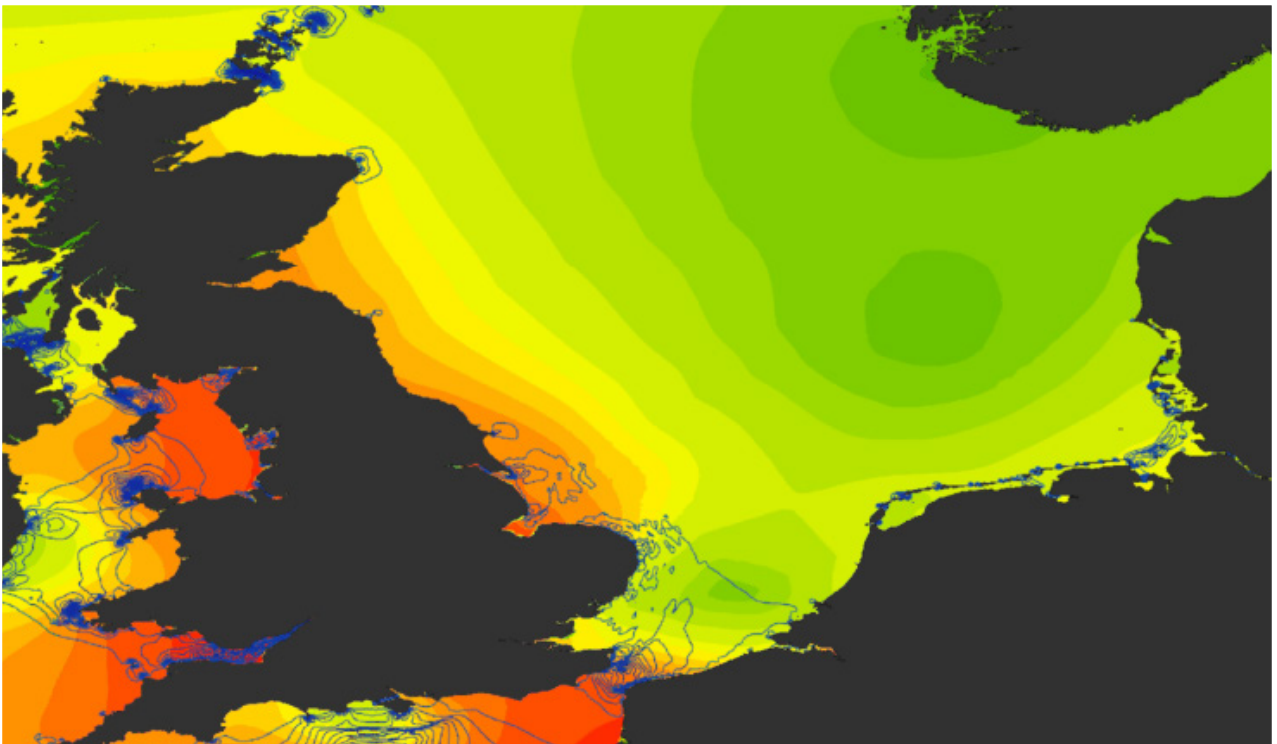# HR Wallingford
Working with water

Growth and Innovation in the Ocean Economy: North Sea Checkpoint

Progress Report 1

DLS0342-RT001-R01-00   May 2014

Document information

Document permissions	Unrestricted
Project number	DLS0342
Project name	Growth and Innovation in the Ocean Economy: North Sea Checkpoint
Report title	Progress Report 1
Report number	RT001
Release number	R01-00
Report date	May 2014
Client	European Commission DG MARE
Client representative	Ian Shepherd
Project manager	Marie Pendle
Project director	Keiran Millard

Document history

Date	Release	Prepared	Approved	Authorised	Notes
09 May 2014	01-00	MPN	KLH	TKM	

Document authorisation

Prepared	Approved	Authorised

© HR Wallingford Ltd

Contents

1. Introduction

This is progress report 1 for the EMODnet North Sea Basin Checkpoint (NSCP) covering the period up to 30th April 2014.

This period of the project has focussed on putting systems in place and refining the precise scope of project outputs. There have been particular considerations as to what the persistent outputs of the checkpoint projects should be and how the checkpoints work alongside the EMODnet thematic portals. This was scoped to some extent in the project brief and our project proposal, however upon commencing the work, further discussing with the Mediterranean Sea Checkpoint (MSCP) and discussing the project with other stakeholders, this needs to be refined.

The progress during this reporting period is summarised below, and further detail is provided in the remaining sections.

1.1. NSCP challenges

The oil platform leak challenge has been set up and is ready for a rehearsal. The marine protected areas, climate and coastal protection and wind farm challenges are all at the data gathering phase, whilst fisheries management, river inputs and marine environment are not timetabled for activity to commence yet.. This is discussed further in Section 3.

1.2. Literature review

The literature review is underway and is using the free web-based tool Mendeley as a collaborative area to share and comment on papers between project partners. The literature review is providing a contextual view on the blue economy and data gaps that are limiting its development. This is discussed more in Section 4 of the report.

1.3. Dataset registry

A scalable data registry model has been established to easily relate datasets to challenges and the value they contribute to solving the challenge. This data model could easily be exposed as a Resource Descriptor Format (RDF) or linked data allowing the outputs of NSCP to be discoverable by search engines. This is covered in Section 5 of this report.

1.4. Data portal

In order to communicate the outputs from the NSCP in a consistent way with other EMODnet activities, the NSCP has agreed with the EMODnet secretariat to use the existing EMODnet content management system as a service platform. This offers a number of advantages to EMODnet and is discussed further in Section 6 of this report.

