

UNIVERSIDADE FEDERAL DO RIO GRANDE
CENTRO DE CIÊNCIAS COMPUTACIONAIS
PROGRAMA DE PÓS-GRADUAÇÃO EM MODELAGEM COMPUTACIONAL
CURSO DE MESTRADO EM MODELAGEM COMPUTACIONAL

Dissertação de Mestrado

**Utilização de Distribuições de Probabilidades em
Variáveis de Simulação Baseada em Agentes: Um Estudo
de Caso para a Curva de Crescimento do Bacilo da
Tuberculose**

Marcilene Fonseca de Moraes

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Modelagem Computacional da Universidade Federal do Rio Grande, como requisito parcial para a obtenção do grau de Mestre em Modelagem Computacional

Orientador: Prof^a. Dr^a. Diana Francisca Adamatti
Co-orientador: Prof. Dr. Adriano Velasque Werhli

Rio Grande, 2016

Ficha catalográfica

M827u Moraes, Marcilene Fonseca de.

Utilização de distribuições de probabilidades em variáveis de simulação baseada em agentes: um estudo de caso para curva de crescimento do bacilo da tuberculose / Marcilene Fonseca de Moraes. – 2016.

66 f.

Dissertação (mestrado) – Universidade Federal do Rio Grande – FURG, Programa de Pós-graduação em Modelagem Computacional, Rio Grande/RS, 2016.

Orientador: Dr^a. Diana Francisca Adamatti.

Coorientador: Dr. Adriano Velasque Werhli.

1. Tuberculose 2. Simulação baseada em agentes 3. Distribuição
4. Probabilidade I. Adamatti, Diana Francisca II. Werhli, Adriano Velasque
III. Título.

CDU 004.891

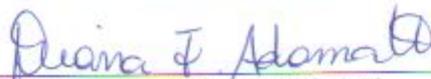
MARCILENE FONSECA DE MORAES

“Utilização de Distribuições de Probabilidades em Variáveis de Simulação Baseada em Agentes: Um Estudo de Caso para a Curva de Crescimento do Bacilo da Tuberculose”

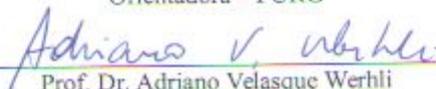
Dissertação apresentada ao Programa de Pós Graduação em Modelagem Computacional da Universidade Federal do Rio Grande - FURG, como requisito parcial para obtenção do Grau de Mestre. Área de concentração: Modelagem Computacional.

Aprovada em

BANCA EXAMINADORA



Prof. Dr.ª Diana Francisca Adamatti
Orientadora – FURG



Prof. Dr. Adriano Velasque Werhli
Co-orientador - FURG



Prof. Dr.ª Viviane Leite Dias de Mattos
FURG



Prof. Dr. Fabio Yoshimitsu Okuyama
IFRS



Prof. Dr. Cleo Zanella Billa
FURG

Rio Grande - RS
2016

*Dedico este trabalho a minha mãe, Maria Tereza,
e ao meu namorado Mário Roberto,
sem vocês nada disso teria sido possível.*

AGRADECIMENTOS

A Deus, pela proteção e cuidado incondicional.

A minha família, pelo amor e paciência dedicados, e por entender minha ausência.

Aos amigos, pelas palavras de carinho e incentivo durante esta jornada.

À orientadora, Diana Adamatti, pelo apoio, ensinamentos e paciência.

Ao co-orientador, Adriano Werhli, pela disponibilidade e ideias.

Ao colaborador que tive no desenvolvimento desta dissertação, Albano Borba, obrigada pela parceria.

Ao curso de Modelagem Computacional por me acolher em seu programa e tornar possível a conclusão de mais esta etapa.

À CAPES pelo apoio financeiro.

E, finalmente, aos colegas de mestrado pelo carinho, apoio, pelos bons momentos compartilhados e pela troca de conhecimentos.

“Você pode encarar um erro como uma besteira a ser esquecida, ou como um resultado que aponta uma nova direção”.

STEVE JOBS

RESUMO

MORAES, Marcilene Fonseca de. **Utilização de Distribuições de Probabilidades em Variáveis de Simulação Baseada em Agentes: Um Estudo de Caso para a Curva de Crescimento do Bacilo da Tuberculose**. 2016. 66 f. Dissertação (Mestrado) – Programa de Pós-Graduação em Modelagem Computacional. Universidade Federal do Rio Grande, Rio Grande.

Mesmo tratável e podendo ser prevenida com medicamentos, a tuberculose (TB) continua a infectar e causar mortes em toda população, especialmente nas de países pobres e em partes mais vulneráveis de países ricos. Diante dessa situação, o estudo da curva de crescimento do *Mycobacterium tuberculosis*, causador da tuberculose, pode ser um forte aliado contra a TB, seja para a elaboração de novos fármacos, ou até mesmo no desenvolvimento da teoria. Este trabalho modela a curva de crescimento do *Mycobacterium tuberculosis* com a utilização de simulação baseada em agentes, e tem como objetivo simular a curva com o mínimo de erro possível com relação à real. Para implementar esse modelo, foi definido um meio onde agentes representam as bactérias em seu *habitat*, a forma como interagem umas com as outras e com o ambiente. E para tornar o modelo mais similar a um modelo de crescimento real, os valores das variáveis dos agentes são gerados a partir de distribuições de probabilidades. O modelo baseado em agentes desenvolvido apresentou resultados satisfatórios, uma vez que a curva de crescimento modelada revelou-se similar a curva real. Essa similaridade torna o modelo bastante útil para verificação de hipóteses, já que as simulações levam minutos, em oposição aos experimentos *in vitro* que levariam dias.

Palavras-chave: *Mycobacterium tuberculosis*, Simulação Baseada em Agentes, Distribuições de Probabilidades.

ABSTRACT

MORAES, Marcilene Fonseca de. **Using Probability Distributions in Variables at Agent-Based Simulations: A Case Study For The Tuberculosis Bacillus Growth Curve.** 2016. 66 f. Dissertação (Mestrado) – Programa de Pós-Graduação em Modelagem Computacional. Universidade Federal do Rio Grande, Rio Grande.

Even treatable and prevented with medicines, tuberculosis (TB) continues to infect and cause deaths in the whole population, especially in poor countries and in vulnerable parts of the rich countries. In this way, the study of the growth curve of *Mycobacterium tuberculosis* can be a strong ally against TB, to develop new drugs or a theory. This paper models the growth curve of *Mycobacterium tuberculosis* with using agent-based simulation, and it aims to simulate the curve with minimum error possible comparing to real. To implement this model, we have defined agents as the bacteria in their *habitat*, how they interact each other and with the environment. And to make the model more similar to a real growth model, the values of the agents variables are generated from probability distributions. The agent-based model developed presented satisfactory results, since the growth curve modeled proved similar to real curve. This similarity makes the model high useful for verification of hypotheses, since simulations take minutes to execute on contrary of *in vitro* experiments that could take days.

Keywords: *Mycobacterium tuberculosis*, Agent-Based Simulation, Probability Distributions.

LISTA DE FIGURAS

Figura 1	Curva de crescimento e suas fases (TODAR, 2013).	17
Figura 2	Curvas reais obtidas em experimento (GROLL, 2010).	19
Figura 3	Distribuição uniforme no intervalo $[\alpha, \beta]$	21
Figura 4	F.d.p de uma variável aleatória normal como média μ e desvio padrão σ	22
Figura 5	Distribuições normais com variações na μ e no σ	23
Figura 6	Simulação de variável com distribuição uniforme.	24
Figura 7	Simulação de variável com distribuição normal.	25
Figura 8	Etapas de um processo de simulação (ADAMATTI, 2011).	26
Figura 9	Tela de abertura NetLogo.	27
Figura 10	Wolf Sheep Predation, livreria NetLogo (WILENSKY, 1999).	28
Figura 11	Exemplo de código: Wolf Sheep Predation, livreria NetLogo (WILENSKY, 1999).	29
Figura 12	Fluxograma do ciclo de vida dos agentes (WERLANG, 2013).	35
Figura 13	Interface do modelo desenvolvido no NetLogo.	36
Figura 14	Função responsável pela metabolização dos agentes.	37
Figura 15	Função responsável pela movimentação dos agentes.	38
Figura 16	Função responsável pela reprodução dos agentes.	39
Figura 17	Função responsável por resolver a etapa de sinalização dos agentes.	40
Figura 18	Comparação de crescimento para diferentes valores da variável funções vitais. Em (a) utiliza-se o valor 0,5 e em (b) 0,7. As outras variáveis mativeram-se com parâmetros fixos nas duas simulações.	41
Figura 19	Comparação de crescimento para diferentes valores da variável energia reproduz. Em (a) utiliza-se o valor 200 e em (b) 250. Todas as outras variáveis tem valores iguais nas duas simulações.	42
Figura 20	Comparação de crescimento para diferentes valores da variável sensor sinal. Em (a) utiliza-se o valor 2000 e em (b) 3000. Todas as outras variáveis tem valores iguais nas duas simulações.	43
Figura 21	Curvas experimentais representadas por agentes.	44
Figura 22	Simulação de GC 01-2522.	48
Figura 23	Simulação de GC 02-2761.	49
Figura 24	Simulação de GC 03-0850.	49
Figura 25	Simulação de GC 02-2922.	50
Figura 26	Simulação de H37Rv.	50
Figura 27	Curvas simuladas	51
Figura 28	Comparação das curvas real e simulada de GC 01-2522.	52

Figura 29	Comparação das curvas real e simulada de GC 02-2761.	53
Figura 30	Comparação das curvas real e simulada de GC 03-0850.	53
Figura 31	Comparação das curvas real e simulada de GC 03-2922.	54
Figura 32	Comparação das curvas real e simulada de H37Rv.	54
Figura 33	Comparação dos crescimentos GC 01-2522.	57
Figura 34	Comparação dos crescimentos GC 02-2761.	58
Figura 35	Comparação dos crescimentos GC 03-0850.	58
Figura 36	Comparação dos crescimentos GC 03-2922.	59
Figura 37	Comparação dos crescimentos GC H37Rv.	59

LISTA DE TABELAS

Tabela 1	Cepas utilizadas e suas origens (GROLL, 2010).	18
Tabela 2	Parâmetros em comum a todas as populações de agentes.	45
Tabela 3	Parâmetros das variáveis dos agentes comuns a todas as cepas. A média e desvio padrão de cada variável refere-se a uma determinada distribuição normal.	46
Tabela 4	Parâmetros das variáveis que diferem de cepa para cepa. A média e desvio padrão são entradas de uma distribuição normal.	47
Tabela 5	Tempo de adaptação, crescimento e população de cada uma das cepas.	51
Tabela 6	Tempo de adaptação, crescimento e população de GC 01-2522. . . .	55
Tabela 7	Tempo de adaptação, crescimento e população de GC 02-2761. . . .	55
Tabela 8	Tempo de adaptação, crescimento e população de GC 03-0850. . . .	55
Tabela 9	Tempo de adaptação, crescimento e população de GC 03-2922. . . .	56
Tabela 10	Tempo de adaptação, crescimento e população de GC H37Rv.	56

LISTA DE ABREVIATURAS E SIGLAS

f.d.p	Função de densidade de probabilidade
GU	Growth Units
MGIT	Mycobacteria Growth Indicator Tube
QS	Quorum sensing
TB	Tuberculose
v.a	Variável aleatória
WHO	World Health Organization

SUMÁRIO

1	INTRODUÇÃO	14
1.1	Objetivos	15
1.2	Estrutura do Texto	15
2	REFERENCIAL TEÓRICO	16
2.1	Tuberculose	16
2.1.1	Curva de Crescimento Bacteriano	16
2.1.2	<i>Quorum sensing</i>	17
2.1.3	Curva de Crescimento do <i>Mycobacterium tuberculosis</i>	18
2.2	Variáveis Aleatórias	19
2.2.1	Modelos Probabilísticos	20
2.3	Sistema Baseado em Agentes	25
2.3.1	Linguagem NetLogo	27
2.4	Trabalhos Relacionados	29
3	MODELO DESENVOLVIDO	32
3.1	Modelo de Crescimento da Curva do <i>Mycobacterium tuberculosis</i>	32
3.2	Implementação do Modelo	36
3.3	Estudo de Sensibilidade em Relação as Principais Variáveis que Modelam a Curva de Crescimento	40
3.3.1	Funções Vitais	40
3.3.2	Energia de Reprodução	41
3.3.3	Sensor Bactéria	42
3.4	Interpretação dos Valores	44
4	ANÁLISE DOS RESULTADOS	45
4.1	Parametrização do Ambiente	45
4.2	Resultados das Simulações para Cada Cepa	48
4.3	Validação do Modelo	52
4.4	Comparativo com os Resultados Obtidos por WERLANG (2013)	57
5	CONCLUSÕES E TRABALHOS FUTUROS	61
	REFERÊNCIAS	64

1 INTRODUÇÃO

A tuberculose (TB) é um grande problema de saúde pública, que afeta predominantemente países de baixa e média renda, se desenvolvendo também entre imigrantes, e partes mais pobres e vulneráveis de países ricos (LÖNNROTH et al., 2015). Segundo BURGOS; PYM (2002), o *Mycobacterium tuberculosis*, causador da TB, é um dos patógenos bacterianos mais bem sucedidos na história da humanidade.

Diante dessas circunstâncias, o estudo da curva de crescimento do *Mycobacterium tuberculosis* torna-se extremamente importante, já que através desse estudo podem-se testar hipóteses, verificar reações do bacilo a fármacos e ainda pode ajudar no desenvolvimento de novos (GROLL, 2010).

No entanto, o bacilo da tuberculose possui uma taxa de crescimento populacional muito lenta, esse comportamento faz com que experimentos *in vitro* sejam algo demorado e de alto custo.

Para REBONATTO (2000), métodos de simulações computacionais têm se apresentado eficientes em situações que envolvam custos e riscos elevados. As simulações computacionais permitem o estudo dos mais diversos problemas com eficácia, já que possibilitam, na maioria das vezes, a visualização do comportamento, e de detalhes específicos do objeto de estudo.

Neste contexto, surgem os sistemas baseados em agentes, área da inteligência artificial que permite, por meios de suas ferramentas, simular regras de comportamento de determinado sistema. De acordo com GARCIA; SICHMAN (2005), agentes são personagens computacionais que atuam de acordo com o programa definido, direta ou indiretamente, por um usuário. Eles podem atuar sozinhos ou em comunidades, formando sistemas multiagentes.

Muitos fenômenos mensuráveis presentes na natureza têm distribuição de probabilidades semelhantes a alguns modelos probabilísticos. Frequentemente, esses modelos são usados para representar a função de densidade de probabilidade de variáveis aleatórias.

Os modelos probabilísticos são úteis em muitas situações reais, para fazer previsões da variável em estudo e auxiliar no apoio a decisões. Dessa forma, acredita-se que as principais variáveis que modelam a curva de crescimento do bacilo *Mycobacterium tu-*

berculosis também se assemelham com algum modelo probabilístico.

Este trabalho modela a curva de crescimento do bacilo da TB, utilizando simulação baseada em agentes, onde os valores das variáveis dos agentes são gerados pelo modelo probabilístico de Gauss, também conhecido como distribuição normal.

1.1 Objetivos

O objetivo geral deste trabalho é modelar a curva de crescimento do bacilo da tuberculose, utilizando sistemas baseados em agentes, onde as variáveis da simulação utilizam distribuição de probabilidades, tornando assim, o modelo de crescimento desenvolvido mais similar ao modelo de crescimento real.

O modelo também possui como objetivo obter resultados que mantenham o mesmo nível de precisão que dos experimentos *in vitro*. Dessa maneira, o mesmo poderá ser útil para validar experimentos e testar hipóteses. No entanto, é necessário que a curva de crescimento modelada seja fiel à curva real.

Para atingir os objetivos principais, os seguintes objetivos específicos devem ser alcançados:

- Estudar modelos de distribuições de probabilidades;
- Estudar modelagem e simulação baseada em agentes;
- Implementar o modelo proposto em ferramenta para simulação baseada em agentes;
- Validar os resultados obtidos com a simulação baseada em agentes comparando-os com os resultados obtidos por experimento *in vitro*.

1.2 Estrutura do Texto

O texto foi estruturado em 5 capítulos. No capítulo 2 é apresentada a fundamentação teórica para realização deste trabalho. No capítulo 3 está a proposta do trabalho: como foi feita a modelagem da curva de crescimento do *Mycobacterium tuberculosis*, explícita a inserção de cada variável no modelo e a maneira como cada uma influencia no desenvolvimento da curva. No capítulo 4 é mostrado o ambiente elaborado, os resultados provenientes do modelo e ainda uma comparação com dados reais. Por fim, no capítulo 5 são apresentadas as principais considerações deste trabalho e algumas sugestões para pesquisas futuras, que podem ser desenvolvidas para dar continuidade ao estudo.

2 REFERENCIAL TEÓRICO

2.1 Tuberculose

A tuberculose é uma doença infectocontagiosa transmitida pelo *Mycobacterium tuberculosis*. A doença foi descoberta em 24 de março de 1882 pelo cientista alemão Robert Koch, e por isso o bacilo da tuberculose, *Mycobacterium tuberculosis*, também é conhecido como Bacilo de Koch.

O *Mycobacterium tuberculosis* é um patógeno intracelular que pode afetar várias espécies animais, embora os seres humanos sejam os principais hospedeiros. Ele cresce com mais sucesso em tecidos que contenham níveis elevados de oxigênio, como os pulmões (LAWN; ZUMLA, 2011).

A TB é uma doença antiga que continua a ser um dos maiores problemas de saúde em todo mundo. Durante a história da humanidade, ela recebeu diferentes nomes e foi frequentemente associada com altas taxas de contágio e mortalidade (GROLL, 2010).

PAIXÃO; GONTIJO (2007) relatam que embora prevenível e tratável com medicamentos, a tuberculose vem se apresentando com intensidade de contaminação, com repercussões nos níveis de saúde e mortalidade.

Em um relatório publicado em 2015 pela *World Health Organization* (WHO, 2015), consta que no ano 2014 foram 9,6 milhões de novos casos de TB e 1,5 milhões de mortes, e ainda, juntamente com o HIV, a tuberculose apresenta-se como um dos maiores causadores de mortes por doenças infecciosas.

