



score



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003534



DOCUMENT TRACKS DETAILS

Project acronym	SCORE
Project title	Smart Control of the Climate Resilience in European Coastal Cities
Starting date	01.07.2021
Duration	48 months
Call identifier	H2020-LC-CLA-2020-2
Grant Agreement No	101003534

Deliverable Information	
Deliverable number	5.1
Work package number	5
Deliverable title	Analysis of the Technical Features of Existing Databases
Lead beneficiary	UCD
Author	Yuri Calleo, Jose Barron, Francesco Pilla, Filippo Giannetti, Emilio Laino, Sebastiano Pinna, Alberto Ortolani, Luca Baldini, Olympia Papadopoulou
Due date	31/12/21
Actual submission date	20/12/2021
Type of deliverable	Report
Dissemination level	Public

VERSION MANAGEMENT

Revision table			
Version	Name	Date	Description
V 0.1	Francesco Pilla, UCD	13/12/201	First draft
V 0.2	Cécil Meulenberg, ZRS	18/12/2021	Updated draft internally reviewed
V 0.3	Neslihan Beden, Samsun	20/12/2021	Updated draft internally reviewed
V1.0	Francesco Pilla, UCD	20/12/2021	Final version
V1.1	Salem Gharbia, IT Sligo	20/12/2021	Coordinator review and Final version



All information in this document only reflects the author's view. The European Commission is not responsible for any use that may be made of the information it contains.

LIST OF ACRONYMS AND ABBREVIATIONS

Acronym / Abbreviation	Meaning / Full text
API	Application Programming Interface
CCLL	Coastal City Living Lab
CDMS	Climate Data Management System
CRM	Coastal Resilience Model
DIY kits	“Do it yourself” kits
EAA	European Environmental Agency
EBA	Ecosystem-Based Approach
EC	European Commission
ELT	extract, load and transform
EU	European Union
FAIR	Findable, Accessible, Interoperable, and Reusable
GIS	Geographic Information System
H2020	Horizon 2020
ICT	Information, Communication and Technology
NBS	Nature based Solutions
NNBFs	Natural and Nature-Based Features
OAL	Open-Air Laboratory
RCSMs	Regional Climate System Models
SIP	SCORE ICT Platform
WMO	World Meteorological Organization
WP	Work Package



BACKGROUND: ABOUT THE SCORE PROJECT

SCORE is a four-year EU-funded project aiming to increase climate resilience in European coastal cities.

The intensification of extreme weather events, coastal erosion and sea-level rise are major challenges to be urgently addressed by European coastal cities. The science behind these disruptive phenomena is complex, and advancing climate resilience requires progress in data acquisition, forecasting, and understanding of the potential risks and impacts for real-scenario interventions. The Ecosystem-Based Approach (EBA) supported by smart technologies has potential to increase climate resilience of European coastal cities; however, it is not yet adequately understood and coordinated at European level.

SCORE outlines a co-creation strategy, developed via a network of 10 coastal city 'living labs' (CCLs), to rapidly, equitably and sustainably enhance coastal city climate resilience through EBAs and sophisticated digital technologies.

The 10 coastal city living labs involved in the project are: Sligo and Dublin, Ireland; Barcelona/Vilanova i la Geltrú, Benidorm and Basque Country, Spain; Oeiras, Portugal; Massa, Italy; Piran, Slovenia; Gdansk, Poland; Samsun, Turkey.

SCORE will establish an integrated coastal zone management framework for strengthening EBA and smart coastal city policies, creating European leadership in coastal city climate change adaptation in line with The Paris Agreement. It will provide innovative platforms to empower stakeholders' deployment of EBAs to increase climate resilience, business opportunities and financial sustainability of coastal cities.

The SCORE interdisciplinary team consists of 28 world-leading organisations from academia, local authorities, RPOs, and SMEs encompassing a wide range of skills including environmental science and policy, climate modelling, citizen and social science, data management, coastal management and engineering, security and technological aspects of smart sensing research.



EXECUTIVE SUMMARY

This document is a deliverable of the SCORE project, funded under the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003534.