1.5. Data Products

All of the challenges involve developing some kind of data product. Across the challenges these vary in their representation (maps, grids, statistics) and geographical extent (local site, basin wide). Clearly this has implications on a "single size" data service to deliver these data products. In addition, there is considerable scope for variability in the data products between each check point. For this reason, an agreement between the checkpoints on the data products is needed – including how they may vary between checkpoints.

2. Management

2.1. Interface with EMODnet

We have attended various meetings with EMODnet, the first being the nineteenth meeting of Marine Observation and Data Expert Group (MODEG) on 4th and 5th July 2013, prior to contract signature. At this meeting an initial presentation of the project was given. Subsequent to the signing of contract on 31st October 2013, we attended the first EMODnet Steering Committee meeting on 16th-17th December 2013. We were unable to attend the inauguration event of the Steering Committee, but we have subsequently attended a meeting for the EMODnet, checkpoints and SeaDataNet entitled "Unlock the potential of marine knowledge by making easier and less costly the access to real-time and archived data" on 29th April 2014. Additionally, we have stayed in email contact with the EMODnet secretariat and various EMODnet partners, as well as the EC representative, Dr Iain Shepherd.

2.2. Interface with MedSea checkpoint

The project manager of the North Sea checkpoint project attended the MSCP kick-off meeting by means of Webex. Subsequently, a series of meetings has been held alongside the MedSea checkpoint. At this stage we are working as separate projects, but sharing approaches to distil best practice and highlight items that need to be addressed collectively.

2.3. Project meetings

The initial kick-off meeting for the North Sea Checkpoint project was held in December 2013 at the HR Wallingford campus in Oxfordshire, England, following signature of the contract. All project partners attended and a representative from the EMODnet secretariat was also welcomed for one day of the two day meeting. Subsequent communications have mostly relied on teleconference meetings, to minimise expenditure on travel and travelling time. However, another face to face meeting is to be arranged subsequent to the next EMODnet steering committee meeting to discuss anything arising from that meeting and the response of DG MARE to this progress report.

3. Challenges

Most of the work package challenges are either yet to start, or still within the early data gathering phase. Those within the data gathering phase are wind farm siting, climate and coast protection, and marine protected areas. No outputs from these challenges have yet been produced.

Those which are not timetabled to yet commence are fisheries management, marine environment and river inputs. Despite no timetabled activity for the marine environment, links have been made with the OSPAR Eutrophication working group and one of their meetings was attended by the leader of the marine environment challenge to strengthen the link and ensure that OSPAR activities are considered during the challenge.

The Oil Platform leak is one of the challenges (Work Package 4) in the project. IMARES have been preparing to achieve the level of readiness that is needed to perform the challenge. The contract allowed about six months preparation time and these have now passed. It is considered that the challenge is now at the stage where the intended rehearsal can be undertaken.

The project plan does not specify who will deliver the specifications for the rehearsal oil platform leak and how this will be done. We would prefer that the specifications be supplied by DG-MARE in a fashion similar to the real thing and would welcome discussions about this activity.

3.1. Choices and preparations made

In the original proposal document and the project plan IMARES envisaged using the DREAM-model, included in MEMW (Marine Environmental Modelling Workbench) to stand up to the Oil Platform Leak challenge. As a result of past contributions to the development of MEMW, IMARES have enjoyed the use of this model for many years. Conditions applied to their use of the model during this time. However in between writing the bid, award of the project and the start of project, SINTEF (the owners of MEMW) have decided to take this software from a development phase to a production phase and change the licensing conditions and costs. IMARES has not been successful in ensuring access to DREAM at an affordable cost to the project, as the present cost for a one year license is €25,000. Therefore an alternative approach for modelling the oil spill was needed.

A number of alternative modelling approaches have been considered, e.g. MEDSLIK II , DREAM-MEMW, OILTRANS etc. References have been identified that point to the strengths and weaknesses of these models and this will be used in the reporting on the challenge. An very useful example is a 2004 report[1] that compares four models, including DREAM and GNOME. As MEDSLIK will already be used and tested in the Mediterranean (Lot 2 of the DG-MARE tender), it was felt that it would not be helpful to use the same model for this challenge (Lot 1, North Sea), as it is considered beneficial to be able to compare different modelling approaches and data usages between the two projects. The selected model is GNOME, which is available free from the National Oceanic and Atmospheric Administration (NOAA) (USA). This choice has strongly been influenced by the fact that GNOME is the operational model for oil spill modelling from the NOAA. It is also used operationally elsewhere.

---

[1] Jong, K. de, 2004. A Comparison of Four Particle Models A Comparison of DemWaq, DREAM, GNOME, and SIMPAR. RIKZ/OS-2004.121w, The Hague. Available at: http://edepot.wur.nl/174587 [Accessed April 14, 2014].

IMARES have identified sources for the data required to feed into and run the GNOME model to give an oil spill track.

Data requirements include currents (sourced from MyOcean) and winds (partially sourced from MyOcean = historical, recent past). Wind predictions are not available from MyOcean, but can be sourced from other weather sites. Presently we are looking to source these as GRIB-files from Global Marine Network or GMN (http://www.globalmarinenet.com/grib_downloads.php)

Recently some alternatives have been found that may be better for the purpose of the work package, although these are still under evaluation. zyGrib allows for quick downloads from NOAA servers of GRIB files. These also have wind information on land, so may be better to interpolate between land and sea for inshore areas. Similarly UGRIB is a program that sources the GRIB-information from ECMWF, so may be slightly better in capturing winds on and around (on land) the North Sea.

NetCDF and GRIB data do not come in formats that can directly be fed into GNOME. GNOME expects NetCDF as input for both currents and winds, but has assumptions on variable names. Using R, IMARES have developed scripts that allow the NetCDF variables to be renamed, so they can be used in GNOME. A similar script is available to join past wind (from NetCDF) with future wind (from GRIB).

