

André Rodrigues da Cruz

Uma Metodologia Multiobjetivo para o Controle de Epidemias através de Vacinação Impulsiva via Algoritmo Genético com Operador de Busca Local Baseado em Aproximação Quadrática Convexa e Validação Estocástica

Dissertação submetida à banca examinadora designada pelo Colegiado do Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal de Minas Gerais como requisito parcial para obtenção do grau de Mestre em Engenharia Elétrica.

Orientador: Ricardo H. C. Takahashi
Co-orientador: Rodrigo T. N. Cardoso

Universidade Federal de Minas Gerais
Belo Horizonte, 31 de Janeiro de 2011.

Dedicatória

Para Cecília.

E a todos aqueles que sonham chegar lá.

[Página deixada em branco intencionalmente.]

Agradecimentos

Primeiramente, agradeço ao Senhor DEUS, criador da vida, fonte de sabedoria e de amor.

Aos meus queridos pais, Olga e Eustáquio, minhas referências de carinho, ternura, disciplina, amor, fé e vitória.

Aos meus inspiradores irmãos Nilton, Nilcemar e Alexsandro, sempre presentes quando preciso e não preciso.

À minha linda Cecília, pela força, paciência e carinho.

Aos meus grandes orientadores Ricardo Takahashi e Rodrigo Cardoso, por todo o conhecimento que adquiri, ao longo desses maravilhosos anos, através das inúmeras e produtivas discussões.

Aos meus amigos do grupo de pesquisa, Alexandre Celestino, Eduardo Carrano, Elizabeth Wanner, Flávio Martins, Luciana Rocha, Oriane Magela, Rodney Saldanha e Sérgio Serqueira, pelo exemplo de profissionalismo, docência e pesquisa.

Aos meus avaliadores, Felipe Campelo e Luiz Duczmal, pelas valiosas críticas e observações efetuadas em meu trabalho.

A todos os meus amigos e guerreiros da Matemática Computacional, em especial a Alexandre Marinho, André Rocha, David Capachi, Felipe Lemos, Gustavo Zeferino e João Júnior.

E a todos aqueles que contribuíram para que esse trabalho fosse realizado.

Resumo

A Epidemiologia é a ciência que estuda os padrões de saúde e doença e os respectivos fatores associados em uma população. Ela se fundamenta na pesquisa em saúde pública para identificar os riscos de uma doença e determinar abordagens de tratamento ideal na medicina clínica e preventiva. A Epidemiologia Matemática modela a dinâmica da proliferação e permite o estudo quantitativo, oferecendo ferramentas para determinar intervenções efetivas para o controle de doenças. A modelagem matemática contribui para o projeto e análise de pesquisas epidemiológicas, sugere quais são os dados cruciais que devem ser coletados, identifica tendências, permite gerar previsões gerais, analisar previamente uma possível intervenção e determinar incertezas. Nessa dissertação é apresentada uma metodologia multiobjetivo para otimização e validação de um conjunto de soluções contendo políticas de controle que minimizam a população de infectados e o custo de implementação de campanhas de vacinação em um horizonte de tempo finito. O modelo epidemiológico que governa o sistema durante a otimização é o Suscetível-Infectado-Recuperado (SIR). As soluções apresentam o número de campanhas, a quantidade necessária de vacinas em cada campanha e os intervalos de tempo entre campanhas. O motor de otimização é o NSGA-II, acrescido de um operador de busca local baseado na otimização de somas ponderadas das aproximações quadráticas convexas das funções objetivo sobre uma vizinhança de pontos. A aproximação do conjunto Pareto é validada em um Modelo Baseado em Indivíduos (MBI) através de um procedimento Monte Carlo. Informações sobre a probabilidade de erradicação e intervalos de confiança são extraídas.

Abstract

Epidemiology is the science that studies patterns of health and disease and its associated factors in a population. It is based on public health research to identify risks for disease and to determine optimal treatment approaches in clinical and preventive medicine. Mathematical Epidemiology models the dynamics of the spread of diseases and allows the quantitative study, providing tools to determine effective interventions for disease control. Mathematical modeling contributes to the design and analysis of epidemiological studies, suggests what are the crucial data to be collected, identifies trends, generates predictions, helps to analyze possible pre-interventions and determines uncertainties. This dissertation presents a multiobjective methodology to optimize and validate a set of solutions containing control policies that minimize the infected population and the implementation costs of vaccination campaigns in a finite time horizon. The epidemiological model that governs the system during the optimization is the Susceptible-Infected-Recovered (SIR). The solutions contain the number of campaigns, the necessary number of vaccine doses in each campaign and the time intervals between campaigns. The optimization engine is the NSGA-II, with an embedded local search operator based on the optimization of weighted sums of convex quadratic approximations of the objective functions on a neighborhood of points. The approximation of the Pareto set is validated in an Individual Based Model (IBM) through a Monte Carlo procedure. Information about probability of eradication and confidence intervals are extracted.

Conteúdo

Conteúdo	vi
Lista de Figuras	vii
Lista de Tabelas	viii
Lista de Símbolos	ix
1 Introdução	1
1.1 Revisão Bibliográfica	3
1.2 Objetivos	7
1.3 Contribuições do Trabalho	9
2 Epidemiologia Matemática	10
2.1 Modelagem Matemática	10
2.2 O Modelo SIR	13
2.3 O Modelo MBI	18
3 Otimização Multiobjetivo	26
3.1 Problema de Otimização Multiobjetivo	27
3.2 Conjunto Pareto-Ótimo	28
3.3 Algoritmos Evolutivos Multiobjetivo	31
3.4 NSGA-II	33
3.5 Algoritmos Meméticos	40
3.6 Aproximação Quadrática Convexa	42
3.7 NSGA-II com Operador de Busca Local	44
4 Metodologia	48
4.1 Modelo de Otimização	49
4.2 Configuração do NSGA-II	51
4.3 S-Metric	52
4.4 Simulação Monte Carlo	53

4.5	Dominância Estocástica	54
5	Resultados	56
5.1	Otimização utilizando o modelo SIR	56
5.2	Validação utilizando o modelo MBI	60
6	Conclusões	77
6.1	Discussão	78
6.2	Trabalhos Futuros	79
6.3	Produção Bibliográfica	79

Lista de Figuras

1.1	O diagrama geral do modelo SIR.	3
2.1	Exemplos do uso de modelos em epidemiologia.	11
2.2	Diagramas de fase para o modelo clássico SIR endêmico.	16
2.3	Comportamento do modelo SIR endêmico.	18
2.4	Equivalência SIR/MBI.	25
3.1	Imagem do conjunto factível \mathcal{F}_x	31
3.2	Cálculo da crowding distance.	37
3.3	Funcionamento geral do NSGA-II.	40
4.1	Região delimitada para o cálculo da S-Metric.	53
5.1	Fronteira não dominada retornada pelo NSGA-II memético.	57
5.2	União das soluções finais não dominadas.	58
5.3	Fronteira não dominada final.	59
5.4	Box-plot da S-Metric para as soluções finais não dominadas.	60
5.5	Box-plot da S-Metric pela geração.	61
5.6	Mediana da S-Metric pela geração.	62
5.7	Box-plot do tempo das otimizações.	63
5.8	Histograma dos números de campanhas.	64
5.9	Histograma dos custos de implementação das campanhas.	65
5.10	Histograma das integrais de infectados das soluções.	66
5.11	Custo de implementação pelo número de campanhas.	67
5.12	Solução mais barata (SIR).	68
5.13	Solução de custo mediano (SIR).	69
5.14	Solução mais cara (SIR).	70
5.15	Fronteira não dominada estocasticamente final.	71
5.16	Fronteira não dominada estocasticamente final (mediana).	72
5.17	Solução com menor probabilidade de erradicação.	75
5.18	Solução com maior probabilidade de erradicação.	76

Lista de Tabelas

2.1	Parâmetros do modelo SIR endêmico usados no trabalho. . . .	18
4.1	Parâmetros do NSGA-II com busca local.	52
5.1	Dados das soluções não dominadas estocasticamente.	73

Lista de Símbolos

SIR	Suscetível-Infectado-Recuperado
MBI	Modelo Baseado em Indivíduo
t	Tempo
$s(t)$	Taxa de suscetíveis no tempo
$i(t)$	Taxa de infectados no tempo
$r(t)$	Taxa de recuperados no tempo
$S(t)$	Número de suscetíveis no tempo
$I(t)$	Número de infectados no tempo
$R(t)$	Número de recuperados no tempo
R_0	Número básico de reprodução
σ	Número de contato
ϱ	Número de substituição
N	Número de indivíduos no modelo epidemiológico
β	Taxa de infecção
η	Coefficiente de ação de massa
λ	Força de infecção
γ	Taxa de imunização
μ	Taxa de nascimento/mortalidade
T	Plano de fase si

$(s_0 \ i_0 \ r_0)$	Condição inicial (taxa)
$(S_0 \ I_0 \ R_0)$	Condição inicial (valor absoluto)
$(s_e \ i_e \ r_e)$	Ponto de equilíbrio endêmico (taxa)
$(S_e \ I_e \ R_e)$	Ponto de equilíbrio endêmico (valor absoluto)
\vec{x}	Vetor de variáveis
\vec{f}	Vetor de objetivos
n	Dimensão do espaço de variáveis
M	Dimensão do espaço de objetivos
J	Número de restrições de desigualdade
K	Número de restrições de igualdade
\mathcal{D}	Conjunto das variáveis de decisão
\mathcal{F}_x	Região factível
\mathcal{Y}	Espaço de objetivos
\prec	Operador de dominância
\mathcal{P}'	Conjunto de soluções não dominadas
\mathcal{X}_l	l -ésima fronteira não dominada
\mathcal{X}^*	Conjunto Pareto-Ótimo Global
\mathcal{Y}^*	Fronteira Ótima
AE	Algoritmo evolucionário
AG	Algoritmo genético
AM	Algoritmo memético
NSGA-II	Non-Dominated Sorting Genetic Algorithm II
Π	Quantidade de indivíduos no algoritmo evolucionário
\prec_n	Operador de comparação de aglomeração
q	Função quadrática convexa

s	Quantidade de parâmetros da função quadrática
m	Quantidade de pontos gerado em uma nuvem
T_{ctrl}	Horizonte de tempo de controle
τ	Instantes de tempo
C	Número de campanhas
Γ	Conjunto dos intervalos de tempo entre campanhas
P	Conjunto dos percentuais de suscetíveis a serem vacinados em cada campanha
$tmin, tmax$	Valores de intervalo de tempo mínimo e máximo entre duas campanhas consecutivas
$pmin, pmax$	Valores de percentual de vacinas mínimo e máximo permitido por campanha
I_{tol}	Valor máximo de infectados tolerado
T_{ini}	Tempo inicial para a ativação da restrição do número máximo de infectados tolerado

Capítulo 1

Introdução

A *Epidemiologia* é o estudo da distribuição e dos determinantes relacionados aos estados de saúde em uma população, e à aplicação de ações para a prevenção e controle de problemas de saúde [1]. A epidemiologia se originou a partir das observações de Hipócrates há mais de 2000 anos atrás, o qual notou que fatores ambientais influenciam na ocorrência de doenças. No entanto, foi durante o século XIX, com os trabalhos de John Snow sobre a cólera na Inglaterra, que a distribuição de doenças em grupos específicos da população humana foi medida em grande escala. A partir desse momento, observou-se que os estudos epidemiológicos devem ser realizados sob um aspecto de um conjunto diversificado de pesquisas que relacionam processos físicos, químicos, biológicos, sociológicos e políticos [2].

A necessidade de compreender e modelar a dinâmica da proliferação de doenças gerou uma nova área de investigação, a *Epidemiologia Matemática* [3]. Esta área de pesquisa estuda quantitativamente a distribuição dos fenômenos de saúde e doença, além de avaliar a eficácia das intervenções em saúde pública [4]. Esta área oferece ferramentas importantes para analisar a propagação e para executar o controle de doenças infecciosas. A modelagem matemática contribui para a formulação e análise de inquéritos epidemiológicos, sugere dados cruciais que devem ser coletados, identifica tendências, faz previsões em geral e estima incertezas [5].

O processo de formulação do modelo esclarece premissas, variáveis e parâmetros. Além disso, os modelos fornecem resultados conceituais tais como limiares, números básicos de reprodução, números de contato e números de substituição. *Modelos Matemáticos* e simulações em computador são ferramentas experimentais úteis para construir e testar teorias, avaliando conjecturas quantitativas, respondendo a perguntas específicas, determinando sensibilidade a mudanças nos valores de parâmetros e estimativas de parâmetros-chave de dados. Compreender as características de transmissão de doenças

infecciosas em comunidades, regiões e países, pode levar às melhores abordagens para a diminuição da transmissão de doenças. Os modelos matemáticos são utilizados para comparar, planejar, implementar, avaliar e otimizar a detecção, prevenção e o tratamento dos programas de controle [6].

Os modelos matemáticos inicialmente ajudam a refinar as questões de estudo por expressar visualmente sistemas complexos e na identificação de variáveis sensíveis. Na fase de desenho do projeto, os modelos podem ser usados para testar estratégias de amostragem, estimar o tamanho da amostra e para prever resultados de estudos inviáveis devido ao tempo ou por considerações éticas. Quando os dados são coletados, os modelos podem auxiliar na interpretação dos resultados, na exploração de caminhos causais e na análise combinada dos dados de múltiplas fontes. Finalmente, os modelos são comumente utilizados no processo de aplicação dos resultados de pesquisas para a prática de saúde pública, estimando o risco para a população, prevendo os efeitos das intervenções e contribuindo para a avaliação dos programas em curso [7].

Uma *Epidemia* é definida como a ocorrência de casos em excesso do que é normalmente esperado em uma comunidade ou região. Ao descrever uma epidemia, o período de tempo, a região geográfica e as características da população na qual ocorrem os casos devem ser especificadas. Uma *Endemia* é definida como uma doença transmissível que possui um padrão relativamente estável de ocorrência em uma determinada área geográfica ou grupo populacional com prevalência e incidência relativamente alta [2].

Existem dois tipos de modelos matemáticos para analisar o comportamento de doenças, o *Modelo Epidêmico* e o *Modelo Endêmico*. O modelo epidêmico é usado para descrever surtos rápidos que ocorrem em um tempo menor que um ano. O modelo endêmico é usado para estudar doenças de longa duração sobre uma população, na qual há renovação de suscetíveis através de nascimentos [6].

O presente trabalho tratará somente com modelos matemáticos epidemiológicos endêmicos, que são:

- *Suscetível-Infetado-Recuperado* (SIR), que é um modelo determinístico baseado em um sistema de equações diferenciais; e o
- *Modelo Baseado em Indivíduo* (MBI), que é um modelo estocástico adaptado para ter um comportamento médio igual ao modelo SIR, que simula a interação entre cada indivíduo em uma população.

O modelo epidemiológico de transmissão SIR [8, 6], proposto por Kermack e McKendrick, é um dos mais usados para representar a dinâmica de

doenças infecciosas. Este modelo classifica os indivíduos em três estados compartimentais baseado na condição de saúde:

- *Suscetível* (S): um indivíduo que é capaz de contrair a doença a partir de uma pessoa infectada através do contato;
- *Infectado* (I): um indivíduo que possui a doença e pode transmiti-la para os suscetíveis através do contato; e
- *Recuperado* (R): um indivíduo que foi curado ou que recebeu vacina, e se tornou imune à doença.

A Figura 1.1 mostra o diagrama de fluxo, em que os compartimentos são as categorias da população (suscetíveis, infectados ou recuperados) e as setas representam o fluxo de indivíduos entre quaisquer dois compartimentos.

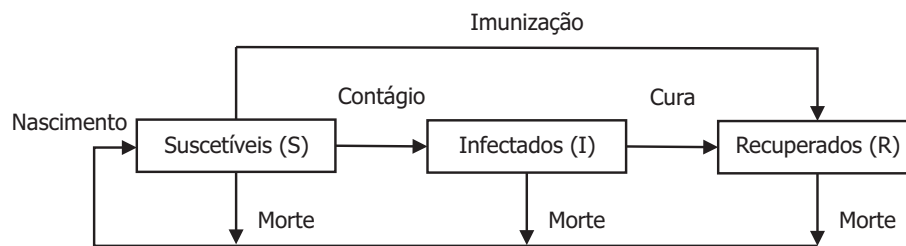


Figura 1.1: O diagrama geral do modelo SIR com as classes Suscetíveis (S), Infectados (I) e Recuperados (R).

1.1 Revisão Bibliográfica

O modelo SIR, tanto endêmico quando epidêmico, e os respectivos variantes são amplamente estudados na literatura. O trabalho [7] descreve a diversidade de aplicações dos modelos matemáticos no campo da epidemiologia e fornece uma visão geral dos métodos de modelagem. O autor de [9] discute como os modelos matemáticos ajudam a melhorar a compreensão das epidemias. No artigo [10] realiza-se uma análise qualitativa dos modelos de doenças transmissíveis, mostrando as interpretações epidemiológicas. O trabalho [11] analisou o modelo epidemiológico SIRS com dinâmica vital em uma população de tamanho variável, no qual se mostra a não existência de soluções periódicas.

O artigo [12] derivou e estudou um modelo SIR com retardo de tempo com uma taxa de incidência geral para investigar a estabilidade e para encontrar o equilíbrio trivial e endêmico. O trabalho [13] estudou o impacto

da forma não-linear da taxa de incidência de doenças infecciosas sobre a dinâmica de modelos epidemiológicos e mostrou que, sob uma população de tamanho constante, esses modelos exibem estados de equilíbrio assintoticamente estáveis. Também sobre o pressuposto do tamanho da população ser constante, o autor em [14] provou que a concavidade da função não-linear da taxa de transmissão, com relação ao número de infectados, garante a unicidade e a estabilidade global do estado de equilíbrio endêmico positivo. O artigo [15] considera as propriedades globais dos modelos clássicos SIR, SIRS e SEIR e constrói funções de Lyapunov que permitem encontrar condições biologicamente realistas, suficientes para garantir a existência e a unicidade de um estado de equilíbrio globalmente assintoticamente estável (endêmico ou livre de infecção, dependendo do valor do número básico de reprodução). O artigo [16] mostra o estudo de propriedades globais dos modelos SIR e SEIR como, por exemplo, a existência de vários estágios paralelos infectantes.

A partir do fim da década de 1980, o método de simulação denominado Modelo Baseado em Indivíduo (MBI), também conhecido por Modelo Baseado em Agentes, tem sido usado, com maior intensidade, para formular hipóteses mais realistas do que modelos baseados equações diferenciais, de modo a permitir melhor a compreensão de propriedades dos sistemas que emergem da interação entre cada indivíduo em uma população [17]. O MBI fornece resultados importantes para a teoria ecológica, provendo uma abordagem estrutural única para diversas aplicações, incluindo a dinâmica de epidemias. Essa abordagem opera no menor nível de organização considerado das atividades dos indivíduos. A simulação das ações dos organismos simples permite estudar o comportamento das entidades de maior nível ecológico como enxames, populações, conexões de rede e padrões de distribuição regionais. Diferentemente de outras abordagens de trabalho em níveis mais elevados de abstração, os MBIs podem representar relações funcionais estruturais semelhantes aos padrões de conhecimentos ecológicos disponíveis [18].

Encontra-se na literatura diversas aplicações com o MBI para simular as interações dos indivíduos em populações. Em ecologia, por exemplo, implementações foram usadas para simular a dinâmica das comunidades de peixes [19], para estudar a propagação de doenças em plantas [20] e a propagação de uma doença hemorrágica viral em coelhos selvagens [21].

Em epidemiologia, o MBI é uma ferramenta importante para simular e analisar a aleatoriedade em uma população de indivíduos suscetíveis, infectados e recuperados. Pode-se extrair a probabilidade de erradicação da doença e o intervalo de confiança do comportamento de cada população, durante um horizonte de tempo dado.

O trabalho [22] apresentou resultados obtidos a partir de um modelo de

epidemia, concebido para simular a peste bubônica em uma população de roedores com um alto nível de resolução espacial e temporal. Desenvolveu-se neste trabalho um modelo baseado em indivíduo estruturalmente realista. No mesmo, foram necessários conhecimentos sobre o comportamento de pulgas e ratos, transmissão interindividual das pragas, além da evolução da doença em organismos de maneira individual, de modo que o modelo refletisse a forma como o sistema real opera.

Em [23], um modelo baseado em indivíduo foi aplicado para computar o número de nascimentos e mortes dos indivíduos, a geração de novas famílias e a carreira escolar dos indivíduos, a fim de investigar a dinâmica da hepatite viral do tipo A nas áreas italianas mais afetadas.

O trabalho [24] utilizou um modelo baseado em indivíduos para estudar os efeitos da epidemia de gripe com simulações estocásticas em grande escala, usando os diferentes cenários de propagação da doença que foram notificados na Polônia.

Um dos grandes desafios em saúde pública está relacionado com a promoção de campanhas de vacinação, que visa à erradicação de uma doença com o mínimo custo possível [8]. Usando o modelo SIR como regra para descrever a dinâmica do sistema, o objetivo é encontrar a melhor solução de um problema de controle não-linear, uma vez que o sistema é descrito por um conjunto de equações diferenciais algébricas não-lineares.

Esse tipo de problema tem mostrado, em muitas vezes, possuir vários mínimos locais, tornando-se inadequado o uso de algoritmos determinísticos para a busca do ótimo global. Neste contexto, é comum e mais aplicável o uso de algoritmos de busca estocástica para este ideal. Porém, a otimalidade das soluções não é garantida, uma vez que essa técnica é uma *Heurística* [25].

Muitos trabalhos encontrados na literatura têm buscado soluções para determinar as políticas de controle em modelos epidemiológicos. No artigo [26], consideram-se dois conceitos para uma estratégia de vacinação, uma com o número de vacinações pelo tempo e outra pelo número reprodução. O mesmo trabalho tratou os problemas de encontrar a estratégia com o menor número de reprodução dado um custo, e de encontrar a estratégia com o menor custo, dado o número reprodução.

Um *Sistema Impulsivo* consiste em um sistema dinâmico contínuo sujeito a perturbações muito rápidas, cuja duração é muito pequena comparada com o tempo de evolução do sistema. Essas perturbações podem ser tratadas como instantâneas. Tais perturbações instantâneas são chamadas de *Impulsos*. Outra característica importante é que esses impulsos ocorrem descontinuamente, intervindo na dinâmica apenas de tempos em tempos. Esse tipo de análise é ideal para tratar sistemas que sofrem alterações bruscas em certos períodos [27].

Um *Controle Impulsivo* é uma intervenção externa que altera o estado do sistema através de impulsos. Este controle visa levar o sistema para algum estado desejado. Este tipo de controle pode ser muito interessante em sistemas que não suportam um distúrbio contínuo, além de representar uma economia no número de intervenções necessárias para controlar um certo sistema real [27].

Os autores de [28] mostraram usando o modelo epidemiológico SIR, que sob o regime planejado de *Vacinação Impulsiva* ou vacinação por pulsos, o sistema converge para uma solução estável. Neste caso, o número de indivíduos infecciosos tende a zero. Mostrou-se também que esta política leva para a erradicação da epidemia, caso respeitado certas condições a respeito da magnitude da proporção de vacinação e sobre o período dos pulsos.

O artigo [29] analisou os fundamentos da estratégia de vacinação por pulsos em um modelo SIR. Foi demonstrado pelo modelo que vacinando a população suscetível repetidamente, em uma série de pulsos, é possível erradicar o sarampo na população. Também foram derivadas as condições para a erradicação da epidemia sob diversas restrições e mostrado a dependência dos parâmetros em relação ao modelo da epidemia.

O trabalho [30] estudou a dinâmica do modelo epidêmico SIR com *Transmissão Horizontal*¹ e *Transmissão Vertical*², sob vacinação constante e por pulsos. Os resultados teóricos mostram que, com a vacinação constante, o comportamento dinâmico é semelhante à ausência de vacinação. Mas, sob um regime previsto de vacinação por pulsos, o sistema converge para uma solução estável e periódica com o número de indivíduos infecciosos igual a zero.

O autor em [31] analisou o problema da aplicação da estratégia de vacinação impulsiva para a erradicação estável de uma classe relevante de doenças infecciosas. Essa análise foi realizada em termos do estudo da estabilidade assintótica local e global da solução para o modelo epidemiológico SEIR. Em outro trabalho do mesmo autor [32], foi provada a estabilidade assintótica global da solução para a erradicação de doenças com transmissão vertical através da estratégia de vacinação por pulsos. Foi demonstrado que a condição para a *Estabilidade Local* garante também a *Estabilidade Global*.

O trabalho [33] descreve o estudo da teoria da análise de estabilidade para encontrar o equilíbrio para o modelo SIR. Em seguida, a fim de atingir o controle da doença, considera-se um problema de controle em relação ao modelo SIR, através da vacinação da população suscetível. É mostrado

¹A transmissão horizontal ocorre através de contato entre indivíduos infectados e suscetíveis à doença em questão.

²A transmissão vertical ocorre quando uma mulher grávida transmite a doença ao feto.

que existe um controle ótimo e são descritas simulações numéricas usando o *Método de Runge-Kutta* [34] de quarta ordem. Os mesmos autores introduziram a estratégia de controle ótimo do modelo epidemiológico SIR com atraso de tempo em [35].

Os autores de [36] analisaram três estratégias diferentes de vacinação e tratamento no modelo epidemiológico SIR com força infecciosa saturada e com transmissão vertical. A dinâmica dos modelos de epidemias a nível mundial foi investigada usando a teoria de Floquet [37] e o teorema da comparação de equações diferenciais impulsivas. Os limites foram identificados e os resultados sobre a estabilidade global foram provados. Para cada estratégia de tratamento e vacinação, a solução periódica foi obtida com a condição livre de doença do sistema impulsivo. Essa solução é globalmente e assintoticamente estável quando o número básico de reprodução é inferior a um.

O trabalho [38] investigou um modelo epidemiológico SIR com atraso em que foi aplicada a vacinação impulsiva. Foi concluído que a solução livre de infecção periódica é globalmente atrativa, caso a taxa de vacinação seja grande o suficiente. Foi provado também que a doença é permanente, caso a proporção de vacinação seja menor que um valor crítico sob condições apropriadas.

O artigo [39] apresenta uma otimização através da vacinação impulsiva sob o modelo SIR estruturado em estágios. Algumas condições suficientes para a tendência global de uma solução livre de infecção periódica e a permanência deste sistema foram apresentadas.

Em [40], os autores apresentaram uma metodologia para a modelagem da propagação de agentes patogênicos ao longo de uma população e para a geração de políticas que minimizem o impacto da enfermidade. A técnica foi utilizada para estudar a propagação de vírus em humanos entre as cidades através de viagens de avião. Ela combina a simulação baseada em agentes, análise matemática e um algoritmo evolutivo. O objetivo deste estudo foi desenvolver ferramentas que determinem a distribuição ótima de um suprimento de vacinas.

1.2 Objetivos

Uma questão importante a ser analisada é a relação entre a eficácia das políticas de controle de epidemias e o respectivo custo de implementação. Esta é uma relação conflituosa porque existe um compromisso entre os objetivos quando há um número diferente de campanhas de vacinação, cada uma sendo implementada após subintervalos de tempo diferentes, havendo

uma quantidade diferente de vacinas a serem aplicadas em cada estágio na população suscetível, tendo em conta um horizonte de tempo finito.

O presente trabalho apresenta uma metodologia multiobjetivo para determinar as melhores soluções não dominadas que representam políticas de controle de epidemias, através de vacinação impulsiva, que possuem um compromisso entre o custo e a eficácia.

Dado um horizonte de tempo, uma *Política de Controle* é representada pelo número de campanhas de vacinação, os momentos em que cada uma delas deve ser realizada, bem como o percentual da população suscetível que deve ser imunizado em cada campanha. Os intervalos entre duas campanhas de vacinação consecutivas não são necessariamente iguais, o otimizador irá decidir o tamanho de cada subintervalo. Da mesma forma, o número de vacinas que serão aplicadas não é necessariamente o mesmo para todas as campanhas.

O custo de implementação do controle é minimizado e a eficácia do mesmo é maximizada concomitantemente. O modelo epidemiológico SIR endêmico, com parâmetros fixos, governa a dinâmica da epidemia durante o processo de otimização.

O mecanismo de otimização usado no trabalho é o algoritmo evolutivo multiobjetivo NSGA-II³ [41], incorporado com uma busca local baseada na otimização de somas ponderadas das aproximações quadráticas convexas das funções objetivo originais. Essa busca local melhora a qualidade das soluções.

A aproximação do conjunto Pareto [42] é retornada pelo algoritmo evolutivo multiobjetivo com as políticas de controle. Como as soluções são baseadas no modelo SIR, que é descrito por um sistema de equações diferenciais, ao número de infectados nunca será atribuído o valor zero quando uma condição inicial não trivial for fornecida. Desse modo, o uso do modelo epidemiológico determinístico não pode fornecer uma probabilidade de erradicação, tão pouco um intervalo de confiança do comportamento das populações de suscetíveis, infectados e recuperados no tempo [43].