No Brasil, a tuberculose é um sério problema da saúde pública, com profundas raízes sociais. A cada ano, são notificados aproximadamente 70 mil casos novos e ocorrem 4,6 mil mortes em decorrência da doença. O Brasil ocupa o 17º lugar entre os 22 países responsáveis por 80% do total de casos de tuberculose no mundo (SAÚDE, 2016).

2.1.1 Curva de Crescimento Bacteriano

Quando o *Mycobacterium tuberculosis* é inoculado em um meio que contenha todos os nutrientes necessários para sua sobrevivência, as bactérias tendem a duplicar-se. Inicialmente, elas ajustam-se ao novo meio (fase de latência) até que elas possam começar o

processo de divisão regularmente (fase exponencial). Quando seu crescimento torna-se limitado, as células param de se dividir (fase estacionária), até que finalmente eles morrem pela inviabilidade do ambiente (TODAR, 2013), conforme apresenta a Figura 1.

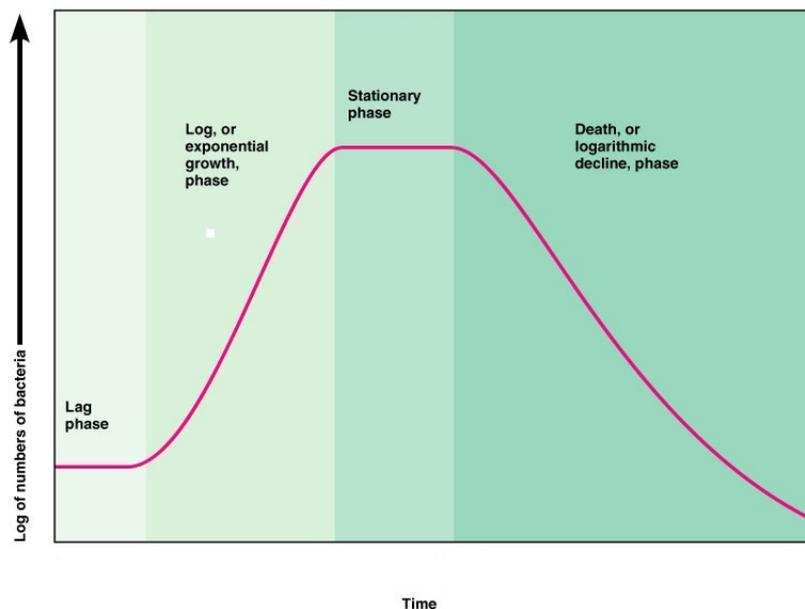


Figura 1: Curva de crescimento e suas fases (TODAR, 2013).

1. Fase de Latência: Imediatamente após a inoculação em um meio, a população permanece temporariamente inalterada. Embora não haja aparente divisão celular, as bactérias podem estar crescendo em volume ou massa, sintetizando enzimas, proteínas e RNA.
2. Fase Exponencial: Logo após a adaptação elas sofrem um crescimento rápido e consomem o máximo de nutrientes do ambiente. Também nessa fase, liberam resíduos e moléculas químicas sinalizadoras no ambiente.
3. Fase Estacionária: Após o crescente aumento populacional, as bactérias começam a parar de crescer devido a exaustão de nutrientes disponíveis, acúmulo de resíduos no ambiente e pela falta de espaço. Nessa fase, elas entram em fase de dormência onde poupam energia para aumentar o tempo de sobrevivência.
4. Fase de Decréscimo (morte): A última fase da curva é a de morte, onde elas começam a morrer por falta de nutrientes.

2.1.2 *Quorum sensing*

Por muitos anos, pesquisadores acreditavam que as bactérias existiam como células individuais, agiam como populações de células independentes e se reproduziam quando

em condições favoráveis. No entanto, nas últimas décadas, diversos estudos mostraram que esses microrganismos podem se comunicar através de sinais (ANTUNES, 2003).

Os sinais são denominados *Quorum sensing* (QS). O processo baseia-se na produção e difusão de pequenas moléculas sinalizadoras, que podem ser detectadas pelas bactérias. Esse processo acontece quando há percepção de alta densidade celular, possibilitando a toda população iniciar uma ação, uma vez que a concentração crítica tenha sido alcançada (WHITEHEAD et al., 2001).

Esse sistema de sinalização já foi identificado em muitos gêneros bacterianos e permite a coordenação do comportamento em relação ao meio ambiente, regulando a expressão de genes especializados, em resposta à densidade populacional, além da intervenção em diversos processos fisiológicos.

Quando essa sinalização atinge o nível crítico, os microrganismos passam a agir como um único organismo multicelular, sendo capaz de organizar respostas unificadas favoráveis a sobrevivência da população (RUMJANEK; FONSECA; XAVIER, 2004).

O estudo do *Quorum sensing* representa não só uma possibilidade para o conhecimento mais aprofundado de organismos com os quais convivemos, seja de forma amistosa ou competitiva, durante toda a nossa vida, mas também uma forma promissora para o controle de doenças infecciosas que representam um grande risco à saúde e ao bem-estar (ANTUNES, 2003).

2.1.3 Curva de Crescimento do *Mycobacterium tuberculosis*

GROLL (2010) padronizou um método para determinação da curva de crescimento baseado em um sistema que obtém o crescimento bacteriano através do monitoramento do meio líquido, por meio de um sensor de oxigênio que emite fluorescência, sendo que este ambiente é limitado e controlado. As curvas de crescimento foram geradas através do MGIT (*Mycobacteria Growth Indicator Tube*), em português, tubo indicador de crescimento bacteriano.

Os dados obtidos são referentes as cepas que tem origem em regiões geográficas distintas e são resistentes a diferentes fármacos, conforme Tabela 1.

Tabela 1: Cepas utilizadas e suas origens (GROLL, 2010).

Identificação das cepas	Origens
GC 01-2522	Georgia
GC 02-2761	Bangladesh
GC 03-0850	Georgia
GC 03-2922	Georgia
H37Rv	ATCC

Para cada uma das cepas foram feitos dois experimentos com mesma solubilidade. O monitoramento foi realizado por 25 dias e a cada hora o equipamento media o crescimento bacteriano, este expresso por unidades de crescimento, Growth Units (GU). A Figura 2, expressa a média dos experimentos de cada cepa.

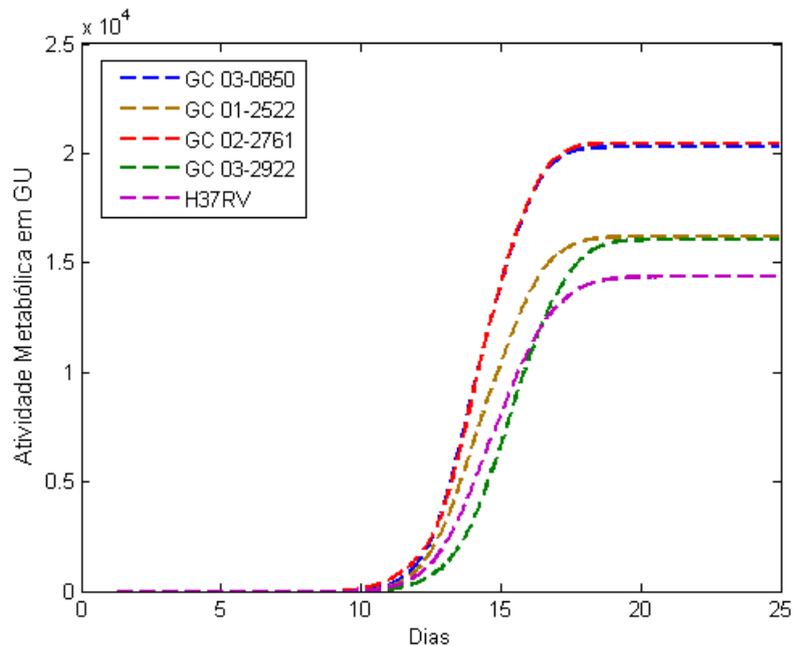


Figura 2: Curvas reais obtidas em experimento (GROLL, 2010).

No entanto com o método utilizado não é possível obter por quanto tempo as bactérias mantiveram-se na fase estacionária ou a taxa de decréscimo da população na fase de morte. Isso porque a luminosidade provém da atividade metabólica das mesmas, logo só pode ser monitorada enquanto houver crescimento populacional.

2.2 Variáveis Aleatórias

Em um experimento há diversas características que podem ser observadas ou medidas. Geralmente, cada resultado de um experimento é associado a um número, especificando-se uma regra de associação.

Essa regra de associação é denominada variável aleatória (v.a), variável porque é possível obter diferentes valores numéricos, e aleatória porque o valor esperado depende de qual dos resultados possíveis do experimento é obtido (DEVORE; SILVA, 2006).

Uma variável aleatória X representa um valor numérico, cuja mensuração está associada a cada um dos resultados de um experimento probabilístico. Segundo LARSON; FARBER (2005), essa variável aleatória será contínua se houver um número incontável de resultados possíveis.

Para cada variável aleatória contínua X há duas funções que podem ser associadas: a

função de densidade de probabilidade (f.d.p) ou distribuição de probabilidade, simbolizada por $f(x)$, e a função de distribuição acumulada, representada por $F(x)$.

Qualquer função só pode ser considerada como uma função de densidade de probabilidade se forem satisfeitas as condições apresentadas nas Equações 1 e 2.

$$f(x) \geq 0, \quad (1)$$

e

$$\int_{-\infty}^{+\infty} f(x)dx = 1. \quad (2)$$

A função $f(x)$ é aquela cuja integral de a até b , com $b > a$, tem como resultado a probabilidade de que X assuma valores compreendidos no intervalo $[a, b]$, conforme Equação 3.

$$P(a \leq X \leq b) = \int_a^b f(x)dx \quad (3)$$

A função de distribuição acumulada $F(x)$ é apresentada na Equação 4.

$$P(X \leq x) = \int_{-\infty}^x f(x)dx \quad (4)$$

Ao se calcular a média de uma distribuição de probabilidades, obtém-se o valor médio que esperaríamos ter se pudéssemos repetir as provas indefinitivamente. Ela representa o valor provável de uma v.a, por isso, muitas vezes, é chamada de valor esperado ou esperança, conforme Equação 5.

$$E(X) = \int_{-\infty}^{\infty} xf(x)dx. \quad (5)$$

Já o desvio-padrão apresenta uma medida de quanto a distribuição de probabilidades se dispersa em torno da média. Um grande desvio-padrão reflete dispersão considerável, enquanto que um desvio-padrão menor traduz menor variabilidade, com valores relativamente mais próximos da média.

$$DP(X) = E(X^2) - [E(X)]^2. \quad (6)$$

2.2.1 Modelos Probabilísticos

Os modelos probabilísticos são úteis para representar situações reais, ou para descrever um experimento aleatório. Os estudos de simulação tentam reproduzir num ambiente controlado o que se passa em um ambiente real (BUSSAB; MORETTIN, 2005)

As distribuições de frequências construídas a partir de observações podem ser representadas através de fórmulas matemáticas. Essas fórmulas são utilizadas para a idealização dos dados reais e são chamadas de distribuições teóricas.

Dada uma variável aleatória estudada, interessa saber qual função de densidade de probabilidade da mesma. Em teoria das probabilidades existem muitos modelos probabilísticos prontos que descrevem a distribuição de probabilidades de diversas variáveis.

Entre os modelos mais simples e mais importantes, respectivamente, estão a distribuição uniforme e a distribuição normal, sendo que a última é aplicável em diversos fenômenos e constantemente utilizada para o desenvolvimento teórico da inferência estatística. Para DEVORE; SILVA (2006), a distribuição normal é um modelo fundamental em probabilidade e inferência estatística.

2.2.1.1 O modelo Uniforme

É o modelo mais simples para variáveis aleatórias contínuas. Nesse modelo, a probabilidade de gerar qualquer ponto em um intervalo é proporcional ao tamanho dele. A distribuição uniforme é muito usada em situações em que não há razão para atribuir probabilidades diferentes a um conjunto de possíveis valores da v.a estudada.

i) Definição: Suponha que X seja distribuído uniformemente em um intervalo $[\alpha, \beta]$ logo sua função de densidade de probabilidade é pela Equação 7.

$$f(x; \alpha, \beta) = \begin{cases} \frac{1}{\beta - \alpha}, & \text{se } \alpha \leq x \leq \beta, \\ 0, & \text{caso contrário.} \end{cases} \quad (7)$$

ii) Gráfico: ver Figura 3.

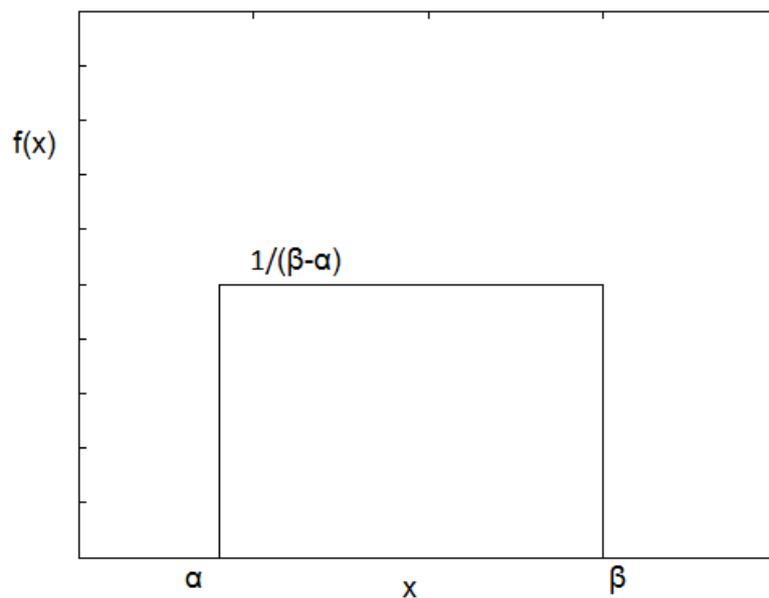


Figura 3: Distribuição uniforme no intervalo $[\alpha, \beta]$.

iii) Momentos: ver Equações 8 e 9.

$$E(X) = \frac{\alpha + \beta}{2}, \quad (8)$$

$$Var(X) = \frac{(\beta - \alpha)^2}{12}. \quad (9)$$

iv) Função de distribuição acumulada: ver Equação 10.

$$P(X \leq x) = \int_{-\infty}^x f(x)dx = \begin{cases} 0, & \text{se } x < \alpha, \\ \frac{1}{\beta - \alpha}, & \text{se } \alpha \leq x \leq \beta, \\ 1, & \text{se } x \geq \beta. \end{cases} \quad (10)$$

2.2.1.2 O modelo Normal

É umas das mais importantes distribuições de probabilidades, sendo aplicável em diversos fenômenos e constantemente utilizada para o desenvolvimento teórico da inferência estatística.

i) Definição: Diz-se que a variável aleatória X tem distribuição normal com parâmetros μ (média de X) e σ (desvio padrão de X), se sua função de densidade é dada pela Equação 11.

$$f(x; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \frac{1}{2\sigma^2}(x - \mu)^2, -\infty < x < \infty. \quad (11)$$

ii) Gráfico: ver Figura 4.

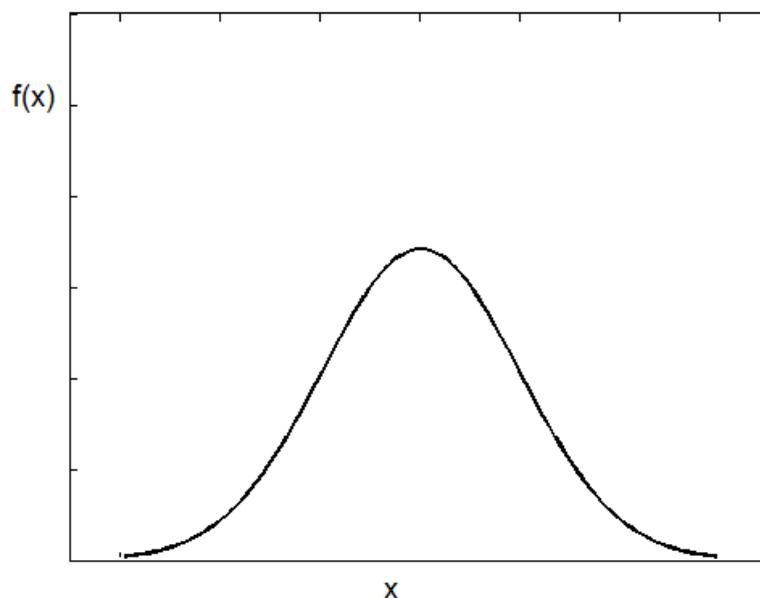


Figura 4: F.d.p de uma variável aleatória normal como média μ e desvio padrão σ .

iii) Momentos: ver Equações 12 e 13.

$$E(x) = \mu, \quad (12)$$

$$Var(x) = \sigma^2. \quad (13)$$

iv) Função de distribuição acumulada: ver Equação 14.

$$P(X \leq x) = \int_{-\infty}^x f(x)dx = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^x \exp \frac{1}{2\sigma^2}(x - \mu)^2. \quad (14)$$

A integral da Equação 14 não pode ser calculada analiticamente, e portanto a probabilidade indicada só poderá ser obtida, aproximadamente, por meio de integração numérica. Por essa razão usa-se a transformação da Equação 15.

$$Z = \frac{X - \mu}{\sigma} \quad (15)$$

O cálculo da variável reduzida Z faz uma transformação dos valores reais em valores codificados. Uma vez calculada a variável reduzida Z, consulta-se a tabela normal padronizada para identificar a probabilidade acumulada à esquerda de Z, ou seja, a probabilidade de ocorrerem valores menores ou iguais ao valor de Z consultado.

Sensibilidade a variações dos parâmetros do modelo:

A distribuição normal é unimodal, simétrica em relação à sua média e tende cada vez mais ao eixo horizontal à medida que se afasta da média (ver Figura 5).

