

**ПРИМЕНЕНИЕ ИСКУССТВЕННЫХ НЕЙРОННЫХ СЕТЕЙ
ДЛЯ СИСТЕМ ПРОГНОЗА ЗАГРЯЗНЕНИЯ
ВОЗДУШНОГО БАСЕЙНА**

Н.С. Попов, О.В. Кондраков

Кафедра «Химия и промышленная экология», ТГТУ

Представлена членом редколлегии профессором В.И. Коноваловым

Ключевые слова и фразы: искусственная нейронная сеть; корреляционно-спектральный анализ; моделирование загрязнения воздуха.

Аннотация. Рассматривается задача моделирования загрязнения воздушного бассейна на основе искусственной нейронной сети (ИНС). Проведено сравнение двух моделей: на основе ИНС и на основе авторегрессионной модели. Показано, что точность аппроксимации временного ряда выше для модели на основе ИНС.

Для решения задач прогноза загрязнения тропосферы и управления качественным состоянием воздушного бассейна актуальным является построение экспериментальных математических моделей на основе информации, поступающей с автоматизированных систем контроля загрязнения воздуха (АСК ЗВ) в диспетчерский пункт.

В зависимости от постановки задач моделирование окружающей среды использует разнообразные подходы. Выбор подхода часто зависит от временной шкалы решаемой задачи.

Во многих случаях сегодня используют модели в частных производных [1], однако для оперативных задач широко применяют и статистические подходы, основанные на методах авторегрессии, линейной регрессии, линейной множественной регрессии, МГУА [2].

В последние годы все чаще стали использовать искусственные нейронные сети (ИНС). Они находят свое применение в системах распознавания образов, адаптивного управления, прогнозирования, экспертных систем, при аппроксимации функционалов [3-7].

В связи с этим мы решили применить ИНС для решения задачи диспетчерского управления состоянием воздушного бассейна, связанной с прогнозом значения концентрации примеси.

В данной работе поставлены следующие задачи:

а) применение искусственной нейронной сети для прогноза значений концентрации примесей;

б) сравнение двух методов статистического моделирования на основе авторегрессионной модели и на основе искусственной нейронной сети.

Критерием для сравнения двух методов являлась ошибка прогноза, которая вычислялась по формуле:

$$\varepsilon = \frac{1}{N} \sum_{n=1}^N \left(\frac{y_n^3 - y_n^p}{y_n^3} \right)^2 100\% , \quad (1)$$

где $y_n^э$ – экспериментальное значение концентрации примеси; $y_n^р$ – расчетное значение концентрации; N – количество экспериментальных данных.

Также оценивалась трудоемкость применения обоих методов: по времени, затраченном на создание программ расчета τ_1 ; по времени выдачи результата готовой программой τ_2 .

Исходные данные, необходимые для моделирования, представляют собой временные ряды концентраций аммиака, зарегистрированные во время эксперимента на одной из контрольно-замерных станций (КЗС) г. Тамбова. Значения концентраций выражались в $\text{мг}/\text{м}^3$. Одновременно с концентрациями примесей регистрировались: метеопараметры – скорость v (м/с) и направление ветра φ (рад). Временной ряд включал в себя 70 значений. Интервал регистрации данных составлял 20 минут.

Нейронная сеть представляет собой совокупность элементов (нейронов), определенным образом соединенных друг с другом и с внешней средой с помощью связей, определяемых весовыми коэффициентами [1, 2].

Каждый нейрон состоит из элементов 3-х типов: умножителей (синапсов), сумматора и нелинейного преобразователя.

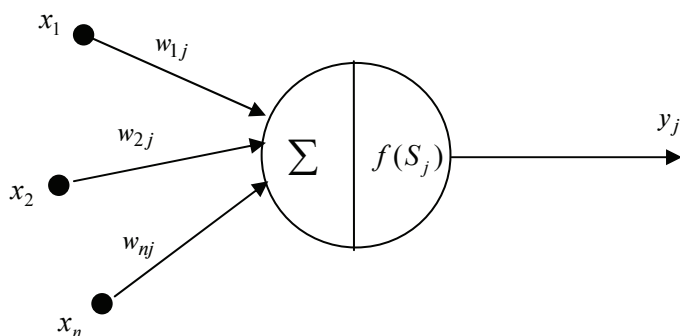


Рис. 1 Общий вид нейрона

Синапсы осуществляют связь между нейронами, умножают входной сигнал x_i на число, характеризующее силу связи или вес синапса w_{ij} . Сумматор выполняет сложение сигналов, поступающих по синаптическим связям от других нейронов, и внешних входных сигналов. Нелинейный преобразователь реализует нелинейную функцию одного аргумента – выхода сумматора.

