Assimilation of SST observations with the new ECMWF Ensemble-variational Ocean DA system

Marcin Chrust, Hao Zuo, Philip Browne, Magdalena Alonso Balmaseda, Patricia de Rosnay, Eric de Boisseson, Beena Balan Sarojini

Thanks to *NEMOVAR team, especially Anthony Weaver*

ECMWF

Ocean DA at ECMWF: NEMOVAR

- **The “NEMOVAR” assimilation system used in ECMWF.**
  - Ensemble-Variational DA 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 background-error covariances formulation

General B formulation in NEMOVAR

$$\mathbf{B} = \beta_m^2\left(\mathbf{B}_{\mathrm{m}_1} + \mathbf{B}_{\mathrm{m}_2} + ..\right) + \beta_e^2\mathbf{B}_{\mathrm{e}} + \beta_E^2\mathbf{B}_{EOF}$$

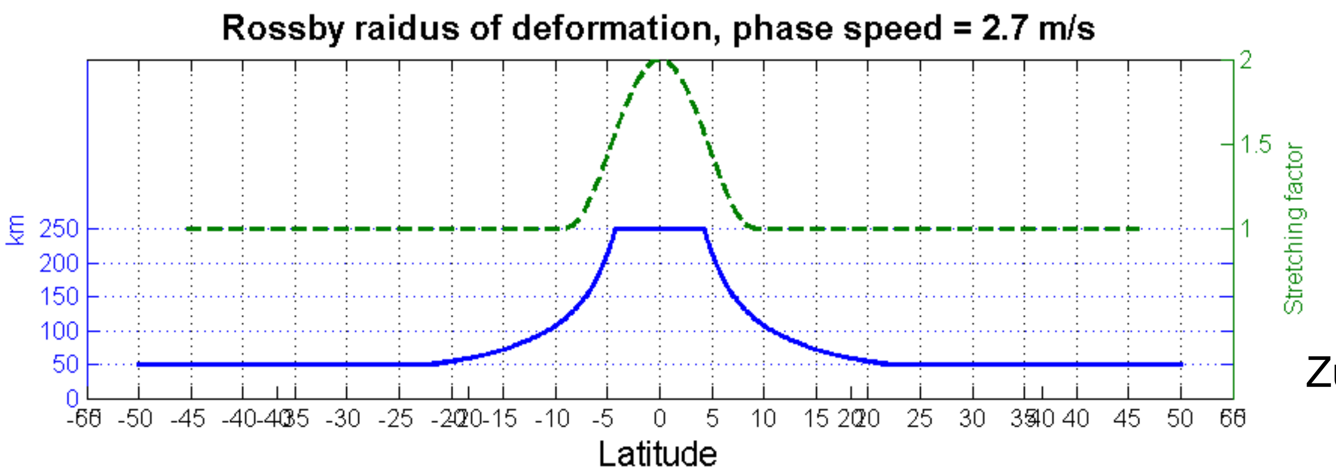
$$\mathbf{B}_{\mathrm{m_i}} = \mathbf{K}_{\mathrm{b}}\mathbf{D}_{\mathrm{i}}^{1/2}\mathbf{C}_{\mathrm{m_i}}\mathbf{D}_{\mathrm{i}}^{1/2}\mathbf{K}_{\mathrm{b}}^{\mathrm{T}}$$

- $\boldsymbol{B_m}$ is a modelled covariance matrix (can use multiple models to represent different scales)
- $\boldsymbol{B_e}$ is a localized ensemble-based covariance matrix
- $\boldsymbol{B_{EOF}}$ is an EOF-based covariance matrix
- $\boldsymbol{C_m}$ is a correlation matrix (including diffusion operator)
- $\boldsymbol{D_m}$ is a diagonal matrix of variances

$$\mathbf{C}_X^{1/2} = \boldsymbol{\Gamma}_X^{1/2}\boxed{\mathbf{L}_X^{1/2}}\mathbf{W}_X^{-1/2}$$

diffusion operator use diffusion tensor $\boldsymbol{\kappa_m}$ to represent a particular de-correlation length-scales

Horizontal correlation length-scales used in ORAS5



Zuo et al., 2015

Ensemble-variational DA with hybrid B

A cost-effective way to account for errors of the day is to use ensemble to estimate both variances ($\boldsymbol{D_m}$ -> $\boldsymbol{D_e}$) and the Local Correlation Tensor ($\boldsymbol{\kappa_m} \rightarrow \boldsymbol{\kappa_e}$ in $\boldsymbol{C_m}$) in a **modelled covariance** matrix $\boldsymbol{B_m}$

