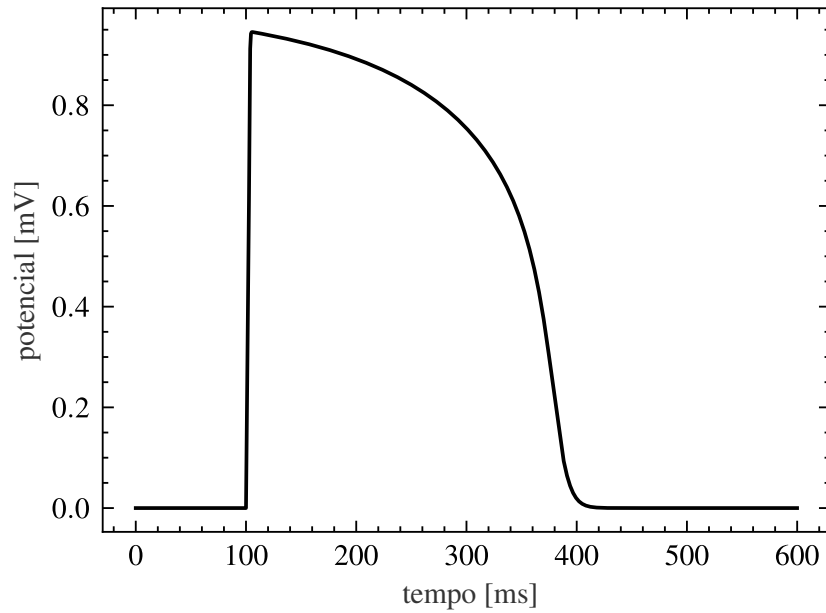


Figura 5 – Potencial de Ação do modelo MMS com uma corrente aplicada em $t = 100$ ms.

comportamento eletrofisiológico de uma única célula pode ser descrito de forma geral com a seguinte EDO:

$$C_m \frac{dV}{dt} = -(I_{ion} + I_{stim}), \quad (2.8)$$

onde V é a tensão, t é o tempo, I_{ion} é a soma de todas as correntes iônicas transmembrânicas, I_{stim} é a corrente de estímulo aplicada externamente e C_m é a capacitância da membrana celular dada por unidade de área de superfície.

Em particular, no modelo ventricular humano proposto por Tusscher e Panfilov (2006), o termo I_{ion} , que descreve as correntes iônicas do modelo, é dado por:

$$I_{ion} = I_{Na} + I_{KI} + I_{to} + I_{Kr} + I_{Ks} + I_{CaL} + I_{NaCa} + I_{NaK} + I_{pCa} + I_{pK} + I_{bCa} + I_{bNa}, \quad (2.9)$$

onde I_{Na} , I_{KI} , I_{to} , I_{Kr} , I_{Ks} , I_{CaL} , I_{NaCa} , I_{NaK} , I_{pCa} , I_{pK} , I_{bCa} e I_{bNa} são as correntes iônicas consideradas no modelo.

Algumas correntes citadas, como as correntes I_{Ks} e I_{CaL} , serão descritas a seguir:

$$I_{Ks} = G_{Ks} x_s^2 (V - E_{Ks}),$$

$$I_{CaL} = G_{CaL} d f f_2 f_{cass} 4 \frac{(V - 15) F^2}{RT} \frac{0.25 Ca_{ss} e^{\frac{2(V-15)F}{RT}} - Ca_o}{e^{\frac{2(V-15)F}{RT}} - 1},$$

onde G_{Ks} e G_{CaL} são as condutâncias destes canais iônicos e os demais termos são constantes (como R e F), parâmetros (como, por exemplo, Ca_o) ou variáveis do modelo (como, por exemplo, V , d , f , f_2 , f_{cass} e x_s).

Pode-se perceber a importância das condutâncias em modelos celulares deste tipo, pois elas determinam a intensidade máxima das correntes iônicas que, por sua vez, moldam

o perfil temporal do potencial de ação. O modelo TP possui muitos outros parâmetros, entretanto, neste trabalho, assim como em outros da literatura, fez-se a escolha dos 10 parâmetros que representam as condutâncias máximas dos canais iônicos dentre os muitos parâmetros do modelo Tusscher e Panfilov (2006) para a análise de sensibilidade.

A Tabela 2 mostra os valores base dos parâmetros do modelo Tusscher e Panfilov (2006), conforme a publicação original, e utilizado neste trabalho. Além disso, a tabela apresenta os intervalos de distribuições uniformes utilizadas nas análises de sensibilidade deste trabalho, os quais foram centrados nos valores base, resultando em um coeficiente de variação de aproximadamente 17%.

Tabela 2 – Valores base dos parâmetros do modelo TP e o intervalo de uma distribuição uniforme utilizada para análise de sensibilidade.

Parâmetros	Valores base	Distribuição
G_{CaL}	$3,98 \times 10^{-5}$	$U(2,786; 5,174) \times 10^{-5}$
G_{K1}	5,405	$U(3,7835; 7,0265)$
G_{Kr}	0,153	$U(0,1071; 0,1989)$
G_{Ks}	0,392	$U(0,2744; 0,5096)$
G_{Na}	14,838	$U(10,3866; 19,2894)$
G_{bCa}	$5,92 \times 10^{-4}$	$U(4,144; 7,696) \times 10^{-4}$
G_{bNa}	$2,9 \times 10^{-4}$	$U(2,030; 3,770) \times 10^{-4}$
G_{pCa}	0,1238	$U(0,08666; 0,16094)$
G_{pK}	0,0146	$U(0,01022; 0,01898)$
G_{to}	0,294	$U(0,2058; 0,3822)$

A Figura 6 apresenta o potencial de ação do modelo TP, gerado após a aplicação de um estímulo externo em $t = 100\text{ms}$, considerando os parâmetros do modelo em seus valores base, como mostrado na Tabela 2.

2.5 Modelo de Mahajan-Shiferaw

No modelo de potencial de ação ventricular de coelho proposto por Mahajan e Shiferaw (2008), o termo I_{ion} , que descreve as correntes iônicas do modelo, é dado por:

$$I_{ion} = I_{Na} + I_{to,f} + I_{to,s} + I_{Kr} + I_{Ks} + I_{K1} + I_{NaK} + I_{Ca} + I_{NaCa}, \quad (2.10)$$

onde I_{Na} , $I_{to,f}$, $I_{to,s}$, I_{Kr} , I_{Ks} , I_{K1} , I_{NaK} , I_{Ca} e I_{NaCa} são as correntes iônicas consideradas no modelo.

A título de exemplo, duas dessas correntes, I_{Na} e I_{Ks} , têm a seguinte forma:

$$\begin{aligned} I_{Na} &= G_{Na} m^3 h j (V - E_{Na}), \\ I_{Ks} &= G_{Ks} x_s^2 (V - E_{Ks}), \end{aligned}$$

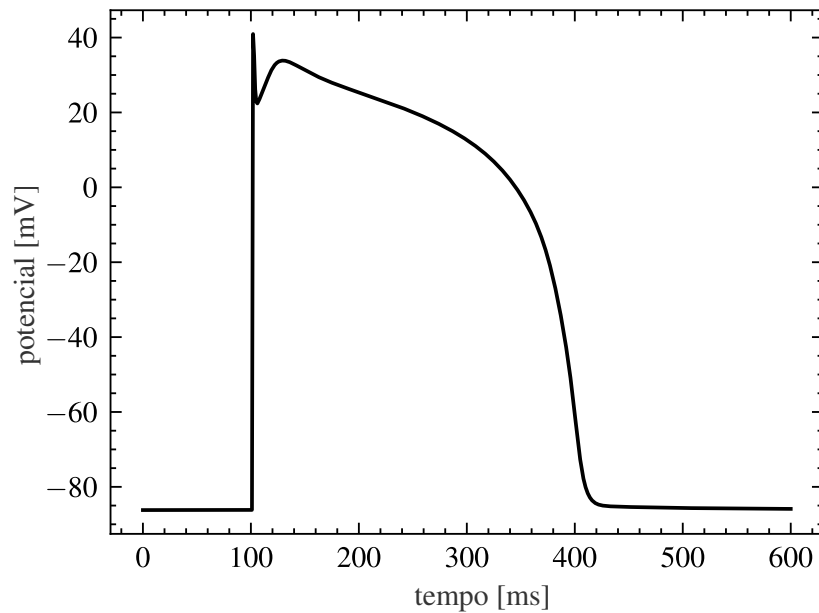


Figura 6 – Simulação numérica do potencial de ação do modelo TP resultante após a aplicação de uma corrente externa aplicada em $t = 100$ ms.

onde G_{Na} e G_{Ks} são as condutâncias destes canais iônicos e os demais termos são variáveis do modelo (como, por exemplo, V , m , h , j e x_s). As demais correntes seguem formalismo semelhante e podem ser encontradas na publicação original de Mahajan e Shiferaw (2008). Note que, assim como no modelo Tusscher e Panfilov (2006) apresentado anteriormente, os parâmetros das condutâncias determinam a intensidade máxima das correntes iônicas que, por sua vez, moldam o perfil temporal do potencial de ação.

A Figura 7 apresenta o potencial de ação do modelo Mahajan, gerado após a aplicação de um estímulo externo em $t = 100$ ms, considerando os parâmetros do modelo em seus valores base, como mostrado na Tabela 3 adiante.

O Modelo de Mahajan e Shiferaw (2008) possui muitos parâmetros, no entanto, neste trabalho, fez-se a escolha de 13 parâmetros para avaliar a sensibilidade dos mesmos na ocorrência de alternans neste estudo. Os parâmetros selecionados representam as condutâncias dos canais iônicos, as concentrações intracelulares de potássio e sódio e as concentrações extracelulares de cálcio, potássio e sódio. A Tabela 3 mostra os valores base dos parâmetros do modelo Mahajan e Shiferaw (2008) utilizado neste trabalho e as distribuições utilizadas neste trabalho para a análise de sensibilidade, as quais foram centradas nos valores base, resultando em um coeficiente de variação de aproximadamente 15%.

2.6 Conceitos de eletrofisiologia cardíaca

Apresentados os modelos utilizados neste trabalho, introduzem-se a seguir alguns conceitos fundamentais da eletrofisiologia cardíaca necessários para as análises desenvolvi-

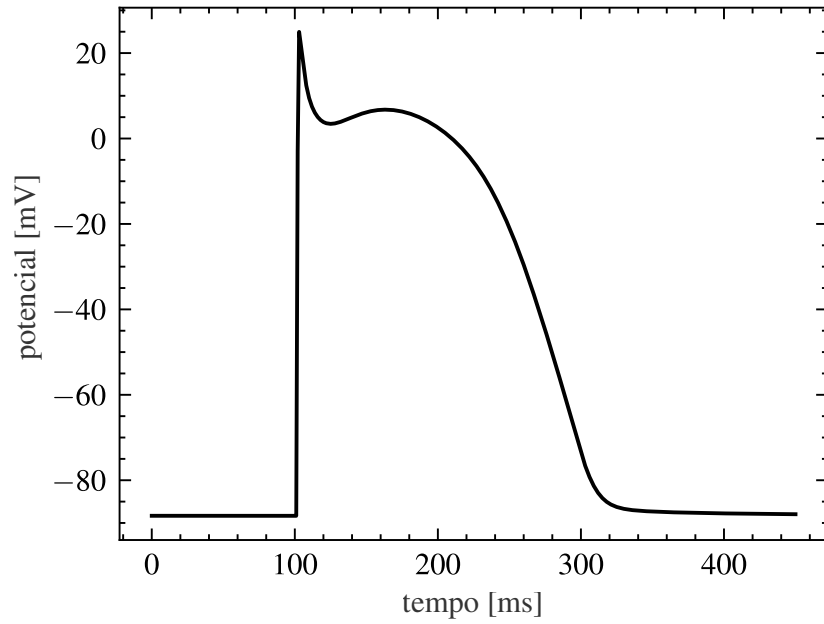


Figura 7 – Simulação numérica do potencial de ação do modelo Mahajan resultante após a aplicação de uma corrente externa aplicada em $t = 100$ ms.

Tabela 3 – Valores base dos parâmetros do modelo Mahajan e o intervalo de uma distribuição uniforme utilizada para análise de sensibilidade.

Parâmetros	Valores base	Distribuição
K_i	140,0	$U(105,0; 175,0)$
Ca_o	1,3	$U(0,975; 1,625)$
G_{CaL}	150,0	$U(112,5; 187,5)$
G_{K1}	0,22	$U(0,165; 0,275)$
G_{Kr}	0,0125	$U(9,375; 15,625) \times 10^{-3}$
G_{Ks}	0,40	$U(0,30; 0,50)$
G_{Na}	7,0	$U(5,25; 8,75)$
G_{NaCa}	0,54	$U(0,405; 0,675)$
G_{tof}	0,11	$U(8,25; 13,75) \times 10^{-2}$
G_{tos}	0,04	$U(0,03; 0,05)$
K_o	5,4	$U(4,05; 6,75)$
Na_i	12,3	$U(9,225; 15,375)$
Na_o	136,0	$U(102,0; 170,0)$

das nos capítulos seguintes.

2.6.1 Duração do potencial de ação

A Duração do Potencial de Ação descreve o intervalo de tempo durante o qual um potencial de ação cardíaco se estende desde sua despolarização inicial até a sua repolarização completa.

A Figura 8 ilustra as fases de um potencial de ação típico cardíaco: a fase 1 corresponde à despolarização rápida, mediada pelo rápido influxo de íons de sódio, a qual pode ser caracterizada pelo instante de tempo t_{dp} ; a fase 2 é a fase de platô, sustentada pelo influxo de cálcio e pelo equilíbrio do efluxo de potássio; a fase 3 é a repolarização final, caracterizada pelo fechamento dos canais de cálcio e pela abertura adicional dos canais de potássio, caracterizada pelo instante de tempo t_{rep} ; enquanto a fase 4 representa o momento onde o potencial da membrana celular está novamente em repouso.

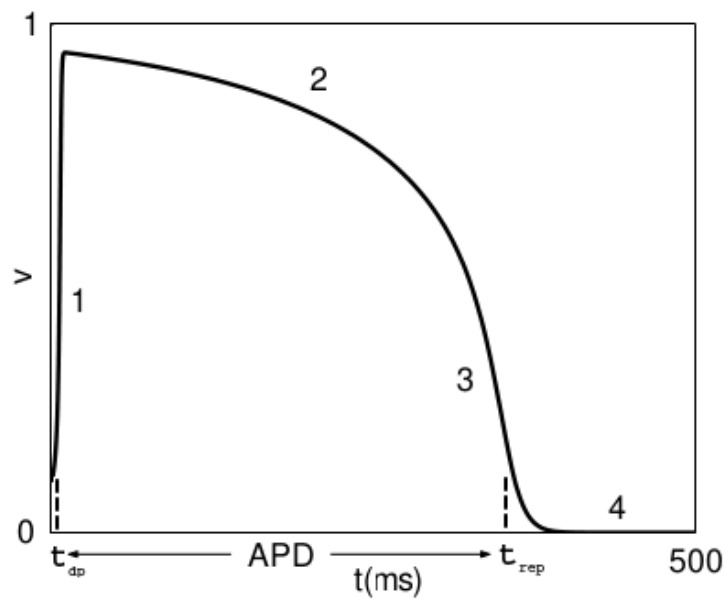


Figura 8 – Potencial de Ação típico do modelo MMS e suas fases: (1) despolarização; (2) platô; (3) repolarização e (4) repouso.

A duração do potencial de ação pode ser medida simplesmente calculando-se a diferença do tempo de repolarização t_{rep} pelo tempo de despolarização t_{dp} , isto é:

$$APD = t_{rep} - t_{dp}. \quad (2.11)$$

Os tempos t_{dp} e t_{rep} podem ser determinados de diferentes formas e isso resulta em diferentes métricas para o APD. Em particular, neste trabalho optou-se por utilizar APD como a medida conhecida como APD90 que representa o tempo que leva para que o potencial de ação decaia até 90% de seu valor máximo.

2.6.2 Curva de restituição

A Restituição na eletrofisiologia cardíaca se refere à relação entre o APD de um miócito cardíaco e o intervalo de tempo entre os batimentos cardíacos, também chamado de intervalo diastólico (DI, da sigla em inglês *diastolic interval*). Esse conceito é fundamental para entender a dinâmica do comportamento elétrico do coração e suas implicações para a estabilidade do ritmo cardíaco.

A Curva de Restituição é utilizada para descrever a relação entre o APD e a frequência de estimulação, pois, à medida que a frequência de estimulação é aumentada, o APD é reduzido para manter uma relação de resposta, um para um. A Figura 9 apresenta uma curva de restituição típica do modelo MMS, a partir da qual pode-se observar a relação do APD com DI.

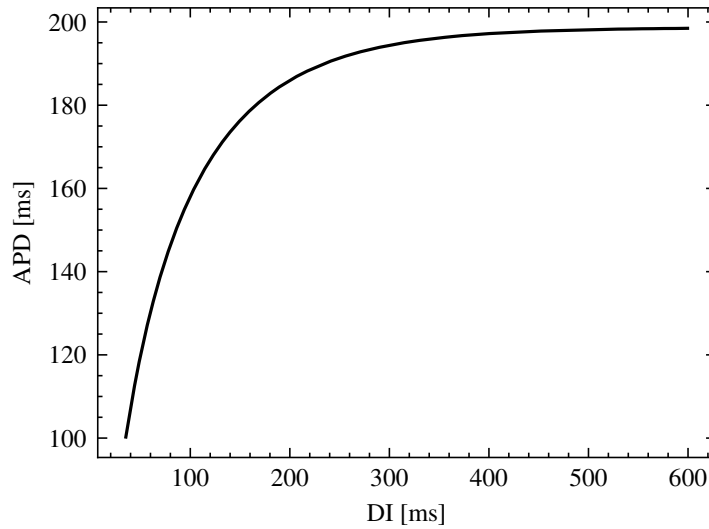


Figura 9 – Curva de Restituição ($DI \times APD$) do modelo MMS.

Nesse contexto, algumas definições são importantes. Define-se BCL (da sigla em inglês de *basic cycle length*) como o intervalo de tempo entre o início de dois batimentos cardíacos consecutivos, sendo a soma da duração do potencial de ação APD e do intervalo diastólico DI , que é o tempo de repouso entre os potenciais de ação, isto é:

$$BCL = APD + DI. \quad (2.12)$$

Para o caso particular do modelo MMS, foi mostrado em Corrado e Niederer (2016) a seguinte relação entre APD e DI com base nos parâmetros do modelo:

$$APD_{i+1} = f(DI_i) = \tau_{close} \ln \left(\frac{1 - \left(1 - h_{min}^{MMS}\right) e^{-\frac{DI_i}{\tau_{open}}}}{h_{min}^{MMS}} \right), \quad (2.13)$$

onde h_{min}^{MMS} é dado por:

$$h_{min}^{MMS} = \left(1 + \frac{\tau_{out}}{4\tau_{in}} (1 - v_{gate})^2 \right)^{-1}. \quad (2.14)$$

De forma simples, a Curva de Restituição vista na Figura 9 também pode ser gerada pela expressão analítica para o modelo MMS dada pela Equação (2.13). Sendo assim, neste modelo para verificar a ocorrência de alternans para um conjunto de parâmetros específico, basta examinar se a Equação (2.13) atende ao critério proposto por Nolasco e Dahlen (1968):

$$\max |f'(DI)| > 1. \quad (2.15)$$

Em resumo, para o modelo MMS, é possível realizar a análise da ocorrência do alternans ou não para um determinado conjunto de parâmetros do modelo a partir do estudo da equação (2.13) através do critério determinado pela equação (2.15).

Para modelos celulares mais complexos, como é o caso do modelo TP e da grande maioria dos modelos biofísicos detalhados, equações que descrevem a relação $DI \times APD$ não estão disponíveis. Portanto, para estes modelos a curva de restituição do APD é determinada através de simulações numéricas utilizando-se de protocolos específicos para tal fim. Os protocolos para a curva de restituição serão apresentados adiante na Seção 2.6.4.

2.6.3 Alternans elétrico

Na eletrofisiologia cardíaca, o alternans elétrico refere-se à variação na amplitude e duração dos APs celulares, que ocorrem, batimento-a-batimento, como pode-se ver na Figura 10, a qual foi obtida a partir da simulação numérica do modelo MMS. A figura mostra o alternans elétrico emergindo a partir de um determinado instante de tempo ($t = 265000\text{ms}$) em que podemos observar a oscilação do APD, alternando em um potencial de ação com duração longa e outro com duração curta.

Alternans podem ser observados e até mesmo induzidos em simulações de células cardíacas através dos protocolos de restituição, que são comumente usados em eletrofisiologia cardíaca para investigar o APD e as propriedades de adaptação de frequência das células cardíacas. Os protocolos de restituição do potencial de ação serão mostrados na Seção 2.6.4 a seguir.

Como explicado anteriormente pela Seção 2.6.2, há uma bifurcação que surge a partir do valor de $|f'(DI)| = 1$, ou seja, o alternans pode ser observado também na curva de restituição como uma bifurcação. A Figura 11 mostra um exemplo de curva de restituição típica para o modelo MMS que resultou em alternans, o qual pode ser verificado por volta de 450 ms, indicando que, a cada 2 estímulos, dois potenciais de ação com durações diferentes são gerados, um curto e um longo.

A Figura 12 mostra um exemplo da Curva de Restituição para o modelo Tusscher e Panfilov (2006), para um conjunto de parâmetros específico e, onde nota-se, por volta de 360ms, o alternans elétrico.

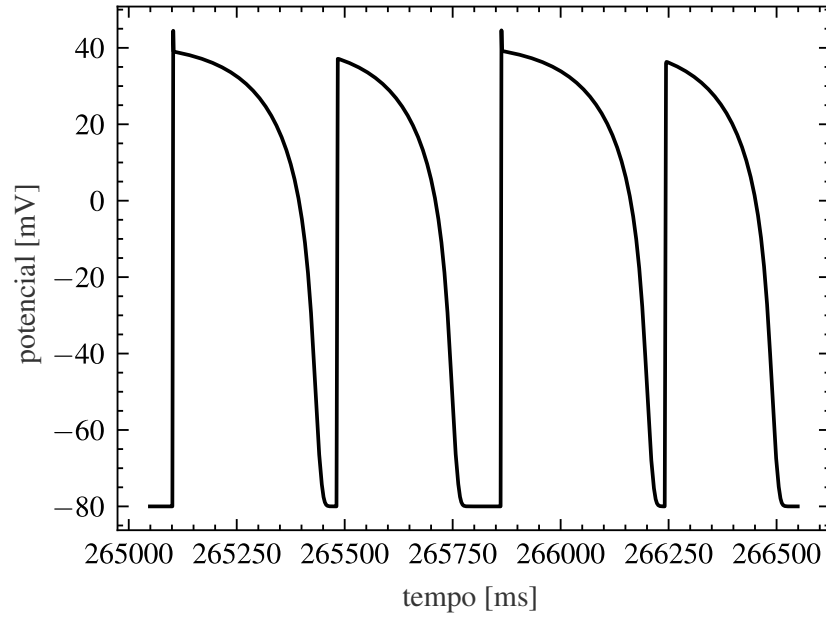


Figura 10 – APs no modelo MMS e a ocorrência do alternans, para o seguinte conjunto de parâmetros do modelo MMS: $\tau_{in} = 0,1$, $\tau_{out} = 9,0$, $\tau_{open} = 100$, e $\tau_{close} = 120$.

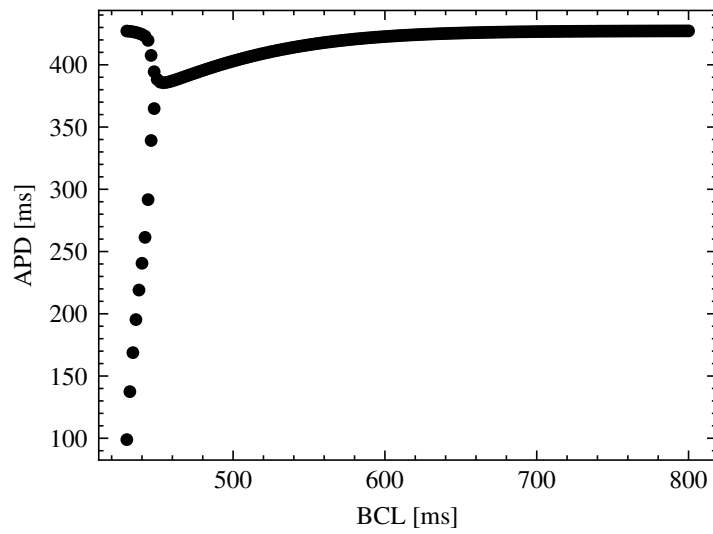


Figura 11 – A curva de restituição ($BCL \times APD$) do modelo MMS mostra alternans em um BCL de 450ms. Parâmetros do modelo: $\tau_{in} = 0,1$, $\tau_{out} = 4,3$, $\tau_{open} = 50,1$ e $\tau_{close} = 177,5$.

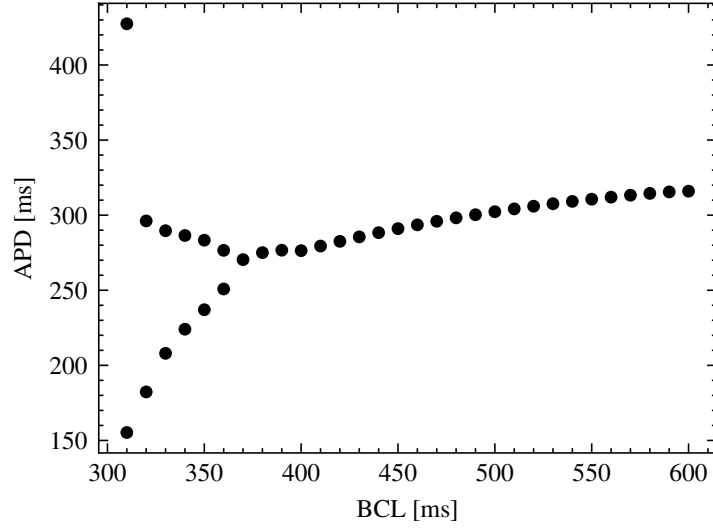


Figura 12 – Curva de restituição (BCL $\times$ APD) do modelo TP e exemplo da ocorrência de alternans. Parâmetros: $G_{CaL} = 2,1 \times 10^{-5}$, $G_{K1} = 4,060$, $G_{Kr} = 0,094$, $G_{Ks} = 0,210$, $G_{Na} = 14,558$, $G_{bCa} = 5,9 \times 10^{-4}$, $G_{bNa} = 4,1 \times 10^{-4}$, $G_{pCa} = 0,135$, $G_{pK} = 0,011$, $G_{to} = 0,184$.

2.6.4 Protocolo de restituição

Os protocolos de restituição são projetados para revelar diferentes características da restituição do potencial de ação. Basicamente, existem 2 protocolos para gerar a curva de restituição: o protocolo S1S2 e o dinâmico. O protocolo S1S2 é usado para verificar a predisposição a arritmias, enquanto o protocolo dinâmico reflete transições mais lentas, mostrado no trabalho de Plank et al. (2021). Os experimentos computacionais deste trabalho utilizaram o protocolo dinâmico. O protocolo dinâmico será brevemente detalhado a seguir tendo em vista a sua utilização no simulador OpenCARP.

No protocolo de restituição dinâmico, o comprimento do ciclo de estímulo $S1$ é aplicado por um número de batimentos previamente escolhido, N_{beats} , e então decrescido sucessivamente, repetindo-se esse procedimento até que se atinja um intervalo de acoplamento (*coupling interval*, CI) mínimo, denominado CI_0 . O protocolo é ilustrado na Figura 13.

É importante destacar que em ambos os protocolos uma etapa de preparação denominada *prepacing* é usualmente realizada com o objetivo de atingir uma condição de estado estacionário do modelo celular antes de se iniciar o protocolo de restituição.

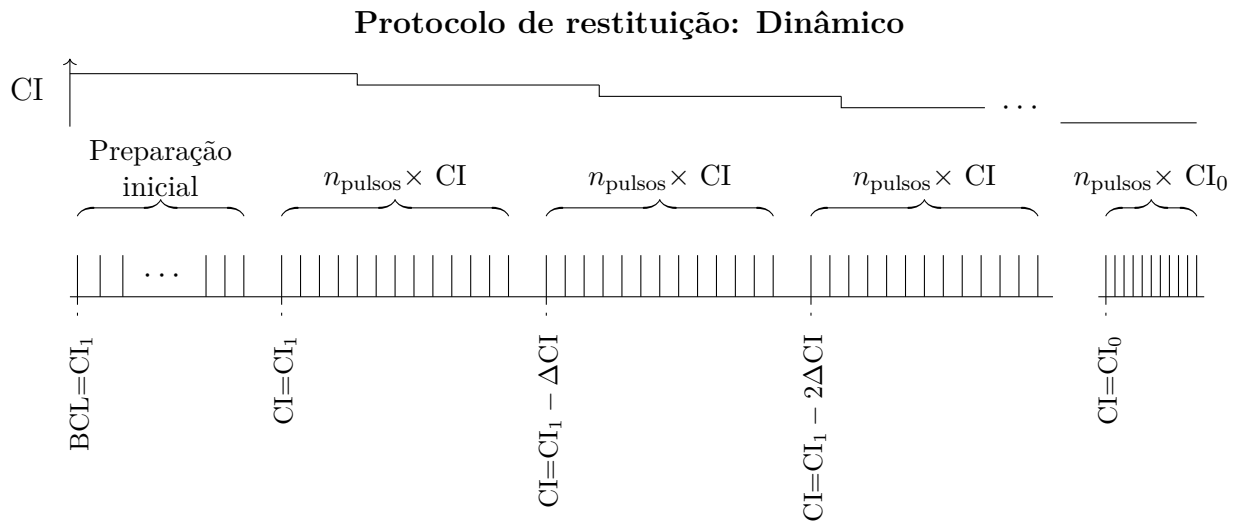


Figura 13 – Exemplo de protocolo de restituição dinâmico.

3 ANÁLISE DE SENSIBILIDADE E MODELOS SUBSTITUTOS

Este capítulo destina-se a apresentar os métodos de análise de sensibilidade que serão utilizados neste trabalho, como análise de sensibilidade local, análise de sensibilidade global baseada nos índices de Sobol e a técnica de Filtragem de Monte Carlo. Além disso, as quantidades de interesse (QoI) serão definidas. E, por fim, será feita uma apresentação breve da utilização de modelos substitutos (*surrogate model*) baseados na Expansão em Caos Polinomial, visando a aceleração das técnicas de análise de sensibilidade.

Este capítulo pressupõe que o leitor tenha conhecimento prévio em probabilidade e estatística, como nos trabalhos de Casella e Berger (2024), Ross (1976). Uma compreensão básica desses conceitos é essencial para a apreciação adequada dos métodos apresentados, especialmente no que diz respeito a valor esperado, variância, distribuição de probabilidade, correlação, função densidade de probabilidade e função de distribuição acumulada.

3.1 Análise de Sensibilidade

A análise de sensibilidade (SA) consiste no estudo de como a variação (incerteza) nos resultados de um modelo matemático pode ser distribuída, de forma qualitativa ou quantitativa, entre as diferentes fontes de variação das entradas do modelo. Esse procedimento permite identificar os fatores mais influentes no comportamento do sistema de interesse. Além disso, a SA pode ser utilizada para determinar a necessidade de pesquisas adicionais para reduzir a incerteza associada aos parâmetros de entrada, antes da tomada de decisões subsequentes, como no trabalho de Sobol (2001).

A aplicação de análise de sensibilidade inicia-se com a escolha de um modelo computacional cujo comportamento se pretende compreender em maior profundidade. Definem-se os parâmetros de entrada a serem analisados e, em seguida, atribuem-se distribuições de probabilidade a esses parâmetros. A variabilidade das entradas é então propagada para as saídas, tipicamente, por meio de múltiplas execuções do modelo.

A análise de quantificação de incertezas envolve a avaliação de como a variabilidade nos parâmetros de entrada se reflete nas saídas do modelo. Por sua vez, a análise de sensibilidade visa determinar a contribuição relativa da variabilidade de cada parâmetro de entrada em relação à variabilidade total observada nas saídas. Diversos métodos de análise de sensibilidade estão disponíveis, sendo que a escolha do método mais adequado depende da natureza do problema e do modelo em questão. A definição clara da importância de um fator ou parâmetro requer a utilização de critérios objetivos, assegurando uma interpretação inequívoca de sua relevância.

A análise de sensibilidade, apesar de essencial para avaliar a influência de parâmetros em modelos complexos, pode apresentar um elevado custo computacional. Sendo assim,

é comum a utilização de modelos substitutos como, por exemplo, aqueles baseados na Expansão em Caos Polinomial (PCE), para se obter uma abordagem eficiente, permitindo reduzir significativamente o número de simulações necessárias sem comprometer a precisão dos resultados, algo feito também em Sudret e Mai (2015).

3.1.1 Análise de sensibilidade local

A análise de sensibilidade local investiga a resposta do modelo em torno de um ponto nominal do espaço de parâmetros de entrada. Formalmente, a sensibilidade de Y em relação a um parâmetro X_i é definida como a derivada parcial da saída em relação àquele parâmetro, avaliada no ponto nominal (MCCLARREN, 2018):

$$S_i = \left. \frac{\partial Y}{\partial X_i} \right|_{X^0}. \quad (3.1)$$

Na prática, adotou-se a técnica *one-at-a-time* (OAT), na qual apenas um parâmetro X_i , com i fixo, é perturbado a partir de seu valor nominal, mantendo-se os demais fixos, e a variação resultante na saída Y é atribuída exclusivamente àquele parâmetro. Esse procedimento corresponde a uma aproximação por perturbação finita da derivada acima.

Apesar de simples e de baixo custo computacional, a análise de sensibilidade local avalia a influência dos parâmetros apenas em uma vizinhança restrita do ponto nominal, não capturando efeitos de interação entre parâmetros nem o comportamento do modelo ao longo de todo o espaço de entrada. Tais limitações motivam o uso de métodos de análise de sensibilidade global, apresentados a seguir.

3.1.2 Análise de sensibilidade global baseada nos índices de Sobol

A análise de sensibilidade global examina os efeitos das variações simultâneas em todos os parâmetros de entrada, abrangendo todo o espaço de entrada de interesse. Essa análise pode ser conduzida por meio de métodos baseados em análise de variância, tradicionalmente utilizando o método de Monte Carlo. Além disso, outros métodos comumente implementados em pacotes, como o método dos Índices de Sobol proposto por Sobol (2001), também são amplamente empregados para essa finalidade.

O método dos Índices de Sobol é amplamente utilizado na literatura devido à sua capacidade de avaliar as interações entre os parâmetros de entrada e suas respectivas contribuições para a variância da saída. O índice de sensibilidade de um parâmetro quantifica seu impacto na incerteza dos resultados, medindo a parcela da variância da saída que pode ser atribuída à variabilidade direta desse parâmetro de entrada (índice de Sobol principal) e possíveis interações com os demais parâmetros (índice de Sobol total).

Neste trabalho foram calculados os índices de sensibilidade Sobol de primeira ordem ou principal (S_1) e os índices totais (S_T) para quantidades de interesse dos modelos de

eletrofisiologia celular. O método utilizado para calcular os índices de Sobol é aquele descrito por Saltelli et al. (2008), cuja implementação foi baseada na biblioteca **SALib** de Iwanaga, Usher e Herman (2022), disponível para a linguagem de programação Python.

