

ИНФОРМАЦИОННЫЕ СИСТЕМЫ И ТЕХНОЛОГИИ В ЭКОЛОГИИ

УДК 004.8.032.26: 556.531

О. С. Хилько, С. П. Кундас

Международный государственный экологический университет имени А.Д. Сахарова, г. Мисск

МЕТОДИКА ПОДГОТОВКИ ОБУЧАЮЩЕЙ ВЫБОРКИ ИСКУССТВЕННОЙ НЕЙРОННОЙ СЕТИ ДЛЯ ПРОГНОЗИРОВАНИЯ МИГРАЦИИ РАДИОНУКЛИДОВ НА ПОВЕРХНОСТИ ПОЧВЫ

Применение искусственных нейронных сетей предполагает наличие достаточной обучающей выборки, состоящей из предварительно обработанных данных, представленных в численном виде. С целью подготовки данных для нейросетевого прогнозирования миграции радионуклидов на поверхности почвы была разработана методика, включающая следующие способы и методы: способ расчета параметра наклона поверхности; метод замены символьных значений направления ветра их числовыми характеристиками с проведением предварительных расчетов; способ статистической обработки данных климатических факторов по сезонам с динамическим заданием границ; метод выбора пар точек, параметры которых включаются в обучающую выборку.

- **Ключевые слова:** обучение искусственных нейронных сетей, подготовка обучающих данных, латеральная миграция радионуклидов.

Прогнозирование латеральной миграции радионуклидов является актуальной задачей экологического мониторинга, так как позволяет на основе результатов математического и компьютерного моделирования делать выводы о состоянии окружающей среды после масштабных радиоактивных выбросов в результате техногенных катастроф, о возможности использования природных ресурсов территорий, подвергнувшихся загрязнению [1].

Одним из современных инструментов прогнозирования миграции радионуклидов в горизонтальном направлении являются искусственные нейронные сети (ИНС). Нейросетевой подход позволяет учесть зависимости, не поддающиеся формализованному математическому описанию [2], а также проводить анализ влияния отдельных факторов на значение целевой функции [2], в частности на перераспределение загрязнения на поверхности почвы [1].

Использование нейросетевого метода прогнозирования требует наличия достаточной обучающей выборки. Однако данные для обучения не всегда представлены в требуемом формате и зачастую представляют собой набор символьных характеристик, которые не могут восприниматься ИНС. Перед использованием нейронной сетью данные должны быть предварительно обработаны, представлены в цифровом виде (во многих случаях характеристики растительного покрова, ландшафта, рельефа, типа почв, направления ветра описываются словесно) и нормированы.

Для приведения экспериментальных данных к виду, удобному для восприятия искусственной нейронной сетью, предложена методика, применение которой позволяет: ускорить процесс решения задачи прогнозирования миграции радионуклидов на поверхности почвы путем сокращения времени, затрачиваемого на предварительную обработку данных; уменьшить количество входных параметров, что увеличивает точность прогноза; исключить из обучающей выборки наборы данных, относящиеся к точкам, между которыми не происходит процесс переноса радионуклидов.

Структура входного и выходного слоев искусственной нейронной сети для прогнозирования миграции радионуклидов на поверхности почвы в микромасштабе

При проведении моделирования распространения радионуклидов на поверхности почвы на микроуровне особенно важно учитывать рельефные особенности местности [3–7]. Содержание

^{137}Cs в пахотном горизонте различных элементов рельефа склоновых земель различается в 1,5–3 раза [5–7]. Поэтому при использовании ИНС для прогнозирования миграции радионуклидов предложено учитывать взаимное влияние значений радиоактивности близко расположенных точек, вследствие чего входной слой ИНС включает пространственно зависимые данные двух соседних точек, одна из которых является базовой (точка, на которую оказывается влияние), другая – влияющей.

