

УДК 004.65

В. А. Усольцев¹, В. П. Часовских², Е. Н. Стариков²¹ Ботанический сад УрО РАН, г. Екатеринбург, Российская Федерация² Уральский государственный экономический университет, г. Екатеринбург, Российская Федерация

Исследование методов и обработка баз данных о биомассе лесов Евразии как нейронных сетей. Часть 1. Системный анализ базы данных для трансформации в нейронные сети искусственного интеллекта

Аннотация. Оценка углерододепонирующей способности лесов выходит в настоящее время на глобальный уровень, актуальной является оценка поглощения парниковых газов на карбоновых полигонах. Авторами разработаны и опубликованы три авторские базы данных о биологической продуктивности лесов Евразии. Разработан алгоритм оценки приходной части углеродного цикла (чистой первичной продукции) лесного покрова Уральского региона, образующий основу генерации нейронной сети из начальных баз данных. Дальнейшему развитию моделей и алгоритмов применения названных баз данных, трансформированных в нейронные сети для решения актуальных глобальных проблем современности, связанных с изменением климата, посвящена настоящая работа.

Ключевые слова: база данных; биомасса; углерододепонирование; экология; нейронная сеть.

Финансирование: Работа выполнена при финансовой поддержке Министерства науки и высшего образования РФ, проект FEUZ-2021-0014.

Дата поступления статьи: 4 июня 2022 г.

Для цитирования: Усольцев В. А., Часовских В. П., Стариков Е. Н. Исследование методов и обработка баз данных о биомассе лесов Евразии как нейронных сетей. Часть 1. Системный анализ базы данных для трансформации в нейронные сети искусственного интеллекта // Цифровые модели и решения. 2022. Т. 1, № 1. DOI: 10.29141/2782-4934-2022-1-1-2. EDN: DNJRGT.

V. A. Usoltsev¹, V. P. Chasovskikh², E. N. Starikov²¹Botanical Garden of the Ural Branch of the Russian Academy of Sciences, Yekaterinburg, Russia²Ural State University of Economics, Yekaterinburg, Russia

Research of methods and processing of databases on the biomass of Eurasian forests as neural networks. Part 1. System analysis of a database for transformation into neural networks of artificial intelligence

Abstract. The assessment of the carbon storage capacity of forests is now reaching the global level, and the assessment of greenhouse gas absorption at carbon landfills is relevant. The authors have developed and published three author's databases on the biological productivity of Eurasian forests. An algorithm for estimating the incoming part of the carbon cycle (NWP) of the forest cover of the Ural region has been developed, which forms the basis for generating a neural network from initial databases. The present work is devoted to the further development of models and algorithms for the use of these databases, transformed into neural networks to solve urgent global problems of our time related to climate change.

Key words: database; biomass; carbon deposition; ecology; neural network.

Funding: The work was supported by the Ministry of Science and Higher Education of the Russian Federation, project FEUZ-2021-0014.

Paper submitted: June 4, 2022

For citation: Usoltsev V. A., Chasovskikh V. P., Starikov E. N. Research of methods and processing of databases on the biomass of Eurasian forests as neural networks. Part 1. System analysis of a database for transformation into neural networks of artificial intelligence. Digital models and solutions. 2022. Vol. 1, no. 1. DOI: 10.29141/2782-4934-2022-1-1-2. EDN: DNJRGT.

Введение

Современная климатическая проблема отличается особенно сложным многокомпонентным характером в сравнении с другими глобальными экологическими проблемами, она в той или иной мере оказывает влияние на все сферы деятельности человеческой цивилизации, а также изменения в структуре и среде обитания биоты и ее основного наземного компонента – лесного покрова. Хотя считается, что древние цивилизации вымерли в основном из-за вырубки лесов, в действительности опустынивание районов, некогда покрытых лесом, было результатом изменения климата. Недавно 11 258 ученых из 153 стран в своем докладе заявили, что планета Земля находится в чрезвычайной

климатической ситуации, и существует моральное обязательство предупредить человечество о реальной катастрофической угрозе. Необходимы международные усилия для предотвращения глобального потепления и снижения выбросов CO₂, поскольку изменение климата создает огромные проблемы для человечества и биоты. При этом растет осведомленность о потенциальной роли лесов в смягчении последствий изменения климата и адаптации к нему, что делает еще более насущным задачу совершенствования оценки и мониторинга лесов [1, с. 10].