Assim, um MBI é ajustado ao modelo SIR e um procedimento Monte Carlo é executado para cada solução do conjunto Pareto, retornado pelo processo de otimização. As funções objetivo originais são reavaliadas e analisadas, contabilizando o número de erradicações e conseqüentemente a probabilidade da doença se extinguir. O intervalo de confiança para cada população no horizonte de tempo pode ser extraído. Ao final, é executada uma ordenação pela não-dominância estocástica com o propósito de eliminar as soluções sensíveis à variação de parâmetros [44].

³Outro algoritmo evolutivo multiobjetivo pode ser usado ao invés do NSGA-II.

1.3 Contribuições do Trabalho

As principais contribuições do presente trabalho são:

- A aplicação da metodologia multiobjetivo para determinar, usando o modelo SIR endêmico e algoritmos evolucionários multiobjetivo, os melhores planejamentos de campanhas de vacinação para o controle de epidemias, através da minimização concomitante dos custos de implementação e da integral da população de infectados no horizonte de tempo em análise;
- O desenvolvimento e a aplicação de uma busca local baseada na otimização das aproximações quadráticas convexas das funções objetivo originais para melhorar a qualidade das soluções em otimização multiobjetivo; e
- A validação das soluções encontradas usando o MBI e a aplicação da dominância estocástica aproximada, a fim de eliminar soluções sensíveis à variação de parâmetros e extrair informações sobre probabilidade de erradicação e intervalos de confiança.

A presente dissertação está organizada da seguinte maneira: o Capítulo 2 apresenta a teoria básica de epidemiologia matemática e os modelos usados no trabalho; o Capítulo 3 apresenta a teoria básica de otimização multiobjetivo, a descrição do algoritmo evolutivo multiobjetivo usado no trabalho para determinar as políticas ótimas de controle de epidemia usando o modelo SIR; em especial é apresentado o NSGA-II, o algoritmo genético utilizado com o operador de busca local, também explanado; o Capítulo 4 detalha a metodologia do projeto experimental do trabalho; o Capítulo 5 apresenta os resultados encontrados; e o Capítulo 6 apresenta as conclusões, trabalhos futuros e as considerações finais.

Capítulo 2

Epidemiologia Matemática

2.1 Modelagem Matemática

A *Modelagem Matemática* é um conjunto de técnicas, ferramentas e equações que podem ser adaptadas para disciplinas específicas. Em epidemiologia, os modelos matemáticos geralmente definem as interações entre indivíduos, populações e ambientes. Ao definir as regras que descrevem essas interações e traduzi-las em equações, os processos de um conjunto complexo podem ser decompostos em componentes e quantificados. Um modelo pode ser usado para explorar as relações entre a população modelada, testar o impacto de regras dentro de um sistema e os respectivos componentes, além de examinar os resultados dos vários eventos que poderiam ocorrer [7]. O processo de modelagem já fez contribuições significativas para as ciências da saúde, incluindo medicina clínica e saúde pública, e disciplinas relacionadas como biologia, matemática, estatística, bioinformática além de outras áreas [45]. Algumas áreas afins de pesquisa em epidemiologia, que usam modelos, são apresentadas na Figura 2.1.

Os *Modelos Matemáticos* são ferramentas úteis para abordar questões sobre o processo de infecção e imunização. Eles permitem a investigação da relação matemática entre os eventos invisíveis, como a transmissão da infecção entre indivíduos infectados e suscetíveis, e outras mais visíveis, por exemplo, o processo de vacinação. Uma vez que o mecanismo é conhecido, tais modelos permitem a previsão e a antecipação de eventos futuros, como o impacto de uma intervenção para o controle de uma infecção. Uma maior familiaridade os modelos matemáticos em pesquisas epidemiológicas permitirá que os mesmos sejam utilizados mais amplamente e corretamente. Deste modo, será mais acessível uma gama mais ampla de aplicações para que epidemiologistas possam fazer uma análise mais crítica.

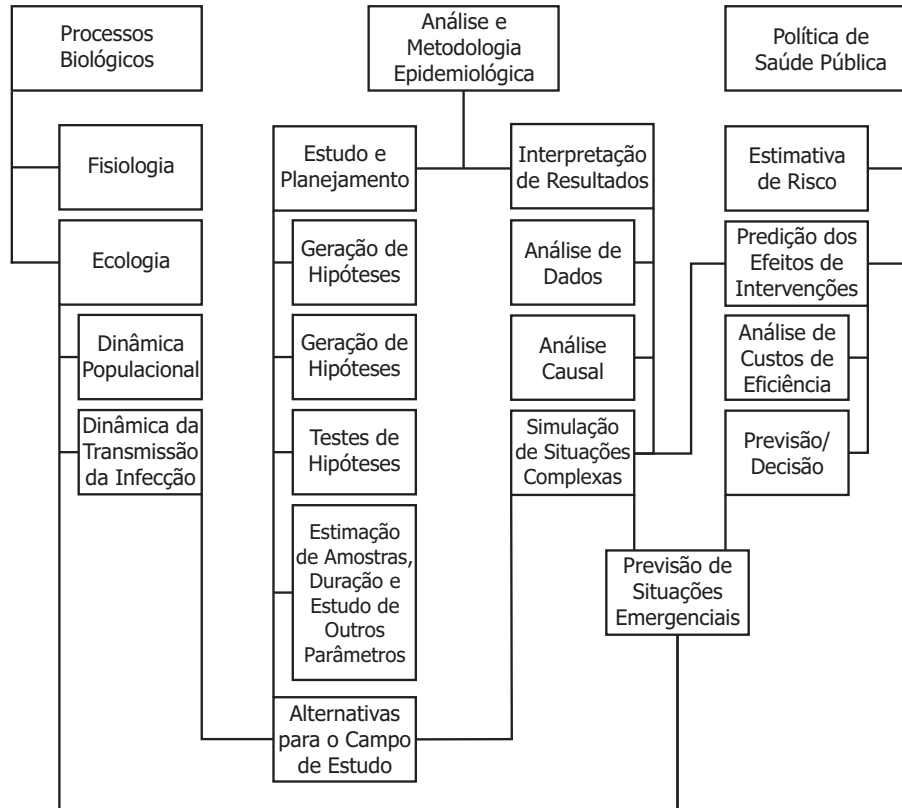


Figura 2.1: Exemplos do uso de modelos em epidemiologia [7].

A principal razão para se estudar modelos matemáticos de propagação da doença é a esperança de se obter uma melhor compreensão do mecanismo de transmissão para obter estratégias de controle mais eficazes [4].

Uma das primeiras contribuições formais em modelagem de epidemias foi executada por Daniel Bernoulli, que teve como objetivo avaliar a eficácia de uma determinada técnica de variolização contra a varíola [46]. Ronald Ross, em outro trabalho, modelou a transmissão da malária [47]. Kermack e McKendrick executaram um estudo mais geral e rigoroso em modelagem epidemiológica [48]. Estes primeiros modelos foram todos determinísticos. Alguns dos primeiros modelos estocásticos são os trabalhos de Bartlett [49] e Kendall [50]. Algumas aplicações recentes podem ser encontradas em [28, 6, 51, 52, 23, 24].

Um modelo aceitável é geralmente aquele que é simples o suficiente para dar respostas que são aplicáveis a um contexto geral, mas também complexo o suficiente para que ele possa abordar as questões mais relevantes [9]. Os dois modelos epidemiológicos utilizados neste trabalho, SIR e MBI adaptado

ao SIR, consideram que em um dado momento uma pessoa é exclusivamente suscetível, infectado ou recuperado. Essas são as *Classes de Estados* ou compartimentais dos indivíduos. As seguintes suposições são feitas: todos os indivíduos nascem como suscetíveis; apenas indivíduos suscetíveis podem se tornar infectados e, depois de ter sido infecciosos durante algum tempo, se curam e se tornam completamente imunes a doença em questão; quando alguém morre outra suscetível nasce, mantendo assim o número de pessoas constante, e não há imigração ou emigração, significando ser uma *Comunidade Fechada* [43]. Uma consequência das hipóteses é que um indivíduo pode fazer os movimentos mudando o respectivo estado: de S para I (processo de infecção), de I para R (processo de recuperação), a partir de S, I ou R para S (morte e nascimento), e de S para R (processo de imunização). Esse esquema de *Mobilidade do Estado de Saúde* é apresentado na Figura 1.1.

Existem muitos modelos epidemiológicos e as escolhas dos compartimentos a se incluir dependem das características da doença em particular que está sendo modelada. MSEIR, MSEIRS, SEIR, SEIRS, SIRS, SEI, SEIS, SI e SIS são outros exemplos de modelos epidemiológicos. O estado M representa a classe dos indivíduos passivamente imunes e o estado E representam àqueles que possuem a doença em período de latência. As classes S, I e R são as mesmas como definidas previamente [6]. Modelos que permitem números de nascimentos e óbitos diferentes, além de imigração e emigração são conhecidos por possuir demografia ou *Comunidade Dinâmica* [43].

O valor limite denominado *Número Básico de Reprodução*, R_0 , é definido como o número médio de infecções secundárias¹ produzidas quando um indivíduo infectado é introduzido em uma população, em que todos são suscetíveis [53]. Em demografia e ecologia, esse termo analogamente significa o número esperado de fêmeas que nascem de uma única fêmea durante toda a vida [54]. Para diversos modelos epidemiológicos determinísticos, uma infecção pode começar em uma população, completamente suscetível, se e somente se $R_0 > 1$, desta forma a epidemia cresce rapidamente. Quando $R_0 \leq 1$ haverá apenas um pequeno surto. Assim, o número básico de reprodução R_0 é muitas vezes considerado como a quantidade limite que determina quando uma infecção pode invadir e persistir em uma população nova [6].

O *Número de Contato*, σ , é definido como o número médio de contatos adequados de uma pessoa doente típica, durante o período infeccioso [10, 55]. Um contato adequado é aquele que é suficiente para a transmissão, ou seja, se o indivíduo infectante entra em contato com pelo menos um suscetível [6].

O *Número de Substituição*, ρ , é definido como sendo o número médio de infecções secundárias produzidas por um indivíduo infeccioso típico, durante

¹Uma infecções secundária é

todo o período em que esteve doente [10].

Embora R_0 seja somente definido no momento da invasão da doença, σ e ρ são definidos para todos os momentos. Para a maioria dos modelos, o número de contato σ mantém-se constante com a progressão da infecção, por isso é sempre igual ao número básico de reprodução R_0 . Nesses modelos σ e R_0 podem ser trocados entre si (sem fazer qualquer diferença) e os teoremas de invasão podem ser declarados em termos de quantidade. O número de substituição ρ é o número real de casos secundários a partir de um infeccioso típico de modo que, após a infecção invadir uma população, nem todos continuam suscetíveis. O valor de ρ é sempre menor ou igual do que o número básico de reprodução R_0 . Além disso, após a invasão da doença, a fração de suscetíveis é inferior a 1, de modo que nem todos os contatos resultam em um novo caso. Assim, o número de substituição ρ é sempre menor ou igual que o número de contato σ após a invasão. A combinação desses resultados implica em $R_0 \geq \sigma \geq \rho$, com a igualdade das três quantidades valendo no momento da invasão. Deve-se notar que $R_0 = \sigma$ para a maioria dos modelos e $\sigma > \rho$ após a invasão de todos os modelos [6].

Os modelos determinísticos visam, essencialmente, uma população com um número grande de indivíduos. Eles podem generalizar o comportamento médio das populações compartimentais. Isso é possível, porque é aceito o pressuposto de uma *Comunidade Homogênea e Uniforme*. Em alguns casos, esse ponto de vista não é adequado. Por exemplo, se for considerar uma pequena comunidade de uma escola com um surto, parece razoável assumir incerteza e aleatoriedade no número final de infectados. Mesmo em uma comunidade grande e a doença possuindo número básico de reprodução $R_0 > 1$, é possível que o surto inicie com apenas poucos infecciosos e a epidemia nunca venha a crescer. Isso indica que os modelos estocásticos podem fornecer informações que não se pode extrair dos determinísticos, como a probabilidade não nula de erradicação da doença, tempo de persistência da epidemia, intervalos de confiança, além de análise de redes sociais [43].

As seções 2.2 e 2.3 apresentam como funcionam, respectivamente, os modelos SIR endêmico e o MBI, particularmente adaptado ao modelo determinístico em média. São apresentados também os parâmetros usados nesse trabalho e o comportamento que esses modelos apresentam.

2.2 O Modelo SIR

O *Modelo SIR Endêmico* é utilizado neste trabalho para descrever a dinâmica dos indivíduos suscetíveis, infectados e recuperados em uma população durante um período considerável de tempo. Esse modelo pode descrever doenças

com contágio viral como sarampo, caxumba e rubéola [6]. Essas doenças compartilham o padrão pelo qual o vírus infecta indivíduos suscetíveis, que se tornam infecciosos, e depois de algum tempo se recuperam e tornam-se imunes a reinfeção. Quando uma vacina existe e é aplicada em um indivíduo suscetível, esse se torna um recuperado por causa da *Imunidade Adquirida*.

A *Incidência Horizontal* mostrada na Figura 1.1 é a taxa de infecção de indivíduos suscetíveis através dos contatos com indivíduos infecciosos. Se no tempo t , $S(t)$ é o número de suscetíveis, $I(t)$ é o número de infecciosos, $R(t)$ é o número de recuperados e N é o tamanho da população total, então $s(t) = S(t)/N$, $i(t) = I(t)/N$ e $r(t) = R(t)/N$ são as frações suscetíveis, infectados e recuperados, respectivamente. Se o valor β é a *Taxa de Infecção*, ou o número médio de contatos adequados, (isto é, contatos suficientes para a transmissão) de uma pessoa por unidade de tempo, consequentemente $\beta I/N = \beta i$ é o número médio de contatos com infecciosos por unidade de tempo de um suscetível, e $\beta IS/N = \beta N i s$ é o número de novos casos da doença por unidade de tempo. Esta forma de incidência horizontal é denominada *Incidência Padrão*, pois é formulada a partir destes princípios básicos [10, 56].

A lei de ação de massa $\eta IS = \eta(Ni)(Ns)$, com η sendo o *Coefficiente de Ação de Massa*, por vezes tem sido utilizada para a incidência horizontal. O parâmetro η não possui interpretação epidemiológica direta, mas observando a formulação padrão verifica-se que $\beta = \eta N$, dessa forma assumindo-se implicitamente que a taxa de contato β aumenta linearmente com o tamanho da população. Ingenuamente, poderia ser plausível que a densidade populacional e, portanto, a taxa de contato fosse aumentar com o tamanho da população, mas os padrões de contato diário das pessoas são geralmente semelhantes em grandes e pequenas comunidades, cidades e regiões. Portanto, o conceito de que as pessoas são infectadas através de seus encontros diários é altamente independente do tamanho da comunidade dentro de um determinado país [6].

A *Força de Infecção*, $\lambda = \beta i$, é o número médio de contatos com infectados por unidade de tempo [6].

Os indivíduos infectados se recuperam, tornando-se imunes a doença a uma *Taxa de Imunização* γ . Assim, a quantidade de indivíduos que transitam do estado infectado para o recuperado a cada tempo t é γI . O valor $1/\gamma$ representa o *Período Médio de Infecção* [6].

Considera-se nesse trabalho, por simplicidade, que o número N de indivíduos da população se mantém constante. Dessa maneira, a *Taxa de Nascimento* será igual à *Taxa de Mortalidade*, que será designada por μ . Todos os indivíduos que nascem são suscetíveis e todos os indivíduos de qualquer classe morrem com a mesma taxa. Portanto, μN é a quantidade de novos

indivíduos suscetíveis que nascem e $\mu(S + R + I) = \mu N$ é a quantidade de indivíduos que morrem no sistema a cada tempo t . O *Tempo Médio de Vida* é dado por $1/\mu$ [6].

O modelo clássico SIR endêmico com dinâmica vital (nascimentos e mortes) é descrito pelo sistema de equações diferenciais apresentado pela Equação 2.1 [6].

$$\begin{aligned} \frac{dS}{dt} &= \mu N - \mu S - \frac{\beta IS}{N}, & S(0) &= S_0 \geq 0, \\ \frac{dI}{dt} &= \frac{\beta IS}{N} - \gamma I - \mu I, & I(0) &= I_0 \geq 0, \\ \frac{dR}{dt} &= \gamma I - \mu R, & R(0) &= R_0 \geq 0, \end{aligned} \quad (2.1)$$

Dividindo-se a Equação 2.1 por N obtém-se a Equação 2.2, em que se assume $r(t) = 1 - s(t) - i(t)$ [6].

$$\begin{aligned} \frac{ds}{dt} &= \mu - \mu s - \beta is, & s(0) &= s_0 \geq 0, \\ \frac{di}{dt} &= \beta is - \gamma i - \mu i, & i(0) &= i_0 \geq 0, \end{aligned} \quad (2.2)$$

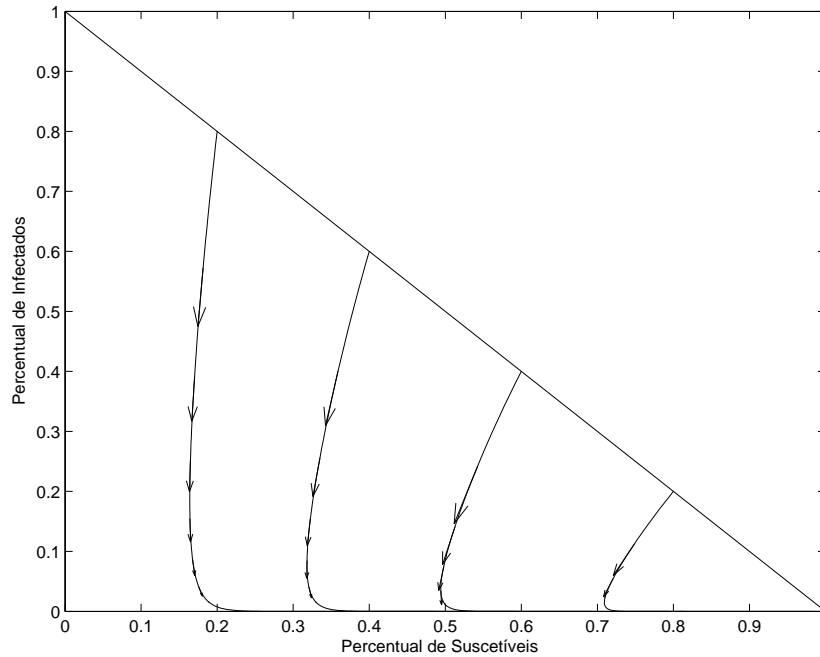
O triângulo T que determina o plano de fase si , dado pela Equação 2.3, é positivamente invariante e possui uma única solução, de modo que o modelo é matematicamente e epidemiologicamente bem posto [10].

$$T = \{(s, i) | s \geq 0, i \geq 0, s + i \leq 1\} \quad (2.3)$$

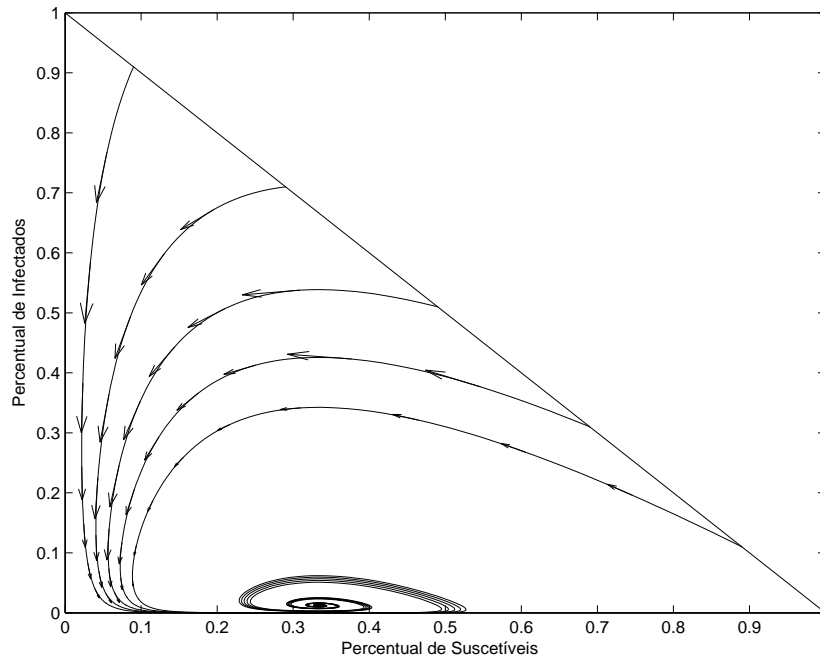
No modelo SIR endêmico, o número de contato σ permanece sempre igual ao número básico de reprodução R_0 . Isso acontece porque nenhuma nova classe de suscetíveis ou infectados ocorre após a invasão da doença. Para esse modelo, a quantidade limite é dada pela Equação 2.4, que é a taxa contato vezes o período médio ajustado do fim da infecção [6].

$$R_0 = \sigma = \frac{\beta}{\gamma + \mu} \quad (2.4)$$

Em [10, 5] é provado o teorema que diz que dado $(s(t), i(t))$, uma solução de 2.2 em T , se $\sigma \leq 1$ ou $i_0 = 0$ então os caminhos das soluções em t levam para o ponto de equilíbrio dado por $(s_e, i_e) = (1, 0)$. Se $\sigma > 1$, então todos os caminhos das soluções com $i_0 > 0$ levam para o equilíbrio endêmico $s_e = 1/\sigma$ e $i_e = \mu(\sigma - 1)/\beta$. Como neste trabalho usa-se o modelo SIR endêmico, pode-se substituir σ por R_0 , uma vez que são iguais.



(a) Exemplo de comportamento do modelo SIR endêmico para $\sigma = R_0 = 0,5$, com $\gamma = 0,6$, $\mu = 1/90$ e $\beta = 0,3056$.



(b) Exemplo de comportamento do modelo SIR endêmico para $\sigma = R_0 = 3$, com $\gamma = 0,6$, $\mu = 1/90$ e $\beta = 1,8333$.

Figura 2.2: Amostras do diagrama de fase para o modelo clássico SIR endêmico, com número de contato σ menor e maior que 1, média do período infeccioso $1/\gamma = 1,67$ unidades de tempo, e expectativa de vida dos indivíduos igual a $1/\mu = 90$ unidades de tempo.

A Figura 2.2 apresenta exemplos das duas possibilidades de pontos de equilíbrio endêmico. Se $R_0 = \sigma \leq 1$, então o número de substituição σs é inferior a 1 quando $i_0 > 0$, de forma que a população de infectados decresça, tendendo o limite para zero. Se $R_0 = \sigma > 1$, i_0 for pequeno, e s_0 for grande, com $\sigma s_0 > 1$, então $s(t)$ diminui e $i(t)$ aumenta até um pico e depois diminui, assim como seria para uma epidemia. No entanto, caso a fração de infectados diminua para um nível baixo, com a lentidão dos processos de morte de recuperados e nascimento de novos indivíduos, de modo que a fração da população de suscetíveis seja grande o suficiente, outra epidemia menor ocorrerá. Esse processo de alternância de rápidas epidemias e lenta regeneração de suscetíveis continua até o sistema entrar no equilíbrio endêmico, apresentado pelo teorema. No equilíbrio endêmico o número de substituição $\sigma s_e = 1$, o que é plausível, pois se o número de substituição for maior ou menor que 1, a fração infectante $i(t)$ deve aumentar ou diminuir, respectivamente [6].

Quando uma doença infecciosa possui taxa básica de reprodução $R_0 > 1$, o nível de infectados no equilíbrio endêmico pode ser alto, o que causa diversos transtornos para a sociedade. O interesse desse trabalho é desenvolver uma solução para o controle, e se possível para a erradicação, de doenças com $R_0 > 1$, através da determinação de políticas de vacinação. Nesse trabalho serão utilizados os parâmetros presentes na Tabela 2.1, que descrevem uma doença infecciosa fictícia para ser controlada. Esses parâmetros não representam uma doença do mundo real, mas são valores que representam um cenário difícil de ser controlado. A expectativa de vida dessa população é de 90 unidades de tempo e o tempo médio de persistência da doença é de 50 unidades de tempo. Nesse caso, a taxa básica de reprodução $R_0 = 4,8214 > 1$, o ponto de equilíbrio endêmico em termos de taxa é $(s_e \ i_e \ r_e) = (0,2074 \ 0,2831 \ 0,5095)$ e em termos de valores absolutos é $(S_e \ I_e \ R_e) = (207,4074 \ 283,0688 \ 509,5238)$. No equilíbrio endêmico quase 30% da população está infectada, o que gera diversos transtornos e gastos para a sociedade. Isso justifica o combate à doença para que a mesma seja controlada ou mesmo erradicada.

O comportamento das populações de suscetíveis, infectados e recuperados é descrito pelo método de Runge-Kutta de quarta ordem [34]. A ilustração do comportamento das populações no tempo, com os parâmetros da Tabela 2.1 é apresentado na Figura 2.3.

A metodologia proposta no presente trabalho se adequa para qualquer combinação de parâmetros para o modelo SIR endêmico. Dessa forma os valores da Tabela 2.1 poderiam assumir outros valores.

Parâmetro	Valor	Descrição
N	1000	Número de indivíduos na população
μ	1/90	Taxa de nascimento/mortalidade
β	0,15	Taxa de infecção
γ	1/50	Taxa de recuperação
s_0	0,99	Percentual de suscetíveis inicial na população
i_0	0,01	Percentual de infectados iniciais na população

Tabela 2.1: Parâmetros do modelo SIR endêmico usados no trabalho.

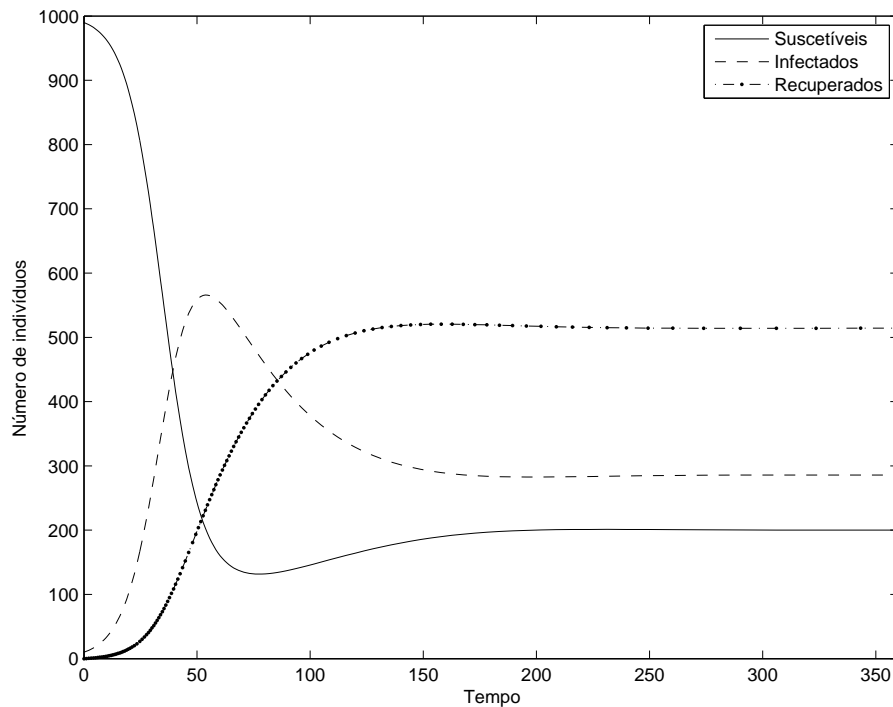


Figura 2.3: Comportamento do modelo SIR endêmico com os parâmetros da Tabela 2.1.

2.3 O Modelo MBI

O Modelo Baseado em Indivíduo (MBI), também conhecido por *Modelo Baseado em Agente*, é uma ferramenta bem conhecida e é utilizada desde a década de 1970 para simular a propagação de doenças infecciosas [57, 58]. A partir do trabalho [59] os MBIs tornaram-se mais conhecidos. Outras aplicações desse tipo de modelo abrangem ecologia [60, 61], ciências sociais [62], economia [63], demografia [64], geografia [65] e ciências políticas [66].

Os MBIs permitem aos pesquisadores estudarem como as propriedades em nível de sistema emergem através do comportamento adaptativo dos indivíduos, bem como a forma como o sistema os atinge [67]. Os MBIs são importantes tanto para a teoria quanto para gerenciamento, porque eles permitem que se considerem aspectos geralmente ignorados nos modelos analíticos como, por exemplo, variabilidade entre os indivíduos, interações locais, ciclos de vida completos e, em particular, comportamento individual adaptativo em virtude de alterações internas e externas no ambiente [68].

Na metodologia apresentada pelo MBI é muito simples representar sistemas complexos físicos e sociais, embora isto possa exigir grandes quantidades de dados, especialmente quando se deseja reproduzir uma estrutura social mais realista de um determinado país ou território. Uma descrição do MBI pode facilmente refletir a heterogeneidade de contatos sociais e introduzir a possibilidade de implementação de rede social dinâmica [24].

