

Plans for verification of ocean fields at ECMWF

Kristian Mogensen (and a lot of others)

Senior Scientist

krisitan.Mogensen@ecmwf.int



A bit of a disclaimer:

- ECMWF = European Centre for Medium-Range Weather Forecasts
- That means ECMWF's primary mission is not to provide forecasts of the ocean state
- Our primary mission is to provide the best medium-range (15 days) weather forecasts possible
 - As well as sub-seasonal (6 weeks) and seasonal (7 to 12/24 months)
 - In all cases the ocean acts as the lower boundary conditions of the atmosphere for ~70% of the earth
- We and other institutions have shown that the ocean impacts the weather and sub-seasonal to seasonal predictions
 - For longer time-scales the ocean is a strong driver of predictability for e.g. ENSO, MJO
 - For medium range the ocean impacts
 - Large scale circulation is impacted by tropical instability waves
 - Strong wind events introduce Ekman pumping
 - Especially important in tropical cyclones, but also present in mid-latitude weather systems
 - Sea-ice freezing and melting happens on a time scale of days and changes lower boundary of the atmosphere
- In the Earth System Modelling (and data assimilation approach) we want to evaluate all components
 - In our upgrade decision making process the atmosphere performance is the main driver, but we believe that in the long-term optimizing each component against observable variables will lead to better forecasts of the atmosphere

Outline of the rest of the talk

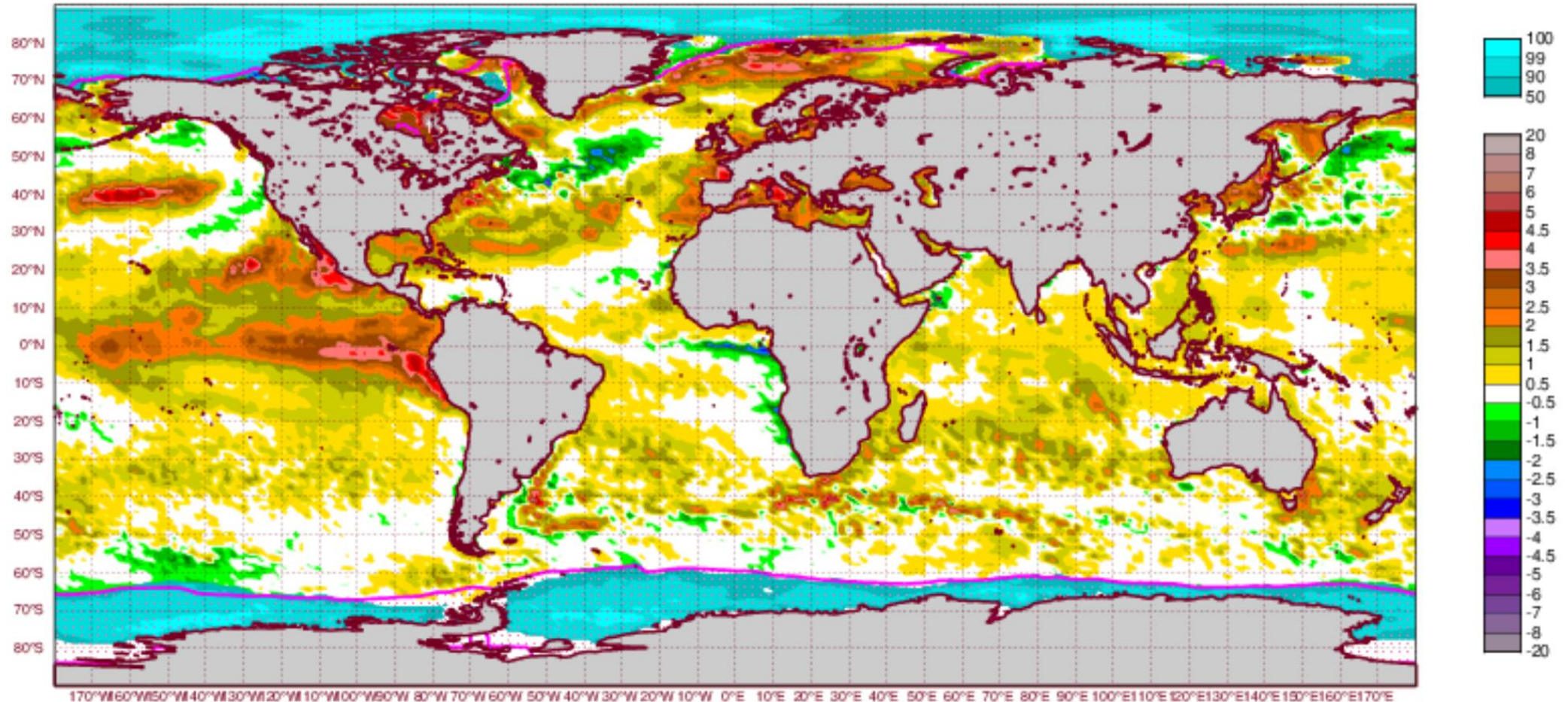
- Making plots of fields available to internal analysts and external users
 - We have a series of charts which are both available to the world and member/corporating states users only
- Operational routine monitoring of the ocean data assimilation system
 - Are the observations available as they should be?
- Outline on how we verify the atmospheric model:
 - Routine operational validation against own analysis and observations
 - Research evaluation of future contributions to new model versions
 - It is important to assess contributions to make decisions on which contributions to include in the next model upgrade
 - Ideally researchers should have access to the same tools as operations to ensure comparable evaluations
- Where we are with ocean operational monitoring and verification
- Verification of reanalysis (focused on the ocean in this talk)
- Working together with the people providing observations
 - Performing OSE's or OSSE's
 - Use of campaign data for process evaluation

Operational ocean plots (so far)

<https://charts.ecmwf.int>

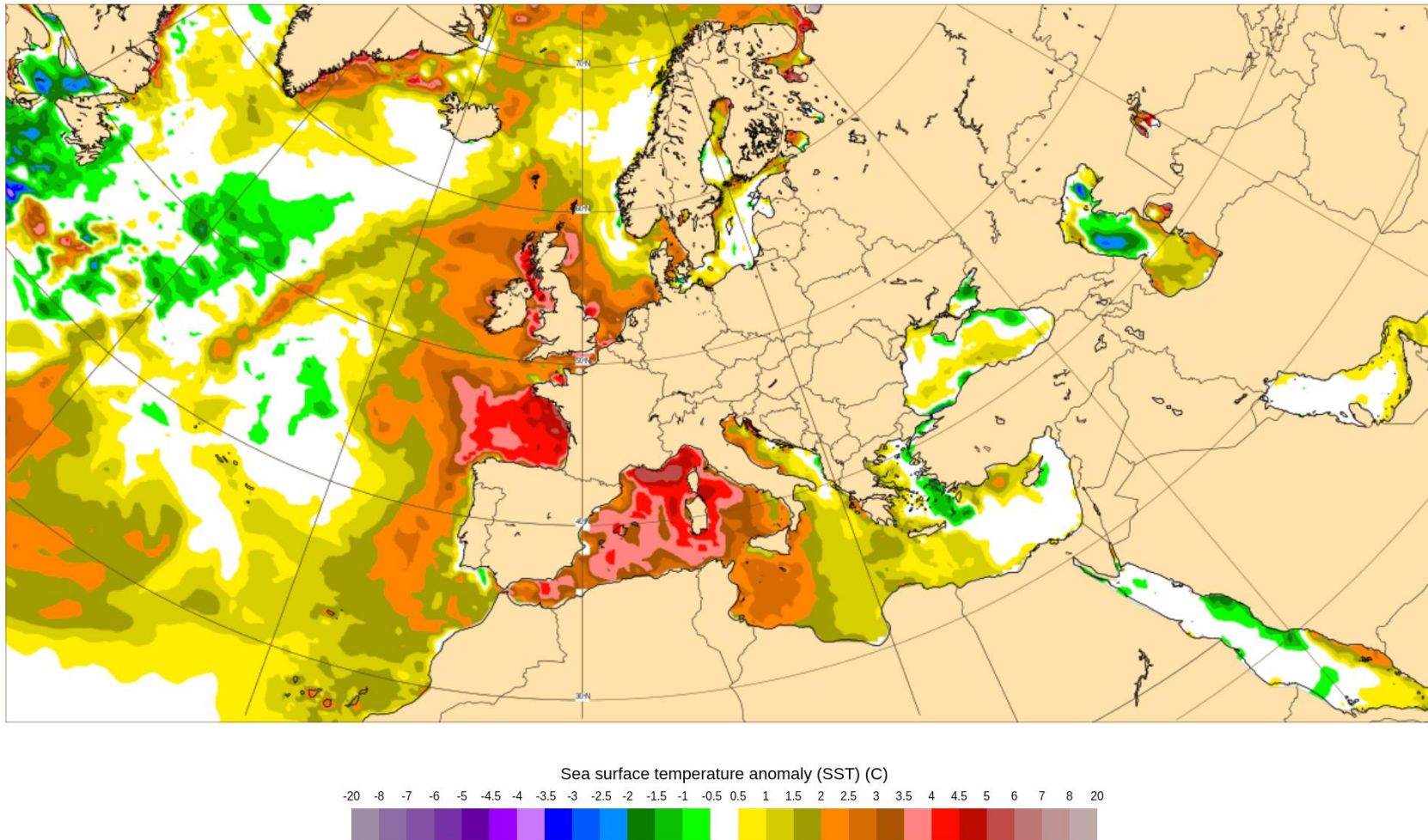
OpenCharts example (<https://charts.ecmwf.int/permalinks/sea-surface-temperature-14235>)

Monday 06 July 2026 0000 UTC ECMWF t+0 VT Monday 06 July 2026 0000 UTC
SST anomaly and Sea Ice. Sea Surface Temperature anomaly (C) in control run (bottom legend, 2C contour interval near extrema).
Control Run Sea Ice cover (top legend >= 50% only). Climatological Sea Ice cover in magenta (>= 50%).

