

Explorando o uso da Teoria Fuzzy no software R: estudo de caso para avaliar a força de infecção da Anemia Infecciosa Equina (AIE)

Gabriela Servidone¹

Sônia Ternes²

Marcelo Rossi³

Resumo: A Anemia Infecciosa Equina (AIE) é uma doença grave no Pantanal Sul-Mato-Grossense, e uma das possíveis vias de infecção é a iatrogênica. A obtenção de dados do processo infeccioso no campo que permitam estimar a força de infecção da transmissão da AIE é uma etapa onerosa, pois envolve altos custos e o bem estar animal. Para contornar a escassez de tais dados, implementou-se um sistema baseado em regras fuzzy para mimetizar a força de infecção correspondente ao reúso de possíveis agulhas contaminadas nos cavalos. Este parâmetro epidemiológico foi gerado por diferentes t-normas, verificando-se sua importância nos respectivos resultados do modelo matemático.

Palavras-chave: teoria de conjuntos fuzzy, força de infecção, epidemiologia.

Introdução

O cenário científico mundial mostra o crescimento vertiginoso da capacidade de aquisição de dados de diferentes naturezas. Com relação à pesquisa agropecuária, seus diversos setores não podem ser tratados de maneira

¹ Estudante de Matemática Aplicada e Computacional na Universidade Estadual de Campinas (Unicamp), bolsista PIBIC da Embrapa Informática Agropecuária, Campinas, SP.

² Pós-doutor do Laboratoire d'Ingénierie pour les Systèmes Complexes (LISC-Cemagref), Pesquisadora da Embrapa Informática Agropecuária, Campinas, SP.

³ Doutor do Programa Inter unidades em Biotecnologia da Universidade de São Paulo (USP), bolsista PNPd/CNPq da Embrapa Informática Agropecuária, Campinas, SP.

homogênea no que diz respeito à medição de variáveis e parâmetros de interesse.

Ao trabalhar com dados com incerteza, a informação deixa de ser representada por um valor determinado e passa a ser representada por um conjunto de dados imprecisos. Desta forma, o uso da teoria clássica dos conjuntos torna-se inviável devido à sua ineficiência no tratamento dessas informações. Entretanto, essas incertezas podem ser estudadas e modeladas, de forma mais robusta, utilizando a teoria dos conjuntos nebulosos, também conhecida como teoria dos conjuntos fuzzy (MEIRELLES et al., 2007).

Conforme Bönisch et al. (2004), a lógica fuzzy é capaz de modelar, de maneira aproximada, o raciocínio humano, visando manipular informações em um ambiente de incerteza e imprecisão. Esse ferramental pode fornecer uma resposta aproximada, baseada em um conhecimento inexato, incompleto ou não totalmente confiável.

Este trabalho tem como objetivo realizar um experimento utilizando lógica fuzzy, buscando avaliar quais t-normas e t-conormas são mais adequadas para representar a força de infecção do possível uso de agulhas contaminadas no modelo determinístico da dinâmica de propagação da AIE (TERNES; VILAMIU, 2013). Esse modelo é representado pelo seguinte sistema de equações diferenciais ordinárias:

$$\left\{ \begin{array}{l} \frac{dS_h}{dt} = \Phi - (\lambda_h + \lambda_n + \mu)S_h \\ \frac{dI_h}{dt} = (\lambda_h + \lambda_n)S_h - \mu I_h \\ \frac{dS_v}{dt} = \varphi(S_v + I_v) + \varepsilon I_v - (\lambda_v - \zeta)S_v \\ \frac{dI_v}{dt} = \lambda_v S_v - (\varepsilon + \zeta)I_v \end{array} \right. \quad \text{onde} \quad \lambda_h = \frac{\alpha N_v I_v}{N_h}, \quad \lambda_v = \beta I_h, \quad \lambda_n = \gamma I_v$$

No modelo acima, as variáveis de interesse representam a população de animais suscetíveis (S_h), equinos infectados (I_h), mutucas infectivas (S_v) e mutucas não infectivas (I_v). Os parâmetros λ_h , λ_v e λ_n representam, respectivamente, a força de infecção da AIE em equinos transmitida por mutucas; força de infecção em mutucas, devido à alimentação do sangue contaminado de equinos e a força de infecção em equinos, a partir do possível reúso de agulhas contaminadas, sendo esta última o objetivo de interesse deste estudo.