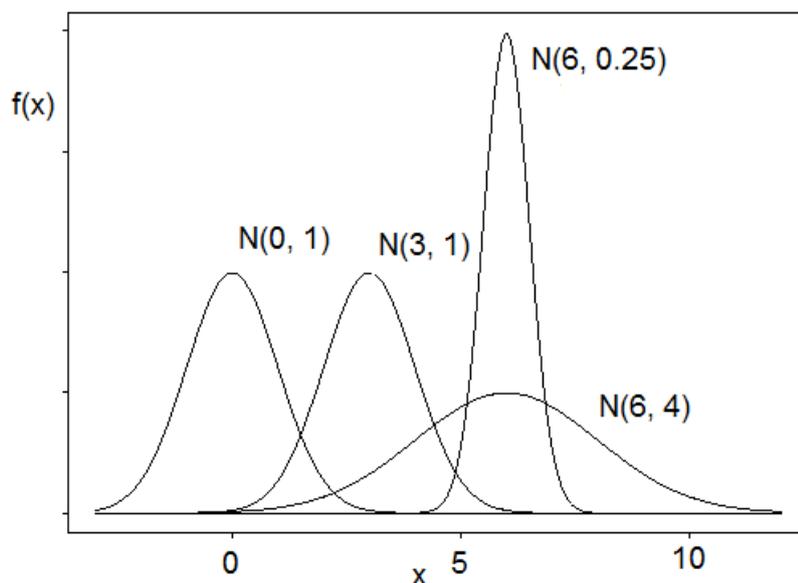


Figura 5: Distribuições normais com variações na μ e no σ .

1. Da distribuição $N(0,1)$ para a $N(3,1)$ muda a tendência central, mas a variabilidade é constante;

2. Da distribuição $N(6,0,25)$ para a $N(6,4)$ muda a variabilidade, mas a tendência central é constante;
3. Entre outras combinações das distribuições da Figura 5, mudam as tendências centrais e as variabilidades.

2.2.1.3 Simulações

Simulação é um ensaio realizado com o auxílio de modelos durante determinado período de tempo, de fenômeno, situação ou processo do mundo real. A simulação envolve a geração de uma história artificial de um sistema, podendo-se, a partir dela, estimar como o sistema funciona em situação real.

Os frequências dos histogramas das Figuras 6 e 7 foram obtidas computacionalmente. Para cada um deles, foi gerado uma quantidade de números aleatórios distribuídos por modelos probabilísticos.

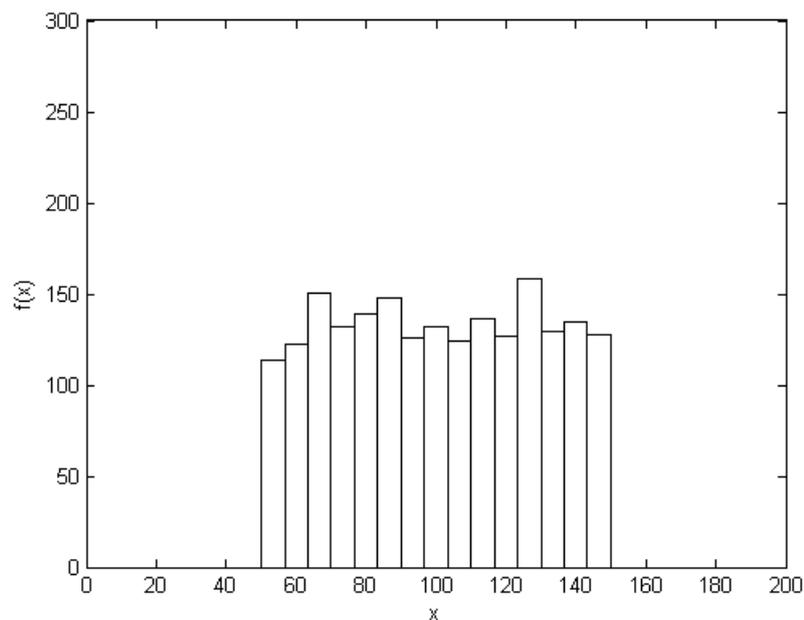


Figura 6: Simulação de variável com distribuição uniforme.

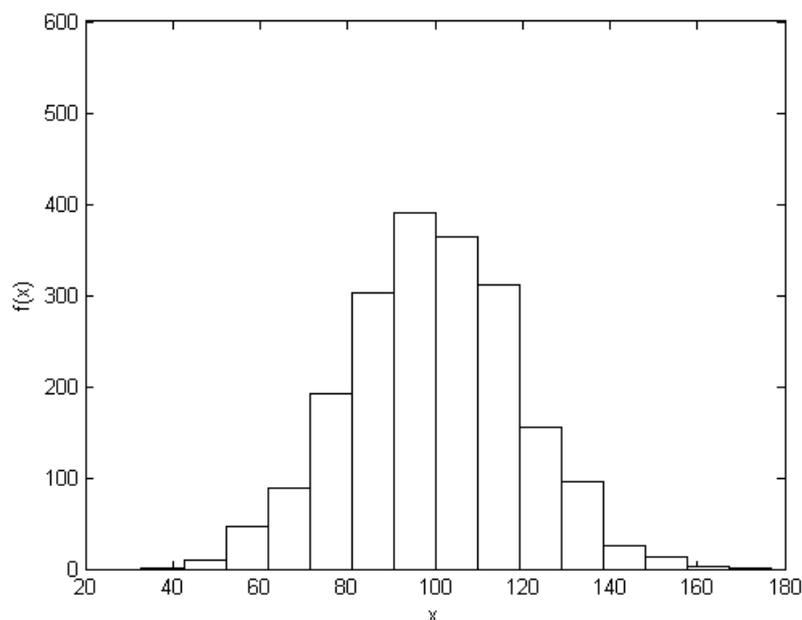


Figura 7: Simulação de variável com distribuição normal.

Os dados da Figura 6 foram distribuídos uniformemente, ou seja, todos os possíveis valores da v.a tinham a mesma probabilidade de ocorrer. Já os dados da Figura 7 são distribuídos normalmente, sendo atribuídas diferentes probabilidades ao conjunto de possíveis valores da v.a estudada.

2.3 Sistema Baseado em Agentes

A simulação computacional é uma das ferramentas mais viáveis disponíveis para projetar, planejar, controlar e avaliar novas alternativas e/ou mudanças de estratégias em sistemas do mundo real. Para fazer uso da simulação computacional, necessita-se de softwares que representem o sistema do mundo real em questão e copiem seu funcionamento (REBONATTO, 2000).

Para AZEVEDO (2007), o uso de simulações no computador tem sido bastante apropriado para descrever sistemas reais, pois de modo geral, a simulação consiste em reproduzir artificialmente uma situação ou fenômeno natural, no qual hipóteses podem ser verificadas, porém com muito mais facilidade e sem riscos.

Um importante tipo de simulação é a modelagem baseada em agentes. Nesse tipo de simulação é possível recriar uma população de um sistema real em um artificial, onde cada indivíduo da população é representado por um agente e cada um deles possui um conjunto de regras específicas que ditam seu comportamento, como eles irão interagir uns com os outros e com o ambiente onde estão inseridos (COLELLA; KLOPFER; RESNICK, 2001).

Em um sistema de agentes, um fenômeno real é decomposto em um conjunto de elementos e suas interações. Cada elemento é modelado como um agente e o modelo

geral é o resultado das interações entre os agentes. STRACK(1984), conforme citado por ADAMATTI (2011) relata que a simulação pode ser dividida em três etapas:

- Etapa da Modelagem: construir o modelo do fenômeno a ser estudado;
- Etapa de Experimento: aplicar variações sobre o modelo construído, alterando parâmetros que influenciam no processo de resolução;
- Etapa de Validação: comparar dados experimentais obtidos com o modelo e a realidade, permitindo a análise dos resultados obtidos.

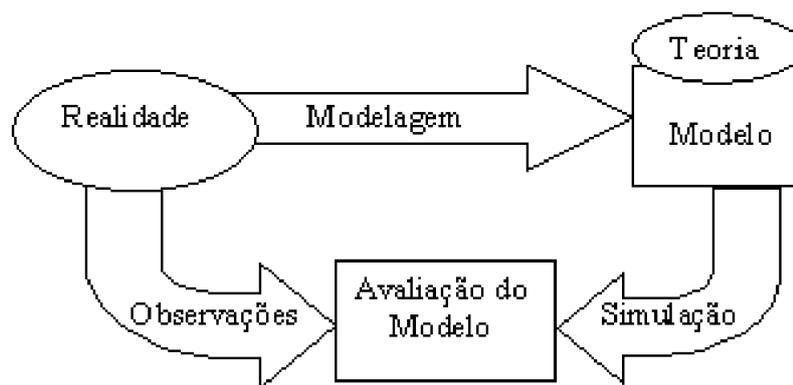


Figura 8: Etapas de um processo de simulação (ADAMATTI, 2011).

Para simular computacionalmente um problema real, faz-se necessário um estudo rigoroso para poder abstrair todas as variáveis e relações que definem o modelo. Esse estudo geralmente é feito por meio de observação e análise do fenômeno real.

Capturar todos os componentes do modelo de simulação não é uma tarefa fácil, e quanto maior o número de variáveis e mais detalhado for o modelo a ser simulado, maior será o trabalho computacional.

Depois de definido o modelo e feita a simulação, os resultados obtidos são comparados com os observados no fenômeno natural, com a finalidade de avaliar sua equivalência, ou seja, seu compromisso com a realidade.

A escolha por simulação baseada em agentes para a modelagem da curva de crescimento do *Mycobacterium tuberculosis* se deu pela necessidade de inserção de diferentes comportamentos entre os agentes, pelo dinamismo que esse tipo de modelagem pode possibilitar, como a interação entre agentes, entre agentes e ambiente, além da flexibilidade do modelo a modificações e ampliações.

Para criar uma simulação baseada em agentes, faz-se necessário dispor de ferramentas capazes de simular fenômenos. Existem diversos ambientes de programação projetados especialmente para a modelagem baseada em agentes, entre eles o NetLogo, que será o ambiente usado neste trabalho para simular o modelo de crescimento do bacilo da tuberculose.

2.3.1 Linguagem NetLogo

O NetLogo (Figura 9) é um ambiente de modelagem programável, e tem por finalidade analisar o comportamento de fenômenos, especialmente os naturais e sociais. O NetLogo possui uma linguagem de programação simples, sendo facilmente aprendida por pesquisadores de áreas diversas da computação e matemática (AZEVEDO, 2007).

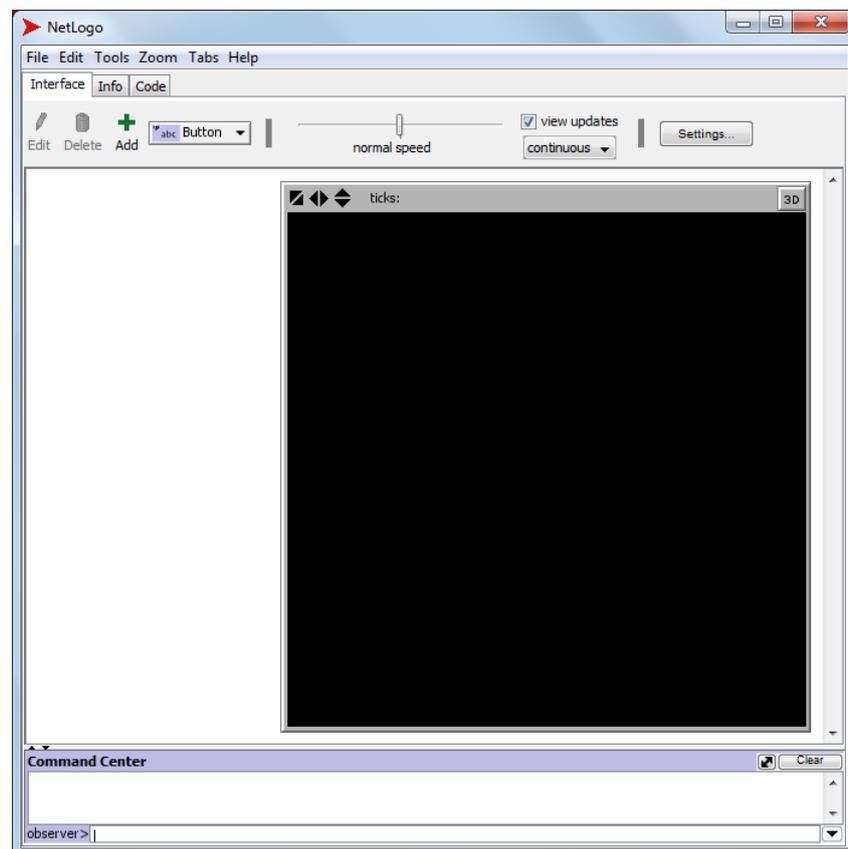


Figura 9: Tela de abertura NetLogo.

O ambiente NetLogo apresenta na sua interface principal três abas: *Interface*, *Info* e *Code*, sendo que elas não podem ser mostradas simultaneamente, mas basta clicar na aba desejada que haverá mudança.

A aba de *Interface* é onde se visualiza a execução do modelo e também o comportamento dos agentes. A barra de ferramentas da janela de interface contém botões que permitem ao usuário criar *buttons*, *sliders*, *switchs*, *monitors* ou *plots*. Veja a Figura 10.

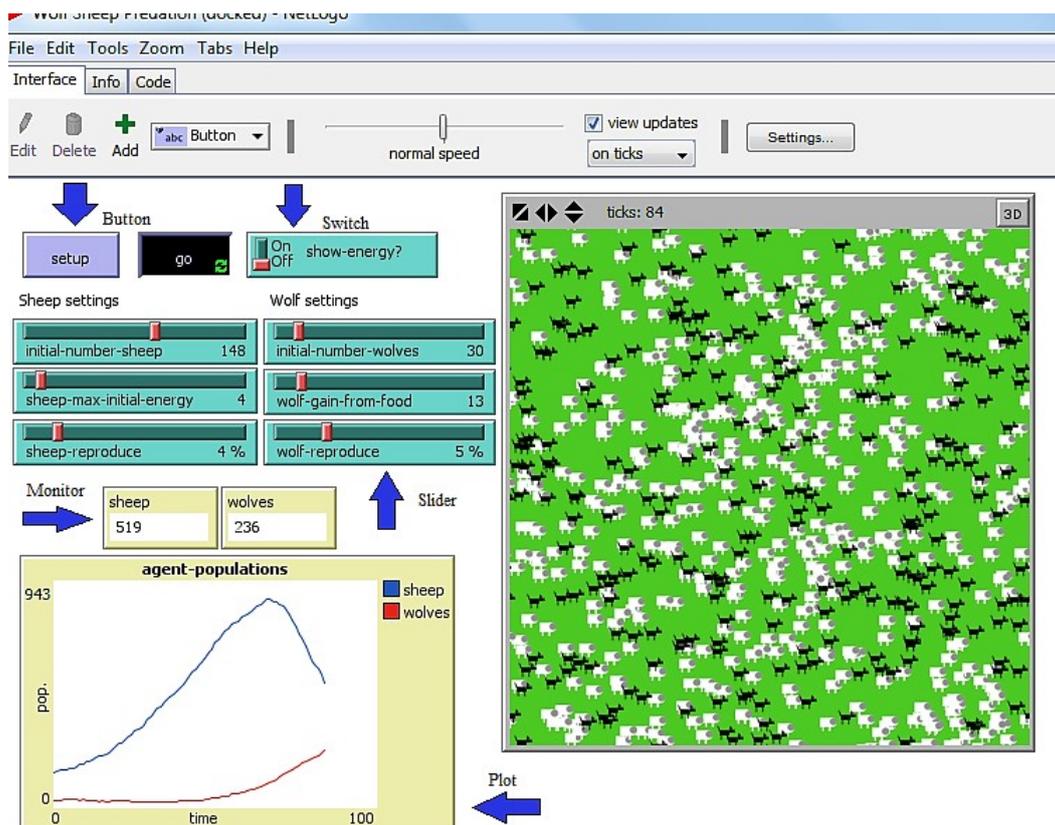


Figura 10: Wolf Sheep Predation, livraria NetLogo (WILENSKY, 1999).

O *Button* serve para iniciar um conjunto de procedimentos.

Os *Sliders* oferecem a possibilidade de inserção de parâmetros que possam ser facilmente ajustados, basta mover a barra central do *slider* e os valores são modificados sem necessidade de alterar o programa e gravar o procedimento.

Plots são usados para mostrar gráficos gerados por dados do modelo em tempo real.

Os *Switches* são ferramentas visuais de verdadeiro ou falso para as variáveis. O usuário tem a opção de realizar ou não a execução de um evento no modelo.

Através dos *Monitors* é possível mostrar qualquer expressão ou simplesmente pode ser usado como um contador em tempo real de simulação.

A aba *info* é utilizada para descrever uma introdução do modelo, explicar seu funcionamento, modo de usar, o que pode ser explorado, possíveis extensões e características sobre o modelo e o ambiente. Ela é mais de efeito informativo, uma primeira leitura é recomendada antes de iniciar a exploração do modelo.

Já a aba *Code* é onde se escreve o programa do modelo (Figura 11). É através da escrita do código que se definem as regras de comportamento dos agentes.

```

;; Sheep and wolves are both breeds of turtle.
breed [sheep a-sheep] ;; sheep is its own plural, so we use "a-sheep" as the singular.
breed [wolves wolf]
turtles-own [energy] ;; both wolves and sheep have energy

to setup
  clear-all
  ask patches [ set pcolor green ]
  set-default-shape sheep "sheep"
  create-sheep initial-number-sheep ;; create the sheep, then initialize their variables
  [
    set color white
    set size 1.5 ;; easier to see
    set label-color blue - 2
    set energy 1 + random sheep-max-initial-energy
    setxy random-xcor random-ycor
  ]
  set-default-shape wolves "wolf"
  create-wolves initial-number-wolves ;; create the wolves, then initialize their variables
  [
    set color black
    set size 1.5 ;; easier to see
    set energy random (2 * wolf-gain-from-food)
    setxy random-xcor random-ycor
  ]
  display-labels
  reset-ticks
end

to go
  if not any? turtles [ stop ]
  ask sheep [
    move

```

Figura 11: Exemplo de código: Wolf Sheep Predation, livraria NetLogo (WILENSKY, 1999).

