On reanalysis production and quality assurance

Dinand Schepers

Bill Bell, Paul Berrisford, Hans Hersbach, Andras Horanyi, Joaquin Munoz-Sabater, Julien Nicolas, Paul Poli, Raluca Radu, Adrian Simmons, Cornel Soci

And many others that contributed to ERA5

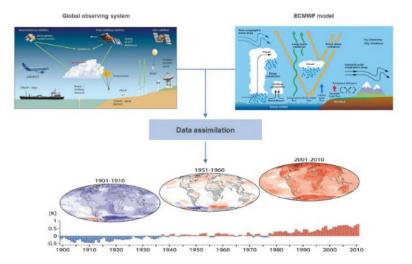


Reconstruction of the past weather&climate:

- Input: integrator of all observations we have
- ✓ **Output:** convenient and as accurate as possible 'maps without gaps' of 3D atmosphere (+ other domains)

State-of-the-art:

- Redo historical weather using a modern but fixed NWP system
- ✓ For extended period back in time, but at lower resolution
- Made available to users in a convenient way
- Maintained close to NRT



All observations we have

0.75 (1979) – 26 Million (2021) obs per day Over 200 types of reports

Reprocessed data sets (In blue)

Radiances: SSM/I brightness temp from CM-SAF MSG from EUMETSAT Atmospheric motion vector winds: METEOSAT, GMS/GOES-9/MTSAT, GOES-8 to 15, AVHRR METOP and NOAA Scatterometers: ASCAT-A (EUMETSAT), ERS 1/2 soil moisture (ESA) Radio Occultation: COSMIC, CHAMP, GRACE, SAC-C, TERRASAR-x (UCAR) Ozone: NIMBUS-7, EP TOMS, ERS-2 GOME, ENVISAT SCIAMACHY, Aura MLS, OMI, MIPAS, SBUV Wave Height: ERS-1,ERS-2, Envisat, Jason

Improved data usage compared to ERA-I (in blue)

all-sky vs clear-sky assimilation, latest radiative transfer function, corrections, extended variational bias control

Marslist: 0001 ea oper marser mfb 1978-01-01 2015-12-3 Ground-Based Radar:2500 Ground-Based Radar:25001 BUFR DRIFTING BUOYS:16084 MOORED BUOYS:16083 VIGOS AMDAR-16082 BUFR LAND SYNOP:1607 BUFR SHIP SYNOP:16074 BUFR LAND PILOT:16068 ACARs with mixing ratio:16067 Automatic METAR:16065 BUFR SHIP TEMP:16046 BUFR LAND TEMP:16045 ACARS:16030 AMDAR:16029 AIREP:16026 and TEMP:16022 obile TEMP:16021 DROP Sonde:16020 TEMP SHIP:16019 European Wind Profiler 1601 American Wind Profiler 2:1601 American Wind Profiler 1.1601 PILOT SHIP:16014 Land PILOT:16013 DRIBU-TESAC:1601 DRIBU-BATHY:16010 Abbreviated SHIP:16009 SHIP:16008 Reduced SHIP:1600 Automatic SHIP:16006 DRIBU:16005 METAR:16004 Manual Land SYNOP:16002 Automatic Land SYNOP:16001 GOES 16 GEOS Rad GOES 15 GEOS Rad. GOES 14 GEOS Rad GOES 13 GEOS Rad GOES 12 GEOS Rad GOES 11 GEOS Rad GOES 10 GEOS Rad GOES 9 GEOS Rad. GOES 8 GEOS Rad METEOSAT 11 GEOS Rad. Allsky METEOSAT 10 GEOS Rad. Allsky METEOSAT 9 GEOS Rad. Allsky METEOSAT 8 GEOS Rad. Allsky IETEOSAT 7 GEOS Rad. METEOSAT 5 GEOS Rad. Himawari 8 GEOS Rad. MTSAT-2 GEOS Rad. MTSAT-1R GEOS Rad NOAA-20 CRIS Rad. METOP-B IASI Rad. EUROPEAN CENTRE FOR MEDIUM-RANGE WEATHER FORECASTS Multi-spectral infrare

Latest instruments

IASI, ASCAT, ATMS, CrIS, MWHS, Himawari, ... TAMDAR, MODE-S

ERA5T vs ECMWF NWP operations:

ERA5 only: AURA MLS ERA5 not possible: Saphir, Aeolus, TEMP descent, SPIRE

using a modern but fixed NWP system (I)

Every new generation of reanalysis benefits from years of research at ECMWF by about 110 staff in ECMWF Research Department

Projected benefits of ERA6 over ERA5

- Improved realism of near-surface quantities and radiative forcing
 - **vegetation** cover and type, **LAI**, **lake** cover and properties, the **urban tile**
 - New, and more species of, aerosols and GHG's

Improved ocean wave physics

wave physics upgrade

4D-Var DA developments

- Dynamically evolving **B**_{cli}
- EDA developments

Land DA developments

Revised conventional observation feedback for T2m, RH2m and snow analyses

Ocean DA developments

ERA6 will be forced by ORAS6 and OCEAN6 fields

using a modern but fixed NWP system (II)

Add additional, specific functionalities

• Support for non-current observations Revive and maintain IFS-support for *old*, *rescued* or *reprocessed* observation types

• Optimization and setup System optimization for maximum throughput and high-volume analyses archiving

Postprocessing and archiving

E.g. monthly means, vertical integrals, accumulations and high-frequency archiving of analyses of an extended parameter set.

