



Министерство науки и высшего образования Российской Федерации
Федеральное государственное бюджетное образовательное учреждение
высшего образования
«Московский государственный технический университет
имени Н.Э. Баумана
(национальный исследовательский университет)»
(МГТУ им. Н.Э. Баумана)

ФАКУЛЬТЕТ ИНФОРМАТИКА И СИСТЕМЫ УПРАВЛЕНИЯ

КАФЕДРА СИСТЕМЫ ОБРАБОТКИ ИНФОРМАЦИИ И УПРАВЛЕНИЯ

РАСЧЕТНО-ПОЯСНИТЕЛЬНАЯ ЗАПИСКА К НАУЧНО-ИССЛЕДОВАТЕЛЬСКОЙ РАБОТЕ

НА ТЕМУ:

**Анализ точности получения данных Lidar для
вычисления возраста**

Студент ИУ5-31М
(Группа)

К.Р. Савченко
(Подпись, дата) (И.О.Фамилия)

Руководитель

Ю.Е. Гапанюк
(Подпись, дата) (И.О.Фамилия)

2023 г.

**Министерство науки и высшего образования Российской Федерации
Федеральное государственное бюджетное образовательное учреждение
высшего образования
«Московский государственный технический университет имени Н.Э. Баумана
(национальный исследовательский университет)»
(МГТУ им. Н.Э. Баумана)**

УТВЕРЖДАЮ
Заведующий кафедрой ИУ5
(Индекс)
В.И. Терехов
(И.О.Фамилия)
« 04 » сентября 2023 г.

З А Д А Н И Е
на выполнение научно-исследовательской работы

по теме Анализ точности получения данных Lidar для вычисления возраста

Студент группы ИУ5-31М

Савченко Ксения Руслановна
(Фамилия, имя, отчество)

Направленность НИР (учебная, исследовательская, практическая, производственная, др.)
ИССЛЕДОВАТЕЛЬСКАЯ

Источник тематики (кафедра, предприятие, НИР) КАФЕДРА

График выполнения НИР: 25% к нед., 50% к нед., 75% к нед., 100% к нед.

Техническое задание Проанализировать возможность использования Lidar для сбора данных о деревьях на примере определения возраста. Описать рекомендации по сбору данных, сделать заключение

Оформление научно-исследовательской работы:

Расчетно-пояснительная записка на 18 листах формата А4.

Перечень графического (иллюстративного) материала (чертежи, плакаты, слайды и т.п.)

Дата выдачи задания « 04 » сентября 2023 г.

Руководитель НИР

(Подпись, дата)

Ю.Е. Гапанюк

(И.О.Фамилия)

Студент

(Подпись, дата)

К.Р. Савченко

(И.О.Фамилия)

Примечание: Задание оформляется в двух экземплярах: один выдается студенту, второй хранится на кафедре.

Оглавление

| | |
|---|----|
| 1. Введение..... | 4 |
| 2. Основная часть | 6 |
| Характеристика датасета..... | 6 |
| Вычисление возраста дерева в зависимости от породы..... | 11 |
| 3. Рекомендации | 16 |
| 4. Заключение | 17 |
| 5. Список использованных источников | 18 |

Введение

В современном мире растущего осознания экологических проблем и необходимости сохранения природных ресурсов, вопросы, связанные с устойчивым использованием лесов и сохранением их биоразнообразия становятся все более актуальными. Одним из ключевых аспектов в этом контексте является понимание возраста деревьев, что в свою очередь позволяет не только оценить их состояние, но и принимать обоснованные решения относительно их восстановления и сохранения.

Но как собрать большой объем данных о лесе? Технология Lidar [1] может в этом помочь. Преимущества:

- **Высокая точность:** Lidar позволяет получить очень точные данные о форме, размере и расположении деревьев, что может быть полезно для оценки их состояния и определения необходимых мер по уходу.
- **Быстрота:** Lidar может сканировать большие площади за короткий промежуток времени, что позволяет собрать данные о большом количестве деревьев за относительно короткий период времени.
- **Неинвазивность:** Lidar не требует физического контакта с деревьями, что снижает риск повреждения их стволов или веток.

