Ajuste Evolutivo de um Modelo Baseado em Autômatos Celulares para Representar a Dinâmica de Propagação de Incêndios Florestais

Lucas Vieira Murilo



Universidade Federal de Uberlândia
Faculdade de Computação
Programa de Pós-Graduação em Ciência da Computação

Uberlândia / MG
2025

Lucas Vieira Murilo

Ajuste Evolutivo de um Modelo Baseado em Autômatos Celulares para Representar a Dinâmica de Propagação de Incêndios Florestais

Dissertação de mestrado apresentada ao Programa de Pós-graduação da Faculdade de Computação da Universidade Federal de Uberlândia como parte dos requisitos para a obtenção do título de Mestre em Ciência da Computação.

Área de concentração: Inteligência Artificial

Orientador: Luiz Gustavo Almeida Martins
Coorientador: Gina Maira Barbosa de Oliveira

Uberlândia / MG
2025

Ficha Catalográfica Online do Sistema de Bibliotecas da UFU
com dados informados pelo(a) próprio(a) autor(a).

M977 2025

Murilo, Lucas Vieira, 1995-
Ajuste Evolutivo de um Modelo Baseado em Autômatos Celulares para Representar a Dinâmica de Propagação de Incêndios Florestais [recurso eletrônico] / Lucas Vieira Murilo. - 2025.

Orientador: Luiz Gustavo Almeida Martins.
Coorientadora: Gina Maira Barbosa de Oliveira.
Dissertação (Mestrado) - Universidade Federal de Uberlândia, Pós-graduação em Ciência da Computação.
Modo de acesso: Internet.
Disponível em: http://doi.org/10.14393/ufu.di.2025.21
Inclui bibliografia.
Inclui ilustrações.

1. Computação. I. Martins, Luiz Gustavo Almeida ,1974-, (Orient.). II. Oliveira, Gina Maira Barbosa de ,1967-, (Coorient.). III. Universidade Federal de Uberlândia. Pós-graduação em Ciência da Computação. IV. Título.

CDU: 681.3

Bibliotecários responsáveis pela estrutura de acordo com o AACR2:

Gizele Cristine Nunes do Couto - CRB6/2091
Nelson Marcos Ferreira - CRB6/3074

UNIVERSIDADE FEDERAL DE UBERLÂNDIA
Coordenação do Programa de Pós-Graduação em Computação
Av. João Naves de Ávila, 2121, Bloco 1A, Sala 243 - Bairro Santa Mônica, Uberlândia-MG, CEP 38400-902
Telefone: (34) 3239-4470 - www.ppgco.facom.ufu.br - cpgfacom@ufu.br

ATA DE DEFESA - PÓS-GRADUAÇÃO

Programa de Pós-Graduação em:	Ciência da Computação				
Defesa de:	Dissertação, 55/2025, PPGCO				
Data:	31 de janeiro de 2025	Hora de início:	9:05	Hora de encerramento:	11:27
Matrícula do Discente:	12122CCP011				
Nome do Discente:	Lucas Vieira Murilo				
Título do Trabalho:	Ajuste Evolutivo de um Modelo Baseado em Autômatos Celulares para Representar a Dinâmica de Propagação de Incêndios Florestais				
Área de concentração:	Ciência da Computação				
Linha de pesquisa:	Inteligência Artificial				
Projeto de Pesquisa de vinculação:	Cerrado Resiliente: Computação Bio-inspirada e Monitoramento por VANTs na Prevenção e Combate a Incêndios Florestais e Urbanos.				

Reuniu-se por videoconferência, a Banca Examinadora, designada pelo Colegiado do Programa de Pós-graduação em Ciência da Computação, assim composta: Professores Doutores: Claudiney Ramos Tinoco - FACOM/UFU, Renato Tinós - FFCLRP/USP e Luiz Gustavo Almeida Martins - FACOM/UFU, orientador do candidato.

Os examinadores participaram desde as seguintes localidades: Renato Tinós de Ribeirão Preto/SP e os outros membros da banca da cidade de Uberlândia/MG. O aluno Lucas Vieira Murilo participou de Rio Claro/SP.

Iniciando os trabalhos, o presidente da mesa, Prof. Dr. Luiz Gustavo Almeida Martins, apresentou a Comissão Examinadora e o candidato, agradeceu a presença do público, e concedeu ao discente a palavra para a exposição do seu trabalho. A duração da apresentação da discente e o tempo de arguição e resposta foram conforme as normas do Programa.

A seguir o senhor presidente concedeu a palavra, pela ordem sucessivamente, aos examinadores, que passaram a arguir ao candidato. Ultimada a arguição, que se desenvolveu dentro dos termos regimentais, a Banca, em sessão secreta, atribuiu o resultado final, considerando o candidato:

Aprovado

Esta defesa faz parte dos requisitos necessários à obtenção do título de Mestre.

O competente diploma será expedido após cumprimento dos demais requisitos, conforme as normas do Programa, a legislação pertinente e a regulamentação interna da UFU.

Nada mais havendo a tratar foram encerrados os trabalhos. Foi lavrada a presente ata que após lida e achada conforme foi assinada pela Banca Examinadora.

Documento assinado eletronicamente por **Luiz Gustavo Almeida Martins**, **Professor(a) do Magistério Superior**, em 31/01/2025, às 13:36, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do Decreto nº 8.539, de 8 de outubro de 2015.

Documento assinado eletronicamente por **Claudiney Ramos Tinoco**, **Professor(a) do Magistério Superior**, em 31/01/2025, às 14:21, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do Decreto nº 8.539, de 8 de outubro de 2015.

Documento assinado eletronicamente por **Renato Tinós**, **Usuário Externo**, em 31/01/2025, às 14:46, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do Decreto nº 8.539, de 8 de outubro de 2015.

A autenticidade deste documento pode ser conferida no site https://www.sei.ufu.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **5989048** e o código CRC **82BA6E27**.

Referência: Processo nº 23117.089537/2024-84 SEI nº 5989048

Este trabalho é dedicado à minha esposa Lisiane aos meus pais e ao meu irmão que me incentivaram a sempre seguir em frente e acreditar em meus sonhos.

Agradecimentos

A Deus pelo dom da vida e por todas as bençãos que tem derramado diariamente.

Ao meu orientador professor Dr. Luiz Gustavo Almeida Martins pelas contribuições ao longo desta pesquisa, pelo incentivo e suporte em todos os momentos mesmo nos que achei que não seria possível chegar até aqui.

A minha co-orientadora professora Dra. Gina Maira Barbosa de Oliveira (*in memorian*), pelas valorosas contribuições, pela dedicação e pelos ensinamentos.

Aos meus parceiros do LCBio: Samuel, Claudiney, à Débora e, em especial, à Maria Eugênia pelo apoio na reprodução do seu trabalho.

À Universidade Estadual Paulista (Unesp) e ao professor Milton por cederem o laboratório para finalização dos experimentos.

À FACOM (Faculdade de Computação) e ao CNPq (Processo 423105/2021-3) pelo apoio financeiro durante o desenvolvimento desta pesquisa.

Aos professores do Programa de Pós-Graduação da Faculdade de Computação (PPGCO) por contribuírem significativamente para à minha formação. À Secretária, pela presteza e dedicação nos momentos necessários.

Sonhe, acredite e, então, realize, pois, "O impossível é só uma questão de opinião."
(Charlie Brown Jr.)

Resumo

Devido às mudanças climáticas, a ocorrência de incêndios em diversos biomas, adaptados ou não ao fogo, tem aumentado significativamente nas últimas décadas, causando danos ambientais sendo, alguns, irreparáveis. Nesse contexto, modelos computacionais baseados em Autômatos Celulares (AC) são capazes de simular e prever o comportamento da propagação de incêndios, auxiliando as autoridades na tomada de medidas de controle e prevenção. A construção e o ajuste de parâmetros desses modelos são tarefas complexas devido à quantidade de variáveis envolvidas. Características do ambiente influenciam diretamente o comportamento dos incêndios. Assim, uma ferramenta capaz de realizar o ajuste de forma automática baseada em Algoritmos Genéticos (AG) tem se mostrado promissora. Sendo assim, o objetivo desta pesquisa é produzir um AG capaz de ajustar adequadamente um conjunto de parâmetros de um modelo de propagação de incêndios florestais baseado em AC, de modo a replicar a dinâmica das chamas representada em um conjunto de dados de incêndios de uma região monitorada. Adicionalmente, é proposto um novo parâmetro "taxa de amostragem" (SR), a fim de melhorar a adaptabilidade do modelo AC à dinâmica de propagação do fogo, independentemente do intervalo de amostragem presente nos dados de referência. Também foi avaliado o impacto de diferentes funções de aptidão na convergência do algoritmo. Durante os experimentos, conjuntos de dados artificiais foram utilizados como referência, os quais foram gerados a partir de simulações com nosso próprio modelo de AC, bem como por outro modelo de propagação de incêndios da literatura, a fim de validar a robustez da abordagem. Os resultados mostraram que o AG conseguiu encontrar parâmetros adequados para o modelo de propagação de incêndios, inclusive o número de passos do AC necessários para representar a evolução do fogo em conformidade com os dados de referência gerados por ambos os modelos.

Palavras-chave: Algoritmos Genéticos, Ajuste de Parâmetros, Autômatos Celulares, Modelo de Propagação de Incêndios, Taxa de Amostragem.

Abstract

Due to climate change, the occurrence of fires in various biomes, whether fire-adapted or not, has significantly increased in recent decades, causing environmental damage, some of which is irreparable. In this context, computational models based on Cellular Automata (CA) can simulate and predict fire propagation behavior, assisting authorities in making control and prevention decisions. However, constructing and tuning the parameters of these models are complex tasks due to the number of variables involved. Environmental characteristics directly influence fire behavior. Thus, a tool capable of performing automatic parameter tuning based on Genetic Algorithms (GA) has proven to be promising. Therefore, the objective of this research is to develop a GA capable of properly adjusting a set of parameters of a forest fire propagation model based on CA to replicate the fire dynamics represented in a dataset of monitored region fires. Additionally, a new parameter, "sampling rate" (SR), is proposed to improve the adaptability of the CA model to fire propagation dynamics, regardless of the sampling interval in the reference data. The impact of different fitness functions on the algorithm's convergence was also evaluated. During the experiments, artificial datasets were used as references, generated from simulations with our own CA model as well as another fire propagation model from the literature, to validate the robustness of the approach. The results showed that the GA successfully found appropriate parameters for the fire propagation model, including the number of CA steps required to represent fire evolution in accordance with the reference data generated by both models.

Keywords: *Genetic Algorithms, Parameter Tuning, Cellular Automata, Fire Propagation Model, Sampling Rate.*

Lista de ilustrações

Figura 1 – Incêndios registrados em 2024 26
Figura 2 – Incêndios e mudanças climáticas 26
Figura 3 – Diferentes formas e dimensões de um AC. 34
Figura 4 – Autômato celular unidimensional e evolução temporal. 35
Figura 5 – Espaço celular de um AC bidimensional e tipos de vizinhança 35
Figura 6 – Padrões *Game of Life.* 36
Figura 7 – Exemplo regra determinística. 37
Figura 8 – Regra 30. 38
Figura 9 – Regra 30 com diferentes padrões para condições iniciais. 39
Figura 10 – Fluxograma de um AG. 40
Figura 11 – Representação cromossômica de um AG. 41
Figura 12 – Representação de uma população. 41
Figura 13 – Seleção por roleta. 43
Figura 14 – Seleção por torneio. 43
Figura 15 – Cruzamento com 1 ponto de corte. 45
Figura 16 – Cruzamento com 2 pontos de corte. 45
Figura 17 – Tipos de mutação. 46
Figura 18 – Reinserção por elitismo. 47
Figura 19 – Estados do AC e evolução temporal. 57
Figura 20 – Matriz W e vetor resultante. 58
Figura 21 – Dinâmica do modelo em função da regra de recuperação da vegetação adotada (determinística ou estocástica). 60
Figura 22 – Possíveis estados e evolução temporal modelo de validação. 61
Figura 23 – Cromossomo 64
Figura 24 – Impacto da taxa de amostragem na avaliação 65
Figura 25 – Divisão reticulado e focos iniciais incêndio. 66
Figura 26 – Interação entre autômato celular e o algoritmo genético 67
Figura 27 – Esferas cromáticas e mapa de calor. 76

Figura 28 – Matriz de preferência vento: Fogo. 80
Figura 29 – Evolução temporal e frequência contabilizando-se Fogo. 80
Figura 30 – Evolução temporal e frequência FQ. 82
Figura 31 – Matriz de preferência vento: Fogo, Queimado. 83
Figura 32 – Evolução temporal e frequência FQV. 84
Figura 33 – Matriz de preferência vento: Fogo, Queimado e Vegetação 84
Figura 34 – Melhores matrizes de preferência. 86
Figura 35 – Evolução temporal autômato celular e mapa de calor. 86
Figura 36 – Evolução temporal autômato celular. 89
Figura 37 – Evolução temporal mapa de calor. 90
Figura 38 – Matrizes preferência vento do experimento sensibilidade. 90
Figura 39 – GA e GA-SR evolução temporal e mapa de calor. 92
Figura 40 – Modelo prob. matriz de preferência. 95
Figura 41 – Evolução temporal recuperação probabilística. 97
Figura 42 – Mapa de calor recuperação probabilística. 98
Figura 43 – Evolução autômato celular e mapa de calor: Fogo 99
Figura 44 – Evolução autômato celular e mapa de calor: Fogo, Queimado. 100
Figura 45 – Matriz de preferência modelo referência. 101
Figura 46 – GA e GA-SR evolução temporal configuração modelo referência. 103
Figura 47 – Reticulado e mapa de calor Oeste. 104
Figura 48 – Reticulado e mapa de calor Nordeste. 104
Figura 49 – Reticulado e mapa de calor Sul. 105

Lista de tabelas

Tabela 1 – Melhores resultados: Fogo 79
Tabela 2 – Melhores resultados: Fogo e Queimado 81
Tabela 3 – Melhores resultados: Fogo, Queimado e Vegetação 83
Tabela 4 – Melhores funções por cenário avaliado 85
Tabela 5 – Referência e melhores parâmetros encontrados pelo AG para cada *dataset* 88
Tabela 6 – Teste de robustez 91
Tabela 7 – Comparação abordagem tradicional e proposta 94
Tabela 8 – Teste de sensibilidade 94
Tabela 9 – Teste funções: Fogo 99
Tabela 10 – Teste funções: Fogo, Queimado 100
Tabela 11 – Análise sensibilidade modelo referência 101
Tabela 12 – Teste de adaptabilidade 102

Lista de siglas

AC Autômato Celular

AG Algoritmo Genético

DS *Dataset* (Conjunto de Dados)

EVOCA-WiSM *Evolutionary Optimization of Cellular Automata-based Wildfire Spreading Models* (Otimização evolutiva de modelos de propagação de incêndios florestais baseados em autômatos celulares)

GA-SR *Genetic Algorithm - Sample Rating* (Algoritmo Genético - Taxa de Amostragem)

INPE Instituto Nacional de Pesquisas Espaciais

ONU Organização das Nações Unidas

SR *Sample Rating* (Taxa de Amostragem)

Lista de variáveis

α Constante de Recuperação

λ Intensidade Global das Chamas - Caloria

δ Influência do Vento

ϵ Limiar

d Decaimento

I Intensidade

il Intensidade Local das Chamas

LQ Tempo de Queima

LR Tempo de Recuperação

N Instantes de Propagação

p Período

pb Probabilidade Base de Queima

S Estados do AC

S_t Estágios de Queima

S_t Estágios de Recuperação

t Tempo

T Passos de Tempo

Tb Estágios de Queima $Brasa$

Tf_e Estágios de Queima $Fogo_{estável}$

Tf_i Estágios de Queima $Fogo_{inicial}$

ts_r Tempo de Recuperação

W Matriz de Preferência do Vento

WC Matriz de Combustão

Sumário

1 INTRODUÇÃO 25
1.1 Motivação 29
1.2 Objetivos e Questões de Pesquisa 30
1.3 Contribuições 31
1.4 Organização da Dissertação 32

2 FUNDAMENTAÇÃO TEÓRICA 33
2.1 Autômatos Celulares 33
2.1.1 ACs Unidimencionais 34
2.1.2 ACs Bidimencionais 35
2.1.3 Regras de Transição de um AC 37
2.2 Algoritmos Genéticos 39
2.2.1 População 40
2.2.2 Função Objetivo 41
2.2.3 Seleção dos Pais 42
2.2.4 Operadores Genéticos 43
2.2.5 Reinserção 46
2.3 Avaliação de Aptidão 47
2.3.1 Fit9P 47
2.3.2 F1 e F2 47
2.4 Normalização de Dados 48
2.5 Trabalhos Correlatos 49
2.5.1 Modelos de Propagação de Incêndios 49
2.5.2 Algoritmos Genéticos no Ajuste de Parâmetros de ACs 52

3 MODELOS PROPAGAÇÃO DE INCÊNDIOS 57
3.1 Modelo de Referência 57
3.1.1 Regras de Transição 59

3.1.2 Modelo Estocástico e Determinístico 59
3.2 Modelo de Validação 60
3.2.1 Matriz de Ventos 60
3.2.2 Regras de Transição 61

4 ABORDAGEM EVOLUTIVA PROPOSTA: EVOCA-WISM . 63
4.1 Representação do Indivíduo 63
4.2 Funções de Avaliação 65
4.2.1 F1P9 e F2P9 69
4.2.2 Funções Normalizadas Pelo Máximo 69

5 EXPERIMENTOS E ANÁLISE DOS RESULTADOS 71
5.1 Configuração Experimental 71
5.1.1 Hiperparâmetros do AC 72
5.1.2 Hiperparâmetros do AG 73
5.2 Geração das Bases de Dados 74
5.3 Análise Mapa de Calor Propagação de Fogo 75
5.4 Avaliação da Abordagem Proposta 76
5.5 Modelo com Recuperação Determinística 77
5.5.1 Desempenho Diferentes Funções de Aptidão 78
5.5.2 Análise Sensibilidade 86
5.5.3 Análise de Robustez 90
5.6 Modelo com Recuperação Probabilística 92
5.6.1 Análise de Sensibilidade 93
5.6.2 Análise de Robustez 96
5.6.3 Análise Sensibilidade FitMax 100
5.6.4 Análise de Adaptabilidade 102

6 CONCLUSÃO 107
6.1 Trabalhos Futuros 108
6.2 Contribuições em Produção Bibliográfica 110

REFERÊNCIAS 111

Capítulo 1

Introdução

Recentemente, devido às mudanças climáticas, vários focos de incêndios em diferentes locais do globo têm tomado proporções alarmantes, sendo prejudiciais tanto para vegetações quanto para áreas urbanas (EDO et al., 2024). Temperaturas mais altas ressecam a paisagem e ajudam a criar o ambiente perfeito para incêndios florestais mais extensos e mais frequentes.

Em 2024, os incêndios florestais resultaram na perda de 30,8 milhões de hectares queimados no Brasil, afetando diversos biomas como a Amazônia, Cerrado, Pantanal, Mata Atlântica, Pampa e Caatinga. A ocorrência de incêndios florestais representa um aumento de 79% (13,6 milhões de hectares) em relação a 2023, (MAPBIOMAS Brasil, 2025). O boletim apresentado pelo Instituto Nacional de Pesquisas Espaciais (INPE) mostra que o Brasil registrou mais de 278,3 mil focos de incêndio (Figura 1), cidades do triângulo mineiro como Uberlândia e do interior de São Paulo como Ribeirão Preto sofreram devido a imensas nuvens de fumaça decorrentes das queimadas (CARDOSO, 2024). Ao longo do globo a situação também se agravado incêndios na Califórnia, Los Angeles, Canadá, Portugal e em outros países causaram diversos danos, autoridades da Organização das Nações Unidas (ONU) alertam que incêndios florestais vão ficar cada vez mais intensos e prolongados (BIERNATH, 2025).

Um estudo realizado pela *Global Forest Watch* (MACCARTHY et al., 2025), representado graficamente na Figura 2, mostra que as queimadas levam a um ciclo vicioso, pois, conforme as emissões de carbono aumentam, o clima se torna mais quente e seco, o que favorece a incidência de queimadas em áreas cada vez maiores e por um período prolongado, o que gera mais emissão de carbono. O ciclo vai se repetindo e, a cada ano, os danos oriundos da propagação de incêndios são maiores, além de fatores relacionados à intervenção humana que agravam ainda mais o quadro.

Dentre os esforços ligados à prevenção e combate dos incêndios estão aqueles relativos à ampliação do entendimento de como esse fenômeno se propaga em diferentes paisagens. Esta é, de fato, uma questão complexa uma vez que, entre as dificuldades no combate a incêndios estão a falta de acesso às áreas atingidas e o entendimento de como áreas com

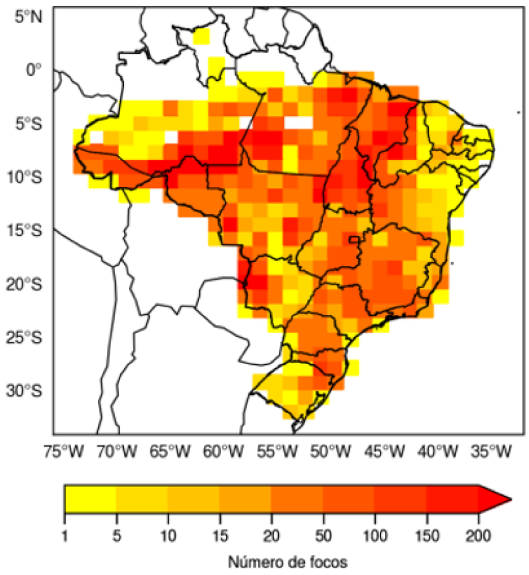


Figura 1 – Incêndios registrados em 2024
Fonte: (Instituto Nacional de Pesquisas Espaciais (INPE), 2024)



Figura 2 – Incêndios e mudanças climáticas
Fonte: (MACCARTHY et al., 2025)

maior potencial de queima são influenciadas por elementos como relevo, vegetação, solo e vento no episódio de queima (CARRILLO et al., 2023).

Modelos computacionais desempenham um papel cada vez maior no avanço da ciência, sendo aplicados a uma ampla variedade de temas. O seu objetivo é replicar o comportamento de sistemas com base nas propriedades reais e conhecidas de seus componentes, fornecendo um complemento importante para estudos experimentais e teóricos ao incrementar sua base de conhecimento (BRODLAND, 2015).

Tais modelos tiveram grandes avanços partindo de áreas de engenharia e física, especialmente nos domínios estrutural, aeroespacial, mecânico, eletromagnético, fluidodinâmico, químico, de controle e elétrico (ZIENKIEWICZ; TAYLOR, 2005). Nesse contexto, houve um extenso desenvolvimento de algoritmos e da teoria matemática necessários para assegurar a confiança aos cálculos, sendo as características destes modelos descritas por equações (BRODLAND, 2015). Tais equações muitas vezes são diferenciais e descrevem a evolução dos modelos ao longo do tempo, porém o principal desafio é encontrar aquelas capazes de descrever as variáveis do sistema e suas interações já que cada uma delas possui características próprias (FRAGA; OLIVEIRA; MARTINS, 2020).

Autômatos Celulares (ACs) têm sido utilizados como alternativa à complexidade apresentada no modelamento tradicional com equações diferenciais (MARTÍNEZ et al., 2013). Os ACs podem ser definidos como modelos computacionais que representam sistemas dinâmicos a partir de componentes simples que, ao atuarem em conjunto, são capazes de produzir padrões de comportamento complexos (computação emergente). Eles são caracterizados por uma computação descentralizada, onde cada célula evolui com base no seu estado atual e de sua vizinhança, através de regras de transição simples que representam a interação local da vizinhança, mas que, num espaço de tempo discreto, proporcionam uma dinâmica global (PACKARD; WOLFRAM, 1985). Tais características tornam ACs uma técnica versátil capaz de representar fenômenos físicos, químicos, naturais, entre outros. Dessa forma, seu uso tem se tornado cada vez mais presente na modelagem da dinâmica de espalhamento do fogo em incêndios florestais, tais como os modelos de (CATCHPOLE; CATCHPOLE; ROTHERMEL, 1993) um dos pioneiros no estudo de incêndios, (LOUZADA; FERREIRA, 2008), (LIMA; LIMA, 2014), (LEI; ZHIQIANG, 2017), (FREIRE; DACAMARA, 2019), (FERREIRA et al., 2022), (FERREIRA et al., 2023), (BRASIEL; LIMA, 2023) e entre outros.

Entretanto, para que os ACs que sejam capazes de representar esses fenômenos de forma adequada, é necessário ajustar corretamente os parâmetros do modelo. A estimação precisa dos valores para esses parâmetros não é uma tarefa trivial, exigindo um elevado conhecimento do fenômeno em estudo, e costuma demandar um tempo considerável. Sendo assim, uma ferramenta capaz de ajustar automaticamente o conjunto de parâmetros de um modelo para reproduzir adequadamente a dinâmica presente em dados históricos é fundamental para o estudo desses fenômenos, auxiliando no planejamento de

políticas públicas de combate e prevenção de incêndios florestais (FRAGA; OLIVEIRA; MARTINS, 2020).

Abordagens baseadas em algoritmos evolutivos, como os algoritmos genéticos, são amplamente usadas na literatura para ajustar os parâmetros de modelos complexos, como classificação de densidade (PAREDIS et al., 1997), definição de parâmetros de redes neurais (TIRUMALA, 2020), problemas de escalonamento (SHEN; GUO; LI, 2020), resolução de problemas reais como otimização de abastecimento de água (SILVA et al., 2003) e vias de transporte (MUGNELA, 2021), dentre outros.

Algoritmos Genéticos (AGs) são algoritmos de busca, inspirados na teoria da evolução de Darwin, que empregam conceitos de seleção natural, combinando a sobrevivência do mais apto e a propagação de seus genes para as gerações futuras, a fim de direcionar a exploração pelo espaço de busca (GOLDBERG, 1989). Embora não haja garantia de obtenção do resultado ótimo, o AG costuma encontrar boas soluções (ótima ou próxima dela) para o problema investigado, usando seus operadores genéticos ao longo do processo evolutivo para fugir de mínimos ou máximos locais, a depender da natureza do problema, e conduzir a busca para regiões promissoras do espaço de soluções (LINDEN, 2006).

Este trabalho tem como objetivo desenvolver uma abordagem evolucionária usando algoritmos genéticos para ajustar automaticamente os parâmetros de um modelo de propagação de incêndios florestais baseado em autômato celular probabilístico, tendo como base a dinâmica e os padrões de espalhamento de fogo observados em conjuntos de dados históricos das queimadas na região monitorada. Nossa abordagem é uma extensão do modelo proposto em (FERREIRA et al., 2022). O trabalho original se mostra eficiente na reprodução de incêndios a partir de dados gerados artificialmente pelo próprio modelo de autômato celular ajustado pelo AG. Entretanto, apresenta dificuldades no ajuste dos parâmetros a fim de representar a dinâmica do fogo a partir de dados gerados usando outro modelo de propagação.

A principal contribuição de nossa pesquisa é propor mecanismos para melhorar a adaptabilidade da abordagem aos vários padrões presentes nos conjuntos de dados. Inicialmente, investigamos diferentes formas de avaliação usadas pelo algoritmo genético para aferir a aptidão das possíveis soluções (configurações do modelo de propagação). Além da função objetivo em si, também avaliamos os estados das células que são considerados no cálculo das diferenças entre os cenários esperados e os reticulados do AC ao longo da simulação. Também introduzimos um novo parâmetro a ser configurado pelo AG, chamado taxa de amostragem (SR - Sample Rate), o qual determina o intervalo de simulação entre as avaliações, ou seja, quantos passos de evolução do autômato celular são executados para gerar cada reticulado que será comparado com uma amostra temporal presente no conjunto de dados de incêndios da região monitorada. Tais melhorias visam uma aferição mais precisa da similaridade entre a evolução temporal do AC e a dinâmica presente no conjunto de dados analisado.

A capacidade do AG em encontrar valores adequados para os parâmetros do AC é analisada considerando diferentes intervalos de amostragem entre os dados representados no conjunto de dados usado como referência. O desempenho da abordagem também é validado a partir de dados gerados artificialmente com outro modelo de propagação, confirmando o aumento da sua robustez e adaptabilidade em relação ao modelo original (FERREIRA et al., 2022). Resultados experimentais mostram que os valores dos parâmetros sugeridos pelo AG resultam em uma dinâmica de espalhamento das chamas semelhante à dinâmica representada em cada conjunto de dados de referência, demonstrando a robustez e adaptabilidade do método.

1.1 Motivação

O fogo nem sempre é prejudicial, podendo ser benéfico em alguns biomas, como o caso do cerrado brasileiro, onde algumas plantas florescem apenas após a passagem de um incêndio como o caso da *Bulbostylis paradoxa* (PILON et al., 2023). Os efeitos do fogo são bastante conhecidos, porém, devido às ações humanas como desflorestamento e mudanças climáticas têm tornado seu comportamento imprevisível (SANTOS et al., 2021). Neste cenário, controlar e prever a dinâmica de espalhamento pela paisagem é uma forma de preservar o ambiente e evitar maiores consequências, tal prática é denominada manejo do fogo. No Brasil, este manejo considera diferentes usos do fogo sob aspectos sociais ecológicos como queimas relacionadas às práticas de manejo agrícola, conservacionista ou, ainda, para caça e rituais religiosos realizados por povos indígenas e povos e comunidades tradicionais (IBAMA, 2023). Entretanto, este uso deve ser controlado e realizado em épocas que não exista risco de uma queimada descontrolada.

Nos últimos anos devido às mudanças climáticas, focos de incêndio têm se tornado mais frequentes e, a cada ano, a intensidade e danos causados têm sido cada vez mais graves, especialmente em áreas com alto valor biológico, como áreas de preservação, onde os incêndios são um risco para a fauna e a flora, ou de alta densidade demográfica, prejudicando a população e a economia local (ARCA; GHISU; TRUNFIO, 2015).

Nesse contexto, modelos de simulação são ferramentas extremamente úteis na representação de sistemas naturais, como a propagação de incêndios florestais, sendo aliados importantes em práticas de manejo de fogo, uma vez que são capazes de fornecer informações importantes sobre o comportamento do fogo na região estudada, auxiliando na tomada de decisão dos órgãos responsáveis pela prevenção e combate de incêndios, visando evitar grandes desastres e prejuízos ao ecossistema da área monitorada.

Modelos baseados em autômatos celulares têm sido aplicados em diferentes áreas do conhecimento, desde criptografia, artes, até modelagem e simulação de sistemas complexos. Na última área, entretanto, às vezes a capacidade de representar adequadamente dinâmicas complexas, mas distintas, que ocorrem dentro de uma determinada área, é

limitada pela necessidade de modelos de calibração nos quais o número de parâmetros necessários aumenta. Custos ocultos relacionados à identificação de valores específicos ou faixas plausíveis para parâmetros podem se tornar opressores (GARCÍA-DURO et al., 2018).

No caso de incêndios florestais, para que os autômatos celulares (ACs) sejam capazes de simular o fenômeno de forma adequada, sua parametrização envolve fatores como clima, relevo, quantidade de material combustível, intensidade do vento e muitos outros parâmetros. O ajuste adequado desses parâmetros não é uma tarefa trivial, a depender do número de variáveis a serem consideradas, demandando um conhecimento especializado e um tempo considerável (FRAGA; OLIVEIRA; MARTINS, 2020). Além disso o ajuste depende da qualidade dos dados e do intervalo de amostragem, que precisam ter um tamanho adequado e devem conter as informações necessárias para a identificação dos padrões de espalhamento durante a avaliação. Assim, uma ferramenta capaz de ajustar os parâmetros de forma automática é de extrema utilidade.

Este trabalho combina AC e AG para reproduzir o comportamento de incêndios florestais, encontrando valores para os parâmetros do modelo de propagação, de modo a simular adequadamente o espalhamento do fogo de acordo com os padrões presentes no conjunto de dados de referência. A validação da abordagem foi realizada por meio da análise comparativa entre as simulações resultantes do modelo configurado pelo AG e o conjunto de dados analisado.

1.2 Objetivos e Questões de Pesquisa

O objetivo da pesquisa é desenvolver uma abordagem evolutiva capaz de ajustar os parâmetros de um modelo de propagação do fogo baseado em autômatos celulares probabilísticos para representar adequadamente os comportamentos observados em dados históricos de incêndios florestais.

Nesse sentido, este trabalho busca atingir os seguintes objetivos específicos:

- ❏ Estudar e implementar um modelo de propagação de incêndios baseado em autômatos celulares, o qual será usado como estudo de caso na avaliação do módulo de ajuste de parâmetros.
- ❏ Estudar e implementar uma abordagem evolutiva de configuração de parâmetros do modelo de propagação de incêndio, utilizando Algoritmos Genéticos (AG).
- ❏ Investigar e avaliar outras funções de *fitness* que considerem mais informações além da existência do fogo como área não queimada e em recuperação.
- ❏ Incluir a taxa de amostragem entre as avaliações como parâmetro a ser configurado pelo AG.

- ❏ Analisar a eficiência da abordagem evolutiva ao ajustar o comportamento do modelo de referência a dados produzidos por outros modelos de propagação de incêndios.

Desta forma, buscou-se responder às seguintes questões de pesquisa:

- ❏ Qual função de avaliação permite aferir com maior precisão a similaridade entre a dinâmica simulada pelo modelo de propagação e aquela representada no conjunto de dados de referência?
- ❏ Quais estados do autômato celular são relevantes na avaliação dos indivíduos do AG, a fim de melhorar a qualidade da dinâmica simulada?
- ❏ Qual a quantidade mínima de registros no conjunto de dados necessária para que o algoritmo genético seja capaz de identificar o padrão representado e ajustar adequadamente o modelo de propagação de incêndios?
- ❏ Qual a sensibilidade do algoritmo em relação à taxa de amostragem, ou seja, qual é o limiar máximo de amostragem dos dados do conjunto de referência que o AG consegue reconhecer os padrões, sem comprometer a qualidade da dinâmica reproduzida?
- ❏ As modificações propostas nesta pesquisa proporcionam uma abordagem evolutiva mais adaptativa, capaz de reproduzir de forma satisfatória diferentes comportamentos de incêndios florestais?

1.3 Contribuições

Neste trabalho foi desenvolvida uma abordagem evolutiva para ajuste automático dos parâmetros de modelos de propagação de incêndios baseados em autômatos celulares bidimensionais probabilísticos que representam a vegetação florestal da região monitorada. O modelo de AC configurado deve ser capaz de reproduzir a dinâmica e padrões de espalhamento do fogo observados no conjunto de referência, que representa o histórico de incêndios florestais. Utilizando como inspiração modelos usados em sistemas epidemiológicos, foi desenvolvida uma função de avaliação capaz de considerar informações importantes no processo de espalhamento do fogo, além da presença ou ausência de chamas em uma célula, fazendo com que nossa abordagem alcance um refinamento mais preciso dos parâmetros do modelo em estudo. Por fim, a principal contribuição deste trabalho é a introdução de um novo parâmetro utilizado na avaliação do modelo de AC, denominado taxa de amostragem, que permite ao AG tornar-se mais adaptável a diferentes cenários ao ajustar os passos de evolução entre as amostras de dados para capturar adequadamente a dinâmica de fogo presente no conjunto de referência.

1.4 Organização da Dissertação

O restante do documento está organizado da seguinte maneira:

- Capítulo 2: são apresentados os fundamentos e conceitos relacionados as técnicas de computação bio-inspirada investigadas neste trabalho que são autômatos celulares e algoritmos genéticos. Também são apresentadas as técnicas de normalização e padronização de dados utilizadas na combinação das diferenças calculadas na função de avaliação, com escalas distintas, em um único valor que descreve a aptidão de uma possível solução. Por fim, são apresentados trabalhos correlatos a ACs e AGs que fornecem outros conceitos fundamentais para o entendimento desta pesquisa e serviram de inspiração para o desenvolvimento da abordagem.

- Capítulo 3: é apresentado de forma detalhada a reprodução dos modelos baseados em autômatos celulares utilizados como referência e validação da abordagem desenvolvida para ajuste automático dos parâmetros dos modelos por meio de algoritmos genéticos.

- Capítulo 4: descreve em detalhes a abordagem evolutiva proposta utilizando algoritmos genéticos e suas configurações.

- Capítulo 5: são apresentados os resultados alcançados nos experimentos efetuados, bem como suas respectivas análises.

- Capítulo 6: são apresentadas as considerações adicionais, conclusões e perspectivas de trabalhos futuros.

