

Benchmark de Emuladores para Simulação Cardíaca, Quantificação de Incerteza e Análise de Sensibilidade

Yan B. Werneck¹, Lucas T. Oliveira¹, Bernardo M. Rocha^{1,2},
Rafael Oliveira³, Joventino O. Campos², Rodrigo W. dos Santos^{1,2}

¹ Programa de Pós Graduação em Modelagem Computacional -
Universidade Federal de Juiz de Fora

² Departamento de Ciência da Computação - Universidade Federal de Juiz de Fora

³ Departamento de Ciência da Computação - Universidade Federal de São João Del Rei

yanbwerneck@outlook.com

Abstract. *This work presents a benchmark of three families of emulators—polynomial chaos expansions, Gaussian processes, and neural networks—applied to four representative problems in cardiac modeling involving ventricular mechanics and cellular electrophysiology. The emulators are trained to approximate the mapping from model input parameters to quantities of interest (QoIs). The models are evaluated in terms of predictive accuracy, computational cost, and fidelity in reproducing uncertainty quantification and global sensitivity analyses. Results show that the emulators achieve relative errors typically below 2% while providing speedups of up to 10^9 compared to the original models. Additionally, the results highlight qualitative differences in performance across the different families depending on problem complexity.*

Resumo. *Este trabalho apresenta um benchmark de três famílias de emuladores — expansões de caos polinomial, processos gaussianos e redes neurais — aplicados a quatro problemas representativos de modelagem cardíaca envolvendo mecânica ventricular e eletrofisiologia celular. Os emuladores são treinados para aproximar o mapeamento entre parâmetros de entrada e quantidades de interesse (QoIs), caracterizando um problema de regressão. Os modelos são avaliados em termos de precisão preditiva, custo computacional e fidelidade na reprodução de análises de quantificação de incerteza e sensibilidade global. Os resultados mostram que os emuladores alcançam erros relativos tipicamente inferiores a 2% e proporcionam acelerações de até 10^9 em relação aos modelos originais. Além disso, evidenciam diferenças qualitativas no desempenho das diferentes famílias conforme a complexidade do problema.*

1. Introdução

Doenças cardiovasculares permanecem como a principal causa de mortalidade no mundo, impondo grande impacto sobre sistemas de saúde e a sociedade. Avanços em modelagem computacional têm permitido o desenvolvimento de modelos cardíacos personalizados, ou *digital twins* cardíacos [Cluitmans et al. 2024, Jaffery et al. 2024], capazes de

reproduzir o comportamento eletrofisiológico de um paciente a partir de dados clínicos específicos. Esses modelos apresentam grande potencial para auxiliar diagnóstico e planejamento terapêutico [Bhagirath et al. 2023], incluindo a avaliação *in silico* de intervenções como ablação ou terapias farmacológicas [Rodero et al. 2023, Waight et al. 2025].

Grande parte desses modelos é baseada em descrições mecanicistas da atividade elétrica e mecânica cardíaca por meio de sistemas de equações diferenciais derivados de princípios biofísicos [Niederer et al. 2019, Trayanova et al. 2024]. A resolução desses sistemas frequentemente envolve a integração de equações rígidas e não lineares em múltiplas escalas espaciais e temporais, tornando simulações tridimensionais realistas computacionalmente custosas, podendo demandar horas ou até dias de computação mesmo em ambientes de alto desempenho [Sundnes et al. 2006, Barros et al. 2012, Augustin et al. 2021]. Esse custo torna-se particularmente limitante em aplicações que exigem múltiplas avaliações do modelo, como otimização de parâmetros, quantificação de incerteza e análise global de sensibilidade. Do ponto de vista matemático, um simulador cardíaco pode ser interpretado como um operador que mapeia parâmetros fisiológicos em quantidades de interesse (QoIs): $\mathbf{y} = \mathcal{M}(\mathbf{x})$, onde $\mathbf{x} \in R^d$ representa os parâmetros do modelo e $\mathbf{y} \in R^q$ os outputs correspondentes. Como cada avaliação de \mathcal{M} envolve a solução de modelos computacionalmente caros, como equações diferenciais ordinárias (ODEs), equações diferenciais parciais (PDEs) e simulações por elementos finitos (FEM), é comum introduzir modelos substitutos ou emuladores que aproximam essa relação por meio de funções mais baratas computacionalmente.

Historicamente, emuladores aplicados à modelagem cardíaca foram frequentemente construídos a partir de técnicas de *reduced-order modeling* (ROM), que buscam aproximar sistemas de alta dimensionalidade por representações de menor dimensão capazes de preservar suas principais dinâmicas [FitzHugh 1961]. Mais recentemente, avanços em aprendizado de máquina e modelagem estatística têm impulsionado abordagens baseadas em dados que aprendem diretamente o comportamento do sistema a partir de dados de simulação [Gherman et al. 2023, Kudela et al. 2022]. Nesse contexto, um emulador pode ser formalmente representado como $\hat{\mathbf{y}} = \mathcal{S}(\mathbf{x}; \boldsymbol{\theta})$, onde $\boldsymbol{\theta}$ representa os parâmetros do emulador. O treinamento utiliza um conjunto de dados $\{\mathbf{x}_{\text{train}}^i, \mathbf{y}_{\text{train}}^i\}_{i=1}^{N_S}$ obtido a partir de avaliações do modelo completo, para otimizar $\boldsymbol{\theta}$. Uma vez treinado, o emulador \mathcal{S} permite aproximar rapidamente a resposta do modelo, viabilizando análises que seriam impraticáveis utilizando apenas o simulador original. Entre as abordagens baseadas em dados mais utilizadas destacam-se as Expansões de Caos Polinomial (PCE), os Processos Gaussianos (GP) e as Redes Neurais (NN). Comparações sistemáticas entre essas técnicas têm sido amplamente investigadas em áreas como dinâmica dos fluidos, mecânica estrutural e modelagem ambiental [Liu et al. 2020, Tsokanas et al. 2022, Shahzadi et al. 2021], mas permanecem relativamente limitadas no contexto da modelagem cardíaca.

