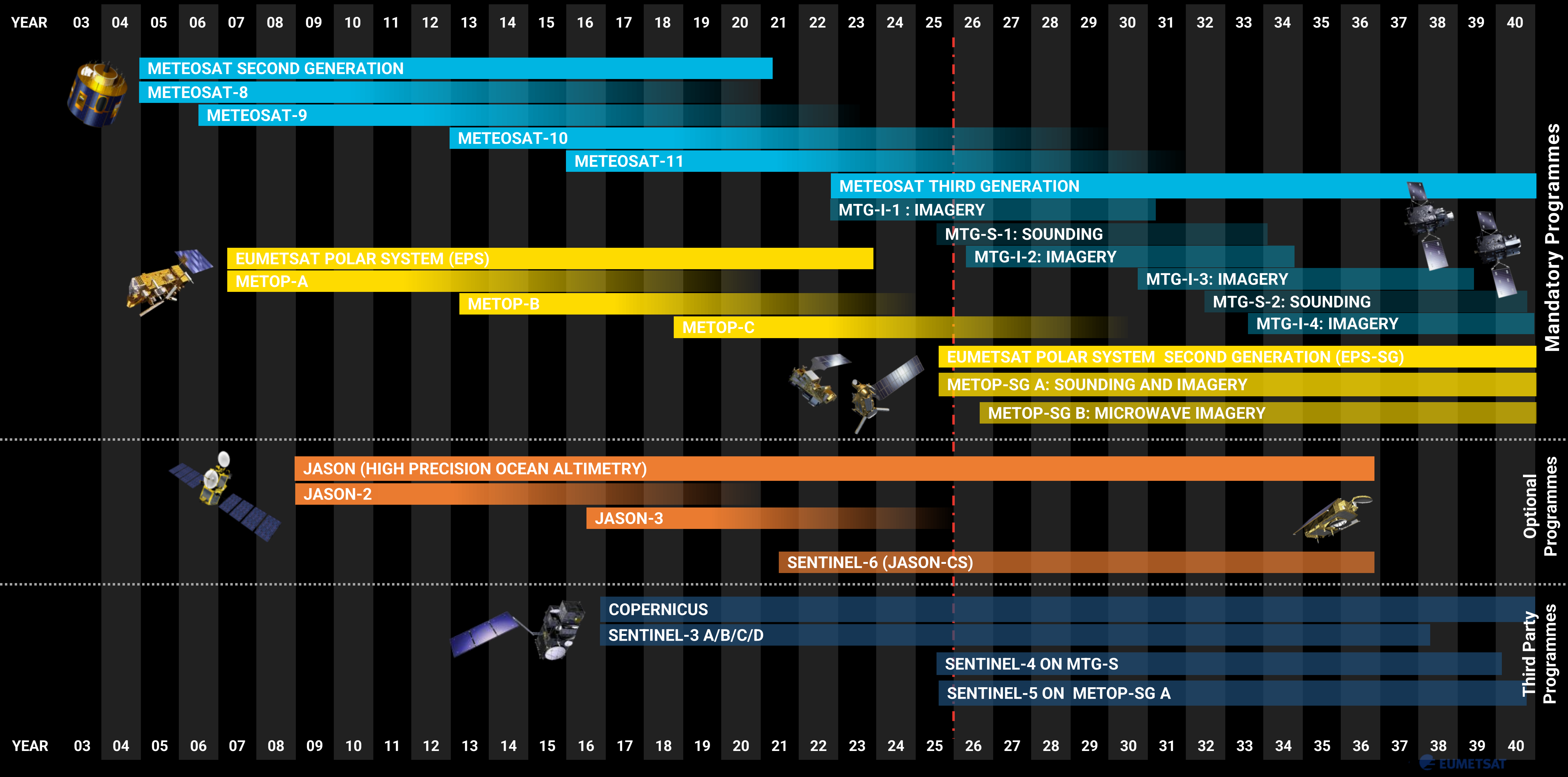
EUMETSAT
Data-driven EO for ML

Dr Lauren Biermann
obo Dr Roope Tervo

European Organisation for the Exploitation of Meteorological Satellites

- EU•MET•SAT is an operational organisation formed in 1986, services the needs of member states for timely, accurate weather and climate data.
- Currently:
  - 30 member states
  - India, China, Japan, Korea, Canada and the USA.
  - 12 active satellites

EUMETSAT Mission Planning:



Polar Orbiting LEO and Geostationary Satellite Fleet:

METOP –B, -C , SGA1
(Local Equator Crossing Time - 9:30 AM)

LOW EARTH, SUN-SYNCHRONOUS ORBIT
EUMETSAT POLAR SYSTEM (EPS) / FIRST AND SECOND GENERATION

SENTINEL-3A & -3B (98.65° incl.)

LOW EARTH, SUN-SYNCHRONOUS ORBIT
COPERNICUS SATELLITES DELIVERING MARINE AND LAND OBSERVATIONS

JASON-3 & SENTINEL-6 (63° incl.)

LOW EARTH, NON-SYNCHRONOUS ORBIT
OCEAN SURFACE TOPOGRAPHY MISSION, SHARED WITH CNES/NOAA/EU

METEOSAT-10, 11, 12, 13

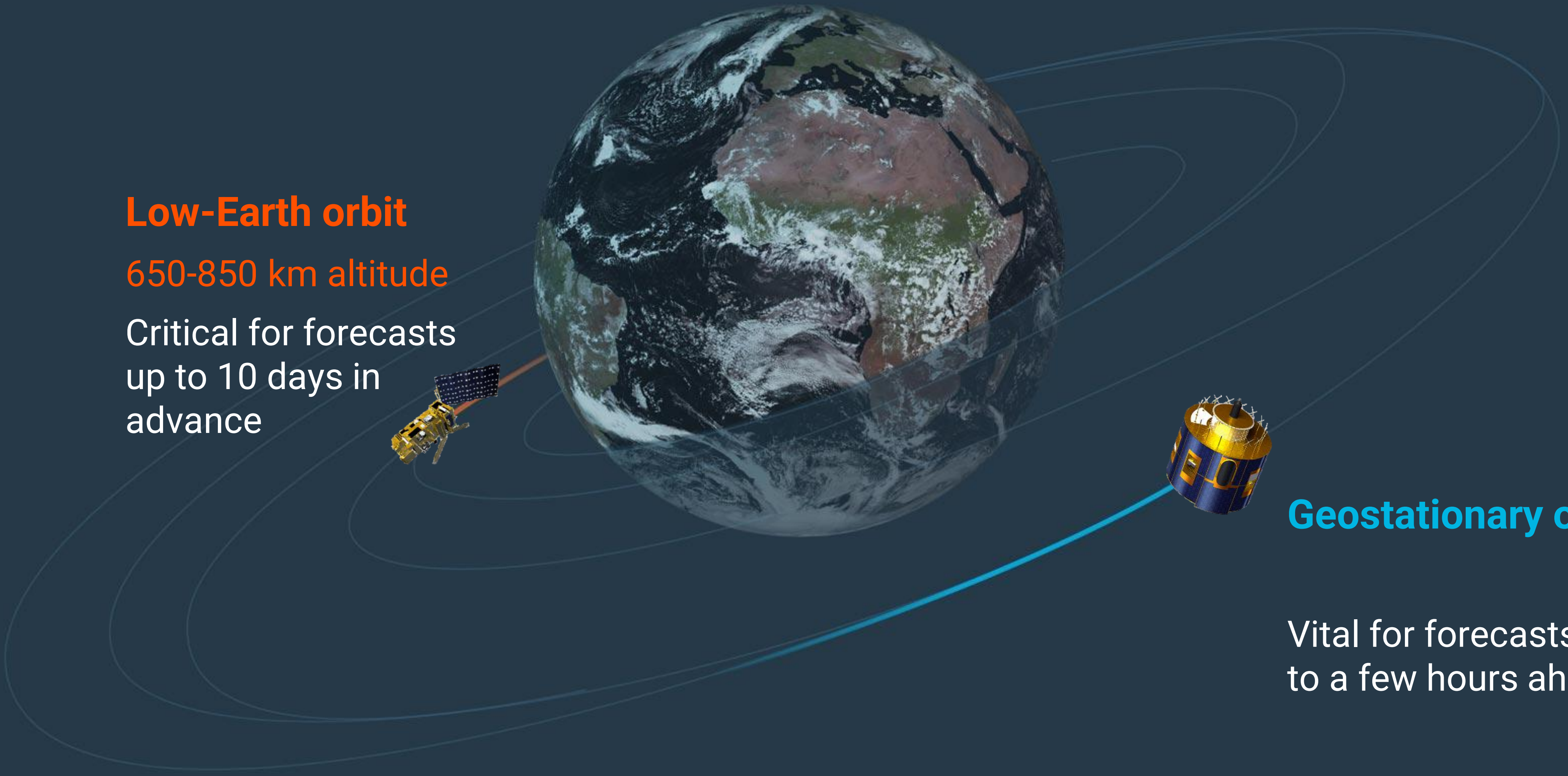
GEOSTATIONARY ORBIT	MULTI-SATELLITE SYSTEM
METEOSAT 2ND GENERATION	FULL DISC IMAGERY MISSION (15 MINS) (METEOSAT-11 (9.5°E)) RAPID SCAN SERVICE OVER EUROPE (5 MINS) (METEOSAT-10 (0.5° W))
METEOSAT 3RD GENERATION	FULL DISC IMAGERY MISSION (10 MINS) (METEOSAT-12 (0°)) ATMOSPHERIC SOUNDING IR (30 MINS) UV (60 MINS) (METEOSAT-13 (0°))

METEOSAT-9 (46° E)

GEOSTATIONARY ORBIT
METEOSAT 2ND GENERATION: PROVIDING INDIAN OCEAN DATA COVERAGE SERVICE (IODC)

The need for two types of EUMETSAT satellites

www.eumetsat.int



Low-Earth orbit

650-850 km altitude

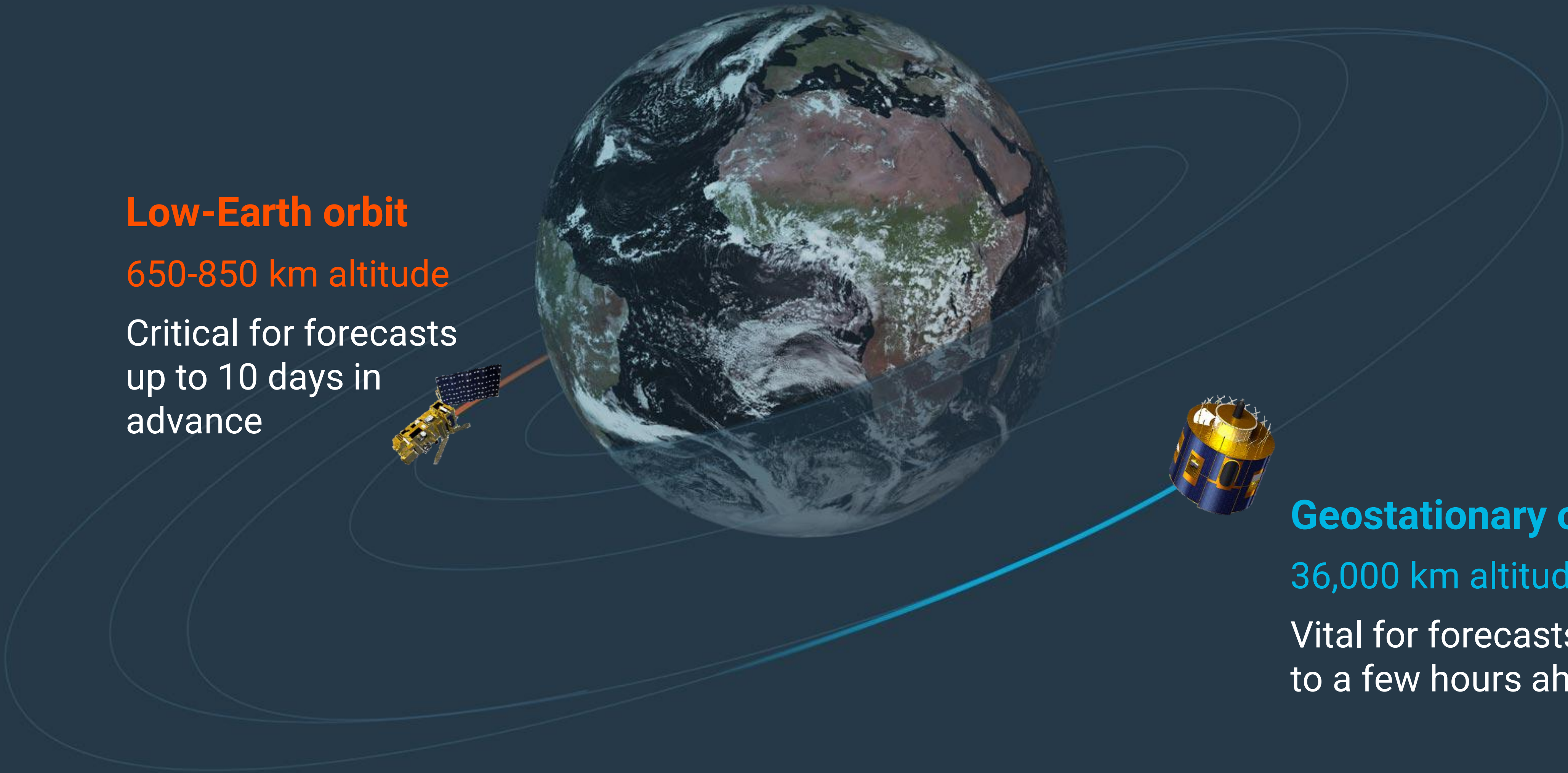
Critical for forecasts up to 10 days in advance

Geostationary orbit

Vital for forecasts up to a few hours ahead

The need for two types of EUMETSAT satellites

www.eumetsat.int



Low-Earth orbit

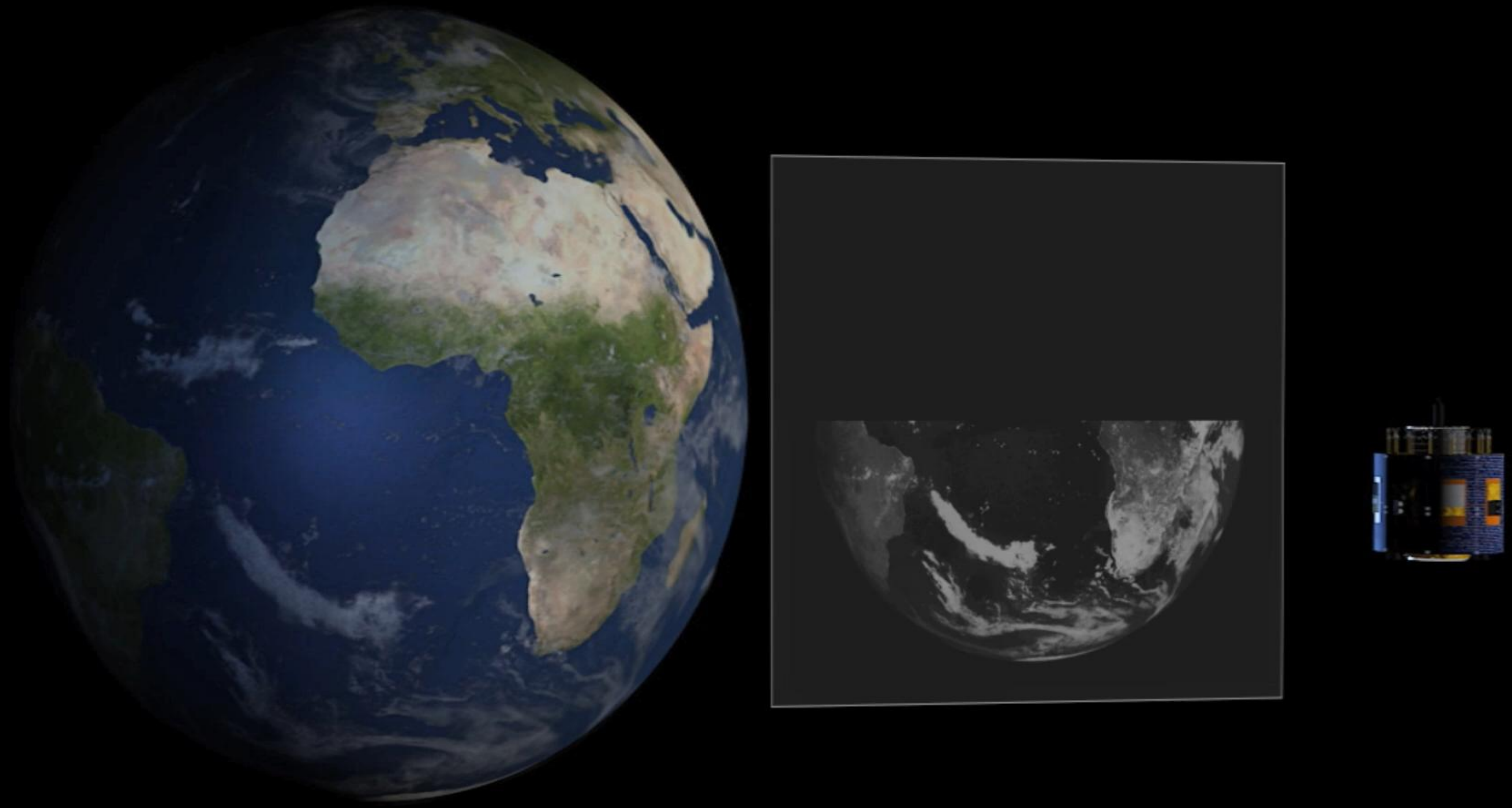
650-850 km altitude

Critical for forecasts up to 10 days in advance

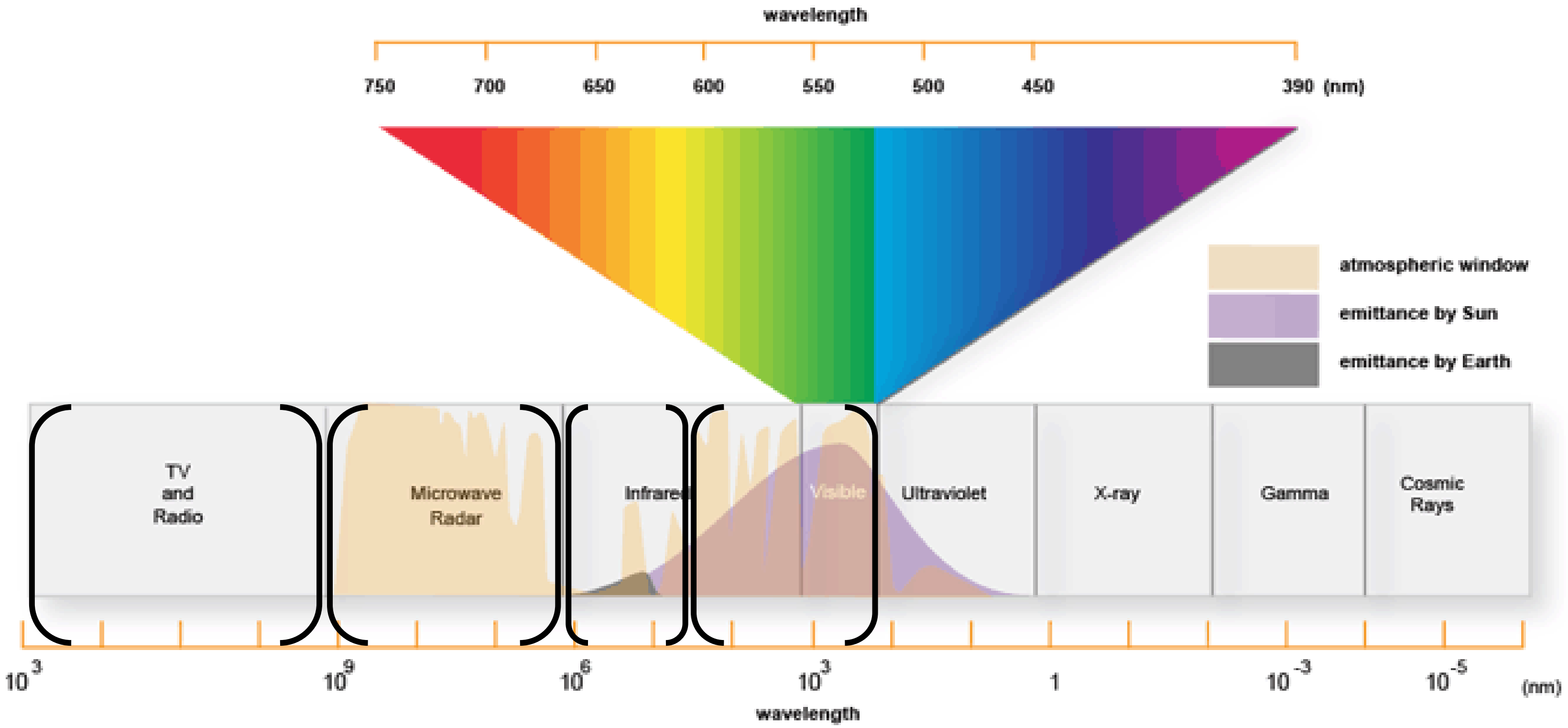
Geostationary orbit

36,000 km altitude

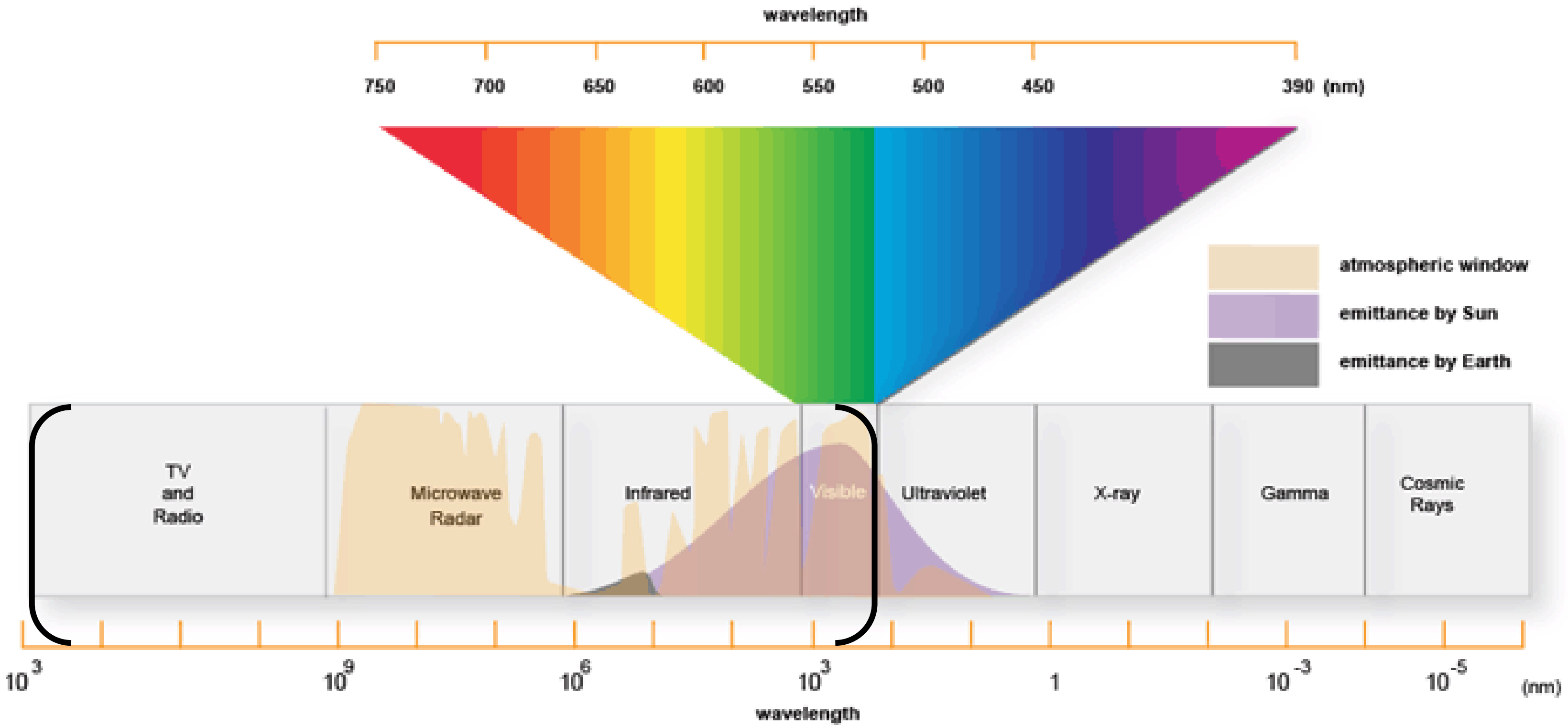
Vital for forecasts up to a few hours ahead



All satellite-based EO uses sections of the electromagnetic spectrum, and each region tells us different things about the properties of surfaces being observed.



