

Система усвоения данных для оценки эмиссии парниковых газов на основе спутниковых данных и глобальной транспортной модели

Платонова М.В., Котлер В.Д. ,
Климова Е.Г.

ФИЦ ИВТ СО РАН



SDM-2025

Оценка эмиссии парниковых газов:

- Разработка современной системы усвоения реальных спутниковых данных для оценки эмиссии парниковых газов очень важна ввиду ряда экологических и экономических факторов.
- Используется подход, основанный на применении систем усвоения данных с привлечением математических моделей распространения пассивных примесей в атмосфере.
- Под усвоением данных принято понимать совместное использование данных моделирования и данных наблюдений для получения оптимальной оценки состояния моделируемого процесса.

Постановка задачи: специфика

- Разработка алгоритма усвоения данных для трехмерной модели распространения примеси в атмосфере с использованием спутниковых данных
- В экспериментах используется модель MOZART-4, в качестве данных наблюдений использованы спутниковые данные AIRS (Aqua L2 Standard Physical Retrieval)
- Производится сравнение получаемых оценок эмиссии метана с данными из базы реанализа CAMS global greenhouse gas reanalysis (EGG4)

Оценка значений приземной концентрации (математическая модель)

The logo for NCAR | UCAR, featuring the text "NCAR" above "UCAR" in white on a dark blue background, with a vertical line to the right of the text.

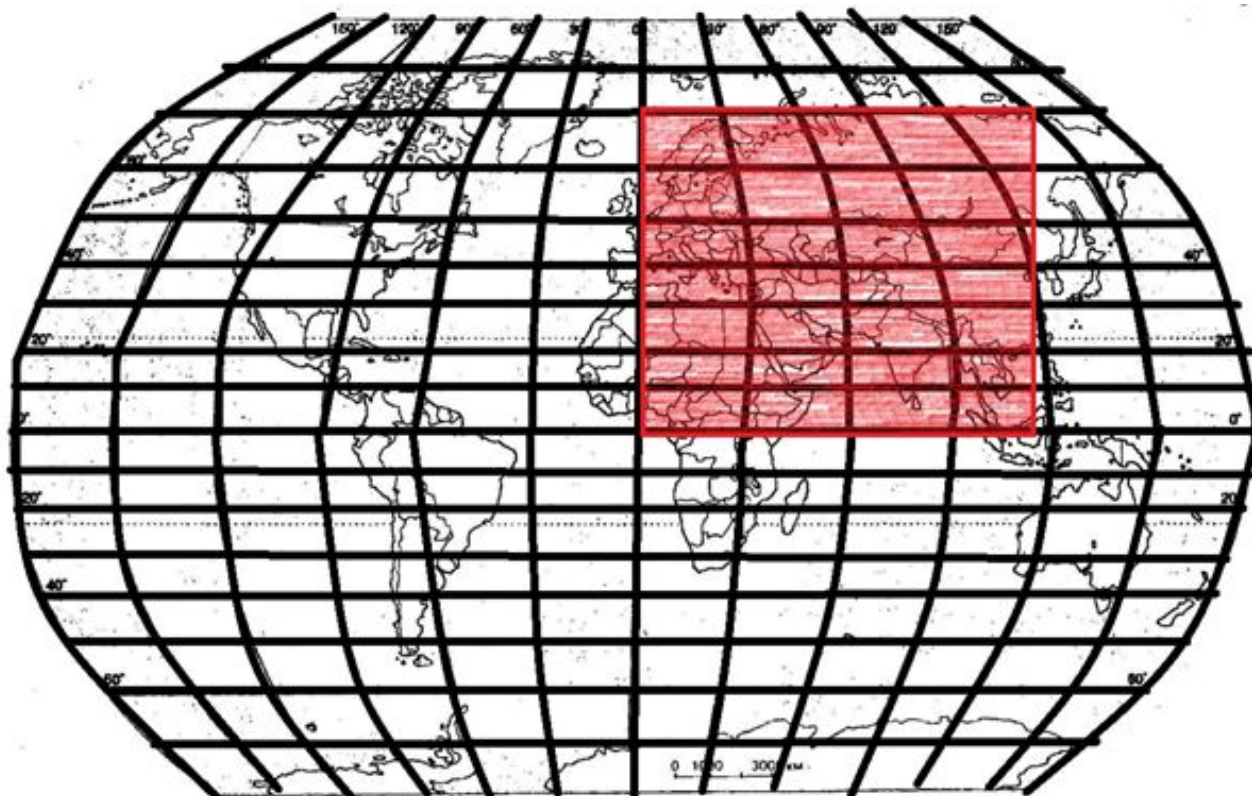
Глобальная трехмерная транспортная модель (MOZART-4)