O índice de Sobol de primeira ordem, S_1^i , quantifica o quanto a variação de um parâmetro específico (X_i) contribui diretamente para a variabilidade do resultado (Y). Em outras palavras, ele mede o impacto isolado de um parâmetro sem levar em conta suas interações com outros parâmetros. A sua fórmula é:

$$S_1^i = \frac{V[E(Y|X_i)]}{V(Y)}, \quad (3.2)$$

onde $E(Y|X_i)$ é o valor esperado (média) de Y dado um valor específico de X_i , e $V(E(Y|X_i))$ representa a variância desse valor esperado. Note ainda que $V(Y)$ é a variância total de Y . Assim, o índice de primeira ordem indica a fração da variância total que pode ser atribuída ao parâmetro X_i sozinho.

A principal propriedade associada aos índices de primeira ordem ou principal é:

$$S_1^i \in [0, 1], \quad \text{e} \quad \sum_i S_1^i \leq 1, \quad (3.3)$$

ou seja, S_1^i está sempre entre 0 e 1. Um valor de $S_1^i = 0$ indica que o parâmetro X_i não afeta a saída diretamente, enquanto $S_1^i = 1$ significa que o parâmetro X_i é o único determinante da variabilidade da saída, sem influência de outros parâmetros. Além disso, a soma de todos os índices de primeira ordem nunca ultrapassa o valor 1, já que S_1^i representa apenas o efeito direto de cada parâmetro, ignorando interações entre os parâmetros.

Por outro lado, o índice de Sobol total, S_T^i , mede não apenas a contribuição direta de um parâmetro, mas também as interações desse parâmetro com todos os outros parâmetros do modelo. Ele avalia o impacto global de um parâmetro na variabilidade do resultado, considerando todos os efeitos combinados. Este índice é dado por:

$$S_T^i = \frac{E[V(Y|X_{\sim i})]}{V(Y)} = 1 - \frac{V[E(Y|X_{\sim i})]}{V(Y)}, \quad (3.4)$$

onde $X_{\sim i}$ indica todos os parâmetros exceto X_i . A expressão $E[V(Y|X_{\sim i})]$ é o valor esperado da variância de Y quando se fixam todos os parâmetros exceto X_i . O índice de Sobol total mostra a fração de variação que resta no resultado Y ao se excluir o parâmetro X_i , com o valor de $1 - \frac{V[E(Y|X_{\sim i})]}{V(Y)}$ refletindo a contribuição total do parâmetro X_i à variância de Y , incluindo interações.

Vale destacar que, se S_1^i e S_T^i são aproximadamente iguais, o parâmetro X_i influencia a saída principalmente de forma direta, com pouca ou nenhuma interação com outros parâmetros. No entanto, se houver uma diferença substancial entre S_1^i e S_T^i , isso indica que o parâmetro X_i participa de interações importantes com outros parâmetros.

3.1.3 Exemplo índices de Sobol

Nesta seção é apresentado um exemplo da análise de sensibilidade pelos índices de Sobol aplicada à função de Rosenbrock. Em sua forma bidimensional, a função de Rosenbrock é definida por

$$f(x_1, x_2) = (a - x_1)^2 + b(x_2 - x_1^2)^2, \quad (3.5)$$

em que a e b são constantes, adotadas neste trabalho como $a = 1$ e $b = 100$. Neste caso, considerou-se ambos os parâmetros x_1 e x_2 do modelo no intervalo de $[-2, 2]$ e o experimento foi implementado na linguagem de programação Python utilizando-se da biblioteca **SALib** para a análise e cálculo dos índices de Sobol. Foi utilizado um total de 6144 amostras para este estudo.

Pode-se resumir a aplicação da análise de sensibilidade através dos índices de Sobol na biblioteca **SALib** da seguinte forma:

- Especificação dos dados de entrada (parâmetros) e seus limites. No caso deste exemplo: x_1 e x_2 que estão ambos delimitados no intervalo de $[-2, 2]$;
- Execução da função que gera as amostras;
- Avaliação do modelo em cada amostra de entrada gerada, ou seja, calcular $Y[i] = \text{rosenbrock}(X[i,0], X[i,1])$, considerando Y como o vetor de saídas, X a matriz dos parâmetros e **rosenbrock()** a função que implementa o modelo de Rosenbrock;
- Execução da função que faz o cálculo dos índices de Sobol usando o método de Saltelli et al. (2010);
- Análise dos resultados.

Os resultados obtidos a partir do cálculo dos índices de Sobol são apresentados na Figura 14. Pode-se perceber uma influência maior do parâmetro x_1 do que x_2 no valor final da Função Rosenbrock. Resultado este esperado e confirmado, uma vez que, x_1 na expressão matemática da Função Rosenbrock possui um expoente maior (4) do que o do parâmetro x_2 . Nota-se também que os dois parâmetros, x_1 e x_2 , têm interação de alta ordem, também esperada, uma vez que, na formulação matemática existe a presença de um termo $(x_2 - x_1^2)^2$, que resulta em termos contendo o produto destes parâmetros.

3.2 Filtragem de Monte Carlo (MCF)

As técnicas de análise de sensibilidade apresentadas até então (OAT e Sobol) visam quantificar o impacto de variações nos parâmetros de entrada na saída do modelo. Entretanto, existem situações em que o foco é em uma parte ou região específica da distribuição da variável de saída Y . Pode-se estar interessado em analisar se Y está acima

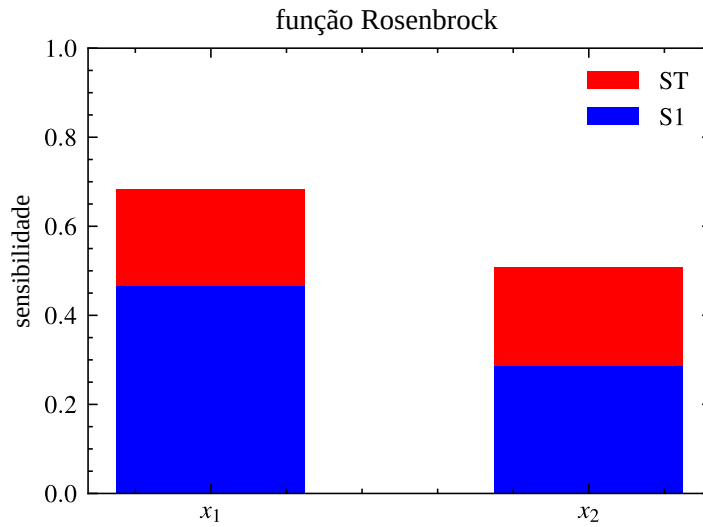


Figura 14 – Resultados do cálculo dos índices de Sobol de primeira ordem e de ordem total para a saída da função Rosenbrock.

ou abaixo de um valor limite, ou se Y apresenta um determinado comportamento ou não. Nesse tipo de cenário divide-se ou se classifica as realizações de Y entre aquelas que apresentam o comportamento alvo, ou, não apresentam, usando o método de Saltelli et al. (2004). Esse tipo de análise é conhecido como Filtragem de Monte Carlo ou também de *Regionalised Sensitivity Analysis* (RSA).

Neste tipo de análise, é preciso uma definição qualitativa do comportamento do sistema e uma classificação binária das saídas do modelo com base na definição especificada de comportamento. Qualifica-se uma saída da simulação como comportamental, B , se a saída do modelo estiver dentro dos limites ou definições, e como não comportamental, $\bar{B}$, caso contrário, como ilustrado de forma esquemática na Figura 15. Em particular, neste trabalho, o comportamento alvo B é a ocorrência de alternans nos modelos cardíacos através da simulação da curva de restituição do APD, enquanto a situação não comportamental $\bar{B}$ é a não ocorrência de alternans.

O método MCF pode ser descrito da seguinte forma. Define-se k parâmetros do modelo X_i , para $i = 1, \dots, k$, e realiza-se um grande número de avaliações do modelo utilizando o método Monte Carlo (MC). Cada saída do modelo está associada a um conjunto de parâmetros de entrada, e as saídas Y são classificadas de forma binária como B ou $\bar{B}$ gerando dois subconjuntos: $(X_i|B)$ de m elementos e $(X_i|\bar{B})$ de n elementos, tal que $N = m + n$ é o número de execuções do método Monte Carlo.

Em seguida, um teste estatístico é realizado para cada parâmetro independentemente para verificar a distância máxima entre as funções de distribuição acumuladas de $(X_i|B)$ e $(X_i|\bar{B})$. Portanto, se para um determinado parâmetro X_i , as duas funções de probabilidade acumulada marginal $F_m(X_i|B)$ e $F_n(X_i|\bar{B})$ forem significativamente diferentes, então X_i é considerado um fator influente no comportamento pré-determinado do modelo

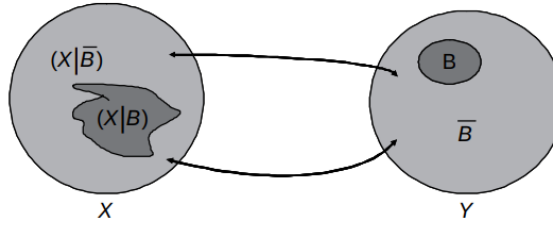


Figura 15 – Representação esquemática da Filtragem de Monte Carlo.

Y . Caso contrário, X_i não é influente. O teste de Smirnov (ou de Kolmogorov-Smirnov) para duas amostras é usualmente realizado para cada parâmetro independentemente. O teste estatístico é dado pelo cálculo de:

$$d_{n,\bar{n}}(X_i) = \sup \|F_n(X_i|B) - F_{\bar{n}}(X_i|\bar{B})\| \quad (3.6)$$

onde F são as funções de probabilidade acumulada marginais. De forma prática, quanto maior o valor da estatística $d_{n,\bar{n}}$, mais importante é o parâmetro na condução do comportamento B do modelo. A Figura 16 apresenta um exemplo de análise no qual fica claro a diferença entre as CDFs de B e $\bar{B}$ e, portanto, pode-se concluir que o parâmetro X_i apresentado na figura possui influência significativa para resultar no comportamento B .

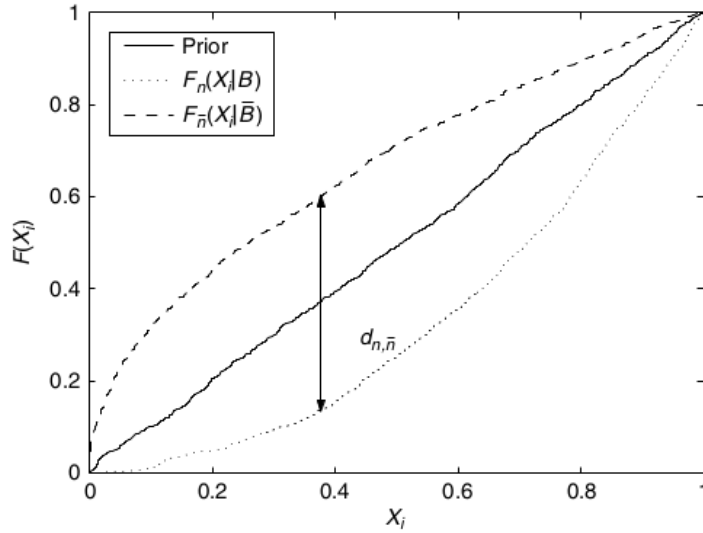


Figura 16 – Representação gráfica do teste de Smirnov para um parâmetro influente X_i . A barra vertical é a estatística $d_{n,\bar{n}}$. Adaptado de Saltelli et al. (2008).

3.2.1 Exemplo MCF 1

Para ilustrar a aplicação desta técnica de forma mais completa, um exemplo simples é apresentado a seguir. Seja um modelo simples dado por $Y = X_1 + X_2$, com $X_i \in U(0, 1)$, isto é, as variáveis aleatórias X_1 e X_2 seguem distribuição uniforme de 0 a 1, enquanto a variável aleatória Y representa a sua soma. Define-se o comportamento alvo B do

modelo a ser analisado como: $Y > 1$, e, caso contrário, quando o modelo não exibe o comportamento, denota-se por $\bar{B}$.

O resultado comparativo das CDFs para os casos B e $\bar{B}$ deste Exemplo 1 é apresentado na Figura 17. Pode-se perceber que ambos os parâmetros X_i são influentes no comportamento B do modelo, pois, os seus testes de Smirnov estão em valores de aproximadamente 0,50.

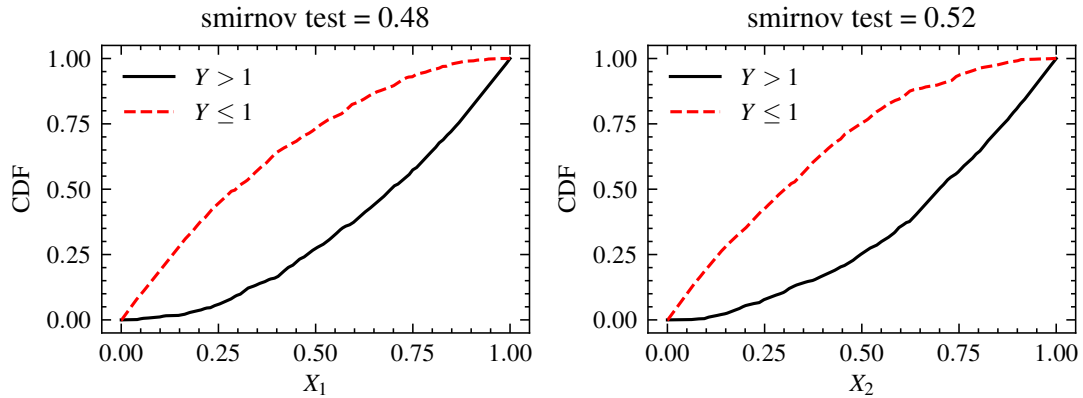


Figura 17 – Representação gráfica dos testes de Smirnov das variáveis X_1 e X_2 para o modelo simples $Y = X_1 + X_2$, onde $Y > 1$ é o comportamento alvo a ser alcançado.

Além disso, ao se analisar a correlação dos parâmetros X_1 e X_2 , pode-se compreender melhor o papel de ambos como resultantes do comportamento B . A Figura 18 apresenta um gráfico de dispersão dos valores de X_1 e X_2 que resultaram no comportamento B , isto é, para os quais o valor de $Y > 1$. Pode-se notar a tendência no espaço dos parâmetros que desencadeiam o comportamento B e, além disso, uma correlação negativa significativa de $cc = -0,53$ entre X_1 e X_2 no subconjunto B .

3.2.2 Exemplo MCF 2

Seja um modelo simples dado pela equação $Y = X_1 + X_2$, com $X_1 \sim N(0, 1)$ e $X_2 \sim N(0, 3)$, isto é, X_1 é uma variável aleatória normal com média zero e desvio padrão unitário e X_2 é uma normal com média zero e desvio padrão igual a três. Define-se o comportamento alvo B do modelo como: $Y > 1$, e, consequentemente $\bar{B}$ ocorre para $Y \leq 1$. Então, a partir da técnica MCF de análise de sensibilidade, deseja-se identificar a influência dos parâmetros de maneira individual para o modelo alcançar o comportamento alvo definido previamente.

A Figura 19 mostra os resultados deste caso, cuja implementação e análise foram feitas na linguagem de programação Python. Pela estatística de Smirnov, o parâmetro X_1 é em torno de 4 vezes menos influente no comportamento $Y > 1$ do que o parâmetro X_2 . Este resultado era esperado, e tem relação com a forma como foram criados estes parâmetros, que apesar de possuírem a mesma média tem-se que X_2 possui um desvio padrão maior. O

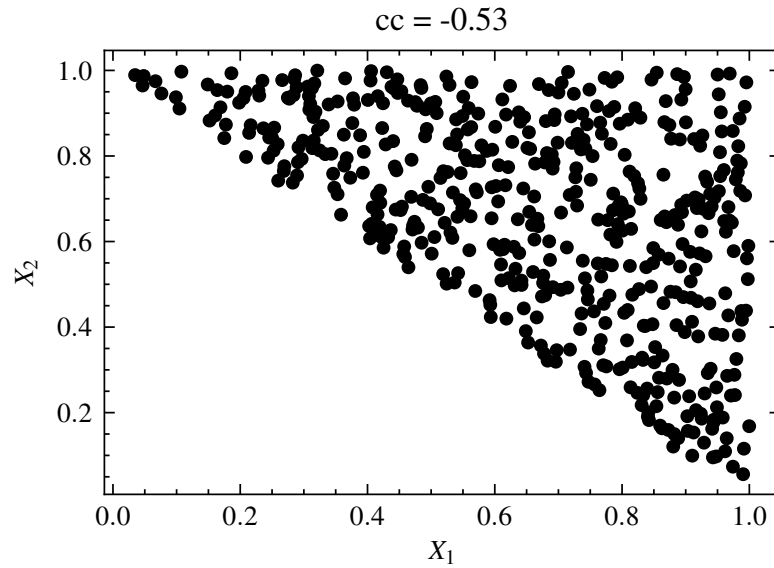


Figura 18 – Projeção bidimensional da subamostra B para o modelo simples $Y = X_1 + X_2$. O coeficiente de correlação entre X_1 e X_2 na amostra MC é $cc = -0,53$.

coeficiente de correlação dos dois parâmetros para que ocorra o comportamento alvo $Y > 1$ é de $cc = -0.37$, ou seja, existe uma correlação significativa entre os dois parâmetros.

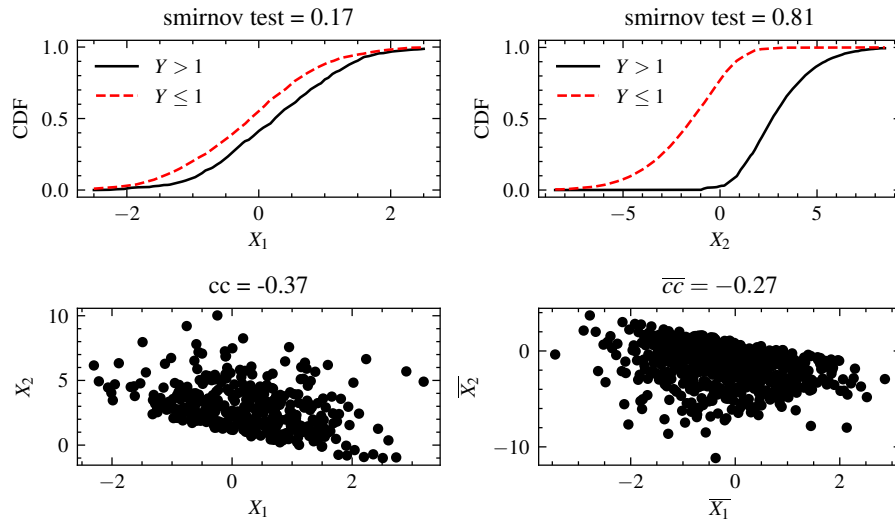


Figura 19 – Resultados da análise MCF para o Exemplo 2: na parte superior os gráficos das CDFs e na parte inferior as correlações dos parâmetros para os casos B e $\bar{B}$.

3.3 Modelos Substitutos e aprendizado de máquina

As técnicas de análise de sensibilidade apresentadas anteriormente requerem, tipicamente, um grande número de avaliações do modelo. Quando o modelo é computacionalmente caro, como no caso dos modelos de eletrofisiologia celular para alternans, essa exigência torna as análises inviáveis em termos de custo. Uma alternativa amplamente empregada consiste em substituir o modelo original por um modelo substituto (*surrogate*

model, ou emulador), de baixo custo de avaliação, construído a partir de um conjunto limitado de simulações do modelo completo (SUDRET, 2008; MARELLI; SUDRET,).

A Figura 20 apresenta de forma esquemática a utilização de emuladores ou modelos substitutos e suas vantagens e aplicações.

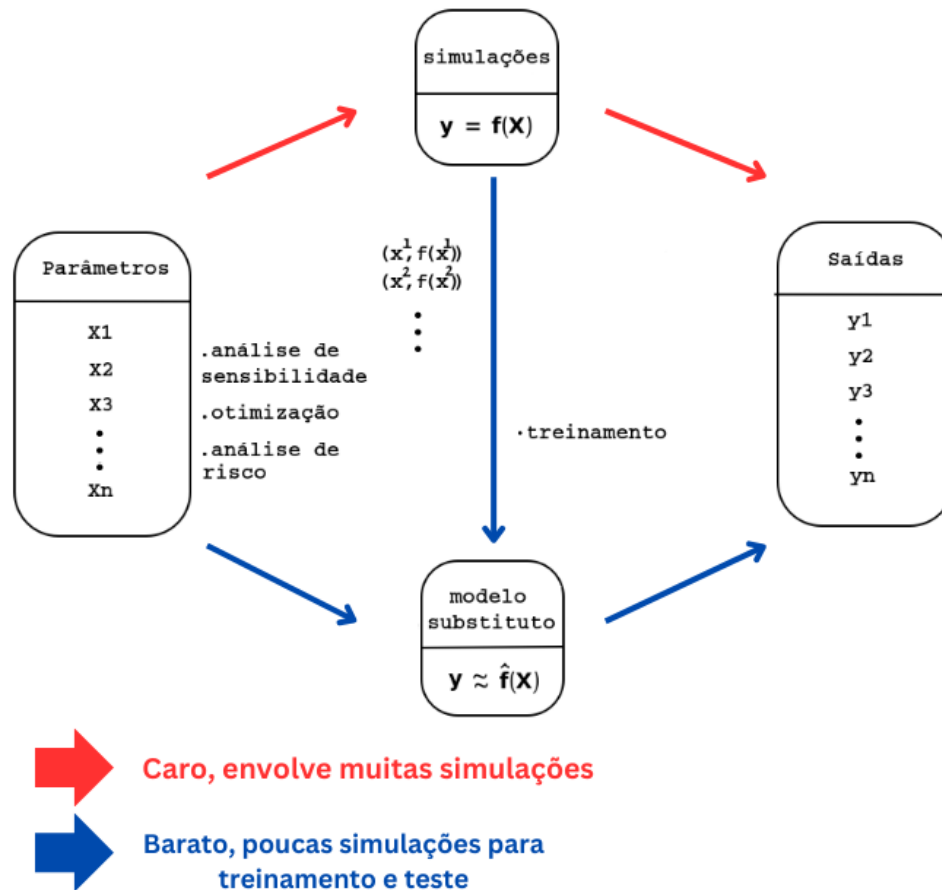


Figura 20 – Esquema ilustrativo da construção e utilização de modelo substituto para análise de sensibilidade.

Neste trabalho foram exploradas três técnicas para a construção desses modelos substitutos: a Expansão em Caos Polinomial (PCE), os Processos Gaussianos (GP), e as Florestas Aleatórias (RF). Cada uma delas foi empregada em duas tarefas distintas: na regressão, para aproximar quantidades de interesse contínuas dos modelos eletrofisiológicos, como o APD e a máxima inclinação da curva de restituição; e na classificação, para distinguir as simulações que apresentam alternans daquelas que não o apresentam. Os resultados do uso do PCE como emulador são apresentados no Capítulo 5, enquanto uma comparação de desempenho entre as três técnicas, em ambas as tarefas, é apresentada no Capítulo 6. As três técnicas são descritas a seguir.

3.3.1 Expansão em Caos Polinomial (PCE)

A base teórica da Expansão em Caos Polinomial (PCE, da sigla em inglês *polynomial chaos expansion*) tem origem no trabalho de Wiener (1938) sobre caos homogêneo e foi desenvolvida posteriormente para aplicações em processos estocásticos, como no trabalho de Xiu e Karniadakis (2002). PCE é um método comumente utilizado para quantificação de incerteza e aproximação de sistemas complexos, especialmente aqueles governados por muitas equações diferenciais ou processos estocásticos. Quando aplicado como um emulador ou modelo substituto (*surrogate model*), o PCE aproxima a saída de um modelo computacionalmente caro através de um polinômio cujas variáveis de entrada são aleatórias. Por se tratar de uma forma de aproximação polinomial, ele se destaca por ser de baixo custo computacional. Portanto, esse tipo de ferramenta tem sido muito utilizada em estudos de análise de sensibilidade e quantificação de incertezas, onde são requeridas muitas avaliações do modelo.

Emuladores PCE têm sido utilizados com sucesso em diversos estudos sobre quantificação de incertezas e análises de sensibilidades, tanto em modelos cardíacos, como nos trabalhos de Campos et al. (2020), Campos et al. (2023), quanto em outras áreas de aplicação, como proposto por Ciriello et al. (2013), Valdez et al. (2020), Sudret (2008).

Seja $f(\mathbf{Z})$ o simulador de uma quantidade de interesse Y , onde $\mathbf{Z}$ é um vetor com os parâmetros de entrada considerados como variáveis aleatórias. Nesta abordagem, a quantidade Y é expandida em uma série de polinômios ortogonais Ψ_j com entradas aleatórias. Ou seja, o PCE para Y é dado por:

$$Y = \sum_{j=0}^{\infty} b_j \Psi_j(\{Z_n\}_{n=1}^{\infty}) = f_{PCE}(\mathbf{Z}), \quad (3.7)$$

onde b_j são os coeficientes a serem determinados e Ψ_j são os polinômios ortogonais com suas respectivas variáveis aleatórias $Z_1, Z_2, \dots, Z_n$.

Na prática, a expansão é truncada em um número finito de termos, e a aproximação de Y , obtida pelo emulador f_{PCE} , e denotada por $\bar{Y}$, pode ser expresso por:

$$Y \approx \bar{Y} = \sum_{j=0}^{N_p-1} c_j \Psi_j(\mathbf{Z}) = f_{PCE}(\mathbf{Z}), \quad (3.8)$$

onde c_j são os coeficientes da expansão a serem determinados e $\Psi_j(\mathbf{Z})$ são os polinômios ortogonais. A expansão polinomial da equação acima tem grau p para D parâmetros de entrada, e o número de termos N_p é dado por:

$$N_p = \frac{(D+p)!}{D!p!}. \quad (3.9)$$

Em resumo, para treinamento do emulador PCE é necessária a execução do modelo completo N_p vezes. Entretanto, também podem ser utilizadas mais amostras para a

construção do mesmo e geralmente esta abordagem resulta em melhores resultados. Veja mais detalhes sobre a construção de emuladores PCE nos trabalhos de Campos (2019), Sudret (2008), Sudret e Mai (2015).

3.3.2 Processos Gaussianos (GP)

Os Processos Gaussianos (do inglês, *Gaussian Processes – GP*), também conhecidos como *Kriging* (RASMUSSEN; WILLIAMS, 2005) para aprendizado de máquina, constituem uma abordagem não paramétrica amplamente utilizada para a construção de modelos substitutos em simulações computacionalmente caras. Diferentemente da Expansão em Caos Polinomial, que aproxima explicitamente a resposta do modelo por uma combinação de polinômios ortogonais, um GP modela diretamente a distribuição de probabilidade sobre o conjunto de funções capazes de representar os dados observados.

Em um modelo substituto baseado em GP, assume-se que a resposta do emulador pode ser representada por uma realização de um processo estocástico gaussiano,

$$Y(\mathbf{x}) \sim \mathcal{GP}(m(\mathbf{x}), k(\mathbf{x}, \mathbf{x}')), \quad (3.10)$$

onde $m(\mathbf{x})$ representa a função média e $k(\mathbf{x}, \mathbf{x}')$ é a função de covariância, ou *kernel*, responsável por descrever a correlação entre diferentes pontos do espaço de entrada. É comum assumir uma média nula nesta correlação escrita acima, após a normalização dos dados, concentrando a modelagem do problema na escolha do kernel, pois isto, segundo Santner (2010), é o que determina propriedades como suavidade, continuidade e comprimento de correlação da função aproximada Y .

Dado um conjunto de amostras de treinamento composto por pares entrada-saída $\{(\mathbf{x}_i, \mathbf{y}_i)\}_{i=1}^N$, o GP utiliza a estrutura de covariância para prever a resposta em novos pontos do espaço de parâmetros. Além da estimativa da média preditiva, o método fornece naturalmente a variância associada à predição, permitindo quantificar a incerteza do modelo substituto. Essa característica representa uma das principais vantagens dos Processos Gaussianos em relação a outros métodos (FORRESTER; SOBSTER; KEANE, 2008).

Formalmente, agrupando as observações de treinamento no vetor $\mathbf{y} = (y_1, \dots, y_N)^\top$, a predição do GP em um novo ponto $\mathbf{x}^*$ é ela própria uma distribuição gaussiana, cuja média e variância são dadas por:

$$\hat{\mu}(\mathbf{x}^*) = m(\mathbf{x}^*) + \mathbf{k}(\mathbf{x}^*)^\top \mathbf{K}^{-1} (\mathbf{y} - \mathbf{m}), \quad (3.11)$$

$$\hat{\sigma}^2(\mathbf{x}^*) = k(\mathbf{x}^*, \mathbf{x}^*) - \mathbf{k}(\mathbf{x}^*)^\top \mathbf{K}^{-1} \mathbf{k}(\mathbf{x}^*), \quad (3.12)$$

onde $\mathbf{K}$ é a matriz de covariância do conjunto de treinamento, com entradas $K_{ij} = k(\mathbf{x}_i, \mathbf{x}_j)$, o vetor $\mathbf{k}(\mathbf{x}^*)$ reúne as covariâncias entre o ponto $\mathbf{x}^*$ e os pontos de treinamento, e $\mathbf{m}$ é o vetor da função média avaliada nos pontos de treinamento. A média preditiva $\hat{\mu}(\mathbf{x}^*)$

fornece a estimativa do emulador, enquanto a variância preditiva $\hat{\sigma}^2(\mathbf{x}^*)$ quantifica a incerteza associada.

A escolha do kernel define as propriedades da função aproximada. Neste trabalho adotou-se a função de covariância de Matérn com parâmetro de suavidade $\nu = 3/2$, na forma anisotrópica:

$$k(\mathbf{x}, \mathbf{x}') = \sigma^2 \left(1 + \sqrt{3} r\right) \exp\left(-\sqrt{3} r\right), \quad r = \sqrt{\sum_{i=1}^D \frac{(x_i - x'_i)^2}{\theta_i^2}}, \quad (3.13)$$

onde σ^2 é a variância do processo, D é o número de parâmetros de entrada e θ_i são os comprimentos de correlação. A forma anisotrópica atribui um comprimento de correlação θ_i distinto a cada parâmetro de entrada, permitindo que o kernel represente diferentes escalas de variação em cada direção do espaço de parâmetros.

Cabe destacar que a escolha do kernel exerce forte influência sobre o comportamento do emulador GP, afetando propriedades como suavidade e capacidade de generalização da resposta aproximada. Neste trabalho, os emuladores GP para regressão foram construídos com o *framework* UQLab (MARELLI; SUDRET,), adotando-se as configurações *default* da ferramenta, as quais apresentaram bom desempenho em testes iniciais.

Além da função de covariância, a construção do emulador GP envolve a definição de uma função de tendência (*trend*), que representa o comportamento médio da resposta, e a estimação dos hiperparâmetros do kernel. Neste trabalho adotou-se o *ordinary Kriging*, no qual a função média é assumida como uma constante desconhecida $m(\mathbf{x}) = \beta$, cujo valor é estimado a partir dos dados de treinamento.

A estimação dos hiperparâmetros do kernel, reunidos no vetor $\boldsymbol{\vartheta} = (\sigma^2, \theta_1, \dots, \theta_D)$, é formulada como um problema de otimização, no qual se busca o conjunto de valores que minimiza um critério de qualidade do ajuste do emulador aos dados de treinamento:

$$\boldsymbol{\vartheta}^* = \arg \min_{\boldsymbol{\vartheta}} \mathcal{J}(\boldsymbol{\vartheta}). \quad (3.14)$$

Neste trabalho, adotou-se a validação cruzada como critério $\mathcal{J}$, cujo problema de otimização foi resolvido por meio de um algoritmo genético.

3.3.3 Random Forests (RF)

O algoritmo Random Forest (RF) se consolidou como uma das técnicas de aprendizado supervisionado mais utilizadas nas últimas duas décadas devido à sua elevada capacidade preditiva, robustez estatística e flexibilidade para lidar com diferentes tipos de dados. Essas características explicam sua ampla aplicação em áreas como bioinformática, genômica, ecologia, medicina, finanças e modelagem computacional, como mostrado em Maindonald (2021). Diferentemente dos métodos tradicionais de regressão, que frequentemente assumem relações lineares ou previamente especificadas entre variáveis de entrada

e saída, o RF é um método não-paramétrico, capaz de representar relações altamente não-lineares, capturar interações complexas entre variáveis e acomodar simultaneamente atributos quantitativos e qualitativos. Além disso, apresenta reduzida sensibilidade à presença de ruído, elevada capacidade de generalização e baixo risco de sobreajuste (*overfitting*) tornando-se particularmente adequado para problemas envolvendo grande número de variáveis explicativas ou relações desconhecidas entre os dados (MAINDONALD, 2021).

O objetivo consiste em construir uma função capaz de aprender o relacionamento existente entre essas variáveis e utilizá-la posteriormente para classificar novas observações ainda desconhecidas. Formalmente, considera-se um conjunto de treinamento composto por pares ordenados

$$\mathcal{L} = \{(\mathbf{x}_1, y_1), (\mathbf{x}_2, y_2), \dots, (\mathbf{x}_n, y_n)\}, \quad (3.15)$$

onde $\mathbf{x}_i = (x_{i1}, x_{i2}, \dots, x_{ip})$ representa o vetor contendo as p variáveis explicativas da i -ésima observação, enquanto y_i corresponde à variável resposta associada a essa observação. No fenômeno abordado neste trabalho, cada vetor de entrada é constituído pelos parâmetros eletrofisiológicos amostrados durante as simulações computacionais, enquanto a variável resposta corresponde à presença ou ausência do alternans.

A unidade fundamental de um modelo RF é a árvore de decisão, também conhecida como árvore CART (do inglês, *Classification and Regression Trees*) (BREIMAN et al., 1984). Uma árvore de decisão consiste em uma estrutura hierárquica composta por nós internos, ramos e nós terminais (*folhas*) conforme ilustrado na Figura 21. Cada nó interno realiza uma pergunta referente ao valor de uma variável de entrada, dividindo iterativamente o conjunto de dados em subconjuntos cada vez mais homogêneos. Ao final desse processo, cada folha representa uma classe prevista pelo modelo.

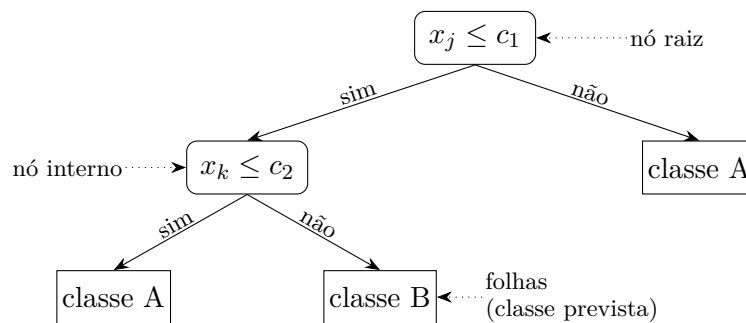


Figura 21 – Estrutura de uma árvore de decisão. Cada nó interno aplica uma condição sobre uma variável de entrada, dividindo o conjunto de dados em dois ramos; as folhas correspondem à classe prevista pelo modelo.

