

МЕТОДОЛОГИЯ И МЕТОДИКА ИССЛЕДОВАНИЙ

УДК 550.35: 614.876

АНАЛИЗ ВЛИЯНИЯ ПРИРОДНЫХ И ТЕХНОГЕННЫХ ФАКТОРОВ НА ПЛОТНОСТЬ ПОТОКА РАДОНА В МОСКВЕ С ПРИМЕНЕНИЕМ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ

© 2025 г. С. Г. Гаврильев¹, Т. Б. Петрова^{2,*}, П. С. Микляев^{1,**}, Е. А. Карфидова¹

¹Институт геоэкологии им. Е.М. Сергеева РАН, Уланский пер. 13, стр. 2, Москва, 101000 Россия

²Московский государственный университет имени М.В. Ломоносова, химический ф-т, кафедра радиохимии, Ленинские горы 1, стр. 3, ГСП-1, Москва, 119991 Россия

*E-mail: tbp111@inbox.ru

**E-mail: peterm7@inbox.ru

Поступила в редакцию 12.09.2024 г.

После доработки 17.11.2024 г.

Принята к публикации 25.12.2024 г.

При выполнении инженерных изысканий на строительных площадках г. Москва в ходе плановых измерений плотности потока радона (ППР) были обнаружены аномально высокие значения ППР, превышающие 400 мБк/(м²с) и достигающие >5000 мБк/(м²с). Причины выявленных аномалий до конца не ясны, так как на исследуемой территории нет активных тектонических разломов, считающихся каналами дегазации земной коры, или высоких концентраций урана и радия в геологической среде, достаточных для формирования подобных радоновых аномалий. В представленном исследовании использованы алгоритмы машинного обучения, чтобы попытаться найти ответ на вопрос о природе столь высоких значений ППР. Для прогнозирования аномалий ППР в масштабах города был использован алгоритм Random Forest. В качестве предикторов использовались данные крупномасштабного геологического картирования Москвы и радиационно-экологических изысканий на территории города. Данные для обучения включали результаты измерений ППР на 931 участке, из которых 112 были классифицированы как аномальные (ППР > 400 мБк/(м²с)). На основе полученных прогнозов выделены факторы, влияющие на ППР в фоновых и аномальных условиях. Анализ значимости предикторов показал, что наиболее важные факторы, обуславливающие фоновые значения ППР, — литологический тип грунта, концентрация в грунтах ²²⁶Ra и глубина залегания подземных вод. Аномальные значения ППР проявляют пространственную связь с геодинамически активными зонами, склонами крутизной более 8°, а также с техногенными факторами, такими как линии железных дорог, что, вероятно, обусловлено их вибрационным воздействием на грунт при движении поездов.

Ключевые слова: радон, машинное обучение, Random Forest, плотность потока радона, потенциальная радоноопасность, прогноз

DOI: 10.31857/S0869780925020085 EDN: EPZKGU

ВВЕДЕНИЕ

Радон — природный радиоактивный газ, образующийся в земной коре при распаде радия. В статье речь пойдет о наиболее распространенном и гигиенически значимом изотопе ²²²Rn, дочернем продукте ²²⁶Ra, входящем в цепочку распада ²³⁸U. Благодаря относительной распространенности урана в земной коре, радон всегда присутствует в газовой фазе грунтов и, выделяясь из грунтов, может накапливаться в зданиях в опасных концентрациях. Радон считается вторым по значимости

фактором риска развития рака легких после курения [25]. Наибольшую опасность представляет не столько сам радон, сколько его дочерние продукты распада (ДПР) — радиоактивные изотопы полония, свинца и висмута, осаждающиеся в легких. В связи с этим крайне важно контролировать концентрацию радона и его дочерних продуктов распада в помещениях.

Согласно рекомендациям Международной комиссии по радиологической защите (МКРЗ) максимальная объемная активность радона в жи-

лых помещениях не должна превышать 300 Бк/м³ [11], что соответствует максимальной годовой дозе 10 мЗв. В Российской Федерации нормируется эквивалентная объемная активность (ЭРОА) радона — взвешенная сумма эквивалентных активностей основных дочерних продуктов радона в воздухе. Согласно требованиям Норм радиационной безопасности (НРБ-99/2009) ЭРОА радона в помещениях новых зданий не должна превышать 100 Бк/м³, для эксплуатируемых зданий — 200 Бк/м³. Кроме того, Основными санитарными правилами обеспечения радиационной безопасности (ОСПОРБ-99/2010) регулируется плотность потока радона (ППР) с поверхности грунта на участках строительства, которая не должна превышать 80 мБк/(м²с) для жилых и общественных зданий, для производственных зданий — 250 мБк/(м²с). На участках, где ППР превышает указанные значения, в проекте зданий необходимо предусматривать противорадоновую защиту. За исключением относительно редких случаев, когда в ходе строительства зданий были использованы материалы с повышенной концентрацией природных радионуклидов, основным источником радона в зданиях являются грунты основания.

Известно, что наиболее важные условия, определяющие эксхалицию (выделение) радона с поверхности грунта, — содержание ²²⁶Ra в грунтах и проницаемость грунтов. Последняя, в свою очередь, зависит от пористости и влажности грунта, а также наличия трещин [16]. Кроме того, на эксхалицию радона влияют метеорологические факторы: температура воздуха и грунта, атмосферное давление, количество осадков [9, 14, 16, 24]. Радон, выделяясь из грунтов, может проникать в здания, где накапливается, как правило, на нижних этажах. Накопление радона в зданиях — сложный процесс, зависящий, кроме прочего, от особенностей конструкции здания, типа фундамента, износа конструкций подземной части, состояния врезок коммуникаций, системы вентиляции, типов оконных рам, этажности и многого другого, включая даже такие непредсказуемые факторы, как привычки жителей проветривать помещения [6]. В этой связи прогноз содержания радона в каждом здании очень сложен, однако можно выделить территории, на которых по геологическим условиям вероятность повышенного поступления радона в здания может быть высокой. Такая вероятность называется в отечественной практике потенциальной радоноопасностью территории. В мировой литературе аналогичный показатель называют Geogenic Radon Hazard Index (GRHI) [6]. Основными признаками потенциальной радоноопасности территории служат значения ППР, превышающие допустимый предел (80 мБк/(м²с)), а также повышенные значения концентрации радона в грунтовой газе или содержание (объемной

активности) ²²⁶Ra в грунтах в сочетании с их высокой проницаемостью.

Перенос радона в геологической среде осуществляется посредством молекулярной диффузии благодаря градиенту концентрации на границе литосферы и атмосферы, а также за счет адвективного переноса в потоках газа, формирующихся в крупных порах и трещинах за счет градиентов давления [7].

Плотность потока радона с поверхности грунта может быть оценена с использованием относительно простой формулы, являющейся решением дифференциального **уравнения диффузионного переноса** радона в пористой среде [24]:

$$J_D = C_{Ra} \lambda f \rho_s (1 - \varepsilon) L, \quad L = \sqrt{D_e / \lambda}, \quad (1)$$

где J_D — ППР с поверхности грунта, Бк/(м²с); C_{Ra} — удельная активность ²²⁶Ra в грунтах, Бк/кг; λ — постоянная распада ²²²Rn, 1/с; f — коэффициент эманирования (доля радона, выделяющаяся в поровое пространство), д.ед.; ρ_s — плотность частиц грунта, кг/м³, ε — пористость грунта, д.ед.; L — длина диффузии радона в грунте, м.