- Данные расчета на всей поверхности Земли
- Сетка с шагом $\sim 2,8^\circ$
- Модельная область 64 узла на 128 узла
- Расчет концентрации газа на 28 уровнях (можно задавать)
- Исходный код в открытом доступе

Наблюдения (спутниковые данные AIRS)

- Атмосферный инфракрасный зонд на спутнике NASA
- Данные обеспечивают трехмерные измерения температуры и водяного пара в атмосферном столбе, а также множество газовых примесей.
- Используются центрами прогнозирования погоды по всему миру для улучшения своих прогнозов. Они также используются для оценки климатических моделей и различных приложений, от обнаружения вулканических шлейфов до прогнозирования засухи.

Наблюдения (спутниковые данные AIRS)



масштаб подобласти, покрываемой
спутниковыми наблюдениями AIRS за 9 часов

Данные для экспериментов:

Информация о концентрации - спутники AIRS

Информация о прогнозах - модель MOZART-4

Данные предоставили коллеги:

- Лагутин Анатолий Алексеевич
доктор ф-м наук, профессор
- Мордвин Егор Юрьевич
кандидат ф-м наук, доцент

[Мордвин Е.Ю., Лагутин А.А. «Метан в атмосфере западной Сибири», Барнаул, 2016]

АНСАМБЛЕВЫЙ ФИЛЬТР КАЛМАНА

первое приближение \rightarrow
генерируем ансамбль
первых приближений

$$x_0^{F(i)} = x_0^f + \sigma_i; \sum \sigma_i = 0$$

наблюдения \rightarrow генерируем
ансамбль наблюдений

$$y_n^i = y_n + \varepsilon^i; \sum \varepsilon_i = 0$$

ШАГ ПРОГНОЗА

$$x_{n+1}^{F(i)} = A_n x_n^{A(i)} \text{ ансамбль прогнозов}$$

$$\bar{x}_{n+1}^F = \frac{\sum_{i=1}^k x_{n+1}^{F(i)}}{k} \text{ средний прогноз}$$

$$P_{n+1}^F = \frac{1}{k+1} \sum_{i=1}^k (x_{n+1}^{F(i)} - \bar{x}_{n+1}^F)(x_{n+1}^{F(i)} - \bar{x}_{n+1}^F)^T$$

матрица ковариаций ошибок прогноза

ШАГ АНАЛИЗА

$$K_n = P_n^F H^T (H P_n^F H^T + R_n)^{-1} \text{ матрица усиления}$$

$$x_{n+1}^{A(i)} = x_{n+1}^{F(i)} + K_n (y_{n+1}^{(i)} - H x_{n+1}^{F(i)})$$

ансамбль оценок вектора состояния

$$\bar{x}_{n+1}^A = \frac{\sum_{i=1}^k x_{n+1}^{A(i)}}{k} \text{ средняя оценка ансамбля анализов}$$

$$P_{n+1}^A = \frac{1}{k+1} \sum_{i=1}^k (x_{n+1}^{A(i)} - \bar{x}_{n+1}^A)(x_{n+1}^{A(i)} - \bar{x}_{n+1}^A)^T$$

матрица ковариаций ошибок оценки

LETKF

(LOCAL ENSEMBLE TRANSFORM KALMAN FILTER)

$$\mathbf{x}_a = \mathbf{x}_f + \mathbf{D}\mathbf{x}_f \tilde{\mathbf{P}}_a \left(\mathbf{H}\mathbf{D}\mathbf{x}_f \right)^T \mathbf{R}^{-1} (\mathbf{y}_0 - \mathbf{H}\mathbf{x}_f)$$
$$\tilde{\mathbf{P}}_a = \left[(N-1)\mathbf{I} + \left(\mathbf{H}\mathbf{D}\mathbf{x}_f \right)^T \mathbf{R}^{-1} \mathbf{H}\mathbf{D}\mathbf{x}_f \right]^{-1}$$

