Practical challenges using observations in operational NWP

Peter Lean

Thanks to : Niels Bormann, Alan Geer, Mohamed Dahoui, Tony McNally, Tomas Kral
Jenny Rourke, Axel Bonet, Cristiano Zanna, Andrew Bennett , Cornel Soci
Massimo Bonavita, Elias Hólm, Heikki Järvinen, Phil Browne, Giovanna De Chiara

ECMWF

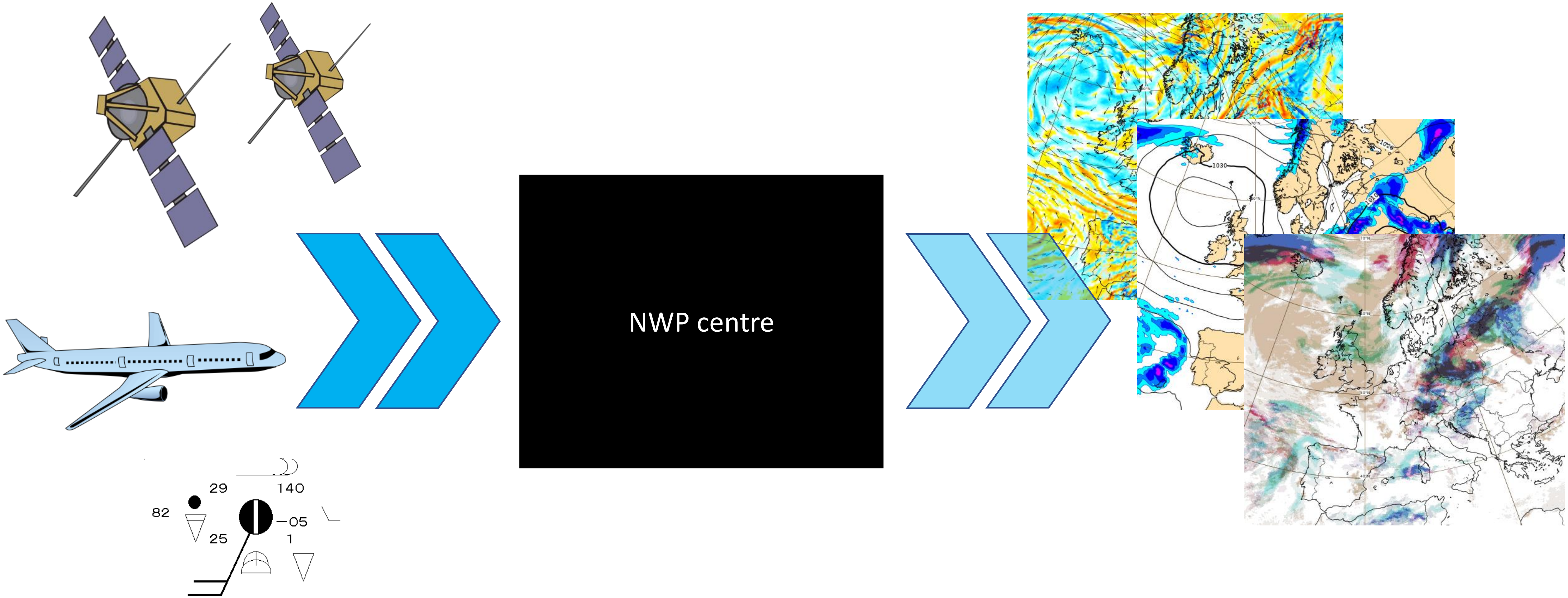
Annual Seminar - September 2021

ECMWF

Outline

- **Time critical** aspects for real time NWP systems
- Importance of **robustness** for operational observation processing
- Handling ever increasing **variety** of observations

Top level schematic of an NWP centre



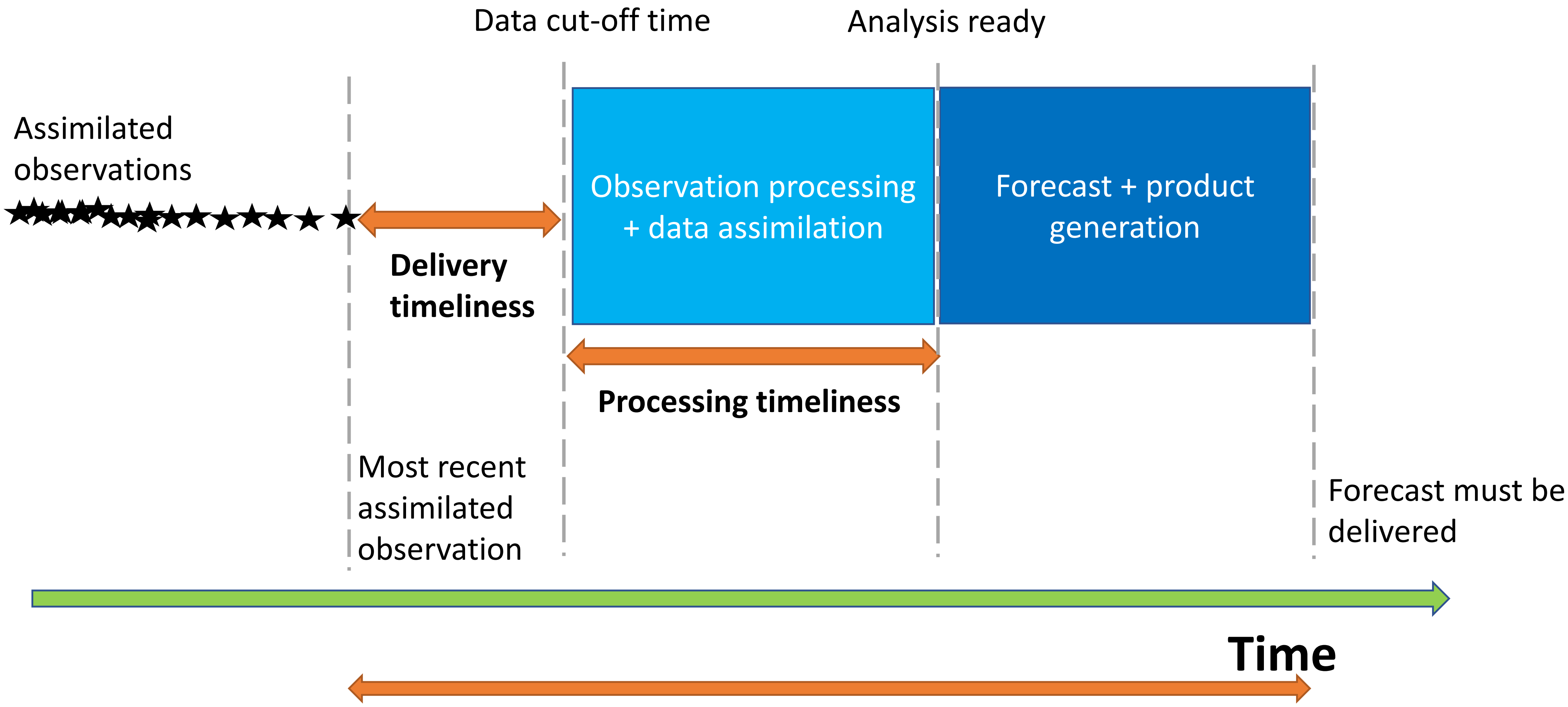
Top level schematic of an NWP centre

- In an ideal world, forecasts would be produced instantaneously and use all observations that have been made up until that time.

- But in the real world, there are practical challenges for real-time forecast production:

  - **Observation latency**; delay between when they are made, and when delivered
  - **Processing time**; generating an analysis and forecast is computationally expensive

A journey along the time critical path



By the time a forecast is released, the observations upon which it is based may be several hours old

Value of recent observations

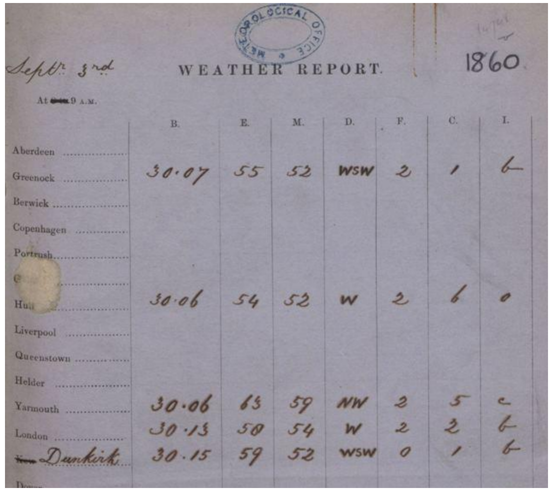
- To maximise value from observational data for NWP we need good:
  - Delivery timeliness
  - Processing timeliness
- Improving either of these allows more recent observations to be assimilated, improving the quality of the analysis

Historical context; extracting value for real time forecasting

- Delivery timeliness

High quality meteorological observations have been made for hundreds of years, but data could not be delivered fast enough to provide value for real time forecasting.

It took a technological advance (invention of the telegraph) to improve the delivery timeliness and make real time forecasting feasible (mid-1800s)


Sept. 3rd WEATHER REPORT. 1860

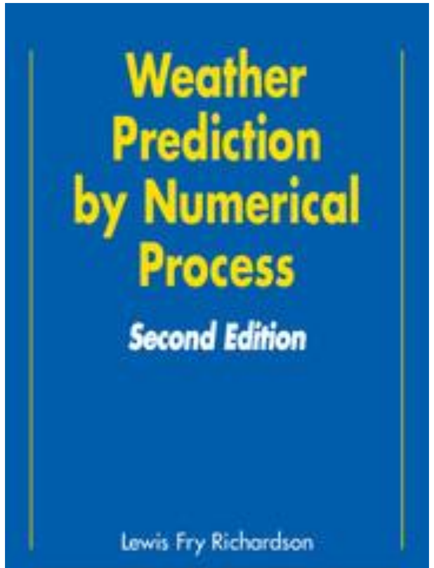
At 9 A.M.

	B.	E.	M.	D.	F.	C.	I.
Aberdeen							
Greenock	30.07	55	52	WSW	2	1	6
Berwick							
Copenhagen							
Portrush							
Hull	30.06	54	52	W	2	6	0
Liverpool							
Queenstown							
Helder							
Yarmouth	30.06	63	59	NW	2	5	c
London	30.13	58	54	W	2	2	6
Dunkirk	30.15	59	52	WSW	0	1	6

- Processing timeliness

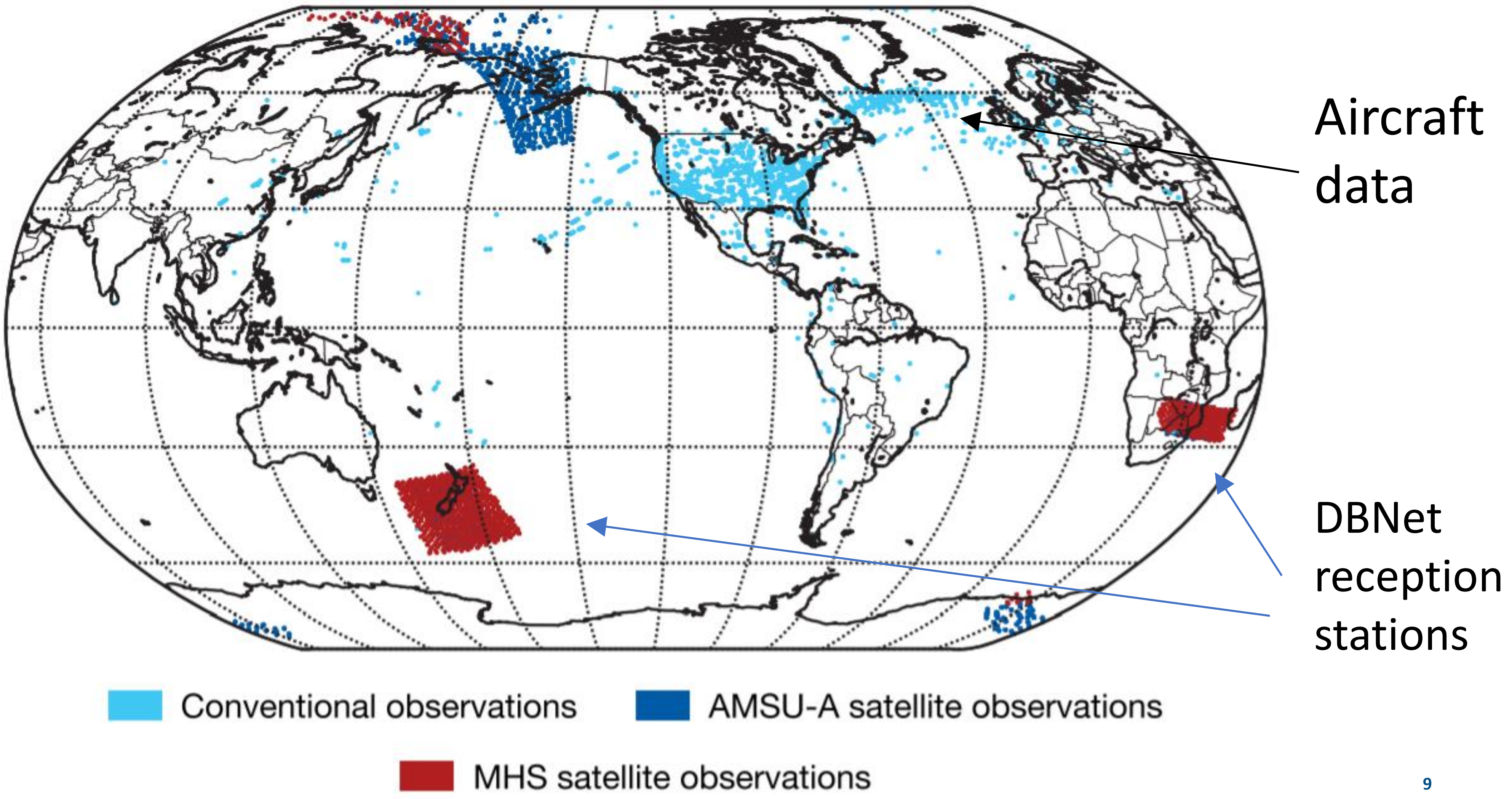
Richardson, L.F. (1922) made an NWP forecast to t+6h, but it took 6 weeks to process the data by hand!

It took a technological advance (invention of the computer) for faster processing to be possible; ENIAC took 24 hours to process a t+24h forecast (1950)



Observation timeliness (delivery)

Observation coverage in final 15 minutes of assimilation window



Value of recent observations

Received: 27 March 2019 | Revised: 29 May 2019 | Accepted: 21 June 2019 | Published on: 29 July 2019
DOI: 10.1002/qj.3596

RESEARCH ARTICLE

Quarterly Journal of the Royal Meteorological Society RMetS

On the sensitivity of a 4D-Var analysis system to satellite observations located at different times within the assimilation window

Anthony P. McNally

European Centre for Medium-Range Weather Forecasts, Reading, UK

Correspondence
Anthony P. McNally, ECMWF, Shinfield Park, Reading RG2 9AX, UK.
Email: anthony.mcnally@ecmwf.int

Funding information
EUMETSAT project
EUM/RSP/SOW/15/814210

Abstract

This study quantifies the extent to which the ECMWF 4D-Var displays differential (heightened) sensitivity to observations located near the end of the 12-hr assimilation time window compared to observations located near the start of the window. Using dedicated satellite data denial experiments, it is shown that the lattermost 3 hr of observations are significantly more influential on the quality of the assimilation and forecasting system than the first 3 hr of data. Furthermore, it is found that the last 3 hr of data even outperforms the 6 hr of data (i.e. twice the number of observations) located in the first half of the window. The heightened importance of late window data is discussed in terms of these measurements being our most up-to-date information on the atmosphere, but also their ability to provide additional dynamical information to the assimilation system via feature advection wind tracing. The implications of this sensitivity are discussed. Firstly, it leads to the existence of influential (late window) satellite orbits, the location of which can have a strong bearing on the impact of observations from different satellites in different regions. Secondly, this sensitivity reinforces the need for data providers to minimize dissemination delays to ensure that crucial late window data reach users in time to be assimilated. Finally, numerical weather prediction (NWP) centres (who run 4D systems) must ensure that these lattermost observations are being captured and used effectively. Some suggestions for this are proposed.

