Ocean Data Assimilation

DA training course 2019

Hao Zuo

With input from M Balmaseda, K Mogensen, M Chrust, P Browne, E de Boisseson, R Buizza and Coupled DA team

hao.zuo@ecmwf.int

ECMWF

Outline

1. **General remarks**
2. **Ocean DA in ECMWF**
   - NEMOVAR
   - DA: In-situ and bias correction
   - DA: Sea-level anomaly
   - DA: SST and SIC
3. **Application of ODA in ECMWF**
4. **Further development of ODA**

Why do we do Ocean DA?

- **Forecasting: initialization of coupled models**
  - NWP, monthly, seasonal, decadal
  - Seasonal forecasts need calibration
- **Towards coupled DA system (weakly -> quasi-strong -> strong …)**
  - **See Phil's presentation**
- **Climate application: reconstruct & monitor the ocean *(re-analysis)***
- ***Verification/evaluation of Global Ocean observing network (OSE/OSSE)***
- **Monitor/forecast the ocean mesoscale and biochemistry**
  - Defence, commercial applications (oil rigs …), safety and rescue, environmental (algii blooms, spills)

Ocean versus Atmosphere: some facts

- **Spatial/time scales** The radius of deformation in the ocean is small (~30km) compared to the atmosphere (~3000km).

  Smaller spatial scales and Longer time scales

- **Ocean is a data sparse system**, in-situ observation is limited and mostly covers upper ocean only, satellite observation only covers ocean surface.

- **The ocean is forced at the surface and land boundary,** by the wind/waves, heating/cooling and fresh-water fluxes

  Uncertainty in forcing fluxes contributes to uncertainty in model results.

- **The ocean is strongly stratified in the vertical,** although deep convection also occurs

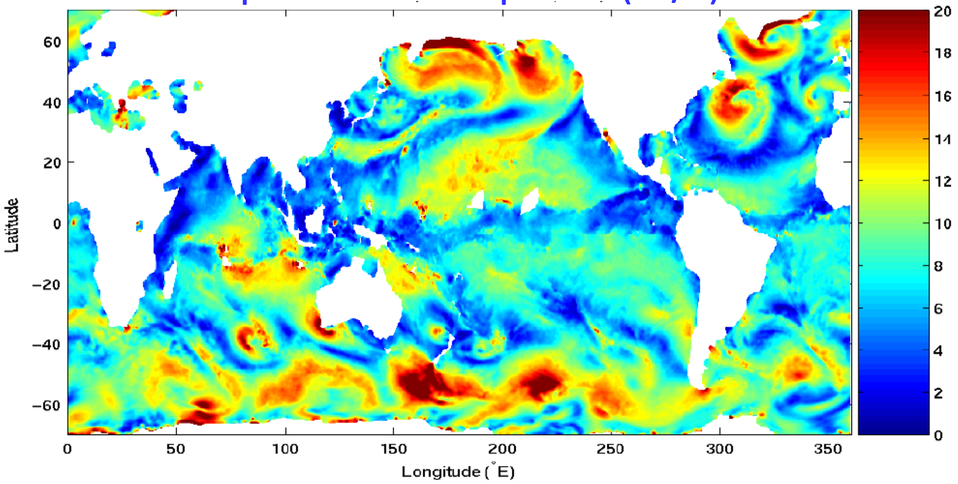
  Density is determined by Temperature and Salinity

- **The ocean has continental boundaries;** dealing with them is not trivial in data assimilation

Ocean spatial scales

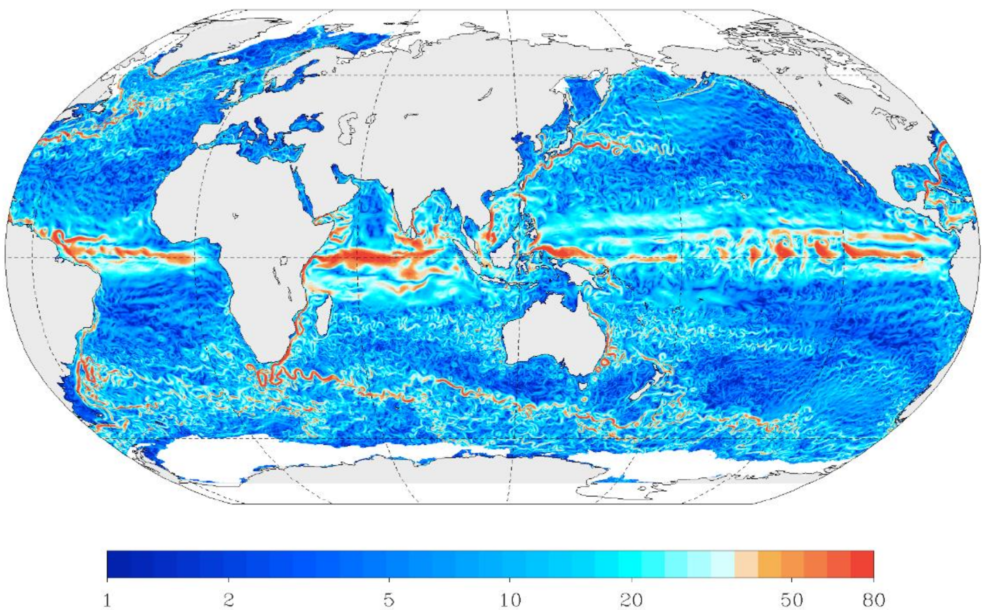
(Zhang et al.,)

Atmospheric wind speed (m/s)



Wind speed up to 20m/s
With large spatial scales

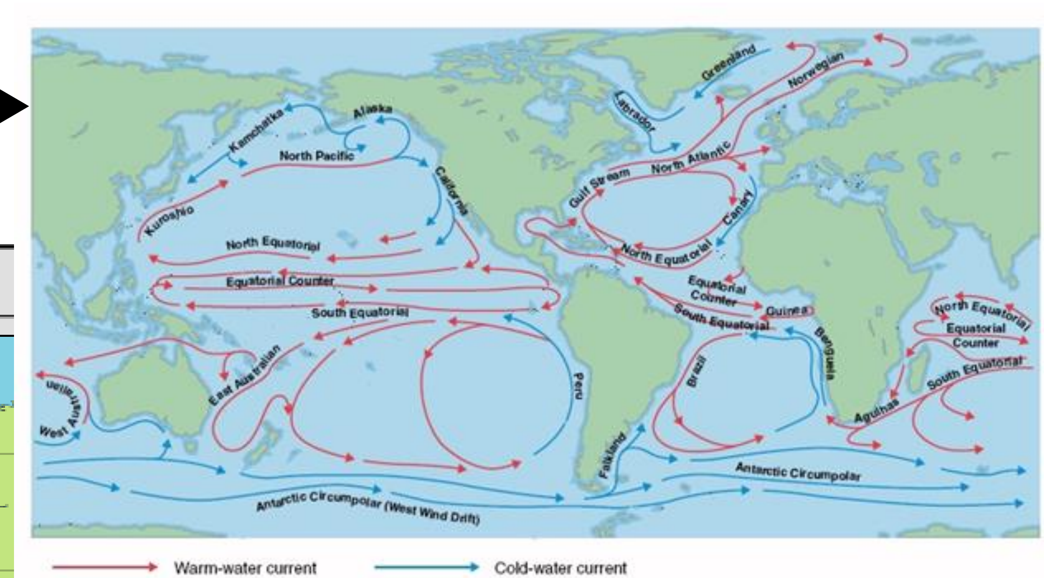
1/12 NEMO Ocean current speed (cm/s)

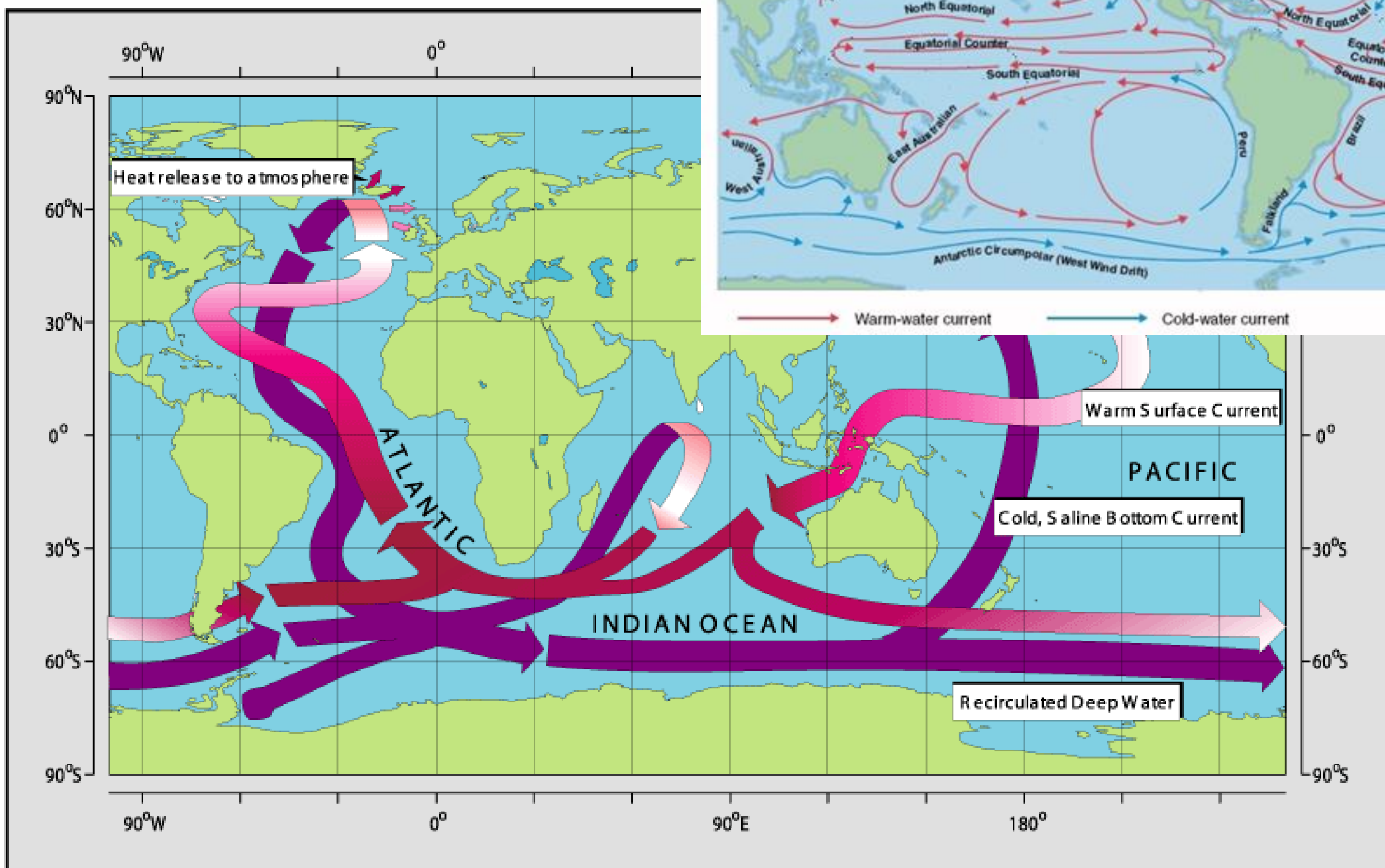


Ocean current speed up to 1m/s
With smaller spatial scales

Ocean time scales: from hours to centuries

Wind Driven: Gyres, Western Boundary Currents, Upwelling regions (coastal, equatorial), Ekman pumping and subduction





Density Driven:

Thermohaline

Circulation

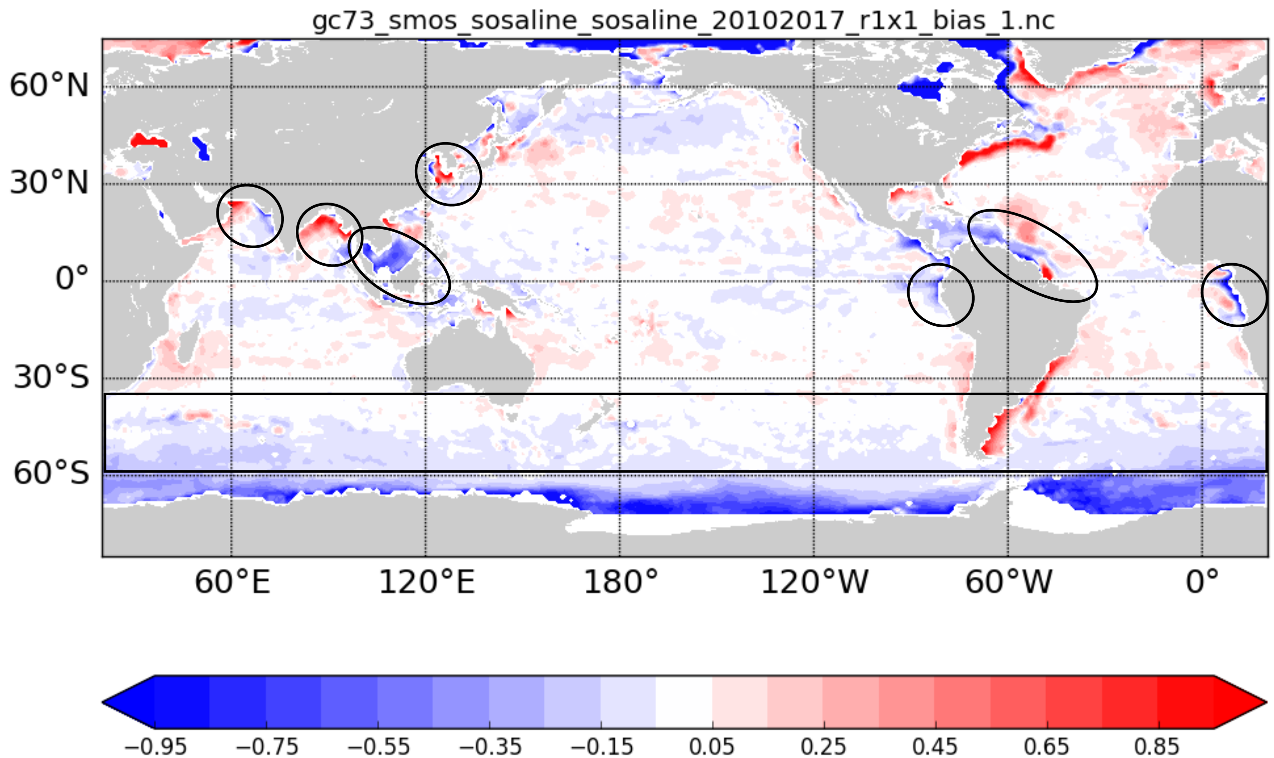
Ocean is forced by external forcings

Simulated Sea Surface Salinity error is affected by

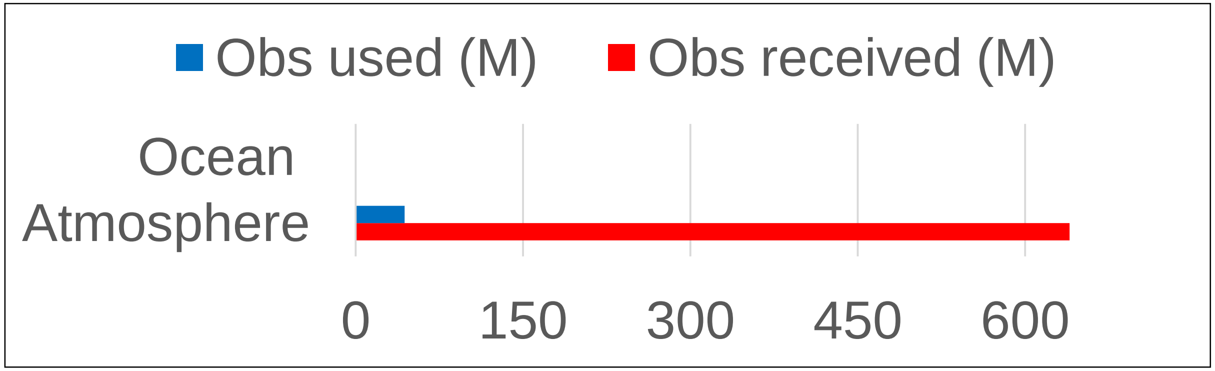
- Precipitation error in ERA-int forcing
- River runoff error in land freshwater input

