

Avaliação da Exposição de Contaminantes no Solo Através de Técnicas de Mineração de Dados

Sérgio Fred Ribeiro Andrade¹, Fermin C. Garcia Velasco², Francisco Heriberto Martinez Luzardo³

^{1,2,3} Departamento de Ciências Exatas e Tecnológicas – Universidade Estadual de Santa Cruz (UESC)
Ilhéus – Ba – Brasil

sergiof@uesc.br, fermin@uesc.br, fmartinezluzardo@gmail.com

Abstract. *The assessment of chemical contaminant exposure is an important and necessary process for knowing the ills caused by them to environment and managing their mitigating actions. In this paper are illustrated an application and comparison of algorithms data mining in order to classify them, aiming to evaluate the environmental risk of its exposure, considering physical properties and the migration of contaminants in sediments. The method is based on the approach of soil contaminant adsorption for previous supervised training and techniques of multi classification of risk categories through statistical learning. In applying the algorithms on simulated samples were achieved statistical indicators with full and satisfactory classification among the presented categories, pointing techniques are viable for use in the assessment of risk exposure and environmental impact studies.*

Resumo. *A avaliação da exposição por contaminantes químicos é um processo importante e necessário para conhecimento dos males provocados ao ambiente e para a gestão de suas ações mitigadoras. Neste artigo são ilustradas a aplicação e a comparação de algoritmos de mineração de dados para classificação, visando avaliação da exposição de riscos ambientais, onde são consideradas as propriedades físicas e a migração dos contaminantes em sedimentos. A metodologia baseia-se na abordagem da adsorção de contaminantes pelo solo para o prévio treinamento supervisionado e em técnicas da multiclassificação de categorias de risco através da aprendizagem estatística. Na aplicação dos algoritmos nas amostras simuladas foram alcançados indicadores estatísticos com plena e satisfatória classificação entre as categorias apresentadas, apontando serem técnicas viáveis para aplicação nos processos de avaliação da exposição de risco e em estudos de impactos ambientais.*

1. Introdução

Constitui um importante problema ambiental a contaminação por substâncias químicas em sedimentos nos diversos meios, principalmente, nos solos superficiais e subterrâneos através de dejetos industriais e outros. A avaliação da exposição por contaminantes é um processo complexo que exige estudo interdisciplinar e juízo científico, para a obtenção de resultado consistente para a gestão e mitigação das áreas afetadas.

Existem vários modelos de avaliação que podem ser adotados para estimar a exposição de seres humanos e biota às substâncias nocivas nos diversos compartimentos ambientais. Segundo a WHO (2010), a exposição pode ser estimada por um exato período de tempo ou integrada ao longo de um tempo que varia de minutos a uma vida inteira. As saídas podem gerar valores médios ou por mediana, parâmetros de distribuição ou de probabilidades, incorrendo em modelos de exposição que variam amplamente em complexidade, abordagem, diversidades de variáveis de entradas e saídas.

Em geral, observa-se que o emprego de técnicas estatísticas de regressão para representações de distribuições probabilísticas, são generalizações de grande escala e podem ser encontrados largamente na literatura. Por outro lado, além das técnicas de regressão, existem outros métodos que também podem ser empregados nos processos de avaliação ambiental, como o reconhecimento de padrão, análise de *clusters* e classificação.

Sobre a classificação, especificamente para categorização de risco ambiental muito pouco tem sido registrado. Segundo Spadotto (2002), os impactos ambientais podem ser classificados segundo os critérios de valor, ordem, espaço, tempo, dinâmica ou contaminação por produtos nocivos. Particularmente, sobre os produtos químicos em sedimentos e solos, o conhecimento das propriedades físico-químico do meio é uma exigência necessária.

Segundo Jacques et. al. (2006), o transporte e destino de contaminantes subterrâneos devem ser modelados para longa escala temporal e espacial, para garantir o conhecimento mais apropriado da heterogeneidade dos parâmetros físicos e geoquímicos, como exemplo das alterações na infiltração devido à precipitação e irrigação, períodos de evaporação, mudanças de temperatura, influência do carbono no solo, alterações no pH da água do solo, e outros, que afetam soluções em meios porosos e geram novas condições de troca catiônicas.