All satellite-based EO uses sections of the electromagnetic spectrum, and each region tells us different things about the properties of surfaces being observed.



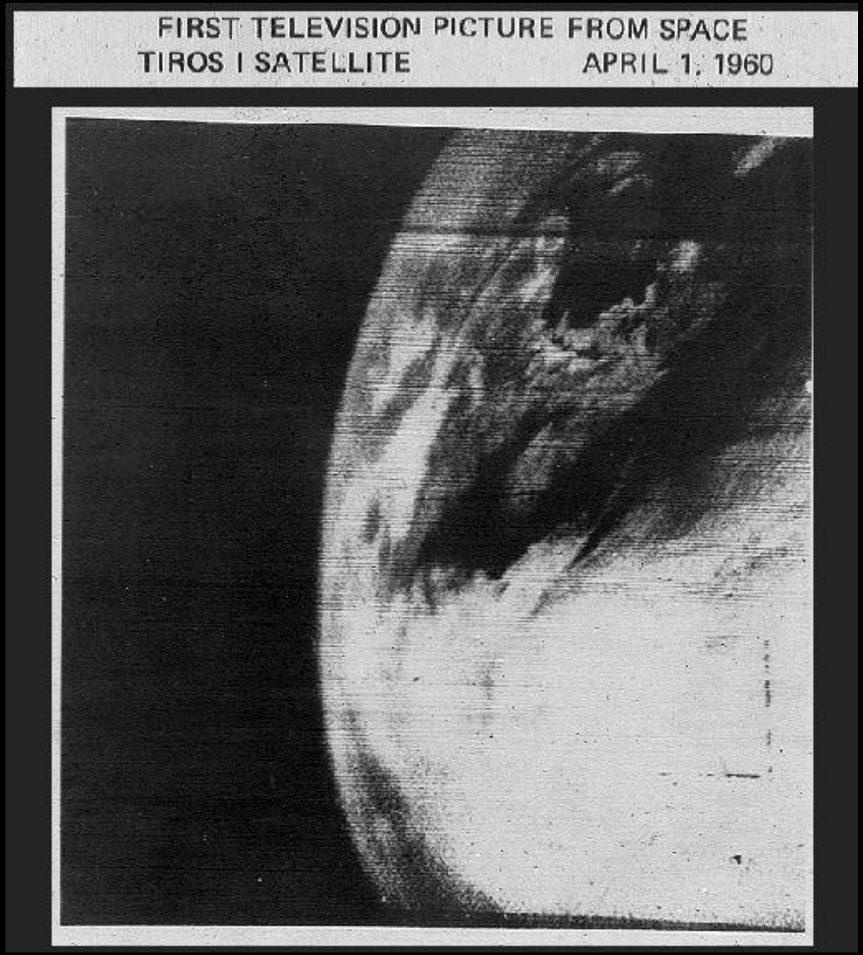
- **Passive** and **active** sensors are used for Earth Observation applications.
- Different bands can be used to observe different features in unique ways

HOW IT ALL BEGAN…

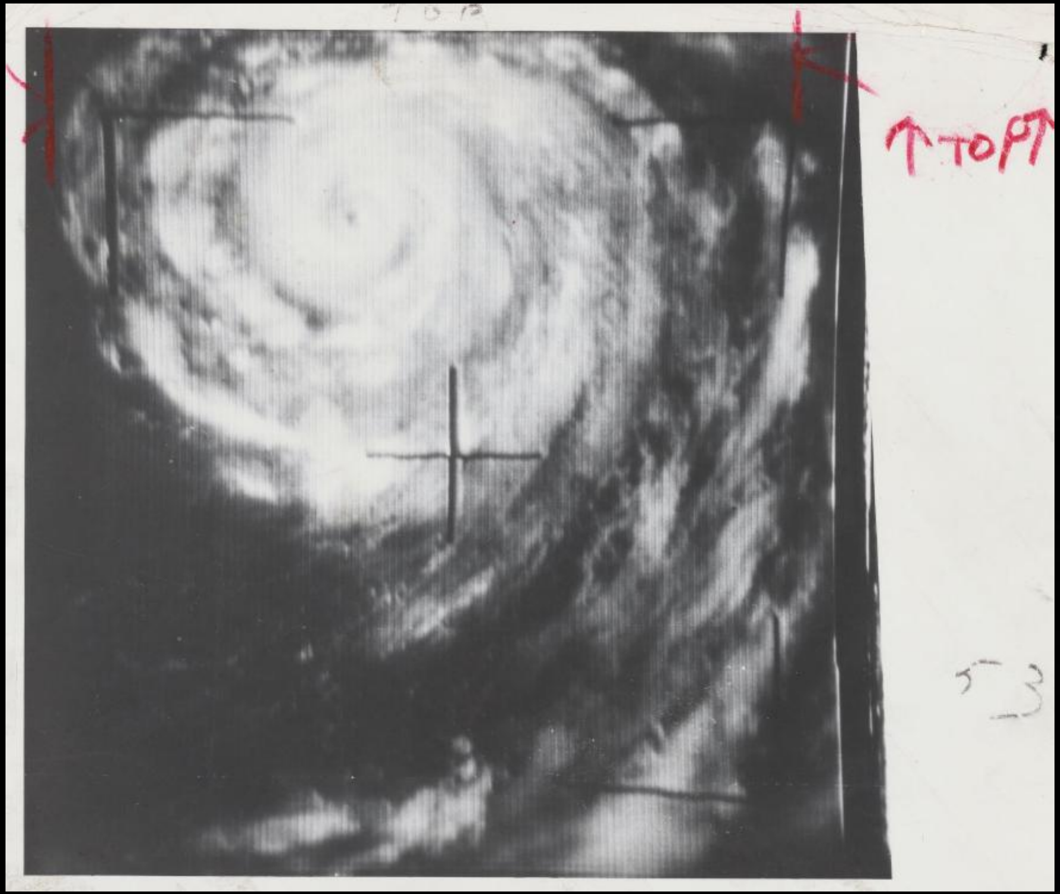
Sputnik 1 the first satellite

S- ('fellow') -putnik ('traveler') corresponding to Latin root satelles ('guard' or 'companion'), which is the origin of English word "satellite".

TIROS 1
the first meteorological satellite



TIROS 1
the first meteorological satellite



TIROS 1

the first meteorological satellite

NIMBUS foundation for Earth Observation

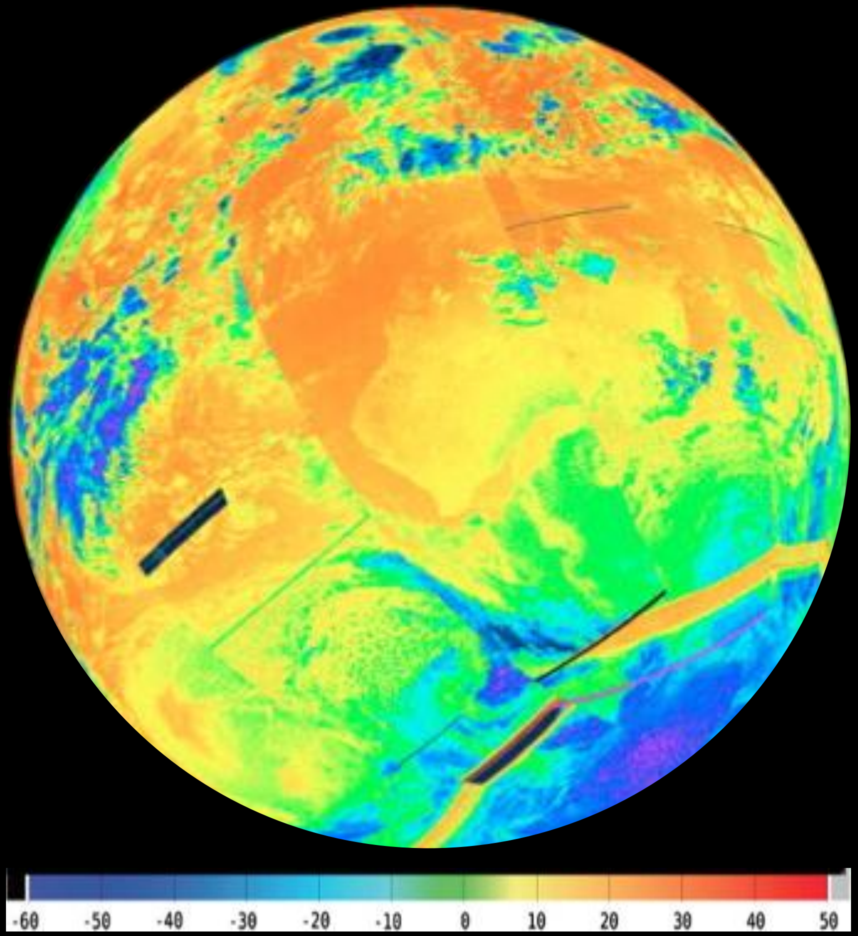


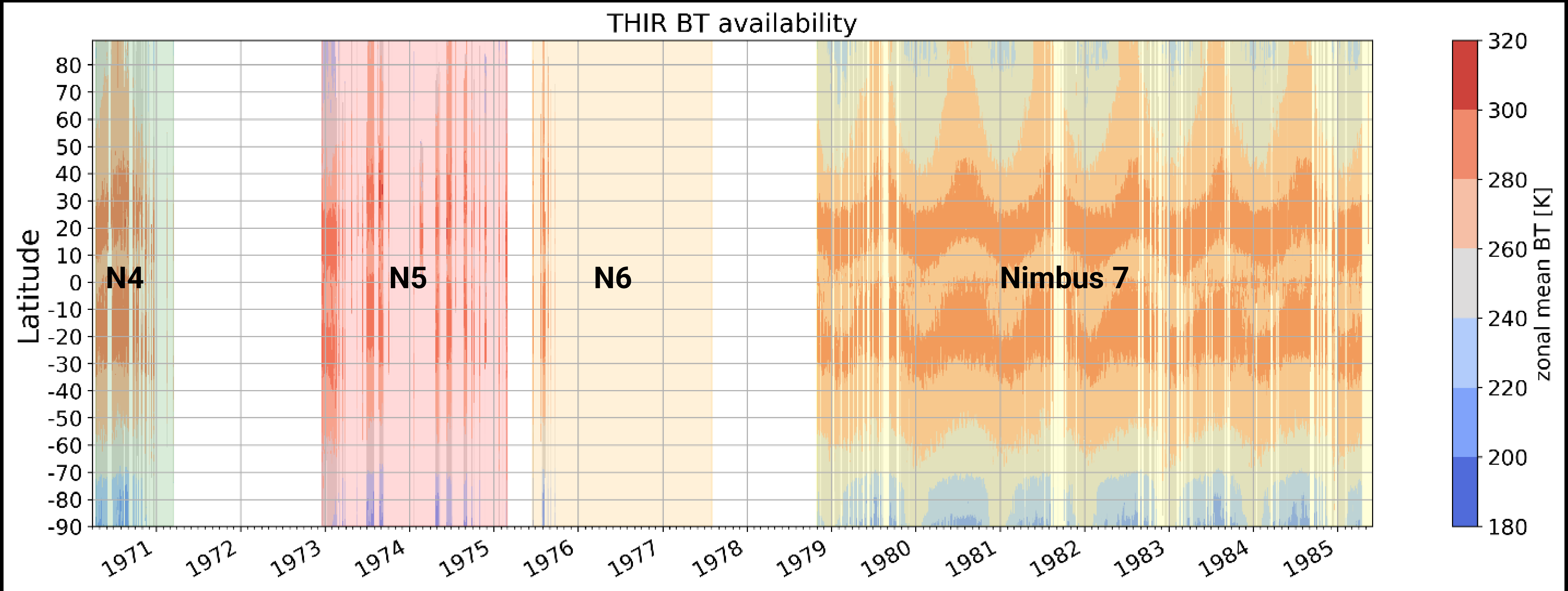
Figure from Gallaher et al. “The rescue of the early Nimbus satellite data”. 2015.

NIMBUS foundation for Earth Observation

- From the mid-1970s, EO datasets (esp. over data-sparse oceans) have substantially improved numerical weather predictions.
- WMO reports that improved early warning systems have contributed to a decline in global weather-related deaths:
  - >50,000 per year in the 1970s
  - <20,000 per year in the 2010s

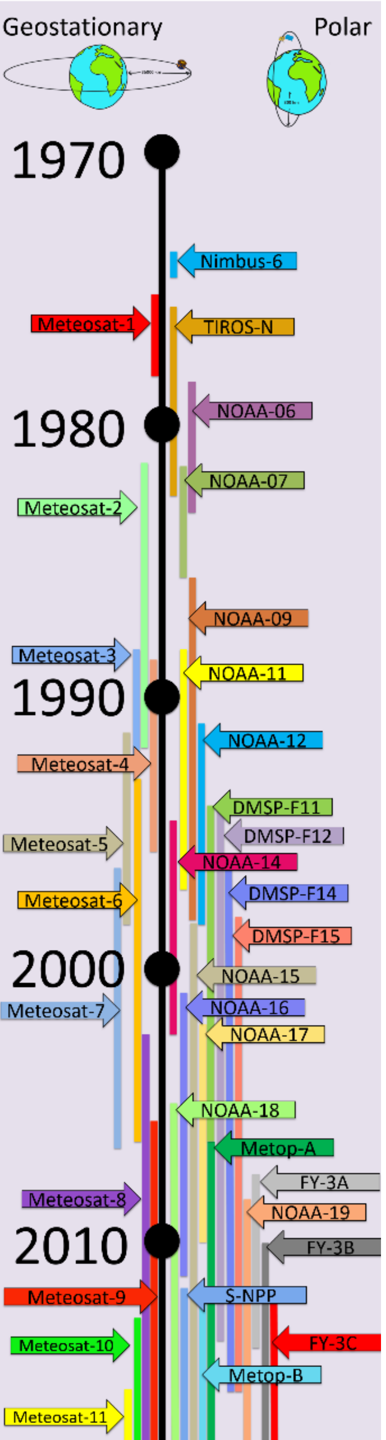
https://wmo.int/about-us/world-meteorological-day/wmd-2022/early-warnings-save-lives-and-livelihoods

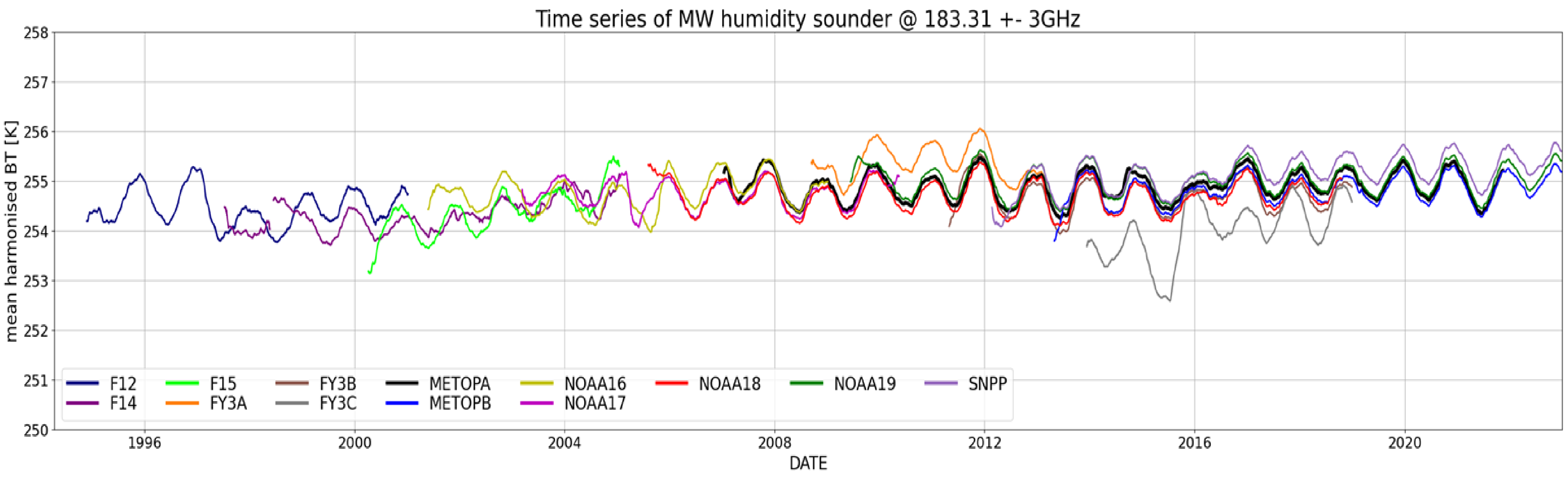
THIR Temperature Humidity Infrared Radiometer



- From Nimbus-4 onwards, wind estimates were derived from experimental cloud tracking, demonstrating the feasibility of satellite-derived Atmospheric Motion Vectors.

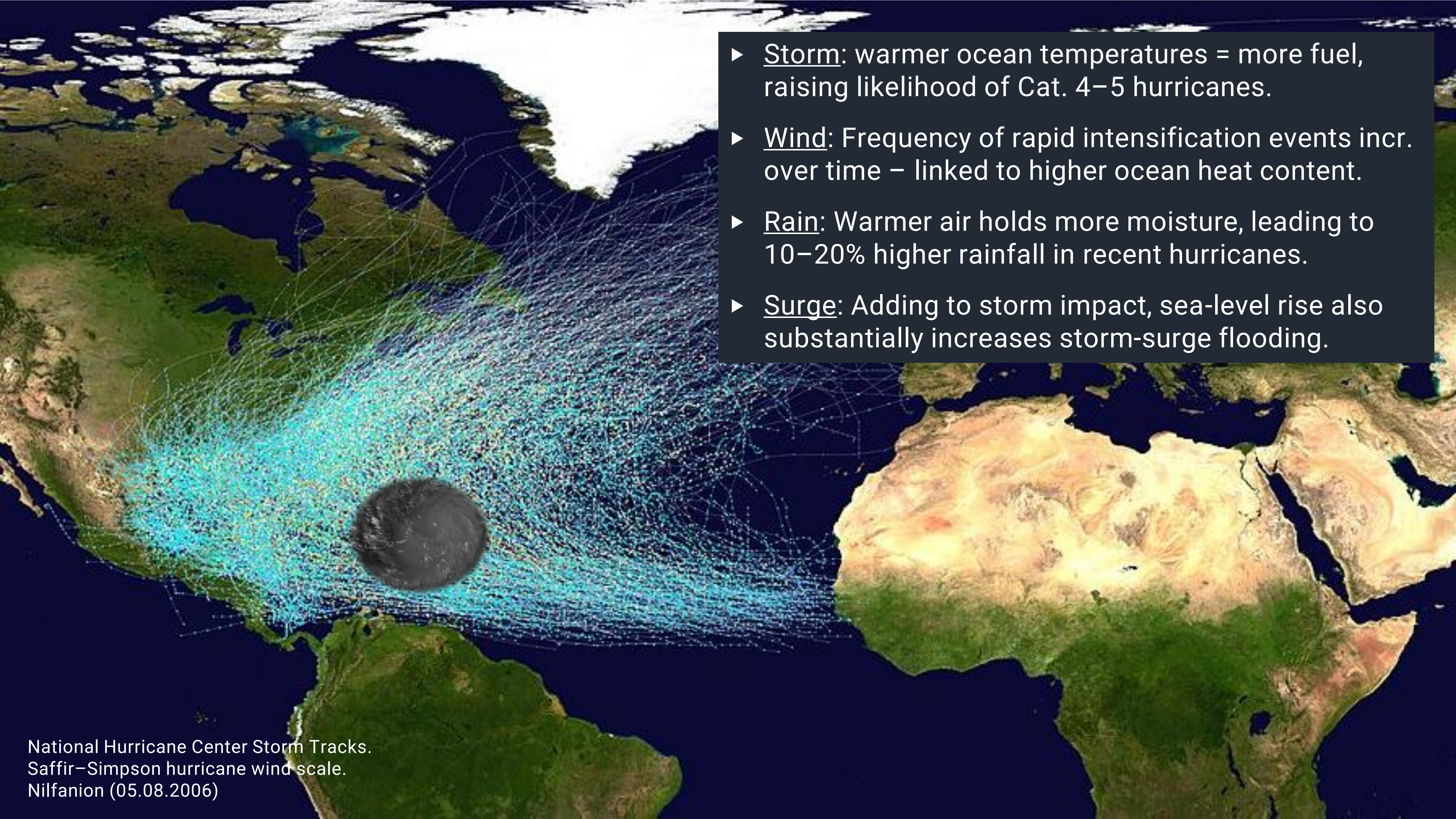
Harmonisation: 183.31 ± 3 GHz





Common reference definition:

Mapping of all observations onto an agreed “virtual channel” centered on the 183.31 ± 3 GHz water vapour absorption line.