The available data is at 0.5 degree resolution, but it was noticed that GNOME hardly moved an oil spill between a grid point and the shoreline with this resolution. To overcome this, the data can be interpolated to achieve 0.25 degree resolution.

In conclusion the required data can be sourced, and scripts to prepare it for use in the model have been developed, including stitching together past winds/currents to future wind/currents. An automated download script (written in Python) has been coded. The Motu-client.py script (as supplied from MyOcean) is used inside our own script.

GNOME does not cover all the needs that the challenge sets: it does not encompass ecology or sensitive coastal habitats. IMARES intends to tackle that part of the challenge by post-analysis of the oil spill track information as modelled by GNOME in GIS (ArcGIS). Checks and analysis that we have prepared for are earlier grounding of the oil spill than noticeable in GNOME (high water/shore line) e.g. on sand banks and mud flats, by checking against water depth, available as EMODnet Bathymetry. Sensitive coastal habitats and species can be checked using the Natura2000 GIS information available from the European Environment Agency, or from national nature agency datasets, such as that provided by the UK Joint Nature Conservation Council, which may capture potential sites as well as currently designated sites. Bathing water information datasets are to be included to assess potential leisure use impacts and/or public perceptions.

The most appropriate output for use in the post-analysis in GIS has been determined to be the NetCDF LE File for Model Run. After checking against other output options this gives the most detailed information and has the best options for scripting the conversion. A Python script has been prepared to process the GNOME output to a CSV output that could be read by GIS. However it is intended to improve the script and directly convert to a GIS-dataset.

3.2. Remaining issues to address

- There seems to be a data gap between (re-analysed) current and weather data and forecast data. The main gap is "yesterday". Depending on the actual challenge this may not be needed. The best available data for "yesterday" is yesterday's forecast (when in it was the "today" of the forecast). However from the data sources that we have obtained, yesterday's forecast is not available. One possible way to overcome

this is by further automating the data acquisition process, so that a library of some past forecasts is automatically built to fill the data gap. If this path is chosen, the stitching together of data from NetCDF and GRIB-files may have to be revisited to properly patch all available data together.

■ The NetCDF and GRIB sources for wind appear to disagree on the wind direction. Although the file specifications point to it containing – in both cases - u (eastward bound) and v (northward bound) components, a 180 degree change in wind direction is noticeable at the change between sources.

■ GNOME does not do weathering of oil and only has a limited data base of oil types. It can output a wind data file which can be used in conjunction with a related tool from NOAA called ADIOS (Automated Data Inquiry for Oil Spills), an oil weathering program.

■ Literature and references. The most relevant references have already been uploaded into the Mendeley on-line database that is available for the MareGaps-project, but more will be added.

■ Specifications for the rehearsal of the model, as mentioned previously. It is preferred that these come from DG MARE.

4. Literature survey and data gathering

At the outset of the literature survey (Work Package 1) consideration has been given to how to provide not only the overview of data availability and adequacy, but also an understanding of the usability by those who will be part of the blue economy. The role as 'pseudouser' of the check points is one that we consider of paramount importance if the aim of "an identification of how well the present data collection and monitoring programmes meet the needs of users" is to be met.

4.1. Considerations

Much of the monitoring, data collection and aggregation has been undertaken by public bodies and the academic community, whose efforts have been extensive. However, the driving factors in the blue economy of commercial users may not directly match to the experiences of the academic and public sectors who are already very involved in the EC initiatives. This is not to say that the latter sectors will be ignored, indeed, data requirements of the Marine Strategy Framework Directive (MSFD) are considerable and the use of existing monitoring programmes and data sets will be essential to measure the progress over time of the steps taken to meet the aspirations of the MSFD.   But academic and public sectors have extensive linkages to the data initiatives, so their data accessing behaviour is likely to be different to that of commercial concerns. It was therefore decided that a 'two-pronged' approach should be taken, looking at both the data initiative reporting, and also seeking commercial or practical user reports.

The literature survey does not in itself appraise data attributes (independent of third party published findings), this is to be done in the context of the challenges. This is because the value of the data can only be assessed in the specific context of the challenge. No two usages of data are likely to be exactly the same, at the very least because of temporal and geographic alterations. Therefore the report of data adequacy will necessarily begin by stating what the checkpoint assumptions have been, with regard to judging both overview of data availability and completeness of data. At this stage, the tentative assumption is that the availability and completeness will be based on the purely geographic nature of the project, so the coverage and accessibility of marine datasets within the bounds of the North Sea as delineated in the proposal, with an additional examination of the temporal extents of data.

4.2. Activities

The tasks being undertaken are:

- looking at the reports from the data provision initiatives, to answer the queries posed in the tender document regarding data adequacy, but also;

- doing a literature and internet search to find any studies from users, giving their experiences with accessing data or other interfacing with data initiatives. This has its limitations, as the commercial world is less likely to use scientific literature to report such findings, but conference proceedings or specialised trade publication articles or even website articles can contain this type of information;

- testing the discoverability of the various data initiatives (not the data themselves) through use of different search engines and search terms, to understand whether commercial and private sector users with no previous link to these activities would find them quickly and easily when identifying a need to access North Sea data.

The literature survey therefore has the purpose to better understand data factors affecting the blue economy as well as identifying data sets and data services. However, there is sufficient time left within the literature survey activity to refocus if this approach is considered unnecessary by DG MARE.

