

УДК 528.854.4

DOI:10.38028/ESI.2022.26.2.011

Подход к построению углеродной карты Ханты-Мансийского автономного округа на основе оценки биомассы с помощью данных дистанционного зондирования Земли

Бредихин Арсентий Игоревич¹, Мельников Андрей Витальевич²

¹Югорский государственный университет,

²Югорский научно-исследовательский институт информационных технологий,

Россия, г. Ханты-Мансийск, *a_bredihin@ugrasu.ru*

Аннотация. Ханты-Мансийский автономный округ-Югра обладает большой площадью лесных территорий (более 90% от всей площади округа). А лесная растительность естественным образом рано или поздно отмирает, вследствие чего из органического вещества происходит выделение углекислого газа в атмосферу и усиление глобального потепления.

Для недопущения повышения глобальной температуры необходимо вести учет запасов углерода. Помимо выбросов от сжигания топлива, диоксид углерода (CO₂), который является парниковым газом, образуется при разложении мертвой растительности, поскольку растения накапливают и хранят в себе углерод. При этом Ханты-Мансийский автономный округ-Югра (ХМАО-Югра) обладает большой площадью лесов: по состоянию на 2020 год она составляет 503990 км² при общей площади территории округа в 534,8 тыс. км² [1, с. 608-610], что составляет 94,2% площади всего округа. Поэтому для учета запаса углерода в рамках округа необходимо оценить данный запас в виде количества растительной (лесной) биомассы.

Один из способов оценки растительной биомассы – создание углеродных карт с применением методов дистанционного зондирования Земли (ДЗЗ). Применение полученных с помощью методов ДЗЗ спутниковых снимков и методов их обработки позволит получить карту округа с полным охватом всей территории, а применение моделей машинного обучения позволит разработать модель, с помощью которой будет возможно получать углеродную карту округа с заданной точностью.

В данной работе приведен обзор решений зарубежных ученых за последние 5 лет в области ДЗЗ, направленных на создание углеродных карт. На основании данного обзора предложена программа исследований, которая позволит разработать подход, позволяющий получать цифровую углеродную карту ХМАО с необходимой точностью.

Ключевые слова: дистанционное зондирование Земли, спутниковый снимок, углерод, растительная биомасса, машинное обучение

Цитирование: Бредихин А.И. Подход к построению углеродной карты Ханты-Мансийского автономного округа на основе оценки биомассы с помощью данных дистанционного зондирования Земли / А.И. Бредихин, А.В. Мельников // Информационные и математические технологии в науке и управлении. – 2022. – № 2(26). – С. 115-126. – DOI:10.38028/ESI.2022.26.2.011.

Введение. Ханты-Мансийский автономный округ-Югра является крупным нефте- и газодобывающим российским регионом. Он характеризуется большими объемами выбросов парниковых газов в атмосферу. Так, на территории округа имеется большое количество болот, которые являются источником метана (CH₄), а нефте- и газодобыча является источником диоксида углерода (CO₂). Кроме того, углерод содержится и в растительности (в т.ч. лесной).

Одним из способов оценки лесной биомассы округа является применение спутниковых изображений оптического и инфракрасного диапазонов совместно с лидарными снимками. Здесь примечательным является подход к решению данной задачи американскими учеными [2], который основан на применении полевых измерений с совместными данными НАСА. Суть подхода заключается в последовательном получении полевых измерений – массы деревьев на определенных участках, определении согласованности данных с сигналами лидара, обучении модели «случайный лес» на отдельных участках карты и вычисления с ее помощью биомассы на остальных участках карты.

В работе [2] описана усовершенствованная версия данного подхода из статьи [3]. В обоих случаях тот же коллектив ученых разрабатывал модели в континентальном масштабе – для Африки, Азии и Южной Америки и использовал для этого изображения MODIS с разрешением 500 м.

Однако для нашего случая подход необходимо модифицировать. Ведь разрешение изображений в 500 м оказывается недостаточным для построения углеродной карты в рамках региона РФ. Также следует принять во внимание тот факт, что различия в видах растительности, обусловленные различиями климата, также могут влиять на точность модели.

Данная статья рассматривает проблему получения углеродной карты. Она разбита на 4 части: в первой части будут подробно рассмотрены методы сбора полевых данных и преобразования их в значения биомассы. Вторая часть будет посвящена анализу сигналов лидара и проверке согласованности полевых данных с метриками сигналов лидара. В третьей части будет уделено внимание построению модели машинного обучения для вычисления значений биомассы различных участков будущей углеродной карты округа. В четвертой части работы будут приводиться предложения по изменению подхода к созданию углеродных карт для адаптации его к условиям северного региона.

1. Сбор полевых данных. Это первый этап создания углеродной карты. Для начала выбираются участки, соответствующие «пятнам» лидара GLAS.

«Пятна», т.е. отпечатки лидара GLAS, имеют диаметр около 70 метров и расположены на расстоянии 170 метров друг от друга. Съёмка ведется вдоль поверхности Земли с помощью двух каналов с длинами волн 532 и 1064 нм [4]. Согласно данным из [5], архив данных лидара GLAS за 3 года (с 10.2018 по 01.2022) содержит примерно 5468 файлов общим размером ~11 Тб для территории ХМАО.

