SOCIB Balearic Islands Coastal Observing and Forecasting System

Data Products Catalog
User's Manual

SOCIB-Data Center Facility

Document type:	Product User Manual
Date:	2019-07-03

Description:	Data Products Catalog user manual
Authors:	protllan, jfernandez
Supervision:	jfernandez
Involved Personnel:	Data Center
URL:	http://repository.socib.es/repository/entry/show?entryid=927c8211-610f-45fd-b63f-d844f135655c

www.socib.es

Parc Bit. Ctra. Valldemossa, km. 7,4 · Edifici Naorte, Bloc A · Pl. 2a, P. 3 · 07121 Palma (Illes Balears, Espanya) · Tel: +34 971 43 99 98 · Fax: +34 971 43 99 79

DOCUMENT VERIFICATION LIST

Date:	Checked by (name)	SOCIB division	Ref.

DOCUMENT DISTRIBUTION LIST

Date:	Distribution to:
2019-07-03	All SOCIB organization

CHANGE RECORD

#	Date	Description	Author	Checked by
0.1	2019-05-17	First version document	jfernandez	
1.0	2019-07-03	Detailed documentation.	protllan	jfernandez

www.socib.es

Parc Bit. Ctra. Valldemossa, km. 7,4 · Edifici Naorte, Bloc A · Pl. 2a, P. 3 · 07121 Palma (Illes Balears, Espanya) · Tel: +34 971 43 99 98 · Fax: +34 971 43 99 79

Index of contents

Introduction **4**

Overall description **4**

Main features **6**

Parc Bit. Ctra. Valldemossa, km. 7,4 · Edifici Naorte, Bloc A · Pl. 2a, P. 3 · 07121 Palma (Illes Balears, Espanya) · Tel: +34 971 43 99 98 · Fax: +34 971 43 99 79

1. Introduction

The "**S**istema de **O**bservación y Predicción **C**ostero de las **I**slas **B**aleares" (SOCIB) - Coastal Ocean Observing and Forecasting System of the Balearic Islands - is a multi-platform distributed and integrated system that provides streams of oceanographic data and modelling services. It supports operational oceanography in a European and international framework and contributes to the needs of marine and coastal research in a global change context. SOCIB coordinates the deployment and data management of a wide range of equipments and models from 8 facilities. It also manages data from external international institutions and collaborates with international aggregators for the dissemination of ocean data.

The data coming from SOCIB multiplatform system is processed (quality control included), stored and distributed in NetCDF format. The operation covers three main datastreams (delayed time, delayed mode and real time) and for each, two or more processing levels (L0, L1 etc) that are later exposed through a variety of public services. In this line, an Application Programming Interface (SOCIB API) has been released recently to ease SOCIB overall data discovery. Whereas APIs are widely known among developers, most regular users are not used to them. Therefore, some in-between move is needed in order to tackle both user profiles and boost the use of the underlying data.

SOCIB Data Products Catalog (http://apps.socib.es/data-catalog/), a browser-like interface built on top of the services rooted in the SOCIB API, provides an easy and friendly gateway for regular users to find, download and plot the data available at SOCIB. It stands also as a living example of how-to interrogate and exploit SOCIB API services for developers to reproduce and customize. Next, a detailed description of such application is provided.

2. Overall description

SOCIB Data Products Catalog is a browser-like application built on top of the SOCIB API services. In particular, Data Products Catalog focus mainly on the data-products and data-sources concepts to ease the discovery, exploration and download the of deployment-based/platform-based datasets available at SOCIB:

- Data-Products refer collections of datasets produced in the frame of a given project, campaign or observation program. The grouping of datasets in

www.socib.es

Parc Bit. Ctra. Valldemossa, km. 7,4 · Edifici Naorte, Bloc A · Pl. 2a, P. 3 · 07121 Palma (Illes Balears, Espanya) · Tel: +34 971 43 99 98 · Fax: +34 971 43 99 79

products serves a single purpose: to compress the underlying whole stock to ease discovery (up to 786 datasets vs 167 products).

- Data-Sources refers the datastreams produced by each multi-platform elements deployed in the ocean to recover data over time. All the data produced by a data-source is preserved at SOCIB as collections of NetCDF files and all-in-one are considered to coform a single dataset.

Pivoting on these two main concepts, Data Catalog comprehends two main pages:

- The landing page.
  This page is the defaulting access point for users to discover all SOCIB products. It is actually a paginated product listing that exposes for each one some distinctive metadata and enables filtering. The overall layout defines, no doubt, the application name ("Data Catalog").

- The product detail view.
  This page focus on a particular product, exposing key information to the whole set of datasets (product metadata) and each particular one (dataset metadata and data) composing the product. This very same view can be found under a different domain (doi.socib.es) as part of the product metadata enables traceability: citation, unique identifier etc.

A more detailed description of each of the above views can be found in the "main features" section below.

Be aware that in both, the landing page and the product detail view, a legal notice is provided always at the top-right corner in a collapsible panel linked to a *legal-notice* button (Figure 1).

Figure 1. Legal-notice collapsible panel and button

www.socib.es

Parc Bit. Ctra. Valldemossa, km. 7,4 · Edifici Naorte, Bloc A · Pl. 2a, P. 3 · 07121 Palma (Illes Balears, Espanya) · Tel: +34 971 43 99 98 · Fax: +34 971 43 99 79

3. Main features

The *product-paginated* list and *form* (Figure 2) are the landing page main components.