Диффузионный перенос радона непрерывно протекает на границе раздела литосферы с атмосферой, так как концентрации радона в массиве составляют десятки тысяч Бк/м³, а в атмосферном воздухе около десяти Бк/м³. Как видно из (1), интенсивность диффузионного переноса (плотность потока) радона зависит, главным образом, от концентрации ²²⁶Ra и длины диффузии радона в грунтах. Длина диффузии радона — среднее расстояние, на которое перемещаются атомы радона в среде с момента своего образования из радия до радиоактивного распада, она зависит от проницаемости грунта, определяющей эффективной пустотностью и степенью водонасыщения грунтов, изменяется от 0.2–0.4 м в слабопроницаемых и/или водонасыщенных грунтах до 1.5–2.0 м в хорошо проницаемых сухих отложениях. Как показывают наши предыдущие исследования [3], используя (1), можно с достаточно высокой точностью оценить средние значения ППР на территориях, где выделение радона из грунтов обусловлено преимущественно диффузионным переносом.

Адвективный перенос — поток радона в объеме движущейся газовой фазы. Применяя закон Дарси к переносу радона, можно получить формулу для оценки адвективной составляющей ППР:

$$J = C_a v_D / \varepsilon_a, \quad (2)$$

где J — плотность потока радона, Бк/(м²с); C_a — объемная активность радона в поровом газе, Бк/м³; v_D — скорость адвективного переноса газа, м/с; ε_a — степень заполнения пор воздухом, д.ед.

Адвективный перенос радона распространен sporadически в пространстве и непостоянен во времени. В пространственном отношении адвективные потоки приурочены, как правило, к зонам повышенной проницаемости геологической среды, где возможны объемные потоки газов. Временная неоднородность адвективного потока грунтового газа обусловлена кратковременным или периодическим действием градиентов давления. Вместе с тем значения ППР за счет адвекции, как правило, значительно (более чем на порядок) выше, чем при диффузионном переносе. Частным случаем адвективного переноса является температурная конвекция в поле силы тяжести.

Конвективный перенос газов, называемый также “эффектом дымовой трубы”, обычно происходит при соблюдении трех условий [15]: 1) наличие проницаемой среды, где возможен объемный перенос газов (зоны трещиноватости, пещеры, горные выработки); 2) наличие перепада высот между точками, где проницаемые зоны выходят на дневную поверхность; 3) наличие перепада температур между грунтовым и атмосферным воздухом. В предыдущих исследованиях [14, 15] был зафиксирован прямой и обратный эффект дымохода в зависимости от времени года, вызывающий очень сильные сезонные колебания радона в зонах повышенной трещиноватости у подножия склонов горных массивов.

Выявление и картирование территорий с высокой потенциальной радоноопасностью для принятия первоочередных мер по защите населения от облучения радоном является приоритетной задачей во всем мире. В последнее время в этих целях все более активно используются возможности машинного обучения, в том числе для прогнозирования эксгаляции радона с поверхности грунта. Больше число работ посвящено изучению методов машинного обучения и выбору наиболее эффективных из них, соответствующих задачам оценки потенциальной радоноопасности.

Такие методы, как Multivariate Adaptive Regression Splines (MARS), Random Forest (RF), Support Vector Machines (SVM), Artificial Neural Networks (ANN), Group Methods of Data Handling (GMDH)¹ и др., ранее использовались для прогнозирования радоноопасности, концентраций радона в помещениях или других, связанных с радоном показателей. Методы машинного обучения также могут быть использованы для прогнозирования временных рядов радона: изменение во времени концентрации радона в почвенном газе для целей прогнозирования землетрясений [22] или для

реконструкции пробелов во временных рядах наблюдений при оценке облучения человека [12]. Возможности прогнозирования и построения карт потенциальной радоноопасности с помощью методов машинного обучения продемонстрированы в работах [8, 19–21].

В качестве исходных данных принимались геологическое строение (тип горных пород или почв), гидрогеологические условия, свойства грунтов, включая плотность, проницаемость, пористость, содержание урана (радия). Прогнозируемой величиной была концентрация радона либо в грунтовом газе, либо в помещениях. В большинстве перечисленных работ методы RF и ANN рассматриваются как лучшие и наиболее популярные для прогнозирования параметров радона, в то время как методы MARS, GBM, SVM и GMDH менее популярны. Однако это не исключает их полностью, поскольку некоторые из этих методов иногда превосходят методы ANN и RF по своей точности.

Цель данной работы — совершенствование методологических аспектов картирования плотности потока радона с использованием методов машинного обучения, а также выяснение факторов, влияющих на выделение радона из грунтов на территории, сложенной мощной толщей рыхлых песчано-глинистых отложений на территории г. Москва. Для целей прогноза плотности потока радона с поверхности грунта был протестирован метод RF, эффективность которого была показана нами ранее [10].

ГЕОЛОГИЧЕСКИЕ УСЛОВИЯ ТЕРРИТОРИИ

Территория Москвы — уникальный объект по объему доступных геологических данных, результатов измерения ППР и других радиационных параметров. Город расположен в центре Московской синеклизы Восточно-Европейской древней платформы. Территория города сложена с поверхности толщей четвертичных отложений, представленных на водоразделах в основном переслаиванием моренных суглинков и флювиогляциальных песков, а в долинах рек преимущественно песчаными аллювиальными отложениями. В юго-западной и южной частях города, южнее долины р. Москва, а также на севере (примерно севернее долины р. Лихоборка) поверхности водоразделов и наиболее высокой III террасы р. Москва перекрыты слоем покровных суглинистых образований мощностью 2–10 м. Общая мощность четвертичных отложений колеблется примерно от 10 м в долинах рек до 60–80 м на юго-западе города в пределах Теплотанского возвышенности. Ниже залегают мезозойские песчано-глинистые отложения, постилающиеся палеозойской карбонатно-терригенной толщей. Наибольшая мощность мезозойских отложений наблюдается в юго-западной части города в районе

¹ Многомерные адаптивные сплайны регрессии (MARS), метод случайного леса (RF), метод опорных векторов (SVM), искусственные нейронные сети (ANN), методы группового учета аргументов (GMDH).

Теплостанской возвышенности, а вдоль современных и древних долин р. Москвы и ее притоков эти отложения частично или полностью размыты эрозией.

На территории города отсутствует как современная тектоническая активность, так и активные разломы в их классическом понимании с наличием современных направленных подвижек вдоль их крыльев. Вместе с тем на территории города выделяются геодинамически активные зоны (ГДАЗ), представляющие собой пространственно локализованные линейные или изометричные объемы (участки) земной коры разного масштаба, в которых, в силу различных причин, имеются или могут возникать условия для концентрации и разрядки дополнительных тектонических напряжений и повышенных градиентов движений и деформированности горных пород [1, 2]. Наиболее крупная система геодинамически активных зон протягивается вдоль долины Москвы-реки (Москворецкая ГДАЗ). Эта система прослеживается далеко на юго-восток вплоть до г. Рязань и далее, по геофизическим данным, сбросовые деформации в этой зоне фиксируются в верхней части земной коры, в том числе и в отложениях чехла. Современная активность зоны в виде деформаций растяжения с правосдвиговой компонентой была выявлена повторными GPS-измерениями [2]. На севере Москвы выделяется также менее крупная Лихоборская ГДАЗ субширотного простирания, а также ряд геодинамически активных зон более мелкого ранга.

