TOWARDS THE ESTIMATION OF DOC FROM SPACE IN THE OPEN OCEAN

ANA GABRIELA BONELLI

HUBERT LOISEL

VINCENT VANTREPOTTE

DANIEL JORGE

ANTOINE MANGIN

JULIEN DEMARIA

DOC: DESCRIPTION AND RELEVANCE

Second largest carbon pool in the ocean

Largest exchangeable reservoir of carbon in the marine environment

Dynamic component of the carbon budget in surface and deep waters of the open ocean

Great contribution to the biological carbon pump

YET, DOC spatial-temporal variability over open ocean is still not well known.

SOURCE

COASTAL WATER

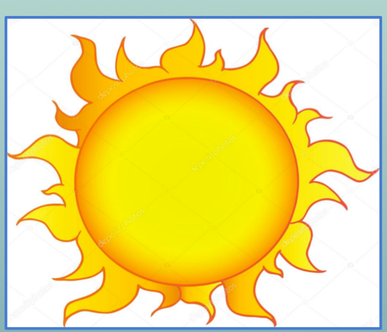
- Heteretrophic activity.
- Degradation of organic mater mainly with terrestrial origin (rivers runoff groundwater discharge)

OPEN OCEAN

- Phytoplankton and bacterial excretion.
- Non-trophic processes (viral lysis, grazing)

SINK

Microbial remineralization



Photochemical degradation

Other processes:

- abiotic aggregation into microparticles
- abiotic degradation via free radical reactions with oxygen

DOC TYPES

LABILE
- rapid mineralization
- < 1% of the total ocean DOC.
- Residence times: minutes to days

SEMI-LABILE
- More biologically resistant fraction.
- It accumulates in the surface ocean and part of it is exported to deeper water (<500 m)
- Residence time: months to years

SEMI-REFRACTORY:
- Part of the semi-labile DOC pool that reaches the deep ocean
- It is more recalcitrant than the semi-labile DOC, but more labile than the refractory.
- Residence time: decades to centuries.

REFRACTORY::
- 94% of the oceanic DOC
- Resistant to short term microbial mineralization
- Is exported to deep ocean
- Residence time: 12,500 years

CARACTERISTICS OF CDOM

COASTAL WATER

- Origin: Terrestrial (rivers discharge and land washing)
- Highly concentrated.
- Strongly correlated to DOC.

OPEN OCEAN

- Local biological origin.
- Low concentration (complicates its estimation with remote sensors).
- No direct correlation with DOC.

A good estimation of CDOM absorption from space in open ocean combined with other ocean color products could improve the estimation of DOC with remote sensors.

OBJECTIVES

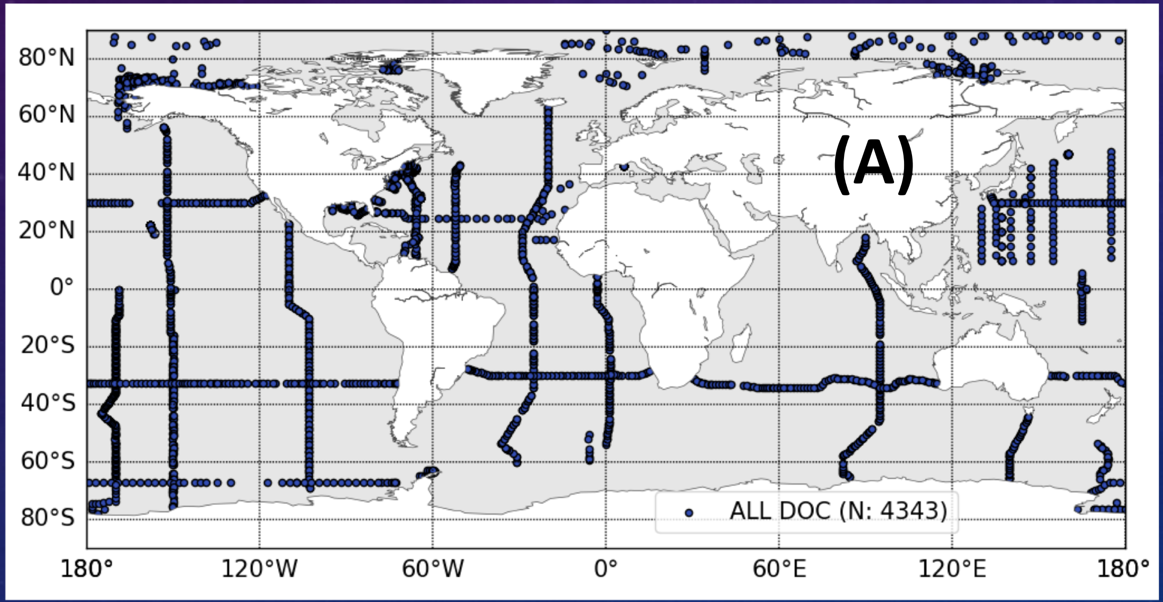
- Develop an algorithm to estimate labile/semi-labile DOC from space in open ocean based on a neural network approach.
  - Identify the pertinent parameters needed for the estimation of DOC from space.
  - Find the time lag between DOC and NN inputs to better account for the DOC dynamic.

