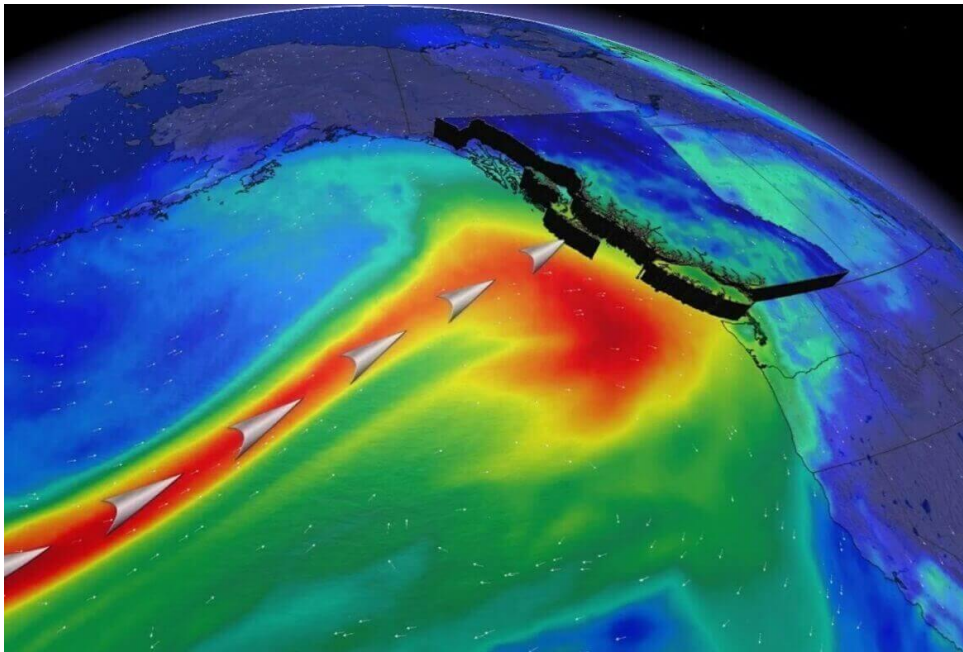
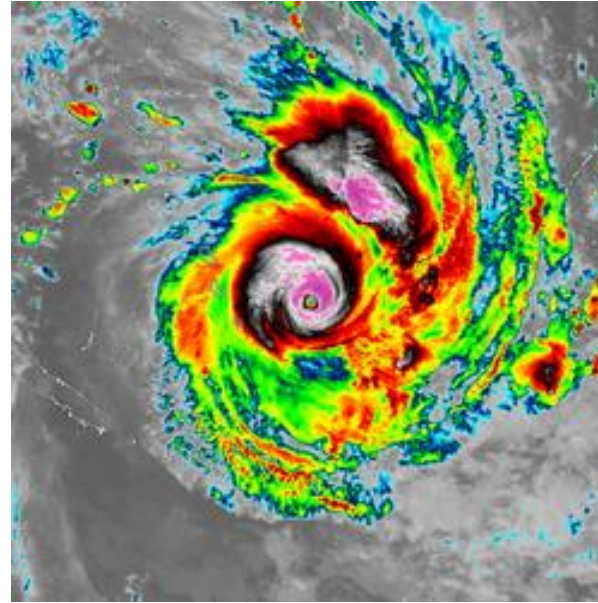
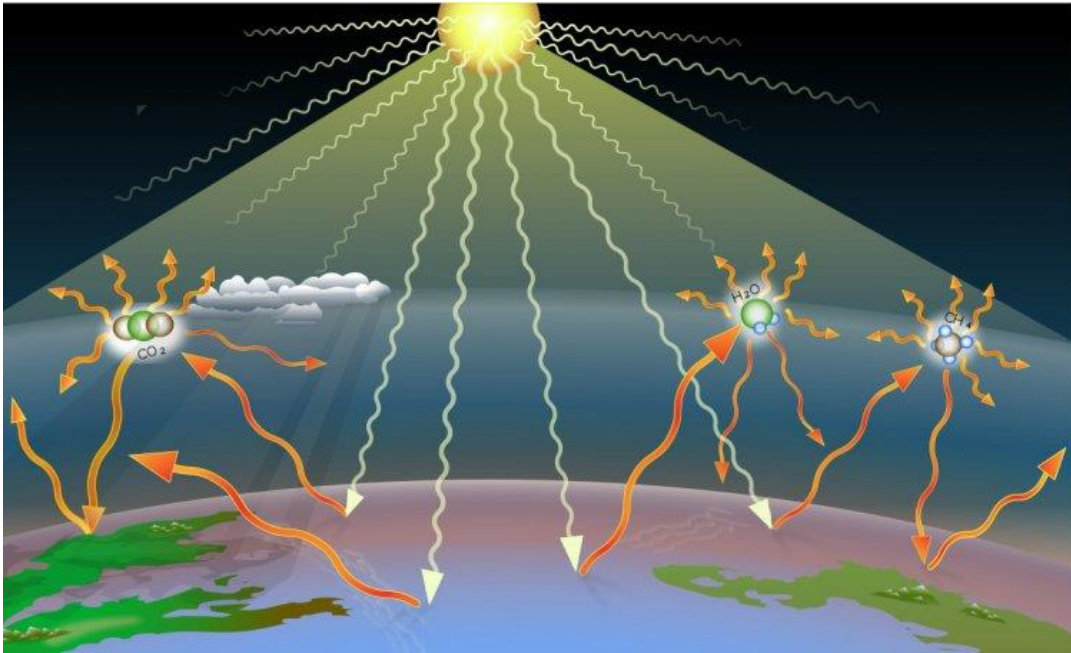



# Развитие многолетней базы данных интегрального содержания и потоков водяного пара в атмосфере

*Пашинов Е.В.*, Втюрин С.А., Ермаков Д.М. (ИКИ РАН)

# Водяной пар в атмосфере



<https://fireras.su/tpw/>



## СПУТНИКОВОЕ РАДИОТЕПЛОВИДЕНИЕ: ДИНАМИКА ГЕОФИЗИЧЕСКИХ ПОЛЕЙ

Анимация Поля Икар Методика Публикации Программы Новости Контакты

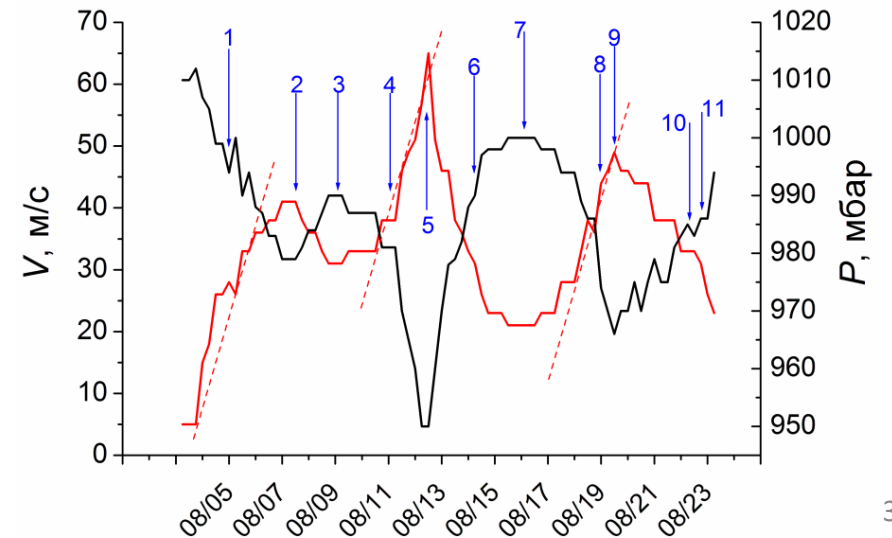
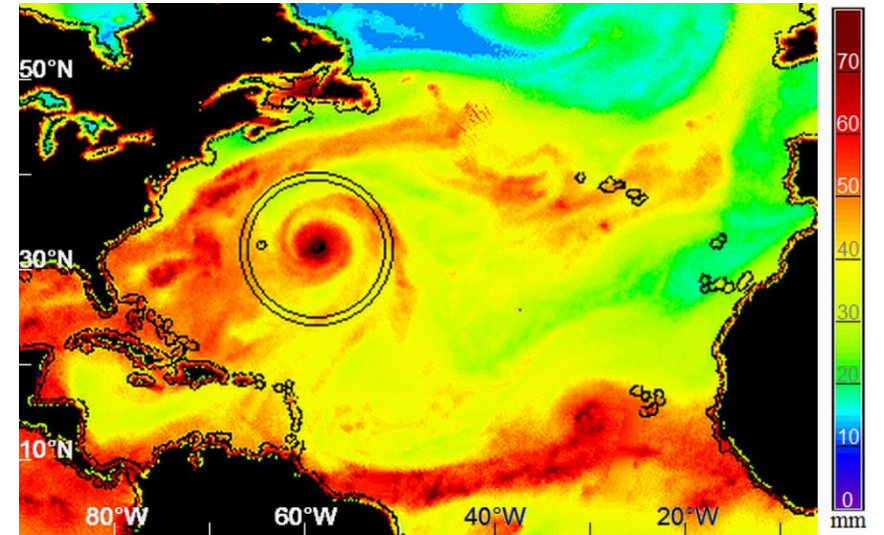
2024 Параметер: TPW

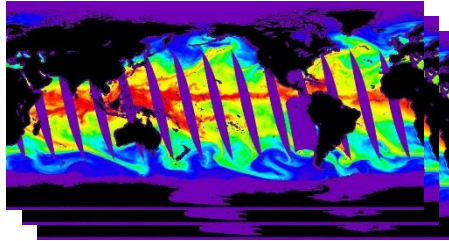
### АНИМИРОВАННЫЕ ГЕОФИЗИЧЕСКИЕ ПОЛЯ

Динамика поля геофизического параметра атмосферы над акваториями Мирового океана: интегральное влагосодержание атмосферы (TPW), водозапас облаков (CLW), скорость приповерхностного ветра (WND) и скорость адвекции водяного пара в нижней тропосфере (ADV). Шаг по времени 3 часа - для полей TPW, CLW, WND; 6 часов - для полей ADV. Шаг сетки - 0,25°. Цветовая шкала значений приведена справа.

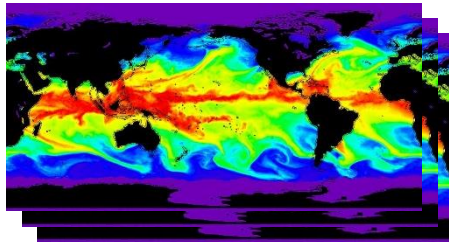


0:00 / 4:53





«МГНОВЕННЫЕ» ( $\delta t = 0$ ) значения геофизических полей в точках наблюдений

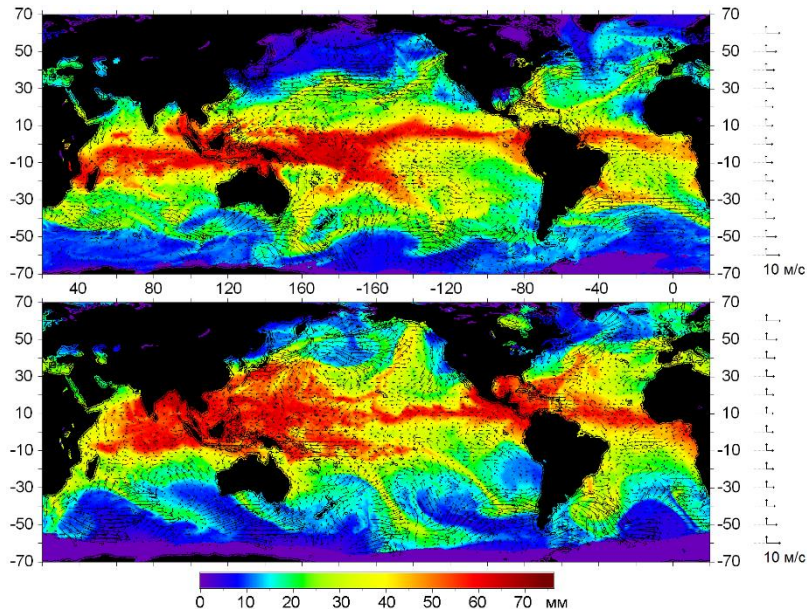


сшивка лакун

геофизические поля системы «океан-атмосфера» полного покрытия с периодом 12 часов («опорные»)

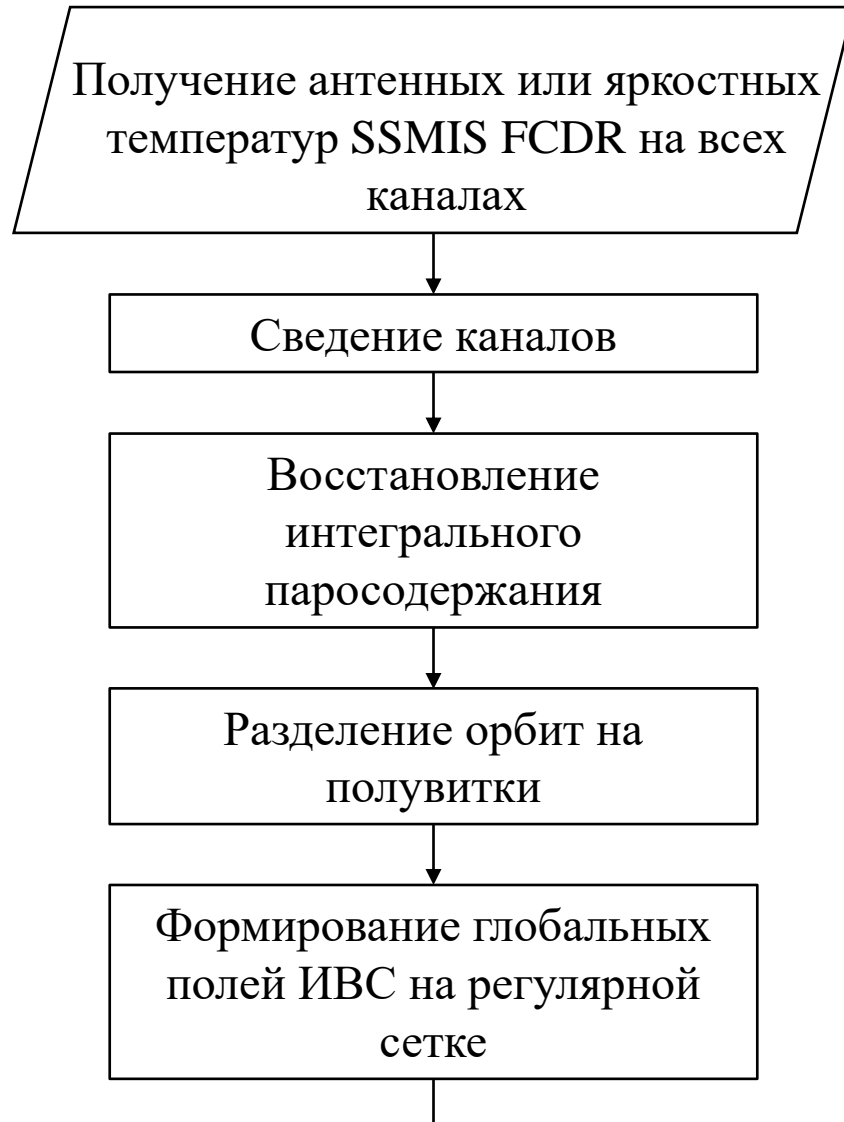
оценка и компенсация движения

геофизические поля системы «океан-атмосфера» полного покрытия, интерполированные к любому моменту времени в промежутке между «опорными» полями ( $0 < \delta t < 12$  часов) и синхронные с ними векторные поля скорости адвекции