В настоящее время разработана и успешно осуществляется международная программа UN-REDD, инициированная FAO, Программой ООН по окружающей среде (UNEP) и другими международными организациями, а также ее продвинутая версия REDD+ (Reducing Emissions from Deforestation and Forest Degradation), цель которой – минимизация последствий изменения климата, вызванного деградацией лесов, путем снижения эмиссий, повышения способности лесного покрова к депонированию и консервации атмосферного углерода, а также увеличения биоразнообразия и повышения экологических функций лесов [2, с. 138].

В вопросах стабилизации климата большие надежды связаны с углерододепонирующей способностью лесов и разработкой и испытаниями технологий контроля баланса климатических активных газов, прежде всего CO₂, в природных экосистемах (карбоновые полигоны). Оценка биологической продуктивности, или углерододепонирующей способности лесов, выходит в настоящее время на глобальный уровень, и ее повышение является одним из основных факторов стабилизации климата. Климатическая система планеты уже утратила стационарность, в результате чего климатические эксцессы все учащаются, а климатические сценарии становятся непредсказуемыми. Различные сценарии глобальной динамической модели растительности показывают существенную неопределенность в отношении будущего накопления углерода в наземной биоте. Согласно прогнозам, изменения углеродного пула к концу нынешнего столетия варьируют от минус 106 до плюс 201 Пг (млрд т). Таким образом, достоверно неизвестен даже знак изменения углеродного пула на планете, и мы до конца не знаем, является ли биота планеты источником или накопителем углерода [3, с. 101].

Проблеме углерододепонирующей способности лесов посвящено множество публикаций в мировой литературе. Поскольку эта проблема имеет глобальный характер, для ее решения необходима разработка прогностических моделей климатически обусловленного изменения биомассы на основе соответствующих мировых баз данных о биологической продуктивности лесов, и научное сообщество заявляет о наступлении «эры больших данных». Сегодня формируются не просто базы данных, а базы баз данных [4, с. 142] как второе поколение объединения данных. Выполненный авторами системный анализ природных баз данных (TRY) показал их свойство как технологию Big Data, что определило целесообразность перехода к искусственным семантическим сетям искусственного интеллекта. Однако полный потенциал Big Data ограничен проблемами в охвате данных, заполнение которых стало непростой задачей, поскольку они касаются 50 % биологических видов во всем мире [5, с. 63]. В последние годы разра-