A construção da árvore ocorre por meio de sucessivas divisões binárias do conjunto de treinamento. Em cada etapa, o algoritmo procura selecionar simultaneamente a variável e o ponto de corte que produzam a melhor separação entre as classes presentes naquele nó. O objetivo de cada divisão é tornar os dois grupos resultantes o mais puros possível,

isto é, concentrar em cada lado observações de uma mesma classe. À medida que a árvore cresce, os grupos tornam-se progressivamente mais homogêneos.

Para quantificar essa homogeneidade são utilizados critérios de impureza. O mais empregado para o RF é o índice de Gini, definido por

$$G(t) = 1 - \sum_{k=1}^K p_k^2,$$

onde p_k representa a proporção de observações pertencentes à classe k no nó t , enquanto K corresponde ao número total de classes. O índice de Gini assume valor igual a zero quando todas as observações do nó pertencem à mesma classe, aumentando à medida que cresce a heterogeneidade entre as classes. Outros critérios podem se basear no conceito da entropia de Shannon, mas não foram utilizados neste trabalho.

A cada nó, o par variável–ponto de corte é escolhido por busca exaustiva: avaliam-se os cortes possíveis de cada variável e seleciona-se aquele que produz a maior redução de impureza entre o nó e seus dois filhos. O procedimento é guloso, pois otimiza cada divisão localmente, sem garantir a otimalidade global da árvore.

Apesar da simplicidade e da fácil interpretação, as árvores de decisão apresentam elevada variância: pequenas alterações no conjunto de treinamento podem produzir árvores muito diferentes, com previsões distintas. Isso as torna propensas ao sobreajuste, comprometendo sua capacidade de generalização diante de novos dados.

O *Random Forest* (BREIMAN, 2001) contorna essa limitação combinando um conjunto de B árvores em vez de uma única. Para que as árvores sejam distintas entre si, introduz-se aleatoriedade em duas etapas: cada árvore é construída sobre uma amostra *bootstrap* (uma reamostragem com reposição do conjunto de treinamento) e, em cada nó, a melhor divisão é buscada não entre todas as p variáveis, mas apenas em um subconjunto aleatório delas. Como resultado, obtêm-se árvores pouco correlacionadas, cujos erros tendem a se cancelar quando agregados. A predição final da floresta é dada pela combinação das predições das B árvores: por voto majoritário, no caso de classificação, ou pela média, no caso de regressão. Essa agregação reduz substancialmente a variância associada a uma árvore isolada, conferindo ao modelo maior capacidade de generalização. A Figura 22 ilustra esse esquema.

Neste trabalho, o RF foi empregado tanto em regressão, para aproximar o valor contínuo do MaxSlope, quanto em classificação binária, para distinguir as simulações com e sem alternans. A construção da floresta é essencialmente a mesma nos dois casos. As diferenças estão no critério de divisão dos nós, baseado na impureza de Gini na classificação e na redução da variância na regressão, e na forma de agregação das árvores, feita por voto majoritário na classificação e por média na regressão. No contexto deste trabalho, o RF é integrado no procedimento de filtragem de Monte-Carlo, permitindo avaliar a contribuição dos parâmetros eletrofisiológicos na discriminação entre os regimes com e sem alternans.

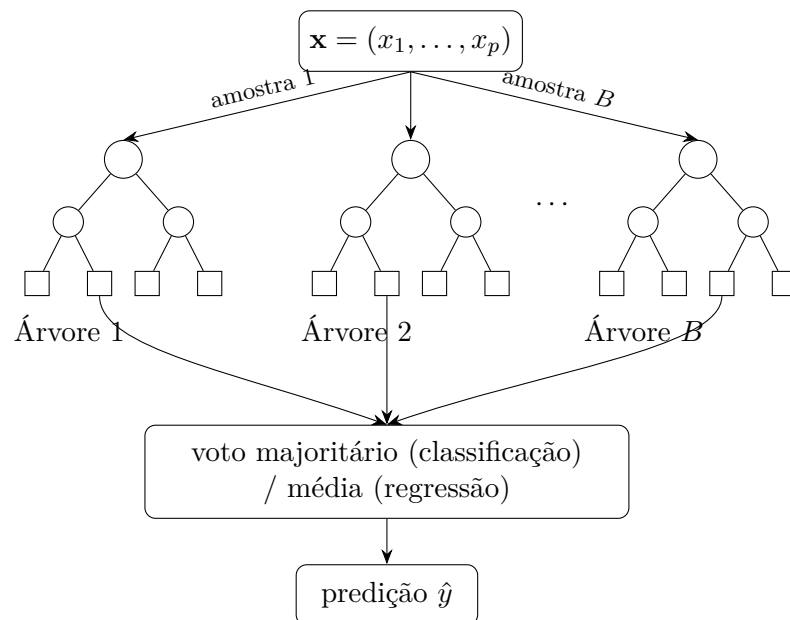


Figura 22 – Esquema ilustrativo de uma *Random Forest*: o vetor de entrada $\mathbf{x}$ é avaliado por um conjunto de B árvores de decisão, cada uma treinada sobre uma amostra *bootstrap* distinta do conjunto de dados. As predições individuais das árvores são combinadas por voto majoritário (classificação) ou por média (regressão), resultando na predição final $\hat{y}$.

4 MÉTODOS

Enquanto os capítulos anteriores apresentaram os modelos da eletrofisiologia cardíaca e as técnicas de análise de sensibilidade e de modelos substitutos de forma geral, este capítulo descreve como esses elementos são combinados para compor a abordagem proposta neste trabalho. A essência do capítulo está em detalhar a metodologia empregada para identificar os parâmetros mais relevantes à ocorrência de alternans, articulando as quantidades de interesse, os modelos substitutos e as técnicas de análise de sensibilidade. Para isso, apresentam-se, em sequência: as quantidades de interesse adotadas nas análises (APD e MaxSlope); a estratégia de combinação dos emuladores com as técnicas de índices de Sobol e de Filtragem de Monte Carlo; e, por fim, as métricas de erro e de qualidade utilizadas para avaliar e comparar as diferentes abordagens.

4.1 Quantidades de interesse

As análises de sensibilidade deste trabalho são conduzidas sobre duas quantidades de interesse (QoI, do inglês *quantity of interest*) associadas ao comportamento de alternans, obtidas a partir das simulações dos modelos celulares: a duração do potencial de ação (APD) e a máxima inclinação da curva de restituição (MaxSlope).

A primeira QoI é o APD, definido na Seção 2.6.2 como a medida APD90, isto é, o tempo até que o potencial de ação atinja 90% de sua repolarização. O APD é uma quantidade central no estudo de alternans, uma vez que o próprio fenômeno se manifesta como uma variação, batimento a batimento, na duração do potencial de ação. Para computá-lo, cada modelo celular foi submetido a um *pre pacing* em um comprimento de ciclo básico fixo de 1000 ms até atingir o estado estacionário, tomando-se então o APD correspondente ao último pulso aplicado.

A segunda QoI é o MaxSlope, definido como a inclinação máxima da curva de restituição do APD, isto é,

$$\text{MaxSlope} = \max |f'(DI)|, \quad (4.1)$$

onde f é a relação entre o APD e o intervalo diastólico DI . A escolha desta QoI justifica-se por ela estar diretamente ligada ao critério de predição de alternans introduzido na Seção 2.6.2: segundo o critério de Nolasco e Dahlen (1968), a ocorrência de alternans está associada à condição $\text{MaxSlope} > 1$, enquanto $\text{MaxSlope} \leq 1$ corresponde à sua ausência.

Para obter o MaxSlope, empregou-se o protocolo de restituição dinâmico para gerar a curva de restituição de cada modelo. Nesse protocolo, o comprimento do ciclo de estímulo $S1$ é decrementado sequencialmente após um número previamente escolhido de batimentos N_{beats} , partindo de um intervalo de acoplamento inicial CI_1 até que um intervalo de acoplamento mínimo CI_0 seja atingido. O intervalo de acoplamento é o tempo

entre o início de um dado estímulo e o início do potencial de ação precedente. Neste trabalho, adotou-se $N_{\text{beats}} = 20$ para todos os modelos; para o modelo MMS, os valores de CI_1 e CI_0 foram fixados em 1000 e 200, respectivamente, enquanto para os modelos TP e Mahajan utilizou-se $CI_1 = 600$ e $CI_0 = 200$.

Assim, enquanto o APD fornece uma medida direta da grandeza que oscila no alternans, o MaxSlope fornece um critério preditor da ocorrência dessa oscilação. Ambas as quantidades são empregadas nas análises de sensibilidade, e o critério $\text{MaxSlope} > 1$ é adotado para a classificação binária entre os regimes com e sem alternans, base para as análises de Filtragem de Monte Carlo apresentadas adiante.

4.2 Índices de Sobol

Para o cálculo dos índices de Sobol de APD e MaxSlope seguiu-se o fluxo usual de análise de sensibilidade baseada em emuladores, no qual estes são construídos a partir de um conjunto de dados de treinamento e com eles os índices de Sobol são determinados, conforme ilustrado na Figura 23. Nesse caso, aplicaram-se os emuladores baseados em PCE e GP para obter estimativas dos índices de Sobol ao invés de utilizá-los para uma análise de Monte-Carlo tradicional. Mais detalhes destes procedimentos podem ser encontrados na literatura (SUDRET; MAI, 2015; MARELLI; SUDRET,).

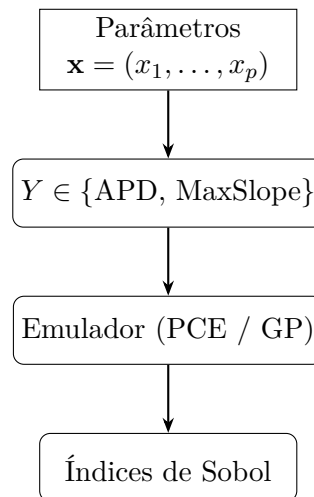


Figura 23 – Fluxo para o cálculo dos índices de Sobol. Cada quantidade de interesse, APD ou MaxSlope, é aproximada por um emulador (PCE ou GP), a partir do qual os índices de Sobol são obtidos.

4.3 Filtragem de Monte-Carlo

Para a análise de sensibilidade do índice MaxSlope utilizando a técnica MCF, foram investigadas diferentes estratégias de construção de modelos substitutos, conforme ilustrado na Figura 24. Em todas as abordagens, o ponto de partida consiste no conjunto de parâmetros de entrada $\mathbf{x}$, a partir do qual se busca identificar a ocorrência de alternans.

Duas estratégias distintas foram consideradas. Na primeira, emprega-se um modelo de regressão para aproximar diretamente o valor de MaxSlope (Y). Em seguida, a classificação entre os casos com e sem alternans é obtida por meio do critério $Y > 1$, sendo as amostras então utilizadas na etapa de Filtragem de Monte Carlo. Nessa categoria foram avaliados modelos substitutos baseados em PCE, GP e RF.

Na segunda estratégia, a tarefa de classificação é realizada diretamente pelo modelo substituto, dispensando a etapa intermediária de predição de MaxSlope. Nesse caso, os classificadores fornecem diretamente a classe correspondente à presença ou ausência de alternans, que é posteriormente empregada pela técnica de MCF. Para essa abordagem foram considerados modelos baseados em Processos Gaussianos (GP) e Florestas Aleatórias (RF).

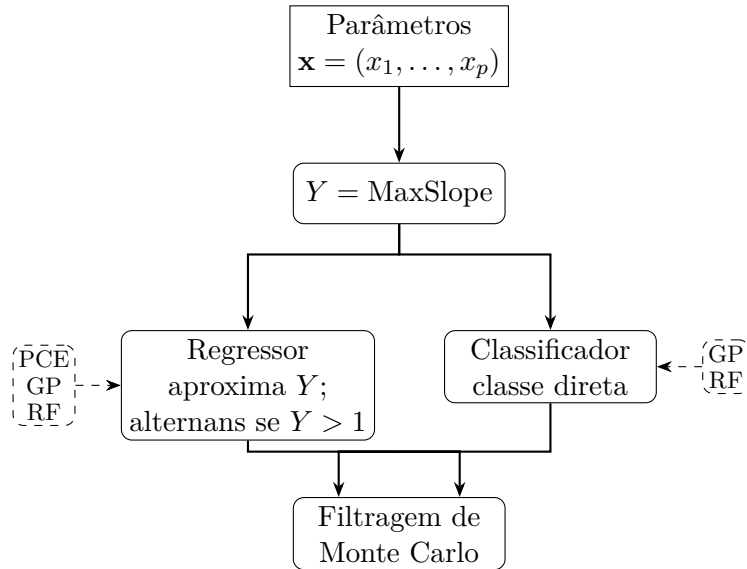


Figura 24 – Fluxo para a detecção de alternans via Filtragem de Monte Carlo. A partir do MaxSlope, um regressor aproxima Y e aplica o critério $Y > 1$, ou um classificador prevê a classe diretamente. As técnicas indicadas ao lado de cada ramo mostram que PCE atua apenas como regressor, enquanto GP e RF atuam tanto na regressão quanto na classificação.

4.4 Métricas

A seguir são definidas as métricas utilizadas neste trabalho para avaliar a qualidade dos modelos substitutos baseados em regressão ou classificação.

4.4.1 Métricas de precisão na regressão

A qualidade dos emuladores de regressão foi avaliada por meio de duas métricas: o coeficiente de determinação R^2 e o erro de validação cruzada *leave-one-out* (LOO).

O coeficiente de determinação mede a fração da variância dos dados que é explicada pelo emulador, sendo dado por

$$R^2 = 1 - \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - \bar{y})^2}, \quad (4.2)$$

onde y_i são os valores de referência obtidos diretamente do modelo, $\hat{y}_i$ são os valores preditos pelo emulador, $\bar{y}$ é a média dos valores de referência e N é o número de amostras de validação. Valores de R^2 próximos de 1 indicam uma aproximação de alta qualidade.

O erro LOO estima a capacidade de generalização do emulador sem a necessidade de um conjunto de validação independente. Para um conjunto de N amostras, cada amostra i é sucessivamente retirada do conjunto de treinamento, o emulador é reconstruído com as $N - 1$ amostras restantes, e o valor da amostra retirada é predito, obtendo-se $\hat{y}_{-i}$. O erro LOO é então definido como

$$\text{LOO} = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_{-i})^2, \quad (4.3)$$

de modo que valores menores indicam melhor capacidade preditiva.

4.4.2 Métricas de precisão na classificação

Para os métodos empregados na classificação entre os regimes com e sem alternans, foi utilizada a acurácia balanceada (BA, do inglês *balanced accuracy*). Para isso, considera-se como classe positiva a ocorrência de alternans.

A acurácia balanceada é definida como a média entre a sensibilidade e a especificidade do classificador:

$$\text{BA} = \frac{1}{2} \left(\frac{\text{TP}}{\text{TP} + \text{FN}} + \frac{\text{TN}}{\text{TN} + \text{FP}} \right), \quad (4.4)$$

onde TP são os verdadeiros positivos (alternans predito como alternans), TN os verdadeiros negativos (não-alternans predito como não-alternans), FN os falsos negativos (alternans predito como não-alternans) e FP os falsos positivos (não-alternans predito como alternans). O uso da acurácia balanceada, em vez da acurácia simples, justifica-se pelo desbalanceamento entre as classes: como os casos de alternans são menos frequentes, a acurácia balanceada evita que o bom desempenho na classe majoritária mascare um desempenho ruim na classe minoritária.

A AUC corresponde à área sob a curva ROC (*Receiver Operating Characteristic*), que relaciona a taxa de verdadeiros positivos e a taxa de falsos positivos ao longo de todos os limiares de decisão possíveis. Ela pode ser interpretada como a probabilidade de que uma amostra da classe positiva (alternans), escolhida aleatoriamente, receba um *score* maior do que o de uma amostra da classe negativa (não-alternans). A AUC varia entre 0 e 1: valores próximos de 1 indicam alta capacidade discriminativa, enquanto 0,5 corresponde a uma classificação aleatória. Diferentemente da acurácia balanceada, que depende de um limiar fixo, a AUC avalia a qualidade do ordenamento (*ranking*) produzido pelo método.

4.4.3 Implementação Computacional

As simulações numéricas da eletrofisiologia cardíaca utilizando os modelos MMS, TP e Mahajan-Shiferaw deste trabalho foram todas realizadas utilizando-se o software openCARP (PLANK et al., 2021; openCARP consortium et al., 2024). Para isso, foi necessário implementar um módulo adicional do software que descreve as equações do modelo MMS, uma vez que o software não dispunha deste modelo. Por outro lado, os modelos TP e Mahajan-Shiferaw já fazem parte do módulo de modelos celulares existentes dentro deste ambiente de simulação e foram utilizados sem alteração. Todo o processamento subsequente (amostragem dos parâmetros, execução das simulações, coleta das saídas e avaliação das métricas) foi implementado em linguagem Python.

Os modelos substitutos foram construídos com duas ferramentas computacionais complementares: o *framework* UQpyLab (MARELLI; SUDRET,), empregado na construção dos emuladores baseados em PCE e GP para regressão, e a biblioteca *scikit-learn* (PEDREGOSA et al., 2011), empregada nos classificadores baseados em GP e em RF. As configurações específicas de cada técnica são detalhadas nos capítulos de resultados correspondentes. As métricas de erro e de qualidade descritas na Seção 4.4 são calculadas de forma uniforme sobre as saídas de todos os emuladores, independentemente do *framework* utilizado.

A análise de Filtragem de Monte Carlo foi implementada em código próprio Python utilizando os emuladores. O procedimento comum a todos os experimentos deste trabalho é composto por quatro etapas. Primeiramente, constrói-se um modelo substituto (PCE, GP ou RF) para a quantidade de interesse MaxSlope a partir do conjunto de treinamento. Em seguida, um grande conjunto de 10000 amostras dos parâmetros de entrada é gerado por amostragem por hipercubo latino (LHS) e avaliado pelo modelo substituto, obtendo-se as respostas Y a um custo computacional desprezível. Esse tamanho de amostra foi escolhido por ser suficientemente grande para garantir a convergência das estatísticas de interesse nas análises de Monte Carlo, sendo viabilizado pelo baixo custo de avaliação dos emuladores. Cada amostra é então classificada em um dos dois regimes, isto é, alternans, se $Y > 1$, ou não-alternans, caso contrário, no caso dos regressores, ou diretamente pelo classificador, no caso das abordagens de classificação. A partir dessa separação, calcula-se, para cada parâmetro de entrada, a estatística de Smirnov entre as distribuições dos dois regimes, quantificando a influência do parâmetro na ocorrência de alternans e permitindo a análise de sensibilidade regional descrita na Seção 3.2.

5 RESULTADOS PARA OS MODELOS MMS E TP

Neste capítulo apresentam-se os detalhes da descrição dos experimentos computacionais, os detalhes da amostragem e os resultados correspondentes à aplicação das técnicas de sensibilidade para cada modelo cardíaco abordado por este trabalho. As análises serão apresentadas utilizando o método Monte Carlo clássico como referência, assim como através da combinação da Filtragem de Monte Carlo com emuladores PCE para acelerar as análises. Os índices de Sobol também são empregados. Os resultados referentes aos modelos MMS e TP foram publicados em Pigozzo, Santos e Rocha (2025).

5.1 Descrição dos experimentos

Os experimentos computacionais foram realizados da seguinte forma. Para cada modelo estudado, foram feitas as análises de sensibilidade utilizando o método MC com um alto número de amostras para se obter soluções e respostas de referência, as quais foram comparadas com os resultados das análises obtidas através de emuladores PCE. Para cada modelo, foram realizadas análises de sensibilidade utilizando a técnica de filtragem de Monte Carlo (MCF) e através do cálculo dos índices de Sobol. Adicionalmente, para cada modelo, uma análise de sensibilidade do tipo *one-at-a-time* (OAT) é apresentada de forma inicial como forma de complementar o entendimento dos parâmetros de cada modelo.

Os experimentos computacionais deste estudo foram realizados utilizando a linguagem de programação Python, aprimorada para computação científica, aproveitando-se das bibliotecas NumPy e SciPy (VIRTANEN et al., 2020). Adicionalmente foram utilizadas as bibliotecas SALib (IWANAGA; USHER; HERMAN, 2022) e UQpyLab (SUDRET, 2008).

Neste trabalho foram calculados os índices de sensibilidade Sobol de primeira ordem e ordem total para as QoIs escolhidas em estudos realizados para os modelos MMS e TP. Na análise com o método de MC, os índices foram calculados utilizando o método de Saltelli et al. (2010) disponível na implementação da biblioteca SALib (IWANAGA; USHER; HERMAN, 2022). Por outro lado, quando as análises foram feitas através dos emuladores PCE, os índices de Sobol foram calculados a partir de seus coeficientes, como implementado na biblioteca UQpyLab.

A aplicação da técnica MCF foi implementada em código próprio utilizando apenas dos recursos da linguagem Python. A análise baseada no MC utilizou os dados de todas as amostras geradas, enquanto a análise baseada nos emuladores executou a filtragem através da avaliação dos emuladores polinomiais PCE.

5.1.1 Escolha dos parâmetros

Para as análises do modelo MMS foram escolhidos os quatro parâmetros que correspondem às 4 constantes de tempo do modelo conhecidos como τ_{in} , τ_{out} , τ_{open} e τ_{close} . Por outro lado, como o modelo TP tem muitos parâmetros, o que tornaria a análise computacional altamente onerosa, foram escolhidos apenas os parâmetros relacionados às condutâncias dos canais iônicos para análise, como feito no trabalho de Corso e Viola (2020). Os parâmetros do modelo TP utilizados foram as condutâncias dos canais iônicos: G_{CaL} , G_{K1} , G_{Kr} , G_{Ks} , G_{Na} , G_{bCa} , G_{bNa} , G_{pCa} , G_{pK} e G_{to} .

As Tabelas 1 e 2, apresentadas no Capítulo 2, mostram, respectivamente, o espaço de parâmetros utilizados para a realização dos estudos dos modelos MMS e TP. Esses intervalos foram selecionados tomando os valores base como média das distribuições uniformes e adicionando uma perturbação ao redor destes para determinar os limites inferior e superior. Esta perturbação foi adicionada de forma que nos dois modelos o coeficiente de variação (CoV) dado por $CoV = \sigma/\mu$, onde σ é o desvio padrão e μ a média, fosse fixo para cada parâmetro do modelo. Para o modelo MMS, foi considerado um CoV de aproximadamente 38,5%, enquanto que, para o modelo TP, esse CoV foi de, aproximadamente 17%.

5.1.2 Amostragem

Para gerar uma base de dados para as análises de MC e também para a construção dos emuladores PCE, geraram-se $N_s = N(2D + 2)$ amostras para os parâmetros de entrada, onde D é o número de parâmetros de entrada e N um tamanho de amostras a ser escolhido. Os emuladores PCE foram construídos com um número de amostras N_{PCE} significativamente menor do que N_s e escolhidas aleatoriamente dentro deste conjunto de amostras.

Para as análises de ambos os modelos, MMS e TP, foi utilizado $N = 1024$. Como o modelo MMS possui $D = 4$ parâmetros de entrada, isso resultou em um total de $N_s = 10240$ amostras e simulações para cada QoI. Para o modelo TP, como foram escolhidas as dez condutâncias dos canais iônicos deste modelo para análise, tem-se que $D = 10$ e, portanto, resultou em $N_s = 22528$ amostras e simulações para cada QoI.

5.1.3 Construção dos emuladores PCE

Para a construção dos emuladores PCE para cada QoI, selecionou-se aleatoriamente N_{PCE} amostras e suas respectivas saídas do conjunto de dados de N_s amostras geradas para análise pelo método MC. Foram construídos emuladores com diferentes valores de N_{PCE} para se avaliar a sua precisão, conforme será relatado adiante. Para a construção de ambos os emuladores PCE, dos modelos MMS e TP, um total de 75% de amostras

N_{PCE} foram utilizadas para treinamento, enquanto as demais 25% das amostras foram reservadas para teste.

Na Tabela 4 são apresentadas mais informações sobre os modelos MMS e TP, grau escolhido para a construção dos emuladores PCE e o método utilizado para emulador. A biblioteca **UQpyLab** oferece diversos métodos para a construção de emuladores, como OLS, LARS, OMP, entre outros. Os métodos OLS (*Ordinary Least Squares*), LARS (*Least Angle Regression*) e OMP (*Orthogonal Matching Pursuit*) são métodos de regressão usados para estimar os coeficientes de um modelo linear, onde o OLS é o mínimos quadrados usual que minimiza a soma dos erros quadráticos, enquanto LARS e OMP são métodos de regressão esparsa que selecionam um subconjunto de variáveis relevantes em cenários onde o número de variáveis explicativas é grande. Para mais detalhes sobre estes métodos veja Luethen, Marelli e Sudret (2021) e Sudret e Mai (2015). Em uma análise preliminar, observou-se que o método LARS apresentou melhor desempenho para a variável APD, enquanto o método OMP mostrou-se mais adequado para MaxSlope. A seleção do grau polinomial para cada caso foi fundamentada em uma análise exploratória preliminar, levando em consideração o equilíbrio entre o custo computacional e a precisão do emulador. Como resultado, adotaram-se emuladores de grau 4 para APD e de grau 7 para MaxSlope.

Tabela 4 – Dados sobre a construção dos emuladores PCE para os modelos MMS e TP.

Modelo	N_p	N_s	N_{PCE}	Grau	Método
MMS	4	10240	1000 (APD)	4	LARS
			1000 (MaxSlope)	7	OMP
TP	10	22528	1000 (APD)	4	LARS
			1000 (MaxSlope)	7	OMP

Realizou-se uma análise do erro LOO e do coeficiente R^2 dos emuladores para cada variável, com o objetivo de determinar o número adequado de amostras para sua construção.

A Figura 25 apresenta, para diferentes valores de N_{PCE} , os resultados dessas duas métricas para ambos os modelos analisados. Observa-se que, em todos os casos, para N_{PCE} superiores a 1000, o coeficiente R^2 permanece próximo de 1 (como desejado), e o erro se mantém dentro de limites satisfatórios. Com base nessa análise, foi escolhido $N_{PCE} = 1000$ (conforme indicado na Tabela 4) para a construção de todos os emuladores, buscando reduzir o custo computacional no treinamento e aumentar a precisão dos modelos.

5.2 Modelo MMS

Inicialmente são apresentados os resultados para o modelo MMS, que, além de mais simples, permite a comparação dos resultados numéricos obtidos através de simulações

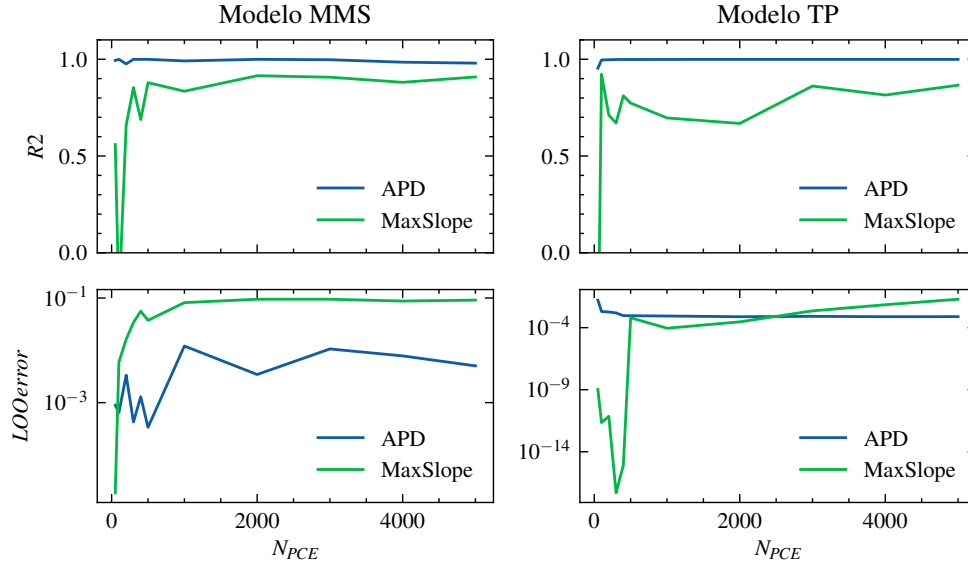


Figura 25 – Análise da qualidade dos emuladores PCE para os modelos MMS e TP em função do tamanho do conjunto de dados N_{PCE} usado em sua construção: em cima, coeficientes R^2 , e em baixo, erro LOO.

com as expressões analíticas desenvolvidas para o modelo original de Mitchell e Schaeffer (2003), que também são válidas para o modelo modificado para o caso de alternans elétrico.

5.2.1 Análise de sensibilidade local

A Figura 26 mostra o resultado da análise OAT dos parâmetros do modelo MMS para o APD. Note que, para o APD, pode-se perceber que 3 dos 4 parâmetros analisados deste modelo, τ_{in} , τ_{out} e τ_{close} , influenciam de forma significativa o APD, como era de se esperar tendo em vista a Eq. (2.13). A influência dos parâmetros, no entanto, é diferente, pois, um aumento no valor de τ_{in} , reduz o valor do APD. Esse comportamento é o contrário com os outros dois parâmetros, τ_{out} e τ_{close} . Por outro lado, temos que o parâmetro τ_{open} não tem influência sobre o APD.

A análise OAT dos parâmetros do modelo MMS para a inclinação máxima da curva de restituição (MaxSlope) é apresentada na Figura 27. Pode-se observar novamente que os mesmos três parâmetros, τ_{in} , τ_{out} e τ_{close} , apresentam influência mais significativa na inclinação da parte final da curva de restituição, enquanto que o τ_{open} apresenta uma influência baixa na inclinação do final da curva de restituição.

5.2.2 Índices de Sobol

A Figura 28 a seguir apresenta o resultado da análise de sensibilidade através dos índices de Sobol dos parâmetros do modelo MMS, para as quantidades de interesse, APD e MaxSlope, respectivamente, onde $S1$ representa os índices de primeira ordem e ST representa os índices de Sobol de ordem total. O gráfico de barras à esquerda da

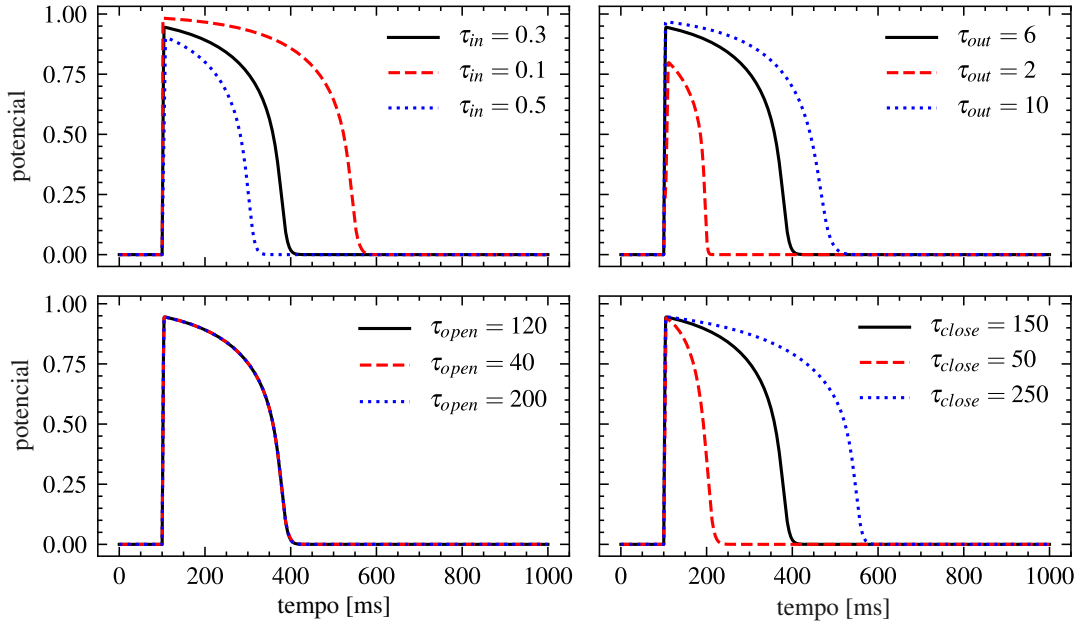


Figura 26 – SA local OAT dos parâmetros do modelo MMS para o APD.

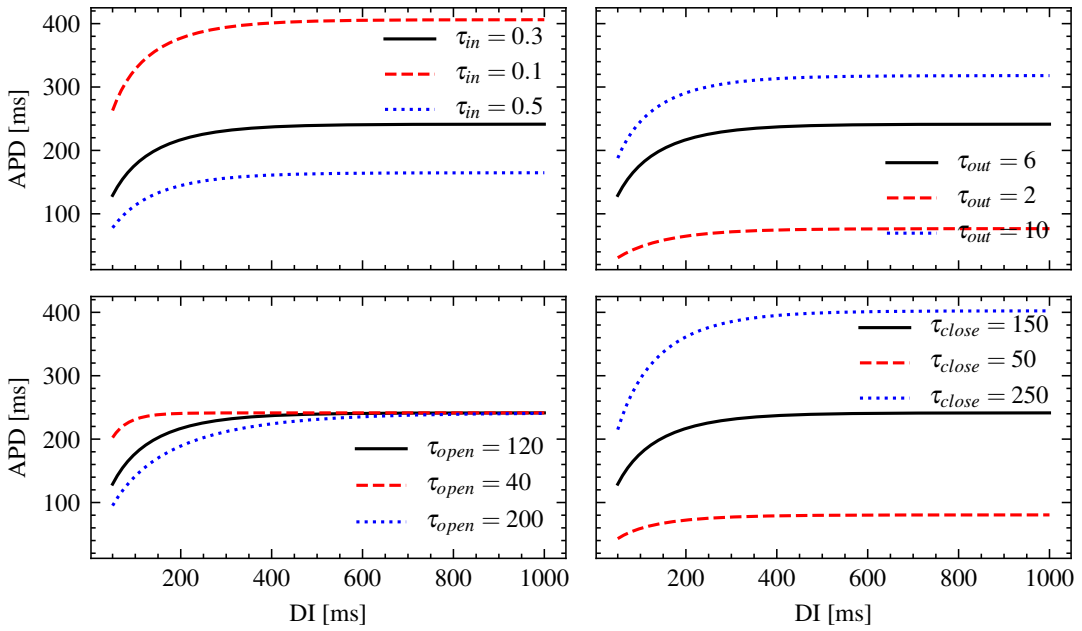


Figura 27 – SA local OAT dos parâmetros do modelo MMS para o MaxSlope.

Figura 28 apresenta o resultado para o APD, enquanto o gráfico de barras à direita mostra o resultado para o MaxSlope.

Para o APD, pode-se observar que τ_{close} é o parâmetro mais influente, seguido dos parâmetros τ_{out} e τ_{in} que também exercem alguma influência no comportamento do APD. Por esta análise, observa-se que o parâmetro τ_{open} não influencia no APD de forma significativa, confirmando aquilo que foi verificado também na análise OAT apresentada na Figura 26 anteriormente.

No gráfico à direita da Figura 28 apresenta-se os índices de Sobol dos parâmetros do MMS para a quantidade MaxSlope. Os resultados mostram tendência similar ao observado para APD, uma vez que, o parâmetro τ_{close} tem a maior influência na mudança da inclinação da curva.