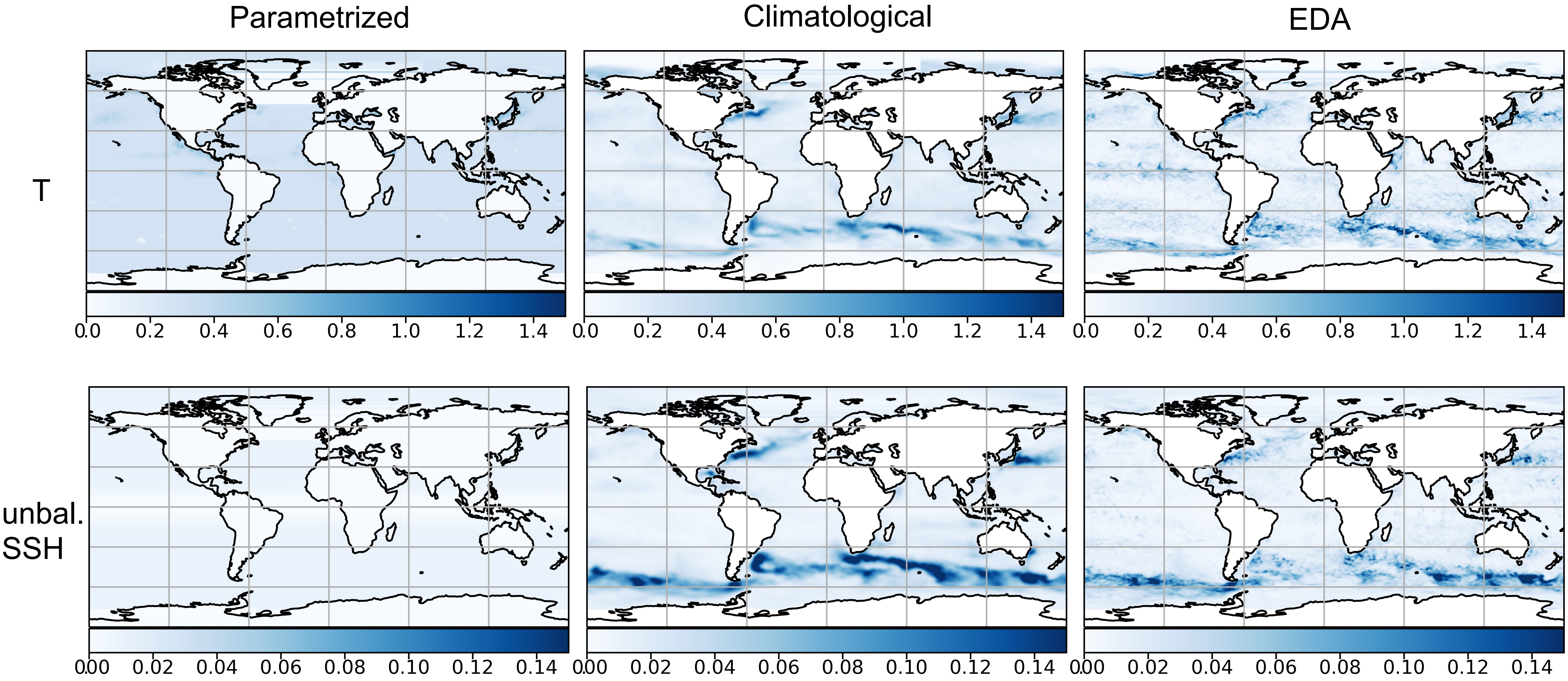
$$\mathsf{B}_{\mathrm{m}} = \mathsf{K}_{\mathrm{b}}\,\mathsf{D}_{\mathrm{m}}^{1/2}\,\mathsf{C}_{\mathrm{m}}\,\mathsf{D}_{\mathrm{m}}^{1/2}\,\mathsf{K}_{\mathrm{b}}^{\mathrm{T}}$$

Given an EDA ensemble we would like to construct an ensemble of perturbations to estimate the parameters of the modelled **B** matrix to capture errors of the day.

- We address sampling errors using objective spatial filter of *Ménétrier et al*. (2015)
- Local Correlation Tensor (LCT) is computed using ensemble-gradient method

Being able to estimate flow dependent **variances** and the **local correlation tensor** using EDA also allowed us to produce climatological estimates of parameters.

Hybrid background errors standard deviations – ORCA025 (15 member ensemble)



Ensemble-variational DA with hybrid B

Thanks to C3S ERGO project!

NEMOVAR covariance formulation

$$B_m = K_b D_m^{1/2} C_m D_m^{1/2} K_b^T ,$$

$\boldsymbol{B_m}$ is the modelled covariance matrix; $\boldsymbol{C_m}$ is the correlation matrix (using diffusion operator $\boldsymbol{\kappa}$); $\boldsymbol{D_m}$ is a block-diagonal matrix of variances (σ^2).

Hybrid background error std formulation (because our current ens spread is sub-optimal)

$$\sigma_h = \frac{1}{h}\log(e^{hw_m\sigma_m} + e^{hw_e\sigma_e} - 1),$$

and $\sigma_m = \frac{1}{h}\log(e^{hw_c\sigma_c} + e^{hw_p\sigma_p} - 1)$

Here w_c, w_p, w_m, w_e are dimensionless weighting factors for *clim/param/modelled/ensemble* components of variances

Hyrbid diffusion tensor formulation

$$\kappa_h = \gamma_m^2 \kappa_m + \gamma_e^2 \kappa_e$$

and $\kappa_m = \gamma_p^2 \kappa_p + \gamma_c^2 \kappa_c$

Here $\gamma_c^2, \gamma_p^2, \gamma_m^2, \gamma_e^2$ are 3D (+direction) weighting factors for *clim/param/modelled/ensemble* components of diffusion tensor

Hybrid background error variances (σ_h^2)

Hybrid background error variances $\boldsymbol{\sigma_h^2}$ contain *parameterized* $\boldsymbol{\sigma_p^2}$ *and climatology* $\boldsymbol{\sigma_c^2}$, with option to add "error-of-the-day" estimated from *ensemble spread* ($\boldsymbol{\sigma_e^2}$).

Parameterized $\boldsymbol{\sigma_p^2}$ | Climatology $\boldsymbol{\sigma_c^2}$ | Hybrid $\boldsymbol{\sigma_h^2}$