Разработанная авторами интегрированная топология ИНС для прогнозирования миграции радионуклидов ^{137}Cs и ^{90}Sr на поверхности почвы с шагом в один год включает значения следующих факторов в качестве входных нейронов [8, 9]:

- тип радионуклида ($^{137}\text{Cs} - 1, ^{90}\text{Sr} - 0$);
- значения удельной (Бк/кг) или поверхностной (Бк/м²) активности $C_{i\text{ start}}$ и $C_{\text{ start}}$ в соседних точках на начало периода;
- характеристики углов наклона поверхности в соседних точках;
- относительные или абсолютные высоты соседних точек (м);
- пространственно зависимые характеристики для каждой из точек в паре (характеристики растительного покрова: тип растительности в виде коэффициента балльности, удельная активность радионуклида в растениях (Бк/кг), толщина подстилки (см); почвенные характеристики: процентное содержание песка, ила, глины, гумуса (%), кислотность (единицы рН), емкость катионного обмена (мг-экв/100 г или смоль+/кг);
- общее количество осадков (мм), средняя температура (в градусах), средняя влажность воздуха (%) для каждого сезона, марта и апреля; суммы положительных температур за период с начала снеготаяния до полного исчезновения снежно-ледяного покрова (в градусах);
- количество затоплений (если рассматриваемый участок находится вблизи водотока) за период с начала снеготаяния;
- коэффициент увлажнения по Г. Н. Высоцкому;
- характеристики хозяйственной деятельности человека на исследуемой территории (коэффициент противоэрозионных мероприятий; коэффициент интенсивности возделывания);
- характеристики двух преобладающих направлений ветра и их средние скорости.

В качестве одного выходного нейрона в набор обучающих данных включается удельная (Бк/кг) или поверхностная (Бк/м²) активность радионуклида в базовой точке на конец моделируемого периода.

Методика подготовки данных обучающей выборки для искусственной нейронной сети

Среди выделенных факторов, на основе которых формируется обучающая выборка, имеются как численные, так и символьные значения. Факторы, значения которых представлены в виде физических величин, без труда выражаются численно (активности и тип радионуклида, высоты точек, количество затоплений, приведенные коэффициенты, указанные пространственно зависимые характеристики) и могут быть использованы искусственной нейронной сетью «как есть» после нормирования.

Следует отметить, что такие факторы, как уровень наклона поверхности, количество осадков, температура, влажность воздуха, хоть и представлены в численном виде, однако требуют предварительной обработки. В свою очередь, направления ветра обычно выражены в румбах в символьном виде, что требует замены указанных значений их числовыми характеристиками с проведением предварительных расчетов для дальнейшего использования искусственной нейронной сетью.

В ходе анализа пространственных данных в микромасштабе может быть выделено несколько точек, оказывающих влияние на значение активности радионуклида в рассматриваемой (базовой) точке на поверхности почвы. Для базовой точки предложено выбирать единственную пару, которая будет включена в обучающую выборку.

Определение характеристики наклона поверхности

Для характеристики наклона поверхности предложено использовать функцию синуса угла наклона. Так как движение влаги по склонам осуществляется под действием гравитационных сил, то вещество, переносимое с поверхностными водами, будет перемещаться от выше- к ниже расположенным точкам. Разность высот между вышележащей (оказывающей влияние) и ниже лежащей (подверженной влиянию) точками будет положительной. С другой стороны, лежащая ниже точка не будет оказывать влияние на активность радионуклида в точке, расположенной выше. Разность высот между ниже лежащей и вышележащей точками в рассматриваемом случае будет отрицательной.

