

РАЗРАБОТКА НЕЙРОННЫХ СЕТЕЙ ДЛЯ ПРОГНОЗИРОВАНИЯ МИГРАЦИИ ХИМИЧЕСКИХ ВЕЩЕСТВ В ПОЧВЕ И АЛГОРИТМОВ ИХ ОБУЧЕНИЯ

Докт. техн. наук, проф. КУНДАС С. П., КОВАЛЕНКО В. И., ХИЛЬКО О. С.

Международный государственный экологический университет имени А. Д. Сахарова

Анализ процессов [1–5], влияющих на миграцию химических веществ в почве, показывает, что для осуществления моделирования и прогнозирования в этой области необходимо учитывать большое количество факторов, связанных как с особенностями самого загрязняющего вещества, так и с условиями его распространения. В силу этого большинство разработанных моделей либо значительно упрощают реальные процессы, не учитывая при этом отдельные факторы, либо для уменьшения погрешности моделирования учитывают все наиболее влияющие факторы, что приводит к их значительному усложнению. Однако усложнение модели связано со значительными проблемами в формировании их входных параметров и граничных условий. Причем некоторые входные параметры и в сложных феноме-

нологических моделях можно получать только экспериментальным путем (например, изотерма сорбции водяного пара), некоторые требуют проведения дополнительных расчетов (тепловая задача, расчет движения влаги и т. п.). А это в свою очередь приводит к проблеме численной реализации и применимости модели, так как даже если предложенная модель и позволяет получать корректные результаты, для ее применения необходимо иметь достаточно объемную базу данных входных параметров, что не всегда представляется возможным.

Проведенный анализ существующих математических моделей миграции веществ в почве показывает, что наибольшее распространение получили два класса моделей: эмпирические модели [1–3 и др.], основанные на решении уравнения конвективной диффузии, и модели,

которые базируются на положениях феноменологической теории массопереноса и законах термодинамики [4, 5 и др.]. Первый класс моделей отличается относительной простотой и достаточной для практического использования точностью, но применим для условий, при которых получены эмпирические коэффициенты, входящие в состав уравнения. С учетом современных достижений в области вычислительной математики более перспективным является второй класс моделей, базирующихся на математическом описании реальных процессов, которые протекают в окружающей среде. Они позволяют проводить анализ и осуществлять прогнозирование переноса веществ в почвенно-грунтовых системах при широком варьировании пространственно-временных масштабов, климатических и гидрологических факторов.

Достоинство феноменологической модели заключается в учете большинства из факторов и процессов, влияющих на миграцию веществ в почве. К недостаткам относятся сложность численной реализации и медленная скорость расчета. Достоинством эмпирической группы моделей является их относительная простота, однако для задач прогнозирования в широких временных масштабах они не применимы.

С учетом перечисленных недостатков поставленную задачу можно успешно решить с применением подхода, базирующегося на использовании искусственных нейронных сетей (ИНС). Основной проблемой в применении этого подхода может быть отсутствие необходимой базы для обучения ИНС. Однако к настоящему времени собран большой объем информации по миграции радиоактивных веществ, тяжелых металлов, пестицидов и гербицидов, нефти и нефтепродуктов и других веществ, в силу чего при переходе на нейросетевой подход эту проблему можно решить. При использовании ИНС появляется возможность выполнения быстрых расчетов, что расширяет масштабность применения таких моделей. Применение экспериментальных данных при обучении нейронных сетей позволяет получать результаты, более приближенные к реальным, которые учитывают большинство из факторов, влияющих на миграцию веществ.

В настоящее время существуют некоторые попытки применения искусственных ИНС для

прогнозирования миграции веществ [6–10]. Анализ применения моделей на базе нейронных сетей показал, что они используются в основном для решения задач аппроксимации и классификации. А это либо «универсальные аппроксиматоры», цель которых классифицировать набор собранных на некотором обособленном участке (группа реперных точек) данных [6, 11], либо классификаторы физико-химических характеристик почв, прогнозирующие информацию об искомом параметре [12]. В частности, известны варианты использования ИНС для прогнозирования коэффициентов диффузии и дисперсии в зависимости от имеющихся характеристик почв при решении уравнения конвективной диффузии. Таким образом, известные разработки в этой области не решают в полном объеме задачи радиационного мониторинга и прогнозирования.

