

UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ  
CÂMPUS CURITIBA  
ENGENHARIA DE CONTROLE E AUTOMAÇÃO

ÁLISSON LUÍS MATEUS MADRUGA  
LUIS FERNANDO BARATIERI STOCCO  
NICOLAS FREDERICE MALNARCIC

**FILTRO DE PARTÍCULAS APLICADO AO RASTREAMENTO E  
PREVISÃO DE EPIDEMIAS**

TRABALHO DE CONCLUSÃO DE CURSO

CURITIBA

2020

ÁLISSON LUÍS MATEUS MADRUGA  
LUIS FERNANDO BARATIERI STOCCO  
NICOLAS FREDERICE MALNARCIC

**FILTRO DE PARTÍCULAS APLICADO AO RASTREAMENTO E  
PREVISÃO DE EPIDEMIAS**

Trabalho de Conclusão de Curso de Graduação, apresentado à disciplina de Metodologia Aplicada ao TCC, do curso de Engenharia de Controle e Automação do Departamento Acadêmico de Eletrotécnica (DAELT) da Universidade Tecnológica Federal do Paraná (UTFPR) como requisito para obtenção do título de Engenheiro de Controle e Automação.

Orientador: Prof. Dr. Victor Baptista Frencl

CURITIBA  
2020

Álison Luís Mateus Madruga  
Luis Fernando Baratieri Stocco  
Nicolas Frederice Malnarcic

## Filtro de Partículas aplicado ao rastreamento e previsão de epidemias

Este Trabalho de Conclusão de Curso de Graduação foi julgado e aprovado como requisito parcial para a obtenção do Título de Engenheiro de Controle e Automação, do curso de Engenharia de Controle e Automação do Departamento Acadêmico de Eletrotécnica (DAELT) da Universidade Tecnológica Federal do Paraná (UTFPR).

Curitiba, 25 de novembro de 2020.

---

Prof. Paulo Sérgio Walenia, Esp.  
Coordenador de Curso  
Engenharia de Controle e Automação

---

Prof. Amauri Amorin Assef, Dr.  
Responsável pelos Trabalhos de Conclusão de Curso  
de Engenharia de Controle e Automação do DAELT

### ORIENTAÇÃO

---

Prof. Dr. Victor Baptista Frencl  
Universidade Tecnológica Federal do Paraná  
Orientador

### BANCA EXAMINADORA

---

Prof. Dr. Elder Oroski  
Universidade Tecnológica Federal do Paraná

---

Prof. Dr. Luiz Ledo Mota Melo Junior  
Universidade Tecnológica Federal do Paraná

---

Prof. Dr. Victor Baptista Frencl  
Universidade Tecnológica Federal do Paraná

A folha de aprovação assinada encontra-se na Coordenação do Curso de Engenharia de Controle e Automação.

## RESUMO

MADRUGA, Álisson L. M., STOCCO, Luis, F. B., MALNARCIC, Nicolas F.. FILTRO DE PARTÍCULAS APLICADO AO RASTREAMENTO E PREVISÃO DE EPIDEMIAS. 75f. Trabalho de Conclusão de Curso - Engenharia de Controle e Automação, Universidade Tecnológica Federal do Paraná. Curitiba 2020.

Um dos objetivos da filtragem estocástica é obter estimativas confiáveis o mais próximo possível dos valores exatos. Devido a essa característica, buscou-se aplicar métodos de filtragem estocástica ao problema de previsão e rastreamento de epidemias. Na literatura, três tipos de modelos que representam dinâmicas populacionais diferentes durante epidemias podem ser encontrados: modelo SIS (Suscetíveis Infectados Suscetíveis), modelo SIR (Suscetíveis Infectados Recuperados) e modelo SEIR (Suscetíveis Expostos Infectados Recuperados). Esses modelos são não lineares, motivando a aplicação de filtros estocásticos não lineares para a obtenção das estimativas. O foco desse trabalho é a aplicação do Filtro de Partículas aos modelos SIS, SIR e SEIR. Para efeito de comparação de desempenho, o Filtro de Kalman Estendido também foi estudado. Ambos os filtros estocásticos foram implementados através do *software* MATLAB, o qual também foi utilizado para gerar os dados dos modelos. Os dados dos modelos foram gerados de duas formas diferentes: (i) a partir do próprio modelo matemático, denominado GS (Geração Sintética); e (ii) através do algoritmo aberto MOSES (do inglês, *MATLAB-Based Open-Source Stochastic Epidemic Simulator*). Para efeito de comparação, obteve-se o desempenho dos filtros implementados com base no RMSE (do inglês, *Root Mean Squared Error*) das estimativas em relação aos valores exatos/simulados. Os resultados obtidos ficaram, em 4 das 5 simulações, dentro do que era esperado, com o FP apresentando menos erro que o FKE.

**Palavras-chave:** Epidemia, Modelos não lineares, Modelo SIS, Modelo SIR, Modelo SEIR, Filtragem Estocástica, Filtro de Kalman Estendido, Filtro de Partículas.

## ABSTRACT

MADRUGA, Álisson L. M., STOCCO, Luis, F. B., MALNARCIC, Nicolas F.. PARTICLE FILTERS APPLIED TO TRACKING AND PREDICTION OF EPIDEMICS. 75p. Engenharia de Controle e Automação, Universidade Tecnológica Federal do Paraná. Curitiba 2020.

One of stochastic filtering objectives is to obtain more reliable and trustworthy data. Due to that characteristic, it was decided to apply some of the most well-known stochastic filtering methods to predicting and tracking epidemics. Through literature, three different models that represent populational dynamics in epidemics were obtained. The models are: SIS (Susceptible Infected Susceptible) model, SIR (Susceptible Infected Recovered) model and SEIR (Susceptible Exposed Infected Recovered) model. In all these cases, white noise was added to the data set in order to simulate a real measurement sample. Since the models are all non-linear, all the stochastic filters applied to the models are non-linear as well. This work is focused on the application of the Particle Filter to the aforementioned models. However, for matters of comparison, the Extended Kalman Filter was also implemented. Both filters were implemented using MATLAB software. The data obtained from the models was also generated in MATLAB, but through two different methods. The first method was with the data being generated based on the model itself, which was named Synthetic Generation. The second method was through a code written in MATLAB called MOSES (MATLAB-Based Open-Source Stochastic Epidemic Simulator), which was obtained from literature. Finally, so that all of the results that were obtained could be compared, the RMSE (Root Mean Squared Error) of each data set obtained through the stochastic filtering process. The results were within what was expected, with the Particle Filter outperforming de Extended Kalman Filter 4 out of 5 times.

**Keywords:** Epidemic, Non-linear models, SIS model, SIR model, SEIR model, Stochastic filtering, Extended Kalman Filter, Particle Filter.

## LISTA DE ABREVIACOES

<b>AIDS</b>	<i>Acquired Immunodeficiency Syndrome</i>
<b>ASI</b>	Amostragem Sequencial por Importncia
<b>CSW</b>	<i>Cumulative Sum of Weights</i>
<b>CTMC</b>	<i>Continuous Time Markov Chain</i>
<b>Covid-19</b>	<i>Coronavirus Disease - 2019</i>
<b>DTMC</b>	<i>Discrete Time Markov Chain</i>
<b>EDO</b>	Equao Diferencial Ordinria
<b>FDP</b>	Funo Densidade de Probabilidade
<b>FK</b>	Filtro de Kalman
<b>FKE</b>	Filtro de Kalman Estendido
<b>FKU</b>	Filtro de Kalman <i>Unscented</i>
<b>FP</b>	Filtro de Partculas
<b>GS</b>	Gerao Sinttica
<b>HIV</b>	<i>Human Immunodeficiency Virus</i>
<b>IMM</b>	<i>Interacting Multiple Model</i>
<b>MC</b>	Monte Carlo
<b>MDF</b>	Mtodo das Diferenas Finitas
<b>MOSES</b>	<i>MATLAB-Based Open-Source Stochastic Epidemic Simulator</i>
<b>NASA</b>	<i>National Aeronautics and Space Administration</i>
<b>PDF</b>	<i>Probability Density Function</i>
<b>RMSE</b>	<i>Root Mean Squared Error</i>

<b>SARS</b>	<i>Severe Acute Respiratory Syndrome</i>
<b>SDE</b>	<i>Stochastic Differential Equation</i>
<b>SEIR</b>	Suscetíveis Expostos Infectados Recuperados
<b>SEQIJR</b>	<i>Susceptible Exposed Quarantined Infected Isolated Recovered</i>
<b>SIR</b>	Suscetíveis Infectados Recuperados
<b>SIS</b>	Suscetíveis Infectados Suscetíveis
<b>SMC</b>	<i>Sequential Monte Carlo</i>
<b>TCC</b>	Trabalho de Conclusão de Curso
<b>UT</b>	<i>Unscented Transformation</i>

## LISTA DE TABELAS

TABELA 1	– Tabela de parâmetros das simulações. ....	51
TABELA 2	– Tabela do erro em número de indivíduos do modelo SIS gerado pelo MOSES. ....	56
TABELA 3	– Tabela do erro em número de indivíduos do modelo SIR gerado pelo MOSES. ....	60
TABELA 4	– Tabela do erro em número de indivíduos do modelo SIS gerado sinteticamente. ....	63
TABELA 5	– Tabela do erro em número de indivíduos do modelo SIR gerado sinteticamente. ....	65
TABELA 6	– Tabela do erro em número de indivíduos do modelo SEIR gerado sinteticamente. ....	67



## LISTA DE FIGURAS

FIGURA 1	– Várias realizações de um processo aleatório. ....	22
FIGURA 2	– Diagrama de estados do modelo SIS. Linhas sólidas representam trocas de estado, linhas pontilhadas apontando para dentro dos compartimentos representam nascimentos e as linhas pontilhadas apontando para fora dos compartimentos representam mortes. ....	23
FIGURA 3	– Diagrama dos estados do modelo SIR. Linhas sólidas representam trocas de estado, linhas pontilhadas apontando para dentro dos compartimentos representam nascimentos e as linhas pontilhadas apontando para fora dos compartimentos representam mortes. ....	25
FIGURA 4	– Diagrama dos estados do modelo SEIR. Linhas sólidas representam trocas de estado, linhas pontilhadas apontando para dentro dos compartimentos representam nascimentos e as linhas pontilhadas apontando para fora dos compartimentos representam mortes. ....	32
FIGURA 5	– Diagrama de estados mostrando as transições entre diferentes etapas e “superetapas” e variáveis de relação no simulador MOSES. Linhas tracejadas indicam transições entre etapas iguais. ....	35
FIGURA 6	– Representação dos sorteios e associação com a soma acumulada. ....	48
FIGURA 7	– Um ciclo de um Filtro de Partículas. ....	48
FIGURA 8	– Gráfico do resultado do cenário 1. ....	52
FIGURA 9	– Gráfico do resultado do cenário 2. ....	52
FIGURA 10	– Gráfico da aplicação do FKE ao modelo SIS (Cenário 1). ....	54
FIGURA 11	– <i>Zoom</i> nas etapas de transição da aplicação do FKE ao modelo SIS (Cenário 1) dos números de indivíduos de cada população. ....	55
FIGURA 12	– RMSE, erro absoluto e médio do FKE em relação aos valores reais. ....	55
FIGURA 13	– Gráfico da aplicação do FP ao modelo SIS (Cenário 1). ....	56
FIGURA 14	– <i>Zoom</i> nas etapas de transição da aplicação do FP ao modelo SIS (Cenário 1) dos números de indivíduos de cada população. ....	57
FIGURA 15	– RMSE, erro absoluto e médio do FP em relação aos valores reais. ....	57
FIGURA 16	– Gráfico da aplicação do FKE ao modelo SIR (Cenário 2). ....	58
FIGURA 17	– <i>Zoom</i> nas etapas de transição da aplicação do FKE ao modelo SIR (Cenário 2) dos números de indivíduos de cada população. ....	59

FIGURA 18	– RMSE, erro absoluto e médio do FKE em relação aos valores reais. ....	60
FIGURA 19	– Gráfico da aplicação do FP ao modelo SIR (Cenário 2). ....	61
FIGURA 20	– <i>Zoom</i> nas etapas de transição da aplicação do FP ao modelo SIR (Cenário 2) dos números de indivíduos de cada população. ....	61
FIGURA 21	– RMSE, erro absoluto e médio do FP em relação aos valores reais. ....	62
FIGURA 22	– Gráfico do modelo SIS. ....	63
FIGURA 23	– Gráfico da aplicação do FKE ao modelo SIS. ....	64
FIGURA 24	– Gráfico da aplicação do FP ao modelo SIS. ....	64
FIGURA 25	– Gráfico do modelo SIR. ....	65
FIGURA 26	– Gráfico da aplicação do FKE ao modelo SIR. ....	66
FIGURA 27	– Gráfico da aplicação do FP ao modelo SIR. ....	66
FIGURA 28	– Gráfico do modelo SEIR. ....	67
FIGURA 29	– Gráfico da aplicação do FKE ao modelo SEIR. ....	68
FIGURA 30	– Gráfico da aplicação do FP ao modelo SEIR. ....	68

## SUMÁRIO

<b>1 INTRODUÇÃO</b>	<b>12</b>
1.1 TEMA	12
1.1.1 Delimitação do Tema	14
1.2 PROBLEMAS E PREMISSAS	14
1.3 OBJETIVOS	15
1.3.1 Objetivo Geral	15
1.3.2 Objetivos Específicos	15
1.4 JUSTIFICATIVA	16
1.5 PROCEDIMENTOS METODOLÓGICOS	16
1.6 ESTRUTURA DO TRABALHO	17
<b>2 MODELOS EPIDÊMICOS</b>	<b>18</b>
2.1 MODELOS MATEMÁTICOS	18
2.1.1 Modelos Lineares	18
2.1.2 Modelos Não Lineares	19
2.1.3 Modelos Dinâmicos	19
2.1.4 Modelos de Espaço de estados	19
2.1.5 Modelos Estocásticos	20
2.2 MODELOS MATEMÁTICOS DE EPIDEMIAS	22
2.2.1 Modelo SIS	23
2.2.1.1 Discretização do modelo SIS	24
2.2.2 Modelo SIR	25
2.2.2.1 Discretização do modelo SIR	26
2.2.3 Modelo DTMC SIS	27
2.2.4 Modelo DTMC SIR	29
2.2.5 Variações de modelos	30
2.2.6 Modelo SEIR	31
2.2.7 Modelo MOSES para MATLAB	33
2.3 SÍNTESE DO CAPÍTULO	35
<b>3 FILTROS ESTOCÁSTICOS</b>	<b>36</b>
3.1 FILTRO DE KALMAN	36
3.1.1 Modelo dinâmico linear	36
3.1.2 Predição	37
3.1.3 Atualização	38
3.1.4 Pseudocódigo para o FK	39
3.2 FILTRO DE KALMAN ESTENDIDO	39
3.2.1 Predição	40
3.2.2 Atualização	41
3.2.3 Pseudocódigo para o FKE	42
3.3 FILTRO DE PARTÍCULAS	42
3.3.1 Integração de Monte Carlo	43
3.3.2 Amostragem Sequencial Por Importância	45

3.3.3 Problema da Degeneração .....	47
3.3.4 Reamostragem Sequencial Por Importância .....	47
3.3.5 Pseudocódigo para o FP .....	49
3.4 SÍNTESE DO CAPÍTULO .....	49
<b>4 SIMULAÇÕES E RESULTADOS .....</b>	<b>50</b>
4.1 SIMULAÇÃO DO MODELO MOSES .....	50
4.2 APLICAÇÃO DO FKE — MODELO DE EPIDEMIA SIS E DADOS DO MODELO MOSES .....	53
4.3 APLICAÇÃO DO FP — MODELO DE EPIDEMIA SIS E OS DADOS DO MODELO MOSES .....	56
4.4 APLICAÇÃO DO FKE — MODELO DE EPIDEMIA SIR E OS DADOS DO MODELO MOSES .....	58
4.5 APLICAÇÃO DO FP — MODELO DE EPIDEMIA SIR E OS DADOS DO MODELO MOSES .....	60
4.6 ANÁLISE FKE E FP ATRAVÉS DA GS DE DADOS SIMPLES .....	62
4.7 SÍNTESE DO CAPÍTULO .....	69
<b>5 CONCLUSÕES .....</b>	<b>70</b>
<b>REFERÊNCIAS .....</b>	<b>72</b>

# 1 INTRODUÇÃO

## 1.1 TEMA

Epidemias são capazes de afetar a sociedade e os indivíduos que a compõe de diversas maneiras. A forma como as epidemias afetam os indivíduos varia desde uma forma direta, como doenças capazes de assolar milhares de pessoas, até de formas indiretas, como pragas que afetam plantações e, conseqüentemente, a alimentação e a economia de uma comunidade (YORINORI et al., 2005). A pandemia do Coronavírus evidencia esse impacto que uma epidemia de nível global tem no dia a dia de todos (WHO, 2020). Ser capaz de rastrear e prever efetivamente o progresso destas epidemias pode ajudar, por exemplo, as autoridades de saúde pública a monitorar, planejar e intervir de forma mais eficiente, com o objetivo de controlá-las (CAZELLES; CHAU, 1997).

Para tornar possível a previsão e o rastreamento dessas epidemias, é necessário definir um modelo matemático capaz de representar adequadamente o seu comportamento. De maneira geral, a modelagem matemática que melhor descreve a dinâmica de uma epidemia possui características não lineares <sup>1</sup> (OGATA; YANG, 2002).

Nos modelos matemáticos não lineares que representam comportamentos epidêmicos, considera-se a existência de perturbações ou ruídos, os quais estão relacionados às mudanças do ambiente e à dinâmica entre indivíduos de uma comunidade (VLADAR; PEN, 2007). Este tipo de modelo que leva em consideração os ruídos aleatórios nas suas equações dinâmicas e/ou de observações é chamado de modelo estocástico. Nesse modelo, os ruídos são representados por variáveis aleatórias que dependem do tempo, definidas como processos estocásticos (BHATTACHARYA; WAYMIRE, 2009).

A “contaminação” dos modelos pelos ruídos prejudicam a obtenção de informações precisas sobre o sistema, tornando-se necessário o uso de ferramentas matemáticas que lidem adequadamente com estes sistemas ruidosos. Estas ferramentas são os filtros estocásticos. No

---

<sup>1</sup>Um modelo é classificado como não linear se o princípio de superposição não se aplicar a ele

caso de modelos lineares, os métodos de filtragem estocástica utilizam a equação de observação e da dinâmica do sistema, além de um ganho determinado pela matriz de covariância do erro de estimativa. Já nos modelos não lineares existem diversos filtros estocásticos, dentre eles aquele que utiliza a linearização do modelo e obtém a estimativa de estado e o que realiza uma estimativa de distribuição condicional do sinal. Filtrar um determinado processo estocástico consiste em remover ou diminuir a interferência de ruídos indesejados, tais quais inexatidões advindas de distúrbios na fonte de medição e das incertezas do modelo matemático sendo utilizado (BUCY; JOSEPH, 2005).

O filtro estocástico mais comumente utilizado para tratar modelos matemáticos lineares (discretos e discretizados), estes representados em espaços de estado, é o Filtro de Kalman (FK). Este filtro é capaz de estimar grandezas do sistema modelado com o objetivo de obter estimativas mais próximas possíveis dos valores exatos<sup>2</sup>, com base em cálculos de valores esperados e matrizes de covariâncias, ambos condicionados às medidas (KALMAN, 1960). Na literatura existem diversos filtros estocásticos, cada um com sua aplicabilidade. Para situações em que o sistema é descrito por modelos não lineares há o Filtro de Kalman Estendido (FKE) (BAR-SHALOM et al., 2004), o Filtro de Kalman *Unscented* (FKU), o Filtro IMM (do inglês, *Interacting Multiple Model*) (LI; JILKOV, 2005) e o Filtro de Partículas (FP).

Com o aumento da complexidade dos modelos não lineares e de ruídos com distribuições diferentes da gaussiana, demanda-se um método de filtragem estocástica que possa lidar com esse tipo de situação. O filtro que atende a estes requisitos é o FP, o qual opera com o processamento de  $N$  partículas a cada instante de tempo  $k$ . Como o número de partículas costuma ser elevado, é necessário uma alta capacidade de processamento, que é possível com os computadores atuais. Em essência, o FP opera com um grupo de dados amostrados aleatoriamente. Esses valores amostrados, são propagados ao longo do tempo para obter a distribuição de probabilidade *a posteriori* do estado em cada instante de tempo. Assim, a cada partícula, atribui-se um peso por importância em relação a sua distribuição de probabilidade *a posteriori* (GORDON et al., 2004).

O uso do FP para tratar de modelos estocásticos não lineares é comum em diversas áreas, por exemplo, na geolocalização (GUSTAFSSON, 2010), no rastreamento de multidões (FREITAS et al., 2016), em tratamento de dados econômicos (ARGUETA, 2013), entre outros.

Filtros estocásticos já foram usados para monitoramento de doenças como a malária na Índia (BHADRA et al., 2011), síndrome respiratória aguda grave ou SARS (do inglês, *Severe Acute Respiratory Syndrome*) na China (ANDERSON et al., 2004) e AIDS/HIV (do

---

<sup>2</sup>Sem a presença de ruído.

inglês, *Acquired Immunodeficiency Syndrome/Human Immunodeficiency Virus*) nos anos 1990 (CAZELLES; CHAU, 1997). A implementação do estudo de previsão e rastreamento de epidemias faz-se muito relevante, principalmente em um país como o Brasil que sofre de surtos constantes de epidemias como dengue, zika e chikungunya, por exemplo (VALLE et al., 2016) e também da pandemia do Coronavírus (WHO, 2020).

### 1.1.1 Delimitação do Tema

A partir da modelagem matemática de epidemias, é possível descrever a evolução do número de pessoas impactadas por uma doença de uma determinada população e em uma certa janela de tempo. Os modelos abordados no trabalho são SIS (do inglês, *Susceptible Infected Susceptible*), SIR (do inglês, *Susceptible Infected Recovered*) e SEIR (do inglês *Susceptible Exposed Infected Recovered*), sendo que este último tem maior destaque, pois descreve o comportamento de doenças mais complexas e que ainda não possuem cura, como é o caso do HIV e do Coronavírus (LACERDA, 2020).

Uma vez definido os possíveis modelos matemáticos, é possível fazer análises qualitativas e quantitativas sobre as estimativas resultantes do FP, inclusive comparando-o com o outros filtros estocásticos da literatura, como o FKE. A aplicação do FP ao modelo de epidemia é adequada, pois fornece uma estimativa do estado da epidemia, envolvendo diversos parâmetros, como o número de pessoas suscetíveis, infectadas e recuperadas, por exemplo (SKVORTSOV; RISTIC, 2012).