O grande potencial dos MBIs possui um custo. Os MBIs são necessariamente mais complexos em estrutura, são mais difíceis de analisar, compreender e comunicar do que os modelos analíticos. Eles devem ser implementados e executados em computadores [17]. O trabalho [69] apresentou e discutiu sete razões para a aplicação de MBIs em ecologia, que são:

1. A inerente complexidade de sistemas ecológicos, impassíveis de uma análise matemática formal;
2. Processos populacionais são fenômenos emergentes, resultando das interações entre os elementos constituintes (indivíduos) e destes com o meio;
3. Poder de predição;
4. Adoção definitiva, por parte da ecologia, de uma visão evolutiva;
5. Indivíduos são entidades discretas;
6. Interações são localizadas no espaço; e
7. Indivíduos diferem entre si.

Um MBI consiste de uma estrutura que simula os relacionamentos de indivíduos, cujo comportamento é definido por um conjunto de regras. À medida que os relacionamentos ocorrem, as características de cada indivíduo podem mudar. O presente trabalho utiliza o modelo elaborado em [70], que foi aperfeiçoado a partir do trabalho [71], desenvolvido de maneira a reproduzir as premissas utilizadas na formulação do modelo SIR endêmico.

O MBI, nesse caso, executa um tratamento estocástico dos contatos entre os indivíduos da população. As seguintes premissas epidemiológicas foram adotadas:

- *População constante*: a população possui um tamanho constante N ;
- *Característica do indivíduo*: cada indivíduo assume uma única característica, a cada momento, que é o estado em relação à epidemia;
- *Distribuição estatística*: A mortalidade e recuperação dos indivíduos seguem uma distribuição uniforme em cada intervalo de tempo e, além disso, todos os indivíduos têm a mesma probabilidade de contato;
- *Mudança de estado*: Em cada instante de tempo, cada indivíduo pode mudar de estado conforme as seguintes regras:
 - $S, I, R \rightarrow S$: renovação dos indivíduos, ou seja, eles morrem e, para que a população permaneça com tamanho constante, nasce o mesmo número de indivíduos suscetíveis;
 - $S \rightarrow I$: um indivíduo suscetível entra em contato com um infectado, adquirindo a doença;
 - $I \rightarrow R$: recuperação de um indivíduo infectado, tornando-se imune a um outro processo de infecção;
 - $S \rightarrow R$: vacinação de um indivíduo infectado, tornando-se imune a um novo processo de infecção.

A evolução do sistema, a cada passo ΔT no tempo, para cada indivíduo da população, é dado da seguinte maneira:

- Sorteia-se aleatoriamente, com distribuição uniforme, a renovação (morte e nascimento de um indivíduo suscetível) com probabilidade $\mu\Delta T$;
- Se o indivíduo estiver infectado e não tiver ocorrido a renovação do mesmo, sorteiam-se com probabilidade $\frac{\beta\Delta T}{N-1}$ entre os indivíduos suscetíveis aqueles que serão infectados por ele; e
- Ainda se o indivíduo estiver infectado, sorteia-se aleatoriamente com distribuição uniforme, com probabilidade $\gamma\Delta T$ a recuperação do mesmo.

O MBI foi elaborado para convergir, em média, para o comportamento do modelo SIR endêmico. Para isso, os valores de μ , β e γ do MBI devem ser iguais aos do modelo determinístico. A prova segue abaixo [72]:

Definem-se os eventos ocorridos no intervalo de tempo $(k\Delta T, (k+1)\Delta T)$:

- $\Delta(k)$: o evento de um indivíduo da classe I se recuperar, passando para a classe R;
- $\Omega(k)$: o evento de um indivíduo morrer e, conseqüentemente, nascer outro na classe S;
- $\Gamma(k)$: o evento de se sortear um indivíduo da classe S;
- $\Phi(k)$: o evento de um indivíduo qualquer ter um contato com um indivíduo da classe I;
- $\Lambda(k)$: o evento de um indivíduo da classe S passar para a classe I, devido ao contato com outro indivíduo da classe I.

A probabilidade de ocorrer a recuperação de um indivíduo infectado é dada pela Equação 2.5.

$$P(\Delta(k)) = \gamma \Delta T. \quad (2.5)$$

A probabilidade de ocorrer a renovação de um indivíduo qualquer é dada pela Equação 2.6.

$$P(\Omega(k)) = \mu \Delta T. \quad (2.6)$$

A probabilidade de novas infecções é a probabilidade de ocorrerem simultaneamente os eventos independentes Γ e Φ , que é apresentada pela Equação 2.7.

$$P(\Lambda(k)) = P(\Gamma(k) \cap \Phi(k)) = P(\Gamma(k))P(\Phi(k)). \quad (2.7)$$

A probabilidade de sortear um indivíduo suscetível é dada pela Equação 2.8.

$$P(\Gamma(k)) = \frac{S(k)}{N}. \quad (2.8)$$

A probabilidade de acontecer um contato com um indivíduo infectado é proporcional ao número de infectados presentes no momento, que é dada pela Equação 2.9.

$$P(\Phi(k)) = \frac{\beta \Delta T}{N-1} I(k). \quad (2.9)$$

Das Equações 2.7, 2.8 e 2.9 chega-se a Equação 2.10.

$$P(\Lambda(k)) = \frac{\beta \Delta T}{N-1} \frac{S(k)I(k)}{N}. \quad (2.10)$$

O número médio de indivíduos, por unidade de tempo ΔT , da classe I que se recuperam da infecção e passam para a classe R, calculado a partir da Equação 2.5, é dado pela Equação 2.11.

$$\gamma\Delta TI(k). \quad (2.11)$$

Os números médios de indivíduos das classes S, I e R que morrem e são substituídos por outros da classe S, a cada unidade de tempo ΔT , são calculados a partir da Equação 2.6, e são dados pela Equação 2.12.

$$\begin{aligned} \mu\Delta TS(k) \\ \mu\Delta TI(k) \\ \mu\Delta TR(k) \end{aligned} \quad (2.12)$$

Se um experimento é realizado X vezes, com X suficientemente grande, e verifica-se que certo evento ocorre E vezes, então, segundo [73], a probabilidade desse evento ocorrer se aproxima da Equação 2.13.

$$P = \frac{E}{X} \quad (2.13)$$

No MBI, em cada intervalo de tempo, executa-se o algoritmo descrito uma vez para cada indivíduo, ou seja, N vezes. A probabilidade da infecção ocorrer é dada por $P(\Lambda(k))$ na Equação 2.10. Fazendo $X = N$ e $E = P(\Lambda(k))$ na Equação 2.13, conclui-se que o número médio de indivíduos, por unidade de tempo ΔT , que passam da classe S para a classe I, é dado pela Equação 2.14.

$$N \frac{\beta\Delta T}{N-1} \frac{S(k)I(k)}{N} \quad (2.14)$$

Para valores de N razoavelmente grandes, digamos a partir de 1000, pode-se considerar $N/(N-1) \approx 1$, e dessa maneira reescreve-se a Equação 2.14 como sendo apresentado pela Equação 2.15.

$$\frac{\beta\Delta TS(k)I(k)}{N} \quad (2.15)$$

Através das Equações 2.11, 2.12 e 2.15 descreve-se a dinâmica do MBI:

- Suscetíveis: recebe o acréscimo de todos os indivíduos renovados ($\mu\Delta TS + \mu\Delta TI + \mu\Delta R$), lembrando que $S(k) + I(k) + R(k) = N$ é dado por $\mu\Delta TN$, e terá a perda dos seus próprios indivíduos renovados ($\mu\Delta TS$) e dos indivíduos infectados ($\beta\Delta TI(k)S(k)/N$);

- Infectados: recebe o acréscimo de novas infecções ($\beta\Delta TI(k)S(k)/N$) e terá a perda dos seus próprios indivíduos recuperados ($\gamma\Delta TI(k)$) e renovados ($\mu\Delta TI(k)$);
- Recuperados: recebe os indivíduos que se recuperam da infecção ($\gamma\Delta TI(k)$) e terá a perda dos seus próprios indivíduos que se renovam ($\mu\Delta TR(k)$).

Assim, a Equação 2.16 descreve a dinâmica associada ao comportamento médio do MBI.

$$\begin{aligned} S(k+1) &= S(k) + \mu\Delta TN - \mu\Delta TS(k) - \frac{\beta\Delta TI(k)S(k)}{N} \\ I(k+1) &= I(k) + \frac{\beta\Delta TI(k)S(k)}{N} - \mu\Delta TI(k) - \gamma\Delta TI(k) \\ R(k+1) &= R(k) + \gamma\Delta TI(k) - \mu\Delta TR(k) \end{aligned} \quad (2.16)$$

Lembrando que k representa o instante de tempo $k\Delta T$ e usando o limite apresentado na Equação 2.17, segue que a Equação 2.16 coincide com a equação de diferenças da Equação 2.1 utilizada na obtenção de soluções do modelo SIR endêmico por integração numérica.

$$\frac{df}{dt} = \lim_{\Delta T \rightarrow 0} \frac{f(t + \Delta T) - f(t)}{\Delta T} \quad (2.17)$$

Assim, é provado que o MBI converge em média para o modelo SIR endêmico. Para uma mesma doença, os parâmetros dos modelos SIR e MBI são os mesmos. Dessa maneira, os parâmetros do MBI usados nesse trabalho são os mesmos presentes na Tabela 2.1. É usado nesse trabalho $\Delta T = 1$.

A Figura 2.4 ilustra a equivalência entre o modelo baseado em equações diferenciais SIR e o modelo baseado em equações de diferenças MBI.

O Algoritmo 1 apresenta um código em alto nível descrevendo o funcionamento do MBI. A função *InitializePopulation()* inicializa a população de acordo com a condição inicial com relação ao número de suscetíveis, infectados e recuperados na população. Depois é simulado para cada tempo, variando-se ΔT , a interação entre os indivíduos. Para cada indivíduo, indexado pelo inteiro i , a função *AnalyzeDeath(P.ind(i))* analisa se esse vai morrer e nascer outro suscetível, a função *AnalyzeVaccine(P.ind(i))* analisa se deve-se aplicar a vacina, a função *AnalyzeInfection(P.ind(i))* analisa se o mesmo vai se tornar um infectado, e por último, é verificado se o indivíduo (infectado) se tornará um recuperado na função *AnalyzeRecovery(P.ind(i))*.

O presente trabalho utiliza o MBI para validar os resultados provenientes da otimização multiobjetivo, que usa o modelo SIR, a fim de auxiliar no

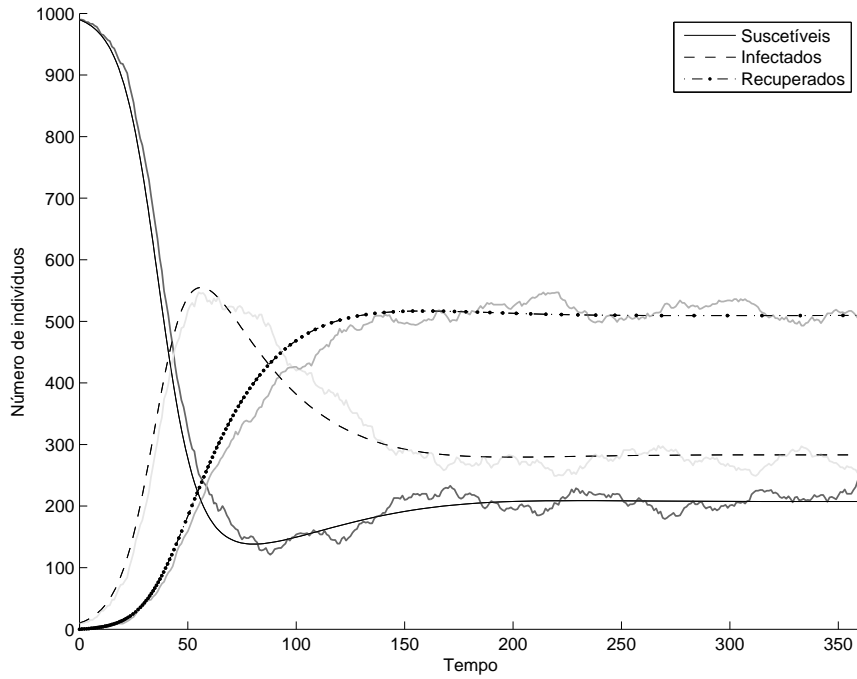
Algorithm 1 MBI

```

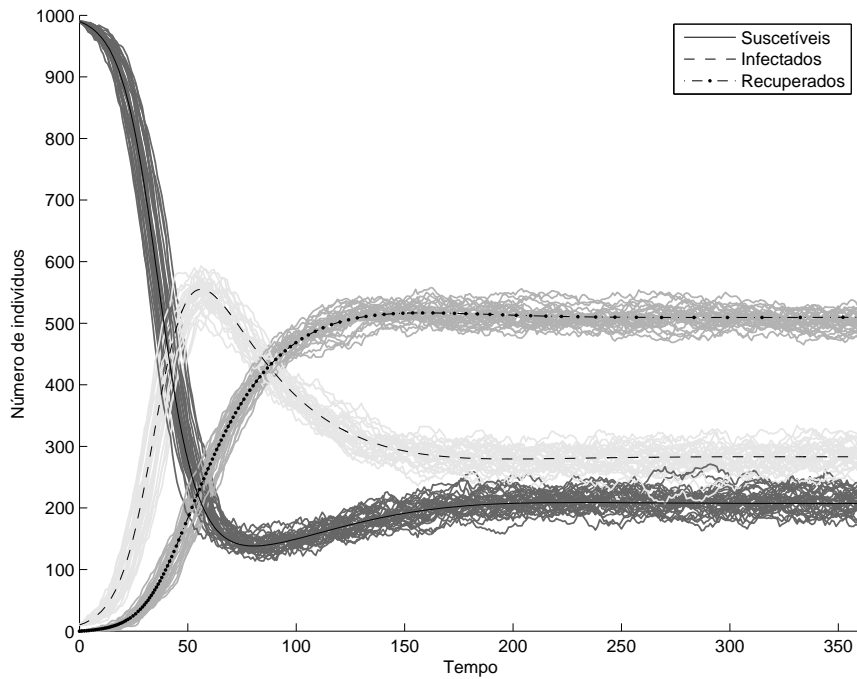
P ← InitializePopulation()
for t from  $T_0$  to  $T_f$  step  $\Delta T$  do
  for i from 1 to N step 1 do
    AnalyzeDeath(P.ind(i))
    if It is in campaign and P.ind(i) is susceptible then
      AnalyzeVaccine(P.ind(i))
    end if
    if P.ind(i) is susceptible then
      AnalyzeInfection(P.ind(i))
    end if
    if P.ind(i) is infected then
      AnalyzeRecovery(P.ind(i))
    end if
  end for
end for

```

processo de tomada de decisão que deve ser realizado para a escolha de uma única solução do conjunto Pareto ótimo. Pretende-se, através de um procedimento de simulação Monte Carlo [74], simular a proliferação de uma doença, testando-se a eficácia de políticas de vacinação com números diferentes de campanhas, possuindo taxas de vacinação e intervalos de tempo não fixados entre cada campanha. Essa validação não é possível utilizando o modelo SIR. Mais especificamente, extrai-se após o processo de simulação, a informação sobre a probabilidade de erradicação da doença, intervalos de confiança com relação aos custos com as campanhas, com a integral da população de infectados e sobre o comportamento das populações de suscetíveis, infectados e recuperados no tempo. Removem-se também as soluções que são sensíveis a variação de parâmetros, causado pelo processo aleatório da simulação, através da aplicação da dominância estocástica [44].



(a) Comportamento de uma execução do MBI e uma do modelo SIR.



(b) Comportamento de trinta execuções do MBI e uma do modelo SIR.

Figura 2.4: Equivalência SIR/MBI, utilizando os parâmetros da Tabela 2.1. As linhas pretas representam o modelo SIR e as linhas em tons de cinza o modelo MBI.

Capítulo 3

Otimização Multiobjetivo

A *Otimização* é processo de encontrar e comparar soluções factíveis, cada vez melhores, até que nenhuma outra possa ser determinada. A otimização minimiza ou maximiza funções objetivo (ou simplesmente objetivos). Quando a modelagem de um problema de otimização considera somente a busca de uma única solução ótima de uma função objetivo, essa tarefa é denominado *Otimização Mono-Objetivo*. Quando um problema de otimização envolve mais do que uma função objetivo, a tarefa de encontrar as soluções é conhecido como *Otimização Multiobjetivo* [42, 75].

Uma grande parcela das aplicações e pesquisas em otimização consideram somente um único objetivo, embora a maioria dos problemas do mundo real envolvam, naturalmente, múltiplos objetivos conflitantes. Neste caso, não se deve buscar somente a solução extrema (máxima ou mínima, quando for o caso) de uma função objetivo enquanto as outras funções objetivo são também importantes. Deve-se buscar o melhor conjunto de soluções em que cada uma delas possui uma relação de compromisso com relação à todos os objetivos. Em uma solução qualquer deste conjunto, uma melhoria em um objetivo implicaria no detrimento de pelo menos um outro objetivo. Todas as soluções que pertencem a este conjunto não podem ser consideradas melhores ou piores, em plenitude, quando comparadas entre si. Isto ocorre porque para qualquer duas soluções diferentes há pelo menos um objetivo que é melhor em uma e pelo menos um outro objetivo que é melhor na outra solução [75].

Uma *Restrição* descreve um determinado limite, intervalos ou conjuntos de valores que uma certa propriedade de um sistema pode assumir. O tratamento de restrições em otimização é muito importante na prática, uma vez que problemas de otimização do mundo real podem possuir limitações físicas, monetárias, além de outros tipos [42, 75].

Um procedimento de otimização multiobjetivo ideal deve primeiro encontrar as soluções ótimas de acordo com as relações de compromisso entre os

diversos objetivos. Posteriormente, deve-se decidir e selecionar uma solução que se adeque às necessidades do usuário. Para isso, deve-se analisar, de forma quantitativa e qualitativa, a relação de compromisso das soluções das funções objetivo.

Existem basicamente duas metodologias para a otimização multiobjetivo. Uma é baseada na otimização mono-objetivo de diversas ponderações das funções objetivo, através da parametrização de escalares cuja soma é convexa. Deve-se fornecer previamente valores de importância para cada função objetivo, o que requer uma análise prévia para direcionar a informação. É necessário também executar o otimizador para cada conjunto de parâmetros. Essa técnica, apesar de ser simples, possui as dificuldades de se determinar o grau de importância de cada objetivo, é sensível a variação dos parâmetros e ainda existe a possibilidade da solução encontrada não pertencer ao conjunto das melhores soluções¹. A outra metodologia busca concomitantemente o conjunto ótimo, respeitando a relação de compromisso entre as funções objetivo, em uma única execução do algoritmo. Desse modo, é necessário decidir qual solução do conjunto se deve escolher. Esse tipo de otimização é menos subjetivo, é mais metódico e prático [42] e será usado nesse trabalho.

3.1 Problema de Otimização Multiobjetivo

Um *Problema de Otimização Multiobjetivo* possui um determinado número de funções que devem ser minimizadas ou maximizadas. Esse tipo de problema geralmente possui um determinado número de restrições que toda solução factível, incluindo o conjunto de soluções ótimas, devem satisfazer. A forma geral de um problema de otimização (em particular, de minimização) multiobjetivo é apresentado na Equação 3.1.

$$\begin{aligned} \min_{\vec{x}} \quad & f_e(\vec{x}) && e = 1, \dots, M && (3.1) \\ \text{sujeito a:} & && && \\ & g_j(\vec{x}) \leq 0 && j = 1, \dots, J && \\ & h_k(\vec{x}) = 0 && k = 1, \dots, K && \\ & x_i^{(L)} \leq x_i \leq x_i^{(U)} && i = 1 \dots n && \end{aligned}$$

Na Equação 3.1, a solução $\vec{x} \in \mathfrak{R}^n$ é um vetor com n variáveis de decisão: $\vec{x} = (x_1, \dots, x_n)^T$. O último conjunto de restrições são os limites das

¹Quando a imagem do conjunto de soluções não dominadas ótimas for não convexa alguns métodos determinísticos baseados em somas ponderadas podem falhar[75].

variáveis, restringindo cada variável de decisão x_i entre os valores limites $x_i^{(L)}$ e $x_i^{(U)}$. Estes limites constituem o conjunto das variáveis de decisão $\mathcal{D} \in \mathfrak{R}^n$. Associado ao problema podem existir J restrições de desigualdade e K restrições de igualdade. Os termos $g_j : \mathfrak{R}^n \rightarrow \mathfrak{R}$ e $h_k : \mathfrak{R}^n \rightarrow \mathfrak{R}$ são denominados funções de restrição. As restrições de desigualdade são tratadas como do tipo “menor ou igual a”, embora se possa tratar também como “maior ou igual a” simplesmente multiplicando a equação por -1 . Uma solução \vec{x} que não satisfaz alguma das $J + K$ restrições ou algum dos $2n$ limites é denominada *Solução Infactível*. Por outro lado, se uma solução \vec{x} satisfaz todas as restrições e todos os limites, então a mesma é chamada de *Solução Factível*. Deve-se perceber que, na presença de restrições até mesmo a própria região \mathcal{D} pode se tornar infactível. O conjunto das soluções factíveis $\mathcal{F}_x \subset \mathcal{D}$ é chamado de *Região Factível* [75].

São consideradas M Funções Objetivo na formulação do problema, apresentado na Equação 3.1. O vetor de objetivos $\vec{f} \in \mathfrak{R}^M$, ou escrito como $\vec{f}(\vec{x}) = (f_1(\vec{x}), \dots, f_M(\vec{x}))^T$, possui funções que podem ser minimizadas ou maximizadas. Pelo princípio da dualidade [75, 76], no contexto da otimização, pode-se converter um problema de maximização em um de minimização multiplicando-se a função objetivo por -1 . O princípio da dualidade permite o tratamento misto de objetivos de maneira muito mais simples. Em geral, os algoritmos de otimização são construídos para solucionar um tipo de otimização. Assim, se um algoritmo somente minimiza os objetivos, então pelo princípio da dualidade pode-se converter as funções originais que precisam ser maximizadas em outras, que possuem as mesmas soluções ótimas, para serem minimizadas.

Uma outra diferença entre a otimização mono-objetivo e multiobjetivo, é que as funções objetivo constituem um novo espaço multidimensional. Esse Espaço de Objetivos é designado pelo conjunto $\mathcal{Y} \subset \mathfrak{R}^M$. Para cada solução \vec{x} no espaço de variáveis de decisão, existe um ponto no espaço de objetivos, denotado por $\vec{f}(\vec{x}) = \vec{z} = (z_1, \dots, z_M)^T$. Esse mapeamento leva um vetor n -dimensional em um outro vetor M -dimensional. A otimização multiobjetivo é às vezes chamada de *Otimização Vetorial* pelo fato de um vetor de objetivos ser otimizado [42].

3.2 Conjunto Pareto-Ótimo

A grande maioria dos algoritmos de otimização multiobjetivo utilizam o conceito de *Dominância* na pesquisa das soluções. Nesses algoritmos duas soluções são comparadas e o resultado informa se uma domina ou não a outra. Suponha que em um problema de otimização multiobjetivo existam

M funções objetivo que se desejam minimizar. Usaremos o símbolo \prec para denotar que uma solução é melhor do que a outra, como apresentado na Definição 3.1.

Definição 3.1. Uma solução \vec{x} domina uma outra solução \vec{y} , $\vec{x} \prec \vec{y}$, se as condições 1 e 2 são verdadeiras:

1. A solução \vec{x} não é pior do que \vec{y} em nenhum objetivo, o que implica $f_e(\vec{x}) \leq f_e(\vec{y})$ para $e = 1, \dots, M$.
2. A solução \vec{x} é estritamente melhor do que \vec{y} em pelo menos um objetivo, o que implica existir $e \in \{1, \dots, M\}$ tal que $f_e(\vec{x}) < f_e(\vec{y})$.

Se alguma condição for violada então \vec{x} não domina a solução \vec{y} . O fato de uma solução \vec{x} não dominar \vec{y} , não implica que \vec{y} domina \vec{x} . Existem três possibilidades de resultados para um teste de dominância:

- \vec{x} pode dominar \vec{y} ;
- \vec{y} pode dominar \vec{x} ; e
- não existe relação de dominância entre \vec{x} e \vec{y} .

A relação de dominância não é reflexiva, é assimétrica (consequentemente não anti-simétrica) e é transitiva. A presença de transitividade em uma relação binária possibilita haver uma relação de ordenamento [77]. Pelo fato da relação de dominância ser não reflexiva e não anti-simétrica, ela assume somente uma relação de *Ordem Parcial Estrita* [42].

Para um conjunto finito de soluções, pode-se comparar todos os possíveis pares e determinar quem domina quem e quais soluções não são dominadas. Ao final da otimização, espera-se ter um conjunto de soluções que não se dominam. Em particular, este conjunto deve possuir a propriedade de dominar todas as soluções que não pertençam a ele. Esse é o conjunto das soluções não-dominadas, apresentado na Definição 3.2.

Definição 3.2 (Conjunto Não Dominado). Seja \mathcal{P} um conjunto de soluções. O conjunto de soluções não-dominadas \mathcal{P}' é constituído por aquelas que não são dominadas por nenhum membro de \mathcal{P} .

Quando \mathcal{P} é próprio conjunto factível, $\mathcal{P} = \mathcal{F}_x$, o conjunto não-dominado resultante \mathcal{P}' é chamado *Conjunto Pareto-Ótimo Global*, \mathcal{X}^* , descrito na Definição 3.3.

Definição 3.3 (Conjunto Pareto-Ótimo Global). O conjunto não-dominado de toda a região factível \mathcal{F}_x é o conjunto Pareto-Ótimo Global \mathcal{X}^* . Ou seja, $\vec{x}^* \in \mathcal{X}^* \subset \mathcal{F}_x$ é uma solução Pareto-ótimo, se não existe qualquer outra solução $\vec{x} \in \mathcal{F}_x$ tal que $f_e(\vec{x}) \leq f_e(\vec{x}^*)$ para todo $e = 1, \dots, M$ e $\vec{f}(\vec{x}) \neq \vec{f}(\vec{x}^*)$.

Em muitas ocasiões, o conjunto Pareto-Ótimo Global é simplesmente chamado de conjunto Pareto-Ótimo. Uma vez que as soluções desse conjunto não são dominadas por algum membro da região factível, essas são consideradas as soluções ótimas do problema de otimização multiobjetivo. A imagem do conjunto \mathcal{X}^* define a *Fronteira Ótima* ou *Fronteira de Pareto* no espaço de objetivos, \mathcal{Y}^* .

Definição 3.4 (Conjunto Pareto-Ótimo Local). Seja $\vec{x} \in \mathcal{F}_x$. É dito que \vec{x} é uma solução do conjunto Pareto-Ótimo Local em uma vizinhança se, para um dado $\varepsilon > 0$, não existir $\vec{y} \in \mathcal{F}_x$ com $\|\vec{x} - \vec{y}\| < \varepsilon$, de modo que $f_e(\vec{y}) \leq f_e(\vec{x})$ para todo $e = 1, \dots, M$ e $\vec{f}(\vec{y}) \neq \vec{f}(\vec{x})$.

A Definição 3.4 apresenta o conceito de Conjunto *Pareto-Ótimo Local*. Quando uma solução nesse conjunto é perturbada localmente no espaço de variáveis de decisão, nenhuma solução factível que domina qualquer membro desse conjunto pode ser encontrada. Assim, é verdade que todo conjunto Pareto-Ótimo Global é também um Pareto-Ótimo Local, porém o inverso não é verdade.

Definição 3.5 (Vetor de Objetivos Utópico). A solução utópica ou vetor de objetivos utópico \vec{z}^* do problema de minimização multiobjetivo é definido como $z_e^* = f_e(\vec{x}^{(e)})$, com $e = 1, \dots, M$ e $\vec{x}^{(e)} = \arg \min_{\vec{x} \in \mathcal{F}_x} f_e(\vec{x})$.

A Definição 3.5 descreve a *Solução Utópica* de um problema de minimização multiobjetivo [75]. Esse vetor possui o mínimo global de cada função objetivo em cada coordenada. Esse vetor do espaço de objetivos é a solução ideal, porém ela não existe no espaço de variáveis de decisão quando o problema de otimização é realmente multiobjetivo. Desse modo, não existe $\vec{x} \in \mathcal{F}_x$ que implique na solução utópica em \mathcal{Y} . Se tal solução existisse o problema seria mono-objetivo [42].

Ao contrário da ideia da solução utópica, que representa o limite inferior de cada função no espaço de objetivos, a *Solução Nadir* representa o limite superior de cada objetivo sobre todo conjunto Pareto-ótimo. O vetor de objetivos nadir não pode ser confundido com o vetor encontrado através dos piores (máximos) valores encontrados no espaço de busca para cada objetivo. A solução nadir pode representar uma solução existente, dependendo das características do conjunto Pareto-ótimo.