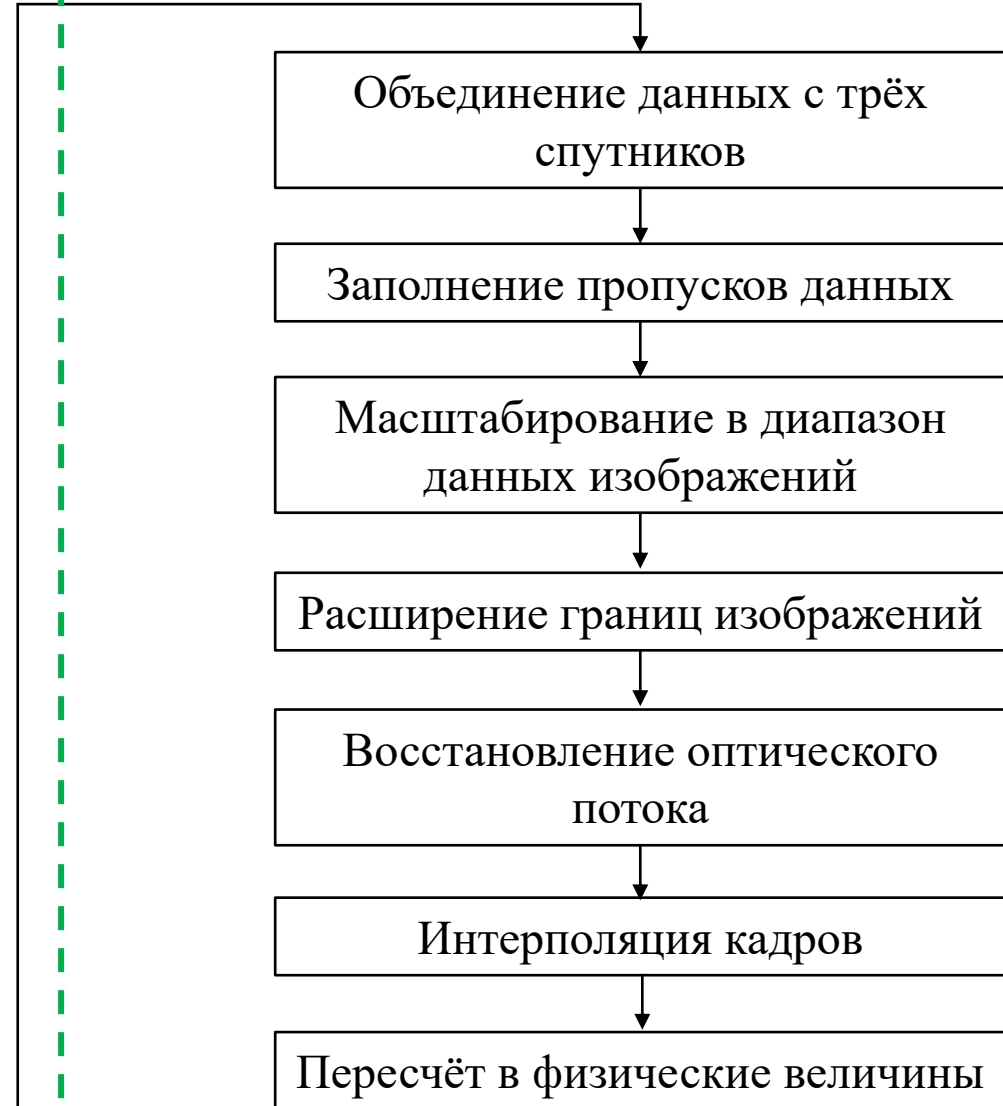


# Схема обработки данных

## Блок радиометрии



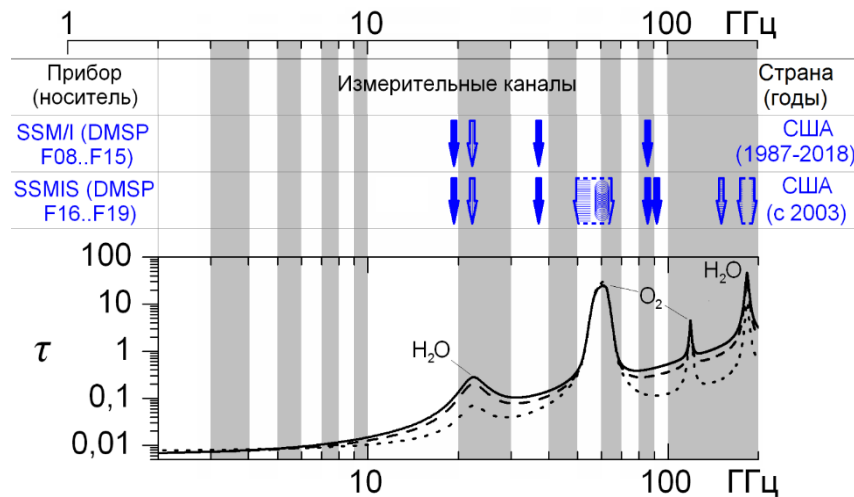
## Блок динамики



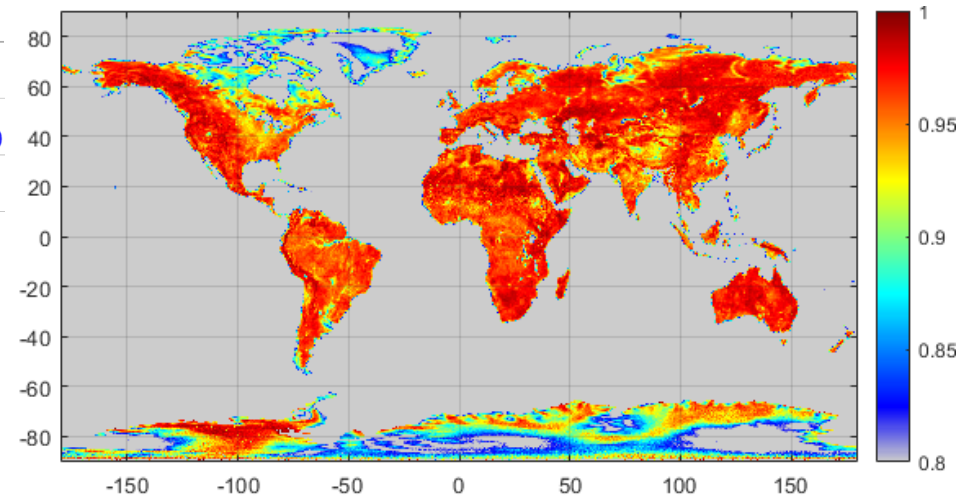
# Данные многоканального радиотеплового спутникового мониторинга

Программа МО США – DMSP, приборы - Special Sensor Microwave...

- измерения ИВП над океаном с 1987 г. – SSM/I
- возможность измерений над всеми поверхностями - SSMIS 2004 – н.в.

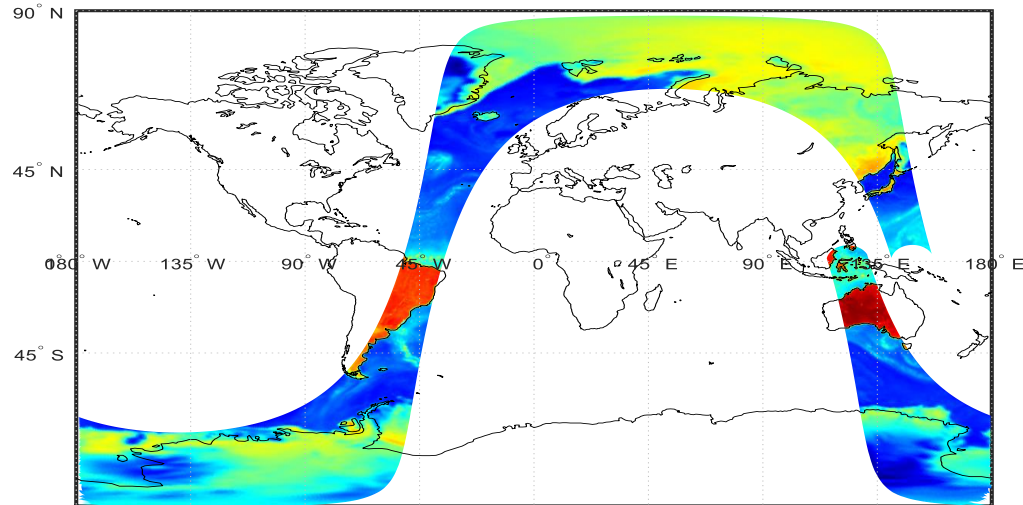


Излуч. способность SSM/I 22.235 ГГц (B)



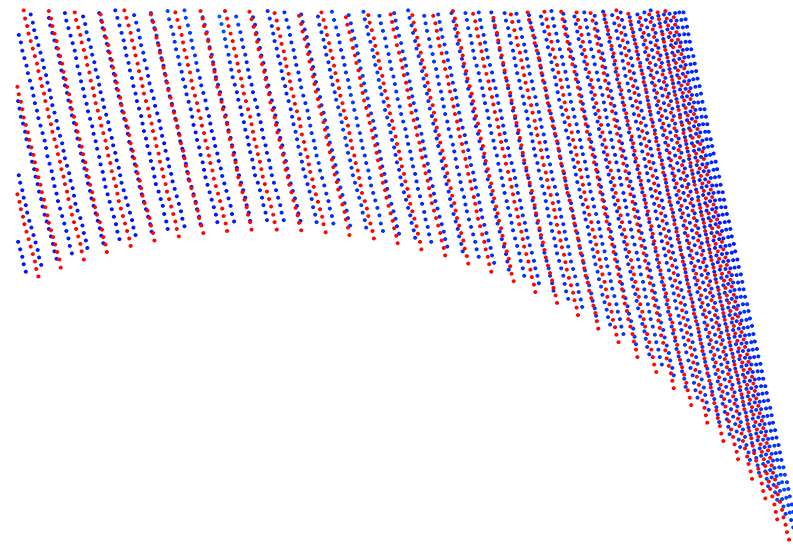
- восстановление интегрального паросодержания атмосферы над сушей (всеми типами поверхности) возможно по данным SSMIS.

# Описание радиометрических данных



**Полная доступность только в виде - BASE Temperature Data Record (TDR) dataset from Colorado State University (CSU).**

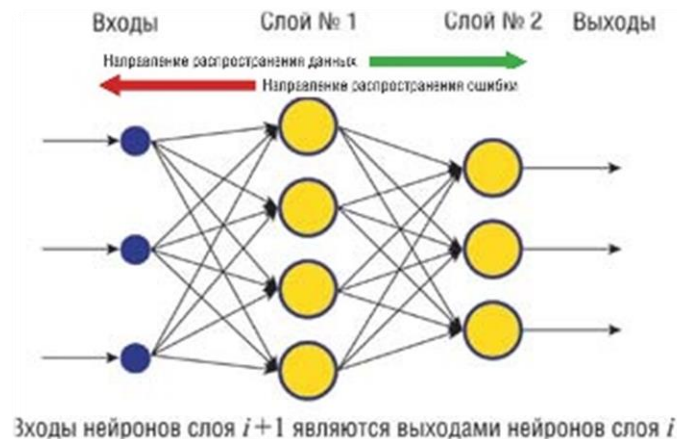
Антенные температуры L1A до 2021. Частичная доступность в виде Fundamental Climate Data Record с 2021



| Ch. No. | Center Frequency (GHz)      | Pass Band (MHz) | Polarization | NE?T (K) |
|---------|-----------------------------|-----------------|--------------|----------|
| 1       | 50.3                        | 400             | H            | 0.4      |
| 2       | 52.8                        | 400             | H            | 0.4      |
| 3       | 53.596                      | 400             | H            | 0.4      |
| 4       | 54.4                        | 400             | H            | 0.4      |
| 5       | 55.5                        | 400             | H            | 0.4      |
| 6       | 57.29                       | 350             | *            | 0.5      |
| 7       | 59.4                        | 250             | *            | 0.6      |
| 8       | 150.0                       | 1500            | H            | 0.88     |
| 9       | 183.31±6.6                  | 2500            | H            | 1.2      |
| 10      | 183.31±3                    | 100             | H            | 1.0      |
| 11      | 183.31±1                    | 500             | H            | 1.25     |
| 12      | 19.35                       | 400             | H            | 0.7      |
| 13      | 19.35                       | 400             | V            | 0.7      |
| 14      | 22.235                      | 400             | V            | 0.7      |
| 15      | 37.0                        | 1500            | H            | 0.5      |
| 16      | 37.0                        | 1500            | V            | 0.5      |
| 17      | 91.655                      | 3000            | V            | 0.9      |
| 18      | 91.655                      | 3000            | H            | 0.9      |
| 19      | 63.283248 ±0.285271         | 3               | V+H          | 2.4      |
| 20      | 60.792668 ±0.357892         | 3               | V+H          | 2.4      |
| 21      | 60.792668 ±0.357892 ±0.002  | 6               | V+H          | 1.8      |
| 22      | 60.792668 ±0.357892 ±0.0055 | 12              | V+H          | 1.0      |
| 23      | 60.792668 ±0.357892 ±0.016  | 32              | V+H          | 0.6      |
| 24      | 60.792668 ±0.357892 ±0.050  | 120             | V+H          | 0.7      |

- Выборка совмещённых спутниковых и подспутниковых измерений за 2020 год.
- Подспутниковые измерения – реанализ ECMWF Era5. Совмещение по времени в пределах  $\pm 15$  мин. Всего – 8 млн случайных измерений за год для каждого спутника.
- Случайное разбиение на обучающую и тестовую выборки. Обучающая – 5 млн измерений, тестовая – 3 млн измерений.
- Нейросеть – прямого распространения с одним скрытым слоем.

