Requirements for data to be used in marine spatial planning

Deliverable 3.1.
BONUS BASMATI
Requirements for data to be used in marine spatial planning

August 2018

Authors:
Kerstin Schiele, Wanda Holzhüter, Miriam von Thenen, Institute for Baltic Sea Research Warnemünde
Hanna Luhtala, University of Turku
Solvita Strāķe, Maija Viška, Kristīne Pakalniņe, Latvian Institute of Aquatic Ecology
Christian Koski, Finnish Geospatial Research Institute
Andrea Morf, Nordregio
This report is a publicly accessible deliverable of the BONUS BASMATI project. The present work has been carried out within the project ‘Baltic Sea Maritime Spatial Planning for Sustainable Ecosystem Services (BONUS BASMATI)’, which has received funding from BONUS (art. 185), funded jointly by the EU, Innovation Fund Denmark, Swedish Research Council Formas, Academy of Finland, Latvian Ministry of Education and Science and Forschungszentrum Jülich GmbH (Germany).

This report may be downloaded from the internet and copied, provided that it is not changed and that it is properly referenced. It may be cited as:

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>BONUS BASMATI in Brief</td>
<td>5</td>
</tr>
<tr>
<td>Report Summary</td>
<td>6</td>
</tr>
<tr>
<td>1 Introduction</td>
<td>7</td>
</tr>
<tr>
<td>1.1 Purpose and scope</td>
<td>7</td>
</tr>
<tr>
<td>1.2 Interaction with other work packages</td>
<td>7</td>
</tr>
<tr>
<td>2 Data requirements</td>
<td>9</td>
</tr>
<tr>
<td>2.1 General aspects</td>
<td>9</td>
</tr>
<tr>
<td>2.2 Planner and stakeholder perspective</td>
<td>11</td>
</tr>
<tr>
<td>2.3 Case studies perspective</td>
<td>13</td>
</tr>
<tr>
<td>2.4 The Baltic Explorer perspective</td>
<td>15</td>
</tr>
<tr>
<td>3 Proposals for data requirements</td>
<td>17</td>
</tr>
<tr>
<td>3.1 Data categories</td>
<td>17</td>
</tr>
<tr>
<td>3.2 Relations between data categories</td>
<td>18</td>
</tr>
<tr>
<td>3.3 Descriptive data properties</td>
<td>19</td>
</tr>
<tr>
<td>3.4 Guideline on metadata</td>
<td>21</td>
</tr>
<tr>
<td>4 Conclusion</td>
<td>23</td>
</tr>
<tr>
<td>References</td>
<td>24</td>
</tr>
</tbody>
</table>
BONUS BASMATI In Brief

BONUS call 2015:
Blue Baltic

Project coordinator:
Henning Sten Hansen, Aalborg University, Denmark

Project partners:
Aalborg University, Denmark (AAU)
Aarhus University, Denmark (AU)
Finnish Geospatial Research Institute, Finland (FGI)
Latvian Institute of Aquatic Ecology, Latvia (LIAE)
Leibniz Institute for Baltic Sea Research Warnemünde, Germany (IOW)
Nordregio, Sweden (Nordregio)
University of Turku, Finland (UTU)

Duration:
3 years, 7/2017 – 6/2020

Key theme addressed:
Theme 4.3 Maritime spatial planning from local to Baltic Sea region scale

Subthemes:
Theme 2.3 Integrated approaches to coastal management and Theme 4.1 Governance structures, policy performance and policy instruments
https://www.bonusportal.org/projects/blue_baltic_2017-2020

Project abstract:
Maritime Spatial Planning (MSP) requires a spatially explicit framework for decision-making, and on that background the overall objective of BONUS BASMATI is to develop integrated and innovative solutions for MSP from the local to the Baltic Sea Region scale. This is to be realized through multi-level governance structures and interactive information technology aiming at developing an ecologically and socio-economically sound network of protected marine areas covering the Baltic Sea. Based on the results of former MSP projects, the BONUS BASMATI project sets out to analyse governance systems and their information needs regarding MSP in the Baltic Sea region in order to develop an operational, transnational model for MSP, while maintaining compliance with existing governance systems. It also develops methods and tools for assessments of different plan-proposals, while including spatially explicit pressures and effects on maritime ecosystem services in order to create the Baltic Explorer, which is a spatial decision support system (SDSS) for the Baltic Sea region to facilitate broad access to information. During the project running until 2020, new data will be produced and tested in assessments corresponding to policy goals. The data will support the combined analysis of the three elements of the concept of ecosystem services: the capacity, flow and benefit of provisioning, regulating and cultural services. A central aim of the project is to facilitate cross-border collaboration, and the project is carried out in close cooperation with relevant stakeholders in the BSR. The impact of the project will be facilitated and assessed in transnational case studies, where integrated solutions are required. The local scale will consist of case study areas in the South-West Baltic, the Latvian territorial and EEZ waters including open part of the Baltic Sea and the Gulf of Riga, and across the region, a pan-Baltic case study will be performed.
This document provides recommendations to structure information within datasets for Marine Spatial Planning (MSP) based on scientific literature, reports of associated MSP projects in the Baltic Sea region (BSR), documents issued by the HELCOM-VASAB Maritime Spatial Planning Working Group and expert knowledge within the BONUS BASMATI project.

The aim is to reach a better interoperability and harmonisation for the inclusion of ecosystem services in MSP processes as well as a joint data management within the case studies anchored in the BONUS BASMATI project. Further, the proposed requirements on data can improve transboundary communication and foster coherent marine spatial plans.

The requirements on data described in this document are reflected and discussed from different perspectives including the planner and stakeholder perspective, the case studies and a technical perspective (Baltic Explorer). Accessibility and understandability of data are central aspects for planners and stakeholders. Highly, harmonised quality standards allow data to be easily and flexibly applicable by the user for different planning purposes and phases in the MSP process. Depending on the perspective, information on spatial dimension, time frame, origin of data, potential conflicts and synergies as well as structuring data is likely to be weighted differently.

