

Methodology for the collection and processing of national data for analysing user needs in coastal areas and mapping existing products and services.

Action 2021-2-47: Coastal coordination of user needs and methodologies (AD AIR Centre, CNES, ISPRA, IHCantabria; Portugal, France, Italy, Spain)

Table of contents

| | |
|---|--|
| Abstract: | 3 |
| General guidelines for survey design..... | 3 |
| Surveys to collect users' needs | 7 |
| Analysis and prioritization of users' needs..... | 8 |
| Surveys to gather information on existing national services..... | 9 |
| Analysis of state of art of existing national services..... | 10 |
| Identification of knowledge gaps and design of new services | 10 |
| Conclusions and final remarks | 10 |
| Glossary | Errore. Il segnalibro non è definito. |
| Bibliography | 11 |

Abstract

This document describes a methodology, based on previous experiences at national/local level, to collect users' needs for services providing data and information on parameters of interest for stakeholders in coastal areas, and to survey existing services and identifying gaps between demand and offer in order to design new services. The main improvement offered by this methodology is that it is shared among partners belonging to different EU Member States, so that results collected at country level may be easily summed up at Union level.

General guidelines for survey design

Surveys are the most functional tool to collect data and information from a wide range of stakeholders. Therefore, it is imperative to formulate the survey questions with careful consideration of a pre-established plan for analysing the data within the responses. This approach aims to avoid constraints that may arise during the data analysis, such as missing or incomplete answers. The final goal is to ensure the quality of the collected data by mitigating potential misunderstandings, specifically addressing the accuracy of the responses. Additionally, this methodology ensures that the data obtained from various surveys conducted in distinct Member States remains comparable.

Concerning the necessity of avoiding constraints that will make it difficult or impossible to analyse the answers, open-ended questions should be avoided. Multiple-choice questions should be preferred because they yield structured data that's easy to analyse. When it is not possible to use multiple-choice questions, one should provide a complete set of reference terms to be used to formulate the answer (i.e. provide a list of variables names to be used) and clearly specify how to formulate the answer (i.e., ranking the importance of a user's need using numbers between 0 and 10). leaving no room for the inventiveness of the consulted stakeholder. This will enable a homogeneous processing of all questions, facilitating the comparison of all answers and leaving no space for the creative or innovative input from the consulted stakeholder. Nonetheless, it is still possible to include open-ended questions in the surveys, but they should be limited to very specific and peculiar themes that cannot be tackled in other ways. Since it is not possible to analyze them in a quantitative way, they should only be used to collect qualitative information on issues that are complementary to the surveys or only functional to specific national issues.

It is also important to ensure the quality of collected answers to the questionnaires. Therefore, it is very important to clearly define the meaning of each term and of each request in the survey. For this, a dedicated Glossary section should be included in the survey to define and clarify any technical term, acronym, or specialized wording used, ensuring a shared understanding of the survey content and minimizing confusion, such as terms regarding Copernicus services and related terminology. To ensure consistency into the terminology and language commonly used within the marine domain is recommended to utilize existing resources as reference glossary, such as the one provided by Copernicus Marine^{1,2}.

Finally, for avoiding contradictions in the answers and facilitating the comprehension of the questions, queries in the survey should be grouped according to themes or topics. For example, categorizing questions related to specific Copernicus services, user's preferences, challenges faced, and suggestions for improvements. Surveys are generally completed in autonomy, anyway in specific case it is possible to answer them during an interview. During an interview the interviewer asks to the user the questions of the survey, but the interviewer may also provide support by better explaining what is not clear to the user.

¹ <https://marine.copernicus.eu/explainers/glossary>

² <https://marine.copernicus.eu/glossary>

Concerning the homogenisation of surveys among different initiatives in different Member States, it is important to create a common set of reference terms and reference features so that all the answers are homogeneous. We are not suggesting that each survey should tackle all the topics discussed below. But it is important that, when collecting feedback to a questionnaire on one of such specific topics, the answers are all homogeneous.