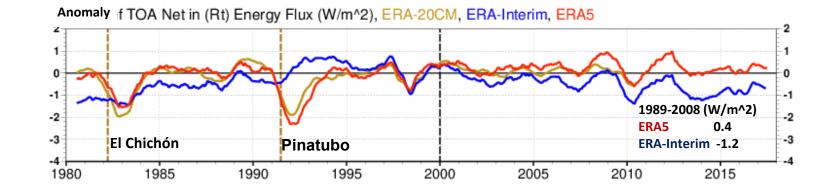
The challenge is to do this in a maintainable and sustainable way

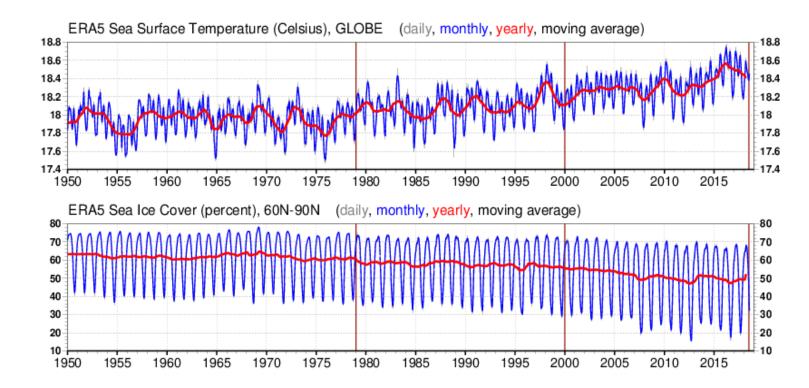
Accurate and for extended period back in time

CMIP5 recommended data sets

Total solar irradiance, greenhouse gases, ozone, aerosols (including volcanic)

(Prepared in the ERA-CLIM project, ERA-20CM)





SST and sea ice cover

Carefully selected from OSTIA, OSI-SAF and HadISST2 (Hadley Centre, *ERA-CLIM*)

Different ensemble members use different SST realizations

(Hirahara et. al., 2016)

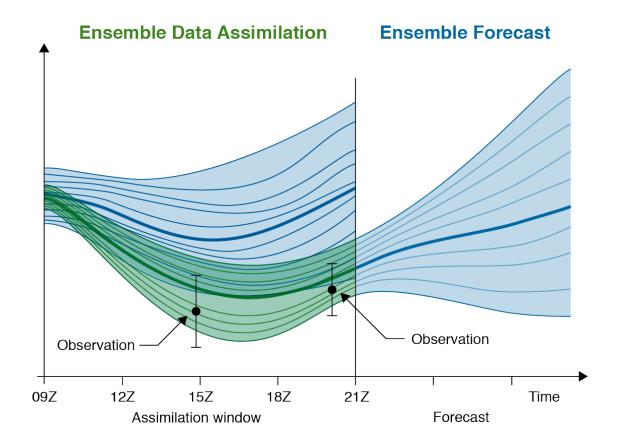
System design

• Deterministic (HRES)

Maximum affordable resolution

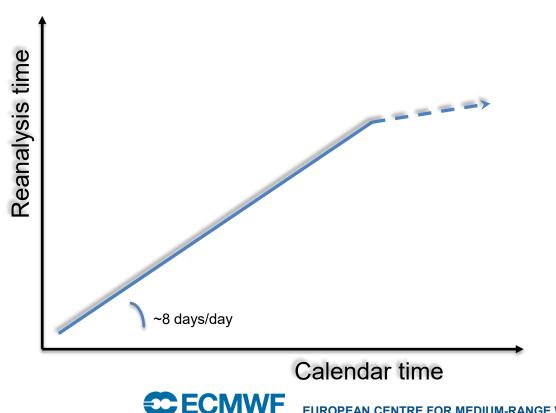
Ensemble Data assimilation (EDA)

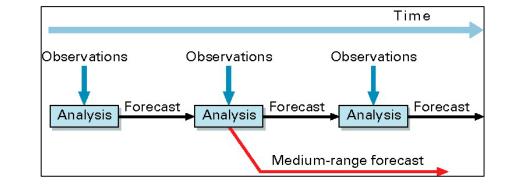
Reduced resolution, relatively few members Provides flow-dependent B for the HRES Provides uncertainty product in and of itself



Projected production speed and resource requirements

Identify and resolve bottlenecks for continuous analysis cycling, with the goal of maximizing throughput and efficiency





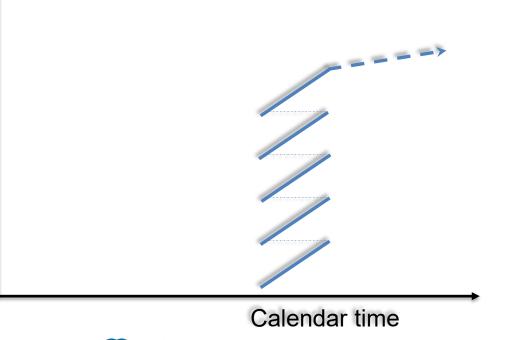


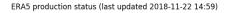
Projected production speed and resource requirements

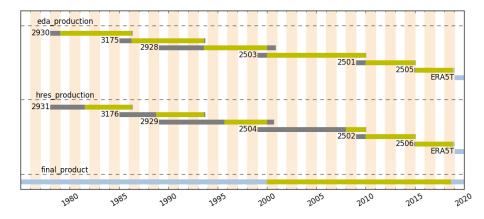
Production of modern reanalyses is generally parallelized into production streams

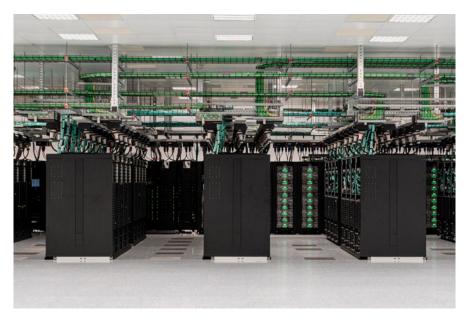
Trading overall production time against instantaneous (compute) resource requirements