Характеристика угла наклона ($\alpha < 90^\circ$) может быть рассчитана по разности высот Δh (противолежащий катет) и расстоянию между точками r (прилежащий катет). Для вычисления может быть выбрана любая из тригонометрических функций. Однако функции тангенса и котангенса могут принимать значения в промежутке $(-\infty; \infty)$, что затрудняет их использование на этапе подготовки данных для применения в искусственных нейронных сетях (при нормировании данных должны быть заданы пределы, в которых находится величина, а предел бесконечности в численном виде выразить невозможно). Функция синус является нечетной, т. е. ее значение положительно при $\alpha > 0^\circ$ и отрицательно при $\alpha < 0^\circ$. Учитывая тот факт, что при $\Delta h > 0$ параметры точки оказывают влияние на активность радионуклида в соседней, а при $\Delta h < 0$ – такое влияние отсутствует, то функция синус подходит для характеристики угла наклона, в отличие от функции косинуса угла, при использовании которой вследствие ее четности невозможно учесть указанную особенность. Таким образом, синус угла наклона поверхности в точке может быть рассчитан как:

$$\sin \alpha_i = \frac{\Delta h_i}{\sqrt{r_i^2 + \Delta h_i^2}} \quad (1)$$

где r_i – расстояние между двумя соседними точками; Δh_i – разность высот между оказывающей влияние (X_1) и подверженной влиянию точкой (X_2).

Расстояния между местами отбора проб предложено рассчитывать по координатам по методике независимой лаборатории ГИС путем вычисления расстояния и начального азимута между двумя точками на земной поверхности [10]. Корректность расчетов при применении указанной методики была проверена путем вычислений расстояний между 10 парами точек на картах Google.

Рис. 1 позволяет наглядно представить способ расчета угла наклона по формуле (1).

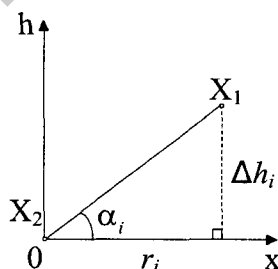


Рис. 1. Угол наклона поверхности в точке

Следует также отметить, что на перемещение вещества с водой не влияет ориентация склона, поэтому точки, равноудаленные от вершины с одинаковыми углами наклона, будут испытывать одинаковое воздействие с ее стороны. Функция $\sin(\alpha)$ имеет одинаковый знак в I и II четвертях единичной окружности, поэтому ее использование позволяет также учесть данную особенность перемещения вещества. Функция $\cos(\alpha)$ имеет различные знаки в I и II четвертях, вследствие чего ее использование не позволяет учесть указанную особенность.

Определение границ сезонов для вычисления статистических значений климатических факторов

Так как ежемесячные измерения активностей как на поверхности почвы, так и в растениях проводятся достаточно редко, традиционным является шаг измерений в 1 год. Значения климатических факторов, в свою очередь, измеряются гораздо чаще. К примеру, гидрометеоцентр Республики Беларусь публикует данные о погоде с шагом в 3 часа. Поэтому

данные о температуре, количестве выпавших осадков, влажности воздуха и т. д. должны быть представлены в усредненном виде за некоторый период. Таким периодом является сезон, так как особенности миграции радионуклидов на поверхности почвы и вдоль почвенного профиля связаны с количеством и агрегатным состоянием выпавших осадков, процессами промерзания почвы, снеготаяния, инфильтрации, испарения и другими [1], которые имеют различные пики интенсивности именно по сезонам.

Для определения временных границ начала сезонов предложено использовать следующие значения пороговых температур для весны, лета, осени и зимы соответственно: $0\text{ }^{\circ}\text{C}$, нормальная температура июня ($T_{\text{норм июнь}}$), нормальная температура сентября ($T_{\text{норм сент}}$), $0\text{ }^{\circ}\text{C}$, с которыми сравнивается вычисленное значение средней фактической температуры за неделю (T). Таким образом, началом сезона считается начало той недели, среднее значение температуры которой впервые за рассматриваемый год удовлетворяет соответствующему неравенству: $T > 0\text{ }^{\circ}\text{C}$ (начало весны), $T > T_{\text{норм июнь}}$ (начало лета), $T < T_{\text{норм сент}}$ (начало осени), $T < 0\text{ }^{\circ}\text{C}$ (начало зимы).