4.3. Use of web collaboration tools

It was identified that a system of efficient collaboration was required to carry out Work Package 1 (the literature survey) effectively. Three different web collaboration and referencing tools were considered. These were EndNote, Zotero and Mendeley. The previous personal experience of some users of EndNote suggested that they were not keen to continue using EndNote, as they found it difficult to use and unintuitive to the extent that they were unaware of some of the functionalities because they could not easily be discovered, so this was not pursued. Zotero had not been previously used by any members of the team, so its viability was uncertain. There were some issues over whether Mendeley should be used, given the ability to infringe copyright that the tool possesses, but it was decided that this would be the tool used and that each team member should refrain from uploading material for which copyright violation may be an issue.

Legal issues were pursued further and it was decided that IMARES, as part of WUR and with full co-operation of their library (of Wageningen University), could offer to host this Mendeley Group, in a properly licensed way and are entitled to invite any collaborator to join our group. In other words for this use and this project IMARES has legal coverage.

Mendeley is a free reference manager and academic social network that can help organize research, collaborate with others online, and discover the latest research. It is available from the website http://www.mendeley.com/

With Mendeley Groups the project group gains the advantage of being able to quickly and efficiently collaborate on the literature survey. All group members have access and the ability to add references, access PDFs of references (if made available, something that should not be done for confidential sources or anything which might count as copyright violation), and add annotations and notes on the literature being reviewed. Mendeley Desktop software can be installed by each team member, with the benefit being some additional functionality, certain accesses and use of the Citation-plugin (for MS-Word). Mendeley is said to work well with on-line searches in Scopus and ScienceDirect.

All project team members working on the literature survey are being asked to provide opinion on the most important pieces of information (articles, reports, websites, books) they have to offer that are relevant to the project and the challenges that have been set, through Mendeley.

5. Central project repository (data register)

In our project proposal we presented a conceptual overview as to how we would create a register of data sets that would provide input to each of the challenges. During the first phases of the project we have refined this into a solution that can be directly implemented. Furthermore it can be exposed to search engines such that the knowledge gained in the check points can be utilised in standard web searches. This is shown below in Figure 5.1and described below.
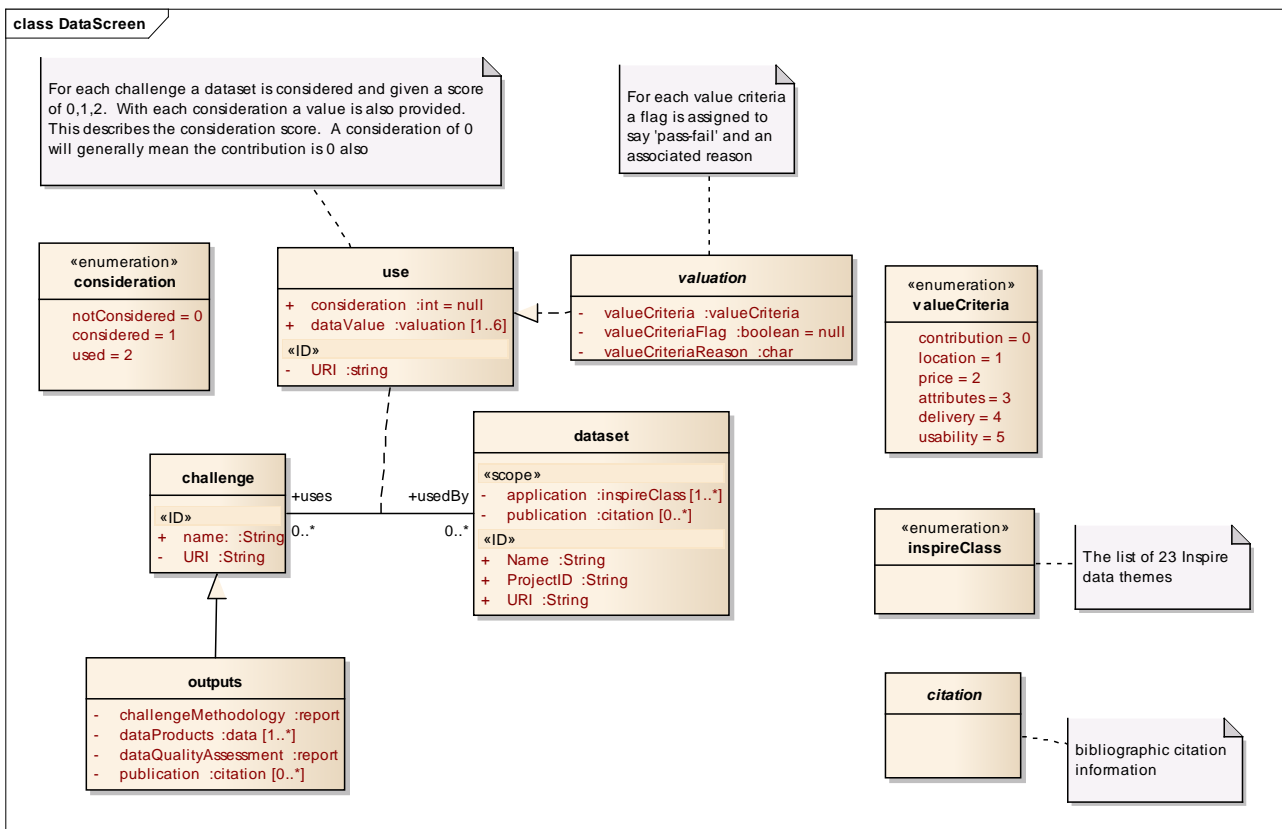

Figure 5.1: Data Repository Conceptual Model

5.1. Entities in data register

5.1.1. Dataset

A dataset identified by a given URI and meaningful name. Each data set has a broad classification according to the Inspire Annexes and also associated to a paper that may cite it. A dataset can be used by many challenges.