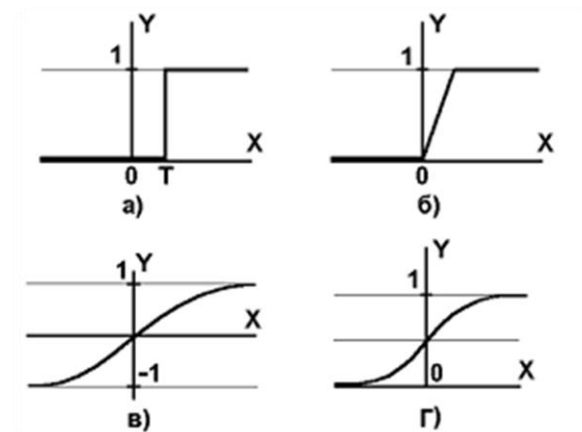
Формула:



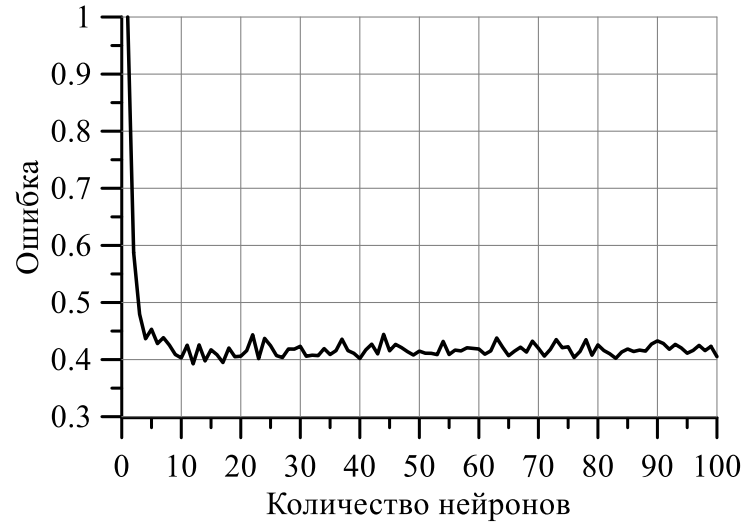
Математическая модель нейрона

$$x_{(k)}^{(i+1)} = f\left(\sum_{j=1}^N w_j^{(k)} x_j^{(i)}\right)$$

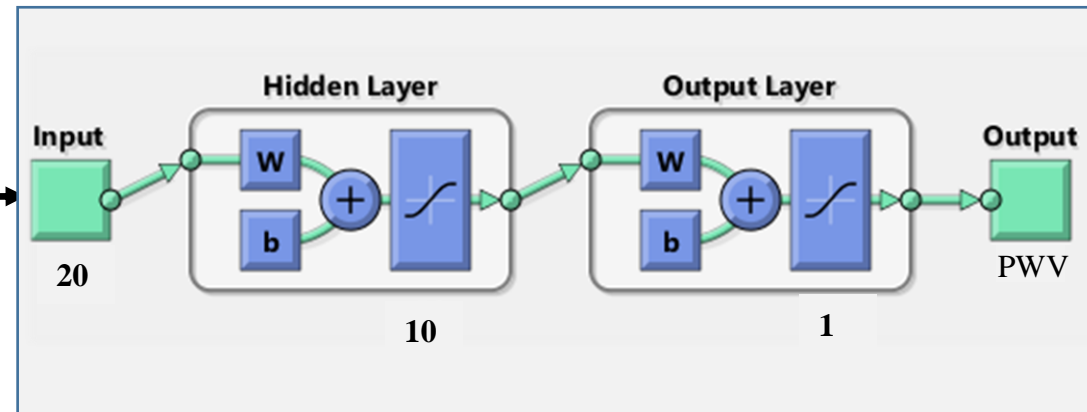
Выход  $k$ -го нейрона слоя  $i+1$  рассчитывается как взвешенная сумма всех его входов со слоя  $i$ , к которой применена функция активации, нормализующая выходной сигнал



# Восстановление интегрального паросодержания



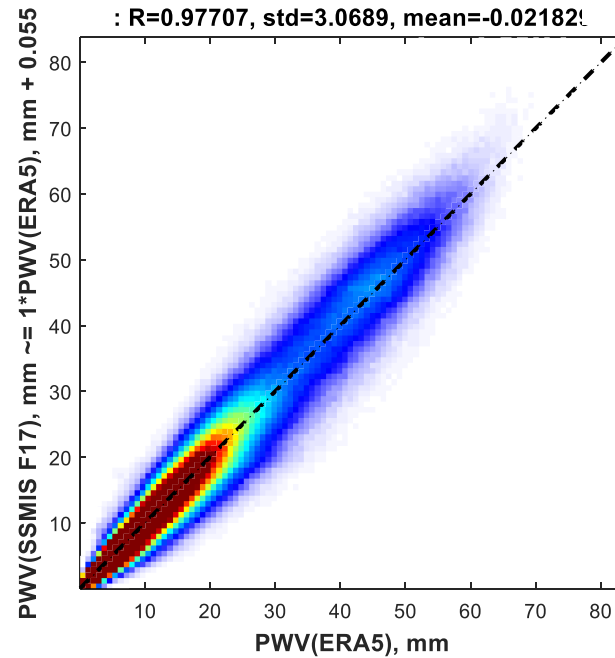
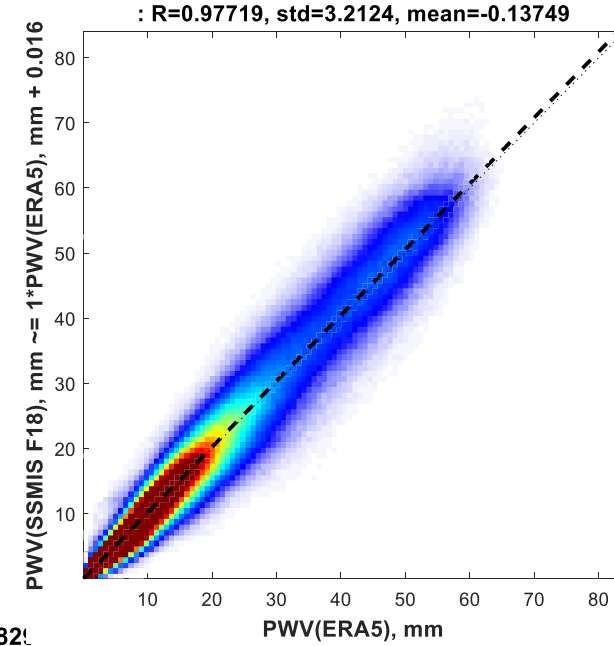
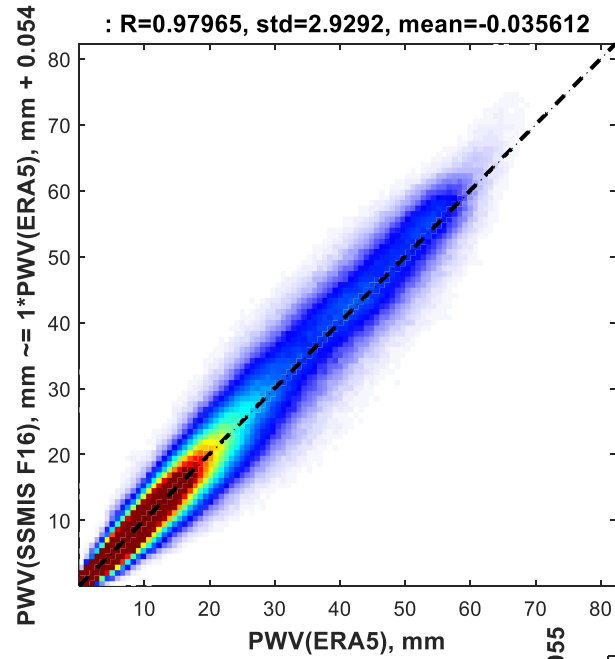
| Ch. No. | Center Frequency (GHz) | Pass Band (MHz) | Polarization | NE?T (K) |
|---------|------------------------|-----------------|--------------|----------|
| 1       | 50.3                   | 400             | H            | 0.4      |
| 2       | 52.8                   | 400             | H            | 0.4      |
| 3       | 53.596                 | 400             | H            | 0.4      |
| 4       | 54.4                   | 400             | H            | 0.4      |
| 5       | 55.5                   | 400             | H            | 0.4      |
| 6       | 57.29                  | 350             | *            | 0.5      |
| 7       | 59.4                   | 250             | *            | 0.6      |
| 8       | 150.0                  | 1500            | H            | 0.88     |
| 9       | 183.31±6.6             | 2500            | H            | 1.2      |
| 10      | 183.31±3               | 100             | H            | 1.0      |
| 11      | 183.31±1               | 500             | H            | 1.25     |
| 12      | 19.35                  | 400             | H            | 0.7      |
| 13      | 19.35                  | 400             | V            | 0.7      |
| 14      | 22.235                 | 400             | V            | 0.7      |
| 15      | 37.0                   | 1500            | H            | 0.5      |
| 16      | 37.0                   | 1500            | V            | 0.5      |
| 17      | 91.655                 | 3000            | V            | 0.9      |
| 18      | 91.655                 | 3000            | H            | 0.9      |



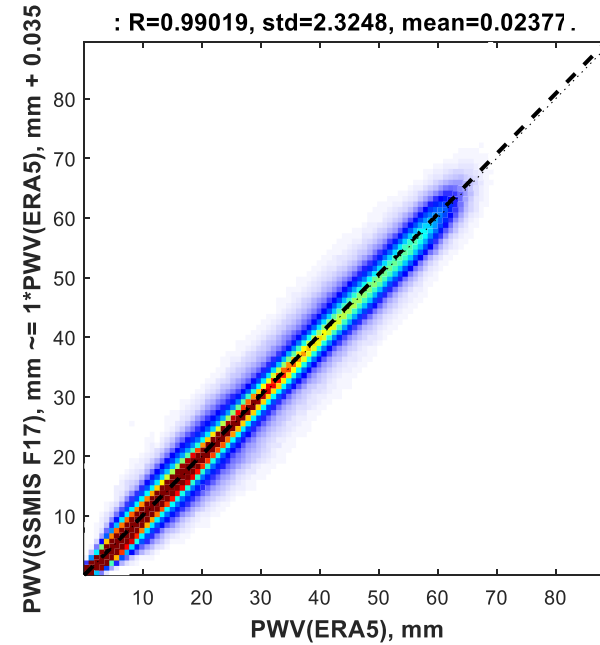
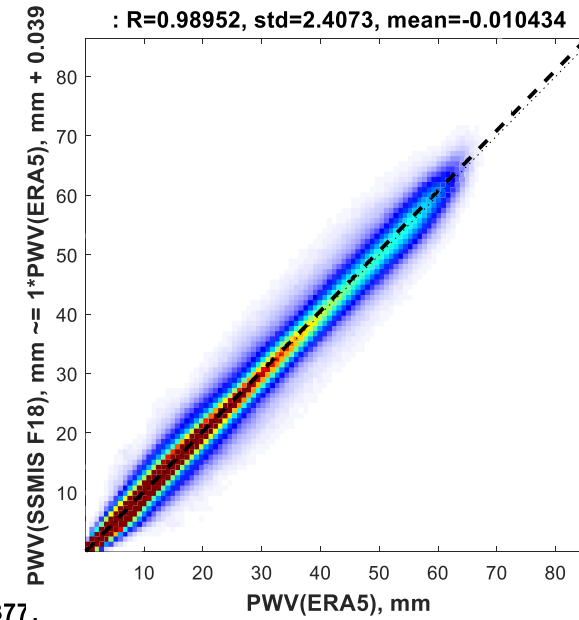
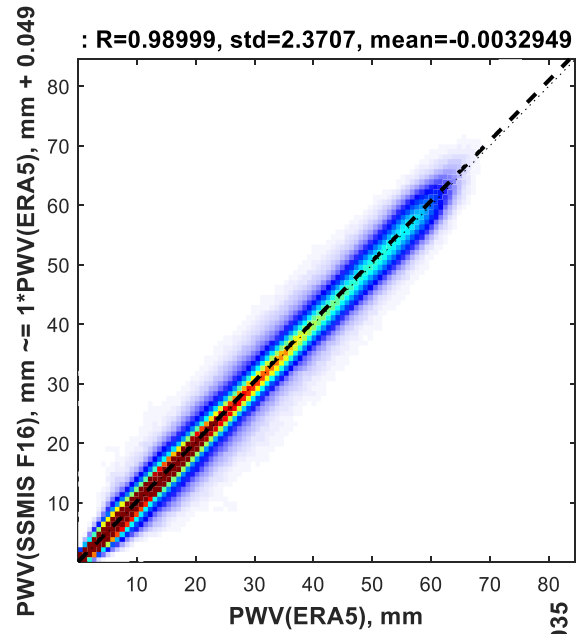
+