Temperature

Unbal. Salinity

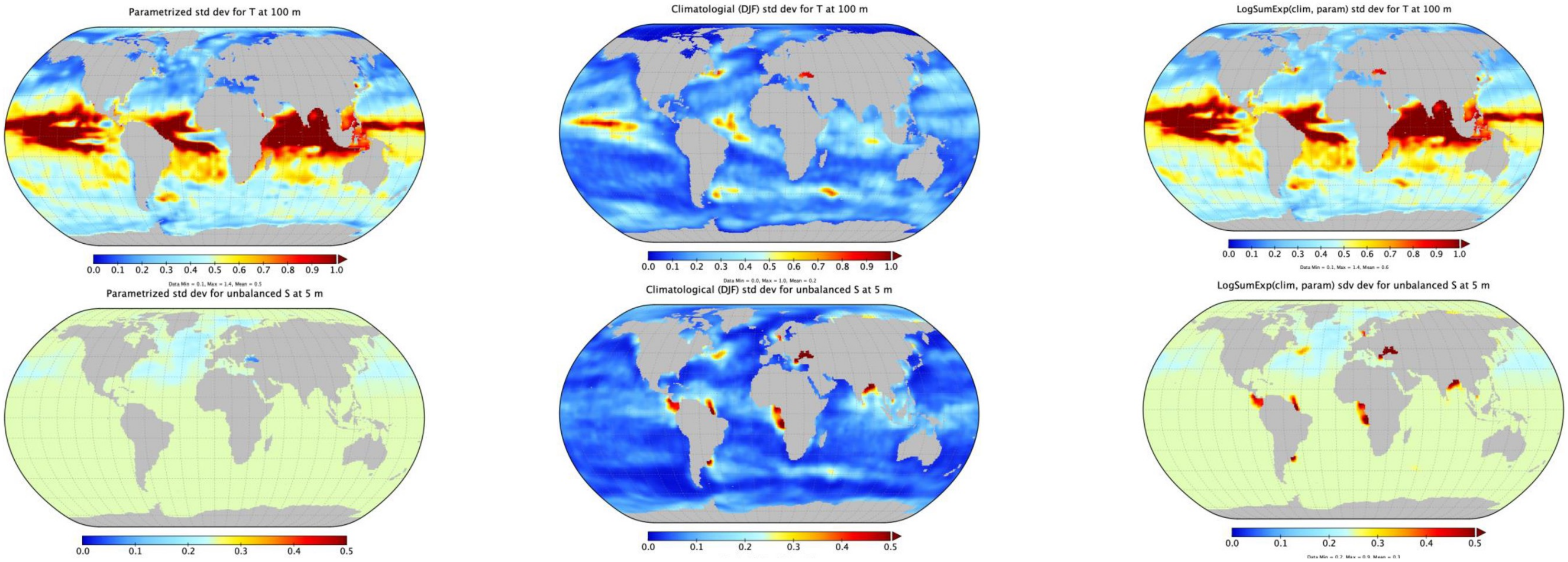
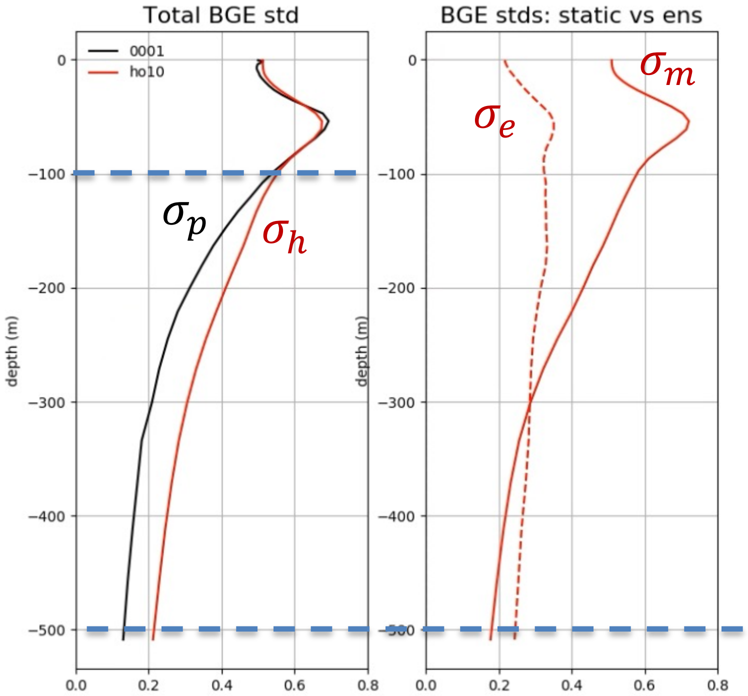


Figure by Anthony

Consolidate hybrid variance (σ_h^2)

- Global mean hybrid T variance in the upper ocean is tuned to be close to OCEAN5
- Hybrid T variance below the thermocline is larger than OCEAN5 due to softmax hybrid formulation
- Larger hybrid σ_b in the WBCs mostly comes from climatological σ_c

Global mean Temperature BGE std



σ_h^2: hybrid. var.

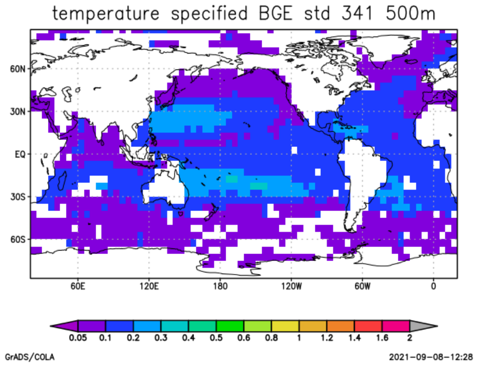
Static σ_m^2 : σ_p^2: param. var.; σ_c^2: clim. var.

Flow-dep. var.: σ_e^2: ens. var.

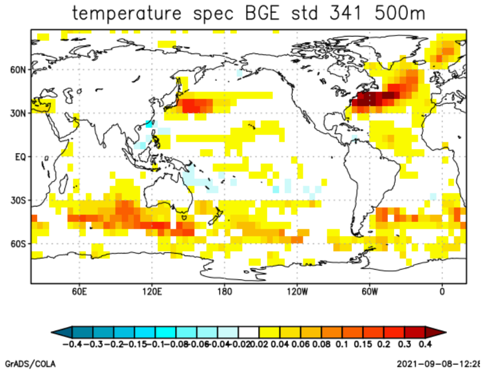
Effective Temperature BGE std

500m

OCEAN5



Hybrid – O5



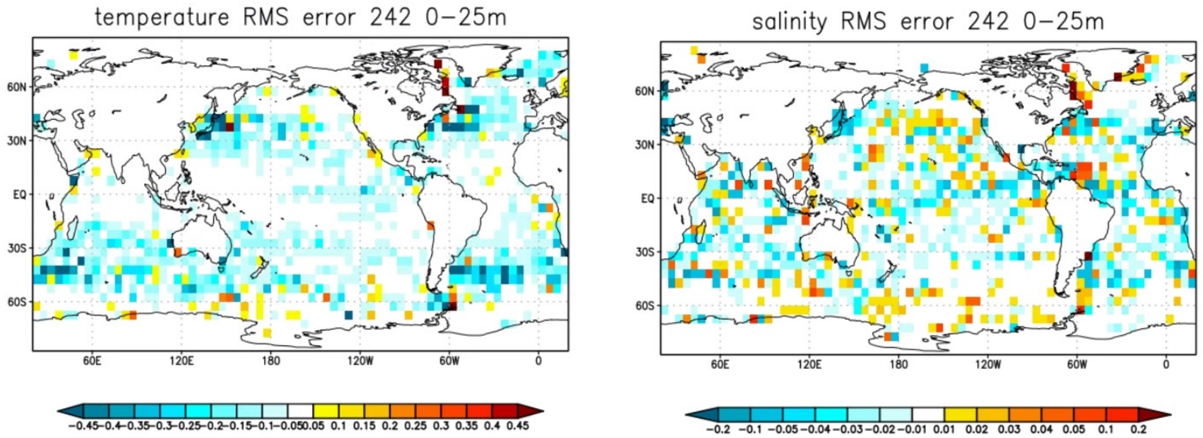
OCEAN5 use only parametrized variance: σ_p

