

Um Estudo de Autômatos Celulares com Sistemas Difusos para Modelos Compartimentais do Tipo SIR

Walley da Costa^{1,2}, Lílíam Medeiros¹, Sandra Sandri¹

¹Instituto Nacional de Pesquisas Espaciais – INPE

²Fundação Centro de Pesquisa e Desenvolvimento em Telecomunicações (CPqD)

{walleycosta@gmail.com, lcastro@dpi.inpe.br, sandra.sandri@inpe.br}

Abstract. *This paper proposes a representation of the dynamics of epidemics through a compartmental SIR model, simulated by the combined use of cellular automata and fuzzy systems. Furthermore, we present an application of the model on the spread of dengue, with simulations done using data from the Governor's Island (RJ).*

Resumo. *Este trabalho propõe uma representação da dinâmica de epidemias, através de um modelo compartimental SIR, simulado por meio do uso conjunto de autômatos celulares e sistemas difusos. Apresenta-se também uma aplicação do modelo proposto sobre propagação de dengue, com simulações efetuadas utilizando-se dados da Ilha do Governador (RJ).*

Palavras-chave: *Epidemias, Modelo Compartimental SIR, Autômatos Celulares, Sistemas Difusos, Propagação de Dengue, Ilha do Governador (RJ).*

1. Introdução

A epidemiologia matemática [YANG 2001] é uma nova área da ciência que propõe modelos que ajudam a traçar políticas de controle de doenças infecciosas. Modelos compartimentais, em particular, têm sido bastante utilizados para estudar a epidemiologia destas doenças [HETHCOTE 2000]. Um dos modelos compartimentais mais utilizados na modelagem da dinâmica de transmissão de doenças infecciosas é o modelo *SIR* (Suscetíveis-Infetados-Recuperados) [KERMACK and MCKENDRICK 1927].

Os *Autômatos Celulares* são capazes de descrever a evolução espaço-temporal de sistemas complexos, como a modelagem de propagação de epidemias [OLIVEIRA et al. 2008], mediante a aplicação de um conjunto de regras de transição¹, que alteram os atributos de cada célula do autômato em função dos atributos das células vizinhas. Porém, com frequência, as regras para transições locais em modelos de sistemas reais não são conhecidas com exatidão [MRAZ et al. 1999]. A utilização da *Teoria dos Conjuntos Difusos* [ZADEH 1965], permite a criação da função de transição de autômatos celulares, ou seja, do comportamento de uma célula, através da modelagem do conhecimento de um especialista sobre o comportamento real do sistema. A capacidade dos conjuntos difusos em expressar graus intermediários de pertinência possibilita, além de uma representação significativa e poderosa de incertezas, uma representação valiosa de conceitos vagos expressos em linguagem natural [KLIR and YUAN 1995].

¹Neste trabalho, também nos referimos às regras de transição de autômatos celulares como função de transição, para evitar confusões com as regras utilizadas nos sistemas difusos.

O objetivo deste trabalho é propor uma representação da dinâmica de epidemias, através de um modelo compartimental SIR, simulado por meio do uso conjunto de autômatos celulares e sistemas difusos. Os efeitos locais e não locais de doenças infecciosas são considerados, tendo a taxa de transmissão como saída de um sistema difuso. Neste trabalho, apresenta-se também uma aplicação do modelo proposto sobre propagação de dengue, com simulações efetuadas utilizando-se dados da Ilha do Governador (RJ).

2. Proposta de Modelos SIR com Autômatos Celulares Difusos

Caso as informações para as regras de transição de um autômato celular sejam incertas, dificultando a construção de expressões analíticas, este trabalho propõe a utilização de um Sistema de Inferência Difuso [KLIR and YUAN 1995, NGUYEN and WALKER 2006] para obtenção de valores numéricos precisos que podem ser usados como parâmetros em funções de transição. Também é formulado um modelo SIR geral que é simulado através de um autômato celular difuso.