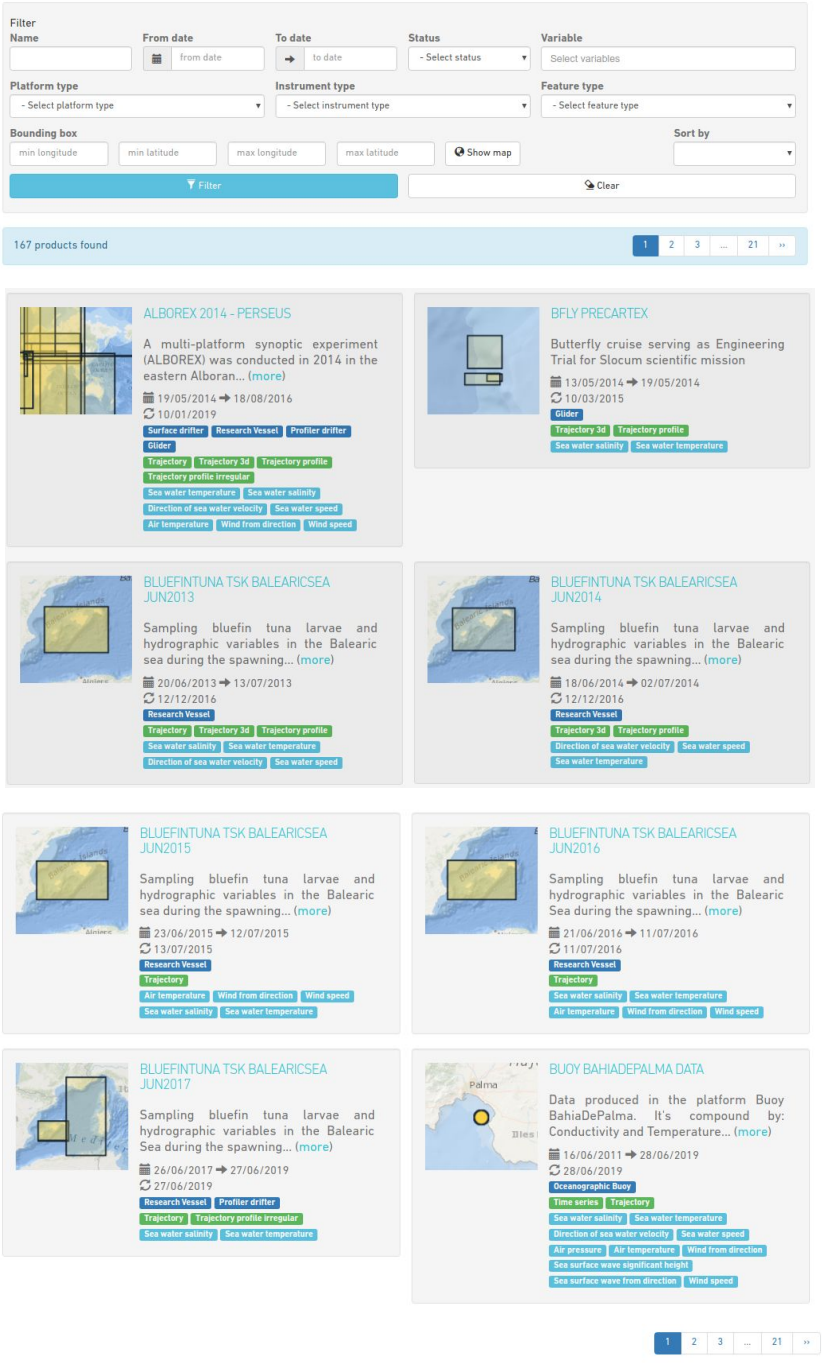

Figure 2. Landing page main components.

www.socib.es

Parc Bit. Ctra. Valldemossa, km. 7,4 · Edifici Naorte, Bloc A · Pl. 2a, P. 3 · 07121 Palma (Illes Balears, Espanya) · Tel: +34 971 43 99 98 · Fax: +34 971 43 99 79

The *form* (Figure 3) comprehends several elements to set restrictions over the defaulting *product-paginated list* given below by:

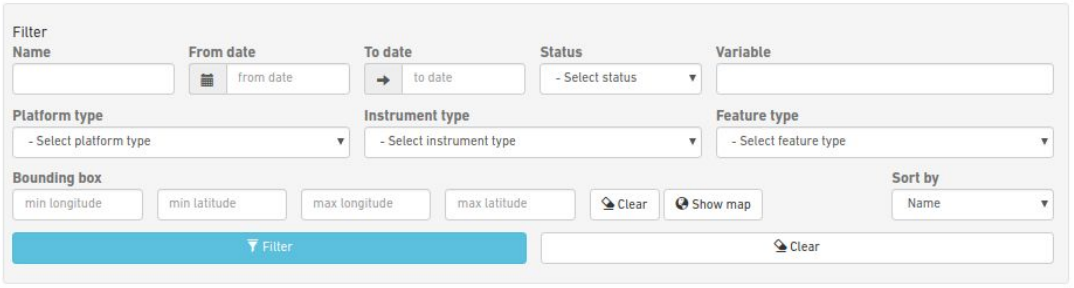

Figure 3. Form elements.

- *name*: input to introduce some text. On form submit, only those product names matching completely or partially the text introduced here will be preserved in the paginated-product list below.

- *temporal coverage*: two inputs labeled as '*From date*' '*To date*' to introduce dates formatted as DD/MM/YYYY. A calendar shortcut is provided on clicking inside the input (Figure 4). On form submit, only those products containing datasets overlapping completely or partially dates introduced here will be preserved in the paginated-product list below.

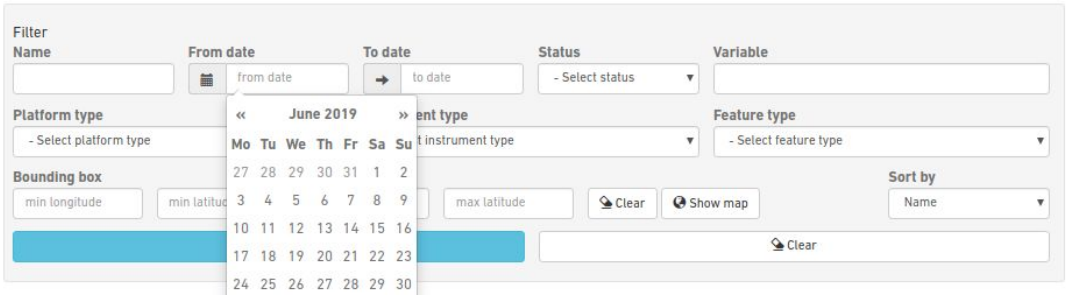

Figure 4. Calendar shortcut displayed on temporal coverage inputs click

- *status*: single select option listing all possible status of the datasets available at SOCIB. On form submit, only those products containing datasets in such status will be preserved in the paginated-product list below.

