



**ISTITUTO NAZIONALE DI GEOFISICA E VULCANOLOGIA**

**Exploiting SeaDataCloud Temperature and Salinity time series data collections and comparing with Copernicus - a novel approach with SOURCE tool**

**P. Oliveri, S. Simoncelli, P. Di Pietro, S. Durante**  
*[paolo.oliveri@ingv.it](mailto:paolo.oliveri@ingv.it)*



# Problem and solution

## Problems

- Observational data are fragmented between different repositories and infrastructures!
- Data can be fragmented within the same infrastructure due to the data ingestion process from the different data providers!

## Solution

Use of SOURCE's merging tool to process the data and maximize the data available for usage!

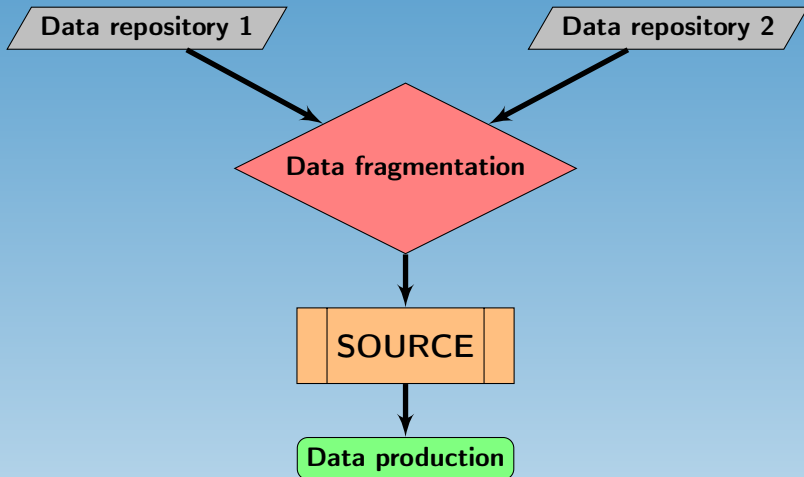


# Outline

- 1 Generalities on data infrastructures (CMEMS and SeaDataCloud) and related issues;
- 2 SOURCE tool description;
- 3 SeaDataCloud Sea Temperature and Salinity time series database analysis;
- 4 Merging of SeaDataCloud database with CMEMS;
- 5 Merging case test;
- 6 New INGV web service application (VIDEO).



## Solution







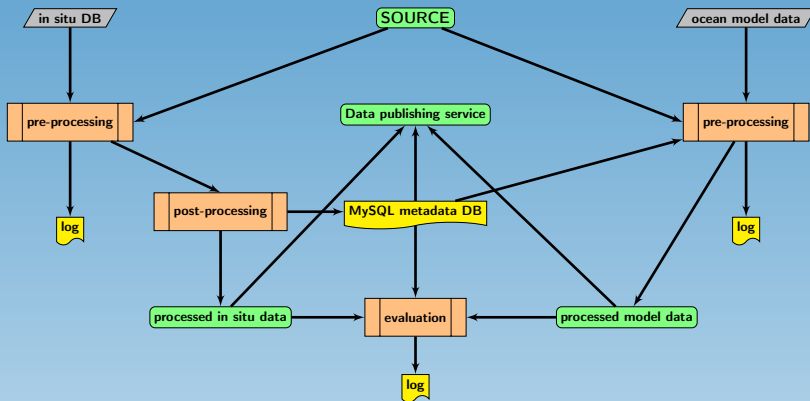
# SOURCE (Sea Observations Utility for Reprocessing, Calibration and Evaluation)

SOURCE is an open source software developed at INGV and written in Python that handles the pre-processing of the data coming from the different infrastructures, merging them in one final database. Post-processing consists of duplication removal and data quality control using several tools that follow the ocean best practices guidelines.

SOURCE was previously developed to deal with the Coperincus Marine Service (CMEMS) in situ TAC Near Real Time database.

The first infrastructures data merging test has been done between SeaDataCloud V2 and CMEMS sea temperature and salinity data collections.

# SOURCE flow chart





# SeaDataCloud data collection

## SeaDataCloud ODV Database

A Global Ocean data time series collection in ODV (Ocean Data View) format has been exported in netCDF file and then analyzed.

The analyzed data sub-set covers the Mediterranean Sea and part of the Atlantic Ocean.