A motivação para esse estudo decorreu da proposição de comparar técnicas de descoberta do conhecimento, através de algoritmos de mineração de dados, para avaliação da exposição por contaminantes em solos, empregando métodos de classificação em riscos ambientais, visando conhecer os modelos mais apropriados e as ameaças de riscos potenciais para a saúde humana e o meio ambiente.

2. Métodos

Os dados obtidos para aplicação das técnicas no presente trabalho foram provenientes de Jacques et. al. (2002, 2008) e Seuntjens (2000), que fizeram simulação do transporte e retenção de contaminante no solo, em profundidade de até 1 m, com pluviosidade aproximada de 100 mm/dia, para 30 anos, em 7 amostras, com solo arenoso de Lommel e Kempen, na Bélgica, contaminado pela ação da indústria local onde houve lixiviação dos metais pesados Cd, Pb, Zn e Al e de cátions Na, K, Mg e Ca, por deposição atmosférica.

A metodologia seguida para avaliar exposição de contaminantes, neste trabalho, contempla o roteiro ilustrado na Figura 1, e está descrito posteriormente.

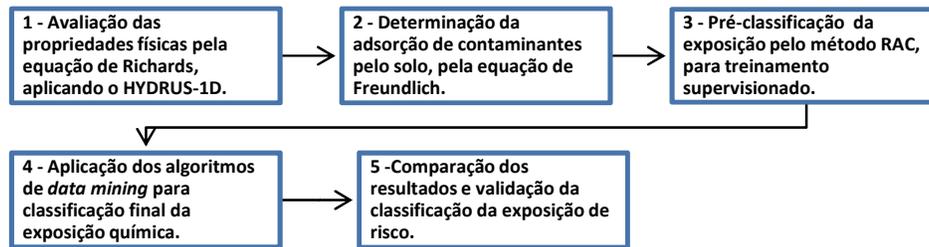


Figura 1. Etapas da metodologia para avaliação da exposição de contaminantes através das técnicas de mineração de dados.

2.1. Avaliação das propriedades físicas

A avaliação das propriedades físicas do solo é necessária para o conhecimento dos fatores que interferem no transporte do soluto em meio aquoso e sua retenção no solo quando da percolação de contaminantes químicos em sedimentos. Pode-se considerar a formação do solo insaturado, parcialmente saturado ou saturado, em região composta por meios porosos uniformes ou não, em várias camadas, na direção vertical, horizontal ou com inclinação (SIMUNEK et. al., 2012).

Os fundamentos para os parâmetros que servem como entradas para o modelo de propriedades físicas são originadas das variáveis hidráulicas e do transporte de solutos no solo, de acordo com as equações de Van Genuchten e Richards, especificadas em Simunek (2005), onde:

$$\frac{\partial(h)}{\partial t} = \frac{\partial}{\partial z} [K(h) \left(\frac{\partial h}{\partial t} + 1\right)] - S \quad (1)$$

$$\frac{\partial \theta R c}{\partial t} = \frac{\partial}{\partial z} [\theta D \left(\frac{\partial c}{\partial z}\right) - q c] - \varphi \quad (2)$$

$$\frac{\partial C(\theta) T}{\partial t} = \frac{\partial}{\partial z} [\lambda(\theta) \left(\frac{\partial T}{\partial z}\right) - C_w q T] - \varphi \quad (3)$$

Os parâmetros referidos na equação (1) correspondem a: z é a coordenada gravitacional [L], t é o tempo, h é o potencial de pressão [L], φ é a unidade volumétrica da água no solo [L³/L³], S a concentração de soluto adsorvido [M/M] e $K(h)$ em [LT⁻¹] é a função de condutividade hidráulica insaturada, dado como o produto da condutividade hidráulica relativa Kr e a condutividade hidráulica saturada Ks .

Na equação (2), conhecida como a equação de convecção-dispersão, c é a concentração da solução, R é o fator de retardamento da adsorção, D é o coeficiente de difusão [L²/T], q é a velocidade de escoamento do fluido [L/T], e θ o potencial de pressão [L]. Na equação (3), T é a temperatura, λ é a condutividade térmica aparente e, C e C_w são capacidades volumétricas térmicas do solo e do fluido líquido, respectivamente. As unidades de medidas citadas acima estão especificadas em Simunek et. al. (2012).

Os resultados das mensurações das propriedades físicas citadas são valores das variáveis de retenção do conteúdo hidráulico, saturação de conteúdo hidráulico, coeficiente de condutividade hidráulica, porosidade do solo e densidade do solo, que estão ilustrados na Tabela 3, deste.