Reanalysis time







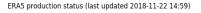


Consideration for parallel production

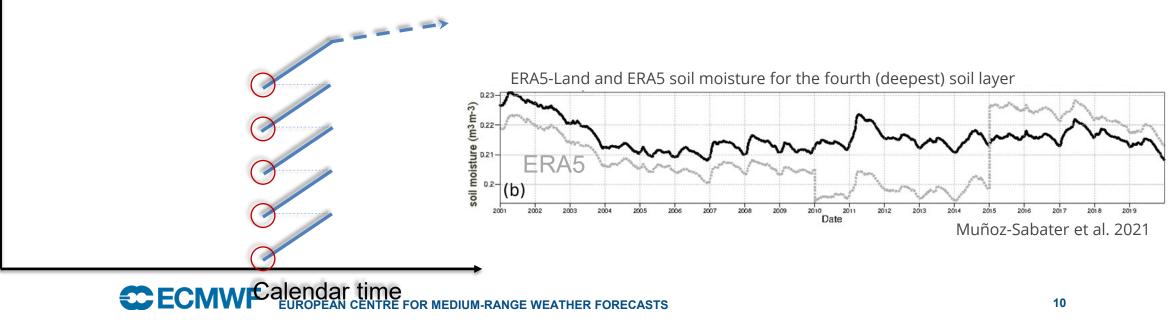
Added costs of spin-up for each stream

Reanalysis time

Initialization of poorly observed and/or slowly evolving parameters (e.g. deep soil moisture)



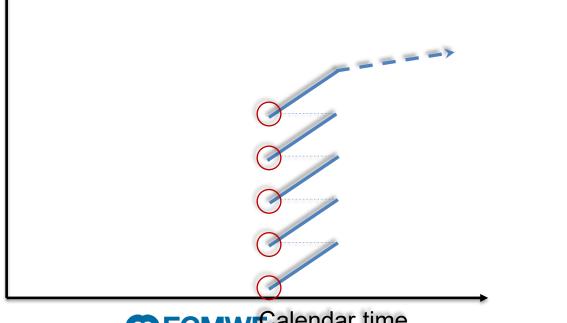




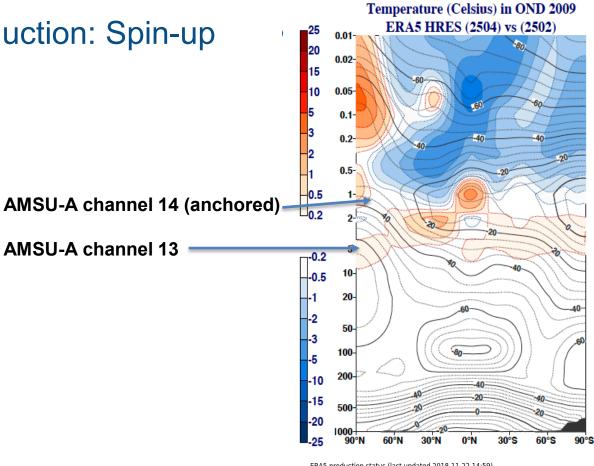
Consideration for parallel production: Spin-up

AMSU-A channel 13

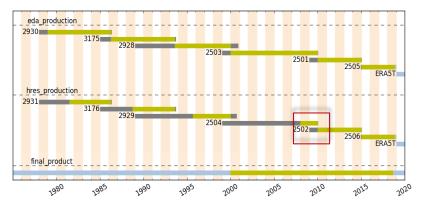
Convergence of VarBC bias estimates is not guaranteed, even after allowing for spin-up



ECMUE alendar time EUROPEAN CENTRE FOR MEDIUM-RANGE WEATHER FORECASTS

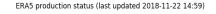


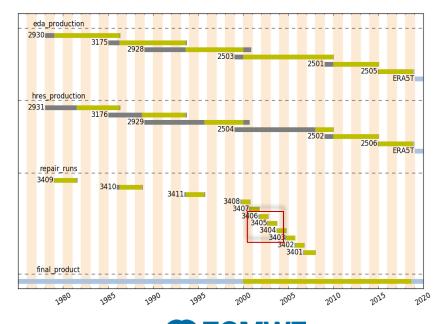
ERA5 production status (last updated 2018-11-22 14:59)

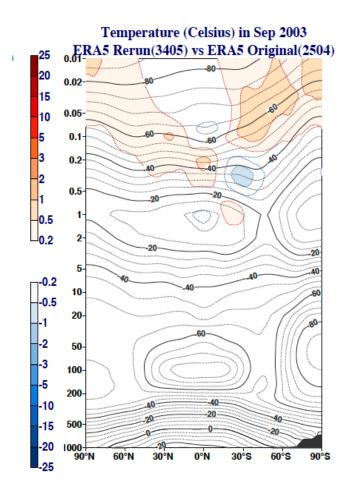


Consideration for parallel production: Spin-up

If we force convergence of VarBC bias estimates, by prescribing them to be equal, discontinuities at the seam virtually disappear









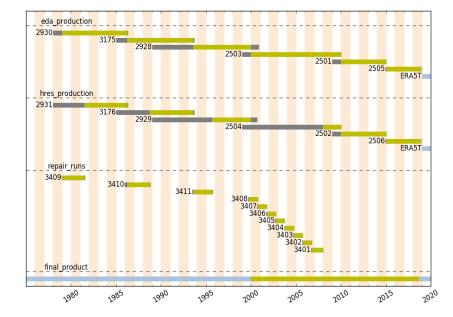
CASE STUDY: Baltic summer sea ice

The prescribed sea ice field tends to exhibit sea ice in the Baltic Sea in summertime

Corrective action:

The spurious sea ice was removed from the prescribed field and HRES repaired in a series of repair reruns

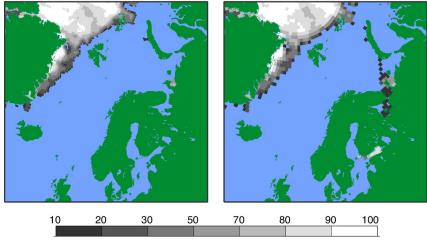
The ERA5 ensemble was left as-is. The impact on the ensemble spread is found to be small.