The names of parameters, services, or products should be as much as possible homogeneous and in line with Copernicus Services' terminology. Furthermore, we propose the following reference terms and categories related to products and services in marine and coastal areas (Table 1).

| | | | | | |
|---------------------|------------------------------------|--------------------|----------------------|---------------------------|--------------------------------|
| Physical parameters | Air temperature | Chemical | BOD5 | Geo-morphological | Backscattering |
| | Atmospheric pressure | | CO2 | | Bathymetry |
| | Gravity acceleration | | COD | | Coastal defences |
| | Hydrostatic pressure | | Colour | | Coastal flooding |
| | Images | | Conductibility | | Coastline |
| | Inherent Optical Properties (IOPs) | | Dissolved Oxygen | | Deposition |
| | Irradiance | | Fluorescence | | Emerged coast characterization |
| | Light intensity | | Hydrocarbons | | Ground motion |
| | Mixed Layer Depth | | Marine litter | | Habitat extent |
| | Noise | | Nutrients | | Kd490 |
| | Other | | Oil Spill | | Land use/Land cover |
| | Radiative fluxes | | Other | | Particulate matter |
| | Rain | | pCO2 | | RRS |
| | Relative humidity | | pH | | Sediment |
| | River flow | | Pollutants | | Soil sealing |
| | Saltwater intrusion | | Salinity | | SSM |
| | Sea current | | Sea surface salinity | | Topography |
| | Sea level | | TRIX | | Transparency |
| | Sea surface current | | Biological | | Abundance of life forms |
| | Sea surface density | Algal blooms | | TSS | |
| | Sea surface temperature | Aquatic fauna | | Turbidity | |
| | Sea bottom temperature | Aquatic flora | | Vegetal coverage | |
| | Sound speed | CDOM | | Fishing vessels | |
| | Stokes drift | Chlorophyll a | | Maritime transport: goods | |
| | Video | Cyanobacteria | | Other | |
| | Water density | Faecal bacteria | | Port traffic | |
| | Water exchange | Phytoplankton | | Raw Satellite data | |
| | Water temperature | Primary production | | Vessel detection | |
| | Wave braking | Yellow substances | | | |
| | Waves | | | | |
| | Wind | | | | |

The technology used to produce data and information may be differentiated according to the following categories listed in Table 2:

| Technology | Description |
|---------------------|--|
| In-situ | This includes all the data collected by using an instrument that is in direct contact with the sample to be measured. This encompasses two scenarios: one where an instrument is moved in direct contact with the sample, and another where a portion of the mean is collected first, followed by the contact between the instrument and the sample occurring at a distance from the origin of the sample. |
| Remote sensing | This refers to all the data collected using an instrument not in direct contact with the sample to be measured. The most recognized data type within this category is derived from satellite observations. Observations from sensors equipped on drones or vehicles belong to this category too. We also consider remote sensing data collected by stationary instruments not in direct contact with the sample, such as coastal radar or coastal cameras used for monitoring beaches. |
| Mathematical Models | This category includes all information obtained from the processing of data through mathematical formulas. There is no direct or indirect contact with samples. The main example of this category are numerical models used to forecast oceanographic parameters or sea state. |

When designing the survey, it is also important to refer to the same features of data and products. For instance, if two different surveys refer to different categories of spatial resolution, analysing them together becomes challenging.