- *variable:* multiple select option listing all variables contained in the datasets available at SOCIB. On form submit, only those products containing datasets in such variables will be preserved in the paginated-product list below. It solves as an *'and'*.

www.socib.es

Parc Bit. Ctra. Valldemossa, km. 7,4 · Edifici Naorte, Bloc A · Pl. 2a, P. 3 · 07121 Palma (Illes Balears, Espanya) · Tel: +34 971 43 99 98 · Fax: +34 971 43 99 79

- *platform*: single select option listing all possible platforms deployed at SOCIB. On form submit, only those products containing datasets whose data-source comprises such platform will be preserved in the paginated-product list below.

- *instrument*: single select option listing all possible instruments deployed at SOCIB. On form submit, only those products containing datasets whose data-source comprises such instrument will be preserved in the paginated-product list below. Be aware that platform-based datasets will be overlooked right away (these kind of datasets are instrument-agnostic).

- *features*: single select option listing all possible features or sampling geometries produced by SOCIB. On form submit, only those products containing datasets with at least one NetCDF matching this feature will be preserved in the paginated-product list below.

- *bounding-box*: four inputs elements corresponding to latitude minima, latitude maxima, longitude minima and longitude maxima geographical coordinates. To be introduced as integers or floats (dot-separated). A shortcut is provided via "*show-map*", where this bounding-box coordinates can be set drawing a rectangle over a map (Figure 5). On form submit, only those products containing datasets with at least one NetCDF geographically overlapping this bounding-box will be preserved in the paginated-product list below. Be aware that If any inputs is left empty, this restriction is ignored.

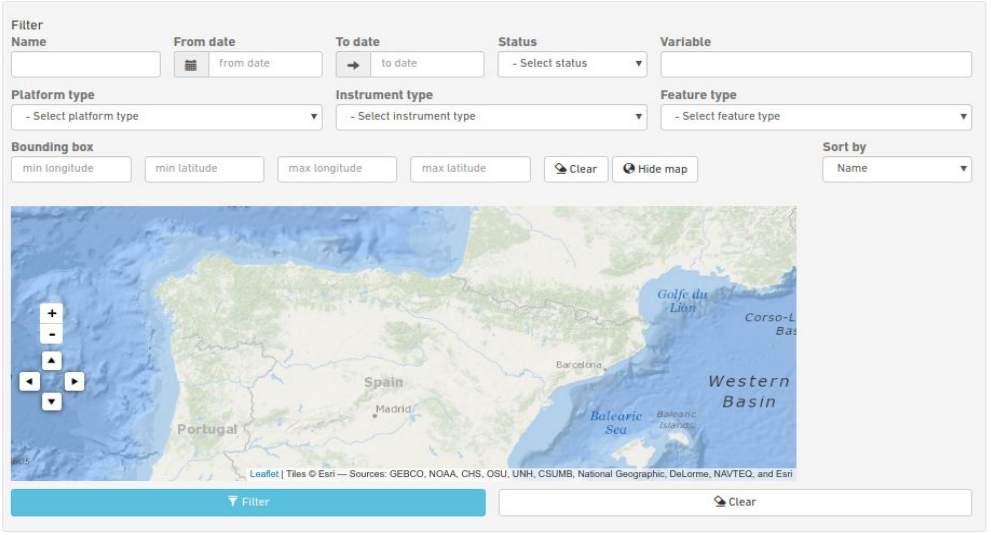

Figure 5. Map shortcut displayed on "show-map" click.

www.socib.es

Parc Bit. Ctra. Valldemossa, km. 7,4 · Edifici Naorte, Bloc A · Pl. 2a, P. 3 · 07121 Palma (Illes Balears, Espanya) · Tel: +34 971 43 99 98 · Fax: +34 971 43 99 79

A sorting utility is also provided via form to re-order the paginated product-list as single-select labeled as “sort-by”. Form submission is triggered by clicking on ”Filter” bottom button. Reset is triggered by clicking on “Clear” bottom button.

Below the form, the paginated product list contains:

- two page navigation elements located before and after the list of product panels. As the layout is set to contain a maxima of 8 product panels per page, this element will have as many tabs/pages as groups of 8 products can be paged together.

- a product-matching counter next to the top page-navigation element (Figure 6) that announces the number of products matching the restrictions set via the just-submitted as: “*{number} products found*”. This *{number}* will change to reflect overall form restrictions matches on above form submission (on “Filter” click) or resetting (on “Clear” click).

Figure 6. Product-matching counter detail.

- a maxima of 8 product panels (Figure 7) per page showcasing distinctive product metadata such as name, collapsible-like description, time coverage , spatial coverage (rendering on a map the geographical bounding box of the datasets composing the product), freshness (last dataset update) and overall involved platforms, feature types and variables in colored-like labels (unique lists computed over all datasets composing the product).

Figure 7. Product panel detail

www.socib.es

Parc Bit. Ctra. Valldemossa, km. 7,4 · Edifici Naorte, Bloc A · Pl. 2a, P. 3 · 07121 Palma (Illes Balears, Espanya) · Tel: +34 971 43 99 98 · Fax: +34 971 43 99 79

The detail view (Figure 8) of the product is structured in 2 main components:

- the product metadata section
- the data-sources metadata & data (dataset) section

