living planet symposium BONN 2022

esa

# Towards probabilistic analyses and predictions of the Green Ocean using a stochastic NEMO-PISCES modelling system


Mikhail Popov[1], Jean-Michel Brankart[2], Pierre Brasseur[2], Arthur Capet[3], Emmanuel Cosme[1]

[1]IGE/UGA, Grenoble, France, [2]IGE/CNRS, Grenoble, France, [3]University of Liège, Belgium



In the frame of the **H2020 SEAMLESS project**, ensemble generation methods are being developed with the aim to improve the service through better data assimilation / inversion methods. A stochastic version of the NEMO-PISCES model has been developed and implemented in a global ocean configuration at $1/4°$ inherited from the CMEMS global Monitoring and Forecasting Center. The 40-member ensemble represents a probabilistic view of the 2019 seasonal cycle in the global and North Atlantic ocean. This ensemble is used to generate a posterior distribution of the BGC state variables and associated ecological indicators using a 4D, non Gaussian inversion scheme.

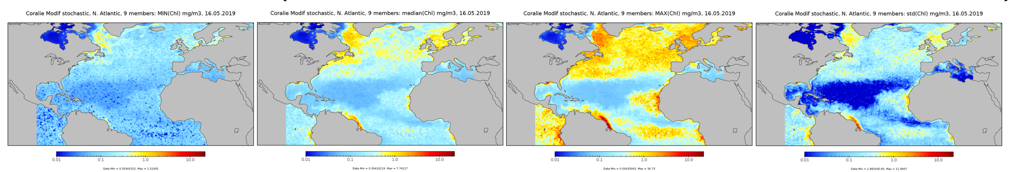

## Context / motivations / goals

- H2020 SEAMLESS general objective and motivation : provide CMEMS with **robust modelling/assimilation methods** to deliver useful **indicators** of climate-change impacts and food security in marine ecosystems.
- Among the blocking points : Many **CMEMS MFC products** describing ocean ecosystems and BGC currently **do not include robust uncertainty estimates**.
- IGE team goals : Explore **innovative inversion methods to unlock pitfalls of CMEMS operational systems**, with a focus on GLO/IBI MFC "Green Ocean" applications, through:
  - ✔ Transition **from deterministic to probabilistic ocean BGC modelling** based on stochastic parameterizations of uncertainty sources, and
  - ✔ Development of **ensemble-based inversion methods** dealing with non Gaussian pdfs to assimilate CMEMS L3 Ocean Colour data.

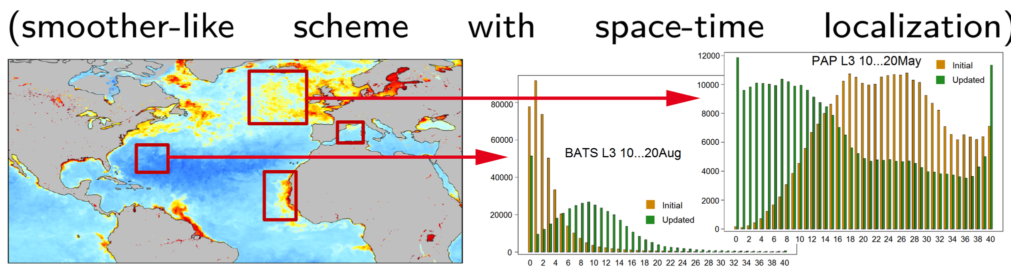
## Methodology

**Approach** : ***Decoupling between (i) prior pdf generation using full-complexity physical/BGC model, and (ii) Bayesian inversion step*** (*including local anamorphic transformations, Brankart et al., 2012*)

(i) Prior pdf : 2019 GLO NEMO-PISCES 40-member ensemble NEMO-PISCES based on stochastic perturbations, assuming uncertain bio parameters, mesoscale feature locations and subgrid-scale processes (Garnier et al., 2016; Leroux et al., 2022).



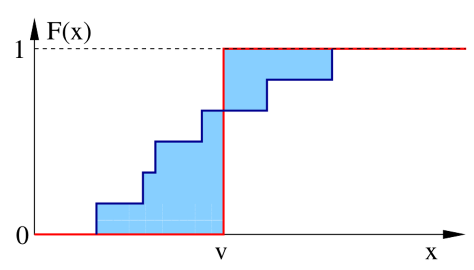
(ii) Posterior pdf : 4D multivariate regional inversions of L3 CMEMS OC data using LETKF/SEEK (smoother-like scheme with space-time localization).




