Observations: data rescue and reprocessing

Climate Change

Paul Poli (ECMWF/C3S)

Acknowledgements

Reanalysis

Hans Hersbach, Bill Bell, Adrian Simmons, Cornel Soci, Dinand Schepers, Raluca Radu, Paul Berrisford, Joaquin Sabater, Patricia de Rosnay

Satellite observations

Rob Roebeling, Viju John, Marie Doutriaux-Boucher, Joerg Schulz, Alessio Lattanzio, Kristina Petraityte, Mike Grant, Timo Hanschmann, Jaap Onderwaater, Oliver Sus, Roger Huckle, Dorothee Coppens, Bertrand Theodore, Thomas August, Jonathan Mittaz, Thomas Hall, Jerome Vidot, Pascal Brunel, James Johnson, Emily Zamkoff, Atheer Al-Jazrawi, Asghar Esfandiari, Irina Gerasimov, Pascal Prunet, Andrzej Klonecki, Carsten Standfuss

In-situ observations

Peter Thorne, Simon Noone, Kevin Healion, Leopold Haimberger, Federico Ambrogi, Ulrich Voggenberger, Fabio Madonna, Monica Proto, Emanuele Tramutola, Simone Gagliardi, Fabrizio Marra, Alessandro Fasso, Kalev Rannat, Patricia Rodriguez, Jean-Philippe Rannou, Thierry Carval, Dominique Obaton, Tanguy Szekely , Jerome Gourrion, Sylvie Pouliquen, Gilles Reverdin, Ed Hawkins

Copernicus Europe's eyes on Earth

European Commission

Climate Change

Outline

I. Introduction: On the importance of observations in reanalysis

II. Reanalyze, Rescue, Reprocess

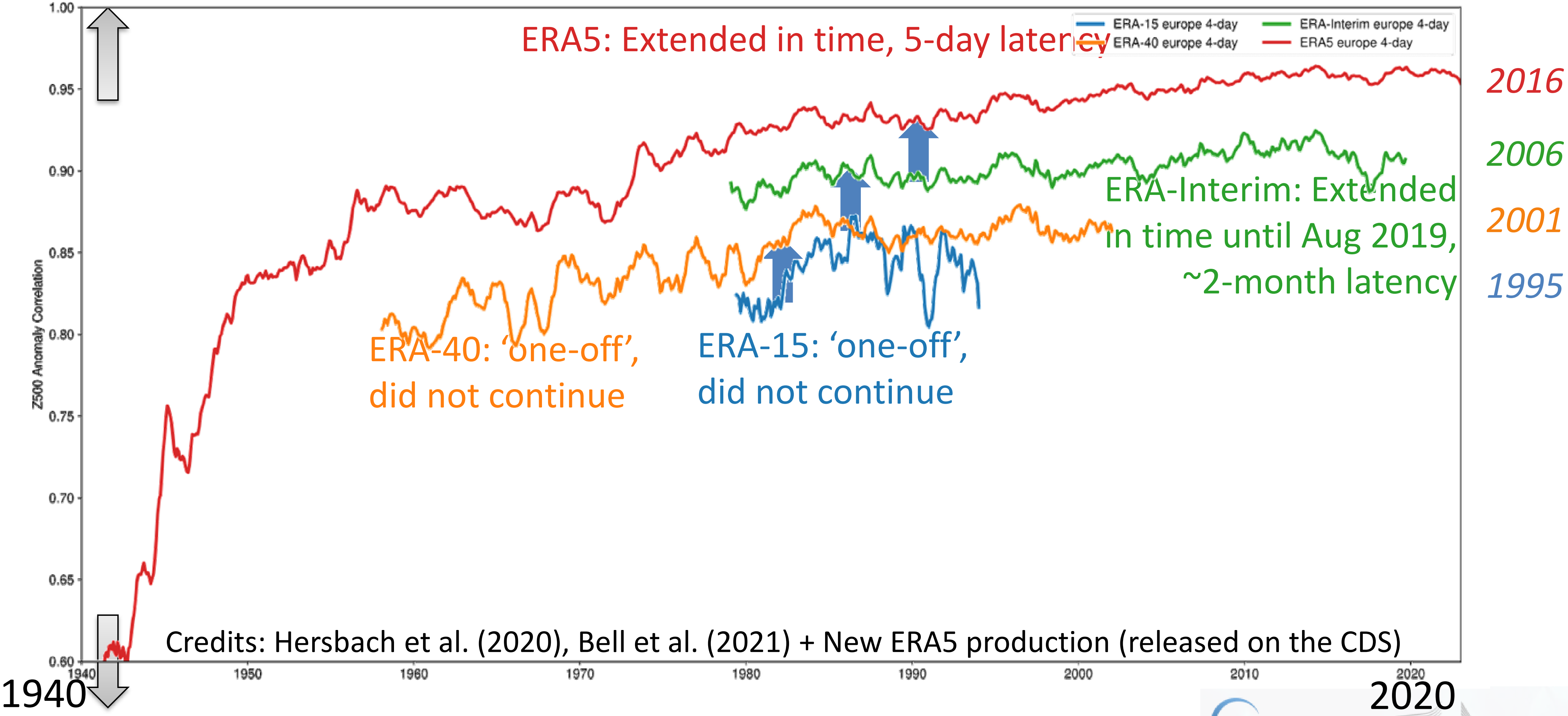
III. Axes of improvement
   1) Improving the availability
   2) Improving the quality
   3) Improving the usage

IV. Discussion and conclusions

Climate
Change

Part I. On the importance of observations in reanalysis

“Perfect forecast”



ERA5: Extended in time, 5-day latency

ERA-Interim: Extended in time until Aug 2019, ~2-month latency

ERA-40: ‘one-off’, did not continue

ERA-15: ‘one-off’, did not continue

Credits: Hersbach et al. (2020), Bell et al. (2021) + New ERA5 production (released on the CDS)

Correlation of 4-day forecasts of geopotential height anomalies at 500 hPa with corresponding analyses

How useful are the observations? Let’s look at *explained variance*

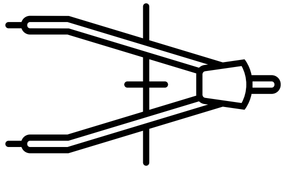
$d_b = y^0 - h(x^b)$ $\qquad$ $y^0 = y^t + \varepsilon_0$ $\qquad$ $d_b = \varepsilon_0 - \varepsilon_b - \varepsilon_h$

$$EV_b = \frac{Var(y^0) - Var(d_b)}{Var(y^0)}$$

← Explained Variance
(*in the observations*, *by the background*)

$$EV_b = \frac{\sigma_t^2 + \sigma_0^2 - \sigma_b^2 - \sigma_h^2 - \sigma_0^2}{\sigma_t^2 + \sigma_0^2}$$

$$EV_b = \frac{\sigma_t^2 - \sigma_b^2 - \sigma_h^2}{\sigma_t^2 + \sigma_0^2}$$



$$EV_a = \frac{\sigma_t^2 - \sigma_a^2 - \sigma_h^2}{\sigma_t^2 + \sigma_0^2}$$

Explained Variance Gain

$$EGV = EV_a - EV_b$$

d	Observation departure
y^0	Observation
x_b	Background
x_a	Analysis
$h(\,)$	Observation operator
y^t	Truth in obs. space
ε_0	Observation error
ε_b	Background error
ε_h	Representativeness error
ε_a	Analysis error
σ	Standard deviation
σ'	*Assumed* standard deviation

Notes:

1. Assuming normal distributions, uncorrelated error sources
2. Normalization by assumed obs. error standard deviation, σ_0', makes it possible to run this diagnostic across variables, across geographical regions, and across vertical levels.
3. For a large data sample, one would hope that $\sigma_t^2 \gg \sigma_0^2$
(or else the observations would be of little value!)

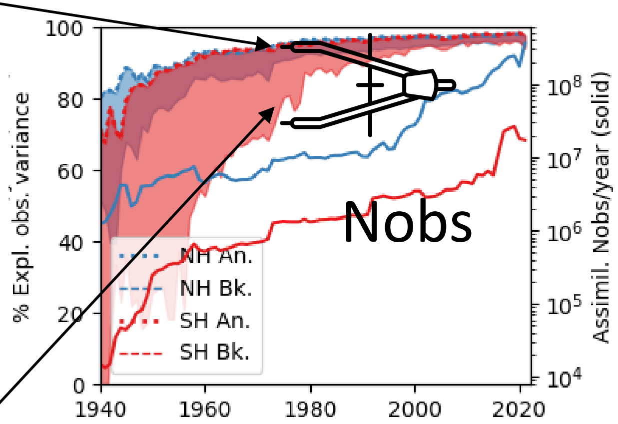
$$d_b^n = \frac{y^0 - h(x^b)}{\sigma_0'}$$

How useful are the observations? Explained variance gains in ERA5

Climate Change

% observation variance explained by ERA5 **analysis**

Troposphere, zonal wind

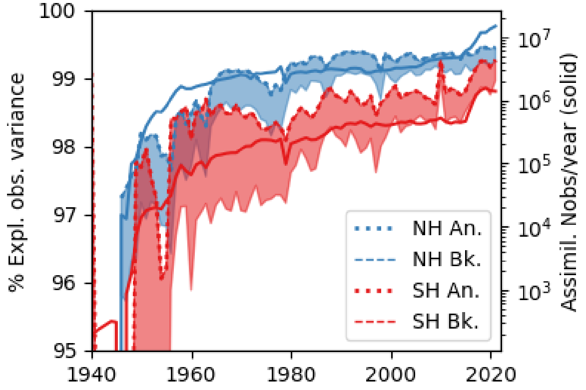


% observation variance explained by ERA5 **background**

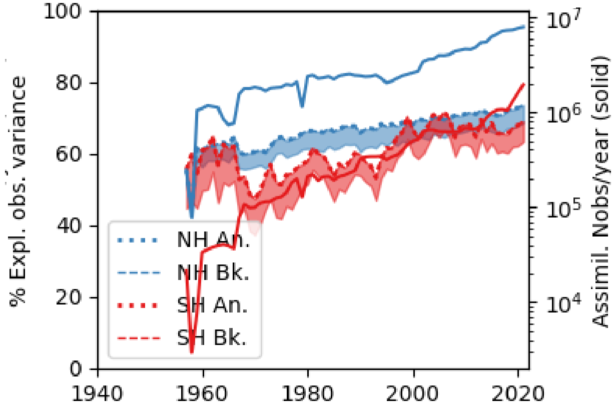
20ºN-90ºN

20ºS-90ºS

Stratosphere, temperature



Near-surface, humidity



Copernicus Europe's eyes on Earth

European Commission

Part II. Why aren't all observations already online and ready to use?

- Observations were (are) acquired <u>**in all cases**</u> for specific purpose(s) (*decision making*).
    - This purpose was <u>**generally not**</u>, 'documenting the present for eternity'
    - Observations do not escape principles of <u>**Intellectual Property**</u>
    - Observations got lost, or got saved, for a myriad of different reasons.
        - With the notable exception of climatological archives, observations were rarely kept *by design*, more *by circumstance.* (*e.g.,* as accepted good practice, as not for any individual to decide on destruction).
        - This lack of intention/long-term planning explains why so many archives got lost when institutions moved or changed guardianship
- Why so little interest in (what is perceived as) "investing in the past"?
    - The phrase says it all. It's about costs and timeliness. (the more) recent and present are the data → (the faster) the reaction time that is expected to handle them – in order to *make a decision*. Conversely, the more distant (the older) the data → the lower the priority
    - However, from a pure cost point of view,
      the expenses of making the observation were already made.
    - <u>Saving</u> what still exists today, only costs a **fraction** of what it did cost at the time to acquire these observations in the first place.
- Unsurprisingly, there are *rather few* large-scale data rescue programmes.
    - But there are many bottom-up initiatives!

US Army signal service,
Cape Mendocino (1888)
Credits: NOAA

→ For more on this topic:
Griffin, 2015
DOI:10.1016/j.grj.2015.02.004

Climate Change

Data rescue: Four myth-busters

✗ "*Data rescue consists in turning back the clock*"

- ✓ Data rescue consists in placing ancient data in the context of modern data and science.

✗ "*Old data can't possibly teach us anything we don't know already*"

- ✓ Forgotten observations, by definition, cannot have been confronted with present knowledge.

✗ "*We have many observations today (in digital form), why bother with adding just a few (from analog records)*"

- ✓ Any switchover date (analog to digital records) is an artefact of data management practices (or change in obs. system), and the Earth system may have changed before or after.

✗ "*Old data with their unknown errors can only corrupt large-sample-statistics obtained from modern data*"

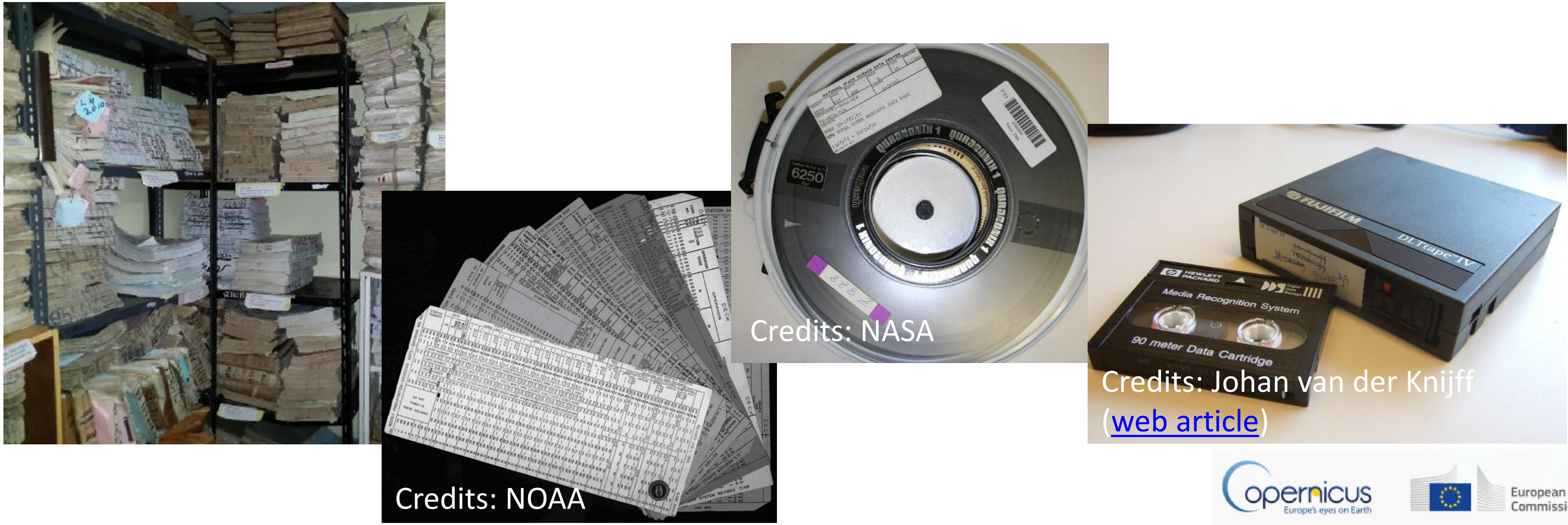
- ✓ The value of observations that all agree is far inferior to that of outliers: these indicate that science needs to be refined.

Copernicus Europe's eyes on Earth
European Commission

Data rescue: One golden priority

Golden priority: safeguard data "at risk" from complete loss

- e.g., ink fade, discs demagnetization, reading device (often specialized items at the time) obsolescence, destruction (war, earthquake, flood, fire…), unavailability (war, building unsafe, asbestos…), retirement of individuals with unique knowledge of how to read/interpret the data…



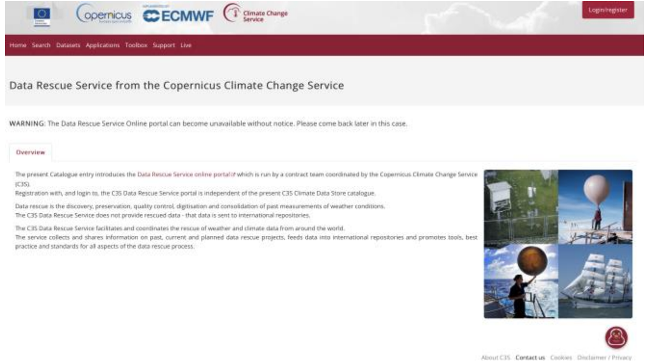
Credits: NOAA

Credits: NASA

Credits: Johan van der Knijff (web article)

Some data rescue programmes and activities

- Euro-Climhist: weather, climate, phenology, socio-political data, …
- Atmospheric Circulation Reconstructions over the Earth (ACRE)
- International Environmental Data Rescue Organization (IEDRO)
- NASA: Nimbus data
- US Geological Survey: Landsat data, bird phenology, glaciers, …
- Copernicus Climate Change Service Data Rescue Service (C3S)
- National Meteorological Services, under WMO-IDARE



Data Rescue Service from the Copernicus Climate Change Service

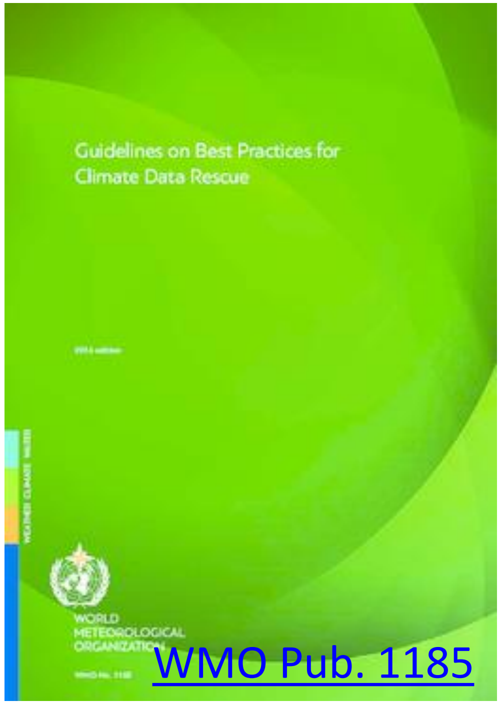
WARNING: The Data Rescue Service Online portal can become unavailable without notice. Please come back later in this case.

Overview

The present Catalogue entry introduces the Data Rescue Service online portal which is run by a contract team coordinated by the Copernicus Climate Change Service (C3S).
Registration with, and login to, the C3S Data Rescue Service portal is independent of the present C3S Climate Data Store catalogue.

Data rescue is the discovery, preservation, quality control, digitisation and consolidation of past measurements of weather conditions.
The C3S Data Rescue Service does not provide rescued data - that data is sent to international repositories.

The C3S Data Rescue Service facilitates and coordinates the rescue of weather and climate data from around the world.
The service collects and shares information on past, current and planned data rescue projects, feeds data into international repositories and promotes tools, best practice and standards for all aspects of the data rescue process.


