
**Estratégias Evolutivas para o Ajuste de
Parâmetros de um Modelo Epidemiológico
Baseado em Autômatos Celulares
Probabilísticos**

Larissa Maiara Fraga



UNIVERSIDADE FEDERAL DE UBERLÂNDIA
FACULDADE DE COMPUTAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

Uberlândia
2021

Larissa Maiara Fraga

**Estratégias Evolutivas para o Ajuste de
Parâmetros de um Modelo Epidemiológico
Baseado em Autômatos Celulares
Probabilísticos**

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

2021

Dados Internacionais de Catalogação na Publicação (CIP)
Sistema de Bibliotecas da UFU, MG, Brasil.

F811e
2021 Fraga, Larissa Maiara, 1985-
Estratégias evolutivas para o ajuste de parâmetros de um modelo epidemiológico baseado em autômatos celulares probabilísticos [recurso eletrônico] / Larissa Maiara Fraga. - 2021.

Orientador: Luiz Gustavo Almeida Martins.
Coorientadora: Gina Maira Barbosa deOliveira.
Dissertação (Mestrado) - Universidade Federal de Uberlândia,
Programa de Pós-Graduação em Ciência da Computação.
Modo de acesso: Internet.
Disponível em: <http://doi.org/10.14393/ufu.di.2022.5042>
Inclui bibliografia.
Inclui ilustrações.

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

CDU: 681.3

André Carlos Francisco
Bibliotecário - CRB-6/3408


UNIVERSIDADE FEDERAL DE UBERLÂNDIA

Coordenação do Programa de Pós-Graduação em Ciência da Computação
 Av. João Naves de Ávila, nº 2121, Bloco 1A, Sala 243 - Bairro Santa Mônica, Uberlândia-MG, CEP 38400-902
 Telefone: (34) 3239-4470 - www.ppgco.facom.ufu.br - cpqfacom@ufu.br


ATA DE DEFESA - PÓS-GRADUAÇÃO

| | | | | | |
|------------------------------------|---|-----------------|----------|-----------------------|----------|
| Programa de Pós-Graduação em: | Ciência da Computação | | | | |
| Defesa de: | Mestrado Acadêmico, 15/2021, PPGCO | | | | |
| Data: | 05 de julho de 2021 | Hora de início: | 14h02min | Hora de encerramento: | 17h26min |
| Matrícula do Discente: | 11822CCP006 | | | | |
| Nome do Discente | Larissa Maiara Fraga | | | | |
| Título do Trabalho: | Estratégias Evolutivas para o Ajuste de Parâmetros de um Modelo Epidemiológico Baseado em Autômatos Celulares Probabilísticos | | | | |
| Área de concentração: | Ciência da Computação | | | | |
| Linha de pesquisa: | Inteligência Artificial | | | | |
| Projeto de Pesquisa de vinculação: | - | | | | |

Reuniu-se, por videoconferência, a Banca Examinadora, designada pelo Colegiado do Programa de Pós-graduação em Ciência da Computação, assim composta: Professores Doutores: Murillo Guimarães Carneiro - FACOM/UFU; Luiz Henrique Alves Monteiro - USP; Gina Maira Barbosa de Oliveira - FACOM/UFU (coorientadora) e Luiz Gustavo Almeida Martins - FACOM/UFU, orientador da candidata.

Os examinadores participaram desde as seguintes localidades: Luiz Henrique Alves Monteiro - São Paulo/SP; Murillo Guimarães Carneiro, Gina Maira Barbosa de Oliveira e Luiz Gustavo Almeida Martins - Uberlândia/MG. A discente participou da cidade de Araxá/MG.

Iniciando os trabalhos o presidente da mesa, Prof. Dr. Luiz Gustavo Almeida Martins, apresentou a Comissão Examinadora e a candidata, agradeceu a presença do público, e concedeu à 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 a candidata. Ultimada a arguição, que se desenvolveu dentro dos termos regimentais, a Banca, em sessão secreta, atribuiu o resultado final, considerando a candidata:

Aprovada.

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 06/07/2021, às 10:34, 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 **Murillo Guimarães Carneiro, Professor(a) do Magistério Superior**, em 06/07/2021, às 13:05, 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 **Luiz Henrique Alves Monteiro, Usuário Externo**, em 06/07/2021, às 18:47, 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 **Gina Maira Barbosa de Oliveira, Professor(a) do Magistério Superior**, em 07/07/2021, às 17:40, 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 **2880194** e o código CRC **AE837CC7**.

*Dedico este trabalho a minha filha Alice,
meu combustível infinito de motivação.*

Agradecimentos

Agradeço a Deus por me guiar e proteger, principalmente nas viagens diárias entre Araxá e Uberlândia.

À minha família pelo o apoio e incentivo durante essa caminhada. Sem vocês seria impossível.

À minha filha Alice que me motiva a querer ser uma pessoa melhor para o mundo.

À minha mãe Beth por cuidar da minha filha enquanto eu estava fora, me dando tranquilidade para correr atrás do meus sonhos.

Ao meu pai Danilo pelo suporte, incentivo e participação.

Ao Lucas por nunca soltar a minha mão, por todo amor, ajuda e incentivo. Por muitas vezes colocar as minhas prioridades como dele e ser alicerce quando necessário.

À minha avó Anita, cuja partida foi responsável pela minha primeira pausa. Sei que lá do céu, ela está orgulhosa com o meu recomeço.

Aos meus orientadores Luiz Gustavo e Gina, por todo interesse, disposição e generosidade em compartilhar seus conhecimentos e experiências. Por toda paciência e compreensão diante de todas as curvas dessa caminhada, a contribuição de vocês na minha vida foi muito além da acadêmica.

À Coordenação do Programa de Pós-Graduação em Computação pela ajuda oferecida sempre que precisei.

À Fundação de Amparo à Pesquisa de Minas Gerais (FAPEMIG) pelo apoio financeiro durante parte da pesquisa, por meio da bolsa de mestrado.

E, por fim, à Universidade Federal de Uberlândia (UFU), em especial à Faculdade de Computação (FACOM), pela receptividade, pela infraestrutura, e pelo ensino gratuito e de qualidade.

*“Se você quiser alguém em quem confiar
Confie em si mesmo
Quem acredita sempre alcança.”
(Renato Russo)*

Resumo

Modelos confiáveis permitem a simulação de processos críticos que podem servir de base para o planejamento e definição de políticas públicas. A modelagem de sistemas dinâmicos é um importante instrumento de pesquisa que pode prever e avaliar o impacto de decisões tomadas por organizações e governos. Autômatos celulares têm sido usados como alternativa para esses tipos de modelagem, pois são sistemas dinâmicos, discretos e capazes de descrever fenômenos complexos a partir de um conjunto de regras simples e iterações locais. Uma vez definida a representação básica do sistema, uma das principais dificuldades na modelagem é o ajuste dos diversos parâmetros que o compõem. Considerando que os algoritmos evolutivos têm se mostrado uma poderosa técnica de busca adaptativa, neste trabalho investigamos duas estratégias baseadas em algoritmos genéticos (AGs) para ajustar parâmetros de um modelo, visando aproximar o resultado das simulações ao comportamento dos dados a serem modelados. A primeira estratégia consiste na aplicação de um algoritmo genético típico para resolver o problema, enquanto a segunda adota múltiplos estágios de evolução, no qual cada estágio é responsável por ajustar um subconjunto dos parâmetros. Ambas as estratégias adotam como estudo de caso um modelo que descreve a evolução de uma população de insetos vetores responsáveis pela doença de Chagas (SLIMI et al., 2009). Esse modelo é baseado em autômatos celulares probabilísticos e foi originalmente proposto para reproduzir dados reais coletados em uma aldeia no México. Aqui, ele é usado para gerar o conjunto de dados usada como referência pelo algoritmo genético. Resultados experimentais mostram que os parâmetros definidos em ambas as estratégias evolutivas reproduzem um comportamento semelhante ao modelo de referência tanto na quantidade de insetos quanto na sua distribuição espacial.

Palavras-chave: Algoritmos Genéticos. Autômatos Celulares. Evolução multi-estágio. Computação evolutiva. Modelagem de sistemas dinâmicos. Ajuste de parâmetros. Doença de Chagas.

Abstract

Reliable models allow the simulation of critical processes that can serve as a basis for planning and defining public policies. Dynamic systems modeling is an important research tool that can predict and assess the impact of decisions taken by organizations and governments. Cellular automata have been used as an alternative for these types of modeling, as they are dynamic, discrete systems capable of describing complex phenomena from a set of simple rules and local iterations. Once the basic representation of the system is defined, one of the main difficulties in modeling is the adjustment of the various parameters that compose it. Since the evolutionary algorithms have been shown to be a powerful adaptive search technique, in this work we investigate two strategies based on genetic algorithms (GAs) to adjust parameters of a model, aiming to approximate the simulation results to the behavior found in the data to be modeled. The first strategy consists of applying a typical genetic algorithm to solve the problem, while the second adopts multiple stages of evolution, where each stage is responsible for adjusting a subset of the parameters. Both strategies adopt as a case study a model that describes the evolution of a population of insect vectors responsible for Chagas disease (SLIMI et al., 2009). This model is based on probabilistic cellular automata and was originally proposed to reproduce real data collected in a village in Mexico. Here, it is used to generate the dataset used as reference by the genetic algorithm. Experimental results show that the parameters defined in both evolutionary strategies reproduce a behavior similar to the reference model with respect of the number of insects and their spatial distribution.

Keywords: Genetic Algorithms. Cellular Automata. Multi-stage evaluation. Evolutionary computing. Dynamic systems modeling. Parameter adjustment. Chagas disease.

Lista de ilustrações

| | |
|--|----|
| Figura 1 – Espaço celular: (a) unidimensional (b) bidimensional e (c) tridimensional. | 36 |
| Figura 2 – Possíveis formatos das células de um Autômato Celular: (a) quadrado (b) triângulo e (c) hexágono. | 37 |
| Figura 3 – Tipos de borda mais comuns em um Autômato Celular unidimensional: (a) periódica, circular ou cíclica, (b) reflexiva, (c) fixa. | 37 |
| Figura 4 – Vizinhanças mais comuns em um Autômato Celular: (a) von Neumann e (b) Moore. | 38 |
| Figura 5 – Vizinhança de um Autômato Celular unidimensional de acordo com o raio (\mathcal{R}): (a) $\mathcal{R} = 1$ e (b) $\mathcal{R} = 2$ | 39 |
| Figura 6 – Vizinhança de um Autômato Celular bidimensional de acordo com o raio (\mathcal{R}): (a) von Neumann com $\mathcal{R} = 1$, (b) von Neumann com $\mathcal{R} = 2$, (c) Moore com $\mathcal{R} = 1$ e (d) Moore com $\mathcal{R} = 2$ | 39 |
| Figura 7 – Autômato Celular unidimensional “Regra 30”: (a) regra de transição (b) reticulado inicial e (c) evolução espaço-temporal por 7 passos de tempo. | 41 |
| Figura 8 – Exemplo de reticulado do AC “Jogo da Vida”, com uma de suas vizinhanças em destaque. | 42 |
| Figura 9 – Histórico de evolução do AC “Jogo da Vida” nos instantes de tempo (t): (a) $t = 1$, (b) $t = 2$, (c) $t = 3$, (d) $t = 4$ e (e) $t = 5$ | 43 |
| Figura 10 – Tipos mais comuns de representação de indivíduos de um AG: (a) binária (b) inteira e (c) real. | 46 |
| Figura 11 – Representação de uma população. | 47 |
| Figura 12 – Fluxograma de um Algoritmo Genético. | 49 |
| Figura 13 – Aptidão e probabilidade dos indivíduos da população. | 50 |
| Figura 14 – Seleção - Roleta. | 50 |
| Figura 15 – Seleção por torneio. | 50 |
| Figura 16 – Cruzamento com 1 ponto de corte. (a) pais com um ponto de corte aleatório selecionado e (b) filhos gerados pela combinação de fragmentos. | 52 |

| | |
|---|----|
| Figura 17 – Cruzamento com 2 pontos de corte. (a) pais com 2 pontos de corte aleatórios selecionados e (b) filhos gerados pela combinação de fragmentos. | 52 |
| Figura 18 – Cruzamento uniforme. (a) pais e máscara de cruzamento e (b) filhos gerados a partir da máscara. | 53 |
| Figura 19 – Cruzamento aritmético. | 53 |
| Figura 20 – Mutação binária ou <i>bit flip</i> . | 54 |
| Figura 21 – Mutação por permutação. | 54 |
| Figura 22 – Exemplo de mutação para representações numéricas. | 54 |
| Figura 23 – Tipos de mutação: (a) indivíduo e (b) gene. | 55 |
| Figura 24 – Evolução do espalhamento dos insetos adultos simulada pelo AC a partir dos valores de referência do artigo original (retirada de (SLIMI et al., 2009)). | 65 |
| Figura 25 – Dinâmica populacional dos insetos ao longo de 10 anos simulados pelo autômato celular, de acordo com o modelo adotado. | 67 |
| Figura 26 – Distribuição espacial dos adultos nos 5º, 90º, 180º e 270º dias, de acordo com a simulação do modelo adotado. | 68 |
| Figura 27 – Distribuição espacial das larvas nos 5º, 90º, 180º e 270º dias, de acordo com a simulação do modelo de referência. | 68 |
| Figura 28 – Representação de um indivíduo da população. | 71 |
| Figura 29 – Distribuição espacial adotada na avaliação regional. | 72 |
| Figura 30 – Ordem de execução dos estágios do MSES. | 74 |
| Figura 31 – Fluxograma da estratégia evolutiva multi-estágio. | 78 |
| Figura 32 – Dinâmica populacional de insetos ao longo de 10 anos simulados pelo modelo usando as configurações definidas pelo PGA. | 80 |
| Figura 33 – Distribuição espacial dos adultos nos 5º, 90º, 180º e 270º dias de simulação do modelo, usando as configurações definidas pelo PGA. | 81 |
| Figura 34 – Distribuição espacial das larvas nos 5º, 90º, 180º e 270º dias de simulação do modelo, usando as configurações definidas pelo PGA. | 83 |
| Figura 35 – Evolução do erro para o melhor, pior e indivíduo mediano da população, considerando a melhor execução com o PGA. | 85 |
| Figura 36 – Dinâmica populacional de insetos ao longo de 10 anos simulados pelo modelo usando as configurações definidas pelas estratégias MSES-SR. | 88 |
| Figura 37 – Distribuição espacial dos adultos nos 5º, 90º, 180º e 270º dias de simulação do modelo, usando as configurações definidas pelas estratégias MSES-SR. | 89 |
| Figura 38 – Distribuição espacial das larvas nos 5º, 90º, 180º e 270º dias de simulação do modelo, usando as configurações definidas pelas estratégias MSES-SR. | 90 |

| | |
|--|-----|
| Figura 39 – Evolução do erro para o melhor, pior e indivíduo mediano da população, considerando a melhor execução com as estratégias MSES-SR. | 92 |
| Figura 40 – Dinâmica populacional de insetos ao longo de 10 anos simulados pelo modelo, usando as configurações definidas pelas estratégias MSES-R20. | 93 |
| Figura 41 – Distribuição espacial dos adultos nos 5 ^o , 90 ^o , 180 ^o e 270 ^o dias de simulação do modelo, usando as configurações definidas pelas estratégias MSES-R20. | 94 |
| Figura 42 – Distribuição espacial das larvas nos 5 ^o , 90 ^o , 180 ^o e 270 ^o dias de simulação do modelo, usando as configurações definidas pelas estratégias MSES-R20. | 95 |
| Figura 43 – Evolução do erro para o melhor, pior e indivíduo mediano da população, considerando a melhor execução com as estratégias MSES-R20. | 97 |
| Figura 44 – Dinâmica populacional de insetos ao longo de 10 anos simulados pelo modelo, usando as configurações definidas pelas estratégias MSES-RBEST. | 98 |
| Figura 45 – Distribuição espacial dos adultos nos 5 ^o , 90 ^o , 180 ^o e 270 ^o dias de simulação do modelo, usando as configurações definidas pelas estratégias MSES-RBEST. | 99 |
| Figura 46 – Distribuição espacial das larvas nos 5 ^o , 90 ^o , 180 ^o e 270 ^o dias de simulação do modelo, usando as configurações definidas pelas estratégias MSES-RBEST. | 100 |
| Figura 47 – Evolução do erro para o melhor, pior e indivíduo mediano da população, considerando a melhor execução com as estratégias MSES-RBEST. | 102 |
| Figura 48 – Dinâmica populacional de insetos ao longo de 10 anos simulados pelo modelo, usando as configurações definidas pelas estratégias evolutivas com a avaliação regional. | 104 |
| Figura 49 – Distribuição espacial dos adultos nos 5 ^o , 90 ^o , 180 ^o e 270 ^o dias de simulação do modelo, usando as configurações definidas pelas estratégias evolutivas com a avaliação regional. | 105 |
| Figura 50 – Distribuição espacial das larvas nos 5 ^o , 90 ^o , 180 ^o e 270 ^o dias de simulação do modelo, usando as configurações definidas pelas estratégias evolutivas com a avaliação regional. | 106 |
| Figura 51 – Evolução do erro para o melhor, pior e indivíduo mediano da população, considerando a melhor execução das estratégias evolutivas com a avaliação regional. | 107 |
| Figura 52 – Dinâmica populacional de insetos ao longo de 10 anos simulados pelo modelo de referência, de acordo com o tamanho do reticulado e os raios adotadas na etapa de dispersão. | 110 |

| | |
|---|-----|
| Figura 53 – Distribuição espacial dos adultos nos 5º, 90º, 180º e 270º dias de simulação, de acordo com a configuração usada no modelo de referência. | 111 |
| Figura 54 – Dinâmica populacional de insetos ao longo de 10 anos simulados com o reticulado 60 × 60, usando as configurações definidas pelo PGA. | 112 |
| Figura 55 – Evolução do erro para o melhor, pior e indivíduo mediano da população, considerando a melhor execução do PGA para o reticulado 60 × 60. | 112 |
| Figura 56 – Dinâmica populacional de insetos ao longo de 10 anos simulados com reticulado 60 × 60, usando as configurações definidas pelo MSES. | 113 |
| Figura 57 – Evolução do erro para o melhor, pior e indivíduo mediano da população, considerando a melhor execução do MSES para o reticulado 60 × 60. | 113 |
| Figura 58 – Distribuição espacial dos adultos nos 5º, 90º, 180º e 270º dias de simulação do modelo com reticulado 60 × 60 e raios 3 e 1. | 115 |
| Figura 59 – Distribuição espacial dos adultos nos 5º, 90º, 180º e 270º dias de simulação do modelo com reticulado 60 × 60 e raios 6 e 2. | 116 |
| Figura 60 – Distribuição espacial das larvas nos 5º, 90º, 180º e 270º dias de simulação do modelo com reticulado 60 × 60 e raios 3 e 1. | 117 |
| Figura 61 – Distribuição espacial das larvas nos 5º, 90º, 180º e 270º dias de simulação do modelo com reticulado 60 × 60 e raios 6 e 2. | 118 |

Lista de tabelas

| | | |
|-----------|---|-----|
| Tabela 1 | – Regra de transição do AC “Regra 30”. | 40 |
| Tabela 2 | – Valores dos coeficientes e raios nos períodos de infestação e não infestação. | 64 |
| Tabela 3 | – Melhor indivíduo e o erro robusto resultantes das execuções com o PGA. | 80 |
| Tabela 4 | – Erro médio e desvio padrão resultantes das execuções com o PGA. | 84 |
| Tabela 5 | – Melhor indivíduo e o erro robusto resultantes das execuções com as estratégias evolutivas: (1) PGA-FULL, (2) MSES-SR-50-25-25, (3) MSES-SR-50-10-10-10-10-10 e (4) MSES-SR-50-30-20. | 87 |
| Tabela 6 | – Erro médio e desvio padrão das execuções com as estratégias MSES-SR. | 91 |
| Tabela 7 | – Melhor indivíduo e erro robusto resultantes das execuções com as estratégias evolutivas: (1) PGA-FULL, (2) MSES-SR-50-30-20, (3) MSES-R20 - 50-25-25, (4) MSES-R20 - 50-10-10-10-10-10 e (5) MSES-R20 - 50-30-20. | 91 |
| Tabela 8 | – Erro médio e desvio padrão das execuções com as estratégias MSES-R20. | 96 |
| Tabela 9 | – Melhor indivíduo e o erro robusto resultantes das execuções com as abordagens evolutivas: (1) PGA-FULL, (2) MSES-SR-50-30-20, (3) MSES-R20 - 50-25-25, (4) MSES-RBEST-50-25-25, (5) MSES-RBEST-50-10-10-10-10-10 e (6) MSES-RBEST-50-30-20. | 98 |
| Tabela 10 | – Erro médio e desvio padrão das execuções com as estratégias MSES-RBEST. | 101 |
| Tabela 11 | – Melhor indivíduo e erro robusto resultantes das execuções com as abordagens evolutivas usando a avaliação regional. | 103 |
| Tabela 12 | – Erro médio e desvio padrão das execuções com a avaliação regional. | 108 |
| Tabela 13 | – Melhor indivíduo e erro robusto resultantes das execuções com as abordagens evolutivas para os modelos com reticulado 60×60 | 114 |
| Tabela 14 | – Erro médio e desvio padrão das execuções de acordo com a estratégia evolutiva e a configuração do modelo (tamanho do reticulado e raios de dispersão). | 115 |

Lista de Siglas

AC Autômatos Celulares

AG Algoritmos Genéticos

CBR Raciocínio Baseado em Casos (*Case Based Reasoning*)

CE Computação Evolutiva

DCT Tarefa de Classificação de Densidade (*Density Classification Task*)

GSCA Autômato Celular Estocástico Global (*Global Stochastic Cellular Automata*)

RL Aprendizagem por Reforço (*Reinforcement Learning*)

MSES Estratégia Evolutiva Multi-Estágio (*Multi-Stage Evolutionary Strategy*)

MSES-SR Multi-Estágio Sem Repovoamento

MSES-R20 Multi-Estágio com Repovoamento baseado nos 20 melhores indivíduos

MSES-RBEST Multi-Estágio baseado no melhor indivíduo

PGA Algoritmo Genético Simples (*Plain Genetic Algorithm*)

RNA Redes Neurais Artificiais

RBS Sistema Baseado em Regras (*Rule Based System*)

SMA Sistemas Multi-Agentes

SIR Suscetíveis Infectados Recuperados

SI Inteligência de Enxame (*Swarm Intelligence*)

SR Sensoriamento Remoto

Sumário

| | | |
|----------|---|-----------|
| 1 | INTRODUÇÃO | 27 |
| 1.1 | Motivação | 29 |
| 1.2 | Justificativa | 30 |
| 1.3 | Objetivos | 32 |
| 1.4 | Hipótese | 32 |
| 1.5 | Contribuições | 33 |
| 1.6 | Organização da Dissertação | 33 |
| 2 | AUTÔMATOS CELULARES | 35 |
| 2.1 | Histórico e Definição | 35 |
| 2.1.1 | Espaço e Dimensão de um Autômato Celular | 36 |
| 2.1.2 | Estado e Vizinhança | 37 |
| 2.1.3 | Regras de Transição | 38 |
| 2.2 | Autômato Celular Probabilístico | 40 |
| 2.3 | Exemplos | 40 |
| 2.3.1 | Autômato celular unidimensional - Regra 30 | 40 |
| 2.3.2 | Autômato celular bidimensional - Jogo da vida | 41 |
| 3 | ALGORITMOS GENÉTICOS | 45 |
| 3.1 | Histórico | 45 |
| 3.2 | Representação do Indivíduo | 46 |
| 3.2.1 | População | 46 |
| 3.3 | Função Objetivo - <i>Fitness</i> | 47 |
| 3.4 | Fluxo de Execução | 48 |
| 3.5 | Seleção | 48 |
| 3.5.1 | Roleta | 49 |
| 3.5.2 | Torneio | 50 |
| 3.5.3 | Elitismo | 51 |

| | | |
|------------|---|-----------|
| 3.6 | Operadores Genéticos | 51 |
| 3.6.1 | Reprodução - <i>Crossover</i> | 51 |
| 3.6.2 | Mutação | 53 |
| 3.7 | Reinserção - Sobrevivência de Indivíduos | 55 |
| 3.8 | Critério de Parada | 55 |
| | | |
| 4 | TRABALHOS CORRELATOS | 57 |
| 4.1 | Autômatos Celulares e Modelos Epidemiológicos | 58 |
| 4.2 | Algoritmos Genéticos na configuração de parâmetros de Autômatos Celulares | 59 |
| | | |
| 5 | MODELO EPIDEMIOLÓGICO | 61 |
| 5.1 | Estudo de Caso | 61 |
| 5.1.1 | Demografia | 62 |
| 5.1.2 | Dispersão | 63 |
| 5.1.3 | Análise | 64 |
| 5.2 | Reprodução do Modelo | 66 |
| | | |
| 6 | ESTRATÉGIAS EVOLUTIVAS PROPOSTAS | 69 |
| 6.1 | Algoritmo Genético Simples | 70 |
| 6.1.1 | Representação do Indivíduo | 70 |
| 6.1.2 | Função Objetivo - <i>Fitness</i> | 71 |
| 6.1.3 | Seleção | 73 |
| 6.1.4 | Operadores Genéticos | 73 |
| 6.1.5 | Reinserção | 73 |
| 6.1.6 | Avaliação Robusta | 73 |
| 6.1.7 | Divisão do Modelo | 73 |
| 6.2 | MSES - Estratégia Evolutiva Multi-Estágio | 74 |
| 6.2.1 | Estratégia sem Repovoamento (MSES-SR) | 75 |
| 6.2.2 | Repovoamento Baseado nos 20 Melhores (MSES-R20) | 75 |
| 6.2.3 | Repovoamento Baseado no Melhor Indivíduo (MSES-RBEST) | 76 |
| 6.2.4 | Visão Geral do Algoritmo Multi-Estágio | 77 |
| | | |
| 7 | EXPERIMENTOS E ANÁLISE DOS RESULTADOS | 79 |
| 7.1 | Método para a Avaliação | 79 |
| 7.2 | Algoritmo Genético Simples | 79 |
| 7.3 | MSES - Estratégias Evolutivas Multi-Estágio | 86 |
| 7.3.1 | Análise da Estratégia MSES-SR | 87 |
| 7.3.2 | Análise da Estratégia MSES-R20 | 91 |
| 7.3.3 | Análise da Estratégia MSES-RBEST | 96 |
| 7.4 | Avaliação Regional | 101 |

| | | |
|------------|---|------------|
| 7.5 | Análise de Robustez | 108 |
| 7.5.1 | Algoritmo Genético Simples | 109 |
| 7.5.2 | Estratégia Multi Estágio com Repovoamento Baseado no Melhor Indiví- duo (MSES-RBEST) | 111 |
| 7.5.3 | Análise e Distribuição Espacial | 114 |
| 8 | CONCLUSÃO | 119 |
| 8.1 | Trabalhos Futuros | 120 |
| 8.2 | Produção Bibliográfica Originada da Dissertação | 121 |
| | REFERÊNCIAS | 123 |

Introdução

Mudanças cada vez mais intensas vêm alterando o funcionamento do nosso ecossistema. Essas mudanças podem ser de causas naturais, provocadas pela atividade humana ou uma combinação desses dois. Entretanto, estudos apontam que as alterações mais drásticas nos biomas são causadas principalmente pelas ações humanas (SALAZAR; NOBRE; OYAMA, 2007). Essas mudanças ambientais têm consequência direta na integridade do meio ambiente e na qualidade de vida das pessoas. Por exemplo, o aumento da temperatura global pode implicar em explosões populacionais de vetores de doenças tropicais e resultar em epidemias globais. Mudanças na vegetação têm aumentado as áreas de deserto e, cada vez mais, espécies de animais são ameaçadas de extinção.

As pessoas responsáveis por projetar e estabelecer políticas públicas precisam de ferramentas de modelagem que sejam confiáveis e capazes de simular a dinâmica e o impacto dessas mudanças ambientais para nortear seu planejamento e tomada de decisão (LAMBIN, 1994; KAIMOWITZ; ANGELSEN, 1998; Turner II et al., 1995). A simulação de processos naturais ou das interações humano-ambiente são instrumentos de pesquisa de impactos e previsão (BATISTELLA; MORAN, 2005) que têm potencial como meio de detectar, identificar, mapear e monitorar alterações nos ecossistemas, independentemente dos seus agentes causais (COPPIN et al., 2004). Essa simulação pode ser feita para vários tipos de sistemas, como: ecológicos (CHEN; MYNETT, 2006; TAN, 2007; GARCÍA-DURO et al., 2018), climáticos (CHANG; LO; YU, 2005), públicos de saúde (SLIMI et al., 2009; BONE; DRAGICEVIC; ROBERTS, 2006), hidrológicos (D'AMBROSIO et al., 2001) e de uso do solo (D'Ambrosio et al., 2007; D'Ambrosio; SPATARO, 2007). Modelos computacionais que reproduzem de uma forma satisfatória o fenômeno natural sob estudo contribuem para o ganho do conhecimento científico no que diz respeito ao seu funcionamento e este conhecimento pode servir de alicerce para o planejamento e definição de políticas públicas.

Uma das formas de se modelar um sistema é por meio de equações diferenciais. Essas equações são ferramentas matemáticas que descrevem a evolução dos modelos em relação ao tempo. Um conjunto de equações pode descrever o comportamento de um sistema,

mesmo que de forma simplificada. O principal desafio dessa modelagem é deduzir as equações que descrevem as variáveis do sistema e suas interações, já que cada uma delas possui características próprias e a leitura de seus comportamentos depende do olhar criterioso de um especialista da área em questão. Outro desafio encontrado é o custo computacional associado às simulações de sistemas complexos, que costuma ser elevado.

Autômatos celulares (ACs) têm sido usados como alternativa de modelagem e simulação de sistemas em que as equações diferenciais envolvem implementações complexas e computacionalmente caras (TOFFOLI, 1984). Eles são sistemas dinâmicos e discretos, que são capazes de descrever fenômenos complexos a partir de um conjunto de regras simples (TOFFOLI; MARGOLUS, 1987). Um autômato celular é formado por um conjunto de células distribuídas espacialmente que podem ser unidimensionais, bidimensionais ou tridimensionais, possuem características de estado e interagem entre si sendo capaz de alterar o estado das células vizinhas. Essa evolução de estados acontece a partir da definição de regras que são aplicadas em intervalos de tempo discretos, ou seja, a cada iteração de tempo. As regras são aplicadas em todo o reticulado e o estado de cada uma das células é afetado de acordo com suas características pessoais e as características das células vizinhas de maneira síncrona. Dessa forma, a representação de problemas complexos evoluem a partir da interação de regras simples, o que facilita a construção do modelo se comparado com a construção de modelos utilizando equações diferenciais (TOFFOLI, 1984). Uma aplicação importante desses conceitos acontece no ramo da epidemiologia. Autômatos Celulares possuem uma estrutura adequada para estudar fenômenos de espalhamento espacial.

Uma das principais doenças infecciosas que preocupam a saúde da América do Sul é a doença de Chagas, causada pelo protozoário *Trypanosoma cruzi*, transmitida principalmente através de insetos vetores, como os barbeiros (COURA; DIAS, 2009). Não há vacinas para essa doença e o método mais eficaz de prevenção é o controle de vetores. Um modelo baseado em ACs para a doença de Chagas é proposto em (SLIMI et al., 2009). Nesse trabalho os autores buscam desenvolver uma descrição espaço-temporal da dinâmica da população vetorial em uma escala de aldeia. O principal objetivo desse modelo é desvendar como a demografia e dispersão do vetores dessa doença interagem para produzir variações anuais na abundância de insetos na escala proposta. O modelo foi investigado de acordo com informações coletadas sobre o fenômeno real na península de Yucatan, no México. Nessa área os vetores são muito pouco ou não domesticados, ou seja, a sua ocorrência dentro das casas (área urbana) depende de uma forte migração sazonal do habitat peri doméstico e/ou florestal (SLIMI et al., 2009). Os dados coletados foram utilizados na validação do modelo e estimativa dos parâmetros.

A configuração de um modelo é uma tarefa difícil, necessita da expertise de um especialista além de demandar um tempo considerável. Portanto é interessante um método automático que configure seus múltiplos parâmetros de forma assertiva, e é na busca dessa

contribuição que focaremos esse trabalho.

AGs são uma classe particular de algoritmos evolutivos que utilizam do processo de evolução natural como um paradigma de solução de problemas, como os de otimização. Eles são inspirados no princípio Darwiniano da evolução das espécies e na reprodução genética (GOLDBERG, 1989). De acordo com esse princípio, a seleção de indivíduos é feita de modo a privilegiar aqueles mais aptos à sobrevivência, que, conseqüentemente, têm a maior probabilidade de reprodução. O indivíduo que se reproduz, transmite seu código genético a seus descendentes e com isso suas características são mantidas na população. A manipulação dessas características define o perfil da população.

Os AGs são probabilísticos e constituem uma técnica de busca heurística, altamente adaptativa. Os princípios biológicos como hereditariedade, mutação, seleção natural e recombinação (*crossover*) são codificados em algoritmos computacionais que buscam solucionar da melhor forma um determinado problema evoluindo populações direcionadas por uma função objetivo. A vantagem mais notória da computação evolutiva está na generalização do método para a resolução de problemas gerais (meta heurísticas), possibilitando a sua solução a partir da definição da forma de representação e avaliação de uma possível solução para o problema e dos seus operadores genéticos usados no processo de busca, sem a necessidade de especificar os passos que levam ao resultado, os quais são específicos de cada problema. Isso proporciona ao algoritmo robustez e flexibilidade o que faz com que seus procedimentos genéricos e adaptáveis sejam aplicados na solução de problemas complexos.

Nesse trabalho será desenvolvida uma abordagem baseada em algoritmos genéticos para a configuração automática de modelos baseados em autômatos celulares. Para isso, será reproduzido um modelo que trata da dispersão do vetor barbeiro, descrito em (SLIMI et al., 2009), como estudo de caso. O ajuste dos parâmetros desse modelo será feito por meio de uma abordagem utilizando algoritmos genéticos. Esse ajuste consiste na busca pelo conjunto de valores que melhor configura uma simulação que se assemelha aos dados reais. Cada indivíduo da população desse algoritmo genético representa o conjunto de parâmetros do modelo. A validação se dará a partir de experimentos que comparam o comportamento do modelo com o real. Para testar a robustez da ferramenta, o grau de dificuldade do ajuste é potencializado aumentando o tamanho do reticulado e introduzindo ruídos brancos aos dados.

1.1 Motivação

A sustentabilidade de ecossistemas tornou-se um objetivo primordial na atualidade trazendo como uma de suas conseqüências, a contínua necessidade de precisão e atualização dos dados do ambiente (COPPIN et al., 2004). Dentre os principais problemas em modelagem ambiental, pode-se destacar a configuração (calibração dos parâmetros)

dos modelos desenvolvidos. O desenvolvimento de modelos de simulação confiáveis não pode prescindir das etapas de calibração e validação. Dependendo da complexidade do modelo e do método escolhido, o tempo despendido na execução dessa calibração pode ser bem elevado. Esse custo do tempo da execução influencia diretamente na evolução do modelo, pois qualquer alteração visando uma melhoria em seu desenvolvimento terá uma análise mais demorada. Em particular, ao desenvolver modelos espacialmente explícitos baseados em autômatos celulares, o número de parâmetros a serem ajustados pode ser muito elevado e a função de ajuste do modelo pode apresentar uma relação não linear com os seus parâmetros, que também podem ser complexos ou multimodais. Por fim, é comum que modelos ambientais possuam parâmetros estocásticos ou mistos (modelos que englobam valores de configuração fixos e estocásticos), fazendo com que a função de avaliação da qualidade do modelo seja estocástica, ao invés de determinística. Todas essas características podem tornar a tarefa de calibração em modelagem ambiental um problema relativamente complexo do ponto de vista da otimização, em que técnicas heurísticas podem ser indicadas.

Modelos ambientais desenvolvidos são validados, geralmente, por meio de análises que comparam os resultados obtidos em simulações com aqueles realmente coletados acerca do fenômeno estudado. Quanto mais parecido o resultado (comportamento) simulado for do real, melhor representado está o problema e o modelo resultante tem maior segurança para prever um comportamento futuro desse fenômeno. Para que um modelo tenha bons resultados, ele deve levar em consideração as peculiaridades do problema. Todos os fatores que contribuem no comportamento do fenômeno em questão devem estar presentes, muitos deles como parâmetros do modelo criado. O ajuste de parâmetros de modelos consiste em encontrar a combinação de valores que contribui para que o resultado encontrado na simulação seja o mais parecido possível ao comportamento real do fenômeno.

Um dos principais problemas nesse tipo de modelagem diz respeito à identificação e uso dos parâmetros que o projetista (desenvolvedor) julga relevante para o modelo. Uma possível forma de identificar esses parâmetros é utilizando problemas inversos (TARRANTOLA, 2005). A solução de um problema inverso consiste em determinar causas baseando-se na observação dos seus efeitos (VELHO, 2001). É possível e interessante resolver problemas desse cunho utilizando algoritmos genéticos (TOMINAGA; KOGA; OKAMOTO, 2000). No contexto da modelagem, isso seria eficiente pois observando-se as características do fenômeno poderíamos descobrir quais parâmetros seriam relevantes, definir possíveis valores e com isso melhorar os resultados.

1.2 Justificativa

Um algoritmo genético (AG) é uma técnica de procura utilizada para encontrar soluções aproximadas em problemas de otimização e busca. Eles consistem numa classe

particular de algoritmos evolutivos que usam técnicas inspiradas na biologia evolutiva como hereditariedade, mutação, seleção natural e recombinação.

O AG é uma opção interessante como ferramenta de uso geral para o ajuste de parâmetros em modelagem ambiental pelas razões listadas a seguir:

- ❑ Modelos baseados em autômatos celulares, pela própria natureza de modelagem espacialmente distribuída, em geral apresentam muitos parâmetros desconhecidos a serem ajustados.
- ❑ Modelos desenvolvidos em sistemas ambientais em geral apresentam erros não lineares em relação aos parâmetros de configuração, gerando funções de erro não convexas e multimodais (possuem vários ótimos locais) (CORTES; SAAVEDRA, 2004).
- ❑ Os parâmetros de configuração dos modelos podem ser variados, ou seja, alguns parâmetros podem assumir valores contínuos, enquanto outros são discretos, tornando necessário um método de ajuste flexível e capaz de lidar com esses diferentes tipos de variáveis.
- ❑ É comum que sistemas ambientais possuam comportamentos estocásticos, de modo que suas simulações produzam resultados diferentes a partir de um mesmo conjunto de valores para os parâmetros do sistema. Portanto, o problema de ajuste desses parâmetros torna-se um problema de otimização em que a função objetivo apresenta ruído. Algoritmos evolutivos, no geral, têm se mostrado eficientes no contexto de otimização de funções com ruído.
- ❑ A avaliação da qualidade (erro de ajuste) de um dado conjunto de valores para os parâmetros do sistema normalmente requer a simulação de modelos complexos, o que costuma ser uma tarefa computacionalmente cara. Por exemplo, a calibração por métodos puramente estocásticos, como simulação de Monte Carlo, demanda elevado esforço computacional, pois exige a análise de uma quantidade elevada de estados do sistema (ROBERT; CASELLA; CASELLA, 1999). Por outro lado, o fato da busca dos algoritmos evolutivos ser direcionada por uma heurística baseada na “aptidão” das soluções já avaliadas (pressão seletiva), torna o AG um método de busca “estocástico informado”, ou seja, direcionado pela pressão seletiva e acumulativa da operação de seleção.

Nesse contexto, um módulo de calibração que atenda a todas as peculiaridades de um fenômeno modelado se faz necessário. Ele deve ser, principalmente, eficiente no quesito de tempo de retorno de resultados. Ele também deve contribuir com a configuração dos parâmetros levando em conta a distribuição espacial dos autômatos, podendo generalizar algumas características desses parâmetros e obter melhores resultados. Para isso é

interessante analisar e levar em consideração, dentre outras coisas, o impacto que cada célula do autômato gera em seus vizinhos. Espera-se que o modelo baseado em AG seja capaz de escolher os valores dos parâmetros de modo similar a um especialista da área do modelo, porém, de uma forma bem mais rápida.

1.3 Objetivos

O objetivo principal desse trabalho é propor uma abordagem de ajuste que seja capaz de definir valores adequados para os parâmetros de modelos baseados em autômatos celulares, afim de reproduzir o comportamento apresentado nos dados observados. Para isso, será implementado um módulo de ajuste genérico baseado em algoritmos genéticos (MITCHELL, 1998; EIBEN; SMITH et al., 2003).

Durante o desenvolvimento de nossa pesquisa, pretende-se alcançar os seguintes objetivos específicos:

- ❑ Estudo e implementação do modelo de reprodução e dispersão do vetor barbeiro para controle da doença de chagas (SLIMI et al., 2009), o qual será usado como o estudo de caso para a avaliação do módulo de ajuste de parâmetros proposto.
- ❑ Desenvolver um método de ajuste de parâmetros de modelos baseados em autômatos celulares (AC), utilizando Algoritmos Genéticos (AG).
- ❑ Validar o método proposto, analisando seu desempenho em reproduzir a dinâmica comportamental gerada a partir do modelo usado como estudo de caso (modelo de referência).
- ❑ Avaliar a eficácia e robustez do método proposto em relação a variações nos dados a serem modelados. Essas variações podem ser obtidas a partir de pequenas modificações no modelo de referência, como, por exemplo, mudança no tamanho do reticulado do autômato celular.