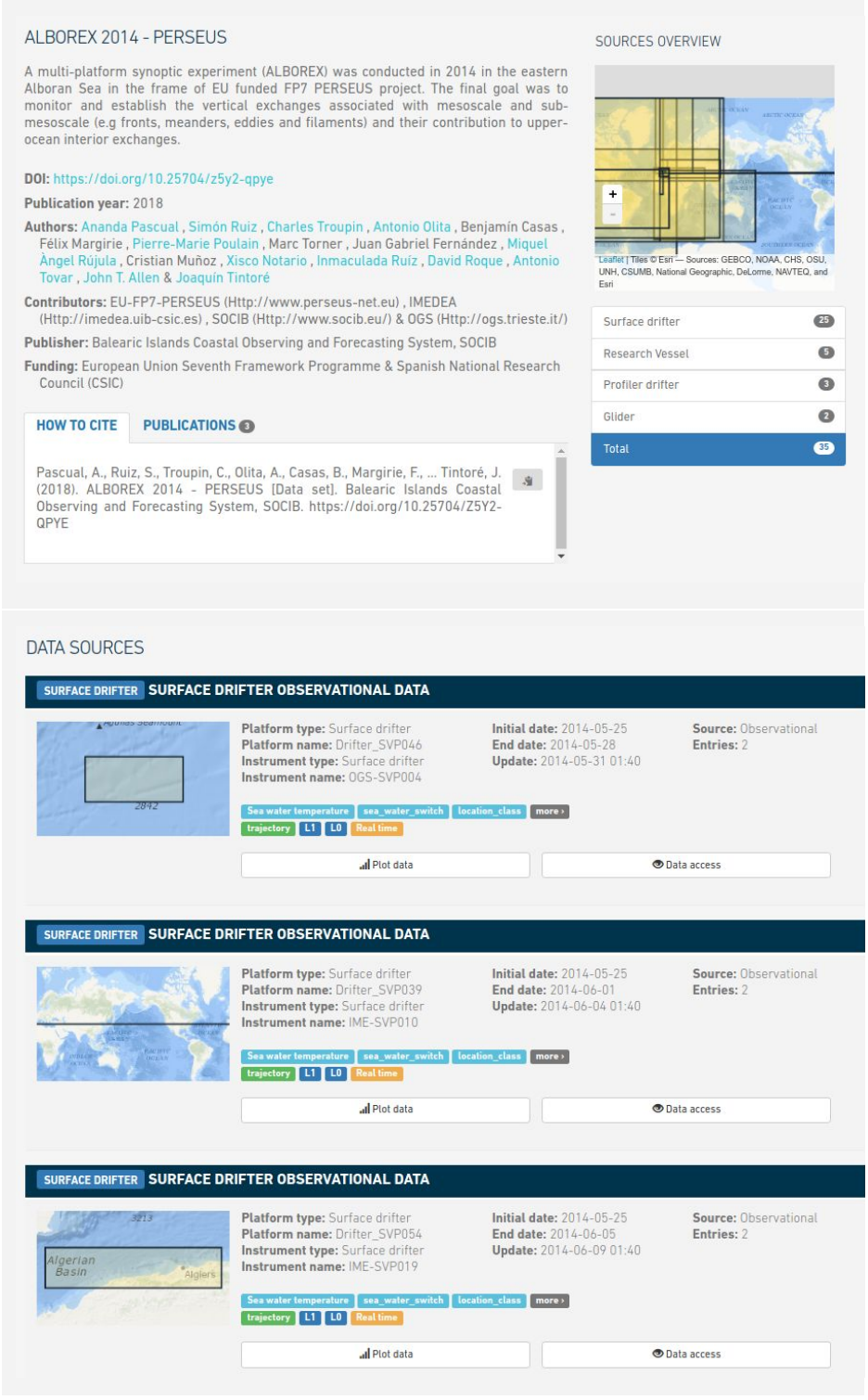

Figure 8. Product detailed view main sections

www.socib.es
Parc Bit. Ctra. Valldemossa, km. 7,4 · Edifici Naorte, Bloc A · Pl. 2a, P. 3 · 07121 Palma (Illes Balears, Espanya) · Tel: +34 971 43 99 98 · Fax: +34 971 43 99 79

The product metadata section (Figure 9) is dedicated to expose:

- a **description** or brief contextualization of the dataset grouping (project, campaign, observation program etc).

- a **minima overview of the data-sources** composing the product: counter/filter of platforms types and its bounding boxes. The counter acts as the only filter available to perform subsetting platforms type over the data-sources listed below.

- **traceability**: unique identifier (DOI), citation rules, publications etc.
  This metadata is common to all products but is meant to be more comprehensive in case of data products with DOIs. Data products with DOI are subject of an strict versioning policy in order to guarantee reproducibility any time such data product is requested. Therefore, a new version of the data product will be issued only if the values or sampling-rule of an already published data is modified/altered. That way, **no versioning** will be issued in case of additions of new files, new variables, new appended data (evolving datasets). The different versions of a data product will include the time range in which such version is valid. In addition, each version will include a *changelog* file with appropriate information about the changes that were applied respecting the previous version. Only the latest version of the product will be available directly. Older versions will be available under request.

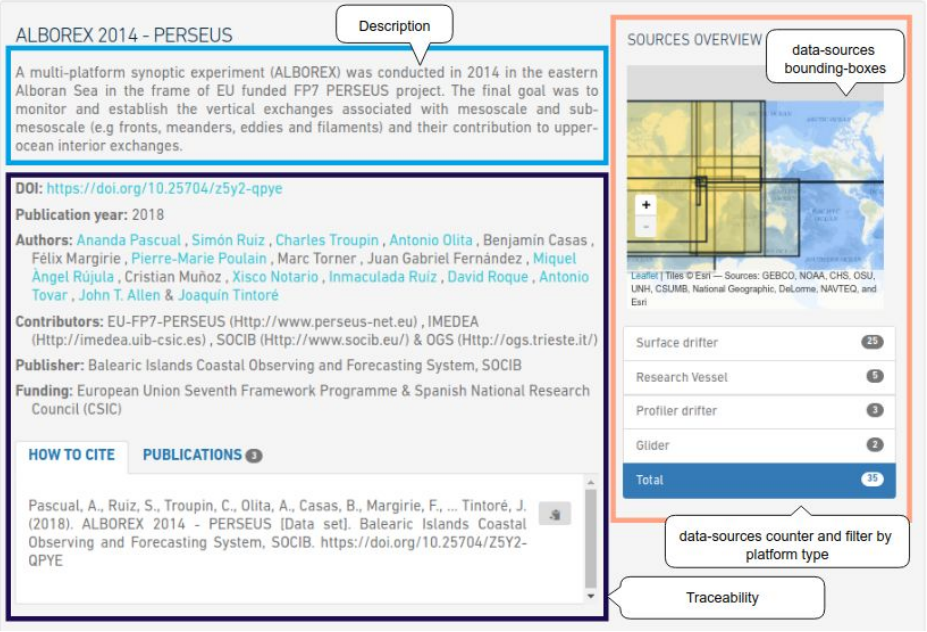

Figure 9. Product metadata ensuring traceability

The section dedicated to the data-sources metadata and data (dataset) is emplaced right below the product metadata and preceded by a '*DATA SOURCES*' header (Figure 10).