## 1.2 PROBLEMAS E PREMISSAS

Uma das dificuldades encontradas é modelar, da melhor maneira possível, as dinâmicas populacionais em caso de uma epidemia (ANDERSON et al., 2004). O modelo a ser escolhido deve ser complexo o suficiente para representar o comportamento de disseminação de uma epidemia de acordo com as interações sociais da população, mas simples ao ponto de possibilitar a aquisição de novos dados, conforme estes se tornam disponíveis, e uma estimação fiel nos estágios iniciais da simulação, quando poucos parâmetros estão disponíveis ou nem se quer são conhecidos.

O modelo a ser escolhido deve levar em conta três fatores importantes: o pico de amplitude da epidemia, o instante em que isso ocorre e a duração total da epidemia. Outros dados derivados podem ser obtidos através de aproximações desses três parâmetros principais (SKVORTSOV; RISTIC, 2012).

Para a implementação do código computacional, deve-se levar em conta que o FP demanda um processamento de um número elevado de partículas a cada instante do tempo da recursão. Mesmo com a tecnologia computacional disponível atualmente, dependendo de como o código for elaborado, irá requerer uma demanda razoável de processamento. Dessa forma, há a necessidade da elaboração de um código computacionalmente sucinto e eficiente, tentando otimizar o processamento, tornando a resposta do código rápida (ARULAMPALAM et al., 2002). Como auxílio para a implementação do código diversas referências da literatura poderão ajudar na implementação, como TULSYAN et al., SPEEKENBRINK e CARPENTER et al..

### 1.3 OBJETIVOS

#### 1.3.1 Objetivo Geral

Implementar computacionalmente o FP em MATLAB, de tal forma a prever e rastrear a quantidade de indivíduos suscetíveis, expostos, infectados e recuperados em uma epidemia modelada matematicamente de forma não-linear.

#### 1.3.2 Objetivos Específicos

Os objetivos específicos para este Trabalho de Conclusão de Curso (TCC) estão listados nos itens a seguir:

- Realizar a revisão bibliográfica e o estudo sobre filtros estocásticos, modelagem matemática em espaço de estados e discretização de sistemas representados em espaço de estados;
- Pesquisar outros modelos matemáticos que possam definir o comportamento dinâmico de uma epidemia;
- Modelar matematicamente epidemias capazes de reproduzir dados simulados verossímeis para possibilitar a aplicação de filtros estocásticos;
- Implementar o FKE e o FP em MATLAB;
- Utilizar os filtros estocásticos implementados com os modelos de epidemia;



- Comparar o desempenho do FP implementado com os outros filtros estocásticos implementados, e fazer uma análise comparativa dos resultados de forma qualitativa e quantitativa.

#### 1.4 JUSTIFICATIVA

Nem sempre uma população infectada busca os médicos quando se depara com os sintomas da doença contraída. Muitas vezes os sintomas nem se manifestam em certas pessoas, mas nem por isso deixa de ser infecciosa, vide os casos de Coronavírus em crianças e adolescentes (WHO, 2020). Ocasionalmente, os dados disponíveis não representam a real escala que uma epidemia tem ou pode vir a ter devido a uma serie de fatores. Uma ferramenta que seja capaz de, a partir de um conjunto de dados, prever e rastrear o progresso de uma epidemia, torna-se relevante para as autoridades de saúde (JÉGAT et al., 2008).

Como epidemias muitas vezes comportam-se de maneira não linear de disseminação (SKVORTSOV et al., 2007), é recomendada a aplicação de uma ferramenta adequada como a filtragem estocástica.

Utilizando um modelo SIS, mais simples, ou variações mais complexas como as que podem ser feitas no MOSES, torna-se possível a análise de epidemias em várias situações hipotéticas. Além disso, é possível executar esses modelos em computadores com especificações acessíveis hoje em dia sob o ponto de vista de processamento. (SKVORTSOV et al., 2007).

#### 1.5 PROCEDIMENTOS METODOLÓGICOS

Os primeiros passos tiveram como objetivo a fundamentação teórica da modelagem matemática de modelos não lineares, de processos estocásticos e filtros estocásticos, que ocorreu praticamente durante todo o processo de desenvolvimento do TCC.

Com base em referências encontradas na literatura especializada e nos modelos matemáticos que foram utilizados, o FP foi implementado computacionalmente. Além disso, o FP desenvolvido teve seus resultados analisados a partir de variações de parâmetros do filtro e dos modelos (SKVORTSOV; RISTIC, 2012). Analogamente ao FP foi implementado FKE com a finalidade de comparar os erros de suas estimativas. O *software* a ser utilizado para implementação e análise de resultados será o MATLAB.

Uma vez que os modelos foram gerados e os filtros foram aplicados, os resultados obtidos foram comparados, visando a confirmar se o FP realmente tem um desempenho melhor que o FKE. Essa análise comparativa foi feita em cima dos valores de erro obtido dos processos de filtragem.

## 1.6 ESTRUTURA DO TRABALHO

O trabalho segue com mais quatro capítulos.

- No segundo capítulo, são apresentados os possíveis modelos matemáticos de epidemias e são escolhidos dentre eles um a ser utilizado;
- O terceiro capítulo aborda os filtros estocásticos, sendo o FP o principal a ser tratado no presente trabalho;
- No quarto capítulo são apresentadas as implementações e simulações em MATLAB e os resultados obtidos da aplicação do FP e os outros filtros estocásticos da literatura;
- Por fim, no quinto capítulo, é discutido a conclusão do tema, ponderando-se acerca dos resultados obtidos e comparando-os.

## 2 MODELOS EPIDÊMICOS

### 2.1 MODELOS MATEMÁTICOS

Modelos matemáticos são uma das peças fundamentais para realizar a análise de um sistema e são usados para verificar saídas do sistema de acordo com diferentes entradas e outros fatores, como perturbações e ruídos. Dessa forma, é possível analisar as propriedades dinâmicas do sistema ou até mesmo controlá-lo. Estes modelos são classificados de acordo com os sistemas que eles modelam, podendo ser lineares ou não lineares, variantes ou invariantes no tempo, contínuos ou discretos.

Modelos em espaço de estados, discretos ou contínuos, têm como principal vantagem o fato de que algumas variáveis podem estar diretamente relacionadas com o sistema de variáveis medidas. Este tipo de representação de modelos é a utilizada para a aplicação de filtros estocásticos.

#### 2.1.1 Modelos Lineares

Os modelos lineares são a representação para sistemas cuja estrutura inerente é linear por natureza e em que o sistema obedece aos princípios da superposição. Esse princípio afirma que para todo sistema linear, a resposta obtida na saída de um sistema, causada por uma ou mais entradas, é equivalente a soma das respostas causadas por cada entrada individualmente. Na área de Controle e Automação, modelos desenvolvidos para descrever processos em linhas de produção foram aplicados com sucesso na indústria. Em comunicação e eletrônica, os modelos lineares são rotineiramente usados na análise e projeto de circuitos, sistemas de comunicação e teoria de filtros (BILLINGS, 2013).

Um exemplo de sistema linear estático é uma equação de reta com coeficiente linear nulo. Para uma entrada  $x$ , gera-se uma saída  $y$ , de acordo com a (1):

$$y = 3x. \tag{1}$$

Se  $x_1 = 3$ , então  $y_1 = 9$ ; se  $x_2 = 2$ , então  $y_2 = 6$ . Considerando um terceiro valor  $x_3 = x_1 + x_2$ , a saída  $y_3$  deve ser igual à soma  $y_1 + y_2$ . De fato,  $y_3$  satisfaz a soma  $y_1 + y_2$  para a entrada  $x_3 = x_1 + x_2$ .

### 2.1.2 Modelos Não Lineares

Apesar dos modelos lineares serem utilizados para a simplificação de estudo e análise de sistemas, a grande maioria dos sistemas do mundo real possuem comportamentos não lineares, como no ambiente industrial. Ao contrário do modelo linear, um modelo é classificado como não linear se o princípio de superposição apresentado na Seção 2.1.1 não se aplicar a ele. Assim, para um sistema não linear, não se pode obter a resposta a duas entradas simultâneas considerando as entradas individualmente e somando os resultados (OGATA; YANG, 2002). Um exemplo de sistema não-linear é a equação que segue:

$$y = x^3 + \sqrt[8]{x}, \quad (2)$$

### 2.1.3 Modelos Dinâmicos

O modelo dinâmico é composto por um conjunto de equações diferenciais, podendo ser lineares ou não lineares. Os sistemas não lineares costumam representar mais fielmente a dinâmica de sistemas reais. Normalmente, são modelos mais complexos que os lineares, em que a resposta não é proporcional a entrada do sistema.

Os sistemas dinâmicos são importantes em todos os ramos da Ciência, devido a sua capacidade de modelar fenômenos naturais de grande complexidade (CIPOLLI, 2012).

Para um exemplo de sistema dinâmico linear, pode-se citar a EDO (Equação Diferencial Ordinária) na equação (3):

$$\tau \dot{y}(t) + y(t) + Ku(t) = 0, \quad (3)$$

na qual  $\tau$  e  $K$  são constantes invariantes no tempo.

### 2.1.4 Modelos de Espaço de estados

Os modelos representados em espaço de estados podem ser classificados como a forma mais genérica de representação de sistemas dinâmicos. Uma das vantagens desse tipo de

representação é o fato das variáveis de estado poderem representar variáveis físicas no sistema em estudo (AGUIRRE, 2015).

Na equação (4), tem-se a representação em espaço de estados de um modelo linear contínuo:

$$\begin{aligned} \dot{x} &= Ax + Bu, \\ y &= Cx + Du, \end{aligned} \quad (4)$$

nas quais  $A$  é a matriz de estado,  $B$  a matriz de entrada do sistema,  $C$  a matriz de saída e  $D$  a matriz de transmissão direta<sup>1</sup>.

Para o caso de uma representação em espaço de estados discreto, tem-se as equações em (5):

$$\begin{aligned} x_{k+1} &= Gx_k + Hu_k, \\ y_k &= Cx_k + Du_k, \end{aligned} \quad (5)$$

em que, de forma similar ao conjunto de equações (4),  $G$  é a matriz de estado discreta,  $H$  é a matriz de entrada discreta,  $C$  é a matriz de saída discreta e  $D$  a matriz de transmissão direta discreta do sistema.

### 2.1.5 Modelos Estocásticos

Até o momento foram tratados modelos classificados como determinísticos. Um modelo determinístico é aquele que é representado de forma precisa e exata por meio de uma fórmula ou equação matemática ou um conjunto de equações. Isso permite assumir que o que acontece no sistema é perfeitamente conhecido e que isso está especificado sem erros e sem aleatoriedade nas fórmulas, tornando possível avaliar diretamente as fórmulas para qualquer ponto do tempo.

Um outro tipo de classificação de modelo é o estocástico, em que os estados que descrevem o modelo têm comportamentos aleatórios. Partindo do princípio de que os valores iniciais são conhecidos, encontram-se probabilidades do sistema estar definido em diferentes estados futuros possíveis. Além disso, se o processo for executado várias vezes, não serão obtidos resultados idênticos. Execuções diferentes de um processo estocástico são frequentemente chamados de realizações do processo (PARZEN, 1999).

Os modelos estocásticos são muitas vezes mais realistas quando comparados a

---

<sup>1</sup>Efeito da entrada diretamente na saída.

modelos determinísticos, por exemplo, em um caso de uma fila de supermercado. Em um modelo determinístico, apenas um número exato de clientes entram, ou saem, dessa fila num determinado intervalo de tempo. Já num modelo estocástico, o número de clientes que entram na fila e saem dela varia com o tempo de forma não exata. Existe uma incerteza em relação a esse número, sendo necessário utilizar a probabilidade desse evento ocorrer. Por esse motivo, é uma representação mais fiel da dinâmica real de uma fila de supermercado. Essa probabilidade pode estar relacionada a fatores como dia do mês, dia da semana e horário, entre outros.

Considerando-se um experimento aleatório dentro do espaço amostral  $S$  e as respectivas probabilidades desses eventos, além de seus resultados especificados como  $\zeta$ ; a cada resultado de experimento, é atribuída uma função do tempo de acordo com a seguinte representação:

$$X(t, \zeta) \quad t \in \mathbb{R}, \quad (6)$$

em que  $X$  é um experimento aleatório,  $t$  é o tempo e  $\mathbb{R}$  é o conjunto dos números reais. O gráfico da função de  $X(t, \zeta)$  por  $t$ , Figura 1, para diferentes  $\zeta$ , é chamado realização do processo estocástico contínuo. Quando é criada uma família ou conjunto de variáveis aleatórias indexadas pelos parâmetros  $t \in \mathbb{R}$ , classifica-se esse conjunto como processo aleatório ou estocástico, representado por  $X_t$ .

Para o processo estocástico discreto no tempo, substitui-se o índice  $t \in \mathbb{R}$  por  $n \in \mathbb{N}$ . Nesse caso, a família de variáveis aleatórias discretas no tempo é representada como  $X_n$  (LEON-GARCIA, 2008).

Um modelo realista para o comportamento dinâmico de um sistema normalmente leva em consideração uma estimativa inicial imperfeita e uma evolução previsível com um distúrbio que interfere nessa evolução. Esse distúrbio é chamado de ruído aleatório ou estocástico.

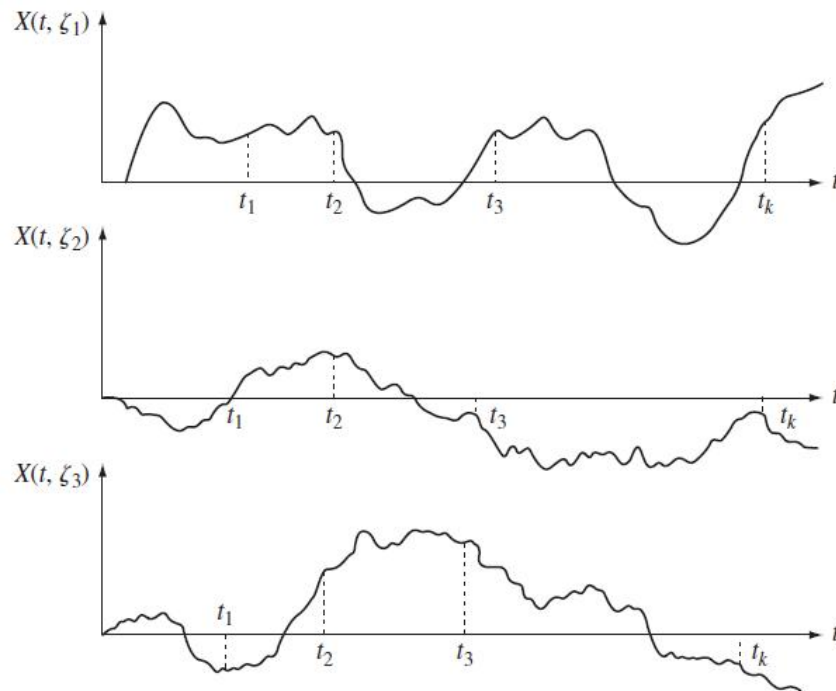
O modelo de medição para o sistema é obtido de tal maneira que apenas alguns componentes, ou uma combinação deles, são observados. Além disso, essas observações também são consideradas imprecisas, devido à presença do ruído, que modela matematicamente tal imprecisão (BAR-SHALOM et al., 2004).

O modelo de evolução desse sistema representado no espaço de estados é descrito na equação (7):

$$\begin{aligned} x_{k+1} &= f(x_k) + w_k, \\ y_k &= g(x_k) + v_k, \end{aligned} \quad (7)$$

em que  $x_k$  é o estado,  $y_k$  é a medida do estado atual,  $f(x_k)$  é uma função linear ou não linear do

**Figura 1: Várias realizações de um processo aleatório.**



**Fonte: (LEON-GARCIA, 2008).**

estado atual,  $g(x_k)$  é uma função que descreve a medição do sistema,  $w_k$  é o ruído do processo e  $v_k$  é o ruído da medição.

Os modelos estocásticos podem representar o comportamento de diversos sistemas sofisticados, como o de epidemias, os quais têm sido desenvolvidos e estudados na literatura (BOLDOG et al., 2020). Eles incorporam fatores como o número de contatos que uma pessoa infectada tem com pessoas não infectadas, as chances de que a doença seja transmitida durante um desses contatos, entre outros. Esses modelos ajudam a entender os fatores que determinam se uma epidemia ocorrerá e afetará uma grande proporção da população ou se é mais provável que afete apenas algumas pessoas e depois desapareça. Além disso, os modelos também podem ser usados para medir efeitos de intervenções, como quanto uma epidemia é afetada pela vacinação da população.

## 2.2 MODELOS MATEMÁTICOS DE EPIDEMIAS

Os modelos matemáticos de epidemias normalmente são baseados nos modelos SIS ou SIR. Em ambos os modelos, a população parte da etapa “Suscetível” e passa para a etapa “Infectado”. Porém, no modelo SIS, o indivíduo retorna à etapa “Suscetível” quando sai da etapa “Infectado”. No modelo SIR, o indivíduo passa para a etapa “Recuperado”. Este

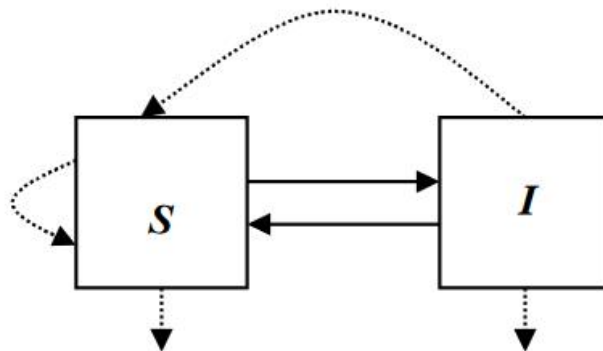
último seria um modelo fiel para doenças que o indivíduo dificilmente contrai duas vezes, como catapora e sarampo, por exemplo. Já na última etapa, no modelo SIS, o indivíduo retorna à etapa suscetível. Essa é uma representação válida nos casos de doenças que podem ser contraídas mais de uma vez. Um exemplo seriam doenças transmitidas sexualmente, como o herpes.

Ambos os modelos podem ser estendidos incluindo novas etapas, como é o caso do modelo SEQIJR<sup>2</sup> (do inglês, *Susceptible, Exposed, Quarantined, Infected, Isolated and Recovered*), que considera as etapas “Suscetível”, “Exposto”, “Em Quarentena”, “Infectado”, “Isolado” e “Recuperado”. Os seguintes desenvolvimentos matemáticos baseiam-se no apresentado em (ALLEN, 2008).

### 2.2.1 Modelo SIS

Em um modelo SIS, um indivíduo “Suscetível” pode se tornar “Infectado” e infeccioso após contato com um indivíduo infectado, mas não desenvolve imunidade. Além de não existir transmissão vertical, significando que a doença não é transmitida hereditariamente, todos os indivíduos nascem suscetíveis, não existindo morte relacionada à doença. O diagrama da Figura 2 representa a dinâmica do modelo SIS, em que  $S$  representa os indivíduos “Suscetíveis” e  $I$  os indivíduos “Infectados”.

**Figura 2: Diagrama de estados do modelo SIS. Linhas sólidas representam trocas de estado, linhas pontilhadas apontando para dentro dos compartimentos representam nascimentos e as linhas pontilhadas apontando para fora dos compartimentos representam mortes.**



**Fonte: Adaptado de (ALLEN, 2008).**

De forma geral, as equações diferenciais que regem o modelo SIS, que é um modelo matemático não linear, são:

<sup>2</sup>A letra “J” é usada para que a letra “I” não seja repetida no acrônimo, evitando confusões.



$$\begin{aligned}
\dot{S}(t) &= -\frac{\beta}{N}S(t)I(t) + (b + \gamma)I(t) \\
\dot{I}(t) &= \frac{\beta}{N}S(t)I(t) - (b + \gamma)I(t) \\
N(t) &= S(t) + I(t),
\end{aligned} \tag{8}$$

sendo que  $S(t)$  e  $I(t) \in \mathbb{N}$  e  $\beta > 0$  é a taxa de contato entre indivíduos da população,  $\gamma > 0$  é a taxa de recuperação,  $b \geq 0$  é a taxa de natalidade e  $N(t)$  é o total da população. As condições iniciais satisfazem as seguintes relações:

$$S(0) > 0, \quad I(0) > 0 \quad \text{e} \quad S(0) + I(0) = N(0). \tag{9}$$

Assumindo que a população, representada por  $N(t)$ , seja constante, para que a taxa de natalidade seja igual a de mortalidade, tem-se que  $\dot{N}(t) = 0$ . Como declarado em Hethcote (1976), as dinâmicas do modelo são regidas pelo número básico de reprodução  $\mathcal{R}_0$ . Esse número é definido da seguinte maneira:

$$\mathcal{R}_0 = \frac{\beta}{b + \gamma}. \tag{10}$$

A fração  $1/(b + \gamma)$  é a duração do período de infecção, ajustada para o número de mortes. Portanto, o modelo representado pelas equações em (8) pode ser resumido de acordo com o Teorema 1:

**Teorema 1.** Considerando que  $S(t)$  e  $I(t)$  são soluções para as equações diferenciais em (8), as seguintes implicações são verdadeiras.

$$\begin{aligned}
i) \text{ Se } \mathcal{R}_0 \leq 1, \text{ então } \lim_{t \rightarrow \infty} (S(t), I(t)) &= (N, 0) \quad (\text{equilíbrio livre de doença}), \\
ii) \text{ Se } \mathcal{R}_0 > 1, \text{ então } \lim_{t \rightarrow \infty} (S(t), I(t)) &= \left( \frac{N}{\mathcal{R}_0}, N \left( 1 - \frac{1}{\mathcal{R}_0} \right) \right) \quad (\text{equilíbrio endêmico}).
\end{aligned} \tag{11}$$

sendo “equilíbrio livre de doença” o estado em que o sistema tende a erradicação da doença, e o “equilíbrio endêmico” é quando a doença não é erradicada.