And model error

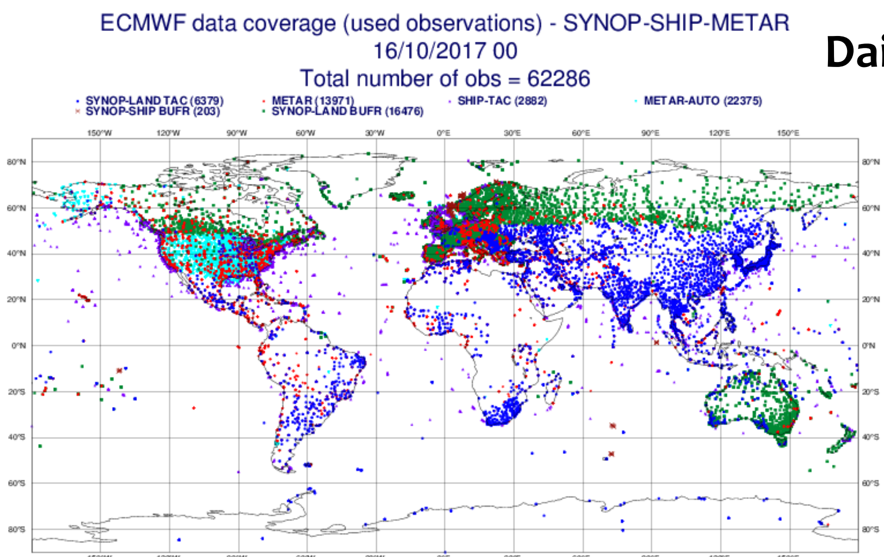
Mean SSS bias in ORAS5 (Zuo et al., 2019)



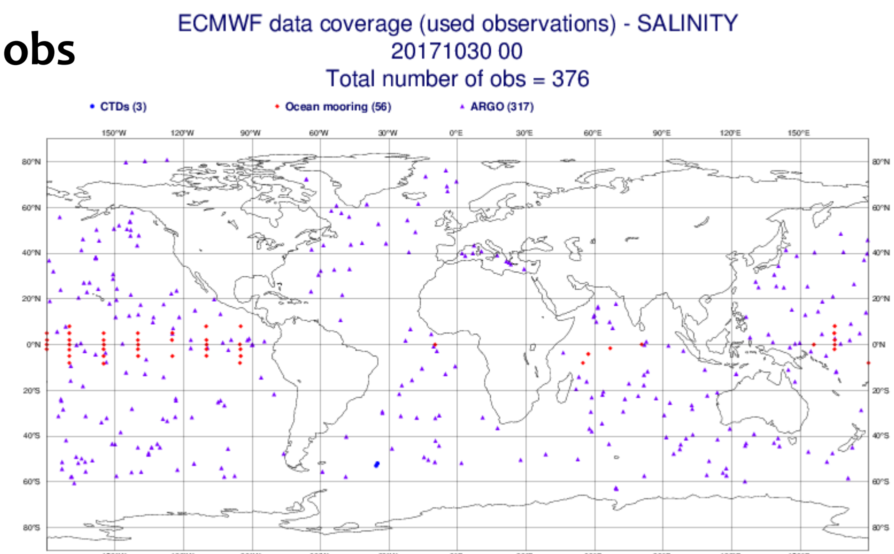
Ocean is a data sparse system



Ocean obs is about 1/1000 to 1/10000 smaller than Atmospheric obs



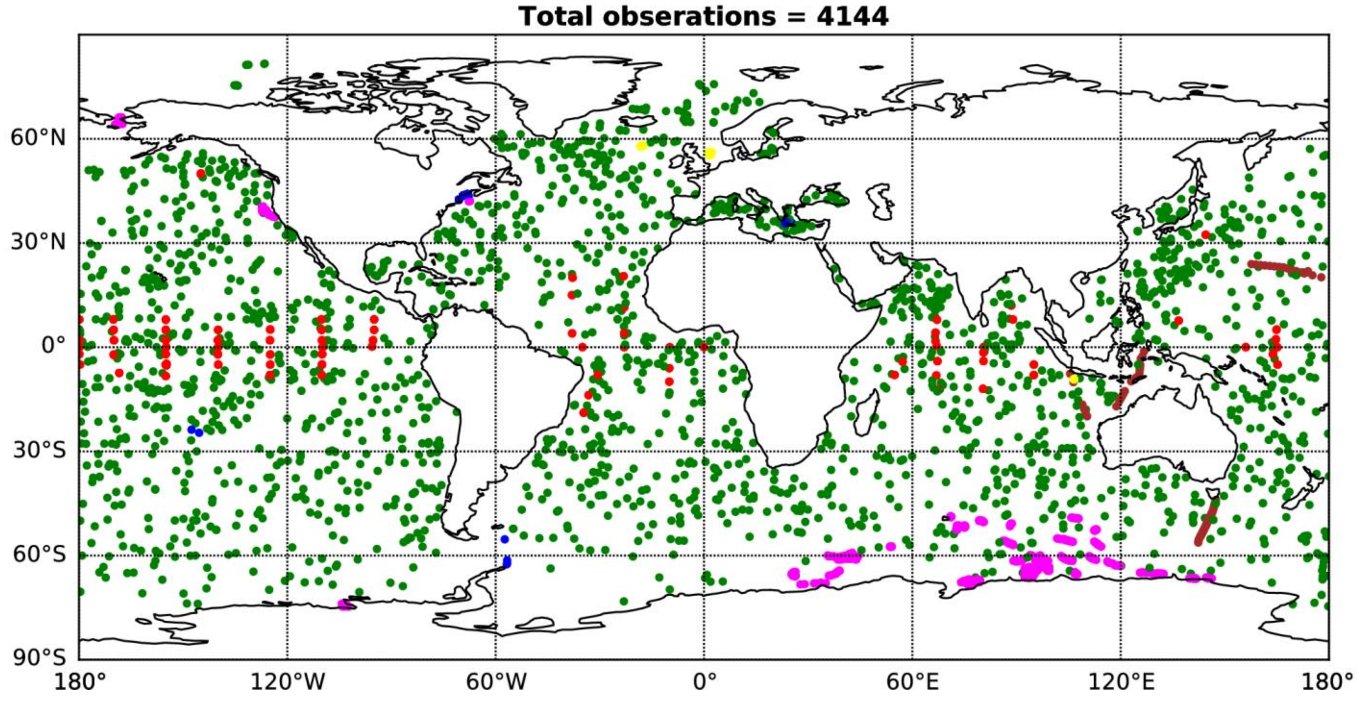
Daily obs



The Global Ocean Observing System

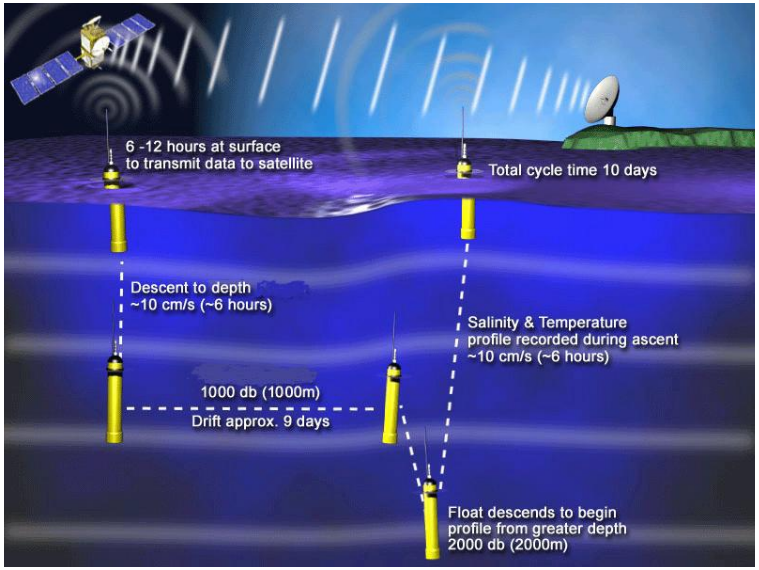
Ocean in-situ observations used in OCEAN5 (5-days in Feb 2019)

CTD:450	APB:0	UOR:0	Seaglider:124	hres CTD:0
Argo:2149	XBT:90	Mammals:931	Mooring:400	MBT:0

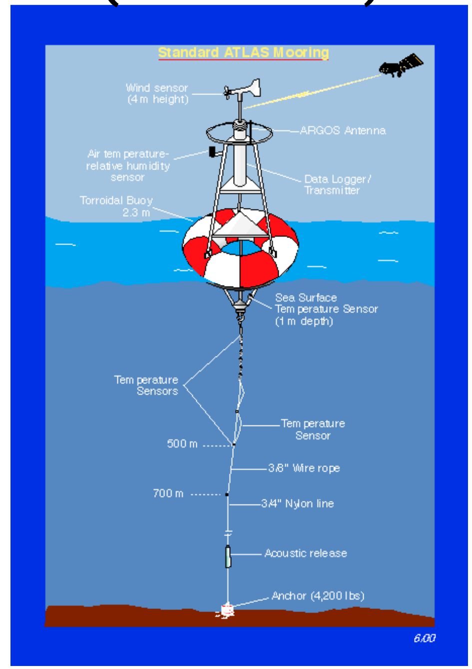


Ocean in-situ observations

Argo

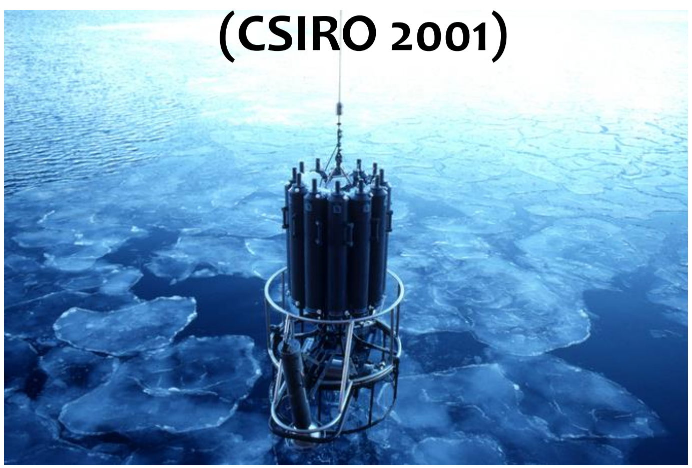


Moored buoy (PMEL 2018)



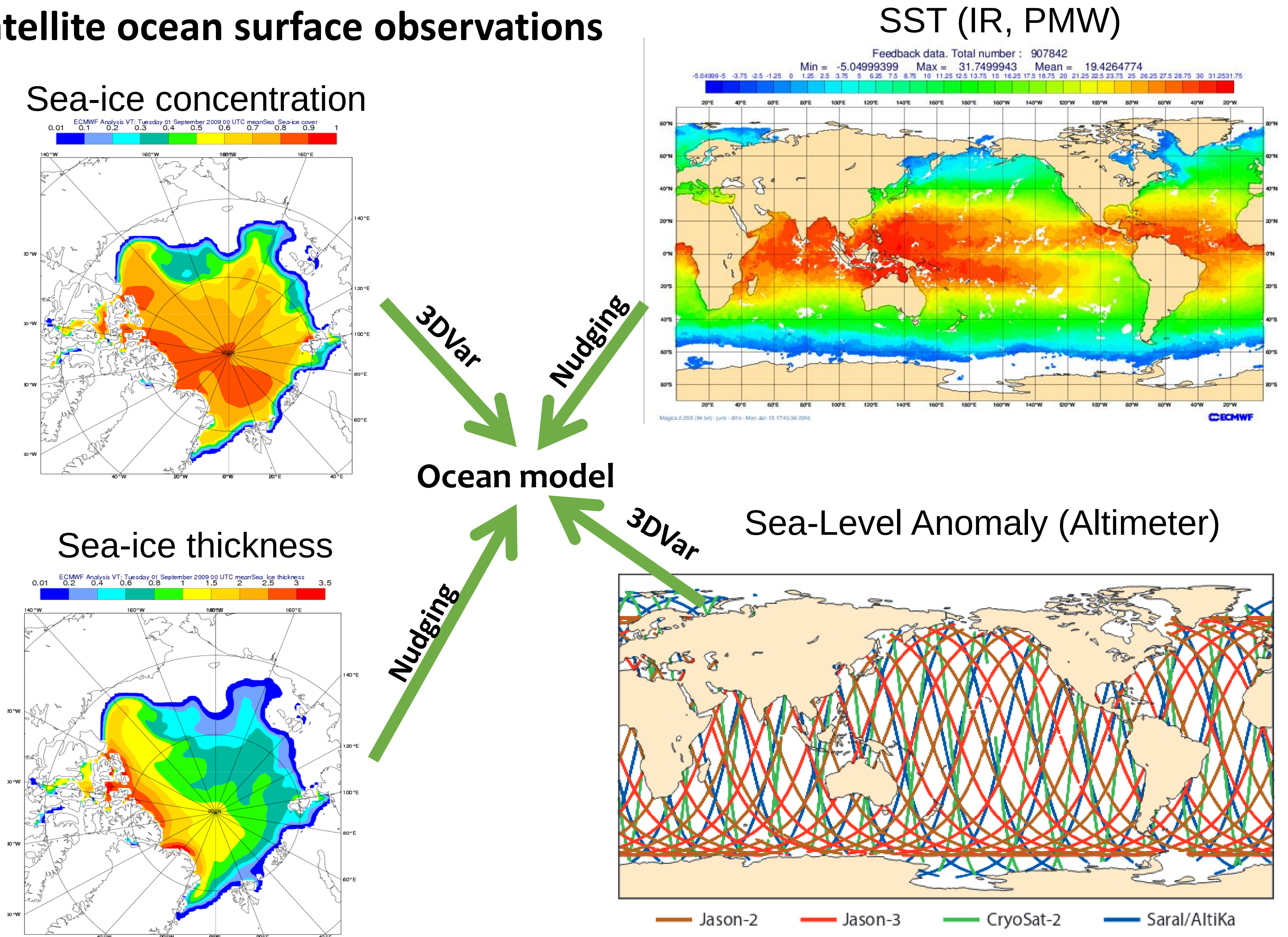
XBT

CDT (CSIRO 2001)



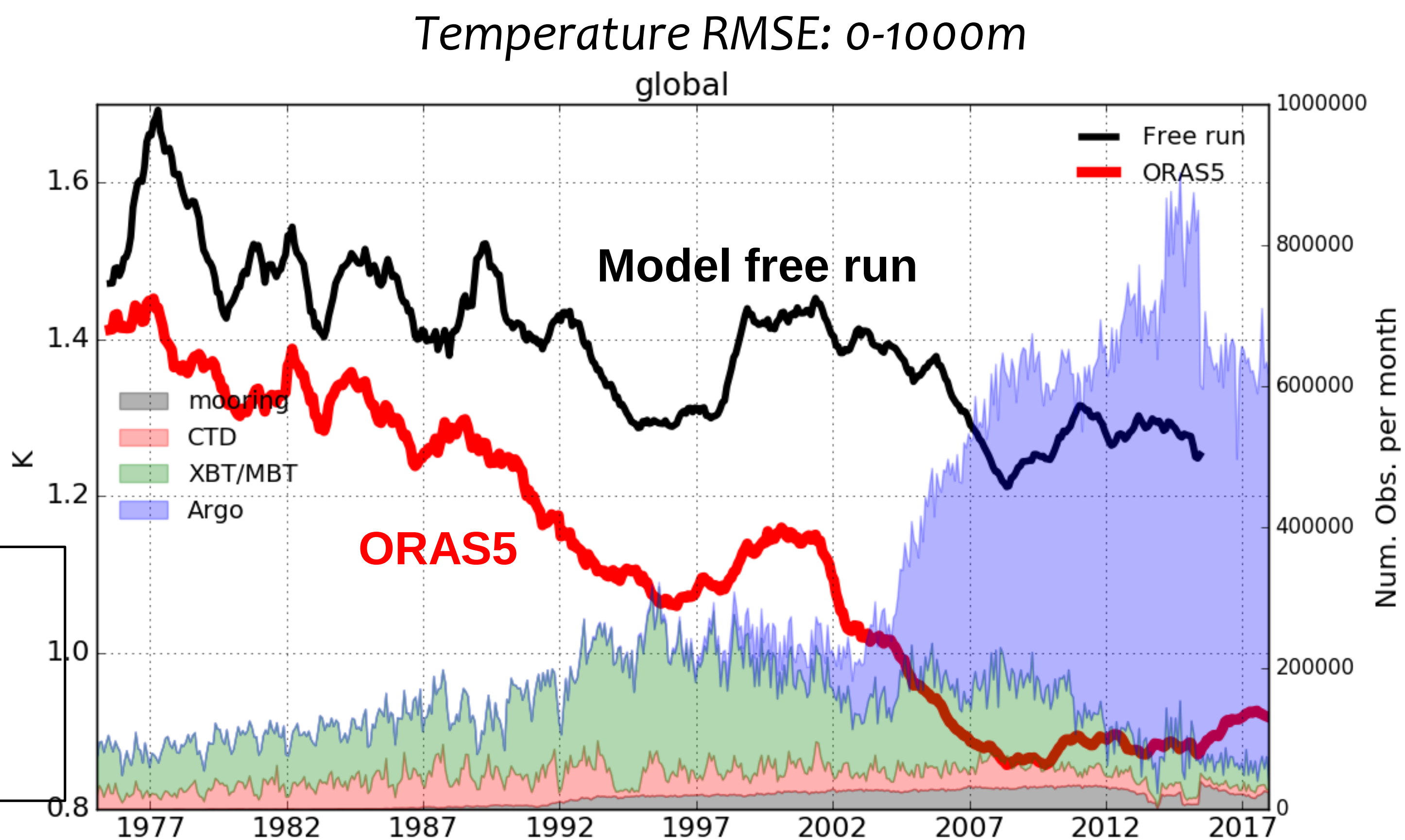
Mammal (MEOP et al., 2015)