Capítulo 2

Fundamentação Teórica

Nesta seção serão abordadas as características das técnicas utilizadas no desenvolvimento desta pesquisa. Sendo assim, a Seção 2.1 aborda os conceitos de Autômatos Celulares, a Seção 2.2 apresenta os Algoritmos Genéticos, técnica evolutiva utilizada na calibragem dos modelos. A seção 2.4 apresenta algumas técnicas de normalização de dados utilizadas na unificação do cálculo de *fitness*. Por fim, uma visão geral dos principais trabalhos que estão sendo realizados utilizando Autômato Celular (AC) e Algoritmo Genético (AG) é apresentada na Seção 2.5.

2.1 Autômatos Celulares

A proposta inicial dos autômatos celulares (AC) é creditada a John Von Neumann e Stannislaw Ulam. Na década de 50, Neumann tinha como objetivo projetar agentes computacionais capazes de auto-reprodução; na mesma época, Ulam investigava jogos matemáticos dispostos em uma grade bidimensional onde o estado de cada ponto era atualizado com base no seu estado e no dos vizinhos. A partir desta sugestão, Von Neumann construiu o primeiro modelo de AC e este autômato resolvia a questão da auto-reprodução. Ele foi capaz de encontrar uma máquina universal de Turing encapsulada em uma grade de células, usando 29 estados por célula e cinco células como vizinhança (OLIVEIRA, 2003).

Em 1982, Wolfram estudou a mecânica estatística e o funcionamento dos autômatos celulares e a formação de padrões. Devido à complexidade dos resultados, concluiu que eles poderiam ser utilizados como modelos dinâmicos complexos, presentes em sistemas físicos, biológicos, entre outros. Em 2002, apresentou os resultados de 20 anos de pesquisa no livro *A New Kind of Science* (WOLFRAM, 2002).

Um autômato celular é definido pelo seu espaço celular (reticulado) e por sua regra de transição. O espaço celular é composto por um conjunto de células dispostas de maneira N-Dimensional (1D, 2D, 3D), em que cada célula se conecta de maneira igual com sua vizinhança, sendo facilmente configurados em qualquer estrutura geométrica, conforme

ilustrado na Figura 3. Sendo mais comuns em 2D o triângulo (3 vizinhos) quadrado (4 vizinhos) e o hexágono (6 vizinhos); em 3D temos por exemplo o cubo (6), o prisma hexagonal (8), o dodecaedro rômbico (12) (WOLFRAM, 2002). Classicamente, os ACs trabalham com espaços de estados discretos, mas em alguns estudos são utilizados estados contínuos como dinâmcia de fluidos ou otimização global (SECK-TUOH-MORA et al., 2021).

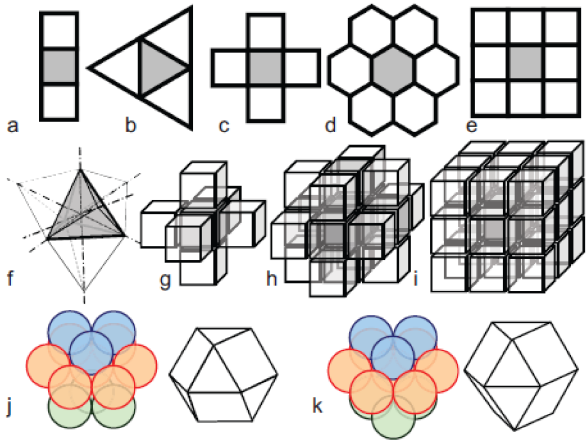
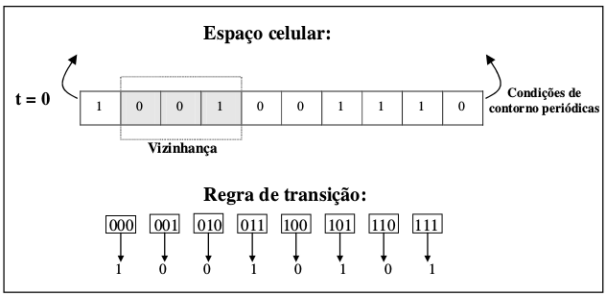


Figura 3 – Diferentes formas e dimensões de um AC.
Fonte: (GOBRON; BONAFOS; THALMANN, 2010)

2.1.1 ACs Unidimencionais

O tipo mais simples de um AC é o autômato unidimensional, descrito como um vetor de células, onde a cada intervalo de tempo discreto, as células assumem um dos possíveis estados de um conjunto finito. O estado seguinte é definido pela combinação de seu estado atual e de seus vizinhos da esquerda e direita, que estão dentro de um raio de alcance, através de uma regra de transição local ou regra local, igual para todas as células (SARKAR, 2000). A condição de contorno denota como as células das extremidades preenchem os vizinhos imediatos faltantes para a aplicação da regra de transição. Diferentes estratégias podem ser aplicadas, como replicação das bordas, adoção de um valor fixo ou das células na extremidade oposta. O conjunto de estados das células em um ponto é denominado configuração ou estado global do AC e descreve sua evolução a partir de uma configuração inicial dos estados das células no tempo $t = 0$ (reticulado inicial). Na Figura 4a é apresentado um exemplo de autômato celular unidimensional com sua regra de transição e condição de contorno. Na Figura 4b são apresentados quatro passos de evolução temporal do AC ($t = 1, 2, 3, 4$), partindo do reticulado inicial ($t = 0$) aplicando as regras de transição em cada instante. Tomando como exemplo a célula central e sua vizinhança destacadas na Figura 4a (001) no estado t=0 ao aplicar as regras de transição $001 \rightarrow 0$ a célula se mantêm nesse estado, se os seu vizinhos fossem ambos a célula passaria para esse estado ($101 \rightarrow 1$), dessa forma a regra é aplicada em cada célula do reticulado ao longo



(a) AC1D e regras de transição.

t = 0	1	0	0	1	0	0	1	1	1	0
t = 1	0	0	0	0	0	0	1	1	0	1
t = 2	0	1	1	1	1	0	1	0	1	0
t = 3	0	1	1	1	0	1	0	1	0	0
t = 4	0	1	1	0	1	0	1	0	0	1

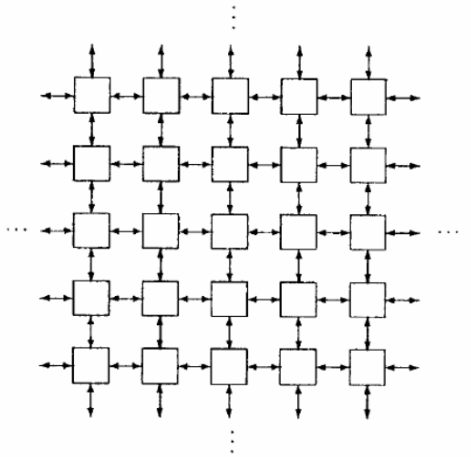
(b) Evolução temporal.

Figura 4 – Autômato celular unidimensional e evolução temporal.
Fonte: (OLIVEIRA, 2003)

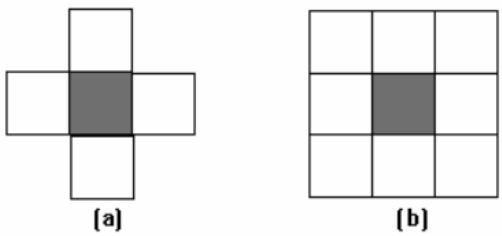
do tempo, produzindo assim a evolução temporal mostrada na Figura 4b até o passo de tempo $t = 4$.

2.1.2 ACs Bidimencionais

Em autômatos bidimensionais, o espaço celular (reticulado) é formado por células dispostas em uma matriz ou grade, conforme ilustrado na Figura 5a. Nesse tipo de AC, as duas vizinhanças mais comum são: von Neumann e Moore. Considerando um raio de alcance unitário ($r = 1$), a vizinhança de von Neumann é composta por 5 células, conforme ilustrado na Figura 5b-[a], isto é, envolve a célula central e suas vizinhas cardinais (direita, esquerda, acima e abaixo). A vizinhaça de Moore é formada por 9 células, sendo as mesmas células da vizinhança anterior, mais as quatro células diagonais, como mostrado na Figura 5b-[b]. Assim como em um AC unidimensional, o estado de uma célula é evoluído a partir de uma regra de transição que considera o estado da célula atual e dos seus vizinhos imediatos.



(a) Espaço celular AC2D.
Fonte: Sarkar (2000)



(b) Vizinhança Von Neuman [a] e Moore [b].
Fonte: Oliveira (2003)

Figura 5 – Espaço celular de um AC bidimensional e tipos de vizinhança

Segundo Castro e Castro (2008) um exemplo de AC bidimensional bem conhecido na literatura é o "jogo da vida" ("game of life", em inglês) proposto por John Holton Conway em 1970. Ele utiliza um reticulado formado por células binárias, isto é, que assumem apenas 2 estados possíveis: morta (0) ou viva (1). A mudança dos estados das células ao longo da evolução do AC, considera uma vizinhança de Moore de raio 1 e é determinada por uma regra de transição totalística, que considera a quantidade de células em cada estado na vizinhança, como segue:

1. **Uma célula viva com 2 ou 3 vizinhos vivos, permanece viva ;**
2. **Uma célula viva com apenas 1 ou 0 vizinhos vivos, morre (solidão);**
3. **Uma célula viva com 4 ou mais vizinhos, morre (sufocada);**
4. **Uma célula morta com exatamente 3 vizinhos vivos, renasce.**

Embora adote uma regra de transição simples, esse AC é capaz de gerar padrões interessantes e complexos (OLIVEIRA, 2003). Na Figura 6a são apresentados alguns exemplos de padrões estacionários (*block, beehive, boat, loaf e ship*), enquanto a Figura 6b ilustra dois padrões oscilatórios (*blinker e toad*). A Figura 6c ilustra a evolução temporal de uma estrutura oscilatória propagante, chamada *glider*, a qual tende a retornar ao seu estado inicial após um período de tempo, com um pequeno deslocamento espacial.

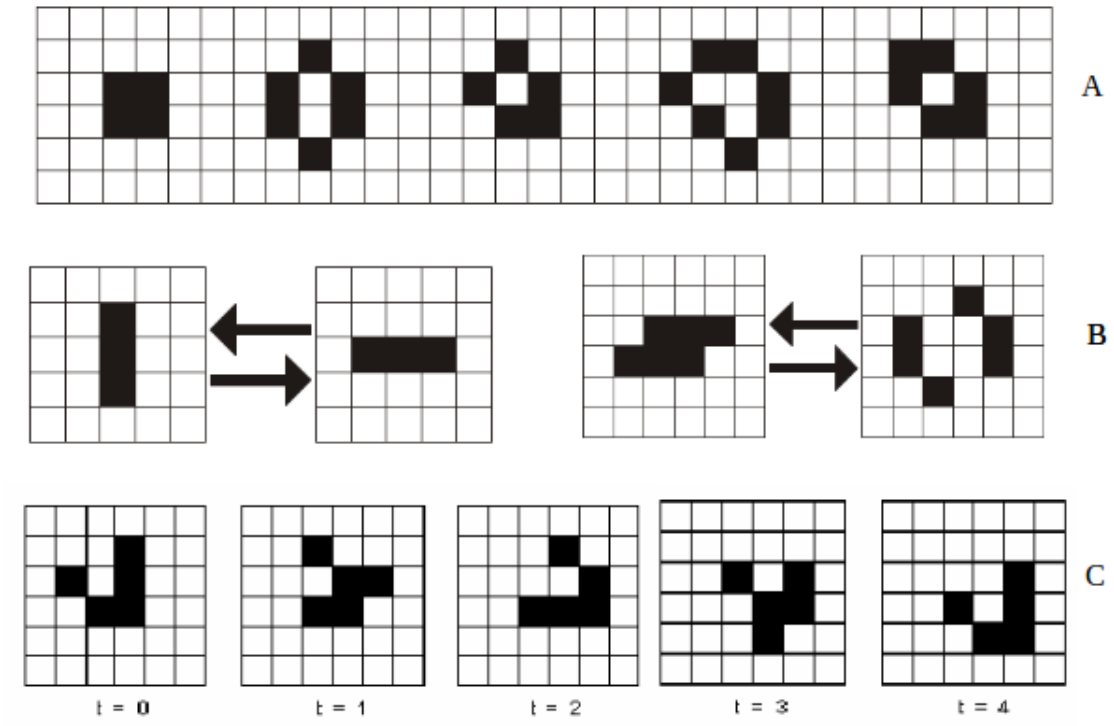


Figura 6 – Padrões *Game of Life.*
Fonte: A e B adaptado de Castro e Castro (2008) e C de Oliveira (2003).

2.1.3 Regras de Transição de um AC

Wolfram (2002) cita em seu livro que se a ciência teórica é possível de alguma forma, então os sistemas estudados precisam seguir regras definidas em algum nível. Tradicionalmente equações matemáticas ou programas baseados em regras complexas são projetados para executar tarefas específicas. Entretanto, fenômenos naturais geralmente podem ser descritos a partir de regras simples que, quando combinadas, resultam em um comportamento complexo. Nesse contexto, os autômatos celulares vem surgindo como uma boa alternativa aos métodos tradicionais para a modelagem desses fenômenos, devido a sua simplicidade (WOLFRAM, 2002).

Qualquer programa em algum grau consiste de conjuntos de regras que especificam o que deve ser feito em cada instrução (WOLFRAM, 2002). Há diferentes formas de especificar as regras que determinam as mudanças de estados das células de um autômato celular com base em sua vizinhança. De acordo com suas características, essas regras podem ser categorizadas em determinísticas, totalísticas e probabilísticas.

As regras determinísticas são a forma mais básica e simples de definir a dinâmica de um autômato celular, sendo utilizadas nos primeiros experimentos de Wolfram. Essas regras são constantes e não incluem elementos de aleatoriedade. O estado de uma célula no próximo passo de tempo $t+1$ é inteiramente definido pelo seu estado atual da célula e de sua vizinhança, garantindo que o valor central seja sempre o mesmo se a condição inicial for a mesma. As regras determinísticas frequentemente produzem padrões estacionários ou ciclos repetitivos. Estas formas básicas podem ser visualizadas graficamente na Figura 7, a qual mostra 5 passos de evolução de um AC unidimensional, considerando a aplicação da regra 254, ilustrada no topo da figura.

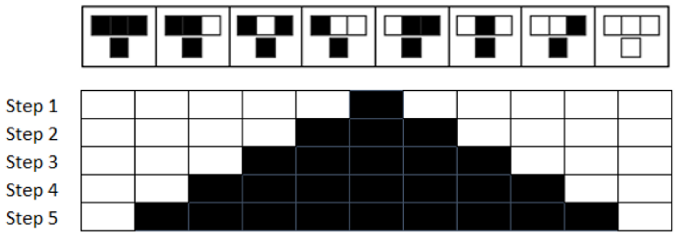


Figura 7 – Exemplo regra determinística.
Fonte: Adaptado de Wolfram (2002)

Algumas regras de transição são mais complexas pois envolvem mais que dois estados possíveis ou consideram uma quantidade maior de células na vizinhança (dimensões ou raios maiores). Nesses casos, não é viável determinar a transição individual para cada possível configuração de vizinhança, sendo necessária uma forma mais eficiente de representação. As regras totalísticas são um tipo específico de regra de transição para autômatos celulares, caracterizadas por dependerem apenas do quantitativo de estados das células vizinhas e, às vezes, da própria célula, ignorando a disposição espacial da

vizinhança para definir o estado da célula no próximo instante de tempo $(t+1)$. Um dos exemplos de regras totalísticas mais conhecidos é o "Game of Life".

Conforme abordado por Wolfram (2002), a aleatoriedade é uma característica comum em fenômenos naturais. Muitos sistemas diferentes possuem esta característica podendo à causa ser diferente para cada fenômeno observado. Em seus estudos, Wolfram concluiu que há três fatores principais responsáveis pela aleatoriedade.

- **Ambiente:** quando a aleatoriedade é introduzida nas regras de transição, de modo que o valor do atributo é escolhido para cada célula em todo passo. Dessa forma, assume-se que há um ambiente externo que continuamente afeta o sistema, injetando aleatoriedade, e uma contínua interação entre eles.

- **Condições Iniciais:** quando não há interação do sistema com o ambiente. Nesse caso, as condições iniciais do sistema costumam ser escolhidas aleatoriamente, mas a evolução temporal do AC seguem regras determinísticas, sem qualquer fator estocástico. Portanto, o comportamento do sistema depende sensivelmente da forma das condições iniciais. Tanto neste modelo quanto no anterior a aleatoriedade vem de entradas externas do sistema, não importando a origem

- **Intrínseco:** nenhuma entrada aleatória é requerida, sendo gerada internamente no próprio sistema, onde cada célula contribui ativamente para a aleatoriedade. Um exemplo é a regra 30, ilustrada na Figura 8. Embora não exista aparentemente qualquer fator externo que introduza aleatoriedade, essa regra produz intrinsecamente um comportamento que parece em muitos aspectos aleatório. A partir de um reticulado inicial específico, a evolução do AC resultará sempre na mesma sequência temporal. Entretanto, conforme a condição inicial é alterada, diferentes padrões são gerados, como mostrado na Figura 9.

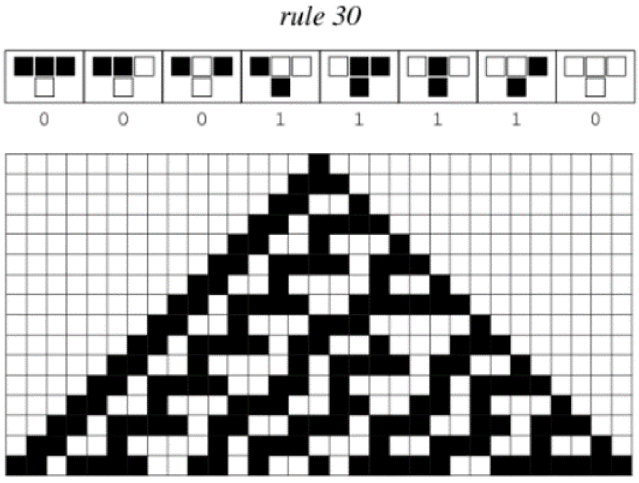


Figura 8 – Regra 30.
Fonte: (WOLFRAM, 2002)

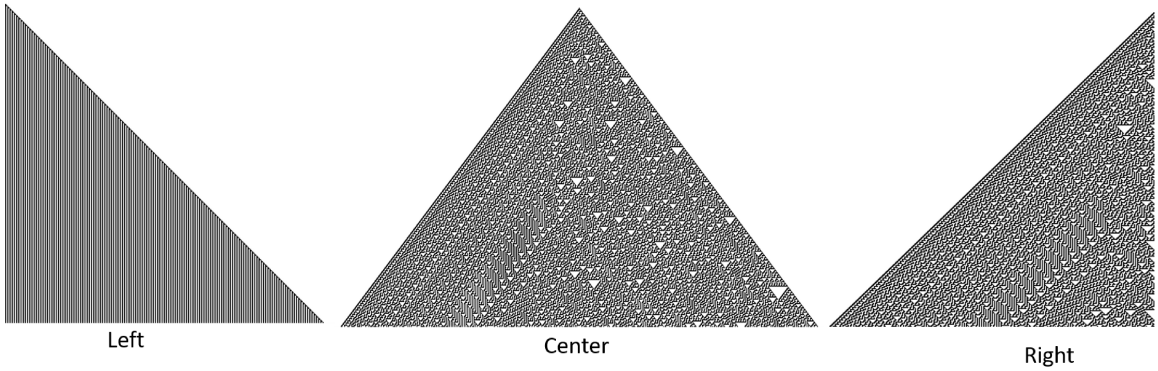


Figura 9 – Regra 30 com diferentes padrões para condições iniciais.
Fonte: (ACKER, 2014)

Nos sistemas que implementam modelos probabilísticos, uma determinada regra de um autômato celular será aplicada se a geração de um número aleatório dentro de uma faixa específica, geralmente 0 a 1 ou 0 a 100, atender a um valor predeterminado. Por exemplo, no modelo de referência (descrito na Seção 3.1) a probabilidade corresponde à chance de uma célula pegar fogo.

2.2 Algoritmos Genéticos

Algoritmos genéticos são métodos de busca e otimização proposto por Holland (1975), mas que se tornaram conhecidos pelo trabalho de (GOLDBERG, 1989). Tal técnica tem como inspiração biológica a teoria da evolução de Darwin, cuja ideia principal que a adaptação de uma população de indivíduos a algum ambiente é dada pela sobrevivência do mais apto (seleção natural) e a transferência das características úteis para as gerações futuras (evolução genética).

AGs se diferenciam de outros métodos tradicionais em quatro fatores (GOLDBERG, 1989):(i) eles trabalham com a codificação do conjunto de parâmetros e não com os parâmetros propriamente ditos; (ii) busca considera vários pontos do espaço de soluções e não apenas em um único; (iii) se valem de funções objetivo que não dependem de derivadas ou outra forma de conhecimento prévio da dinâmica do problema; e (iv) e usam estratégias de evolução probabilística em vez de métodos exatos e determinísticos.

O funcionamento básico de um algoritmo genético típico é ilustrado na Figura 10, sendo um processo estocástico interativo sem convergência garantida que mantém uma população constante, a qual é inicialmente gerada de forma aleatória ou a partir de alguma heurística previamente estabelecida com base no problema investigado. A cada iteração os indivíduos são avaliados e recebem um valor de aptidão (*fitness*), no processo de seleção de pais é feita a escolha de indivíduos que combinados no processo de *crossover*, darão origem a novas soluções chamados de filhos. Eles podem ter suas características alteradas por meio do processo de mutação, responsável por introduzir um ruído na busca com o intuito

de evitar que as soluções fiquem presas em ótimos locais e tenham uma convergência prematura adicionando aleatoriamente novos pontos no espaço de busca. Após o processo de mutação os filhos são avaliados pela função de aptidão, e com base nesse valor se decide quais quais indivíduos irão compor a nova geração no processo denominado reinserção. O processo continua até atingir o critério de parada, quando a busca de soluções por um AG deve cessar. Em geral, o processo encerra alcançando um número definido de gerações, um valor de erro, determinada quantidade de gerações sem melhoria (diferença de erro entre gerações) ou entre outras a depender da especificidade do problema.

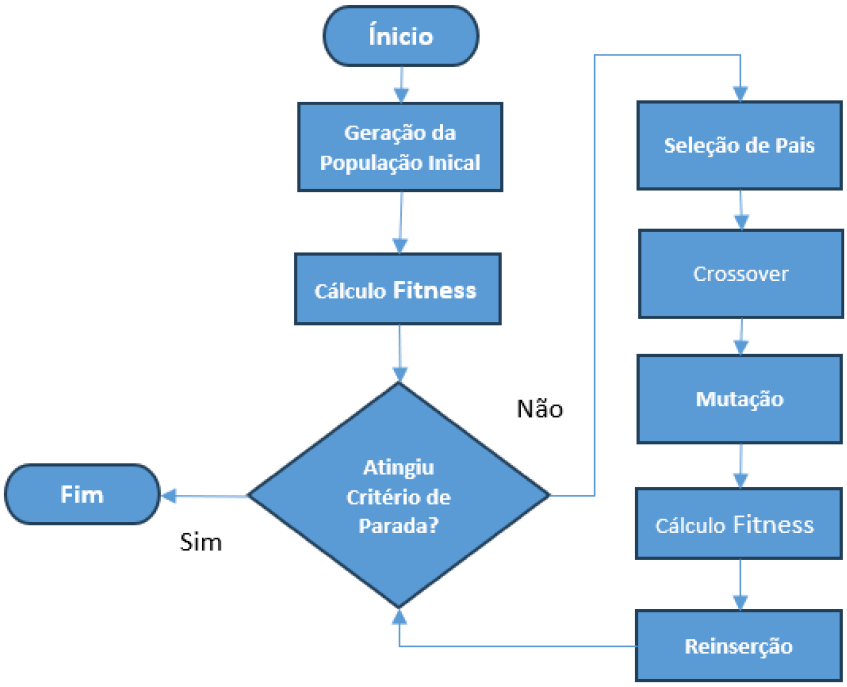


Figura 10 – Fluxograma de um AG.

2.2.1 População

A ligação entre o problema e o AG se dá por meio do indivíduo, que representa uma solução viável para o problema através de um mapeamento adequado. A representação cromossomial é fundamental para o AG, pois ela consiste em uma forma de traduzir a informação do problema em uma forma viável de ser tratada pelo computador, devendo ser simples e capaz de representar restrições caso existam (LINDEN, 2006). Um dos aspectos mais importantes de um algoritmo genético é a representação da solução para um dado problema. Essa informação é representada dentro de uma estrutura de dados adequada, geralmente vetores compondo o que biologicamente é denominado de cromossomo, sendo que suas partes indivisíveis são chamadas de genes. A informação representada por cada gene pode ser especificada de várias formas, sendo que os mais usuais são valores binários, inteiros ou reais, como ilustrado na Figura 11.

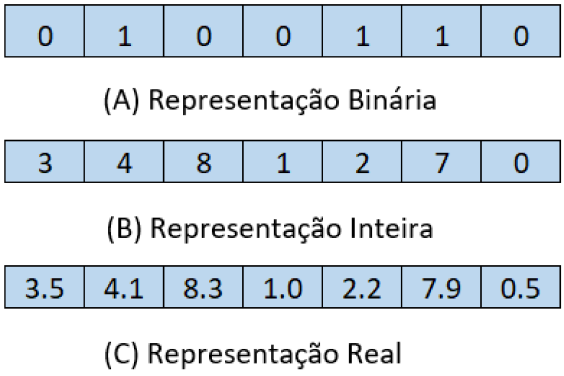


Figura 11 – Representação cromossômica de um AG.

O indivíduo é a unidade básica de um AG pois estão presentes todas as informações (genes) relevantes que irão compor a solução (cromossomo) para um dado problema. Como não existe apenas uma solução, mas todo um conjunto possível de soluções é dado o nome "população" (Figura 12), esse conjunto representa diferentes pontos no espaço de solução, que evoluem ao longo de gerações do AG. A população inicial pode ser gerada aleatoriamente ou de forma heurística, garantindo que a representação cromossômica atenda às especificações do problema, por exemplo que soluções válidas sejam geradas em problemas que tenham restrições.

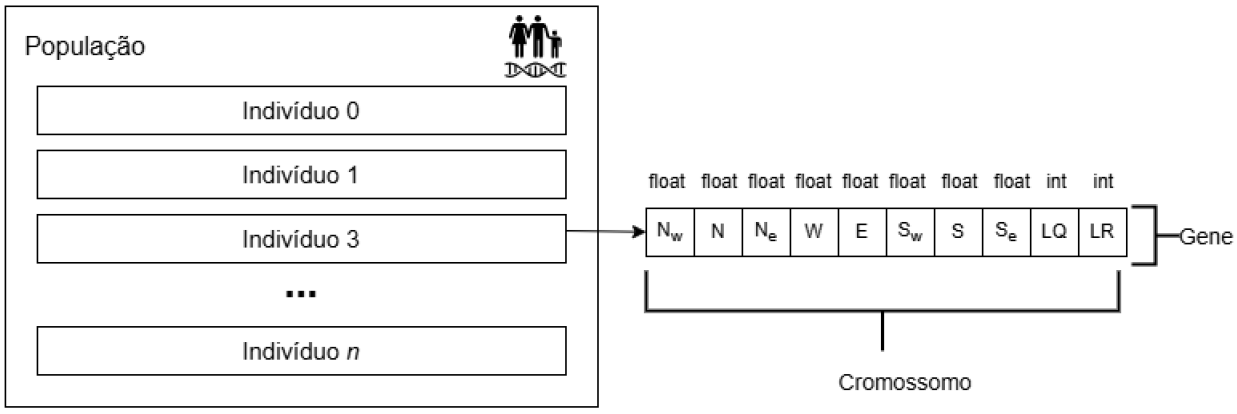


Figura 12 – Representação de uma população.

2.2.2 Função Objetivo

Assim como um indivíduo deve ser capaz de conter as características do problema na representação da solução, é necessário algum mecanismo capaz de dizer o quão boa esta solução se apresenta, o qual é denominado "função objetivo", "função de avaliação" ou "*fitness*". Linden (2006) define, de forma bem simples, a função de avaliação como a nota dada ao indivíduo na resolução do problema, sendo a forma de diferenciar entre boas e más soluções para um problema. Este valor é utilizado na escolha da melhor solução, na seleção de pais e reinserção.

2.2.3 Seleção dos Pais

O método de seleção é usado para simular o mecanismo de seleção natural que atua sobre as espécies biológicas em que os indivíduos mais capazes têm uma descendência maior, embora os cromossomos menos aptos também possam procriar. Sendo assim, indivíduos com uma avaliação alta são privilegiados porém, aqueles com uma avaliação baixa não devem ser desprezados, uma vez que podem possuir características que sejam favoráveis para a geração de um indivíduo que seja mais adequado à solução do problema. Se deixarmos apenas os melhores indivíduos se reproduzirem, a população será composta cada vez mais por indivíduos semelhantes faltando diversidade para que a evolução siga de forma satisfatória, a este processo denominamos convergência genética (LINDEN, 2006).

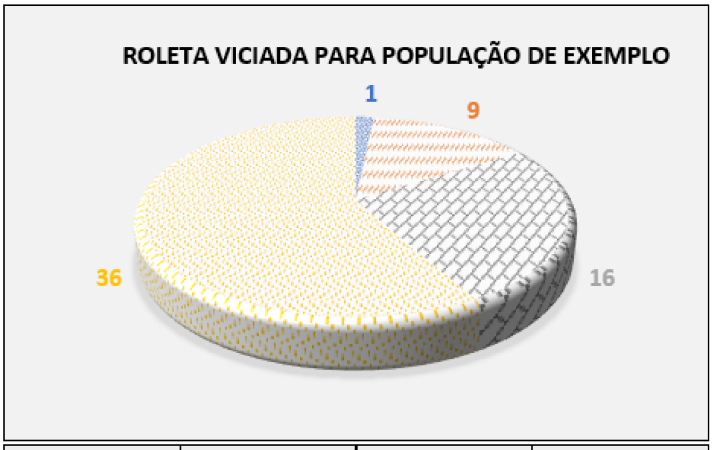
Existem diferentes formas de selecionar os indivíduos (pais) que participarão do *crossover*, sendo que os mais comumente utilizados são os métodos da roleta e por torneio.

A seleção por roleta foi proposta por (HOLLAND, 1975), em que os indivíduos candidatos a pais participam do sorteio por meio da simulação de uma roleta em que cada um recebe um pedaço proporcional à sua avaliação. Após a criação da roleta é feito um sorteio para definir qual indivíduo será selecionado, indivíduos com melhores avaliações recebem porções maiores da roleta, dessa forma possuem uma chance maior de serem selecionados, porém indivíduos com parcelas menores, isto é, com aptidões menores ainda tem a oportunidade de serem escolhidos (LINDEN, 2006). A Figura 13 exemplifica a construção de uma roleta, em que o indivíduo mais apto 0110 recebe mais da metade da role (58,07%) e o com pior aptidão 0001 recebe apenas 1,61%, sendo assim, ao se sortear aleatoriamente uma parcela na roleta aquele que ocupa mais espaço tendem a ser escolhidos como pais pais.

Na seleção por torneio, são selecionados alguns indivíduos da população de forma aleatória e, então, eles são comparados em relação aos seus respectivos *fitness*, sendo escolhido aquele mais apto.

Neste método, existe um parâmetro denominado tamanho do torneio (K) que define quantos indivíduos irão competir sendo o tamanho mínimo de 2 indivíduos. Não existe nenhum favorecimento para melhores indivíduos como acontece no método da roleta, a única vantagem é que, quando selecionados, estes indivíduos provavelmente serão os vencedores.

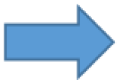
Na Figura 14 é mostrado um exemplo de seleção por torneio com K igual a 3 para um problema de maximização. Logo, indivíduos com um *fitness* mais alto são os melhores para a solução do problema. No primeiro torneio, foram selecionados aleatoriamente os indivíduos X1, X7 e X8, e o primeiro (X1) é considerado o vencedor, pois seu *fitness* é o maior entre eles. O mesmo processo é aplicado aos demais torneios, sendo selecionados os indivíduos X3, X6, X1 e X3 como vencedores, respectivamente.



Indivíduo	Avaliação	Pedaço da Roleta (%)	Pedaço da Roleta (°)
0001	1	1,61	5,8
0011	9	14,51	52,2
0100	16	25,81	92,9
0110	36	58,07	209,1
Total	62	100	360

Figura 13 – Seleção por roleta.
Fonte: Adaptado de Linden (2006)

Indivíduo	Fitness
X_1	200
X_2	100
X_3	9500
X_4	100
X_5	100
X_6	10000
X_7	1
X_8	40



X_1	X_7	X_8
X_2	X_3	X_5
X_6	X_4	X_4
X_2	X_7	X_1
X_3	X_4	X_2

Figura 14 – Seleção por torneio.
Fonte: Adaptado de Linden (2006)

2.2.4 Operadores Genéticos

Os operadores genéticos são os responsáveis pela busca realizada por um AG produzindo novos indivíduos para compor as soluções da geração seguinte. Ao aplicá-los nos indivíduos da população atual, parte das informações são mantidas, mas novas são adicionadas por meio de alteração dos valores atuais (mutação) ou recombinação dos já existentes (*crossover*). Eles são aplicados ao longo das gerações, onde a cada iteração a ação destes conduz a uma busca pelo espaço de soluções, busca esta que seria realizada pela evolução natural se os organismos ficassem em um ambiente imutável. É importante ressaltar que a evolução natural não é um processo dirigido. Desta forma, um algoritmo genético não é um algoritmo de busca pela solução ótima de um problema, mas sim uma

heurística que encontra boas soluções a cada execução, mas não necessariamente a mesma (LINDEN, 2006).

2.2.4.1 *Crossover*

O operador de *crossover* ou recombinação é o responsável por transmitir a carga genética, as características, dos pais para os filhos de modo a manter essas informações para a próxima geração e também permitir a exploração de novos espaços de soluções. Existem diversas formas de se realizar *crossover* que pode ser complexo e dependente da representação cromossômica e do tipo de problema abordado. Por isso, existem tipos recomendados para cada situação que garantem um melhor desempenho do AG (LAMBORA; GUPTA; CHOPRA, 2019).

Dentre os tipos de *crossover*, podemos citar o *crossover* uniforme que utiliza uma máscara gerada aleatoriamente com o mesmo comprimento dos cromossomos dos pais para realizar o cruzamento. Cada gene dos filhos é copiado dos pais de acordo com a máscara, se o *bit* da máscara for 1 o *bit*, é copiado do primeiro pai se for 0 do segundo. No *crossover* de três pais, três indivíduos são escolhidos aleatoriamente para gerar o gene dos filhos. Cada gene do primeiro e segundo pai são comparados; se forem iguais, o gene será copiado para o filho; se forem diferentes, será escolhido o gene do terceiro pai. O PMX (*Partially Matched or Mapped Crossover*) foi proposto por Goldberg e Lingle em 1985, para o problema do caixeiro viajante. Dois pontos são escolhidos aleatoriamente e acontece a troca dos cromossomos. Em seguida, é feita uma correção dos genes para que não haja uma repetição. O *crossover* cíclico (CX), é utilizado para problemas em que o cromossomo representa uma permutação. Neste tipo de *crossover*, há uma limitação de quais genes venham de cada pai, e é selecionado uma posição inicial aleatória no primeiro pai por onde o *crossover* começa. Este valor é procurado no segundo pai; quando é encontrado verifica-se qual o valor correspondente desta posição no primeiro. Isto acontece até que o ciclo se repita finalizando o *crossover*. O *crossover* aritmético é usado para cromossomos de representação real e combina linearmente dois pais, sendo seus descentes a mistura linear de ambos (KORA; YADLAPALLI, 2017).

O *crossover* de ponto simples e N-pontos são amplamente utilizados em diversos trabalhos, sendo o *crossover* dois pontos de corte utilizado nesta pesquisa. Eles se diferenciam pela quantidade de pontos de corte em que a troca de informações e a mistura de genes dos pais ocorre para produzir a prole.

O *crossover* de ponto simples é um dos operadores mais utilizados principalmente por sua simplicidade pois, apenas uma localidade de troca de informações denominada "ponto de corte" é escolhida dentre as possíveis. Em seguida, os pais são divididos em partes à direita ou à esquerda deste ponto e os novos filhos são compostos pelas combinação das partes, sendo o primeiro composto pelo lado esquerdo do primeiro pai e à direita do segundo; já o segundo filho, segue a ordem contrária, conforme Figura 15.