For each parameter we propose to consider the following features:

| Feature | Categories | Description |
|--------------------|---|--|
| Spatial resolution | 0m-5m (coastal engineering, design of infrastructures) | The spatial resolution is a feature describing the distance in space between one data and the closest one. It is important to distinguish between different ranges of resolutions because they refer to different uses of the data or of the information. In the proposed bins the right end is included in the range and the left end is not. |
| | > 5m-50m (local studies at beach/local level) | |
| | > 50m-250m (local studies physiographic units, or municipality level) | |
| | > 250m-500m (studies at Region/Country level) | |
| | > 500m-1Km (local seas, sub-basins) | |
| | > 1km-5Km (basin level) | |
| | More than 5Km (open ocean, National studies) | |

| | | |
|-----------------------|--|---|
| Time resolution | 0'-30' (early warning systems for quick events, i.e.. tsunamis) | The time resolution is a feature describing the distance in time between one data and the next one. It is important to distinguish between different ranges of resolutions because they refer to different uses of the data or of the information. In the proposed bins the right end is included in the range and the left end is not. |
| | > 30'-1h (early warning systems related to waves or flooding) | |
| | > 1h-3h (monitoring of marine parameters) | |
| | > 3h-6h (monitoring of slow varying marine parameters) | |
| | > 6h-1d (management of daily events) | |
| | > 1d-1m (monitoring of intra-seasonal features) | |
| | > 1m-1y (monitoring of seasonal features) | |
| | more than 1y (monitoring of annual trends) | |
| Spatial coverage | | The precise extent of the area covered by the data should be specified |
| Time span covered | | Period of time covered by the data and information |
| Update Time frequency | | This feature, that is only relevant for some type of data, identifies how often an existing data is replaced by a new one for the same area and the same date, i.e., forecasts are often update more than once per day. |
| Accuracy | | This provides and estimation of the error associated with the measured data or processed information. It is important to include within the possible answers the option “not evaluated”, given the great number of national products and services that are not validated yet. |
| Modelling strategy | Re-analysis/analysis (modelled data postprocessed with observations to improve their quality) | This feature regards models-: it is essential to differentiate not only the output of the model but also understand how this information has been produced and its connection to observations. |
| | Hindcast (modelled data for past periods, forcings may be observations) | |
| | Forecast (modelled data for the next future, i.e., last link with observations is at some time in the near past) | |
| | Long term Projections (modelled data for a far future) | |

| | | |
|--|--|--|
| | scenario, i.e., generally no link with observations) | |
|--|--|--|

Ensuring a Representative Sampling

To ensure that the sample surveyed is representative of the ecosystem, it is necessary to ensure that the response rate is sufficiently high to minimize the risk of a biased representation.

To achieve this, preliminary work is required to identify the target population and each user group and the size of the total population. Then, one approach could be to achieve a sample of 10% of the total identified population, this rate must be adapted according to the characteristics of the target population.

This enables analysis of the specific needs of each user group, facilitating the design and targeted improvement of services in coastal areas. By adhering to this methodological principle, it is possible to ensure that the portion of the ecosystem being questioned is representative and yields reliable and generalizable results by sector of interest. These recommendations must be considered in the survey dissemination strategy.

Surveys to collect users' needs

Involvement of users and their participation in the co-development of the services is a key element to ensure a successful implementation of the services and the end-user acceptance. As different stakeholders represent different requirements, the participatory process (survey or interview) should include, as far as possible, the involvement of regulators, companies, and scientists.

In order to collect data from the user filling the survey, the survey/interview must consider at least 6 different issues:

- A general introduction about the aim of the survey/interview. This section must also collect information about the participants (such as name, contact data) and acknowledge their acceptance about the use of personal data for the project.
- Data about his/her affiliation, identification of his/her role (according to the following classes: institutional, research/academia, private non-commercial, commercial) and his/her relationship with the coastal zone.
- Information about the marine sectors they are related to and the services of their interest

| | |
|--|---|
| Coastal and Marine sectors of interest: | General |
| | Fisheries |
| | Emergency response |
| | Climate change |
| | Research and development services |
| | Educational and outreach services |
| | Environmental monitoring |
| | Aquaculture |
| | Tourism and recreational activities |
| | Cultural heritage protection |
| | Maritime transport routes and traffic flows |
| | Urban and landscape planning and management |
| | Design and management of infrastructures |
| Species conservation and protected areas | |

| | |
|--|---|
| | Energy sector (hydrocarbons and renewable energies) |
| | Coastal protection |
| | Raw material extraction |
| | Other |

