

ИНФОРМАТИКА

COMPUTER SCIENCES

Вестник Сыктывкарского университета.

Серия 1: Математика. Механика. Информатика. 2024.

Выпуск 4 (53)

Bulletin of Syktывkar University.

Series 1: Mathematics. Mechanics. Informatics. 2024; 4 (53)

Научная статья

УДК 004.94

https://doi.org/10.34130/1992-2752_2024_4_39

ПОСТРОЕНИЕ МОДЕЛЕЙ ВЛИЯНИЯ КАЧЕСТВА ОКРУЖАЮЩЕЙ СРЕДЫ НА ЗДОРОВЬЕ НАСЕЛЕНИЯ

Виктор Александрович Рыбак

Белорусский государственный университет информатики
и радиоэлектроники, v.gubak@bsuir.by

Аннотация. В работе рассматриваются вопросы построения различных информационных моделей для анализа и прогнозирования влияния загрязнения атмосферного воздуха и почвенного покрова на здоровье детского населения. Исходными данными выступают числовые ряды, сформированные при исследованиях областных центров Республики Беларусь. Уровни заболеваемости получены с учётом деления территорий городов на районы обслуживания поликлиник. Для оценки и анализа качества окружающей среды выполнены натурные исследования с отбором проб и дальнейшим построением картосхем вероятностного распределения концентраций загрязняющих веществ. В случае атмосферного воздуха средствами геоинформационных технологий выполнена обработка данных об основных источниках загрязнения: промышленных предприятий и автотранспорта, и их объёмов. Для анализа степени загрязнения почв разработаны сетки отбора проб, результаты геохимических исследований

интерполированы в соответствии с новыми алгоритмами для получения непрерывных полей концентраций. Для выявления взаимосвязей в системе «Окружающая среда — здоровье населения» построены модели линейной и квадратичной регрессии. Обосновано применение искусственных нейронных сетей совместно с теорией нечётких множеств. Разработанная нейро-нечёткая система оказалась точнее двух выше указанных моделей. Полученные результаты, кроме научной новизны, обладают практической значимостью и могут быть использованы для оценки эффективности градостроительных и природоохранных мероприятий.

Ключевые слова: нейронные сети, регрессионные модели, анализ качества окружающей среды, нейро-нечёткая модель

Для цитирования: Рыбак В. А. Построение моделей влияния качества окружающей среды на здоровье населения // *Вестник Сыктывкарского университета. Сер. 1: Математика. Механика. Информатика*. 2024. Вып. 4 (53). С. 39–51. https://doi.org/10.34130/1992-2752_2024_4_39

Article

CONSTRUCTION OF MODELS OF THE IMPACT OF ENVIRONMENTAL QUALITY ON PUBLIC HEALTH

Viktor A. Rybak

Belarusian State University of Informatics
and Radioelectronics, v.rybak@bsuir.by

Abstract. The paper considers the issues of constructing various information models for analyzing and forecasting the impact of air and soil pollution on the health of children. The initial data are numerical series generated during studies of regional centers of the Republic of Belarus. Morbidity levels were obtained taking into account the division of city territories into service areas of polyclinics. To assess and analyze the quality of the environment, field studies were carried out with sampling and subsequent construction of maps of the probability distribution of pollutant concentrations. In the case of atmospheric air, data on the main sources of pollution, industrial enterprises and motor transport, and their volumes were processed by means of geoinformation technologies. To analyze the degree of soil pollution, sampling grids were developed, the results

of geochemical studies were interpolated in accordance with new algorithms to obtain continuous concentration fields. Linear and quadratic regression models were built to identify relationships in the „Environment – population health“ system. The use of artificial neural networks in conjunction with the theory of fuzzy sets is substantiated. The developed neuro-fuzzy system turned out to be more accurate than the two above-mentioned models. The obtained results, in addition to scientific novelty, have practical significance and can be used to assess the effectiveness of urban development and environmental protection measures.

