

Подход вычисления неопределенности спутниковых оценок запаса углерода лесов и других их характеристик на основе подхода вариации обучающей выборки

Хвостиков С.А., Барталев С.А., Ворушилов И.И., Егоров В.А., Ховратович Т.С.

khvostikov@d902.iki.rssi.ru

Институт Космических Исследований РАН

Работа выполнена в рамках реализации важнейшего инновационного проекта государственного значения

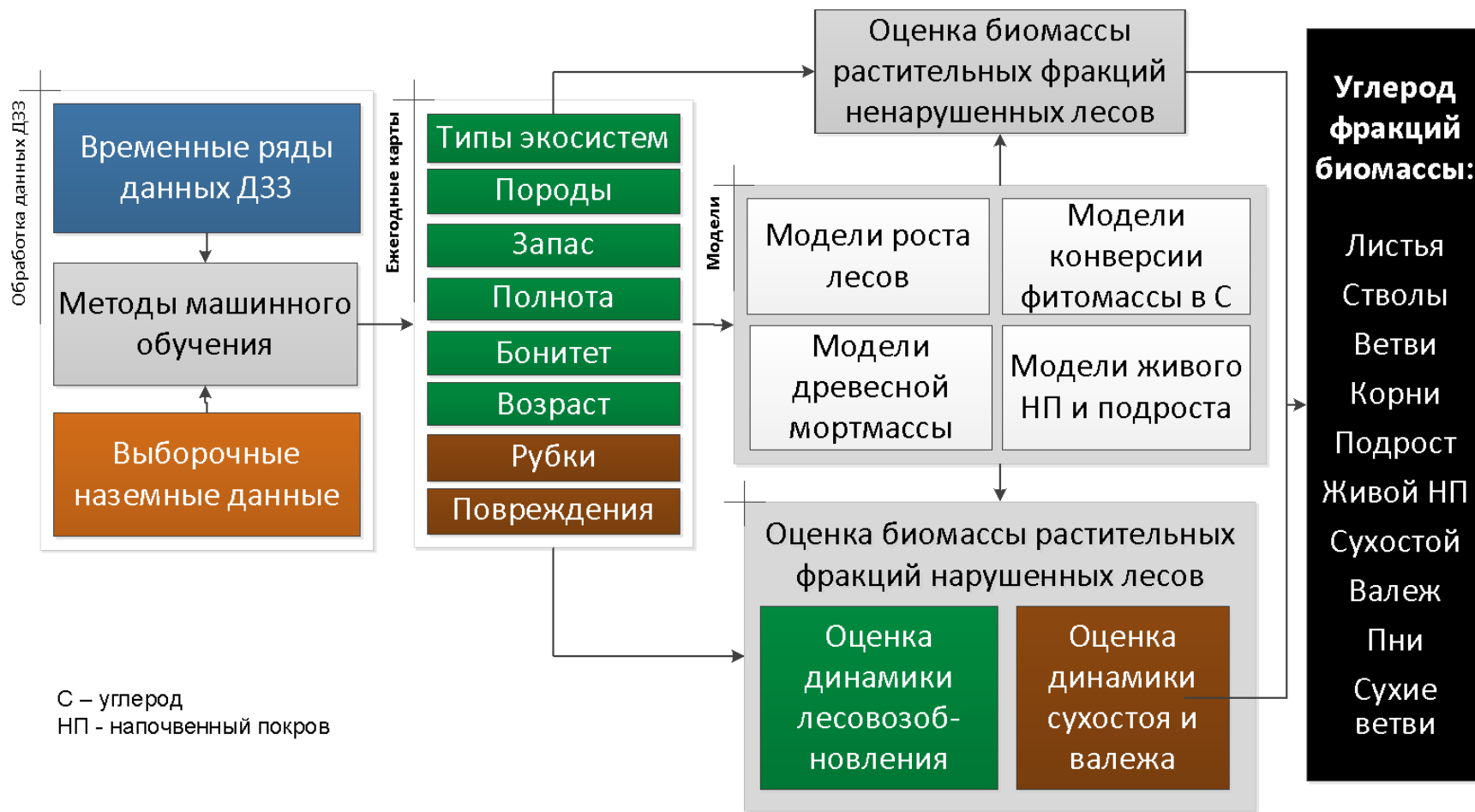
"Разработка системы наземного и дистанционного мониторинга пулов углерода и потоков парниковых газов на территории Российской Федерации, обеспечение создания системы учета данных о потоках климатически активных веществ и бюджете углерода в лесах и других наземных экологических системах»
(рег. № 124060500032-7.).

Двадцать вторая международная конференция
«Современные проблемы дистанционного зондирования Земли из космоса»

ИКИ РАН 11 – 15 ноября 2024 г.

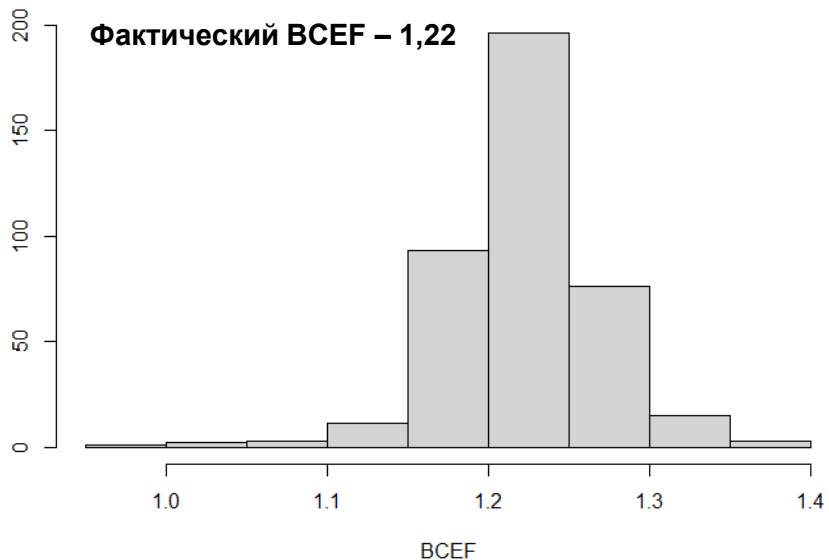
Методика оценки углерода

Оценка углерода основывается на комбинировании спутниковых карт характеристик лесов и набора моделей.