1.4 Hipótese

Este trabalho busca tratar as seguintes hipóteses: Aplicando-se métodos bio-inspirados, mais especificamente AGs, é possível construir uma abordagem de ajuste automático de parâmetros de modelos ambientais? Quantos parâmetros a abordagem seria capaz de prever com robustez? O ajuste por meio da abordagem proposta traz algum impacto considerável na distribuição espacial do modelo do AC?

1.5 Contribuições

A maior contribuição deste trabalho é investigar uma solução automática e robusta para uma das etapas mais difíceis, trabalhosas e demoradas da criação de um modelo ambiental que é o ajuste de seus parâmetros. Com um algoritmo eficaz, o projetista do modelo pode não só obter bons valores para os parâmetros de forma mais rápida, como também, a partir daí, concluir quais parâmetros são ou não relevantes. Uma solução com economia de tempo permite um processo cíclico de desenvolvimento com uma variedade maior de alternativas e análises.

1.6 Organização da Dissertação

O restante deste documento está organizado da seguinte forma:

- ❑ Capítulo 2: introduz os principais conceitos relacionados aos Autômatos Celulares, descrevendo seu funcionamento básico e apresentando dois exemplos clássicos de ACs.
- ❑ Capítulo 3: descreve os componentes principais de Algoritmos Genéticos, detalhando seu funcionamento e opções de configuração.
- ❑ Capítulo 4: apresenta o estado da arte, descrevendo os principais trabalhos nas áreas correlatas.
- ❑ Capítulo 5: descreve detalhadamente o modelo proposto em (SLIMI et al., 2009), o qual é utilizado como estudo de caso em nossos experimentos.
- ❑ Capítulo 6: detalha as configurações e estratégias evolutivas utilizadas para atingir os objetivos propostos.
- ❑ Capítulo 7: apresenta os resultados dos experimentos com nossas estratégias evolutivas para o ajuste dos parâmetros do modelo ao conjunto de dados de referência, bem como as respectivas análises de desempenho e robustez.
- ❑ Capítulo 8: descreve as principais conclusões obtidas a partir dos desempenhos alcançados durante o trabalho, bem como as publicações geradas e algumas propostas para a continuidade da pesquisa.

Autômatos Celulares

Nesse capítulo são apresentados definições e um breve histórico de Autômatos Celulares (AC). Na sequência, são detalhados os conceitos de espaço celular, dimensão, estado, vizinhança e regras de transição. Autômatos celulares determinísticos e probabilísticos são discutidos. Ao final, um exemplo de autômato celular unidimensional é ilustrado (Regra 30).

2.1 Histórico e Definição

Autômatos celulares têm sido usados como alternativa de modelagem e simulação de sistemas, nos quais as equações diferenciais envolvem simulações complexas e computacionalmente caras (TOFFOLI, 1984). Eles são sistemas dinâmicos, discretos, que são capazes de descrever fenômenos complexos a partir de um conjunto de regras simples (TOFFOLI; MARGOLUS, 1987), facilitando a construção do modelo e tornando o AC uma importante ferramenta de modelagem. Uma importante aplicação desses conceitos ocorre no campo da epidemiologia, em que os ACs possuem uma estrutura adequada para estudar fenômenos de disseminação espacial.

A teoria de autômatos celulares foi introduzida em um trabalho conjunto por dois matemáticos, o húngaro John von Neumann e o polaco Stanislaw Ulam, na década de 1940. Neumann trabalhava com foco na auto-reprodução, enquanto Ulam estudava o crescimento de cristais, usando uma rede bidimensional simples no seu modelo. Ulam sugeriu que von Neumann descrevesse uma abstração matemática semelhante à que ele usava, assim, descobriram que uma simulação de auto reprodução poderia ser descrita por um algoritmo (NEUMANN, 1966).

Em 1970, o inglês John Horton Conway propôs um modelo de autômato celular chamado “Jogo da Vida” (ou *Game of Life*) (CONWAY, 1970), que é o mais conhecido modelo de AC até a atualidade. Ele reproduz, por meio de regras simples, alterações em grupos de células, onde cada célula nasce ou morre de acordo com a sua vizinhança. Em 1982 Stephen Wolfram, apresentou um estudo de funcionamento e formação de padrões

de ACs (WOLFRAM, 1982). Ele conduziu vários outros estudos mostrando que regras locais simples são capazes de reproduzir padrões complexos e desenvolveu uma teoria de classificação de AC baseada no padrão espaço-temporal (WOLFRAM, 1986).

2.1.1 Espaço e Dimensão de um Autômato Celular

Autômatos Celulares (AC)s são sistemas discretos no tempo, espaço e estado. Eles são formados por um conjunto de células distribuídas espacialmente em uma rede unidimensional (linear), bidimensional (plano) ou tridimensional (espacial), que seguem um padrão de conexão baseado na vizinhança.

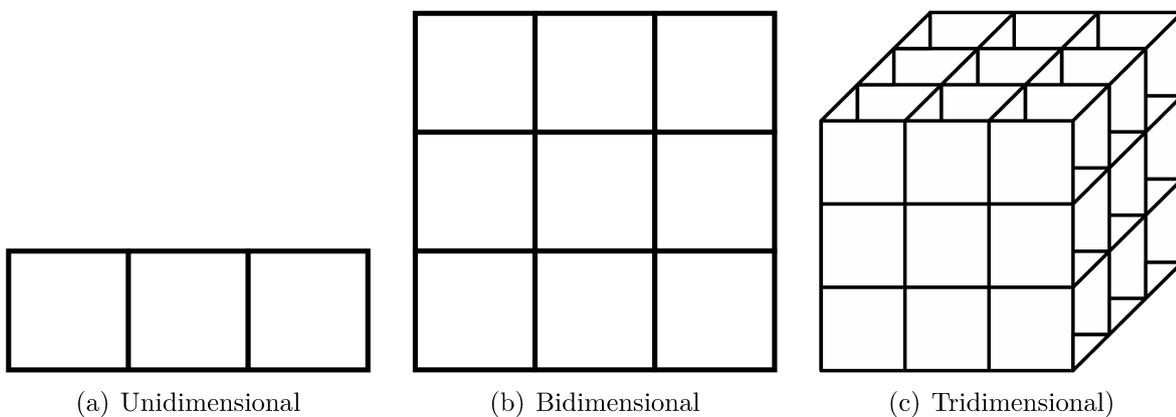


Figura 1 – Espaço celular: (a) unidimensional (b) bidimensional e (c) tridimensional.

O espaço celular é formado por um conjunto de células espacialmente distribuídas, também conhecido por reticulado, que seguem o mesmo padrão de conexão com seus vizinhos mais próximos. Cada célula possui um número definido de estados e a evolução desses estados acontece a partir da definição de regras de transição que são aplicadas em intervalos de tempo discretos, ou seja, a cada iteração de tempo. As regras são aplicadas em todo o reticulado e o estado de cada célula no tempo $t + 1$ é calculado de acordo com seu próprio valor e o das células vizinhas no tempo t (MITCHELL, 1998).

As células que compõem o espaço celular do AC podem ser de formas geométricas variadas, desde que todas elas tenham o mesmo formato, conforme ilustrado na Figura 2.

Na prática, autômatos celulares são implementados com tamanho finito. Para delimitar seu espaço celular são definidas regiões de borda ou fronteiras e as células dessa região são tratadas de maneira diferente das outras. Essas células servem principalmente para tratar a vizinhança incompleta das células próximas a borda, conforme ilustrado na Figura 3. Esse tratamento pode considerar que as extremidades do espaço celular estão conectadas, chamada borda periódica, cíclica ou circular (Figura 3(a)). Outro tipo de borda é a reflexiva (Figura 3(b)) onde as células da borda são repetidas completando a vizinhança. Já na borda fixa, as células a serem completadas já tem um estado pré definido representado por “X” na Figura 3(c).

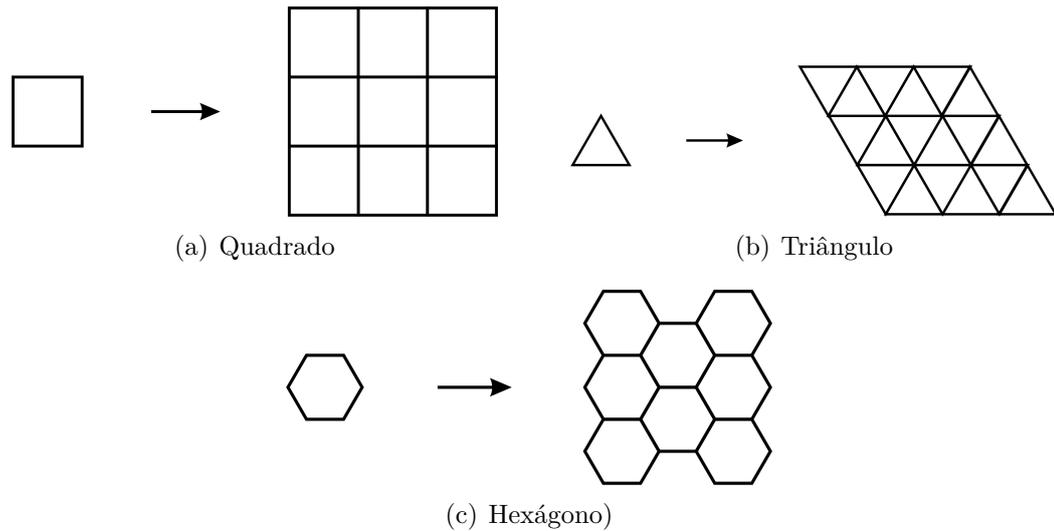


Figura 2 – Possíveis formatos das células de um Autômato Celular: (a) quadrado (b) triângulo e (c) hexágono.

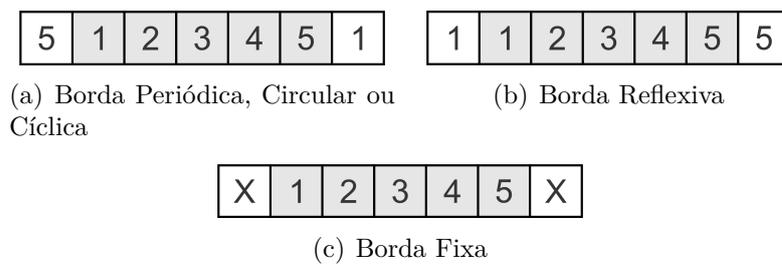


Figura 3 – Tipos de borda mais comuns em um Autômato Celular unidimensional: (a) periódica, circular ou cíclica, (b) reflexiva, (c) fixa.

Considerando:

- $\Sigma \rightarrow$ um conjunto finito de estados que cada uma das células (c) podem assumir;
- $\Phi \rightarrow$ a vizinhança de uma célula (c);
- $f \rightarrow$ as regras de transição;
- $\mathcal{D} \rightarrow$ a dimensão do reticulado.

Pode-se formalizar um autômato celular por meio da quádrupla (OLIVEIRA; OLIVEIRA; OMAR, 2001):

$$\mathcal{A} = (\Sigma, \Phi, f, \mathcal{D}) \quad (1)$$

2.1.2 Estado e Vizinhança

Em um AC, cada célula tem um estado definido dentro de um conjunto finito de estados (Σ). A cada passo de tempo, todas as células podem mudar de estado ou permanecer

no mesmo, dependendo do seu estado atual e do estado de sua vizinhança. Considera-se vizinhança de uma célula, aquelas que têm uma proximidade física pré definida, lógica ou arbitrária, cujas características vão influenciar no estado no próximo instante de tempo.

Os ambientes para mapeamento de propagação são melhor modelados por um AC bidimensional, que comumente emprega dois tipos de vizinhança: von Neumann e Moore (SARKAR, 2000). A Figura 4 mostra as células consideradas em cada tipo. A vizinhança de von Neumann é formada pela célula central e suas quatro vizinhas nas direções cardinais, totalizando 5 células, conforme ilustrado na Figura 4(a). A vizinhança de Moore compreende todas as células da vizinhança de von Neumann (central e vizinhas cardinais), acrescidas das quatro células diagonais, totalizando 9 células, como mostrado na Figura 4(b).

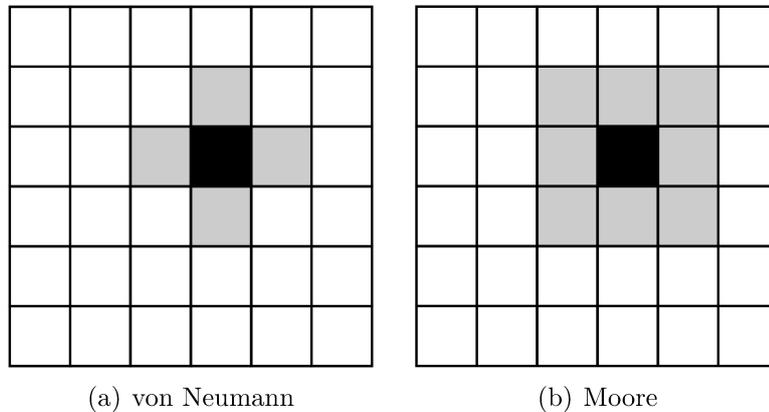


Figura 4 – Vizinhanças mais comuns em um Autômato Celular: (a) von Neumann e (b) Moore.

A vizinhança Φ de uma célula é sempre composta por um conjunto finito de células. A cardinalidade desse conjunto é definida pelo tipo de vizinhança e pelo seu raio \mathcal{R} . O raio pode ser definido como o tamanho do passo (quantidade de células) que será considerada vizinhança, em cada uma das direções. Na Figura 5, vemos que no caso unidimensional(Figura 5(a)), a vizinhança de 3 células é definida pelo raio 1, enquanto a vizinhança de 5 células (5(b)) é definida pelo raio 2. A Figura 6 apresenta o raio associado a diferentes tamanhos de vizinhança de Von Neumann e Moore em um espaço bidimensional. A cardinalidade de Φ , para um AC bidimensional com vizinhança de Moore, pode ser calculada por meio da equação Eq. 2.

$$|\Phi| = (2 * \mathcal{R} + 1)^2 - 1 \quad (2)$$

2.1.3 Regras de Transição

A evolução dos estados acontece a partir da definição de regras de transição (f), que são aplicadas a todas as células do reticulado em intervalos de tempo discretos. Essas regras de

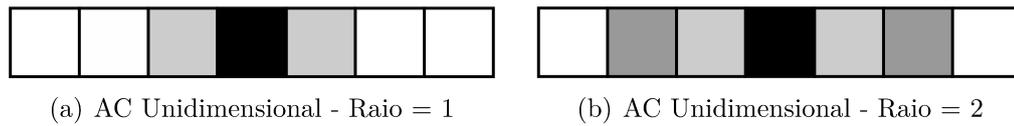


Figura 5 – Vizinhança de um Autômato Celular unidimensional de acordo com o raio (\mathcal{R}): (a) $\mathcal{R} = 1$ e (b) $\mathcal{R} = 2$.

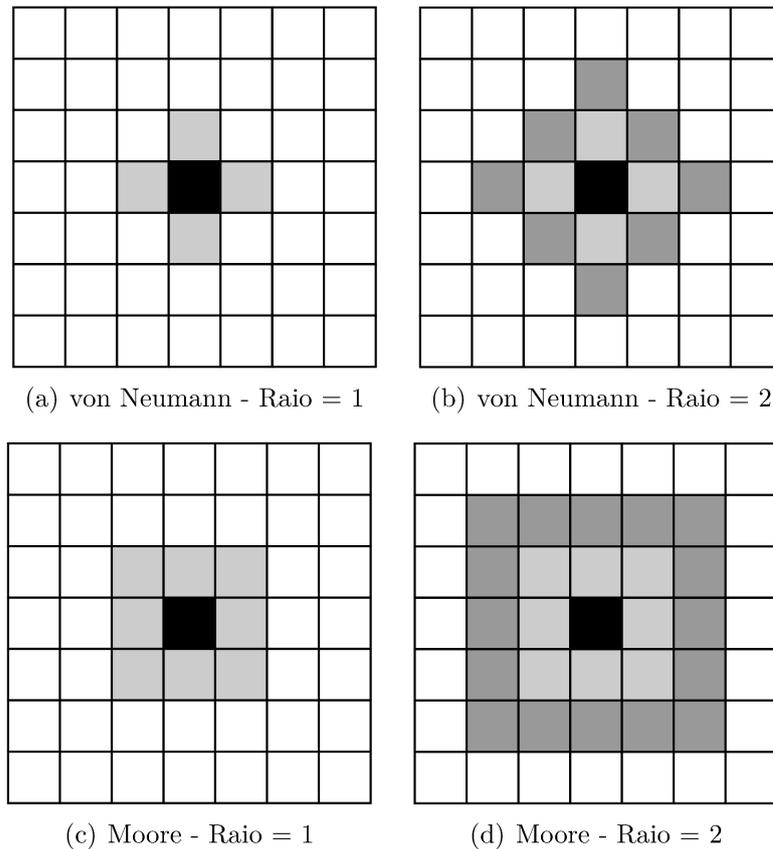


Figura 6 – Vizinhança de um Autômato Celular bidimensional de acordo com o raio (\mathcal{R}): (a) von Neumann com $\mathcal{R} = 1$, (b) von Neumann com $\mathcal{R} = 2$, (c) Moore com $\mathcal{R} = 1$ e (d) Moore com $\mathcal{R} = 2$.

transição, normalmente representadas por tabelas, são compostas por um conjunto finito de regras que contemplam todas as combinações possíveis entre o estado da própria célula e o estado da vizinhança para prever o próximo estado no intervalo de tempo seguinte. As regras de transição conduzem a evolução temporal do AC. A seção 2.3 apresenta exemplos de regras de transição de ACs unidimensionais e bidimensionais.

Algumas regras levam em consideração a quantidade de células vizinhas que estão em um determinado estado, denominadas regras totalísticas. Nesse caso não é necessário especificar todas as combinações possíveis das células da vizinhança, já que o que determinará o estado é a quantidade de vizinhos contabilizados em cada estado possível e não a posição que eles estão. Quando o estado da própria célula não faz parte da regra de transição, ela é denominada totalística externa.

2.2 Autômato Celular Probabilístico

Os ACs tradicionalmente são determinísticos, ou seja, as regras de transição determinam saídas específicas para cada caso determinando somente um estado de saída. Porém, existem modelos que utilizam regras de transição regidas por probabilidades, o que reflete em uma maior proximidade com os fenômenos naturais. As regras probabilísticas possuem mais de um estado resultado possível para uma mesma configuração de vizinhança. Autômatos celulares probabilísticos têm sido estudados como um modelo possível para a simulação de vários sistemas biológicos, como os modelos epidemiológicos. Nesses últimos, tais modelo possibilitam estudar a disseminação espacial de doenças e seus vetores (SLIMI et al., 2009; GAGLIARDI; ALVES, 2010; CISSÉ; YACOUBI; GOURBIERE, 2014; PEREIRA; SCHIMIT, 2018).

2.3 Exemplos

Nessa seção, apresentaremos dois exemplos de ACs, um unidimensional e outro bidimensional, para ilustração dos conceitos.

2.3.1 Autômato celular unidimensional - Regra 30

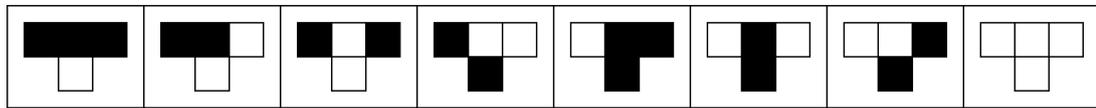
O AC denominado “Regra 30” é unidimensional e determinístico. O conjunto de seus estados é binário, ou seja, $\Sigma = \{0,1\}$. A vizinhança de uma célula é composta pelas células que se localizam a esquerda e direita dela, logo, seu raio é 1 ($\mathcal{R}= 1$). As regras de transição (f) levam em conta os estados da vizinhança (3 células), logo, $|f| = 2^3 = 8$. Considerando o quadrado preto como o estado celular 1 e o quadrado branco como o estado celular 0, podemos representar a regra de transição como na Figura 7. Podemos representar essa regra de transição também por meio da Tabela 1.

Tabela 1 – Regra de transição do AC “Regra 30”.

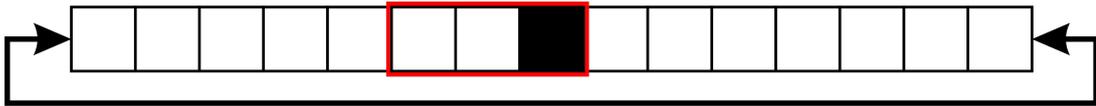
| Vizinho Esquerdo (c_{i-1}) | Célula (c_i) | Vizinho Direito (c_{i+1}) | Resultado (c_i) ^{t+1} |
|--------------------------------|------------------|-------------------------------|------------------------------------|
| 1 | 1 | 1 | 0 |
| 1 | 1 | 0 | 0 |
| 1 | 0 | 1 | 0 |
| 1 | 0 | 0 | 1 |
| 0 | 1 | 1 | 1 |
| 0 | 1 | 0 | 1 |
| 0 | 0 | 1 | 1 |
| 0 | 0 | 0 | 0 |

Na Figura 7(b) temos a configuração inicial do reticulado unidimensional, com uma região destacada em vermelho. Essa região corresponde à análise de uma vizinhança para

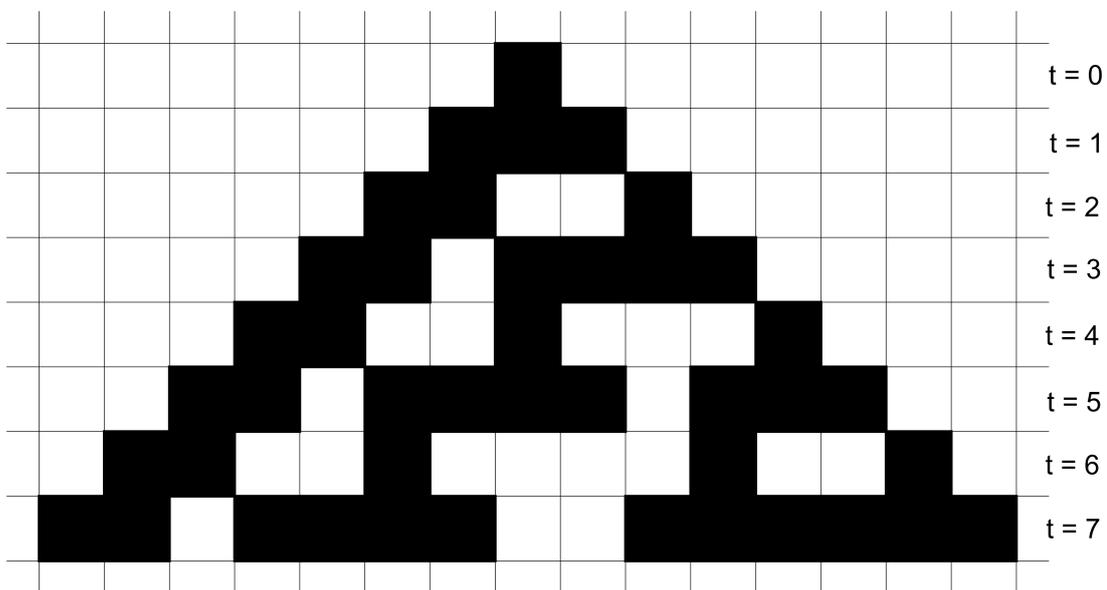
a aplicação da regra, no caso a vizinhança 0 – 0 – 1 que retorna o valor 1. Na Figura 7(c) temos o histórico da evolução do reticulado do AC nas 7 primeiras iterações de tempo, a partir da configuração inicial apresentada na Figura 7(b).



(a) Regras - Regra 30



(b) AC unidimensional - Vizinhança



(c) Histórico - Regra 30

Figura 7 – Autômato Celular unidimensional “Regra 30”: (a) regra de transição (b) reticulado inicial e (c) evolução espaço-temporal por 7 passos de tempo.

2.3.2 Autômato celular bidimensional - Jogo da vida

O Jogo da Vida (*Game of Life*) é o autômato celular mais conhecido do mundo. Proposto por John Conway (CONWAY, 1970), se trata de um AC bidimensional, com conjunto de estados binários, $\Sigma = \{0, 1\}$ que usa vizinhança de Moore com raio 1 (como na Figura 6(c)), logo, a cardinalidade da vizinhança é:

$$|\Phi| = (2 * \mathcal{R} + 1)^2 - 1 = (3^2) - 1 = 8. \quad (3)$$

As regras de transição f são totalísticas e definidas de forma que:

- uma célula viva com 2 ou 3 vizinhos, permanece viva;

- ❑ uma célula viva com menos de 2 vizinhos, morre de solidão;
- ❑ uma célula viva com mais de 3 vizinhos, morre sufocada;
- ❑ uma célula morta com exatamente 3 vizinhos vivos, se torna uma célula viva.

Na Fig 8, apresentamos um reticulado inicial, com uma vizinhança destacada em vermelho. O somatório de células vivas é 1, logo, no próximo instante de tempo o estado da célula analisada será morta.

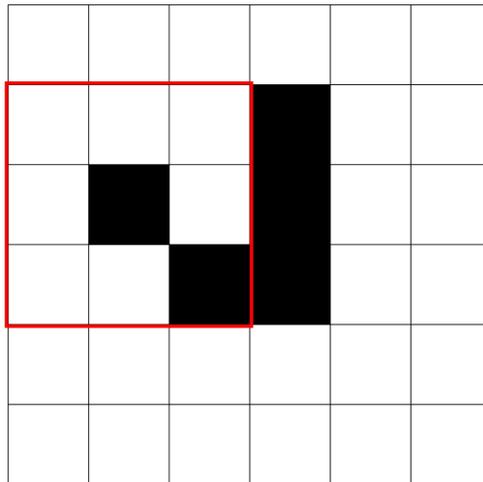


Figura 8 – Exemplo de reticulado do AC “Jogo da Vida”, com uma de suas vizinhanças em destaque.

Na Fig 9, temos a evolução do AC “Jogo da Vida”, em 5 iterações de tempo, a partir do reticulado inicial da Figura 8. É possível observar padrões que se movem pelo reticulado (gliders), repetindo seu formato com a posição deslocada em relação a diagonal, se comparado com o anterior. Estruturas desse tipo possuem ciclos, pois de acordo com a evolução do reticulado, voltam a se repetir. A Figura 9(a) e a Figura 9(e) são exemplos de um mesmo padrão deslocado na direção da diagonal inferior direita.

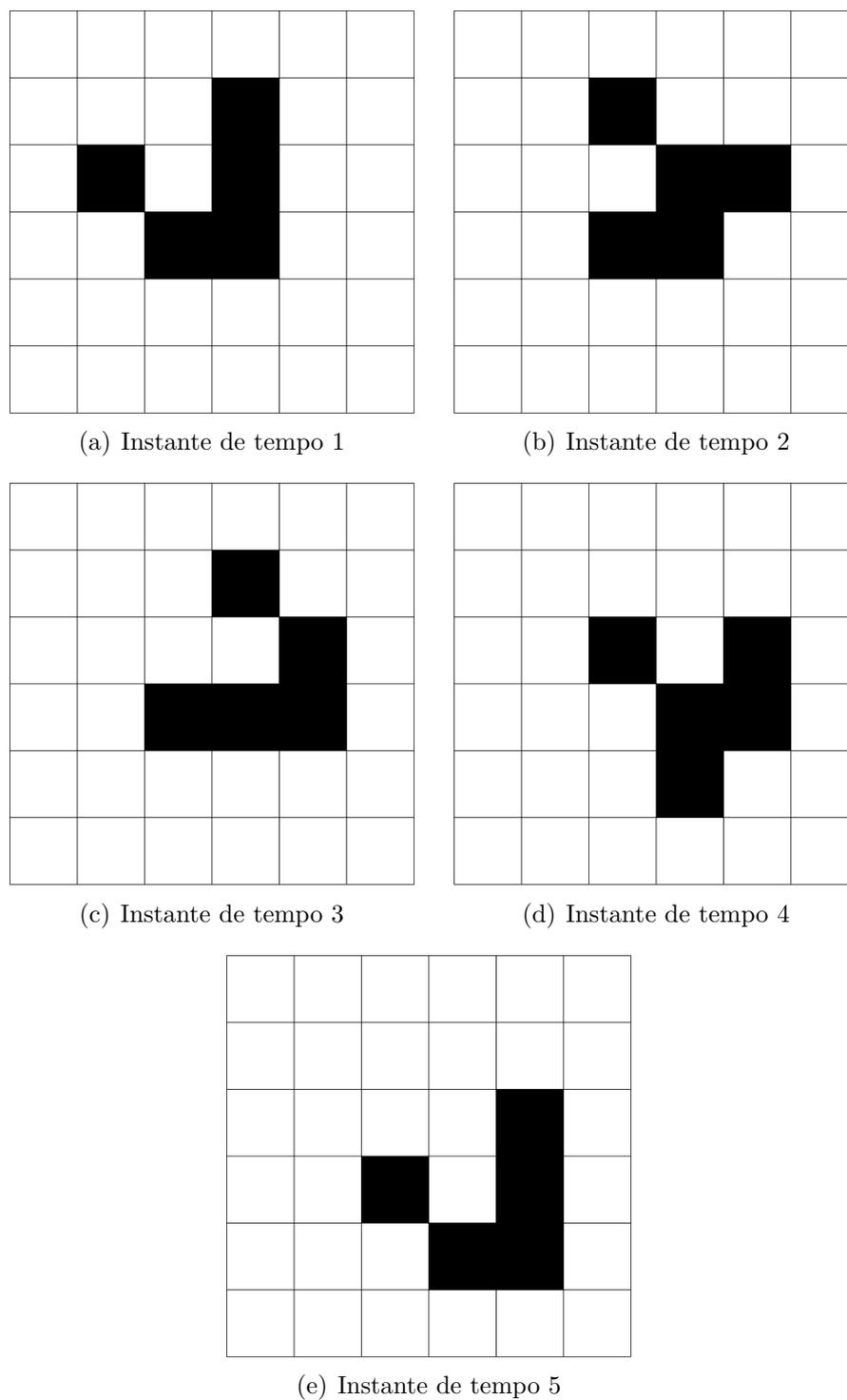


Figura 9 – Histórico de evolução do AC “Jogo da Vida” nos instantes de tempo (t): (a) $t = 1$, (b) $t = 2$, (c) $t = 3$, (d) $t = 4$ e (e) $t = 5$.

Algoritmos Genéticos

Nesse capítulo, serão apresentados os conceitos de Algoritmos Genéticos (AG), além de um breve histórico seguido das principais etapas.

3.1 Histórico

A Computação Evolutiva (CE) é um ramo de pesquisa da área de Inteligência Artificial que propõe um paradigma para solução de problemas inspirados na teoria da seleção natural, proposta por Charles Darwin em 1859 (DARWIN, 1859). Os algoritmos evolutivos se inspiram na ideia básica da teoria de Darwin, na qual os indivíduos melhor adaptados têm vantagem evolutiva sobre os demais. AG é uma classe particular dos algoritmos evolutivos que utiliza o processo de evolução natural como um paradigma de solução de problemas de busca e otimização (MITCHELL, 1998). Eles são inspirados no princípio Darwiniano da evolução das espécies (DARWIN, 1859) e na reprodução genética (HOLLAND, 1975; GOLDBERG, 1989). Algoritmos Genéticos foram introduzidos por John Holland (HOLLAND, 1975) em 1975, quando ele desenvolveu uma forma de incorporar os conceitos evolutivos em ferramentas computacionais adaptáveis. As ideias de Holland foram posteriormente aprimoradas por David E. Goldberg (GOLDBERG; HOLLAND, 1988; GOLDBERG, 1989).

Os Algoritmos Genéticos são estocásticos e constituem uma técnica de busca heurística, altamente paralela e adaptativa. Os princípios biológicos como hereditariedade, mutação, seleção natural e recombinação (ou *crossover*) são codificados em algoritmos computacionais que buscam solucionar da melhor forma um determinado problema evoluindo populações direcionadas por uma função objetivo.

De acordo com esse princípio, a seleção de indivíduos é feita de modo a privilegiar aqueles mais aptos à sobrevivência, que, conseqüentemente, têm a maior probabilidade de reprodução. O indivíduo que se reproduz, transmite seu código genético a seus descendentes e com isso suas características são mantidas na população. A manipulação dessas características define o perfil da população. A implementação se dá por meio de uma

simulação em que uma população de soluções candidatas sofre as mudanças resultantes da aplicação dos processos evolutivos, direcionadas pelo objetivo pré estabelecido.

A vantagem mais notória dos AGs está na generalização da ferramenta de resolução de problemas, possibilitando a sua aplicação a partir da simples descrição matemática da solução desejada, pois se houvesse a necessidade de especificar os passos que levam ao resultado, eles certamente seriam específicos de cada caso. Isso proporciona ao algoritmo robustez e flexibilidade o que faz com que seus procedimentos genéricos e adaptáveis sejam aplicados na solução de problemas complexos.

3.2 Representação do Indivíduo

Para a construção dos elementos que compõem o AG, o primeiro e um dos mais importantes passos é definir como será a representação do problema para o qual se busca solução. Um indivíduo é uma representação de uma abstração que contempla uma dessas possíveis soluções do problema. Normalmente, os indivíduos são implementados por meio de uma estrutura de dados, como vetores de tamanho fixo. Os tipos mais comuns de representações são: binária, inteira ou real (MITCHELL, 1998; HOLLAND et al., 1992). Cada posição dessa representação da solução é equivalente a um gene, do ponto de vista evolutivo, enquanto o indivíduo equivale a um cromossomo. A Figura 10 exemplifica as principais representações encontradas na literatura: binária, inteira e real. Esse conjunto de valores compõe a informação genética do indivíduo, e é por meio disso que se perpetuam conceitos como hereditariedade, adaptação, etc do processo evolutivo.

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 0 | 1 | 1 | 0 | 1 | 0 | 0 | 1 |
|---|---|---|---|---|---|---|---|

(a) Representação Binária

| | | | | | | | |
|---|----|---|---|---|---|---|---|
| 5 | 14 | 2 | 7 | 3 | 8 | 1 | 6 |
|---|----|---|---|---|---|---|---|

(b) Representação Inteira

| | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|
| 2,2 | 1,7 | 4,0 | 3,1 | 8,6 | 6,0 | 2,8 | 5,4 |
|-----|-----|-----|-----|-----|-----|-----|-----|

(c) Representação Real

Figura 10 – Tipos mais comuns de representação de indivíduos de um AG: (a) binária (b) inteira e (c) real.

3.2.1 População

População é o conjunto de indivíduos que sofrerão os processos evolutivos do AG. A população compõe o espaço de busca e seu tamanho é um parâmetro do AG que deve ser

ajustado de forma consciente pois afeta o desempenho, eficiência e tempo de execução do algoritmo. Na Figura 11 temos uma representação gráfica de uma população, composta por um conjunto de cromossomos, que por sua vez é formado por um conjunto de genes.

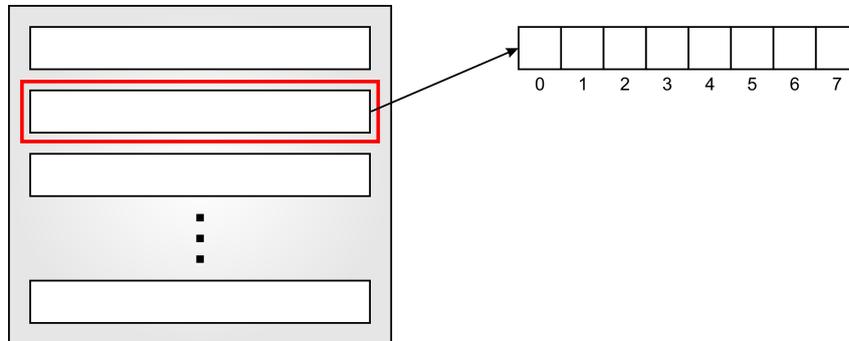


Figura 11 – Representação de uma população.

3.2.1.1 Geração da População Inicial

O processo evolutivo se inicia com a geração de uma população inicial, onde cada indivíduo dessa população é um candidato a solução. Essa geração inicial pode acontecer de forma totalmente aleatória ou guiada por alguma heurística pertinente ao problema. No caso de problemas com restrições, essa geração pode garantir que somente possíveis soluções válidas sejam criadas ou deixar soluções inválidas nesse momento, que serão descartadas durante o processo evolutivo.

3.3 Função Objetivo - *Fitness*

A função objetivo ou *fitness*, é responsável por guiar o processo evolutivo. Ela mapeia a representação de um indivíduo em um valor numérico associado à adaptação, dimensionando o quanto um indivíduo é bom ou ruim (MITCHELL, 1998). Dessa forma é selecionado quem permanece ou não na população, quem tem preferência nas operações, quem é melhor adaptado. Essa aptidão é determinada pelo cálculo de uma função objetivo que é específica do problema investigado. A aptidão de um indivíduo é recalculada sempre que ele sofre alguma alteração.

Um problema pode ter como característica mais de um objetivo para sua solução. Nesse caso, chamado otimização multiobjetivo, a solução pode ser composta por uma combinação de pesos entre as funções, identificando e privilegiando aquelas que têm maior relevância. Existem outros tipos de problemas que mudam de objetivo durante o processo de busca por solução, nesse caso, a função objetivo muda durante o processo, o que caracteriza o AG como um processo extremamente adaptável.

3.4 Fluxo de Execução

O processo de evolução executado por um algoritmo genético corresponde a um processo de busca em um espaço de soluções potenciais para um problema. Esse processo de busca é multidirecional uma vez que manipula uma população de soluções candidatas, guiadas pela função objetivo. AGs apresentam um balanço entre aproveitamento de melhores soluções e exploração do espaço de busca.

São parâmetros de configuração de um AG:

- ❑ Tamanho da população,
- ❑ Número de gerações,
- ❑ Probabilidade de Reprodução Pr (taxa de reprodução),
- ❑ Probabilidade de Mutação Pm (taxa de mutação)

O funcionamento típico de um AG é apresentado na Figura 12. Uma vez definidos os parâmetros de configuração, o processo começa com a geração de uma população inicial, cujo os indivíduos (possíveis soluções) são avaliados e classificados (ranqueados) de acordo com sua aptidão para resolver o problema. Em seguida inicia-se o processo iterativo, ou seja, as gerações do AG. Enquanto o critério de parada não for satisfeito, o algoritmo realiza a seleção de indivíduos, os quais se reproduzem entre si. Os filhos resultantes podem ou não sofrer mutações de seus genes e são avaliados e reinseridos na população de acordo com sua aptidão, formando a população que será submetida a nova geração. Nas seções seguintes, serão detalhados os principais processos do fluxograma básico de um AG: seleção, operadores genéticos, reinserção e critério de parada.

3.5 Seleção

O processo de seleção é responsável por escolher dentre todos os indivíduos que compõem a população, aqueles que farão parte das manipulações dos operadores genéticos. Esse operador nunca cria novos indivíduos, apenas faz uma triagem de acordo com um critério pré estabelecido. O objetivo é simular a principal característica do processo evolutivo, a seleção natural, onde os mais aptos sobrevivem e propagam o seu material genético. No geral, os operadores de seleção escolhem com mais frequência aqueles indivíduos que possuem melhor avaliação na função objetivo, mas sem descartar a possibilidade de indivíduos com avaliações piores serem escolhidos, pois assim garantem a diversidade da população.

Os operadores de seleção são normalmente estocásticos e estão associados ao conceito de pressão seletiva. Quanto mais o operador for voltado para selecionar os melhores indivíduos da população, maior pressão seletiva ela sofre. Essa pressão esta diretamente

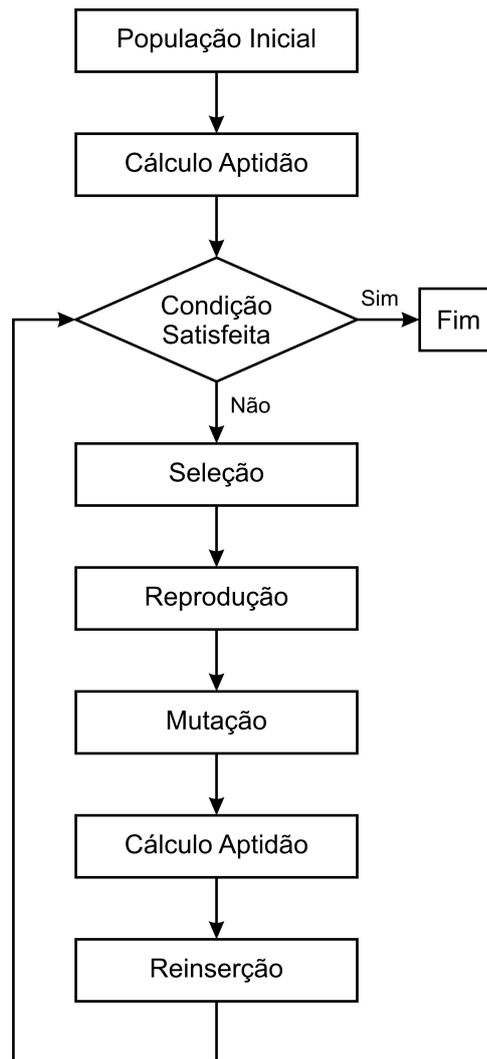


Figura 12 – Fluxograma de um Algoritmo Genético.

ligada com convergência prematura, onde o AG fica preso em um mínimo local. Por outro lado, quanto menor a pressão seletiva, mais diversidade a população possui, podendo demorar muito para convergir ou até mesmo não encontrar uma solução satisfatória.