Materiais e Métodos

Para criar um sistema baseado em regras fuzzy, são necessárias as seguintes informações: a) funções de pertinência das variáveis de entrada e saída; b) base de regras fuzzy; c) definição da t-norma, t-conorma e implicação; d) modelo de inferência; e) processador de saída (defuzzificação).

Considerando o conceito biológico de λ_n , utilizou-se, como variáveis dependentes de entrada, a taxa de injeções aplicadas por ano e o volume morto da agulha, como mostrado nas Figuras 1a e 1b.

Para representar a função de pertinência de saída, a variável independente escolhida é a força de infecção dada por $\gamma = \alpha * \beta$, onde β é a taxa de injeções e α o volume morto da agulha (Figura 1c).

O próximo passo é a definição da base de regras. Neste momento é despendido tempo na verificação das pertinências, no estudo do modelo biológico, para que as regras sejam as mais realistas possíveis.

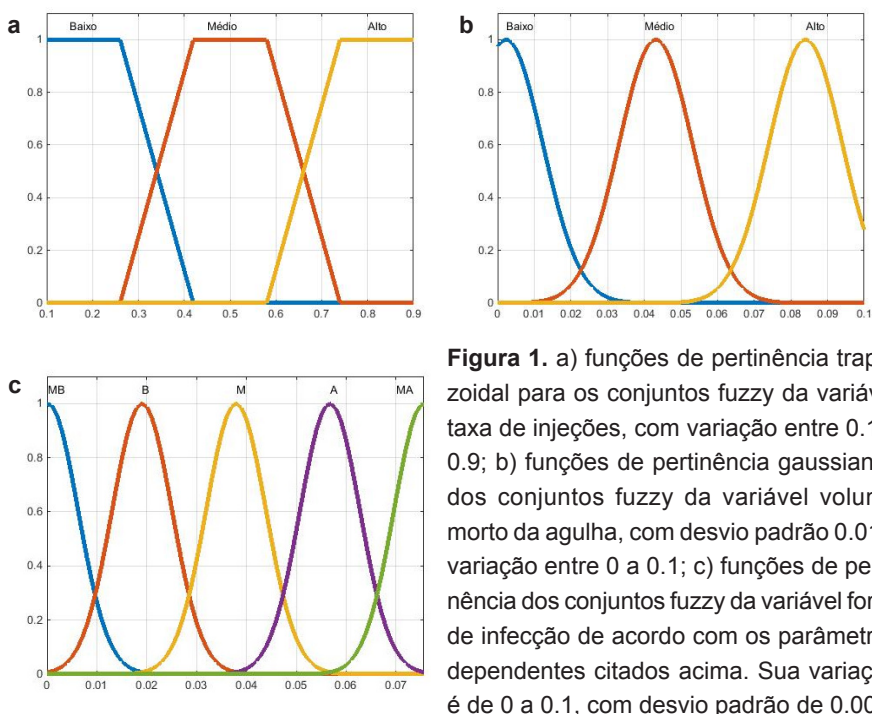


Figura 1. a) funções de pertinência trapezoidal para os conjuntos fuzzy da variável taxa de injeções, com variação entre 0.1 a 0.9; b) funções de pertinência gaussianas dos conjuntos fuzzy da variável volume morto da agulha, com desvio padrão 0.01 e variação entre 0 a 0.1; c) funções de pertinência dos conjuntos fuzzy da variável força de infecção de acordo com os parâmetros dependentes citados acima. Sua variação é de 0 a 0.1, com desvio padrão de 0.006.

Baseado em dados da literatura sobre a infecção pelo Human Immuno deficiency Virus (HIV), que é da mesma família do vírus transmissor da AIE, verificou-se que para agulhas de pequeno volume morto, a possibilidade de infecção é menor. Portanto, tal possibilidade é dependente do volume do conjunto agulha/seringa utilizada por pessoas usuárias de drogas (ABDALA et al., 1999; BOBASHEV; ZULE, 2010). Esta realidade direcionou a escolha das regras para a força de infecção, descritas na Figura 2.

