Carbon dioxide fluxes estimation merging satellite and in-situ data in the Mediterranean Sea (WP-2640*)

Mattia Pecci[1,2], Alcide Giorgio di Sarra[1], and collaborators

[1]ENEA, Italy; [2]Sapienza Univ. of Rome, Italy

IDEAS-QA4EO Cal/Val Workshop#5

11th- 13th June 2024 – Thessaloniki (Greece)

Outline:

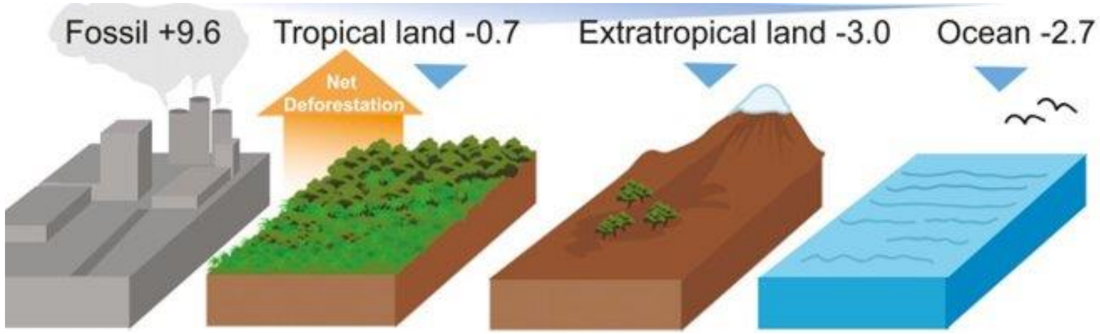
1. Carbon cycle
2. Objectives
3. In situ pCO_2 and CO_2 fluxes
4. Satellite algorithms
5. Results and discussion
6. Conclusions and next steps

Outline:

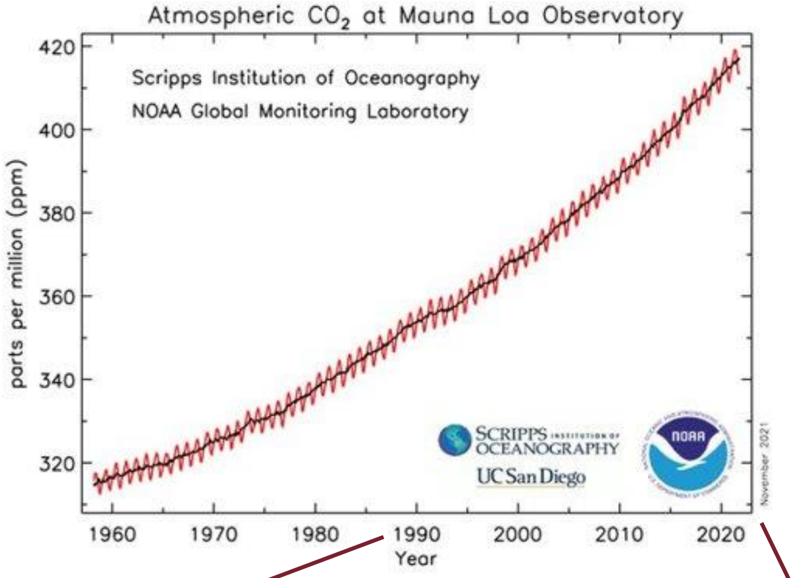
1. **Carbon cycle**
2. Objectives
3. In situ pCO_2 and CO_2 fluxes
4. Satellite algorithms
5. Results and discussion
6. Conclusions and next steps

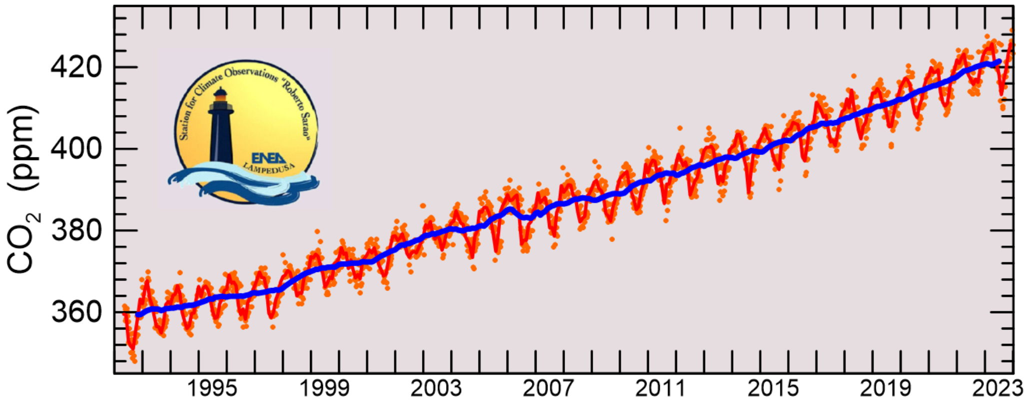
Carbon cycle – brief overview

Atmospheric CO_2 concentration increased from about 280 ppm in pre-industrial times to the actual value of about 420 ppm due to antropogenic activities



Atmospheric increase 4.7 Pg/y





Sellers, P. J. et al., Observing carbon cycle–climate feedbacks from space. *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences. https://doi.org/10.1073pnas.1716613115, 2018

Carbon cycle – brief overview

- Ocean interaction with CO_2 has a great spatio-temporal variability not fully characterised with complex dependencies on phisycal, biological and chemical properties of the ocean
- CO_2 absorption leads to the acidification of ocean waters which can trigger negative feedbacks on absorption efficency
- Climate feedbacks are unknown
- Lack of continuous in situ measurements
- Ocean CO_2 absorption efficiency is strongly related with climate evolution

Monitoring atmosphere-ocean exchanges is crucial

Carbon cycle – ocean-atmosphere fluxes

$$\boxed{\mathbf{F = K_{wa}\ KH\ (\Delta pCO_2)_{sea\text{-}atm}}}$$

- $K_{wa} = 0.251 <U^2> (Sc/660)^{-0.5}$ is the Gas Transfer Velocity for $U < 15$ m/s

- $Sc = A + B*SST + C*SST^2 + D*SST^3 + E*SST^4$ is the Schmidt Number

- $\ln(KH) = A_1 + A_2*(100/SST) + A_3*\ln(SST/100) + SSS*[B_1 + B_2*(SST/100) + B_3*(SST/100)^2]$ is the gas solubility