Only time series data from fixed platforms have been considered.

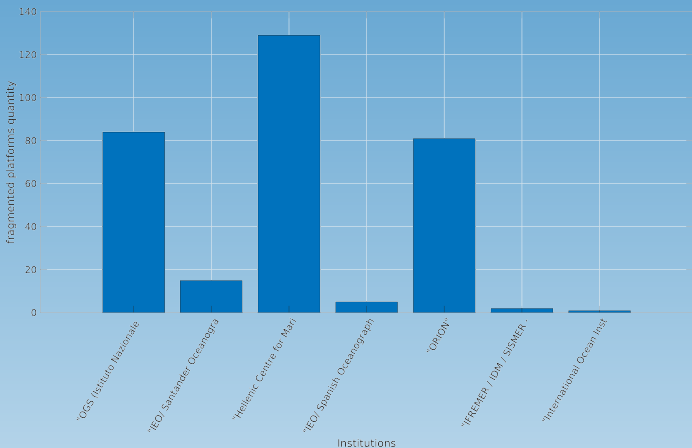


## pre processing issues

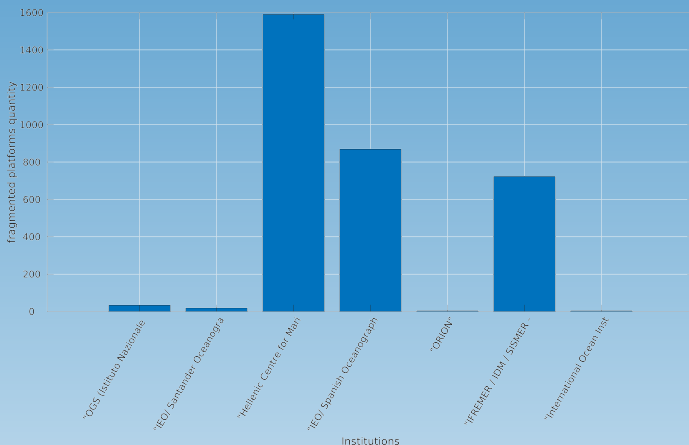
- Finding and aggregating all broken time series using likeness in ID parameter strings
- organizing metadata
- correcting time units
- filtering the data by area of interest or instrument type
- producing information on the original QC scheme by SeaDataNet infrastructure
- producing log files with all warning messages (missing time, depth, data, wrong QC variables, etc.)



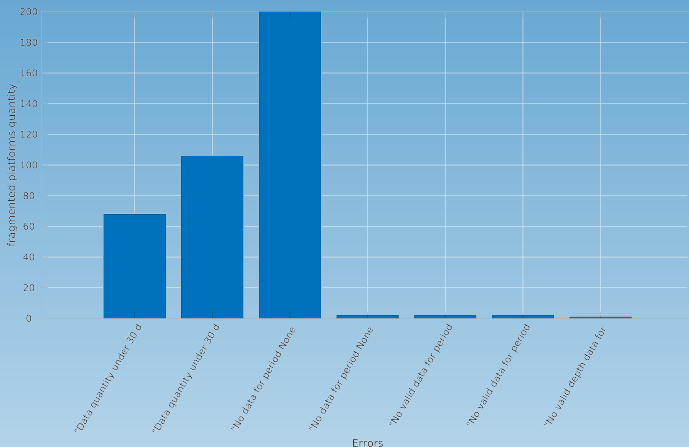
## Amount of fragmented platforms for each institution without repetitions



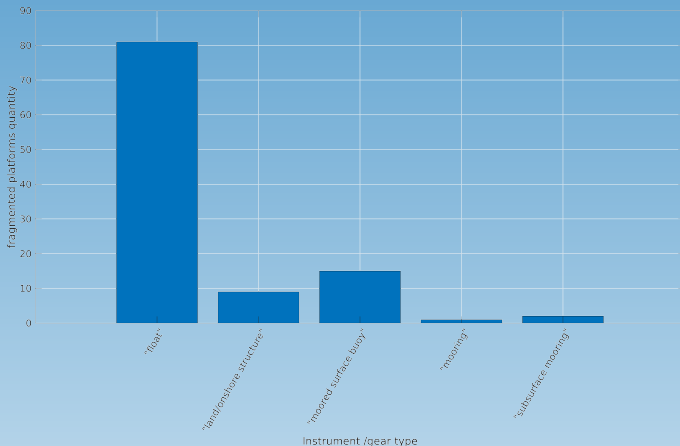
# Amount of fragmented platforms for each institution with repetitions



# Amount of encountered warnings during platform data processing



## Amount of fragmented platforms for each device type without repetitions







## Merging procedure

### Condition

- 1 Horizontal proximity;
- 2 likeliness in metadata identifiers.