Вычисление точного возраста дерева, не срубая его – далеко не простая задача, так как прирост дерева может меняться от года к году. На рост деревьев влияют множество факторов [2], таких как:

- **Почвенные условия:** тип почвы, ее плодородие, влажность и кислотность влияют на рост деревьев.
- **Климатические условия:** температура, влажность, количество осадков и ветер могут оказывать влияние на рост и развитие деревьев.
- **Генетические особенности:** разные виды деревьев имеют различную скорость роста и требования к условиям произрастания.

- Болезни и вредители: они могут замедлить рост дерева или даже привести к его гибели.
- Человеческий фактор: вырубка лесов, изменение ландшафта, загрязнение окружающей среды могут негативно сказаться на росте деревьев.

Почвенные и климатические условия и генетические особенности у деревьев одного леса могут во многом совпадать. В то время как болезни и вредители – непредсказуемые события. Узнать, болеет ли дерево, можно только после его тщательной проверки экспертом.

Использование Lidar для сбора данных о деревьях может повысить точность и скорость измерения их характеристик по сравнению с традиционными методами

Данная исследовательская работа направлена на анализ возможности использования Lidar для сбора данных о деревьях и определения возраста деревьев, в целях повышения точности и скорости измерения их характеристик по сравнению с традиционными методами.

Основная часть

Характеристика датасета

В качестве набора данных мы будем использовать данные собранные с помощью Lidar и преобразованные в табличные данные.

Этот набор данных состоит из данных о деревьях определённого участка леса.

Файл содержит колонки:

- Id - номер дерева, (ГИС/прибит на стволе);
- X - координата дерева по X;
- Y - координата дерева по Y;
- Name - имя файла;
- Diameter_LSMIn, cm - диаметр ствола по методу LSMIn;
- Diameter_LS, cm - диаметр по методу LS;
- Diameter_HLS, cm - диаметр по методу HLS;
- Height, m - высота ствола дерева;
- Length, m - длина дерева;
- Crown_volume, m³ - объем выпуклой оболочки кроны;
- Crown_square, m² - площадь выпуклой оболочки кроны;
- XY_crown_square, m² - площадь проекции выпуклой оболочки кроны на XY;
- XZ_crown_square, m² - площадь проекции выпуклой оболочки кроны на XZ;
- YZ_crown_square, m² - площадь проекции выпуклой оболочки кроны на YZ;
- X_UP_1 - координаты вершины дерева по X;
- Y_UP_1 - координаты вершины дерева по Y;
- Quality - качество сегментации дерева;
- порода – порода дерева;

- Окр. - Полевые измерения окружности ствола;
- Д1,3 - диаметр ствола на высоте 1,3м, высчитывается из Окр.;
- примеч. – дополнительный комментарий по дереву.

Импортируем библиотеки с помощью команды `import`. Как правило, все команды `import` размещают в первой ячейке ноутбука, но мы в этом примере будем подключать все библиотеки последовательно, по мере их использования.

```
import numpy as np
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
%matplotlib inline
sns.set(style="ticks")
```

Загрузим файлы датасета с помощью библиотеки Pandas.

```
# Будем анализировать данные файла all_loc1_p.xlsx
data = pd.read_excel('all_loc1_p.xlsx')
```

Основные характеристики датасета

```
# Первые 5 строк датасета
data.head()
```

| | Id | Veigth | X | Y | Name | Diameter_LSMIn, cm | Diameter_LS, cm | Diameter_HLS, cm | Height, m |
|---|----|--------|---------|--------|---------------|--------------------|-----------------|------------------|-----------|
| 0 | 1 | 1 | -81.824 | 11.979 | tree_0002.pcd | 38.080000 | 17.63 | 19.24 | 14.8572 |
| 1 | 2 | 1 | -81.783 | 12.786 | tree_0001.pcd | 45.460000 | 21.79 | 22.08 | 18.5908 |
| 2 | 3 | 1 | -80.424 | 11.567 | tree_0003.pcd | 29.925845 | 17.40 | 17.88 | 16.2204 |
| 3 | 4 | 1 | -78.511 | 10.504 | tree_0007.pcd | 17.412436 | 14.22 | 16.35 | 17.3342 |
| 4 | 5 | 0 | -73.810 | 11.913 | tree_0042.pcd | 44.420000 | 18.62 | 19.25 | 28.5284 |