Предыдущие исследования [4] показали, что на территории Москвы ППР колеблется в широких пределах. В среднем значения ППР на водоразделах, сложенных глинистыми грунтами, выше, чем в долинах рек, где преобладают песчаные отложения. Средние значения ППР на территориях, сложенных суглинками и песками, составляют 40 и 24 мБк/(м²с) соответственно [4]. Среднее содержание ²²⁶Ra в суглинках — 15–25 Бк/кг, а в песках 10–15 Бк/кг. Вместе с тем на территории Москвы были обнаружены участки, где ППР в отдельных точках принимает аномально высокие значения, превышающие фоновый диапазон и достигающие значений 5000 мБк/(м²с) и более. Эти значения существенно превышают известные в мире уровни и сопоставимы со значениями ППР с поверхности урановых хвостохранилищ, концентрация радия в которых превышает 1000 Бк/кг [24]. Однако зарегистрированные аномалии не связаны с загрязнением территории ураном и радием или повышенным природным содержанием этих радионуклидов в грунтах. На всех аномальных участках удельная активность ²²⁶Ra в грунтах не превышает фоновых значений, характерных для территории города (15–30 Бк/кг). Аномалии составляют примерно 1% от общего количества измерений

ППР, проведенных в Москве [4]. Очевидно, что аномальные значения ППР с поверхности грунта нельзя объяснить диффузионным механизмом. Предположительно, выявленные радоновые аномалии связаны конвективным переносом радона на участках, характеризующихся повышенной проницаемостью грунтов.

Повышенная проницаемость грунтовых массивов в городских условиях может быть обусловлена разными причинами, как природного, так и техногенного характера. Среди природных причин следует упомянуть, прежде всего, развитие на территории Москвы сети геодинамически активных зон и линияментов, которые потенциально могут быть путями переноса радона. Кроме того, в долине р. Москва местами развиты суффозионные и карстово-суффозионные процессы, которые также могут способствовать формированию конвективных потоков радона в геологической среде.

Техногенные причины могут быть связаны, прежде всего, с линиями метро и наземной рельсовой инфраструктурой. Из-за постоянных вибраций структура грунта может деформироваться, а его проницаемость — увеличиваться или уменьшаться. Это особенно актуально для тоннелей метро, которые расположены близко к поверхности и были построены открытым способом выемки грунта, а также для наземных железнодорожных линий. Кроме того, возможно развитие разуплотнения над трассами подземных коллекторов различного назначения, а также над иными подземными сооружениями, не связанными с вибрационными нагрузками. Однако следует отметить, что пространственное воздействие этих факторов может быть очень ограниченным. Возможно сложное сочетание природных и техногенных факторов, влияющих на плотность потока радона из грунта.

МАТЕРИАЛЫ И МЕТОДЫ

Измерение радиационных параметров и применяемые методы

Данные о ППР с поверхности грунта были получены в ходе выполнения инженерно-экологических изысканий для строительства в течение 2002–2008 гг. на 931 участке строительства (до начала строительных работ). Использовались материалы, предоставленные ГК РЭИ и ООО “ГЕОКОН”. Расположение обследованных участков представлено на рис. 1. На каждом участке были выполнены измерения следующих параметров: 1) плотности потока радона, сетка 5 × 10 м, не менее 10 на участок, 2) удельной активности ²²⁶Ra в образцах верхнего слоя почвы с участка, не менее 5 на участок, 3) мощности амбиентного эквивалента дозы (МАЭД) гамма-излучения на высоте 1 м, измерительная сетка 10 × 15 м, не менее 10 на участок.

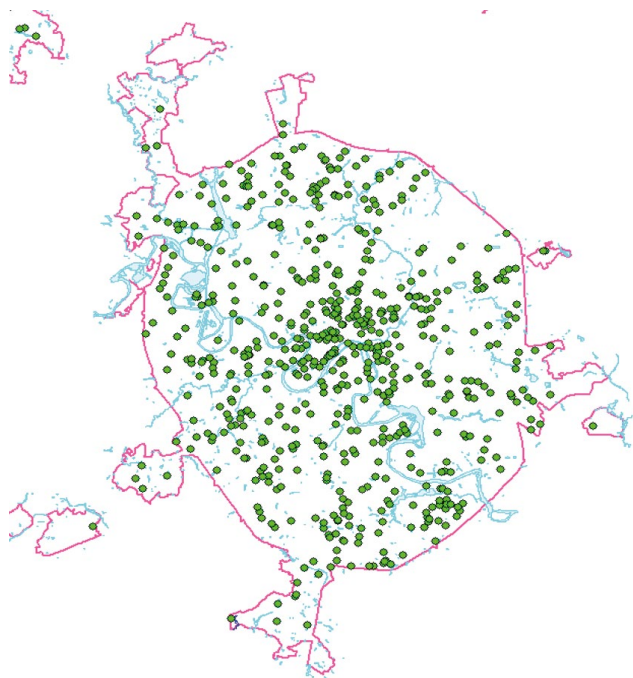


Рис. 1. Схема расположения участков измерений на территории Москвы.

Измерения ППР проводились с помощью измерительного комплекса “Камера-01” с использованием метода открытой камеры с активированным углем [23]. Для измерения ППР на поверхность грунта помещают камеру-накопитель с активированным углем и выдерживают, как минимум, 3 ч для сорбции радона. Затем активированный уголь пересыпают в закрытую колонку и после установления равновесия между радоном и его дочерними продуктами проводят измерения активности радона в угле с использованием бета-радиометра, входящего в состав измерительного комплекса “Камера-01”. Измерительный комплекс позволяет измерять ППР в диапазоне от 3 до 100 000 мБк/(м²с) при температуре окружающей среды от –15°C до 40°C при относительной влажности до 95% с общей неопределенностью результата менее 40%.

Измерения содержания ²²⁶Ra в грунтах проводились с помощью гамма-спектрометрии на основе детектора NaI(Tl) в сосудах Маринелли, объемом 1 л. В пробах предполагалось наличие радиоактивного равновесия между радоном и его дочерними продуктами распада. Всего использованы результаты анализа около 5000 образцов грунта, отобранных на глубине от 0.2 до 20 м.

Измерения МАЭД гамма-излучения проводились с помощью дозиметров гамма-излучения на основе счетчиков Гейгера–Мюллера: ДРГ-01-Т,

ДРГ-012 и ДКГ-07 и т.п. на высоте 1 м над поверхностью земли. Дозиметры могут регистрировать гамма-фотоны с энергией 30–3000 кэВ с погрешностью измерения <30%.

Было сформировано два набора данных, один из которых содержал 819 объектов, включающих только участки с фоновыми значениями ППР, а другой — 931 участок (фоновые + аномальные).

