**МОЛДАВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ**

**ФАКУЛЬТЕТ МАТЕМАТИКИ И ИНФОРМАТИКИ**

**ДЕПАРТАМЕНТ ИНФОРМАТИКИ**

Анализ и визуализация атмосферных данных с использованием Службы мониторинга атмосферы Copernicus (CAMS)

**Индивидуальная работа номер 3**

Доцент университета: Поятэ Анатолий

Автор: студент группы IASD2401 (РУ), Жураковски Владислав

**Кишинёв, 2024**

**СОДЕРЖАНИЕ**

[I. Доступ и загрузка данных 3](#_Toc11796)

[II. Очистка и предобработка данных 3](#_Toc30084)

[III. Объединение и подготовка данных 4](#_Toc14091)

[III. Анализ данных 4](#_Toc30180)

[3.1. Вычисление базовой статистики для поллютантов 4](#_Toc3647)

[3.2. Идентификация тенденций во времени для выбранного поллютанта 5](#_Toc26860)

[3.3. Сравнение уровней загрязнения в разных регионах. 5](#_Toc18010)

[IV. Визуализация данных 6](#_Toc7074)

[V. Интерпретация результатов 6](#_Toc28995)

# I. Доступ и загрузка данных

**Задача**: Получить данные из Службы мониторинга атмосферы Коперника (CAMS) для анализа атмосферных поллютантов.

**Шаги**:

1. Использовался API библиотеки cdsapi для доступа к данным.
2. Были написаны три отдельных скрипта для скачивания данных:
   1. Радиационное воздействие аэрозолей с излучением (radiative\_forcing\_of\_aerosol\_radiation\_interactions).
   2. Радиационное воздействие углекислого газа (radiative\_forcing\_of\_carbon\_dioxide).
   3. Радиационное воздействие метана (radiative\_forcing\_of\_methane).
3. Параметры запроса включали:
   1. Тип воздействия: мгновенное (instantaneous).
   2. Полоса частот: длинноволновая и коротковолновая (long\_wave, short\_wave).
   3. Тип неба: всё небо (all\_sky).
   4. Уровень: поверхность (surface).
   5. Годы: 2003–2018.
   6. Месяцы: все месяцы года.
4. Скачанные данные были сохранены в формате NetCDF.

## II. Очистка и предобработка данных

**Задача**: Преобразовать скачанные данные в таблицу с понятным форматом, выбрать данные для заданного региона и времени.

**Шаги**:

1. Написан код на Python для обработки файлов NetCDF. Извлечение данных осуществлялось с использованием библиотеки netCDF4, выбирались только те данные, которые относятся к заданной географической области (Египет и Судан).
2. Для Египта были заданы географические границы в виде полигона: Египет: [(24.7, 22.0), (35.0, 22.0), (35.0, 31.5), (24.7, 31.5)]. Фильтровались только те координаты, которые попадали внутрь полигона. Сформирована таблица со столбцами: Время — временная метка; Координаты Египта (широта, долгота) — географические координаты; значения поллютантов (например, Радиационное воздействие аэрозолей с излучением, углекислого газа, метана).
3. Для Судана аналогично были заданы географические границы: Судан: [(21.0, 8.5), (39.0, 8.5), (39.0, 22.0), (21.0, 22.0)]. Фильтровались данные только внутри полигона. Сформирована таблица аналогичной структуры, включающая поллютанты

# III. Объединение и подготовка данных

**Задача**: Создать финальные таблицы для анализа данных по Египту и Судану.

**Шаги**:

1. Таблицы для Египта и Судана формировались с использованием функций, где извлекаются и фильтруются данные из NetCDF.
2. Каждая таблица содержит три поллютанта: Радиационное воздействие аэрозолей с излучением, Радиационное воздействие углекислого газа, Радиационное воздействие метана.
3. Обе таблицы сохраняются в формате CSV и готовы для последующего анализа.

# III. Анализ данных

3.1. Вычисление базовой статистики для поллютантов  
Код выполняет анализ данных для определения основных статистических характеристик уровня загрязнения в зависимости от выбранного поллютанта (например, аэрозоли, метан или углекислый газ). Для этого используется функция calculate\_statistics, которая вычисляет следующие показатели:

* Среднее значение (mean);
* Медиана (median);
* Стандартное отклонение (std);
* Минимальное и максимальное значения.

Эти статистические данные предоставляют информацию о центральной тенденции и вариации данных.

3.2. Идентификация тенденций во времени для выбранного поллютанта  
Для анализа временных рядов используется функция analyze\_time\_series. Она группирует данные по временной оси и строит график средней концентрации поллютанта за указанный период.  
Ключевые этапы:

1. Преобразование столбца с датами в формат datetime.
2. Группировка данных по времени и вычисление средних значений.
3. Визуализация графика тенденций для анализа изменений уровня загрязнения с течением времени.

3.3. Сравнение уровней загрязнения в разных регионах.  
Для географического анализа данных реализована функция heatmap\_pollution.  
Этапы:

1. Данные фильтруются по широте и долготе, после чего создается сводная таблица, где:
   * Столбцы представляют долготу;
   * Строки — широту;
   * Значения в ячейках — средняя концентрация поллютанта.
2. Построение тепловой карты с использованием библиотеки seaborn, чтобы визуализировать концентрации загрязнений на географической карте.  
   Это помогает выявить регионы с наибольшим уровнем загрязнения.

**Дополнительное сравнение:**  
Функция bar\_chart\_pollutants строит диаграмму баров, сравнивающую уровни загрязнения в разных регионах. Это используется для количественного анализа различий в уровнях поллютантов между географическими точками.

# IV. Визуализация данных

**Временные ряды:**  
Для каждого поллютанта создается график, показывающий изменение уровня загрязнения с течением времени. Это позволяет выявить:

* Тренды;
* Сезонные изменения;
* Аномальные значения.

**Тепловая карта:**  
Отображение данных на географической карте с использованием тепловой визуализации позволяет увидеть, как концентрация загрязнений варьируется по регионам. Например, тепловая карта уровня углекислого газа в Египте помогает определить наиболее загрязненные районы.

**Гистограммы и диаграммы баров:**  
Сравнительные гистограммы используются для анализа разницы уровней загрязнений между регионами, например, между Египтом и Суданом. Эти графики помогают определить области с более высоким воздействием поллютантов.

# V. Интерпретация результатов

* **Тенденции:** Графики временных рядов показывают, как уровни загрязнений менялись в течение 2003–2018 годов. Например, можно определить рост концентрации углекислого газа в Египте или сезонные колебания метана.
* **Аномалии:** Тепловая карта и временные ряды позволяют быстро обнаружить неожиданные всплески концентрации поллютантов в определенные периоды или регионы.
* **Политика и осведомленность:**  
  Данные могут быть использованы для:
  + Разработки экологических мер по снижению загрязнений в наиболее пострадавших регионах;
  + Увеличения осведомленности населения о критических зонах с высоким уровнем загрязнения;
  + Укрепления международного сотрудничества для борьбы с трансграничными загрязнениями.

Этот анализ служит основой для построения политики охраны окружающей среды и разработки мероприятий для улучшения качества воздуха в рассматриваемых регионах.