The report proposes data requirements to be met to facilitate the inclusion of ecosystem services in MSP processes. First, in order to create a useful structure and to allow an easy search, we suggest to include data categories for datasets corresponding to the ecosystem service framework. Second, we suggest to add a wide range of tags and keywords in datasets to highlight the relation between different categories such as ecosystem services with e.g. associated human activities. Third, we recommend to describe spatial dimension, time frame, origin of data and potential conflicts and synergies in more detail (predefined classes) within datasets. The introduced simple schematic way of structuring information in datasets allows a high flexibility and optional additions by other users, planners and stakeholders depending on the purpose within the MSP process. It fosters transparency about data origin and enhances metadata quality which is crucial for usability of data within MSP. A short and easy understandable general guideline (template) for minimum metadata is included to support the production of high-quality metadata with minimum effort.
1 Introduction

Marine Spatial Planning (MSP) is a complex, data intensive process and its success depends to a large extent on the quality and availability of relevant data and the capacity for analysis (Stamoulis & Delevaux, 2015). Therefore, a good data management is the fundamental pillar for MSP. Data collection, processing, documentation and storage need to be handled with great care. The consent on uniform minimum standards and requirements are a first step towards solid data management providing the grounds for long term and wide applicability of data and processed information.

MSP is challenged to provide comparably detailed information on a wide range of topics that are representative for the MSP process during preparation, planning, adoption and implementation (Carneiro et al., 2017). Carneiro and colleagues note, that this scope of duties is convoyed with inevitable limitations

- Multiple data sources introduce inconsistencies
- Descriptive data fields are less suitable for searching
- Keeping information up to date is a significant challenge
- MSP still suffers from missing data.

Although there have been several initiatives and frameworks to develop common standards within data management, apparently these have never been widely adopted by the community (IOC of UNESCO, 2013).

1.1 Purpose and scope

The aim of this report is to facilitate the inclusion of ecosystem services in marine spatial planning (MSP). MSP-relevant data need to be available, easy to find and usable for further analysis. This report is based on literature, reports of associated MSP projects in the Baltic Sea region (BSR), documents issued by the HELCOM-VASAB Maritime Spatial Planning Working Group and expert knowledge within the BONUS BASMATI project.

This document (D.3.1) is to be understood as a proposal for handling data in MSP processes to enable the inclusion of ecosystem services and as a direct guideline for handling case study data within the project, especially with regard to the project database (D.3.2).

1.2 Interaction with other work packages

Work package 3 is characterised by strong links to several other WPs within BONUS BASMATI (Fig. 1).
WP2 gives requirements on data by stakeholders (in this report via literature and requirements identified by case studies).
The ecosystem service framework (D.4.1) developed in WP4 sets requirements on the availability of data categories reflecting the main themes of the framework.
The Baltic Explorer developed in WP5 is going to use the data supplied by WP3 for pilot processing of data. D.5.1 gives user requirements for SDSS in MSP.
Data generated and analysed in the three case-studies (WP6) will be fed into the WP3 database. All data of the case-studies should comply to the proposed standards. D.6.1 highlights case study data and modelling needs.
Figure 1: Collaboration within the BONUS BASMATI project work packages
2 Data requirements

In the following section explicit requirements on data are discussed from a general perspective as well as from different specific perspectives such as a planner and stakeholder perspective, the case studies and a technical perspective (Baltic Explorer). The following discussion is based on literature and expert knowledge and illustrates the diverse demands each perspective sets on the definition of requirements and the development of standards regarding data in MSP processes. Depending on the perspective, information on spatial dimension, time frame, origin of data, potential conflicts and synergies as well as sorting data into data themes is likely to be weighted differently.

2.1 General aspects

MSP is often said to be an evidence-based process (MSP Data Study, 2016), where data needs to cover a variety of themes and topics and also different types of datasets. Moreover, different information is needed in different phases of the planning cycle, concerning e.g. current situations (stocktaking), future scenarios and visions, as well as policies and planning decisions (Ehler & Douvere, 2009). In general, the MSP process requires information which is usable in policy making. Consequently, not all and any kind of data are useful in MSP as so-called planning evidence. On the other hand, collecting and gathering of datasets and transforming them into evidence useful to answer specific questions of planning is often time-consuming and requires resources. Therefore, in a world of data, where it is possible to gather more and more datasets, data collection should be carefully considered (Ehler & Douvere, 2009). Data collecting and processing may become proportionally expensive if excessive time is spend on those activities, limiting the resources available for other important aspects of spatial planning (Collie et al., 2013; Halpern et al., 2012).

Data are being collected and available in a variety of formats and can be derived from various sources like scientific literature, expert scientific opinions or advice, government sources, local knowledge, and direct field measurement (Ehler & Douvere, 2009). Typically, in a MSP process, most of the datasets used have spatial characteristics but also non-spatial evidence is needed. For instance, economic baseline studies are important to set the objectives in the early stages of an MSP process (MSP Data Study 2016). In a step-by-step approach guide for MSP, Ehler and Douvere (2009) highlight a general rule, according to which data should be up-to-date, objective, reliable, relevant, and comparable. The requirement of relevance, for example, complies with a variety of data types. However, comparability is already a more complicated issue. Besides up-to-date data, also the historical datasets, together with local knowledge and understanding about traditions, can provide valuable insights, for example in defining the changes of the species abundance, diversity or resilience (e.g. Frans & Augé, 2016; von der Heyden, 2017). Moreover, long-term monitoring is essential for an effective and adaptive marine management, enabling adjustments of the plans and measures, guiding future planning processes and identifying new research, method development and data needs, which may improve the next rounds of MSP (Day, 2008; Douvere & Ehler, 2011).