2.2. Determinação da adsorção de contaminantes

Diante das propriedades físicas do solo, do transporte do contaminante em meio aquoso, da saturação e retenção pelo solo, aplicou-se a equação (4) para determinar concentração total de adsorção do contaminante pelo solo, que é baseada na isotérmica de Freundlich e especificada em EPA (1996), onde é proposta a fórmula seguinte cujos parâmetros estão na Tabela 1:

$$C_t = C_w \left(K_d + \frac{\theta_w + \theta_a H'}{\rho_b} \right) \quad (4)$$

Tabela 1. Parâmetros para determinação do coeficiente de adsorção do contaminante pelo solo.

Definição de parâmetro (unidade)	Fonte
C_t /concentração total no nível de solo (mg/kg)	--
C_w /concentração no solo lixiviado (mg/L)	Item 2.0 de EPA (1996)
K_d /coeficiente do conteúdo de água no solo (L/kg)	Parte 5 de EPA (1996)
θ_w /saturação da água no solo (L_{water}/L_{soil})	EPA/ORD
θ_a /saturação do ar no solo (L_{air}/L_{soil})	$n - \theta_w$
n /porosidade total do solo (L_{pore}/L_{soil})	$1 - \rho_b / \rho_s$
ρ_b /densidade do solo (kg/L)	EPA (1991 b)
H' /constante adimensional Lei de Henry	EPA (1991 b)

A equação (4) é usada para calcular a concentração total no solo C_t , considerando as substâncias principais lixiviadas no solo C_w . É assumido que a água do solo, a parte sólida e a gasosa, são conservadas durante a amostragem. Se o gás no solo é perdido durante a amostragem assume-se que θ_a é igual a zero. O K_d varia dependendo do tipo de produto químico contaminante e do solo.

2.3. Pré-classificação da avaliação da exposição de risco

Após a determinação da concentração total adsorvida pelo solo, foi realizada a pré-classificação da exposição do composto químico, visando treinamento dos dados para ajustamento das entradas nos algoritmos considerados, a fim de definir a categoria da exposição de risco por cada contaminante.

Para a classificação prévia aplicou-se o modelo linear generalizado proposto por McCullagh e Nelder (1989), para determinar a distribuição da probabilidade de Y em função das variáveis explanatórias C_t (concentração total do contaminante) e T (tempo de exposição), objetivando conhecer a probabilidade da classe de gravidade da exposição, mencionada anteriormente. A equação geral para classificação prévia é ilustrada a seguir e sua especificação está em EPA (2006).

$$Pr(Y \geq s | C_t, T) = H[\alpha_s + \beta_1 * f(C_t) + \beta_2 * f(T)]. \quad (5)$$

O Pr é a probabilidade da exposição de risco por contaminante em função das categorias dependentes, sendo Y um ordinal diante das categorias s , onde $s = 0,1,2,3,4$. Essas categorias são referenciadas segundo o RAC – *Risk Assessment Code*, proposto por Perin et. al. (1985).

O RAC indica o risco pela exposição de contaminante diante do nível de adsorção pelas trocas iônicas e catiônicas e frações de carbono observadas sobre os

metais. Quanto maior for o percentual de carbono maior o risco, conforme ilustra a Tabela 2. Uma categoria 0 corresponde a menor gravidade quando nenhum efeito adverso é encontrado e 4 corresponde ao efeito letal por exposição à substância nociva.

O H é uma função de probabilidade que resulta valores entre 0 e 1, podendo utilizar três opções do modelo generalizado na regressão linear: *logit*, *probit* ou *loglog*. O parâmetro β_1 determina a dependência de resposta da concentração total (C_t), enquanto que β_2 é o coeficiente da dependência do tempo (T), o parâmetro α é ordenado como $\alpha_1 \geq \alpha_2 \dots \geq \alpha_5$ e corresponde a uma restrição de ponto de corte para quaisquer níveis de C_t e T , sendo:

$$Y = 0 \text{ se } Y \leq \alpha_1; Y = 1 \text{ se } \alpha_1 < Y \leq \alpha_2; Y = 2 \text{ se } \alpha_2 < Y \leq \alpha_3; \dots; Y = 4 \text{ se } Y \geq \alpha_5.$$

As categorias de risco citadas para pré-classificação são necessárias para o aprendizado estatístico supervisionado dos modelos utilizados e, para a etapa final em cada algoritmo de multiclassificação que servem para definir a margem de risco da exposição pelo contaminante.

