

- Thermal and Fluid Science*, 2003, no. 44, pp. 147-160.
6. Mayne D. Constrained model predictive control: stability and optimality. *Automatica*, 2000, no. 36, pp. 789-814.
 7. Miroshnichenko V. Programna realizacija analitichnoї modeli upravlinnja procesom vtorinnogo vodopovitrjanogo oholodzhennja bezperervno-litih zagotivok. Anotatsii dopovidey Mizhn. nauk.-tekhn. konf. «Universitetskaja nauka-2018» [The software for the analytic model of control of secondary air-mist cooling for a steel flat product. Abstracts of Int. Sci.-Tech. Conf. «University science-2018»]. Mariupol, PSTU, 2018, vol. 2, pp. 222-224. (Ukr.)
 8. Miroshnichenko V. Analitichne viznachennja rezhimu oholodzhennja listovogo prokatu pri zastosuvanni vodopovitrjanoї sumishi. [An analytic modeling the rolled sheet air-mist cooling]. *Metallurgicheskaja i gornorudnaja promyshlennost' – Metallurgical and Mining Industry*, 2006, vol. 6, pp. 35-37. (Ukr.)
 9. Mishchenko I.O., Dub A.B., Makarycheva E.V. Modelirovanie i optimizacija temperaturnogo polja nepreryvno-litogo slitka [Temperature field modeling and optimization for continuous cast billets]. *Izvestiia vysshikh uchebnykh zavedenii. Chernaia metallurgija – Izvestiya. Ferrous metallurgy*, 2006, no. 3, pp. 15-21. (Rus.)
 10. Deviatov D. Opredelenie koeficientov teplootdachi v zone vtorichnogo ohlazhdenija MNLZ s pomoshh'ju identificiruemoj matematicheskoj modeli [Heat transfer coefficient calculations in secondary cooling zone of continuous caster by using an identified mathematic model]. *Izvestiia vysshikh uchebnykh zavedenii. Chernaia metallurgija – Izvestiya. Ferrous metallurgy*, 1999, no. 8, pp. 62-65. (Rus.)

Рецензент: С.А. Чичкарьов
д-р техн. наук, проф., ДВНЗ «ПДТУ»

Стаття надійшла 13.03.2018

УДК 622.83

doi: 10.31498/2225-6733.36.2018.142552

© Добровольская Л.А.¹, Ключев Д.С.²

ПРОГНОЗИРОВАНИЕ СТЕПЕНИ ЗАГРЯЗНЕНИЯ АТМОСФЕРНОГО ВОЗДУХА В ПРОМЫШЛЕННОМ РЕГИОНЕ

В статье рассмотрена возможность использования искусственных нейронных сетей (НС) для прогнозирования степени загрязнения атмосферного воздуха в промышленном регионе. Разработана многослойная нейронная сеть, позволяющая спрогнозировать уровень загрязнения на основе данных о текущем качестве воздуха, текущих погодных условиях, прогнозе погоды, времени суток и дня недели. Прогноз составляется для каждого часа, для каждой станции, каждого загрязнителя. Горизонт прогнозирования составляет 6 часов. Определено, что максимальная точность достигается при использовании 30 нейронов на скрытом слое, что является оптимальным решением, дающим лучшую точность прогноза. Подтверждено, что создание одной универсальной нейронной сети, которая будет прогнозировать уровень загрязнений для любой станции региона, не дает точного прогноза, т. к. станции находятся в разных средах.

Ключевые слова: прогнозирование, загрязнение атмосферного воздуха, предикторы, искусственная нейронная сеть, нейроны, настройка нейронной сети, обучение нейронной сети.