- Высота поверхности
- Флаг типа поверхности

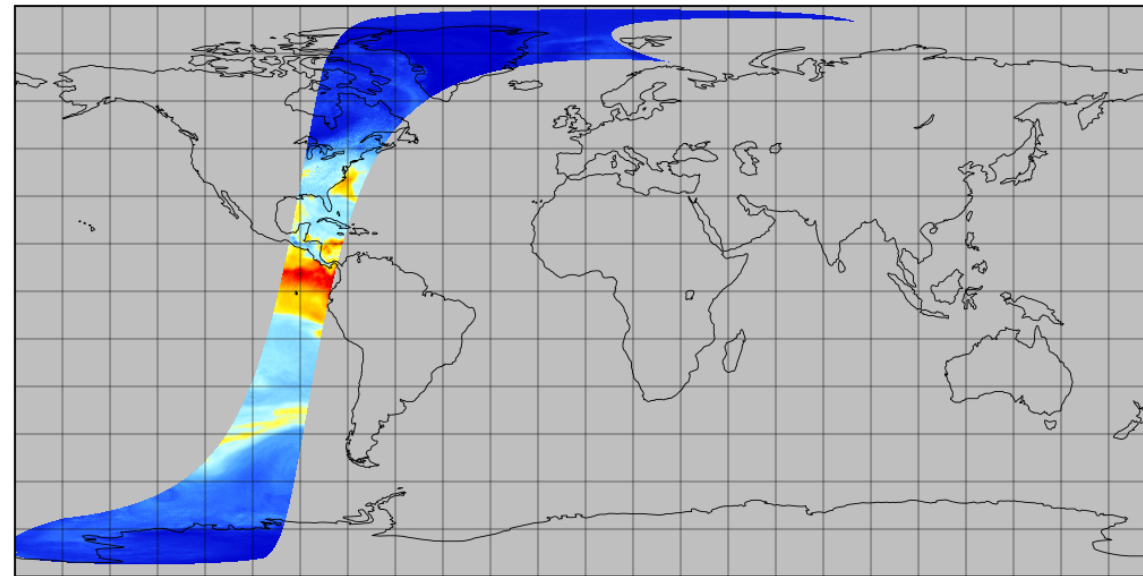
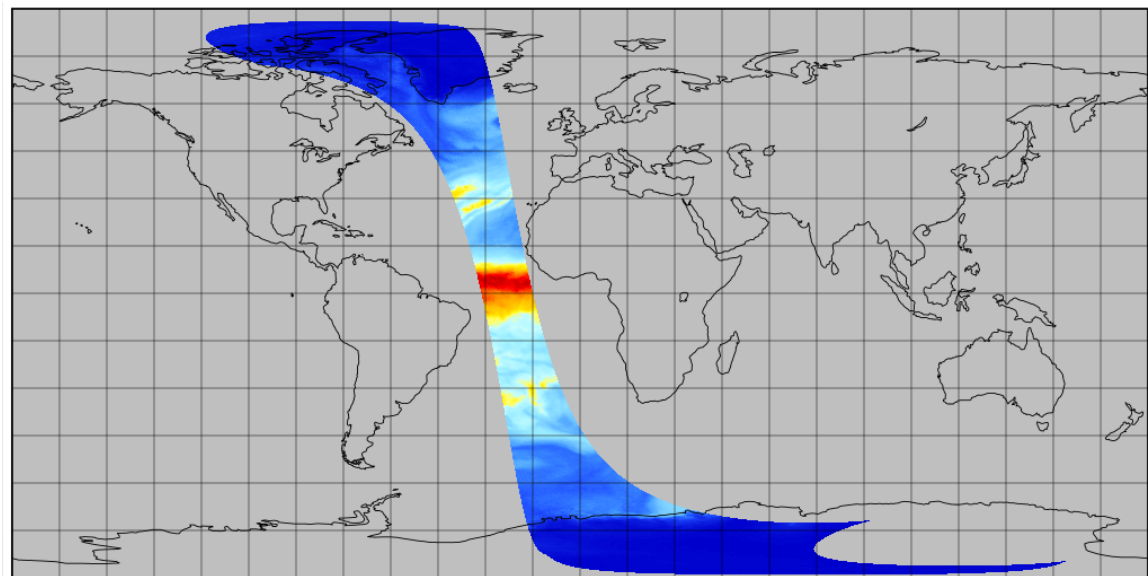
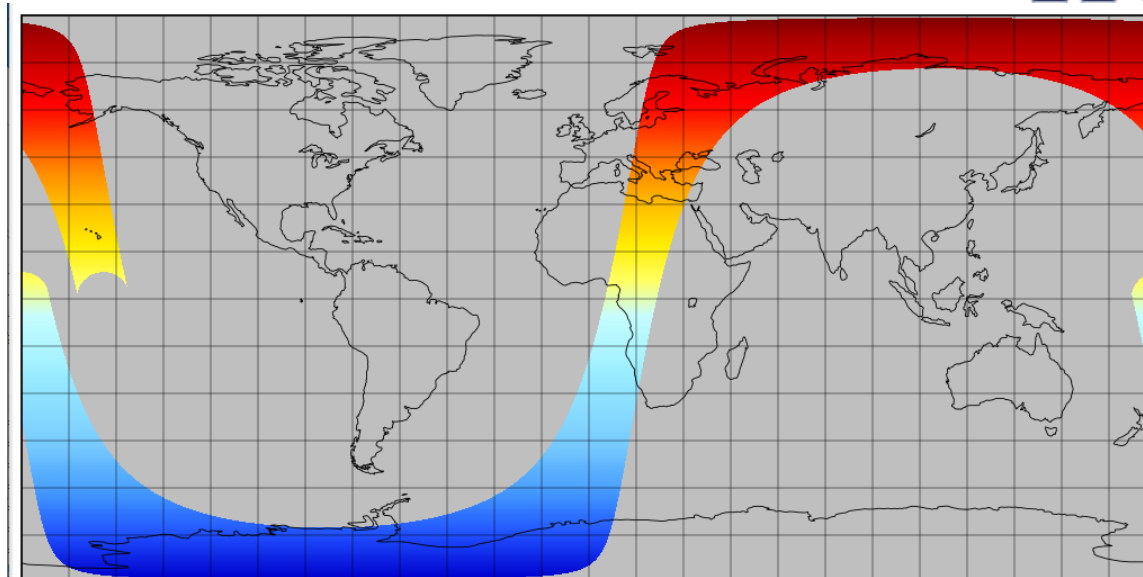
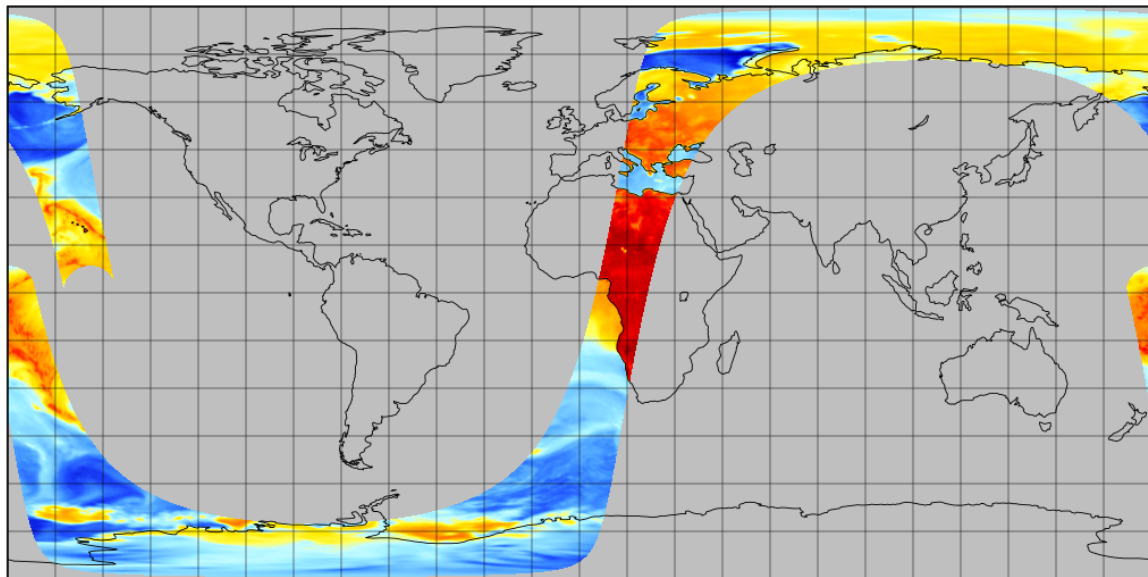
# Восстановление интегрального паросодержания над сушей



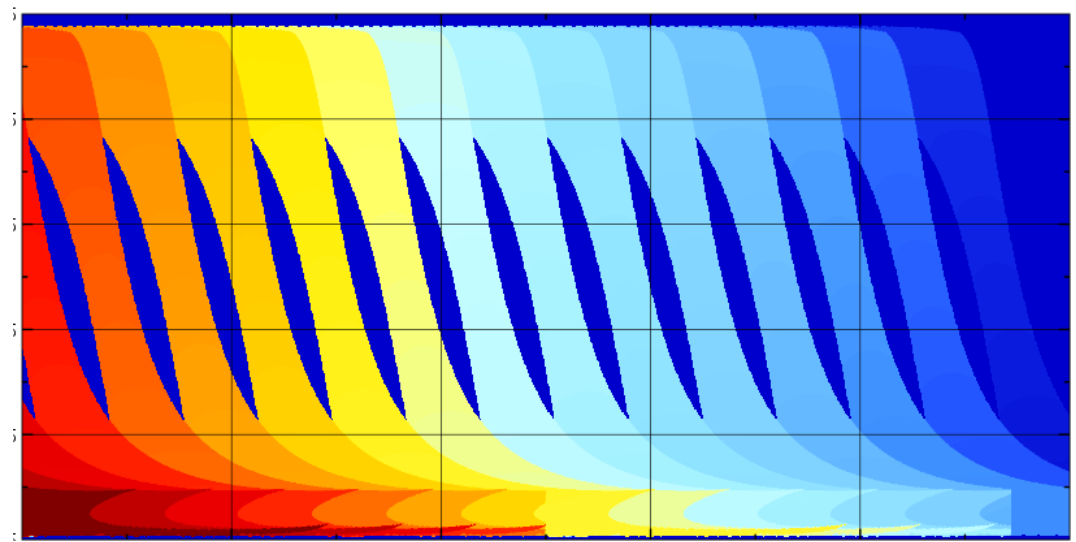
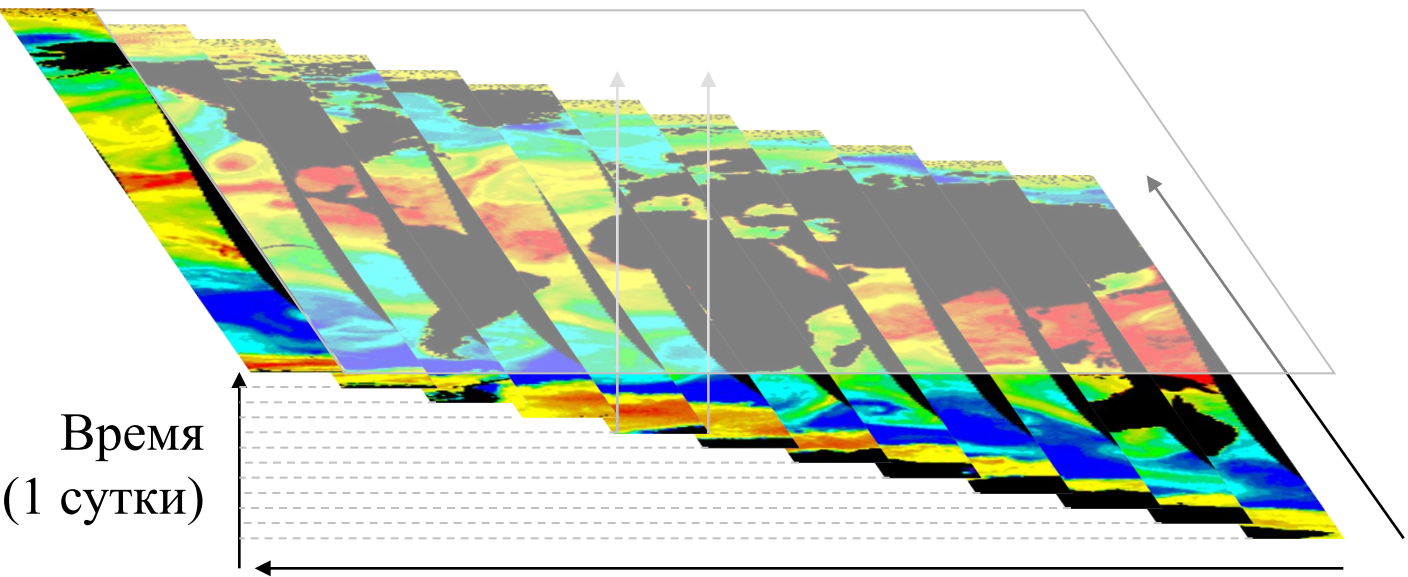
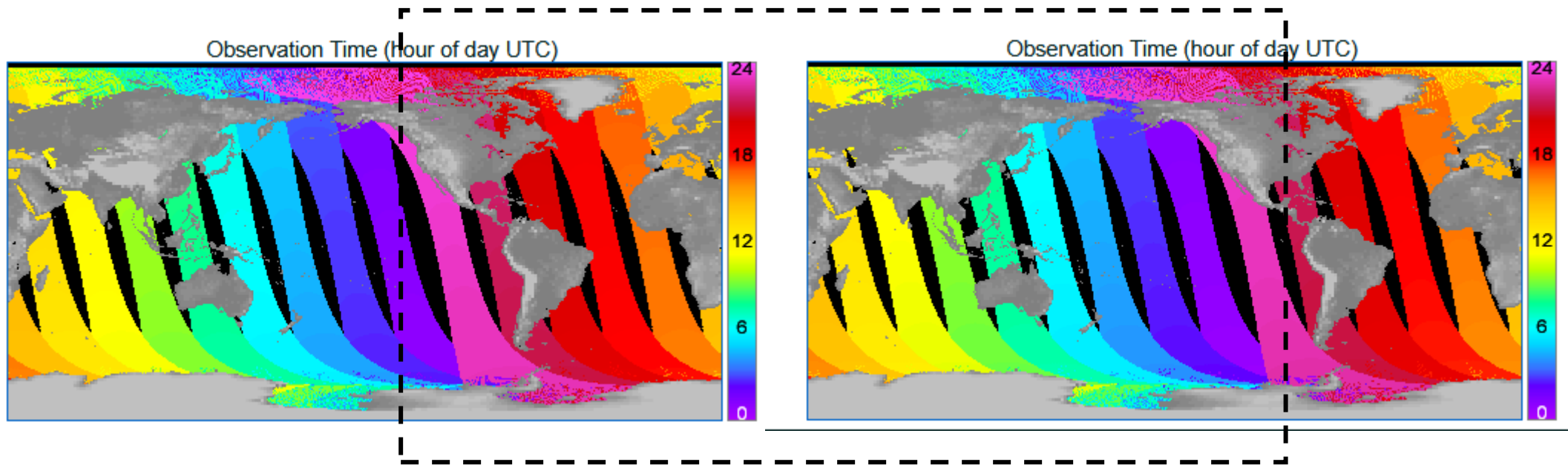
# Восстановление интегрального паросодержания по всему земному шару