Tabela 2. Níveis de risco propostos por Perin et. al. (1985).

Classe	RAC	Critério (%)
0	Nenhum risco	< 1
1	Baixo risco	1 - 10
2	Risco médio	11 - 30
3	Risco alto	31 - 50
4	Risco muito alto	> 50

2.4. Algoritmos aplicados para avaliação da exposição

Tendo sido realizada a pré-classificação pela adsorção dos compostos químicos para treinamento estatístico, a próxima etapa foi a de escolha de algoritmos para técnica de classificação supervisionada.

Os algoritmos aplicados para classificação final e comparação dos resultados foram *LibSVM*¹, *RBFNetwork*², *MultilayerPerceptron*³, *MultiClassClassifier*⁴ e *BayesNet*⁵, que estão implementados no software *Weka*⁶ - *Waikato Environment for Knowledge Analysis*, versão 3.7.5, referenciado em Hall (2009). A escolha por esses algoritmos recaiu por contemplarem modelos para classificação de multicategorias, sem necessidade de ajustes prévios para aplicação, e que são do reconhecimento acadêmico e prático com ampla citação pela literatura específica.

Cada algoritmo foi executado sobre o conjunto de dados referenciado na Tabela 4, devidamente pré-classificado. Em cada execução, foram verificados diversos indicadores como matriz de confusão, instâncias classificadas corretamente, índice estatístico *Kappa* (k), erro médio absoluto e erro quadrático, entre outros disponíveis pelo *Weka*.

¹ <http://weka.wikispaces.com/LibSVM>

² <http://weka.sourceforge.net/doc/weka/classifiers/functions/RBFNetwork.html>

³ <http://weka.sourceforge.net/doc/weka/classifiers/functions/MultilayerPerceptron.html>

⁴ http://bio.informatics.indiana.edu/ml_docs/weka/weka.classifiers.MultiClassClassifier.html

⁵ http://www.cs.waikato.ac.nz/~remco/weka_bn/

⁶ <http://www.cs.waikato.ac.nz/ml/weka/>

No tópico seguinte está ilustrada a aplicação do modelo e comparação dos algoritmos para determinação da margem de exposição de risco segundo o *Code Assessment Risk*, as quais servem para fundamentar a base do presente trabalho e auxiliar na validação da aplicação prática.

3. Dados, Discussões e Resultados

Os parâmetros usados como *input*, contendo as propriedades físicas do solo e de transporte do soluto foram os listados na Tabela 3, onde foi assumido contaminação com os metais percolados no solo a um pH médio de 3,5, conforme Jacques et. al. (2002). Considerou-se também, que houve permuta catiônica com associação à matéria orgânica do solo e carbono e, em cada camada do solo ocorreram diferentes propriedades hidráulicas e capacidades de troca iônica.

A obtenção dos parâmetros para a retenção de água e condutividade hidráulica foram calculados pelo software HYDRUS-1D⁷ onde está implementado um modelo unidimensional de elementos finitos para simular o movimento hidráulico, de calor e de múltiplos solutos em solos (SIMUNEK, 2012).

Tabela 3. Propriedades de retenção e transporte hidráulicos dos solutos no solo.

Horizonte do Solo	Profundidade da Camada (cm)	P (g cm ³)	Carbono Orgânico (%)	θ_r	θ_s	α (cm ⁻¹)	n	K_s (cm dia ⁻¹)
A	0 – 7	1.31	2.75	0.065	0.48	1.6	1.94	1.1 10 ⁻⁵
E	7 – 19	1.59	0.75	0.035	0.42	1.5	3.21	3.6 10 ⁻⁵
Bh1	19 – 24	1.3	4.92	0.042	0.47	1.6	1.52	4.5 10 ⁻⁶
Bh2	24 – 28	1.38	3.77	0.044	0.46	2.8	2.01	1.0 10 ⁻⁴
BC	28 – 50	1.41	0.89	0.039	0.46	2.3	2.99	1.0 10 ⁻⁴
C1	50 – 75	1.52	0.12	0.03	0.42	2.1	3.72	1.0 10 ⁻⁴
C2	75 – 100	1.56	0.08	0.021	0.39	2.1	4.33	1.0 10 ⁻⁴

Esses parâmetros foram originados das equações (1), (2) e (3) e utilizados para o processo geoquímico que simulou a interação entre os cátions e os metais no solo e solutos infiltrados, a troca de íons, a adsorção dos contaminantes no solo e a composição das substâncias pela reação química natural.