Guidelines on Best Practices for Climate Data Rescue

WORLD METEOROLOGICAL ORGANIZATION

WMO Pub. 1185

Key aspects

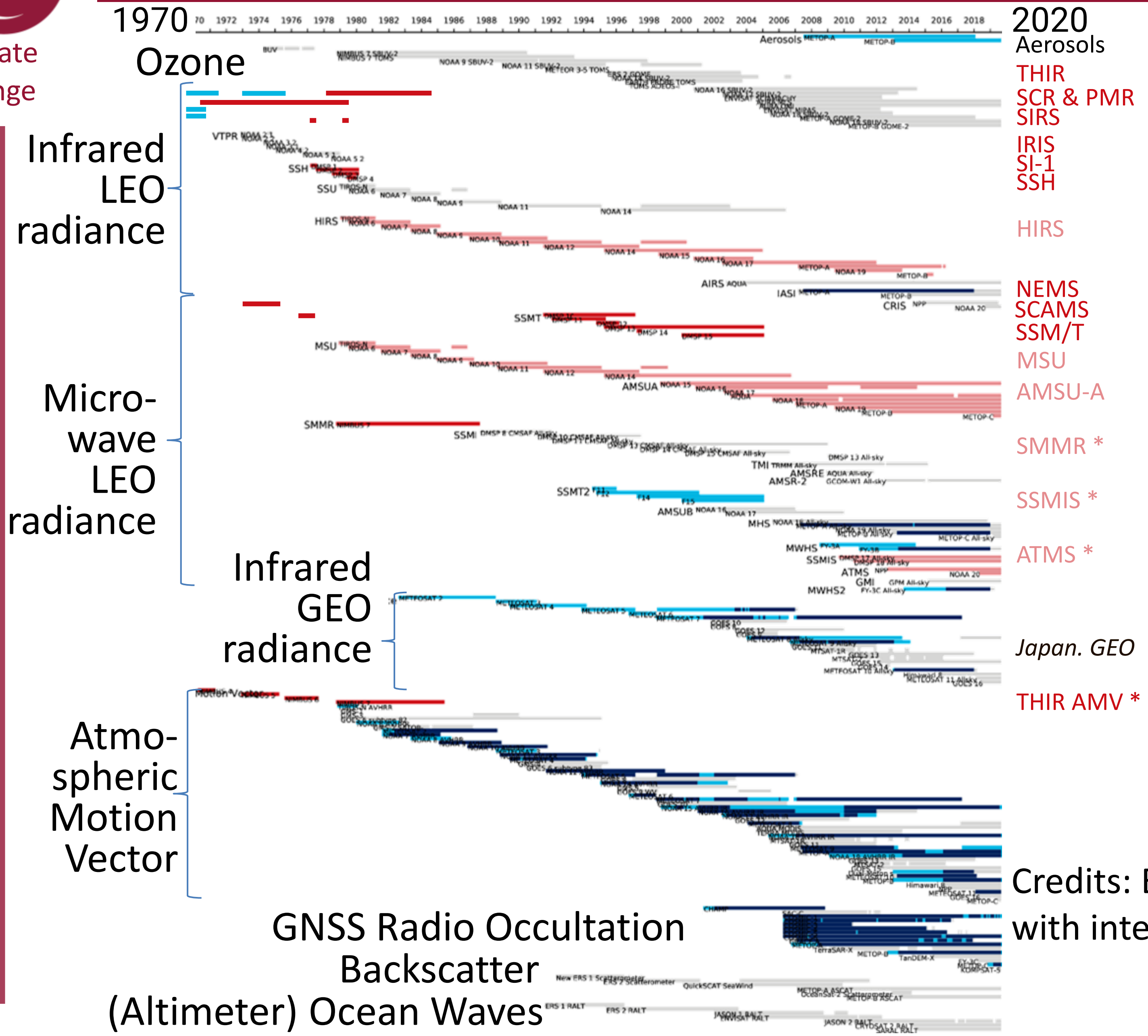
- History of the networks
- Metadata
- Quality Controls
- Data sources, media
- Data formats
- Corrections
- Special value of the observations
- Priorities
- Credits/ attribution

- Archive the sources
- Image the sources
- Digitize the images
- Archive images and databases
- **Make accessible**

→ For more on this topic:
Chimani et al., 2021 DOI:10.1002/gdj3.128

Climate Change

Rescued/reprocessed satellite observations



⬜ Unchanged, as in ERA5

Done in Cop1	Ongoing/plans for Cop2	
🟦	🟥	*New* data
■	■	*Improved* data

* Yet to be confirmed

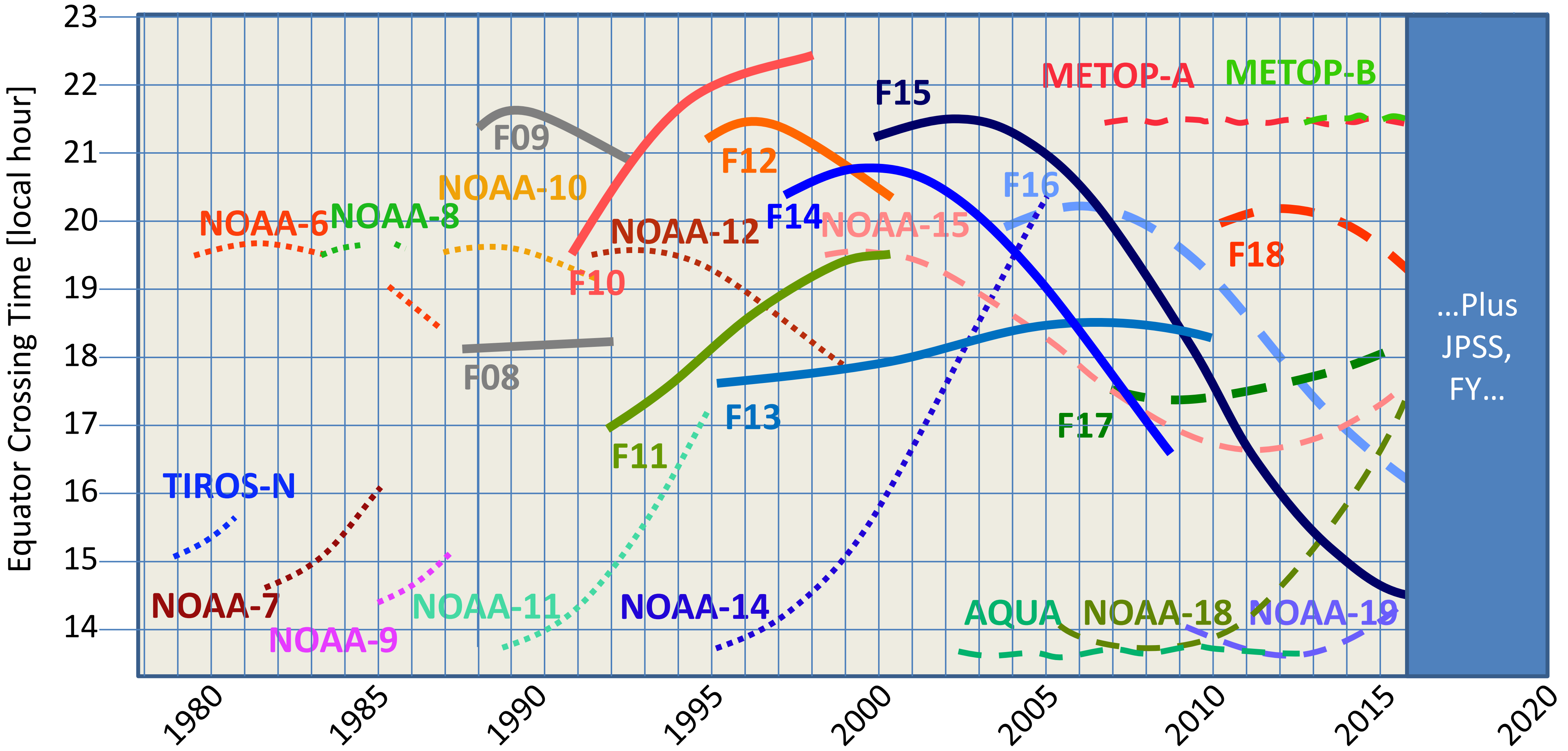
+ For high-resolution reanalyses (Europe & Arctic)

MFG & MFG Rapid-scan radiances 🟥
and AMV (* for MFG) 🟥
Metop AVHRR LAC AMV R3 ■

Credits: EUMETSAT, SPASCIA (C3S), with collaborations with international partners

Climate
Change

Why would an instrument like SSM/T (not used in NWP at the time) help today?



After REMSS and BCU ISR

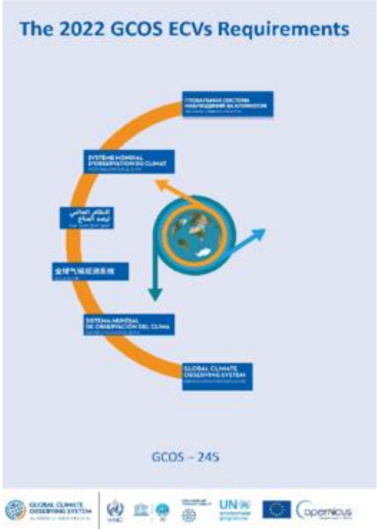
Current radiance data available to global reanalyses for temperature sounding in the band 50-60 GHz:

…. MSU: 4 channels - - AMSU-A or SSMIS: 12 or 13 channels

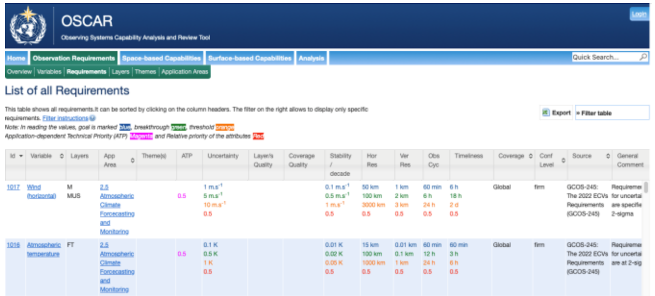
Thanks to NOAA holdings, EUMETSAT (C3S) has reprocessed: - SSMT: 7 channels

Climate
Change

Part III.1: Improving the quantity of observations: GCOS, WMO requirements

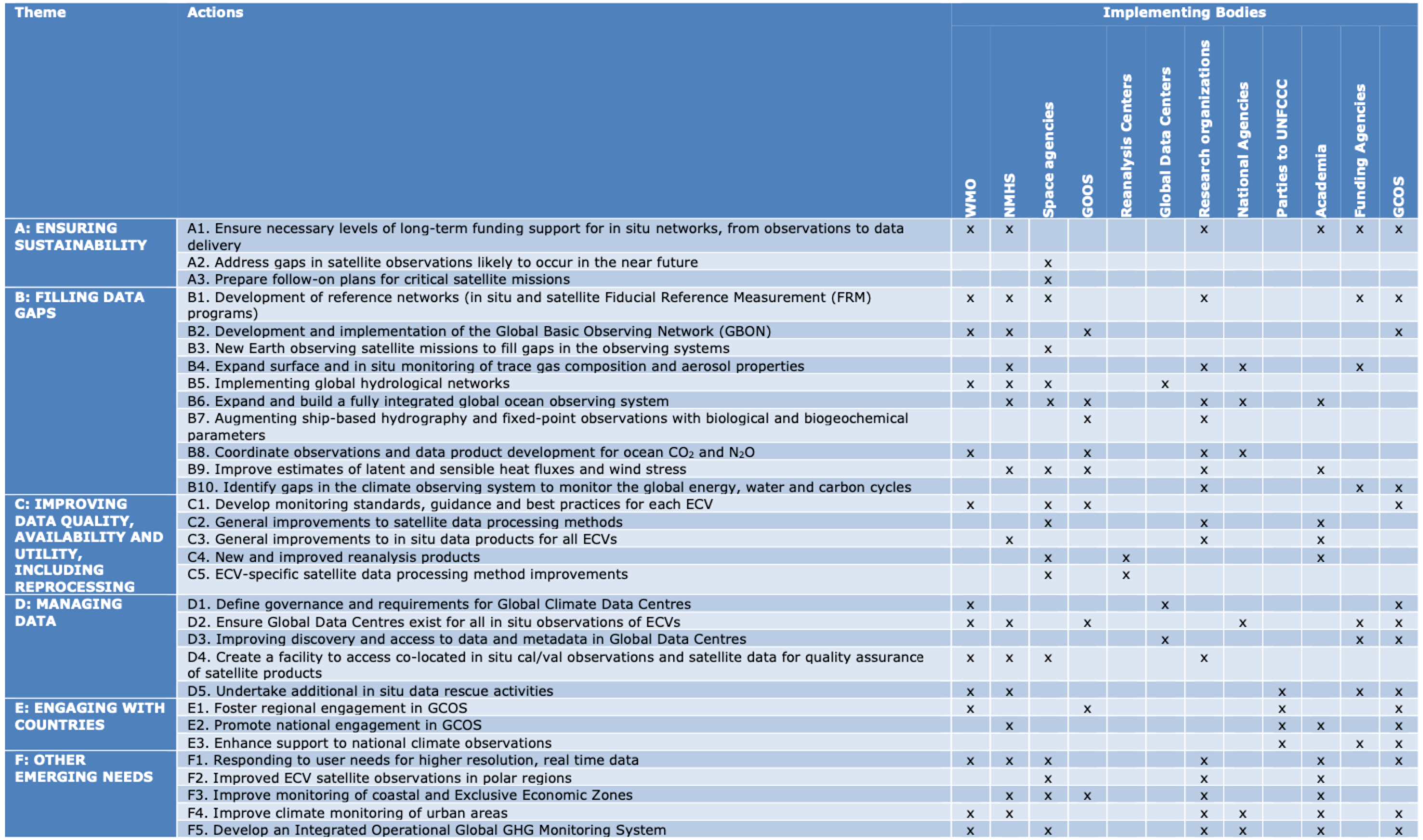


[GCOS requirements]



[WMO OSCAR requirements]

These requirements mostly act for new observing systems ‘going forward’. However, the GCOS implementation plan 2022 includes several actions to help prepare climate archives for the future:



Theme	Actions	Implementing Bodies											
		WMO	NMHS	Space agencies	GOOS	Reanalysis Centers	Global Data Centers	Research organizations	National Agencies	Parties to UNFCCC	Academia	Funding Agencies	GCOS
A: ENSURING SUSTAINABILITY	A1. Ensure necessary levels of long-term funding support for in situ networks, from observations to data delivery	x	x					x			x	x	x
	A2. Address gaps in satellite observations likely to occur in the near future			x									
	A3. Prepare follow-on plans for critical satellite missions			x									
B: FILLING DATA GAPS	B1. Development of reference networks (in situ and satellite Fiducial Reference Measurement (FRM) programs)	x	x	x				x				x	x
	B2. Development and implementation of the Global Basic Observing Network (GBON)	x	x		x								x
	B3. New Earth observing satellite missions to fill gaps in the observing systems			x									
	B4. Expand surface and in situ monitoring of trace gas composition and aerosol properties		x					x	x			x	
	B5. Implementing global hydrological networks	x	x	x			x						
	B6. Expand and build a fully integrated global ocean observing system		x	x	x			x	x		x		
	B7. Augmenting ship-based hydrography and fixed-point observations with biological and biogeochemical parameters				x			x					
	B8. Coordinate observations and data product development for ocean CO_2 and N_2O	x			x			x	x				
	B9. Improve estimates of latent and sensible heat fluxes and wind stress		x	x	x			x			x		
	B10. Identify gaps in the climate observing system to monitor the global energy, water and carbon cycles							x				x	x
C: IMPROVING DATA QUALITY, AVAILABILITY AND UTILITY, INCLUDING REPROCESSING	C1. Develop monitoring standards, guidance and best practices for each ECV	x		x	x								x
	C2. General improvements to satellite data processing methods			x				x			x		
	C3. General improvements to in situ data products for all ECVs		x					x			x		
	C4. New and improved reanalysis products			x		x					x		
	C5. ECV-specific satellite data processing method improvements			x		x							
D: MANAGING DATA	D1. Define governance and requirements for Global Climate Data Centres	x					x						x
	D2. Ensure Global Data Centres exist for all in situ observations of ECVs	x	x		x				x			x	x
	D3. Improving discovery and access to data and metadata in Global Data Centres						x					x	x
	D4. Create a facility to access co-located in situ cal/val observations and satellite data for quality assurance of satellite products	x	x	x				x					
	D5. Undertake additional in situ data rescue activities	x	x							x		x	x
E: ENGAGING WITH COUNTRIES	E1. Foster regional engagement in GCOS	x			x					x			x
	E2. Promote national engagement in GCOS		x							x	x		x
	E3. Enhance support to national climate observations									x		x	x
F: OTHER EMERGING NEEDS	F1. Responding to user needs for higher resolution, real time data	x	x	x				x			x		x
	F2. Improved ECV satellite observations in polar regions			x				x			x		
	F3. Improve monitoring of coastal and Exclusive Economic Zones		x	x	x			x			x		
	F4. Improve climate monitoring of urban areas	x	x					x	x		x		x
	F5. Develop an Integrated Operational Global GHG Monitoring System	x		x				x	x		x		x

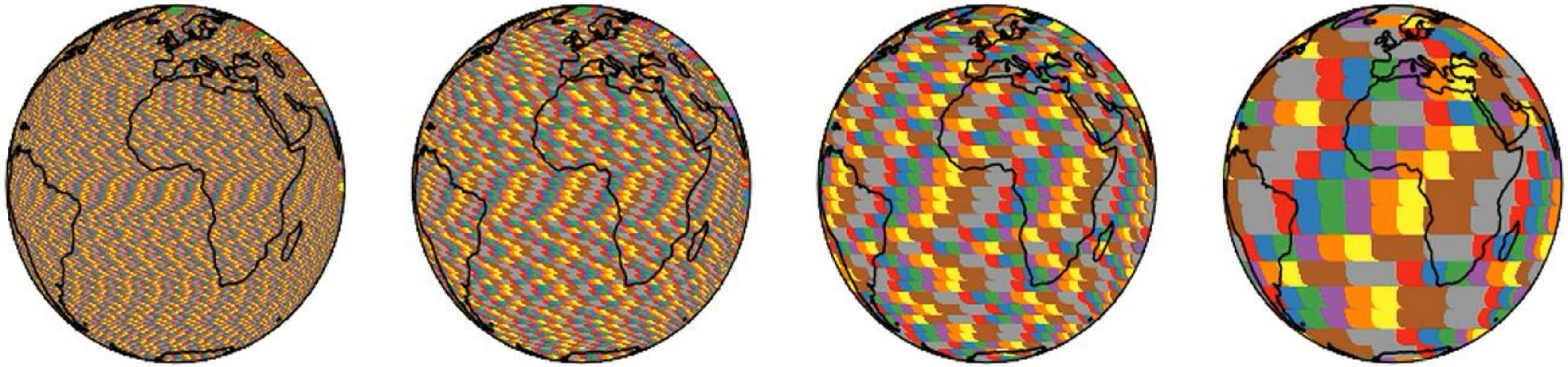
Climate Change

Metrics for measuring observation density

- WMO OSCAR and GCOS requirements:
  - Target horizontal resolution (as a distance, in km),
  - Target observation cycle (as a time difference, in hours)
  - Per variable, per application, and for several levels (threshold or target / breakthrough / goal)
- Looking at these, the following ‘key levels’ are retained:

example for surface pressure

~GCOS breakthrough *~GCOS target*



1 obs every 1° (111 km), every 3 hours

1 obs every 2° (222 km), every 6 hours

1 obs every 500 km, every 12 hours

1 obs every 1000 km, every day

➔ *We count as “1” each bin that has <u>at least</u> one observation. Otherwise, that bin counts as “0”.*