Оценка влияния ветра

Для оценки влияния ветра на перенос радионуклидов на поверхности почвы предложено рассмотреть отклонение направления ветра от вектора, соединяющего две точки в пространстве.

В символьном виде направление ветра нецелесообразно использовать в качестве входного параметра ИНС, так как невозможно учесть особенности переноса вещества в привязке к конкретной местности и ориентации точек в пространстве. Важно учитывать не само направление ветра, а его смещение относительно вектора, соединяющего две соседние точки на поверхности почвы, что позволит определить, является ли склон наветренным и будет ли осуществляться перенос сорбированных почвенными частицами радионуклидов под воздействием силы ветра. Поэтому в первую очередь необходимо выразить направление ветра в градусах, далее привести к единой системе координат с расположением точек в пространстве, а затем выбрать математическую функцию, наилучшим образом подходящую для характеристики влияния направления ветра на перемещение радионуклидов на поверхности почвы и провести необходимые предварительные расчеты согласно способу вычисления выбранной функции.

Традиционно при переводе направления ветра в градусы за 0° принимается северное направление ветра, за 90° – восточное и т. д., т. е. при рассмотрении единичной окружности движение происходит против часовой стрелки. Рис. 2 дает наглядное представление перевода направления ветра в градусы общепринятым способом.

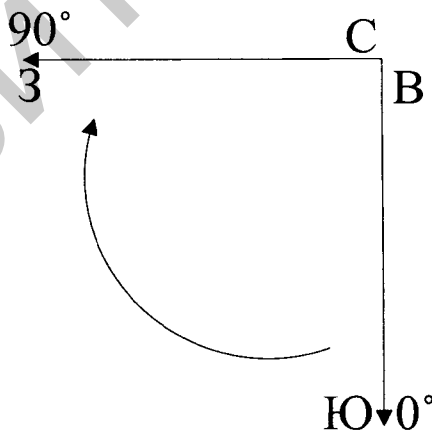


Рис.2. Традиционное представление направления ветра (откуда дует) в градусах

В свою очередь, при анализе географических координат и относительных смещений соседних точек было выявлено, что значение широты возрастает при движении от экватора к полюсам, а в восточном полушарии значение долготы возрастает при удалении от Гринвичского меридиана в восточном направлении. Рис. 3 иллюстрирует систему географических координат, где стрелкой указано направление движения по единичной окружности, OX – направление увеличения долготы, OY – направление увеличения широты.

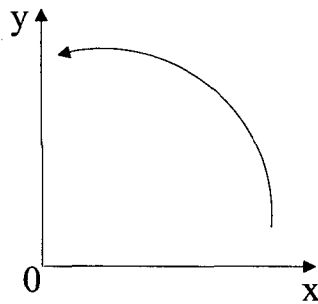


Рис. 3. Система географических координат в северо-восточном полушарии

Отсюда следует, что системы координат для направления ветра и географического положения точек различны и требуют приведения к единому виду. Так как расчет смещений точек друг относительно друга является более сложным, чем численное выражение направления ветра, то за основу принимается система географических координат, а интерпретация направления ветра в градусах переопределяется: за базовое (0°) принимается западное направление ветра, 90° соответствует южное направление, 180° – восточное, 270° – северное.

При анализе влияния ветра на перемещение вещества необходимо учитывать смещение его направления (β) по отношению к вектору, соединяющему соседние точки (γ), между которыми может осуществляться перенос радионуклидов. То есть необходимо рассматривать разность углов ($\beta - \gamma$), характеризующих отклонение направление ветра (β) и вектора расстояния (γ) от некоторой оси координат, за которую предложено принимать географическую параллель, направленную в сторону увеличения долготы (ось OX – см. рис. 3).

Для характеристики влияния направления ветра на перенос радионуклидов между двумя рядом расположенными точками предложено использовать функцию косинуса разности углов ($\beta - \gamma$). Как было показано ранее, функции тангенса и котангенса не удобны в использовании искусственной нейронной сетью при проведении нормирования.