A programação em NetLogo consiste em atribuir ações aos agentes, que são os *patches* e as *turtles*. A linguagem de programação apresenta diversos recursos, como atribuir diferentes formas (*shapes*) para os agentes (*turtles*).

A programação NetLogo também permite definir diferentes espécies (*breeds*) para os agentes, sendo essa uma característica importante, pois uma vez que se tenha definido as espécies, torna-se possível programar para que tenham comportamentos diferenciados.

Pelo fato do NetLogo conseguir fazer distinção entre espécies, é possível que se crie regras de comportamento para uma determinada espécie de uma só vez, visualizando o programa da Figura 12, o usuário define que todos da espécie *sheep* tem formato de uma ovelha branca de tamanho 1,5.

2.4 Trabalhos Relacionados

Uma variedade de técnicas estatísticas e modelos matemáticos são utilizadas com o objetivo de simular o comportamento de variáveis aleatórias. Os modelos probabilísticos são úteis para representar situações reais, ou então para descrever um experimento aleatório.

Com o objetivo de buscar uma função de densidade de probabilidade que mais adequadamente descrevesse a condutividade hidráulica do solo, MESQUITA et al. (2002) verificaram o ajuste das distribuições normal, lognormal, gama e beta. Para os autores a

condutividade hidráulica saturada do solo (K_{sat}) é uma das variáveis de maior importância para estudos de movimento de água e solutos no solo. E ainda, sua determinação em laboratório e campo produz resultados com elevada dispersão, o que pode indicar que essa variável não possui distribuição simétrica. O experimento consistiu em determinar a condutividade hidráulica saturada em amostras de três solos com diferentes texturas da região centro oeste do estado de São Paulo, e submeter os resultados a testes estatísticos para identificação da distribuição assimétrica mais adequada para representá-los. Depois de realizados os testes, concluiu-se que a K_{sat} não apresentou normalidade na distribuição e ajustou-se as distribuições lognormal, gama, beta, sendo que a função de densidade de probabilidade lognormal foi a mais indicada para descrever os dados da variável em estudo.

Outro trabalho que busca aderência de funções de densidade de probabilidade a dados amostrais é o de CARGNELUTTI FILHO; MATZENAUER; TRINDADE (2004). Neste trabalho, são verificados o ajuste da série de dados de radiação solar global média decenal de 22 municípios do Estado do Rio Grande do Sul, as funções de distribuições de probabilidade normal, lognormal, gama, gumbel e weibull. Para os autores, o conhecimento do comportamento da precipitação, temperatura e umidade relativa do ar, evaporação, direção e velocidade do vento, radiação solar global, ocorrência de orvalho, nevoeiro, granizo, geada e neve, é um importante instrumento na tomada de decisões relacionadas às atividades agropecuárias, sendo que o estudo da radiação solar é fundamental. Os dados utilizados nos testes foram gerados em 47 anos de observação diária (1956 a 2003). Estimou-se a média decenal dos 36 decênios do ano, formando 792 séries temporais. Para verificar a aderência do conjunto de dados, os autores fizeram o teste de Kolmogorov-Smirnov, comparando as probabilidades empíricas do conjunto de dados amostrais com as probabilidades geradas pela distribuição em teste.

CARGNELUTTI FILHO; MATZENAUER; TRINDADE (2004) relatam que os dados de radiação solar global média decenal se ajustaram às funções de distribuições de probabilidade normal, log-normal, gama, gumbel e weibull, e apresentaram melhor ajuste à função de distribuição de probabilidade normal.

Com os recentes avanços computacionais tornou-se possível analisar sistemas sociais e econômicos através de estudos de simulação, entre eles, a modelagem baseada em agentes. TERANO (2011) afirma que o método de simulação baseado em agentes é muito importante, uma vez que pode produzir resultados sem suposições ao contrário das abordagens convencionais.

WERLANG (2013) descreve um modelo de curvas de crescimento do *Mycobacterium tuberculosis* baseado em sistemas multiagentes. Para o autor, o estudo dessa bactéria é muito importante, pois possibilita o estudo de suas características e o desenvolvimento de novos fármacos. A realização de testes experimentais com esse bacilo são bastante demorados, levando pelo menos três semanas para mostrar algum resultado, e muitas

vezes falham por causa de contaminação ou desidratação do meio.

No modelo de WERLANG (2013), os agentes que representam as bactérias possuem características individuais, e essas extremamente importantes para que os agentes representem seu papel no ambiente e interajam com ele, da mesma maneira que o *Mycobacterium tuberculosis* interagiria em seu ambiente natural de vida. O autor relata que os resultados obtidos foram muito satisfatórios e as curvas encontradas alcançaram um nível de similaridade bem próximo as curvas reais. Por fim, o autor ainda ressalta que o modelo é bastante útil, uma vez que possibilita a verificação de hipóteses em horas em quanto às realizadas em testes *in vitro* levariam dias.

Na bibliografia pesquisada não foram encontrados trabalhos que utilizam simulação baseada em agentes e distribuições de probabilidades. No entanto, segundo PORCELLIS; MORAES; BERTIN (2015) a escolha pelos modelos, sejam baseados em equações ou em agentes, depende do quão detalhado deseja-se o sistema a ser simulado. E ainda os autores ressaltam que é possível que as duas trabalhem associadas na busca por um resultado geral do modelo baseado em equações aliado ao detalhamento do modelo baseado em agentes.

3 MODELO DESENVOLVIDO

Este capítulo apresenta o modelo de crescimento populacional do *Mycobacterium tuberculosis* baseado em agentes, em que os valores das variáveis tem distribuição de probabilidades. O capítulo explícita as regras de comportamento dos agentes, as variáveis e os parâmetros do ambiente, bem como a relação que cada um exerce no modelo.

3.1 Modelo de Crescimento da Curva do *Mycobacterium tuberculosis*

Quando as curvas de crescimento são obtidas através do MGIT (*Mycobacteria Growth Indicator Tube*), as informações que essas curvas apresentam são o produto de diversos fatores da dinâmica populacional. Logo, não é possível extrair informações isoladas, como quanto consomem ou quanto deixam de consumir depois de atingida a saturação do ambiente, ou ainda qual a proporção de moléculas sinalizadoras é necessária para decretar saturação.

Diante dessas circunstâncias, é necessário inferir quanto as variáveis que interferem no crescimento populacional.

Para simular a dinâmica populacional foi utilizado o ambiente de programação NetLogo. O modelo baseado em agentes implementado simula um ambiente onde os agentes representam o *Mycobacterium tuberculosis*.

Os agentes do modelo possuem regras específicas de comportamento, que são modeladas como variáveis dos agentes. Essas regras são essenciais para que eles representem seu papel no ambiente e interajam com ele.

A simulação tem como divisão temporal o *tick*. A cada *tick*, os agentes realizam uma ou mais ações, essas ações são modeladas por funções estabelecidas no modelo, sendo elas: *alimenta*, *prosegue*, *sinaliza* e *reproduz*.

Assim como muitos fenômenos mensuráveis presentes em nosso dia a dia tendem a se distribuir conforme alguns modelos probabilísticos teóricos, tem-se por hipótese que as principais variáveis que modelam a curva de crescimento do bacilo *Mycobacterium tuberculosis* também podem se distribuir conforme algumas delas.

E como muitas v.a biológicas se ajustam a distribuição normal (CALLEGARI-

JACQUES, 2003), ou seja, os valores centrais são mais frequentes e os extremos, mais raros, sendo os valores muito baixos tão pouco frequentes quanto os muito altos. Assumi-se que as variáveis do modelo também se distribuem normalmente. A seguir estão descritas as variáveis dos agentes.

- Tick de reprodução: Tempo que a bactéria leva para adaptar-se ao ambiente, enquanto esse limiar não é alcançado, ela não consegue reproduzir.
- Energia: Indica o quão saudável o agente está, ele acumula energia toda vez que consome nutrientes.
- Consumo: Quantidade de nutrientes que o agente consome do ambiente a cada *tick*.
- Funções vitais: Indicam a quantidade de energia que será subtraída da variável energia a cada *tick*.
- Energia de reprodução: No modelo proposto, cada agente só sofrerá o processo de divisão quando estiver saudável, ou seja, quando alcançar a quantidade de energia necessária para se reproduzir.
- Reduz energia: Quantidade de energia que o agente gasta para reproduzir-se.
- Sensor de bactérias: Quando a bactéria decreta situação de saturação, ela libera uma molécula sinalizadora, indicando a população que o ambiente está cheio e entra em estado de consumo reduzido.
- Sensor de sinais: Mesmo que bactéria não tenha liberado a molécula sinalizadora, ela pode entrar em consumo reduzido, por detetar um limiar de sinais no ambiente.
- Consumo no consumo reduzido: Esta variável define o quanto será a queda no valor de consumo dos nutrientes quando o agente entrar em estado de consumo reduzido.
- Funções vitais no consumo reduzido: Indicam a quantidade de energia que será gasta a cada *tick*, quando ela estiver em estado de consumo reduzido.

Os agentes estão inseridos num ambiente que é compartilhado por todos. Nesse ambiente eles encontram os nutrientes necessários para sobreviverem. Também nele, depositam os resíduos provenientes de sua metabolização.

Cada espaço do ambiente é chamado de *patch*. Em cada *patch*, há uma quantidade de nutrientes e resíduos. Os nutrientes são utilizados pelos agentes ao longo da simulação para manter suas funções vitais ativas e acumular energia. Os resíduos são depositados pelos agentes após a metabolização dos nutrientes.

Assim como nos experimentos *in vitro* feitos em laboratório, em que uma quantidade de bactérias são inoculadas em um meio, a inicialização do modelo é feita com um determinado número de agentes no ambiente.

Após o início da simulação os agentes começam a percorrer o ambiente em busca de nutrientes. No entanto, nesse início, essa busca tem como único objetivo a sobrevivência do agente, que está se adaptando ao meio ambiente e por isso impossibilitado de reproduzir-se. Esse mesmo, comportamento é observado em experimentos *in vitro*, em que as bactérias precisam de um certo tempo para ajustar suas funções metabólicas ao novo *habitat*.

Posterior ao período de adaptação o agente passa a executar normalmente todas as suas funções, incluindo a reprodução. Ele consome nutrientes e logo os transforma em energia, que será depositada em sua reserva para indicar o quão saudável ele está e para manter as funções vitais ativas.

Cada integrante do modelo tem um momento diferente para reproduzir-se. Esse momento indica a quantidade de energia que o agente deve ter disponível em sua reserva para executar a função *reproduz*. Logo a energia é um dos fatores limitantes para a reprodução.

Existe uma quantidade de energia para manter o agente a cada ciclo. Caso ele venha a consumir menos que o necessário para manter seu metabolismo, ele começará a gastar energia de sua reserva para poder sobreviver e o término desse recurso o levará a morte.

O acúmulo de resíduos nos *patches* faz com que eles tenham dificuldades em absorver os nutrientes. Logo torna-se mais difícil aos agentes sobreviverem em um ambiente muito saturado.

Um outro determinante na modelagem da curva é o sensor de bactérias. Essa variável determina com quantos agentes no ambiente será decretada situação de saturação. Quando esse limiar é atingido, os agentes liberam uma molécula sinalizadora, alertando aos outros que o ambiente esta cheio. Segundo WHITEHEAD et al. (2001), esse processo acontece quando há percepção de alta densidade celular, possibilitando a toda população iniciar uma ação, uma vez que a concentração crítica tem sido alcançada.

No modelo de crescimento proposto a ação realizada pelos agentes é a decisão em reduzir o consumo. Assim que o agente percebe a situação de saturação, além de liberar uma molécula sinalizadora, ele entra em estado de consumo reduzido, onde consome menos nutrientes, gera menos resíduos e se reproduz menos.

O consumo reduzido é uma variável que só pode receber dois valores, verdadeiro ou falso. Uma vez que o consumo reduzido seja verdadeiro, a bactéria reduz a quantidade de nutrientes que irá absorver do ambiente por *tick* e também diminui proporcionalmente a quantidade de energia necessária para manter-se viva.

Na inicialização do modelo, todos os agentes tem valor falso para consumo reduzido e ele só recebe verdadeiro quanto o agente detecta uma quantidade significativa de agentes ou sinais no ambiente. Essa informação é armazenada pelas variáveis sensor de sinais ou sensor de bactérias.

A redução de consumo de nutrientes e da energia para manter funções vitais do agente tem por objetivo fazer com que eles cresçam menos, e assim possam sobreviver por mais

tempo no ambiente (ver fluxograma do ciclo de vida dos agentes - Figura 12).

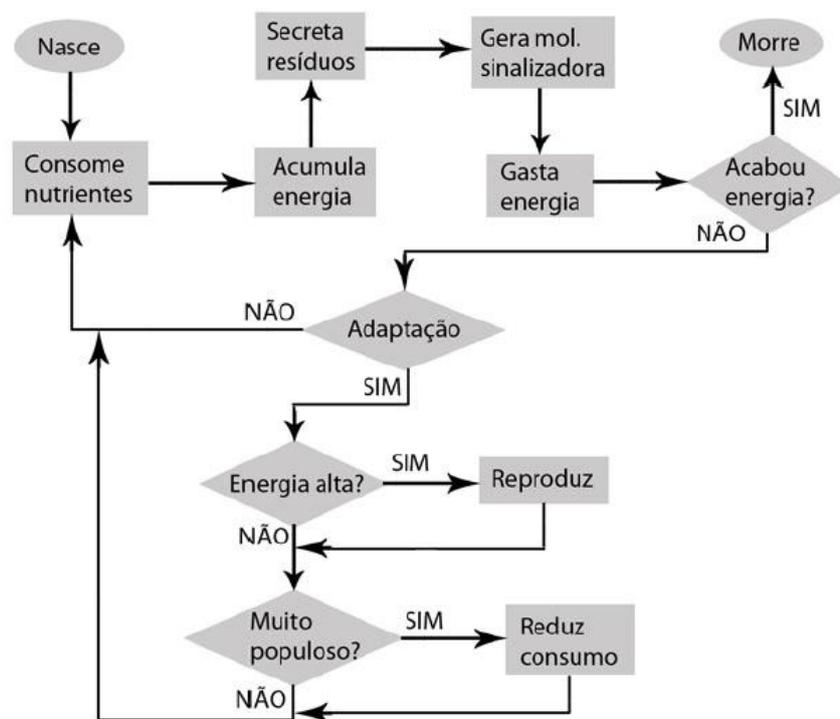


Figura 12: Fluxograma do ciclo de vida dos agentes (WERLANG, 2013).

O agente inicia o ciclo consumindo nutrientes do *patch* onde se encontra, a taxa de consumo é dada pela variável *consumo*. Os nutrientes absorvidos são transformados em energia que será depositada na reserva do agente. E após a metabolização dos nutrientes, ele secreta os resíduos provenientes no ambiente.

Depois de desempenhadas essas tarefas, uma quantidade de energia é subtraída da reserva do agente. Essa quantidade é dada pela variável *func_vitais*. Se a energia do agente, quantificada pela variável *energia*, for menor que 1, o agente morre; caso contrário, ele prossegue para as outras etapas da modelagem.

Passado o período de adaptação, estipulado pela variável *tick_reprodução*, o agente finalmente está apto a se reproduzir. No entanto, um fator limitante para a reprodução é a energia disponível na sua reserva, se a energia for maior ou igual a estabelecida pela variável *energia_reproduz*, o agente reproduz; se não, ele continua suas funções sem reproduzir-se, até obter a energia necessária.

A última etapa do ciclo consiste na verificação da densidade populacional do ambiente. Caso o agente decreta situação de saturação, dada pela variável *sensor_bacteria*, ele libera um molécula sinalizadora indicando aos outros que o ambiente está cheio e logo reduz seu consumo.

3.2 Implementação do Modelo

Depois de elaborada a dinâmica populacional, o próximo passo foi implementar o modelo em NetLogo. Através da intervenção do usuário, é possível configurar alguns parâmetros do ambiente, como o número inicial de agentes e a quantidade de nutrientes disponibilizados. A Figura 13 ilustra a interface do modelo criado e os parâmetros citados.

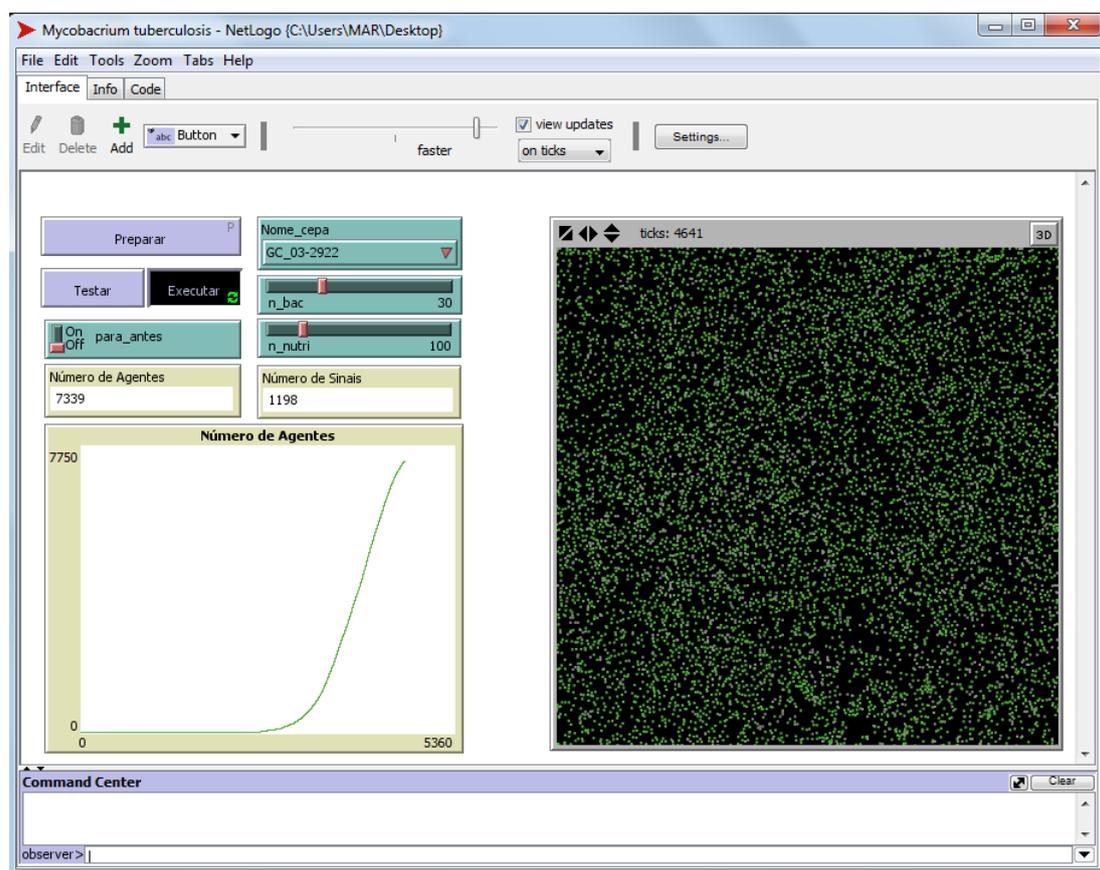


Figura 13: Interface do modelo desenvolvido no NetLogo.