## Metrics definitions

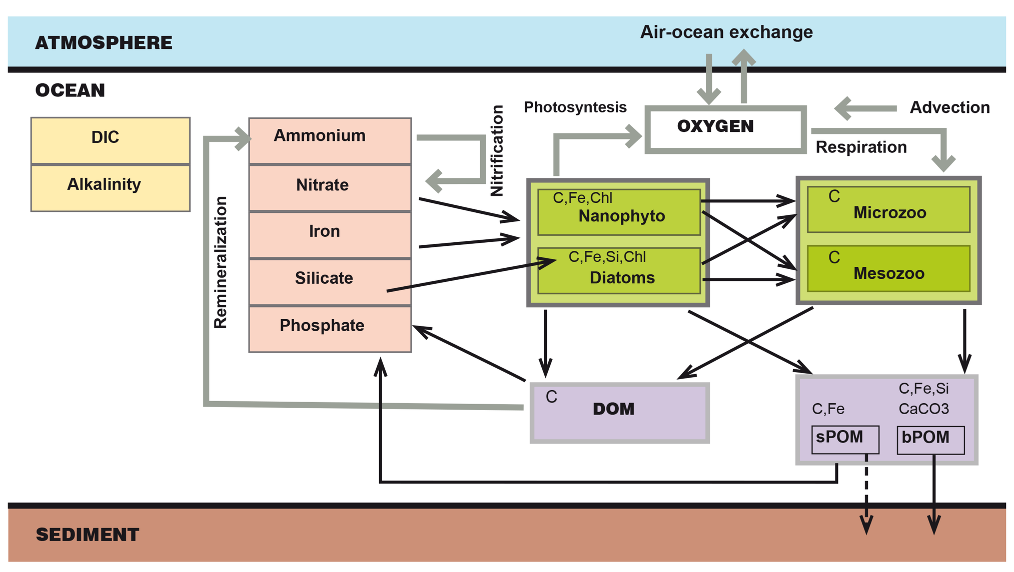
- Rank histograms: each observation is ranked relatively to its location within the sorted ensemble.
- Continuous Rank Probability Score (CRPS): misfit between stepwise probability distribution of a variable and Heaviside function, increasing by 1 at the true value of the variable.

$$CRPS(v) = \int_{-\infty}^{+\infty} |F(x) - \mathbb{1}(v-x)|dx$$




## NEMO-PISCES global configuration

- NEMO 4.0-r13720
- $1/4°$ resolution & 75 $z$-levels
- Initial conditions : MERCATOR GLORYS2V4 (Global Ocean Eddy-Permitting Physical Reanalysis)
- Forcing : ERA5 dataset (only surface level is used)
- Biogeochemical model PISCES (Aumont et al., 2015): 24 variables
- Initial date = 01.01.2017; spin-up = 2 years; target year = 2019




## Stochastic parameterization of uncertainty sources

1. Parameter uncertainties

$$\left.\frac{\partial C}{\partial t}\right|_{bio} = \text{SMS}\,(C, u, p \cdot \exp\left[\xi(t)\right], t)$$

$$p' = p \cdot \exp\left[\xi(t)\right] \sim \log\mathcal{N}(\mu = 0, \sigma = 0.4) \approx \mathcal{N}(\mu = 1, \sigma = 0.4)$$

Autoregressive processes: $\xi(t_{n+1}) = a\,\xi(t_n) + bw + c$

- $\alpha^N$ – photosynthetic efficiency of nanophytoplankton ;
- $\alpha^D$ – the same for diatoms;
- $\mu^0_{max}$ – growth rate of nanophytoplankton at $0°$C;
- $b$ – temperature sensitivity of phytoplankton growth;
- $b'_Z$ – temperature sensitivity of grazing by zoo;
- $f^N_{day}$ – day length dependence for nanophytoplankton;
- $f^D_{day}$ – the same for diatoms.

2. Unresolved scales

$$\left.\frac{\partial C}{\partial t}\right|_{bio} = \frac{1}{2}\left[\text{SMS}\,(C + C\xi(t), u, p, t) + \text{SMS}\,(C - C\xi(t), u, p, t)\right]$$

$C\xi(t) \equiv \delta C(t)$ – fluctuations, not resolved by the mesh.
Stochastic processes $\xi(t)$ with $\sigma = 0.2$ are applied to 20 of 24 passive tracers (except of Dissolved Inorganic Concentration (DIC), Alkalinity, $O_2$ and $CaCO_3$ concentrations).