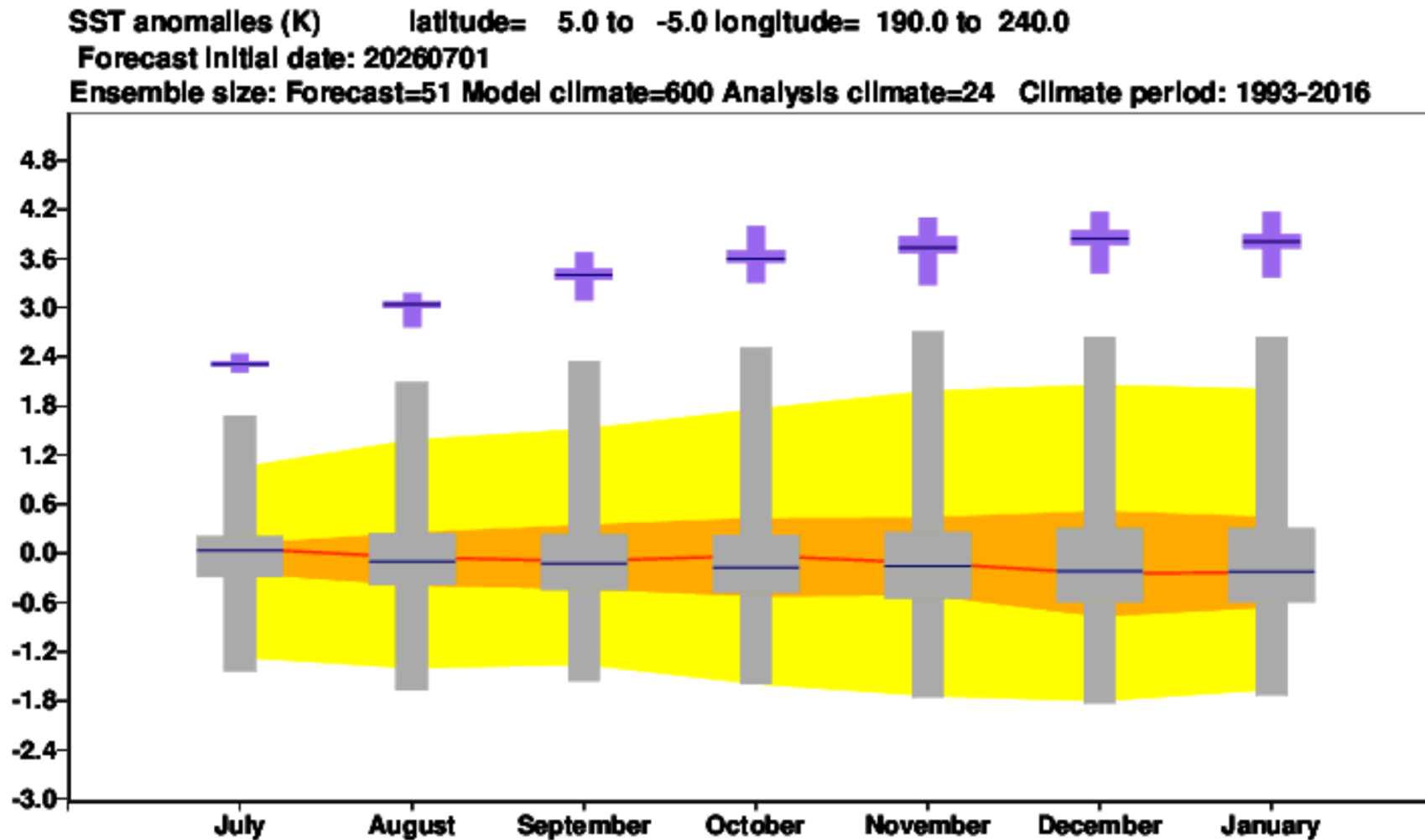


MHW around Europe: (<https://charts.ecmwf.int/permalinks/sea-surface-temperature-91720>)

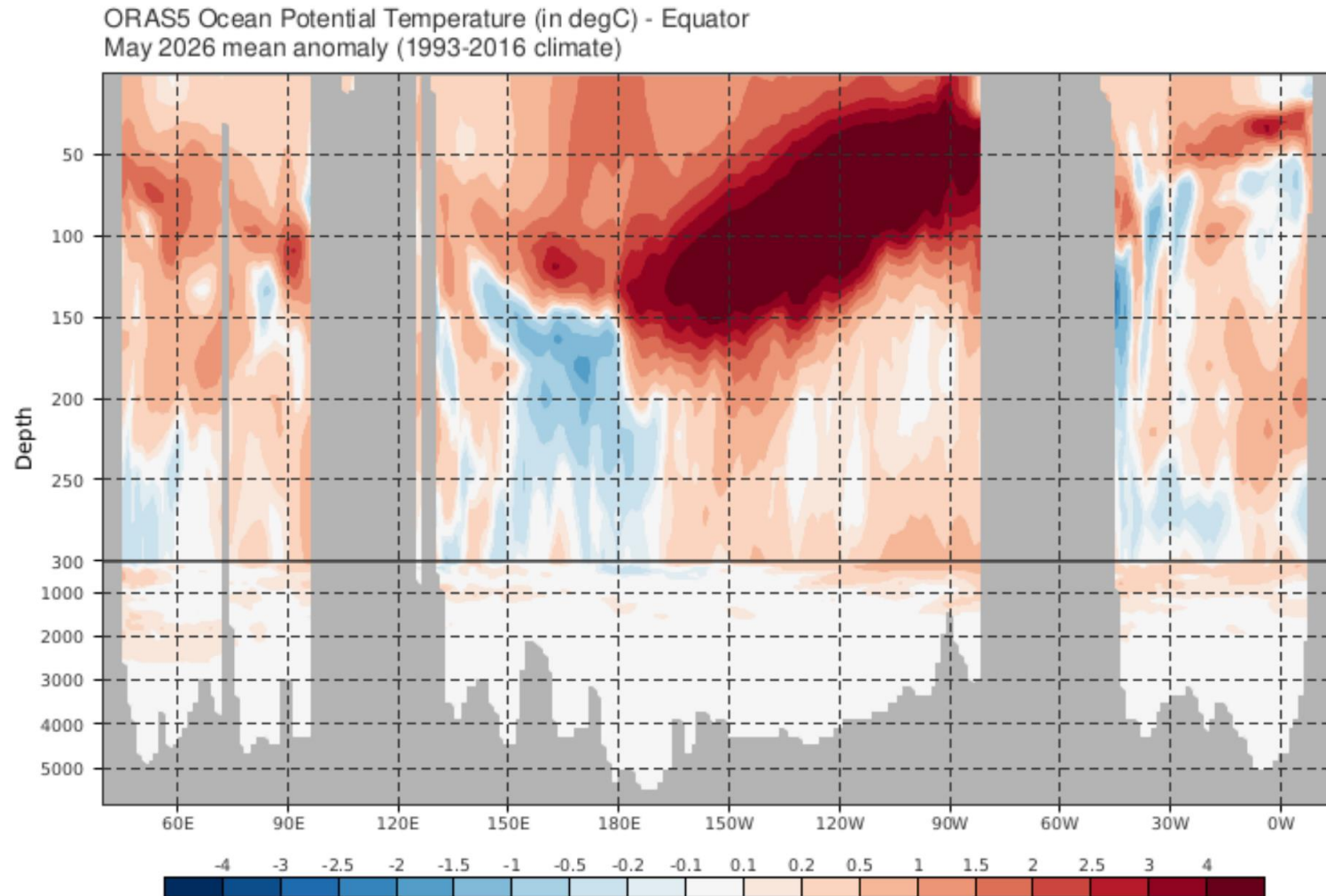
Base time: Mon 06 Jul 2026 00 UTC Valid time: Thu 16 Jul 2026 00 UTC (+240h) Parameter : sst_anomaly Area : Europe



El Nino forecasts: <https://charts.ecmwf.int/permalinks/sea-surface-temperature-area-averages-seasonal-forecast-seas5-47240>



Ocean heat content along equator from ORAS5: <https://charts.ecmwf.int/catalogue/packages/oras5/>



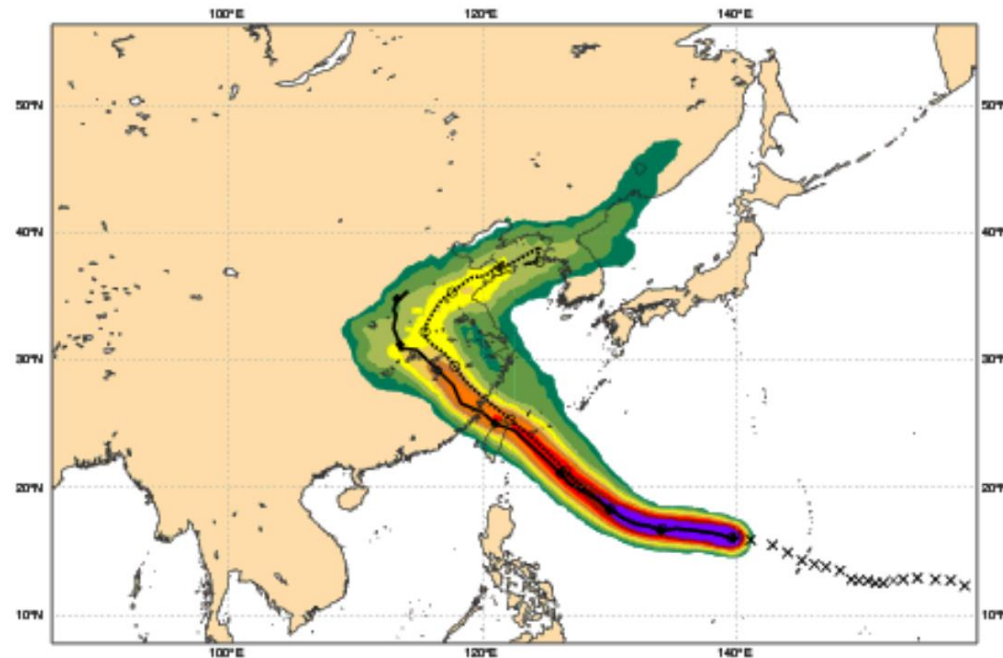
TC Bavi atmospheric prediction:

Date 20260707 00 UTC @ECMWF

Probability that **BAVI** will pass within 120 km radius during the next 240 hours

tracks: **solid**=CTRL; **dot**=Ens Mean [reported minimum central pressure (hPa) **935**]

5-10 10-20 20-30 30-40 40-50 50-60 60-70 70-80 80-90 > 90%

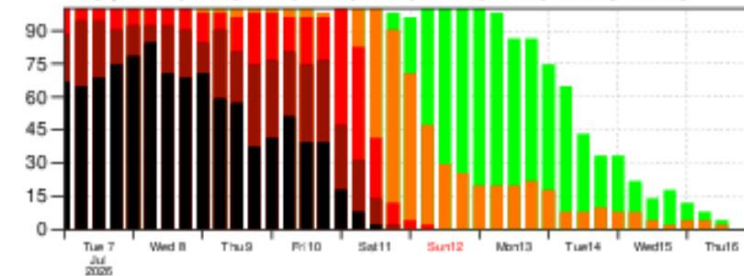


List of ensemble members numbers forecast Tropical Cyclone

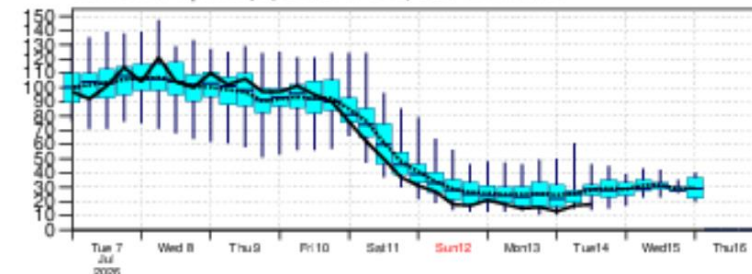
Intensity category in colours: **TD**[up to 33] **TS**[34-63] **HR1**[64-82] **HR2**[83-95] **HR3**[> 95 kt]

+024 h: 01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50
 +048 h: 01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50
 +072 h: 01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50
 +096 h: 01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50
 +120 h: 01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50
 +144 h: 01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50
 +168 h: 01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50
 +192 h: 02 07 17 18 20 21 23 24 26 27 28 29 30 32 33 34 36 37 40 45 47
 +216 h: 20 29 30 34 36 37 40 45 47
 +240 h: 20

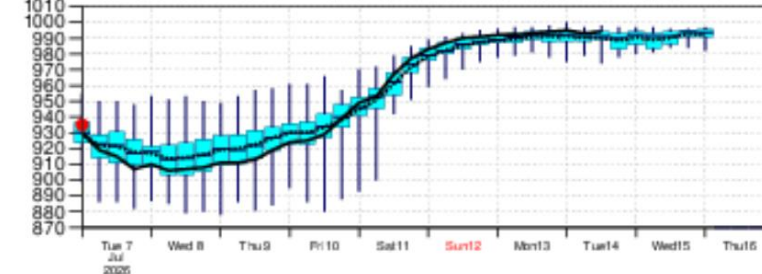
Probability (%) of Tropical Cyclone Intensity falling in each category
TD[up to 33] **TS**[34-63] **HR1**[64-82] **HR2**[83-95] **HR3**[> 95 kt]



