A neural network approach to estimate water-column nutrient concentrations and carbonate system variables in the Mediterranean Sea: **CANYON-MED**

Marine Fourrier[1], Laurent Coppola[1], Fabrizio D'Ortenzio[1], Hervé Claustre[1]

(1) Sorbonne Université, CNRS, Laboratoire d'Océanographie de Villefranche, LOV, F-06230 Villefranche-sur-Mer, France.

Contact:
fourrier@obs-vlfr.fr

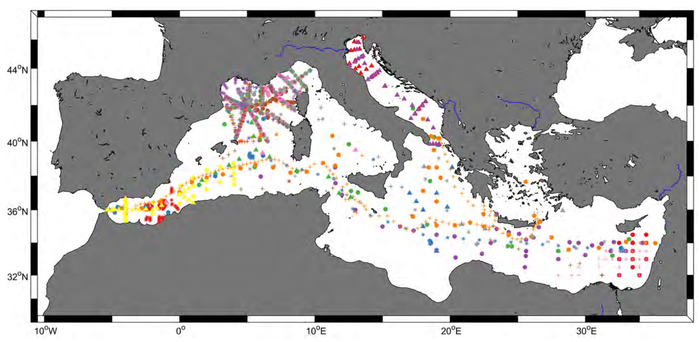
1– Introduction

The Mediterranean Sea (MedSea) is characterized by a short residence time[1], an elevated **salinity** and is one of the largest oceanic nutrient-depleted areas[2] exhibiting an eastward-increasing **oligotrophy** gradient[3].

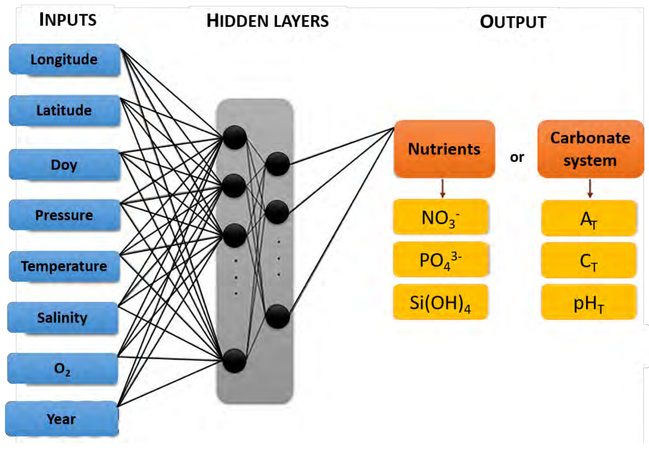
In the context of a critically undersampled ocean, the development and intensive use of instrumented *in situ* **autonomous platforms** allows to densify the measurements of some biogeochemical variables (remaining however far from exhaustive).

2– CANYON-MED method

We compiled a **homogeneous quality-controlled** dataset of *in situ* measurements from 1981 to 2018, including samples of core variables such as dissolved oxygen (O_2), temperature (T), salinity (S) as well as **macronutrients** (nitrates: NO_3^-, phosphates: PO_4^{3-}, silicates: $Si(OH)_4$) and **carbonate system variables** (total alkalinity: A_T, total carbon: C_T, and *in situ* pH on the total scale: pH_T).



Map of the 35 cruises and the time-series used for the training and validation of the CANYON-MED neural networks. (~8 000 values for NO_3^-, ~11 000 for PO_4^{3-} and $Si(OH)_4$; and ~6 000 for A_T and pH_T, ~4 000 values for C_T).



Schematic representation of the CANYON-MED neural networks. Doy: day of year

Adapted from [4, 5]

The **CANYON-MED** (for Carbonate system and Nutrients concentration from hYdrological properties and Oxygen using a Neural-network in the MEDiterranean Sea; in prep) **neural network-based** method provides estimations of **nutrients** and **carbonate system** variables from systematically measured oceanographic variables (T, S, O_2, geolocation, sampling date), such as the ones measured by profiling floats.

Statistics (coefficient of determination (R^2), Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE)) of the regressions between the CANYON-MED retrieved values with the corresponding in situ *measurements.*