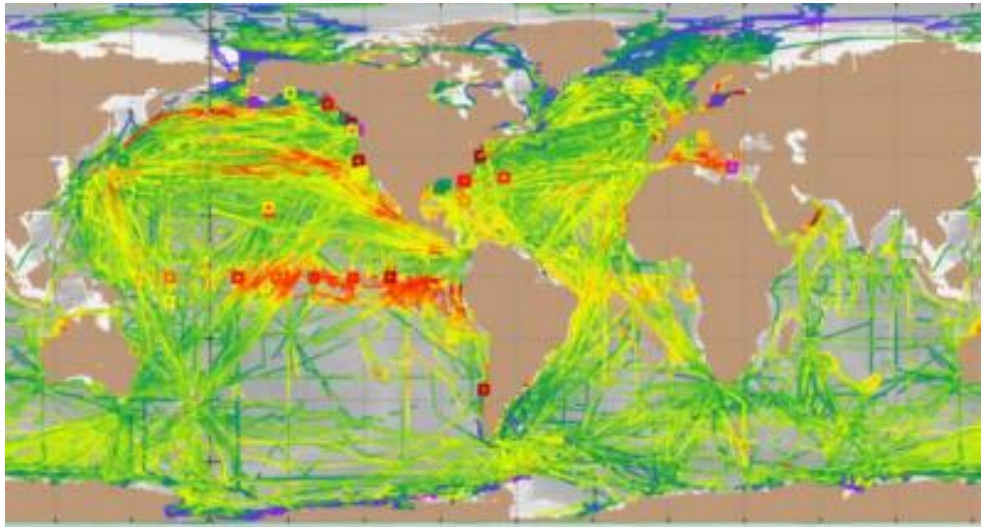
- Sea pCO_2 can be measured or derived

- Air pCO_2 can be measured or derived

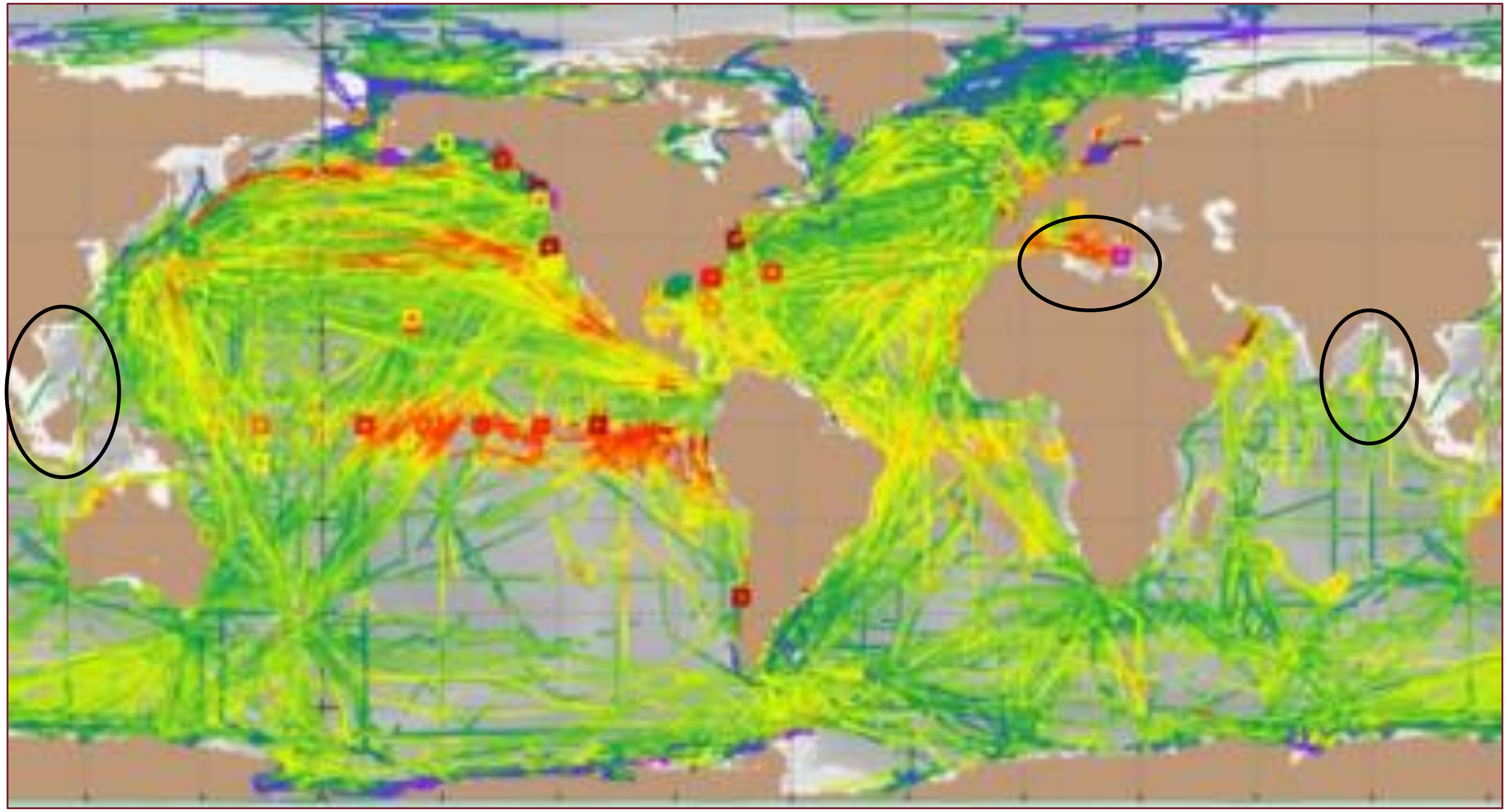
Wanninkhof, R., Relationship between wind speed and gas exchange over the ocean revisited, *Limnol. Oceanogr. Methods*, 12, doi:10.4319/lom.2014.12.351, 2014

Marine carbon cycle

- Current monitoring mainly rely on model-based estimates of pCO_2 and CO_2 fluxes
- Works on satellite-based estimates of pCO_2
- Sparse in situ continuos and naval occasional measurements
- Carbon global monitoring projects and datasets:
    - Global carbon budget (https://www.globalcarbonproject.org/)
    - Surface Ocean CO_2 atlas (https://socat.info/)



Marine carbon cycle



Still missing estimates, including large portion of marginal seas

K. Lee, C.L. Sabine, T. Tanhua, T.W. Kim, R.A. Feely, and H.C. Kim. Roles of marginal seas in absorbing and storing fossil fuel CO2. Energy & Environmental Science, 4:1133–1146, 2011. doi: https://doi.org/10.1039/C0EE00663G.

Outline:

1. Carbon cycle

2. **Objectives**

3. In situ pCO_2 and CO_2 fluxes

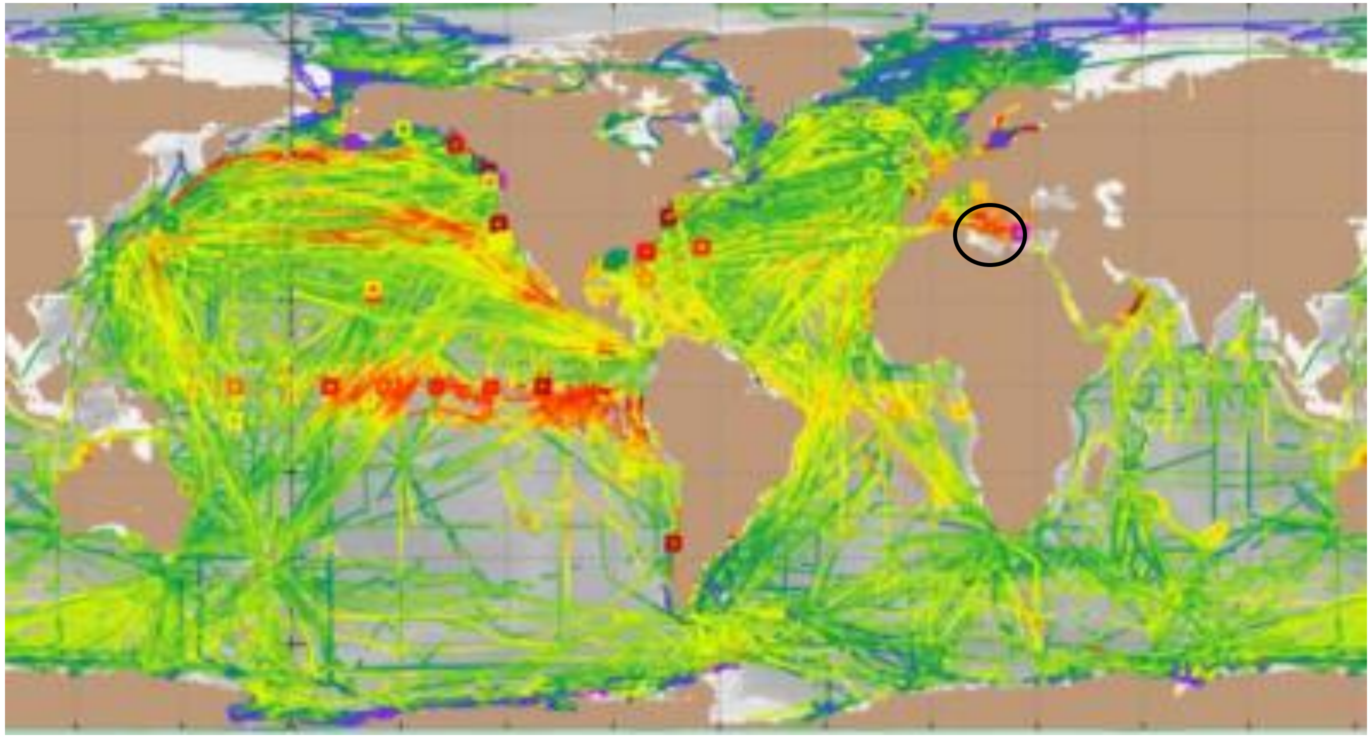
4. Satellite algorithms

5. Results and discussion

6. Conclusions and next steps

WP-2640* objectives

- Characterization of the Central Mediterranean carbon cycle using in situ data
- Satellite pCO_2 estimates using proxies for spatial monitoring
- CO_2 fluxes estimates merging satellite, model and in situ data



Outline:

1. Carbon cycle

2. Objectives

3. **In situ pCO_2 and CO_2 fluxes**

4. Satellite algorithms

5. Results and discussion

6. Conclusions and next steps

Mediterranean Sea

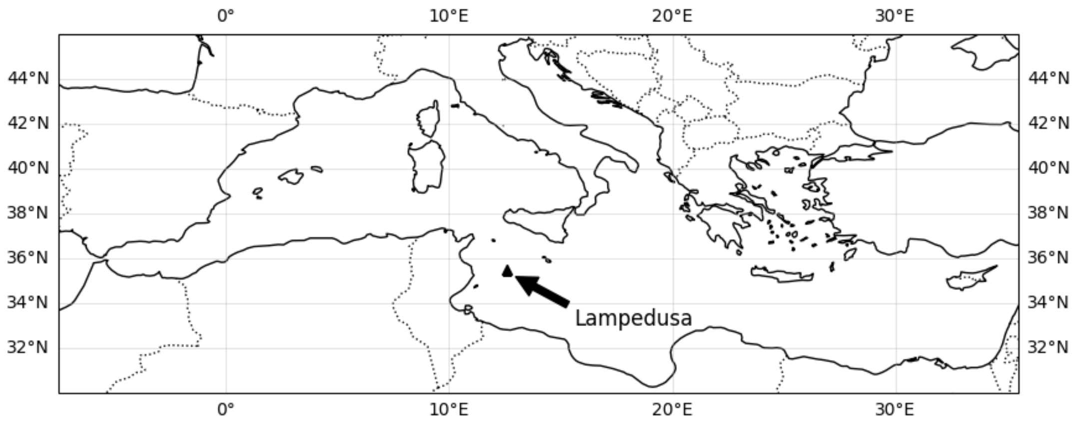
- Climate hotspot
- Semi-enclosed basin under environmental stress
- Few carbon in situ measurements
- Few studies on basin-wide carbon cycle
- Suffering long and intense marine heatwaves (2022-2023; 2023-2024)

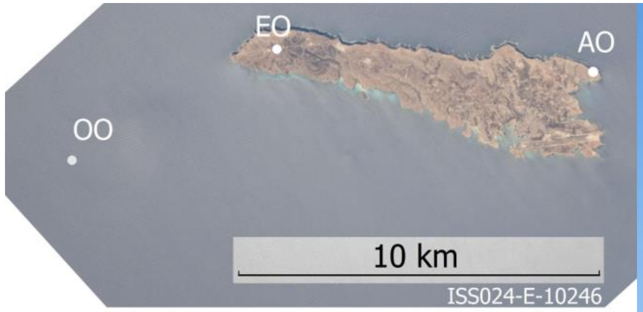
S. Marullo, V. De Toma, A. di Sarra, R. Iacono, A. Landolfi, F. Leonelli, E. Napolitano, D. Meloni, E. Organelli, A. Pisano, R. Santoleri, and D. Sferlazzo. Has the frequency of Mediterranean Marine Heatwaves really increased in the last decades? In EGU General Assembly 2023, Vienna, Austria, 2023. URL https://doi.org/10.5194/ egusphere-egu23-4429.

