Multivariate data assimilation across ocean and sea-ice interface

*Andrea Cipollone[1], **Jan-Erik Tesdal[1]**, **Matteo Broccoli[1]**, Doroteaciro Iovino[1], Cristina Gonzales Haro[2], Antonio Turiel[2], Manuel Arias[3], Evgeniia Makarova[2], Ferran Hernández Macia[2], Carolina Gabarró[2]*

@CP-/DA-TT workshop, Nov 26-30, 2026

1 cmcc Centro Euro-Mediterraneo sui Cambiamenti Climatici

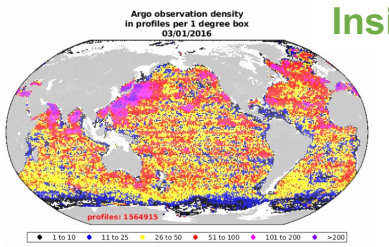
2 Institut de Ciències del Mar CSIC Consejo Superior de Investigaciones Científicas

3 Zenithal Blue Technologies Science

OUTLINE

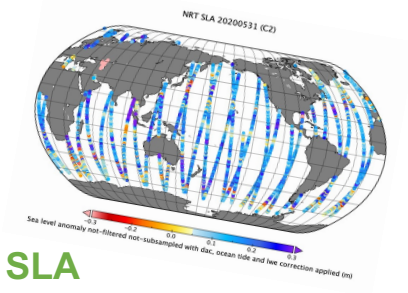
OCEAN AND SEA-ICE COMPONENTS

- ☑ Global operational Reanalysis/Forecast products at CMCC are based on OceanVar, a **3dvar DA scheme** that ingests ocean observations such as T/S/SLA (Dobricic & Pinardi 2008; Storto et al. 2011, etc.)
  - ↘ OceanVar2 (Oddo et al, 2026) entered the testing phase at global scale (see Bigoni's talk)
- ☑ Sea-ice DA introduced in a similar variational background, covarying SIC and SIT and exploiting an anamorphosis transformation to preserve the anisotropy close to the sea-ice edge (Cipollone et al, 2023)
- ↗ Assimilation of diurnal skin/subskin satellite SST retrieval via ML (Broccoli and Cipollone, 2025)
- ↗ Hybrid-BEC 3dvar configuration
- ↗ Strong Coupling in an Earth system environment