3.5.1 Roleta

Proposto por (HOLLAND et al., 1992), o processo de seleção roleta é baseado em proporcionalidade, onde os indivíduos têm a chance de serem escolhidos proporcionalmente ao resultado da avaliação da função objetivo. A probabilidade de seleção de um indivíduo é calculada pela equação 4, onde x é o indivíduo e n o número total de indivíduos.

$$\text{Probabilidade de seleção}(x) = \frac{\text{aptidão}(x)}{\sum_{i=1}^n \text{aptidão}(i)} \quad (4)$$

A Figura 13 apresenta um exemplo com uma população de 5 indivíduos. Para cada indivíduo, é apresentado o valor de sua aptidão e a probabilidade de cada indivíduo ser selecionado. A Figura 14 apresenta graficamente a roleta, na qual o tamanho de cada fatia

é proporcional à probabilidade de seleção dos indivíduos, que são representados na cor roxa (indivíduo 1), amarela (indivíduo 2), azul (indivíduo 3), rosa (indivíduo 4) e verde (indivíduo 5). A ideia do processo de seleção se assemelha a uma roleta que é girada e para aleatoriamente em uma posição, onde os valores de *fitness* mais promissores ocupam um espaço maior com mais chance de serem selecionados.

| Indivíduo | Aptidão | Probabilidade |
|-----------|---------|---------------|
| 1 | 10 | 25% |
| 2 | 3 | 7,5% |
| 3 | 15 | 37,5% |
| 4 | 5 | 12,5% |
| 5 | 7 | 17,5% |

Figura 13 – Aptidão e probabilidade dos indivíduos da população.

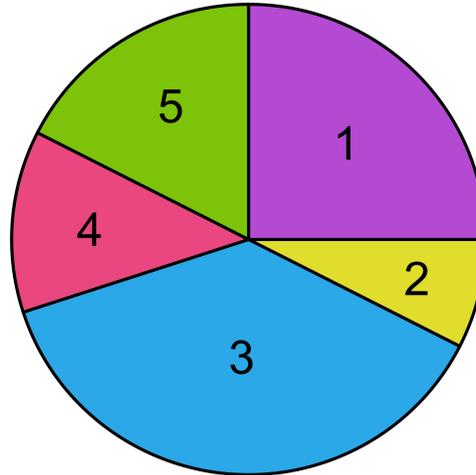


Figura 14 – Seleção - Roleta.

3.5.2 Torneio

No torneio, uma amostra de n indivíduos é selecionada aleatoriamente, ou de acordo com alguma heurística, e dentro dessa amostra os indivíduos competem pela seleção final de um indivíduo para participar da reprodução. Essa seleção final pode ser determinística ou probabilística. O processo é repetido até o número de indivíduos necessários para o processo de reprodução serem selecionados. Na Figura 15, três indivíduos da população são selecionados e competem pelo valor de sua aptidão. A figura ilustra um problema de maximização e, nesse caso, o indivíduo com maior valor da função objetivo vence o torneio e é selecionado para a reprodução. Esse é um exemplo de seleção determinística.

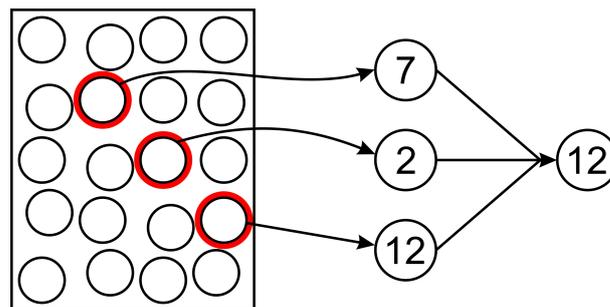


Figura 15 – Seleção por torneio.

3.5.3 Elitismo

No elitismo, uma parte da população considerada promissora é preservada e mantida na próxima geração da população, com o objetivo de evitar que bons indivíduos se percam no processo evolutivo aumentado assim a pressão seletiva. O restante da população da próxima geração é formado por *crossover*, utilizando-se algum outro método para selecionar os pais (como a roleta e o torneio).

3.6 Operadores Genéticos

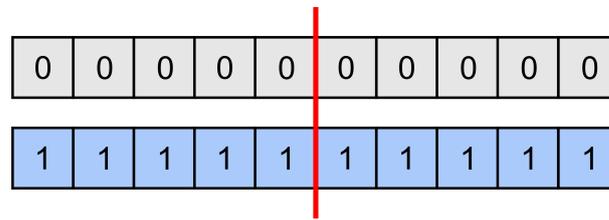
Os operadores genéticos são responsáveis pela geração de novos indivíduos na população. Esses novos indivíduos são provenientes de operações que modificam aqueles que foram selecionados da população corrente, mantendo parte de sua carga genética. A quantidade de indivíduos da população que sofrerá essas modificações é determinada por dois parâmetros de probabilidades: probabilidade de reprodução (Pr) e probabilidade de mutação (Pm).

3.6.1 Reprodução - *Crossover*

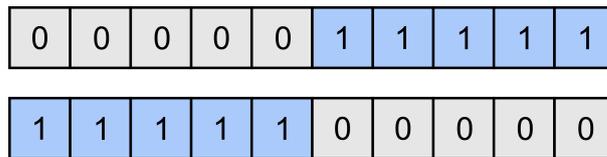
A reprodução ou *crossover* produz novos indivíduos a partir da combinação de indivíduos pais. Essa combinação é feita a partir de trocas de fragmentos dos cromossomos, garantindo a hereditariedade. O tipo de reprodução mais simples é o cruzamento com um único ponto de corte (Figura 16) (HOLLAND et al., 1992). Nele, dois indivíduos, denominados “pais” são selecionados da população corrente e uma posição da representação do indivíduo é selecionada aleatoriamente (Figura 16(a)). Essa divisão gera dois fragmentos de cromossomo em cada pai que são combinados entre si gerando dois novos indivíduos denominados “filhos” (Figura 16(b)). O processo de reprodução com os mesmos pais, pode gerar vários filhos diferentes, uma vez que a fragmentação acontece de maneira aleatória.

Um cruzamento pode ter múltiplos pontos de corte. Um exemplo com dois pontos de corte é representado na Figura 17. Os pais selecionados têm duas posições de seus cromossomos sorteadas aleatoriamente (Figura 17(a)) e geram filhos resultantes da combinação desses fragmentos (Figura 17(b)).

O cruzamento uniforme ilustrado na Figura 18 utiliza uma estrutura denominada “máscara”, para definir os pontos de corte. Ela é do tamanho do cromossomo e seleciona as posições do cromossomo que serão transmitidas pelo primeiro pai, sendo o segundo pai responsável pelas posições que complementam a máscara. Na Figura 18, temos os pais selecionados para o cruzamento e na sequência a máscara que define a configuração dos filhos. As posições da máscara com o valor 0, buscam o gene do cromossomo do primeiro

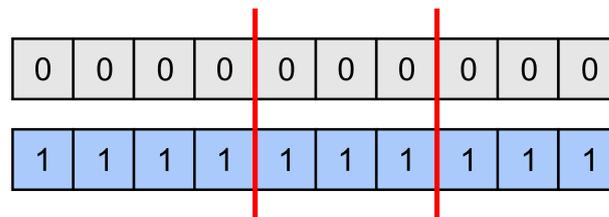


(a) Pais

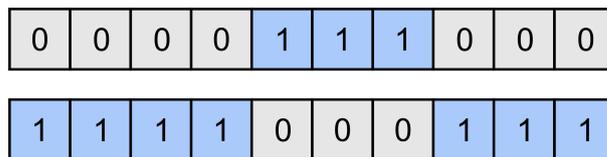


(b) Filhos

Figura 16 – Cruzamento com 1 ponto de corte. (a) pais com um ponto de corte aleatório selecionado e (b) filhos gerados pela combinação de fragmentos.



(a) Pais



(b) Filhos

Figura 17 – Cruzamento com 2 pontos de corte. (a) pais com 2 pontos de corte aleatórios selecionados e (b) filhos gerados pela combinação de fragmentos.

pai, enquanto as posições com valor 1 buscam o gene do cromossomo do segundo pai. O segundo filho é formado pelo complemento do primeiro (Figura 18(b)).

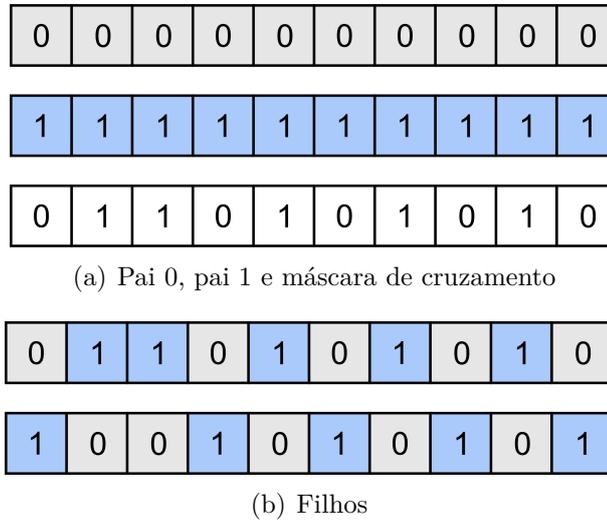


Figura 18 – Cruzamento uniforme. (a) pais e máscara de cruzamento e (b) filhos gerados a partir da máscara.

O cruzamento aritmético é feito com indivíduos representados por valores reais. Nele, dois pais geram um filho e o valor de cada posição do cromossomo filho, equivale à média aritmética dos elementos que estão na mesma posição dos pais (Figura 19).

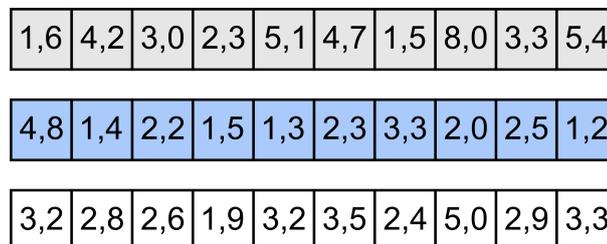


Figura 19 – Cruzamento aritmético.

3.6.2 Mutação

A Mutação é o operador genético responsável por causar uma perturbação no indivíduo gerado pela reprodução. Algumas características são alteradas de forma aleatória ou guiada de acordo com uma probabilidade Pm (GOLDBERG; HOLLAND, 1988), geralmente baixa em relação a Pr (probabilidade de reprodução), gerando um descendente mutante. Essa alteração faz com que o filho tenha alguma característica diferente daquelas encontradas nos pais, podendo ser melhor ou pior que elas, o que aumenta a diversidade da população fazendo com que ela possa visitar outros espaços de busca, saindo de mínimos locais. Quando essa perturbação é muito elevada, a convergência do AG fica potencialmente comprometida. O tipo mais simples de mutação é a binária (*Bit Flip Mutation*). A

Figura 20 apresenta essa mutação aplicada em indivíduos de representação binária, onde uma posição (locus) do cromossomo é sorteada e o valor binário dessa posição é invertido (0 passa a ser 1 ou 1 passa a ser 0) (HOLLAND et al., 1992).

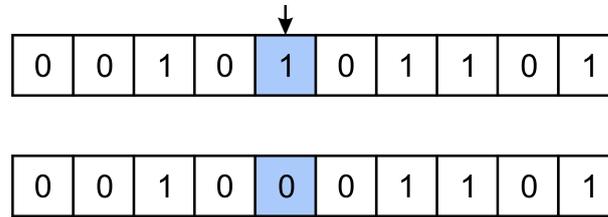


Figura 20 – Mutação binária ou *bit flip*.

A mutação baseada em permutação (Figura 21) opera trocando genes de lugar para manter as restrições da representação dos indivíduos. Nela, 2 ou mais posições são sorteadas e os valores dessas posições são trocados entre elas.

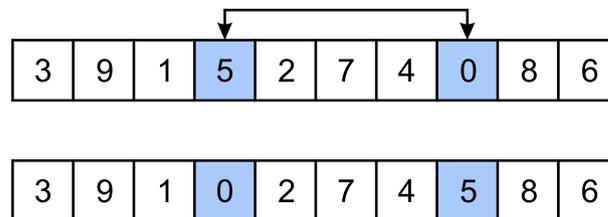


Figura 21 – Mutação por permutação.

Para representações de indivíduos com números inteiros ou reais, uma alternativa de mutação pode ser a soma de um valor positivo ou negativo ao gene sorteado. Na Figura 22 a quarta posição do indivíduo foi a sorteada e seu valor (2,3) recebeu um acréscimo de 0,5.

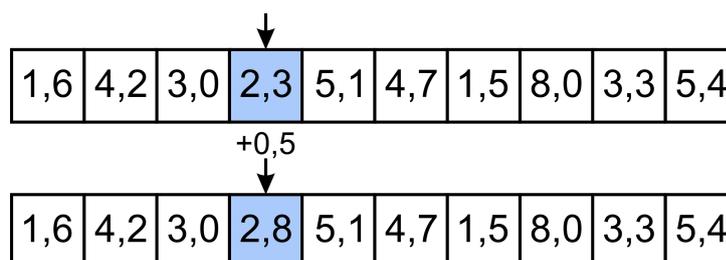


Figura 22 – Exemplo de mutação para representações numéricas.

A mutação ainda pode acontecer por indivíduo ou gene. A Figura 23 apresenta as diferenças, em que a cor azul sinaliza o gene que sofrerá mutação. Na mutação por indivíduo (Figura 23(a)), um gene ou um conjunto pequeno pré definido de genes são selecionados, aleatoriamente ou não, e podem sofrer mutação de acordo com uma probabilidade. Na mutação por gene (Figura 23(b)), todos os genes do cromossomo que representam o indivíduo são suscetíveis a mutação, mas a quantidade de genes que são efetivamente alterados depende da probabilidade de mutação. A mutação por gene pode causar uma perturbação bem maior que a mutação por indivíduo.

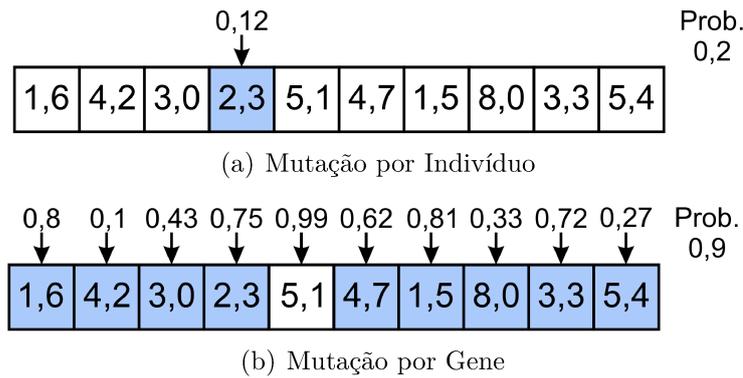


Figura 23 – Tipos de muta o: (a) indiv duo e (b) gene.

3.7 Reinscri o - Sobreviv ncia de Indiv duos

Ao final de uma gera o, ap s o AG passar por todos os operadores gen ticos, a popula o deve permanecer do mesmo tamanho. Para isso, deve-se escolher os indiv duos que v o permanecer na popula o, analisando aqueles que j  estava na popula o somados com os que foram gerados. A defini o dessa nova popula o, chamada reinscri o, pode acontecer de acordo com alguns crit rios. Na reinscri o pura os filhos substituem os pais sem considerar nenhum tipo de avalia o. Na reinscri o por elitismo, um n mero x de indiv duos mais aptos s o selecionados arbitrariamente para continuar na popula o. Na reinscri o por aptid o o conjunto (popula o + indiv duos gerados)   ordenado de acordo com a fun o objetivo e os melhores formam a nova popula o, enquanto na reinscri o uniforme um procedimento de sele o   aplicado para escolher os sobreviventes.

3.8 Crit rio de Parada

O crit rio de parada   uma condi o preestabelecida para finalizar a busca do AG. Os mais comuns s o:

- At ngir um n mero espec fico de gera es.
- At ngir determinado valor para a fun o objetivo (*fitness*).
- At ngir uma quantidade predefinida de gera es sem qualquer melhoria significativa na fun o objetivo.

Trabalhos Correlatos

Neste capítulo são apresentados os trabalhos relacionados ao modelo baseado em autômatos celulares que é discutido nessa dissertação e ao uso de algoritmos genéticos no ajuste de parâmetros de ACs. Inicialmente foi feito um levantamento bibliográfico detalhado do estado da arte em dois tópicos de pesquisa: o uso de Autômato Celulares (AC) na modelagem de vetores de doenças, principalmente doença de Chagas, e o uso de algoritmos evolutivos na mineração de parâmetros de modelos com ênfase em modelos de simulação baseados em Autômatos Celulares.

Uma revisão da aplicação de técnicas de inteligência artificial e computacional em sistemas de modelagem ambiental é apresentada em (CHEN; JAKEMAN; NORTON, 2008), cobrindo raciocínio baseado em casos, lógica fuzzy, redes neurais artificiais, algoritmos genéticos, sistemas multiagentes, autômatos celulares e inteligência de enxame (*Swarm Intelligence* (SI)). Para cada caso específico, uma técnica de inteligência artificial é adequada para a modelagem ambiental. Para processos complexos e/ou pouco compreendidos podem favorecer uma abordagem “caixa preta”, como Raciocínio Baseado em Casos (CBR), Redes Neurais Artificiais (RNA) ou AGs. As RNAs aprendem a partir de casos de treinamento e captura de relacionamentos entre os dados. Sistemas de inteligência de enxame (SI) envolvem simples agentes, resolvendo coletivamente um problema por interação local. Para problemas com processos bem compreendidos, o Sistema Baseado em Regras (RBS) pode ser aplicado, muito utilizado em diagnóstico de plantas e animais. Autômatos celulares (AC) são usados para entender e tentar prever comportamentos. Os Sistemas Multi-agentes (SMA) empregam agentes interagindo para resolver um problema comum, geralmente são adequados para o gerenciamento de recursos naturais, explorando estratégias de gerenciamento com as partes interessadas. Em contraste, Aprendizagem por Reforço (RL) modela a aprendizagem de tentativa e erro de agentes individuais. Lógica fuzzy pode lidar com dados imprecisos e incompletos, também por isso é frequentemente combinada com outras técnicas para formar sistemas híbridos.

4.1 Autômatos Celulares e Modelos Epidemiológicos

Em (WHITE; REY; SÁNCHEZ, 2007) é apresentado um modelo teórico para simular a propagação de uma epidemia baseado em autômatos celulares bidimensionais. Ele é um modelo do tipo Suscetíveis-Infectados-Recuperados (SIR), que divide a população em três classes: suscetível à doença, infectado e recuperado, e o estado de cada célula representa a porção dessas classes de indivíduos na célula em cada etapa do tempo. As simulações obtidas por meio dos parâmetros escolhidos artificialmente, foram de acordo com o esperado, considerando o comportamento de epidemias reais. O comportamento espacial de pandemias é investigado em (FU; MILNE, 2003). O modelo produz padrões de comportamento que imitam intuitivamente o observado na natureza. O trabalho não apresenta novas conclusões epidemiológicas porém investiga o comportamento espacial de doenças infecciosas que possuem apenas dados estatísticos.

MIKLER; VENKATACHALAM; ABBAS (2005) propõem um paradigma de Autômato Celular Estocástico Global (GSCA), que incorpora interações baseadas em dados geográficos e demográficos. O modelo GSCA é orientado para populações heterogêneas e pode incorporar interações baseadas em geografia, demografia, meio ambiente e padrões de migração. Esse modelo foi implementado para modelar a progressão de três doenças: resfriado comum, conjuntivite e gripe. Ele foi apresentado como alternativa para tratar as limitações dos modelos tradicionais de ACs determinísticos e SIR.

Um modelo de autômato celular estocástico para simular a propagação da dengue foi desenvolvido em (MEDEIROS et al., 2011). Populações humanas e de mosquitos foram monitoradas individualmente durante as simulações. Os resultados do modelo em estudo, associou a infestação de vetores à contagem de ovos de *Aedes Aegypti*, na cidade de Recife, Brasil.

Os padrões de mortalidade de árvores causados por insetos florestais são investigados em (BONE; DRAGICEVIC; ROBERTS, 2006). O modelo desenvolvido foi baseado em padrões de mortalidade de árvores causados por surtos de besouro do pinheiro da montanha (MPB), no interior central da Colúmbia Britânica, no Canadá. O modelo desenvolvido foi um AC baseado em GIS que incorporava a teoria dos conjuntos fuzzy para obter informações de imagens de Sensoriamento Remoto (SR) de alta resolução. Os padrões de mortalidade de árvores causados pelos insetos florestais fornecem um bom caso para AC, pois as interações locais levam a mudanças no nível da paisagem. Esse estudo enfatizou a utilidade de integrar a teoria dos conjuntos difusos na modelagem a fim de reconhecer e representar a incerteza espacial inerentes aos processos referentes a ela. Em (ORTIGOZA; BRAUER; LORANDI, 2019) modelos de autômatos celulares propostos para simular a propagação de doenças transmitidas por mosquitos foram revisados e classificados. Os modelos simulam Dengue, Zika, Mayaro, Chikungunya e Febre Amarela. Eles são classificados quanto a mobilidade humana, tipo de compartimento, simulações de cenários hipotéticos ou aqueles que usam dados reais para ajustar o modelo, uso de

sistemas de informação, densidades populacionais, sazonalidade e estratégias de controle para prevenir a propagação de doenças.

4.2 Algoritmos Genéticos na configuração de parâmetros de Autômatos Celulares

Algoritmos evolutivos têm sido amplamente usados para desenvolver regras e comportamentos de autômatos celulares. Packard (PACKARD, 1988) propôs o emprego de AGs para desenvolver regras de ACs para realizar uma tarefa computacional denominada Tarefa de Classificação de Densidade (DCT). Em (MITCHELL; CRUTCHFIELD; HRABER, 1994) essa ideia foi melhor explorada e também utilizaram algoritmos genéticos para buscar regras de ACs capazes de executar a tarefa de classificação da densidade. Foi feita uma análise detalhada sobre os mecanismos evolutivos que produzem o comportamento do AG nessa tarefa e os impedimentos encontrados por ele. Diversos trabalhos continuam a investigar o DCT por meio da pesquisa evolutiva (MITCHELL; CRUTCHFIELD; HRABER, 1994; JUILLE; POLLACK, 1998; OLIVEIRA; BORTOT; OLIVEIRA, 2002). Esta ideia foi adaptada a outros desafios computacionais como tarefa de sincronização (OLIVEIRA; OLIVEIRA; OMAR, 2001) e problema de tabuleiro de damas (BREUKELAAR; BÄCK, 2005). Nesse último, é proposta uma abordagem para gerar regras de transição para ACs usando algoritmos genéticos que evoluem o comportamento global com regras locais, assim como na natureza. Três experimentos são realizados usando topologias de ACs multidimensionais e os resultados mostram um grande potencial da ferramenta para resolver problemas do mundo real. A busca evolutiva também foi utilizada para modelos de ACs aplicados a problemas clássicos da computação, tais como, escalonamento (SEREDYNSKI; ZOMAYA, 2002; CARNEIRO; OLIVEIRA, 2013) e criptografia (OLIVEIRA et al., 2010; SZABAN; SEREDYNSKI; BOUVRY, 2006).

A busca evolutiva também foi aplicada na configuração de modelos físicos e biológicos baseados em regras de autômatos celulares. D'Ambrosio et al. (2007) utilizam algoritmos genéticos paralelos para a calibração de um modelo de deslizamento de terras baseado em autômatos celulares. A calibração foi feita utilizando os dados de maio de 1998, em Campania, na Itália depois de um período de chuva abundante. O processo de avaliação de um AG, é a etapa que mais demanda recursos computacionais, por isso, se for feita de forma paralela, pode agilizar a execução. Portanto, foi investigado o paralelismo em hardware para a implementação desse modelo. A função de aptidão compara a extensão da área da simulação com a real. O artigo conclui que o modelo apresenta uma melhoria em relação ao sua versão anteriormente publicada, porém ainda necessita de ajustes. A calibração ocorreu de forma satisfatória para aquele estudo de caso específico. Algoritmos genéticos paralelos também são usados para calibração de modelos de processos geológicos tais como fluxo de detritos e lava em (D'Ambrosio; SPATARO, 2007). A calibração

mostrou melhoria no desempenho do modelo, porém não foi possível fazer a validação porque não conseguiram dados reais para comparação.

O uso de autômatos celulares na modelagem do desenvolvimento de uma cobertura vegetal após incêndios é apresentado em (GARCÍA-DURO et al., 2018). Vários elementos precisam ser considerados como a quantidade de plantas de diferentes espécies, a interação entre as espécies, os processos envolvidos no desenvolvimento de cada uma delas, como dispersão de sementes, etc. Alguns fatores do processo foram simplificados devido ao alto custo computacional das simulações. O modelo apresenta um elevado número de parâmetros e um reticulado de grandes dimensões, o que faz que cada incremento de tempo do AC e alterações em seu tamanho/regras, tenha impacto exponencial no custo computacional. Para auxiliar neste processo, os autores propõem uma abordagem de otimização evolutiva de dois níveis para reduzir esse custo, que não foi implementada. Inicialmente, um método baseado em programação genética é usado para definir funções m para a taxa de propagação de cada espécie sem considerar a competição. Em seguida, um AG é usado para escolher a melhor combinação das funções das espécies, a fim de simular de forma mais realista os dados concorrentes.

Finalmente, os algoritmos genéticos também foram aplicados para configurar modelos robóticos baseados em regras de autômatos celulares (LAL; YAMADA; ENDO, 2006; OLIVEIRA et al., 2018).

Modelo Epidemiológico

Nesse capítulo é apresentado o modelo utilizado como estudo de caso, descrito em (SLIMI et al., 2009), que trata da dispersão do vetor barbeiro com o objetivo do controle da doença de Chagas. O passo a passo de seu processo de modelagem é detalhado e também seus resultados. Na sequência apresentamos a nossa representação desse modelo com detalhamento das adaptações e a comparação dos resultados obtidos.

5.1 Estudo de Caso

Um dos principais problemas de saúde pública que afeta a América Latina é a doença de Chagas (CISSÉ; YACOUBI; GOURBIERE, 2014). A doença de Chagas, também conhecida como tripanossomíase americana foi descrita em 1909 pelo brasileiro Carlos Chagas (CHAGAS, 1909). Ela é uma infecção parasitária crônica, sistêmica, potencialmente fatal, causada pelo protozoário *Trypanosoma cruzi*. Sua transmissão se dá principalmente por insetos vetores hematófagos como o barbeiro. Estima-se que 8 milhões de pessoas em todo o mundo estejam infectadas, das quais 30 a 40% têm ou desenvolverão cardiomiopatia, síndromes digestivas, ou ambos (JR; RASSI; MARIN-NETO, 2010). Não existem vacinas para a doença nem tratamento que permita a recuperação após a infecção, portanto, o método de prevenção mais eficaz é o controle do vetor.

O modelo de autômato celular replicado como estudo de caso foi apresentado em (SLIMI et al., 2009). Esse modelo possui uma temática de suma relevância e simula a evolução de uma população de insetos vetores responsáveis pela doença de Chagas com base em regras de autômatos celulares probabilísticos, representando o comportamento espaço-temporal da dinâmica populacional do vetor da doença na escala de uma aldeia. Ele foi investigado de acordo com informações coletadas sobre o fenômeno real, na península de Yucatan, no México. Nessa área, a pulverização de inseticidas demonstrou baixa eficácia e os vetores são muito pouco ou não domesticados, logo, a ocorrência deles no ambiente doméstico depende de uma forte imigração sazonal do habitat florestal. Os dados disponíveis foram usados para validação do modelo e estimativa de parâmetros.

Trata-se de um modelo bidimensional e probabilístico de AC, quadrado, de tamanho 30×30 células que representam uma área de cerca de 1×1 quilômetro quadrado ($33 \times 33 m^2$ por célula) referentes à aldeia em que o estudo foi conduzido. Cada célula pode estar vazia ou ocupada por insetos adultos e/ou larvas. O estado de cada célula é representado por uma dupla de valores, que indicam o número de larvas e de adultos presentes na célula, respectivamente (n_y, n_a) . Esse número de insetos por célula é limitado em 5 adultos e 5 larvas. A célula também possui uma característica de habitat, denotada por G e H , correspondente a jardim e casa. Essa característica de habitat interfere principalmente nas probabilidades, que são parâmetros do modelo. Na implementação adotada, todas as células são do tipo H - casa, portanto as probabilidades serão as mesmas para todas as células. A regra de transição é definida por dois processos: demografia e dispersão.

5.1.1 Demografia

A demografia é o processo responsável por definir a quantidade de insetos no reticulado. Ela é dividida em 3 etapas, reprodução, sobrevivência e desenvolvimento.

5.1.1.1 Reprodução

Um inseto adulto se reproduz com uma probabilidade (p_r) gerando F larvas. Neste trabalho, assim como no artigo original, foi adotado $F=1$. O número de larvas geradas (n_{ry}) é descrito por uma variável aleatória após uma distribuição binomial $B(Fn_a, p_r)$. Após a reprodução, o estado da célula é:

$$(n_y + n_{ry}, n_a) \quad (5)$$

5.1.1.2 Sobrevivência

Descreve a quantidade de larvas e adultos sobreviventes $(n_{sy}$ e $n_{sa})$, no próximo instante de tempo, por célula, guiada pelas probabilidades de sobrevivência p_{sy} (larvas) e p_{sa} (adultos). A variável aleatória que descreve essa quantidade de indivíduos sobreviventes se dá por meio de uma distribuição binomial da forma $B(n, p_s)$. O estado da célula após a sobrevivência é:

$$(n_{ry} + n_{sy}, n_{sa}) \quad (6)$$

5.1.1.3 Desenvolvimento

Cada larva pode se tornar adulta no próximo instante de tempo de acordo com uma probabilidade (p_d) . Dentre as larvas sobreviventes n_{sy} , a quantidade delas que se trans-

formarão em adultas é definida por uma função binomial da forma $B(n_{sy}, p_d)$. O estado da célula após o desenvolvimento é:

$$(n_{ry} + n_{sy} - n_{da}, n_{sa} + n_{da}) \quad (7)$$

No final da Demografia, o estado da célula é dado por:

$$(n_{ry} + n_{sy} - n_{da}, n_{sa} + n_{da}) \quad (8)$$

5.1.2 Dispersão

A dispersão é o processo responsável pela movimentação dos insetos no reticulado. Somente adultos podem se dispersar e as regras de transição representam a entrada (n_{in}) e a saída (n_{out}) de insetos na célula, de acordo com uma probabilidade de dispersão p . O modelo possui um período de infestação e um período de não infestação. O período de infestação dura 3 meses (de abril a junho) e nesse período acontece a migração de insetos da floresta para a aldeia. Logo, presume-se que célula da floresta esteja ocupada por Nf insetos que definem a condição de fronteira. Como os insetos podem vir de qualquer célula da borda, consideramos sempre $Nf = 1$. Os insetos adultos podem migrar para a aldeia com a probabilidade

$$D_f = \frac{Qf}{4M} \quad (9)$$

onde D_f é o coeficiente de dispersão nas células da floresta, Qf é o número de insetos que chegam na aldeia por dia e M a dimensão do reticulado. No período de não infestação, não acontece a migração de insetos da floresta para a aldeia.

As regras de transição da dispersão se referem à entrada e saída de insetos na célula. A movimentação dos insetos presentes no reticulado acontece considerando a vizinhança de Moore. Cada adulto presente na rede tem a possibilidade de se dispersar de sua célula de origem com uma probabilidade p , calculada usando o coeficiente de dispersão D e o raio de infestação r a cada passo do tempo e cada célula também pode receber insetos adultos vindos de sua vizinhança.

1. Cada adulto pode deixar a célula e escolher aleatoriamente uma da vizinhança com probabilidade:

$$p = \frac{D}{(2r + 1)^2 - 1} \quad (10)$$

2. Cada adulto pode entrar em uma célula vindo de qualquer célula da vizinhança com a mesma probabilidade.

O coeficiente de dispersão e o raio de infestação possuem valores diferentes de acordo com o período apresentados na Tabela 2:

Tabela 2 – Valores dos coeficientes e raios nos períodos de infestação e não infestação.

| Período | Coeficiente | Raio |
|----------------|----------------|--------------|
| Infestação | $D_i = 0.9$ | $r_i = 3$ |
| Não Infestação | $D_{ni} = 0.1$ | $r_{ni} = 1$ |

Finalmente, o estado da célula após a dispersão é:

$$(n_y, n_a - n_{out} + n_{in}) \quad (11)$$

5.1.3 Análise

Para analisar o número total de insetos (Na - adultos e Ny - larvas), vamos considerar toda a aldeia como uma única célula mantendo as dinâmicas locais do reticulado e as interações com a floresta, da seguinte forma:

$$Na = \sum_{c \in \mathcal{L}} n_a(c) \quad Ny = \sum_{c \in \mathcal{L}} n_y(c) \quad (12)$$

Onde o total é composto pelo somatório de larvas (n_y) e adultos (n_a) presentes em cada célula (c) do reticulado \mathcal{L} . Após o processo demográfico, o número total de larvas e adultos é definido por:

$$Ny^{(t+1)} = p_{sy}(1 - p_d)Ny^{(t)} + p_rNa^{(t)} \quad (13)$$

$$Na^{(t+1)} = p_dp_{sy}Ny^{(t)} + p_{sa}Na^{(t)} \quad (14)$$

A equação 14, mostra que a quantidade de larvas é composta daquelas que sobrevivem e não viram adultas acrescida daquelas geradas pela reprodução dos adultos ($p_rNa^{(t)}$). Mostra também que a quantidade de adultos é composta pelos adultos que sobrevivem acrescidos das larvas que se tornaram adultas ($p_dp_{sy}Ny^{(t)}$). Após o processo de dispersão, as quantidades são definidas por:

$$Ny^{(t+1)} = p_{sy}(1 - p_d)Ny^{(t)} + p_rNa^{(t)} \quad (15)$$

$$Na^{(t+1)} = p_dp_{sy}Ny^{(t)} + p_{sa}Na^{(t)} + Qf - Qout \quad (16)$$

Na equação 16, o número de larvas não sofre alteração porque somente os adultos se dispersam. Já a quantidade de adultos recebe a interferência dos insetos que entram no reticulado (Qf) e dos insetos que saem do reticulado ($Qout$) no processo de dispersão. Vale ressaltar que durante o período de não infestação, os insetos não migram da floresta para a aldeia, logo $Qf = 0$.

De acordo com as informações disponíveis em (SLIMI et al., 2009), para modelar a dinâmica populacional observada a partir dos dados coletados, os parâmetros do modelo receberam os seguintes valores:

- (i) probabilidade de reprodução (p_r) = 0,004111;
- (ii) probabilidade de sobrevivência das larvas (p_{sy}) = 0,90272518;
- (iii) probabilidade de sobrevivência de insetos (p_{sa}) = 0,9828095;
- (iv) probabilidade de desenvolvimento (p_d) = 0,004158;
- (v) média do número de insetos que chegam à aldeia por dia (Q_f) = 50.

A evolução do espalhamento dos insetos adultos no reticulado (distribuição espacial) é apresentada na Figura 24. Essa figura foi retirada de (SLIMI et al., 2009) e mostra os reticulados referentes à configuração inicial e aos 5º, 90º, 365º dias de simulação do AC, usando os valores de referência originais (com $Q_f = 50$). Os insetos vetores têm um comportamento cíclico anual, onde sua quantidade aumenta no período de infestação, devido a migração sazonal dos adultos e decai durante o período de não infestação, onde eles se dispersam gradativamente da aldeia até o início de um novo ciclo. A configuração inicial do AC é um reticulado vazio que representa a aldeia, com células da cor verde que indicam ausência de insetos, enquanto as bordas vermelhas representam os insetos presentes na floresta ao redor da aldeia. O modelo se inicia no período de infestação. No 5º dia podemos observar os insetos começando a infestar a aldeia representados pela cor azul, inicialmente povoando as células da borda e se espalhando localmente, destacamos que o centro da aldeia nesse momento está ausente de insetos. No 90º dia acontece o pico da infestação. Os insetos estão presentes em maior quantidade, espalhados por toda a aldeia. Esse é o último dia em que a borda vermelha se faz presente, indicando o fim do período de infestação. No 365º dia, o reticulado se encontra novamente vazio, assim como na configuração inicial. Os insetos passaram pelos processos de demografia e dispersão, deixando o reticulado gradativamente até que ele ficasse vazio.

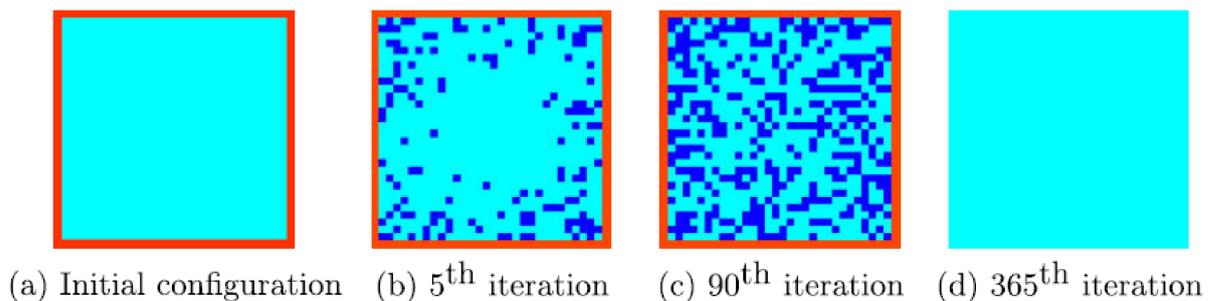


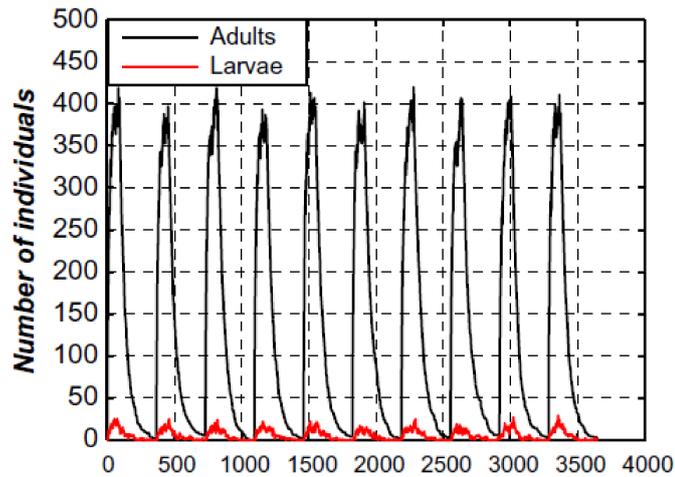
Figura 24 – Evolução do espalhamento dos insetos adultos simulada pelo AC a partir dos valores de referência do artigo original (retirada de (SLIMI et al., 2009)).

5.2 Reprodução do Modelo

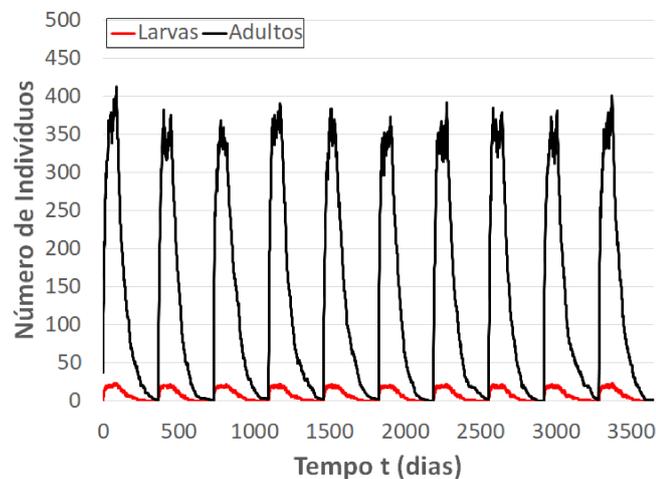
Inicialmente, o modelo baseado em AC descrito em (SLIMI et al., 2009) foi reproduzido, implementado e executado usando os mesmos valores de parâmetros apresentados na Seção 5.1 com o objetivo de criar dados de referência, pois não dispomos da base de dados utilizada para a comparação de resultados. As simulações obtiveram um comportamento cíclico de vetores semelhante ao modelo original, porém, as quantidades diárias e a distribuição espacial dos insetos observadas nos reticulados do AC foram ligeiramente diferentes dos números apresentados em (SLIMI et al., 2009). Com base em experimentos empíricos, o número médio de insetos que chegam na aldeia por dia (Qf) foi modificado para 95 a fim de minimizar essas diferenças e reproduzir com mais precisão a dinâmica populacional apresentada em (SLIMI et al., 2009). Os demais parâmetros foram mantidos com os mesmos valores usados no modelo original, como a seguir: $p_r = 0,004111$; $p_{sy} = 0,90272518$; $p_{sa} = 0,9828095$; $p_d = 0,004158$; $D_i = 0,9$; $D_{ni} = 0,1$; e $Qf = 95$.

