Coupled multi-grid stochastic modelling and data assimilation and their impact on regional/coastal forecasting in the Bay of Biscay

Pierre De Mey-Frémaux (DR CNRS, LEGOS, Toulouse, France)
Vassilios Vervatis (Prof., NK U. Athens, Greece)
Bénédicte Lemieux-Dudon (Contractor, LAERO, Toulouse, France)

10th meeting of the OceanPredict Coastal Ocean and Shelf Seas Task Team
Ifremer, Brest, 24-26 September 2024

Copernicus Marine Service / Service Evolution projects

- **SCRUM (2016-18) – Stochastic Coastal/Regional Uncertainty Modelling**

  Focus: Regional Ensemble consistency verification, Ensemble data assimilation

  – Vervatis, V.D., P. De Mey-Frémaux, et al., 2021a: *Assessment of a regional physical-biogeochemical stochastic ocean model. Part 1: Ensemble generation.* Ocean Modelling, 160, 101781.

  – Vervatis, V.D., P. De Mey-Frémaux, et al., 2021b: *Assessment of a regional physical–biogeochemical stochastic ocean model. Part 2: Empirical consistency.* Ocean Modelling, 160, 101770.

- **SCRUM2 (2018-20) – Stochastic Coastal/Regional Uncertainty Modelling II**

  Focus: Use of atmospheric Ensembles, age of errors, probabilistic forecasting

  – Vervatis, V.D., P. De Mey-Frémaux, et al., 2025: *Regional ocean model uncertainties using stochastic parameterizations and a global atmospheric ensemble.* Ocean Modelling, 194.

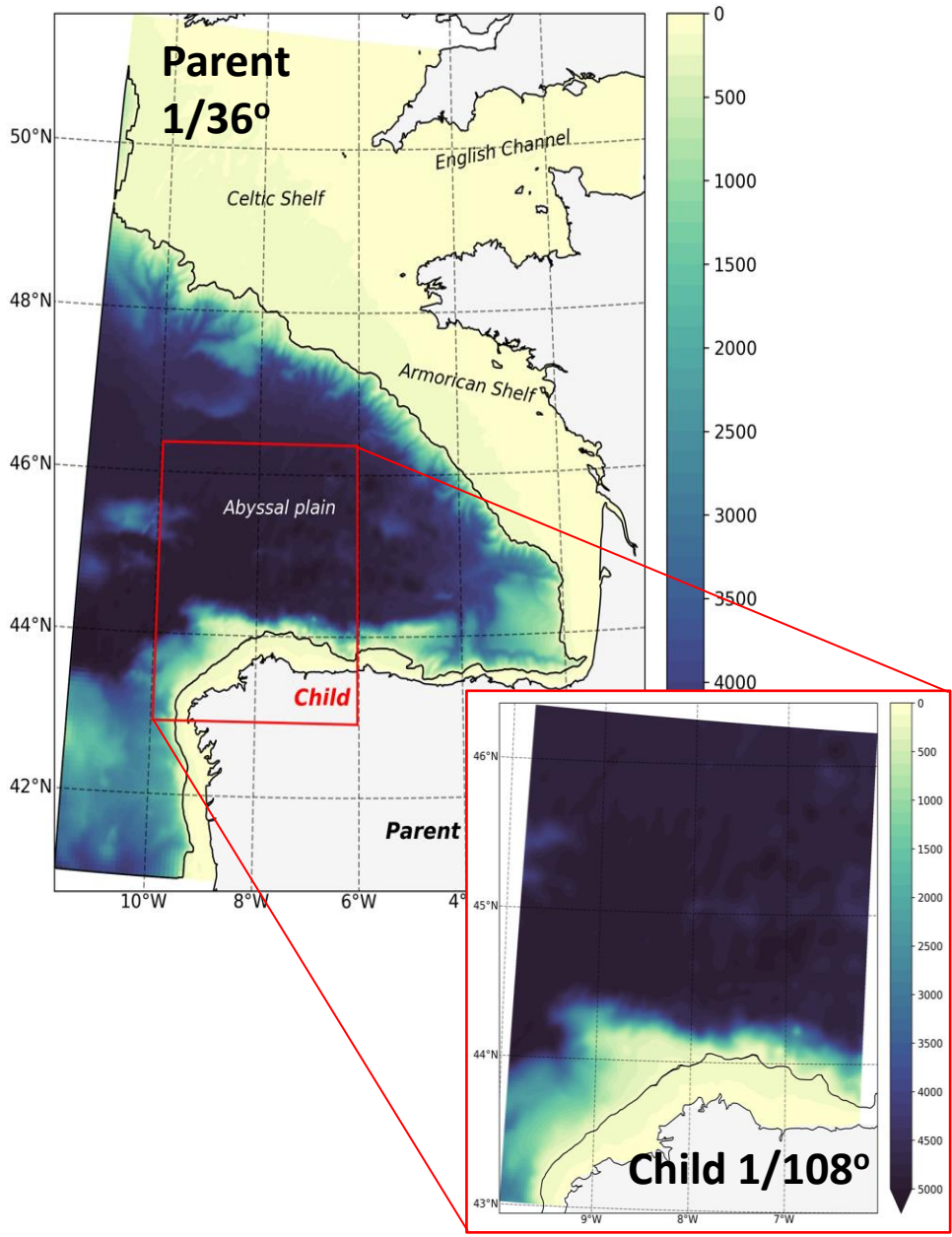
- **MultiCast (2022-24) – Stochastic multi-grid ocean forecasting**

  Focus: Ensemble-based multiresolution regional/coastal forecasting, Ensemble-based multigrid data assimilation

  – De Mey-Frémaux, P., V. Vervatis, B. Lemieux-Dudon, 2025: *Coupled multi-grid stochastic modelling and data assimilation and their impact on regional/coastal forecasting in the Bay of Biscay.* In prep.

Dual-grid stochastic model generation in MultiCast

- Stochastic models provide consistent Ensembles, which are in turn used in (1) Ensemble DA, and (2) probabilistic forecast skill metrics
- Stochastic models are built from deterministic models: BISCAY36 parent + nested 1/108° child domain, both using NEMO 4.2, with AGRIF coupling
- SPPT-AR1 surface wind vector perturbations
- Ensembles: 20 (7) members + CONTROL member
- Two periods in February and June 2017
- Free or assimilated Ensemble runs
- 30-day medium-range Ensemble forecasts.