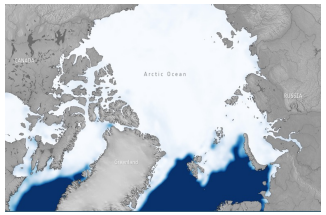


Insitu

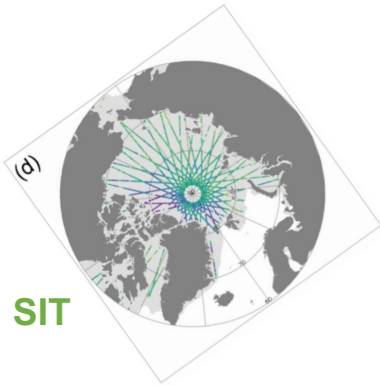
@https://argo.ucsd.edu/



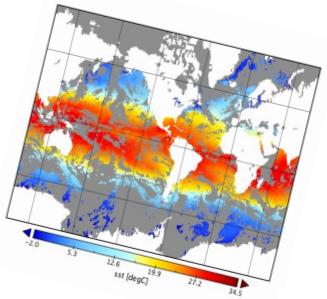
SLA



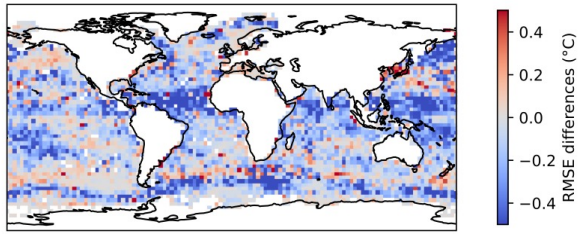
SIC



SIT

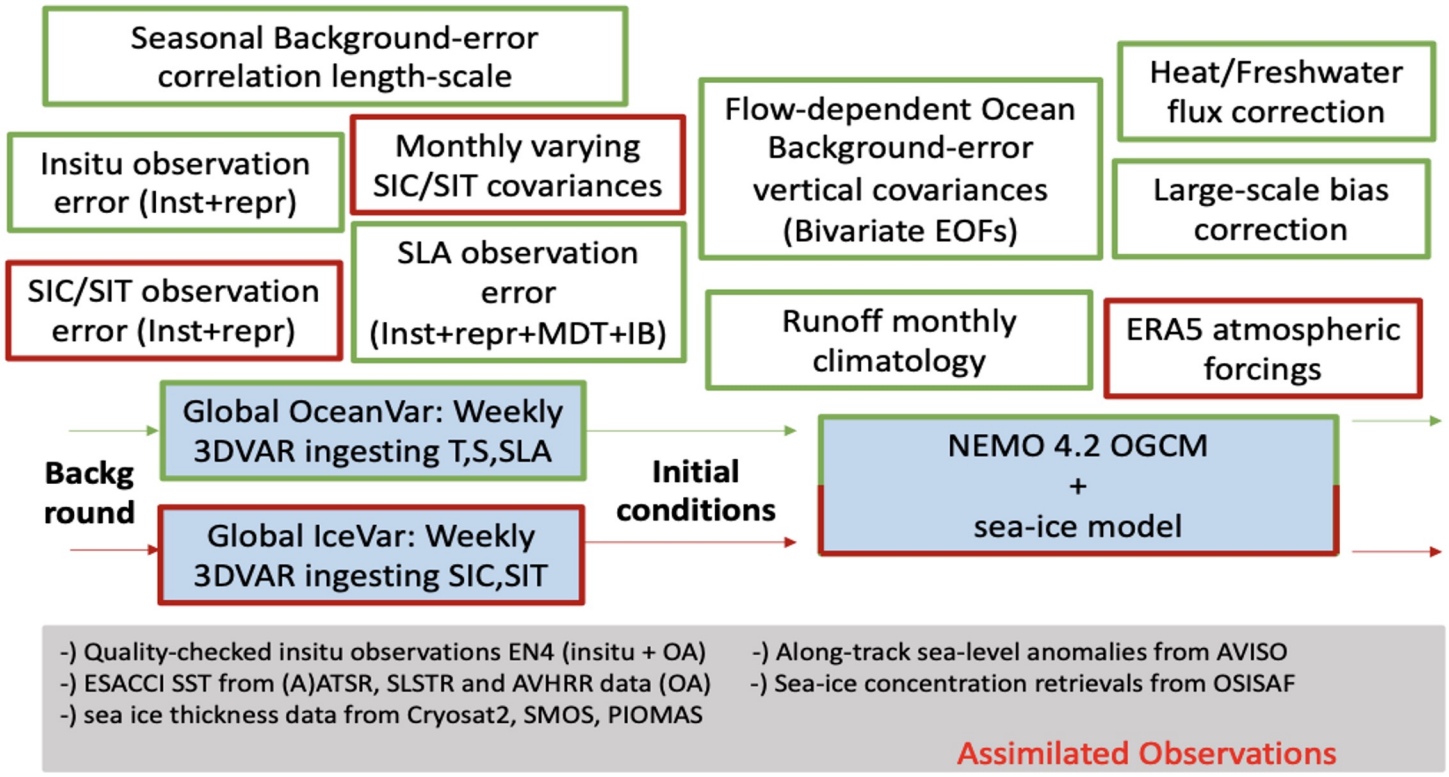


SST

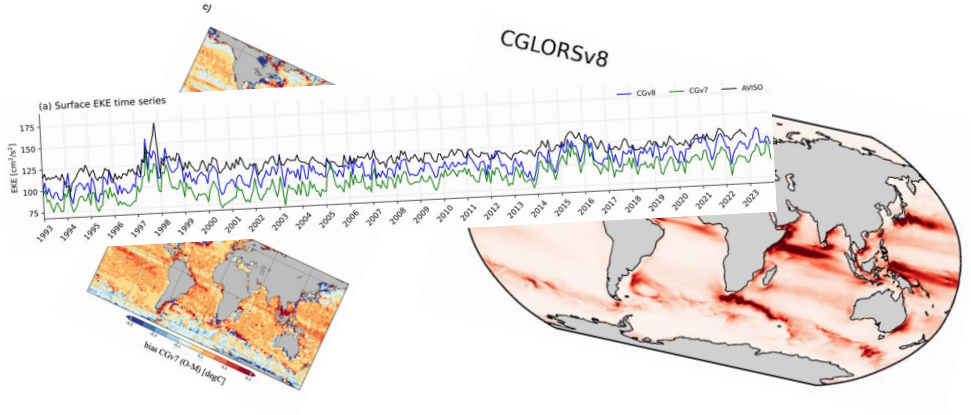


CMCC GLobal Ocean and sea ice Reanalysis System (C-GLORS)

The CMCC Global Ocean Physical Reanalysis System (C-GLORS) is used at CMCC to simulate the state of the ocean in the last decades at ¼°. New timeseries since 1993 is out!



New C-GLORSv8 is out!



- ☑ Improved thermal structures (bias) compared to previous version
- ☑ More realistic hydrography in the polar regions
- ☑ Smoother and consistent deep and abyssal circulation

ADVANCES IN PRESENT IN-HOUSE SYSTEM

Complexity: Strongly Coupled 3Dvar

$$J(v) = \frac{1}{2}\boldsymbol{v}^{\mathrm{T}}\boldsymbol{v} + \frac{1}{2}(\mathbf{H}\mathbf{V}\boldsymbol{v} - \boldsymbol{d})^{\mathrm{T}}\mathbf{R}^{-1}(\mathbf{H}\mathbf{V}\boldsymbol{v} - \boldsymbol{d})$$

$$\boldsymbol{B} = \boldsymbol{V}\boldsymbol{V}^{\mathrm{T}} \qquad \boldsymbol{V} = \boldsymbol{V}_{\mathrm{gICE}\rightarrow\mathrm{ICE}}\boldsymbol{V}_h\boldsymbol{V}^{\mathrm{gSIC:gSIT:}T:S}$$

Observation Operator $\boldsymbol{H} = (\boldsymbol{D}^{\mathrm{SLA}}, \boldsymbol{I}^{T}, \boldsymbol{I}^{S}, \boldsymbol{I}^{\mathrm{SSS}}, \boldsymbol{H}^{\mathrm{SST}}_{\mathrm{ML}}, \boldsymbol{H}^{\mathrm{SIT}}_{\mathrm{ML}}, \mathrm{H}^{\mathrm{SIC}}_{\mathrm{ML}})$

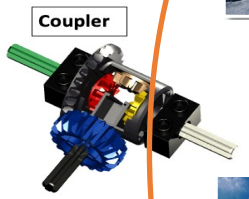
Collab with UniBo

Completed/ Tuning phase

Earth system Model



Atmosphere
ERA5 hourly data





Sea Ice
Community Ice Code
CICE6

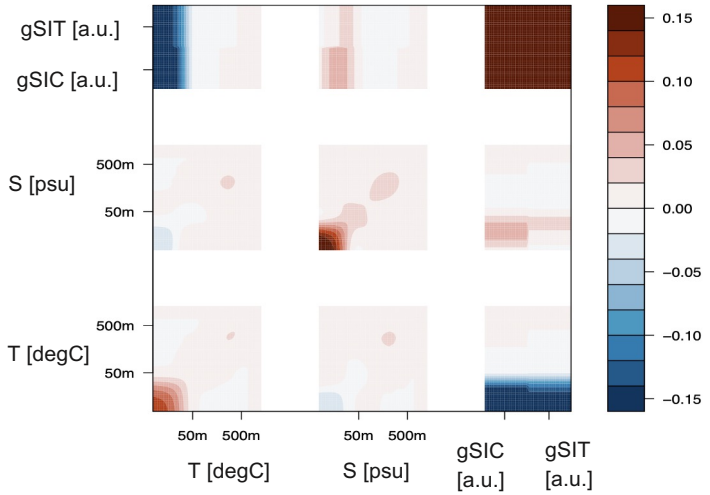


Land / Vegetation
Runoff climatology

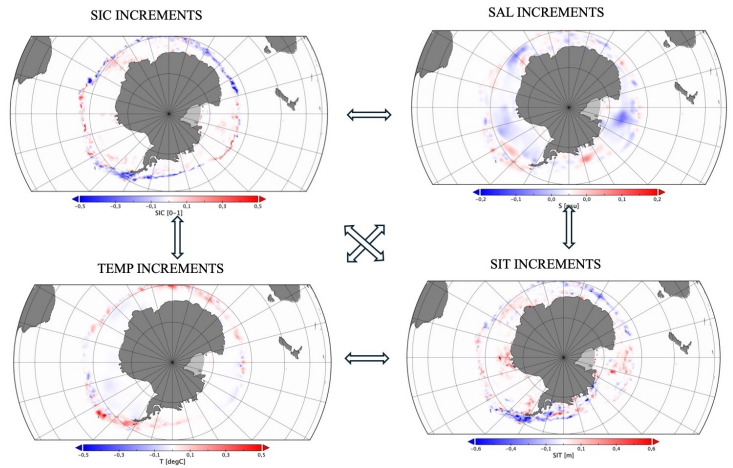


Ocean
Nucleus for European Modelling of the Ocean
NEMO 4.2

Data Assimilation



Increments



ASSIMILATION OF DIURNAL SST DATA

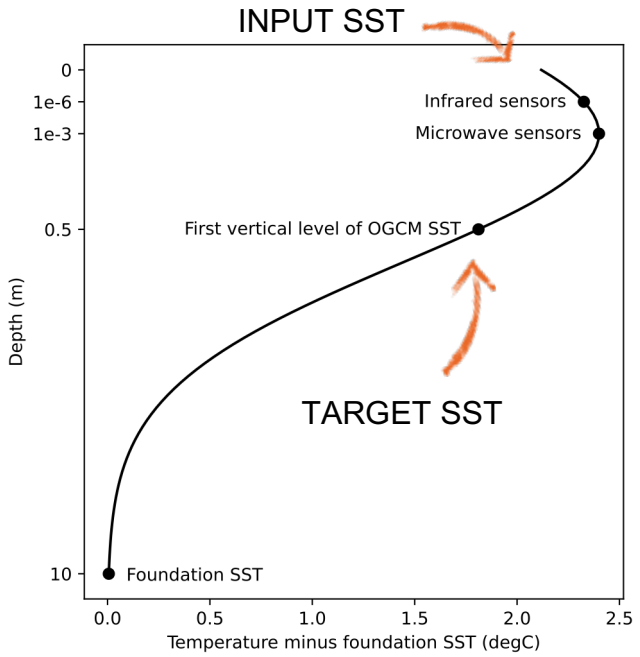
Why assimilating SST data ?