Study site

In situ measurements made at **Lampedusa**

- Small island with small pollution sources
- Far from land
- Representative for background conditions
- Host observatories (AO, OO, EO) for climate studies and carbon monitoring (within the Integrated Carbon Observation System - **ICOS infrastructure**)





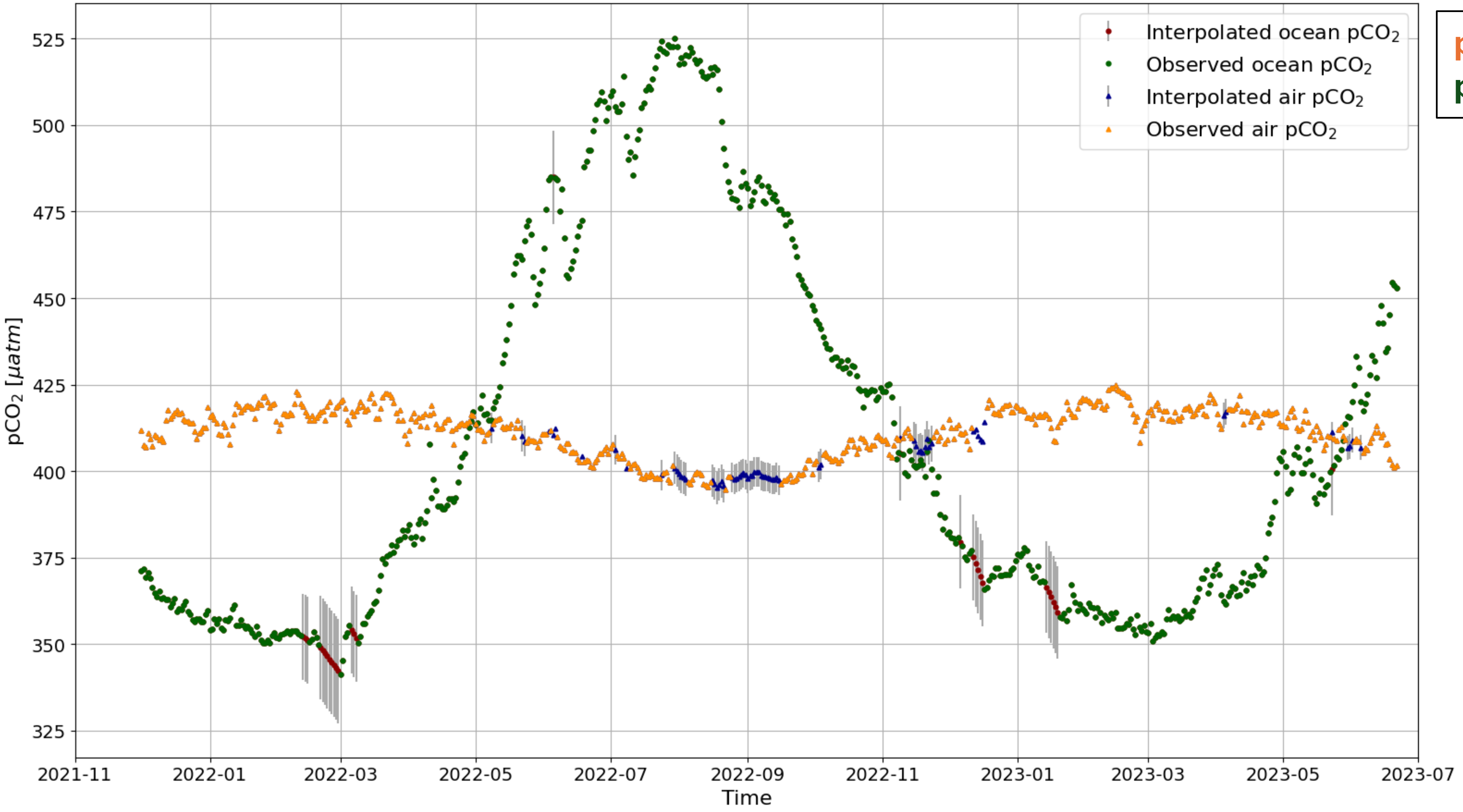
- Ocean pCO_2, temperature and salinity are available at 5 m depth from October 2021 at OO
- Wind speed, atmospheric pressure, atmospheric CO_2 concentration are available at AO

Study site

Variable	Instrument	Accuracy	Height (asl)
Sea pCO_2	ProOceanus CO_2 Pro-CV	**± 3 ppm**	-5m (OO)
SST	CTD SBE16+	**±0.005°C**	-5m (OO)
SSS	CTD SBE16+	**±0.01 PSU**	-5m (OO)
Wind speed	Gill windsonic sensor	**±2%**	10m (OO)
Wind speed	Vaisala WS425	**±3%**	60m (AO)
Atm. pressure	Vaisala BARO-1	**±0.25 hPa**	52m (AO)
Atm. CO_2 conc.	Picarro G2401	**±0.1 ppm**	57m (AO)

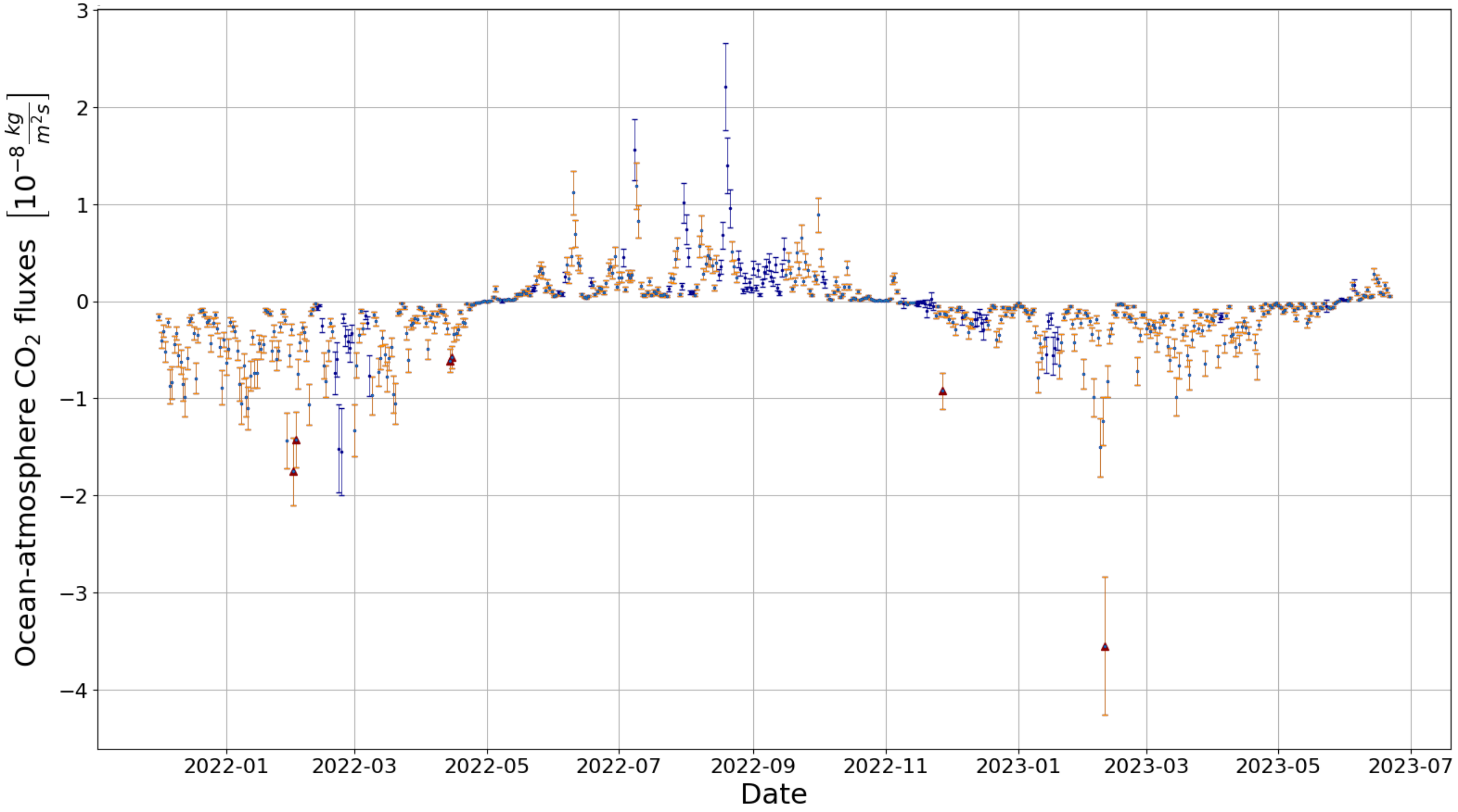
pCO_2 issue with the a/d zero measurements between March and July 2022. An empirical correction was applied with an increased associated uncertainty.