Keywords: neural networks, regression models, environmental quality analysis, neuro-fuzzy model

For citation: Rybak V. A. Construction of models of the impact of environmental quality on public health. *Vestnik Syktyvkar'skogo universiteta. Seriya 1: Matematika. Mekhanika. Informatika* [Bulletin of Syktyvkar University, Series 1: Mathematics. Mechanics. Informatics], 2024, no 4 (53), pp. 39–51. (In Russ.) https://doi.org/10.34130/1992-2752_2024_4_39

1. Введение

С развитием информационных технологий и ростом вычислительных мощностей становится возможным создание сложных моделей, отражающих взаимное влияние показателей друг на друга.

По мере возрастания обеспокоенности общества ухудшением качества окружающей среды учёные пытаются выявить статистически значимые закономерности, описывающие влияние степени загрязнения на здоровье населения [1; 2]. Сегодня концепция предельно допустимых концентраций подвергается справедливой критике, хотя она по-прежнему занимает важную часть основы санитарно-гигиенической науки.

Таким образом, с одной стороны, есть заинтересованность граждан получать актуальную информацию о качестве среды обитания и рисках неблагоприятного воздействия, а с другой — существующая система мониторинга не позволяет это осуществить в полном объёме. Кроме того, не существует единой утверждённой методики оценки вредного воздействия на здоровье населения.

Данная проблема может быть решена с использованием элементов искусственного интеллекта для обработки больших объёмов статистической информации [3].

Исследования в области создания первых элементов искусственного интеллекта начались в 40-х годах прошлого века. Благодаря повышению качества оптических приборов учёным удалось установить точную структуру нейронов головного мозга, после чего была создана упрощённая модель нервной клетки. Группируя и объединяя их в сеть, удалось создать инструмент для выполнения логических операций. Далее Дональд Хебб обосновал правила компьютерного обучения [4; 5].

После некоторого прорыва в развитии нейронных сетей в 50-е годы, связанного с их первым использованием в работе компьютеров и созданием алгоритма распознавания образов, научный интерес в данной сфере несколько снизился, что было обусловлено выявленными проблемами по дальнейшей реализации из-за ограниченных вычислительных мощностей того времени. Только к 2000 году мощь вычислительной техники выросла настолько, что позволила реализовать многие ранее предложенные методы и алгоритмы [6; 7].

Кроме нейронных сетей эффективным инструментом для решения нетривиальных задач является теория нечётких множеств (ТНМ), которая берёт своё начало с момента публикации статьи Л. Заде [8], в которой вместо традиционных характеристических функций были предложены новые функции принадлежности элемента подмножеству со значениями в интервале $[0, 1]$. Дальнейшее развитие ТНМ было связано с предложенным Л. Заде принципом обобщения, который ввёл общий метод адаптации стандартных математических средств к обработке нечётких данных [9]. ТНМ и нечёткая логика составляют основу лингвистического подхода. При этом каждое лингвистическое значение характеризуется меткой или, как ещё говорят, первичным термом (например, «высокий») и смыслом, который определяется степенью принадлежности. Такой подход позволяет оперировать близкими к пониманию человека терминами, например «высокий рост», «низкая зарплата», «среднее качество». Вместе с тем предложенные научные методы позволяют осуществлять переход от нечётких множеств к чётким значениям. Этот процесс принято называть «дефузификацией» [10].

Цель работы — разработка и компьютерная реализация моделей с повышенной точностью прогнозирования влияния загрязнения окружающей среды на здоровье населения.

2. Материалы и методы

Для установления и подтверждения закономерностей в системе «Окружающая среда — здоровье населения» в рамках корректировок генеральных планов областных городов Республики Беларусь было выполнено исследование качества основных природных сред и уровня заболеваемости людей. В результате обработки проб почвы, воздуха и воды в среде ArcView Gis новыми алгоритмами были построены карто-схемы вероятностного распределения концентраций. При этом фактические концентрации, измеряемые в мг/м³, обрабатывались с учётом функционально-экологического зонирования территорий для получения оценки по пятибалльной шкале. Фрагмент исходных данных, характеризующих качество атмосферного воздуха, приведён в табл. 1.