KEYWORDS
4D-Var; data assimilation; NWP; tools and methods

1 | INTRODUCTION

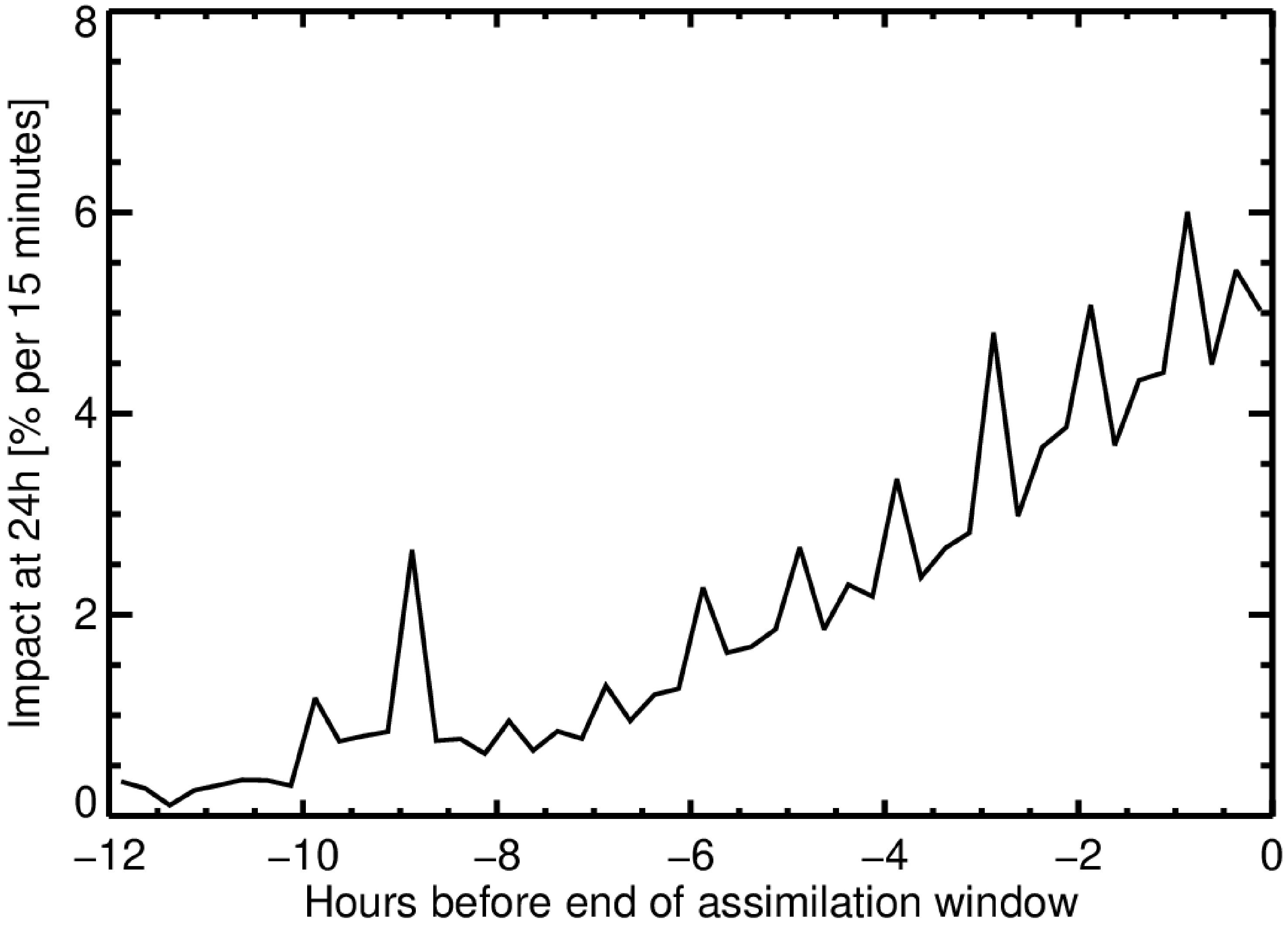
Numerical weather prediction (NWP) systems require an accurate estimate of the current atmospheric conditions (or initial state) from which a forecast model can be run to pre-

is obtained by analysing information from the global network of meteorological observations. However, these observations may be distributed randomly (and incompletely) over the model spatial domain, made at a variety of different times during the day (e.g. not at 1200 UTC) and indeed measure

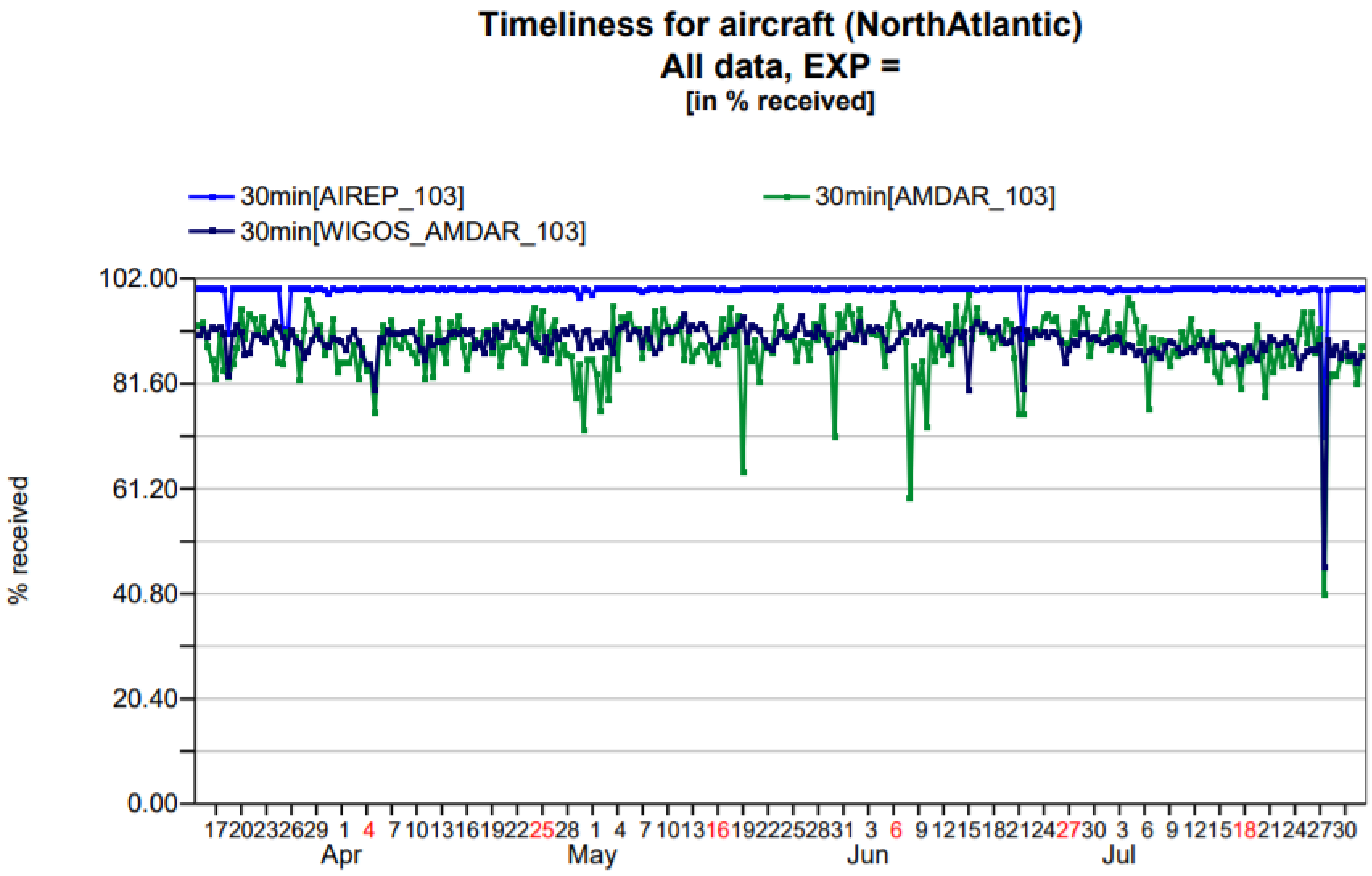
- McNally (2019)
- Observations at the end of the assimilation window are significantly more influential than less recent observations earlier in the window.

Forecast Sensitivity Observation Impact

Thanks: **Alan Geer**



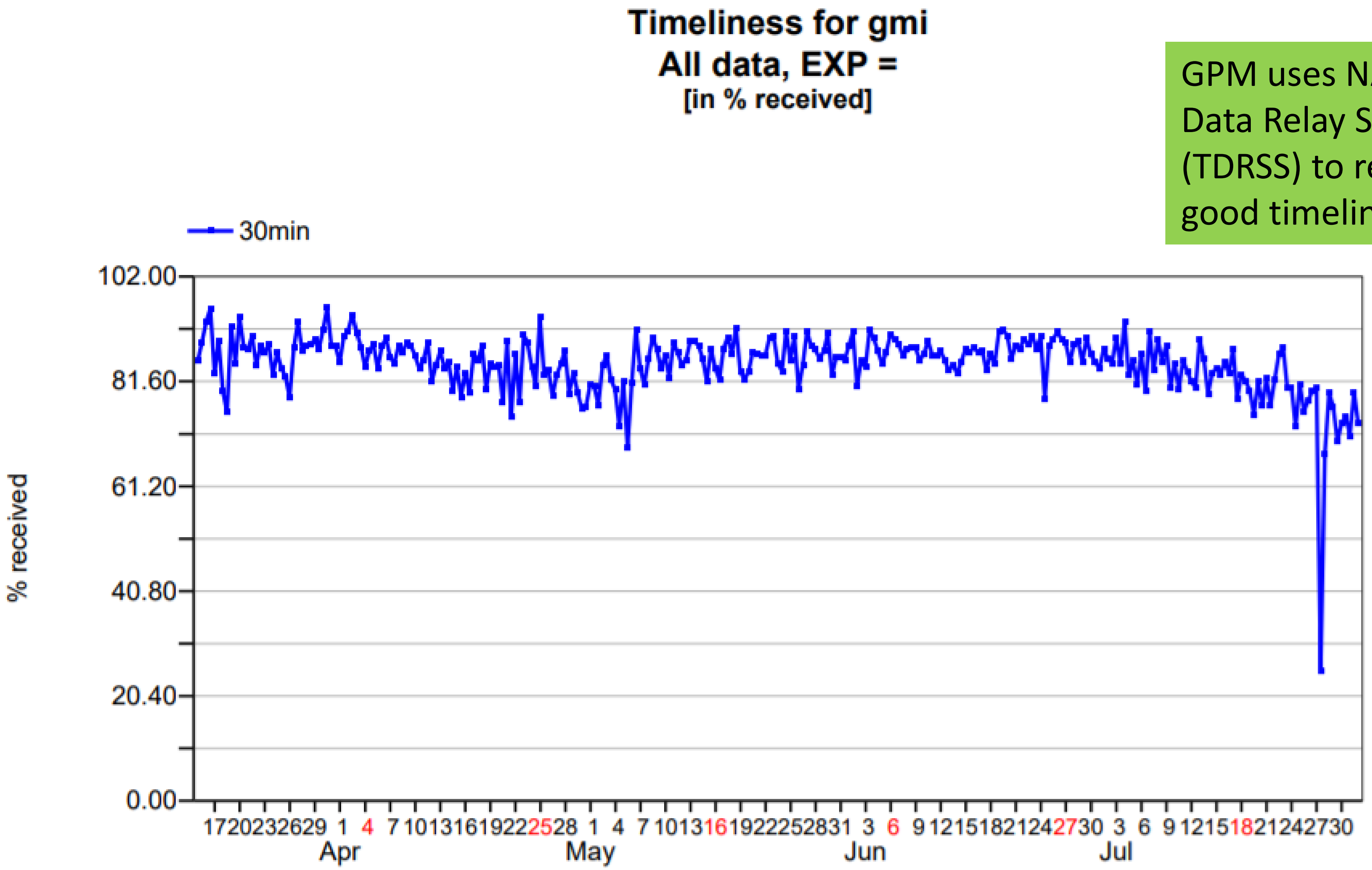
Timeliness: % delivered within 30 minutes of observation time
Aircraft



ECMWF EUROPEAN CENTRE FOR MEDIUM-RANGE WEATHER FORECASTS

Thanks: **Mohamed Dahoui, Tomas Kral**

Timeliness: % delivered within 30 minutes of observation time
GMI

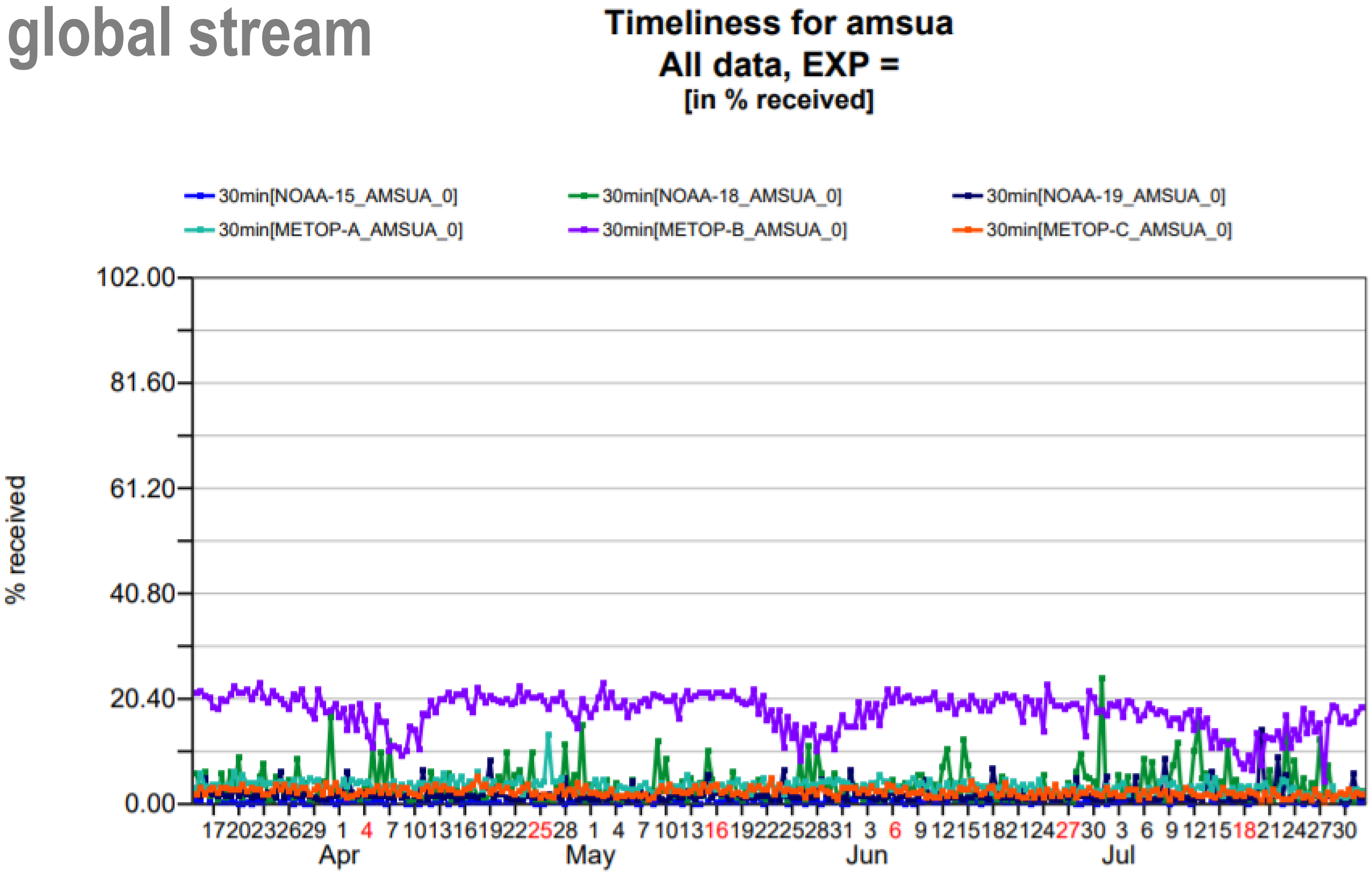


GPM uses NASA's Tracking and Data Relay Satellite System (TDRSS) to relay data with very good timeliness