Структура нейронных сетей для прогнозирования миграции веществ в почве. Нейронная сеть – это лишь формальный язык, описывающий взаимосвязь между входными и выходными параметрами модели с помощью известных и понятных функциональных зависимостей. Рассмотрим возможность применения для решения поставленной задачи наиболее распространенной нейросетевой архитектуры – многослойной нейронной сети. Основная ее идея состоит в расположении формальных нейронов слоями. Наименьшее число слоев три: входной, скрытый и выходной. Число нейронов во входном и выходном слоях зависит от размерности данных. Следует отметить, что входной слой состоит не совсем из формальных нейронов. Скорее, это просто точки ветвления, распределяющие входной сигнал на входы нейронов скрытого слоя. Иногда входной слой трактуют как препроцессор входных данных (например, линейное преобразование данных к интервалу $(-1, 1)$ или любое другое нормирование). Число скрытых слоев и количество нейронов в каждом из них может быть любым и зависит от сложности решаемой задачи. Точка ветвления (выход) каждого нейрона распределяет сигнал по всем входам нейронов последующего слоя. Наличие хотя бы одного скрытого слоя является принципиально важным, так как совместно с нелинейными элементами, входящими в их состав, позволяет решать лю-

бую линейно неразделимую задачу. Доказательство этого утверждения следует из основной теоремы нейронных сетей – теоремы Стоуна [13].

Исходя из изложенного выше для прогнозирования миграции химических веществ в почве предлагается общая структура многослойной нейронной сети, представленная на рис. 1.

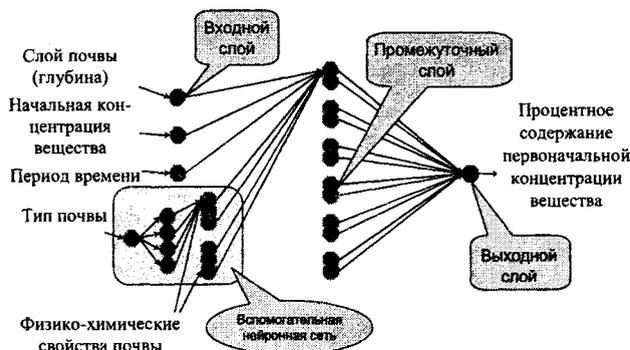


Рис. 1. Структура многослойной нейронной сети для прогнозирования миграции веществ в почве

Как показано на рис. 1, в упрощенном варианте сети на вход подается четыре сигнала. Первый из них – это слой в почве, или глубина, измеряемая в метрах (однако, как было оговорено ранее, необходимо нормализовать эти и все остальные значения), на которой производятся текущие замеры концентрации вещества. Второй – это концентрация веществ на поверхности в начальный момент времени. Третий – характеризует период времени, измеряемый в месяцах (годах), прошедший с момента появления первых результатов измерений концентрации прогнозируемого вещества на поверхности почвы до интересующей пользователя даты. Четвертый – характеризует тип почвы (например, чернозем, дерново-подзолистая, глеевая и т. п.).

Ввиду того что при проведении замеров концентраций химических веществ в большинстве случаев определяются также и физико-химические характеристики почв и грунтов, целесообразным является построение вспомогательной нейронной сети. Она могла бы осуществлять прогнозирование физико-химических характеристик почв исходя из их типов, состава, условий окружающей среды и подавала бы на вход основной нейронной сети прогнозные значения. На рис. 1 она изображается как вспомогательная нейронная сеть.

При анализе предметной области было выявлено, что одними из немаловажных факторов, которые необходимо обязательно учитывать при прогнозировании миграции веществ в почве, являются рельефно-ландшафтные и климатические характеристики [14]. На рис. 2 они составляют отдельные группы данных, подаваемых на входной слой нейронной сети предложенной структуры.

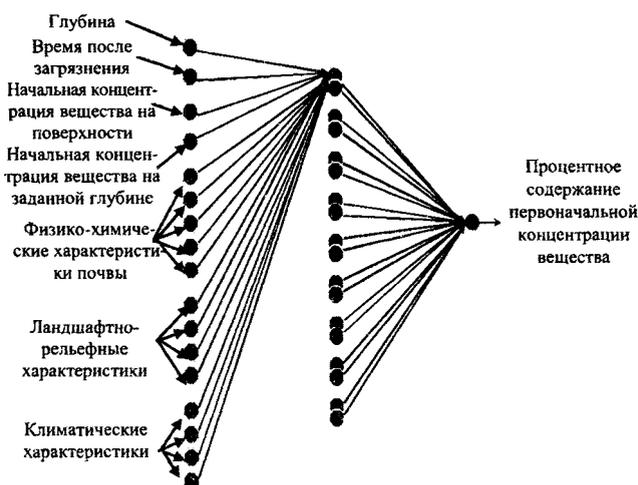


Рис. 2. Уточненная структура многослойной нейронной сети для прогнозирования миграции загрязнителей в почве