Таблица 1

**Итоговая балльная оценка качества воздуха г. Могилёва
(для каждого подрайона), фрагмент**

Код	Подрайон	Площадь га	Балльная оценка загрязнения атмосферного воздуха			
			Р	ПТ	СЗ	сводная
03_05	Юго-западный узел	442,27	3,07	0,23	0,33	3,15
04_02	Речной порт	168,52	3,12	0,10	0,17	3,15
04_19	Габровская – Лавсановское кладбище (усадебная застройка)	159,84	2,43	0,17	1,47	3,02
01_03	Дубровенка	116,70	2,82	0,87	0,26	3,05
03_06	Фатина – Димитрова	235,51	2,14	1,54	0,42	3,10
04_14	Холмы (усадебная застройка)	78,35	1,49	0,14	2,21	3,03
05_02	Могилёв-II – Мелькомбинат	284,34	2,81	0,24	0,40	2,83
06_12	Областная больница	103,19	2,48	0,60	0,50	2,72
04_17	Зона отдыха у озера Святое	135,27	2,69	0,57	0,15	2,86
03_03	Карабановка	315,16	2,10	0,53	0,95	2,84
04_14	Южный узел	604,44	2,65	0,22	0,11	2,80
02_07	Подниколье	109,65	2,73	0,20	0,20	2,64
04_25	Гребенёво (усадебная застройка)	273,17	2,38	0,46	0,41	2,67
05_13	Зелёный Луг (зона перспективной застройки)	196,01	2,40	0,23	0,38	2,51
06_21	Габровская – Гребенёва (коммунально-складская зона)	94,86	2,63	0,16	0,23	2,54
05_05	Юго-западный промузел	442,27	3,04	0,16	0,31	3,20

Примечание. Р – балльные оценки по комплексному показателю загрязнения атмосферы;

ПТ – дополнительные баллы по территориям с повышенными требованиями;

СЗ – баллы по расположению в санитарно-защитных зонах.

Для примера на рис. 1 представлена картосхема загрязнённости атмосферного воздуха г. Могилёва.

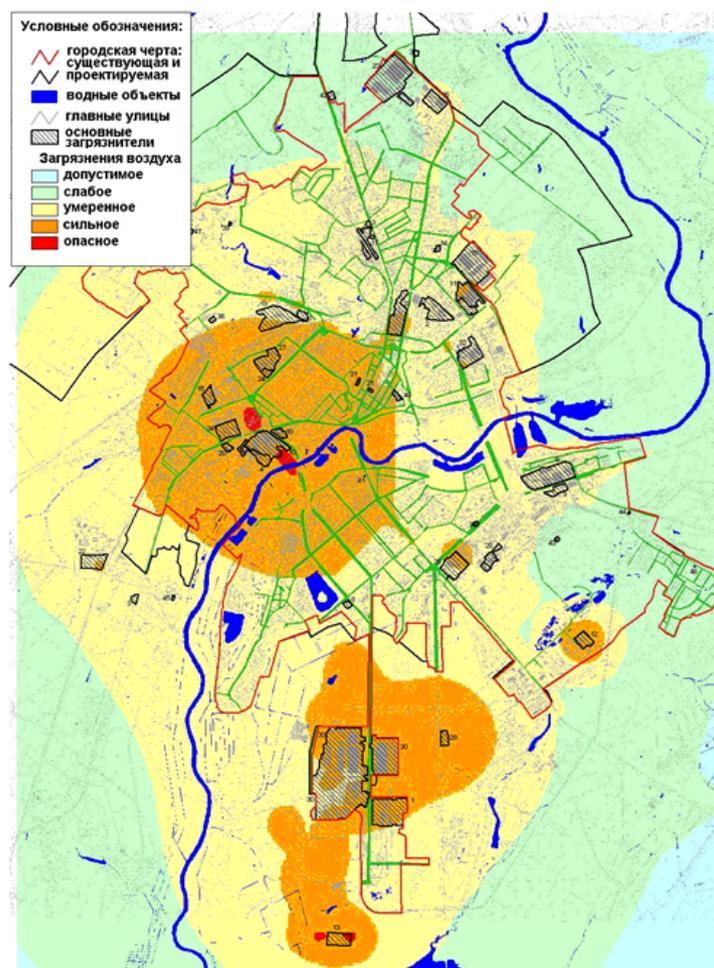


Рис. 1. Степень загрязнённости атмосферного воздуха г. Могилёва

При этом задача получения непрерывных полей по точкам сетки отбора проб и деление на качественные степени не является тривиальной. Для почвенного покрова, например, был разработан следующий алгоритм:

- 1) построение полей концентраций тяжёлых металлов по кригинг-методу, оптимизирующему процедуру интерполяции на основе анализа статистической природы поля;
- 2) расчёт вспомогательного поля на основе комплексного показателя Z_c с учётом его градаций (Z_c — суммарный показатель загрязне-

ния, учитывающий отношение фактической концентрации вещества к геохимическому фону);

- 3) итоговый расчёт балльной оценки с учётом ограничений на концентрацию тяжёлых металлов, зависящих от местоположения и вида использования территорий.

Для анализа качества атмосферного воздуха был разработан новый алгоритм, основанный на официальной методике ОНД-86 [11] с дополнительными модификациями, обусловленными функциональным использованием территорий. Это позволило учитывать определённые требования к уровню загрязнённости для медицинских учреждений, школ, детских садов и др.

Информация о заболеваемости населения ведётся районными поликлиниками и аккумулируется в соответствующих отделах исполкомов. Хотя деление территории города на районы обслуживания поликлиниками является условным, оно позволяет соотносить показатели заболеваемости и уровня загрязнённости (рис. 2) [12]. Сформированные таким образом несколько числовых рядов, характеризующие как показатели здоровья, так и уровень загрязнённости окружающей среды, обрабатывались для выявления взаимосвязей, при этом заболеваемость представлена количеством случаев на 1000 человек, а загрязнённость – балльная оценка, полученная на основании концентраций, выраженных в $\text{мг}/\text{м}^3$.

В исследованиях были использованы показатели детской заболеваемости, так как дети, в отличие от взрослых, в меньшей мере подвержены влиянию производственной среды и, таким образом, находясь в районе проживания, являются более точным индикатором качества окружающей среды.

3. Результаты

Традиционным способом анализа и прогноза является вычисление линейной регрессии. Исходными данными при этом выступают три числовых ряда одинакового размера: x_1 – балльная оценка качества атмосферного воздуха для территории обслуживания поликлиники, усреднённая с учётом вклада нескольких оценок с учётом площади, x_2 – балльная оценка качества почвенного покрова, x_3 – показатель общей детской заболеваемости (случаи на 1000 человек за год). Фрагмент данных представлен в табл. 2.

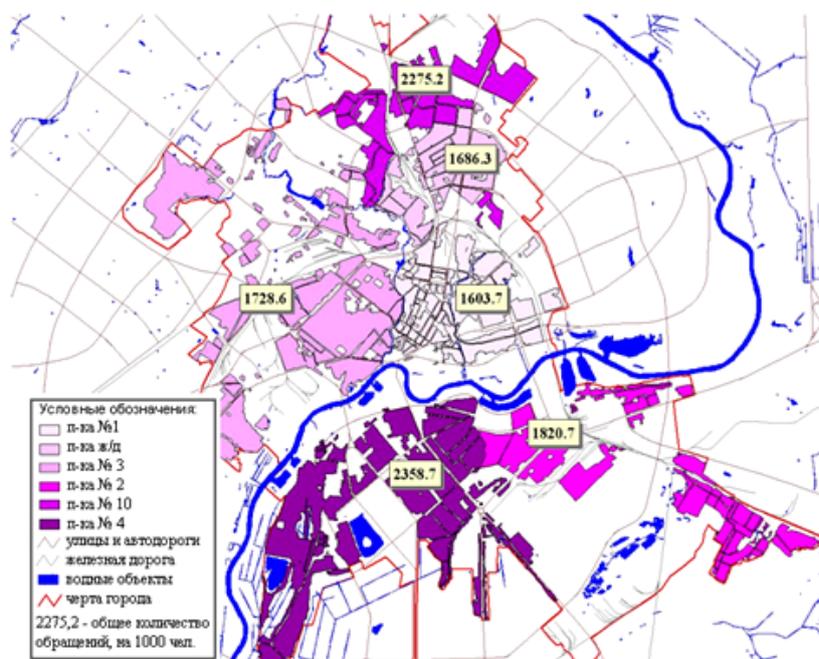


Рис. 2. Показатели общей заболеваемости для участков обслуживания детских поликлиник г. Могилёва, случаи на 1000 человек