Математическая модель нейрона:

$$s = \sum_{i=1}^n x_i \cdot w_i + b, \quad (2)$$

$$y_j = f(s_j), \quad (3)$$

где w_i – вес синапса, $i = 1, \dots, n$; b – значение смещения; s – результат суммирования; x_i – компонент входного вектора (входной сигнал); y_j – выходной сигнал нейрона; n – число входов нейрона; f – нелинейное преобразование именуемое как функция активации.

Смещение – параметр нейрона, который суммируется со взвешенными входами нейрона, образуя выходную величину (аргумент) для функции активации, чтобы усилить нелинейность.

На входной сигнал нелинейный преобразователь отвечает выходным сигналом, представляющим собой выход нейрона.

В качестве активационной функции применяются многие выражения. Приведем пример некоторых из них:

а) гиперболический тангенс $f(s) = \frac{\exp(s) - \exp(-s)}{\exp(s) + \exp(-s)}$;

б) линейная функция $f(s) = s$;

в) экспоненциальная функция $f(s) = e^{-s}$.

Одной из наиболее распространенных и удобных для применения в ИНС является нелинейная функция активации с насыщением, так называемая логистическая функция (сигмоид)

$$f(s) = \frac{1}{1 + e^{-s}}. \quad (4)$$

Из выражения сигмоида видно, что выходное значение нейрона лежит в диапазоне (0, 1). Одно из ценных свойств сигмоидальной функции – простое выражение для ее производной, применение которой будет рассмотрено в дальнейшем. Сигмоидальная функция дифференцируема на всей оси абсцисс, что очень удобно при использовании ее в алгоритмах обучения. Кроме того, она обладает свойством усиливать слабые сигналы лучше, чем большие и сглаживает большие сигналы. Данную функцию активации мы использовали при работе персептрона.

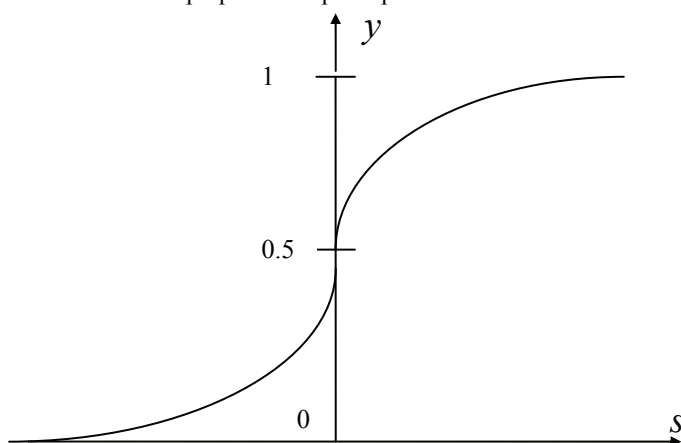


Рис. 2 Сигмоидальная функция активации

Отметим роль нелинейности функции активации. Если бы функция активации была бы линейной, то результат функционирования ИНС сводился бы только к умножению входного сигнала на весовой коэффициент.

Среди различных структур ИНС одной из наиболее известных является многослойная, в которой каждый нейрон произвольного слоя связан со всеми аксонами нейронов предыдущего слоя или, в случае первого слоя, со всеми входами ИНС. Такие ИНС называются полносвязными.

Для решения своей задачи мы выбрали искусственную полносвязную нейронную сеть персептрона Розенблатта [2]. Общий вид сети показан на рис. 3. Персептрон содержал 3 слоя.