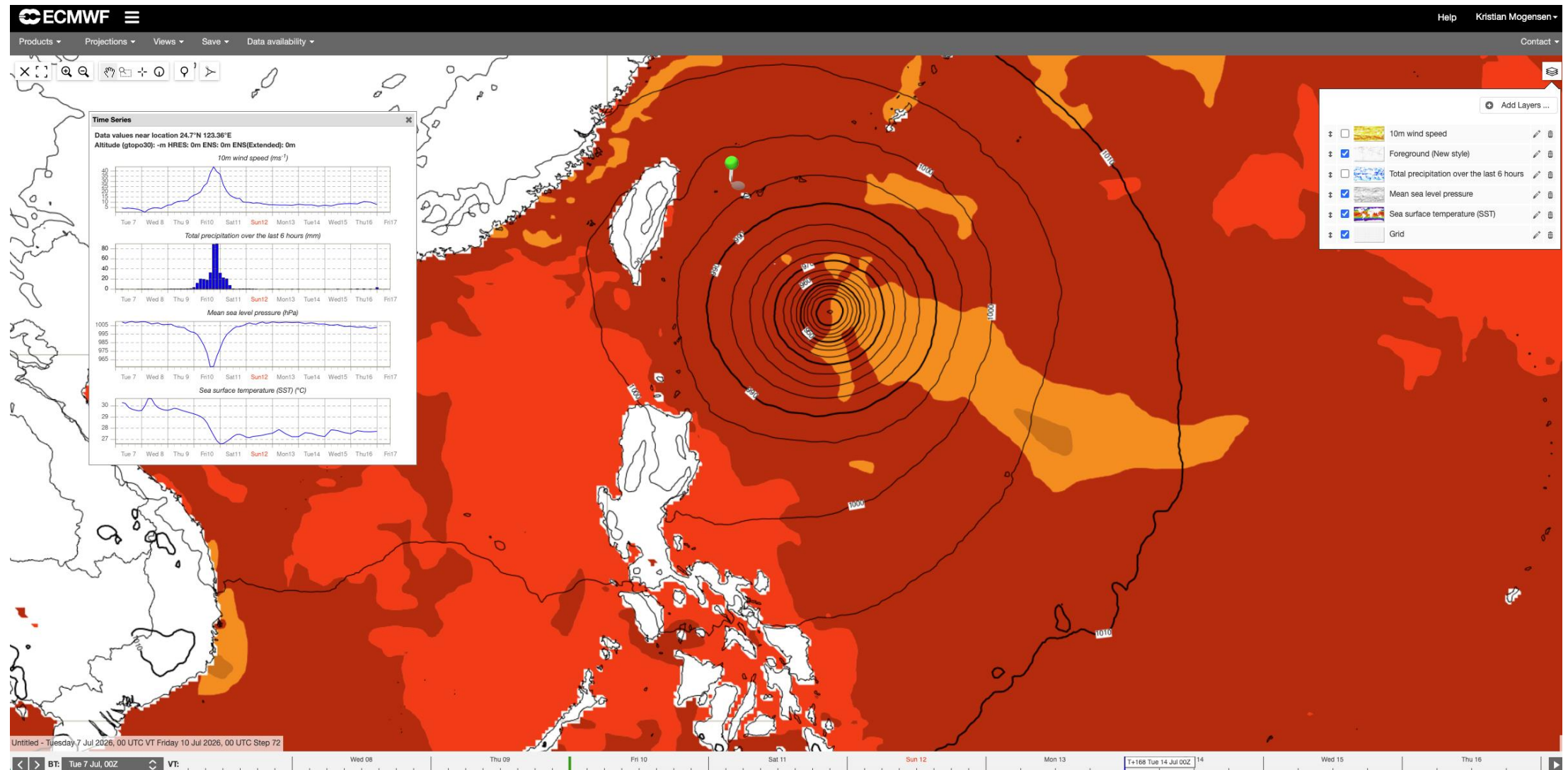
10m Wind Speed (kt) **solid**=CTRL; **dot**=Ens Mean



Mean Sea Level Pressure in Tropical Cyclone Centre (hPa) **solid**=CTRL; **dot**=Ens Mean



On demand eccharts for ECMWF member and corporation states



Ocean output plots outlook

- We will plan to add more ocean plots to both the opencharts and eccharts systems
- Should help weather forecasters in situations where the ocean matter for the atmospheric predictions
- For TC Bavi having upper ocean heat content or tropical cyclone heat potential could have been useful for forecasters

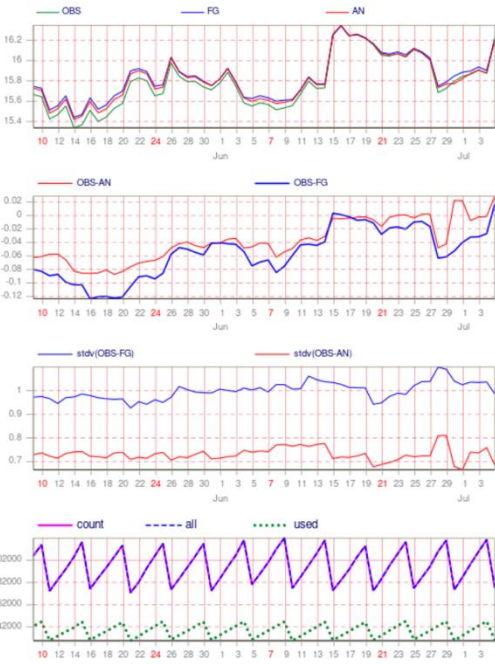
Monitoring of the ocean data assimilation system

<https://www.ecmwf.int/en/forecasts/quality-our-forecasts/monitoring-observing-system#Ocean>

Ocean observation monitoring

Potential Temperature

STATISTICS FOR POTENTIAL TEMPERATURE FROM ARGO (Globe)
DEPTH=50.0_100.0m All DATA (TIME STEP=24 HOURS)
Area Globe
Exp=0001 LAST TIME WINDOW (2026070600)

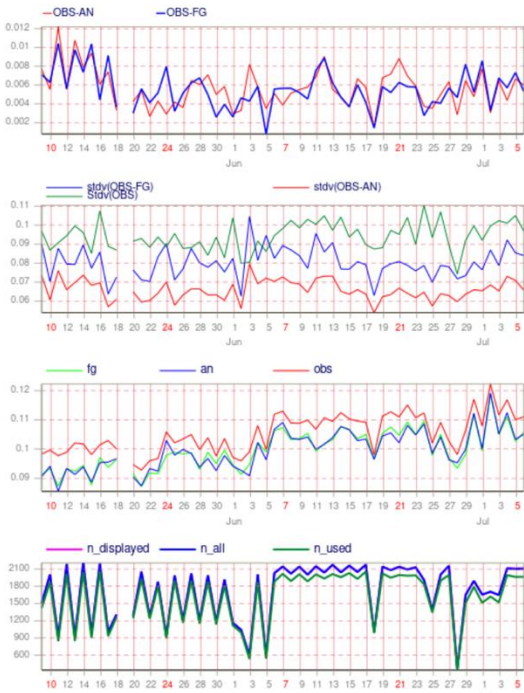


ARGO

Cryosat

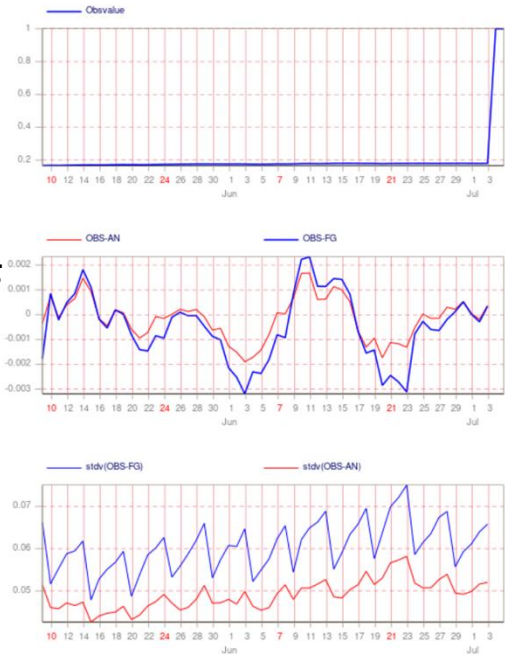
Series of statistics for Sea level anomaly

STATISTICS FOR Sea level anomaly FROM CRYOSAT (Globe)
All DATA (TIME STEP=24 HOURS)
Area 90.N/-90.S/0.W/360.E (Over \$surface)
Exp=0001 LAST TIME WINDOW (2026070700)



Sea Ice Concentration

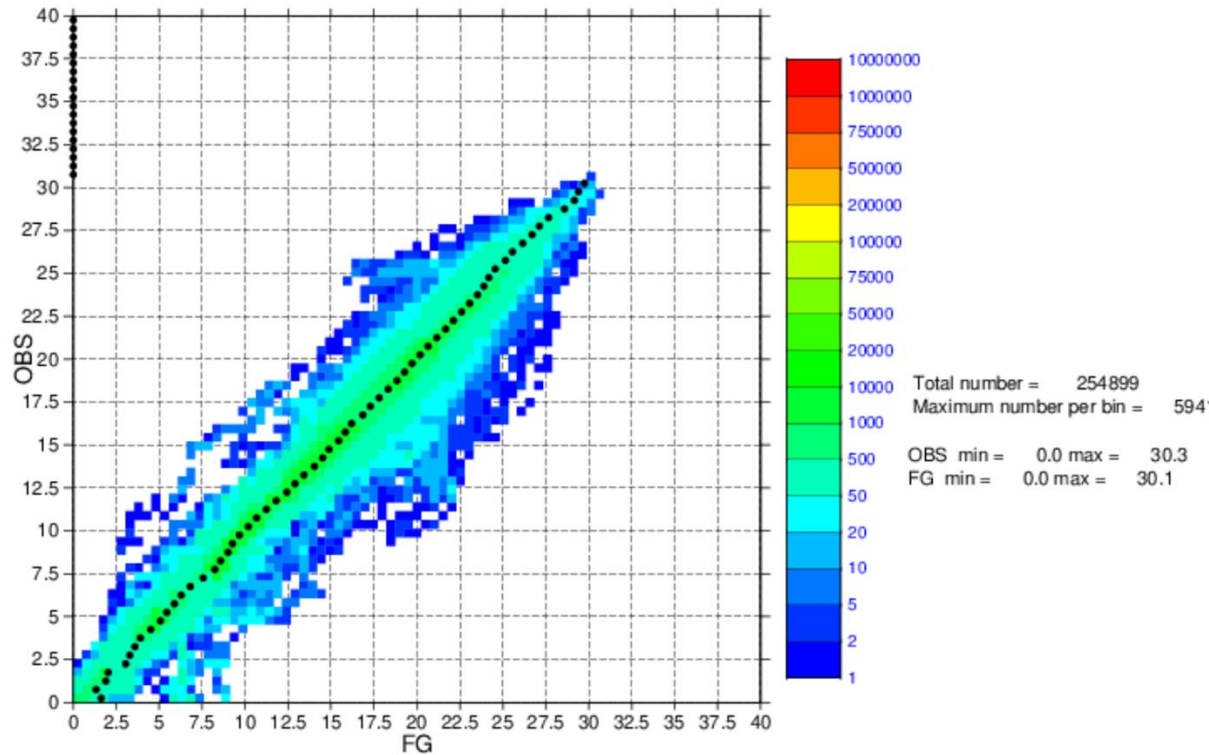
STATISTICS FOR Sea Ice Concentration [%] (Globe)
DEPTH=\$level All DATA (TIME STEP=24 HOURS)
Area Globe
Exp=0001 LAST TIME WINDOW (2026070600)