Таблица 2

Пример части исходных данных для построения модели

	1	2	3	4	5	6	7	8	9	10
x_1	3,14	4,12	4,41	3,69	3,6	3,9	2,34	3,87	4,23	4,74
x_2	4,87	4,21	2,03	4,29	3,38	4,3	3,13	2,03	3,56	3,51
x_3	1729	2281	1861	1823	1372	1935	1108	1229	1960	2098

Расчёт влияния качества атмосферного воздуха и почвенного покрова на уровень заболеваемости детей в среде MATLAB показал, что средняя квадратичная ошибка получаемой модели $x_3 = 0,373 - 0,098x_1 + 0,458x_2$ на обучающей и тестовой выборках составляют 2,02 и 2,08 соответственно, что не превышает значение 0,2 % от среднего значения заболеваемости.

Модель квадратичной регрессии $x_3 = -25,564 + 0,87x_1 + 0,314x_2 - 0,008x_1^2 + 0,004x_2^2 - 0,002x_1x_2$ имеет ошибки 1,94 и 1,99 соответственно.

Статистическая значимость разработанных моделей была подтверждена с использованием t -критерия Стьюдента и находится на уровне значимости ошибки 0,1 %.

Для повышения точности моделирования обратимся к нейронным сетям и теории нечётких множеств. В отличие от задач, для решения которых происходит разработка и программная реализация точного алгоритма, нейронные сети, подобно живым организмам, «приобретают знания» через обучение. Такая их особенность является решающим отличием от обычных алгоритмов, наделяя их рядом преимуществ. Процесс обучения при этом состоит не в накоплении и хранении данных или знаний, а в подборе регулирующих работу нейронов коэффициентов. При этом нейронные сети способны не только находить очевидные связи между входными и выходными данными, но и выявлять зависимости, о которых исследователь даже не подозревал. Кроме того, данный инструмент эффективно используется для прогнозирования временных рядов, распознавания образов, классификации, кластеризации и др. Дополнительным преимуществом искусственных нейронных сетей является возможность обрабатывать неполные, частично искажённые или «зашумленные» данные [13].

Когда вместо обычной функции активации нейронов используется функция принадлежности нечёткого числа, получаемую систему принято называть нейро-нечёткой. Такие решения являются универсальными аппроксиматорами и были использованы в описываемых исследованиях. С учётом количества числовых рядов была спроектирована нейронная сеть, состоящая из входного, выходного и полносвязного слоя нечётких правил.

Для обучения спроектированной нечёткой системы типа Сугено были использованы обучающая выборка из 120 записей и тестовая из 30. Для этого выполним следующие команды, по которым в рабочей области формируются обучающая выборка (*trn_s*) и тестовая выборка (*tst_s*) [14]:

```
data=load('c:\data.d');
illness_number=length(data);
test_index=[6:10:illness_number 7:10:illness_number];
tr_index=setdiff(1:illness_number, tst_index);
trn_s=ill(tr_index, [1 2 3]);
tst_s=ill(test_index, [1 2 3]);
```

После настройки ошибка обучения (средний квадрат разницы между входными и выходными сигналами) составляет 1,89. Ошибка на тестовой выборке — 1,87.

4. Обсуждение

Для установления зависимостей в системе «Окружающая среда — здоровье населения» были сформированы числовые ряды, отражающие уровень заболеваемости детского населения, степень загрязнения атмосферного воздуха и почвенного покрова для крупных городов и промышленных центров Республики Беларусь. Далее выполнена разработка моделей для обработки числовых данных.