2.1. Descrição do Modelo Geral

A dinâmica da doença em uma comunidade é modelada por um autômato celular H , bidimensional, composto por uma rede de células quadradas, e que representa os espaços dos humanos. Cada célula $H(i, j)$ é um vetor que simboliza um conjunto de domicílios onde habitam pessoas. Entre os dados armazenados em uma célula $H(i, j)$ do autômato H estão o total de humanos $NH_{i,j}$, o total de humanos suscetíveis $NS_{i,j}$, o total de humanos infectados $NI_{i,j}$ e o total de humanos recuperados $NR_{i,j}$. O total de humanos infectados $NI_{i,j}$ pode ser dividido em k estágios de infecção, de acordo com a doença ou a necessidade. As pessoas em cada estágio estão contabilizadas em $NI1_{i,j}, NI2_{i,j}, \dots, NIk_{i,j}$. Durante as simulações do modelo, é considerada uma população constante de agentes transmissores e humanos. Os agentes transmissores da doença estão implícitos no modelo através das pessoas infectadas e pode-se dizer que são em número igual em todas as células. No caso dos humanos, a quantidade de pessoas pode variar em cada célula, mas no modelo não há renovação, ou seja, não são levados em consideração nascimentos, mortes, imigrações e emigrações. Logo, se as pessoas se tornam imunes à doença considerada na execução do modelo, neste ambiente os indivíduos suscetíveis se esgotam e a epidemia termina acompanhada de um alto percentual de pessoas imunizadas. Variações climáticas não são consideradas no modelo, assim como a população humana não é estruturada por idade.

Entre os parâmetros de entrada do modelo estão o percentual ρ_{ASS} de indivíduos com infecções assintomáticas (sem sintomas), que está relacionado à fatores ambientais, individuais, do agente transmissor e do próprio vírus, e a taxa de mobilidade humana ρ_{MOB} a células vizinhas (vizinhança de Moore com condição de contorno nula). Ao final apenas os indivíduos infectados assintomáticos, juntamente com indivíduos recuperados e suscetíveis, que podem sair de suas células locais para células vizinhas, uma vez que os demais ficariam impossibilitados devido às suas condições físicas. Outros parâmetros de entrada são a quantidade k de estágios da infecção da doença considerada, e o intervalo de tempo Δt que define o período em que as pessoas permanecem em cada um dos estágios da infecção. Dessa forma, a cada Δt as pessoas em $NI1$ passam para $NI2$, as pessoas em $NI2$ passam para $NI3$ e assim, sucessivamente, até que as pessoas em NIk vão para NR .

O percentual de transmissão PT é a taxa de suscetíveis que se tornam infectados, através do qual se determina a quantidade de pessoas que passam de NS para $NI1$, em cada Δt . Este percentual é obtido através da saída inferida por um sistema difuso de Mamdani com decodificação por centro de área [KLIR and YUAN 1995]. As entradas numéricas do sistema difuso são a porcentagem ρ_I de pessoas infectadas e a porcentagem ρ_R de pessoas recuperadas, ambas na faixa de valores entre 0 e 1. Estas porcentagens são obtidas dividindo-se o número de pessoas infectadas ou recuperadas de uma área considerada pelo número total de pessoas da mesma área. No modelo, esta área pode ser local (apenas uma célula do autômato considerada), possuir diferentes níveis (vizinhança de uma célula considerada, cujo valor do raio varia conforme o nível), ou atingir a região inteira (conjunto de todas as células do autômato, diferentes de uma célula considerada). Cada uma das áreas da influência de infecção representa um risco para indivíduos suscetíveis de uma célula considerada do autômato, onde localmente estes se encontram ameaçados pelos agentes transmissores que vivem na mesma célula e os que se movem para as demais células são ameaçados pelos agentes transmissores que habitam naqueles espaços celulares. A Figura 1 ilustra as diferentes áreas, onde a célula considerada está em vermelho e as células da área da influência de infecção estão em verde.