Além disso, nestas duas análises, observa-se que, tanto para o APD quanto para o MaxSlope, os parâmetros mais influentes apresentam interações relevantes com os demais, evidenciadas pela diferença entre os índices de Sobol totais e os de primeira ordem. Esse efeito é mais pronunciado para o parâmetro τ_{close} : no APD, seu índice total é próximo de 0,6, enquanto o de primeira ordem é de aproximadamente 0,4; no MaxSlope, a diferença é ainda maior, com índice total em torno de 0,7 contra cerca de 0,5 de primeira ordem. Para os demais parâmetros, a diferença entre os dois índices é menor, indicando contribuições predominantemente individuais.

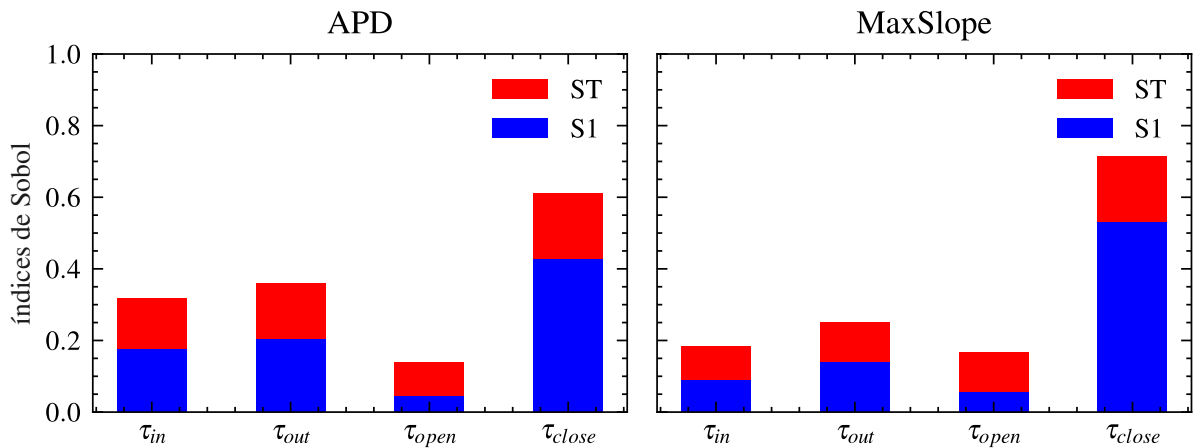


Figura 28 – SA global pelo cálculo dos Índices de Sobol dos parâmetros do modelo MMS para as QoI: APD e MaxSlope.

5.2.3 Filtragem de Monte Carlo

Em seguida procedeu-se à análise pelo MCF, que visa classificar e indicar quais os parâmetros são mais influentes para gerar determinado comportamento do modelo, que para o caso deste trabalho, é o alternans do APD. A Figura 29 apresenta o resultado da análise de sensibilidade feita utilizando a técnica MCF através do cálculo da estatística

de Smirnov ($d_{n,\bar{n}}$) para cada um dos parâmetros do modelo MMS, onde o resultado do valor do teste mede diretamente a influência de cada parâmetro na ocorrência do alternans elétrico.

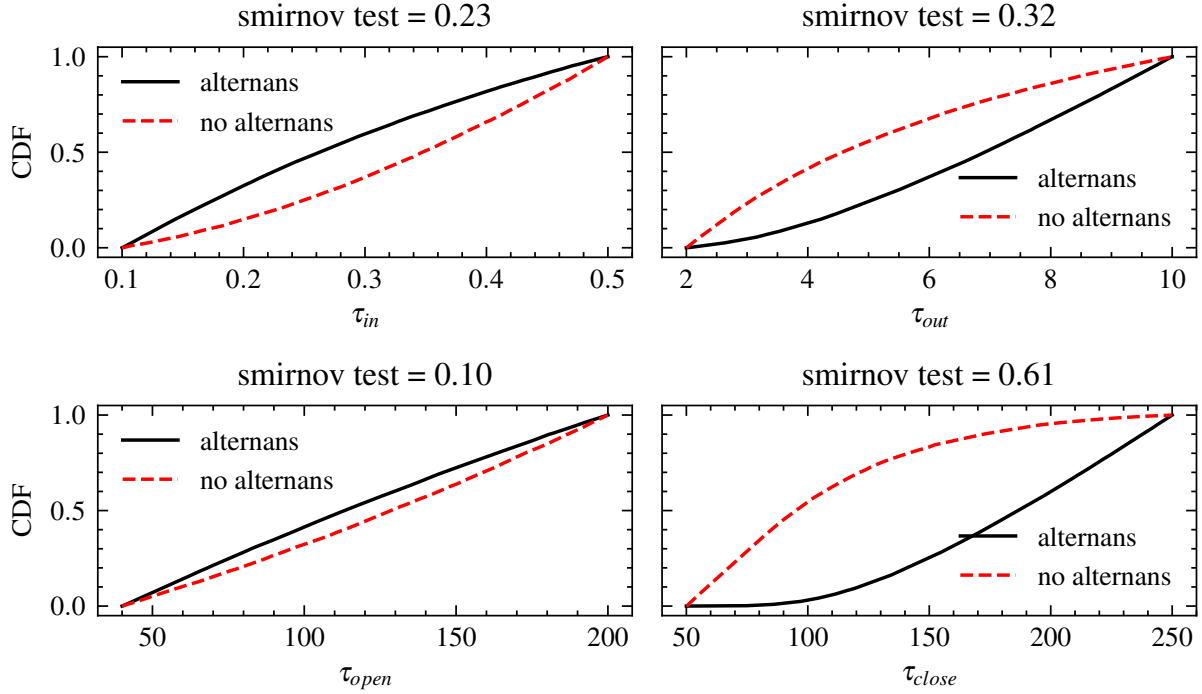


Figura 29 – Resultados do MCF e os valores da estatística de Smirnov para os parâmetros τ_{in} , τ_{out} , τ_{open} e τ_{close} do modelo MMS obtidos através de simulação computacional do modelo no OpenCARP.

Considerando que o parâmetro τ_{close} obteve o maior valor para o Teste de Smirnov, com um valor de 0.61, pode-se dizer que é o mais influente para a ocorrência de alternans. Os parâmetros, τ_{out} e τ_{in} , tiveram valores menores para o Teste de Smirnov, mas com alguma relevância e, portanto, pode-se dizer, que estes possuem alguma influência para a ocorrência de alternans, enquanto o parâmetro τ_{open} é o de menor importância na ocorrência deste fenômeno.

Estes resultados estão em boa concordância com o que foi obtido pela análise OAT e pelo cálculo dos Índices de Sobol feitos anteriormente. Entretanto, vale ressaltar que na técnica MCF os resultados do modelo são classificados dentre aqueles que geraram o comportamento (alternans) e os que não geraram, para então se realizar o Teste de Smirnov. Portanto, diferentemente das análises anteriores, essa análise leva em conta se ocorreu alternans ou não, enquanto as anteriores apenas avaliam QoI que estão associadas ao comportamento de alternans (como APD e MaxSlope).

Para aprofundar a análise, a Figura 30 apresenta os resultados dos gráficos de dispersão dos parâmetros do modelo MMS, dois-a-dois, assim como apresenta o cálculo do coeficiente de correlação (cc) entre os mesmos, que indica a possível contribuição conjunta dos parâmetros do modelo MMS na influência da ocorrência do alternans. Pode-se

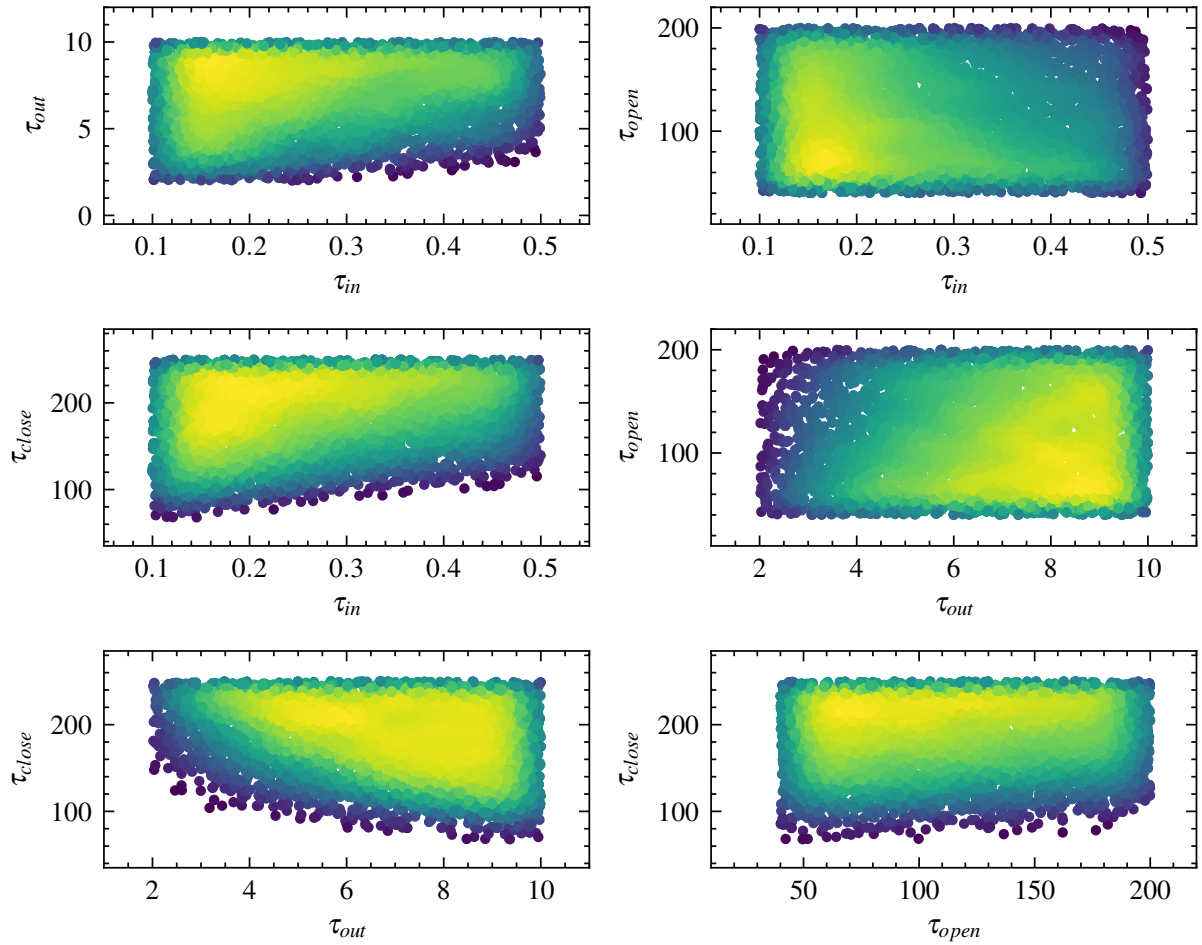


Figura 30 – Correlações (cc) de pares de parâmetros para os casos de ocorrência de alternans do modelo MMS. A cor indica a densidade local de pontos, das regiões menos densas (roxo) às mais densas (amarelo); as áreas amarelas correspondem, portanto, às combinações de parâmetros mais frequentes entre as amostras que produzem *alternans*.

destacar, portanto, a partir desta análise que os pares de parâmetros (τ_{in}, τ_{out}) , $(\tau_{in}, \tau_{close})$ e $(\tau_{out}, \tau_{close})$, que obtiveram coeficientes de correlação de 0,16, 0,20 e 0,25, respectivamente, contribuem, de maneira conjunta, para a ocorrência de alternans, apesar dessa correlação ser moderada.

A Figura 31 apresenta a estabilidade da estatística de Smirnov obtida pela análise MCF em função do tamanho amostral N_s , para cada um dos quatro parâmetros do modelo MMS na quantidade MaxSlope. Para cada valor de N_s , subconjuntos de amostras foram sorteados aleatoriamente do conjunto total e a estatística de Smirnov foi recalculada, sendo mostradas a média (curvas) e o desvio-padrão (faixas sombreadas) sobre as repetições. Observa-se que a estatística de cada parâmetro se estabiliza à medida que N_s aumenta, com as faixas de incerteza estreitando-se progressivamente, o que indica que o tamanho amostral empregado nas análises é suficiente para uma estimativa estável.

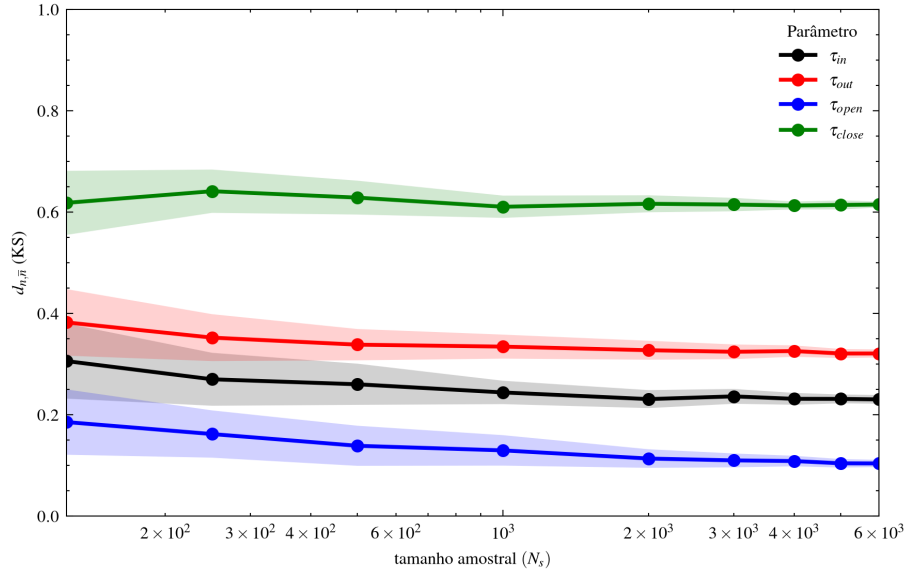


Figura 31 – Convergência da estatística de Smirnov ($d_{n,n}$) da análise MCF em função do tamanho amostral N_s , para cada parâmetro do modelo MMS na quantidade MaxSlope. As curvas mostram a média e as faixas sombreadas o desvio-padrão sobre repetições da subamostragem; as linhas correspondem, de cima para baixo, a τ_{close} (verde), τ_{out} (vermelho), τ_{in} (preto) e τ_{open} (azul).

5.2.4 Resultados com emuladores PCE do modelo MMS

Inicialmente, antes de apresentar os resultados pelos emuladores PCE, serão apresentados resultados a respeito da qualidade e capacidade de previsão dos emuladores criados para APD e MaxSlope. O conjunto de amostras de teste, descrito anteriormente na Seção 5.1.3, foi utilizado neste estudo para verificar a precisão dos emuladores PCE.

A Figura 32 apresenta o resultado do histograma do modelo MMS original em comparação com o seu emulador PCE em relação aos valores do APD calculados. Note-se que resposta obtida pelo emulador PCE do modelo MMS é satisfatória, tanto ao se analisar o histograma (à esquerda) de valores verdadeiros e previstos pelo emulador, quanto pelos resultados do gráfico de paridade (à direita), uma vez que os valores do APD alcançados pelo emulador PCE são bem próximos aos verdadeiros obtidos diretamente pelo modelo MMS. Essa boa precisão alcançada pelo PCE no gráfico de paridade (à direita) fica clara na obtenção do coeficiente R^2 de aproximadamente 0,9913, o que representa uma boa aproximação.

Para o emulador PCE da quantidade MaxSlope, os resultados são apresentados na Figura 33 para o modelo MMS em comparação com o emulador PCE. Note-se que a resposta obtida pelo emulador PCE do modelo MMS, em geral, é satisfatória. No histograma apresentado à esquerda para MaxSlope nota-se boa concordância, entretanto pelo gráfico de paridade (à direita) é possível perceber que o emulador PCE foi um pouco menos preciso do que para o caso da quantidade APD. Ainda assim, o emulador PCE

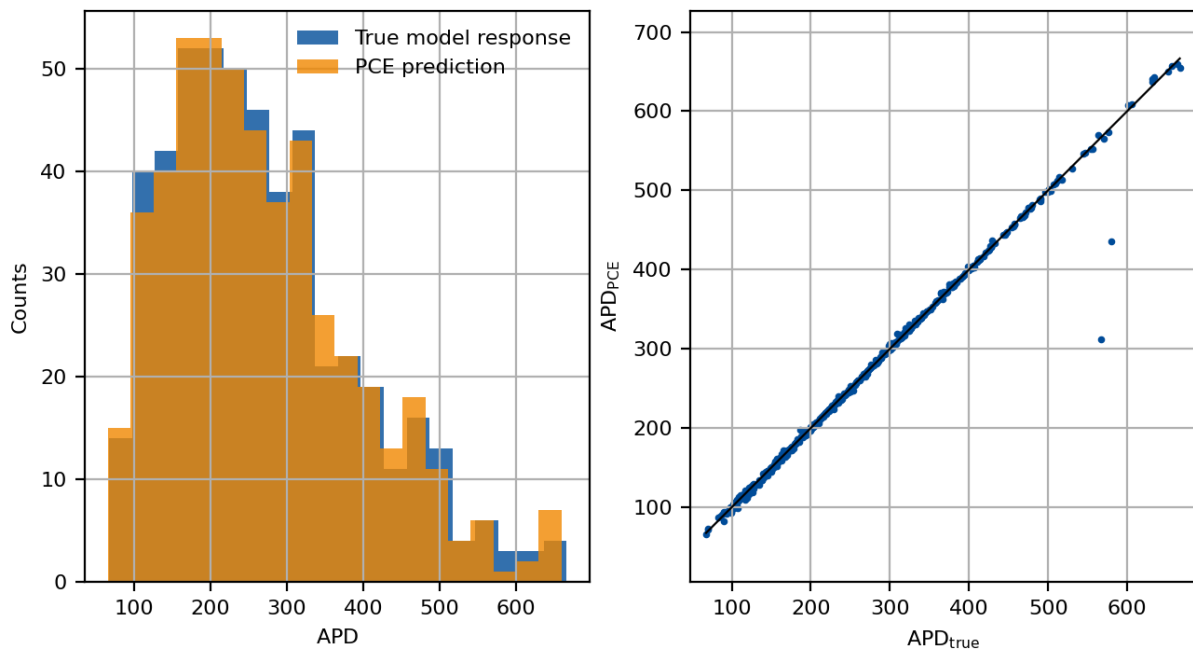


Figura 32 – Esquerda: histograma do APD obtido para um conjunto de amostras através do modelo MMS e do emulador PCE. Direita: gráfico de paridade dos valores do APD obtidos pelo modelo MMS e pelo emulador PCE.

respondeu bem quantitativamente tendo seu coeficiente R^2 calculado dando um valor de aproximadamente 0,8349.

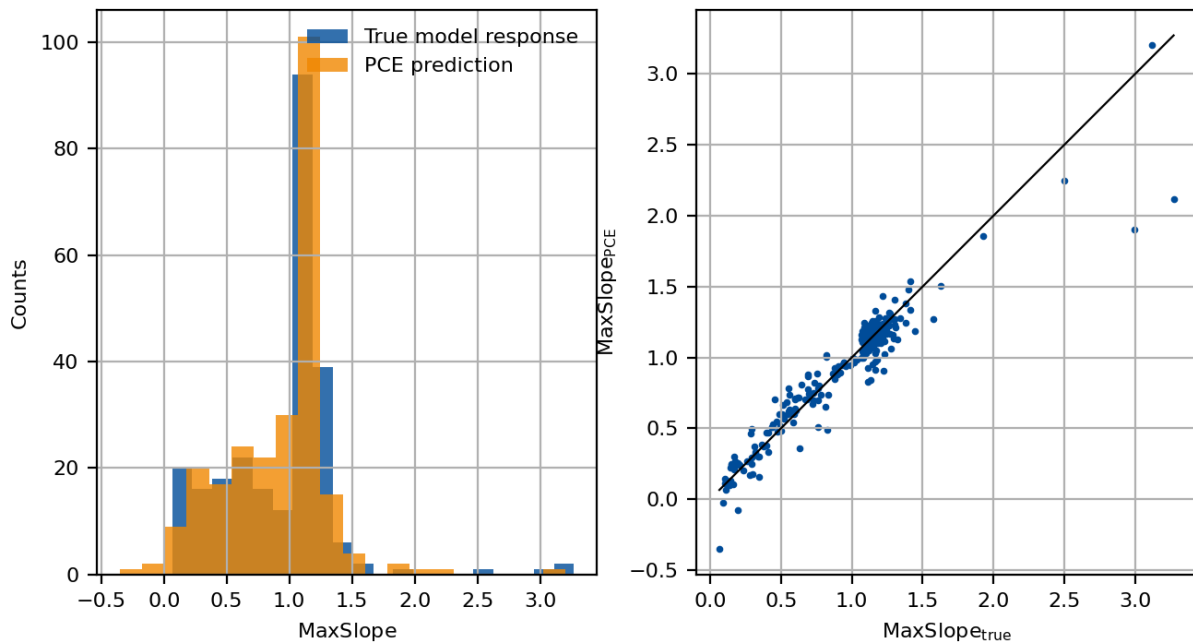


Figura 33 – Lado esquerdo está o histograma do emulador PCE do modelo MMS para o MaxSlope com uma predição boa do modelo original. Lado direito está a amostra, por meio dos gráficos de dispersão, dos valores do MaxSlope do modelo MMS original que estão próximos em comparação aos do modelo emulado pelo PCE.

A Tabela 5 apresenta os resultados do cálculo dos índices de Sobol, tanto de

primeira ordem quanto de ordem total, para os parâmetros do modelo MMS obtidos pelos emuladores PCE para o APD e MaxSlope, assim como o resultado do teste de Smirnov. Pode-se observar que os valores dos índices de Sobol $S1$ e ST encontrados para cada parâmetro estão próximos dos valores encontrados através do método de Saltelli mostrados anteriormente na Figura 28. Entretanto, vale destacar que, para o cálculo dos índices de Sobol apresentados na Figura 28 foram realizadas 10240 simulações para se calcular o APD e MaxSlope, que representa um alto custo computacional, sobretudo para a quantidade MaxSlope. Por outro lado, através do emulador PCE apenas 1000 simulações, para o APD e MaxSlope, foram realizadas para construir o emulador PCE, que uma vez construído pode facilmente calcular os índices de Sobol analiticamente com custo computacional baixo, conforme detalhado em Sudret (2008).

Tabela 5 – Cálculo dos Índices de Sobol e Teste de Smirnov para o modelo MMS através dos emuladores PCE.

Parâmetros	APD		MaxSlope		Alternans
	S1	ST	S1	ST	Teste de Smirnov
τ_{in}	0,15	0,18	0,08	0,18	0,20
τ_{out}	0,21	0,23	0,13	0,20	0,36
τ_{open}	0,00	0,00	0,04	0,20	0,08
τ_{close}	0,58	0,62	0,52	0,69	0,62

Por fim, a Tabela 6 apresenta o comparativo dos resultados alcançados a partir do cálculo dos índices de Sobol do modelo MMS e seu emulador PCE para o APD e para o MaxSlope. Pode-se notar que os resultados obtidos pelo PCE estão próximos aos do modelo, para cada parâmetro e para cada índice S1 e ST calculados.

Tabela 6 – Comparativo entre o cálculo dos Índices de Sobol do modelo MMS e do seu emulador PCE para o APD e o MaxSlope.

Parâmetros	APD		APD PCE		MaxSlope		MaxSlope PCE	
	S1	ST	S1	ST	S1	ST	S1	ST
τ_{in}	0,17	0,31	0,15	0,18	0,09	0,18	0,08	0,18
τ_{out}	0,20	0,36	0,21	0,23	0,14	0,25	0,13	0,20
τ_{open}	0,04	0,14	0,00	0,00	0,05	0,16	0,04	0,20
τ_{close}	0,42	0,61	0,58	0,62	0,53	0,71	0,52	0,69

5.2.5 Verificação e comparativo das análises com a fórmula analítica do modelo MMS

Aqui serão apresentados os resultados obtidos pela técnica MCF para o modelo MMS utilizando-se da fórmula analítica do modelo para a curva de restituição dada na Eq. (2.13) que permite determinar se irá ocorrer alternans, a partir da sua máxima inclinação. O objetivo deste estudo é verificar que a abordagem feita e apresentada anteriormente baseada em simulações numéricas através do OpenCARP está correta e consistente com os resultados desta fórmula obtida pela análise do modelo MMS.

A Figura 34 mostra a análise de sensibilidade utilizando a técnica MCF através do cálculo do teste de Smirnov para cada um dos parâmetros do modelo MMS, baseado na fórmula analítica da Eq. (2.13). Os resultados mostram que o parâmetro τ_{close} é muito importante e influente na ocorrência de alternans, com um valor de 0,55 para a estatística de Smirnov, seguido por τ_{out} , com 0,37, e τ_{in} , com 0,27, enquanto o parâmetro τ_{open} é o de menor relevância.

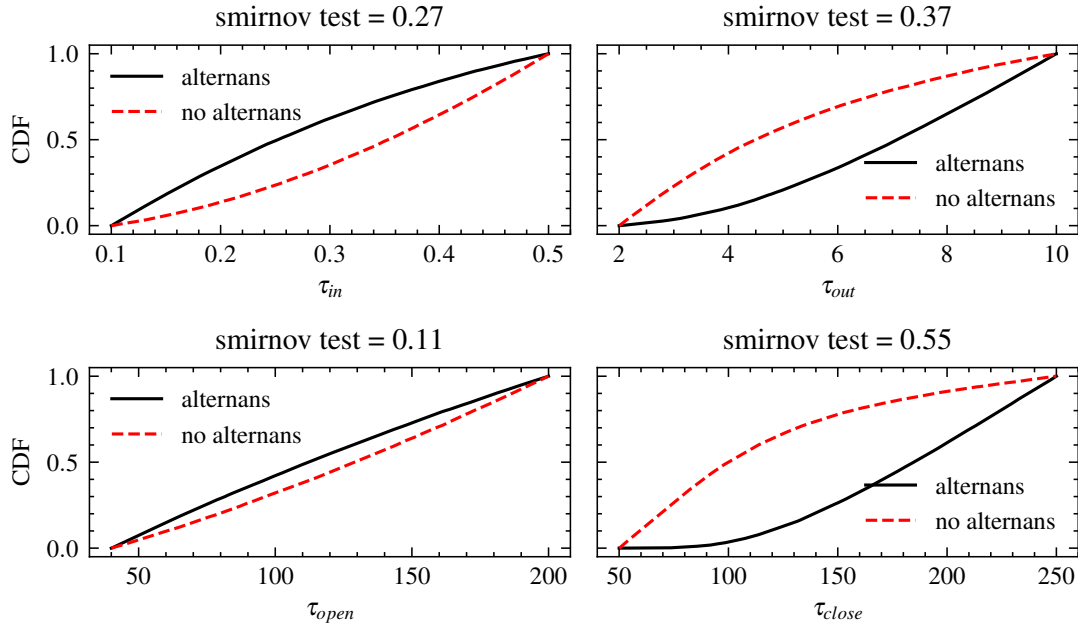


Figura 34 – Resultados do MCF e os valores da estatística de Smirnov para os parâmetros τ_{in} , τ_{out} , τ_{open} e τ_{close} do modelo MMS obtido a partir das expressões analíticas.

Em seguida, na Figura 35 são apresentados mapas de cores descrevendo as correlações, juntamente com o coeficiente de correlação (cc), de pares de parâmetros do modelo MMS com base na classificação da ocorrência de alternans através das expressões analíticas do modelo. Pode-se destacar, portanto, que os pares (τ_{in}, τ_{out}) , $(\tau_{in}, \tau_{close})$ e $(\tau_{out}, \tau_{close})$, com os coeficientes 0,24, 0,19 e 0,20, respectivamente, contribuem, de maneira conjunta, para a ocorrência de alternans. Os outros pares de parâmetros praticamente não contribuem conjuntamente para o alternans, uma vez que apenas correlações inferiores a 0,08 foram observadas.

Em seguida, será apresentado um comparativo das análises feitas pelas fórmulas analíticas, simulações numéricas pelo OpenCARP e os resultados obtidos através dos emuladores PCE tanto para os índices de Sobol quanto para o Teste de Smirnov. A Tabela 7 apresenta o comparativo dos resultados do MCF (Teste de Smirnov) alcançados pelas diferentes abordagens para o modelo MMS. Pode-se perceber uma boa concordância entre os valores dos parâmetros para cada abordagem (simulação numérica e emulador PCE), pois os resultados são próximos do que foi produzido pela fórmula analítica deste modelo. Adiante será apresentado um comparativo de desempenho computacional de cada

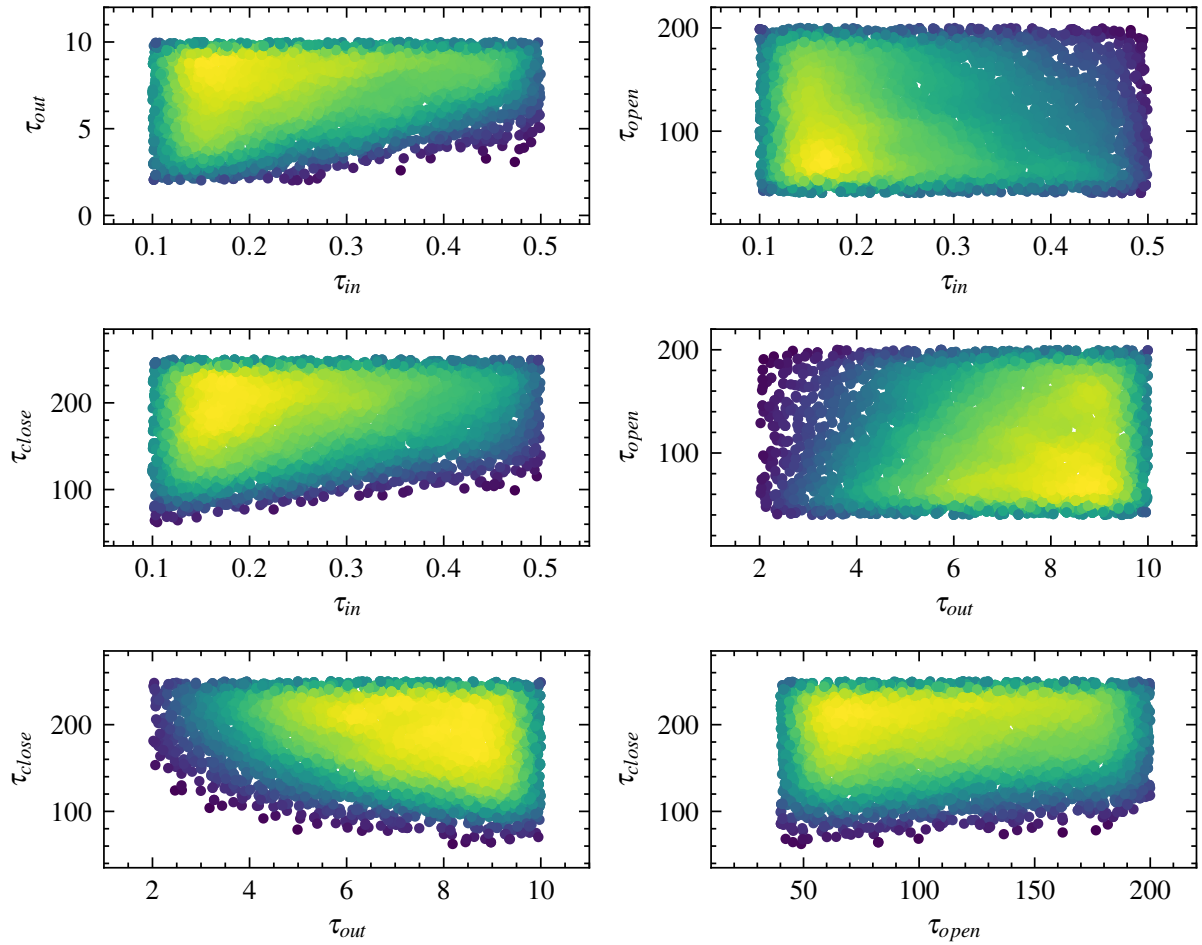


Figura 35 – Projeções bidimensionais das amostras associadas à ocorrência de *alternans* para cada par de parâmetros do modelo MMS. A cor indica a densidade local de pontos, das regiões menos densas (roxo) às mais densas (amarelo); as áreas amarelas correspondem, portanto, às combinações de parâmetros mais frequentes entre as amostras que produzem *alternans*.

abordagem.

Tabela 7 – Comparativo entre os valores do Teste de Smirnov para o modelo MMS obtidos de forma analítica, numérica (simulado pelo OpenCARP) e pelo emulador PCE.

Parâmetros	MMS Analítico	MMS Numérico	MMS PCE
τ_{in}	0,27	0,23	0,20
τ_{out}	0,37	0,32	0,32
τ_{open}	0,11	0,10	0,08
τ_{close}	0,55	0,61	0,62

A Tabela 8 apresenta o comparativo dos resultados alcançados a partir do cálculo dos coeficientes de correlação do modelo MMS numérico e analítico. Aqui mostra-se a consistência das técnicas aplicadas nas implementações, uma vez que, os resultados obtidos pelos experimentos computacionais estão próximos dos obtidos por meio das formulações analíticas do MMS. Pode-se notar que existe uma pequena interação entre os parâmetros

τ_{in} , τ_{out} e τ_{close} para a ocorrência de alternans e que ela é praticamente nula quando o parâmetro τ_{open} é envolvido.

Tabela 8 – Comparativo entre o cálculo dos coeficientes de correlação (cc) do modelo MMS por meio analítico e pelas simulações numéricas do OpenCARP.

Correlações	MMS Analítico	MMS Numérico
(τ_{in}, τ_{out})	0,24	0,16
$(\tau_{in}, \tau_{close})$	0,19	0,20
$(\tau_{out}, \tau_{close})$	0,20	0,25
demais pares	$\leq 0,08$	$\leq 0,08$

5.3 Modelo TP

Inicialmente os resultados da análise de sensibilidade local OAT para o modelo TP são apresentados nas Figuras 36 e 37 para quantidades de interesse APD e MaxSlope, respectivamente. Através dos resultados mostrados nas figuras, pode-se notar que, apenas três parâmetros G_{CaL} , G_{Kr} e G_{Ks} de todas as 10 condutâncias analisadas apresentam alguma influência significativa no APD do modelo TP: G_{CaL} e G_{Ks} de forma mais significativa e G_{Kr} de forma menos significativa. Os outros parâmetros não apresentaram influência significativa no APD através desta análise OAT.