Dual-grid stochastic model classes

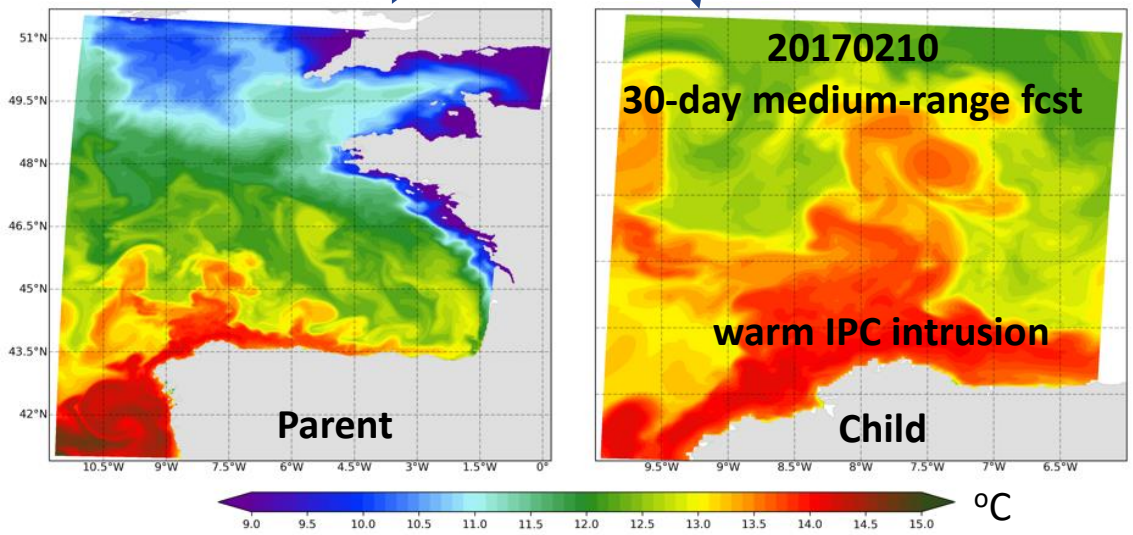
- The MultiCast project was funded under the Copernicus Marine Service.
- We address target classes of operational protocols >>> enhancing fine-scale forecasts in certain key areas.
- We use probabilistic scores to assess the forecast skill of stochastic models.

Classes of Target Operational Protocols (TOPs) addressed

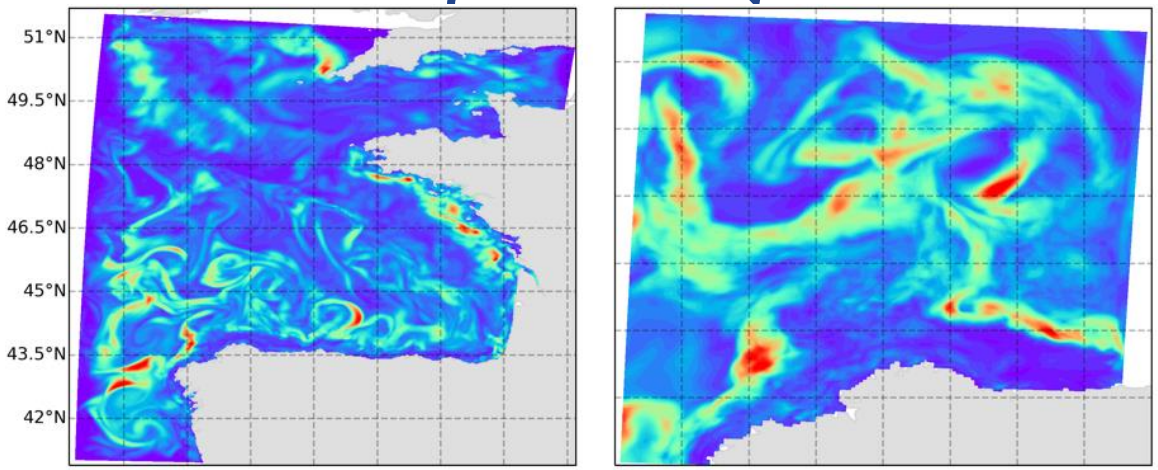
Class	Class definition	Stochastic approach
TOP-0	**1-way nesting, stochastic parent** • Parent d.o.f's from **intrinsic parent errors** • Child d.o.f's only from **downscaling**	• Wind perturbed in parent
TOP-1	**1-way nesting, stochastic parent & child** • Parent d.o.f's from **intrinsic parent errors** • Child d.o.f's both from **downscaling** and **intrinsic child errors**	• Wind perturbed in parent and child
TOP-2	**2-way nesting, stochastic parent & child** • Parent d.o.f's both from **upscaling** and **intrinsic parent errors** • Child d.o.f's both from **downscaling** and **intrinsic child errors**	• Wind perturbed in parent and child