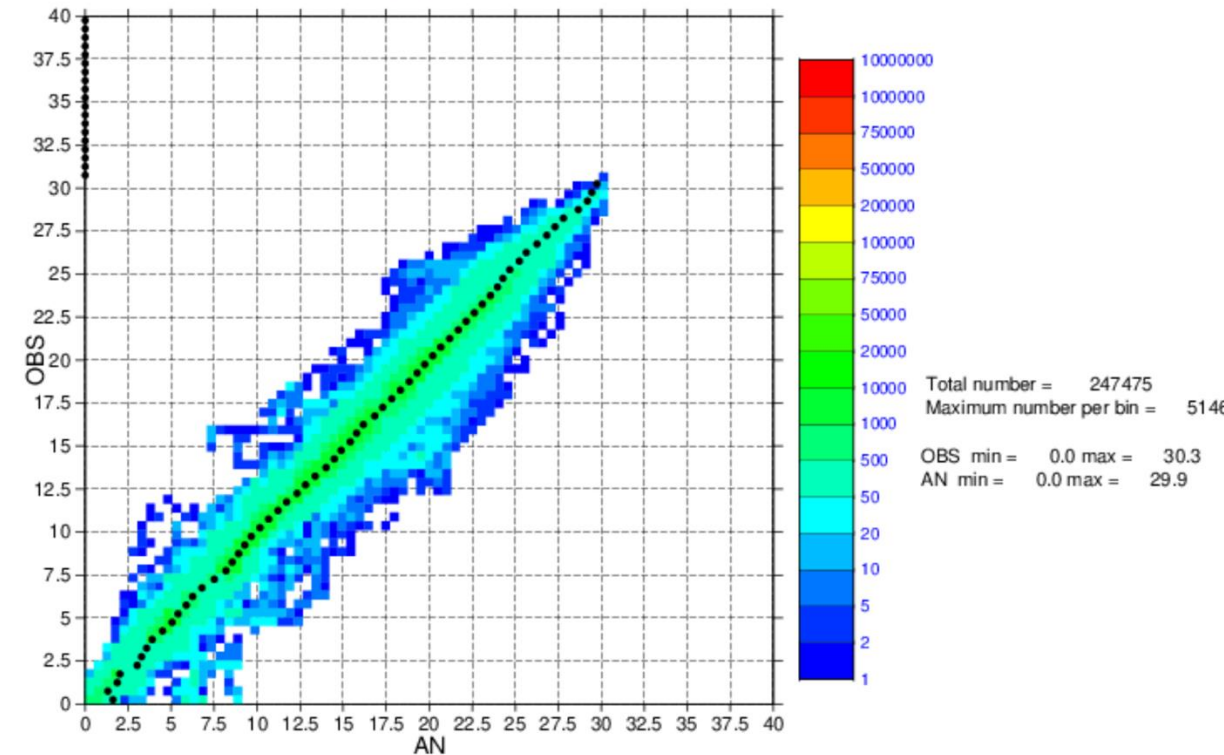
OSI-SAF

Observation monitoring: DA performance of ARGO

ISTICS FOR POTENTIAL TEMPERATURE FROM ARGO, DEPTH 100.00 - 200.00 M
EXP = 0001 ; PERIOD = 2024010100 - 2024011600
ALL - GLOBE



ISTICS FOR POTENTIAL TEMPERATURE FROM ARGO, DEPTH 100.00 - 200.00 M
EXP = 0001 ; PERIOD = 2024010100 - 2024011600
ALL - GLOBE

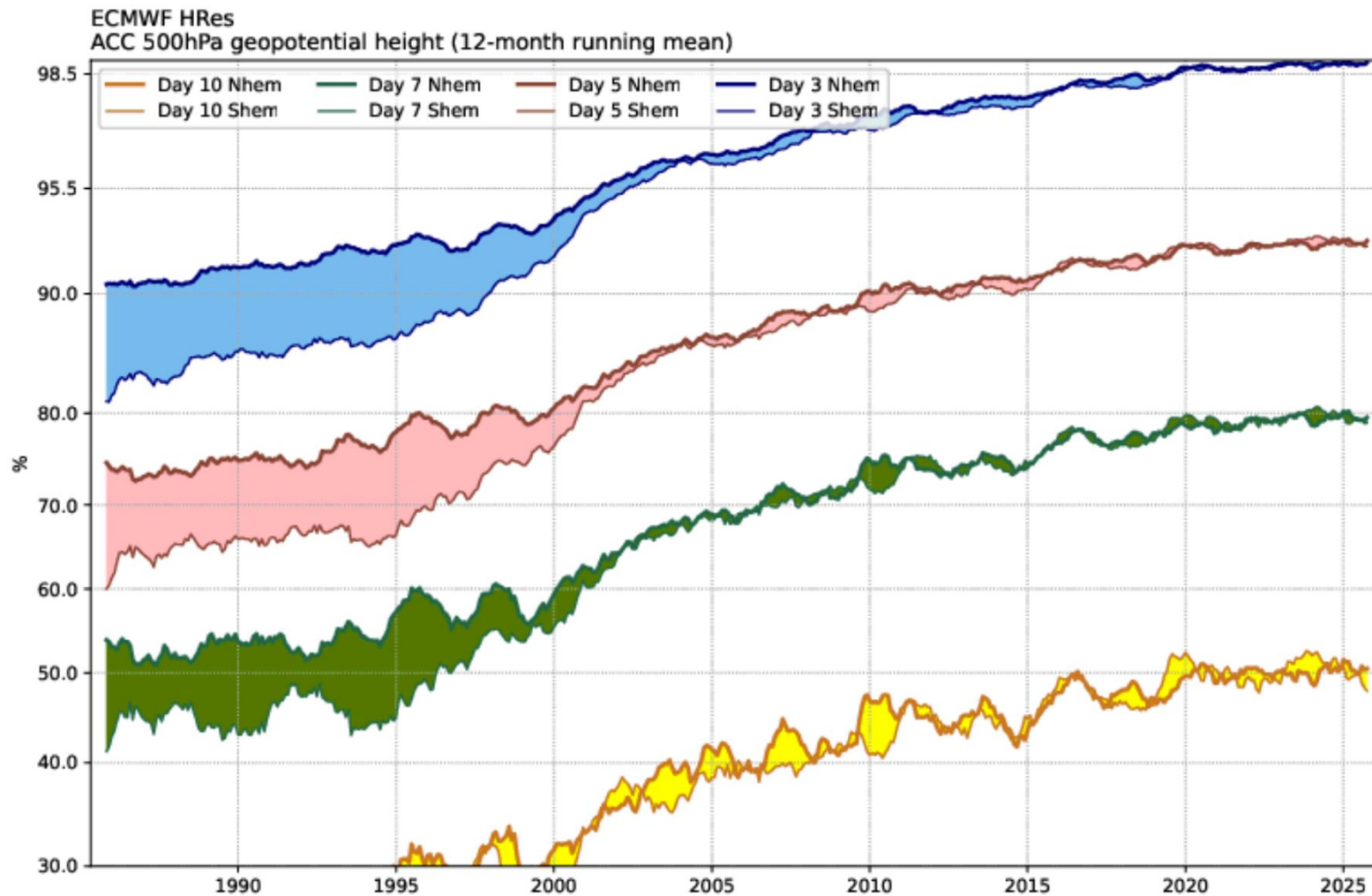


The analysis (AN) versus observation is narrower than the first guest (FG) versus observations