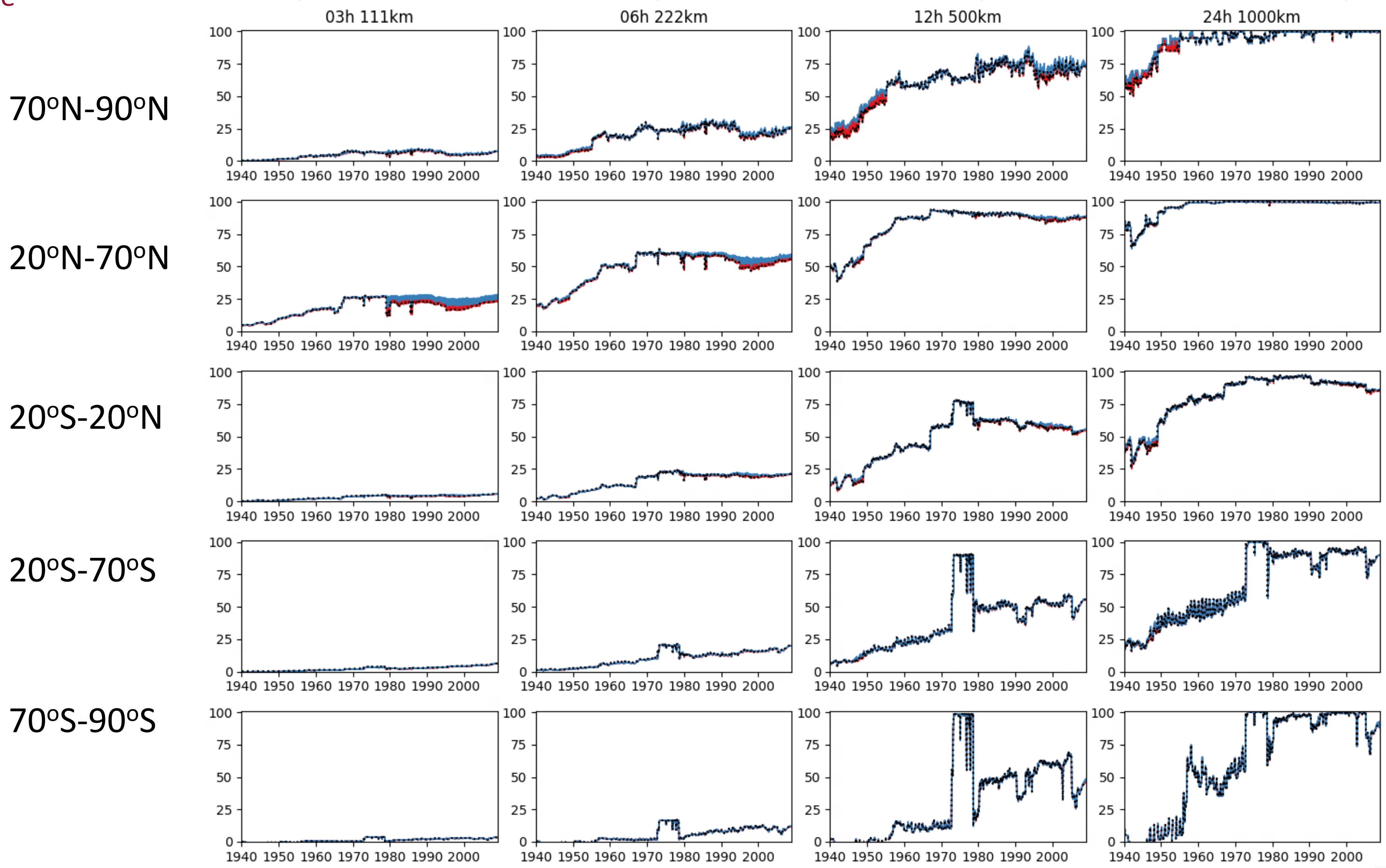
➔ *Doing this over all possible bins in a month yields **monthly percentage of coverage, for a given target resolution.***

Equal-area boxes (Leopardi, 2009)

Climate Change

Improvements in coverage: surface (or mean-sea-level) pressure obs. density

1 obs every 1° (111 km), every 3 hours | 1 obs every 2° (222 km), every 6 hours | 1 obs every 500 km, every 12 hours | 1 obs every 1000 km, every day



🟦 Impact of adding ISPD v4.7 over land

🟥 Impact of adding C3S comprehensive surface dataset

. . . . . ERA5 (baseline)

Climate
Change

Improvements in coverage: 2-metre temperature obs. density

1 obs every 1° (111 km), every 3 hours

1 obs every 2° (222 km), every 6 hours

1 obs every 500 km, every 12 hours

1 obs every 1000 km, every day

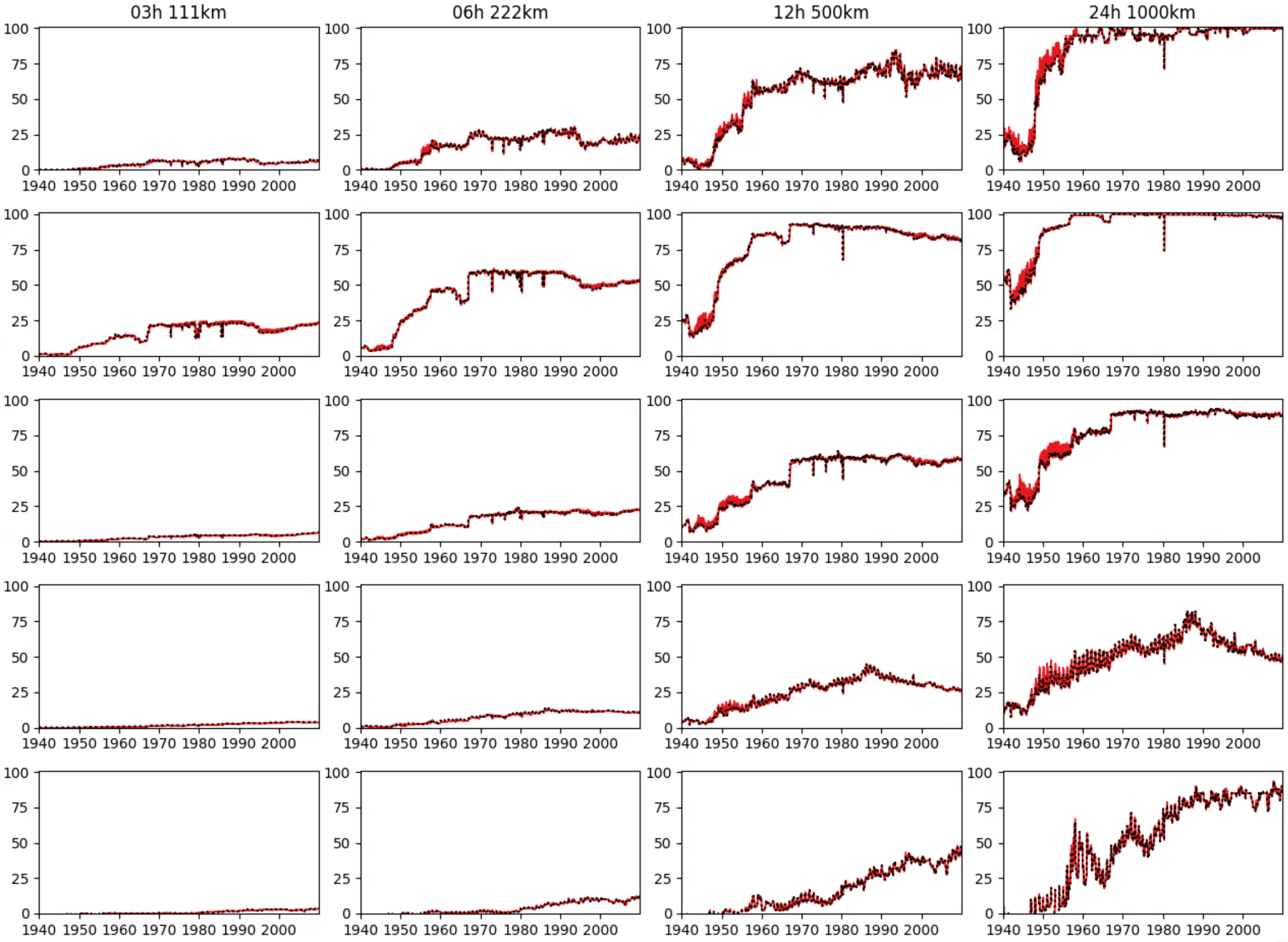
70°N-90°N

20°N-70°N

20°S-20°N

20°S-70°S

70°S-90°S



🟥 Impact of adding C3S comprehensive surface dataset

. . . . . ERA5 (baseline)

Climate Change

Improvements in coverage: 2-metre dew point temperature obs. density

1 obs every 1° (111 km), every 3 hours | 1 obs every 2° (222 km), every 6 hours | 1 obs every 500 km, every 12 hours | 1 obs every 1000 km, every day

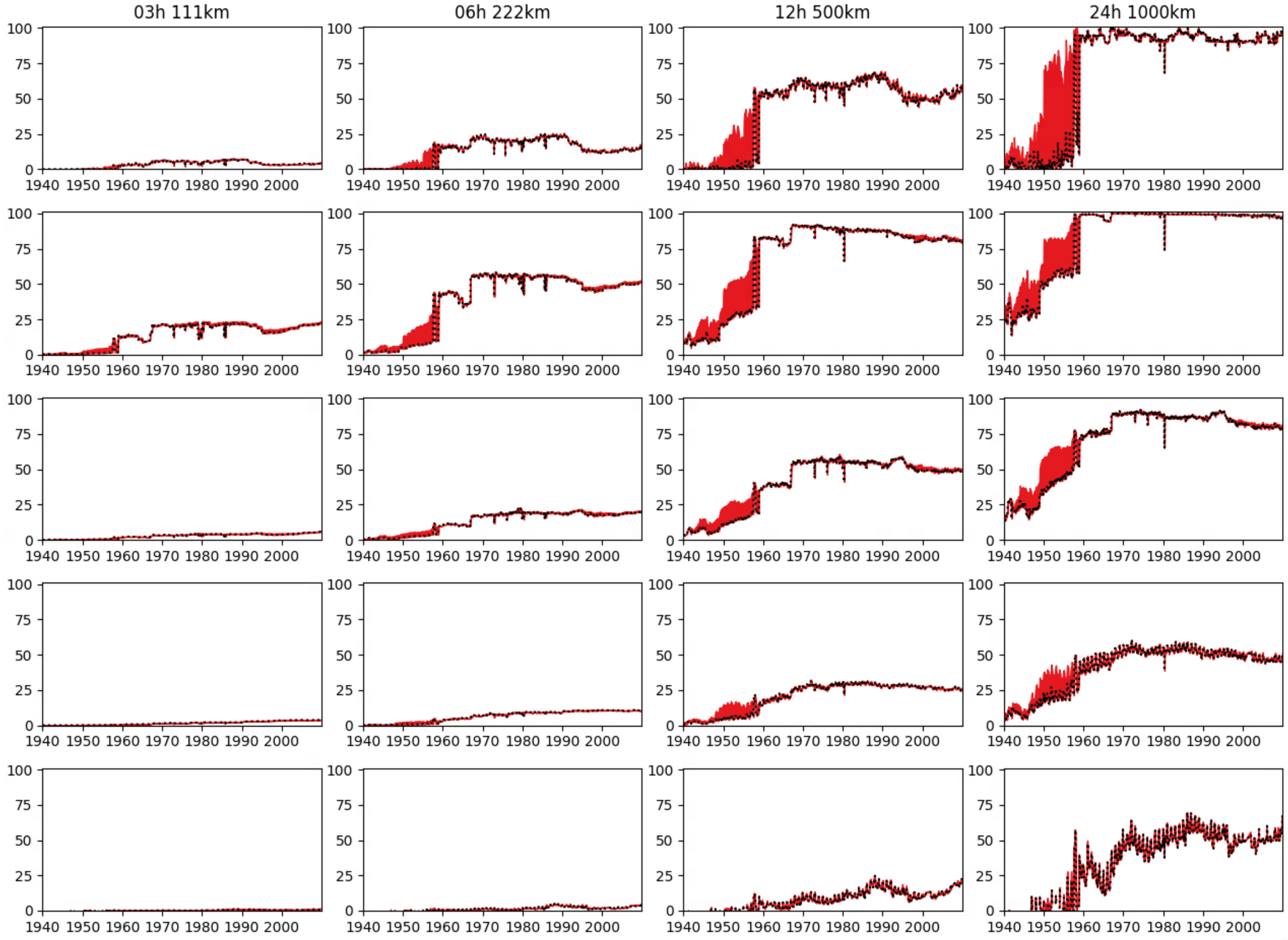
70°N-90°N

20°N-70°N

20°S-20°N

20°S-70°S

70°S-90°S

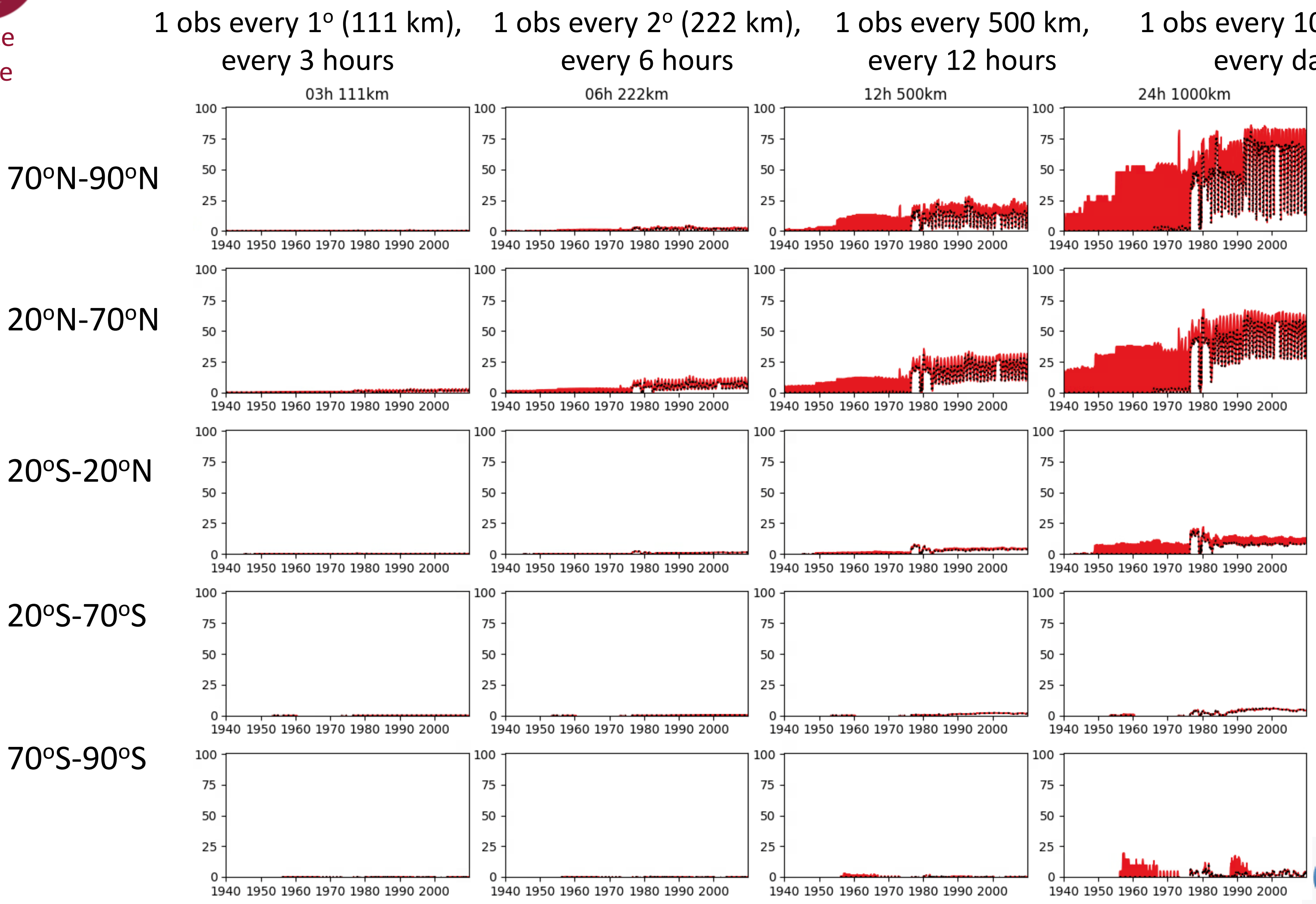


🟥 Impact of adding C3S comprehensive surface dataset

. . . . . ERA5 (baseline)

Climate Change

Improvements in coverage: snow depth obs. density

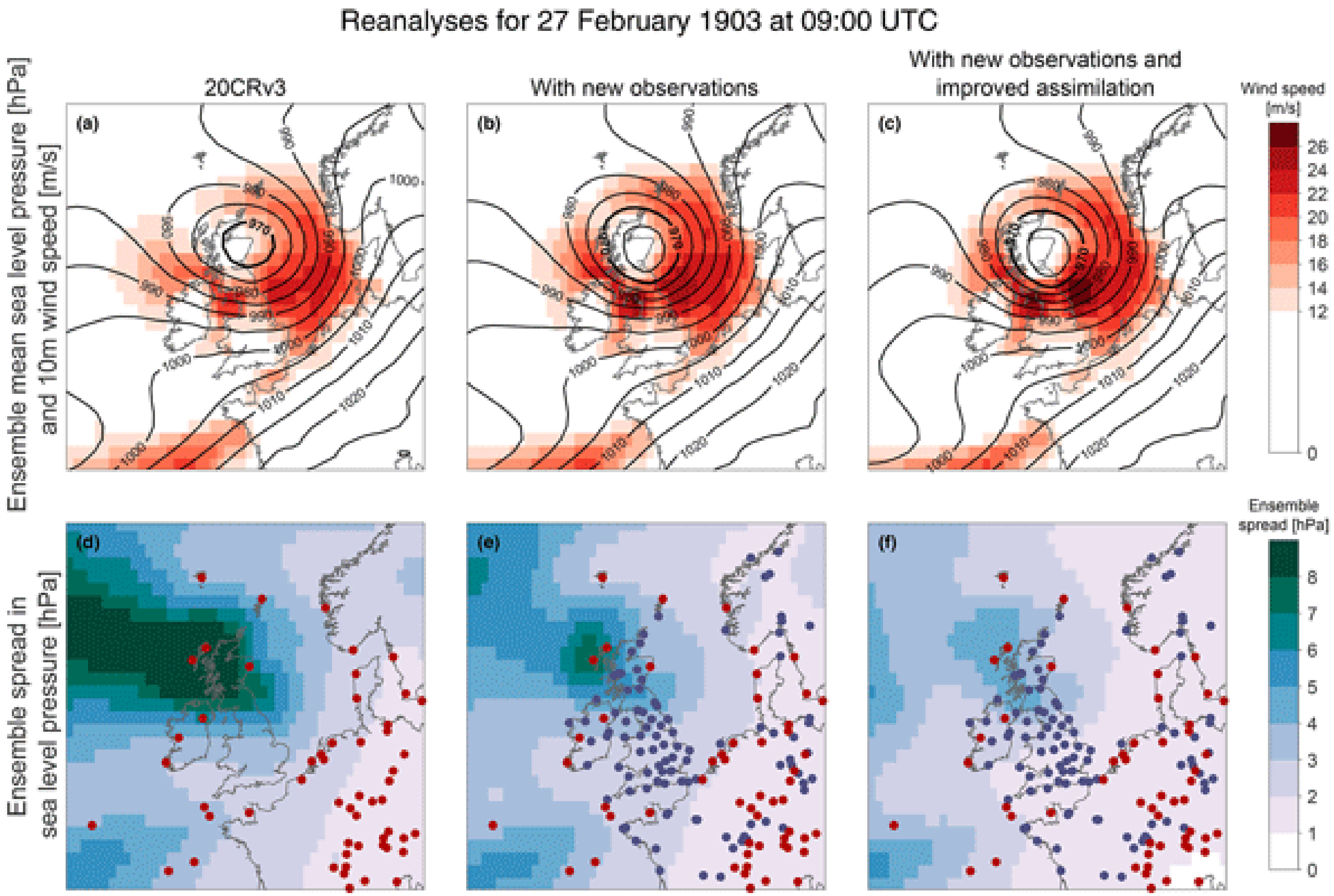


🟥 Impact of adding C3S comprehensive surface dataset

. . . . . ERA5 (baseline)