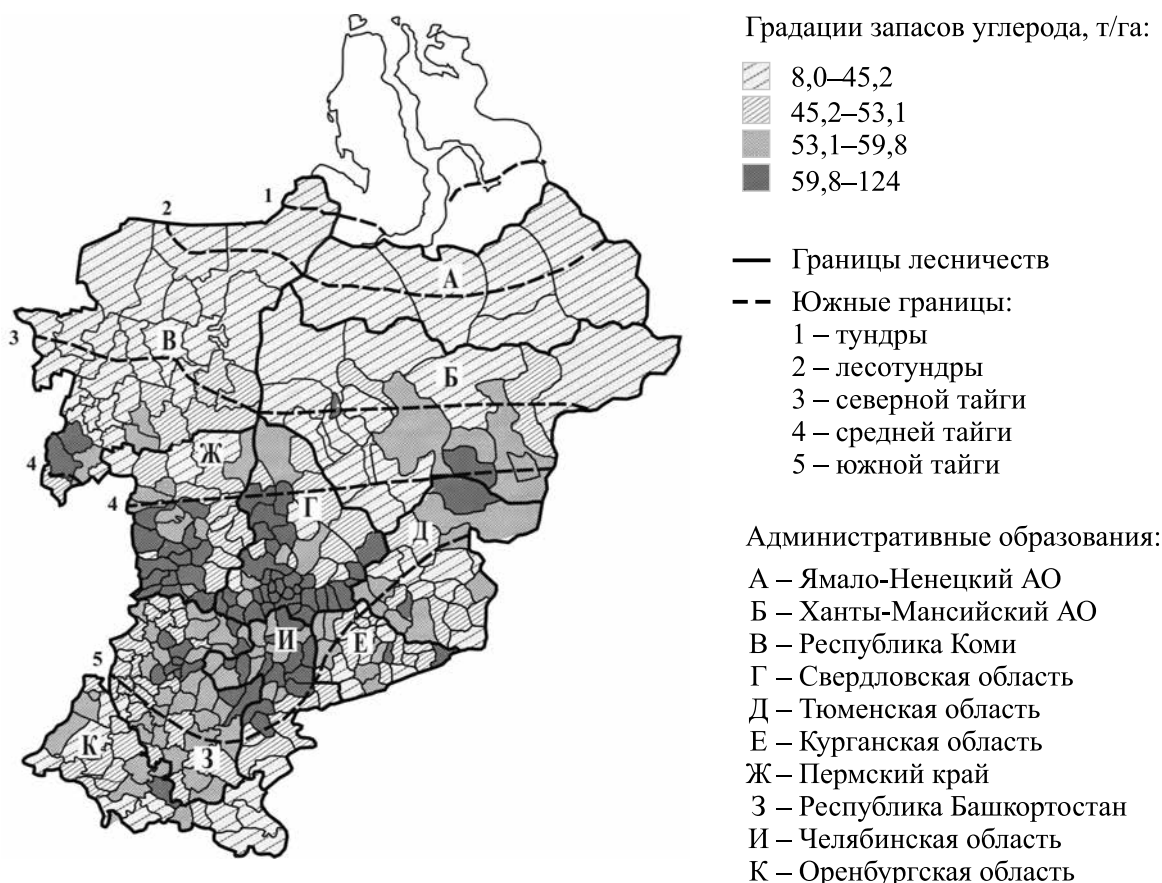
ботаны и опубликованы три базы данных о биологической продуктивности лесов Евразии. Одна из них включает информацию о биомассе и чистой первичной продукции (ЧПП) лесобразующих древесных видов [6, с. 38], вторая – информацию о биомассе деревьев лесобразующих видов, третья – о квалитетических показателях (плотности и содержании сухого вещества в компонентах биомассы деревьев) лесобразующих видов. Методике, модели, алгоритму применения первой из названных баз данных, трансформированной в нейронные сети для решения актуальных глобальных проблем современности, связанных с изменением климата, посвящена настоящая работа.

Методы, алгоритмы и результаты исследований

Появление первых компьютеров в нашей стране означало качественный информационный скачок с уровня деревянных костяшек-счетов на уровень, казалось бы, безграничных возможностей моделирования исследуемых процессов. Спустя некоторое время обнаружилось, что оптимизм оказался преждевременным, возлагаемые на компьютеры надежды не оправдываются, и для работы с компьютером нужны программы и их создатели, т. е. программисты. Спустя еще некоторое время обнаружилось, что программистам нужны алгоритмы, которые могут дать только специалисты, знающие природу исследуемого процесса. Конструирование адекватных алгоритмов сегодня представляет собой одну из основных проблем при моделировании сложных систем. Важность построения корректного алгоритма модели покажем на примере оценки приходной части углеродного цикла лесных экосистем с использованием первой из упомянутых баз данных.

Эта база данных о биомассе и ЧПП [6, с. 52] включает два раздела: в первом содержатся данные только о биомассе, а во втором – о биомассе и ЧПП. Поскольку получить фактические данные о ЧПП многократно сложнее, чем о биомассе, то объем данных второго раздела на порядок меньше, чем первого. В ходе расчетов приходной части углеродного цикла (ЧПП) было применено два разных алгоритма регрессионных моделей одной общей идеологии. Согласно одному из них [7, с. 60], регрессионные модели ЧПП компонентов древостоя (стволов, ветвей, хвои, корней) сопрягались по рекурсивному принципу напрямую с данными Государственного учета лесного фонда (ГУЛФ) по схеме: ЧПП → ГУЛФ. При разработке другого алгоритма авторами данной статьи был принят во внимание тот факт, что данных о ЧПП на порядок меньше, чем данных о биомассе. Поэтому в предыдущий алгоритм было включено промежуточное звено (модель биомассы) и построена цепочка последовательных рекурсивных сопряжений: ЧПП → биомасса → ГУЛФ. Алгоритм был реализован в среде Natural и СУБД ADABAS (см. рисунок).