Пример вероятностных распределений в пикселе

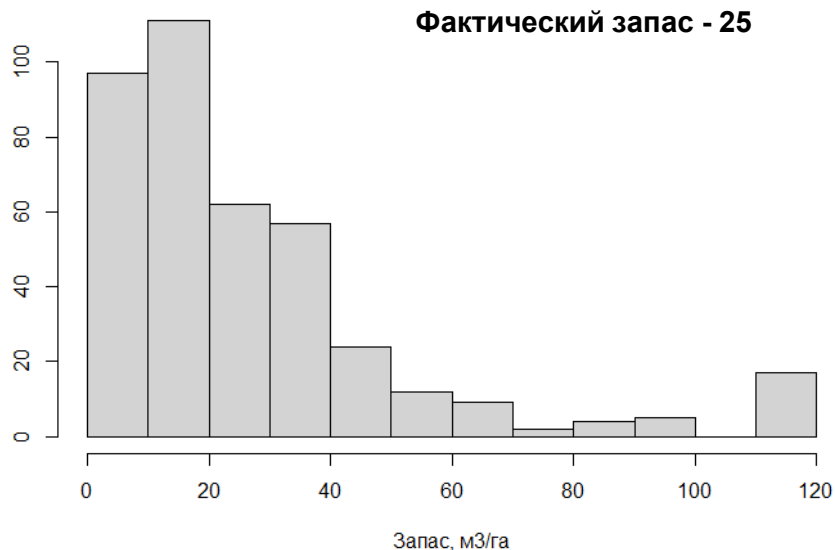
Вероятностное распределение конверсионного коэффициента



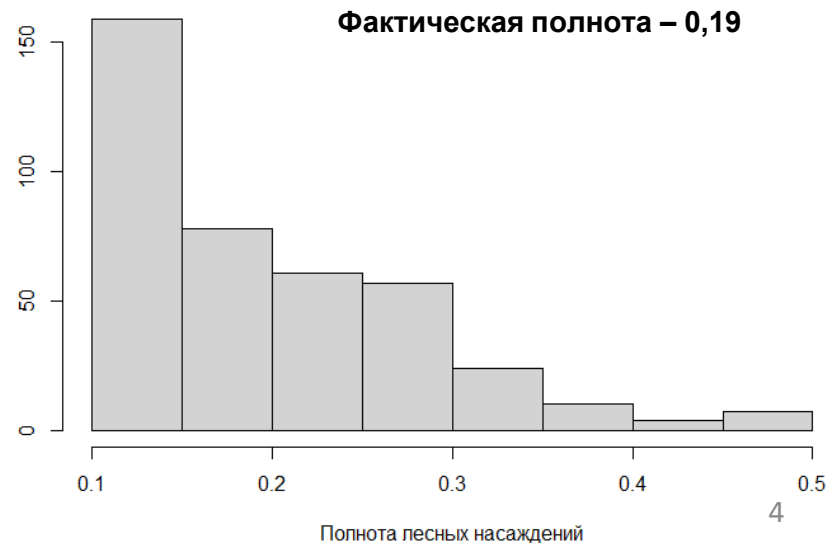
Распределение пород

Порода	Доля %
Сосна	3
Листв.	92,75
Береза	2,25
Ель	0,75
Кедр	1,25

Вероятностное распределение запаса



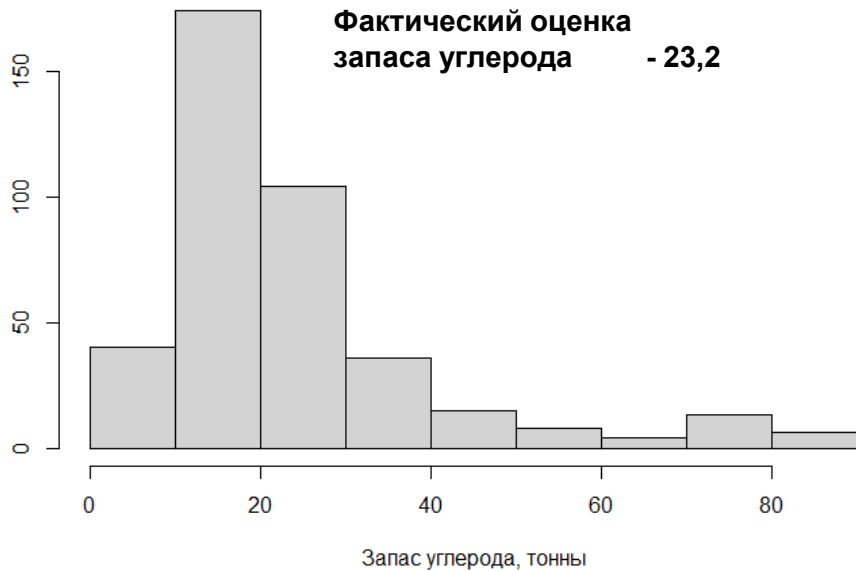
Вероятностное распределение полноты



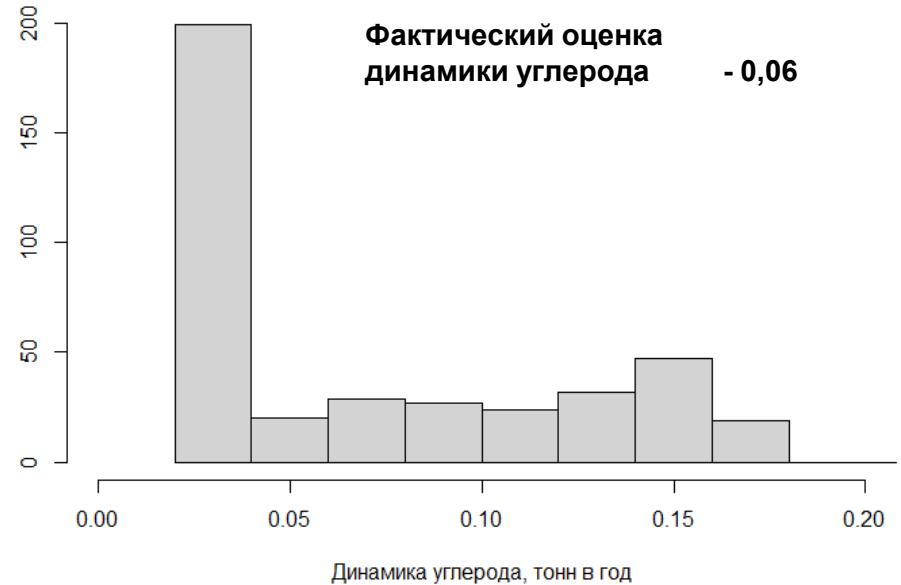
Неопределенность углерода в пикселе

Выполнение большого числа итераций с генерацией входных параметров и модельных коэффициентов по их распределением и выполнением оценки углерода позволяет получить вероятностное распределение запаса и динамики углерода.

