WATER RESEARCH 45 (2011) 5535–5544

Available at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/watres

Prediction of contamination potential of groundwater arsenic in Cambodia, Laos, and Thailand using artificial neural network

Kyung Hwa Cho [a,1], Suthipong Sthiannopkao [b,1], Yakov A. Pachepsky [c], Kyoung-Woong Kim [d], Joon Ha Kim [d,e,*]

[a] Department of Civil and Environmental Engineering, University of Massachusetts, Amherst, 130 Natural Resources Road, Amherst, MA 01003, USA

[b] Department of Environmental and Occupational Health, National Cheng Kung University, Tainan City, Taiwan, ROC

[c] USDA-ARS, Environmental Microbial & Food Safety Laboratory, 10300 Baltimore Ave., Beltsville, MD 20705, USA

[d] School of Environmental Science and Engineering, Gwangju Institute of Science and Technology (GIST), 261 Cheomdan-gwagiro, Buk-gu, Gwangju 500-712, Republic of Korea

[e] Sustainable Water Resource Technology Center (SWRTC) at GIST, 261 Cheomdan-gwagiro, Buk-gu, Gwangju 500-712, Republic of Korea

ARTICLE INFO

Article history:
Received 10 February 2011
Received in revised form
29 May 2011
Accepted 8 August 2011
Available online 27 August 2011

Keywords:
Multiple linear regression
Principal component regression
Artificial neural network
Principal component-artificial neural network

ABSTRACT

The arsenic (As) contamination of groundwater has increasingly been recognized as a major global issue of concern. As groundwater resources are one of most important freshwater sources for water supplies in Southeast Asian countries, it is important to investigate the spatial distribution of As contamination and evaluate the health risk of As for these countries. The detection of As contamination in groundwater resources, however, can create a substantial labor and cost burden for Southeast Asian countries. Therefore, modeling approaches for As concentration using conventional on-site measurement data can be an alternative to quantify the As contamination. The objective of this study is to evaluate the predictive performance of four different models; specifically, multiple linear regression (MLR), principal component regression (PCR), artificial neural network (ANN), and the combination of principal components and an artificial neural network (PC-ANN) in the prediction of As concentration, and to provide assessment tools for Southeast Asian countries including Cambodia, Laos, and Thailand. The modeling results show that the prediction accuracy of PC-ANN (Nash–Sutcliffe model efficiency coefficients: 0.98 (traning step) and 0.71 (validation step)) is superior among the four different models. This finding can be explained by the fact that the PC-ANN not only solves the problem of collinearity of input variables, but also reflects the presence of high variability in observed As concentrations. We expect that the model developed in this work can be used to predict As concentrations using conventional water quality data obtained from on-site measurements, and can further provide reliable and predictive information for public health management policies.

© 2011 Elsevier Ltd. All rights reserved.

* *Corresponding author*. School of Environmental Science and Engineering, Gwangju Institute of Science and Technology (GIST), 261 Cheomdan-gwagiro, Buk-gu, Gwangju 500-712, Republic of Korea. Tel.: +82 62 970 3277; fax: +82 62 970 243.
E-mail address: joonkim@gist.ac.kr (J.H. Kim).

[1] Cho and Sthiannopkao contributed equally to this work and are listed in alphabetical order.

0043-1354/$ – see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.watres.2011.08.010