Замечание. В работе [3] используются датасеты, в которых диаметр «пятен» равен 65 м; участок имеет форму квадрата со сторонами 40 метров.

Затем выбираются участки лесных массивов, целиком лежащие в «пятнах» лидара GLAS. Схема выбора участка в «пятне» лидара показана на рисунке 1. После чего в границах каждого участка измеряется диаметр ствола всех деревьев на высоте 1,3 м от земли. При вычислении биомассы деревьев используется среднее значение удельного веса дерева (плотность) на основе вида.

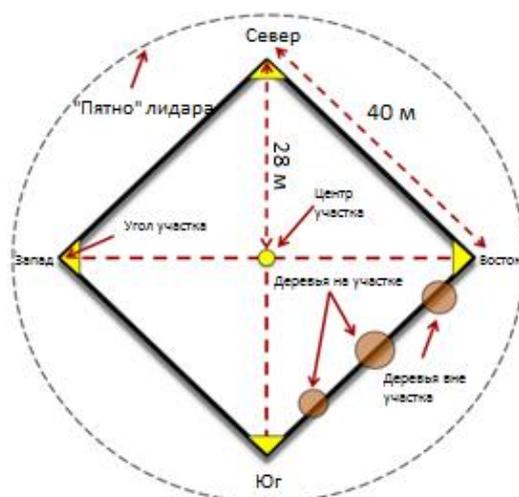


Рис. 1. Схема выбора участка в «пятне» лидара GLAS

В работе [2] для оценки биомассы предлагается регрессионная модель следующего вида:

$$\ln(AGB) = \alpha + \beta_1 \ln(D) + \beta_2 \ln(H) + \beta_3 \ln(\rho), \quad (1)$$

где AGB – значение биомассы (англ. above-ground biomass), D и H – диаметр и высота дерева, а ρ – плотность древесины. Однако на практике высота дерева не всегда доступна. Поэтому, как считают авторы [2], необходимо отойти от ее использования в модели:

$$\ln(AGB) = \alpha + b \ln(D) + c(\ln(D))^2 + d(\ln(D))^3 + \beta_3 \ln(\rho) \quad (2)$$

При логарифмировании данных окончательная оценка будет смещаться и, таким образом, нескорректированные оценки биомассы будут недооценивать реальные значения. Поэтому оценку необходимо умножать на поправочный коэффициент:

$$CF = \exp\left(\frac{RSE^2}{2}\right), \quad (3)$$

где RSE – метрика полученной модели (англ. root square error – корень квадратичной ошибки). Авторы [2] утверждают, что разновидности данной модели в основном применяются для оценки биомассы широколиственных и тропических деревьев.

Лидары применяются для оценки биомассы и в аналогичных исследованиях. Так, в работе [6] авторы применяют гиперспектральные и лидарные данные для оценки биомассы кукурузы. Однако вместо спутниковых лидаров используется бортовой лидар Leica с дискретным возвратом, предназначенный для съемок на малой высоте. При этом в работе [6] процедура полевых измерений схожа с таковой в работе [3] – измерение высоты и массы стеблей кукурузы.

Способ получения полевых измерений в работе [7] полностью аналогичен основной работе ([3]) с тем отличием, что дополнительно измеряется высота дерева с помощью лазерного гипсометра. Там же приведены аллометрические уравнения, которые учитывают еще и высоту дерева. При этом уравнения предназначены для определения биомассы ствола, ветвей, листьев и плодов (sic!) дерева.

В работе [8] применяются лидарные данные и гиперспектральные изображения для определения влияния размера выбранного участка на точность определения биомассы. Причем оценивается не только надземная, но еще и подземная биомасса.

Работа [9] посвящена разработанному авторами методу оценки объема ствола под названием Outer Hull Model – ОНМ. Данный метод применяется для оценки биомассы отдельных деревьев, что в нашем случае может быть слишком затратно. Согласно полученным в [9] результатам, для хвойных деревьев биомасса хвои и ветвей дерева может достигать до 20% биомассы всего дерева.

В работе [10], посвященной оценке лесной биомассы, используются полученные в ходе миссии GEDI лидарные данные. Миссия GEDI была начата сравнительно недавно на момент публикации данной статьи – 5 декабря 2018 г. Согласно данным из [11], лидар миссии GEDI установлен на международной космической станции (МКС), имеет разрешение 1 км и собирает данные в диапазоне с 51.6 ю.ш. до 51.6 с.ш. Упомянуется в [10] и об оптимальном разрешении снимков для создания углеродных карт – оно принято равным 100 м.

В работе [12] процедура сбора полевых данных также ничем не отличается от таковых в рассмотренных работах. В ней упоминается об использовании значения диаметра дерева на уровне груди в 10 см. Авторы [12] утверждают, что это значение может быть слишком велико, поскольку на деревья с меньшим диаметром может приходиться значительная доля биомассы (произведена оценка биомассы в Кении).

Работа [13] посвящена оценке лесной биомассы на острове Калимантан. В данной работе представляет интерес необычный подход к определению биомассы на участке – он основан на методе Монте-Карло.