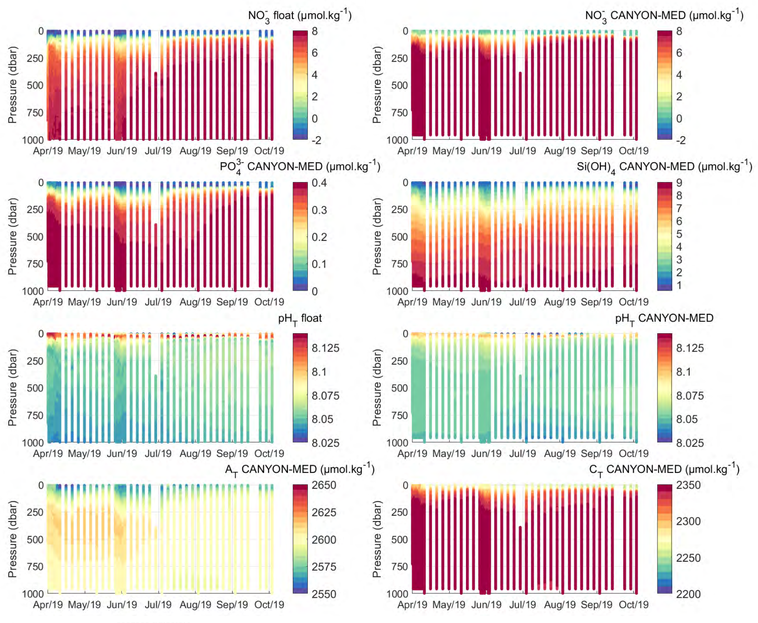
	CANYON-MED		
	R^2	MAE	RMSE
NO_3^- ($\mu mol.kg^{-1}$)	0.95	0.47	**0.77**
PO_4^{3-} ($\mu mol.kg^{-1}$)	0.91	0.028	**0.047**
$Si(OH)_4$ ($\mu mol.kg^{-1}$)	0.96	0.40	**0.70**
A_T ($\mu mol.kg^{-1}$)	0.95	7	**11**
C_T ($\mu mol.kg^{-1}$)	0.89	9	**14**
pH_T	0.86	0.010	**0.016**

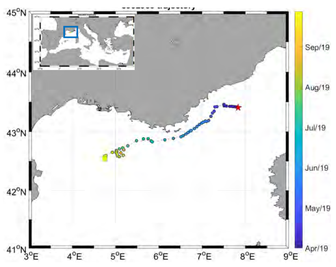
In order to better understand nutrients and pH dynamics, we test CANYON-MED's predictions on **3 BGC-Argo floats** equipped with **nitrates** and **pH** sensors, in contrasted areas of the **MedSea** (North-Western Mediterranean Sea, Adriatic and Ionian Seas, Levantine Basin), along the W-E oligotrophy gradient.

3– Application on MedSea BGC-Argo floats

For each panel: Timeseries of NO_3^- float data, CANYON-MED NO_3^-, CANYON-MED PO_4^{3-}, CANYON-MED $Si(OH)_4$, pH_T float data, CANYON-MED pH_T, CANYON-MED A_T and CANYON-MED C_T. Location of the BGC-Argo float, the red star and yellow square representing the location of the 1st and last profiles.

NW Mediterranean **6902969**



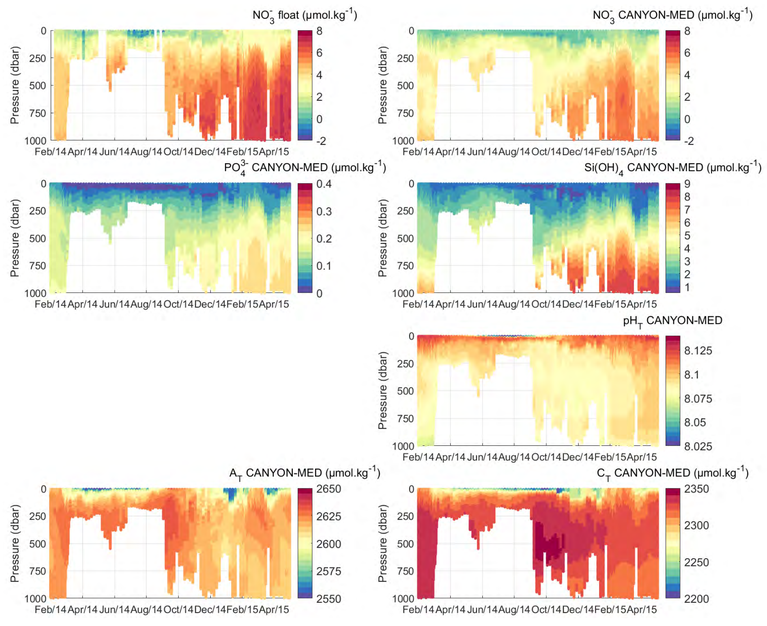


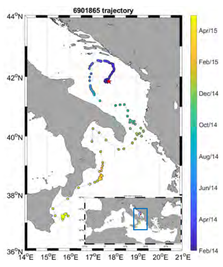
Good agreement between float and CANYON-MED NO_3^-.

For pH_T, small difference in surface values. PO_4^{3-}, $Si(OH)_4$, A_T and C_T coherent with bottle data.

→ Dynamics well reproduced.