¹ канд. техн. наук, доцент, ГВУЗ «Приазовский государственный технический университет», г. Мариуполь, ludmila_dobrovolskaya@mail.ru

² студент, ГВУЗ «Приазовский государственный технический университет», г. Мариуполь

Добровольська Л.О., Ключев Д.С. Прогнозування ступеня забруднення атмосферного повітря у промисловому регіоні. У статті розглянута можливість використання штучних нейронних мереж (НМ) для прогнозування ступеня забруднення атмосферного повітря в промисловому регіоні. Нейромережевий підхід є дуже популярним і досить ефективним в рішенні задач прогнозування. Нейронні мережі дозволяють моделювати лінійні залежності в разі великого числа змінних. Завдання прогнозування за допомогою НМ полягає в побудові оптимальної НМ на основі вихідних даних, її навчання за різними алгоритмами, донавчання (при необхідності) і побудови прогнозу. Для побудови якісної мережі досить кілька сотень або тисяч спостережень. Навчання мережі являє собою підгонку моделі, яка реалізується мережею, до наявних навчальних даних. Система, що розробляється, працює в режимі реального часу і в простій і наочній формі виробляє досить якісний прогноз рівня забруднення. Система має можливість охоплювати всі станції моніторингу промислового регіону, збирає і аналізує дані не менш ніж з 100 станцій моніторингу (про забруднення щогодини, про погоду – кожні 15 хвилин). Система працює 24 години на добу. Розроблена багатошарова нейронна мережа, що дозволяє спрогнозувати рівень забруднення на основі даних про поточну якість повітря, поточні погодні умови, прогноз погоди, часу доби і дня тижня, оскільки якість повітря залежить від цих характеристик. Прогноз складається для кожної години, для кожної станції, кожного забруднювача. Горизонт прогнозування становить 6 годин (це стандарт). Протестована точність прогнозу за допомогою нейронної мережі з різною кількістю шарів. Отримано залежності точності прогнозу забруднення повітря від кількості шарів нейронної мережі, звідки видно, що максимальна точність прогнозу досягається з використанням двох зовнішніх і одного прихованого шарів нейронної мережі. Було визначено, що максимальна точність досягається при використанні 30 нейронів на прихованому шарі, що є оптимальним рішенням, що дає кращу точність прогнозу. Підтверджено, що створення однієї універсальної нейронної мережі, яка буде прогнозувати рівень забруднень для будь-якої станції регіону, не дає точного прогнозу, тому що станції знаходяться в різних середовищах.

Ключові слова: прогнозування, забруднення атмосферного повітря, предиктори, штучна нейронна мережа, нейрони, настройка нейронної мережі, навчання нейронної мережі.

L.O. Dobrovolskaya, D.S. Kluev. Forecasting the degree of air pollution in the industrial region. *The article considers the possibility of using artificial neural networks to predict the degree of atmospheric air pollution in the industrial region. The neural network approach is very popular and quite effective in solving forecasting problems. Neural networks make it possible to model linear dependencies in the case of a large number of variables. The task of forecasting with the help of NN consists in constructing an optimal NN based on the initial data, teaching it various algorithms, pre-training (if necessary), and making a forecast. To build a high-quality network, several hundred or thousands of observations are sufficient. Network training is a fitting of the model, which is implemented by the network, to the available training data. The developed system works in real time and in a simple and clear form produces quite a high-quality forecast of the pollution level. The system is able to cover all monitoring stations in the industrial region, collecting and analyzing data from at least 100 monitoring stations (as to pollution-every hour, as to the weather – every 15 minutes). The system works 24 hours a day. A multilayer neural network has been developed that allows to predict the level of pollution based on data on the current air quality, current weather conditions, weather forecast, time of day and day of the week, since air quality depends on these characteristics. The forecast is made for each hour, for each station, each pollutant. The forecasting horizon is 6 hours (this is the standard). The accuracy of the prediction for a neural network with a different number of layers was tested. Dependences of the air pollution forecast accuracy on the number of layers of the neural network have been obtained, which show that the maximum accuracy of the forecast is achieved with the use of two external layers and*

one hidden layer of the neural network. It was determined that the maximum accuracy is achieved with the use of 30 neurons on the hidden layer, which is the optimal solution that ensures a better prediction accuracy. It was confirmed that the creation of one universal neural network, which will predict the level of pollution for any station in the region, does not give an accurate forecast, because stations are in different environments.