Работа [14] в этом списке стоит особняком: здесь описывается подход к оценке биомассы травянистой растительности, растущей вблизи болот. В отличие от вычисления биомассы лесов здесь не требуется использование лидарных данных.

Работа [15] также отличается от остальных: в ней описывается подход оценки биомассы в урбанизированных территориях. Особенность данного подхода заключается в использовании высокоточных спутниковых снимков (разрешение 5 м), а для измерения высоты деревьев при сборе полевых данных используется лазерный высотомер. Подход включает в себя метод стратификации лидарных снимков, т.е. деление снимков на слои по высоте.

Исходя из представленного обзора методов сбора полевых данных, можно видеть, что во всех работах получение данных о растительности связано либо с измерениями на местах (деревья), либо со сбором и обработкой образцов растительности (травы). При агрегировании данных о биомассе наиболее подходящим в нашем случае способом агрегации данных является вычисление среднего значения биомассы по классам растительности, аналогично проделанным действиям в работах [3, 15].

2. Анализ сигналов лидара. Следующий этап создания углеродной карты – анализ формы сигнала лидара GLAS. Это необходимо для установления статистической взаимосвязи между полевыми оценками плотности надземного углерода (англ. aboveground carbon density, ACD). В работе [2] предполагается, что отношение массы углерода к общей биомассе составляет 1:2.

Здесь же подробно изложен процесс анализа «пятен» лидара GLAS. Так, перед непосредственной обработкой данных авторы отбросили «пятна», которые имеют менее 2 пиков, максимальную высоту, превышающую шум менее чем в 2 раза и расхождение со значениями высоты от SRTM более 25 м.

Подробное описание алгоритма оценки биомассы по снимкам лидара (LVIS) дано в работе [16]. Согласно ему, сигнал с лидара переводится в профиль высоты покрова. При этом используется методика из статьи [17]. По сути, дальнейший анализ «пятна» лидара связан с анализом гистограммы профиля покрова.

Методика перевода сигнала с лидара в профиль высоты покрова из статьи [17] состоит из следующих шагов: определение и отсечение шума, вычисление кумулятивной «функции покрова», преобразование «функции покрова» по методологии МакАртура-Хорна, обратное преобразование ее в плотность вероятности. Ее схема показана на рисунке 2.

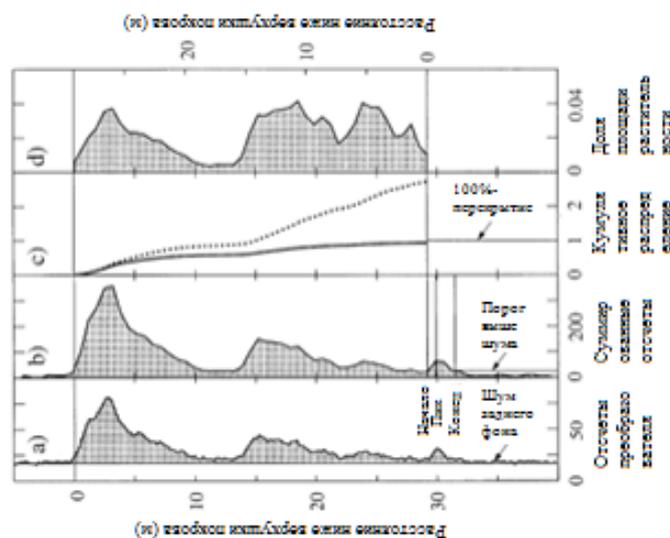


Рис. 2. Схема перевода сигнала лидара в профиль высоты покрова

Профиль высоты покрова является ничем иным, как функцией плотности распределения высот в «пятне» лидара. Далее из профиля высоты извлекаются значения перцентилей функции распределения высот и другая дополнительная информация.

Применение вышеописанных алгоритмов в работе [3] позволило получить по 2 значения биомассы на каждый участок, по которым была построена линейная регрессионная модель.

В других работах по составлению карты плотности углерода проводились аналогичные исследования по установлению связи между полевыми данными биомассы и метриками формы сигнала лидара. В работе [7], в отличие от подхода из [3], при анализе лидарных данных для каждого "пятна" вместо профиля высоты создается цифровая модель земной поверхности с помощью воксельного подхода. А в работе [8] для получения цифровой модели земной поверхности используется метод интерполяции триангулированной нерегулярной сети. Более подробно подход составления цифровой модели высоты покрова описан в [21]. В работе [9] для оценки объема ствола хвойного дерева так же, как и в работе [7], применяется воксельный подход при анализе сигналов лидара. При анализе сигналов лидара в работе [10] используется метод случайного объема над землей (Random Volume over Ground – RvOG), который применяется к полученным с помощью InSAR изображениям. В работе [12] заслуживают внимания приемы обработки лидарных данных. В ней рассказывается об обработке лидарных снимков: здания, линии электропередач и выбросы (высокие точки) фильтруются с помощью ПО Terrascan, LAStools (Rapidlasso GmbH) и ручного редактирования. Приводятся аллометрические уравнения для разнородных лесных участков, а также указаны справочные материалы для определения плотности древесины. Поиск наиболее точных регрессионных моделей производится путем перебора различных комбинаций из 1-3 предикторов.