Причем:

$$\mathbf{P}_f = \mathbf{D}\mathbf{x}_f \mathbf{D}\mathbf{x}_f^T \quad \text{ковариационная матрица ошибок прогноза}$$

$$\mathbf{D}\mathbf{x}_f = \frac{1}{\sqrt{N}} [\mathbf{d}\mathbf{x}_f^1, \dots, \mathbf{d}\mathbf{x}_f^N]^T \quad \text{ансамбль ошибок прогноза}$$

[HOUTEKAMER, FUQING ZHANG: Review of the Ensemble Kalman Filter for Atmospheric Data Assimilation, 2016]

LETKF (LOCAL ENSEMBLE TRANSFORM KALMAN FILTER)

Вычисляемые
весовые матрицы

Невязка

$$\mathbf{x}_a = \mathbf{x}_f + \mathbf{D}\mathbf{x}_f \tilde{\mathbf{P}}_a (\mathbf{H}\mathbf{D}\mathbf{x}_f)^T \mathbf{R}^{-1} (\mathbf{y}_0 - \mathbf{H}\mathbf{x}_f)$$

Искомая оценка

прогноз по модели

наблюдения

$$\tilde{\mathbf{P}}_a = [(\mathbf{N} - 1)\mathbf{I} + (\mathbf{H}\mathbf{D}\mathbf{x}_f)^T \mathbf{R}^{-1} \mathbf{H}\mathbf{D}\mathbf{x}_f]^{-1}$$

локальность алгоритма позволяет:

- использовать разбиение по подобластям
- разбиение поверхности Земли на квадраты равной площади (1000 км на 1000 км) и вычислять алгоритм внутри них
- оценка среднее по подобласти значение приземной концентрации
- проводить усвоение независимо друг от друга в отдельных узлах

Оценка приземной концентрации: шаг прогноза и анализа

Оценка значений концентраций по данным наблюдений и прогнозу производится по стандартной формуле фильтра Калмана (шаг анализа)

$$\mathbf{x}_a = \mathbf{x}_f + \mathbf{K}[\mathbf{y}_0 - H(\mathbf{x}_f)]$$

- \mathbf{x}_a , \mathbf{x}_f – оцениваемые и смоделированные концентрации парниковых газов
- \mathbf{y} – данные наблюдений
- \mathbf{K} – матрица усиления

Первичная обработка данных модели MOZART-4

Обработка данных, и организация структуры хранения:

- Используется подход, основанный на декомпозиции области модели
- Данные группируются по областям 1000км на 1000км
- Интерполяция по высоте в точки наличия наблюдений
- Данные MOZART-4 рассчитываются каждые 6 часов (окно усвоения ± 4.5 часа с момента прогноза по модели)

Первичная обработка спутниковых данных AIRS

Обработка данных, и организация структуры хранения:

- Формирование списка данных наблюдений для поиска оценки
- Вычисление невязок:
 - 1) окно усвоения 4.5 часа
 - 2) расстояние не более 50 км
от точки MOZART-4 до точки измерения AIRS
- Контроль:

при уровне ошибок в данных наблюдений $\sigma=20$ ppb, данные отбрасываются и не используются в усвоении

Схема численных экспериментов

1) В алгоритмах обработки больших объемов спутниковых данных для оценки эмиссии парниковых газов принято проводить оценку для заданного временного интервала (например, неделя), считая, что значения эмиссии постоянными в течение этого временного периода.

Эмиссия считается постоянной, работа алгоритма производится в заданном интервале.

2) Поверхность Земного шара разбивается на регионы, для которых производится оценка (благодаря локальности используемого алгоритма).

Процедура оценки эмиссии парниковых газов проводится отдельно для заданных подобластей.