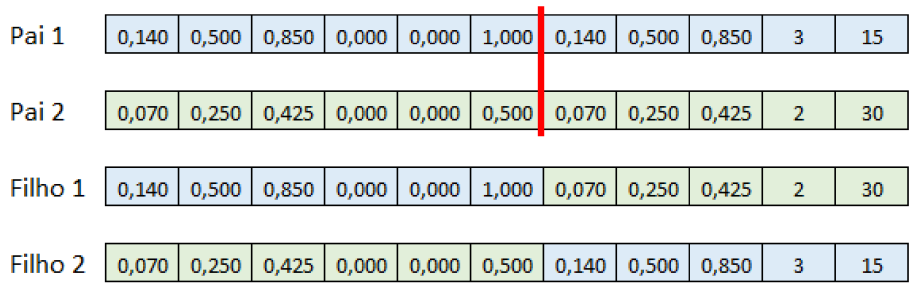


Figura 15 – Cruzamento com 1 ponto de corte.

O *crossover* de ponto duplo funciona de forma análoga ao anterior porém, são selecionados dois pontos de corte entre o primeiro e o penúltimo ponto de corte. Os filhos serão compostos pelas partes fora dos pontos de corte do primeiro pai e das partes entre os dois pontos do segundo pai para o primeiro filho e o segundo pelas partes que não foram utilizadas no primeiro, conforme Figura 16.

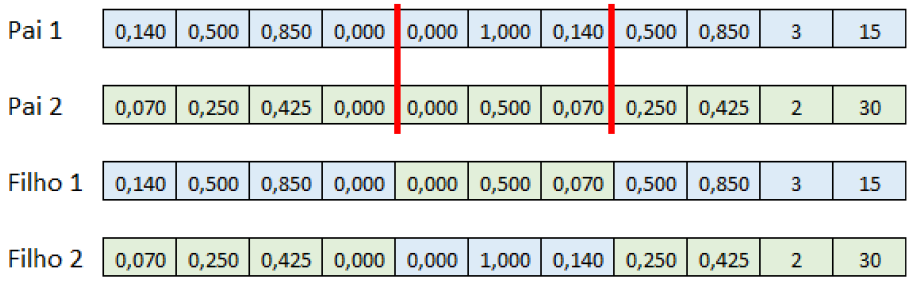


Figura 16 – Cruzamento com 2 pontos de corte.

2.2.4.2 Mutação

O operador de mutação, assim como acontece na natureza, é o responsável por introduzir pertubações nas características que compõem uma população. Este é um processo natural que não é guiado, podendo ser benéfico ou não para uma espécie dependendo do impacto que ele terá na relação do indivíduo com o ambiente. Em sistemas bio-inspirados, a mutação tem o mesmo intuito de inserir esta pertubação porém, com o propósito de varrer outros pontos no espaço de busca que podem auxiliar na busca pela solução, além de evitar que a população se torne uniforme e incapaz de evoluir (HOLLAND, 1975). Para decidir se um gene deverá sofrer mutação, há o sorteio de um valor entre 0 e 1. Se o valor for menor do que um valor previamente determinado (taxa de mutação), então um ou mais genes são alterados. Caso contrário, o indivíduo permanece inalterado. Essa ação se repete para cada um dos filhos gerados no processo de *crossover*.

Quando apenas um gene escolhido aleatoriamente é alterado, ocorre uma mutação simples; já quando para cada gene que compõe um indivíduo, existe uma probabilidade de sofrer mutação; ocorre uma mutação por gene. Dependendo, dos valores dos genes e do tipo de problema, a mutação pode ser aplicada de diferentes formas, seja rearranjando

as posições dos genes, seja alterando os valores de um gene por um fator (ABDOUN; ABOUCHABAKA; TAJANI, 2012).

Para cromossomos binários, o bit é alterado pelo seu complemento (Figura 17a) e essa foi umas das primeiras formas de mutação descritas no trabalho de Holland (1975). Se o cromossomo não permitir repetição como em problemas de permutação, a operação de mutação envolve a troca dos genes escolhidos aleatoriamente de posição (Figura 17b) no cromossomo (LIU et al., 2000). Por fim, para cromossomos com números reais ou inteiros que permitem repetição, a mutação pode envolver o acréscimo ou decréscimo de um valor aleatório, dentro de um intervalo previamente estabelecido, ao gene (Figura 17c).

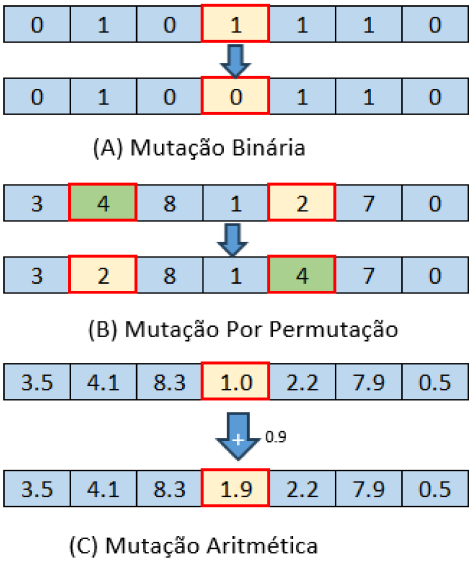


Figura 17 – Tipos de mutação.

2.2.5 Reinserção

O número de indivíduos que compõem uma população não deve ser alterado entre as gerações. Para isso é necessário definir um critério para seleção daqueles que estarão presentes na próxima iteração. Classicamente os filhos sempre substituem os pais na geração seguinte devendo ser gerados filhos suficientes para tal. Entretanto, em algumas situações seja pelo número de descendentes não ser suficiente, por decisão do projetista ou para garantir a permanência da melhor solução encontrada até o momento, é necessário manter uma parcela de indivíduos da geração anterior. Esse processo é denominado "reinserção por elitismo" (Figura 18). Na reinserção ordenada, a população de pais e filhos é ordenada pelo *fitness*, e os mais aptos são escolhidos até que seja alcançado o tamanho da população; enquanto na reinserção uniforme um procedimento de seleção é aplicado para escolher os sobreviventes (FRAGA; OLIVEIRA; MARTINS, 2020).

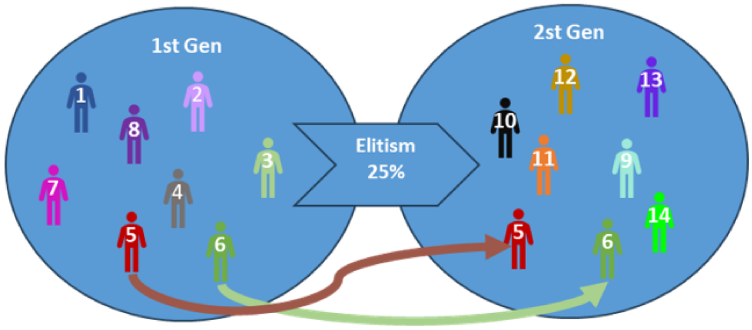


Figura 18 – Reinserção por elitismo.

2.3 Avaliação de Aptidão

Esta seção apresenta as funções de aptidão utilizadas como inspiração para as propostas nesta pesquisa. Foram utilizados dois trabalhos da literatura sendo Fit9P a função utilizada no trabalho de referência; já F1, F2 foram retiradas de outro trabalho que lida com modelos epidemiológicos.

2.3.1 Fit9P

É a função proposta no trabalho de Ferreira et al. (2022), nela em cada passo de tempo i do AC é considerada a diferença absoluta em cada partição q do valor simulado e o valor referência da base gerada e, ao final, é feita a média dos resultados obtidos ao dividir o resultado por nove que é o número total de divisões do reticulado. Na Equação 1, Ind se refere aos valores simulados com base no indivíduo produzido pelo AG e Ref é a base de referência considerada.

$$Fit9P = \frac{\sum_{q=1}^{9} \sum_{i=1}^{50} |Ind_{iq} - Ref_{iq}|}{9} \tag{1}$$

2.3.2 F1 e F2

São utilizadas no trabalho de Monteiro, Gandini e Schimit (2020) para modelos SIR (Suscetível (S), Infectado (I), Recuperado (R)) para avaliação de indivíduos imunes no processo de contágio e recuperação de varicela na Bélgica e Itália nos períodos de pré-vacinação. Em que S^* e I^* representam os valores médios da simulação e os valores alvos são S^{tar} e I^{tar}. Os valores de L_s e L_I são obtidas nas Equações 4 e 5 e o Limiar (ϵ) foi definido pelos autores como $10^5(0,00001)$.

$$F1 = \frac{1}{\sqrt[2]{L_s + L_I}} \tag{2}$$

$$F2 = \frac{1}{L_s + L_I} \tag{3}$$

$$L_S = max\left(\epsilon, \left|\log\left(\frac{S^*}{S^{tar}}\right)\right|\right) \tag{4}$$

$$L_I = max\left(\epsilon, \left|\log\left(\frac{I^*}{I^{tar}}\right)\right|\right) \tag{5}$$

2.4 Normalização de Dados

Técnicas de normalização têm sido utilizadas para melhorar o desempenho de classificadores em diferentes áreas, ao transformar ou redimensionar dados brutos de modo que tenham uma contribuição uniforme, valendo-se de medidas estatísticas como média, desvio padrão e medianas (SINGH; SINGH, 2020).

Nesta pesquisa a normalização será utilizada para combinar o *fitness* calculado para informações diferentes como quantidade de chamas, vegetação e área em recuperação em um único *fitness* de modo que seja possível avaliar o quão bom um indivíduo é em reproduzir o comportamento geral da solução. Foram avaliadas cinco formas de normalização; uma forma baseada na média e desvio padrão Z-score, três baseadas no valor mínimo e máximo (Minimax, Max, Amplitude) e uma transformação de dados pela raiz quadrada inspirada no trabalho de (MONTEIRO; GANDINI; SCHIMIT, 2020).

- Z-Score: usa a média (μ) e desvio padrão (σ) para re-escalar os dados (x) de modo que os novos valores tem uma média igual a zero e desvio padrão igual a um. Seu cálculo é dado por:

$$Z - Score = \frac{x - \mu}{\sigma} \tag{6}$$

- Minimax: re-escala os dados em um intervalo de 0 a 1 ou de -1 a 1 com base no valor máximo e mínimo, conforme Equação 7. Do valor a ser normalizado subtrai-se o valor mínimo e, em seguida o resultado é dividido pela amplitude da amostra (diferença entre os valores máximo e mínimo).

$$Minimax = \frac{x - x_{min}}{x_{max} - x_{min}} \tag{7}$$

- Max: é uma técnica variante do minimax que traz os dados para uma escala de -1 a 1, dividindo-se o valor a ser normalizado pelo máximo da amostra, como segue:

$$Max = \frac{x}{x_{max}} \tag{8}$$

- Amplitude: é uma técnica de normalização utilizada em análise de dados (ADEYEMO; WIMMER; POWELL, 2018) e aprendizagem de máquina (VENTOURAS et al., 2014), (BHANJA; DAS,) para colocar valores em uma mesma escala. O trabalho

de (JAIN, 1986) descreve essa normalização como a divisão de cada item por uma função que produz um resultado mais uniforme. Nesse caso, a função é a diferença entre o valor máximo e mínimo do *fitness*, sendo a amplitude calculada por:

$$Amplitude = \frac{x}{x_{max} - x_{min}} \tag{9}$$

- Raíz e Logarítmo Natural: Osborne (2002) descreve que transformações nos dados como o logaritmo natural e a raiz quadrada deixam os dados em uma escala única de modo que seja possível realizar sua análise. Essas normalizações são calculadas respectivamente por:

$$Raíz = \sqrt{x} \tag{10}$$

$$Logarítimico = \ln x \tag{11}$$

2.5 Trabalhos Correlatos

Nesta seção é apresentada uma revisão da literatura que inclui alguns trabalhos importantes relacionados tanto aos modelos baseados em autômatos celulares para simular a propagação de incêndios florestais apresentados na Seção 2.5.1, quanto aos algoritmos genéticos empregados para otimizar os parâmetros de tais modelos (Seção 2.5.2).

2.5.1 Modelos de Propagação de Incêndios

A abordagem proposta em Chopard e Droz (1998) é uma referência para modelos de propagação de incêndios usando autômatos celulares. É uma rede AC bidimensional composta por células de três estados e regras de transição estocásticas, que usa a vizinhança de Moore e simula a recuperação de células queimadas e a combustão espontânea.

Em Yongzhong et al. (2004), foi proposto um modelo baseado em AC com células hexagonais integradas como um modelo Rothermel para simular a propagação temporal e espacial de incêndios florestais. Dados reais de uma província chinesa e dados artificiais foram usados em experimentos que mostraram sua capacidade de reproduzir incêndios em paisagens heterogêneas.

No trabalho de Yassemi, Dragićević e Schmidt (2008), o AC foi integrado ao GIS (Sistemas de Informação Geográfica) para fornecer simulações de propagação de incêndios em que elementos topográficos, combustíveis florestais e variáveis climáticas são considerados. A proposta foi avaliada em dados reais de uma cidade canadense e os autores concluíram que ela pode ser aplicada para reproduzir incêndios realistas e outras aplicações espaço-temporais.

Usando um estudo de caso relacionado a um incêndio florestal na Ilha Spetses em 1990, um modelo de propagação de incêndio baseado em AC foi desenvolvido em Alexandridis et al. (2008), onde uma abordagem de otimização não linear de caixa preta é usada para ajustar os parâmetros de dados GIS reais. Os resultados mostraram que o evento foi bem representado e que a abordagem pode ser usada para gerenciar riscos em incêndios futuros.

Em Ghisu et al. (2015) foi apresentado um modelo de propagação de incêndio baseado em AC que aplica uma abordagem de otimização numérica para encontrar valores que correlacionam os parâmetros do modelo. De acordo com os autores, os resultados encontrados foram próximos às abordagens clássicas.

No trabalho de Xuehua et al. (2016), um modelo baseado em AC foi usado para avaliar a influência de um conjunto de fatores, como materiais combustíveis, vento, temperatura e terreno, na dinâmica das chamas. Os autores concluíram que a abordagem proposta foi capaz de simular satisfatoriamente as tendências de comportamento das chamas, auxiliando a entender os principais fatores de propagação de um incêndio florestal sob diferentes condições.

Rashid et al. (2020) desenvolveram um modelo de monitoramento de incêndios florestais utilizando modelagem computacional e drones sociais que coletam informações de redes sociais de localidades próximas a incêndios que se comunicam entre si coletando dados em tempo real do incêndio e de uma base de dados meteorológica. E essas informações são enviadas para o modelo de AC para prever o comportamento do espalhamento do incêndio.

Pavlova, Rubtsov e Telyatnikov (2020) desenvolveram um modelo de autômato celular com uma matriz triangular com obstáculos naturais e artificiais para propagação de incêndios em terrenos acidentados explorando como modelos podem descrever o comportamento de incêndios em diferentes terrenos, tendo como base um mapa de krasnodar obtido por meio de dados de sistemas geográficos. Segundo os autores o modelo desenvolvido pode ser utilizado em conjunto com outros modelos para fornecer informações sobre a propagação de incêndios, seu impacto e liberação de poluentes para tomada de decisão em situações de emergência relacionadas a previsão e espalhamento de incêndios.

No trabalho de Cerruti, Dutto e Murru (2020) foi proposta uma implementação de AC cujas regras determinam qual célula é o melhor candidato para reproduzir e o operador de *crossover* determina o resultado dos filhos. Os autores conceituam que, com esta característica o sistema pode ser tanto um Autômato Celular quanto um Algoritmo Genético, e o aplicam para reproduzir o "Jogo da Vida" e outros problemas clássicos.

Uma abordagem híbrida foi apresentada em Sun et al. (2021), em que o AC é combinado com outras técnicas que calculam um índice de propagação usando dados meteorológicos de sensoriamento remoto para ajustar o passo de evolução do autômato à propagação do incêndio. De acordo com os resultados, o modelo foi capaz de prever a

dinâmica das chamas, garantindo precisão e adaptabilidade nas simulações.

Zhang et al. (2022) utilizam a fórmula de velocidade de propagação de incêndios do modelo de Rothermel (fatores como vegetação, umidade, vento e topografia são utilizados) de forma simplificada em conjunto com autômatos celulares para produzir uma abordagem multidimensional. O trabalho buscou, desta forma produzir previsões e simulações mais precisas sendo capazes de produzir bons resultados em laboratório se aproximando dos dados de espalhamento, entretanto a diminuição de erro ao longo do tempo ainda será avaliado.

Os autores Xu et al. (2022) propõem um modelo de autômato celular baseado em aprendizado de máquina em um ambiente GIS para modelar a propagação de incêndios florestais denominado "LSSVM-CA", em que os mínimos quadrados suportam máquinas de vetores (LSSVM) são combinados com uma estrutura tridimensional de AC para incêndios florestais. O LSSVM foi utilizado para encontrar as regras de transformação para a propagação do incêndio por meio de treinamento em dados reais. Os resultados apontados pelos autores relatam que a abordagem foi capaz de simular a probabilidade e propagação de incêndios com um bom desempenho.

Meng et al. (2023) desenvolvem um modelo baseado em autômato celular para propagação de incêndios em um cenário virtual 3D que permite a visualização da propagação e extinção de um incêndio, foi criada uma biblioteca de plantas 3D onde cada célula representa uma vegetação, isto permite observar a combustão em escala unitária ou florestal. Os resultados do estudo apontaram uma melhora na imersão e realismo de uma simulação de incêndios florestais possibilitando uma melhor análise do comportamento e mudança vegetal durante a propagação de um foco de incêndio.

No estudo de Brasiel e Lima (2023), foi desenvolvido um AC bidimensional para reproduzir o comportamento de incêndios no cerrado brasileiro, tendo como parâmetros de intensidade de combustão, presença de vento, fontes de água, vegetações homogêneas e heterogêneas. No trabalho, foi avaliado como esses diferentes parâmetros podem afetar a dinâmica de espalhamento do fogo através da vegetação. Os resultados obtidos pelos autores apontam que os parâmetros avaliados afetam diretamente o espalhamento do foco especialmente a intensidade em vegetações com grande quantidade de matéria seca e menor em locais com fontes de água que atuam como uma barreira física para propagação.

Boters-Pitarch et al. (2024) propõem uma evolução de modelos anteriores baseados em autômatos celulares para simular incêndios de uma forma mais realista combinando aprendizado de máquina com o autômato celular de forma que a rede possa aprender com a vizinhança do modelo. Os dados utilizados são de uma localidade real permitindo a integração de informações climáticas e distribuição espacial do incêndio para calcular a vizinhança. A abordagem se mostrou eficiente para melhorar a precisão de simulações e espera-se que a arquitetura proposta possa ser adaptada para outros fenômenos que dependam de dados espaciais e climáticos.

Os autores Liu, Liu e Tian (2024) combinam modelos para propagação de incêndios, fluxo de ar e dinâmica de voo da aeronave e modelo de controle automático por meio de modelos físicos para simular a dinâmica de voo de aeronaves de combate a incêndios florestais. Tais modelos são integrados por meio de um *framework* para produzir informações a respeito do espalhamento de incêndios em uma determinada área para se construir um plano de combate de combate pelas aeronaves, o automato celular é usado para discretizar as área da floresta e mapear os focos de incêndio para ser utilizado nos outros modelos e produzir a saída do *framework*.

Perello et al. (2024) utilizam como caso de estudo a cidade de Ligúria na Itália e fazem, o uso da ferramenta *PROPAGATOR*, desenvolvida com um paradigma baseado em autômato celular pela CIMA (Departamento de Defesa Civil Italiana) para identificar locais em que planos de incêndios controlados podem ser aplicados para mitigar riscos de uma queimada descontrolada. Simulações de Monte Carlo foram utilizadas para avaliar a eficácia do plano de ação indicado, levando em conta material combustível, focos de incêndio e fatores meteorológicos. O trabalho buscou ressaltar a importância de modelos baseados em AC como ferramenta para o desenvolvimento de planos de combate eficientes pelo governo de Ligúria.

Para os autores Ferreira et al. (2022), um modelo 2D probabilístico baseado em AC está sendo ajustado para melhorar a qualidade das simulações de propagação do incêndio. A regra de transição de AC considera características como intensidade do vento, tempo de queima e recuperação da vegetação; em que a probabilidade de uma célula no estado de árvore começar a queimar é calculada com base nos vizinhos em chamas e na direção e intensidade do vento. O modelo de propagação de incêndio baseado em um AC probabilístico 2D é proposto por Tinoco et al. (2022). O modelo adota diferentes estados de fogo e uma função de recuperação de vegetação não linear, além de modelar características importantes que afetam a dinâmica das chamas como o tempo de queima, a intensidade do fogo, a intensidade e direção do vento e a presença de obstáculos. Os dois últimos modelos foram usados como estudo de caso nesta pesquisa. A abordagem apresentada em Ferreira et al. (2022) foi adotada como o modelo padrão a ser configurado pelo AG, enquanto o modelo proposto em Tinoco et al. (2022) foi usado para gerar conjuntos de dados de referência para os experimentos.

2.5.2 Algoritmos Genéticos no Ajuste de Parâmetros de ACs

Algoritmos genéticos são uma técnica bio-inspirada que têm sido aplicados a uma grande variedade de problemas e áreas devido a sua adaptabilidade e capacidade de produzir soluções ótimas. Como exemplo, podemos citar áreas da medicina especialmente em iniciativas de combate e identificação de infecção por coronavírus (COVID-19), bioinformática, aprendizado de máquina, por exemplo para encontrar a melhor combinação de parâmetros para uma rede, jogos, problemas de otimização, sistemas dinâmicos, proble-

mas de engenharia, calibragem de modelos, especificação de regras dentre outros. Sendo assim os parágrafos abaixo trazem uma revisão do uso de algoritmos genéticos para calibragem de modelos baseados em autômatos celulares que é o foco desta pesquisa assim como outras aplicações desta técnica.

Em Handa et al. (1997) foi proposta uma abordagem evolutiva inspirada na relação hospedeiro-parasita para busca efetiva de informações utilizando dois AGs no qual o primeiro busca por soluções no espaço de busca (hospedeiro) e o segundo (parasita) procura subestruturas que podem ter informações úteis para buscar soluções de forma mais eficaz.

No trabalho de Paredis et al. (1997) o algoritmo genético foi usado para busca de autômatos celulares que resolvem o problema de classificação de densidade e o AG co-evolui duas populações que se comportam como predador e presa. Na natureza, as interações predador-presa fornecem uma força motriz para a complexidade. Isso porque há uma forte pressão evolutiva para que as presas se defendam melhor, por exemplo, correndo mais rápido, criando escudos maiores, melhor camuflagem entre outros e os predadores para conseguirem superar estas defesas.

Para os autores Kim, Kim e Kim (2000) foi construído um algoritmo genético capaz de lidar simultaneamente com problemas de balanceamento e produção de uma variedade de produtos em uma linha de montagem. Tais problemas são fortemente correlacionados, porém na maioria dos casos, são tratados separadamente. Foi adotado uma interação localizada dentro e entre populações, e desenvolvidos métodos de seleção de parceiros simbióticos e avaliação de aptidão levando em consideração as características específicas dos problemas. Os autores chegam à conclusão que a abordagem produz bons resultados e que é mais eficiente que as abordagens já produzidas na época.

No trabalho de Cai e Peng (2002) foi desenvolvido um Algoritmo Genético Adaptativo para planejamento de caminhos de sistemas de robôs multi-móveis cooperativos usados em ambientes e tarefas complexas que não são adequadas para serem executadas por seres humanos como montagem e manufatura.

Para os autores García-Pedrajas, Hervás-Martínez e Muñoz-Pérez (2003) foi desenvolvido um novo modelo cooperativo para a evolução de redes neurais artificiais, sub-redes devem cooperar para formar uma solução para um problema específico, em vez de evoluir redes completas. A combinação dessas sub-redes faz parte de um processo co-evolutivo no qual, várias subpopulações de sub-redes co-evoluem cooperativamente e isoladas geneticamente. O indivíduo de cada subpopulação é combinado para formar redes inteiras. O desempenho das redes foi testado aplicando-as em problemas de classificação.

Monteiro et al. (2009) avalia o efeito de escala espacial na previsão de séries temporais em um estudo de caso de varicela com um modelo SIR baseado em autômato celular no Arizona (Estados Unios) entre os anos de 1994 a 2004. O AG foi utilizado para encontrar os parâmetros do modelo que varia de acordo com o raio e tamanho da matriz. Os autores fizeram experimentos com matrizes e raios distintos, com e sem restrições, cujos resultados

indicaram que o tamanho do reticulado impacta diretamente na qualidade solução, assim como restrições pois os experimentos sem restrição levaram a melhores resultados porém, esses não foram compatíveis com os dados reais.

Para Gerardi, Monteiro et al. (2011) casos de espalhamento de dengue são investigados na cidade do Rio de Janeiro no ano de 2007 através do uso de modelos SIS (Suscetível-Infectado-Suscetível) e SIR (Suscetível-Infectado-Recuperado) baseados em autômatos celulares. O algoritmo genético foi utilizado para identificar a probabilidade de transição de indivíduos suscetíveis para infectados com base na série temporal de 2007. Após isto, a combinação de AC e AG foi utilizada para prever casos de infecção do ano seguinte.

O autor Amol (2020) enfatiza que a modelagem de incêndios requer muitos parâmetros que nem sempre são conhecidos e um algoritmo genético é utilizado para encontrar esses parâmetros a partir de dados históricos e calibrar o modelo ao final da execução. A população de AG é avaliada utilizando uma métrica chamada FFDI (*Forest Fire Danger Index*), que é calculada com base em fatores de variáveis climáticas como temperatura, umidade relativa, velocidade do vento e fator de seca.

No trabalho de Monteiro, Gandini e Schimit (2020), é avaliada a influência de indivíduos imunes no processo de contágio e recuperação de varicela na Bélgica e na Itália nos períodos de pré-vacinação. O modelo epidemiológico é baseado em um AC com três parâmetros obtidos por um algoritmo genético por meio de dados históricos das duas cidades, sendo possível predizer a ocorrência da doença com um erro de dois (2) a quatro (4) por cento. Concluiu-se que a metodologia pode ser usada em outras doenças como COVID-19, mas que o papel de indivíduos imunes deve ser delimitado.

Fraga, Oliveira e Martins (2020) desenvolveram um algoritmo genético para o ajuste de parâmetros de um modelo epidemiológico da doença de Chagas em uma vila na península de Yucatam no México. Por ter apresentado bons resultados no ajuste do modelo o algoritmo foi adaptado para uma abordagem multi-estágio no trabalho de (FRAGA; OLIVEIRA; MARTINS, 2021) de modo que diferentes estágios se especializam em um conjunto de parâmetros específico produzindo uma calibragem mais precisa pois, quanto maior a quantidade de parâmetros ajustado, maior a complexidade da solução a ser encontrada pelo algoritmo genético.

Shen, Guo e Li (2020) construíram um modelo matemático para o problema de escalonamento de projeto de *software* multiobjetivo em larga escala e propõem um algoritmo genético multiobjetivo co-evolutivo cooperativo para resolver o modelo estabelecido. Foram incorporadas estratégias de decomposição de variáveis , alocação de recursos computacionais aprimorados e o otimizador de subcomponentes específicos do problema. Para teste da solução, foram realizados testes em 15 instâncias de projetos gerados aleatoriamente com até 2048 variáveis de decisão e três instâncias derivadas de projetos de *software* do mundo real.

Tirumala (2020) propõe o uso de duas técnicas de co-evolução (cooperativa e compe-

titiva) em duas populações diferentes nas etapas de treinamento e aprendizado de redes evoluindo os neurônios de modo que ao final sejam produzidas redes profundas otimizadas. Novos parâmetros foram propostos para evitar a convergência prematura e manter a diversidade genética. Os testes foram realizados em *datasets* já conhecidos para testes de redes neurais sendo eles o MNIST e o IRIS.

No trabalho de (PEREIRA; SCHIMIT; BEZERRA, 2021) foi utilizado um modelo baseado em AC para simular a propagação de uma epidemia. A partir deste modelo, foi desenvolvido um outro baseado em *deep learning* para prever a eficácia de diferentes parâmetros no comportamento do AC de modo que estes se adaptassem a dados reais. O Algoritmo Genético foi então utilizado para ajustar os parâmetros do AC utilizando o modelo substituto para avaliar cada conjunto de parâmetros produzidos.

Tinoco, Vizzari e Oliveira (2021) utilizaram algoritmo genético para otimizar os parâmetros de um modelo de coordenação de enxames de robôs que combina ACs e colônia de formigas invertidas. Os resultados obtidos mostraram que o AG foi capaz de produzir resultados melhores que ajustes empíricos do modelo feitos inicialmente pelos autores.

Os autores Ferreira et al. (2022), usaram um AG para ajustar os parâmetros de um modelo de propagação de incêndio que considera elementos como intensidade e direção do vento e cenários de vegetação homogênea e heterogênea com base em dados históricos em uma determinada área, sendo capaz de reproduzir bons resultados em diferentes experimentos.

No trabalho de Jr et al. (2023) o AG é utilizado para ajustar um modelo hidrológico-hidrodinâmico e de qualidade de água utilizando dados de uma bacia hidrográfica como descargas e quantidades de poluentes. Os resultados apontaram que o algoritmo foi capaz de encontrar um conjunto de parâmetros quase ideal para as informações observadas e que um conhecimento prévio da bacia capaz é capaz fornecer um conjunto mais preciso de ranges de parâmetros que reduza a chance de encontrar parâmetros com equifinalidade.

Na abordagem descrita por Carrillo et al. (2023) uma estratégia de duas etapas é utilizada para calibrar o modelo, em que a qualidade da solução é avaliada comparando os dados da simulação com os dados do incêndio real (queimado no incêndio real, queimado no incêndio simulado ou não queimado). Em seguida, é construída uma tabela 2×2 com quatro possibilidades (*Hits, False Alarms, Misses, Correct Negatives*). Por fim, é calculada a aptidão do indivíduo. Os autores concluíram que houve uma melhora significativa no desempenho sem comprometer o tempo de execução.

Por fim, um AG paralelo com vizinhança variável e múltiplas subpopulações é proposto por Li et al. (2024) com o objetivo de ajustar os parâmetros do modelo para otimizar as rotas de combate a incêndios e minimizar os danos resultantes da propagação simulada.

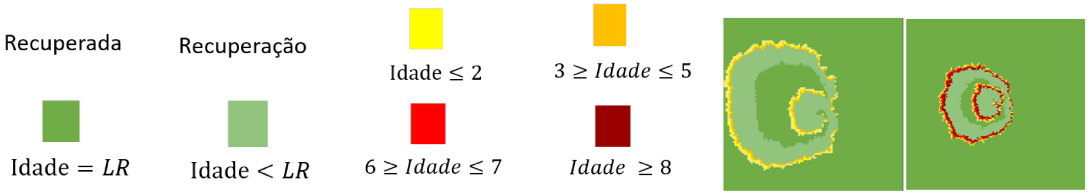
Capítulo 3

Modelos Propagação de Incêndios

Neste trabalho foram reproduzidos dois modelos de propagação de incêndios baseados em autômatos celulares, sendo o modelo de Ferreira et al. (2022) utilizado como modelo de referência, cujo os parâmetros são ajustados automaticamente pela abordagem evolutiva descrita no Capítulo 4; o modelo de Tinoco et al. (2022) foi usado para avaliar a robustez da abordagem proposta.

3.1 Modelo de Referência

O modelo de Ferreira et al. (2022) é um AC probabilístico bidimensional composto por uma rede de células de tamanho 100×100, evoluídas ao longo de 50 passos de tempo, onde, cada célula é definida pelo par $S = (Tipo, Idade)$. Cada célula pode ser do tipo árvore ou fogo. Células em chamas (tipo fogo) passam pelos estágios de queima $S_f \in \{1...LQ\}$ (Fig. 19b), onde LQ representa o tempo de queima. Células queimadas (tipo árvore) passam pelos estágios de recuperação $S_t \in \{1...LR\}$ (Fig. 19a), onde LR representa o tempo de recuperação, antes de se tornarem suscetíveis a novos incêndios. Desta forma, há $LQ + LR$ possíveis estados distribuídos em dois tipos de células. As regras de transição controlam a evolução do AC (Fig.19c), isto é, define o estado da célula central no próximo passo de tempo baseado em seu estado atual e de sua vizinhança. A condição de contorno é reflexiva, isto é, cada célula da borda é refletida para ser utilizada como parte do limite do AC.

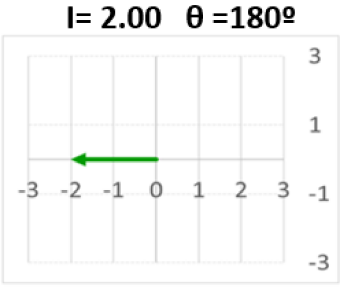


(a) Estados de Vegetação. (b) Estados de Fogo. (c) Evolução temporal.

Figura 19 – Estados do AC e evolução temporal.

N_w	N	N_e
0.140	0.500	0.850
W	C	E
0	0	1
S_w	S	S_e
0.140	0.500	0.850

(a) Matriz $W_{3\times3}$.



(b) Vetor resultante.

Figura 20 – Matriz W e vetor resultante.

Fonte: Adaptado de Ferreira et al. (2022)

.

A matriz de preferência do vento W é uma matriz quadrada 3×3, usada para quantificar a probabilidade de uma determinada célula se inflamar através da propagação do fogo de suas células vizinhas, considerando a Vizinhança de Moore com raio 1. Essa matriz representa a influência do vento, isto é, cada uma de suas oito (8) posições externas representa uma direção do vento de acordo com o plano cartesiano (noroeste, norte, nordeste, leste, sudeste, sul, sudoeste e oeste) e indica o impacto daquela célula vizinha no estado da célula central.

Cada célula da matriz W (Figura 20a) pode assumir um valor real entre [0,1], cuja composição determina a direção, intensidade e velocidade das chamas na simulação (Figura 20b). Tais valores são obtidos pelo vetor resultante da decomposição vetorial de força dos elementos da matriz, sendo sua intensidade i denotada pela Equação 12 e sua inclinação θ pelas Equações 13 e 14.

$$i = \sqrt{\left(\left(\sum \frac{W_{(ij)x}}{\sqrt{2}}\right)^2 + \left(\sum \frac{W_{(ij)y}}{\sqrt{2}}\right)^2\right)} \tag{12}$$

$$\theta_1 = arctang\left(\frac{\sum \frac{W_{(ij)x}}{\sqrt{2}}}{\sum \frac{W_{(ij)y}}{\sqrt{2}}}\right) \tag{13}$$

$$\theta_2 = \frac{\pi}{2} + |\theta_1|, \theta_3 = \frac{3\pi}{2} - |\theta_1|, \theta_4 = \frac{3\pi}{2} + |\theta_1| \tag{14}$$

A célula central $W_{2\times2}$ representa a origem do plano cartesiano, W_{ij} a célula atual da matriz e $W_{(ij)x}$, $W_{(ij)y}$ os componentes verticais e horizontais da célula. A célula central foi considerada como a origem do plano cartesiano, e cada célula adjacente foi tratada como um vetor. As componentes vetoriais foram somadas levando em conta tanto a direção quanto a intensidade. Para determinar a intensidade, utilizou-se a hipotenusa do triângulo formado pelos vetores resultantes como catetos, conforme descrito no Teorema de Pitágoras. O cálculo do ângulo θ variou de acordo com o quadrante no qual o vetor foi plotado no plano cartesiano.

3.1.1 Regras de Transição

A regra de transição do modelo de propagação de incêndios considera características da vegetação, como tempo de queima (LQ) e tempo de recuperação (LR), que representam, respectivamente, o tempo que uma célula permanecerá no estado de fogo e o tempo necessário para que uma célula queimada esteja suscetível a queimar novamente. Essa dinâmica de evolução do AC implementado segue as seguintes regras:

I. Células do tipo árvore começam com o valor máximo de LR;

II. Uma célula do tipo fogo evolui por diferentes estágios até atingir o valor máximo LQ;

III. Ao atingir o valor máximo de LQ, uma célula do tipo fogo se transforma em uma do tipo árvore, com idade igual a 1;

IV. Uma célula do tipo árvore tem probabilidade P de pegar fogo definida como:

$$P = \frac{Idade}{LR} \times \frac{\overline{IW_{fogo}}}{LQ} \tag{15}$$

$\overline{IW_{fogo}}$ é a média do produto entre o estágio de queima e a influência do vento (W) de cada vizinho ($\forall x_{ij} \, \epsilon \, \eta_{ij}^{8}$) em estado de fogo ($S_{f_{ij}}$) e é definido pela Equação:

$$\overline{IW_{fogo}} = \sum_{ij}^{\eta} S_{f_{ij}} \times W_{ij} \tag{16}$$

3.1.2 Modelo Estocástico e Determinístico

A Regra V da Seção 3.1.1 define a probabilidade de uma célula passar do estado de vegetação para o de fogo. Entretanto, há uma diferença no comportamento do modelo quando ele tem um tempo de recuperação determinístico ou estocástico, conforme apresentado nas evoluções temporais do modelo ilustradas na Figura 21.