Satellite ocean surface observations



Observations impact on the ocean state estimation

Observations are essential for improving initialization



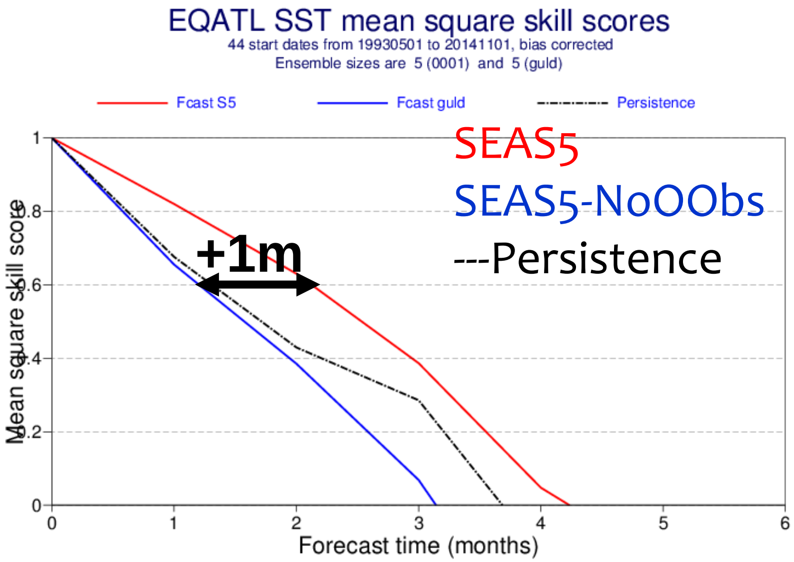
MRB: moored buoy
OSD: CTD sonde
XBT: Expendable bathythermograph
PFL: Argo float

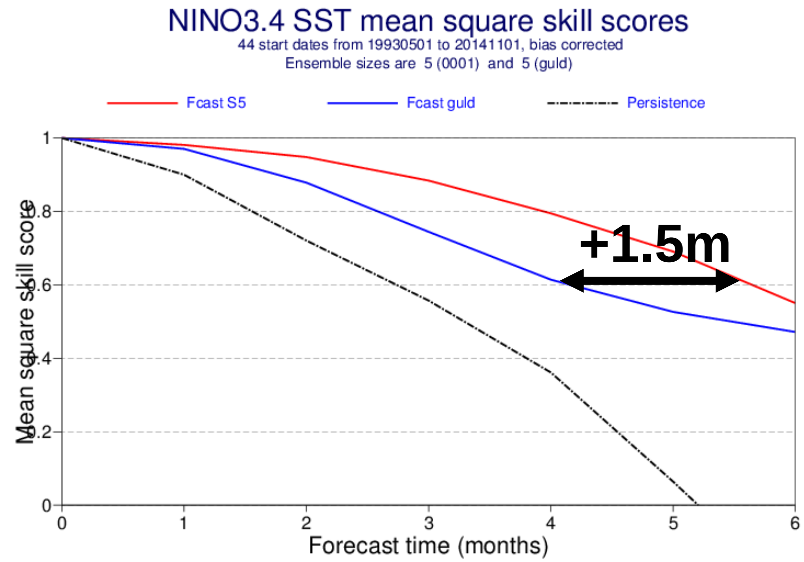
Assimilation of ocean in-situ observations helps to constrain the 3D ocean, therefore providing better estimation of the ocean initial condition for the coupled forecasting system

Impact of ODA in Seasonal Forecast

A proper initialisation played a key role in seasonal forecasts

ENSO forecast





SEAS5 is the new ECMWF seasonal forecasting systems (Johnson et al 2018, GMD)
SEAS5 initialized by Ocean Reanalyses ORAS5 (Zuo et al, 2018)

SEAS5-NoOObs is initialized by an "Ocean Simulation" where Ocean observations have are not assimilated (Only winds and SST)

NEMOVAR

- **The "NEMOVAR" assimilation system used in ECMWF.**
  - Variational system as a collaborative project among **CERFACS**, **ECMWF**, **INRIA** and the **Met Office** for assimilation into the NEMO ocean model.
    - Solves a linearized version of the full non-linear cost function.
    - Incremental **3D-Var FGAT** running operational, 4D-Var in research model
  - Background correlation model based **diffusion operators**
  - Background errors are correlated between different variables through **balance operator**
- To avoid initialization shock increments are typically applied via Incremental Analysis Update (**IAU**) which applies the increments as a forcing term over a period of time.

NEMOVAR: 3D-Var FGAT

Daget et al 2009

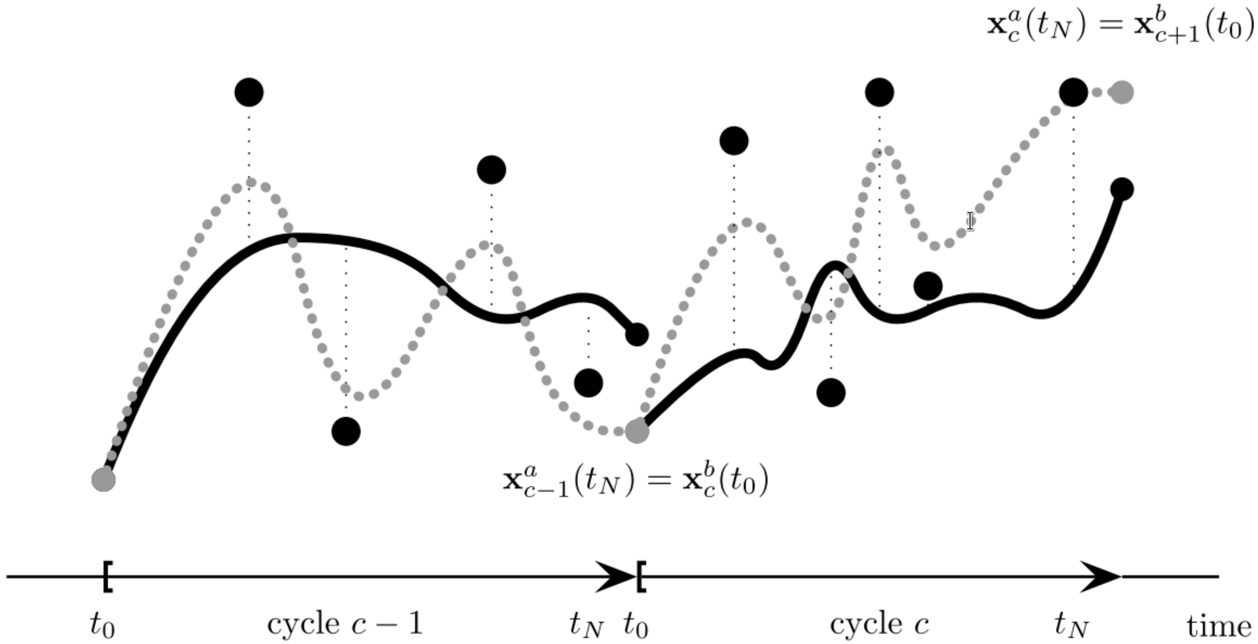


Figure 1: Schematic illustration of the procedure used to cycle 3D-Var. On each cycle c, the model is integrated from t_0 to t_N starting from a background initial condition $\mathbf{x}_c^b(t_0)$ (grey dots) to produce the background trajectory $\mathbf{x}_c^b(t_i)$ (black solid curve). The difference between the observations $\mathbf{y}_{c,i}^o$ (black dots) and their background counterpart ($\mathbf{H}_{c,i}\mathbf{x}_c^b(t_i)$) is computed (represented by the vertical thin dotted lines) for use in the 3D-Var FGAT minimization. After minimization, the model integration is repeated from the same initial condition ($\mathbf{x}_c^b(t_0)$) but with the analysis increment applied using IAU. This produces the analysis trajectory $\mathbf{x}_c^a(t_i)$ (grey dashed curve). The updated model state $\mathbf{x}_c^a(t_N)$ at the end of cycle c is then used as the background initial condition for the next cycle $c+1$ (grey dots).

NEMOVAR: Linearized Cost function

Weaver et al 2003,2005
Daget et al 2009
Mogensen et al 2012
Balmaseda et al 2013

$$J[\delta\mathbf{w}] = \frac{1}{2}\delta\mathbf{w}^{\mathrm{T}}\mathbf{B}^{-1}\delta\mathbf{w} + \frac{1}{2}(\mathbf{G}\delta\mathbf{w}-\mathbf{d})^{\mathrm{T}}\mathbf{R}^{-1}(\mathbf{G}\delta\mathbf{w}-\mathbf{d})$$

$\mathbf{y}^{\mathrm{o}} = \{(\mathbf{y}_0^{\mathrm{o}})^{\mathrm{T}} \cdots (\mathbf{y}_i^{\mathrm{o}})^{\mathrm{T}} \cdots (\mathbf{y}_N^{\mathrm{o}})^{\mathrm{T}}\}^{\mathrm{T}}$ $\longrightarrow$ 4D observation array

$\delta\mathbf{w} = \mathbf{w} - \mathbf{w}^{\mathrm{b}}$ $\longrightarrow$ w is the control vector

$\mathbf{d} = \mathbf{y}^{\mathrm{o}} - G(\mathbf{w}^{\mathrm{b}})$ $\longrightarrow$ Departure vector

$$G(\mathbf{w}) = \begin{pmatrix} \vdots \\ G_i(\mathbf{w}) \\ \vdots \end{pmatrix} = \begin{pmatrix} \vdots \\ H_i[M(t_i, t_0)\{K(\mathbf{w})\}] \\ \vdots \end{pmatrix}$$

Balance operator

$$\mathbf{w}^{\mathrm{b}} \equiv K^{-1}\{\mathbf{x}^{\mathrm{b}}(t_0)\}$$

$$(T, S_{\mathrm{U}}, \eta_{\mathrm{U}}, u_{\mathrm{U}}, v_{\mathrm{U}})^{\mathrm{T}} \longleftarrow (T, S, \eta, u, v)^{\mathrm{T}}$$

linearly independent

Solution:

$$\delta\mathbf{w}^{\mathrm{a}} \approx \mathbf{B}\,\mathbf{G}^{\mathrm{T}}\left(\mathbf{G}\,\mathbf{B}\,\mathbf{G}^{\mathrm{T}} + \mathbf{R}\right)^{-1}\mathbf{d}.$$

$$\delta\mathbf{x}^{\mathrm{a}} = K\left(\mathbf{w}^{\mathrm{b}} + \delta\mathbf{w}^{\mathrm{a}}\right) - K\left(\mathbf{w}^{\mathrm{b}}\right) \approx \mathbf{K}\,\delta\mathbf{w}^{\mathrm{a}}$$

$$\mathbf{x}^{\mathrm{a}}(t_i) = M(t_i, t_{i-1})\left[\mathbf{x}^{\mathrm{a}}(t_{i-1}), F_i\,\delta\mathbf{x}^{\mathrm{a}}\right]$$

IAU,Bloom et al 1996

- **Balance operator**: convert to w space, B becomes block diagonal, representing the spatial covariance model.
- **Diffusion operator**: The spatial covariances is specified by diffusion operator (Weaver and Courtier 2001)

NEMOVAR: Linearized Balance Operator

Define the balance operator symbolically by the sequence of equations

$$
\begin{array}{llccccccccc}
\text{Temperature} & \delta T & = & \delta T & & & & & & \\
\text{Salinity} & \delta S & = & \mathrm{K}^{\mathrm{b}}_{S,T}\,\delta T & + & \delta S_{\mathrm{U}} & = & \delta S_{\mathrm{B}} & + & \delta S_{\mathrm{U}} \\
\text{SSH} & \delta \eta & = & \mathrm{K}_{\eta,\rho}\,\delta \rho & + & \delta \eta_{\mathrm{U}} & = & \delta \eta_{\mathrm{B}} & + & \delta \eta_{\mathrm{U}} \\
\text{u-velocity} & \delta u & = & \mathrm{K}_{u,p}\,\delta p & + & \delta u_{\mathrm{U}} & = & \delta u_{\mathrm{B}} & + & \delta u_{\mathrm{U}} \\
\text{v-velocity} & \delta v & = & \mathrm{K}_{v,p}\,\delta p & + & \delta v_{\mathrm{U}} & = & \delta v_{\mathrm{B}} & + & \delta v_{\mathrm{U}}
\end{array}
\left.\vphantom{\begin{array}{c}1\\1\\1\\1\\1\end{array}}\right\}
$$

Treated as approximately mutually independent without cross correlations

$$
\begin{array}{llccccc}
\text{Density} & \delta \rho & = & \mathrm{K}^{\mathrm{b}}_{\rho,T}\,\delta T & + & \mathrm{K}^{\mathrm{b}}_{\rho,S}\,\delta S \\
\text{Pressure} & \delta p & = & \mathrm{K}_{p,\rho}\,\delta \rho & + & \mathrm{K}_{p,\eta}\,\delta \eta
\end{array}
\left.\vphantom{\begin{array}{c}1\\1\end{array}}\right\}.
$$

Weaver et al., 2005, QJRMS

NEMOVAR: B matrix

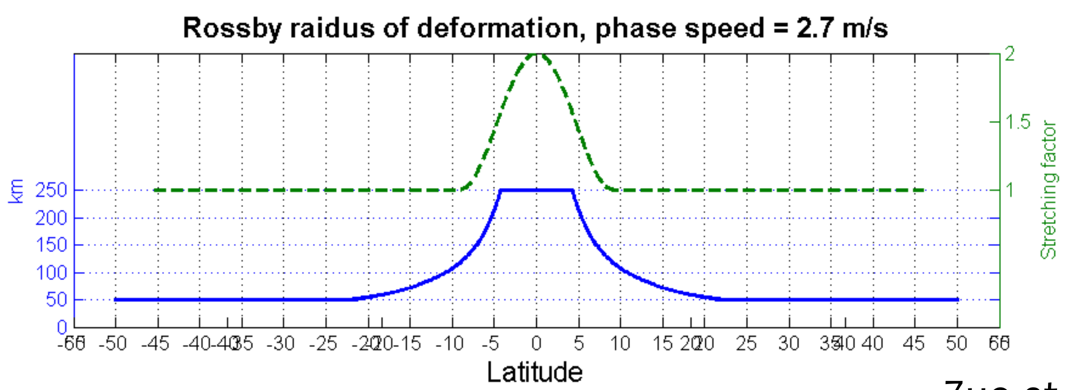
General B formulation in NEMOVAR

$$\mathbf{B} = \alpha\,\mathbf{B}_{\mathrm{m}} + \beta\,\mathbf{B}_{\mathrm{e}}$$

$$\mathbf{B}_{\mathrm{m}} = \mathbf{K}_{\mathrm{b}}\,\mathbf{D}_{\mathrm{m}}^{1/2}\,\mathbf{C}_{\mathrm{m}}\,\mathbf{D}_{\mathrm{m}}^{1/2}\,\mathbf{K}_{\mathrm{b}}^{\mathrm{T}}$$