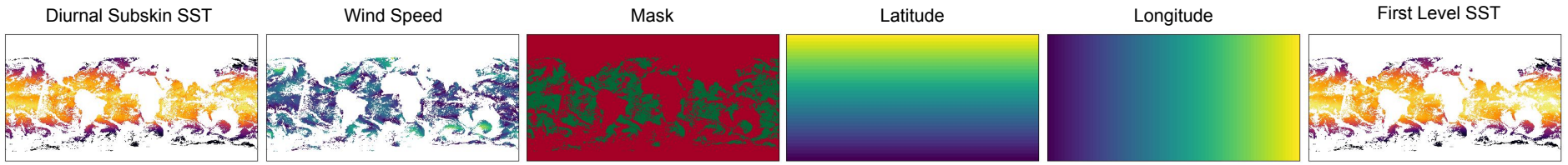
- The idea is to assimilate diurnal/nocturnal SST data that can **possibly** replace the SST nudging **or acts coherently** with it (Target dataset: ESACCI SST, daily mean)

How to use deep learning algorithms?

- Use machine learning to construct:

1. Forward operator and assimilate **unbiased** (ESACCI-like) subskin SST (metop B)
2. Inverse operator employed as **observation operator** to assimilate subskin SST





Input

- Diurnal L3 global **subskin SST** from AVHRR's infrared channels on MetOp satellites produced by OSI SAF
- **Wind speed** at 10m
- **Mask** for land and clouds
- **Latitude** and **longitude** grids

Ground truth

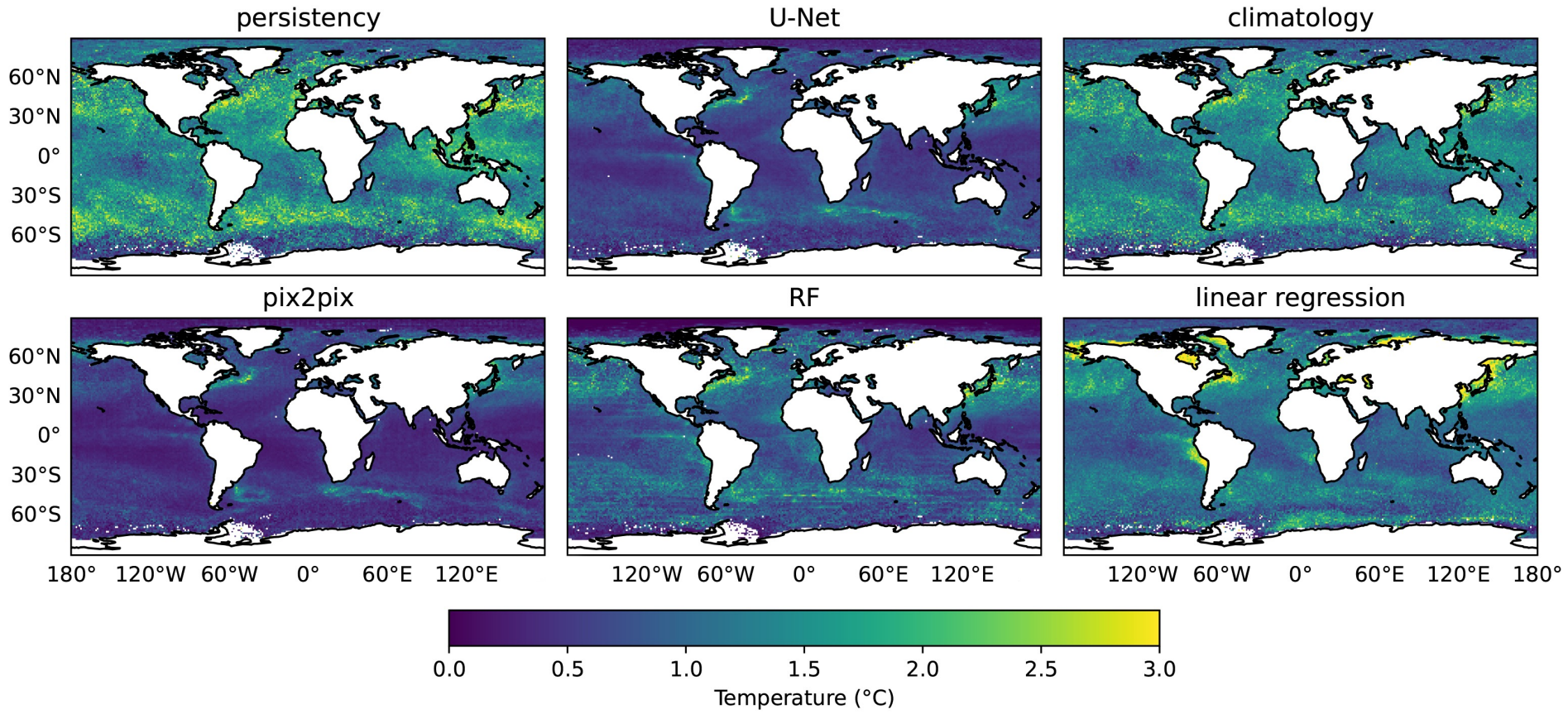
- L4 global **SST** from ESA SST CCI by CMEMS
  - Masked as input data

Training on one year of data (2017)
80% training, 20% testing

ASSIMILATION OF DIURNAL SST DATA

Different approaches considered to construct the operator:

- U-Net
- pix2pix
- Random forest
- Persistency (subskin)
- Multivariate linear regression
- Climatological bias estimation