При совпадении направлений указанных векторов влияние будет проявляться в наибольшей степени. А, как известно, $\cos 0^\circ = 1$, что является максимальным значением, которое может принимать эта функция.

Кроме того, функция косинус является четной и принимает положительные значения в первой и четвертой четвертях единичной окружности. При анализе влияния направления ветра было определено, что при отклонении вектора расстояния от направления ветра при прочих равных условиях не имеет значения, в какую сторону это отклонение происходит. Также, чем больше указанное отклонение, тем в меньшей степени направление ветра влияет на перенос вещества между рассматриваемыми соседними точками. Эта закономерность также отражается функцией косинус, так как ее значение уменьшается от 1 к 0 при увеличении модуля угла.

Если указанное отклонение будет более 90° , то перенос радионуклидов будет осуществляться в противоположном направлении, т. е. не от начальной точки вектора к конечной, а наоборот. Функция $\cos(\gamma)$ принимает отрицательные значения для углов $-180 < \gamma < -90^\circ$ и $90^\circ < \gamma < 180^\circ$, что подтверждает ее выбор для оценки влияния направления ветра на перенос радионуклидов на поверхности почвы.

Таким образом, функция косинус подходит для характеристики влияния направления ветра на перемещение адсорбированных почвенными частицами радионуклидов между точками в пространстве.

Косинус смещения направления ветра относительно вектора, соединяющего точки в пространстве, предложено вычислить по тригонометрической формуле косинуса разности:

$$\cos(\beta - \gamma) = \cos(\beta)\cos(\gamma) + \sin(\beta)\sin(\gamma), \quad (2)$$

где β – сдвиг направления ветра по отношению к западному, γ – сдвиг вектора, соединяющего точки в пространстве, относительно оси OX (направления западного ветра и оси OX совпадают – см. рис. 3).

Определение пар точек, включаемых в обучающую выборку

Как было показано выше, для базовой точки предложено выбирать единственную пару для включения в обучающую выборку, влияющая точка которой оказывает наибольшее среди остальных точек влияние на изменение активности радионуклида в базовой точке.

При прогнозировании миграции радионуклидов с поверхностным стоком выбор пары точек предложено осуществлять по максимальному значению критерия $\frac{C_{i\ start}}{C_{\ start}} \sin \alpha_i \frac{R - r_i}{R}$ $i = \{1, \dots, N\}$, где N – количество пар, включающих базовую точку; $C_{i\ start}$ и $C_{\ start}$ – концентрации радионуклида во влияющей и базовой точках соответственно; r_i – линейные расстояния между точками в каждой паре; α_i – угол наклона поверхности.

При рассмотрении влияния ветра на перенос адсорбированных почвенными частицами радионуклидов выбор пары точек предложено осуществлять по максимальному значению критерия $\cos(\beta_j - \gamma_i)$.

Алгоритм подготовки данных для обучающей выборки

Резюмируя все вышеизложенное, для практического применения описанной методики предложен алгоритм подготовки данных для обучающей выборки, предусматривающий выполнение следующих действий:

1. Определяются два преобладающих направления ветра за требуемый период, рассчитываются статистические значения климатических факторов по сезонам с динамическим заданием границ (весна – $T > 0$ °C, лето – $T > T_{\text{норм июнь}}$, осень – $T < T_{\text{норм септ}}$, зима – $T < 0$ °C, где T – средняя температура в °C за неделю).

2. Выбираются критерии оценки влияния (пороговые расстояния $R_{\text{вод}}$ и $R_{\text{ветр}}$).

3. Для двух преобладающих направлений ветра определяются $\sin \beta_j$ и $\cos \beta_j$, $j = \{1, 2\}$.

4. Для всех пар точек рассчитываются разность высот Δh , разности широты Δlat и долготы $\Delta long$, расстояние r (по координатам или относительным расстояниям).