Hybrid variance: $\sigma_h = \frac{1}{h}\log(e^{hw_m\sigma_m} + e^{hw_e\sigma_e} - 1)$

Consolidate Hybrid tensor (κ_h)

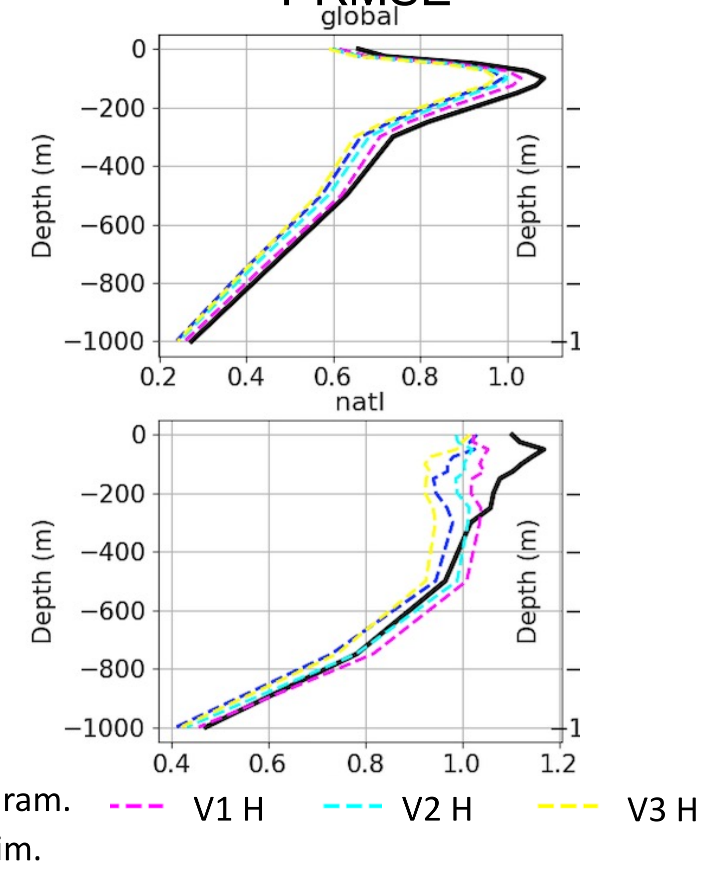
- Consolidation of hybrid-B configurations with our new ocean EDA system has been completed.
- Among all tested configurations, using a retuned hybrid horizontal diffusion tensor (parameterized tensor in tropics + climatological tensor in extra-tropics) together with an ensemble-based vertical diffusion tensor that updates every cycle gave the best performance

Hybrid tensor - parameterized tensor



Changes in O-B RMSE (hybrid - param. B), verified against all in-situ observations

T RMSE



Direct SST DA with Ocean EDA

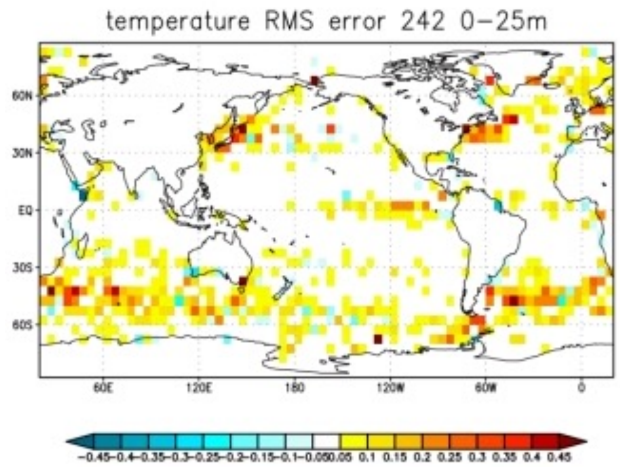
Assimilation of SST with ocean EDA and hybrid-B

Key developments for SST assimilation:

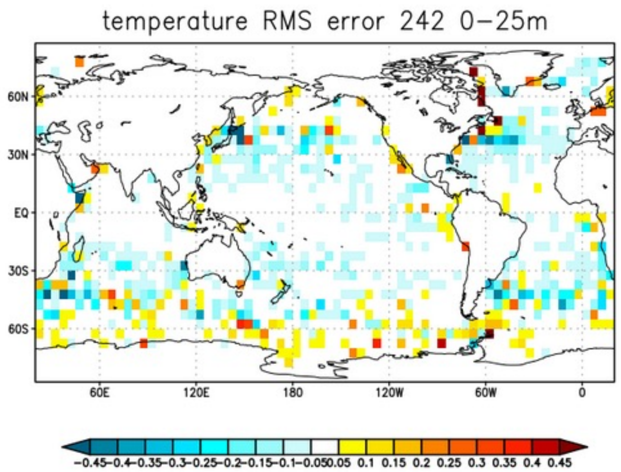
- Flow dependent vertical correlation scales are critical, thanks to the factorized formulation of normalization factors;
- Replaced randomized approximate estimation of vertical normalization factors with exact, much more efficient direct computation;

Changes in O-B RMSE: SST DA – SST nudging verified against all in-situ observations

- SST DA performance is much worse than SST nudging with parameterized diffusion tensor (as in OCEAN5)