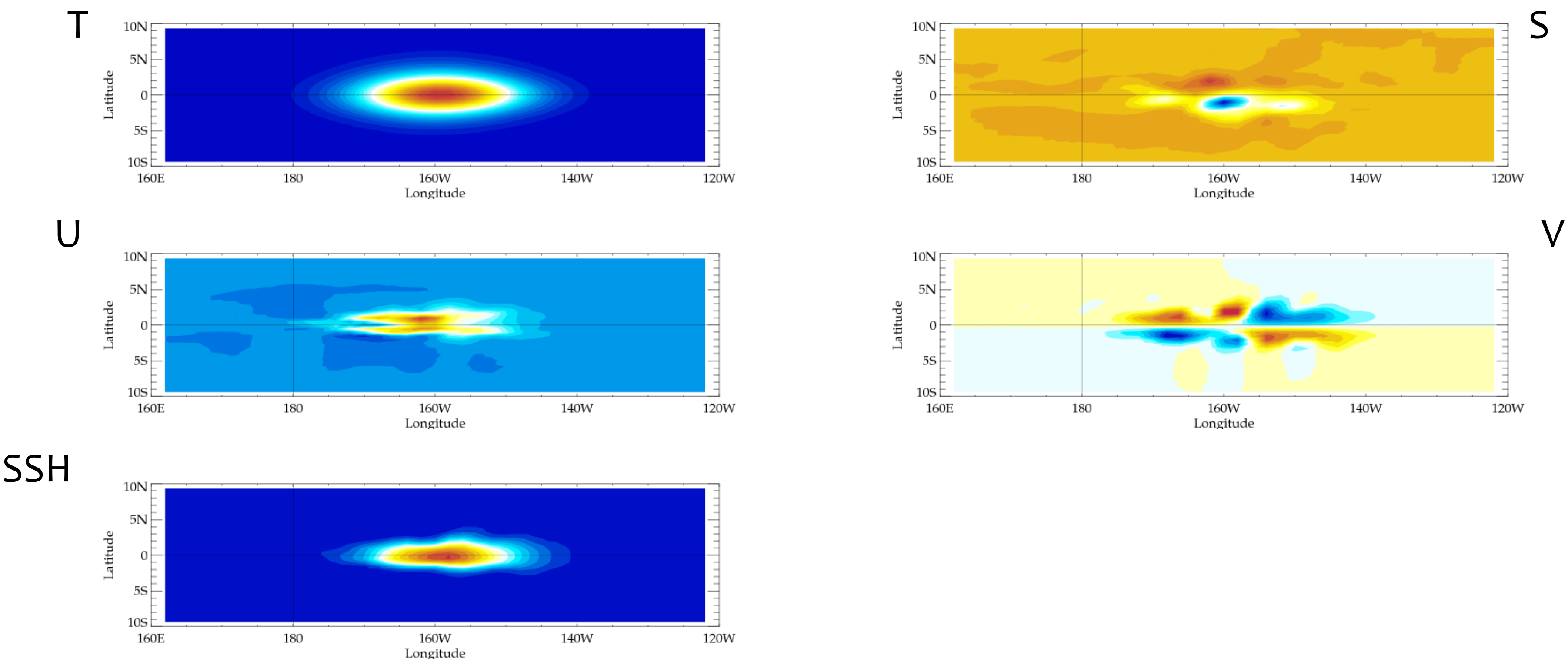
B_m is covariance model for each variable, **C**_m is correlation matrix (including diffusion operator), and **D**_m is a diagonal matrix of variance

The background error correlation length-scales



Zuo et al., 2015

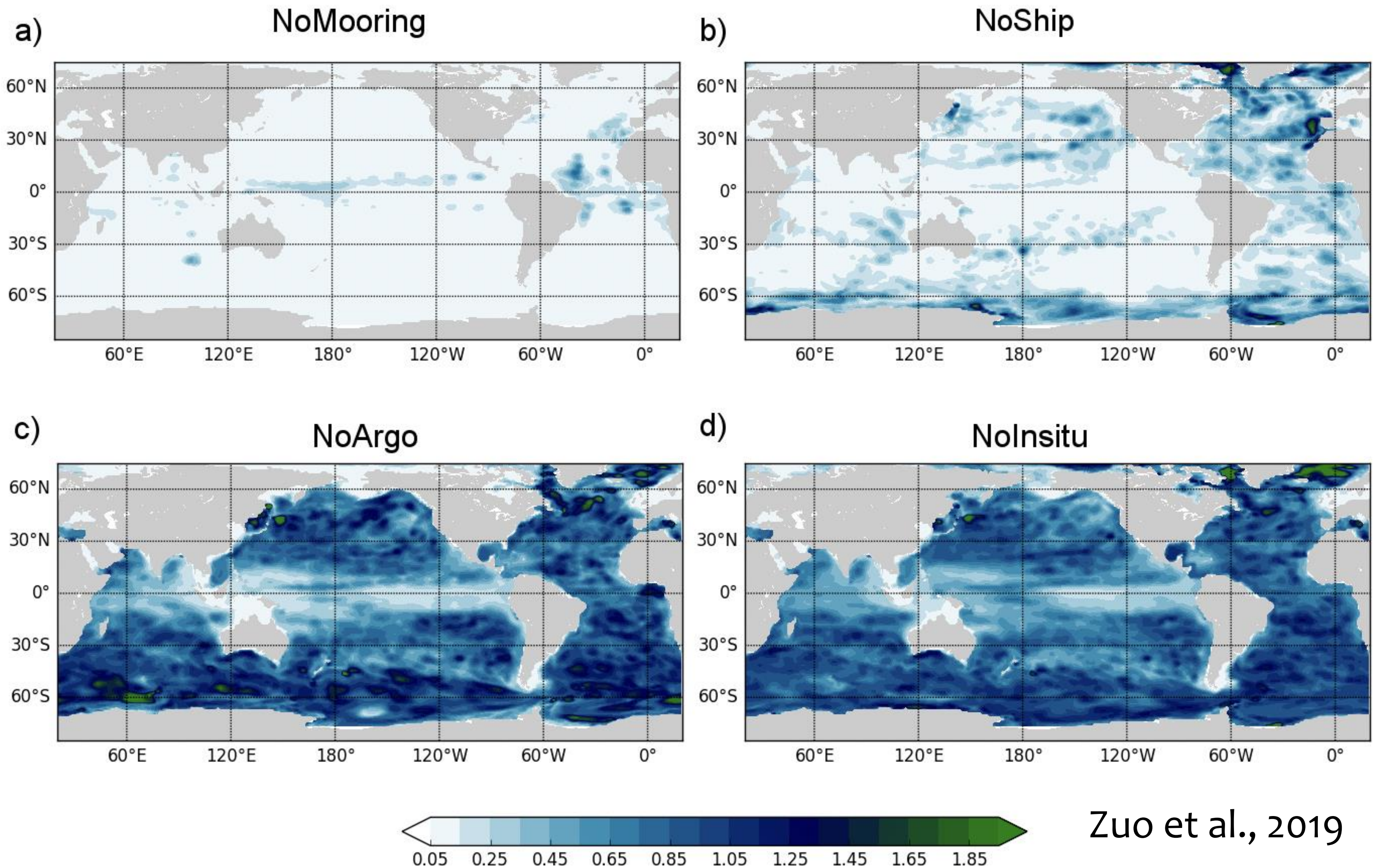
Horizontal cross-correlation of T at 100m



- From single observation of temperature experiment.
- The specific background determines the shape due to the balance relations.
- S, U, V, SSH increments are from balance with T only.

Assimilation of in-situ observations

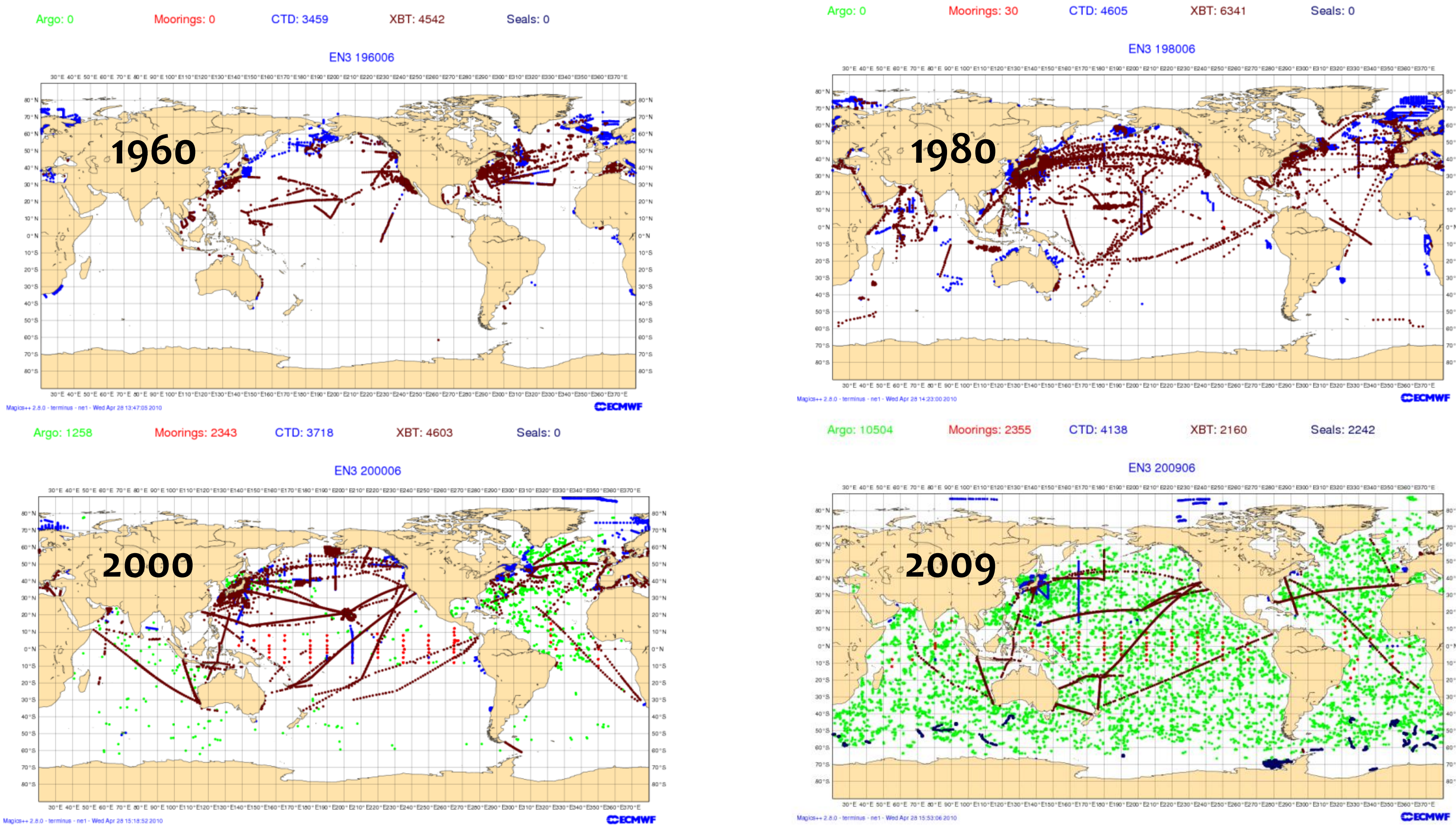
Maps of normalized RMSD of Temperature in OSEs



Zuo et al., 2019

Evolving ocean observing network

Changes in the observing system can be damaging for the representation of the inter-annual variability, leading to spurious climate signals



Bias Correction Algorithm

- To correct systematic errors in models/forcing
- To mitigate changes in the observing system

$$\mathbf{b}^{f}_{k} = \overline{\mathbf{b}}_{k} + \mathbf{b}'^{f}_{k}$$

Seasonal term, estimated offline from rich-data (Argo) Period

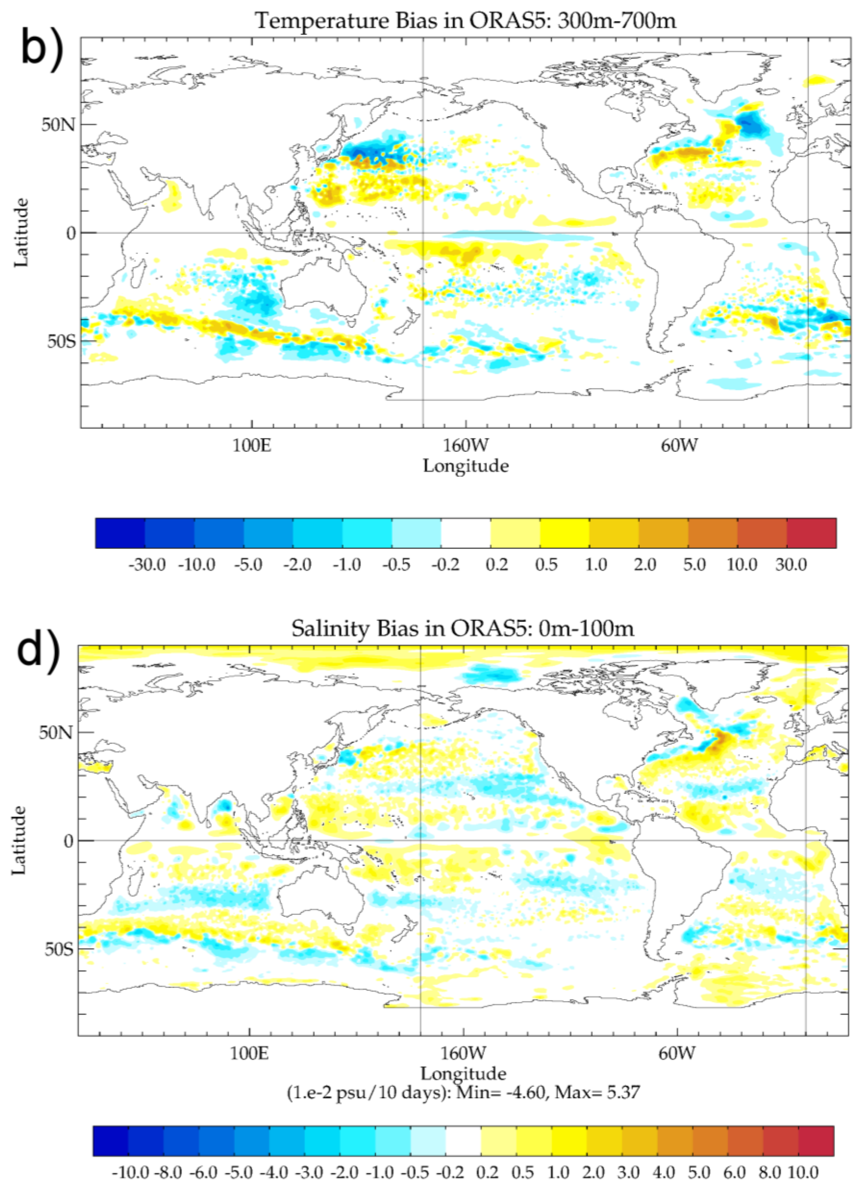
Slow varying term, estimated online from assimilation increments $\mathbf{d}_k$

$$\mathbf{b}'_{k} = \alpha \mathbf{b}'_{k-1} + \mathbf{A}(y)\beta \mathbf{d}_{k}$$

A(y**)**: Partition of bias into T/S and pressure gradient.

(Balmaseda et al 2007)

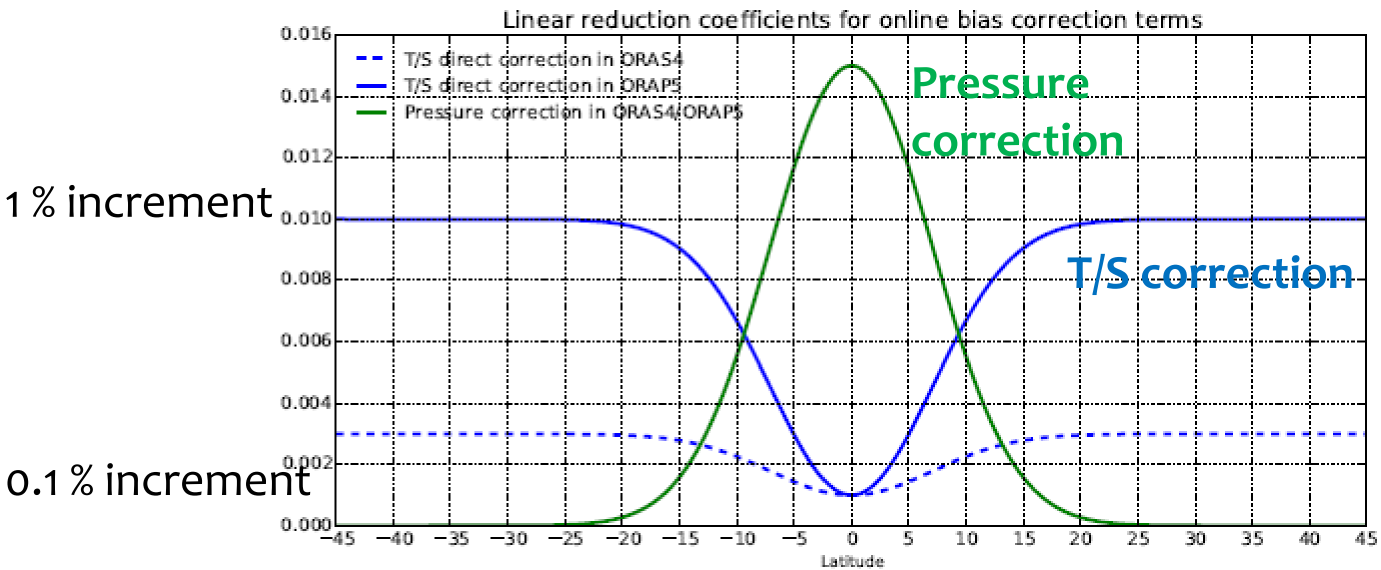
Offline bias term in ORAS5



(Zuo et al 2018)

Online bias correction term

The latitude dependent partition coefficients determine the proportion of online bias corrections applied directly on T/S, and on pressure term. These values ensure that at low latitude the dominant bias term is pressure correction (green solid line).