Результат Монте-Карло - суммарный углерод



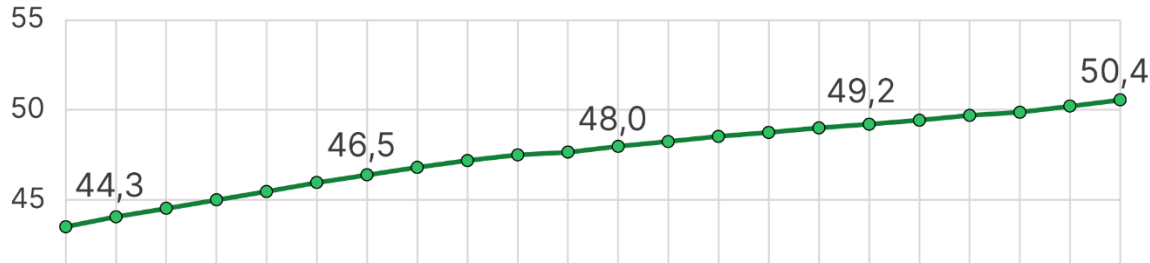
Результат Монте-Карло - динамика



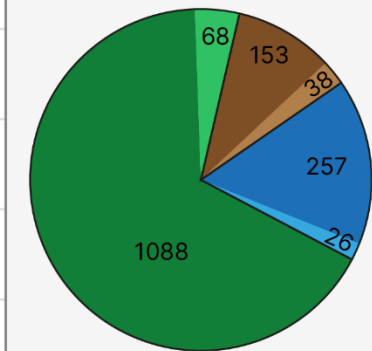
Необходимые оценки углерода

Запас углерода в растительной биомассе лесов и редиин

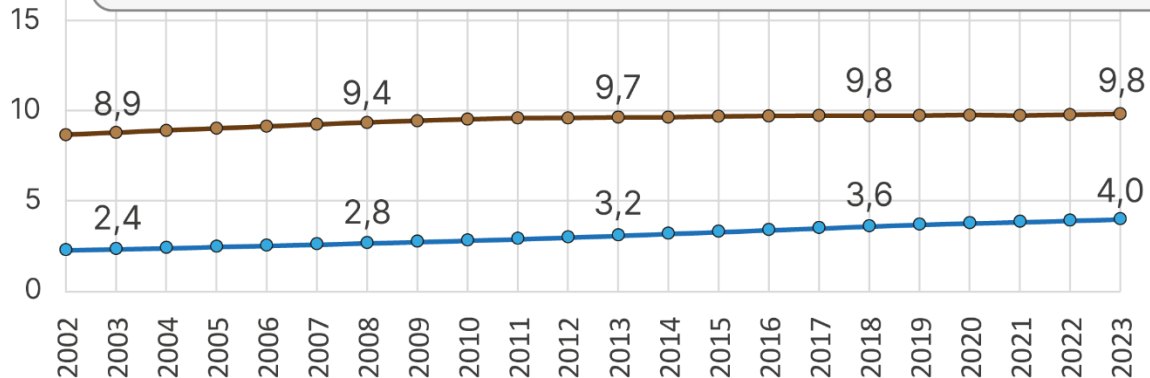
Результаты оценки нетто-поглощения углерода лесами



Среднее нетто-поглощение 2002-2023, млн т CO₂ в год:



	Управляемые	Резервные	Неучтенные	
Леса	1088,5	153,3	256,6	1498,4
Редины	68,4	38,2	26,4	133,0
	1156,9	191,5	283,0	1631,5



Нужна погрешность для суммарных оценок запаса углерода и их динамики

Оценка погрешности нужна в разрезе различных типов лесов и субъектов РФ

Неопределенность оценок суммы характеристик

Сумма нормальных распределений зависит от корреляции между ними

$$N(0, \sigma_1) + N(0, \sigma_2) = N(0, \sqrt{\sigma_1^2 + \sigma_2^2 + 2 * \sigma_1 * \sigma_2 * cor})$$

Для двух пикселей с одинаковой погрешностью:

$$cor = 1 \quad , \quad RMSE = 2\sigma$$

$$cor = 0 \quad , \quad RMSE = \sqrt{2}\sigma$$

Для множества пикселей с одинаковой погрешностью:

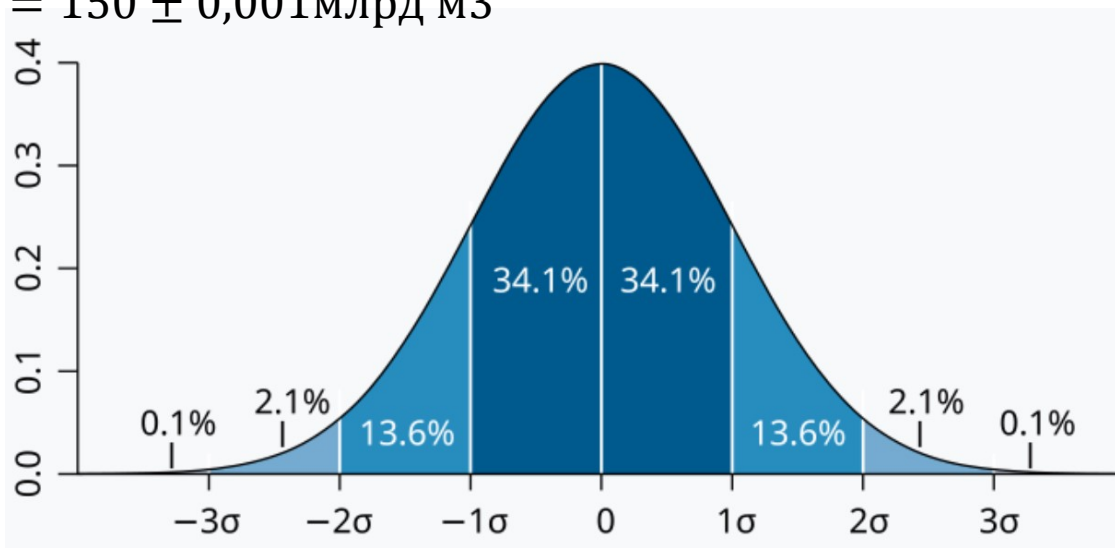
$$cor = 1 \quad , \quad RMSE = n\sigma$$