5.1.2.   Challenge

The name of the challenge to which a dataset may be applied, together with a URI to a webpage describing the challenge. A challenge delivers a range of outputs such as reports. A challenge uses many datasets.

5.1.3.   Use

The use of a particular data set in a challenge. Each data set has three consideration levels for a challenge ('notConsidered', 'considered' and 'used'). A 'valuation' is associated with a consideration instance to both qualify and quantify the consideration. The valuation is expressed in terms of an enumerated list of value criteria. There is a pass/fail flag associated with each criteria and an optional narrative reason for this flag. So for example, a data set may be 'consideration::notConsidered' for a challenge because valueCriteria::contribution has valueCriteriaFlag:="False" because valueCriteriaReason::"buildings dataset does not contain offshore structures". A data set used on a challenge would have valueCriteriaFlag:="True" for all valueCriteria.

This is illustrated further in Tables 5.1 to 5.3. This considers three data sets for use on the wind farm siting challenge where a high resolution bathymetry is needed. In the first case the data set is not considered as it contains no useful information. In the second case the data is considered for a challenge, but it is not used as the data does not meet the requirements of the challenge. In the last case the data does meet the requirements of the challenge.

Table 5.1: Example use record for a data set "not considered" for a challenge

Use	URI (links dataset URI to a challenge URI)	
Consideration	notConsidered	
ValueCriteria	**ValueCriteriaFlag**	**ValueCriteriaReason**
Contribution	F	Dataset contains no information that can be used on the challenge
Location	Null	
Price	Nul	
Attributes	Null	
Delivery	null	
Usability	null	

Table 5.2: Example use record for a data set "considered" for a challenge

Use	URI (links dataset URI to a challenge URI)	
Consideration	Considered	
ValueCriteria	**ValueCriteriaFlag**	**ValueCriteriaReason**
Contribution	T	Dataset contains gridded bathymetry which is required to develop a model of the seafloor for scour assessment
Location	F	Gridded bathymetry outside of study area
Price	T	Free to download
Attributes	F	Grid resolution of 10arcSec too coarse for modelling
Delivery	T	Web download
Usability	T	OGL

Table 5.3: Example use record for a data set "used" for a challenge

Use		URI (links dataset URI to a challenge URI)
Consideration	Used	
ValueCriteria	**ValueCriteriaFlag**	**ValueCriteriaReason**
Contribution	T	Dataset contains gridded bathymetry which is required to develop a model of the seafloor for scour assessment
Location	T	Gridded bathymetry within study area
Price	T	Licence fee €35
Attributes	T	Grid resolution of 1arcSec. Better than needed
Delivery	T	Web download
Usability	T	Raster grid format

6. Data portal development

6.1. Website

The NSCP data portal is implemented as a website. The website is to sit alongside other projects in the EMODnet family and therefore it is appropriate that it maintains EMODnet look, feel and branding. Given that there will ultimately be seven checkpoint projects (and therefore seven websites) some consistency in navigation between check point websites would also seem appropriate. Furthermore the implementation platform of the websites needs to be considered such that the user experience is consistent across the EMODnet family, for example ensuring browser compatibility.

With this in mind, the NSCP agreed with VLIZ (managing the central EMODnet website on behalf of DG MARE) that they would release the EMODnet website as a platform for other EMODnet projects. The current EMODnet website is built on the Drupal content management system (CMS) and HR Wallingford have been granted access to edit and publish pages under a top level "emdonet/northsea" domain. There is very limited effort in VLIZ allowing this as all the content is still created by the NSCP. The advantage is that the NSCP inherit the branding and styling from the main EMODnet page. In addition, migration issues at the end of the NSCP project are eased as the website is already being hosted by the EMODnet secretariat.

We do not suggest that all check point websites should be clones of one another. Each sea basin is unique and this will require information presented in ways that best reflect the issues within each basin. That said, there should be a level of consistency on the sort of information contained within each website, not least the outputs from each of the challenges.

The one point to bear in mind is that there is no formal contractual relationship between NSCP and EMODnet secretariat to share the EMODnet CMS. We would suggest therefore that approval of this progress report acts to formalise this co-operation.

6.2. Web services

The websites will provide a client to some web services to both discover and view data. At this stage we are still in discussion about the precise form that these web services should take and their ultimate purpose. Our

proposal suggested using a formal 'OGC approach' based on CSW (catalogue services), WMS/TS (view service) and FTP/Atom (Download service). This can be done, however we feel it is useful to reflect on what these services are trying to achieve and the target user community.

6.2.1. Checkpoints and data services

The check points, unlike the EMODnet thematic portals and the data services of Copernicus, are not data services first and foremost and so should not be replicating or duplicating what these services are doing. We feel the NSCP should be signposting users to these existing catalogue services and via the challenges highlighting how the data they contain can (or cannot) be used. We appreciate some of these data will be via catalogue services, others as simple file downloads. Where the NCSP is creating unique content then this could be accessed via the portal, however very little unique content will be created (and updated) and so formal web services are not considered to be necessary.

At this stage we have not reached a conclusion on the best option for the checkpoints and are presenting in this report options for consideration at the next steering committee. We consider the demand for web services using the wind farm siting challenge as an example.

6.2.2. Wind farm Spatial Data

The data for wind farm siting is extensive and comprehensive. Figure 6.1 shows an example of much of the spatial data overlaid in a particular study area; this excludes the non-spatial data such as time series and climates of oceanographic and meteorological data. Displaying all this in a Web Map Service (WMS) is cumbersome and although technically easy is not the best environment for spatial analysis – this is best performed in a desktop GIS. Given that the primary use case is "ease of download" then industry-standard data download services should be used (see Figure 6.2). This will allow users to 'clip, snip and ship' the relevant data package they need. This would be a good solution if we were developing a "wind farm siting package" for the North Sea basin, but we are not; we are looking at an aggregated data product for a single site. A third option is a 'Google maps' type approach where the map provides a context, with links to the services that can deliver the data. These links would be based on the value assessment conducted by the check point.