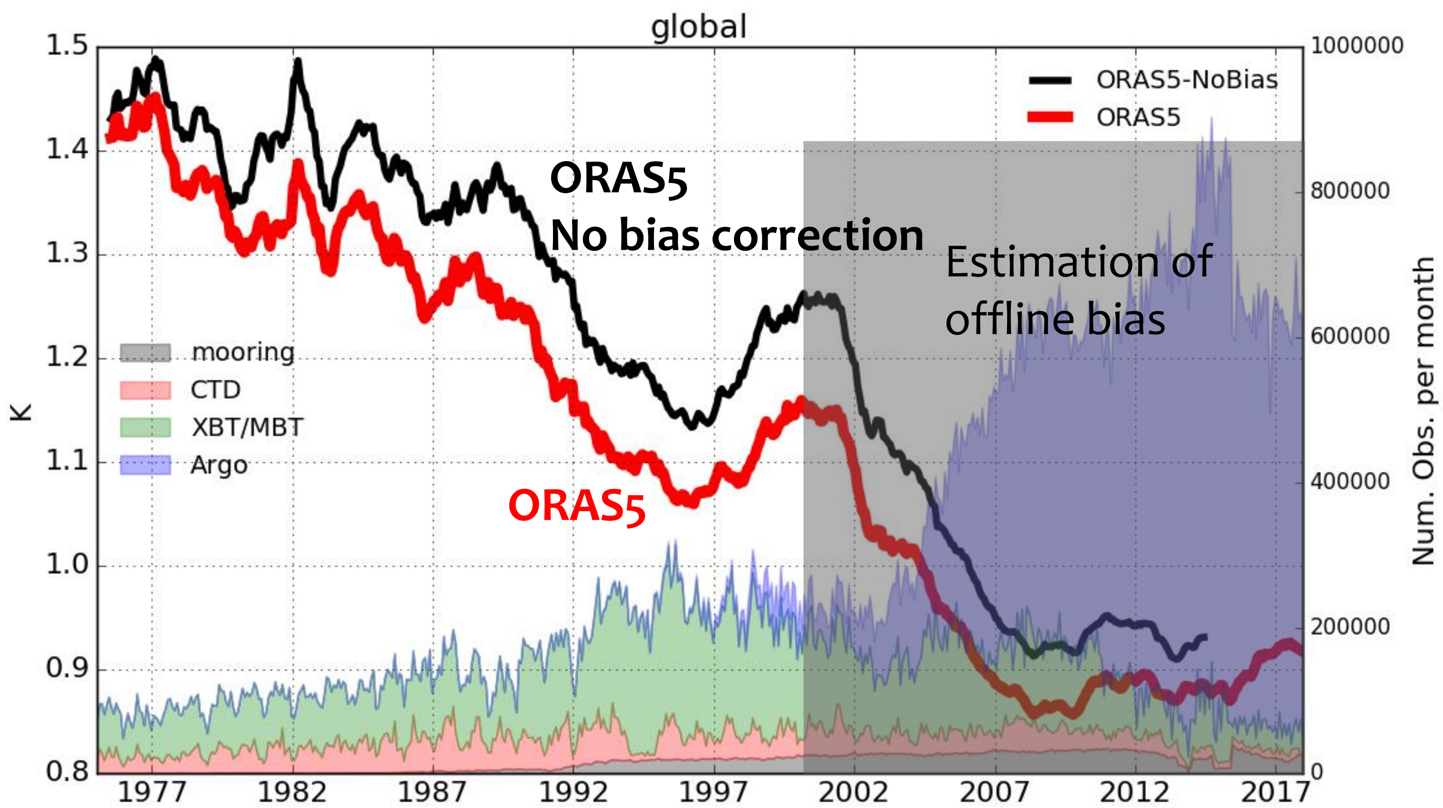


Zuo et al., 2015

Figure 3: Latitude-dependent linear reduction coefficients as applied on online bias correction terms in equations 6 and 7: blue line - $a^{tr,T/S}$, reduction coefficients that apply to direct temperature and salinity corrections (different for ORAS4 and ORAP5); and green line - $a^{p,T/S}$, reduction coefficients that apply to pressure bias correction.

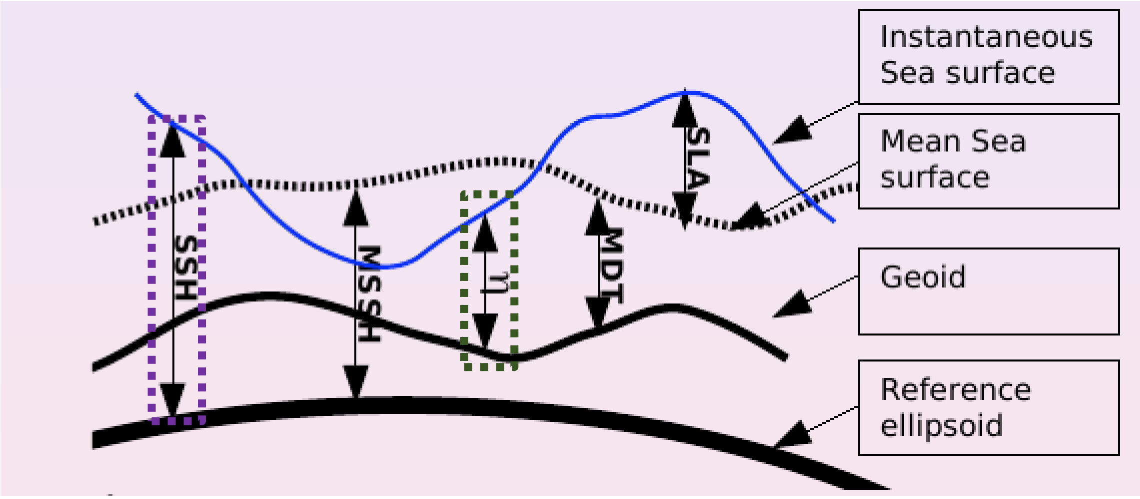
Effect of bias correction on ODA

Temperature RMSE: 0-1000m



- ➢ Bias correction in ODA is essential, and in particular important for mitigating spurious signals introduced due to changes in the observing system

Assimilation of SLA



Altimeter measures SSH (respect reference ellipsoide)
Model represents η (ssh referred to the Geoid)

SSH-Geoid= η

Geoid was poorly known (not any longer) and changes in time (*)

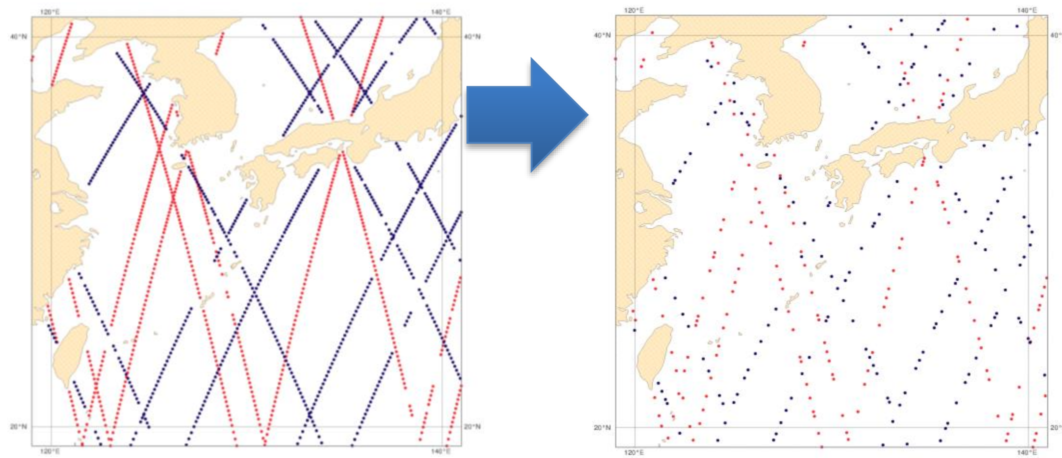
Alternative: Assimilate Sea Level Anomalies (SLA) respect a time mean

Obs: SSH anomalies = SSH-MSSH = Obs SLA
Mod: η anomalies = η – MDT = Mod SLA

Where: MSSH= Temporal Mean SSH ;
MDT = Temporal Mean of model SL Mean Dynamic Topography
MSSH – Geoid = MDT

Assimilation of SLA

- **The SLA along track data has very high spatial (9-14km) resolution for the operational ocean assimilation systems.**
  - Features in the data which the model can not represent
  - "Overfitting" to SLA obs
- **This can be dealt with in different ways:**
  - Inflate the observation error
  - Construction of "superobs" or thinning



ERS-2: 13005 Envisat: 0 Envisat N: 0

Jason-1 N: 0 Jason-2: 0 T/P: 13219

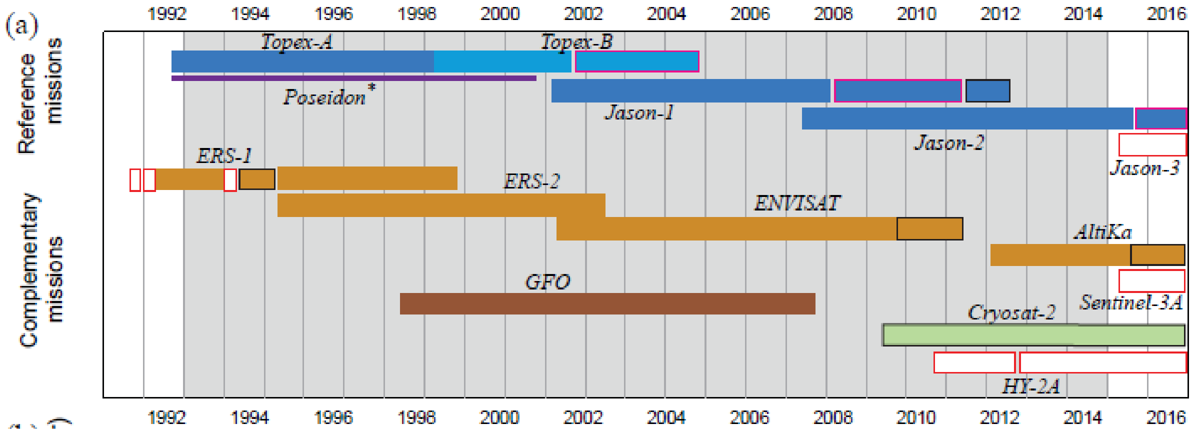
Impact of SLA assimilation

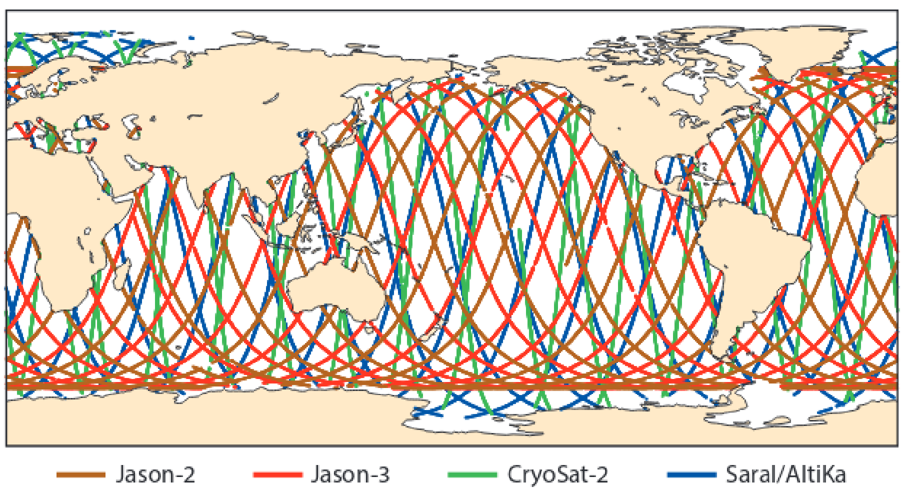
CryoSat2
Jason-3
Jason-2
AltiKa

Jason-3 Sentinel-3A

25 June 2016

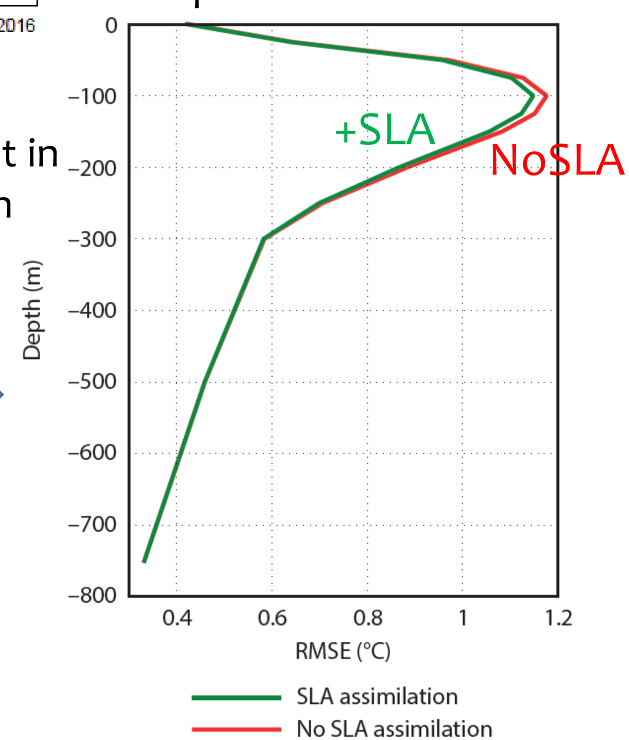
24 Dec 2016





Improvement in ORAS5 ocean initial conditions

Temperature RMS errors

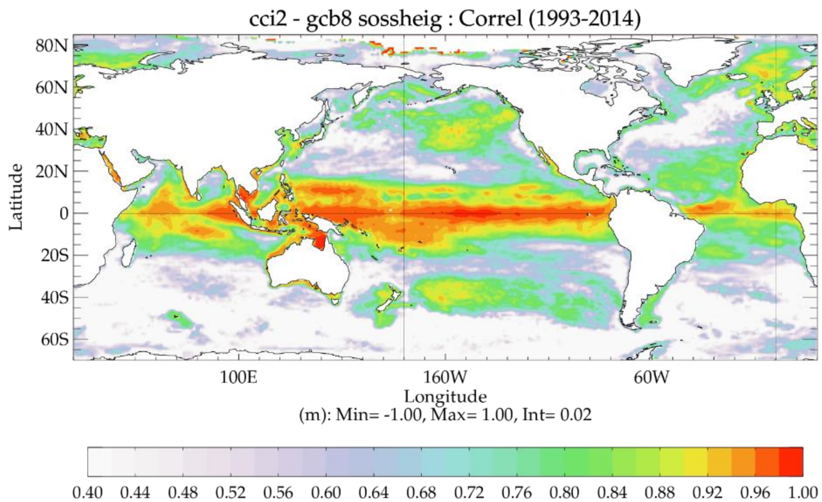


Impact of SLA assimilation

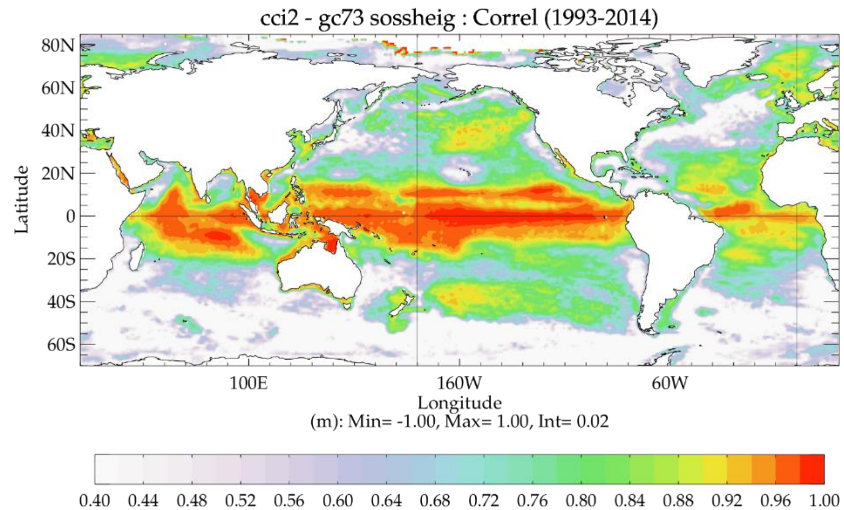
Temporal correlation (monthly) between ORA and ESA SL CCI data