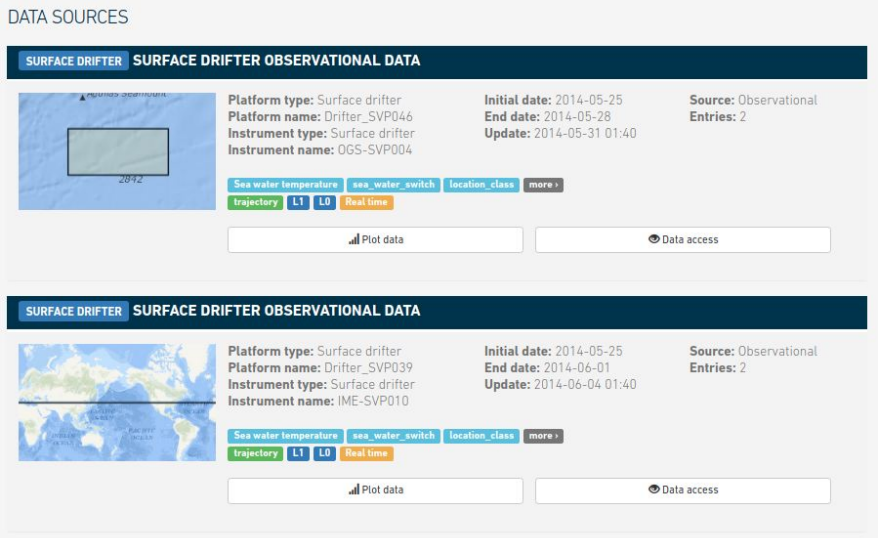

Figure 10. Section dedicated to a given product data-sources (metadata and data)

A navy blue colored header (Figure 11) mark each data-source overview referring both, the data-source platform (light blue) and the resulting dataset name (white).

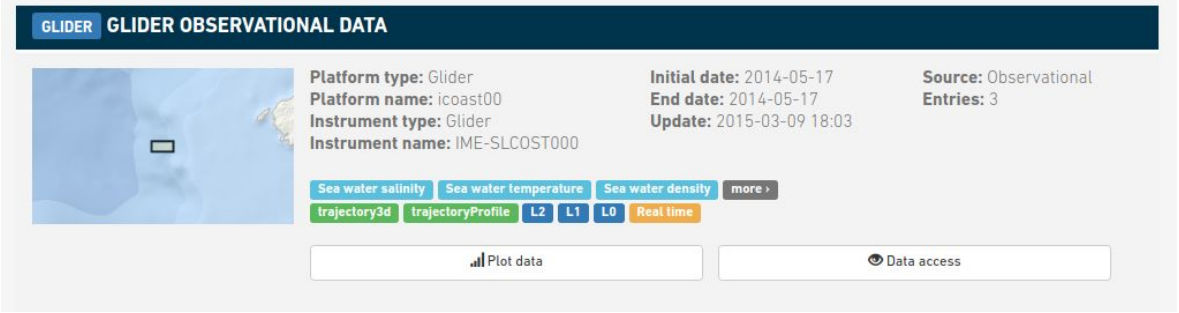

Figure 11. A given data-source metadata and data section details

At the data-source level, the metadata is presented similarly to the product panels in the landing page; exposing each resulting dataset time coverage, spatial coverage (rendering on a map the geographical bounding box of the dataset), freshness (last dataset update) and overall available *variables*, *processing levels*, *data modes* and *feature types* in colored-like labels.

Based on the data-sources metadata is possible to distinguish three kind of datasets:

- deployment-based datasets results from an specific instrument and platform combo deployed somewhere somewhile. For this kind of datasets, both the instrument type and instrument name will be therefore showcased on the metadata section (figure 11).

- aggregation datasets results from the concatenation overtime of the deployment-based datasets by disregarding the use of an specific instrument unit in each deployment. Aggregations are therefore same platform and same instrument type long time series. For this kind of datasets, no instrument name is showcased on the metadata section (figure 12). Aggregated datasets will appear always at the very bottom of the overall dataset list.

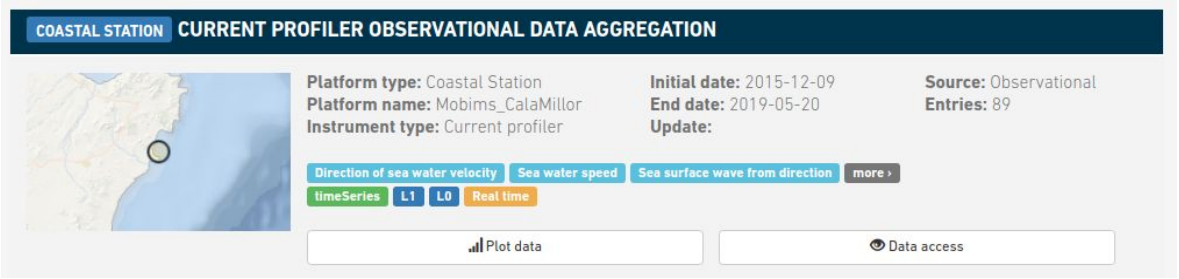

Figure 12. Aggregations metadata and data section details

- platform-based datasets results from the combination of several datastreams coming from the different instruments attached to the same platform. This kind of dataset is also an aggregation but, its data is computed over more than one instrument. In this case, no instrument specification (no name or type) is therefore showcased on the metadata section (figure 13). This is the typical situation of observations performed through different instruments such as the *sea level observations* (barometer and tide gauge).

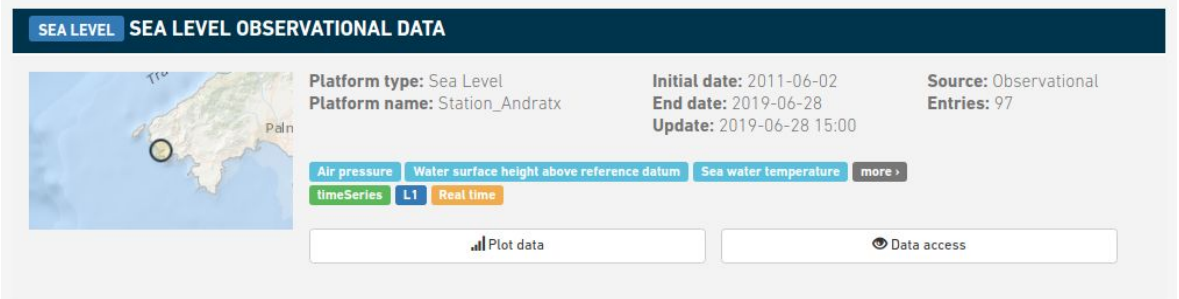

Figure 13. Metadata and data section details for platform-based datasets

For a quick access to all available aggregations (platform-based and instrument-based) use the shortcut *'Aggregations',* available at the sources overview (figure 14).