Since data play such an important role in MSP, there have also been several efforts to study and list requirements related to data needs in MSP. In the Baltic Sea region, there have been many transnational research projects, which have, among other things, addressed the complexity of data issues. Moreover, a recent study of MSP data examines comprehensively the data and information needs of the EU member states in the MSP process (MSP Data Study 2016). In general, the study identified four broad categories of data that have been used in the existing plans and their planning process:

1. Administrative boundaries,
2. Description of the geophysical environment and biological/ ecological features,
3. Data relating to the relevant human activities and sectors,
4. Socio-economic and policy-related data.
Furthermore, the MSP Data Study (2016) highlights that the data needs in transboundary planning are different from the needs in national planning. In cross-border work, the scope and level of detail in data may be much simpler, but the coherence and harmonisation of these datasets remains a major challenge. The Baltic Sea Region Maritime Spatial Planning Data Expert Sub-group (BSR MSP Data Group), which acts as a sub-group to HELCOM-VASAB Maritime Spatial Planning Working Group (HELCOM-VASAB MSP WG), has agreed on a list of most important datasets considering specifically the aspects of transboundary cooperation. With some additions, their list is based on the Directive 2014/89/EU on Maritime Spatial Planning. The list of datasets included 13 planning issues, under which 52 datasets to be used in transboundary MSP were further proposed:

- aquaculture areas,
- fishing areas,
- installations and infrastructures for the exploration, exploitation and extraction of oil, gas and other energy resources, of minerals and aggregates, and for the production of energy from renewable sources,
- maritime transport routes and traffic flows,
- military training areas,
- nature and species conservation sites and protected areas,
- raw material extraction areas,
- scientific research,
- submarine cable and pipeline routes,
- tourism,
- underwater cultural heritage,
- administrative and jurisdictional borders,
- other themes or uses that are not indicated at the moment, but might be considered as relevant to transboundary / cross-border planning issues.

Because MSP is still a relatively ‘young’ area of administrative practice, a lot of the data and information utilised to develop planning evidence for MSP is produced by other actors (public and private) than the responsible planning authorities (MSP Data Study 2016). This raises the issue of requirements for data quality and reliability as decisions need to be based on sound information irrespective of its source of origin. Harmonisation of cross-border data is challenging because of different data collection protocols and data formats. Further problems arise by the different languages used around the Baltic Sea region (MSP Data Study 2016). According to HELCOM-VASAB MSP WG, the underlying data should be as uniform as possible to obtain coherence in MSP. Similarly, it is especially important that the datasets and their contents are well and transparently described, which refer to great importance of proper metadata documentation.

There is a considerable number of international efforts to establish coherence among datasets in different areas of data collection on different geographical and institutional scales. For example, the International Oceanographic Data and Information Exchange (IODE) of the Intergovernmental Oceanographic Commission (IOC), established already in 1961. Since then, there have been numerous programs and organisations build around the issues of ocean data exchange and management. One example is the Ocean Data Standards Pilot Project (ODS), established and implemented jointly between the Joint Commission for Oceanography and Marine Meteorology (JCOMM) and IODE. The pilot project and its successors aimed at developing and implementing international standards in managing data. They have provided, for example, recommendations to adopt ISO standard based coding of dates, times, and countries, as well as submitted proposals for using standardised terms in metadata.

In Europe, a major development is the Infrastructure for Spatial Information in the European Community (INSPIRE). The European Union’s INSPIRE directive came into force in 2007. It aims to improve consistency, availability, and re-use of spatial information to support environmental policies. Among other things, the directive makes requirements with respect to metadata, data sharing, and
Proposals of data standards to be used in MSP

interoperability of spatial datasets and services. Thus, it aims to make a transboundary exchange of data among the EU member states possible. INSPIRE covers 34 spatial data themes. According to a report of BSR MSP Data Group (2017), 21 themes relate to maritime areas and are relevant to MSP. However, INSPIRE does not consider all thematic aspects related to data needs in MSP, and therefore, it does not provide an exclusive solution to the data challenges of transboundary MSP (MSP Data Study 2016). The importance of clear data properties regarding spatial dimension, time and data origin are repeatedly discussed in the literature as pressing needs in MSP (Ehler, 2008, Hattam et al., 2015, MSP Data Study 2016). Yet, the actual implementation in databases remains challenging (Shucksmith & Kelly, 2014).

Relevant sources of information, especially for transnational MSP are provided by a number of international data portals. The European Marine Observation and Data Network (EMODnet), for example, provides a valuable source of harmonised transboundary data, even though it does not have a MSP focus as such. Currently there are seven EMODnet data portals, all of which are fully INSPIRE compliant (MSP Data Study 2016). The thematic data portals cover topics on bathymetry, geology, seabed habitats, chemistry, biology, physics, and human activities. At the Baltic Sea region level, the HELCOM Map and Data Service is an important source of data and information assembled across national borders. The MSP Data Study (2016) lists it as the only data infrastructure that manages a comprehensive set of analyses which address interactions in the marine area.

The Copernicus Marine environment monitoring service (CMEMS) and WISE Marine – Marine Information System for Europe are recent initiatives by the EU to share marine data, the latter still being under development in 2018. Other data portals are provided by OSPAR and ICES. In general, the focus of these data portals is environmental assessment, not planning. Still, environmental data are a prerequisite for developing sensible planning alternatives.

Open data, meaning publicly available data for anybody (authorities or stakeholders) can help to facilitate MSP processes, decisions and acceptance. Currently, only a small percentage of available data are shared in data portals. Often, these contain little information e.g. processed data about the environmental status of a region, required to be shared due to reporting guidelines under EU-law or data with a very coarse resolution, rarely suitable to assess planning alternatives. Original data gathered by authorities may be accessed by other authorities but might be more difficult or even impossible to access by other stakeholders.

If the most promising way to gain access to data is to know who to ask, data gathering can be a challenge within a country. Accessing data from other countries is even more challenging as a foreign language and a different administrative culture increase possible difficulties.

In transnational MSP, all data available to the planning authorities should also be disclosed to all other countries to ensure equal participation.

2.2 Planner and stakeholder perspective

Many of the aspects relevant from a planner’s perspective have already been raised in the introduction, citing documents reflecting on MSP experiences so far. Besides the above cited documents, the base of this (in relation to project-specific activities still initial) specification of needs of planners and stakeholders includes research on participation and observations by the Nordregio team from the EU-financed cross-border MSP projects (Baltic SCOPE project and Pan Baltic SCOPE).

Overall, one can say that marine and coastal planners and data users with a stakeholder perspective need to be able to easily understand what data are about, how they are organised and can be used. However, the perspectives and needs and ability to work with data and a data management structure might differ considerably.