Em consideração ao objetivo deste trabalho e ao pouco espaço textual, não foram detalhados os processos geoquímicos, as especiações dos compostos e o processo de mudança das concentrações químicas, ocorridas pelas variações do clima e presença de outros elementos naturais.

Para categorização prévia da exposição dos compostos nocivos foram calculadas as concentrações totais (C_t) pela adsorção dos contaminantes, para o processo de treinamento supervisionado das técnicas de mineração de dados, resultando nas classes da exposição de risco, conforme o RAC.

Os elementos, os valores de sete amostras das simulações e as pré-classes da exposição de risco resultantes da equação (5), formam a matriz de dados que serviram como entrada para os algoritmos de mineração de dados e, resultaram as informações

⁷ <http://www.pc-progress.com/en/Default.aspx>

para o quadro comparativo da acurácia na aplicação das técnicas e estão ilustrados na Tabela 4, a seguir.

Tabela 4. Amostras por contaminantes e a pré-classificação da exposição de risco.

Elementos	Amostras ($\mu\text{mol/l}$)							Pré-Classes
	1	2	3	4	5	6	7	
Ca	98.0	65.0	35.0	55.0	75.0	75.0	9.8	0
Na	450.0	450.0	450.0	450.0	450.0	450.0	450.0	0
Mg	5.0	5.0	5.0	5.0	5.0	5.0	5.0	0
Zn	50.0	24.0	22.0	32.0	36.0	36.0	0.0	3
Cd	0.8	0.38	0.42	0.63	0.35	0.35	0.0	1
Pb	2.5	1.2	1.1	1.6	9.0	9.0	0.0	2
K	120.0	120.0	120.0	120.0	120.0	120.0	120.0	0
Al	220.0	220.0	220.0	220.0	220.0	220.0	220.0	3

As sete amostras foram referenciadas de acordo com cada elemento químico percolado no solo, sem levar em consideração as suas especificações, como citado anteriormente. A pré-classificação evidenciou os metais pesados em razão das especificações do RAC que mede a influência do carbono sobre os metais, sinalizando risco. Por outro lado, os elementos Ca, Na, Mg e K são encontrados naturalmente na água e no solo, não caracterizando perigo ambiental nas condições avaliadas.

Os *scripts* utilizados para execução de cada algoritmo na ferramenta *Weka*, estão ilustrados na Figura 2.

LibSVM	weka.classifiers.functions.LibSVM -S 0 -K 2 -D 3 -G 0.0 -R 0.0 -N 0.5 -M 40.0 -C 1.0 -E 0.001 -P 0.1 -model "C:\Program Files (x86)\Weka-3-7" -seed 1
RBFNetwork	weka.classifiers.functions.RBFNetwork -B 2 -S 1 -R 1.0E-8 -M -1 -W 0.1
MultiClassClassifier	weka.classifiers.meta.MultiClassClassifier -M 0 -R 2.0 -S 1 -W weka.classifiers.functions.Logistic -- -R 1.0E-8 -M -1
MultilayerPerceptron	weka.classifiers.functions.MultilayerPerceptron -L 0.3 -M 0.2 -N 500 -V 0 -S 0 -E 20 -H a
BayesNet	weka.classifiers.bayes.BayesNet -D -Q weka.classifiers.bayes.net.search.local.K2 -- -P 1 -S BAYES -E weka.classifiers.bayes.net.estimate.SimpleEstimator -- -A 0.5

Figura 2. Script's de execução dos algoritmos no Weka 3.7.5.

Os resultados mais significativos ilustrados na Tabela 5 foram alcançados pelos algoritmos de classificação LibSVM e MultiClassClassifier que obtiveram precisão de 100% e 87,5% e índice estatístico *kappa* (k) de 1,0 e 0,8, respectivamente.

O método e algoritmo da LibSVM baseado em máquinas de vetores para multiclassificação pela abordagem um-contra-todos, obteve plena classificação na concordância entre as observações e os valores esperados, considerando um intervalo de confiança de 95%, não obteve erro médio absoluto e nem erro quadrático médio, indicando ser a técnica mais apropriada para esse tipo de análise.