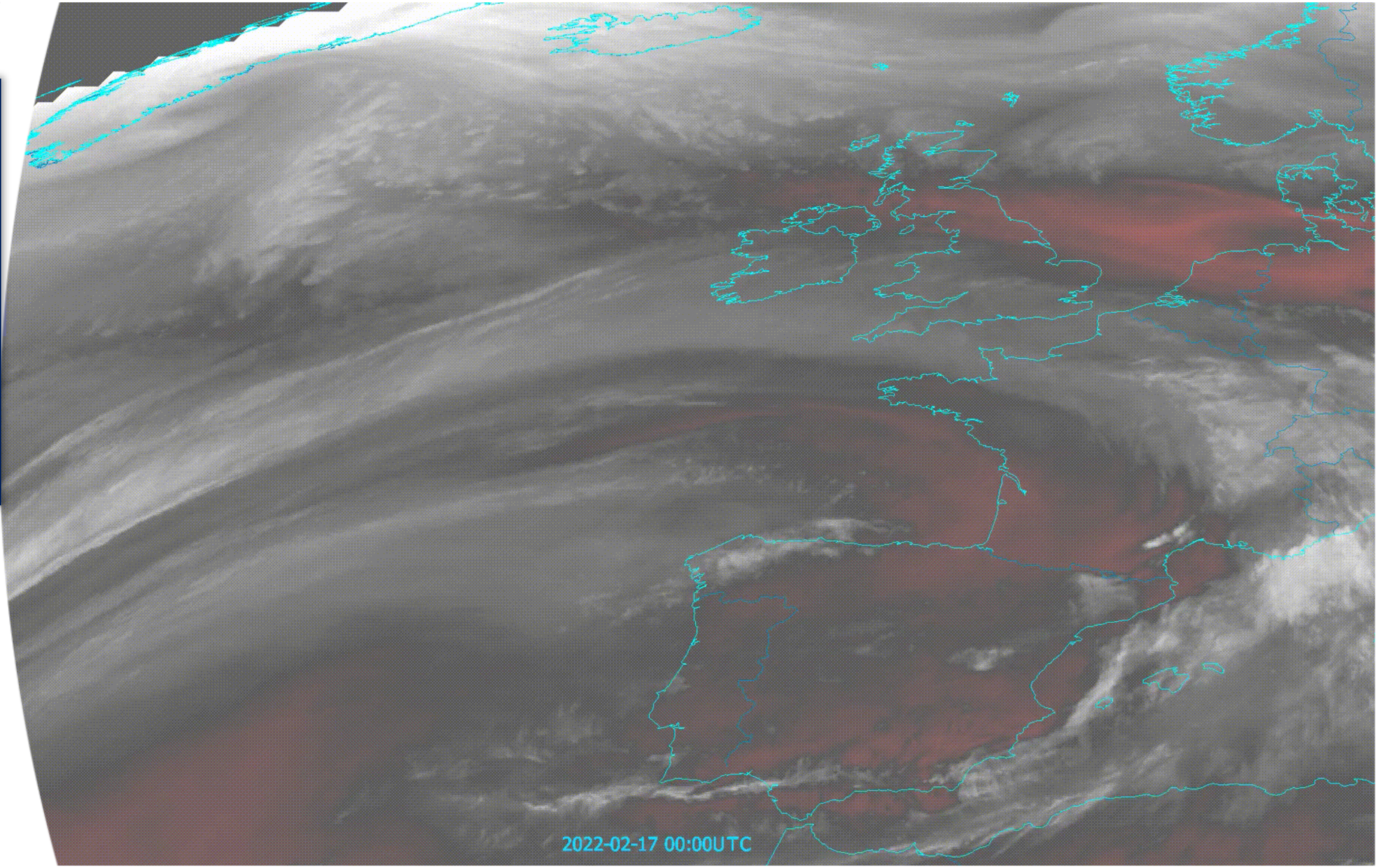
- Storm: warmer ocean temperatures = more fuel, raising likelihood of Cat. 4–5 hurricanes.
- Wind: Frequency of rapid intensification events incr. over time – linked to higher ocean heat content.
- Rain: Warmer air holds more moisture, leading to 10–20% higher rainfall in recent hurricanes.
- Surge: Adding to storm impact, sea-level rise also substantially increases storm-surge flooding.

National Hurricane Center Storm Tracks.
Saffir–Simpson hurricane wind scale.
Nilfanion (05.08.2006)

Storm Eunice over Europe in February 2022

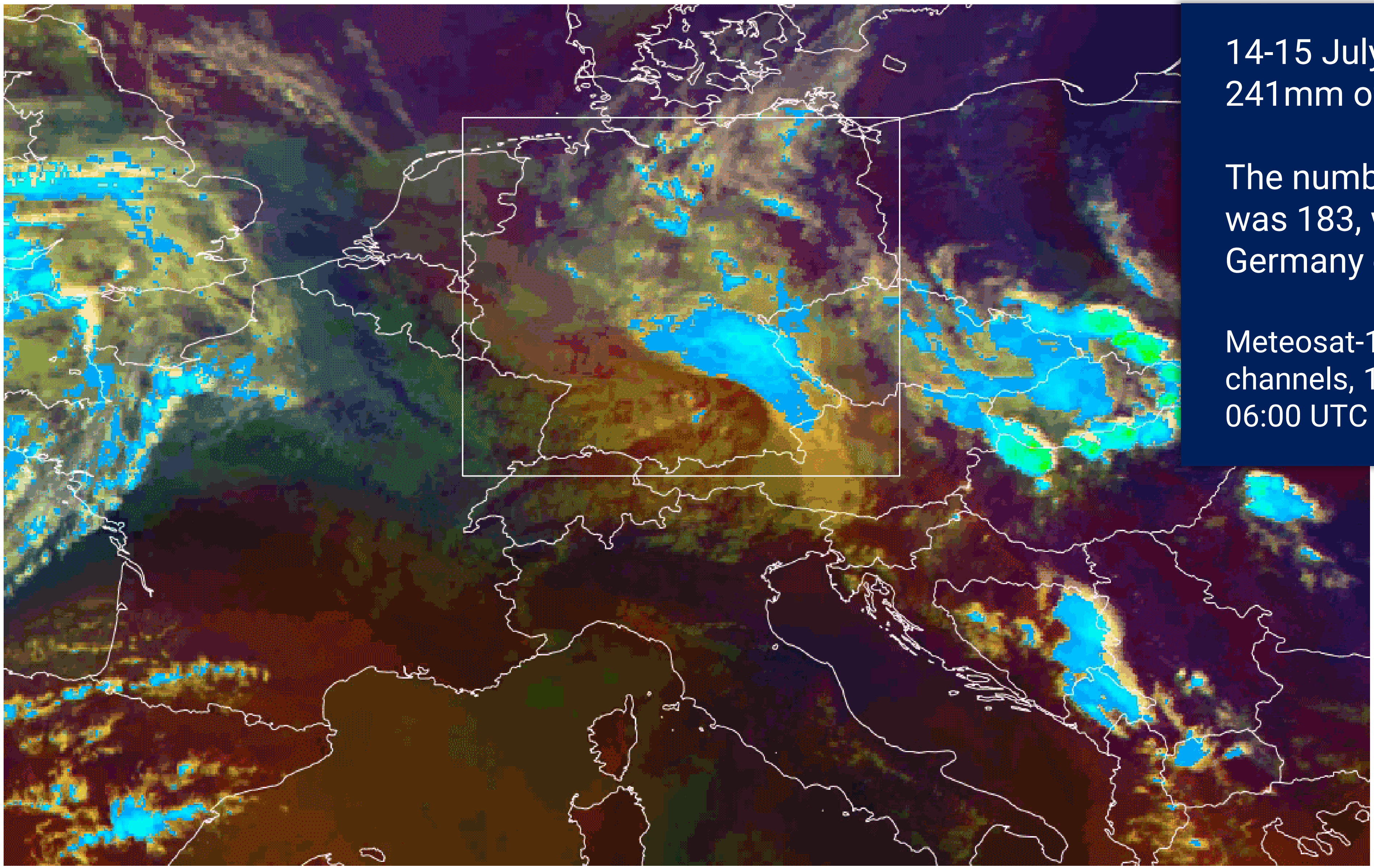
Very severe Atlantic storm, Storm Eunice, hit parts of Western Europe 17-18 February 2022.

Viewed by Meteosat-11 water vapour channel, 17 February 00:00 – 18 February 10:00 UTC



Devastating floods in Western Europe in July 2021

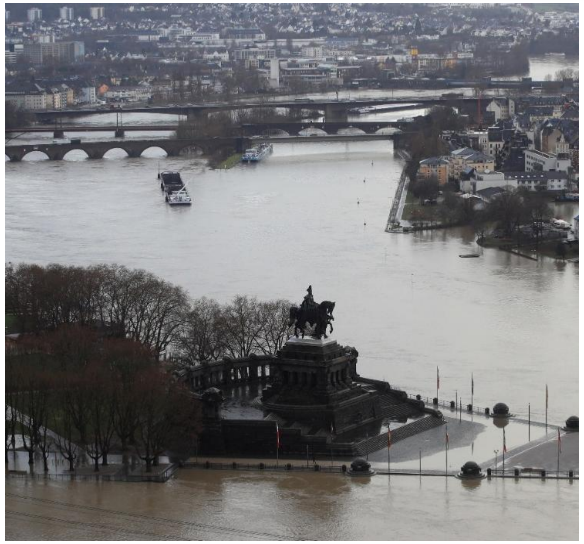
www.eumetsat.int



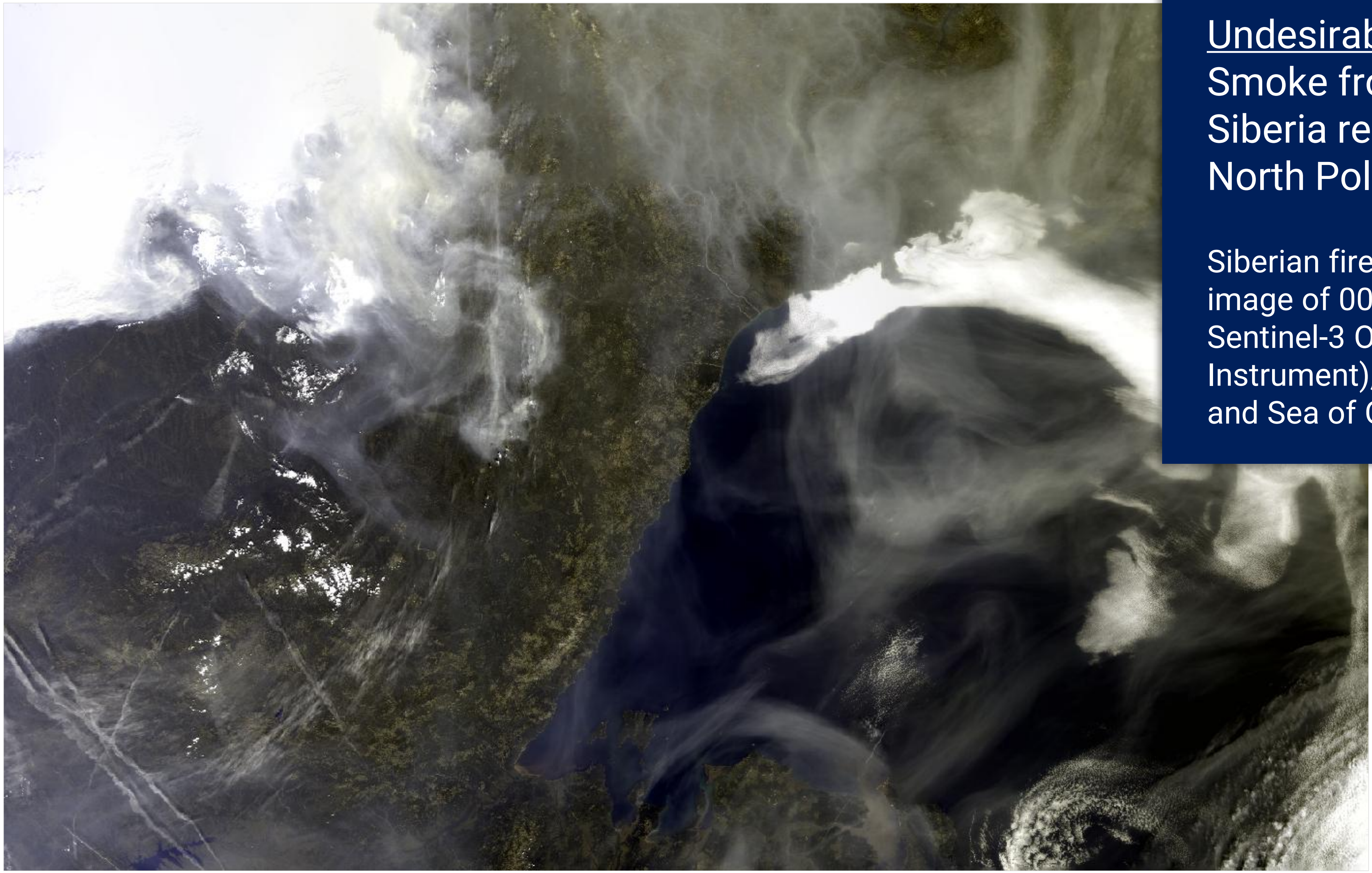
14-15 July 2021 Germany received 241mm of rainfall in 22 hours.

The number of deaths reported was 183, with economic losses in Germany exceeding US$20 billion.

Meteosat-11 precipitation rate from IR channels, 11 July 12:00 UTC – 16 July 06:00 UTC (Credit: H SAF)

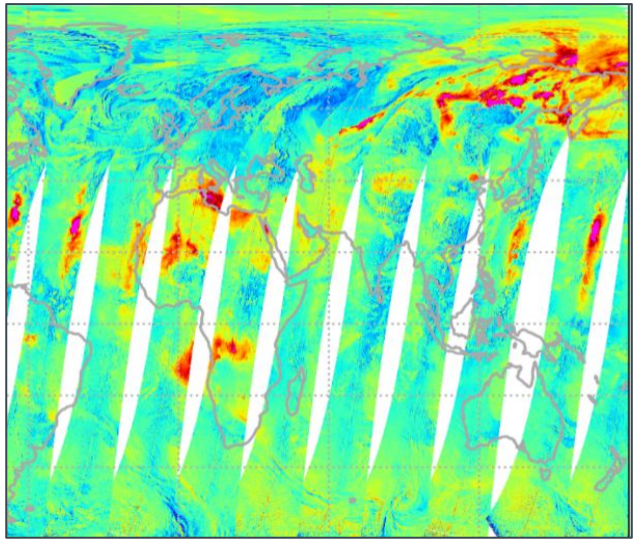


Unprecedented wildfires in Siberia spread smoke over the Arctic

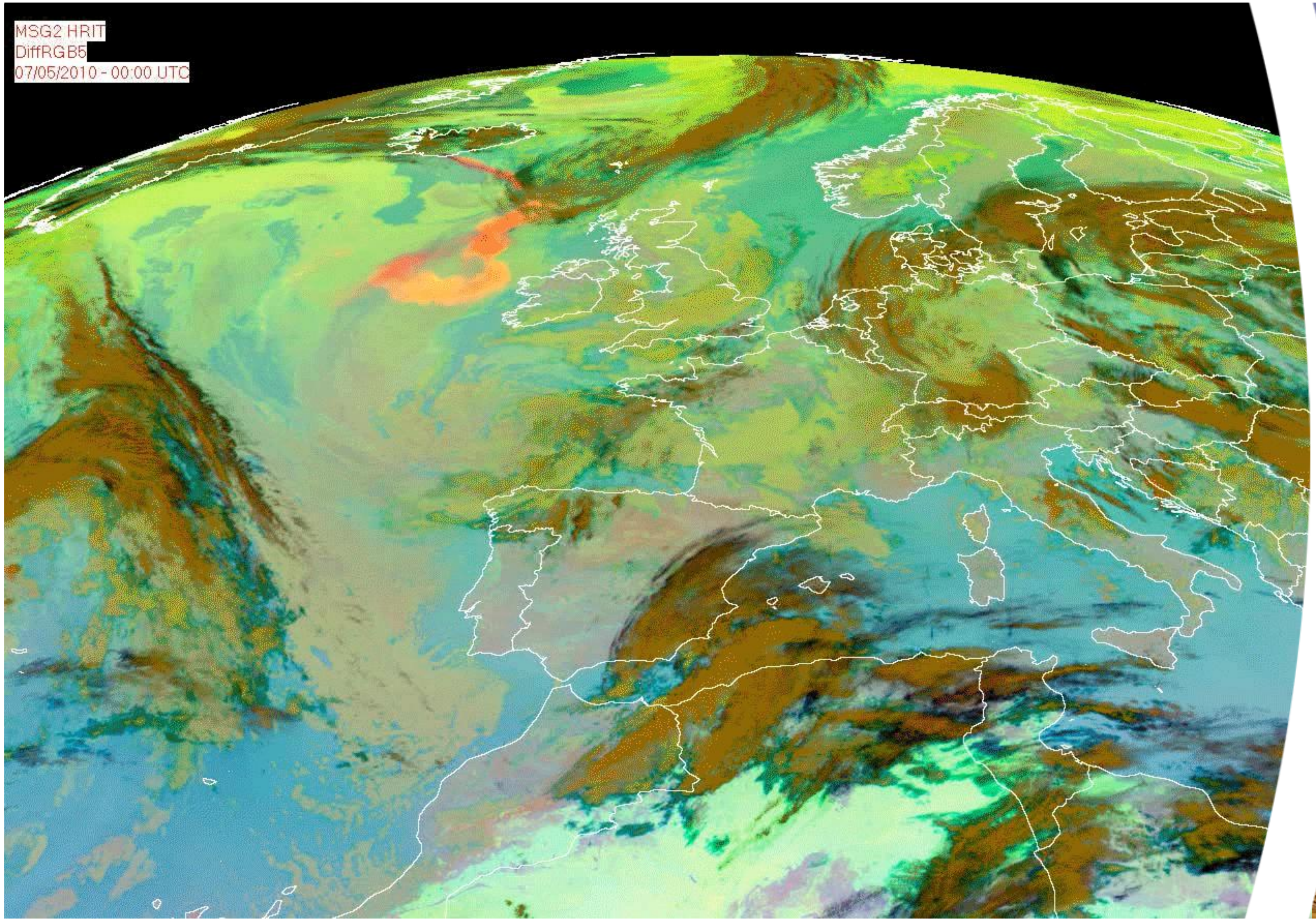


<u>Undesirable world first</u>:
Smoke from wildfires in part of Siberia reached the geographic North Pole in August 2021.

Siberian fires, 12 August 2021, Composite image of 00:30 and 01:32 UTC Copernicus Sentinel-3 OLCI (Ocean and Land Colour Instrument), showing the Dzhigda region and Sea of Okhotsk.



Meteosat for transport safety in aviation



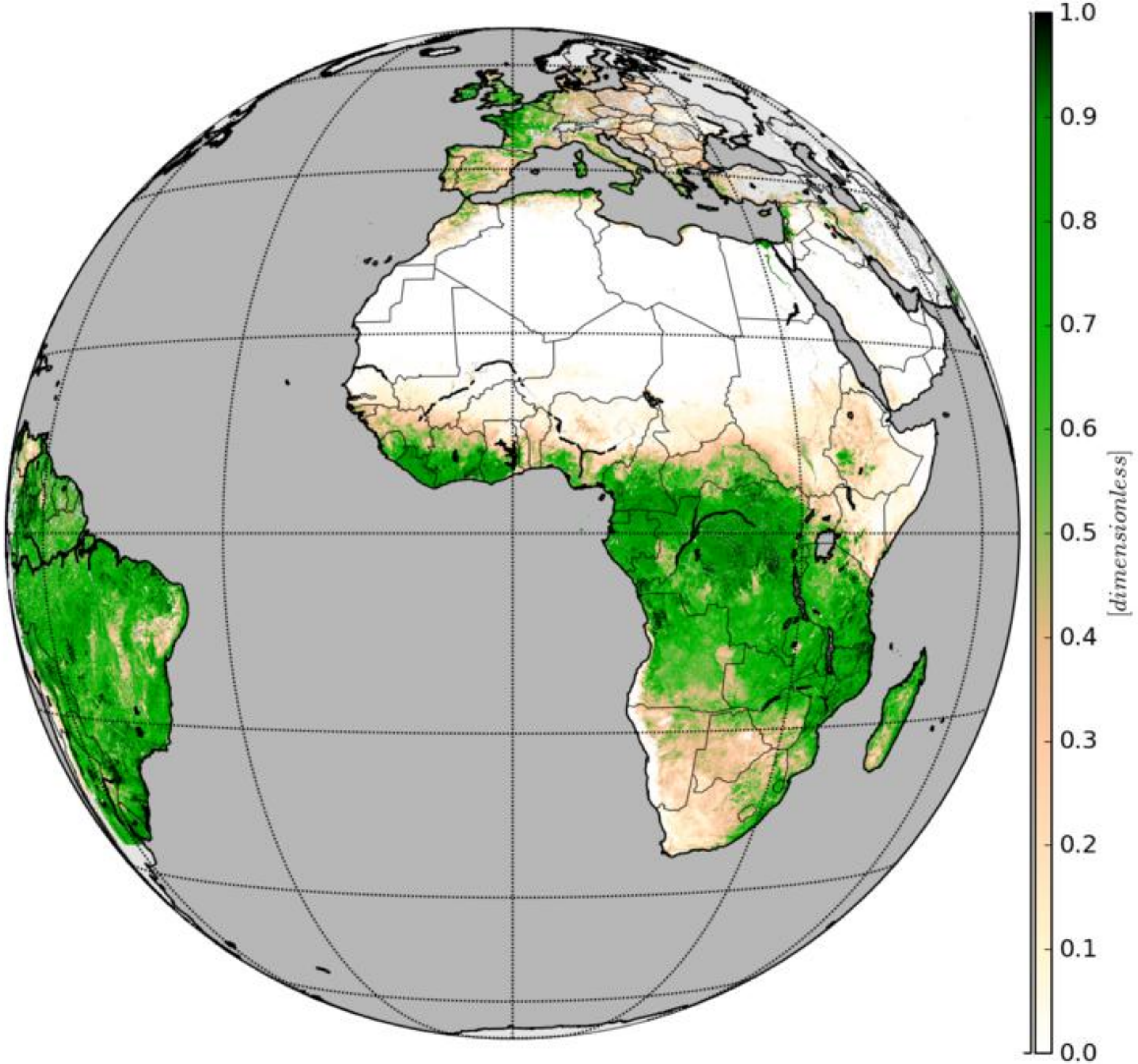
Iceland 7-11 May 2010

Volcanic ash cloud (orange) from the second eruption of Eyjafjallajökull tracked by MSG geostationary satellites.