Marine and coastal planners should be able to easily a) access and extract data, b) complement with their own data, c) share data with others and d) structure and (re)organize them according to their own purposes and questions.

Based on the above needs of planner’s and stakeholder’s data and database should provide and facilitate the following:
• **Accessibility & understandability:** a data repository and the data included should be technically easy to use and not require complicated extra training. This also includes transferability between different organisations and countries (possibility to share easily).

• **Quality:** as high quality as possible and basic information allowing to judge the validity and reliability of data.

• **Applicability and transparency:** information on a minimum what the original intention of use is and possibly how far data can be applied to other situations. Transparency in relation how data was produced and assembled (if already combined): a planner should be able to decide whether it is possible to use specific data for new purposes and possibly be able to produce complementary evidence and if possible even contact information (where to learn more, get further information).

• **Scalability:** the possibility to compress and zoom out or to zoom into a dataset when needed, depending on the temporal and spatial scale of the planning question (important keywords: resolution, vector data).

• **Coherence/harmonisation:** From a cross-border and cross-level perspective (i.e. including different administrative/political levels of planning), coherence and harmonisation cannot be underemphasised in relation to quality, data format, and resolution in time and space, and that similar types of data are actually available to develop planning evidence to answer a specific planning question (e.g. a specific use sector).

• **Flexibility and possibility to combine data in new ways:** Planning implies a need to think interdisciplinary and across sectors and reacting to new questions and issues. Therefore, it is extremely important that data can be combined in different ways and easily so. Here, relevant metadata allowing this are important, not just to assess the quality (validity and reliability) of a specific dataset, but also to judge how a specific dataset can be used and whether it is possible to combine at all with other datasets. Moreover, flexibility should include both to adapt the data and the database structure (flagging etc.) to the purposes and problems at hand (the data has to answer relevant planning questions and suit the biogeographical and societal and economic circumstance).

Regarding the needs of MSP stakeholders, the situation is more complicated than for planners to define actual ‘requirements. As there are two components, the stakeholders and their data, this requires besides data processing aspects also pedagogical and process facilitation related reflections. Stakeholders’ needs are highly diverse and difficult to specify at this stage of the project (without asking specific stakeholders). Instead of specific requirements, we therefore rather provide some general points to consider, and terms of concrete requirements rather refer to the basic mapping in relation to the SDSS and future project activities in relation to the case studies below that are likely to lead to more specific requirements through a recursive process.

Ideally, it should be easy for different types of stakeholders to provide their own data input and share it with others in a MSP process (and the BASMATI related experiments). Sharing of data implies mutual “translation” and education, data processing, quality marking and management, which requires both data compiling and pedagogical and process facilitation skills and structures. Beyond the basics of well-established practice of stakeholder involvement in land-based planning, needs will have to be defined based on case studies and specific situations. These needs are both stakeholder and situation dependent.

**Stakeholder diversity:** Firstly, stakeholders in MSP are highly diverse, even within sector “groups” (e.g. fishers differ in gear and vessel types, target species, harbours). Education level and familiarity with different types of data plays an important role, too. Stakeholders also tend to have a highly detailed perspective (high spatial resolution) in the area they have a connection to and interact with on a regular basis, but may not be as informed about aspects that are not part of their perspective (e.g. what is under water or what is not easily visible by eyes, what is changing slowly and imperceptibly). Interpretation of material may also vary considerably. Depending on the situation and the resolution, a line on a map can both calm down or escalate a latent conflict situation.

**Purpose and phase of process:** Secondly, the needs may depend highly on the planning situation stakeholders are involved in and the purposes and scope of involvement (the step in the planning
process: initial or late stage and the activity e.g. problem mapping, needs mapping, trends mapping, data collection, scenario building, evaluation of different alternatives, process evaluation) and on the education background of the individual but also the whole group interacting in the planning situation (what do they know, what do they understand). For instance, data and resulting planning evidence used for communication with and among stakeholders needs to be developed aware of language/terminology, pedagogical purposes and perception aspects, including possible biases. Collection of data from and through stakeholders has to consider more technical perception and collection aspects.

**Easy system:** Generally, if data and planning evidence is to be accessible for a highly diverse group, it should be easily understandable, connecting to situations all can relate to and well visualised. It should also be easy to see how reliable the information is and where the knowledge gaps are (markers). Lastly, it is important to clarify what the implications of maps and other types of data are – i.e. decisive or more indicative and where. A workable strategy is to (using experiences from land-based planning), find a common denominator in relation to colours, scale etc. and start with known maps and flight/satellite images and well-established ways to visualise and perceive data, and practically go forward testing what works and what doesn’t.

**Flexible logical framework:** If working with a participatory process, where stakeholders are part of the problem and causes mapping and defining the solutions, there is a need to have a logical framework and data system flexible enough to construct situation specific problem and solution trees and be able to relate relevant data to these, if there are to be maps and visualisation but also possibly scenario work. Here, we refer to work e.g. in connection with the Open Standards for the Practice of Conservation and the related MIRADI software (see www.conservationmeasures.org/ and for an analysis of the method for MSP see Morf et al. (forthcoming) and literature on participatory modelling and work done within the BONUS project BaltCoast.

**Black box, trust and transparency:** A final question when designing data bases and sharing systems and compiling data for planning evidence is also which data and level of compilation should be accessible to which degree for what type of user. This implies issues of transparency and trust (in relation to process and access to data) vs. understandability and complexity (who can actually understand and work with the data) that needs to be discussed with those involved (stakeholders and planners).

### 2.3 Case studies perspective

The BONUS BASMATI case studies have different thematic scopes and spatial range from the local and regional scale of the Latvian and Danish-German studies to the Baltic-wide needs of the Pan-Baltic case study. D.6.1 ‘Specifications on data and modelling needs’ report has identified that each case study requires physical, chemical, and biological information as well as information on the current and planned sea uses and on the restrictions of the coastal and marine environment. Spatial and socio-economic aspects are identified according to the specific needs of each case study.