$$cor = 0 \quad , \quad RMSE = \sqrt{n}\sigma$$

Для запаса, 200 000 000 пикселей MODIS, погрешность 10%

$$cor = 1 \quad , \quad GSV = 150 \pm 15 \text{ млрд м}^3$$

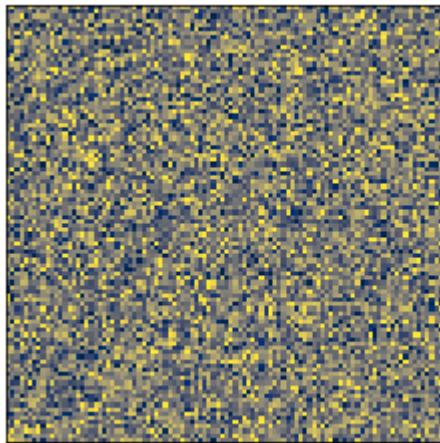
$$cor = 0 \quad , \quad GSV = 150 \pm 0,001 \text{ млрд м}^3$$



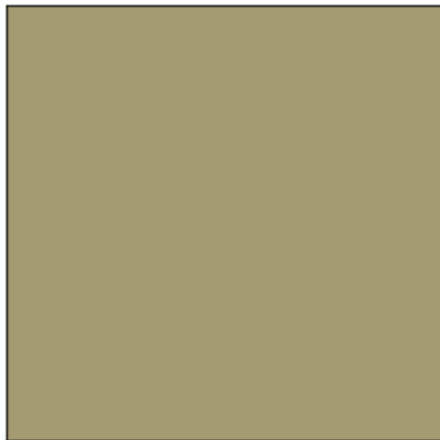
Неопределенность оценок суммы характеристик

Фактическая корреляция между пикселями находится между двумя крайностями

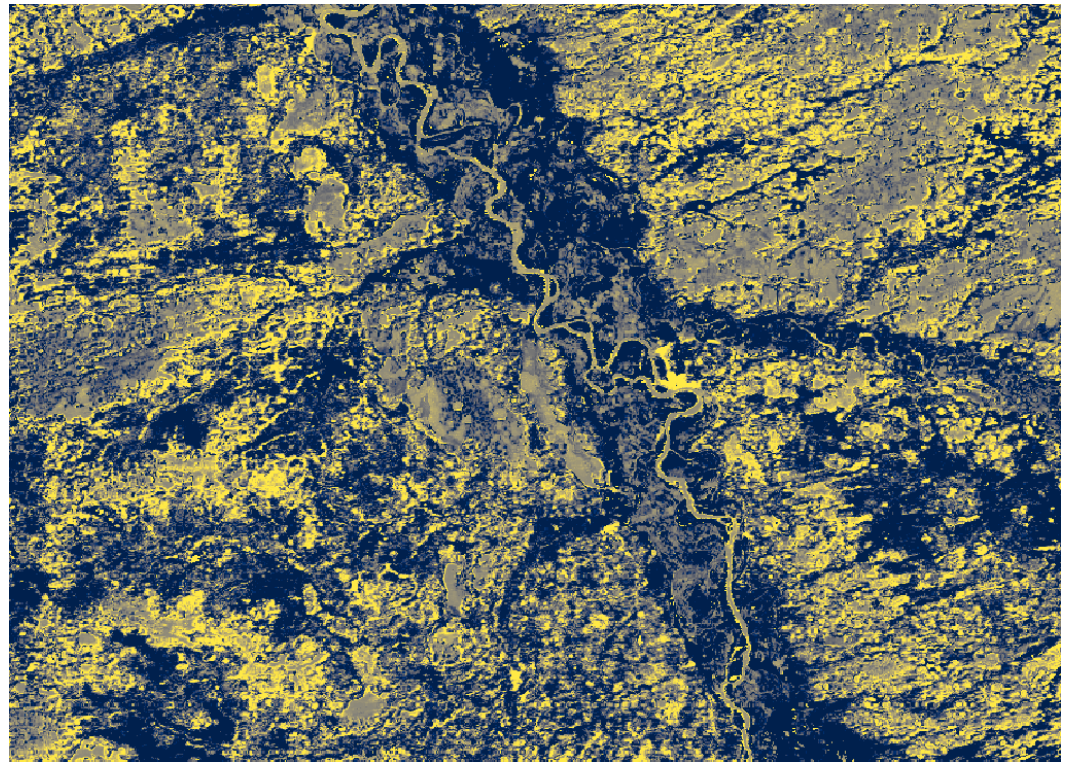
Cor = 0



Cor = 1



Разность между картой запаса и обучающей выборкой

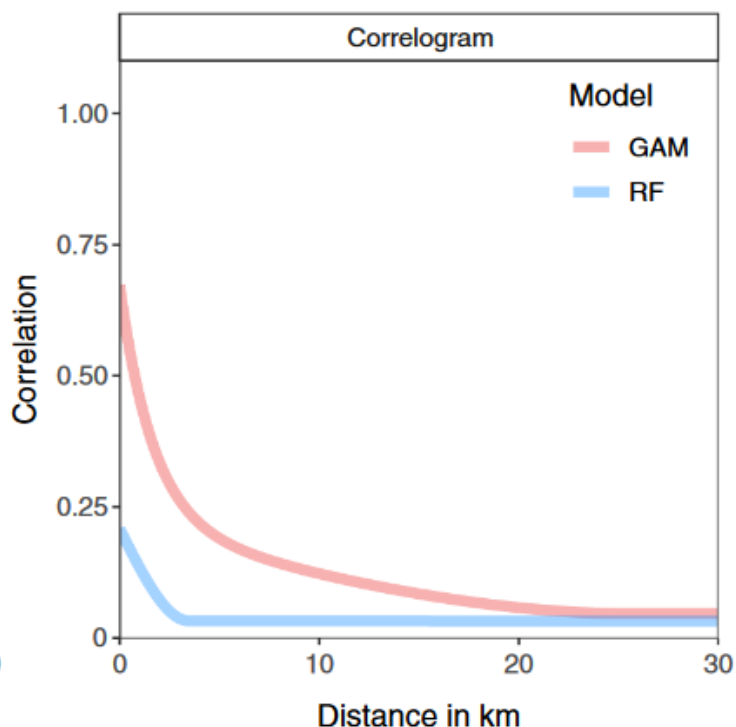


Неопределенность суммы - численная оценка

С учетом пространственных связей неопределенность можно описать следующим образом:

$$RMSE = \sqrt{\sum \sigma_i^2 + \sum \sum \sigma_i * \sigma_j * cor(i, j)}$$

Для корреляции можно предположить, что она зависит только от расстояния, и считать ее по вариограмме:



Для вычисления можно использовать численные методы вроде расчета по Монте-Карло на ограниченной части пикселей.