# Разделение орбит на полувитки

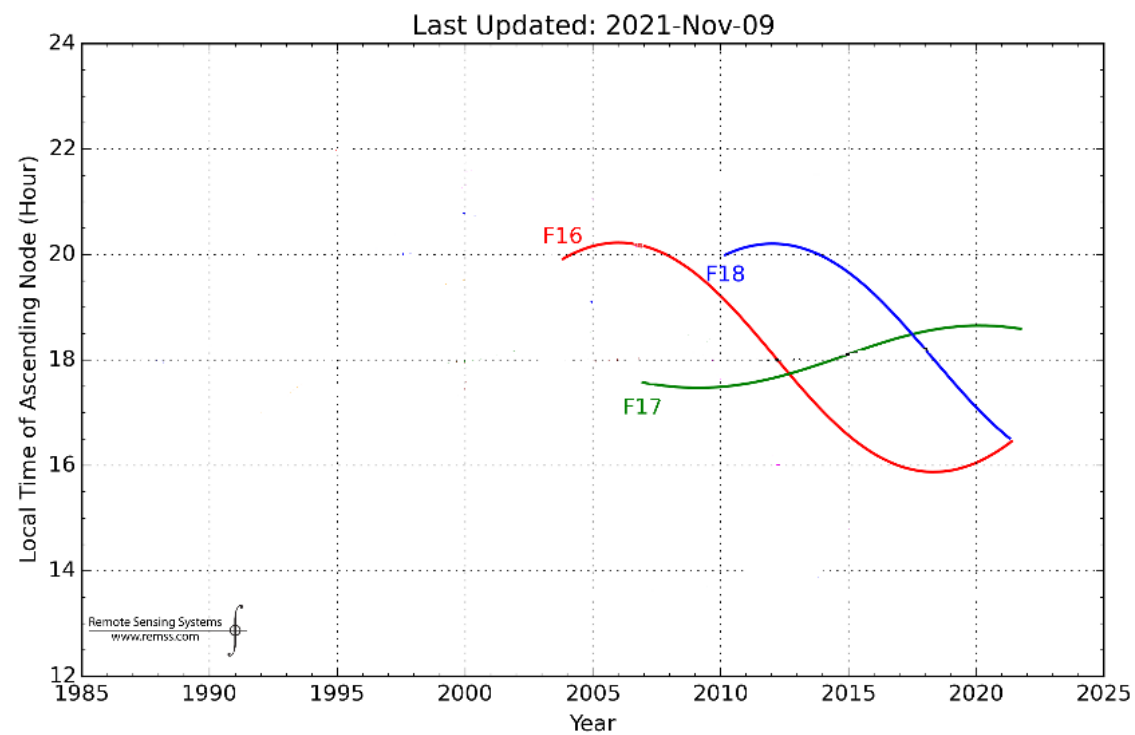
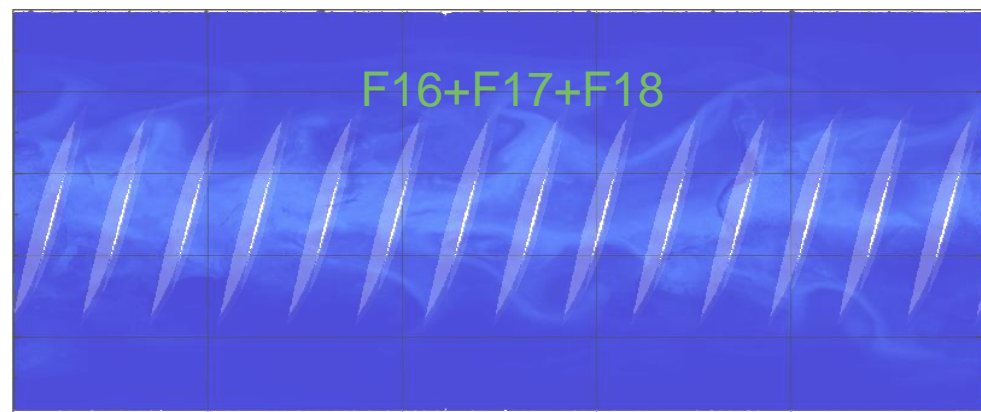
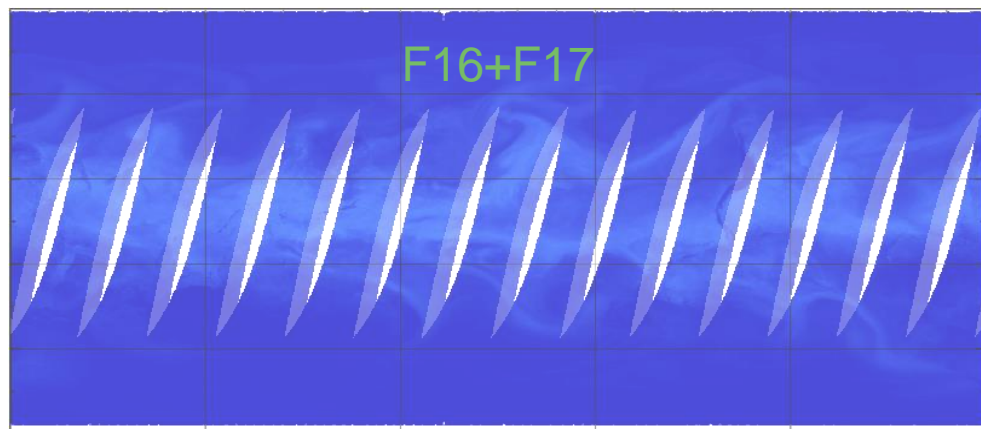
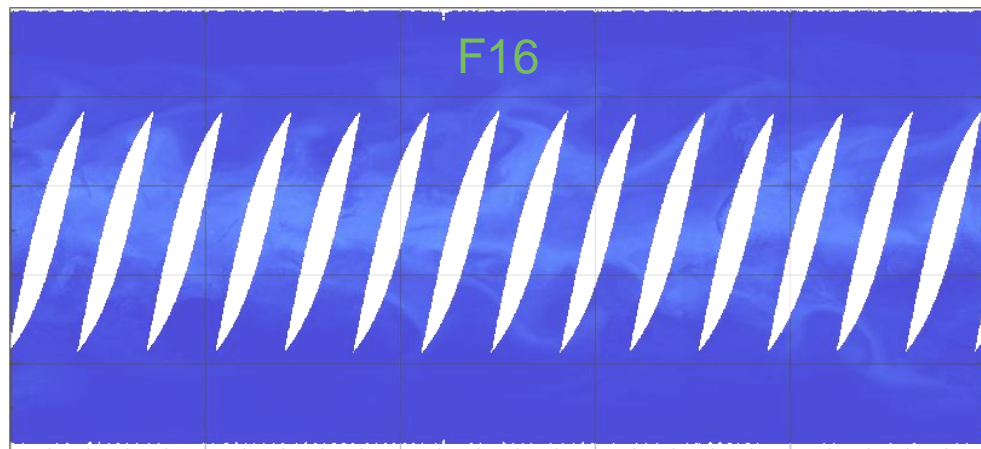


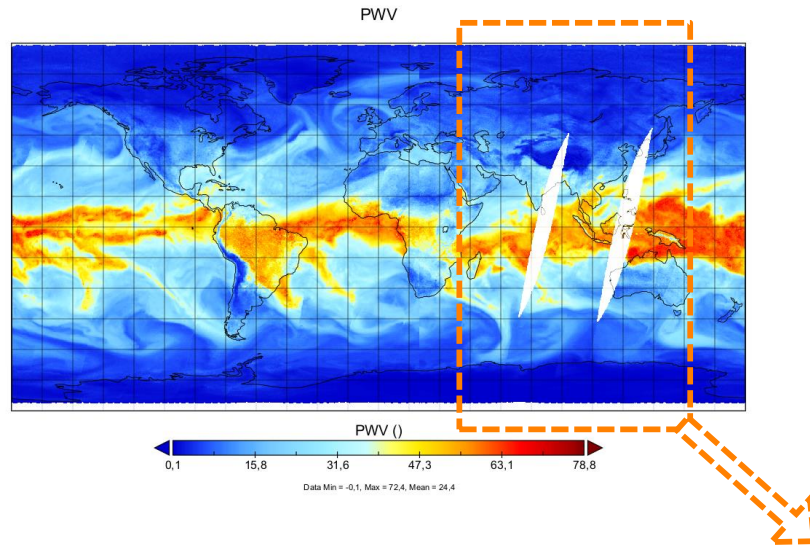
# Формирование глобальных полей ИВС на регулярной сетке



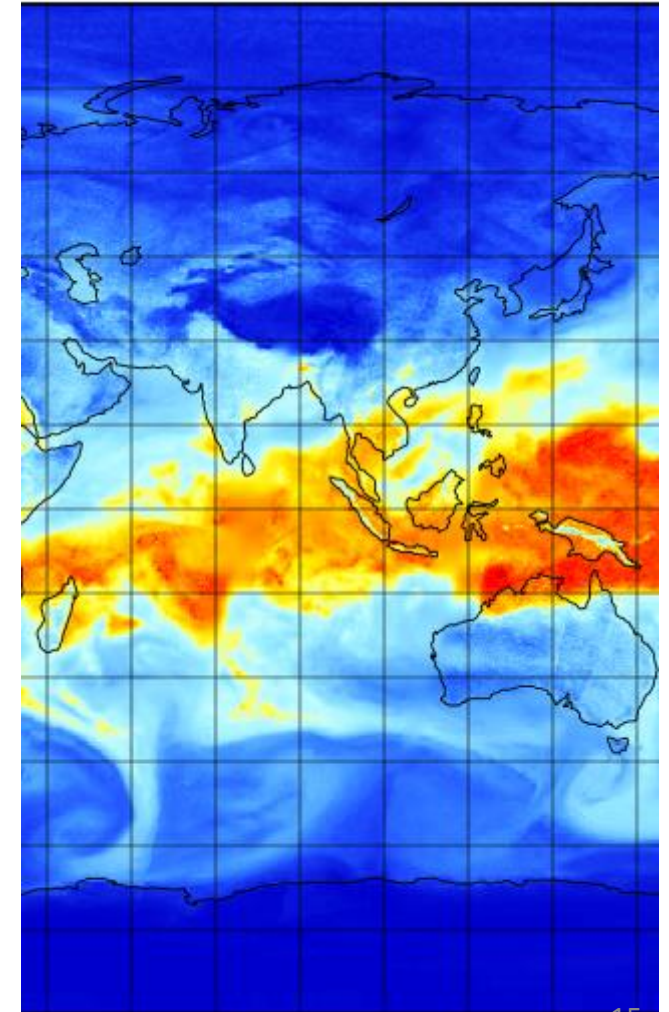
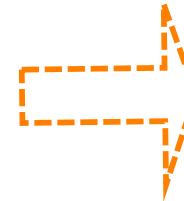
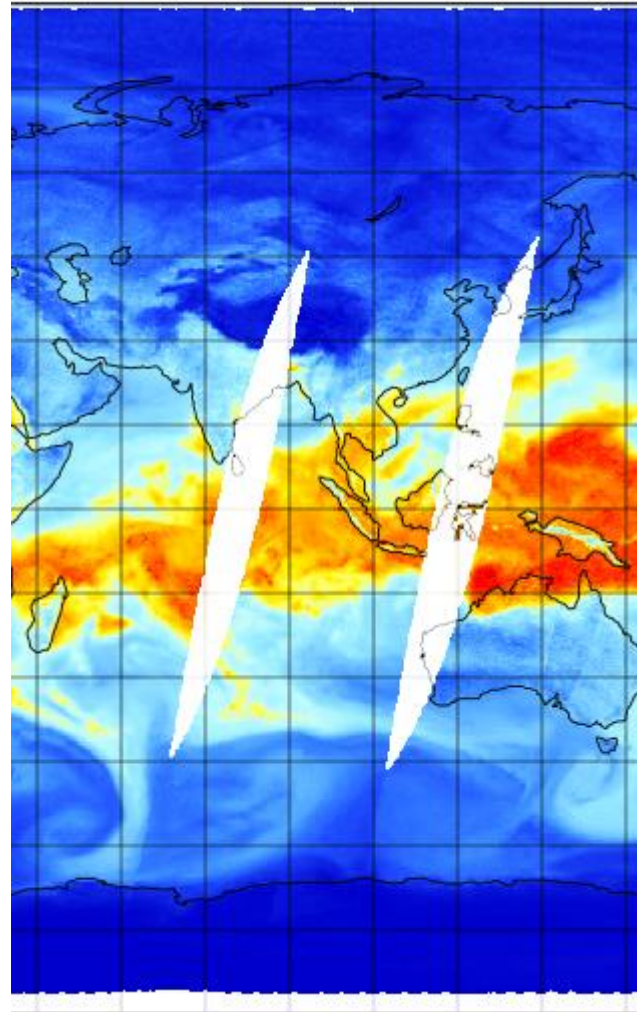
Покрытие по долготе за счет суточного вращения Земли

# Объединение данных с трёх спутников





Области, не содержащие данные ни с одного спутника восстанавливаются при помощи контекстного заполнения. Они остаются отмеченными как “нет реальных данных”, но заполнение улучшает работу алгоритма оптического потока и построения интерполированных кадров