**Danish-German case study**

The Danish-German case study investigates opportunities for aquaculture in the south-western Baltic Sea. With eutrophication being one of the main environmental issues in the Baltic Sea, it was chosen to focus on mussel farming as mussels can contribute to the mitigation of eutrophication. In order to choose potential mussel farm locations, a tiered site selection process was adopted, consisting of

i) the mapping of constraining human activities and enabling environmental conditions in the chosen area

ii) identification of external requirements

iii) identification of a number of potential sites and

iv) an evaluation of the sites based on ecosystem services.
The first three steps – relying heavily on data and GIS analyses – have been concluded and provided some insights into data requirements. The first thing to note is that data on maritime uses had to be gathered from a number of different sources, e.g. the HELCOM Map and Data Service, the environmental GIS ("Miljøgis") from the Danish Natura Agency, and the GeoSeaPortal from the Federal Maritime and Hydrographic Agency. The datasets were not only projected in different coordinate systems but also in different formats, so one of the first questions raised were:

- Which coordinate system is most suitable for the case study area?
- Are the transformations adequate?
- Which format, i.e. vector or raster, is most suitable with regard to the needed GIS analyses?

Another issue that surfaced was the timeliness of the data. Some datasets were last updated several years ago. Thus, to validate e.g. the status of the wind parks (which was only discovered when looking at the attribute table) required an internet search to obtain whether the respective wind park is still under construction, already fully commissioned or have plans been put on hold.

With regard to data on environmental conditions, further issues emerged, such as the appropriate resolution for the study area and suitable resampling methods. In order to document all the choices made in the site selection process, the INSPIRE metadata template available in ArcCatalogue was quite useful but it also raised further questions, e.g. the choice of keywords, the detailedness of the layer description and so forth.

These are just a few examples that illustrate where common data standards and requirements are very useful and needed. It facilitates a faster workflow when preparing MSP relevant data (e.g. when an appropriate coordinate system or resolution is already proposed). Furthermore, it facilitates that data users can see at a glance what the dataset contains, how it has been processed and how up-to-date it is.

**Latvian case study**

The Latvian case study addresses the need for the assessment and comparison of environmental impacts, costs, and benefits of alternative sea use options / scenarios in relation to the designation of Marine Protected Areas (MPA) to provide support for discussions with stakeholders and political decision-making for MSP. The case study utilises integrated assessment where spatial data layers on Drivers-Pressures-State-Impacts (DPSI) components are needed.

- ‘State’ is characterized by variables related to the benthic habitats. Data about the current abundance of flora and fauna are needed to characterise the present state of the marine environment.
- ‘Drivers’ are human activities using the sea and impacting these benthic habitats, and
- ‘Pressures’ are caused by these activities. Hence, spatial information about the current and future human uses of the sea is needed.
- ‘Impacts’ are characterised by the ecosystem services provided by the benthic habitats and the human welfare gains (benefits) from these ecosystem services, where data on relevant indicators allowing their assessment are needed.

The Latvian case study will aim to utilise accessible environmental data sources (e.g. HELCOM Map and Data Service, EMODnet), where countries provide national information regularly. Only those data that are considered of sufficient quality will be used. As for environmental layers, the Latvian case study will need the point measurement data tables, modelled data from observation points or polygons etc. Further, data from MSP documents, official reports as well as Latvian national monitoring program will serve as substantial resource for the case study needs. The availability of raw data will be governed by the policies of the respective databases. The socioeconomic analysis planned as part of the case study will build on data characterising benefits and values from the ecosystem services and changes in their provision in various planning scenarios, as well as
socioeconomic data on relevant maritime activities. The data will come from national statistics, published economic (monetary) valuation studies, earlier studies and literature, as well as expert knowledge. Most of the necessary socioeconomic data are not collected regularly but are gathered as part of dedicated studies. Thus, lack of the data is the main challenge. The next challenge is spatial distribution of the data (e.g. on the benefits and economic (monetary) values from the ecosystem services) since such data production has been addressed on limited extent in the research so far. Thus efforts will be made to create spatial data based on available data, qualitative research and expert knowledge.

In order to evaluate data quality, information about the type of data and origin plays a central role during the data collecting process. Homogeneous descriptions would significantly facilitate the time-consuming evaluation process.

**Pan-Baltic case study**

The Pan-Baltic case study investigates stakeholder involvement from a business sector perspective. It studies stakeholder perceptions and requirements concerning their involvement in the MSP process. The study will produce information based on the expert knowledge of the respondents and interviewees. Data and information about the human uses and the current environmental status of the sea areas may be utilised to support the discussion with the business sector representatives. The case study relies on existing data sources, such as international data portals and official reports. Specifically, the Pan-Baltic case study focuses on two international and transboundary business sectors, namely maritime transport and marine tourism. The Baltic-wide perspective of the case study poses problems for the data requirements regarding their spatial scale. On one hand, information needs to cover the entire sea basin and thus it must be rather generalised. For example, the main shipping lanes are well known and presented by AIS data. On the other hand, more detailed information might be needed if local conflicts and synergies are examined. The datasets also need to specify their time frame. As some anthropogenic uses, especially regarding marine and coastal tourism, are strongly linked with specific seasons, data should cover seasonal or monthly time scale. Similarly, it is important to separate regular activities, such as passenger ferry operations, from one-time events. In a longer time line, it is important to have information both on the current and the expected future activities at the sea. For co-location analyses, potential conflicts and synergies among sea uses need to be defined, and there, information on the temporal differences in the occurrence and timing of the activities at the sea is relevant. The quality and coverage of metadata description play an important role whenever utilising datasets produced by others. In a Baltic-wide case study, good metadata is especially crucial if datasets from various sources are combined to reach better pan-Baltic data coverage. New data created in the Pan-Baltic case study will be based on the expert views and perspectives about the stakeholder involvement in MSP and are mainly qualitative. Therefore, raw data, such as recordings of interviews, will not be saved in public archives due to privacy issues. Suitable format for publishing the analysed results will be considered and decided as the case study proceeds.