Методы машинного обучения

Предыдущее исследование продемонстрировало эффективность алгоритмов ANN и RF для прогнозирования плотности потока радона на территории Москвы с использованием геологических данных [10]. В этом исследовании использовались алгоритм Random Forest из пакета “Scikit-learn Python”. Для целей обучения и тестирования использовалось соотношение 80 : 20%, что является наиболее распространенной конфигурацией подобного разделения для целей машинного обучения. Категориальные данные, такие как, например, типы четвертичных отложений, были однократно закодированы цифрами. Алгоритм Random Forest использует случайные деревья-решения, которые усиливаются в процессе обучения, если результат прогнозирования близок к реальным данным. Затем полученные деревья, которые являются наиболее близкими, объединяются для получения окончательного прогноза.

Геологические данные

В исследовании использовались крупномасштабные геологические карты территории Москвы в границах МКАД масштаба 1 : 10 000, полученные под научно-методическим руководством Института геоэкологии им. Е.М. Сергеева РАН в ходе реализации проекта “Тематическое геологическое крупномасштабное картирование территории Москвы”. Карты были построены на основе цифровой обработки данных более 85 000 скважин, пробуренных в городе в разные годы, среди которых 13 000 глубоких [5, 17]. Для создания карт использовалась единая городская картографическая основа г. Москвы. В данной работе использовались карта абсолютных отметок кровли четвертичных и дочетвертичных отложений, карта четвертичных отложений, структурно-геоморфологическая карта (м-б 1 : 25 000), карта абсолютных отметок залегания уровня грунтовых вод.

Кроме того, на основе анализа результатов инженерно-экологических изысканий авторами была построена карта среднего содержания радия в четвертичных отложениях и карта мощности дозы гамма-излучения [10]. Для чтения, экспорта и редактирования всех геологических данных использовался пакет QGIS 3.24 Tisler. Чтобы собрать эти данные в модель машинного обучения, при-

меняли Python и TensorFlow, а также несколько библиотек Python, таких как Pandas, Geopandas, Matplotlib, Pyearth, Scikit-learn, Numpy и Scipy.

В итоге в качестве предикторов (исходных данных для прогноза) для применения методов машинного обучения использовались следующие данные (рис. 2):

- 1) тип четвертичных отложений;
- 2) удельная активность ^{226}Ra в приповерхностных грунтах, Бк/кг;
- 3) высота над уровнем моря, м;
- 4) расстояние до бровок склонов крутизной более 8° , м;
- 5) расстояние до линий железных дорог и линий метро, м;
- 6) распределение мощности амбиентного эквивалента дозы по площади, мЗв/ч;
- 7) расстояние до геодинамически активных зон, м;
- 8) расстояние до линеаментов, м;
- 9) глубина залегания подземных вод, м.

РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

Статистическая обработка полного набора данных показывает, что пространственное распределение ППР на территории города подчинено логнормальному закону. Среднее геометрическое значение ППР составляет 29 мБк/(м²с), медиана — 26 мБк/(м²с), геометрическое стандартное отклонение — 2.4. Согласно правилу “трех сигм” диапазон фоновых колебаний ППР составляет от 2 до 400 мБк/(м²с). Важно отметить, что нормируемый уровень ППР с поверхности грунта составляет 80 мБк/(м²с), согласно НРБ-99/2009. В данный диапазон укладываются значения ППР, которые могут быть объяснены преимущественно диффузионным переносом радона из грунта. Участки, на которых максимальные значения ППР > 400 мБк/(м²с), были отнесены к аномальным. Всего было выделено 112 таких участков. В ходе тестирования было обнаружено, что для целей машинного обучения набора данных, включающего только аномальные участки, недостаточно из-за малого объема обучающих данных, поэтому было принято решение использовать комбинацию аномальных и фоновых данных. В результате было создано два набора данных: один включал 819 фоновых значений ППР, а другой — 931 значение и содержал как фоновые, так и аномальные данные. Результирующие карты ППР в Москве как для фоновых данных, так и для полного набора данных показаны на рис. 3.

Модель Random Forest продемонстрировала надежные прогностические возможности. Среднеквадратическая ошибка модели составила 14.9, коэффициент детерминации $R^2 = 0.97$, что отражает высокую корреляцию прогнозных и калибровочных значений. Эти показатели свидетельствуют, что Random Forest эффективно улавливает сложные закономерности и взаимодействия в наборе данных, обеспечивая точные и надежные прогнозы уровней ППР. Примечательно, что полученное значение R^2 — весьма высокое по сравнению с результатами аналогичных исследований, например, в работе [18] $R^2 = 0.24$, а в [13] этот показатель был равен 0.62. Вместе с тем неизвестно, связано ли полученное нами высокое значение R^2 с более точной картой или с чрезмерной подгонкой, осуществляемой алгоритмом Random Forest, что является одним из его недостатков.

Карта прогнозной ППР с применением алгоритма Random Forest и использованием только фонового набора данных представлена на рис. 3а, на котором видно, что более высокие значения ППР в соответствии с прогнозом наблюдаются в юго-западной части города в районе Теплостанской возвышенности, а также в несколько меньшей степени и в северной части города. Области повышенных значений ППР, в целом, совпадают с распространением моренных и покровных суглинков, характеризующихся относительно высоким содержанием ^{226}Ra . Полученная прогнозная карта ППР хорошо согласуется с литологической картой, а также картой радонового поля Москвы, построенной ранее методом интерполяции данных [4], что можно рассматривать как некоторую дополнительную верификацию принятой модели машинного обучения.

Интересно, что искусственные сооружения, такие как железнодорожная инфраструктура, линии метро, а также природные особенности рельефа (тальвеги, крутые склоны (перепады высот), геодинамически активные зоны и линеаменты) не оказывают заметного влияния на фоновую ППР. Корреляция между ППР и расстоянием до перечисленных выше объектов не является значимой, коэффициенты корреляции не превышают значения $r = \pm 0.06$. Вместе с тем анализ алгоритма и результатов прогноза показывает, что значимыми предикторами для прогноза фоновых значений ППР являются тип четвертичных отложений ($r = 0.12$), содержание ^{226}Ra в приповерхностных грунтах ($r = 0.14$), глубина залегания подземных вод ($r = 0.12$) и абсолютная высота участка ($r = 0.19$). Все перечисленные параметры проявляют статистически значимую (при $\alpha = 0.05$) прямую корреляцию с ППР с поверхности грунта. Однако следует отметить, что все эти параметры определенным образом связаны между собой. Возвышенные участки (Теплостанская и Смолен-