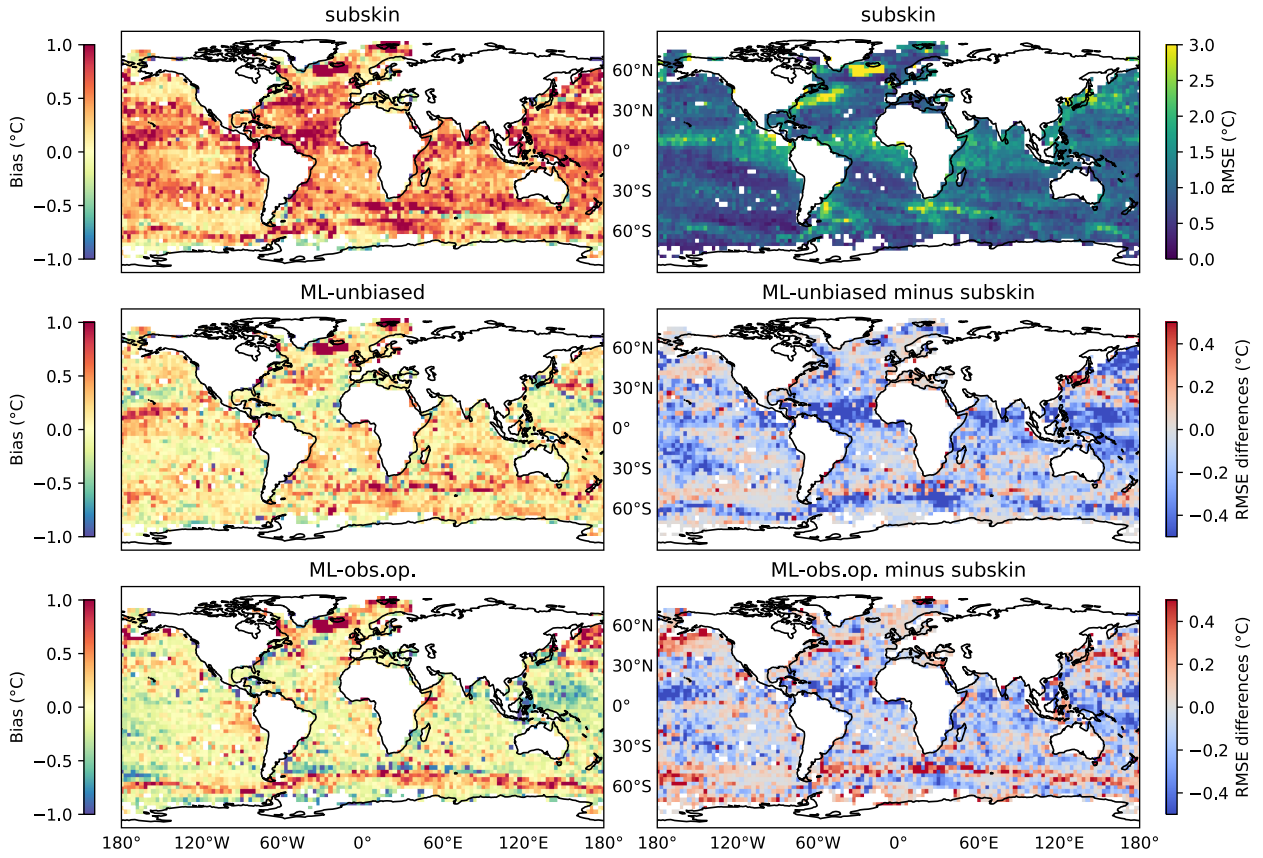
RMSE between the predictions of the different models and the ground-truth on the test set

ASSIMILATION OF DIURNAL SST DATA

Exp. Name	Analysis Increments	Observations	Length	Description
subskin	T, S	Skin SST, SLA,T,S	1y	Weakly coupled (OceanVar only), daily assim
ML-unbias	T, S	ML-unbias SST, SLA,T,S	1y	Weakly coupled (OceanVar only), daily assim
ML-obs	T, S	ML-operator SST,SLA,T,S	1y	Weakly coupled (OceanVar only)

Temperature maps of bias and RMSE [0-100m] against in situ observations

- ML approaches manage to reduce the bias (against insitu observations) from subskin experiment
- ML-unbiased scheme is mostly beneficial in Tropical region, improvement exceeding 0.4° in some areas
- ML-obs.op. less beneficial in the Tropics and it generates higher errors in extra-tropic regions with strong mesoscale activities (ACC, Kuroshio, etc)



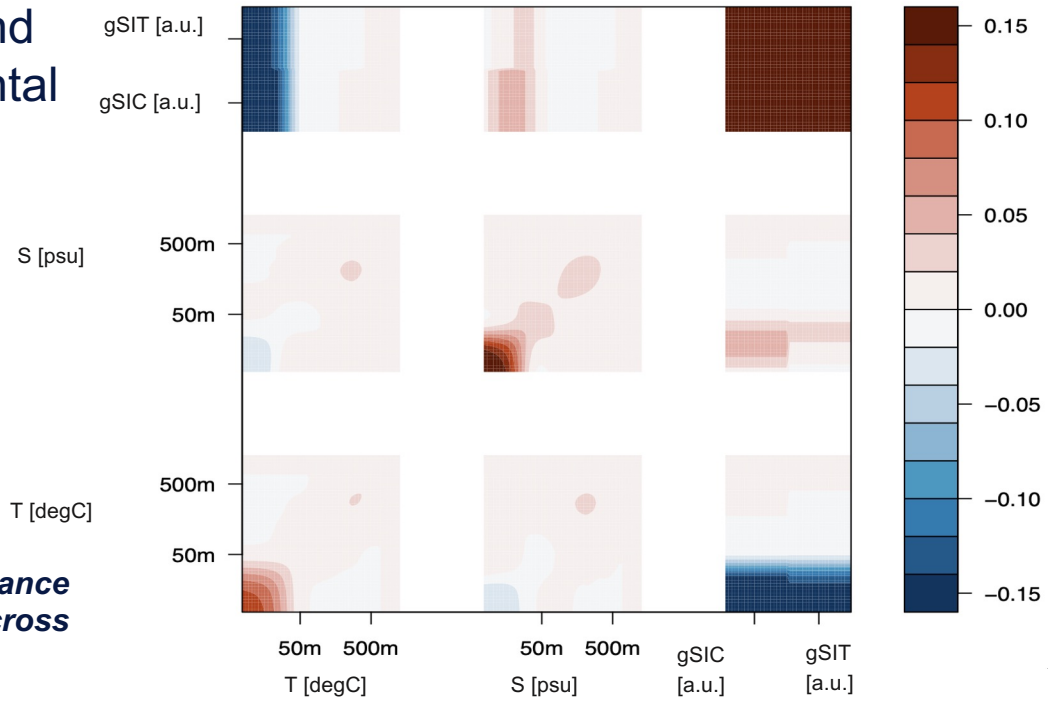
***blue** means the T RMSE is lower than direct subskin assim, **red** means it is higher*

ASSIMILATION OF SEA-ICE DATA ACROSS INTERFACE

How we assimilate sea-ice characteristics

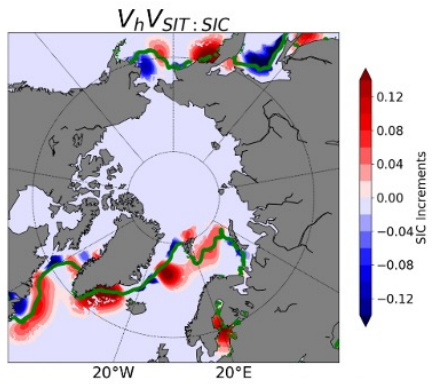
- We use an anamorphism (**quantile mapping**) to minimize into gaussian space (gSIC/gSIT) preserving the physical correlation without disrupting the strong anisotropy close to sea-ice edge

- gSIC/gSIT are covaried with T and S for each horizontal point

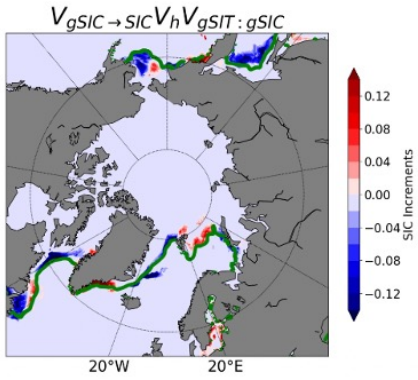


Example of vertical covariance matrix in one point (across vertical column)