Cada uma das cepas simuladas possuem diferentes momentos para iniciar suas fases. Dessa forma, cada uma recebe diferentes valores para as variáveis. Para não ter que criar cinco simulações, referentes a cada uma delas, foi criado um artifício em que as cinco simulações podem ser executadas em um mesmo código. Basta escolher a cepa que se deseja simular e iniciá-la no botão *Nome_cepa*.

Como já mencionado anteriormente, os resultados obtidos por GROLL (2010) representam as curvas de crescimento do *Mycobacterium tuberculosis* em apenas dois momentos, a adaptação e a fase exponencial. Com a finalidade de facilitar as comparações entre o crescimento real e o simulado, foi criada uma chave que possibilita ao usuário simular curvas com as mesmas etapas de crescimento, ou seja, assim que é detectado que a curva não está crescendo, a simulação é encerrada.

Para ampliar a compreensão entre a teoria e o modelo desenvolvido na prática, a se-

guir serão apresentados trechos do código fonte. Os trechos são referentes as principais funções executadas pelo agentes, assim como as ações e decisões que esses tomam a cada *tick*.

```
1  to metaboliza
2
3  if nutrientes > 0
4  [
5    set energia energia + (nutrientes / (nutrientes + sujeira)) * consumo
6    set nutrientes nutrientes - consumo
7    set sujeira sujeira + consumo
8  ]
9  set energia energia - func_vitais
10
11 end
```

Figura 14: Função responsável pela metabolização dos agentes.

Pela Figura 14, pode-se observar que se o *patch* onde o agente está situado possuir nutrientes, ele absorverá do mesmo a quantidade de nutrientes estabelecidos pelo consumo, e a energia gerada será determinada pelo consumo e saturação do *patch* (linhas de 3 a 6). Após alimentar-se, o agente secreta os resíduos gerados pela metabolização dos nutrientes (linha 7).

Independente do *patch* conter ou não nutrientes, o agente gasta energia para manter-se vivo, conforme analisado na linha 9.

Se posterior a execução da função *alimenta*, a energia do agente for menor ou igual 1, então ele morre. Veja as linhas 3 e 5 da Figura 15.

```

1  to prossegue
2
3  ifelse energia <= 1
4  [
5    die
6  ]
7  [
8    if energia > energia_max
9    [
10     set energia energia_max
11    ]
12    if sensor_sinal <= count sinal
13    [
14     set consumo_reduzido true
15     set flag_sinaliza false
16    ]
17    if consumo_reduzido = true and flag_consumo = false
18    [
19     set consumo consumo_red
20     set func_vitais func_vitais_red
21     set flag_consumo true
22    ]
23  ]
24  if any? neighbors with [not any? bacteria-here and nutrientes > 0]
25  [
26    face one-of neighbors with [not any? bacteria-here and nutrientes > 0]
27    fd consumo
28  ]
29
30 end

```

Figura 15: Função responsável pela movimentação dos agentes.

Caso o agente mantenha-se vivo, sua energia é regulada e ele é questionado se o número de sinais no ambiente já não são suficientes para que ele reduza seu consumo (linhas 10 e 12).

Se o limiar de sinais já tenha sido atingido, então o agente entra em estado de consumo reduzido, reinicializando sua taxa de consumo e funções vitais com valores mais baixos (linhas 14 a 20).

Após, o agente procura um *patch* vizinho ao dele que não esteja ocupado e que tenha nutrientes disponíveis, movimentando-se para o mesmo, conforme as linhas 24 e 26.

```

1 to reproduz
2
3   if tick_reprodução = ticks
4     [
5       set flag_reprodução true
6     ]
7   if any? neighbors with [not any? bacteria-here] and energia > energia_reproduz and
8     flag_reprodução = true
9     [
10    hatch 1
11      [
12        move-to one-of neighbors with [not any? bacteria-here]
13        set flag_consumo false
14        set consumo_reduzido false
15        set flag_sinaliza true
16        set energia random-normal (item 0 energia_global) (item 1 energia_global)
17        set consumo random-normal consumo (consumo * 0.025)
18        set func_vitais random-normal func_vitais (func_vitais * 0.025)
19        set sensor_sinal random-normal sensor_sinal (sensor_sinal * 0.025)
20        set sensor_bac random-normal sensor_bac (sensor_bac * 0.025)
21        set energia_reproduz random-normal energia_reproduz (energia_reproduz * 0.025)
22        set energia_max random-normal energia_max (energia_max * 0.025)
23        set consumo_red random-normal consumo_red (consumo_red * 0.025)
24        set func_vitais_red random-normal func_vitais_red (func_vitais_red * 0.025)
25        set reduz_energia random-normal reduz_energia (reduz_energia * 0.025)
26      ]
27    set energia energia * reduz_energia
28  ]
29 ]
30 end

```

Figura 16: Função responsável pela reprodução dos agentes.

Pela Figura 16, se o agente já atingiu a etapa de adaptação então ele já está pronto para reproduzir-se (linha 6).

Caso o agente tenha todos os pré-requisitos para a reprodução, que são: ter passado a fase de adaptação, possuir energia maior ou igual que a definida pela *energia_reproduz* e ainda ter um dos *patches* vizinhos vazio, então ele vai reproduzir um agente semelhante a ele ao seu redor (linhas 7, 8 e 10).

A semelhança dos agentes é definida quando a inicialização das variáveis do novo agente são feitas baseadas nas variáveis do agente reprodutor. Isso é possível utilizando a função *random-normal* do NetLogo, e os argumentos da função são a média e o desvio padrão.

Para que o novo agente seja semelhante ao seu anterior, os valores das variáveis são sorteados tendo distribuição normal, com média igual ao valor da variável do reprodutor e um desvio padrão muito baixo (Linhas 16 a 25).

Cada vez que o agente reproduz, ele perde energia com uma taxa dada pela variável *reduz_energia* (linha 27).

```

1  to sinaliza
2
3      if flag_sinaliza = true and sensor_bac <= count bacteria
4          [
5              ask patch-here
6              [
7                  if any? neighbors with [not any? sinal-here]
8                      [
9                          sprout-sinal 1
10                         [
11                             set color gray
12                             set shape "dot"
13                             move-to one-of neighbors with [not any? sinal-here]
14                         ]
15                     ]
16             ]
17             set consumo_reduzido true
18             set flag_sinaliza false
19         ]
20
21 end

```

Figura 17: Função responsável por resolver a etapa de sinalização dos agentes.

Por fim, se o número de agentes no ambiente for maior ou igual ao da variável *sensor_bacteria*, que define o número máximo de agentes para decretar saturação do ambiente, então o agente libera uma molécula sinalizadora e entra em estado de consumo reduzido.

3.3 Estudo de Sensibilidade em Relação as Principais Variáveis que Modelam a Curva de Crescimento

Ao testar variações nos parâmetros das variáveis dos agentes, foi observado que algumas variáveis exercem mais influência na modelagem da curva de crescimento que outras. Diante do conhecimento das três principais, foi feito um estudo de impacto analisando o que a variação nos parâmetros dessas variáveis causam na curva de crescimento.

3.3.1 Funções Vitais

A variável funções vitais influencia diretamente no comportamento da curva de crescimento. As funções vitais indicam a quantidade de energia que será subtraída da variável energia a cada *tick*.

Quanto menor seu valor, maior será a taxa de crescimento dos agentes e mais rápido eles irão se reproduzir. Esse comportamento se justifica pelo fato da reprodução estar relacionada à quantidade de energia, quanto mais energia disponível na reserva do agente maior será a taxa de reprodução.

No modelo proposto, cada agente só sofrerá o processo de divisão quando estiver saudável, ou seja, quando alcançar a quantidade de energia necessária para se reproduzir, quantidade esta estabelecida pela variável energia reproduz.

A variável funções vitais está indiretamente relacionada a reprodução. Por exemplo,

se o valor das funções vitais for 0,5, isso significa que a cada tick o agente gastará 0,5 de sua energia para manter-se vivo, e o restante da reserva poderá ser utilizada para medir o quão saudável o agente está, estando apto ou não a reproduzir.

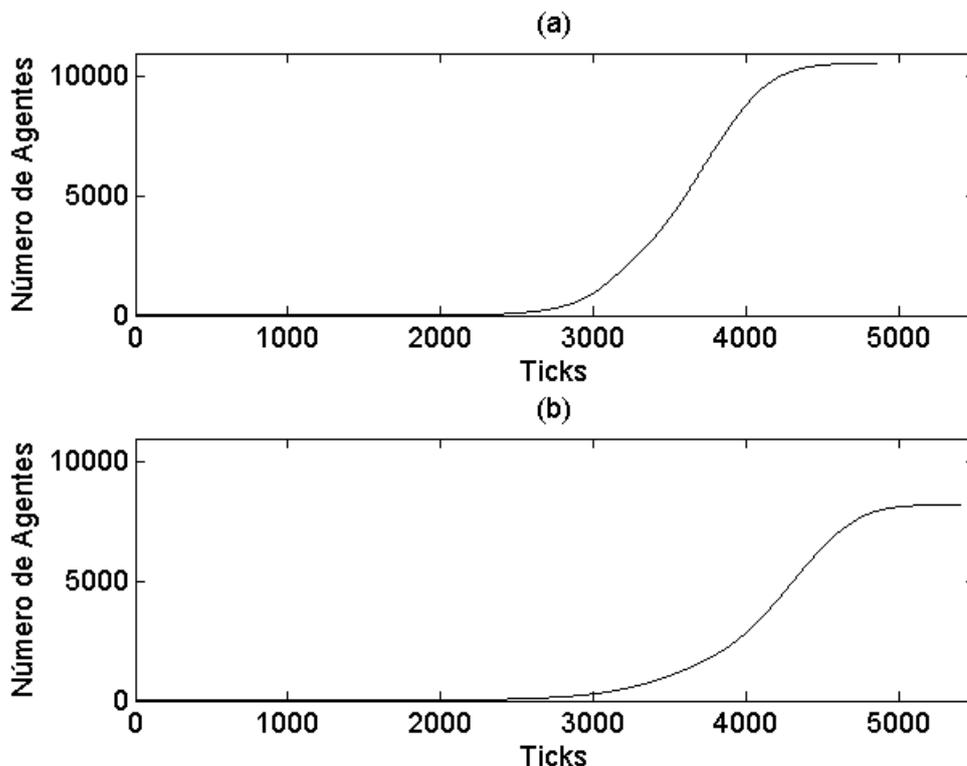


Figura 18: Comparação de crescimento para diferentes valores da variável funções vitais. Em (a) utiliza-se o valor 0,5 e em (b) 0,7. As outras variáveis mativeram-se com parâmetros fixos nas duas simulações.

Pela Figura 18, pode-se ver que (a) cresceu mais rápido que (b), e obteve um pico maior de população. E ainda, como (a) cresceu mais que (b), os nutrientes disponibilizados no ambiente foram consumidos mais depressa, e conseqüentemente o tempo de vida da população, medido em ticks, foi menor em (a) do que em (b).

Logo, quanto maior a taxa de consumo de energia para manter as funções vitais, menor será a energia depositada em sua reserva e taxa de reprodução dos agentes.

3.3.2 Energia de Reprodução

A energia de reprodução é outra variável que exerce influência na modelagem da curva de crescimento do *Mycobacterium tuberculosis*. A variável, energia de reprodução, é a quantidade de energia necessária para que o agente possa se reproduzir.

A quantidade de energia disponível na reserva do agente indica o quão saudável ele está. Se esta energia for maior ou igual à estabelecida pela variável energia de reprodução, o agente está pronto para duplicar-se.

Caso o agente não esteja saudável, ou seja, com energia menor do que a especificada pela energia de reprodução, ele continuará a exercer todas as suas funções no ambiente, exceto reproduzir.

O aumento no valor da energia de reprodução faz com que os agentes levem mais ciclos para se reproduzir, gerando menos filhos.

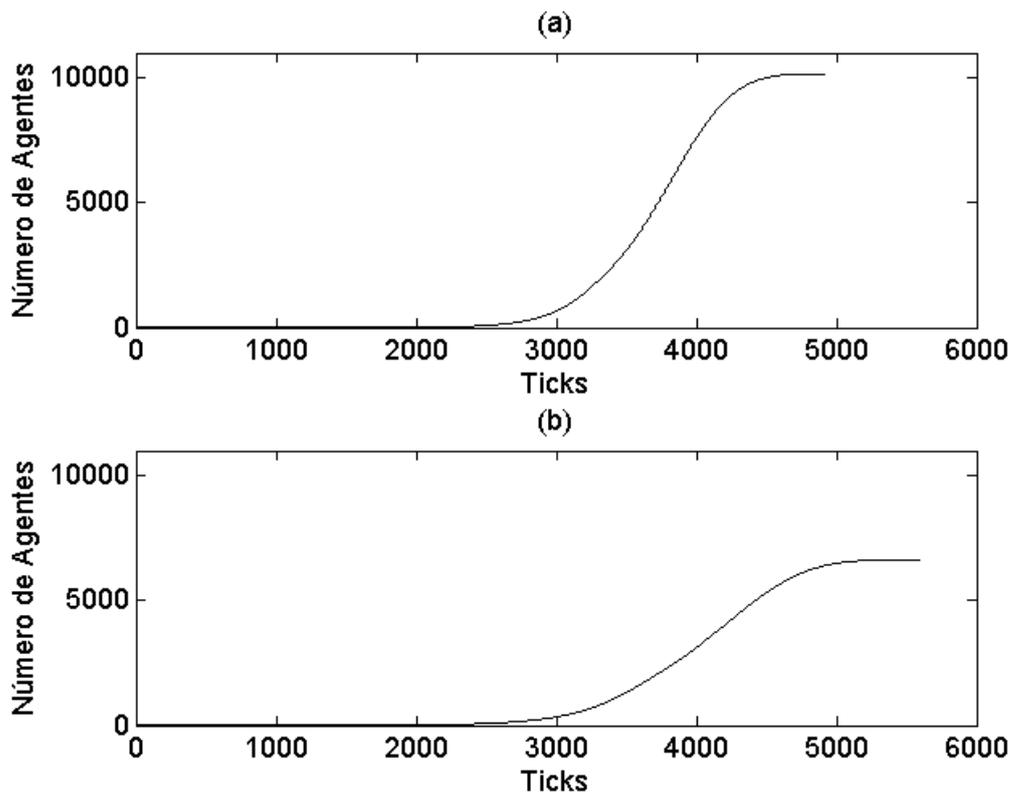


Figura 19: Comparação de crescimento para diferentes valores da variável energia reproduz. Em (a) utiliza-se o valor 200 e em (b) 250. Todas as outras variáveis tem valores iguais nas duas simulações.

Pela Figura 19, nota-se que a população de (a) é desmedidamente maior do que a de (b), e que (a) sobrevive por menos tempo no ambiente.

Assim, quanto maior for a energia necessária para reproduzir, menor será o número de agentes no ambiente e por mais tempo os agentes sobreviverão.

3.3.3 Sensor Bactéria

O sensor bactéria é uma das quatro variáveis que influenciam diretamente no modelo de crescimento bacteriano proposto. O sensor bactéria estabelece o momento em que a bactéria deve reduzir seu consumo.

Quando a bactéria decreta situação de saturação, ela libera uma molécula sinalizadora indicando a população que o ambiente está cheio, e entra em estado de consumo reduzido.

Caso a detecção do ambiente estar superpopuloso não ocorra, as bactérias consumirão os nutrientes do ambiente desenfreadamente, acumulando muita energia em sua reserva, e por isso se duplicando ligeiramente. Esse comportamento faz com que os nutrientes acabem rapidamente, resultando em uma morte prematura de todos os agentes do ambiente.

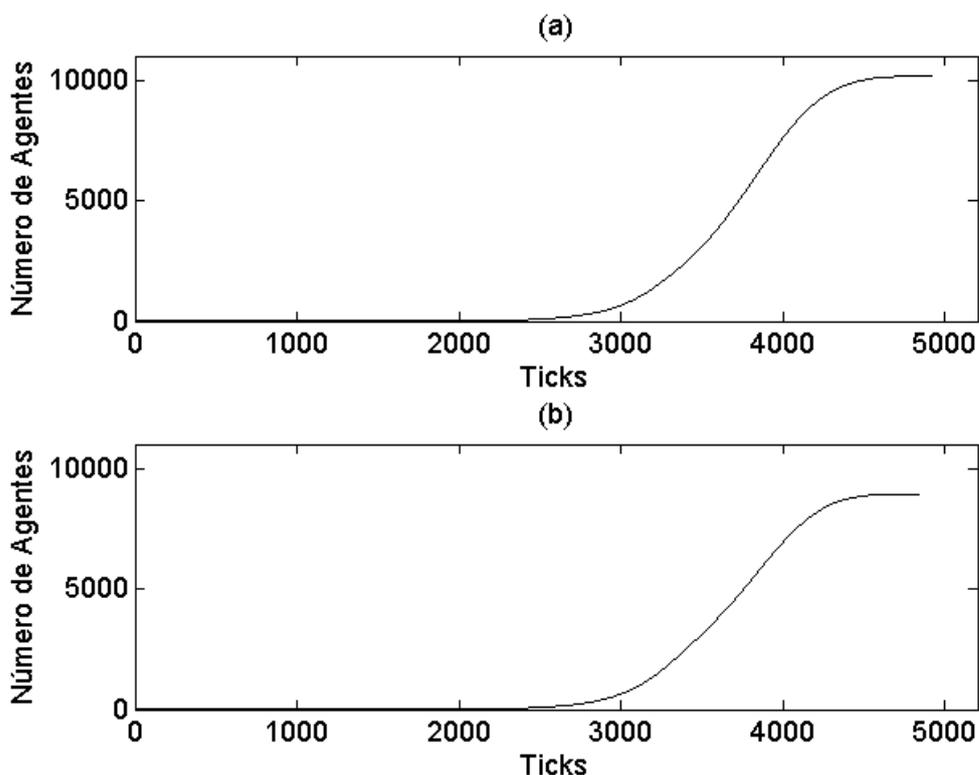


Figura 20: Comparação de crescimento para diferentes valores da variável sensor sinal. Em (a) utiliza-se o valor 2000 e em (b) 3000. Todas as outras variáveis tem valores iguais nas duas simulações.