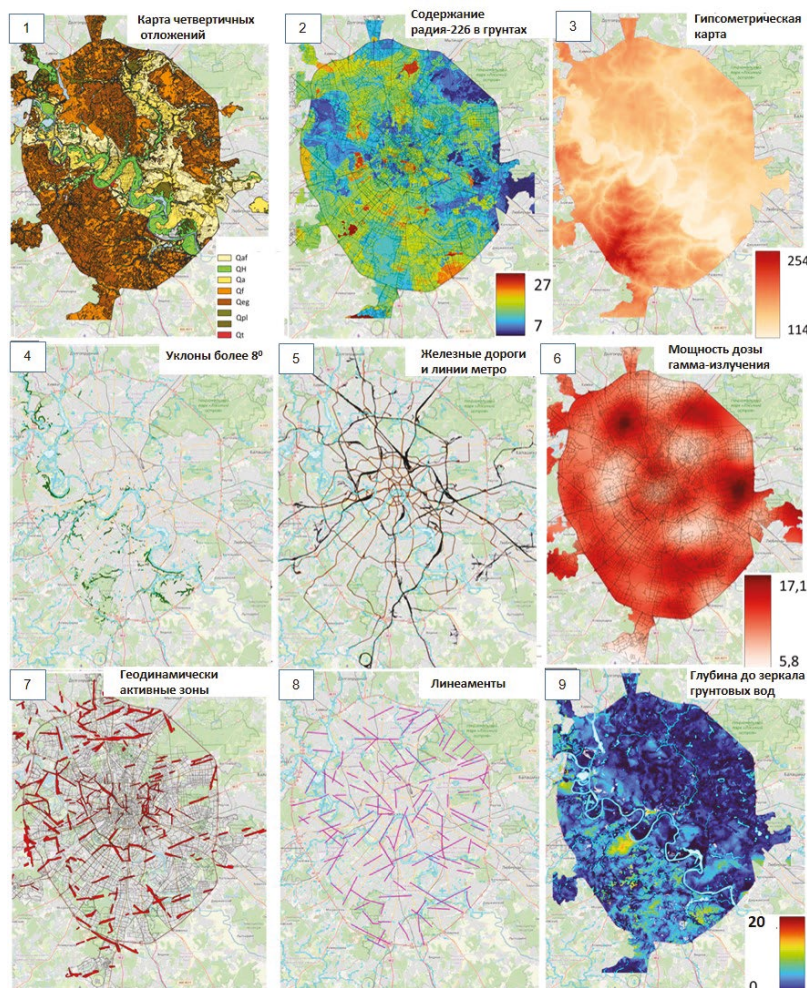


Рис. 2. Карты первичных предикторов, используемые для прогнозов с применением методов машинного обучения (пояснения в тексте).

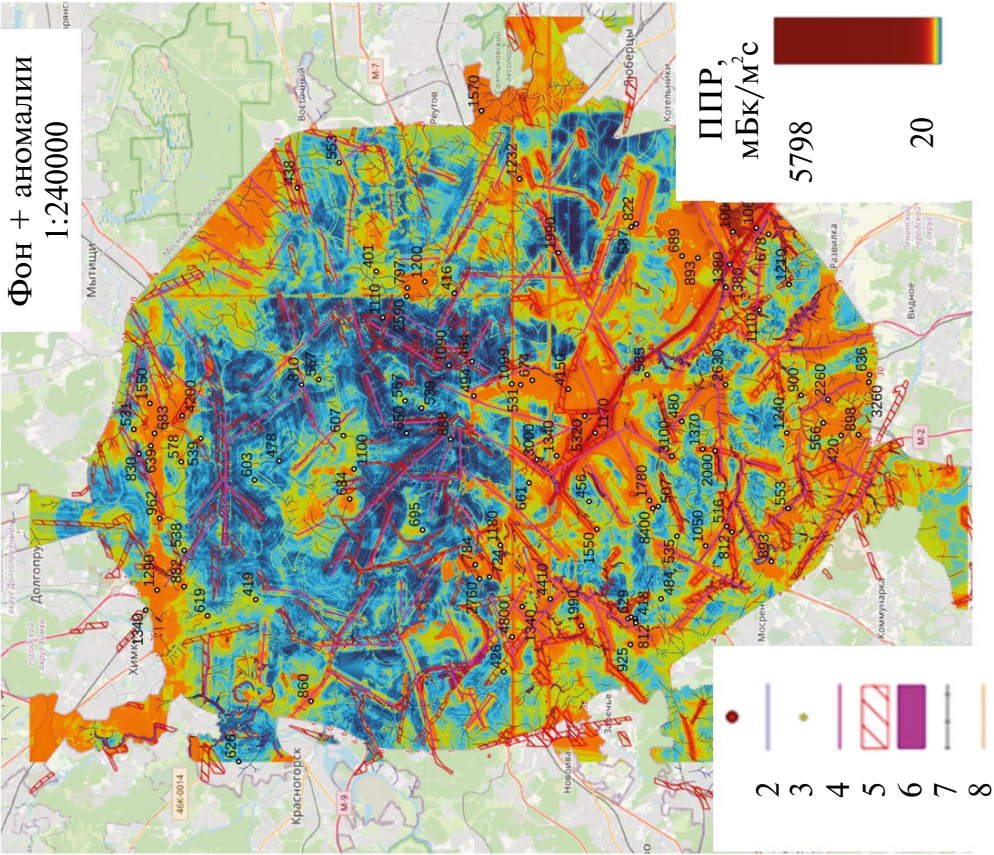
ско-Московская возвышенности) сложены преимущественно суглинками, содержащими более высокие концентрации ^{226}Ra , и характеризуются более глубоким залеганием подземных вод. В то время как долины рек, а также Мещерская низменность на востоке города сложены преимущественно слаборадиоактивными песками с неглубоким залеганием подземных вод. Установленный комплекс факторов свидетельствует, что фоновая ППР формируется преимущественно за счет диффузионного выноса радона из приповерхностных грунтов.

Следует отметить, что авторам пока еще не удалось довести алгоритм до совершенства, и, например, в южной части карт (см. рис. 3а,б) можно видеть горизонтальную линию, пересекающую весь город от Солнцево до Люберец. К сожалению, эта линия объясняется несовершенством работы алгоритма, когда алгоритм воспринял заданные координаты Y как предиктор более высокой плотности потока радона. В настоящий момент

ведется работа по совершенствованию алгоритма для устранения данного эффекта.

Картина меняется с введением в набор данных по аномальным значениям ППР (см. рис. 3б). В этом случае алгоритм воспринимает в качестве значимых предикторов такие природные и техногенные факторы, как геодинамически активные зоны, градиенты уклонов земной поверхности, а также линии железной дороги. Геодинамически активные зоны и градиенты уклонов рельефа оказывают заметное влияние, которое, однако, в значительной степени зависит от местоположения участков. На возвышенностях воздействие этих факторов сильнее, чем в долинах рек. Наблюдается относительно тесная значимая отрицательная корреляция между ППР на аномальных участках и расстоянием до ближайшей геодинамически активной зоны ($r = -0.23$) и бровки склонов, крутизной более 8° ($r = -0.22$). Среди техногенных факторов в качестве наиболее сильного предиктора, влияющего на формирование аномальных

(б)



(а)