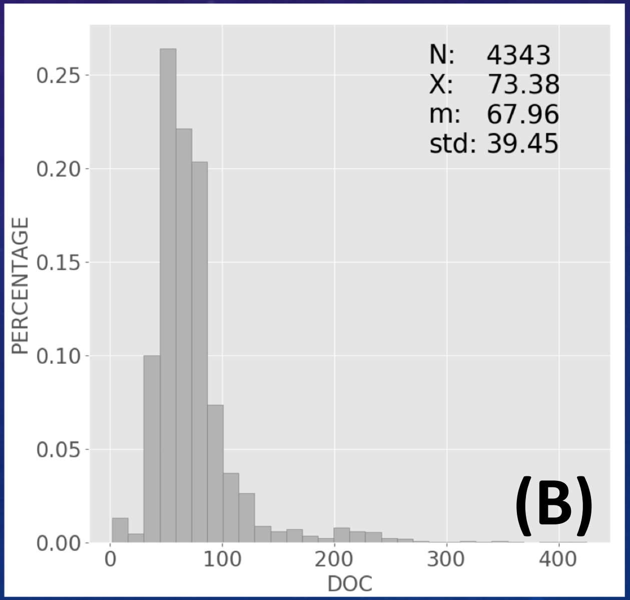
DOC IN SITU DATASET

SOURCES

- ANTARES
- BATS
- GLODAP
- GOCAD
- GoMX
- HOTS
- K2
- POLARSTERN
- TRANSDRIFT_VB 4
- TRANSDRIFT_YS

The DOC in situ dataset gathers data from all over the world reaching a wide cover of the global oceans. Covers a range form almost 0 to > 400 $\mu mol \cdot kg^{-1}$ with a slight negative asymmetry. The majority of the samples (~80%) present concentrations of DOC between 40 and 100 $\mu mol \cdot kg^{-1}$.





In situ DOC distribution (A) and its histogram (B). X corresponds to the mean, m is the median and std the standard deviation.

DOC MATCH UP DATASET

PAR Rrs [400, 412, 443, 490, 510, 590, 665]	GlobColour L3 merged database Temporal resolution: Daily Spatial resolution: 4km
SSS	ESA Sea Surface Salinity Climate Change Initiative Temporal resolution: Weekly Spatial resolution: 25 km
SST	NOAA Optimum Interpolation (OI) Sea Surface Temperature (SST) V2 Temporal resolution: Weekly Spatial resolution: 1 degree
MLD	MILA GPV (MIxed Layer data set of Argo, Grid Point Value) Temporal resolution: 10 days averaged Spatial resolution: 1 degree
CHL aCDOM	Both calculated from Globecolour L3 merged Rrs OC4 algorithm was used for Chl-a BL1 algorithm was used for aCDOM [443nm](publication in process)

For all cases the match up was searched for lag 0, 1, 2, 3 and 4. Being each times 7 days difference.

The neural network does not accept missing data. Based on the matchup with Rrs [412, 443, 490, 510, 590, 665], when any of the other variables presented gaps, they were filled.

PAR	European Union Open Data Portal (ODP) GMIS - SeaWiFS Monthly climatology photosynthetically available radiation (9km)
SSS	ISAS-15 temperature and salinity gridded fields Global data between 2002 and 2015 from the ARGO floats Monthly data
MLD	Ifremer/Los Mixed Layer Depth Climatology website Temporal resolution: monthly climatology Spatial resolution: 2 degrees

aCDOMBL1[m-1, 443nm] (BL1)

Semi-analytical model to estimate CDOM in clear water.
Adapted from Loisel et al. (2014) model developed to estimate CDOM in coastal water.

$$a_{CDOM}(443) = 10^{[0.99\,X - 0.052]}$$
$$X = \Delta_{kd} - \Delta p$$
$$\Delta p = 10^{[-0.906(\mathrm{Log10}(\Delta_{kd}) - 0.526]}$$
$$\Delta_{kd} = \left(kd(443) - kw(443)\right) - \left(kd(560) - kw(560)\right)$$