Wadoux A. M. J. -C., Heuvelink G. B. M. Uncertainty of spatial averages and totals of natural resource maps // Methods Ecol Evol. 2023. Т. 14. № 5. С. 1320–1332.

Можно использовать статистические подходы (сравнение с эталоном).

Olofsson P. и др. Good practices for estimating area and assessing accuracy of land change // Remote Sensing of Environment. 2014. Т. 148. С. 42–57.

Методы оценки погрешности - проблемы

Представленный ранее подход может использоваться для оценки погрешности, например, суммарного запаса древесины.

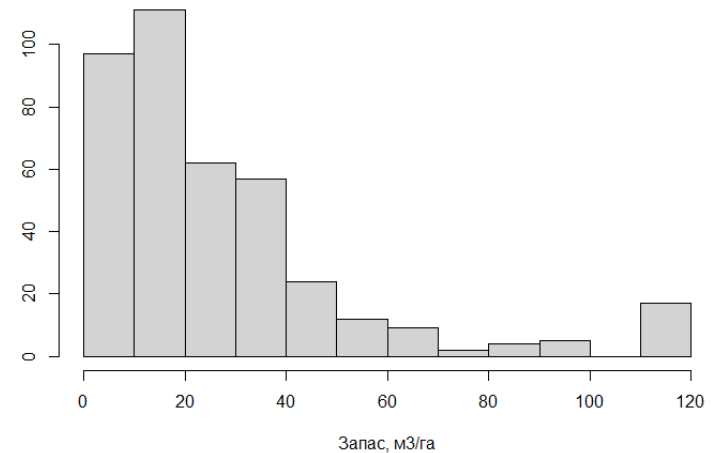
Но он не может использоваться для оценки погрешности динамики (поглощения) углерода лесов ввиду отсутствия обучающей выборки для этой характеристики.

Метод оценки погрешности суммы: фиксированные группы

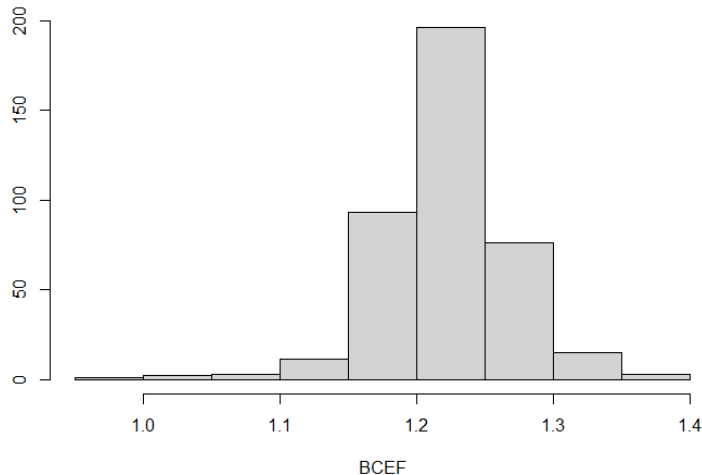
Можно выбрать определенные группы лесов, и считать, что внутри них погрешности связаны ($cor=1$), а между ними – независимы ($cor=0$)

Брались группы по породе, бонитету, относительной полноте и классу возраста. Далее по методу Монте-Карло случайно выбирались погрешности в пикселе по их распределениям.