A Figura 25 apresenta a dinâmica populacional dos vetores (larvas e adultos) na área da aldeia ao longo de 10 anos (3650 dias). A Figura 25(a) mostra a evolução temporal dos insetos apresentados em (SLIMI et al., 2009), enquanto os dados obtidos usando nosso modelo de referência são apresentados na Figura 25(b). Pode-se notar que ambos os modelos apresentam uma dinâmica periódica com ciclos anuais na evolução da população de larvas e adultos, e uma variação no número de insetos entre os períodos de infestação e não infestação. Como esperado, no período de infestação ocorre um aumento considerável no número de vetores, atingindo as maiores quantidades próximas ao 90º dia. Por outro lado, ocorre a dispersão dos insetos da área urbana durante o período de não infestação, retornando ao estado inicial da rede AC (sem insetos) próximo ao 365º dia. Pode-se observar que ambos os modelos resultaram em dinâmicas populacionais semelhantes. Apesar dos picos ligeiramente mais altos obtidos no modelo de referência, a ordem de magnitude e as formas das curvas para larvas e adultos são muito semelhantes.

A Figura 26 corrobora a similaridade dinâmica entre os modelos. Mostra reticulados em instantes de tempo que representam a distribuição espacial de insetos adultos. Estes foram obtidos por meio da simulação dos modelos de AC, considerando tanto a temporada de infestação (5º e 90º dias) quanto a temporada de não infestação (180º e 270º dias). Os mesmos dias são apresentados em (SLIMI et al., 2009) para viabilizar a comparação. Ao 5º dia, os insetos estão mais concentrados nas bordas da rede, pois é o início da migração sazonal da floresta para a aldeia. No 90º dia (pico de infestação), a área da vila é densamente povoada por vetores da doença. Embora ainda dispersos pela aldeia, o número de adultos diminui consideravelmente no 180º dia e praticamente termina em 270º dia. Na Figura 26(a), células brancas indicam ausência de insetos, células cinzas apenas um adulto e células pretas 2 ou mais adultos. A borda cinza no 5º e 90º dias representa a presença de adultos na área da floresta. Ela desaparece durante o período de não infestação. Na Figura 26(b), as células azuis indicam a ausência de insetos, enquanto



(a) Modelo Original (retirado de (SLIMI et al., 2009))



(b) Modelo de Referência

Figura 25 – Dinâmica populacional dos insetos ao longo de 10 anos simulados pelo autômato celular, de acordo com o modelo adotado.

as verdes, amarelas, laranja, magenta e vermelhas indicam a presença de 1, 2, 3, 4 e 5 adultos, respectivamente. Ao comparar os reticulados, o modelo de referência resulta em uma quantidade ligeiramente maior de insetos. No entanto, a distribuição espacial dos adultos dentro da aldeia é muito semelhante, indicando que nosso modelo foi capaz de se reproduzir satisfatoriamente o modelo original descrito em (SLIMI et al., 2009).

Na Figura 27 temos o comportamento das larvas obtidos nas simulações do modelo de referência, nos mesmos instantes de tempo que os reticulados dos adultos. As cores das células seguem a mesma legenda da Figura 26(b). Para as larvas, não temos a distribuição espacial no modelo original (SLIMI et al., 2009) para comparação. Sabemos que as quantidades e o comportamento cíclico são semelhantes baseado no gráfico de quantidades da Figura 25. Ao 5º dia, os insetos estão mais presentes na região de borda. No 90º dia (pico de infestação), as larvas estão mais distribuídas pela vila porém em regiões de maior concentração. No 180º dia o número de larvas diminui e já se esgotou no 270º dia.

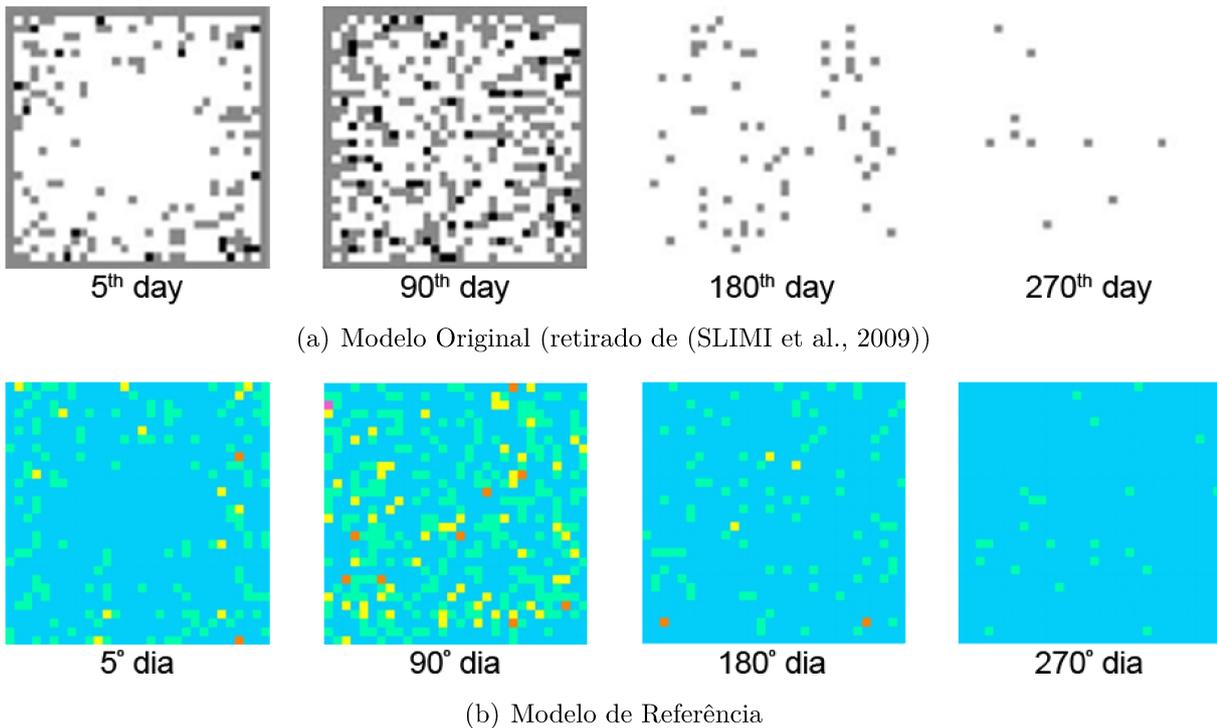


Figura 26 – Distribuição espacial dos adultos nos 5^o, 90^o, 180^o e 270^o dias, de acordo com a simulação do modelo adotado.

A presença das larvas nas células é totalmente dependente da reprodução dos adultos, já que elas não se dispersam, por isso sua distribuição no reticulado vai acompanhar o espalhamento dos adultos.

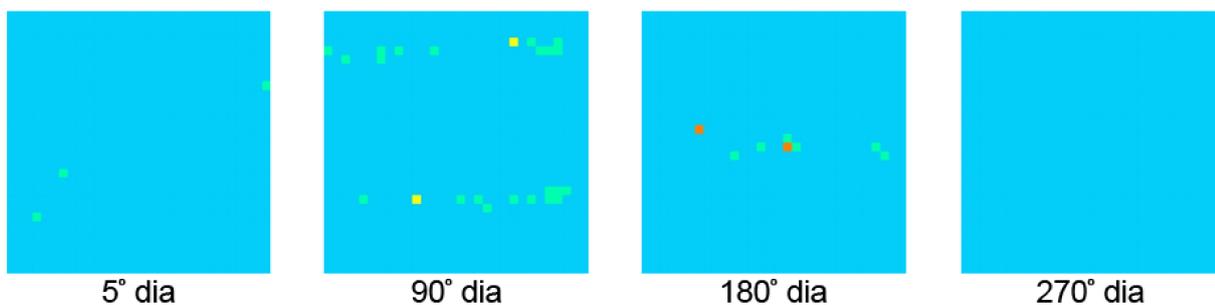


Figura 27 – Distribuição espacial das larvas nos 5^o, 90^o, 180^o e 270^o dias, de acordo com a simulação do modelo de referência.

O modelo de referência foi utilizado para gerar a série de dados artificiais (dados de referência) que utilizaremos como base de dados de comparação. Essa série de dados é nossa reprodução dos dados coletados em campo e é composta pelas quantidades de larvas e adultos presentes no reticulado, dia por dia, durante um ciclo (1 ano - 365 dias). A quantidade de insetos é definida pela média de 1000 execuções do modelo baseado em AC, com os valores de probabilidades, coeficientes e raios descritos em (SLIMI et al., 2009), ou seja, executamos o modelo 1000 vezes e calculamos a média dia a dia, de larvas e adultos, para os 365 dias do ciclo.

Estratégias Evolutivas Propostas

O primeiro desafio do trabalho é escolher um modelo para ser implementado como estudo de caso inicial. Após um levantamento, o modelo escolhido foi o descrito no capítulo 5 baseado em (SLIMI et al., 2009), que trata da dispersão do vetor barbeiro, com o objetivo do controle da doença de Chagas. Esse modelo foi escolhido porque possui uma temática de suma relevância, o artigo apresenta com detalhes (passo a passo) o processo de modelagem e seus resultados. Esse modelo tem complexidade relevante pois envolve diversos parâmetros, que retratam: as probabilidades de reprodução, desenvolvimento e sobrevivência do inseto; a quantidade de insetos que chegam na área por dia; em que período do ano eles migram; por quanto tempo dura a migração; a probabilidade de dispersão dos insetos adultos; o tamanho do raio da vizinhança do Autômato Celular por período de infestação/não infestação (para configurar a dispersão); dentre outros.

Após a implementação do AC que representa o modelo, foi feito um estudo e implementação de um módulo de mineração baseado em Algoritmo Genético (AG) que objetiva buscar uma configuração adequada de parâmetros para modelos epidemiológicos baseados em regras estocásticas de AC. Como o estudo de caso se trata de um modelo já ajustado e definido, a base de comparação já está delimitada, facilitando a avaliação dos resultados encontrados pelo AG. O objetivo principal é obter resultados de simulação para a população de vetores (larvas e adultos) exibindo dinâmica próxima à observada em simulações utilizando os parâmetros definidos no modelo de referência (SLIMI et al., 2009). O principal desafio dessa etapa é definir a função de aptidão usada para avaliar os indivíduos do Algoritmo Genético. A simulação do modelo em AC retorna as quantidades de larvas e adultos em cada célula do reticulado (e o total de jovens/adultos/dia) por dia em um período de um ano. Como se trata de uma modelagem com probabilidades, sua simulação não é determinística, ou seja, seus resultados variam entre as execuções com um mesmo conjunto de parâmetros. Então, avaliar a proximidade desses resultados (tanto nas quantidades quanto na distribuição espacial) constitui um desafio para o cálculo da aptidão.

Neste capítulo são descritas, de forma mais detalhada, as estratégias empregadas na

configuração dos parâmetros do modelo epidemiológico baseado em autômato celular reproduzido a partir de (SLIMI et al., 2009).

6.1 Algoritmo Genético Simples

A primeira versão do método de configuração dos parâmetros para o modelo epidemiológico consiste de um algoritmo genético típico, o qual foi chamado *Plain Genetic Algorithm* (PGA). O fluxograma de execução do PGA é o mesmo apresentado na Figura 12, no capítulo 3. Em nossa implementação, o AG manipula uma população de 100 indivíduos ao longo de 100 gerações, sendo que, a cada geração, todo o cromossomo do indivíduo é evoluído, ou seja, todos os parâmetros (genes) de uma solução podem ser afetados nas operações genéticas. Essa versão é considerada mono-estágio, enquanto a versão apresentada na seção 6.2 é multi-estágio, uma vez que diferentes sub-conjuntos de parâmetros são ajustados em diferentes estágios da evolução..

6.1.1 Representação do Indivíduo

O indivíduo é uma possível configuração do modelo e é representado por um vetor com os sete valores dos parâmetros explorados, na seguinte ordem:

- (i) probabilidade de reprodução (p_r);
- (ii) probabilidade de sobrevivência das larvas (p_{sy});
- (iii) probabilidade de sobrevivência de adultos (p_{sa});
- (iv) probabilidade de desenvolvimento (p_d);
- (v) coeficiente de dispersão durante o período de infestação (D_i);
- (vi) coeficiente de dispersão durante o período de não infestação (D_{ni}); e
- (vii) média do número de insetos que chegam à aldeia por dia (Q_f).

A Figura 28 descreve a representação de um indivíduo modelo, com os valores considerados como referência, baseados na nossa representação de (SLIMI et al., 2009).

Os primeiros seis parâmetros podem armazenar qualquer valor ponto flutuante (número real) no intervalo $[0, 0; 1, 0]$. O último pode assumir um valor inteiro entre $[5; 100]$. Na geração da população inicial, aplicamos uma distribuição uniforme considerando 3 faixas de valores para os parâmetros reais $[0, 0; 0, 33]$, $[0, 34; 0, 66]$ e $[0, 67; 1, 00]$. Essa estratégia foi utilizada para proporcionar uma maior diversidade da população, pois gera indivíduos com combinações de valores de parâmetros (alto, médio e baixo). Os 100 indivíduos gerados para a população inicial são avaliados e classificados de acordo com sua aptidão.

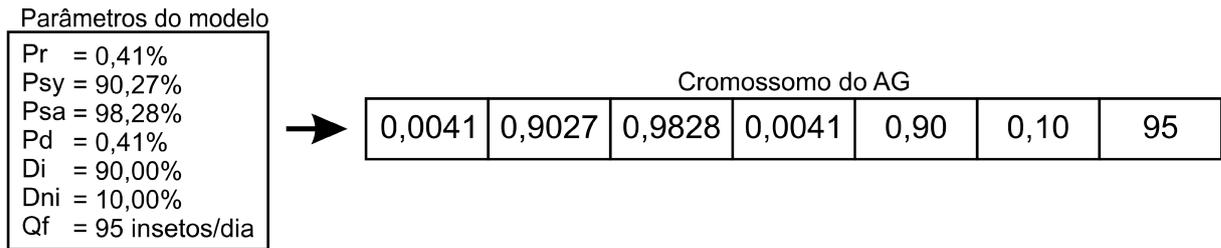


Figura 28 – Representação de um indivíduo da população.

6.1.2 Função Objetivo - *Fitness*

A seguir são apresentadas duas abordagens de avaliação utilizadas nesse trabalho, uma considerando os valores absolutos de insetos e outra levando em consideração a posição de espalhamento.

6.1.2.1 Diferença Relativa

O fitness de um indivíduo é calculado considerando o número médio diário de insetos resultante das várias execuções de cada modelo (referência e avaliado). Para o modelo de referência, este valor médio é calculado apenas uma vez usando 1.000 execuções do AC usando a configuração de parâmetros padrão da representação de (SLIMI et al., 2009). Para o modelo avaliado, o valor médio é calculado a partir de 10 execuções do AC usando a configuração de parâmetros fornecida pelo indivíduo do AG que está sendo avaliado. Experimentos de sensibilidade foram realizados para escolher o número de execuções, considerando o ponderamento entre o tempo de execução do AG e o desempenho do modelo AC. A função de aptidão de i -ésimo indivíduo (Fit_i) é a soma das diferenças entre o número de larvas (Y) e adultos (A) encontrados em todo o reticulado gerado por dia (considerando 365 dias) para o modelo AC ajustado com os parâmetros avaliados (AG individual) e aquele que usa a configuração de referência (os valores dos parâmetros baseados na representação de (SLIMI et al., 2009)). A soma da diferença no número total de adultos e larvas para cada dia é ponderada pela quantidade máxima diária de insetos do reticulado observada na base de referência (Y_{max} e A_{max}), como segue:

$$Fit_i = \frac{\sum_1^{365} |Y_{ref} - Y_{ga}|}{Y_{max}} + \frac{\sum_1^{365} |A_{ref} - A_{ga}|}{A_{max}} \quad (17)$$

6.1.2.2 Avaliação Regional

Na avaliação regional, a análise da quantidade de insetos leva em conta a região do reticulado em que eles se encontram. A Figura 29 ilustra as divisões que realizamos no reticulado para identificar três regiões diferentes. A cor verde destaca a região com maior quantidade de células. Ela se concentra nas bordas do reticulado, onde se inicia a migração de insetos. A cor vermelha destaca a região intermediária do reticulado enquanto

as células em azul destacam o centro, região com menor número de células. Essa divisão de regiões foi configurada empiricamente, repetindo a proporção de um terço de células em cada região (linhas e colunas) partindo do centro do reticulado.

O fitness é calculado a partir da comparação entre o modelo de referência e o avaliado, com uma média de 10 avaliações, como apresentado na seção 6.1.2.1. A função de aptidão de i -ésimo indivíduo ($Fit(Regional)_i$) é a soma das diferenças entre o número de larvas (Y) e adultos (A), por região do reticulado dia a dia (considerando 365 dias) para o modelo AC ajustado com os parâmetros avaliados (AG individual) e aquele que usa a configuração de referência (os valores dos parâmetros baseados na representação de (SLIMI et al., 2009)).

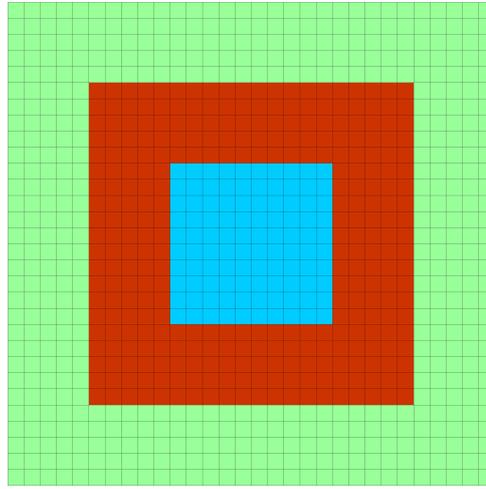


Figura 29 – Distribuição espacial adotada na avaliação regional.

A soma da diferença no número total de larvas e adultos para cada dia é dada pelas equações 18 e 19, onde YB e AB são os insetos presentes na região de borda, YI e AI presentes na região intermediária e YC e AC os presentes na região central. Esse somatório é ponderado pela quantidade máxima diária de insetos da rede (Y_{max} e A_{max}), como segue na equação 20:

$$Dif_{(larvas)} = \sum_1^{365} \left(\frac{|YB_{ref} - YB_{ga}|}{YB_{max}} \right) + \sum_1^{365} \left(\frac{|YI_{ref} - YI_{ga}|}{YI_{max}} \right) + \sum_1^{365} \left(\frac{|YC_{ref} - YC_{ga}|}{YC_{max}} \right) \quad (18)$$

$$Dif_{(adultos)} = \sum_1^{365} \left(\frac{|AB_{ref} - AB_{ga}|}{AB_{max}} \right) + \sum_1^{365} \left(\frac{|AI_{ref} - AI_{ga}|}{AI_{max}} \right) + \sum_1^{365} \left(\frac{|AC_{ref} - AC_{ga}|}{AC_{max}} \right) \quad (19)$$

$$Fit(Regional)_i = Dif_{(larvas)} + Dif_{(adultos)} \quad (20)$$

6.1.3 Seleção

A seleção de pais é feita utilizando o torneio simples ($tour = 3$), ou seja, 3 indivíduos da população são sorteados e aquele que possui melhor (menor) valor de fitness é selecionado como pai. A seleção é feita uma vez para cada um dos pais.

6.1.4 Operadores Genéticos

Em cada geração, um operador de crossover de um ponto de corte foi definido com 90% de probabilidade para sua aplicação a um par de indivíduos (pais). A mutação é feita por gene, com taxa de 15% aplicada aos filhos. Para cada gene mutado, o parâmetro é alterado somando um valor aleatório positivo ou negativo. Para os parâmetros ponto flutuantes, o valor pode variar entre $[-0, 2; 0, 2]$, e para o parâmetro inteiro, o valor varia entre 5 ou -5 .

6.1.5 Reinservação

A reinservação é feita por elitismo, onde ordenamos a população total (pais + filhos) e mantemos os 100 (tamanho da população) melhores indivíduos para a próxima geração.

6.1.6 Avaliação Robusta

Na última geração, os 10 indivíduos principais da população final são reavaliados. Nesse processo, o modelo de CA gerado utilizando cada configuração (individual) é submetido a uma avaliação mais robusta (100 simulações ao invés de 10) e o fitness é a soma das diferenças entre o número médio de adultos (\bar{A}) e larvas (\bar{Y}) do reticulado por dia em relação aos valores do modelo de referência. Esta nova avaliação beneficia em certa medida a configuração dos insetos adultos, pois eles têm uma escala muito maior do que as larvas:

$$Fit_i = \sum_1^{365} \frac{|\bar{Y}_{ref} - \bar{Y}_{ga}| + |\bar{A}_{ref} - \bar{A}_{ga}|}{2} \quad (21)$$

6.1.7 Divisão do Modelo

Os parâmetros do modelo a serem ajustados são divididos em três subconjuntos:

- S_{disp} : subconjunto formado pelos 3 parâmetros (Cdi , $Cdni$ e Qf),
- S_{demo} : subconjunto formado pelos 4 parâmetros ($Prep$, PSY , PSA e PD),
- S_{full} : subconjunto formado pelo conjunto completo (todos os 7 parâmetros).

Enquanto os parâmetros S_{demo} estão relacionados às probabilidades envolvidas na fase demográfica do modelo de CA, os parâmetros S_{disp} estão relacionados à fase de dispersão.

A estratégia PGA manipula todos esses subconjuntos, onde, o PGA-DISP configura os parâmetros de S_{disp} ; o PGA-DEMO configura os parâmetros de S_{demo} e o PGA-FULL: configura os parâmetros de S_{full} .

6.2 MSES - Estratégia Evolutiva Multi-Estágio

A estratégia evolutiva multi-estágio (*Multi-Stage Evolutionary Strategy* (MSES)) busca refinar a configuração dos parâmetros do modelo baseado em AC a partir da adoção de diferentes estágios de evolução. A ideia é baseada no conceito de dividir para conquistar, em que parâmetros do modelo a serem ajustados continuam utilizando a divisão dos três subconjuntos (S_{disp} , S_{demo} e S_{full} - Seção 6.1.7) na configuração de seus estágios, onde em cada um deles, um dos três subconjuntos é escolhido para ser evoluído. Além disso, o subconjunto escolhido varia entre os estágios.

Foram testadas estratégias multi-estágio utilizando 3 e 6 estágios com e sem repovoamento da população, sendo todas com a configuração de tamanho da população 100, taxa de cruzamento 0,9 e taxa de mutação 0,15. O conjunto S_{full} é ajustado nas 50 primeiras gerações de todas as configurações do MSES investigadas nesse trabalho. A partir dela utilizamos algumas configurações de congelamento de parâmetros, em que um subconjunto é ajustado enquanto outro permanece estático. O número de gerações utilizadas nas próximas etapas varia de acordo com a estratégia adotada, mas respeita a seguinte ordem:

1. evolução do subconjunto S_{demo} (com os 4 parâmetros da demografia),
2. evolução do subconjunto S_{disp} (com os 3 parâmetros de dispersão).

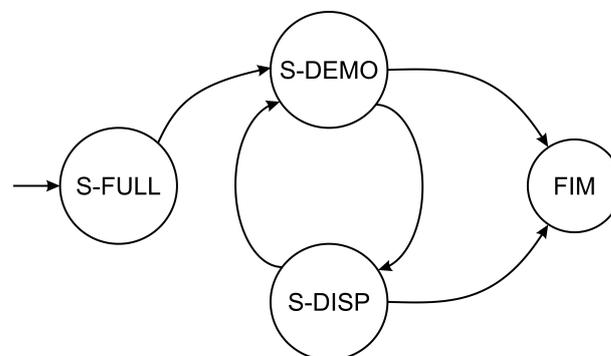


Figura 30 – Ordem de execução dos estágios do MSES.

Essa ordem é repetida ciclicamente até que todos os estágios sejam desenvolvidos, de acordo com a Figura 30. A estratégia MSES é representada pelo número de gerações

usadas em cada estágio evolutivo, separado por “-”. Para cada uma dessas estratégias, utilizamos as configurações de estágio 50–25–25, 50–10–10–10–10–10 e 50–30–20, onde esse conjunto de valores diz respeito a quantidade de gerações que uma determinada configuração de ajuste de parâmetros é aplicada. Por exemplo, a configuração 50–30–20 indica que 50 gerações são usadas para evoluir o conjunto escolhido no primeiro estágio (S_{full}), as próximas 30 gerações são usadas para ajustar o conjunto escolhido no segundo estágio (S_{demo}) e nas últimas 20 gerações o conjunto de parâmetros escolhido será desenvolvido no terceiro estágio (S_{disp}). Na estratégia 50–10–10–10–10–10, temos um total de 100 gerações divididas em seis estágios: (i) S_{full} será desenvolvido pelas primeiras 50 gerações; (ii) S_{demo} será evoluído por 10 gerações (até a 60ª geração); (iii) S_{disp} será evoluído por 10 gerações (até 70ª); (iv) S_{demo} será evoluído por 10 gerações (até 80ª); (v) S_{disp} será evoluído por 10 gerações (até 90ª); e, por fim, (vi) S_{demo} será evoluído pelas últimas 10 gerações (até a 100ª).

Considerando que cada estágio tem uma meta evolutiva diferente (o sub-conjunto de parâmetros escolhido para ser ajustado), a forma como a população é reinserida entre os estágios pode impactar no desempenho do algoritmo. Para avaliar tal influência, avaliamos três estratégias para repovoar a população entre os estágios: sem repovoamento (MSES-SR); com repovoamento baseado nos 20 melhores indivíduos (MSES-R20); e com repovoamento baseado no melhor indivíduo (MSES-RBEST).

6.2.1 Estratégia sem Repovoamento (MSES-SR)

A estratégia multi-estágio mais simples é a sem repovoamento, onde a população da última geração de um estágio é mantida para o próximo, sem nenhum tipo de interferência externa. São testados os 3 estágios para essa estratégia:

- MSES-SR (50 – 25 – 25),
- MSES-SR (50 – 10 – 10 – 10 – 10 – 10),
- MSES-SR (50 – 30 – 20).

6.2.2 Repovoamento Baseado nos 20 Melhores (MSES-R20)

Nessa estratégia, uma nova população é gerada cada vez que o algoritmo muda de estágio. O objetivo desse repovoamento é aumentar a diversidade nos parâmetros que serão ajustados na próxima etapa (parâmetros ativos). As estratégias de repovoamento adotadas baseiam-se em uma espécie de elitismo. A reconfiguração da população é feita na mudança de estágio, preservando sempre os 20 melhores indivíduos da população, ou seja, aqueles que possuem os 20 melhores valores de avaliação. Como as equações de aptidão definem um problema de minimização, serão aqueles com os menores valores retornados

pela equação 17 (TOP-20). Os parâmetros que não estão sendo ajustados naquele estágio recebem os valores dos parâmetros desses 20 indivíduos preservados, de forma que cada um deles seja replicado cinco vezes na população. Para os parâmetros que serão ajustados no estágio, preservamos os 20 melhores indivíduos e geramos valores aleatórios para os 80 seguintes. Reavaliamos toda a população depois disso. O intuito dessa reconfiguração é aumentar a diversidade daqueles parâmetros que serão ajustados ao mesmo tempo que restringimos a diversidade dos parâmetros congelados para valores previamente ajustados como bons, buscando assim uma nova combinação de parâmetros que seja mais adequada ao problema.

As três configurações de estágio apresentadas anteriormente são testadas para essa estratégia.

- MSES-R20 (50 – 25 – 25),
- MSES-R20 (50 – 10 – 10 – 10 – 10 – 10),
- MSES-R20 (50 – 30 – 20).

6.2.3 Repovoamento Baseado no Melhor Indivíduo (MSES-RBEST)

Esse repovoamento executa a avaliação robusta para os 10 melhores indivíduos, a fim de assegurar a manutenção do melhor dentre eles. Uma vez escolhido o indivíduo, seus valores são replicados para os parâmetros que não serão evoluídos (parâmetros fixos) em todos os indivíduos da nova população. Os valores dos demais parâmetros são definidos de acordo com a seguinte estratégia de repovoamento:

- Preservação parcial da carga genética dos 20 melhores indivíduos da última população (elitismo), onde são copiados apenas os valores para os parâmetros a serem ajustados no estágio corrente.
- Geração de N indivíduos a partir de uma única mutação no cromossomo do melhor indivíduo da última população. Essa mutação consiste na adição de um número aleatório entre $[-0, 2; 0, 2]$ ao valor original de um dos parâmetros a ser evoluído. Considerando que são realizadas 10 mutações por parâmetro, o valor de N varia de acordo com o subconjunto a ser tratado no estágio (40 para o S_{demo} e 30 para o S_{disp}).
- Geração aleatória dos valores dos parâmetros para o restante da população (40 indivíduos para os estágios que ajustam S_{demo} e 50 indivíduos quando S_{disp} é evoluído). Nesse processo, é usada a mesma estratégia empregada na geração da população inicial.

Após o processo de repovoamento, os indivíduos da população são então avaliados e ranqueados e o processo evolutivo típico do algoritmo genético é retomado.

Essa estratégia de repovoamento também é aplicada às três configurações multi-estágio avaliadas nesta pesquisa:

- MSES-RBEST (50 – 25 – 25),
- MSES-RBEST (50 – 10 – 10 – 10 – 10 – 10),
- MSES-RBEST (50 – 30 – 20).

6.2.4 Visão Geral do Algoritmo Multi-Estágio

A integração das estratégias multi-estágio e de repovoamento ao algoritmo genético é ilustrada na Figura 31. Como pode ser observado, o fluxograma do algoritmo multi-estágio (MSES) é semelhante ao PGA em muitos aspectos. Após a população ser iniciada e avaliada, verifica-se se a condição de parada é satisfeita. Caso não tenha sido, verifica-se se é o momento de uma mudança de estágio, a qual pode ocorrer com ou sem repovoamento, de acordo com a estratégia adotada. Após a devida configuração, o algoritmo segue o fluxo típico do AG, com os operadores respeitando os subconjuntos de ajuste inerentes a cada etapa.

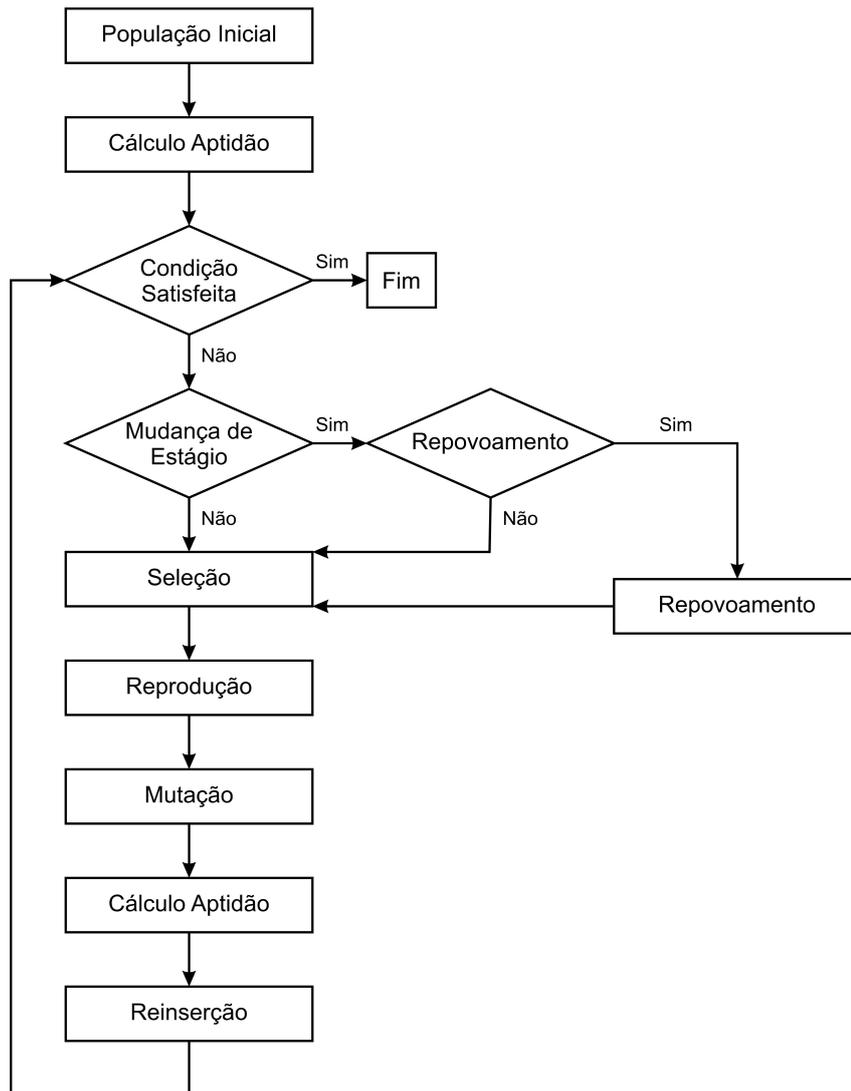


Figura 31 – Fluxograma da estratégia evolutiva multi-estágio.

Experimentos e Análise dos Resultados

Este capítulo apresenta os experimentos realizados a fim de avaliar a eficácia da nossa estratégia de ajuste baseada em AG para os parâmetros de um modelo epidemiológico baseado em regras de transição probabilísticas de ACs.

7.1 Método para a Avaliação

Para todos os experimentos, o AG foi executado 30 vezes. Em todas as simulações, o modelo baseado em AC começa com um reticulado vazio, ou seja, sem insetos (larvas ou adultos) na área da aldeia. Em nossos experimentos, a análise comparativa foi realizada utilizando a melhor configuração encontrada pelo AG, ou seja, aquela que resultou no menor erro durante a avaliação robusta (baseada em 100 simulações) entre todas as 30 execuções do AG.

7.2 Algoritmo Genético Simples

Os experimentos do PGA são realizados em três configurações:

1. PGA-DEMO: configura apenas os parâmetros de S_{demo} , fixando os valores de referência para os outros parâmetros ($D_i = 0,9$; $D_{ni} = 0,1$; e $Qf = 95$) por todas as gerações,
2. PGA-DISP: configura apenas os parâmetros de S_{disp} , fixando os valores de referência para os outros parâmetros ($p_r = 0,004$; $p_{sy} = 0,903$; $p_{sa} = 0,983$; $p_d = 0,004$) por todas as gerações,
3. PGA-FULL: configura os parâmetros de S_{full} , ou seja, todos os sete parâmetros do modelo e, portanto, sem utilizar os valores de referência.

A configuração de parâmetros que apresentou o melhor resultado dentre as 30 execuções para cada um dos experimentos é descrita na Tabela 3. As células destacadas em

cinza realçam os valores que foram pré-definidos, enquanto as células em branco retratam o que foi de fato ajustado pelo AG. O PGA-DEMO ajusta 4 parâmetros, o PGA-DISP ajusta 3 e o PGA-FULL todo o conjunto de 7 parâmetros.

Tabela 3 – Melhor indivíduo e o erro robusto resultantes das execuções com o PGA.

| Configuração | p_r | p_{sy} | p_{sa} | p_d | D_i | D_{ni} | Q_f | Erro |
|--------------|---------|----------|----------|---------|---------|----------|-------|------|
| Referência | 0,00411 | 0,90272 | 0,98280 | 0,00415 | 0,9 | 0,1 | 95 | |
| PGA-DEMO | 0,07509 | 0,16223 | 0,97419 | 0,66626 | 0,9 | 0,1 | 95 | 2,63 |
| PGA-DISP | 0,00411 | 0,90272 | 0,98280 | 0,00415 | 0,93475 | 0,80599 | 93 | 2,31 |
| PGA-FULL | 0,05386 | 0,19861 | 0,97204 | 0,56577 | 0,41054 | 0,66327 | 53 | 3,13 |

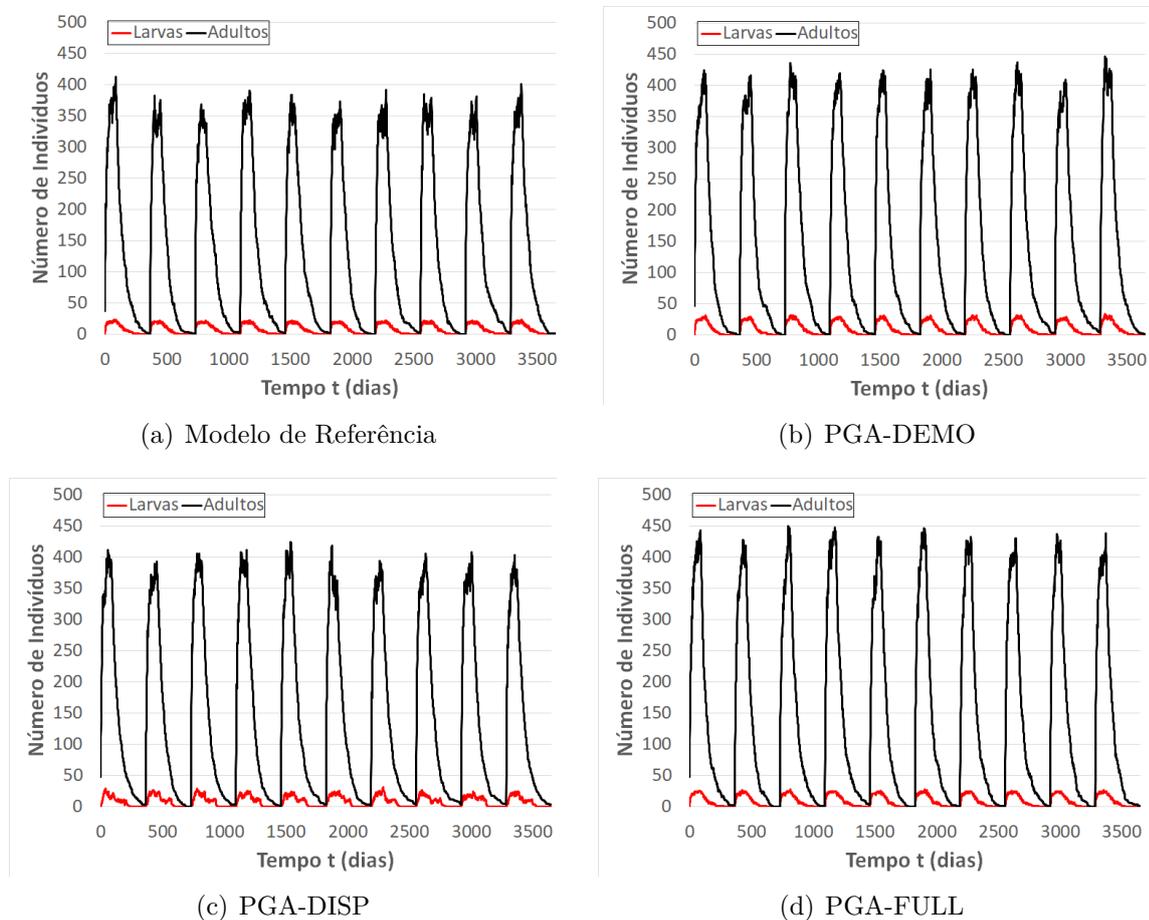


Figura 32 – Dinâmica populacional de insetos ao longo de 10 anos simulados pelo modelo usando as configurações definidas pelo PGA.

A Figura 32 mostra a dinâmica populacional de larvas e adultos resultante das três versões do experimento em um período de 10 anos, comparada com os valores de referência (Figura 32(a)). O número de insetos varia dentro dos períodos de infestação e não infestação conforme o esperado. Embora os valores de alguns parâmetros evoluídos difiram significativamente daqueles do modelo de referência descritos na Tabela 3 (ou seja, p_{sy} e p_d), é possível observar que ambas as versões do AG foram capazes de equaci-

onar uma boa composição de valores de parâmetros (e suas iterações) para atingir uma configuração satisfatória do modelo baseado em AC. Todas as três evoluções temporais apresentam comportamentos muito semelhantes de larvas e adultos em relação ao modelo de referência da Figura 32(a).

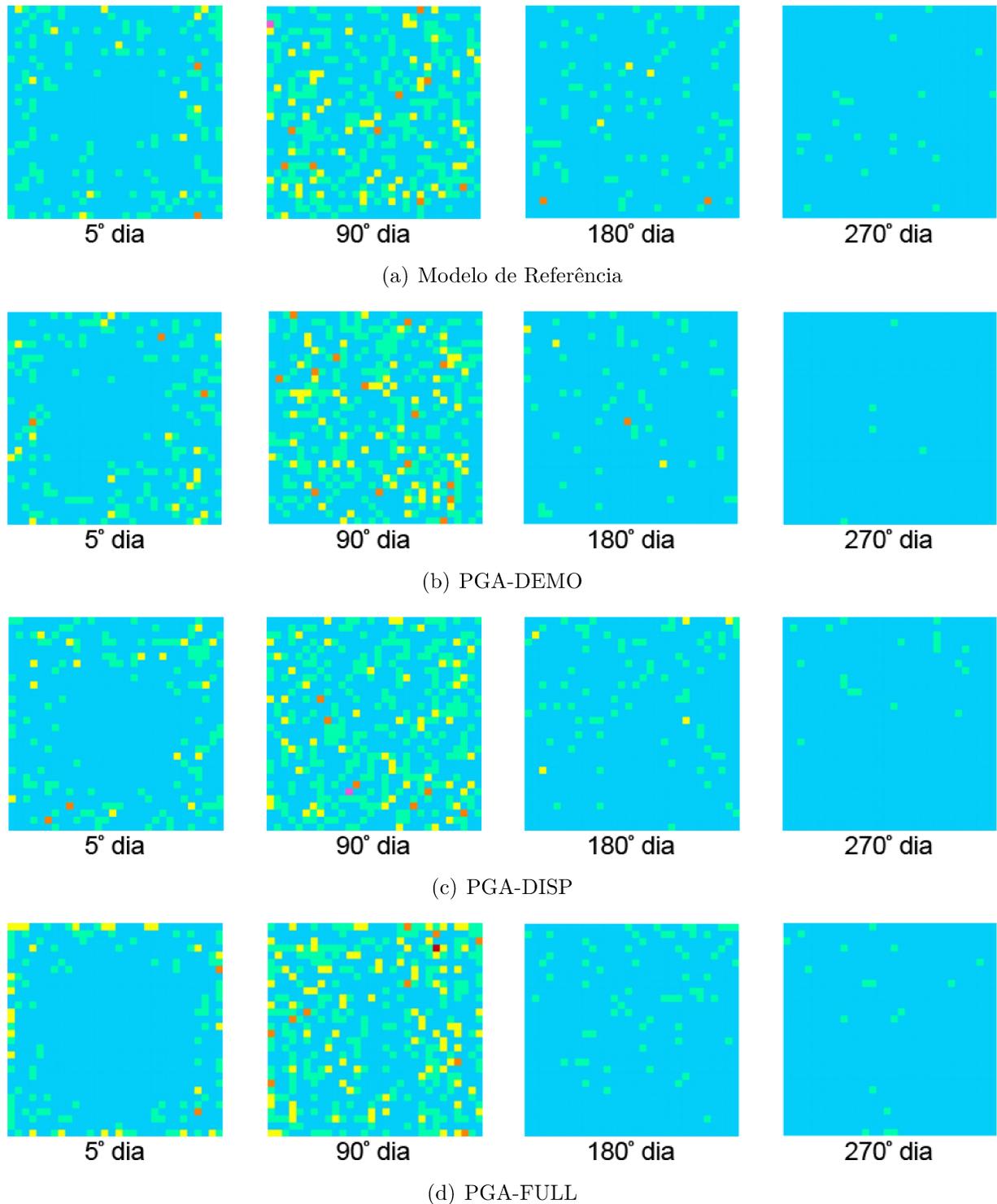


Figura 33 – Distribuição espacial dos adultos nos 5^o, 90^o, 180^o e 270^o dias de simulação do modelo, usando as configurações definidas pelo PGA.