Sendo assim, calcula-se ρ_I e ρ_R , em relação a uma célula $H(i, j)$, de acordo com a influência considerada:

1. Influência local:

(a) $\rho_I = NI_{i,j}/NH_{i,j}$; (b) $\rho_R = NR_{i,j}/NH_{i,j}$.

2. Influência não local:

(a) $total_I = (NI_{i,j} \cdot \rho_{MOB} \cdot \rho_{ASS}) + \sum_{\alpha \in V_r} NI_{\alpha}$;

(b) $total_R = (NR_{i,j} \cdot \rho_{MOB}) + \sum_{\alpha \in V_r} NR_{\alpha}$;

(c) $total_H = ((NS_{i,j} + NR_{i,j}) \cdot \rho_{MOB}) + (NI_{i,j} \cdot \rho_{MOB} \cdot \rho_{ASS}) + \sum_{\alpha \in V_r} NH_{\alpha}$;

(d) $\rho_I = total_I / total_H$; (e) $\rho_R = total_R / total_H$,

onde $total_H$, $total_I$ e $total_R$ são a quantidade total de pessoas, pessoas infectadas e pessoas recuperadas que saem da célula $H(i, j)$, em um intervalo de tempo, somadas com a quantidade total de pessoas, pessoas infectadas e pessoas recuperadas das células vizinhas de $H(i, j)$, e V_r é o conjunto de todos os vizinhos da célula $H(i, j)$ para um raio de vizinhança qualquer.



Figura 1. Diferentes áreas da influência de infecção presentes no modelo.

No modelo, em cada Δt primeiramente é aplicado a influência local, depois com o vetor de atributos resultante da célula se aplica as demais influências não locais, sempre por ordem de abrangência. É importante ressaltar que apenas as pessoas suscetíveis que se deslocam em um intervalo de tempo, podem ser infectadas devido a uma influência não local.

Uma implementação bastante geral de um autômato celular difuso para o modelo acima, e utilizado neste trabalho, tem duas variáveis linguísticas de entrada: I (indivíduos infectados) e R (indivíduos recuperados), cada qual associada ao mesmo conjunto de termos

linguísticos: B (Baixa), M (Média), A (alta). A base de regras foi definida de forma empírica e está descrita abaixo:

1. **Se** I é B , **Então** PT é B ;
2. **Se** I é M , **Então** PT é M ;
3. **Se** I é A , **Então** PT é A ;
4. **Se** R é B , **Então** PT é A ;
5. **Se** R é A , **Então** PT é B .

Ao se observar as regras, percebe-se que neste modelo as pessoas que estão recuperadas (ou imunes contra a doença) servem como uma barreira à transmissão do vírus, uma vez que a probabilidade de um agente transmissor infectar um humano suscetível irá diminuir.

3. Aplicação do Modelo na Difusão da Infecção pelo Vírus da Dengue

O modelo geral descrito na Seção anterior foi implementado e aplicado para o caso da dengue, usando o ambiente de programação para modelagem dinâmica espacial TerraME [CARNEIRO 2006]. O TerraME tem como características fornecer uma ligação ao banco de dados geográficos TerraLib [CAMARA and et al. 2000], e uso da linguagem Lua [IERUSALIMSKY 2006] para criação de modelos. Para implementar o sistema de inferência difuso foi utilizada a biblioteca LuaFuzzy², totalmente escrita na linguagem Lua e que implementa o sistema de Mamdani. Devido à inexistência da função sigmoidal na biblioteca LuaFuzzy, esta função foi implementada e adicionada à biblioteca.

3.1. A Construção do Ambiente Celular

A área de estudo selecionada para a aplicação do modelo proposto neste artigo foi a Ilha do Governador, situada na cidade do Rio de Janeiro, RJ. A partir do TerraView³ foi criado um Banco de Dados Geográfico (BDG) para a Ilha do Governador. O BDG da área de estudo possui uma camada que contém o mapa de setores censitários e as tabelas estáticas ligadas ao mapa, que contém dados de identificação como o geocódigo e nome do bairro para cada setor censitário.