Sea ice cover (percent) on 27 July 2006

(a) ERA5 HRES

(b) ERA5 EDA control



Hersbach et al. 2020



ECMUF EUROPEAN CENTRE FOR MEDIUM-RANGE WEATHER FORECASTS

Reconstruction of the past weather&climate:

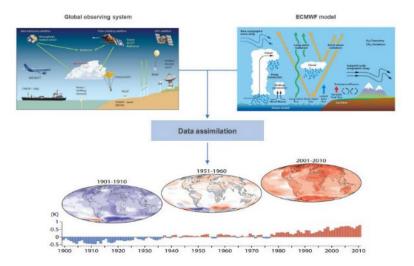
- Input: integrator of all observations we have
- Output: convenient and as accurate as possible 'maps without gaps' of 3D atmosphere (+ other domains)

State-of-the-art:

- Redo historical weather using a modern but fixed NWP system
- ✓ For extended period back in time, but at lower resolution
- Maintained close to NRT
- Made available to users in a convenient way

Multiple classes of applications:

- ✓ Study of **specific events** or phenomena:
 - accurate (3D) synoptic situation; i.e., the weather of the day
- Climate monitoring:
 - Accurate recent synoptic situation + **consistent 30-year climate**
- ✓ **Climate** applications:
 - low-frequency variability of the mean state
 - Statistics, e.g., of extremes
 ECMWF
 EUROPEAN CENTRE FOR MEDIUM-RANGE WEATHER FORECASTS



Operational monitoring during production

Observations

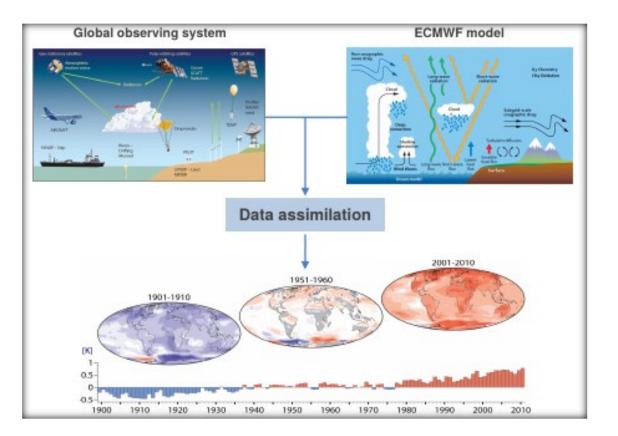
Observation system completeness

Assimilation

Assimilation diagnostics and feedback

Reanalysis product

• Quality of the final product



Routine monitoring – Observations space (I)

Completeness and consistency of observational input

We routinely monitor the observation types (reportypes) we assimilate, by comparing against a known reference

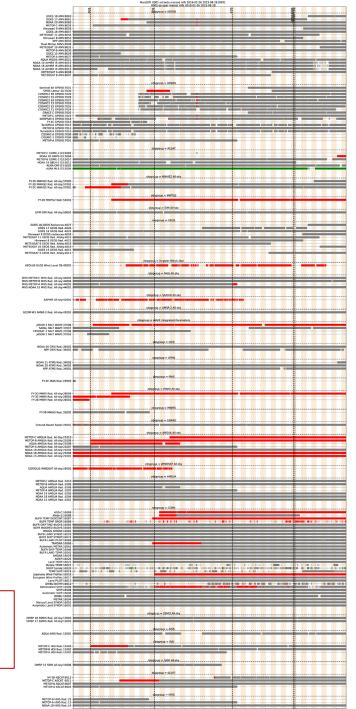
Observations used in ERA-Interim, NWP operations Provides a check on which observations make it into the assimilation

Observations used in ERA5 EDA

Provides a check on consistency with the underlying EDA

All unpacked observations Provides a check on the fetching, block listing

> Missing w.r.t. reference Additional w.r.t. reference Present in in both



CECMWF

EUROPEAN CENTRE FOR MEDIUM-RANGE WEATHER FORECASTS

Routine monitoring – Observations space (II)

Completeness and consistency of observational input

We routinely monitor the observation types (reportypes) we assimilate, by comparing against a known reference

Powerful

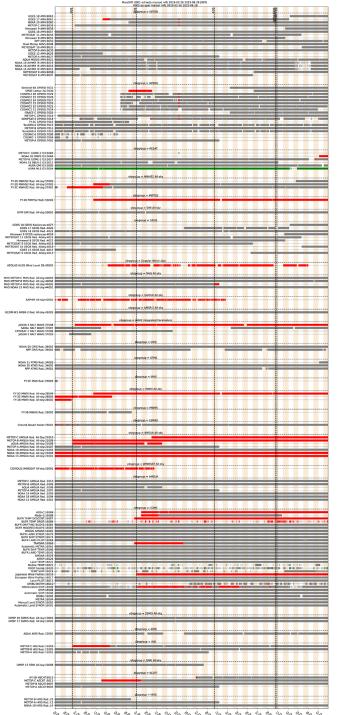
Provedes a high-level check through one single plot

Lightweight Based on MARS *LIST*, functionality in its default form

Flexible

Can ingest full ODB feedback for more granularity

Missing w.r.t. reference Additional w.r.t. reference Present in in both



CECMWF

EUROPEAN CENTRE FOR MEDIUM-RANGE WEATHER FORECASTS

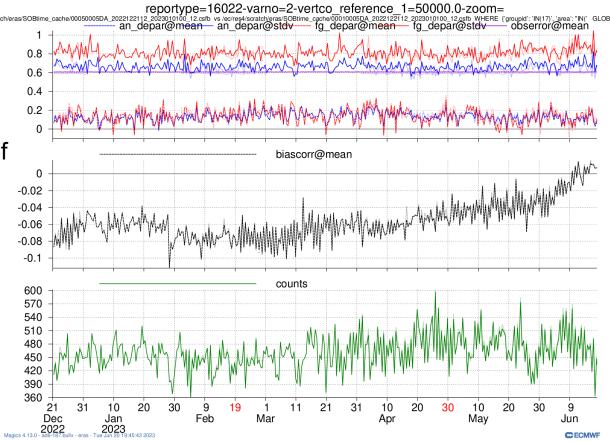
Routine monitoring – Assimilation space (I)