При оценке приходной части углеродного цикла (ЧПП) лесного покрова Уральского региона по двум названным алгоритмам оказалось, что первый из них дает 2–3-кратное завышение [8, с. 83]. Тот же показатель ЧПП был рассчитан по двум независимым альтернативным методам: в основе первого лежала оценка по хлорофилльному индексу [9, с. 394], а базой для второго были данные не ГУЛФ, а таблицы хода роста древосто-



*Распределение годовичного депонирования углерода
в биомассе лесов в расчете на лесопокрытую площадь согласно ГУЛФ [8, с. 71]*

ев [10, с. 24]. Корректность нашего (второго из названных) алгоритма подтверждается тем, что результаты его реализации отличаются от результатов, полученных по альтернативным методам [9, с. 395]), всего на 5–8%. Таким образом, корректные алгоритмы альтернативных методов могут дать близкие результаты, а некорректный алгоритм может обусловить существенное смещение результата по отношению к модели той же идеологии, но построенной по корректному алгоритму.

Было доказано, что точность результатов повышается, если базы данных реализуются как Big Data. В качестве источников Big Data было предложено использовать научные материалы для космических многозональных съемок сверхвысокого разрешения, разработанных в НИР РФФИ-06-01-08082 (научный руководитель В. Г. Лабунец), а базу данных рассматривать как инфраструктуру центра обработки данных через конфигурационные файлы (Infrastructure as Code, IaC). Предложенные методы дистанционного зондирования основываются на математической теории гиперкомплексных инвариантов, быстрых преобразованиях Фурье – Клиффорда и агрегационных операторах [11, с. 13; 12, с. 5]. Предложено в алгебро-геометрическом подходе каждый гиперспектральный пиксель рассматривать не как K -мерный вектор, а как K -мерное гиперкомплексное число (где K – число спектральных каналов). Изменения в окружающем мире, вызванные гиперцветовыми и геометрическими искажениями наблюдаемых объектов, трактуются не как матричные преобразования физического и гиперцветового про-

странств, а как действие некоторых чисел Клиффорда (спиноров) в этих пространствах. Было доказано, что аппарат гиперкомплексных алгебр Клиффорда более адекватно описывает процессы обработки и распознавания гиперспектральных изображений, чем векторно-матричный математический аппарат, и что на его основе можно разрабатывать более эффективные алгоритмы цифровой обработки и распознавания гиперспектральных изображений, что и было сделано применительно к выбранным аэрокосмическим снимкам лесных секторов. Появилась возможность создавать искусственные нейронные сети искусственного интеллекта для обработки сверхбольших массивов данных аэрокосмического мониторинга и рационального использования и сохранения лесов.

Изменение климата приводит к снижению биологической продуктивности лесов, что, в свою очередь, может повлиять на биосферную функцию лесного покрова и его углерододепонирующую способность. Таким образом, современная наука сталкивается с необходимостью одновременного анализа двух взаимосвязанных процессов, но нет ответа на вопрос, какой из них первичен. Изменения в биоте на региональном и биомном уровнях оцениваются обычно по ширине годичных колец, связанной с метеорологическими показателями региональных метеостанций. В подобных случаях моделируется прямая связь (воздействие – реакция), например, летней температуры с шириной годичного кольца соответствующего года [13, с. 60].

Применительно к биомассе дерева подобный подход не является корректным, поскольку его компоненты (фракции) формируются в течение периодов разной длительности: ствол – 50–100 лет, ветви – 30–60 лет, хвоя – 5–8 лет у хвойных и в течение 1 года у лиственных древесных видов. Каждое дерево и каждый древостой при прочих равных условиях характеризуются специфической структурой биомассы, т. е. соотношением биомассы стволов, ветвей, хвои, корней, которая формировалась в данном экотопе под влиянием температур и осадков в течение тысячелетий. Эти территориально распределенные показатели температур и осадков также формировались в течение тысячелетий, и поэтому логичным является использование алгоритма, позволяющего сопрягать структуру биомассы деревьев и древостоев не с метеорологическими показателями метеостанций, а с фактическими территориально распределенными показателями температур и осадков на уровне континента, усредненными за десятки или сотни лет. Но это возможно только при наличии фактических данных о биомассе на территории не региона, не биома, а всего континента. И такая возможность появилась после публикации наших трансконтинентальных баз данных, в нашем случае – баз данных о биомассе и ЧПП древостоев, о биомассе деревьев и о квалиметрии деревьев того или иного вида (рода), покрывающих территорию Евразии.