В свою очередь в состав климатических характеристик входят такие факторы, как: климатический пояс; средние температуры зимы, весны, лета и осени; средние осадки зимы, весны, лета и осени. К рельефно-ландшафтным характеристикам относятся: ландшафт; тип местности и ее рельеф; высота над уровнем моря. В состав физико-химических характеристик почв внесены: гранулометрические фракции песка, глины, ила; процентное содержание органических веществ; емкость катионного обмена; кислотность и ее способ определения; плотность. При этом для возможности проведения процедур обучения на экспериментальных, а также расчетных данных добавляется дополнительный параметр, характеризующий начальную концентрацию вещества на указанной глубине. Это позволяет использовать нейронную сеть для уточненного прогноза распространения загрязняющего вещества и в большей степени адаптировать нейронную сеть под закономерности происходящих при этом процессов.

Проведенные исследования [14, 15] позволили уточнить структуры основной и вспомогательной нейронных сетей, добавив в их состав дополнительные влияющие факторы (рис. 2, 3).

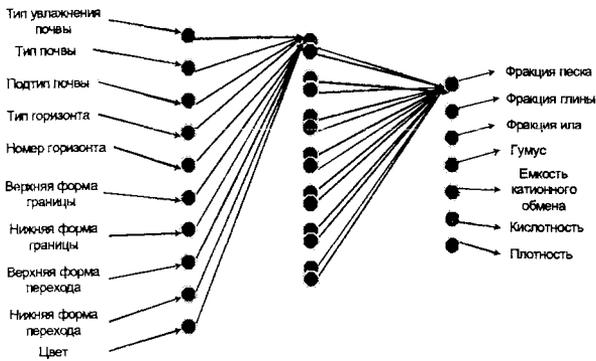


Рис. 3. Уточненная структура вспомогательной сети для прогнозирования физико-химических характеристик почв

Показанная на рис. 3 расширенная структура вспомогательной нейронной сети позволяет также решать задачи перехода от лингвистических форм данных к их численным характеристикам. Такое преобразование не даст точных оценок определения искомых параметров, однако позволяет расширить круг применения предложенного нейросетевого подхода. Благодаря этому появляется возможность использования описательных характеристик почв для проведения оценочных прогнозов распространения химических веществ, а также для решения задачи классификации физико-химических характеристик почв регионов.

Следуя рекомендациям [13], при проведении полных циклов обучения предложенных нейронных сетей число нейронов во входном слое выбирается равным количеству поступающих на нейронную сеть параметров, а число выходных нейронов – соответственно равным количеству искомых параметров.

Особенности обучения разработанной нейронной сети. Для детальной оценки и оптимизации структуры нейронной сети необходимо определить ее соответствующие параметры.

В предложенных структурах ИНС количество слоев больше или равно трем, поэтому анализ проведен для ИНС, состоящих из трех слоев. Их работа описывается следующим выражением:

$$Y = F(F(XW_1 - T_1)W_2 - T_2) - T_3, \quad (1)$$

где $Y = (y_1, y_2, \dots, y_n)$ – вектор выходных сигналов; $X = (x_1, x_2, \dots, x_n)$ – то же входных сигналов; F – оператор нелинейного преобразования; W_i – вектор весовых коэффициентов i -го слоя многослойной нейронной сети (в данном случае $i = \{1, 2\}$); T_i – то же порогов нейронов i -го слоя.

При выборе нелинейной функции активации наиболее применимы: сигмоидная; биполярная сигмоидная; гиперболический тангенс; логарифмическая и др. [9, 13, 16, 17]. Критерием выбора той или иной функции служит: во-первых, возможность получения аналитического решения производной n -го порядка этой функции, а во-вторых, степень приближенности выбранной функции к аппроксимируемой целевой. В силу того что в рассматриваемом случае информация о характеристиках целевой функции отсутствует, исходя из соображений простоты алгоритмизации аналитического решения производной, выбираем сигмоидную функцию активации, выходное значение которой описывается выражением:

$$y_j = \frac{1}{1 + e^{-S_j}}, \quad (2)$$

где y_j – выходное значение j -го нейронного элемента;

$$S_j = \sum_i w_{ij} y_i - T_j, \quad (3)$$

где w_{ij} – i -й весовой коэффициент j -го нейрона; T_j – пороговое значение j -го нейрона.