Operational validation of atmospheric forecasts

<https://charts.ecmwf.int/?facets=%7B%22Type%22%3A%5B%22Verification%22%5D%7D>

Headline scores: z500 anomaly correlation coefficient

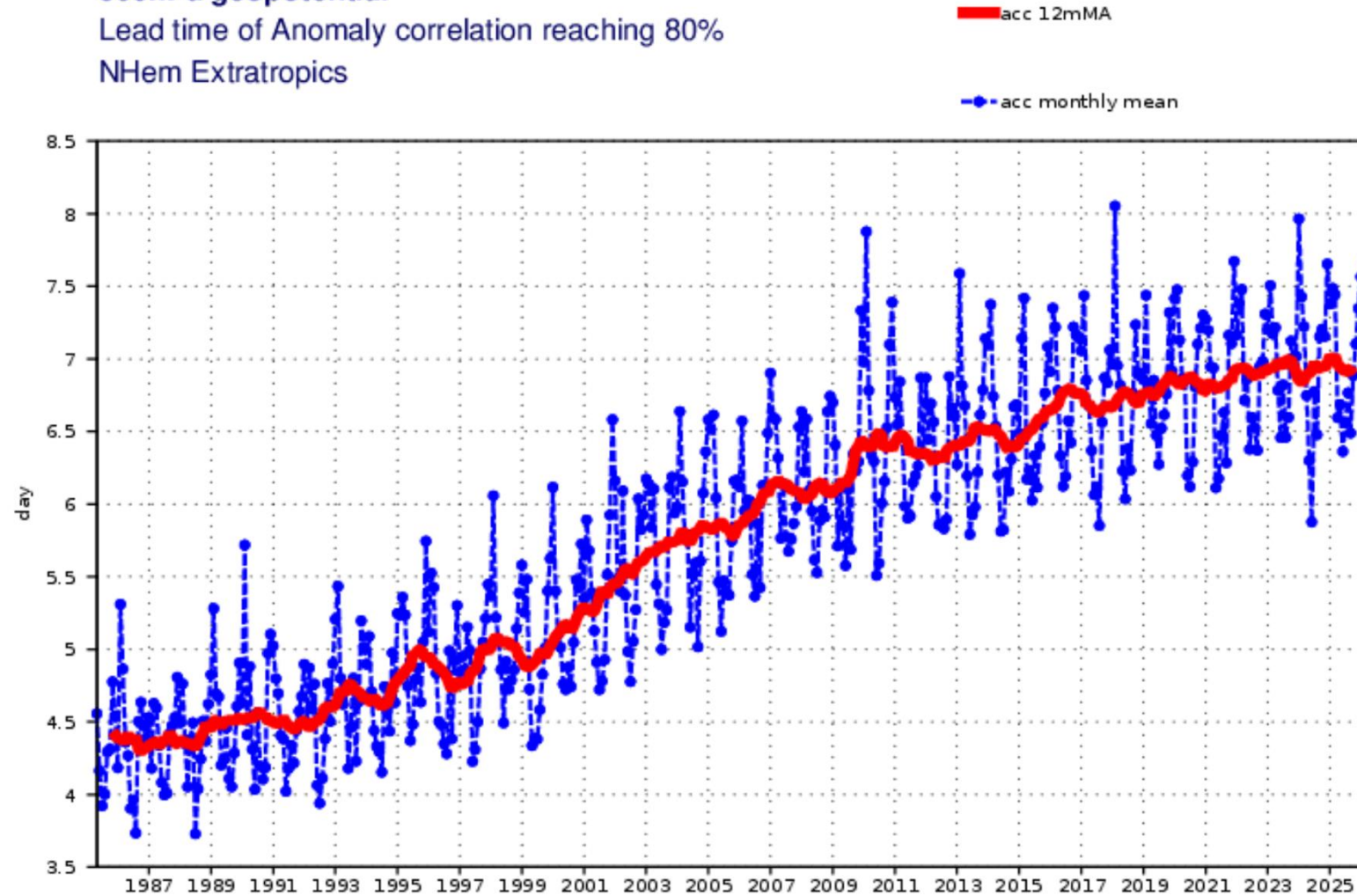


Headline scores: z500 anomaly correlation coefficient

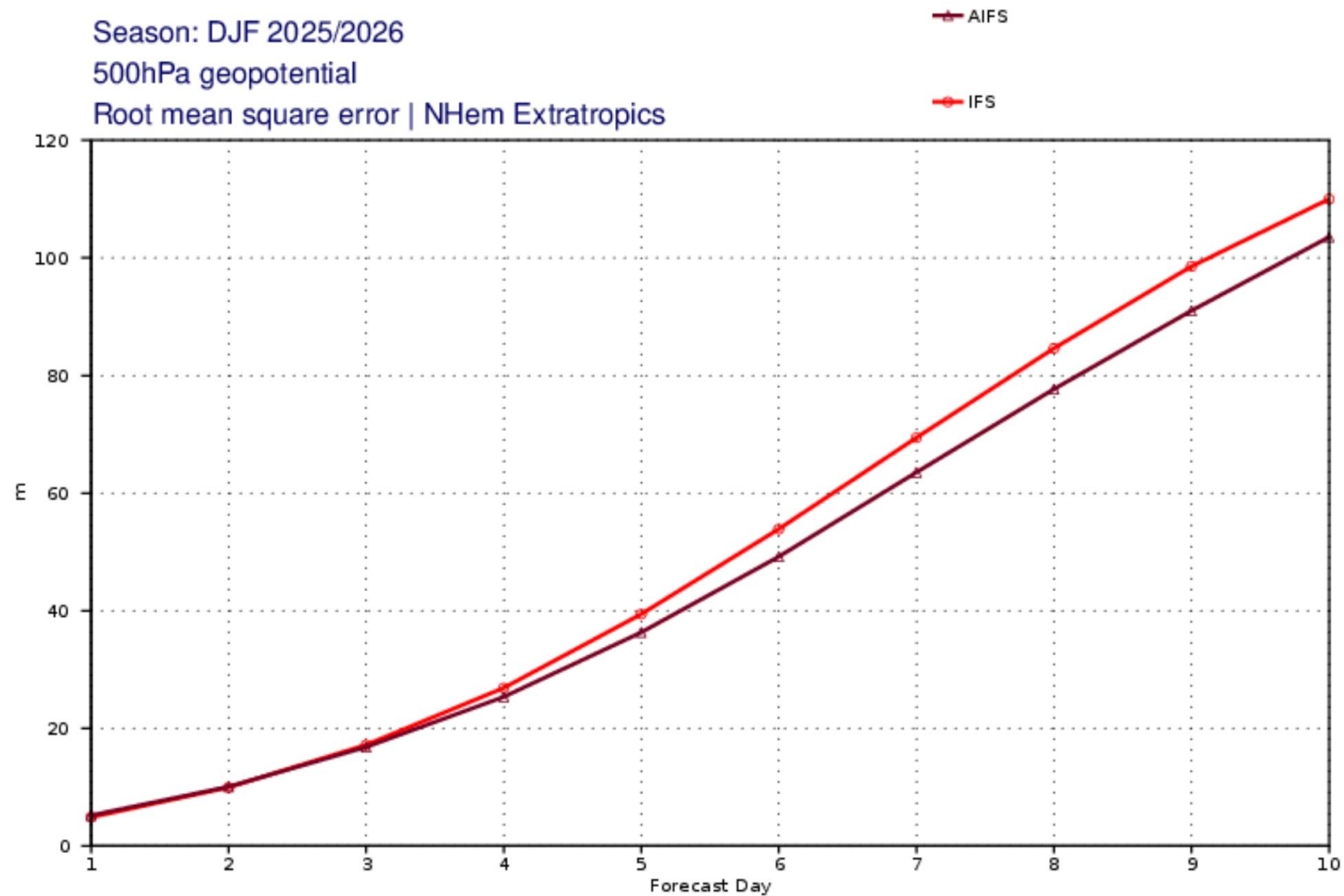
500hPa geopotential

Lead time of Anomaly correlation reaching 80%

NHem Extratropics



Validation of AIFS against IFS:



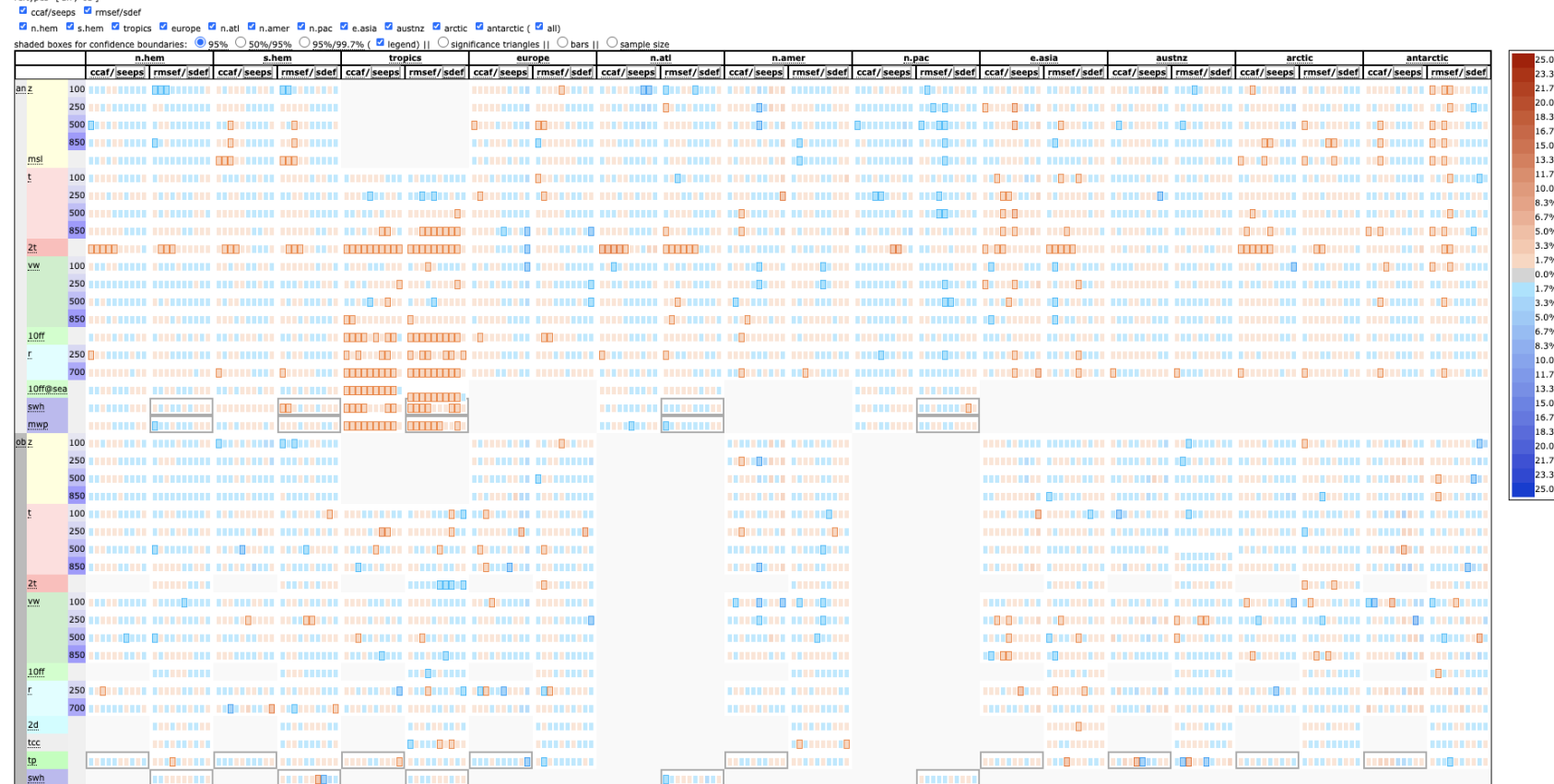
Tools used for developments

Scorecards for medium-range and sub-seasonal forecasts

Impact of removing ARGO in ocean DA on NWP 1

ARGO_0pct vs REF for ALL start dates. scorecard

dates=[2021-06-01 00:00:00,2021-06-02 00:00:00,2021-06-03 00:00:00,...,2023-05-31 00:00:00,2023-06-01 00:00:00]
 steps=[24, 48, 72, 96, 120, 144, 168, 192, 216, 240]
 vstreams=['grdx_ne1_det_an', 'grdx_ne1_det_ob']
 classes=r0
 streams=wave
 expvars=(cntrl:ld5p, exper:icstu)
 reftypes=['an', 'ob']



ccaf=Anomaly correlation,rmsef=Root mean square error,sdef=Standard deviation of forecast error,seeps=Stable Equitable Error in Probability Space

red = the experiment (esute) is worse than the control. blue = the experiment (esute) is better than control.

purple = the experiment is more active than the control. green = the experiment is less active than control.

Colour hue and saturation are proportional to normalised difference value. A [colour-framed cell](#) indicate the differences are statistically significant on 95% confidence level. However, where the differences are not significant on at least 0% level they are masked by a light grey box.

For more details:

- Hover the mouse cursor over a cell and a small information window pops up with the difference values.
- Click on a cell and a graph of normalized differences will show.

Screenshot

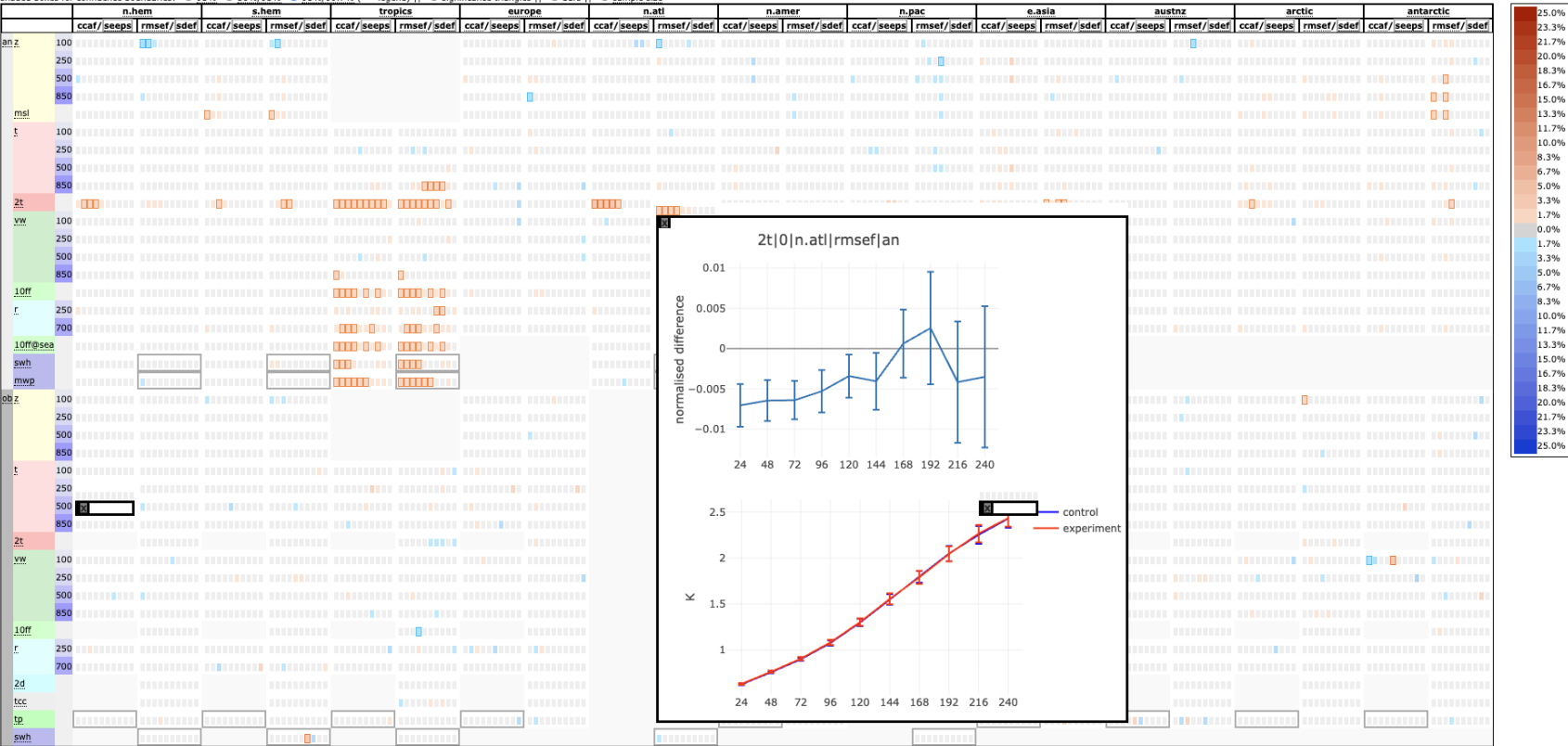
Top half: verification
against operational
atmospheric analysis;

Bottom half: Against
observations

Impact of removing ARGO in ocean DA on NWP 2

ARGO_0pct vs REF for ALL start dates. scorecard

dates=[2021-06-01 00:00:00,2021-06-02 00:00:00,2021-06-03 00:00:00,...,2023-05-31 00:00:00,2023-06-01 00:00:00]
steps=[24, 48, 72, 96, 120, 144, 168, 192, 216, 240]
vstreams=['grdx_ne1_det_an', 'grdx_ne1_det_ob']
classes=rd
streams=wave
expvars=(ctrl:idsp, exper:ictu)
reftypes=['an', 'ob']
☒ ccaf/seeps ☒ rmsef/sdef
☒ n.hem ☒ s.hem ☒ tropics ☒ europe ☒ n.atl ☒ n.amer ☒ n.pac ☒ e.asia ☒ austnz ☒ arctic ☒ antarctic (☒ all)
shaded boxes for confidence boundaries: ☐ 95% ☐ 50%/95% ☒ 95%/99.7% (☒ legend) || ☐ significance triangles || ☐ bars || ☐ sample size



ccaf=Anomaly correlation,rmsef=Root mean square error,sdef=Standard deviation of forecast error,seeps=Stable Equitable Error in Probability Space

red = the experiment (esuite) is worse than the control. blue = the experiment (esuite) is better than control.
purple = the experiment is more active than the control. green = the experiment is less active than control.
Colour hue and saturation are proportional to normalised difference value. A colour-framed cell indicate the differences are statistically significant on 99.7% confidence level. However, where the differences are not significant on at least 95% level they are masked by a light grey box.