Para o modelo com uma recuperação determinística, uma célula que tenha sofrido queima passará pelos LR estágios de recuperação até alcançar o valor máximo e estar apta a sofrer uma nova queima. Isso permite uma maior previsibilidade no tempo que uma área sofrerá uma nova queima e permite uma observação mais precisa do foco principal do incêndio e da propagação das ondas secundárias.

Por outro lado, uma recuperação estocástica permite uma adaptação a cenários mais complexos e realísticos, uma vez que uma área pode queimar novamente após uma onda recente de chamas, enquanto outra com a mesma vegetação pode demandar um tempo maior para acumular novamente material combustível. Em outras palavras, no modelo estocástico, a vegetação não precisa estar totalmente recuperada como no modelo determinístico para voltar a queimar, pois outras fontes podem servir de combustíveis para o fogo.

Essas diferenças no comportamento do modelo em função da regra de recuperação serão melhor exploradas no Capítulo 5.

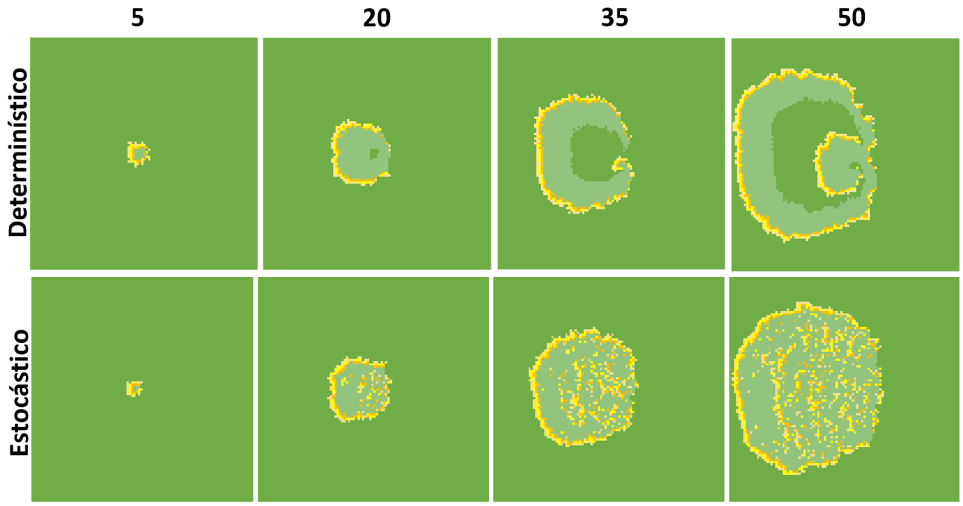


Figura 21 – Dinâmica do modelo em função da regra de recuperação da vegetação adotada (determinística ou estocástica).

Fonte: Adaptado de Ferreira et al. (2022).

3.2 Modelo de Validação

O modelo utilizado para validar a adaptabilidade e robustez da nossa abordagem evolutiva foi proposto por Tinoco et al. (2022). Ele também é um modelo de propagação de incêndios baseado em ACs bidimensionais composto por uma grade de células de tamanho 100×100, com evolução estocástica ao longo de 300 passos de tempo.

3.2.1 Matriz de Ventos

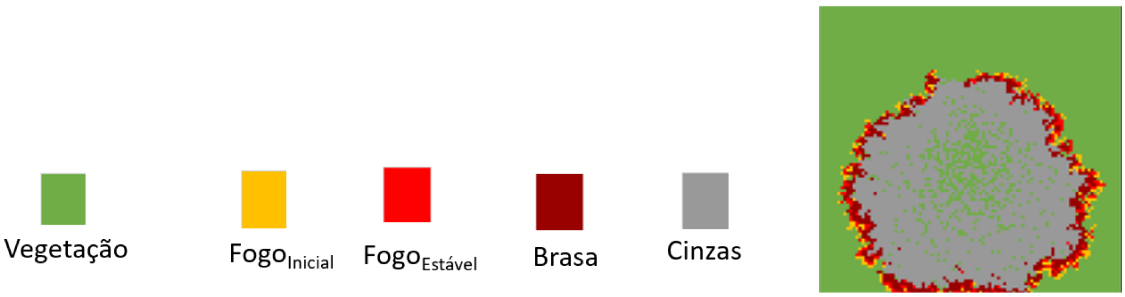
Esse modelo considera diferentes fatores que influenciam no comportamento e espalhamento de incêndios, os quais são modelados como coeficientes nas regras de transição, sendo um coeficiente calórico (λ) para representar a intensidade global das chamas e um coeficiente (δ) para a influência do vento (intensidade e direção). A intensidade local do incêndio (il) é definida como de acordo com os estágios de queima. O tempo de permanência em cada estágio de queima também varia, sendo 3 passos nos estágios de fogo inicial e estável e 10 passos no estágio de brasa.

Assim como no modelo anterior, uma matriz de combustão (WC), de dimensão 3×3, representa a influência do vento e da vizinhança em chamas na célula central atual. Essa matriz é calculada como descrito na Equação 17, onde a variável pb se refere à probabilidade base de uma vegetação sofrer queima e é determinada de acordo com cada tipo de vegetação e o decaimento que determina a intensidade do vento ao redor da célula, como se fossem sensores que informam em qual região a influência do vento é maior.

$$WC = \begin{bmatrix} \lambda \times (pb - (d \times 1)), & \lambda \times (pb - (d \times 0)), & \lambda \times (pb - (d \times 1)) \\ \lambda \times (pb - (d \times 2)), & 0, & \lambda \times (pb - (d \times 2)) \\ \lambda \times (pb - (d \times 3)), & \lambda \times (pb - (d \times 4)), & \lambda \times (pb - (d \times 3)) \end{bmatrix} \tag{17}$$

3.2.2 Regras de Transição

Os possíveis estados (S) de cada célula $S \in \{Vegetação, Fogo, Cinza\}$, são mostrados na Figura 22a. Células de vegetação com material combustível completo são representadas em verde e os estados seguintes são do tipo Fogo, sendo $Fogo_{inicial}$ (amarelo), $Fogo_{estável}$ (vermelho) e $Brasa$ (vermelho escuro). Após passar por todos os estágios de Fogo, a célula se torna $Cinzas$ (cinza). Nesse estado, o fogo se extingue e não se espalha para outras células nem entra em ignição novamente, a menos que a célula tenha se recuperado e entrado em combustão novamente. Neste modelo, é proposto uma função de recuperação não linear baseada em uma função probabilística exponencial para que células entrem em ignição, descrita na Equação 19 . O valor de $\alpha = 6$ utilizado na equação de recuperação responsável pela distribuição dos dados ao longo dos passos de tempo, foi encontrado pelos autores por meio de experimentos preliminares.



(a) Estados das células. (b) Evolução temporal.

Figura 22 – Possíveis estados e evolução temporal modelo de validação.

Fonte: Adaptado de Tinoco et al. (2022)

No modelo implementado, a dinâmica dos incêndios é determinada pela seguinte regra de transição do AC:

I. Células no estado vegetação sem vizinhos no estado fogo mantém-se nesse estado;

II. Se o estado da célula é vegetação; e algum de suas vizinhas está no estado fogo, há a probabilidade P de mudar para fogo conforme:

$$P_{Fogo_{inicial}} = WC \times il \times \lambda; \tag{18}$$

III. Se a célula está no estado $Fogo_{inicial}$, o estado se mantém por Tf_i passos de tempo e, no passo seguinte muda para $Fogo_{estável}$;

IV. Se a célula está no estado $Fogo_{estável}$, o estado se mantém por Tf_e passos de tempo e, no passo seguinte muda para $Brasa$;

V. Se a célula está no estado $Brasa$, o estado se mantém por Tb passos de tempo e, no passo seguinte muda para $Cinzas$;

VI. Se a célula está no estado $Cinzas$, ela pode mudar para o estado de vegetação de acordo com a probabilidade de recuperação, definida da seguinte forma:

$$P_r(x_{ij}) = \begin{cases} 0.0, & Se\ Ocioso \\ \frac{(ts_r)^2}{10^a}, & Senão\ se\ ts_r \geq 1\ e\ \alpha \geq 1 \end{cases} \tag{19}$$

Capítulo 4

Abordagem Evolutiva Proposta: EVOCA-WiSM

Este capítulo apresenta uma abordagem evolutiva utilizando algoritmos genéticos (AGs), denominada como *Evolutionary Optimization of Cellular Automata-based Wildfire Spreading Models (Otimização evolutiva de modelos de propagação de incêndios florestais baseados em autômatos celulares) (EVOCA-WiSM)*, que tem como propósito realizar o ajuste dos parâmetros do modelo de propagação de incêndios descrito na Seção 3.1 para produzir adequadamente o comportamento de espalhamento do fogo observado nas amostras do conjunto de referência, o qual representa dados históricos de incêndios na região a ser monitorada. Inicialmente, é apresentada a estrutura de representação dos indivíduos do AG, que codifica uma possível solução para o problema. Em seguida, são descritas as diferentes formas de avaliação investigadas durante a pesquisa. Por fim, são especificados os métodos e operadores genéticos utilizados no algoritmo genético, bem como os valores dos parâmetros de configuração adotados na abordagem evolutiva.

4.1 Representação do Indivíduo

Cada indivíduo é uma configuração possível para o modelo baseado em AC composto por 11 genes, conforme mostrado na Figura 23. Os primeiros oito genes têm valores reais entre $[0, 1]$ representando a matriz de vento W que compõe a vizinhança de Moore (azul). Como uma analogia a rosa dos ventos, cada um dos oito genes representa a influência de uma direção do vento na célula central, sendo respectivamente noroeste - GNw, norte - GN, nordeste - GNe, oeste - GW, leste - GE, sudoeste - GSw, sul - GS e sudeste - GSe. A célula central (cinza) é sempre considerada zero, portanto não é incluída no cromossomo. Os próximos dois genes são valores inteiros representando o valor de queima LQ com valores entre $[0, 10]$, o valor de recuperação LR variando entre $[0, 1000]$.

O último parâmetro *Sample Rating* (Taxa de Amostragem) (SR), também é um inteiro pertencente ao intervalo $[0, 10]$. Este parâmetro é a principal contribuição desta pesquisa;

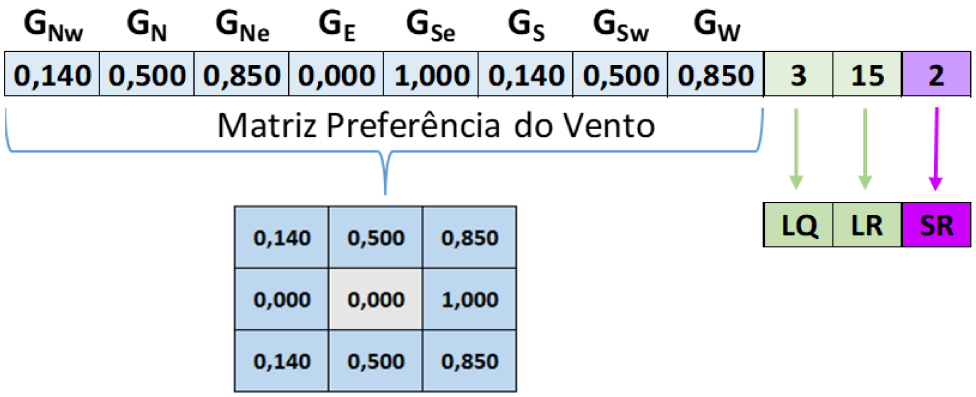
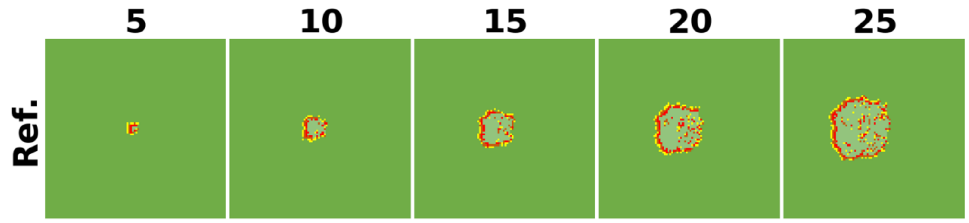


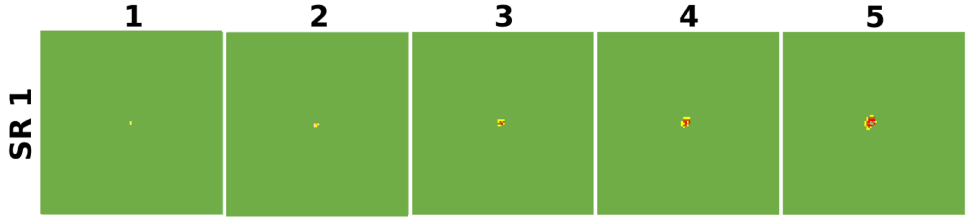
Figura 23 – Cromossomo

ele permite uma maior adaptabilidade do algoritmo genético a diferentes entradas, independente do tamanho e do intervalo entre cada registro. Em uma abordagem tradicional o período de amostragem é sempre unitário, isto é, o registro é sempre comparado a cada passo de tempo do AC para verificar a qualidade da reprodução, entretanto nem sempre os registros estarão armazenados dessa forma, ou conterão todos os dados.

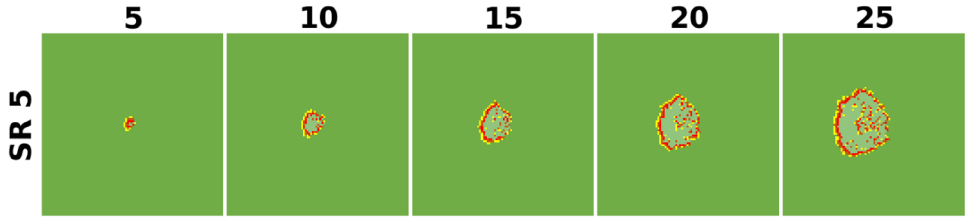
Considere um conjunto de referência hipotético com 5 amostras de um incêndio (Figura 24a), cujos registros foram gravados com frequência de 5 passos de tempo. Em seguida, executou-se o algoritmo utilizando-a como entrada. Desta execução, foram selecionadas evoluções com diferentes taxas de amostragem $SR = \{1, 5, 10\}$. Os resultados são apresentados na Figura 24, onde a numeração acima do reticulado indica o passo de tempo do AC e ao lado a taxa de amostragem encontrada pelo AG. Note que um SR igual a 1 (Figura 24b) equivale à abordagem tradicional sem a taxa, neste cenário como cada registro da base foi gravado com $p = 5$ a e evolução temporal produzida no primeiro reticulado é muito diferente do esperado, pois não houve tempo de evolução suficiente para que os reticulados pudessem ser comparados e a solução devidamente avaliada. Por outro lado, quando a taxa é bem determinada, o AG é capaz de identificar a evolução do fenômeno e apresentar uma evolução bem similar àquela presente no conjunto de dados de referência, como o mostrado na Figura 24c. Por fim, se o valor da taxa for muito alto (Figura 24d), o algoritmo irá superestimar a evolução temporal e produzir um espalhamento muito maior do que o desejado. Essa diferença de evolução temporal aumenta a cada amosta do conjunto de dados. Um exemplo seria comparar a segunda amostra da Figura 24b que está no segundo passo do AC com a segunda amostra da referência que esta no passo 10. Identificar o padrão de amostragem adequado para reproduzir o comportamento de espalhamento presente na base de referência foi o grande desafio solucionado com a inserção deste novo parâmetro ajustado pelo AG em conjunto com os demais.



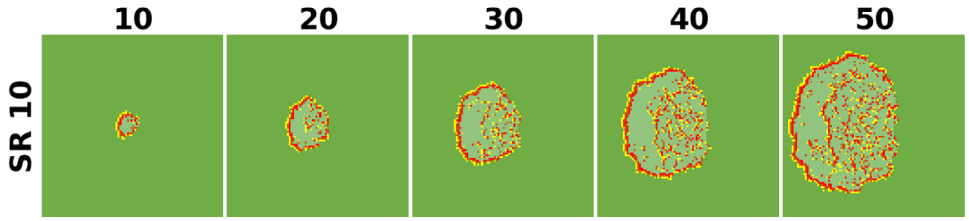
(a) Referência ($p = 5$).



(b) Taxa de amostragem 1.



(c) Taxa de amostragem 5.



(d) Taxa de amostragem 10.

Figura 24 – Impacto da taxa de amostragem na avaliação

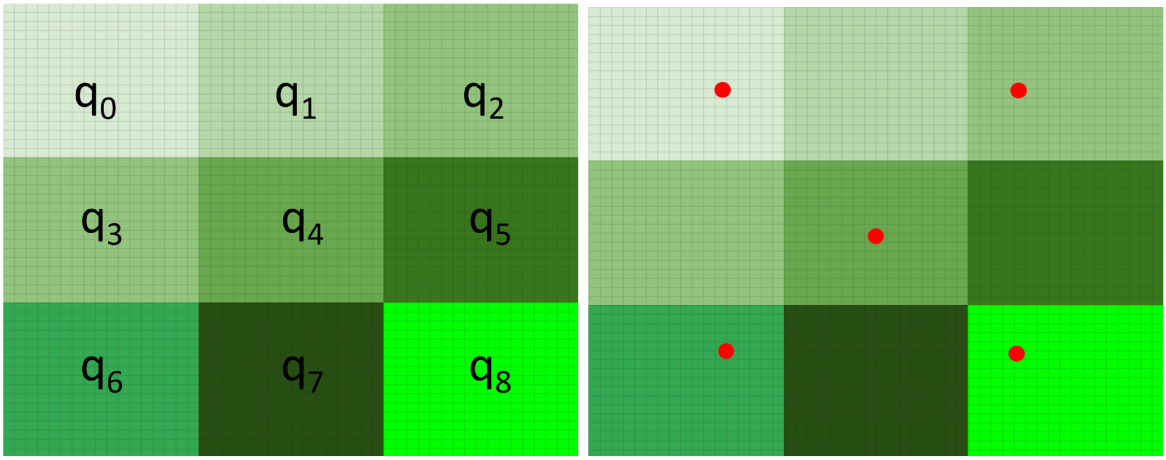
.

4.2 Funções de Avaliação

O cálculo de *fitness* de cada indivíduo é feito através da simulação do AC utilizando o cromossomo definido pelo AG. Com o intuito de melhorar a captura da dinâmica de propagação do incêndio ao longo do reticulado, ele é dividido em nove partições (Figura 25a). Cada função de avaliação compara os resultados da simulação produzidos utilizando os genes do indivíduo em questão em cada uma das nove partições com a base de referência do cenário escolhido.

Um detalhe importante são os focos iniciais de incêndio, uma vez que, a depender da partição que se inicia um foco e da direção do vento, esse pode facilitar ou dificultar a identificação de qual o sentido de propagação do incêndio. Por exemplo, incêndios que têm o foco no meio do reticulado ($Q4$) são mais fáceis de identificar, pois, tem todo o reticulado para propagar independente da direção. Já incêndios que começam nas extremidades, como $Q0$, atingem a borda do reticulado mais rapidamente o que dificulta

a calibragem dos parâmetros. A fim de contornar essa situação, foram inseridos mais focos de incêndio, sendo eles $(24, 24), (24, 74), (74, 24), (74, 74), (49, 49)$ conforme a Figura 25b. Como o AC é estocástico, mesmo usando a igual combinação de parâmetros, cada execução pode produzir reticulados diferentes para a mesma evolução temporal e, a fim de diminuir o erro relacionado a essa aleatoriedade, cada ponto é executado 5 vezes e o resultado por partição é o valor médio das execuções.



(a) Partições. (b) Focos iniciais.

Figura 25 – Divisão reticulado e focos iniciais incêndio.

Dado um conjunto de dados de referência composto por N instantes de propagação de fogo, essas amostras podem ser geradas por outros modelos de propagação ou mesmo extraídas de uma sequência de imagens obtidas de um incêndio do mundo real. O número de passos de tempo T necessários no modelo de simulação que serão usados para reproduzir o comportamento presente no conjunto de dados de referência é dado por $SR \times N$. Um conjunto de T reticulados será gerado a partir da evolução temporal do modelo AC utilizando os parâmetros codificados no indivíduo da população. O parâmetro SR representa o intervalo de passos de tempo que será aplicado à amostragem do conjunto de reticulados usados para avaliar a similaridade da simulação em relação aos registros de propagação de incêndio presentes no conjunto de dados de referência. Por exemplo, se SR for 5 e o conjunto de dados de referência tiver $N = 10$ registros de uma propagação de incêndio florestal, o conjunto de aptidão usado para avaliar o indivíduo será composto pelos reticulados do AC nos passos de tempo em {5, 10, 15, 20, ..., 50} ($T = 50$).

A Figura 26 apresenta como é feita a interação entre o AC e AG. O cromossomo selecionado para avaliação é usado como entrada para o AC, em que, os 8 primeiros genes irão compor a matriz W que representa a matriz de influência do vento, e os seguintes o tempo de queima LQ e e recuperação LR. $Steps(S)$ representa o tempo máximo de evolução do AC e é definido como 50 passos. As regras de transição são definidas na Seção 3.1.1 e, indicam como cada célula irá alterar seu estado ao longo do tempo. A taxa de amostragem (SR) dita o intervalo em que cada reticulado é selecionado para avaliação

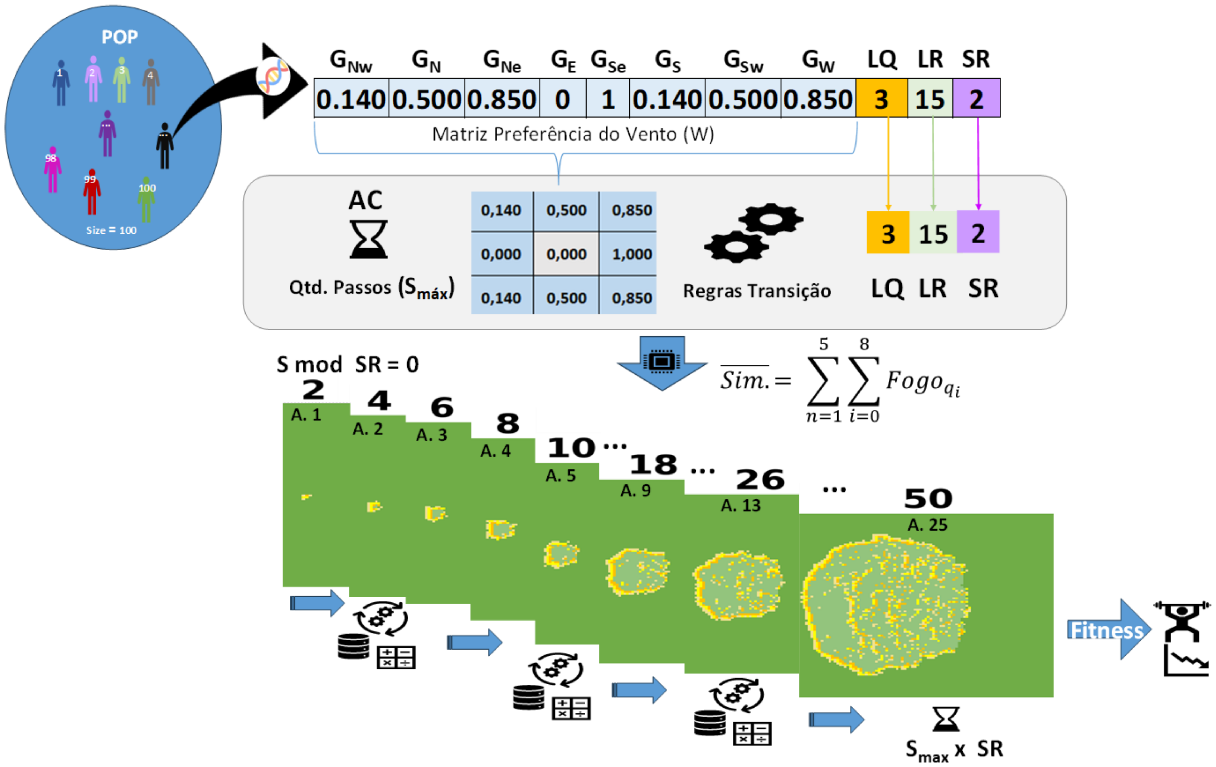


Figura 26 – Interação entre autômato celular e o algoritmo genético

conforme a operação de módulo do estado atual S pelo valor do parâmetro SR (S mod $SR = 0$). Se o resultado da operação for igual a zero se avalia este reticulado, caso contrário a evolução do autômato continua até que o processo avaliar todos os reticulados presentes na base de dados. A quantidade de amostras depende do número de passos de tempo, taxa de amostragem e da quantidade de registros mínima definida para a base (nos experimentos). Desta forma, a Amostra 1 (A. 1) não corresponde exatamente ao primeiro passo de tempo do AC, mas do primeiro reticulado que obedece ao critério de seleção (passo de tempo igual a 2) e o processo se repete até que tenham sido analisadas todas as entradas.

Se a análise for do modelo de referência em relação aos dados produzidos por ele, considera-se a quantidade máxima de passos $S_{máx}$ e o intervalo SR entre cada registro para a reprodução do comportamento. Já para casos em que não se conhece a quantidade de passos, isto é, a entrada pode ser proveniente ou de outro modelo ou de um incêndio real se multiplica o valor da taxa de amostragem pela quantidade máxima de passos do modelo de referência $S_{máx} \times SR$. Desta forma, o AG é capaz de reproduzir o comportamento descrito na base além de definir taxa de amostragem mais adequada e a quantidade de passos que devem ser considerados na análise.

Os detalhes de como realizar o cálculo de aptidão utilizando a taxa de amostra descritos estão sintetizados no Algoritmo 1. Após inicializar as variáveis (linhas 1 a 6), o passo seguinte é a execução do AC utilizando os genes do indivíduo em análise (linhas 7 a

9), durante o processo é contabilizado a quantidade de células que estão em chamas ou se recuperando em cada partição; este processo é realizado 5 vezes e a média dessas execuções (linha 10) é comparada com a base de referência no cálculo de aptidão na linha 15. Em cada partição se avalia todos os reticulados selecionados de acordo com a taxa de amostragem (linha 14) até se atingir o tamanho máximo da base de referência. O valor desse cálculo por cada partição é armazenado (linha 15) para compor o valor final de *fitness*, que considera todas as partições, obtendo assim a aptidão em relação ao reticulado completo.

Algoritmo 1 Cálculo de Aptidão

```
1: var MaxPassos ← 50
2: var passos ← MaxPassos × SR
3: var size ← Sizeof(Dataset)
4: var particao ← 8
5: var q[particao][size]
6: var grid[100][100]
7: for times ← 0, times++, while times < 5 do
8:    grid ← grid + Model(genes) {Obtem o reticulado com os genes do A! (A!)}
9: end for
10: grid ← AVG(grid) {Obtém a média das 5 execuções}
11: for i ← 0, i++, while i < particao do
12:    var S ← 0
13:    for d ← 0, d++, while d < passos do
14:       if d mod SR = 0 then
15:          q[i][d] ← F(grid[i][d], Dataset[i][d]) {Aptidão no Intevalo}
16:          S++
17:          if S = size then
18:             BREAK
19:          end if
20:       end if
21:    end for
22: end for
23: for s ← 0, s++, while s < size do
24:    fitness ← (q[0][s] + ... + q[7][s])/X {X é definido com função adotada }
25: end for
```

A seguir, são mostradas as funções consideradas na avaliação de cada indivíduo; Fit9P proposta no trabalho de referência (FERREIRA et al., 2022) e considera a diferença da simulação com relação à referência em cada uma das 9 partições. Esta ideia de divisão do reticulado foi combinada com as funções F1 e F2 do trabalho de (MONTEIRO; GANDINI; SCHIMIT, 2020) para produzir as funções F1P9 e F2P9. Devido ao fato de F1 e F2 serem utilizadas em modelos epidemiológicos, o conceito de SIR (Suscetível, Infectado, Recuperado) foi feito um um paralelo para incêndios ao considerar fogo como infectado, vegetações que sofreram queima como em recuperação e vegetação intacta como suscetíveis.

4.2.1 F1P9 e F2P9

F1P9 e F2P9 são funções inspiradas em modelos SIR (Suscetível, Infectado, Recuperado), pois, além da quantidade de fogo (indivíduos infectados), outros fatores são considerados como vegetação ainda não queimada (indivíduos suscetíveis) e vegetação que sofreu queima e está se regenerando (indivíduos recuperação). Os valores de Fs se referem aos valores de fogo simulados pelos parâmetros do AG, Fr é a quantidade de focos de incêndio na referência, de modo análogo Vs é a vegetação simulada sem queima e Vr a de referência, Qs é a vegetação simulada que sofreu queima e está em recuperação e Qr a de referência. O epsilon (ϵ) atua como um filtro pois é o limiar mínimo de que os valores podem assumir no cálculo das funções. Foram feitos testes para encontrar o valor mais adequado para o problema (0,1). Esse valor também foi utilizado para substituir o denominador em casos que é zero e produz um *log* infinito.

O valor de F, Q, V é encontrado nas Equações 22, 23 e 24 dado pelo valor máximo do limiar ϵ e o módulo do *log* da divisão do valor simulado pelo de referência. Uma vez encontrados os valores de F, Q e V em F1P9, eles são utilizados conforme a Equação 20 para produzir o *fit* ao tirar a raiz quadrada da soma dia a dia em cada partição. No final, o valor médio é obtido dividindo-o por nove que é a quantidade de partições. Para F2P9, na Equação 20 o cálculo é feito de modo análogo a F1P9, porém, sem a raiz quadrada.

$$F1P9 = \frac{\sum_{q=1}^{9}\sum_{i=1}^{50}\sqrt[2]{F_{iq}+Q_{iq}+V_{iq}}}{9} \tag{20}$$

$$F2P9 = \frac{\sum_{q=1}^{9}\sum_{i=1}^{50}F_{iq}+Q_{iq}+V_{iq}}{9} \tag{21}$$

$$F = max\left(\epsilon, \left|log\left(\frac{Fs}{Fr}\right)\right|\right) \tag{22}$$

$$Q = max\left(\epsilon, \left|\log\left(\frac{Qs}{Qr}\right)\right|\right) \tag{23}$$

$$V = max\left(\epsilon, \left|\log\left(\frac{Vs}{Vr}\right)\right|\right) \tag{24}$$

4.2.2 Funções Normalizadas Pelo Máximo

Como ao se considerar diferentes estados do AC no cálculo de aptidão produz *fitness* com diferentes dimensões, se avaliou técnicas de normalização e padronização (Z-Score, Minimax, Max, Amplitude, Log, Raíz) para unificar os resultados dos diferentes estados em um único valor de *fitness*. O melhor resultado obtido foi utilizando-se a normalização pelo valor máximo dos estados de estados Fogo (F), Queimado(Q) e Vegetação (V) em cada partição.

A função função foi inspirada na F1P9 explicada na seção anterior e desta maneira, o cálculo de estados Fogo (F), Queimado(Q) e Vegetação (V) se mantém. Entretanto, a forma de combinar os valores de erro em cada dimensão são diferentes porque, ao invés da raíz, utilizamos a normalização pelo valor máximo. Assim sendo para cada quadrante, o valor máximo de cada dimensão é utilizado para normalizar o valor.

Para a função Fit9PMax (Equação 25) continua-se realizando a diferença entre simulação e referência porém para cada estado analisado de normaliza e ao final e feita a média do valor para as 9 partições. Para FitMax (Equação 26) e Fit2Max (Equação 27) o cálculo dos estados de Fogo, Recuperação e Vegetação (Equações 22, 23, 24) se mantém. Entretanto para FitMax ao final, os valores são somados e a média entre os os valores é feita. D se refere ao valores utilizados. Por exemplo, se forem considerados apenas fogo e queimado, D será igual a dois; se foram considerados fogo, vegetação e queimado, D será igual a três.

$$Fit9PMax = \frac{\sum_{q=1}^{9}\sum_{i=1}^{50}\left|\frac{F_{iq}-Ref_{iq}}{QMAXF_q}+\frac{Q_{iq}-Ref_{iq}}{QMAXQ_q}+\frac{V_{iq}-Ref_{iq}}{QMAXV_q}\right|}{9} \tag{25}$$

$$FitMax = \frac{\sum_{q=1}^{9}\sum_{i=1}^{50}\frac{F_{iq}}{QMAXF_q}+\frac{Q_{iq}}{QMAXQ_q}+\frac{V_{iq}}{QMAXV_q}}{D} \tag{26}$$

$$Fit2Max = \sum_{q=1}^{9}\sum_{i=1}^{50}\frac{F_{iq}}{QMAXF_q}+\frac{Q_{iq}}{QMAXQ_q}+\frac{V_{iq}}{QMAXV_q} \tag{27}$$

Capítulo 5

Experimentos e Análise dos Resultados

Neste capítulo, são descritos os experimentos realizados para o ajuste do modelo de propagação de incêndios baseado em Autômato Celular utilizando uma abordagem evolutiva baseada em Algoritmos Genéticos [1]. A Seção 5.2 descreve o processo de geração das bases de dados sintéticas utilizadas no processo experimental e diferentes dados foram gerados de acordo com o objetivo do experimento. Para permitir a visualização do comportamento médio de uma solução, a frequência de queima em uma área é avaliada por meio de uma abordagem de mapa de calor descrito na Seção 5.3. Os critérios para avaliação da abordagem são descritos na seção 5.4.

Devido ao fato do modelo de propagação proposto por Ferreira et al. (2022), utilizado como referência, possuir um comportamento diferente para o cenário de recuperação determinística e probabilística, a fim de avaliar essa influência na calibragem do modelo e reprodução do fenômeno de propagação, os experimentos foram divididos em duas seções: uma com ênfase na recuperação determinística (Seção 5.5) e outra com enfoque na probabilística (Seção 5.6).

Para essas Seções, foram feitos experimentos para avaliar qual função e base de dados mais adequadas, de Sensibilidade em que se avalia a capacidade de adaptação do AG a diferentes entradas sem a perda de qualidade no ajuste e de Robustez, cujo objetivo é avaliar a capacidade de generalização do algoritmo ao utilizar informações provenientes de outro modelo de propagação baseado em AC.

5.1 Configuração Experimental

Esta seção apresenta o detalhamento dos valores utilizados para configurar os parâmetros do AC (Seção 5.1.1) e AG (Seção 5.1.1) utilizados na execução dos experimentos.

[1] O código fonte e o conjunto de dados usados nos experimentos estão disponíveis no repositório https://github.com/lucasmurilo/Evol-Adj-of-a-CA-based-Model-for-WF-SD.git

5.1.1 Hiperparâmetros do AC

Foram feitos experimentos para determinar a melhor combinação dos parâmetros para os autômatos celulares, de modo que as chamas alcancem a bordas do reticulado em tempo suficiente para que o AG seja capaz de reconhecer os padrões de espalhamento, pois uma vez que as chamas sejam muito intensas e atinjam o limite muito rápido o algoritmo pode não ser capaz de encontrar uma boa solução. De modo similar se as chamas se espalharem de forma muito lenta serão necessários muitos passos de tempo tornando a solução mais custosa computacionalmente.

Para o modelo utilizado como referência (FERREIRA et al., 2022), a matriz utilizada $W_{3\times3} = \{\{0.140, 0.500, 0.850\},\{0.000, 0.000, 1.00\}, \{0.140, 0.500, 0.850\}\}$ produz ventos com uma **I!** (**I!**) igual a 2.0 e inclinação $\theta = 180°$. O tempo de queima LQ foi definido como 3 em todos os experimentos. Já o tempo de recuperação LQ foi definido como 15 para a Seção 5.5 com a estratégia de recuperação determinística e para Seção 5.6 com recuperação probabilística esse valor foi definido como 30.

Para o modelo utilizado como validação (TINOCO et al., 2022), nos experimentos de robustez adotou-se a seguinte configuração para os parâmetros:

- **pb!** (**pb!**): 0,16
- **d!** (**d!**): 0,03
- Intensidade Global das Chamas - Caloria (λ): 1,0
- **il!** (**il!**):
  - $Fogo_{inicial}$: 0,6
  - $Fogo_{estável}$: 1,0
  - $Brasa$: 0,2
- Estágios de Queima $Fogo_{inicial}$ (Tf_i): 3
- Estágios de Queima $Fogo_{estável}$ (Tf_e): 3
- Estágios de Queima $Brasa$ (Tb): 10
- Constante de Recuperação (α): 6

Para o teste de adaptabilidade (Seção 5.6.4) o valor de lambda foi definido como 25% mais intenso ($\lambda = 1.25$) para as direções nordeste e oeste. Na direção sul este valor foi 25% menos intenso ($\lambda = 0.75$) logo são necessários mais passos de tempo para o espalhamento do incêndio.