ECMWF EUROPEAN CENTRE FOR MEDIUM-RANGE WEATHER FORECASTS

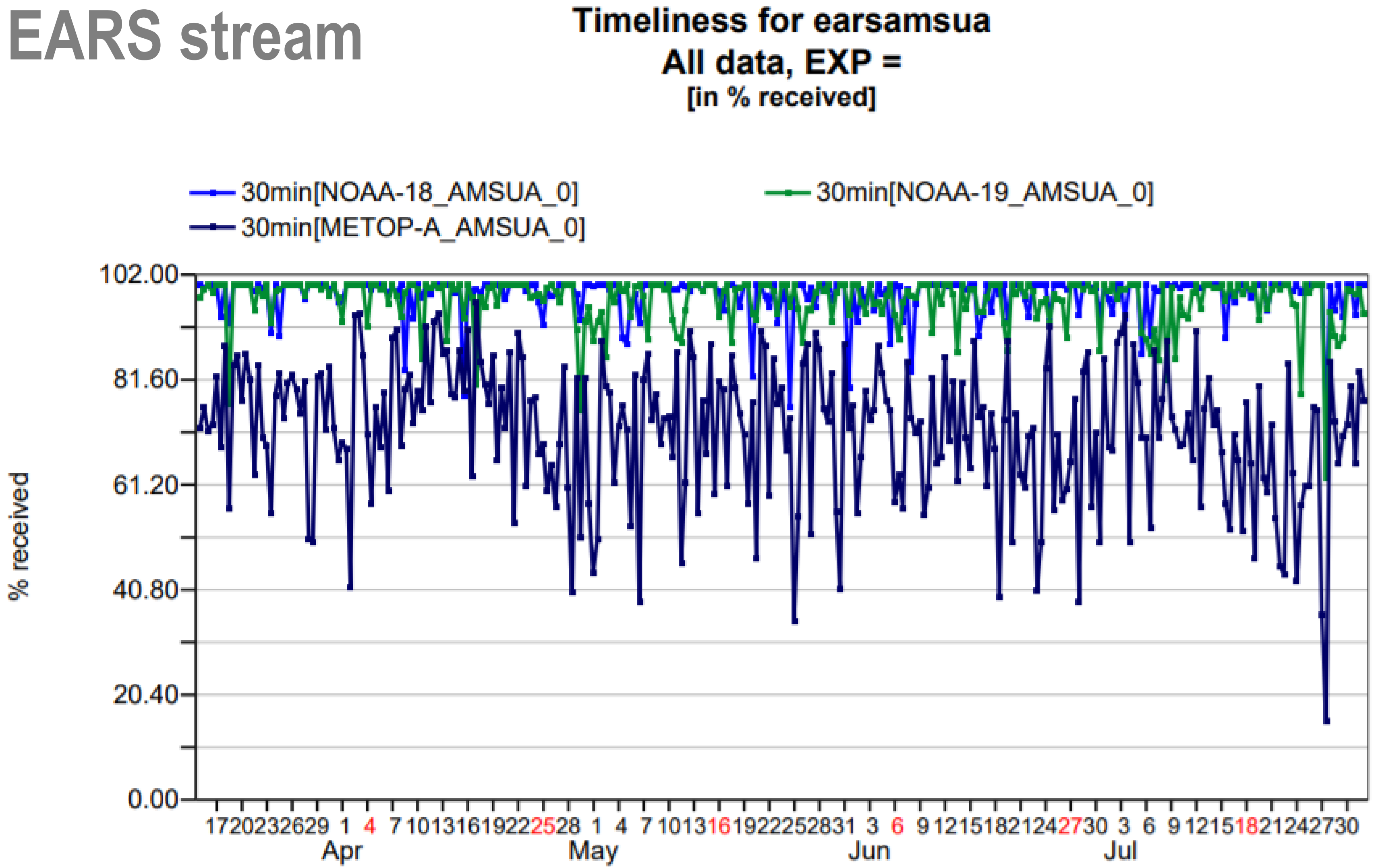
Thanks: **Mohamed Dahoui, Tomas Kral**

Timeliness: % delivered within 30 minutes of observation time AMSU-A global stream



Thanks: **Mohamed Dahoui, Tomas Kral**

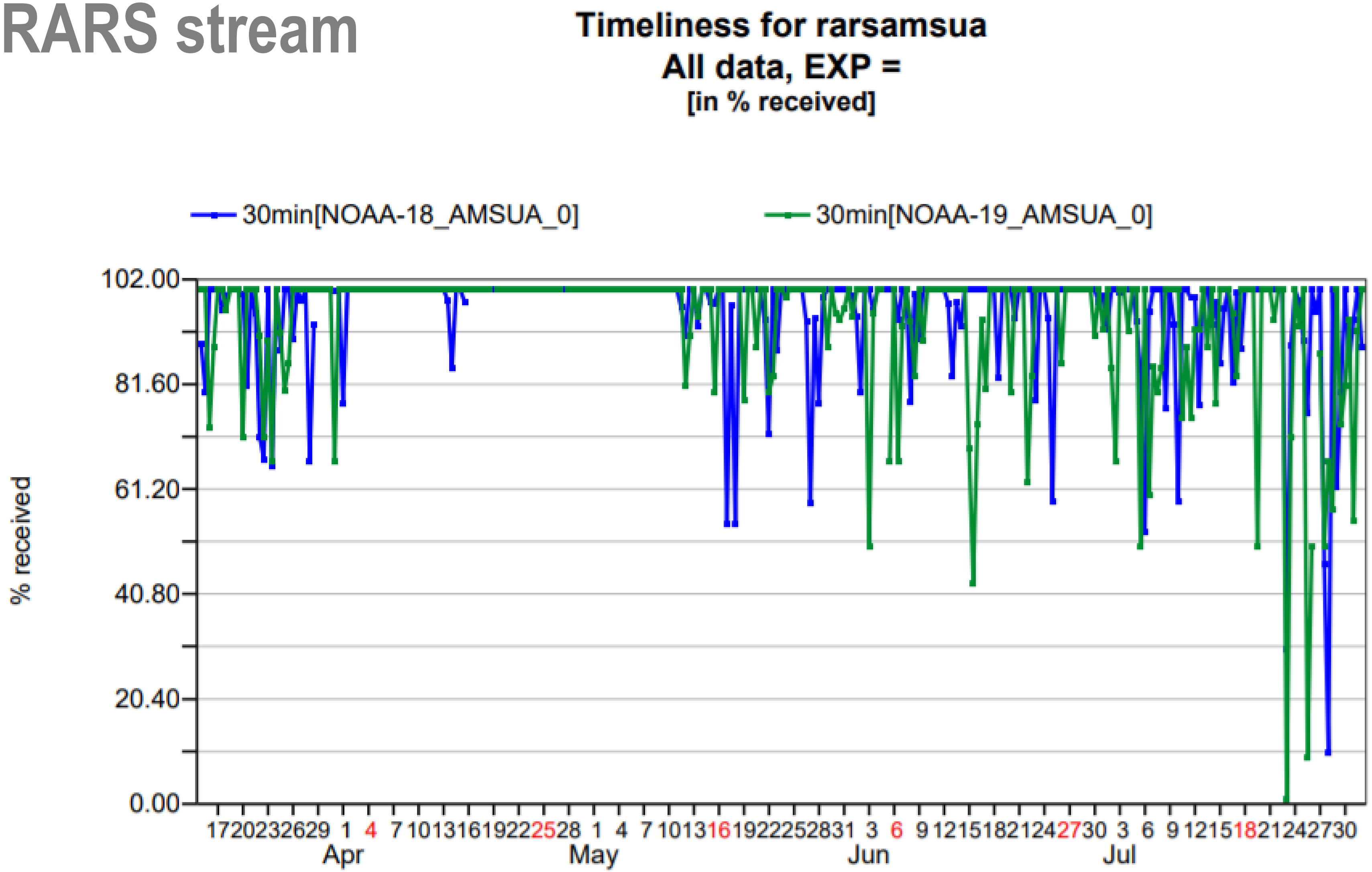
Timeliness: % delivered within 30 minutes of observation time
AMSU-A EARS stream



ECMWF EUROPEAN CENTRE FOR MEDIUM-RANGE WEATHER FORECASTS

Thanks: **Mohamed Dahoui, Tomas Kral**

Timeliness: % delivered within 30 minutes of observation time AMSU-A RARS stream



ECMWF EUROPEAN CENTRE FOR MEDIUM-RANGE WEATHER FORECASTS

Thanks: **Mohamed Dahoui, Tomas Kral**

DBNet

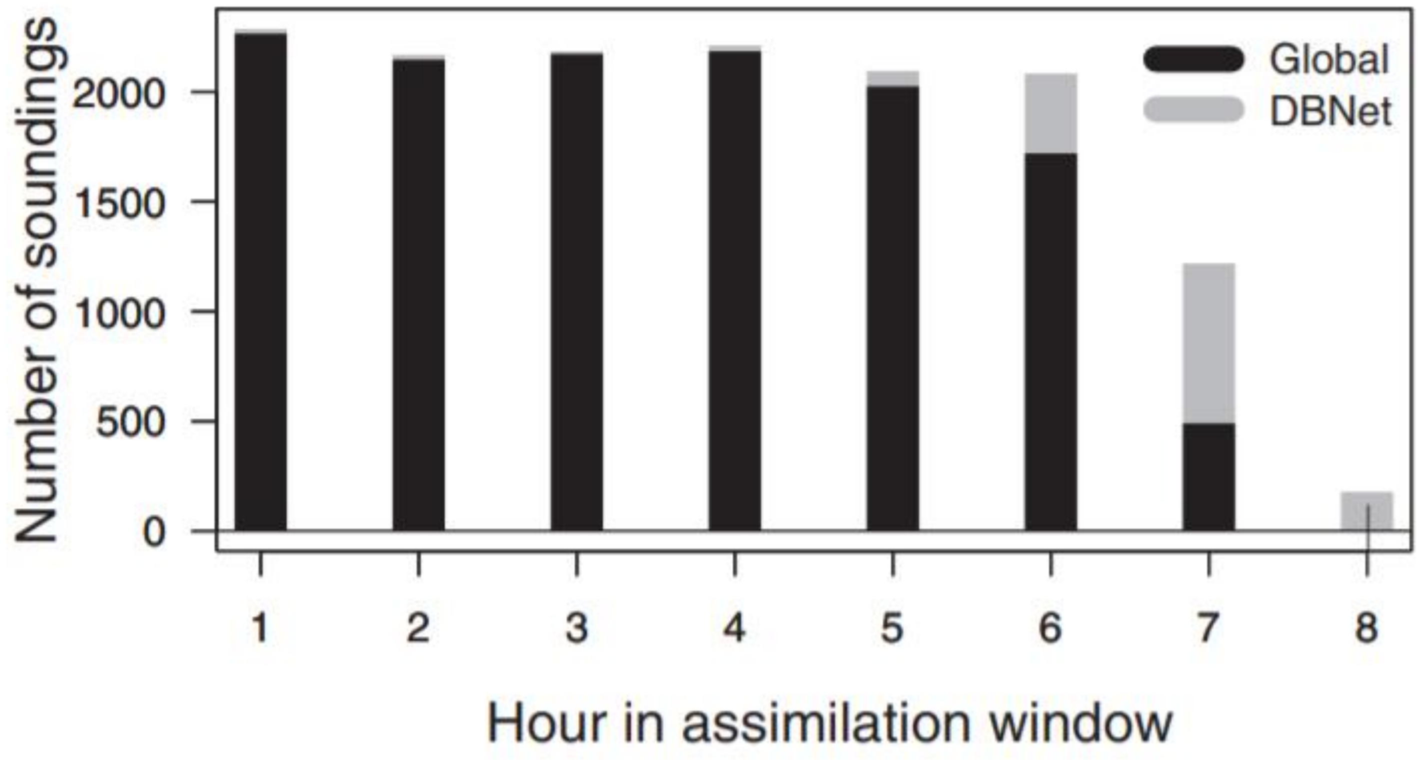
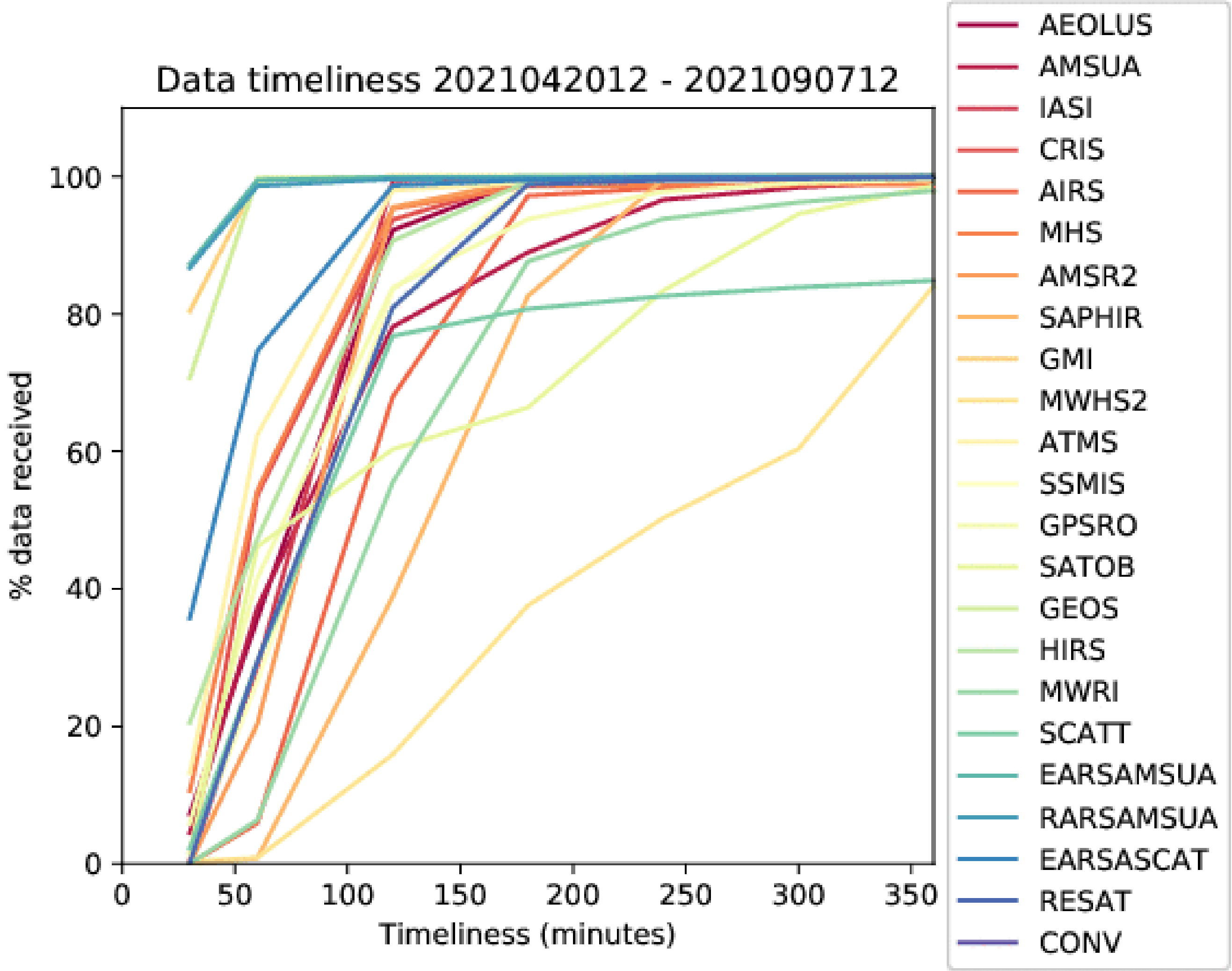