- Switching to hybrid tensor with ensemble vertical tensor improves WBCs and ACC

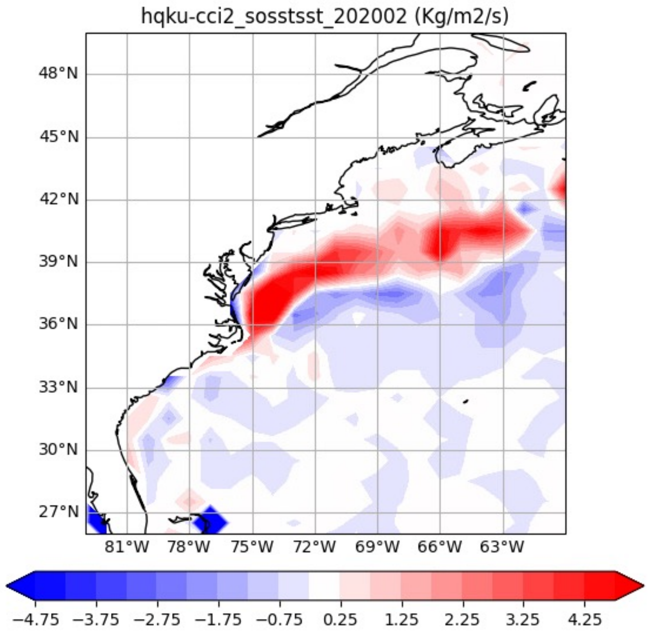


SST assimilation with EDA: Gulf Stream

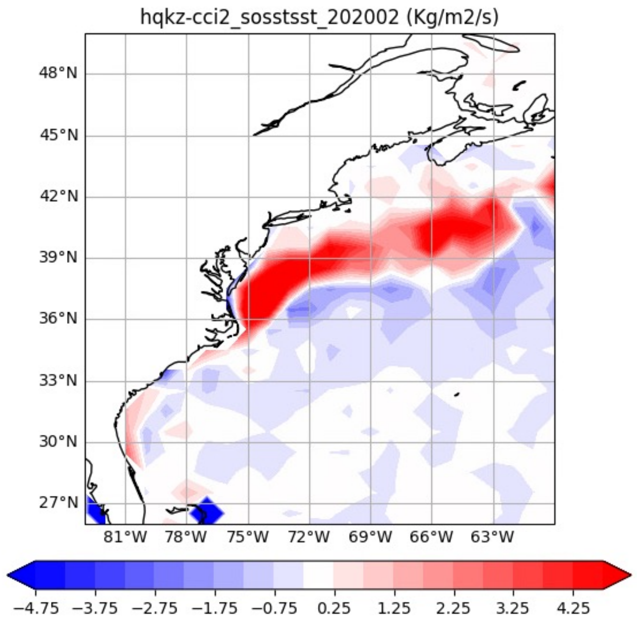
- Assimilation of L4 SST with *parameterized* tensor increased SST biases in the Gulf Stream region
- Using an *ensemble-based* V-tensor in hybrid-B has greatly reduced SST biases, especially in the Gulf Stream regions where the ¼ degree NEMO model has persistent bias.

SST biases in the GS region after 2-months of DA

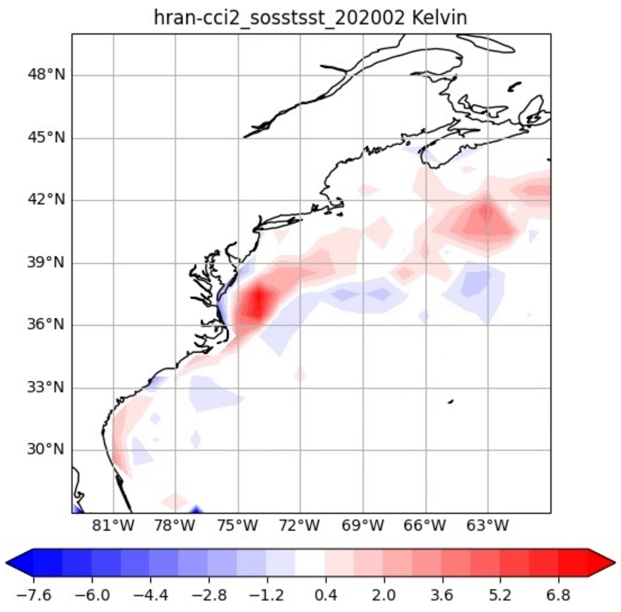
SST Nudging (O5-like)



SST DA (param. tensor)



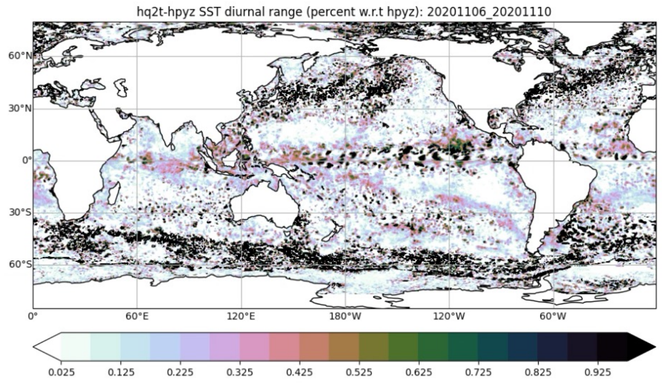
SST DA (V3 hybrid)



SST assimilation with EDA: diurnal cycle

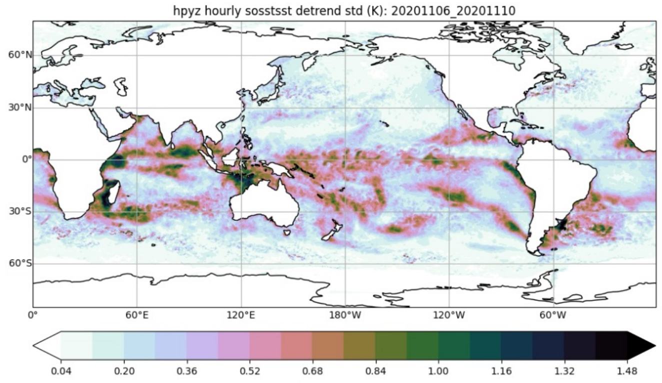
- Nudging to daily mean SST damps SST diurnal cycle (switched off in coupled DA)
- Direct assimilation of L4 SST with Ocean EDA system enhance the diurnal cycle (~15%) of analysis SST

SST DA – nudging (norm)

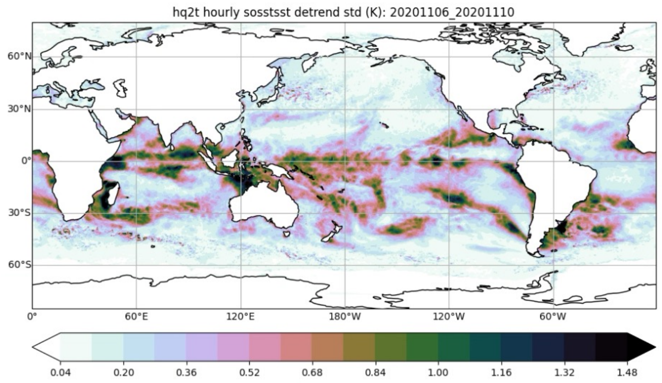


SST diurnal range (in K, 5-day mean)