### 2.2.1.1 Discretização do modelo SIS

Para a utilização das equações diferenciais que descrevem o modelo SIS nos filtros estocásticos é necessária realizar a discretização do sistema. A discretização é efetuada utilizando o Método das Diferenças Finitas (MDF) *forward-difference* (BURDEN; FAIRES,

2008). A equação resultante da discretização é obtida em (12):

$$\frac{dx}{dt} \simeq \frac{\Delta x}{\Delta t} = \frac{x_{k+1} - x_k}{t_{k+1} - t_k}, \quad (12)$$

em que  $\Delta t$  é o tempo de amostragem. Utilizando a equação (12), foi obtida uma discretização do conjunto de equações (8), resultando no conjunto de equações (13):

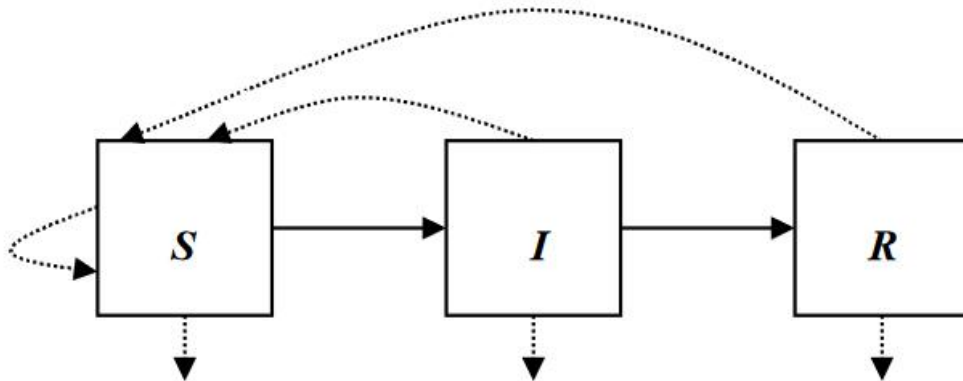
$$\begin{cases} S_{k+1} = S_k + \Delta t \left( -\frac{\beta}{N} S_k I_k + (b + \gamma) I_k \right) \\ I_{k+1} = I_k + \Delta t \left( \frac{\beta}{N} S_k I_k - (b + \gamma) I_k \right) \\ N = S_k + I_k \end{cases} \quad (13)$$

Esse modelo discretizado poderá ser utilizado nas equações dos filtros estocásticos.

### 2.2.2 Modelo SIR

No modelo SIR, os indivíduos desenvolvem imunidade a doença, portanto não voltam para a etapa “Suscetível”, mas, sim, passam para a etapa “Recuperado”. O diagrama da Figura 3 ilustra a relação entre etapas nesse modelo, em que  $S$  representa os indivíduos “Suscetíveis”,  $I$  os indivíduos “Infectados” e  $R$  representa os indivíduos “Recuperados”.

**Figura 3: Diagrama dos estados do modelo SIR. Linhas sólidas representam trocas de estado, linhas pontilhadas apontando para dentro dos compartimentos representam nascimentos e as linhas pontilhadas apontando para fora dos compartimentos representam mortes.**



Fonte: Adaptado de (ALLEN, 2008).

De forma semelhante, as equações diferenciais que regem o modelo SIR são:

$$\begin{aligned}\dot{S}(t) &= -\frac{\beta}{N}S(t)I(t) + b(I(t) + R(t)), \\ \dot{I}(t) &= \frac{\beta}{N}S(t)I(t) - (b + \gamma)I(t), \\ \dot{R}(t) &= \gamma I(t) - bR(t), \\ N(t) &= S(t) + I(t) + R(t).\end{aligned}\tag{14}$$

As condições iniciais satisfazem as seguintes relações:

$$S(0) > 0, \quad I(0) > 0, \quad R(0) > 0 \quad \text{e} \quad S(0) + I(0) + R(0) = N(0).\tag{15}$$

Assumindo que a população, representada por  $N$ , seja constante, para que a taxa de natalidade seja igual a de mortalidade, tem-se que  $\dot{N}(t) = 0$ . O número básico de reprodução em (8) e a taxa de natalidade  $b$  também regem a dinâmica do modelo em (14). Dessa forma, o modelo pode ser resumido de acordo com o Teorema 2:

**Teorema 2.** Sejam  $S(t)$ ,  $I(t)$  e  $R(t)$  soluções para o modelo em (14).

i) Se  $\mathcal{R}_0 \leq 1$ , então  $\lim_{t \rightarrow \infty} I(t) = 0$  (equilíbrio livre de doença),

ii) Se  $\mathcal{R}_0 > 1$ , então

$$\lim_{t \rightarrow \infty} (S(t), I(t), R(t)) = \left( \frac{N}{\mathcal{R}_0}, \frac{bN}{b+\gamma} \left( 1 - \frac{1}{\mathcal{R}_0} \right), \frac{\gamma N}{b+\gamma} \left( 1 - \frac{1}{\mathcal{R}_0} \right) \right) \quad (\text{equilíbrio endêmico}),$$

iii) Assumindo  $b = 0$ . Se  $\mathcal{R}_0 \frac{S(0)}{N} > 1$ , então existe um aumento inicial no número de casos de infecção  $I(t)$  (epidêmico), mas se  $\mathcal{R}_0 \frac{S(0)}{N} \leq 1$ , então  $I(t) \frac{S(0)}{N} > 1$  diminui monotonicamente a zero (equilíbrio livre de doença).

(16)

A quantidade  $\mathcal{R}_0 S(0)/N$  é chamada de número inicial de reposição. Ao longo do tempo, essa quantidade passa a ser chamada de número de reposição e representada por  $\mathcal{R}_0 S(t)/N$ . Essa quantidade representa o número de infecções secundárias, que é o número de pessoas que um único indivíduo infectado pode infectar, produzidas por um indivíduo infectado. Se esse número for menor que 1, a doença irá desaparecer da população. Porém, se for maior que 1, a população irá vivenciar uma epidemia.

### 2.2.2.1 Discretização do modelo SIR

Da mesma forma que o modelo SIS, a discretização do modelo SIR é realizada utilizando a equação (12), resultando no conjunto de equações (17):

$$\begin{cases} S_{k+1} = S_k + \Delta t \left( -\frac{\beta}{N} S_k I_k + b(I_k + R_k) \right) \\ I_{k+1} = I_k + \Delta t \left( \frac{\beta}{N} S_k I_k - (b + \gamma) I_k \right) \\ R_{k+1} = R_k + \Delta t (\gamma I_k - b R_k) \\ N = S_k + I_k + R_k. \end{cases} \quad (17)$$

Esse conjunto de equações já pode ser usado como base para os filtros estocásticos.

### 2.2.3 Modelo DTMC SIS

Para o funcionamento do MOSES (do inglês, *MATLAB-Based Open-Source Stochastic Epidemic Simulator*), um dos meios de geração utilizado nesse TCC, deve ser realizada mais uma etapa de equacionamento, descrita nesta seção e na 2.2.4. Esse equacionamento, porém, não tem impacto direto no processo de filtragem estocástica. Essa etapa adicional descreve apenas o método matemático em que o MOSES se baseia. No processo de filtragem estocástica serão usadas apenas as medidas finais geradas pelo MOSES.

O desenvolvimento dos modelos SIR e SIS se dá a partir do método Cadeia de Markov a Tempo Discreto (DTMC, do inglês *Discrete Time Markov Chain*).

Considera-se que  $\mathfrak{S}(k)$ ,  $\mathfrak{I}(k)$  e  $\mathfrak{R}(k)$  sejam processos estocásticos discretos relacionados ao número de indivíduos suscetíveis, infectados e imunes num determinado tempo  $t$ . Num modelo epidêmico DTMC,  $k \in \{0, \Delta k, 2\Delta k, \dots\}$ , que representa múltiplos dos intervalos de período de amostragem do modelo  $\Delta t$ , e as variáveis aleatórias discretas pertencem ao conjunto  $\mathfrak{S}(k), \mathfrak{I}(k), \mathfrak{R}(k) \in \{0, 1, 2, \dots, N\}$ .

Em um modelo epidêmico SIS, só existe uma variável aleatória discreta,  $\mathfrak{I}(k)$ , visto que  $\mathfrak{S}(k) = N(k) - \mathfrak{I}(k)$ , em que  $N(k)$  é constante, ou seja,  $N(k) = N$ . O processo estocástico  $\mathfrak{I}(k)$  tem uma função massa de probabilidade dada em (18):

$$p_i(t) = \mathbb{P}\{\mathfrak{I}(k) = i\}, \quad (18)$$

para  $i = 0, 1, 2, \dots, N$ , sabendo que

$$\sum_{i=0}^N p_i(t) = 1, \quad (19)$$

sendo  $p(t) = [p_0(t) \ p_1(t) \ \dots \ p_N(t)]^T$  o vetor de probabilidade associado à  $\mathfrak{I}(k)$ . O

processo estocástico  $\mathcal{J}(k)$  possui a propriedade markoviana se a igualdade abaixo é satisfeita.

$$\mathbb{P}\{\mathcal{J}(k + \Delta t) | \mathcal{J}(0), \mathcal{J}(\Delta t), \dots, \mathcal{J}(k)\} = \mathbb{P}\{\mathcal{J}(k + \Delta t) | \mathcal{J}(k)\}. \quad (20)$$

A propriedade markoviana significa que o processo, avaliado num tempo  $k + \Delta t$ , depende apenas do processo na etapa de tempo anterior  $k$ . Para completar a formulação do modelo epidêmico DTMC SIS, é necessário estabelecer a relação entre os processos estocásticos  $\mathcal{J}(k)$  e  $\mathcal{J}(k + \Delta t)$ . A probabilidade de transição de etapa é

$$p_{ji}(k + \Delta t, t) = \mathbb{P}\{\mathcal{J}(k + \Delta t) = j | \mathcal{J}(k) = i\}. \quad (21)$$

Além disso, assume-se que o número de transições efetuadas num passo de tempo  $\Delta t$  seja de, no máximo, apenas uma.

As probabilidades de transição para o modelo epidêmico DTMC satisfazem as equações em (22):

$$p_{ji}(\Delta t) = \begin{cases} \frac{\beta i(N-i)}{N} \Delta t, & j = i + 1, \\ (b + \gamma) i \Delta t, & j = i - 1, \\ 1 - \left[ \frac{\beta i(N-i)}{N} + (b + \gamma) i \right] \Delta t, & j = i, \\ 0, & \text{caso contrário,} \end{cases} \quad (22)$$

sendo  $\beta i(N-i)\Delta t/N$  a probabilidade de uma nova infecção,  $(b + \gamma) i \Delta t$  a probabilidade de morte ou recuperação e  $1 - [\beta i(N-i)/N + (b + \gamma) i] \Delta t$  a probabilidade de não haver troca de estado.

Simplificando a notação, a probabilidade de transição para uma nova infecção é representada por  $b(i)\Delta t$  e a probabilidade de morte ou recuperação por  $d(i)\Delta t$ . Então, a equação (22) passa a ser escrita como:

$$p_{ji}(\Delta t) = \begin{cases} b(i)\Delta t, & j = i + 1, \\ d(i)\Delta t, & j = i - 1, \\ 1 - [b(i) + d(i)]\Delta t, & j = i, \\ 0, & j \neq i + 1, i, i - 1. \end{cases} \quad (23)$$

Aplicando a propriedade de Markov e as propriedades de transição anteriores, as probabilidades  $p_i(k + \Delta t)$  podem ser descritas em termos das probabilidades num tempo  $t$ . No

tempo  $k + \Delta t$ , tem-se:

$$p_i(k + \Delta t) = p_{i-1}(k)b(i-1)\Delta t + p_{i+1}(k)d(i+1)\Delta t + p_i(k)(1 - [b(i) + d(i)]\Delta t), \quad (24)$$

para  $i = 1, 2, \dots, N$ , sendo  $b(i) = \beta i(N-i)/N$  e  $d(i) = (b + \gamma)i$ .

A matriz de transição  $P(\Delta t)$  é formada quando os estados são ordenados de 0 a  $N$ . Nela, o elemento  $(i, j)$  é a probabilidade de transição do estado  $i$  ao estado  $j$ , sendo  $0 \leq i \leq N$  e  $0 \leq j \leq N$ .

$$P(\Delta t) = \begin{bmatrix} 1 & d(1)\Delta t & 0 & \dots & 0 & 0 \\ 0 & 1 - (b+d)(1)\Delta t & d(2)\Delta t & \dots & 0 & 0 \\ 0 & b(1)\Delta t & 1 - (b+d)(2)\Delta t & \dots & 0 & 0 \\ 0 & 0 & b(2)\Delta t & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & d(N-1)\Delta t & 0 \\ 0 & 0 & 0 & \dots & 1 - (b+d)(N-1)\Delta t & d(N)\Delta t \\ 0 & 0 & 0 & \dots & b(N-1)\Delta t & 1 - d(N)\Delta t \end{bmatrix} \quad (25)$$

sendo que  $(b+d)(i) = [b(i) + d(i)]$ . Com isso, o processo estocástico para o modelo DTMC SIS está completamente formulado em (ALLEN, 2008).

#### 2.2.4 Modelo DTMC SIR

Como no modelo anterior, temos que  $\mathfrak{S}(k)$ ,  $\mathfrak{I}(k)$  e  $\mathfrak{R}(k)$  são processos estocásticos discretos para o número de indivíduos “Suscetíveis”, “Infectados” e “Recuperados” num determinado tempo  $k$ . Esse processo possui dois processos estocásticos independentes,  $\mathfrak{S}(k)$  e  $\mathfrak{I}(k)$  e o processo estocástico  $\mathfrak{R}(k) = N(k) - \mathfrak{S}(k) - \mathfrak{I}(k)$ , com  $N(k) = N$ . O par de processos  $\mathfrak{S}(k)$ ,  $\mathfrak{I}(k)$  tem uma probabilidade conjunta dada por:

$$p_{(s,i)}(k) = \mathbf{P}\{\mathfrak{S}(k) = s, \mathfrak{I}(k) = i\}. \quad (26)$$

Esse processo possui as propriedades markovianas. Novamente, considera-se para o modelo DTMC SIR que, no máximo, uma troca de estado ocorra por intervalo de tempo  $\Delta t$ . As

probabilidades de transição são descritas por

$$P_{(s+k,i+j),(s,i)}(\Delta t) = P\{(\Delta \mathfrak{S}, \Delta \mathfrak{J}) = P(k, j) | (\mathfrak{S}(k), \mathfrak{J}(k)) = (s, i)\}, \quad (27)$$

sendo  $\Delta \mathfrak{S} = \mathfrak{S}(k + \Delta t) - \mathfrak{S}(t)$  e  $\Delta \mathfrak{J} = \mathfrak{J}(k + \Delta t) - \mathfrak{J}(t)$ . Então,

$$P_{(s+k,i+j),(s,i)}(\Delta t) = \begin{cases} \frac{\beta is}{N\Delta t}, & (k, j) = (-1, 1) \\ \gamma i \Delta t, & (k, j) = (0, -1) \\ bi \Delta t, & (k, j) = (1, -1) \\ b(N - s - i) \Delta t, & (k, j) = (1, 0) \\ 1 - \frac{\beta is}{N\Delta t} - [\gamma i + b(N - s)] \Delta t, & (k, j) = (0, 0) \\ 0, & \text{caso contrário.} \end{cases} \quad (28)$$

Como, nesse caso, as etapas são ordenadas em pares, a matriz de transição é mais complexa para esse modelo do que para o modelo SIS. Portanto, aplicando a propriedade de Markov, a equação a diferenças satisfeita pela probabilidade  $p_{(s,i)}(t + \Delta t)$  pode ser expressa em termos de probabilidades de transição, como representado em (29):

$$\begin{aligned} p_{(s,i)}(k + \Delta t) = & p_{(s+1,i-1)}(k) \frac{\beta}{N} (i-1)(s+1) \Delta t + p_{(s,i+1)}(k) \gamma (i+1) \Delta t \\ & + p_{(s-1,i+1)}(k) b(i+1) \Delta t + p_{(s-1,i)}(k) b(N-s+1-i) \Delta t \\ & + p_{(s,i)}(t) \left( 1 - \left[ \frac{\beta}{N} is + \gamma i + b(N-s) \right] \Delta t \right). \end{aligned} \quad (29)$$

Existem ainda outros métodos que podem ser aplicados a esses modelos, como a Cadeia de Markov Tempo Contínuo (CTMC, do inglês *Continuous Time Markov Chain*) e Equação Diferencial Estocástica (SDE, do inglês *Stochastic Differential Equation*) (ALLEN, 2008).

Sob o ponto de vista da filtragem estocástica, os elementos relevantes são as medidas finais obtidas através do MOSES e os modelos inicialmente representados nas Seções 2.2.1 e 2.2.2 pelas equações (8) e (14), respectivamente discretizadas em (13) e (17).

### 2.2.5 Variações de modelos

Os modelos epidêmicos podem ser utilizados para representar diferentes doenças, como a Hepatite B (KHAN et al., 2019). A partir desses modelos, porém, pode-se desenvolver variações que se adéquem melhor a um determinado comportamento epidêmico e represente-

os com uma fidelidade maior. É o caso, por exemplo, do modelo de Djordjevic et al. (2018) para representar o comportamento da transmissão do HIV (do inglês, *Human Immunodeficiency Virus*).

Nesse modelo, não é considerada a etapa “Recuperado”. Isso se deve ao fato de, mesmo não tendo uma cura definitiva atualmente, o indivíduo que possui o HIV pode manter o vírus sob controle sem nunca desenvolver a AIDS (do inglês, *Acquired Immunodeficiency Syndrome*). Em vez da etapa “Recuperado”, utilizam-se as etapas “Crônico”, representado pela letra C, e “AIDS”, representado pela letra A. Na etapa “Crônico”, o indivíduo já está sob tratamento e com carga viral baixa, apresentando poucos ou nenhum sintoma da AIDS. Já na etapa “AIDS”, o indivíduo apresenta os sintomas da doença.

Outro exemplo de doença que pode ser melhor representada por uma variação dos modelos SIR e SIS é a dengue. O espalhamento dessa doença se dá através do mosquito *Aedes aegypti*, e o vírus pode ser encubado tanto no mosquito quanto no ser humano (MURAD et al., 2018).

Mais recentemente o modelo SEIR foi utilizado para descrever o comportamento do Coronavírus na população. Tal modelo é interessante para representar uma doença como a Covid-19, pois conta com a etapa “Exposto”, que inclui os indivíduos que foram contaminados com o vírus, mas ainda não apresentaram sintomas próprios. Essa dinâmica é muito semelhante à fase assintomática da Covid-19 (BOLDOG et al., 2020).

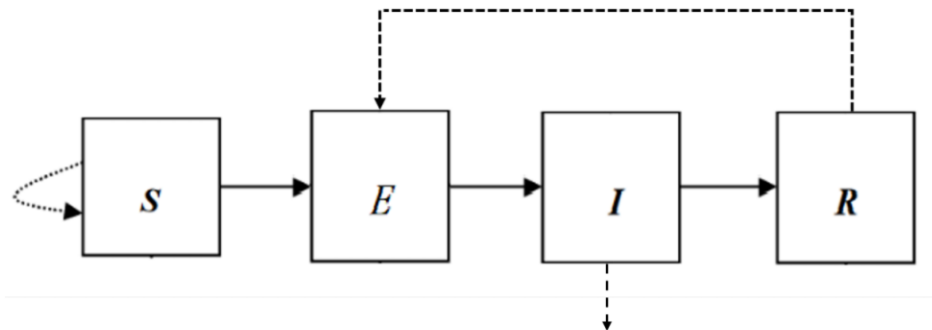
#### 2.2.6 Modelo SEIR

No modelo SEIR, assim como no modelo SIR, os indivíduos desenvolvem imunidade à doença, passando para a etapa “Recuperado”. No entanto, esse modelo também relaciona o número de indivíduos que estão infectados, mas ainda não demonstram sintomas. Essas características são observadas em doenças com um período de incubação longo.

Este atraso entre a contaminação e a etapa de “Infectado” pode ser incorporado ao modelo SIR adicionando uma população latente, classificada como etapa “Exposto”, sendo assim, uma etapa intermediária entre “Suscetível” e “Infectado”. O modelo em questão é bastante utilizado nas pesquisas de comportamento de epidemia (LACERDA, 2020).



**Figura 4: Diagrama dos estados do modelo SEIR. Linhas sólidas representam trocas de estado, linhas pontilhadas apontando para dentro dos compartimentos representam nascimentos e as linhas pontilhadas apontando para fora dos compartimentos representam mortes.**



**Fonte: Adaptado de (FANG et al., 2019).**

As equações diferenciais que regem o modelo SEIR são:

$$\begin{aligned}
 \dot{S}(t) &= -\alpha I(t)S(t) \\
 \dot{E}(t) &= \alpha I(t)S(t) - \beta E(t) \\
 \dot{I}(t) &= \beta E(t) - \gamma I(t) \\
 \dot{R}(t) &= \gamma I(t) \\
 N(t) &= S(t) + E(t) + I(t) + R(t).
 \end{aligned} \tag{30}$$

As condições iniciais satisfazem as seguintes relações:

$$S(0) > 0, \quad E(0) > 0, \quad I(0) > 0, \quad R(0) > 0 \quad \text{e} \quad S(0) + E(0) + I(0) + R(0) = N(0), \tag{31}$$

em que  $E(t)$  representa a etapa “Exposto” e  $\alpha$  é o período médio de incubação. Uma vez que a latência ou tempo de reação atrasa o início do período infeccioso do indivíduo, a disseminação de um indivíduo infectado ocorrerá mais tarde em comparação ao modelo SIR, que não tem latência. Portanto, um período de latência mais longo resultará em um crescimento inicial mais lento do surto epidêmico.

Após o rápido crescimento inicial, a epidemia esgota a população suscetível. Eventualmente, o vírus não consegue encontrar pessoas suscetíveis o suficientes e, então, morre (FOUNDATION, 2020).

As notações  $S_k, E_k, I_k, R_k$  são funções discretizadas no tempo.

$$\begin{cases} S_{k+1} = S_k - \left(1 - \frac{u}{100}\right) \left(\frac{\beta}{N}\right) S_k I_k \\ E_{k+1} = E_k + \left(1 - \frac{u}{100}\right) \left(\frac{\beta}{N}\right) S_k I_k - \alpha E_k \\ I_{k+1} = \alpha E_k - \gamma I_k \\ R_{k+1} = \gamma I_k \\ N = S_k + E_k + I_k + R_k. \end{cases} \quad (32)$$

Neste conjunto de equações em (32),  $u$  representa a taxa de isolamento da população de indivíduos<sup>3</sup> e  $\alpha$  é a taxa de incubação, que indica o número de indivíduos latentes que são infectados.