FIGURE 13 Number of assimilated NOAA-19 AMSU-A soundings in each hour of the early-delivery assimilation window

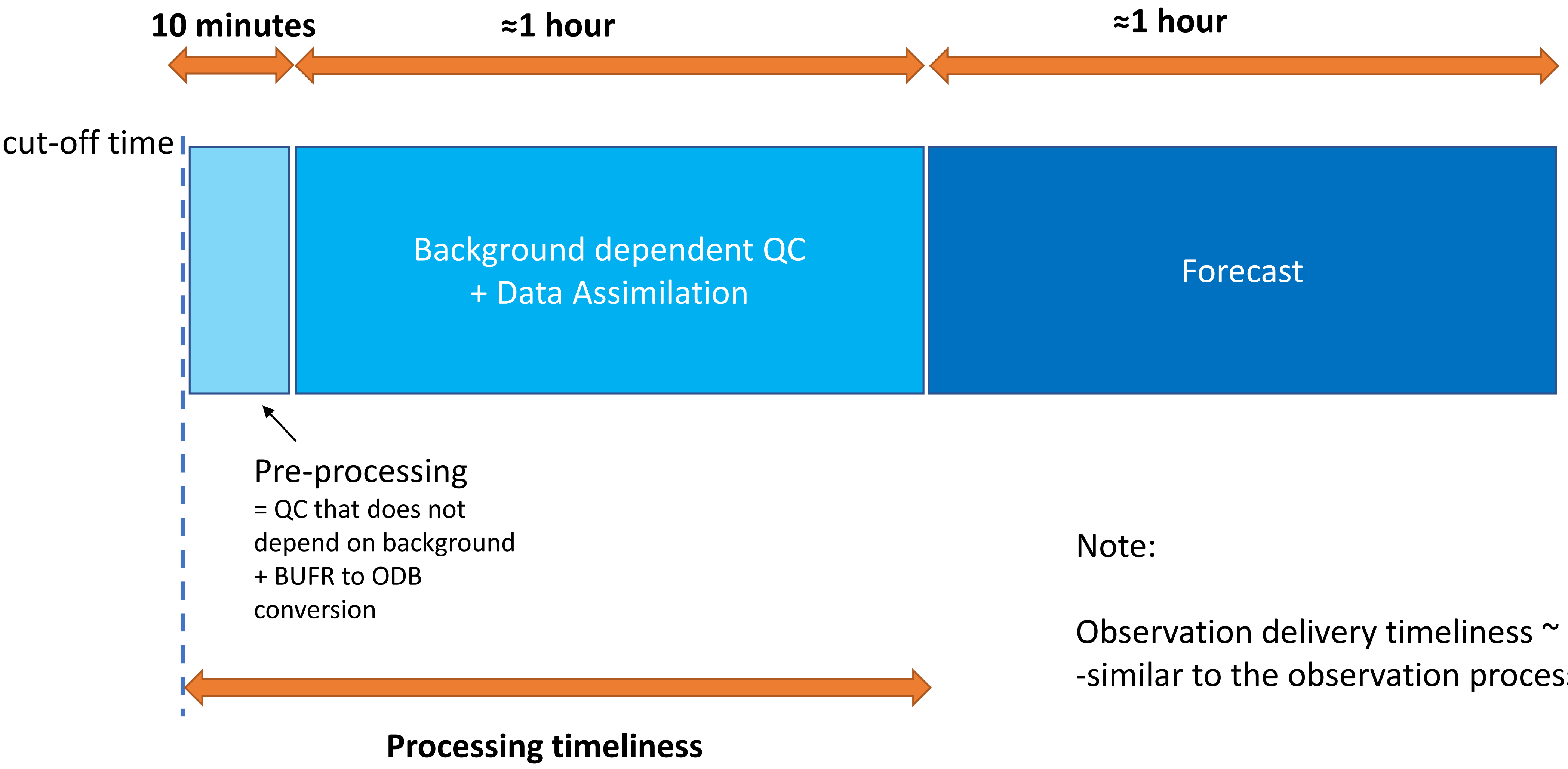
Delivery timeliness

- Overall:
  - approx. 50% delivered within 1-hour
  - approx. 90% delivered within 2-hours



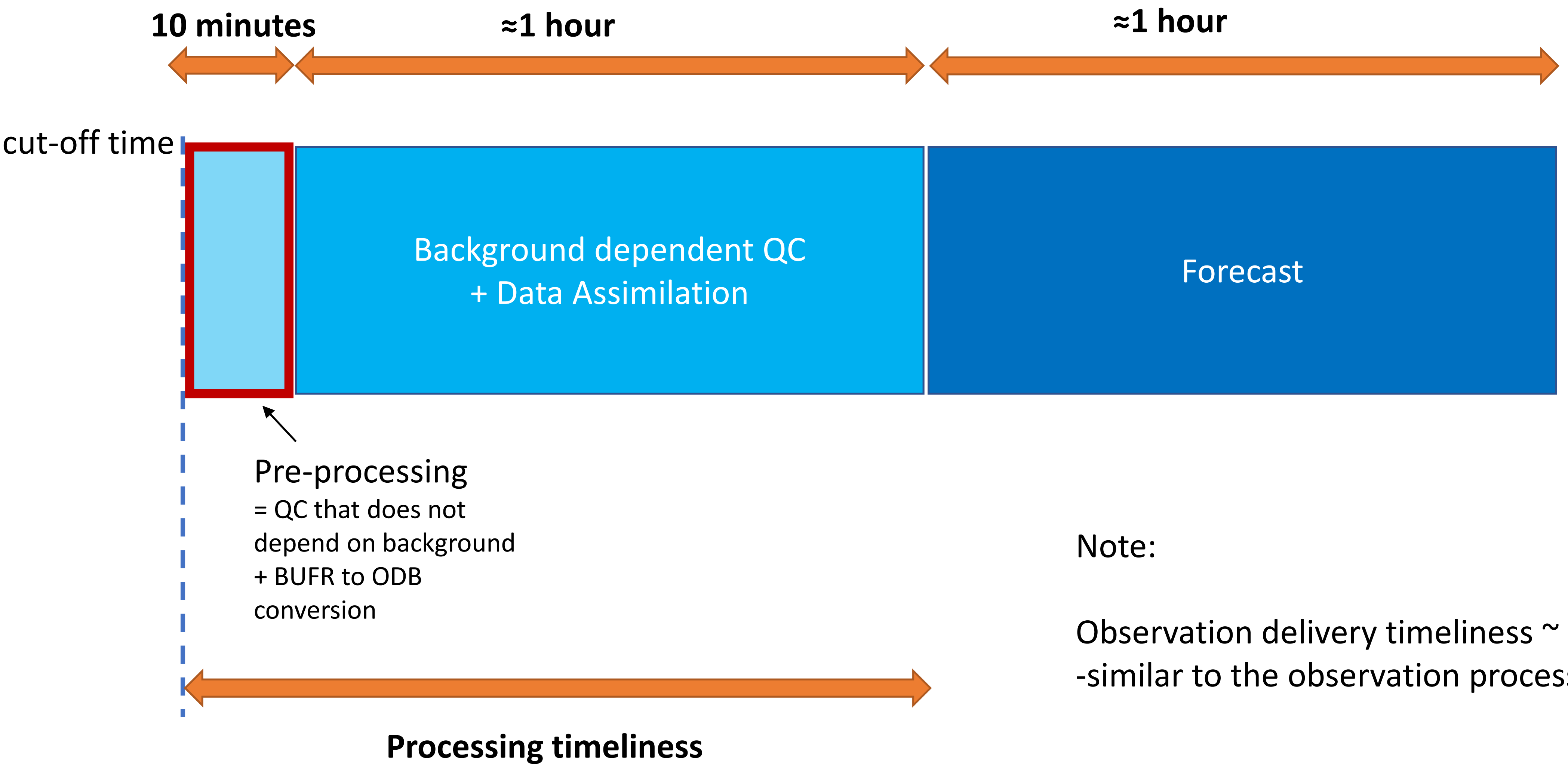
Initiatives to improve the delivery timeliness greatly increase the value of those observations for real-time NWP

Observation processing timeliness



Note:

Observation delivery timeliness ~ 1-2 hours
-similar to the observation processing timeliness

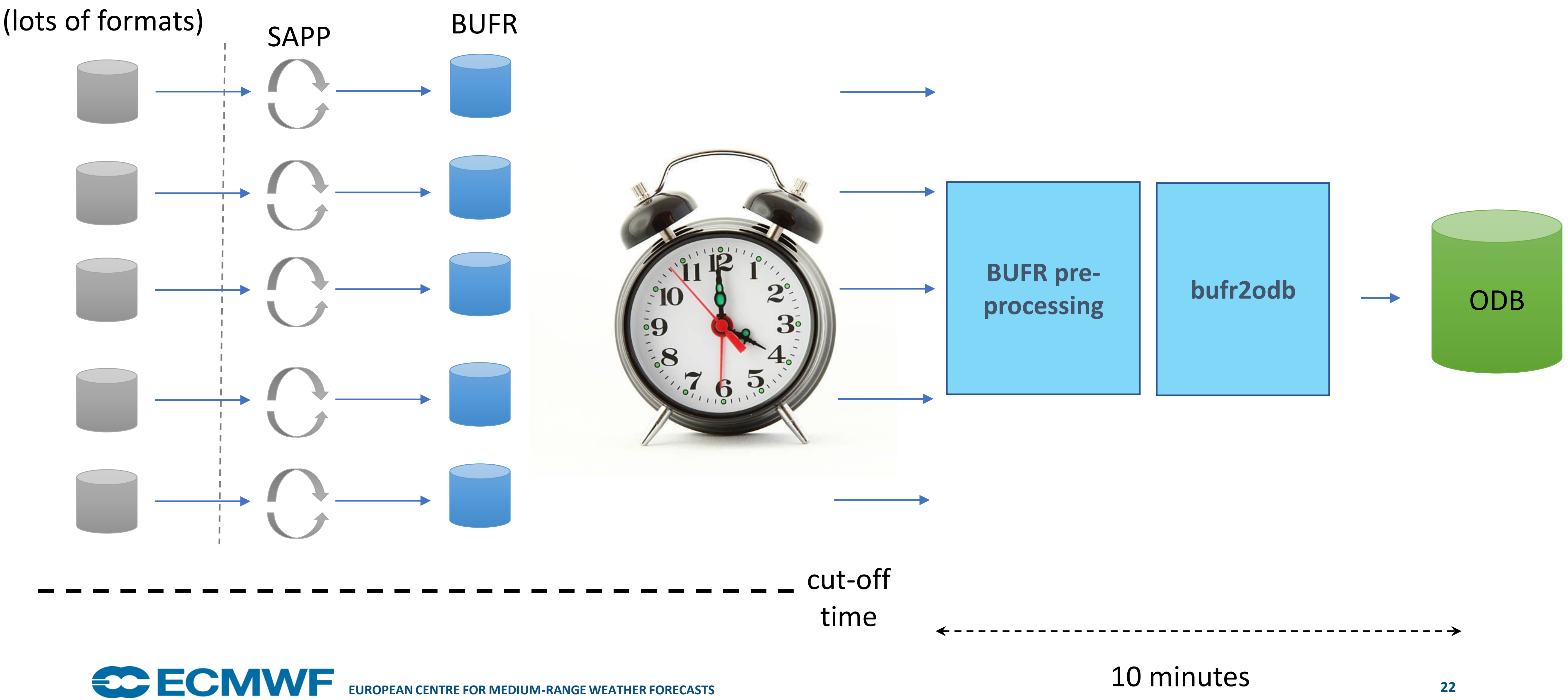


Note:

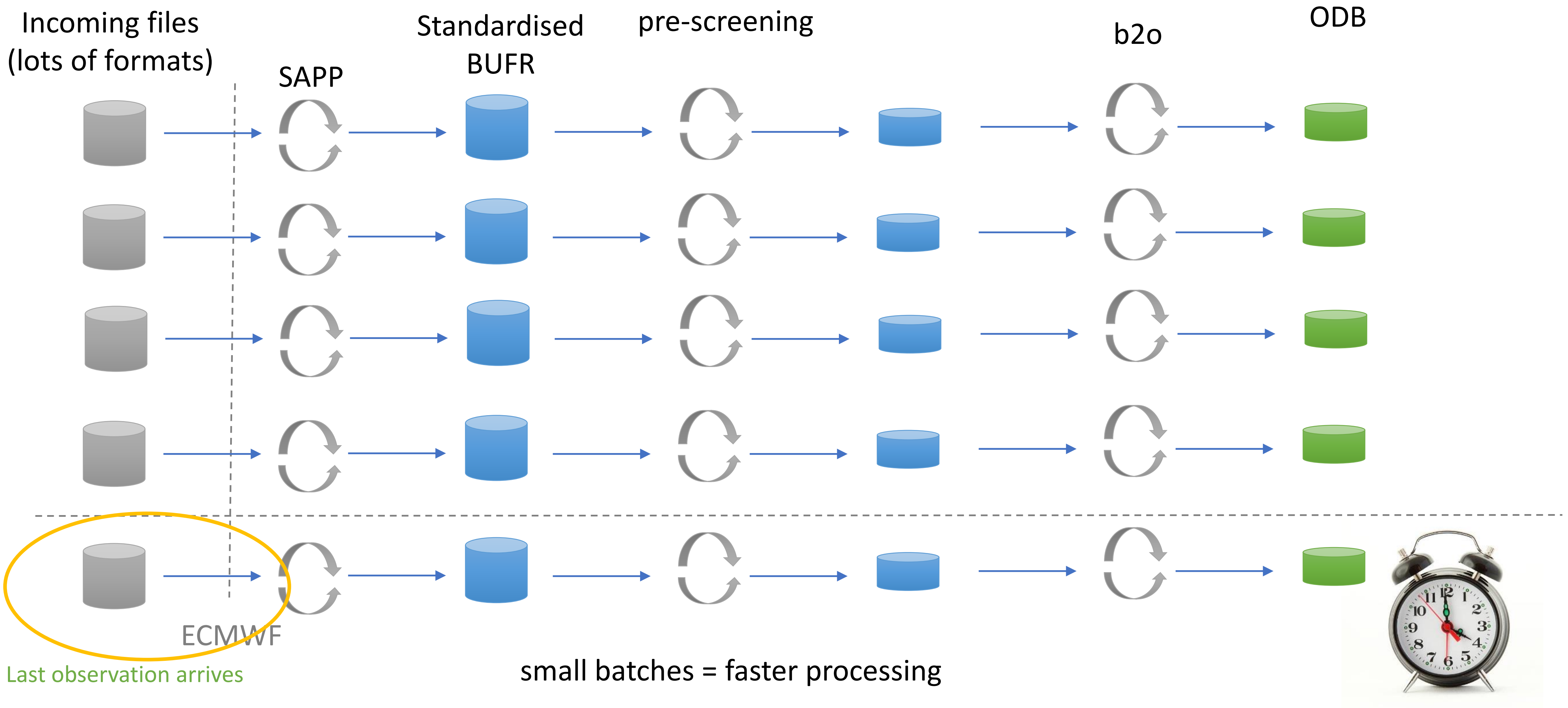
Observation delivery timeliness ~ 1-2 hours
-similar to the observation processing timeliness

Pre-processing

None of this relies on the model background -why wait for the cut-off time?