If two platforms verifies the condition, the available fields and available depths for both platforms will be concatenated. Given  $z_1, \dots, z_n$  and  $z'_1, \dots, z'_m$  the recorded depths for the same field for two platforms, the merged time series will have the depths  $z''_1, \dots, z''_p$ , such that:

$$z'' = \begin{cases} z' & \text{if } z \notin \{z_1, \dots, z_n\}; \\ z & \text{if } z \notin \{z'_1, \dots, z'_m\}; \\ z(=z') & \text{otherwise,} \end{cases}$$



## Post processing procedure

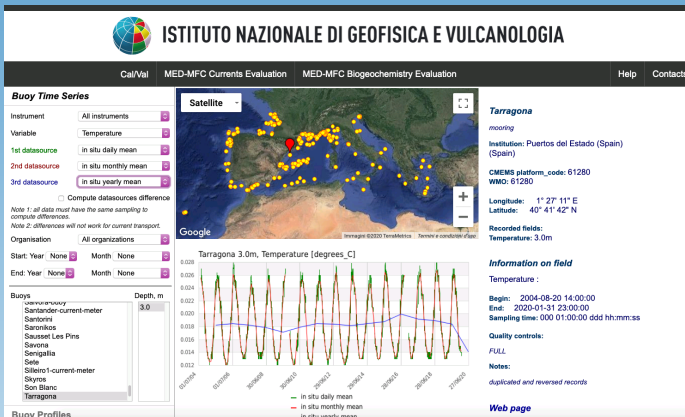
Merged data before the publishing phase **must** be checked for duplicates and reprocessed by passing several tests:

### SOURCE Q/C procedures

- Global range check;
- Spike test;
- Stuck value test;
- Out of statistics and low probable iterative tests.

# Pre-operational data publishing service

Upcoming web service deployed at INGV → Easy access, view and download processed data. More functionalities are planned.






## Merging case test: the Lesvos station

Example of data merging at the HCMR Lesvos Platform.

- Parameter: sea temperature
- Depth: 3.0m
- Time range: 2005-01-01 → 2007-12-31
- Data type: in situ daily mean

# CMEMS data at Lesvos station



## ISTITUTO NAZIONALE DI GEOFISICA E VULCANOLOGIA

Cal/Val
MED-MFC Currents Evaluation
MED-MFC Biogeochemistry Evaluation
Help
Contacts

### Buoy Time Series

Instrument: All instruments

Variable: Temperature

1st datasource: in situ daily mean

2nd datasource: None

3rd datasource: None

☐ Compute datasources difference

Note 1: all data must have the same sampling to compute differences.  
Note 2: differences will not work for current transport.


Organisation: All organizations

Start: Year: 2005 Month: None

End: Year: 2008 Month: None

Buoys	Depth, m
Lesvos	3.0
Lion	10.0
Marseille	20.0
Mykonos	30.0
Nice	35.0
Nice et Dyfamed	40.0
Pyllos	45.0
Santorini	
Saronikos	
Sete	
Silleiro2-current-meter	

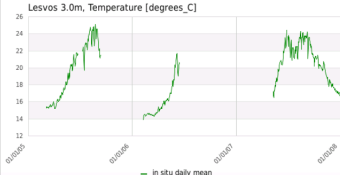
Satellite



Google

Immagini ©2020 TerraMetrics | Tempi in condizioni d'uso

### Lesvos 3.0m, Temperature [degrees\_C]



— in situ daily mean

### Lesvos

*moored surface buoy*

Institution: HELLENIC CENTER FOR MARINE RESEARCH (HCMR) (Greece)

platform\_code: LESVO  
WMO: 6101004

Longitude: 25° 48' 25" E  
Latitude: 39° 09' 22" N

**Recorded fields:**  
Temperature: 3.0 10.0 20.0 30.0 35.0 40.0 45.0m  
Salinity: 3.0m

### Information on field

Temperature :

Begin: 2001-01-29 12:00:00  
End: 2012-07-28 12:00:00  
Sampling time: 000 03:00:00 ddd hh:mm:ss

**Quality controls:**  
FULL  
Notes:

# SeaDataCloud data at Lesvos station



ISTITUTO NAZIONALE DI GEOFISICA E VULCANOLOGIA

Cal/Val

MED-MFC Currents Evaluation

MED-MFC Biogeochemistry Evaluation

Help

Contacts

## Buoy Time Series

Instrument: All instruments

Variable: Temperature

1st datasource: in situ daily mean

2nd datasource: None

3rd datasource: None

☐ Compute datasources difference

Note 1: all data must have the same sampling to compute differences.