Схема численных экспериментов

- 3) При наличии информации задаются начальные (климатические) значения эмиссии в регионах, а также матрицы ковариаций ошибок наблюдений и модели. (считаем эмиссию = 0, можно брать оценку с прошлого шага или первое приближение)
- 4) Данные наблюдений разбиваются на блоки по временным интервалам, а затем по подобластям.
- 5) Модель переноса и диффузии рассчитывает прогноз по заданным начальным значениям приземных концентраций в течение временного периода с заданным временным шагом, затем производится интерполяция в точку наблюдений и момент времени, в который производится наблюдение.
- 6) Производится шаг анализа и находятся оценки

Численные эксперименты по оценке приземной концентрации

- взяты подобласти $1000\text{км} \times 1000\text{км}$
- производился поиск оценок приземной концентрации
- использовались данные расчёта модели MOZART-4 за 2016 год
- использовались спутниковые данные за 2016 год:
AIRS/Aqua L2 Standard Physical Retrieval (AIRS-only) V006 (AIRS2RET)

Численные эксперименты с модельными данными

- задается «истинное» значение оцениваемых и моделируемых величин
- модельные данные получаются путем добавления к «истинному» значению случайных возмущений с нормальным распределением (с нулевым математическим ожиданием и заданной дисперсией), которые моделируют ошибки наблюдений в реальных данных
- для вычисления прогноза значение концентрации в приземной слое увеличивалось равномерно во всей поверхности Земного шара
- ошибка наблюдений и введенная погрешность концентрации метана в модельных экспериментах были выбраны и настроены, в соответствии смыслу физических аналогов в экспериментах с реальными данными
- для оценки достоверности полученных результатов вычислялась норма отклонения оценки от принятого за реальное («истинное») состояния системы значения

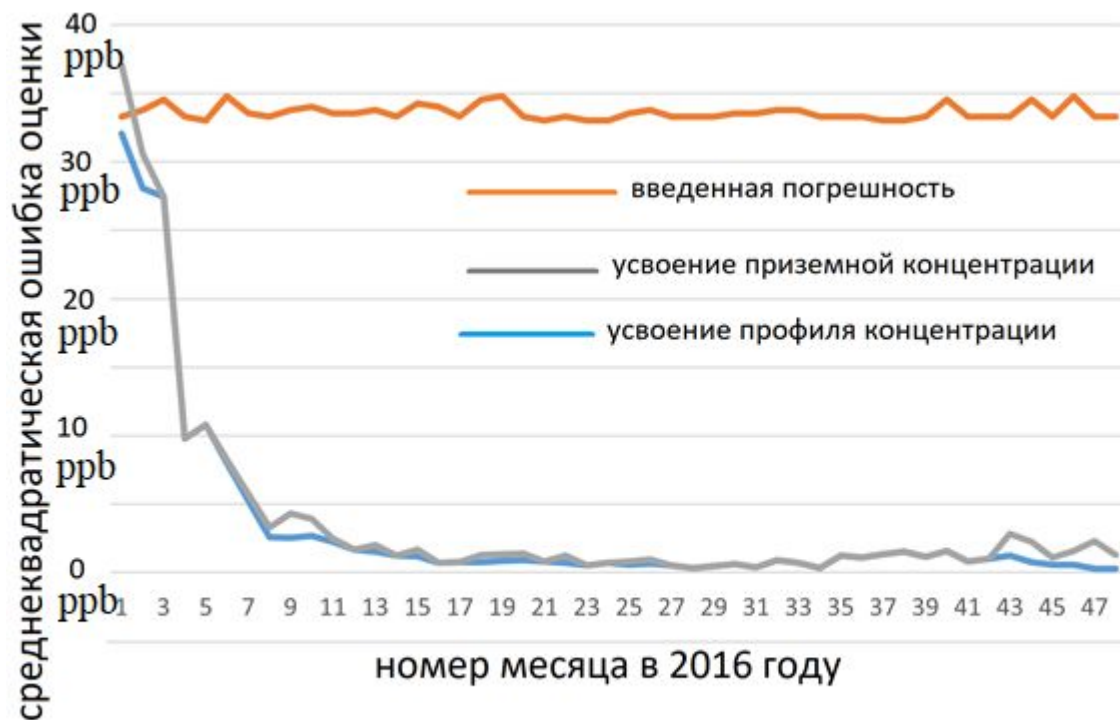
Численные эксперименты с модельными данными для процедуры усвоения данных