SIC increments without ananmorphism

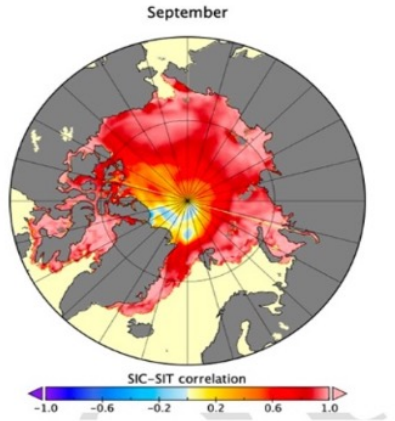


SIC increments with ananmorphism

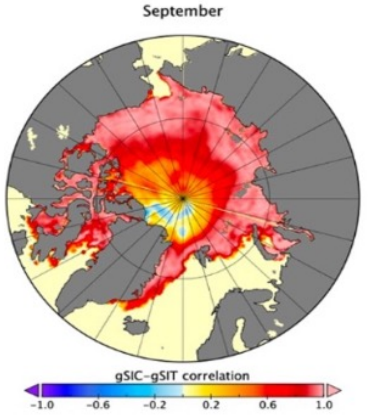


H-lenght=200km

SIC/SIT correlation without



SIC/SIT correlation with



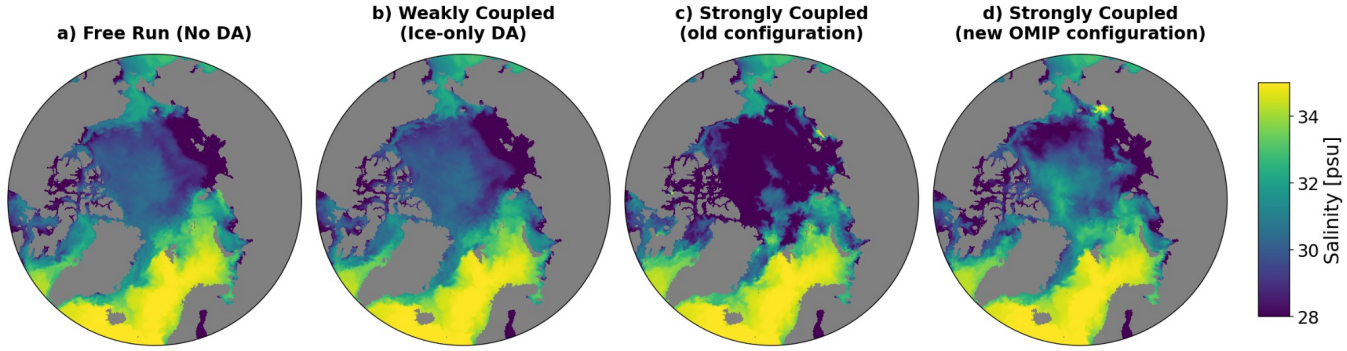
ASSIMILATION OF SEA-ICE DATA ACROSS INTERFACE

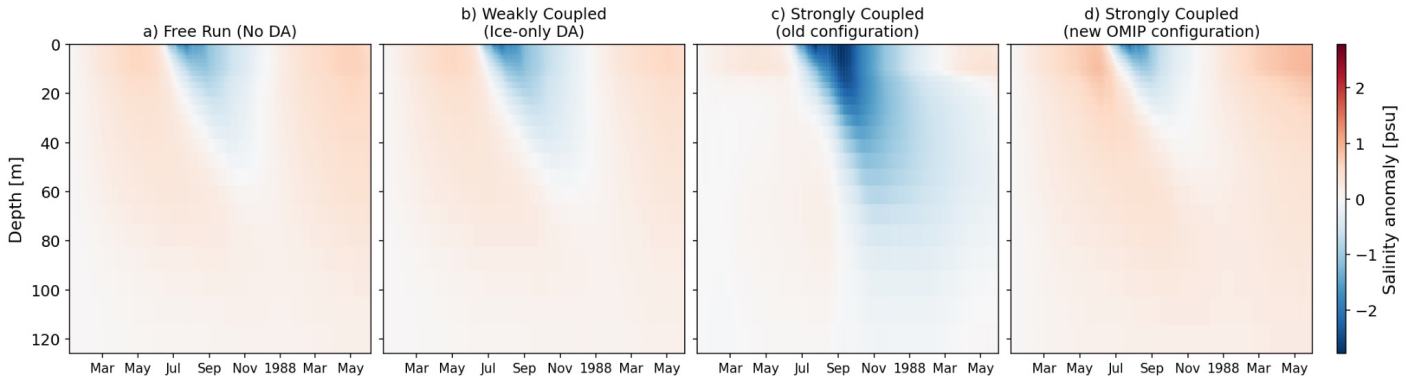
Addressing spurious salinity corrections

Exp. ID	Analysis Increments	Observations	Length	Description
Weak	SIC SIT	SIC, SIT	1y	BEC from 10y statistics
Strong-old	SIC, SIT,T,S	SIC, SIT	1y	BEC from 10y statistics
Strong new	SIC, SIT, T, S	SIC, SIT	1y	BEC from 60y statistics

- Old config. produces substantially fresher surface waters in the interior Arctic at the end of the melting season
- Increasing the statistics of BEC, reduces the strong spurious correlation among ocean/sea-ice variables
- In the new configuration isolated regions remain where salinity is potentially still too affected by the sea-ice corrections