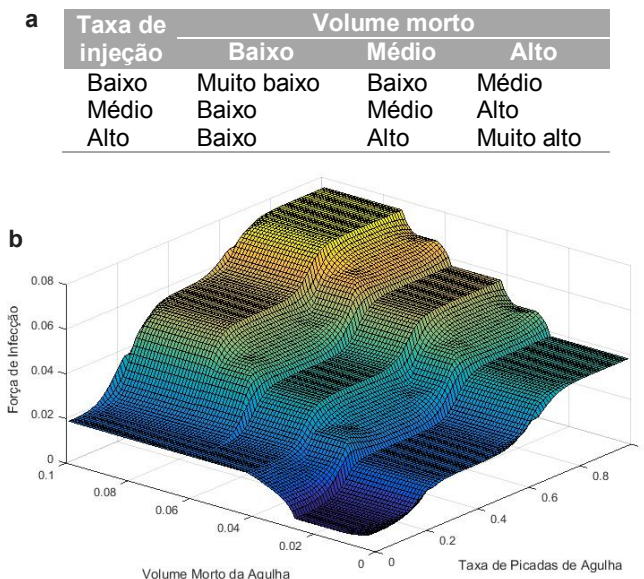


Figura 2. a) base de regras definida na Tabela de acordo com o modelo biológico estudado: Anemia Infecciosa Equina; b) relação gráfica dos valores obtidos aplicados a essa base de regras com t-norma MIN e t-conorma MAX.

O software R (R DEVELOPMENT CORE TEAM, 2015) possui o pacote Fuzzy Rule-Based Systems (FRBS) (RIZA et al., 2015), o qual possui em sua estrutura as diferentes t-normas e t-conormas que foram testadas para o sistema fuzzy que modela a força de infecção do modelo compartimental determinístico. Estas normas são comentadas, de forma reduzida, na Tabela 1.

Importante observar que cada t-norma tem sua respectiva t-conorma para o cálculo. Por fim, utilizando o modelo de inferência de Mamdani e o cálculo do centroide como método de defuzzificação, os conjuntos fuzzy resultantes do

Tabela 1. Listagem das t-normas e t-conormas utilizadas no experimento.

	t-norma	t-conorma
Min/Max	$\min(x_1, x_2)$	$\max(x_1, x_2)$
Hamacher	$(x_1 \cdot x_2) / (x_1 + x_2 - x_1 \cdot x_2)$	$(x_1 + x_2 - 2x_1 \cdot x_2) / (1 - x_1 \cdot x_2)$
Yager	$1 - \min(1, ((1 - x_1) + (1 - x_2)))$	$\min(1, (x_1 + x_2))$
Product	$(x_1 \cdot x_2)$	$(x_1 + x_2 - x_1 \cdot x_2)$
Bounded	$\max(0, x_1 + x_2 - 1)$	$\min(1, x_1 + x_2)$

processo de implicação da base de regras são defuzzificados para a variável de saída “força de infecção”.

Resultados e Discussão

Após o teste das cinco diferentes combinações de t-normas e t-conormas, encontramos a variação descrita na Figura 3. Uma matriz 10x10 com parâmetros igualmente espaçados foi construída com valores estimados de força de infecção e alimentado o respectivo parâmetro nas simulações do modelo. O gráfico demonstra a prevalência dos cavalos infectados no período de 4 anos.

Claramente, quando analisamos entre YAGER e MIN/MAX, vemos que a t-norma de MIN é biologicamente mais aceitável por refletir uma variação

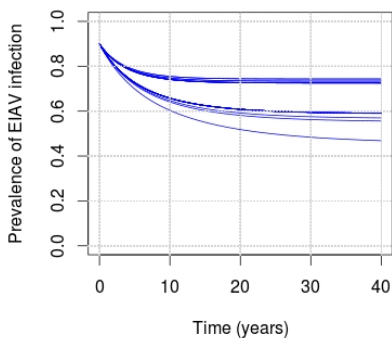
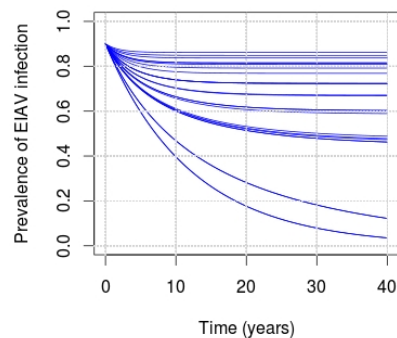
a Prevalence of infected horse - YAGER**b** Prevalence of infected horse - MIN