In situ pCO_2 and CO_2 fluxes

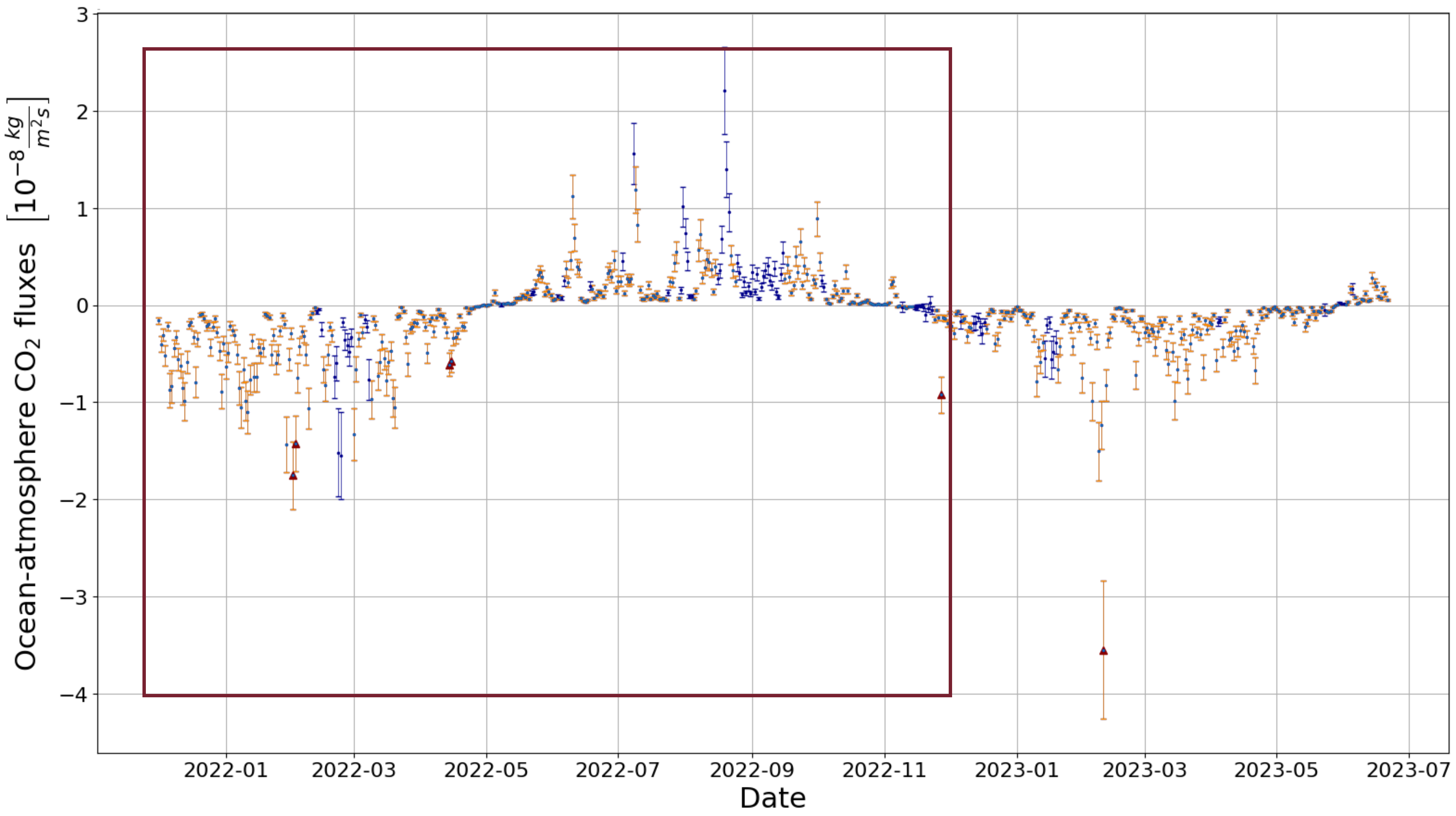


pCO_2 ATM
pCO_2 OCE

In situ pCO_2 and CO_2 fluxes

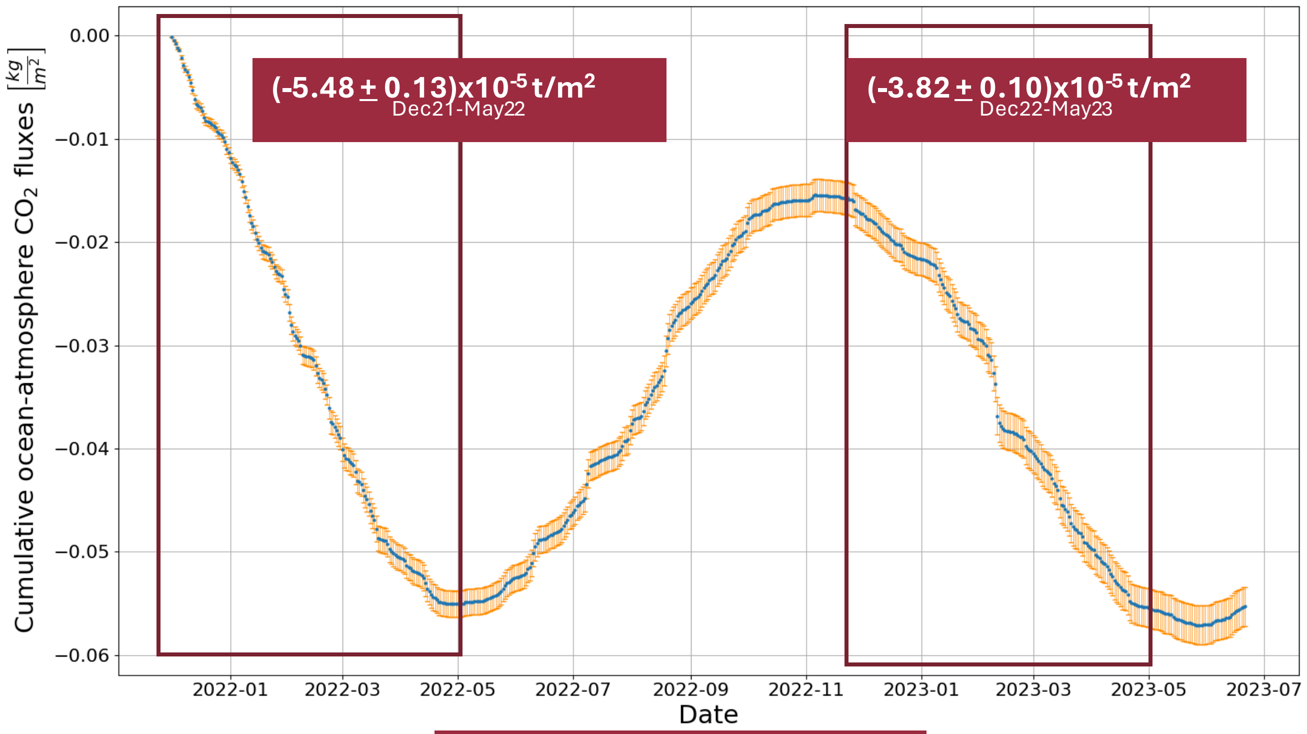


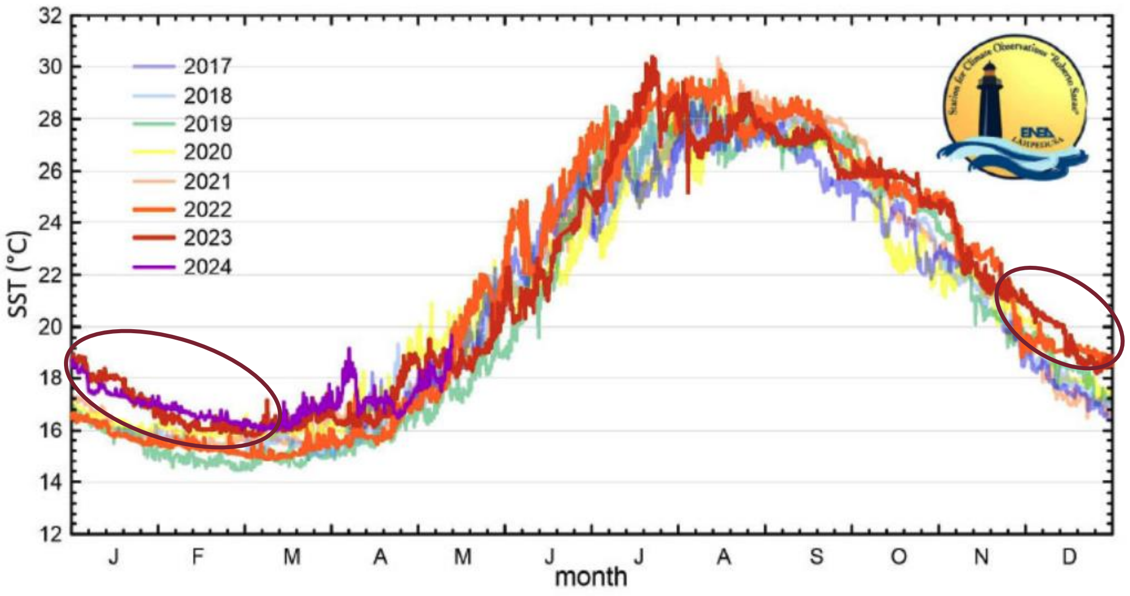
In situ pCO_2 and CO_2 fluxes



$(-1.73 \pm 0.16)\times10^{-5}$ t/m^2
Dec21-Dec22

In situ pCO_2 and CO_2 fluxes





Reduction of about 30% in ocean absorption

Hourly SST data; orange and red refers to 2022 and 2023

A. Mignot, K. von Schuckmann, P. Landschützer, F. Gasparin, S. van Gennip, C. Perruche, J. Lamouroux, and T. Amm. Decrease in air-sea CO_2 fluxes caused by persistent marine heatwaves. Nature Communications, 13(1):4300, 2022. doi: https://doi.org/10.1038/s41467-022-31983-0.

Outline:

1. Carbon cycle
2. Objectives
3. In situ pCO_2 and CO_2 fluxes
4. **Satellite algorithms**
5. Results and discussion
6. Conclusions and next steps

Satellite algorithms

Regional regression algorithms are introduced to estimate pCO_2 with satellite measurable proxies:

- Use of variables linked to physical and biological processes affecting the marine carbon cycle
- Introduction of new variables in the regression models
- Comparison of selected satellite data with in situ data to evaluate the reliability
- Use of model or in situ-adapted data for variable with large deviation from the observations
- **First validation of OLCI** (Sentinel-3 satellite) **PAR** product in the Central Mediterranean
- Use of pCO_2 uncertainty as weight in the regression models **(«weighted» models)**