### 2.2.7 Modelo MOSES para MATLAB

Para o objetivo de aplicação, um dos modelos mais interessantes é o MOSES. O código *open source*, disponibilizado por Varol (2016), utiliza o modelo SEQIJR citado anteriormente para simulação de epidemias. Esse modelo pode ser simplificado ao zerar alguns parâmetros, podendo ser transformado num modelo mais simples, como os modelos SIR ou SIS, por exemplo. Além disso, a variação de diversos valores alteram o comportamento da simulação, permitindo modelar epidemias com diversos comportamentos diferentes.

O modelo MOSES, representado pelo fluxograma da Figura 5, é inicialmente constituído por quatro “superetapas”, sendo elas: “Suscetível”, “Exposto”, “Infectado” e “Imunizado”. Dentro dessas “superetapas”, existem etapas, que se relacionam através de algumas taxas.

Dentro da “superetapa” “Suscetível”, existem as etapas “Suscetível” e “Vacinado”. Um indivíduo passa da etapa “Suscetível” para vacinado de acordo com a taxa de vacinação ( $vr$ ) determinada. Uma vez na etapa “Vacinado”, o indivíduo passa, de acordo com a taxa de imunização de vacina ( $vir$ ) e , para a etapa de “Imunizado por vacina” ou retorna à etapa “Suscetível”. A etapa “Imunizado por vacina”, juntamente com as etapas “Imunizado Maternalmente” e “Imunizado por Recuperação”, fazem parte da “superetapa” “Imunizado”.

O indivíduo que pertence à etapa “Suscetível” pode também transitar para as “superetapas” “Exposto” ou “Infectado”, de acordo com as respectivas taxas de exposição ( $\epsilon_{exp}$ )

<sup>3</sup>Quando  $u = 100$ , o isolamento compreende toda a população.

e infecção ( $\beta_{inf}$ ). É importante notar que, para que seja modelada uma epidemia com ou sem período de exposição, uma dessas taxas deve ser zerada no início da simulação. A “superetapa” “Exposto” possui as etapas “Exposto” e “Em Quarentena”, já a “superetapa” “Infectado” possui as etapas “Infectado” e “Isolado”.

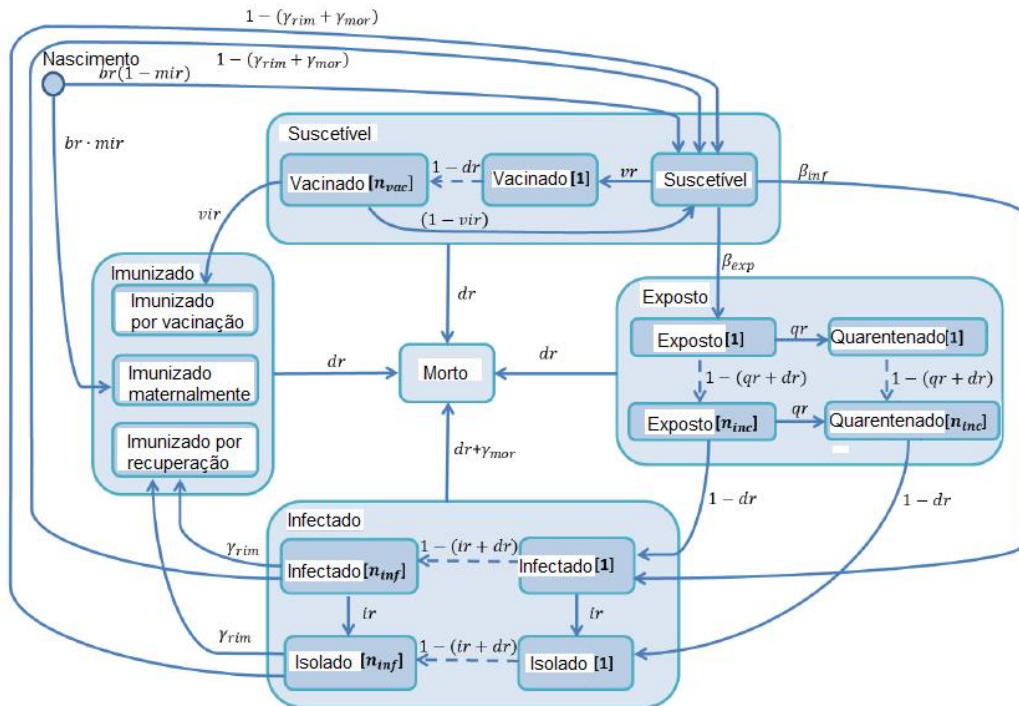
A taxa de quarentena ( $\varepsilon_{qua}$ ) rege quantos indivíduos passam da etapa “Exposto” à etapa “Em Quarentena”. Dessa forma, indivíduos que permanecem na etapa “Exposto”, após um número determinado de interações, passam à etapa “Infectado”. Por sua vez, aqueles que foram à etapa de “Em Quarentena” passam para a etapa “Isolado”. Parte desses indivíduos infectados passam para a etapa “Isolado”, de acordo com a taxa de isolamento ( $\varepsilon_{iso}$ ). Após determinado número de iterações, o indivíduo da “superetapa” “Infectado” passa para um de três etapas: “Imunizado por Recuperação”, “Suscetível” ou “Morto”. A taxa de imunizado por recuperação ( $\gamma_{rim}$ ) e a taxa de mortalidade por infecção ( $\gamma_{mor}$ ) regem essas transições.

Indivíduos das etapas “Infectado”, “Isolado”, “Em Quarentena” ou “Exposto” podem infectar um indivíduo da etapa “Suscetível”, porém não com a mesma intensidade. Um indivíduo da etapa “Infectado” tem muito mais chance de infectar alguém “Suscetível”. Tendo isso em vista, o modelo MOSES pesa as taxas de transmissão utilizando coeficientes de redução para as etapas “Isolado”, “Em Quarentena” e “Exposto”.

Vale ressaltar que esse modelo considera, além da população inicial, taxa de natalidade ( $br$ ) e taxa de mortalidade natural ( $dr$ ). O indivíduo ao nascer será da etapa “Imunizado” ou, de acordo com a taxa de imunização materna ( $mir$ ), da etapa “Imunizado Maternalmente”. Da mesma forma, a taxa de mortalidade simula a morte de indivíduos em qualquer “superetapa”, porém, a taxa de mortalidade natural da “superetapa” “Infectado” é acrescida da taxa de mortalidade por infecção (VAROL, 2016).

No programa MOSES, as seguintes abreviações tem os respectivos significados:  $br$  é a taxa de nascimento anual,  $dr$  é a taxa de mortalidade anual,  $vr$  é a taxa de vacinação diária,  $vir$  é a taxa de imunização de vacinação,  $mir$  é a taxa de imunização materna,  $\beta_{exp}$  é a constante de transmissão de “Exposto” para “Suscetível”,  $\beta_{inf}$  é a constante de transmissão de “Suscetível” para “Infectado”,  $qr$  é a taxa de quarentena diária,  $ir$  é a taxa de isolamento diária,  $\varepsilon_{exp}$  é a taxa de transmissão de “Exposto” comparado a “Infectado”,  $\varepsilon_{qua}$  é a taxa de transmissão de “Em Quarentena” comparado a “Infectado”,  $\varepsilon_{iso}$  é a taxa de transmissão de em “Isolado” comparado a “Infectado”,  $\gamma_{mor}$  é a taxa de mortalidade de “Infectado”,  $\gamma_{rim}$  é a taxa de transição de “Infectado” para “Recuperado por Imunização”,  $t_{exp}$  é o período de exposição em dias,  $t_{inf}$  é o período de infecção em dias,  $t_{vac}$  é o período de imunização por vacinação em dias.

**Figura 5: Diagrama de estados mostrando as transições entre diferentes etapas e “superetapas” e variáveis de relação no simulador MOSES. Linhas tracejadas indicam transições entre etapas iguais.**



**Fonte: Adaptado de (VAROL, 2016).**

### 2.3 SÍNTESE DO CAPÍTULO

Neste capítulo foram definidas algumas classificações de modelos matemáticos, passando pelos modelo lineares, não lineares, dinâmicos, espaço de estados e, por fim, estocásticos. Estes combinam ruídos estocásticos ao modelo de espaço de estados.

Após as definições de modelos matemáticos, tem-se as aplicações ligadas a modelagem de epidemias com as demonstrações matemáticas dos modelos SIS e SIR, bem como as suas discretizações, chegando no modelo MOSES. Por meio de ajustes em seus parâmetros, é possível representar os modelos SIS, SIR, SEIR e até mesmo o modelo SEQIJR.

A partir das definições dos modelos matemáticos e a aplicação nos modelos de epidemias, torna-se possível simular os mais diversos tipos de epidemias com comportamentos variados. É possível simular desde uma gripe que infecta intermitentemente o indivíduo, o comportamento da AIDS, que acompanha o indivíduo afetado até o fim de sua vida ou até uma doença que possui um período de incubação como a Covid-19 (WHO, 2020).

### 3 FILTROS ESTOCÁSTICOS

#### 3.1 FILTRO DE KALMAN

O Filtro de Kalman (FK) possui contribuições importantes na história da tecnologia. Um dos primeiros casos de utilização do FK foi na resolução do problema de estimação de trajetórias das naves do Projeto Apollo. O FK acabou sendo incorporado no computador de navegação da nave, colaborando para que a missão de levar o homem à Lua fosse um sucesso (MCGEE; SCHIMIDT, 1985). O FK também foi aplicado em sistemas de navegação de submarinos e nos sistemas de navegação de mísseis de cruzeiro como o BGM-109 Tomahawk e o AGM-86 ALCM das forças armadas americanas. Era também utilizado na navegação dos Ônibus Espaciais da NASA (do inglês, *National Aeronautics and Space Administration*) e está presente no controle de atitude da Estação Espacial Internacional (GAYLOR DAVID; LIGHTSEY, 2003).

O FK estima os estados de um sistema dinâmico linear discreto ou discretizado no domínio do tempo a partir de uma série de medições ruidosas (KALMAN, 1960). O FK é um estimador recursivo, ou seja, a estimativa do estado no instante de tempo anterior e a medição atual são utilizados para calcular a estimativa do estado atual.

##### 3.1.1 Modelo dinâmico linear

A modelagem dinâmica estocástica é uma proposta de análise e previsão de séries do ponto de vista estocástico. O modelo dinâmico linear híbrido, que envolve parcelas determinísticas e estocásticas, assume que o estado atual no tempo  $k$  é obtido através do estado anterior em  $k - 1$  de acordo com a equação (33):

$$x_k = F_{k-1}x_{k-1} + B_{k-1}u_{k-1} + w_{k-1}, \quad (33)$$

em que  $F_k$  é a matriz dinâmica,  $x_k \in \mathbb{R}^n$  é o vetor de estados,  $B_k \in \mathbb{R}^n$  é a matriz das entradas de

controle, aplicado no vetor de entrada de controle  $u_k$  e  $w_k$  é o ruído do processo, cuja distribuição é gaussiana multivariada de média nula e matriz de covariância  $Q_k$ , ou seja,  $w_k \sim \mathcal{N}(0, Q_k)$ .

Além de se basear no modelo dinâmico linear, o FK também precisa saber como é o modelo de medições do sistema. Quanto ao modelo de medição no tempo  $k$ , medida  $z_k$  do vetor de estados  $x_k$  possui a seguinte relação:

$$z_k = H_k x_k + v_k, \quad (34)$$

sendo que  $H_k$  é a matriz de observação, e  $v_k \in \mathbb{R}^n$  é o ruído da observação, assumido como sendo uma distribuição gaussiana multivariada de média nula e matriz de covariância  $R_k$ , ou seja,  $v_k \sim \mathcal{N}(0, R_k)$ .

O FK tem a sua estrutura baseada em duas etapas distintas: predição e atualização. A fase de predição usa a estimativa do vetor de estados do passo anterior para obter uma previsão do vetor de estados no tempo atual. Esta predição é chamada de predição *a priori*, pois ainda não inclui a informação vinda da medição do estado atual. Na fase de atualização, a predição *a priori* é atualizada para refinar a estimativa do estado. A estimativa refinada é chamada de estimativa *a posteriori*.

### 3.1.2 Predição

Como citado na Seção 3.1, a predição utiliza a estimativa do vetor de estados do passo anterior para obter a previsão do vetor de estado no tempo atual. Sendo assim, tem-se que a predição do estado é dada por:

$$\hat{x}_{k|k-1} = F_{k-1} \hat{x}_{k-1|k-1} + B_{k-1} u_{k-1}, \quad (35)$$

em que  $\hat{x}_{k|k-1}$  representa a previsão do estado e  $\hat{x}_{k-1|k-1}$  a estimativa do estado no instante  $k-1$ .

O indicativo de incerteza é dada pela matriz de covariância do erro de previsão, dada por:

$$P_{k|k-1} = F_{k-1} P_{k-1|k-1} F_{k-1}^\top + Q_{k-1}, \quad (36)$$

sendo que  $P_{k-1|k-1}$  é a matriz de covariância do erro de estimativa no instante  $k-1$ .

### 3.1.3 Atualização

Para a etapa de atualização do estado, segue-se com a inovação, a qual expressa a diferença entre o vetor de estados e o vetor de estados previsto, dada por:

$$\tilde{z}_k = x_k - \hat{x}_{k|k-1}. \quad (37)$$

Calcula-se, então, a matriz de covariância da inovação, a qual é dada por

$$S_k = H_k P_{k|k-1} H_k^T + R_k. \quad (38)$$

Para ajustar a diferença do vetor de estado predito para o vetor de estados atualizado é utilizado o ganho ótimo de Kalman, descrito pela seguinte equação:

$$K_k = P_{k|k-1} H_k^T S_k^{-1}. \quad (39)$$

Depois de encontrada a inovação e o ganho ótimo de Kalman, através das equações (37) e (39), respectivamente, faz-se a atualização do vetor de estados obtendo como resultado o vetor de estados estimado  $\hat{x}_{k|k}$ :

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + K_k \tilde{z}_k. \quad (40)$$

Analogamente ao estado estimado, com a obtenção do ganho ótimo de Kalman, é possível obter a matriz de covariância do erro de estimativa:

$$P_{k|k} = (I - K_k H_k) P_{k|k-1}, \quad (41)$$

sendo  $I$  uma matriz identidade com dimensões adequadas.

### 3.1.4 Pseudocódigo para o FK

**Dados de entrada:**  $x_k, z_k, Q_k$  e  $R_k$

**Para**  $i = 1$  até  $N$  **faça**

Fase de predição:

Calcule a predição do estado:

$$\hat{x}_{k|k-1} = F_{k-1}\hat{x}_{k-1|k-1} + B_{k-1}u_{k-1}$$

Calcule a matriz de covariância do erro de previsão:

$$P_{k|k-1} = F_{k-1}P_{k-1|k-1}F_{k-1}^T + Q_{k-1}$$

Fase de atualização:

Calcule a inovação:  $\tilde{z}_k = x_k - \hat{x}_{k|k-1}$

Calcule a matriz de covariância da inovação:  $S_k = H_k P_{k|k-1} H_k^T + R_k$

Calcule o ganho de Kalman:  $K_k = P_{k|k-1} H_k^T S_k^{-1}$

Calcule a estimativa do estado:  $\hat{x}_{k|k} = \hat{x}_{k|k-1} + K_k \tilde{z}_k$

Calcule a matriz de covariância do erro de estimativa:  $P_{k|k} = (I - K_k H_k) P_{k|k-1}$

**Fim Para**

## 3.2 FILTRO DE KALMAN ESTENDIDO

O FK é um filtro estocástico ótimo sob as hipóteses de linearidade dos modelos matemáticos e de distribuição gaussiana no ruídos de processo e de observações. Na natureza, boa parte dos sistemas costumam ser modelados por equações não lineares e, portanto, foram necessárias adaptações nas equações de filtragem para aplicar esse método de filtragem a sistemas não lineares. A maior parte deste trabalho foi realizada em uma instalação de testes e centro de pesquisas da NASA (MCELHOE, 1966). Surgiu, então, o FKE, que utiliza expansões via série de Taylor para linearizar um modelo sobre um ponto de operação.

Para a aplicação do FKE, faz-se necessária a linearização do modelo em espaço de estados. Para obter a predição do vetor estados  $\hat{x}_{k|k-1}$ , o vetor de estados  $x_k$ , da equação (41), é expandido em série de Taylor e, por sua vez, são calculadas as jacobianas do modelo dinâmico e do modelo de medidas. Uma vez que as funções são linearizadas, a mecânica de propagação do valor previsto é equivalente a do FK (WAN; MERWE, 2000).



### 3.2.1 Predição

A formulação básica do FKE é dada pelas seguintes equações de estado e de medidas:

$$\begin{aligned} x_k &= f_{k-1}(x_{k-1}) + w_{k-1}, \\ z_k &= h_k(x_k) + v_k. \end{aligned} \quad (42)$$

A função não linear  $f_k$  pode ser usada para calcular o estado previsto a partir da estimativa anterior e, da mesma forma, a função não-linear  $h_k$  pode ser usada para calcular a medida prevista a partir do estado previsto.

O estado previsto é determinado por:

$$\hat{x}_{k|k-1} = f_{k-1}(\hat{x}_{k-1|k-1}), \quad (43)$$

Após ser atualizado,  $\hat{x}_{k|k-1}$  é obtido pela avaliação da função  $f_k$  no vetor de estados estimado no instante anterior.

A matriz de covariância do erro de previsão é expressa por:

$$P_{k|k-1} = \mathcal{F}_{k-1} P_{k-1|k-1} \mathcal{F}_{k-1}^\top + Q_{k-1}, \quad (44)$$

sendo que  $\mathcal{F}_{k-1}$  é a jacobiana calculada em  $\hat{x}_{k-1|k-1}$  dada por:

$$\mathcal{F}_{x_k} = \left[ \frac{\partial f_i}{\partial x_j} \right]_{x_{k-1} = \hat{x}_{k-1|k-1}}, \quad (45)$$

o qual  $f_i$  e  $x_j$  são expressos por:

$$\vec{f} = \begin{bmatrix} f_1 \\ \vdots \\ f_i \\ \vdots \\ f_n \end{bmatrix}, \quad \vec{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_j \\ \vdots \\ x_n \end{bmatrix}. \quad (46)$$

### 3.2.2 Atualização

A matriz de covariância da inovação é obtida por

$$S_k = \mathcal{H}_{x_k} P_{k|k-1} \mathcal{H}_{x_k}^T + R_k, \quad (47)$$

em que  $\mathcal{H}_{x_k}$  é a matriz jacobiana da função  $h_k$ , calculada em :

$$\mathcal{H}_{x_k} = \left[ \frac{\partial h_i}{\partial x_j} \right]_{x=\hat{x}_{k|k-1}}, \quad (48)$$

o qual  $h_i$  e  $x_j$  são obtidos por:

$$\vec{h} = \begin{bmatrix} h_1 \\ \vdots \\ h_i \\ \vdots \\ h_n \end{bmatrix}, \quad \vec{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_j \\ \vdots \\ x_n \end{bmatrix}. \quad (49)$$

A partir das equações (43), (44) e (47), torna-se possível obter o ganho do filtro de Kalman, definido por:

$$K_k = P_{k|k-1} \mathcal{H}_{x_k}^T S_k^{-1}. \quad (50)$$

A estimativa de estado é composta assim como a equação (37), sendo que a única diferença é no cálculo da inovação ( $z_k - h_k(\hat{x}_{k|k-1})$ ):

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + K_k [z_k - h_k(\hat{x}_{k|k-1})]. \quad (51)$$

Por fim, o indicativo de incerteza da matriz de covariância do erro de estimativa é expressa por

$$P_{k|k} = P_{k|k-1} - K_k S_k K_k^T. \quad (52)$$

### 3.2.3 Pseudocódigo para o FKE

**Dados de entrada:**  $x_k, z_k, Q_k$  e  $R_k$

**Para**  $i = 1$  até  $N$  **faça**

Fase de predição:

Calcule o estado previsto:  $\hat{x}_{k|k-1} = f_{k-1}(\hat{x}_{k-1|k-1})$

Calcule a matriz de covariância do erro de previsão:

$$P_{k|k-1} = \mathcal{F}_{k-1} P_{k-1|k-1} \mathcal{F}_{k-1}^\top + Q_{k-1}$$

Fase de atualização:

Calcule a matriz de covariância da inovação:  $S_k = \mathcal{H}_{x_k} P_{k|k-1} \mathcal{H}_{x_k}^\top + R_k$

Calcule o ganho de Kalman:  $K_k = P_{k|k-1} \mathcal{H}_{x_k}^\top S_k^{-1}$

Calcule a estimativa do estado:  $\hat{x}_{k|k} = \hat{x}_{k|k-1} + K_k [z_k - h_k(\hat{x}_{k|k-1})]$

Calcule a matriz de covariância do erro de estimativa:

$$P_{k|k} = P_{k|k-1} - K_k S_k K_k^\top$$

**Fim Para**

## 3.3 FILTRO DE PARTÍCULAS

O Filtro de Partículas (FP) surgiu como uma aproximação numérica para o problema de filtragem bayesiana não linear nos anos 90 e desde então, tornou-se então popular. O FP é uma abordagem probabilística geral para estimar recursivamente uma a Função Densidade de Probabilidade (FDP) desconhecida ao longo do tempo usando medições ruidosas e um modelo matemático definido (GUSTAFSSON, 2010). O FP é capaz de fornecer uma aproximação numérica aos problemas não lineares de filtragem estocástica utilizando um método de seleção adaptativa capaz de determinar automaticamente os pontos relevantes do conjunto de medidas de um sistema (GUSTAFSSON, 2010).

O FP foi utilizado, com sucesso, em diversas áreas do conhecimento como no aprendizado de máquina ou *machine learning*, utilizando robôs com inteligência artificial para mapear e posicionar-se, em tempo real, num determinado ambiente (MONTEMERLO et al., 2002) e rastreamento de múltiplos objetos em vídeos (ISARD; BLAKE, 1998). Mais recentemente, os FP tem sido usado como modelo de cognição humana, como a forma que humanos aprendem a categorizar objetos (SANBORN et al., 2010), como detectam e fazem previsões de mudanças de padrões (BROWN; STEYVERS, 2009) e como tomam decisões (YI et al., 2009).

### 3.3.1 Integração de Monte Carlo

Os FP são baseados na Amostragem Sequencial por Importância (ASI), que consiste em representar FDP *a posteriori* a partir de um conjunto de amostras aleatórias, associadas a certos pesos denominados pesos por importância (DOUCET et al., 2001).