5.1.2 Hiperparâmetros do AG

Para determinar a melhor combinação de hiperparâmetros do AG, foram realizados experimentos preliminares, partindo-se dos valores adotados no artigo original (FERREIRA et al., 2022). Nesses experimentos, fixava-se alguns parâmetros enquanto outros eram variados, a fim de determinar o melhor; ao encontrar o valor mais adequado, este era mantido e se alterava outros hiperparâmetros até se determinar os melhores para o problema de reprodução da propagação de incêndios florestais. Foram avaliados crossovers de ponto simples e duplo, mutação por indivíduo e por gene, reinserção por elitismo e ordenado, seleção por torneio com *tour* 2, 3 e 4. Após a selecão das melhores técnicas se avaliou quais taxas produziam a melhor solução.

O algoritmo genético tem como critério de parada a execução de 100 gerações e uma população inicial de 100 indivíduos, gerada em primeiro momento de forma aleatória e, com o intuito de promover uma maior diversidade, algumas técnicas foram avaliadas. Como primeira abordagem, manteve-se os parâmetros da matriz de vento gerados de forma aleatória e os parâmetros de LQ e LR foram variados em faixas. Para LQ, foram definidas três faixas variando de 1 a 4, 5 a 7, e 8 a 10 e, para LR foram definidas quatro faixas 1 a 25, 26 a 50, 51 a 75 e 76 a 100. Foi fixado uma faixa de LQ e gerados seis indivíduos para cada faixa de LR correspondente, tendo um total de 72 indivíduos gerados com uma distribuição normal e, os 28 indivíduos restantes foram gerados utilizando valores aleatórios de LQ e LR.

Como a primeira abordagem obteve bons resultados, a segunda abordagem buscou conseguir uma distribuição mais ampla para valores de LQ de modo que exista um valor de LQ para cada faixa de LR, além de buscar uma distribuição mais uniforme das matrizes de vento garantindo que haja representatividade para cada elemento da rosa dos ventos. Nesta abordagem, os valores de LR foram divididos em dez faixas 1-100, 101-200, 201-300, 301-400, 401-500, 501-600, 601-700, 701-800, 801-900, 901-1000 em que cada valor de LQ será gerado em cada uma das faixas de LR. Os genes que correspondem à matriz de vento podem assumir valores entre $[0, 1]$, sendo assim, o intervalo foi dividido inicialmente em quatro faixas e, em seguida, em cinco faixas. Esta divisão foi a que obteve os melhores resultados sendo 0-200, 201-400, 401-600, 601-800, 801-1000.

A seleção dos pais para a próxima geração é feita por uma seleção com torneio simples de $Tour = 3$, o *Crossover* é feito por ponto duplo com probabilidade de 90% pelo qual, são selecionados dois pontos de corte aleatoriamente sem repetição em que o material genético dos dois pais é trocado gerando dois filhos. A mutação realizada é por gene e, para cada um, é feito o sorteio de uma probabilidade e, quando o valor sorteado é menor que 20%, o gene é mutado (Equação 28). Para genes com valores reais, é selecionado um valor aleatório por meio de uma distribuição uniforme no intervalo de $[-0.2, 0.2]$, o valor atual do gene é então incrementado ou decrementado de acordo com o valor sorteado. Para os valores inteiros, ocorre um incremento ou decremento de 1.

$$valor\_mutado = valor\_atual + incremento \tag{28}$$

A reinserção é feita de forma pura com um elitismo 25%, isto é, ao final de cada geração são mantidos os 25 melhores pais da geração atual e o restante da população é preenchida pelos 75 melhores filhos dentre os 90 gerados no processo de *crossover*. A configuração descrita é apresentada de forma resumida a seguir:

- ❏ Tamanho População: 100
- ❏ Gerações: 100
- ❏ Seleção Pais: Torneio Simples Tour=3
- ❏ *Crossover* de dois pontos: 90%
- ❏ Mutação por Gene: 20%
- ❏ Elitismo: 25%

5.2 Geração das Bases de Dados

Dado um conjunto de dados de referência composto por N instantes de propagação de fogo, essas amostras podem ser geradas por outros modelos de propagação de fogo ou mesmo extraídas de uma sequência de imagens obtidas em uma propagação do mundo real.

Usamos dados artificiais gerados por simulações de incêndio como conjunto de dados de referência. Os experimentos iniciais usam o mesmo modelo para gerar esses conjuntos de dados. No experimento subsequente, um modelo de espalhamento baseado em AC diferente foi usado para gerar os dados de referência. Para ambos os casos, avaliamos a capacidade do AG de especificar um conjunto adequado de parâmetros para reproduzir o comportamento geral observado nas imagens capturadas da área durante a evolução de um incêndio.

Inicialmente, o AC descrito na Seção 3.1 foi reproduzido. Ele foi executado 100 vezes usando os parâmetros da matriz W correspondentes ao vento para o oeste, $LQ = 3$ e $LR = 15$ em 50 passos de tempo. O valor médio para cada passo de tempo foi computado considerando as 100 simulações de AC e *datasets* (DS) de referência com tamanhos diferentes foram criados. O primeiro *dataset* armazena os valores médios relacionados a todos os 50 passos de tempo. O segundo banco de dados é composto por apenas 25 valores médios em que, foram armazenados apenas os passos de tempo pares $\{2, 4, 6, ..., 50\}$. De forma semelhante, outros quatro bancos de dados foram construídos para armazenar 16, 10, 7 e 5 valores médios. Assim, conjuntos de dados de referência (DS) foram confeccionados usando períodos de amostragem distintos $p = \{1, 2, 3, 5, 7, 10\}$, resultando em

diferentes tamanhos de conjuntos de dados $N = \{50, 25, 16, 10, 7, 5\}$, respectivamente. Esses conjuntos de referência foram chamados de *Dataset* i ou DSi em que, i indica o i-ésimo elemento de p e S. Embora esses conjuntos de dados caracterizem o mesmo cenário de incêndio florestal, uma vez que foram construídos com base em grupo igual de 100 simulações do modelo de AC, eles representam os dados de um mesmo fenômeno com frequência de amostragem diferente. Generalizando para um incêndio florestal do mundo real, as imagens de propagação são armazenadas com frequência de captura diferente.

Para o experimento seguinte, a fim de avaliar a robustez da abordagem, um segundo modelo de propagação, descrito na Seção 3.2, foi utilizado para gerar três conjuntos de dados com 20 registros (em cada um) executando-o com diferentes taxas de amostragem $SR \in \{5, 7, 10\}$, sendo nomeados Conjunto de dados 7 (DS7, SR=5), Conjunto de dados 8 (DS8, SR=7) e Conjunto de dados 9 (DS9, SR=9), considerando-se a média de 100 simulações do CA.

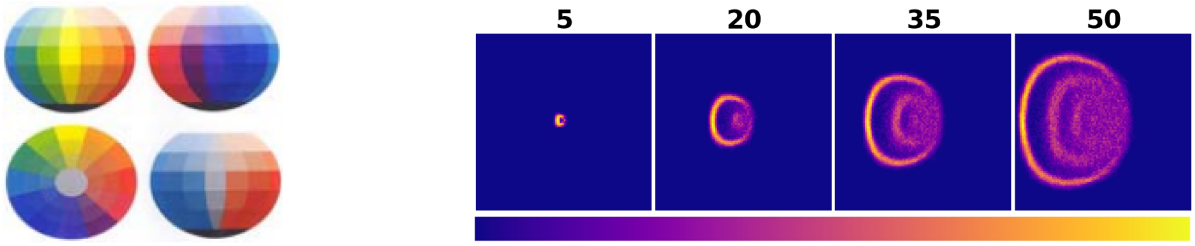
Seguindo-se esse raciocínio foram gerados um total de 21 bases de dados utilizando sempre a média de 100 simulações do autômato celular. Desde o DS10 ao DS21 os experimentos foram gerados para avaliar a sensibilidade, adaptabilidade e capacidade da proposta em configurar o modelo com recuperação probabilística.

5.3 Análise Mapa de Calor Propagação de Fogo

Durante o espalhamento de um incêndio, as chamas tendem a espalhar mais em uma determinada direção devido às características do ambiente como tipo de vegetação, quantidade de material combustível, intensidade e direção do vento. Considerando esse fenômeno representado por meio de modelos baseados em AC descritos no Capítulo 3, foi desenvolvido um mapa de calor que representa a distribuição e concentração de focos de incêndio na área de estudo durante um período de tempo específico por meio de cores com diferentes temperaturas e é um conceito que se refere ao aquecimento ou esfriamento de um corpo, representado como uma esfera (Figura 27a) ou gradiente cromático. Cores baseadas em vermelho, amarelo, roxo são cores quentes enquanto cores baseadas em verde, azul e violeta são cores frias; cores neutras são baseadas em verde-amarelo e vermelho-violeta. Estas podem ser quentes ou frias dependendo das cores próximas, propriedades de cores como intensidade e saturação influenciam no surgimento e percepção de temperatura de uma cor (YELIZAVETA; TAT-SENG; IRINA, 2005).

O mapa de calor (Figura 27b) foi gerado utilizando o gradiente abaixo dos reticulados em que as cores frias (mais à esquerda) representam áreas que sofreram pouca ou nenhuma queima e, conforme as cores se tornam quentes (mais à direita), isso indica que o pixel foi foco de incêndio com alta ou baixa intensidade. Como o intuito do mapa de calor é representar o comportamento médio geral do incêndio fornecendo uma análise mais precisa da direção e concentração dos focos, já que determinadas sementes se mostram

mais ou menos promissoras para propagação, foram feitas 100 simulações com sementes fixas de 0 a 99 com foco inicial central pois permite uma melhor observação da propagação. Portanto, o fogo tem o reticulado todo para se espalhar, pois, caso ele se inicie perto das extremidades, rapidamente alcança o limite do reticulado dificultando a observação.



(a) Esferas cromáticas. (b) Mapa de calor.

Figura 27 – Esferas cromáticas e mapa de calor.
Fonte: Figura 27a retirada de Yelizaveta, Tat-Seng e Irina (2005).

5.4 Avaliação da Abordagem Proposta

Considerando que a aleatoriedade em regras de recuperação produz comportamentos diferentes no modelo de referência descrito na Seção 3.1.2, foram feitos experimentos considerandos a regra determinística e estocástica com o intuito de verificar a influência dessa característica na evolução do AG e no cálculo de aptidão. Diferentes funções de avaliação foram investigadas com o intuito de verificar qual produz um resultado mais assertivo na configuração dos parâmetros do modelo em cenários homogêneos, mas também, avaliou-se se a inserção de outras informações além da existência de focos de incêndio, como a quantidade de matéria disponível para queima ou em recuperação, pode levar a um ajuste mais preciso.

Foi proposto um novo parâmetro para o algoritmo genético denominado *Taxa de Amostragem* que define quantos passos do AC são necessários entre cada avaliação de *fitness* e espera-se, desta forma, tornar o AG adaptável a bases distintas, além de promover uma maior eficiência computacional ao evitar processamentos desnecessários na avaliação. Para que o algoritmo seja capaz de convergir, é necessário uma quantidade de registros mínima em uma base de dados, cujo intervalo seja suficientemente espaçado de modo que represente corretamente a evolução do fenômeno em estudo. Sendo assim, experimentos de sensibilidade foram conduzidos com o intuito de avaliar qual é o limiar capaz de otimizar o tempo de execução da abordagem sem que o ajuste perca a qualidade. Por fim, objetivando testar a robustez da abordagem, esta foi aplicada a uma base de dados produzida por outro modelo com a finalidade de verificar a capacidade do algoritmo genético se adaptar aos dados.

Para cada experimento, foram feitas 30 execuções do algoritmo genético: os melhores resultados de cada execução foram agrupados e ranqueados pelo *fitness* e o melhor

resultado foi usado para configurar o modelo de simulação no cenário investigado. A fim de avaliar a qualidade da simulação produzida pela configuração encontrada pelo AG, o autômato foi executado com semente fixa 1 no ponto central 49-49, pois propicia uma melhor análise da propagação já que os focos podem se espalhar melhor pelo reticulado do que se estivessem em um dos pontos próximos das extremidades que alcançam a borda mais rápido. Por fim com intuito de avaliar o comportamento médio das simulações foi gerado um mapa de calor com o valor médio dos focos de incêndios considerando-se 100 execuções com sementes de 0 a 99. Em seguida foi feita uma análise do vetor de força encontrado pelo AG *versus* a referência, pois valores de LQ, LR e um vetor de força próximos tendem a produzir um bom comportamento.

Foi feita uma análise dos valores de genes encontrados nessas execuções para avaliar a aproximação das soluções aos valores desejados, além da quantidade de boas soluções produzidas por uma abordagem a fim de verificar a estabilidade mas, a análise gráfica terá o maior peso na escolha da melhor função. O enfoque é na reprodução do comportamento e pode acontecer que uma função não seja tão estável, porém, produza soluções mais assertivas ao ser executada algumas vezes mas também, para dados reais ou simulados por outros modelos não se conhece os valores adequados para os parâmetros essa se mostra uma boa estratégia.

Como método de avaliação foram considerados os seguintes critérios:

1. Capacidade de aproximação dos parâmetros de referência para queima (LQ) e recuperação (LR);

2. capacidade de reprodução da intensidade e inclinação do vetor de matriz de preferência do vento;

3. capacidade de reprodução da evolução temporal do autômato celular nos instantes de tempo definidos;

4. capacidade de apresentar uma evolução temporal média adequada (mapa de calor) nos instantes de tempo definidos.

5.5 Modelo com Recuperação Determinística

O objetivo desta seção é realizar uma análise do comportamento de um incêndio ao se espalhar por uma área, com enfoque principal na primeira onda de queima e em sequência na ocorrência de novas ondas de requeima. Com estas informações, foram geradas várias bases de dados contendo informações de células em ignição, ignição e em processo de recuperação e; por fim, ignição, recuperação e que não sofreram queima. Tais bases com diferentes características do fenômeno foram utilizadas para a avaliação da capacidade das funções analisadas em representar o fenômeno (Seção 5.5.1), e quais informações produzem

uma melhor aproximação do fenômeno no cálculo de *fitness*. Definida a melhor função de avaliação e formato da base de dados, na Seção 5.5.2, foram feitos testes para avaliar a influência de diferentes entradas na adaptabilidade do AG e a robustez da abordagem é avaliada na Seção 5.5.3 ao aplicar o algoritmo para reproduzir dados provenientes de outro modelo da literatura.

5.5.1 Desempenho Diferentes Funções de Aptidão

A fim de comparar o desempenho de cada função investigada no cálculo de aptidão (Fit9P, F1, F2) foram feitas execuções considerando uma a três informações, sendo apenas a contabilização células no estado de fogo denominadas como **F**, correspondendo ao estado infectado de uma função SIR. Para as duas informações além da contabilização de células no estado de Fogo, considera-se o estado queimado **Q**: tal estado não existe no modelo propriamente dito, ele refere a células do tipo árvore que acabaram de sofrer queima e tem idade 1 ou que estão em processo de recuperação e não atingiram o valor máximo de LR para estarem suscetíveis a nova queima. Para as três informações, além dos dois estados anteriores, a quantidade de células do tipo árvore com valor máximo de LR também é contabilizada, isto é, são células que estão suscetíveis **V** a sofrerem queima no próximo passo de tempo do AC. A matriz foi fixada com vento da direção oeste com intensidade 2 e ângulo de 180^{o} com valor de queima (LQ) igual a 3 e de recuperação (LR) igual a 15. Considerando o foco central de queima, foi feita uma análise numérica e gráfica das funções a fim de determinar qual obteve o melhor resultado com base nos 4 critérios da Lista 5.4. Cada uma das funções foram executadas 30 vezes para cada estado em análise sendo apenas fogo, fogo e áreas em recuperação e por fim fogo, recuperação e áreas que não sofreram queima.

Na Tabela, 1 são apresentados os melhores cromossomos encontrados para cada função analisada em relação aos parâmetros utilizados como Referência para geração da base de dados que contabilizam apenas a ocorrência de fogo em uma célula. A linha *fitness* da tabela se refere à avaliação produzida pelos genes mostrados e a linha Máximo se refere à pior avaliação encontrada para cada função. Como a forma de avaliação é diferente em cada uma, o valor máximo foi utilizado para normalizar o valor do *fitness* de modo que possam estar em uma escala única e o erro encontrado por cada função seja comparado. Também foi considerada a direção do vetor de intensidade do vento e a reprodução dos reticulados nos instantes de tempo $t = \{7, 20, 35, 50\}$, tendo como ponto inicial de queima a posição central do reticulado $(49, 49)$ e uma semente fixa em 1 de modo que o resultado produzido pelas funções possa ser comparado.

Em relação ao erro produzido pelas funções as que obtiveram o melhor erro, Fit9P e F2P9 obtiveram o menor erro. Entretanto no pior caso, F2P9 foi a que obteve o pior desempenho. Da mesma maneira, Fit9P obteve um desempenho geral melhor. Um fator importante na representação do comportamento de um incêndio é o tempo de propagação

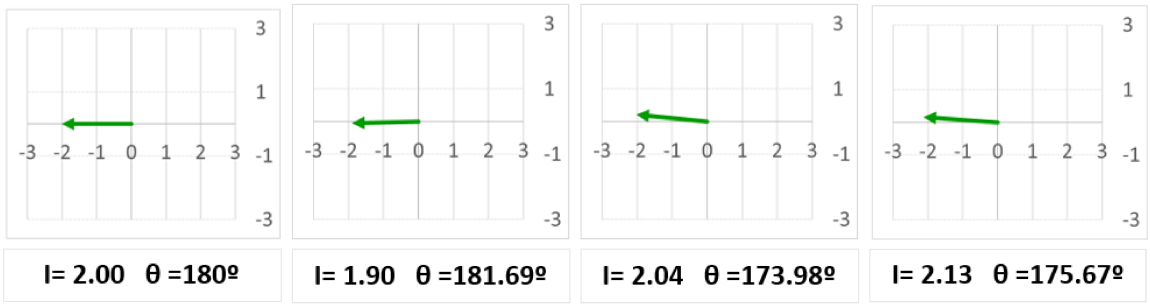
Tabela 1 – Melhores resultados: Fogo

Parâmetros	Referência	Fit9P	F1P9	F2P9
G:0	0.140	0.125	0.169	0.003
G:1	0.500	0.515	0.525	0.752
G:2	0.850	0.916	0.832	0.696
G:3	0.000	0.025	0.03	0.234
G:4	1.000	0.811	0.971	1.272
G:5	0.140	0.15	0.089	0.014
G:6	0.500	0.425	0.706	0.79
G:7	0.850	0.939	0.959	0.886
LQ	3	3	3	3
LR	15	15	15	15
Fitness	*	3.926	16.756	5.914
Máximo	*	56.368	59.396	125.291
Normalizado	*	0.070	0.282	0.047

(LQ) e recuperação (LR), encontrado por todas as funções foi encontrado, pois esses parâmetros descrevem o tempo que uma célula permanece em chamas e quanto tempo será necessário para que essa esteja suscetível a uma nova queima. A combinação desses fatores faz com que os espalhamento seja mais acelerado ou gradual, uma vez que valores mais altos de LQ produzem uma queima mais gradual e mais baixos uma queima mais acelerada; de modo similar, quanto mais baixo o valor de LR, mais rápido uma área estará propícia à queima, levando a uma propagação interna ao do foco principal.

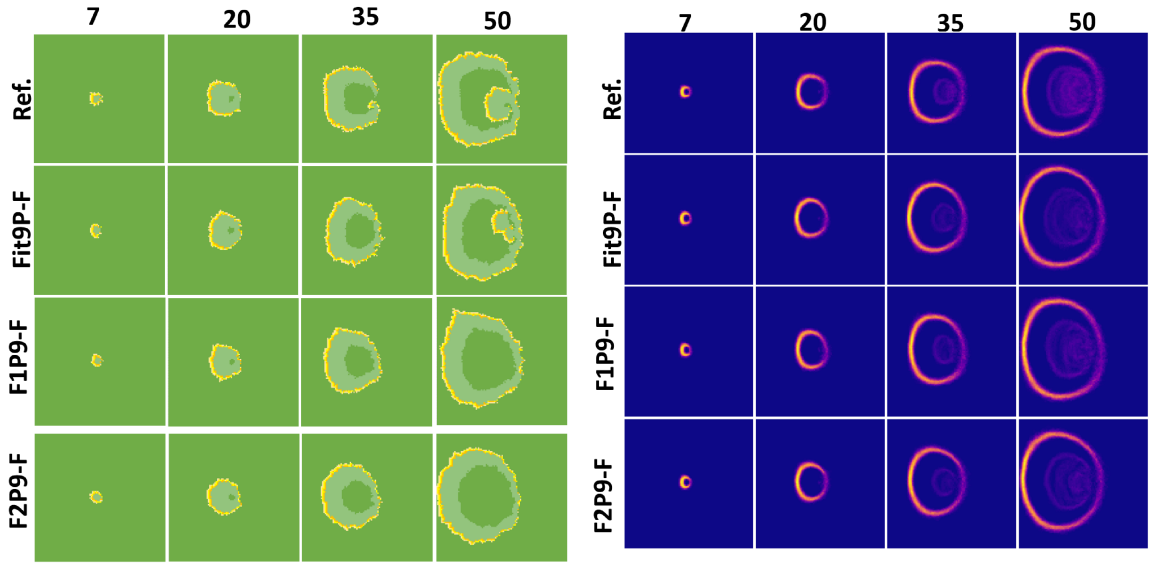
Uma característica que leva a uma boa reprodução da base de dados é a capacidade da função em capturar a combinação de valores que produzem um vetor de influência do vento adequado, pois a combinação deles deve produzir uma força resultante com intensidade e angulação adequada, uma vez que a intensidade em conjunto com LQ ditam o ritmo de espalhamento do incêndio na vegetação e a inclinação a direção da propagação. Na Figura, 28 são apresentados os vetores resultantes das matrizes de vento encontradas por cada função: o vetor de referência possui uma intensidade 2 e inclinação de 180º graus na direção Oeste (o vento sopra do Leste para o Oeste a partir de posições com valores mais altos para mais baixos). As funções foram capazes de encontrar a direção adequada do vento porém com ligeiras diferenças na inclinação e intensidade. A função Fit9P encontrou um ângulo de 181,69º e intensidade de 1,90 (uma diferença de 1,69º e 0,10 em relação à inclinação e intensidade de referência); já a função FP91, obteve uma inclinação de 173,98º e intensidade de 2,04 uma diferença de 6,02º e 0,04 para os valores base; por fim, F2P9 obteve um ângulo de 175,20º e intensidade de 2,15 sendo a diferença de 4,8º e 0,15 para a referência. Desta forma, pode-se afirmar que a função com melhor desempenho neste critério continua sendo Fit9P.

Uma vez analisado o desempenho dos parâmetros, o passo final é a análise gráfica dos resultados produzidos por cada função investigada. Deste modo, na Figura 29a, a evolução temporal do AC é apresentada nos passos de tempo definidos. No reticulado de



(a) Ref. (b) Fit9P (c) F1P9 (d) F2P9

Figura 28 – Matriz de preferência vento: Fogo.



(a) Evolução do autômato celular. (b) Mapa de calor.

Figura 29 – Evolução temporal e frequência contabilizando-se Fogo.

referência, há dois arcos, um principal que representa o maior espalhamento do incêndio e quase toca as bordas no dia 50 e um secundário mais interno que representa uma segunda propagação do incêndio nas áreas da vegetação que já se recuperaram após a passagem da primeira onda. Todas as funções foram capazes de representar essa primeira onda e a F2P9 foi a função que melhor a representou, seguida por F1P9 e Fit9P que tiveram uma primeira onda mais afastada da borda, mas que, entretanto, foi a única que produziu uma segunda onda, considerada a que mais se aproximou do reticulado de referência.

Por causa de algumas sementes mostrarem serem mais favoráveis em representar determinados comportamentos e, a fim de obter uma comparação capaz de capturar essas singularidades, foram utilizados mapas de calor pois permitem observar regiões com maior ocorrência de focos de incêndio identificando o comportamento geral dos reticulados. Com a finalidade de permitir uma comparação, o autômato foi executado 100 vezes com sementes de 0 a 99, com foco inicial central. Toda vez que uma célula estava em chamas o valor foi contabilizado e, no final o valor médio que ela sofreu queima foi obtido e, tal

Tabela 2 – Melhores resultados: Fogo e Queimado

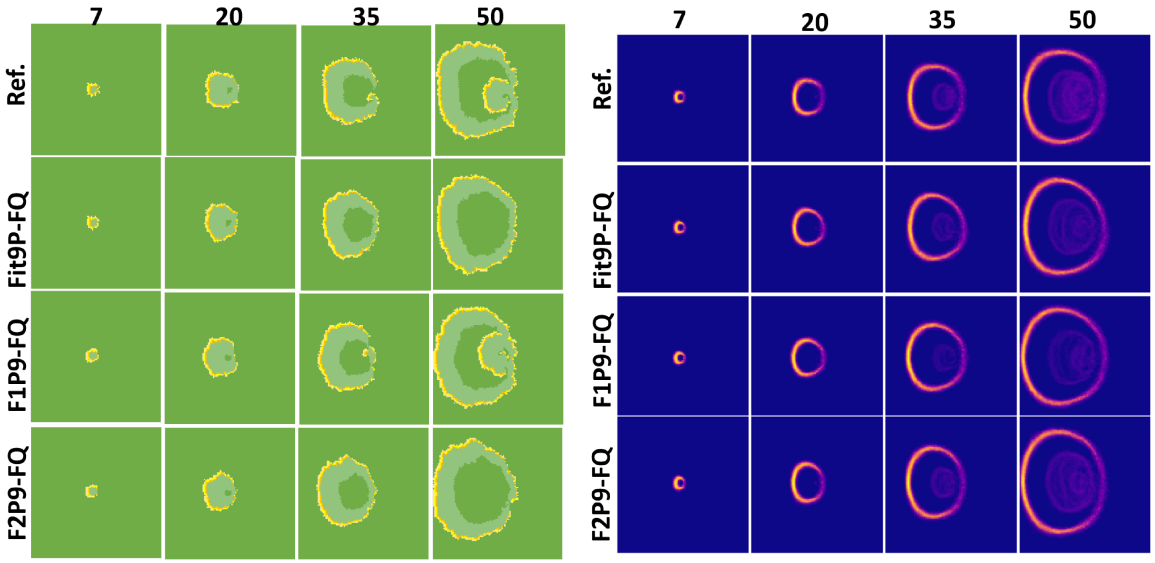
Parâmetros	Referência	Fit9P	F1P9	F2P9
G:0	0.140	0,111	0,073	0,114
G:1	0.500	0,538	0,604	0,573
G:2	0.850	0,847	0,818	0,809
G:3	0.000	0,165	0,179	0,075
G:4	1.000	0,942	1,193	1,205
G:5	0.140	0,022	0,034	0,075
G:6	0.500	0,766	0,814	0,711
G:7	0.850	0,892	0,926	0,919
LQ	3	3	3	3
LR	15	15	15	15
Fitness	*	2,572	23,574	11,231
Fit Máximo	*	95,971	29,298	19,175
Fit Normalizado	*	0,027	0,805	0,586

valor foi utilizado para produzir o mapa de calor apresentado na Figura 29b. Conforme se observa no mapa de calor na média, as funções obtiveram um comportamento adequado tanto no raio externo quanto interno, porém com pequenas inclinações devido aos vetores de força produzidos, havendo uma leve inclinação para noroeste em F1P9 e para sudoeste em F2P9. A Fit9P pela inclinação estar quase alinhada com a de referência não houve inclinação perceptível no mapa de calor.

Devido ao desempenho em todos os critérios adotados para análise ao se considerar apenas o cenário de incêndio e pela simplicidade no cálculo o que leva a um melhor desempenho computacional, a Fit9P foi escolhida como a melhor função.

Assim como na análise anterior, serão utilizados três critérios de análise (análise dos genes e *fitness* do cromossomo, evolução do autômato celular e mapa de calor) para definir qual função produziu a melhor configuração ao contabilizar a existência de focos de incêndio e áreas queimadas. Sendo assim, a Tabela 2 apresenta os melhores cromossomos encontrados pelas funções de aptidão considerando essas duas informações.A função Fit9P obteve um *fitness* mais baixo do que o cenário de contabilização de apenas fogo, entretanto para o pior resultado obtido o valor quase dobrou passando de 56,368 para 95,971 sendo a função que obteve a maior diferença entre a melhor e pior solução quando comparada a F1P9 que obteve uma diferença de 5,724 e F2P9 com 7,944. Dessa forma, F2P9 e F1P9 se mostraram mais estáveis que Fit9P ao se acrescentar uma nova informação ao cálculo, especialmente F1P9, que apresentou a menor diferença sendo a melhor nesta análise, mesmo não apresentando o melhor valor na normalização pela diferença entre os valores ser muito próxima.

Comparando as matrizes de preferência do vento do experimento anterior, ao se incluir mais uma informação, percebe-se que houve um ligeiro aumento na inclinação para a direção Noroeste, conforme apresentado na Figura 31. Essa pequena diferença na inclinação não foi significativa, pois, as funções conseguiram uma boa aproximação em que,



(a) Evolução do autômato aelular: FQ. (b) Mapa de calor: FQ.

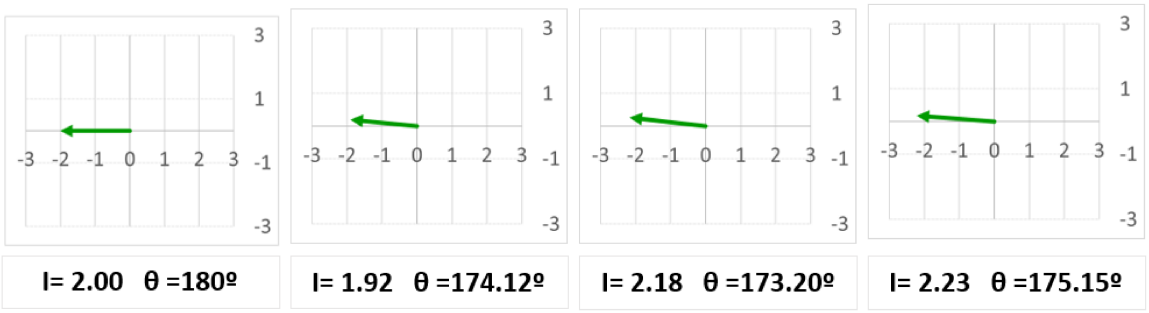
Figura 30 – Evolução temporal e frequência FQ.

Fit9P encontrou um vetor de 174,12º e intensidade de 1,92 havendo uma diferença de 5,88º e 0,08 em relação a referência; F1P9 encontrou um valor de 173,20º e 2,18 com uma diferença de 6,80º e 0,18, F2P9 encontrou um vetor de 175,15º e 2,23 uma diferença de 4,85 e 0,23. A função que teve o ângulo mais próximo do de referência foi F2P9 entretanto a que obteve a intensidade mais próxima foi Fit9P. Nesse quesito, não houve uma função que apresentou uma solução muito superior.

Em relação a evolução temporal, há uma diferença na evolução devido a intensidades diferentes pois as funções que obtiveram maior intensidade de vento houve uma propagação mais alongada e rápida no foco principal, como em F2P9, que possui a maior intensidade seguida por F1P9 e por Fit9P que obteve a intensidade mais próxima. Quanto à segunda onda de propagação apenas F1P9 obteve esse resultado, portanto, na análise gráfica dos reticulados foi considerada a com melhor desempenho.

Na média geral apresentada pelo mapa de calor, o desempenho foi muito similar havendo diferenças bem pequenas entre elas e todas foram capazes de produzir as duas ondas de propagação porém, com intensidades diferentes. Do mesmo modo, observado na análise dos reticulados, a F2P9 obteve ondas externa e interna mais acentuadas com maiores raios; Fit9P e F1P9 apresentaram formas bem similares, porém, por F1P9 possuir uma maior intensidade, o foco principal foi ligeiramente mais acentuado mas o foco interno em Fit9P estava mais preenchido. Pela proximidade dos reticulados no mapa de calor, não houve uma função com um comportamento muito superior nesse quesito, mas, por ter sido melhor em dois critérios de avaliação (análise numérica e dos reticulados), para a dimensão analisada, F1P9 foi selecionada como a de melhor desempenho.

A Tabela 3 mostra o resultado obtido pelas funções ao se considerar três informações no cálculo de *fitness* em que, células em chamas, que sofreram queima e que estão suscetíveis à



(a) Ref. (b) Fit9P (c) F1P9 (d) F2P9

Figura 31 – Matriz de preferência vento: Fogo, Queimado.

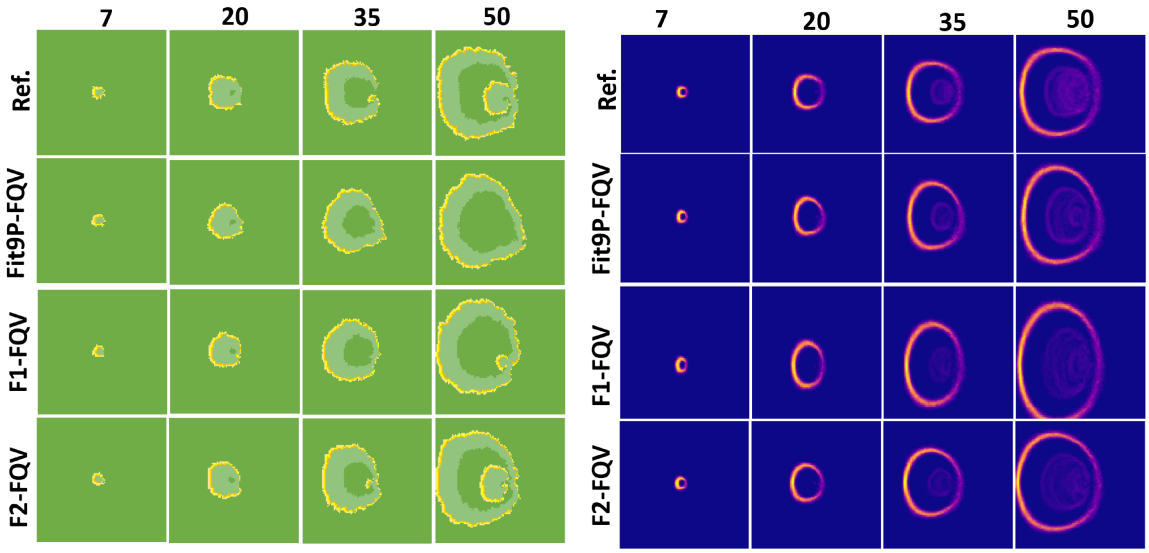
Tabela 3 – Melhores resultados: Fogo, Queimado e Vegetação

Parâmetros	Referência	Fit9P	F1P9	F2P9
G:0	0.140	0,161	0,087	0,109
G:1	0.500	0,381	0,625	0,57
G:2	0.850	0,904	0,758	0,827
G:3	0.000	0,119	0,144	0,136
G:4	1.000	0,787	1,394	1,365
G:5	0.140	0,015	0,046	0,03
G:6	0.500	0,943	0,868	0,891
G:7	0.850	0,855	0,741	0,871
LQ	3	3	3	3
LR	15	15	15	15
Fitness	*	4,859	28,42	16,54
Máximo	*	19,648	34,063	25,484
Normalizado	*	0,247	0,834	0,649

ignição. Novamente, a função Fit9P foi a que obteve o menor erro normalizado, entretanto, ela também foi a com maior diferença entre a melhor e pior solução sendo de 14,789; em F1P9 o erro normalizado foi o segundo maior mas a diferença entre os extremos foi de 5,643; por fim, F2P9 obteve o segundo menor *fitness* normalizado e erro entre a melhor e pior solução (8,944) sendo mais estável que Fit9P ao se utilizar 3 dimensões e, por esse motivo, foi escolhida como melhor função neste quesito.