Adriatic & Ionian **6901865**



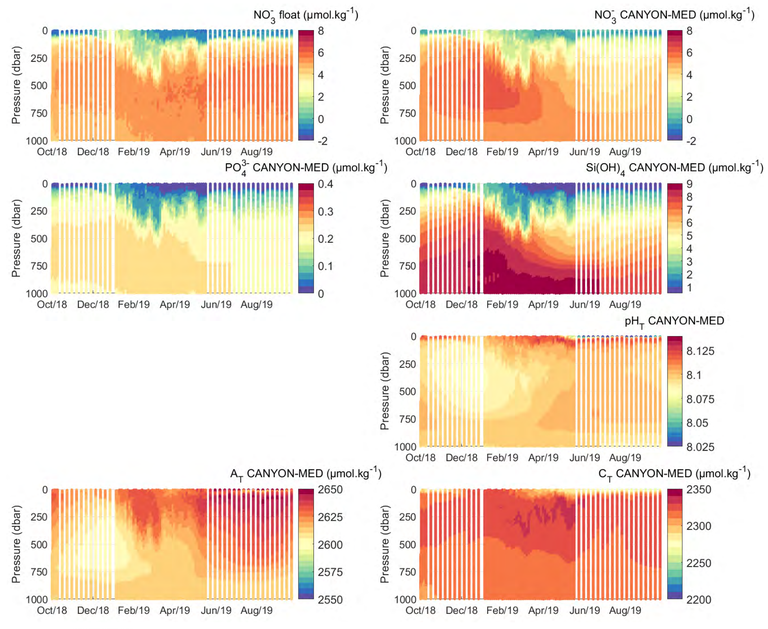


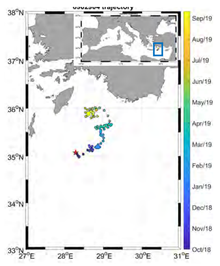
Good agreement between float and CANYON-MED NO_3^-.

pH_T, PO_4^{3-}, $Si(OH)_4$, A_T and C_T coherent with bottle data.

→ Satisfying results in relatively coastal waters!

Levantine basin **6902904**





Similar tendencies between float and CANYON-MED NO_3^-.

pH_T, PO_4^{3-}, $Si(OH)_4$, A_T and C_T in the same range as bottle data.

→ Promising results in under-represented area of CANYON-MED's training.

4 – Conclusions and Perspectives

The regional CANYON-MED neural networks produce satisfactory results: accuracies of **0.77**, **0.047**, and **0.70 µmol.kg^{-1}** for **NO_3^-**, **PO_4^{3-}** and **$Si(OH)_4$** respectively, and of **0.016**, **11 µmol.kg^{-1}** and **14 µmol.kg^{-1}** for **pH_T**, **A_T** and **C_T** respectively, comparable to the ones obtained by the NO_3^- (±1 µmol.kg^{-1}) and pH_T (±0.005) sensors mounted on BGC-Argo floats. Applied to the large Argo network, CANYON-MED will contribute to **Quality Control** and **Delayed Mode** processing for NO_3^- and pH_T as well as to increase the amount of biogeochemical data in the MedSea, dramatically improving our understanding of nutrients, pH and pCO_2 variability. Furthermore, CANYON-MED will generate "virtual" data of parameters not yet measured by BGC-Argo floats (namely PO_4^{3-}, $Si(OH)_4$, A_T, and C_T).

References: /**1**/ Somot, S., Sevault, F., & Déqué, M. (2006). Transient climate change scenario simulation of the Mediterranean Sea for the twenty-first century using a high-resolution ocean circulation model. Climate Dynamics, 27(7–8), 851–879. DOI: 10.1007/s00382-006-0167-z /**2**/ Ignatiades, L. (2005). Scaling the trophic status of the Aegean Sea, eastern Mediterranean. Journal of Sea Research, 54(1), 51–57. DOI: 10.1016/j.seares.2005.02.010 /**3**/ Pujo-Pay, M., Conan, P., Oriol, L., Cornet-Barthaux, V., Falco, C., Ghiglione, J.-F., et al. (2011). Integrated survey of elemental stoichiometry (C, N, P) from the western to eastern Mediterranean Sea. Biogeosciences, 8(4), 883–899. DOI: 10.5194/bg-8-883-2011 /**4**/ Sauzède, R., Bittig, H. C., Claustre, H., Pasqueron de Fommervault, O., Gattuso, J.-P., Legendre, L., et al. (2017). Estimates of Water-Column Nutrient Concentrations and Carbonate System Parameters in the Global Ocean: A Novel Approach Based on Neural Networks. Frontiers in Marine Science 4. DOI: 10.3389/fmars.2017.00128 /**5**/ Bittig, H. C., Steinhoff, T., Claustre, H., Fiedler, B., Williams, N. L., Sauzède, R., et al. (2018). An Alternative to Static Climatologies: Robust Estimation of Open Ocean CO2 Variables and Nutrient Concentrations From T, S, and O2 Data Using Bayesian Neural Networks. Frontiers in Marine Science 5. DOI: 10.3389/fmars.2018.00328.

Acknowledgments:
This work was supported by the E-ARISE H2020, WG4.