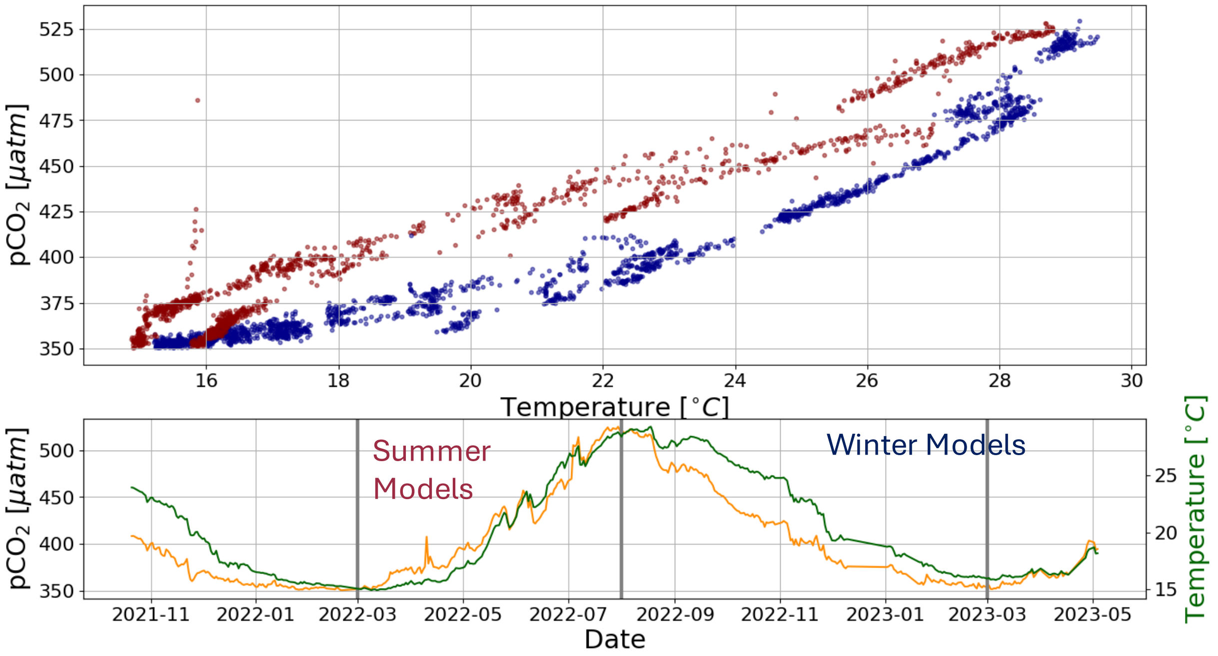
Satellite and model data

Variable	Dataset	Data type	Reference
SST	High Resolution and Ultra High-Resolution L3S SST	Satellite	CMEMS - Mediterranean Sea - High Resolution and Ultra High Resolution L3S Sea Surface Temperature. https://doi.org/10.48670/moi-00171
SSS	Ocean Reanalysis System 5 (ORAS5)	Model	C3S - ORAS5 global ocean reanalysis monthly data from 1958 to present. *Copernicus Climate Change Service (C3S) Climate Data Store (CDS).* https://doi.org/10.24381/cds.67e8eeb7
Wind Speed	Global Ocean Sea Surface Winds from Scatterometer	Satellite	CMEMS - Global Ocean Daily Gridded Sea Surface Winds from Scatterometer. Retrieved September 2023, from https://doi.org/10.48670/moi-00182
PAR	OLCI – Sentinel-3 satellite	Satellite	Pecci, M. et al. (2024). Validation of photosynthetically active radiation by OLCI on Sentinel-3 against ground-based measurements in the central Mediterranean and possible aerosol effects. *European Journal of Remote Sensing*, *57*(1). https://doi.org/10.1080/22797254.2024.2307617
PAR	MODIS – Aqua satellite	Satellite	NASA Ocean Biology Processing Group - Aqua Daily Photosynthetically Active Radiation https://oceancolor.gsfc.nasa.gov/l3/
CHL	Med. Sea, Bio-Geo-Chemical, Satellite Observations	Satellite	CMEMS - Mediterranean Sea, Bio-Geo-Chemical, L3, daily Satellite Observations (1997-ongoing) https://doi.org/10.48670/moi-00299
Atm xCO_2	In situ data-based fit	In situ-adapted	
Atmospheric pressure	ERA5 hourly reanalysis	Model	Hersbach, H., et al., (2023). ERA5 hourly data on single levels from 1940 to present. *Copernicus Climate Change Service (C3S) Climate Data Store (CDS).* https://doi.org/10.24381/cds.adbb2d47

Satellite algorithms

Traditional regression algorithms to estimate pCO_2 with satellite measurable proxies:

- Use of least-squares method
- Use of different functional forms, including **multiple parameters and non-linear terms**
- First selection of functional forms on smoothed datasets to reduce noise
- Training and test on in situ **daily** datasets
- Use of **seasonal regression** to exploit the pCO_2 hysteresis
- Best performing models applied to satellite data



Satellite algorithms

Machine learning approach to estimate pCO_2 with satellite measurable proxies:

- Use of eXtreme Gradient Boosting (**XGBoost**) algorithm
- Trained on **bi-hourly** data and tested on daily data
- First selection of input parameters on smoothed datasets to reduce noise
- Use of default and **cross-validated** settings
- Best performing models applied to satellite data

Training and test set

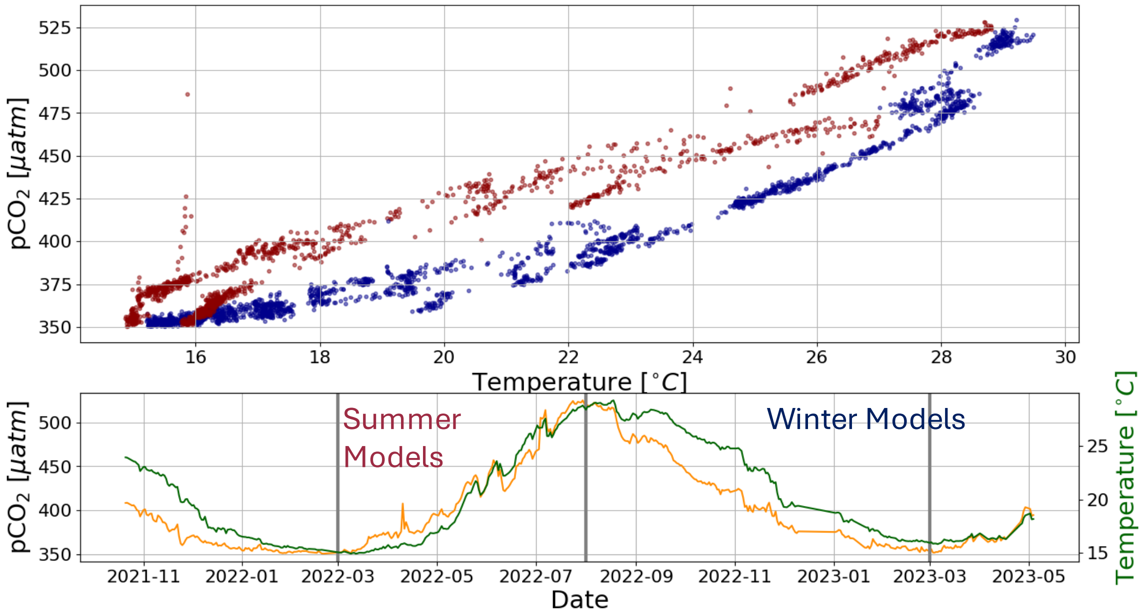
The entire dataset spans from December 2021 to June 2023 (18 months):

- Single regression for the whole dataset («**Annual models**»)
    - Training set is composed of 12 months of data (Dec21 to Dec22)
    - Test set is composed of 6 months of data (Jan23-June23)

Dataset	Bi-hourly	Daily
Training set	2400 data pairs	200 data pairs
Test set	/	130 data pairs

Training and test set

- Traditional regression models divided to follow the branches of the hysteresis («**Seasonal models**»):
  - Summer models (Mar-Aug)
    - Training set: Mar22-Aug22
    - Test set: Mar23-Aug23
  - Winter models (Aug-Mar)
    - Training set: Dec21-Mar22 and Aug22-Dec22
    - Test set: Dec22-Mar23



Dataset	Summer models	Winter Models
Training set	120 data pairs	180 data pairs
Test set	80 data pairs	50 data pairs

Performance metrics

Bias = $\frac{1}{N}\sum_i (y_i - x_i)$

RMSD = $\sqrt{\frac{1}{N}\sum_i (y_i - x_i)^2}$

$$\mathrm{R}^2 = 1 - \frac{\sum_i (y_i - x_i)^2}{\sum_i (y_i - \bar{y})^2}$$

$$\bar{R}^2 = 1 - (1 - \mathrm{R}^2)\frac{n-1}{n-k-1}$$

Where

- y is the predicted values
- x is the observed value
- n is the dataset size
- k is the number of parameters used in the regression

Outline:

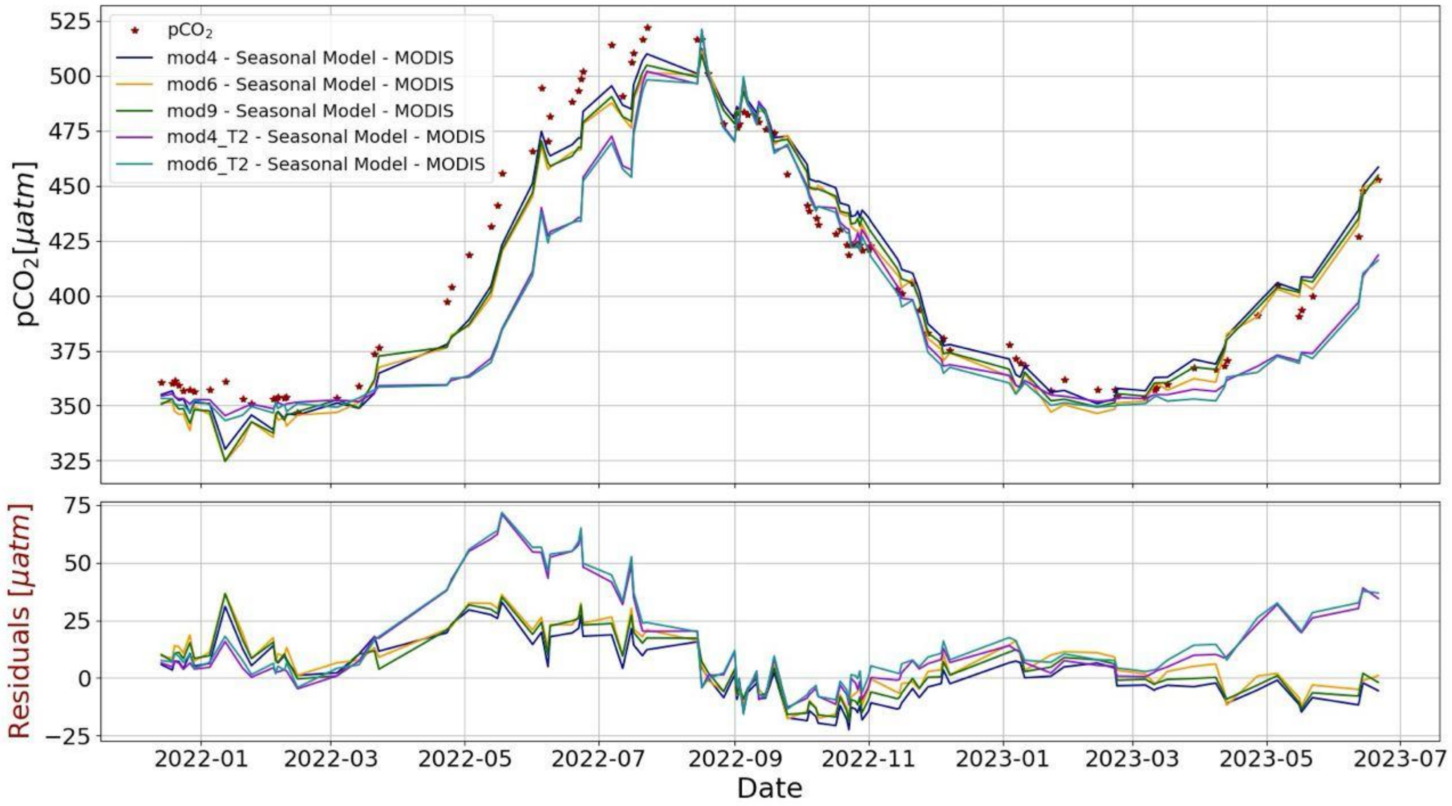
1. Carbon cycle
2. Objectives
3. In situ pCO_2 and CO_2 fluxes
4. Satellite algorithms
5. **Results and discussion**
6. Conclusions and next steps