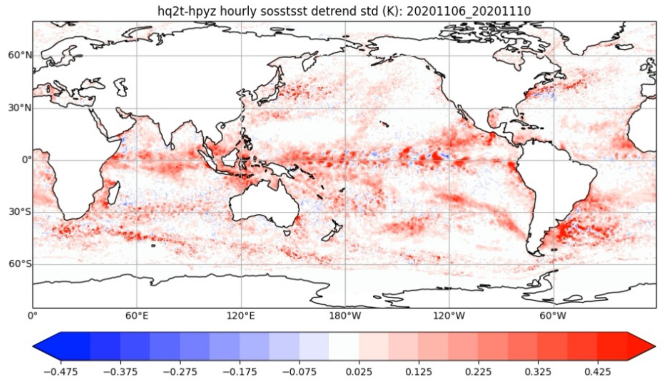
SST nudging



SST DA



SST DA - nudging



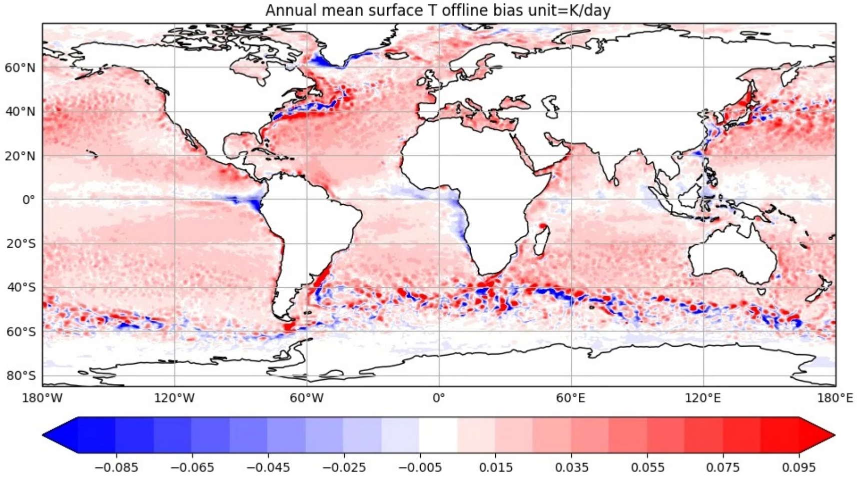
ECMWF

SST assimilation: impact of BC

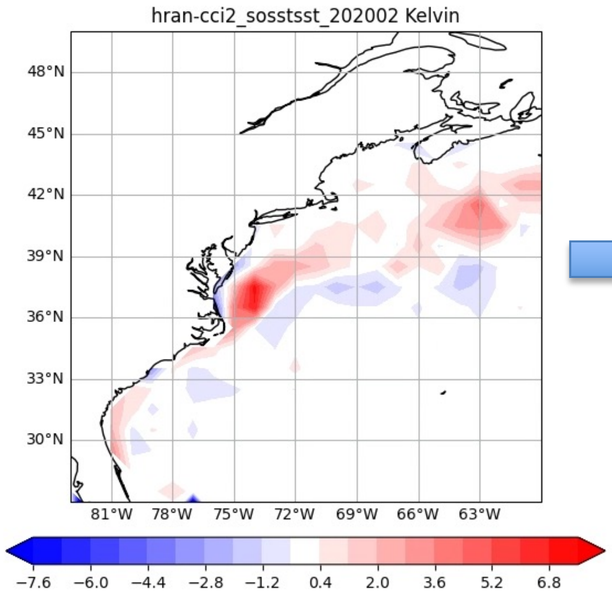
The BC scheme used in ECMWF ODA system includes a-priori biases with seasonal variations

- With direct assimilation of SST, SST bias is visible in surface temperature increments, and can be included when constructing a-priori T bias term
- BC can further improve the GS performance

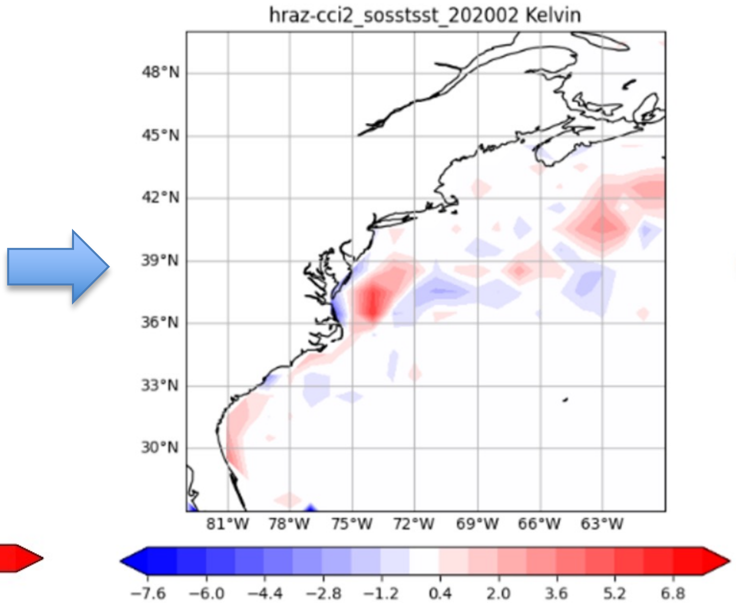
Surface T bias including SST increments



SST DA (hybrid-B)



SST DA (hybrid-B + BC)



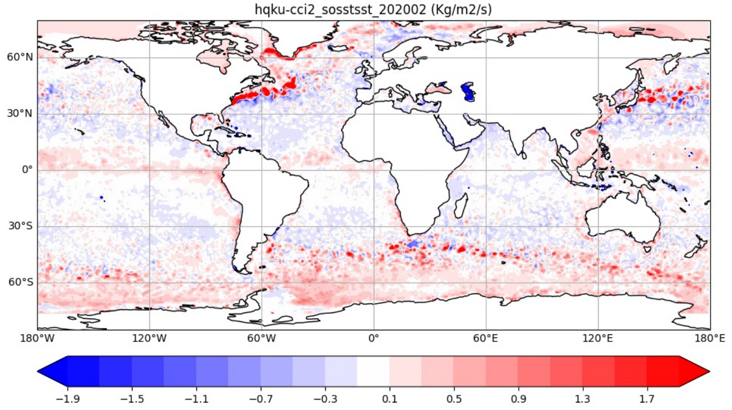
ECMWF

SST assimilation with EDA

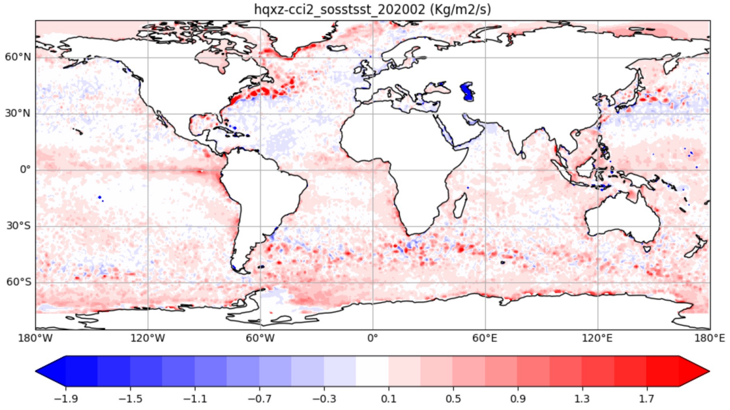
- Warm biases are noticeable with SST DA, which is related with pre-processing of L4 SST data. This warm bias can be reduced by assigning different Local Solar Time (LST) to SST obs.
- Work is on-going to explore L2 SST, and L1 radiance data through a coupled DA approached at ECMWF.