Первый слой нейронной сети – входной – на рис. 3 отмечен черными кружками. В нем содержатся входные переменные. Число скрытых слоев, а также нейронов в них в процессе обучения персептрона менялось. Оптимальным был определен один скрытый слой с количеством нейронов равным 20. При таком количестве слоев и нейронов в них нейронная сеть лучше всего аппроксимирует экспериментальные данные, выда-

ет меньше всего ошибок при адаптивной работе и избегает переобучения. Выходной слой содержал один нейрон.

Экспериментальные данные делили на две части. Одна часть использовалась для обучения, а вторая – для контрольной проверки результата функционирования ИНС.

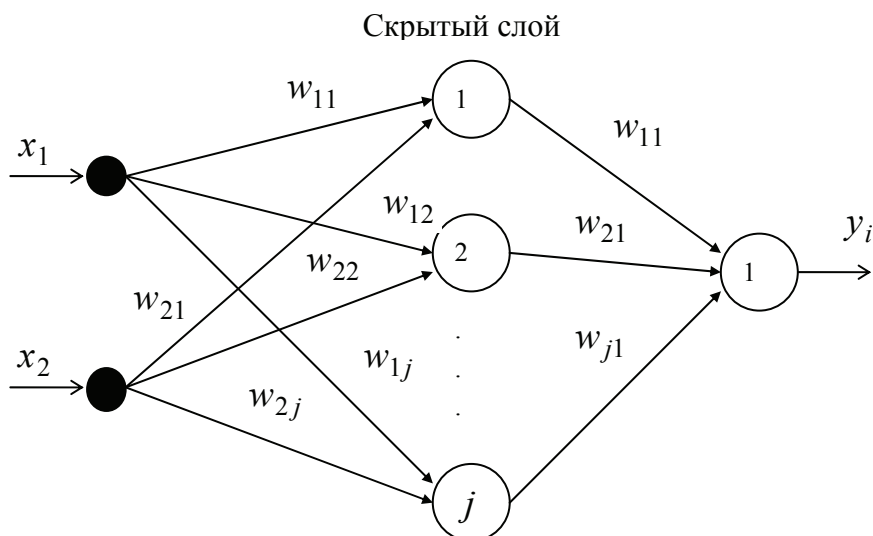


Рис. 3 Архитектура нейронной сети

Обучение ИНС – настройка архитектуры и весов связей для эффективного выполнения поставленной задачи, т.е. поиск закономерности между входными и выходными значениями. Обучить нейронную сеть можно несколькими методами: методом обратного распространения ошибки, методом быстрого распространения, методом упругого распространения, методом масштабируемых сопряженных градиентов.

Мы использовали классический алгоритм настройки персептрона, предложенный Розенблаттом [4]. Его работа сводится к следующему.

Шаг 1. Задание синаптических весов w_{ij} и смещения b случайными числами из диапазона (0, 1).

Шаг 2. Задание персептрону нового входного вектора $X_i = \{x_{1i}, x_{2i}\}$ и выходного сигнала y_i .

Шаг 3. Вычисление выходного сигнала персептрона (2), (3).

Шаг 4. Расчет ошибки $\delta_i^{(n)}$ для выходного слоя n

$$\delta_i^{(n)} = (y_i^{(n)} - y_i) \cdot \frac{dy_i}{ds_i}. \quad (5)$$

Шаг 5. Расчет изменения весов $\Delta w_{ij}^{(n)}$ выходного слоя

$$\Delta w_{ij}^{(n)} = -\eta \cdot \delta_j^{(n)} \cdot y_i^{(n-1)}. \quad (6)$$

Шаг 6. Расчет соответственно $\delta_j^{(n-1)}$ и $w_j^{(n-1)}$ для скрытого слоя

$$\delta_j^{(n-1)} = \left[\sum_k \delta_k^{(n)} \cdot w_{jk}^{(n)} \right] \cdot \frac{dy_j}{ds_j}; \quad (7)$$

$$\Delta w_{ij}^{(n-1)} = -\eta \cdot \delta_j^{(n-1)} \cdot y_i^{(n-2)}. \quad (8)$$

Шаг 7. Корректировка всех весов в ИНС

$$w_{ij}^{(n)}(t) = w_{ij}^{(n)}(t-1) + \Delta w_{ij}^{(n)}(t), \quad (9)$$

где t – номер итерации.

Веса корректируются до тех пор, пока $y^p \approx y^*$. В этом случае цикл завершается и ИНС считается обученной.