5 rows × 28 columns

```
# Размер датасета
data.shape
```

(624, 28)

```
total_count = data.shape[0]
print('Всего строк: {}'.format(total_count))
```

Всего строк: 624

```
# Список колонок
data.columns
```

```
Index(['Id', 'Veigth', 'X', 'Y', 'Name', 'Diameter_LSMIn, cm',
      'Diameter_LS, cm', 'Diameter_HLS, cm', 'Height, m', 'Length, m',
      'Crown_volume, m3', 'Crown_square, m2', 'XY_crown_square, m2',
      'XZ_crown_square, m2', 'YZ_crown_square, m2', 'X_UP_0', 'Y_UP_0',
      'X_UP_1', 'Y_UP_1', 'X_UP_2', 'Y_UP_2', 'X_UP_3', 'Y_UP_3', 'Quality',
      'порода', 'Окр.', 'Д1,3', 'примеч.'],
      dtype='object')
```

```
# Список колонок с типами данных
data.dtypes
```

| | |
|---------------------|---------|
| Id | int64 |
| Veigth | int64 |
| X | float64 |
| Y | float64 |
| Name | object |
| Diameter_LSMIn, cm | float64 |
| Diameter_LS, cm | float64 |
| Diameter_HLS, cm | float64 |
| Height, m | float64 |
| Length, m | float64 |
| Crown_volume, m3 | float64 |
| Crown_square, m2 | float64 |
| XY_crown_square, m2 | float64 |
| XZ_crown_square, m2 | float64 |
| YZ_crown_square, m2 | float64 |
| X_UP_0 | float64 |
| Y_UP_0 | float64 |
| X_UP_1 | float64 |
| Y_UP_1 | float64 |
| X_UP_2 | float64 |
| Y_UP_2 | float64 |
| X_UP_3 | float64 |
| Y_UP_3 | float64 |
| Quality | int64 |
| порода | object |
| Окр. | float64 |
| Д1,3 | float64 |
| примеч. | object |
| dtype: | object |


```
# Проверим наличие пустых значений
# Цикл по колонкам датасета
for col in data.columns:
    # Количество пустых значений - все значения заполнены
    temp_null_count = data[data[col].isnull()].shape[0]
    print('{} - {}'.format(col, temp_null_count))
```

```
Id - 0
Veigth - 0
X - 0
Y - 0
Name - 0
Diameter_LSMIn, cm - 0
Diameter_LS, cm - 0
Diameter_HLS, cm - 0
Height, m - 0
Length, m - 0
Crown_volume, m3 - 0
Crown_square, m2 - 0
XY_crown_square, m2 - 0
XZ_crown_square, m2 - 0
YZ_crown_square, m2 - 0
X_UP_0 - 0
Y_UP_0 - 0
X_UP_1 - 0
Y_UP_1 - 0
X_UP_2 - 0
Y_UP_2 - 0
X_UP_3 - 0
Y_UP_3 - 0
Quality - 0
порода - 0
Окр. - 0
Д1,3 - 0
примеч. - 575
```

```
# Основные статистические характеристики набора данных
data.describe()
```