Neste trabalho apresentamos um benchmark sistemático dessas três abordagens de emuladores aplicadas a quatro problemas representativos de modelagem cardíaca envolvendo eletrofisiologia e mecânica ventricular. Os problemas considerados cobrem dois dos principais componentes da dinâmica cardíaca, amplamente utilizados na construção de digital twins, e abrangem diferentes níveis de complexidade e dimensionalidade. Os emuladores são treinados para aproximar a relação entre parâmetros dos modelos cardíacos e as saídas simuladas correspondentes. Cada abordagem é avaliada em ter-

mos de precisão preditiva, custo computacional e escalabilidade, bem como pelo impacto dessas aproximações em análises de quantificação de incerteza, sensibilidade global e problemas inversos.

2. Métodos

2.1. Modelos cardíacos de benchmark

Consideramos dois problemas de benchmark em modelagem cardíaca representando domínios fisiológicos distintos: mecânica ventricular e eletrofisiologia celular.

Problema mecânico: enchimento passivo do ventrículo esquerdo Usamos um modelo da resposta passiva do ventrículo esquerdo (VE) durante o enchimento, descrita pelo modelo constitutivo de Holzapfel–Ogden (HO). Consideramos duas formulações com diferentes níveis de anisotropia: A formulação transversalmente isotrópica, denominada Problema A_M , considera apenas a contribuição das fibras e resulta em um modelo com quatro parâmetros $\{a, b, a_f, b_f\}$, com energia de deformação

$$\Psi = \frac{a}{2b} \left(e^{b(I_1-3)} - 1 \right) + \frac{a_f}{2b_f} \left(e^{b_f((I_{4f}-1))^2} - 1 \right), \quad (1)$$

onde $I_1 = \text{tr}(\mathbf{C})$ mede a distorção volumétrica e $I_{4f} = \mathbf{f}_0 \cdot (\mathbf{C}\mathbf{f}_0)$ representa o alongamento na direção das fibras \mathbf{f}_0 . E a formulação anisotrópica completa, denominada Problema B_M , que inclui também interações entre folhas e entre fibras e folhas, introduzindo os invariantes $I_{4s} = \mathbf{s}_0 \cdot (\mathbf{C}\mathbf{s}_0)$ e $I_{8fs} = \mathbf{f}_0 \cdot (\mathbf{C}\mathbf{s}_0)$. A energia de deformação correspondente é

$$\Psi = \frac{a}{2b} e^{b(I_1-3)} + \sum_{i \in \{f,s\}} \frac{a_i}{2b_i} \left(e^{b_i(\max(I_{4i},1)-1)^2} - 1 \right) + \frac{a_{fs}}{2b_{fs}} \left(e^{b_{fs}I_{8fs}^2} - 1 \right), \quad (2)$$

resultando em uma formulação com oito parâmetros $\{a, b, a_f, b_f, a_s, b_s, a_{fs}, b_{fs}\}$.

As simulações por elementos finitos seguem o protocolo de [Campos et al. 2023], inflando o VE até 2.7 kPa a partir de um volume inicial de aproximadamente 50 mL. De cada simulação são extraídas seis QoIs que caracterizam a resposta mecânica passiva: dois pares de coeficientes do tipo Klotz obtidos a partir dos ajustes exponenciais das curvas pressão–volume e pressão–deformação, (α_1, β_1) e (α_2, β_2) , além do volume diastólico final (EDV) e do alongamento médio das fibras λ_f . A Figura 1 demonstra essas QoIs. Os dados de treinamento são gerados amostrando cada parâmetro dentro de $\pm 30\%$ dos valores de referência e resolvendo o modelo completo de elementos finitos para cada realização.