- процедура усвоения данных основана на алгоритме LETKF и работает циклично
- цикл начинается с расчета по модели на 7 дней, далее находится оценка приземной концентрации метана для средних значений по ансамблю за недельный период (7 дней), далее оценка усредняется для подобластей
- усреднение для подобластей помогает сгладить случайные ошибки в наблюдениях, а также выявить динамику активности подобластей за длительный период, облегчая анализ больших объемов данных
- перед следующим циклом моделирования значения приземной концентрации метана обновляются, используя полученную оценку (обновление данных раз в неделю)

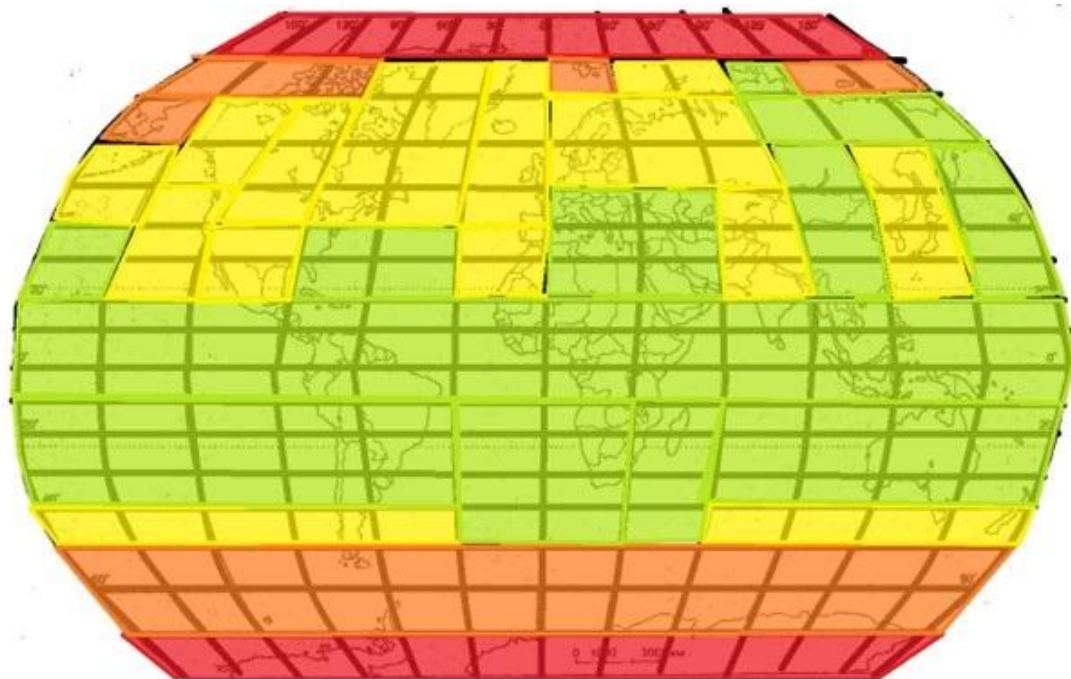
Численные эксперименты с модельными данными для процедуры усвоения данных

- результаты численных экспериментов с модельными данными (работа модели в течении одной недели)
48 циклов усвоения за 2016 год
- при введенной погрешности в размере 30-40 ppb, получаемая среднеквадратическая ошибка оценки приземной концентрации метана меньше вносимой погрешности, что является достаточно хорошим показателем проведенного эксперимента
- в экспериментах с реальными данными достичь таких результатов невозможно, так как присутствует этап интерполяции данных, данные наблюдений доступны не в каждый момент времени и не в каждой точке сетки модели и т.д.

Среднеквадратическая ошибка оценки приземной концентрации метана (модельные эксперименты)



Результаты модельных экспериментов с использованием модельных данных, адаптированных под характеристики спутниковых данных AIRS: средний процент ошибки оценки, получаемой с 3-ей по 48 неделю 2016 года по подобластям