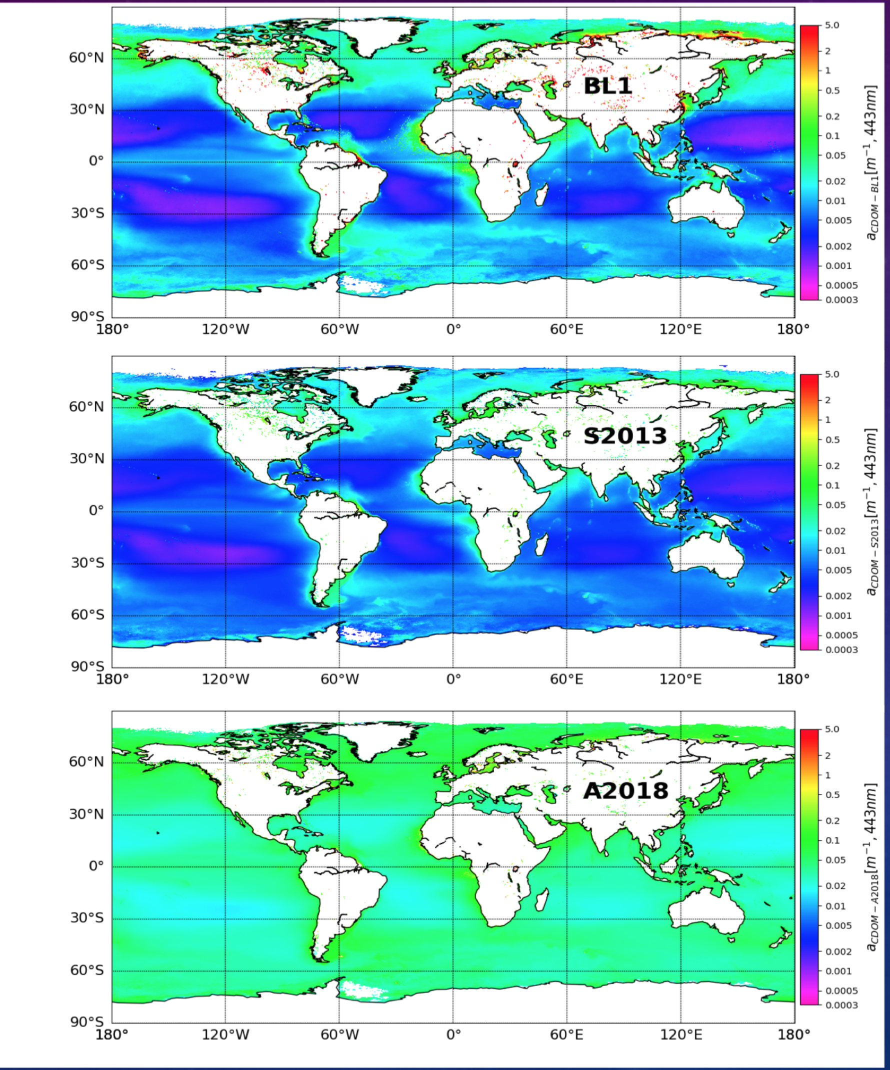
Assumptions:

- **Attenuation of light is driven by water, particulate matter and CDOM**
- **CDOM dose not absorb at long wavelengths.**

Δ_{kd} is estimated with a neural network adapted from Jamet et al. (2012).

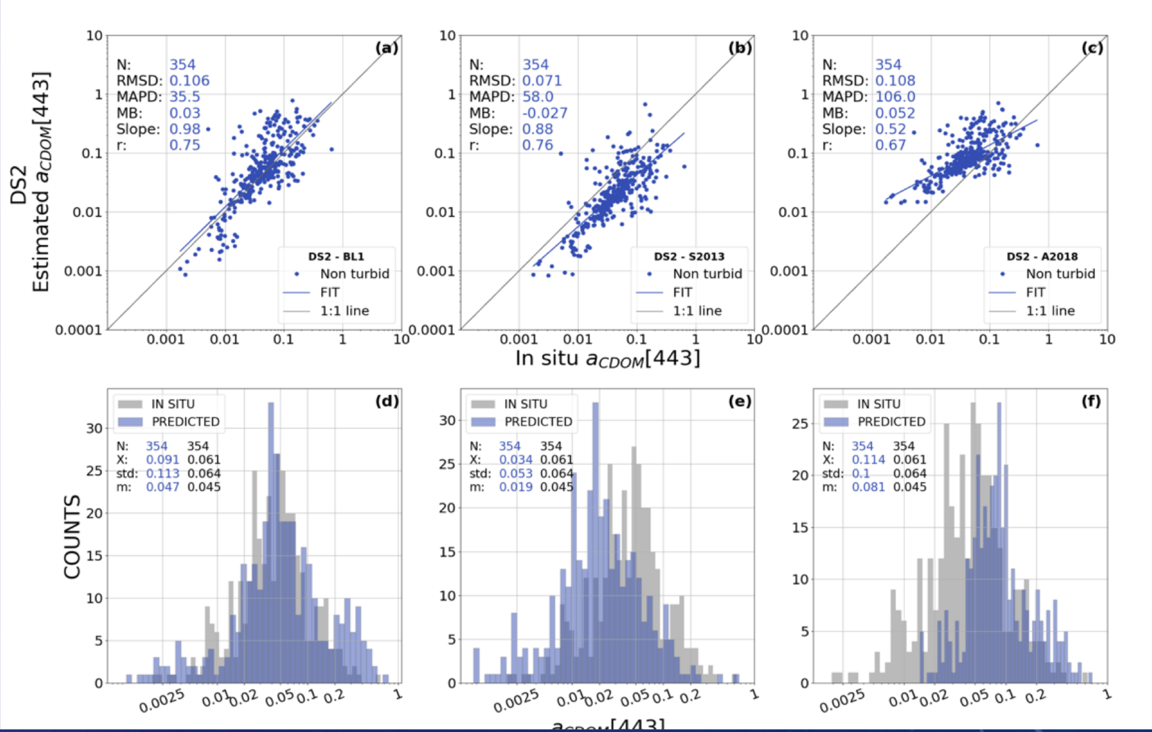
BL1 validation and comparison with other existing models