Ensure optimal assimilation of observations

We routinely monitor the assimilation system through *observation feedback*

Based on pre-computed, area-averaged Timeseries of

- Number of observations
- applied bias correction
- First guess and analysis departures



Routine monitoring – Assimilation space (II)

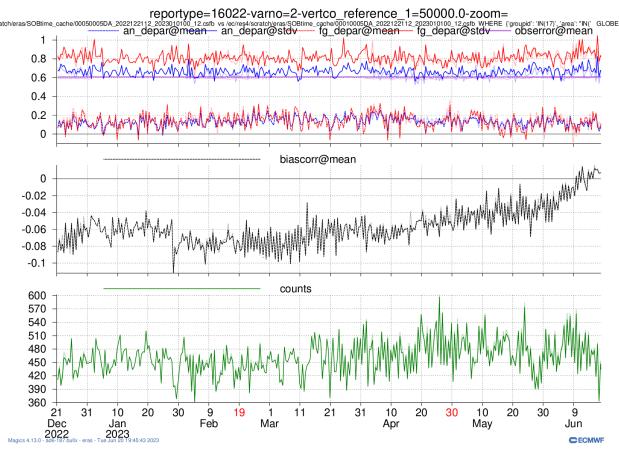
Ensure optimal assimilation of observations

We routinely monitor the assimilation system through observation feedback

Flexible, detailed and granular

- 6 predefined regions
- Per, instrument, channel, variable, datum_status
- Requires significant expertise to interpret

Typically, useful when a new observation or instrume introduced, or in case observation quality degradatio (rare)



Routine monitoring – Model space (I)

Ensure an optimal reanalysis product

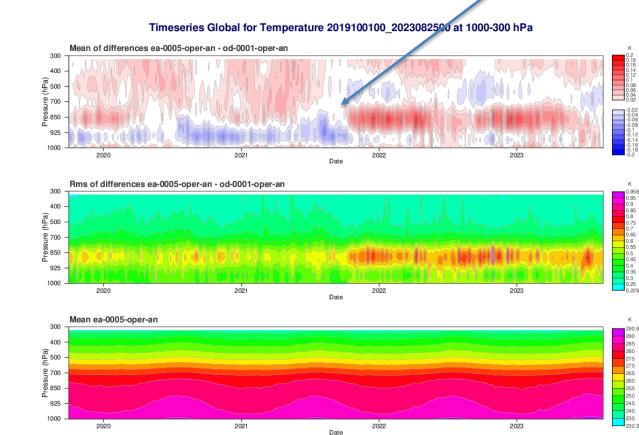
We routinely monitor the actual reanalysis product mainly through comparing the *area-averaged* analysis to a reference

ERA-Interim, NWP operations, ...

Provides a handle on gross errors in the mean state and consistency in time*

The ERA5 background forecast

Provides quasi-increments, a particularly sensitive metric for monitoring mean state



* new NWP cycle 47R3

Routine monitoring – Model space (II)

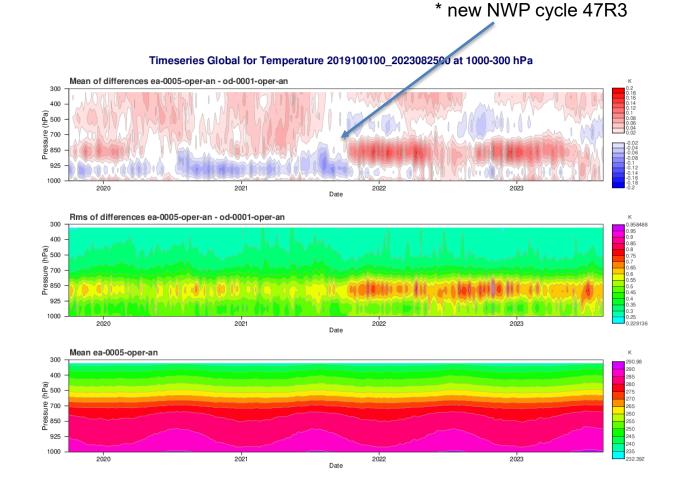
Ensure an optimal reanalysis product

We routinely monitor the actual reanalysis product mainly through comparing the *area-averaged* analysis to a reference

High level, aggregated in model space

- Area-averaged
- pair-wise pre-computed
- Fairly straight-forward to interpret

Typically, useful for catching gross errors or timeinconsistencies in the analyzed state



ECMUF EUROPEAN CENTRE FOR MEDIUM-RANGE WEATHER FORECASTS

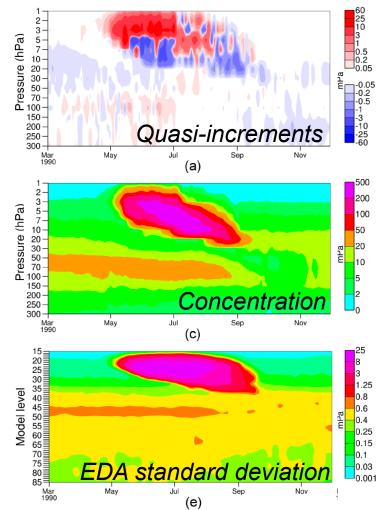
CASE STUDY: Anomalous ozone in Antarctic polar night

In polar night conditions during the 80's and 90's ERA5 produced physically unrealistic amounts of ozone in the high stratosphere

Diagnosis:

Assimilating SBUV observations above 5 hPa induces increments over the *unobserved* winter pole, adding ozone and simultaneously inflating the EDA spread. Establishing a positive feedback loop.