В работе [14] выполняется классификация травянистой растительности на основании объектно-ориентированного подхода. Для создания детализированной (фракционной) карты болотистой растительности использовались снимки NAIP с разрешением 1 м, после чего данные биомассы с первой карты были использованы для сведения полученных значений к разрешению спутниковых снимков Sentinel (разрешение 30 м). В работе [15], как уже говорилось ранее, обработка лидарных снимков включает в себя стратификацию лидарных снимков, т.е. деления снимков на слои по высоте. Для вычисления общей биомассы используются усредненные значения биомассы по классам. Полученные в работе [15] значения коэффициентов корреляции Пирсона высоты покрова с биомассой возрастают с ростом высоты слоя: для высот 2-5 м коэффициент корреляции равен 0.3, для 5-10 м – 0.5, а для высот более 10 м – 0.7.

Исходя из представленного обзора методов обработки лидарных данных, видно, что во всех случаях применение характеристик сигналов лидара приводит к повышению точности оценки биомассы.

В ряде рассмотренных работ для определения формы растительности используются бортовые лидары [9, 12, 13] или лазерные высотомеры [15]. При этом в работе [13] утверждается, что модели оценки биомассы на основе данных бортовых лидаров показывают более высокую точность, чем модели оценки биомассы на основе данных спутниковых лидаров. В нашем случае, скорее всего, не будет возможности применения бортовых лидаров или лазерных высотомеров, поэтому в нашей задаче придется задействовать данные спутниковых лидаров, прежде всего GLAS.

3. Создание карты биомассы. В работе [3] создание карты плотности углерода сводится к созданию карты лесной биомассы. Разработанные для создания карт биомассы методы включают в себя:

- классификацию типов растительного покрова, где каждому классу присваивается среднее значение плотности биомассы на основе справочных оценок или данных инвентаризации весов;
- определение взаимосвязей между плотностью биомассы и характеристиками пикселей спутниковых изображений, которые могут быть отображены на больших пространствах карты.

В работах [3,16] для создания непрерывных карт использованы изображения со спектрорадиометра MODIS среднего разрешения (500 м). Для получения высококачественных снимков авторы [3] составили композицию из изображений, взятых за период в 2 полных календарных года. После получения свободного от облаков композитного изображения для всех его пикселей были вычислены индексы растительности EVI2 и NDII. Кроме того, для тех же наблюдений извлечены данные о температуре поверхности суши (MODIS LST) и вычислены средние значения. Также применены данные о высоте (SRTM) с разрешением 90 м.

Таким образом, для обучения модели машинного обучения «случайный лес», которая генерирует оценки плотности углерода в каждом пикселе изображения, использовались такие переменные как каналы 1, 2, 4, 5, 6, 7 спутникового изображения MODIS, значения отклонений в 7 канале MODIS, средняя температура и значения ее отклонений (LST), коэффициенты растительности EVI2, NDII и SRTM.

В работе [2] установлено, что каналы 1 (620-670 нм) и 7 (2105-2155 нм) MODIS, скорее всего (sic!), являются наиболее важными переменными для различий в плотности углерода в наземной части. В частности, канал красного цвета позволяет отличать области с растительностью от областей без нее, а канал 7 позволяет идентифицировать регионы с высокой плотностью углерода.

Важность применения именно ИК-изображений для анализа лесного покрова подтверждается и в других работах [18, 19, 20]. В последней [20] есть замечательная иллюстрация (рисунок 3), демонстрирующая разделение наземной растительности от хвойных деревьев.

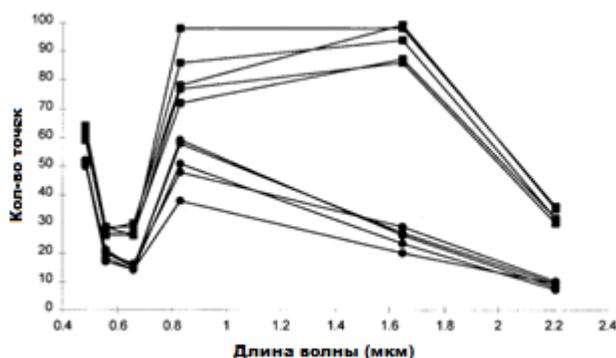


Рис. 3. Разделимость наземной растительности от хвойных деревьев на ИК-каналах

На рисунке 3 видно, что наилучшая разделимость хвойной и наземной растительности проявляется на длинах волн ~1650 нм.

Помимо прочего, авторы [2] указали общие требования к участкам, на которых проводятся полевые измерения. Так, для всех участков, где проводились измерения, помимо базовой информации (размер, форма, измерения деревьев), необходимо также знать информацию об участках, которые располагаются более чем на одном пикселе изображения. Здесь также говорится, что чем больше изменчивость плотности углерода в пикселе, тем больше участков необходимо взять в пределах данного пикселя. Это актуально для пикселей с изменяющимся ландшафтом (см. рисунок 4).