SST biases after 2-months of DA

SST Nudging (O5-like)

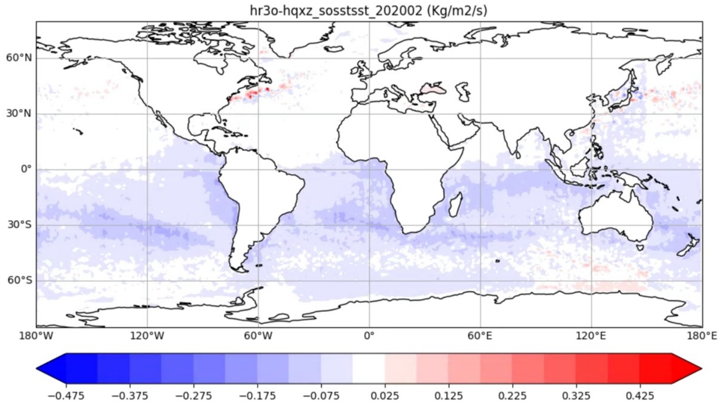


SST DA (hybrid-tensor)



different obs pre-processing

LST10:00 – LST06:00



Summary

- ECMWF is developing the 6th generation of ocean and sea-ice ensemble reanalysis-analysis system – OCEAN6; major updates include a new Ensemble based variational DA system with a hybrid-B approach.
- Compared to SST nudging, direct assimilation of SST data improves the SST diurnal cycle. This approach also greatly reduces the SST biases in the Gulf Stream regions, but only when an ensemble based vertical diffusion tensor is used in the EDA system.
- This has been found that SST performance is also very sensitive to the forward model used in observation operator, as well as observation data quality control and pre-processing strategy.
- Other on-going ocean DA developments at the ECMWF (and friends!): 4D-Var, weak constraint/model error correction; new ensemble generation methods; multi-scale B; correlated observation errors; etc.