Através do TerraView, o mapa pertencente à camada do BDG foi dividido em células quadradas de resolução 134 x 116, o que deu origem à uma nova camada no BDG formada por estas células. Os atributos de cada célula (por exemplo, o nome do bairro ao qual a célula pertence) foram definidos através da Classe Majoritária do TerraView, que é um operador que determina dentre todos os polígonos que fazem intersecção com a célula (no caso, os setores censitários) aquele que tem a maior área de intersecção com a mesma. Assim sendo, os atributos de cada célula são os mesmos atributos do setor censitário que tem a maior área de intersecção com a mesma. E com base em uma imagem de satélite, foram definidas manualmente algumas células que não contêm nenhuma pessoa (florestas, aeroporto, etc.). Também para as células cuja porcentagem total de intersecção com os polígonos dos setores censitários fosse menor que 0.5 (50%), foi considerado que não fazem parte de nenhum bairro e que não possuem, portanto, pessoas.

Os dados contendo a população e o número de casos de dengue na Ilha do Governador (RJ) foram obtidos a partir da Secretaria Municipal de Saúde e Defesa Civil da Prefeitura do Rio de Janeiro⁴. Estes estão divididos por bairros, os casos de dengue são referentes ao ano 2001, e a população segundo o censo do ano 2000. Com relação aos casos de

²<http://luaforge.net/projects/luafuzzy/>

³<http://www.dpi.inpe.br/terraview/index.php>

⁴<http://www.rio.rj.gov.br/web/smsdc>

dengue, foi considerado que dentre todos os possíveis casos, apenas 40% foram notificados ao SINAN (Sistema de Informação de Agravos de Notificação), correspondentes aos sintomáticos de cama. Logo, temos que: $C_{NOT} = 0.4.C_T \Rightarrow C_T = 2.5.C_{NOT}$, onde C_{NOT} é o número de casos notificados e C_T é o número de total de casos. O número total de casos foi dividido em casos de pessoas recuperadas (todos os casos até novembro de 2001) e casos de pessoas infectadas (dezembro de 2001). A população e estes números de casos de pessoas infectadas e recuperadas foram, então, distribuídos uniformemente nas células que possuem pessoas, e segundo o bairro ao qual pertencem.

3.2. Particularidades do Modelo

Para a simulação da difusão da infecção pelo vírus da dengue na Ilha do Governador foram considerados três estágios do período de infecção, logo se tem o: número de humanos infectados há um curto prazo $NI1$, número de humanos infectados há um médio prazo $NI2$ e número de humanos infectados há um longo prazo $NI3$. O número de suscetíveis em cada célula é obtido através da diferença entre total de humanos e a soma entre humanos infectados e recuperados: $NS = NH - (NI + NR)$.

Foram consideradas duas áreas da influência de infecção: local (própria célula onde habitam as pessoas) e a de nível um (células vizinhas, de raio um, à célula onde habitam as pessoas). Neste modelo o agente transmissor (que está implícito no modelo) é o mosquito *Aedes aegypti*, e por simplificação, foi considerado que existe somente um sorotipo circulando, uma vez que na epidemia de 2001 no Rio de Janeiro predominou o sorotipo DENV3 [PASSOS et al. 2004].

A Figura 2 ilustra o comportamento de uma célula $H(i, j)$ do autômato em uma iteração. No mundo real ocorrem simultaneamente os tipos de influência, mas por motivos de implementação e esquematização, decidiu-se aplicar a influência local e, posteriormente, não local (conforme descrito na subseção 2.1) para encontrar os valores de ρ_I e ρ_R . Utilizando estes valores como entrada no sistema difuso, este retorna como saída precisa o percentual de transmissão decodificado. Este percentual é, então, usado para se calcular os novos valores de suscetíveis, infectados e infectados no 1º estágio. Ao final de uma iteração, as pessoas infectadas, no início da iteração, passam para o estágio seguinte da doença, e as pessoas no último estágio da doença se tornam recuperadas contra esta.