Pela Figura 20, mesmo a população de (a) sendo maior do que a de (b), a mesma consegue permanecer por um tempo maior no ambiente.

Logo, quanto mais cedo o agente identificar superpopulação e entrar em estado de consumo reduzido, melhor será o seu aproveitamento do ambiente, podendo sua população crescer mais e permanecer viva por mais tempo.

Através dos resultados obtidos com a variação dos parâmetros de cada variável pode-se fazer uma comparação para cada uma delas, e entender como cada uma influencia na modelagem da curva de crescimento. Com a finalidade de obter similaridade com os resultados experimentais de (GROLL, 2010), os parâmetros dessas variáveis serão estimados.

3.4 Interpretação dos Valores

As curvas provenientes do MGIT expressam o resultado em unidades crescimento (*Growth Units-GU*). Por isso, torna-se necessário uma medida arbitrária de equivalência que relacione o GU com o número de agentes da simulação. No modelo desenvolvido, foi utilizado um agente para representar cada *dois* GU's.

Da mesma forma, pode-se representar as unidades de crescimento por agentes usando a mesma equivalência. A Figura 21 é a mesma Figura 2, só que com unidades de crescimento e tempo diferentes. Veja que a equivalência foi de 20.000 GU's para 10.000 agentes.

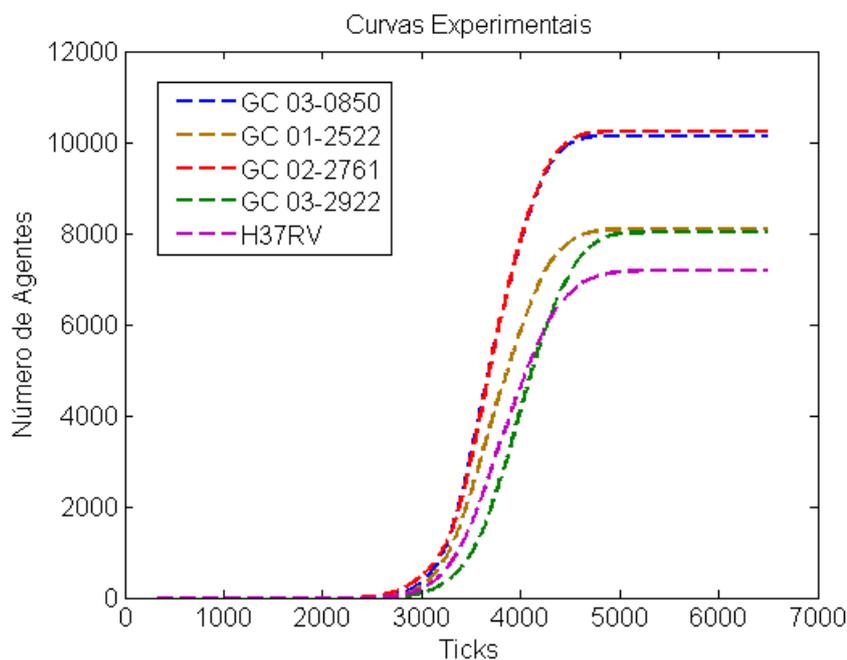


Figura 21: Curvas experimentais representadas por agentes.

Pela mesma lógica, uma equivalência arbitrária também foi obtida para o tempo de crescimento. Como a curva experimental expressa os resultados ao decorrer dos dias, e conforme visto no modelo proposto a simulação tem como divisão temporal os *ticks*, foi utilizada a equivalência de 260 *ticks* para cada dia de simulação.

4 ANÁLISE DOS RESULTADOS

Neste capítulo são apresentados os resultados obtidos após a implementação do modelo proposto. Inicialmente, apresenta-se como foi feita a parametrização do ambiente. Em um segundo momento, mostra-se as curvas geradas pelo modelo, e posteriormente é feita a validação das curvas, comparando-as com as reais e por fim tem-se um comparativo entre as curvas geradas neste trabalho e as obtidas por WERLANG (2013).

4.1 Parametrização do Ambiente

Depois de modelar a curva de crescimento do *Mycobacterium tuberculosis*, a próxima etapa foi calibrar as características e os parâmetros das variáveis do modelo buscando uma curva semelhante ao de um crescimento bacteriano. Essa busca torna-se necessária, pois, dependendo das características do ambiente e das variáveis do agente, essa semelhança nem sempre é encontrada.

Após descoberta a curva similar, neste caso, a curva com os mesmos padrões de crescimento das curvas reais, alguns parâmetros e variáveis foram fixados. Esses são comuns as cinco populações (cepas) de agentes a serem simuladas. Ver Tabelas 2 e 3.

Tabela 2: Parâmetros em comum a todas as populações de agentes.

Parâmetro	Descrição	Valor
<i>n_bac</i>	Número inicial de agentes	30 agentes
<i>n_nutri</i>	Nutrientes em cada <i>patch</i>	100 unidades

Tabela 3: Parâmetros das variáveis dos agentes comuns a todas as cepas. A média e desvio padrão de cada variável refere-se a uma determinada distribuição normal.

Variável	Descrição	Média	Desvio Padrão
<i>energia</i>	Quantidade de energia inicial do agente	100	7
<i>consumo</i>	Quantidade de nutrientes que o agente consome do ambiente a cada tick	1	0,07
<i>sensor_sinal</i>	Mesmo que bactéria não tenha liberado a molécula sinalizadora, ela pode entrar em consumo reduzido por causa detecção de um limiar de sinais no ambiente	1000	100
<i>consumo_red</i>	Define de quanto será a queda no valor de consumo dos nutrientes quando o agente entrar em estado de consumo reduzido	0,45	0,0315
<i>reduz_energia</i>	Quantidade de energia que o agente gasta para reproduzir-se	0,7	0,049

Considerando que o ambiente criado será utilizado para simular cinco populações, cada uma referente a uma cepa, torna-se necessário variações nas características, variáveis dos agentes, para possibilitar a representação de todas.

Devido ao modelo proposto ser composto de diversas variáveis, e cada uma delas podendo receber uma quantidade infinita de valores em sua inicialização, torna-se inviável trabalhar com todas. Diante dessas circunstâncias, foi feito um estudo para selecionar quais seriam as melhores para utilizá-las como função objetivo.

Para cada uma das variáveis foram testados diversos valores para seus parâmetros. Algumas demonstraram-se mais sensíveis a variação do que outras, e outras mais suscetíveis a variação por serem previsíveis.

Dentre as possíveis escolhas, foram selecionadas as que apresentavam simultaneamente maior impacto na modelagem da curva e previsibilidade com a variação de seus parâmetros. A Tabela 4 demonstra quais as variáveis e os valores de inicialização de cada uma das cepas.

Tabela 4: Parâmetros das variáveis que diferem de cepa para cepa. A média e desvio padrão são entradas de uma distribuição normal.

Variável	Descrição	GC 03-0850		GC 02-2761		GC 01-2522		GC 03-2922		H37Rv	
		Média	DP	Média	DP	Média	DP	Média	DP	Média	DP
<i>tick_reprodução</i>	Tempo que a bactéria leva para adaptar-se ao ambiente. Enquanto esse limiar não é alcançado, ela não consegue reproduzir	2400	72	2350	70,5	2516	75,48	2611	78,33	2505	75,15
<i>func_vitais</i>	Indicam a quantidade de energia que será subtraída da variável energia a cada tick	0,55	0,0385	0,55	0,0385	0,50	0,035	0,50	0,035	0,50	0,35
<i>energia_reproduz</i>	No modelo proposto, cada agente só sofrerá o processo de divisão quando estiver saudável, ou seja, quando alcançar a quantidade de energia necessária para reproduzir-se	200	14	200	14	225	15,75	230	16,1	235	16,45
<i>sensor_bac</i>	Quando a bactéria decreta situação de saturação, ela libera uma molécula sinalizadora, indicando a população que o ambiente está cheio e entra em estado de consumo reduzido	2000	200	1950	195	3000	300	2600	260	2600	260

Cabe ressaltar que o *tick_reprodução* é a única variável do modelo que pode ser estimada com base na observação dos dados reais. Com o método padronizado por GROLL (2010) é possível determinar por quanto tempo as bactérias ficam no ambiente até iniciar a fase de reprodução. Logo para obter o tempo de adaptação deve-se fazer a conversão de dias para *ticks*.

4.2 Resultados das Simulações para Cada Cepa

Nesta seção serão apresentados os resultados obtidos e alguns momentos das curvas, como o tempo de adaptação, crescimento máximo, crescimento e população. O resultados obtidos foram gerados com as entradas estabelecidas anteriormente para cada variável e cepa.

As Figuras 22 a 26 apresentam os resultados gerados para todas as populações (cepas). Os valores gerados foram obtidos pela média de 10 (dez) simulações, que é o número usual de simulações para trabalhos da área, realizadas para cada uma das cepas. A área sombreada em cinza representa o desvio padrão, que mede quanto os resultados se distanciaram da média.

De todas a simulações, a que apresentou maior variação foi cepa CG 02-2761, conforme pode ser visualizado na Figura 23. O desvio padrão mais alto é um indicador de maior flexibilidade com relação as possíveis curvas a serem geradas, lembrando que quanto mais alto o desvio padrão menor o nível de robustez (estabilidade) do modelo.

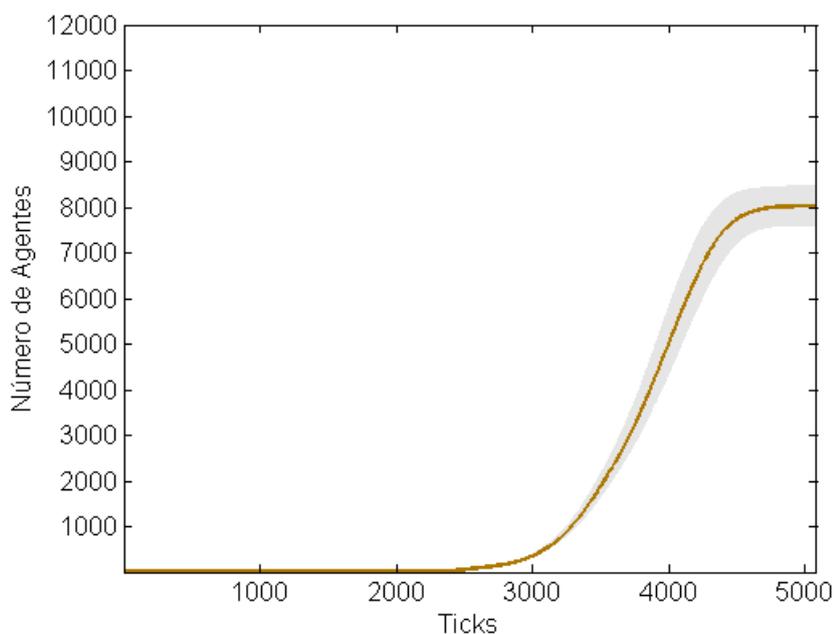


Figura 22: Simulação de GC 01-2522.

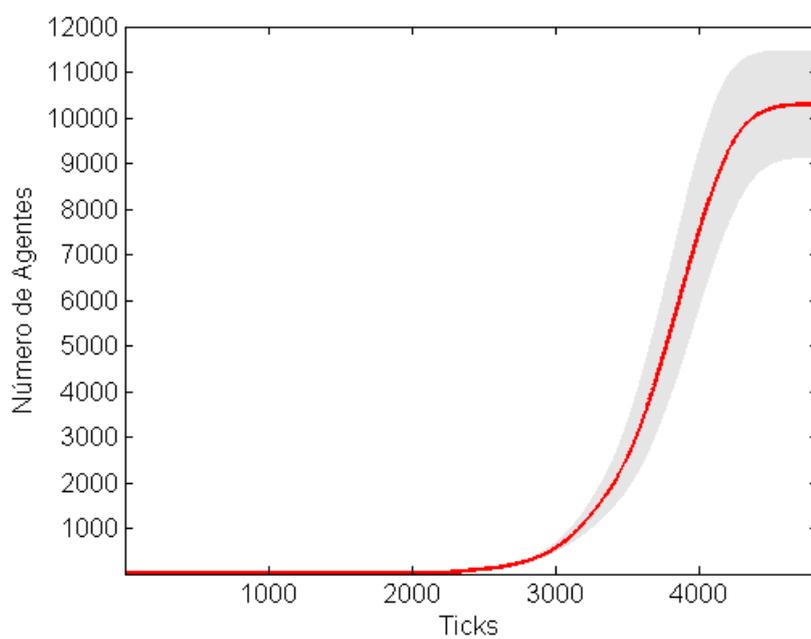


Figura 23: Simulação de GC 02-2761.

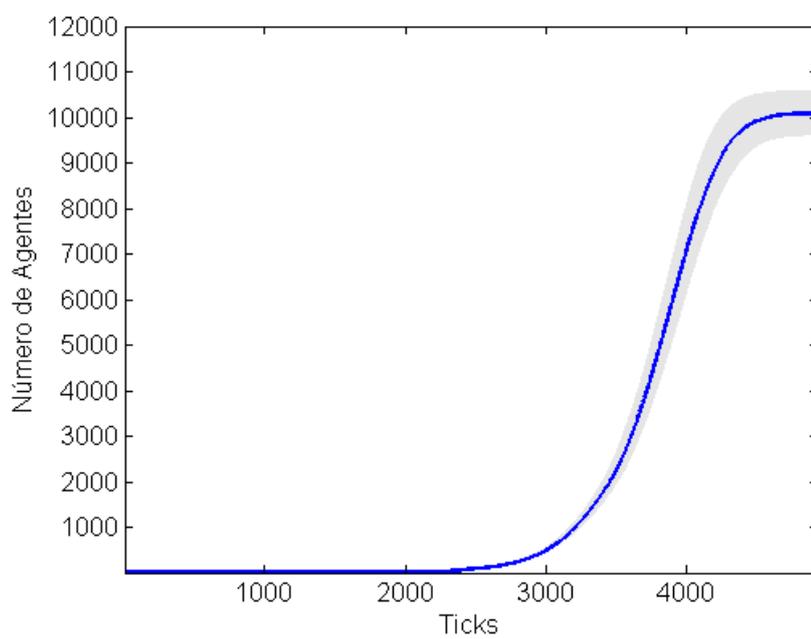


Figura 24: Simulação de GC 03-0850.

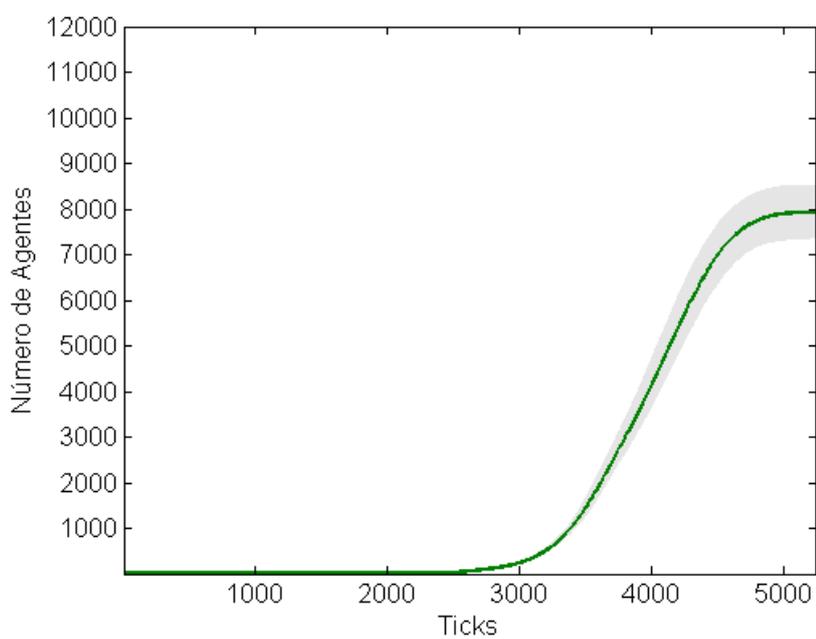


Figura 25: Simulação de GC 02-2922.

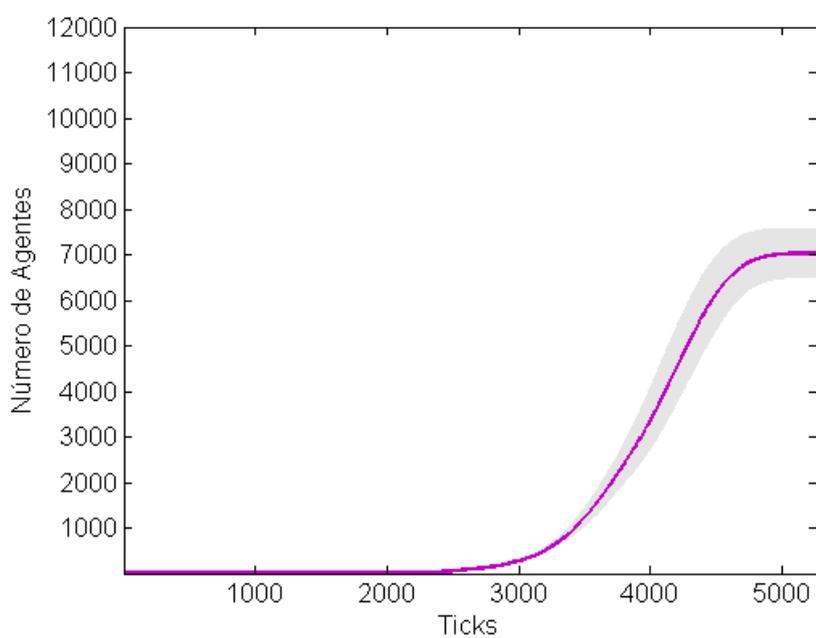


Figura 26: Simulação de H37Rv.