Problema eletrofisiológico: Modelo de Potencial de Ação Ventricular (AP) Também consideramos duas parametrizações do modelo de miócito ventricular humano de ten Tusscher e Panfilov de 2004 (TP) [Ten Tusscher et al. 2004] com níveis crescentes de complexidade. A primeira, Problema A_E , corresponde a um modelo reduzido focado em condições isquêmicas, contendo três parâmetros de entrada que modulam cinco processos iônicos do modelo TP: a condutância de sódio (G_{Na}), a condutância de cálcio tipo L (G_{CaL}), as concentrações intracelular e extracelular de potássio (K_i, K_o) e a concentração

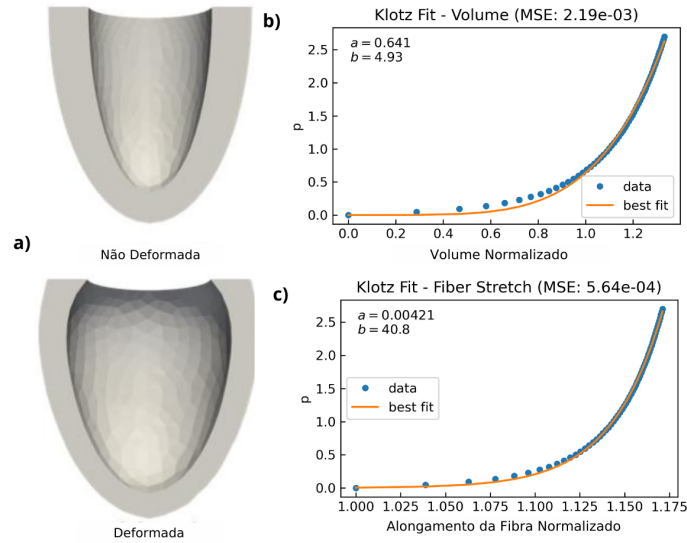


Figura 1. Quantidades de interesse dos problemas A_M e B_M . O volume diastólico final e o alongamento médio das fibras λ_f são calculados a partir da geometria deformada (a). Os parâmetros (α_1, β_1) são obtidos pelo ajuste exponencial da curva pressão–volume (b), enquanto (α_2, β_2) são obtidos a partir da relação pressão–deformação (c).

de ATP. Esses parâmetros capturam os principais mecanismos associados às alterações eletrofisiológicas induzidas por isquemia, conforme descrito em [Lawson et al. 2020].

A segunda, Problema B_E , configuração considera doze parâmetros de entrada correspondentes a coeficientes de escala aplicados às condutâncias do modelo TP [Ten Tusscher et al. 2004], representando variabilidade populacional. Além dos cinco parâmetros anteriores, incluem-se sete condutâncias adicionais: corrente transiente de potássio (G_{to}), correntes retificadoras tardias rápida e lenta (G_{Kr} , G_{Ks}), corrente retificadora interna de potássio (G_{K1}), condutância de cálcio de fundo (G_{bCa}), e correntes de bomba de cálcio e potássio (G_{pCa} , G_{pK}).

Seis QoIs são extraídas das formas de onda dos potenciais de ação simulados (Fig. 2) para caracterizar o comportamento eletrofisiológico celular: o potencial de repouso da membrana (V_{rest}), o pico do potencial transmembrana (V_{peak}), a velocidade máxima de despolarização (dV/dt_{max}) e as durações do potencial de ação em 30%, 50% e 80% de repolarização (APD_{30} , APD_{50} e APD_{80}). Todas as medições são realizadas após 20 ciclos de estimulação para garantir regime permanente. Os dados de treinamento são gerados amostrando uniformemente cada parâmetro de entrada, entre 0 e 1 para A_E e 0.75 e 1.25 para B_E , e resolvendo o modelo TP utilizando o método de Euler com passo temporal $dt = 0.01$.

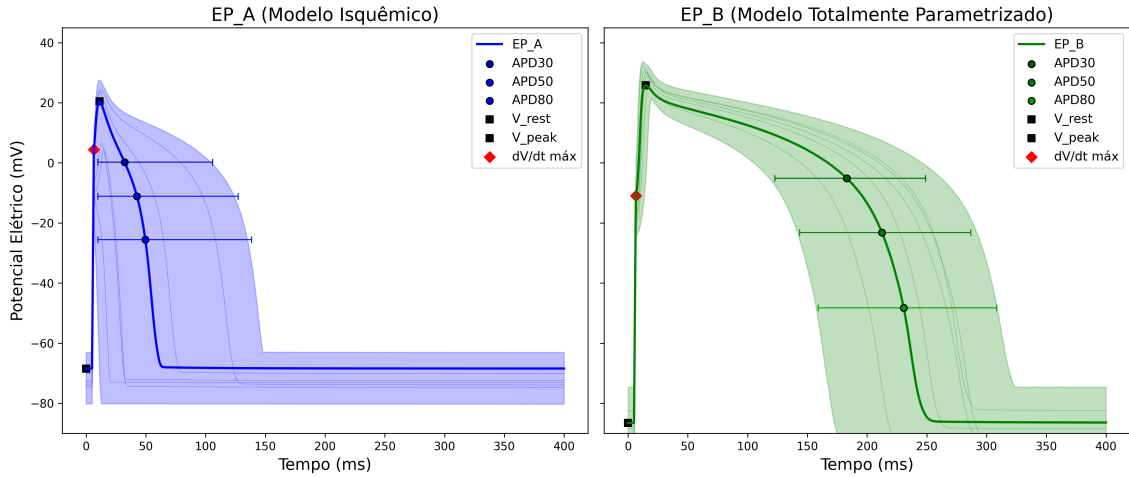


Figura 2. Potenciais de ação simulados para dois modelos eletrofisiológicos. A região sombreada mostra o envelope (min–max) das amostras de parâmetros, enquanto curvas tênues representam trajetórias individuais. A curva destacada é a mais próxima da média do conjunto. Quadrados pretos indicam os potenciais de repouso e de pico (V_{rest} , V_{peak}), o losango vermelho marca a velocidade máxima de subida ($\frac{dV}{dt}_{max}$), e marcadores coloridos indicam APD30, APD50 e APD80, com barras mostrando sua distribuição no espaço de parâmetros.