| | Id | Veigth | X | Y | Diameter_LSMIn, cm | Diameter_LS, cm | Diameter_HLS, cm | Height, m | Length, m | Crown_volume, m3 |
|-------|--------------|------------|-------------|------------|--------------------|-----------------|------------------|------------|------------|------------------|
| count | 624.000000 | 624.000000 | 624.000000 | 624.000000 | 624.000000 | 624.000000 | 624.000000 | 624.000000 | 624.000000 | 624.000000 |
| mean | 608.328526 | 0.097756 | -48.228715 | 42.462213 | 21.100849 | 15.345737 | 16.681747 | 20.386347 | 20.502025 | 260.883955 |
| std | 1899.194714 | 0.447997 | 23.960367 | 22.715228 | 10.282168 | 3.527443 | 3.686223 | 8.510649 | 8.461428 | 235.897239 |
| min | 1.000000 | 0.000000 | -100.632000 | 1.602000 | 3.607999 | 0.000000 | 0.000000 | 0.000000 | 0.000000 | 0.000000 |
| 25% | 156.750000 | 0.000000 | -66.639500 | 22.815500 | 12.597862 | 13.235000 | 14.535000 | 12.492275 | 12.686395 | 64.854260 |
| 50% | 313.500000 | 0.000000 | -46.461500 | 42.291000 | 20.763215 | 15.835000 | 17.265000 | 24.650750 | 24.736265 | 188.347980 |
| 75% | 469.250000 | 0.000000 | -29.157000 | 62.396500 | 27.948422 | 17.532500 | 18.975000 | 28.105275 | 28.170127 | 401.555540 |
| max | 20301.000000 | 3.000000 | -0.319000 | 86.804000 | 55.500000 | 27.450000 | 39.660000 | 31.953300 | 31.962830 | 1720.222590 |

| X_UP_3 | Y_UP_3 | Quality | Окр. | Д1,3 |
|-------------|------------|------------|------------|------------|
| 624.000000 | 624.000000 | 624.000000 | 624.000000 | 624.000000 |
| -48.130309 | 42.676021 | 0.798077 | 68.414904 | 21.788186 |
| 23.973321 | 22.691016 | 0.401757 | 33.843417 | 10.778158 |
| -101.273003 | -0.278900 | 0.000000 | 16.200000 | 5.159236 |
| -66.548948 | 23.058425 | 1.000000 | 38.475000 | 12.253185 |
| -46.586451 | 41.816000 | 1.000000 | 67.050000 | 21.353503 |
| -28.582325 | 62.331050 | 1.000000 | 94.775000 | 30.183121 |
| -0.071500 | 90.592003 | 1.000000 | 188.600000 | 60.063694 |

```
# Определим уникальные значения
data['порода'].unique()
```

```
array(['Е', 'ОС', 'С', '??', 'Б', 'Ряб', 'РЯБ', 'Д', 'ЛП', 'Яб', 'Лещ'],
      dtype=object)
```

- Е - ель
- ОС - осина
- С - сосна
- ??
- Б - берёза
- Ряб - рябина
- Д - дуб
- ЛП - липа
- Яб - яблоня
- Лещ - лещина (орешник)

```
for i in data['порода'].unique():
    print('{} - {}'.format(i, len(data[data['порода']==i])))
```

```
Е - 252
ОС - 8
С - 185
?? - 1
Б - 116
Ряб - 32
РЯБ - 6
Д - 20
ЛП - 1
Яб - 2
Лещ - 1
```

```
# Изменим РЯБ на Ряб, чтобы избежать избытка данных
for i in range(data.shape[0]):
    if data.iloc[i]['порода']=='РЯБ':
        data.loc[i,'порода']='Ряб'
```

```
data['порода'].unique()
```

```
array(['Е', 'ОС', 'С', '??', 'Б', 'Ряб', 'Д', 'ЛП', 'Яб', 'Лещ'],
      dtype=object)
```

```
data.iloc[40]['порода']
```

```
'Ряб'
```

```
for i in data['порода'].unique():
    print('{} - {}'.format(i, len(data[data['порода']==i])))
```

```
Е - 252
ОС - 8
С - 185
?? - 1
Б - 116
Ряб - 38
Д - 20
ЛП - 1
Яб - 2
Лещ - 1
```

Вычисление возраста дерева в зависимости от породы

Проведем анализ данных в датасете, чтобы лучше понять характер и свойства имеющихся данных. Для определения возраста нам будет важно знать породу дерева, длину ствола, диаметр ствола.

Так как есть породы, численное количество деревьев которых мало (диаграмма 1), поэтому для вычислений возьмём ель (252), сосну (185), берёзу (116) и рябину (38).