In western Africa, tracking droughts by satellite saves lives

In combination with in-situ measures, Meteosat-11 data is supporting local organisations to monitor rainfall. Most crises in the Sahel region are linked to drought.

The Land Surface Analysis Satellite Application Facility's **fractional vegetation** product can be used to track the progress of growing seasons.



Where we are today: MSG

Meteosat 2nd Generation (MSG)
'Next Gen' EUMETSAT geostationary satellites

The main instruments carried by the MSG satellites include:

- Geostationary Earth Radiation Budget **GERB**
- Spinning Enhanced Visible and Infrared Imager **SEVIRI**.

GERB observes the reflected sunlight and thermal emissions from the Earth in 2 broadband channels.

Where we are today: MSG

Meteosat 2nd Generation (MSG)
'Next Gen' EUMETSAT geostationary satellites

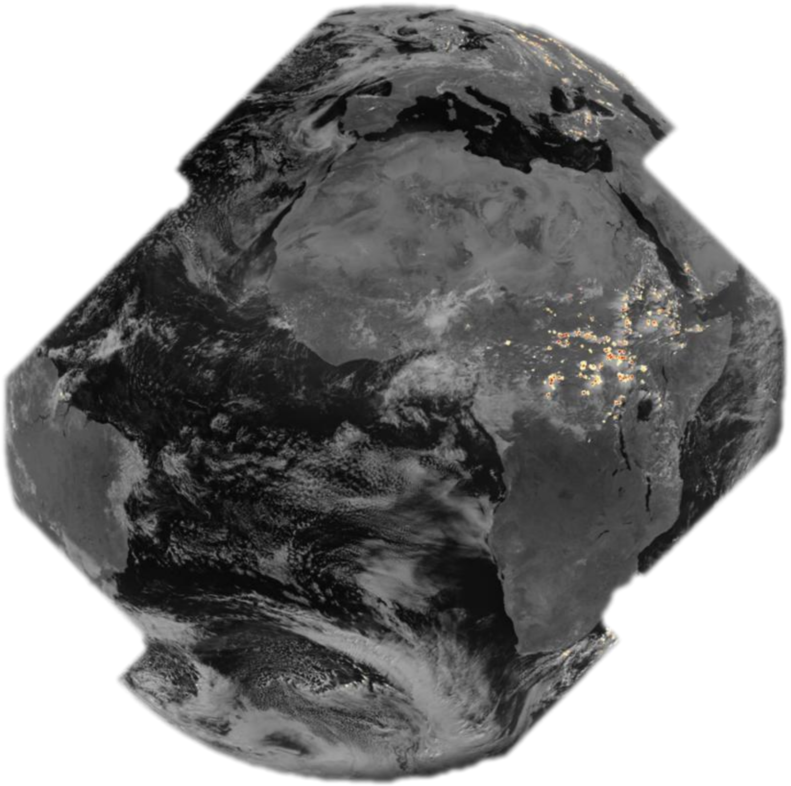
SEVIRI has 12 spectral channels:

- Thermal IR to measure cloud, land, sea surface temperature.
- Channels for ozone, water vapour, and CO_2 to analyse air masses and reconstruct 3D a view of the atmosphere.
- Visible channel at 1 km horizontal resolution to help weather forecasters detect, track, predict severe weather.

Where we are today: MTG-I

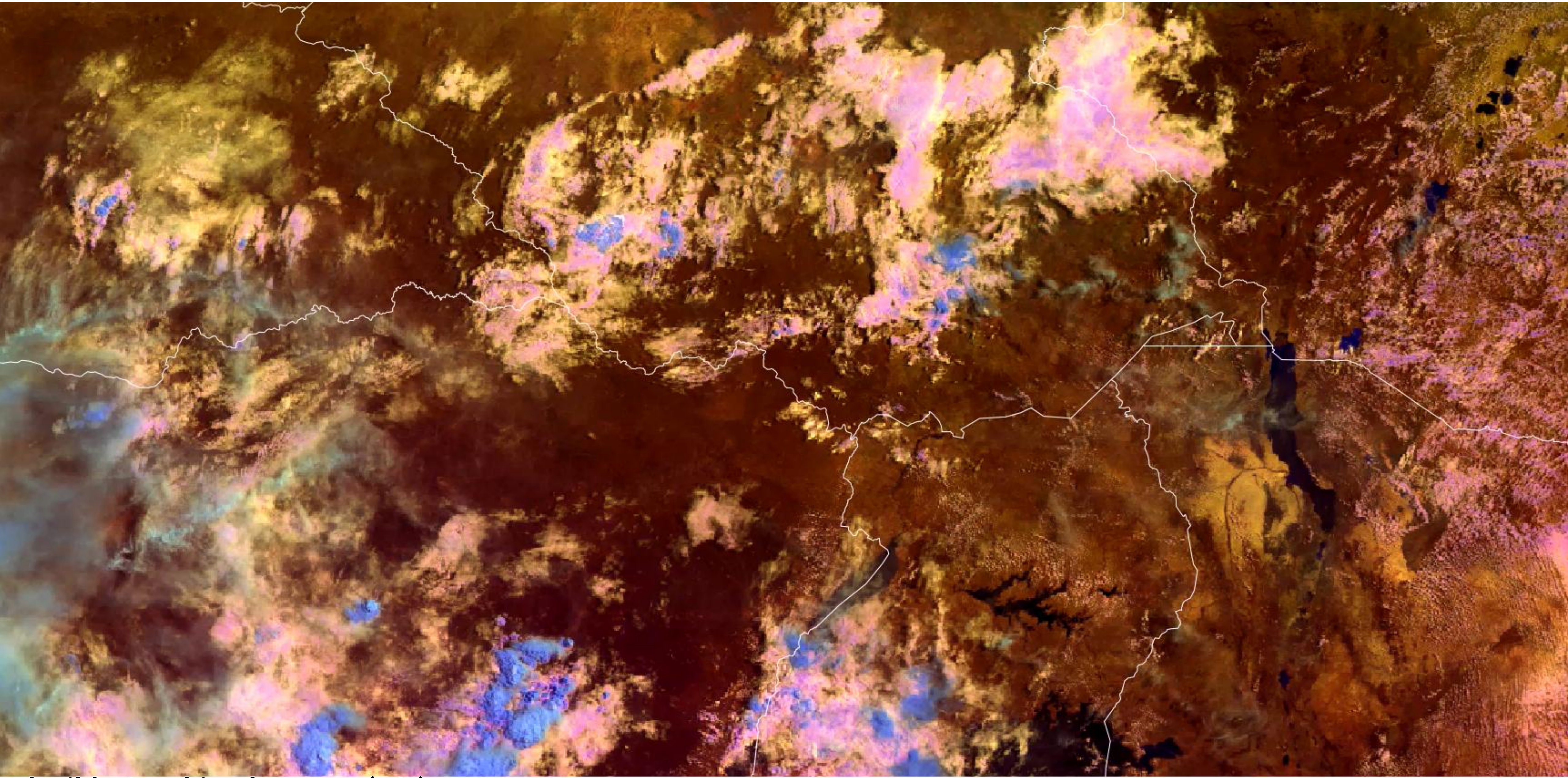
Meteosat Third Generation (MTG)
'Next Gen' EUMETSAT geostationary satellites

FCI bands	
VIS 0.4	1.0 km
VIS 0.5	1.0 km
VIS 0.6	**1.0 km** **0.5 km (HR)**
VIS 0.8	1.0 km
VIS 0.9	1.0 km
NIR 1.3	1.0 km
NIR 1.6	1.0 km
NIR 2.2	1.0 km 0.5 km (HR)
IR 3.8	2.0 km 1.0 km (HR)
WV 6.3	2.0 km
WV 7.3	2.0 km
IR 8.7	2.0 km
IR 9.7	2.0 km
IR 10.5	2.0 km 1.0 km (HR)
IR 12.3	2.0 km
IR 13.3	2.0 km



Lightning Imager

Where we are today: FCI Vis 0.6



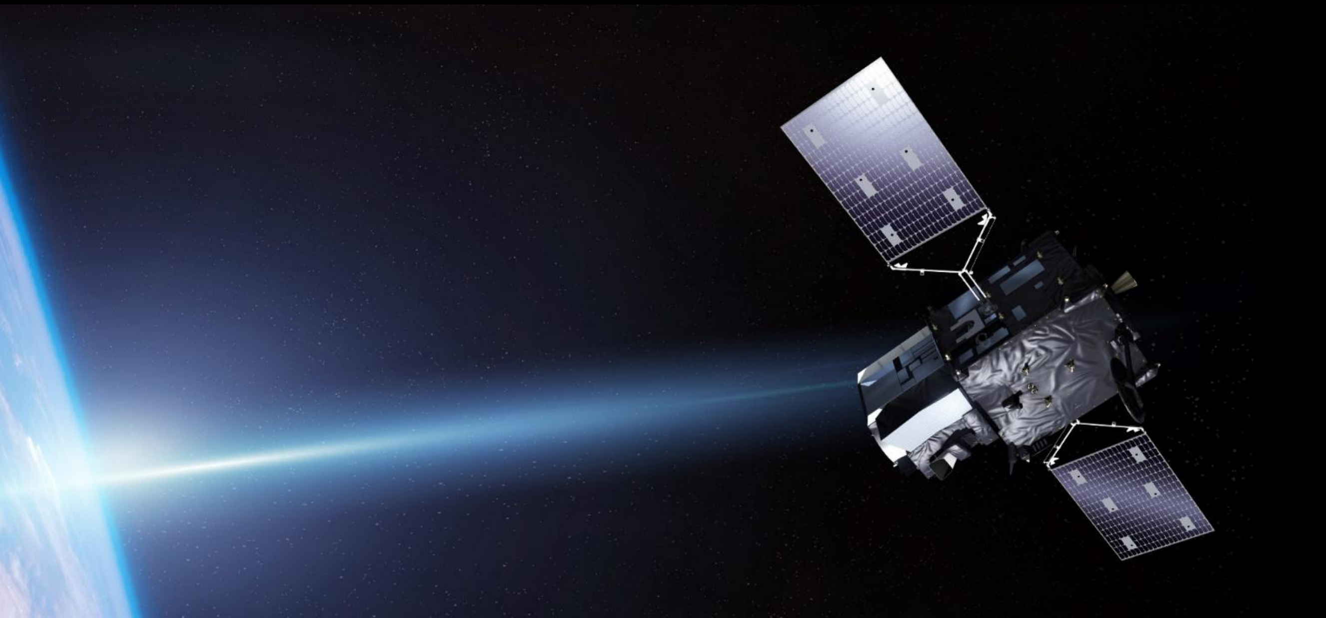
Flexible Combined Imager (FCI)

2025-05-02 08:00:00 UTC

Where we are today: LI

Where we are today: MTG-S

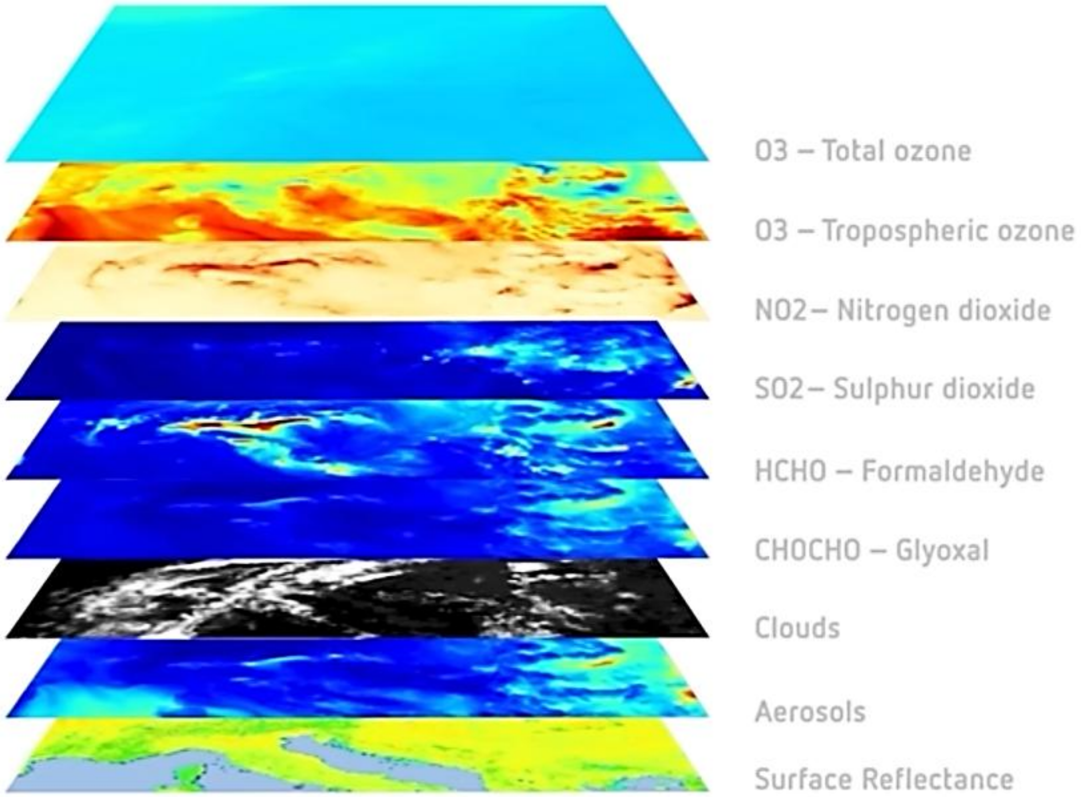
- Carries two key missions: Infrared Sounder and Copernicus Sentinel-4 Ultraviolet Visible Near-infrared spectrometer.
- **Infrared Sounder** will scan nearly 2,000 thermal IR wavelengths every 30 minutes to build vertical profiles of temperature, humidity, trace gases.
- Crucial for detecting fast-developing convective weather by revealing sudden shifts in instability, moisture, wind – before clouds begin to form.



Combined with data from MTG’s imager satellites, the IR Sounder will offer a continuous view of a storm’s full life cycle, enhancing nowcasts, forecasts, models, and long-term climate monitoring.

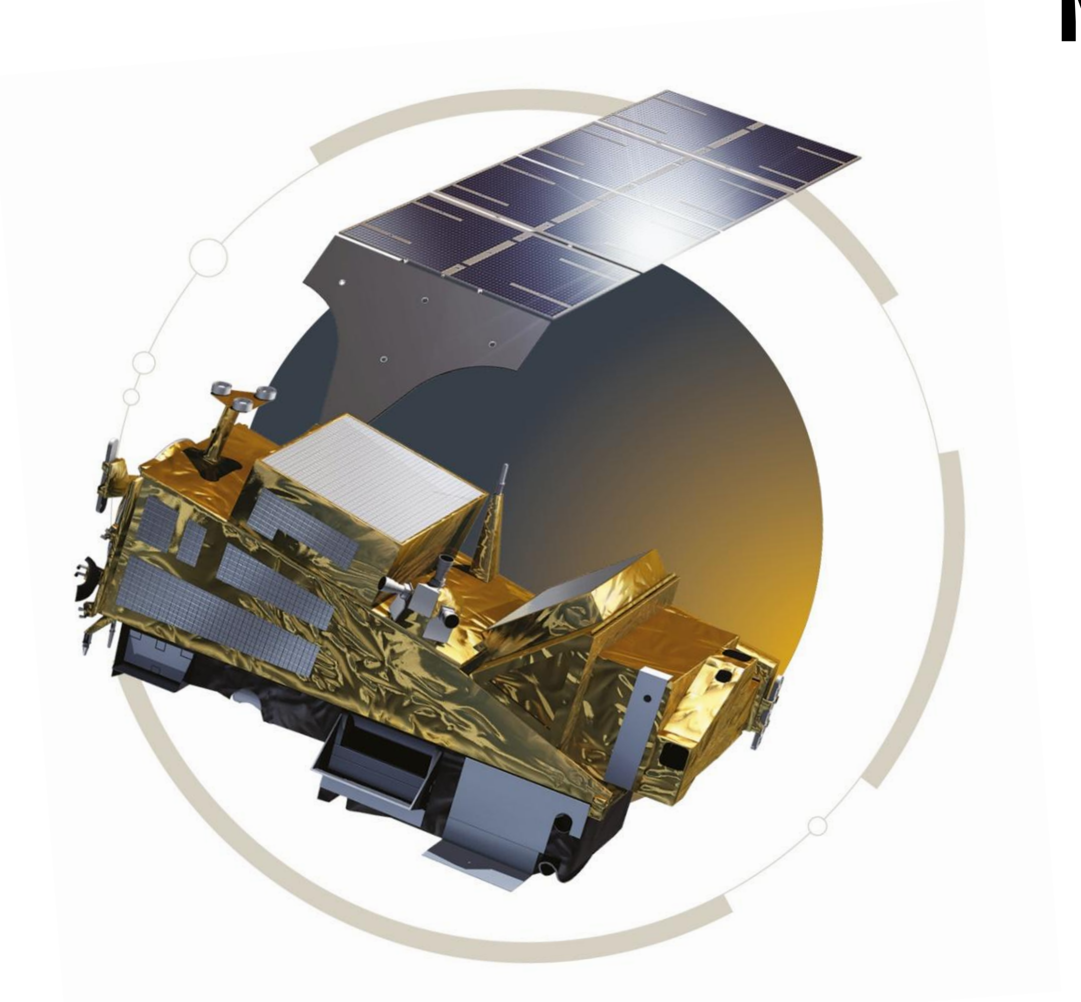
Where we are today: Sentinel-4 UVN

- Hourly data on pollutants and aerosols – including from wildfires and volcanic eruptions.
- Measurements will enhance monitoring of emissions and air quality forecasts, as well as support public health and environmental policy.



https://www.youtube.com/watch?v=sMW62W-0G78

Where we are today:
EUMETSAT Polar System Second Generation (EPS-SG)
Metop-SG



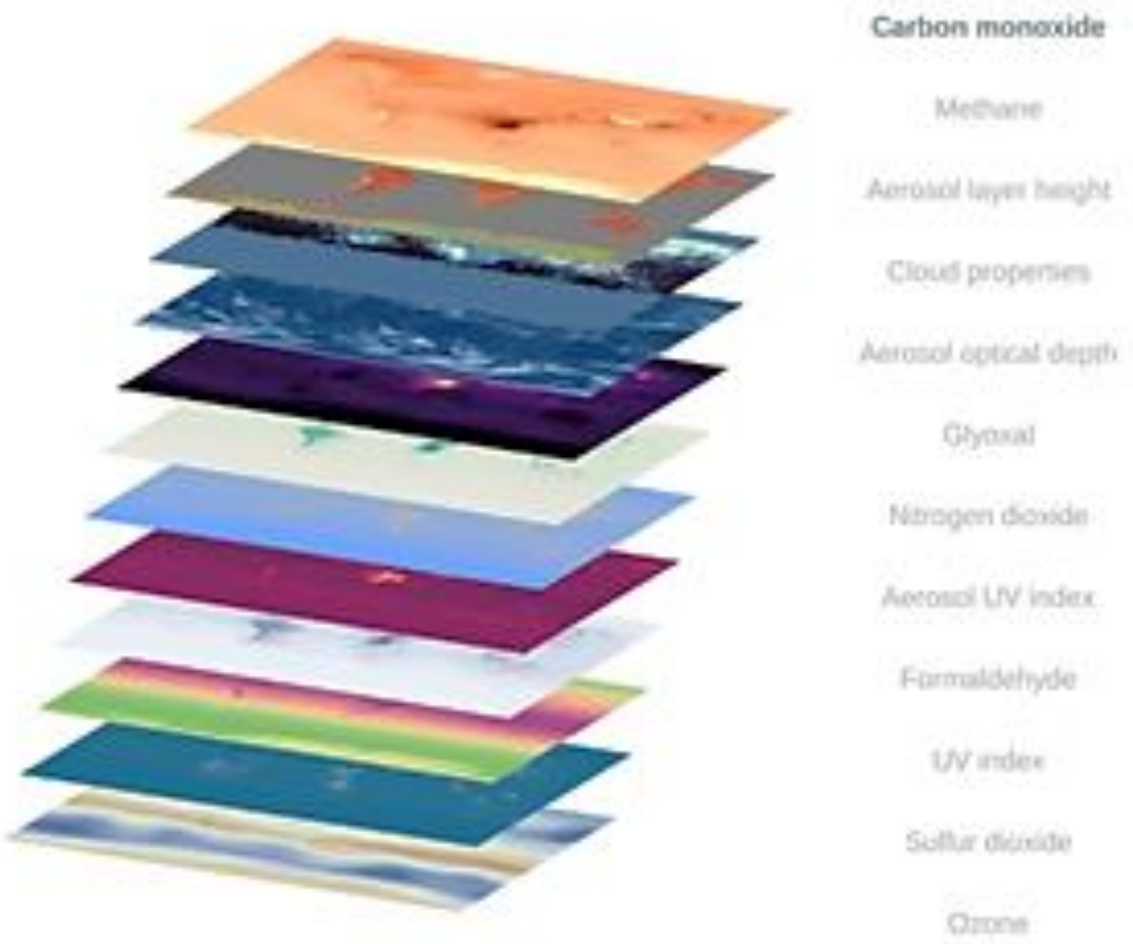
Metop Second Generation (Metop-SG)
'Next Gen' EUMETSAT polar-orbiting satellites

Data to improve global weather monitoring and predictions:

- Enhanced IR, MW, and radio-occultation soundings of temperature and humidity.
- Polar atmospheric motion vectors extracted from optical imagery.
- Novel precipitation and cloud measurements in the optical, sub mm and MW spectra.
- High-resolution ocean surface wind-vector and soil moisture measurements extracted from scatterometer observations.