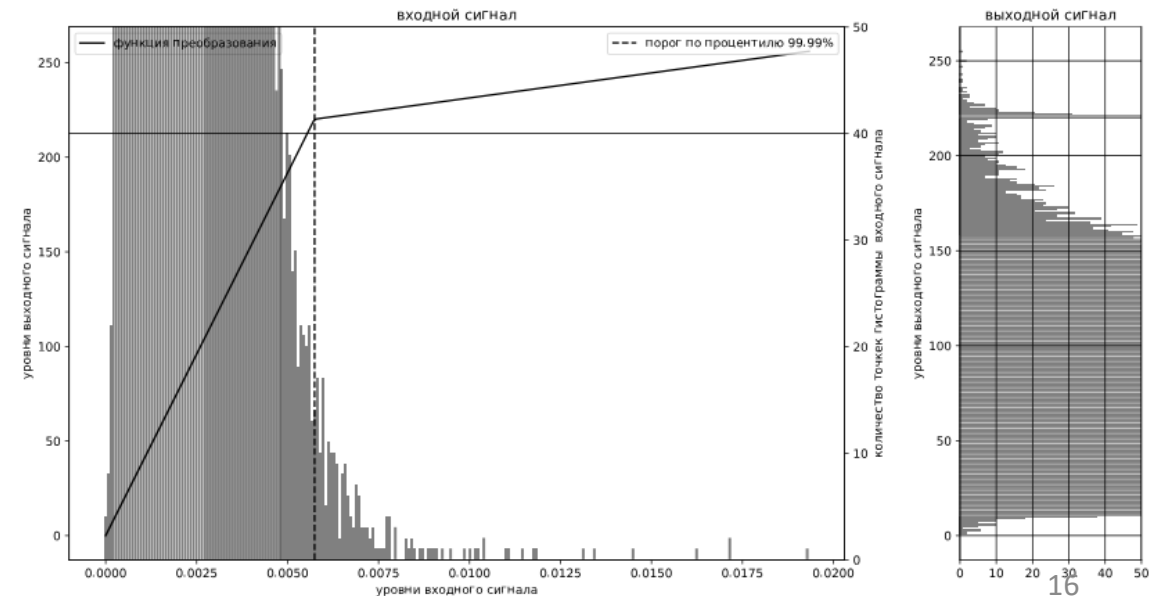


- Используемая реализация оптического потока требует входных данных в формате однобайтовое целое.
- В новых версиях для учёта редких экстремальных, но значимых величин при преобразовании данных в нужную точность применяется кусочно-линейная компрессия по порогу, определяемому заданным процентилем.
- И масштабирование и компрессия применяется к паре обрабатываемых кадров с общими параметрами, что улучшает оценку смещений.

$$k_1 = \frac{T}{p_1}, \quad k_2 = \frac{1 - T}{s_{\max} - p_1}$$

$$S_{\text{compressed}} = \text{int8} \left( 255 \cdot \begin{cases} k_1 \cdot s, & s \leq p_1 \\ T + k_2 \cdot (s - p_1), & s > p_1 \end{cases} \right)$$

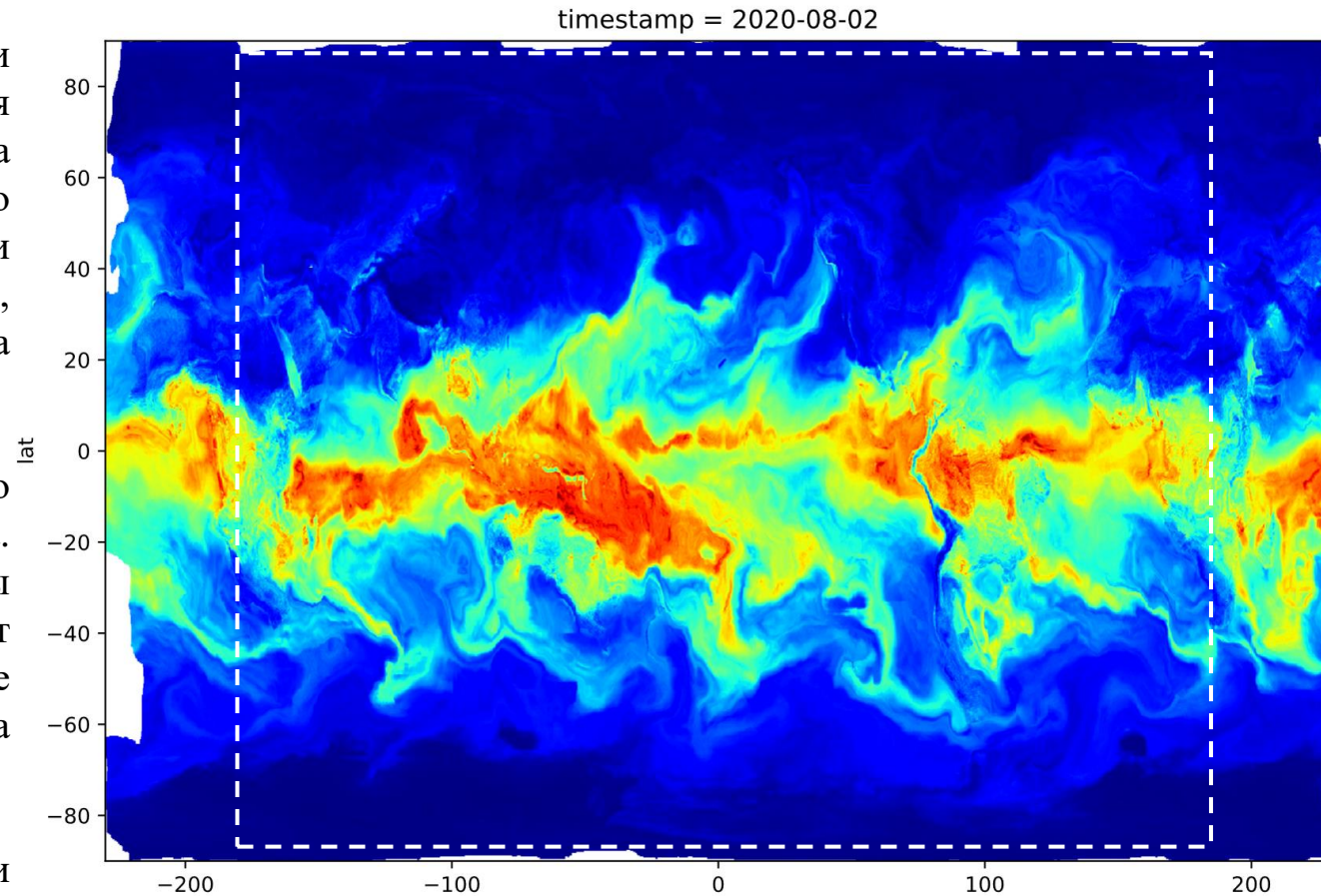
где:  $S$  — исходный сигнал,  $p_1 = Q_\alpha(S)$  -  $\alpha$ -й процентиль сигнала, где  $\alpha$  — выбранный уровень процентиля.  
 $s_{\max}(S)$  — максимальное значение сигнала,  $T$  — пороговое значение (outThreshold).



При оценке оптического потока и построении интерполяционных кадров важен контекст для каждого участка. Чтобы качество работы на границах не снижалось, к кадру временно присоединяется фрагмент кадра предыдущих и следующих суток по долготе. В последующем, после обработки, области, выходящие за реальные координаты снова отрезаются.

На рисунке показана реализация вычисленного смещения для условного тестового кадра. Видно, что области, в которые перенесены пиксели из-за условных границ- пусты. За счёт расширения кадра такие дефекты не затрагивают реальные границы кадра (показаны пунктиром).

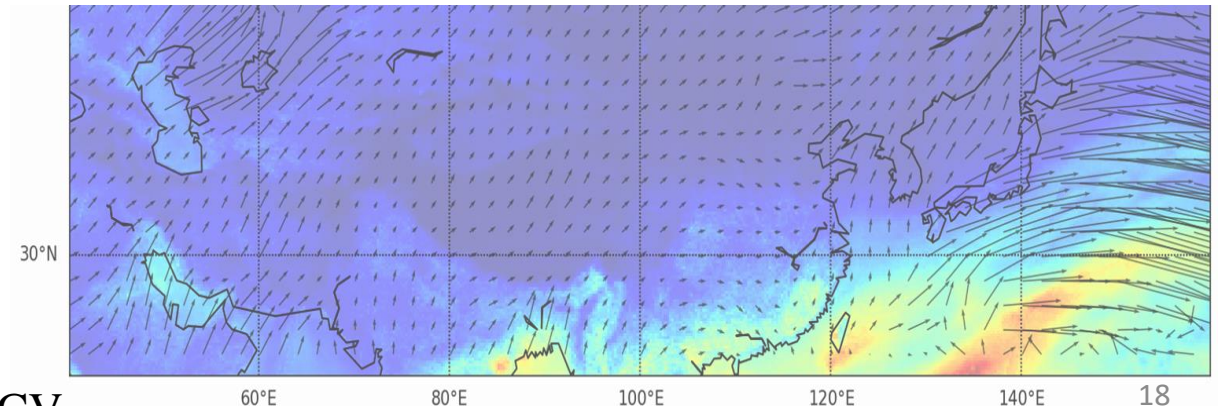
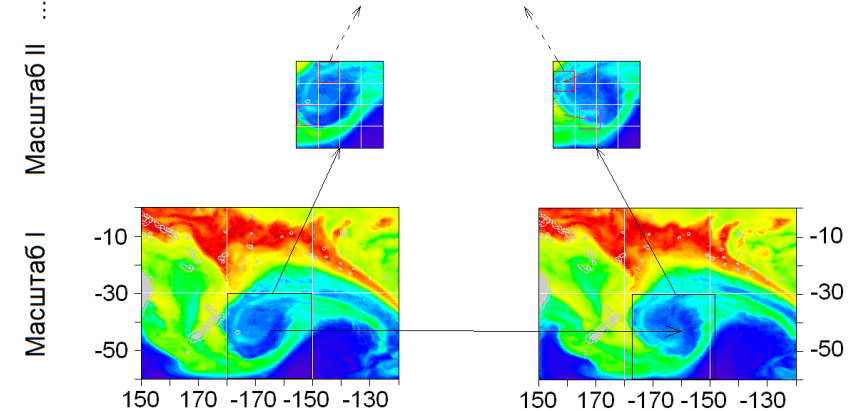
В последней экспериментальной версии аналогичная процедура выполняется для полярных областей.



Используется алгоритм DIS (*англ.* DenseInverseSearch), предложенный в работе (Kroeger et al., 2016).

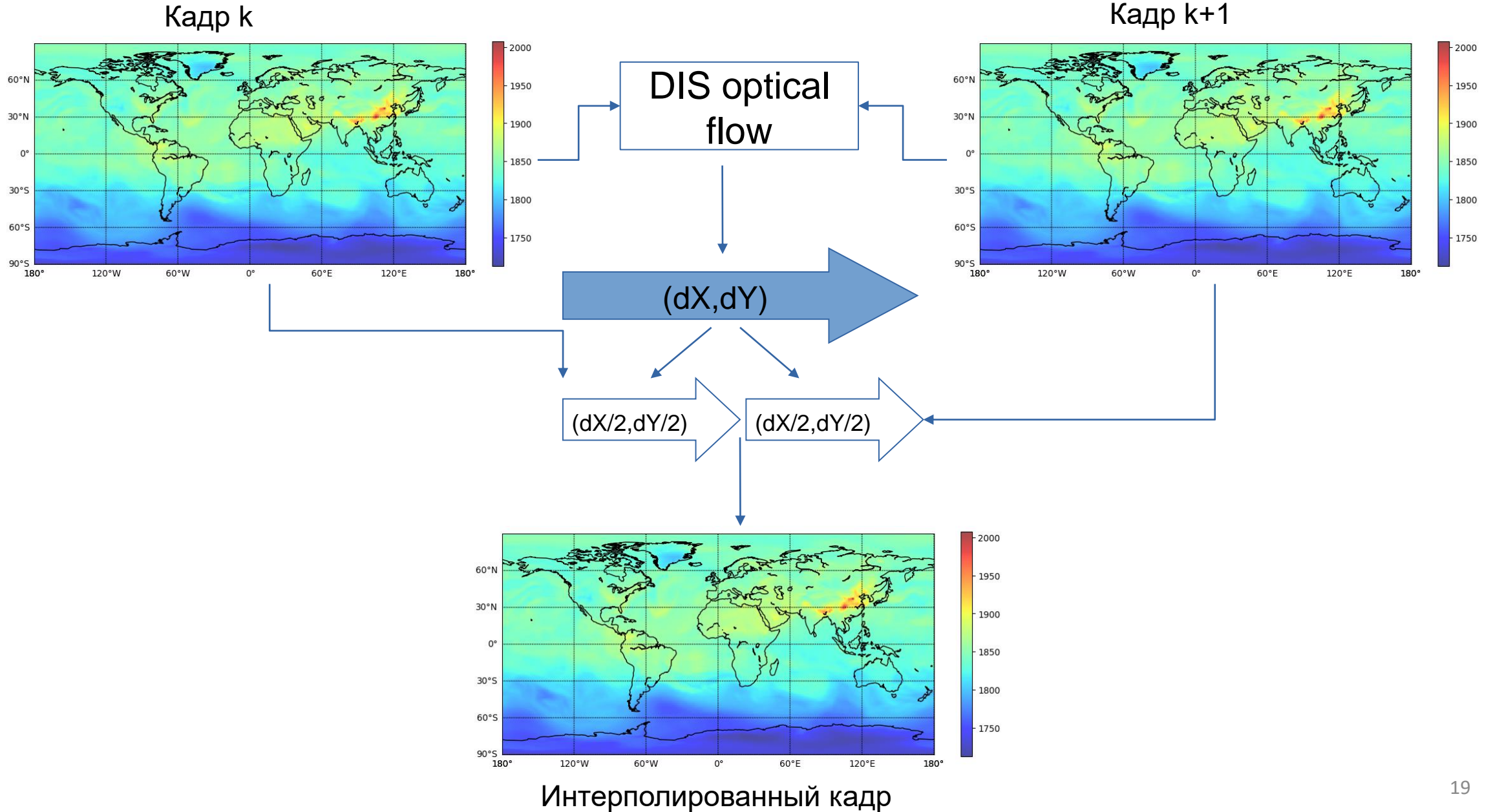
Преимущества для данной задачи:

- Вычисляет плотный оптический поток (в каждой точке исходного изображения)
- На каждой итерации рассчитывается промежуточное плотное поле перемещений
  - повышает сглаженность
  - улучшает производительность
- Вариационная корректировка с критерием гладкости улучшает фильтрацию «выбросов»
- возможность анализа большого диапазона масштабов смещений
  - субпиксельная интерполяция для малых масштабов
  - анализ пирамиды изображений для больших масштабов