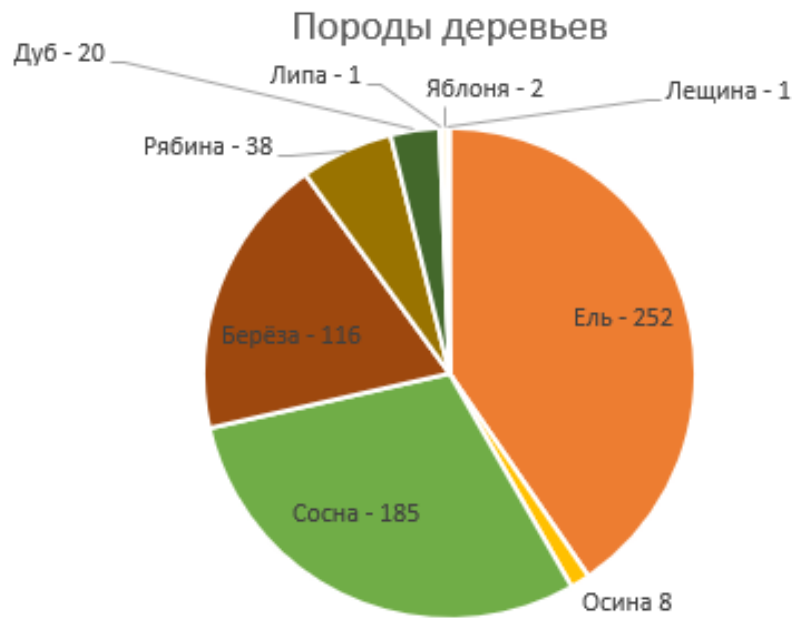


Диаграмма 1. Численное соотношение деревьев по породам.

Чтобы проверить эффективность использования данного датасета, попробуем определить возраст деревьев.

Для вычисления возраста воспользуемся двумя способами и сравним их.

1. Вычисление возраста по высоте дерева

В среднем, деревья растут до 50 см в год [3]. Данную информацию и будем использовать для вычисления возраста.

2. Вычисление возраста по диаметру ствола дерева

Средний прирост дерева в диаметре немало зависит от его породы. Проведя исследование, было вычислено, что прирост в год ели и сосны 0.9 см, рябины 0.5 см, берёзы 1.5 см [4].

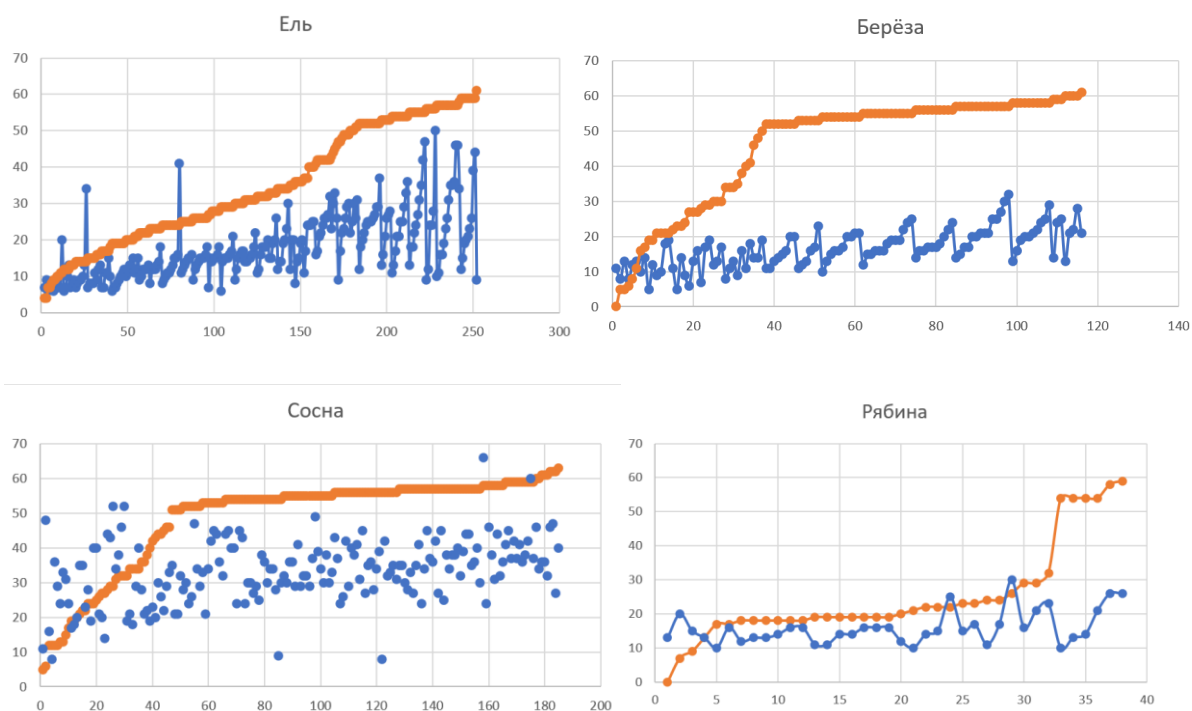
Итого были получены такие данные.

```
df[["age_height", "age_dia"]]
```