A Figura 33 apresenta a evolução espacial dos insetos adultos nos modelos evoluídos

pelo AG, mostrando a população adulta para os mesmos dias descritos na Figura 32. As células azuis indicam a ausência de vetores, enquanto as verdes, amarelas, alaranjadas, rosas e vermelhas indicam a presença de 1, 2, 3, 4 e 5 adultos, respectivamente. A Figura 33(b) mostra a distribuição espacial resultante da configuração do conjunto S_{demo} de probabilidades. A Figura 33(c) mostra a distribuição espacial resultante do AC configurado a partir do conjunto de parâmetros S_{disp} . Na Figura 33(d), o AC foi simulado usando os 7 parâmetros retornados pela versão mais completa do GA. Conforme observado em modelos anteriores (original e de referência), os reticulados do AC apresentam comportamentos espaciais semelhantes. Quando o período de infestação começa (5º dia), os adultos estão concentrados perto dos limites da aldeia, uma vez que estão migrando da floresta para a aldeia. Pode-se notar que a versão completa do AG tem uma distribuição mais esparsa do que o PGA-DISP. Ou seja, no caso da PGA-FULL os adultos estão mais concentrados próximo às bordas, enquanto no PGA-DISP é possível observar adultos em uma região intermediária, o que aproxima esse comportamento mais dos resultados obtidos com o modelo de referência. Por volta de 90º dia ocorre o pico de infestação em todos os modelos. Os insetos estão bem distribuídos e em grande quantidade ao longo do reticulado. No período de não infestação (180º e 270º dias) observa-se uma redução considerável no número de vetores. Os reticulados apresentam comportamentos espaciais semelhantes no que diz respeito à dinâmica do modelo. As quantidades e concentrações de adultos por região variam um pouco entre os modelos mas não a ponto de alterar a dinâmica. No 270º dia, essas alterações ficam mais evidentes, onde os adultos aparecem em menor quantidade no PGA-DEMO, mais concentrados na parte superior do reticulado no PGA-DISP e mais esparsos no PGA-FULL.

A Figura 34 mostra a evolução espacial das larvas, nas abordagens analisadas. As coloração das células segue a mesma definição da Figura 33. Observamos comportamentos espaciais semelhantes no que diz respeito à dinâmica do modelo e às quantidades e proporções de insetos observadas na Figura 32. As concentrações de larvas por região sofrem variações mas dentro do esperado para o tipo de modelagem. A diferença de concentração fica mais evidente no 90º dia, sendo que nos modelos de referência e PGA-DISP as larvas estão mais aglomeradas e nas demais abordagens há um espalhamento maior.

A Figura 35 apresenta a evolução do erro para o melhor, mediano e pior indivíduo, a cada geração. É possível observar que nas três abordagens, a avaliação tanto do melhor indivíduo, quanto da média da população, experimenta um decréscimo contínuo, à medida que as gerações evoluem. Esse é um comportamento esperado para um AG simples que usa um método de reinserção que preserva o melhor indivíduo e uma função de avaliação que não é dinâmica. Ou seja, com o passar das gerações, os melhores indivíduos podem apresentar uma melhoria nos valores de avaliação (menor erro), sendo que nunca observamos uma deterioração nas avaliações, no máximo uma estabilização. Entretanto, é nítido que tanto no PGA-DEMO quanto no PGA-DISP, que otimizam um número menor

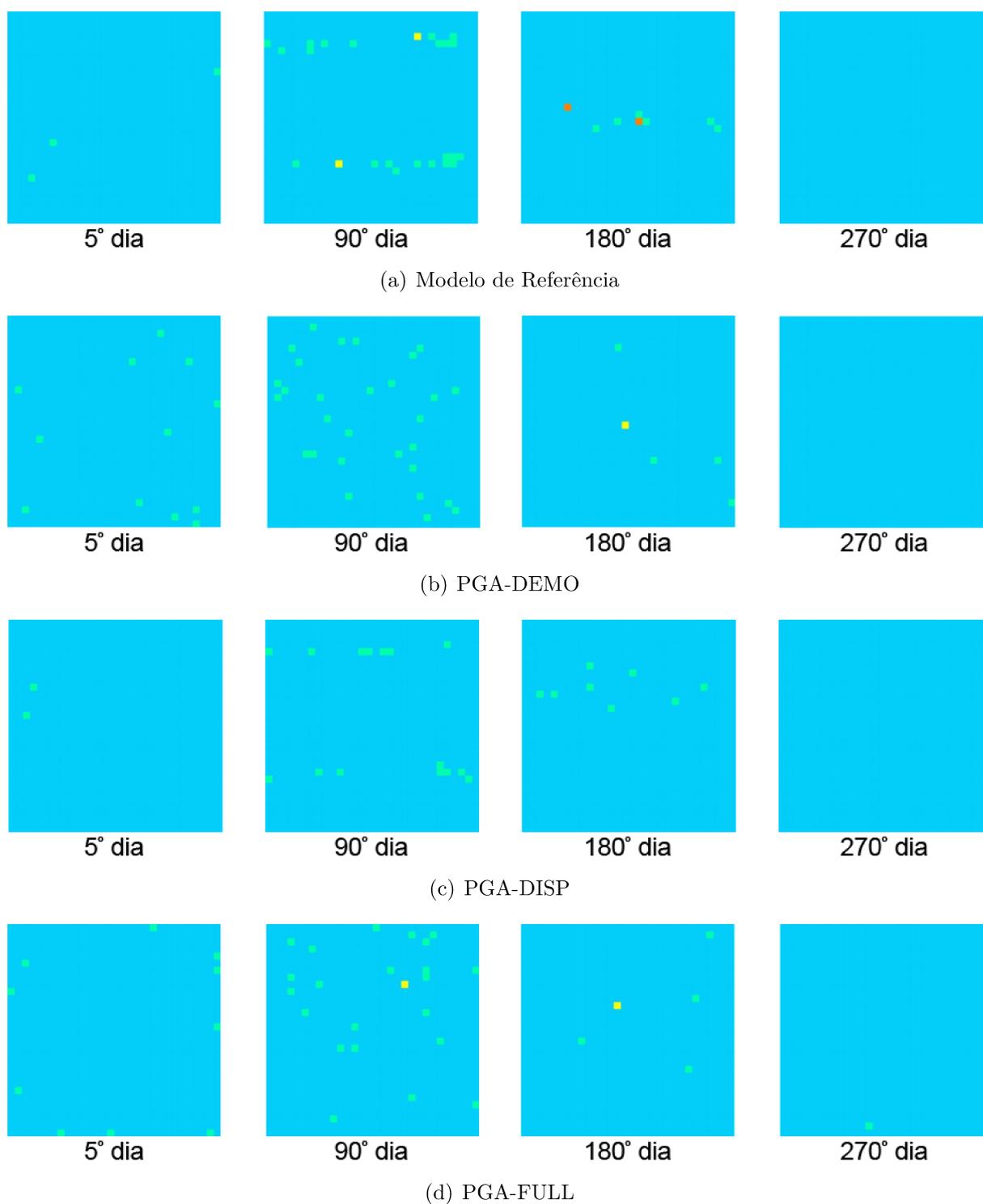


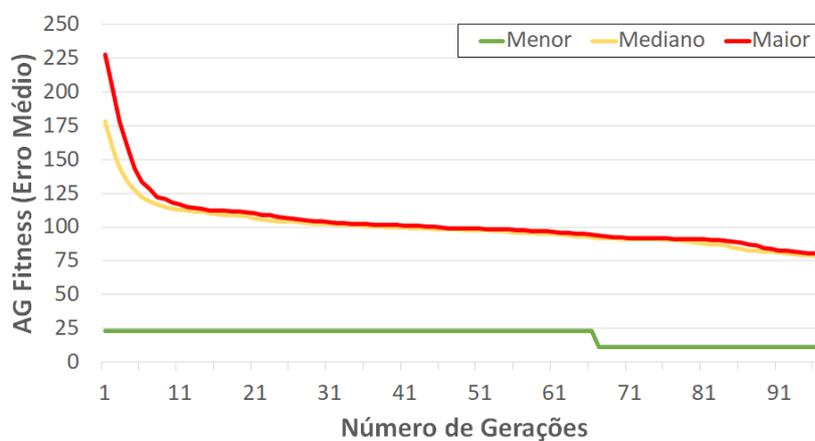
Figura 34 – Distribuição espacial das larvas nos 5^o, 90^o, 180^o e 270^o dias de simulação do modelo, usando as configurações definidas pelo PGA.

de parâmetros e são pré-ajustados para os demais parâmetros com os valores de referência, esse decréscimo do erro se dá de uma forma muito rápida, quando comparados ao PGA-FULL que otimiza todos os parâmetros a partir de uma inicialização aleatória. Esse comportamento realça uma conclusão que já era esperada: quando apenas um subconjunto dos parâmetros é utilizado o problema de busca é relativamente simples e o AG converge rapidamente (praticamente nas 10 primeiras gerações). Além disso, entre os dois sub-problemas sendo otimizados, fica claro que o ajuste dos parâmetros da fase demográfica é muito mais crítico e quando já partimos desses valores ajustados pelo modelo de (SLIMI et al., 2009), como na estratégia PGA-DISP, os melhores indivíduos já partem da população inicial com valores de erro muito baixos e o AG praticamente não contribui na melhoria desses melhores indivíduos iniciais gerados aleatoriamente. Na estratégia PGA-DISP apresentada na Figura 35(b), a convergência do AG apenas atua na melhoria da média populacional, conseguindo fazer com que a média da população praticamente se iguale ao melhor indivíduo, o que representa muito provavelmente que todos os indivíduos da população acabam por se tornar cópias desse melhor indivíduo que é rapidamente encontrado nas gerações iniciais. Por outro lado, no PGA-DEMO, que ajusta apenas esses parâmetros críticos da fase demográfica, embora a convergência observada na Figura 35(a) mostre que a média da população também cai rapidamente, ela não se iguala ao valor do melhor indivíduo, demonstrando que a população ainda mantém uma diversidade maior até o final da execução. O melhor indivíduo do PGA-DEMO encontrado na geração inicial, permanece por quase 70 gerações e só é melhorado muito provavelmente por uma mutação benéfica. Entretanto, no caso da estratégia PGA-FULL apresentada na Figura 35(c), vemos um gráfico de evolução em que não só a média da população tem um decaimento mais prolongado, como a avaliação do melhor indivíduo passa por várias mudanças abruptas, denotando que o melhor indivíduo foi atualizado várias vezes durante a evolução. Portanto, no PGA-FULL é possível identificar uma boa convergência do AG que mesmo partindo de um problema mais complexo, consegue aos poucos aproximar a avaliação média da aptidão do melhor indivíduo e, embora parta de um indivíduo que representa uma configuração bem desfavorável ao modelo (com um erro alto), consegue ajustar esses valores por adaptações sucessivas que ocorrem pela aplicação dos operadores genéticos.

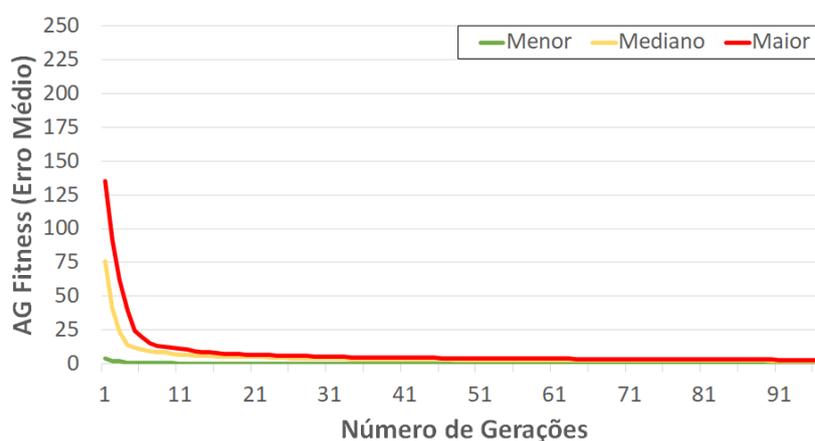
Tabela 4 – Erro médio e desvio padrão resultantes das execuções com o PGA.

| Abordagem | Menor Erro | Média | Desvio Padrão |
|-----------|------------|-------|---------------|
| PGA-DEMO | 2,63 | 3,81 | 1,62 |
| PGA-DISP | 2,31 | 2,40 | 0,07 |
| PGA-FULL | 3,13 | 7,77 | 3,59 |

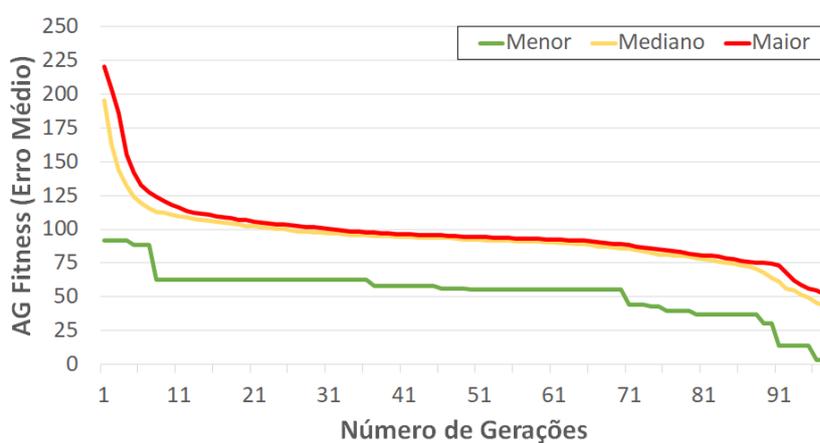
Como podemos observar, os experimentos feitos com as abordagens parciais (PGA-DEMO e PGA-DISP), que ajustam 2 subconjuntos dos parâmetros, utilizando parte dos



(a) PGA-DEMO



(b) PGA-DISP



(c) PGA-FULL

Figura 35 – Evolução do erro para o melhor, pior e indivíduo mediano da população, considerando a melhor execução com o PGA.

parâmetros fixados iguais ao modelo de referência, resultam em um erro menor que a abordagem que ajusta um conjunto equivalente à união dos dois subconjuntos anteriores (PGA-FULL). A complexidade do problema aumenta à medida que aumentam as possíveis combinações de parâmetros. Analisando os erros robustos alcançados na melhor execução de cada versão do AG, observamos que a PGA-FULL retornou um erro de 3,13, enquanto o PGA-DISP obteve um erro de 2,63 na simulação de um ano. Considerando que o número médio de insetos (larvas e adultos) no mesmo período é de 30.700, quantitativamente a diferença entre os modelos de referência e os baseados em AG é desprezível. Ou seja, no cenário mais complexo (PGA-FULL), essa diferença é de aproximadamente 3 insetos na simulação de um ano (em torno de 0,01%). Comparando a média aritmética e o desvio padrão do erro robusto obtido pelos melhores indivíduos em todas as execuções do AG, há uma maior diferenciação entre as versões (PGA-DISP e PGA-FULL), embora os resultados de ambos ainda estejam em um nível aceitável. O ajuste de PGA-DISP retornou um erro médio de $3,81 \pm 1,62$, enquanto a versão completa resultou em um erro médio de $7,77 \pm 3,59$. As abordagens evolutivas multi-estágio com e sem repovoamento foram estratégias desenvolvidas para diminuir o erro procurando dividir o problema em outros de complexidade menor ao mesmo tempo que testamos diferentes configurações de diversidade. Embora não tenhamos utilizado nenhuma métrica para avaliação espacial dos dados gerados a partir de modelos baseados em AG, é possível observar uma boa semelhança entre os reticulados apresentados nas figuras 26 e 33. Assim, consideramos que o método proposto foi capaz de modelar adequadamente os dados de referência em todas as abordagens investigadas.

7.3 MSES - Estratégias Evolutivas Multi-Estágio

Esta seção apresenta os experimentos realizados com a estratégia evolutiva multi-estágio. Os valores de referência são apresentados de acordo com a nossa representação do modelo descrito em (SLIMI et al., 2009). Apresentamos os resultados da abordagem MSES-SR, MSES-R20 e MSES-RBEST, separadamente. Um comparativo dos valores dos parâmetros de todas as abordagens, erro, média do erro e desvio padrão do erro dos experimentos é apresentado ao final de cada subseção. Para cada abordagem foram executados 30 experimentos, usando reticulados 30×30 , que foi o mesmo tamanho adotado em (SLIMI et al., 2009). Todos os experimentos foram realizados em uma população de 100 indivíduos durante 100 gerações. As próximas sub-seções detalham os resultados dos experimentos usando diferentes estratégias para a população durante a mudança de estágio.

7.3.1 Análise da Estratégia MSES-SR

A estratégia multi-estágio sem repovoamento (MSES-SR) não altera a população entre os estágios da evolução. A configuração de parâmetros para cada uma das estratégias multi-estágio sem repovoamento que apresentou o melhor resultado dentre as 30 execuções é descrita na Tabela 5. A coluna “Erro” mostra o resultado da avaliação robusta da melhor das execuções.

Tabela 5 – Melhor indivíduo e o erro robusto resultantes das execuções com as estratégias evolutivas: (1) PGA-FULL, (2) MSES-SR-50-25-25, (3) MSES-SR-50-10-10-10-10-10 e (4) MSES-SR-50-30-20.

| Configuração | p_r | p_{sy} | p_{sa} | p_d | D_i | D_{ni} | Qf | Erro |
|--------------|---------|----------|----------|---------|---------|----------|------|------|
| Referência | 0,00411 | 0,90272 | 0,98280 | 0,00415 | 0,9 | 0,1 | 95 | |
| 1 | 0,05386 | 0,19861 | 0,97204 | 0,56577 | 0,41054 | 0,66327 | 53 | 3,13 |
| 2 | 0,02278 | 0,70207 | 0,96335 | 0,25165 | 0,83271 | 0,14574 | 100 | 3,79 |
| 3 | 0,05646 | 0,62292 | 0,93870 | 0,9922 | 0,78677 | 0,88614 | 98 | 4,29 |
| 4 | 0,05067 | 0,25667 | 0,97438 | 0,24806 | 0,75753 | 0,23469 | 80 | 2,50 |

Na Figura36 podemos observar a dinâmica populacional de larvas e adultos para um período de 10 anos. No geral, o padrão de comportamento cíclico se assemelha ao esperado. Os experimentos com as estratégias MSES-SR-50-25-25 e MSES-SR-50-10-10-10-10 são ilustrados nas Figura36(b) e Figura36(c), respectivamente, e apresentam picos de infestação de adultos levemente maiores que o modelo de referência, enquanto o experimento com a estratégia MSES-SR-50-30-20 (Figura36(d)) mantém seu picos no limite inferior.

A Figura 37 apresenta a evolução espacial dos insetos adultos nos modelos evoluídos pelo AG, mostrando a população adulta para os mesmos dias descritos na Figura 36. Todos os experimentos ficaram semelhantes em relação às quantidades. Podemos destacar no experimento com a estratégia MSES-SR-50-25-25 ilustrado na Figura 37(b), que no 5º dia, os insetos não conseguiram atingir regiões do centro do reticulado como no modelo de referência (Figura 37(a)), enquanto no 180º dia, estiveram menos espalhados e no 270º dia em quantidade ligeiramente menor. No experimento MSES-SR-50-10-10-10-10-10, representado na Figura 37(c) os insetos apresentaram um espalhamento semelhante em boa parte do ciclo, porém dispersam um pouco mais cedo, evidenciado na imagem do 270º dia. No experimento MSES-SR-50-30-20 as quantidades e espalhamentos estão semelhantes ao modelo de referência nos 2 primeiros reticulados, já nos 2 últimos, o espalhamento ficou mais concentrado e deslocado para a região superior direita.

A Figura 38 destaca o comportamento das larvas no mesmo período considerado nas análises anteriores. As quantidades de insetos são bem semelhantes ao modelo de referência (Figura 38(a)), já a distribuições dos mesmos no reticulado é um pouco menos concentrada que o esperado. No experimento com a estratégia MSES-SR-50-30-20 (Figura 38(d)) a dispersão é um pouco mais lenta evidenciada pela imagem do 270º dia.

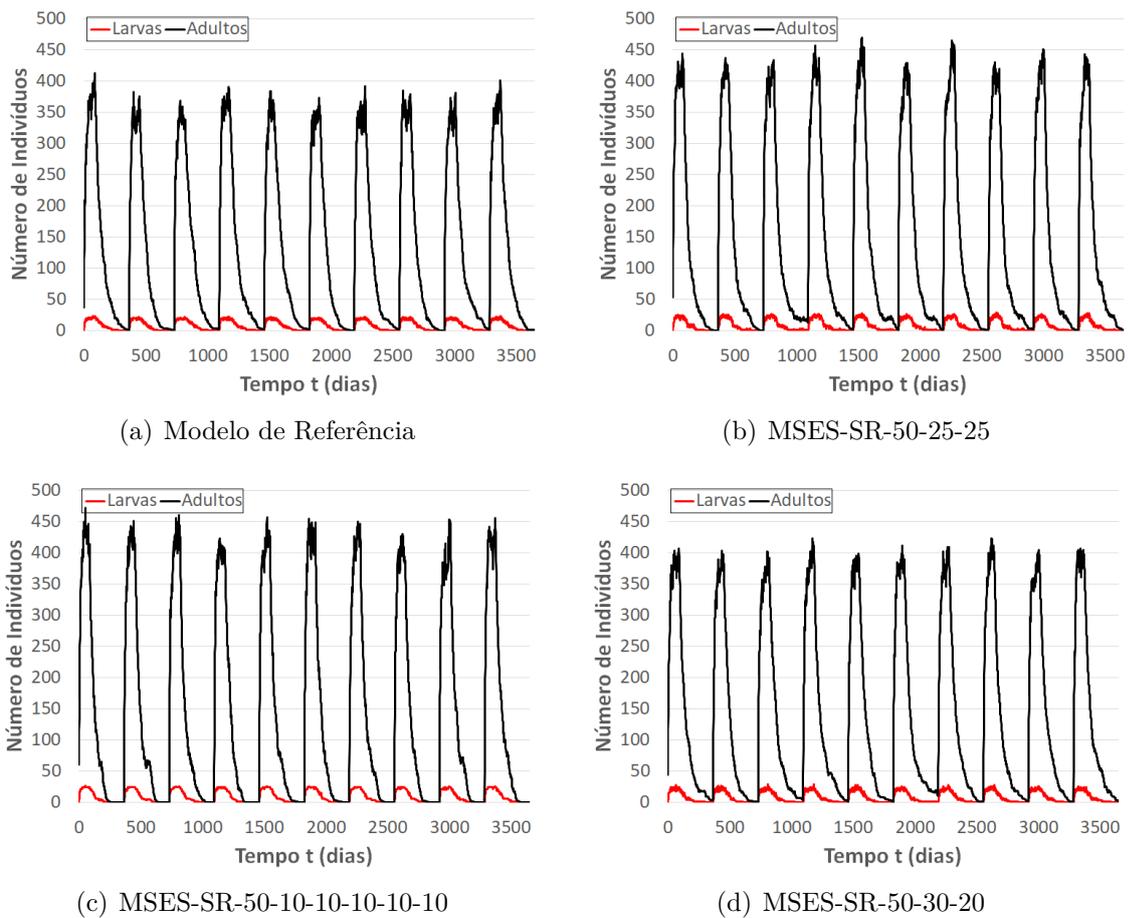


Figura 36 – Dinâmica populacional de insetos ao longo de 10 anos simulados pelo modelo usando as configurações definidas pelas estratégias MSES-SR.

A Figura 39 apresenta a evolução do erro das execuções para o melhor, mediano e pior indivíduo, a cada geração e para cada estratégia MSES-SR avaliada. Nos 3 gráficos da Figura 39, podemos observar um comportamento similar, o melhor indivíduo parte de um valor mais alto de erro e apresenta diversos saltos de melhoria, até chegar em um erro final abaixo de 5. Por outro lado, a média populacional decai até que nas últimas gerações se aproxima do melhor indivíduo, mostrando uma baixa diversidade a partir da geração 90. Ou seja, apesar de todas as execuções serem finalizadas em 100 gerações, dificilmente teríamos uma melhoria do melhor indivíduo caso tivéssemos um número maior de gerações. Na Figura 39(a) observamos que a abordagem MSES-SR-50-25-25 apresenta uma diminuição dos erros equilibrada ao longo das gerações, enquanto na Figura 39(b) observamos um decaimento mais acentuado no início seguido de uma estabilização e proximidade do fitness médio da população. Na Figura 39(c) temos a evolução do erro mais constante, com um distanciamento maior do fitness do melhor indivíduo inicialmente, seguido de uma maior proximidade dos fitness médio da população no final da execução. O comportamento desses gráficos nas primeiras 50 gerações é similar ao gráfico da PGA-FULL apresentada Figura 35(c) e, de fato, nas primeiras 50 gerações

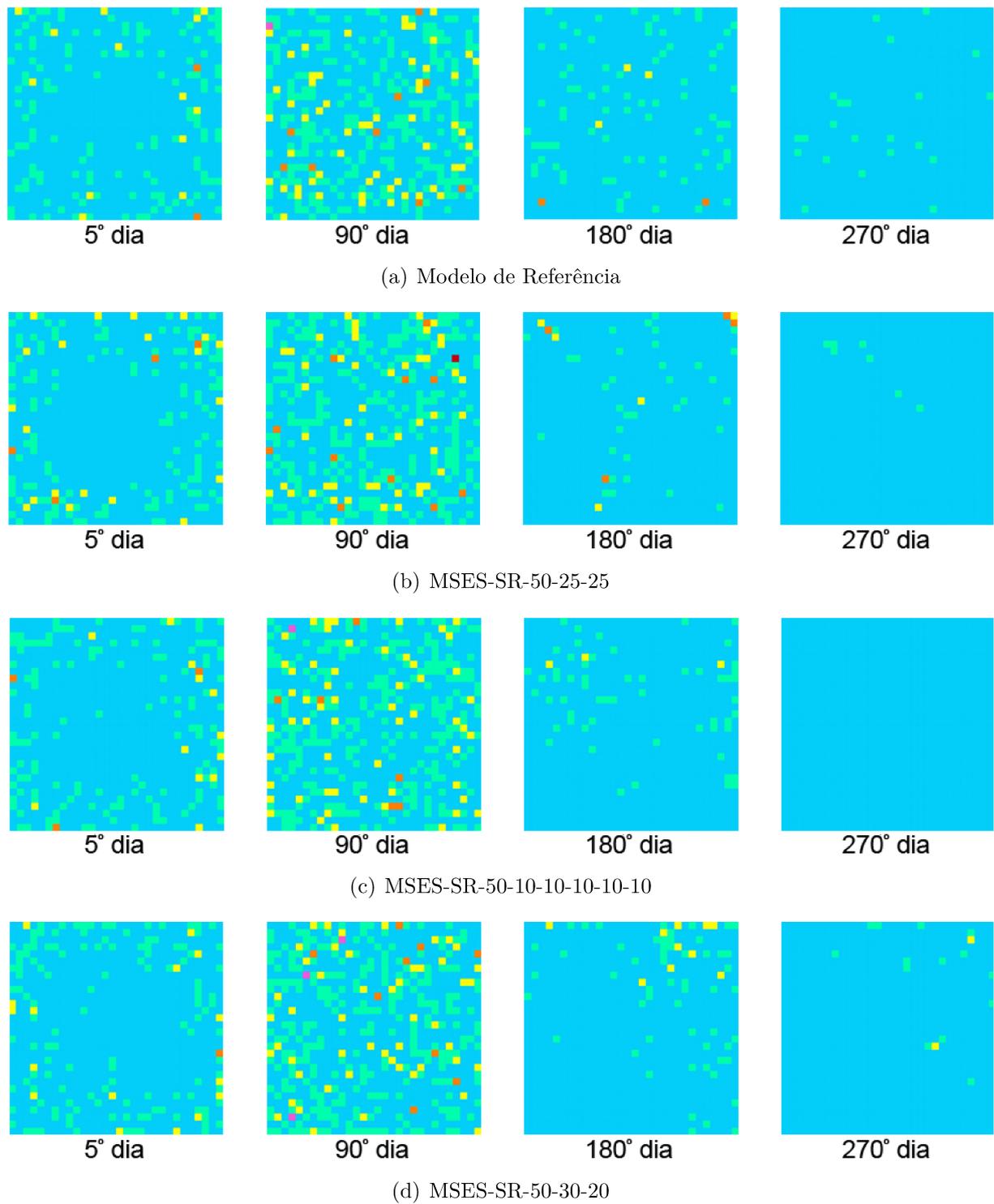


Figura 37 – Distribuição espacial dos adultos nos 5^o, 90^o, 180^o e 270^o dias de simulação do modelo, usando as configurações definidas pelas estratégias MSES-SR.

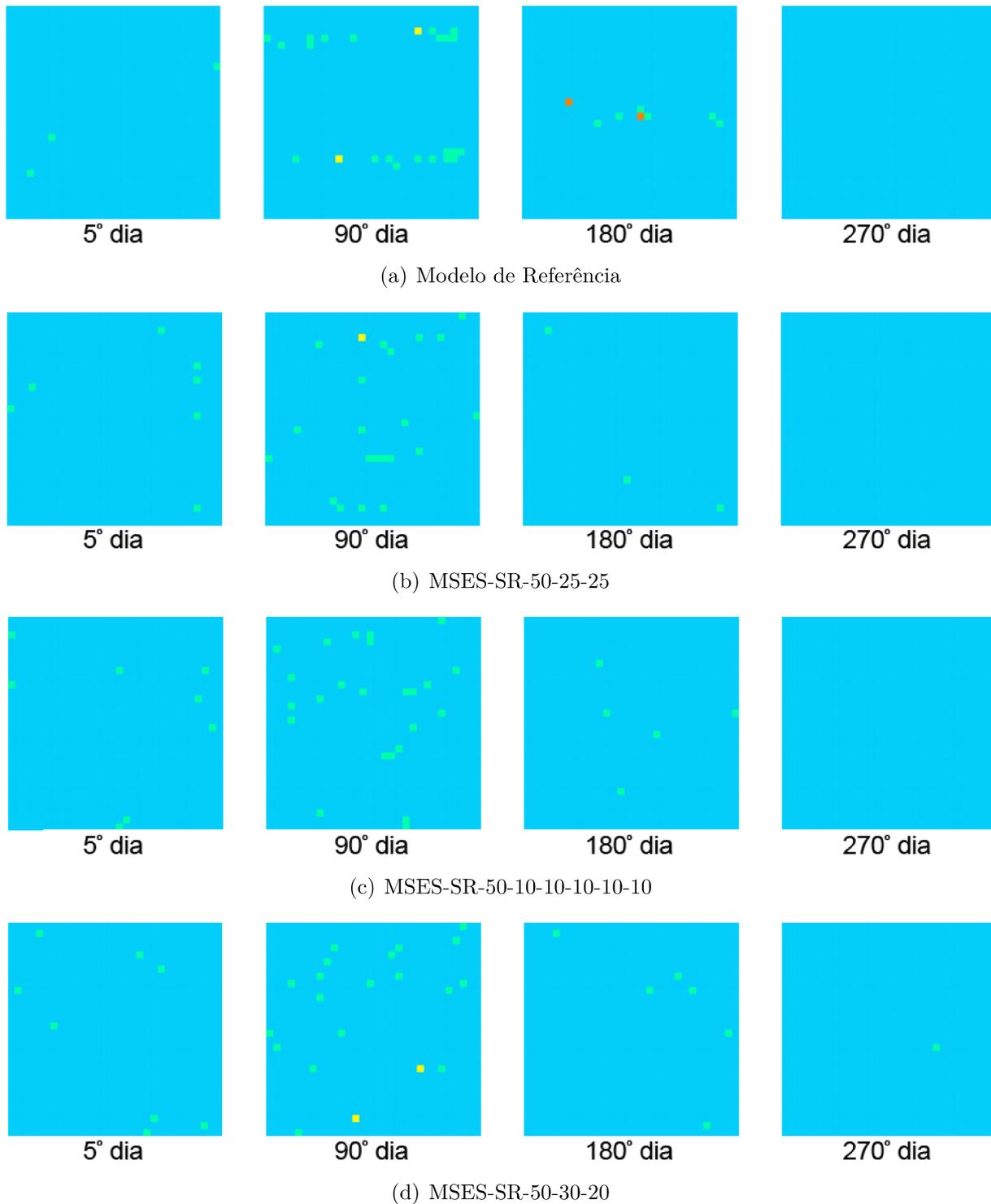


Figura 38 – Distribuição espacial das larvas nos 5º, 90º, 180º e 270º dias de simulação do modelo, usando as configurações definidas pelas estratégias MSES-SR.

todas as estratégias são similares e fazem a evolução dos 7 parâmetros do modelo a partir de uma amostra aleatória. A partir da geração 50, passamos a ter um comportamento diferente entre as estratégias e a estratégia simples (PGA-FULL), onde podemos perceber que o erro cai mais rapidamente, tanto para o melhor indivíduo, quanto para a média. Entretanto, o valor final do erro obtido pela PGA-FULL é menor do que os obtidos nas estratégias MSES-SR.

A Tabela 6 detalha os erros robustos alcançados na melhor execução de cada versão do MSES-SR. Observamos que o menor erro robusto e a menor média foram apresentados pela abordagem MSES-SR-50-30-20, com os valores 2,50 e 9,21, respectivamente. Já o menor desvio padrão foi encontrado na estratégia MSES-SR-50-25-25. A partir disso concluímos que a configuração da estratégia MSES-SR que obteve o melhor resultado foi a 50-30-20. Podemos destacar também que essa estratégia foi a única (dentre as três MSES-SR avaliadas) que encontrou uma solução melhor que a encontrada pelo PGA-FULL, apesar da média de resultados ser pior.

Tabela 6 – Erro médio e desvio padrão das execuções com as estratégias MSES-SR.

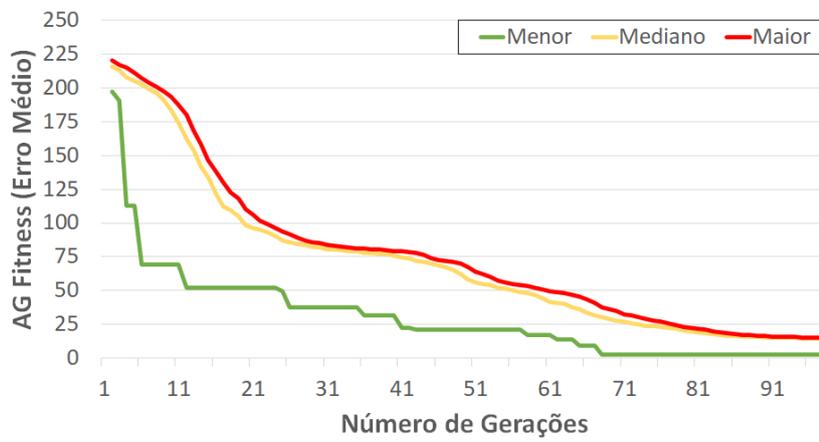
| Estratégia | Estágios | Menor Erro | Média | Desvio Padrão |
|------------|-------------------|------------|-------|---------------|
| PGA-FULL | - | 3,13 | 7,77 | 3,59 |
| MSES-SR | 50-25-25 | 3,79 | 15,39 | 3,81 |
| MSES-SR | 50-10-10-10-10-10 | 4,29 | 14,05 | 4,10 |
| MSES-SR | 50-30-20 | 2,50 | 9,21 | 3,95 |

7.3.2 Análise da Estratégia MSES-R20

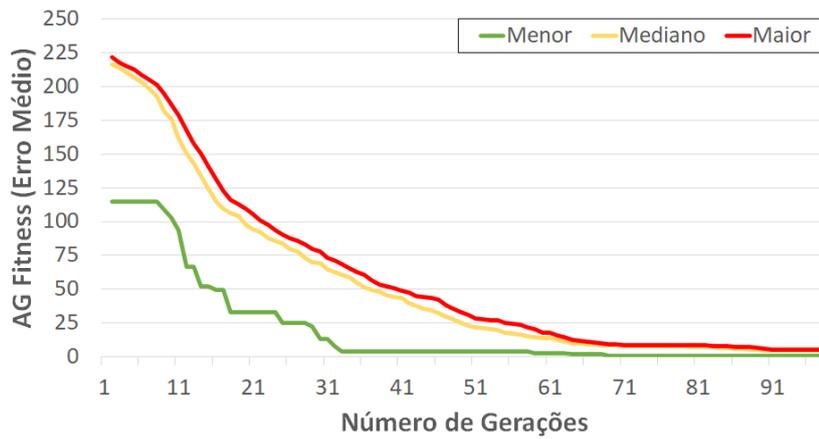
A estratégia multi-estágio com repovoamento baseado nos 20 melhores indivíduos (MSES-R20), altera a população a cada mudança de estágio. A configuração de parâmetros para cada uma das estratégias MSES-R20 que apresentou o melhor resultado dentre as 30 execuções é descrita na Tabela 7. Além disso, os resultados obtidos nos experimentos PGA-FULL e MSES-SR-50-30-20 são replicados para facilitar a comparação.

Tabela 7 – Melhor indivíduo e erro robusto resultantes das execuções com as estratégias evolutivas: (1) PGA-FULL, (2) MSES-SR-50-30-20, (3) MSES-R20 - 50-25-25, (4) MSES-R20 - 50-10-10-10-10-10 e (5) MSES-R20 - 50-30-20.

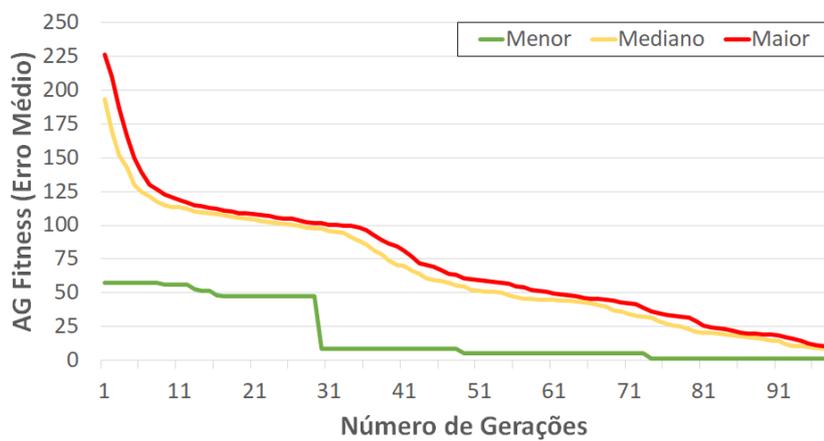
| Configuração | p_r | p_{sy} | p_{sa} | p_d | D_i | D_{ni} | Qf | Erro |
|--------------|---------|----------|----------|---------|---------|----------|------|------|
| Referência | 0,00411 | 0,90272 | 0,98280 | 0,00415 | 0,9 | 0,1 | 95 | |
| 1 | 0,05386 | 0,19861 | 0,97204 | 0,56577 | 0,41054 | 0,66327 | 53 | 3,13 |
| 2 | 0,05067 | 0,25667 | 0,97438 | 0,24806 | 0,75753 | 0,23469 | 80 | 2,50 |
| 3 | 0,05354 | 0,090324 | 0,97883 | 0,4818 | 0,6165 | 0,6822 | 65 | 2,63 |
| 4 | 0,07151 | 0,14038 | 0,97358 | 0,901 | 0,99296 | 0,29038 | 100 | 2,67 |
| 5 | 0,09358 | 0,01686 | 0,97901 | 0,26273 | 0,84703 | 0,3092 | 91 | 2,84 |



(a) MSES-SR-50-25-25



(b) MSES-SR-50-10-10-10-10-10



(c) MSES-SR-50-30-20

Figura 39 – Evolução do erro para o melhor, pior e indivíduo mediano da população, considerando a melhor execução com as estratégias MSES-SR.