O fundamento por trás do FP é a integração de Monte Carlo, que torna possível a criação das amostras distribuídas. A integral multidimensional é representada por:

$$I = \int f(x)\pi(x)dx, \quad (53)$$

e a média das amostras  $f(x^i)$  é obtida através da aproximação da integral em (53). A equação (53) pode ser interpretada como o valor esperado de  $f(x)$ , sendo  $\pi(x)$  a FDP de  $x$ . Dessa forma, a equação (53) pode ser aproximada pela média amostral, dada por:

$$I_N = \frac{1}{N} \sum_{i=1}^N f(x^i). \quad (54)$$

No contexto da estimação bayesiana, normalmente não é possível aplicar, de forma eficiente, a amostragem para a distribuição *a posteriori*, representada por  $\pi(x)$ . A função de distribuição *a posteriori* é uma combinação da distribuição *a priori* e da FDP, que informa quais são as informações contidas nos dados observados. Para contornar o problema, uma possível solução é utilizar o método de amostragem por importância ao invés da função de distribuição *a posteriori*. A Amostragem por importância é um método geral de integração de Monte Carlo para realizar a filtragem não linear.

Idealmente, deve-se gerar amostras através da FDP *a posteriori*  $\pi(x)$ , porém, devido aos problemas citados anteriormente, as amostras devem ser geradas a partir de uma FDP  $q(x)$ , similar a  $\pi(x)$ . Esta FDP  $q(x)$  é conhecida como densidade por importância, pode ser expressa pela seguinte inferência:

$$\pi(x) > 0 \Rightarrow q(x) > 0 \quad \text{para todo } x \in \mathbb{R}^{n_x}. \quad (55)$$

Sendo assim, por meio desta relação, a integral de Monte Carlo na equação (53) pode ser reescrita da seguinte forma:

$$I = \int f(x)\pi(x)dx = \int f(x) \frac{\pi(x)}{q(x)} q(x)dx. \quad (56)$$

A estimativa da integral  $I$  por Monte Carlo é computada ao gerar  $N \geq 1$  amostras independentes, distribuídas de acordo com  $q(x)$  formando a soma ponderada  $I_N$ :

$$I_N = \frac{1}{N} \sum_{i=1}^N f(x^i) \bar{w}(x^i), \quad (57)$$

na qual o peso de importância é definido por:

$$\bar{w}(x^i) = \frac{\pi(x^i)}{q(x^i)}. \quad (58)$$

**Prova:** Continuando o processo, a integral de Monte Carlo deve ser estimada considerando a função  $\pi'(x) = c\pi(x)$ , com  $c$  desconhecido. Dessa forma tem-se que a estimativa  $\hat{I}$  da integral de Monte Carlo é:

$$\hat{I} = \int f(x) \pi'(x) dx.$$

Multiplicando e dividindo por  $q(x)$ , tem-se  $\int f(x) \frac{\pi'(x)}{q(x)} q(x) dx$ ,  
pode-se aproximar então por  $\frac{1}{N} \sum_{i=1}^N f(x^i) \frac{\pi'(x^i)}{q(x^i)} = c \frac{1}{N} \sum_{i=1}^N f(x^i) \frac{\pi(x^i)}{q(x^i)}$ .

Portanto:

$$\hat{I} = cI_N.$$

Como o fator de normalização da FDP  $\pi(x)$  é desconhecido, é necessário executar uma normalização dos pesos por importância. Sem perda de generalidade, considera-se  $f(x) = 1$ . Dessa forma, tem-se que:

$$I = \int 1 \pi'(x) dx \cong \frac{1}{N} \sum_{i=1}^N 1 \frac{\pi'(x^i)}{q(x^i)} \cong c \frac{1}{N} \sum_{i=1}^N 1 \frac{\pi(x^i)}{q(x^i)}.$$

Lembrando-que  $\pi(x)$  é uma FDP, ou seja, satisfaz a seguinte igualdade:

$$\int \pi(x) dx = 1,$$

tem-se que:

$$\frac{1}{N} \sum_{i=1}^N \frac{\pi(x^i)}{q(x^i)} = 1 \text{ e } \frac{1}{N} \sum_{i=1}^N \frac{\pi(x^i)}{q(x^i)} = c.$$

Dessa forma, dividindo  $\hat{I}$  por  $c$ , é possível normalizar tal valor:

$$\hat{I}_N = \frac{\frac{1}{N} \sum_{i=1}^N f(x^i) \frac{\pi(x^i)}{q(x^i)}}{\frac{1}{N} \sum_{j=1}^N \frac{\pi(x^j)}{q(x^j)}} = \frac{c \frac{1}{N} \sum_{i=1}^N f(x^i) \frac{\pi(x^i)}{q(x^i)}}{c \frac{1}{N} \sum_{j=1}^N \frac{\pi(x^j)}{q(x^j)}} = \frac{\sum_{i=1}^N f(x^i) \bar{w}(x^i)}{\sum_{j=1}^N \bar{w}(x^j)}.$$

Por fim, chega-se no seguinte valor de  $I_N$ :

$$I_N = \sum_{i=1}^N f(x^i) w(x^i),$$

no qual a normalização dos pesos por importância é dado por:

$$w(x^i) = \frac{\bar{w}(x^i)}{\sum_{j=1}^N \bar{w}(x^j)}.$$

### 3.3.2 Amostragem Sequencial Por Importância

O método de ASI é um método computacional que consiste na implementação recursiva do filtro bayesiano pela simulações de Monte Carlo, representando FDP *a posteriori* requerida por meio de amostras aleatórias com associação de pesos, estimada com base nessas mesmas amostras e pesos.

Quando o número de amostras torna-se muito elevado, essa caracterização torna-se uma representação equivalente da FDP *a posteriori* e o filtro de ASI se aproxima do estimador ótimo bayesiano (GORDON et al., 2004).

Antes do desenvolvimento das equações referentes ao algoritmo do filtro, é necessário definir algumas variáveis e representações. Primeiramente, tem-se  $X_k = \{x_j, j = 0, \dots, k\}$  representando todos os estados  $x_j$  até o tempo  $k$ . O conjunto de FDPs *a posteriori* no tempo  $k$  é representada como  $p(X_k|Z_k)$  e sua FDP marginal é  $p(x_k|Z_k)$ ;  $\{X_k^i, w_k^i\}_{i=1}^N$  representa as medidas aleatórias caracterizadas pela densidade de probabilidade conjunta *a posteriori*  $p(X_k|Z_k)$ . Os pesos são normalizados de modo que a soma desses pesos é unitária. Então, a FDP *a posteriori* no tempo  $k$  pode ser aproximada da seguinte forma:

$$p(X_k|Z_k) \approx \sum_{i=1}^N w_k^i \delta(X_k - X_k^i), \quad (59)$$

em que  $\delta$  é a função delta de Dirac.

Os pesos normalizados  $w_k^i$  são escolhidos usando o princípio da amostragem por importância. Além disso, se as amostras  $X_k^i$  forem adquiridas através da FDP  $q(X_k|Z_k)$ , a seguinte proporcionalidade é válida:

$$w_k^i \propto \frac{p(X_k^i|Z_k)}{q(X_k^i|Z_k)}. \quad (60)$$

Para o tempo  $k - 1$  temos as amostras se aproximando de  $p(X_{k-1}|Z_{k-1})$ . Após a aquisição de medida  $z_k$  no tempo  $k$ , deseja-se aproximar  $p(X_k|Z_k)$  do novo conjunto de amostras. Se a FDP for escolhida para fatorar de modo que a igualdade:

$$q(X_k|Z_k) = q(x_k|X_{k-1}, Z_k) q(X_{k-1}|Z_{k-1}), \quad (61)$$

é válida, então é possível obter as amostras  $X_k^i \sim q(X_k|Z_k)$ , aumentando cada uma das amostras  $X_{k-1}^i \sim q(X_{k-1}|Z_{k-1})$  com o novo estado  $x_k^i \sim q(x_k|X_{k-1}, Z_k)$ . Para obter a equação de atualização dos pesos por importância, a FDP  $p(X_k|Z_k)$  é expressa em termos de  $p(X_{k-1}|Z_{k-1})$ ,  $p(z_k|x_k)$  e  $p(x_k|x_{k-1})$ :

$$p(X_k|Z_k) = \frac{p(z_k|x_k) p(x_k|x_{k-1})}{p(z_k|Z_{k-1})} p(X_{k-1}|Z_{k-1}), \quad (62)$$

sendo assim possível chegar-se na equação de atualização do peso por importância em (63):

$$w_k^i \propto w_{k-1}^i \frac{p(z_k|x_k^i) p(x_k^i|x_{k-1}^i)}{q(x_k^i|X_{k-1}^i, Z_k)}. \quad (63)$$

Além disso, através da propriedade markoviana, a densidade por importância depende apenas da amostra  $x_{k-1}$  e da medida  $z_k$ . Nesta situação, é necessário armazenar apenas  $x_k^i$ , tornando possível o descarte de  $X_{k-1}$  e do histórico de observação  $Z_{k-1}$ . Sendo assim, a densidade *a posteriori* filtrada  $p(x_k|Z_k)$  é aproximada pelo trem de impulsos:

$$p(x_k|Z_k) \approx \sum_{i=1}^N w_k^i \delta(x_k - x_k^i). \quad (64)$$

Assim, a filtragem via ASI consiste na propagação recursiva dos pesos por importância  $w_k^i$  e dos estados  $x_k^i$  conforme cada medição é recebida sequencialmente.

### 3.3.3 Problema da Degeneração

A degeneração do algoritmo do FP é inevitável e ocorre na etapa de ASI, como provado pelo teorema de Kong-Liu-Wong (KONG et al., 1994). Na prática isso significa que, após a evolução dinâmica das partículas em um ciclo, todas as partículas, com exceção de uma, terão os respectivos pesos por importância nulos.

Para limitar o efeito dessa degeneração deve ser escolhida uma função que minimiza a variância dos pesos por importância (DOUCET et al., 2000). Um bom indicador da degeneração de um algoritmo é o tamanho efetivo da amostra  $\hat{N}_{eff}$ , dado por:

$$\hat{N}_{eff} = \frac{1}{\sum_{i=1}^N (w_k^i)^2} . \quad (65)$$

Da equação (65), obtém-se diretamente duas condições:

(i) Se os pesos forem uniformes, isto é,  $w_k^i = 1/N$ , então  $\hat{N}_{eff} = N$ .

(ii) Se  $j \in \{1, \dots, N\}$  de modo que  $w_k^j = 1$ , e  $w_k^i = 0$  para todo  $i \neq j$ , então  $\hat{N}_{eff} = 1$ .

Portanto, quando  $\hat{N}_{eff}$  assume um valor baixo, há um indício de uma degeneração severa e, quando  $\hat{N}_{eff}$  assume um valor elevado, o algoritmo do FP apresenta uma baixa degeneração (GORDON et al., 2004).

### 3.3.4 Reamostragem Sequencial Por Importância

O processo de reamostragem deve ser considerado quando ocorre quando uma degeneração significativa no algoritmo do FP. A reamostragem tem o objetivo de eliminar amostras de baixa importâncias e multiplica amostras de alta importâncias. Isso envolve mapeamento de pesos uniformes  $\{x_k^i, w_k^i\}$ :  $\{x_k^i, 1/N\}$ . As novas amostras  $\{x_k^{i*}\}_{i=1}^N$  são geradas a partir de reamostragem de uma representação discreta aproximada de  $p(x_k|Z_k)$  dado por

$$p(x_k|Z_k) \approx \sum_{i=1}^N w_k^i \delta(x_k - x_k^i), \quad (66)$$

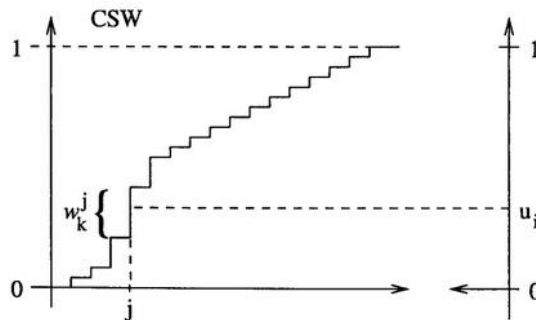
sendo que  $P\{x_k^{i*} = x_k^j\} = w_k^j$  e os novos pesos das partículas são uniformes. A seleção  $x_k^{i*} = x_k^j$  é baseada na soma acumulada ilustrada na Figura 6.

Na Figura 6,  $u_i$ , que é uma variável aleatória uniforme no intervalo  $[0,1]$ , é mapeado no índice  $j$ ; a partícula correspondente  $x_k^j$  tem uma chance maior de ser selecionada devido ao



seu alto valor  $w_k^j$ , representado pela altura do degrau do gráfico da soma acumulada. Ou seja, essa amostra será replicada, devido a sua alta importância, ao passo que outras amostras menos significativas serão descartadas.

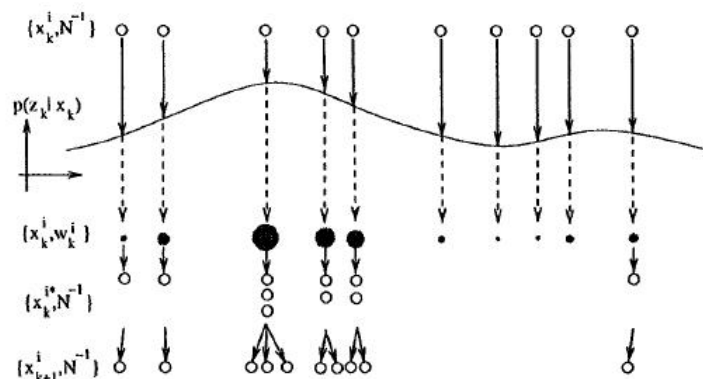
**Figura 6: Representação dos sorteios e associação com a soma acumulada.**



**Fonte: (GORDON et al., 2004).**

Uma representação gráfica de um FP de  $N=10$  amostras que utiliza como função de importância  $q(x_k|x_{k-1}^i, z_k) = p(x_k|x_{k-1}^i)$  é ilustrado na Figura 7. Inicialmente, no topo, tem-se um valor aleatório com peso uniforme  $\{x_k^i, N^{-1}\}$ , que aproxima a FDP  $p(x_k|Z_{k-1}^i)$ . Então, é computado, para cada partícula, seu peso de importância a partir do recebimento da medida  $z_k$ . O resultado é o par  $\{x_k^i, w_k^i\}$ , que gera uma aproximação de  $p(x_k|Z_k)$ . Se uma degeneração significativa ocorrer, a etapa de reamostragem é executada, em que as partículas de maior importância são selecionadas para obter a medida de pesos uniformes  $\{x_k^{i*}, N^{-1}\}$  que, por sua vez, aproxima  $p(x_k|Z_k)$ . A última etapa é a previsão, que introduz a variedade advinda do ruído do processo e resulta no par  $\{x_{k+1}^i, N^{-1}\}$ , que aproxima  $p(x_{k+1}|Z_k)$ .

**Figura 7: Um ciclo de um Filtro de Partículas.**



**Fonte: (GORDON et al., 2004).**

Apesar de reduzir os efeitos de degeneração, esse processo pode gerar algumas complicações práticas. Devido as características da reamostragem, as partículas de alta importância são estatisticamente escolhidas mais vezes. Isso gera uma perda de diversidade

de amostras, já que as mesmas amostras irão conter muitas partículas repetidas. Esse problema é conhecido como empobrecimento de amostra, sendo mais grave em casos em que o ruído do processo é baixo em relação ao sinal. Isso faz com que todas as partículas se aproximem de um mesmo ponto em poucas iterações.

De forma geral, a precisão de qualquer estimativa de uma função de distribuição só pode ser reduzida como resultado de reamostragem. Portanto, a média e a covariância das amostras devem ser calculadas antes do processo de reamostragem (GORDON et al., 2004).

### 3.3.5 Pseudocódigo para o FP

**Dados de entrada:**  $\{\mathbf{x}_{k-1}^i, w_{k-1}^i\}$  e  $\mathbf{z}_k$

**Para**  $i = 1$  até  $N$  **faça**

Sortear  $\mathbf{x}_k^i$  com base na distribuição por importância:  $\mathbf{x}_k^i \sim q(\mathbf{x}_k | \mathbf{x}_{k-1}, \mathbf{z}_k)$

Avalie os pesos por importância:  $\bar{w}_k^i = \frac{p(\mathbf{z}_k | \mathbf{x}_k^i) p(\mathbf{x}_k^i | \mathbf{x}_{k-1}^i)}{q(\mathbf{x}_k^i | \mathbf{x}_{k-1}^i, \mathbf{z}_k)} w_{k-1}^i$

**Fim Para**

Calcule o peso total:  $S = \sum_{i=1}^N \bar{w}_k^i$

**Para**  $i = 1$  até  $N$  **faça**

Normalizar os pesos:  $w_k^i = \bar{w}_k^i / S$

**Fim Para**

## 3.4 SÍNTESE DO CAPÍTULO

Neste capítulo foram apresentados o FK, o FKE e o FP. Como exemplo de filtro para aplicações em modelos lineares há, o FK, o qual é capaz de estimar grandezas do sistema modelado com o objetivo de obter estimativas o mais próximo possível dos valores exatos (sem ruído). Isso é obtido através cálculos de valores esperados e matrizes de covariâncias, ambos condicionados às medidas.

Para modelos matemáticos não lineares uma das escolhas é o FKE que, através de linearizações por séries de Taylor, torna possível utilizar as equações do FK, ainda que adaptadas.

Por fim, o FP consegue tratar os modelos matemáticos não lineares através das amostras de partículas e seus respectivos pesos por importância. Isso torna possível a sua aplicação em qualquer sistema com medidas ruidosas, apresentando distribuições gaussianas ou não.

## 4 SIMULAÇÕES E RESULTADOS

Para implementar computacionalmente o processo de filtragem estocástica, foi necessário, inicialmente, descrever os modelos que regem as dinâmicas de epidemias e gerar as medidas relacionadas ao número de indivíduos numa população. Isso foi possível através dos equacionamentos dos modelos SIS, SIR e SEIR, descritos nas Seções 2.2.1, 2.2.2 e 2.2.6 respectivamente, e a geração dos dados populacionais equivalentes utilizando dois métodos distintos, o MOSES e a GS (Geração Sintética). Além disso, foi necessário “contaminar” esses dados com ruídos aleatórios, criando, assim, as medidas simuladas para serem aplicadas aos filtros estocásticos. Para a GS, os dados foram gerados a partir do próprio modelo, ou seja, o modelo que descreve a dinâmica do sistema é exatamente o mesmo utilizado nos filtros.

Para adquirir os resultados, realizaram-se simulações no MOSES para obter as medidas referentes aos modelos SIS e SIR. Geraram-se as medidas no MOSES e, em seguida, adicionou-se ruído aleatório a esses valores para simular medidas imprecisas. Com a geração sintética de dados foi possível utilizar os modelos descritos anteriormente, aplicando o ruído aos valores da mesma forma que a geração pelo MOSES. Por fim, essas medidas ruidosas passaram pelos filtros FKE e um FP, ambos implementados em MATLAB.

### 4.1 SIMULAÇÃO DO MODELO MOSES

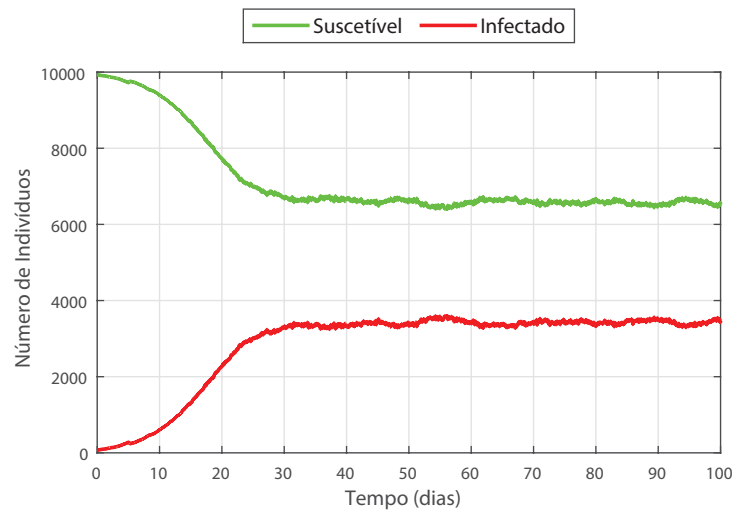
Nesta seção serão apresentadas duas simulações realizadas no MATLAB utilizando o modelo MOSES. O cenário 1 baseia-se no modelo SIS e o cenário 2 baseia-se no modelo SIR. Os dados gerados foram representados por gráficos que permitem uma análise visual da dinâmica do modelo, sendo possível observar a evolução da epidemia por um número pré-definido de indivíduos ao longo dos dias. A simulação foi efetuada realizando uma amostragem por hora durante 100 dias. Os parâmetros utilizados na simulação encontram-se na Tabela 2.

**Tabela 1: Tabela de parâmetros das simulações.**

<b>Parâmetro</b>	<b>Cenário 1</b>	<b>Cenário 2</b>
Taxa de nascimento anual $br$	0	0
Taxa de mortalidade anual $dr$	0	0
Taxa de vacinação diária $vr$	0	0,05
Taxa de imunização de vacinação $vir$	0	0,9
Taxa de imunização materna $mir$	0	0
Constante de transmissão de Exposto para Suscetível $\beta_{exp}$	0	0
Taxa de quarentena diária $qr$	0	0
Constante de transmissão de Suscetível para Infectado $\beta_{inf}$	0,3	0,6
Taxa de isolamento diária $ir$	0	0
Taxa de transmissão de Exposto comparado a Infectado $\epsilon_{exp}$	0	0
Taxa de transmissão de Em Quarentena comparado a Infectado $\epsilon_{qua}$	0	0
Taxa de transmissão de Isolado comparado a Infectado $\epsilon_{iso}$	0	0
Taxa de mortalidade dos Infectados $\gamma_{mor}$	0	0
Taxa de transição de Infectado para Recuperado por Imunização $\gamma_{rim}$	0	0,5
Período de exposição em dias $t_{exp}$	0	0
Período de infecção em dias $t_{inf}$	5	4
Período de imunização por vacinação em dias $t_{vac}$	6	3
População inicial Suscetíveis	9930	9910
População inicial Expostos	0	0
População inicial Infectados	70	70

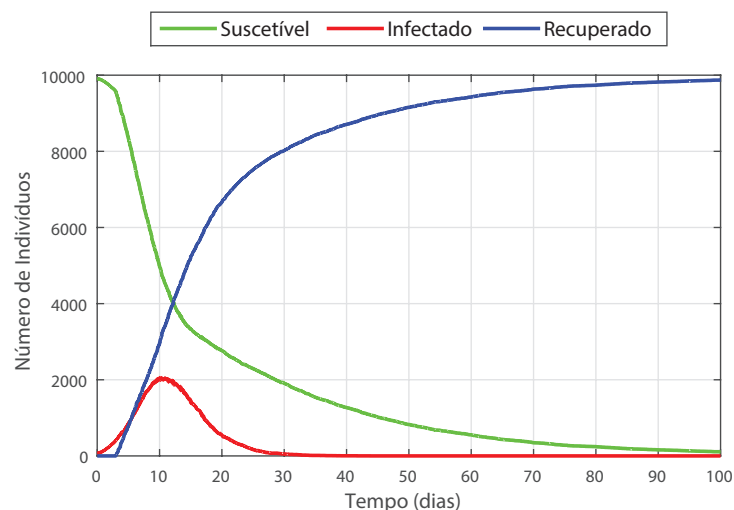
**Fonte: Adaptado de (VAROL, 2016).**

No cenário 1 foi realizada a simulação apenas com indivíduos suscetíveis a uma certa doença com constante de transmissão de suscetível para infectado de 0,3 (30% de probabilidade). Com isso, é possível observar, na Figura 8, que a população inicial de 10000 indivíduos, após o 30º dia, desenvolveu uma certa resistência à doença.