Climate Change

Going full circle: reanalyzing rescued observations for the Ulysses storm

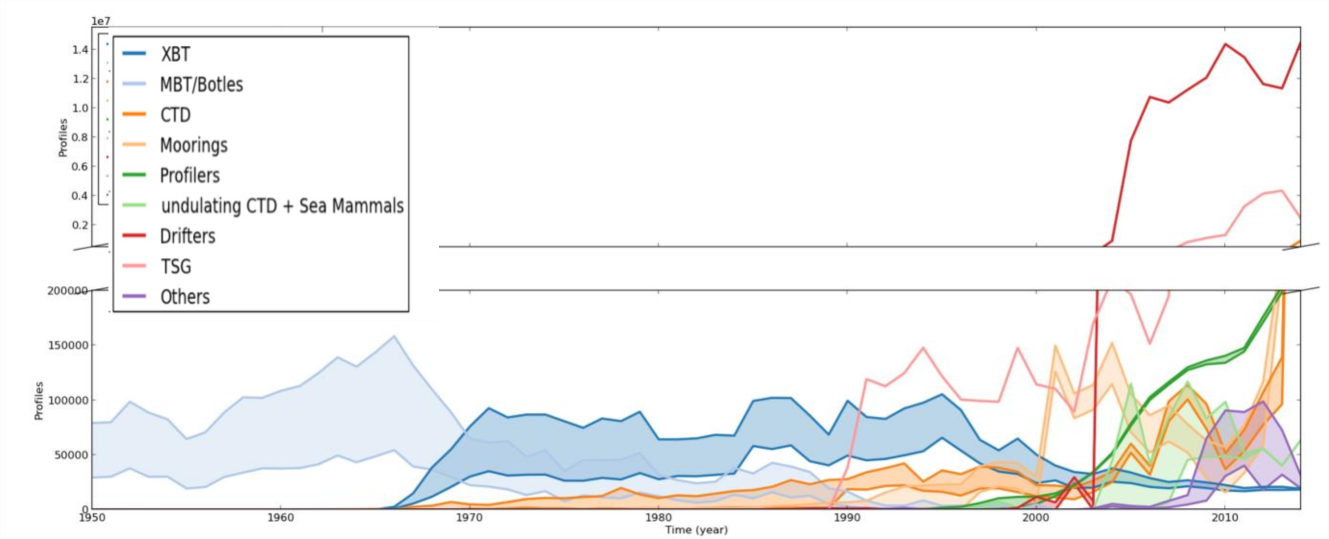


→ For more on this topic: Hawkins et al., 2023 [DOI:10.5194/nhess-23-1465-2023](https://doi.org/10.5194/nhess-23-1465-2023)

Climate Change

Merging archives: more complex than it seems, but well worth the effort

2018 merged datasets: CORA and EN4



+1.4 million XBT	+80 000 profilers	+2.0 million MBT/bottles
+72 000 moorings	+600 000 CTD	+ 80 000 drifters
+36 000 Sea mammals	+45 000 scanfish	

Lower line show CORA alone, upper line show CORA + EN4 (after CORA data removed). Shaded areas show the improvement in terms of number of profiles

N.B. data is cleaned from duplicates

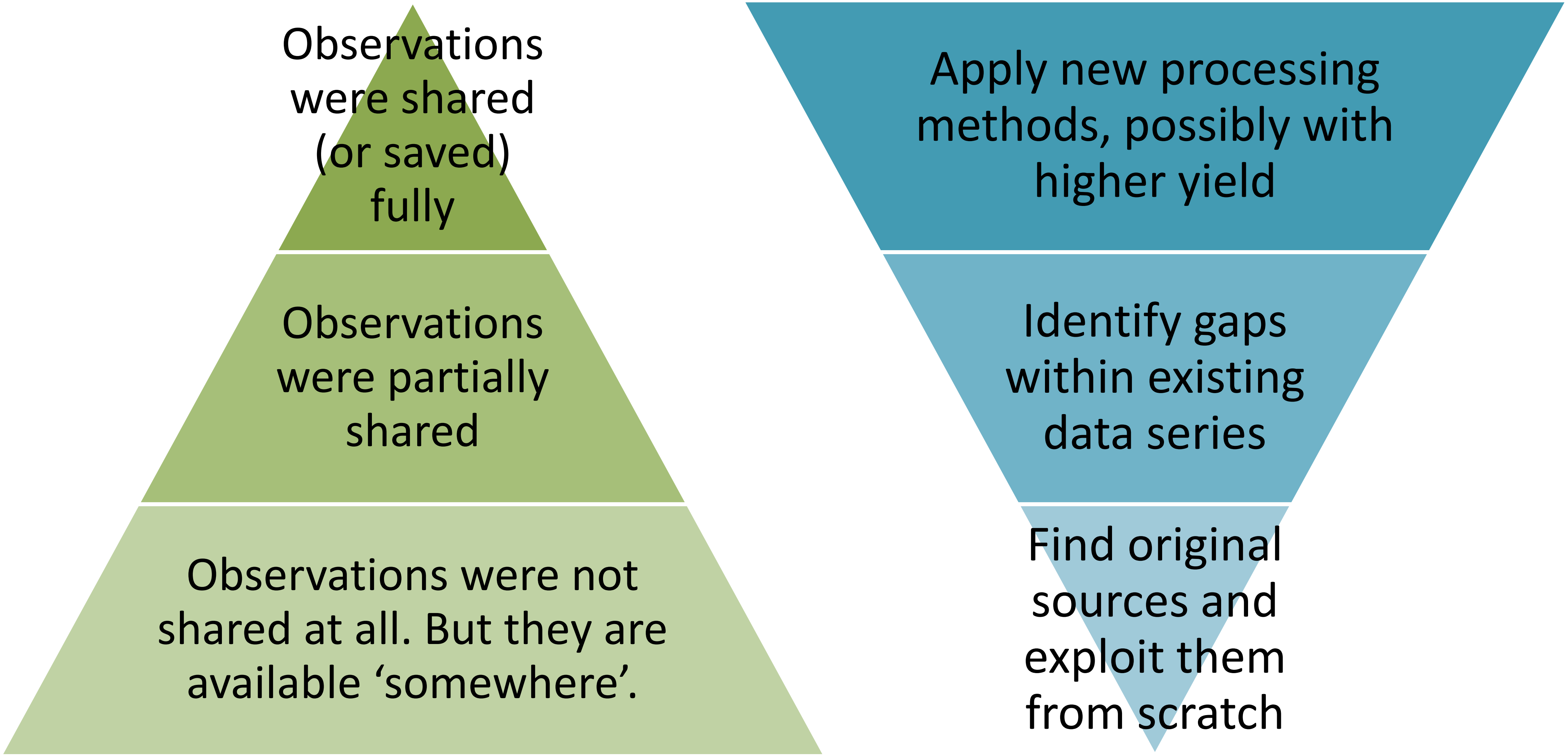
DOI:10.17882/46219

Credits: Copernicus Marine Environment Monitoring Service

Co-funded by the Horizon 2020 Framework Programme of the European Union

Part III.2: Processes to obtain more (better) observations

“Data reprocessing”



Observations were shared (or saved) fully

Observations were partially shared

Observations were not shared at all. But they are available ‘somewhere’.

Apply new processing methods, possibly with higher yield

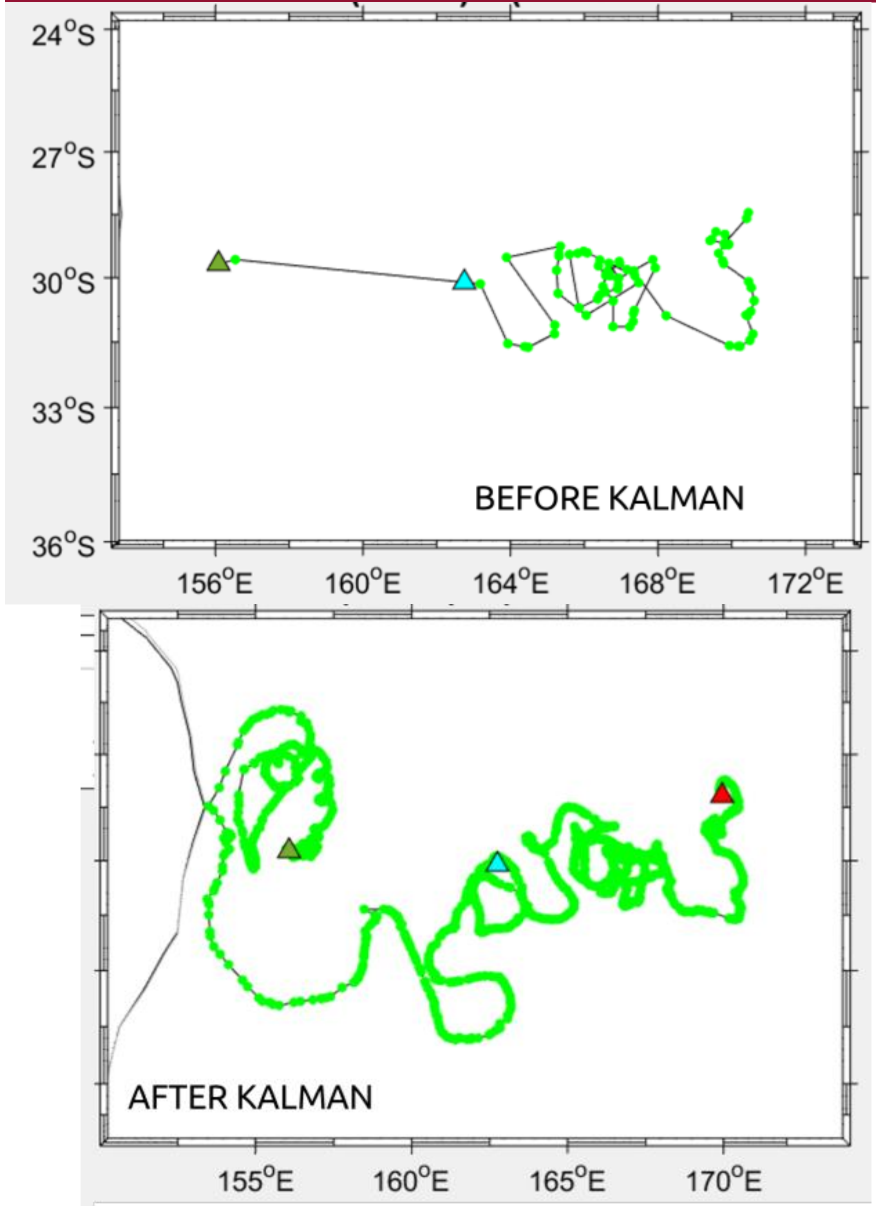
Identify gaps within existing data series

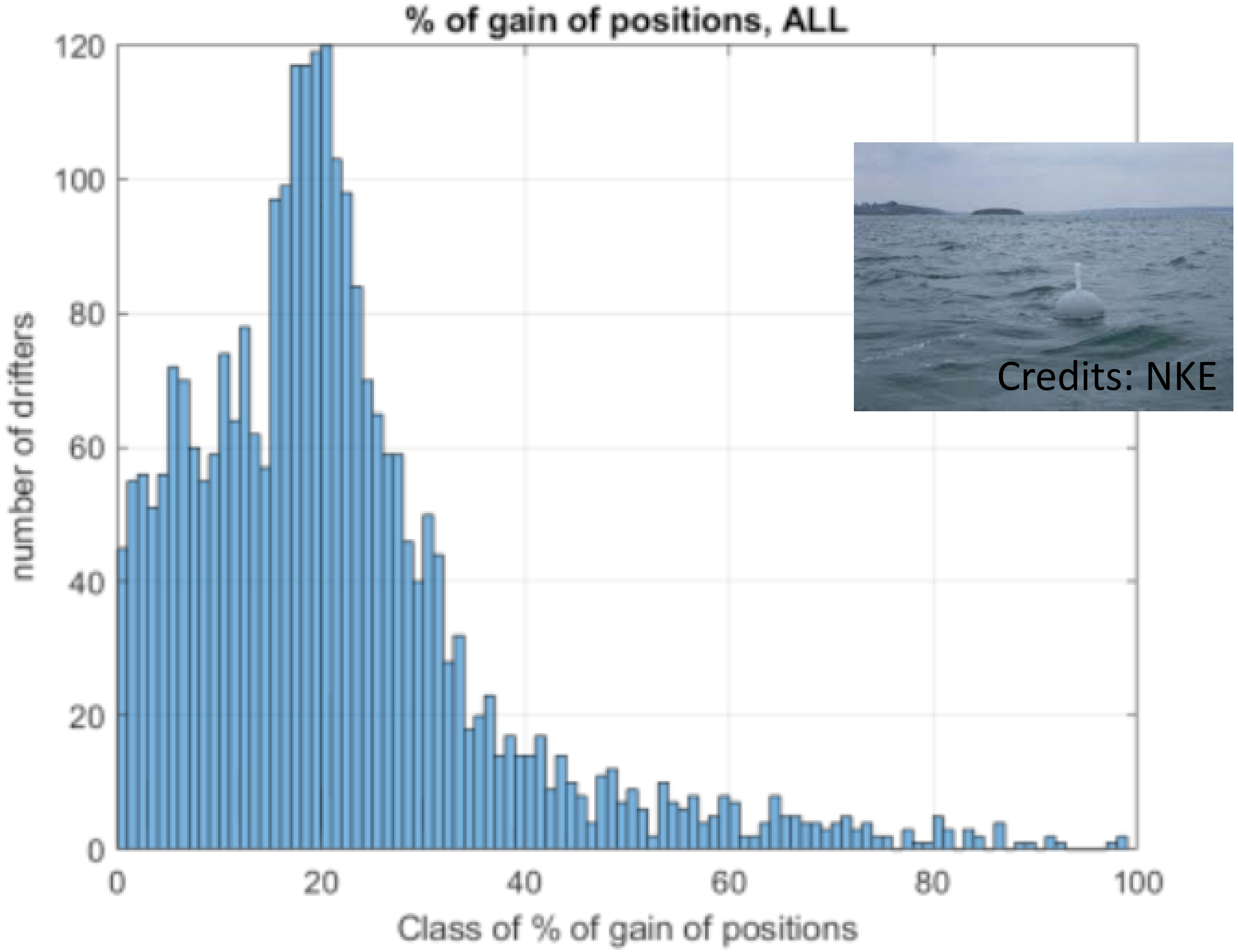
Find original sources and exploit them from scratch

“Data rescue”

Climate Change

Improving geolocation quality: reprocessing drifting buoy trajectories

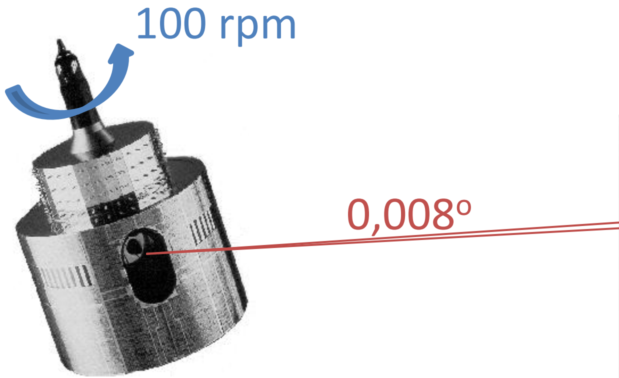




Credits: NKE

Credits: Rodriguez and Rannou, 2022. https://doi.org/10.13155/92124 Copernicus In Situ EEA C-RAID project

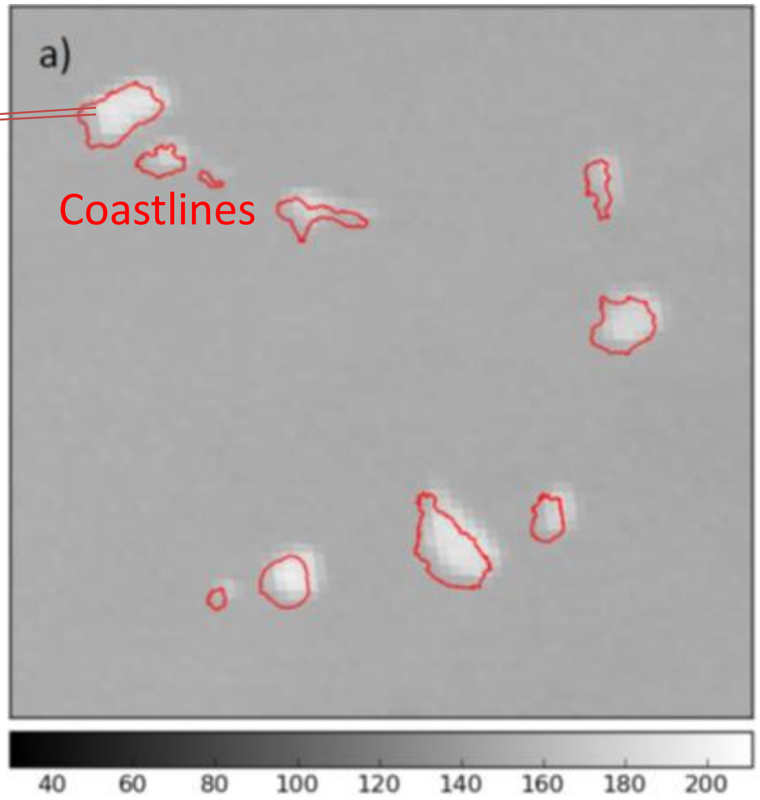
Improving geolocation quality: satellite



European Space Operations Centre in 1978 ([credits: ESA](#))

Credits: EUMETSAT (C3S)

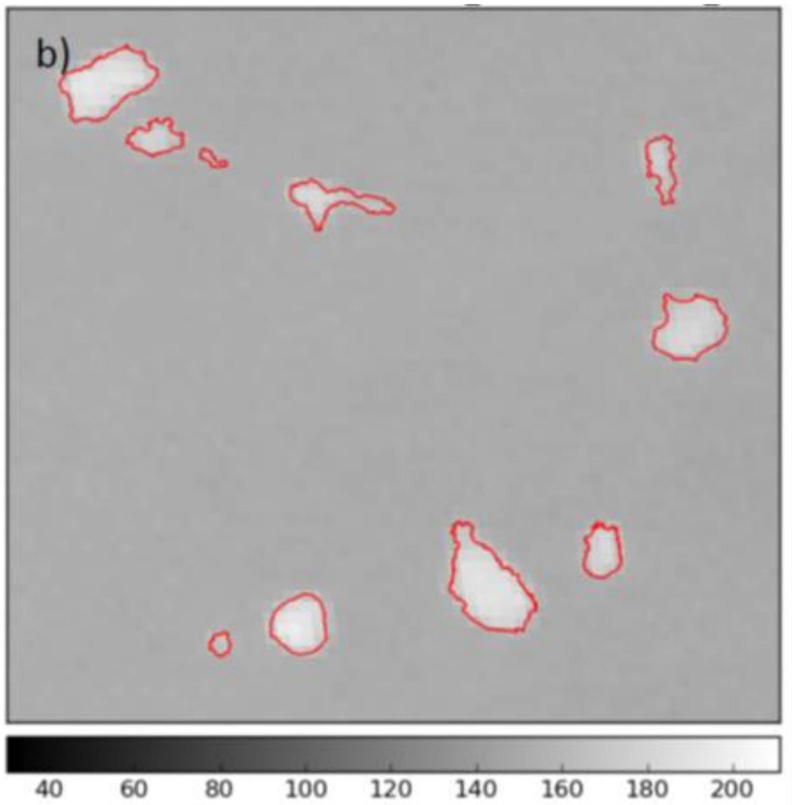
Original data



Sensor count (0-254; 255=missing)

Meteosat-2, 13 June 1985

Reprocessed



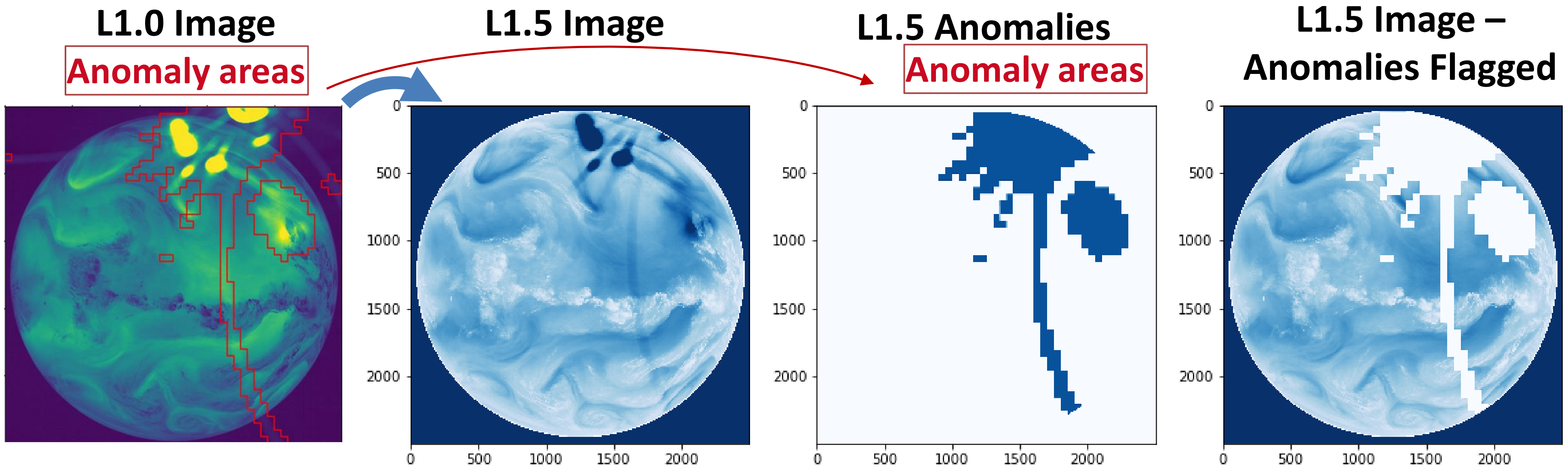
➔ Differences on the order of 1-2 pixels, i.e. 5-10 km