Parent-child ensemble spread – winter

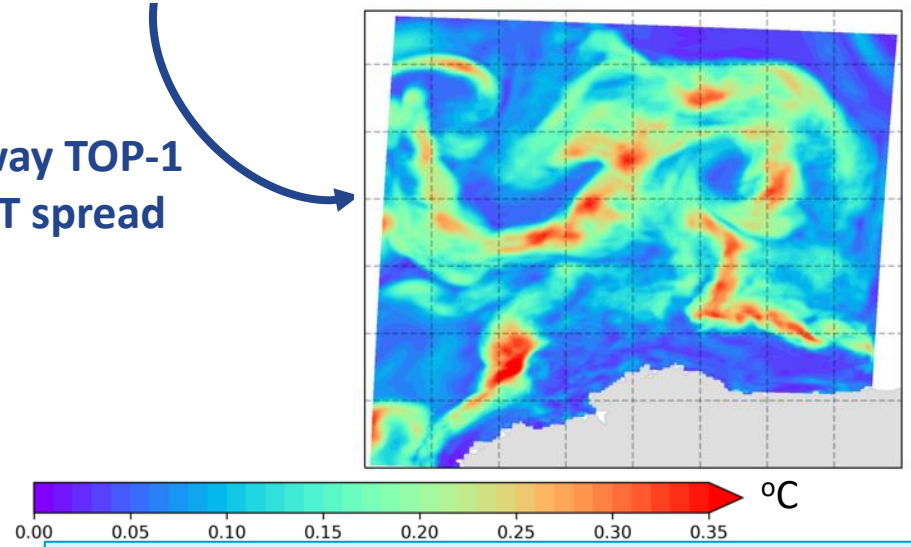
2-way SST reference



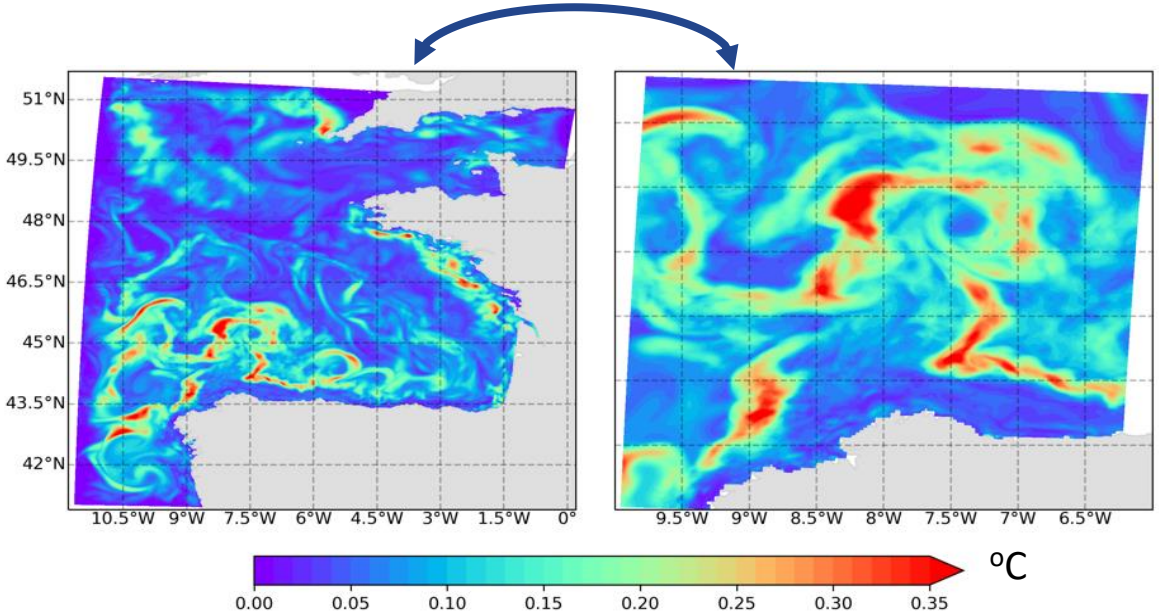
1-way TOP-0 SST spread



1-way TOP-1 SST spread



2-way TOP-2 SST spread



TOP-0: 1-way nesting with stochastic parent
TOP-1: 1-way nesting with stochastic parent & child
TOP-2: 2-way nesting with stochastic parent & child

Probabilistic skill scores

Brier score associated to event **E** : $X \leq x_t$ being given a parameter X and a threshold x_t :

$$\mathrm{BS}(\mathrm{x_t}) = \frac{1}{K} \sum_{k=1}^{K} \left(P_k^f(x_t) - O_k(x_t) \right)^2$$

Having K realizations of **E**, the Brier score compares:

- ➔ P^f_k the probability predicted by the stochastic model / Ensemble members $(i=1,\ldots,N)$
- ➔ O_k the occurrence of **E** captured by the verifying observations.

$$P_k^f \left(X_{i,k}^f \leq x_t \right) = p_i = \frac{i}{N} \in [0,1]$$

$$O_k(X_k^o \leq x_t) = 0 \text{ or } 1$$

DRPS (*Discrete Ranked Probability Score*): extension of the BS over several thresholds

$$< \mathrm{DRPS} >= \frac{1}{L} \sum_{l=1}^{L} \mathrm{BS}(x_t^l) = \frac{1}{L} \sum_{l=1}^{L} \frac{1}{K} \sum_{k=1}^{K} \left(P_k^f(x_t^l) - O_k(x_t^l) \right)^2$$

CRPS (*Continuous Ranked probability Score*): extension to a “continuous” range of thresholds

$$< \mathrm{CRPS} >= \int_{-\infty}^{+\infty} \mathrm{d}\, x_t\, \mathrm{BS}(x_t) = \int_{-\infty}^{+\infty} \mathrm{d}\, x_t\, \frac{1}{K} \sum_{k=1}^{K} \left(P_k^f(x_t) - O_k(x_t) \right)^2$$

Probabilistic skill scores: the Hersbach (2000) CRPS decomposition

(Based on the Murphy, 1973 Brier score decomposition)

An effective Ensemble-based prediction system should have the following skills :

- **Reliability**: statistical consistency between the probability predicted by the Ensemble and the frequency of occurrence captured by the verifying observations.
- **Resolution**: ability of the EPS to perform better than a climatological forecast.

$$\overline{\mathrm{CRPS}} = \overline{\mathrm{Reli}} + \overline{\mathrm{CRPS\_pot}}$$

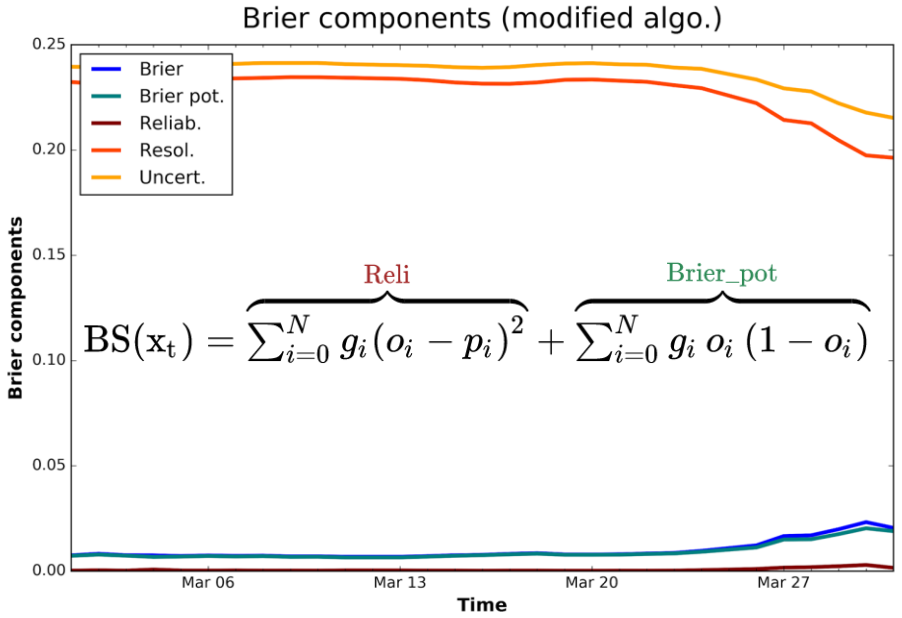
$$\overline{\mathrm{CRPS\_pot}} = -\overline{\mathrm{Reso}} + \overline{\mathrm{Uncert}}$$

- The smaller the better, except Resolution.
- Uncertainty only depends on the verifying observations (base rate variance).