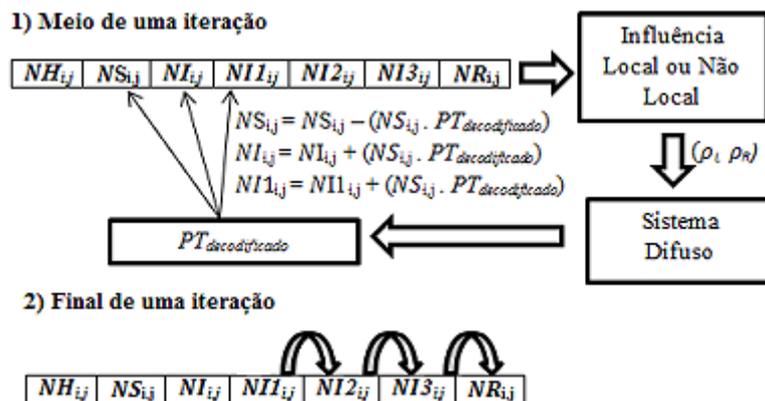


Figura 2. Comportamento de uma célula $H(i, j)$ do autômato em uma iteração.

3.3. Resultados

Nas simulações, assim como na construção do ambiente celular, foi considerado que 40% dos casos de dengue são sintomáticos, ou seja, $\rho_{ASS} = 0.6$. Este modelo possui três

estágios de infecção ($k = 3$), conforme mencionado na subseção anterior. O período em que as pessoas permanecem em um dos estágios da infecção é de 2 dias ($\Delta t = 2$), ou seja, uma pessoa permanece infectada num total de 6 dias, antes de se recuperar. É considerado que no tempo $t = 0$, todas as pessoas infectadas do ambiente celular (conforme construído na subseção 3.1), estão no primeiro estágio da infecção. Na simulação, a taxa de mobilidade humana para células vizinhas é de 30% ($\rho_{MOB} = 0.3$). As variáveis de entrada do Sistema de Inferência Difuso e a faixa de valores numéricos definidos para a variável de saída (0 a 0.05) foram definidas empiricamente. Nas simulações foram utilizadas funções gaussianas e sigmoidais, conforme ilustra a Figura 3. É bom ressaltar que ao se modificar estas, poderá ocasionar em resultados diferentes.

O resultado para esta simulação está ilustrado no item "b" da Figura 4, através de gráficos para as populações de suscetíveis, infectados e recuperados. Para efeito de comparação, o item "a" da Figura 4 ilustra a evolução temporal de um modelo SIR simples no qual não há renovação de humanos (população constante, sem mortes e emigrações ou nascimentos e imigrações). É importante ressaltar aqui que uma vez que o mosquito *Aedes aegypti* está implícito através de humanos infectados, este modelo serve apenas como um exemplo inicial ou ilustração do uso do modelo SIR por meio de autômatos celulares difusos, e não tem a pretensão de servir como meio de previsões, planejamento de controle da doença ou afins.



Figura 3. Variáveis linguísticas de entrada e saída das simulações.

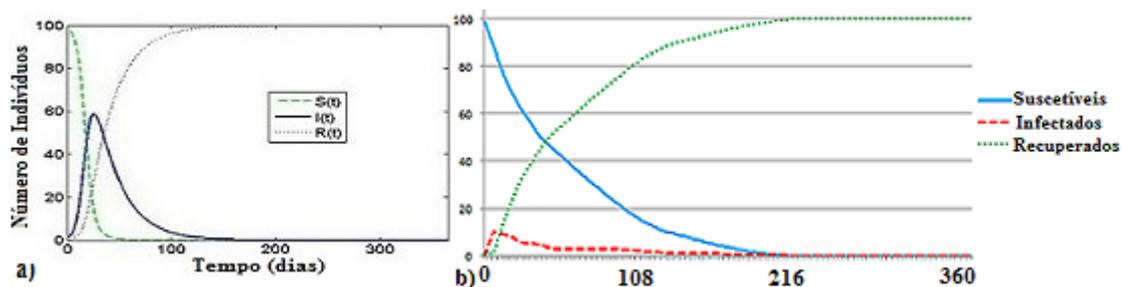


Figura 4. a) Gráfico de um modelo SIR onde não há renovação de humanos [BIOMAT04]; b) Resultados da simulação.