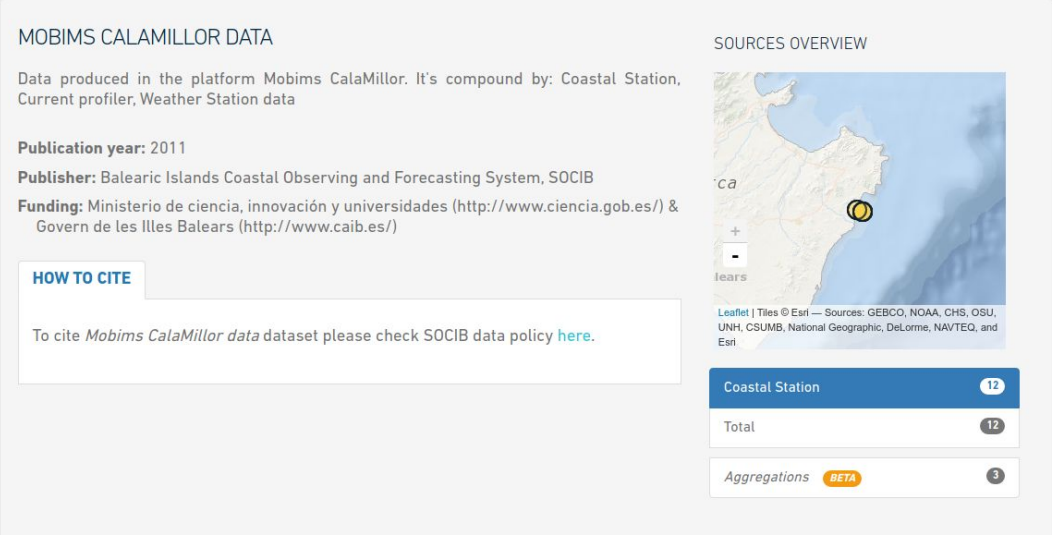

Figure 14. "Aggregations" shortcut on the sources overview at the product metadata section

Furthermore, the data-source section enables also the data exploration and download via the "Plot data" and "Data access" buttons respectively:

- Plot data button.
  When clicking the "Plot data" button a panel to query and plot the variables available in the dataset is provided (Figure 15).

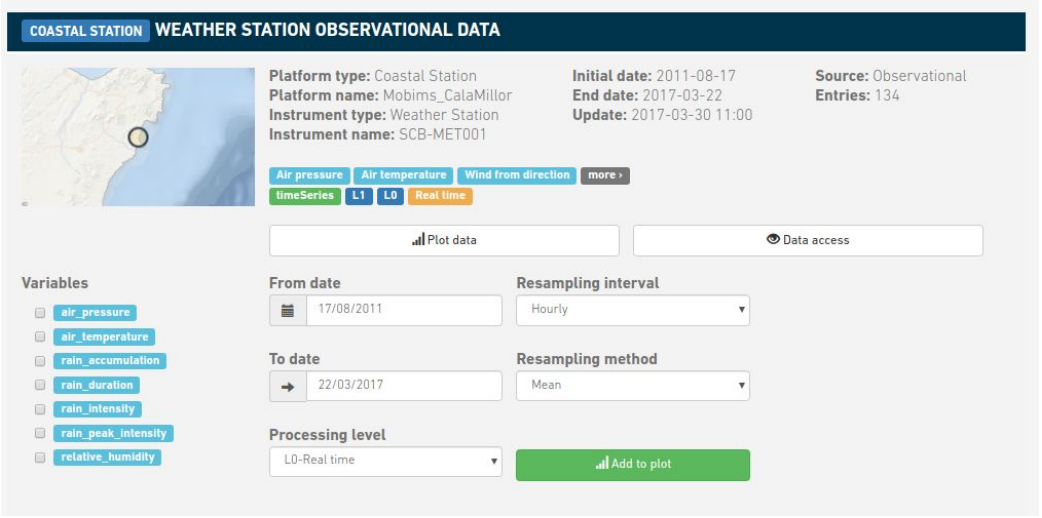

Figure 15. "Plot data" panel detail

www.socib.es
Parc Bit. Ctra. Valldemossa, km. 7,4 · Edifici Naorte, Bloc A · Pl. 2a, P. 3 · 07121 Palma (Illes Balears, Espanya) · Tel: +34 971 43 99 98 · Fax: +34 971 43 99 79

Users can select a given variable from a specific data mode and processing level (see that variable list updates when selecting a different and processing-level & data-mode combo) and a certain time range to visualize the data in interactive charts. Resampling intervals and methods selectors enables users to perform also on-the-fly operations over the data, like for example, the monthly mean of a given variable (Figure 15). For trajectory-linked observations (instruments anchored on profilers, gliders, drifters, vessel etc) a map based plot is provided.

Be aware that gridded data (High Frequency radar data so far) display is yet to be implemented (Figure 16).