- Type of user (end-user or intermediate user) and the components of the data used in activities (in situ, modelling, remote sensing)
- Information about the needs and requirements in relation with the spatial and temporal resolution of data and specific parameters, according to the tables in the previous section.
- Information about access and exploitation of the data, including the technical level of the participants and information about data accessibility
- The main problems identified in the use of the data, differentiating between knowledge problems (i.e., selection of the most suitable dataset) and technological problems (i.e., problems in downloading, accessing, or manipulating the data)

Analysis and prioritization of users' needs

When analysing the data collected through the surveys (questionnaires or interviews), the frequency of mention is the most immediate solution to prioritise the requests. Anyway, this may not be the most valuable way. A more detailed analysis of the answers, differentiating by type of user and area of interest, may lead to a more effective design of improvement plans for existing products or the development of new ones. For instance, it is important to differentiate between the request of an intermediate user that has a deep knowledge of Copernicus products, and the one from an end-user that may be unaware of existing products or technological limits. On the other hand, a representative of the institutions will formulate requests to improve the production of public products, while a representative of commercial entities will favour the market turnover of the European investment in Copernicus. Hence, we suggest to not only prioritize the requests according to the overall frequency, but also to analyse the partial frequency in a specific subset of users, according to the specific goal of the analysis. Hence, it is crucial to distinctly identify the users while collecting their requests.

To analyse the gaps raised in the survey, a semi-quantitative method can be used in order to categorise the quality of the data offered. The following aspects should be analysed:

| Technological gaps | Knowledge gaps |
|--|--|
| Complexity of the data | Data reliability |
| Data format | Heterogeneous methodologies in data collection |
| Heterogeneous sources | Incomplete Temporal distribution |
| Data inaccessible or unavailable | Incomplete Spatial distribution |
| Lack of tools to manipulate and visualise the data | Unsuitable resolution |
| | |
| | |

It is important to note that all the gaps raised by the survey should be considered, regardless of their frequency.

Surveys to gather information on existing national services

Surveys for collecting information on existing services and products must be designed according to precise terms of reference. This approach ensures the ability to generalize the findings, and involves the classification of services (in-situ, remote sensing, modelling), the categorization and naming conventions of parameters of interest, reference classes for spatial/temporal resolution, and the identification of producers (institutional, research, commercial, NGOs). In addition to the features already highlighted, it is important to specify:

| | | |
|-----------------------------|--|--------|
| Kind of product supplied | Maps, datasets, reports, decision support tool, other to specify | |
| Availability of the service | Operational since: start date | |
| | Available from start date to end date | |
| Completeness of data | Presence of spatial gaps (yes/no) | |
| | Presence of temporal gaps (yes/no) | |
| Topic | (Same options of the table “Coastal and Marine sectors of interest”) | |
| Data Access | On demand - free of charge (data are made available to the user upon request to the owner, who needs to approve the request before delivering the data) | |
| | On demand – Commercial (data are made available to the user upon request to the owner, who needs to approve the request before delivering the data after payment) | |
| | On demand - mixed (data may be free of charge or commercial depending on the user requesting the data or the declared use of the data) | |
| | Free and open (data are free to be accessed on a web portal, even if the identification of the user is required, i.e., registration required only to identify the user for statistical issues or for the acceptance of a data policy that will not prevent the user from using the data, like done by the Copernicus Services CMEMS, CLMS, etc.) | |
| Delivery tool | Not present (data will be directly sent by producer) | |
| | Only data illustration or post-processed data may be selected and downloaded by the user. Raw data are not accessible directly. | |
| | Post-processed data or raw data may be selected and downloaded by the user. | |
| | The user has the option to directly process the producer's data using the delivery tool and subsequently download either the results or the original data. | |
| | The user may upload data and process them with producer's data directly on the delivery tool and then download the results and/or original producer's data | |
| | In-situ Component/EmodNet | Yes/No |

| | | |
|---|---|----------------------------|
| Linked to Copernicus/Emodnet (Specify if Copernicus data are used as input for producing the product) | Space Component | Yes (specify satellite)/No |
| | Service Component | Yes (specify service)/No |
| User engagement strategies | None, user surveys, workshops and training sessions, online forum, other to specify | |

Analysis of state of art of existing national services

As for user requirements, national products should be analysed according to the frequency of different features.