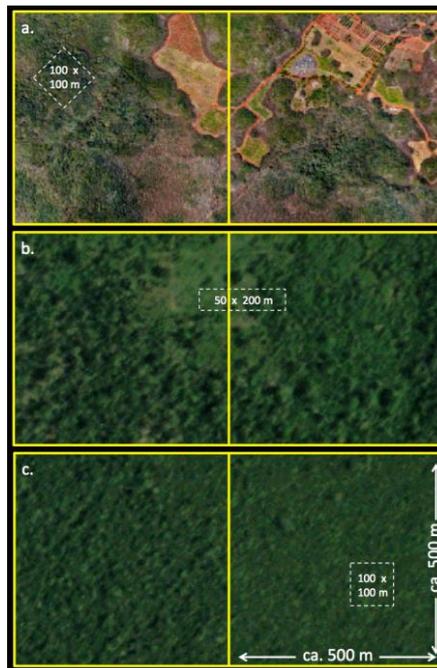


Рис. 4. 3 различных варианта лесного ландшафта (сверху вниз): деградированный лес, деградированный густой лес, нетронутый густой лес

Упомянуется и про необходимость использования карт с одинаковым разрешением, поскольку изменение разрешения карт может приводить к ухудшению результатов. Однако на практике доступность карт одного и того же разрешения зависит от их наличия.

Результаты. Как сообщается в приложении к статье [2], полученная модель случайного леса позволила объяснить 83%, 78% и 71% различий в ACD на тестовых данных для тропической Америки, Африки и Азии соответственно.

В работе [6] (оценка биомассы кукурузы) проведено исследование по определению индекса растительности, дающего наибольшее значение R^2 итоговой регрессионной модели, в котором показаны очень высокие значения корреляции индексов растительности друг с другом. Также подтверждается надежность применения лидаров для оценки биомассы ($R^2 = 0,835$), а сочетание лидарных снимков с индексами растительности повышает значение R^2 ($R^2 = 0,883$).

И в работе [7] (оценка биомассы хвойных лесов) сочетание различного рода характеристик сигналов лидара приводит к повышению R^2 : так, на характеристиках лидаров дискретного возврата наибольшее достигнутое значение R^2 равно 0,702, на характеристиках лидара полной волны – 0,760, а на сочетании характеристик лидаров дискретного возврата и полной волны – 0,815.

В работе [8] (оценка надземной и подземной биомассы лесов) наибольшие значения R^2 для подземной/надземной/общей биомассы составили 0,742/0,874/0,860, 0,513/0,545/0,552 и 0,785/0,893/0,882 для только лидарных, только гиперспектральных и лидарных + гиперспектральных данных соответственно.

Результаты работы [10] подтверждают результаты предыдущих работ: применение данных TDX (мультиспектральных) и GEDI (лидар) на трех участках позволило получить значения R^2 от 0,82 до 0,90, в то время как только для данных GEDI – от 0,58 до 0,90, что еще раз показывает важность совместного применения лидарных и мультиспектральных снимков для оценки биомассы.

В работе [21] полученные результаты (R^2 -коэффициент) следующие: для уравнений объема ствола с использованием только общей модели покрова, общей модели покрова с

DSM (дискретный возврат) и общей модели покрова с DSM (полная волна) значения составили 0,93, 0,94 и 0,95 соответственно. А для уравнений объема биомассы – 0,87, 0,88 и 0,91 соответственно.

В работе [12] применение сегментации с переменным значением масштаба и алгоритма классификации kNN позволило получить значения R^2 -коэффициентов для моделей оценки биомассы в горных районах и низинах, равные 0,93 и 0,89 соответственно.

В работе [13] точность полученных моделей оценки биомассы, по утверждению авторов, не является высокой (значения R^2 -коэффициентов при этом не приведены). Они же утверждают, что углеродные карты на основе данных бортовых лидаров показывают более лучшие результаты, чем углеродные карты на основе данных спутниковых лидаров.

А при оценке биомассы травянистой растительности на болотах в работе [14] результаты показали, что плотность травянистой растительности в 95% случаях не превышает 1,56 кг/м² (разумеется, для северного региона данная цифра может отличаться). Значения общей точности классификации растительности составили от 80.5% до 98%. Однако для задачи регрессии полученные значения R^2 невысокие: от 0.36 до 0.61.

В работе [15] для оценки биомассы растительности на городских территориях использовались BRT-модели. Значения R^2 -коэффициентов для 300 BRT-моделей составили 0.77-0.89 для данных ландшафта и 0.42-0.65 для моделей формы города.

Исходя из представленных результатов аналогичных работ, видно, что совместное применение мультиспектральных данных и данных лидара при создании карт плотности углерода приводит к наиболее точным оценкам биомассы, а углеродные карты на основе данных бортовых лидаров имеют более высокую точность. Различий в результатах при использовании различных способов обработки сигналов лидара и моделей регрессии не замечено, поэтому в нашем случае будет необходима проверка этих двух способов на более высокую точность оценки биомассы. При оценке биомассы лесов практически во всех работах применяются регрессионные модели для оценки согласованности сигналов лидара с полевыми измерениями.

4. Адаптация подхода к северному региону. Как говорилось ранее, разрешение снимка в 500 метров оказывается слишком грубым для анализа региона. Поэтому были рассмотрены (с учетом ограничений по длине волны и необходимому разрешению) альтернативные варианты спутниковых снимков, которые могут подойти в нашем случае (таблица 1).