3. Location uncertainties

$$\Delta x_i(t) = \Delta x_i^0\left[1 + \xi_i(t)\right] \qquad \xi_i(t + \Delta t) = a\,\xi_i(t) + b\,w$$

$a$, $b$ define the standard deviation and the correlation length.

## 4D inversion

- Analysis stage of the Ensemble Kalman Filter (EnKF)

$$\mathbf{x}_m^{pos} = \mathbf{x}_m + \hat{K}\left(\mathbf{y}^o - \hat{H}\mathbf{x}_m\right) \qquad \hat{K} = P\,\hat{H}^T\left(\hat{H}\,P\,\hat{H}^T + R\right)^{-1}$$

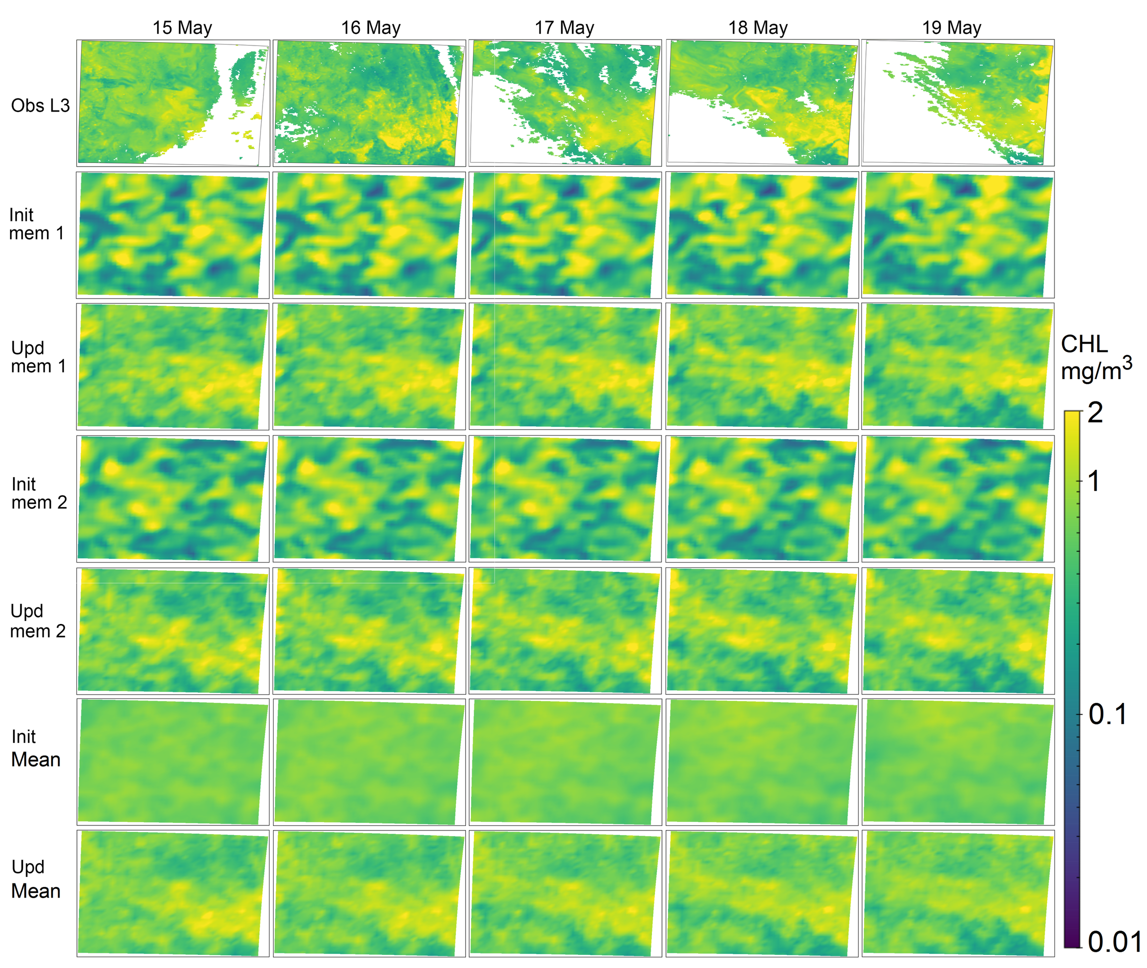
$P$ – ensemble covariance matrix; $R$ – error covariance matrix.