Traditional regression models

Model	**Functional Form**	**SST** $[\frac{\mu atm}{°C}]$	**SST²** $[\frac{\mu atm}{°C^2}]$	**SSS** $[\frac{\mu atm}{PSU}]$	**CHL** $[\frac{\mu atm \cdot L}{\mu g}]$	**PAR** $[\frac{\mu atm \cdot m^2}{W}]$	**Wspd** $[\frac{\mu atm \cdot s}{m}]$	**Const.** $[\mu atm]$
Mod4	A·SST+B·SSS+C·CHL+ D·PAR+E	9.20	0	10.35	-99.12	0.13	0	-146.36
Mod6	A·SST+B·CHL+C·PAR+ D·**wspd**+E	9.23	0	0	-98.99	0.11	-0.66	247.17
Mod9	A·SST+B·CHL+C·PAR+D	9.28	0	0	-99.99	0.14	0	238.68
Mod4_T2	A·SST+B·**SST²** +C·SSS+D·CHL+E·PAR+F	25.31	-0.36	18.06	-66.98	0.06	0	-599.96
Mod6_T2	A·SST+B·**SST²**+C·CHL+ D·PAR+E·**wspd**+F	24.56	-0.34	0	-68.94	0.05	-0.78	90.43

Models functional form and coefficient for the «seasonal summer» models weighted with the reciprocal of the pCO_2 uncertainty

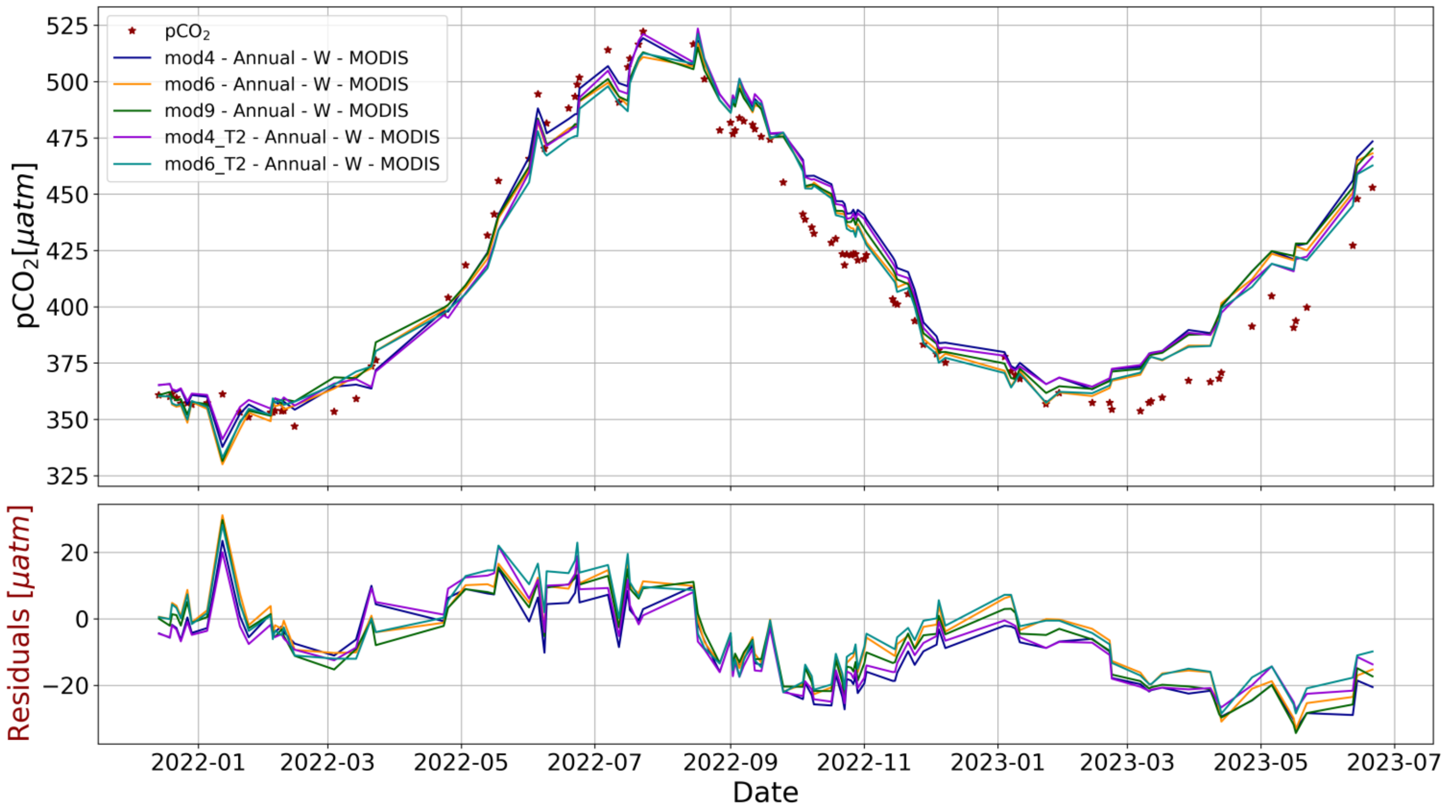
Traditional regression models



«Seasonal models» with MODIS PAR product non weighted regression

	Model	$\bar{R}^2$	RMSD [μatm]	Bias [μatm]
NW	**Mod4**	**0.94**	**13.1 (3%)**	**-1.5 (< 1%)**
	Mod6	0.92	14.4 (3%)	-5.8 (2%)
	Mod9	**0.93**	**13.7 (3%)**	**-4.3 (1%)**
	Mod4_T2	0.75	24.1 (6%)	-13.1 (3%)
	Mod6_T2	0.74	25.1 (6%)	-14.8 (3%)
W	**Mod4**	**0.93**	**13.2 (3%)**	**-1.2 (< 1%)**
	Mod6	0.92	14.6 (3%)	-5.9 (2%)
	Mod9	0.92	14.0 (3%)	-4.0 (1%)
	Mod4_T2	0.72	25.0 (6%)	-13.4 (3%)
	Mod6_T2	0.71	26.1 (6%)	-15.5 (3%)