Keywords: forecasting, atmospheric air pollution, predictors, artificial neural network, neurons, neural network tuning, neural network training.

Постановка проблеми. В современном мире проблемы прогнозирования загрязнений атмосферного воздуха приобретают первостепенное значение, особенно в крупных промышленных регионах. Чрезмерная концентрация населения в городах, транспорта, промышленных предприятий приводят к химическому изменению и загрязнению воздушной атмосферы. По данным ВОЗ, ежегодно 7 миллионов смертей связаны с воздействием загрязненного воздуха, каждая 8 смерть в мире вызвана сердечно-сосудистыми заболеваниями в результате загрязнения воздуха, из-за загрязнений увеличивается риск заболевания астмой на 15%. Загрязнение воздуха – серьезная проблема для здоровья людей. 98% объема выбросов приходится на 4 загрязняющих компонента: взвешенные частицы; диоксид углерода; оксиды азота и оксиды углерода.

Украина представлена крупными промышленными регионами. Здесь функционируют предприятия горнодобывающей, металлургической, химической промышленности, энергетики, тяжелого машиностроения, эксплуатируется около 300 месторождений полезных ископаемых. Высокая концентрация крупных промышленных предприятий, транспортной инфраструктуры в сочетании со значительной плотностью населения регионов создают огромную нагрузку на биосферу и влияют на состояние воздушного бассейна. Поэтому особо важными и актуальными являются задачи контроля концентраций загрязняющих веществ и прогнозирование состояния воздушного бассейна. Решение этих задач позволит рационально подходить к вопросу размещения новых предприятий в регионе, принимать административные решения в области экологической безопасности, разрабатывать эффективные мероприятия по снижению уровня загрязнения атмосферного воздуха.

Прогнозирование уровня загрязнения атмосферы позволит предупредить население о возможных опасностях и усилить экологический контроль со стороны общества. Используя системы мониторинга, организации смогут корректировать график работы сотрудников на открытом воздухе, а государственные органы предупреждать об опасности проведения различных мероприятий в дни, когда нормы превышены, а промышленные предприятия могут регулировать уровень выбросов вредных веществ в атмосферу [1].

Анализ последних исследований и публикаций. Интерес к вопросам загрязнения атмосферного воздуха возник еще в середине прошлого века, что инициировало исследования данного направления в различных странах мира.

Начало изучению вопроса прогнозирования загрязнений атмосферного воздуха положено научными представителями Соединенных Штатов Америки, которые проводили опытное и оперативное прогнозирование с использованием простейших критериев. Это работы Миллера, Коупера и Смита. На основании их работ была создана система краткосрочного прогноза загрязнения атмосферы на территории страны. Но в их прогнозах отсутствовала возможность прогноза высокого уровня загрязнений воздуха вне ситуации застоя, так как не учитывались закономерности влияния метеорологических условий на загрязнения.

В развитие направления прогнозирования загрязнений атмосферного воздуха внесли вклад представители различных стран мира. В странах СНГ вопросами моделирования прогнозов загрязнений атмосферного воздуха занимались профессора М.Е. Берлянд, Л.Р. Сонькин, Е.Л. Генихович и Р.И. Оникул. Была разработана формула расчета концентрации загрязнителей воздуха в зависимости от геометрической высоты источника.

В Украине, России, Белоруссии действует методика ОНД-86, которая является общепризнанным нормативным документом в большинстве стран СНГ. В ряде западных стран используются официальные программы Агентства защиты ОПС (EPA) [2].