- Anamorphosis transformations $x \to \eta \in \mathcal{N}(0,1)$

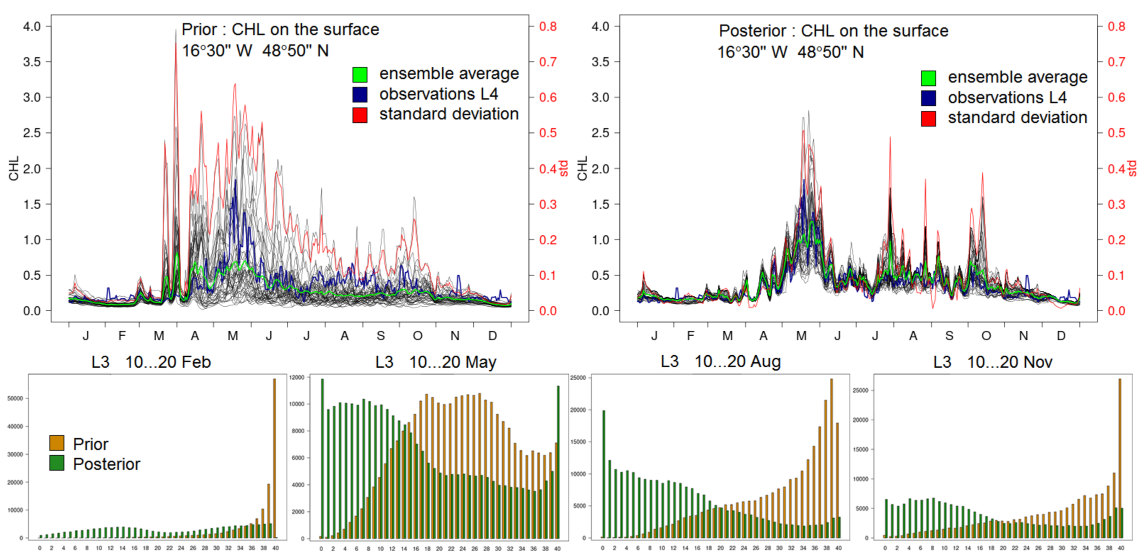
$$\Psi(x) = \begin{cases} \eta_1\,, & x < x_1\,, \\ \eta_k + \frac{\eta_{k+1} - \eta_k}{x_{k+1} - x_k}(x - x_k)\,, & x_k \le x \le x_{k+1}\,, \\ \eta_s\,, & x > x_s\,. \end{cases}$$

- Domain localization
  - ✔ In space : length scale $\sim 40$ km in horizontal plane, no localization in vertical direction;
  - ✔ In time : time scale $\sim 10$ days.

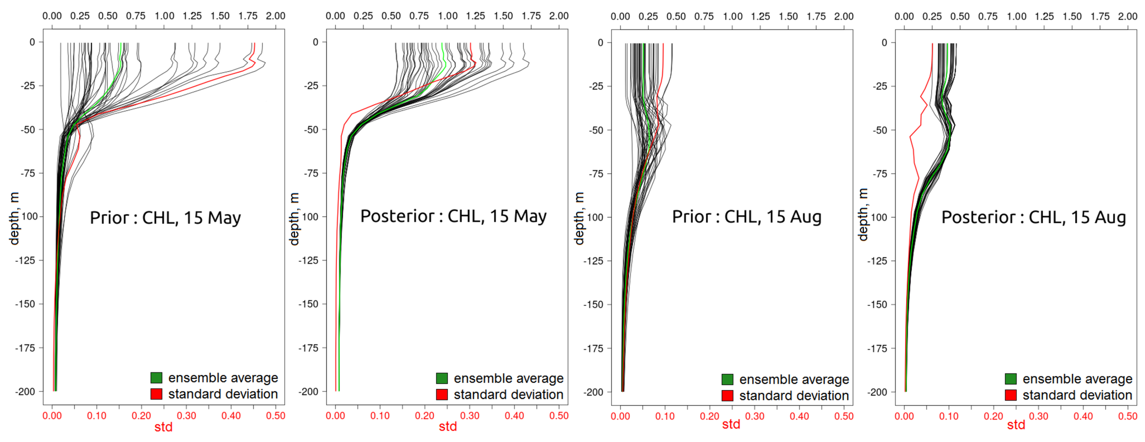
## Surface chlorophyll concentration in PAP region (1100 km×720 km centered on 16°30' W, 48°50' N)