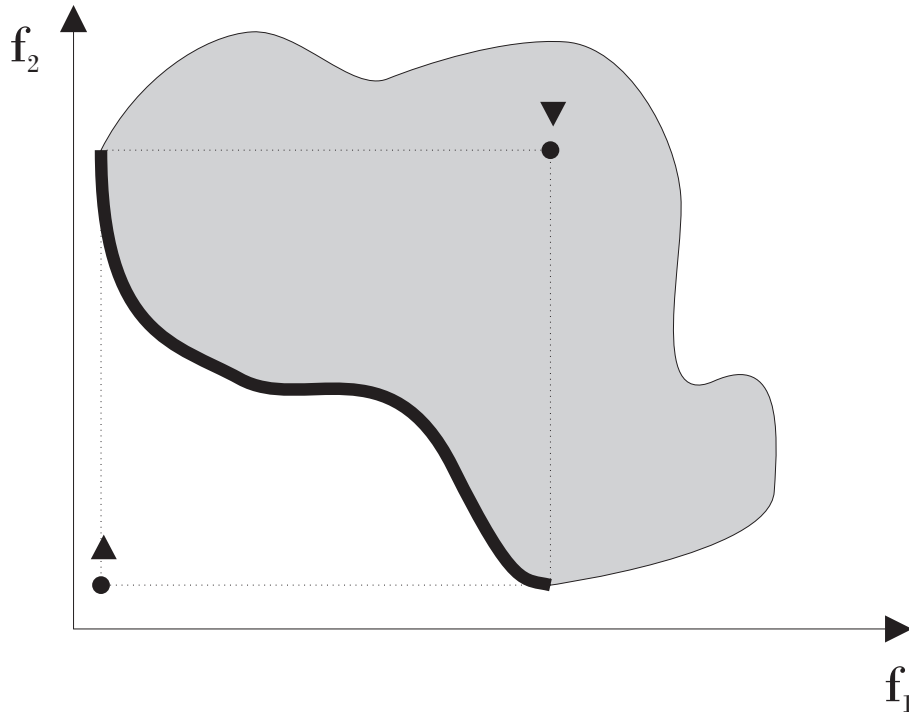


Figura 3.1: Imagem \mathcal{Y} do conjunto factível \mathcal{F}_x e, em destaque, da fronteira ótima \mathcal{Y}^* , imagem de \mathcal{X}^* . A solução utópica é representada por um triângulo com a ponta para cima e a nadir por um triângulo com a ponta para baixo.

A Figura 3.1 apresenta uma ilustração da imagem \mathcal{Y} do conjunto \mathcal{F}_x para um problema de minimização de dois objetivos. São apresentadas em destaque a fronteira ótima \mathcal{Y}^* do conjunto Pareto-ótimo \mathcal{X}^* , a solução utópica e a solução nadir.

3.3 Algoritmos Evolutivos Multiobjetivo

Segundo o naturalista Charles Darwin [78], a *Seleção Natural* é definida como a preservação de indivíduos mais adaptados ao ambiente. Ou seja, na natureza, estes indivíduos possuem maiores chances de competir, sobreviver e reproduzir. Durante a evolução, indivíduos com características mais fracas geralmente são eliminados. Tais características são controladas por unidades, denominadas *Genes*, que formam um conjunto chamado *Cromossomo*. Após gerações subsequentes, não somente os melhores indivíduos sobrevivem, mas também os melhores genes são transmitidos para os descendentes durante o processo de *Recombinação Sexual*. Essa analogia entre os mecanis-

mos de seleção natural e o processo de aprendizagem leva ao desenvolvimento dos *Algoritmos Evolucionários* (AEs), no qual se enquadram os *Algoritmos Genéticos* (AGs).

Os AGs, introduzidos por John Holland na década de 1960 [79, 80], são métodos computacionais de otimização baseados no mecanismo genético e na evolução natural. Nesse algoritmo, um conjunto de soluções tentativas (população) envolve regras probabilísticas inspiradas em metáforas biológicas, de modo que, em cada geração, os indivíduos tendem a ficar melhores ao longo do período em que o processo de evolução continua. Em geral, AGs possuem as seguintes características [81, 82]:

- AGs trabalham com um conjunto de pontos (população);
- AGs trabalham com um espaço de busca codificado;
- AGs necessitam somente da informação sobre a função objetivo para cada membro da população; e
- AGs executam transições probabilísticas.

Em qualquer AG, alguns operadores comuns são: inicialização, avaliação da função de adequabilidade, seleção, mutação e recombinação. Em alguns outros, por exemplo, podem haver busca local, nicho, dizimação, etc., dependendo do objetivo da aplicação. Goldberg [81] citou quatro razões que fazem os AGs atrativos para aplicações:

- AGs podem resolver problemas difíceis em um pequeno tempo de forma confiável;
- A interface para a construção de AGs com os modelos de problemas existentes é, geralmente, simples;
- AGs são extensíveis; e
- AGs são fáceis de hibridizar.

Uma das diferenças mais impressionantes entre os métodos clássicos determinísticos de otimização e os AEs é que esses últimos usam uma população de soluções em cada iteração, ao invés de uma solução única. Da mesma maneira que a população de soluções é processada em cada iteração, a saída de um AE é também uma população de soluções. Se for aplicado um AE em um problema de otimização mono-objetivo, espera-se que a população convirja para a solução ótima. Agora, se o problema possui múltiplos objetivos, espera-se que a população convirja para as múltiplas soluções ótimas. Como

os AEs são técnicas heurísticas, não é garantido que os mesmos encontrem o ótimo global sempre, mas sempre é esperado que o ótimo ou outras soluções boas sejam retornadas [42].

Os algoritmos evolucionários de otimização multiobjetivo são ferramentas computacionais de grande importância para a resolução de problemas com múltiplas funções objetivo conflitantes. Esse fato ocorre, em parte, porque os algoritmos evolutivos aperfeiçoam um conjunto de soluções que constitui uma estimativa do conjunto Pareto-ótimo do problema em uma única execução. Além disso, técnicas de otimização multiobjetivo evolucionária podem tratar também funções objetivo nos quais os métodos determinísticos apresentam dificuldades como descontinuidade, não convexidade, multimodalidade, presença de ruídos e assim por diante. Porém, a maioria dos algoritmos determinísticos para otimização multiobjetivo tratam funções bem comportadas e encontram um ponto somente a cada execução, que pode ainda não ser do conjunto Pareto-ótimo. Assim, esses métodos são inconvenientes para solucionar problemas de otimização multiobjetivo. Por essas razões os métodos evolucionários se mostram mais adequados para a solução desse tipo de problema e é uma área de pesquisa que recebe muitas contribuições desde a década de 1990 [83, 84, 85, 86, 87, 42].

3.4 NSGA-II

O *Non-Dominated Sorting Genetic Algorithm II* (NSGA-II) [41] é o algoritmo genético multiobjetivo base, usado nesse trabalho para determinar as soluções não dominadas. As principais características do NSGA-II são:

- O procedimento rápido para ordenação pela não dominância (*Fast Non-Dominated Sorting Procedure*), que classifica os indivíduos pelo nível da fronteira de Pareto de acordo com a relação de dominância. Por exemplo, os indivíduos que são dominados por nenhuma outra solução recebem o índice de fronteira 1, os indivíduos que são dominados somente por soluções da fronteira 1 recebem o índice 2, aqueles que são dominados somente pela fronteira 1 e 2 recebem o índice 3. Essa operação se repete até que não haja indivíduos para se classificar com relação a fronteira de Pareto. O Algoritmo 2 apresenta o pseudocódigo desse algoritmo que possui ordem de complexidade $\mathcal{O}(M\Pi^2)$, em que M é o número de objetivos e Π o número de indivíduos na população de soluções do algoritmo genético.
- A implementação de *Elitismo Multiobjetivo* para pesquisa das soluções, usando uma metodologia de preservação. O elitismo é introduzido ao

armazenar as Π melhores soluções, dentro de um limite especificado de acordo com a qualidade e espalhamento no espaço de objetivos, para passar para a próxima geração na população do algoritmo genético. Isso é feito em todas as gerações. O elitismo aprimora propriedades de convergência para a obtenção do conjunto Pareto-ótimo.

- O mecanismo de *Preservação da Diversidade* sem a necessidade de parâmetro. Diversidade e espalhamento das soluções é garantido sem o uso de parâmetros compartilhados. O NSGA-II usa uma medida de aglomeração (*Crowding Distance*), que estima a densidade das soluções, dentro de uma camada da fronteira no espaço de objetivos, e um operador de comparação, que guia o processo de seleção para que as soluções apresentem uma fronteira de Pareto uniformemente espalhada.
- O método de *Tratamento de Restrições* não usa parâmetros de penalidade, que era necessário na versão anterior desse algoritmo. O NSGA-II utiliza um vetor que computa o valor de penalidade para cada restrição, ou armazena o valor zero caso não a infringir. Cada dimensão desse vetor armazena o quanto ele restringiu com relação a cada função de restrição. O somatório dos valores absolutos desse vetor indica a intensidade da penalidade. Dessa forma, quanto maior for o somatório mais infactível é considerado. Obviamente, se o somatório for zero então o indivíduo é factível.
- Na implementação original, utilizada nesse trabalho, a seleção usada é baseada em *Torneio Binário*, de modo que dois pares de indivíduos são selecionados aleatoriamente da população. Para cada par, é realizado um torneio que usa os seguintes critérios para a vitória de um indivíduo²:
 - Se dois indivíduos são infactíveis, ganha aquele menos infactível;
 - Se um indivíduo é infactível e o outro é factível, ganha o factível;
 - Se dois indivíduos são factíveis, ganha aquele que está na melhor fronteira de Pareto, por exemplo, um indivíduo da fronteira 2 ganha daquele está na fronteira 4;
 - Se os dois indivíduos são da mesma fronteira então ganha o menos aglomerado (análise da crowding distance);
 - E se ainda assim houver empate, sorteia-se de maneira uniforme com chances de 0,5 para cada.

²Os critérios estão em ordem de prioridade, por exemplo um indivíduo da fronteira 1 ganha o torneio de um indivíduo infactível.

- A *Recombinação* ou *Cruzamento*, acontece por um determinado valor de probabilidade e é realizado com os dois pais selecionados através do torneio binário. A recombinação é baseado no *Simulated Binary Cross-Over*. Usa-se um índice de distribuição, que é um parâmetro que define a forma da distribuição de probabilidade. Para valores pequenos, pontos distantes dos pais possuem mais chances de serem escolhidos, enquanto que para grandes valores, pontos próximos dos pais possuem maior probabilidade de serem selecionados [88].
- A *Mutação*, que ocorre por um determinado valor de probabilidade é baseado no *Simulated Binary Mutation*, que imita o comportamento da mutação binária no domínio das variáveis contínuas. Usa-se um índice de distribuição, que define a forma da distribuição de probabilidade para o operador de mutação. Para valores pequenos, a perturbação gera pontos distantes, enquanto que para valores grandes é provável que estejam mais próximos do ponto original.

O Algoritmo 2 apresenta o Fast Non-Dominated Sorting, que atribui a todos os indivíduos da população, previamente avaliados pelas funções objetivo, os valores de nível de fronteira Pareto. Em resumo, esta atribuição é feito da seguinte forma:

- Inicialmente, para cada solução p calcula-se a contagem de dominância n_p (o número de soluções que a dominam) e o conjunto S_p de soluções que essa solução domina.
- A próxima fronteira a se classificar será indexada por $i = 1$;
- Serão classificadas como pertencentes à i -ésima fronteira não dominada \mathcal{X}_i , todas as soluções com $n_p = 0$;
- Enquanto houver solução não classificada, faz-se o seguinte:
 - Para todo $p \in \mathcal{X}_i$ e toda solução $q \in S_p$, faz-se $n_q \leftarrow n_q - 1$, removendo assim a informação da contagem de dominância da fronteira i ;
 - A próxima fronteira a se classificar será indexada por $i = i + 1$;
 - Serão classificadas como pertencentes à i -ésima fronteira não dominada \mathcal{X}_i , todas as soluções com $n_q = 0$.

O pior caso, com relação ao custo computacional do Fast Non-Dominated Sorting, será quando houver Π fronteiras, com um indivíduo em cada. Nesse caso, para o indivíduo da primeira fronteira é necessário fazer $\Pi - 1$ visitas

(uma para cada indivíduo dominado por ele, para decrementar os n_q 's). Para o segundo indivíduo (da segunda fronteira), são necessárias $\Pi - 2$ visitas. Continuando, para o $(\Pi - 2)$ -ésimo indivíduo (da $(\Pi - 2)$ -ésima fronteira) são necessárias 2 visitas e para o $(\Pi - 1)$ -ésimo indivíduo (da $(\Pi - 1)$ -ésima fronteira) é necessária 1 visita. No total são $\Pi(\Pi - 1)/2$ visitas, e como em cada verificação de dominância são necessárias M comparações de funções, o custo computacional será $\mathcal{O}(M\Pi^2)$.

Algorithm 2 FastNonDominatedSorting(P)

```

for all  $p \in P$  do
   $S_p \leftarrow \emptyset$ 
   $n_p \leftarrow 0$ 
  for all  $q \in P$  do
    if  $p \prec q$  then
       $S_p \leftarrow S_p \cup \{q\}$ 
    else if  $q \prec p$  then
       $n_p \leftarrow n_p + 1$ 
    end if
  end for
  if  $n_p = 0$  then
     $p_{\text{rank}} \leftarrow 1$ 
     $\mathcal{X}_1 \leftarrow \mathcal{X}_1 \cup \{p\}$ 
  end if
end for
 $i = 1$ 
while  $\mathcal{X}_i \neq \emptyset$  do
   $Q \leftarrow \emptyset$ 
  for all  $p \in \mathcal{X}_i$  do
    for all  $q \in S_p$  do
       $n_q \leftarrow n_q - 1$ 
      if  $n_q = 0$  then
         $q_{\text{rank}} \leftarrow i + 1$ 
         $Q \leftarrow Q \cup \{q\}$ 
      end if
    end for
  end for
   $i \leftarrow i + 1$ 
   $\mathcal{X}_i \leftarrow Q$ 
end while

```

Para calcular o valor da medida de aglomeração (crowding distance) de uma solução é necessário calcular uma distância entre os dois pontos de cada lado dessa solução para cada objetivo. Essa distância é calculada como sendo o perímetro do cubóide que é formado pelos vértices das soluções vizinhas mais próximas, como é apresentado pela Figura 3.2.

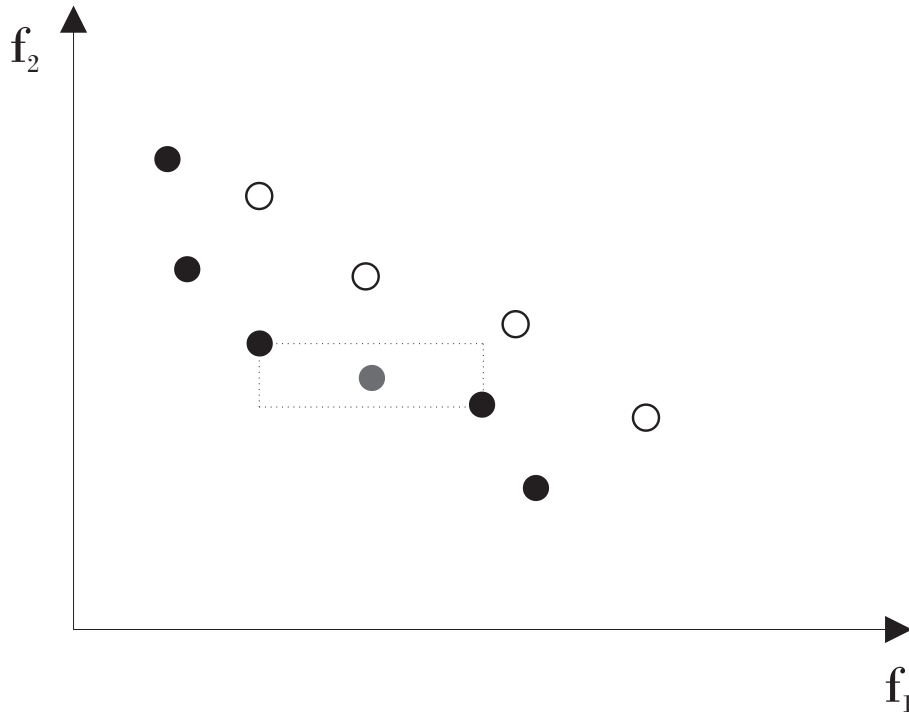


Figura 3.2: Cálculo da crowding distance. Os pontos preenchidos são soluções da mesma fronteira não dominada.

O cálculo da crowding distance requer a ordenação da população de acordo com cada função objetivo em ordem crescente. Para cada função objetivo, as soluções limites (com menor e maior valor) são assinaladas com o valor de distância infinito, pois elas obrigatoriamente serão preservadas para a próxima geração. Todas as outras soluções intermediárias recebem o valor de distância igual à diferença absoluta normalizada dos valores de função de duas soluções adjacentes. Essa computação é realizada para as outras funções objetivo. O valor final da crowding distance é calculado como a soma das distâncias individuais para cada objetivo. Cada função objetivo é normalizada antes de se calcular a crowding distance.

O Algoritmo 3 mostra como calcular a crowding distance para todas as soluções de um conjunto não dominado \mathcal{I} . Nesse algoritmo, $\mathcal{I}[i].e$ refere-se ao e -ésimo valor de função objetivo do i -ésimo indivíduo e os parâmetros f_e^{\max} e

f_e^{\min} são os valores máximos e mínimos do e -ésimo objetivo encontrados. A complexidade desse procedimento é governada pelo algoritmo de ordenação. Uma vez que há M ordenações independentes, usando um método eficiente, a complexidade computacional será $\mathcal{O}(M\Pi \log \Pi)$ para Π soluções.

Algorithm 3 CrowdingDistance(\mathcal{I})

```

size ← | $\mathcal{I}$ |
for all  $i \in \mathcal{I}$  do
   $\mathcal{I}[i]_{\text{distance}} \leftarrow 0$ 
end for
for all objective  $e$  do
   $\mathcal{I} \leftarrow \text{sort}(\mathcal{I}, e)$ 
   $\mathcal{I}[1]_{\text{distance}} \leftarrow \infty$ 
   $\mathcal{I}[\text{size}]_{\text{distance}} \leftarrow \infty$ 
  for  $i$  from 2 to size – 1 step 1 do
     $\mathcal{I}[i]_{\text{distance}} \leftarrow \mathcal{I}[i]_{\text{distance}} + (\mathcal{I}[i + 1].e - \mathcal{I}[i - 1].e) / (f_e^{\max} - f_e^{\min})$ 
  end for
end for

```

Uma solução que possui valor baixo de crowding distance está mais aglomerada. É interessante que se duas soluções da mesma fronteira forem comparadas, deve ganhar aquela que possuir o valor mais alto de medida de aglomeração. Desse modo, a diversidade das soluções será mantida. Por esse motivo, é apresentado na Definição 3.6 o operador de comparação de aglomeração (*Crowded-Comparison Operator*), que é usado dentro do NSGA-II. Durante o processo de seleção entre duas soluções que se diferem do nível de fronteira, prefere-se aquela com menor (melhor) fronteira. Entre duas soluções da mesma fronteira, prefere-se a menos aglomerada. O operador de comparação de aglomeração guia o NSGA-II em diversos estágios fazendo com que a fronteira de Pareto fique uniformemente espalhada.

Definição 3.6 (Operador de Comparação de Aglomeração (Crowded-Comparison Operator)). Sejam i e j duas soluções das fronteiras i_{rank} e j_{rank} , com medidas de aglomeração i_{distance} e j_{distance} . O Operador de Comparação de Aglomeração (\prec_n) é definido como:

$$i \prec_n j \equiv (i_{\text{rank}} < j_{\text{rank}}) \text{ ou } ((i_{\text{rank}} = j_{\text{rank}}) \text{ e } (i_{\text{distance}} > j_{\text{distance}}))$$

No módulo principal do NSGA-II, uma população aleatória de pais P_1 é criada com Π indivíduos. Essa população inicial é ordenada pela não dominância, obtendo-se as l primeiras fronteiras $\mathcal{X}^{(1)} = (\mathcal{X}_1, \dots, \mathcal{X}_l)$ da primeira geração, com l camadas. A cada solução (indivíduo) é assinalado um valor

de *Adequabilidade* (“*fitness*” ou “*rank*”) igual ao nível da fronteira de não dominação (1 para o melhor, 2 para o próximo e assim por diante).

Para cada geração g das ng gerações, é realizado torneio binário, recombinação e mutação para se criar uma nova população de filhos Q_g de tamanho Π . Uma vez que o elitismo é introduzido, é comparado a população de 2Π indivíduos de pais e filhos, R_g , para se extrair as Π melhores soluções para se tornar a próxima população P_{g+1} de pais da próxima geração. A extração dos Π melhores indivíduos, ilustrada pelo Algoritmo 4, é executada da seguinte maneira:

- Ordena-se R_g pela não dominância, classificando as soluções em novas fronteiras $\mathcal{X}^{(g)} = (\mathcal{X}_1, \dots, \mathcal{X}_l)$ da primeira geração, com l camadas;
- Deve-se extrair as Π melhores soluções, das 2Π de R_g , para se tornar a próxima população de pais P_{g+1} . Deve-se incluir todas as soluções das primeiras fronteiras por completo em P_{g+1} até não exceder Π soluções. Se por acaso a j -ésima fronteira for a última que deve ser incluída por completo, não preenchendo P_{g+1} com Π soluções, e se a inclusão do total de soluções da $(j + 1)$ -ésima fronteira exceder Π , então:
 - Deve-se ordenar as soluções da $(j + 1)$ -ésima fronteira pelo operador de comparação de aglomeração (\prec_n) de forma crescente; e finalmente
 - Inserir as primeiras soluções ordenadas da $(j + 1)$ -ésima fronteira até completar Π indivíduos em P_{g+1} .

Algorithm 4 FillNondominated(R_g, \mathcal{X})

```

 $P \leftarrow \emptyset$ 
 $\mathcal{X} \leftarrow \text{FastNonDominatedSorting}(R_g)$ 
 $i = 1$ 
while  $|P| + |\mathcal{X}_i| \leq \Pi$  do
   $P \leftarrow P \cup \mathcal{X}_i$ 
   $i \leftarrow i + 1$ 
end while
 $\mathcal{X}_i \leftarrow \text{sort}(\mathcal{X}_i, \prec_n)$ 
 $P \leftarrow P \cup \mathcal{X}_i(1 : \Pi - |\mathcal{X}_i|)$ 

```

Dessa maneira, a evolução do NSGA-II ocorre até um critério de parada desejado, em que no caso particular é simular a evolução de ng gerações. O funcionamento geral do NSGA-II é ilustrado pela Figura 3.3 e pelo Algoritmo

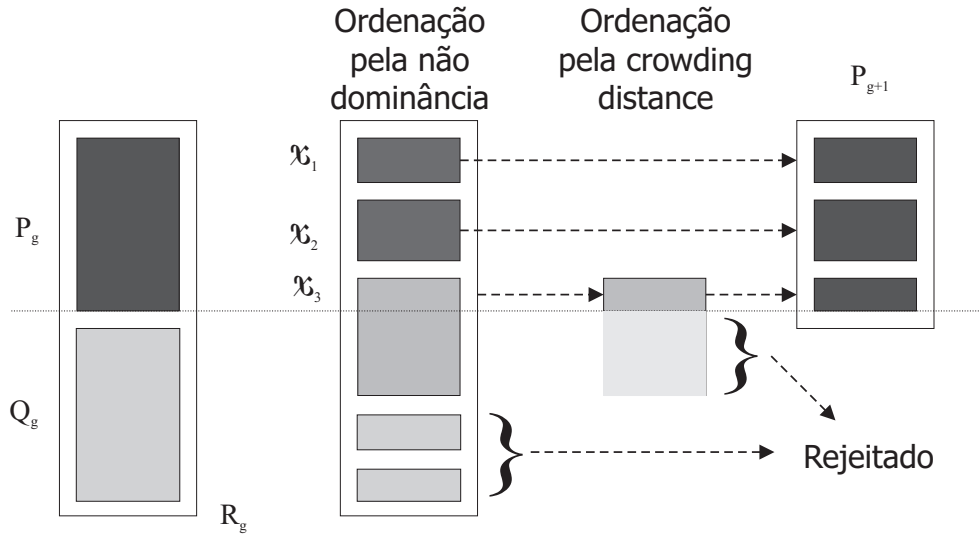


Figura 3.3: Funcionamento geral do NSGA-II.

5. Ao final, espera-se que a população tenha convergido para soluções em uma única fronteira não dominada, uniformemente espaçada no espaço de objetivos e que esteja, pelo menos, próxima do conjunto Pareto-ótimo.

Algorithm 5 NSGA-II

```

 $P_1 \leftarrow \text{InitialPopulation}()$ 
 $P_1 \leftarrow \text{EvaluatePopulation}(P_1)$ 
 $\mathcal{X}^{(1)} \leftarrow \text{FastNonDominatedSorting}(P_1)$ 
 $\mathcal{X}^{(1)} \leftarrow \text{CrowdingDistance}(\mathcal{X}^{(1)})$ 
for  $g$  from 1 to  $ng$  step 1 do
   $Q_g \leftarrow \text{SelectionAndRecombination}(P_g)$ 
   $Q_g \leftarrow \text{Mutation}(Q_g)$ 
   $Q_g \leftarrow \text{EvaluatePopulation}(Q_g)$ 
   $R_g \leftarrow P_g \cup Q_g$ 
   $\mathcal{X}^{(g)} \leftarrow \text{FastNonDominatedSorting}(R_g)$ 
   $P_{g+1} \leftarrow \text{FillNondominated}(R_g, \mathcal{X}^{(g)})$ 
end for

```

3.5 Algoritmos Meméticos

É sabido que os métodos de otimização determinísticos possuem propriedades de convergência rápida e precisa, mas são limitados apenas para encontrar

ótimos locais. Isso é um problema quando as funções objetivo não são unimodais, uma situação muito comum em problemas de otimização do mundo real. Os chamados *Algoritmos Meméticos* (AMs), uma hibridização entre algoritmos evolucionários e métodos determinísticos de busca, que possui o objetivo de melhorar as propriedades de busca local, surgiu como uma ideia promissora. Alguns trabalhos relacionados que utilizam AMs para otimizar um ou mais objetivos são [89, 90, 91, 92, 93, 94, 95, 96].

Na proposta original dos AMs, cada população de filhos gerada por um operador evolutivo (recombinação ou mutação, por exemplo) era submetida imediatamente para uma fase de busca que leva a um ótimo local. Os ótimos locais gerados dessa forma constituem a população que será submetida aos operadores evolutivos na próxima geração. Os AMs, definidos dessa maneira, apresentam um bom desempenho em vários contextos. No entanto, a combinação de métodos de pesquisa global de base populacional com algoritmos de *Busca Local*, que realizam uma pesquisa até atingir uma solução que apresente otimalidade local para cada novo indivíduo, em geral, apresenta-se com um alto custo computacional. Isso se deve ao elevado número de avaliações das funções objetivo na fase de busca local. Esse fato gera uma grande dificuldade para a aplicação desse tipo de AM em muitos casos. Tais preocupações levaram ao relaxamento da definição dos Algoritmos Meméticos, permitindo que somente alguns indivíduos da população, ao invés de toda, fossem selecionados para serem submetidos ao procedimento de busca local [92].

Outras estratégias também relaxam a exigência da a busca local poder encontrar pontos que são, de fato, ótimos locais, devido ao fato dessa tarefa poder ser computacionalmente cara e, em vários casos, poder ser difícil a caracterização da otimalidade local. Essas estratégias, portanto, empregam motores de busca locais que melhoram os indivíduos, sem a garantia de encontrar ótimos locais [92].

O presente trabalho elaborou um algoritmo memético, que é o algoritmo genético multiobjetivo NSGA-II acoplado com um operador de busca local baseado na otimização de somas ponderadas das aproximações quadráticas convexas das funções objetivo originais. A aproximação quadrática convexa apresentada nesse trabalho possui Hessiana diagonal semidefinida positiva, que é mais adequada para problemas de dimensões mais alta, conforme apresentado em um trabalho do autor [97].

A técnica para a realização de aproximações quadráticas convexas é apresentada na Seção 3.6 e a busca local baseada na otimização de somas ponderadas das aproximações quadráticas convexas das funções objetivo originais é apresentada na Seção 3.7.

3.6 Aproximação Quadrática Convexa

Uma possível estratégia para a realização de buscas locais em funções contínuas é determinar aproximações locais simples que são fáceis de otimizar. O uso de funções quadráticas para aproximar os dados no \mathfrak{R}^n tem sido extensivamente estudado [98, 99, 100, 101], podendo ser baseado em interpolação de um número específico de pontos de fornecidos [102] ou através da minimização do erro, em uma norma adequada, da aproximação quadrática [100]. Nesses métodos, a aproximação quadrática pode ser indefinida e erros nos dados podem gerar um grande efeito sobre a precisão da aproximação [97]. É provado que a aproximação quadrática, pelo método dos mínimos quadrados, de uma função multivariada por um conjunto finito de pontos não é necessariamente convexa, mesmo que a função original seja convexa [103]. Muitos trabalhos recentes apontam essas dificuldades, o que sugere que o requerimento da convexidade deva ser imposta no processo de aproximação quadrática [104, 105, 106, 107]. Nesses trabalhos, uma aproximação quadrática convexa foi aplicada para aproximar um conjunto de mínimos locais, como parte do método iterativo para encontrar o mínimo global.