**Figura 8: Gráfico do resultado do cenário 1.****Fonte: Autoria própria.**

No cenário 2 foi acrescentado um novo parâmetro que é significativo considerando uma epidemia: o parâmetro de vacinação. Considerando este parâmetro, surgem os indivíduos imunizados. A taxa de imunização por vacinação é de 90%, ou seja, após a imunização este indivíduo não contrai novamente a doença, reduzindo a probabilidade de morte pela infecção para 0%. Este parâmetro, gerou ao fim do período, uma população 100% imunizada, erradicando a epidemia, vide Figura 9.

Ainda no cenário 2, é possível notar que de acordo com o parâmetro  $t_{vac}$ , tempo para a imunização por vacinação, ocorre a redução total do número de infectados e a imunização de toda a população restante que não foi morta pela infecção.

**Figura 9: Gráfico do resultado do cenário 2.****Fonte: Autoria própria.**

## 4.2 APLICAÇÃO DO FKE — MODELO DE EPIDEMIA SIS E DADOS DO MODELO MOSES

De acordo com o que foi descrito na Seção 3.2 sobre o FKE, implementou-se, em MATLAB, o algoritmo do filtro estocástico, baseado nas equações (42)–(52).

Na aplicação do filtro FKE com modelo SIS, na Seção 2.2.1, utilizou-se o conjunto de equações em (13). Este conjunto descreve um modelo determinístico exato da população, não possuindo erros que poderiam ser originados, por exemplo, de sistemas de medições reais. Por este motivo, foi necessário “contaminar” os estados com ruídos aleatórios, com distribuição gaussiana de média nula e variância conhecida, para que fosse possível a utilização do filtro estocástico.

Os parâmetros utilizados para chegar ao resultado mostrado na Figura 10, com um maior detalhamento na Figura 11, foram os seguintes:

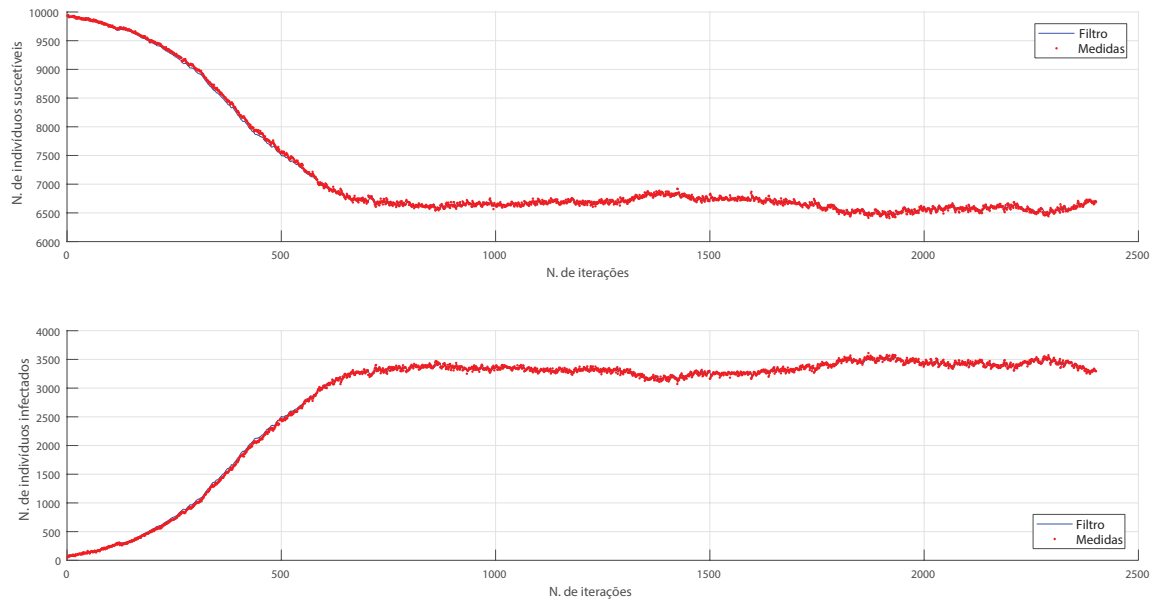
$$Q = \begin{bmatrix} 20 & 0 \\ 0 & 20 \end{bmatrix}, \quad R = \begin{bmatrix} 100 & 0 \\ 0 & 100 \end{bmatrix}, \quad \hat{x}_{0|0} = \begin{bmatrix} 9930 \\ 70 \end{bmatrix}, \quad P_{0|0} = \begin{bmatrix} 99 & 0 \\ 0 & 50 \end{bmatrix}, \quad (67)$$

em que  $Q$  representa a matriz de covariância do ruído do processo,  $R$  representa a matriz covariância do ruído da observação (definidas empiricamente),  $\hat{x}_{0|0}$  e  $P_{0|0}$  as condições iniciais de estado e a matriz de covariância do erro de estimativa, respectivamente. Além disso, o parâmetro  $b$ , que representa a taxa de natalidade, é nulo de maneira que não haja variações no número total de indivíduos da população.

Aumentando os valores da matriz  $Q$ , observou-se que o FKE tende a perder confiança no modelo dinâmico, aproximando-se, portanto, das medidas. Para corrigir este problema, ajustou-se um valor de equilíbrio entre os parâmetros  $\beta$  e  $\gamma$  do modelo dinâmico e os valores da matriz  $Q$ . A matriz de covariância do ruído de observação  $R$  é preenchida com as variâncias das medidas em sua diagonal principal. Por fim, os valores escolhidos para as condições iniciais  $\hat{x}_{0|0}$  e  $P_{0|0}$  foram próximos à entrada no modelo MOSES, porém foi possível observar que estes valores não tiveram tanta influência no comportamento do FKE, visto que a correção do filtro ocorre rapidamente, evitando a propagação do erro.

O modelo MOSES foi utilizado para a geração dos valores de medidas. Dessa forma, as medidas “contaminadas” foram geradas através dos dados exatos do cenário 1, conforme os parâmetros na Tabela 2. O ruído, que contaminou os valores reais para a simulação das medidas, segue a distribuição gaussiana  $v_k \sim \mathcal{N}(0, 10)$ .

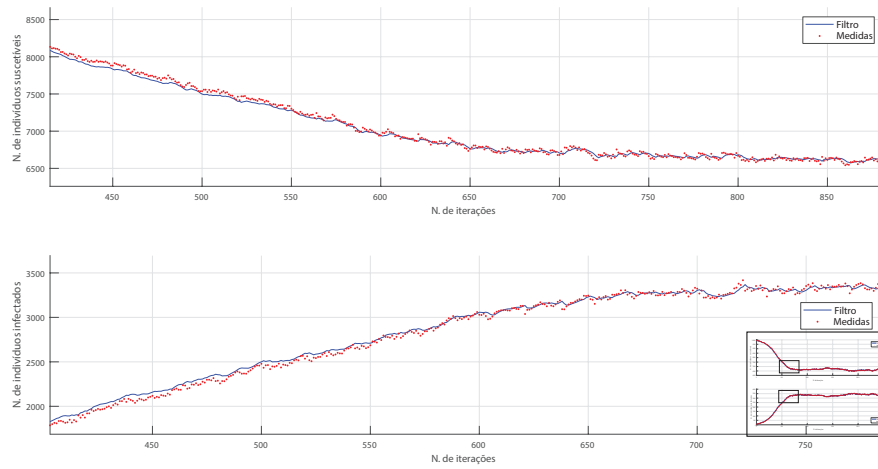
**Figura 10: Gráfico da aplicação do FKE ao modelo SIS (Cenário 1).**



**Fonte: Autoria própria.**

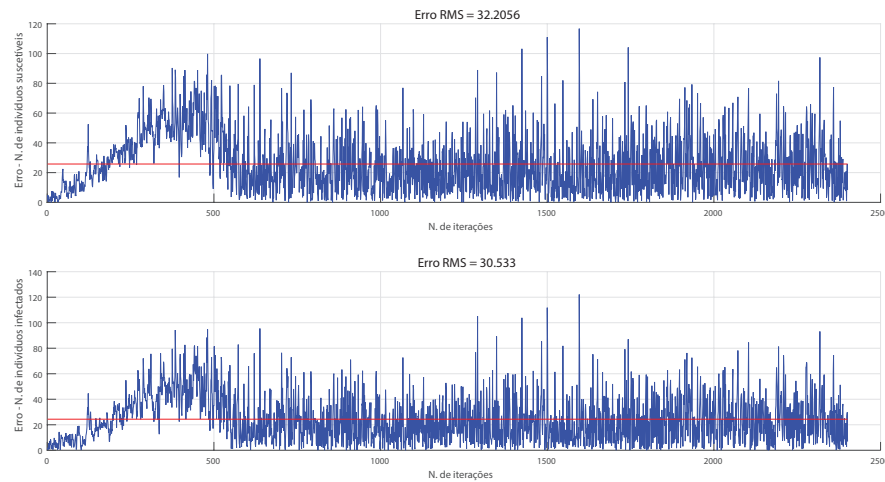
O erro foi calculado pelo valor absoluto da diferença entre os valores reais e os valores filtrados. Observa-se na Figura 12 que o erro é elevado na etapa de transição do gráfico, que pode ser observada na Figura 10, entre as iterações 1 e 750. Fica evidente nas Figura 11 o distanciamento entre os valores medidos e os valores estimados pelo FKE durante a transição. Essa é a região em que a dinâmica populacional é mais ativa, ou seja, existe um volume consideráveis de transições entre etapas a cada iteração. Quando a curva atinge algo semelhante a uma espécie de regime permanente, o erro de “regime permanente” diminui. Essa etapa representa o período em que a epidemia entra em equilíbrio entre o número de indivíduos suscetíveis e infectados, ou seja, as transições de etapas são mínimas ou nulas.

**Figura 11: Zoom nas etapas de transição da aplicação do FKE ao modelo SIS (Cenário 1) dos números de indivíduos de cada população.**



**Fonte: Autoria própria.**

**Figura 12: RMSE, erro absoluto e médio do FKE em relação aos valores reais.**



**Fonte: Autoria própria.**



### 4.3 APLICAÇÃO DO FP — MODELO DE EPIDEMIA SIS E OS DADOS DO MODELO MOSES

Com o FP, foram utilizadas as equações (53)–(66), descritas na Seção 3.3 sobre FP. O modelo SIS utilizado foi o mesmo utilizado na Seção 4.2 para a aplicação do FKE. Assim como na Seção anterior, o modelo MOSES foi utilizado para gerar os dados simulados. De forma semelhante, o ruído gaussiano adicionado para “contaminar” os dados segue a mesma distribuição da seção anterior.

Novamente, os filtros apresentaram comportamentos esperados, com o FP tendo um desempenho melhor do que o FKE. Em relação ao FKE, verifica-se que o FP demonstrou um comportamento melhor através de todo o processo de filtragem, não apresentando um aumento do RMSE nas regiões de transição. Essa diferença pode ser percebida quando compara-se as Figuras 12 e 15 ou as Figuras 11 e 14.

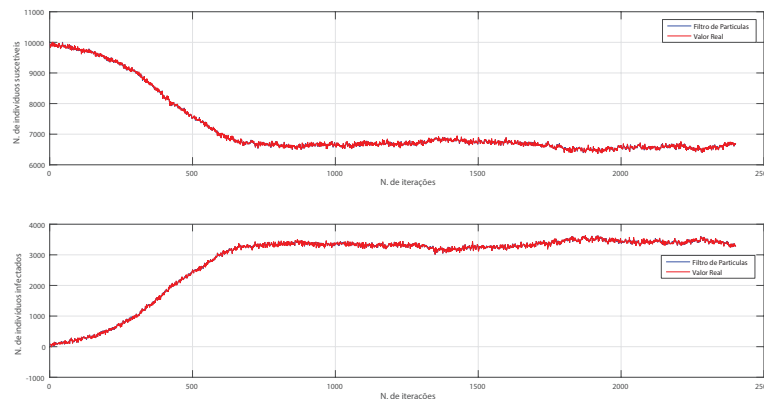
**Tabela 2: Tabela do erro em número de indivíduos do modelo SIS gerado pelo MOSES.**

Filtro	RMSE Suscetível	RMSE Infectado
FKE	32,2056	30,533
FP	15,5277	16,6222

**Fonte: Autoria própria.**

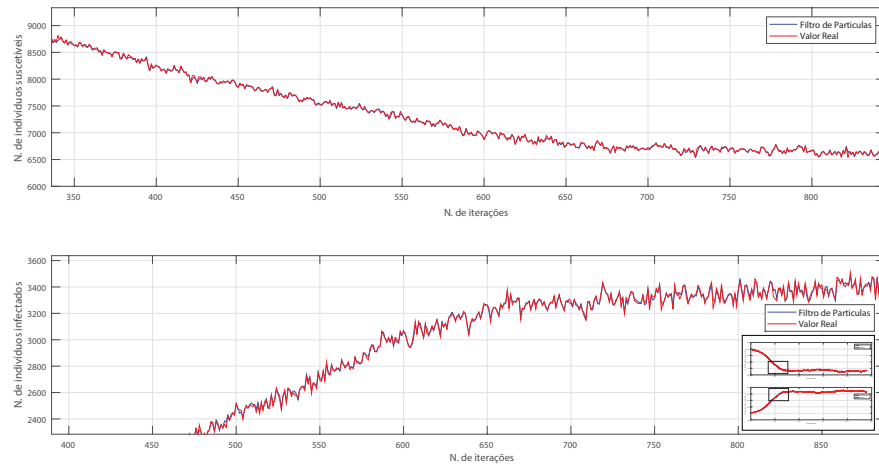
A Tabela 2 mostra o RMSE da quantidade de indivíduos em cada etapa. Nela, observa-se que o erro do FKE é aproximadamente o dobro do FP. O comportamento do filtro pode ser visto na Figura 13.

**Figura 13: Gráfico da aplicação do FP ao modelo SIS (Cenário 1).**



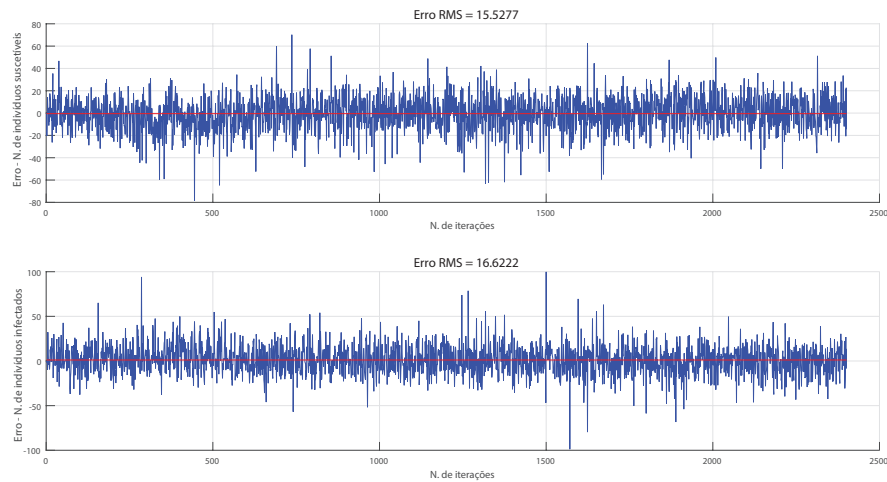
**Fonte: Autoria própria.**

**Figura 14: Zoom nas etapas de transição da aplicação do FP ao modelo SIS (Cenário 1) dos números de indivíduos de cada população.**



**Fonte: Autoria própria.**

**Figura 15: RMSE, erro absoluto e médio do FP em relação aos valores reais.**



**Fonte: Autoria própria.**

#### 4.4 APLICAÇÃO DO FKE — MODELO DE EPIDEMIA SIR E OS DADOS DO MODELO MOSES

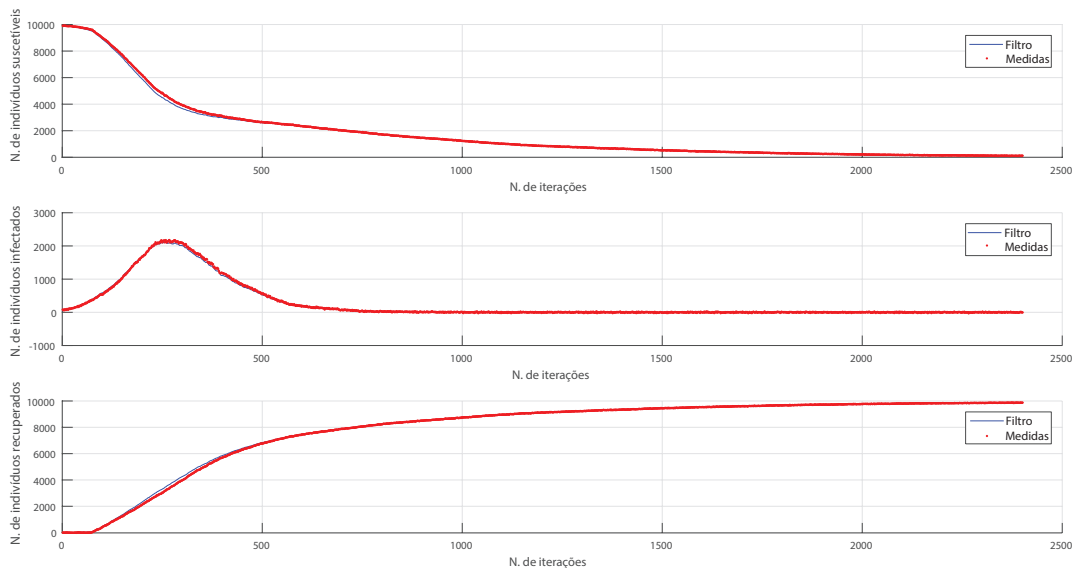
De maneira similar ao modelo SIS, aplicou-se ao modelo SIR ao FKE. Para esse modelo, utilizou-se o conjunto de equações em (17). Assim como para o modelo SIS, o conjunto de medidas para essa aplicação também recebeu a adição de ruído gaussiano. Os parâmetros inseridos no FKE foram os seguintes:

$$Q = \begin{bmatrix} 25 & 0 & 0 \\ 0 & 25 & 0 \\ 0 & 0 & 25 \end{bmatrix}, R = \begin{bmatrix} 100 & 0 & 0 \\ 0 & 100 & 0 \\ 0 & 0 & 100 \end{bmatrix}, \hat{x}_{0|0} = \begin{bmatrix} 9930 \\ 70 \\ 0 \end{bmatrix}, P_{0|0} = \begin{bmatrix} 99 & 0 & 0 \\ 0 & 50 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \quad (68)$$

Os parâmetros foram ajustados para sintonizar o FKE e gerar o melhor comportamento possível das estimativas, minimizando os erros, visto que, como citado anteriormente, para o modelo SIS, existe uma diferença entre o modelo utilizado para gerar as medidas e o modelo dinâmico utilizado na filtragem.

Os gráficos com as estimativas encontram-se nas Figuras 16 e 17.

**Figura 16: Gráfico da aplicação do FKE ao modelo SIR (Cenário 2).**



**Fonte: Autoria própria.**

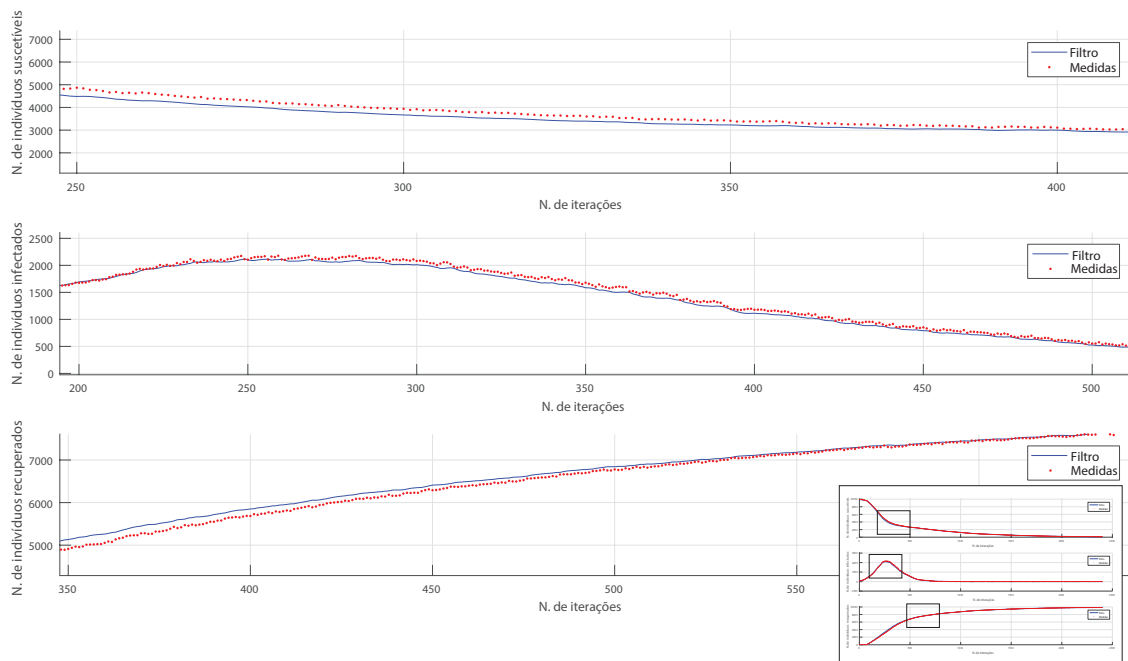
O erro da estimativa do modelo SIR, assim como para o modelo SIS, foi calculado a partir da diferença absoluta entre os valores reais e os valores estimados. No gráfico da Figura 17, pode-se observar que, de forma semelhante ao modelo SIS, o erro é maior na região de maior mudança no número de indivíduos de cada estado. É possível verificar a diferença

entre os valores medidos e os valores filtrados nas Figuras 17 e 18. Vale ressaltar que esse comportamento, obtido na filtragem de ambos os modelos, é esperado: como o filtro opera através de previsões e atualizações, valores com pouca variação em amplitude e em sequência são mais facilmente filtrados do que valores com uma grande variação.