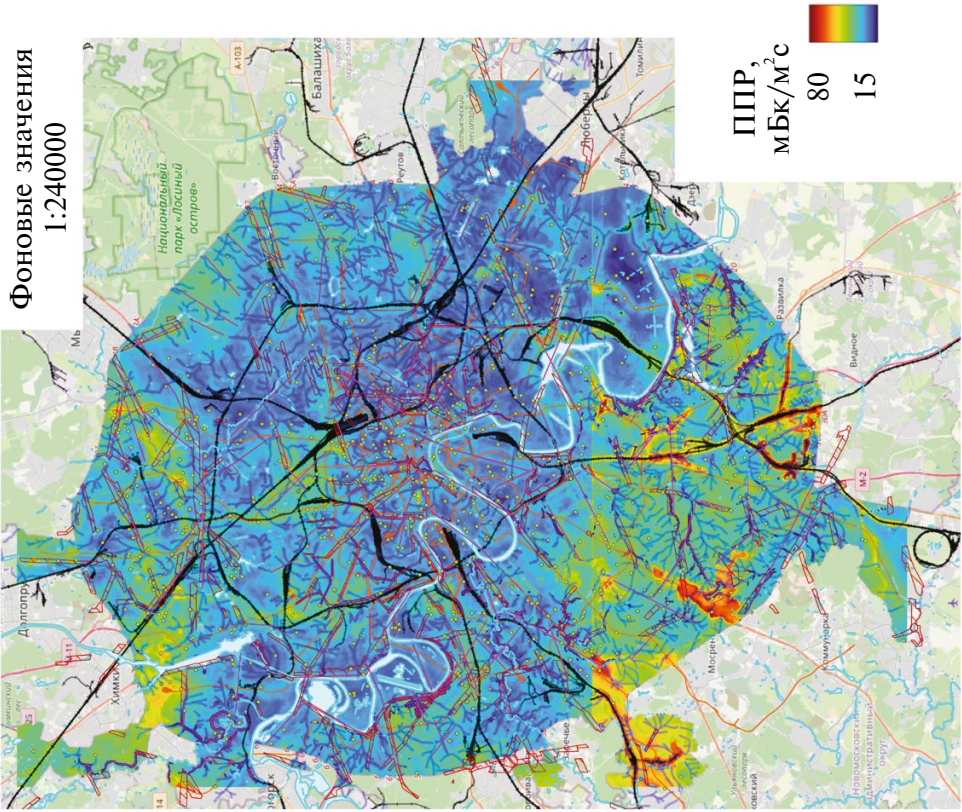


Рис. 3. Карта прогноза ППР, составленная с использованием: а – фоновых значений и алгоритма Random Forest; б – полного набора данных (аномальные и фоновые участки). Алгоритм Random Forest с квантильным цветовым масштабированием. Условные обозначения: 1 — участки с аномалиями ППР, 2 — тальвеги, 3 — участки с фоновыми значениями ППР, 4 — линейные элементы, 5 — геодинамически активные зоны, 6 — уклоны более 8°, 7 — железные дороги, 8 — линии метро.

значений ППР, алгоритм рассматривает расстояние до железнодорожных линий ($r = -0.15$). Тогда как трамвайные линии и линии метро, как это ни странно, не рассматриваются алгоритмом как важные, значимая корреляция между ППР и расстоянием до этих объектов техногенной инфраструктуры отсутствует. Следует отметить, что, несмотря на существенную разницу в результатах прогноза ППР с учетом радоновых аномалий и без их учета, в обоих случаях прослеживаются в целом более высокие значения ППР на возвышенностях и низкие в долинах рек.

Таким образом, как показывает анализ результатов прогноза ППР с применением методов машинного обучения, в качестве основных факторов, способствующих формированию ее аномальных значений на территории Москвы, по-видимому, выступают геодинамически активные зоны, являющиеся зонами повышенной трещиноватости и проницаемости среды, а также наличие уклонов поверхности (перепада высот) и техногенного вибрационного воздействия железнодорожного транспорта. Возможно сочетание этих факторов, например, радоновые аномалии часто встречаются на участках, где железные дороги проходят в глубоких природных или искусственных понижениях (выемках), а также в областях пересечения геодинамически активных зон со склонами речных долин и оврагов. Данные факты свидетельствуют о вероятном формировании аномальных значений ППР в результате конвективного переноса радона в проницаемых зонах.

Качество и объем данных имеют решающее значение для применения алгоритмов машинного обучения. Это еще более важно для целей выявления радоновых аномалий, поскольку они обычно намного превышают фоновые значения выделения радона, а наличие лишь небольшого числа точек данных приводит к низкому качеству получаемых прогнозов. Это привело к появлению некоторых ложных линий на карте прогноза ППР, а также к необоснованному увеличению средней ППР по всей территории при прогнозе с учетом ее аномальных значений. Согласно карте на рис. 3б, большая часть южной части города будет классифицирована как радоноопасная с очень значительным превышением допустимых значений ППР ($80 \text{ мБк}/(\text{м}^2\text{с})$), что явно является завышенным прогнозом. В настоящее время работа над усовершенствованием алгоритмов прогноза ППР с использованием методов машинного обучения продолжается.

ВЫВОДЫ

Проведенные исследования показали, что методы машинного обучения могут с успехом применяться для составления карт потенциальной

радоноопасности городских территорий. Полученные карты плотности потока радона с поверхности грунта с учетом только фоновых значений ППР ($<400 \text{ мБк}/(\text{м}^2\text{с})$) показывают хорошую сходимость с картами, построенными с применением обычных методов интерполяции данных, но при этом более точно соответствуют геологическим и литологическим границам. Анализ значимости предикторов подтверждает, что наиболее важными факторами, обуславливающими фоновые значения плотности потока радона, являются:

- литологический тип грунта,
- концентрация в грунтах ^{226}Ra ,
- глубина залегания грунтовых вод.

Анализ полного набора данных, включающего как фоновые, так и аномальные значения ППР, свидетельствует, что в формировании аномальных радоновых полей существенную роль играют следующие факторы:

- близость к геодинамически активным зонам,
- близость к склонам крутизной более 8° ,
- близость к линиям железных дорог.

Влияние геодинамически активных зон, по-видимому, в значительной степени зависит от точного местоположения и сопутствующих геологических и ландшафтных факторов (тип грунта, наличие уклонов поверхности и т.п.). В целом, влияние техногенных факторов, за исключением линий железных дорог, на поток радона в условиях города не слишком велико. Возможна суперпозиция перечисленных выше факторов. Изучение механизмов влияния природных и техногенных факторов на аномальные выбросы радона — тема нашей дальнейшей работы.

Установленная связь радоновых аномалий с геодинамически активными зонами (участками повышенной проницаемости геологической среды) и склонами (перепадом высот) свидетельствует о возможном формировании аномалий ППР в результате конвективной циркуляции воздуха в проницаемых областях верхней части грунтовых массивов за счет перепада температур между грунтовым массивом и атмосферой [15]. Этот механизм может объяснить наличие контрастных радоновых аномалий на платформенных территориях, где отсутствует активная разрывная тектоника, распространены рыхлые четвертичные отложения большой мощности и близко к поверхности залегают подземные воды.

Следует отметить, что разработанные алгоритмы прогноза плотности потока радона еще далеки от совершенства и обладают рядом существенных недостатков, к которым можно отнести появление на картах ложных линий, связанных с техниче-

скими проблемами анализа данных, завышение средних прогнозных значений ППР при учете аномальных значений, появление на картах областей повышенных или пониженных значений ППР, которые явно не связаны с какими-то природными или техногенными факторами, а объясняются особенностями алгоритма.

Эти проблемы требуют решения в ходе дальнейших исследований, включая усовершенствование модели и ее верификацию. Однако, несмотря на все ограничения, полученная модель может быть полезна в качестве инструмента для выделения областей с потенциально высокими потоками радона и определения влияния определенных геологических и техногенных факторов на формирование плотности потока радона.

Работа выполнена при поддержке Российского научного фонда, грант № 24-17-00217.