| | age_height | age_dia |
|-----|------------|---------|
| 0 | 29.0 | 21.0 |
| 1 | NaN | NaN |
| 2 | 32.0 | 38.0 |
| 3 | 34.0 | 23.0 |
| 4 | NaN | NaN |
| ... | ... | ... |
| 619 | 18.0 | 13.0 |
| 620 | 13.0 | 13.0 |
| 621 | 54.0 | 13.0 |
| 622 | 23.0 | 5.0 |
| 623 | 16.0 | 9.0 |

624 rows × 2 columns

Для визуализации построим графики по каждой породе, понимая, что при идеальных условиях графики должны накладываться друг на друга.



По оси X – индекс, Y – возраст.

Оранжевый график – возраст, вычисленный по высоте ствола, синий график – возраст, вычисленный по диаметру ствола.

В имеющейся таблице у нас ещё есть столбцы «Quality» и «примеч.», отвечающие за качество собранных Lidar данных и полезные примечания.

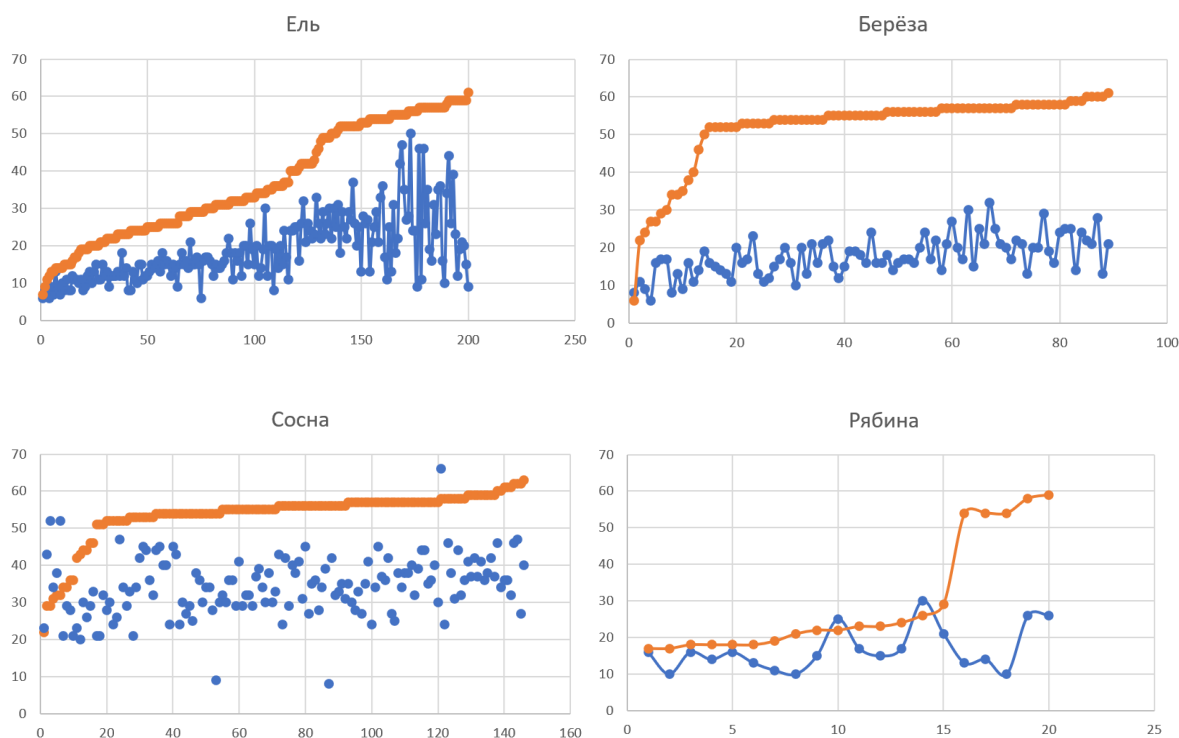
Если со столбцом, отвечающим за качество, всё понятно, то какие же примечания могут быть?

```
data['примеч.'].unique()

array([nan, 'остолоп', 'сух.', 'она же 231, двойной номер', 'СУХ.',
       'ПАСЫНОК - ???', 'ДВУХВЕРШИННАЯ', 'сухостой, наклонная к Ю',
       'наклонная к СЗ', 'без вершины',
       'кора обглодана на 80%, потенциальный сухостой',
       'наклонный в сторону окна', 'наклонная'], dtype=object)
```