Методика расчета концентраций ОНД-86 ориентирована на расчет максимальной концентрации загрязняющего вещества на максимальном расстоянии от источника выбросов. Данная методика использует эмпирические коэффициенты, полученные для стран СНГ, что не по-

зволяет использовать методику в других странах. Большое количество источников выбросов делает невозможным восстановление целостной картины загрязнения атмосферы региона. Несмотря на указанные недостатки, методика ОНД-86 стала основой большинства отечественных программных продуктов, используемых для расчета загрязнения атмосферного воздуха промышленными предприятиями и получения разрешений на выбросы [3].

Прогнозирование, в основном, осуществляется при помощи статистических методов, базирующихся на основе метода линейной и логарифмической регрессии, корреляционного анализа, также применялись временная авторегрессионная модель и метод группового учета аргументов.

Цель статьи – на основе существующих теоретических и практических материалов разработать интеллектуальную автоматизированную систему прогнозирования загрязнений атмосферного воздуха промышленного региона.

Изложение основного материала. Данные о загрязнении атмосферного воздуха обычно представляются в значениях индексов качества воздуха. Для вычисления индекса загрязненности используют информацию о концентрации загрязнителей воздуха за определенный период времени, полученную в результате мониторинга воздуха. Для преобразования концентраций в индекс загрязненности используют различные специальные математические функции, поскольку различные вещества опасны в разной степени. Полученные данные оцениваются по бальной шкале. Значения индексов в ней группируются в определенном диапазоне. Каждому диапазону присваивается название, цветовой код и описание возможных последствий для здоровья. По значениям этих индексов власти могут давать различные рекомендации населению для уменьшения вредных воздействий на организм.

Единого индекса загрязнения нет. Наибольшее распространение получил индекс AQI (Air Quality Index) США, как наиболее простой и информативный для понимания. Индекс AQI позволяет определить концентрации следующих загрязняющих веществ: приземный озон (O_3), оксид углерода (CO), диоксид серы (SO_2), диоксид азота (NO_2) и взвешенные частицы (PM) размером от 2,5 до 10 мкм. Значение данного индекса принимаются от 0 до 500. Шкала AQI делится на 6 категорий, каждой категории присваивается свой цвет, чтобы люди могли понять насколько загрязнен воздух.

Поэтому важной проблемой при решении задач прогнозирования является учет сложного характера связей между концентрациями и метеорологическими факторами. Большое значение имеет выбор метода, который позволит максимально учесть вид указанных связей.

При разработке схем прогноза загрязнений атмосферного воздуха в промышленном регионе необходимо правильно выбрать предикторы, то есть те характеристики, которые в наибольшей степени определяют изменения концентраций примесей в воздухе. Правильно выбранные предикторы позволят учесть особенности региона по сезонам, связанные с климатическими условиями, рельефом, выбросами и т. п.

В качестве предикторов могут быть использованы: V – скорость ветра (м/с); d – направление ветра; V_1 – скорость ветра на уровне изобарической поверхности 925 гПа (м/с); T – температура воздуха у земли (град С); H – толщина слоя перемешивания (м); P – интегральный показатель загрязнения воздуха, рассчитывается как для совокупности примесей, так и для отдельных примесей и др.

При выборе предикторов для разработки схем прогноза по материалам наблюдений за предыдущие годы проводится анализ связей между загрязнениями воздуха в регионе и отдельными метеорологическими параметрами. Отбираются те предикторы, которые в наибольшей степени коррелированы с характеристиками загрязнения.

Предлагается для проблемы прогнозирования применить нейронные сети (НС). Нейросетевой подход является очень популярным и достаточно эффективным в решении задач прогнозирования. Нейронные сети позволяют моделировать линейные зависимости в случае большого числа переменных.

Задача прогнозирования с помощью НС состоит в построении оптимальной НС на основе исходных данных, её обучения по различным алгоритмам, дообучении (при необходимости) и построении прогноза. Для построения качественной сети бывает достаточно несколько сотен или тысяч наблюдений.