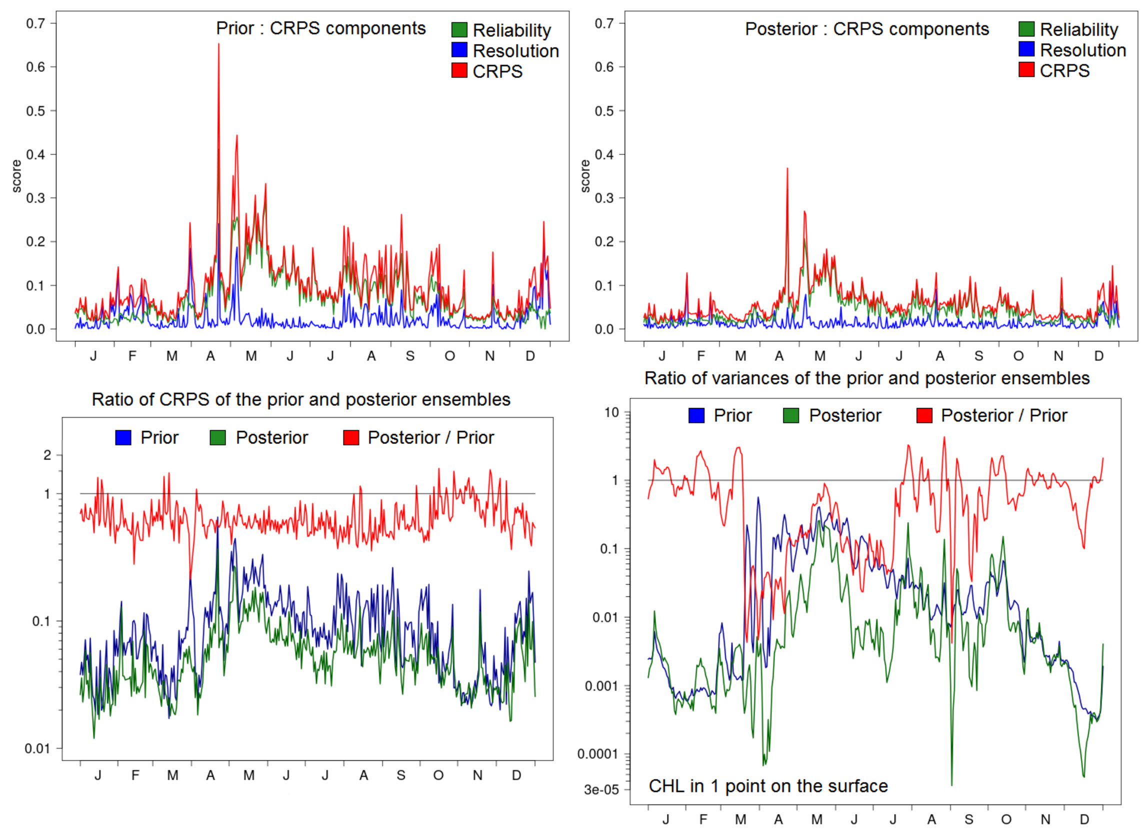

## Time series and rank histograms



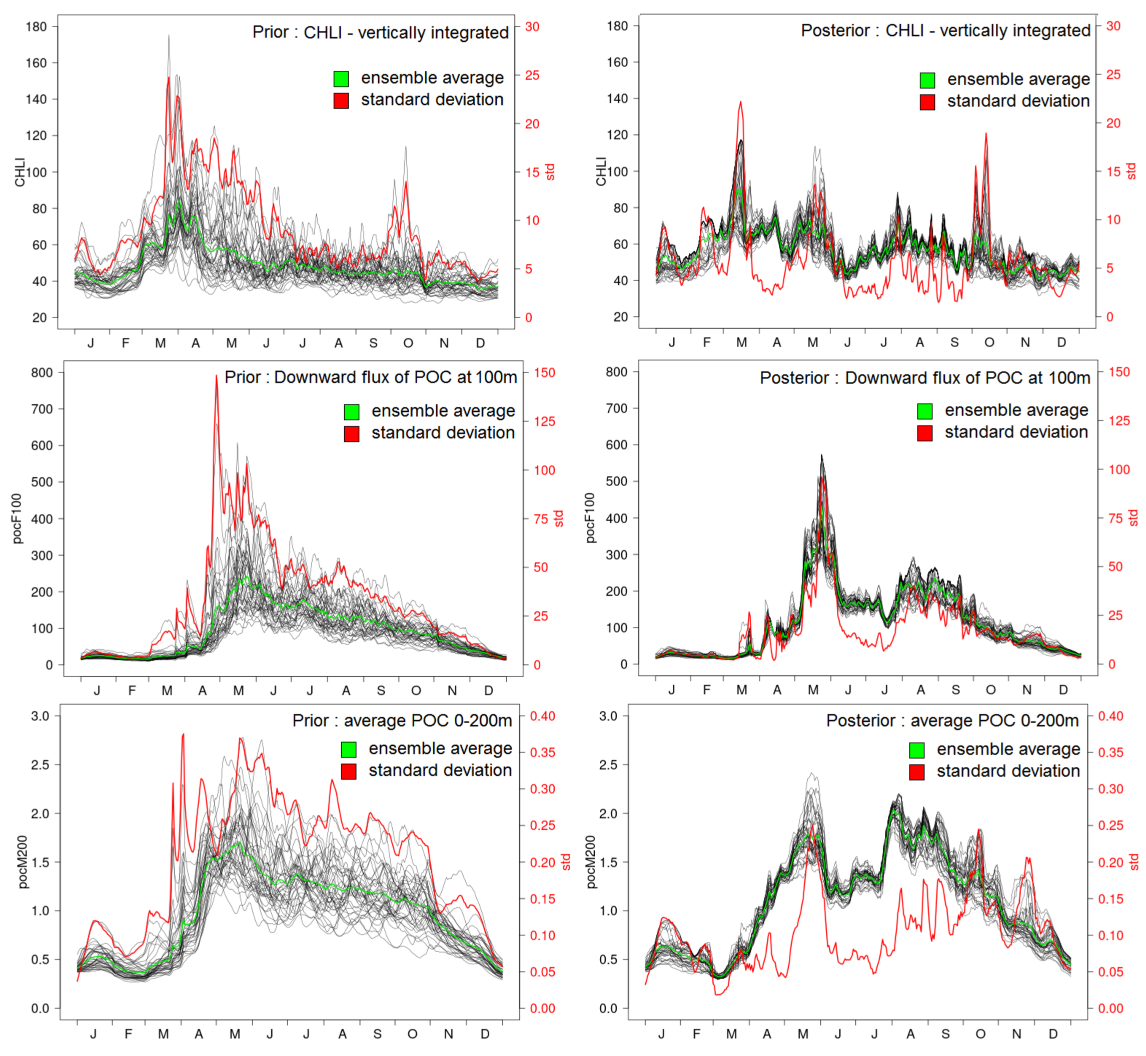

## Projection of surface OC on the vertical




## CRPS and ensemble variance at 16°30' W, 48°50' N




## Uncertainty reduction for selected indicators




## Conclusions...

- A **new 4D space-time scheme** has been developed as a natural extension to sequential ensemble analysis/forecast in place today (such as LETKF) in CMEMS MFCs.
- **Controlability of key indicators** (POC, NPP, trophic efficiency) is demonstrated in PAP region, except for specific time periods. Other results (not shown here) suggest lower performance in BATS region.
- Accounting of additional (or **revising assumptions** about the) uncertainty sources in models and assimilated data is **part of the process**.
- The overall approach provides a **methodology to help decide whether to faithfully catalog a new product** with objective added value to users and scientists.

## ...and perspectives

- ✔ Ongoing : exploration of the skill of the method for **probabilistic forecasts** (and associated predictability time scales).
- ✔ Next step : **joint inversion of satellite ocean color and altimetric data**, bringing additional constraints and further reduction of uncertainties on estimated quantities.
- ✔ **Sensitivity to observation error statistics** needs further investigation.


#LPS22 This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 776480


SEAMLESS IGE UGA Université Grenoble Alpes cnrs LIÈGE université