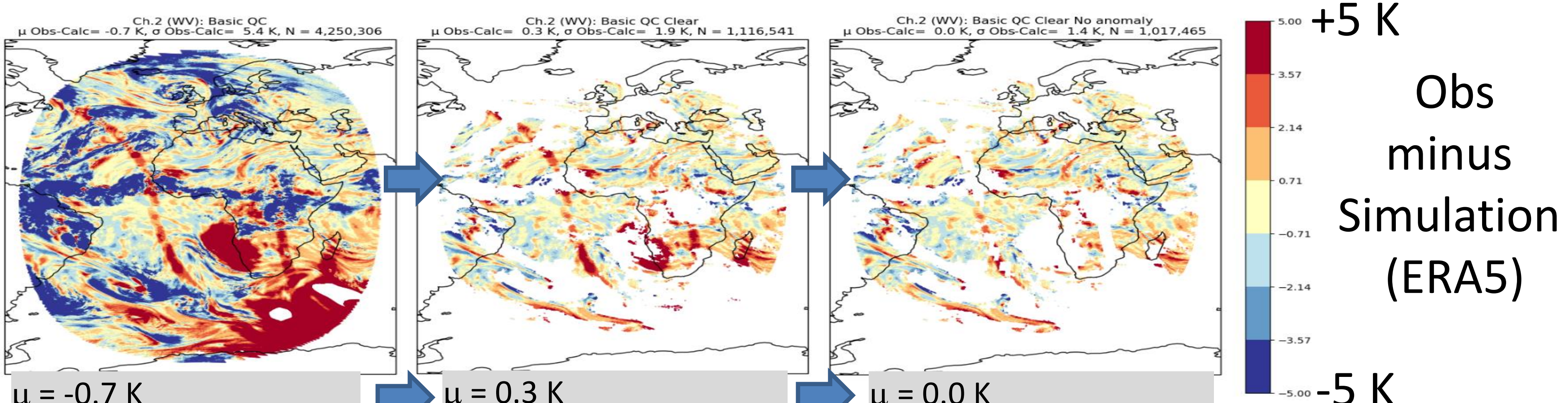
Climate Change

Improving data quality: satellite radiances

MET5
1996/10/16
00 UTC

WV channel





μ = -0.7 K
σ = 5.4 K (4.3 M pixels)

μ = 0.3 K
σ = 1.9 K (1.1 M pixels)

μ = 0.0 K
σ = 1.4 K (1.0 M pixels)

Leave out cloudy areas

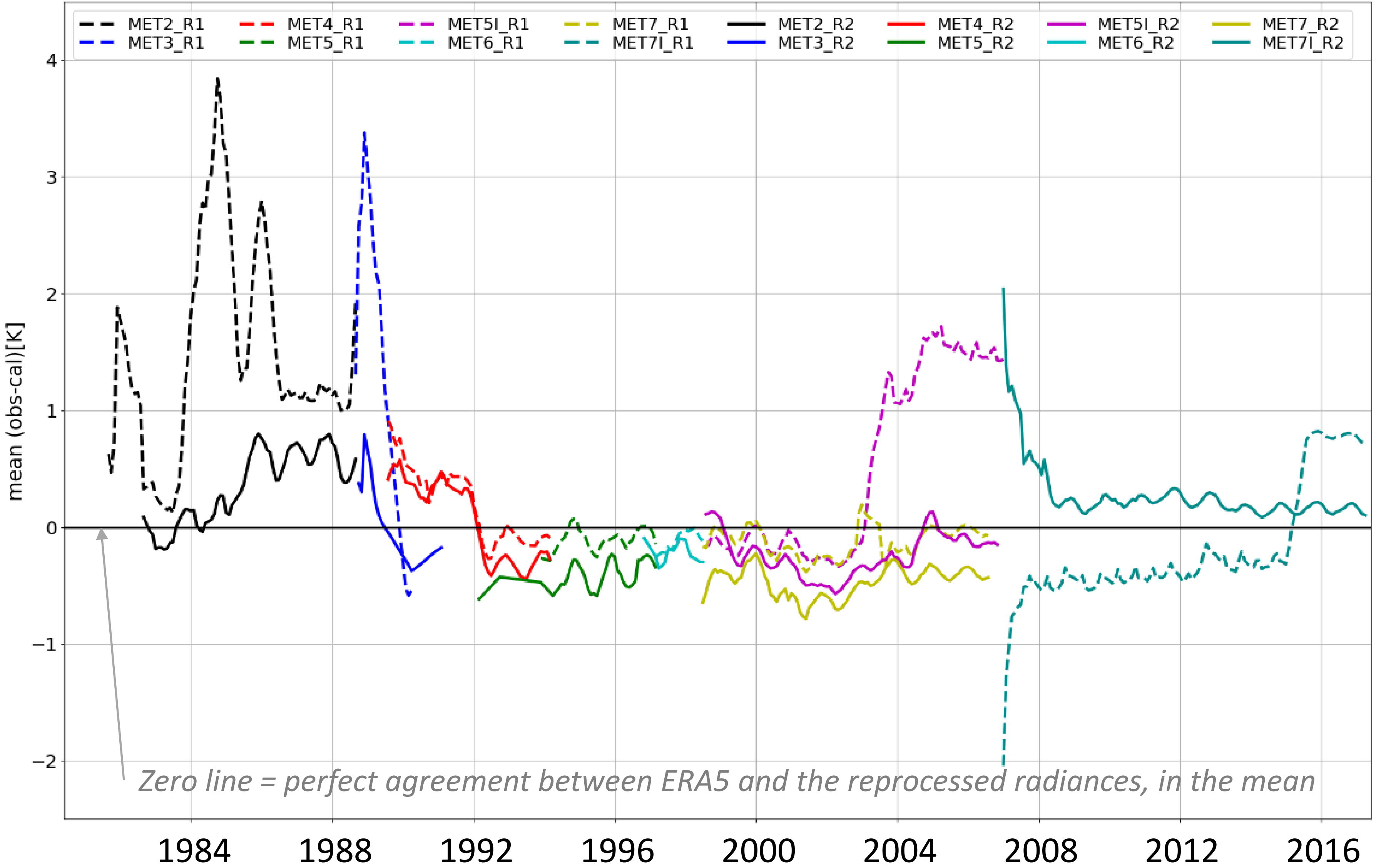
Leave out anomaly-flagged areas

→ For more on this topic: Poli al., 2023 DOI:10.22541/essoar.167591063.34033446/v1

Credits: EUMETSAT (C3S)

Climate Change

Improving data quality: improvement between 2 major reprocessings



Zero line = perfect agreement between ERA5 and the reprocessed radiances, in the mean

-- 1st Reproc.

– 2nd Reproc.

3 factors contributing to improvement:

✓*Geolocation*

✓*Handling image anomalies*

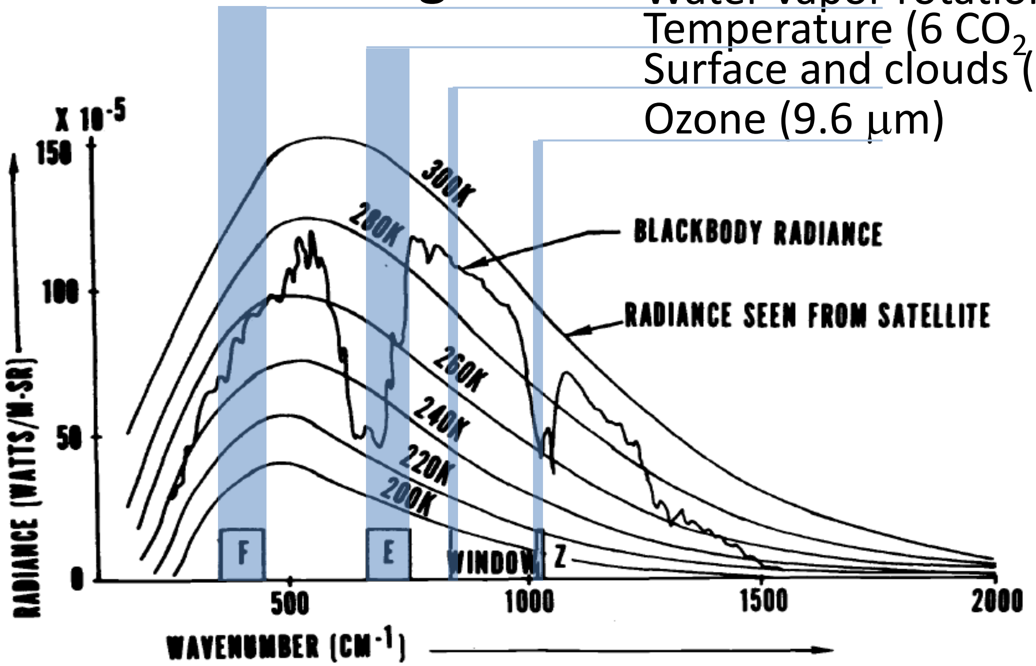
✓*Calibration*

Meteosat water vapor channel, observations compared with radiative transfer simulations using ERA5

Credits: EUMETSAT (C3S)

Recovering data from a ‘forgotten’ instrument: DMSP SSH

- Example with the Special Sensor H, also known as Multichannel Filter Radiometer (MFR), flown on 4x DMSP block 5D-1 satellites F-1 to F-4 (1976-1980)
- Important sensor to potentially improve reanalyses in the late 1970s
- Issue: the data documentation is inconsistent about channel ordering



4 IR regions

Water vapor rotational transitions (8 ch.)

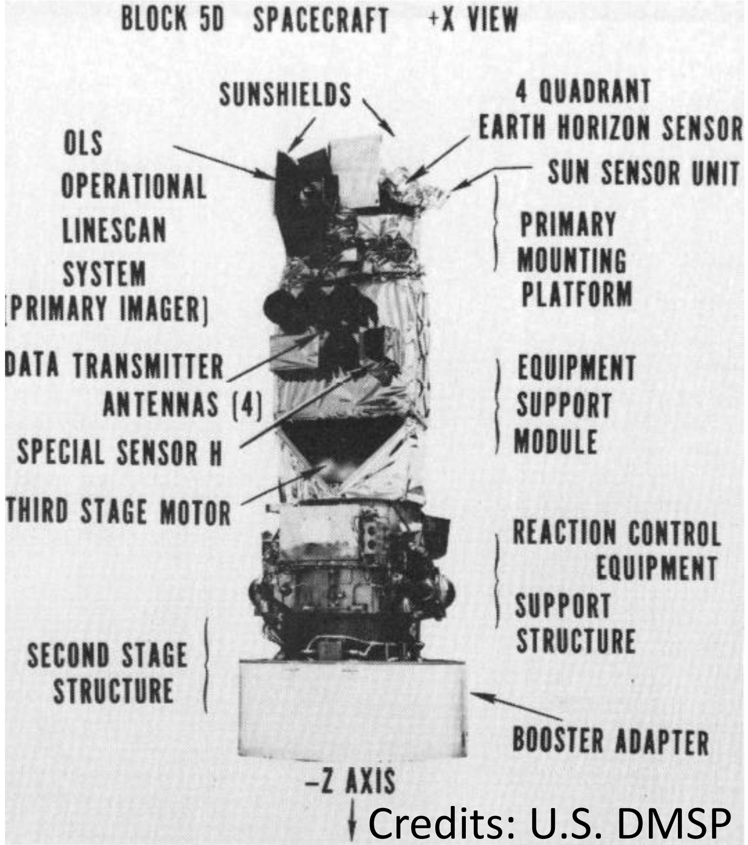
Temperature (6 CO_2 channels)

Surface and clouds (window channel)

Ozone (9.6 μm)

Fig. 1. Atmospheric Absorption Bands.

Excerpt from Nichols (1975)



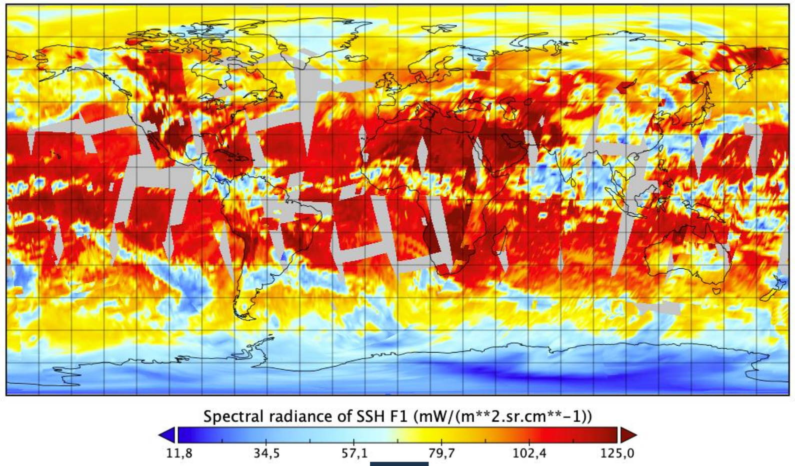
Credits: U.S. DMSP

Later instrument: SSH-2, but **no data source known** at present.

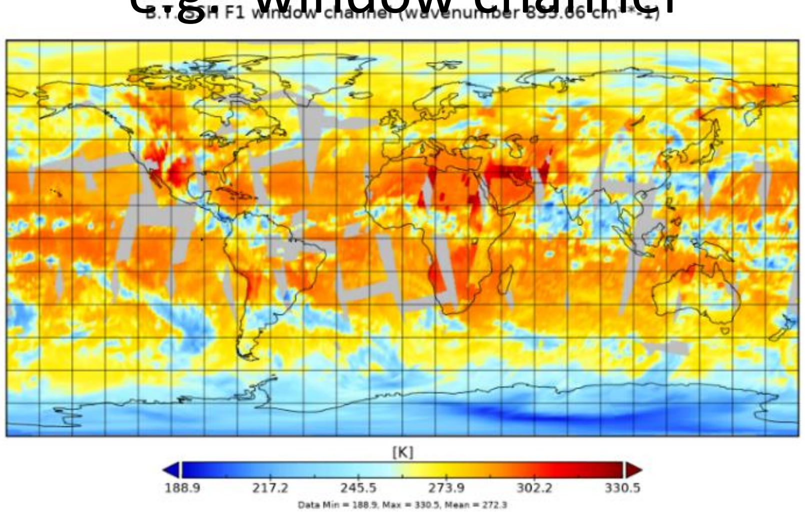
DMSP SSH sampling: F1 1977/06/20

Data encoded with FRAMIS database running on VAX 11/780 computer with a VMS operating system. NASA restored the data from ageing tapes. Acknoweldgements to the Nuclear Energy Agency (NEA) of the Organisation for Economic Co-operation and Development (OECD) for having kept a copy of this code.

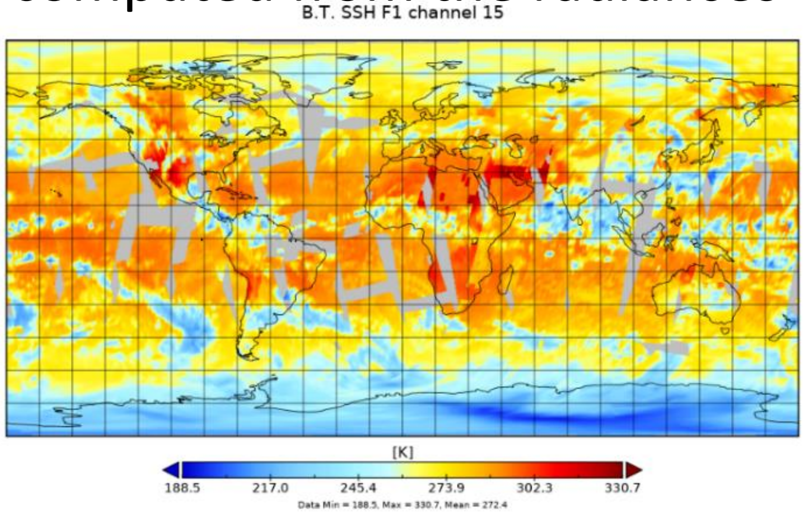
Radiance in the original data (for all channels)



B.T. in the original data, but for a few channels <u>only</u>, e.g. window channel



B.T. computed from the radiances



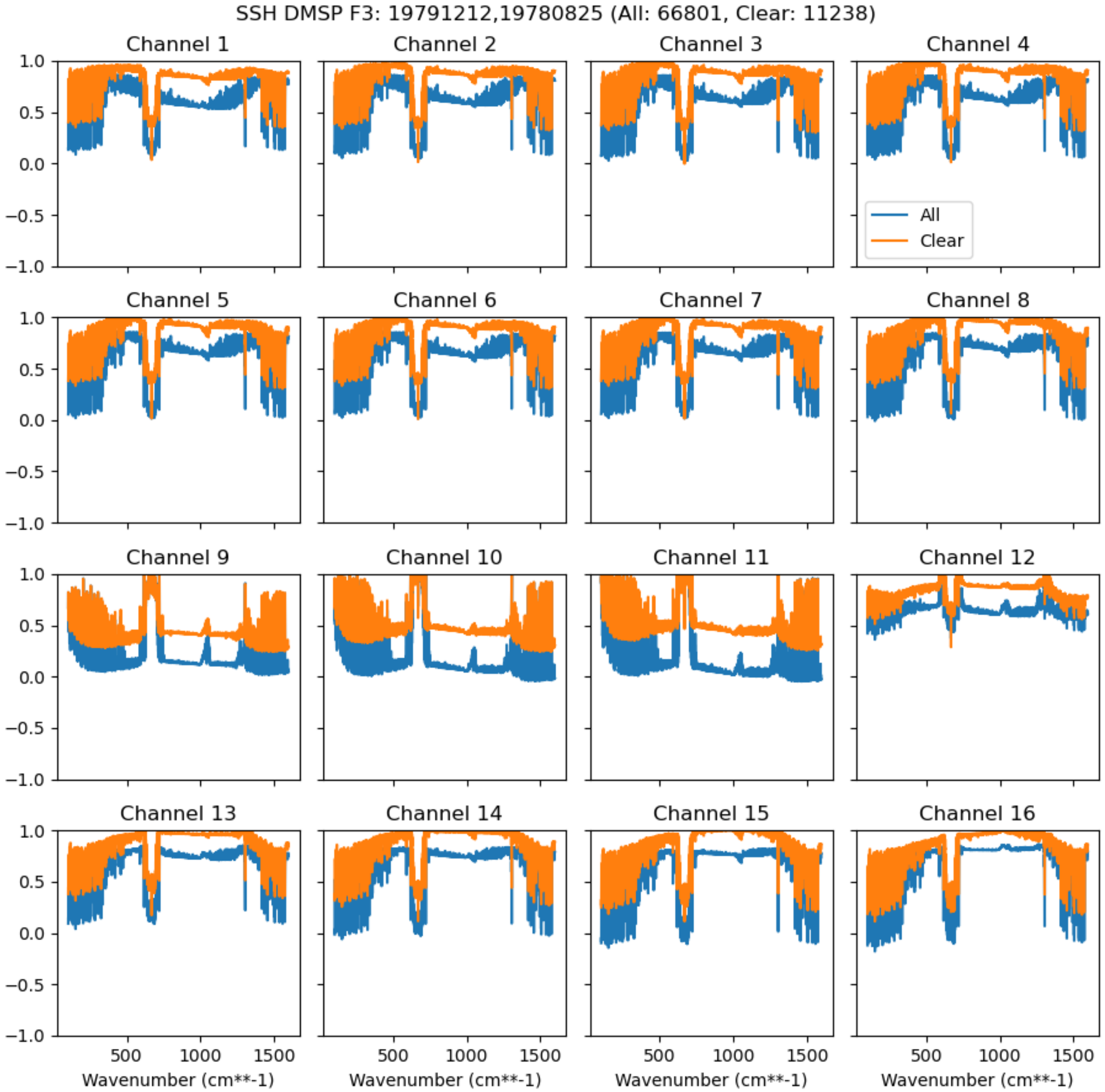
→ We are positive this is the window channel for cloud detection