Обучение сети представляет собой подгонку модели, которая реализуется сетью, к

имеющимся обучающим данным. Для этого используется целый ряд алгоритмов обучения сети, а именно: алгоритм обратного распространения, метод спуска по сопряженным градиентам, метод Левенберга–Маркара, метод Квази-Ньютона и другие [4].

Преимущество нейросетевого подхода – то, что он позволяет воспроизводить сложные нелинейные зависимости и выполнять прогноз на любое число шагов (соизмеримое с количеством начальных данных) [4].

Разрабатываемая система должна работать в режиме реального времени и в простой и наглядной форме производить качественный прогноз уровня загрязнения.

Система должна иметь возможность охватывать все станции мониторинга промышленного региона, собирать и анализировать данные не менее чем со 100 станций мониторинга (о загрязнении ежечасно, о погоде – каждые 15 минут). Система должна работать 24 часа в сутки.

Уровень загрязнения будет спрогнозировано на основе данных о текущем качестве воздуха, текущих погодных условиях, прогнозе погоды, времени суток и дня недели, поскольку качество воздуха зависит от этих характеристик. Прогноз составляется для каждого часа, для каждой станции, каждого загрязнителя. Горизонт прогнозирования составляет 6 часов (это стандарт). Принято решение использовать аппарат многослойных нейронных сетей.

Построение модели прогнозирования. Обучающие и тестовые выборки. Модель прогнозирования строится на 4 наборах данных за текущий час (t) и за следующий ($t + 1$): качество воздуха для t , метеорологические данные для t , прогноз погоды для ($t + 1$), час дня и день недели. Входные векторы НС: $W_{(t+1)}$ – состояние погоды, температуры, скорость ветра, направление ветра, влажность; $D_{(t+1)}$ – день недели, время дня ($AQI_{(t+1)}$); $W_{(t+2)}$ – состояние погоды, температура и т. д. ($AQI_{(t+2)}$).

На выходе НС получаем значение AQI загрязнителя для ($t + 1$).

Отработка исходных данных предварительная. НС работает со значениями временных рядов от 0 до 1. Поскольку значения собранных данных выходят за этот диапазон, то для использования полученных данных в НС необходимо отмасштабировать исходные временные ряды. Данная процедура выполняется путем деления значения входного массива на максимальное значение данного параметра.

Настройка НС. После определения масштабов входов определим необходимое количество слоев в НС. Имеется набор данных за 3 месяца и 13 характеристик. Целесообразным будет использование трехслойной НС, то есть НС, имеющую один скрытый слой, т. к. двухслойной НС будет недостаточно для моделирования требуемой функции, а увеличение до четырех слоев ухудшит результаты. Все три слоя НС предсказателя будут полносвязными (Dense). В качестве функции активации эвристическим путем принято решение использовать функцию (ReLU). Структура НС представлена на рис. 1.

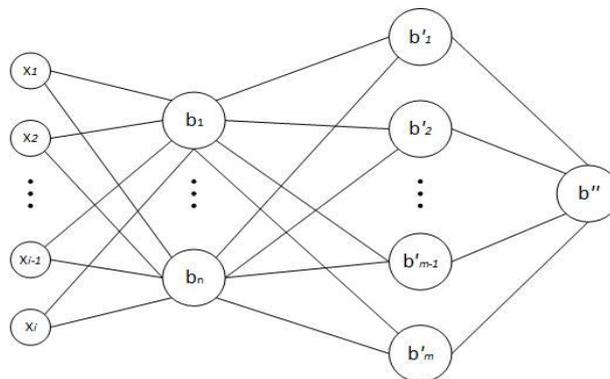


Рис. 1 – Структура НС

На вход НС подается $i = 13$ значений, содержащих информацию о погоде и загрязнении для момента времени t и прогноз погоды для ($t + 1$). Входной слой НС содержит $n = 13$ нейронов, скрытый слой содержит $m = 30$ нейронов. Выходной слой состоит из одного нейрона, так как на выходе необходимо получить одно спрогнозированное значение, которое является AQI загрязнителя для $t + 1$.