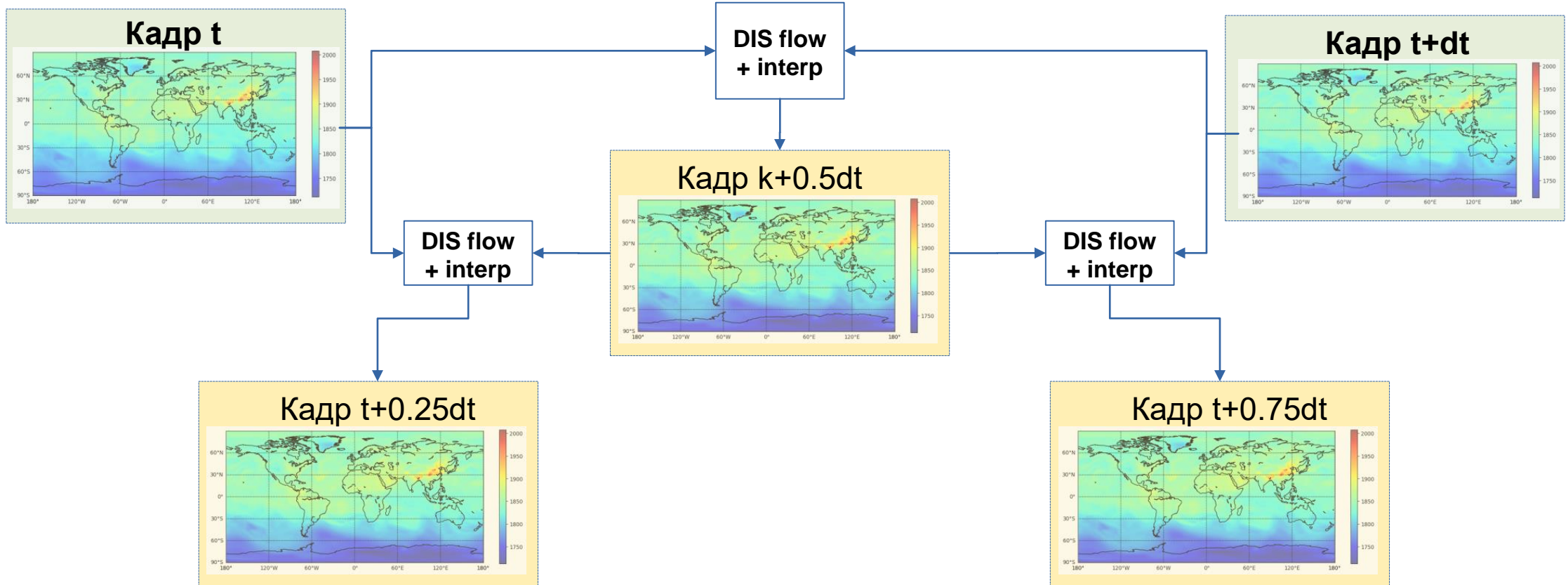


Используется реализация из открытой библиотеки openCV

# Построение интерполированных кадров



# Построение интерполированных кадров



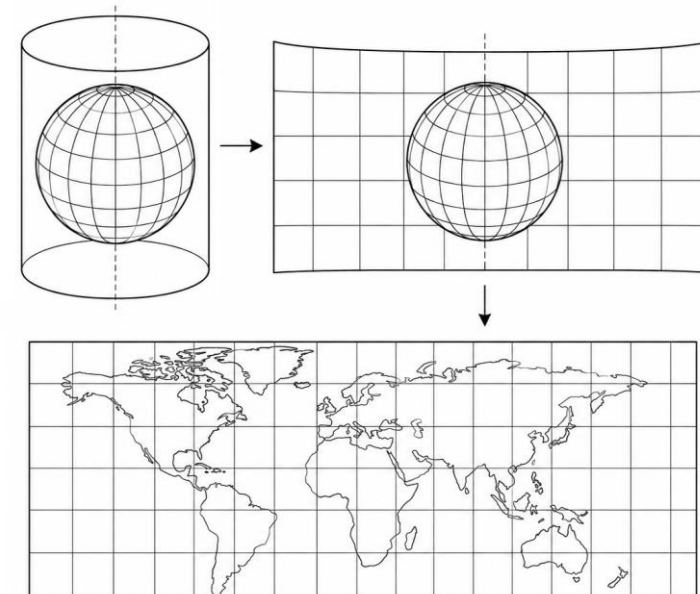
- Сырой результат оценки оптического потока — векторы смещений пикселей в проекции заданного изображения. При этом плоское изображение представляет собой цилиндрическую проекцию, следовательно, для перехода к реальным линейным размерам требуется преобразование

- Пикселы преобразуются в **градусы**, соответственно принятой равномерной сетке

- По формулам гаверсина, оптимизированной для цикла по матрице вычисляются соответствующие дистанции в **метрах**.

$$d = 2r \arcsin \left( \sqrt{\sin^2 \left( \frac{\phi_2 - \phi_1}{2} \right) + \cos(\phi_1) \cos(\phi_2) \sin^2 \left( \frac{\lambda_2 - \lambda_1}{2} \right)} \right)$$

- Для получения скорости в **м/с** используется промежуток времени, между условными временами кадра.

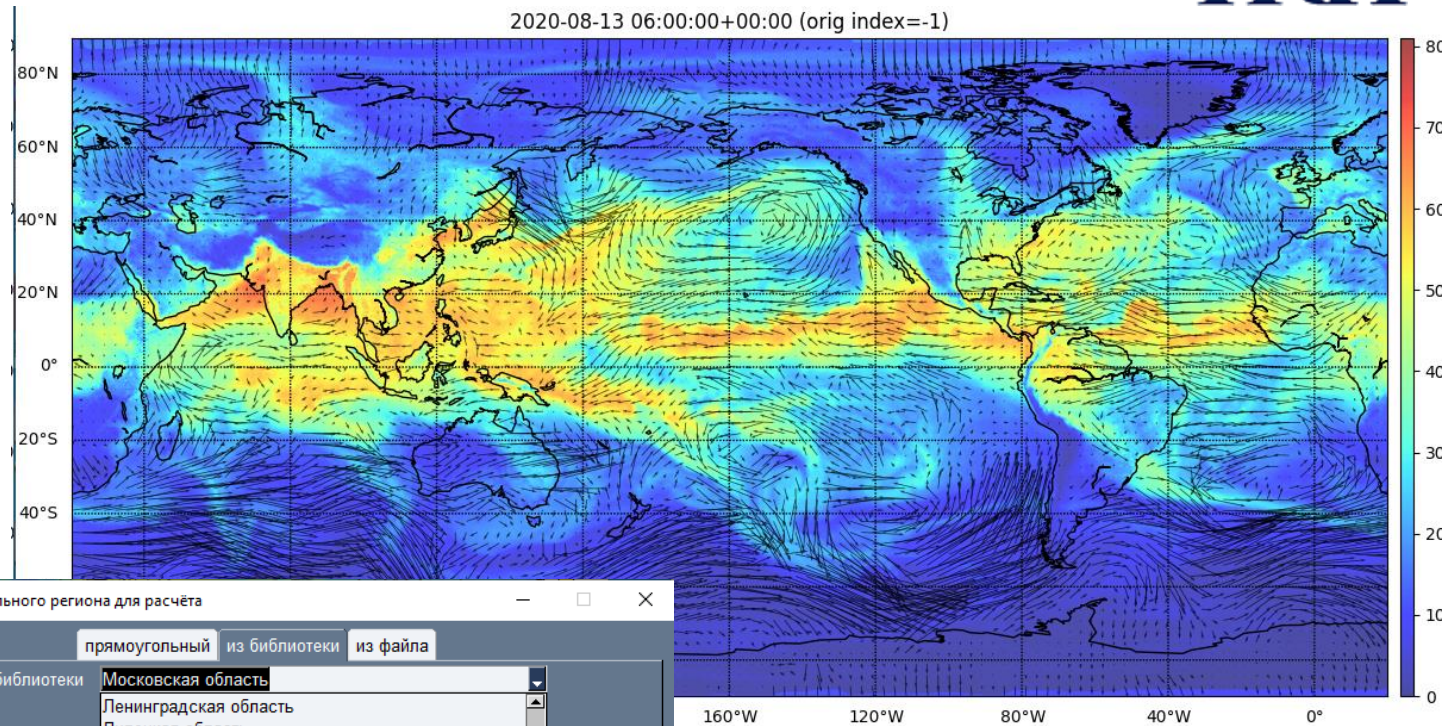


# База данных и инструмент работы с ней



| Local File      | Local File      | Local File |
|-----------------|-----------------|------------|
| htime           | htime           | 1D         |
| index           | index           | 1D         |
| lat             | lat             | 1D         |
| lon             | lon             | 1D         |
| PWV             | PWV             | Geo2D      |
| stime           | stime           | —          |
| Time            | Time            | Geo2D      |
| timeshort       | timeshort       | 1D         |
| timeshortFields | timeshortFields | —          |
| timestamp       | timestamp       | 1D         |
| U               | U               | Geo2D      |
| V               | V               | Geo2D      |

ViewFlow



Вычисление по нескольким кадрам

Доступные даты: 2012-01-01 00:00:00 – 2024-12-31 00:00:00

ОТ даты: 01.01.2012 00:00:00    ДО даты: 01.01.2012 12:00:00

прореживание (оставьте поле пустым чтобы выключить прореживание):

set region...

напоминание: регион для расчёта будет округляется до ближайшего узла сетки!

конвергенция    перенос    сумма и баланс    значение в точках    экспорт видео    параметры региона    скользящее окно

результаты (предпросмотр): (могут быть показаны не полностью)

экспортировать в csv    экспортировать в xlsx    в .MAT    в netCDF

close

Выбор произвольного региона для расчёта

прямоугольный    из библиотеки    из файла

из библиотеки

- Московская область
- Ленинградская область
- Липецкая область
- Луганская Народная Республика
- Магаданская область
- Москва
- Московская область
- Мурманская область
- Новгородская область
- Новосибирская область
- Омская область

preview

Московская область

```

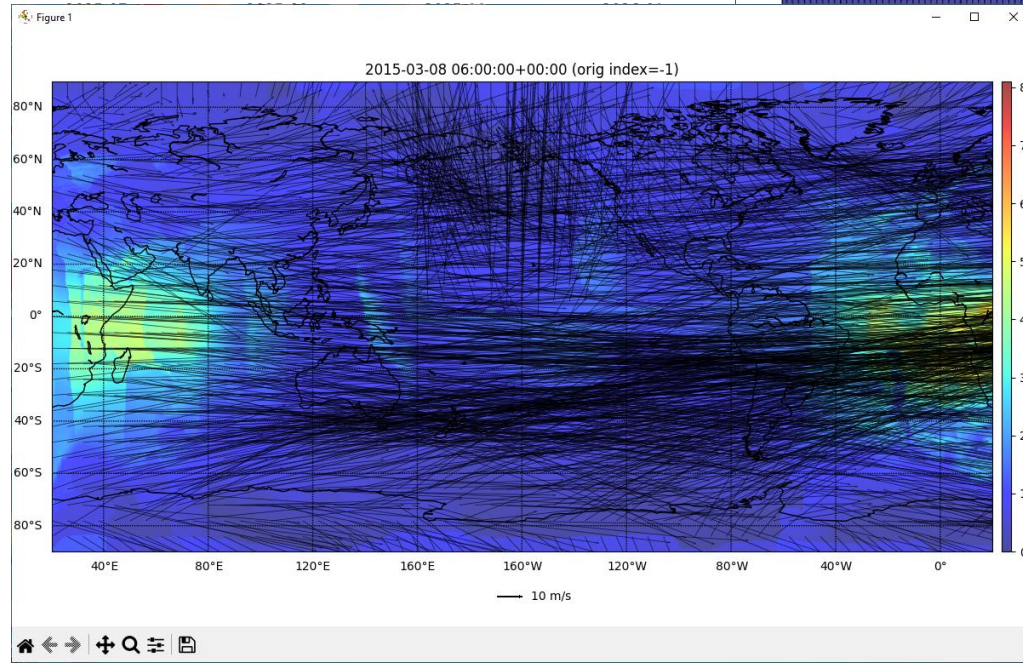
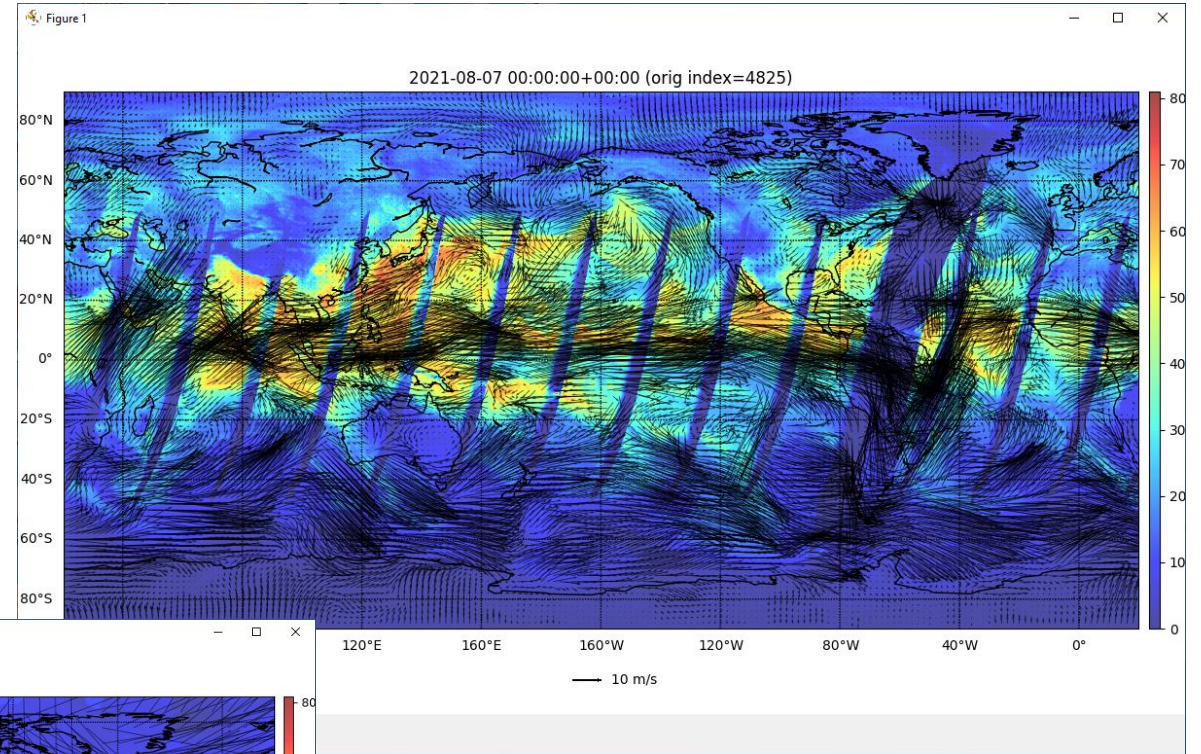
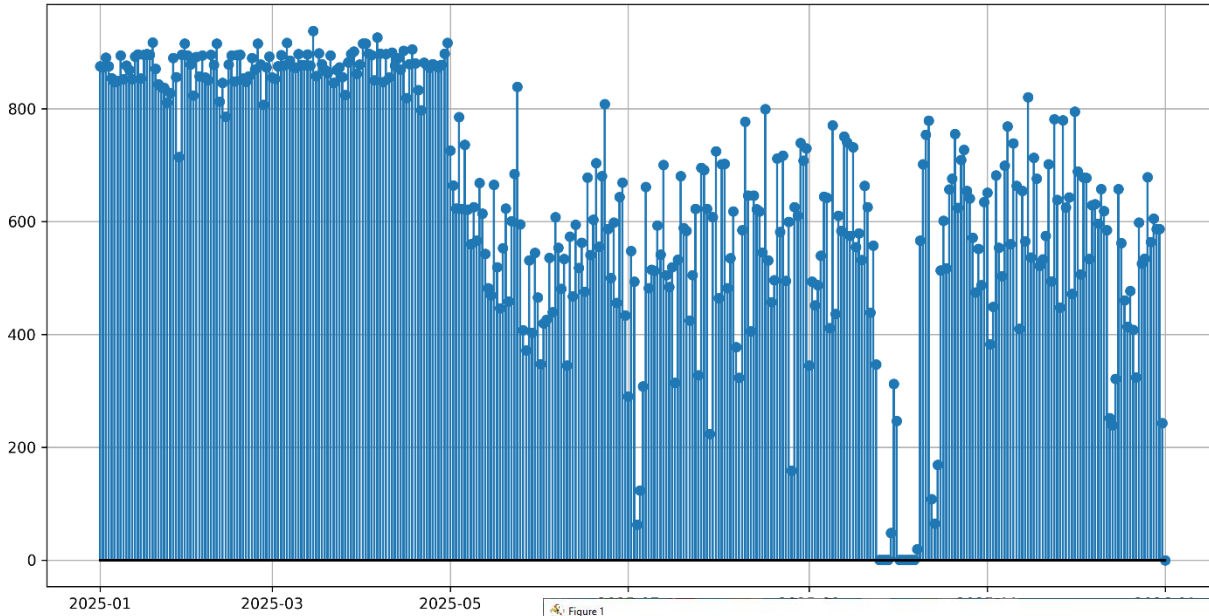
[[37.625 56.875]
 [37.375 56.875]
 [37.125 56.875]
 [37.125 56.625]
 [36.875 56.625]
 [36.875 56.375]
 [36.625 56.375]
 [36.375 56.375]
 [36.125 56.375]
 [35.875 56.375]