Zuo et al., 2018

ORAS5-NoAlti



ORAS5



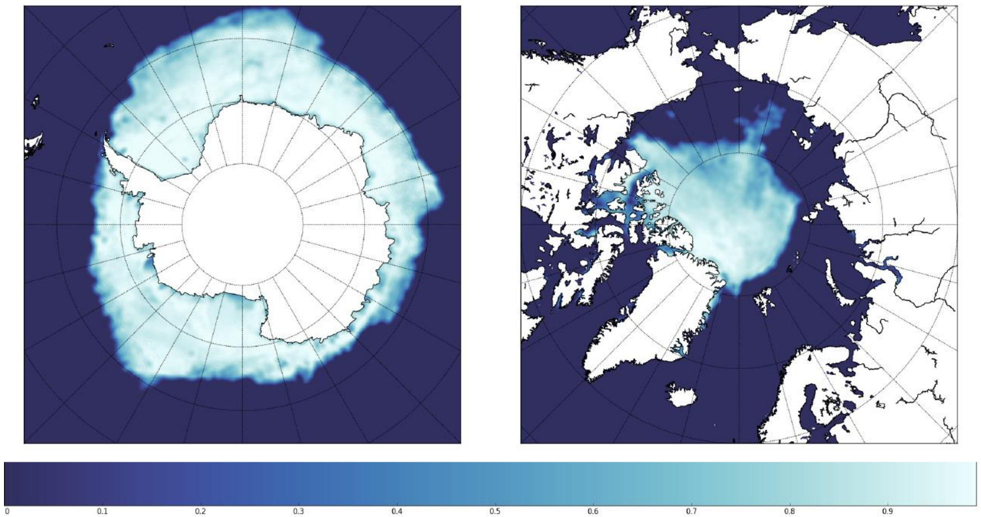
Assimilating altimeter has a large impact on the spatial distribution and magnitude of Sea Level variations in Ocean Reanalyses. This impacts the ocean circulation and the seasonal/decadal forecasts

Assimilation of SIC

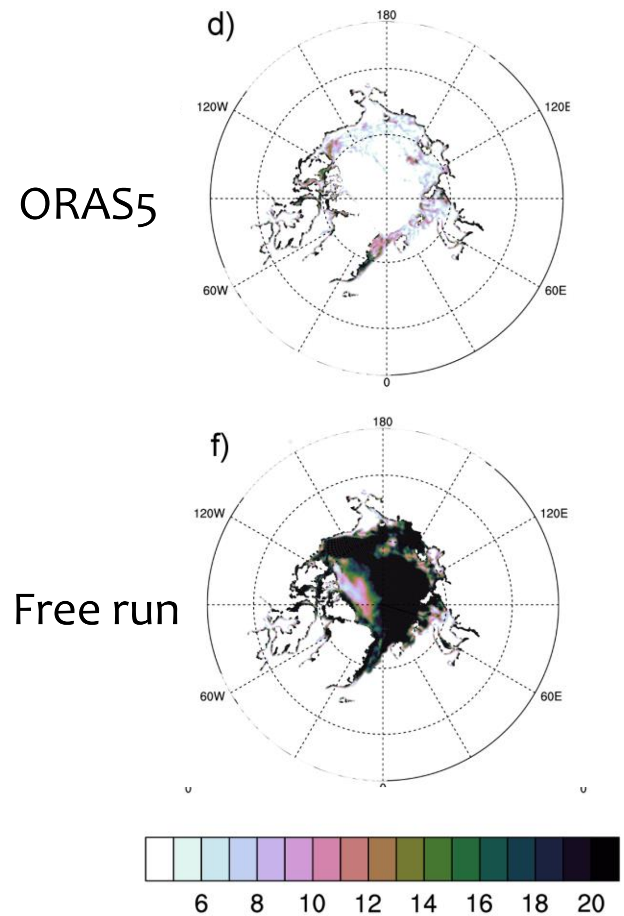
Sea-Ice Concentration data from OSTIA L4 analysis is assimilated

- Via NEMOVAR 3DVar-FGAT
- Treated as univariate
- Assimilated through outer-loop coupling
- Positive impact on both SIC and SIT

Latest L4 sea ice concentration observations from OSTIA (20180912)



SIC RMSE in Sep (1993-2008)



Zuo et al., 2019

Assimilation of SST: nudging

SST nudging scheme

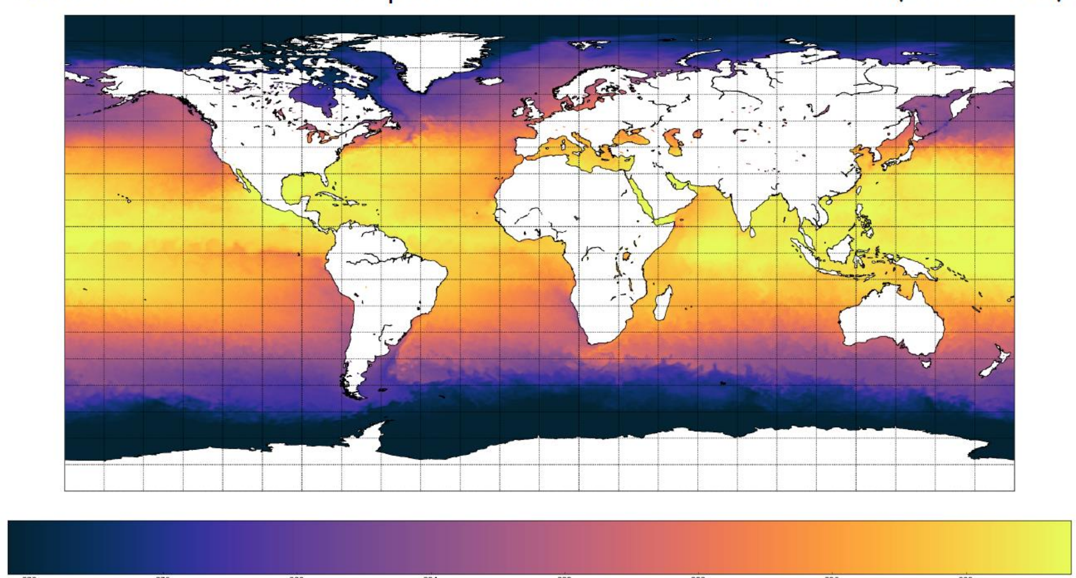
Haney 1917

$$Q_{ns} = Q_{ns}^{o} + \frac{dQ}{dT}(SST_{MODEL} - SST_{TARGET})$$

<u>non-solar total heat flux</u> <u>Fixed negative feedback coefficient</u>

Nudging to OSTIA L4 objective analysis data

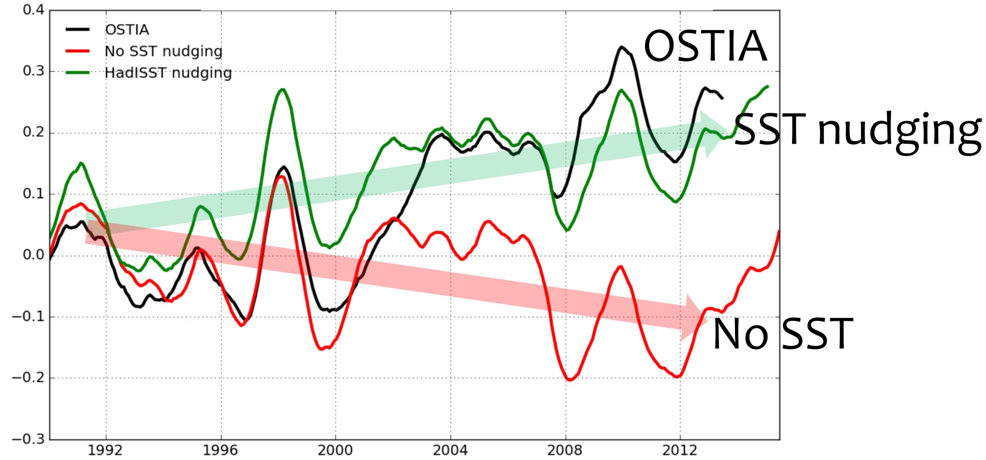
Latest L4 sea-surface temperature observations from OSTIA (20180912)



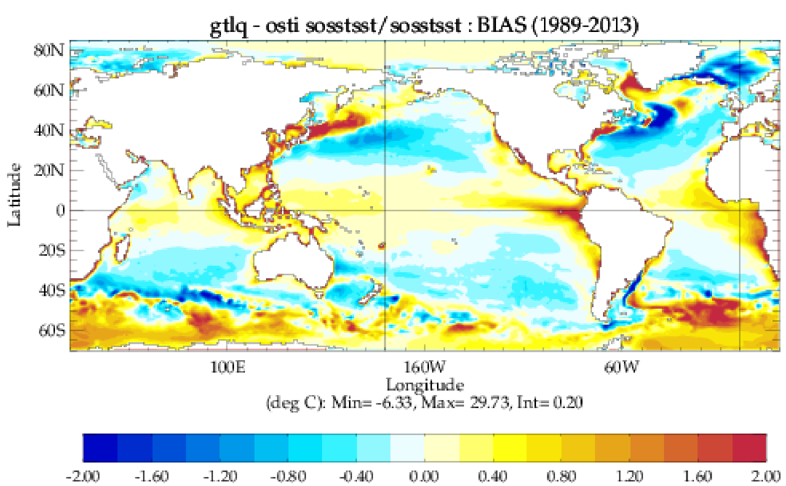
Impact of SST nudging

- Overall very effective
- Not accounting complicated error characteristics in the L4 SST analysis
- Not accounting vertical correlation when apply SST constrain in the surface

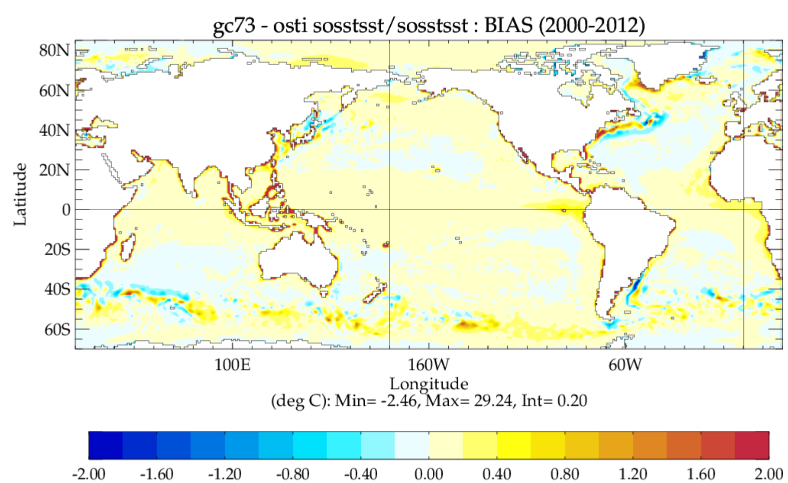
Global mean SST



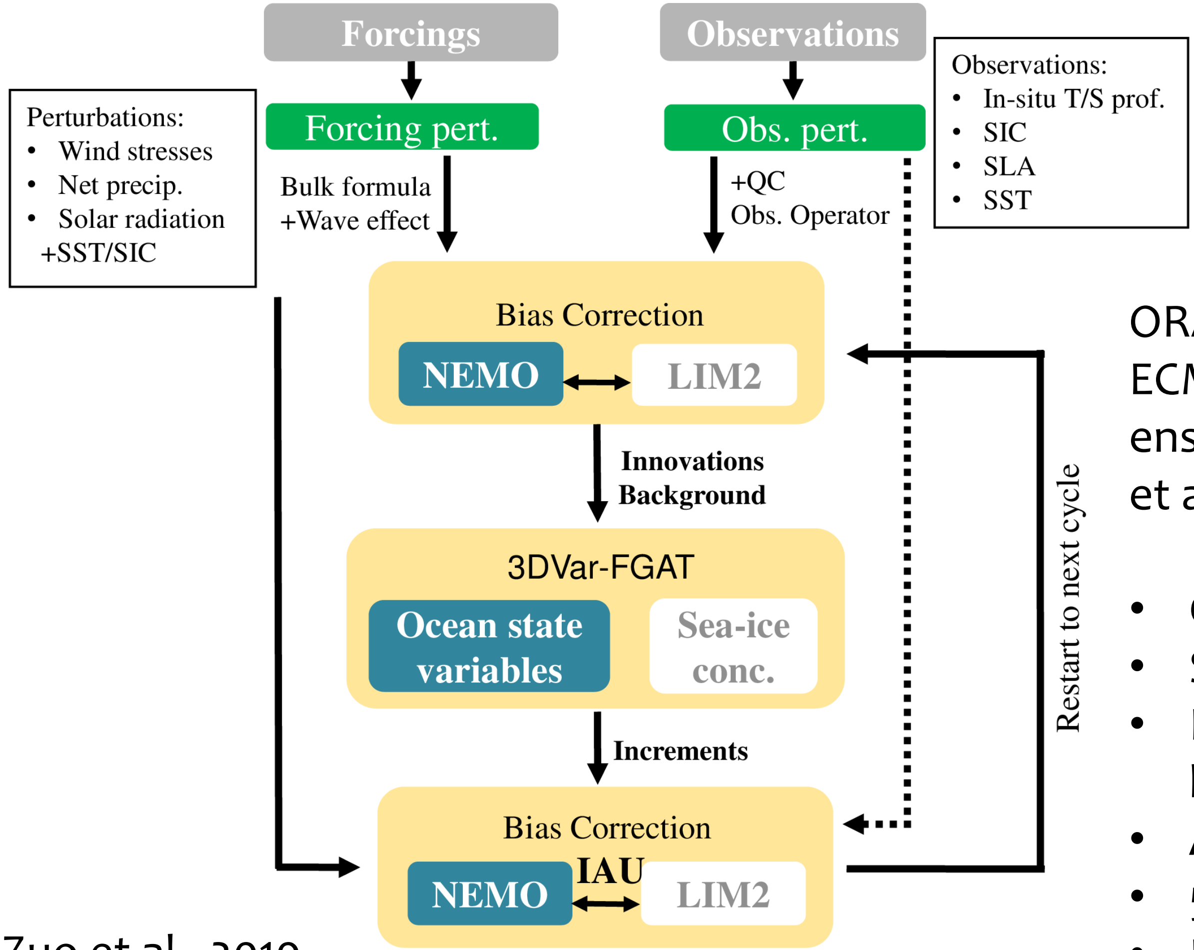
SST: CNTL - OSTIA



SST: ORAS5 - OSTIA



ECMWF Ocean Reanalysis-analysis system



Zuo et al., 2019

Overview of the ORAS5 setup

ORAS5 is the 5th generation of ECMWF ocean and sea-ice ensemble reanalysis system (Zuo et al., 2018, 2019).

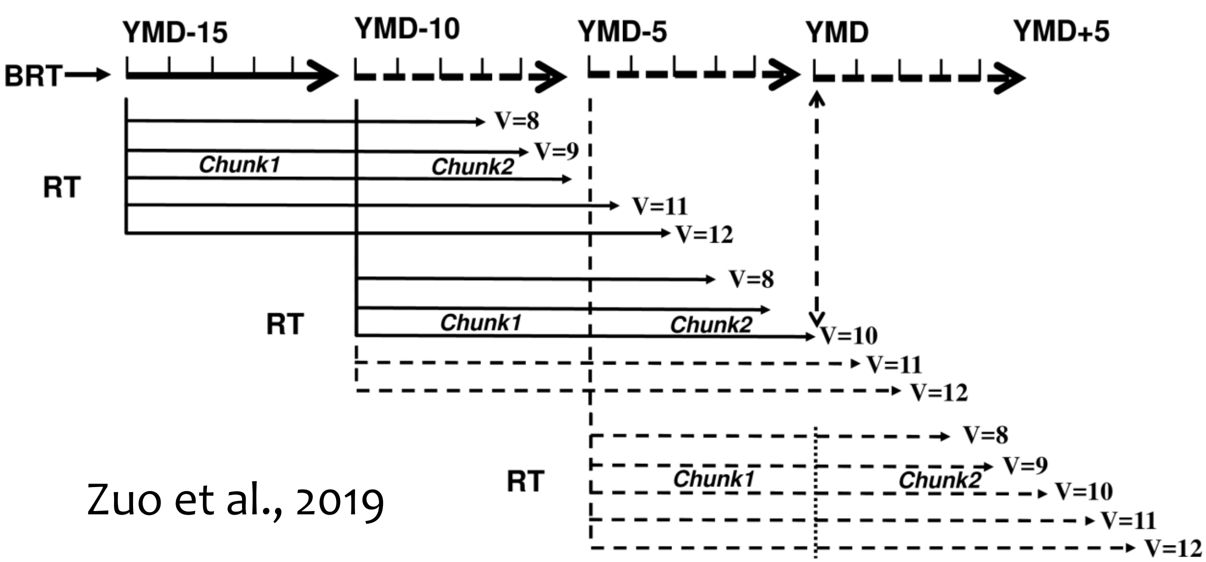
- Ocean: NEMOv3.4
- Sea-ice: LIM2
- Resolution: ¼ degree with 75 levels
- Assimilation: 3DVAR-FGAT
- 5 ensemble member
- Forcing: ERA-int

Real-Time ocean analysis

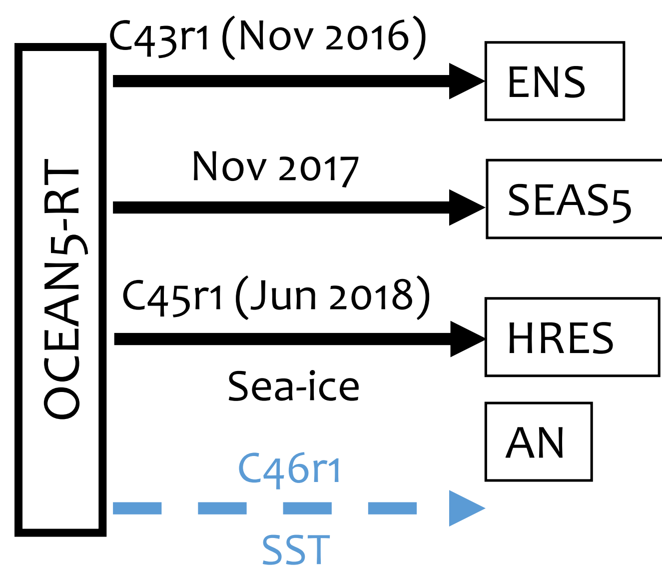
ECMWF OCEAN5 comprises BRT and RT stream, and provide initial conditions for the ocean and sea-ice components of all ECMWF coupled forecasts (ENS, SEAS5, HRES).