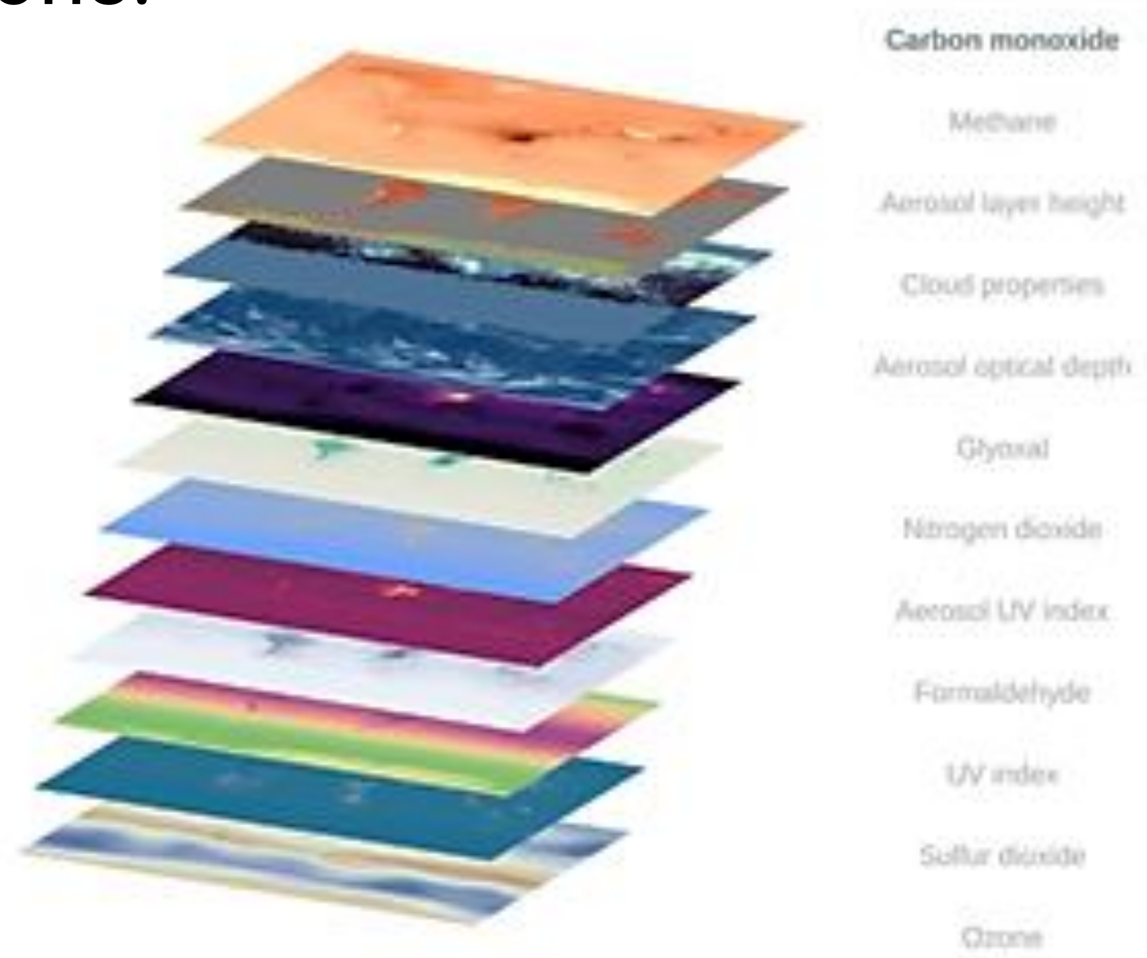
Where we are today: Sentinel-5 UVNS

- Daily snapshot of pollutants and aerosols – including wildfires and volcanic eruptions.
- Measurements will enhance monitoring of emissions and air quality forecasts, as well as support public health and environmental policy.



Where we are today: Sentinel-5 UVNS

- Daily snapshot of pollutants and aerosols – including wildfires and volcanic eruptions.
- Measurements will enhance monitoring of emissions and air quality forecasts, as well as support public health and environmental policy.



EPS-Sterna secures support of all member states

E-AI Data Curation WG "Gap Analysis"

Lauren Biermann, Roope Tervo and EUMETNET community contributors

E-AI Working Group on Data Curation: “Gap Analysis”

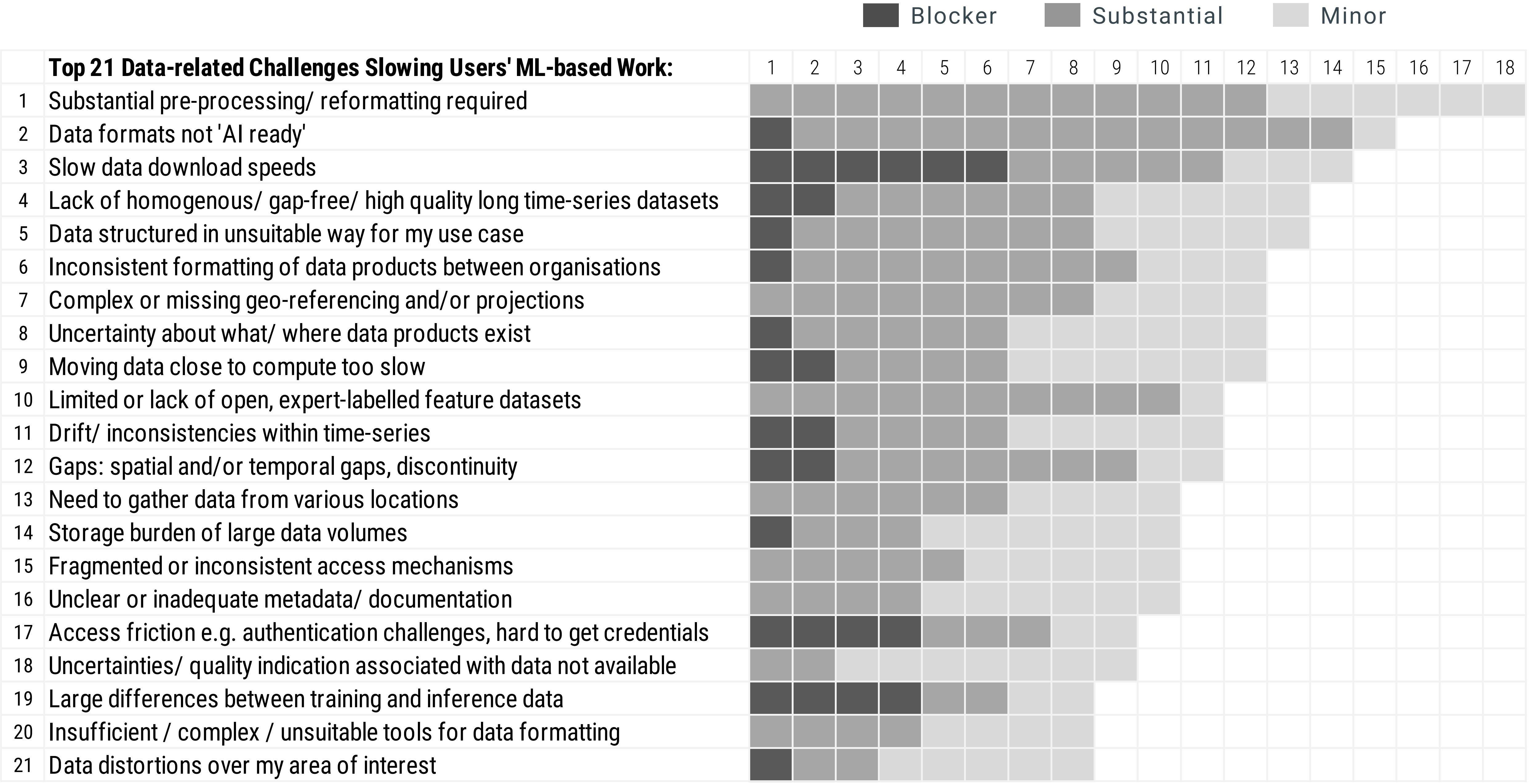
- GAP ANALYSIS: structured questionnaire to capture information on **data challenges that were preventing or slowing domain experts** from seamless ML-driven workflows.
- Via the E-AI mailing list, a request was sent to **400+** members to participate in a 1-hour interview, during which time I would ask structured questions and complete the questionnaire “in real time”.
- Six months later, I have responses from **18 domain experts** across nowcasting and forecasting (n=10), products and services (n=5) and climate (n=3)… a response rate that reflects the universal scarcity of expert time.

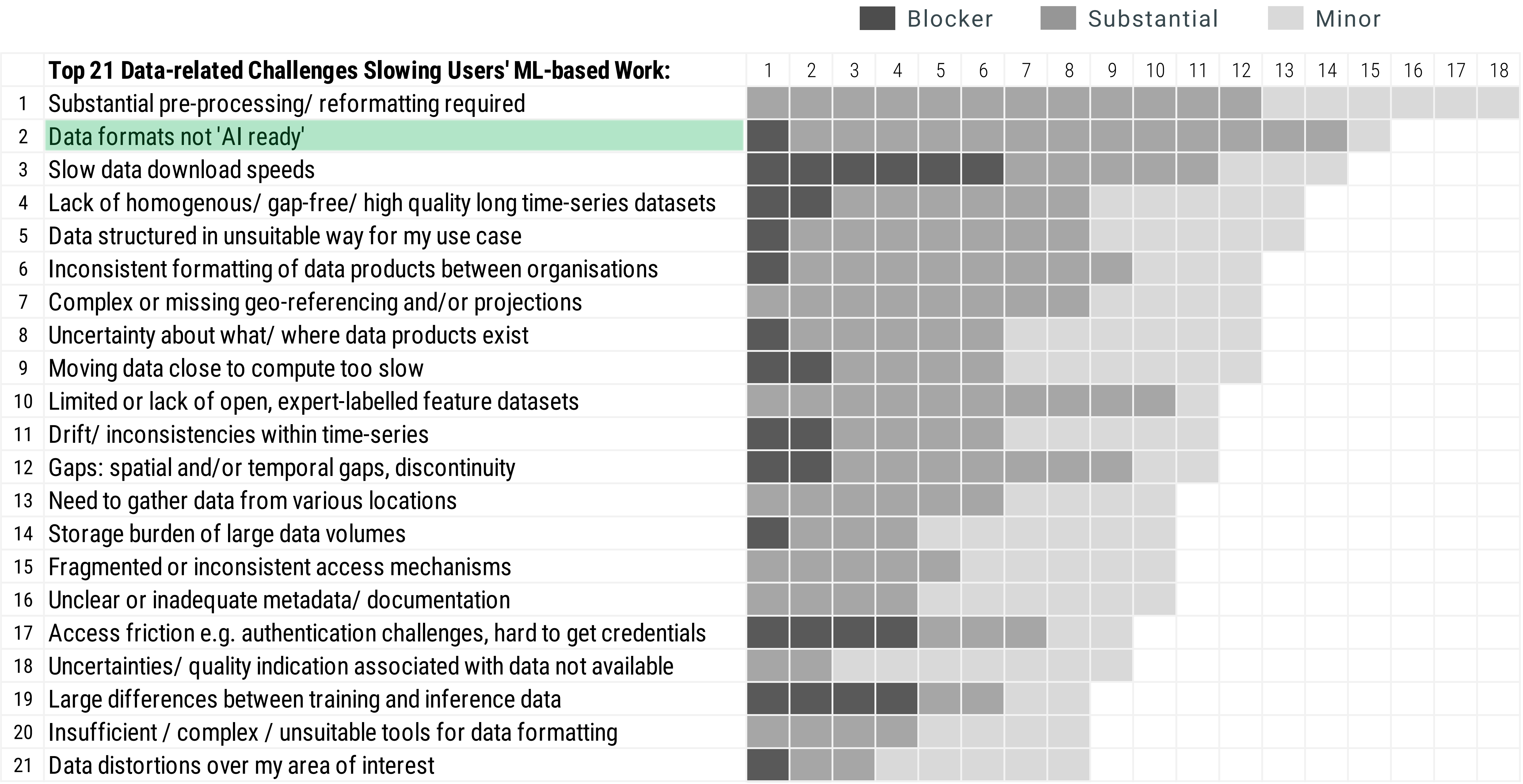
E-AI Working Group on Data Curation: “Gap Analysis”

- GAP ANALYSIS: structured questionnaire to capture information on **data challenges that were preventing or slowing domain experts** from seamless ML-driven workflows.
- Via the E-AI mailing list, a request was sent to 400+ members to participate in a 1-hour interview, during which time I would ask structured questions and complete the questionnaire “in real time”.
- Six months later, I have responses from **18 domain experts** across nowcasting and forecasting (n=10), products and services (n=5) and climate (n=3)… a response rate that reflects the universal scarcity of expert time.

Contributors to Gap Analysis:

- Belgium
- Czechia
- Denmark
- Finland
- France
- Germany
- Ireland
- Portugal
- Spain
- Switzerland
- United Kingdom
- IBM, Innoflair


Blocker
Substantial
Minor
Top 21 Data-related Challenges Slowing Users' ML-based Work:
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18
1 Substantial pre-processing/ reformatting required
2 Data formats not 'AI ready'
3 Slow data download speeds
4 Lack of homogenous/ gap-free/ high quality long time-series datasets
5 Data structured in unsuitable way for my use case
6 Inconsistent formatting of data products between organisations
7 Complex or missing geo-referencing and/or projections
8 Uncertainty about what/ where data products exist
9 Moving data close to compute too slow
10 Limited or lack of open, expert-labelled feature datasets
11 Drift/ inconsistencies within time-series
12 Gaps: spatial and/or temporal gaps, discontinuity
13 Need to gather data from various locations
14 Storage burden of large data volumes
15 Fragmented or inconsistent access mechanisms
16 Unclear or inadequate metadata/ documentation
17 Access friction e.g. authentication challenges, hard to get credentials
18 Uncertainties/ quality indication associated with data not available
19 Large differences between training and inference data
20 Insufficient / complex / unsuitable tools for data formatting
21 Data distortions over my area of interest



Blocker
Substantial
Minor
Top 21 Data-related Challenges Slowing Users' ML-based Work:
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18
1 Substantial pre-processing/ reformatting required
2 Data formats not 'AI ready'
3 Slow data download speeds
4 Lack of homogenous/ gap-free/ high quality long time-series datasets
5 Data structured in unsuitable way for my use case
6 Inconsistent formatting of data products between organisations
7 Complex or missing geo-referencing and/or projections
8 Uncertainty about what/ where data products exist
9 Moving data close to compute too slow
10 Limited or lack of open, expert-labelled feature datasets
11 Drift/ inconsistencies within time-series
12 Gaps: spatial and/or temporal gaps, discontinuity
13 Need to gather data from various locations
14 Storage burden of large data volumes
15 Fragmented or inconsistent access mechanisms
16 Unclear or inadequate metadata/ documentation
17 Access friction e.g. authentication challenges, hard to get credentials
18 Uncertainties/ quality indication associated with data not available
19 Large differences between training and inference data
20 Insufficient / complex / unsuitable tools for data formatting
21 Data distortions over my area of interest

Destination Earth Framework for Preparing AI-Ready Data (DEFAIR)

Motivation for DEFAIR

- **Preparing EO data for AI/ML** is often complex and time-consuming.
- Datasets differ in **format**, **structure**, **projection**, and **metadata** conventions.
- Multi-source analysis requires **harmonisation** and **data collocation**.
- Large-scale datasets require **scalable processing** methods that can fit into memory.
- Data preparation **workflows** are often difficult to reuse, **integrate** into processing chains, and deploy at scale across different environments.

Destination Earth data lake Framework for preparing AI-Ready Data (DEFAIR)

Key Design Features

- Dask Lazy Evaluations
- YAML / Python API
- Pipeline Definition
- Plugin Architecture
- Modular

Example pipeline

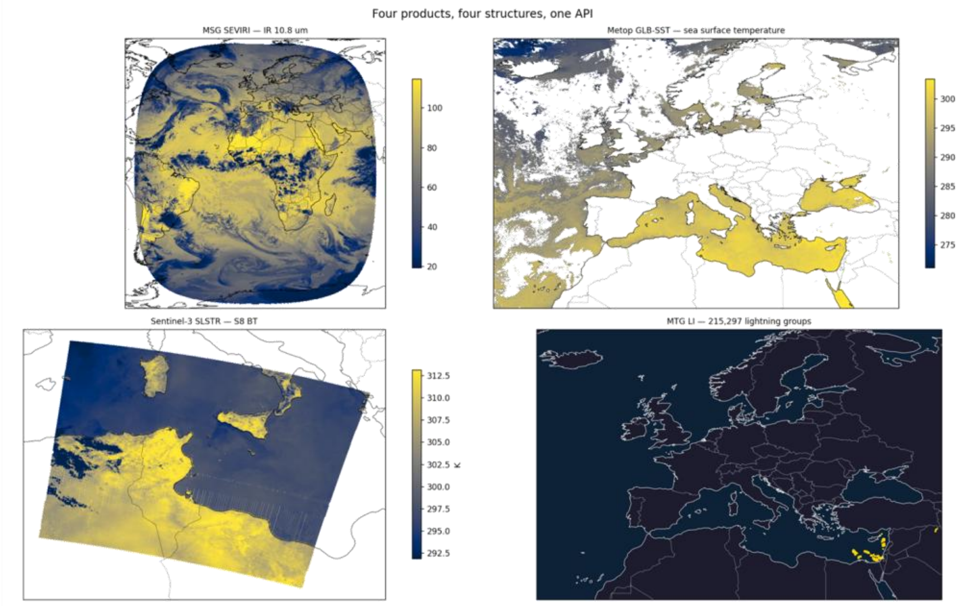
```python
pipeline = (
        Dataset.from_source(input_mtg, auto_discover_chunks=True)
        .transform("content_filter", include_vars=["ir_3.8", "wv_7.3"]),
        .transform("spatial_filter", lat_min=30.0, lat_max=60.0, lon_min=-10.0,
         lon_max=30.0, drop=True)
        .transform("reprojection", target="EPSG:4326", resolution=0.1,
        resampling="bilinear", drop=True)
)
```

Key Components

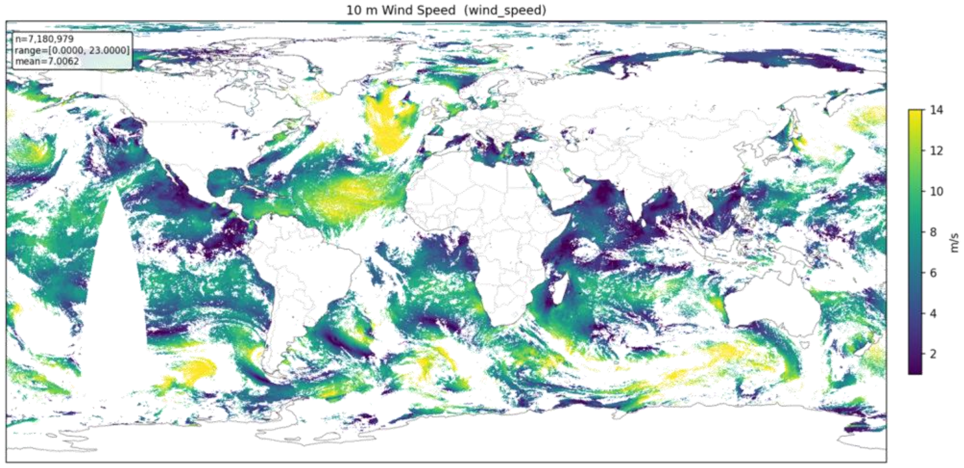
DEFAIR provides a unified API for reading EUMETSAT satellite data from DestinE Data Portfolio, from different data origins and applying operations to generate AI-Ready data cubes, which can be exported to variety of output formats.