A Figura 40 apresenta a quantidade de insetos resultantes da aplicação das configurações do MSES-R20 testadas. Todas as configurações conseguiram reproduzir bem o comportamento cíclico, com picos de adultos levemente inferiores.

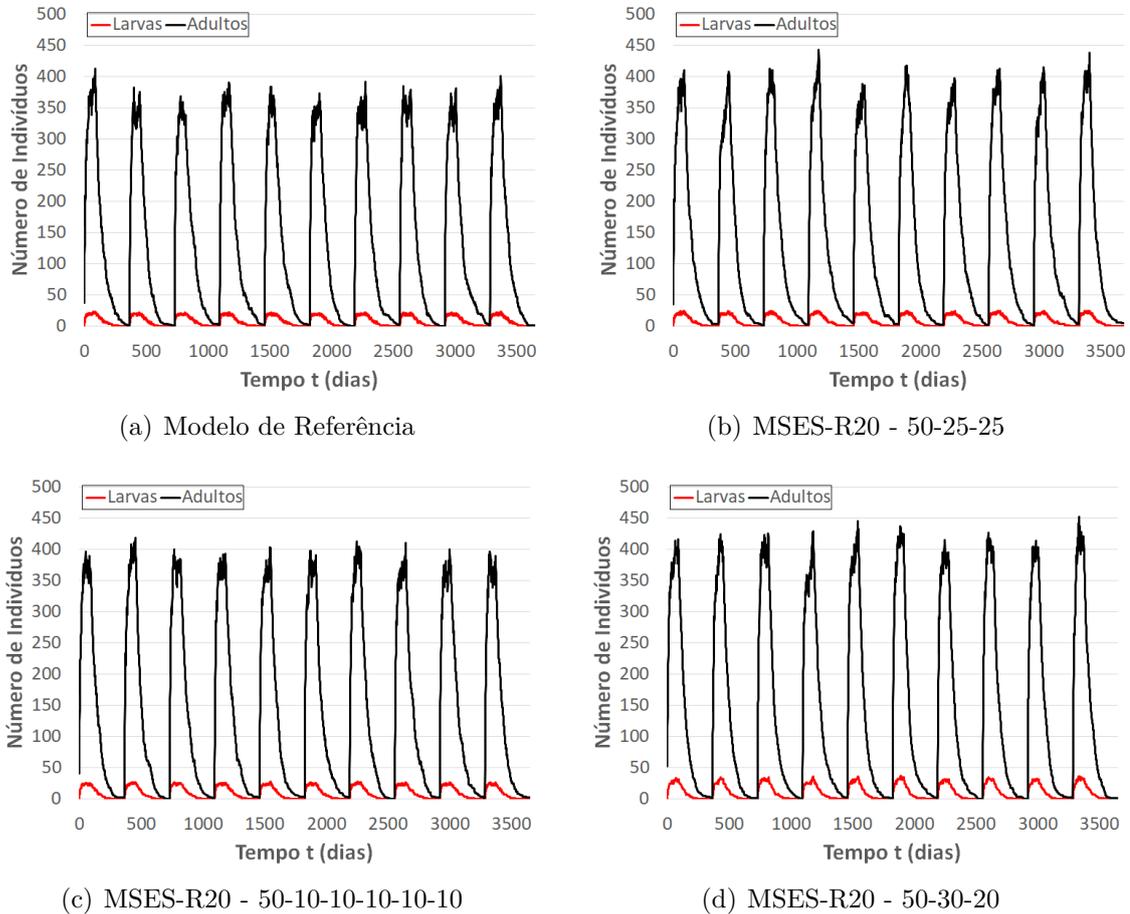


Figura 40 – Dinâmica populacional de insetos ao longo de 10 anos simulados pelo modelo, usando as configurações definidas pelas estratégias MSES-R20.

A distribuição espacial dos adultos apresentada na Figura 41 e das larvas apresentada na Figura 42 mostra um comportamento análogo ao esperado durante o ciclo. Nos adultos a estratégia que mais se aproximou do modelo de referência foi a MSES-R20-50-30-20 (Figura 41(d)) tanto em espalhamento quanto no tempo de permanência dos insetos. Já nas larvas, todas as estratégias tiveram comportamentos semelhantes.

A Figura 43 apresenta a evolução do erro das execuções para o melhor, mediano e pior indivíduo, a cada geração e para cada estratégia MSES-R20 analisada. É possível observar que nas 3 estratégias, os gráficos se comportaram de forma similar aos anteriores (MSES-SR e PGA-FULL), uma vez que essa etapa inicial os 7 parâmetros são ajustados a partir de uma amostragem aleatória. Ou seja, o erro do melhor indivíduo decresce em alguns saltos, sendo que quando se parte de um melhor indivíduo com um maior erro, esses saltos são mais frequentes e acentuados. Ao mesmo tempo, a média do fitness da população decai de forma constante. Entretanto, a partir da geração 50, é possível

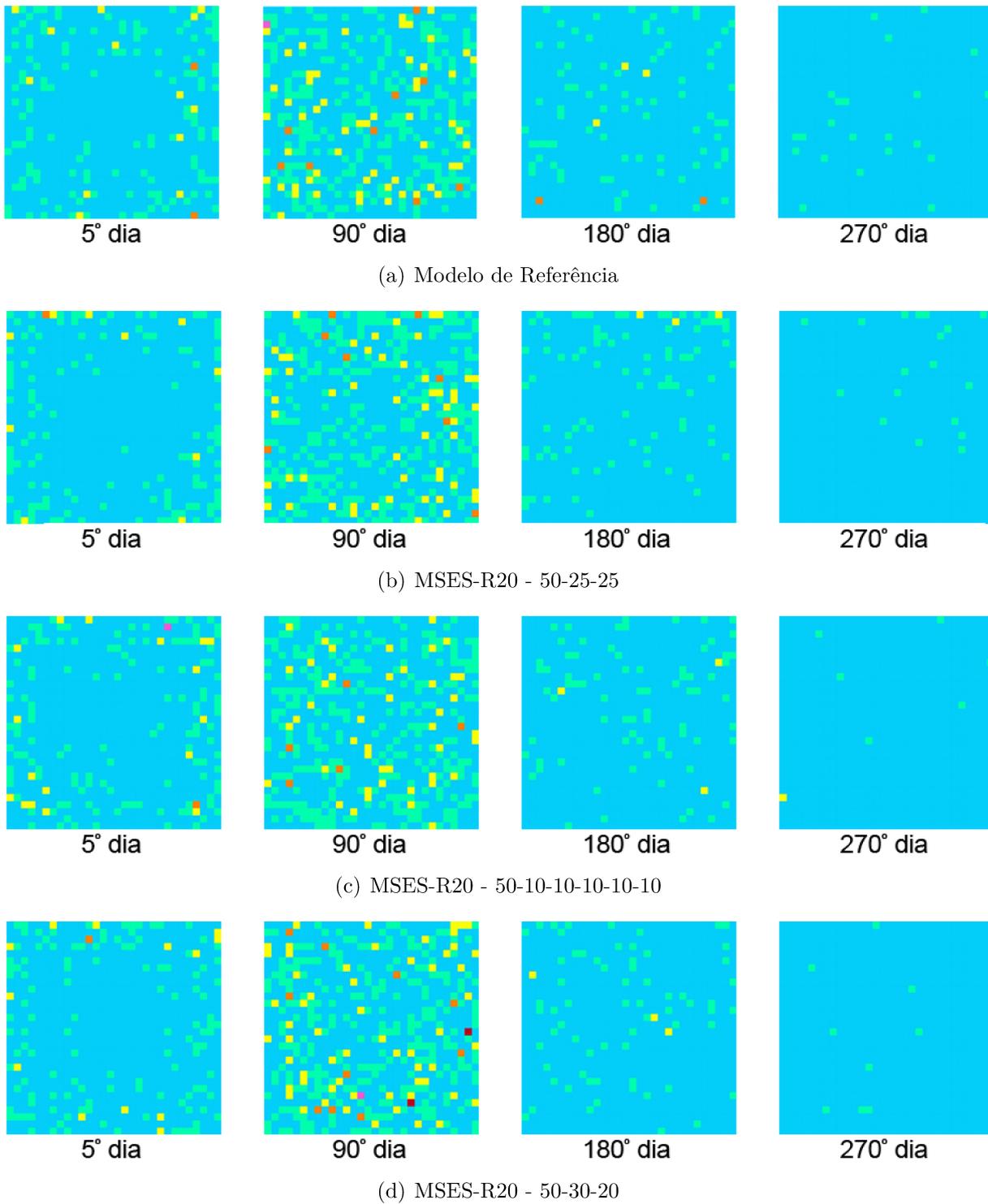


Figura 41 – Distribuição espacial dos adultos nos 5º, 90º, 180º e 270º dias de simulação do modelo, usando as configurações definidas pelas estratégias MSES-R20.

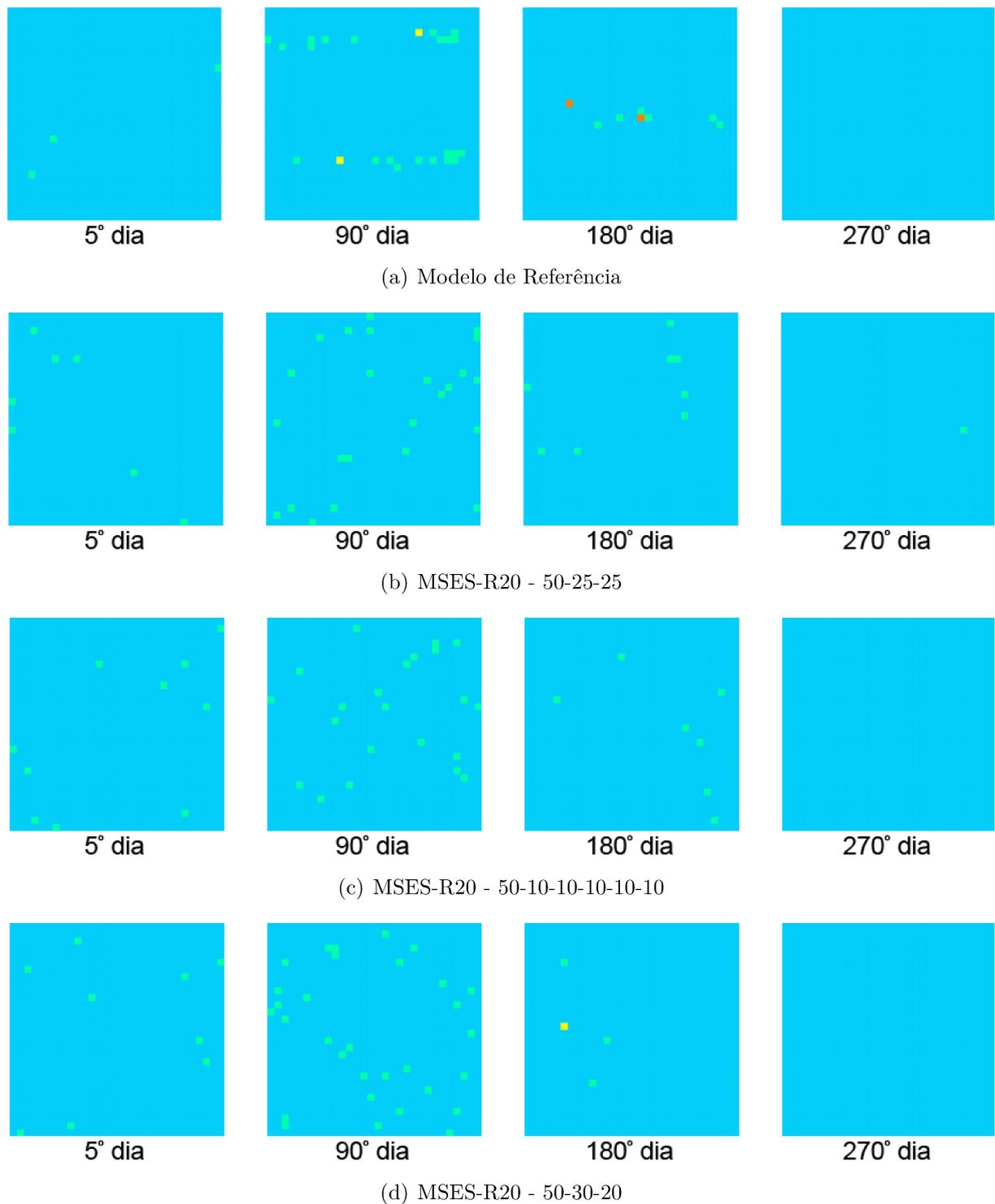


Figura 42 – Distribuição espacial das larvas nos 5º, 90º, 180º e 270º dias de simulação do modelo, usando as configurações definidas pelas estratégias MSES-R20.

observar que a variação MSES-R20 provoca alguns saltos de acréscimo na aptidão média da população, exatamente nos pontos em que uma nova etapa é reiniciada e a estratégia de repovoamento é aplicada. Após cada acréscimo, a aptidão média volta a cair até que uma nova etapa seja iniciada. Dessa forma, é possível perceber que essa variação injeta uma variedade maior na população a partir da geração 50. O erro do melhor indivíduo, por sua vez, mantém sua curva descendente uma vez que a estratégia de repovoamento mantém os melhores indivíduos na população.

A Tabela 8 detalha os erros robustos alcançados na melhor execução de cada versão do MSES-R20. Observamos que o menor erro robusto e o menor desvio padrão foram apresentados pela abordagem MSES-SR-50-25-25, com os valores 2,63 e 3,73, respectivamente. Já a menor média foi encontrada na estratégia MSES-SR-50-10-10-10-10-10. A partir disso concluímos que a configuração da estratégia MSES-R20 que obteve o melhor resultado foi a 50-25-25. Todas essas estratégias encontraram soluções melhores que a encontrada pelo PGA-FULL, apesar da média de resultados novamente ser pior. Se compararmos com a abordagem MSES-SR-50-30-20, que não realiza repovoamento, nenhuma das estratégias encontrou solução melhor e suas média também foram piores.

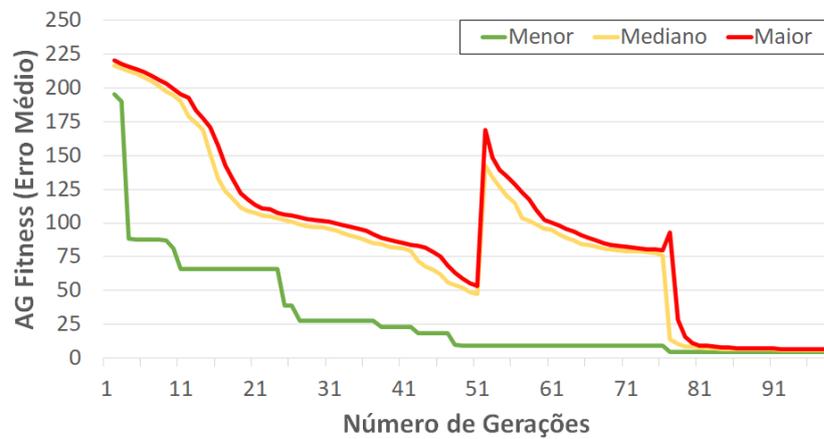
Tabela 8 – Erro médio e desvio padrão das execuções com as estratégias MSES-R20.

| Estratégia | Estágios | Menor Erro | Média | Desvio Padrão |
|------------|-------------------|------------|-------|---------------|
| PGA-FULL | - | 3,13 | 7,77 | 3,59 |
| MSES-SR | 50-30-20 | 2,50 | 9,21 | 3,95 |
| MSES-R20 | 50-25-25 | 2,63 | 12,09 | 3,73 |
| MSES-R20 | 50-10-10-10-10-10 | 2,67 | 11,55 | 3,76 |
| MSES-R20 | 50-30-20 | 2,84 | 11,85 | 4,92 |

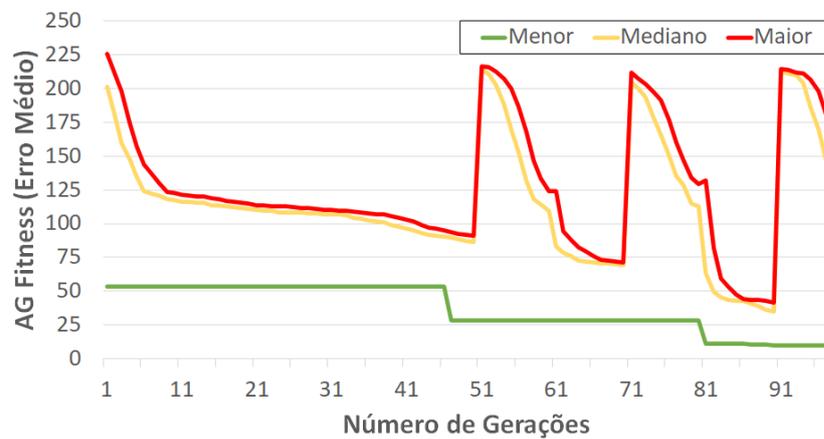
7.3.3 Análise da Estratégia MSES-RBEST

A estratégia multi-estágio baseada no melhor indivíduo (MSES-RBEST), altera a população entre os estágios da evolução, baseada nos valores de parâmetros do melhor indivíduo da população. A configuração de parâmetros para cada uma das estratégias MSES-RBEST que apresentou o melhor resultado dentre as 30 execuções é descrita na Tabela 9. Além disso, replicamos os resultados do modelo de referência e das estratégias PGA-FULL, MSES-SR-50-30-20 e MSES-R20-50-25-25 para facilitar as comparações.

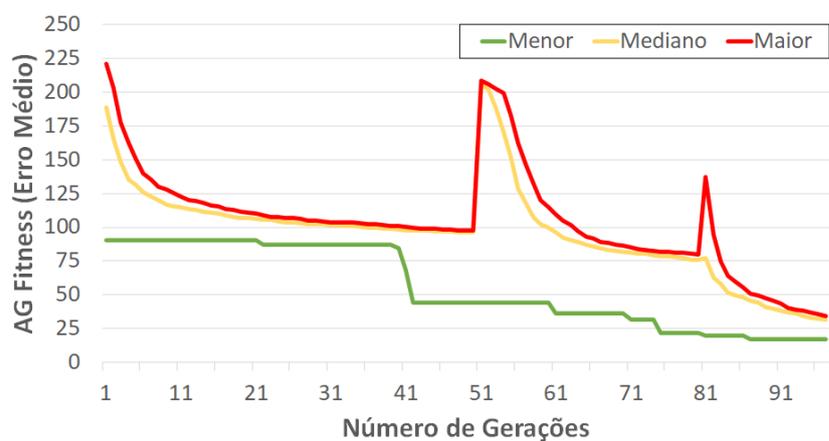
Na Figura44 podemos observar a dinâmica populacional de larvas e adultos para um período de 10 anos. No geral, o padrão de comportamento cíclico se assemelha ao esperado. O experimento com a estratégia MSES-RBEST-50-25-25 (Figura44(b)) alcançou picos de adultos bem semelhantes ao modelo de referência. Nos experimentos com as estratégias MSES-RBEST-50-10-10-10-10-10 e MSES-RBEST-50-30-20 ilustrados nas figuras Figura44(c) e Figura44(d) apresentam picos de infestação de adultos levemente meno-



(a) MSES-R20 - 50-25-25



(b) MSES-R20 - 50-10-10-10-10-10



(c) MSES-R20 - 50-30-20

Figura 43 – Evolução do erro para o melhor, pior e indivíduo mediano da população, considerando a melhor execução com as estratégias MSES-R20.

Tabela 9 – Melhor indivíduo e o erro robusto resultantes das execuções com as abordagens evolutivas: (1) PGA-FULL, (2) MSES-SR-50-30-20, (3) MSES-R20 - 50-25-25, (4) MSES-RBEST-50-25-25, (5) MSES-RBEST-50-10-10-10-10-10 e (6) MSES-RBEST-50-30-20.

| Configuração | p_r | p_{sy} | p_{sa} | p_d | D_i | D_{ni} | Q_f | Erro |
|--------------|---------|----------|----------|---------|---------|----------|-------|------|
| Referência | 0,00411 | 0,90272 | 0,98280 | 0,00415 | 0,9 | 0,1 | 95 | |
| 1 | 0,05386 | 0,19861 | 0,97204 | 0,56577 | 0,41054 | 0,66327 | 53 | 3,13 |
| 2 | 0,05067 | 0,25667 | 0,97438 | 0,24806 | 0,75753 | 0,23469 | 80 | 2,50 |
| 3 | 0,05354 | 0,090324 | 0,97883 | 0,4818 | 0,6165 | 0,6822 | 65 | 2,63 |
| 4 | 0,04971 | 0,15614 | 0,975 | 0,33131 | 0,74577 | 0,7047 | 83 | 2,57 |
| 5 | 0,05431 | 0,13303 | 0,97626 | 0,28489 | 0,94839 | 0,77879 | 100 | 2,38 |
| 6 | 0,05495 | 0,07808 | 0,97907 | 0,48088 | 0,9823 | 0,2631 | 99 | 2,37 |

res que o modelo de referência. O comportamento da larvas em todos os experimentos apresenta a mesma ordem de magnitude.

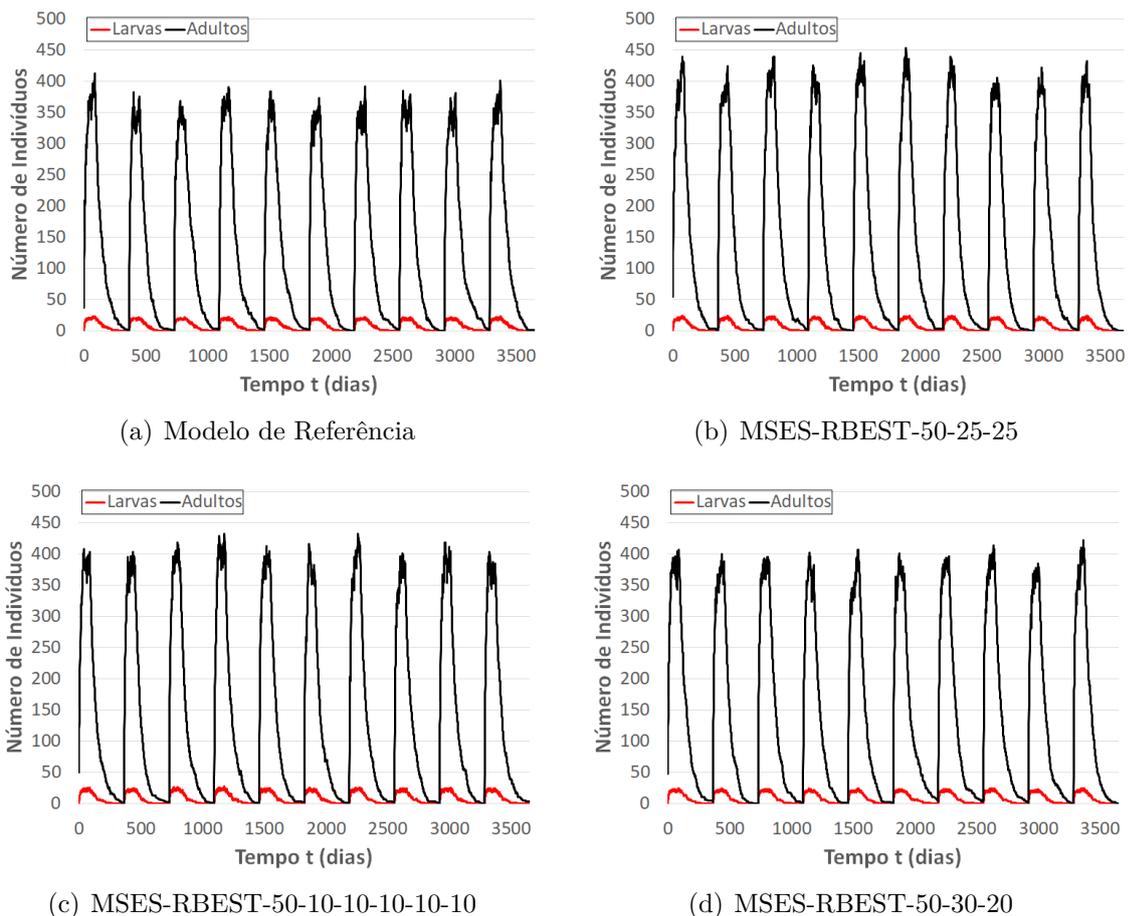


Figura 44 – Dinâmica populacional de insetos ao longo de 10 anos simulados pelo modelo, usando as configurações definidas pelas estratégias MSES-RBEST.

A evolução espacial dos adultos pode ser observada na Figura 45, nas abordagens MSES-RBEST analisadas. Podemos destacar que os reticulados apresentam comportamentos espaciais semelhantes no que diz respeito a dinâmica do modelo, com discretas

concentrações de adultos nas janelas de tempo dos dias 180º e 270º. Por outro lado, a Figura 46 mostra a evolução das larvas, onde todos os experimentos têm comportamentos espacialmente semelhantes.

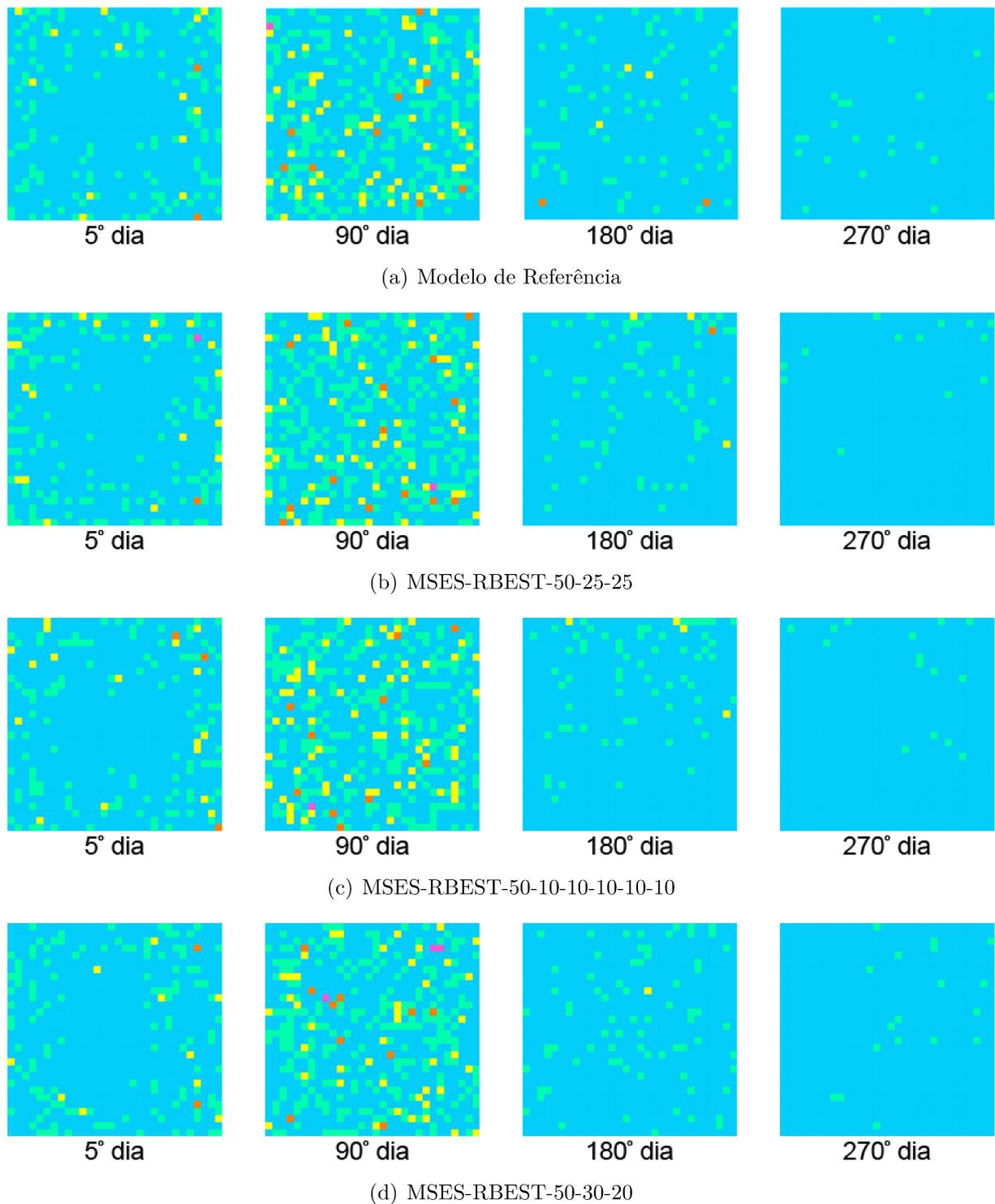


Figura 45 – Distribuição espacial dos adultos nos 5º, 90º, 180º e 270º dias de simulação do modelo, usando as configurações definidas pelas estratégias MSES-RBEST.

A Figura 47 apresenta a evolução do erro das execuções para o melhor, mediano e pior indivíduo, a cada geração e para cada estratégia MSES-RBEST analisada. De forma

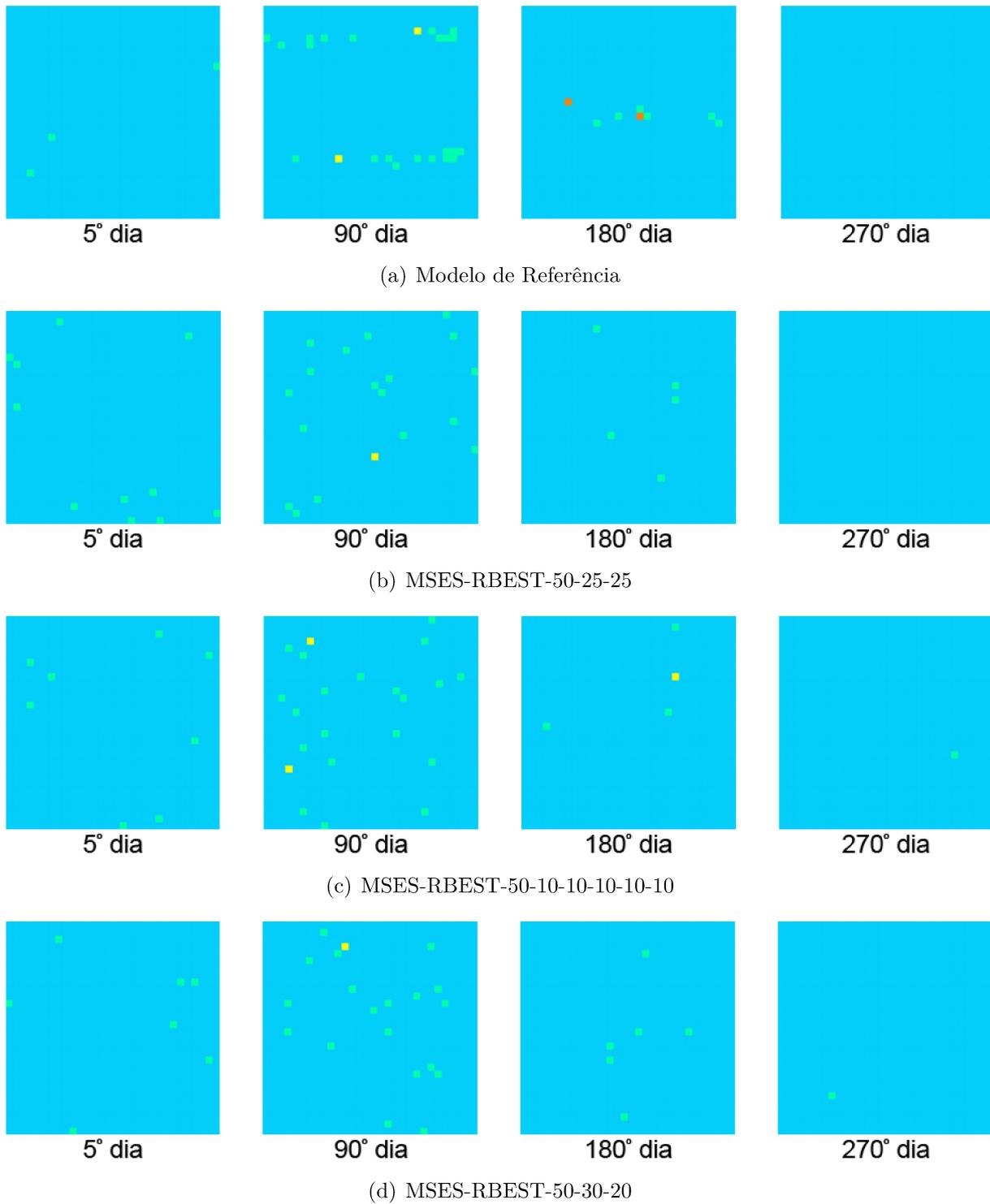


Figura 46 – Distribuição espacial das larvas nos 5º, 90º, 180º e 270º dias de simulação do modelo, usando as configurações definidas pelas estratégias MSES-RBEST.

geral, as curvas se comportam de forma similar ao observado na MSER-R20 apresentada na Figura 43. Ou seja, a curva do erro médio apresenta um declínio do erro até a geração 50 e depois disso, apresenta saltos com grandes aumentos de erro, exatamente nos momentos em que uma nova etapa é iniciada. Entretanto, os saltos de erro da MSER-RBEST, de uma forma geral, são um pouco menores que o observado na MSER-R20. Provavelmente, essa observação está relacionada ao fato que todos indivíduos preservam os parâmetros que são congelados idênticos a um único indivíduo (o Best), enquanto os demais parâmetros que serão ajustados na etapa são obtidos a partir de mutações aplicadas nos valores dos 20 melhores indivíduos. No MSER-20, por outro lado, são utilizados os parâmetros fixos de 20 indivíduos (ao invés de apenas o Best) e os parâmetros ajustáveis são reinicializados de forma aleatória. Portanto, embora apenas o melhor indivíduo seja totalmente preservado, os demais indivíduos são reinicializados de uma forma mais próxima ao Best e aos melhores indivíduos que já existiam na população. Assim, acreditamos que o repovoamento na MSER-RBEST foi implementado de uma forma um pouco mais conservadora que na MSER-20.

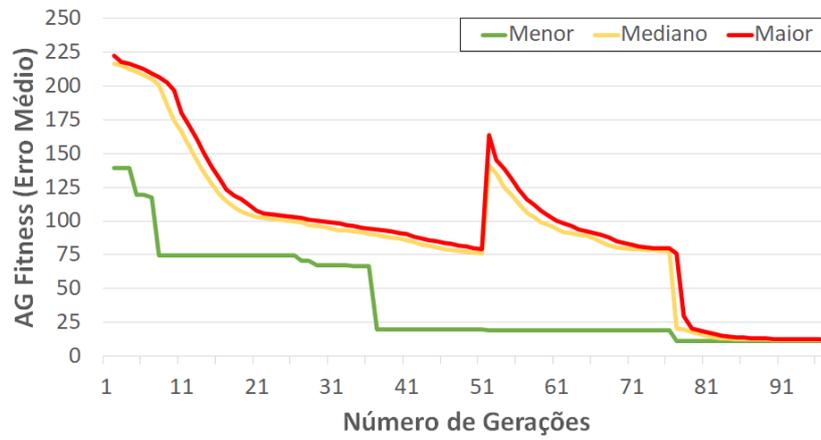
A Tabela 10 detalha os erros robustos alcançados na melhor execução de cada versão do MSES-RBEST. Observamos que o menor erro robusto, a menor média e o menor desvio padrão foram apresentados pela abordagem MSES-RBEST-50-30-20, com os valores 2,37, 7,69 e 4,22, respectivamente. Essa estratégia, apresentou solução melhor que a encontrada pelo PGA-FULL nas 3 configurações. A estratégia MSES-RBEST-50-30-20 apresentou o melhor resultado dentre todas as estratégias testadas e apresentou também a média mais baixa dentre todas as execuções que ajustam 7 parâmetros, sendo assim a melhor configuração encontrada para o ajuste completo dos parâmetros do modelo.

Tabela 10 – Erro médio e desvio padrão das execuções com as estratégias MSES-RBEST.

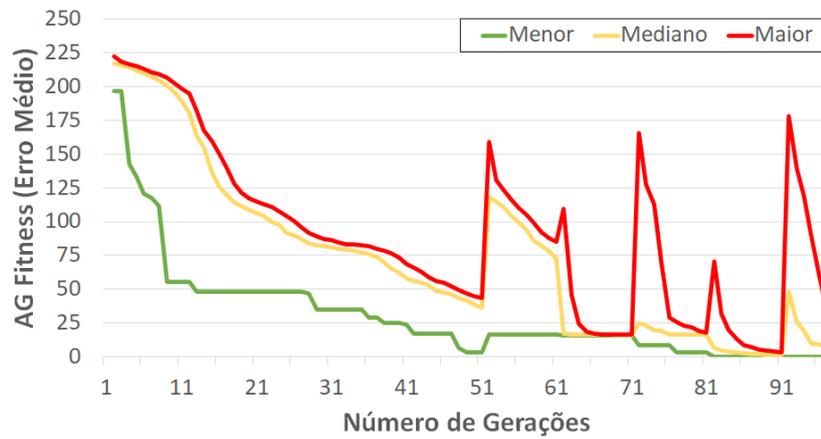
| Estratégia | Estágios | Menor Erro | Média | Desvio Padrão |
|------------|-------------------|------------|-------|---------------|
| PGA-FULL | - | 3,13 | 7,77 | 3,59 |
| MSES-SR | 50-30-20 | 2,50 | 9,21 | 3,95 |
| MSES-R20 | 50-25-25 | 2,63 | 12,09 | 3,73 |
| MSES-RBEST | 50-25-25 | 2,57 | 12,53 | 4,36 |
| MSES-RBEST | 50-10-10-10-10-10 | 2,38 | 13,47 | 5,17 |
| MSES-RBEST | 50-30-20 | 2,37 | 7,69 | 4,22 |

7.4 Avaliação Regional

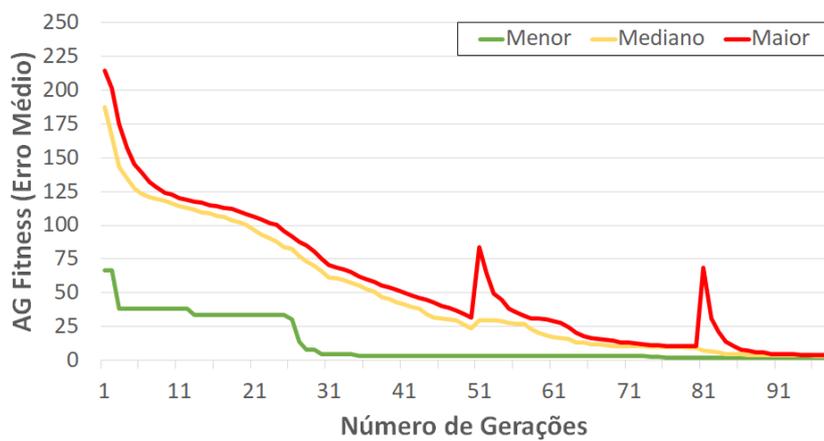
Essa seção apresenta uma variação na forma de avaliar a aptidão dos indivíduos do algoritmo genético, a fim de considerar a distribuição espacial dos insetos pelo reticulado. Isto é, a nova função de avaliação, descrita na seção 6.1.2.2, tem como objetivo avaliar não só a quantidade de insetos presentes no reticulado, mas também a sua localização.



(a) MSES-RBEST-50-25-25



(b) MSES-RBEST-50-10-10-10-10-10



(c) MSES-RBEST-50-30-20

Figura 47 – Evolução do erro para o melhor, pior e indivíduo mediano da população, considerando a melhor execução com as estratégias MSES-RBEST.

Nesse experimento foram utilizadas as estratégias PGA-FULL e MSES-RBEST-50-30-20, as quais serão chamadas de PGA e MSES, por simplicidade.

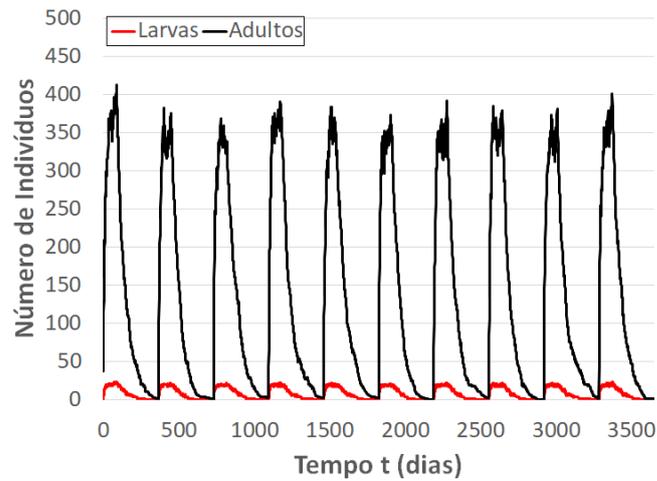
A ideia é analisar o impacto do uso da avaliação regional no desempenho das estratégias evolutivas (PGA e MSES). A configuração de parâmetros para cada uma das estratégias que apresentou o melhor resultado dentre as 30 execuções é descrita na Tabela 11. Como pode ser observado na tabela, exceto pela probabilidade de desenvolvimento (p_d), todos os demais parâmetros e o erro robusto estão em faixas próximas de valores, comparando-se as duas abordagens evolutivas.