In this way a user could select 'wind farm siting' as a theme, and be able to see which data could be used and how it had been used (see Figure 6.3).
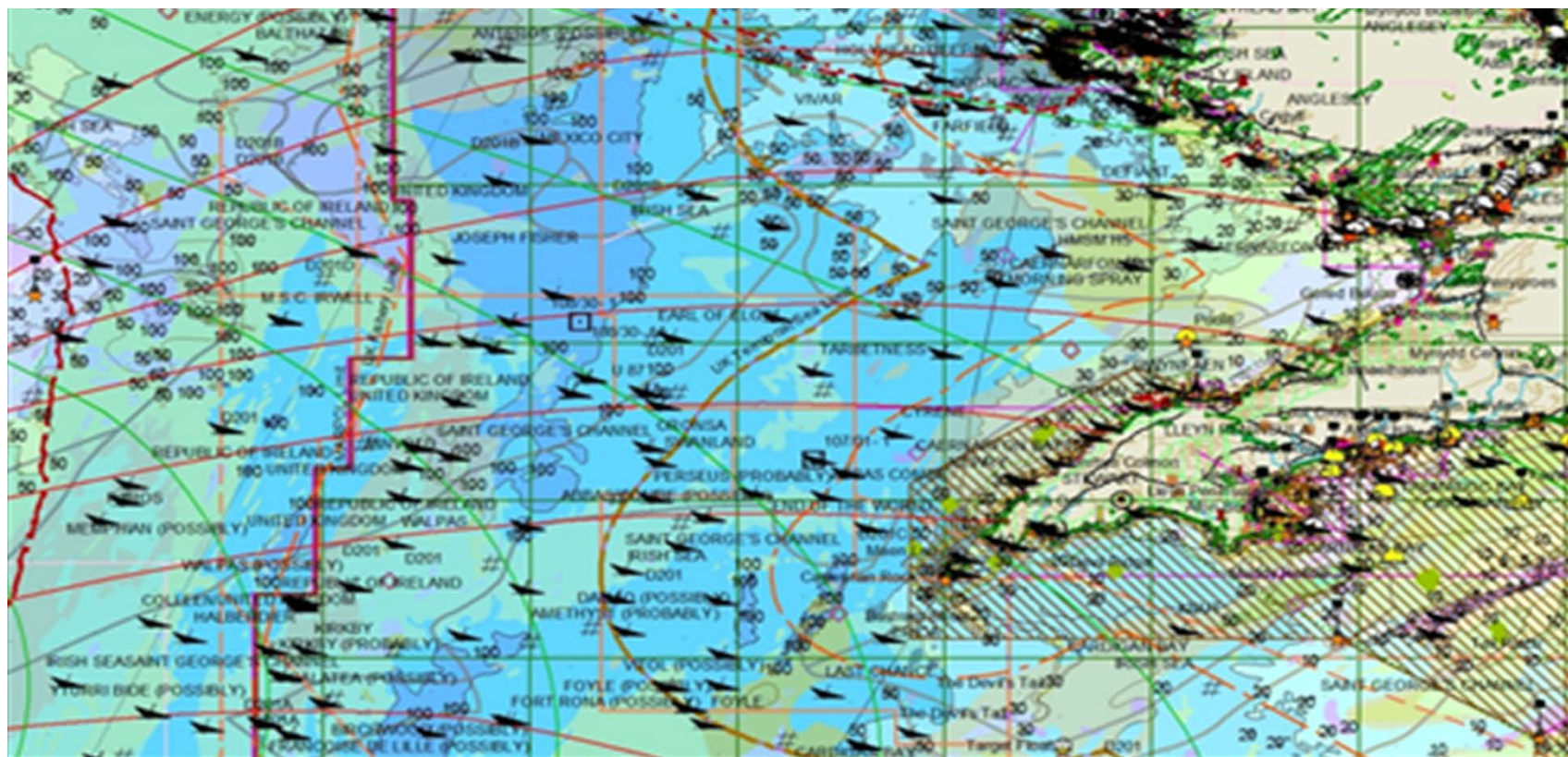
Figure 6.1: Example of spatial data for wind farm siting