2.2. Emuladores e treinamento

Três famílias de emuladores foram investigadas: Expansões de Caos Polinomial, Processos Gaussianos e Redes Neurais. No caso das PCE, o emulador é representado por

$$\hat{y}(\mathbf{x}) = \sum_{i=0}^{N_P} c_i \Phi_i(\mathbf{x}),$$

onde $\Phi_i(\mathbf{x})$ são polinômios ortogonais multivariados e c_i os coeficientes estimados por regressão. Utilizamos bases de polinômios de Legendre com graus $k = \{2, 3, 5\}$ implementadas na biblioteca ChaosPy [Feinberg and Pinkus 2018].

Os modelos GP definem um processo estocástico sobre funções cuja predição média posterior é dada por

$$\hat{y}(\mathbf{x}) = m(\mathbf{x}) + k(\mathbf{x}, X)K^{-1}(y - m(X)),$$

onde $m(\mathbf{x})$ é a função média e $k(\mathbf{x}, X)$ o kernel de covariância avaliado entre o ponto de teste e os dados de treinamento. Os modelos foram implementados em GPyTorch [Gardner et al. 2018] com três configurações: GP_S com média linear e kernel RBF, GP_M com média linear e kernel composto RBF–Matérn ($\nu = 2.5$), e GP_L com média constante e kernel composto RBF–Matérn ($\nu = 1.5$).

Para redes neurais, o emulador é modelado como uma função paramétrica

$$\hat{y}(\mathbf{x}) = f_{\theta}(\mathbf{x}),$$

onde f_{θ} corresponde a uma série de camadas sucessivas, cada uma delas composta por uma transformação linear e uma função de ativação não linear. As redes foram implementadas em PyTorch [Paszke et al. 2019] com três arquiteturas: NN_S (1 camada oculta com

16 neurônios), NN_M (2 camadas ocultas com 32 neurônios) e NN_L (4 camadas ocultas com 64 neurônios), utilizando funções de ativação SiLU.

Os conjuntos de treinamento $\{\mathbf{x}_{\text{train}}^i, \mathbf{y}_{\text{train}}^i\}_{i=1}^{N_S}$ foram gerados por amostragem Latin Hypercube, considerando tamanhos $N_S = \{100, 200, 500, 1000\}$ para todos os problemas. Um conjunto independente de validação com 5000 amostras foi utilizado para avaliar a precisão preditiva. Os modelos NN e GP foram treinados com o otimizador Adam [Kingma and Ba 2015], utilizando redução adaptativa da taxa de aprendizado e *early stopping* baseado na perda de validação. Durante a inferência, NN e GP realizam avaliações em lote utilizando GPU via PyTorch, enquanto as PCE são avaliadas em lote em CPU utilizando ChaosPy e multi-threading.

2.3. Métricas de avaliação

O desempenho dos emuladores foi avaliado segundo três critérios: (i) precisão preditiva na aproximação do mapeamento do modelo completo, (ii) custo computacional de treinamento e inferência, e (iii) fidelidade na reprodução dos resultados de quantificação de incerteza (UQ) e análise global de sensibilidade (SA).

Precisão preditiva A precisão foi quantificada pelo erro absoluto relativo médio (MARE),

$$\text{MARE} = \frac{1}{N_q N_v} \sum_{j=1}^{N_q} \sum_{i=1}^{N_v} \left| \frac{\hat{y}_i^{(j)} - y_i^{(j)}}{y_i^{(j)}} \right|,$$

onde $y_i^{(j)}$ representa o valor da QoI j produzido pelo modelo completo para a amostra de validação i , $\hat{y}_i^{(j)}$ a predição do emulador, N_q o número de QoIs e N_v o número de amostras de validação.

Quantificação de incerteza e análise de sensibilidade Após o treinamento, os emuladores foram utilizados para realizar quantificação de incerteza (UQ) e análise global de sensibilidade (SA). Para UQ, um conjunto de 10 000 amostras de parâmetros foi gerado por Latin Hypercube Sampling [McKay et al. 1979], e os emuladores foram avaliados para estimar a média e o desvio padrão de cada QoI.

Para SA, foram calculados índices de sensibilidade de Sobol [Sobol' 2001] utilizando os estimadores de Monte Carlo propostos por Saltelli [Saltelli et al. 2008]. Os índices de primeira ordem e de efeito total requerem um número total de avaliações dado por

$$N_{\text{total}} = N_{\text{base}}(d + 2),$$

onde d é o número de parâmetros do modelo e N_{base} o tamanho da amostra base (utilizamos $N_{\text{base}} = 128$). A fidelidade dos emuladores na análise de sensibilidade foi quantificada pelo erro absoluto médio (MAE) entre os índices de Sobol estimados a partir das avaliações do emulador e aqueles obtidos com o modelo de referência,

$$\text{MAE}_S = \frac{1}{N_q N_p} \sum_{j=1}^{N_q} \sum_{i=1}^{N_p} \left| \hat{S}_i^{(j)} - S_i^{(j)} \right|,$$