Как видно из (2) и (3), в этом случае вычисление производной функции y_i по S_j значительно проще [13], чем операции с гиперболическим тангенсом и биполярной сигмоидной функцией в существующих разработках [13, 16].

На рис. 4 условно представлена схема нейрона, работающего в режиме прогнозирования, на который подается вектор входных сигналов $X(x_{t,1}, x_{t,2}, x_{t,3})$ и снимается выходной сигнал y . Для наглядности число нейронов, с которыми связан текущий нейрон на предыдущем слое, равняется трем, что соответствует переменной k . Переменная t обозначает текущий слой, в котором находится нейрон. Вектор входных сигналов взаимодействует с соответствующим

вектором весовых коэффициентов W_i , вектор выходных сигналов взаимодействует с вектором весовых коэффициентов нейронов последующего слоя.

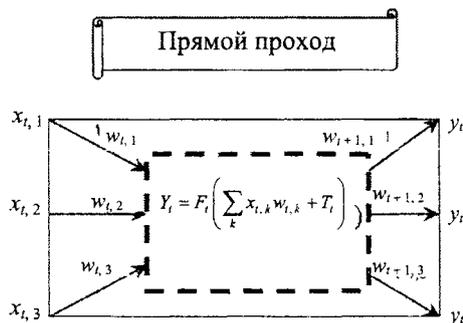


Рис. 4. Схема работы нейрона в режиме прогнозирования

Одной из важнейших задач при разработке многослойной нейронной сети является ее обучение. В настоящее время для этих целей широко применяется алгоритм обратного распространения ошибки [9, 13, 16 и др.]. В его основе лежит метод градиентного спуска в пространстве весовых коэффициентов и порогов нейронной сети. Основу обучения составляют прямой и обратный проходы. Прямой проход представляет собой прямое вычисление выходов, определяемое выражением (1). Обратный проход заключается в подстройке весов текущего слоя в зависимости от ошибки последующего слоя.

Подстройка весов выходного слоя k отличается от аналогичной задачи для скрытого слоя. Это связано с тем, что в выходном слое величина ошибки вычисляется явно. Выход нейрона слоя k , вычитаемый из целевого значения t_j , дает сигнал ошибки. Умножив его на производную сжимающей функции $[y_j(1 - y_j)]$, вычисленную для этого нейрона слоя k , получим

$$\delta = y_j(1 - y_j)(t_j - y_j), \quad (4)$$

где δ – величина ошибки выходного слоя; y_j – рассчитанные значения выходов.

Величина δ , необходимая для вычисления ошибки скрытого слоя, получается суммированием всех таких произведений и умножением на производную сжимающей функции

$$\delta_j = y_j(1 - y_j) \left[\sum_q \delta_{q,k} w_{q,k} \right], \quad (5)$$

где δ_j – величина передаваемой ошибки текущего слоя; $\delta_{q,k}$ – то же последующего слоя k нейронного элемента q ; $w_{q,k}$ – значение весового коэффициента нейронного элемента q слоя k .

Умножая δ на величину y_j нейрона j , можно получить значение весового коэффициента. Это произведение в свою очередь умножается на коэффициент скорости обучения η (обычно от 0,01 до 1,0) и результат прибавляется к весу. Такая же процедура выполняется для веса каждого нейрона от скрытого слоя к выходному.

Представленные ниже уравнения иллюстрируют это вычисление:

$$\Delta w_k = \eta \delta_j y_k; \quad (6)$$

$$w_k(n+1) = \Delta w_k + w_k(n), \quad (7)$$

где Δw_k – изменение величины веса в слое k ; η – коэффициент скорости обучения; δ_j – величина δ для нейрона в последующем слое j ; $w_k(n)$ – величина весового коэффициента одного из нейронов в предыдущем слое на шаге n до коррекции (индекс k относится к слою, в котором заканчивается данный вес, т. е. с которым он объединен); $w_k(n+1)$ – величина веса на шаге $n+1$ (после коррекции); y_k – величина выхода нейрона в слое k .

Таким образом, работу нейронов при обратном проходе можно свести к случаю, показанному на рис. 5.