A diferença principal em relação a filtragem dos modelos SIS e SIR é a grandeza dos valores. Na etapa de transição, a filtragem considerando o modelo SIR apresentou um erro maior do que a filtragem considerando o modelo SIS. Já com a população estabilizada, o erro da filtragem com o modelo SIR foi menor do que a filtragem com o modelo SIS nessa etapa.

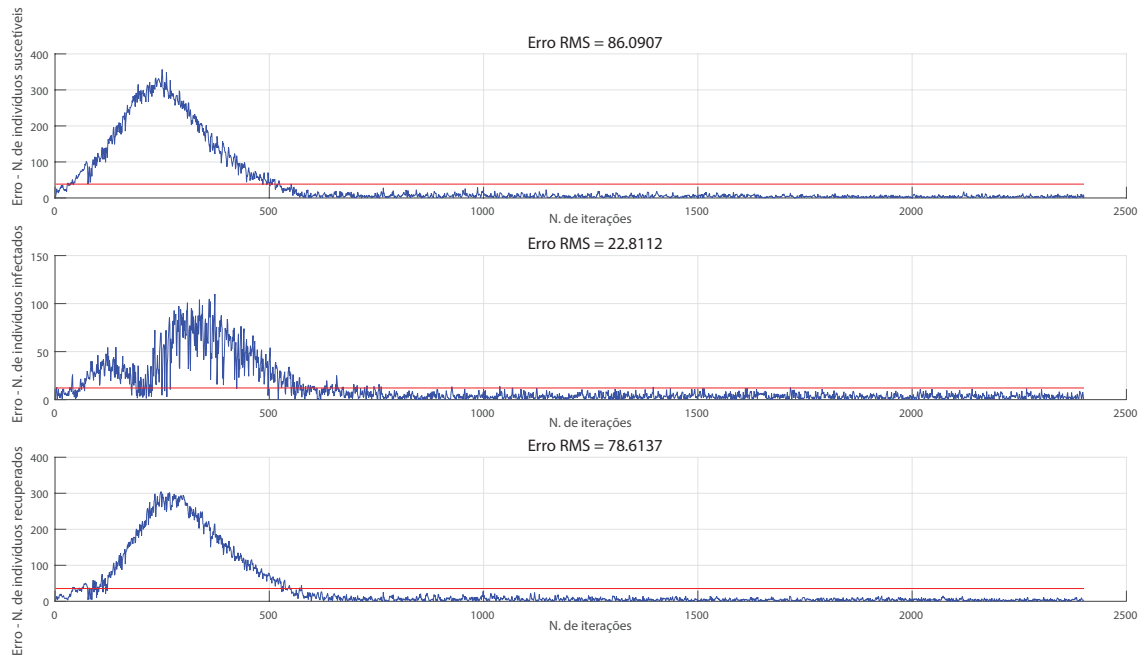
Os valores de erro foram os menores observados até o momento, dada a aplicação do modelo MOSES ao FKE. Isso se deu, pois dos modelos gerados pelo MOSES, o modelo SIS é o mais próximo do modelo dinâmico que alimenta os filtros, sendo assim como o modelo SIR gerado pelo MOSES é o que mais se difere dos que alimentam os filtros, gerando os maiores erros.

**Figura 17: Zoom nas etapas de transição da aplicação do FKE ao modelo SIR (Cenário 2) dos números de indivíduos de cada população.**



**Fonte: Autoria própria.**

**Figura 18: RMSE, erro absoluto e médio do FKE em relação aos valores reais.**



**Fonte: Autoria própria.**

#### 4.5 APLICAÇÃO DO FP — MODELO DE EPIDEMIA SIR E OS DADOS DO MODELO MOSES

O resultado esperado para esse processo de filtragem era novamente de uma performance superior do FP. Porém, diferentemente dos outros casos, o FKE apresentou um melhor resultado do que o FP. O desempenho do FKE foi, aproximadamente, duas vezes melhor do que o desempenho do FP, como é mostrado na Tabela 3.

**Tabela 3: Tabela do erro em número de indivíduos do modelo SIR gerado pelo MOSES.**

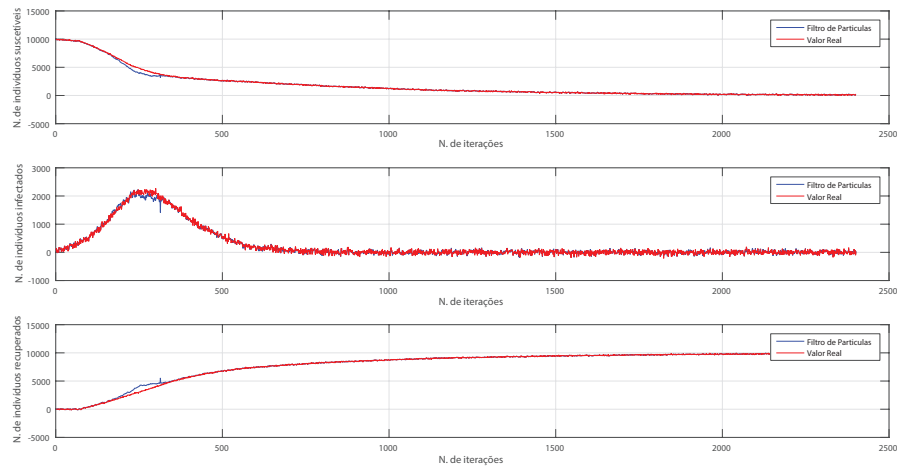
<b>Filtro</b>	<b>RMSE Suscetível</b>	<b>RMSE Infectado</b>	<b>RMSE Recuperado</b>
<b>FKE</b>	86,0907	22,8112	78,6137
<b>FP</b>	151,4758	57,3257	174,5098

**Fonte: Autoria própria.**

Observa-se na Figura 19 e com mais detalhes nas áreas de transição na Figura 20 que houve um grande distanciamento entre os valores filtrados e os valores reais nesta região. Esse comportamento não ocorreu com tanta intensidade no modelo SIS, filtrado pelo FP. Comparando as Figuras 18 e 21 verifica-se que, principalmente nas etapas suscetível e recuperado, que a amplitude do erro nas regiões de transição é significativamente maior para a filtragem com o FP do que com a filtragem pelo FKE.

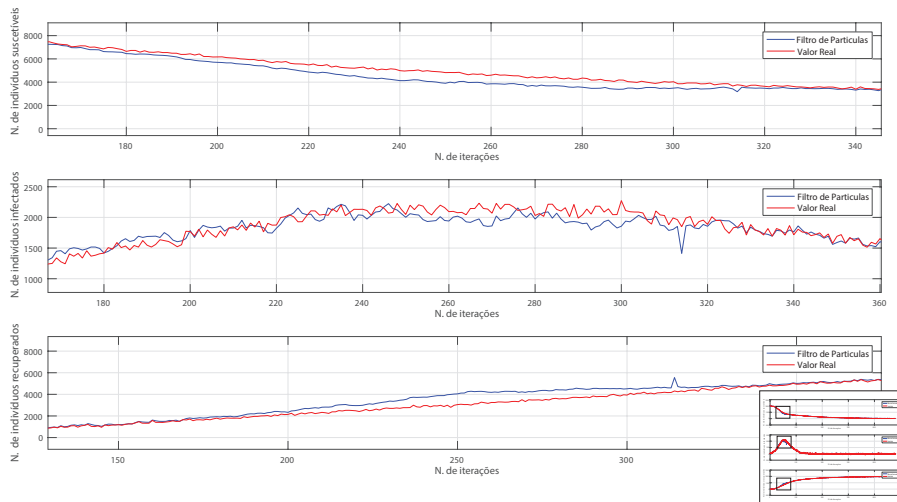
Apesar do resultado ser inesperado, não é incorreto. Devido à escolha feita para o valor dos parâmetros, pode-se acabar não chegando em sua melhor performance.

**Figura 19: Gráfico da aplicação do FP ao modelo SIR (Cenário 2).**



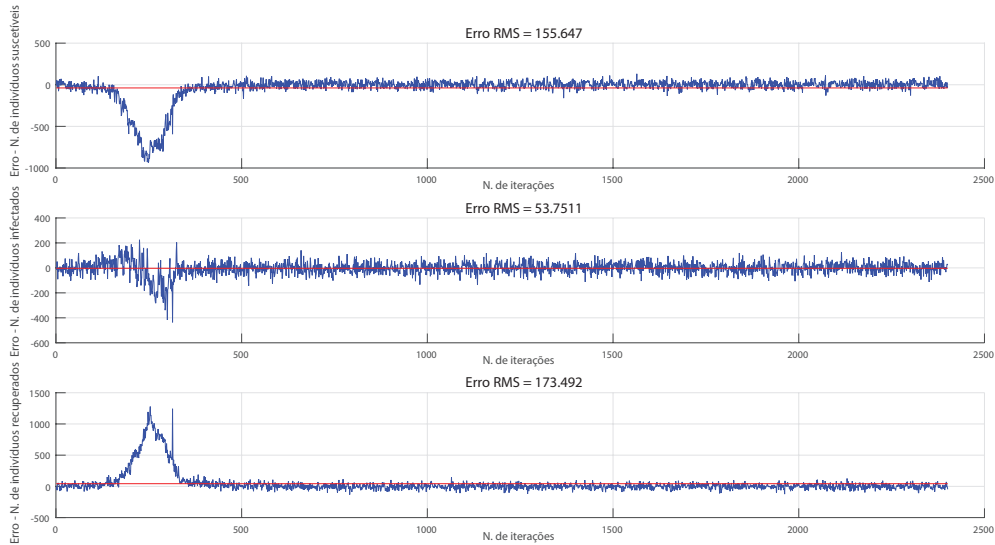
**Fonte: Autoria própria.**

**Figura 20: Zoom nas etapas de transição da aplicação do FP ao modelo SIR (Cenário 2) dos números de indivíduos de cada população.**



**Fonte: Autoria própria.**

**Figura 21: RMSE, erro absoluto e médio do FP em relação aos valores reais.**

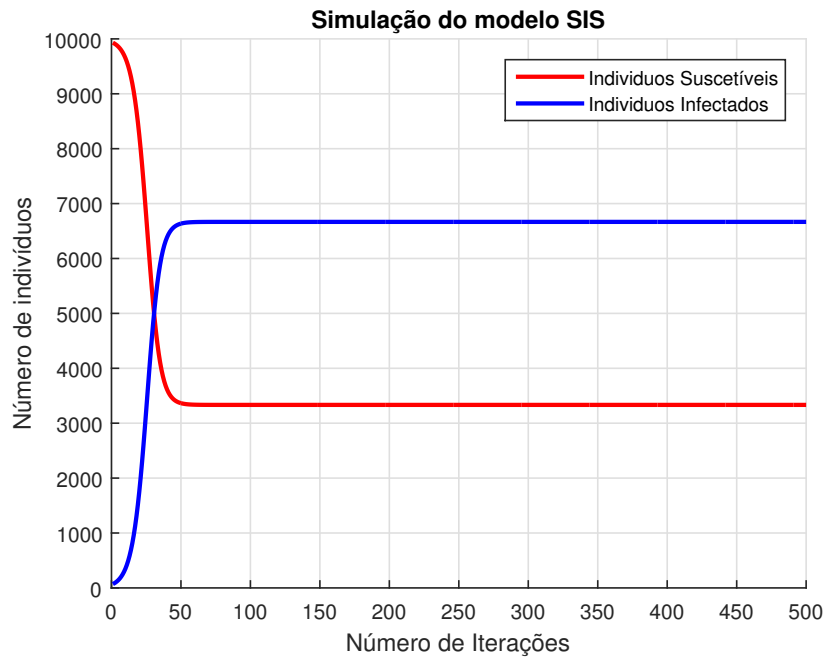


**Fonte: Autoria própria.**

#### 4.6 ANÁLISE FKE E FP ATRAVÉS DA GS DE DADOS SIMPLES

A geração de dados de forma sintética é realizada com base no do próprio modelo matemático em (13), (17) e (32). Espera-se que para esse tipo de geração de dados, haja um bom desempenho de ambos os filtros, pois o equacionamento que é utilizado para gerar o conjunto de dados é conhecido em sua totalidade pelo filtro, diferentemente dos dados gerados pelo MOSES.

Observa-se, na Figura 22, a simulação de um cenário com modelo SIS. Esses valores, como comentado anteriormente, foram gerados a partir do próprio modelo, utilizando as equações (13), (17) e (32). Os dados gerados representam a dinâmica populacional real de um grupo de 10000 indivíduos.

**Figura 22: Gráfico do modelo SIS.**

**Fonte: Autoria própria.**

Esse modelo foi então submetido a dois tipos diferentes de filtragem estocástica estudados neste trabalho: o FKE e o FP. Os gráficos das estimativas obtidas pelos dois filtros encontram-se nas Figuras 23 e 24, respectivamente.

Visualmente, ambos os métodos de filtragem apresentam o mesmo desempenho. Ambos os métodos de filtragem apresentaram um bom desempenho, ao analisar os RMSE de ambos, verifica-se na Tabela 4 que o FP apresenta um desempenho superior ao do FKE. Vale notar que o RMSE foi calculado através dos erros calculados da diferença entre o valor exato e o filtrado a cada instante da simulação.

**Tabela 4: Tabela do erro em número de indivíduos do modelo SIS gerado sinteticamente.**

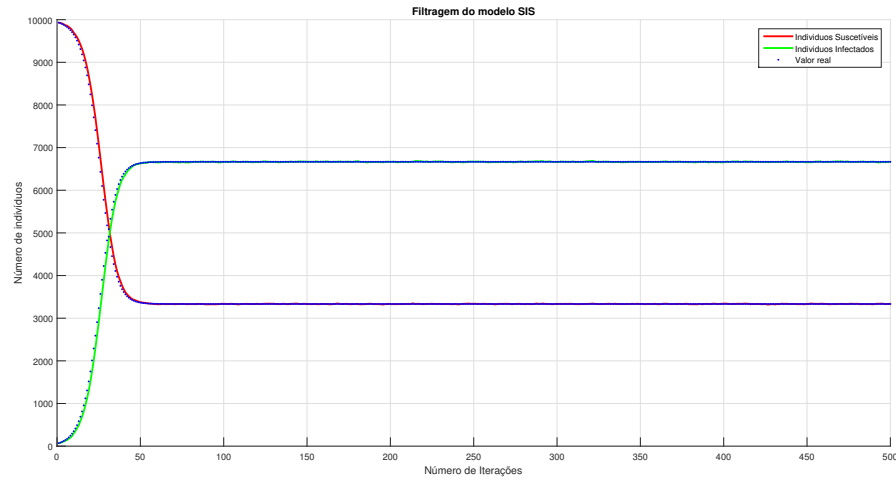
Filtro	RMSE Suscetível	RMSE Infectado
<b>FKE</b>	53,997	54,346
<b>FP</b>	8,680	9,338

**Fonte: Autoria própria.**

Seguindo com o que era previsto, o FP apresentou um desempenho superior ao do FKE, obtendo um RMSE aproximadamente 6 vezes menor para ambas as etapas de indivíduos do modelo SIS.

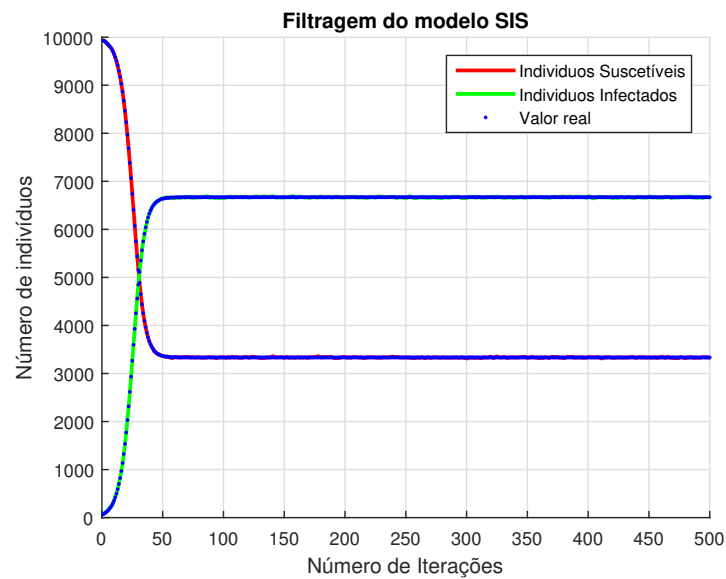


**Figura 23: Gráfico da aplicação do FKE ao modelo SIS.**



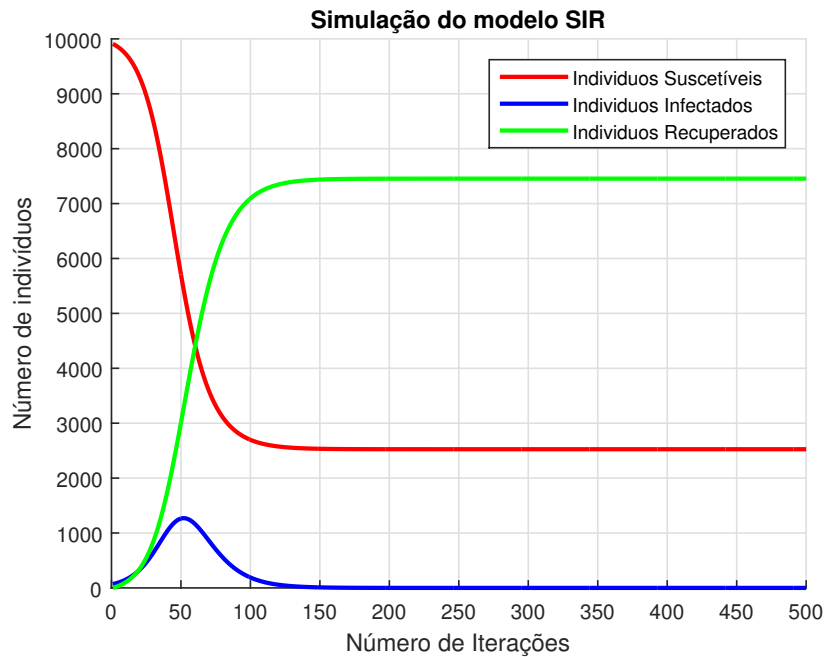
**Fonte: Autoria própria.**

**Figura 24: Gráfico da aplicação do FP ao modelo SIS.**



**Fonte: Autoria própria.**

Em seguida, realizou-se a simulação de um modelo SIR. O resultado dessa simulação pode ser observado na Figura 25. Assim como no modelo anterior, os dados também foram gerados com base no próprio modelo. Neste novo cenário, também foi utilizada uma população de 10000 para a simulação.

**Figura 25: Gráfico do modelo SIR.**

**Fonte: Autoria própria.**

Assim como no modelo SIS, ambos os filtros apresentaram pouco erro, apresentando uma leve variação, perceptível apenas nos momentos de mudanças mais bruscas que acontecem entre a primeira e a centésima iteração, como representado nos gráficos das Figuras 26 e 27.

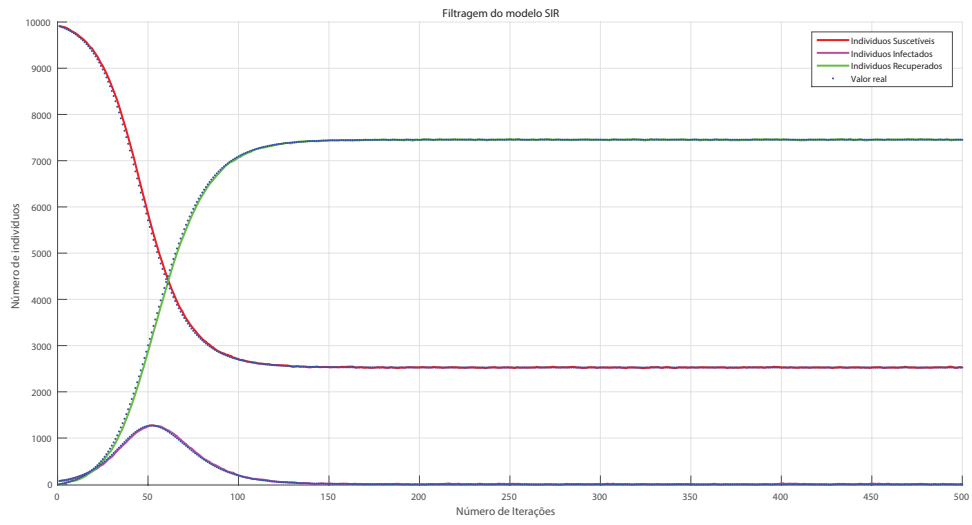
Como esperado o valor do RMSE do FP, por ser um filtro melhor, foi menor do que o RMSE do FKE, como exposto na Tabela 5. Comparativamente ao modelo SIS, o FKE performou melhor e o FP apresentou desempenho semelhante.