Таблица 1. Альтернативные варианты спутниковых снимков

Спутник	Канал (длина волны, нм)	Разрешение, м
Landsat-7	band 5 (1550-1750)	30
	band 7 (2080-2350)	30
Landsat-8	band 6 (1560-1660)	30
	band 7 (2100-2300)	60
WorldView3 (коммерческий)	SWIR-3 (1640-1680)	3,7
	SWIR-5 (2145-2185)	3,7
	CAVIS Snow (1620-1680)	30
	Aerosol-3 (2105-2145)	30
Sentinel 2	B4 (665)	10
	B11 (1610)	20
	B12 (2190)	20

Также следует иметь в виду отличие видов растительности северного региона от видов растительности экваториальных районов Земли, для которых и проводились описанные в работах [2, 3] исследования. Поскольку в экваториальных районах преобладают тропические, мангровые леса, а также пальмы, тогда как в ХМАО преобладают лиственные и хвойные леса.

Предлагаемая программа исследований. Исходя из анализа представленной литературы, предлагается выполнение следующих исследований по созданию углеродной карты ХМАО:

- 1) исследование лидарных данных GLAS и мультиспектральных данных представленных в таблице 1 спутников. Цель: определение наиболее подходящих и одновременно доступных спутниковых изображений;
- 2) определение классов земного покрова и участков, соответствующих пятнам лидара. Проведение полевых измерений. Цель: определение средних значений биомассы на участках для разных классов земного покрова;
- 3) исследование различных способов обработки сигналов лидара и регрессионных моделей перевода характеристик сигнала в значения биомассы. Цель: получение способа обработки сигналов лидара, дающего наибольшую точность вычисления биомассы;
- 4) исследование моделей классификации земного покрова по мультиспектральным спутниковым изображениям. Цель: получение модели классификации земного покрова с наибольшей точностью и создание на ее основе углеродной карты ХМАО.

Заключение. В данной работе проанализирован один из подходов [2-3] к построению карт плотности углерода, которые предназначены для оценки объема растительности и содержащегося в ней углерода. Подход основан на применении биометрических данных, спутниковых снимков, расчете и анализе групп пикселей с помощью множественной линейной регрессии. Данный подход применен для построения карты плотности углерода для экваториальных районов Африки, Азии, Южной Америки.

В работах [2-3] приведено подробное описание используемых математических методов перевода биометрических данных в значения плотности углерода, указаны требования к используемым данным, а также приведены результаты тестирования алгоритма и другие рекомендации по использованию данных.

Помимо указанных работ, проведен детальный анализ аналогичных существующих подходов к оценке биомассы с помощью дистанционного зондирования Земли. На основе анализа определены основные методы и подходы, используемые как для оценки биомассы лесов, так и для оценки биомассы травяной растительности болот и оценки биомассы растительности в городских территориях. Полученный обзор методов и подходов позволил составить программу исследований именно для нашего случая.

Также спланированы шаги по адаптации подхода из работы [3] к одному из регионов Крайнего Севера России – ХМАО. Так, был произведен поиск действующих на январь 2022 года спутников, а также подходящих каналов съемки. Что касается лидарных данных, то для ХМАО имеется более 5000 файлов со спутника ICESat-2 за более чем 3 года.

Также необходимо учитывать различие в видах растительности для экваториальных лесов и тайги (тундры), в зоне которой расположен ХМАО. Для них придется обучать новые регрессионные модели, которые, вероятнее всего, будут отличаться набором переменных и коэффициентами при них.