Data Sources	Data Products & Formats	Operations	Output Formats
❑ Local Network File Systems ❑ S3 Object Storage ❑ DEDL Harmonised Data Access API (STAC-compliant)	❑ Operational feed & reprocessed data ❑ MSG SEVIRI ❑ MTG FCI & LI ❑ METOP ❑ Sentinel-3	❑ Content filtering ❑ Temporal filtering ❑ Spatial filtering ❑ Reprojection (WGS84, HEALPix, Lambert, …) ❑ Collocation & Alignment (under testing)	❑ NetCDF ❑ Zarr V2 & V3 ❑ GeoTIFF ❑ PNG ❑ JPEG ❑ …

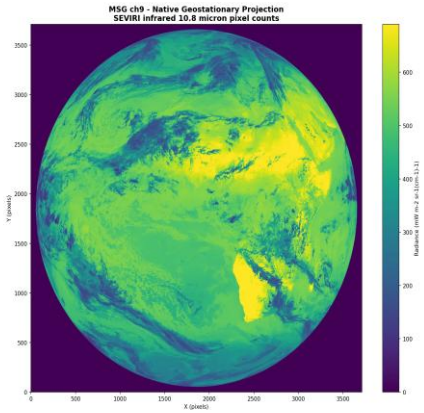
DEFAIR Processing Tool - Capabilities



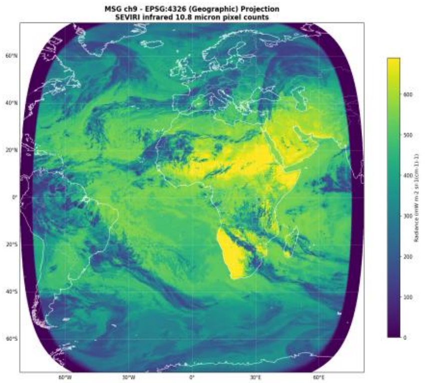
50+ Readers



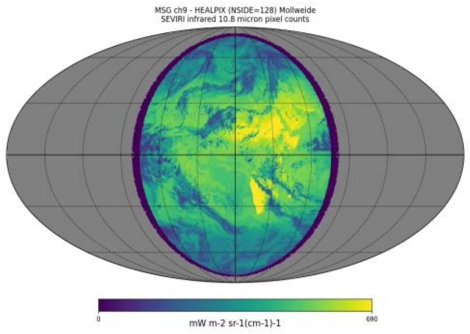
Reprojections



Native Geostationary

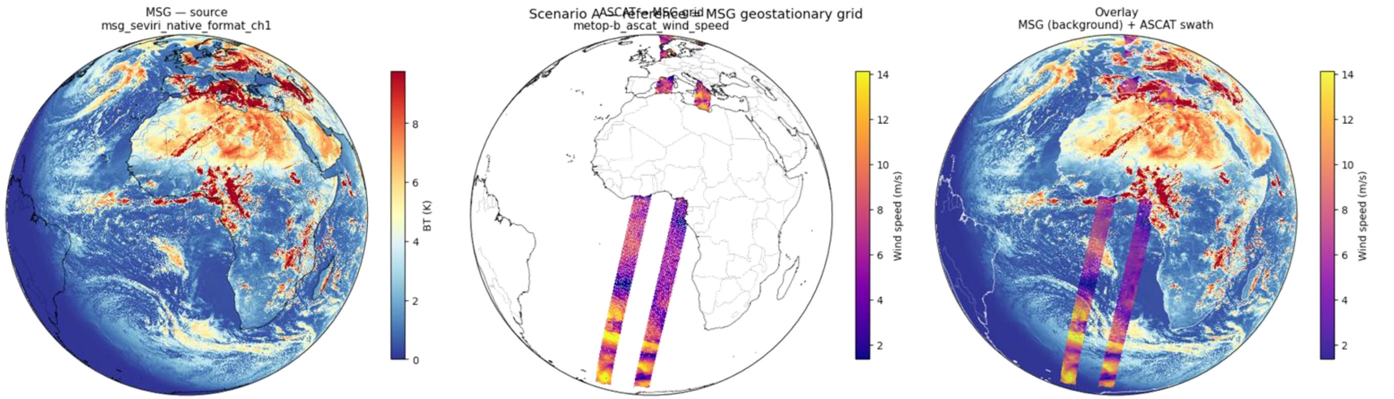


EPSG:4326 Geographic

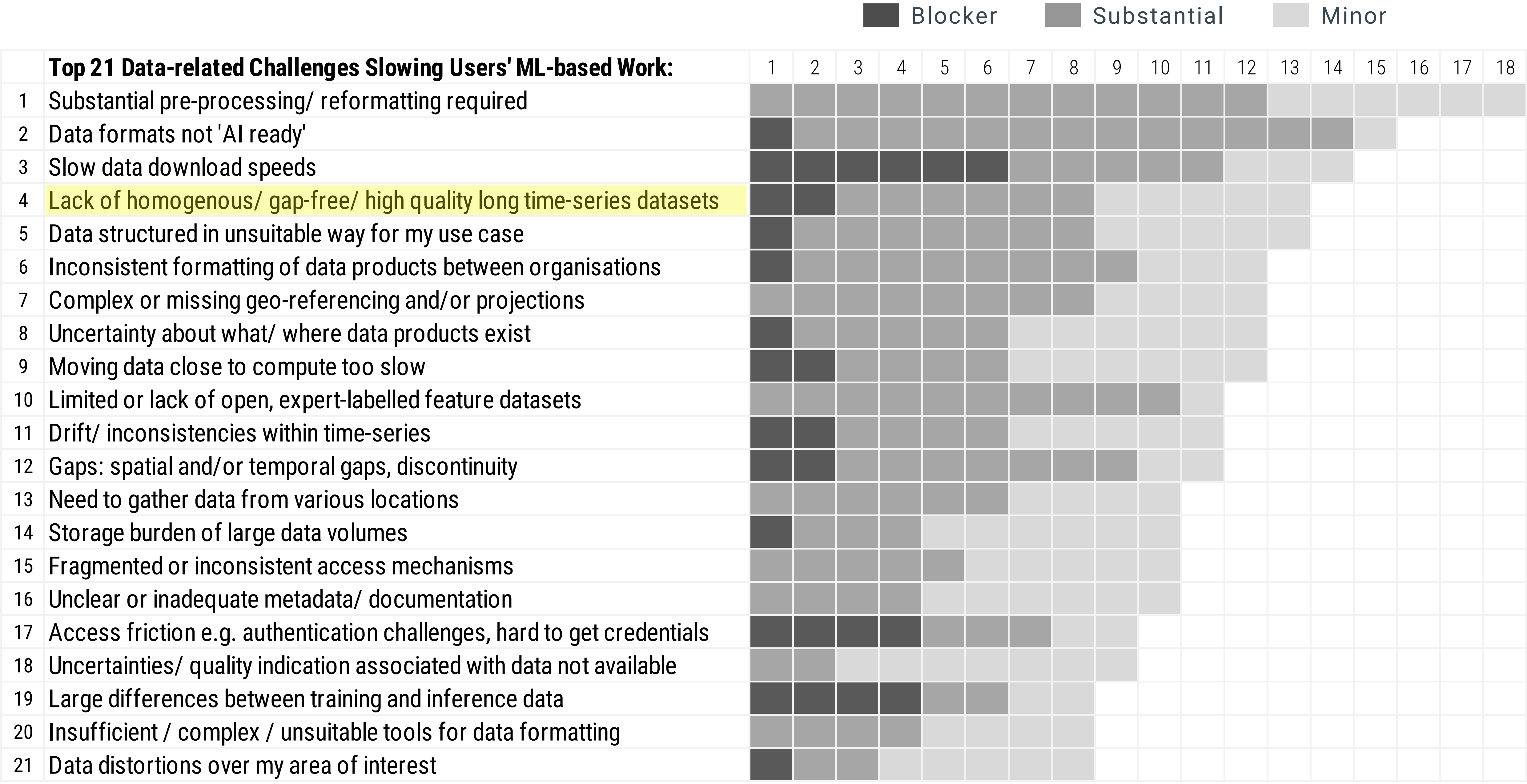


HEALPix (Equal Area)

Collocation Plugin



Funded by the European Union Destination Earth



Blocker
Substantial
Minor
Top 21 Data-related Challenges Slowing Users' ML-based Work:
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18
1 Substantial pre-processing/ reformatting required
2 Data formats not 'AI ready'
3 Slow data download speeds
4 Lack of homogenous/ gap-free/ high quality long time-series datasets
5 Data structured in unsuitable way for my use case
6 Inconsistent formatting of data products between organisations
7 Complex or missing geo-referencing and/or projections
8 Uncertainty about what/ where data products exist
9 Moving data close to compute too slow
10 Limited or lack of open, expert-labelled feature datasets
11 Drift/ inconsistencies within time-series
12 Gaps: spatial and/or temporal gaps, discontinuity
13 Need to gather data from various locations
14 Storage burden of large data volumes
15 Fragmented or inconsistent access mechanisms
16 Unclear or inadequate metadata/ documentation
17 Access friction e.g. authentication challenges, hard to get credentials
18 Uncertainties/ quality indication associated with data not available
19 Large differences between training and inference data
20 Insufficient / complex / unsuitable tools for data formatting
21 Data distortions over my area of interest

Extra·ordinary datasets: CDRs

- EUMETSAT and our SAFs generate Climate Data Records from past and current satellite data.
- State-of-art data processors ensure the data coming from different sources form a coherent series.
- Equivalent to Level-3 data.

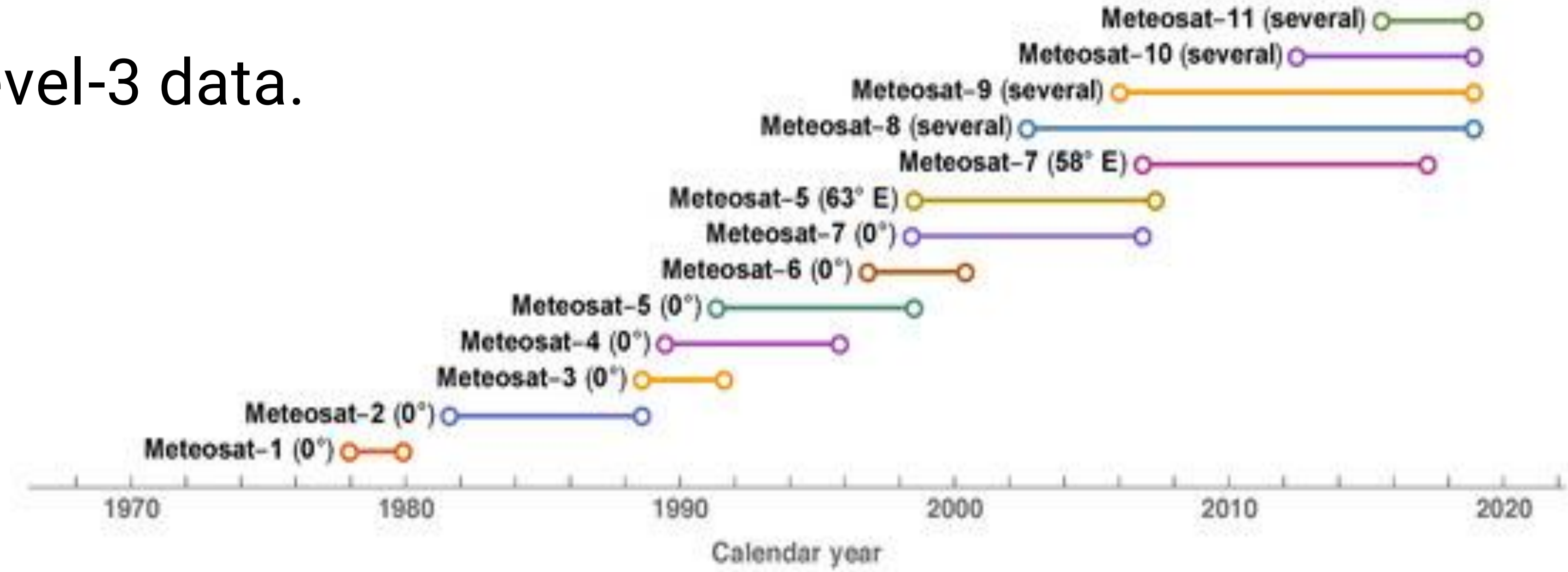


Fig. from Quast et al. (2019). Climate Data Records from Meteosat First Generation Part II: **Retrieval of the In-Flight Visible Spectral Response.**

Extra·ordinary datasets: CDRs

LEO	IASI	Ozone (O3)	Metop-A and B	2007 - 2023
LEO	IASI	All Sky Temperature and Humidity Profiles	Metop-A and B	2007 - 2023
LEO	IASI	Sulphur Dioxide (SO2)	Metop-A and B	2007 - 2023
LEO	IASI	Carbon Monoxide FORLI-CO	Metop-A and B	2007 - 2023
LEO	AVHRR	CLARA-A3 cLoud, Albedo, surface RAdiation	Metop + NOAA	1979 - 2020
LEO	SSMIS	Global Sea Ice Concentration	Metop + NOAA	2023 - 2026
GEO	SEVIRI	Cloud Properties	MSG	2004 - 2019
GEO	GEOring	GIRAFE Global Interpolated RAinFall	Multi	2002 - 2022
GEO	MVIRI + SEVIRI	SARAH Surface Radiation DataSet Heliosat	MSG + MFG	1983 - 2020
GEO	SEVIRI	Atmospheric Motion Vectors	MSG	2008 - 2020

EUMETSAT network of satellite application facilities (SAFs)

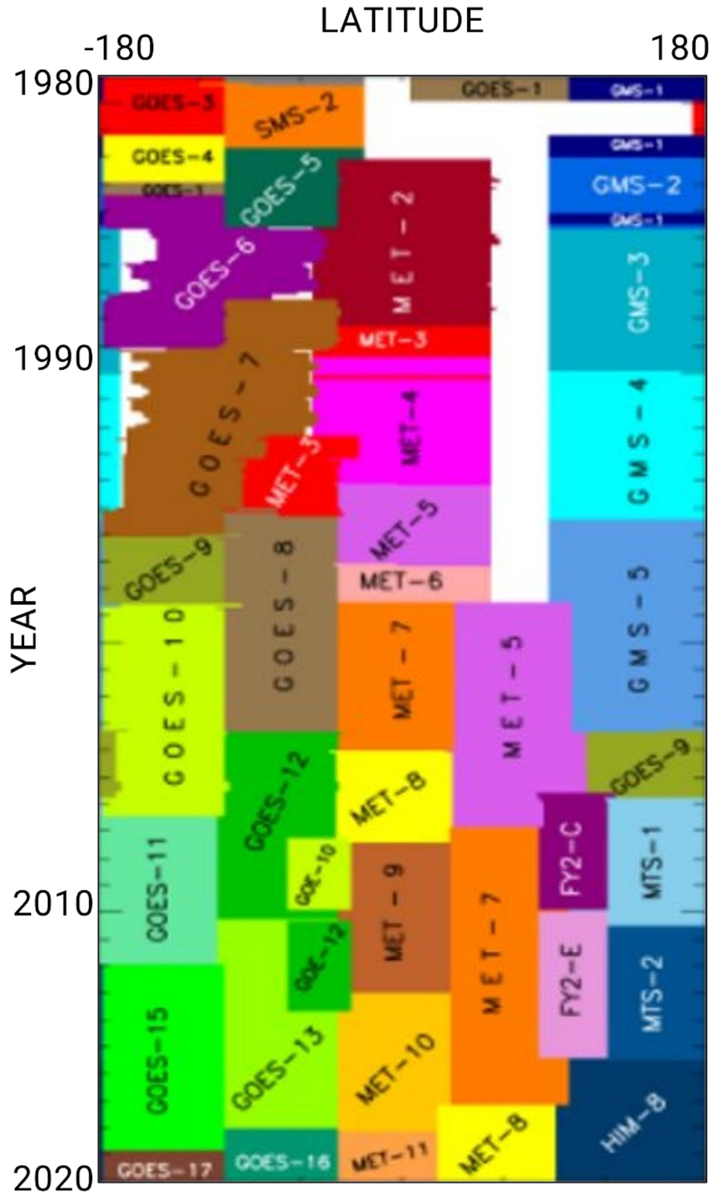
www.eumetsat.int

EUMETSAT SAFs are dedicated centres of excellence for EO data processing, making use of specialist expertise in our member states.

EUMETSAT's SAFs provide operational data and software products, each one for a dedicated user community and application area.

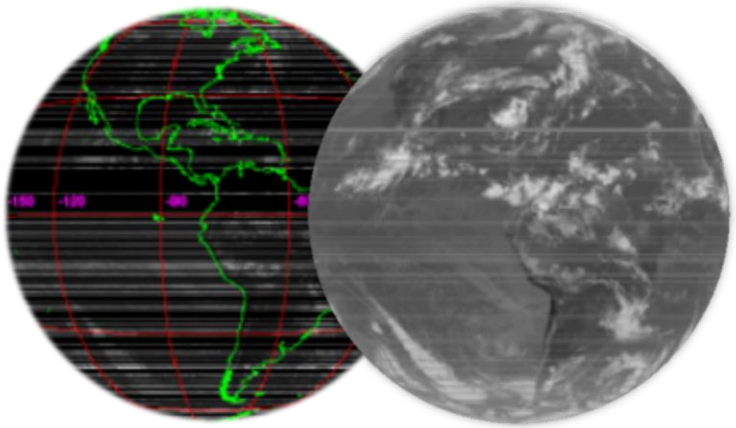
Extra·ordinary datasets: GEO-Ring

- 1983: Int. Satellite Cloud Climatology Project (ISCCP) pioneered use of the GEO-Ring by combining all GEOstat data.
- We are currently working to produce GEO-Ring CDRs by addressing consistent quality control and recalibration.
- Basis for gridded level-1 data record (L1g) composed of radiance measurements.



Extra·ordinary datasets: GeoRing

Data Rescue → Quality Control → Recalibration Cross-Calibration → Validation Evaluation

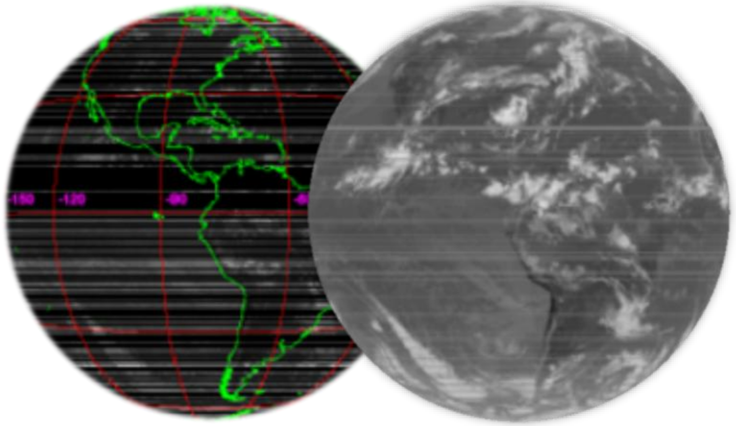


NOAA CLASS
GOES-5 IR
17.08.1981

SSEC reprocessed
GOES-5 IR
17.08.1981

Extra·ordinary datasets: GeoRing

Data Rescue → Quality Control → Recalibration Cross-Calibration → Validation Evaluation

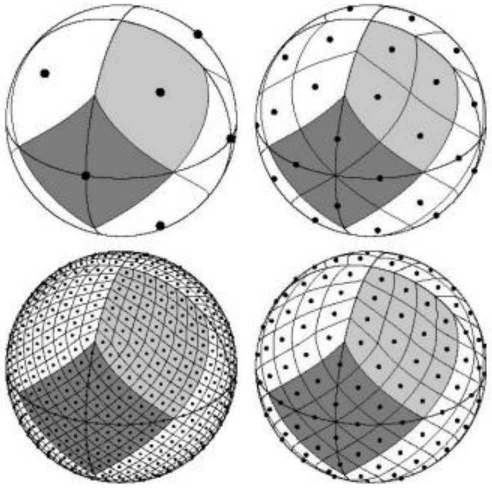


NOAA CLASS GOES-5 IR 17.08.1981

SSEC reprocessed GOES-5 IR 17.08.1981

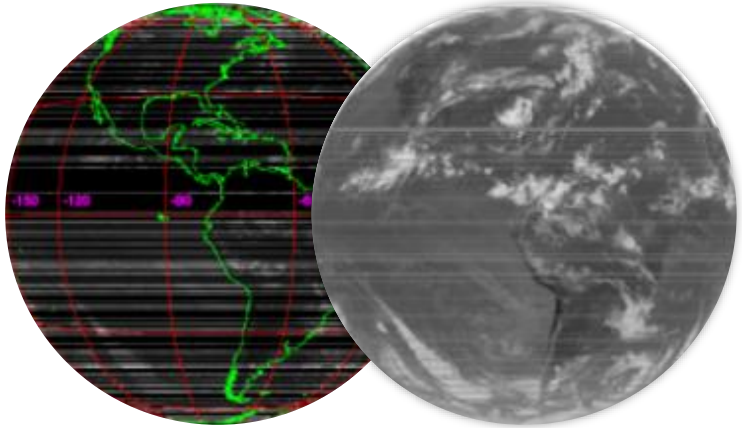
↓↓↓↓↓

Product Generation



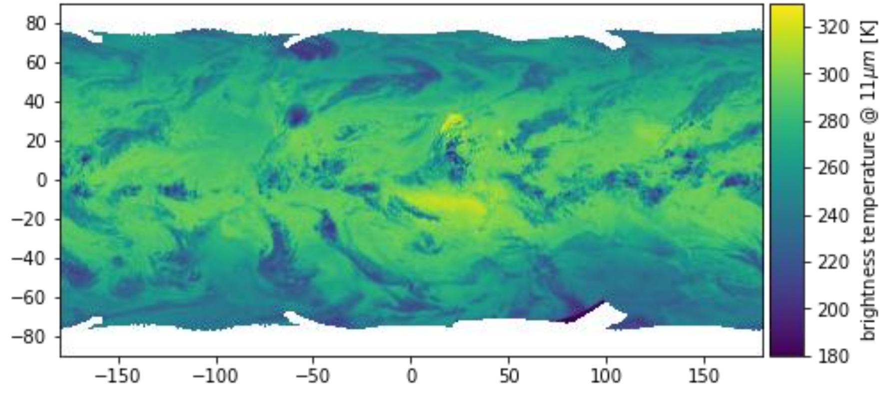
Extra·ordinary datasets: GeoRing

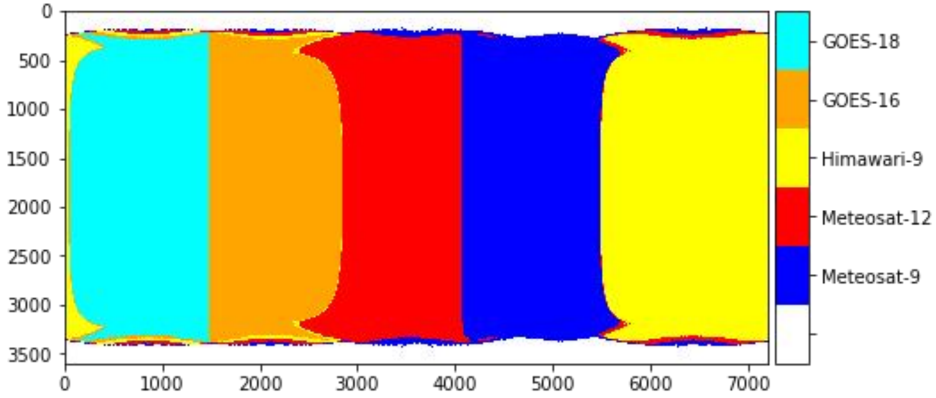
Data Rescue → Quality Control → Recalibration Cross-Calibration → Validation Evaluation