Area-averaged (Antarctic) ozone-



CASE STUDY: Anomalous ozone in Antarctic polar night

In polar night conditions during the 80's and 90's ERA5 produced physically unrealistic amounts of ozone in the high stratosphere

Diagnosis:

Assimilating SBUV observations above 5 hPa induces increments over the *unobserved* winter pole, adding ozone and simultaneously inflating the EDA spread. Establishing a positive feedback loop.

Corrective action:

Partial blacklisting of SBUV in ERA5 (above 5 hPa) and full blacklisting in the EDA was shown to prevent occurrence. And the corrected fields were released to the public.

Work has been done on the IFS to prevent reoccurrence in future reanalysis

Are re (hPa) Quasi-increments (hPa) 200 Concentratio 250 A standard deviat

Routine monitoring – Implementation for ERA5

Spaces ~ Calendars Create

✓ ERA5 Monitoring Tools

- ERA5 MONITORING: ALL LINKS
- > ERA5 MONITORING: MPs
- ✓ ERA5 MONITORING BRIEFINGS
- 1940 suite (scout: 3889)
- > 1950 suite (EDA: 3588/3646, HRES: 3589/3647)
- > 1958 suite (EDA:3584/3648, HRES: 3585/3649)
- > 1965 suite (EDA: 3582/3650, HRES: 3583/3651)
- 1967 test (3994/3995)
- > 1972 suite (EDA: 3580/3644 HRES: 3581/3645 Scout: 3563)
- > 1989 suite (EDA: 2928, HRES: 2929)
- BX rerun: 1939 suite (EDA: 4216, HRES: 4217)
- * BX rerun: 1944 suite (EDA: 4214, HRES: 4215)
- BX rerun: 1949 suite (EDA: 4212, HRES: 4213)
- BX rerun: 1954 suite (EDA: 4210, HRES: 4211)
- BX rerun: 1959 suite (EDA: 4043, HRES: 4044)
- BX rerun: 1964 suite (EDA: 4052, HRES: 4053)
- BX rerun: 1969 suite (EDA: 4050 ;HRES: 4051)
- BX rerun: 1974 suite (EDA: 4054, HRES: 4055)
- > COMP: ETED: 1978 suite (EDA: 2930, HRES: 2931)
- > COMPLET ID: 1935 5 lite (SD: 0.2175, FR IS: 3170)
- > COMPLETED: 1995 suite (EDA: 3177, HRES: 3178)
- > CON PL 2" EF : 19 99 SUI13 "20 A12 (U3: HRE 30.5.)4
- > COMPLETED: 2009 suite (EDA: 2501; HRES: 2502)
- > CON PLETES I E -RUN FUITES
- → Near Real-Time (NRT) suite (EDA: 2505; HRES: 2506, former 2443/2445)
- NRT suite issues
- NRT suite notes 1 (September, 2016 March, 2017)
- · NRT suite notes 2 (April, 2017 September, 2017)
- NRT suite notes 3 (October, 2017 March, 2018)
- * NRT suite notes 4 (April, 2018 October, 2018)
- NRT suite notes 5 (October, 2018 April, 2019)
- NRT suite notes 6 (April, 2019)
- NRT suite notes 7 (December 2019 December 2020)
- NRT suite notes 8 (January 2021)
- NRT suite notes 9 (January 2023)
- Summary of ERA5 infrared instruments-related issues

Summary of ERA5 microwave instruments-related issues

- * Summary of ERA5 ozone-related monitoring findings (January, 2017)
- Summary of ERA-Interim issues
- ERA5 Monitoring suite(s)
- > Tool maintenance and upgrading

Increments

Nothing new to report:

- Surface: https://intraplots.ecmwf.int/packages/era5_monitoring_model/products/ainc-sfc-0005_DA/overview/part
- Upper-air: https://intraplots.ecmwf.int/packages/era5_monitoring_model/products/ainc-upper-0005_DA?levt

Differences

Nothing new to report

- Surface switch of colour labels depending on param (already reported) also diff colours depending on the time windows i

 it seems that it is 'only' a label issue (e.g. CI 90 days vs 30 days)
- Upper-air: https://intraplots.ecmwf.int/packages/era5_monitoring_model/products/diff-upper-0005_DA?expver_ct

Upper-air anomalies

• not updated: https://intraplots.ecmwf.int/packages/era5_monitoring_model/products/upperair_anomalies?component=hresE

30-08-2023 Paul

Status: ERA5T and ERA5-LandT one day behind RT

Observations

Nothing new to report:

- Intraplots not updated: https://intraplots.ecmwf.int/packages/era5_monitoring_obs/products/obslines?stream=s0005DA8typerations/products/obslines?stream=stream=stream=stream=stream=stream=stream=stream=stream=stream=stream=stream=stream=stream=stream=stream=stream=stream=stream</pream=stre
- ERA5T HRES vs Ops: evince /perm/eras/era5_popnew/plots/OBSLINE/obsline_0005DA_mfb_vs_OPER_mfb.pdf
- ERA5T HRES vs EDA: evince /perm/eras/era5_popnew/plots/OBSLINE/obsline_0005DA_mfb_vs_0005ENDA_mfb.pd

Departures

- /t c/t ra/era5/DA_eras/sfb/anim/AAnim_2022-2023_Global_ftype-var-level.mp4 (viewed with Quicktime on mac)
- · Cornel's note from last week: MetOp-C ch8 steady noise increase even after 5/07 when obs errors were changed (see Ome
- BUFR LAND PILOT meridional wind 50 hPa: spike in SD OmB 22 August 2023 (seq 6m3s).
- Trirrat A R-) ((P Sec.) 3.4.1() m iarg i mean spike 20 August 2023 (seq 8m12s); 34 km, 22 August 2023 (seq 8m15)

Increments

Nothing new to report

- Surface: https://intraplots.ecmwf.int/packages/era5_monitoring_model/products/ainc-sfc-0005_DA/overview/param?ndays-te
- Upper-air: https://intraplots.ecmwf.int/packages/era5_monitoring_model/products/ainc-upper-0005_DA?levtype=stratpl8.nd

Differences

Nothing new to report:

- Surface: https://intraplots.ecmwf.int/packages/era5_monitoring_model/products/diff-sfc-0005_DA?expver_ctl=od-00018nd04
- Upper-air: https://intraplots.ecmwf.int/packages/era5_monitoring_model/products/diff-upper-0005_DA?expver_ctl=od-000

Upper-air anomalies

• not updated: https://intraplots.ecmwf.int/packages/era5_monitoring_model/products/upperair_anomalies?component=hree

Like Be the first to like this

Write a comment...

