Analysis of factors contributing to uncertainty in BCF predictions for chemical substances

Riku Maeda
Kyushu Institute of Technology

The bioconcentration factor (BCF) is an indicator of the accumulation potential of chemical substances in organisms. A higher BCF value indicates a higher concentration potential in organisms, which in turn implies a greater impact on ecosystems and human health. Until now, BCF prediction models based on molecular structure have been evaluated solely based on prediction accuracy, with little systematic evaluation of the reliability of predictions or the identification of outliers. In this study, we aimed to improve the practicality of BCF prediction models by identifying the physicochemical factors associated with BCF prediction uncertainty. Specifically, we developed a regression model based on physicochemical descriptors calculated from SMILES, compared the predicted values with actual BCF values, and constructed a classifier for chemicals with significantly deviating predicted values (overestimated group and underestimated group). We then analyzed the contribution of features. The results showed that, in the underestimated group, topological descriptors (Kappa2, Kappa3), hydrophobicity (LogP), and the presence or absence of specific functional groups had a strong influence on prediction. The results of this study provide specific information on which types of chemical substances require careful interpretation of BCF prediction results and are expected to contribute to the streamlining of future environmental risk assessments.