A presente seção apresenta as técnicas para construir as *Aproximações Quadráticas Convexas* que são usadas no operador de busca local. Esse operador é acoplado ao NSGA-II para melhorar a qualidade do conjunto de soluções eficientes para o problema multiobjetivo dado.

A aproximação quadrática convexa apresenta as seguintes vantagens:

- A convexidade da aproximação permite que as poderosas ferramentas da análise convexa sejam aplicáveis. Existem métodos computacionais eficazes para resolver problemas de otimização convexa quadrática [76].
- É uma boa ferramenta para aproximar localmente a função original em uma grande parcela do domínio [95].
- Se a função original possui uma tendência global e o conjunto de pontos é bem escolhido, a aproximação pode ser usada para direcionar a pesquisa para o ótimo global [107].

Definição 3.7 (Função Convexa). Uma função $g : \mathfrak{R}^n \rightarrow \mathfrak{R}$ é convexa se, para quaisquer $\vec{x}, \vec{y} \in \mathfrak{R}^n$ e para todo $0 \leq \alpha \leq 1$, é verdade que

$$g(\alpha\vec{x} + (1 - \alpha)\vec{y}) \leq \alpha g(\vec{x}) + (1 - \alpha)g(\vec{y}). \quad (3.2)$$

A partir da Definição 3.7, pode-se afirmar sobre *Funções Convexas* as seguintes propriedades:

- Para todo $\alpha \geq 0$ a função αg é convexa [76];

- Para toda função convexa g_1, \dots, g_M e todo $\alpha_e \geq 0$, $e = 1, \dots, M$, tem-se que $g = \sum_{i=1}^M \alpha_e g_e$ é também convexa e, portanto, o ótimo de g pode ser determinado via programação quadrática convexa [76];
- Sendo g convexa, tem-se $g(\vec{y}) \geq g(\vec{x}) + \nabla g(\vec{x})^T(\vec{y} - \vec{x})$, em que ∇g é o gradiente de g e $\vec{x}, \vec{y} \in \mathfrak{R}^n$ [76, 75];
- A matriz Hessiana da função convexa g é definida ou semidefinida positiva para todo \vec{x} [76, 75];
- Um mínimo local de uma função convexa é também o mínimo global [76, 75].

Considere uma função real $f : \mathfrak{R}^n \rightarrow \mathfrak{R}$, com uma amostra distinta de dados $\{(\vec{x}^{(k)}, f(\vec{x}^{(k)}))\}$, para $k = 1, \dots, m$, em que m é o número de pontos disponível para ser usado no ajuste de uma função real quadrática convexa real $q : \mathfrak{R}^n \rightarrow \mathfrak{R}$, de modo que $f(\vec{x}^{(k)}) \cong q(\vec{x}^{(k)})$. Uma função quadrática é convexa se a matriz Hessiana da mesma é, pelo menos, positiva semidefinida [76]. Uma aproximação convexa força a existência de uma fatoração de Cholesky da matriz Hessiana da função q [97, 34].

A fórmula de uma *Função Quadrática* é apresentada na Equação 3.3, em que $\vec{x} \in \mathfrak{R}^n$ é um vetor no espaço, $h_0 \in \mathfrak{R}$ é um escalar, $\nabla \vec{h} \in \mathfrak{R}^n$ é o vetor gradiente e $H \in \mathfrak{R}^{n \times n}$ é a Hessiana. A aproximação quadrática convexa pode ser descrita por uma função $q(\vec{p}, \vec{x})$, em que $\vec{p} \in \mathfrak{R}^s$ é o vetor dos parâmetros da função quadrática q , sendo s a quantidade de parâmetros.

$$q(\vec{x}) = h_0 + \nabla \vec{h}^T \vec{x} + 0,5 \vec{x}^T H \vec{x} \quad (3.3)$$

Nesse trabalho, será usada uma aproximação quadrática convexa com Hessiana diagonal definida ou semidefinida positiva, obtida através da solução de um problema de programação linear. Este modelo se apresenta mais viável em problemas de dimensão alta, com relação ao custo computacional, do que o modelo com Hessiana simétrica cheia (com elementos não-nulos fora da diagonal principal) [97]. Dessa maneira, a aproximação pode ser simplificada, obtendo-se a fórmula quadrática apresentada na Equação 3.4, em que $\vec{x} \in \mathfrak{R}^n$, $d_0 \in \mathfrak{R}$, $\nabla \vec{d} \in \mathfrak{R}^n$ e $D \in \mathfrak{R}^{n \times n}$ é a matriz Hessiana com elementos d_1, \dots, d_n na diagonal principal, com $d_i \geq 0$ para todo $i = 1 \dots n$. Assim, o número de parâmetros da aproximação quadrática convexa é $s = 1 + n + n = 2n + 1$. Supõe-se que o número de pontos fornecidos é suficiente, de modo que $m \geq s$. Em particular, foi adotado nesse trabalho $m = 2s$, para que a aproximação

quadrática convexa seja mais consistente.

$$q(d_0, \nabla \vec{d}, D; \vec{x}) = d_0 + \nabla \vec{d}^T \vec{x} + 0,5 \vec{x}^T D \vec{x} = d_0 + \sum_{i=1}^n d_i x_i + 0,5 \sum_{i=1}^n d_i x_i^2, \quad (3.4)$$

Para um vetor fixo \vec{x} , q é linear em relação a d_0 , \vec{d} , e D . Consequentemente, para encontrar a aproximação soluciona-se o problema de otimização linear [97] apresentado na Equação 3.5. O problema de minimização possui $2n + 1$ variáveis e $m + n$ restrições de desigualdade. Esse problema de programação linear pode ser solucionado com algoritmos eficientes para programação linear, como por exemplo os métodos baseados em pontos interiores ou simplex [108].

$$\min_{d_0, \vec{d}, D} \sum_{k=1}^m \left(f(\vec{x}^{(k)}) - q(d_0, \vec{d}, D; \vec{x}^{(k)}) \right) \quad (3.5)$$

sujeito a:

$$\begin{aligned} f(\vec{x}^{(k)}) &\geq q(d_0, \vec{d}, D; \vec{x}^{(k)}), & k = 1, \dots, m \\ d_i &\geq 0, & i = 1, \dots, n \end{aligned}$$

3.7 NSGA-II com Operador de Busca Local

Considere o problema geral de minimização multiobjetivo apresentado pela Equação 3.1. Uma implementação memética do NSGA-II com a busca local, baseada na otimização de somas ponderadas das aproximações quadráticas convexas das funções objetivo originais, foi elaborada para a realização da otimização multiobjetivo. O algoritmo híbrido foi construído tendo o NSGA-II como algoritmo base, com a seguinte busca local:

- Um conjunto de r pontos é escolhido a partir do conjunto das soluções não dominadas até a presente geração;
- Um conjunto de m pontos é gerado aleatoriamente, com distribuição de gaussiana com média zero e variância positiva próxima de zero³, em torno de cada ponto escolhido na última etapa. Cada nuvem gerada possui $m = 2s$ pontos, em que s é o número de parâmetros da função quadrática aproximada. No presente caso $s = 2n + 1$, em que n é a dimensão do espaço de variáveis;

³De maneira a gerar pontos na vizinhança. Deve-se ter atenção com relação a escala da região factível.

- Todos os novos m pontos, $\{\vec{x}^{(1)}, \dots, \vec{x}^{(m)}\}$, são avaliados nas funções objetivo originais f_1, \dots, f_M ;
- As aproximações quadráticas convexas q_1, \dots, q_M são determinadas, através da solução da Equação 3.5, respectivamente, para cada função objetivo f_1, \dots, f_M , e para cada um dos r conjuntos de $m + 1$ pontos;
- As funções de aproximação quadráticas convexas são utilizadas para gerar p estimativas de soluções Pareto-ótimo das funções objetivo originais, utilizando o procedimento de soma ponderada e um algoritmo de programação quadrática, como por exemplo o método Quasi-Newton [75]. Ou seja, são fornecidos p conjuntos distintos de pesos $\alpha_e \geq 0$ com $\sum_{e=1}^M \alpha_e = 1$ de forma que a solução do problema apresentado pela Equação 3.6 seja solução do Pareto-ótimo do problema de minimização das aproximações quadráticas [75];
- Os $r \times p$ pontos novos ótimos das aproximações quadráticas são avaliados de acordo com as funções objetivo originais e são adicionados a população atual do algoritmo genético.

$$\min_{\vec{x}} \quad \tilde{q}(\vec{x}) = \sum_{e=1}^M \alpha_e q_e(\vec{x}) \quad e = 1, \dots, M \quad (3.6)$$

sujeito a:

$$\begin{aligned} g_j(\vec{x}) &\leq 0 & j &= 1, \dots, J \\ h_k(\vec{x}) &= 0 & k &= 1, \dots, K \\ x_i^{(L)} &\leq x_i \leq x_i^{(U)} & i &= 1 \dots n \end{aligned}$$

O pseudocódigo da busca local baseada na otimização de somas ponderadas das aproximações quadráticas convexas das funções objetivo originais é apresentado no Algoritmo 6. Para cada solução i selecionada de um conjunto, é gerado em torno da mesma uma nuvem de pontos através de *GeneratePoints*(i). Essa nuvem é avaliada em *Evaluate*($U, \{f_1, \dots, f_M\}$) com relação a todas as funções objetivo do problema de otimização multiobjetivo. Em *ObtainConvexQuadApprox*($\{F_1, \dots, F_M\}$) é obtido as aproximações quadráticas convexas de cada função objetivo. Assim, é obtido no laço de repetição interno os p pontos ótimos das somas ponderadas das quadráticas convexas obtidas previamente.

Sendo b o número de vezes em que a busca local ocorre durante o processo de otimização, o número de funções objetivo avaliadas no NSGA-II com busca

Algorithm 6 LocalSearch(\mathcal{X})

```

 $X \leftarrow \text{SelectIndividuals}(\mathcal{X}, r)$ 
for all  $i \in X$  do
   $U \leftarrow \text{GeneratePoints}(i)$ 
   $[F_1, \dots, F_M] \leftarrow \text{Evaluate}(U, \{f_1, \dots, f_M\})$ 
   $[q_1, \dots, q_M] \leftarrow \text{ObtainConvexQuadApprox}(\{F_1, \dots, F_M\})$ 
   $L \leftarrow \emptyset$ 
  for  $j$  from 1 to  $p$  step 1 do
     $s \leftarrow \text{Solve Equation 3.6 using } \{\alpha_1, \dots, \alpha_M\}^{(j)}$ 
     $L \leftarrow L \cup \{s\}$ 
  end for
end for

```

local é apresentado na Equação 3.7.

$$F_{aval}^{(gap)} = ng \cdot \Pi + b \cdot r (m + p) \quad (3.7)$$

No presente trabalho, o número de buscas locais realizadas dentro do algoritmo otimizador foi $b = 5$. O número de pontos selecionados para a busca é $r = 2$. Um dos pontos selecionados é um dos extremos da primeira fronteira da solução não dominada, possuindo cada uma chances iguais de serem escolhidas. Isso é feito, principalmente, para aumentar a abrangência das soluções no espaço de objetivos. O outro ponto é selecionado aleatoriamente da primeira fronteira. O número de pontos “ótimos” obtidos por cada conjunto de funções aproximadas foi escolhido como sendo $p = \Pi/2$ (2 divisor de Π), em que Π é tamanho da população do algoritmo genético. Isso significa que, a cada vez que a busca local é efetuada, Π pontos são gerados, e o processo de seleção de indivíduos, em $\text{FillNondominated}(R_g, \mathcal{X}^{(g)})$, para a próxima geração é executada sobre 3Π indivíduos, no qual 2Π são provenientes da iteração usual do NSGA-II e Π da busca local.

Assim, o código em alto nível do NSGA-II é apresentado no Algoritmo 7.

Algorithm 7 NSGA-II com Aproximação Quadrática

```

 $P_1 \leftarrow \text{InitialPopulation}()$ 
 $P_1 \leftarrow \text{EvaluatePopulation}(P_1)$ 
 $\mathcal{X}^{(1)} \leftarrow \text{NonDominatedSorting}(P_1)$ 
 $\mathcal{X}^{(1)} \leftarrow \text{CrowdingDistance}(\mathcal{X}^{(1)})$ 
for  $g$  from 1 to  $ng$  step 1 do
   $Q_g \leftarrow \text{SelectionAndRecombination}(P_g)$ 
   $Q_g \leftarrow \text{Mutation}(Q_g)$ 
  if It is desired to execute local search then
     $L \leftarrow \text{LocalSearch}(\mathcal{X}_1^{(g)})$ 
     $Q_g \leftarrow Q_g \cup L$ 
  end if
   $Q_g \leftarrow \text{EvaluatePopulation}(Q_g)$ 
   $R_g \leftarrow P_g \cup Q_g$ 
   $\mathcal{X}^{(g)} \leftarrow \text{NonDominatedSorting}(R_g)$ 
   $P_{g+1} \leftarrow \text{FillNondominated}(R_g, \mathcal{X}^{(g)})$ 
end for

```

Capítulo 4

Metodologia

O presente trabalho propõe um controle impulsivo de epidemias utilizando um procedimento de otimização dinâmica com variáveis contínuas em um intervalo aberto. Assim, esse capítulo apresenta o planejamento do projeto experimental do presente trabalho. É descrita a metodologia usada nos testes computacionais e em análises para:

1. Determinar as melhores políticas de planejamento de campanhas de vacinação para o controle de epidemias, através da minimização concomitante dos custos de implementação e da integral da população de infectados em um horizonte de tempo dado; e
2. Validar as soluções encontradas, no item anterior, com o MBI e aplicar a dominância estocástica, a fim de eliminar as soluções sensíveis à variação de parâmetros e extrair informações sobre probabilidade de erradicação e intervalos de confiança.

Para gerar as políticas de controle não dominadas deve-se construir um modelo de otimização multiobjetivo, a partir do qual serão geradas as soluções pelo NSGA-II híbrido descrito no capítulo anterior. Essas soluções não dominadas serão validadas através do MBI com uma simulação Monte Carlo, em que as funções objetivo serão reavaliadas na presença de aleatoriedade. Ao final, será aplicada a dominância estocástica utilizando os quantis 0,25, 0,50 e 0,75 das funções objetivo. Sobre as soluções não dominadas estocasticamente serão extraídos os intervalos de confiança dos comportamentos das populações de suscetíveis infectados e recuperados e as probabilidades de erradicação de cada solução.

4.1 Modelo de Otimização

A principal ideia do *Controle Impulsivo* [27] é dividir o intervalo de tempo contínuo em um número finito de estágios, executando as ações de controle impulsivamente em alguns instantes de tempo. O sistema dinâmico mantém a dinâmica autônoma nos intervalos abertos de tempo entre as consecutivas ações de controle impulsivas.

A solução do problema de controle de programação dinâmica impulsiva é determinado em *Malha Aberta*, utilizando o algoritmo evolutivo que permite o uso de funções objetivo e restrições arbitrárias, e é uma alternativa ao uso de algoritmos enumerativos, que tem complexidade computacional proibitiva [109].

O horizonte de tempo $[0, T_{ctrl}]$ é particionado em um conjunto de instantes de tempo, não previamente fixados $\Gamma = \{\tau_0, \dots, \tau_C\}$, de modo que $\tau_k < \tau_{k+1}$, $\tau_0 = 0$ e $\tau_C = T_{ctrl}$. Esses intervalos não são necessariamente equidistantes. O instante de tempo τ_k^+ é definido como o momento “logo depois” da ação impulsiva em τ_k .

As variáveis de otimização do problema são:

- O número de campanhas (estágios) de vacinação C ;
- Os intervalos de tempo entre cada campanha $\Gamma = \{\tau_0, \dots, \tau_C\}$; e
- Os percentuais da população de suscetíveis que devem ser vacinados em cada campanha $P = \{p[1], \dots, p[C]\}$, para cada instante de tempo em Γ .

A notação $x[k] = x(\tau_k)$ e $x[k^+] = x(\tau_k^+)$ de variável discreta no tempo é considerada para cada estado Suscetível (S), Infectado (I) ou Recuperado (R). Uma vez que as ações de vacinação agem somente nas taxas de suscetíveis, considera-se a Equação 4.1 para a atualização do sistema.

$$\begin{aligned}
 S[k^+] &= S[k](1 - p[k]); \\
 I[k^+] &= I[k]; \\
 R[k^+] &= R[k] + S[k]p[k]; \\
 k &= 0, 1, \dots, C - 1.
 \end{aligned} \tag{4.1}$$

Cada problema de valor inicial na Equação 2.1 é válido durante o intervalo de tempo aberto em que não há ação de controle (vacinação) no sistema, de modo que o sistema apresenta dinâmica autônoma em tais intervalos. Uma nova condição inicial é estabelecida para o sistema em cada estágio de acordo

com a Equação 4.1, ligando o estágio ao próximo. O conjunto de condições iniciais no tempo $t = 0$ deve ser previamente conhecido caso uma simulação explícita do sistema deva ser executada.

As restrições do problema foram modeladas da seguinte forma:

- Cada intervalo de tempo entre os estágios $\Delta\tau_k = \tau_{k+1} - \tau_k$ deve obedecer a relação $0 < tmin \leq \Delta\tau_k \leq tmax < T_{ctrl}$.
- A soma dos $\Delta\tau_k$'s não pode exceder o tempo de controle T_{ctrl} e, dessa maneira, determina-se o número de campanhas C .
- A taxa de vacinação deve seguir a regra $0 < pmin \leq p[k] \leq pmax \leq 1$.
- É desejado que o número de infectados, a partir do tempo $t = T_{ini}$, $T_{ini} \leq t \leq T_{ctrl}$, limite-se ao valor $I_{tol} > 0$, ou seja, $I(t) \leq I_{tol}$.

O modelo de otimização biobjetivo possui duas funções que devem ser minimizadas, que são:

- F_1 - a integral da população infectada durante o horizonte de tempo dado para se efetuar o controle; e
- F_2 - o custo total com as campanhas (custo com as vacinas adicionado a um custo fixo por cada campanha, por exemplo, transporte, propaganda, etc.).

Deve-se notar que diferentes soluções podem possuir quantidades de campanhas distintas, ou seja, o problema de otimização trabalha com vários espaços de variáveis de decisão com diferentes dimensões. Por exemplo, $C \in \{[T_{ctrl}/tmax], [T_{ctrl}/tmax] + 1, \dots, [T_{ctrl}/tmin]\}$ na função $F_2 : \mathfrak{R}^{2C+1} \rightarrow \mathfrak{R}$. Isso ocorre por causa da flexibilidade do modelo de otimização que será proposto e é um desafio para o otimizador. Esse também é outro motivo para se utilizar algoritmos evolutivos, que são capazes de lidar com essa tarefa de trabalhar com diferentes espaços de variáveis de decisão ao mesmo tempo.

Assim, o *Modelo Biobjetivo de Otimização* pode ser formulado como apresentado pela Equação 4.2.

$$\min_{C, \Gamma, P} \begin{cases} F_1 = \int_0^{T_{ctrl}} I(t) dt \\ F_2 = c_1 \cdot C + c_2 \cdot \sum_{k=0}^C S[k] \cdot p[k] \end{cases} \quad (4.2)$$

sujeito a:

$$\left\{ \begin{array}{l}
 \frac{dS}{dt} = \mu N - \mu S - \frac{\beta IS}{N}, \quad S(0) = S_o \geq 0; \\
 \frac{dI}{dt} = \frac{\beta IS}{N} - \gamma I - \mu I, \quad I(0) = I_o \geq 0; \\
 \frac{dR}{dt} = \gamma I - \mu R, \quad R(0) = R_o \geq 0; \\
 t \in (\tau_k^+, \tau_{k+1}); \\
 S[\tau_k^+] = S[k^+] = S[k](1 - p[k]); \\
 I[\tau_k^+] = I[k^+] = I[k]; \\
 R[\tau_k^+] = R[k^+] = R[k] + S[k]p[k]; \\
 k = 0, 1, \dots, C - 1; \\
 0 < tmin \leq \Delta\tau_k \leq tmax < T_{ctrl}; \\
 0 < pmin \leq p[k] \leq pmax \leq 1; \\
 \sum_{k=1}^C \Delta\tau_k \leq T_{ctrl}; \\
 I(t) \leq I_{tol}, \quad t \in (T_{ini}, T_{ctrl}); \\
 C \in \{ \lfloor T_{ctrl}/tmax \rfloor, \lfloor T_{ctrl}/tmax \rfloor + 1, \dots, \lfloor T_{ctrl}/tmin \rfloor \}.
 \end{array} \right.$$

Os parâmetros considerados nesse trabalho são $T_{ctrl} = 360$ (ut), $tmin = 5$ (ut), $tmax = 20$ (ut), $pmin = 0,4$, $pmax = 0,8$, $T_{ini} = 100$ (ut), $I_{tol} = 5$, $c_1 = 1000$, $c_2 = 1$. Os outros valores são apresentados na Tabela 2.1.

4.2 Configuração do NSGA-II

Para gerar as soluções não dominadas, mais próximas quanto possível do conjunto Pareto-Ótimo, foi elaborada a versão especialmente adaptada do NSGA-II acoplado com uma busca local (*gap*). Os parâmetros do algoritmo são apresentados pela Tabela 4.1. A busca local ocorre de acordo com o que foi apresentado na Seção 3.7.

Em cada execução do NSGA-II memético, a busca local ocorrerá nas gerações 5, 10, 20, 40 e 80. Esse algoritmo será executado 30 vezes. Ao final desse processo, as soluções não dominadas de cada execução serão reunidas em um único conjunto. Esse conjunto final será submetido a uma ordenação pela não dominância, após a qual será extraída a fronteira não dominada.

Parâmetros	Valor
Número de gerações	100
Número de indivíduos no NSGA-II	40
Taxa de recombinação	0,8
Índice de distribuição para recombinação	10
Taxa de mutação	0,05
Índice de distribuição para mutação	10

Tabela 4.1: Parâmetros do NSGA-II com busca local.

Para cada execução do NSGA-II memético, o número de avaliações efetuadas da função objetivo, de acordo com a Equação 3.7, será $F_{aval}^{(gap)} = 40 \cdot 100 + 5 \cdot 2(4n + 2 + 20)$. Aqui, assume-se que n é a maior dimensão possível para o problema, que é o piso do tempo de controle dividido pelo menor intervalo entre cada campanha de vacinação, ou seja, $n = \lfloor T_{ctrl}/tmin \rfloor = 360/5 = 72$. Assim, $F_{aval}^{(gap)} = 7100$.

Para ilustrar a eficiência do otimizador com busca local, o NSGA-II canônico (*can*) será executado 30 vezes, para comparar as soluções dos dois algoritmos. Os parâmetros do NSGA-II sem busca serão os mesmos da Tabela 4.1, com exceção do número de gerações que será 178. Essa alteração no número de gerações se faz necessária para uma comparação justa com relação ao número de funções objetivo avaliadas no total. Dessa maneira, ocorrerá $F_{aval}^{(can)} = 7120$ avaliações de funções objetivo para o NSGA-II simples. A medida de qualidade *S-Metric*, explicada na próxima sessão, será usada para comparar os conjuntos não dominados de acordo com o número final (e equivalente) de avaliações de funções objetivo. Esta medida também é usada para ilustrar a evolução da qualidade dos conjuntos não dominados com o passar das gerações.

A união das soluções finais retornadas pelo NSGA-II com e sem busca local será submetida a uma ordenação pela não dominância. Do conjunto final de soluções não dominadas, será contabilizado o número de soluções provenientes de cada método. Esse conjunto final de soluções será submetido à validação pela simulação Monte Carlo através do MBI.

4.3 S-Metric

Para medir a qualidade das fronteiras não dominadas, com o objetivo de comparar o NSGA-II memético, será utilizada a métrica de qualidade conhecida como *S-Metric* [110]. Essa métrica calcula o hipervolume de uma região multidimensional qualquer englobada por um conjunto de pontos e por um

ponto fixo de referência. A Figura 4.1 ilustra em duas dimensões a área usada para calcular o valor da S-Metric, para os pontos v_1, v_2, v_3 e usando o ponto de referência z .

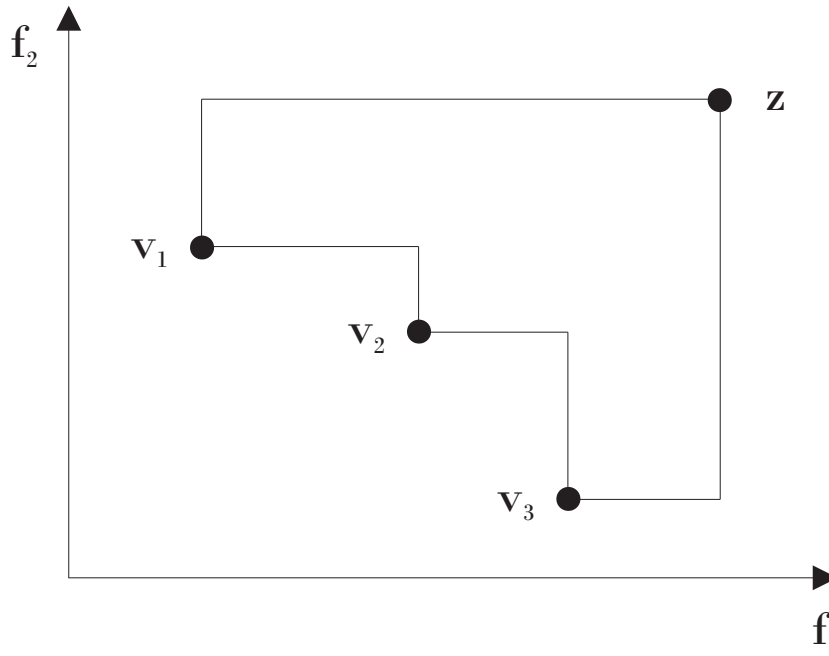


Figura 4.1: Região delimitada para o cálculo da S-Metric em duas dimensões.

No presente caso, será calculado o valor da S-Metric para as fronteiras não dominadas das versões do NSGA-II com e sem operador de busca local. Assim, será calculado um valor aproximado de área gerado pelas soluções não dominadas e um ponto de referência no espaço de objetivos. Deve-se notar que cada coordenada do ponto de referência deve ser um limite superior para cada função objetivo.

Uma fronteira não dominada com maior valor de S-Metric é melhor do que uma outra com valor menor, pois a primeira pode possuir mais pontos não dominados e ser mais espaçada no espaço de objetivos.

4.4 Simulação Monte Carlo

Após executado o processo de otimização utilizando o modelo endêmico SIR, o conjunto final das soluções não dominadas será submetida a um processo de simulação pelo MBI, conforme descrito pela Seção 2.3. Para cada política de campanhas de vacinação obtida no processo de otimização serão realizadas 100 simulações. Obviamente, será executada uma *Simulação Monte Carlo*

para analisar o comportamento da dinâmica da epidemia na presença de ações de controle nos tempos mais próximos daqueles descritos pela solução. O percentual de vacinas a serem aplicadas na população suscetível, em cada campanha, é interpretado neste contexto como a probabilidade de cada indivíduo suscetível vacinar.

Em cada simulação serão recalculadas as funções objetivo da Equação 4.2 e armazenadas como quantis para uma análise posterior.

É contabilizada também a quantidade de vezes em que o número de infectados atingiu o valor zero, ou seja, em que ocorreu a erradicação da doença. Isso é executado para o cálculo da probabilidade de erradicação, que é efetuado utilizando a Equação 2.13, em que no caso o número de experimentos é o número de simulações realizadas.

4.5 Dominância Estocástica

Em um problema de minimização, uma variável aleatória X_1 domina estocasticamente outra variável aleatória X_2 se e somente se $F_{X_1}(x) \geq F_{X_2}(x)$ para todo $x \in \mathfrak{R}$, em que F_X é a função de distribuição cumulativa de X [44].

Um ponto x^k define um quantil q_δ de X se e somente se $P(X \leq x^k) = \delta$ e $P(X > x^k) = 1 - \delta$.

Uma aproximação discreta da Dominância Estocástica de Primeira Ordem, simplesmente chamada aqui de dominância estocástica aproximada, que é requerida para o processamento computacional, pode ser construída empregando um número finito de valores da função de distribuição de cumulativa. Isso leva a uma metodologia de intervalo, como apresentada pela Definição 4.1.

Definição 4.1 (Dominância Estocástica Aproximada). Seja X_1 uma variável aleatória e seja $x_1^1 = q_{\frac{1}{n+1}}, \dots, x_1^i = q_{\frac{i}{n+1}}, \dots, x_1^n = q_{\frac{n}{n+1}}$ os n quantis igualmente espaçados de X_1 . Seja também X_2 outra variável aleatória e seja $x_2^1 = q_{\frac{1}{n+1}}, \dots, x_2^i = q_{\frac{i}{n+1}}, \dots, x_2^n = q_{\frac{n}{n+1}}$ os mesmos quantis para X_2 . É dito que X_1 domina estocasticamente X_2 se e somente se $x_1^i \leq x_2^i$ para todo $i \in \{1, \dots, n\}$ e $x_1^i < x_2^i$ para algum $i \in \{1, \dots, n\}$.