Continuous Observation Pre-processing (COPE)

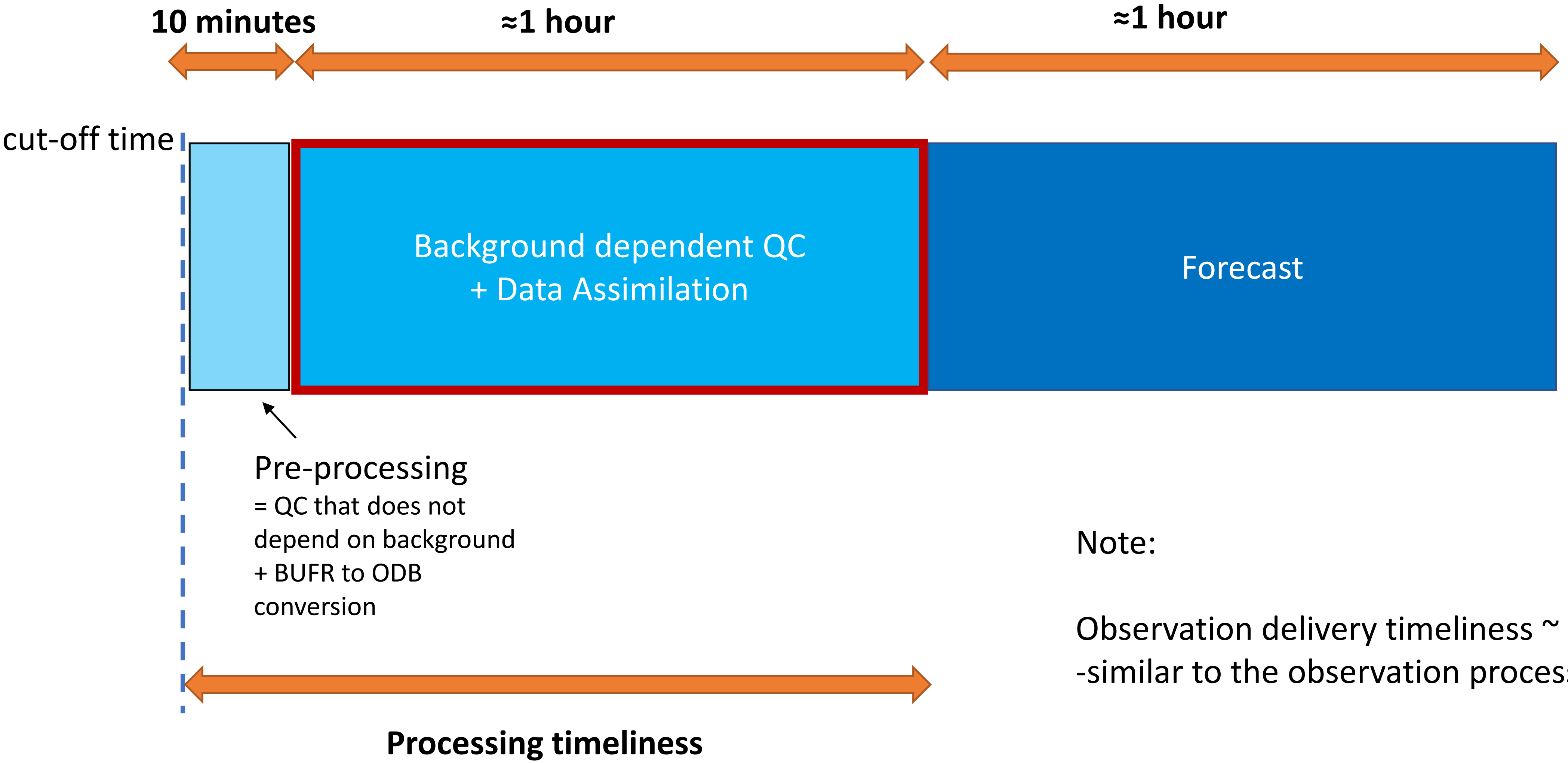


Quasi-continuous observation pre-processing

- Upcoming IFS cycle (48r1) upgrades the pre-processing code to provide a clean framework for incremental pre-processing

- Following IFS cycle will introduce quasi-continuous pre-processing (hourly) which will remove the majority of the pre-processing from the time critical path

Jenny Rourke, Tomas Kral, Axel Bonet, Cristiano Zanna, Andrew Bennett



Note:

Observation delivery timeliness ~ 1-2 hours
-similar to the observation processing timeliness

Cost of radiative transfer in IFS

- Satellite radiance observations currently make up around 95% of total observation numbers

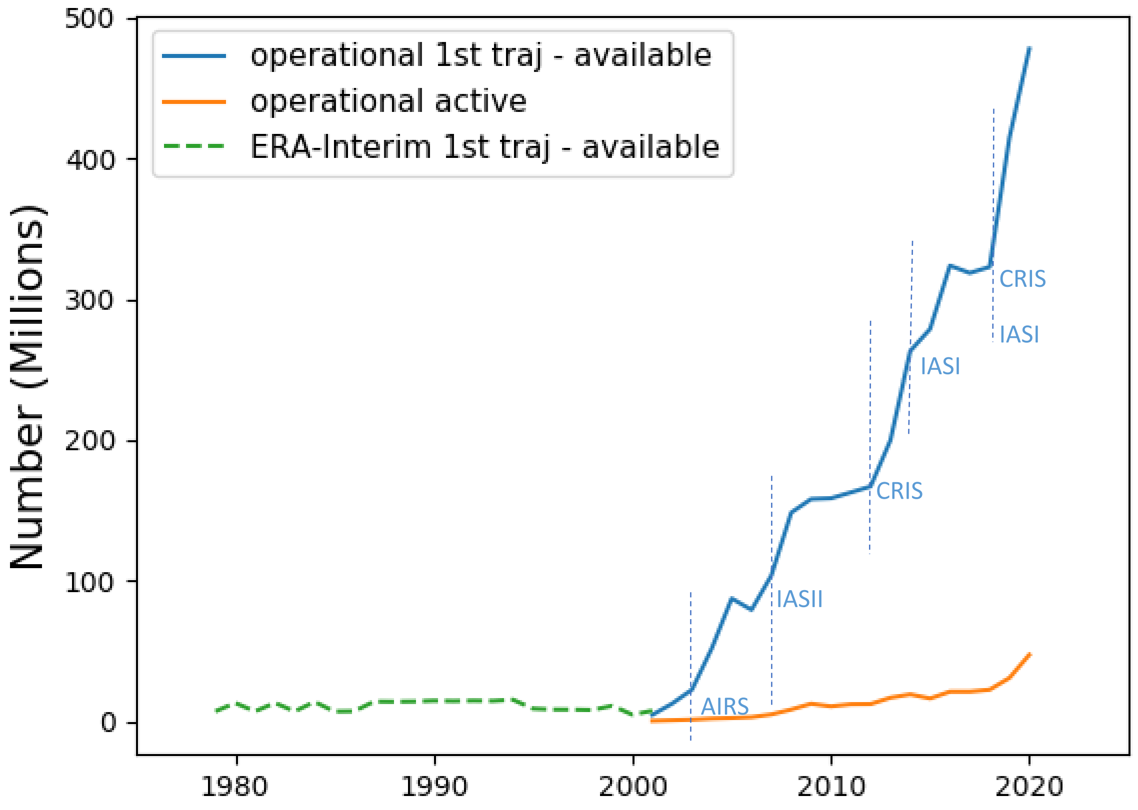
- How much time is spent running radiative transfer calculations inside IFS in HRES configuration?

  - Only around **7 seconds** is spent running RTTOV for clear sky radiances

  - Around **10 seconds** running RTTOV SCATT

(0.5% of total 1 hour runtime of IFS in the HRES configuration)

Context: trends in observation numbers



- Explosion in availability of satellite data since 2001.
- Driven primarily from infra-red hyper-spectral sounders.
- Only around 5-10% are actively assimilated:
  - most are cloud-screened, thinned or blacklisted.
- Trend expected to continue with MTG-IRS, IASI-NG

Context: size of model increasing *faster* than number of observations

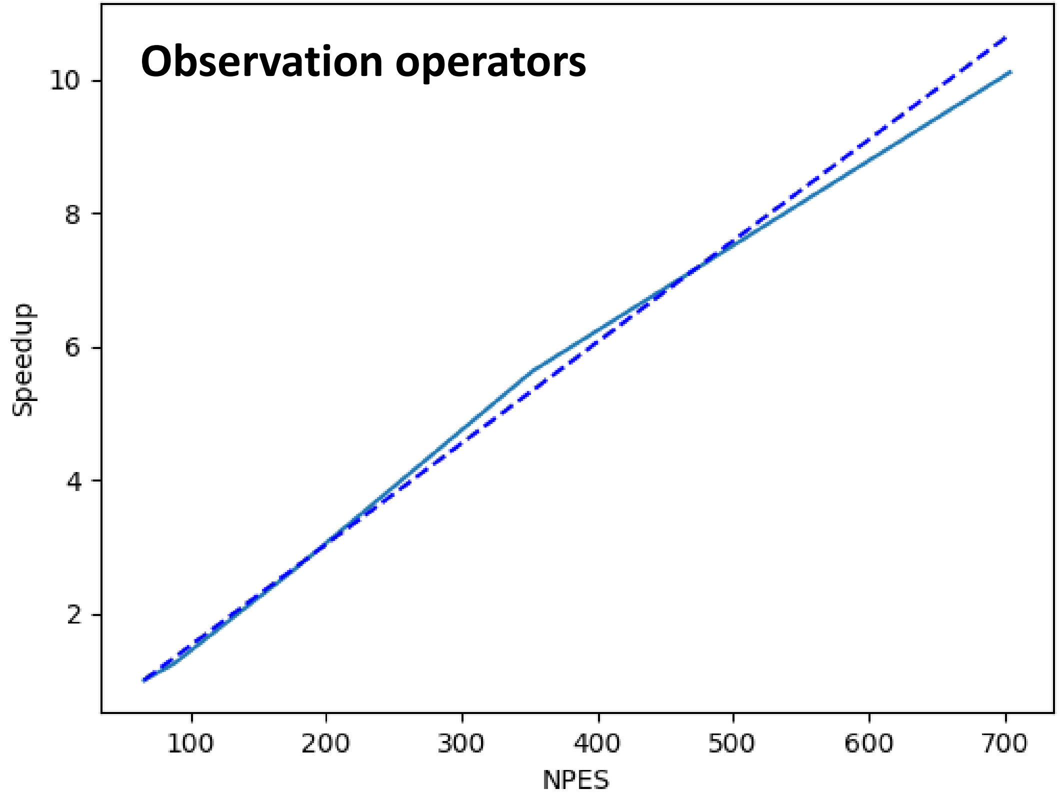
- Between 1997 (T213 L31) and 2016 (TCo1279 upgrade)

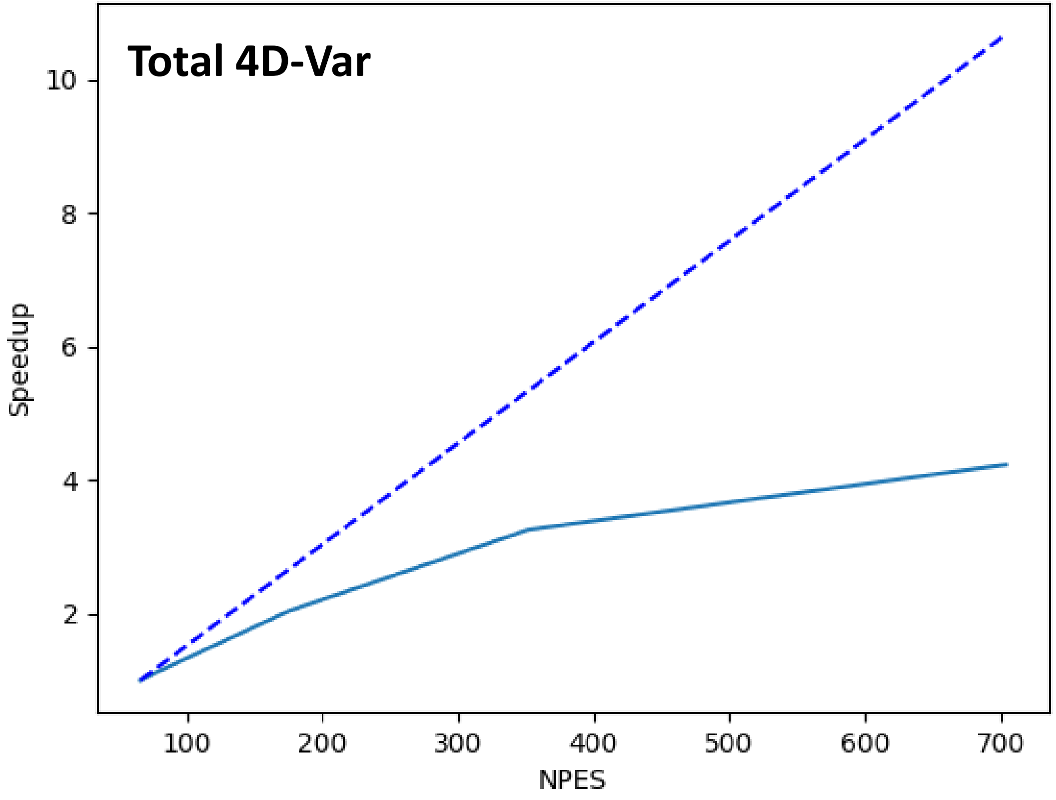
  - the number of model grid points * timesteps has increased **x843**
  - the number of cores used to run IFS has increased **x960**

  - the number of available observations has increased **x321**
  - the number of assimilated observations has increased **x256**

By 2016, each core was only processing around 33% the number of observations that were processed in 1997.