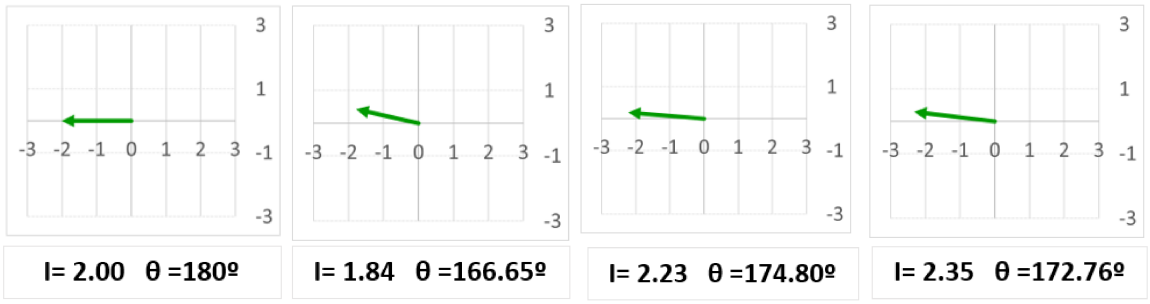
Ao se considerar as matrizes de vento produzidas na Figura 33, aumentando uma nova dimensão, a inclinação produzida foi maior assim como a intensidade, produzindo reticulados mais distantes da representação de referência. A função Fit9P foi a que apresentou a pior inclinação, encontrando um valor de 166,65º e intensidade de 1,84, uma diferença de 13,35º e 0,16 em relação à angulação e força do *baseline*. F1P9 foi a que mais se aproximou da inclinação ideal tendo encontrado um vetor de 174,80º e intensidade de 2,23 sendo a diferença de 5,2° e 0,23. Já F2P9, apresentou a maior intensidade com um valor de 2,35 e angulação de 172,76º, um erro de 0,35 e 7,24º. Nessa análise, a função que mais se aproximou da matriz ideal foi F1P9.

Na Figura 32a, pode-se notar que Fit9P, por ter a inclinação mais acentuada produ-



(a) Evolução temporal FQV. (b) Mapa de calor FQV.

Figura 32 – Evolução temporal e frequência FQV.



(a) Ref. (b) Fit9P (c) F1P9 (d) F2P9

Figura 33 – Matriz de preferência vento: Fogo, Queimado e Vegetação

ziu um reticulado com essa característica tendo um foco de incêndio muito acentuado a Noroeste; já F1P9 apresentou um comportamento na propagação principal adequado mas a segunda onda foi muito leve e quase imperceptível; por fim, F2P9 foi a única função com um comportamento adequado para a propagação da primeira e segunda onda. Desta forma,foi a melhor função neste critério.

Ao se analisar o mapa de calor da Figura 32b, houve uma piora significativa no comportamento médio das funções, aliado ao fato da inclinação e intensidade dos vetores serem mais acentuados. Isso pode ser verificado em F1P9 que produziu o pior comportamento médio tendo a onda principal ultrapassando a borda sul do reticulado e tanto esta como a onda interna com uma forma bastante ovalada em detrimento do comportamento radial esperado. F2P9 e Fit9P foram as funções com melhor resultado. Entretanto, a inclinação de Fit9P é mais acentuada e, por esse motivo, F2P9 foi escolhida como melhor nesse cenário e, por ter sido escolhida no critério anterior, foi selecionada como a melhor ao se trabalhar com três dimensões.

Os melhores resultados de cada análise considerando-se uma (fogo), duas (fogo e quei-

Tabela 4 – Melhores funções por cenário avaliado

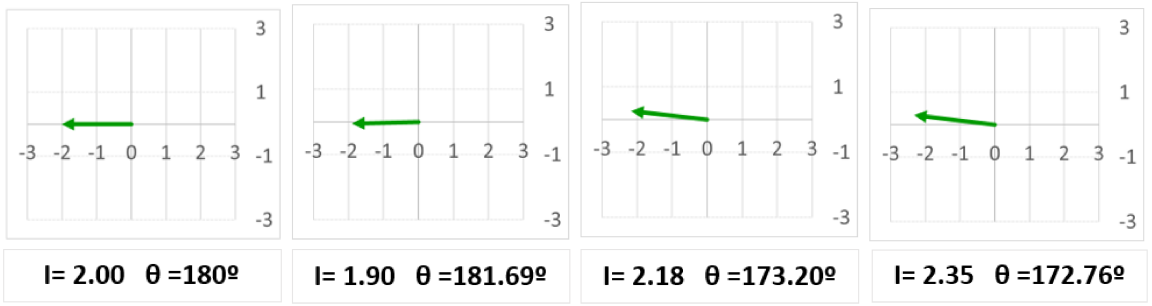
Parâmetros	Referência	Fit9P-F	F1P9-FQ	F2P9-FQV
G:0	0.140	0,125	0,073	0,109
G:1	0.500	0,515	0,604	0,57
G:2	0.850	0,916	0,818	0,827
G:3	0.000	0,025	0,179	0,136
G:4	1.000	0,811	1,193	1,365
G:5	0.140	0,15	0,034	0,03
G:6	0.500	0,425	0,814	0,891
G:7	0.850	0,939	0,926	0,871
LQ	3	3	3	3
LR	15	15	15	15
Fitness	*	3,926	23,574	16,54

mado) e três (fogo, queimado, vegetação) informações no cálculo de *fitness*, foram reunidos em uma tabela e nos *plots* de evolução temporal e frequência com o intuito de permitir uma melhor visualização e comparação do desempenho entre as funções com o intuito de se avaliar a melhor.

A Tabela 4 mostra as melhores soluções encontradas para cada função considerando-se diferentes informações no cálculo de aptidão. Ao se considerar apenas a quantidade de células no estado de fogo (uma informação) Fit9P foi a melhor. Incluindo-se mais informação no cálculo de aptidão, funções com inspiração SIR obtiveram um melhor resultado. Incluindo a informação de queimado no cálculo, isto é, duas informações (fogo, queimado), F1P9 foi a melhor função. Ao se considerar três informações (fogo, queimado, vegetação) F2P9 apresentou o melhor desempenho.

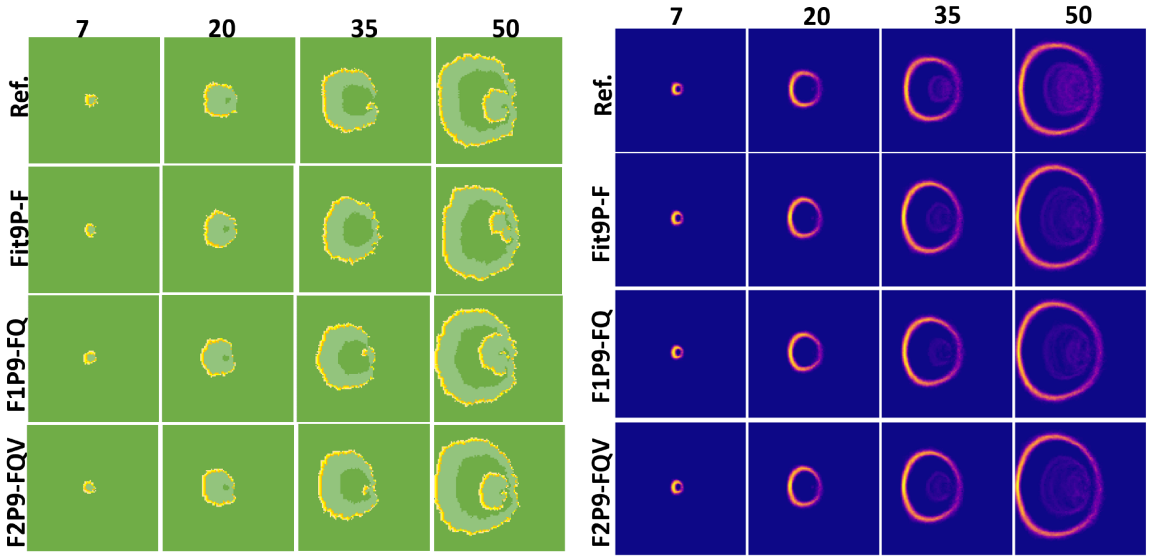
Ao compararmos as matrizes de vento da Figura 34 podemos notar que, conforme novas dimensões são adicionadas, menos precisa é a angulação e intensidade e, deste modo, a função Fit9P-F foi a que apresentou o vetor com melhor inclinação e intensidade. Do ponto de vista de representação da evolução temporal do AC apresentado na Figura 35a, a função que apresentou a melhor representação em relação à referência foi a função F2P9-FQV, logo após F1P9-FQ e por fim, Fit9P-F. Porém, ao se estender a análise para o mapa de calor, percebemos que o comportamento médio foi muito parecido em todas as funções, não havendo ganhos significativos ao se utilizar múltiplas dimensões dada a influência da matriz de preferência de vento e sua intensidade os mapas tem ligeiras inclinações de acordo com o vetor encontrado e, desta forma, o mapa de calor que está mais alinhado com o de referência é o da função Fit9P-F.

Diante dos resultados encontrados, não há ganhos em se utilizar mais informações no cálculo do *fitness* e essa complexidade adicional leva a uma matriz de ventos mais imprecisa, além de ser necessário mais poder computacional nos cálculos, optou-se por manter Fit9P com apenas a informação da existência de focos de incêndio como a função de referência.



(a) Ref. (b) Fit9P-F (c) F1P9-FQ (d) F2P9-FQV

Figura 34 – Melhores matrizes de preferência.



(a) Melhores evoluções autômato celular. (b) Melhores evoluções mapa de calor.

Figura 35 – Evolução temporal autômato celular e mapa de calor.

5.5.2 Análise Sensibilidade

Após a definição da melhor função de aptidão e das informações a serem consideradas no cálculo de *fitness*, com o objetivo de avaliar a capacidade do algoritmo em calibrar os parâmetros e sua adaptabilidade a uma determinada entrada, foram realizados testes de sensibilidade variando o tamanho do conjunto de dados. O objetivo foi determinar o tamanho mínimo necessário para a convergência do algoritmo sem perder a qualidade do ajuste. Essa abordagem torna o algoritmo mais adaptável às fontes de informação e otimiza o tempo computacional ao ser capaz de calibrar o modelo com menos registros.

Inicialmente, o autômato celular descrito na Seção 3.1 foi executado 100 vezes utilizando os parâmetros da matriz W correspondente ao vento para o Oeste, $LQ = 3$ e $LR = 15$ ao longo de 50 etapas de tempo. O valor médio para cada etapa de tempo foi calculado considerando as 100 simulações de AC e bancos de dados de referência com tamanhos diferentes foram criados. O primeiro banco de dados armazena os valores médios relacionados a todas as 50 etapas de tempo. O segundo banco de dados é composto por

apenas 25 valores médios e que, foram armazenadas apenas as etapas de tempo pares {2, 4, 6, ..., 50}. Da mesma forma, outros quatro bancos de dados foram construídos para armazenar 16, 10, 7 e 5 valores médios. Assim, seis conjuntos de dados de referência (DS) foram construídos usando diferentes períodos de amostragem $p = \{1, 2, 3, 5, 7, 10\}$, resultando em diferentes tamanhos de conjunto de dados $N = \{50, 25, 16, 10, 7, 5\}$ respectivamente. Esses conjuntos de dados de referência foram chamados de *Dataset i* ou DSi, em que i indica o i-ésimo elemento de p e S. Embora esses conjuntos de dados caracterizem o mesmo cenário de incêndio florestal, já que foram construídos com base no mesmo grupo de 100 simulações do modelo de AC, eles representam os dados de um mesmo fenômeno com diferentes frequências de amostragem. Generalizando para um incêndio florestal no mundo real, eles armazenam uma sequência de imagens da propagação com diferentes frequências de captura.

A matriz de vento (W) na Figura 20a produz um vetor resultante mostrado na Figura 20b, que representa a intensidade do vento ($I = 2$) e a direção Oeste (θ =180º). Esse vetor é obtido por meio da decomposição vetorial em que, a célula central W_{2x2} representa a origem do plano cartesiano e cada célula da matriz W_{ij} é decomposta em componentes verticais e horizontais que são utilizados para gerar o vetor resultante da matriz conforme demonstrado na Seção 3.1, mesmo que não sejam encontrados os valores exatos dos parâmetros da matriz; uma boa reprodução do comportamento do vento pode ser alcançada a partir de uma configuração capaz de produzir um vetor resultante semelhante, ou seja, com intensidade e direção do vento próximas.

O AG foi aplicado para otimizar os parâmetros do modelo de AC para os conjuntos de dados usados como referência (DS1, DS2, DS3, DS4, DS5 e DS6). Foram realizadas 30 execuções do AG para cada conjunto de dados. Na Tabela 5, os parâmetros de referência usados nas simulações do modelo baseado em AC para a geração do conjunto de dados original com 50 grades são apresentados na coluna denominada "Referência". As melhores soluções encontradas pelo AG para cada conjunto de dados são apresentadas nas outras colunas. Os genes G_{Nw} a G_W que compõem a matriz W para cada experimento são mostrados. Esses valores são utilizados para produzir o vetor resultante com sua intensidade e inclinação, os valores para os parâmetros de tempo de queima (LQ), tempo de recuperação (LR) e taxa de amostragem (SR) também são apresentados, assim como a aptidão alcançada pelo melhor indivíduo do AG em cada conjunto de dados.

Como pode ser observado, o AG alcançou os valores exatos dos parâmetros LR, LQ e SR para todos os conjuntos de dados. É importante notar que os valores de aptidão alcançados pelo melhor indivíduo em cada experimento não são comparáveis entre os experimentos DS1 a DS6, já que o cálculo dos valores depende do número de registros pertencentes a cada banco de dados.

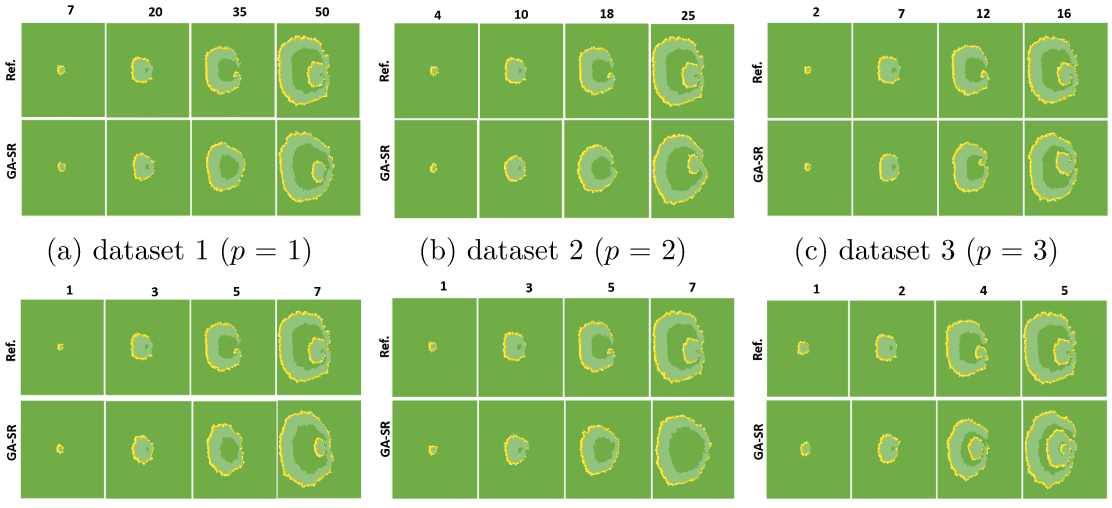
Considerando o experimento DS1, o maior conjunto de dados com 50 registros foi usado, então o AG acessa todas as informações de propagação (50 grades), tornando

Tabela 5 – Referência e melhores parâmetros encontrados pelo AG para cada *dataset*

Parâmetros	Referência	DS1	DS2	DS3	DS4	DS5	DS6
G_{Nw}	0.140	0.102	0.036	0.015	0.117	0.113	0.030
G_N	0.500	0.441	0.626	0.858	0.583	0.497	0.626
G_{Ne}	0.850	0.929	0.816	0.746	0.859	0.954	0.839
G_E	0.000	0.125	0.247	0.102	0.101	0.044	0.245
G_{Se}	1.000	0.750	1.235	1.062	0.897	0.834	1.039
G_S	0.140	0.117	0.007	0.139	0.096	0.169	0.014
G_{Sw}	0.500	0.371	0.715	0.430	0.464	0.384	0.677
G_W	0.850	1.057	0.679	0.905	0.952	0.871	0.814
LQ	3	3	3	3	3	3	3
LR	15	15	15	15	15	15	15
SR	-	1	2	3	5	7	10
Fitness	-	4.315	3.01	1.717	0.959	0.840	0.353

o caso mais fácil de ajuste. O valor de SR igual a 1 foi encontrado, e a matriz W produziu valores próximos aos originais, conforme mostrado na Figura 38a, com um vetor de intensidade de 1,87 e uma inclinação de 179,04º, que é próxima ao vetor resultante relacionado à matriz original W. As amostras da evolução temporal escolhidas para o plotagem gráfica foram $t = \{7, 20, 35, 50\}$. A Figura 36a apresenta a evolução temporal de (i) o DS1 de referência (gráfico superior); e (ii) a simulação baseada nos parâmetros do melhor indivíduo evoluído pelo AG (gráfico inferior). Para todas as outras evoluções temporais apresentadas na Figura 36, o gráfico superior foi feito com base no conjunto de dados usado como referência, e o gráfico inferior foi gerado a partir da simulação utilizando o melhor indivíduo do AG evoluído para esse conjunto de dados. O comportamento observado na Figura 36a foi apropriado, com áreas queimadas muito semelhantes para ambas as ondas principal e secundária. Além disso, essa observação pode ser generalizada para várias simulações, conforme mostrado no mapa de calor na Figura 37a, com uma leve tendência a evoluir para Noroeste devido ao grau de diferença no vetor, mas nada muito significativo.

No experimento DS2, a metade dos registros (25) foram utilizados, portanto o AG tem menos informações sobre a propagação do fogo - amostrando o fenômeno a cada dois intervalos de tempo - o que aumenta a dificuldade do AG em encontrar um bom ajuste de parâmetros. Apesar disso, ele foi capaz de determinar que o SR mais adequado é 2. O AG também obteve bons resultados em relação à matriz W (Figura 38b), com um vetor de intensidade de 2,01 e uma inclinação de 180,8º e foi o melhor até agora comparado à referência. As amostras escolhidas para a plotagem gráfica foram $t = \{4, 10, 18, 25\}$. A simulação utilizando o melhor indivíduo apresentou uma evolução temporal adequada na Figura 36b, quando comparada ao comportamento dos dados, especialmente na onda principal, enquanto a segunda onda existe, mas é ligeiramente menor que a referência. Considerando o mapa de calor na Figura 37b, o comportamento médio foi bastante similar



(a) dataset 1 ($p = 1$) (b) dataset 2 ($p = 2$) (c) dataset 3 ($p = 3$)

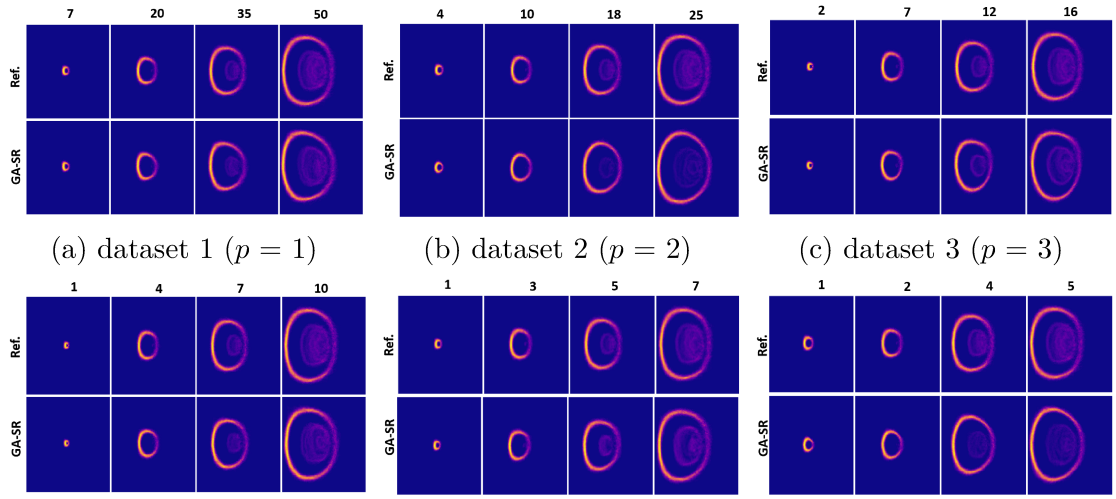
(d) dataset 4 ($p = 5$) (e) dataset 5 ($p = 7$) (f) dataset 6 ($p = 10$)

Figura 36 – Evolução temporal autômato celular.

para ambas as ondas externas e internas, embora nela o traço seja um pouco mais fraco que a referência, mas o tamanho é muito próximo.

Considerando os experimentos restantes (DS3, DS4, DS5 e DS6), o melhor indivíduo encontrado para cada conjunto de dados de referência determina o SR mais adequado: 3, 5, 7 e 10, respectivamente. Em relação à matriz W, os valores são apropriados, produzindo vetores resultantes semelhantes à referência, conforme mostrado na Figura 30, embora os experimentos DS3, DS4 e DS5 tenham mostrado uma inclinação mais pronunciada para o Sudoeste (6,44º, 7,27º e 4,01º, respectivamente). Essa diferença levou a uma propagação suave do fogo nessa direção (Figura 36) para esses experimentos. No geral, apesar da inclinação, o comportamento observado nos experimentos DS3 e DS4 foi apropriado. Considerando o experimento DS5, cujo conjunto de dados de referência tem 7 registros, a evolução temporal na Figura 36e mostra que a simulação com o melhor indivíduo evoluído foi capaz de reproduzir a onda de propagação principal, mas a segunda onda não existe. O comportamento médio observado na Figura 37 para os experimentos DS3, DS4 e DS5 foi adequado. DS6 foi o experimento com a menor base contendo apenas 5 registros. Embora uma boa configuração tenha sido encontrada para a matriz W, na evolução temporal (Figura 36f), pode-se notar que as ondas internas foram bastante aparentes e ainda mais intensas que a referência. No comportamento médio mostrado pelo mapa de calor na Figura 37f, são observados valores com comportamento semelhante à onda externa, mas a onda interna é mais suave do que nos experimentos anteriores, mostrando que a evolução temporal não refletiu o comportamento médio adequadamente.

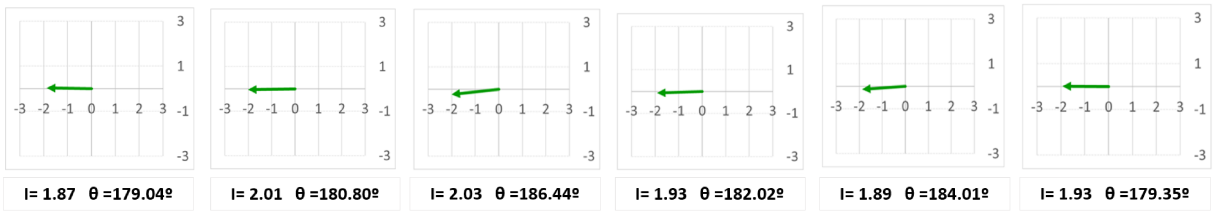
Em geral, os experimentos mostraram que o AG foi capaz de encontrar boas soluções e o SR mais adequado, com exceção de bases com os menores tamanhos de registros (DS5 e DS6) que não produziram resultados tão bons quanto os outros, com ondas internas inexistentes ou muito intensas.



(a) dataset 1 ($p = 1$) (b) dataset 2 ($p = 2$) (c) dataset 3 ($p = 3$)

(d) dataset 4 ($p = 5$) (e) dataset 5 ($p = 7$) (f) dataset 6 ($p = 10$)

Figura 37 – Evolução temporal mapa de calor.



(a) DS1 (b) DS2 (c) DS3 (d) DS4 (e) DS5 (f) DS6

Figura 38 – Matrizes preferência vento do experimento sensibilidade.

5.5.3 Análise de Robustez

Uma vez verificada a sensibilidade da abordagem a diferentes entradas, para validar a robustez e adaptabilidade, devido à falta de imagens de incêndios florestais do mundo real, empregamos um segundo modelo de propagação de fogo baseado em AC proposto por Tinoco et al. (2022) para gerar conjuntos de dados artificiais. Este modelo possui outros parâmetros além do modelo discutido na Seção 3.2, como umidade e vários estágios de queima, bem como os passos de tempo do AC representam um tempo de evolução do fogo diferente do modelo de propagação baseado em AC discutido anteriormente. Devido às características distintas entre os modelos, os melhores valores para os parâmetros do modelo não são conhecidos, o que se aproxima do problema de reproduzir as dinâmicas de incêndios florestais do mundo real com um modelo de simulação. O desempenho da nossa abordagem será avaliado com base na comparação entre as dinâmicas do fogo resultantes do modelo configurado pelo AG e os registros dos conjuntos de dados.

Três *datasets* foram gerados com 20 registros (em cada um) executando o modelo de referência baseado em AC com diferentes taxas de amostragem $SR \in \{5, 7, 10\}$, sendo nomeados como *Dataset* 7 (DS7, SR=5), *Dataset* 8 (DS8, SR=7) e *Dataset* 9 (DS9, SR=9). Essas taxas de amostragem foram escolhidas porque são mais desafiadoras para a busca do algoritmo genético, como observado nos experimentos na Seção anterior.

Tabela 6 – Teste de robustez

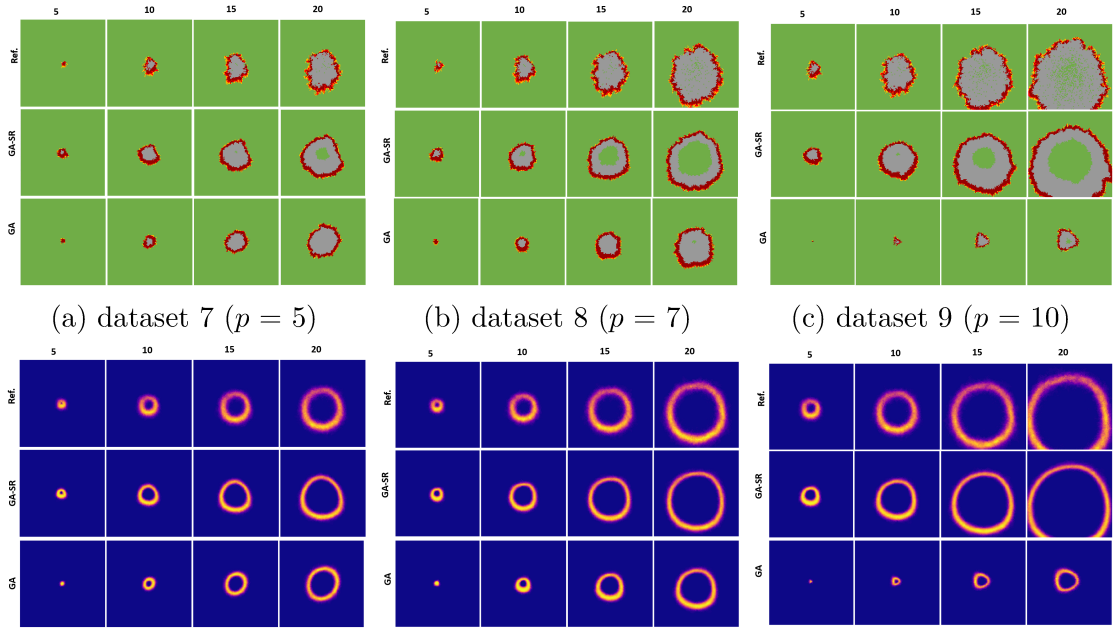
Params.	GA-DS7	GA-SR-DS7	GA-DS8	GA-SR-DS8	GA-DS9	GA-SR-DS9
G_{Nw}	0.265	0.744	0.329	0.790	0.064	1.077
G_N	0.786	0.907	0.990	1.330	0.060	1.450
G_{Ne}	0.367	0.706	0.084	0.628	0.330	0.931
G_E	0.903	0.645	0.149	0.297	0.333	0.523
G_{Se}	0.540	0.887	0.811	0.762	0.153	0.976
G_S	0.465	0.093	0.271	0.364	0.057	0.357
G_{Sw}	0.661	0.636	0.438	0.184	0.289	0.290
G_W	0.216	0.148	0.202	0.216	0.401	0.205
LQ	10	10	10	10	9	10
LR	91	33	57	23	28	37
SR	-	3	-	4	-	5
Fitness	70.474	4,223	103.086	5,848	134.999	6.997

Para validar a maior precisão da abordagem proposta, os conjuntos de dados artificiais foram submetidos ao ajuste dos parâmetros do modelo baseado em AC usando duas versões diferentes do algoritmo genético. A primeira versão não utiliza o parâmetro SR na codificação do indivíduo, enquanto a segunda versão do AG usa o parâmetro proposto no código genético, conforme apresentado no Capítulo 4.

Os resultados de cada abordagem do AG são apresentados e analisados para destacar o efeito da inclusão do parâmetro SR. Para cada versão do AG e cada conjunto de dados, foram realizadas 30 execuções, e elas foram analisadas usando a mesma abordagem utilizada nos experimentos anteriores DS1 a DS6. Os novos experimentos com o *Dataset* 7 foram nomeados GA-DS7 e, o AG sem o parâmetro SR foi utilizado, GA-SR-DS7 em que, o parâmetro proposto SR é ajustado junto com os outros parâmetros no indivíduo. O mesmo foi feito para o *Dataset* 8 e o *Dataset* 9, resultando nos experimentos nomeados GA-DS8, GA-DS9, GA-SR-DS8, GA-SR-DS9. Para a plotagem, são apresentados os mesmos passos de tempo ($t = \{5, 10, 15, 20\}$).

Considerando o experimento **GA-DS7**, a Tabela 6 mostra que o melhor indivíduo evoluído pelo AG mais simples tem $LQ = 10$ (valor máximo possível) e $LR = 91$, produzindo uma evolução temporal com uma área queimada menor, como mostrado na Figura 39a. Por outro lado, no experimento **GA-SR-DS7**, o melhor indivíduo evoluído tem $LQ = 10$ e $LR = 33$, resultando em uma área queimada maior, que é mais próxima da referência, especialmente em $t = 20$. É importante notar que existe um processo de recuperação no modelo (TINOCO et al., 2022) que não foi capturado pelo AG, mas o GA-SR foi capaz de encontrar uma recuperação semelhante usando $SR = 3$. Considerando o comportamento médio mostrado na Figura 39d, destaca-se que o GA-SR-DS7 apresentou uma dinâmica mais próxima da referência.

Em relação ao experimento GA-DS8, a Tabela 6 mostra que o melhor indivíduo evoluído pelo AG mais simples tem $LQ = 10$ e $LR = 57$. Devido ao fato de que o período



(a) dataset 7 ($p = 5$) (b) dataset 8 ($p = 7$) (c) dataset 9 ($p = 10$)

(d) dataset 7 ($p = 5$) (e) dataset 8 ($p = 7$) (f) dataset 9 ($p = 10$)

Figura 39 – GA e GA-SR evolução temporal e mapa de calor.

de amostragem é mais longo, a evolução temporal do GA-DS8 na Figura 39b não foi adequada, pois a área queimada é significativamente menor do que a referência. Por outro lado, GA-SR-DS8 foi capaz de encontrar uma melhor adaptação, aumentando o valor de SR para 4, com $LQ = 10$ e $LR = 23$. Assim, a evolução temporal foi semelhante à referência, tanto nas áreas queimadas quanto nas áreas de recuperação. No comportamento médio mostrado na Figura 39e, essa diferença entre GA-DS8 e GA-SR-DS8 é destacada. Acreditamos que a diferença é muito significativa, pois com uma taxa de amostragem mais alta, o GA conseguiu se aproximar mais da solução.

Finalmente, no experimento **GA-DS9**, a Tabela 6 mostra que o melhor indivíduo evoluído pelo AG mais simples tem $LQ = 9$ e $LR = 28$, produzindo uma evolução temporal muito diferente (Figura 39a) quando comparada à referência. Por outro lado, GA-SR-DS9 definiu $LR = 10$, $LR = 37$ e $SR = 5$ no melhor indivíduo, reproduzindo melhor as dinâmicas adequadamente, com uma área queimada (vermelha), em recuperação (cinza) e recuperada (verde) muito semelhantes. O comportamento médio na Figura 39f deixa claro a melhor adaptabilidade do GA-SR em relação a uma abordagem de GA sem a taxa de amostragem.

5.6 Modelo com Recuperação Probabilística

Os experimentos anteriores objetivam verificar como um incêndio primário se espalha ao longo da vegetação, não sendo considerada a probabilidade de requeima a menos que a vegetação esteja totalmente recuperada, o que não ocorre em incêndios reais já que há presença de material combustível e não tendo, necessariamente, vegetação recuperada,

pois este requer um tempo considerável, que pode ser desde resquícios de uma vegetação que não queimou totalmente, dejetos humanos como plástico, pneus ou outro material. Diante disso, a regra em que uma célula só pode estar suscetível à queima com LR máximo foi retirada proporcionando um comportamento mais aleatório mais próximo do fenômeno real. Ao ser inserido mais estocasticidade no modelo, a calibragem torna-se mais desafiadora para o Algoritmo genético e, com o intuito de verificar a robustez das funções no ajuste, foi escolhido o período $p = 10$, com frequência de gravação dos registros do *dataset* mais espaçado. Serão feitas análises de sensibilidade, configuração do modelo de referência, validação e adaptabilidade.

5.6.1 Análise de Sensibilidade

Nesta seção, o intuito é fazer uma análise da capacidade da abordagem em produzir uma boa aproximação do comportamento previsto no *dataset* em um modelo com reprodução probabilística sem que se perca a qualidade do ajuste. De modo similar a Seção 5.5.2, será avaliada a capacidade do algoritmo em se adaptar a diferentes entradas sem que perca a qualidade da solução. Sendo assim, foram geradas bases com diferentes períodos de amostragem $p = \{1, 2, 5, 7, 10\}$, resultando em diferentes tamanhos de conjunto de dados $N = \{50, 25, 10, 7, 5\}$. Esses períodos foram escolhidos, pois os dois primeiros são os que possuem maior quantidade de informação para o AG sendo menos desafiadores e os três últimos por serem mais complexos. A matriz de preferência do vento W considerada será a mesma dos experimentos anteriores com direção Oeste, intensidade $I = 2.00$ e inclinação $\theta = 180^{o}$. O parâmetro de queima se mantém $LQ = 3$, porém o de recuperação foi ajustado para o novo comportamento do modelo sendo definido como $LR = 30$.

A numeração de *dataset* continuará sempre que um conjunto de dados novo for gerado e, na Tabela 7, são apresentados os resultados encontrados pelo AG para os *datasets* DS10 e DS11. Foram feitos experimentos com duas abordagens com (GA-SR-DSi) e sem (GA-DSi) o parâmetro Taxa de amostragem (SR). Nos experimentos realizados com DS10 cujos registros são salvos a cada passo de tempo do AC, tanto a abordagem tradicional que não utiliza SR (GA-DS10) quanto nossa abordagem (GA-SR-DS10) encontraram bons resultados. Porém, ao passar para um SR com período igual a 2 a abordagem clássica já não foi capaz de encontrar valores adequados para os para o tempo de queima e recuperação, tendo encontrado respectivamente $LQ = 2$ e $LR = 48$.

Uma vez verificado que períodos de amostragens diferentes de unitários não produzem bons resultados na abordagem clássica em cenários mais complexos, os experimentos seguintes com taxas mais esparsas (Tabela 8) sempre se considerou o parâmetro SR. Em todos os experimentos a abordagem foi capaz de se aproximar de um valor aceitável para os parâmetros com erros próximos a mais ou menos 1 para LR, sendo encontrado o valor exato de LQ e LR para GA-SR-DS13 e com um erro ligeiramente maior para o período de recuperação em GA-SR-DS14.