Quando n tende para infinito a dominância estocástica aproximada se torna equivalente a Dominância Estocástica de Primeira Ordem.

A dominância estocástica aproximada será usada nesse trabalho, após a simulação do MBI, para validar as soluções não dominadas provenientes da otimização. Será determinado um novo conjunto de soluções não dominadas utilizando-se os quantis 0,25, 0,50 (mediana) e 0,75 das funções objetivo reavaliadas durante a simulação. Dessa maneira, soluções que são sensíveis

a variação de parâmetros devem se tornar dominadas. Soluções robustas devem ficar no conjunto de soluções não dominadas estocasticamente.

Capítulo 5

Resultados

5.1 Otimização utilizando o modelo SIR

Após efetuar as 30 execuções do NSGA-II memético, conforme apresentado no Algoritmo 7 e configurado conforme mostrado na Seção 4.2, encontrou-se a fronteira apresentada na Figura 5.1. Os eixos representam a integral dos indivíduos infectados durante o tempo de controle, F_1 , normalizado pelo total de pontos retornados pelo método de Runge-Kutta [34], e o custo com o controle, F_2 . Cada solução é representada por um sinal de mais e a solução utópica é representada por um losango. O conjunto não dominado final possui 84 soluções.

Para se ter uma ideia da qualidade do NSGA-II com operador de busca local, executou-se também o NSGA-II canônico 30 vezes, conforme especificações apresentadas no Capítulo 4. Fez-se a união das soluções finais e selecionou-se o conjunto não dominado final. Ou seja, fez-se o mesmo procedimento de busca das soluções não dominadas do NSGA-II memético.

Unido-se as soluções finais dos dois algoritmos (*qap* e *can*), como pode-se observar na Figura 5.2, e aplicando uma ordenação pela não dominância observa-se o comportamento apresentado na Figura 5.3. As soluções representadas pelo símbolo mais são provenientes do NSGA-II com busca local (*qap*) e as soluções representadas por triângulos são provenientes do NSGA-II canônico (*can*). Observa-se na fronteira final 87 soluções, sendo que 77 são provenientes do NSGA-II memético e apenas 10 do NSGA-II canônico. Ou seja, 88,5% das soluções não dominadas finais são provenientes do NSGA-II com operador de busca local.

O fato das soluções provenientes do NSGA-II com operador de busca local dominarem a grande maioria das soluções do NSGA-II canônico já é um bom indicativo da qualidade do otimizador construído para este trabalho.

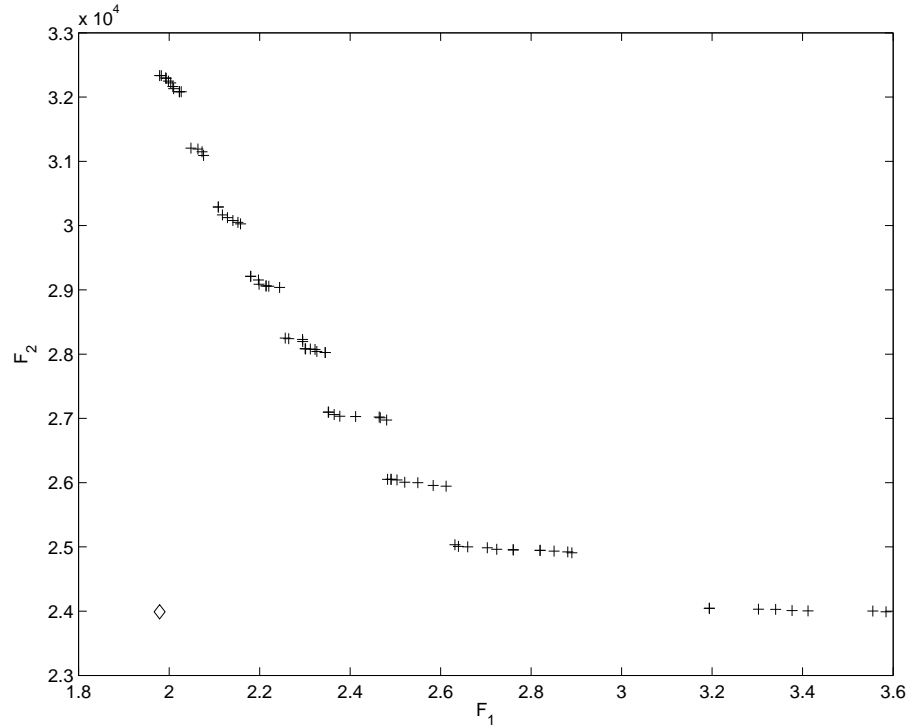


Figura 5.1: Fronteira não dominada com soluções para o Problema 4.2 retornada pelo NSGA-II com operador de busca local em 30 execuções. F_1 representa a integral da população infectada durante o controle e F_2 representa o custo total com as campanhas de vacinação.

Os valores da S-Metric das fronteiras não dominadas pelo número de funções objetivo avaliadas no final das execuções também demonstram a superioridade do NSGA-II memético, conforme apresentado pela Figura 5.4. Para se ter uma ideia, a diferença entre o quantil 0,75 da versão *gap* e da versão *can* é igual a $4,2270 \cdot 10^4$, que é uma diferença de área significativa.

Com o propósito de ilustrar a evolução da qualidade das soluções, não comparando os dois métodos¹, as Figuras 5.5 e 5.6 apresentam os valores da S-Metric das soluções não dominadas pela geração através de box-plots e gráfico das medianas para os dois otimizadores. Todos estes gráficos possuem escala igual no eixo vertical. Pode-se observar que após a busca local ser executada ocorre uma melhoria considerável na qualidade do conjunto não dominado. Uma observação gráfica parece sugerir que há uma menor variabilidade nos valores da S-Metric no otimizador híbrido nas gerações finais.

¹Uma vez que o número de avaliações de funções objetivo por geração é diferente para os dois métodos.

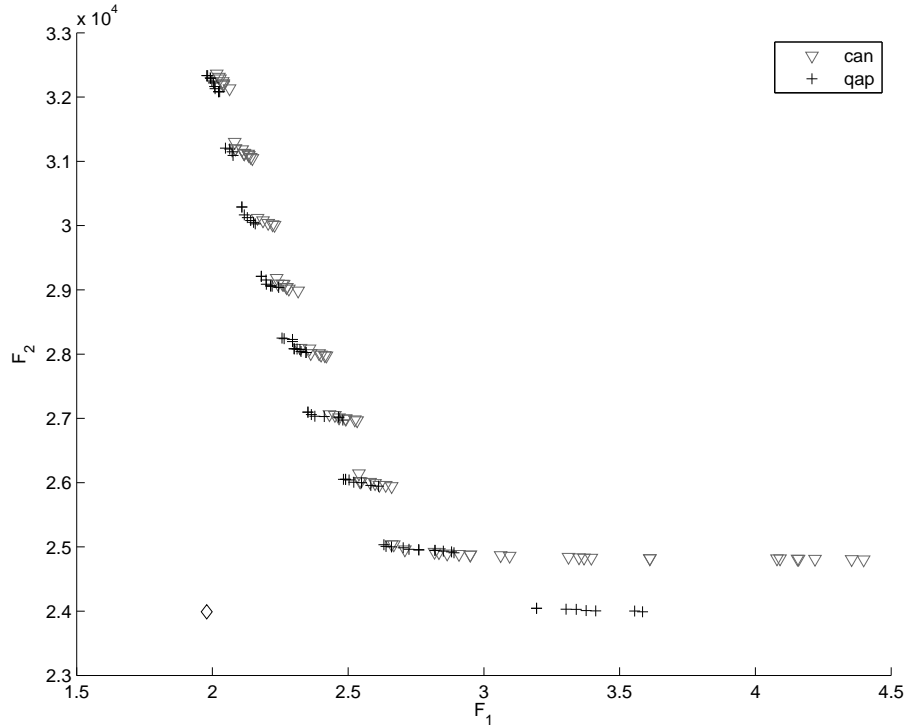


Figura 5.2: União das soluções finais não dominadas do NSGA-II com busca local e do NSGA-II canônico para o Problema 4.2. F_1 representa a integral dos infectados durante o controle e F_2 representa o custo total com o controle. As soluções representadas pelo símbolo mais são provenientes do NSGA-II com busca local e as soluções representadas por triângulos são provenientes do NSGA-II canônico.

Embora mais eficiente, o NSGA-II memético possui um custo computacional maior do que o canônico devido a busca local. Essa afirmação pode ser observada na Figura 5.7, que apresenta o box-plot do tempo, em segundos, de cada otimizador em todas as execuções. A diferença entre o quantil 0,75 do *gap* e do *can* é de apenas 110,0991 segundos.

O conjunto de soluções não dominadas final, apresentado na Figura 5.3, obedece as restrições do Problema 4.2. O número de campanhas variou entre 20 e 28 estágios. O custo de implementação das campanhas de vacinação variou entre $2,3991 \cdot 10^4$ e $3,2335 \cdot 10^4$ unidades monetárias. A integral da população infectada ficou entre 1,9790 e 3,5844. Para se ter uma ideia das distribuições dessas variáveis no conjunto de soluções não dominadas final são apresentados histogramas nas Figuras 5.8, 5.9 e 5.10. A Figura 5.11 apresenta o custo de implementação de acordo com o número campanhas

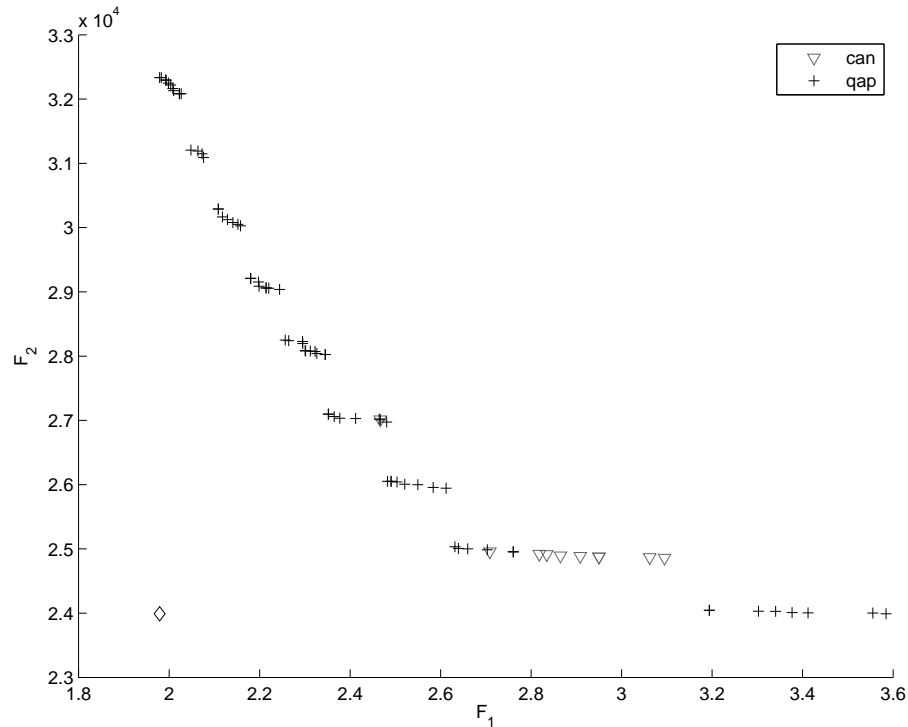


Figura 5.3: Fronteira não dominada final com soluções do NSGA-II com busca local e do NSGA-II canônico para o Problema 4.2. F_1 representa a integral dos infectados durante o controle e F_2 representa o custo total com o controle. As soluções representadas pelo símbolo mais são provenientes do NSGA-II com busca local e as soluções representadas por triângulos são provenientes do NSGA-II canônico.

que o conjunto de soluções não dominadas oferece.

Para ilustrar algumas soluções sobre o modelo endêmico determinístico SIR são apresentadas as soluções de custo mais barato, de custo mediano e a de custo mais caro da fronteira não dominada, apresentada na Figura 5.3. Essas soluções são ilustradas, respectivamente, nas Figuras 5.12, 5.13 e 5.14. É apresentado em cada figura o comportamento da população de suscetíveis, infectados e recuperados no tempo, o número vacinas que devem ser aplicadas em cada campanha nos suscetíveis e o percentual de suscetíveis que devem ser vacinados em cada campanha.

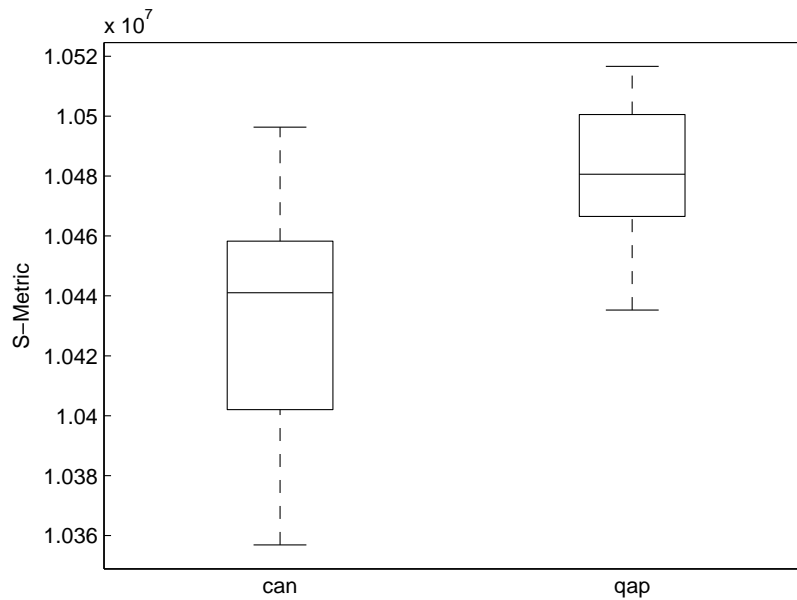
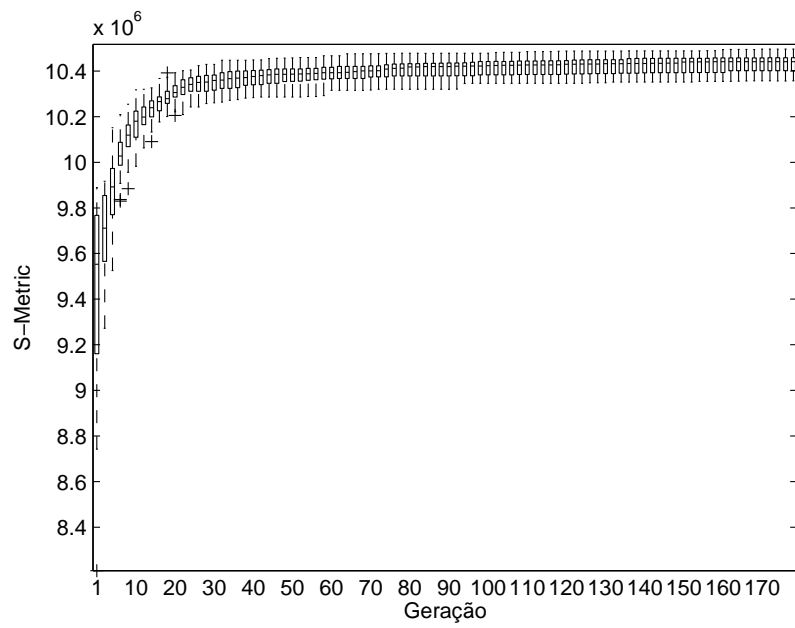


Figura 5.4: Box-plot dos valores da S-Metric para as soluções não dominadas após 7120 avaliações de funções objetivo para o NSGA-II canônico (*can*) e 7100 avaliações para o NSGA-II com operador de busca local (*qap*) em 30 execuções.

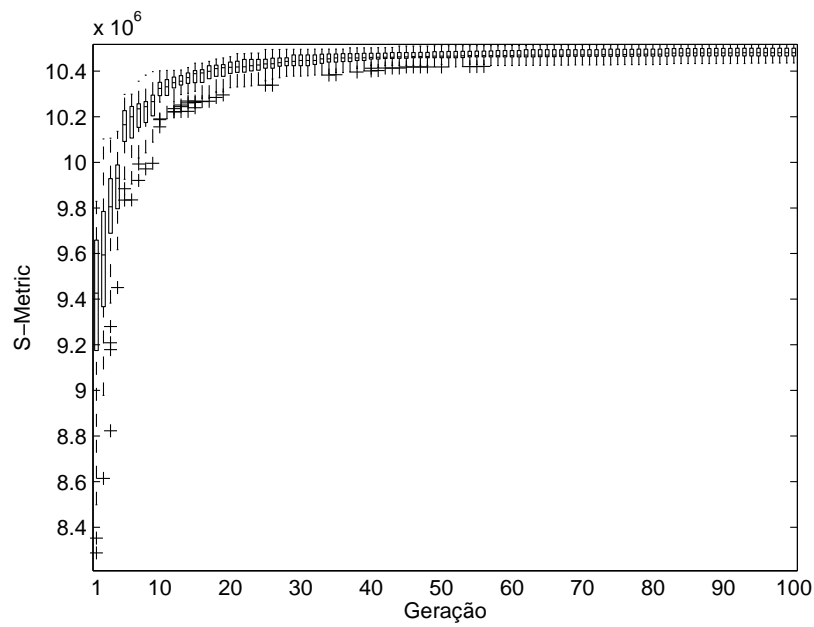
5.2 Validação utilizando o modelo MBI

Cada uma das 87 soluções não dominadas provenientes da otimização com o NSGA-II usando o modelo SIR, apresentada na Figura 5.3, foi simulada 1000 vezes no MBI. Em cada execução do MBI foram reavaliadas as duas funções objetivo F_1 e F_2 do Problema 4.2. Foi contabilizado também quantas vezes em que o número de infectados se anulou, para o cálculo das probabilidades de erradicação. Para cada solução, durante processo de Monte Carlo, foram selecionados os quantis 0,25, 0,50 e 0,75 de cada função objetivo reavaliada. Com esses valores foi aplicada a ordenação pela não dominância estocástica, obtendo-se uma nova fronteira de soluções não dominadas, que são robustas em relação à variação de parâmetros.

A nova fronteira é apresentada na Figura 5.15. Os eixos são os mesmos apresentados pela Figura 5.1. O novo conjunto Pareto considera um intervalo de confiança, determinado pelos quantis extraídos, em que os quantis 0,25 são os menores valores, a mediana é ponto de cruzamento dos eixos e os quantis 0,75 são os maiores valores. Com isso, essa informação de confiabilidade pode



(a)



(b)

Figura 5.5: Box-plot da S-Metric pela geração das soluções não dominadas para o otimizador canônico (*can*) (a) e para o otimizador memético (*qap*) (b).

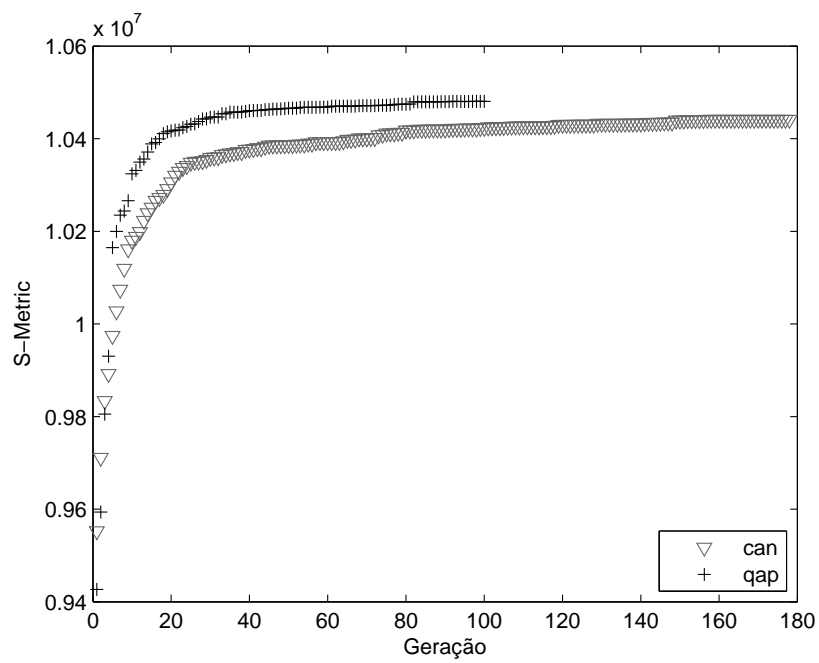


Figura 5.6: Mediana dos valores da S-Metric pela geração para o otimizador canônico (*can*) (a) e para o otimizador memético (*gap*).

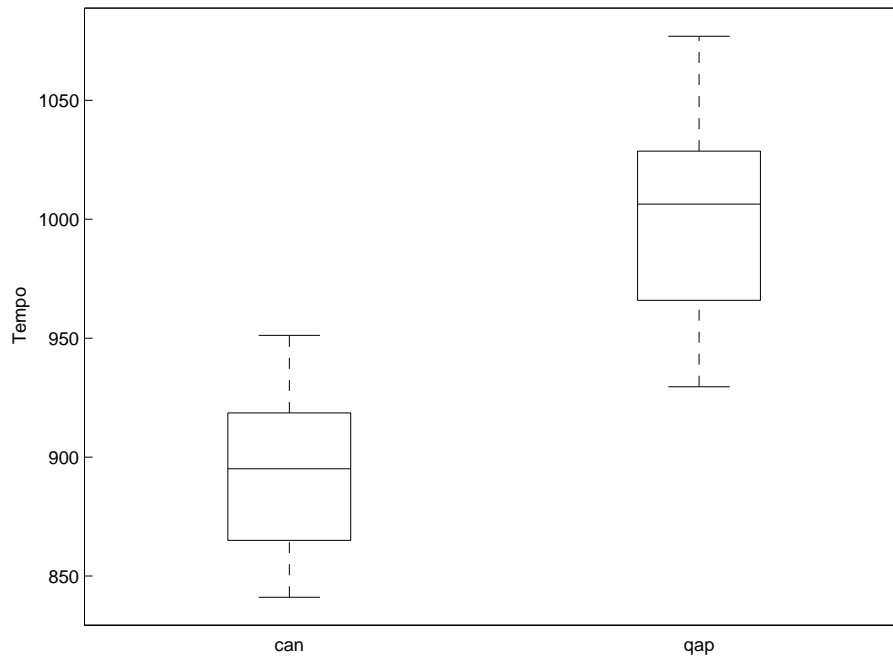


Figura 5.7: Box-plot do tempo, em segundos, das 30 execuções do NSGA-II canônico (*can*) para 178 gerações e do NSGA-II com operador de busca local (*qap*) para 100 gerações.

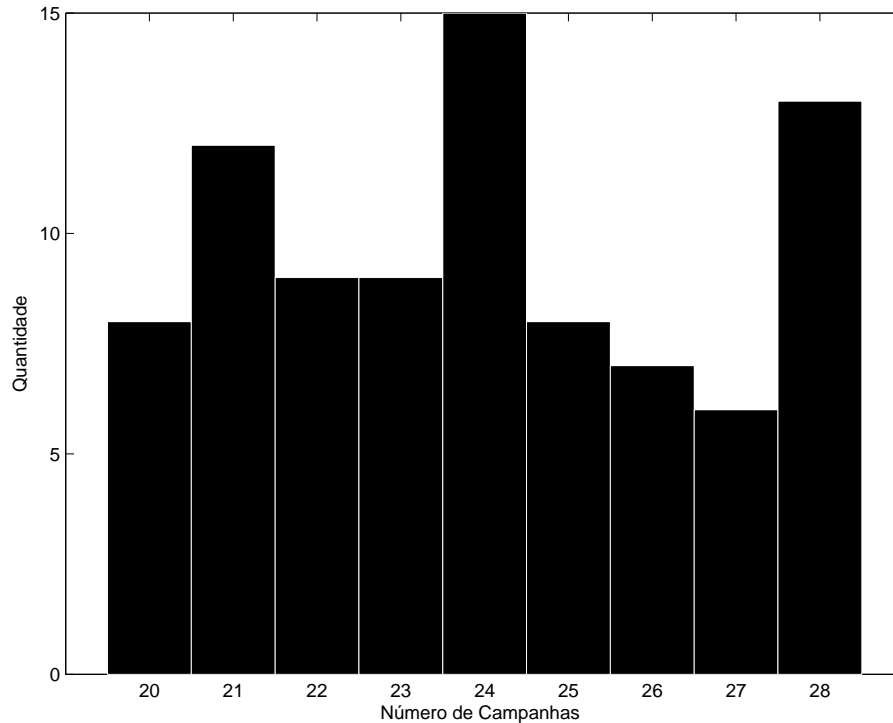


Figura 5.8: Histograma da quantidade de soluções pelo número de campanhas de vacinação das soluções não dominadas provenientes da otimização utilizando o modelo epidemiológico determinístico.

ser usada como fator de escolha para o decisor. Esse tipo de informação não pode ser extraída utilizando o modelo SIR. Uma visão simplificada dessa fronteira não dominada estocasticamente é apresentada na Figura 5.16, em que são mostrados somente os valores das medianas das funções objetivo.

Das 87 soluções, provenientes da otimização utilizando o modelo SIR endêmico, ficaram 31 soluções na fronteira não dominada estocasticamente. Para cada uma dessas soluções não dominadas, os quantis das funções objetivo, conforme apresentados na Figura 5.15, e as probabilidades da doença ser erradicada em um tempo menor ou igual a $T_{ctrl} = 360$ unidades de tempo são apresentadas na Tabela 5.1. Observa-se que pelo comportamento rígido e de difícil controle da epidemia, obtiveram-se resultados satisfatórios para a probabilidade de erradicá-la. A informação sobre probabilidade de erradicação não pode ser extraída através do modelo determinístico SIR.

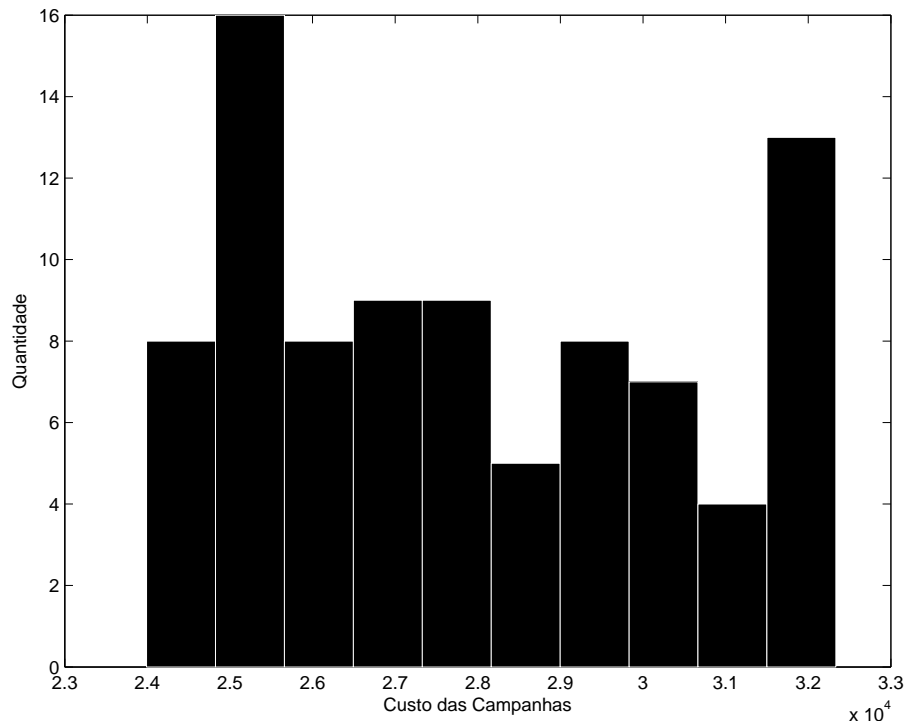


Figura 5.9: Histograma da quantidade de soluções pelo custo de implementação das campanhas de vacinação das soluções não dominadas provenientes da otimização utilizando o modelo epidemiológico determinístico.