Обучение НС и тестовый прогноз. Пусть имеются данные за период с 10.02.2017 по 10.05.2017. Для обучения и проверки точности прогноза разобьем выборку на тренировочную и тестовую в соотношении 70/30, т. е. процесс обучения осуществляется на обучающей выборке (включает в себя данные с 10.02.2017 по 13.04.2017), а в качестве тестовой выборки взяты данные с 14.04.2017 по 10.05.2017.

В качестве функции ошибки используется функция RMSE (the rootmean squared error), а для оптимизации НС используют функцию RMSProp. Оптимальное количество эпох будет равно 100. Для прогнозирования качества воздуха в моменты $(t + 2), \dots, (t + 6)$ в качестве текущего значения AQI загрязнителя подаем предсказанное значение AQI данного загрязняющего вещества в моменты времени $(t + 1), \dots, (t + 5)$.

Расчет точности прогноза. Для определения эффективности разрабатываемой модели необходимо вычислить точность прогнозирования. Рассчитываем точность предсказания по формуле:

$$P = 1 - \frac{\sum_i |y'_i - y_i|}{\sum_i y_i},$$

где y'_i – предсказанное значение AQI; y_i – истинное значение.

Оптимальность топологии НС. Протестирована точность прогноза на НС с разным количеством слоев. Получен график зависимости точности прогноза загрязнения воздуха от количества слоев НС (рис. 2), откуда видно, что максимальная точность прогноза достигается с использованием двух внешних и одного скрытого слоев НС.

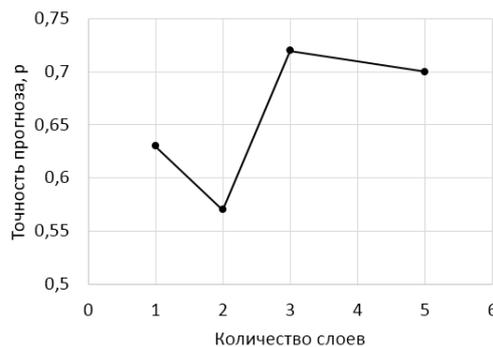


Рис. 2 – Зависимость точности прогноза от количества слоев

Для определения оптимального количества нейронов в скрытом слое проведен ряд тестов расчета прогноза уровня загрязнения воздуха с разным количеством нейронов во втором слое (рис. 3). Было определено, что максимальная точность достигается при использовании 30 нейронов на скрытом слое, что является оптимальным решением, дающим лучшую точность прогноза.

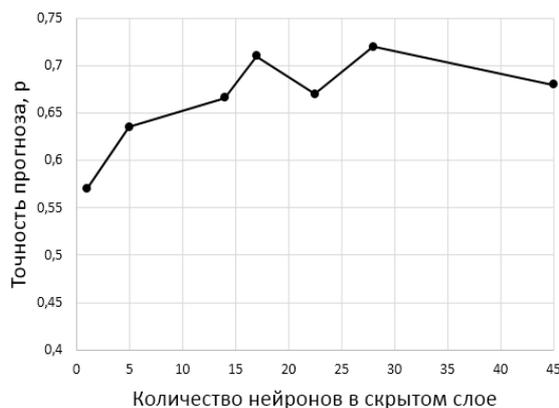


Рис. 3 – Зависимость точности прогноза от количества нейронов

Варианты обучения. Протестировано 3 варианта обучения НС и определено, что оптимальным вариантом будет обучение сети для нескольких близлежащих станций (рис. 4).