Шаг 8. Переход к шагу 2.

По мере обучения сети ошибка обучения убывает, как и ошибка на контрольном множестве. Если контрольная ошибка не убывает, следовательно, ИНС переобучилась (т.е. стала слишком близко аппроксимировать данные). В этом случае рекомендуется уменьшить число нейронов в скрытом слое [4]. Если контрольная ошибка перестала уменьшаться, но при этом $y^p \approx y^*$, следовательно, ИНС является недостаточно адекватной для решения данной задачи, тогда следует увеличить число нейронов в скрытом слое.

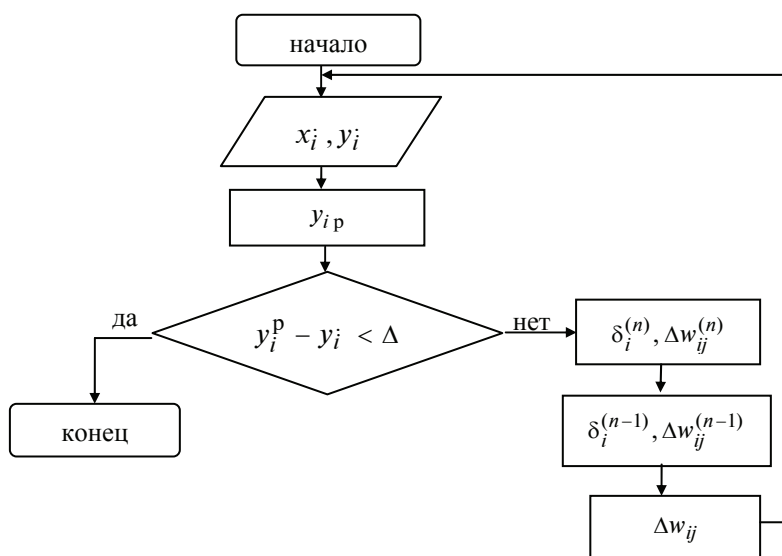


Рис. 4 Блок- схема работы персептрона

На вход персептрона подавали вектор $X_i = \{x_{1i}, x_{2i}\} = \{C_i, x_i\}$, где C_i – значение концентрации примеси в i -й момент времени, $x_i = v_i \cdot \cos \varphi_i$. На выходе персептрона получали значение концентрации в $i + 1$ -й момент времени (C_{i+1}).

Экспериментальные данные перед заданием в персептрон были нормализованы по формуле:

$$x_{i\text{norm}} = 2 \left(\frac{(x_i - x_{\min})}{(x_{\max} - x_{\min})} \right) - 1. \quad (10)$$

Выходные данные подвергались обратному преобразованию

$$y_i = \left(\frac{(y_{i\text{norm}} + 1) * (y_{\max} - y_{\min})}{2} \right) + y_{\min}. \quad (11)$$

После обучения сети проверили ее адаптивную работу на контрольной выборке, рассчитав при этом ошибку прогноза.

Регрессионная модель имела следующий вид [10]:

$$\hat{y}(k+1) = \alpha_0 y(0) + \alpha_1 y(k) + \alpha_2 y(k-1) + \alpha_3 \varphi(k) + \alpha_4 v(k) + \varepsilon(k+1), \quad (12)$$

где $\hat{y}(k+1)$ – оценка прогноза в $(k+1)$ -й момент времени; $y(k), y(k-1)$ – действительные значения концентраций примесей в моменты времени k и $k-1$; φ, v – значения метеопараметров в $i+1$ -й момент времени; ε – ошибка прогноза; $y(0)$ – фиктивная переменная равная 1; α_i – параметры модели ($i = 0, 1, \dots$).