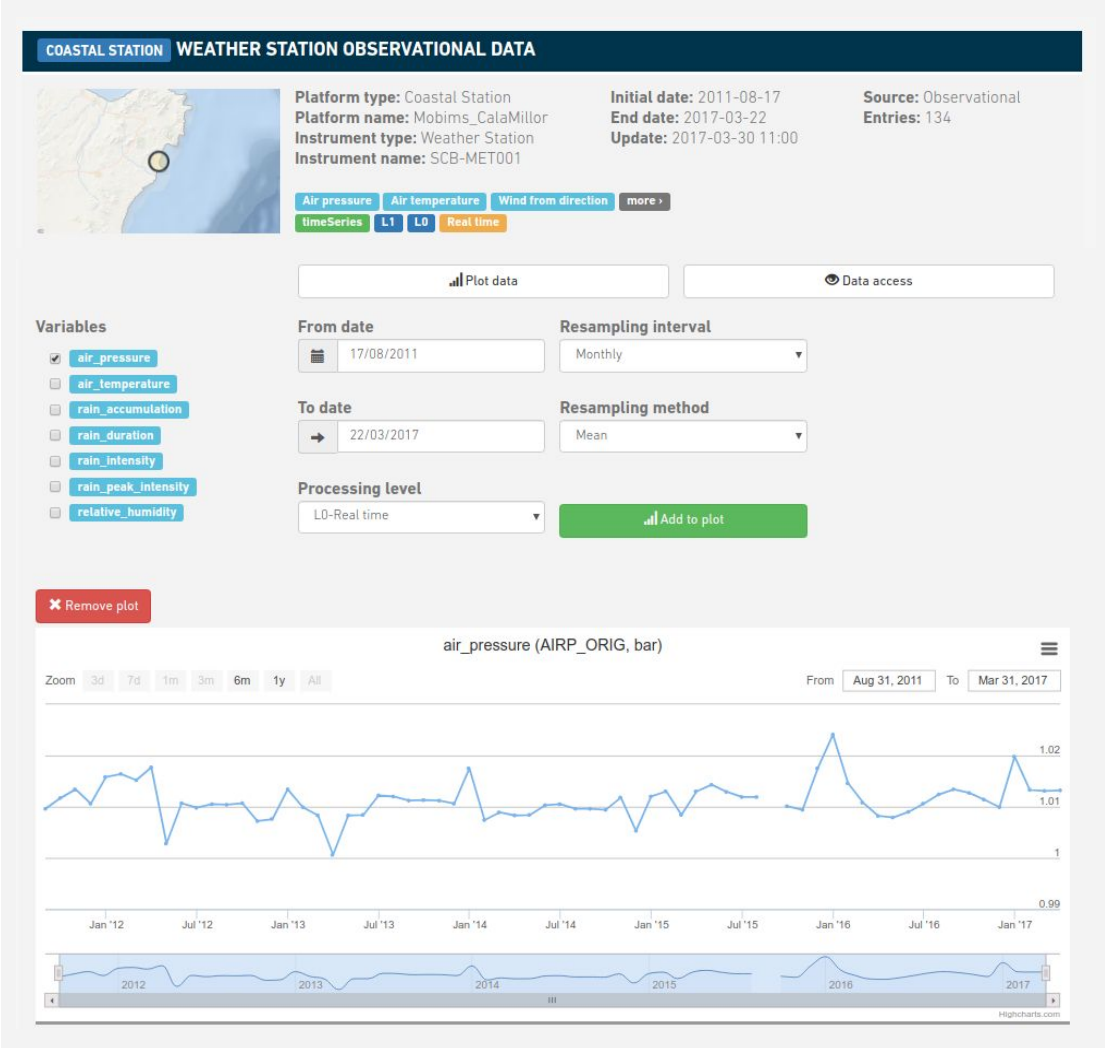

Figure 15. Plotting the air pressure monthly mean of a real time (L0) dataset

www.socib.es

Parc Bit. Ctra. Valldemossa, km. 7,4 · Edifici Naorte, Bloc A · Pl. 2a, P. 3 · 07121 Palma (Illes Balears, Espanya) · Tel: +34 971 43 99 98 · Fax: +34 971 43 99 79

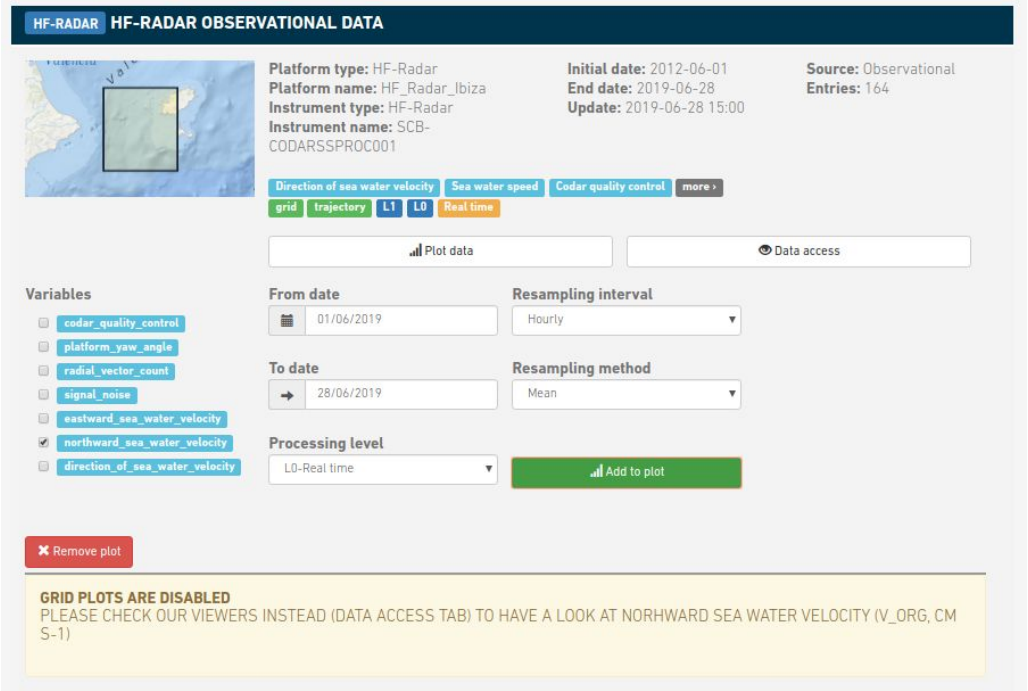

Figure 16. Gridded data plotting disabled

- Data access button.
  When clicking the "*Data access*" a panel listing all the NetCDFs (referred as "entries") conforming the data-source's dataset are displayed and ordered firstly, by processing level and secondly, by services/viewers (Figure 17).