Why is radiative transfer so cheap in IFS?

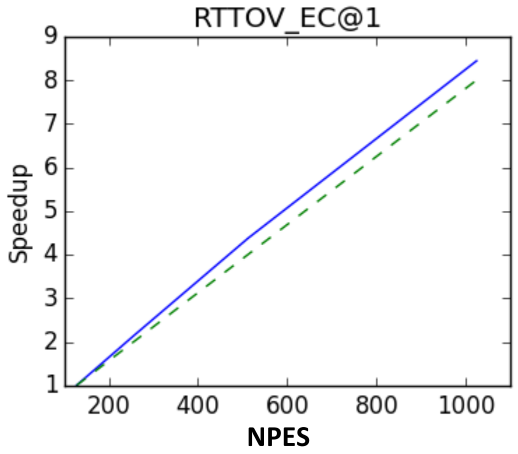


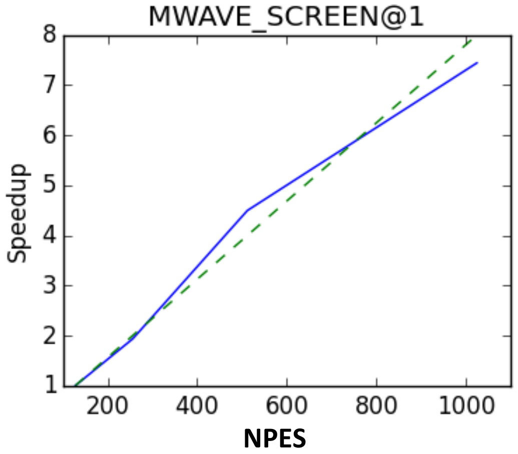


Perfect scalability

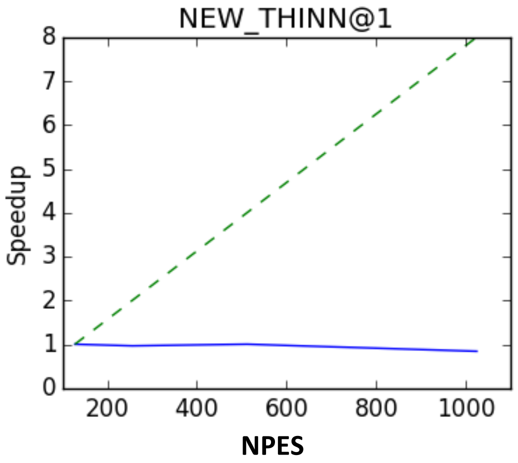
As NPES is increased, the total cost becomes dominated by those aspects which do not scale well

Most observation processing scales very well





- But, some aspects do not scale:



Bottleneck : single PE global sort of the data

Other observation-related processing inside IFS

Direct observation processing costs:

- IO to read observations **5s**
- Further observation pre-processing **15s**
- MPI communication between model data layout and observation layout **20s**
- Observation operators to calculate observation equivalents from model **25s**
- Background dependent quality control + thinning **30s**
- IO to write feedback data (departures, status flags etc) **10s**

Indirect observation costs:

- Runtime of 4D-Var minimisation is driven by the number of iterations required to produce a good fit to the observations

Summary of observation cost in IFS

- Observation numbers have increased dramatically
- **BUT** the size of the model has increased *even more* dramatically
- Number of nodes is driven by the size of the model
- Most observation processing is inherently scalable and so the total cost of IFS becomes dominated by other aspects which do not scale so well
- Caveat : in lower resolution configurations, the model cost is decreased and the observation cost is increased (running on fewer nodes)

e.g. in Ensemble of Data Assimilations observations account for a larger percentage of the total cost

Predicting computational cost in coming years

- *<u>If current trends continue</u>*, then the direct cost of observation processing in IFS will continue to reduce

- **But what if current trends don't continue? e.g.**

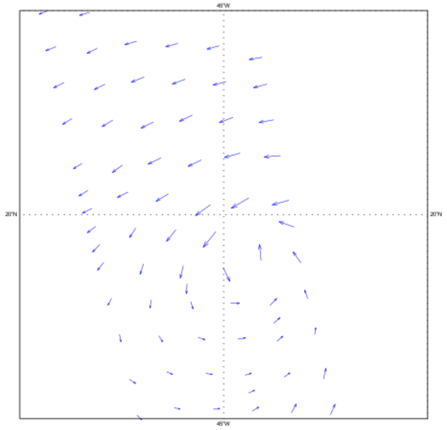
  - If we assimilate a greater proportion of the available data;

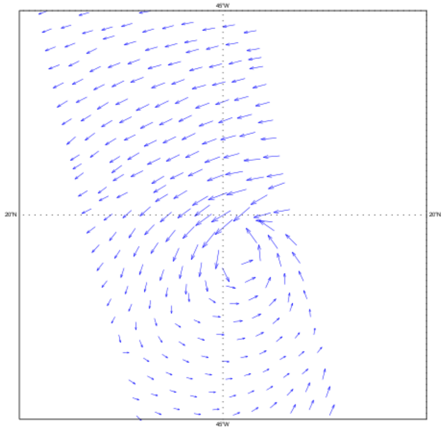
    - Currently, we assimilate only a small percentage (~5%) of the available observations

    - Representation of spatial/temporal error correlations may lead to a greater percentage being used in 4D-Var

  - If the scalability of other parts of the system is improved, then the observation components will take up a greater proportion of the total cost

  - If there is a rapid increase in the number of observations

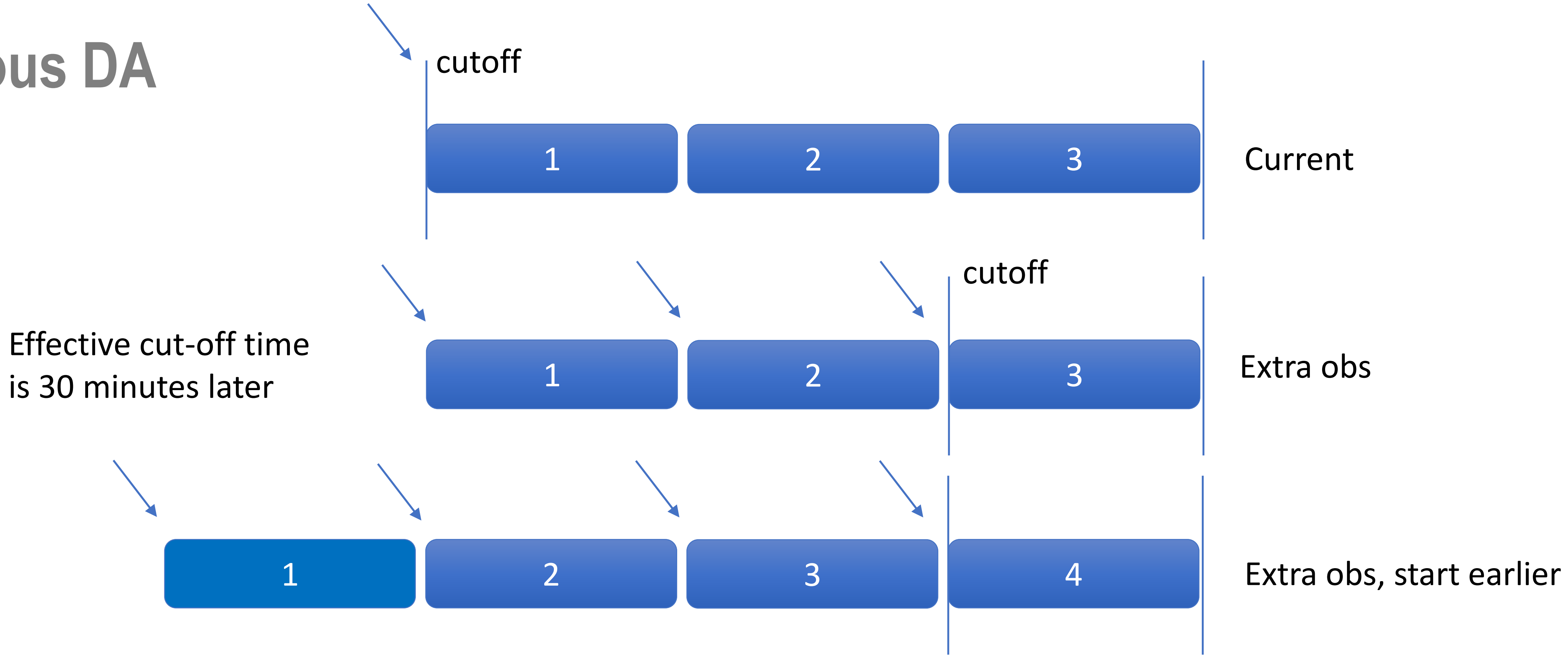




Thanks: Giovanna De Chiara

If IFS takes 1-hour to run can anything else be done to allow more recent observations to be assimilated?

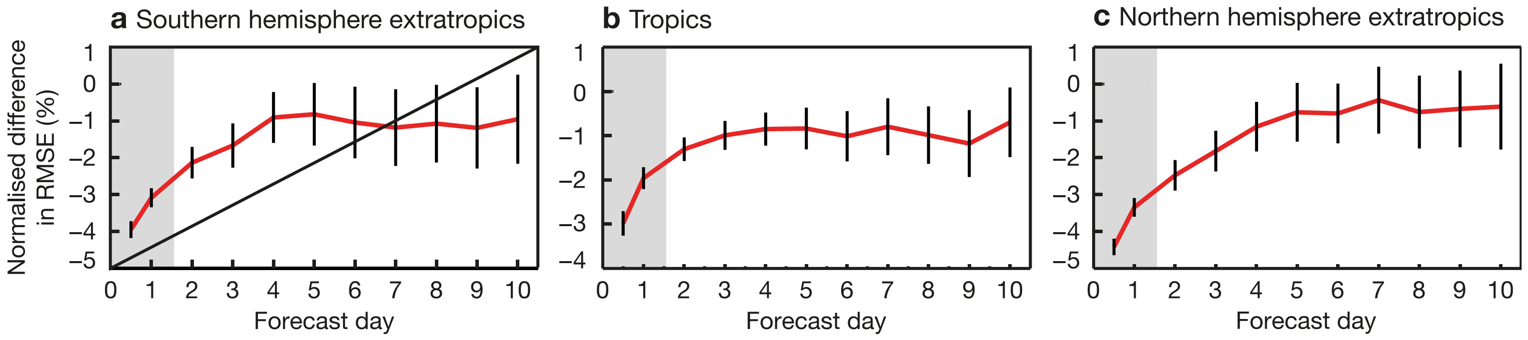
Continuous DA



- Start running data assimilation **before** all of the observations have arrived
  - Most of the assimilation is removed from the time critical path
  - DA can run for longer, starting earlier e.g. can add extra outer loops

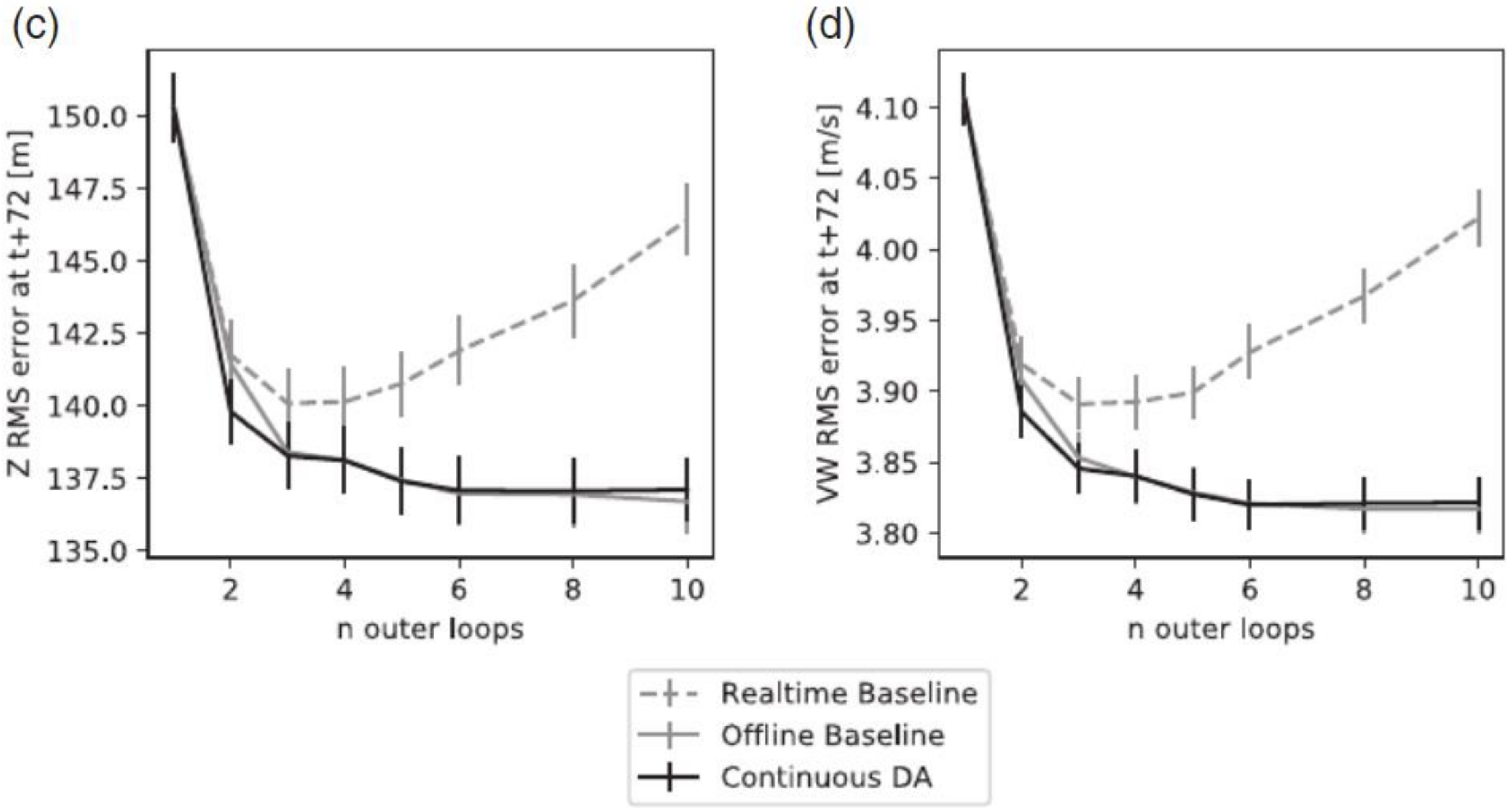
Continuous DA

- Introduced in IFS CY46R1 (2019)
- At the time that the analysis is complete, the most recent assimilated observation is 35 minutes old compared to 120 minutes previously.
- Improved medium range forecast scores approx. 2% reduction in RMS error