It may be important to assess which are the most frequently parameters provided in output by national services what are the most widely used technologies or what are the topics with more dedicated services. Next to this also a crossed analysis may be of interest, i.e., verify the most used technology to produce a certain parameter or a certain class of parameters.

Identification of knowledge gaps and design of new services

When the survey has the goal of assessing knowledge gaps, the following aspects should be considered:

- identify the technical parameters that do not meet the required threshold for delivering useful service (i.e., spatial and temporal resolution, spectral bands).
- identify the base parameters that are not being offered by Copernicus Services, but that are necessary to feed the considered downstream services.
- Assess whether the data delivery system meets at least the user basic needs, and the data format is suitable for their applications.

Otherwise, it is possible to identify proxy parameters for matching users' needs and existing services. One of the most functional proxies is spatial resolution, that differentiates among users operating in different domains and, on the other side, is related to different sensors/mathematical tools to produce the data. Time resolution may be useful to identify specific gaps for management of events (i.e., emergencies).

It is important to analyse not only the correspondence between requests and available products, but also to verify the ease of use of existing products with respect to the technological skills of users in a certain domain/topic.

Conclusions and final remarks

Mapping national user needs and existing products is an important task to identify information gaps for marine and land-based applications in coastal areas, which can be used as a baseline to promote recommendations for future development. Scaling up these activities at EU level helps to provide a broader and homogeneous picture of demand and offer and to harmonize the future evolution of Copernicus products to meet the needs of coastal users.

The strength and novelty of the proposed methodology lies in the fact that it's generated from synergic interactions among different discussion groups on coastal areas applications from several EU countries, who

capitalized on the national experiences for agreeing in a common way of collect user needs and analyse the state of art of existing national services.

It is of outmost importance to guarantee in all the countries the heterogeneity, representativity, and transdisciplinary nature of the surveyed users and of their main thematic tasks to reach national comprehensive perspectives that can be subsequently compared and assembles smoothly at the EU level overcoming regional/sectorial fragmentation.

In particular, we highlight the following features of the proposed methodology:

- a detailed identification of users allows to take into account factors such as affiliation and sector of interest. This leads to an in-depth analysis that ensures a comprehensive understanding of user needs.
- a clear definition and a dedicated glossary would improve the accuracy and reliability of survey results, supporting informed decision-making.
- the prioritization of user needs allows the effective design of improvement plans and new services, aligned with user preferences and technological advances.

It is also important to set up a mechanism for continuous review of user needs and feedback loops to promote the sustainability and relevance of methodologies, adapting to technological developments and changing needs.

The collected requirements and mapping of existing services should then be considered onto the design of new services alongside with the following considerations:

- Ensure a smooth user experience on platform navigation and data access, that should be streamlined and user friendly
- Use [common protocols](#) of data access to favour interoperability of dissemination platforms.
- Avoid the redundancy with other services, ensuring that the offered data has been required, and it is not being provided by other sources
- Use a common approach of [data checking, standardization, and quality-control](#).
- incorporating user feedback into the refinement of new services

Last, we highlight the importance of further sharing and agreeing on a common methodology to survey user needs and mapping existing services with other Member States and with Copernicus Entrusted Entities.

Bibliography

Fernandez, F., Álvarez-Fanjul, E., Delgado, J., Díaz-Sánchez, R., Valcárcel, N., Carcía-Flecher, A., Molina, R., Medina, R. (2020). Data services for intermediate users at the coastal zone: analysis of gaps and recommendations for evolution from the Spanish perspective.