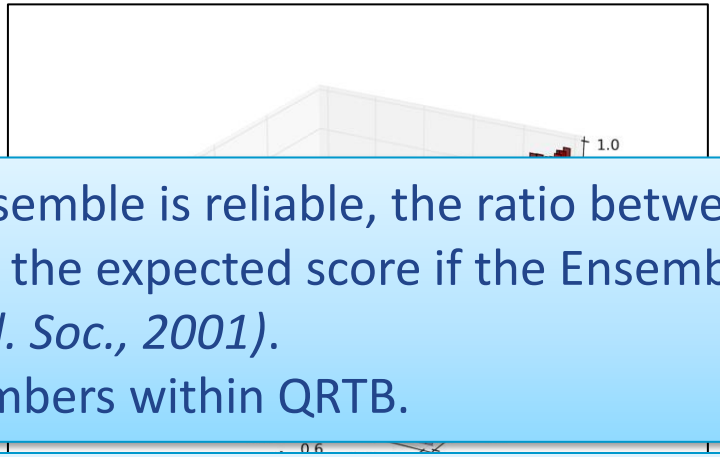
Synthetic observations: quasi-reliable validation framework

QRTB (Quasi-Reliable Test-Bed) – split single Ensemble:

- Synthetic verifying observations generated from one or several randomly drawn Ensemble member(s) at observation locations using observation operator **H** and noisified using observational error covariance matrix **R**.
- Run fcst skill assessment on rest of Ensemble.
- High Reliability skills expected
  ➔ Under QRTB, scores will mostly discriminate **Resolution**.



Reliability diagram



SCRUM2 upw

- Brier (bin
- Synthetic
  drawn fro

Useful property: CRPS sampling error: When the Ensemble is reliable, the ratio between the expected CRPS of the m-member-based forecast and the expected score if the Ensemble was of infinite size is $1+1/m$ *(Richardson, Q. J. R. Meteorol. Soc., 2001)*.
➔ E.g., we can expect +/-5% CRPS error with 20 members within QRTB.

➔ High Reliability (as expected) and Resolution skills
➔ Fair, “linear” Reliability diagram, within some sampling errors

SST CRPS for dual-grid stochastic models (free = no assim)