Список источников

1. Государственный доклад «О состоянии и об охране окружающей среды Российской Федерации в 2020 году». – 2021. – URL: https://www.mnr.gov.ru/docs/gosudarstvennye_doklady/gosudarstvennyu_doklad_o_sostoyanii_i_ob_okhrane_okruzhayushchey_sredy_rossiyskoy_federatsii_v_2020/?PAGEN_2=2 (дата обращения: 14.01.2022).
2. Vaccini A. et al. Tropical forests are a net carbon source based on aboveground measurements of gain and loss. *Science*, 2017, vol. 358, no. 6360, pp. 230-234.
3. Vaccini A. et al. Estimated carbon dioxide emissions from tropical deforestation improved by carbon-density maps. *Nature climate change*, 2012, vol. 2, no. 3, pp. 182-185.
4. NASA: ICESat, 2021, available at: <https://icesat.gsfc.nasa.gov/icesat/glas.php> (accessed: 01/20/2022).
5. ATLAS/ICESat-2 L2A Global Geolocated Photon Data, Version 5 | National Snow and Ice Data Center, 2021, available at: <https://nsidc.org/data/ATL03/versions/5> (accessed: 01/20/2022).
6. Wang C. et al. Estimating the biomass of maize with hyperspectral and LiDAR data. *Remote Sensing*, 2017, vol. 9, no. 1, 11 p.
7. Nie S. et al. Above-ground biomass estimation using airborne discrete-return and full-waveform LiDAR data in a coniferous forest. *Ecological Indicators*, 2017, vol. 78, pp. 221-228.
8. Luo S. et al. Fusion of airborne LiDAR data and hyperspectral imagery for aboveground and belowground forest biomass estimation. *Ecological Indicators*, 2017, vol. 73, pp. 378-387.
9. Stovall A. E. L. et al. Non-destructive aboveground biomass estimation of coniferous trees using terrestrial LiDAR. *Remote Sensing of Environment*, 2017, vol. 200, pp. 31-42.
10. Qi W. et al. Forest biomass estimation over three distinct forest types using TanDEM-X InSAR data and simulated GEDI lidar data. *Remote Sensing of Environment*, 2019, vol. 232, p. 111283.
11. Global Ecosystem Dynamics Investigation | Wikipedia, 2022. Available at: https://en.wikipedia.org/wiki/Global_Ecosystem_Dynamics_Investigation (accessed: 02/11/2022).
12. Pellikka P. K. E. et al. Impact of land cover change on aboveground carbon stocks in Afrotropical landscape in Kenya. *Applied Geography*, 2018, vol. 94, pp. 178-189.
13. Jucker T. et al. Estimating aboveground carbon density and its uncertainty in Borneo's structurally complex tropical forests using airborne laser scanning. *Biogeosciences*, 2018, vol. 15, no. 12, pp. 3811-3830.
14. Byrd K. B. et al. A remote sensing-based model of tidal marsh aboveground carbon stocks for the conterminous United States. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2018, vol. 139, pp. 255-271.
15. Mitchell M. G. E. et al. Identification of fine scale and landscape scale drivers of urban aboveground carbon stocks using high-resolution modeling and mapping. *Science of the total Environment*, 2018, vol. 622, pp. 57-70.
16. Drake J. B. et al. Sensitivity of large-footprint lidar to canopy structure and biomass in a neotropical rainforest. *Remote Sensing of Environment*, 2002, vol. 81, no. 2-3, pp. 378-392.
17. Harding D. J. et al. Laser altimeter canopy height profiles: Methods and validation for closed-canopy, broadleaf forests. *Remote Sensing of Environment*, 2001, vol. 76, no. 3, pp. 283-297.
18. Cohen W. B., Spies T. A. Estimating structural attributes of Douglas-Fir. *Remote sensing of environment*, 1992, vol. 41, no. 1, pp. 1-17.
19. Gemmill F. M. Effects of forest cover, terrain, and scale on timber volume estimation with Thematic Mapper data in a Rocky Mountain site. *Remote Sensing of Environment*, 1995, vol. 51, no. 2, pp. 291-305.
20. Puhr C. B., Donoghue D. N. M. Remote sensing of upland conifer plantations using Landsat TM data: a case study from Galloway, south-west Scotland. *International Journal of Remote Sensing*, 2000, vol. 21, no. 4, pp. 633-646.
21. Allouis T. et al. Stem volume and above-ground biomass estimation of individual pine trees from LiDAR data: Contribution of full-waveform signals. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2012, vol. 6, no. 2, pp. 924-934.

Бредихин Арсентий Игоревич. 1996 г.р., аспирант 2 курса ЮГУ направления «Информатика и вычислительная техника» (профиль – «Системный анализ, управление и обработка информации»), инженер-программист УЦР ЮГУ, AuthorID – 1012558, SPIN-код: 9495-5105, a_bredihin@ugrasu.ru, Россия, г. Ханты-Мансийск.

Мельников Андрей Витальевич. 1956 г.р., директор АУ ХМАО-Югры «Югорский научно-исследовательский институт информационных технологий», д.т.н., профессор. AuthorID – 179248, melnikovav@uriit.ru, Россия, г. Ханты-Мансийск.

UDC 528.854.4

DOI:10.38028/ESI.2022.26.2.011

An approach to the creation of a carbon density map of the Khanty-Mansiysk Autonomous Okrug based on the assessment of biomass using Earth remote sensing data

Arcenty I. Bredikhin¹, Andrey V. Melnikov²

¹Yugra State University (UGRASU),

²Yugra Research Institute of Information Technologies (URIIT),

Russia, Khanty-Mansiysk, *a_bredihin@ugrasu.ru*

Abstract. Khanty-Mansi Autonomous Okrug-Yugra has a large area of forest territories (more than 90% of the total area of the district). And forest vegetation naturally dies sooner or later, as a result of which carbon dioxide is released into the atmosphere from organic matter and global warming increases.

In order to prevent an increase in global temperature, it is necessary to estimate the carbon stock in the form of the amount of vegetation biomass.

One of the ways to assess vegetation biomass is to create carbon maps using of the Earth remote sensing methods. The use of satellite images obtained with remote sensing methods and methods of their processing will make it possible to obtain a map of the district with full coverage of the entire territory, and the use of machine learning models will make it possible to develop a model with which it will be possible to obtain a carbon map of the district with a given accuracy.

This paper provides an overview of the decisions of foreign scientists over the past 5 years in the field of remote sensing aimed at creating carbon maps. Based on this review, a research program has been proposed that will allow us to develop an approach that allows us to obtain a digital carbon map of the KhMAO with the necessary accuracy.