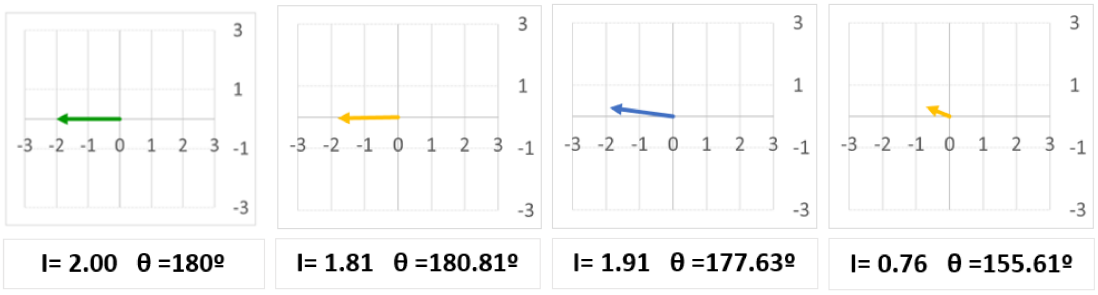
Tabela 7 – Comparação abordagem tradicional e proposta

Parâmetros	GA-DS10	GA-SR-DS10	GA-SR-DS11	GA-DS11
Gnw	0.122	0.060	0.145	0.775
Gn	0.410	0.627	0.516	0.420
Gne	0.894	0.782	0.935	0.306
Ge	0.081	0.086	0.014	0.216
Gse	0.946	0.863	0.738	0.792
Gs	0.060	0.149	0.123	0.228
Gsw	0.836	0.462	0.429	0.723
Gw	0.748	0.890	1.009	0.874
LQ	3	3	3	2
LR	31	29	31	48
SR	1	-	2	-
Fitness	6.072	6.607	3.007	87.506

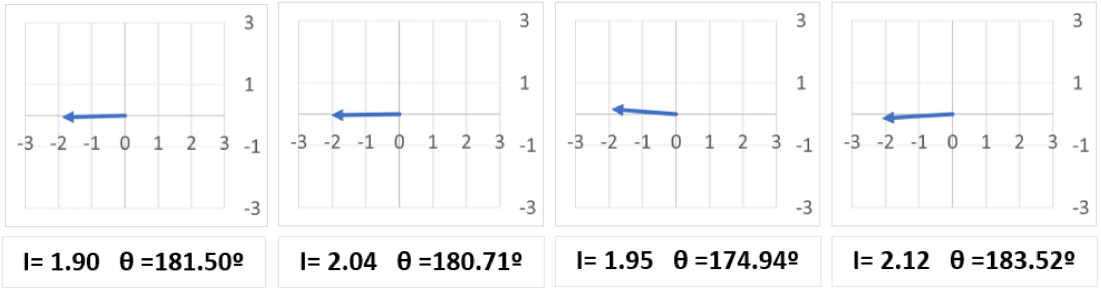
Tabela 8 – Teste de sensibilidade

Parâmetros	GA-SR-DS12	GA-SR-DS13	GA-SR-DS14
Gnw	0.139	0.172	0.027
Gn	0.547	0.348	0.798
Gne	0.824	0.895	0.828
Ge	0.028	0.013	0.177
Gse	1.057	0.919	1.144
Gs	0.113	0.127	0.067
Gsw	0.513	0.572	0.595
Gw	0.862	0.866	0.891
LQ	3	3	3
LR	31	30	34
SR	5	7	10
Fitness	1.451	0.905	0.383

Com relação à matriz de preferência do vento, para o experimento com *dataset* DS10 (Figura 40b), cujo período de amostragem é única, isto é, todas as informações de evolução dos incêndios produzidas pelo AC estão disponíveis e são consideradas no cálculo de *fitness*. A taxa de amostragem não tem um impacto na reprodução do fenômeno já que não é necessário fazer esse ajuste, logo, nesse cenário, a abordagem sem o parâmetro teve um desempenho ligeiramente melhor que a que utiliza (Figura 40c), pois, nela é requerido calibrar, além dos parâmetros do modelo, a taxa. Por outro lado, quando partimos para um cenário em que a frequência de captura não é mais unitária, sendo a cada 2 passos de tempo (DS11), a abordagem sem o parâmetro (Figura 40d) teve uma intensidade muita abaixo que a desejada obtendo 0,76 quando se busca uma intensidade próxima a 2 um erro de 1,24 a menor e, quanto a inclinação, foi obtido um valor de 155,61º que é um erro de quase 25º. A abordagem com o parâmetro de amostragem produziu uma matriz bem similar à referência tanto em intensidade quanto em inclinação com um desempenho melhor do que a abordagem com parâmetro e $p = 1$.



(a) Ref. (b) GA-DS10 (c) GA-SR-DS10 (d) GA-DS11



(e) GA-SR-DS11 (f) GA-SR-DS12 (g) GA-SR-DS13 (h) GA-SR-DS14

Figura 40 – Modelo prob. matriz de preferência.

Para o DS12 (Figura 40f) foi encontrado o melhor resultado para a matriz W com uma intensidade de 2,04 e inclinação de 180,71º. Os experimentos com GA-SR-DS13 (Figura 40g) e GA-SR-DS14 (Figura 40h) obtiveram boas aproximações, porém com uma pequena inclinação para Noroeste de 5,06º e intensidade de 1,95 e apenas 0,05 com relação à referência. Por fim, GA-SR-DS14 apresentou uma leve inclinação para sudoeste de 3,52º e foi 0,12 mais intenso que o valor de referência.

A evolução temporal do AC ajustado pelo AG com $p = 1$ (Figura 41a) mostra que ambas as abordagens GA-DS10 e GA-SR-DS10 foram capazes de encontrar uma boa solução para a dinâmica de espalhamento do incêndio descrito na base de dados, havendo um comportamento médio bem similar conforme mostrado no mapa de calor (Figura 42a). Tanto para a onda primária quanto para as ondas internas, GA-SR-DS10 apresentou uma leve inclinação devido à composição da matriz de preferência do vento que tende a empurrar as chamas levemente para Noroeste.

Entretanto, a partir do período $p = 2$ a abordagem clássica sem o parâmetro de taxa de amostragem não foi capaz de produzir uma boa evolução (Figura 41b). Tal fato se torna mais evidente ao observamos o mapa de calor (Figura 42b) em que o comportamento mostrado pelo GA-DS11 teve um espelhamento muito intenso e com um formato muito ovalado diferente do ideal e, por isso, demostrando que um modelo com recuperação estocástica é mais desafiador para o AG pois há mais incerteza no processo de que no determinístico.

Por demonstrar que, sem o parâmetro SR, o AG não é capaz de se adaptar às informações presentes na entrada, os períodos $p = \{5, 7, 10\}$ apresentam apenas a evolução

temporal e o mapa de calor do AG com esse gene. GA-SR-DS12 com $p = 5$ foi o que apresentou melhor reprodução da matriz de preferência do vento que, combinados com bons valores LQ e LR levam a uma evolução temporal (Figura 41c) bem próxima à referência. De modo análogo, o mapa de calor (Figura 42c) tem ondas primárias e secundárias de propagação adequadas. Os experimentos seguintes GA-SR-DS13 e GA-SR-DS14 foram capazes de produzir uma boa evolução temporal (Figuras 41d e 41e), mas com uma forma um pouco mais alongada que a referência devido às matrizes encontradas possuírem uma ligeira inclinação, porém, de uma forma geral, o comportamento médio foi adequado conforme indicado pelos mapas de calor nas Figuras 42d e 42e.

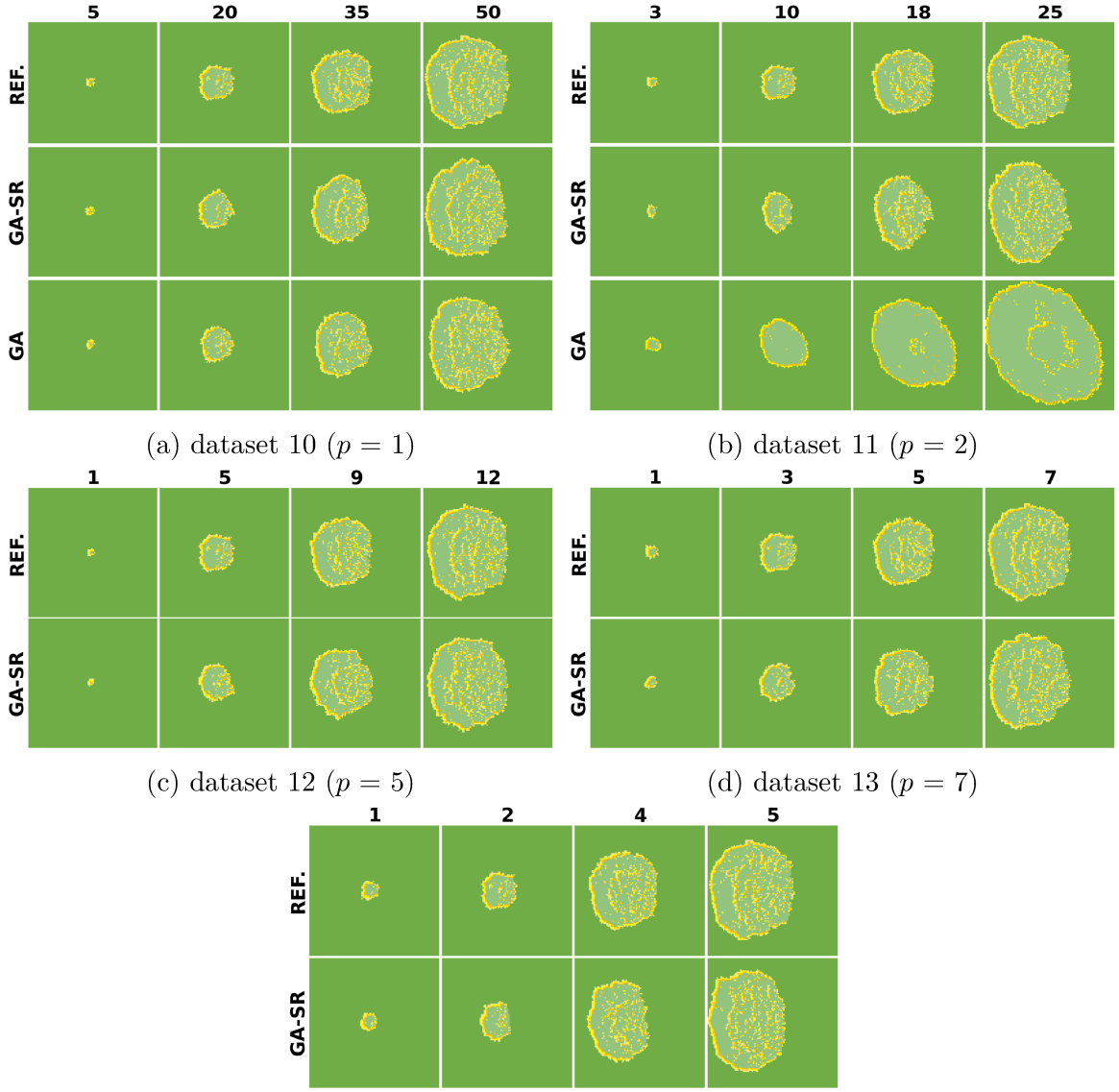
Experimentos com uma taxa de amostragem superior a $p = 10$ e com uma quantidade de informações na base de dados inferior a dez (10) registros começam a perder a qualidade do ajuste, pois é necessário um balanço adequado entre intervalo de amostragem e quantidade, já que os registros contidos devem ser capazes de descrever corretamente a evolução do fenômeno. Tal fato pode ser observado em $p = 5$ que produziu a melhor combinação de parâmetros do modelo, pois a base continha o tamanho mínimo necessário para a convergência e o intervalo de gravação estava com espaçamento adequado entre os registros facilitando, assim, para o AG convergir. A partir de $p = 7$, a qualidade do ajuste começou a decair havendo erro na inclinação do vetor de forças e na reprodução temporal, pois a quantidade de registros era menor que o estipulado mas, ainda assim, o AG foi capaz de encontrar uma boa aproximação, porém, o resultado é mais assertivo quando o balanço entre amostragem e quantidade de registros é garantido.

5.6.2 Análise de Robustez

Uma vez verificada a sensibilidade da abordagem em se adaptar a diferentes entradas, o passo seguinte é examinar a robustez reproduzindo o comportamento de um incêndio com base em dados produzidos pelo modelo de Tinoco et al. (2022) utilizado como validação de modo a se aproximar de um incêndio real, assim como na Seção 5.5.3.

Devido ao fato de mais aleatoriedade no processo de recuperação deixar a calibragem do modelo pelo AG mais complexa, foi gerado um *dataset* (DS15) que contabiliza apenas a ocorrência de incêndios, considerando o período de amostragem $p = 10$, intervalo esse mais desafiador nos experimentos anteriores para testar a capacidade da melhor função Fit9P em produzir uma boa solução.

Entretanto, conforme apresentado na Tabela 9, Fit9P não apresentou uma taxa de amostragem adequada tendo encontrado o valor de 1, LQ foi um valor próximo do ideal porém LR, que se refere ao tempo de recuperação foi encontrado um valor de 100 que, em conjunto a valores que produzem a matriz de preferência do vento, levaram a uma propagação bem diferente da referência conforme mostrado na Figura 43a. Tal diferença fica mais evidente no mapa de calor (Figura 43b) no qual o comportamento médio se assemelha mais a uma chama de um isqueiro do que ao espalhamento em uma área.



(a) dataset 10 ($p = 1$)

(b) dataset 11 ($p = 2$)

(c) dataset 12 ($p = 5$)

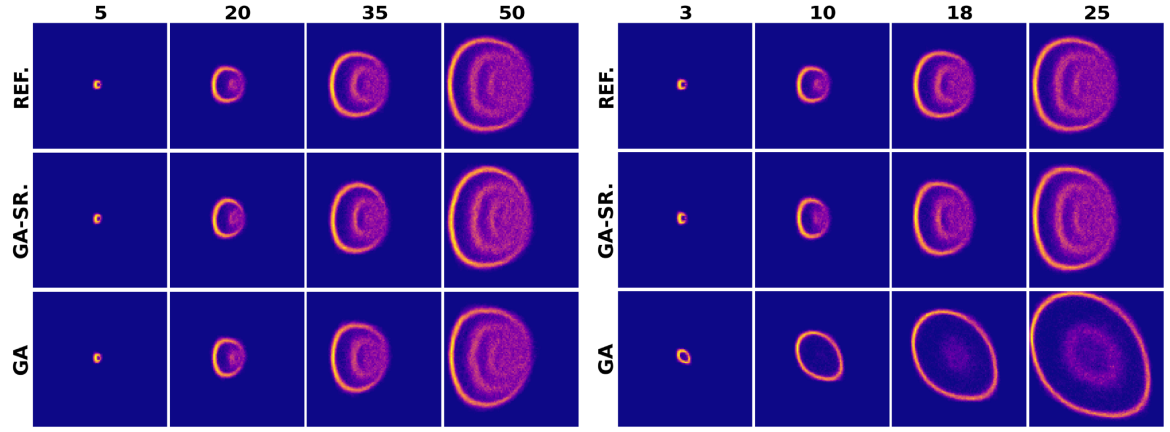
(d) dataset 13 ($p = 7$)

(e) dataset 14 ($p = 10$)

Figura 41 – Evolução temporal recuperação probabilística.

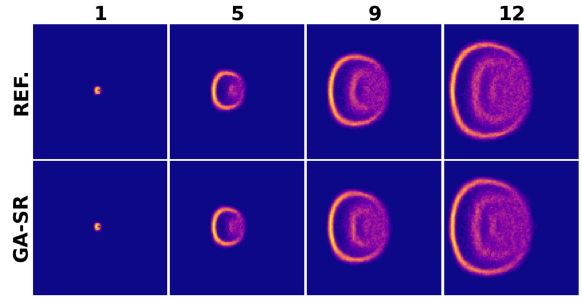
Por causa da função utilizada como referência no cenário de recuperação determinística não ter sido capaz de reproduzir corretamente o comportamento de espalhamento descrito na base de dados, fez-se necessário uma re-avaliação de desempenho das demais funções investigadas na Seção 5.5.1 considerando-se o impacto de diferentes informações (estados do AC) no cálculo de similaridade.

Considerando a quantidade de células no estado fogo para uma informação e para duas informações além de fogo, contabiliza-se a quantidade de células no estado de recuperação. Devido ao fato de a informação de células no estado fogo e em recuperação possuírem dimensões diferentes foram realizados experimentos preliminares para definir qual a melhor forma de unificar as duas informações e produzir um *fitness* único. Sendo escolhida como melhor a normalização pelo máximo, nela o valor máximo de fogo e recuperação em cada quadrante em análise é utilizado para normalizar os valores calculados que são

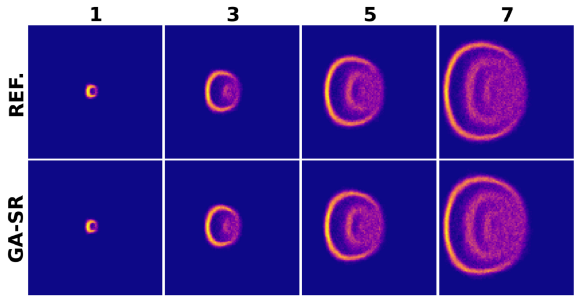


(a) dataset 10 ($p = 1$)

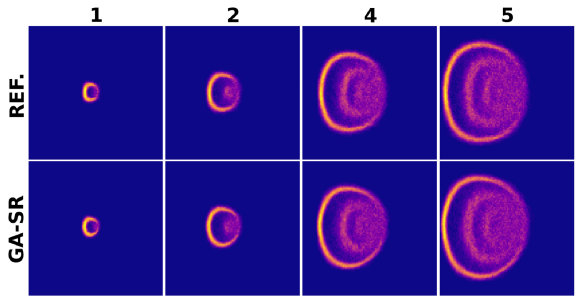
(b) dataset 11 ($p = 2$)



(c) dataset 12 ($p = 5$)



(d) dataset 13 ($p = 7$)



(e) dataset 14 ($p = 10$)

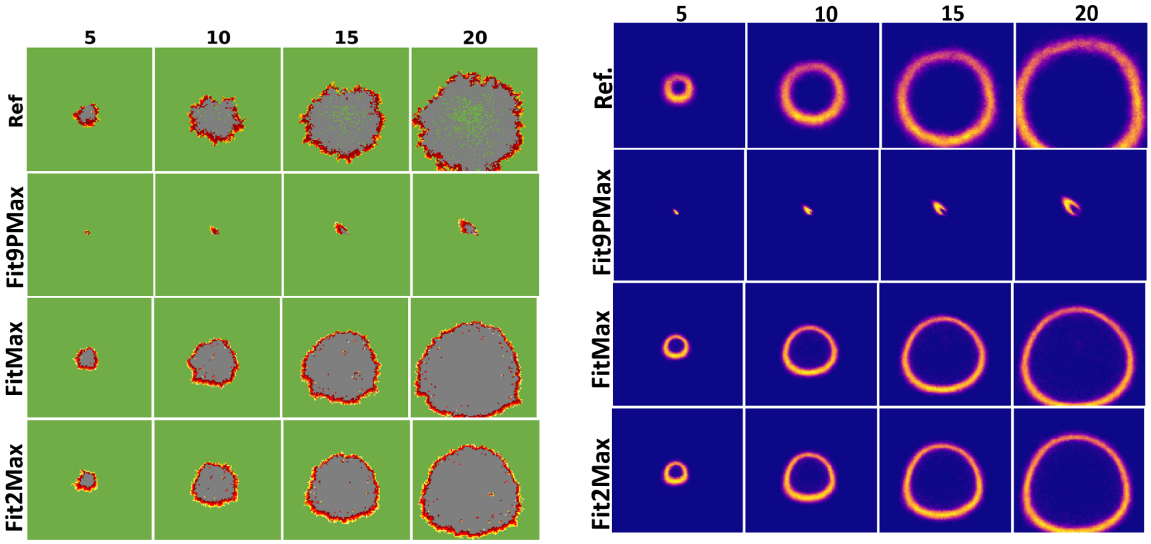
Figura 42 – Mapa de calor recuperação probabilística.

unificados conforme o número de informações analisadas. A forma detalhada de cálculo para cada função analisada é explicada na Seção 4.2.2, em que Fit9P com a normalização pelo máximo passou a se chamar Fit9PMax, F1 se tornou FitMax e F2 foi nomeado como Fit2Max de modo a melhor identificar qual função está sendo utilizada nos cálculos.

Diferente de Fit9PMax, as funções FitMax e Fit2Max contabilizando a quantidade de células no estado fogo foram capazes de produzir uma evoluções temporais próximas à referência conforme apresentado na Figura 43a e um comportamento médio adequado (Figura 43b). Fit2Max apresentou uma evolução um pouco mais alongada ao Sul do que a referência que tem um espalhamento na mesma direção, porém mais elíptico. A Tabela 9 mostra os melhores valores encontrados por cada função avaliada; a taxa de amostragem definida por FitMax e Fit2Max foi de 5, metade do tempo do modelo utilizado para geração da base de dados e foi adequada para reproduzir o comportamento descrito no

Tabela 9 – Teste funções: Fogo

Parâmetros	FitMax	Fit2Max	Fit9PMax
Gnw	0.451	0.663	0.007
Gn	0.993	0.712	0.009
Gne	0.590	0.848	0.003
Ge	0.665	0.366	0.000
Gse	0.576	0.346	1.051
Gs	0.055	0.130	0.017
Gsw	0.381	0.246	0.257
Gw	0.083	0.193	2.624
LQ	8	9	8
LR	650	984	100
SR	5	5	1
Fitness	2.328	0.164	58.874



(a) Evolução reticulado. (b) Mapa de calor.

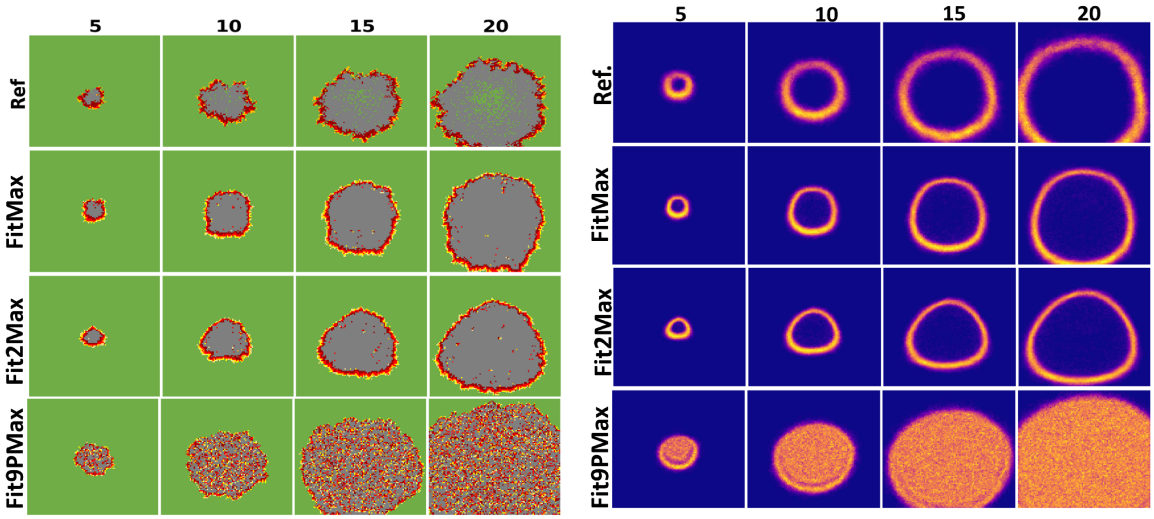
Figura 43 – Evolução autômato celular e mapa de calor: Fogo

dataset. Com relação ao tempo de queima FitMax encontrou um valor de 8 e Fit2Max de 9, bem próximos ao valor máximo de queima dez. Para o tempo de recuperação, FitMax encontrou um valor de 650, pouco mais da metade do valor máximo permitido de 1000; já Fit2Max encontrou o valor de 984. Tendo como base os resultados analisados, a função que mais se aproximou da evolução temporal e frequências adequadas foi FitMax.

Foi gerado um novo *dataset* (DS16) e inserida mais uma dimensão na análise considerando a existência de Fogo e áreas que sofreram queima e estão em processo de recuperação, permitindo que a nossa abordagem seja mais adaptável a modelos mais complexos sendo capaz de representar um número maior de características do fenômeno. A adição de mais complexidade no cálculo de *fitness*, levou a uma melhora do desempenho da função Fit9PMax mas, ainda assim não foi capaz de apresentar um desempenho satisfatório ao

Tabela 10 – Teste funções: Fogo, Queimado

Parâmetros	FitMax	Fit2Max	Fit9PMax
Gnw	0.632	0.632	0.059
Gn	0.963	0.441	0.400
Gne	0.762	0.808	0.087
Ge	0.188	0.507	0.274
Gse	0.200	0.691	0.225
Gs	0.236	0.081	0.029
Gsw	0.108	0.535	0.112
Gw	0.243	0.038	0.010
LQ	9	9	10
LR	710	812	33
SR	5	5	15
Fitness	1.537	0.166	2.101



(a) Evolução reticulado. (b) Mapa de calor.

Figura 44 – Evolução autômato celular e mapa de calor: Fogo, Queimado.

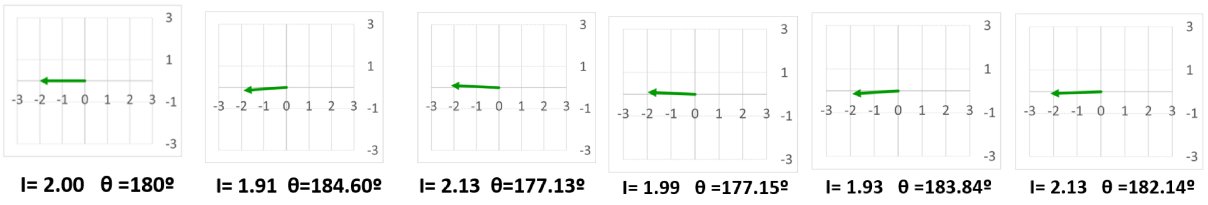
superestimar a ocorrência de chamas levando a uma área maior de espalhamento além de focos internos bastante evidentes. Tal fato fica ainda mais evidente no mapa de calor ao ser possível observar que a área está preenchida de laranja. Por outro lado, FitMax e Fit2Max obtiveram boas aproximações; Fit2Max teve um comportamento triangular com arestas arredondas; e FitMax continuou tendo o melhor desempenho mantendo o comportamento radial com uma área de espalhamento próxima à referência tanto na evolução temporal quanto no mapa de calor.

5.6.3 Análise Sensibilidade FitMax

Uma vez que a função FitMax foi capaz de reproduzir o comportamento do modelo de validação, nesse experimento o objetivo é verificar se ela é capaz de reproduzir corretamente o comportamento do utilizado como referência. Para isso, foram gerados *datasets*

Tabela 11 – Análise sensibilidade modelo referência

Parâmetros	DS17	DS18	DS19	DS20	DS21
Gnw	0.023	0.069	0.110	0.100	0.073
Gn	0.740	0.537	0.601	0.686	0.829
Gne	0.875	0.887	0.838	0.793	0.731
Ge	0.234	0.157	0.107	0.074	0.091
Gse	0.981	1.182	1.133	0.980	1.205
Gs	0.042	0.034	0.051	0.119	0.073
Gsw	0.605	0.743	0.854	0.494	0.669
Gw	0.830	0.782	0.679	0.863	0.845
LQ	3	3	3	3	3
LR	32	35	31	32	35
SR	1	2	5	7	10
Fitness	2.870	1.361	0.597	0.435	0.302



(a) Ref (b) DS17 (c) DS18 (d) DS19 (e) DS20 (f) DS21

Figura 45 – Matriz de preferência modelo referência.

considerando células em ignição e recuperação para os períodos $p = \{1, 2, 5, 7, 10\}$ utilizados no experimento de sensibilidade. Os valores utilizados como referência para a matriz W e $LQ = 3$ se mantiveram e ajustou-se $LR = 30$.

Os cromossomos produzidos pela função em cada cenário são apresentados na Tabela 11. O tempo de queima LR foi encontrado em todos os experimentos; com relação ao tempo de recuperação LR os valores foram adequados, com variação ligeiramente maior para DS18 e DS21 que encontraram o valor de 35 (5 unidades mais intenso que a referência). As taxas de amostragem também foram corretamente calibradas em todos os experimentos.

As matrizes de preferência do vento encontradas também foram boas soluções: DS17 (Figura 45b) obteve uma intensidade de 1,91 e inclinação de 184,60º, uma diferença de 0,09 e 4,60º; DS18 (Figura 45c) obteve uma intensidade de 2,13 e inclinação de 177,13°, uma diferença de 0,13 e 2,87; DS19 (Figura 45d) foi a melhor matriz encontrada com intensidade 1,99 e inclinação de 177,15º, uma diferença de 0,01 e 2,85º; DS20 (Figura 45e) produziu uma intensidade de 1,93 e inclinação de 183,84º, uma diferença de 0,07 e 3,84º; e DS21 (Figura 45f) obteve a inclinação mais próxima com um erro de 2,14º tendo encontrado uma inclinação de 182,14º e intensidade de 2,13, uma diferença de 0,13.

Devido ao fato de terem sido encontrados bons arranjos cromossômicos ao se utilizar a função FitMax, foram produzidas matrizes de preferência do vento, tempos de queima

e recuperação capazes de levar a frequências e evoluções temporais capazes de reproduzir adequadamente a dinâmica de propagação de incêndio descritas nos *datasets* conforme Figura 46, havendo um leve destaque para DS19 pois foi o que mais se aproximou da solução ideal nas métricas numéricas e gráficas.

5.6.4 Análise de Adaptabilidade

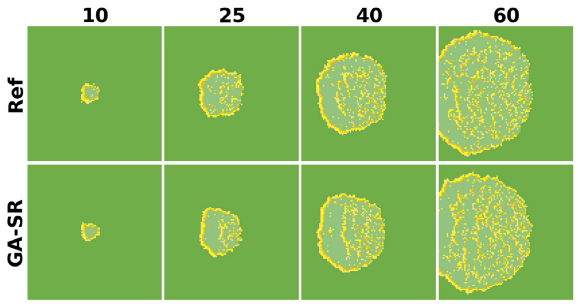
Após identificar qual a função mais adequada para cenários de recuperação determinística e probabilística, com o intuito de verificar se a abordagem é capaz de reproduzir incêndios com diferentes intensidades e direções para o modelo de avaliação da robustez, foram avaliadas três novas direções Nordeste e Oeste 25% mais intensas e direção Sul com uma intensidade menor o que requer mais tempo para a evolução.

A Tabela 12, apresenta os cromossomos encontrados para cada experimento; para a direção Nordeste foi encontrado o tempo de queima $LQ = 9$, quase o máximo, e o tempo máximo de recuperação $LR = 1000$; Já Oeste definiu o tempo de queima $LQ = 10$ como máximo e recuperação próximo $LR = 973$; a direção Sul por ter uma intensidade de propagação menor produziu um tempo de queima menos intenso $LQ = 7$ e recuperação alta $LR = 978$. Para a taxa de amostragem os valores encontrados foram em média próximos a 5, um valor adequado para a reprodução da dinâmica e que nos leva a crer que o modelo de referência necessita da metade do tempo para obter mesmo comportamento da entrada.

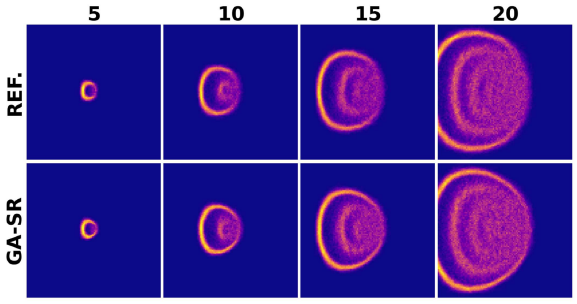
Tabela 12 – Teste de adaptabilidade

Parâmetros	Nordeste	Oeste	Sul
Gnw	0.704	0.324	0.729
Gn	0.222	0.573	1.318
Gne	0.231	0.857	0.726
Ge	0.584	0.012	0.720
Gse	0.170	0.998	0.749
Gs	1.351	0.296	0.098
Gsw	0.815	0.555	0.453
Gw	0.771	0.830	0.197
LQ	9	10	7
LR	1000	973	978
SR	5	6	6
Fitness	1.112	1.131	1.498

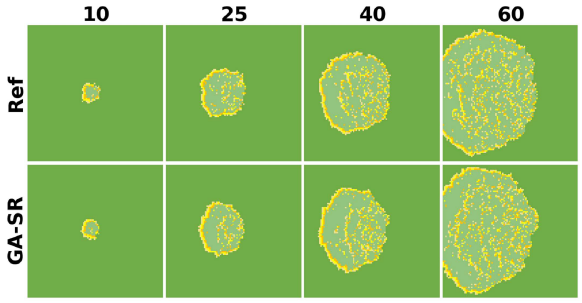
A evolução temporal, considerando a direção Oeste, é apresentada na Figura 47a e pode-se perceber que a área de espalhamento é bem similar e também que o AG foi capaz de encontrar a direção representada no *dataset*. A frequência da área de espalhamento é muito similar à referência com o mesmo tamanho praticamente.



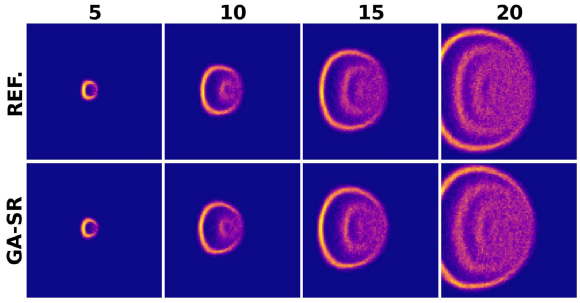
(a) dataset 17 ($p = 1$)



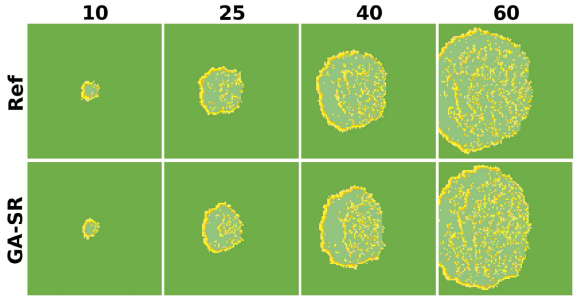
(b) dataset 17 ($p = 1$)



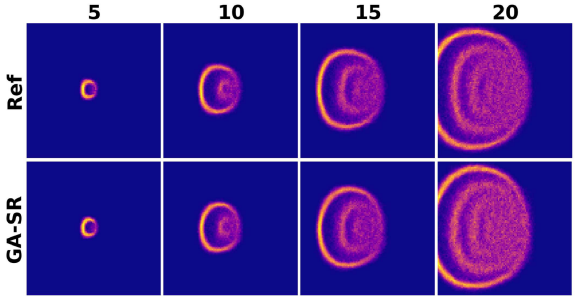
(c) dataset 18 ($p = 2$)



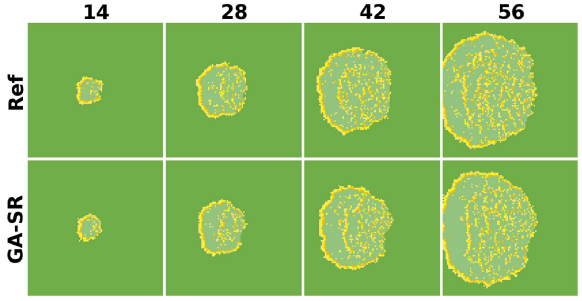
(d) dataset 18 ($p = 2$)



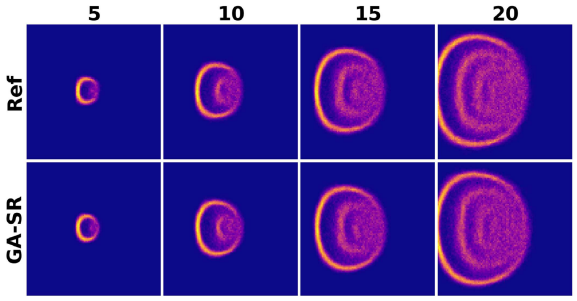
(e) dataset 19 ($p = 5$)



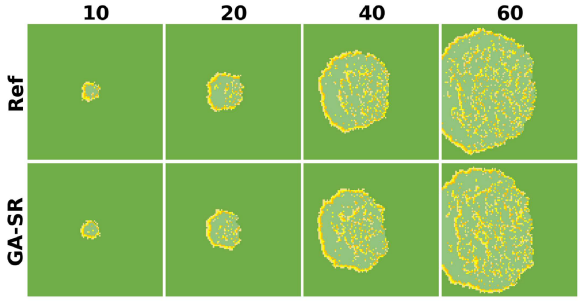
(f) dataset 19 ($p = 5$)



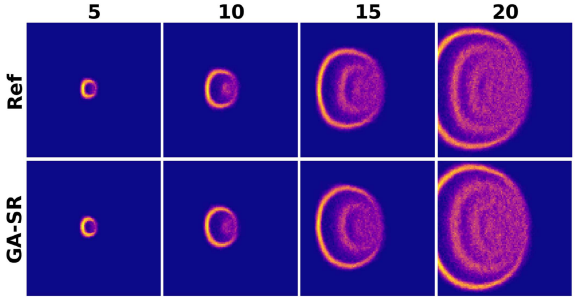
(g) dataset 20 ($p = 7$)



(h) dataset 20 ($p = 7$)

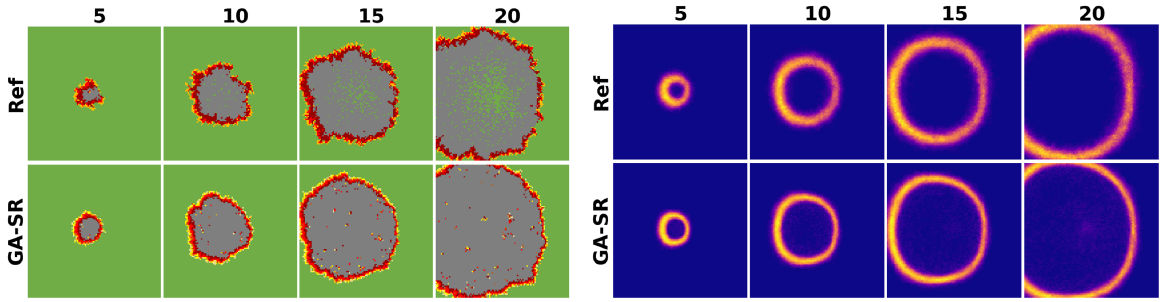


(i) dataset 21 ($p = 10$)



(j) dataset 21 ($p = 10$)

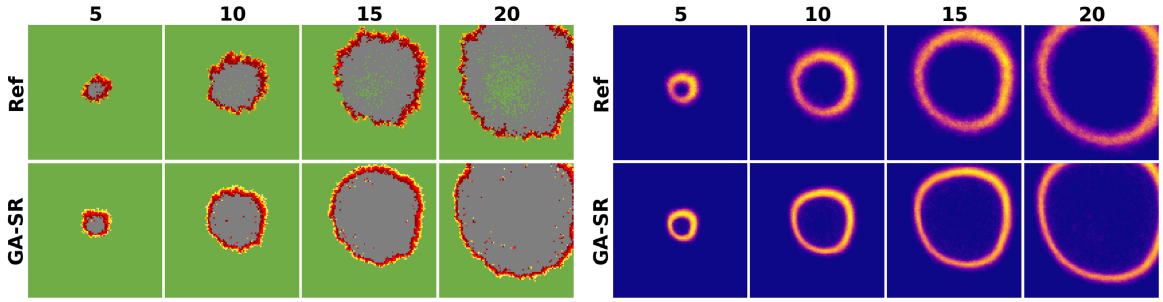
Figura 46 – GA e GA-SR evolução temporal configuração modelo referência.