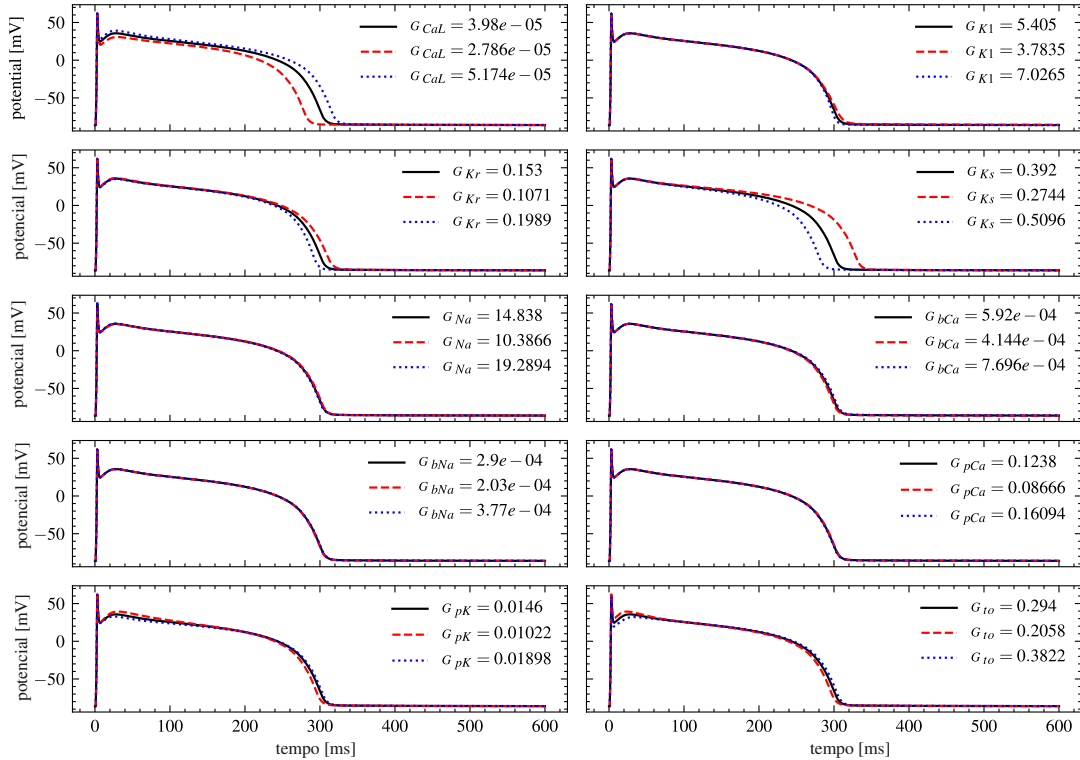


Figura 36 – Análise de sensibilidade local OAT para os parâmetros do modelo TP com relação a quantidade APD.

Através dos resultados da análise para MaxSlope, apresentados na Figura 37, pode-se perceber um comportamento semelhante dos parâmetros ao anterior, uma vez que apenas 4 dos 10 parâmetros geraram algum impacto nas curvas de restituição ao longo de toda a sua forma. Em particular, os efeitos mais significativos foram observados para os parâmetros G_{CaL} , G_{K1} , G_{Kr} e G_{Ks} , dentre os quais estão os 3 mais influentes para o comportamento do APD.

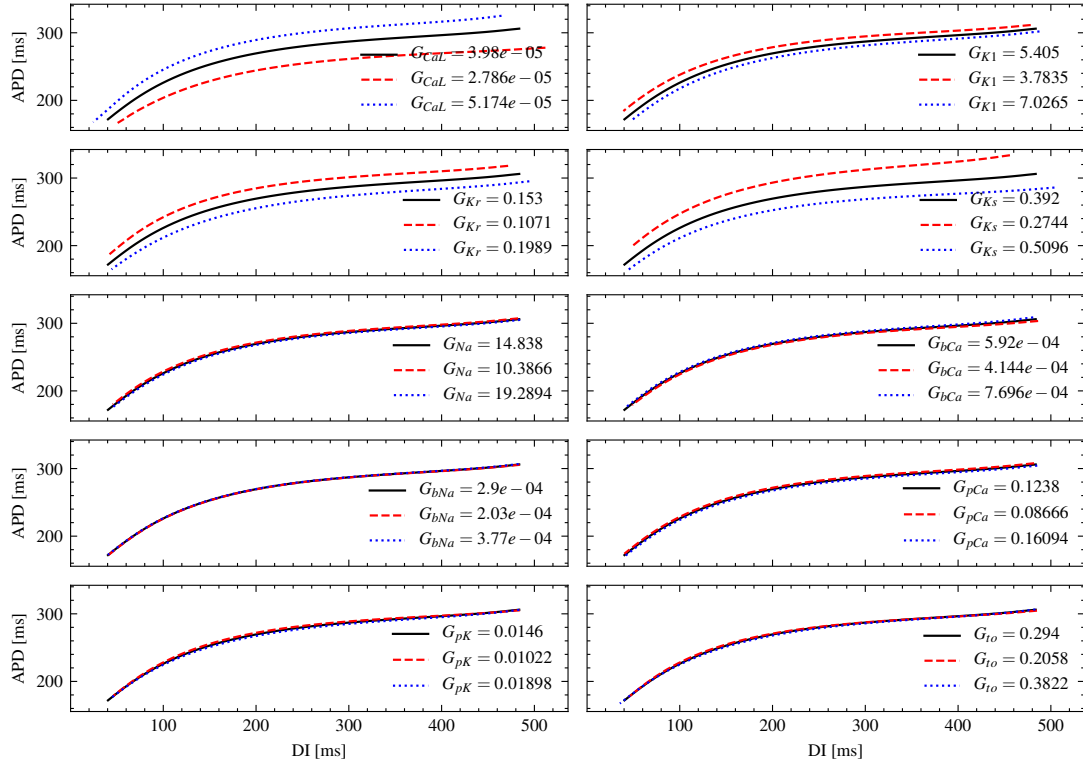


Figura 37 – SA local OAT dos parâmetros do modelo TP para o MaxSlope.

5.3.1 Índices de Sobol

A Figura 38 apresenta o resultado do cálculo dos índices de Sobol para parâmetros do modelo TP, onde $S1$ representa os índices de primeira ordem e ST representa os índices de ordem total para o APD e MaxSlope.

É possível notar de forma quantitativa que os parâmetros G_{CaL} e G_{Ks} são aqueles que mais impactam no APD, enquanto o G_{CaL} é aquele que mais influencia no máximo da inclinação da curva de restituição. É interessante observar também que interações de alta ordem não foram observadas entre os parâmetros para a quantidade APD, enquanto para a quantidade MaxSlope interações de alta ordem significativas entre quase todos os parâmetros foram observadas.

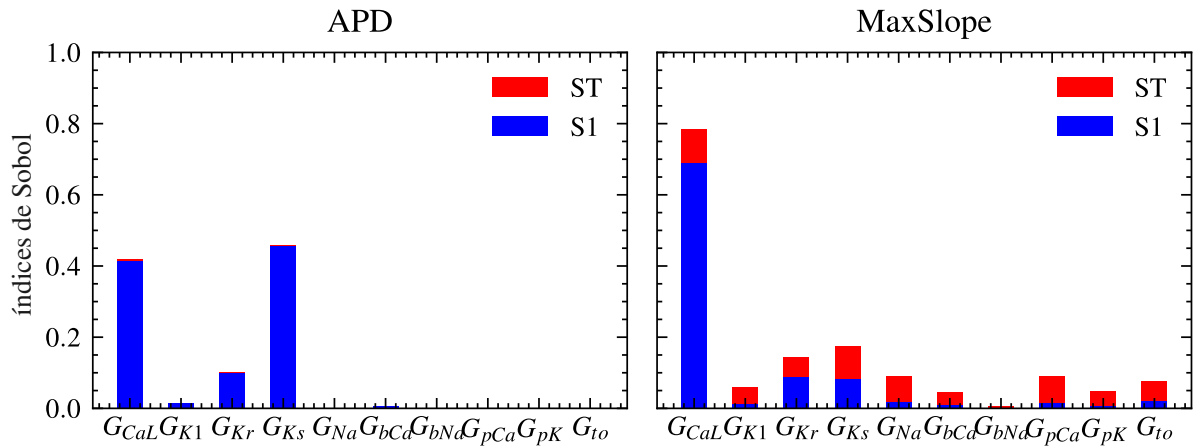


Figura 38 – SA global pelo cálculo dos Índices de Sobol dos parâmetros do modelo TP para as QoI: APD e MaxSlope.

5.3.2 Filtragem de Monte Carlo

A Figura 39 mostra os resultados obtidos pela análise de sensibilidade com a técnica MCF e o cálculo do Teste de Smirnov para cada um dos parâmetros do modelo TP, mostrando, assim, a influência de cada um deles na ocorrência do alternans. Pode-se destacar, portanto, a partir desta análise que os parâmetros G_{CaL} , G_{Kr} e G_{Ks} , de maneira individual, contribuem para o alternans acontecer, sendo G_{CaL} o mais influente destes com uma estatística de Smirnov de 0,73 e os outros dois com valores próximos de 0,15. Os demais parâmetros obtiveram valores muito pequenos e, portanto, conclui-se que estes são pouco determinantes para a ocorrência de alternans no modelo TP.

Também foi feito o cálculo do coeficiente de correlação para pares de parâmetros, para avaliar se a contribuição conjunta destes parâmetros do modelo TP influenciam na ocorrência de alternans. Entretanto, observou-se que não houve correlações significativas entre os parâmetros que contribuíssem para a ocorrência de alternans. Sendo assim, de forma ilustrativa a Figura 40 apresenta o resultado das correlações para apenas dois pares de parâmetros como exemplo. Todos os pares dos parâmetros apresentaram um cc baixo, sempre menores do que 0,02.

5.3.3 Resultados com emuladores PCE

Em seguida, foram construídos emuladores PCE para as quantidades APD e MaxSlope do modelo TP utilizando os parâmetros de condutância e as configurações da Tabela 2.

A Figura 41 apresenta o resultado do histograma do modelo TP original em comparação com o seu emulador PCE em relação aos valores do APD calculados. A resposta obtida pelo emulador PCE do modelo TP é satisfatória, uma vez que o histograma dos valores verdadeiros e dos estimados pelo PCE são próximos. Além disso, nota-se que

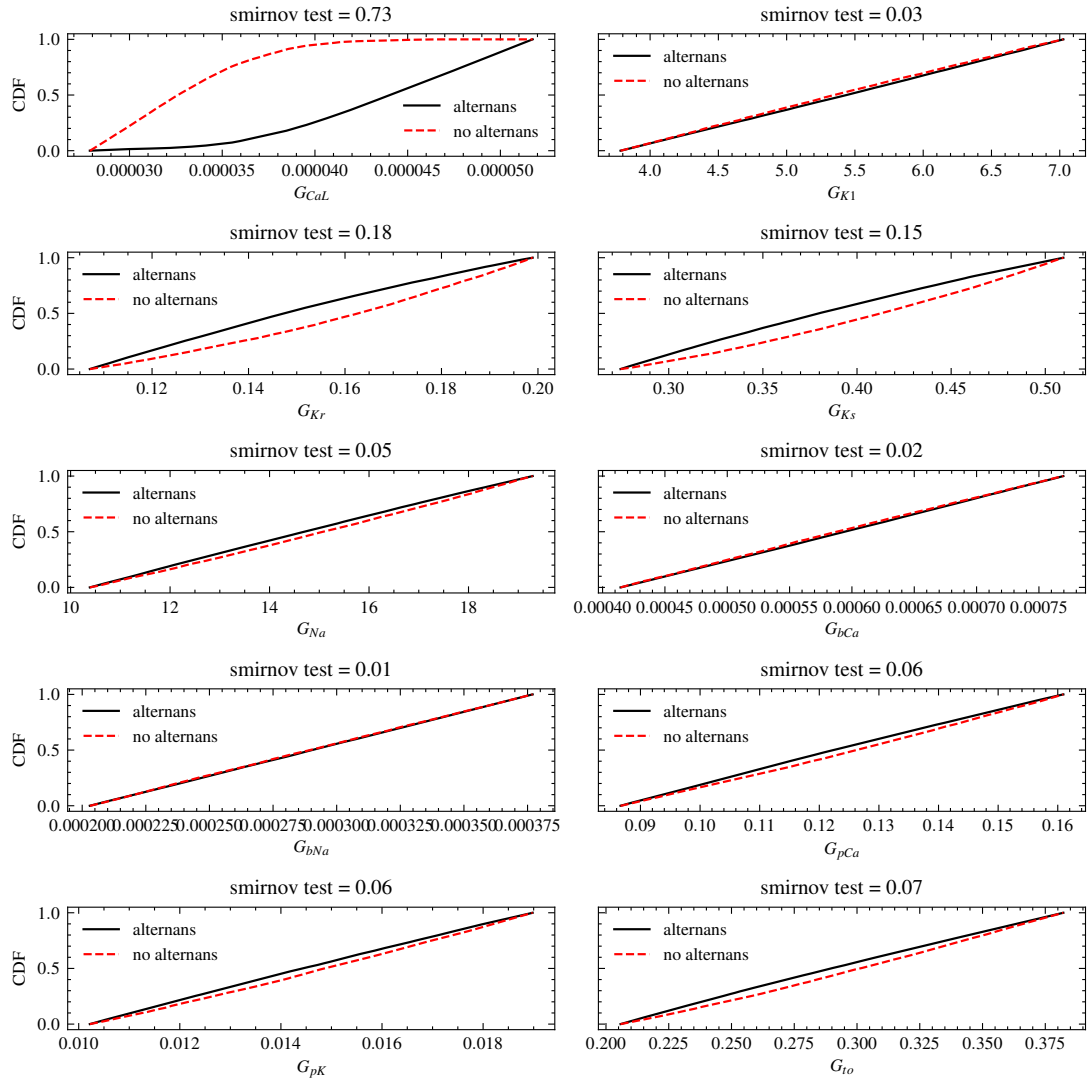


Figura 39 – Resultados do MCF: CDF das amostras que geram alternans (preto) e não geram alternans (vermelho) para os parâmetros do modelo TP, através das quais se faz o cálculo do Teste de Smirnov.

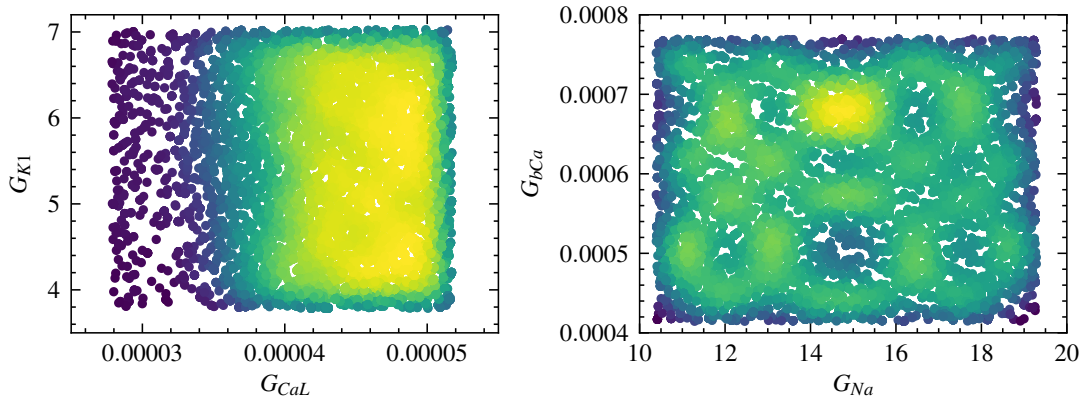


Figura 40 – Correlação entre pares de parâmetros do modelo TP para a ocorrência de *alternans*. A cor indica a densidade local de pontos, das regiões menos densas (roxo) às mais densas (amarelo). Mostram-se apenas dois casos, dado o tamanho do espaço de parâmetros e o fato de que todos apresentaram correlações muito baixas, inferiores a 0,02.

o gráfico de paridade (à direita) deixa claro que o emulador PCE para esta quantidade é bastante preciso, obtendo um coeficiente R^2 de aproximadamente 0,9989, o que tipicamente representa uma boa capacidade de aproximação.

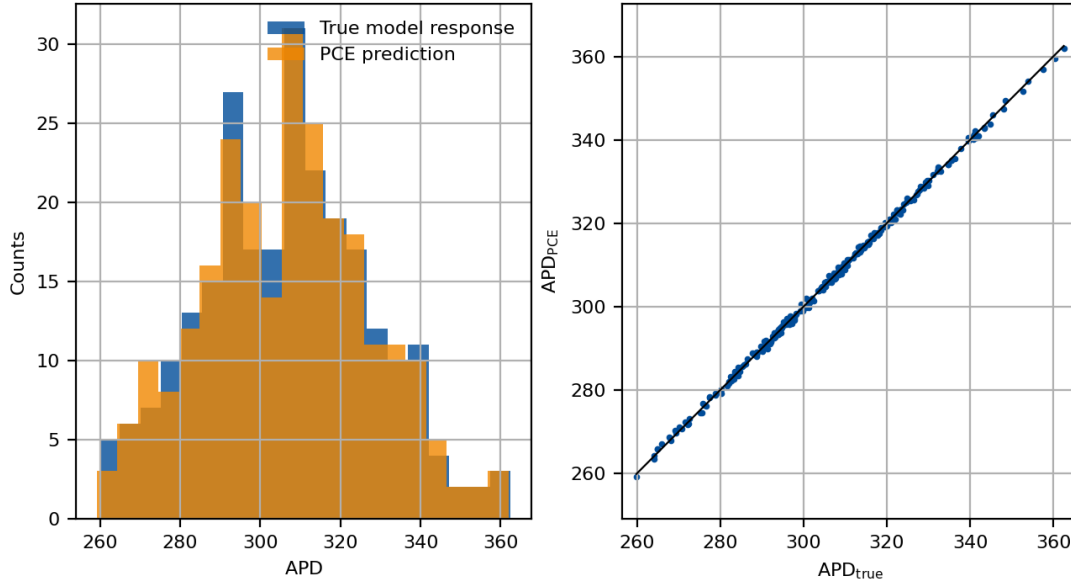


Figura 41 – Histograma (esquerda) e gráfico de paridade (direita) dos valores verdadeiros e os preditos pelo emulador PCE do modelo TP para APD.

Tem-se ainda que a Figura 42 apresenta o resultado do histograma do modelo TP original em comparação com o seu emulador PCE em relação aos valores da máxima inclinação da curva de restituição (MaxSlope). A resposta obtida pelo emulador PCE do modelo TP é boa. No lado esquerdo da Figura 42 pode-se observar os histogramas das respostas verdadeiras e aquelas obtidas pelo PCE. O gráfico de paridade (à direita) mostra uma dispersão maior entre os dados verdadeiros e os estimados pelo PCE, mostrando uma dificuldade maior em emular esta quantidade. Ainda assim, ressalta-se que um coeficiente R^2 de aproximadamente 0,6967 foi obtido.

A Tabela 9 apresenta os resultados do cálculo dos índices de Sobol, de primeira ordem e de ordem total, para os parâmetros do modelo TP emulado pelos PCE de APD e MaxSlope. Pode-se observar que, em geral, os valores dos índices, $S1$ e ST , encontrados para cada parâmetro estão próximos dos valores encontrados pelo cálculo dos índices de Sobol para o modelo TP da Figura 38, que foram obtidos pelo método de Monte Carlo. Pela diferença entre o índice total ST e o índice de primeira ordem $S1$, pode-se notar, como nos resultados obtidos por Monte Carlo, que há interações de alta ordem significativas para a quantidade MaxSlope.

A Tabela 10 apresenta o resultado dos valores da estatística de Smirnov para os parâmetros do modelo TP emulado pelo PCE para MaxSlope. Observa-se aqui também um resultado com valores do Teste de Smirnov para cada parâmetro do emulador bem próximos aos valores do modelo TP obtidos anteriormente pelo método de Monte Carlo.

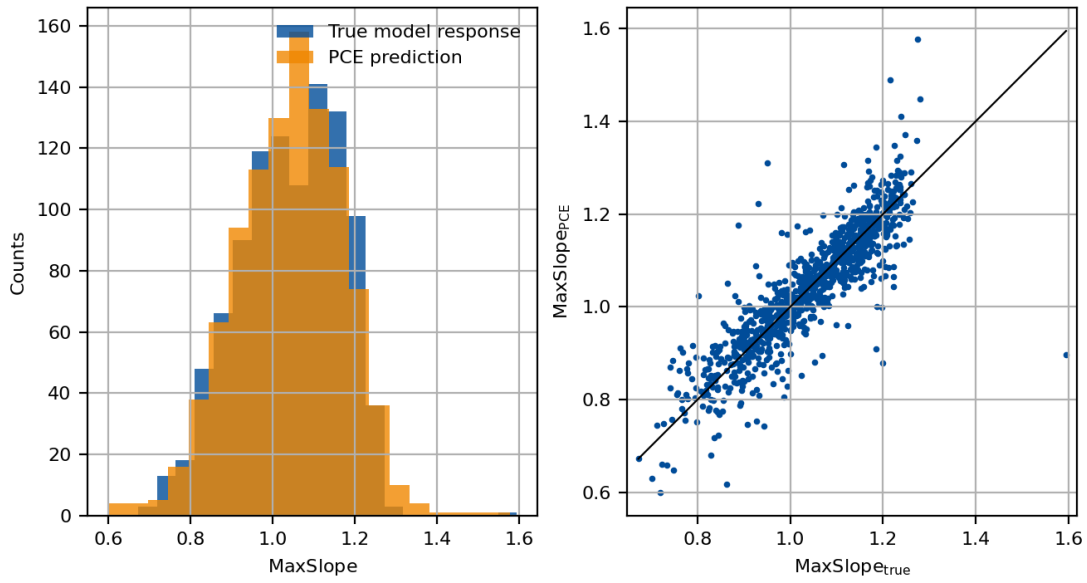


Figura 42 – Histograma (esquerda) e gráfico de paridade (direita) dos valores verdadeiros e os preditos pelo emulador PCE do modelo TP para MaxSlope.

Tabela 9 – Índices de Sobol ($S1$ e ST) para o modelo TP com resultados obtidos por simulação numérica e pelos emuladores para APD e MaxSlope.

Parâmetros	S1 APD	ST APD	S1 MaxSlope	ST MaxSlope
G_{CaL}	0,414	0,418	0,64	0,72
G_{K1}	0,016	0,016	0,00003	0,0337
G_{Kr}	0,101	0,102	0,091	0,145
G_{Ks}	0,453	0,457	0,086	0,120
G_{Na}	0,001	0,001	0,005	0,053
G_{bCa}	0,005	0,005	0,001	0,034
G_{bNa}	0,0002	0,0003	0,00003	0,0309
G_{pCa}	0,001	0,001	0,005	0,053
G_{pK}	0,0003	0,001	0,006	0,047
G_{to}	0,001	0,001	0,010	0,057

Tabela 10 – Estatística de Smirnov calculada pela técnica MCF para os resultados do modelo TP emulado pelo PCE.

Parâmetros	Teste de Smirnov
G_{CaL}	0,66
G_{K1}	0,02
G_{Kr}	0,22
G_{Ks}	0,16
G_{Na}	0,04
G_{bCa}	0,03
G_{bNa}	0,01
G_{pCa}	0,04
G_{pK}	0,04
G_{to}	0,07

Como feito para modelo MMS aqui nesta Seção 5.3 não foram colocados em uma Tabela os valores dos coeficientes de correlação dos parâmetros do modelo TP emulado pelo PCE, pois os valores são baixos, como reproduzidos pelo modelo TP.

5.3.4 Comparativo das Análises do Modelo TP

Por fim, aqui será apresentado um comparativo entre os resultados obtidos nas análises numéricas executadas pelo OpenCARP para o modelo TP e as análises aceleradas pelos emuladores PCE do mesmo.

A Tabela 11 apresenta o comparativo dos resultados alcançados a partir dos cálculos dos Testes de Smirnov, feitos a partir da técnica MCF, para o modelo TP implementado numericamente e por meio do seu emulador PCE, onde, pode-se perceber que os resultados obtidos pelos emuladores PCE são bem próximos daqueles obtidos pelo método Monte Carlo, porém com um custo computacional bem menor, como será discutido adiante.

Tabela 11 – Comparativo entre os Testes de Smirnov calculados, a partir da técnica MCF, para o modelo TP e do seu emulador PCE.

Parâmetros	Modelo TP Teste de Smirnov	PCE Teste de Smirnov
G_{CaL}	0,73	0,66
G_{K1}	0,03	0,02
G_{Kr}	0,18	0,22
G_{Ks}	0,15	0,16
G_{Na}	0,05	0,04
G_{bCa}	0,02	0,03
G_{bNa}	0,01	0,01
G_{pCa}	0,06	0,04
G_{pK}	0,06	0,04
G_{to}	0,07	0,07

A Tabela 12 mostra o comparativo dos resultados alcançados neste trabalho para o modelo TP por meio do método Monte Carlo e pelos emuladores PCE. Novamente, pode-se perceber uma boa concordância qualitativa e quantitativa entre os índices de Sobol, tanto os de primeira ordem, quanto os de ordem total. O maior erro relativo observado entre os índices de Sobol obtidos pelas diferentes técnicas foi da ordem de 1.43×10^{-1} , que, no caso, foi o índice de Sobol de ordem total calculado para a condutância G_{K1} para o caso do APD.

5.4 Desempenho computacional

Nesta seção, a partir dos resultados das análises de sensibilidade que foram apresentados e comparados nas Subseções 5.2.5 e 5.3.4 para os modelos MMS e TP, respecti-

Tabela 12 – Comparativo entre o cálculo dos Índices de Sobol do modelo TP e do seu emulador PCE para o APD e o MaxSlope.

Parâmetros	APD		APD PCE		MaxSlope		MaxSlope PCE	
	S1	ST	S1	ST	S1	ST	S1	ST
G_{CaL}	0,412	0,418	0,414	0,418	0,69	0,78	0,64	0,72
G_{K1}	0,154	0,158	0,016	0,016	0,01	0,06	0,00003	0,0337
G_{Kr}	0,01	0,102	0,101	0,102	0,08	0,14	0,091	0,145
G_{Ks}	0,455	0,458	0,453	0,457	0,08	0,17	0,086	0,120
G_{Na}	0,001	0,001	0,001	0,001	0,01	0,009	0,005	0,053
G_{bCa}	0,005	0,005	0,005	0,005	0,009	0,04	0,001	0,034
G_{bNa}	0,0003	0,0003	0,0002	0,0003	0,0009	0,005	0,00003	0,0309
G_{pCa}	0,001	0,001	0,001	0,001	0,01	0,009	0,005	0,053
G_{pK}	0,001	0,001	0,0003	0,001	0,007	0,04	0,006	0,047
G_{to}	0,001	0,001	0,001	0,001	0,02	0,07	0,010	0,057

vamente, é realizada também uma comparação de desempenho computacional. Para isso, analisou-se o tempo de execução das análises e da geração dos dados.

O emulador PCE acelerou bastante o tempo de execução das análises, tanto para o modelo MMS, quanto para o modelo TP. O emulador do modelo MMS para o APD ficou em torno de 12 vezes mais rápido que o modelo MMS original, enquanto que, para o MaxSlope, ficou em torno de 10 vezes mais rápido. Já no modelo TP, para o APD, o emulador produziu resultados em torno de 18 vezes mais rápido, enquanto que, para o MaxSlope, resultados foram obtidos em torno de 22 vezes mais rápido.

A Tabela 13 apresenta o tempo computacional (em horas e minutos) para as análises de cada modelo por meio do método Monte Carlo ou dos emuladores PCE. Para complementar a análise, são apresentadas informações como: número de parâmetros, número de amostras (ou simulações do modelo celular realizadas) para o método Monte Carlo, denotado por N_s , assim como o número de amostras usadas para criar o PCE, N_{PCE} , e o grau do mesmo. É importante ressaltar que o tempo reportado para o PCE contabiliza todo o procedimento, incluindo a geração das $N_{PCE} = 1000$ amostras de treinamento a partir do modelo celular original, o ajuste do emulador e a execução da análise de sensibilidade sobre ele, de modo que os tempos são diretamente comparáveis aos do método Monte Carlo.

Para o modelo MMS, o tempo de execução de 10240 amostras para o APD é de aproximadamente 6 horas, e para o MaxSlope é de aproximadamente 91 horas, enquanto que para o modelo TP o tempo de execução de 22528 amostras foi de aproximadamente 18 horas para o APD e aproximadamente 169 horas para o MaxSlope.

Tabela 13 – Comparativo entre os Desempenhos Computacionais dos modelos MMS e TP e seus respectivos emuladores PCE.

Modelo	N_p	Monte Carlo		PCE		
		N_s	Tempo	N_{PCE}	Grau	Tempo
MMS	4	10240 (APD)	5h41m	1000 (APD)	4 (APD)	34m30s
		10240 (MaxSlope)	34h1m	1000 (MaxSlope)	7 (MaxSlope)	3h21m
TP	10	22528 (APD)	17h57m	1000 (APD)	4 (APD)	48m36s
		22528 (MaxSlope)	168h57m	1000 (MaxSlope)	7 (MaxSlope)	7h30m

5.5 Discussões

Este trabalho utilizou de diferentes técnicas de análise de sensibilidade aplicadas em dois modelos celulares da eletrofisiologia cardíaca que descrevem a geração do potencial de ação e podem representar alternans de APD. Foram escolhidos dois modelos com características diferentes: um modelo qualitativo e outro quantitativo, sendo o primeiro um modelo simplificado com apenas 4 variáveis (modelo MMS), e o outro um modelo biofísico detalhado com 19 variáveis (modelo TP). A seguir, algumas discussões a respeito dos resultados preliminares deste trabalho são apresentadas.

5.5.1 Verificação dos experimentos para o modelo MMS

Na seção 5.2.5 foi apresentada uma verificação das análises feitas para o modelo MMS utilizando o OpenCARP com a fórmula analítica da curva de restituição deste modelo. Optou-se por essa abordagem para posteriormente se utilizar das expressões analíticas como forma de verificação da proposta deste trabalho. Ressalta-se que para modelos mais complexos, fórmulas que descrevem a curva de restituição do modelo não estão disponíveis e, portanto, somente através das simulações computacionais o modelo poderia ser analisado.

5.5.2 Comparação com a literatura

Nesta seção, os resultados das análises de sensibilidade apresentados na Seção 5.3.4 para o modelo TP são comparados com aqueles de estudos prévios disponíveis na literatura, os quais utilizaram diferentes metodologias e abordaram outros objetivos. O intuito dessa comparação é verificar a coerência e validade dos resultados obtidos no presente trabalho.

O trabalho de Corso e Viola (2020) apresenta diversas análises e utiliza emuladores baseados em processos gaussianos. Em particular, um dos estudos apresentados em seu artigo analisa a sensibilidade do APD através dos índices de Sobol com relação a diversos parâmetros de entrada, inclusive as 10 condutâncias do modelo TP. A Tabela 14 apresenta os índices de Sobol principal e total para os parâmetros incluídos na análise do trabalho de Corso e Viola (2020) e as obtidas no presente trabalho.

Tabela 14 – Comparativo entre o cálculo dos Índices de Sobol do modelo TP neste trabalho e do modelo TP abordado no trabalho de Corso e Viola (2020), para o APD.

Parâmetros	Presente Trabalho		Corso e Viola (2020)	
	S1	ST	S1	ST
G_{CaL}	0,414	0,418	0,301	0,305
G_{K1}	0,016	0,016	0,020	0,023
G_{Kr}	0,101	0,102	0,090	0,095
G_{Ks}	0,453	0,457	0,250	0,259
G_{Na}	0,001	0,001	0,001	0,001
G_{bCa}	0,005	0,005	0,003	0,003
G_{bNa}	0,0002	0,0003	0,001	0,001
G_{pCa}	0,001	0,001	0,001	0,001
G_{pK}	0,0003	0,001	0,011	0,014
G_{to}	0,001	0,001	0,001	0,001
Ca_O	–	–	0,191	0,195
Na_O	–	–	0,120	0,124
K_O	–	–	0,025	0,028

Em primeiro lugar, é importante observar que os resultados apresentam boa concordância para os parâmetros das condutâncias. A análise do presente trabalho não incluiu os parâmetros Ca_O , K_O e Na_O como feito no trabalho de Corso e Viola (2020). Logo, como, na análise feita do modelo TP pelo trabalho de Corso e Viola (2020), as 10 condutâncias tiveram seus índices de Sobol calculados para o APD juntamente com estes outros 3 parâmetros, ocorreu uma redistribuição no valor absoluto das influências das condutâncias no APD para os outros 3 parâmetros analisados, explicando assim as maiores diferenças observadas em alguns parâmetros.

Entretanto, vale ressaltar que o trabalho de Corso e Viola (2020) não realiza um estudo de análise de sensibilidade que avalia o impacto de um parâmetro na ocorrência (ou não) de um determinado comportamento, como feito no presente trabalho para o alternans de APD através da técnica de filtragem de Monte Carlo.

5.5.3 Sobre os modelos e seus parâmetros

Sobre os resultados para o modelo MMS nota-se claramente a relevância do parâmetro τ_{close} individualmente para a ocorrência de alternans, enquanto os parâmetros τ_{out} e τ_{in} , também contribuem para a ocorrência de alternans, porém de forma moderada, como pode ser comprovado pelos índices de Sobol para MaxSlope e pela estatística de Smirnov para o modelo MMS. Esses resultados foram obtidos tanto via análise Monte Carlo, quanto através dos emuladores PCE, como esperado, entretanto a um custo computacional muito inferior pelos emuladores PCE, como apresentado na Tabela 13.

Sobre as análises para o modelo TP é importante notar que os resultados alcançados confirmam a grande influência do parâmetro G_{CaL} individualmente para a ocorrência

do alternans, prevista em parte pela análise OAT, pelo cálculo dos índices de Sobol e pela estatística de Smirnov. Observou-se também o papel dos parâmetros G_{Ks} e G_{Kr} na ocorrência de alternans pela estatística de Smirnov e também nos índices de Sobol para a máxima inclinação da curva de restituição. Os demais parâmetros não apresentaram relevância significativa para o comportamento de alternans. Os resultados obtidos pelo método MC e PCE estão consistentes, como mostrou a Tabela 12. E ainda, a diferença no desempenho computacional das análises pelo método MC comparada com os emuladores PCE também é significativa para o modelo TP, como também evidenciado pela Tabela 13.

No modelo TP, é relevante destacar que a proposta apresentada por (TUSSCHER; PANFILOV, 2006) constitui uma evolução do modelo anterior dos mesmos autores (TUSSCHER et al., 2004). A nova versão do modelo de célula ventricular humana apoia-se em recentes medições experimentais de restituição do APD humano e oferece uma descrição mais abrangente da dinâmica intracelular de cálcio. Além disso, foram realizados estudos detalhados sobre a curva de restituição do APD e a instabilidade associada à geração de alternans, algo não explorado no modelo anterior. Os resultados do presente trabalho confirmam que o parâmetro G_{CaL} , seguido por G_{Kr} e G_{Ks} são fundamentais na ocorrência de alternans, em conformidade com as modificações e o que foi apresentado em (TUSSCHER; PANFILOV, 2006).

Sobre o resultado dos índices de Sobol calculado sobre cada parâmetro do modelo TP e seu comparativo com o PCE mostrado na Tabela 12, é correto observar que este resultado obtido reitera a importância destacada para o parâmetro G_{CaL} na influência das QoIs, e, por consequência, no alternans, mas diferente do analisado na estatística de Smirnov, essa análise de Sobol coloca o parâmetro G_{Ks} em igual importância e influência nas QoIs.

5.5.4 Comparativo das técnicas de análise de sensibilidade

Por fim, apresenta-se nesta seção uma breve discussão sobre as técnicas de análise de sensibilidade utilizadas neste trabalho em vista dos resultados obtidos. Foram utilizadas as seguintes técnicas de análise de sensibilidade: OAT, índices de Sobol e a Filtragem Monte Carlo.

Inicialmente é importante destacar que a técnica OAT é uma técnica de SA local onde se avalia o impacto de pequenas variações em um único parâmetro de entrada por vez, mantendo os demais constantes. Esta técnica se concentra em perturbações locais no espaço dos parâmetros, sem explorar o comportamento global do modelo. Foi utilizada como um passo inicial devido à sua simplicidade e baixo custo computacional, mas deve-se notar que ela não é capaz de capturar possíveis interações entre os parâmetros do modelo.