Source: SeaZone HydroSpatial One in ArcGIS showing the multitude of spatial data layers needed
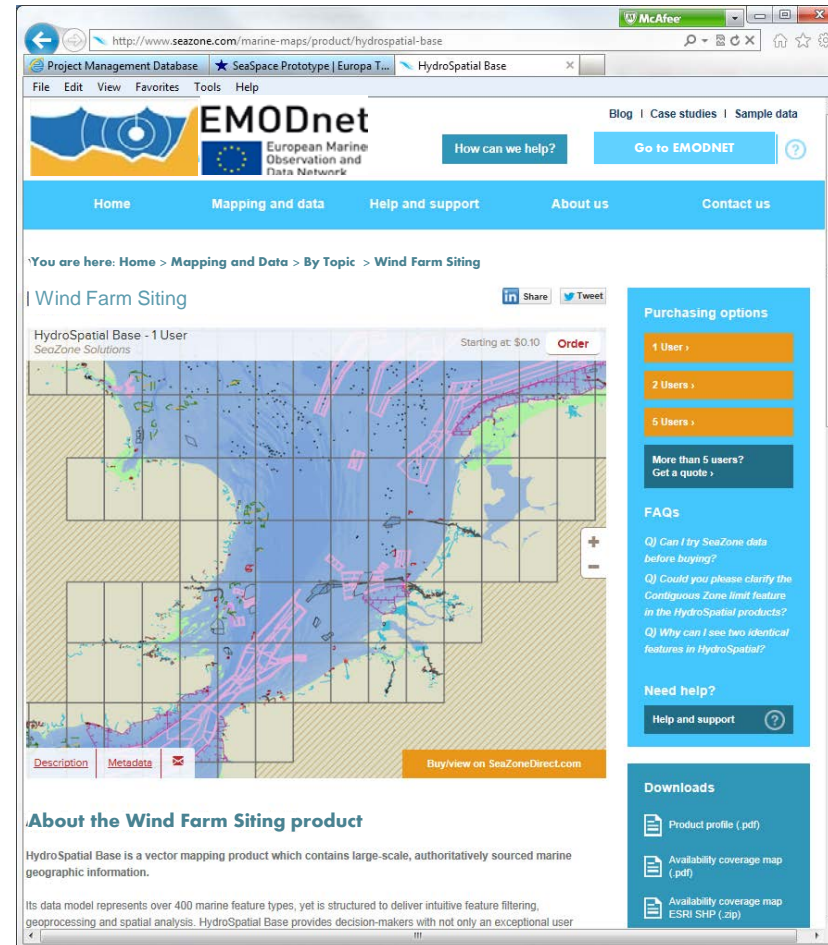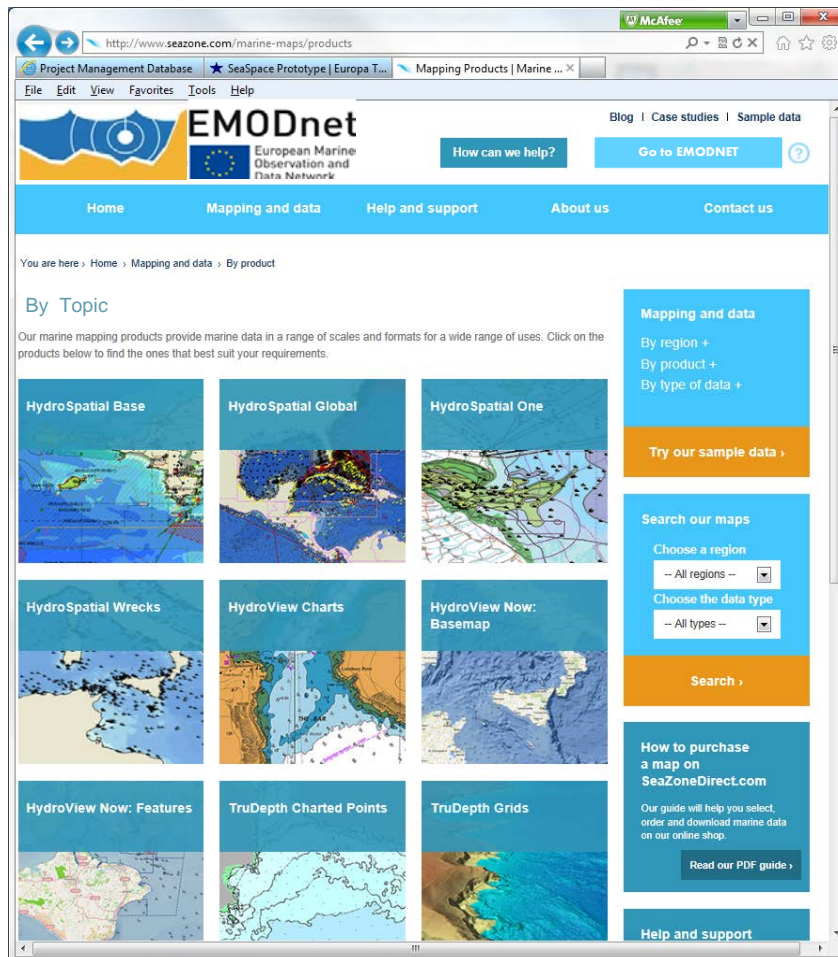
Figure 6.2: Example of a possible NSCP data download service