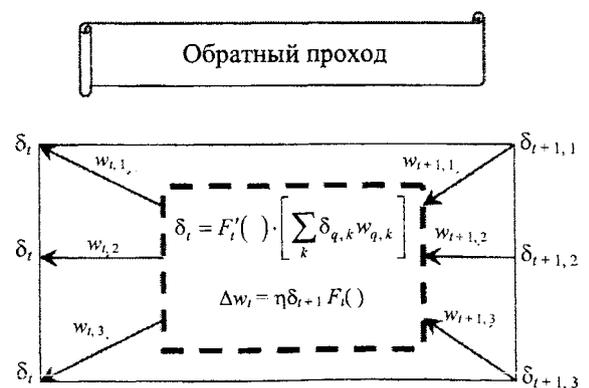


Рис. 5. Схема работы нейронов при обратном проходе

Необходимо также отметить, что важную роль играет величина случайно инициализируемых синаптических связей (весовых коэффициентов). Проведенные исследования [13] показывают, что наиболее эффективным явля-

ется выбор весовых коэффициентов случайным образом в диапазоне $(-0,05; 0,05)$, а также (при большом числе нейронов в слое) в диапазоне

$$\left[\frac{-1}{\sqrt{n(i)}}; \frac{1}{\sqrt{n(i)}} \right], \quad (8)$$

где $n(i)$ – число нейронных элементов в i -м слое.

В предложенных ИНС в силу того что число нейронов в слое небольшое, весовые коэффициенты инициализируются случайным образом в диапазоне $(-0,05; 0,05)$.

Проведенный анализ показал, что применение представленного выше алгоритма связано с рядом проблем. К ним относятся:

- неопределенность в выборе числа слоев и количества нейронов в слое;
- медленная сходимость градиентного метода;
- невозможность определения зон локального и глобального минимума целевой функции;
- сложность получения оптимальной скорости обучения;
- влияние случайной инициализации весовых коэффициентов нейронной сети на поиск минимума функции среднеквадратической ошибки.

Однако, несмотря на указанные проблемы, этот алгоритм на сегодняшний день является одним из лучших. Поэтому его практическое использование является целесообразным при условии минимизации приведенных выше недостатков. С этой целью авторами сделаны попытки разработки модифицированных алгоритмов [13, 16 и др.].

Для увеличения сходимости градиентного спуска и нейтрализации «скатывания» в точки локального минимума используется ряд методов и подходов. Например, описано так называемое «больцмановское» обучение, позволяющее системе делать случайный шаг в направлении, изменяющем целевую функцию, и тем самым выходить из локальных минимумов [13].

Предложен метод «тяжелого шарика», согласно которому в подстройку весов добавляется дополнительный множитель – «моментный терм» [16]. Его введение позволяет выходить из небольших локальных минимумов.

Однако эти меры не дают полной гарантии того, что процесс обучения сойдется. Это не-

приемлемо для практического использования ИНС в предметной области. Они лишь позволяют увеличить вероятность того, что сеть обучится, но при этом значительно увеличивают время, затрачиваемое на процесс обучения, так как возникает необходимость дополнительного просчета состояния сети на каждом итерационном цикле обучения. Одной из эффективных мер в этом случае является использование адаптивного шага обучения, который ускоряет процесс обучения, но также не гарантирует сходимость.

Существует еще целая серия алгоритмов по совершенствованию сходимости метода обратного распространения ошибки, такие как алгоритм Левенберга – Маркара [6], послонного обучения [13] и др.

Для решения указанных выше проблем авторами предложен метод улучшения сходимости обучения разработанных ИНС. Он заключается в подаче значительного дополнительного сигнала на все веса отдельного слоя, начиная с последнего:

$$w_k(n+1) = w_k(n) + \Delta w_k + \xi, \quad (9)$$

где ξ – дополнительный сигнал, величина которого зависит от повторяемости величины ошибки слоя.

При нормальной работе величина сигнала равна нулю. Эта величина неслучайная и непостоянная. Подача сигнала производится только в тех случаях, если суммарная ошибка сети достаточно велика и незначительно изменяется в пределах нескольких итераций обучения. Такая мера приводит к выходу отдельного слоя из состояния как локального, так и глобального минимума. Сигнал при этом сохраняется с торможением в течение нескольких итераций, поэтому слой успевает реабилитироваться и войти с большой вероятностью в состояние глобального минимума для целевой функции.

ВЫВОДЫ

1. Проведен обзор существующих методов и подходов к решению задач прогнозирования миграции химических веществ в почве. Показаны соответствующие достоинства и недос-

татки эмпирического, феноменологического и нейросетевого подходов.