Normalised difference in RMS error for vector wind

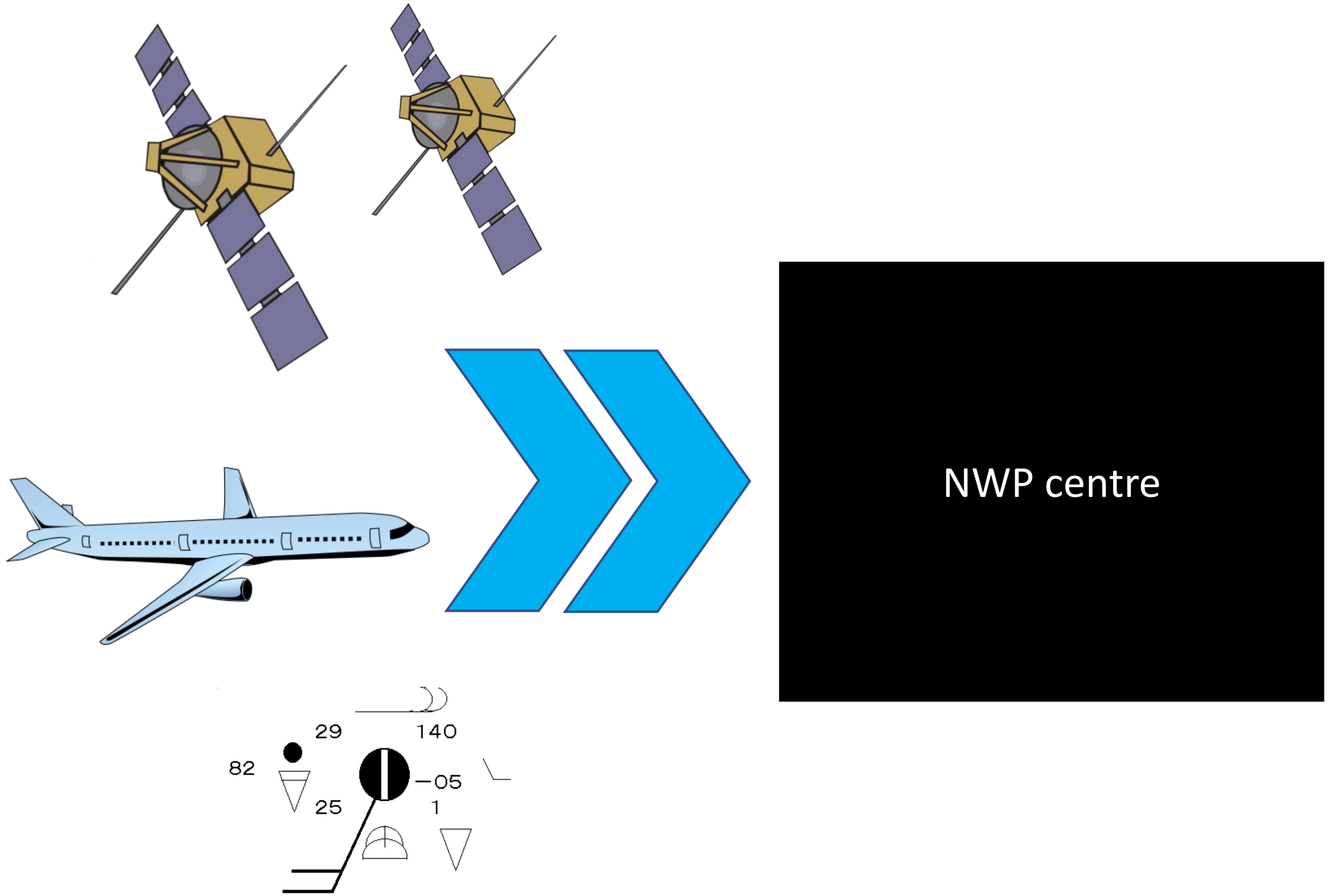
Lean, P, Hólm, EV, Bonavita, M, Bormann, N, McNally, AP, Järvinen, H. Continuous data assimilation for global numerical weather prediction. *Q J R Meteorol Soc*. 2021; 147: 273– 288.



Lean, P, Hólm, EV, Bonavita, M, Bormann, N, McNally, AP, Järvinen, H. Continuous data assimilation for global numerical weather prediction. *Q J R Meteorol Soc*. 2021; 147: 273– 288.

Robustness of observation processing systems

Incoming observations



- We have no control over what observations arrive
- Corrupt / unusual / unexpected data have potential to cause failures unless processing systems are **robust**
- Operations is the first suite to process these incoming observations
- Delays in forecast production reduce the value of those forecasts to users

Thanks: Tomas Kral

Examples of failures caused by observations

- In 2014, during transition to BUFR data, a station operator increased the sampling frequency during radiosonde ascents
- Started sending 50,000+ values in each ascent (typically tens or hundreds)
- Parts of the assimilation code ran out of stack memory for the arrays processing this ascent leading to a hard to debug memory error.

Now, additional checks + vertical thinning prevent this from happening again

Thanks: Cornel Soci

Examples of failures caused by observations

- Variational Bias Correction system (VarBC) stores bias correction coefficients for each ship in ASCII files for use in the next cycle
- A ship report was sent with a carriage return character at the end of the station identifier
- This caused a new line to be generated midway through the VarBC coefficient file leading to a corruption

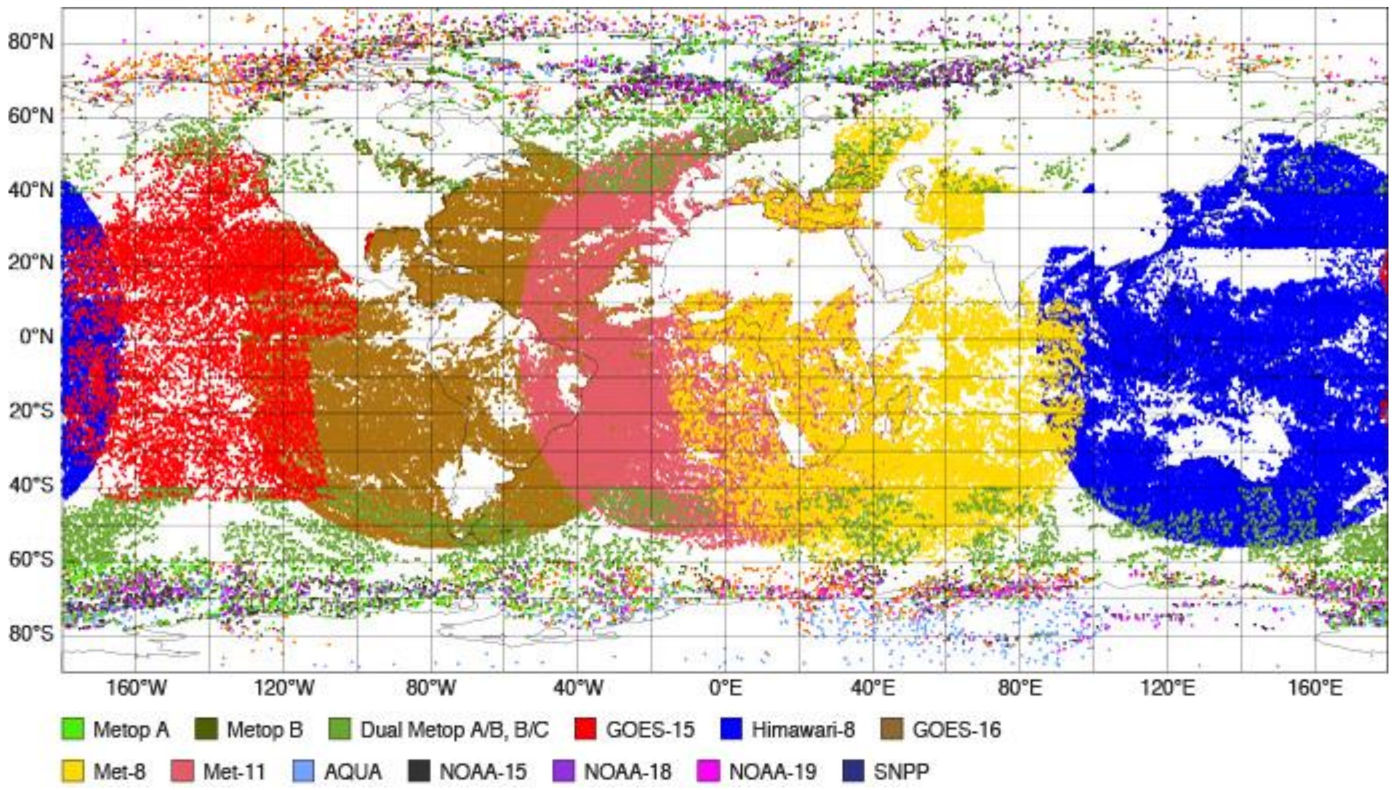
Now, identifiers undergo cleaning to remove special characters

Thanks: Francis Warwick

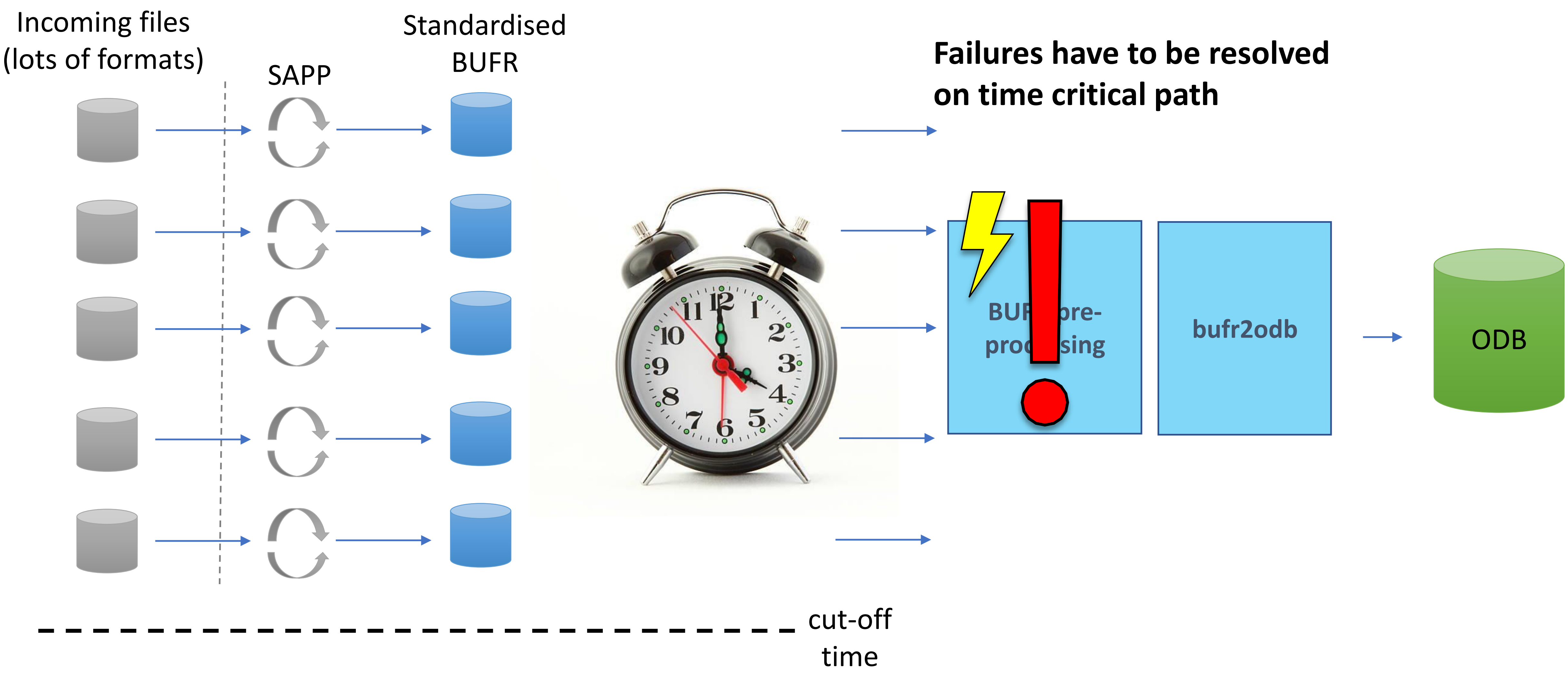
Examples of failures caused by observations

- Atmospheric Motion Vector report came in with an assigned pressure of 0.
- Part of the IFS code performed LOG(0) leading to a Floating Point Exception

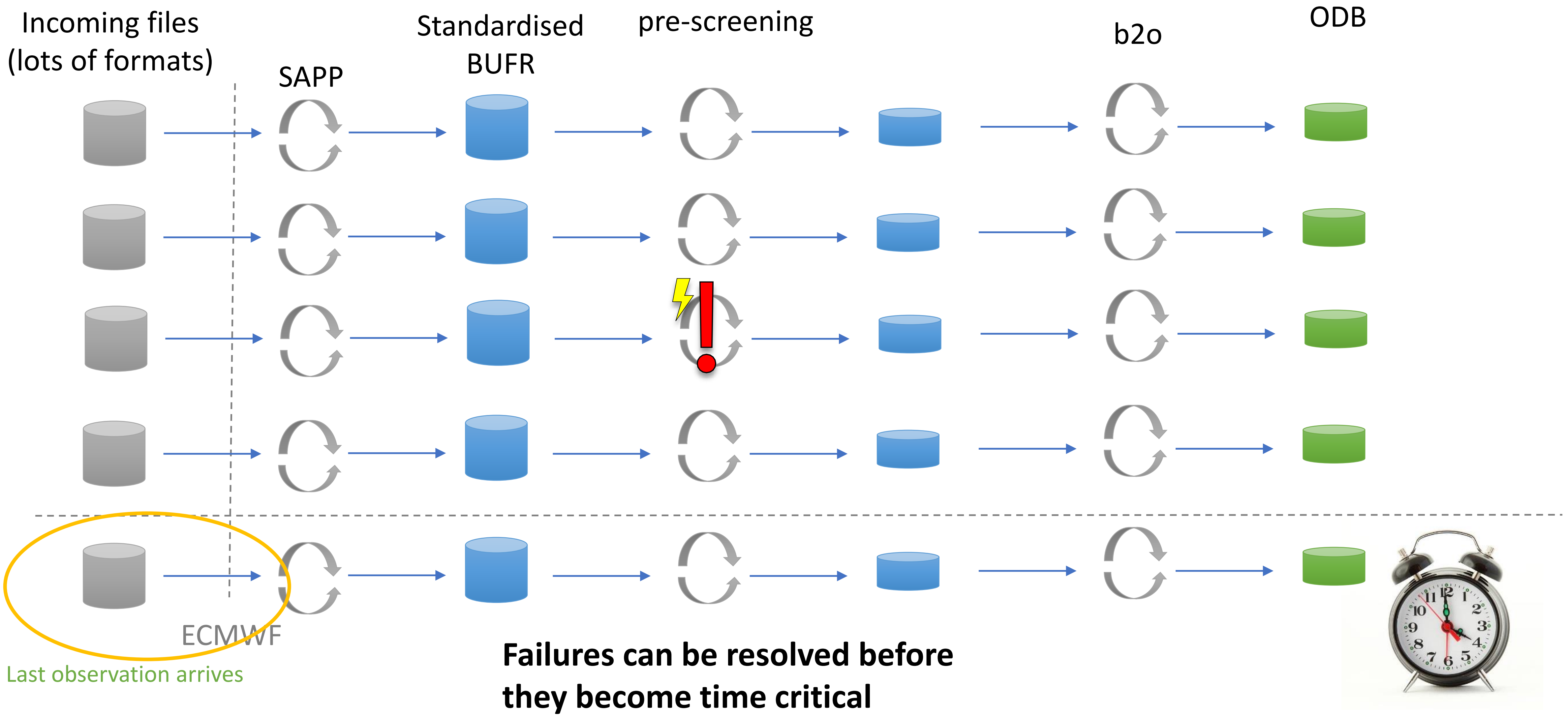
Now the ranges of assigned pressures undergo a sanity check to prevent the data flowing through that part of the code