Figura 3. Gráfico da porcentagem de cavalos infectados por anos. (a) t-norma e t-conorma YAGER. (b) t-norma MIN e t-conorma MAX. A simulação foi feita para 100 diferentes valores de entrada (força de infecção).

da porcentagem de cavalos infectados em relação às diferentes forças de infecção utilizadas. Resultados semelhantes foram obtidos quando testadas as normas PRODUCT e HAMACHER. Já a BOUNDED, não teve um comportamento esperado, pois todos os valores retornados pelos diferentes inputs foram 0.04.

Considerações Finais

Para a organização do experimento, foi primeiramente avaliado se as definições de t-normas, t-conormas, implicação, defuzzificação e tipo do modelo eram as preferíveis para a utilização. Desse modo, foi possível fazer um levantamento do que o pacote FRBS do software R podia oferecer como recurso. Inicializando os parâmetros do modelo como constantes, foi analisado como essas definições se comportavam quando eram variadas. O resultado obtido foi o esperado: mesmo sendo a norma do mínimo e a conorma do máximo as mais simples, elas foram as que retornaram uma abrangência de valores os quais são biologicamente aceitáveis. Outras normas como a do produto e a Hamacher retornaram resultados iguais, mas como seus cálculos possuem grau de complexidade maior, mais uma vez o uso das normas de mínimo e máximo se mostra mais adequado.

Como trabalho futuro, visualiza-se testar o efeito da quantidade de conjuntos fuzzy (granularidade) no particionamento das variáveis, bem como diferentes formatos das funções de pertinência.

Agradecimentos

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), pelo suporte financeiro (Processo 160212/2015-4).

Referências

ABDALA, N.; STEPHENS, P. C.; GRIFFITH, B. P.; HEIMER, R. Survival of HIV-1 in syringes. *Journal of Acquired Immune Deficiency Syndromes*, v. 20, n. 1, p. 73-80, Jan.1999.

BOBASHEV, G. V.; ZULE, W. A. Modeling the effect of high dead space syringes on the human immunodeficiency virus (HIV) epidemic among injecting drug users. **Addiction**, v. 105, n. 8, p. 1439-47, Aug. 2010. DOI: 10.1111/j.1360-0443.2010.02976.x.

BÖNISCH, S.; ASSAD, M. L. L.; CÂMARA, G.; MONTEIRO, A. M. V. Representação e propagação de incertezas em dados de solos. I. Atributos categóricos. **Revista Brasileira de Ciências do Solo**, v. 28, n. 1, p. 21-32, jan./fev. 2004.

MEIRELLES, M. S. P.; MOREIRA, F. R.; CAMARA, G. Técnicas de inferência espacial. In: MEIRELLES, M. S. P.; CAMARA, G.; ALMEIDA, C. M. de (Ed). **Geomática: modelos e aplicações ambientais**. Brasília, DF : Embrapa Informação Tecnológica, 2007. Cap. 3.

R DEVELOPMENT CORE TEAM. **R: a language and environment for statistical computing**. Vienna: R Foundation for Statistical Computing, 2015. Disponível em: <<http://www.R-project.org/>>. Acesso em: 15 out. 2015.

RIZA, L. S.; BERGMEIR, C.; HERRERA, F.; BENÍTEZ, J. M. Frbs: fuzzy rule-based systems for classification and regression in R. **Journal of Statistical Software**, v. 65, n. 6, p. 1-30, May 2015. Disponível em: <<http://www.jstatsoft.org/v65/i06/>>. Acesso em: 15 out. 2015.

TERNES, S.; VILAMIU, R. **Epidemiologia matemática da Anemia Infecciosa Equina**. In: WORKSHOP PROJETO, 1., 2013, Corumbá. Corumbá: [s.n.], 2013. Slides. Trabalho não publicado.