For more details:

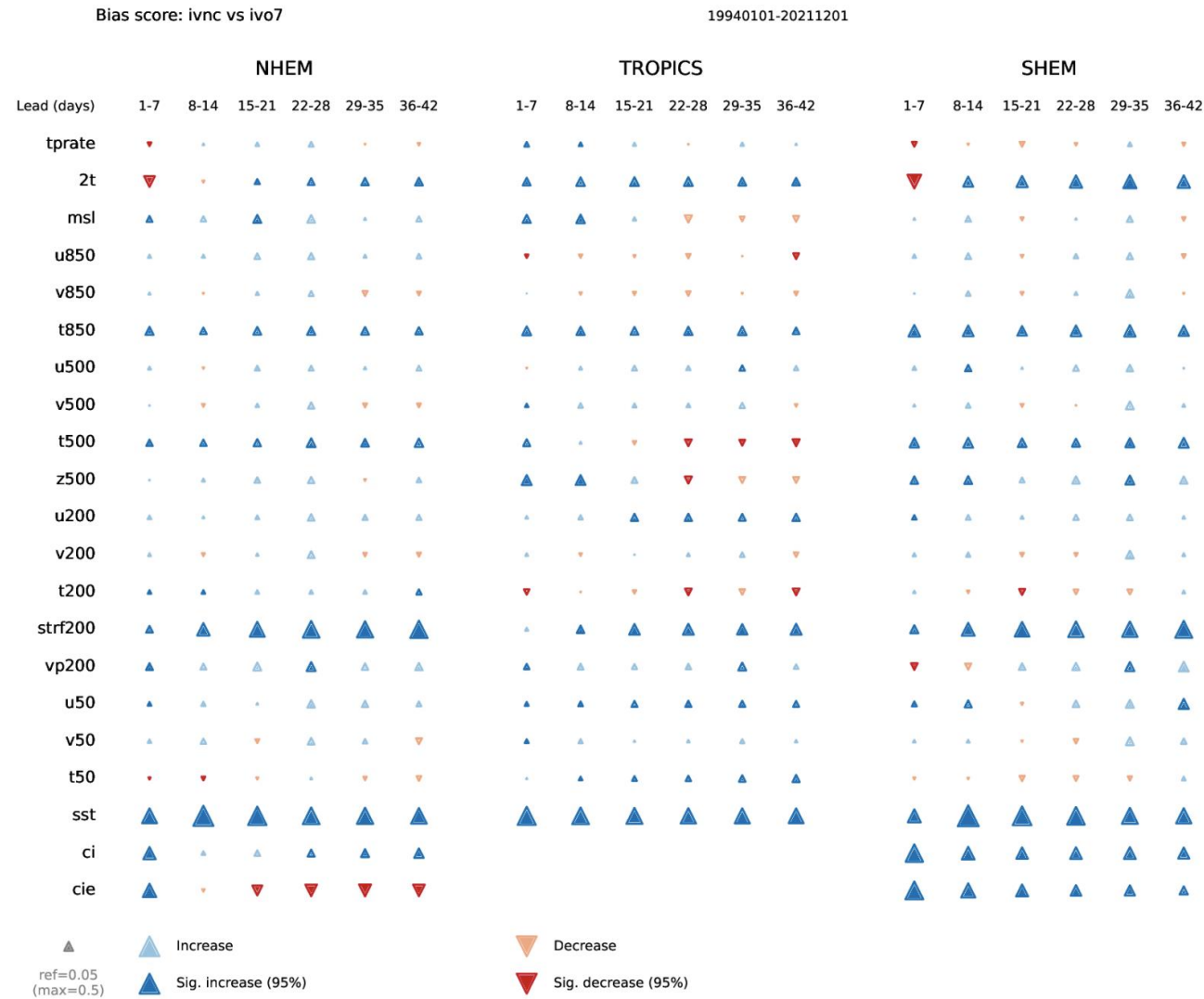
- Hover the mouse cursor over a cell and a small information window pops up with the difference values.
- Click on a cell and a graph of normalized differences will show.

Screenshot

Top half: verification
against operational
atmospheric analysis;

Bottom half: Against
observations

Scorecards for S2S model evaluation has some ocean fields



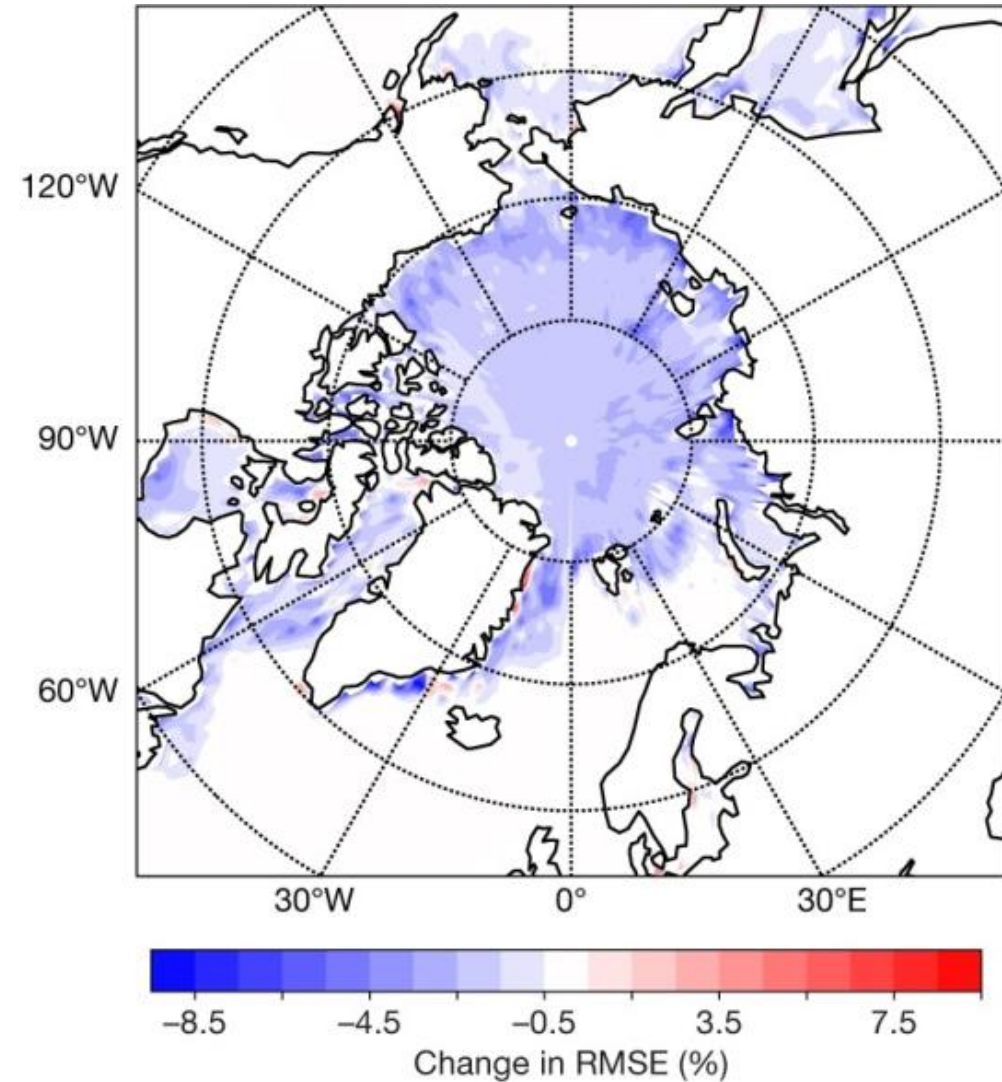
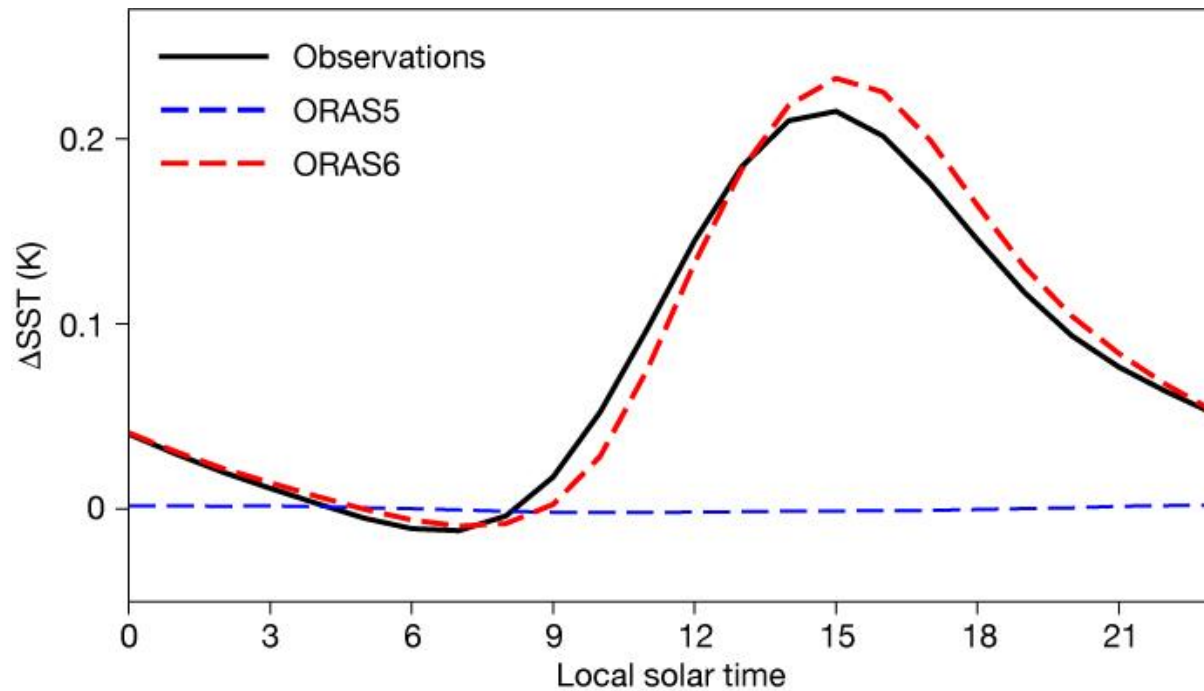
Remarks on our tools