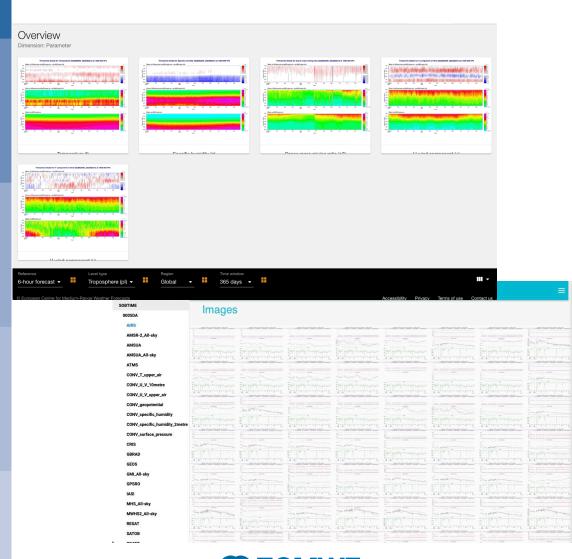
Weekly monitoring team meetings with one or more designated 'monitoring person(s)' or MP

- Up to 6 production experiments per MP
- In the observation-rich era, 1,000s of plots/metrics to monitor, per stream
- Findings are documented, discussed, and actioned if needed

Full 'production log' for all ERA5 streams

ECMWF EUROPEAN CENTRE FOR MEDIUM-RANGE WEATHER FORECASTS

Routine monitoring – Reflection and lessons



- The vast majority of issues logged were identified and actioned in manner that has not negatively affected the final ERA5 product
- Time for investigations and options for corrective actions are often limited
- Monitoring is effectively done per-stream, complicating early assessment of long-term trends and slowly varying (climate) signals
- This manner of monitoring is labor intensive, and it does not scale well
- Practically, it becomes a challenge to quickly and efficiently access all monitored metrics

To get better, we need to learn from our mistakes...

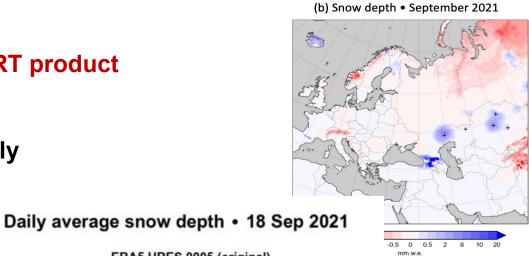


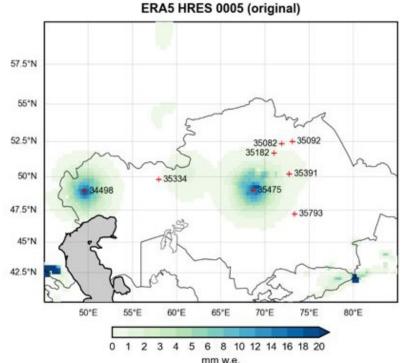
CASE STUDY: Anomalous snow depth in ERA5T

Anomalous snow depth in the *preliminary* ERA5 NRT product

<u>Hourly analysis</u> of snow depth in ERA5 inadvertently Interfered with the observation quality control

Anomalous snow depth observations were assimilated





CASE STUDY: Anomalous snow depth in ERA5T

EUROPEAN CENTRE FOR MEDIUM-RANGE WEATHER FOR

Anomalous snow depth in the preliminary ERA5 NRT

<u>Hourly analysis</u> of snow depth in ERA5 inadvertently Interfered with the observation quality control

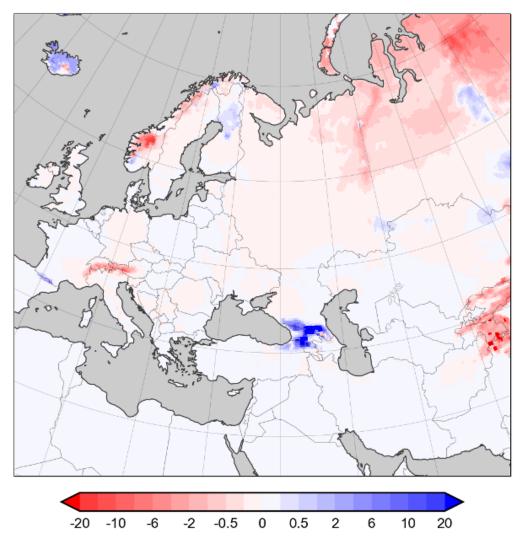
Anomalous snow depth observations were assimilated

Corrective action:

Adopt NWP Quality Control assessments, applying them to ERA5T – preventing the anomalous obs from being assimilated

Rerun and repair the affected *preliminary* released data and inform users

(b) Snow depth • September 2021



mm w.e.

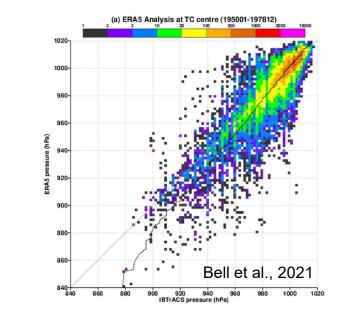
CASE STUDY: Unrealistically strong cyclones in early ERA5

Unrealistically strong cyclones up to 840 hPa and ocean waves up to 31 metres

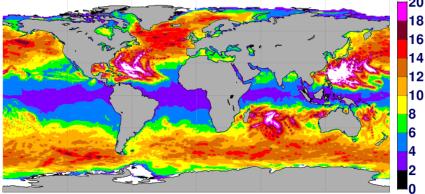
In the original back extension, IBTrACS observations had been given high weight

The data was well assimilated, departure statistics at obs location looked ok.