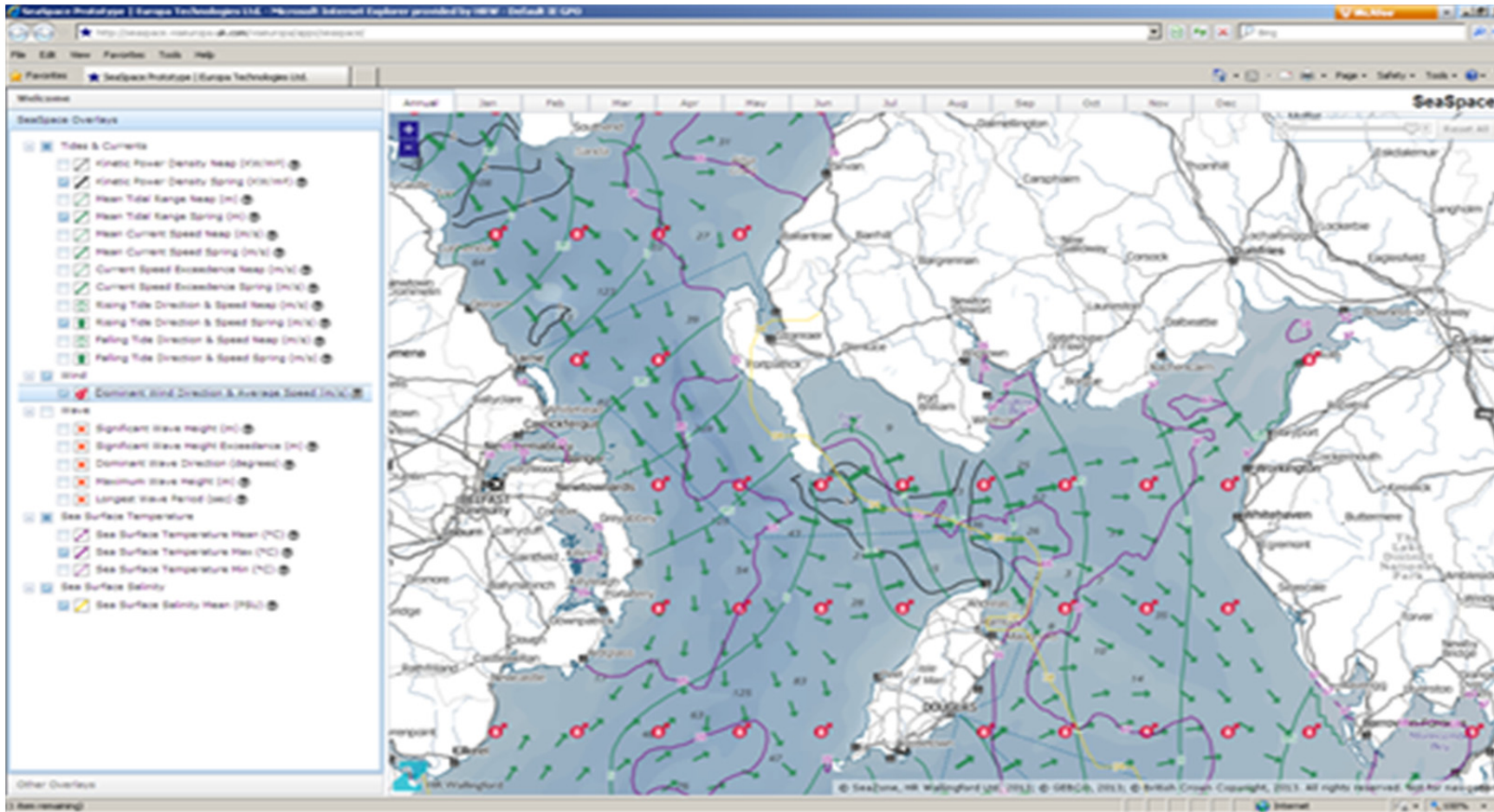*Source: This is an example of what a customer-facing data download service for the NSCP could look like*

Figure 6.3: Example context map for marine data

7. Key points for further discussion

- The oil platform leak challenge is ready for rehearsal; would DG-MARE be prepared to supply specifications in a fashion similar to the real thing?
- Is the broadening of the literature survey as we have outlined regarded as a useful activity?
- Sharing the EMODnet CMS between NSCP and EMODnet secretariat formally would be useful.
- The best output options for the checkpoints have not yet been decided, thus we are presenting options for consideration at the next steering committee in this report.

8. Conclusions

This is the first progress report for the North Sea Check Point (NSCP). It covers the 4 months period up to the end of April 2014.

The work on the project challenges has commenced. The literature review is in progressing and the bibliographic tool Mendeley is used to share and comment on papers and other publications. The tool will also be used for internal project document sharing. In the context of the actual challenges themselves, work is underway on the marine protected areas, climate and coastal protection and wind farm siting challenges. The oil platform leak challenge is further advanced and has been set up and is ready for a rehearsal.

Emphasis during this phase of the project has also been placed on the project infrastructure such as the website, web services and data repository. Within the scope of the contract there are many ways this could be done and although some variability is good, too much difference between the NSCP and MSCP and future checkpoints would be confusing to users. With this in mind the NSCP has presented in this report its options for the website, web services and data repository. These have been shared with the MSCP and are to be presented at the EMODNET steering group meeting in June 2014 to agree the best way forward.

HR Wallingford

Working with water
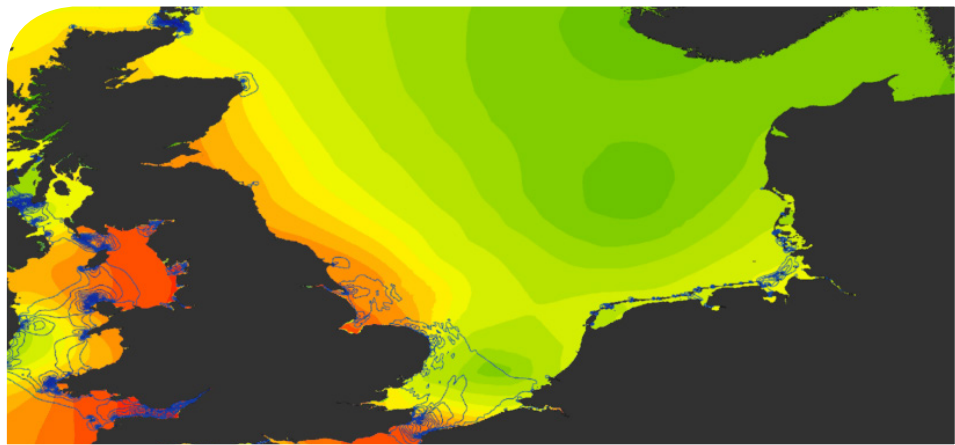
HR Wallingford is an independent engineering and environmental hydraulics organisation. We deliver practical solutions to the complex water-related challenges faced by our international clients. A dynamic research programme underpins all that we do and keeps us at the leading edge. Our unique mix of know-how, assets and facilities includes state of the art physical modelling laboratories, a full range of numerical modelling tools and, above all, enthusiastic people with world-renowned skills and expertise.

HR Wallingford, Howbery Park, Wallingford, Oxfordshire OX10 8BA, United Kingdom
tel +44 (0)1491 835381 fax +44 (0)1491 832233 email info@hrwallingford.com
www.hrwallingford.com