Down: *Annual climatology of (a) BL1, (b) S2013 and (c) A2018 produced with the 10 years archive of GlobColour L3 merged 25km 8 day composite data from 23rd April 2002 to 13th April 2012.*



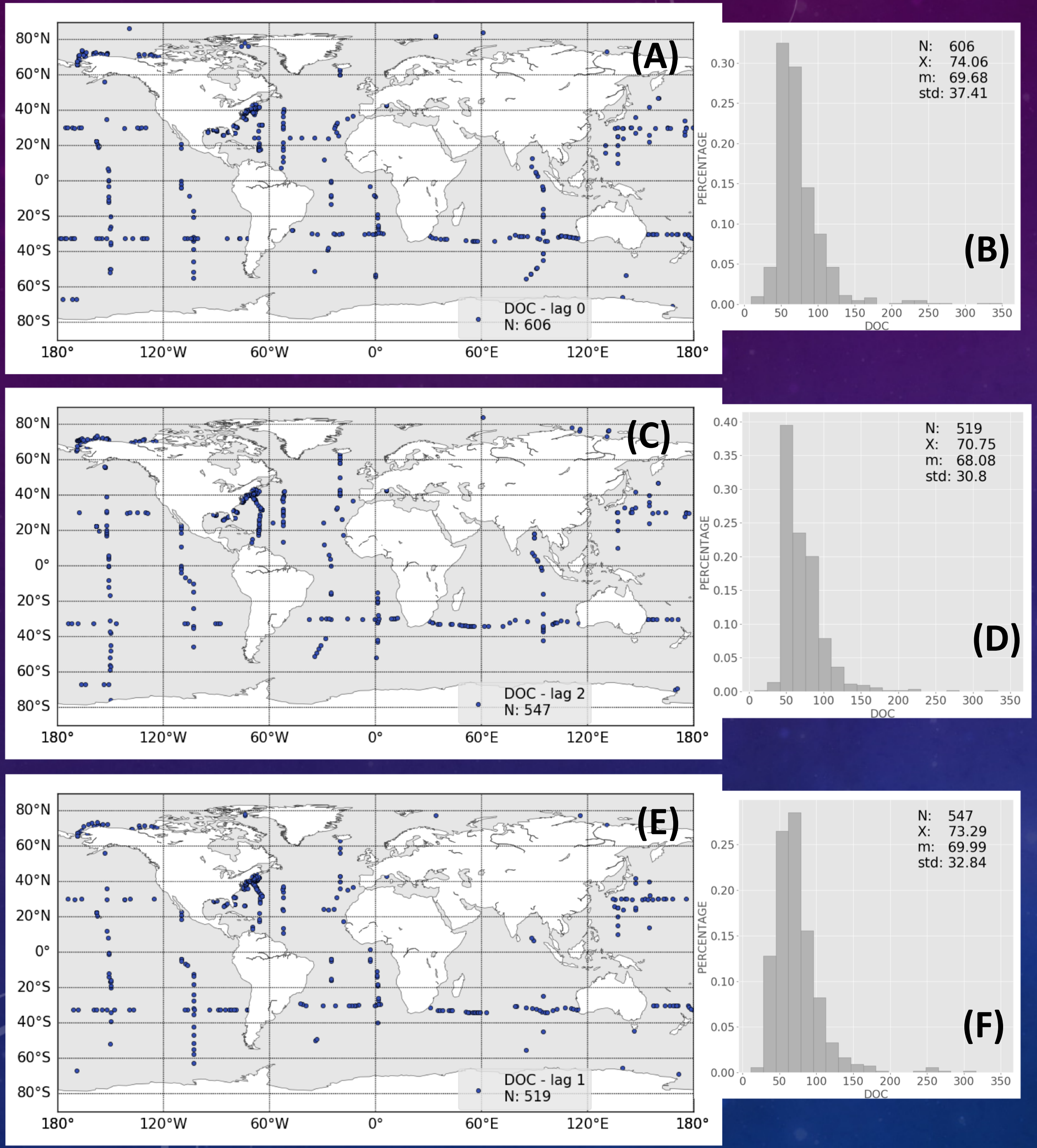
TESTED MODELS

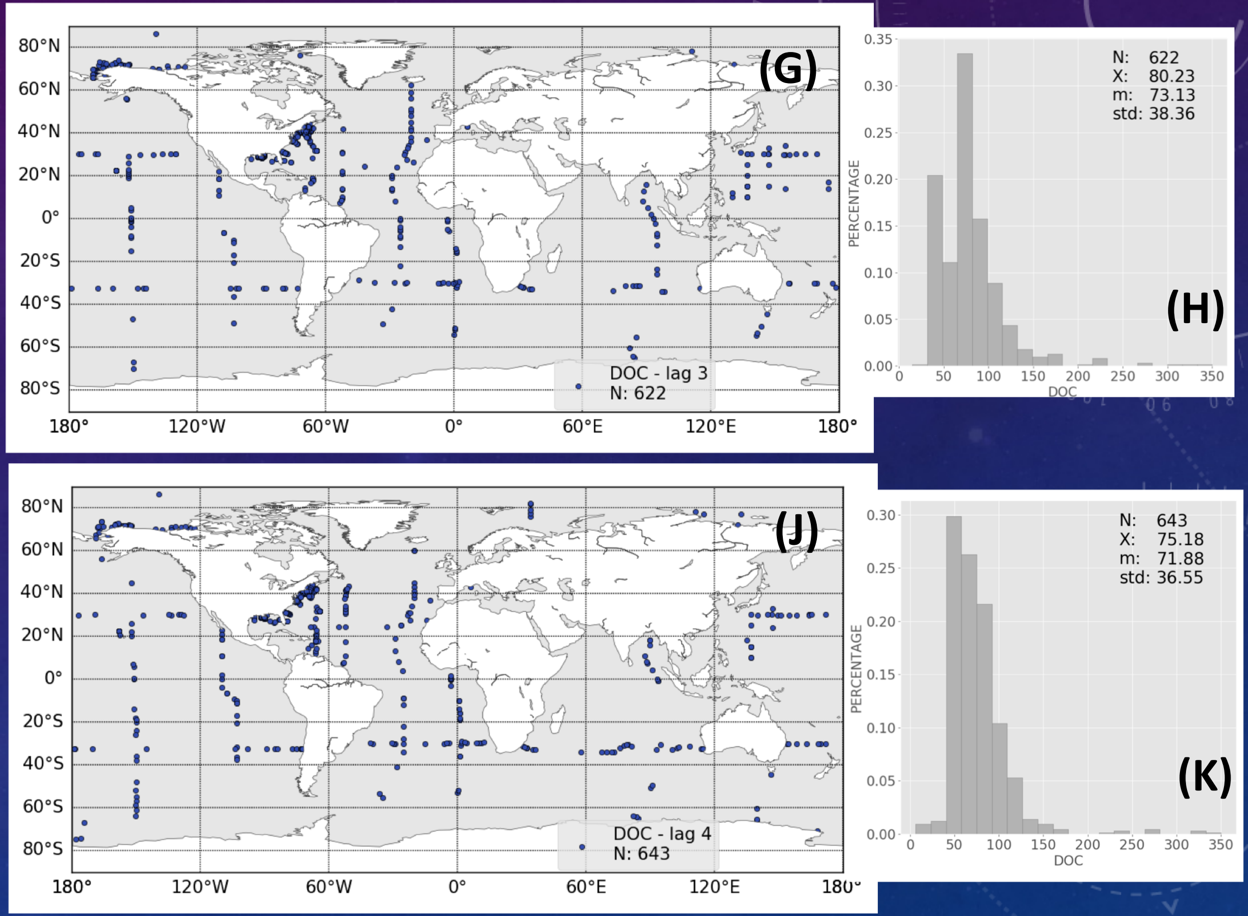
- BL1
- Swan et al.(2013; S2013)
- Aurin et al. (2018; A2018)



Up: *Validation of the three models over DS2 in situ non-turbid waters validation data set: S2013 (a), BL1 (b) and A2018 (c) and their respective histograms (d, e, f). In gray the histogram corresponding to aCDOM(443) measured in situ and in blue, the values estimated with each model.*

DOC DISTRIBUTION MATCH UP DATA SET





DOC distribution on the match up dataset for lag 0, 1, 2, 3 and 4 weeks (A, C, E, G, I). And their respective percentage histograms (B, D, F, H, K).

DOC NEURAL NETWORK

Multi-layer perceptron structure:

- Activation function: Rectified Linear Unit
- Optimizer: Adam optimizer
- Loss calculation: mean square error
- Density layers
- 2 hidden layers