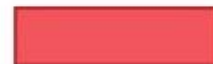
ошибка оценки 0%-20%



ошибка оценки 20%-40%

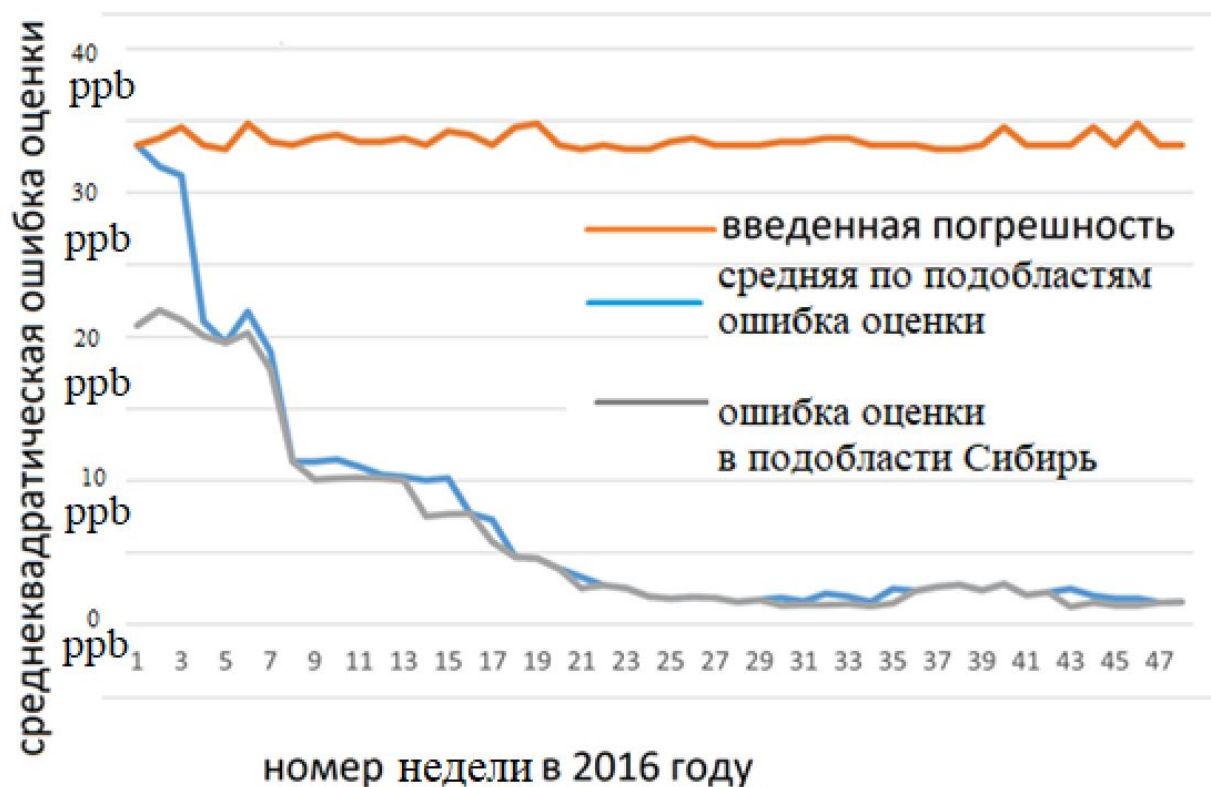


ошибка оценки 40%-60%



ошибка оценки 60%-80%

График среднеквадратической ошибки оценки, получаемой при использовании имитации спутниковых данных AIRS



Численные эксперименты с реальными данными AIRS

- взяты подобласти 1000км*1000км
- производился поиск оценок приземной концентрации с использованием модели MOZART-4 (за 2016 год)
- использовались спутниковые данные за 2016 год:
AIRS/Aqua L2 Standard Physical Retrieval (AIRS-only) V006
(AIRS2RET)

Расчет эмиссии подобласти с использованием оценки приземной концентрации

F - поток загрязняющих веществ на единицу площади и за единицу времени

Q - общее количество загрязняющих веществ, поступающих в воздушную массу, с искомой территории подобласти за время τ