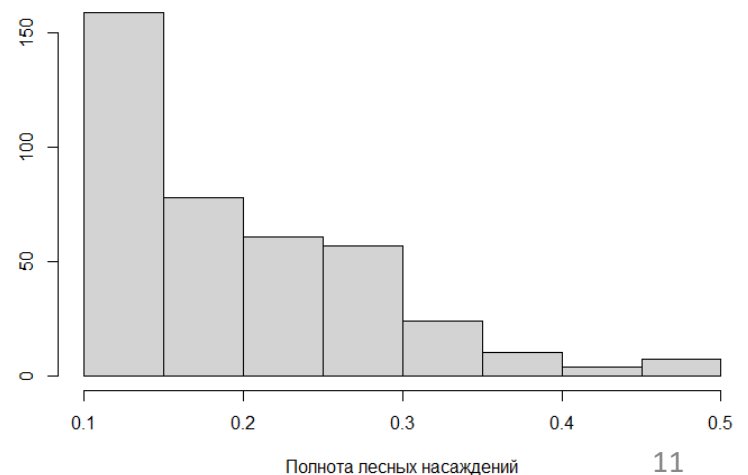
Вероятностное распределение запаса



Вероятностное распределение конверсионного коэффициента



Вероятностное распределение полноты

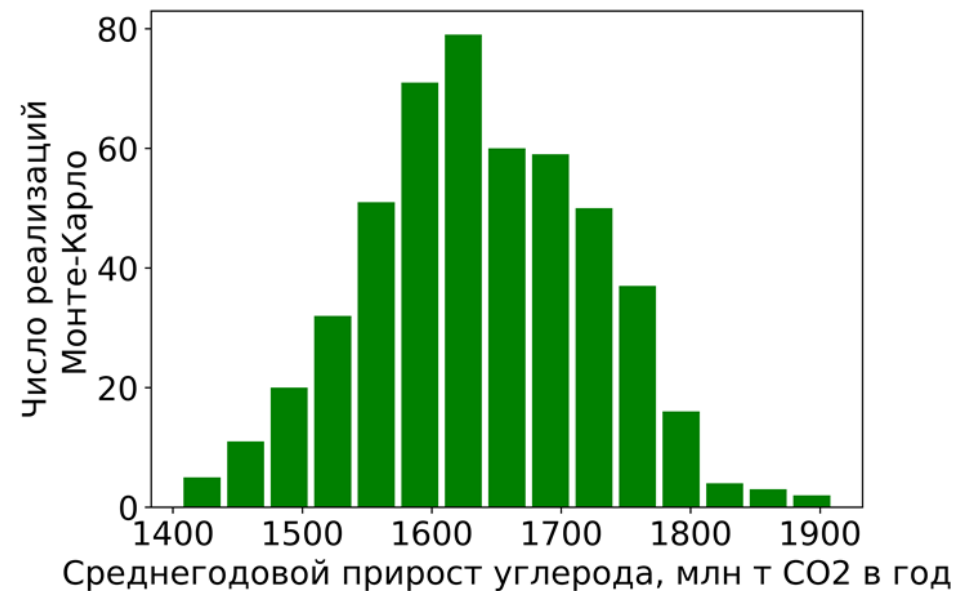
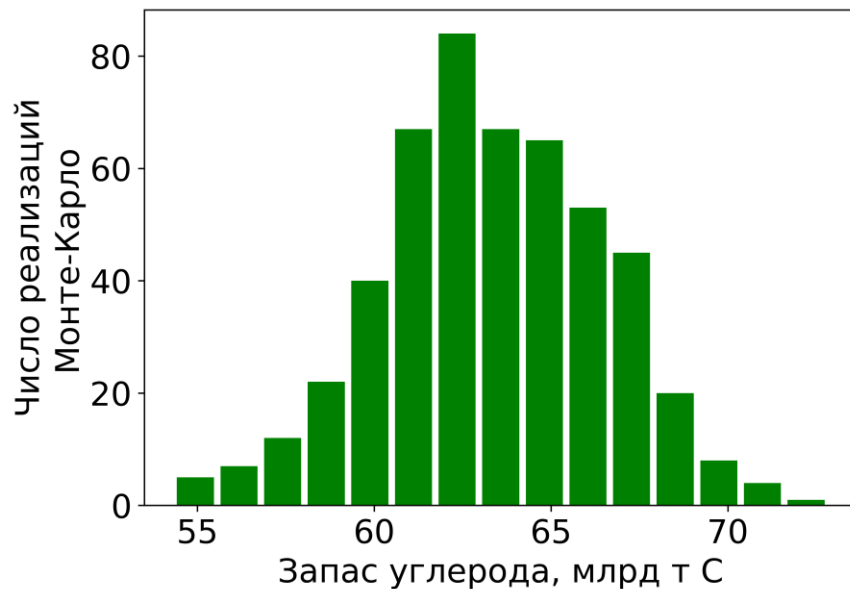


Метод оценки погрешности суммы: фиксированные группы

Такой подход позволяет получить адекватные оценки неопределенности

Но в данном подходе выбор групп является произвольным

	Запас углерода	Баланс углерода
Среднеквадратическое отклонение	3,1 млрд. Т С	87,8 млн. Т CO ₂ в год
Доверительный интервал 95%	56,9 – 69,1 млрд. Т С	1468 – 1801 млн. Т CO ₂ в год



Метод оценки суммы и средних: вариация выборки

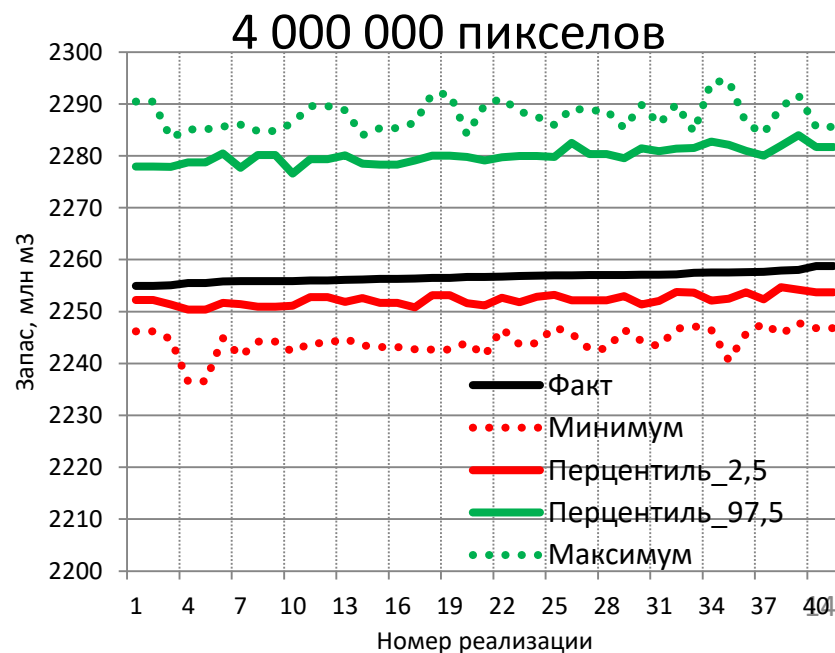
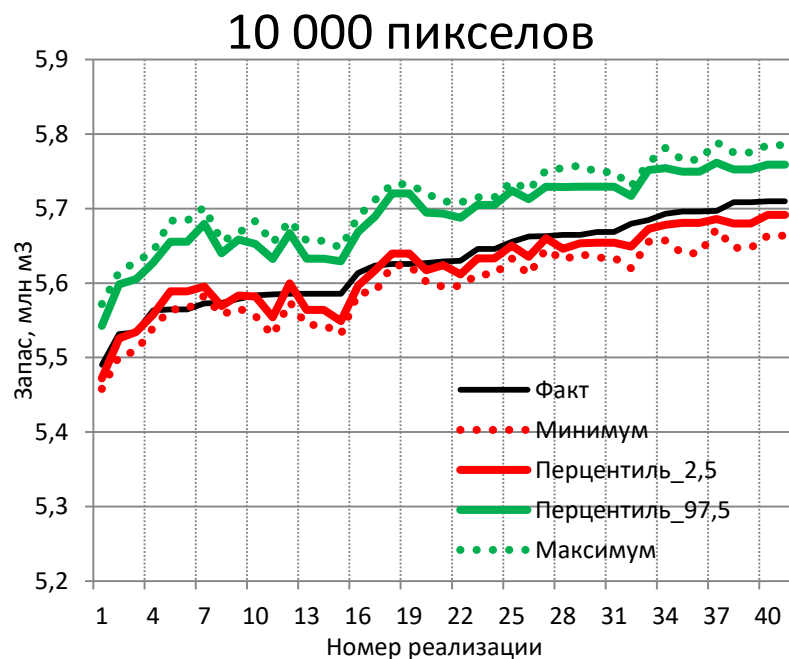
- 1) Выполнить множество итераций, на каждой из которых:
 - 1.1) Варьировать обучающую выборку, выбрав ее часть
 - 1.2) Обучить по ней регрессию и построить карту
- 2) На основе полученного множества карт построить гистограмму, стандартное отклонение, 95% интервал, как на уровне страны, так и для отдельных территорий.
- 3) Полученные карты также могут быть поданы на вход модуля расчета углерода для оценки его неопределенности

Анализ литературы ниже показывает, что этот метод позволяет учесть только часть погрешности, связанную с недостатком данных или знаний (epistemic, knowledge uncertainty).