- The score card for medium-range forecasts uses both an atmospheric analysis and observations for evaluation of forecasts
 - Quality control of observations is a key thing to get right
 - Our first attempt to use SST observations from ships and buoys based on our operational archive failed due to errors in the observations
 - Using data from the global drifter programme (<https://www.aoml.noaa.gov/global-drifter-program/>) is being investigated and seems more successful
- We will investigate using the IV-TT collected ocean observations with these tools

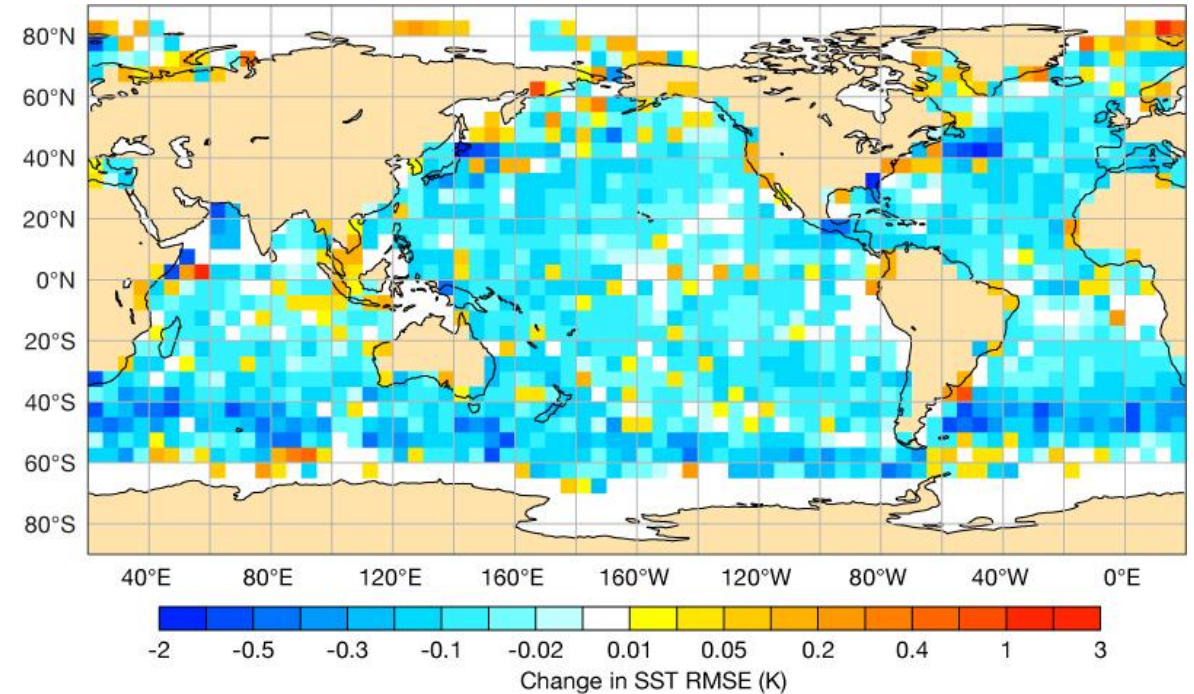
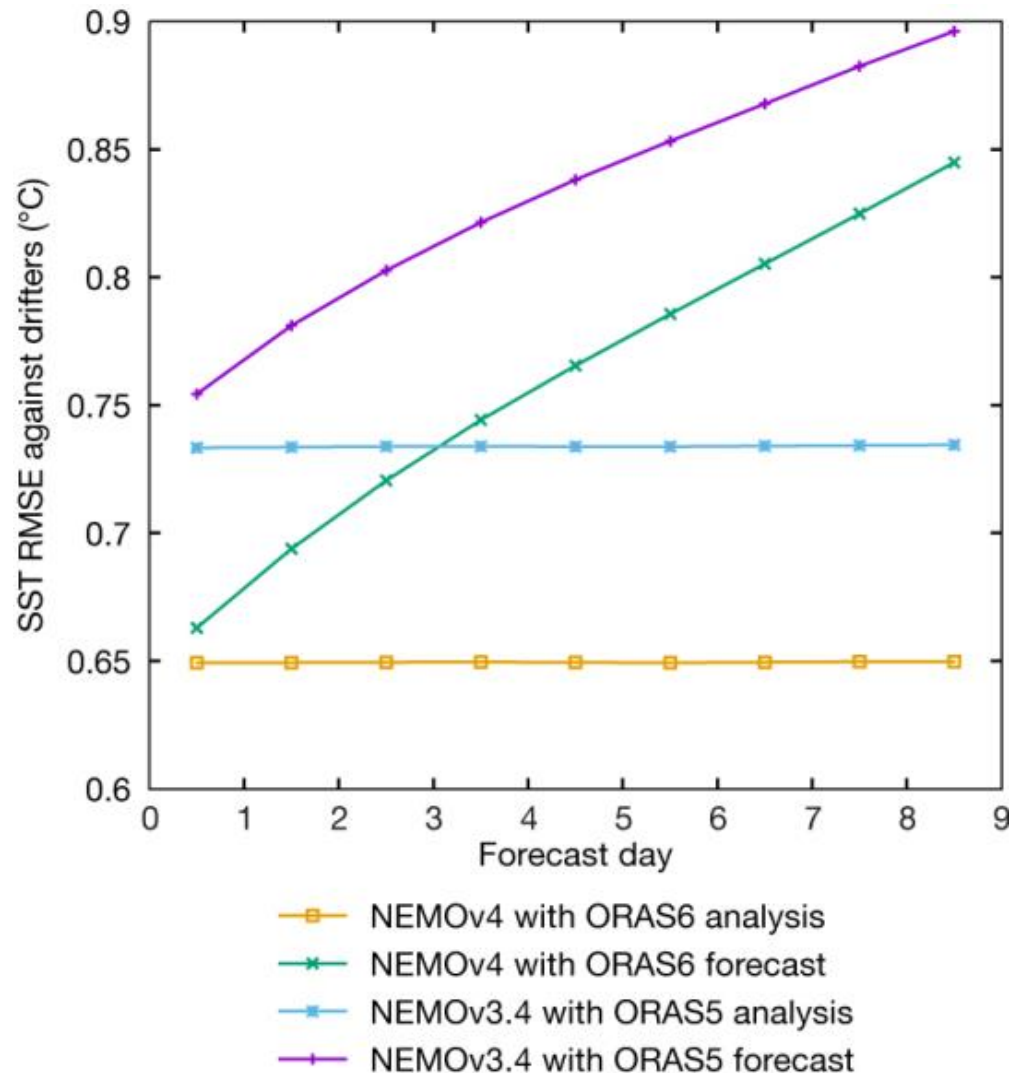
Evaluation of ocean reanalyses

<https://www.ecmwf.int/en/newsletter/180/earth-system-science/ecmwfs-next-ensemble-reanalysis-system-ocean-and-sea-ice>

New features in ORAS6: diurnal cycle and better sea-ice model/DA



Validation of ORAS6 compared to ORAS5 against drifters



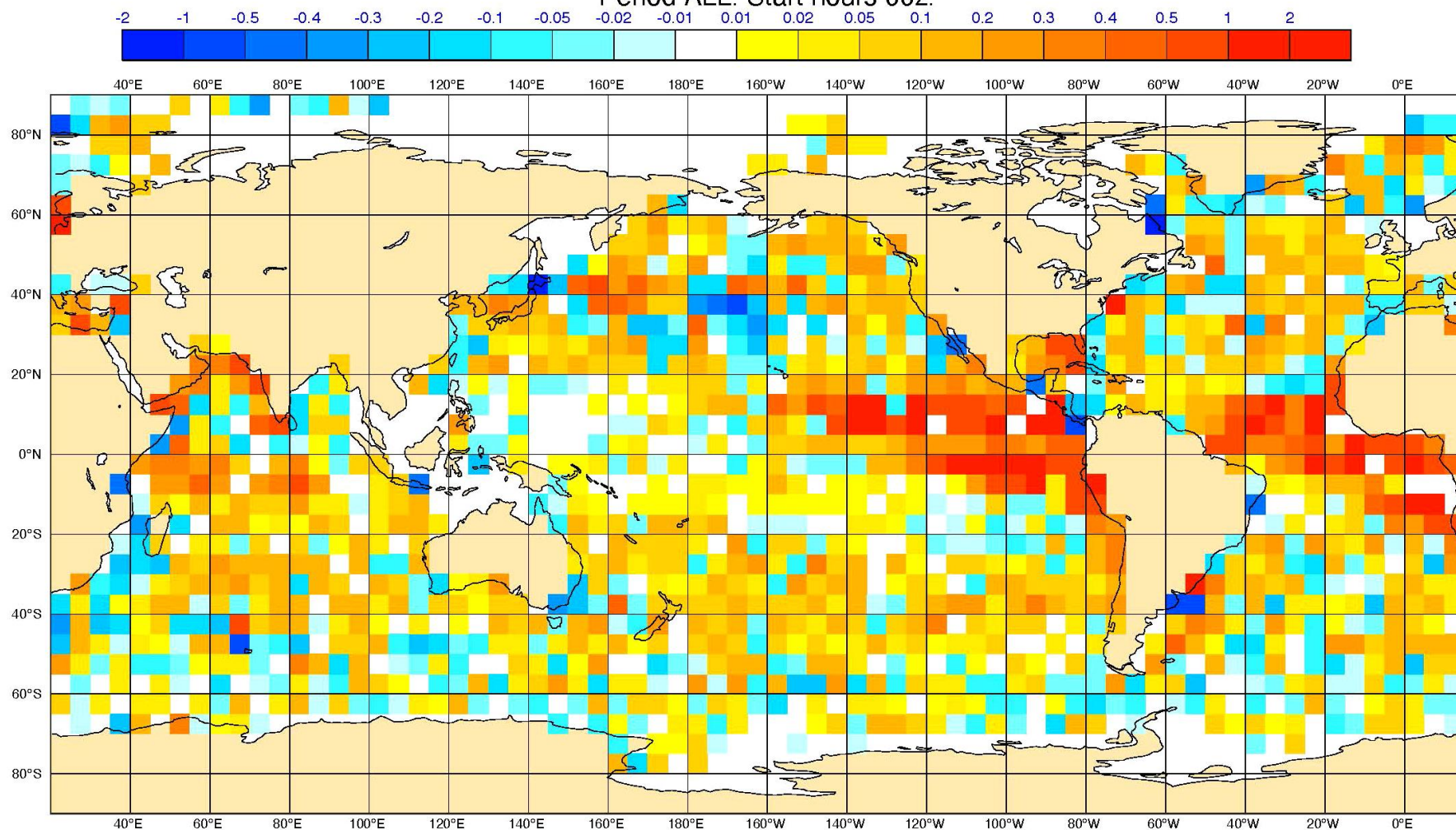
- Blue means ORAS6 is better than ORAS5
- The observations were not directly assimilated, but were still used as calibration for the SST products used

Working with the observational communities

Working with the observational community: ARGO for NWP

Argo Temp RMSE diff NoArgo fc - CNT fc. Level 50m. Fcst day 10.

Period ALL. Start hours 00z.