Parent SST CRPS

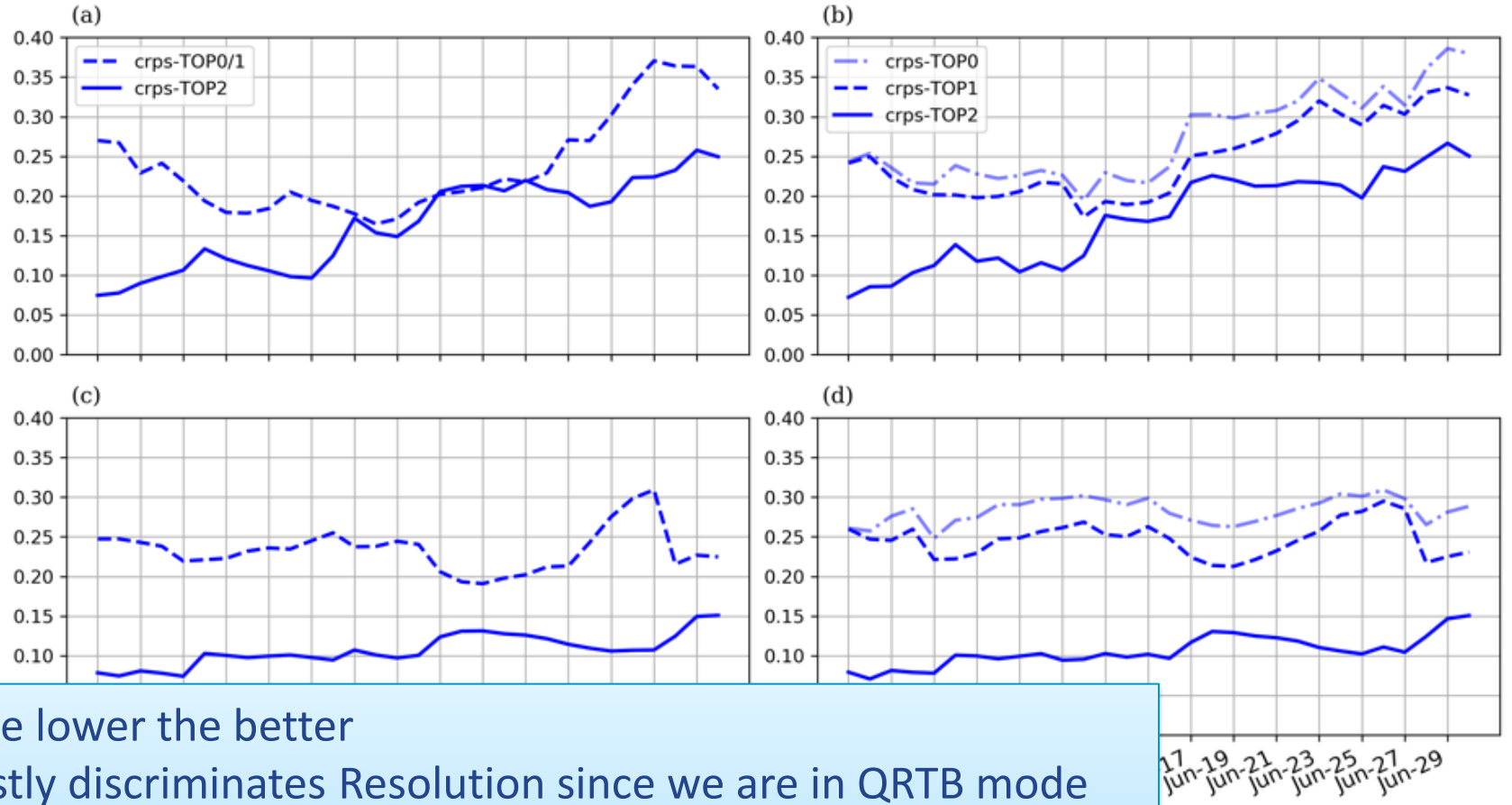
Child SST CRPS

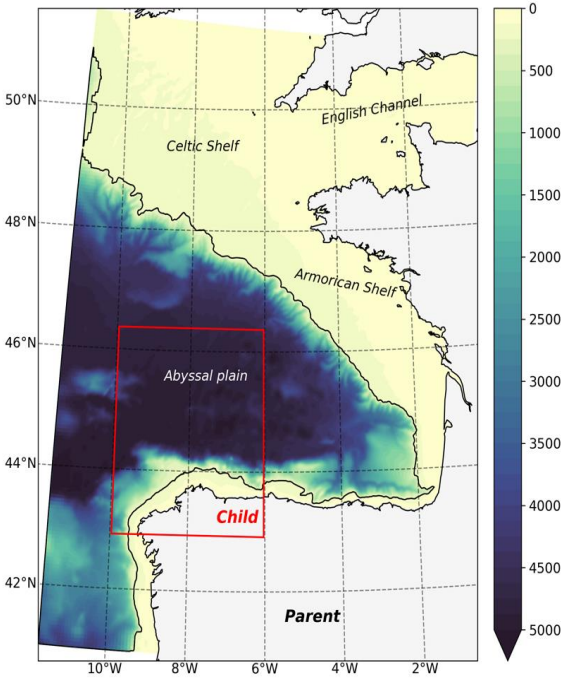
Summer 2017

QRTB

Open ocean

Coastal strip





- CRPS: the lower the better
- This mostly discriminates Resolution since we are in QRTB mode
- We verify that the TOP-2 free model (2-way nesting with stochastic parent & child) shows the best performance among the three models
- The free model performance is better, even for the parent, when a child is two-way coupled to the parent

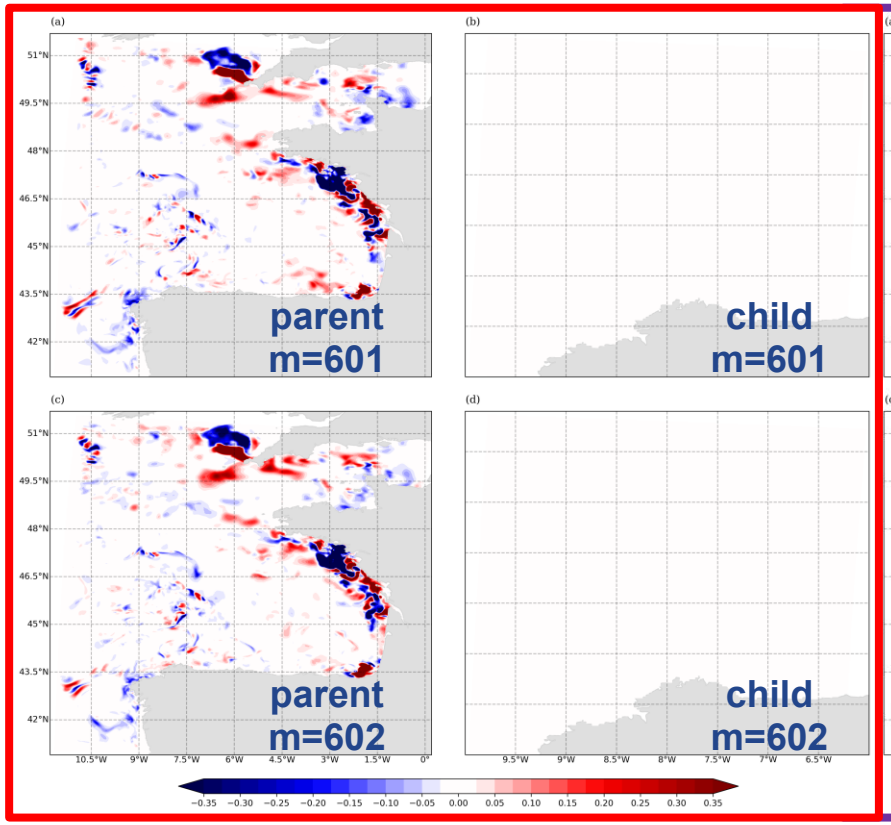
TOP-0: 1-way nesting with stochastic parent
TOP-1: 1-way nesting with stochastic parent & child
TOP-2: 2-way nesting with stochastic parent & child

Multigrid DA configurations

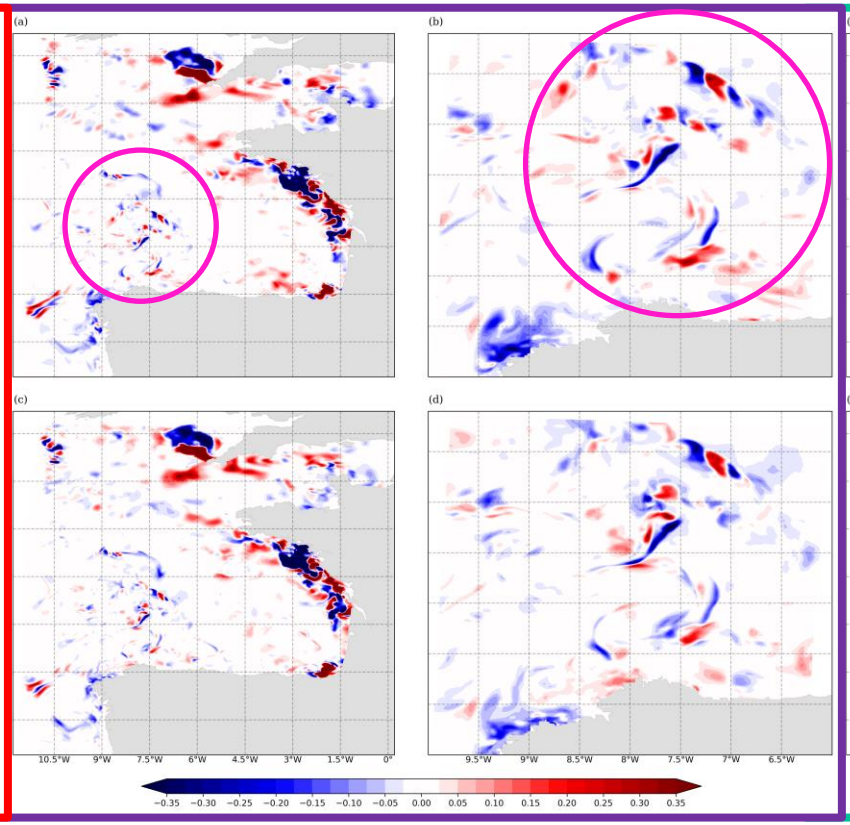
- We now test the impact of various DA options in the parent/child system, in TOP-2 class stochastic models only (TOP-2 = 2-way nesting with stochastic parent & child).
- In the DA runs, we use multigrid stochastic EnKF (stochastic models, Ensembles and covariances are multigrid).
- Quasi-reliable mode (HR SST, SLA maps).
- DA options tested:
  - **TOP-2.1**: Observations in parent only, covariances in parent only
  - **TOP-2.2**: Observations in parent only, multigrid covariances
  - **TOP-2.3**: Observations in child only, covariances in child only
  - **TOP-2.4**: Observations in child only, multigrid covariances
  - **TOP-2.5**: Observations in parent + child, multigrid covariances