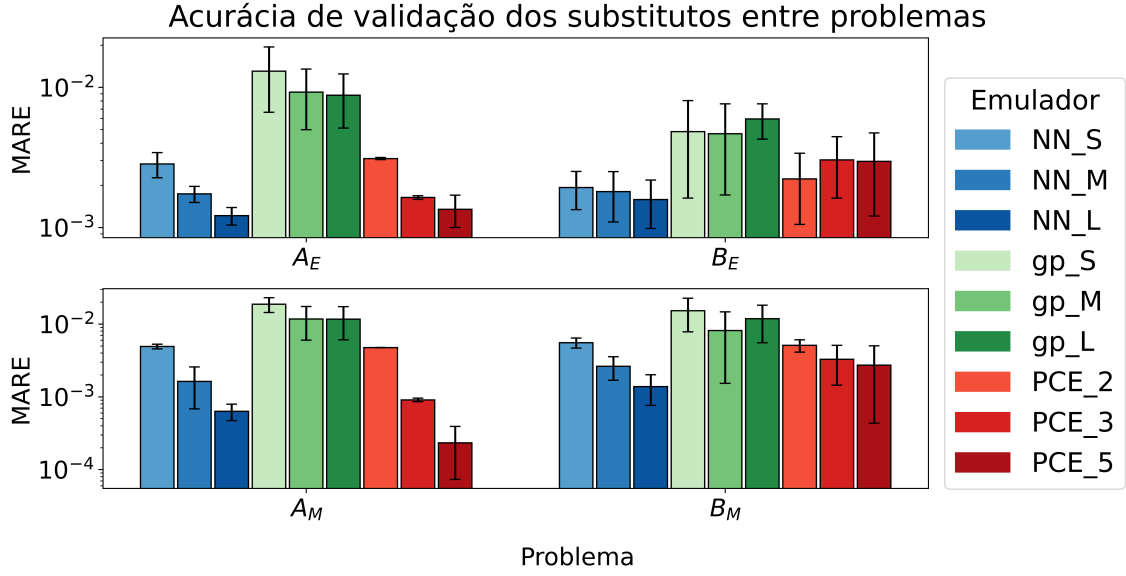


Figura 3. MARE médio de validação para todos os emuladores nos quatro problemas de benchmark. Barras representam o erro médio e as linhas indicam a variação obtida com diferentes tamanhos de conjunto de treinamento.

onde $S_i^{(j)}$ representa o índice de Sobol do parâmetro i para a QoI j calculado com o modelo completo, $\hat{S}_i^{(j)}$ o valor correspondente obtido com o emulador, N_q o número de QoIs e N_p o número de parâmetros.

Eficiência computacional O custo computacional foi avaliado em termos de tempo de treinamento e inferência. A inferência foi medida como o tempo necessário para avaliar um conjunto representativo de 10^5 amostras, típico de estudos de Monte Carlo. O ganho computacional foi quantificado pelo fator de aceleração

$$S = \frac{T_{\text{full}}(N_s)}{T_{\text{sur}}(N_s)},$$

onde $T_{\text{full}}(N_s)$ representa o tempo de execução do modelo completo para N_s avaliações e $T_{\text{sur}}(N_s)$ o tempo correspondente do emulador.

3. Resultados

Precisão preditiva A Figura 3 resume o erro médio relativo de validação (MARE) para todos os emuladores nos quatro problemas de benchmark. Redes neurais apresentam os menores erros e maior robustez frente ao aumento da dimensionalidade e complexidade dos problemas, mantendo erros tipicamente inferiores a 1%. PCE obtêm bom desempenho nos casos de menor dimensionalidade (A_M e A_E), mas sofrem degradação significativa à medida que o número de parâmetros aumenta, refletindo o crescimento combinatório da base polinomial. Processos gaussianos apresentam precisão média inferior às redes neurais, porém mostram comportamento mais estável com relação à dimensionalidade, sendo sua precisão principalmente limitada pela quantidade de dados de treinamento.