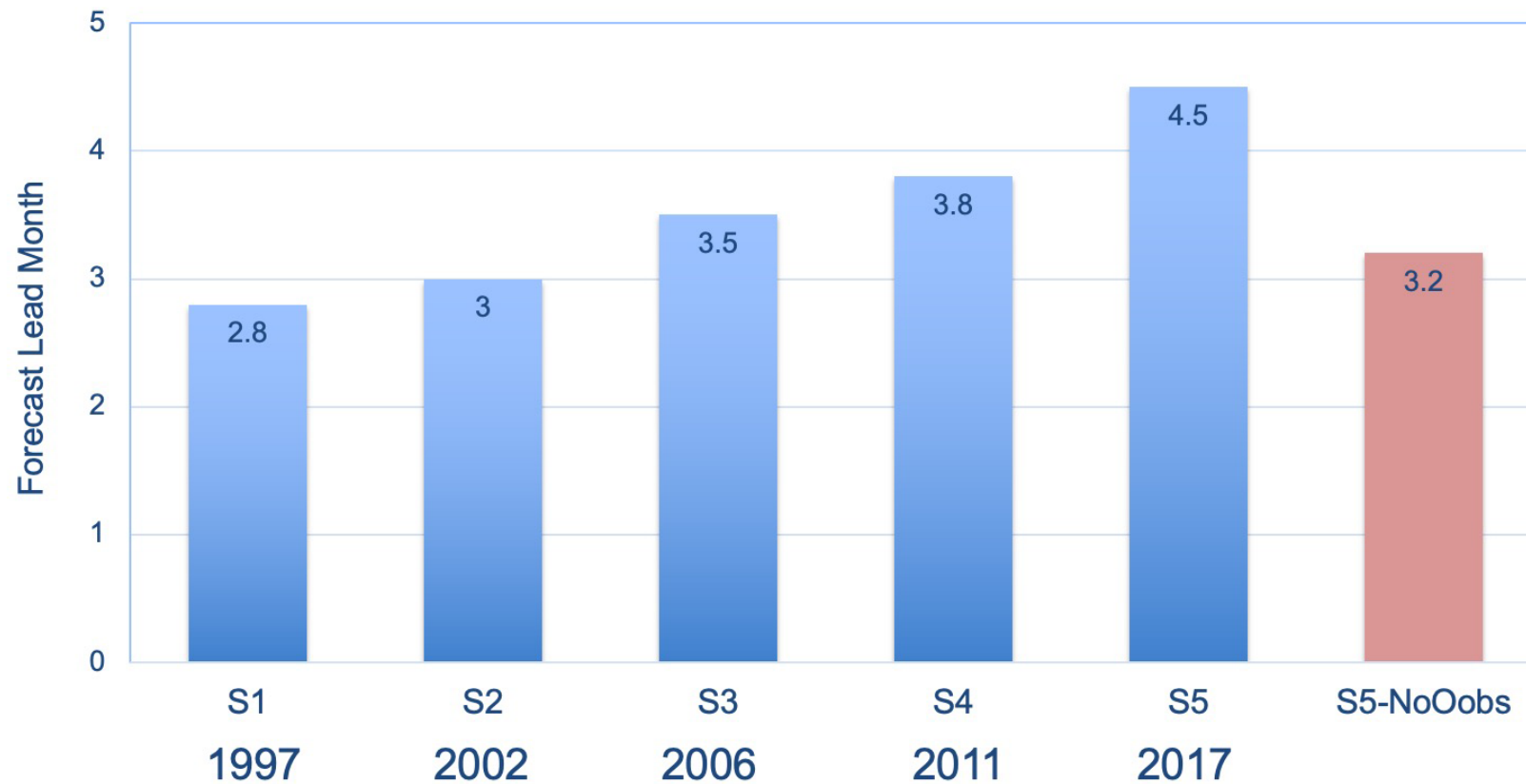
Magics 4.15.4 - ad6-108.bulx - ne1 - Tue Jul 2 18:05:14 2024



<https://www.frontiersin.org/journals/marine-science/articles/10.3389/fmars.2025.1593904/full>

Working with the observational community: ARGO for seasonal forecasting

Forecast Lead Month with correlation above 0.9 in Nino3.4

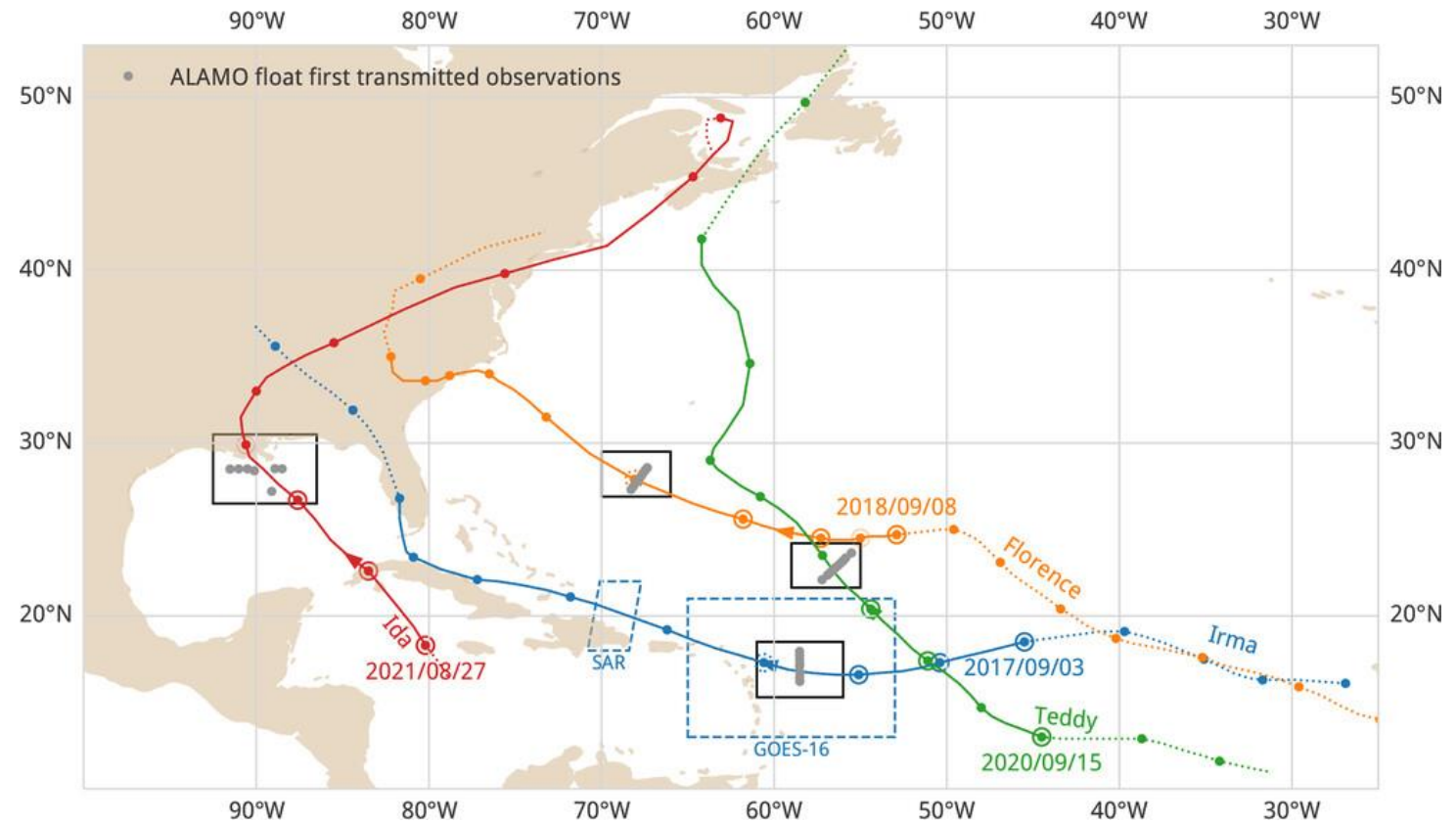


ECMWF Seasonal Forecasting Systems

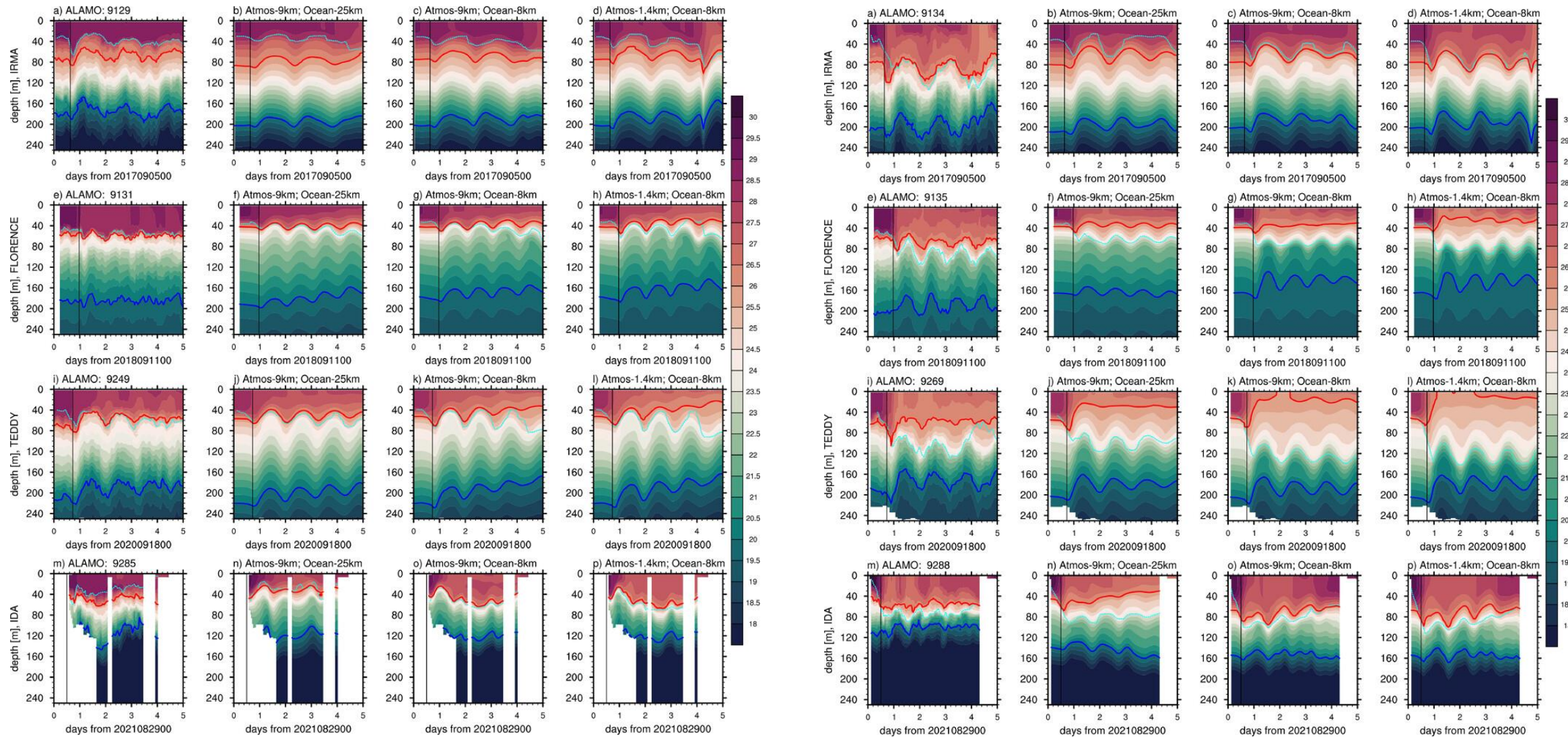
<https://www.frontiersin.org/journals/marine-science/articles/10.3389/fmars.2025.1593904/full>

Using ALAMO floats for validation of ocean TC response 1

- ALAMO floats are micro profilers which measures temperature and salinity
- High repetition frequency (~2 hours) down to a few hundreds' meters
- Air dropped for observational campaigns
- <https://journals.ametsoc.org/view/journals/mwre/153/11/MWR-D-24-0104.1.xml>



Using ALAMO floats for validation of ocean response 2



Conclusions and output

- A very limited set of ocean fields are available for users to look at, but we are working on it
 - Scores are great, but having eyes on fields are useful as well
- Monitoring of ocean data assimilation is in place and works well
 - Coupled to an automatic alarm system so missing data can be detected
- Standardizing tools across our different modelling activities to also include ocean variable will be nice
 - Use of IV-TT observations could be a good option
 - Our new ocean model upgrade will have output in GRIB2 which will also help us here
- The plan is to contribute to the IV-TT's activities and provide input to the class-4 diagnostics

Kristian Mogensen (who takes all the blame)

Senior Scientist

kristian.Mogensen@ecmwf.int