4. Conclusão e Trabalhos Futuros

O modelo compartimental SIR (Suscetíveis - Infectados - Recuperados) é utilizado na modelagem da dinâmica de transmissão de doenças infecciosas. No entanto, o SIR clássico não leva em consideração explicitamente a dimensão espacial de transmissão da doença. Ao implementar este modelo por meio de um autômato celular é possível descrever a evolução espaço-temporal da doença. Mas devido à complexidade na forma de transmissão de certas doenças, principalmente aquelas que envolvem vetores transmissores, a criação da função de transição dos autômatos celulares pode não ser trivial. O uso de sistemas difusos permite utilizar o conhecimento de especialistas para modelar a função

de transição, apesar da incerteza e imprecisão inerentes ao conhecimento humano. Ao observar os resultados obtidos da simulação do modelo geral proposto neste trabalho na difusão da infecção pelo vírus da dengue, percebe-se que este é capaz de representar qualitativamente o comportamento dos modelos epidemiológicos compartimentais do tipo Suscetíveis-Infetados-Recuperados [KEELING and DANON 2009, WHITE et al. 2007]. A abordagem aqui proposta considera duas importantes características para modelos de transmissão de doenças infecciosas do tipo SIR: o percentual dos recuperados e o percentual dos assintomáticos. Quanto mais pessoas recuperadas da doença (ou seja, imunizadas), mais lentamente a doença será transmitida, já que o número de contatos de suscetíveis com infectados diminui. Já os assintomáticos são os responsáveis diretos pela continuidade em massa da transmissão da doença, pois estes continuam suas rotinas diárias de mobilidade normalmente.

Ao simular um modelo SIR por meio de um autômato celular difuso é possível representar a dinâmica de uma epidemia e criar um ambiente onde se pode estudar o espalhamento da doença, e observar o padrão espacial de sua prevalência, ao mesmo tempo em que se permite lidar mais facilmente com diferentes níveis de incerteza e imprecisão que ocorrem ao utilizar conceitos vagos e parâmetros de natureza subjetiva. No caso das doenças transmitidas por mosquitos vetores, uma grande variedade de fatores influenciam a dinâmica espacial e temporal das populações de mosquitos e, conseqüentemente, os padrões de transmissão da doença nas populações humanas [GUBLE et al. 2001]. Estudar os fatores influenciadores nos padrões espaciais de transmissão de doenças vectoriais como a dengue é, de certa forma, complicado, já que seriam necessários dados refinados tanto de coleta dos vetores e suas fases aquáticas quanto de dados climáticos. O cumprimento das funções de vigilância epidemiológica, no entanto, depende da disponibilidade e qualidade de dados que sirvam para subsidiar o processo de produção de informação para a ação[SAÚDE]. Assim, a simulação via modelagem computacional torna-se um importante instrumental no estudo dos cenários e padrões espaciais de transmissão de doenças, como também no auxílio à tomada de decisões. Para suprir a falta de dados espacializados refinados, uma alternativa para a compreensão melhor do tema é utilizar o conhecimento, por exemplo, de especialistas, inclusive com relação à área de estudo. Desta forma, a abordagem por modelagem computacional utilizando lógica difusa mostra-se relevante neste tipo de estudo.