**Tabela 5: Tabela do erro em número de indivíduos do modelo SIR gerado sinteticamente.**

<b>Filtro</b>	<b>RMSE Suscetível</b>	<b>RMSE Infectado</b>	<b>RMSE Recuperado</b>
<b>FKE</b>	39,744	12,538	38,521
<b>FP</b>	9,572	8,924	9,355

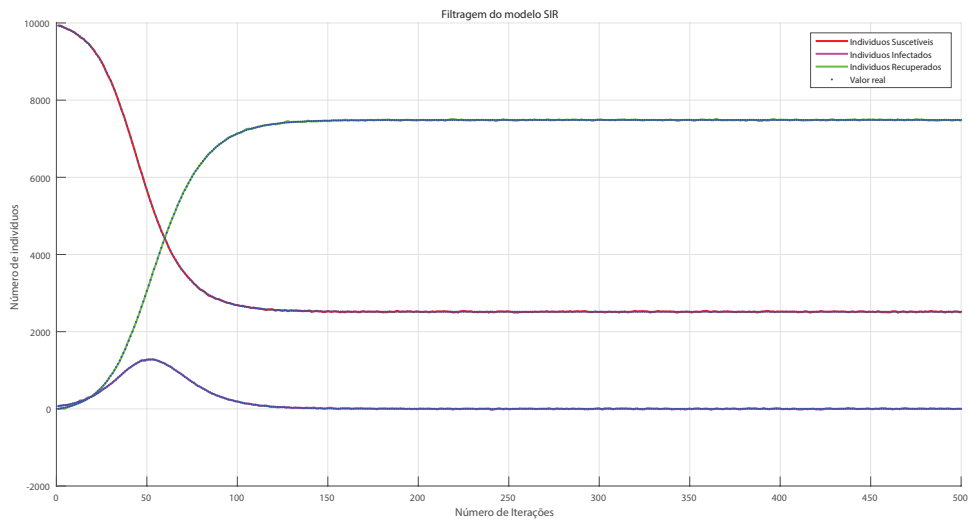
**Fonte: Autoria própria.**

**Figura 26: Gráfico da aplicação do FKE ao modelo SIR.**



**Fonte: Autoria própria.**

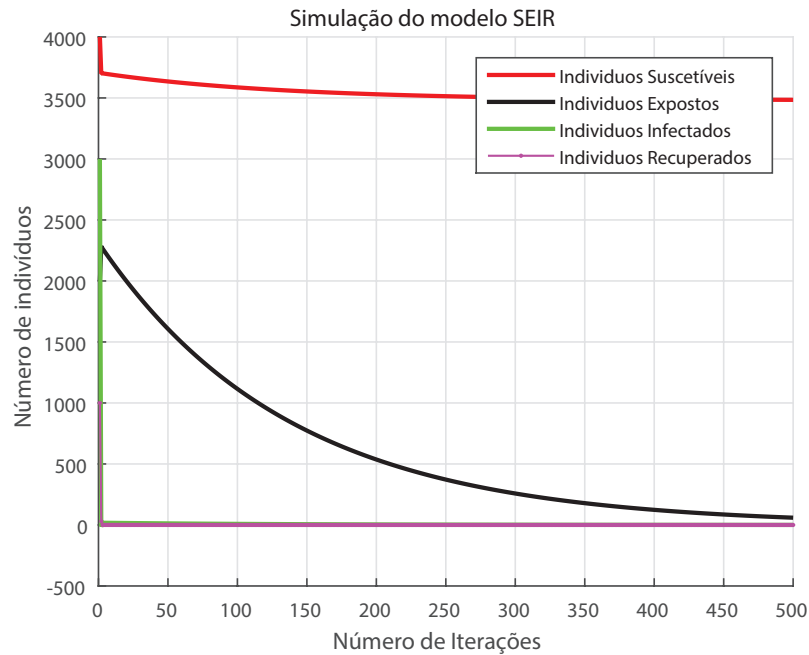
**Figura 27: Gráfico da aplicação do FP ao modelo SIR.**



**Fonte: Autoria própria.**

Por fim, realizou-se a filtragem do modelo SEIR. Devido ao maior número de etapas e maior exigência de processamento, o número total de indivíduos da população para esse modelo foi reduzido de 10000 para 4000, como pode ser visto na Figura 28.

**Figura 28: Gráfico do modelo SEIR.**



**Fonte: Autoria própria.**

Novamente, os filtros apresentaram bom desempenho, evidenciado nas Figuras 29 e 30 e apresentado na Tabela 6. Para a etapa de indivíduos recuperados os valores do RMSE foram muito parecidos entre os métodos de filtragem. Já para as outras etapas os valores foram de grandeza semelhantes aos obtidos com o modelo SIR.

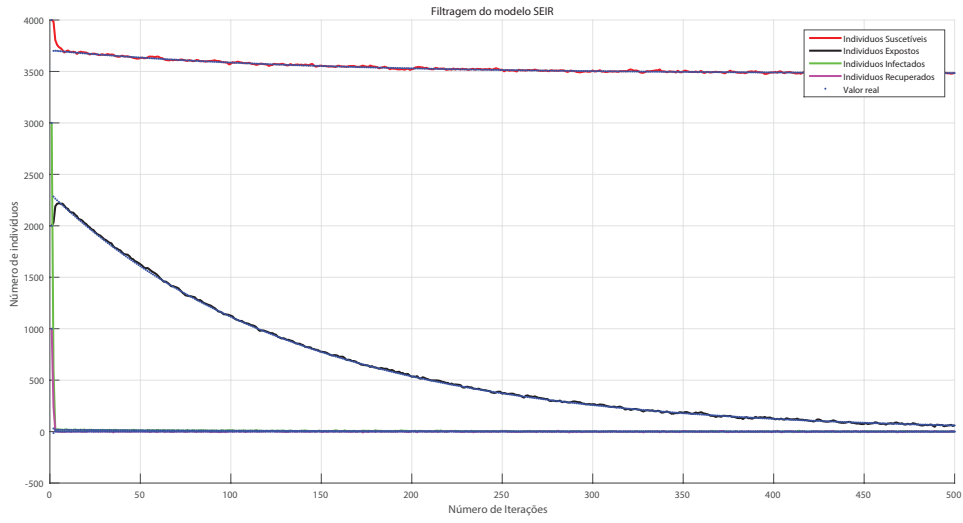
**Tabela 6: Tabela do erro em número de indivíduos do modelo SEIR gerado sinteticamente.**

Filtro	RMSE Suscetível	RMSE Exposto	RMSE Infectado	RMSE Recuperado
<b>FKE</b>	16,166	14,328	27,486	9,629
<b>FP</b>	9,236	9,410	6,645	6,645

**Fonte: Autoria própria.**

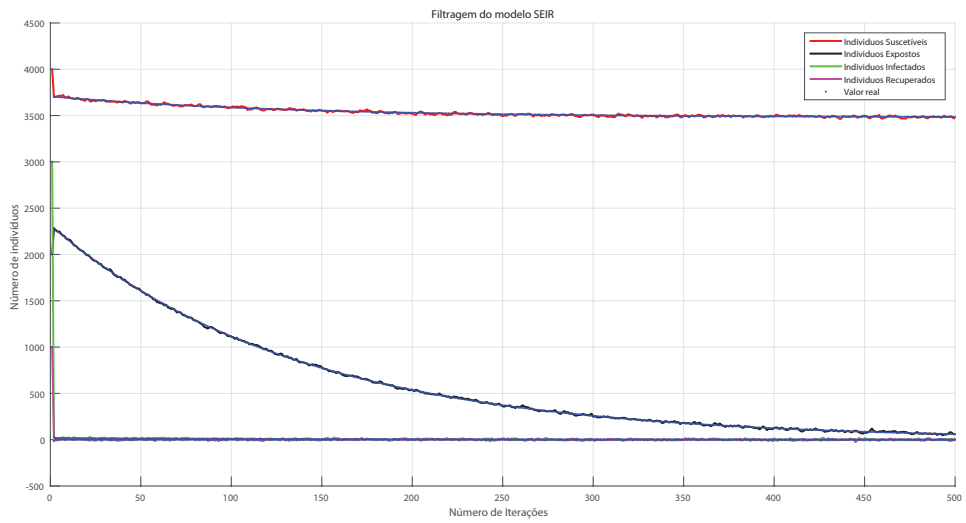
No gráfico apresentado na Figura 28 os indivíduos infectados (representados pela cor verde) estão sobrepostos pelos indivíduos recuperados (representados pela cor magenta).

**Figura 29: Gráfico da aplicação do FKE ao modelo SEIR.**



**Fonte: Autoria própria.**

**Figura 30: Gráfico da aplicação do FP ao modelo SEIR.**



**Fonte: Autoria própria.**

#### 4.7 SÍNTESE DO CAPÍTULO

Neste capítulo, foram realizadas as implementações em MATLAB dos filtros FKE e FP com os modelos epidemiológicos SIS e SIR utilizando os dados gerados pelo MOSES, além da implementação da geração de dados sintética que tornou possível a utilização do modelo SEIR.

Para a geração das medições, utilizou-se os dados gerados pelo MOSES com a adição de um ruído gaussiano para simular as medidas efetuadas em cada modelo epidêmico e os dados gerados pelo próprio modelo. Da mesma forma, foi adicionado o mesmo tipo de ruído aos valores criados pela geração sintética de dados.

Obtidas as medições, implementou-se o FKE e o FP e aplicou-se nos modelos SIS, SIR e SEIR discretizados. A aplicação do filtro nos dados gerados externamente foi mais trabalhosa, visto que o MOSES possui mais parâmetros para descrever as epidemias que os modelos SIS e SIR. As medidas resultantes destoaram levemente dos valores reais, sendo 173,402 o maior valor de RMSE dentro de uma população de 10000 indivíduos.

## 5 CONCLUSÕES

Foram estudados modelos matemáticos que representassem a dinâmica populacional que seriam aplicados os métodos de filtragem estocástica. Os modelos que mais se adequaram à proposta foram os modelos SIS e SIR. Esses são modelos simples o suficiente para representar, de forma genérica, boa parte das dinâmicas infecciosas existentes. Por esse motivo, tais modelos são largamente utilizados no universo de estudos epidemiológicos. No decorrer da elaboração do trabalho, com o surgimento da COVID-19, incorporou-se também ao trabalho a análise do modelo SEIR, que descreve uma dinâmica semelhante à pandemia do coronavírus.

Utilizou-se duas formas de geração de dados a partir dos modelos definidos nesse trabalho. A primeira forma recebeu o nome de GS de dados. Nela, os dados são gerados com base no modelo matemático e o ruído é adicionado ao conjunto de dados para ser utilizado como medidas. A geração de dados dessa forma torna o erro menor pelo motivo de que o filtro estocástico está utilizando um modelo que descreve perfeitamente o sistema a partir do qual foram obtidas as medidas.

A segunda forma de geração é feita pelo MOSES. O MOSES é um *software open-source* de geração de modelos epidêmicos, com ele é possível gerar modelos simples, como o SIS e o SIR, mas também criar modelos mais complexos adicionando etapas, como a etapa “Quarentenado” e diferentes transições entre estas e as demais etapas. Os filtros estocásticos apresentam um rendimento abaixo do esperado quando aplicados aos modelos gerados pelo MOSES. Encontrou-se dificuldades na integração do mesmo com os filtros, visto que os dados eram gerados sem depender do modelo.

Dois filtros estocásticos foram escolhidos para serem aplicados nos modelos gerados. O primeiro deles foi o FKE e depois o FP, filtro estocástico mais geral. Como já era esperado, o FP demonstrou um desempenho superior ao do FKE em 75% das simulações. O único caso em que o FP apresentou um desempenho inferior ao FKE foi na filtragem com o modelo SIR baseado na geração externa de medidas pelo MOSES. Em relação aos diferentes modelos, de forma geral, o FP apresentou desempenho semelhante, independentemente do modelo utilizado

quando utilizando, dados gerados sinteticamente. Quando os dados eram gerados externamente através do MOSES, o desempenho se tornou irregular.

Tendo em vista a atualidade do assunto tratado neste TCC, enfatiza-se a importância e a necessidade dos dados durante essas situações como a que se vive hoje, na época em que está sendo escrito este TCC. A informação e o conhecimento são as ferramentas mais poderosas no combate, não só do Coronavírus, mas de todas as possíveis pandemias subsequentes que poderão surgir futuramente no mundo. Dessa forma, esse trabalho pode ser continuado por outros alunos e pesquisadores de forma a especificar cada vez mais os diferentes tipos de modelos para que represente, de forma fiel, cada variante de epidemia que possa surgir. Além disso, esses modelos podem levar em consideração não só as peculiaridades do agente infeccioso, mas também o efeito que diferentes políticas sanitárias públicas e comportamentos sociais tem no desenrolar de uma epidemia.



## REFERÊNCIAS

- AGUIRRE, L. **Introdução à Identificação de Sistemas**. 2015. ISBN 978-85-423-0079-6.
- ALLEN, L. J. An introduction to stochastic epidemic models. In: **Mathematical epidemiology**. Springer, 2008. p. 81–130.
- ANDERSON, R. M.; FRASER, C.; GHANI, A. C.; DONNELLY, C. A.; RILEY, S.; FERGUSON, N. M.; LEUNG, G. M.; LAM, T. H.; HEDLEY, A. J. Epidemiology, transmission dynamics and control of sars: the 2002-2003 epidemic. **Philosophical Transactions of the Royal Society B: Biological Sciences**, The Royal Society, v. 359, n. 1447, p. 1091, 2004.
- ARGUETA, L. M. A. Particle filtering estimation for linear and nonlinear state-space models. Universitat Politècnica de Catalunya, 2013.
- ARULAMPALAM, M. S.; MASKELL, S.; GORDON, N.; CLAPP, T. A tutorial on particle filters for online nonlinear/non-gaussian bayesian tracking. **IEEE Transactions on Signal Processing**, v. 50, n. 2, p. 174–188, Feb 2002. ISSN 1053-587X.
- BAR-SHALOM, Y.; LI, X. R.; KIRUBARAJAN, T. **Estimation with applications to tracking and navigation: theory algorithms and software**. John Wiley & Sons, 2004.
- BHADRA, A.; IONIDES, E. L.; LANERI, K.; PASCUAL, M.; BOUMA, M.; DHIMAN, R. C. Malaria in northwest india: Data analysis via partially observed stochastic differential equation models driven by lévy noise. **Journal of the American Statistical Association**, Taylor & Francis, v. 106, n. 494, p. 440–451, 2011.
- BHATTACHARYA, R. N.; WAYMIRE, E. C. **Stochastic processes with applications**. Siam, 2009.
- BILLINGS, S. A. **Nonlinear system identification: NARMAX methods in the time, frequency, and spatio-temporal domains**. John Wiley & Sons, 2013.
- BOLDOG, P.; TEKELI, T.; VIZI, Z.; DÉNES, A.; BARTHA, F. A.; RÖST, G. Risk assessment of novel coronavirus covid-19 outbreaks outside china. **Journal of Clinical Medicine**, MDPI AG, v. 9, n. 2, p. 571, Feb 2020. ISSN 2077-0383.
- BROWN, S. D.; STEYVERS, M. Detecting and predicting changes. **Cognitive psychology**, Elsevier, v. 58, n. 1, p. 49–67, 2009.
- BUCY, R. S.; JOSEPH, P. D. Filtering for stochastic processes with applications to guidance. **American Mathematical Soc.**, 2005.
- BURDEN, R. L.; FAIRES, J. D. **Análise numérica**. Cengage Learning, 2008.
- CARPENTER, J.; CLIFFORD, P.; FEARNHEAD, P. Improved particle filter for nonlinear problems. **IEE Proceedings - Radar, Sonar and Navigation**, v. 146, n. 1, p. 2–7, Feb 1999. ISSN 1350-2395.

- CAZELLES, B.; CHAU, N. Using the kalman filter and dynamic models to assess the changing hiv/aids epidemic. **Mathematical biosciences**, Elsevier, v. 140, n. 2, p. 131–154, 1997.
- CIPOLLI, V. G. Sistemas dinâmicos discretos- análise de estabilidade. UNESP, 2012.
- DJORDJEVIC, J.; SILVA, C. J.; TORRES, D. F. A stochastic sica epidemic model for hiv transmission. **Applied Mathematics Letters**, Elsevier, v. 84, p. 168–175, 2018.
- DOUCET, A.; FREITAS, N. de; GORDON, N. Sequential monte carlo in practice. **Springer-Verlag**, 2001.
- DOUCET, A.; GODSILL, S.; ANDRIEU, C. On sequential monte carlo sampling methods for bayesian filtering. **Statistics and computing**, Springer, v. 10, n. 3, p. 197–208, 2000.
- FANG, M.; LI, L.-N.; YANG, L. Social network public opinion research based on s-seir epidemic model. In: IEEE. **2019 IEEE International Conference on Cloud Computing Technology and Science (CloudCom)**. 2019. p. 374–379.
- FOUNDATION, B. . M. G. Seir and seirs models. 2020. Disponível em: <Disponível em: <https://www.idmod.org/docs/emod/hiv/model-seir.html>>.
- FREITAS, A. D.; MIHAYLOVA, L.; GNING, A.; ANGELOVA, D.; KADIRKAMANATHAN, V. Autonomous crowds tracking with box particle filtering and convolution particle filtering. **Automatica**, Elsevier, v. 69, p. 380–394, 2016.
- GAYLOR DAVID; LIGHTSEY, E. G. Gps/ins kalman filter design for spacecraft operating in the proximity of international space station. AIAA Guidance, Navigation, and Control Conference and Exhibit, 2003.
- GORDON, N.; RISTIC, B.; ARULAMPALAM, S. Beyond the kalman filter: Particle filters for tracking applications. **Artech House, London**, v. 830, p. 3–63, 2004.
- GUSTAFSSON, F. Particle filter theory and practice with positioning applications. **IEEE Aerospace and Electronic Systems Magazine**, IEEE, v. 25, n. 7, p. 53–82, 2010.
- HETHCOTE, H. W. Qualitative analyses of communicable disease models. **Mathematical Biosciences**, Elsevier, v. 28, n. 3-4, p. 335–356, 1976.
- ISARD, M.; BLAKE, A. Condensation—conditional density propagation for visual tracking. **International journal of computer vision**, Springer, v. 29, n. 1, p. 5–28, 1998.
- JÉGAT, C.; CARRAT, F.; LAJAUNIE, C.; WACKERNAGEL, H.; STEINSKOG, D. J.; BERTINO, L. Early detection and assessment of epidemics by particle filtering. **GeoENV VI - Geostatistics for Environmental Applications**, v. 15, 01 2008.
- KALMAN, R. E. A new approach to linear filtering and prediction problems. **Journal of basic Engineering**, American Society of Mechanical Engineers, v. 82, n. 1, p. 35–45, 1960.
- KHAN, T.; ULLAH, Z.; ALI, N.; ZAMAN, G. Modeling and control of the hepatitis b virus spreading using an epidemic model. **Chaos, Solitons & Fractals**, Elsevier, v. 124, p. 1–9, 2019.
- KONG, A.; LIU, J. S.; WONG, W. H. Sequential imputations and bayesian missing data problems. **Journal of the American statistical association**, Taylor & Francis Group, v. 89, n. 425, p. 278–288, 1994.

LACERDA, E. G. N. M. J. Uma nova estratégia de controle para o lockdown na pandemia da covid-19. **Técnicas de modelagem e controle para dirimir o contágio e propagação da Covid-19**, 2020.

LEON-GARCIA, A. Probability, statistics, and random processes for electrical engineering. Pearson Education, 2008.

LI, X. R.; JILKOV, V. P. Survey of maneuvering target tracking. part v. multiple-model methods. **IEEE Transactions on Aerospace and Electronic Systems**, IEEE, v. 41, n. 4, p. 1255–1321, 2005.

MCELHOE, B. A. An assessment of the navigation and course corrections for a manned flyby of mars or venus. **IEEE Transactions on Aerospace and Electronic Systems**, 1966.

MCGEE, L. A.; SCHIMIDT, S. F. Discovery of the kalman filter as a practical tool for aerospace and industry. NASA STI FACILITY ACESS DEPT., 11 1985.

MONTEMERLO, M.; THRUN, S.; KOLLER, D.; WEGBREIT, B. et al. Fastslam: A factored solution to the simultaneous localization and mapping problem. **Aaai/iaai**, v. 593598, 2002.

MURAD, D.; BADSHAH, N.; ALI, S. M. Mathematical modeling and simulation for the dengue fever epidemic. In: IEEE. **2018 International Conference on Applied and Engineering Mathematics (ICAEM)**. 2018. p. 1–3.

OGATA, K.; YANG, Y. **Modern control engineering**. Prentice-Hall, 2002.

PARZEN, E. **Stochastic processes**. SIAM, 1999.

SANBORN, A. N.; GRIFFITHS, T. L.; NAVARRO, D. J. Rational approximations to rational models: alternative algorithms for category learning. **Psychological review**, American Psychological Association, v. 117, n. 4, p. 1144, 2010.

SKVORTSOV, A.; CONNELL, R.; DAWSON, P.; GAILIS, R. Epidemic modelling: Validation of agent-based simulation by using simple mathematical models. In: CITESEER. **MODSIM 2007 International Congress on Modelling and Simulation. Modelling and Simulation Society of Australia and New Zealand**. 2007. p. 657–662.

SKVORTSOV, A.; RISTIC, B. Monitoring and prediction of an epidemic outbreak using syndromic observations. **Mathematical biosciences**, Elsevier, v. 240, n. 1, p. 12–19, 2012.

SPEEKENBRINK, M. A tutorial on particle filters. **Journal of Mathematical Psychology**, Elsevier, v. 73, p. 140–152, 2016.

TULSYAN, A.; GOPALUNI, R. B.; KHARE, S. R. Particle filtering without tears: A primer for beginners. **Computers & Chemical Engineering**, Elsevier, v. 95, p. 130–145, 2016.

VALLE, D.; PIMENTA, D. N.; AGUIAR, R. Zika, dengue e chikungunya: desafios e questões. **Epidemiologia e Serviços de Saúde**, SciELO Public Health, v. 25, p. 419–422, 2016.

VAROL, H. A. Moses: A matlab-based open-source stochastic epidemic simulator. In: IEEE. **2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)**. 2016. p. 2636–2639.

VLADAR, H. P. de; PEN, I. Determinism, noise, and spurious estimations in a generalised model of population growth. **Physica A: Statistical Mechanics and its Applications**, Elsevier, v. 373, p. 477–485, 2007.

WAN, E.; MERWE, R. The unscented kalman filter for nonlinear estimation. In: . 2000. v. 153-158, p. 153 – 158. ISBN 0-7803-5800-7.

WHO. **Coronavirus disease 2019 (COVID-19) Situation Report – 73**. apr 2020. Disponível em: <[https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200402-sitrep-73-covid-19.pdf?sfvrsn=5ae25bc7\\_2](https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200402-sitrep-73-covid-19.pdf?sfvrsn=5ae25bc7_2)>.

YI, M. S.; STEYVERS, M.; LEE, M. Modeling human performance in restless bandits with particle filters. **The Journal of Problem Solving**, v. 2, n. 2, p. 5, 2009.

YORINORI, J.; PAIVA, W.; FREDERICK, R.; COSTAMILAN, L.; BERTAGNOLLI, P.; HARTMAN, G.; GODOY, C.; JR, J. N. Epidemics of soybean rust (*phakopsora pachyrhizi*) in brazil and paraguay from 2001 to 2003. **Plant Disease**, Am Phytopath Society, v. 89, n. 6, p. 675–677, 2005.