Проверка метода вариации выборки

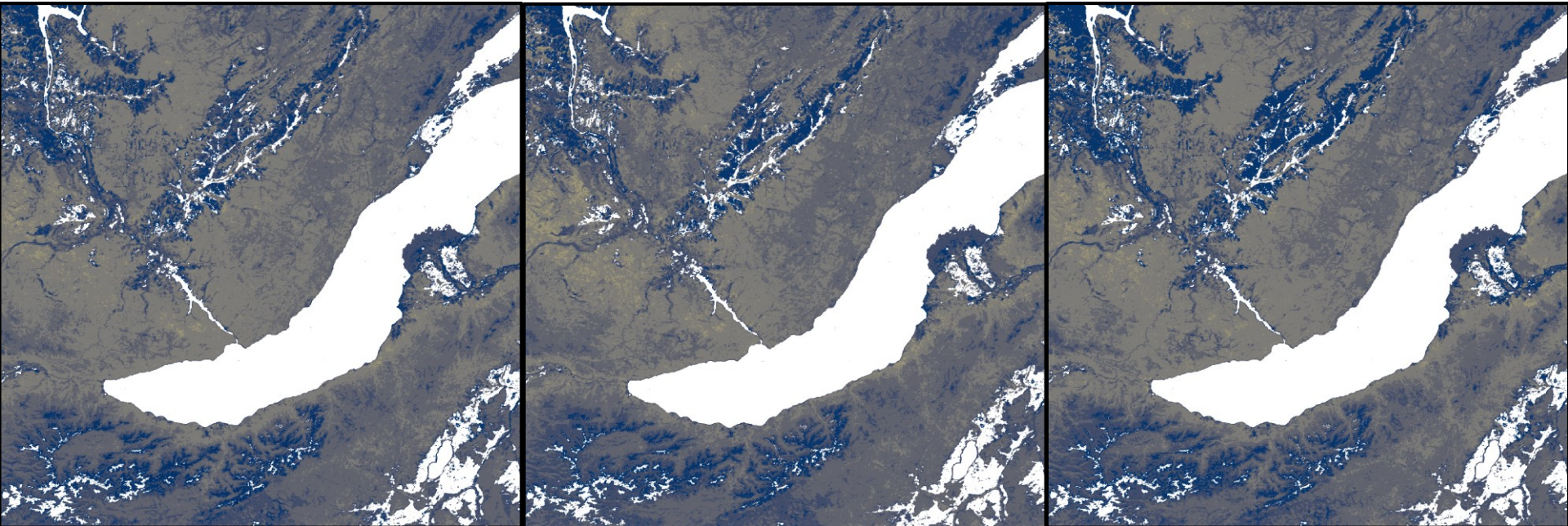
Адекватность подхода проверяется на примере запаса ствольной древесины следующим образом:

- 1) Берется случайная подвыборка из 4 000 000 пикселей, для нее выполняется 400 итераций, на каждой:
 - 1.1) Случайно выбирается обучающая выборка из 10 000 пикселей
 - 1.2) Строится регрессия, по ней оценивается запас для всех пикселей
 - 1.3) Сохраняется суммарный запас
- 2) По всем 400 итерациям строится 95% интервал, минимум, максимум и другие оценки, он сравнивается с фактическим запасом для 4000000 пикселей



Генерация множества карт характеристик лесов

На основе вариации выборки было сгенерировано множество карт ряда характеристик лесов: бонитета, возраста, относительной полноты, запаса стволовой древесины