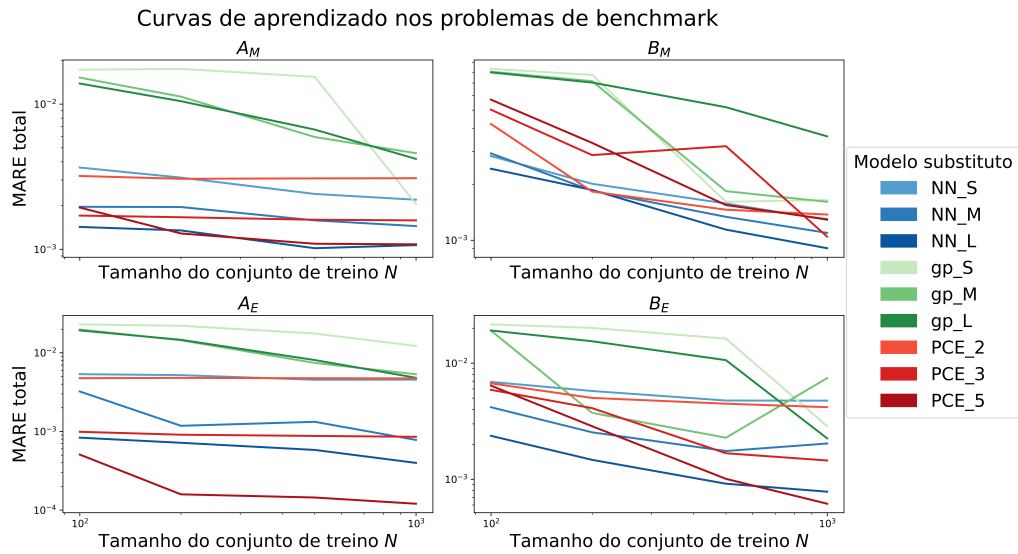


Figura 4. Curvas de aprendizado mostrando o MARE (média das QOIs) em função do tamanho do conjunto de treinamento para cada um dos problemas e emuladores.

As curvas de aprendizado da Figura 4, evidenciam maior dependência de dados nos problemas de maior dimensionalidade (B_M e B_E). PCE apresentam saturação rápida em problemas de baixa dimensão, indicando a limitação da capacidade representacional da base polinomial. Em contraste, GP e NN continuam melhorando com o aumento do número de amostras, especialmente em problemas mais complexos.

Custo computacional A Figura 5 compara o custo computacional de treinamento e inferência dos diferentes emuladores. PCE apresentam forte dependência da dimensionalidade, com aumento significativo de custo em problemas de maior dimensão devido ao crescimento do número de termos da expansão polinomial. Processos gaussianos possuem o maior custo de treinamento, dominado pela construção e fatoração da matriz de covariância, e são também os únicos cuja inferência depende diretamente do tamanho do conjunto de treinamento. Redes neurais apresentam comportamento mais estável, com custo de treinamento moderado e inferência extremamente rápida graças à paralelização em GPU. Nos problemas eletrofisiológicos, as acelerações atingem até 10^9 , enquanto nos problemas mecânicos chegam a aproximadamente 10^6 .

UQ e análise de sensibilidade A Figura 6 mostra a relação entre a precisão preditiva do emulador e o erro na estimativa dos índices de Sobol. Observa-se uma relação aproximadamente linear entre erro de predição e erro na análise de sensibilidade, indicando que imprecisões na aproximação do emulador podem propagar-se na estimativa dos índices. Apesar de apresentarem menor precisão média, os modelos GP mostram desempenho relativamente melhor em SA, sugerindo menor propagação de erros nas estimativas de Sobol. Para UQ, essa dependência é muito menos pronunciada. Mesmo emuladores com erros de predição moderados conseguem estimar corretamente média e desvio padrão das QOIs, indicando maior robustez das análises de propagação de incerteza.

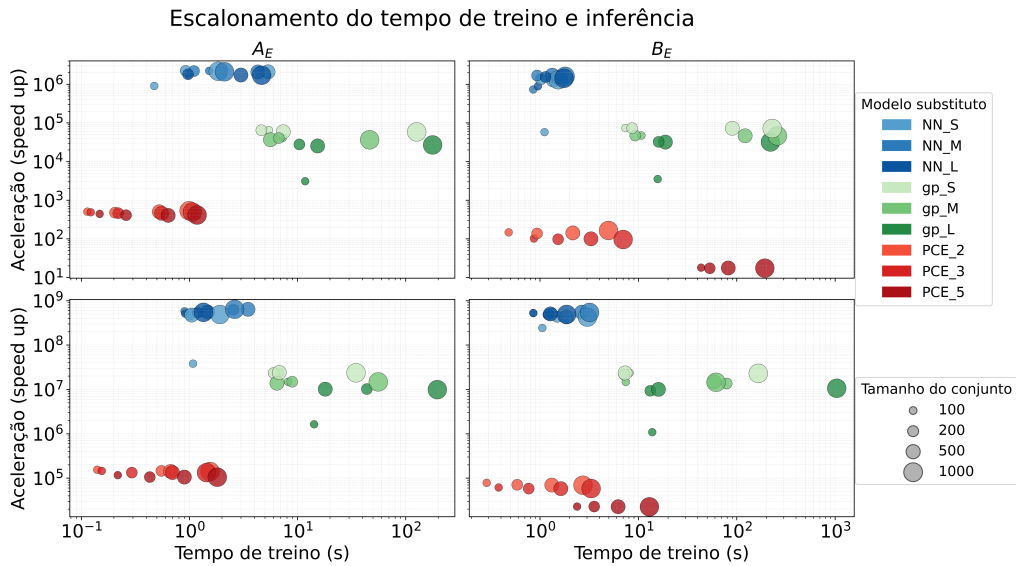


Figura 5. Tempo de treinamento dos emuladores (esq.) e speedup de inferência em relação aos modelos completos (dir.), avaliados em 100k amostras. O tamanho dos marcadores indica o conjunto de treino (100, 500, 1000, 5000). Como referência, problemas mecânicos requerem ~ 2 min por lote de 4 simulações, enquanto eletrofisiologia requer ~ 5 min por lote de 4000 simulações na mesma máquina, correspondendo à latência base e ao throughput máximo de cada caso.