Assimilate SST -- Members 601 and 602 on Jan 21, 2017

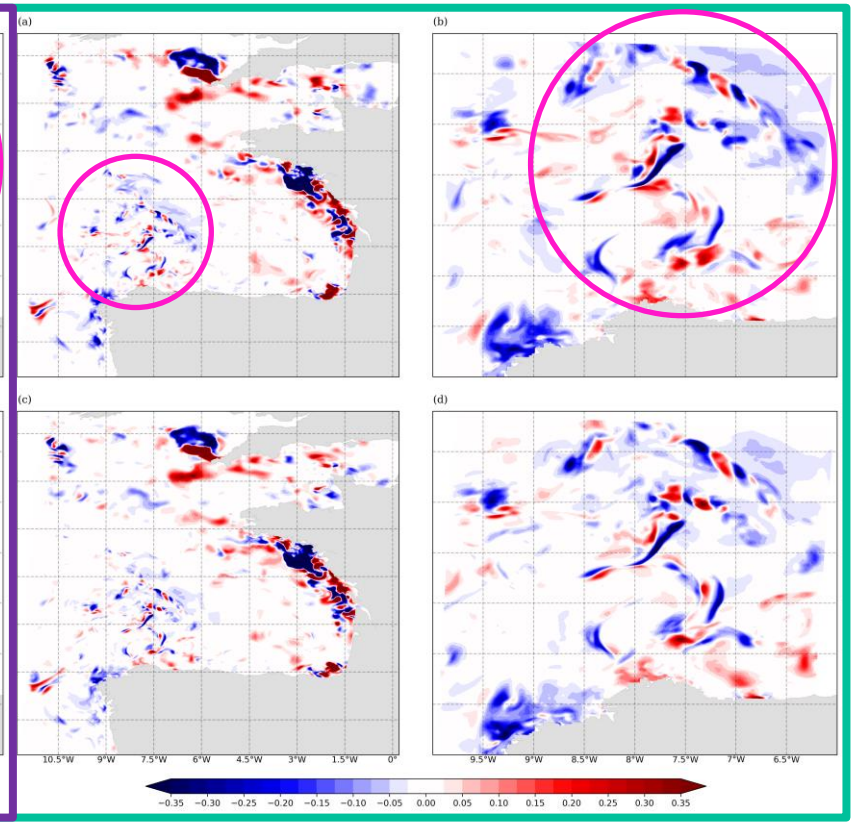
SST increments



TOP-2.1: Observations in parent only, covariances in parent only



TOP-2.2: Observations in parent only, covariances in parent + child



TOP-2.5: Observations in parent + child, covariances in parent + child

➔ When assimilating SST in both the parent and child, and using multigrid covariances, the gain is higher and the level of detail is better.

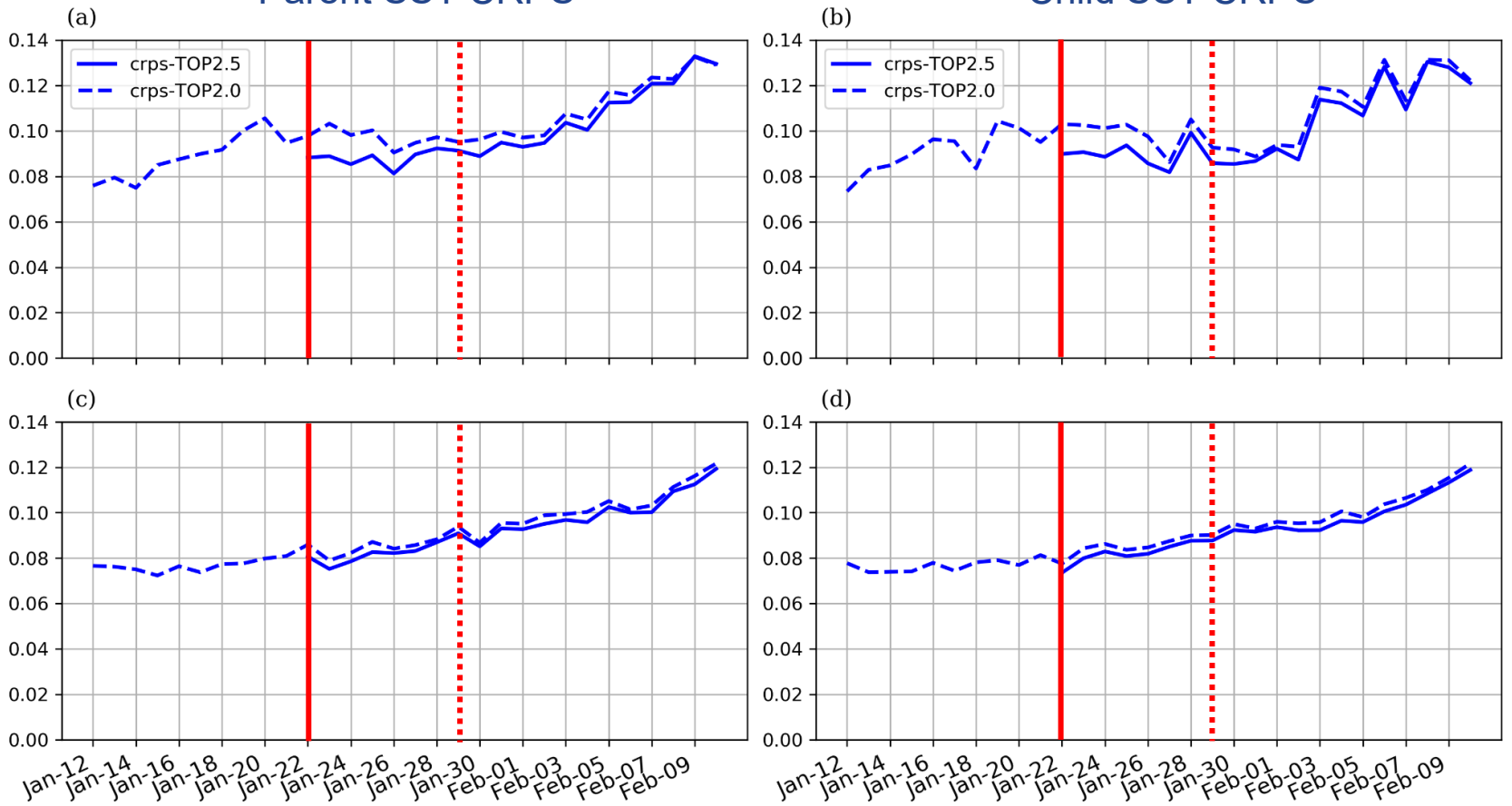
Dual-grid SST CRPS -- assimilated (SST) vs. free -- winter