Extra slides

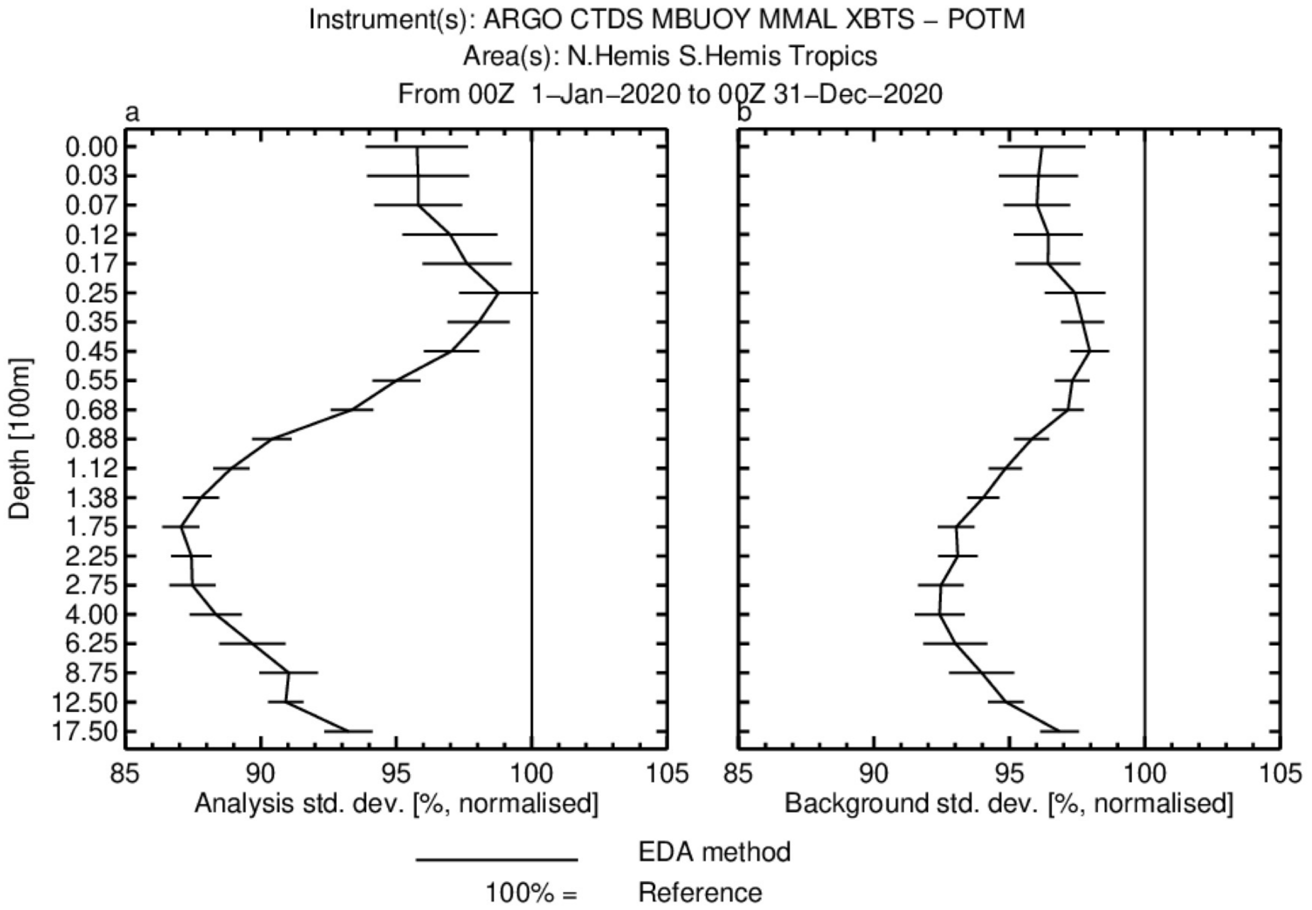


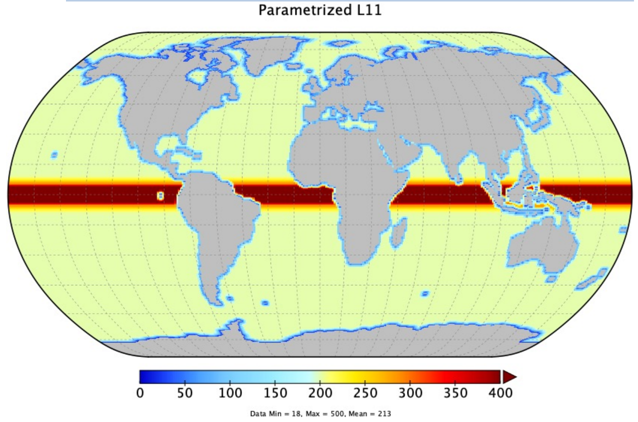
Figure 5: Normalised reduction of misfit between the NEMOVAR analysis and background to in-situ observations when the new ensemble-based vertical length scales (EDA method) are used, compared to the old mixed layer depth-based formulation (100% line, Reference). These are significant results obtained over a whole year (2020).

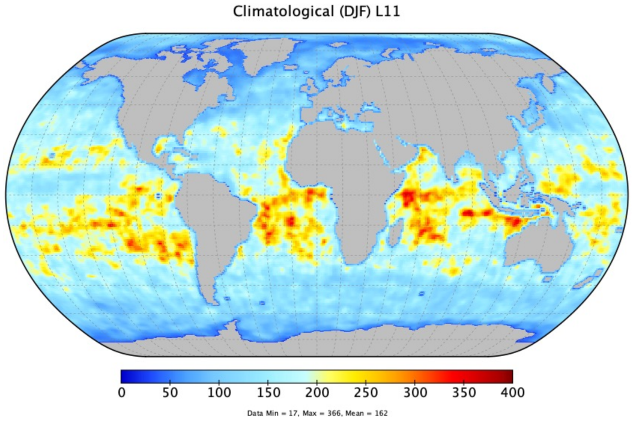
Hybrid diffusion tensor (κ_h)

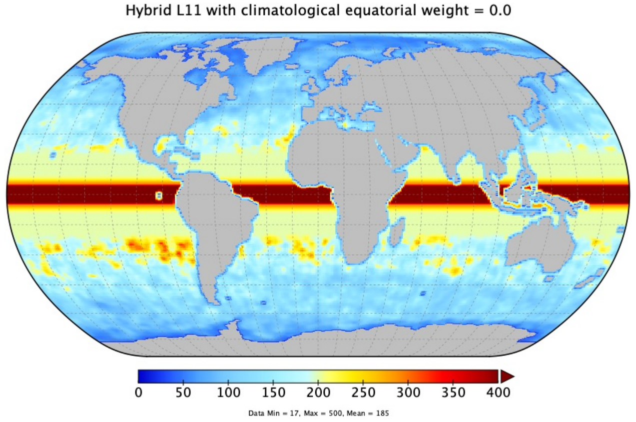
Diffusion tensor κ in NEMOVAR is separated into three directions (κ^{11} - zonal, κ^{22} - meridional and κ^{33} - vertical), and can be specified individually with various hybrid settings.

Parameterized H tensor	climatology H tensor	hybrid H tensor

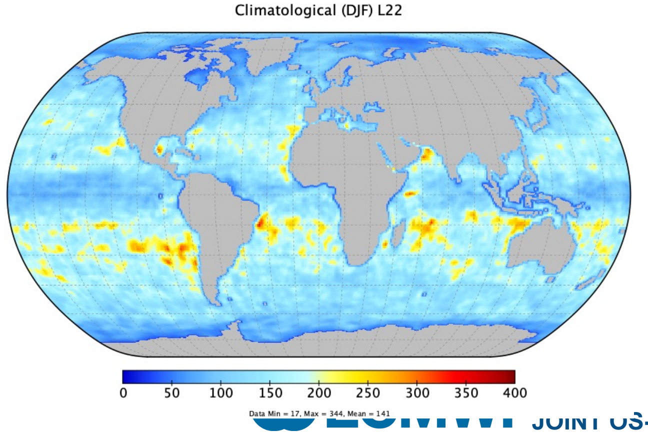
Zonal (km)

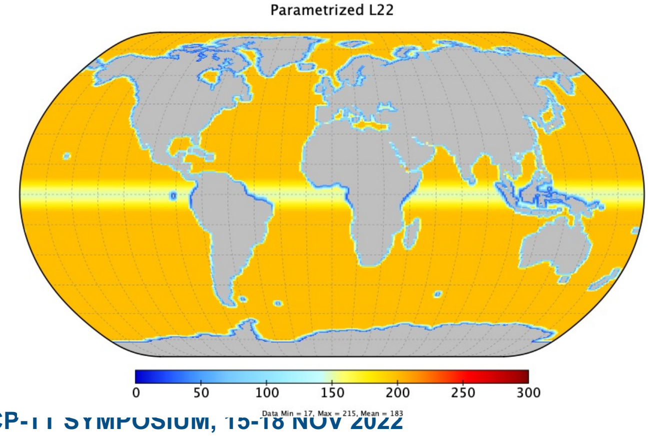






Meridional [km]





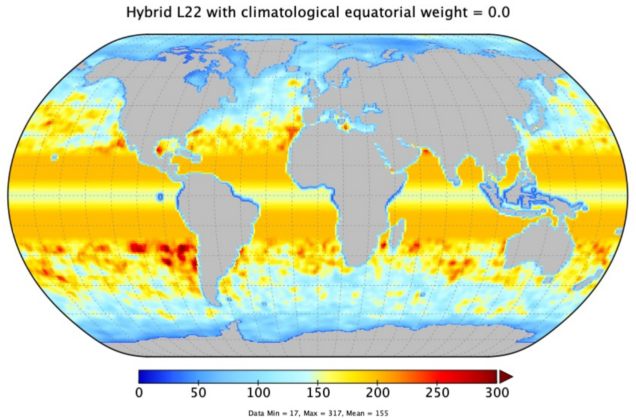


Figure by Anthony