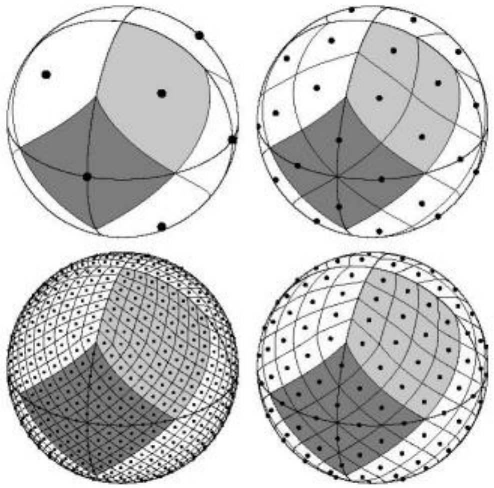
NOAA CLASS
GOES-5 IR
17.08.1981

SSEC reprocessed
GOES-5 IR
17.08.1981

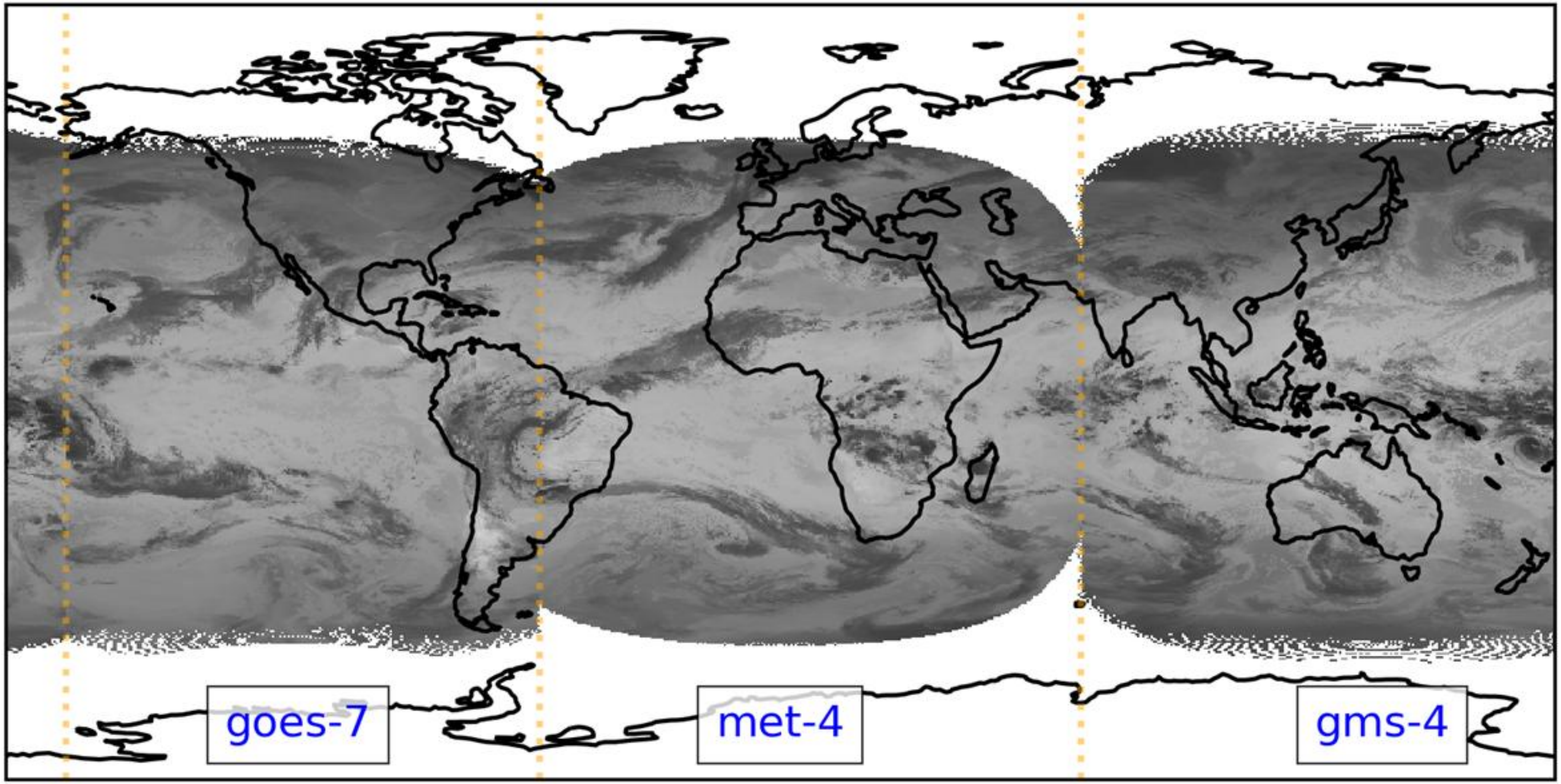




Product Generation

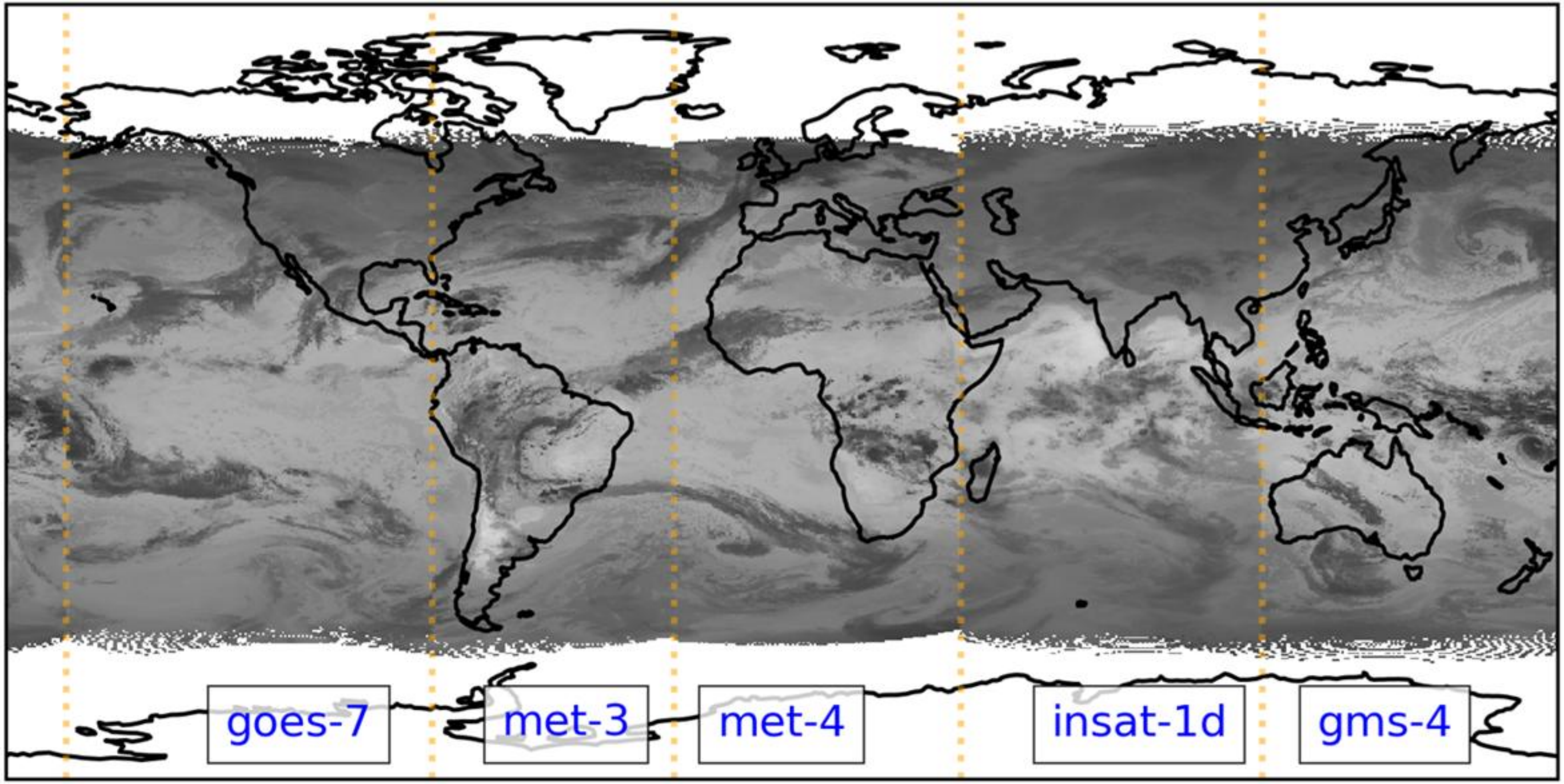


Classical Geostationary Coverage



1993-01-01T18:00

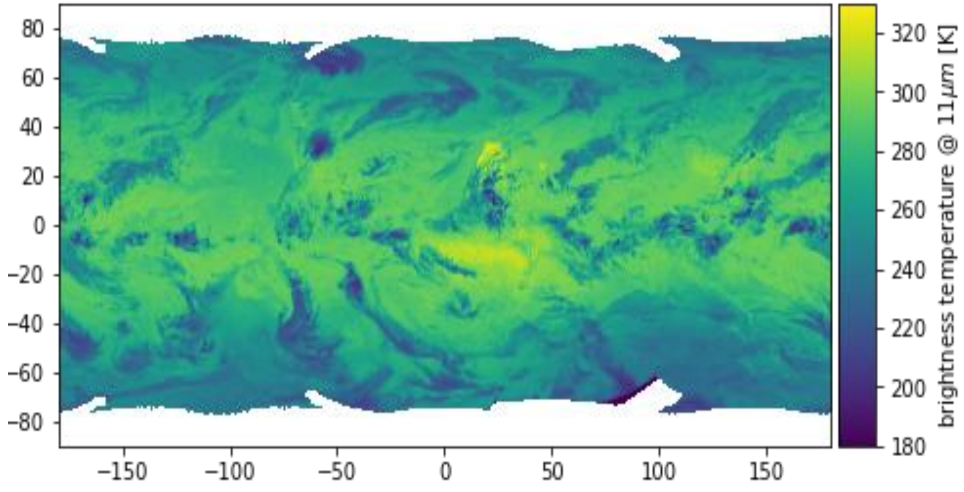
Enhanced Geostationary Coverage



1993-01-01T18:00

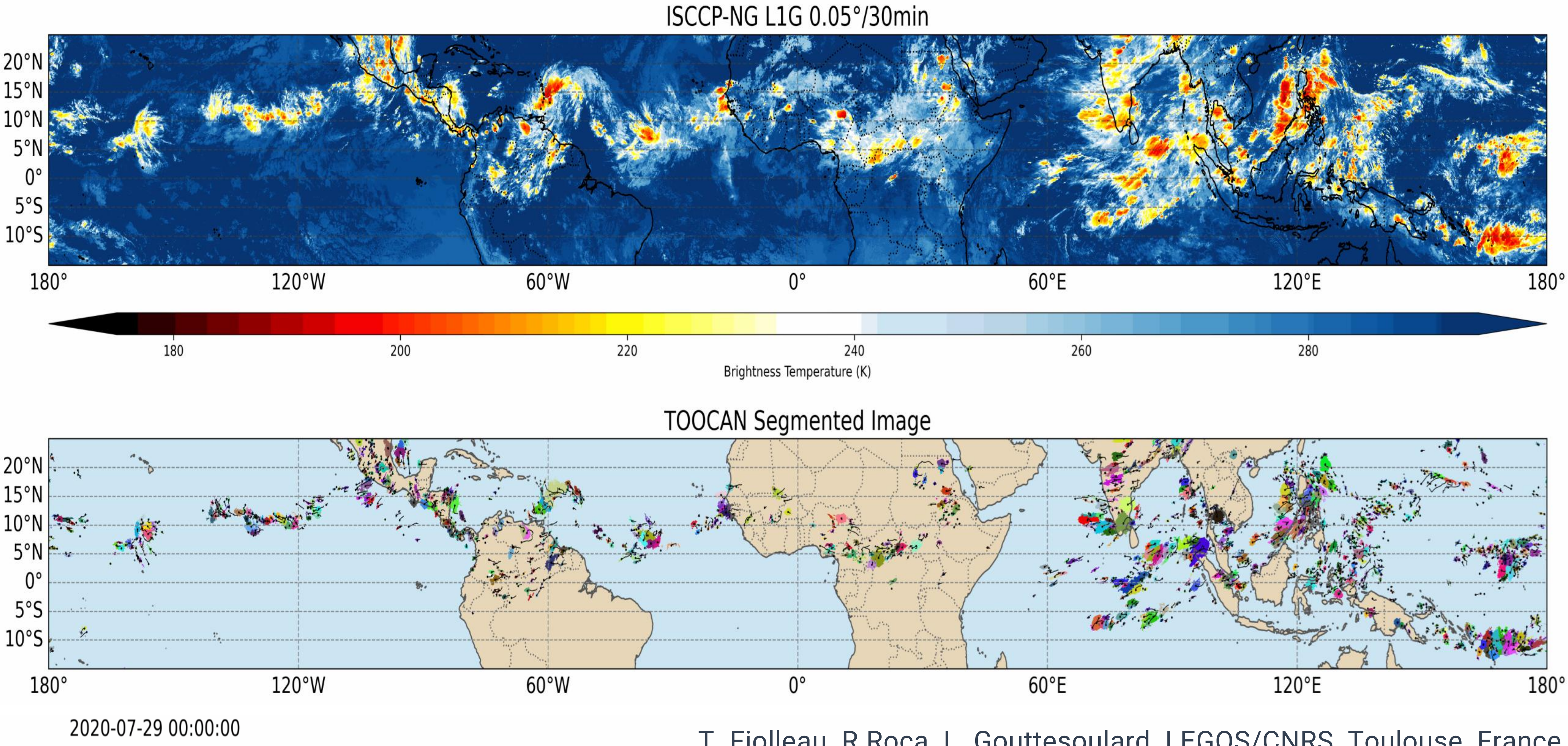
Extra·ordinary datasets: L1 GEO-Ring

Static test data set is available.



- 6 years: 2019 – 2024
- All channels: ABI, AHI and SEVIRI
- 5 km^2/ 0.05° 30-minute sampling
- Equal area HEALPix grid in NetCDF and Zarr, and equal angle grid in NetCDF.
- EWC and EUMETSAT Data Store distribution, NOAA cloud is expected as well.
- Use of data is free of charge, but redistribution isn't permitted.

Extra·ordinary datasets: MCS climatology

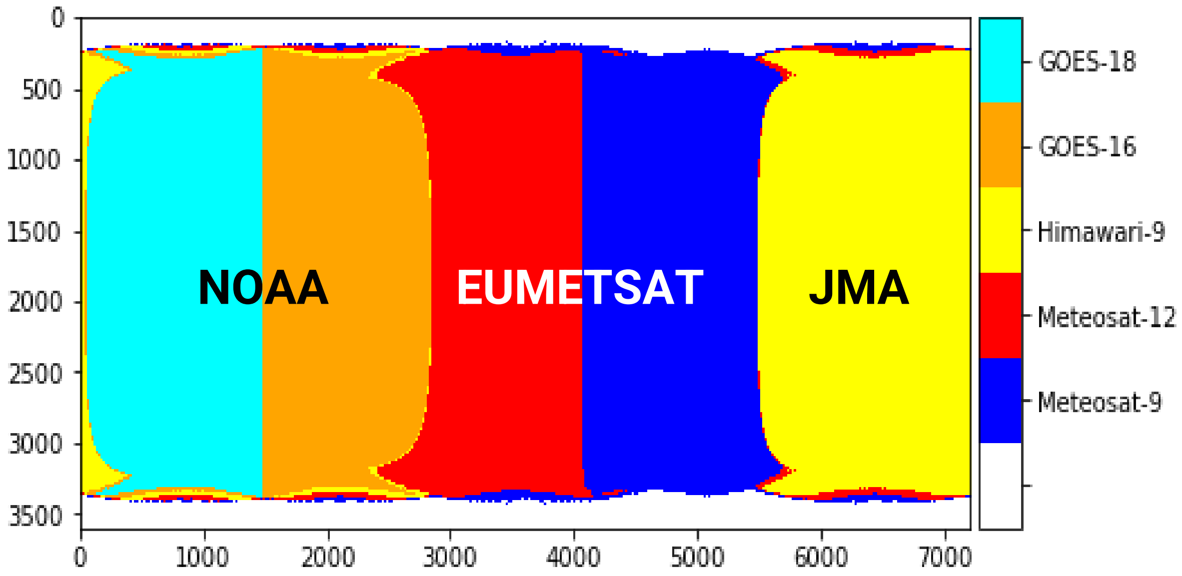


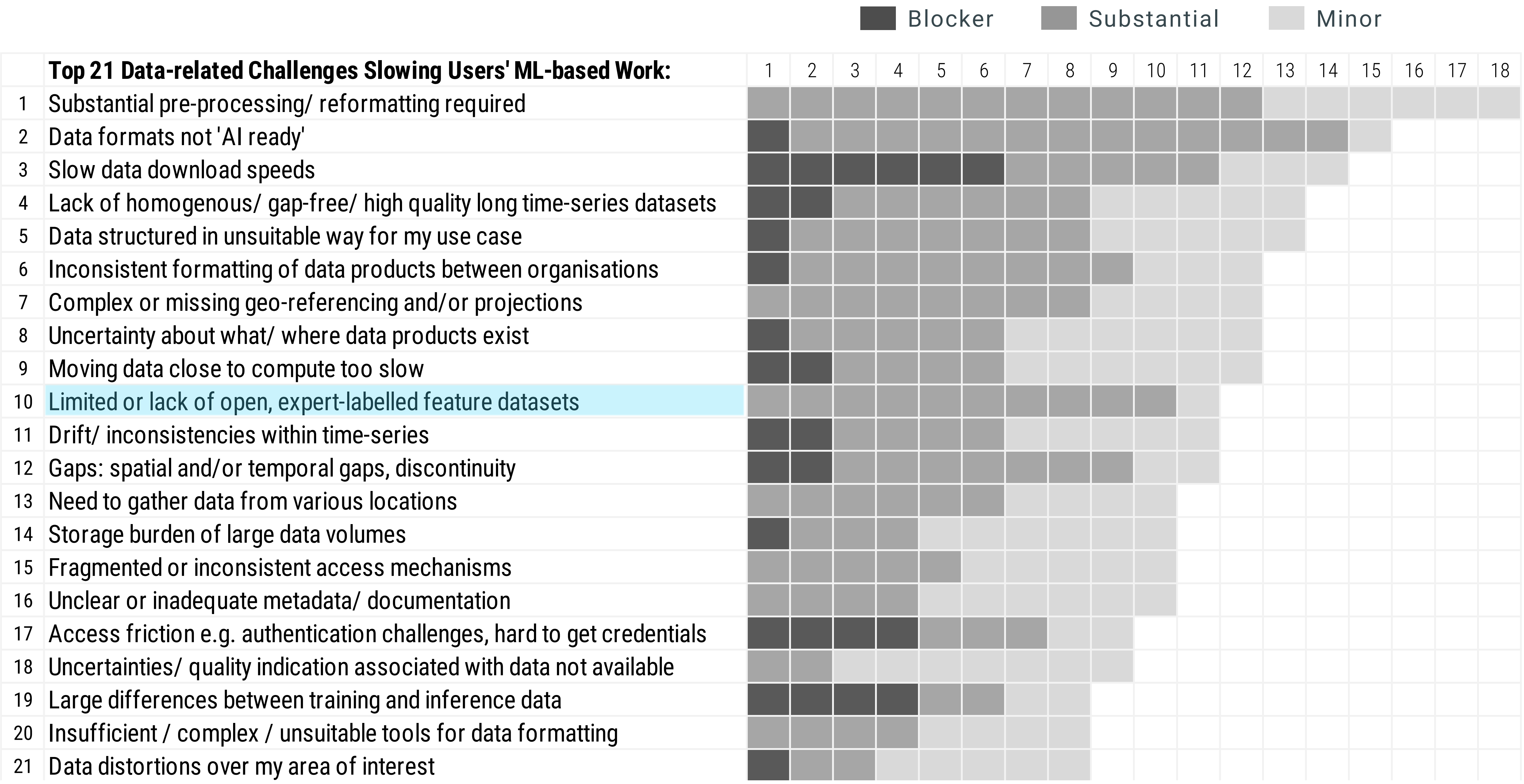
T. Fiolleau, R.Roca, L. Gouttesoulard, LEGOS/CNRS, Toulouse, France

Extra·ordinary datasets: L1 GEO-Ring

Looking ahead:

- Full L1g product spanning 1981-2026 planned for late 2027.
- Low latency delivery planned for ICDR (~5 day)
- Requests for NRT have also been received; this is under consideration.





Blocker
Substantial
Minor
Top 21 Data-related Challenges Slowing Users' ML-based Work:
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18
1 Substantial pre-processing/ reformatting required
2 Data formats not 'AI ready'
3 Slow data download speeds
4 Lack of homogenous/ gap-free/ high quality long time-series datasets
5 Data structured in unsuitable way for my use case
6 Inconsistent formatting of data products between organisations
7 Complex or missing geo-referencing and/or projections
8 Uncertainty about what/ where data products exist
9 Moving data close to compute too slow
10 Limited or lack of open, expert-labelled feature datasets
11 Drift/ inconsistencies within time-series
12 Gaps: spatial and/or temporal gaps, discontinuity
13 Need to gather data from various locations
14 Storage burden of large data volumes
15 Fragmented or inconsistent access mechanisms
16 Unclear or inadequate metadata/ documentation
17 Access friction e.g. authentication challenges, hard to get credentials
18 Uncertainties/ quality indication associated with data not available
19 Large differences between training and inference data
20 Insufficient / complex / unsuitable tools for data formatting
21 Data distortions over my area of interest

Feature•Space

Collaborative Feature Labelling Platform for ML Applications

EUMETSAT
FeatureSpace for Collaborative Labelling

Darmstadt

25 - 27 March 2026

As part of Machine learning (ML) activities for detection and forecasting of extreme events, EUMETSAT is committed to support its member states by providing data, tools, and community platforms. This includes the development of the Joint Working Environment for Feature Identification in the European Weather Cloud, including a Collaborative Label Space for global atmospheric and Earth surface feature tracking.

This in-person training workshop will take place 25-27 March 2026, and will be hosted by EUMETSAT at the headquarters in Darmstadt. After a general introduction to the Joint Working Environment, two sessions will run in parallel — one for developers, the other for labelers.

The event aims to bring together researchers, developers, and stakeholders from the domains of climate and weather from EUMETSAT member states. Preference will be given to national meteorological services, but all meteorological organisations may nominate participants. Researchers from marine, terrestrial, and atmosphere domains are also welcome.

Workshop VI Europe English

Target Audience

Primary: Researchers, developers, and stakeholders from the domains of climate and weather from EUMETSAT member states. Secondary: Researchers from marine, terrestrial, and atmosphere domains.

Hosted by

EUMETSAT

Contact

EUMETSAT
FeatureSpace for Collaborative Labelling

Darmstadt

25 - 27 March 2026

Expected Learning Outcomes

- Leverage the European Weather Cloud for Earth Observation data preparation using S3 buckets.
- Navigate and use tools on the **LabelStudio** platform, including SAM2.
- Accurately label key features in the prepared Earth Observation data.
- Export Feature labels for downstream ML model development on the European Weather Cloud.