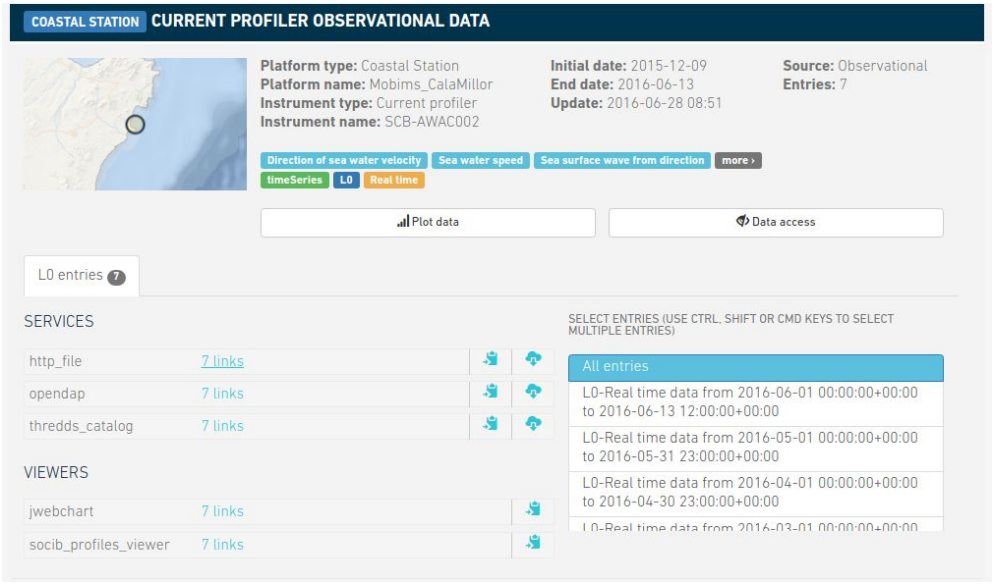

Figure 17. "Data access" panel detail

www.socib.es
Parc Bit. Ctra. Valldemossa, km. 7,4 · Edifici Naorte, Bloc A · Pl. 2a, P. 3 · 07121 Palma (Illes Balears, Espanya) · Tel: +34 971 43 99 98 · Fax: +34 971 43 99 79

For downloading purposes, the service call "*http_file*" (figure 18) is to be used (download will trigger on clicking on the link provided). For exploration purposes, remote services like Opendapp comes handy.

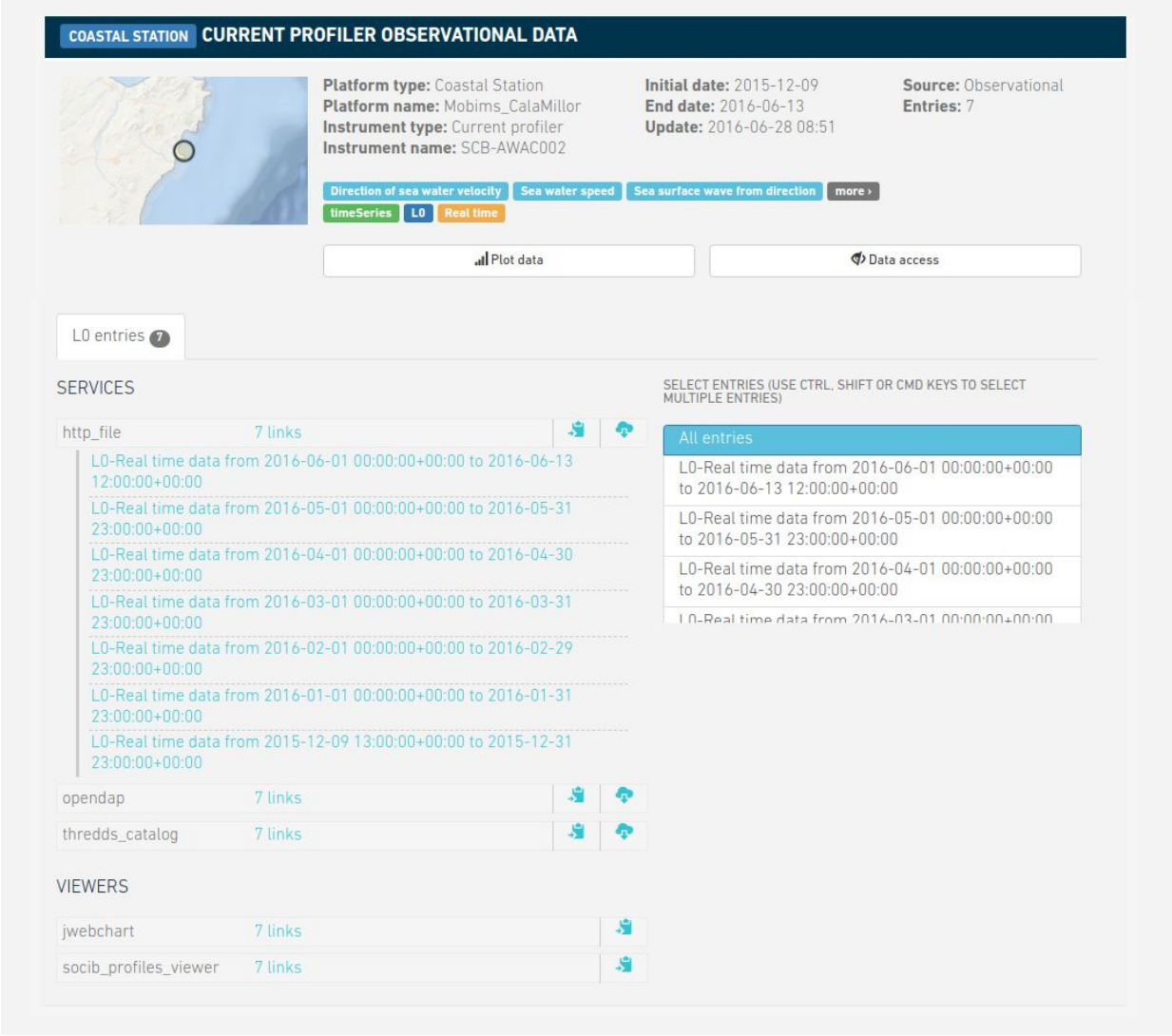

Figure 18. Data Access panel: http_file services list overview.

In addition, a *'viewers'* panel is also provided. This panel comprises a list of links to viewers hosted at SOCIB for getting a quick data glimpse without having to deal with the NetCDF format.

www.socib.es
Parc Bit. Ctra. Valldemossa, km. 7,4 · Edifici Naorte, Bloc A · Pl. 2a, P. 3 · 07121 Palma (Illes Balears, Espanya) · Tel: +34 971 43 99 98 · Fax: +34 971 43 99 79