Проведённые исследования и полученные результаты продемонстрировали преимущество нейро-нечёткой модели в точности описания и прогнозирования влияния уровня загрязнения атмосферного воздуха и почвенного покрова на показатели детской заболеваемости на 4–11 % по сравнению с линейной и квадратичной регрессией.

Полученные результаты могут быть использованы на государственном уровне для планирования и анализа градостроительных и природоохранных мероприятий, расчёта их эффективности и сроков окупаемости.

Список источников

1. Балабина Н. М. Роль загрязнения атмосферного воздуха в развитии железодефицитной анемии у взрослого городского населения // *Санитария и гигиена*. 2006. № 6. С. 12–14.
2. Мун С. А. Бенз(а)пирен в атмосферном воздухе и онкологическая заболеваемость в Кемерово // *Санитария и гигиена*. 2006. № 4. С. 28–29.
3. Рыбак В. А., Рябычина О. П. Аппаратное обеспечение системы для экологической диагностики загрязнения атмосферного воздуха // *Известия высших учебных заведений России. Радиоэлектроника*. 2020. № 3. Т. 23. С. 93–99.
4. Чайка П. Нейронные сети: их применение, работа // *Научно-популярный журнал «Познавайка»*. 2018 [Электронный ресурс]. URL: <http://www.poznavayka.org/nauka-i-tehnika/neuronnyie-seti-ih-primenenie-rabota> (дата обращения: 01.11.2024).
5. Режепа В. Простые слова о комплексе: что такое нейронные сети? // *Новостной портал о технологиях*. 2017 [Электронный ресурс].

- сурс]. URL: <https://gagadget.com/another/27575-prostyimi-slovami-o-slozhnom-cto-takoe-nejronnyie-seti> (дата обращения: 01.11.2024).
6. **Иванько А. Ф., Иванько М. А., Сизова Ю. А.** Нейронные сети: общие технологические характеристики // *Научное обозрение. Технические науки*. 2019. № 2. С. 17–23.
 7. **Дударов С. П.** История нейронных сетей // Портал об искусственном интеллекте. 2013 [Электронный ресурс]. URL: <https://neuronus.com/history/5-istoriya-nejronnykh-setej.html> (дата обращения: 01.11.2024).
 8. **Zadeh L.** Fuzzy sets // *Information and Control*. 1965. Vol. 8. Pp. 338–353.
 9. **Zadeh L.** The concept of a linguistic variable and its applications to approximate reasoning // *Information Sciences*. 1975. Vol. 8. Pp. 199–249.
 10. **Леоненков А. В.** Нечеткое моделирование в среде MATLAB и fuzzyTECH. СПб.: БХВ-Петербург, 2005. 736 с.
 11. Методика расчета концентрации в атмосферном воздухе веществ, содержащихся в выбросах предприятий / ОНД–86: Госгидромет. Л.: Гидрометиздат, 1987. 92 с.
 12. **Науменко Т. Е., Рыбак В. А.** Законодательное обеспечение оценки риска воздействия на здоровье населения качества атмосферного воздуха в Республике Беларусь // *Анализ риска здоровью*. 2013. № 1. С. 30–35.
 13. **Рутковская Д., Пилиньский М., Рутковский Л.** Нейронные сети, генетические алгоритмы и нечёткие системы : пер. с польск. М.: Горячая линия – Телеком, 2006. 452 с.
 14. **Рыбак В. А.** Методологические основы принятия решений для управления природоохранной деятельностью. Минск: РИВШ, 2009. 274 с.

