

МОДЕЛИРОВАНИЕ СВЯЗИ МЕЖДУ СТРУКТУРОЙ И СВОЙСТВАМИ ОРГАНИЧЕСКИХ СОЕДИНЕНИЙ НА ОСНОВЕ ПОДБОРА ОПТИМАЛЬНЫХ АТОМНЫХ ПАРАМЕТРОВ

Яковенко Ю.Ю., Скворцова М.И.

Московская государственная академия тонкой химической технологии им. М.В. Ломоносова, Кафедра высшей и прикладной математики, Россия, 119571, г. Москва, пр. Вернадского, 86, Тел.: (495) 936-88-71, E-mail: yakovenko88.88@mail.ru, skvorivan@mail.ru

Построение и исследование математических моделей связи между структурой и свойствами органических соединений является одной из основных задач математической химии. Найденные закономерности позволяют прогнозировать свойства химических соединений непосредственно по их структуре, минуя эксперимент, и могут быть использованы для планирования целенаправленного поиска соединений с заданными свойствами.

В работе предлагается новый общий метод построения моделей связи «структура-свойство» органических соединений и проводится его исследование. В качестве исходных данных используется некоторая выборка соединений, представленных структурными формулами, для которых известны численные значения y изучаемого свойства. Согласно предложенному методу первоначально проводится некоторая классификация атомов, входящих в структуры изучаемых соединений. Например, атомы могут быть классифицированы по химическим символам и распределениям типов связей. Атомам k -ого класса приписывается некоторый неопределённый вес. Далее предполагается, что модель связи «структура-свойство» имеет следующий вид:

$$y = \frac{\sum x^{(i)} x^{(j)}}{n} + x_0, \quad (1)$$

где $x^{(i)}$, $x^{(j)}$ – веса i -ого и j -ого атомов в молекуле (в соответствии с их классификацией); суммирование распространено на все связи (i, j) , x_0 – некоторая константа, n – число атомов в молекуле. На следующем этапе построения модели эти веса подбираются оптимальным образом так, чтобы модель (1) была бы как можно более точной на исходной выборке соединений. Следует отметить, что: а) классификацию атомов можно проводить с разной степенью детализации, что позволяет увеличить точность аппроксимации; б) в формуле (1) можно использовать другие виды выражений, зависящих от атомных весов.

На основе предложенного метода построены модели связи «структура-свойство» для различных физико-химических свойств и классов соединений; при этом рассматривались различные модификации формулы (1). Эти модели затем использованы для прогнозирования свойств других соединений рассматриваемых классов, не включённых в исходные выборки. Кроме того, проведено сравнение эффективности предложенного метода моделирования связи «структура-свойство» и ряда других методов того же назначения для одних и тех же данных.