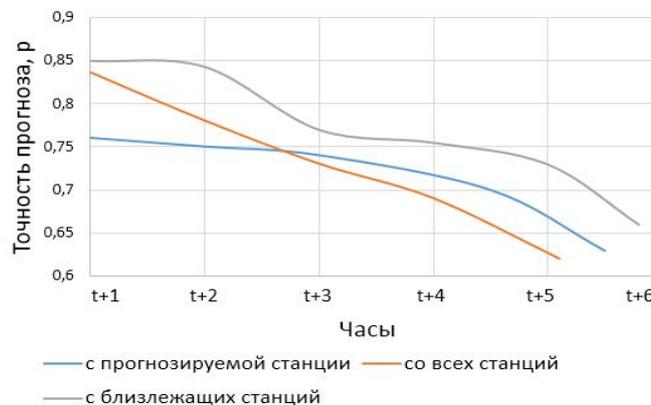


Рис. 4 – Сравнение НС, обученных на разных данных

Создание одной универсальной НС, которая может прогнозировать уровень загрязнений для любой станции города, не дает точного прогноза, т. к. станции находятся в разных средах. Построение прогноза при помощи рассмотренного метода не дает 100% результата, что связано со стохастической природой данных. В дальнейшем можно использовать нейросетевой подход в совокупности с другими моделями.

Выводы

Для прогнозирования степени загрязнения атмосферного воздуха применен аппарат многослойных нейронных сетей. Уровень загрязнения спрогнозирован на основе данных о текущем качестве воздуха, текущих погодных условиях, прогнозе погоды, времени суток и дня недели. Прогноз составляется для каждого часа, для каждой станции, каждого загрязнителя. Горизонт прогнозирования составляет 6 часов. Определено, что максимальная точность достигается при использовании 30 нейронов на скрытом слое, что является оптимальным решением, дающим лучшую точность прогноза. Подтверждено, что создание одной универсальной НС, которая будет прогнозировать уровень загрязнений для любой станции региона, не дает точного прогноза, т. к. станции находятся в разных средах.

Список использованных источников:

1. Сонькин Л.Р. Синоптико-статистический анализ и краткосрочный прогноз загрязнения атмосферы / Л.Р. Сонькин. – Л. : Гидрометеиздат, 1991. – 250 с.
2. ОНД-86. Методика расчета концентраций в атмосферном воздухе вредных веществ, содержащихся в выбросах предприятий. – Введ. 1987-01-01. – Ленинград : Госкомгидромет, 1986. – 68 с.
3. Берлянд М.Е. Прогноз и регулирование загрязнения атмосферы / М.Е. Берлянд. – СПб. : Гидрометеиздат, 1985. – 272 с.
4. Васильев А.Н. Нейросетевое моделирование. Принципы. Алгоритмы. Приложения / А.Н. Васильев, Д.А. Тархов. – СПб. : изд-во СпбПУ, 2009. – 528 с.

References:

1. Son'kin L.R. *Sinoptiko-statisticheskii analiz i kratkosrochnyi prognoz zagriazneniia atmosfery* [Synoptico-statistical analysis and short-term forecast of atmospheric pollution]. Leningrad, Gidrometeoizdat Publ, 1991. 250 p. (Rus.)
2. *Obshchesoiuznyi normativnyi document OND-86. Metodika rascheta kontsentratsii v atmosfernom vozdukhue vrednykh veshchestv, soderzhashchikhsia v vybrosakh predpriatii* [All-Union normative document OND-86. The procedure for calculating the concentrations in the air of harmful substances contained in the emissions of enterprises]. Leningrad, Goskomgidromet Publ., 1986. 68 p. (Rus.)

3. Berlyand M.E. *Prognoz i regulirovanie zagriazneniia atmosfery* [Forecast and regulation of atmospheric pollution]. SPb, Gidrometeoizdat Publ., 1985. 272 p. (Rus.)
4. Vasiliev A.N., Tarkhov D.A. *Neirosetevoe modelirovanie. Printsipy. Algoritmy. Prilozheniia* [Neural network modeling. Principles. Algorithms. Applications]. St. Petersburg, SPbPU Publishing House, 2009. 528 p. (Rus.)

Рецензент: Е.А. Чичкарев
д-р техн. наук, проф., ГВУЗ «ПГТУ»

Статья поступила 15.03.2018