Оценка параметров модели осуществлялась с помощью рекуррентного метода наименьших квадратов:

$$\hat{a}_{i+1} = \hat{a}_i + \eta_i P_{i+1} x_i (y_i - x_i^T \hat{a}_i), \quad (13)$$

$$P_{i+1} = P_i (I - x_i (x_i^T P_i x_i + 1)^{-1} \eta_i x_i^T P_i), \quad (14)$$

где η_i – функция переключения, имеющая вид:

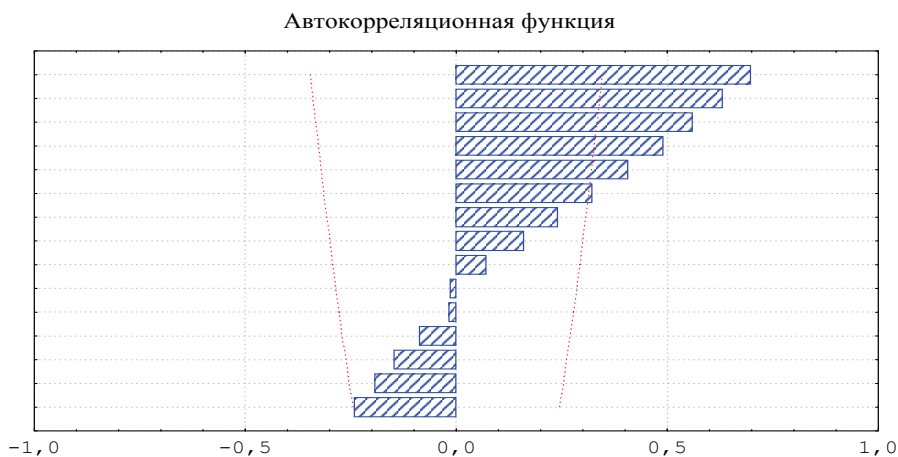
$$\eta_i = 1, \text{ если } \left| \frac{\left[y_i - \hat{y}_i \right]^2}{1 + x_i^T P_{i-1} x_i} \right| > \Delta, \quad (15)$$

в противном случае $\eta_i = 0$.

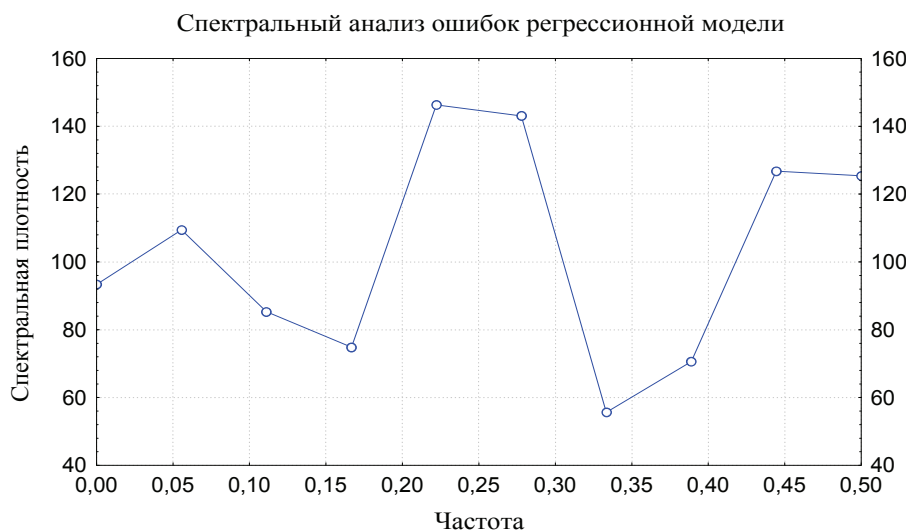
Здесь $\hat{y}_i = x_i^T \hat{a}_i$; I – единичная матрица; Δ – допустимая погрешность модели; $P(0) = I \cdot 10^8$ – ковариационная матрица начального состояния; $\hat{a}_0 = [0, 0, \dots, 0]^T$ – вектор оценок параметров модели; $x^T(y(k), y(k-1), \varphi(k), v(k))$ – вектор независимых переменных; T – знак транспонирования.

Процесс "подгонки" параметров модели считался законченным когда $\eta_i = 0$. В результате прогнозирования по ИНС и по авторегрессионной модели были получены временные ряды ошибок прогноза. Средняя ошибка прогноза по авторегрессионной модели составила 12 %, а по ИНС – 8 %. С временными рядами ошибок прогноза мы проводили корреляционно-спектральный анализ, который показал следующие результаты.