Arctic surface salinity in September 1987





Evolution of average salinity in the whole Arctic Ocean along depth

ASSIMILATION OF SEA-ICE DATA ACROSS INTERFACE

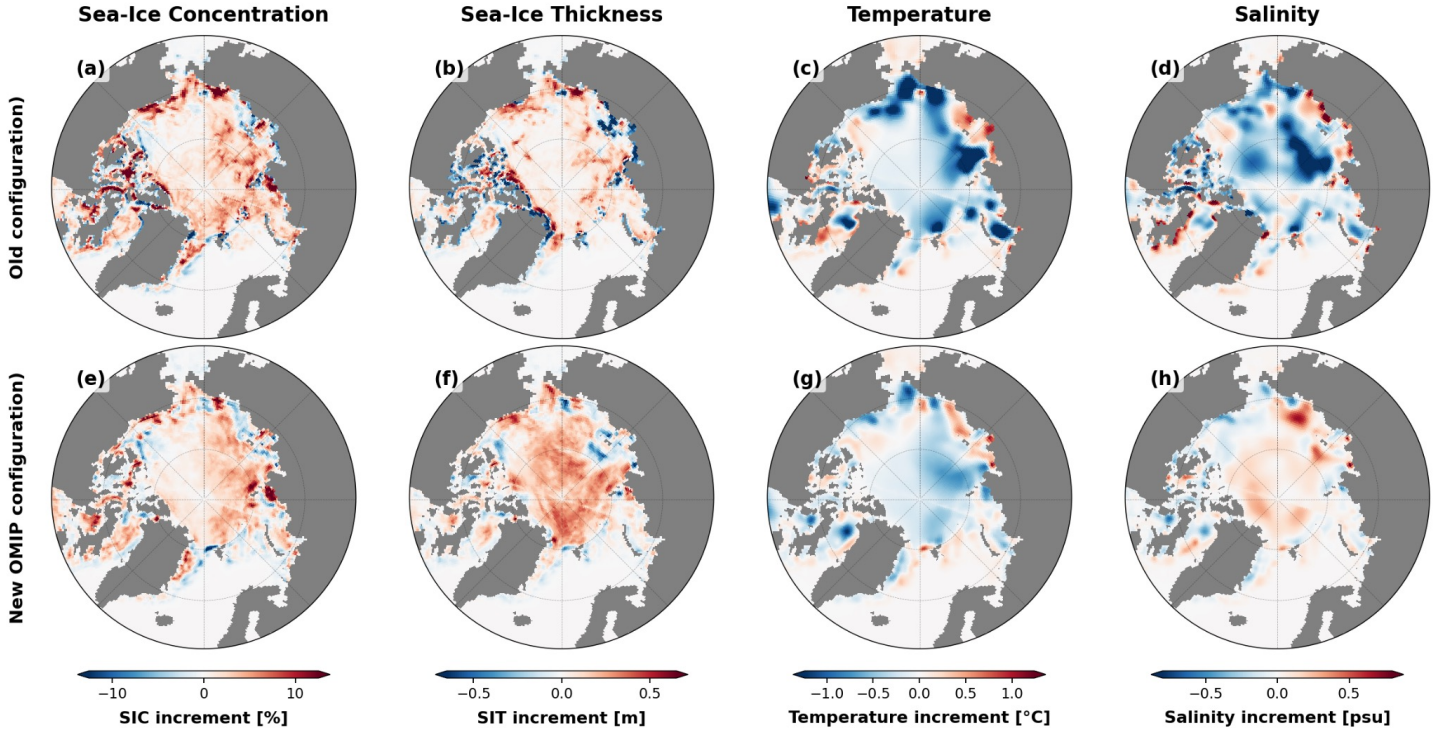
esa COMET

Addressing spurious salinity corrections

Exp. ID	Analysis Increments	Observations	Length	Description
Weak	SIC SIT	SIC, SIT	1y	BEC from 10y statistics
Strong-old	SIC, SIT,T,S	SIC, SIT	1y	BEC from 10y statistics
Strong new	SIC, SIT, T, S	SIC, SIT	1y	BEC from 60y statistics

- Old config. produces substantially fresher surface waters in the interior Arctic at the end of the melting season
- Increasing the statistics of BEC, reduces the strong spurious correlation among ocean/sea-ice variables
- In the new configuration, isolated regions remain where salinity is potentially still too affected by the sea-ice corrections

Averaged increment over 1987

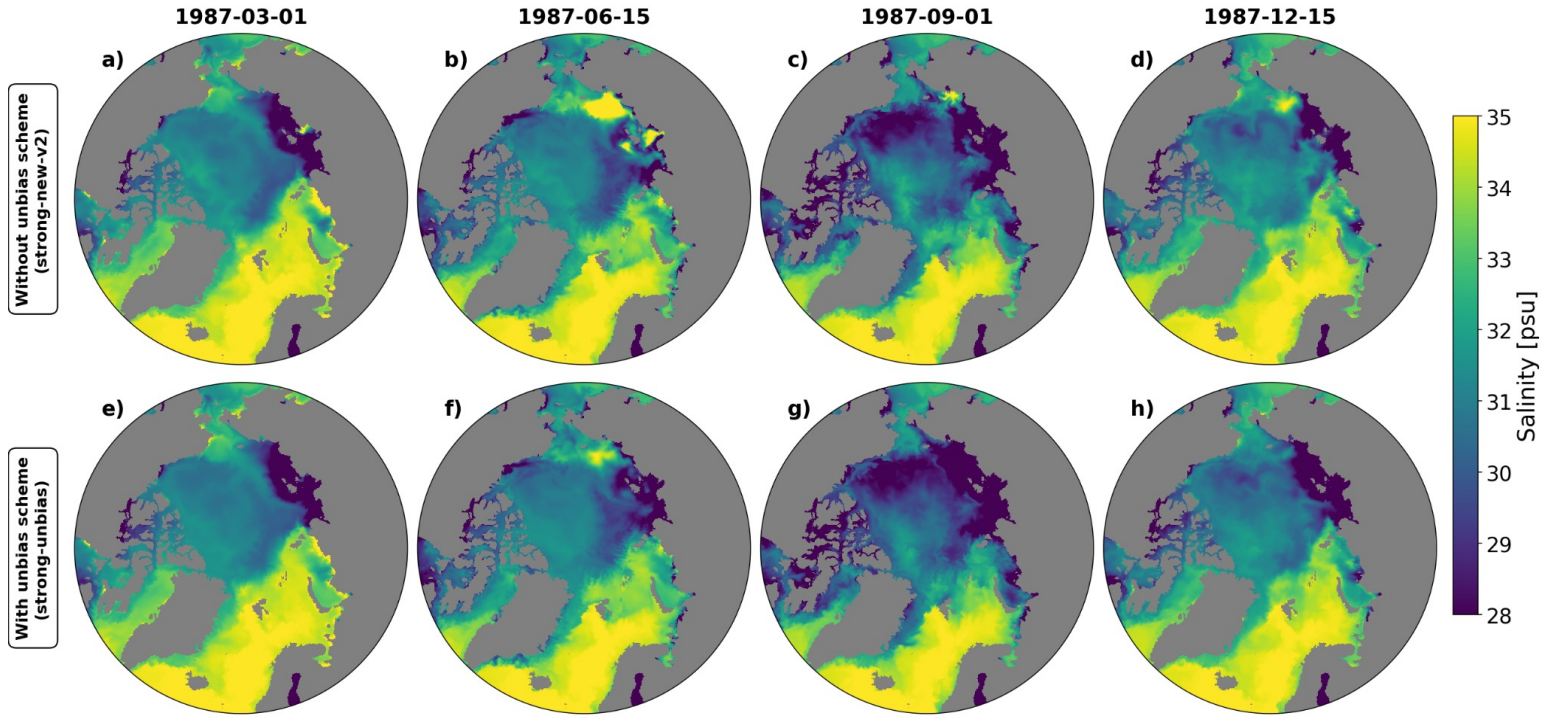


ASSIMILATION OF SEA-ICE DATA ACROSS INTERFACE

Addressing spurious salinity corrections

Exp. ID	Analysis Increments	Observations	Length	Description
Strong new	SIC, SIT, T, S	SIC, SIT	1y	BEC from 60y statistics
Strong new (unbias)	SIC, SIT, T, S	SIC, SIT	1y	BEC from 60y statistics. Climatological unbias by averaging monthly SIC/SIT increments

- Isolated regions remain with **large systematic** bias in different regions
- The application of a **climatological unbias** to SIC and SIT reduces the large local variability although better approaches can be envisaged via ML (see SST)