СПИСОК ЛИТЕРАТУРЫ

1. Макаров В.И., Дорожко А.Л., Макарова Н.В., Макеев В.М. Современные геодинамически активные зоны платформ // Геоэкология. Инженерная геология, гидрогеология, геокриология. 2007. № 2. С. 99–110.
2. Макарова Н.В., Макеев В.И., Дорожко А.Л. и др. Геодинамические системы и геодинамически активные зоны Восточно-Европейской платформы // Бюл. Моск. общества испытателей природы. Отд. геол. 2016. Т. 91. Вып. 4–5. С. 9–22.
3. Маренный А.М., Цапалов А.А., Микляев П.С., Петрова Т.Б. Закономерности формирования радонового поля в геологической среде. М.: “Перо”, 2016. 394 с.
4. Микляев П.С., Макаров В.И., Дорожко А.Л. и др. Радоновое поле Москвы // Геоэкология. Инженерная геология, гидрогеология, геокриология. 2013. № 2. С. 172–187.
5. Осипов В.И. Крупномасштабное геологическое картирование территории г. Москвы // Геоэкология. Инженерная геология, гидрогеология, геокриология. 2011. № 3. С. 195–197.
6. Bossew P., Cinelli G., Ciotoli G.; Crowley Q.G. et al. Development of a Geogenic Radon Hazard Index — Concept, History, Experiences // Int. J. Environ. Res. Public Health. 2020. 17: 4134. <https://doi.org/10.3390/ijerph17114134>
7. Di Carlo C., Maiorana A., Bochicchio F. Indoor Radon: Sources, Transport Mechanisms and Influencing Parameters. 2023. IntechOpen. <https://doi.org/10.5772/intechopen.111710>
8. Friedman J.H. Multivariate adaptive regression splines (with discussion). The Annals of Statistics. 1991. 19:1–141.
9. Gavriliev S., Petrova T., Miklyaev P. Factors influencing radon transport in the soils of Moscow // Environ Sci Pollut Res. 2022, no. 29, pp. 88606–88617. <https://doi.org/10.1007/s11356-022-21919-y>
10. Gavriliev S., Petrova T., Miklyaev P., Karfidova, E. Predicting radon flux density from soil surface using machine learning and GIS data // Science of The Total Environment. 2023. 903:166348, <https://doi.org/10.1016/j.scitotenv.2023.166348>
11. ICRP. Radiological protection against radon exposure. 2014. No. 126. <https://www.icrp.org/publication.asp?id=ICRP%20Publication%20126>
12. Janik M., Bossew P., Kurihara O. Machine learning methods as a tool to analyse incomplete or irregularly sampled radon time series data // Science of The Total Environment. 2018. V. 630. P. 1155–1167.
13. Mair J., Petermann E., Lehné R., Henk A. Can neotectonic faults influence soil air radon levels in the Upper Rhine Graben? An exploratory machine learning assessment // Science of The Total Environment. 2024. 956:177179. <https://doi.org/10.1016/j.scitotenv.2024.177179>
14. Miklyaev P., Petrova T., Marennyy A, et al. High seasonal variations of the radon exhalation from soil surface in the fault zones (Baikal and North Caucasus regions) // Journal of Environmental Radioactivity. 2020. 219, 106271. <https://doi.org/10.1016/j.jenvrad.2020.106271>
15. Miklyaev P.S., Petrova T.B., Shchitov D.V., Sidiyakin P.A. et al. Radon transport in permeable geological environments // Sci. Total Environ. 2022. V. 852:158382. <https://doi.org/10.1016/j.scitotenv.2022.158382>
16. Nazaroff W.W. Radon transport from soil to air // Reviews of Geophysics. 1992. 30(2):137. <https://doi.org/10.1029/92rg00055>
17. Osipov V.I., Burova V.N., Zaikanov V.G., Molodykh I.I. et al. A map of large-scale (detail) engineering geological zoning of Moscow territory // Water Resources. 2012. 39(7):737–746. <https://doi.org/10.1134/S0097807812070093>
18. Petermann E., Bossew P., Kemski J. et al. Development of a high-resolution indoor radon map using a new machine learning- based probabilistic model and German radon survey data // Environ. Health Perspect. 2024.132 (9):97009. <https://doi.org/10.1289/EHP14171>
19. Petermann E., Meyer H., Nussbaum M., Bossew P. Mapping the geogenic radon potential for Germany by Machine Learning. 2020. <https://doi.org/10.5194/egusphere-egu2020-8501>
20. Rezaie F., Panahi M., Bateni S. M., Kim S. et al. Spatial modeling of geogenic indoor radon distribution in Chungcheongnam-do, South Korea using enhanced machine learning algorithms // Environment International. 2023. 171:107724. <https://doi.org/10.1016/j.envint.2022.107724>
21. Timkova J., Fojtikova I., Pacheroova P. Bagged neural network model for prediction of the mean indoor radon concentration in the municipalities in Czech Republic // Journal of Environmental Radioactivity. 2017. 166:398–402. <https://doi.org/10.1016/j.jenvrad.2016.07.008>
22. Torkar D., Zmazek B., Vaupotič J., Kobal I. Application of artificial neural networks in simulating radon levels in soil gas // Chemical Geology. 2010. 270(1–4):1–8. <https://doi.org/10.1016/j.chemgeo.2009.09.017>
23. Tsapalov A., Kovler K., Miklyaev P. Open charcoal chamber method for mass measurements of radon exhalation rate from

- soil surface // Journal of Environmental Radioactivity. 2016. 160:28–35.
<https://doi.org/10.1016/j.jenvrad.2016.04.016>
24. UNSCEAR. Sources and effects of ionizing radiation. 2000. No. 1.
https://www.unscear.org/unscear/en/publications/2000_1.html. Accessed 15 Apr 2022
25. WHO Handbook on indoor radon. WHO Handbook on Indoor Radon: a public health perspective. Hajo Zeeb and Ferid Shannoun (eds), Geneva, WHO Press, 2009.

ANALYSIS OF THE INFLUENCE OF NATURAL AND ANTHROPOGENIC FACTORS ON RADON FLUX DENSITY IN MOSCOW USING MACHINE LEARNING METHODS

S. G. Gavriliev^a, T. B. Petrova^{b,*,#}, P. S. Miklyaev^{a,##}, E. A. Karfidova^a

^a*Sergeev Institute of Environmental Geoscience, Russian Academy of Sciences, Ulanskii per. 13, bld.2, Moscow, 101000 Russia*

^b*Department of Radiochemistry, Faculty of Chemistry, Lomonosov Moscow State University, Leninskie Gory 1, Moscow, 119991 Russia*

*E-mail: thp111@inbox.ru

##E-mail: peterm7@inbox.ru

During routine measurements of radon flux density for construction purposes in Moscow, the areas were found with abnormally high radon flux densities exceeding 400 mBq m⁻² s⁻¹. These values far exceed those average values around 24–40 mBq m⁻² s⁻¹ for local sandy and clay soils, which is compounded with the fact that the area under study doesn't contain any active faults or natural soils rich in uranium. Therefore, the question arises, whether these high values are of technogenic or natural origin. This paper uses machine learning algorithms to find the answer to these questions. Machine learning algorithms including random forest trees and artificial neural networks were used to try and predict radon flux density anomalies on a city scale. Predictors used included maps of geodynamically active areas, lineaments, distances to heavy rail infrastructure such as metro tunnels and surface-level rail. Additionally, normal predictors of radon such as ²²⁶Ra concentration in soil, quaternary soil type and elevation were used for the predictions. Predictions were made for both anomaly-free and anomaly included datasets. Training data included radon flux data for Moscow with both anomalous and background values which included 931 data points, of which 112 was classified as anomalous (more than 400 mBq m⁻² s⁻¹). Based on the predictions obtained, factors which influence radon flux density and those that may produce anomalous values were underlined.