Величины ошибок прогноза по авторегрессионной модели хорошо коррелируют друг с другом. Все значения лежат в диапазоне 95 % доверительного интервала. У значений ошибок прогноза по ИНС корреляция отсутствует. Вид спектральной плотности ошибок прогноза по ИНС показал, что они образуют "белый шум" и не несут в себе полезную информацию. Вид спектральной плотности ошибок прогноза по авторегрессионной модели свидетельствует о неслучайном характере данного временного ряда. Следовательно, этот временной ряд содержит полезную информацию, которую можно извлечь при построении авторегрессионной модели более сложного вида.



а)



б)

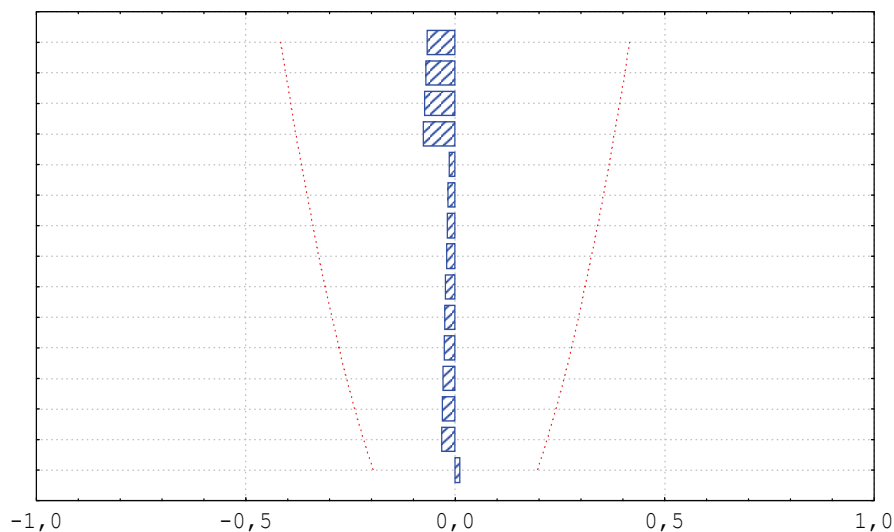
Рис. 4 Корреляционно-спектральный анализ ошибок прогноза по авторегрессионной модели

Анализ результатов показал, что для оперативного прогноза и управления состоянием воздушного бассейна наряду с регрессионными моделями возможно применение ИНС. Ошибка прогноза по авторегрессионной модели примерно в 1,5 раза превышает ошибку прогноза по ИНС. Корреляционно-спектральный анализ показал, что ИНС хорошо описывает экспериментальные данные, а значит ИНС – полезная альтернатива технике традиционного статистического моделирования.

Время создания программы для ИНС τ_1 значительно больше, чем для авторегрессионной модели. Это связано с процессом обучения и построения архитектуры сети. Однако время выдачи результата τ_2 готовой программой по ИНС оказалось меньше, чем по авторегрессионной модели.

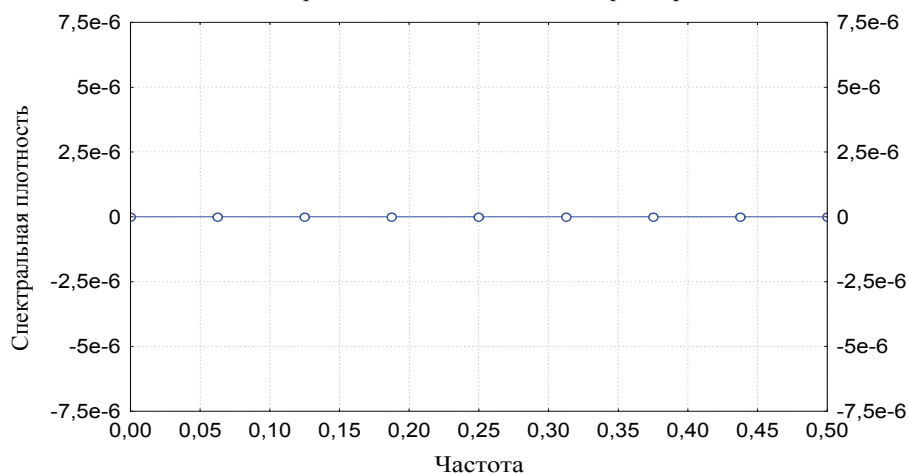
Использование ИНС требует применения быстродействующих вычислительных машин с оперативной памятью не менее 16 мБ и частотой процессора не менее 120 мГц.