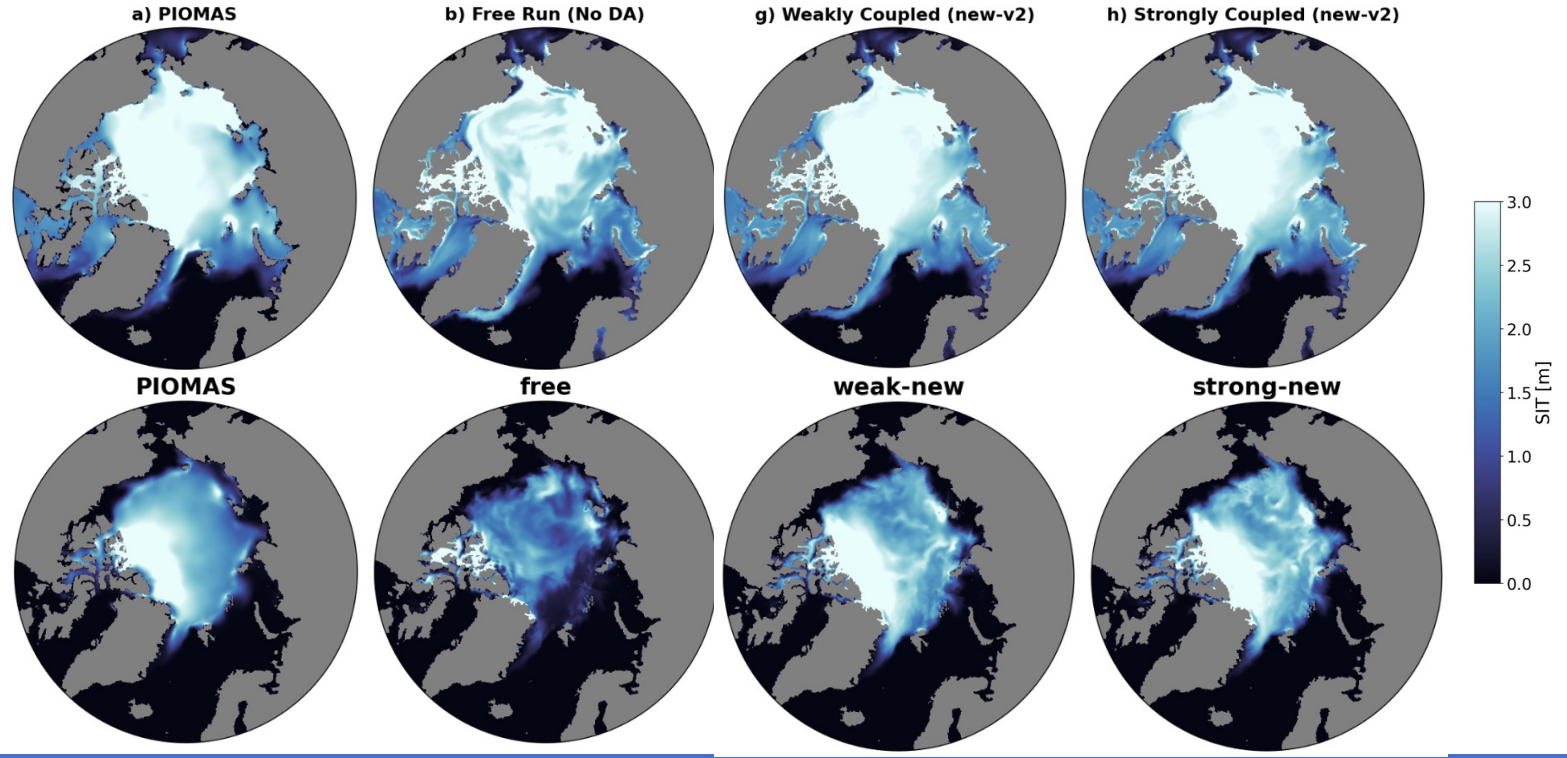
Maps of Arctic surface in different period for strongly coupled experiment with and without unbias procedure

ASSIMILATION OF SEA-ICE DATA ACROSS INTERFACE

Strong vs Weak coupling in sea-ice variables

Exp. ID	Analysis Increments	Observations	Length	Description
free	–	–	1y	Free run with no data assimilation (DA)
weak	SIC, SIT	SIC, SIT	1y	Weakly coupled (OceanIceVar)
strong	SIC, SIT, T, S	SIC, SIT	1y	Strongly coupled (OceanIceVar)

- Maps of sea-ice thickness show a robust and similar improvement w.r.t free-run both for the weak and strong coupling



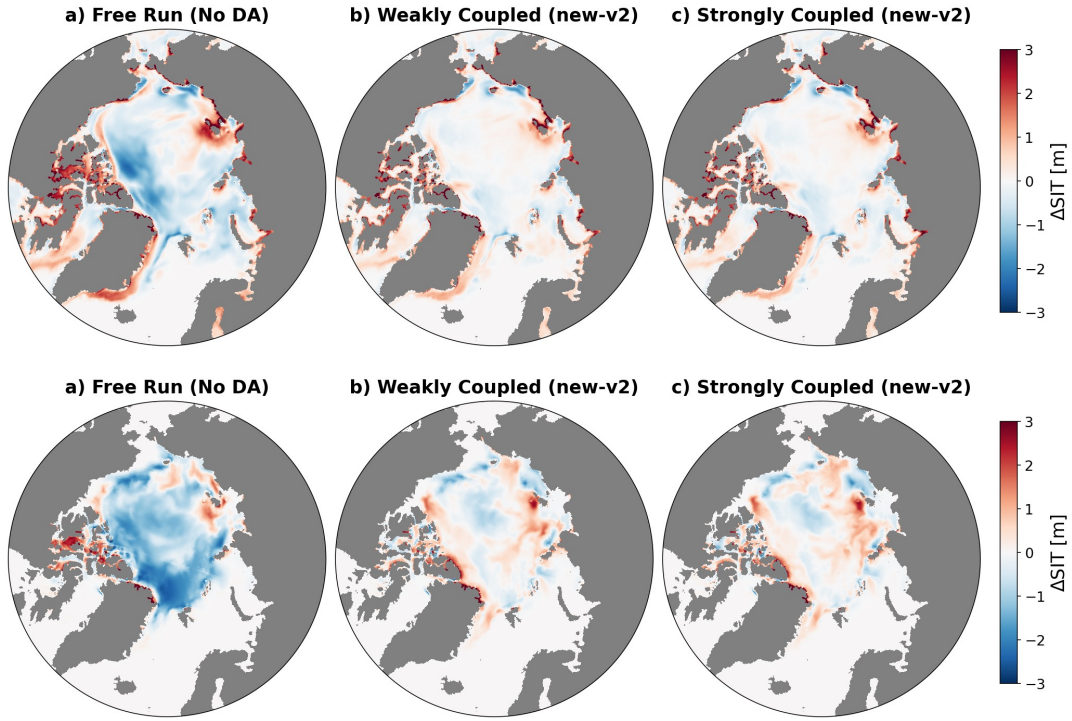
Maps of Arctic SIT in April and September 1987 for different experiments

ASSIMILATION OF SEA-ICE DATA ACROSS INTERFACE

Strong vs Weak coupling in sea-ice variables

Exp. ID	Analysis Increments	Observations	Length	Description
free	–	–	1y	Free run with no data assimilation (DA)
weak	SIC, SIT	SIC, SIT	1y	Weakly coupled (OceanIceVar)
strong	SIC, SIT, T, S	SIC, SIT	1y	Strongly coupled (OceanIceVar)

- Maps of sea-ice thickness bias show similar skill w.r.t free-run both for the weak and strong coupling