Системы управления базами данных сами по себе не представляют ценности для пользователей. Даже данные, хранящиеся в базах, не представляют особой ценности сами по себе. Основной ценностью обладают законченные приложения, дающие пользователям возможность моделировать некоторые аспекты своей деятельности с использованием вычислительной техники.

Мы считаем, что Big Data – это не только база данных, а стек технологий, который содержит и интеллектуальный анализ данных, алгоритмы.

По нашему мнению, необходимость перехода к системам с искусственным интеллектом требует отказаться от догмы о превосходстве реляционной модели данных и сосредоточить основные усилия в исследованиях и разработках альтернативных моделей. Сетевая структура баз данных, трансформируемая в реляционные представления СУБД ADABAS [8, с. 71; 14, с. 22; 15, с. 326], позволяет создавать и обрабатывать (в среде Natural) объектно-ориентированные базы данных TRY. Создаваемые объектно-ориентированные базы данных TRY плотно взаимодействуют с объектно-ориентированными языками программирования Python, C# в среде ASP.NET Core MVC 5. Использование фреймворка Visual Studio 2022 обеспечивает доступ к технологиям больших данных и машинного обучения SQL server 2019.

Заключение

Подводя итог проведенному исследованию, сформулируем наиболее значимые выводы. Во-первых, определено, что проблема углерододепонирующей способности лесов имеет глобальный характер, для ее решения необходима разработка прогностических моделей климатически обусловленного изменения биомассы на основе соответствующих мировых баз данных о биологической продуктивности лесов.

Во-вторых, доказано, что точность результатов повышается, если базы данных реализуются как база баз данных. Было предложено трансформировать их в видеопоток Big Data, рассматриваемый как инфраструктура центра обработки данных через конфигурационные файлы (Infrastructure as Code, IaC).

В-третьих, для видеопотока (Big Data) дистанционного зондирования лесных массивов предложено существенно повысить точность определения депонированного углерода за счет применения при распознавании данных математической теории гиперкомплексных инвариантов, быстрых преобразований Фурье – Клиффорда и агрегационных операторов [11, с. 13; 12, с. 5]. Эффективность данного направления была доказана на обработке реальных лесных баз данных.

В-четвертых, эффективность и точность прогнозных моделей депонируемого углерода в центре обработки данных через конфигурационные файлы достигается за счет применения нейронных сетей технологии искусственного интеллекта.

Представленный в статье подход становится важным как в аналитическом, так и в управленческом смысле для повышения точности расчета карбонового следа от деятельности человека при построении национальной системы мониторинга климатически активных газов.