Автокорреляционная функция



a)

Спектральный анализ ошибок перцептрона



b)

Рис. 5 Корреляционно-спектральный анализ ошибок прогноза по ИНС

Список литературы

1. Берлянд М.Е. Прогноз и регулирование загрязнения атмосферы. – Л.: Гидрометеоздат, 1985. – 272 с.
2. Попов Н.С., Бодров В.И., Перов В.Л. Основные направления в моделировании загрязнения воздушного бассейна за рубежом / Химическая промышленность за рубежом. 1984. – Вып. 3. – С. 28-45.
3. Уоссермен Ф. Нейрокомпьютерная техника. – М.: Мир, 1992. – 270 с.
4. Круглов В.В, Борисов В.В. Искусственные нейронные сети. – М.: Горячая линия-Телеком, 2001. – 382 с.
5. Матвейкин В.Г., Фролов С.В., Третьяков А.А. Решение задачи регрессии на основе нейронных сетей. – Вестник ТГТУ. 1999. Том 5, №3. – С. 332-343.

6. Gardner M.W., Dorling S.R. Artificial neural networks (the multiplayer perceptron) – a review of applications in the atmospheric sciences // *Atmospheric Environment*. Vol.32, №14/15. Pp. 2627-2636, 1998.

7. Gardner M.W., Dorling S.R. Neural network modeling and prediction of hourly NO_x and NO₂ concentrations in urban air in London // *Atmospheric Environment*. Vol.33, №5. – Pp. 709-719, 1999.

8. Ji Ping Shi, Roy M. Harrison. Regression modeling of hourly NO_x and NO₂ concentrations in urban air in London // *Atmospheric Environment*. Vol.31, №24, pp. 4081-4094, 1997.

9. R.L.R. Salcedo, M.S.M. Alvim Ferraz, C.A. Alves, F.G. Martins. Time –series analysis of air pollution data // *Atmospheric Environment*. Vol.33, №, pp. 2361-2372, 1999.

10. Uwe Schlink, Olf Herbarth, Gerd Tetzlaff. A component time - series model for SO₂ data: forecasting, interpretation and modification // *Atmospheric Environment*. Vol.31, №9, pp. 1285-1295, 1997.

11. Попов Н.С., Перов В.И., Бодров В.И., Рыбинский А.Н. Оперативный прогноз загрязнения воздуха на основе авторегрессионных моделей // *Проблемы контроля и защита атмосферы от загрязнения*. - 1985. - Вып.11. –С. 33- 41.

Application of Artificial Neuron Nets for Forecast Systems of Airspace Pollution

N.S. Popov, O.V. Kondrakov

Department of Chemistry and Industrial Ecology, TSTU

Key words and phrases: artificial neuron model; modeling of air pollution; correlation spectrum analysis.

Abstract: The problem of air pollution modeling on the basis of artificial neuron net (ANN) is considered. The comparison of two models – on the basis of ANN and autoregression model is carried out. The accuracy of temporal series approximation is higher for the model on the basis of ANN.

Anwendung von künstlichen Neuronnetzen für das System der Prognostizierung der Verschmutzung des Luftbeckens

Zusammenfassung: Es ist die Aufgabe der Modellierung der Verschmutzung des Luftbeckens auf Grund des künstlichen Neuronnetzes (KNN) betrachtet. Es ist die Vergleichung von zwei Modellen durchgeführt: auf Grund des KNNs und auf Grund des autoregressiven Modells. Es ist gezeigt, daß die Genauigkeit der Approximation der zeitlichen Reihe für das Modell auf Grund des KNNs höher ist.

Application des réseaux artificiels de neurones pour les systèmes de la prévision de la pollution des bassins aériens

Résumé: On examine le problème du modélage de la pollution du bassin aérien à la base du réseau artificiel de neurones (RAN). On a effectué la comparaison de deux modèles: celui à la base de RAN et celui à la base du modèle autorégressif. On a montré que la précision de l'approximation de la série temporelle est plus grande pour le modèle à la base de RAN.