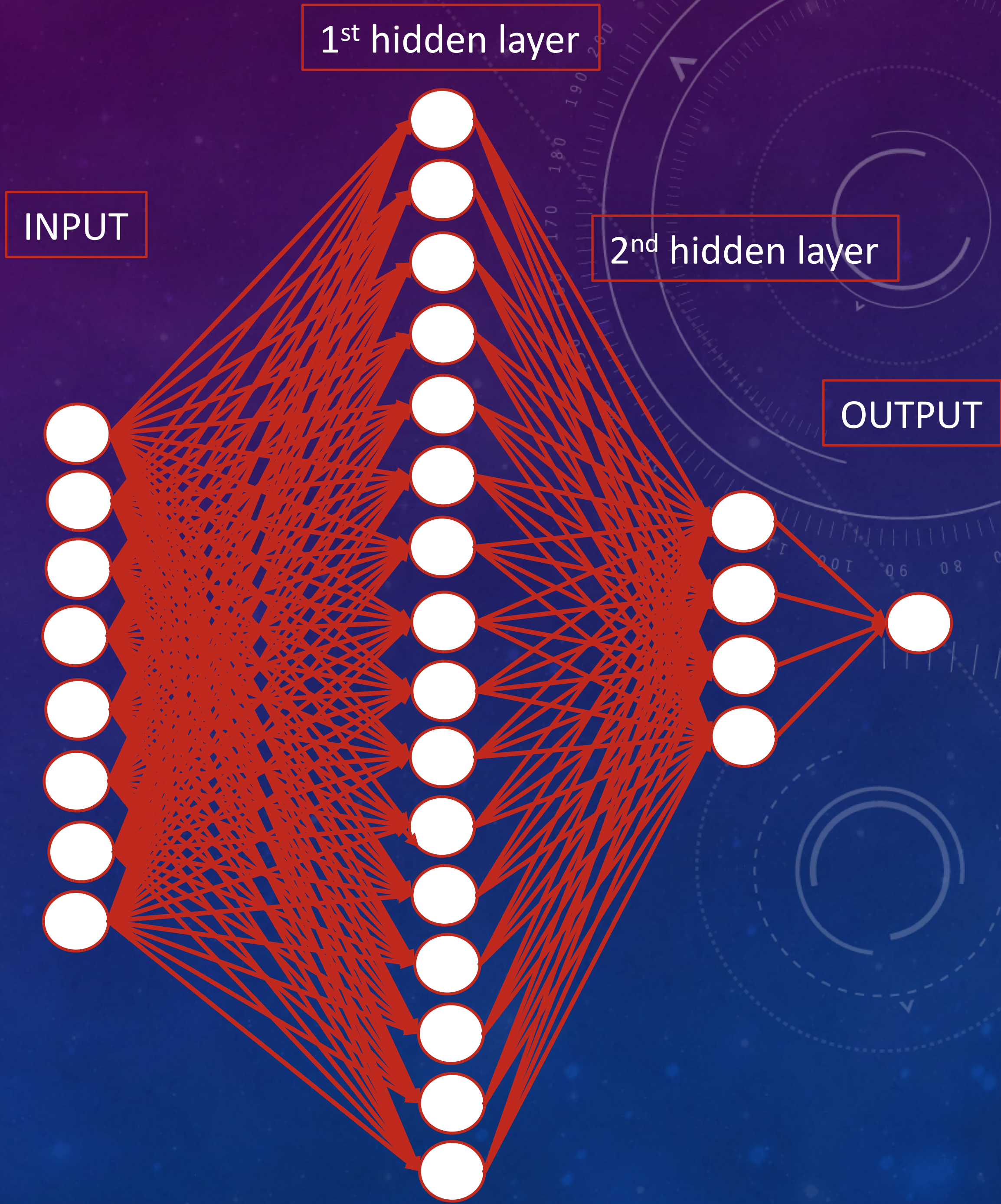
$$\# \, neurons = 2 * (\# \, inputs)$$

$$\# \, neurons = \frac{(\# \, inputs)}{2}$$

Input:

- Latitude
- Longitude
- Chlorophyll a concentration (CHL)
- CDOM absorption (aCDOM)
- Photosynthetical active radiance (PAR)
- Sea Surface Temperature (SST)
- Sea Surface Salinity (SSS)
- Mixed layer depth (MLD)

Output: DOC concentration

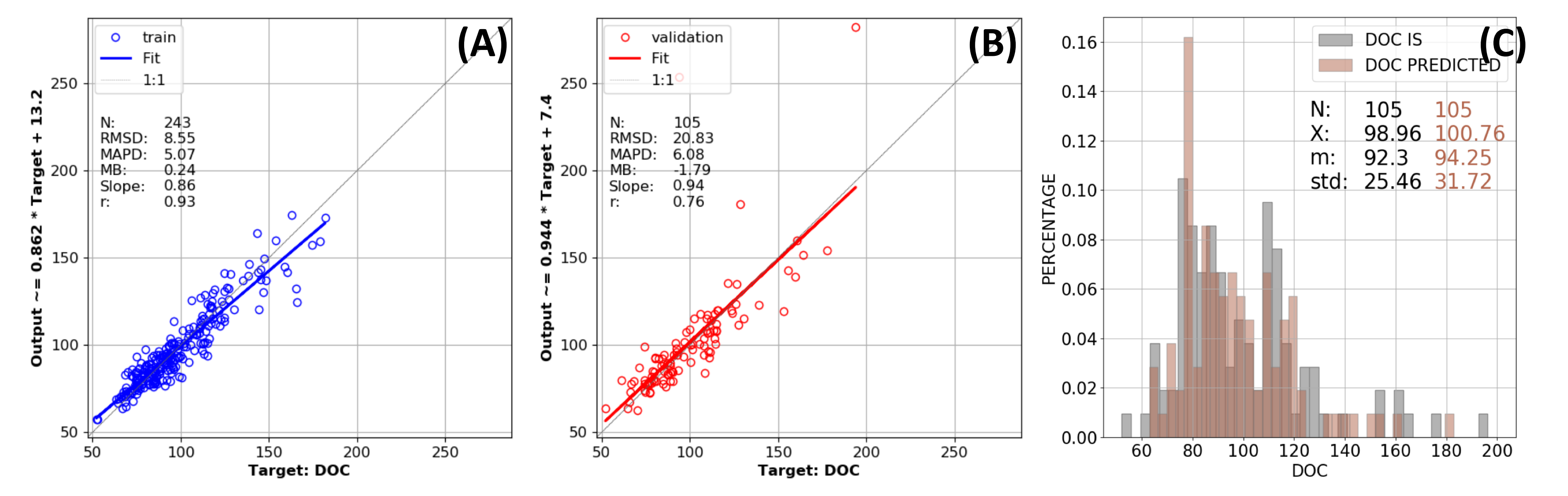


TEST OF THE NEURAL NETWORK WITH IN SITU DATA

IN SITU PRODUCTS

- Lat
- Lon
- SST
- SSS
- CHL
- CDOM [443nm]

Tested in situ variables have the potentiality to be good estimators of DOC with satellite products.