```

- Сформирована многолетняя база данных интегрального содержания и потоков водяного пара за период с 2012 по 2025 год
- *Breakthrough* уровень ВМО по полям ИВС для климатических исследований и глобального численного прогноза погоды
- Технология распространяется на другие малые газовые оставляющие
- Создан инструмент для взаимодействия с базой и всестороннего анализа данных
- База использовалась более чем в 10 научных работах
- С использованием базы было защищено 5 магистерских диссертаций

# Проблемы

Общий объём на дату (MiB)



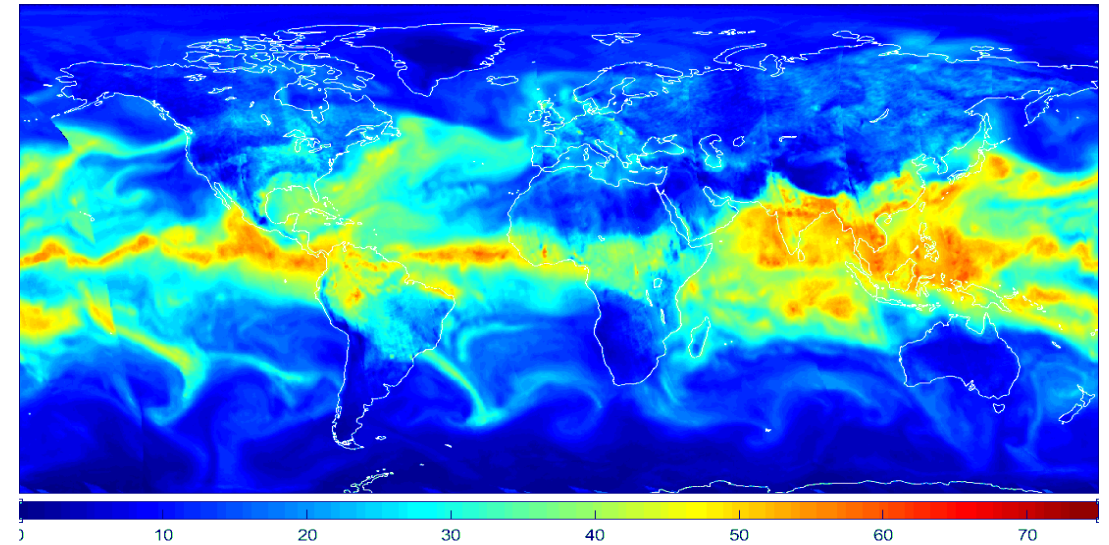
## AMSR3 SWATH 1450 km 2025 - 2032

| Channel number | Central frequency | Bandwidth | Polarisation | NEAT   | IFOV     | Pixel    |
|----------------|-------------------|-----------|--------------|--------|----------|----------|
| 1              | 6.925 GHz         | 350 MHz   | H, V         | 0.34 K | 34x58 km | 10x10 km |
| 2              | 7.3 GHz           | 350 MHz   | H, V         | 0.34 K | 34x58 km | 10x10 km |
| 3              | 10.25 GHz         | 500 MHz   | H, V         | 0.34 K | 22x39 km | 10x10 km |
| 4              | 10.65 GHz         | 100 MHz   | H, V         | 0.7 K  | 22x39 km | 10x10 km |
| 5              | 18.7 GHz          | 200 MHz   | H, V         | 0.7 K  | 12x21 km | 10x10 km |
| 6              | 23.8 GHz          | 400 MHz   | H, V         | 0.6 K  | 14x24 km | 10x10 km |
| 7              | 36.42GHz          | 840 MHz   | H, V         | 0.7 K  | 7x11 km  | 10x10 km |
| 8              | 89.0 GHz (A)      | 3000 MHz  | H, V         | 1.2 K  | 3x5 km   | 5x5 km   |
| 9              | 89.0 GHz (B)      | 3000 MHz  | H, V         | 1.2 K  | 3x5 km   | 5x5 km   |
| 10             | 165.5 GHz         | 3200 MHz  | V            | 1.5 K  | 4x9 km   | 5x5 km   |
| 11             | 183.31±7 GHz      | 1700 MHz  | V            | 1.5 K  | 4x8 km   | 5x5 km   |
| 12             | 183.31±3 GHz      | 1470 MHz  | V            | 1.5 K  | 4x8 km   | 5x5 km   |

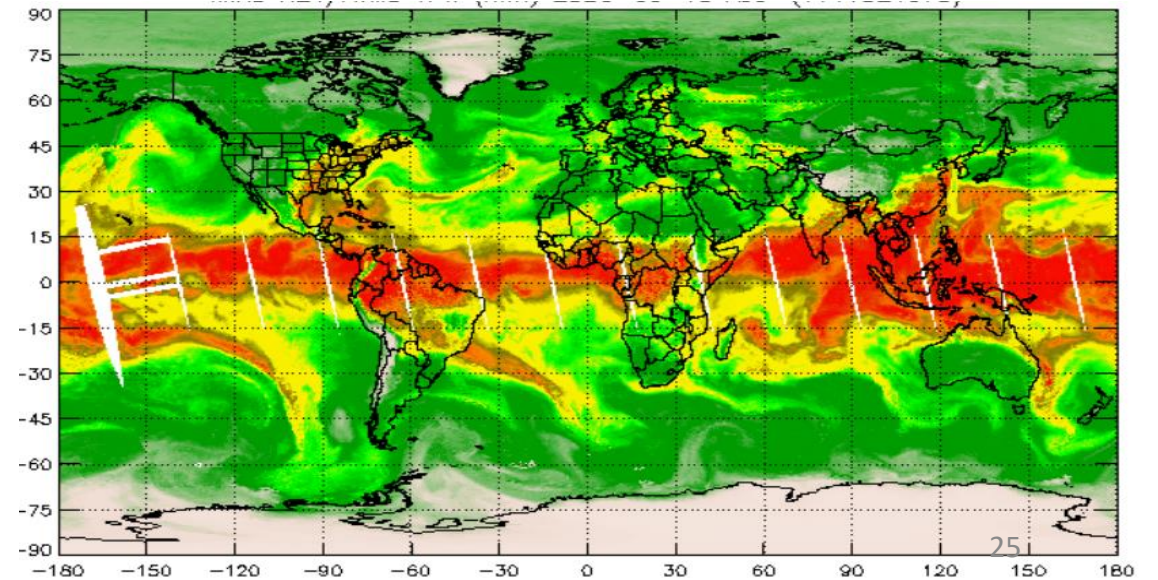
## MWRI-2 SWATH 1400 km 2026 - 2033

| Channel Number | Central frequency (GHz) | Bandwidth (MHz) | Polarisations | NEAT  | IFOV       | Pixel     |
|----------------|-------------------------|-----------------|---------------|-------|------------|-----------|
| 1              | 10.65                   | 180             | V, H          | 0.5 K | 27 x 45 km | 4 x 14 km |
| 2              | 18.7                    | 200             | V, H          | 0.5 K | 18 x 30 km | 4 x 14 km |
| 3              | 23.8                    | 400             | V, H          | 0.5 K | 17 x 29 km | 4 x 14 km |
| 4              | 36.5                    | 900             | V, H          | 0.5 K | 11 x 18 km | 4 x 14 km |
| 5              | 50.3                    | 400             | V, H          | 0.5 K | 9 x 14 km  | 4 x 14 km |
| 6              | 52.61                   | 400             | V, H          | 0.5 K | 9 x 14 km  | 4 x 14 km |
| 7              | 53.24                   | 400             | V, H          | 0.5 K | 9 x 14 km  | 4 x 14 km |
| 8              | 53.75                   | 400             | V, H          | 0.5 K | 9 x 14 km  | 4 x 14 km |
| 9              | 89                      | 3000            | V, H          | 0.5 K | 7 x 11 km  | 4 x 14 km |
| 10             | 118.7503±3.2            | 2 x 500         | V             | 0.8 K | 6 x 10 km  | 4 x 14 km |
| 11             | 118.7503±2.1            | 2 x 400         | V             | 0.8 K | 6 x 10 km  | 4 x 14 km |
| 12             | 118.7503±1.4            | 2 x 400         | V             | 0.8 K | 6 x 10 km  | 4 x 14 km |
| 13             | 118.7503±1.2            | 2 x 400         | V             | 0.8 K | 6 x 10 km  | 4 x 14 km |

## MTB3A-ГЯ SWATH 2200 km

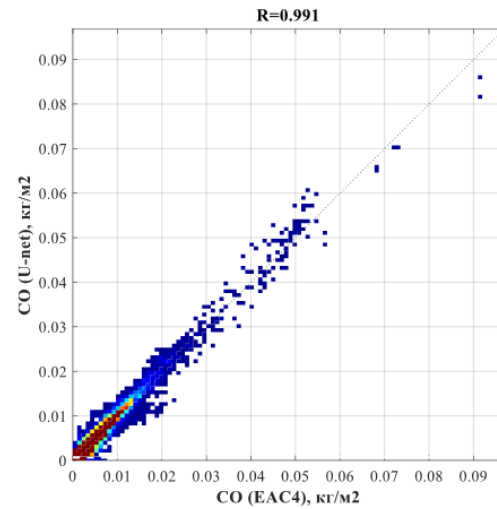
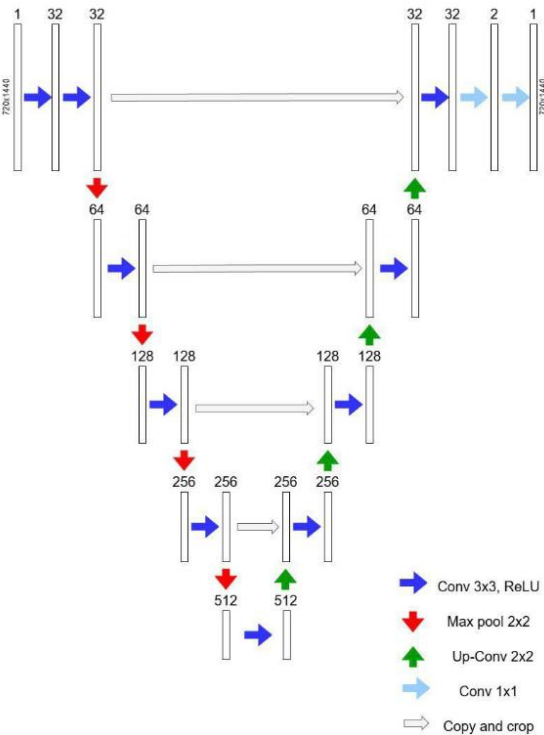


## ATMS SWATH 2200 km

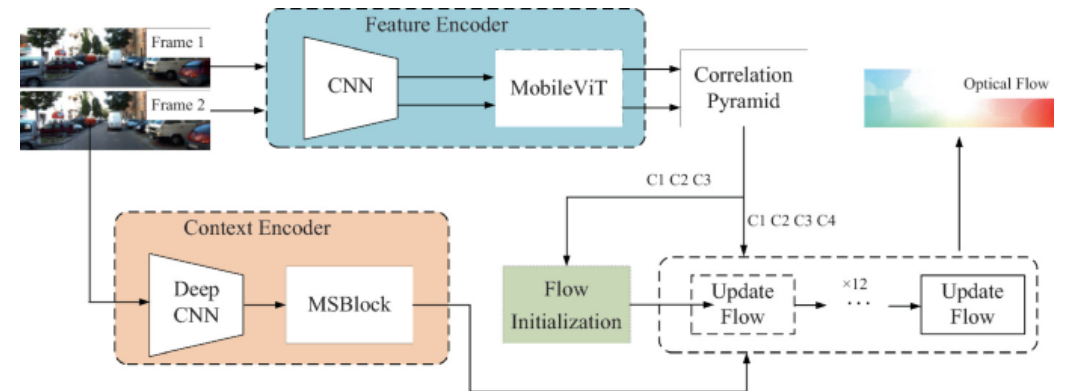


## Применение нейросетей

### U-net для заполнения пропусков



### RAFT для оценки оптического потока



Спасибо за внимание!

Работа выполнена при поддержке темы «Эмиссия» гос. рег. № 12511112878-8