

Thursday, 16 November 2006 - 9:20 AM  
497c

## A Method for Modeling Chemical Multimedia Partitioning with Neural Networks and Classifiers

Izacar Martinez<sup>1</sup>, Gabriela Espinosa<sup>1</sup>, Robert Rallo<sup>2</sup>, Jordi Grifoll<sup>1</sup>, Yoram Cohen<sup>3</sup>, and **Francesc Giral**<sup>1</sup>. (1) Departament d'Enginyeria Química, Universitat Rovira i Virgili, Av. dels Països Catalans 26, Tarragona, 43007, Spain, (2) Universitat Rovira Virgili, Av. dels Països Catalans 26, Tarragona, 43007, Spain, (3) University of California, Los Angeles, Chemical Engineering Department, 5531 Boelter Hall, Los Angeles, CA 90095

Multimedia fate and transport models are typically utilized for screening analysis of the multimedia distribution of chemicals in specific or generic geographical regions. Input to multimedia models consist of physicochemical properties for chemicals under scrutiny. Multimedia models require a variety of physicochemical properties. However, there is a large number of chemicals of concern for which the required physicochemical data are either uncertain or unknown. Accordingly, a methodology is proposed, based on artificial neural networks and classifiers, to estimate transport and fate of organic chemicals when data for the physicochemical properties of chemicals of concern are inaccurate, incomplete or unavailable. In the present approach, several artificial neural network architectures and classifiers have been trained, using data generated from a multimedia fate and transport model, to estimate concentrations in different media from physicochemical and molecular descriptors of the chemicals of interest.

The core of the proposed methodology is to select relevant multimedia chemical distribution patterns that represent regions of the input-output multidimensional space of a given multimedia model, for training networks that can subsequently serve as multivariate function approximators. The relevant patterns to train neural networks were selected via a two step process. In the first step, the input-target space was analyzed with several feature selection algorithms, e.g., filters or wrappers, to reduce the required number of input variables. In the second step, the selected chemicals (represented by input vectors) were clustered with a Self-Organizing Map and placed in either training or test data sets as their corresponding clusters suggest.

Training and testing of different artificial neural networks (backpropagation and RBFs) and classifiers (fuzzy ARTMAP and Support Vector Machines) were performed with the corresponding data sets, followed by comparison of models with respect to their performance. In the present study, multimedia simulations were carried out via a standard multimedia model for a given geographical area and meteorological conditions. In the first stage, the input to the neural networks were sets of physicochemical properties for 490 selected chemicals (332 for training and 158 for testing). In a second stage, the physicochemical properties were replaced by molecular descriptors as input for the neural networks. For the 332 training chemicals selected, seven physicochemical properties and the corresponding multimedia model output concentrations contained sufficient information for training the neural networks and classifiers. That selection was confirmed by an evaluation of the multivariate correlation of the data using the K correlation index. The seven relevant variables were those related to chemical partitioning coefficients and degradation rate parameters in each media. The selected backpropagation final architecture was a 7-20-5 network (7 input variables, one hidden layer with 20 neurons and the concentration in 5 media – air, water, sediments, soil and vegetation - as output) which was able to predict the 5 output concentrations with an mean absolute error of 0.026 in terms of scaled/normalized concentrations. Equally performing models were obtained with RBFs networks, as well as with the classifiers fuzzy ARTMAP and Support Vector Machines. Training models with physicochemical properties as the chemical-specific input variables revealed that the artificial neural network and classifiers based model can be used to estimate chemical concentrations provided that the training data set contained representative patterns. Partitioning concentrations predicted for the above mentioned 5 media with neural networks and classifiers trained with only molecular descriptors of the chemicals of interest will also be presented and discussed.

See more of [#497 - Environmental Fate and Transport Processes I \(09002\)](#)

See more of [Environmental Division](#)

See more of [The 2006 Annual Meeting](#)