5. Для каждой базовой точки выбираются несколько потенциально влияющих, для которых $r_i < R_{\text{вод}}$ или $r_i < R_{\text{ветр}}$, $i = \{1, \dots, N\}$, где N – количество пар, включающих базовую точку.

6. Рассчитываются $\sin \alpha_i$, $\sin \gamma_i$ и $\cos \gamma_i$, согласно формуле (2) вычисляются $\cos(\beta_j - \gamma_i)$,

$$\sin \alpha_i = \frac{\Delta h_i}{\sqrt{r_i^2 + \Delta h_i^2}}, \quad \sin \gamma_i = \frac{\Delta lat_i}{r_i} \quad \text{и} \quad \cos \gamma_i = \frac{\Delta long_i}{r_i},$$

$$\cos(\beta_j - \gamma_i) = \cos(\beta_j) \cos(\gamma_i) + \sin(\beta_j) \sin(\gamma_i),$$

где α_i – угол наклона поверхности; γ_i – угол отклонения вектора, соединяющего точки, от географической параллели, направленной в сторону увеличения долготы.

7. Осуществляется выбор по критерию $\frac{C_{i\ start}}{C_{\ start}} \sin \alpha_i \frac{R - r_i}{R} \rightarrow \max$ или $\cos(\beta_j - \gamma) \rightarrow \max$, где $C_{i\ start}$ и $C_{\ start}$ – активности радионуклида во влияющей и базовой точках.

Применение разработанной методики для обработки климатических данных и прогнозирования активностей радионуклидов нейросетевым методом

Средние и общие величины климатических факторов могут быть рассчитаны вручную (например, используя встроенные функции MS Excel) или при помощи специального программного обеспечения [11, 12]. Однако при использовании табличного процессора не все данные могут быть быстро обработаны в силу следующих особенностей: для расчета средних или суммарных величин за определенные промежутки времени границы периодов необходимо задавать вручную в отдельности для каждого временного отрезка; данные могут содержать пропуски, которые невозможно автоматически обнаружить табличным процессором; в табличных процессорах нет встроенных инструментов подсчета количества вхождений величины и одновременного проведения математических операций над связанными значениями.

Для ускорения процесса подготовки климатических данных был разработан программный модуль на языке Java [12]. Он может использоваться для получения на основании часовых данных метеорологических параметров (значений осадков, температур, влажности, скорости ветра и др.)

их средних, минимальных, максимальных или суммарных значений за различные промежутки времени (день, неделя, месяц, сезон, год). Предусмотрена возможность динамического задания значения границ периодов (даты начала и окончания сезона) в процессе работы.

Следует отметить, что разработанный модуль позволяет определять частоты встречаемости указанных значений параметра за выбранный период, что было использовано для определения преобладающего направления ветра и типа погоды на основании дневных или часовых данных. Также предусмотрена возможность вычисления средней скорости каждого из направлений ветра за определенный период [12].

В результате обработки модулем [12] по данным Полесской метеостанции были получены средние значения за сезон (см. табл. 1) температуры и влажности, построена роза ветров (см. рис. 4) за год (период 08.2011 – 07.2012), вычислены суммарные положительные температуры за март (22.02–16.03 – 28 °С) и начало апреля (17.03–09.04 – 109 °С).

Таблица 1

Средние величины температуры и относительной влажности воздуха по сезонам по метеостанции Полесская (2011–2012 гг.)

| Период | Средняя температура, °С | Средняя влажность, % |
|------------|-------------------------|----------------------|
| Весна 2011 | 7,63 | 71,21 |
| Лето 2011 | 18,70 | 76,44 |
| Осень 2011 | 7,05 | 82,48 |
| Зима 2012 | -4,48 | 86,99 |
| Весна 2012 | 8,80 | 72,57 |
| Лето 2012 | 18,39 | 74,82 |



Рис. 4. Роза ветров (за период 08.2011 – 07.2012)

Предложенная методика подготовки данных и алгоритм были применены при создании модуля для работы с искусственными нейронными сетями в составе программного комплекса SPS v2.0 (Simulation Processes in Soil), используемого для прогнозирования перемещения влаги и загрязняющих веществ в почвенно-грунтовой среде и на поверхности почвы. Работа выполнялась в рамках заданий БРФФИ Ф12М-045 и ГПНИ «Радиация, экология и техносфера».

