Prediction of Henry’s law constant of a group of alkanes using computational descriptors

Zahra Honarmand*,

Abstract In this project, the quantitative structure-Henry’s Law constant relationship of a group of alkanes were studied only by using the topological indices. Henry’s Law constant values of 72 alkanes were used as data set. After calculation of the descriptors, the data set was divided into two groups of training (58 molecules) and test (14 molecules) sets and the multiple linear regression technique with stepwise method was used for modeling Henry’s constant values. Using the method, a model with four topological descriptors (CENT, ZM2V, TL1, S1K) was obtained. The statistics for the model were: R2=0.894, SE=0.561 and F=101.6. Then, to develop a model with higher quality, the ed descriptors and an artificial neural network were employed. In nonlinear modeling, 14 alkanes of the training set was used as the validation set. To develop the model, a multi-layer perceptron neural network and Levenberg–Marquardt and scaled conjugate garadient learning algorithms was applied. The results showed that the best non-linear model is obtained using a 4-14-1 artificial neural network trained by the scaled conjugate garadient. The statistics of the ed artificial neural network model were R2=0.953, MSE=0.541 and F=97.2 for the training set, R2=0.941, MSE=0.704 and F=17.4 for validation set, and R2=0.988, MSE=0.305 and F=88.4 for test set. Based on the results of this research, artificial neural network was superior over multiple linear regression in modeling and predicting Henry’s Law constant values of alkanes.

Keywords : Keywords: Henry’s Law constant, quantitative structure-property relationship, topological descriptor, artificial neural network