Parent SST CRPS

Child SST CRPS

Open ocean

Coastal strip



Assimilation on Jan 21, 2017, then free run

QRTB

CRPS sampling error for 7 members ~.014

- ➔ SST skill enhanced when assimilating SST on both grids in 2-way coupled config, both for the child and the parent slightly above CRPS sampling error as of Richardson (2001).
- ➔ Open ocean: benefit seems to be retained to O(1wk+)
- ➔ Coastal strip: the benefit is not demonstrated.

TOP-2: 2-way nesting with stochastic parent & child

TOP-2.0: Free run

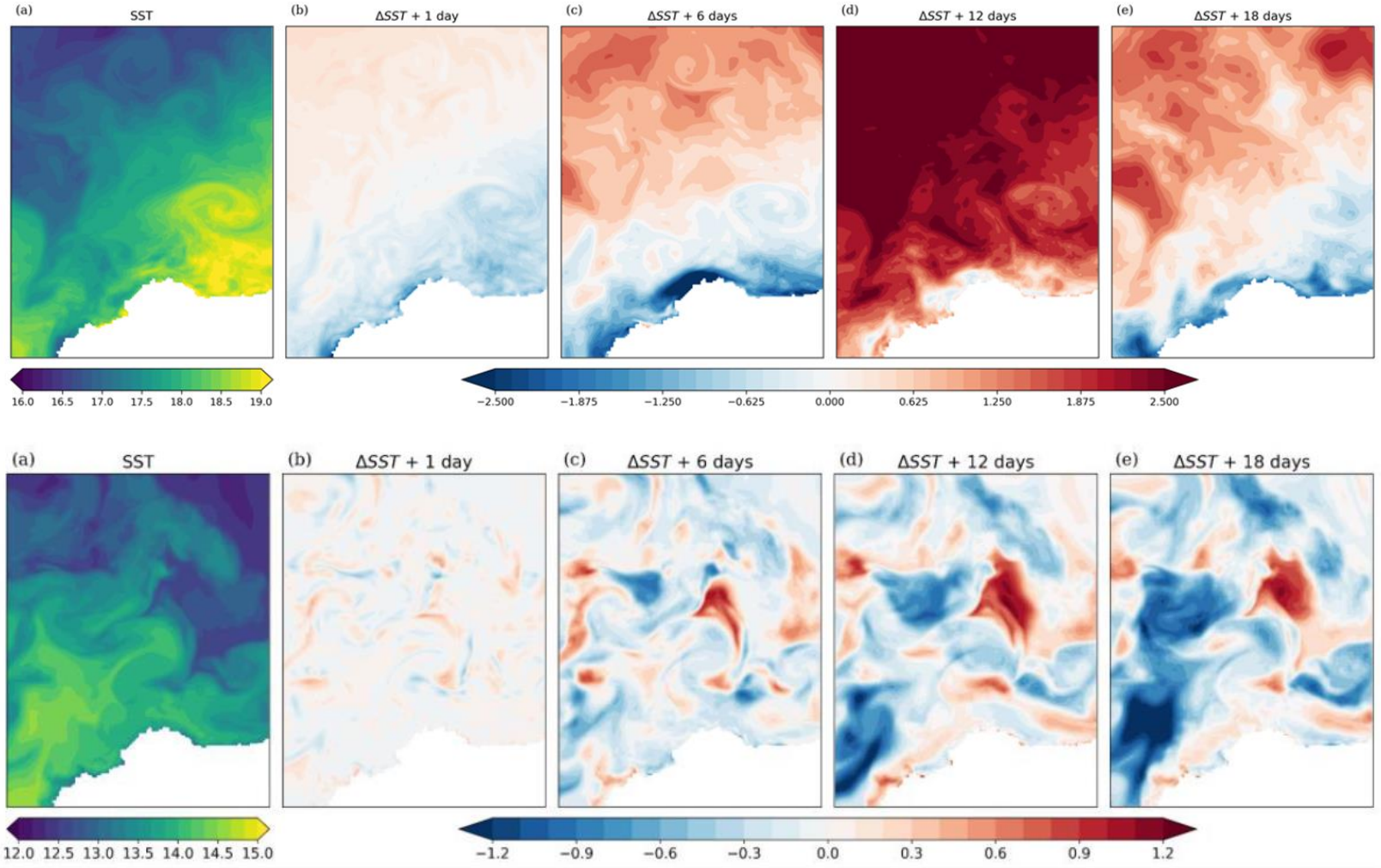
TOP-2.5: Assimilate in both grids

SST trends in TOP-2.5 over various ranges in the child domain

June

Ensemble mean TOP-2.5

Feb



➔ In June, SST trends predominantly atmosphere-driven (trade winds driving coastal upwelling)