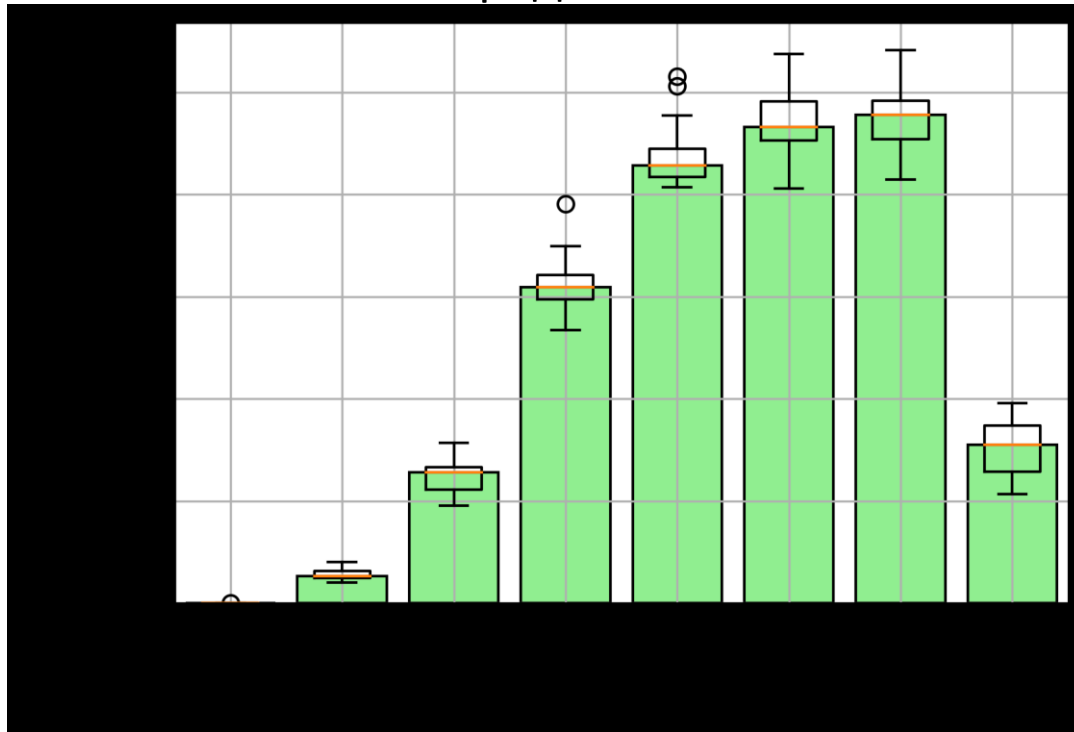
Класс бонитета



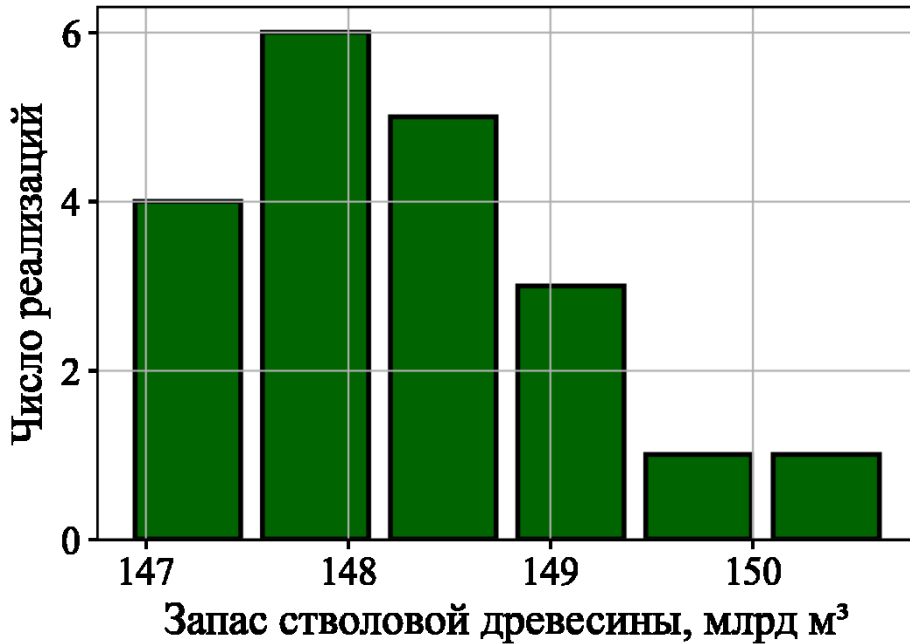
Оценка неопределенности характеристик лесов

Анализ ряда карт позволяет построить неопределенность характеристик леса

Гистограмма бонитетов по данным ДЗЗ и ее неопределенность

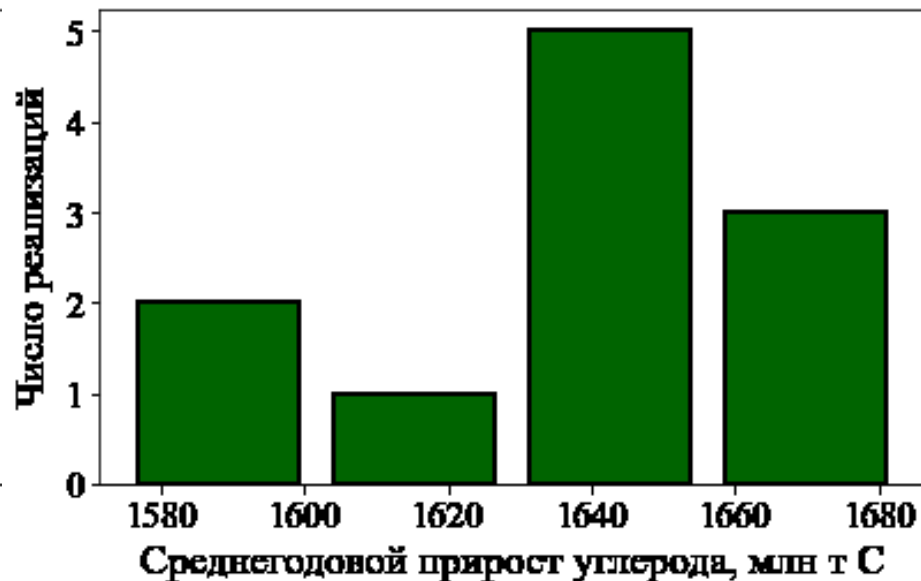
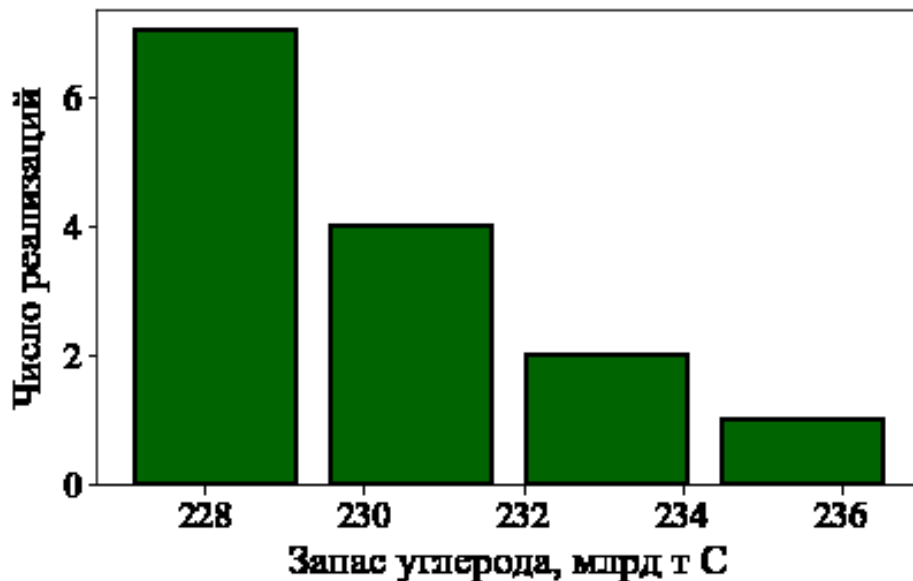


Оценка неопределенности характеристик лесов



Относительная погрешность	Управляемые	Резервные	Неучтенные	Всего
Запас стволовой древесины	0,4%	0,6%	0,4%	0,4%
Прирост за 2002-2023 годы	2,5%	1,0%	2,5%	2,0%

Оценка неопределенности запаса и баланса углерода лесов



Относительная погрешность	Управляемые	Резервные	Неучтенные	Всего
Запас углерода	1,0%	1,7%	1,5%	1,1%
Прирост за 2002-2023 годы	2,2%	2,5%	2,3%	2,1%

Выводы и дальнейшие шаги

Были получены оценки неопределенности ряда характеристик лесов и углерода.

Можно предположить, что предложенный подход не учитывает все источники погрешности.

Процесс генерации множества карт является трудоемким

Подключение данных тестовых полигонов и синтетических тестов может улучшить уверенность в оценках погрешности.

Нужно изучить возможность генерации эталона динамики запаса стволовой древесины на основе моделей хода роста.