**Исследование корреляции между замещёнными ароматическими углеводородами и их токсичностью для водорослей**

***Цю Ваньцяо***

*Студент (магистр)*

*Московский государственный университет имени М.В.Ломоносова*,

*Институт русского языка и культуры, Москва, Россия*

*E-mail: wq710017@163.com*

Замещённые ароматические углеводороды являются важным сырьем в химическом производстве; в то же время они высоко токсичны и могут нанести серьезный ущерб окружающей среде. Но традиционные лабораторные методы оценки биотоксичности требуют больших затрат времени и имеют высокую стоимость.

Важными индикаторами в области токсикологических исследований служат водоросли. Они характеризуются высокой чувствительностью к окружающей среде, широким распространением, малым размером особей и быстрым размножением.

Удобный современный метод быстрой оценки токсичности – это построение модели количественного корреляции “структура-активность” (QSAR) загрязнителя для организма-индикатора [1][2]. Цель данного исследования – построение подобной модели токсичности для замещенных ароматических углеводородов.

Ход исследования.

1. Сбор данных о токсичности замещенных ароматических углеводородов при различных показателях токсичности.
2. Фильтрация данных, выбор организма-индикатора.
3. Выбор подходящего молекулярного дескриптора.
4. Построение модели. Сравнение двух методов моделирования: метода географической взвешенной регрессии (ГВР) и традиционного метода наименьших квадратов (МНК).
5. Прогнозирование токсичности на основе установленных моделей.
6. Оценка моделей.

На ранней стадии исследования были собраны данные о токсичности 186 замещённых ароматических углеводородов для различных видов зелёных водорослей (*Chlorella vulgaris, Tetradesmus obliquus, Selenastrum bibraianum, Dunaliella salina* и др.) в разных концах, а также данные по 47 химическим дескрипторам. После очистки и скрининга данных для проведения исследования были отобраны организм *Chlorella vulgaris,* 67 токсинов, 9 стандартных экспоненциальных дескрипторов токсичности. Токсичность измерялась по трём точкам в течение 96 часов [3].

Основным методом моделирования, используемым в этом исследовании, является географическая взвешенная регрессия (ГВР). Этот метод может учитывать несколько переменных влияния одновременно, и непосредственно и эффективно описывать пространственные неровные отношения между переменными [4]. Процесс анализа метода ГВР в этом исследовании состоит из следующих этапов:

1. Разделение данных на обучающую выборку и тестовую выборку.
2. Отбор дескрипторов и сортировка по значимости.
3. Выбор частичного дескриптора в качестве координаты для расчета веса между различными соединениями. Вес соединения рассчитывается с помощью двойной квадратной функции, а коэффициент регрессии модели оценивается с помощью матрицы пространственных весов.
4. Оценка коэффициента регрессии модели взвешенным методом наименьших квадратов.
5. Создание модели ГВР и прогнозирование токсичности.

В результате на основе метода ГВР для хлореллы (*Chlorella vulgaris)* была создана модель QSAR для 67 замещенных ароматических соединений (токсичность за 96 часов по 3 точкам: IC50, IC20 и LOEC). Коэффициент детерминации R2 для обучающей и тестовой выборки составил 0,79 и 0,79, 0,75 и 0,77, 0,74 и 0,77 соответственно. По сравнению со значениями R2 и AICc, полученными традиционным методом наименьших квадратов (МНК), результаты показали, что метод географической взвешенной регрессии обладает лучшей способностью к подгонке и может обеспечить надежные прогнозы токсичности замещенных ароматических углеводородов для хлореллы в нескольких конечных точках.

В настоящее время в области наук об окружающей среде метод ГВР в основном применяется для изучения взаимосвязи между свойствами окружающей среды и факторами окружающей среды на макроуровне [5], но по-прежнему не используется в анализе токсичности соединений на микроуровне. Уникальные возможности обработки пространственных данных открывают возможности для их применения в исследовательских целях. Модель QSAR, построенная в этом исследовании, может эффективно прогнозировать токсичность замещённых ароматических соединений для хлореллы и служить хорошим примером для анализа воздействия органического вещества на окружающую среду и применения алгоритма ГВР.

**Литература**

1. Liu Tingting.Research on the toxic structure-activity relationship between phenol and aniline organic matter on algae [D].Tianjin:Tianjin University of Science and Technology,2019.
2. Lu Guanghua and Geng Liang.Research on the quantitative relationship between the toxicity of substituted aromatics to algae and structural parameters [J].Environmental scientific research,2004(06):73-75.
3. Fangyou Yan,Tingting Liu,Qingzhu Jia,Qiang Wang.Multiple toxicity endpoint-structure relationships for substituted phenМНК and anilines[J].Science of the Total Environment,2019,663.
4. FOTHERINGHAM A S,BRUNSDON C,CHARLTON M.Geographically Weighted Regression:the analysis of spatially varying relationships [M].Chichester:John Wiley and Sons Inc,2002:42-46.
5. Qu Mingkai,Li Weidong,Zhang Chuanrong,Huang Biao.Geographic weighted regression and its application prospects in soil and environmental science [J].soil,2014,46(01):15-22.