Note 2: differences will not work for current transport.

Organisation: All organizations

Start: Year: 2005  Month: None

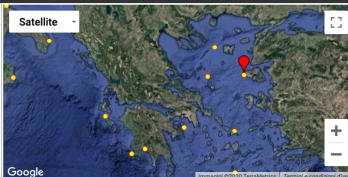
End: Year: 2008  Month: None

Buoys

Lesvos  
Livorno  
MACUB  
MAMBO 10M 00001  
MAMBO1  
Molo Bandiera  
MYKONOS  
Orlona  
Otranto  
Palermo  
Porto Empedocle  
Drahn, Tuzen

Depth, m

3.0



## Lesvos

moored surface buoy

Institution: Hellenic Centre for Marine Research, Hellenic National Oceanographic Data Centre (HCMR/HNODC) (Greece)

platform\_code: 269-  
GN36200707LES\_00001\_269\_D90  
WMO:

Longitude: 25° 48' 29" E  
Latitude: 39° 09' 22" N

Recorded fields:  
Temperature: 3.0m  
Salinity: 3.0m

## Information on field

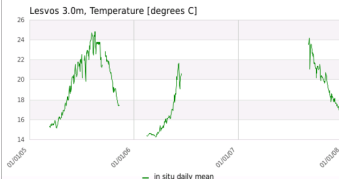
Temperature :

Begin: 2004-09-02 06:04:34  
End: 2011-12-31 20:47:31  
Sampling time: 000 03:12:29 ddd hh:mm:ss

Quality controls:

FULL

Notes:



# Integrated data set at Lesvos station



## ISTITUTO NAZIONALE DI GEOFISICA E VULCANOLOGIA

Cal/Val
MED-MFC Currents Evaluation
MED-MFC Biogeochemistry Evaluation
Help
Contacts

### Buoy Time Series

Instrument: All instruments

Variable: Temperature

1st datasource: in situ daily mean

2nd datasource: None

3rd datasource: None

☐ Compute datasources difference

Note 1: all data must have the same sampling to compute differences.

Note 2: differences will not work for current transport.

Organisation: All organizations

Start: Year 2005 Month None

End: Year 2008 Month None

Buoys	Depth, m
FI3520050601412079	3.0
FI3520050601412200	10.0
Gijon	20.0
Gironde Bordeaux	30.0
Gironde Libourne	35.0
Gironde Pauillac	40.0
Gironde Portets	45.0
Imperia	
Kalamata	
Koper	
Lesvos	

Satellite



### Lesvos

*moored surface buoy*

Institution: Hellenic Centre for Marine Research, Hellenic National Oceanographic Data Centre (HCMR/HNODC) (Greece)

platform\_code: LESVO\_269-GN36200707LES\_00001\_269\_D90

WMO:

Longitude: 25° 48' 27" E

Latitude: 39° 09' 22" N

**Recorded fields:**

Temperature: 3.0 10.0 20.0 30.0 35.0 40.0

45.0m

Salinity: 3.0m

**Information on field**

Temperature :

Begin: 2001-01-29 12:00:00

End: 2012-07-28 12:00:00

Sampling time: 000 03:00:00 ddd hh:mm:ss

**Quality controls:**

FULL

Lesvos 3.0m, Temperature [degrees\_C]



— in situ daily mean



Web service video (please check for version 1 to see the video separately if any problems)

click





## Conclusions

When the web service will be online and the production chained, SOURCE will allow a continuous monitoring of the coastal environment through the widest and highest quality Temperature and Salinity observational time series database.

This is the base to develop new services and applications to serve a variety of different users, such as integrated coastal monitoring systems, early-warning system for coastal environmental protection and preservation.



# Thank you for your attention