Подводя итог, все примечания сигнализируют о «дефекте» дерева, то есть его возраст будет неправильно вычислен методами, описанными выше.

Исключим деревья с Quality = 0 и имеющие примечания и построим графики.



Получается, что экстремумы синих графиков, которые находятся над оранжевым графиком, присутствовали в основном из-за неправильных измерений Lidar или «дефектов» дерева. Да, такие экстремумы ещё остались, их можно объяснить «дефектами» дерева, которые не были учтены. В перспективе, данные деревья должны быть обследованы, для их диагностики.

Рекомендации

Для повышения точности определения возраста с помощью Lidar данных необходимо собирать данные повторно через определённый промежуток времени, например, через год, 5 лет. Данное действие необходимо для отслеживания изменений в структуре и состоянии деревьев. Это позволяет определить динамику роста и развития деревьев, а также выявить возможные проблемы, такие как заболевания или вредители, на ранних стадиях. Кроме того, повторный сбор данных позволяет оценить эффективность мер по уходу за деревьями и при необходимости внести коррективы в планы по их обслуживанию.

Заключение

Таким образом, можно сказать, что использование Lidar для сбора данных о деревьях является перспективным направлением, которое может значительно улучшить точность и скорость измерения характеристик деревьев по сравнению с традиционными методами. Преимущества Lidar включают высокую точность, быстроту и неинвазивность.

Однако, для полного раскрытия потенциала Lidar необходимо провести дополнительные исследования и разработки, чтобы улучшить качество данных и снизить стоимость технологии.

Тем не менее, результаты данной курсовой работы показывают, что Lidar может стать важным инструментом для мониторинга и управления лесными ресурсами, обеспечивая более эффективное и устойчивое использование лесных экосистем в будущем.

Список использованных источников

1. Лидар. URL: <https://ru.wikipedia.org/wiki/Лидар> (дата обращения: 05.12.23)
2. Какие факторы влияют на рост и цветение растений? URL: <https://www.baconline.ru /центр-знаний/блог/3638-kakie-factory-vliaut-na-rost-i-cvetenie-rastenij> (дата обращения: 08.12.23)
3. Таксация леса: теоретические основы вычислений: учебное пособие для подготовки бакалавров по направлению 250100.62 "Лесное дело" / Г.В. Матусевич, Л.В.Стоноженко, Н.Г. Иванов и др.; под общ. ред. Л.В.Стоноженко. – М. : ФГБОУ ВПО МГУЛ, 2012. – 182 с.
4. Оценка продуктивности древостоев / Д. В. Тишин. – Казань: Казанский университет, 2011. – 31 с.