(a) Evolução reticulado. (b) Mapa de calor.

Figura 47 – Reticulado e mapa de calor Oeste.

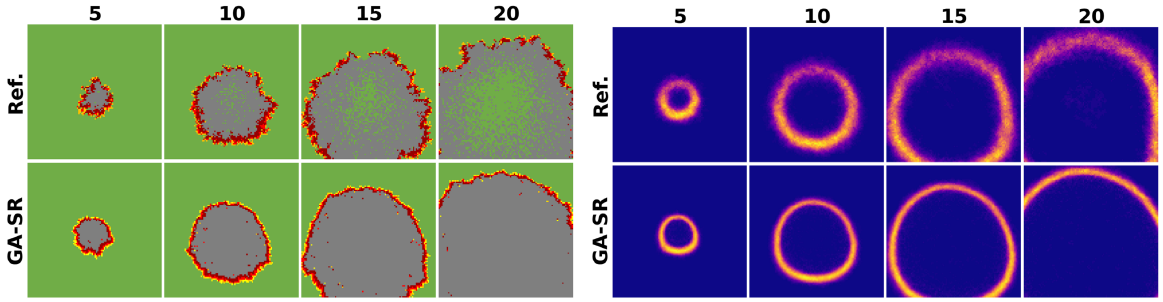
A direção Nordeste (Figura 48) é uma das mais difíceis para o AG identificar, pois, uma vez que o incêndio se propaga em diagonal, ele rapidamente alcança as bordas do reticulado. Conforme evolução temporal do AC (Figura 48a), o AG foi capaz de encontrar uma evolução de direção e intensidade das chamas próximas à referência, sendo bastante evidente no comportamento médio mostrado no mapa de calor (Figura 48b).



(a) Evolução reticulado. (b) Mapa de calor.

Figura 48 – Reticulado e mapa de calor Nordeste.

Para a direção Sul (Figura 49) a intensidade de propagação das chamas foi definida como menos intensa, desta forma, são necessários mais passos de tempo (300 passos) para o fogo se espalhe completamente. Por causa disso para que se contabilize a evolução completa considerando 20 registros na base o período de gravação foi definido como $p = 15$. Desta forma, o intervalo entre os registros é ainda mais esparso que nos experimentos anteriores, ainda assim, a abordagem foi capaz de identificar a correta dinâmica de propagação produzindo uma evolução temporal semelhante à referência (Figura 49a), sendo bastante evidente no comportamento médio mostrado no mapa de calor (Figura 49b).



(a) Evolução reticulado. (b) Mapa de calor.

Figura 49 – Reticulado e mapa de calor Sul.

Capítulo 6

Conclusão

Este trabalho propôs uma abordagem evolutiva baseada em AG (EVOCA-WiSM) para ajuste automático de parâmetros de modelos de propagação de incêndio baseados em Autômatos Celulares, para reprodução da dinâmica observada na propagação de incêndios florestais. O fenômeno a ser representado pode ser uma sequência de reticulados gerados pelo próprio modelo, por outro modelo de propagação de incêndio ou uma sequência de imagens digitalizadas de um incêndio real. Entretanto, é necessário uma quantidade mínima de registros na base de dados que, devem estar corretamente espaçados e representar a evolução do fenômeno ao longo do tempo para que seja possível a convergência sem perder a qualidade do ajuste. Foi possível observar uma redução na eficiência do AG para conjuntos de dados pequenos (tamanho < 10), com 20 registros sendo suficientes para a reprodução do comportamento do fogo mesmo para conjuntos de dados mais esparsos.

Dessa forma, estabelecemos que é relevante nesta investigação que esse ajuste evolutivo de parâmetros possa reproduzir adequadamente a dinâmica considerando diferentes taxas de amostragem para gerar o conjunto de dados de referência. Nesse sentido, um novo parâmetro SR (taxa de amostragem) foi ajustado pelo AG, que determina quais reticulados de uma simulação de AC serão usados para avaliar a similaridade do comportamento em relação aos dados de referência, sendo essa a principal contribuição deste trabalho.

Os experimentos mostraram que a inclusão deste novo parâmetro tornou a abordagem mais versátil, capaz de se adaptar a modelos e entradas distintas, especialmente ao reproduzir conjuntos com dados mais esparsos e permitiu, uma representatividade mais assertiva dos fenômenos estudados. Quando comparado a uma abordagem tradicional sem o uso deste parâmetro em que, o AG não foi capaz de reproduzir o fenômeno corretamente.

Outro fator importante para que um AG seja capaz de encontrar uma boa solução, é a função de aptidão, responsável por avaliar a qualidade da combinação cromossômica na resolução de um problema. Sendo assim, ela guia a busca pelo espaço de soluções com o intuito de verificar qual a mais adequada. Foram avaliadas diferentes funções, Fit9P proposta por Ferreira et al. (2022) e duas funções para modelos SIR (F1 e F2)

provenientes do trabalho de Monteiro, Gandini e Schimit (2020) foram adaptadas para o cenário de propagação de incêndio (F1P9 e F2P9) considerando a divisão do reticulado em nove partes.

Foram inclusas mais informações no cálculo de *fitness* além de a existência de chamas, com o objetivo de avaliar se adicionar mais características do fenômeno possibilita um ajuste mais preciso. Tal premissa se mostrou verdadeira, tendo a função FitMax (utilizando a normalização pelo máximo em cada partição) contabilizando a existência de fogo e áreas em recuperação e foi a que apresentou o melhor resultado tanto em um cenário de recuperação determinística quanto probabilística. Uma vez que a inclusão de aleatoriedade no processo de recuperação tornou a convergência do AG mais complexa e Fit9P não foi capaz de reproduzir adequadamente a dinâmica apresentada no modelo de validação (TINOCO et al., 2022).

6.1 Trabalhos Futuros

Como trabalhos futuros, pretendemos avaliar a inclusão de mais informações, estados do AC, no cálculo de *fitness* da FitMax de modo a verificar se ela consegue capturar estas características. Além disto pretende-se avaliar a inclusão de métricas usadas na avaliação do desempenho do modelo como mapa de calor, que permite visualizar o comportamento médio do incêndio, o que permitiria uma melhoria na qualidade dos resultados alcançados pelo algoritmo. Entretanto em sua geração são necessárias 100 simulações, o que aumentaria consideravelmente o o custo computacional de cada avaliação. Deste modo, é necessária uma investigação de alternativas de implementação que balanceiem o custo computacional e tempo de execução no cálculo de aptidão.

Também pretende-se testar abordagens co-evolutivas ou a multiestágio proposta por (FRAGA; OLIVEIRA; MARTINS, 2021) a fim de verificar se ao evoluir diferentes parâmetros separadamente como por exemplo, a matriz de preferência do vento, e os parâmetros de queima, recuperação e a taxa de amostragem leva a um refinamento mais preciso.

Visando a simplificação do conjunto de parâmetros do modelo sem afetar a representatividade do fenômeno modelado, pretende-se avaliar uma alternativa para representar a influência do vento na simulação que considere dois parâmetros (intensidade, amplitude) diminuindo o custo computacional no ajuste e tornando o algoritmo mais adaptável a diferentes modelos onde não há relação de localidade entre os parâmetros, uma vez que onde os parâmetros da matriz são dispostos de acordo com suas respectivas posições para compor o indivíduo. Nesse sentido, o uso de operadores de recombinação que favorecem uma busca local mais eficiente ao misturar mais as características dos genes como um *crossover* uniforme, ou de uma recombinação própria para números reais como o *crossover* aritmético possibilitam uma exploração local mais eficiente ao realizarem ajustes sutis nos valores dos genes. Isso permite utilizar uma taxa de mutação mais baixa, fazendo com

que o operadores sejam usados mais raramente e para provocar alterações mais significativas, possibilitando a exploração de novas regiões no espaço de busca (exploração global); desta forma, pretende-se avaliar o uso uma mutação normal que permita alterações mais significativas, ou seja, gerar uma gaussiana e aplicar o desvio padrão.

Na abordagem atual a taxa de elitismo de 25% é alta para propiciar a manutenção de características favoráveis para a convergência da abordagem evolutiva permitindo uma melhor combinação de fatores como queima, recuperação e influência do vento, uma vez que, nas gerações iniciais devido a diferentes magnitudes de valores possíveis (especificamente entre LQ e LR) os indivíduos são bem diferentes; mas a medida que o AG evolui as gerações finais essa dissimilaridade não é tão presente. Como alternativa um elitismo adaptativo que começa com taxas mais altas que são decrementadas ao longo das gerações pode proporcionar um refinamento maior das soluções encontradas pelo AG.

Pretende-se avaliar a eficácia da nossa abordagem na reprodução da dinâmica real de propagação de incêndios florestais, sendo necessário a geração dos dados para comparação a partir das imagens, que devem ser discretizadas para gerar os reticulados; e investigar o ajuste de modelos de fogo mais complexos, que têm mais parâmetros e são, consequentemente mais desafiadores para o AG mas, ao mesmo tempo, permitem representar uma gama maior de informações do fenômeno estudado, pois uma das limitações da abordagem se deve as características possíveis de representação utilizando o modelo de referência.

Nesta pesquisa avaliou-se apenas cenários com uma vegetação homogênea, entretanto, em ambientes reais a vegetação tende a ser mais heterogênea, desta forma, investigar essa característica e incluir novas podem melhorar a capacidade de reprodução do fenômeno pelo modelo de referência. Por fim a abordagem pode ser estendida para que o AG seja capaz de planejar ações de combate aos incêndios; por exemplo, escolher mudanças na vegetação (capina e poda de vegetação seca (diminuição de material combustível), molhar determinadas células, inclusão de barreiras, dentre outros) e usar o modelo de AC para avaliar o impacto dessas ações na propagação das chamas.

Os trabalhos futuros descritos, são listados por melhorias em cada técnica (AG e AC):

Algoritmo Genético:

- ❏ Inclusão de mais informações no cálculo do *fitness* considerando mais estados do AC ou mapa de calor;
- ❏ Avaliar a abordagens co-evolutivas ou multiestágio para refinamento dos parâmetros;
- ❏ Avaliar a eficácia da nossa abordagem na reprodução da dinâmica de incêndios florestais reais;
- ❏ Investigar o uso de diferentes modelos como referência, a fim de, verificar de refinar a reprodução do dinâmica de propagação;

- ❏ Avaliar formas de representar a influência do vento considerando-se dois parâmetros que representam a intensidade e amplitude;
- ❏ Avaliar o uso de *crossover* uniforme ou aritmético e de uma mutação normal para tornar o AG mais genérico e adaptável a diferentes modelos e características de um incêndio florestal;
- ❏ Estender a abordagem permitindo que o AG possa planejar ações de combate ao espalhamento de incêndios florestais;
- ❏ Incluir um elitismo adaptativo que permite a abordagem se adequar as características de manutenção de características do problema e promove uma maior diversidade.

Autômato Celular:

- ❏ Estender a abordagem para lidar com vegetação heterogênea, possibilitando o monitoramento de cenários florestais mais realísticos;
- ❏ Incorporar novas características do ambiente no modelo de propagação de incêndio, como umidade, relevo, entre outros, a fim de melhorar sua capacidade de representação de diferentes dinâmicas de espalhamento do fogo;
- ❏ Investigar modelos de propagação mais complexos capazes de representar mais características do fenômeno;
- ❏ Adaptar o modelo para ser compatível com outras formas de representação da influência do vento.

6.2 Contribuições em Produção Bibliográfica

Esta pesquisa resultou na publicação do artigo *"Evolutionary Adjustment of a Cellular Automata-based Model for Wildfire Spreading"* (MURILO; OLIVEIRA; MARTINS, 2025) na 34th *Brazilian Conference on Intelligent Systems* (BRACIS) e, apresenta nossa abordagem evolutiva para ajuste dos parâmetros do modelo de referência com recuperação determinística, destacando a inclusão do novo parâmetro "taxa de amostragem" ajustado pelo AG e sua importância na adaptabilidade da abordagem.

Foi submetido e aguarda-se a resposta do evento IEEE *Congress on Evolutionary Computation* (CEC) o artigo intitulado "*Evolutionary Adjustment of Probabilistic Models based on Cellular Automata applied to Fire Propagation*". No artigo se avalia o impacto da taxa de amostragem e de diferentes funções no cenário de recuperação probabilística do modelo de referência.

Referências

ABDOUN, O.; ABOUCHABAKA, J.; TAJANI, C. Analyzing the performance of mutation operators to solve the travelling salesman problem. **arXiv preprint arXiv:1203.3099**, 2012. Disponível em: <https://doi.org/10.48550/arXiv.1203.3099>.

ACKER, D. **Elementary cellular automaton simulator — devinacker.github.io**. 2014. <https://devinacker.github.io/celldemo/>. [Accessed 20-05-2024].

ADEYEMO, A.; WIMMER, H.; POWELL, L. Effects of normalization techniques on logistic regression in data science. In: **2018 Proceedings of the Conference on Information Systems Applied Research ISSN**. Norfolk, Virginia: ISSN, 2018. p. 2167–1508.

ALEXANDRIDIS, A. et al. A cellular automata model for forest fire spread prediction: The case of the wildfire that swept through spetses island in 1990. **Applied Mathematics and Computation**, v. 204, n. 1, p. 191–201, 2008. Disponível em: <https://doi.org/10.1016/j.amc.2008.06.046>.

AMOL, D. R. Prediction of fire propagation in forest areas using genetic algorithm. In: EDP SCIENCES. **ITM Web of Conferences**. 2020. v. 32, p. 03046. Disponível em: <https://doi.org/10.1051/itmconf/20203203046>.

ARCA, B.; GHISU, T.; TRUNFIO, G. A. Gpu-accelerated multi-objective optimization of fuel treatments for mitigating wildfire hazard. **Journal of Computational Science**, Elsevier, v. 11, p. 258–268, 2015. Disponível em: <https://doi.org/10.1016/j.jocs.2015.08.009>.

BHANJA, S.; DAS, A. Impact of data normalization on deep neural network for time series forecasting. arxiv 2018. **arXiv preprint arXiv:1812.05519**. Disponível em: <https://doi.org/10.48550/arXiv.1812.05519>.

BIERNATH, A. **Los Angeles: incêndios florestais vão ficar cada vez mais intensos e prolongados, alerta cientista da ONU - BBC News Brasil — bbc.com**. 2025. <https://www.bbc.com/portuguese/articles/c2egx388rp2o>. [Accessed 03-02-2025].

BOTERS-PITARCH, J. et al. An intelligent cellular automaton scheme for modelling forest fires. **Ecological Informatics**, Elsevier, v. 80, p. 102456, 2024. Disponível em: <https://doi.org/10.1016/j.ecoinf.2023.102456>.

BRASIEL, H. C.; LIMA, D. A. Exploring the influence of wind, vegetation and water sources on the spread of forest fires in the brazilian cerrado biome using cellular automata. In: SBC. **Anais do XIV Workshop de Computação Aplicada à Gestão do Meio Ambiente e Recursos Naturais**. 2023. p. 61–70. Disponível em: <DOI:https://doi.org/10.5753/wcama.2023.230476>.

BRODLAND, G. W. How computational models can help unlock biological systems. In: ELSEVIER. **Seminars in cell & developmental biology**. 2015. v. 47, p. 62–73. Disponível em: <https://doi.org/10.1016/j.semcdb.2015.07.001>.

CAI, Z.; PENG, Z. Cooperative coevolutionary adaptive genetic algorithm in path planning of cooperative multi-mobile robot systems. **Journal of Intelligent and Robotic Systems**, Springer, v. 33, n. 1, p. 61–71, 2002. Disponível em: <https://doi.org/10.1023/A:1014463014150>.

CARDOSO, D. M. A. **MG: Uberlândia é afetada por fumaça das queimadas no interior de SP | CNN Brasil — cnnbrasil.com.br**. 2024. <https://www.cnnbrasil.com.br/nacional/mg-uberlandia-e-afetada-por-fumaca-das-queimadas-no-interior-de-sp/>. [Accessed 03-02-2025].

CARRILLO, C. et al. Impact of mixed-precision: A way to accelerate data-driven forest fire spread systems. In: **Int. Conf. on Computational Science**. [s.n.], 2023. p. 62–76. Disponível em: <https://doi.org/10.1007/978-3-031-36021-3_5>.

CASTRO, M. L. A.; CASTRO, R. de O. Autômatos celulares: implementações de von neumann, conway e wolfram. **Revista de Ciências Exatas e Tecnologia**, v. 3, n. 3, p. 89–106, 2008. Disponível em: <https://doi.org/10.17921/1890-1793.2008v3n3p89-106>.

CATCHPOLE, E.; CATCHPOLE, W.; ROTHERMEL, R. Fire behavior experiments in mixed fuel complexes. **International Journal of Wildland Fire**, CSIRO Publishing, v. 3, n. 1, p. 45–57, 1993. Disponível em: <https://doi.org/10.1071/WF9930045>.

CERRUTI, U.; DUTTO, S.; MURRU, N. A symbiosis between cellular automata and genetic algorithms. **Chaos, Solitons & Fractals**, Elsevier, v. 134, p. 109719, 2020. Disponível em: <https://doi.org/10.1016/j.chaos.2020.109719>.

CHOPARD, B.; DROZ, M. **Cellular Automata Modeling of Physical Systems**. [s.n.], 1998. Disponível em: <https://doi.org/10.1017/CBO9780511549755>.

EDO, G. I. et al. Impact of environmental pollution from human activities on water, air quality and climate change. **Ecological Frontiers**, Elsevier, 2024. Disponível em: <https://doi.org/10.1016/j.ecofro.2024.02.014>.

FERREIRA, H. F. et al. Stochastic model for wildfire simulation based on the characteristics of the brazilian cerrado. In: SPRINGER. **International Conference on Artificial Intelligence and Soft Computing**. 2023. p. 487–496. Disponível em: <https://doi.org/10.1007/978-3-031-42505-9_41>.

FERREIRA, M. E. A. et al. Automatic evolutionary adjustment of cellular automata model for forest fire propagation. In: SPRINGER. **International Conference on Cellular Automata for Research and Industry**. 2022. p. 235–245. Disponível em: <https://doi.org/10.1007/978-3-031-14926-9_21>.

FRAGA, L. M.; OLIVEIRA, G. M. de; MARTINS, L. G. Adjustment of an epidemiological cellular automata-based model using genetic algorithm. In: IEEE. **2020 IEEE 32nd International Conference on Tools with Artificial Intelligence (ICTAI)**. 2020. p. 589–594. Disponível em: <https://doi.org/10.1109/ICTAI50040.2020.00096>.

_____. Multistage evolutionary strategies for adjusting a cellular automata-based epidemiological model. In: IEEE. **2021 IEEE Congress on Evolutionary Computation (CEC)**. 2021. p. 466–473. Disponível em: <https://doi.org/10.1109/CEC45853.2021.9504738>.

FREIRE, J. G.; DACAMARA, C. Using cellular automata to simulate wildfire propagation and to assist in fire management. **Natural Hazards and Earth System Sciences**, 2019. Disponível em: <https://doi.org/10.5194/nhess-2018-227>.

GARCÍA-DURO, J. et al. Hidden costs of modelling post-fire plant community assembly using cellular automata. In: SPRINGER. **International Conference on Cellular Automata**. 2018. p. 68–79. Disponível em: <https://doi.org/10.1007/978-3-319-99813-8_6>.

GARCÍA-PEDRAJAS, N.; HERVÁS-MARTÍNEZ, C.; MUÑOZ-PÉREZ, J. Covnet: a cooperative coevolutionary model for evolving artificial neural networks. **IEEE Transactions on Neural Networks**, IEEE, v. 14, n. 3, p. 575–596, 2003. Disponível em: <https://doi.org/10.1109/TNN.2003.810618>.

GERARDI, D.; MONTEIRO, L. et al. System identification and prediction of dengue fever incidence in rio de janeiro. **Mathematical Problems in Engineering**, Hindawi, v. 2011, 2011. Disponível em: <https://doi.org/10.1155/2011/720304>.

GHISU, T. et al. An improved cellular automata for wildfire spread. **Procedia Computer Science**, v. 51, p. 2287–2296, 2015. Disponível em: <https://doi.org/10.1016/j.procs.2015.05.388>.

GOBRON, S.; BONAFOS, H.; THALMANN, D. Gpgpu computation and visualization of three-dimensional cellular automata. 08 2010. Disponível em: <https://doi.org/10.1007/s00371-010-0515-1>.

GOLDBERG, D. E. Genetic algorithms in search. **Optimization, and Machine Learning**, Addison-Wesley, 1989.

HANDA, H. et al. Genetic algorithm involving coevolution mechanism to search for effective genetic information. In: IEEE. **Proceedings of 1997 IEEE International Conference on Evolutionary Computation (ICEC'97)**. 1997. p. 709–714. Disponível em: <https://doi.org/10.1109/ICEC.1997.592427>.

HOLLAND, J. H. Genetic algorithms. **Scientific american**, JSTOR, v. 267, n. 1, p. 66–73, 1975. Disponível em: <https://doi.org/10.1038/scientificamerican0792-66>.

IBAMA. **Manejo do fogo**. 2023. <https://www.gov.br/ibama/pt-br/hotsites/amazoniasemincendios/manejo-do-fogo>. [Accessed 24-11-2023].

Instituto Nacional de Pesquisas Espaciais (INPE). **Boletim Mensal Integrado dos Avisos de Desmatamento, Degradação e Queimadas no Brasil**. 2024. <https://dataserver-coids.inpe.br/queimadas/queimadas/Boletins-Integrados/2024/07_2024.pdf>. [Accessed 03-02-2025].

JAIN, J. Amplitude normalization and its application to speech coding. **IEEE**, 1986. Disponível em: <https://doi.org/10.1109/ICASSP.1986.1168986>.

JR, M. N. G. et al. Global optimization-based calibration algorithm for a 2d distributed hydrologic-hydrodynamic and water quality model. **arXiv preprint arXiv:2308.16864**, 2023. Disponível em: <https://doi.org/10.48550/arXiv.2308.16864>.

KIM, Y. K.; KIM, J. Y.; KIM, Y. A coevolutionary algorithm for balancing and sequencing in mixed model assembly lines. **Applied intelligence**, Springer, v. 13, n. 3, p. 247–258, 2000. Disponível em: <http://dx.doi.org/10.1023/A:1026568011013>.

KORA, P.; YADLAPALLI, P. Crossover operators in genetic algorithms: A review. **International Journal of Computer Applications**, Foundation of Computer Science, v. 162, n. 10, 2017. Disponível em: <https://doi.org/10.5120/ijca2017913370>.

LAMBORA, A.; GUPTA, K.; CHOPRA, K. Genetic algorithm-a literature review. In: IEEE. **2019 international conference on machine learning, big data, cloud and parallel computing (COMITCon)**. 2019. p. 380–384. Disponível em: <https://doi.org/10.1109/COMITCon.2019.8862255>.

LEI, W.; ZHIQIANG, D. Modeling and simulation of forest fire detection and fire spread. In: **2017 2nd International Conference on Advanced Robotics and Mechatronics (ICARM)**. [s.n.], 2017. p. 65–69. Disponível em: <https://doi.org/10.1109/ICARM.2017.8273136>.

LI, X. et al. A parallel genetic algorithm with variable neighborhood search for the vehicle routing problem in forest fire-fighting. **IEEE Transactions on Intelligent Transportation Systems**, 2024. Disponível em: <https://doi.org/10.1109/TITS.2024.3395930>.

LIMA, D.; LIMA, H. Autômatos celulares estocásticos bidimensionais aplicados a simulação de propagação de incêndios em florestas homogêneas. In: SBC. **Anais do V Workshop de Computação Aplicada a Gestão do Meio Ambiente e Recursos Naturais**. 2014. p. 15–24. Disponível em: <https://sol.sbc.org.br/index.php/wcama/article/view/10907>.

LINDEN, R. **Algoritmos genéticos: uma importante ferramenta da inteligência computacional**. Rio de Janeiro: Brasport, 2006.

LIU, B. et al. Permutation genetic algorithm for stacking sequence design of composite laminates. **Computer methods in applied mechanics and engineering**, Elsevier, v. 186, n. 2-4, p. 357–372, 2000. Disponível em: <https://doi.org/10.1016/S0045-7825(99)90391-2>.

LIU, H.; LIU, S.; TIAN, Y. Flight simulation of fire-fighting aircraft based on multi-factor coupling modeling of forest fire. **Aerospace**, MDPI, v. 11, n. 4, p. 267, 2024. Disponível em: <https://doi.org/10.3390/aerospace11040267>.

LOUZADA, V. H. P.; FERREIRA, W. Incêndios florestais em autômatos celulares, simples e grandes queimadas. **Biomatemática-Unicamp**, v. 18, p. 15–26, 2008. Disponível em: <https://www.ime.unicamp.br/~biomat/bio18_art2.pdf>.

MACCARTHY, J. et al. **The Latest Data Confirms: Forest Fires Are Getting Worse — wri.org**. 2025. <https://www.wri.org/insights/global-trends-forest-fires>. [Accessed 08-01-2025].

MAPBIOMAS Brasil. **Área queimada no Brasil cresce 79supera os 30 milhões de hectare**. 2025. <https://brasil.mapbiomas.org/2025/01/22/area-queimada-no-brasil-cresce-79-em-2024-e-supera-os-30-milhoes-de-hectares/>. [Accessed 03-02-2025].

MARTÍNEZ, M. J. F. et al. A graph cellular automata model to study the spreading of an infectious disease. In: SPRINGER. **Advances in Artificial Intelligence: 11th Mexican International Conference on Artificial Intelligence, MICAI 2012, San Luis Potosí, Mexico, October 27–November 4, 2012. Revised Selected Papers, Part I 11**. 2013. p. 458–468. Disponível em: <https://doi.org/10.1007/978-3-642-37807-2_39>.

MENG, Q. et al. Forest fire spread simulation and fire extinguishing visualization research. **Forests**, MDPI, v. 14, n. 7, p. 1371, 2023. Disponível em: <https://doi.org/10.3390/f14071371>.

MONTEIRO, L. et al. The effect of spatial scale on predicting time series: a study on epidemiological system identification. **Mathematical Problems in Engineering**, Hindawi, v. 2009, 2009. Disponível em: <https://doi.org/10.1155/2009/137854>.

MONTEIRO, L. H. A.; GANDINI, D.; SCHIMIT, P. H. The influence of immune individuals in disease spread evaluated by cellular automaton and genetic algorithm. **Computer methods and programs in biomedicine**, Elsevier, v. 196, p. 105707, 2020. Disponível em: <https://doi.org/10.1016/j.cmpb.2020.105707>.

MUGNELA, B. S. **GENPOLIS: Prototipagem e aplicação de um simulador de trânsito voltado para otimização de sinalização semafórica por meio de algoritmos genéticos.** Tese (Doutorado) — Universidade de São Paulo, 2021.

MURILO, L. V.; OLIVEIRA, G. M. B.; MARTINS, L. G. A. Evolutionary adjustment of a cellular automata-based model for wildfire spreading. In: PAES, A.; VERRI, F. A. N. (Ed.). **34th BRAZILIAN CONFERENCE ON INTELLIGENT SYSTEMS (BRACIS)**. Cham: Springer Nature Switzerland, 2025. p. 260–275. ISBN 978-3-031-79035-5. Disponível em: <https://doi.org/10.1007/978-3-031-79035-5_18>.

OLIVEIRA, G. Autômatos celulares: aspectos dinâmicos e computacionais. **III Jornada de Mini-cursos em Inteligência Artificial (MCIA)**, v. 8, p. 297–345, 2003.

OSBORNE, J. Notes on the use of data transformations. **Practical assessment, research, and evaluation**, University of Massachusetts Amherst Libraries, v. 8, n. 1, 2002. Disponível em: <https://doi.org/10.7275/4vng-5608>.

PACKARD, N. H.; WOLFRAM, S. Two-dimensional cellular automata. **Journal of Statistical physics**, Springer, v. 38, n. 5, p. 901–946, 1985. Disponível em: <https://doi.org/10.1007/BF01010423>.

PAREDIS, J. et al. Coevolving cellular automata: Be aware of the red queen!. In: CITESEER. **ICGA**. 1997. p. 393–400. Disponível em: <https://www.researchgate.

net/publication/2658208_Coevolving_Cellular_Automata_Be_Aware_of_the_Red_Queen>.

PAVLOVA, A.; RUBTSOV, S.; TELYATNIKOV, I. Using cellular automata in modelling of the fire front propagation through rough terrain. In: IOP PUBLISHING. **IOP Conference Series: Earth and Environmental Science**. 2020. v. 579, n. 1, p. 012104. Disponível em: <https://doi.org/10.1088/1755-1315/579/1/012104>.

PEREIRA, F. H.; SCHIMIT, P. H.; BEZERRA, F. E. A deep learning based surrogate model for the parameter identification problem in probabilistic cellular automaton epidemic models. **Computer Methods and Programs in Biomedicine**, Elsevier, v. 205, p. 106078, 2021. Disponível em: <https://doi.org/10.1016/j.cmpb.2021.106078>.

PERELLO, N. et al. Cellular automata-based simulators for the design of prescribed fire plans: the case study of liguria, italy. **Fire Ecology**, Springer, v. 20, n. 1, p. 7, 2024. Disponível em: <https://doi.org/10.1186/s42408-023-00239-7>.

PILON, N. A. et al. Speedy blooming in cerrado after fire is not uncommon: new records of cyperaceae species flowering 24 h after burning. **Austral Ecology**, Wiley Online Library, v. 48, n. 5, p. 1042–1045, 2023. Disponível em: <https://doi.org/10.1111/aec.13326>.

RASHID, M. T. et al. Compdrone: towards integrated computational model and social drone based wildfire monitoring. In: IEEE. **2020 16th international conference on distributed computing in sensor systems (DCOSS)**. 2020. p. 43–50. Disponível em: <https://doi.org/10.1109/DCOSS49796.2020.00020>.

SANTOS, A. C. D. et al. Managing fires in a changing world: Fuel and weather determine fire behavior and safety in the neotropical savannas. **Journal of environmental management**, Elsevier, v. 289, p. 112508, 2021. Disponível em: <https://doi.org/10.1016/j.jenvman.2021.112508>.

SARKAR, P. A brief history of cellular automata. **Acm computing surveys (csur)**, ACM New York, NY, USA, v. 32, n. 1, p. 80–107, 2000. Disponível em: <https://doi.org/10.1145/349194.349202>.

SECK-TUOH-MORA, J. C. et al. A continuous-state cellular automata algorithm for global optimization. **Expert Systems with Applications**, Elsevier, v. 177, p. 114930, 2021. Disponível em: <https://doi.org/10.1016/j.eswa.2021.114930>.

SHEN, X.; GUO, Y.; LI, A. Cooperative coevolution with an improved resource allocation for large-scale multi-objective software project scheduling. **Applied Soft Computing**, Elsevier, v. 88, p. 106059, 2020. Disponível em: <https://doi.org/10.1016/j.asoc.2019.106059>.

SILVA, F. d. et al. Modelo de calibração de rede de distribuição de água considerando fugas: estudo de caso do setor real monte carlo. **XV Simpósio Brasileiro de Recursos Hídricos**, 2003. Disponível em: <http://dx.doi.org/10.21168/rbrh.v9n1.p37-54>.

SINGH, D.; SINGH, B. Investigating the impact of data normalization on classification performance. **Applied Soft Computing**, Elsevier, v. 97, p. 105524, 2020. Disponível em: <https://doi.org/10.1016/j.asoc.2019.105524>.

SUN, L. et al. Adaptive forest fire spread simulation algorithm based on cellular automata. **Forests**, v. 12, p. 1431, 2021. Disponível em: <https://doi.org/10.3390/f12111431>.

TINOCO, C. R. et al. Wildfire simulation model based on cellular automata and stochastic rules. In: SPRINGER. **International Conference on Cellular Automata for Research and Industry**. 2022. p. 246–256. Disponível em: <https://doi.org/10.1007/978-3-031-14926-9_22>.

TINOCO, C. R.; VIZZARI, G.; OLIVEIRA, G. M. Parameter adjustment of a bio-inspired coordination model for swarm robotics using evolutionary optimisation. In: SPRINGER. **Cellular Automata: 14th International Conference on Cellular Automata for Research and Industry, ACRI 2020, Lodz, Poland, December 2–4, 2020, Proceedings 14**. 2021. p. 146–155. Disponível em: <https://doi.org/10.1007/978-3-030-69480-7_15>.

TIRUMALA, S. S. Evolving deep neural networks using coevolutionary algorithms with multi-population strategy. **Neural Computing and Applications**, Springer, v. 32, n. 16, p. 13051–13064, 2020. Disponível em: <https://doi.org/10.1007/s00521-020-04749-2>.

VENTOURAS, E. M. et al. Amplitude normalization applied to an artificial neural network-based automatic sleep spindle detection system. In: IEEE. **2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society**. 2014. p. 3240–3243. Disponível em: <https://doi.org/10.1109/EMBC.2014.6944313>.

WOLFRAM, S. **A new kind of science. 2002**. Champaign, IL: Wolfram Media, Inc., 2002.

XU, Y. et al. Modeling forest fire spread using machine learning-based cellular automata in a gis environment. **Forests**, MDPI, v. 13, n. 12, p. 1974, 2022. Disponível em: <https://doi.org/10.3390/f13121974>.

XUEHUA, W. et al. A cellular automata model for forest fire spreading simulation. In: **IEEE Symp. Series on Computational Intelligence**. IEEE, 2016. p. 1–6. Disponível em: <https://doi.org/10.1109/SSCI.2016.7849971>.

YASSEMI, S.; DRAGIĆEVIĆ, S.; SCHMIDT, M. Design and implementation of an integrated gis-based cellular automata model to characterize forest fire behaviour. **Ecological Modelling**, v. 210, n. 1-2, p. 71–84, 2008. Disponível em: <https://doi.org/10.1016/j.ecolmodel.2007.07.020>.

YELIZAVETA, M.; TAT-SENG, C.; IRINA, A. Analysis and retrieval of paintings using artistic color concepts. In: IEEE. **2005 IEEE International Conference on Multimedia and Expo**. 2005. p. 1246–1249. Disponível em: <https://doi.org/10.1109/ICME.2005.1521654>.

YONGZHONG, Z. et al. Simulating wildfire spreading processes in a spatially heterogeneous landscapes using an improved cellular automaton model. In: **IEEE Int. Geoscience and Remote Sensing Symp.** Amsterdam, Netherlands: IEEE, 2004. v. 5, p. 3371–3374. Disponível em: <https://doi.org/10.1109/ICME.2005.1521654>.

ZHANG, S. et al. Study on forest fire spread model of multi-dimensional cellular automata based on rothermel speed formula. **CERNE**, SciELO Brasil, v. 27, p. e–102932, 2022. Disponível em: <https://doi.org/10.1590/01047760202127012932>.

ZIENKIEWICZ, O. C.; TAYLOR, R. L. **The finite element method for solid and structural mechanics**. Oxford: Elsevier, 2005.