**2.4 The Baltic Explorer perspective**

The Baltic Explorer will be an intuitive web-based SDSS for MSP with a multi-functional tool design that includes tools for data exploration, cumulative impact assessment (CIA), and suitability analysis. The Baltic Explorer will support interaction by a single expert user, as well as planning by groups in a collaborative setting, where discussions are facilitated through transparent analysis and visualizations on a large touchscreen device. The spatial data in the Baltic Explorer can be divided into three groups, background maps, input data (e.g. pressures and ecosystem components used in CIA), and result data. Background maps are used to give spatial context to the analysis process. Therefore, it is important that the background maps are compatible with the input data, e.g. they should have similar scale, use the same coordinate reference system, and all elements in the background maps should be valid in the same time frame as the input data. The input data is used for the Baltic Explorer's CIA and suitability analysis tools that also produce the result data.
The purpose of the Data explorer tool is to enable users to familiarise and explore the input data for the analyses, and in this way it contributes to the transparency of the system. The Data explorer tool does not set requirements on the categorization, relations, or other data specifications, rather it will be developed and configured to support the data of the BONUS BASMATI project. For example, the categorisation of the data can be supported in the user interface of the Baltic Explorer by visual means, helping the user navigate the user interface fluently and find datasets quickly. Ideally the Data Explorer would be able to use data from multiple sources, or even user provided data. Therefore, standardised categories would enable any additional data to be easily organised in the system. Temporal data could potentially enable highly dynamic and innovative data exploration, such as animations of seasonal variations in data. However, such features require a large amount of frequently collected data, and whether there is a real need for such features is unknown. Data layers representing different levels of the water column can be supported by layer switching. All the datasets currently planned to be used in the Baltic Explorer are two dimensional, but if three dimensional datasets were needed, the tool could also be developed to support three-dimensional data exploration.

Because the functionality provided by the Baltic Explorer’s CIA and suitability analysis tools will be based on the needs of the case studies that are developed in WP6 of the BONUS BASMATI project the requirements set by the Baltic Explorer on the categorization, relations, and other specifications related to the contents of the input data are derived from the case studies. The result datasets from the analyses should be considered in the categorisation and they are also related to their input data. The performance and development of the analysis tools will benefit from data that has been harmonised in advance, especially in regards to technical properties (such as format, origin, grid, coordinate reference system, etc.), as non-harmonised input data will require the development of additional tools for harmonisation. This limits the flexibility of the system and its ease-of-use, as users could be required to make additional efforts to pre-process data with traditional GIS software. However, it enables resources in development to be targeted at more critical functionality. To further enhance the transparency of the system, data that has been processed in advance should include documentation on the methods by which the data has been processed and information on the original datasets that have been used for it. Preferably, the original datasets would be available in the Data explorer. In the current stage of the development of the Baltic Explorer, specifications of the system are still flexibly adjustable to support different, themes, categorizations, relations and other properties of the data.
3 Proposals for data requirements

Derived from the needs identified in Chapter 2 of this report, proposals for data requirements are focusing on the applicability and operability of ecosystem service data in MSP processes. Chapter 3.1 deals with key themes of the ecosystem service framework. In order to intuitively structure different kind of data for easy access and search, we suggest to reflect these themes within data categories of a database. By adding tags or keywords in datasets, the relation of different categories and ecosystem services with e.g. associated human activities etc. can be maintained (Chapter 3.2.). Chapter 3.3 focuses on data compliance with detailed data property descriptions for spatial scale, time frame, type of data, conflicts / sensitivities in order to be useful for planning. Further, transparency of data production is crucial for usability, therefore metadata quality needs to be improved and should be easy to achieve e.g by using a simple template (Chapter 3.4).

3.1 Data categories

Database structure should easily allow to find all relevant data. All key themes addressed in the framework on ecosystem services (D.4.1) should be visible and searchable in the respective data categories. Therefore, we suggest to apply categories reflecting the frameworks regarding the ecosystem service approach and the DPSIR (Driver-Pressure-State-Impact-Response) model (Müller & Burkhard, 2012) as shown in Figure 2. Nevertheless, the naming of these categories needs to be easily understandable in order to be useful for practitioners and stakeholders. Also, these categories need to be structured intuitively for different users looking for certain data. Examples for organising data content within the category ‘ecosystem properties’ is shown in Figure 3.

Figure 2: Data categories derived from the DPSIR management cycle (Müller & Burkhard, 2012)

Depending on the purpose, the user might want to re-arrange data differently based on their specific needs.
In practice, data content may be more complex than the structure suggests. Some data may fit several of the categories or not fit any. These suggestions can be seen as a general proposal. A database should allow optional additional categories to be flexible for the needs of planners and stakeholders. Naming of categories and datasets should be clear and straightforward to support and not hinder finding relevant information.
3.2 Relations between data categories

During an MSP process, data are collected and evaluated not only by experts in the respective discipline. In general analysis must be possible without expert knowledge, relying on the information given.

Interrelations within the marine environment are complex making it difficult to transfer and to maintain the connections and level of relationships between different data categories e.g. the relation of fish being present (ES), fishing (activity), landings of fish (benefits), earning (economic value), fishing gear (effects/pressures) in datasets.

Relationships between different categories should be retraceable within datasets (not left unsaid that still some effects may be hidden or unknown). This is especially relevant when technical tools are used to further analyse and process input data e.g. for cumulative impact assessment or suitability analysis. To maintain these relations between data, datasets should receive tags or keywords to refer to the attributes in their network of affiliations with other categories. Relationships may also be situation, place and case specific.

To support the understanding behind this rationale we propose to draw on the classification after CICES (as proposed in WP4, von Thenen et al., 2018) and expand it for other categories. Meaning that the data topic of the respective dataset is used as a starting point to ask for relations to other themes: e.g. which activities are connected to fish stocks or influence them? Which benefits to humans or effects on the ecosystem are connected to fishing? Which economic values are related to the amount of fish? Often these questions evoke more than a single answer, meaning that several tags can be placed.
3.3 Descriptive data properties

The HELCOM-VASAB Maritime Spatial Planning Working Group (HELCOM-VASAB MSP WG) introduced a set of technical requirements (data specifications) based on INSPIRE land use data specifications to facilitate the interoperability and harmonisation of spatial datasets (HELCOM-VASAB MSP WG, 2018). The recommendations prescribe how MSP data could be structured in a convenient and flexible way to be coherent, effective and used for marine spatial plans across borders (HELCOM-VASAB MSP WG, 2018). With a list of predefined but flexible expandable attributes (code list values) and associated classes (attribute codes) a system is introduced to structure in more detail the cartographic representation of datasets.