$$F = \Delta C_i \frac{H}{\tau}; \quad Q = FS,$$

где S – площадь подобласти

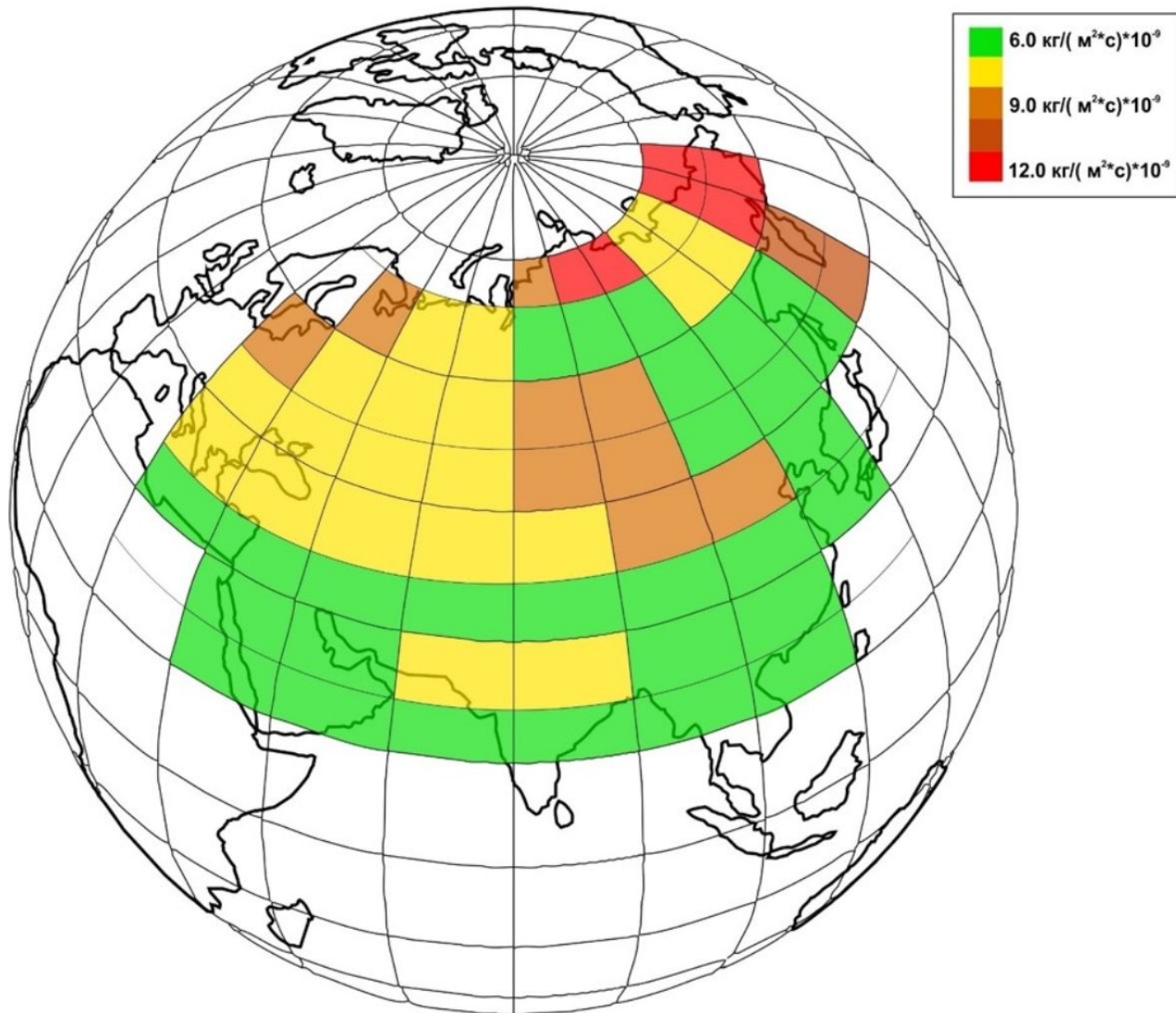
H – высота слоя перемешивания

ΔC_i – изменение концентрации метана в подобласти минус фоновое значение

Глобальный реанализ парниковых газов CAMS (EGG4)