Maps of Arctic SIT bias in April and September 1987 for different experiments

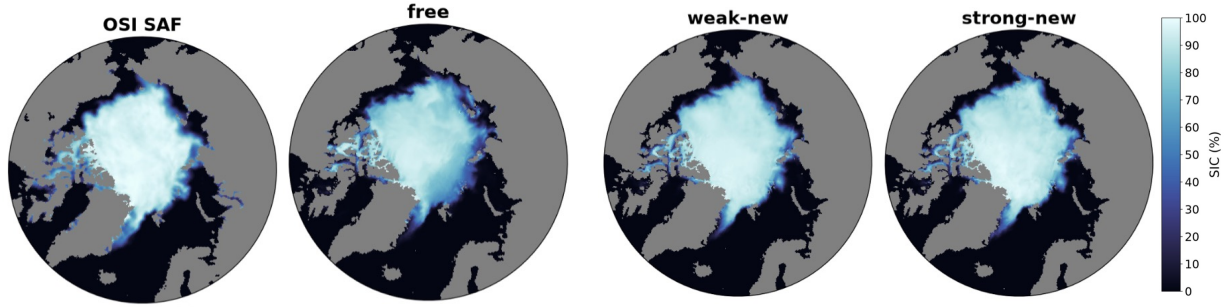
ASSIMILATION OF SEA-ICE DATA ACROSS INTERFACE

Strong vs Weak coupling in sea-ice variables

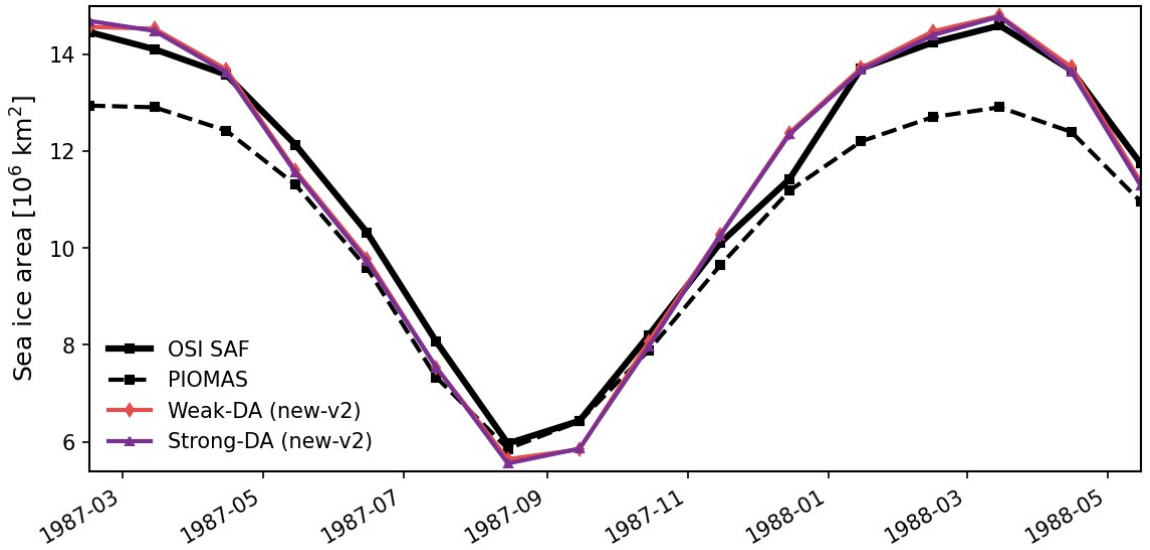
Exp. ID	Analysis Increments	Observations	Length	Description
free	–	–	1y	Free run with no data assimilation (DA)
weak	SIC, SIT	SIC, SIT	1y	Weakly coupled (OceanIceVar)
strong	SIC, SIT, T, S	SIC, SIT	1y	Strongly coupled (OceanIceVar)

- Sea-ice concentration show a robust and similar improvement w.r.t free-run both for the weak and strong coupling

Maps of Arctic SIC in September 1987 for different experiments

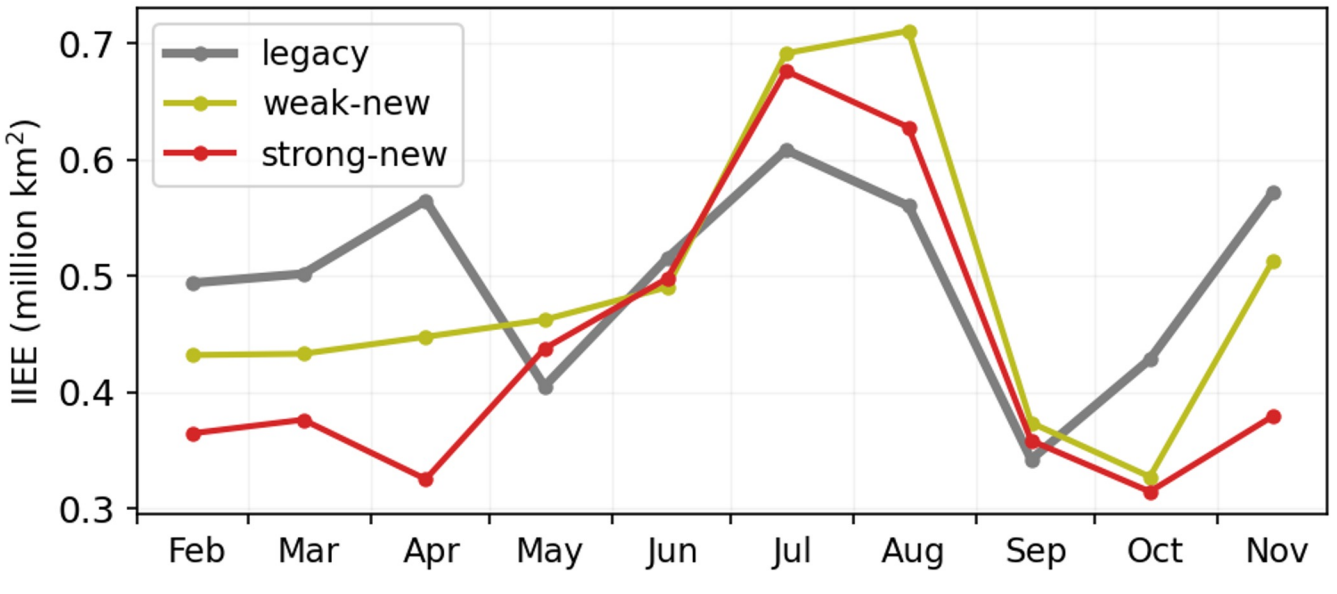


Evolution of SIA for different experiments



WEAK VS STRONG COUPLING (IIEE)

Integrated Ice Edge Error



$$IIEE = O + U$$

O and **U** are the over- and underestimated area of sea ice cover

$$O = \int_A \max(c_{mod} - c_{obs},\ 0)\ dA$$
$$U = \int_A \max(c_{obs} - c_{mod},\ 0)\ dA$$

$\boldsymbol{c_{mod}}$ and $\boldsymbol{c_{obs}}$ are the sea ice extent for the model and the OSI SAF observations

- Ice-edge position largely controlled by ocean temperature
- Strongly coupled DA reduces IIEE vs weakly coupled throughout the period
- During fall and winter months, the strongly coupled run also has lower IIEE compared to the legacy version (i.e, C-GLORSv8)

Website: https://climate.esa.int/en/observations-and-modelling/COMET/

Workshop on climate observation and modeling : January 12-13, 2027 @CSIC

Thank you!