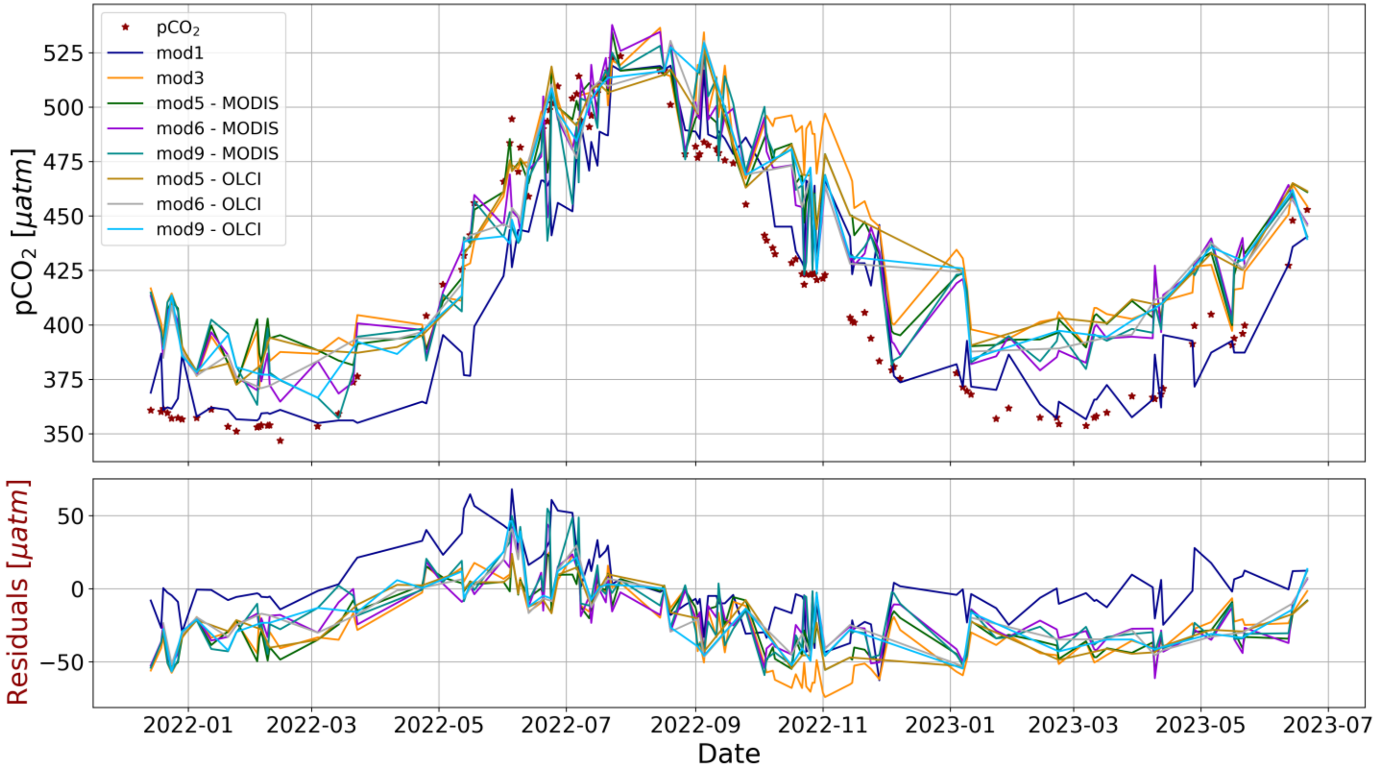
Traditional regression models



«Annual models» with MODIS PAR product weighted regression

	Model	$\overline{R}^2$	RMSD [μatm]	Bias [μatm]
NW	Mod4	0.81	21.2 (5%)	17.2 (4%)
	Mod6	0.82	20.1 (5%)	15.0 (3%)
	Mod9	0.81	20.7 (5%)	15.8 (4%)
	Mod4_T2	0.81	21.2 (5%)	17.2 (4%)
	Mod6_T2	0.83	20.1 (5%)	15.0 (3%)
W	Mod4	0.91	14.9 (3%)	9.2 (2%)
	Mod6	**0.93**	**13.2 (3%)**	**5.3 (1%)**
	Mod9	0.92	13.9 (3%)	6.8 (2%)
	Mod4_T2	0.92	14.2 (3%)	7.9 (2%)
	Mod6_T2	**0.94**	**13.0 (3%)**	**4.4 (1%)**

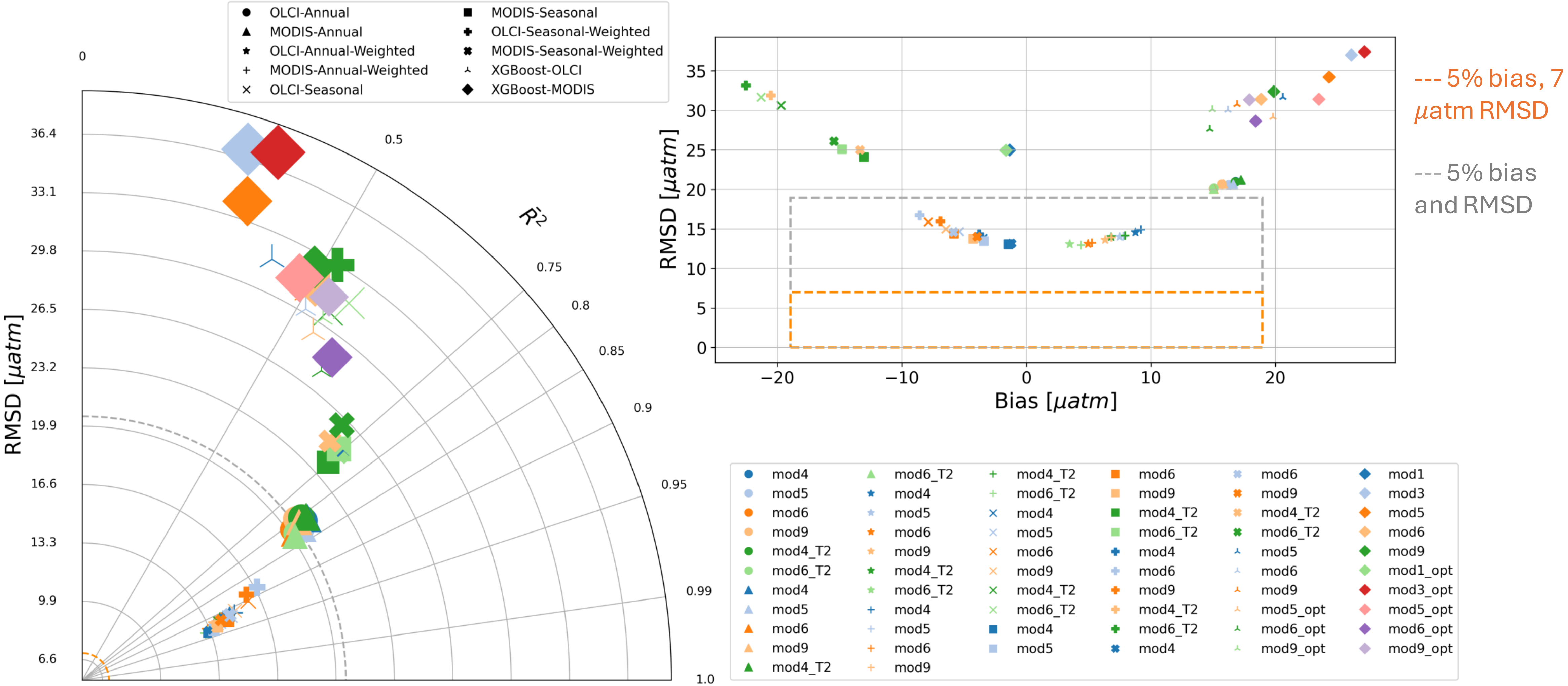
ML models



Model	Input variables
Mod1	SST
Mod3	SST, SSS, CHL
Mod5	SST, SSS, CHL, PAR, WSPD
Mod6	SST, CHL, PAR, WSPD
Mod9	T, CHL, PAR

	Model	$\bar{R}^2$	RMSD $[\mu atm]$	Bias $[\mu atm]$
D	**Mod1**	**0.76**	**25.0 (6%)**	**-1.4 (< 1%)**
	Mod3	0.30	37.0 (8%)	26.1 (6%)
	Mod5	0.33	34.2 (8%)	24.3 (5%)
	Mod6	**0.51**	**31.4 (7%)**	**18.8 (4%)**
	Mod9	0.49	32.4 (7%)	19.9 (4%)
CV	**Mod1**	**0.76**	**25.0 (6%)**	**-1.7 (<1%)**
	Mod3	0.35	37.4 (8%)	27.1 (6%)
	Mod5	0.48	31.4 (7%)	23.5 (5%)
	Mod6	**0.61**	**28.7 (7%)**	**18.4 (4%)**
	Mod9	0.54	31.3 (7%)	17.9 (4%)

Performance summary

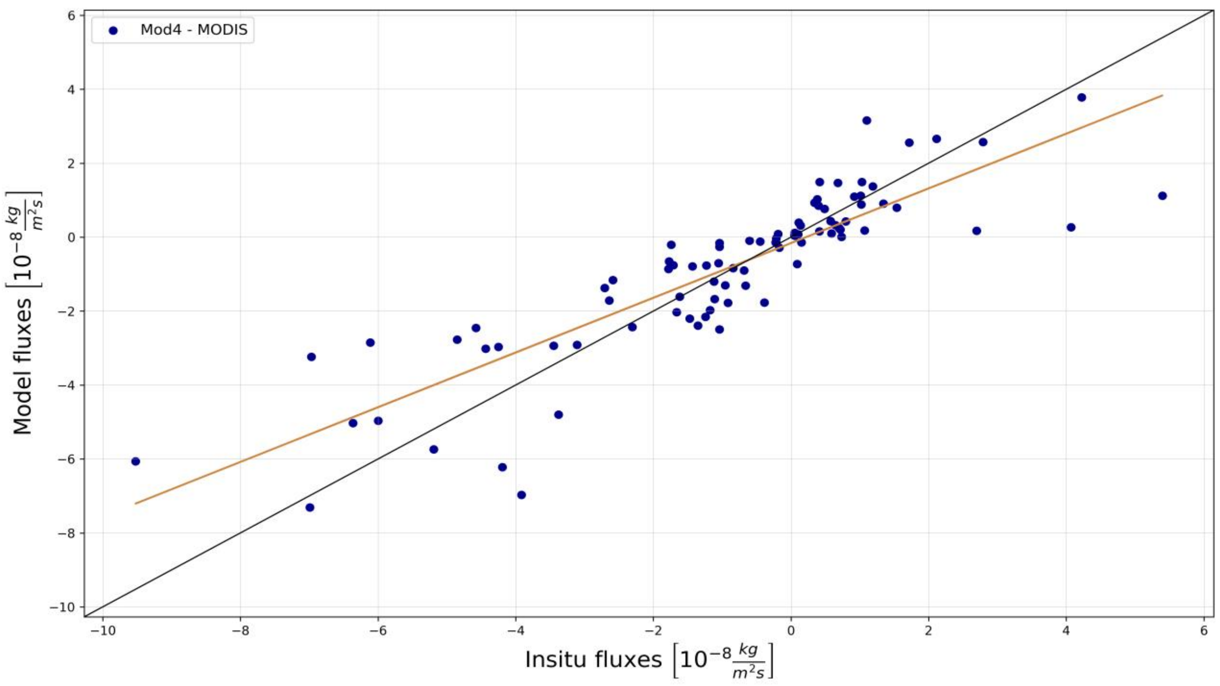


Fluxes estimates

- Fluxes computed using satellite-estimated pCO_2 and satellite/model-based ancillary quantities
- The need of simultaneous data from different dataset leads to a reduced dataset (approximately 100 data pairs for the MODIS-based dataset and 50 for the OLCI-based dataset)