Para o modelo MMS, os resultados da técnica OAT indicam que variações em τ_{in} , τ_{out} e τ_{close} causam mudanças significativas na quantidade MaxSlope (que é determinante para

entender o comportamento de alternans). Entretanto, é difícil distinguir quantitativamente qual desses parâmetros é o que mais influencia na ocorrência de alternans apenas pelos resultados da Figura 27. A mesma situação pode ser observada para o modelo TP, onde a análise OAT indica visualmente que G_{CaL} , G_{K1} , G_{kr} e G_{Ks} possuem uma influência significativa em MaxSlope (veja a Figura 37).

Por outro lado, a análise de sensibilidade através dos índices de Sobol é caracterizada como uma técnica global que avalia a contribuição de cada parâmetro de entrada, e suas interações, na variabilidade da saída de um modelo. Os índices de Sobol consideram o espaço dos parâmetros completo e a variação conjunta dos parâmetros, permitindo uma visão mais abrangente do comportamento do modelo. A SA por Sobol é baseada na decomposição da variância da saída de um modelo em contribuições atribuíveis a cada parâmetro de entrada, bem como às interações entre eles. Sendo assim, essa técnica fornece índices numéricos precisos que quantificam a influência de cada variável de forma mais clara.

Os resultados dos modelos MMS e TP permitiram o ranqueamento de forma quantitativa dos parâmetros de acordo com seus índices de Sobol. No modelo MMS, o parâmetro τ_{close} apresentou o maior impacto em APD e MaxSlope, enquanto no modelo TP, G_{CaL} foi o mais influente para MaxSlope, e tanto G_{CaL} quanto G_{Ks} se destacaram como os mais relevantes para APD. Além disso, foi possível observar através dos índices de Sobol de alta ordem nas Figuras 28 e 38 que ocorrem interações significativas entre os parâmetros afetando a variabilidade da máxima inclinação da curva de restituição (MaxSlope). Esse comportamento não é possível de se constatar pela análise OAT.

Embora as quantidades APD e MaxSlope utilizadas nas análises anteriores sejam relevantes e intimamente relacionadas ao comportamento de alternans do APD, as análises de sensibilidade conduzidas por meio de OAT e índices de Sobol apenas avaliaram essas métricas sem considerar se, para um determinado conjunto de parâmetros, o alternans de fato ocorreu. Nesse contexto, a Filtragem de Monte Carlo é introduzida como uma abordagem complementar: após realizar simulações do modelo (ou do emulador PCE) usando o método de Monte Carlo, ela classifica os resultados entre aqueles com ou sem ocorrência de alternans. Com base nessa distinção, a técnica permite uma análise mais aprofundada, focada em discriminar o papel dos parâmetros do modelo na determinação desse comportamento específico.

Nos resultados deste trabalho para o modelo MMS pelo método MCF usando a estatística de Smirnov observou-se que, de fato, o parâmetro τ_{close} é o principal determinante para alternans de APD, seguido por τ_{in} e τ_{out} . Além disso, a análise revelou correlações significativas entre esses parâmetros para a ocorrência de alternans. Para o modelo TP, os resultados da análise pelo MCF revelaram que G_{CaL} é o principal determinante do comportamento de alternans do APD, dentre os dez parâmetros de condutância analisados.

Outros dois parâmetros G_{Ks} e G_{kr} também indicam alguma relevância para ocorrência de alternans, enquanto os demais são quase insignificantes.

6 ANÁLISE DE SENSIBILIDADE DO MODELO MAHAJAN-SHIFERAW

Enquanto o capítulo anterior avaliou a viabilidade da abordagem proposta empregando emuladores baseados em PCE nos modelos MMS e TP, este capítulo estende a metodologia em duas direções. A primeira é a aplicação ao modelo Mahajan-Shiferaw, biofisicamente mais detalhado e com maior número de parâmetros. A segunda é a ampliação do conjunto de modelos substitutos: além do PCE, são empregados os Processos Gaussianos e as Florestas Aleatórias, tanto na regressão das quantidades de interesse quanto na classificação direta da ocorrência de alternans. Dessa forma, busca-se avaliar como as diferentes técnicas se comportam em um cenário mais realista e desafiador, tanto em termos de qualidade de aproximação quanto de capacidade de discriminação entre os regimes com e sem alternans.

As análises são organizadas em três experimentos de complexidade crescente, nos quais o número de parâmetros de entrada do modelo é progressivamente ampliado, permitindo avaliar a robustez das técnicas à medida que a dimensionalidade do problema aumenta. Para cada experimento, compara-se o desempenho dos emuladores na aproximação do APD e do MaxSlope, os índices de sensibilidade obtidos e a capacidade de detecção de alternans. Por fim, apresenta-se um comparativo direto entre as abordagens de regressão e de classificação, avaliando as vantagens e limitações de cada estratégia.

6.1 Descrição dos experimentos

Para a análise de sensibilidade deste modelo, foram realizados três experimentos com conjuntos de parâmetros de entrada de complexidade crescente, conforme descrito a seguir. Os detalhes dos procedimentos computacionais seguem a abordagem descrita e adotada no capítulo anterior.

No Experimento I, foram analisadas as condutâncias máximas dos canais iônicos do modelo, totalizando 8 parâmetros: G_{CaL} , G_{K1} , G_{Kr} , G_{Ks} , G_{Na} , G_{NaCa} , G_{tof} e G_{tos} . Esta escolha segue a mesma abordagem adotada para o modelo TP no Capítulo 5, permitindo alguma comparação entre os dois modelos.

No Experimento II, o espaço de parâmetros foi ampliado para 11 variáveis, incluindo, além das condutâncias máximas do Experimento I, as concentrações extracelulares de cálcio, potássio e sódio: Ca_o , K_o e Na_o . A inclusão dessas concentrações visa investigar a influência das condições do meio extracelular na ocorrência de alternans e segue o mesmo conjunto de parâmetros explorados por (CORSO; VIOLA, 2020).

Por fim, no Experimento III, foram considerados 13 parâmetros, adicionando ao conjunto do Experimento II as concentrações intracelulares de potássio e sódio, K_i e Na_i , de forma a avaliar também o papel das concentrações intracelulares na dinâmica do

modelo.

6.1.1 Experimento I - 8 parâmetros

Emuladores PCE e GP foram construídos para as quantidades de interesse APD e MaxSlope, e um classificador RF foi treinado diretamente para identificar a ocorrência ou não de alternans. Em todos os casos, utilizaram-se quatro tamanhos de conjunto de treinamento: 500, 1000, 2000 e 3000 amostras, sendo 80% destinadas ao treinamento e os 20% restantes reservados para validação. Essa variação no tamanho do conjunto de treinamento permite avaliar a convergência dos índices de sensibilidade, as métricas de qualidade e o desempenho computacional dos modelos em função do número de amostras disponíveis.

6.1.1.1 Análise para APD

A Tabela 15 a seguir mostra os índices de Sobol de primeira ordem do APD para os diferentes casos, tempo de construção dos emuladores, tempo gasto para avaliação das amostras de validação e as métricas LOO e R^2 . Pode-se observar que, à medida que aumenta o número de amostras utilizadas no treinamento dos emuladores, os índices de Sobol de primeira ordem associados ao APD tendem a convergir para valores estáveis. No entanto, os resultados indicam que, mesmo com 500 amostras de treinamento, as estimativas já se encontram próximas de seus valores assintóticos, sugerindo que a estrutura principal de sensibilidade do modelo é capturada adequadamente mesmo com um conjunto de treinamento relativamente pequeno. Nota-se também que para números de amostras inferiores a 3000 o emulador GP obteve alguns índices de Sobol negativos, porém de magnitude muito pequena e que podem ser considerados desprezíveis, sendo provavelmente decorrentes de erros numéricos ou de estimação, sem impacto relevante na interpretação da importância relativa dos parâmetros.

Com relação ao desempenho computacional, observa-se que o tempo de construção dos emuladores baseados em PCE permanece praticamente constante à medida que aumenta o número de amostras de treinamento. Em contraste, o tempo de construção dos emuladores baseados em GP cresce significativamente com o tamanho do conjunto de treinamento, refletindo o maior custo computacional associado ao ajuste do modelo gaussiano. Por outro lado, o tempo de inferência de ambos os emuladores permanece praticamente constante em todos os casos analisados. Ainda assim, o tempo de avaliação do emulador PCE mostra-se ligeiramente menor que o do emulador GP.

Em termos de precisão, avaliada em termos de LOO e R^2 , ambos os emuladores apresentaram excelente desempenho na aproximação do modelo original, com coeficientes de determinação elevados em todos os casos analisados. O emulador GP apresentou valores de R^2 ligeiramente superiores e erros LOO menores, indicando maior precisão na

Tabela 15 – Experimento I: Índices de Sobol de Primeira Ordem dos parâmetros do modelo Mahajan, custo computacional e métricas de precisão dos emuladores PCE e GP para a quantidade APD, em função do número de amostras de treinamento.

Parâmetro	PCE				GP			
	500	1000	2000	3000	500	1000	2000	3000
G_{CaL}	0,493	0,489	0,493	0,496	0,490	0,496	0,494	0,489
G_{K1}	0,012	0,013	0,015	0,014	0,011	0,029	0,009	0,022
G_{Kr}	0,019	0,017	0,016	0,017	0,012	0,031	0,012	0,023
G_{Ks}	0,000	0,000	0,000	0,000	-0,004	0,015	-0,005	0,007
G_{Na}	0,000	0,000	0,000	0,000	-0,003	0,015	-0,005	0,008
G_{NaCa}	0,055	0,054	0,056	0,055	0,044	0,066	0,044	0,058
G_{tof}	0,315	0,328	0,328	0,325	0,309	0,323	0,310	0,309
G_{tos}	0,032	0,029	0,028	0,028	0,023	0,044	0,023	0,040
Construção (s)	4,09	4,40	3,31	3,46	17,83	35,96	166,36	405,26
Avaliação (s)	0,93	1,09	1,08	1,10	0,92	1,08	1,08	2,66
LOO	0,029	0,025	0,030	0,028	0,003	0,002	0,001	0,001
R^2	0,971	0,970	0,973	0,969	0,996	0,998	0,999	0,999

reconstrução das respostas do modelo. Entretanto, os índices de Sobol estimados por ambos os métodos mostraram-se muito próximos, evidenciando que tanto PCE quanto GP são capazes de capturar adequadamente a estrutura de sensibilidade do modelo.

A Figura 43 apresenta os resultados obtidos pelos emuladores PCE e GP nos painéis (a) e (b), respectivamente, considerando $N = 2000$ amostras no conjunto de validação. Observa-se que o emulador baseado em GP reproduz com elevada fidelidade o comportamento do modelo original. No histograma, as distribuições dos valores verdadeiros e daqueles preditos pelo GP praticamente se sobrepõem, indicando que o emulador preserva adequadamente a média, a dispersão e a forma global da distribuição do APD. No gráfico de paridade, apresentado no painel à direita, os pontos se concentram fortemente ao longo da reta identidade no caso do GP, evidenciando a alta concordância entre os valores preditos e os valores de referência. Em contraste, o emulador PCE apresenta uma leve tendência de superestimação dos valores de APD, o que pode ser observado tanto pelo deslocamento sutil no histograma quanto pela maior dispersão dos pontos em torno da reta identidade no gráfico de paridade.

6.1.1.2 Análise para MaxSlope com PCE e GP

Devido à complexidade da simulação da curva de restituição, foi adotado o seguinte pós-processamento para esta quantidade de interesse. Primeiro, foram descartados os casos em que houve falha no cálculo da curva de restituição. Em seguida, foram descartados os casos com inclinação máxima negativa: como essa quantidade corresponde à maior inclinação ao longo de toda a curva, um valor negativo indicaria uma curva decrescente

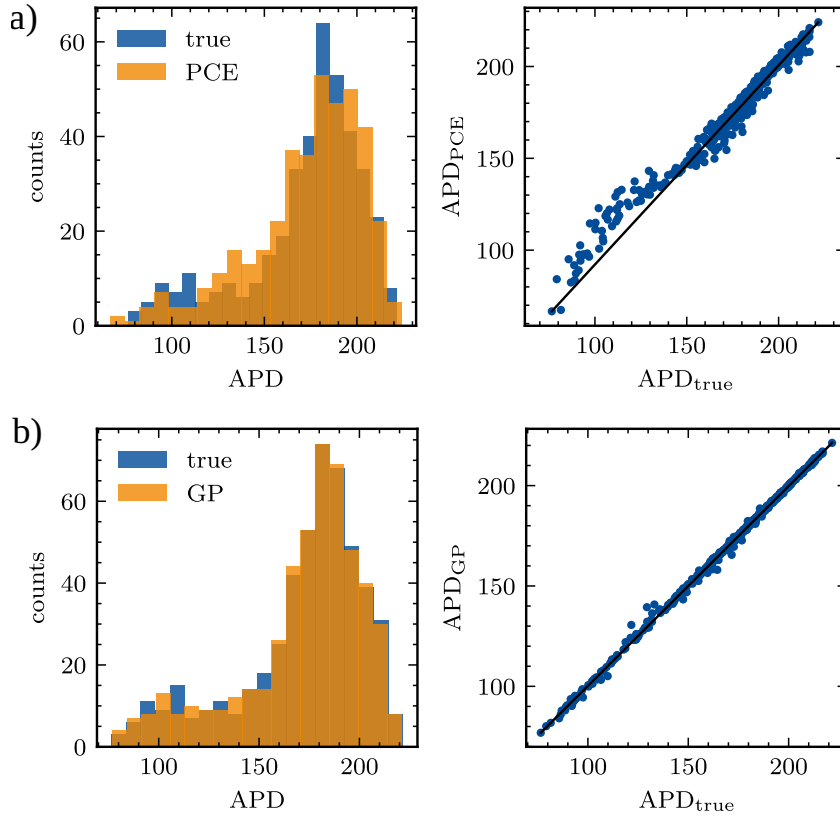


Figura 43 – Experimento I: Comparação dos emuladores PCE (a) e GP (b) no conjunto de dados de validação para a quantidade APD: histograma dos valores verdadeiros e estimados pelo emulador (esquerda) e gráfico de paridade (direita).

em toda a faixa de DI avaliada, comportamento incompatível com a restituição esperada e atribuído a problemas na simulação. Por fim, foram descartados os valores superiores a 3, considerados artefatos numéricos associados a variações abruptas da curva em DIs muito curtos, próximos ao limite de refratariedade. Dessa forma, apenas inclinações máximas no intervalo $(0, 3)$ foram consideradas nas análises. Sendo assim, dados os conjuntos com 500, 1000, 2000 e 3000 amostras, após as simulações e o pós-processamento obteve-se um total efetivo de 401, 813, 1661 e 2471 amostras a partir das quais foram definidos os conjuntos de dados de treinamento e validação seguindo 80% e 20%, respectivamente. O histograma dos dados processados é apresentado na Figura 44.

É importante destacar a complexidade no manejo dos dados gerados por essas simulações, que atingem aproximadamente 98 GB e 134 GB para as quantidades APD e MaxSlope, respectivamente. Tais magnitudes indicam que o aumento do tamanho amostral não é trivial, configurando-se como um fator limitante.

A Tabela 16 apresenta os índices de Sobol de primeira ordem (principal) e totais, assim como as métricas de desempenho (tempo de construção e avaliação) e de erro (LOO e R^2) para os quatro tamanhos de amostras para os emuladores PCE e GP. Pode-se observar, como anteriormente, que os emuladores PCE e GP apresentam boa concordância nos

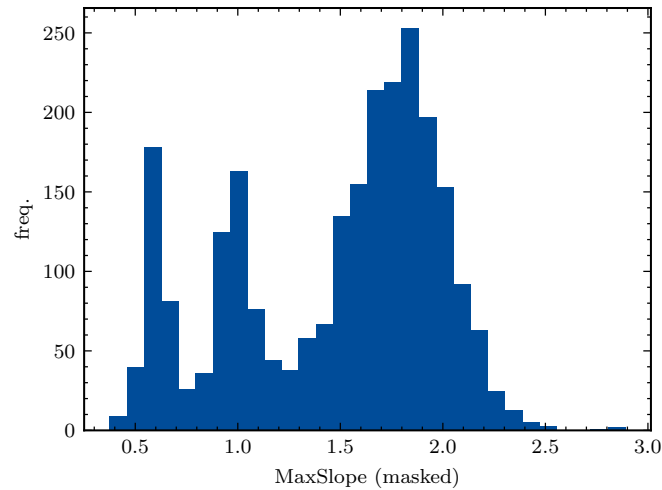


Figura 44 – Experimento I: Histograma das amostras de MaxSlope após pós-processamento, totalizando 2471 amostras válidas.

valores dos índices de Sobol (principal e total) e que convergem à medida que o número de amostras aumenta. Os emuladores GP obtiveram melhores métricas de precisão em comparação aos baseados em PCE, apesar de terem resultados em índices ligeiramente negativos (mas de magnitude desprezível). Em termos computacionais, observou-se, novamente, um maior custo na construção dos emuladores GP.

A Figura 45 apresenta a comparação entre os emuladores PCE e GP no conjunto de validação. Em ambos os casos nota-se uma tendência em superestimar o valor de MaxSlope, apesar dos emuladores terem alcançado coeficientes R^2 moderados. Além disso, observa-se que o emulador baseado em GP reproduz com maior fidelidade o comportamento dos dados, capturando de forma mais acurada a resposta de MaxSlope. Por outro lado, o modelo PCE apresenta limitações importantes, incluindo a predição de valores negativos e a incapacidade de representar adequadamente a tendência dos dados, apesar de exibir métricas globais de precisão razoáveis.

Em seguida, realizou-se a análise MCF utilizando os emuladores. Para esse fim, foram geradas 10000 amostras por meio de Latin Hypercube Sampling (LHS), as quais foram avaliadas pelos emuladores de MaxSlope (PCE e GP). As respostas obtidas, denotadas por Y , foram então classificadas em regimes de alternans e não alternans com base no critério $Y > 1$. Os resultados em termos do teste de Smirnov para os emuladores criados com diferentes tamanhos amostrais são apresentados na Tabela 17. Observa-se que os valores da estatística de Smirnov apresentam uma tendência de estabilização à medida que o tamanho amostral aumenta, indicando convergência dos resultados. Pode-se notar que há boa concordância entre os emuladores PCE e GP, uma vez que ambos são capazes de identificar, por meio da estatística de Smirnov no contexto do MCF, os parâmetros mais relevantes para a ocorrência de alternans. A ordenação dos parâmetros mais relevantes é, em grande parte, preservada entre os métodos, com apenas pequenas variações relativas

Tabela 16 – Experimento I: Índices de Sobol de Primeira Ordem e Totais dos parâmetros do modelo Mahajan, custo computacional e métricas de precisão dos emuladores PCE e GP para a quantidade MaxSlope, em função do número de amostras de treinamento.

Parâmetro	PCE				GP			
	500	1000	2000	3000	500	1000	2000	3000
<i>Índices de Sobol Total</i>								
G_{CaL}	0,772	0,702	0,741	0,739	0,749	0,733	0,768	0,797
G_{K1}	0,084	0,099	0,099	0,101	0,113	0,130	0,121	0,134
G_{Kr}	0,034	0,065	0,044	0,054	0,062	0,114	0,058	0,085
G_{Ks}	0,003	0,000	0,003	0,001	0,002	0,002	0,001	0,002
G_{Na}	0,041	0,053	0,039	0,039	0,078	0,110	0,073	0,043
G_{NaCa}	0,046	0,090	0,068	0,068	0,153	0,081	0,084	0,071
G_{tof}	0,353	0,336	0,338	0,326	0,359	0,399	0,415	0,402
G_{tos}	0,017	0,016	0,033	0,042	0,023	0,090	0,064	0,085
<i>Índices de Sobol de Primeira Ordem</i>								
G_{CaL}	0,534	0,476	0,491	0,502	0,445	0,408	0,396	0,442
G_{K1}	0,026	0,024	0,022	0,021	0,044	0,033	0,021	0,030
G_{Kr}	0,000	0,004	0,000	0,002	0,014	0,014	-0,005	-0,001
G_{Ks}	0,000	0,000	0,000	0,000	0,013	0,011	-0,006	0,007
G_{Na}	0,009	0,033	0,018	0,016	0,018	0,036	0,015	0,018
G_{NaCa}	0,000	0,022	0,014	0,012	0,030	0,013	0,010	0,020
G_{tof}	0,080	0,079	0,090	0,075	0,079	0,061	0,076	0,079
G_{tos}	0,000	0,000	0,000	0,001	0,015	0,015	-0,002	0,008
Construção (s)	4,04	2,79	3,16	3,45	15,10	26,19	103,60	236,27
Avaliação (s)	0,93	0,93	1,09	1,08	0,96	0,96	2,57	1,11
LOO	0,361	0,331	0,339	0,329	0,266	0,245	0,249	0,241
R^2	0,581	0,628	0,707	0,674	0,686	0,779	0,815	0,766

nas posições intermediárias do ranking.

A Figura 46 apresenta as funções de distribuição acumulada após a filtragem do comportamento (alternans e não alternans) para os parâmetros G_{CaL} e G_{K1} com maiores valores da estatística de Smirnov, obtidos com o emulador GP. Resultados análogos foram observados para o emulador PCE, motivo pelo qual não são apresentados. Observa-se que G_{CaL} é, de fato, o parâmetro mais determinante para a ocorrência de alternans. Em particular, para valores mais baixos (inferiores a 150), sua influência está predominantemente associada à não ocorrência de alternans. Os resultados para os demais parâmetros não são exibidos, uma vez que suas CDFs apresentam separação menos pronunciada e comportamento qualitativamente semelhante ao observado para G_{K1} .

De maneira geral, observou-se que ambos os emuladores, PCE e GP, foram capazes de fornecer resultados consistentes, tanto em termos de métricas de precisão quanto na identificação dos parâmetros mais relevantes para a ocorrência de alternans. No entanto,

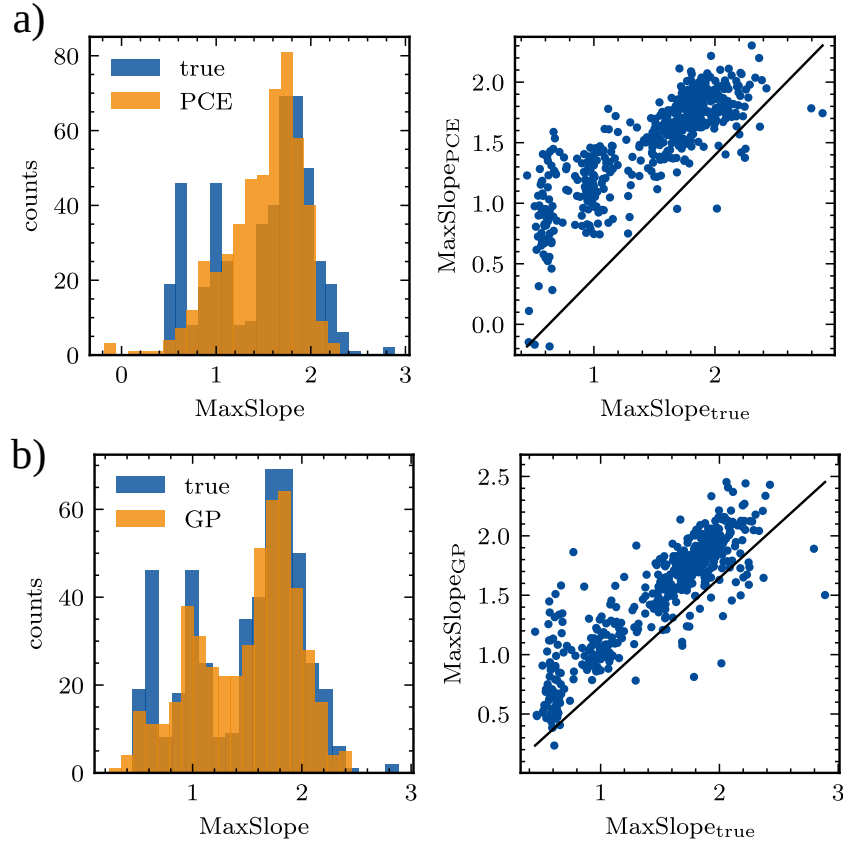


Figura 45 – Experimento I: Comparação dos emuladores PCE (a) e GP (b) no conjunto de dados de validação para a quantidade MaxSlope: histograma dos valores verdadeiros e estimados pelo emulador (esquerda) e gráfico de paridade (direita).

as análises realizadas, em especial para a quantidade MaxSlope, evidenciam limitações na capacidade desses modelos substitutos em capturar adequadamente a forte não linearidade da relação entrada-saída associada ao modelo de Mahajan (MAHAJAN; SHIFERAW, 2008). Essa limitação se manifesta, por exemplo, na dificuldade em reproduzir a dispersão

Tabela 17 – Experimento I: Estatística de Smirnov (KS) obtida pelo método de filtragem de Monte Carlo com emuladores PCE e GP para a quantidade MaxSlope, em função do número de amostras de treinamento.

Parâmetro	PCE				GP			
	500	1000	2000	3000	500	1000	2000	3000
G_{CaL}	0,6200	0,5558	0,5441	0,5918	0,5700	0,5040	0,4738	0,5300
G_{K1}	0,0870	0,0862	0,0596	0,0900	0,1305	0,1542	0,1481	0,1354
G_{Kr}	0,0267	0,0312	0,0178	0,0286	0,0334	0,0544	0,0266	0,0581
G_{Ks}	0,0243	0,0210	0,0161	0,0196	0,0241	0,0176	0,0238	0,0110
G_{Na}	0,1236	0,1162	0,0992	0,1100	0,0800	0,1233	0,0952	0,0755
G_{NaCa}	0,0562	0,0759	0,0662	0,1243	0,0812	0,0298	0,0513	0,0346
G_{tof}	0,1035	0,1813	0,1308	0,1306	0,0978	0,0436	0,1410	0,1004
G_{tos}	0,0483	0,0184	0,0155	0,0298	0,0398	0,0240	0,0740	0,0286

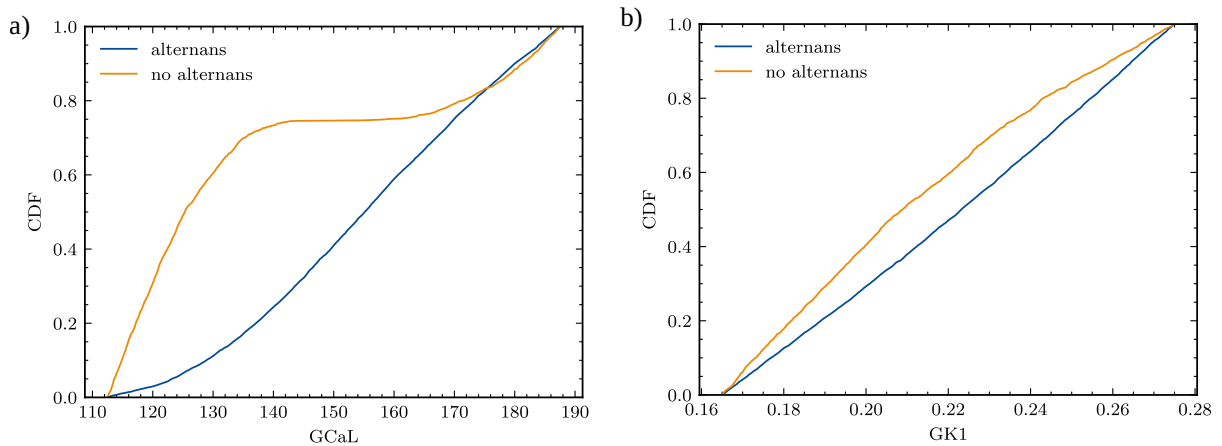


Figura 46 – Experimento I: Funções de distribuição acumulada (CDF) obtidas pela técnica MCF para os parâmetros G_{CaL} (a) e G_{K1} (b) do modelo Mahajan, distinguindo os casos de ocorrência e não ocorrência de alternans.

dos dados e na tendência de suavização das respostas.

6.1.1.3 Análise para MaxSlope com Random Forest

Os emuladores PCE e GP apresentaram limitações na aproximação da quantidade MaxSlope devido à sua forte não-linearidade em função dos parâmetros de entrada. Sendo assim, uma abordagem alternativa baseada no classificador RF foi utilizada, que atua diretamente sobre a classificação binária dos regimes de alternans e não alternans, sem necessidade de emular a quantidade MaxSlope de forma contínua. Dessa forma, o classificador RF é utilizado para guiar a análise de MCF, substituindo a etapa de avaliação do emulador pela predição direta do regime de alternans a partir dos parâmetros de entrada do modelo, como detalhado no Capítulo 4. A importância relativa de cada parâmetro é então estimada por meio da estatística de Smirnov pela abordagem MCF, utilizando as classes de alternans e não alternans previstas pelo classificador RF.

A Tabela 18 apresenta os resultados obtidos pelo classificador RF no contexto da técnica MCF para a quantidade MaxSlope. Em relação à estatística de Smirnov, os resultados confirmam G_{CaL} como o parâmetro mais determinante para a ocorrência de alternans, com valores KS entre 0.55 e 0.66 em todos os tamanhos amostrais, seguido por G_{tof} , com valores entre 0.21 e 0.24. O parâmetro G_{K1} também apresenta relevância moderada, com KS entre 0.14 e 0.20, enquanto G_{Ks} obteve os menores valores em todos os casos, indicando pouca influência na ocorrência de alternans. Esses resultados estão em boa concordância com os obtidos pelos emuladores PCE e GP apresentados na Tabela 17, reforçando a consistência da identificação dos parâmetros mais relevantes entre as diferentes abordagens.

Em relação ao desempenho computacional, o classificador RF apresentou tempos de construção praticamente constantes em torno de 1,1 s independentemente do número

Tabela 18 – Experimento I: Estatística de Smirnov (KS) obtida pelo método de filtragem de Monte Carlo com o classificador RF para a quantidade MaxSlope, em função do número de amostras de treinamento.

Parâmetro	500	1000	2000	3000
G_{CaL}	0,5728	0,6566	0,6011	0,5525
G_{K1}	0,1363	0,1854	0,1811	0,1635
G_{Kr}	0,0960	0,0964	0,0854	0,1202
G_{Ks}	0,0480	0,0281	0,0167	0,0211
G_{Na}	0,1429	0,0658	0,0731	0,1371
G_{NaCa}	0,1550	0,0749	0,0535	0,0733
G_{tof}	0,2079	0,1883	0,2156	0,2446
G_{tos}	0,1036	0,0463	0,0572	0,0804
Construção (s)	1,054	1,399	1,133	1,120
Inferência MCF (s)	0,163	0,231	0,152	0,123
BA	0,647	0,687	0,816	0,774

de amostras de treinamento, refletindo o baixo custo de ajuste do modelo. O tempo de inferência para o MCF também se mostrou estável, em torno de 0,15 s, confirmando que a avaliação do classificador é independente do tamanho do conjunto de treinamento, como esperado. A acurácia balanceada (BA) apresentou tendência geral de melhora com o aumento do número de amostras, partindo de 0,65 para 500 amostras e atingindo 0,82 para 2000 amostras, com uma leve redução para 0,77 em 3000 amostras.

6.1.2 Experimento II - 11 parâmetros

Assim como no Experimento I, emuladores PCE e GP foram construídos para as quantidades de interesse APD e MaxSlope, e um classificador RF foi treinado diretamente para identificar a ocorrência ou não de alternans. Em todos os casos, utilizaram-se quatro tamanhos de conjunto de treinamento: 500, 1000, 2000 e 3000 amostras, sendo 80% destinadas ao treinamento e os 20% restantes reservados para validação.

6.1.2.1 Análise para APD

A Tabela 19 apresenta os índices de Sobol de primeira ordem do APD para os diferentes casos com 11 parâmetros, o tempo de construção dos emuladores, o tempo de avaliação no conjunto de validação e as métricas LOO e R^2 . Pode-se observar que, à medida que aumenta o número de amostras de treinamento, os índices de Sobol de primeira ordem associados ao APD tendem a convergir, embora com oscilações ligeiramente maiores do que as observadas no Experimento I, o que é esperado dado o maior número de parâmetros na análise. Ainda assim, mesmo com 500 amostras, as estimativas já se encontram próximas de seus valores assintóticos. Assim como observado na Tabela 15, o emulador GP produziu alguns índices de Sobol ligeiramente negativos para conjuntos de

treinamento menores que 3000 amostras, valores de magnitude desprezível e provavelmente decorrentes de erros numéricos de estimação.

Tabela 19 – Experimento II: Índices de Sobol de Primeira Ordem dos parâmetros do modelo Mahajan, custo computacional e métricas de precisão dos emuladores PCE e GP para a quantidade APD, em função do número de amostras de treinamento.

Parâmetro	PCE				GP			
	500	1000	2000	3000	500	1000	2000	3000
Ca_o	0,026	0,028	0,034	0,034	0,035	0,030	0,037	0,046
G_{CaL}	0,473	0,443	0,463	0,461	0,463	0,423	0,436	0,434
G_{K1}	0,004	0,009	0,007	0,007	0,000	−0,005	0,006	0,020
G_{Kr}	0,013	0,014	0,010	0,012	0,007	0,000	0,010	0,027
G_{Ks}	0,000	0,000	0,000	0,000	0,000	−0,009	0,002	0,016
G_{Na}	0,000	0,001	0,001	0,000	−0,001	−0,009	0,003	0,016
G_{NaCa}	0,032	0,031	0,028	0,031	0,017	0,013	0,024	0,039
G_{tof}	0,246	0,288	0,262	0,263	0,238	0,245	0,250	0,260
G_{tos}	0,025	0,023	0,022	0,023	0,026	0,014	0,026	0,039
K_o	0,014	0,013	0,015	0,015	0,013	0,010	0,023	0,032
Na_o	0,116	0,093	0,095	0,091	0,078	0,065	0,080	0,099
Construção (s)	4,46	3,96	4,50	4,12	19,85	42,65	191,34	392,02
Avaliação (s)	0,98	1,12	1,13	1,27	1,89	1,13	1,28	1,26
LOO	0,109	0,108	0,108	0,104	0,041	0,038	0,029	0,025
R^2	0,906	0,862	0,906	0,904	0,938	0,949	0,972	0,975

Em relação aos índices de Sobol, G_{CaL} e G_{tof} mantêm-se como os parâmetros mais influentes no APD, em consonância com os resultados do Experimento I. Dentre os três novos parâmetros incluídos neste experimento, Na_o destaca-se como o mais relevante, com índices de primeira ordem entre 0,09 e 0,12, enquanto Ca_o e K_o apresentam contribuições desprezíveis.

Com relação ao desempenho computacional, o comportamento observado é semelhante ao do Experimento I: o tempo de construção dos emuladores PCE permanece praticamente constante em torno de 4 s independentemente do número de amostras e parâmetros, enquanto o tempo de construção dos emuladores GP cresce expressivamente com o tamanho do conjunto de treinamento, atingindo 392 s para 3000 amostras. Vale destacar que o aumento de 8 para 11 parâmetros não impactou de forma significativa o custo de construção e avaliação dos emuladores GP.