Daily Initial condition valid at 00 UTC for the next day

Schematic plot of OCEAN5 BRT and RT components



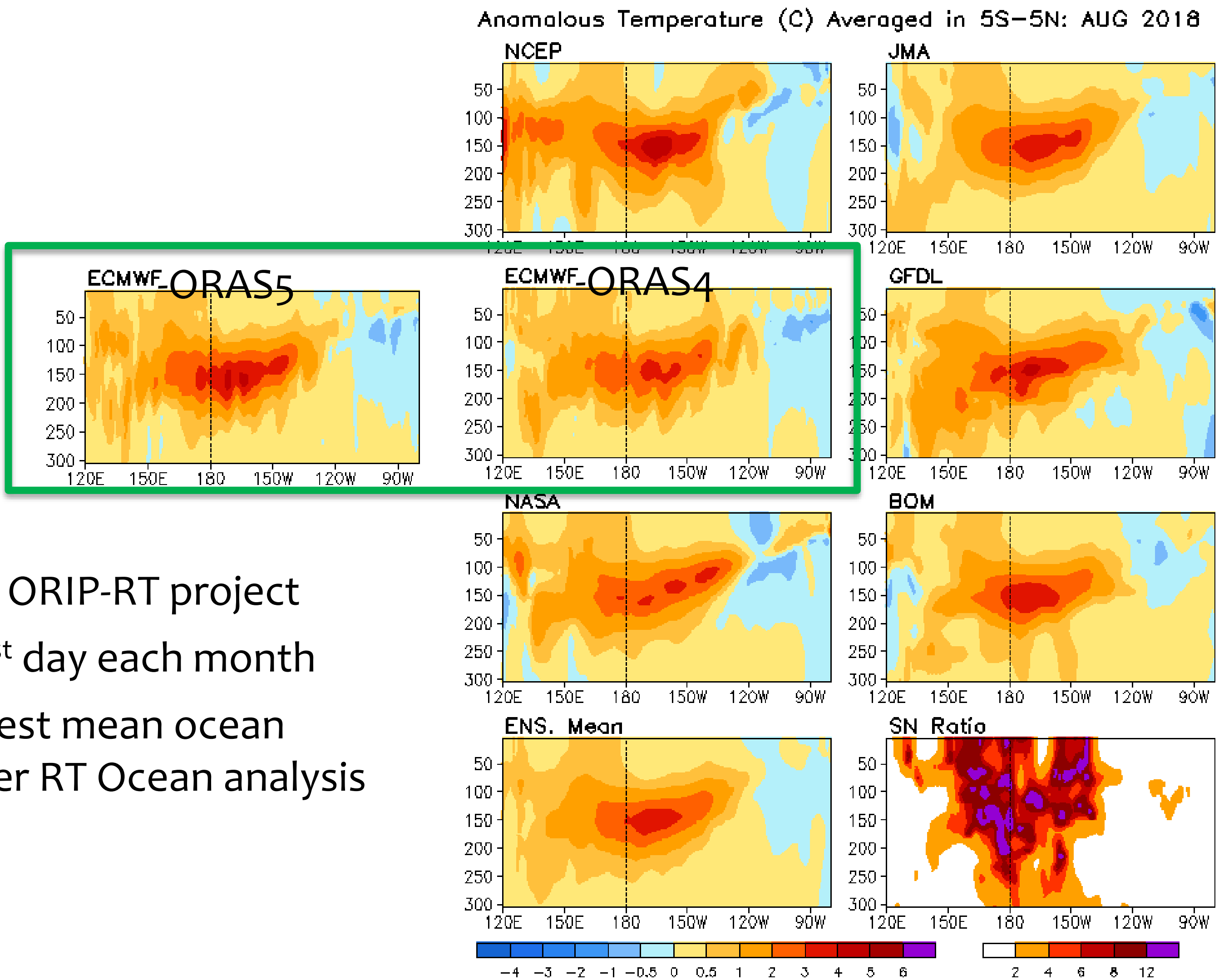
Zuo et al., 2019



Application of RT ocean analysis

Real-Time monitoring of ENSO state

Ref: 1981-2010



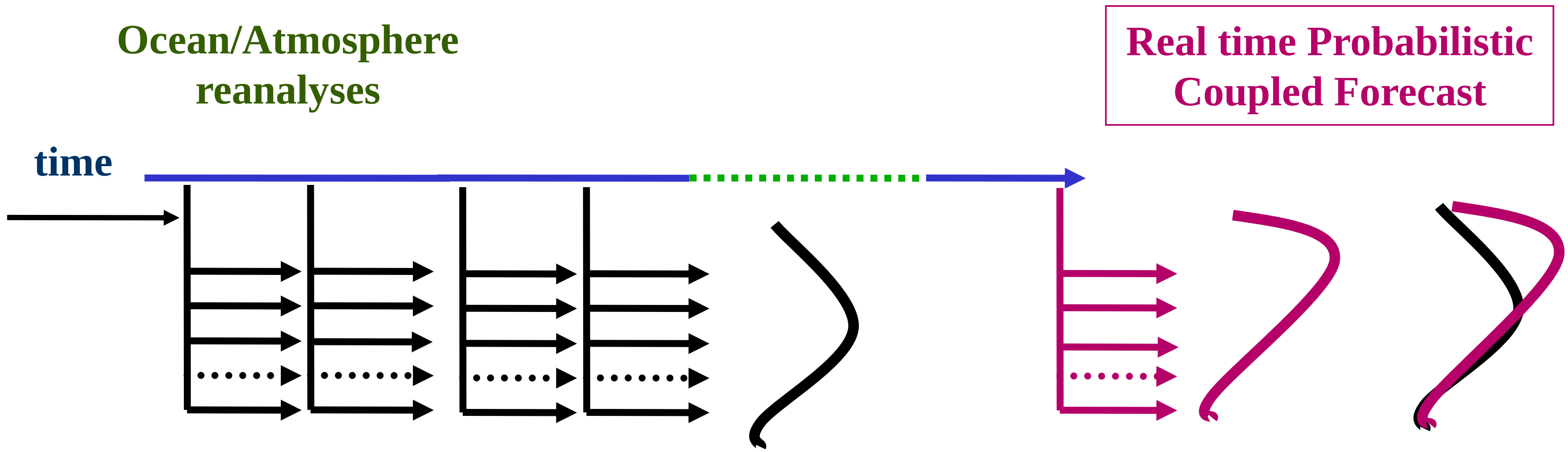
Contribution to the ORIP-RT project

- Update on the 1st day each month
- Compare the latest mean ocean state with 8 other RT Ocean analysis products

Application of Ocean reanalysis

Calibration and reforecasts

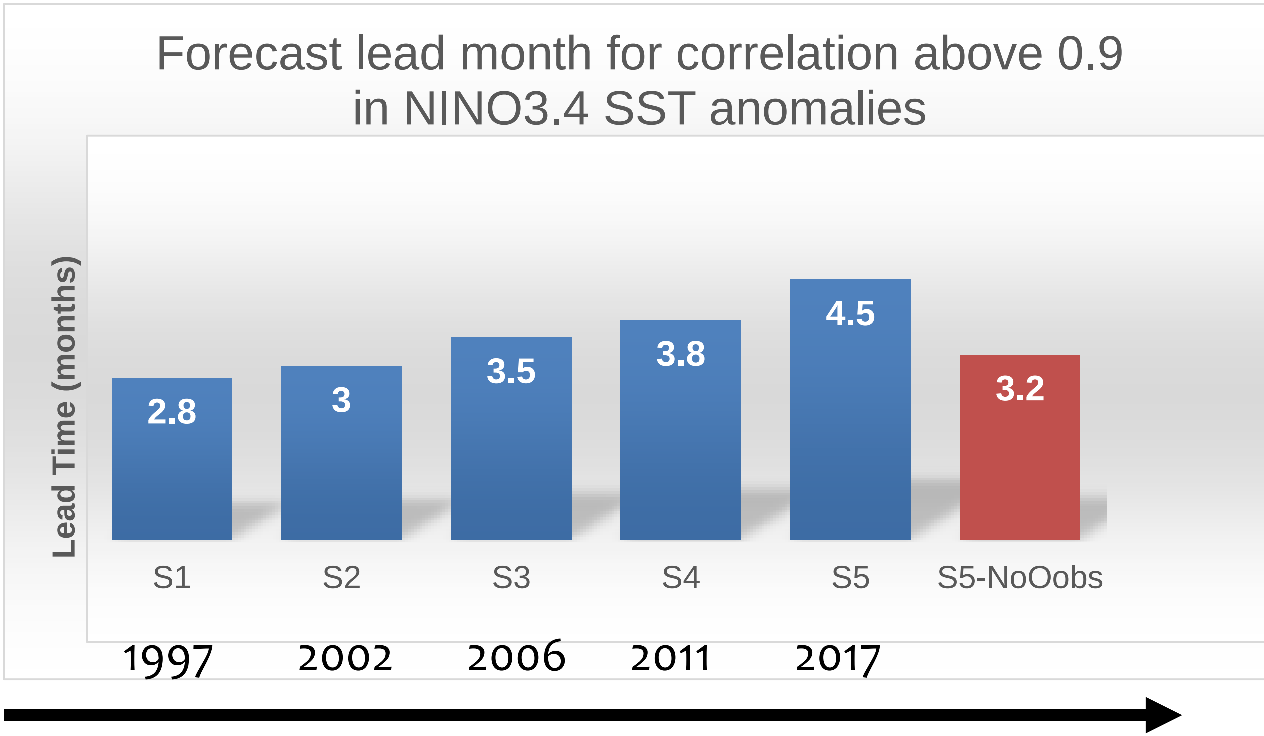
- Correcting model error
- Extreme Events
- Tailored products (health, energy, agriculture)



Hindcasts, needed to estimate climatological PDF, require a historical ocean and atmospheric reanalyses

Consistency between historical and real-time initial conditions is required.

Application of Ocean DA

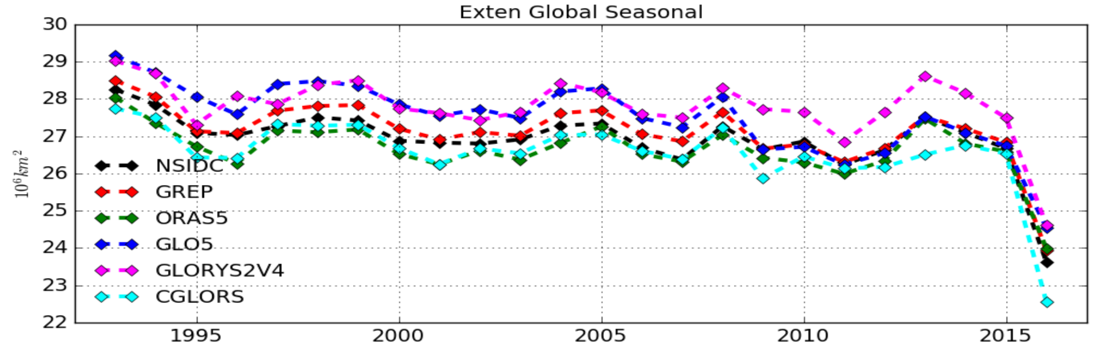


- Gain about 2 months in ENSO prediction
- Without Ocean observation and DA, we would lose about 15 years of progress.

Application of Ocean reanalysis

Climate monitoring and extreme event

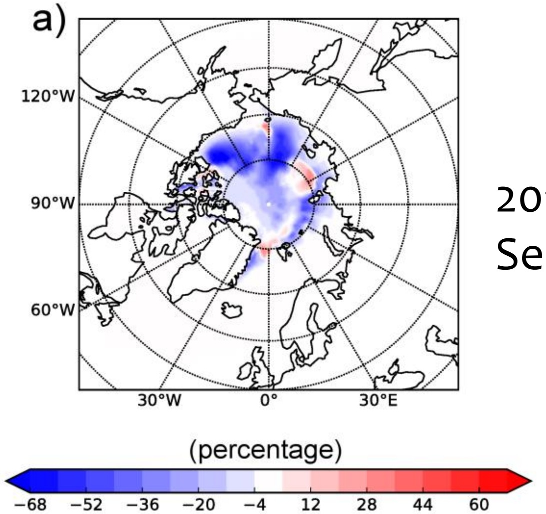
November 2016 has seen the lowest global sea ice cover among 38-year satellite record of ice extent for the month



ORAS5 sea-ice extent anomalies

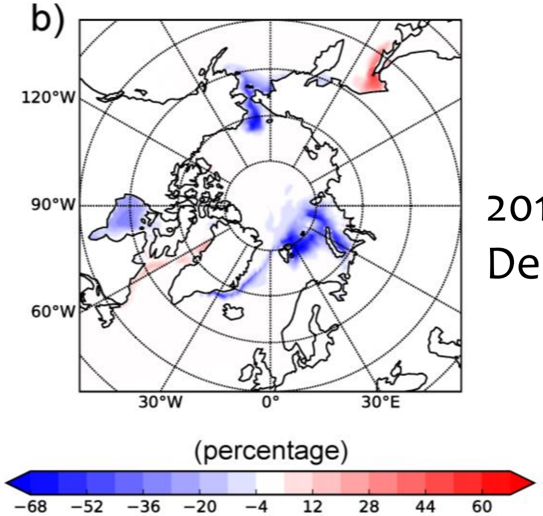
Zuo & Lien, 2018

a)



2016 September

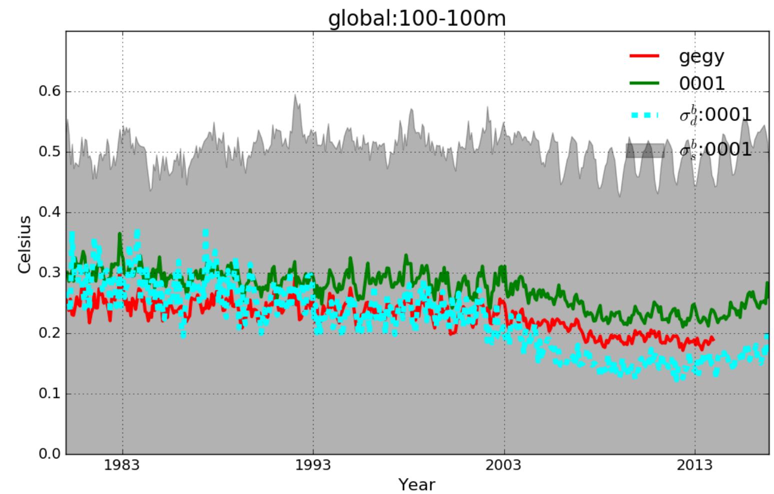
b)



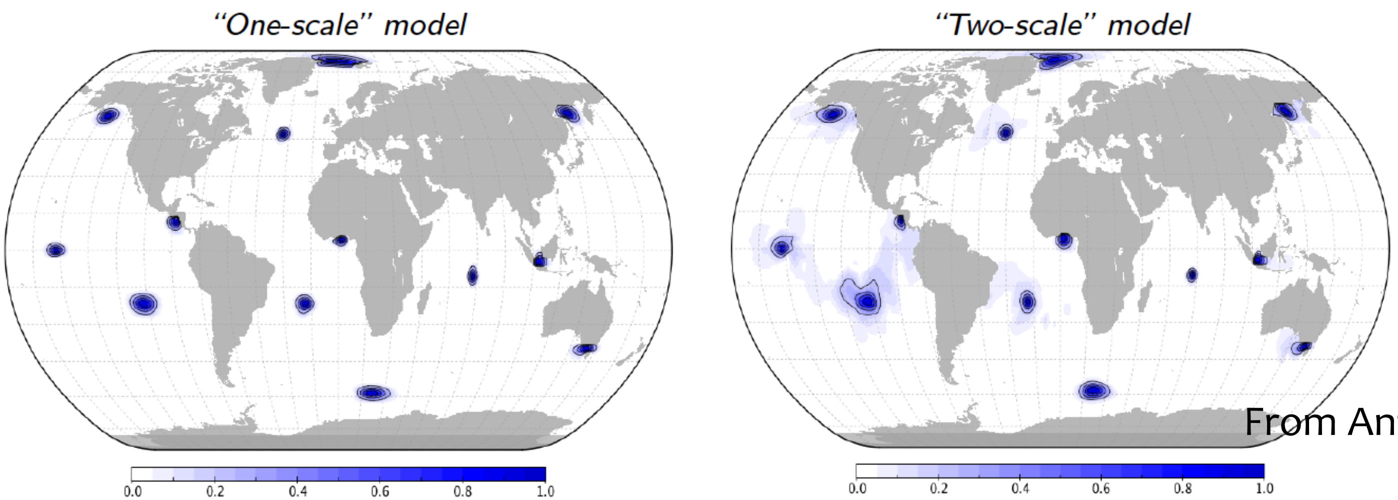
2016 December

Future development for ODA

- Flow-dependent B
- Hyrbid B (B_m + B_e)
- Multi-grid capability
- Multiple spatial scales
- Enhanced perturbation scheme
- Improved SST and SLA assimilation
- Improved SIC assimilation



Diffusion-modelled SSH correlations with parameters estimated from a 20-member ensemble.