Tabela 11 – Melhor indivíduo e erro robusto resultantes das execuções com as abordagens evolutivas usando a avaliação regional.

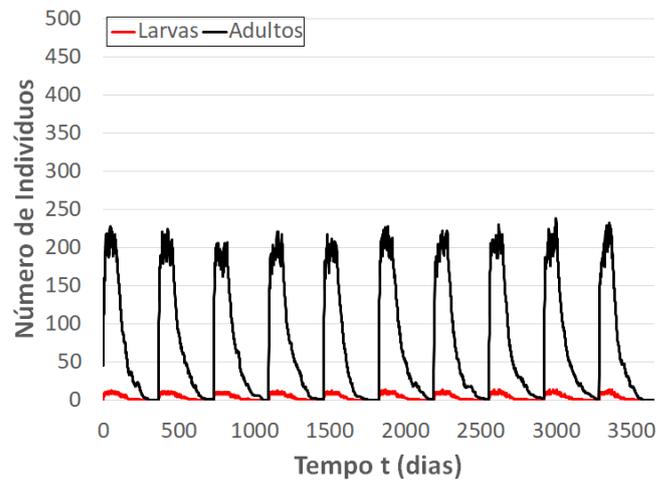
| Configuração | p_r | p_{sy} | p_{sa} | p_d | D_i | D_{ni} | Qf | Erro |
|--------------|---------|----------|----------|---------|---------|----------|------|------|
| Referência | 0,00411 | 0,90272 | 0,98280 | 0,00415 | 0,9 | 0,1 | 95 | |
| PGA | 0,03779 | 0,28074 | 0,97312 | 0,2242 | 0,8931 | 0,18769 | 94 | 1,69 |
| MSES | 0,04227 | 0,14098 | 0,97641 | 0,92962 | 0,90320 | 0,10725 | 95 | 1,57 |

A Figura 48 apresenta a dinâmica populacional de larvas e adultos para um período de 10 anos. O padrão de comportamento cíclico se assemelha ao esperado, porém as quantidades de insetos são bem menores que do modelo de referência (Figura 48(a)). A evolução populacional tanto da estratégia PGA (Figura 48(b)) quanto do MSES (Figura 48(c)), atingem picos de adultos próximos a 250 insetos, ou seja, em torno da metade da quantidade alcançada pelo modelo de referência (Figura 48(a)). Entretanto, o PGA apresenta um intervalo entre ciclos (largura de onda) ligeiramente mais estreito. O comportamento das larvas também apresenta uma redução considerável na quantidade de insetos em ambas as abordagens evolutivas.

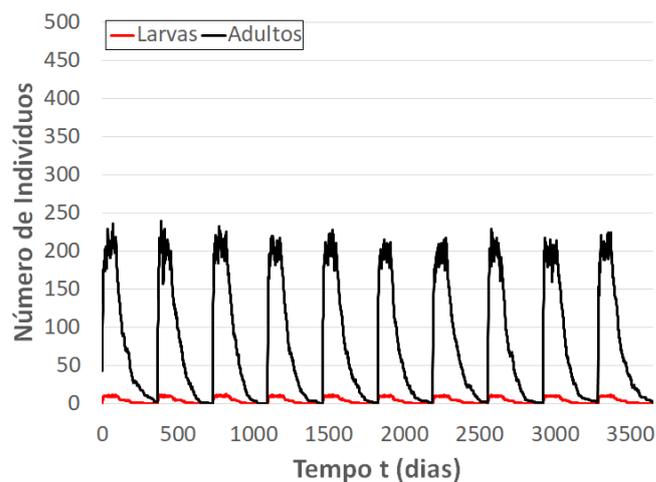
A evolução espacial dos insetos adultos simulada a partir da melhor configuração encontradas por cada abordagem analisada é apresentada na Figura 49. É possível destacar que os reticulados apresentam comportamentos espaciais semelhantes no que diz respeito ao espalhamento dos insetos, porém com redução da quantidade de adultos evidente no reticulado do 90º dia, quando ocorre o pico da infestação. Dentre as estratégias evolutivas, o PGA (Figura 49(b)) apresenta as menores quantidades de insetos em todos os quatro reticulados (instantes de tempo). Além disso, pode-se observar que os insetos se dispersam ligeiramente mais rápido que o comportamento esperado, o que acarreta no intervalo mais estreito entre ciclos observado no respectivo gráfico de evolução populacional (Figura 49(b)). Como pode ser observado na Figura 49(c), os reticulados referentes ao período de não infestação (180º e 270º dias) gerados a partir do MSES apresentam quantidades de adultos mais parecidos com o modelo de referência. Tal comportamento está relacionado com a semelhança na largura das respectivas curvas (intervalo entre ciclos) da dinâmica populacional. Apesar da diferença na quantidade de insetos ser evidente, principalmente no pico da infestação, acreditamos que a dinâmica do MSES consegue



(a) Modelo de Referência



(b) Algoritmo Genético Simples (PGA)



(c) Estratégia Evolutiva Multi-Estágio (MSES)

Figura 48 – Dinâmica populacional de insetos ao longo de 10 anos simulados pelo modelo, usando as configurações definidas pelas estratégias evolutivas com a avaliação regional.

representar, de forma satisfatória, a distribuição espacial dos insetos pelo reticulado. A dinâmica populacional das larvas, mostrada na Figura 50, também apresenta um comportamento espacial semelhante ao observado para os adultos. Isto é, as estratégias evolutivas com uma quantidade menor de insetos, principalmente no ápice da infestação (90º dia) e o PGA com uma dispersão ligeiramente mais rápida que o MSES, o que fica evidente nos reticulados do 180º e 270º dias.

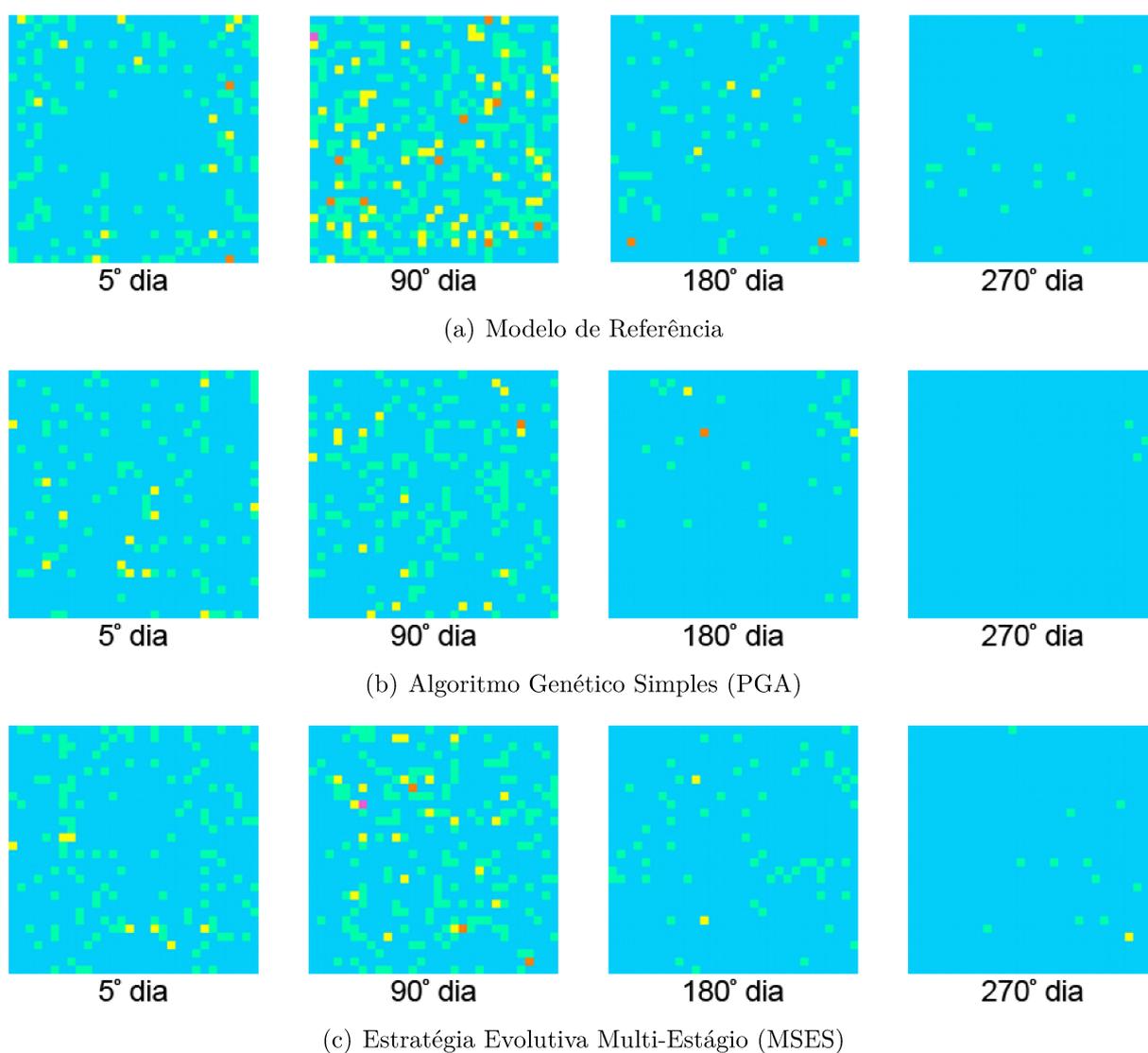


Figura 49 – Distribuição espacial dos adultos nos 5º, 90º, 180º e 270º dias de simulação do modelo, usando as configurações definidas pelas estratégias evolutivas com a avaliação regional.

A Figura 51 apresenta a evolução do erro para o melhor, mediano e pior indivíduos da população, a cada geração do AG. As abordagens apresentam um declínio do erro até a geração 50 para todos os indivíduos. No PGA (Figura 51(a)), observamos um decaimento gradativo do erro ao longo das 100 gerações. A curva do erro do melhor indivíduo apresenta alguns platôs, chegando a manter-se sem alteração por quase 40 gerações, mas voltando a cair na sequência. Ao final da execução, as curvas do pior

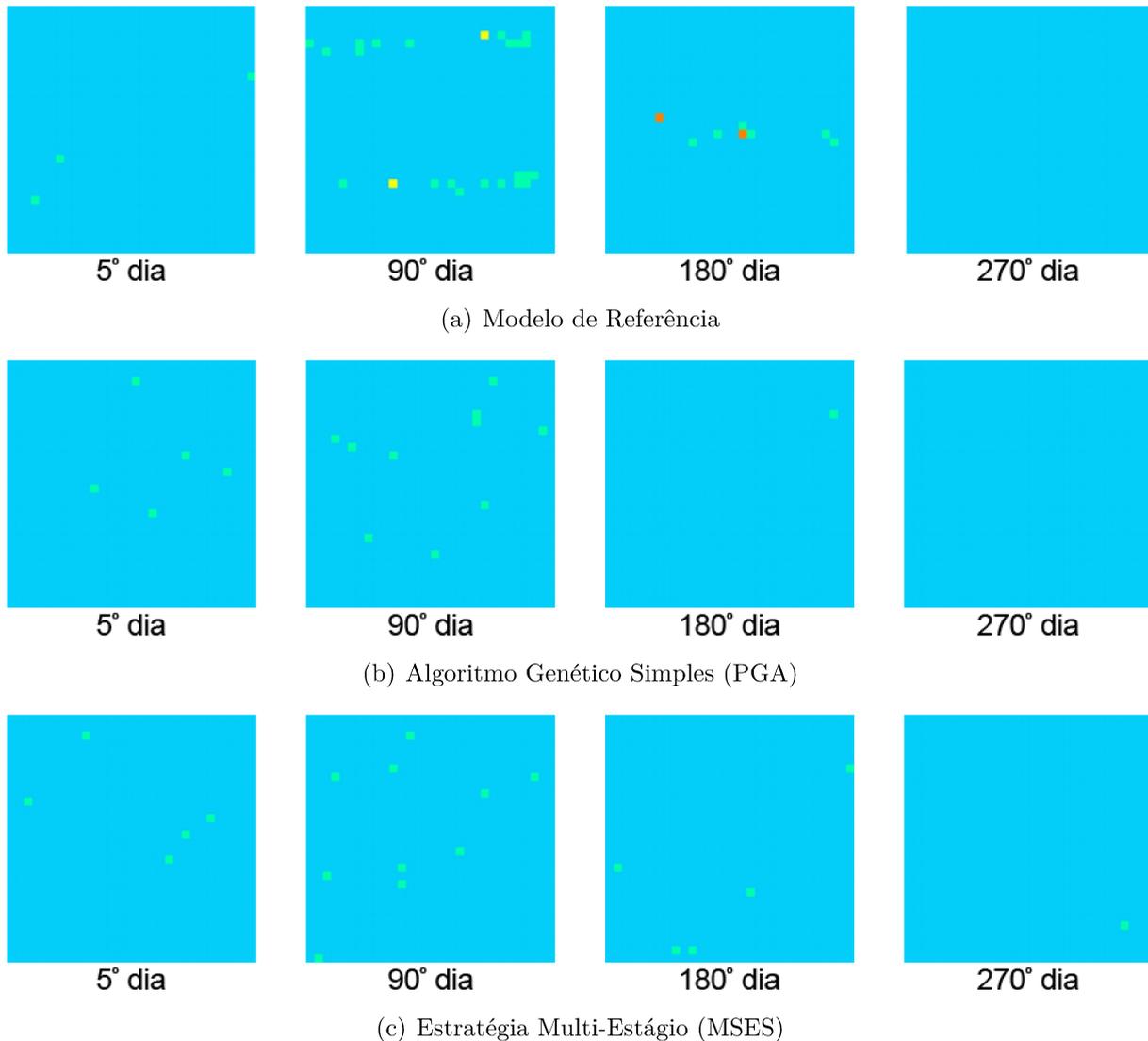
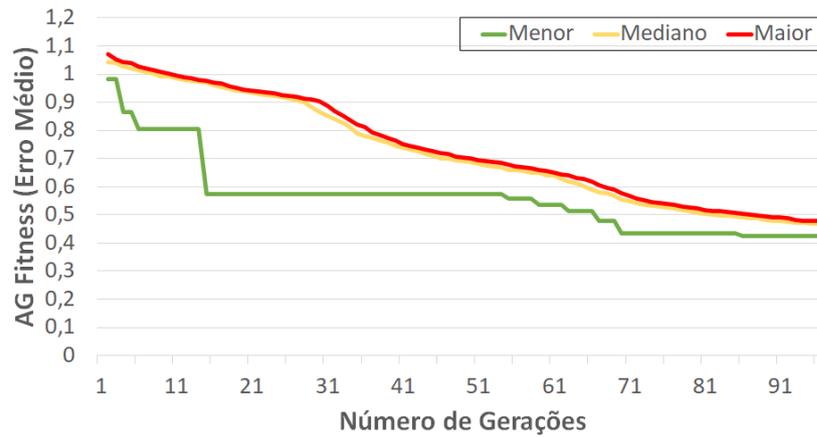


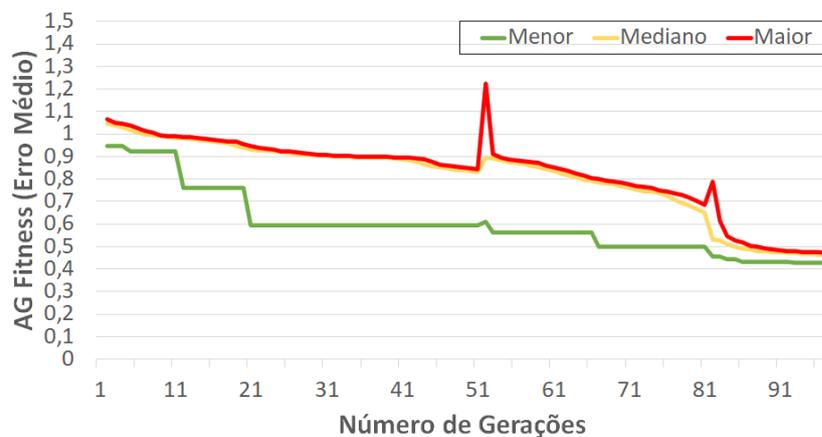
Figura 50 – Distribuição espacial das larvas nos 5°, 90°, 180° e 270° dias de simulação do modelo, usando as configurações definidas pelas estratégias evolutivas com a avaliação regional.

indivíduo e do mediano estão bem próximas e se aproximam do erro do melhor indivíduo do AG. Na estratégia MSES, o decaimento dos erros ocorre até a 50ª geração, quando acontece a primeira mudança de estágio. É possível observar dois picos nas curvas do pior indivíduo e do indivíduo mediano, referentes as mudanças de estágio da estratégia. Após essas mudanças, o algoritmo consegue retomar a convergência do erro, não só alcançando o nível anterior, como também aprimorando-o. Ao final da execução, todos os indivíduos apresentam erros bem próximos.

A Tabela 12 detalha os erros robustos alcançados na melhor execução de cada estratégia evolutiva (PGA e MSES). Observamos que o menor erro robusto foi obtido a partir da configuração gerada pela abordagem MSES, com o valor 1,57. Porém, o PGA apresenta um comportamento mais estável, resultando no menor erro médio e menor desvio padrão, com os valores 2,71 e 0,95, respectivamente.



(a) Algoritmo Genético Simples (PGA)



(b) Estratégia Multi-Estágio (MSES)

Figura 51 – Evolução do erro para o melhor, pior e indivíduo mediano da população, considerando a melhor execução das estratégias evolutivas com a avaliação regional.

Ao contrário do esperado, a adoção da avaliação regional comprometeu o desempenho do método. Embora as configurações evoluídas, principalmente pelo MSES, tenham reproduzido uma dinâmica populacional e distribuição espacial similares ao do modelo de referência, a quantidade de insetos, adultos e larvas, foi afetada significativamente. No geral, as configurações resultantes das abordagens evolutivas, principalmente do MSES, permitiram ao modelo de AC reproduzir uma dinâmica populacional (ciclo de evolução) e distribuição espacial dos insetos próximas àquelas obtidas pelo modelo de referência. Entretanto, as quantidades de insetos (larvas e adultos) são muito diferentes daquelas esperadas, alcançando uma queda em torno de 50% nos insetos adultos no ápice da infestação. Portanto, a avaliação regional se mostrou inadequada para aferir a aptidão das soluções (indivíduos) durante o processo evolutivo.

Tabela 12 – Erro médio e desvio padrão das execuções com a avaliação regional.

| Estratégia | Menor Erro | Média | Desvio Padrão |
|------------|------------|-------|---------------|
| PGA | 1,69 | 2,71 | 0,95 |
| MSES | 1,57 | 3,03 | 1,37 |

7.5 Análise de Robustez

Essa seção apresenta uma variação dos experimentos afim de fazer um teste de robustez da estratégia proposta. Para isso, aumentamos o tamanho do reticulado do AC de 30×30 para 60×60 , ou seja, passamos de uma representação de 900 células para uma representação de 3600 células, quatro vezes maior. Para criar a base de comparação, executamos o modelo em um reticulado 60×60 , com os valores de parâmetros descritos em (SLIMI et al., 2009) e usamos as quantidades retornadas de larvas e adultos (dia a dia) como valor de referência para avaliação da estratégia. O modelo original possui uma configuração de raio igual a 3 no período de infestação e 1 no período de não infestação. Isso quer dizer que um inseto adulto pode se mover da célula atual para uma célula vizinha que esteja a uma distância de até 3 células no período de infestação e 1 célula no período de não infestação. Consideramos dois possíveis cenários de teste, um em que mantemos os valores dos raios dos experimentos com reticulado 30×30 e outro onde dobramos os valores dos raios para possibilitar uma dispersão da mesma proporção observada anteriormente. A primeira configuração mantém a dinâmica do modelo original, porém aumentando a área do reticulado. Nesse caso, é como se o modelo fosse aplicado a uma aldeia de área quatro vezes maior, mas mantendo a característica da floresta no entorno de todo o reticulado. Por outro lado, podemos interpretar a segunda configuração como um aumento da resolução do reticulado, em que manteríamos as quantidades iniciais com uma precisão mais elevada. Para esses dois cenários, vamos executar a estratégia “PGA-FULL” e “MSES-RBEST-50-30-20”. Cada um dos cenários foi executado 30 vezes e os resultados aqui apresentados são referentes a melhor execução dentre elas.

A Figura 52 compara a dinâmica populacional dos vetores obtida a partir da simulação de 10 anos dos respectivos modelos de referência. A Figura 52(a) mostra o comportamento simulado pelo modelo de referência adotado nos experimentos anteriores, ou seja, que usa um reticulado de tamanho 30×30 . A evolução populacional apresentada na Figura 52(b) foi obtida a partir do aumento no tamanho do reticulado para 60×60 e mantendo-se os valores dos raios usados na etapa de dispersão (3 e 1). Como pode ser observado, essa modificação provocou um aumento consideravelmente nos picos de adultos dos ciclos, saindo de aproximadamente 450 insetos para valores próximos aos 650 insetos. Tal comportamento é esperado, uma vez que temos um número maior de células para os vetores (adultos e larvas) se propagarem. Como cada célula tem um limite de ocupação (5 adultos e 5 larvas), seria esperado um aumento de vetores ao se permitir um número

maior de células, com a mesma limitação. Por outro lado, é possível perceber que, como a distância aumentou, o número de dias do período de infestação não é suficiente para que um grande número de insetos chegue à região mais central do reticulado. Por isso, embora os picos tenham aumentado, a proporção não chega a ser 4 vezes do valor observado no modelo anterior. Por outro lado, considerando a dinâmica populacional representada na Figura 52(c), a qual foi obtida a partir do modelo de referência que adota um reticulado de tamanho 60×60 e dobra os valores dos raios da etapa de dispersão (6 e 2), as quantidades de insetos se assemelham ao modelo original (reticulado 30×30 e raios 3 e 1). Isso ocorre porque como dissemos anteriormente, ao aumentar os raios, aumentamos a dispersão e o modelo 60×60 se comporta como se fosse o modelo 30×30 com uma resolução ampliada.

Na Figura 53, analisamos a distribuição espacial dos adultos. Na Figura 53(b) observamos um aumento de insetos em relação a Figura 53(a) e uma concentração maior dos insetos nas bordas, principalmente do começo do período de infestação. Isso acontece porque os valores dos raios são mantidos em reticulados de tamanhos diferentes, logo, os insetos não conseguem alcançar as mesmas distâncias no mesmo intervalo de tempo. Na Figura 53(c), observamos um comportamento de espalhamento mais semelhante a Figura 53(a), devido a correspondência do tamanho do raio. Por outro lado, a cor das células representa uma menor aglomeração de insetos na mesma célula no dia de maior infestação (90º), quando comparamos as Figuras 53(a) e 53(c). Novamente, devemos lembrar que o segundo se comporta de forma similar ao primeiro com uma resolução maior. Assim, se analisarmos cada composição de células em uma janela 4×4 perceberíamos uma maior concentração de vetores no 90º dia, se aproximando ao observado na Figura 53(a).

7.5.1 Algoritmo Genético Simples

As quantidades de insetos gerados pelo modelo nos experimentos PGA-FULL de tamanho 60 podem ser observadas na Figura 54. O comportamento cíclico dos insetos é mantido como esperado com picos de adultos levemente menores em ambos os experimentos. Na Figura 54(b) temos picos de larvas levemente elevados em relação a referência (Figura 54(a)), mas no geral os experimentos reproduziram bem o comportamento esperado.

A Figura 55 apresenta a evolução do erro do melhor, mediano e pior indivíduo da população, geração por geração. Em ambas as execuções, o erro decresce continuamente ao longo das gerações. O experimento da Figura 55(b) converge mais rápido e para um valor de avaliação mais baixo.

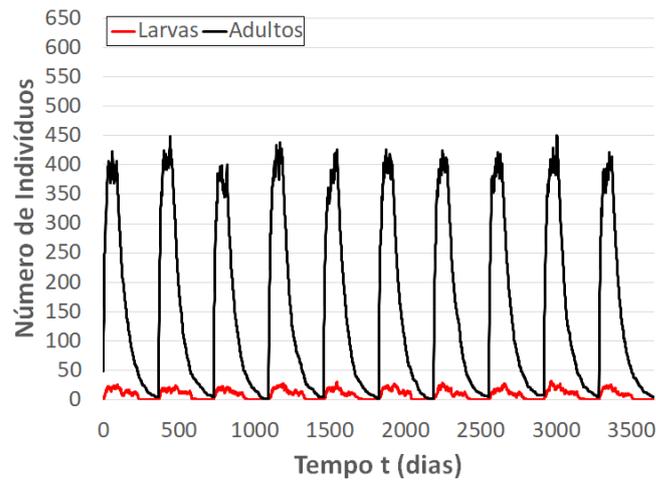
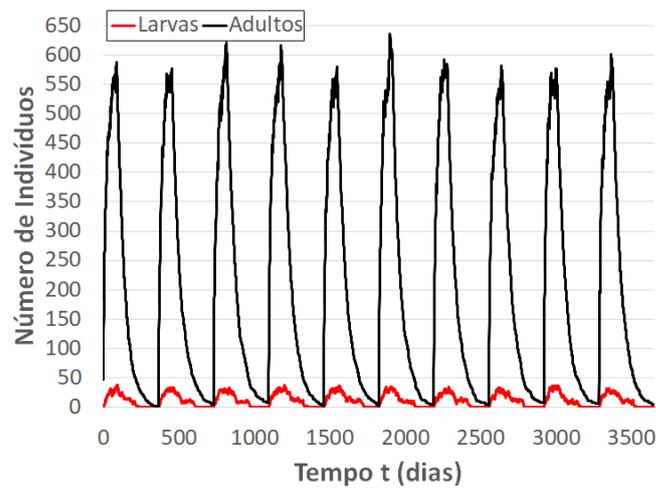
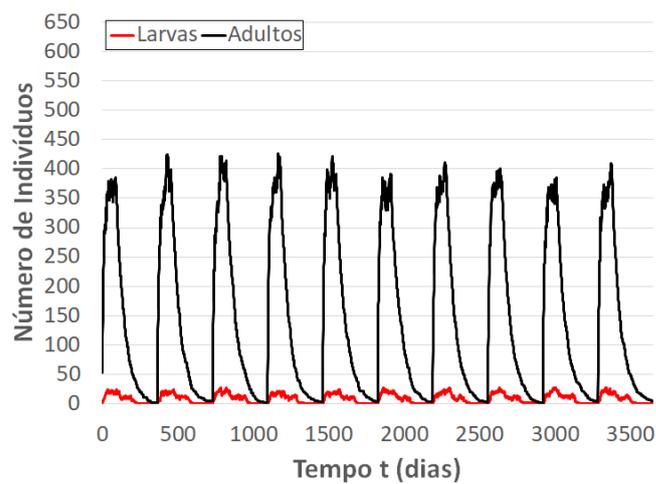
(a) Reticulado 30×30 - raios 3 e 1(b) Reticulado 60×60 - raios 3 e 1(c) Reticulado 60×60 - raios 6 e 2

Figura 52 – Dinâmica populacional de insetos ao longo de 10 anos simulados pelo modelo de referência, de acordo com o tamanho do reticulado e os raios adotadas na etapa de dispersão.

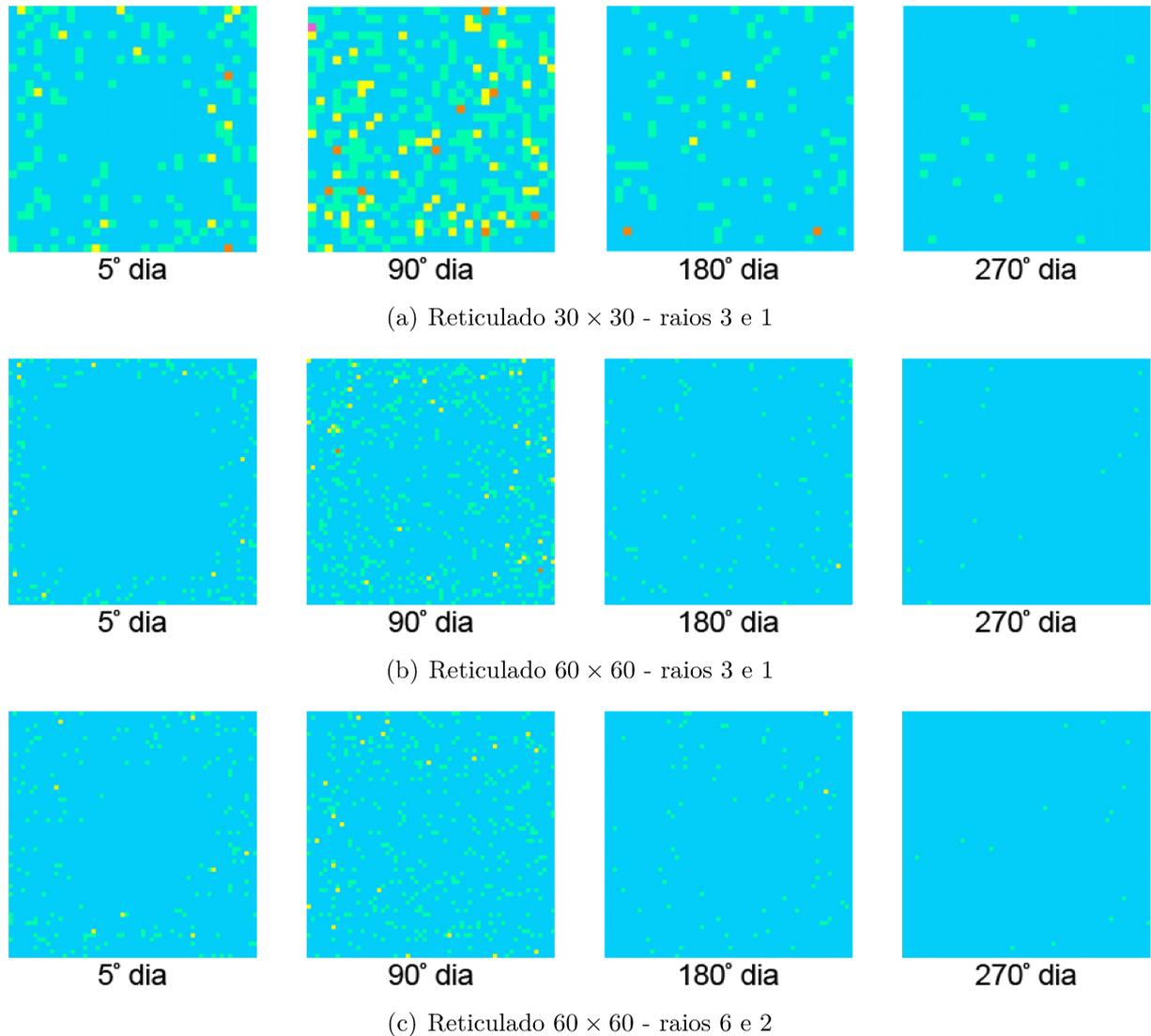
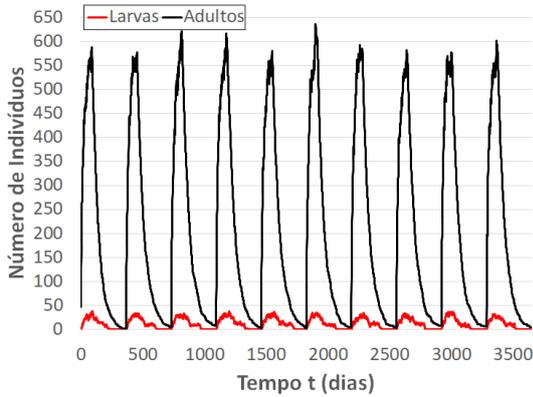


Figura 53 – Distribuição espacial dos adultos nos 5^o, 90^o, 180^o e 270^o dias de simulação, de acordo com a configuração usada no modelo de referência.

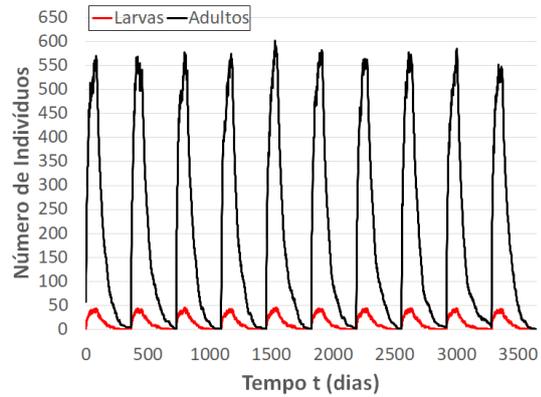
7.5.2 Estratégia Muti Estágio com Repovoamento Baseado no Melhor Indivíduo (MSES-RBEST)

A Figura 56 apresenta as quantidades de insetos gerados pelo modelo nos experimentos MSES-RBEST-50-30-20, novamente usando um reticulado 60x60. O padrão cíclico foi observado em todos os experimentos. Picos de adultos e larvas na mesma ordem de magnitude são apreciados em Figura 56(c) e Figura 56(d). A Figura 56(b) apresenta a curva de encerramento do ciclo mais disforme de todos os experimentos, o que indica que os insetos adultos se mantêm no reticulado por um período maior que o modelo de referência, porém só no final do ciclo.

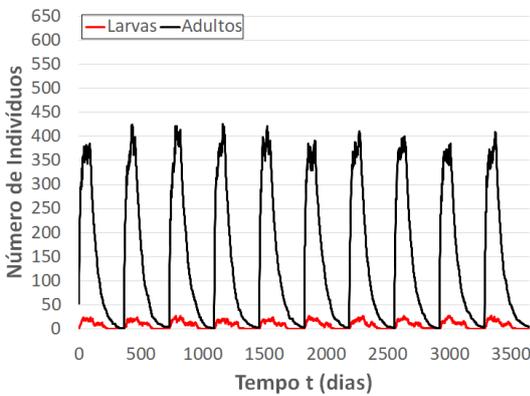
A evolução do erro durante as gerações dos experimentos MSES-RBEST-50-30-20 são apresentadas na Figura 57. Nos dois experimentos observamos uma diminuição contínua dos erros até a geração 50, onde o algoritmo muda de estágio alterando a população que



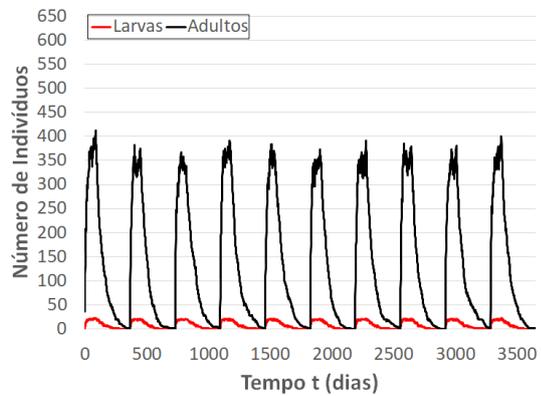
(a) Modelo de Referência - 60×60 e raios 3 e 1



(b) PGA - 60×60 e raios 3 e 1

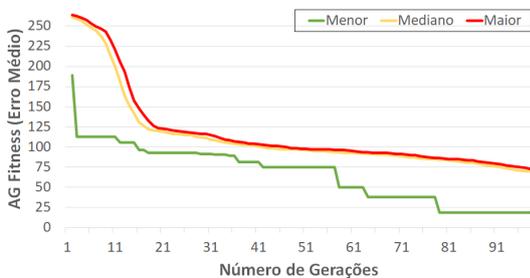


(c) Modelo de Referência - 60×60 e raios 6 e 2

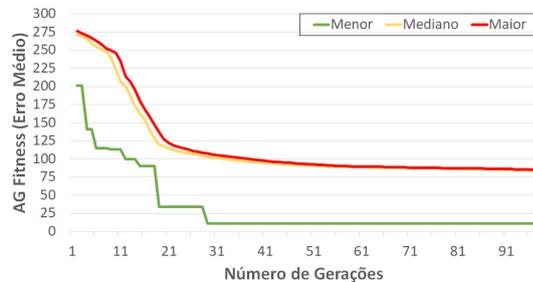


(d) PGA - 60×60 e raios 6 e 2

Figura 54 – Dinâmica populacional de insetos ao longo de 10 anos simulados com o reticulado 60×60 , usando as configurações definidas pelo PGA.



(a) Reticulado 60×60 - raios 3 e 1



(b) Reticulado 60×60 - raios 6 e 2

Figura 55 – Evolução do erro para o melhor, pior e indivíduo mediano da população, considerando a melhor execução do PGA para o reticulado 60×60 .

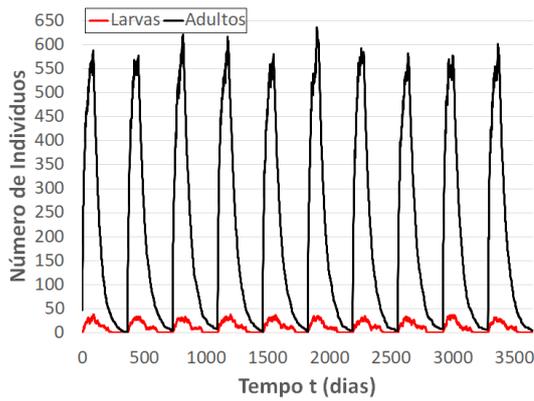
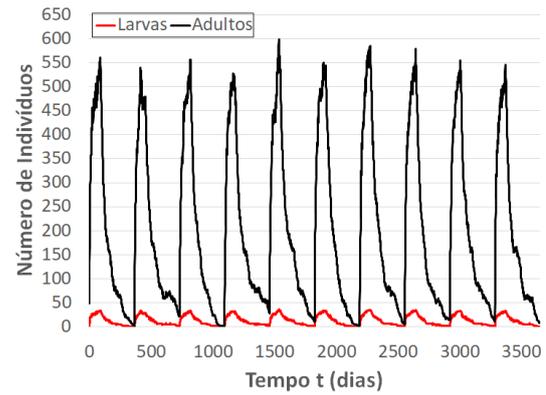
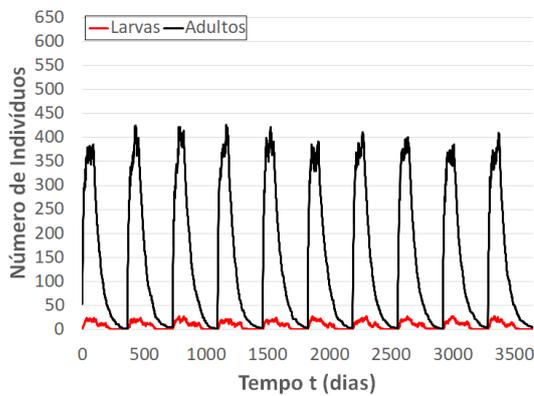
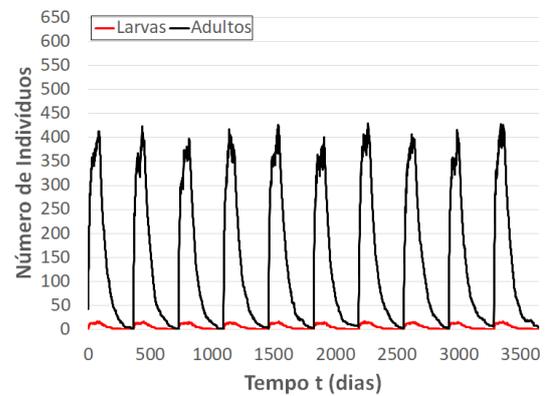
(a) Modelo de Referência - 60×60 e raios 3 e 1(b) MSES - 60×60 e raios 3 e 1(c) Modelo de Referência - 60×60 e raios 6 e 2(d) MSES - 60×60 e raios 6 e 2

Figura 56 – Dinâmica populacional de insetos ao longo de 10 anos simulados com reticulado 60×60 , usando as configurações definidas pelo MSES.

consequentemente sofre uma reavaliação. Como a rotina foi programada para mudar de estágio duas vezes, observamos dois picos de aumento de erro correspondentes a esses dois momentos.

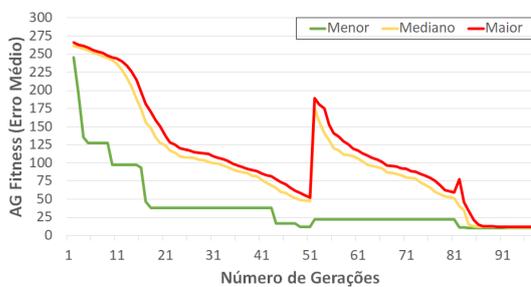
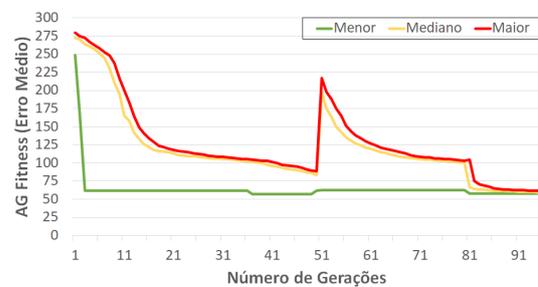
(a) Reticulado 60×60 - raios 3 e 1(b) Reticulado 60×60 - raios 6 e 2

Figura 57 – Evolução do erro para o melhor, pior e indivíduo mediano da população, considerando a melhor execução do MSES para o reticulado 60×60 .