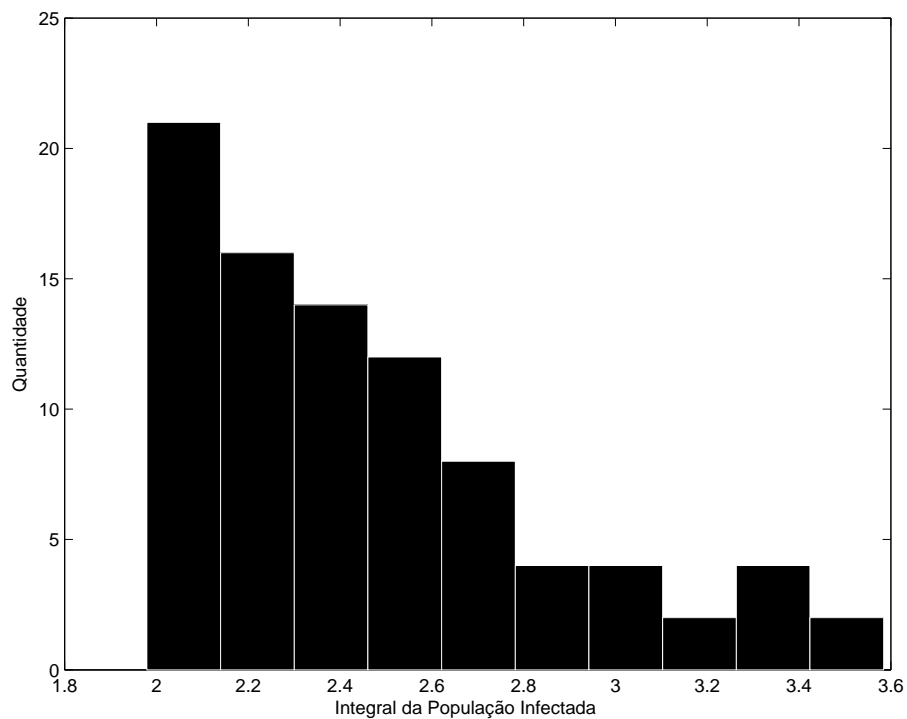


Figura 5.10: Histograma da quantidade de soluções pela integral de infectados das soluções não dominadas provenientes da otimização utilizando o modelo epidemiológico determinístico.

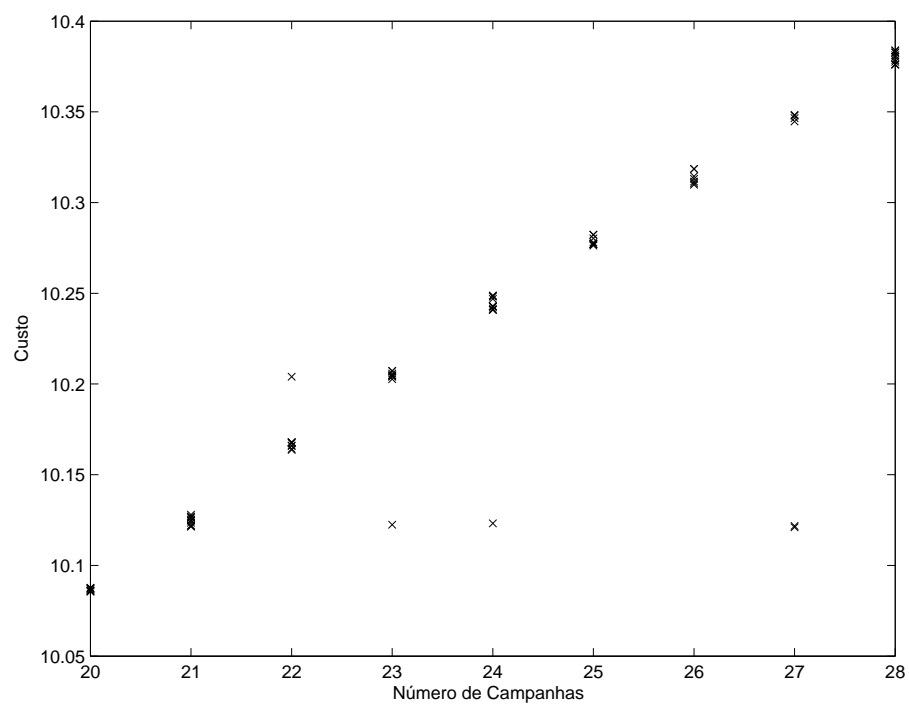
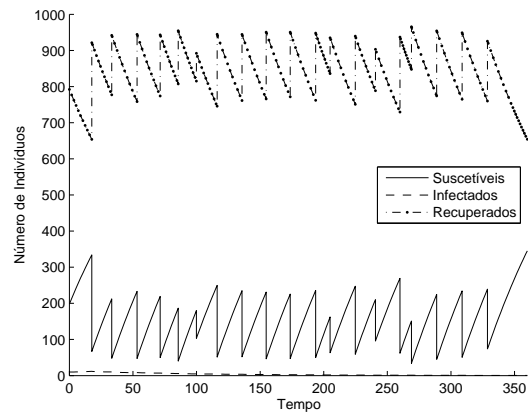
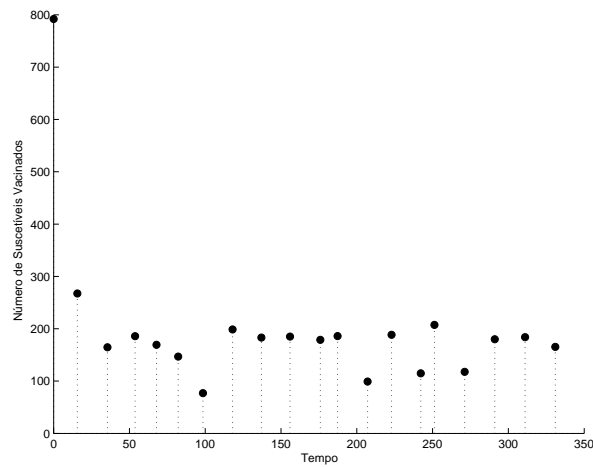


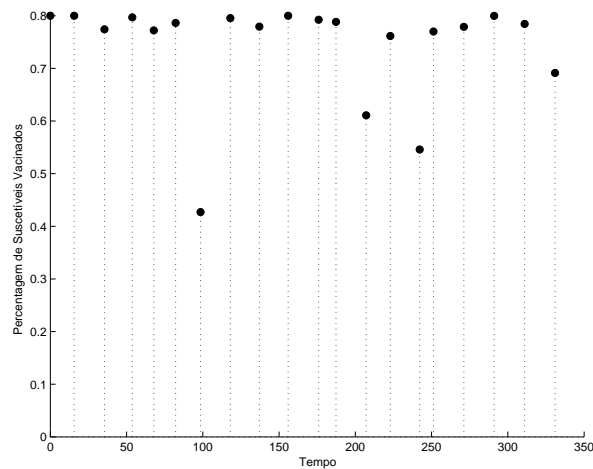
Figura 5.11: Custo de implementação das políticas de vacinação pelo número de campanhas das soluções não dominadas.



(a)

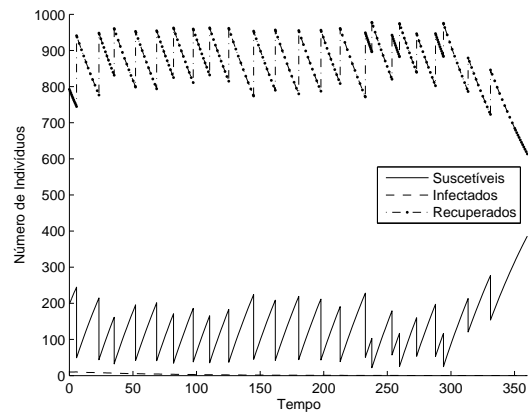


(b)

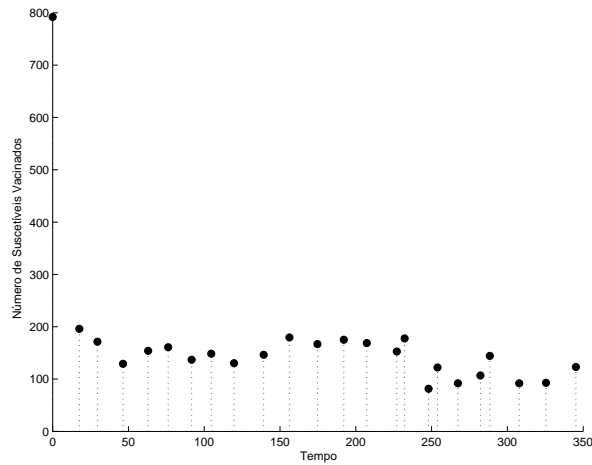


(c)

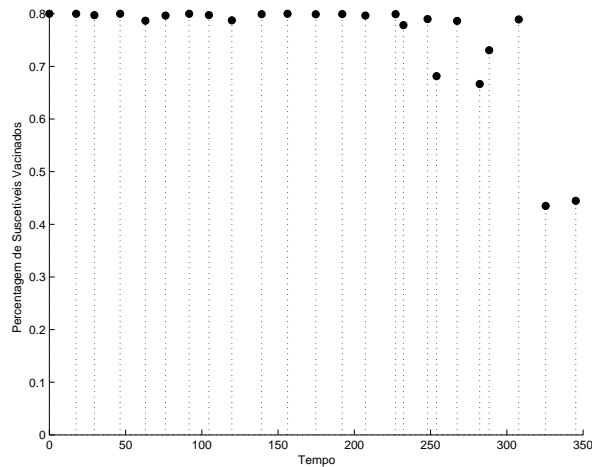
Figura 5.12: Comportamento da solução mais barata da Figura 5.3 com 20 campanhas, custo de implementação $2,3991 \cdot 10^4$ e integral de infectados 3,5844. O gráfico (a) apresenta o comportamento do sistema no tempo, o (b) mostra o número de vacinas em cada campanha e o (c) apresenta o percentual de suscetíveis que devem ser vacinados.



(a)

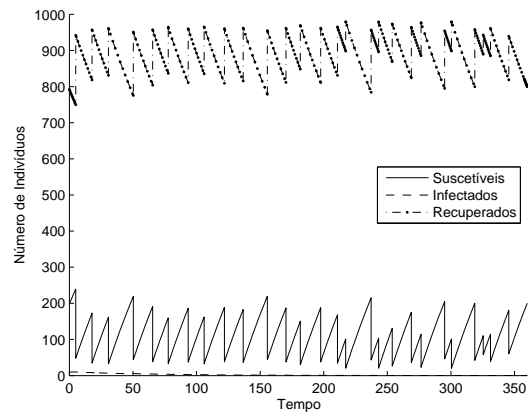


(b)

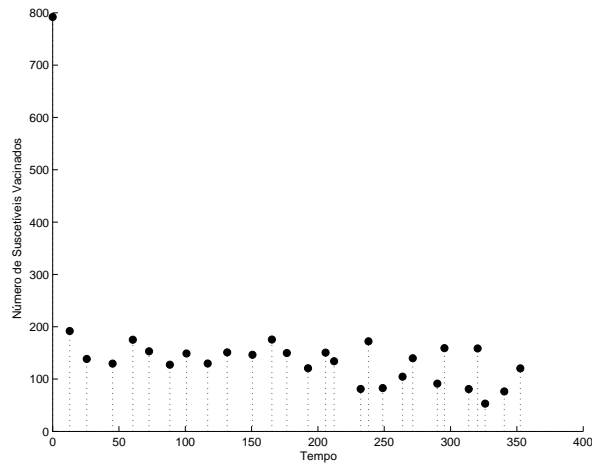


(c)

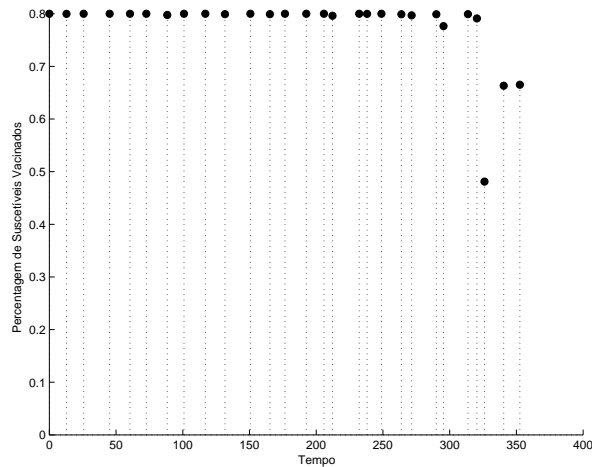
Figura 5.13: Comportamento da solução de custo mediano da Figura 5.3 com 24 campanhas, custo de implementação $2,8041 \cdot 10^4$ e integral de infectados $2,3265$. O gráfico (a) apresenta o comportamento do sistema no tempo, o (b) mostra o número de vacinas em cada campanha e o (c) apresenta o percentual de suscetíveis que devem ser vacinados.



(a)



(b)



(c)

Figura 5.14: Comportamento da solução mais cara da Figura 5.3 com 28 campanhas, custo de implementação $3,2335 \cdot 10^4$ e integral de infectados 1,9790. O gráfico (a) apresenta o comportamento do sistema no tempo, o (b) mostra o número de vacinas em cada campanha e o (c) apresenta o percentual de suscetíveis que devem ser vacinados.

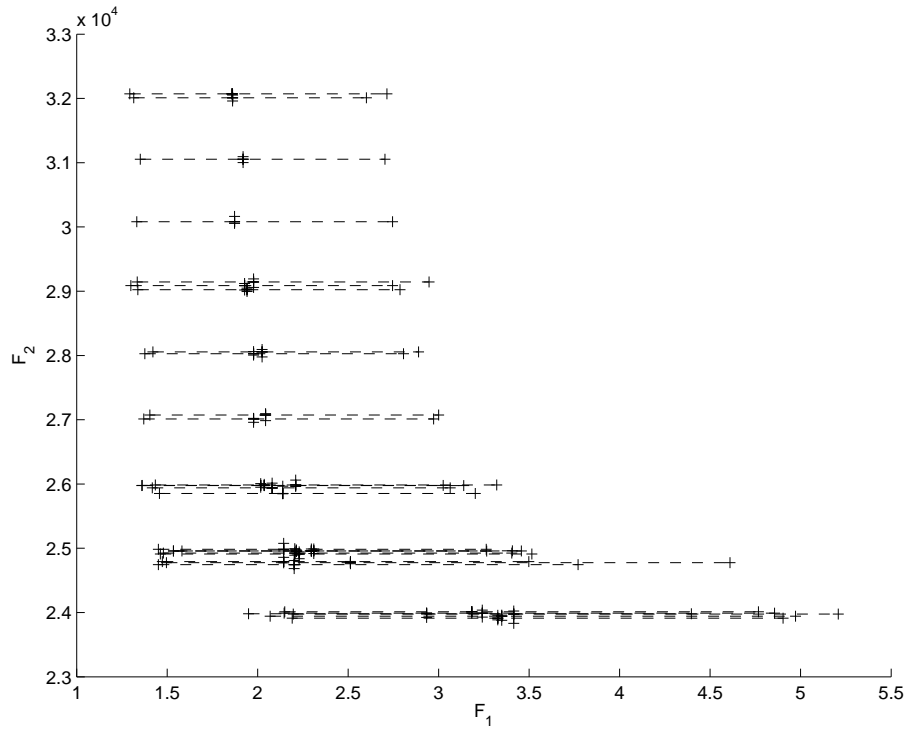


Figura 5.15: Fronteira Pareto representando as soluções não dominadas da Figura 5.1 sob o ponto de vista estocástico. As soluções foram validadas pela dominância estocástica após um procedimento Monte Carlo do MBI.

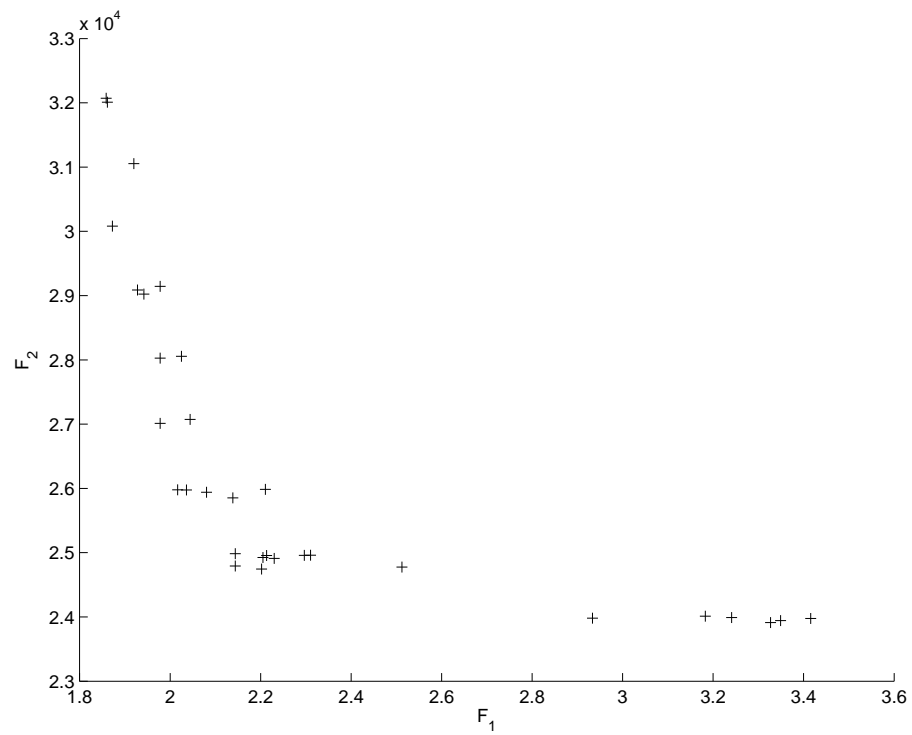


Figura 5.16: Mediana das soluções apresentadas pela Figura 5.15.

Solução	Número de Campanhas	F_1				F_2				$P(I = 0 t \leq 360)$
		$q_{0,25}$	$q_{0,50}$	$q_{0,75}$	$ q_{0,75} - q_{0,25} $	$q_{0,25}$	$q_{0,50}$	$q_{0,75}$	$ q_{0,75} - q_{0,25} $	
1	21	1,4958	2,5125	4,6094	2,0970	24795	24759	24777	18	0,764
2	21	1,4515	2,2022	3,7701	1,5679	24745	24808	24683	62	0,801
3	20	2,1911	3,3269	4,9030	1,5762	23912	23957	23885	27	0,842
4	20	2,1967	3,4155	5,2078	1,7922	23975	24023	23832	143	0,854
5	21	1,4737	2,1440	3,4986	1,3546	24773	24859	24793	20	0,855
6	20	2,1496	3,1828	4,7673	1,5845	24015	24011	23989	26	0,855
7	21	1,4626	2,2299	3,5152	1,2853	24945	24911	24838	107	0,865
8	20	2,0693	3,3490	4,9723	1,6233	23959	23944	23878	81	0,869
9	22	1,3601	2,0360	3,1385	1,1025	25976	25995	25976	0	0,875
10	20	2,1468	3,2410	4,8560	1,6150	24038	23990	23929	109	0,882
11	20	1,9501	2,9335	4,3961	1,4626	23982	23923	23996	14	0,889
12	21	1,4792	2,2050	3,4044	1,1994	24993	24925	24892	101	0,893
13	21	1,5346	2,3102	3,4072	1,0970	24989	24961	24919	70	0,914
14	22	1,4571	2,1385	3,2022	1,0637	25849	25853	25967	118	0,922
15	21	1,4515	2,1440	3,2632	1,1191	25077	24958	24983	94	0,923
16	21	1,5346	2,2133	3,2659	1,0526	24951	24955	24975	24	0,926
17	21	1,5817	2,2964	3,4571	1,1607	24990	24955	24959	31	0,930
18	22	1,4349	2,2105	3,3213	1,1108	25985	26061	25962	23	0,931
19	23	1,3712	1,9778	2,9723	0,9945	27012	27022	26960	52	0,935
20	22	1,4183	2,0803	3,0637	0,9834	26013	25940	25935	78	0,936
21	22	1,3629	2,0166	3,0249	1,0083	25949	25979	26006	57	0,943
22	24	1,4211	2,0249	2,8892	0,8643	28094	27978	28056	38	0,954
23	27	1,3518	1,9197	2,7036	0,7839	31001	31095	31053	52	0,958
24	23	1,4044	2,0443	3,0000	0,9557	27095	26987	27074	21	0,958
25	24	1,3767	1,9778	2,8061	0,8283	28060	28028	28007	53	0,961
26	25	1,3352	1,9778	2,9474	0,9695	29060	29191	29145	85	0,962
27	25	1,2992	1,9280	2,7452	0,8172	29121	29087	29018	103	0,962
28	25	1,3380	1,9418	2,7867	0,8449	29048	28996	29023	25	0,963
29	26	1,3324	1,8726	2,7452	0,8726	30054	30080	30164	110	0,964
30	28	1,2936	1,8587	2,7147	0,8560	32069	32073	32071	2	0,969
31	28	1,3158	1,8615	2,6011	0,7396	32052	31960	32009	43	0,976

Tabela 5.1: Dados das soluções não dominadas estocasticamente.

As Figuras 5.17 e 5.18 apresentam o comportamento estocástico das populações de suscetíveis, infectados e recuperados no tempo das soluções extraídas do conjunto não dominado final que apresentam o menor e o maior valor de probabilidade de erradicação da doença. As curvas em preto representam o comportamento mediano das soluções e as curvas em cinza representam o comportamento dos quantis 0,25 e 0,75 com relação a função objetivo F_1 obtidas nas simulações. Note que estes intervalos possuem importantes informações sobre a previsão do comportamento da doença de acordo com a seleção de uma política de campanha de vacinação a ser adotada. Novamente, esse tipo de informação não pode ser extraída do modelo determinístico SIR.

É observado em algumas soluções que o número de suscetíveis começa a aumentar no final do horizonte de tempo. A maneira como o modelo de otimização foi formulado permite que esse comportamento apareça. Para que esse fato não ocorra, uma simples inserção de uma restrição sobre o número de suscetíveis em um intervalo de tempo no final do controle removeria esse tipo de comportamento. Uma outra alternativa seria manter o percentual de vacinas constante e os intervalos de tempo entre campanhas serem equidistantes ao final do horizonte de tempo, forçando um controle permanente sob um regime estável.

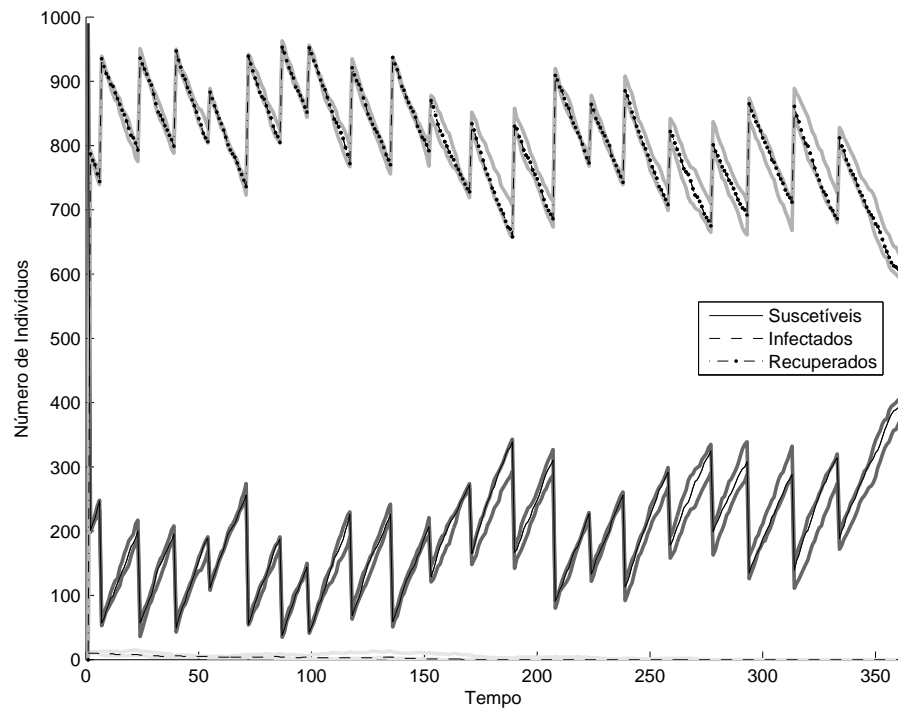


Figura 5.17: Comportamento da solução que apresenta probabilidade de erradicação da doença igual a 0,764. O intervalo de confiança é dado pela linha cinza.

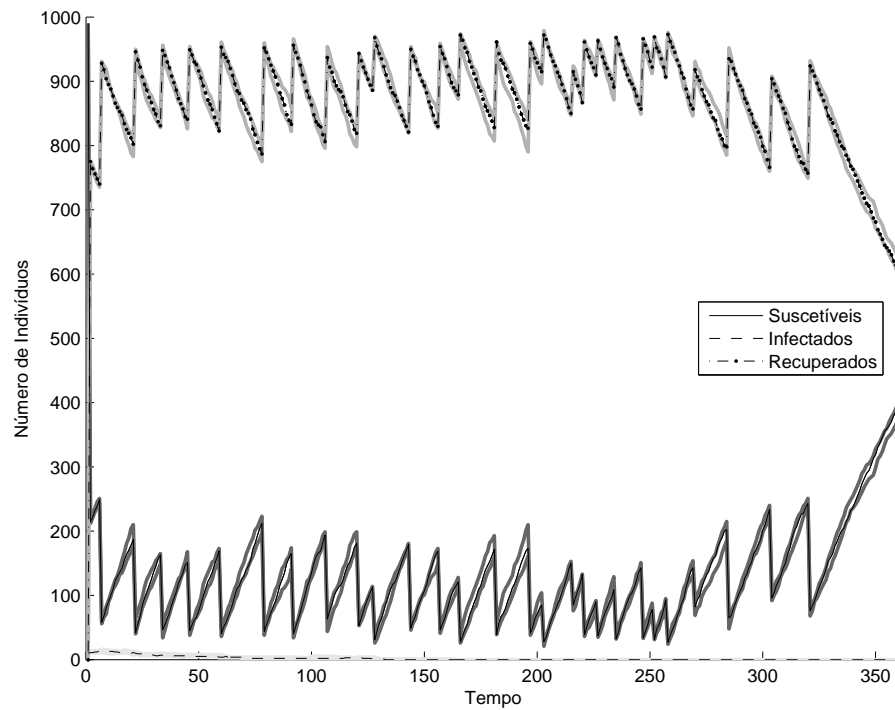


Figura 5.18: Comportamento da solução que apresenta probabilidade de erradicação da doença igual a 0,976. O intervalo de confiança é dado pela linha cinza.

Capítulo 6

Conclusões

O presente trabalho elaborou e apresentou uma metodologia multiobjetivo para determinar, através de otimização e validação, um conjunto de soluções não dominadas que contém políticas de controle que minimizam a integral da população de infectados e o custo com a implementação de campanhas de vacinação em um horizonte de tempo finito fornecido previamente.

O modelo epidemiológico usado para governar o sistema durante a otimização é o Suscetível-Infectado-Recuperado (SIR). Esse modelo baseado em um sistema de equações diferenciais determina o comportamento médio das interações entre as populações de indivíduos suscetíveis, infectados e recuperados. Por ser um modelo determinístico, o mesmo não é capaz de fornecer informações como probabilidades de erradicação da doença e intervalos de confiança.

As soluções retornadas pelo otimizador são políticas de controle que apresentam o número de campanhas de vacinação necessários, a quantidade de vacinas em cada campanha para ser aplicada na população suscetível e os intervalos de tempo entre cada campanha. As quantidades de vacinas usadas em cada campanha podem variar e, da mesma forma, os tempos entre as consecutivas campanhas não são necessariamente os mesmos. O otimizador determinou em malha aberta as melhores configurações de valores para o controle da epidemia.

O motor de otimização foi o NSGA-II, incorporado com uma busca local baseada na otimização de somas ponderadas das aproximações quadráticas convexas das funções objetivo sobre uma vizinhança de pontos. Essa busca local melhorou a qualidade das soluções encontradas.

As soluções retornadas pelo otimizador, pertencentes ao conjunto não dominado que é uma aproximação do Pareto-Ótimo, foram validadas em um Modelo Baseado em Indivíduos (MBI) através de um procedimento Monte Carlo. As informações sobre a probabilidade de erradicação da doença e

os intervalos de confiança foram extraídas tanto para o comportamento das populações no tempo quanto para as variações das funções objetivo.

As principais contribuições do presente trabalho foram:

- A aplicação da metodologia multiobjetivo para determinar, usando o modelo SIR endêmico e algoritmos evolucionários multiobjetivo, os melhores planejamentos de campanhas de vacinação para o controle de epidemias, através da minimização concomitante dos custos de implementação e da integral da população de infectados no horizonte de tempo em análise;
- O desenvolvimento e a aplicação de uma busca local baseada na otimização das aproximações quadráticas convexas das funções objetivo originais para melhorar a qualidade das soluções em otimização multiobjetivo; e
- A validação das soluções encontradas usando o MBI e a aplicação da dominância estocástica, a fim de eliminar soluções sensíveis à variação de parâmetros e extrair informações sobre probabilidade de erradicação e intervalos de confiança.

6.1 Discussão

O presente trabalho propõe uma solução para o controle de epidemias através de vacinação com uma abordagem mais dinâmica, eficaz e confiável.

Dinâmica pelo fato do otimizador decidir quando devem ocorrer as campanhas e quanto de vacinas devem ser investidos em cada uma delas. Essa forma é mais livre e se diferencia da maneira convencional como é feito pela maioria das políticas aplicadas na sociedade, que possui um período entre campanhas fixados.

Eficaz pelo fato de tratar o presente problema de maneira multiobjetivo, evitando desperdícios e buscando o maior proveito dos recursos disponíveis, e reduzindo ao máximo os transtornos causados por uma determinada enfermidade na sociedade.