Keywords: remote sensing, satellite image, carbon, vegetation biomass, machine learning

References

1. State Report "On the state and environmental protection of the Russian Federation in 2020" [Electronic resource]. 2021. Available at: https://www.mnr.gov.ru/docs/gosudarstvennye_doklady/gosudarstvennyy_doklad_o_sostoyanii_i_ob_okhrane_okruzhayushchey_sredy_rossiyskoy_federatsii_v_2020/?PAGEN_2=2 (accessed: 01/14/2022).
2. Baccini A. et al. Tropical forests are a net carbon source based on aboveground measurements of gain and loss. *Science*, 2017, vol. 358, no. 6360, pp. 230-234.
3. Baccini A. et al. Estimated carbon dioxide emissions from tropical deforestation improved by carbon-density maps. *Nature climate change*, 2012, vol. 2, no. 3, pp. 182-185.
4. NASA: ICESat, 2021, available at: <https://icesat.gsfc.nasa.gov/icesat/glas.php> (accessed: 01/20/2022).
5. ATLAS/ICESat-2 L2A Global Geolocated Photon Data, Version 5 | National Snow and Ice Data Center, 2021, available at: <https://nsidc.org/data/ATL03/versions/5> (accessed: 01/20/2022).
6. Wang C. et al. Estimating the biomass of maize with hyperspectral and LiDAR data. *Remote Sensing*, 2017, vol. 9, no. 1, 11 p.
7. Nie S. et al. Above-ground biomass estimation using airborne discrete-return and full-waveform LiDAR data in a coniferous forest. *Ecological Indicators*, 2017, vol. 78, pp. 221-228.
8. Luo S. et al. Fusion of airborne LiDAR data and hyperspectral imagery for aboveground and belowground forest biomass estimation. *Ecological Indicators*, 2017, vol. 73, pp. 378-387.
9. Stovall A. E. L. et al. Non-destructive aboveground biomass estimation of coniferous trees using terrestrial LiDAR. *Remote Sensing of Environment*, 2017, vol. 200, pp. 31-42.
10. Qi W. et al. Forest biomass estimation over three distinct forest types using TanDEM-X InSAR data and simulated GEDI lidar data. *Remote Sensing of Environment*, 2019, vol. 232, p. 111283.
11. Global Ecosystem Dynamics Investigation | Wikipedia, 2022. Available at: https://en.wikipedia.org/wiki/Global_Ecosystem_Dynamics_Investigation (accessed: 02/11/2022).
12. Pellikka P. K. E. et al. Impact of land cover change on aboveground carbon stocks in Afrotropical landscape in Kenya. *Applied Geography*, 2018, vol. 94, pp. 178-189.
13. Jucker T. et al. Estimating aboveground carbon density and its uncertainty in Borneo's structurally complex tropical forests using airborne laser scanning. *Biogeosciences*, 2018, vol. 15, no. 12, pp. 3811-3830.

14. Byrd K. B. et al. A remote sensing-based model of tidal marsh aboveground carbon stocks for the conterminous United States. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2018, vol. 139, pp. 255-271.
15. Mitchell M. G. E. et al. Identification of fine scale and landscape scale drivers of urban aboveground carbon stocks using high-resolution modeling and mapping. *Science of the total Environment*, 2018, vol. 622, pp. 57-70.
16. Drake J. B. et al. Sensitivity of large-footprint lidar to canopy structure and biomass in a neotropical rainforest. *Remote Sensing of Environment*, 2002, vol. 81, no. 2-3, pp. 378-392.
17. Harding D. J. et al. Laser altimeter canopy height profiles: Methods and validation for closed-canopy, broadleaf forests. *Remote Sensing of Environment*, 2001, vol. 76, no. 3, pp. 283-297.
18. Cohen W. B., Spies T. A. Estimating structural attributes of Douglas-Fir. *Remote sensing of environment*, 1992, vol. 41, no. 1, pp. 1-17.
19. Gemmill F. M. Effects of forest cover, terrain, and scale on timber volume estimation with Thematic Mapper data in a Rocky Mountain site. *Remote Sensing of Environment*, 1995, vol. 51, no. 2, pp. 291-305.
20. Pühr C. B., Donoghue D. N. M. Remote sensing of upland conifer plantations using Landsat TM data: a case study from Galloway, south-west Scotland. *International Journal of Remote Sensing*, 2000, vol. 21, no. 4, pp. 633-646.
21. Allouis T. et al. Stem volume and above-ground biomass estimation of individual pine trees from LiDAR data: Contribution of full-waveform signals. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2012, vol. 6, no. 2, pp. 924-934.

Bredikhin Arcenty Igorevich, born 1996, PhD student of the 2nd year of the UGRASU of the direction "Computer Science and Computer Engineering" (profile – "System analysis, management and information processing"), software engineer of the DDD UGRASU, AuthorID – 1012558, SPIN-code: 9495-5105, a_bredihin@ugrasu.ru, Russia, Khanty-Mansiysk.

Melnikov Andrey Vitalievich, born 1956, Director of the AU KhMAO-Yugra "Yugra Research Institute of Information Technologies", Doctor of Technical Sciences, Professor. AuthorID – 179248, melnikovav@uriit.ru, Russia, Khanty-Mansiysk.

Статья поступила в редакцию 05.04.2022; одобрена после рецензирования 09.06.2022; принята к публикации 14.06.2022.

The article was submitted 04/05/2022; approved after reviewing 06/09/2022; accepted for publication 06/14/2022.