We adopt the idea of HELCOM-VASAB MSP WG (2018) in general and propose to include information about four themes into datasets, for a more detailed data description:

- spatial dimension,
- time frame,
- conflict potential / sensitivities and
- origin of data.

The information should be given in predefined subclasses to enable search options and to ensure easy data processing and further analysis.

The spatial dimension describes the indispensable property of data, to be used in the MSP process. Describing a three-dimensional body, data are often interpreted as effecting the whole water column, although e.g. activities like shipping impacts the surface water and upper water column and cables or pipelines on the long-term effect the seafloor (Table 1). To clarify spatial elongations of data for more detailed estimations of synergies and adverse effects, we suggest extending the data descriptions in predefined classes as shown in Table 2.

### Table 1: Descriptive data properties in predefined classes for the vertical spatial dimension

<table>
<thead>
<tr>
<th>spatial dimension</th>
<th>class</th>
<th>definition/description</th>
</tr>
</thead>
<tbody>
<tr>
<td>vertical</td>
<td>above sea surface /air column</td>
<td></td>
</tr>
<tr>
<td></td>
<td>sea surface</td>
<td>air-sea boundary</td>
</tr>
<tr>
<td></td>
<td>surface water</td>
<td>e.g. up to 10 m below sea surface</td>
</tr>
<tr>
<td></td>
<td>water column</td>
<td>whole water column</td>
</tr>
<tr>
<td></td>
<td>bottom water</td>
<td>e.g. up to 10m above seafloor</td>
</tr>
<tr>
<td></td>
<td>seafloor</td>
<td>only seafloor</td>
</tr>
</tbody>
</table>

### Table 2: Descriptive data properties in predefined classes for the horizontal spatial dimension

<table>
<thead>
<tr>
<th>spatial dimension</th>
<th>class</th>
<th>definition/description</th>
</tr>
</thead>
<tbody>
<tr>
<td>horizontal</td>
<td>point</td>
<td>point coordinates available</td>
</tr>
<tr>
<td></td>
<td>local</td>
<td>small spatial scale, a few km or km²</td>
</tr>
<tr>
<td></td>
<td>regional</td>
<td>spatial scale reflecting e.g. ecological, historical, political, climate or morphological zone</td>
</tr>
<tr>
<td></td>
<td>basin wide</td>
<td>spatial scale follows morphological characteristics of sea basins e.g. Bornholm Basin</td>
</tr>
<tr>
<td></td>
<td>Baltic wide</td>
<td>spatial scale reflects the whole geographic elongation of the Baltic Sea</td>
</tr>
</tbody>
</table>
The time frame needs to be considered in order to evaluate e.g. pressure and impact levels on an area correctly. Also, it is necessary in order to identify potential conflicts or synergies for uses in the same area. Whether activities take place once a year or several hours each day makes a difference for planning considerations. Supplementary data descriptions in predefined classes are shown in Table 3. For the spatial scale as well as for the time frame, the best resolution level possible can significantly improve data quality and its applicability on different scenarios.

Table 3: Descriptive data properties in predefined classes for time

<table>
<thead>
<tr>
<th>time frame</th>
<th>class</th>
<th>definition/description</th>
</tr>
</thead>
<tbody>
<tr>
<td>frequency</td>
<td>never</td>
<td></td>
</tr>
<tr>
<td></td>
<td>one-time event</td>
<td></td>
</tr>
<tr>
<td></td>
<td>irregular</td>
<td></td>
</tr>
<tr>
<td></td>
<td>regular</td>
<td></td>
</tr>
<tr>
<td>temporal resolution</td>
<td>frequently and regularly occurring events</td>
<td></td>
</tr>
<tr>
<td>daily</td>
<td>monthly</td>
<td></td>
</tr>
<tr>
<td>seasonal</td>
<td>decadal</td>
<td></td>
</tr>
<tr>
<td>time line</td>
<td>past</td>
<td>e.g. historical data or data of outdated marine spatial plans</td>
</tr>
<tr>
<td></td>
<td>existing, present state</td>
<td>e.g. active data up to date, legally adopted</td>
</tr>
<tr>
<td></td>
<td>planned, future prediction</td>
<td>e.g. planned state or condition, marine spatial plan in progress.</td>
</tr>
</tbody>
</table>

Relying on the available database, to identify potential conflicts these must be extractable from the available datasets. In the first place, conflicts (or synergies) can be approximated by checking first for spatial, and second for temporal overlap. Still, not all conflicts may be detectable by such simple means, especially when cause and effect are spatially disconnected and/or delayed in time. Hence, the identification of conflict or synergy potential remains highly complex and difficult to describe by simply using tags and keywords. Still, potential conflicts can be derived defining sensitivities of selected activities or properties. Some examples are illustrated in Table 4.

Table 4: Descriptive data properties in predefined classes for conflict potential

<table>
<thead>
<tr>
<th>conflict potential</th>
<th>class</th>
<th>definition/description</th>
</tr>
</thead>
<tbody>
<tr>
<td>noise</td>
<td>lethal, hazardous</td>
<td>describes sensitivities of e.g. organisms or human activities to certain noise levels</td>
</tr>
<tr>
<td></td>
<td>high sensitivity</td>
<td></td>
</tr>
<tr>
<td></td>
<td>low sensitivity</td>
<td></td>
</tr>
<tr>
<td></td>
<td>no response</td>
<td></td>
</tr>
<tr>
<td></td>
<td>unknown</td>
<td></td>
</tr>
<tr>
<td>toxins</td>
<td>lethal, hazardous</td>
<td></td>
</tr>
<tr>
<td></td>
<td>high sensitivity</td>
<td></td>
</tr>
<tr>
<td></td>
<td>low sensitivity</td>
<td></td>
</tr>
<tr>
<td></td>
<td>no response</td>
<td></td>
</tr>
<tr>
<td></td>
<td>unknown</td>
<td></td>
</tr>
</tbody>
</table>

In general, data will be used by people who were not involved in producing the dataset, making information about the origin of data a crucial asset. High transparency about data origin is necessary.
to enable adequate interpretation of results. Concerning model-derived data, smudging, and over or underestimation of values can lead to false conclusions and decisions. Therefore, the type and origin of data should be carefully described using for example the classes proposed in Table 5.