Keywords: radon, machine learning, Random Forest, radon flux density, potential radon hazard, forecast

REFERENCES

1. Makarov, V.I., Dorozhko, A.L., Makarova, N.V., Makeev, V.M. [Modern geodynamically active zones of platforms.] *Geoekologiya*, 2007, no. 2, pp. 99–110. (in Russian)
2. Makarova, N.V., Makeev, V.I., Dorozhko, A.L., Sukhanova, T.V., Korobova, I.V. [Geodynamic systems and geodynamically active zones of the East-European platform]. *Byulleten MOIP, ser. Geol*, 2016, vol. 91, pp. 4–5. (in Russian)
3. Marennyy, A.M., Tsapalov, A.A., Miklyaev, P.S., Petrova T.B. [Regularities in radon field formation in geoenvironment]. Moscow, Pero Publ., 2016, 394 p. (in Russian)
4. Miklyaev, P.S., Makarov, V.I., Dorozhko, A.L., Petrova, T.B., Marennyi, M.A., Marennyi, A.M., Makeev V.M. [Radon field of Moscow]. *Geoekologiya*, 2013, no. 2, pp. 172–187. (in Russian)
5. Osipov, V.I. [Large-scale geological mapping of Moscow territory]. *Geoekologiya*, 2011, no. 3, pp. 195–197. (in Russian)
6. Bossew, P., Cinelli, G., Ciotoli, G., Crowley, Q.G., De Cort, M., Elío Medina, J., Gruber, V., Petermann, E., Tollefsen, T. Development of a geogenic radon hazard index—concept, history, experiences. *Int. J. Environ. Res. Public Health*, 2020, no. 17, p. 4134.
<https://doi.org/10.3390/ijerph17114134>
7. Di Carlo, C., Maiorana, A., Bochicchio, F. Indoor radon: sources, transport mechanisms and influencing parameters. *IntechOpen*, 2023. doi: 10.5772/intechopen.111710
8. Friedman, J. H. Multivariate adaptive regression splines (with discussion). *The Annals of Statistics*, 1991, no. 19, pp. 1–141.
9. Gavriliev, S., Petrova, T., Miklyaev, P. Factors influencing radon transport in the soils of Moscow. *Environ. Sci. Pollut. Res.*, 2022, no. 29, pp. 88606–88617.
<https://doi.org/10.1007/s11356-022-21919-y>

10. Gavriliev, S., Petrova, T., Miklyaev, P., Karfidova, E. Predicting radon flux density from soil surface using machine learning and GIS data. *Science of the Total Environment*, 2023, no. 903, p. 166348, <https://doi.org/10.1016/j.scitotenv.2023.166348>
11. Radiological protection against radon exposure. ICRP Publication no. 126, 2014. <https://www.icrp.org/publication.asp?id=ICRP%20Publication%20126>
12. Janik, M., Bossew, P., Kurihara, O.. Machine learning methods as a tool to analyze incomplete or irregularly sampled radon time series data. *Science of the Total Environment*, 2018, no. 630, pp.1155–1167.
13. Mair, J., Petermann, E., Lehné, R., Henk, A. Can neotectonic faults influence soil air radon levels in the Upper Rhine Graben? An exploratory machine learning assessment. *Science of the Total Environment*, 2024, no. 956, p. 177179, <https://doi.org/10.1016/j.scitotenv.2024.177179>
14. Miklyaev, P., Petrova, T., Marennyy, A., Shchitov, D., Sidiyakin, P., Murzabekov, M., Lopatin, M. High seasonal variations of the radon exhalation from soil surface in the fault zones (Baikal and North Caucasus regions). *Journal of Environmental Radioactivity*, 2020, no. 219, pp. 106271.
15. Miklyaev, P.S., Petrova, T.B., Shchitov, D.V., Sidiyakin, P.A., Murzabekov, M.A., Tsebro, D.N., Marennyy, A.M., Nefedov, N.A., Gavriliev, S.G. Radon transport in permeable geological environments. *Sci. Total Environ.*, 2022, vol. 852, p. 158382. DOI: 10.1016/j.scitotenv.2022.158382.
16. Nazaroff, W.W. Radon transport from soil to air. *Reviews of Geophysics*, 1992. vol. 30, issue 2, pp.137–160. <https://doi.org/10.1029/92rg00055>
17. Osipov, V.I., Burova, V.N., Zaikanov, V.G., Molodykh, I.I., Pyrchenko, V.A., Savits'ko, I.S. A map of large-scale (detail) engineering geological zoning of Moscow territory, *Water Resources*, 2012, no. 39(7), pp. 737–746. DOI: 10.1134/S0097807812070093.
18. Petermann, E., Bossew, P., Kemski, J., Gruber, V., Suhr, N., Hoffmann, B., Development of a high-resolution indoor radon map using a new machine learning- based probabilistic model and German radon survey data. *Environ. Health Perspect.*, 2024, no.132 (9), p. 97009. <https://doi.org/10.1289/EHP14171>
19. Petermann, E., Meyer, H., Nussbaum, M., & Bossew, P. Mapping the geogenic radon potential for Germany by machine learning. 2020. <https://doi.org/10.5194/egusphere-egu2020-8501>
20. Rezaie, F., Panahi, M., Bateni, S. M., Kim, S., Lee, J., Lee, J., Yoo, J., Kim, H., Won Kim, S., & Lee, S.. Spatial modeling of geogenic indoor radon distribution in Chungcheongnam-do, South Korea using enhanced machine learning algorithms. *Environment International*, 2023, no. 171, p. 107724. <https://doi.org/10.1016/j.envint.2022.107724>
21. Timkova, J., Fojtikova, I., Pacherova, P. Bagged neural network model for prediction of the mean indoor radon concentration in the municipalities in Czech Republic. *Journal of Environmental Radioactivity*, 2017, no. 166, pp. 398–402. <https://doi.org/10.1016/j.jenvrad.2016.07.008>
22. Torkar, D., Zmazek, B., Vaupotič, J., Kobal, I. Application of artificial neural networks in simulating radon levels in soil gas. *Chemical Geology*, 2010, vol. 270, issues 1–4, pp.1–8. <https://doi.org/10.1016/j.chemgeo.2009.09.017>
23. Tsapalov, A., Kovler, K., Miklyaev, P. Open charcoal chamber method for mass measurements of radon exhalation rate from soil surface. *Journal of Environmental Radioactivity*, 2016, no.160, pp. 28–35. <https://doi.org/10.1016/j.jenvrad.2016.04.016>
24. Sources and effects of ionizing radiation. UNSCEAR, 2000, no.1. https://www.unscear.org/unscear/en/publications/2000_1.html. Accessed April 15, 2022 <https://doi.org/10.1016/j.jenvrad.2020.106271>
25. WHO Handbook on Indoor Radon: a public health perspective. Hajo Zeeb and Ferid Shannoun, Eds., Geneva, WHO Press, 2009