Tabela 5. Comparação dos resultados da multiclassificação e os indicadores de classificação e erro apurados.

Elementos	Dados de treinamento	Comparativo da classificação por técnica				
		LibSVM	MultiClass Classifier	RBF Network	BayseNet	Multilayer Perceptron
Ca	0	0	0	3	0	0
Na	0	0	0	0	0	0
Mg	0	0	0	0	0	0
Zn	3	3	3	0	0	0
Cd	1	1	1	1	1	0
Pb	2	2	2	2	2	0
K	0	0	0	0	0	0
Al	3	3	0	3	0	0
Classificação correta		100%	87,5%	75,0%	75,0%	50,0%
Erro médio absoluto		0,0	0,2916	0,0572	0,1300	0,2442
Erro quadrático médio		0,0000	0,3651	0,1714	0,2630	0,3418
Índice <i>kappa</i> (k)		1,00	0,80	0,62	0,58	0,00

Pelos resultados pode-se verificar que, para Ca, Na, Mg e K as ocorrências foram classificadas como 0 ou nenhum risco, para o Cd houve ocorrência classificada como 1 ou de baixo risco, para Pb a ocorrência obteve classificação como 2 ou risco médio e para Zn e Al as ocorrências foram classificadas como 3 ou alto risco. A classificação 4 ou de risco muito alto não obteve nenhuma ocorrência no estudo.

4. Conclusão

Existem vários modelos de exposição que podem ser adotados para estimar a exposição de indivíduos às substâncias nocivas. As saídas desses modelos podem gerar valores médios ou por mediana, parâmetros de distribuição ou de probabilidades, incorrendo em modelos de exposição que variam amplamente em complexidade, abordagem, diversidades de variáveis de entradas e saídas. Porém, na maioria desses modelos emprega-se a técnica estatística da regressão para representação de distribuições probabilísticas.

Esse estudo decorreu da proposição de comparar técnicas de classificação em mineração de dados num processo de avaliação da exposição por contaminantes, empregando algoritmos para multiclassificação de riscos ambientais.

De acordo com a metodologia indicada, foram considerados dados existentes na literatura, citados em diversos trabalhos, para execução dos algoritmos selecionados, comparação de acurácia e validação dos resultados.

Os resultados obtidos sinalizaram uma forte correlação entre os resultados esperados e os resultados alcançados, plena classificação correta e sem margem de erro para a técnica e algoritmo da LibSVM, com dados simulados das amostras de exposição de contaminantes em cinco classes de riscos consideradas.

Para etapas futuras pretende-se desenvolver metodologia própria para validar com dados reais de uma contaminação importante no solo e aplicar outras técnicas de classificação. Pretende-se generalizar a aplicação para outras atividades impactantes, inclusive como auxílio para prognosticar riscos futuros nos estudos de impactos ambientais e similares.

Referências

- Apeelo, CAJ, Postma D. (2005). "Geochemistry, groundwater and pollution". Leiden: Balkema.
- CY Zhao, Zhang HX, XY Zhang, Liu MC, ZD Hu, BT Fan. (2006). "Aplicação da Máquina de Vetor de Suporte (SVM) para à actividade toxica predição de diferentes conjuntos de dados toxicologia", Volume 217, Issues 2-3, páginas 105-119
- EPA - Environmental Protection Agency (1992). "Final guidelines for exposure assessment". Federal Register, 104: 22888–22938. Washington, DC, United States Environmental Protection Agency.
- EPA - Environmental Protection Agency (1996). "Soil Screening Guidance: Technical Background Document". EPA/540/R95/128. Office of Solid Waste and Emergency Response Washington, DC 2046.EPA/540/R95/128.
- EPA - Environmental Protection Agency (2006). "CatReg Software User Manual". EPA/600/R-04/006. National Center for Environmental Assessment Office of Research and Development, Research Triangle Park, NC 27711.
- Hall Mark, Eibe Frank, Geoffrey Holmes, Bernhard Pfahringer, Peter Reutemann, Ian H. Witten (2009); "The WEKA Data Mining Software: An Update"; SIGKDD Explorations, Volume 11, Issue 1.
- Hertzberg, R.C. and M.L. Dourson. (1993). "Using categorical regression instead of a NOAEL to characterize a toxicologist's judgment in noncancer risk assessment". In: Proceedings, Second International Symposium on Uncertainty Modeling and Analysis, College Park, MD, B.M. Ayyub, ed. IEEE Computer Society Press, Los Alamitos, CA.: 254-261.
- Iowa State University - Department of Computer Science, Ames, IA 50011-1040 USA. Disponível em <http://www.cs.iastate.edu/~yasser/wlsvm/>. Acesso em: 27 maio 2013.
- Jacques D. A., J.; Simunek b, D. Mallants a and M. Th. van Genuchten. (2002). "Multicomponent transport model for variably-saturated porous media: application to the transport of heavy metals in soils". CA, USA.
- Jacques, D., and J. Simunek, (2005). "User Manual of the Multicomponent Variably-Saturated Flow and Transport Model HP1, Description, Verification and Examples, Version 1.0", SCK•CENBLG-998, Waste and Disposal, SCK•CEN, Mol, Belgium, 79 pp.
- Jacques D., J. Simunek, D. Mallants, and M. Th. van Genuchten. (2006). "Operator-splitting errors in coupled reactive transport codes for transient variably saturated flow and contaminant transport in layered soil profiles", *J. Contam. Hydrology*, 88, 197-218.