Scatterplot of (A) the training dataset vs DOC in situ and (B) the validation dataset. (C) Histogram of DOC calculated with the validation dataset and DOC in situ that matches it. X corresponds to the mean, m is the median and std the standard deviation.

NEURAL NETWORK: DECISION SYSTEM

TIME LAGS
t0 = 0
t1 = -1 week
t2 = - 2 weeks
t3 = - 3 weeks
t4 = - 4 weeks

Fix inputs

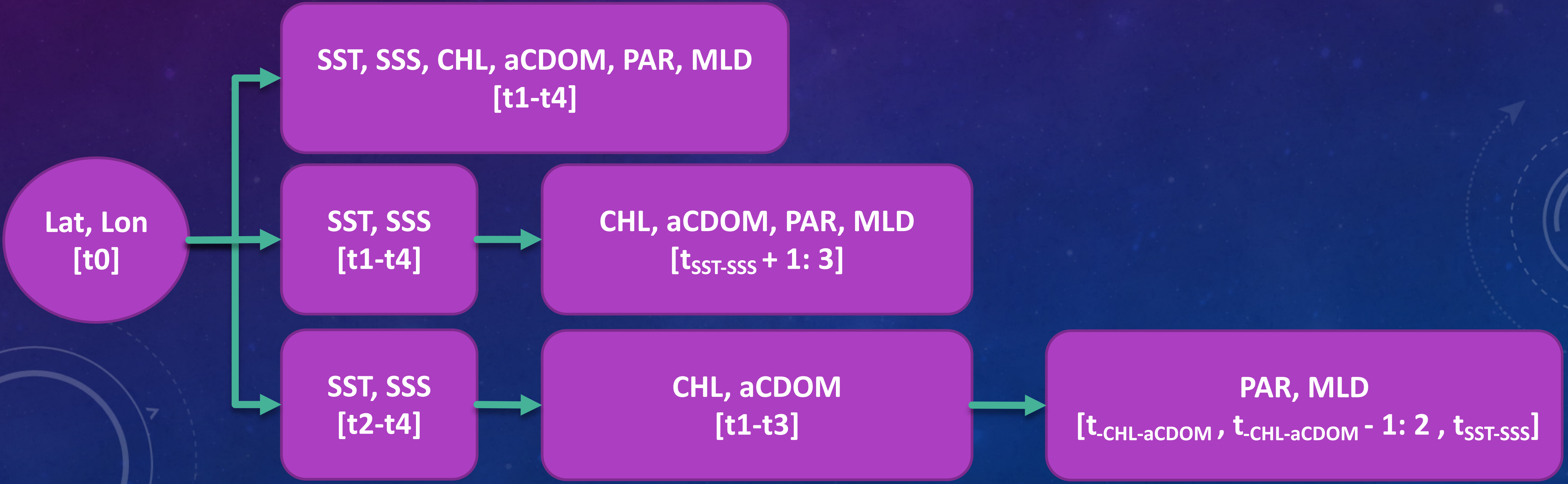
- Latitude
- Longongitude
- SST
- SSS

Mooving inputs

- CHL
- aCDOM
- PAR
- MLD

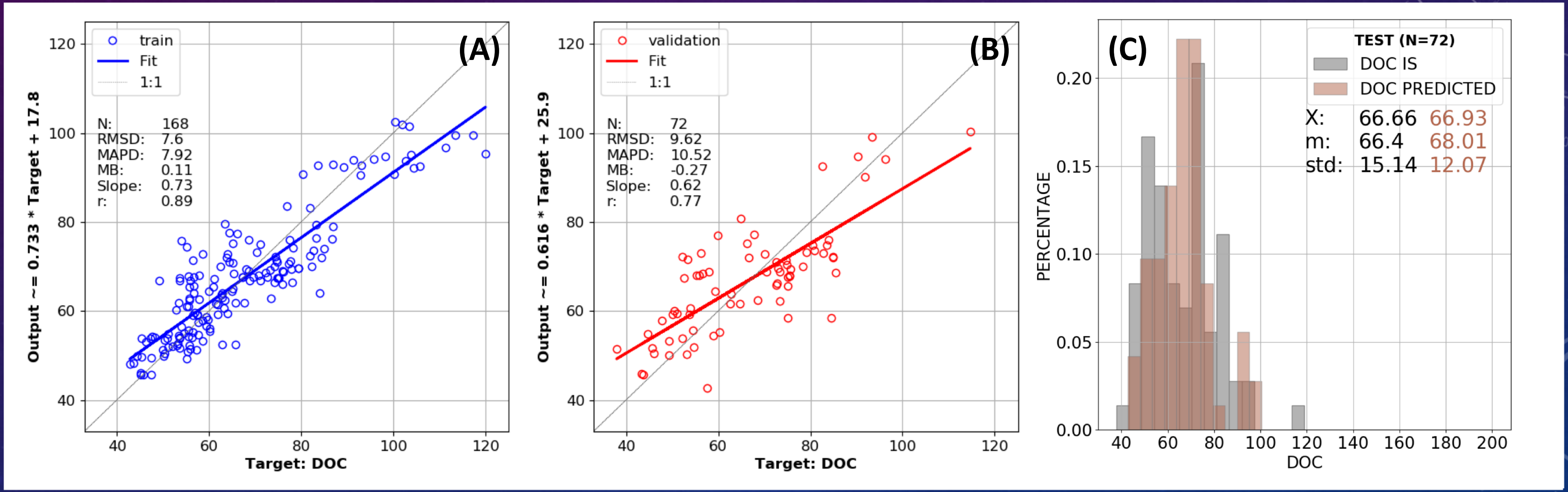
Used alternatively. Best results were obtained when all of them were included