References

1. **Balabina N. M.** The role of air pollution in the development of iron deficiency anemia in the adult urban population. *Sanitariya i gigiyena* [Sanitation and hygiene]. 2006. No 6. Pp. 12–14. (In Russ.)
2. **Moon S. A.** Benz(a)pyrene in the atmospheric air and cancer incidence in Kemerovo. *Sanitariya i gigiyena* [Sanitation and hygiene]. 2006. No 4. Pp. 28–29. (In Russ.)
3. **Rybak V. A., Ryabychina O. P.** Hardware of the system for environmental diagnostics of air pollution. *Izvestiya vysshikh uchebnykh zavedeniy Rossii. Radioelektronika* [News of higher educational institutions of Russia. Radio electronics]. 2020. No 3. Vol. 23. Pp. 93–99. (In Russ.)
4. **Chaika P.** Neural networks: their application, work. *Nauchno-populyarnyy zhurnal „Poznavayka“* [Scientific popular magazine „Poznavayka“]. 2018 [Electronic resource]. Available at: <http://www.poznavayka.org/nauka-i-tehnika/neyronnyie-seti-ih-primenenie-rabota> (accessed: 01.11.2024). (In Russ.)
5. **Rezhepa V.** Simple words about the complex: what are neural networks? *Novostnoy portal o tekhnologiyakh* [Technology news portal]. 2017 [Electronic resource]. Available at: <https://gadget.com/another/27575-prostyimi-slovami-o-slozhnom-chto-takoe-nejronnyie-seti> (accessed: 01.11.2024). (In Russ.)
6. **Ivanko A. F., Ivanko M. A., Sizova Yu. A.** Neural networks: general technological characteristics. *Nauchnoye obozreniye. Tekhnicheskiye nauki* [Scientific review. Technical Sciences]. 2019. No 2. Pp. 17–23. (In Russ.)
7. **Dudarov S. P.** History of neural networks. *Portal ob iskusstvennom intellekte* [Portal about artificial intelligence]. 2013 [Electronic resource]. Available at: <https://neuronus.com/history/5-istoriya-nejronnykh-setej.html> (accessed: 01.11.2024). (In Russ.)
8. **Zadeh L.** Fuzzy sets. *Information and Control*. 1965. Vol. 8. Pp. 338–353.

9. **Zadeh L.** The concept of a linguistic variable and its applications to approximate reasoning. *Information Sciences*. 1975. Vol. 8. Pp. 199–249.
10. **Leonenkov A. V.** *Nechetkoye modelirovaniye v srede MATLAB i fuzzyTECH* [Fuzzy modeling in MATLAB and fuzzyTECH environments]. SPb.: BHV-Petersburg, 2005. 736 p. (In Russ.)
11. *Metodika rascheta kontsentratsii v atmosfernom vozdukh veshchestv, sodержashchikhsya v vybrosakh predpriyatiy* [Methodology for calculating the concentration of substances in the atmospheric air contained in emissions from enterprises]. OND-86: Goshydromet. L.: Gidrometizdat, 1987. 92 p. (In Russ.)
12. **Naumenko T. E., Rybak V. A.** Legislative support for assessing the risk of impact on public health of atmospheric air quality in the Republic of Belarus. *Analiz riska zdorov'yu* [Health risk analysis]. 2013. No 1. Pp. 30–35. (In Russ.)
13. **Rutkovskaya D., Pilinsky M., Rutkovsky L.** *Neyronnyye seti, geneticheskiye algoritmy i nehotkiye sistemy : per. s pol'sk.* [Neural networks, genetic algorithms and fuzzy systems : transl. from Polish]. Moscow: Goryachaya Liniya – Telecom, 2006. 452 p. (In Russ.)
14. **Rybak V. A.** *Metodologicheskiye osnovy prinyatiya resheniy dlya upravleniya prirodookhrannoy deyatel'nost'yu* [Methodological foundations of decision-making for environmental management]. Minsk: RIVSH, 2009. 274 p. (In Russ.)

Сведения об авторе / Information about author

Рыбак Виктор Александрович / Victor A. Rybak

к.т.н., доцент, проректор по учебной работе / Candidate of Technical Sciences, Associate Professor, Vice-Rector for Academic Affairs

Белорусский государственный университет информатики и радиоэлектроники / Belarusian State University of Informatics and Radioelectronics 220013, Республика Беларусь, г. Минск, ул. П. Бровки, 6 / 6, P. Brovki Street, Minsk, 220013, Belarus

Статья поступила в редакцию / The article was submitted 02.11.2024

Одобрена после рецензирования / Approved after reviewing 04.11.2024

Принята к публикации / Accepted for publication 20.11.2024