- Jacques, D., Simunek, J., Mallants, D., van Genuchten, M.Th. (2008). “Modelling coupled water flow, solute transport and geochemical reactions affecting heavy metal migration in a podzol soil”. *Geoderma*, 145, 449 – 461.
- McCullagh, P. and Nelder, J. A. (1989). “Generalized Linear Models, volume 37 of Monographs on Statistics and Applied Probability”. Chapman and Hall, London, 2 edition.
- Perin, G. L, Craboledda, L., Lucchese, M., Cirillo, L., Dotta, M. L., Zanette e A. A. Orio. (1985). “Heavy metal speciation in the sediments of Northern Adriatic Sea-a new approach for environmental toxicity determination”. In: TD Lekkas (Ed.). *Heavy metal in the environmental*, vol 2, p. 456-456.
- Radcliffe, D., and J. Simunek. (2010). “Soil Physics with HYDRUS: Modeling and Applications”, CRC Press, Taylor & Francis Group, ISBN-10: 142007380X, ISBN-13: 9781420073805, pp. 373.
- Spadotto C.A. (2002). “Classificação de Impacto Ambiental”. Comitê de Meio Ambiente, Sociedade Brasileira da Ciência das Plantas Daninhas. Disponível em: <http://www.cnpma.embrapa.br>. Acessado em 23 março 2013.
- Simunek, J. and A. J. Valocchi, (2002). “Geochemical Transport. In: Methods of Soil Analysis, Part 1, Physical Methods”, Chapter 6.9, Eds. J. H. Dane and G. C. Topp, Third edition, SSSA, Madison, WI, 1511-1536.
- Simunek, J.(2005). “Chapter 78: Models of Water Flow and Solute Transport in the Unsaturated Zone”, In: M. G. Anderson and J. J. McDonnell (Editors), *Encyclopedia of Hydrological Sciences*, John Wiley & Sons, Ltd., Chichester, England, 1171-1180.
- Simunek, J., M. Sejna, H. Saito, M. Sakai, and M. Th. van Genuchten. (2012). “The HYDRUS-1D Software Package for Simulating the Movement of Water, Heat, and Multiple Solutes in Variably Saturated Media”, Version 4.15, HYDRUS Software Series 3, Department of Environmental Sciences, University of California Riverside, Riverside, California, USA, pp. 338.
- Suentjens, P. (2000). “Reactive Solute Transport in Heterogeneous Porous Medium: Cadmium Leaching in Acid Sandy Soils”. University of Antwerp, Belgium, Publication Number: AAI3015182; ISBN: 9780493259338; Source: Dissertation Abstracts International, Volume: 62-05, Section: B, page: 2133.; p. 236.
- |
- Wissmeier,L.; Barry D.A. (2008). “Reactive transport in unsaturated soil: Comprehensive modelling of the dynamic spatial and temporal mass balance of water and chemical components”. *Avanços em Recursos Hídricos*, Volume 31, Ed. 5, maio de 2008, Pag. 858-875.
- WHO – World Health Organization /IPCS – International Programme on Chemical Safety. (2010). “Risk Assessment Toolkit: chemical hazards”. USA.