Credits: NASA GES DISC

Credits: EUMETSAT (C3S2_310)

Climate
Change

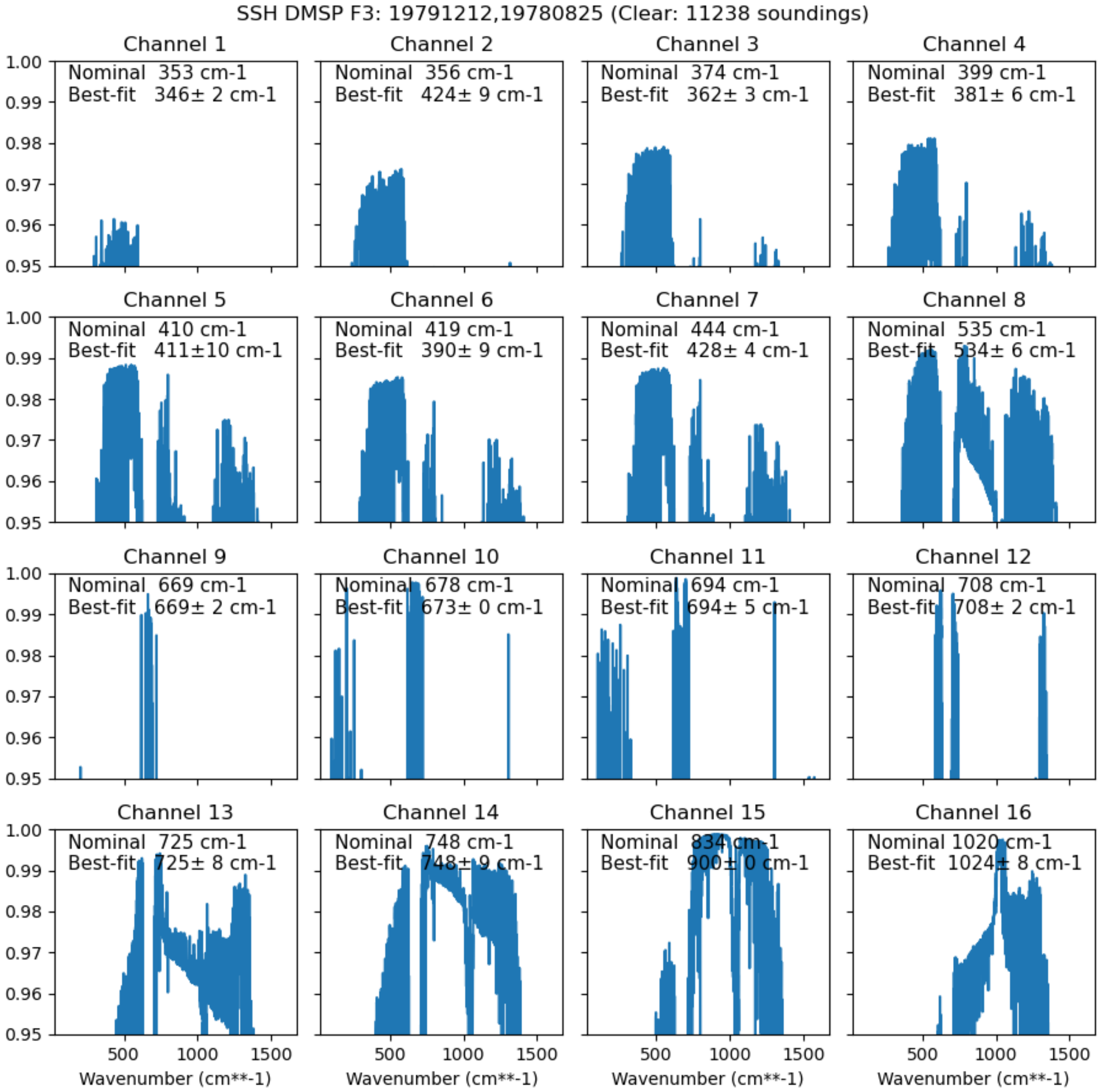
Using ERA5, simulation of the FORUM instrument at 0.3 cm^{-1} resolution



- Correlate the <u>simulated radiances</u> with the <u>observed radiances</u>
  - For each simulated FORUM channel (100-1600 cm^{-1})
  - Results in blue
- Repeat this after applying a window check for cloud detection
  - Obs - Calc in the range [-2K,3K]
  - Results in orange

Credits: EUMETSAT (C3S2_310)

Zoom on high correlations [0.95 – 1.00]



We may then figure out a “most likely” wavenumber

Then, given the limited number of expected wavenumbers from documentation (16), we can find the most likely channel width by applying a moving average of varying width

Beware not to over-interpret the results, given biases in the R.T. simulations!

Credits: EUMETSAT (C3S)

Climate Change

SSH wavenumbers estimated from simulations

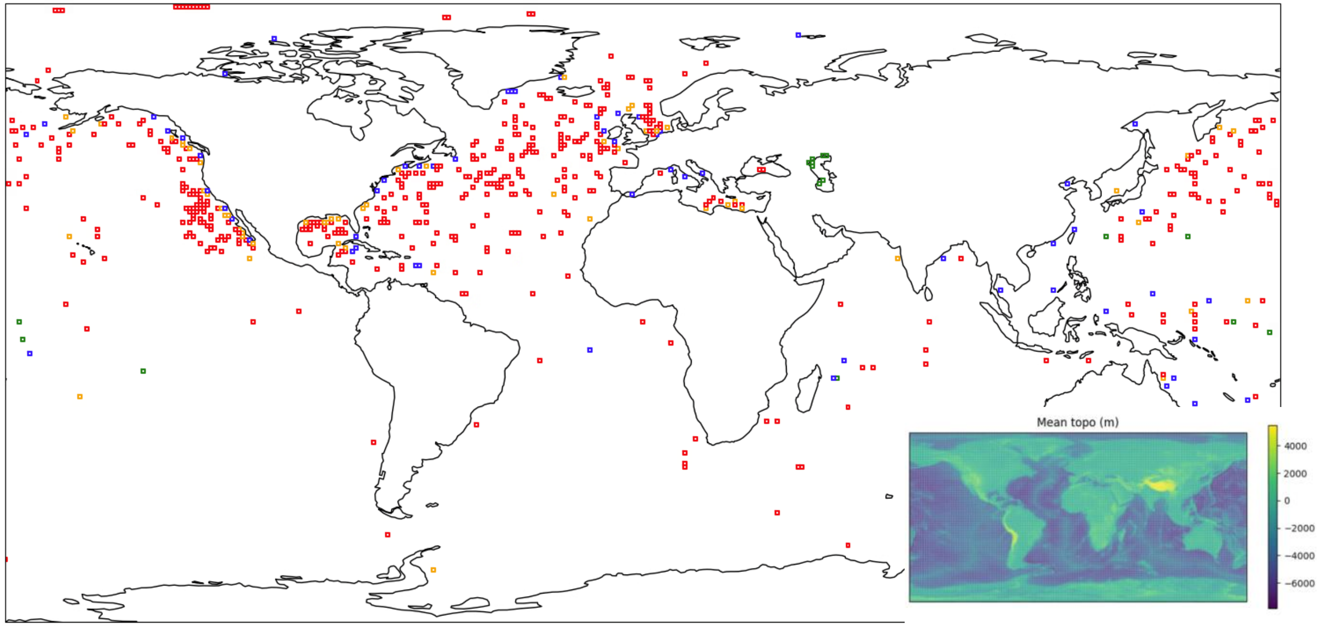
Nominal

Estimated, by comparing with FORUM simulations using ERA5

Satellite	F1		F2		F3		F4	
Channel	$\bar{\nu}_N$	$\bar{\nu}_E$	$\bar{\nu}_N$	$\bar{\nu}_E$	$\bar{\nu}_N$	$\bar{\nu}_E$	$\bar{\nu}_N$	$\bar{\nu}_E$
1	354	347± 4	355	346± 4	353	346± 2	353	346± 2
2	356	424±10	356	361± 6	356	424± 9	355	346± 4
3	373	380± 6	373	362± 8	374	362± 3	375	362± 4
4	398	390± 9	399	405± 1	399	381± 6	399	390±10
5	410	409±10	407	410±10	410	411±10	409	410±10
6	418	428± 4	419	390± 4	419	390± 9	419	405± 2
7	442	439± 5	442	428± 4	444	428± 4	442	428± 4
8	-	-	534	534± 5	535	534± 6	535	534± 6
9	669	669± 4	669	670± 1	669	669± 2	668	669± 2
10	678	673±10	678	680± 0	678	673± 0	679	673± 0
11	694	695± 3	694	694± 5	694	694± 5	694	694± 0
12	708	708± 0	707	707± 2	708	708± 2	707	707± 6
13	726	725± 9	724	724± 6	725	725± 8	724	724± 5
14	747	746± 0	746	746± 2	748	748± 9	749	749± 7
15	839	900± 6	835	900± 6	834	900± 0	835	900± 5
16	1020	1025± 6	1019	1020± 4	1020	1024± 8	1021	1021± 3
Dates	19770702, 19770421		19780825, 19790825		19780825, 19791212		19790825, 19791212	
All	90,418		96,649		66,801		116,358	
Clear	16,210		17,700		11,238		20,662	

Part III.3: Improving how the observations are used: in situ data location

- Station positions: latitude, longitude, elevation. Difficulties:
  1. Referentials have changed over time. Now it is rather stable: WGS84 for latitude, longitude. Still no unique referential for vertical (EGM96 or EGM2008…). However, differences get generally to be small (e.g., RMS 0.5 m between these two)
  2. Precision of encoding. Now WMO mandates 0.001 degrees for metadata but is has not always been the case.
  3. Stations have physically moved. Actual position changes were sometimes not propagated in the data at the same time.
  4. And all this is saying nothing of the installation of instrumentation itself
- Example: International Surface Pressure Databank v4.7: DOI:10.5065/9EYR-TY90, (Compo et al., 2019) Comparison of land stations' altitudes with ETOPO2022 DOI:10.25921/fd45-gt74 (NOAA NCEI, 2022): dataset combining elevation over land and bathymetry over ocean
  - Enables better to spot differences between stations altitudes (often near sea-level) and the bedrock bathymetry.
  - Suspicous land stations from 1940:



red = 466 locations with nearest coastline at least 50 km away;
orange = 58 locations with nearest coastline between 20 and 50 km away;
blue = 60 locations with nearest coastline within 20 km ;
green = 16 locations over a land mass but max. elevation reported to be at least 10 meters minus below sea-level)

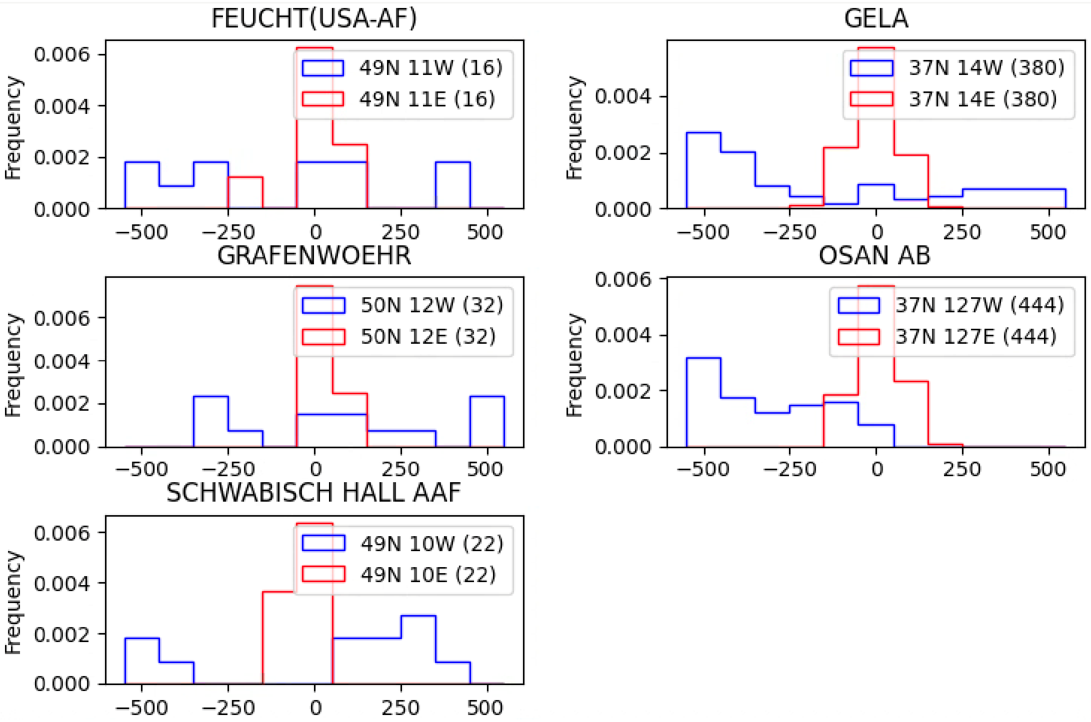
Looking at **each site**, one finds:

Confirmed locations of **moored buoys, met. ships, or ocean platforms** (~85)

Several stations sometimes positioned at **(0,0)** exactly, for a short part of the record

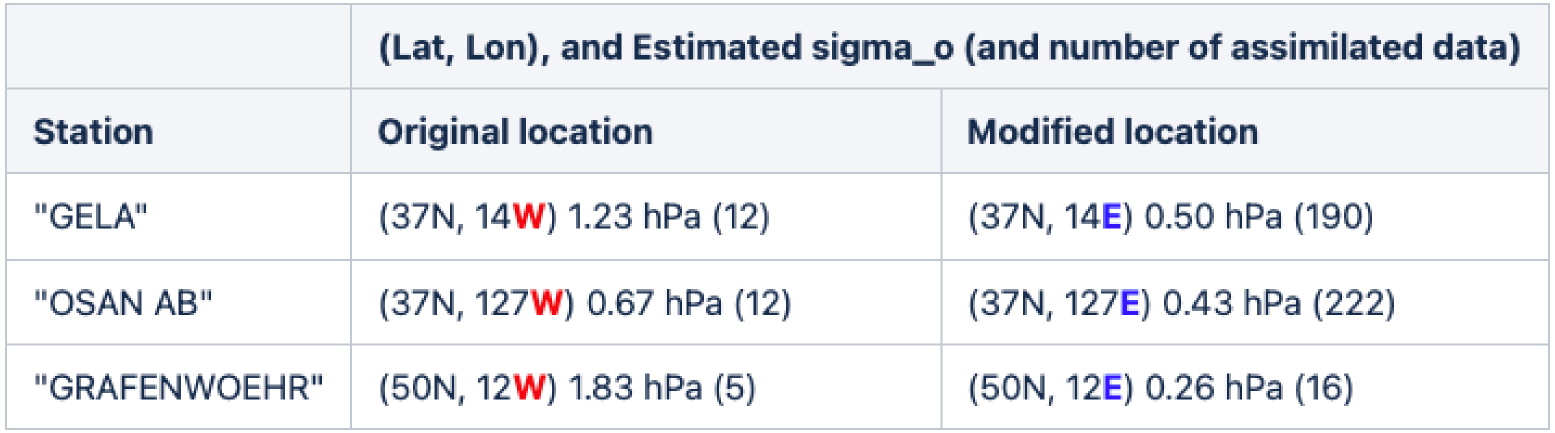
Stations with a matching name found by **swapping the sign of the latitude or longitude** (or both), or by **changing a single digit** in latitude or longitude (~55)

Confirmation of ‘better’ locations: Assimilation of wrong & corrected positions



Considering all data: Easy to see which one has the wrong *vs* corrected position…

ECMWF IFS, June 1980



	(Lat, Lon), and Estimated sigma_o (and number of assimilated data)	
Station	**Original location**	**Modified location**
"GELA"	(37N, 14W) 1.23 hPa (12)	(37N, 14E) 0.50 hPa (190)
"OSAN AB"	(37N, 127W) 0.67 hPa (12)	(37N, 127E) 0.43 hPa (222)
"GRAFENWOEHR"	(50N, 12W) 1.83 hPa (5)	(50N, 12E) 0.26 hPa (16)

Wrong position:

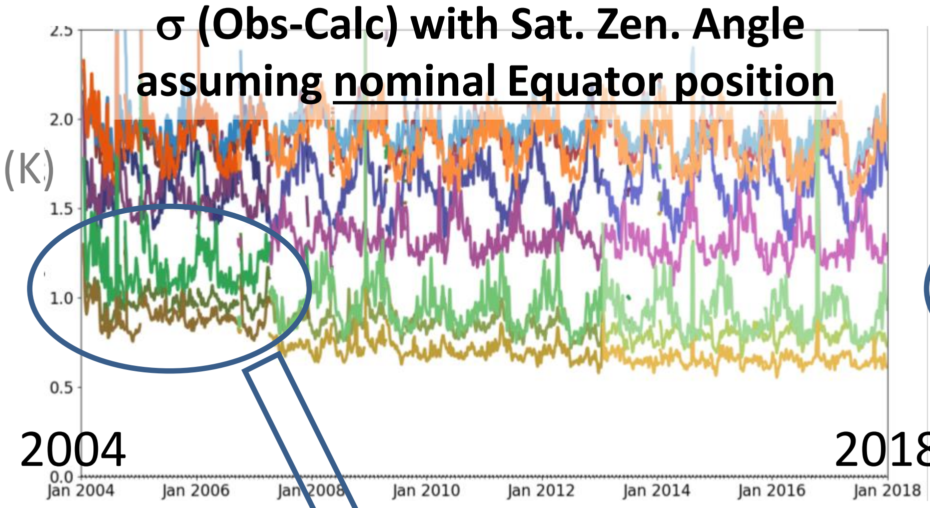
- ➔ Many data rejected
- ➔ Still quite a few data assimilated!