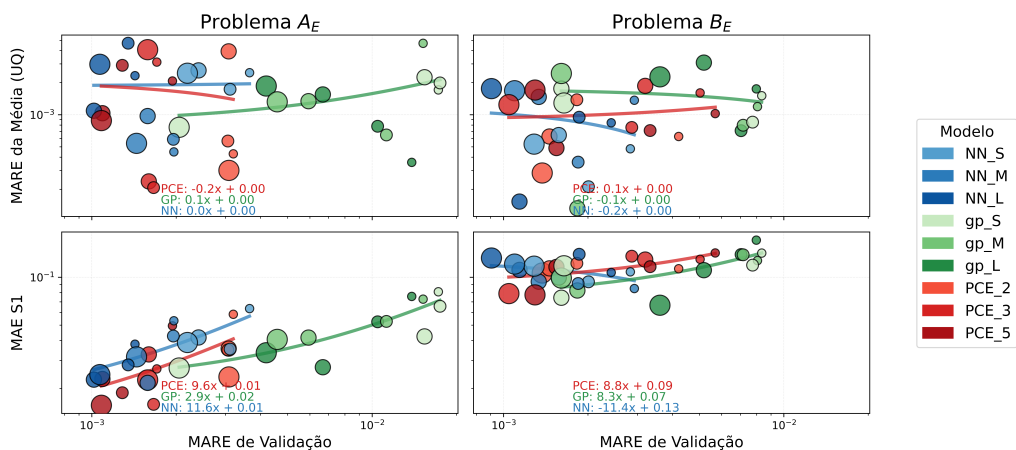


Figura 6. Erro na análise de sensibilidade em função do erro preditivo dos emuladores (superior). Erro nas estatísticas de UQ (média e desvio padrão) em função do erro preditivo dos emuladores (inferior).

4. Conclusão

Este trabalho apresentou um benchmark sistemático das três famílias de emuladores relevantes, aplicados a problemas representativos de modelagem cardíaca envolvendo mecânica ventricular e eletrofisiologia. Os resultados mostram que os emuladores podem reduzir o custo computacional dos modelos completos em várias ordens de magnitude, mantendo boa acurácia relativa.

Redes neurais oferecem o melhor compromisso entre precisão e eficiência computacional, apresentando erros preditivos inferiores a 2%, com fatores de aceleração de até $10^9 \times$ em relação ao solver completo, reduzindo a latência de minutos para milissegundos. As expansões PCE demonstraram bom desempenho em problemas de baixa dimensionalidade, porém sua escalabilidade é limitada devido ao crescimento combinatorial da base polinomial. Processos gaussianos apresentaram desempenho mais estável com o aumento da dimensionalidade, embora com maior custo computacional associado ao treinamento e à inferência. As análises de quantificação de incerteza e sensibilidade indicam que erros de aproximação do emulador podem impactar significativamente os índices de Sobol, enquanto estatísticas globais de incerteza mostraram maior robustez, especialmente no caso dos processos gaussianos, que modelam explicitamente a distribuição preditiva.

Em conjunto, esses resultados evidenciam as vantagens específicas de cada abordagem de emulador e indicam que o aumento da complexidade e da dimensionalidade dos problemas tende a favorecer o uso de redes neurais. Dado que os resultados demonstram ganhos significativos de velocidade e utilidade em tarefas como quantificação de incerteza e problemas inversos, um próximo passo natural é investigar a integração desses emuladores em pipelines de *digital twins* voltados a aplicações clínicas. Além disso, outra direção promissora é explorar emuladores com dados clínicos reais, complementando ou, em alguns casos, reduzindo a dependência de modelos mecanísticos como intermediários no treinamento, ao aprender diretamente a relação entre dados observados e desfechos de interesse.

Agradecimentos

Os autores agradecem à Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG), à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), à Empresa Brasileira de Serviços Hospitalares (Ebserh), ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), ao SINAPAD Santos-Dumont e à Universidade Federal de Juiz de Fora (UFJF) pelo apoio financeiro a este trabalho.

Referências

- Augustin, C. M., Gsell, M. A., Karabelas, E., Willemen, E., Prinzen, F. W., Lumens, J., Vigmond, E. J., and Plank, G. (2021). A computationally efficient physiologically comprehensive 3d–0d closed-loop model of the heart and circulation. *Computer Methods in Applied Mechanics and Engineering*, 386.
- Barros, B. G. D., Oliveira, R. S., Meira, W., Lobosco, M., and Santos, R. W. D. (2012). Simulations of complex and microscopic models of cardiac electrophysiology powered by multi-gpu platforms. *Computational and Mathematical Methods in Medicine*, 2012.