Источники

1. Ripple W. J., Wolf C., Newsome T. M. et al. World scientists' warning of a climate emergency // *BioScience*. 2020. Vol. 70, iss. 1. P. 8–12. DOI 10.1093/biosci/biz088.
2. Knowles T., McCall M., Skutsch M., Theron L. Preparing community forestry for REDD+: engaging local communities in the mapping and MRV requirements of REDD+ // *Pathways for Implementing REDD+: experiences from carbon markets and communities* / ed. by X. Zhu et al. Roskilde: UNEO, 2010. P. 141–151.
3. Schaphoff S., Lucht W., Gerten D. et al. Terrestrial biosphere carbon storage under alternative climate projections // *Climatic change*. 2006. Vol. 74, iss. 1. P. 97–122. DOI 10.1007/s10584-005-9002-5.
4. Kattge J., Bönisch G., Díaz S. et al. TRY plant trait database – enhanced coverage and open access // *Global change biology*. 2020. Vol. 26, iss. 1. P. 119–188. DOI 10.1111/gcb.14904.
5. Tobias J. A. A bird in the hand: global-scale morphological trait datasets open new frontiers of ecology, evolution and ecosystem science // *Ecology letters*. 2022. Vol. 25, iss. 3. P. 573–580. DOI 10.1111/ele.13960.
6. Usoltsev V. A. Forest biomass and primary production database for Eurasia: digital version. The third edition, enlarged. Yekaterinburg: Ural State Forest Engineering University, 2020. URL: https://elar.usfeu.ru/bitstream/123456789/9648/1/Base_v2.xlsx (дата обращения: 03.06.2022).
7. Замолодчиков Д. Г., Уткин А. И. Система конверсионных отношений для расчета чистой первичной продукции лесных экосистем по запасам насаждений // *Лесоведение*. 2000. № 6. С. 54–63.
8. Воронов М. П., Усольцев В. А., Часовских В. П. Исследование методов и разработка информационной системы определения и картирования депонируемого лесами углерода в среде Natural. 2-е изд., испр. и доп. Екатеринбург: Урал. гос. лесотехн. ун-т, 2015. 192 с.
9. Воронин П. Ю., Коновалов П. В., Блондинский В. К., Кайбияйнен Л. К. Хлорофильный индекс и фотосинтетический сток углерода в лесах Северной Евразии // *Физиология растений*. 2004. № 51. С. 390–395.
10. Швиденко А. З., Щепашенко Д. Г., Нильссон С. Материалы к познанию современной продуктивности лесов России // *Базовые проблемы перехода к устойчивому управлению лесами России – учет лесов и организация лесного хозяйства: материалы международного семинара*. Красноярск: Институт леса им. В. Н. Сукачева СО РАН, 2007. С. 5–37.
11. Labunets V. G., Chasovskikh V. P., Smetanin J. G., Ostheimer E. New multi-parameter Golay 2-complementary sequences and transforms // *Journal of Physics: Conference Series*. 2018. Vol. 1096, iss. 1. Art 12043. DOI 10.1088/1742-6596/1096/1/012043.
12. Labunets V. G., Martyugin S. A., Chasovskikh V. P. et al. Intelligent OFDM telecommunication system. Part 4. Anti-eavesdropping and anti-jamming properties of the system, based on many-parameter and fractional Fourier transforms // *Journal of Physics: Conference Series*. 2019. Vol. 1368, iss. 5. Art. 52024. DOI 10.1088/1742-6596/1368/5/052024.

13. Глебов Ф. З., Литвиненко В. И. Динамика ширины годичных колец в связи с метеорологическими показателями в различных типах болотных лесов // Лесоведение. 1976. № 4. С. 56–62.
14. Кузнецов С. Д. Объектно-ориентированные базы данных – основные концепции, организация и управление: краткий обзор. URL: http://citforum.ru/database/articles/art_24.shtml (дата обращения: 03.06.2022).
15. Часовских В. П., Воронов М. П., Фатеркин А. С. Информационные технологии управления: СУБД ADABAS и проектирование приложений средствами Natural. Екатеринбург: Урал. гос. лесотехн. ун-т, 2016. 477 с.

Информация об авторах

Усольцев Владимир Андреевич, доктор сельскохозяйственных наук, профессор, ведущий научный сотрудник. Ботанический сад УрО РАН; 620144, РФ, г. Екатеринбург, ул. 8 Марта, 202а. E-mail: usoltsev50@mail.ru.

Часовских Виктор Петрович, доктор технических наук, профессор, профессор кафедры информационных технологий и статистики. Уральский государственный экономический университет, 620144, РФ, г. Екатеринбург, ул. 8 Марта/Народной Воли, 62/45. E-mail: u2007u@ya.ru

Стариков Евгений Николаевич, кандидат экономических наук, доцент, доцент кафедры информационных технологий и статистики. Уральский государственный экономический университет, 620144, РФ, г. Екатеринбург, ул. 8 Марта/Народной Воли, 62/45. E-mail: starikov_en@usue.ru

Information about the authors

Vladimir A. Usoltsev, doctor of agricultural sciences, professor, leading researcher. Botanical Garden of the Ural Branch of the Russian Academy of Sciences; 620144, Russia, Yekaterinburg, 8 Marta Str., 202a. E-mail: usoltsev50@mail.ru

Victor P. Chasovskikh, doctor of technical sciences, professor, professor of the department of information technologies and statistics. Ural State University of Economics, 620144, Russia, Yekaterinburg, 8 Marta/Narodnoy Voli Str., 62/45. E-mail: u2007u@ya.ru

Evgenii N. Starikov, candidate of economic sciences, associate professor, associate professor of the department of information technologies and statistics. Ural State University of Economics, 620144, Russia, Yekaterinburg, 8 Marta/Narodnoy Voli Str., 62/45. E-mail: starikov_en@usue.ru