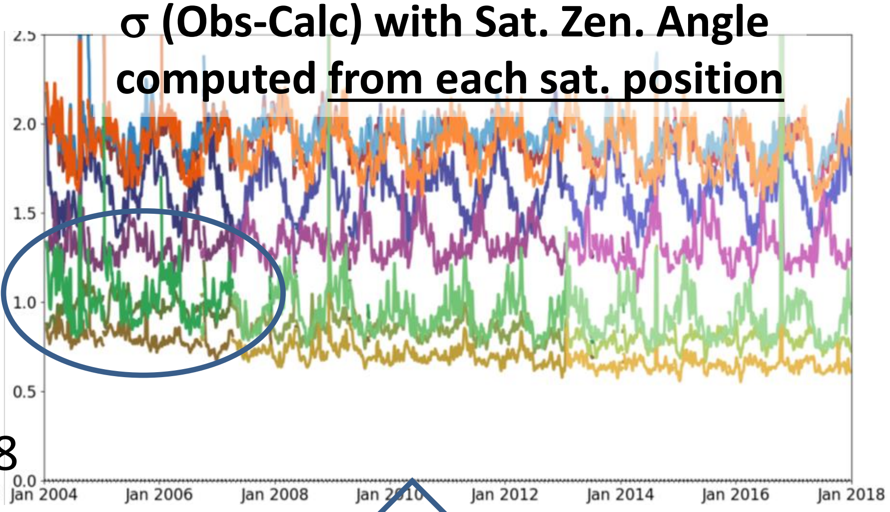
Corrected position:

- ➔ More data assimilated
- ➔ Reduced sigma_o estimate

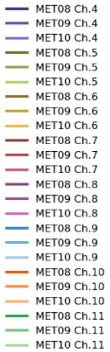
Climate Change

Using the ‘right’ satellite viewing angle: example with Meteosat Sec. Generation

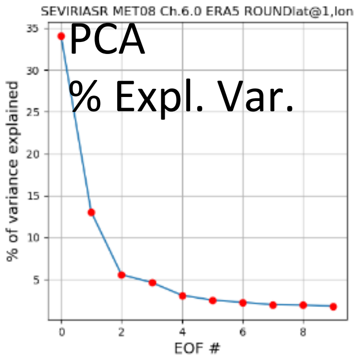

σ (Obs-Calc) with Sat. Zen. Angle assuming nominal Equator position


σ (Obs-Calc) with Sat. Zen. Angle computed from each sat. position

→ For more on this topic: Poli al., 2023 DOI:10.22541/essoar.167591063.34033446/v1

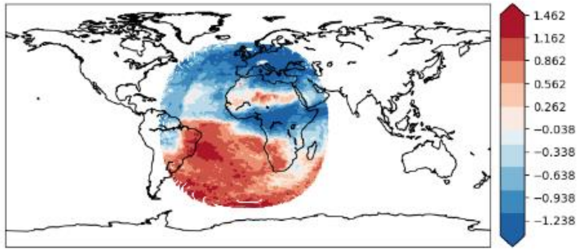


IR 3.9 μm (night only)
IR 6.2 μm
IR 7.3 μm
IR 8.7 μm
IR 9.7 μm
IR 10.8 μm
IR 12.0 μm
IR 13.4 μm

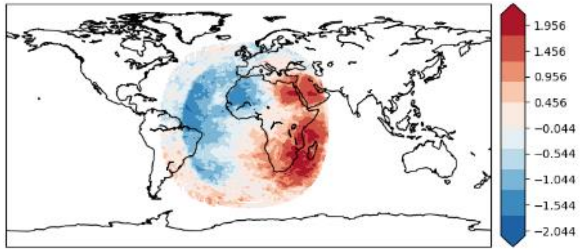

PCA

% Expl. Var.

PCA EOF#1



PCA EOF#2

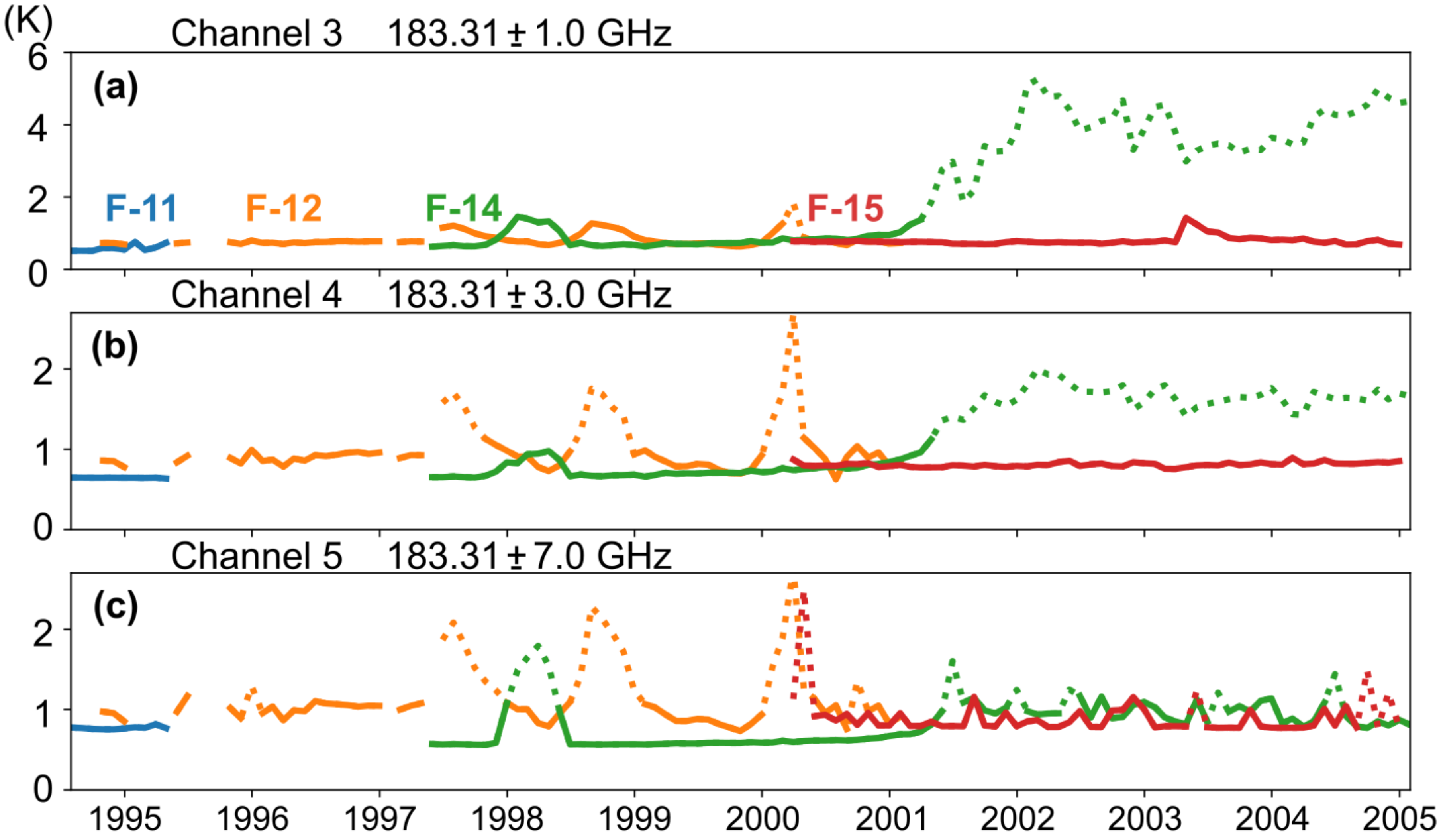


If the changes in the satellite viewing angle are ‘large enough’ to be visible (and explained) in differences between simulations, then it will quite likely propagate into regional applications - unless it is corrected, or accounted for, somehow…

Credits: EUMETSAT (C3S)

Climate Change

Using uncertainty information to avoid 'bad' time periods



Data source: DOI:10.15770/EUM_SEC_CLM_0050

→ For more on this topic: Poli al., 2023 DOI:10.22541/essoar.167591063.34033446/v1

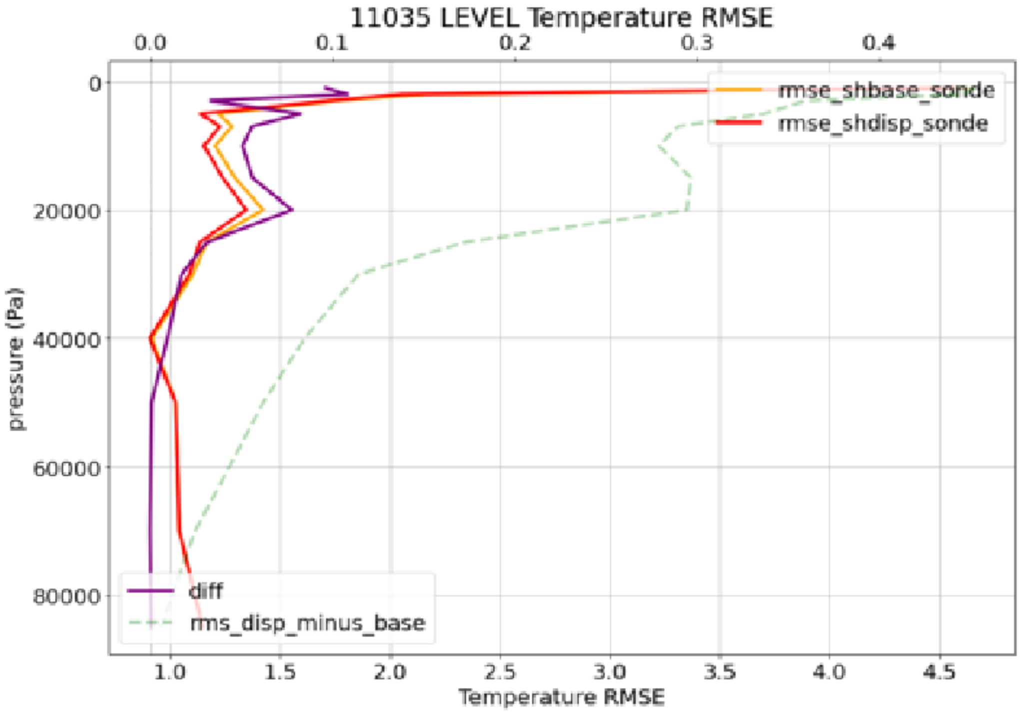
Root sum square of uncertainties estimated using FIDUCEO methodology. Compared to instrument specifications.

Credits: EUMETSAT (C3S)

Climate
Change

Using *better* ‘old’ observations, with ‘new’ methods: balloon drift

Year 1970

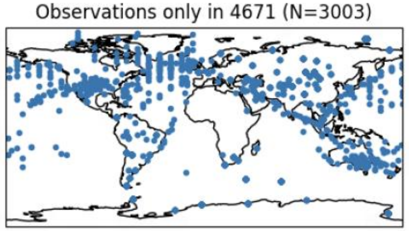


Root Mean Square Error (RMSE) between sonde temperatures and ERA5 (K). The difference between **non-displaced (or base) sonde statistics (in orange)** and **displaced-sonde statistics (in red)** is shown in purple. The **positive values indicated by the purple line** correspond to improvements, i.e., higher consistency between the sonde observations and ERA5.
Data from Vienna Hohe Warte, Austria.

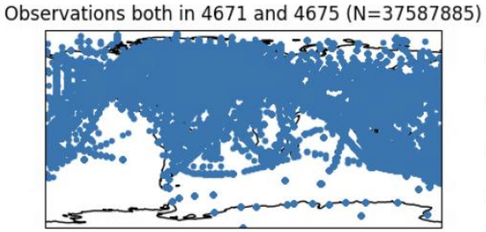
Assimilation trial June-July 1980

Assimilated upper air temperature observations (RS or aircraft)

Data <u>only</u> in control experiment >>

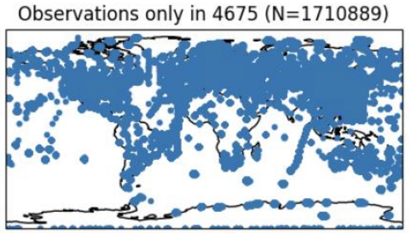


Data in <u>both</u> >>



Estimated σ_0	Control	Balloon drift
RS LAND (*N*=36.3 M)	**1.3 K**	**1.2 K**
RS SHIP (*N*=1.1 M)	**1.3 K**	**1.1 K**

Data <u>only</u> in ‘balloon drift’ experiment >>

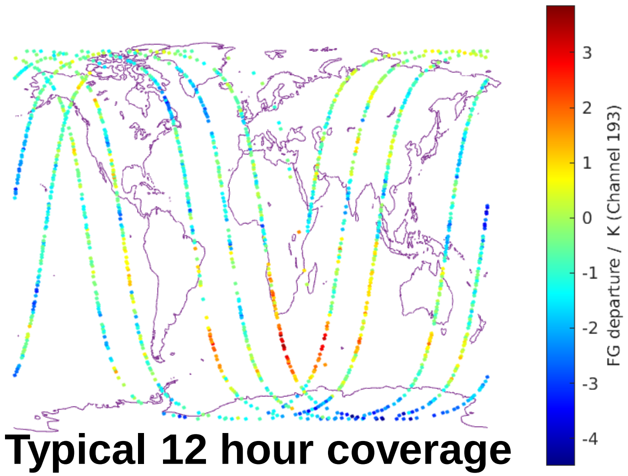


Desroziers et al. (2005), [DOI:10.1256/qj.05.108]

Credits: Haimberger, Voggenberger, Ambrogi, 2023. UNIVIE (C3S)

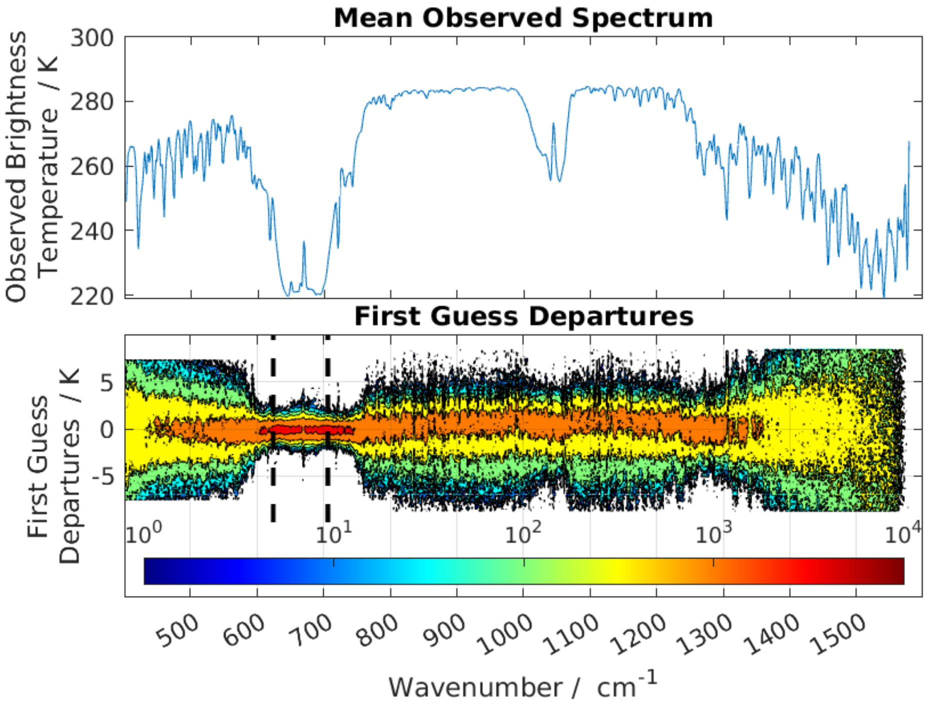
See also [webinar by Ingleby]

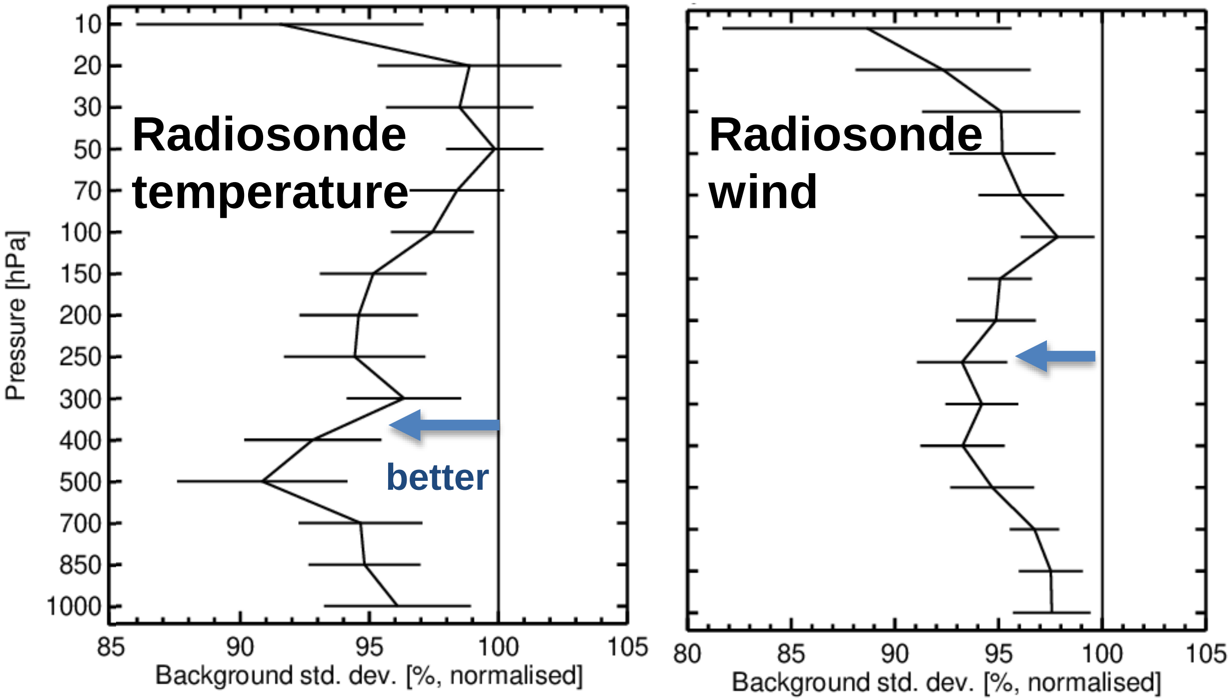
Using early hyperspectral sounder on Nimbus-4 (1970)



Typical 12 hour coverage

- Operated on Nimbus-4, from April 1970 – January 1971
- Nadir only observations. Coverage to 80°N to 80°S
- Spectral range 400 – 1600 cm^{-1}, resolution: 2.53 cm^{-1} to 2.69 cm^{-1}
- 94 km footprint, 13 s measurement time
- Assimilation experiment: TCO399 (25 km res.), L137, 01 June – 12 Aug 1970
- 60 CO_2 temperature sounding channels actively assimilated (624 – 706 cm^{-1})
- Diagonal errors: R = 1.0K, VarBC: Offset and 4 thickness predictors
- McNally & Watts cloud detection, parameters from Poli & Brunel (2016)
- IRIS RTTOV coefficients include several advanced effects (spectral shift, numerical apodisation, self-apodisation due to finite field-of-view…)

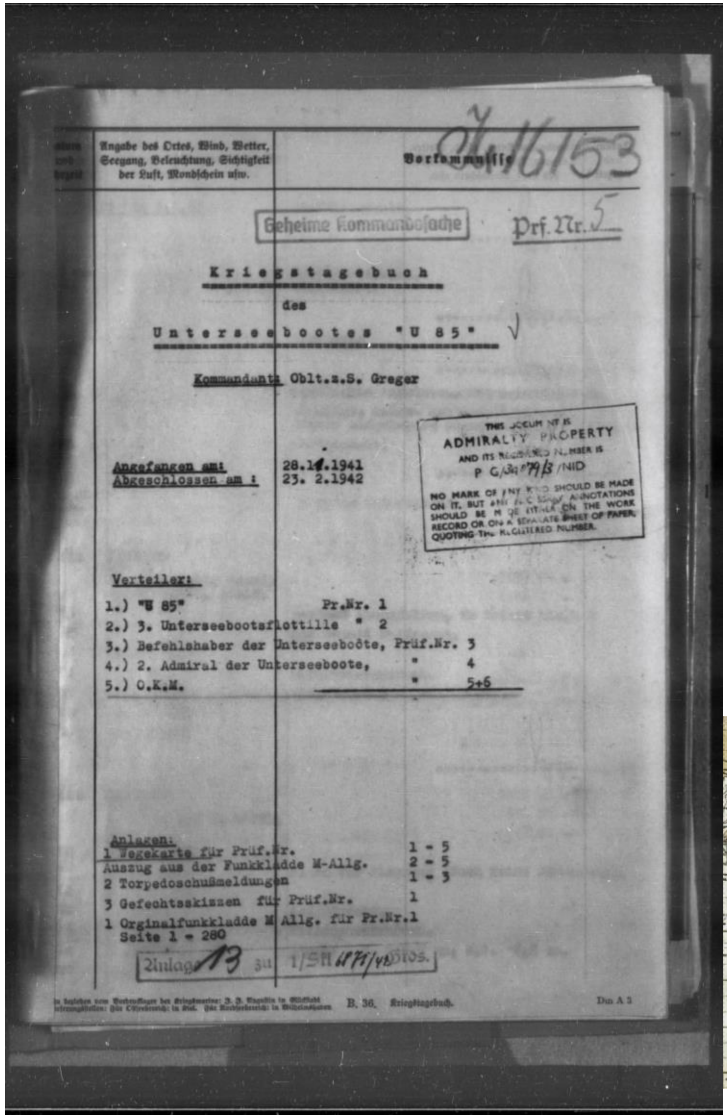




Observation minus Background

Climate Change

Part IV. Is there more to do? Example of ‘never-shared data’: Unterseebooten



Angabe des Ortes, Wind, Wetter, Seegang, Beleuchtung, Sichtigkeit der Luft, Mondschein usw.	Vorkommnisse

16153

Geheime Kommandosache

Prf.Nr. 5

Kriegstagebuch

des

Unterseebootes "U 85"

Kommandant: Oblt.z.S. Greger

THIS DOCUMENT IS ADMIRALTY PROPERTY AND ITS REGISTERED NUMBER IS P G/30079/3 /NID
NO MARK OF ANY KIND SHOULD BE MADE ON IT, BUT ANY NECESSARY ANNOTATIONS SHOULD BE MADE EITHER ON THE WORK RECORD OR ON A SEPARATE SHEET OF PAPER, QUOTING THE REGISTERED NUMBER.