IT HAS BEEN ALSO TESTED THE USE OF THE SAME PRODUCTC AT TWO DIFFERENT TIME LAGS FOR THE SAME RUN



BEST COMBINATION OF TIME LAGS TO ESTIMATE DOC

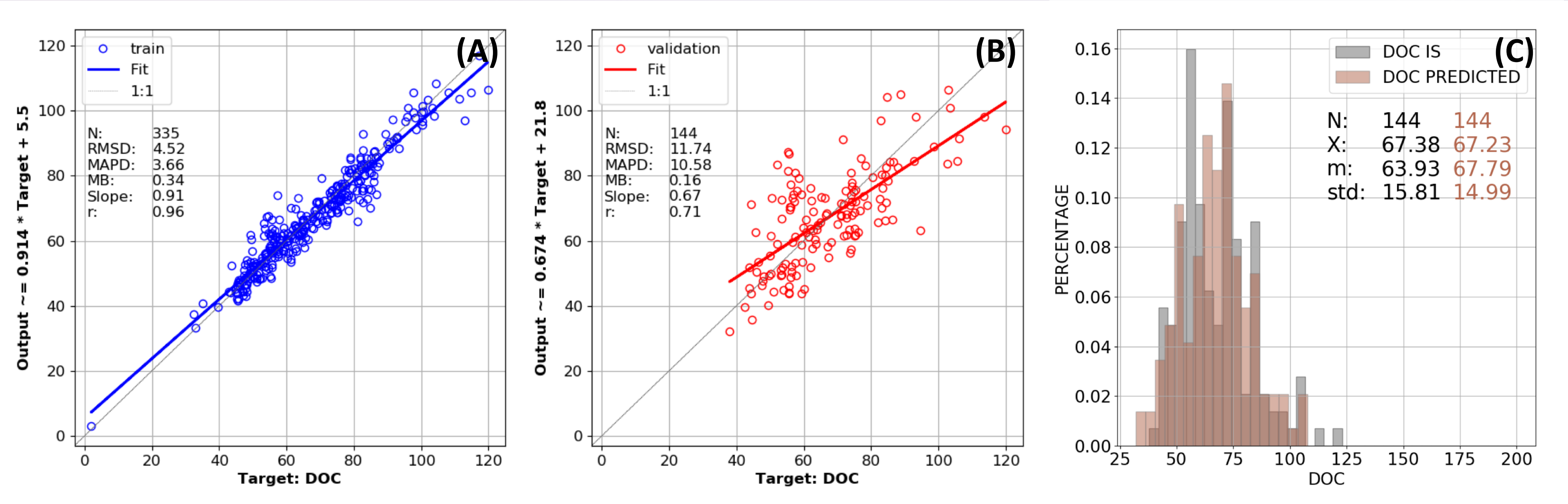
Lat, Lon [t0] + CHL, aCDOM [t1,t3] + PAR, MLD [t3,t4] + SST, SSS [t4]



Scatterplot of (A) the training dataset vs DOC in situ and (B) the validation dataset. (C) Histogram of DOC calculated with the validation dataset and DOC in situ that matches it. X corresponds to the mean, m is the median and std the standard deviation.

STRONG IMPROVEMENT WHEN RRS IS USED INSTEAD OF CHL AND CDOM

Lat, Lon [t0] + RRS [412,443,490,510,560,665] [t1,t3] + PAR, MLD [t3,t4] + SST, SSS [t4]



Scatterplot of (A) the training dataset vs DOC in situ and (B) the validation dataset. (C) Histogram of DOC calculated with the validation dataset and DOC in situ that matches it. X corresponds to the mean, m is the median and std the standard deviation.

CONCLUSIONS AND PERSPECTIVES

- Results are more accurate when the neural network is run directly with Rrs, decreasing the noise added by intermediate calculations.

- Based on preliminary results a neural network developed with Rrs (412, 443, 490, 510, 560, 665), MLD, PAR, SST, SSS, latitude and longitude is promising to estimate labile/semi-labile DOC from space in global scale.

- The use of different lags for the same variables improves the results, probably by linking a tendency of increase/decrease of the variable to the final concentration of DOC.

Next:

- Understand the time lag of the input parameters.
- Improve the neural network with the implementation of different combinations of activation functions and optimizer in order to reduce the error observed.
- Produce global maps and compare results with previous publications.
- Study the dynamic and variability of DOC in global oceans.