Standard pre-processing failures



Continuous Observation Pre-processing (COPE)



Failures can be resolved before they become time critical

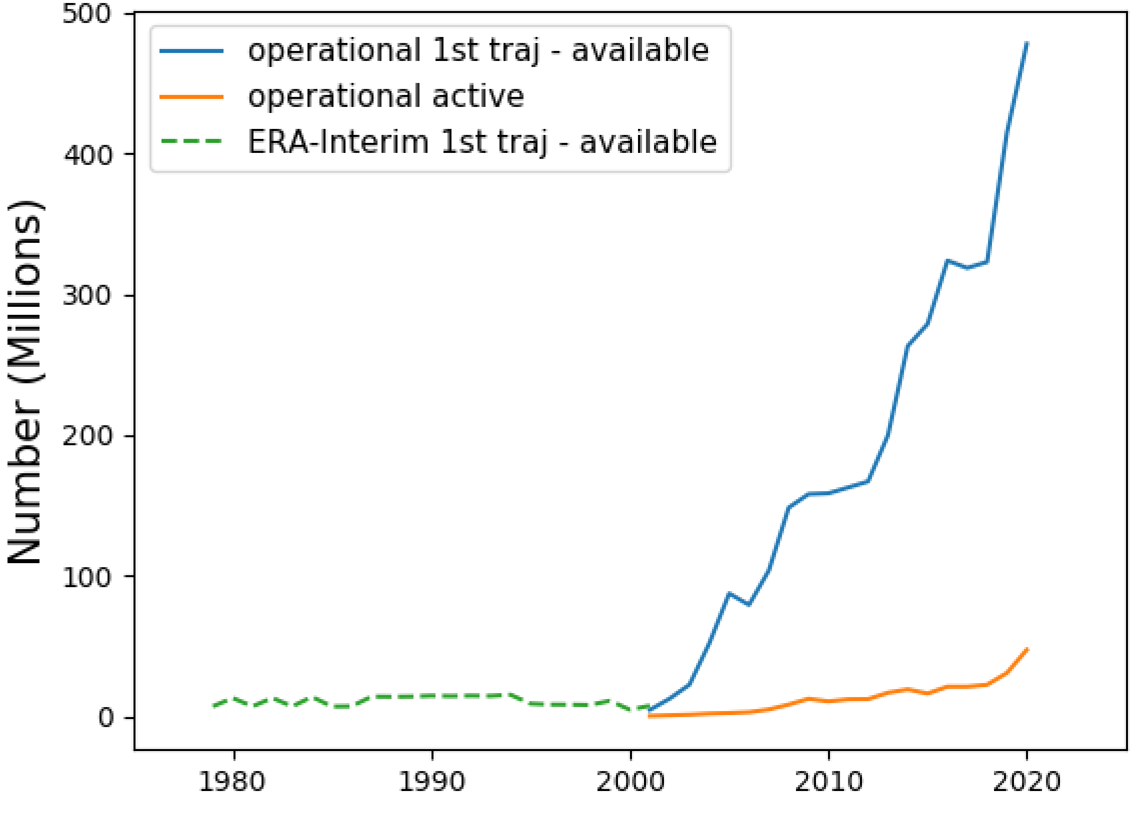
Improved robustness

- Processing systems must be robust to handle the unique challenges that observations pose
- Failures lead to delays in forecast production, which reduces their value to end users
- The upcoming move towards Continuous Observation Processing at ECMWF aims to reduce the number of delays caused by observations

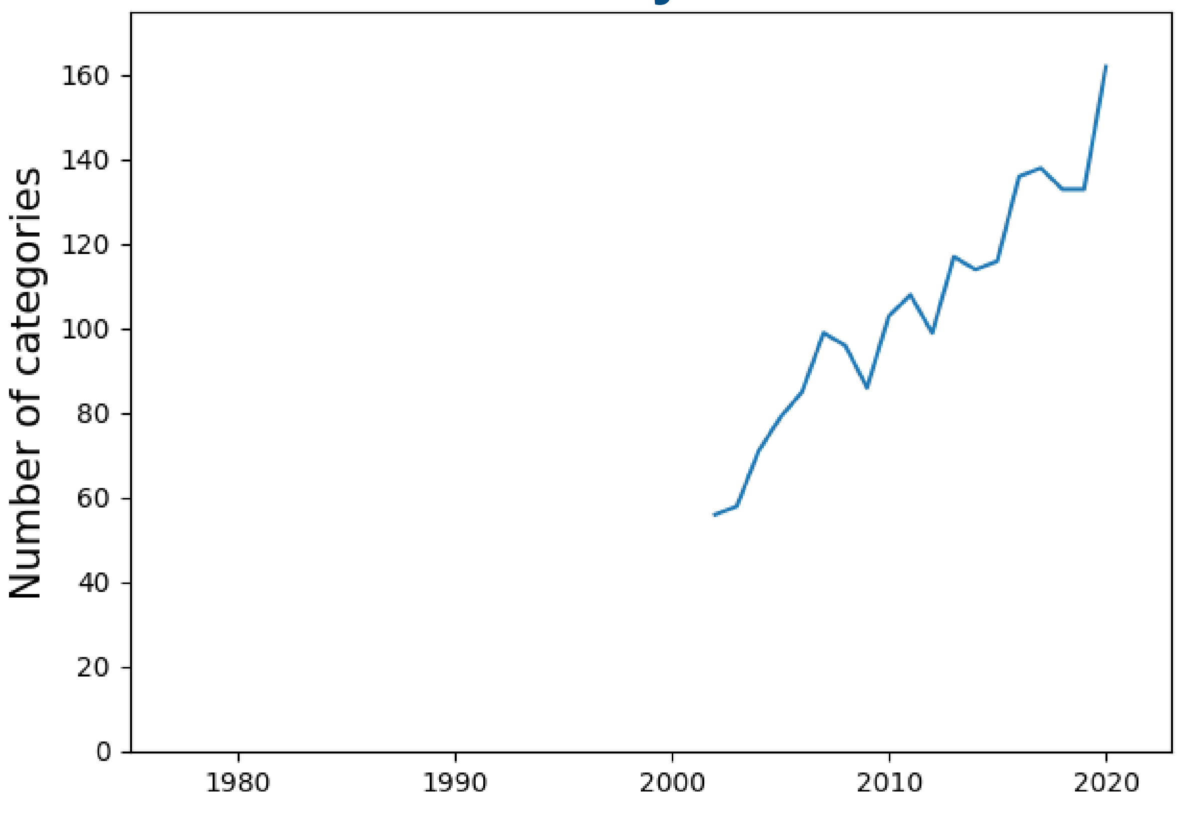
Handling the ever increasing *variety* of observations

Observation trends

Number



Variety



Increasing diversity of observing platforms

- Trend for increasing variety of observations is expected to continue (or accelerate) with the proliferation of small satellites and constellations, citizen obs and other novel platforms

- New platforms may have shorter life times than those in the past

- If we are to take advantage of this valuable new data, we need our systems to be **flexible** and **easy to use** so that new data sources can be added and evaluated rapidly

Thanks: Phil Browne Image: MEOP

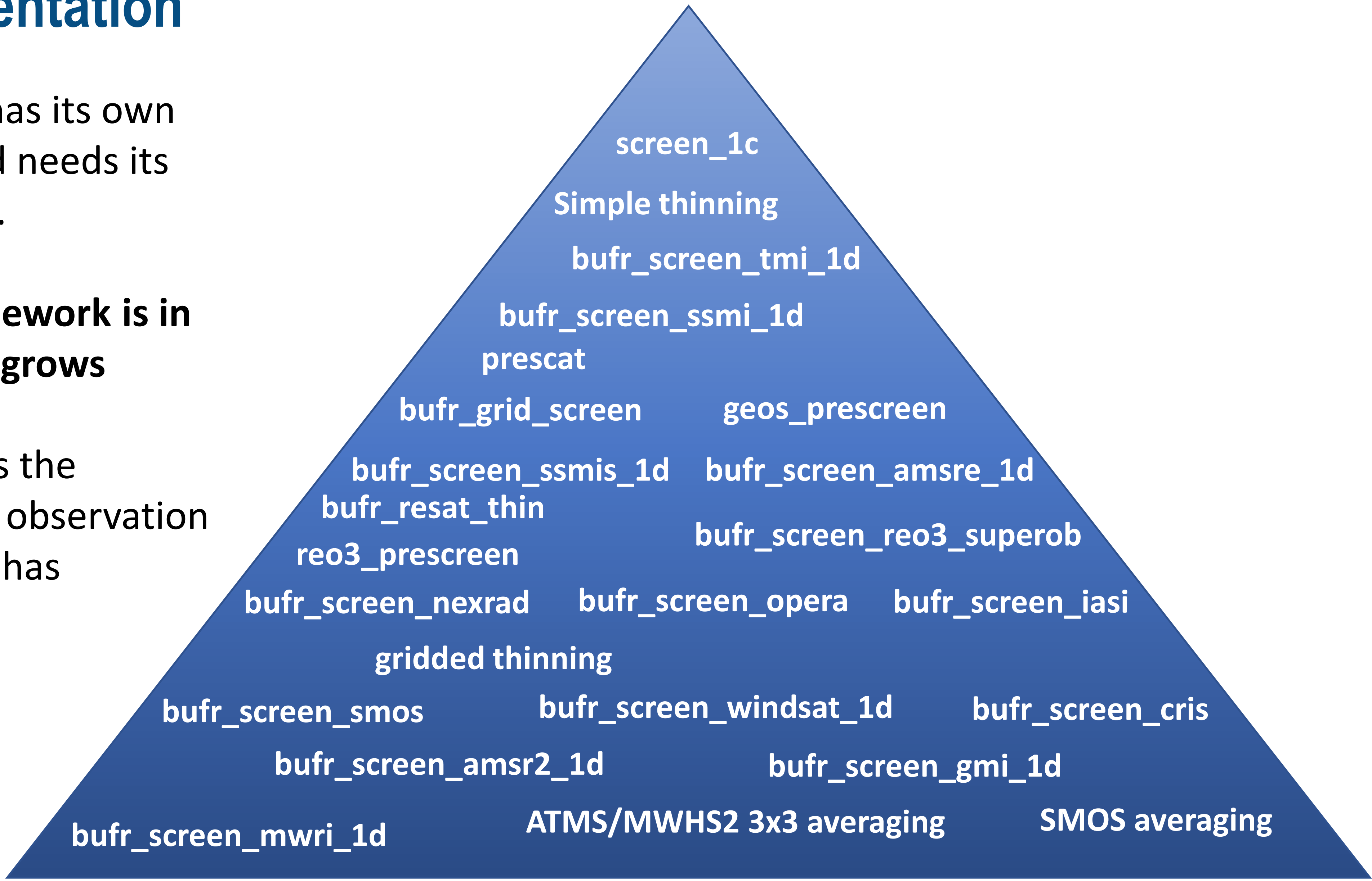
Image: ESA

Increasing fragmentation

- Each observation type has its own unique set of quirks and needs its own custom processing.

- **If no standardised framework is in place, then complexity grows**

- Example: Over the years the complexity of ECMWF's observation pre-processing systems has increased dramatically.



Standardised frameworks: structure + format

- **All observation types are different:**
  - Unique meta-data
  - Data structured differently
  - Often different file formats
- **Standardisation requires governance e.g. WMO level**
  - Balance involved - strong governance can become a bottleneck for rapid developments

ODB

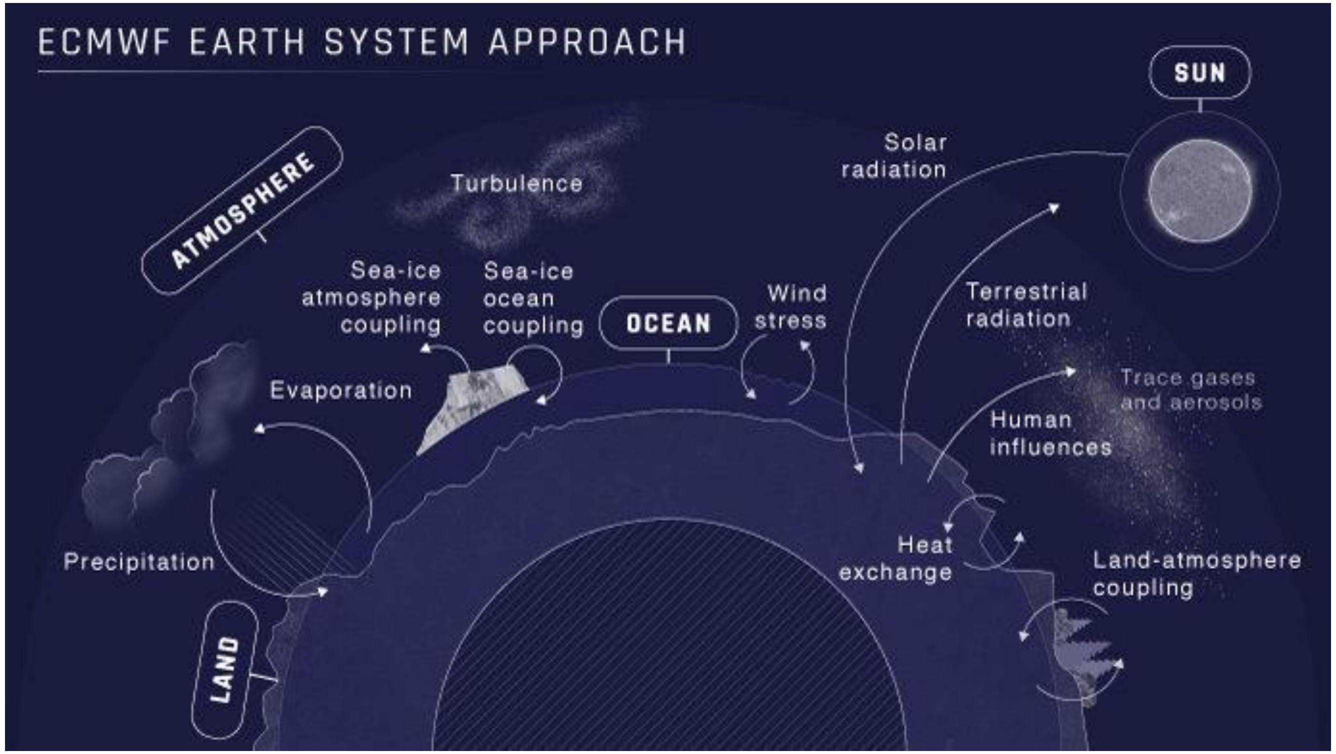
Standardised structure / schema

Standardised file format

Need for standardised frameworks

- As NWP community move towards **coupled Earth System** data assimilation systems we also need to think about the extent to which we standardise ways of storing and interacting with observational data in the different parts of the system

- **Atmosphere**
- **Ocean**
- **Land**
- **Wave**



Summary: observations pose unique challenges for real-time NWP

- Observation delivery timeliness and processing timeliness are both critical and determine the value that can extracted for real-time forecasting
- Most observation processing is inherently scalable; the rate of increase of observation numbers has not kept pace with the increase in model cost meaning that most observation processing is relatively inexpensive (especially observation operators)
- As the only external input to the system, observations have unique opportunities to cause operational failures; systems must be robust
- At ECMWF, the move towards continuous pre-processing saves valuable minutes on the critical path and should lead to improved robustness + fewer delays in forecast production
- Standardised frameworks are needed to handle the increasing variety of observations and to make it easier to rapidly add new data sources into coupled systems