We propose to thoroughly describe data type and origin instead of valuation of data by quality flags. The use of quality flags without context could be misleading in practice. Also, the definition of such quality levels might be biased depending on the application of the data and the intention of the user. E.g. marine biotope data can be suitable for the evaluation of a wind park location whereas the same data can be inadequate to perform an ecological evaluation.

The practice of using model-derived data to inform decision-making processes has progressively increased over the last years. For a transparent evaluation of these data, we advise to provide model-derived data along with

- the underlying measured data point locations,
- the grid on which interpolation was based,
- information about the (spatial) interpolation method,
- statistics about the model quality, likelihood of the prediction e.g. 95th percentile

Concerning a harmonisation of data grids used for spatial interpolations, one option is to provide data in geographical grid systems as applied by the European Environment Agency (EEA) and INSPIRE. Based on an equal area projection (e.g. LAEA), the EEA reference grid is suitable for generalising data, statistical mapping or analytical work whenever a true area representation is required. Recommended grid resolutions are 100 m, 1 km, 10 km and 100 km (Peifer, 2011). Alternative grid cells might be implemented, depending on the real data distribution or level of significance as presented by Korpinen and colleagues, constituting human pressure calculations and their potential impact for the Baltic Sea ecosystem by using 5 km x 5 km grid cells (Korpinen, Meski, Andersen, & Laamanen, 2012).

### Table 5: Descriptive data properties in predefined classes for origin of data

<table>
<thead>
<tr>
<th>origin of data</th>
<th>definition/description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Measurements / observations</td>
<td>scientific based, independent measurements</td>
</tr>
<tr>
<td>official documents</td>
<td>e.g. national or regional reporting</td>
</tr>
<tr>
<td>expert knowledge</td>
<td>assumption by acknowledged experts in the topic</td>
</tr>
<tr>
<td>models</td>
<td>Interpolations and calculations based on modelled or in-situ data</td>
</tr>
<tr>
<td>historical records</td>
<td>historical data and reports</td>
</tr>
<tr>
<td>reports by locals, individuals</td>
<td>historical or recent reports of locals or other individuals</td>
</tr>
<tr>
<td>assumptions by “source”</td>
<td>no solid data foundation, name data source</td>
</tr>
<tr>
<td>by authorities</td>
<td>data provided by authorities</td>
</tr>
<tr>
<td>user data by companies</td>
<td>data provided by companies</td>
</tr>
</tbody>
</table>

Table 1 to 5 show a recurring simple system for structuring information in datasets. It follows an easily assessable and more so flexible scheme. The system can be extended and adjusted to any theme, which might be in focus by a planner or stakeholder depending on the purpose and the phase of a MSP process.

### 3.4 Guideline on metadata

Geospatial data are generally used by a variety of administrative, scientific or other stakeholder personnel and data are provided by a multitude of local, national or international geodata portals. Data and metadata are commonly searched for and found via search engines, therefore metadata should generally be prepared in consideration of being detectable (e.g. refrain from using code
numbers). To enable a far-reaching application of data, metadata need to be precise, short but sufficient and intelligible for a wide community and easy to access.

Two principles are essential for handling metadata:
- only describe your own data
- keep the description up to date.

Guidelines on how to produce metadata are often complex and extensive, hindering the willing user to apply them. But minimum requirements on metadata can be met with little effort. The following guideline is based on the ‘Code of practice - working with high quality metadata’, provided by the GDI-DE by the Ministry for the Agriculture and Environment of Sachsen-Anhalt, Germany. It provides a short and easy understandable template version on the minimum information needed in order to facilitate the production of high quality metadata with minimum effort.

**Data name**
- only generally understandable wording
- no abbreviations or codes

**Coordinate System**
- recommended coordinate reference system by the European Environmental Association is ETRS89-LAEA Europe, EPSG: 3035
- the geodetic datum is the European Terrestrial Reference System 1989 (EPSG: 6258)
- the Lambert Azimuthal Equal Area (LAEA) projection is centred at 10°E, 52°N
- coordinates are based on a false Easting of 4321000 meters and a false Northing of 3210000 meters

**Description of the data**
- running text without formatting
- common terms and definitions
- no abbreviations
- correct orthography
- follow common systematic categories
- follow a gradation - important information first, less important information second
- if applicable: interpolation method, statistics on likelihood

**Administrative information (up to date)**
- author of the data
- host of the data
- contact information
- data security demands
- conditions to access and use

**Key words**
- select key words and catch phrases

**Spatial data service type**
- defined by ISO 19119
- WMS capabilities

**Metadata identification**
- file identifier (UUID or ID)
4 Conclusion

The success of MSP depends to a large extent on the quality and availability of relevant data. The discussion of requirements on data from different perspectives showed that accessibility and understandability of data as well as harmonised quality standards are central aspects when working with data in MSP. A dataset must contain all relevant information for the user to decide about its suitability in the respective phase of the MSP process.

Drawing on these findings we propose

- to include data categories for datasets corresponding to the ecosystem service framework
- to integrate information about relations between data categories via tags or keywords
- to thoroughly describe data properties regarding spatial dimension, time frame, conflict (/synergy) potential and data origin
- to support metadata description by simple guidelines

These recommendations can enhance the inclusion of ecosystem services in MSP processes and provide effective steps towards a more coherent data management.
References


von Thenen, M., Schiele, K., Fredericksen, P., Algars, J., Pakalniete, K., Strake, S., ... Schröder, L.
Proposals of data standards to be used in MSP