7.5.3 Análise e Distribuição Espacial

A Tabela 13 apresenta os valores de parâmetros retornados pela melhor dentre as 30 execuções para cada uma das estratégias testadas (PGA-FULL e MSES-RBEST-50-30-20). A estratégia que apresentou o melhor resultado (menor erro) com reticulado 60×60 para os raios 3 e 1 foi a PGA, com valor 0,88. Para os raios 6 e 2, a que apresentou melhor resultado foi a MSES-RBEST-50-30-20, com valor 0,60. Os valores dos parâmetros ainda diferem dos valores de referência, porém considerando as demais análises podemos concluir que eles se compensam de alguma forma, pois simulam um comportamento bem semelhante ao esperado.

Tabela 13 – Melhor indivíduo e erro robusto resultantes das execuções com as abordagens evolutivas para os modelos com reticulado 60×60 .

| Modelo | Raio | p_r | p_{sy} | p_{sa} | p_d | D_i | D_{ni} | Qf | Erro |
|--------|-------|---------|----------|----------|---------|---------|----------|------|------|
| Refer | – | 0,00411 | 0,90272 | 0,98280 | 0,00415 | 0,9 | 0,1 | 95 | – |
| PGA | 3 e 1 | 0,06845 | 0,23078 | 0,97259 | 0,50397 | 0,84794 | 0,25771 | 94 | 0,89 |
| MSES | 3 e 1 | 0,06332 | 0,5183 | 0,95379 | 0,97212 | 0,7126 | 0,25767 | 79 | 1,47 |
| PGA | 6 e 2 | 0,05221 | 0,1785 | 0,97535 | 0,61178 | 0,93671 | 0,57319 | 95 | 0,70 |
| MSES | 6 e 2 | 0,03478 | 0,04924 | 0,98029 | 0,23179 | 0,83362 | 0,0866 | 90 | 0,60 |

As figuras Figura 58 e Figura59 apresentam a distribuição espacial dos insetos adultos. As figuras Figura60 e Figura61 dizem respeito a distribuição espacial das larvas. As quantidades e espalhamentos das estratégias testadas se assemelham ao esperado, o que classifica as representações como adequadas.

Analisando a Tabela 14, para os experimentos 60×60 com raios 3 e 1, a estratégia PGA-FULL apresentou o menor erro 0,89, porém a estratégia MSES-RBEST-50-30-20 apresentou a menor média e desvio padrão, com valores 4,14 e 1,51, respectivamente. A estratégia MSES-RBEST-50-30-20 teve picos de insetos adultos levemente menores que o modelo de referência e uma curvatura de final de ciclo menos suave, o que indica que os insetos demoram um pouco mais para dispersarem. Isso fica evidente na análise espacial (Figura 60(c)), onde o pico de infestação no 90º dia é menos povoado e no 270º dia a quantidade de insetos é maior do que a do modelo de referência. Nos experimentos 60×60 com raios 6 e 2, o menor erro, média e desvio padrão foram retornados pela MSES-RBEST-50-30-20, com valores 0,61, 3,52 e 1,09. Ela apresenta curvas populacionais mais próximas do modelo de referência e, considerando a distribuição espacial, tem um espalhamento e concentração de insetos também mais semelhantes. Embora haja diferenças de comportamento das abordagens entre os ambientes, ambas conseguem reproduzir de forma satisfatória o comportamento do modelo de referência.

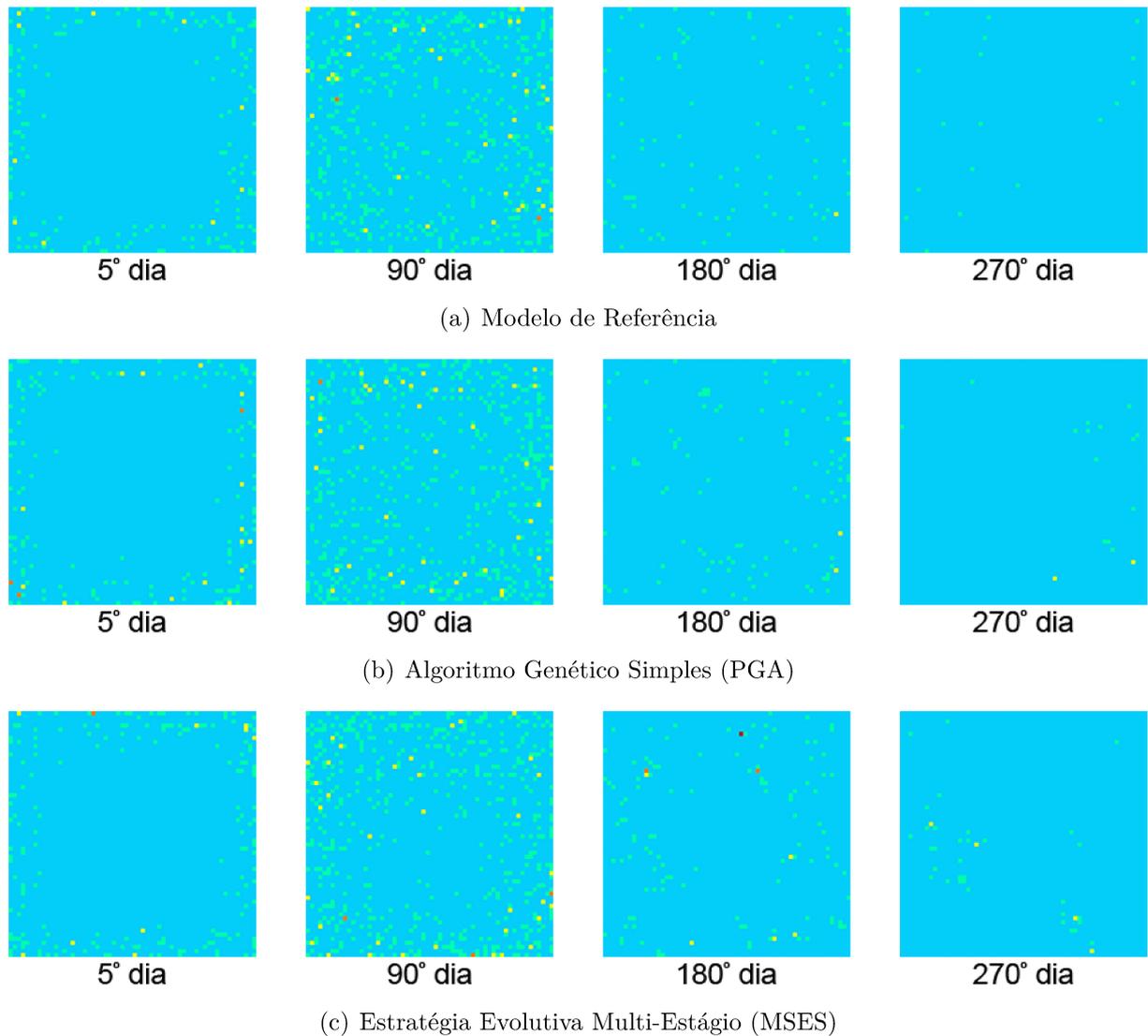


Figura 58 – Distribuição espacial dos adultos nos 5^o, 90^o, 180^o e 270^o dias de simulação do modelo com reticulado 60×60 e raios 3 e 1.

Tabela 14 – Erro médio e desvio padrão das execuções de acordo com a estratégia evolutiva e a configuração do modelo (tamanho do reticulado e raios de dispersão).

| Experimento | Reticulado | Raios | Menor Erro | Média | Desvio Padrão |
|-------------|----------------|-------|------------|-------|---------------|
| PGA | 30×30 | 3 e 1 | 3,13 | 7,77 | 3,59 |
| MSES | 30×30 | 3 e 1 | 2,37 | 7,69 | 4,22 |
| PGA | 60×60 | 3 e 1 | 0,89 | 4,47 | 1,65 |
| MSES | 60×60 | 3 e 1 | 1,48 | 4,14 | 1,51 |
| PGA | 60×60 | 6 e 2 | 0,71 | 3,82 | 1,47 |
| MSES | 60×60 | 6 e 2 | 0,61 | 3,52 | 1,09 |

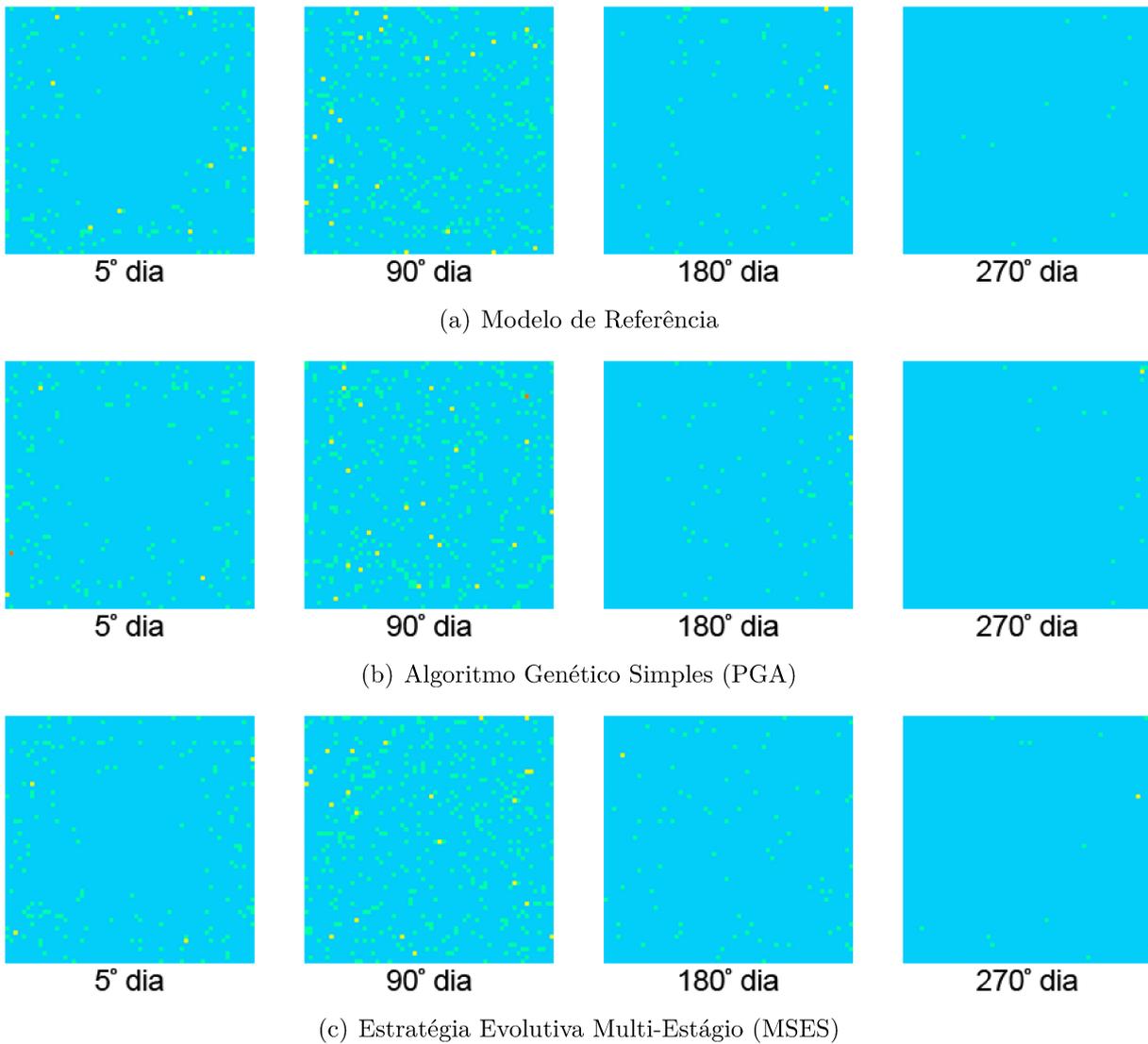


Figura 59 – Distribuição espacial dos adultos nos 5^o, 90^o, 180^o e 270^o dias de simulação do modelo com reticulado 60×60 e raios 6 e 2.

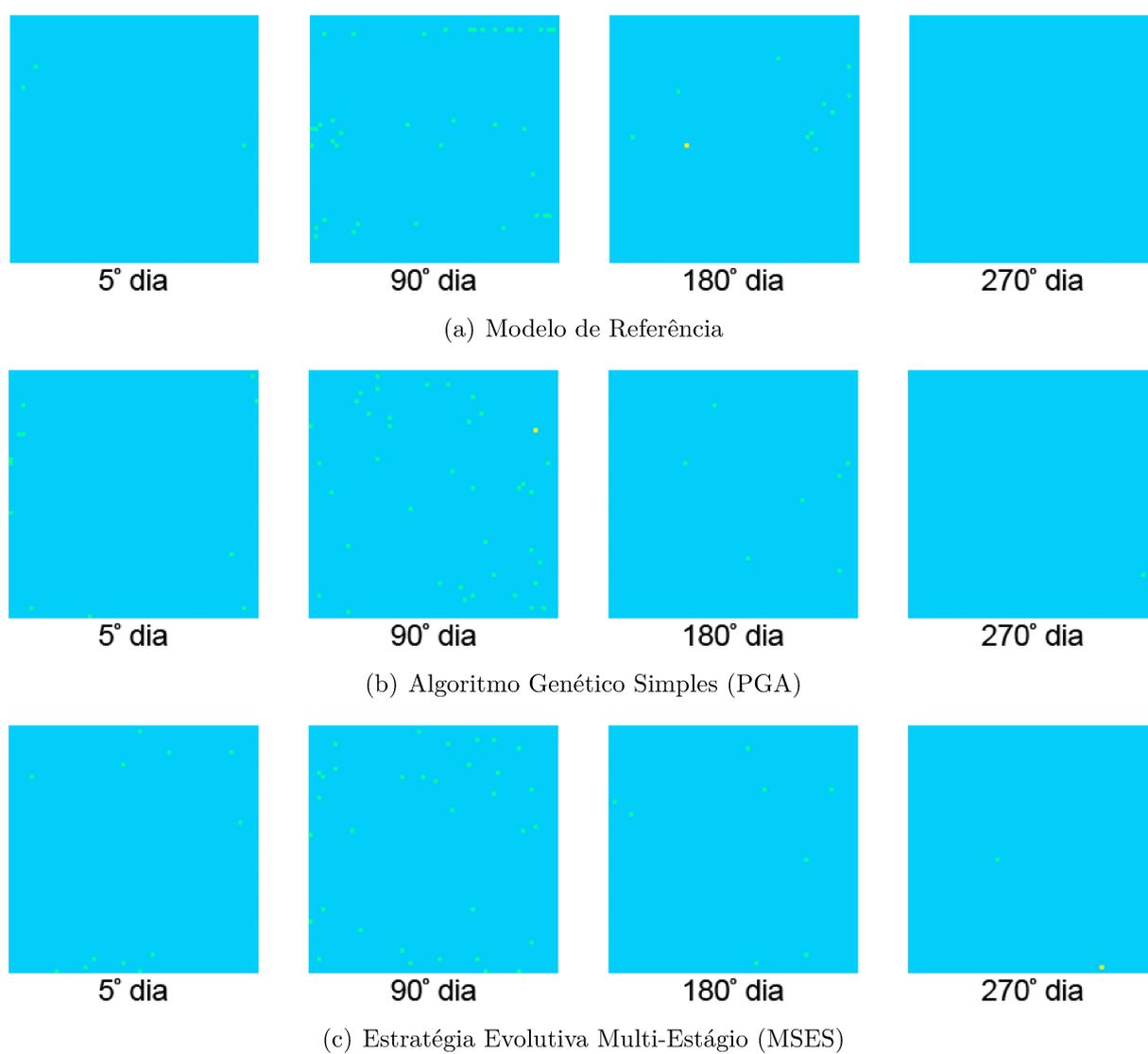


Figura 60 – Distribuição espacial das larvas nos 5^o, 90^o, 180^o e 270^o dias de simulação do modelo com reticulado 60×60 e raios 3 e 1.

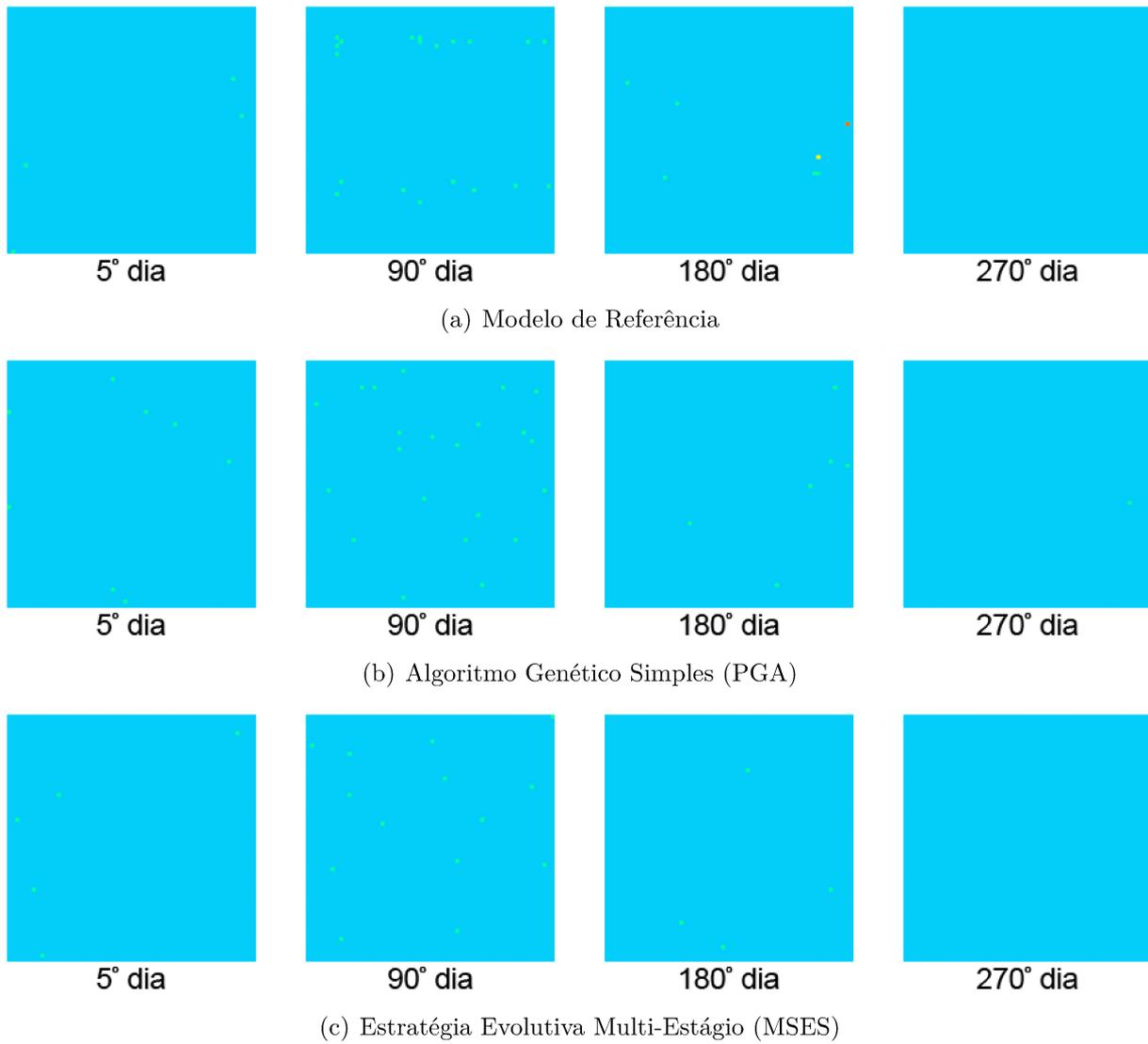


Figura 61 – Distribuição espacial das larvas nos 5º, 90º, 180º e 270º dias de simulação do modelo com reticulado 60×60 e raios 6 e 2.

Conclusão

Neste trabalho desenvolvemos duas abordagens baseadas em algoritmo genético para o ajuste dos parâmetros de um modelo epidemiológico baseado em autômato celular capaz de representar a evolução populacional de vetores da doença de Chagas (SLIMI et al., 2009).

A abordagem mais simples, denominada PGA (*Plain Genetic Algorithm*), encontrou parâmetros para o modelo capazes de reproduzir, de forma satisfatória, o comportamento dos vetores da doença ao longo de uma série temporal sintetizada. Os resultados das simulações mostram uma boa distribuição espacial e um número diário de insetos (larvas e adultos) muito próximo do modelo de referência, a partir do qual foram gerados os dados usados como entrada do algoritmo genético. Os valores dos parâmetros não se ajustam exatamente aos valores mencionados em (SLIMI et al., 2009). Entretanto, acredita-se que os parâmetros que alcançaram os valores mais próximos ao modelo de referência são os mais relevantes para a reprodução do modelo, enquanto os demais podem se adequar de forma combinatória, representando diferentes platôs no espaço de busca. Por exemplo, o parâmetro p_{sa} , relacionado à probabilidade de sobrevivência dos insetos adultos, mostrou-se o parâmetro mais sensível para a reprodução do comportamento, enquanto a probabilidade de desenvolvimento (p_d) apresentou uma maior variabilidade entre seus valores.

Também propusemos uma nova estratégia evolutiva baseada em múltiplos estágios (MSES), e tal proposta foi capaz de evoluir o conjunto completo de parâmetros, retornando um comportamento simulado mais próximo aos dados de referência que sua versão mais simples (PGA). As simulações mantiveram-se com uma distribuição espacial satisfatória e um número de larvas e adultos muito próximo do esperado. Os parâmetros encontrados pelo MSES também não foram ajustados exatamente aos valores de referência, porém alguns deles estão próximos. Dentre as estratégias multi-estágio e de repovoamento, no geral, a abordagem MSES RBEST-50-30-20 foi a que apresentou o melhor desempenho, encontrando a solução com o menor erro robusto e também tendo o menor erro médio do conjunto de experimentos. A configuração encontrada por essa abordagem

tem valores de parâmetros mais próximos da referência, principalmente os relacionados com a etapa de dispersão e a probabilidade de sobrevivência (p_{sa}).

Embora os experimentos aqui relatados explorem a aplicação de uma busca evolutiva em múltiplos estágios para um problema específico - o ajuste de parâmetros para um modelo epidemiológico baseado em AC - o conhecimento da abordagem investigada pode ser utilizado por outros pesquisadores da computação evolutiva em diferentes problemas, relacionados ao ajuste de modelos de parâmetros ou mesmo em aplicações mais distantes. A abordagem de múltiplos estágios pode ser particularmente interessante para problemas em que a busca evolutiva converge rapidamente para valores razoáveis, mas não suficientemente satisfatórios - devido à complexidade intrínseca do espaço de busca. Nesses problemas, a divisão em subproblemas e utilizá-los para realizar a pesquisa genética em etapas pode ser uma boa estratégia para contornar os obstáculos. Além disso, a diversidade incluída no processo de busca por meio das estratégias de repovoamento pode ser uma boa ferramenta para impulsionar a convergência evitando atratores sub-ótimos.

Um teste de robustez foi realizado, no qual se aumentou o tamanho do reticulado e conseqüentemente a complexidade do problema. Os reticulados do AC mudaram de 30×30 para 60×60 , passando de uma representação de 900 células para outra com 3600 células, ou seja, quatro vezes maior. A estratégia MSES-RBEST-50-30-20 retornou o menor erro robusto quando os raios da etapa de dispersão também são dobrados, passando para 6 e 2, respectivamente. Por outro lado, a estratégia PGA-FULL obteve a melhor configuração do modelo quando se mantém os valores originais dos raios (3 e 1). Entretanto, no geral, ambas abordagens evolutivas conseguem reproduzir o modelo de forma satisfatória, tanto em relação às quantidades de insetos, quanto à sua distribuição espacial.

Também foi investigada uma nova forma de avaliação que contabiliza os insetos por regiões do reticulado do AC, a qual foi denominada avaliação regional. Ela busca avaliar não só a quantidade de insetos, mas também a sua distribuição espacial. Os resultados experimentais mostraram uma queda significativa na quantidade de insetos, se tornando mais evidente no pico da infestação (no 90º dia), no qual contabilizamos uma redução em torno de 50% no total de insetos adultos em relação ao modelo de referência. Dentre as abordagens evolutivas, a MSES apresentou o menor erro robusto e o comportamento (populacional e espacial) mais próximo ao desejado. No geral, o método conseguiu uma dinâmica populacional e uma distribuição espacial parecidas com o modelo de referência.

8.1 Trabalhos Futuros

Uma extensão natural desta pesquisa é o refinamento da estratégia evolutiva de múltiplos estágios, investigando especificações de outros estágios e estratégias de repovoamento a serem aplicados durante as mudanças de estágio. Outro aspecto interessante para investigação futura é analisar como a dependência (ou independência) entre os parâmetros de

entrada pode afetar a estratégia evolutiva de múltiplos estágios. Também é importante avaliar o desempenho do MSES no ajuste de parâmetros de outros sistemas dinâmicos baseados em autômatos celulares, tanto na modelagem epidemiológica, quanto em outros campos, como na robótica autônoma, modelagem de crescimento de árvores, entre outros. Considerando que os resultados iniciais com a estratégia de avaliação por região não foram promissores, se faz necessária uma investigação mais aprofundada de funções de aptidão capazes de considerar a distribuição espacial dos insetos pelo reticulado em vez de apenas contabilizar o número total de adultos e larvas. Esse estudo é essencial para entender o que ocasionou a queda de desempenho do modelo no experimento realizado, bem como buscar definir novas formas de avaliação que aumentem a eficiência e eficácia do método.

Por fim, considerando a obtenção de dados reais de regiões endêmicas do Brasil junto a órgãos de saúde e outras instituições de pesquisa, é importante verificar como o método é afetado pelos ruídos presentes nesses dados. Em outras palavras, se ele é capaz de determinar os valores para os parâmetros de configuração a fim de se reproduzir adequadamente o comportamento desses dados.

8.2 Produção Bibliográfica Originada da Dissertação

Nossa pesquisa resultou em dois artigos em eventos científicos internacionais (FRAGA; OLIVEIRA; MARTINS, 2020) e (FRAGA; OLIVEIRA; MARTINS, 2021).

Em (FRAGA; OLIVEIRA; MARTINS, 2020) são apresentados os resultados alcançados pelo nosso método baseado no algoritmo genético simples (PGA) para o ajuste de 3, 4 e 7 parâmetros do modelo epidemiológico estudado.

Em (FRAGA; OLIVEIRA; MARTINS, 2021) é proposto nosso método evolutivo multi-estágio, são avaliadas as diferentes estratégias de evolução e repovoamento, e os resultados são comparados com o PGA.

Referências

BATISTELLA, M.; MORAN, E. F. Dimensões humanas do uso e cobertura das terras na amazônia: uma contribuição do IBA. **Acta Amazonica**, v. 35, n. 2, p. 239–247, 2005. Disponível em: <<https://doi.org/10.1590/S0044-59672005000200014>>.

BONE, C.; DRAGICEVIC, S.; ROBERTS, A. A fuzzy-constrained cellular automata model of forest insect infestations. **Ecological Modelling**, Elsevier, v. 192, n. 1-2, p. 107–125, 2006. Disponível em: <<https://doi.org/10.1016/j.ecolmodel.2005.09.013>>.

BREUKELAAR, R.; BÄCK, T. Using a genetic algorithm to evolve behavior in multi dimensional cellular automata: emergence of behavior. In: **Proceedings of the 7th annual conference on Genetic and evolutionary computation**. ACM Press, 2005. p. 107–114. Disponível em: <<https://doi.org/10.1145/1068009.1068024>>.

CARNEIRO, M. G.; OLIVEIRA, G. Synchronous cellular automata-based scheduler initialized by heuristic and modeled by a pseudo-linear neighborhood. **Natural Computing**, Springer, v. 12, n. 3, p. 339–351, 2013. Disponível em: <<https://doi.org/10.1007/s11047-013-9375-8>>.

CHAGAS, C. Nova tripanozomíase humana: estudos sobre a morfologia e o ciclo evolutivo do *Schizotrypanum cruzi* n. gen., n. sp., agente etiológico de nova entidade morbida do homem. **Memórias do Instituto Oswaldo Cruz**, v. 1, n. 2, p. 159–218, 1909. Disponível em: <<https://doi.org/10.1590/S0074-02761909000200008>>.

CHANG, C.; LO, S.; YU, S. Applying fuzzy theory and genetic algorithm to interpolate precipitation. **Journal of Hydrology**, Elsevier, v. 314, n. 1-4, p. 92–104, 2005. Disponível em: <<https://doi.org/10.1016/j.jhydrol.2005.03.034>>.

CHEN, Q.; MYNETT, A. E. Modelling algal blooms in the dutch coastal waters by integrated numerical and fuzzy cellular automata approaches. **Ecological Modelling**, Elsevier, v. 199, n. 1, p. 73–81, 2006. Disponível em: <<https://doi.org/10.1016/j.ecolmodel.2006.06.014>>.

CHEN, S. H.; JAKEMAN, A. J.; NORTON, J. P. Artificial intelligence techniques: An introduction to their use for modelling environmental systems. **Mathematics and Computers in Simulation**, v. 78, p. 379–400, 2008. Disponível em: <<https://doi.org/10.1016/j.matcom.2008.01.028>>.

- CISSÉ, B.; YACOUBI, S. E.; GOURBIERE, S. The basic reproduction number for chagas disease transmission using cellular automata. In: **Proceedings of the International Conference on Cellular Automata**. Krakow, Poland: Springer, 2014. p. 278–287. Disponível em: <https://doi.org/10.1007/978-3-319-11520-7_29>.
- CONWAY, J. The game of life. **Scientific American**, v. 223, n. 4, p. 4, 1970.
- COPPIN, P. et al. Digital change detection methods in ecosystem monitoring: a review. **International Journal of Remote Sensing**, v. 10, p. 1565–1596, 2004. Disponível em: <<https://doi.org/10.1080/0143116031000101675>>.
- CORTES, O. A. C.; SAAVEDRA, O. R. Estratégias evolutivas paralelas em otimização multimodal. **INFOCOMP**, v. 2, n. 1, p. 63–68, 2004.
- COURA, J. R.; DIAS, J. C. P. Epidemiology, control and surveillance of chagas disease: 100 years after its discovery. **Memórias do Instituto Oswaldo Cruz**, v. 104, p. 31–40, 2009. Disponível em: <<https://doi.org/10.1590/S0074-02762009000900006>>.
- D’AMBROSIO, D. et al. A cellular automata model for soil erosion by water. **Physics and Chemistry of the Earth, Part B: Hydrology, Oceans and Atmosphere**, Elsevier, v. 26, n. 1, p. 33–39, 2001. Disponível em: <[https://doi.org/10.1016/S1464-1909\(01\)85011-5](https://doi.org/10.1016/S1464-1909(01)85011-5)>.
- D’Ambrosio, D. et al. A macroscopic collisional model for debris-flows simulation. **Environmental Modelling & Software**, v. 22, p. 1417–1436, 2007. Disponível em: <<https://doi.org/10.1016/j.envsoft.2006.09.009>>.
- D’Ambrosio, D.; SPATARO, W. Parallel evolutionary modelling of geological processes. **Parallel Computing**, v. 33, p. 186–212, 2007. Disponível em: <<https://doi.org/10.1016/j.parco.2006.12.003>>.
- DARWIN, C. **On the Origin of Species**. London, England: John Murray, 1859.
- EIBEN, A. E.; SMITH, J. E. et al. **Introduction to evolutionary computing**. Berlin, Germany: Springer, 2003. v. 53. Disponível em: <<https://doi.org/10.1007/978-3-662-05094-1>>.
- FRAGA, L. M.; OLIVEIRA, G. M. B.; MARTINS, L. G. A. Adjustment of an epidemiological cellular automata-based model using genetic algorithm. In: **Proceedings of the 2020 IEEE 32nd International Conference on Tools with Artificial Intelligence**. Baltimore, MD, USA: IEEE, 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: **Proceedings of the 2021 IEEE Congress on Evolutionary Computation**. Kraków, Poland: IEEE, 2021. p. 466–473. Disponível em: <<https://doi.org/10.1109/CEC45853.2021.9504738>>.
- FU, S.; MILNE, G. Epidemic modelling using cellular automata. In: **Proceedings of the Australian conference on artificial life**. [S.l.: s.n.], 2003.
- GAGLIARDI, H.; ALVES, D. Small-world effect in epidemics using cellular automata. **Mathematical Population Studies**, v. 17, n. 2, p. 79–90, 2010. Disponível em: <<https://doi.org/10.1080/08898481003689486>>.

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

GOLDBERG, D. E. **Genetic algorithms in search, optimization, and machine learning**. Ma, New York: Addison-Wesley, 1989.

GOLDBERG, D. E.; HOLLAND, J. H. **Genetic algorithms and machine learning**. Kluwer Academic Publishers, 1988. Disponível em: <<https://doi.org/10.1023/A:1022602019183>>.

HOLLAND, J. H. Adaptation in natural and artificial systems. **Control and artificial intelligence**, University of Michigan Press, 1975.

HOLLAND, J. H. et al. **Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence**. Cambridge, MA, United States: MIT press, 1992. Disponível em: <<https://doi.org/10.7551/mitpress/1090.001.0001>>.

JR, A. R.; RASSI, A.; MARIN-NETO, J. A. Chagas disease. **The Lancet**, Elsevier, v. 375, n. 9723, p. 1388–1402, 2010. Disponível em: <[https://doi.org/10.1016/S0140-6736\(10\)60061-X](https://doi.org/10.1016/S0140-6736(10)60061-X)>.

JUILLE, H.; POLLACK, J. B. Coevolving the "ideal" trainer: Application to the discovery of cellular automata rules. In: **University of Wisconsin**. Madison, WI, USA: Citeseer, 1998.

KAIMOWITZ, D.; ANGELSEN, A. **Economic Models of Tropical Deforestation: A Review**. Bogor, Indonesia, 1998.

LAL, S. P.; YAMADA, K.; ENDO, S. Studies on motion control of a modular robot using cellular automata. In: **Proceedings of the Australasian Joint Conference on Artificial Intelligence**. Hobart, Australia: Springer, 2006. p. 689–698. Disponível em: <https://doi.org/10.1007/11941439_73>.

LAMBIN, E. F. **Modelling Deforestation processes, a review**. Luxembourg, 1994.

MEDEIROS, L. C. de C. et al. Modeling the dynamic transmission of dengue fever: investigating disease persistence. **PLOS neglected tropical diseases**, Public Library of Science San Francisco, USA, v. 5, n. 1, p. e942, 2011. Disponível em: <<https://doi.org/10.1371/journal.pntd.0000942>>.

MIKLER, A. R.; VENKATACHALAM, S.; ABBAS, K. Modeling infectious diseases using global stochastic cellular automata. **Journal of Biological Systems**, World Scientific, v. 13, n. 04, p. 421–439, 2005. Disponível em: <<https://doi.org/10.1142/S0218339005001604>>.

MITCHELL, M. **An introduction to genetic algorithms**. Cambridge, Massachusetts, USA: MIT press, 1998. Disponível em: <<https://doi.org/10.7551/mitpress/3927.001.0001>>.

- MITCHELL, M.; CRUTCHFIELD, J. P.; HRABER, P. T. Evolving cellular automata to perform computations: Mechanisms and impediments. **Physica D: Nonlinear Phenomena**, Elsevier, v. 75, n. 1-3, p. 361–391, 1994. Disponível em: <[https://doi.org/10.1016/0167-2789\(94\)90293-3](https://doi.org/10.1016/0167-2789(94)90293-3)>.
- NEUMANN, J. v. **Theory of self-reproducing automata**. Champaign, IL, USA: University of Illinois Press, 1966.
- OLIVEIRA, G. et al. Secret key specification for a variable-length cryptographic cellular automata model. In: **Proceedings of the International Conference on Parallel Problem Solving from Nature**. Krakow, Poland: Springer, 2010. p. 381–390. Disponível em: <https://doi.org/10.1007/978-3-642-15871-1_39>.
- OLIVEIRA, G. M.; OLIVEIRA, P. P. d.; OMAR, N. Definition and application of a five-parameter characterization of one-dimensional cellular automata rule space. In: **Proceedings of the Artificial life**. MIT Press, 2001. v. 7, n. 3, p. 277–301. Disponível em: <<https://doi.org/10.1162/106454601753238645>>.
- OLIVEIRA, G. M. B.; BORTOT, J. C.; OLIVEIRA, P. P. B. de. Multiobjective evolutionary search for one-dimensional cellular automata in the density classification task. In: **Proceedings of the Artificial Life**. [S.l.]: MIT Press, 2002. p. 202–206.
- OLIVEIRA, G. M. B.; OLIVEIRA, P. P. B. de; OMAR, N. Improving genetic search for one-dimensional cellular automata, using heuristics related to their dynamic behavior forecast. In: **Proceedings of the 2001 Congress on Evolutionary Computation**. [S.l.]: IEEE, 2001. v. 1, p. 348–355.
- OLIVEIRA, G. M. B. et al. An evolutionary-cooperative model based on cellular automata and genetic algorithms for the navigation of robots under formation control. In: **Proceedings of the 7th Brazilian Conference on Intelligent Systems**. São Paulo, Brazil: IEEE, 2018. p. 426–431. Disponível em: <<https://doi.org/10.1109/BRACIS.2018.00080>>.
- ORTIGOZA, G.; BRAUER, F.; LORANDI, A. Mosquito-borne diseases simulated by cellular automata: A review. **International Journal of Mosquito Research**, v. 6, n. 6, p. 21–38, 2019.
- PACKARD, N. H. Adaptation toward the edge of chaos. **Dynamic patterns in complex systems**, v. 212, p. 293, 1988. Disponível em: <<https://doi.org/10.1142/0534>>.
- PEREIRA, F. M. M.; SCHIMIT, P. H. T. Dengue fever spreading based on probabilistic cellular automata with two lattices. **Physica A**, v. 499, p. 75–87, 2018. Disponível em: <<https://doi.org/10.1016/j.physa.2018.01.029>>.
- ROBERT, C. P.; CASELLA, G.; CASELLA, G. **Monte Carlo statistical methods**. Paris, France: Springer, 1999. v. 2. Disponível em: <<https://doi.org/10.1007/978-1-4757-3071-5>>.
- SALAZAR, L. F.; NOBRE, C. A.; OYAMA, M. D. Climate change consequences on the biome distribution in tropical South America. **Geophysical Research Letters**, v. 34, p. 1–6, 2007. Disponível em: <<https://doi.org/10.1029/2007GL029695>>.

SARKAR, P. A brief history of cellular automata. **ACM Computing Surveys**, v. 32, n. 1, p. 80–107, 2000. Disponível em: <<https://doi.org/10.1145/349194.349202>>.

SEREDYNSKI, F.; ZOMAYA, A. Sequential and parallel cellular automata-based scheduling algorithms. **IEEE Transactions on Parallel and Distributed Systems**, IEEE, v. 13, n. 10, p. 1009–1023, 2002. Disponível em: <<https://doi.org/10.1109/TPDS.2002.1041877>>.

SLIMI, R. et al. A cellular automata model for chagas disease. **Applied mathematical modelling**, Elsevier, v. 33, n. 2, p. 1072–1085, 2009. Disponível em: <<https://doi.org/10.1016/j.apm.2007.12.028>>.

SZABAN, M.; SEREDYNSKI, F.; BOUVRY, P. Evolving collective behavior of cellular automata for cryptography. In: **Proceedings of the IEEE Mediterranean Electrotechnical Conference**. Malaga, Spain: IEEE, 2006. p. 799–802. Disponível em: <<https://doi.org/10.1109/MELCON.2006.1653219>>.

TAN, R. R. Hybrid evolutionary computation for the development of pollution prevention and control strategies. **Journal of Cleaner Production**, Elsevier, v. 15, n. 10, p. 902–906, 2007. Disponível em: <<https://doi.org/10.1016/j.jclepro.2006.01.011>>.

TARANTOLA, A. **Inverse problem theory and methods for model parameter estimation**. Philadelphia, PA, USA: SIAM, 2005. v. 89. Disponível em: <<https://doi.org/10.1137/1.9780898717921>>.

TOFFOLI, T. Cellular automata as an alternative to (rather than an approximation of) differential equations in modeling physics. **Physica D: Nonlinear Phenomena**, Elsevier, v. 10, n. 1-2, p. 117–127, 1984. Disponível em: <[https://doi.org/10.1016/0167-2789\(84\)90254-9](https://doi.org/10.1016/0167-2789(84)90254-9)>.

TOFFOLI, T.; MARGOLUS, N. **Cellular automata machines: a new environment for modeling**. Cambridge, Massachusetts, USA: MIT press, 1987. Disponível em: <<https://doi.org/10.7551/mitpress/1763.001.0001>>.

TOMINAGA, D.; KOGA, N.; OKAMOTO, M. Efficient numerical optimization algorithm based on genetic algorithm for inverse problem. In: **Proceedings of the 2nd Annual Conference on Genetic and Evolutionary Computation**. Las Vegas, Nevada, USA: [s.n.], 2000. p. 251–258.

Turner II, B. L. et al. **Land-Use and Land-Cover Change Science/Research Plan**. Stockholm and Geneva, 1995.

VELHO, H. F. de C. Problemas inversos: conceitos básicos e aplicações. **Anais do Encontro de Modelagem Computacional**, p. 63–79, 2001.

WHITE, S. H.; REY, A. M. D.; SÁNCHEZ, G. R. Modeling epidemics using cellular automata. **Applied mathematics and computation**, Elsevier, v. 186, n. 1, p. 193–202, 2007. Disponível em: <<https://doi.org/10.1016/j.amc.2006.06.126>>.

WOLFRAM, S. Cellular automata as simple self-organizing systems. **Caltech preprint CALT-68-938**, 1982.

_____. Theory and applications of cellular automata. **World Scientific**, 1986.