A Figura 27 ilustra todas as curvas simuladas, em um único gráfico, como é apresentado pelos testes *in vitro* (ver Figura 2).

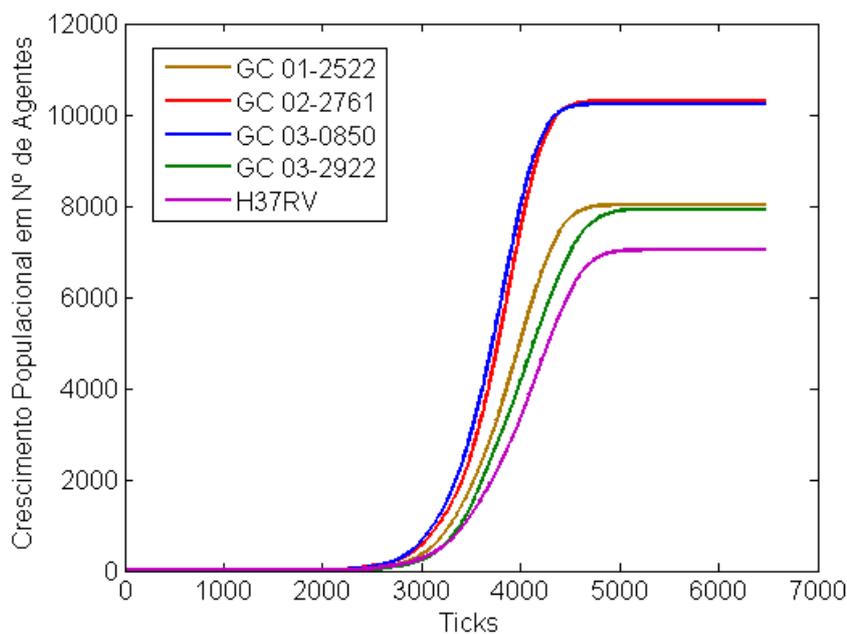


Figura 27: Curvas simuladas

Na Tabela 5 estão representados de forma numérica alguns momentos de cada curva simulada. O tempo de adaptação indica quanto tempo, medido em *ticks*, a curva levou para começar a crescer; o crescimento diz respeito ao número de *ticks* registrados até a população parar de crescer; e a população é o número máximo de agentes detectados no ambiente.

A cada simulação, os valores dos momentos eram registrados e armazenados. Os resultados apresentados são referentes à média de 10 (dez) para cada um deles.

Tabela 5: Tempo de adaptação, crescimento e população de cada uma das cepas.

Momento	GC 01-2522	GC 02-2761	GC 03-0850	GC 03-2922	H37RV
Adaptação	2387	2229	2300	2507	2381
Crescimento	4904	4632	4766	5058	5151
População	8022	10286	10081	7981	7024

4.3 Validação do Modelo

O modelo proposto foi submetido a testes com a finalidade de validá-lo. Os resultados gráficos e numéricos apresentados na seção anterior foram comparados com os dados reais obtidos por GROLL (2010).

Para fazer a comparação foi preciso fazer a conversão dos dados. Os *ticks* foram convertidos para dias e o números de agentes para unidade de crescimento (GU). Lembrando que cada dia equivale a *260 ticks* e cada agente representa *dois* GU's.

As Figuras de 28 a 32 são cada uma composta por duas curvas de crescimento: a linha pontilhada representa a curva real, e a contínua representa a curva simulada. A linha pontilhada trata-se da média de dois experimentos com mesma solubilidade realizados por (GROLL, 2010), a linha contínua é constituída pela média de dez simulações, e a área em cinza em seu contorno representa seus respectivos desvios padrões.

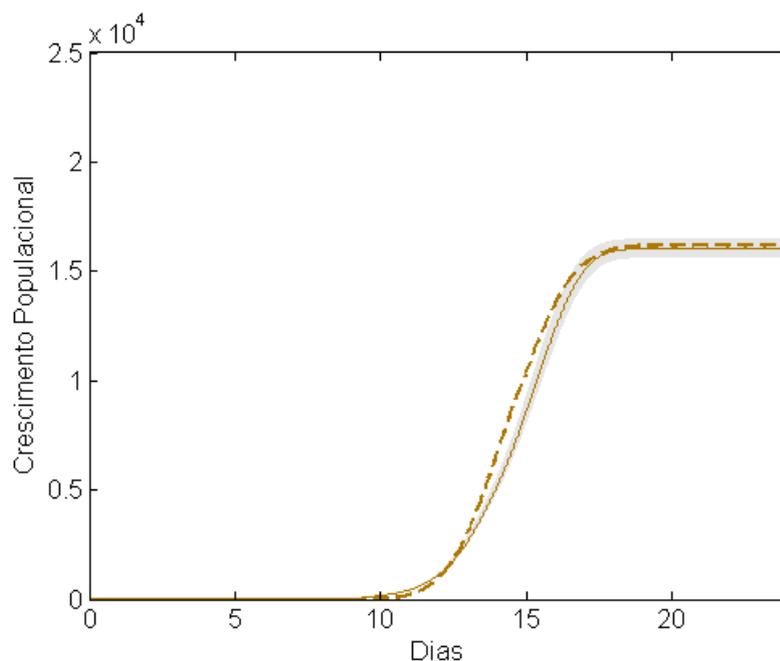


Figura 28: Comparação das curvas real e simulada de GC 01-2522.

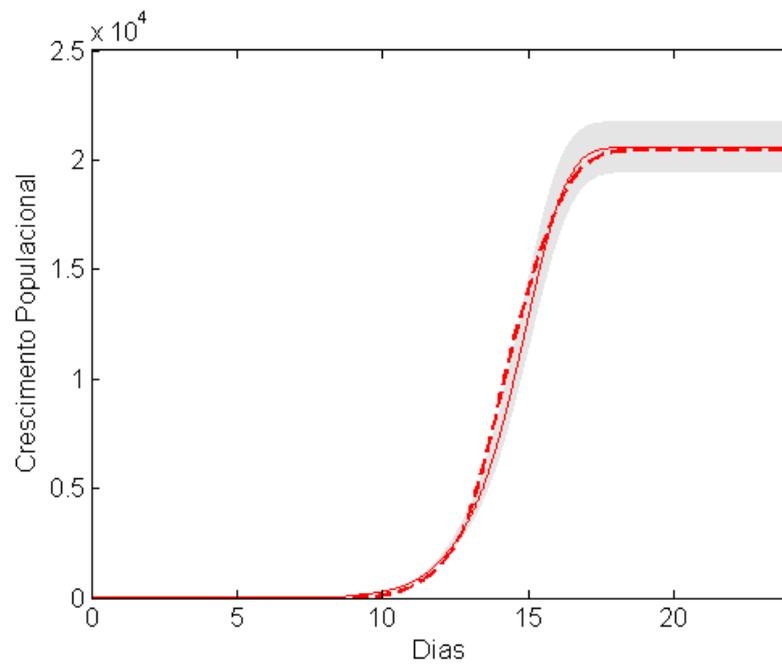


Figura 29: Comparação das curvas real e simulada de GC 02-2761.

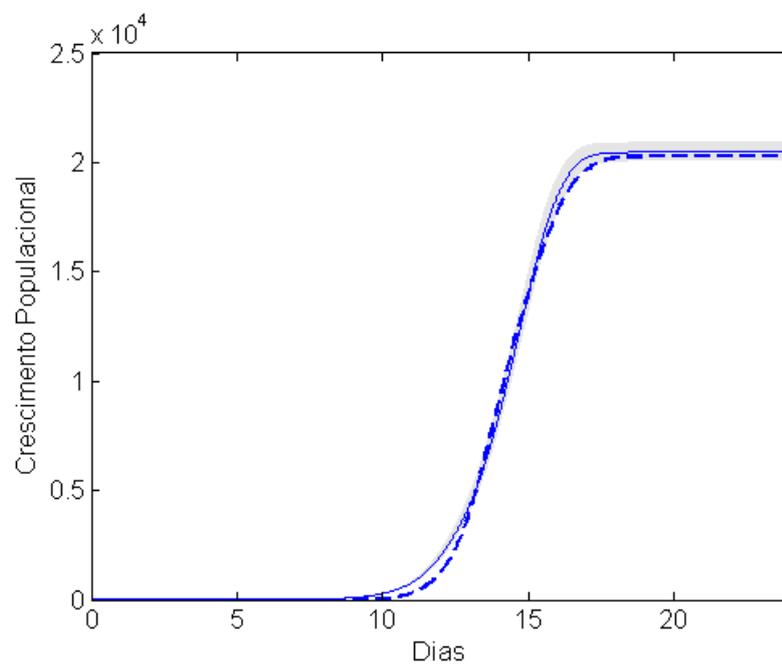


Figura 30: Comparação das curvas real e simulada de GC 03-0850.

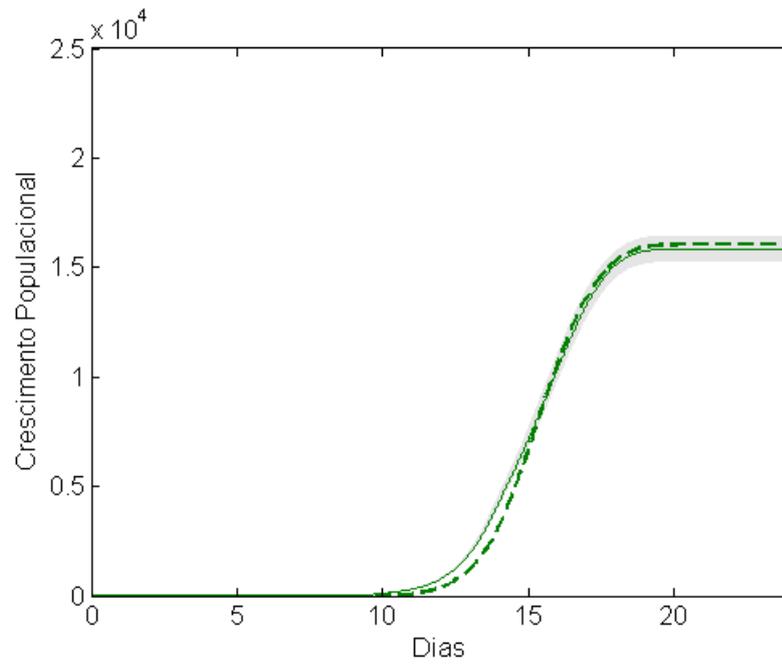


Figura 31: Comparação das curvas real e simulada de GC 03-2922.

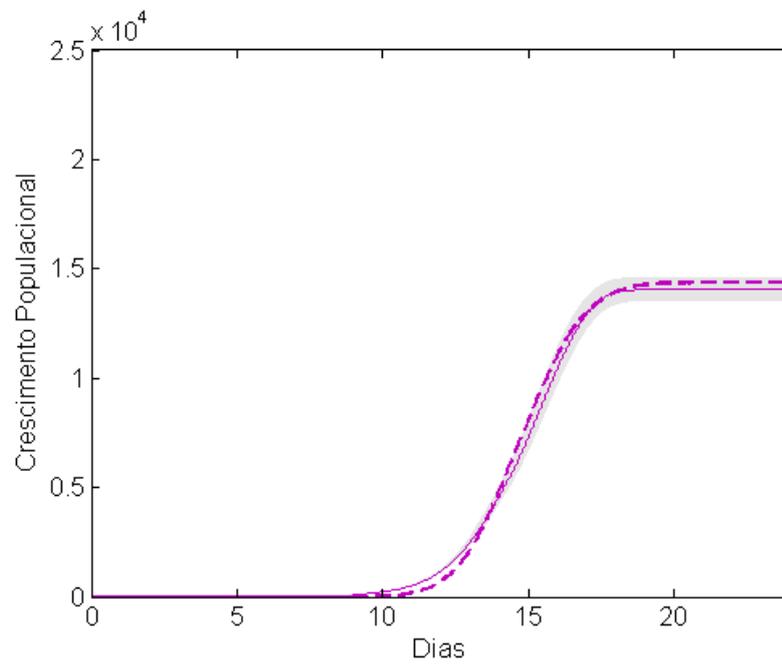


Figura 32: Comparação das curvas real e simulada de H37Rv.

As Tabelas de 6 a 10 representam um comparativo com momentos das curvas simuladas, que já foram mostrados e explicitados anteriormente de forma isolada, com os registrados nas curvas de crescimento reais.

Na segunda coluna estão apresentados os momentos dos dados reais, obtidos pela média de dois experimentos *in vitro*. Na quarta coluna estão ilustrados os momentos das curvas simuladas, gerados pela média de dez simulações. A quarta coluna apresenta o desvio padrão dos dados simulados.

Por fim a última coluna calcula o erro relativo percentual, que permite fazer uma comparação do valor simulado com um valor exato. Ele dá a diferença entre os valores como uma porcentagem do valor exato, ajudando a descobrir o quão próximo esta o simulado do real.

Para fazer as análises numéricas, foi necessário realizar conversão dos dados. No entanto, agora de maneira inversa, pois optou-se por representar os dados por *ticks* e agentes. Observa-se que foram convertidos dias para *ticks* e unidades de crescimento para número de agentes.

Tabela 6: Tempo de adaptação, crescimento e população de GC 01-2522.

Momento	Real	Simulado	Desvio padrão	Erro
Adaptação	2535	2387	40,17	5,84 %
Crescimento	5070	4904	130,11	3,27 %
População	8090	8022	445,44	0,84 %

Tabela 7: Tempo de adaptação, crescimento e população de GC 02-2761.

Momento	Real	Simulado	Desvio padrão	Erro
Adaptação	2384	2229	44,87	6,50 %
Crescimento	4820	4640	128,61	4,06 %
População	10217	10286	1183	0,67 %

Tabela 8: Tempo de adaptação, crescimento e população de GC 03-0850.

Momento	Real	Simulado	Desvio padrão	Erro
Adaptação	2425	2300	28,45	5,15 %
Crescimento	4919	4766	142,68	3,11 %
População	10139	10081	497,36	0,57 %

Tabela 9: Tempo de adaptação, crescimento e população de GC 03-2922.

Momento	Real	Simulado	Desvio padrão	Erro
Adaptação	2620	2507	58,82	4,31 %
Crescimento	5220	5078	129,14	2,72 %
População	8014	7956	341,34	0,72 %

Tabela 10: Tempo de adaptação, crescimento e população de GC H37Rv.

Momento	Real	Simulado	Desvio padrão	Erro
Adaptação	2524	2381	40,94	5,66 %
Crescimento	5299	5151	87,16	2,79 %
População	7179	7024	554,40	2,15 %

As diversas tentativas de ajuste dos parâmetros permitiram que o comportamento das curvas de crescimento simuladas pudessem ser reproduzidas com um nível de semelhança relativamente próximo as curvas de crescimento reais.

Analisando os dados obtidos, percebe-se que algumas representam melhor o início do crescimento; outras o meio do crescimento; e a maioria consegue representar de maneira fiel o final do crescimento. Dessa forma, não pode-se concluir qual obteve melhor representação. No entanto, a simulação da cepa GC 01-2522 é a que mais se afasta do modelo real.

Fazendo uma análise das tabelas, a cepa simulada com menor média de erros foi a CG 03-2922, com média de 2,58%, e a com maior foi a CG 02-2761, com média de 3,74%. E a cepa CG 01-2522, que aparentemente apresenta maior diferença entre as curvas real e simulada, não obteve a maior média de erro. Como essas médias são referentes a momentos da curva, acredita-se que deva-se dar mais confiabilidade aos dados plotados, já que representam um todo e não apenas partes da curva.

4.4 Comparativo com os Resultados Obtidos por WERLANG (2013)

WERLANG (2013) também criou um modelo baseado em sistemas multiagentes para representar as curvas de crescimento do *Mycobacterium tuberculosis*. No entanto, os agentes desse modelo se comportam de maneira idêntica, desprezando a individualidade de cada um. Esta seção apresenta um comparativo entre as curvas de crescimento obtidas por WERLANG (2013) e as curvas simuladas apresentadas neste trabalho.

As Figuras de 33 a 37 apresentam as curvas de crescimento de cada uma das cepas. Foram apresentadas em um mesmo gráfico as curvas de crescimento reais obtidas por GROLL (2010), as simuladas apresentadas por WERLANG (2013) e as modeladas neste trabalho.

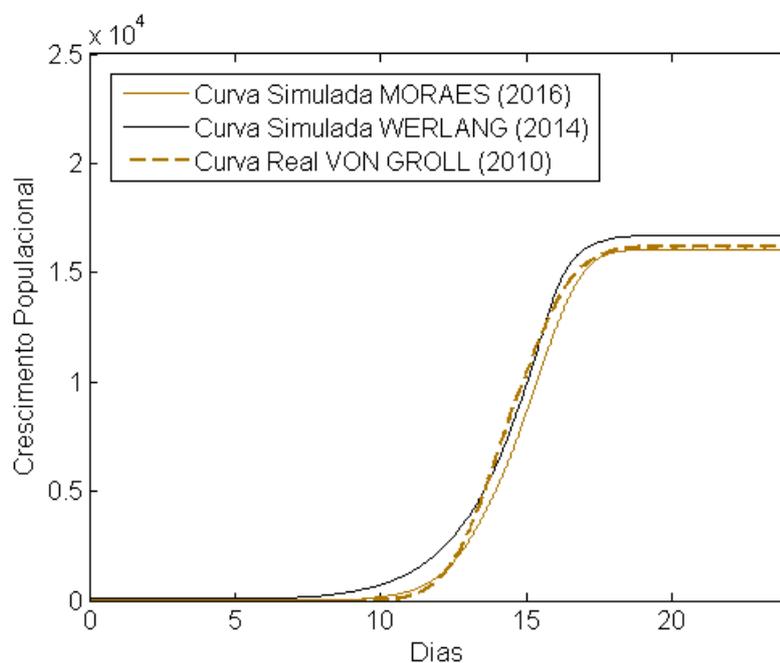


Figura 33: Comparação dos crescimentos GC 01-2522.