However, 4D-Var regularly moved the minimum of extreme cases away from the obs location



(b) Max Sign. wave height (m) for ERA5 Prelim, 1950-1978 Glob 9.1, NH 11, Trop 6.2, SH 11, MIN 0.15, MAX 31

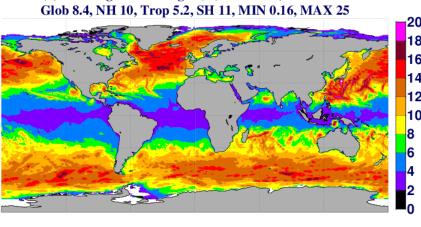


CASE STUDY: Unrealistically strong cyclones in early ERA5

RESOLVED: An improved production, now going back to 1940, resolves this

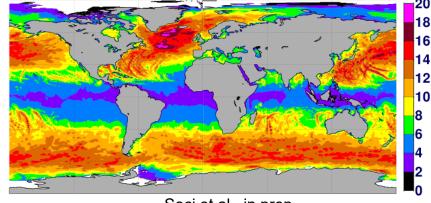
Additional, 'stricter' variational quality control (Huber norm) significantly reduces the weight assigned to IBTrACS observations, and rejects the vert lowest central pressures (< 910hPa

Extreme statistics are now more in line with 1980 onwards



(d) Max Sign. wave height (m) for ERA5, 1940-1981

(f) Max Sign. wave height (m) for ERA5, 1981-2022 Glob 8.7, NH 10, Trop 5.5, SH 11, MIN 0.087, MAX 20



Soci et al., in prep.

29



EUROPEAN CENTRE FOR MEDIUM-RANGE WEATHER FORECASTS

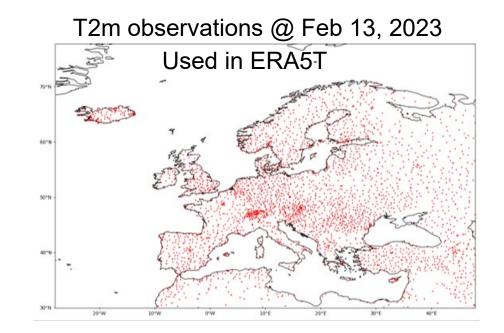
Ongoing maintenance in NRT

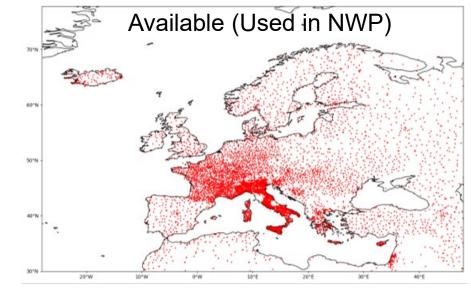
Continuous maintenance needed in Near-Real-Time

- Newly available (or reformatted!) observations
- System- and software maintenance

Requires close cooperation with research and operational teams

Requires dedicated testing and assessment for *each update*







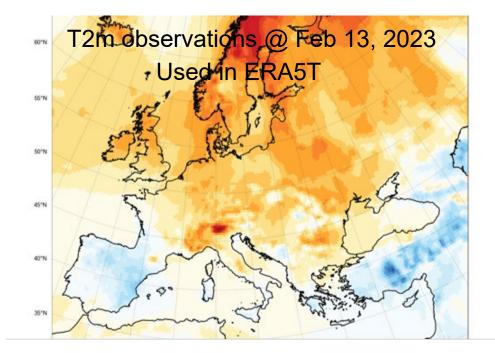
Ongoing maintenance in NRT

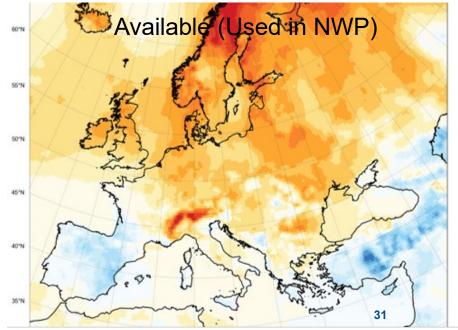
Continuous maintenance needed in Near-Real-Time

- Newly available (or reformatted!) observations
- System- and software maintenance

Requires close cooperation with research and operational teams

Requires dedicated testing and assessment for *each update*







EUROPEAN CENTRE FOR MEDIUM-RANGE WEATHER FORECASTS

In summary (I)

Reanalysis benefits from the work of large and diverse community which take time to collect, assess and integrate

Production of a reanalysis is resource intensive and not a (yet) a routine, fire-and-forget exercise. It requires close and continuous monitoring throughout its production lifetime

Routine monitoring as implemented for ERA5 is providing good handles on the assimilation system and the mean state of the reanalysis product, it has shown weaknesses in detecting localized anomalies, even if severe.

Monitoring and QA methods applied during production of a ERA5 ultimately rely on having eyes on plots and expert judgement, and are therefore weigh heavy on available human resources

With an increased number and diversity of applications for reanalyses, product requirements are diversifying. This spawns the need for monitoring and QA at production-time to adapt.

To reflect upon

Based on what we've seen, can we confidently, positively answer the following questions

Are we monitoring the right things?

- long term trends vs.(?) localized, rare problems
- move towards more complex, coupled reanalysis systems

Are we monitoring enough? Can we monitor enough? Do we want to?

- (human) resources, experience, expert knowledge
- Growing breadth of applications

Are we monitoring in the **right way**? Leveraging the **right tools**?

• Can a machine (learning application) assist?