Em termos de precisão, ambos os emuladores apresentaram bom desempenho, com o GP obtendo valores de R^2 superiores e menores erros LOO em todos os casos. No entanto, comparando com o Experimento I, observa-se uma redução nos valores de R^2 para ambos os emuladores: o PCE passou de aproximadamente 0,97 para 0,90, enquanto o GP passou de aproximadamente 0,998 para 0,96, o que é esperado dado o maior número de parâmetros

e a consequente maior complexidade do espaço de parâmetros. Ainda assim, os índices de Sobol estimados por PCE e GP mostraram-se muito próximos entre si, confirmando que ambas as abordagens capturam adequadamente a estrutura de sensibilidade do modelo mesmo com o espaço de parâmetros ampliado.

A Figura 47 apresenta os resultados obtidos pelos emuladores PCE e GP nos painéis (a) e (b), respectivamente, para $N = 2000$ amostras de validação. O emulador GP reproduz com alta fidelidade o comportamento do modelo original: no histograma, as distribuições dos valores verdadeiros e preditos praticamente se sobrepõem, e no gráfico de paridade os pontos se concentram ao longo da reta identidade, evidenciando boa concordância entre os valores preditos e os de referência. O emulador PCE apresenta uma disparidade mais pronunciada do que a observada no Experimento I: o histograma exibe um deslocamento visível em relação à distribuição verdadeira, e no gráfico de paridade surgem predições próximas de zero ou negativas para valores verdadeiros em torno de 100 ms. Esse comportamento, já presente de forma mais sutil no Experimento I, torna-se mais evidente com o espaço de parâmetros ampliado, e é consistente com a redução do R^2 observada na Tabela 19.

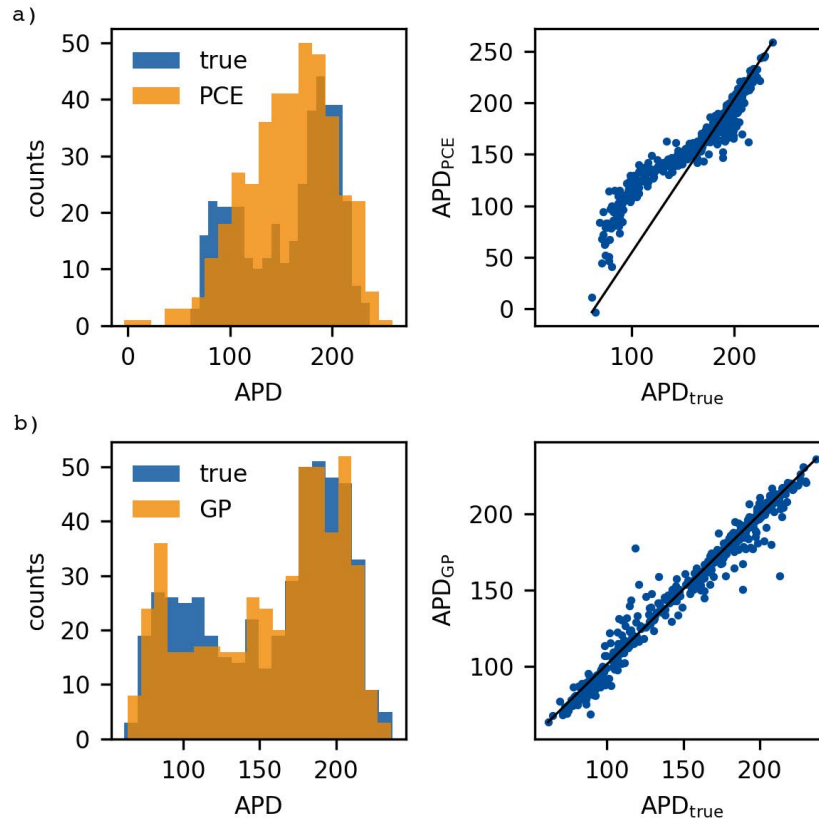


Figura 47 – Experimento II: Comparação dos emuladores PCE (a) e GP (b) no conjunto de dados de validação para a quantidade APD: histograma dos valores verdadeiros e estimados pelo emulador (esquerda) e gráfico de paridade (direita).

6.1.2.2 Análise para MaxSlope com PCE e GP

O pós-processamento adotado foi o mesmo descrito anteriormente: apenas inclinações máximas no intervalo $(0, 3]$ foram consideradas, pelo mesmo motivo de evitar distribuições desbalanceadas no treinamento dos emuladores. Após as simulações e o pós-processamento, obteve-se um total efetivo de 333, 694, 1330 e 2009 amostras válidas para os conjuntos de 500, 1000, 2000 e 3000 amostras originais, respectivamente, das quais 80% foram destinadas ao treinamento e 20% à validação.

A Tabela 20 apresenta os índices de Sobol de primeira ordem, assim como as métricas de desempenho e de erro, para os emuladores PCE e GP, no cenário do Experimento II. Pode-se observar, como anteriormente, que os emuladores PCE e GP apresentam boa concordância nos valores dos índices de Sobol, mas dessa vez, os emuladores convergem de forma mais lenta à medida que o número de amostras aumenta. Os emuladores GP obtiveram melhores métricas de precisão em comparação aos baseados em PCE. Em termos computacionais, observou-se um maior custo na construção dos emuladores GP.

A Figura 48 apresenta a comparação entre os emuladores PCE e GP no conjunto de validação para o Experimento II. No caso, nota-se no emulador PCE uma tendência ainda em superestimar o valor de MaxSlope, apesar de exibir métricas globais de precisão razoáveis, como no Experimento I. O emulador PCE apresenta uma tendência de superestimação do MaxSlope, além de produzir predições negativas visíveis no gráfico de paridade, o que é inconsistente para esta quantidade. O emulador GP, por sua vez, captura com maior fidelidade a forma global do histograma e apresenta menor dispersão no gráfico de paridade, em consonância com os melhores valores de R^2 e LOO obtidos na Tabela 20.

Em seguida, realizou-se a análise de MCF utilizando os emuladores PCE e GP. Os resultados em termos do teste de Smirnov para os emuladores PCE e GP criados com diferentes tamanhos amostrais para os parâmetros do Experimento II são apresentados na Tabela 21. Observa-se que os valores da estatística de Smirnov apresentam uma tendência de estabilização à medida que o tamanho amostral aumenta, indicando convergência dos resultados, mesmo que um pouco mais tardia essa convergência de um emulador para o outro. Pode-se notar que há uma proximidade nos valores da estatística de Smirnov no contexto da Filtragem de Monte Carlo entre os dois emuladores, PCE e GP. Isso os torna capazes de identificar de maneira correta os parâmetros mais relevantes para a ocorrência de alternans no modelo original. A ordenação dos parâmetros mais relevantes é também preservada entre os métodos.

6.1.2.3 Análise para MaxSlope com RF

A seguir são apresentados os resultados do classificador RF aplicado à análise de Filtragem de Monte Carlo para a quantidade MaxSlope no cenário do Experimento II com

Tabela 20 – Experimento II: Índices de Sobol de Primeira Ordem e Totais dos parâmetros do modelo Mahajan, custo computacional e métricas de precisão dos emuladores PCE e GP para a quantidade MaxSlope, em função do número de amostras de treinamento.

Parâmetro	PCE				GP			
	500	1000	2000	3000	500	1000	2000	3000
<i>Índices de Sobol Total</i>								
Ca_o	0,038	0,063	0,078	0,059	0,057	0,070	0,129	0,123
G_{CaL}	0,500	0,620	0,589	0,584	0,662	0,649	0,617	0,616
G_{K1}	0,034	0,010	0,022	0,030	0,022	0,067	0,037	0,054
G_{Kr}	0,000	0,021	0,017	0,023	0,043	0,022	0,023	0,026
G_{Ks}	0,000	0,000	0,000	0,000	0,002	0,002	0,002	0,002
G_{Na}	0,112	0,051	0,046	0,037	0,046	0,033	0,039	0,050
G_{NaCa}	0,049	0,052	0,029	0,039	0,085	0,054	0,047	0,075
G_{tof}	0,491	0,511	0,523	0,522	0,462	0,586	0,627	0,614
G_{tos}	0,039	0,021	0,017	0,022	0,016	0,014	0,026	0,032
K_o	0,196	0,188	0,194	0,218	0,200	0,257	0,214	0,261
Na_o	0,002	0,002	0,000	0,000	0,028	0,004	0,013	0,023
<i>Índices de Sobol de Primeira Ordem</i>								
Ca_o	0,000	0,011	0,021	0,015	0,002	0,003	0,026	-0,008
G_{CaL}	0,298	0,328	0,309	0,300	0,330	0,261	0,241	0,209
G_{K1}	0,025	0,000	0,000	0,002	0,006	-0,001	0,013	-0,005
G_{Kr}	0,000	0,003	0,000	0,002	-0,007	-0,002	0,013	-0,013
G_{Ks}	0,000	0,000	0,000	0,000	-0,005	-0,002	0,011	-0,010
G_{Na}	0,000	0,006	0,007	0,005	0,003	-0,001	0,014	-0,011
G_{NaCa}	0,000	0,000	0,000	0,001	-0,002	0,002	0,014	-0,009
G_{tof}	0,134	0,087	0,111	0,100	0,081	0,067	0,099	0,063
G_{tos}	0,020	0,000	0,000	0,000	-0,005	0,000	0,012	-0,014
K_o	0,054	0,018	0,031	0,033	0,040	0,018	0,037	0,033
Na_o	0,000	0,000	0,000	0,000	-0,003	-0,002	0,010	-0,014
Construção (s)	4,62	4,09	4,49	4,84	10,97	19,91	91,09	207,17
Avaliação (s)	1,06	1,24	1,24	1,23	0,98	1,09	1,14	1,13
LOO	0,361	0,304	0,325	0,344	0,277	0,220	0,205	0,208
R^2	0,618	0,518	0,550	0,641	0,646	0,639	0,662	0,788

11 parâmetros.

A Tabela 22 apresenta as estatísticas de Smirnov, métricas de desempenho e BA obtidas pelo classificador RF. O parâmetro G_{CaL} destaca-se novamente como o mais determinante para a ocorrência de alternans, com valores KS entre 0,43 e 0,44 em todos os tamanhos amostrais, indicando boa convergência do classificador mesmo com poucas amostras. Em seguida, G_{tof} e K_o aparecem como segundo e terceiro parâmetros mais relevantes, com KS em torno de 0,20 e 0,15, respectivamente. O parâmetro G_{Na} também apresenta relevância moderada (0,13), o que não havia sido observado com a mesma

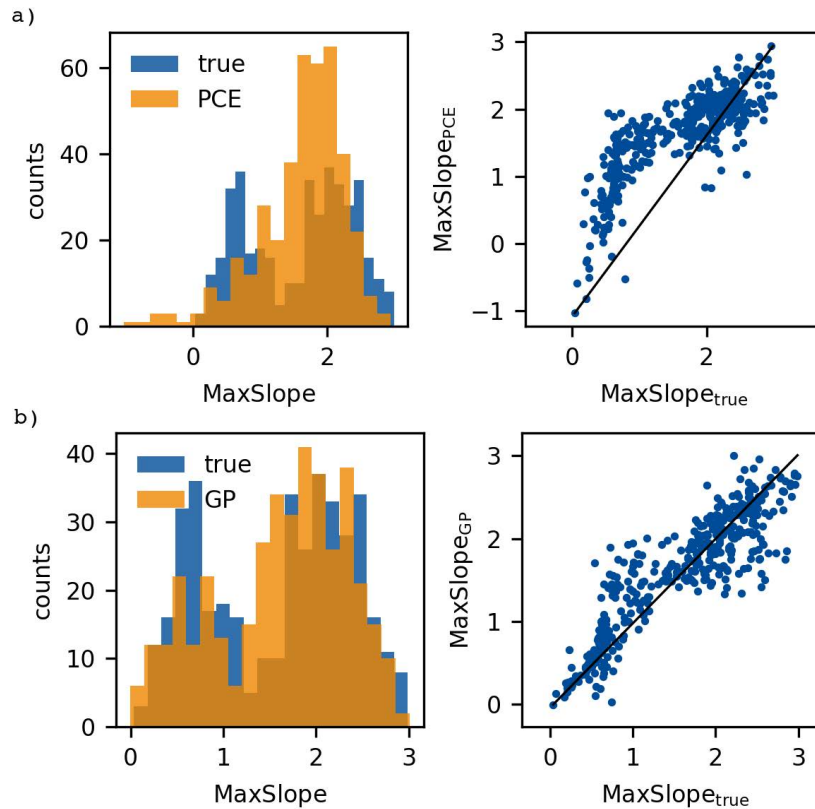


Figura 48 – Experimento II: Comparação dos emuladores PCE (a) e GP (b) no conjunto de dados de validação para a quantidade MaxSlope: histograma dos valores verdadeiros e estimados pelo emulador (esquerda) e gráfico de paridade (direita).

clareza nas análises com PCE e GP, sugerindo que o classificador RF é capaz de capturar influências não lineares que os emuladores de regressão tendem a suavizar.

Em relação ao desempenho computacional, observa-se um crescimento do tempo de construção do RF com o número de amostras, de 1,2 s para 7,9 s, comportamento esperado dado o maior número de parâmetros em relação ao Experimento I. O tempo de inferência do MCF com RF também cresce, de 0,17 s para 1,19 s, mas permanece consideravelmente inferior ao observado para PCE e GP. A acurácia balanceada aumentou de 0,69 para 500 amostras para cerca de 0,78 a partir de 1000 amostras, mantendo-se praticamente estável em torno desse valor para os conjuntos maiores, o que indica boa capacidade discriminativa do classificador e estabilidade a partir de 1000 amostras de treinamento.

6.1.3 Experimento III - 13 parâmetros

Assim como nos experimentos anteriores, no Experimento III foram construídos emuladores PCE e GP para as quantidades APD e MaxSlope, e um classificador RF foi treinado para identificar a ocorrência ou não de alternans. Em relação ao Experimento II, o espaço de parâmetros foi ampliado com a adição das concentrações intracelulares de potássio e sódio, K_i e Na_i , totalizando 13 parâmetros. Utilizaram-se os mesmos quatro tamanhos de conjunto de treinamento: 500, 1000, 2000 e 3000 amostras, sendo 80%

Tabela 21 – Experimento II: Estatística de Smirnov (KS) obtida pelo método de filtragem de Monte Carlo utilizando emuladores PCE e GP, em função do número de amostras de treinamento.

Parâmetro	PCE				GP			
	500	1000	2000	3000	500	1000	2000	3000
Ca_o	0,0425	0,0724	0,0871	0,0435	0,0347	0,0193	0,0263	0,0198
G_{CaL}	0,4276	0,3880	0,3593	0,3846	0,4186	0,3215	0,3473	0,3053
G_{K1}	0,1063	0,0473	0,0495	0,0363	0,0806	0,0335	0,0308	0,0505
G_{Kr}	0,0310	0,0456	0,0220	0,0524	0,0415	0,0266	0,0310	0,0234
G_{Ks}	0,0199	0,0258	0,0436	0,0287	0,0418	0,0164	0,0191	0,0175
G_{Na}	0,0382	0,0721	0,0626	0,0560	0,0734	0,0280	0,0698	0,0553
G_{NaCa}	0,0352	0,0518	0,0470	0,0353	0,0436	0,0425	0,0412	0,0317
G_{tof}	0,2783	0,1893	0,1915	0,2614	0,2708	0,2219	0,2386	0,2274
G_{tos}	0,0543	0,0320	0,0156	0,0181	0,0458	0,0179	0,0284	0,0442
K_o	0,2076	0,1019	0,0959	0,1545	0,1004	0,0805	0,0869	0,0831
Na_o	0,0287	0,0404	0,0365	0,0292	0,0267	0,0276	0,0168	0,0261
MCF(s)	4,615	4,778	5,228	4,752	4,523	3,823	3,764	3,772

Tabela 22 – Experimento II: Estatística de Smirnov (KS) obtida pelo método de Filtragem de Monte Carlo com o classificador RF para a quantidade MaxSlope, em função do número de amostras de treinamento.

Parâmetro	500	1000	2000	3000
Ca_o	0,0426	0,0499	0,0457	0,0404
G_{CaL}	0,4294	0,4336	0,4337	0,4383
G_{K1}	0,1009	0,0989	0,1175	0,1134
G_{Kr}	0,0688	0,0606	0,0727	0,0479
G_{Ks}	0,0214	0,0332	0,0259	0,0192
G_{Na}	0,1313	0,1427	0,1345	0,1127
G_{NaCa}	0,0261	0,0259	0,0306	0,0279
G_{tof}	0,2017	0,1841	0,2074	0,2047
G_{tos}	0,0609	0,0730	0,0637	0,0597
K_o	0,1505	0,1465	0,1586	0,1379
Na_o	0,0366	0,0547	0,0417	0,0472
Construção (s)	1,178	2,060	4,102	7,934
MCF (s)	0,170	0,303	0,573	1,190
BA	0,609	0,567	0,700	0,713

destinadas ao treinamento e os 20% restantes reservados para validação.

6.1.3.1 Análise para APD

A Tabela 23 apresenta os resultados para o Experimento III com 13 parâmetros. Os parâmetros G_{CaL} , G_{tof} e Na_i mantêm-se como os mais influentes no APD, sendo Na_i uma novidade em relação aos experimentos anteriores, com índices de primeira ordem entre

0,15 e 0,23, comparáveis aos de G_{tof} . O parâmetro Na_o , já presente no Experimento II com influência moderada, mantém índices semelhantes (0,05–0,07), enquanto K_i apresenta contribuição pequena mas não nula. Em termos de precisão, o PCE manteve R^2 em torno de 0,86, ligeiramente abaixo dos valores do Experimento II, enquanto o GP melhorou progressivamente com o número de amostras, atingindo 0,951 para 3000 amostras. O custo computacional segue o padrão já observado: construção do PCE constante em torno de 4–6 s, e do GP crescendo de aproximadamente 20 s para 414 s com o aumento do conjunto de treinamento.

Tabela 23 – Experimento III: Índices de Sobol de Primeira Ordem dos parâmetros do modelo Mahajan, custo computacional e métricas de precisão dos emuladores PCE e GP para a quantidade APD, em função do número de amostras de treinamento.

Parâmetro	PCE				GP			
	500	1000	2000	3000	500	1000	2000	3000
K_i	0,016	0,019	0,018	0,016	0,002	0,020	0,026	0,022
Ca_o	0,032	0,025	0,028	0,028	0,028	0,032	0,034	0,029
G_{CaL}	0,353	0,353	0,343	0,361	0,346	0,345	0,331	0,332
G_{K1}	0,007	0,007	0,005	0,005	−0,001	0,012	0,017	0,014
G_{Kr}	0,009	0,009	0,008	0,008	0,005	0,016	0,022	0,018
G_{Ks}	0,000	0,001	0,000	0,000	−0,004	0,010	0,013	0,009
G_{Na}	0,000	0,000	0,000	0,000	−0,004	0,012	0,015	0,010
G_{NaCa}	0,026	0,021	0,022	0,021	0,017	0,028	0,032	0,030
G_{tof}	0,170	0,159	0,167	0,173	0,173	0,165	0,153	0,148
G_{tos}	0,017	0,009	0,016	0,015	0,016	0,025	0,030	0,025
K_o	0,014	0,006	0,007	0,008	0,005	0,022	0,025	0,021
Na_i	0,204	0,228	0,229	0,210	0,151	0,199	0,206	0,194
Na_o	0,051	0,057	0,054	0,055	0,036	0,063	0,068	0,067
Construção (s)	4,415	6,204	3,664	5,335	19,850	44,177	184,594	414,073
Avaliação (s)	0,992	1,119	1,260	1,265	1,117	1,100	2,749	2,783
LOO	0,113	0,120	0,133	0,121	0,079	0,073	0,062	0,054
R^2	0,856	0,862	0,872	0,859	0,851	0,915	0,927	0,951

6.1.3.2 Análise para MaxSlope com PCE e GP

A Tabela 24 apresenta os índices de Sobol e métricas de desempenho para PCE e GP no Experimento III. As tendências observadas nos experimentos anteriores mantêm-se: G_{CaL} e G_{tof} continuam como os parâmetros mais influentes na máxima inclinação da curva de restituição (MaxSlope), com índices totais entre 0,48–0,63 e 0,48–0,58, respectivamente. O parâmetro K_o apresenta relevância moderada (0,11–0,18), enquanto K_i aparece com influência não desprezível (0,18–0,23), reflexo da adição das concentrações intracelulares neste experimento. Os valores de R^2 para o MaxSlope apresentam uma leve degradação

em relação aos experimentos anteriores, com o PCE atingindo valores entre 0,44 e 0,63 e o GP entre 0,55 e 0,73, reflexo do maior número de parâmetros neste experimento.

Tabela 24 – Experimento III: Índices de Sobol de Primeira Ordem e Totais dos parâmetros do modelo Mahajan, custo computacional e métricas de precisão dos emuladores PCE e GP para a quantidade MaxSlope, em função do número de amostras de treinamento.

Parâmetro	PCE				GP			
	500	1000	2000	3000	500	1000	2000	3000
<i>Índices de Sobol Total</i>								
K_i	0,200	0,178	0,208	0,196	0,226	0,155	0,216	0,212
Ca_o	0,106	0,078	0,125	0,110	0,122	0,149	0,156	0,153
G_{CaL}	0,506	0,594	0,537	0,555	0,627	0,624	0,614	0,629
G_{K1}	0,018	0,039	0,025	0,031	0,004	0,028	0,024	0,023
G_{Kr}	0,049	0,009	0,026	0,044	0,016	0,012	0,032	0,040
G_{Ks}	0,007	0,000	0,000	0,003	0,004	0,002	0,000	0,001
G_{Na}	0,000	0,000	0,029	0,027	0,013	0,013	0,027	0,024
G_{NaCa}	0,077	0,072	0,056	0,044	0,058	0,051	0,076	0,076
G_{tof}	0,517	0,524	0,483	0,484	0,581	0,570	0,514	0,536
G_{tos}	0,038	0,024	0,026	0,037	0,012	0,034	0,018	0,038
K_o	0,114	0,143	0,125	0,138	0,140	0,184	0,184	0,179
Na_i	0,018	0,000	0,010	0,005	0,012	0,008	0,010	0,003
Na_o	0,014	0,016	0,006	0,000	0,010	0,052	0,023	0,011
<i>Índices de Sobol de Primeira Ordem</i>								
K_i	0,000	0,000	0,009	0,013	-0,019	-0,011	0,006	0,015
Ca_o	0,048	0,025	0,040	0,027	0,008	-0,006	0,017	0,024
G_{CaL}	0,159	0,184	0,164	0,163	0,147	0,166	0,176	0,162
G_{K1}	0,000	0,000	0,002	0,001	-0,019	-0,015	-0,002	0,011
G_{Kr}	0,011	0,000	0,000	0,004	-0,009	-0,009	-0,001	0,009
G_{Ks}	0,000	0,000	0,000	0,000	-0,018	-0,014	-0,004	0,011
G_{Na}	0,000	0,000	0,002	0,002	-0,019	-0,013	0,000	0,017
G_{NaCa}	0,005	0,004	0,004	0,003	-0,014	-0,010	-0,003	0,009
G_{tof}	0,082	0,081	0,088	0,082	0,071	0,060	0,058	0,064
G_{tos}	0,000	0,000	0,002	0,000	-0,019	-0,011	-0,008	0,011
K_o	0,020	0,023	0,021	0,021	0,008	0,016	0,030	0,039
Na_i	0,000	0,000	0,003	0,001	-0,017	-0,012	-0,003	0,010
Na_o	0,000	0,000	0,000	0,000	-0,016	-0,011	0,000	0,014
Construção (s)	4,389	4,674	4,897	5,226	15,986	23,434	101,137	202,012
Avaliação (s)	0,980	1,130	1,151	1,705	1,006	1,158	1,160	1,304
LOO	0,364	0,408	0,401	0,395	0,321	0,327	0,278	0,277
R^2	0,436	0,463	0,600	0,627	0,618	0,553	0,701	0,733

Os resultados do MCF via PCE e GP são apresentados na Tabela 25, e os do classificador RF na Tabela 26. Em ambos os casos, G_{CaL} mantém-se como o parâmetro mais determinante para a ocorrência de alternans, com KS entre 0,26 e 0,43, seguido por

G_{tof} e K_o . O classificador RF continua apresentando os menores tempos de inferência MCF, abaixo de 0,15 s, em contraste com ~ 4 s do PCE e ~ 5 s do GP, reforçando sua eficiência computacional mesmo com o espaço de parâmetros ampliado.

Tabela 25 – Experimento III: Estatística de Smirnov (KS) obtida pelo método de Filtragem de Monte Carlo utilizando emuladores PCE e GP, em função do número de amostras de treinamento.

Parâmetro	PCE				GP			
	500	1000	2000	3000	500	1000	2000	3000
K_i	0,0526	0,0657	0,0988	0,1151	0,0569	0,0598	0,0629	0,0656
Ca_o	0,1191	0,0788	0,0761	0,0579	0,0734	0,0259	0,0344	0,0382
G_{CaL}	0,2791	0,2769	0,2637	0,2731	0,2753	0,2870	0,2701	0,2729
G_{K1}	0,0212	0,0362	0,0327	0,0301	0,0153	0,0410	0,0398	0,0317
G_{Kr}	0,0561	0,0254	0,0193	0,0515	0,0495	0,0438	0,0198	0,0268
G_{Ks}	0,0269	0,0140	0,0245	0,0239	0,0175	0,0117	0,0146	0,0113
G_{Na}	0,0274	0,0274	0,0373	0,0260	0,0316	0,0434	0,0574	0,0618
G_{NaCa}	0,0421	0,0328	0,0383	0,0290	0,0275	0,0424	0,0214	0,0327
G_{tof}	0,1493	0,1750	0,1771	0,1493	0,2163	0,1708	0,1471	0,1487
G_{tos}	0,0417	0,0290	0,0183	0,0244	0,0210	0,0203	0,0137	0,0189
K_o	0,1163	0,1382	0,1151	0,1145	0,0943	0,1056	0,0977	0,0712
Na_i	0,0324	0,0224	0,0388	0,0230	0,0257	0,0296	0,0297	0,0314
Na_o	0,0202	0,0191	0,0250	0,0206	0,0178	0,0432	0,0242	0,0225
Inferência (s)	4,035	4,016	4,036	3,993	5,678	5,822	5,663	4,306

6.1.3.3 Índices de Sobol de segunda ordem

Além dos índices de primeira ordem e totais, foram calculados os índices de Sobol de segunda ordem para o MaxSlope, que quantificam a contribuição das interações entre pares de parâmetros para a variância da resposta. A Figura 49 apresenta os cinco pares mais relevantes, obtidos com os emuladores PCE e GP para o caso de 3000 amostras.

A interação mais expressiva ocorre entre a condutância do cálcio do tipo L (G_{CaL}) e a corrente transitória de saída rápida (G_{tof}), com índice bastante superior ao dos demais pares em ambos os emuladores. Esse resultado é coerente com os índices de primeira ordem, nos quais G_{CaL} e G_{tof} já figuravam como os dois parâmetros individualmente mais influentes, indicando que os dois parâmetros mais relevantes isoladamente formam também a interação mais forte. Nota-se ainda que o G_{CaL} está presente em três dos cinco pares de maior índice, o que reforça o papel central do cálcio do tipo L não apenas em seu efeito individual, mas também em suas interações com os demais parâmetros.

Os dois emuladores concordam quanto ao ordenamento dos pares mais relevantes, com o PCE atribuindo valores sistematicamente um pouco superiores aos do GP. Essa

Tabela 26 – Experimento III – Estatística de Smirnov (KS) obtida pelo método de Filtragem de Monte Carlo com o classificador RF para a quantidade MaxSlope, em função do número de amostras de treinamento.

Parâmetro	500	1000	2000	3000
K_i	0,1206	0,2498	0,1363	0,1542
Ca_o	0,1127	0,0645	0,0514	0,0743
G_{CaL}	0,4154	0,2648	0,4188	0,4305
G_{K1}	0,0656	0,2779	0,0751	0,0538
G_{Kr}	0,0507	0,0932	0,0490	0,0481
G_{Ks}	0,0531	0,0186	0,0183	0,0339
G_{Na}	0,0922	0,1945	0,1102	0,1151
G_{NaCa}	0,1385	0,0571	0,0631	0,0608
G_{tof}	0,1334	0,2112	0,2160	0,2531
G_{tos}	0,0818	0,0692	0,0657	0,0459
K_o	0,3183	0,0507	0,0435	0,0661
Na_i	0,0348	0,0421	0,1664	0,1382
Na_o	0,0751	0,0190	0,0664	0,0240
Construção (s)	1,331	1,398	1,077	0,624
Inferência MCF (s)	0,147	0,157	0,129	0,133
BA	0,579	0,701	0,672	0,672

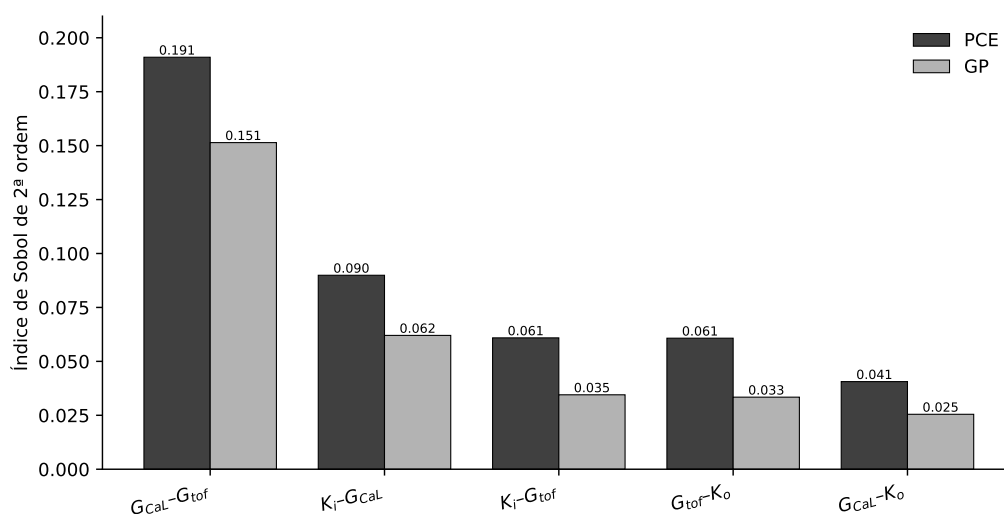


Figura 49 – Cinco maiores índices de Sobol de segunda ordem para o MaxSlope no modelo Mahajan-Shiferaw, obtidos com os emuladores PCE e GP para o caso de 3000 amostras.

concordância entre duas técnicas de natureza distinta reforça a confiabilidade dos resultados obtidos.

Por fim, cabe destacar que, para o MaxSlope, o índice de segunda ordem do par $G_{CaL}-G_{tof}$ supera o índice de primeira ordem de qualquer parâmetro isolado. Isso indica que uma parcela importante da influência dos parâmetros sobre o MaxSlope se manifesta por meio de interações, e não apenas por efeitos individuais. Esse tipo de informação

não é acessível pela Filtragem de Monte Carlo, de caráter univariado, o que evidencia a complementaridade entre as duas abordagens de análise de sensibilidade empregadas neste trabalho.

6.2 Comparativo entre os experimentos

Esta seção apresenta uma análise comparativa dos resultados obtidos nos três experimentos realizados com o modelo Mahajan, considerando o efeito do aumento do espaço de parâmetros sobre a qualidade dos emuladores, o desempenho computacional e a consistência das análises de sensibilidade.

6.2.1 Qualidade dos emuladores

A Tabela 27 resume os valores de R^2 obtidos com 3000 amostras de treinamento para as quantidades APD e MaxSlope, em função do número de parâmetros. Para o APD, observa-se uma degradação monotônica e expressiva à medida que o espaço de parâmetros cresce: o PCE passa de 0,969 no Experimento I para 0,859 no Experimento III, enquanto o GP passa de 0,999 para 0,951. Essa degradação é esperada, pois o maior número de parâmetros aumenta a dimensionalidade do problema e exige mais amostras para uma boa cobertura do espaço de entrada.

Para o MaxSlope, os valores de R^2 já são consideravelmente mais baixos desde o Experimento I, com o PCE obtendo 0,674 e o GP 0,766 mesmo com apenas 8 parâmetros, evidenciando que a forte não linearidade intrínseca desta quantidade é a principal fonte de dificuldade de emulação. Com o aumento do espaço de parâmetros, observa-se uma degradação adicional, embora mais suave do que a observada para o APD: o PCE passa de 0,674 no Experimento I para 0,627 no Experimento III, e o GP de 0,766 para 0,733, indicando que a dimensionalidade contribui para agravar a dificuldade. Em todos os experimentos e quantidades, o GP obteve valores de R^2 superiores aos do PCE, confirmando sua maior capacidade de aproximação para este modelo.

Tabela 27 – Comparativo dos coeficientes de determinação R^2 obtidos com 3000 amostras de treinamento para os emuladores PCE e GP das quantidades APD e MaxSlope do modelo Mahajan em função do número de parâmetros.

Experimento	APD		MaxSlope	
	PCE	GP	PCE	GP
Exp. I (8 parâmetros)	0,969	0,999	0,674	0,766
Exp. II (11 parâmetros)	0,904	0,975	0,641	0,788
Exp. III(13 parâmetros)	0,859	0,951	0,627	0,733

Para PCE e GP, a classificação binária é obtida aplicando-se o limiar $Y > 1$ sobre os valores contínuos preditos pelos emuladores, tornando a comparação direta com o RF

possível. A Tabela 28 apresenta os valores de BA definida na Eq. (4.4) para os três métodos no Experimento III, em função do número de amostras de treinamento.

Tabela 28 – Comparativo da acurácia balanceada (BA) para os emuladores PCE, GP e o classificador RF no Experimento III (13 parâmetros) em função do número de amostras de treinamento – modelo Mahajan.

Método	500	1000	2000	3000
PCE	0,8242	0,7500	0,7527	0,7636
GP	0,8421	0,8068	0,8407	0,8250
RF	0,6087	0,5670	0,7001	0,7129

Em termos de acurácia balanceada, o emulador GP obteve o melhor desempenho de classificação, com valores entre 0,81 e 0,84, seguido pelo PCE, com valores entre 0,75 e 0,82. O classificador RF apresentou os menores valores, entre 0,56 e 0,75, com maior variabilidade entre os tamanhos amostrais e tendência de melhora com o aumento do conjunto de treinamento.

6.2.2 Desempenho computacional

A Tabela 29 apresenta os tempos de construção dos emuladores PCE, GP e do classificador RF em função do número de amostras e do experimento. O PCE mantém tempos praticamente constantes entre 3 s e 6 s em todos os experimentos e tamanhos amostrais, independentemente do número de parâmetros. O GP, por outro lado, apresenta crescimento expressivo com o número de amostras, variando de dezenas de segundos para 500 amostras até centenas de segundos para 3000 amostras, refletindo o custo cúbico do ajuste do modelo gaussiano. O classificador RF apresentou os menores tempos de construção no Experimento I, em torno de 1 s, mas seu custo cresce com o número de parâmetros, atingindo cerca de 8 s no Experimento II para 3000 amostras.

A Tabela 30 apresenta os tempos de inferência MCF. O PCE e o GP apresentam tempos comparáveis entre si, na faixa de 1–8 s, com crescimento observado do Experimento I para os demais. O classificador RF destaca-se pelos menores tempos de inferência em todos os casos, abaixo de 1,2 s no Experimento II, confirmando sua eficiência computacional na etapa de classificação do MCF.