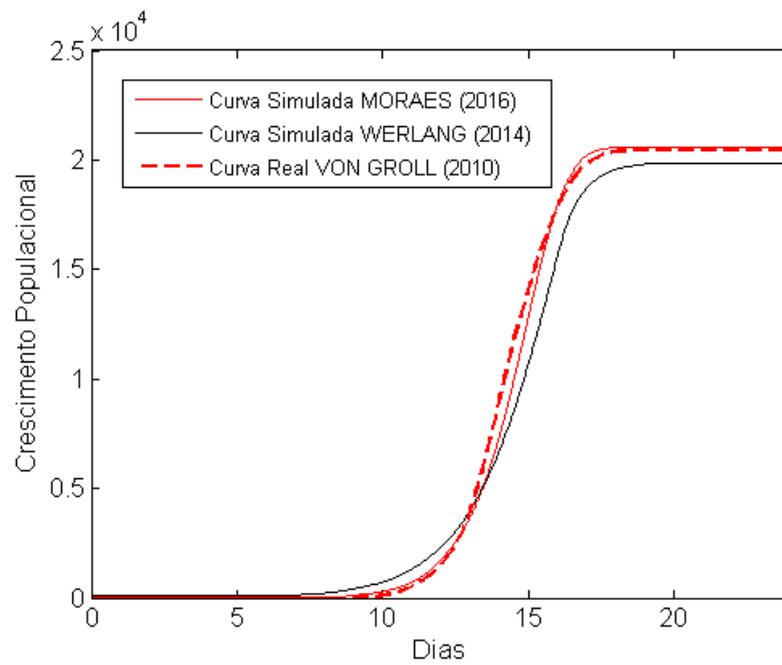


Figura 34: Comparação dos crescimentos GC 02-2761.

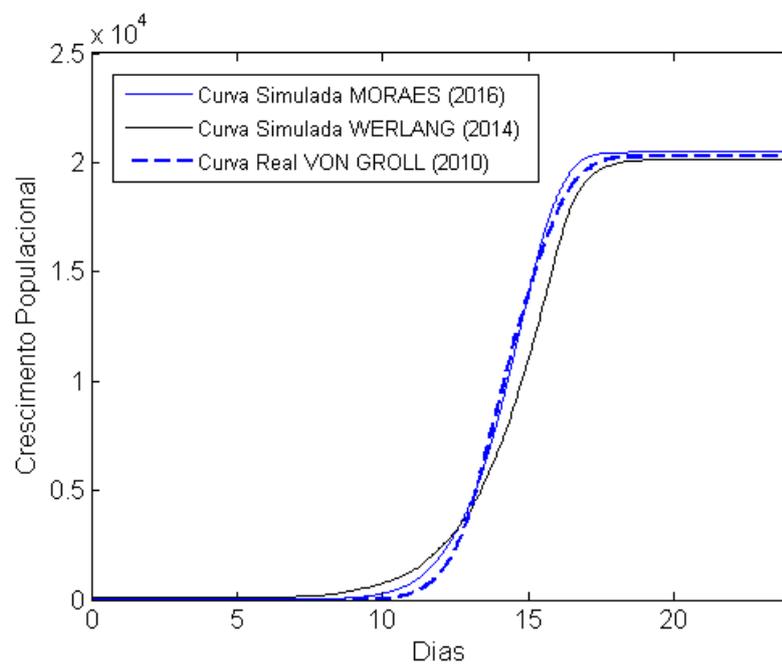


Figura 35: Comparação dos crescimentos GC 03-0850.

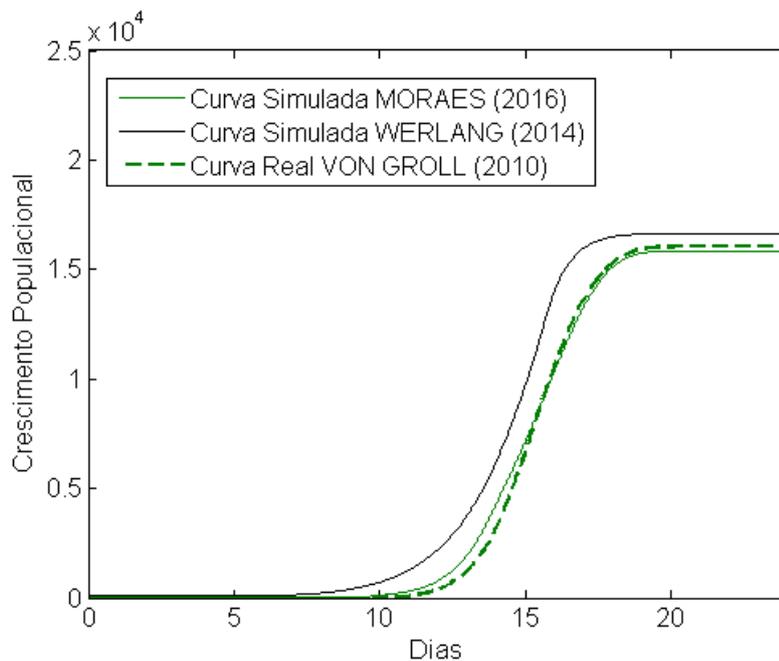


Figura 36: Comparação dos crescimentos GC 03-2922.

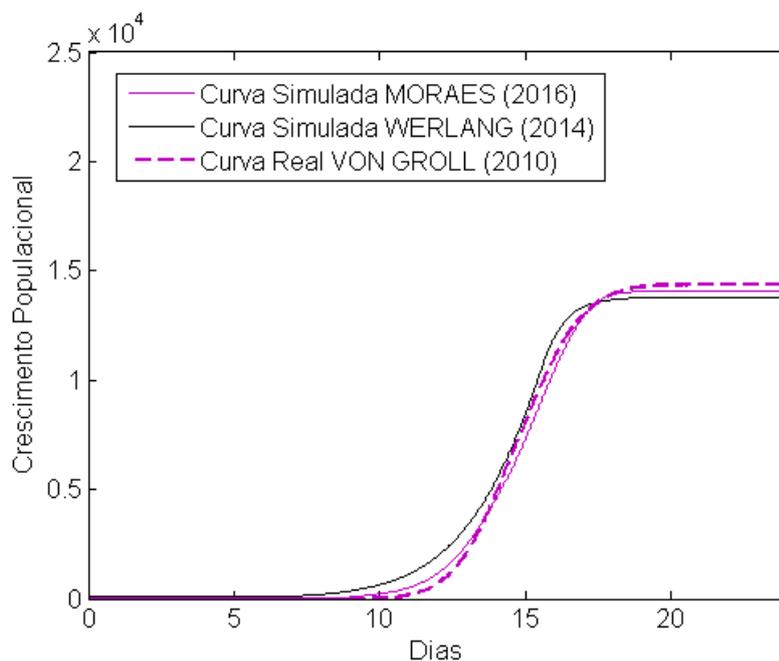


Figura 37: Comparação dos crescimentos GC H37Rv.

A cepa GC 01-2522 foi a que mais se distanciou da curva de crescimento real no modelo desenvolvido neste trabalho. As outras cepas tiveram um comportamento mais aproximado ao das curvas reais, quando comparadas as representações das curvas reais obtidas por WERLANG (2013).

Acredita-se que os resultados deste modelo tenham sido mais aproximados aos das

curvas de crescimento reais, pois utiliza distribuição de probabilidade para os possíveis valores das variáveis. Dessa forma, inserindo um comportamento difuso para cada agente, expressando sua individualidade.

No entanto, não se pode afirmar com certeza os motivos dessa melhor representação, já que a calibragem dos parâmetros foram feitas de maneiras distintas e por diferentes analistas.

5 CONCLUSÕES E TRABALHOS FUTUROS

A tuberculose é uma das enfermidades mais antigas, com ampla distribuição geográfica, constituindo um sério problema de saúde pública no mundo e no Brasil. Os estudos de crescimento populacional do *Mycobacterium tuberculosis*, causador da tuberculose, possibilitam a compreensão de diversos comportamentos do bacilo, como a resistência a fármacos.

Este trabalho teve como principal objetivo o desenvolvimento de um modelo de crescimento populacional do *Mycobacterium tuberculosis*, capaz de reproduzir as curvas reais dadas como entrada. A proposta teve como inspiração os trabalhos de WERLANG (2013), que criou um modelo baseado em multiagentes capaz de representar as curvas reais do *Mycobacterium tuberculosis* com um certo grau de similaridade e GROLL (2010), que padronizou um método para medir o crescimento populacional do bacilo quando colocados em um meio. Os resultados obtidos por GROLL (2010) viabilizaram a realização de testes para verificar a capacidade do modelo em reproduzir as curvas reais.

A curva de crescimento proposta foi modelada com a utilização de simulação baseada em agentes. WOOLDRIDGE (2009) afirma que os sistemas multiagentes são uma poderosa e flexível ferramenta para modelagem de sistemas ambientais/sociais, pois desta maneira pode-se analisar o comportamento de cada indivíduo, ao invés de uma média dos comportamentos dos mesmos.

As regras comportamentais dos agentes foram estipuladas de acordo com o conhecimento adquirido por especialistas e referências bibliográficas da área.

Para tornar o modelo criado mais similar a um modelo de crescimento real, foi proposto que os possíveis valores das variáveis fossem probabilisticamente distribuídos. Dessa forma, abordando a individualidade de cada agente. Tinha-se como hipóteses conseguir estimar a melhor função de densidade de probabilidade possível para as variáveis, tendo como produto final uma curva idêntica, ou com erro mínimo, quando comparado a real.

Após a realização de estudos e pesquisas foi constatado que a distribuição normal é um dos modelos mais utilizados para representar variáveis aleatórias, se aplicando a diversos fenômenos. Logo, as variáveis do modelo também utilizaram distribuição normal para

selecionar seus valores.

Depois de realizadas as primeiras simulações do modelo, pode-se observar que o ajuste dos parâmetros de algumas variáveis exerciam maior impacto no formato da curva. Cada uma destas variáveis foram ajustadas com a finalidade de representar cada uma das cinco cepas a serem simuladas: GC 01-2522, GC 02-2761, CG 03-0850, GC 03-2922 e H37Rv.

Os resultados apresentados foram considerados satisfatórios, uma vez que a curva de crescimento modelada revelou-se similar a curva real. Assim, o modelo mostrou-se eficiente em reproduzir curvas de crescimento do *Mycobacterium tuberculosis* dadas como entrada.

Até o presente momento, esse trabalho gerou as seguintes publicações (aceitas e/ou submetidas):

- Definição Automática de Variáveis em Simulação Baseada em Agentes Utilizando Distribuições de Probabilidades. 14^a Mostra de Produção Universitária da Universidade Federal do Rio Grande, 2015.
- Using Probability Distributions in Parameters of Variables at Agent-Based Simulations: a Case Study for the TB Bacillus Growth Curve. Proposta de Capítulo submetida e aceita para o livro "Multi-Agent Based Simulations Applied to Biological and Environmental Systems". O capítulo completo será submetido até dia 31 de março.
- Modelo de Crescimento Populacional do *Mycobacterium tuberculosis* utilizando Simulação Baseada em Agentes e Distribuições de Probabilidades. Congresso Nacional de Matemática Aplicada e Computacional, a ser apresentado entre 05 e 09 de Setembro de 2016.
- Modelando a curva de crescimento do *Mycobacterium tuberculosis* com a utilização de simulação multiagente: um estudo de caso para a variável *Quorum sensing*. Trabalho submetido ao WESAAC 2016 - 10 Workshop-Escola de Agentes, seus Ambientes e suas Aplicações.

Toda tarefa de busca e otimização possui vários componentes, entre eles destaca-se o espaço de busca, onde são consideradas todas as possibilidades de solução de um determinado problema.

Como a calibragem dos parâmetros das variáveis deste modelo foram feitas manualmente, o espaço de busca teve de ser reduzido, pois se fossem consideradas todas as possibilidades, haveria uma quantidade infinita de possíveis valores a serem levados em consideração, para o número de variáveis apresentadas.

Existem muitos métodos de busca, e propõem-se como trabalhos futuros a utilização de alguns métodos, como algoritmos genéticos, visando soluções ótimas ou aproximadamente ótimas para os valores das variáveis.

REFERÊNCIAS

- ADAMATTI, D. Simulação Baseada em Multiagentes como Ferramenta em Estudos Interdisciplinares. **RENOTE**, Porto Alegre, v.9, n.1, 2011.
- ANTUNES, L. C. M. A Linguagem das bactérias. **Ciência hoje**, Rio de Janeiro, v.33, n.193, p.16–20, 2003.
- AZEVEDO, L. L. **AProSiMA - ambiente de resolução cooperativa de problemas baseado em simulação multiagentes**. 2007. Tese Doutorado — Programa de Doutorado em Engenharia Elétrica, UFES, Vitória.
- BURGOS, M.; PYM, A. Molecular epidemiology of tuberculosis. **European Respiratory Journal**, France, v.20, n.36, p.54s–65s, 2002.
- BUSSAB, W.; MORETTIN, P. **Estatística Básica**. 5.ed. São Paulo: Saraiva, 2005.
- CALLEGARI-JACQUES, S. M. **Bioestatística princípios e aplicações**. Porto Alegre: Artmed, 2003.
- CARGNELUTTI FILHO, A.; MATZENAUER, R.; TRINDADE, J. d. Ajustes de funções de distribuição de probabilidade à radiação solar global no Estado do Rio Grande do Sul. **Revista Agropecuária Brasileira**, Brasília, v.39, n.12, p.1157–1166, 2004.
- COLELLA, V. S.; KLOPFER, E.; RESNICK, M. **Adventures in Modeling: Exploring Complex, Dynamic Systems with Starlogo**. NY: Teachers College Press, 2001.
- DEVORE, J. L.; SILVA, J. P. N. da. **Probabilidade e estatística: para engenharia e ciências**. São Paulo: Pioneira Thomson Learning, 2006.
- GARCIA, A.; SICHMAN, J. Agentes e Sistemas Multiagentes. In: **Sistemas Inteligentes Fundamentos e Aplicações**. 1.ed. São Paulo: Manole, 2005. p.269–304.
- GROLL, A. V. **Fitness of Mycobacterium tuberculosis associated to genotypes and drug resistance: new approaches for understanding the transmission dynamics of tuberculosis**. 2010. Tese Doutorado — Ghent University, Ghent. 137p.

- LARSON, R.; FARBER, E. **Estatística Aplicada**. 2.ed. São Paulo: Pearson Prentice Hall, 2005.
- LAWN, S. D.; ZUMLA, A. I. Tuberculosis. **The Lancet Journal**, [S.l.], v.378, n.9785, p.57–72, 2011.
- LÖNNROTH, K.; MIGLIORI, G. B.; ABUBAKAR, I.; D'AMBROSIO, L.; DE VRIES, G.; DIEL, R.; DOUGLAS, P.; FALZON, D.; GAUDREAU, M.-A.; GOLETTI, D. et al. Towards tuberculosis elimination: an action framework for low-incidence countries. **European Respiratory Journal**, France, v.45, n.4, p.928–952, 2015.
- MESQUITA, M. d. G.; MORAES, S. O.; CORRENTE, J. E. et al. More adequate probability distributions to represent the saturated soil hydraulic conductivity. **Scientia Agricola**, [S.l.], v.59, n.4, p.789–793, 2002.
- PAIXÃO, L. M. M.; GONTIJO, E. D. Perfil de casos de tuberculose notificados e fatores associados ao abandono, Belo Horizonte, MG. **Revista de Saúde Pública**, [S.l.], v.41, n.2, p.205–213, 2007.
- PORCELLIS, D.; MORAES, M.; BERTIN, C. Modelo Baseado em Equações versus Modelo Baseado em Agentes: uma abordagem usando sistema predador-presa. In: IX WORKSHOP-ESCOLA DE SISTEMAS DE AGENTES, SEUS AMBIENTES E APLICACÕES, 2015, Niterói. **Anais...** UFF, 2015. p.100–111.
- REBONATTO, M. T. **Simulação ao paralela de eventos discretos com uso de memória compartilhada distribuída**. 2000. Dissertação (Mestra em Ciência da Computação) — Programa de Pós-Graduação em Computação, UFRGS, Porto Alegre.
- RUMJANEK, N. G.; FONSECA, M.; XAVIER, G. Quorum sensing. **Revista Biotecnologia Ciência & Desenvolvimento-Edição nº**, [S.l.], v.33, p.35, 2004.
- SAÚDE, M. da. **Tuberculose**. Brasília: Portal da Saúde, 2016.
- TERANO, T. Agent-Based Modeling: Introduction and Perspective. In: INTERNATIONAL CONFERENCE ON INNOVATION AND MANAGEMENT, 8., 2011, Wuhan, China. **Proceedings...** Wuhan University of Technology Press, 2011. p.1003–1009.
- TODAR, K. **The growth of bacterial populations**. [S.l.]: Todar's Online Textbook of Bacteriology, 2013.
- WERLANG, P. **Simulação da curva de crescimento do Mycobacterium tuberculosis utilizando sistemas multiagentes**. 2013. Dissertação (Mestrado em Modelagem Computacional) — Programa de Pós-Graduação em Modelagem Computacional, FURG, Rio Grande.

WHITEHEAD, N. A.; BARNARD, A. M.; SLATER, H.; SIMPSON, N. J.; SALMOND, G. P. Quorum-sensing in Gram-negative bacteria. **FEMS microbiology reviews**, [S.l.], v.25, n.4, p.365–404, 2001.

WHO. **Global tuberculosis report 2015**. [S.l.]: World Health Organization, 2015.

WILENSKY, U. **NetLogo**. [S.l.]: Northwestern University, 1999.

WOOLDRIDGE, M. **An Introduction to Multiagent Systems**. England: John Wiley and Sons, 2009.