- Bhagirath, P., Strocchi, M., Bishop, M., Boyle, P., and Plank, G. (2023). From bits to bedside: entering the age of digital twins in cardiac electrophysiology. *Europace*, 26(12):euae295.
- Campos, J., Guedes, R., Werneck, Y., Barra, L., dos Santos, R., and Rocha, B. (2023). Polynomial chaos expansion surrogate modeling of passive cardiac mechanics using the holzapfel–ogden constitutive model. *Journal of Computational Science*, 71:102039.
- Cluitmans, M. J. M., Plank, G., and Heijman, J. (2024). Digital twins for cardiac electrophysiology: state of the art and future challenges. *Herzschrittmachertherapie & Elektrophysiologie*, 35(2):118–123.
- Feinberg, J. and Pinkus, A. (2018). Chaospy: An open source tool for designing methods of uncertainty quantification. *Journal of Computational Science*, 11:46–57.
- FitzHugh, R. (1961). Impulses and physiological states in theoretical models of nerve membrane. *Biophysical Journal*, 1(6):445–466.
- Gardner, J. R., Pleiss, G., Bindel, D., Weinberger, K. Q., and Wilson, A. G. (2018). Gpytorch: Blackbox matrix-matrix gaussian process inference with gpu acceleration. *Advances in Neural Information Processing Systems (NeurIPS)*, 31.
- Gherman, I. M., Abdallah, Z. S., Pang, W., Goroehowski, T. E., Grierson, C. S., and Marucci, L. (2023). Bridging the gap between mechanistic biological models and machine learning surrogates.
- Jaffery, O. A., Melki, L., Slabaugh, G., Good, W. W., and Roney, C. H. (2024). A review of personalised cardiac computational modelling using electroanatomical mapping data. *Arrhythmia & Electrophysiology Review*, 13:e08. Open Access article, PMID 38807744, PMC11131150.
- Kingma, D. P. and Ba, J. (2015). Adam: A method for stochastic optimization. In *International Conference on Learning Representations (ICLR)*.
- Kudela, J., Škřivánková, I., Zahradník, P., et al. (2022). Recent advances and applications of surrogate models for finite element method-based computations: a literature review. *Soft Computing*, 27:1–27. Large survey (180 papers) covering GP, PCE, NN and hybrid approaches; discusses trends, gaps and practical recommendations.
- Lawson, B. A. J., Oliveira, R. S., Berg, L. A., Silva, P. A. A., Burrage, K., and dos Santos, R. W. (2020). Variability in electrophysiological properties and conducting obstacles controls re-entry risk in heterogeneous ischaemic tissue. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 378(2173):20190341.
- Liu, Z., Zhan, L., Pan, R., et al. (2020). Surrogate modelling based on resampled polynomial chaos expansion (rpce): methodology and applications. *Expert Systems with Applications*, 153:113409. Improves robustness of PCE selection procedures; useful when comparing practical PCE pipelines to ML surrogates.
- McKay, M. D., Beckman, R. J., and Conover, W. J. (1979). A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, 21(2):239–245.

- Niederer, S. A., Lumens, J., and Trayanova, N. A. (2019). Computational models in cardiology.
- Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., Desmaison, A., Kopf, A., Yang, E., DeVito, Z., Raison, M., Tejani, A., Chilamkurthy, S., Steiner, B., Fang, L., Bai, J., and Chintala, S. (2019). Pytorch: An imperative style, high-performance deep learning library. *Advances in Neural Information Processing Systems (NeurIPS)*, 32.
- Rodero, C., Baptiste, T. M. G., Barrows, R. K., Keramati, H., Sillett, C. P., Strocchi, M., Lamata, P., and Niederer, S. A. (2023). A systematic review of cardiac in-silico clinical trials. *Progress in Biomedical Engineering*, 5(3):032004. Published by IOP Publishing Ltd.
- Saltelli, A., Chan, K., and Scott, E. M. (2008). *Sensitivity Analysis*. John Wiley & Sons, Chichester, UK.
- Shahzadi, G., Ahmad, W., Nawab, K., et al. (2021). Deep neural network and polynomial chaos expansion based surrogate models for computationally intensive environmental models. *Water*, 13(13):1830. Empirical comparison of deep NN and PCE surrogates for environmental model emulation; highlights data requirements and generalization trade-offs.
- Sobol', I. M. (2001). Global sensitivity indices for nonlinear mathematical models and their monte carlo estimates. *Mathematics and Computers in Simulation*, 55(1-3):271–280.
- Sundnes, J., Lines, G. T., Cai, X., Nielsen, B. F., Mardal, K. A., and Tveito, A. (2006). *Computing the Electrical Activity in the Heart*. Springer.
- Ten Tusscher, K. H. W. J., Noble, D., Noble, P. J., and Panfilov, A. V. (2004). A model for human ventricular tissue. *American Journal of Physiology-Heart and Circulatory Physiology*, 286(4):H1573–H1589. PMID: 14656705.
- Trayanova, N. A., Lyon, A., Shade, J., and Heijman, J. (2024). Computational modeling of cardiac electrophysiology and arrhythmogenesis: toward clinical translation. *Physiological Reviews*, 104(3):1265–1333. Epub 2023 Dec 28.
- Tsokanas, N., Simpson, T., Pastorino, R., Chatzi, E., and Stojadinović, B. (2022). Model order reduction for real-time hybrid simulation: Comparing polynomial chaos expansion and neural network methods. *Mechanics and Machine Theory*, 166:104178. Direct comparative study of PCE and feedforward neural networks for real-time model-reduction in hybrid simulation.
- Waight, M. C., Prakosa, A., Li, A. C., Bunce, N., Marciniak, A., Trayanova, N. A., and Saba, M. M. (2025). Heart digital twins predict features of invasive reentrant circuits and ablation lesions in scar-dependent ventricular tachycardia. *Circulation: Arrhythmia and Electrophysiology*, 18. 18: (2025), forthcoming.