Разработанный модуль использовался для решения задачи прогнозирования миграции ¹³⁷Cs на поверхности почвы и расчета коэффициента поступления указанного радионуклида в бобово-злаковые смеси на экспериментальной площадке СПК «Новое Полесье» Лунинецкого района Брестской области. Для решения задачи прогнозирования перераспределения поверхностной активности и поступления в растения ¹³⁷Cs в зависимости от вида и количества внесенных

удобрений использовалась гомогенная ИНС топологии 12-25-2 с сигмоидальной функцией активации, выходами которой являлись активность ^{137}Cs на поверхности почвы ($A_{\text{почвы}}$) и удельная активность в растениях ($A_{\text{раст}}$). По полученным значениям активностей рассчитывался коэффициент перехода (КП). При верификации результатов были получены следующие значения погрешностей: средняя погрешность расчета КП на верифицирующей выборке – 27,45 %; средняя погрешность расчета $A_{\text{почвы}}$ – 15,58 %; средняя погрешность расчета $A_{\text{раст}}$ – 20,87 %.

Анализ полученных результатов прогнозирования нейросетевым методом показал, что получаемые значения поверхностной активности и коэффициента перехода находятся в пределах доверительных интервалов измеряемых величин (неопределенность измерений для активностей составляет 30 %, для КП – 50 %), что свидетельствует о целесообразности применения многослойного персептрона и предложенной методики для прогнозирования миграции ^{137}Cs на поверхности почвы и его поступления в растения.

Заключение

С целью подготовки данных для нейросетевого прогнозирования миграции радионуклидов на поверхности почвы разработана методика, которая включает:

- способ расчета параметра наклона поверхности с использованием синуса угла наклона поверхности;
- метод замены символьных значений направления ветра их числовыми характеристиками и расчетом значения функции косинуса разности следующих углов: угла отклонения на правления ветра по отношению к западному и угла сдвига вектора, соединяющего точки в пространстве относительно географической параллели, направленной в сторону увеличения долготы;
- способ статистической обработки данных климатических факторов по сезонам с динамическим определением границ;
- метод выбора пар точек, параметры которых включаются в обучающую выборку.

В результате применения предложенной методики из состава входных нейронов ИНС была исключена высота точки, так как этот фактор был учтен при вычислении синуса угла наклона поверхности, что позволило уменьшить размерность вектора входных сигналов. Благодаря исключению из обучающей выборки наборов данных, относящихся к точкам, между которыми не происходит процесс переноса радионуклидов, размер обучающей выборки был уменьшен, что позволило сократить время на обучение и снизить погрешность прогноза.

Методика может быть использована как полностью, так и частично в иных целях: для статистической обработки пространственных данных, подготовки обучающей выборки для прогнозирования вертикальной миграции радионуклидов нейросетевым методом, при нейросетевом анализе поверхностной миграции тяжелых металлов и других вредных веществ.