Em trabalhos futuros, espera-se adicionar no modelo geral proposto uma taxa de renovação de humanos, que vai permitir que a epidemia ocorra em ciclos, uma vez que sempre existirão pessoas suscetíveis à doença. Outra importante melhoria é incluir explicitamente os agentes transmissores no modelo, que também podem migrar para diferentes ambientes. Diferentes áreas da influência de infecção também podem ser definidas e, logo, ser formulados novos meios de se obter a porcentagem de pessoas infectadas e recuperadas. Também a base de regras difusas, que modela a função de transição dos autômatos celulares, pode ser aprimorada a partir, eventualmente, da literatura e de informações provenientes de especialistas, com a possível adição de novas variáveis, novos termos difusos, etc. É importante frisar que o uso de sistemas difusos permite um rápido processo de prototipagem, o que nem sempre ocorre quando a função de transição é analítica, dada sua complexidade. Em relação ao modelo da dengue, nos próximos trabalhos outros sorotipos podem ser incluídos na simulação, bem como a inclusão dos mosquitos *Aedes aegypti*, que interagem com os humanos e se movem pelos espaços

próximos. Outras áreas da influência de infecção também podem ser adicionadas no modelo, assim como podem ser aprimoradas as definições dos termos linguísticos.

Referências

- Biomat04. <http://biomat04.wikispaces.com/conceitos/>.
- CAMARA, G. and et al. (2000). Terralib: Technology in support of gis innovation. *II Brazilian Symposium on Geoinformatics, GeoInfo2000*.
- CARNEIRO, T. G. S. (2006). *Nested-CA: A Foundation for Multiscale Modeling of Land Use and Land Change*. PhD thesis, INPE, São José dos Campos.
- GUBLE, D., REITER, P., EBI, K., YAP, W., NASCI, R., and et al. (2001). Climate variability and change in the united states: potential impacts on vector and rodent-borne diseases. *Environmental Health Perspectives*, 109(2):223–233.
- HETHCOTE, H. W. (2000). The mathematics of infectious diseases. *SIAM Review*, 42(4):599–653.
- IERUSALIMSCHY, R. (2006). *Programming in Lua*. Rio de Janeiro: Lua.org, 2^a edition.
- KEELING, M. and DANON, L. (2009). Mathematical modelling of infectious diseases. *British Medical Bulletin 1-10*: doi: 10.1093/bmb/ldp038.
- KERMACK, W. and MCKENDRICK, A. (1927). A contribution to the mathematical theory of epidemics. *Proceedings of the Royal Society of London Series A: Mathematical and Physical Sciences*, A115(4):700–721.
- KLIR, G. J. and YUAN, B. (1995). *Fuzzy sets and fuzzy logic: theory and applications*. New Jersey: Prentice Hall.
- MRAZ, M., ZIMIC, N., and VIRANT, J. (1999). Intelligent bush fire spread prediction using fuzzy cellular automata. *Journal of Intelligent and Fuzzy Systems*, 7(2):203–207.
- NGUYEN, H. T. and WALKER, E. A. (2006). *A First course in fuzzy logic*. New York: Chapman & Hall/CRC.
- OLIVEIRA, E., LACERDA, M., BARBOSA, A., and NEPOMUCENO, E. (2008). Desenvolvimento de estratégia de controle epidemiológico: Análise espacial e vacinação a partir do foco da doença. *Anais do XVII Congresso Brasileiro de Automática*, pages 1–6.
- PASSOS, M., SANTOS, L., PEREIRA, M., CASALI, C., FORTES, B., VALENCIA, L., ALEXANDRE, A., and MEDRONHO, R. (2004). Diferenças clínicas observadas em pacientes com dengue causadas por diferentes sorotipos na epidemia de 2001/2002, ocorrida no município do rio de janeiro. *Rov. Soc. Bras. Med. Trop.*, 37(4).
- SAÚDE, M. *Guia de vigilância epidemiológica*. 6^a edition.
- WHITE, S., REY, A. M. D., and SANCHEZ, R. (2007). Modeling epidemics using cellular automata. *Appl. Math. Comput.*, 186:193–202.
- YANG, H. M. (2001). *Epidemiologia matemática: Estudos dos Efeitos da Vacinação em Doenças de Transmissão Direta*. Ed. da Unicamp.
- ZADEH, L. A. (1965). Fuzzy sets. *Information and control*, 8:338–353.