2. Предложено несколько взаимодополняющих структур нейронных сетей, предназначенных для решения поставленных задач и учитывающих большинство факторов окружающей среды (физико-химические, климатические, рельефно-ландшафтные, биологические и др.).

3. Проведен анализ возможных методов обучения разработанных ИНС. Показаны недостатки существующих подходов при решении поставленных задач. Предложены способы повышения сходимости разработанных ИНС.

ЛИТЕРАТУРА

1. **Knatko, V. A.** Characteristics of ^{90}Sr , ^{137}Cs and $^{239,240}\text{Pu}$ migration in undisturbed soils of southern Belarus after the Chernobyl accident // *J. Env. Radioactivity*. – 1996. – Vol. 2. – P. 185–196.
2. **Серебряный, Г. З.** Аналитическая модель миграции радионуклидов в пористых средах / Г. З. Серебряный, М. Л. Жемжуров // *ИФЖ*. – 2003. – Т. 6. – С. 146–150.
3. **Жукова, О. М.** Модель миграции радионуклидов в речной системе / О. М. Жукова, Н. М. Ширяева, Н. К. Мышкина // *ИФЖ*. – 2001. – Т. 4. – С. 70–77.
4. **Прохоров, В. М.** Миграция радиоактивных загрязнений в почвах / В. М. Прохоров // *Физико-химические механизмы и моделирование*. – М., 1981.
5. **Overland** water flow and solute transport: Model development and field-data analysis / Abbasi F. [et al.] // *J. of Irrig. and Drainage ASCE*. – 2003. – Vol. 2. – P. 71–81.
6. **Kanevski, M.** Spatial Estimations and Simulations of environment data using geostatistics and artificial neural network / M. Kanevski, V. Demyanov, M. Maignan // *Proceedings of IAMG-97*. – Barc., 1997. – P. 533–538.
7. **Картирование** пространственных данных при помощи многослойного перцептрона и геостатистики / А. С. Кравецкий // *Институт проблем безопасного развития атомной энергетики РАН*. – 1999. – С. 41.
8. **Simunek, J.** The HYDRUS-1D Software Package for Simulating the One-Dimensional Movement of Water, Heat and Multiple Solutes in Variably-Saturated Media. Version 3.0 / J. Simunek, van M. Th. Genuchten, M. Sejna // *Department of Environmental Sciences, University of California*. – 2005. – P. 240.
9. **Wang F.** The use of artificial neural networks in a geographical information system for agricultural land suitability assessment / F. Wang // *Environ. and Plann. A*. – 1994. – Vol. 2. – P. 265–284.
10. **Timonin, V.** Spatial Prediction of Radioactivity using General Regression Neural Network / V. Timonin, E. Saveleva // *Applied GIS*. – 2005. – Vol. 2. – P. 19.1–19.14.
11. **Ланкин, Ю. П.** Моделирование изменений экологических объектов с помощью нейронных сетей / Ю. П. Ланкин, А. П. Лалетин // *Сибирский экологический журнал*. – 1999. – Т. 6. – С. 449–452.
12. **Ермолаев, О. П.** Использование нейронных сетей в геоинформационном анализе рельефа как фактора эрозии почв / О. П. Ермолаев, А. А. Савельев // 14-е Пленар. межвуз. координац. совещание по проблеме эрозионных, русловых и устьевых процессов: материалы и краткие сообщения. – Уфа: МГУ, 1999. – С. 117–119.
13. **Головко, В. А.** Нейронные сети: обучение, организация и применение / В. А. Головко. – Минск, 2001.
14. **Кундас, С. П.** Применение экспертных систем для прогнозирования миграции радионуклидов и химических веществ в почве / С. П. Кундас, В. И. Коваленко // *Материалы 6-й междунар. науч. конф.* – Минск: МГЭУ им. Сахарова, 2006. – Т. 2. – С. 162–166.
15. **Математическое** моделирование процессов переноса вещества и влаги в почве / С. П. Кундас // *Экол. вестник*. – 2007. – С. 62–71.
16. **Bishop C. M.** Theoretical foundation of neural networks / C. M. Bishop // *Neural computing research group*. – Aston Univ., 1996. – P. 8.
17. **Hongbin, L.** Soil water content forecasting by ANN and SVM hybrid architecture / L. Hongbin, X. Deti, W. Wei // *Environmental Monitoring and Assessment*. – 2008. – Vol. 1–3. – P. 187–193.

Поступила 25.05.2009