6.2.3 Consistência da análise de sensibilidade

Em relação à identificação dos parâmetros mais relevantes para a ocorrência de alternans, os três métodos (PCE, GP e RF) apresentaram resultados consistentes entre si em todos os experimentos. O parâmetro G_{CaL} foi unanimemente identificado como o mais determinante, com estatísticas KS médias entre 0,35 e 0,58 dependendo do método e do experimento. O parâmetro G_{tof} aparece consistentemente como o segundo mais

Tabela 29 – Comparativo dos tempos de construção (em segundos) dos emuladores PCE, GP e do classificador RF em função do número de amostras de treinamento.

Método	Experimento	500	1000	2000	3000
PCE	Exp. I (8 param.)	4,09	4,40	3,31	3,46
	Exp. II (11 param.)	4,46	3,96	4,50	4,12
	Exp. III (13 param.)	4,54	3,48	4,03	5,20
GP	Exp. I (8 param.)	17,83	35,96	166,36	405,26
	Exp. II (11 param.)	19,85	42,65	191,34	392,02
	Exp. III (13 param.)	19,96	55,96	189,43	417,00
RF	Exp. I (8 param.)	1,054	1,399	1,133	1,120
	Exp. II (11 param.)	1,178	2,060	4,102	7,934
	Exp. III (13 param.)	1,331	1,398	1,077	0,624

Tabela 30 – Comparativo dos tempos de inferência MCF (em segundos) dos emuladores PCE, GP e do classificador RF em função do número de amostras de treinamento

Método	Experimento	500	1000	2000	3000
PCE	Exp. I (8 param.)	0,93	0,93	1,09	1,08
	Exp. II (11 param.)	4,62	4,78	5,23	4,75
	Exp. III(13 param.)	4,00	7,59	4,92	8,06
GP	Exp. I (8 param.)	0,96	0,96	2,57	1,11
	Exp. II (11 param.)	4,52	3,82	3,76	3,77
	Exp. III(13 param.)	4,10	4,11	4,51	5,35
RF	Exp. I (8 param.)	0,163	0,231	0,152	0,123
	Exp. II (11 param.)	0,170	0,303	0,573	1,190
	Exp. III (13 param.)	0,147	0,157	0,129	0,133

relevante, com KS entre 0,10 e 0,26. Observa-se que os valores KS de G_{CaL} tendem a diminuir ligeiramente com o aumento do número de parâmetros, o que é esperado, uma vez que a variância total é redistribuída entre um número maior de variáveis. A ordenação dos demais parâmetros é preservada em linhas gerais entre os métodos, com pequenas variações nas posições intermediárias do ranking, reforçando a robustez das conclusões independentemente da técnica de emulação utilizada.

6.3 Comparativo entre regressores e classificadores

Esta seção apresenta um comparativo sistemático entre cinco abordagens para a análise de sensibilidade do MaxSlope via MCF:

- três regressores (PCE, GP e RF) seguidos da aplicação do limiar ($\hat{Y} > 1$) para definição da ocorrência de alternans;
- e dois classificadores diretos (RF e GP).

Em particular, diferentemente dos métodos aplicados nos estudos anteriores, aqui foram criados o modelo de regressão de RF (que anteriormente tinha sido criado apenas para classificação) e o modelo de classificação GP, para ampliar as comparações entre as abordagens. Todas as análises foram realizadas no Experimento III (13 parâmetros), permitindo isolar o efeito da estratégia (regressão vs. classificação) dentro de um mesmo método (RF e GP).

6.3.1 Qualidade de classificação

A Tabela 31 apresenta a acurácia balanceada para os cinco métodos. O classificador GP destaca-se como a abordagem de melhor desempenho global, atingindo BA de até 0,91 com 3000 amostras, e exibindo uma tendência clara de melhora com o aumento do conjunto de treinamento. O regressor GP mantém-se estável e competitivo, com BA entre 0,81 e 0,84 em todos os tamanhos amostrais, configurando-se como a opção mais consistente independentemente do volume de dados.

Por outro lado, ambas as estratégias baseadas em RF (regressor e classificador) apresentaram desempenho sensivelmente inferior, com BA entre 0,57 e 0,71, sem uma tendência de melhora clara com o aumento do número de amostras. O PCE, embora superior ao RF, manteve-se abaixo do GP em ambas as estratégias, com BA entre 0,75 e 0,82.

Tabela 31 – Acurácia balanceada (BA) para os cinco métodos no Experimento III (13 parâmetros), em função do número de amostras de treinamento – quantidade MaxSlope, modelo Mahajan.

Método	500	1000	2000	3000
PCE (regressor)	0,8242	0,7500	0,7527	0,7636
GP (regressor)	0,8421	0,8068	0,8407	0,8250
RF (regressor)	0,5789	0,7005	0,6721	0,6724
RF (classificador)	0,6087	0,5670	0,7001	0,7129
GP (classificador)	0,6394	0,8197	0,9098	0,9021

Um resultado interessante é que, para o método GP, a estratégia de classificação direta supera a estratégia de regressão e limiarização à medida que o número de amostras aumenta (BA de 0,90–0,91 contra 0,81–0,84 a partir de 2000 amostras). Já para o RF, não há vantagem clara de uma estratégia sobre a outra já que ambas oscilam na mesma faixa, sugerindo que a limitação está relacionada à natureza do método (estrutura baseada em árvores) e não à estratégia escolhida.

6.3.2 Qualidade de regressão

A Tabela 32 apresenta o R^2 para os três regressores. O GP obteve os melhores resultados em todos os tamanhos amostrais (média de 0,676), seguido pelo PCE (média de 0,549) e pelo RF (média de 0,469), que apresentou o desempenho mais fraco como regressor desta quantidade, reforçando a dificuldade do RF em capturar a relação entre os parâmetros e o MaxSlope tanto como regressor quanto como classificador.

Tabela 32 – Coeficiente de determinação R^2 para os regressores PCE, GP e RF da quantidade MaxSlope no Experimento III (13 parâmetros) em função do número de amostras de treinamento.

Método	500	1000	2000	3000
PCE	0,4369	0,5441	0,6139	0,6003
GP	0,6183	0,6189	0,7591	0,7080
RF	0,3568	0,5068	0,4779	0,5341

6.3.3 Desempenho computacional

A Tabela 33 apresenta os tempos de construção e inferência MCF para os cinco métodos. O RF, tanto como regressor quanto como classificador, é consistentemente o mais rápido em ambas as etapas, com tempos de construção inferiores a 1,2 s e de inferência MCF inferiores a 0,19 s em todos os casos. O PCE mantém tempos de construção moderados (~ 5 s) mas apresenta tempos de inferência MCF relativamente altos (~ 4 – $5,5$ s), associados ao elevado número de avaliações do polinômio no conjunto de 10 000 amostras do MCF. O GP é o método mais custoso na construção, especialmente como regressor, atingindo quase 200 s para 3000 amostras; como classificador, seu custo de construção é mais moderado para conjuntos pequenos, mas cresce de forma acentuada a partir de 2000 amostras (35–56 s).

Tabela 33 – Tempos de construção e inferência MCF (em segundos) para os cinco métodos no Experimento III (13 parâmetros), em função do número de amostras de treinamento – quantidade MaxSlope, modelo Mahajan.

Método	Construção (s)				Inferência MCF (s)			
	500	1000	2000	3000	500	1000	2000	3000
PCE (regressor)	5,852	4,642	4,854	5,191	5,456	5,404	3,989	4,000
GP (regressor)	15,045	24,477	111,520	198,779	3,944	5,608	5,457	5,458
RF (regressor)	0,465	0,340	0,369	0,373	0,086	0,076	0,075	0,075
RF (classificador)	0,781	1,132	0,860	0,633	0,089	0,182	0,145	0,156
GP (classificador)	2,036	2,521	35,445	55,858	0,201	0,286	0,417	0,726

6.3.4 Discussão integrada dos resultados

Considerando conjuntamente qualidade de classificação e custo computacional, observa-se um claro *trade-off*: o classificador GP oferece a melhor qualidade de classificação a partir de 2000 amostras, mas ao custo de tempos de construção elevados ultrapassam 35 s; o regressor GP oferece precisão estável e competitiva com custo de construção também elevado, mas inferência mais rápida; o RF, em ambas as estratégias, oferece o menor custo computacional de toda a comparação, sendo ordens de magnitude mais rápido tanto na construção quanto na inferência, ao custo de uma qualidade de classificação sensivelmente inferior. O PCE posiciona-se como uma alternativa intermediária, com custo de construção baixo e estável, mas tempo de inferência MCF comparável ao do GP.

A seguir é apresentada uma análise visual que agrega as diferentes características dos métodos: tempo de construção, tempo de inferência e precisão. Para possibilitar a representação conjunta dos diferentes indicadores no gráfico radar, os tempos de construção e de inferência foram transformados por meio de uma função logarítmica inversa, enquanto a acurácia balanceada foi normalizada pelo método min-max. Para os tempos, considerando que menores valores são desejáveis, utilizou-se

$$S_t = \frac{1}{1 + \log_{10}(t/t_{\min})}, \quad (6.1)$$

em que t representa o tempo de construção ou de inferência e $t_{\min}$ corresponde ao menor tempo observado entre os métodos. Para a acurácia balanceada, em que maiores valores são desejáveis, utilizou-se

$$S_{BA} = \frac{BA - BA_{\min}}{BA_{\max} - BA_{\min}}. \quad (6.2)$$

Assim, os três indicadores são expressos em uma escala comum, na qual valores próximos de 1 representam melhor desempenho relativo.

Um resumo gráfico considerando a precisão e os tempos de construção e de inferência é apresentado na Figura 50. Observa-se uma relação de compromisso entre essas métricas: métodos com menor custo computacional, como o RF, apresentam menor precisão, enquanto métodos que proporcionam maior precisão tendem a demandar mais tempo, especialmente durante a etapa de construção. Dessa forma, os resultados evidenciam o *trade-off* entre custo computacional e desempenho preditivo dos diferentes métodos avaliados.

Em síntese, para as análises realizadas e para o tamanho dos conjuntos de dados considerados, o classificador GP se mostrou uma alternativa particularmente atrativa, por apresentar elevada precisão e tempos de inferência reduzidos. Entretanto, à medida que o tamanho do conjunto de dados aumenta, o custo computacional associado à etapa de construção do modelo pode se tornar um fator limitante. Nesses cenários, o RF pode constituir uma alternativa mais adequada, em razão de seu menor tempo de construção.

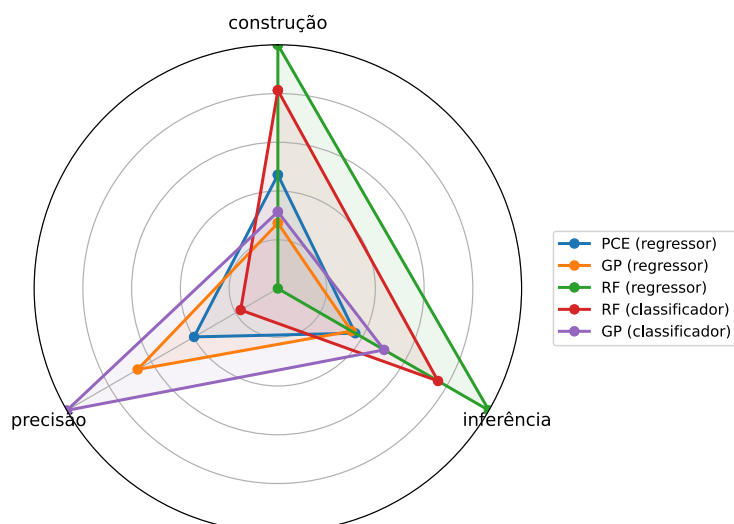


Figura 50 – Gráfico de radar para as quantidades normalizadas: tempo de construção, tempo de inferência da análise MCF e precisão (BA).

Apesar das diferenças observadas em termos de precisão e custo computacional, todos os métodos apresentaram resultados consistentes entre si, não sendo observada, para os casos avaliados, nenhuma limitação que inviabilizasse a utilização de qualquer uma das abordagens na análise de sensibilidade da ocorrência de alternans.

6.4 Discussão e posicionamento na literatura

A utilização de modelos substitutos e técnicas de aprendizado de máquina para a análise de sensibilidade tem sido explorada como uma alternativa para reduzir o custo computacional associado à avaliação repetida de modelos complexos. Em particular, Antoniadis, Lambert-Lacroix e Poggi (2021) discutem o uso de Random Forests como modelos substitutos para análise global de sensibilidade, destacando sua capacidade de representar relações não lineares e interações entre as variáveis de entrada. Trabalhos mais recentes também têm explorado a integração entre análise de sensibilidade e modelos de aprendizado de máquina, como no framework baseado em Random Forests proposto para análise de sensibilidade no contexto de assisted history matching (AULIA et al., 2019).

No contexto de problemas nos quais o interesse está associado à ocorrência de um comportamento específico, a utilização de classificadores em conjunto com técnicas de análise de sensibilidade regional mostra-se particularmente atrativa. Essa abordagem permite que a análise seja direcionada aos parâmetros associados à ocorrência do comportamento de interesse, em vez de considerar exclusivamente a variabilidade global da resposta. Essa característica é especialmente relevante para o problema considerado neste trabalho, no qual o objetivo é identificar os parâmetros associados à ocorrência de alternans cardíaco.

Até onde foi possível identificar na literatura, a combinação de MCF com modelos substitutos baseados em PCE, GP e RF ainda não havia sido aplicada à análise de sensi-

bilidade da ocorrência de alternans cardíaco. Neste trabalho, essa estratégia é explorada sistematicamente, considerando tanto modelos de regressão quanto de classificação. Dessa forma, a abordagem proposta combina a eficiência computacional proporcionada pelos modelos substitutos com a capacidade do MCF de identificar os parâmetros associados especificamente à ocorrência do fenômeno de interesse.

Embora a aplicação apresentada neste trabalho esteja centrada na ocorrência de alternans cardíaco, a metodologia proposta não é restrita a esse contexto, podendo ser empregada em outros problemas nos quais o interesse esteja na identificação dos parâmetros associados à ocorrência de um determinado comportamento. A combinação de MCF com modelos substitutos baseados em PCE, GP ou RF constitui, portanto, uma estratégia geral para análises de sensibilidade direcionadas a eventos ou regiões específicas do espaço de respostas.

7 CONCLUSÕES

Este trabalho apresentou uma abordagem para a análise de sensibilidade de modelos celulares da eletrofisiologia cardíaca, com foco na instabilidade conhecida como alternans da duração do potencial de ação. A contribuição central é o uso de modelos substitutos (emuladores) para acelerar as análises de sensibilidade, tanto por meio da Filtragem de Monte Carlo (MCF) quanto pelos índices de Sobol, permitindo identificar quais parâmetros dos modelos são determinantes para a ocorrência de alternans. Enquanto a Filtragem de Monte Carlo fornece uma visão univariada da influência de cada parâmetro para este fenômeno de interesse, os índices de Sobol complementam essa análise ao quantificar a contribuição de cada parâmetro para a variância da resposta e ao revelar interações entre eles. Sem os emuladores, essas análises exigiriam um número de simulações do modelo completo que as tornaria proibitivas em termos de custo computacional.

As análises foram conduzidas em três modelos de complexidade crescente: o modelo modificado de Mitchell-Schaeffer (MMS), o modelo biofisiologicamente detalhado de ten Tusscher-Panfilov (TP) e o modelo Mahajan-Shiferaw, estes dois últimos com um grande número de parâmetros. O método de Monte Carlo aplicado ao modelo completo foi utilizado como resposta de referência e, no caso do modelo simplificado MMS, as análises foram ainda comparadas com fórmulas analíticas da curva de restituição, o que permitiu validar a abordagem. Para as quantidades de interesse adotadas, a duração do potencial de ação e a máxima inclinação da curva de restituição, os emuladores reproduziram os índices de sensibilidade de referência com boa precisão e a um custo computacional muito menor.

Além do emulador baseado na expansão em caos polinomial, empregado inicialmente, o trabalho estendeu a abordagem a outras duas técnicas de modelagem substituta: os Processos Gaussianos e as Florestas Aleatórias. Essas técnicas foram avaliadas não apenas como regressores das quantidades de interesse, mas também como classificadores diretos da ocorrência de alternans, e seu desempenho foi comparado de forma sistemática em termos de precisão de classificação e de custo computacional. Essa comparação permitiu caracterizar as vantagens e limitações de cada estratégia, oferecendo subsídios para a escolha da técnica mais adequada conforme o modelo e o objetivo da análise. Em particular, a Filtragem de Monte Carlo foi combinada com as três técnicas de emulação (PCE, GP e RF) na detecção de alternans, uma articulação ainda não explorada nesse contexto de modelos cardíacos.

Os resultados identificaram os parâmetros determinantes para o alternans de APD nos modelos estudados, tomando-se como parâmetros as condutâncias máximas dos canais iônicos. Para o modelo TP, biofisiologicamente detalhado, a condutância da corrente de cálcio do tipo L mostrou-se o parâmetro mais relevante para a ocorrência de alternans,

resultado consistente com a literatura e com o papel conhecido do cálcio nesse fenômeno. As diferentes técnicas apresentaram resultados concordantes entre si e, em vários casos, alinhados com dados já relatados na literatura, o que reforça a validade da abordagem.

Vale destacar que os mecanismos eletrofisiológicos subjacentes ao alternans já contam com algum entendimento, com destaque para o papel do cálcio (CLUSIN, 2008; WEISS et al., 2006). Dessa forma, a análise de sensibilidade proposta neste trabalho não visa primariamente à descoberta desses mecanismos, mas cumpre um papel complementar. Por um lado, ela funciona como um instrumento de verificação: ao recuperar, a partir dos mecanismos internos de um modelo biofisicamente complexo, a importância de parâmetros já reconhecidos pela eletrofisiologia, como a condutância do cálcio do tipo L, a análise atesta a coerência do modelo com o conhecimento estabelecido. Por outro, em modelos com grande número de parâmetros e interações, nos quais a intuição fisiológica é insuficiente para hierarquizar as influências e, sobretudo, para identificar interações entre parâmetros, a abordagem quantitativa fornece informação que não seria acessível de outra forma.

De modo geral, a metodologia proposta mostrou-se eficiente e robusta, viabilizando uma análise de sensibilidade aprofundada com economia expressiva de tempo e de recursos computacionais. Embora o alternans tenha sido o comportamento adotado como caso de estudo, a abordagem não é específica a ele: por depender apenas da classificação da resposta do modelo em regimes distintos, ela pode ser aplicada à investigação de outros comportamentos da eletrofisiologia cardíaca, em nível celular ou tecidual, cuja ocorrência se deseje relacionar aos parâmetros do modelo, como bloqueios de condução, pós-despolarizações ou arritmias. Espera-se, assim, que a metodologia possa ser estendida a modelos mais complexos e a diferentes fenômenos de interesse.

7.1 Publicações

Parte dos resultados deste trabalho foi apresentada nas seguintes publicações científicas:

- R. B. Pigozzo, R. W. dos Santos e B. M. Rocha. *Sensitivity Analysis of a Cardiac Electrophysiology Model for the occurrence of electrical alternans*. In: Computational Science and Its Applications — ICCSA 2023 Workshops, Springer Nature Switzerland, 2023, p. 44–58. Pigozzo, Santos e Rocha (2023)
- R. B. Pigozzo, R. W. dos Santos e B. M. Rocha. *Sensitivity analysis of cardiac alternans in electrophysiology cell models: A Monte Carlo filtering approach*. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, v. 35, n. 4, 2025. Pigozzo, Santos e Rocha (2025)

7.2 Limitações e Trabalhos Futuros

Como possibilidades de trabalhos futuros, visando reduzir as limitações e contemplar estudos não realizados neste trabalho, destacam-se:

- a aplicação da abordagem a outros modelos mais complexos da eletrofisiologia cardíaca como o ToR-ORd (TOMEK et al., 2019) para células ventriculares de humanos;
- a construção de emuladores para a curva de restituição completa, e não apenas para a sua máxima inclinação;
- para modelos com muitos parâmetros, a realização de uma redução preliminar do conjunto de parâmetros a serem analisados;
- o uso de estratégias de amostragem adaptativa na construção dos emuladores, refinando as amostras nas regiões de fronteira entre os regimes, de modo a reduzir ainda mais o número de simulações do modelo completo;
- a investigação de abordagens híbridas que combinem regressores e classificadores na detecção de alternans, buscando reunir as vantagens de cada estratégia identificadas na comparação realizada.

REFERÊNCIAS

- ANTONIADIS, A.; LAMBERT-LACROIX, S.; POGGI, J.-M. Random forests for global sensitivity analysis: A selective review. *Reliability Engineering & System Safety*, Elsevier, v. 206, p. 107312, 2021.
- AULIA, A. et al. A random forests-based sensitivity analysis framework for assisted history matching. *Journal of Petroleum Science and Engineering*, Elsevier, v. 181, p. 106237, 2019.
- BARONE, A.; FENTON, F.; VENEZIANI, A. Numerical sensitivity analysis of a variational data assimilation procedure for cardiac conductivities. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, AIP Publishing, v. 27, n. 9, 2017.
- BEELEER, G.; REUTER, H. Reconstruction of the action potential of ventricular myocardial fibers. *The Journal of physiology*, n. 268, p. 177–210, 1977.
- BREIMAN, L. Random forests. *Machine learning*, Springer, v. 45, n. 1, p. 5–32, 2001.
- BREIMAN, L. et al. *Classification and regression trees*. [S.l.]: CRC press, 1984.
- BRITTON, O. J. et al. Experimentally calibrated population of models predicts and explains intersubject variability in cardiac cellular electrophysiology. *Proceedings of the National Academy of Sciences*, National Acad Sciences, v. 110, n. 23, p. E2098–E2105, 2013.
- CAMPOS, J. et al. Uncertainty quantification and sensitivity analysis of left ventricular function during the full cardiac cycle. *Philosophical Transactions of The Royal Society A Mathematical Physical and Engineering Sciences*, v. 378, 05 2020.
- CAMPOS, J. de O. *Impact of uncertainties in cardiac mechanics simulations*. Tese (Tese de Doutorado) — Universidade Federal de Juiz de Fora, Juiz de Fora, Brasil, 2019.
- CAMPOS, J. O. et al. Polynomial chaos expansion surrogate modeling of passive cardiac mechanics using the holzapfel–ogden constitutive model. *Journal of Computational Science*, Elsevier, v. 71, p. 102039, 2023.
- CASELLA, G.; BERGER, R. *Statistical inference*. [S.l.]: CRC Press, 2024.
- CIRIELLO, V. et al. Polynomial chaos expansion for global sensitivity analysis applied to a model of radionuclide migration in a randomly heterogeneous aquifer. *Stochastic Environmental Research and Risk Assessment*, Springer, v. 27, p. 945–954, 2013.
- CLUSIN, W. T. Mechanisms of calcium transient and action potential alternans in cardiac cells and tissues. *American Journal of Physiology-Heart and Circulatory Physiology*, American Physiological Society, v. 294, n. 1, p. H1–H10, 2008.
- CORRADO, C.; NIEDERER, S. A two-variable model robust to pacemaker behavior for the dynamics of the cardiac action potential. *Mathematical biosciences*, v. 281, 08 2016.
- CORSO, G. D.; VIOLA, F. Sensitivity analysis of an electrophysiology model for the left ventricle. *Journal of The Royal Society Interface*, v. 17, p. 20200532, 10 2020.

EASTMAN, J. et al. Using delay differential equations to induce alternans in a model of cardiac electrophysiology. *Journal of theoretical biology*, v. 404, p. 262–272, September 2016. ISSN 0022-5193. Disponível em: <<http://manuscript.elsevier.com/S0022519316301436/pdf/S0022519316301436.pdf>>.

ELSHRIF, M. M.; CHERRY, E. M. A quantitative comparison of the behavior of human ventricular cardiac electrophysiology models in tissue. *PLOS ONE*, Public Library of Science, v. 9, n. 1, p. 1–13, 01 2014. Disponível em: <<https://doi.org/10.1371/journal.pone.0084401>>.

FITZHUGH, R. Mathematical models of threshold phenomena in the nerve membrane. *Bulletin of Mathematical Biology*, v. 17, p. 257–278, 1955. Disponível em: <<https://api.semanticscholar.org/CorpusID:119560125>>.

FORRESTER, A.; SOBESTER, A.; KEANE, A. *Engineering design via surrogate modelling: a practical guide*. [S.l.]: John Wiley & Sons, 2008. ISBN 9780470770801.

GENESER, S. E.; KIRBY, R. M.; SACHSE, F. B. Sensitivity analysis of cardiac electrophysiological models using polynomial chaos. In: IEEE. *2005 IEEE Engineering in Medicine and Biology 27th Annual Conference*. [S.l.], 2006. p. 4042–4045.

GROENENDAAL, W. et al. Cell-specific cardiac electrophysiology models. *PLoS computational biology*, Public Library of Science San Francisco, CA USA, v. 11, n. 4, p. e1004242, 2015.

HODGKIN, A. L.; HUXLEY, A. A quantitative description of membrane current and its application to conduction and excitation in nerve. *The Journal of Physiology*, v. 117, p. 500–544, 1952.

IWANAGA, T.; USHER, W.; HERMAN, J. Toward SALib 2.0: Advancing the accessibility and interpretability of global sensitivity analyses. *Socio-Environmental Systems Modeling*, v. 4, 05 2022.

JEONG, D. U.; MARCELLINUS, A.; LIM, K. M. Sensitivity analysis of cardiac alternans and tachyarrhythmia to ion channel conductance using population modeling. *Bioengineering*, v. 9, n. 11, 2022. ISSN 2306-5354. Disponível em: <<https://www.mdpi.com/2306-5354/9/11/628>>.

KARMA, A. Electrical alternans and spiral wave breakup in cardiac tissue. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, v. 4, n. 3, p. 461–472, 09 1994. ISSN 1054-1500. Disponível em: <<https://doi.org/10.1063/1.166024>>.

KROGH-MADSEN, T.; CHRISTINI, D. Action potential duration dispersion and alternans in simulated heterogeneous cardiac tissue with a structural barrier. *Biophysical journal*, v. 92, p. 1138–49, 03 2007.

LUETHEN, N.; MARELLI, S.; SUDRET, B. Sparse polynomial chaos expansions: Literature survey and benchmark. *SIAM/ASA Journal on Uncertainty Quantification*, SIAM, v. 9, n. 2, p. 593–649, 2021.

MAHAJAN, A.; SHIFERAW, Y. A rabbit ventricular action potential model replicating cardiac dynamics at rapid heart rates. *Biophysical Journal*, 2008 by the Biophysical Society, v. 94, n. 2, p. 392–410, 2008.

- MAINDONALD, J. H. Random forests with r robin genuer & jean-michel poggi springer, 2019, x + 99 pp, £49.99, paperback. isbn: 978-3-030-56484-1. *International Statistical Review*, v. 89, n. 2, p. 422–423, 2021. Disponível em: <<https://onlinelibrary.wiley.com/doi/abs/10.1111/insr.12454>>.
- MARELLI, S.; SUDRET, B. UQLab: A framework for uncertainty quantification in matlab. In: _____. *Vulnerability, Uncertainty, and Risk*. [S.l.: s.n.]. p. 2554–2563.
- MCCLARREN, R. G. *Uncertainty Quantification and Predictive Computational Science: A Foundation for Physical Scientists and Engineers*. 1st. ed. Cham: Springer International Publishing, 2018. ISBN 978-3-319-99524-3.
- MITCHELL, C.; SCHAEFFER, D. A two-current model for the dynamics of cardiac membrane. *Bulletin of mathematical biology*, v. 65, p. 767–93, 10 2003.
- MKOKO, P. et al. Cardiac arrhythmias in low- and middle-income countries. *cardiovasc diagn ther. Cardiovascular Diagnosis and Therapy*, v. 10(2), p. 350–360, 04 2020.
- NAGUMO, J. ichi; ARIMOTO, S.; YOSHIKAWA, S. An active pulse transmission line simulating nerve axon. *Proceedings of the IRE*, v. 50, p. 2061–2070, 1962. Disponível em: <<https://api.semanticscholar.org/CorpusID:51648050>>.
- NIEDERER, S. A. et al. A meta-analysis of cardiac electrophysiology computational models. *Experimental physiology*, Wiley Online Library, v. 94, n. 5, p. 486–495, 2009.
- NOBLE, D. A modification of the hodgkin–huxley equations applicable to purkinje fibre action and pacemaker potentials. *The Journal of physiology*, Wiley, v. 160, n. 2, p. 317, 1962.
- NOLASCO, J. B.; DAHLEN, R. W. A graphic method for the study of alternans. *Journal of Applied Physiology*, v. 25, n. 2, p. 191–196, 1968.
- NOVAES, G. M. et al. An ensemble of parameters from a robust markov-based model reproduces l-type calcium currents from different human cardiac myocytes. *Plos one*, Public Library of Science San Francisco, CA USA, v. 17, n. 4, p. e0266233, 2022.
- openCARP consortium et al. *openCARP*. 2024. Disponível em: <<https://git.opencarp.org/openCARP/openCARP>>.
- PAGANI, S.; MANZONI, A. Enabling forward uncertainty quantification and sensitivity analysis in cardiac electrophysiology by reduced order modeling and machine learning. *International Journal for Numerical Methods in Biomedical Engineering*, Wiley Online Library, v. 37, n. 6, p. e3450, 2021.
- PASTORE, J. M. et al. Mechanism linking t-wave alternans to the genesis of cardiac fibrillation. *Circulation*, Am Heart Assoc, v. 99, n. 10, p. 1385–1394, 1999.
- PATHMANATHAN, P.; CORDEIRO, J. M.; GRAY, R. A. Comprehensive uncertainty quantification and sensitivity analysis for cardiac action potential models. *Frontiers in physiology*, Frontiers Media SA, v. 10, p. 721, 2019.
- PEDREGOSA, F. et al. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, v. 12, p. 2825–2830, 2011.

- PIGOZZO, R.; SANTOS, R. D.; ROCHA, B. Sensitivity analysis of cardiac alternans in electrophysiology cell models: A monte carlo filtering approach. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, AIP Publishing, v. 35, n. 4, 2025.
- PIGOZZO, R. B.; SANTOS, R. W. dos; ROCHA, B. M. Sensitivity analysis of a cardiac electrophysiology model for the occurrence of electrical alternans. In: GERVASI, O. et al. (Ed.). *Computational Science and Its Applications – ICCSA 2023 Workshops*. Cham: Springer Nature Switzerland, 2023. p. 44–58. ISBN 978-3-031-37105-9.
- PLANK, G. et al. The openCARP simulation environment for cardiac electrophysiology. *Computer Methods and Programs in Biomedicine*, v. 208, p. 106223, 2021.
- RASMUSSEN, C.; WILLIAMS, C. *Gaussian Processes for Machine Learning*. [S.l.: s.n.], 2005. ISBN 9780262256834.
- ROSENBAUM, D. et al. Electrical alternans and vulnerability to ventricular arrhythmias. *The New England Journal of Medicine*, v. 330, p. 235–41, 01 1994.
- ROSS, S. M. *A first course in probability*. [S.l.]: Macmillan New York, 1976. v. 2.
- SALTELLI, A. et al. Variance-based sensitivity analysis of model output. design and estimator for the total sensitivity index. *Computer Physics Communications*, v. 181, p. 259–270, 02 2010.
- SALTELLI, A. et al. *Global sensitivity analysis: the primer*. [S.l.]: John Wiley & Sons, 2008.
- SALTELLI, A. et al. *Sensitivity analysis in practice: a guide to assessing scientific models*. [S.l.]: Wiley Online Library, 2004. v. 1.
- SÁNCHEZ, C. et al. Sensitivity analysis of ventricular activation and electrocardiogram in tailored models of heart-failure patients. *Medical & biological engineering & computing*, Springer, v. 56, p. 491–504, 2018.
- SANTNER, T. Design and analysis of computer experiments. *AStA Advances in Statistical Analysis*, 2010.
- SHEN, C.; ZIPES, D. P. Sudden cardiac death: Mechanisms, predictive factors, and management. *Annual Review of Medicine*, Annual Reviews, v. 65, p. 439–456, 2014.
- SOBOL, I. Global sensitivity indices for nonlinear mathematical models and their monte carlo estimates. *Mathematics and Computers in Simulation*, v. 55, p. 271–280, 02 2001.
- SUDRET, B. Global sensitivity analysis using polynomial chaos expansions. *Reliability Engineering & System Safety*, v. 93, n. 7, p. 964–979, 2008. ISSN 0951-8320. Bayesian Networks in Dependability. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0951832007001329>>.
- SUDRET, B.; MAI, C. V. Computing derivative-based global sensitivity measures using polynomial chaos expansions. *Reliability Engineering & System Safety*, Elsevier, v. 134, p. 241–250, 2015.
- TOMEK, J. et al. Development, calibration, and validation of a novel human ventricular myocyte model in health, disease, and drug block. *eLife*, v. 8, 12 2019.

TUSSCHER, K. H. W. J. ten; PANFILOV, A. V. Alternans and spiral breakup in a human ventricular tissue model. *American Journal of Physiology-Heart and Circulatory Physiology*, v. 291, n. 3, p. H1088–H1100, 2006. PMID: 16565318. Disponível em: <<https://doi.org/10.1152/ajpheart.00109.2006>>.

TUSSCHER, K. ten et al. A model of human ventricular tissue. *American Journal of Physiology. Heart and circulatory physiology*, v. 286, p. H1573–89, 05 2004.

VALDEZ, A. R. et al. Uncertainty quantification and sensitivity analysis for relative permeability models of two-phase flow in porous media. *Journal of Petroleum Science and Engineering*, Elsevier, v. 192, p. 107297, 2020.

VIRTANEN, P. et al. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, v. 17, p. 261–272, 2020.

WALMSLEY, J. et al. Application of stochastic phenomenological modelling to cell-to-cell and beat-to-beat electrophysiological variability in cardiac tissue. *Journal of theoretical biology*, Elsevier, v. 365, p. 325–336, 2015.

WEISS, J. N. et al. From pulsus to pulseless: the saga of cardiac alternans. *Circulation research*, Lippincott Williams & Wilkins, v. 98, n. 10, p. 1244–1253, 2006.

WIENER, N. The homogeneous chaos. *American Journal of Mathematics*, Johns Hopkins University Press, v. 60, n. 4, p. 897–936, 1938. ISSN 00029327, 10806377. Disponível em: <<http://www.jstor.org/stable/2371268>>.

WILSON, L.; ROSENBAUM, D. Mechanisms of arrhythmogenic cardiac alternans. *Europace: European pacing, arrhythmias, and cardiac electrophysiology: journal of the working groups on cardiac pacing, arrhythmias, and cardiac cellular electrophysiology of the European Society of Cardiology*, v. 9 Suppl 6, p. vi77–82, 12 2007.

XIU, D.; KARNIADAKIS, G. E. The Wiener-Askey polynomial chaos for stochastic differential equations. *SIAM J. Sci. Comput.*, v. 24, p. 619–644, 2002. Disponível em: <<https://api.semanticscholar.org/CorpusID:264216115>>.

YOU, T. et al. Electrophysiological mechanisms underlying T-wave alternans and their role in arrhythmogenesis. *Frontiers in Physiology*, v. 12, 2021. ISSN 1664-042X.

ZANG, Y.; XIA, L. Cellular mechanism of cardiac alternans: An unresolved chicken or egg problem. *Journal of Zhejiang University. Science. B*, v. 15, p. 201–11, 03 2014.