Список литературы

1. Компьютерное моделирование миграции загрязняющих веществ в природных дисперсных средах / С. П. Кундас, И. А. Гишкелюк, В. И. Коваленко, О. С. Хилько; под общ. ред. С. П. Кундаса – Минск : МГЭУ им. А.Д. Сахарова, 2011. – 212 с.
2. Хайкин, С. Нейронные сети. Полный курс / С. Хайкин. – СПб : Вильямс, 2006. – 1104 с.
3. Алексеев, М. И. Организация отведения поверхностного (дождевого и талого) стока с урбанизированных территорий / М. И. Алексеев, А. М. Курганов. – М. : Изд-во АСВ. – 2000. – 352 с.
4. Агеец, В. Ю. Система радиэкологических контролей в агрофере Беларуси / В. Ю. Агеец. – Гомель : РНИУП «Ин-т радиологии», 2001. – 250 с.
5. Цыбулька, Н. Н. Горизонтальная миграция радионуклидов при эрозии почв / Н. Н. Цыбулька, А. Ф. Черныш, О. М. Жукова // Акт. проблемы экологии : мат-лы I Меж. конф., Гродно, 6-8 окт. 2004 г. : в 2 ч. / ГрГУ; отв. ред.: Н. П. Канунникова. – Гродно, 2005. – Ч. 1. – С. 208–211.
6. Аношко, В. С. Миграция радионуклидов с водно-эрозионными процессами в Белоруссии / В. С. Аношко, А. Ф. Черныш, Н. Н. Цыбулька // 14-е плен. межвуз. коорд. совещание по проблеме эрозионных, русловых и устьевых процессов : мат-лы и краткие сообщ, Уфа, сент.-окт. 1999 г. / МГУ, Баш. гос. ун-т ; под ред. Р.С. Чалова. – Уфа : Баш. гос. ун-т, 1999. – С. 57–59.

7. Руководство по ведению агропромышленного производства в условиях радиоактивного загрязнения земель Республики Беларусь на 1997–2000 гг. ; под ред. И. М. Богдевича. – Минск, 1997. – 76 с.
8. Hilko, O.S. Radionuclides Migration Modelling Using Artificial Neural Networks and Parallel Computing / O.S. Hilko, S.P. Kundas, I.A. Gishkeluk // *European Water*. – 2012. – № 39. – P. 3–13.
9. Хилько, О. С. Применение нейросетевых моделей для прогнозирования горизонтальной миграции радионуклидов с поверхностным стоком / О. С. Хилько, В. И. Коваленко, С. П. Кундас // Сахаровские чтения 2010 года: экол. проблемы XXI века: Мат-лы 10-й межд. научн. конф., Минск, 20-21 мая 2010 г. – Минск : МГЭУ, 2010. – С. 87–88.
10. Вычисление расстояния и начального азимута между двумя точками на сфере [Электронный ресурс] / ГИС и дистанционное зондирование. – Москва, 2001. – Реж. дост.: <http://gis-lab.info/qa/great-circles.html>. – Дата дост. : 25.05.2012.
11. Задорожная, Е. С. Обработка данных перераспределения радионуклидов Cs-137 в почве / Е. С. Задорожная, О. С. Хилько // Сахаровские чтения 2012 года: экол. проблемы XXI века: Мат-лы 12-й межд. науч. конф., Минск, 17–18 мая 2012 г. / МГЭУ. – Минск, 2012. – С. 436–437.
12. Журавлева, Т. А. Обработка метеорологических данных с целью их использования при решении задачи прогнозирования миграции радионуклидов на поверхности почвы / Т. А. Журавлева, О. С. Хилько // Сахаровские чтения 2012 года: экол. пробл. XXI века: Мат-лы 12-й межд. научн. конф., Минск, 17–18 мая 2012 г. – Минск : МГЭУ, 2012. – С. 436.

O. S. Hilko, S. P. Kundas

DATASET PROCESSING METHODS FOR ARTIFICIAL NEURAL NETWORK TRAINING USED TO FORECAST RADIONUCLIDES MIGRATION ON THE SOIL SURFACE

In the article the methodology and algorithm of data processing for artificial neural network training is described, which include: 1) the method of the surface slope calculating, 2) the method of numerical representation of the wind direction, 3) the method of statistical calculation of climatic factors by seasons with the dynamic boundaries, 4) the method of choosing the pairs of points to be included in the training set.