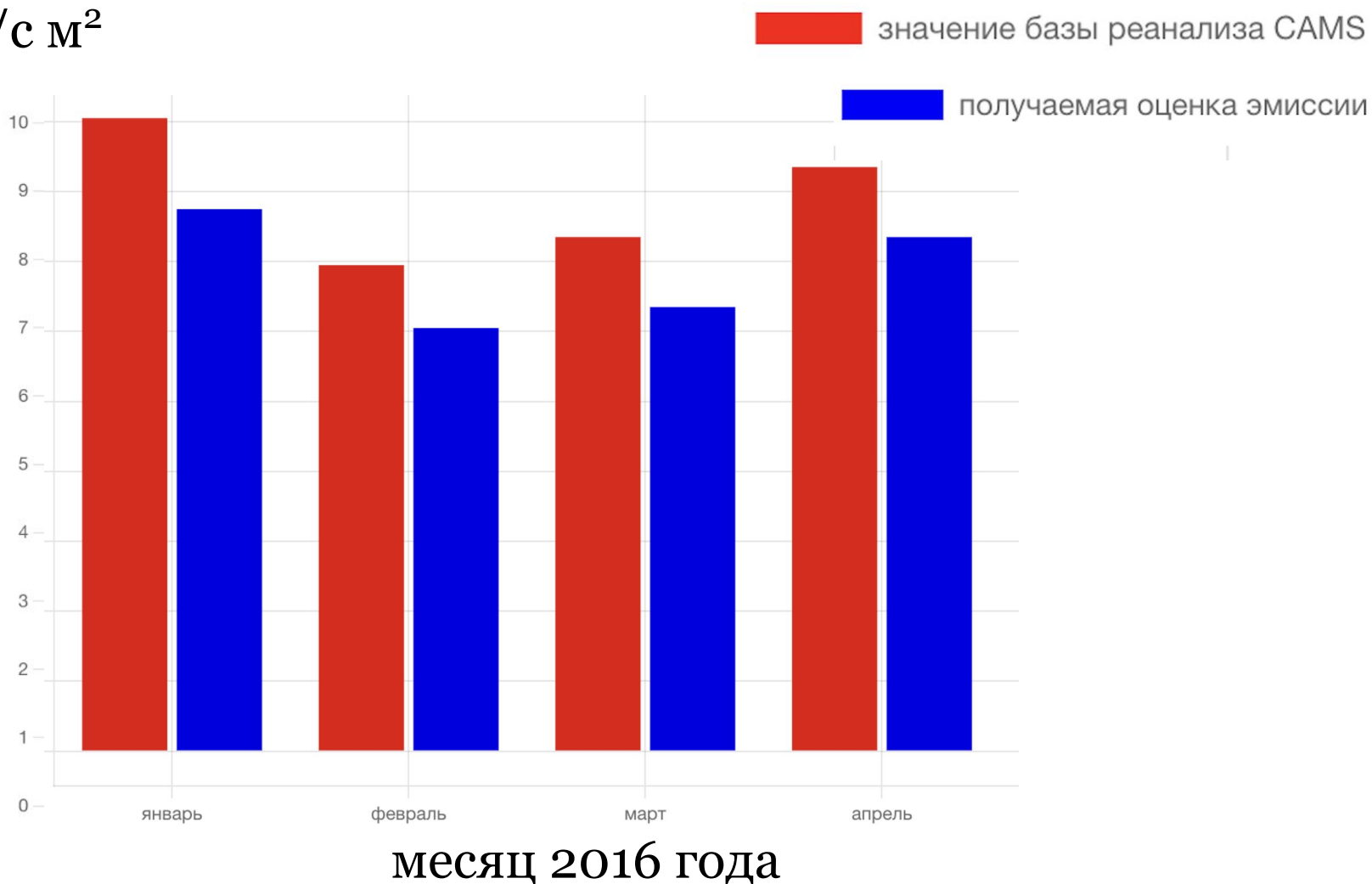
- данные доступны с 2003 года
- данные доступны в открытом доступе на сетке с шагом 0,75 градуса (<https://ads.atmosphere.copernicus.eu/>)
- возможность выбора уровней вертикальной сетки

Оценка эмиссии метана по подоблостям за июнь 2016 год



Сравнительный анализ оценки эмиссии метана для региона Сибирь за 2016 год

e^{15} кг/с м²



Заключение

- Задача оценки состояния окружающей среды по данным наблюдений в настоящее время решается с помощью систем усвоения данных
- Представлен алгоритм оценки эмиссии парниковых газов на с привлечением спутниковых данных и расчетов глобальной транспортной модели
- В докладе представлены результаты численных экспериментов апробации системы усвоения данных с модельными и реальными данными, и проведен сравнительный анализ результатов с данными реанализа

Спасибо за внимание!