Germany

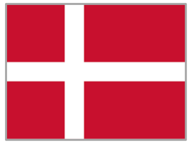

Bulgaria


Norway


Romania


Spain

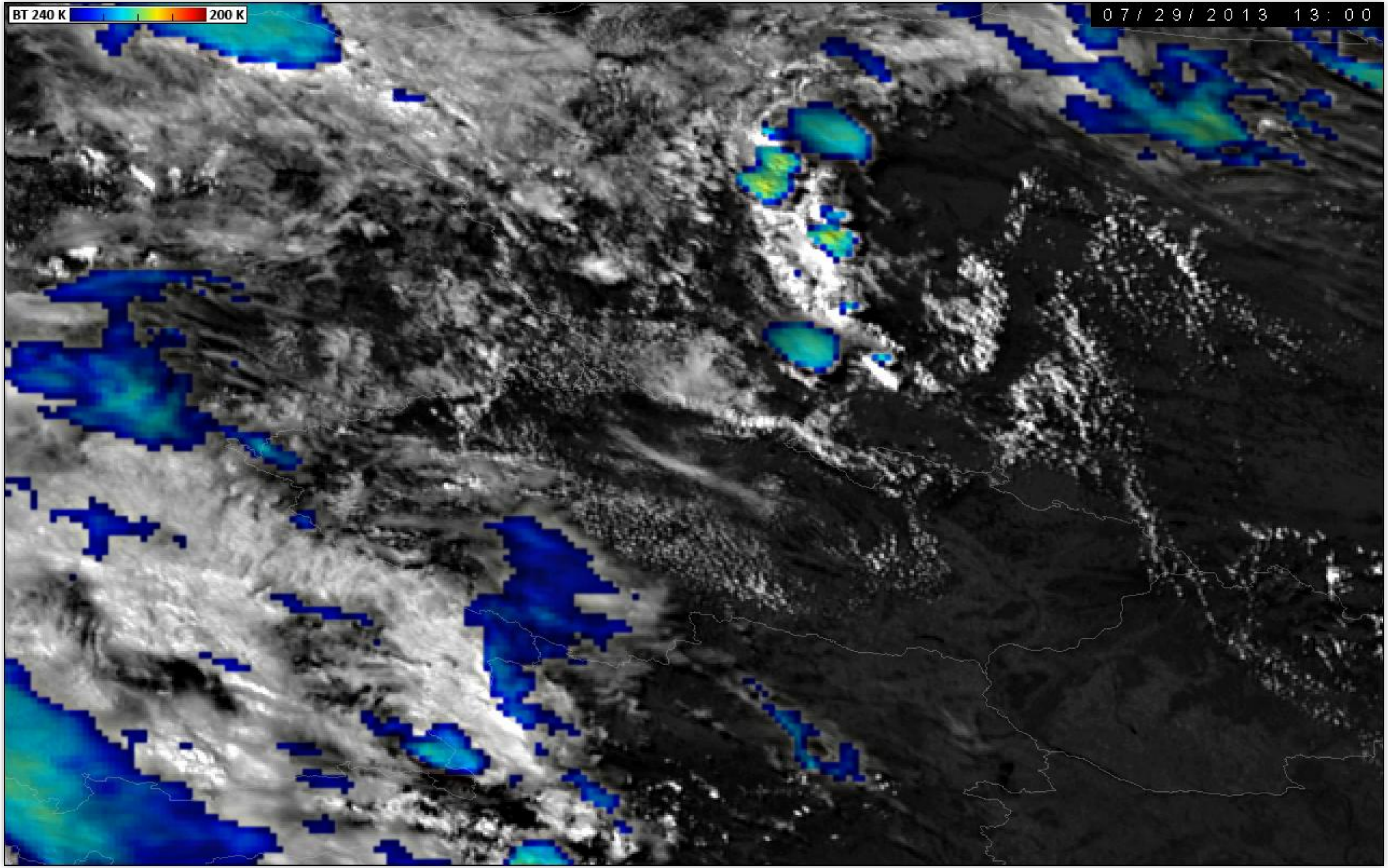

Italy


Denmark


Czechia


Croatia

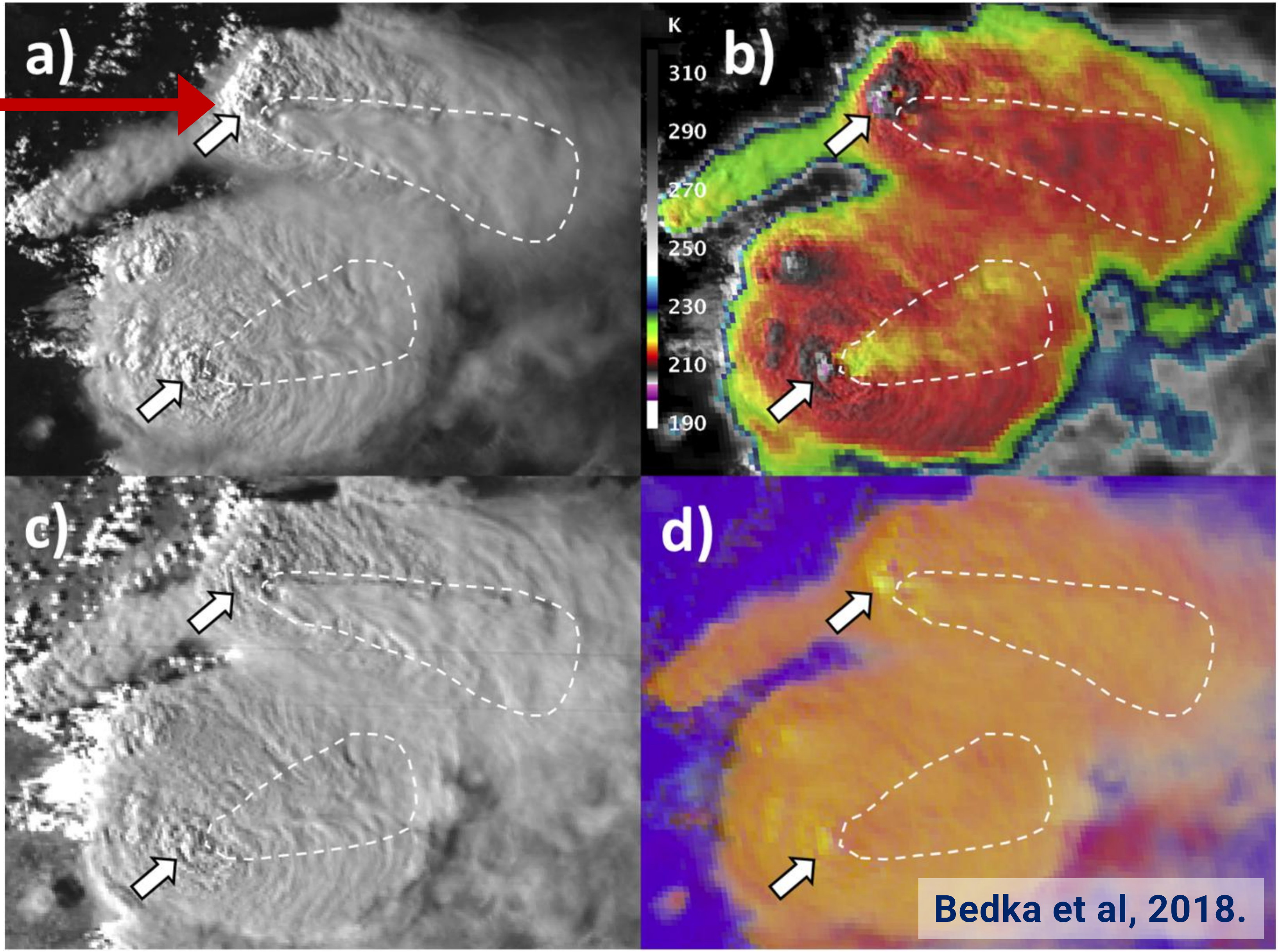
EUMETSAT
FeatureSpace for Collaborative Labelling



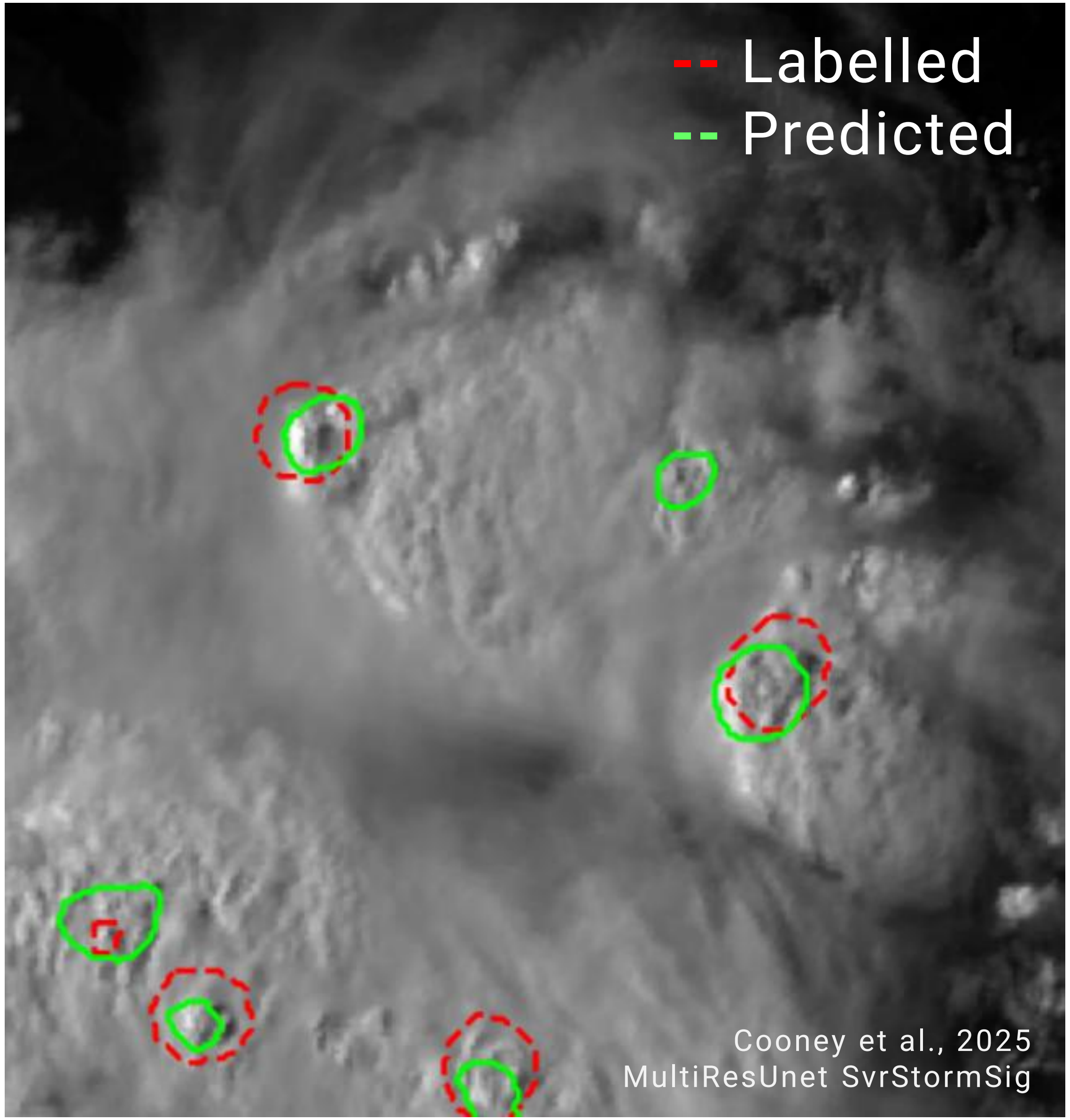
FeatureSpace for Collaborative Labelling

This storm produced hail over 18 cm in diameter…

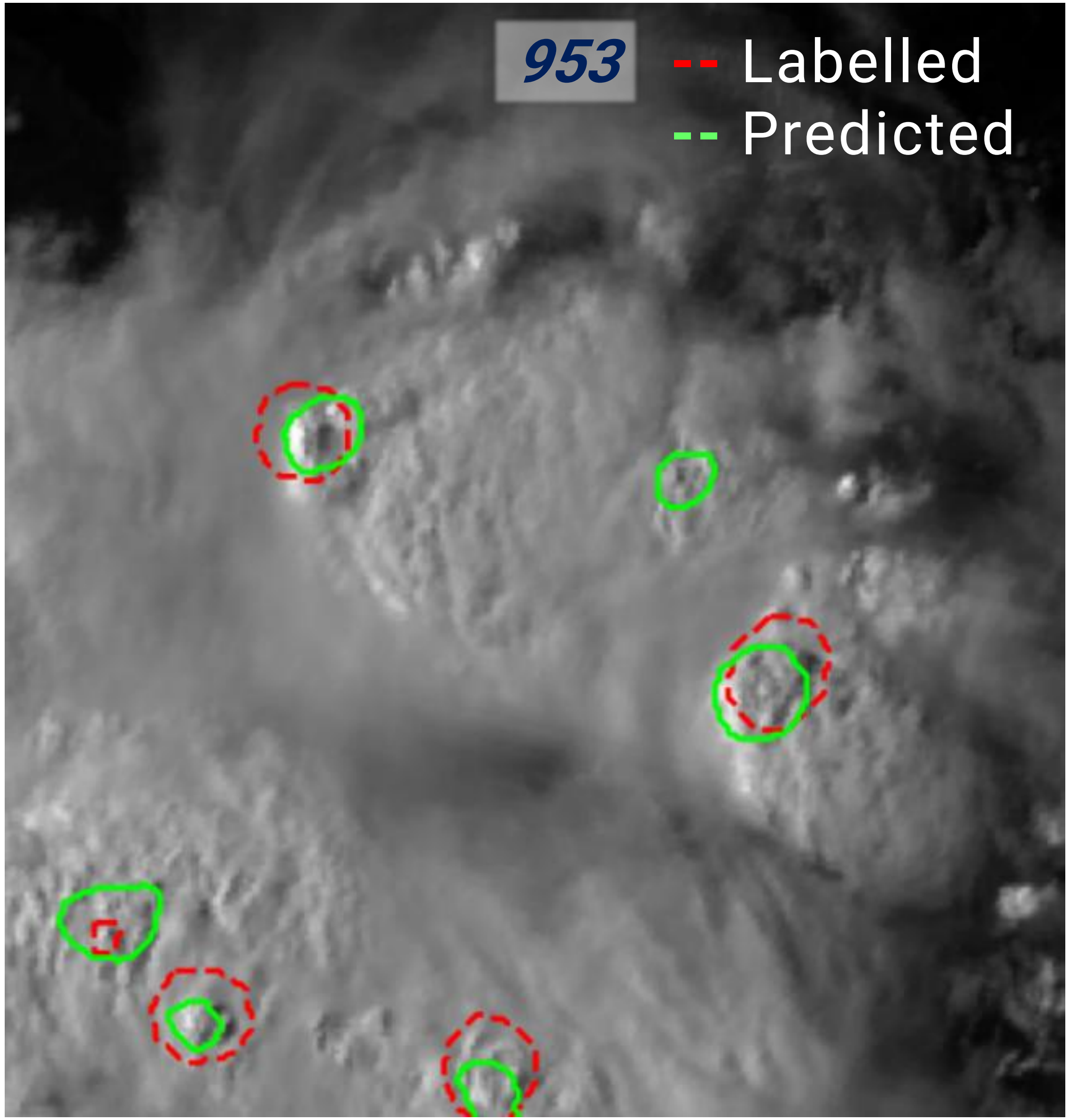
Photo by Victoria Druetta, Argentina, 08.02.2018



FeatureSpace for Collaborative Labelling



FeatureSpace for Collaborative Labelling



ML Applications for Earth System Observations

1. Access to long time series of expert-labelled features, together with the tools to explore them, adds value across weather and climate research domains.
2. These are high-quality training data for ML; supporting more accurate warning generation, improved medium-range forecasting, refined climate research, and **robust evaluation of different model outputs**.
3. Expert labelling of features in freely available EO data should be collaborative where possible.

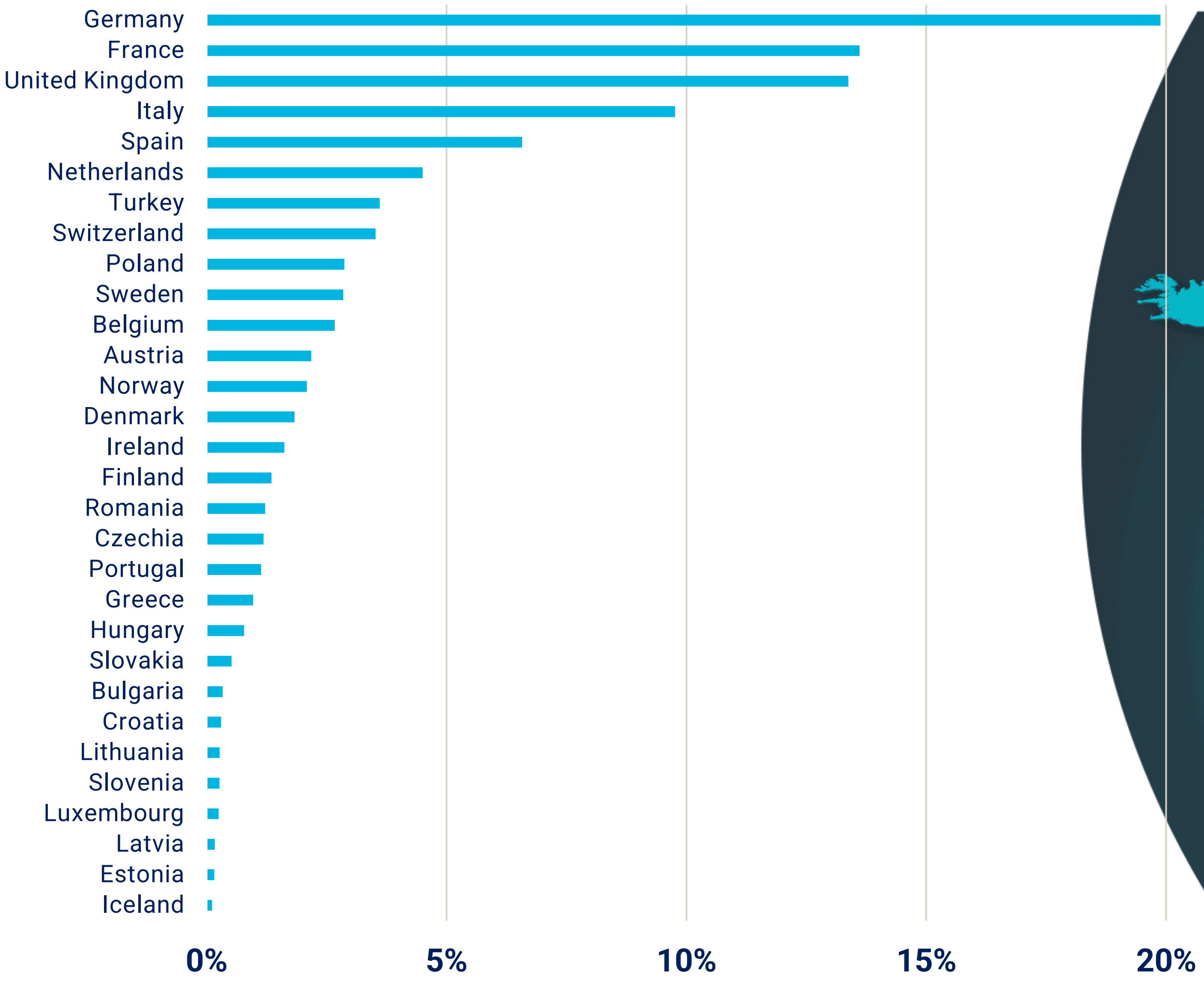
FeatureSpace for Collaborative Labelling

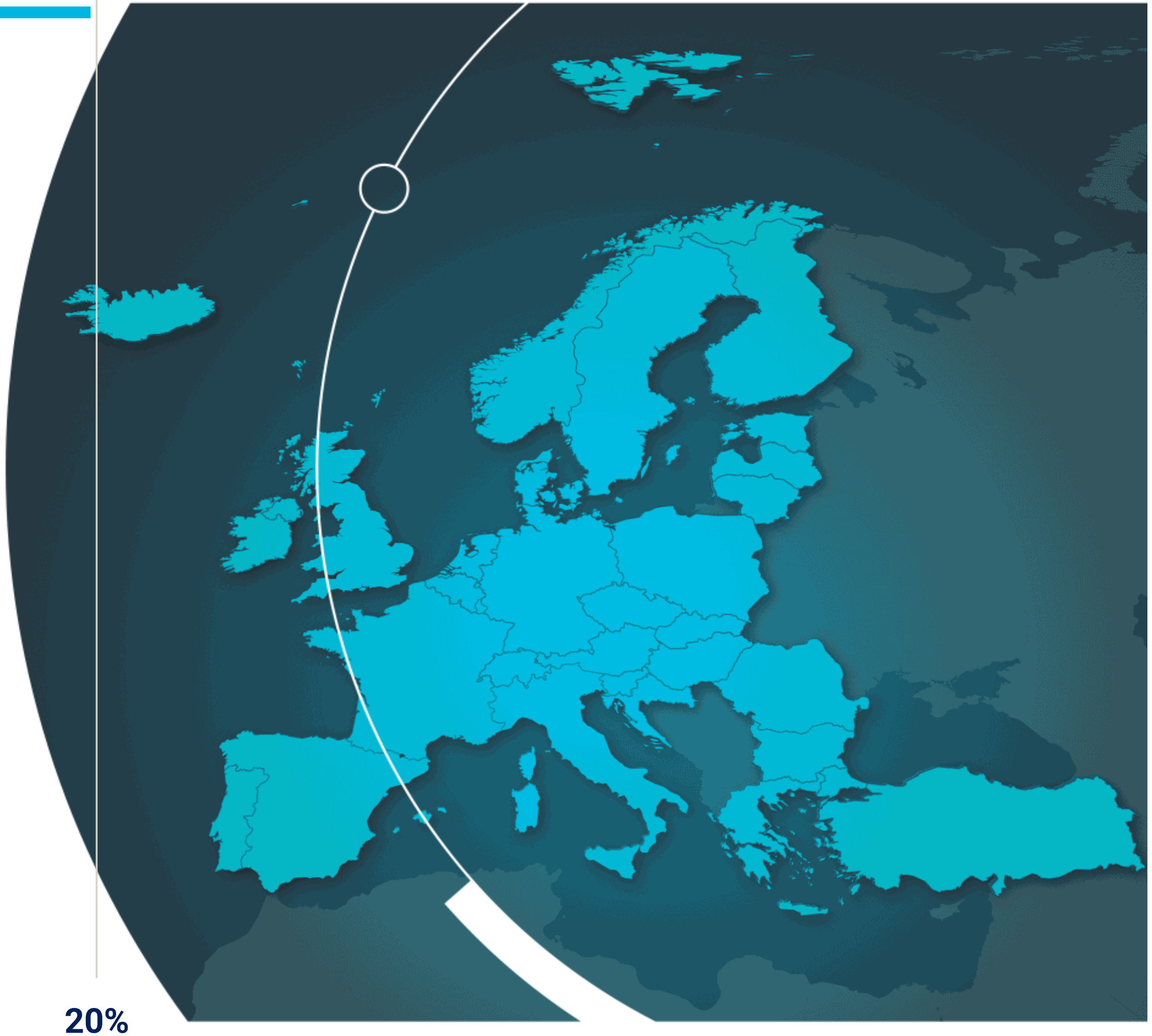
- It is not up to EUMETSAT to decide which features are ‘important’ – if a member user has flagged a feature as valuable, it’s valuable.
- Through Label Studio, the EWC Feature Space provides data, platform, and tools by which users can label their priority features [*and pull into models*].
- EWC Research & Development calls →

Member States Contributions to EUMETSAT

www.eumetsat.int

Scale of Contributions 2024-2026





Thank you!

Questions are welcome.