	Model	R^2	RMSD $[\frac{kg}{m^2 s}]$	Bias $[\frac{kg}{m^2 s}]$
O	**Mod6**	0.74	$1.2 \cdot 10^{-9}$ (100%)	$-2.1 \cdot 10^{-10}$ (18%)
	Mod4_T2	0.74	$1.2 \cdot 10^{-9}$ (100%)	$-2.4 \cdot 10^{-10}$ (20%)
	Mod5_XGBoost	0.57	$1.6 \cdot 10^{-9}$ (140%)	$-6.0 \cdot 10^{-10}$ (50%)
	Mod9_XGBoost	0.56	$1.6 \cdot 10^{-9}$ (140%)	$-5.3 \cdot 10^{-10}$ (47%)
M	**Mod4**	0.75	$1.3 \cdot 10^{-9}$ (110%)	$-8.1 \cdot 10^{-11}$ (7%)
	Mod9	**0.75**	**$1.3 \cdot 10^{-9}$ (110%)**	**$1.0 \cdot 10^{-12}$ (<< 1%)**
	Mod6_T2	0.71	$1.4 \cdot 10^{-9}$ (120%)	$-3.0 \cdot 10^{-10}$ (25%)
	Mod1_XGBoost	0.64	$1.5 \cdot 10^{-9}$ (130%)	$-3.0 \cdot 10^{-10}$ (25%)
	Mod3_XGBoost	0.24	$2.2 \cdot 10^{-9}$ (190%)	$-1.2 \cdot 10^{-9}$ (10%)
	Mod6_XGBoost	0.37	$2.0 \cdot 10^{-9}$ (185%)	$-9.0 \cdot 10^{-10}$ (80%)

Mod4 - MODIS dataset



Outline:

1. Carbon cycle
2. Objectives
3. In situ pCO_2 and CO_2 fluxes
4. Satellite algorithms
5. Results and discussion
6. **Conclusions and next steps**

Conclusion and next steps

- Observed data of pCO_2 and CO_2 fluxes show a net sink effect

- Strong MHW impact on the magnitude of the exchanges

- The use of regional regression algorithms improve the agreement between estimated and observed pCO_2 values (with traditional regressions performing better than ML, probably due to the small dataset)

- CO_2 fluxes estimates using satellite-based pCO_2 and ancillary quantities show a good agreement with fluxes computed using observed data

Dataset	Bias	RMSD	R^2
CMEMS pCO_2	28.4 $[\mu atm]$ (7%)	40.0 $[\mu atm]$ (10%)	0.91
MPNR algorithm	-7.4 $[\mu atm]$ (2%)	31.9 $[\mu atm]$ (7%)	0.59
Mod6_T2	**4.4 $[\mu atm]$ (1%)**	**13.0 $[\mu atm]$ (3%)**	**0.94**
CMEMS Fluxes	$9.8 \cdot 10^{-10}$ $[\frac{kg}{m^2 s}]$ (86%)	$3.1 \cdot 10^{-9}$ $[\frac{kg}{m^2 s}]$ (270%)	0.22
Mod9	**$1.0 \cdot 10^{-12}$ $[\frac{kg}{m^2 s}]$ (<< 1%)**	**$1.3 \cdot 10^{-9}$ $[\frac{kg}{m^2 s}]$ (110%)**	**0.75**

Conclusion and next steps

- Despite being promising, a larger dataset is needed for a more robust statistics:
    - Use of different satellite input (CMEMS L4 salinity, ERA5 or CCMP wind speed, SEVIRI daily PAR) to increase the dataset size for pCO_2 and fluxes estimates
- Carry on the monitoring with a special focus on the SST and MHW impact on ocean absorption efficiency
- Extend the pCO_2 and CO_2 fluxes estimates to a broader area
- Compare the estimates with other Mediterranean carbon datasets (e.g., Integrated Carbon Observation System stations)

Contributions

- Datasets
  - All the mentioned measurements and observatories specifics are available at: https://www.lampedusa.enea.it/

- Journal Papers
  - Pecci, M., Colella, S., Di Iorio, T., Meloni, D., Monteleone, F., Pace, G., Sferlazzo, D., & di Sarra, A. (2024). Validation of photosynthetically active radiation by OLCI on Sentinel-3 against ground-based measurements in the central Mediterranean and possible aerosol effects. European Journal of Remote Sensing, 57(1). DOI: 10.1080/22797254.2024.2307617.

  - Pecci, M., di Sarra, A., et al. (manuscript in preparation). Air-sea CO_2 fluxes in the Central Mediterranean: the first year of measurements at Lampedusa.

  - Pecci, M., di Sarra, A. et al. (manuscript in preparation). Determination of pCO_2 from satellite data in the Central Mediterranean Sea.

Collaborators:

Fabrizio Anello, Giorgia Cinelli, Lorenzo De Silvestri, Alcide Giorgio di Sarra, Tatiana Di Iorio, Toni Iaccarino, Daniela Meloni, Francesco Monteleone, Giandomenico Pace, Mattia Pecci, Salvatore Piacentino, Damiano Sferlazzo

Station for Climate Observations "Roberto Sarao" ENEA LAMPEDUSA

Thank you!

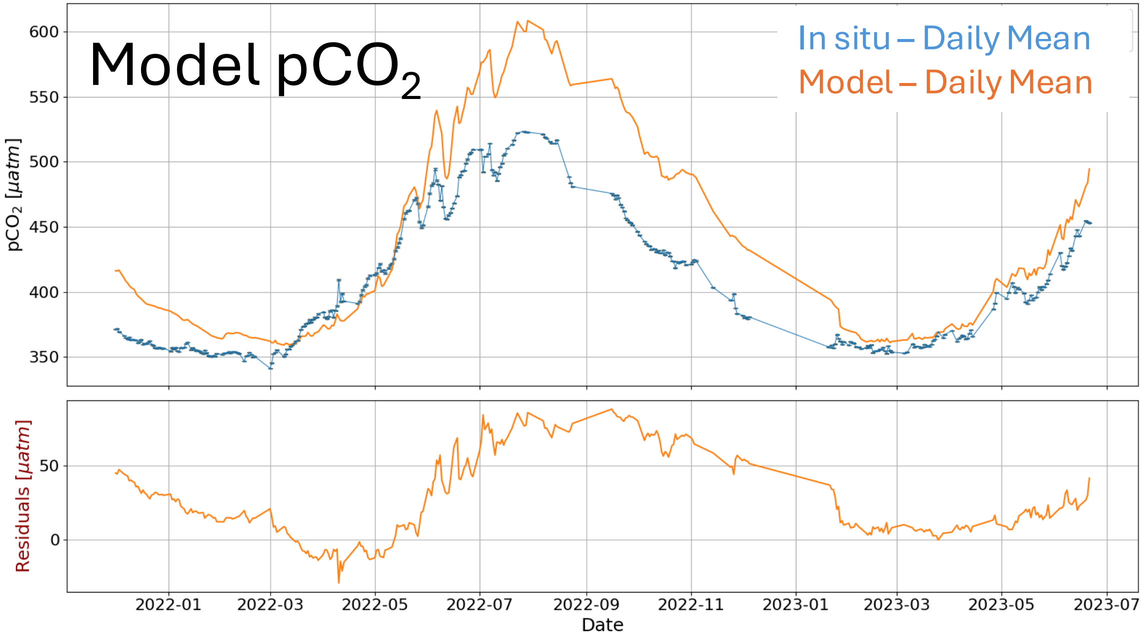
mattia.pecci@enea.it
alcide.disarra@enea.it

https://www.lampedusa.enea.it

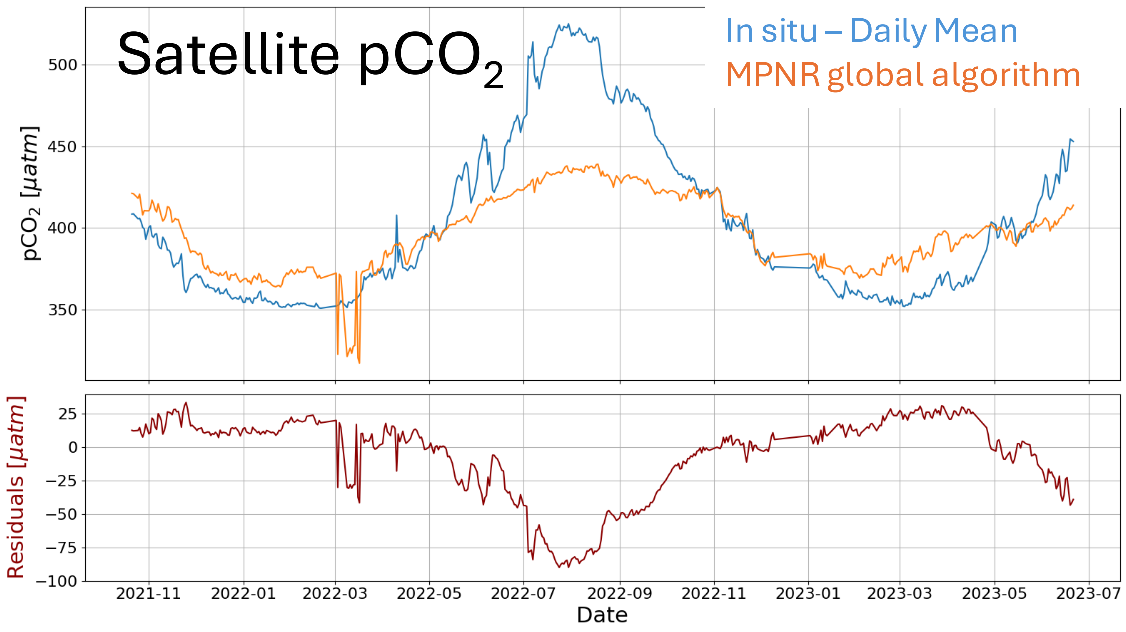
Marine carbon cycle

- Current monitoring mainly rely on model-based estimates of pCO_2 and CO_2 fluxes

- Works on satellite-based estimates of pCO_2



Mediterranean Sea Biogeochemistry Analysis and Forecast - https://doi.org/10.25423/cmcc/medsea_analysisforecast_bgc_006_014_medbfm3



K. V. Krishna, P. Shanmugam and P. V. Nagamani, "A Multiparametric Nonlinear Regression Approach for the Estimation of Global Surface Ocean pCO2 Using Satellite Oceanographic Data," in *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 13, pp. 6220-6235, 2020, doi: 10.1109/JSTARS.2020.3026363.