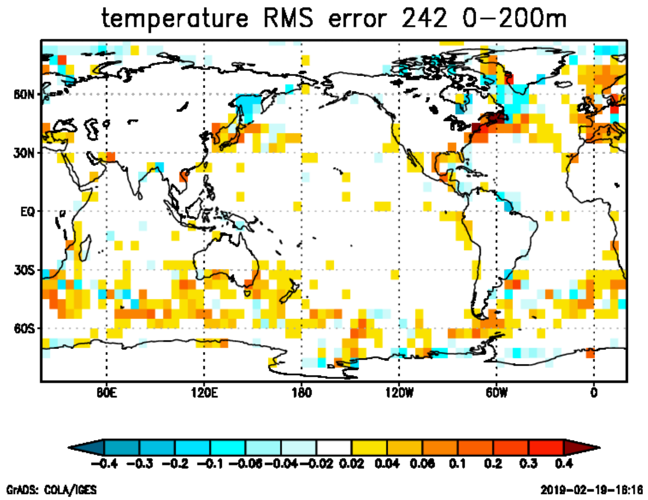


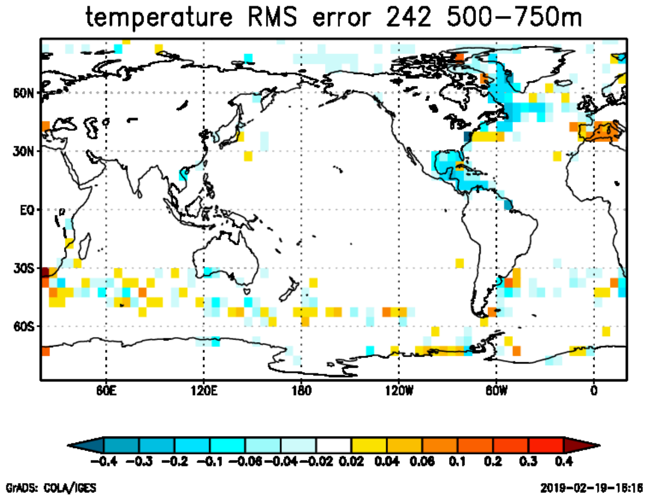
From Anthony weaver

Towards SST assimilation

Twin experiments: SST nudging weak -80 VS strong -200 (in W/m^2)

T RMSE difference: weak-strong nudging





Ref obs: EN4 in-situ

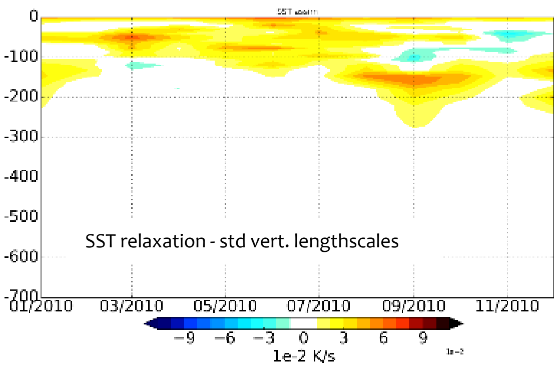
(blue = good, red = bad)

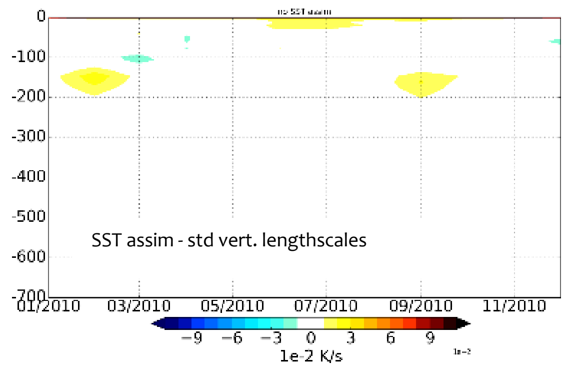
Assimilating L2P SST

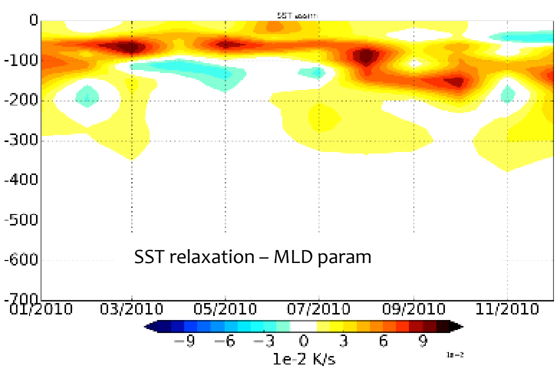
T increments

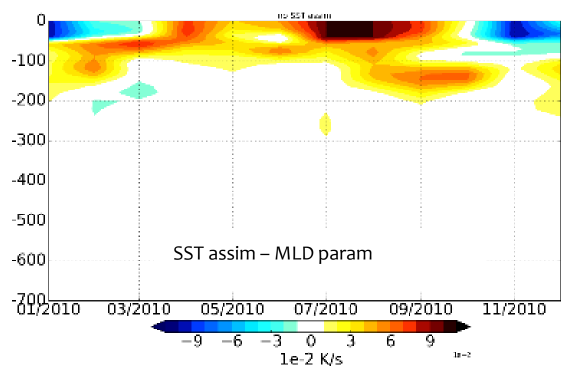
Nudging L4

DA L2P







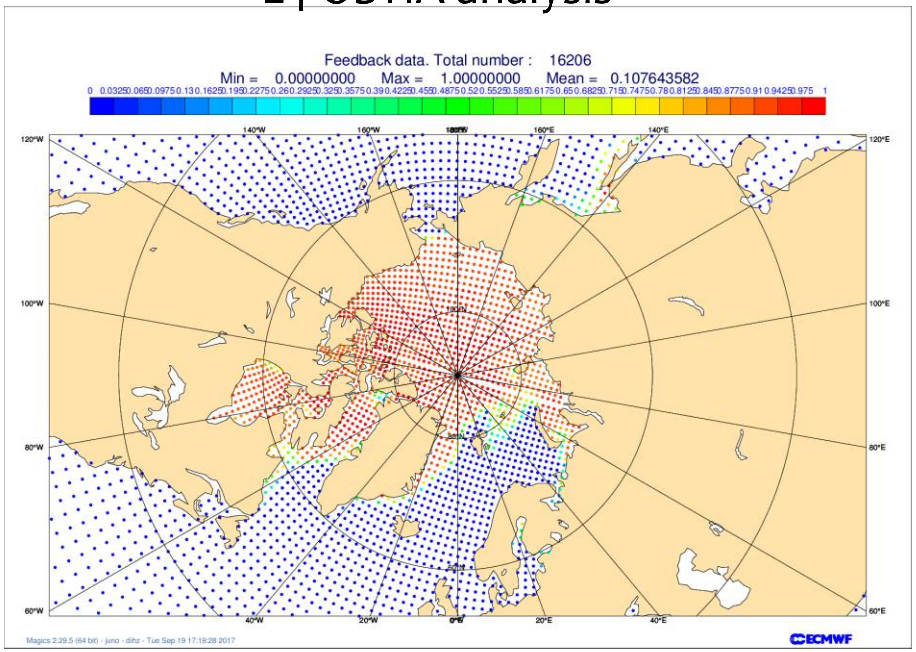


- MLD param allows the propagation of the T incr. down to the thermocline
- Further thinning and increased SST OE reduce the weight given to SST obs. wrt to profiles

Assimilating L3 SIC

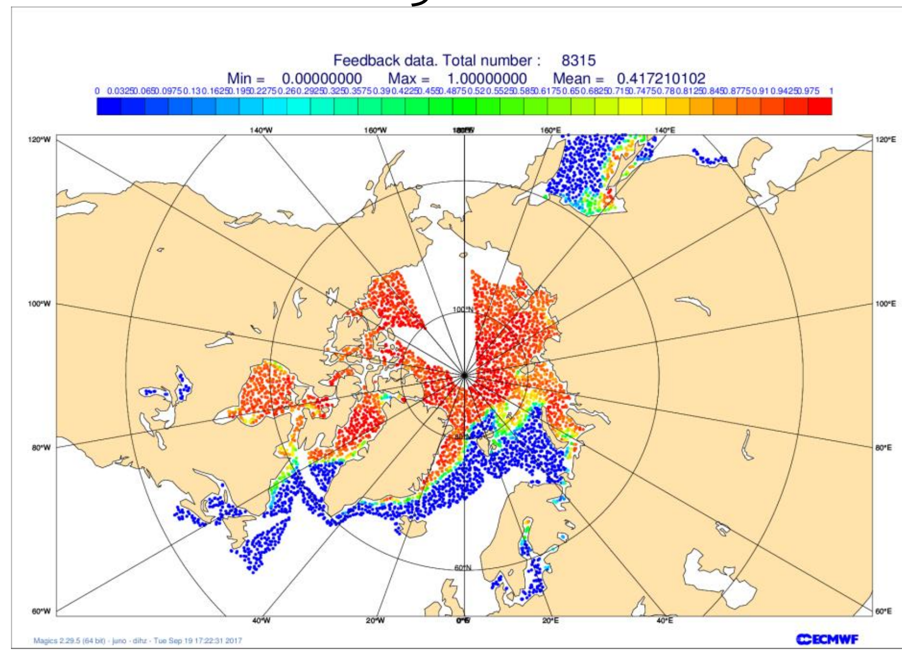
Daily assimilated SIC on 20130118

L4 OSTIA analysis



L4 analysis: with filtering, masking, extrapolation to produce a gap-free product

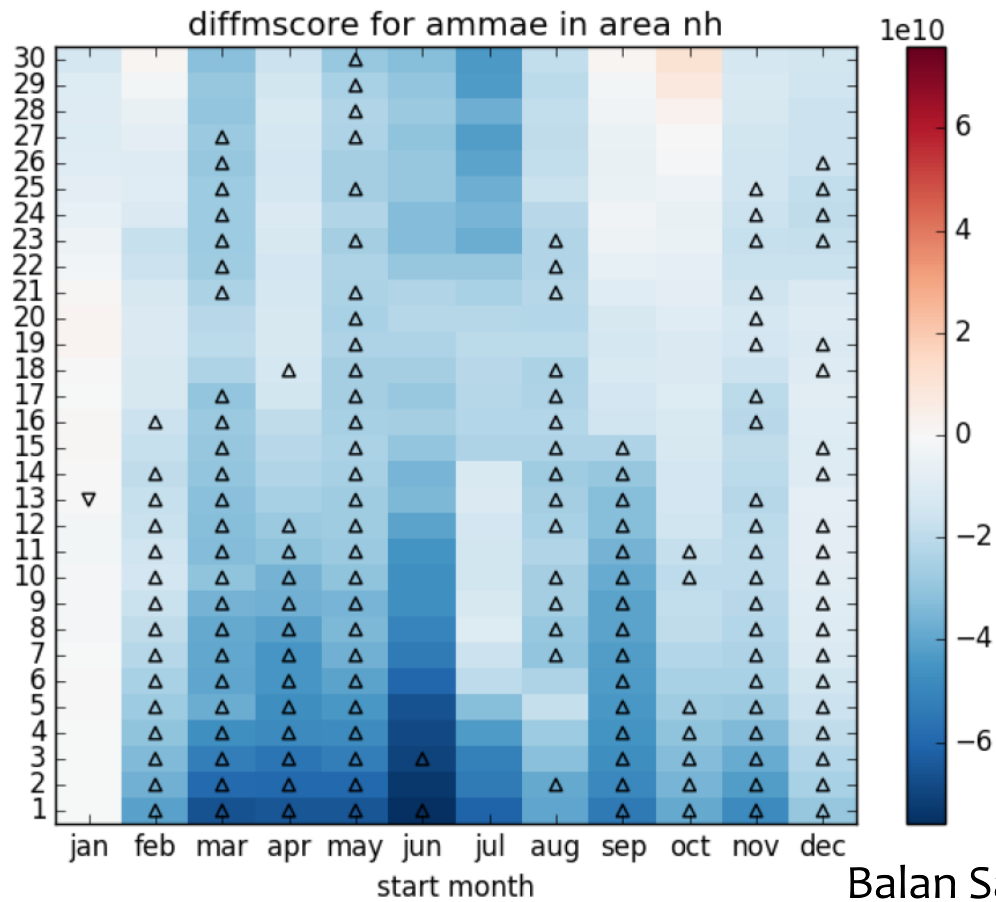
L3 OSI SAF



with 10km resolution there is ~1 milion obs per day from L3 OSI-SAF, obs reduced to ~10,000 per day with a thinning boxes of ~0.5X0.5 degree

Assimilating L3 SIC

Change in Forecasting errors in the NH: L3 – L4 SIC assimilation
(2005-2016, monthly reforecasts for the first 30 days, verified against OSI-SAF)



Balan Sarojini, et al., submitted

Summary

- **Data assimilation in the ocean serves a variety of purposes,** from climate monitoring to initialization of coupled model forecasts and ocean mesoscale prediction.
- This lecture dealt mainly with ocean DA for initialization of coupled forecasts and reanalyses. Global Climate resolution. NEMOVAR as an example.
- Compared to the atmosphere, **ocean observations are sparse.** The main source of information are temperature and salinity profiles (ARGO/moorings/CTDs), sea level from altimeter, SST/SIC/SIT from satellite and in-situ.
- **Assimilation of ocean observations reduces the large uncertainty due to both model and forcing errors.** It improves the initialization of coupled forecasts in NWP, and provides calibration and initialization for reforecast for seasonal forecasts and decadal forecasts.
- **Data assimilation changes the ocean mean state.** Therefore, consistent ocean reanalysis requires an explicit treatment of the bias. More generally, we need a methodology that allows the assimilation of different time scales.

Further Readings

Ocean Data assimilation

- Mogensen, K., Alonso Balmaseda, M., & Weaver, A. (2012). The NEMOVAR ocean data assimilation system as implemented in the ECMWF ocean analysis for System 4. Technical Memorandum (Vol. 668).
- Weaver, A. T., Deltel, C., Machu, É., Ricci, S., & Daget, N. (2005). A multivariate balance operator for variational ocean data assimilation. Quarterly Journal of the Royal Meteorological Society, 131(613), 3605–3625.

Ocean DA and Reanalysis

- Balmaseda, M. A., Mogensen, K., & Weaver, A. T. (2013). Evaluation of the ECMWF ocean reanalysis system ORAS4. Quarterly Journal of the Royal Meteorological Society, 139(674), 1132–1161. http://doi.org/10.1002/qj.2063
- Zuo, H., Balmaseda, M. A., & Mogensen, K. (2015). The new eddy-permitting ORAP5 ocean reanalysis: description, evaluation and uncertainties in climate signals. Climate Dynamics. http://doi.org/10.1007/s00382-015-2675-1
- Zuo, H., Balmaseda, M. A., Tietsche, S., Mogensen, K., and Mayer, M.: The ECMWF operational ensemble reanalysis-analysis system for ocean and sea-ice: a description of the system and assessment, Ocean Sci. Discuss., https://doi.org/10.5194/os-2018-154, in review, 2019.