E, por último, confiável por fornecer informações provenientes de simulações de modelos que podem se aproximar da realidade. Tais modelos são capazes de informar a chance de uma determinada doença erradicar, os intervalos de confiança para análise quantitativa de quanto é benéfica e robusta uma determinada solução, com relação a custos, possíveis prejuízos no decorrer do controle, além da previsão de como as populações podem se comportar durante o tempo.

Agraciado por finalizar esse texto, o autor espera que esse trabalho seja útil a sociedade e para a comunidade científica tanto para o aproveitamento de ideias quanto para a aplicação das mesmas. *E que seja feito na bonança.*

6.2 Trabalhos Futuros

Os seguintes itens são sugestões de continuidade do presente trabalho:

- Aplicação do modelo proposto em doenças presentes no mundo real;
- Estudo minucioso e análise de outras funções de custo para o modelo;
- Aplicação da metodologia empregando outros modelos epidemiológicos, além do modelo SIR;
- Aplicação da otimização empregando o modelo baseado em indivíduo diretamente;
- Aplicação de outras restrições ao modelo que podem ser comuns ou desejadas em diversas situações do mundo real;
- Aplicação de equações diferenciais estocásticas para modelar epidemias e comparar com a metodologia atual; e
- Emprego de outras buscas locais e/ou outros algoritmos de otimização para buscar melhorias na qualidade das políticas de controle.

6.3 Produção Bibliográfica

Durante a graduação até o momento desse honroso trabalho, o autor desenvolveu os seguintes trabalhos:

Artigos Publicados em Periódicos

- Cardoso, R. T. N.; Cruz, A. R.; Wanner, E. F.; Takahashi, R. H. C.; “Multiobjective Evolutionary Optimization of Biological Pest Control with Impulsive Dynamics in Soybean Crops”. *Bulletin of Mathematical Biology*, 2009.

Artigos Aceitos em Conferências

- Cruz, A. R.; Cardoso, R. T. N.; Takahashi, R. H. C.; “Multiobjective Dynamic Optimization of Vaccination Campaigns using Convex Quadratic Approximation Local Search”. *Proceedings of the Sixth International Conference on Evolutionary Multi-Criterion Optimization (EMO 2011)*, Ouro Preto, 2011.

Artigos Publicados em Conferências

- Cruz, A. R.; Magela, O. N.; R. H. C. Takahashi; “Uma Otimização Multiobjetivo para a Definição de uma Linha de Transmissão Elétrica via NSGA-II”. *Simpósio de Mecânica Computacional - SIMMEC*, São João del Rei, 2010.
- Cruz, A. R.; Cardoso, R. T. N.; R. H. C. Takahashi; “Multiobjective Design of Vaccination Campaigns with a Stochastic Validation”. *Simpósio Brasileiro de Pesquisa Operacional*, Porto Seguro, 2009.
- Cruz, A. R.; Cardoso, R. T. N.; R. H. C. Takahashi; “Multi-objective Design with a Stochastic Validation of Vaccination Campaigns”. *IFAC Workshop on Control Applications of Optimisation*, Jyväskylä, 2009.
- Cruz, A. R.; Cardoso, R. T. N.; Takahashi, R. H. C.; “Uma Abordagem Multiobjetivo para o Planejamento de Campanhas de Vacinação via NSGA-II”. *Simpósio Brasileiro de Pesquisa Operacional*, João Pessoa, 2008.
- Cruz, A. R.; Cardoso, R. T. N.; Wanner, E. F.; Takahashi, R. H. C.; “A Multiobjective Non-Linear Dynamic Programming Approach for Optimal Biological Control in Soy Farming via NSGA-II”. *IEEE Congress on Evolutionary Computation*, Singapore, 2007.
- Cruz, A. R.; Cardoso, R. T. N.; Takahashi, R. H. C.; “Uma Abordagem Multiobjetivo para o Problema de Controle Biológico através da Programação Dinâmica Não-Linear via NSGA-II”. *Simpósio Brasileiro de Pesquisa Operacional*, Fortaleza, 2007.
- Cruz, A. R.; Cardoso, R. T. N.; Takahashi, R. H. C.; “Uma Abordagem via Programação Dinâmica Não-Linear para o Problema do Controle Biológico de Pragas numa Lavoura”. *Simpósio Brasileiro de Pesquisa Operacional*, Goiânia, 2006.

- Silva, V. L.; Cruz, A. R.; Carrano, E. G.; Guimarães, F. G.; Takahashi, R. H. C.; “On nonlinear fitness functions for ranking-based selection”. *IEEE Congress on Evolutionary Computation*, Vancouver, 2006.

Resumos Publicados em Conferências

- Cruz, A. R.; Cardoso, R. T. N.; Takahashi, R. H. C.; “Controle Biológico de Pragas em Lavouras de Soja via Otimização Dinâmica Impulsiva Multiobjetivo”. *Congresso Nacional de Matemática Aplicada e Computacional*, 2007, Florianópolis.

Bibliografia

- [1] J. Last, “A dictionary of epidemiology,” *International Journal of Epidemiology*, vol. 15, no. 2, p. 277, 1986.
- [2] R. Bonita, R. Beaglehole, and T. Kjellström, *Basic epidemiology*. WHO, 2006.
- [3] O. Diekmann and J. Heesterbeek, *Mathematical epidemiology of infectious diseases: Model building, analysis, and interpretation*. Wiley, 2000.
- [4] D. Clancy, “Optimal intervention for epidemic models with general infection and removal rate functions,” *Journal of mathematical biology*, vol. 39, no. 4, pp. 309–331, 1999.
- [5] H. Hethcote, “Three basic epidemiological models,” *Applied Mathematical Ecology*, pp. 119–144, 1989.
- [6] H. Hethcote, “The mathematics of infectious diseases,” *SIAM review*, vol. 42, no. 4, pp. 599–653, 2000.
- [7] M. Chubb and K. Jacobsen, “Mathematical modeling and the epidemiological research process,” *European journal of epidemiology*, vol. 25, no. 1, pp. 13–19, 2010.
- [8] R. M. Anderson and R. M. May, *Infectious Diseases of Humans: Dynamics and Control (Oxford Science Publications)*. Oxford University Press, USA, new ed ed., September 1992.
- [9] P. Coen, “How mathematical models have helped to improve understanding the epidemiology of infection,” *Early human development*, vol. 83, no. 3, pp. 141–148, 2007.
- [10] H. Hethcote, “Qualitative analyses of communicable disease models,” *Mathematical Biosciences*, vol. 28, no. 3-4, pp. 335–356, 1976.

- [11] S. Busenberg and P. Driessche, “Analysis of a disease transmission model in a population with varying size,” *Journal of mathematical biology*, vol. 28, no. 3, pp. 257–270, 1990.
- [12] Y. Kyrychko and K. Blyuss, “Global properties of a delayed SIR model with temporary immunity and nonlinear incidence rate,” *Nonlinear Analysis: Real World Applications*, vol. 6, no. 3, pp. 495–507, 2005.
- [13] A. Korobeinikov and P. Maini, “Non-linear incidence and stability of infectious disease models,” *Mathematical Medicine and Biology*, vol. 22, no. 2, p. 113, 2005.
- [14] A. Korobeinikov, “Lyapunov functions and global stability for SIR and SIRS epidemiological models with non-linear transmission,” *Bulletin of mathematical biology*, vol. 68, no. 3, pp. 615–626, 2006.
- [15] A. Korobeinikov, “Global properties of infectious disease models with nonlinear incidence,” *Bulletin of Mathematical Biology*, vol. 69, no. 6, pp. 1871–1886, 2007.
- [16] A. Korobeinikov, “Global properties of SIR and SEIR epidemic models with multiple parallel infectious stages,” *Bulletin of mathematical biology*, vol. 71, no. 1, pp. 75–83, 2009.
- [17] V. Grimm, “Ten years of individual-based modelling in ecology: what have we learned and what could we learn in the future?,” *Ecological modelling*, vol. 115, no. 2, pp. 129–148, 1999.
- [18] B. Breckling, U. Middelhoff, and H. Reuter, “Individual-based models as tools for ecological theory and application: Understanding the emergence of organisational properties in ecological systems,” *Ecological Modelling*, vol. 194, no. 1-3, pp. 102–113, 2006.
- [19] E. van Nes, E. Lammens, and M. Scheffer, “PISCATOR, an individual-based model to analyze the dynamics of lake fish communities,” *Ecological Modelling*, vol. 152, no. 2-3, pp. 261–278, 2002.
- [20] A. Breukers, D. Kettenis, M. Mourits, W. Van Der Werf, and A. Oude Lansink, “Individual-based models in the analysis of disease transmission in plant production chains: An application to potato brown rot,” *Agricultural Systems*, vol. 90, no. 1-3, pp. 112–131, 2006.

- [21] J. Fa, C. Sharples, D. Bell, and D. DeAngelis, “An individual-based model of rabbit viral haemorrhagic disease in European wild rabbits (*Oryctolagus cuniculus*),” *Ecological Modelling*, vol. 144, no. 2-3, pp. 121–138, 2001.
- [22] V. Laperrière, D. Badariotti, A. Banos, and J. Müller, “Structural validation of an individual-based model for plague epidemics simulation,” *Ecological Complexity*, vol. 6, no. 2, pp. 102–112, 2009.
- [23] M. Ajelli and S. Merler, “An individual-based model of hepatitis A transmission,” *Journal of Theoretical Biology*, vol. 259, no. 3, pp. 478–488, 2009.
- [24] F. Rakowski, M. Gruziel, L. Bieniasz-Krzywiec, and J. P. Radomski, “Influenza epidemic spread simulation for poland – a large scale, individual based model study,” *Physica A: Statistical Mechanics and its Applications*, vol. 389, no. 16, pp. 3149 – 3165, 2010.
- [25] W. Esposito and C. Floudas, “Deterministic global optimization in nonlinear optimal control problems,” *Journal of Global Optimization*, vol. 17, no. 1, pp. 97–126, 2000.
- [26] J. Müller, “Optimal vaccination patterns in age-structured populations,” *SIAM Journal on Applied Mathematics*, vol. 59, no. 1, pp. 222–241, 1998.
- [27] T. Yang, “Impulsive control,” *Automatic Control, IEEE Transactions on*, vol. 44, no. 5, pp. 1081–1083, 2002.
- [28] B. Shulgin, L. Stone, and Z. Agur, “Pulse vaccination strategy in the SIR epidemic model,” *Bulletin of Mathematical Biology*, vol. 60, no. 6, pp. 1123–1148, 1998.
- [29] L. Stone, B. Shulgin, and Z. Agur, “Theoretical examination of the pulse vaccination policy in the SIR epidemic model,” *Mathematical and Computer Modelling*, vol. 31, no. 4-5, pp. 207–216, 2000.
- [30] Z. Lu, X. Chi, and L. Chen, “The effect of constant and pulse vaccination on SIR epidemic model with horizontal and vertical transmission,” *Mathematical and computer modelling*, vol. 36, no. 9-10, pp. 1039–1057, 2002.
- [31] A. d’Onofrio, “Stability properties of pulse vaccination strategy in SEIR epidemic model,” *Mathematical biosciences*, vol. 179, no. 1, pp. 57–72, 2002.

- [32] A. d’Onofrio, “On pulse vaccination strategy in the SIR epidemic model with vertical transmission,” *Applied Mathematics Letters*, vol. 18, no. 7, pp. 729–732, 2005.
- [33] G. Zaman, Y. Kang, and I. Jung, “Stability analysis and optimal vaccination of an SIR epidemic model,” *BioSystems*, vol. 93, no. 3, pp. 240–249, 2008.
- [34] F. Campos Filho, *Algoritmos Numéricos*. LTC, 2007.
- [35] G. Zaman, Y. Kang, and I. Jung, “Optimal treatment of an SIR epidemic model with time delay,” *Biosystems*, vol. 98, no. 1, pp. 43–50, 2009.
- [36] S. Liu, Y. Pei, C. Li, and L. Chen, “Three kinds of TVS in a SIR epidemic model with saturated infectious force and vertical transmission,” *Applied Mathematical Modelling*, vol. 33, no. 4, pp. 1923–1932, 2009.
- [37] C. Zou and J. Shi, “Generalization of the Floquet theory,” *Nonlinear Analysis: Theory, Methods & Applications*, vol. 71, no. 3-4, pp. 1100–1107, 2009.
- [38] S. Gao, Z. Teng, and D. Xie, “Analysis of a delayed SIR epidemic model with pulse vaccination,” *Chaos, Solitons & Fractals*, vol. 40, no. 2, pp. 1004–1011, 2009.
- [39] X. Zhang, H. Huo, X. Sun, and Q. Fu, “The differential susceptibility SIR epidemic model with stage structure and pulse vaccination,” *Nonlinear Analysis: Real World Applications*, vol. 11, no. 4, pp. 2634–2646, 2010.
- [40] L. Shaw, W. Spears, L. Billings, and P. Maxim, “Effective vaccination policies,” *Information Sciences*, 2010.
- [41] K. Deb, S. Agrawal, A. Pratap, and T. Meyarivan, “A fast elitist non-dominated sorting genetic algorithm for multi-objective optimization: NSGA-II,” in *Parallel Problem Solving from Nature PPSN VI*, pp. 849–858, Springer, 2000.
- [42] K. Deb, *Multi-objective optimization using evolutionary algorithms*. Wiley, 2001.
- [43] T. Britton, “Stochastic epidemic models: a survey,” *Mathematical Biosciences*, 2010.

- [44] H. Levy, “Stochastic dominance and expected utility: survey and analysis,” *Management Science*, vol. 38, no. 4, pp. 555–593, 1992.
- [45] L. Temime, G. Hejblum, M. Setbon, and A. Valleron, “The rising impact of mathematical modelling in epidemiology: antibiotic resistance research as a case study,” *Epidemiology and Infection*, vol. 136, no. 03, pp. 289–298, 2007.
- [46] D. Bernoulli, “Essai d’une nouvelle analyse de la mortalité causée par la petite vérole et des avantages de l’inoculation pour la prévenir,” *Mémoires de Mathématiques et de Physique, Académie Royale des Sciences, Paris*, pp. 1–45, 1760.
- [47] S. Ross, *The prevention of malaria*. Dutton, 1910.
- [48] W. Kermack and A. McKendrick, “Contributions to the mathematical theory of epidemics-III. Further studies of the problem of endemicity,” *Bulletin of mathematical biology*, vol. 53, no. 1, pp. 89–118, 1991.
- [49] M. Bartlett, “Some evolutionary stochastic processes,” *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 211–229, 1949.
- [50] D. Kendall, “Deterministic and stochastic epidemics in closed populations,” in *Proceedings of the third Berkeley Symposium on mathematical statistics and probability: contributions to the theory of statistics*, p. 149, University of California Press, 1956.
- [51] O. BJØRNSTAD, B. Finkenstädt, and B. Grenfell, “Dynamics of measles epidemics: estimating scaling of transmission rates using a time series SIR model,” *Ecological Monographs*, vol. 72, no. 2, pp. 169–184, 2002.
- [52] J. Dushoff, J. Plotkin, S. Levin, and D. Earn, “Dynamical resonance can account for seasonality of influenza epidemics,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 101, no. 48, p. 16915, 2004.
- [53] K. Dietz, “Transmission and control of arbovirus diseases,” *Proceedings of society for industrial and applied mathematics, Philadelphia*, pp. 104–121, 1974.
- [54] J. Heesterbeek, “A brief history of R_0 and a recipe for its calculation,” *Acta Biotheoretica*, vol. 50, no. 3, pp. 189–204, 2002.

- [55] H. Hethcote and J. Van Ark, “Epidemiological models for heterogeneous populations: proportionate mixing, parameter estimation, and immunization programs,” *Mathematical Biosciences*, vol. 84, no. 1, pp. 85–118, 1987.
- [56] H. Hethcote, “A thousand and one epidemic models,” *Lecture Notes in Biomathematics*, pp. 504–504, 1994.
- [57] J. Fox, L. Elveback, W. Scott, L. Gatewood, and E. Ackerman, “Herd immunity: basic concept and relevance to public health immunization practices,” *American Journal of Epidemiology*, vol. 94, no. 3, p. 179, 1971.
- [58] L. Elveback, J. Fox, E. Ackerman, A. Langworthy, M. Boyd, and L. Gatewood, “An influenza simulation model for immunization studies,” *American Journal of Epidemiology*, vol. 103, no. 2, p. 152, 1976.
- [59] M. Huston, D. DeAngelis, and W. Post, “New computer models unify ecological theory,” *BioScience*, vol. 38, no. 10, pp. 682–691, 1988.
- [60] O. Judson, “The rise of the individual-based model in ecology,” *Trends in Ecology & Evolution*, vol. 9, no. 1, pp. 9–14, 1994.
- [61] V. Grimm and S. Railsback, *Individual-based modeling and ecology*. Princeton University Press, 2005.
- [62] N. Gilbert and K. Troitzsch, *Simulation for the Social Scientist*. Open University Press, 2005.
- [63] L. Tesfatsion, “Agent-based computational economics: Growing economies from the bottom up,” *Artificial life*, vol. 8, no. 1, pp. 55–82, 2002.
- [64] F. Billari and A. Prskawetz, “Agent-based computational demography: Using simulation to improve our understanding of demographic behaviour (contributions to economics),” *Heidelberg: Physica-Verlag*, 2003.
- [65] D. Parker, S. Manson, M. Janssen, M. Hoffmann, and P. Deadman, “Multi-Agent Systems for the Simulation of Land-Use and Land-Cover Change: A Review,” *Annals of the Association of American Geographers*, vol. 93, no. 2, pp. 314–337, 2003.
- [66] R. Huckfeldt, P. Johnson, and J. Sprague, *Political disagreement: The survival of diverse opinions within communication networks*. Cambridge University Press, 2004.

- [67] S. Railsback, “Concepts from complex adaptive systems as a framework for individual-based modelling,” *Ecological modelling*, vol. 139, no. 1, pp. 47–62, 2001.
- [68] V. Grimm, U. Berger, F. Bastiansen, S. Eliassen, V. Ginot, J. Giske, J. Goss-Custard, T. Grand, S. Heinz, G. Huse, *et al.*, “A standard protocol for describing individual-based and agent-based models,” *Ecological Modelling*, vol. 198, no. 1-2, pp. 115–126, 2006.
- [69] H. Giacomini, “Sete motivações teóricas para o uso da modelagem baseada no indivíduo em ecologia,” *Acta Amazonica*, vol. 37, no. 3, 2007.
- [70] L. Alvarenga, “Modelagem de epidemias através de modelos baseados em indivíduos,” Master’s thesis, Universidade Federal de Minas Gerais, 2008.
- [71] E. Nepomuceno, *Dinâmica, Modelagem e Controle de Epidemias*. PhD thesis, Universidade Federal de Minas Gerais, 2005.
- [72] A. Almeida, “Modelagem Dinâmica de Epidemias Baseadas no Modelo SIR sobre Redes Sociais,” 2010. Exame de Qualificação de Doutorado.
- [73] A. Papoulis, S. Pillai, and S. Unnikrishna, *Probability, random variables, and stochastic processes*, vol. 73660116. McGraw-Hill New York, 2002.
- [74] S. M. Ross, *Simulation*. Academic Press, 4 ed., August 2006.
- [75] R. Takahashi, “Otimização escalar e vetorial.” Notas de aula, 2004.
- [76] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, March 2004.
- [77] V. Chankong and Y. Haimes, *Multiobjective decision making: theory and methodology*. North-Holland New York, 1983.
- [78] C. Darwin, *The variation of animals and plants under domestication*. Murray, 1868.
- [79] J. Holland, “Concerning efficient adaptive systems,” *Self-Organizing Systems*, vol. 230, 1962.
- [80] J. Holland, “Outline for a logical theory of adaptive systems,” *Journal of the ACM (JACM)*, vol. 9, no. 3, pp. 297–314, 1962.

- [81] D. Goldberg, *Genetic algorithms in search, optimization, and machine learning*. Addison-wesley, 1989.
- [82] J. Holland, *Adaptation in natural and artificial systems*. MIT Press Cambridge, MA, USA, 1992.
- [83] J. Horn, N. Nafpliotis, and D. E. Goldberg, “A niched pareto genetic algorithm for multiobjective optimization,” *Evolutionary Computation, 1994. IEEE World Congress on Computational Intelligence., Proceedings of the First IEEE Conference on*, pp. 82–87 vol.1, 1994.
- [84] N. Srinivas and K. Deb, “Multiobjective optimization using nondominated sorting in genetic algorithms,” *Evolutionary Computation*, vol. 2, pp. 221–248, September 1994.
- [85] C. M. Fonseca and P. J. Fleming, “An overview of evolutionary algorithms in multiobjective optimization,” *Evolutionary Computation*, vol. 3, pp. 1–16, 1995.
- [86] P. J. Angeline, Z. Michalewicz, M. Schoenauer, X. Yao, and A. Zalzal, eds., *The Pareto Archived Evolution Strategy: A New Baseline Algorithm for Pareto Multiobjective Optimisation*, vol. 1, (Mayflower Hotel, Washington D.C., USA), IEEE Press, June-September 1999.
- [87] E. Zitzler and L. Thiele, “Multiobjective evolutionary algorithms: A comparative case study and the strength pareto approach,” *IEEE transactions on Evolutionary Computation*, vol. 3, no. 4, p. 257, 1999.
- [88] K. Deb and R. Agrawal, “Simulated binary crossover for continuous search space,” *Complex systems*, vol. 9, no. 2, pp. 115–148, 1995.
- [89] P. Moscato, “On evolution, search, optimization, genetic algorithms and martial arts: Towards memetic algorithms,” *Caltech Concurrent Computation Program, C3P Report*, vol. 826, p. 1989, 1989.
- [90] P. Moscato, “Memetic algorithms: a short introduction,” in *New ideas in optimization*, (Maidenhead, UK, England), pp. 219–234, McGraw-Hill Ltd., UK, 1999.
- [91] J. Knowles and D. Corne, “M-PAES: A memetic algorithm for multiobjective optimization,” in *2000 Congress on Evolutionary Computation*, vol. 1, pp. 325–332, Citeseer, 2000.

- [92] H. Ishibuchi, T. Yoshida, and T. Murata, “Balance between genetic search and local search in memetic algorithms for multiobjective permutation flowshop scheduling,” *Evolutionary Computation, IEEE Transactions on*, vol. 7, pp. 204–223, May 2003.
- [93] E. Burke and J. Landa Silva, “The design of memetic algorithms for scheduling and timetabling problems,” *Recent Advances in Memetic Algorithms*, pp. 289–311, 2005.
- [94] V. J. Garcia, P. M. Franca, A. Mendes, and P. Moscato, “A parallel memetic algorithm applied to the total tardiness machine scheduling problem,” *Parallel and Distributed Processing Symposium, 2006. IPDPS 2006. 20th International*, pp. 8 pp.+, 2006.
- [95] E. Wanner, F. Guimarães, R. Takahashi, and P. Fleming, “Local search with quadratic approximations into memetic algorithms for optimization with multiple criteria,” *Evolutionary computation*, vol. 16, no. 2, pp. 185–224, 2008.
- [96] Q. H. Nguyen, Y. S. Ong, and M. H. Lim, “A probabilistic memetic framework,” *Evolutionary Computation, IEEE Transactions on*, vol. 13, no. 3, pp. 604–623, 2009.
- [97] J. Rosen and R. Marcia, “Convex quadratic approximation,” *Computational Optimization and Applications*, vol. 28, no. 2, pp. 173–184, 2004.
- [98] J. Rice, *The approximation of functions. 1. Linear theory*. Addison-Wesley, 1964.
- [99] J. Hayes, *Numerical approximation to functions and data*. Athlone Press University of London, 1970.
- [100] M. Powell, *Approximation theory and methods*. Cambridge University Press, 1981.
- [101] T. Rivlin, *An introduction to the approximation of functions*. Dover Publications, 2003.
- [102] M. Powell, “On the Lagrange functions of quadratic models that are defined by interpolation,” *Optimization Methods and Software*, vol. 16, no. 1, pp. 289–309, 2001.
- [103] D. Hertog, E. Klerk, and J. Roos, “On convex quadratic approximation,” *Open Access publications from Tilburg University*, 2000.

- [104] K. Dill, A. Phillips, and J. Rosen, “CGU: An algorithm for molecular structure prediction,” *IMA Volumes in Mathematics and its Applications*, vol. 94, pp. 1–22, 1997.
- [105] K. Dill, A. Phillips, and J. Rosen, “Protein structure and energy landscape dependence on sequence using a continuous energy function,” *Journal of Computational Biology*, vol. 4, no. 3, pp. 227–239, 1997.
- [106] J. Mitchell, J. Rosen, A. Phillips, and L. Ten Eyck, “Coupled optimization in protein docking,” in *Proceedings of the third annual international conference on Computational molecular biology*, p. 284, ACM, 1999.
- [107] R. Marcia, J. Mitchell, and J. Rosen, “Iterative convex quadratic approximation for global optimization in protein docking,” *Computational Optimization and Applications*, vol. 32, no. 3, pp. 285–297, 2005.
- [108] F. Hillier and G. Lieberman, *Introduction to Operations Research*. McGraw-Hill Science/Engineering/Math, 2004.
- [109] D. P. Bertsekas, *Dynamic Programming and Optimal Control*. Athena Scientific, 1995.
- [110] E. Zitzler, *Evolutionary Algorithms for Multiobjective Optimization: Methods and Applications*. PhD thesis, ETH Zurich, Switzerland, 1999.

Índice

- Adequabilidade, 39
Algoritmos Evolucionários, 32
Algoritmos Genéticos, 32
Algoritmos Meméticos, 40
Aproximações Quadráticas Convexas, 43, 44
- Bartlett, 11
Busca Local, 42, 46, 47
- Classes de Estados, 11
Coeficiente de Ação de Massa, 14
Comunidade Dinâmica, 12
Comunidade Fechada, 12
Comunidade Homogênea e Uniforme, 13
- Conjunto Não Dominado, 29
Conjunto Pareto-Otimo Global, 29
Controle Impulsivo, 5, 49, 50
Cromossomo, 32
- Daniel Bernoulli, 11
Dominância, 28
Dominância Estocástica, 55
- Endemia, 2
Epidemia, 2
Epidemiologia, 1
Epidemiologia Matemática, 1
Espaço de Objetivos, 28
Estabilidade Global, 6
Estabilidade Local, 6
- Força de Infecção, 14
Fronteira de Pareto, 30
- Fronteira Ótima, 30
Função Convexa, 43
Função Quadrática, 44
Funções Convexas, 43
Funções Objetivo, 28
- Genes, 32
- Heurística, 5
Hipócrates, 1
- Impulsos, 5
Imunidade Adquirida, 13
Incidência Horizontal, 14
Incidência Padrão, 14
Infectado, 3
- John Holland, 32
John Snow, 1
- Kendall, 11
Kermack, 2, 11
- Malha Aberta, 50
McKendrick, 2, 11
Método de Runge-Kutta, 6
Mobilidade do Estado de Saúde, 12
Modelagem Matemática, 10
Modelo Baseado em Agente, 17
Modelo Baseado em Indivíduo (MBI), 2, 4, 17
 Algoritmo, 25
 Convergência, 20
Modelo Biobjetivo, 51
Modelo Endêmico, 2

- Modelo Epidêmico, 2
 Modelo SIR, 2, 3
 Modelo SIR Endêmico, 13, 15
 Modelos Matemáticos, 1, 10
 MSEIR, 12
 MSEIRS, 12
 NSGA-II, 34
 Crowded-Comparison Operator, 38
 Crowding Distance, 34, 36, 39
 Cruzamento, 35
 Elitismo Multiobjetivo, 34
 Fast Non-Dominated Sorting Procedure, 34, 35, 37
 Fill Nondominated, 39, 40
 Módulo Principal, 39, 41
 Mutaçãõ, 35
 Preservação da Diversidade, 34
 Torneio Binário, 35
 Tratamento de Restrições, 34
 NSGA-II Memético, 45, 48
 Número Básico de Reprodução, 12
 Número de Contato, 12
 Número de Substituição, 12

 Ordem Parcial Estrita, 29
 Otimização, 26
 Otimização Mono-Objetivo, 26
 Otimização Multiobjetivo, 26
 Otimização Vetorial, 28

 Pareto-Ótimo Local, 30
 Período Médio de Infecção, 14
 Política de Controle, 7
 Problema de Otimização Multiobjetivo, 27

 Recombinação Sexual, 32
 Recuperado, 3
 Região Factível, 28
 Restrição, 26
 Ronald Ross, 11

 S-Metric, 53
 SEI, 12
 SEIR, 12
 SEIRS, 12
 SEIS, 12
 Seleção Natural, 32
 SI, 12
 Simulacao Monte Carlo, 54
 Simulated Binary Cross-Over, 35
 Simulated Binary Mutation, 35
 SIRS, 12
 SIS, 12
 Sistema Impulsivo, 5
 Solução Factível, 28
 Solução Infactível, 28
 Solução Nadir, 30
 Solução Utópica, 30
 Suscetível, 3

 Taxa de Imunização, 14
 Taxa de Infecção, 14
 Taxa de Mortalidade, 14
 Taxa de Nascimento, 14
 Tempo Médio de Vida, 14
 Transmissão Horizontal, 6
 Transmissão Vertical, 6

 Vacinação Impulsiva, 6