➔ In Feb, SST trends driven by the sluggish IPC mesoscale field

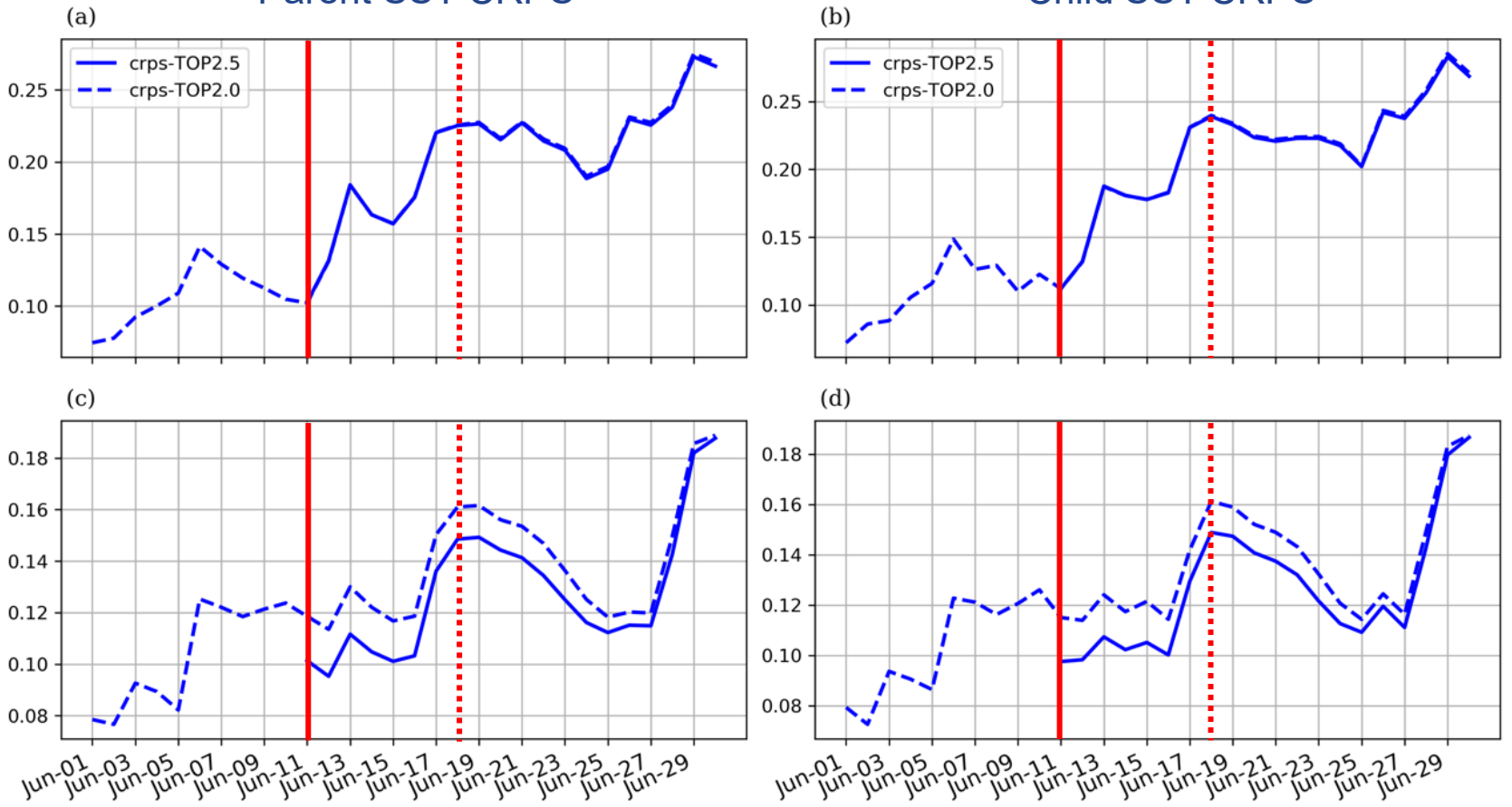
Dual-grid SST CRPS -- assimilated (SLA) vs. free -- summer

Parent SST CRPS

Child SST CRPS

Open ocean

Coastal strip



Assimilation on June 11, 2017, then free run

QRTB

CRPS sampling error for 7 members ~.014

➔ Expected lack of SST forecast skill enhancement in open ocean

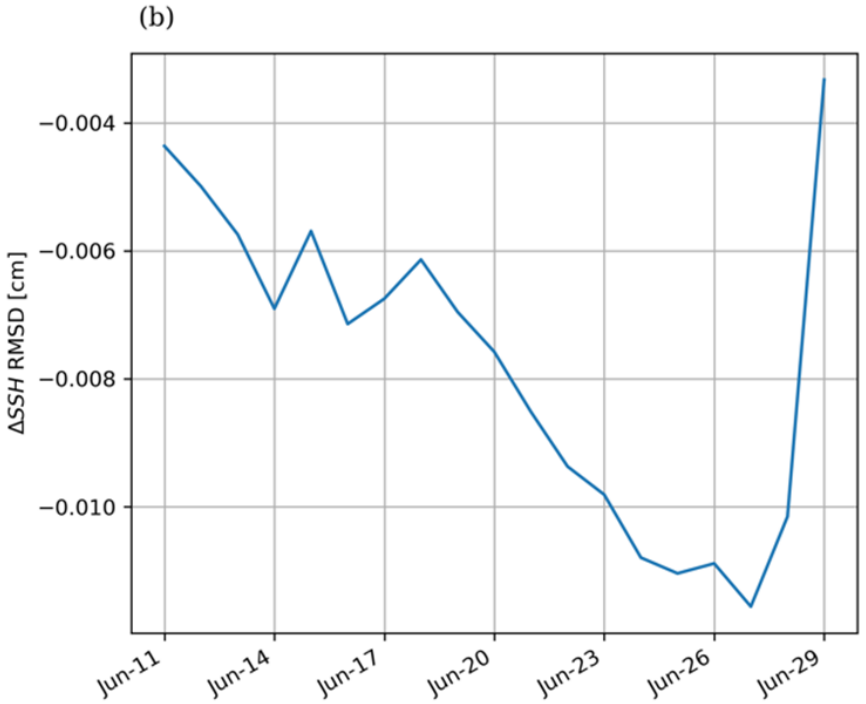
➔ Significant SST forecast skill enhancement in coastal strip (upwelling couples SST and SSH there)

TOP-2: 2-way nesting with stochastic parent & child

TOP-2.0: Free run

TOP-2.5: Assimilate in both grids

Child grid SLA trends -- assimilated(SLA) vs. free -- summer



TOP-2.5 SLA trend *rms* minus TOP-2.0 SLA trend *rms*
as a function of forecast range

➔ TOP-2.5 forecast skill vastly improved for SLA over TOP-2.0 up to O(2 weeks) in child grid

Conclusions so far on the MultiCast project

- Project finished in July, 2024. A paper is being written.
- We assessed the forecasting skill by (1) trends analysis, (2) being better than the climatological forecast (= CRPS Resolution term)
- The 2-way nested stochastic models with assimilation on both grids often led to forecast skill enhancement, in both the parent and child grids, with respect to other cases tested
  - Objective skill enhancement wrt. Richardson's criterion
  - Fcst skill enhancement dependent on observability of model errors (in particular when using surface observations), f(seasonal regime)
- The cross-grid covariances calculated from a 2-way nested Ensemble add definition and variance on top of single-grid covariances
- In a 2-way nested system where the parent assimilates, the benefit of the child also assimilating observations is felt for both the child and the parent.