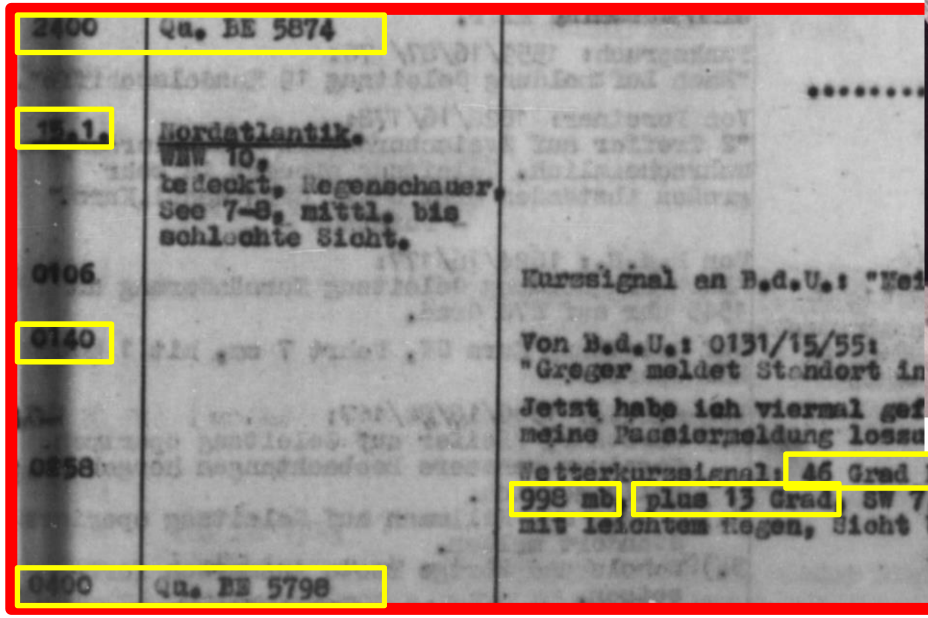
Angefangen am: 28.11.1941
Abgeschlossen am : 23. 2.1942

Verteiler:
1.) "U 85" Pr.Nr. 1
2.) 3. Unterseebootsflottille " 2
3.) Befehlshaber der Unterseeboote, Prüf.Nr. 3
4.) 2. Admiral der Unterseeboote, " 4
5.) O.K.M. " 5+6

Anlagen:
1 Wegekarte für Prüf.Nr. 1 - 5
Auszug aus der Funkkladde M-Allg. 2 - 5
2 Torpedoschußmeldungen 1 - 3
3 Gefechtsskizzen für Prüf.Nr. 1
1 Orginalfunkkladde M Allg. für Pr.Nr.1
Seite 1 - 280

Anlage 13 zu 1/Skl 6871/42 Gkdos.

B. 36. Kriegstagebuch. Din A 3



2400 Qu. BE 5874

15.1. Nordatlantik. WNW 10, bedeckt, Regenschauer, See 7-8, mittl. bis schlechte Sicht.

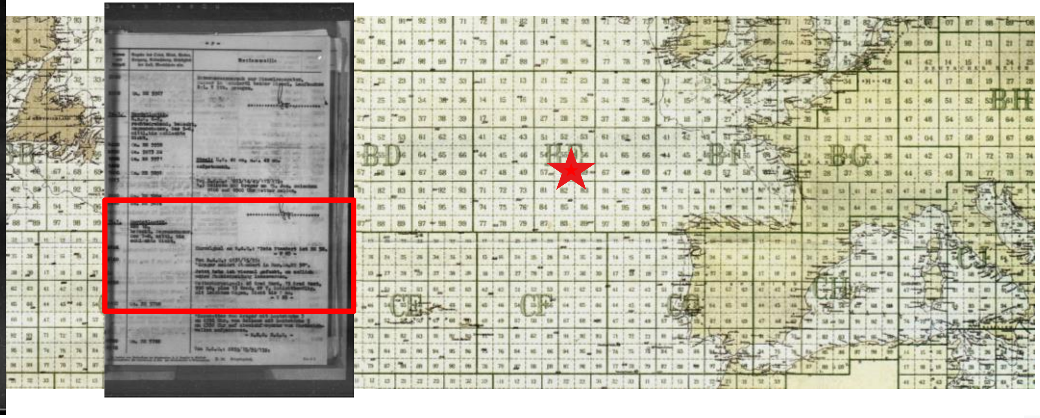
0106 Kurzsignal an B.d.U.: "Wet

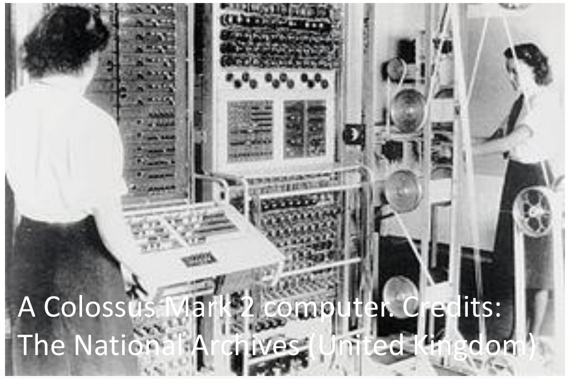
0140 Von B.d.U.: 0131/15/55: "Greger meldet Standort in
Jetzt habe ich viermal gef
meine Passiermeldung lossu

0258 Wetterkurzsignal: 46 Grad Nord, 19 Grad West, 998 mb, plus 13 Grad, SW 7, Schichtbewölkg. mit leichtem Regen, Sicht bis 2 sm.
- U 85 -

0400 Qu. BE 5798

Bletchley Park. Credits: Sir John Dermot Turing.





A Colossus Mark 2 computer. Credits: The National Archives (United Kingdom)

Data source: Records from the German Naval Archives microfilmed by the United States Navy, Office of Naval Intelligence (ONI) at the Admiralty, London. “Kriegestagebuch der Unterseebootes U 85, Kommandant Greger”. PG 30079, National Archives Microfilm Publication T1022, roll 2932.

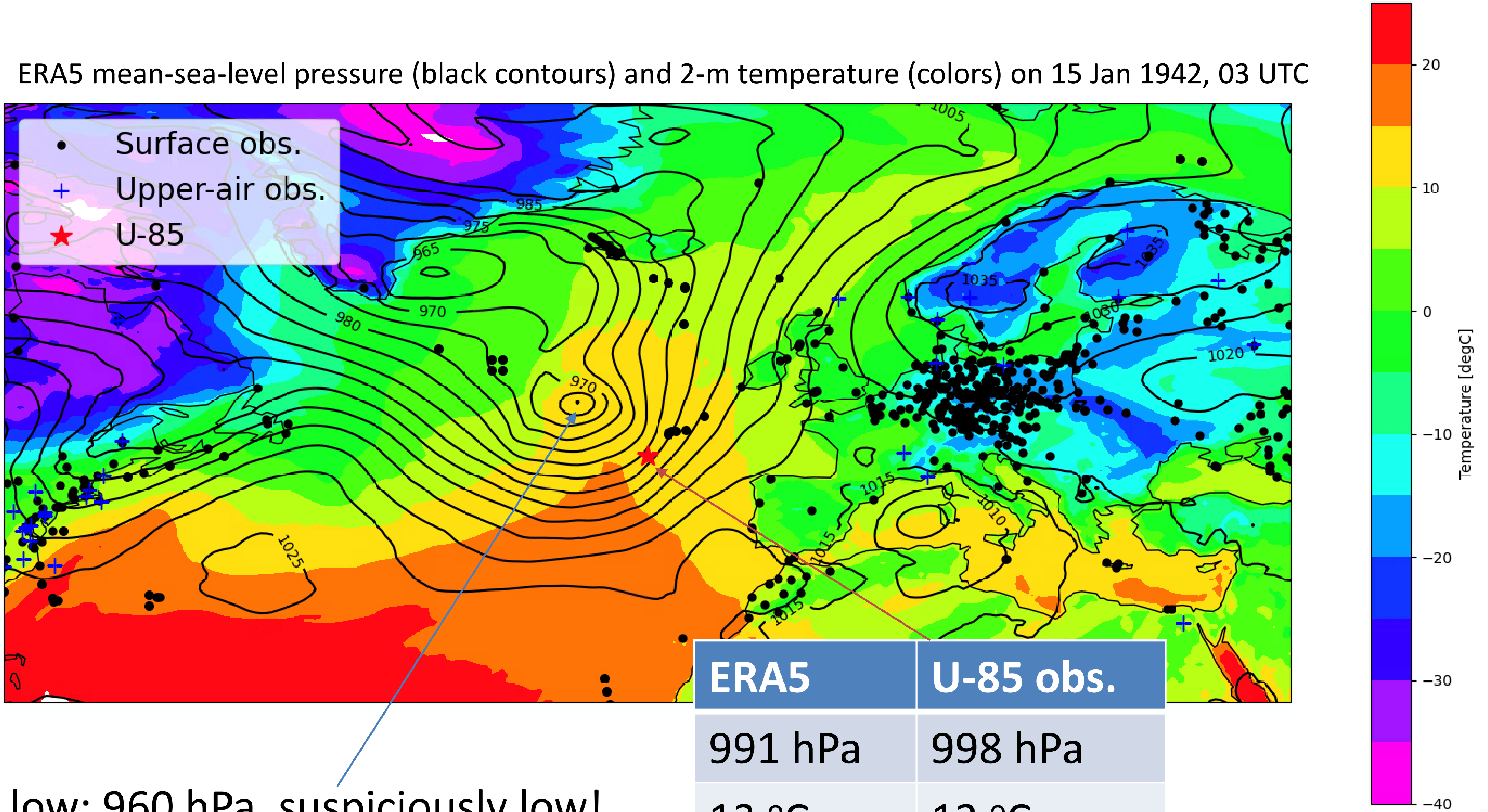
Copernicus Europe's eyes on Earth

European Commission

Climate Change

How would such an observation help?

ERA5 mean-sea-level pressure (black contours) and 2-m temperature (colors) on 15 Jan 1942, 03 UTC



ERA5	U-85 obs.
991 hPa	998 hPa
13 ºC	13 ºC

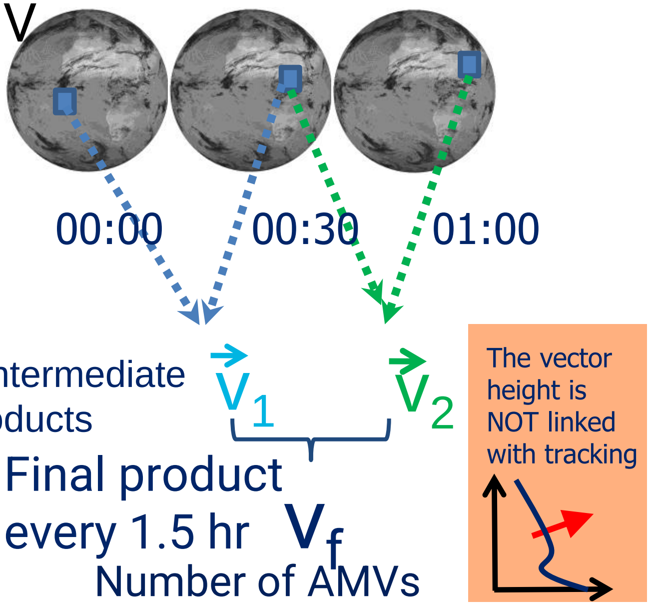
ERA5 low: 960 hPa, suspiciously low!

Improving data yield: Meteosat Atmospheric Motion Vectors (AMV)

Climate Change

Reprocessed AMV
Input to ERA5

Period
0°:1982 – 2001
IODC: 1998 - 2017

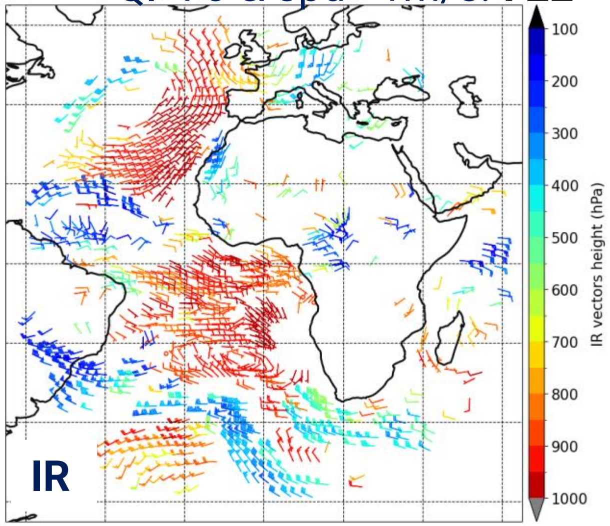


2 intermediate products

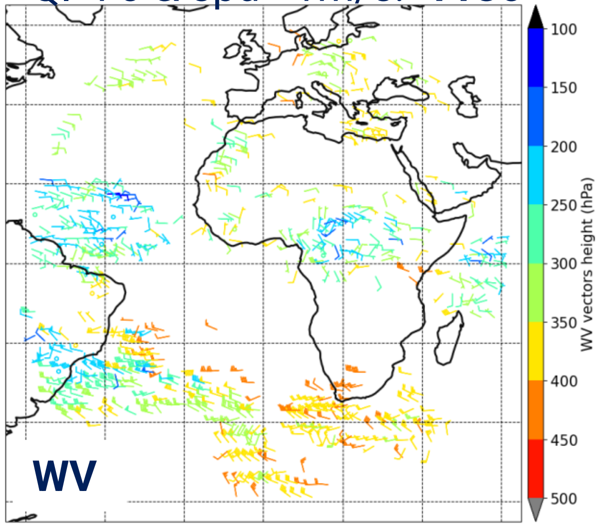
Final product every 1.5 hr V_f

The vector height is NOT linked with tracking

Number of AMVs
QI>70 & spd> 1m/s: **722**

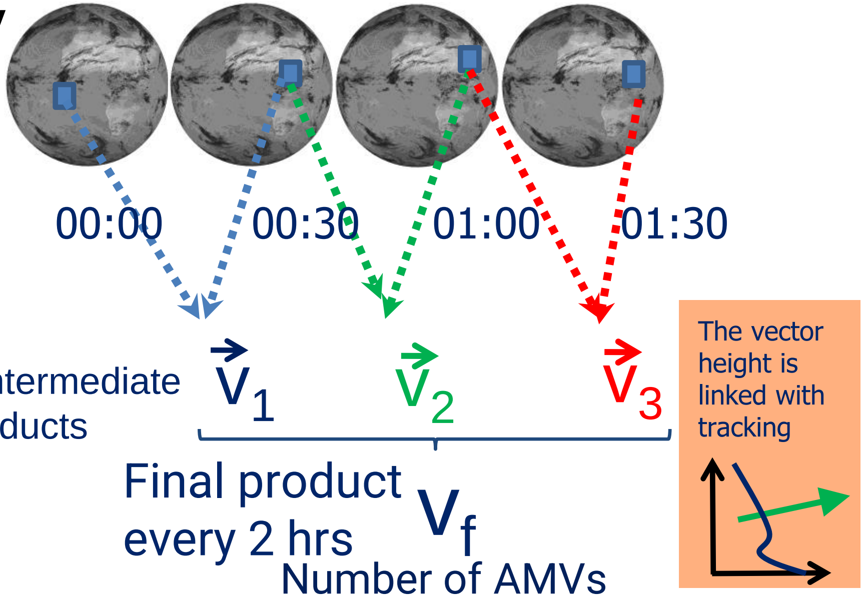


Number of AMVs
QI>70 & spd> 1m/s: **1186**



Reprocessed AMV
For ERA6

Period
0°: 1982 – 2006
IODC: 2007 - 2017

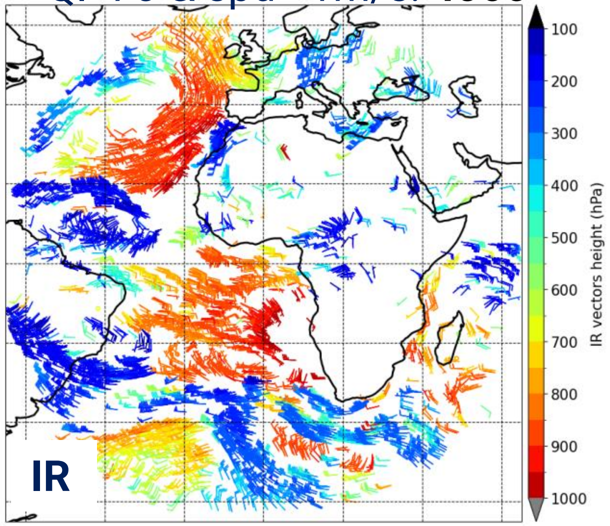


3 intermediate products

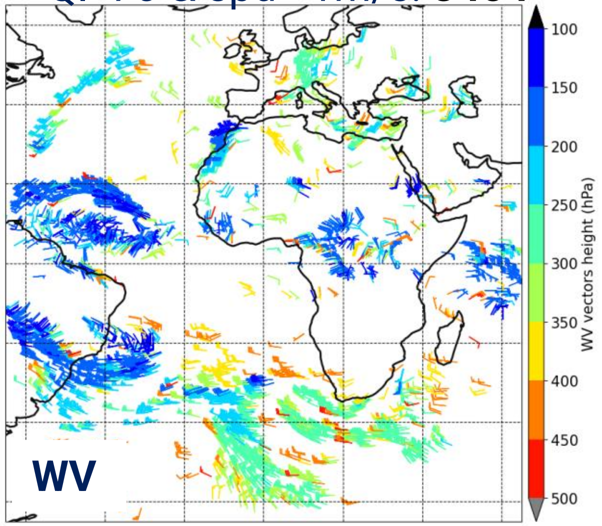
Final product every 2 hrs V_f

The vector height is linked with tracking

Number of AMVs
QI>70 & spd> 1m/s: **4066**



Number of AMVs
QI>70 & spd> 1m/s: **3454**



Meteosat-7, 11 June 1998, 00 UTC

Credits: EUMETSAT

Conclusions

- The pool of past observations is **finite,** and **at risk** -- until secured in modern archives

Picture from https://commons.wikimedia.org/wiki/User:Pkuczynski

- Getting out of this loop requires **long-term data management planning**
  (e.g., ISO/TC 171/SC 1 on “Quality, preservation and integrity of information”)
- Citizens’ science allows to widen the effort base towards data rescue

Old Weather

Help scientists transcribe Arctic and worldwide weather observations recorded in ship's logs since the mid-19th century.

First voyage

Recover hidden weather data collected by the Navy during World War II

The deep

Explore the Arctic of the past from the deck of a whaling ship

Rediscover the historic Arctic voyages of the U.S. Navy and Coast Guard

https://www.oldweather.org

- Assembling datasets requires consolidating the **information gained over time**, keeping **traceability** to the sources, **learning from other efforts,** with pre-assimilation feedback, post-assimilation observation feedback exchange, and sharing of issues found
- **Data reprocessing** allows to extract more information from previous records
- Importance of **quantifying the impact**: Value Of Information (Weatherhead et al., 2017)
- Planning future new systems:
  - **Data assimilation systems, radiative transfer models**: support ‘old’ obs. (e.g., EUMETSAT NWP-SAF RTTOV)
  - **Observing systems**: long-term data management incl. reprocessing, GCOS requirements, high calibration standards

Observations: data rescue and reprocessing

Climate Change

Thank you for your attention!