The aim of this document is to provide a comparative overview of the technical features of the databases produced within previous and current projects dealing with topics of relevance for SCORE, with particular focus on relevant software/hardware platforms, data formats and interfaces.

This deliverable presents the results of the detailed investigation, analysis and comparison of the technical features (e.g. software languages and technical environments, hardware platforms, data management, interfaces, etc.) of existing databases produced within initiatives relating to coastal resilience, EBA deployments and living lab approaches relevant to SCORE activities.

This deliverable used as input documents a series of reports, scientific publications, webpages, information leaflets, etc. about databases developed for both past and current initiatives, giving priority to projects funded by the European Commission involving the participation of multinational consortia.

LINKS WITH OTHER PROJECT ACTIVITIES

This deliverable will serve as a starting point for the design and development of the SCORE database as part of WP5. The links with WP3 and in specific with the Deliverable D3.1 of Task 3.1 have been considered during the preparation of this report. Overall, the database will be used: to enable access, for each CCLL, to both historical data (e.g., time-series of coastal meteo-oceanographic parameters, pre-existing databases and GIS mappings) and data from sensor networks owned/operated by local authorities as part of WP2; to provide access to raw data to be fed into the models and the sharing with other WPs of the processed data about projections and evolution trends as part of WP3; to store data produced by either SCORE sensors deployed in the CCLLs and citizen science activities using DIY kits, low-cost commercial sensors/acquisition devices, smartphones, etc. as part of WP4; to provide access to data to analyse the risk and the socio-economic impact, respectively, of climate changes as part of WP6 and WP7; to complement the digital twin solutions developed therein as part of WP8.



TABLE OF CONTENT

1. Introduction	7
1.1. Scope of the report _____	7
1.2. Framework for analysis and comparison _____	7
1.2.1. Synthesis matrix parameters	8
2. Relevant databases	10
2.1. Review cases _____	10
2.1.1. Synthesis matrix.....	14
2.1.2. Accessibility of data	17
3. Discussion & Conclusions	19
3.1. Similarities and differences between databases _____	19
3.1.1. Recommendations.....	20
3.2. Conclusions _____	20





1. INTRODUCTION

The scope of Work Package 5 is to develop and manage the SCORE ICT Platform (SIP), featuring the SIP together with the relevant interfaces, which will collect, store and share the heterogeneous data acquired and processed during the SCORE project, while ensuring compliance with FAIR principles. In particular, the SIP platform will ensure long-term data storage and will therefore represent an efficient tool for sharing knowledge on EBAs. The first task of this Work Package, Task 5.1, deals with the technical analysis and comparison of existing databases relevant to the SCORE Activities. The scope of this report, which is the deliverable resulting from the activities in Task 5.1, is detailed in the following section.

1.1. Scope of the report

This task is dedicated to being a detailed investigation, analysis and comparison of the technical features (e.g., software languages and software environments, hardware platforms, data management, data interfaces, etc.) of existing databases that have been produced within initiatives about coastal resilience, EBA deployments and living lab approaches, relevant to the SCORE activities. The input documents for Task 5.1 will be: reports, scientific publications, webpages, information leaflets, etc. from both past and current initiatives. Sources of historical data sets about the impacts of climate change and past extreme events on coastal cities, collected during WP1 and to be stored in SIP, will be also considered in the analysis activities. The results of Task 5.1 will produce a document, Deliverable D5.1, containing an overview of the technical features of the databases previously that produced within previous and current projects dealing with topics of relevance for SCORE, with a particular focus on relevant software/hardware platforms, data formats and interfaces, which will be issued at M6.

1.2. Framework for analysis and comparison

The investigation mainly focused on the major European and international research frameworks, such as EU-funded **H2020** (<https://ec.europa.eu/programmes/horizon2020/>), **LIFE** (https://cinea.ec.europa.eu/life_en) and **Interreg** (<https://interreg.eu/>) programs, and other research initiatives supported by international agencies and organizations, such as the **European Space Agency (ESA)**, (<https://www.esa.int/>), the **European Environment Agency (EEA)**, (<https://www.eea.europa.eu/>), the **World Meteorological Organization (WMO)**, (<https://public.wmo.int/en>), the **World Climate Research Programme (WCRP)**, (<https://www.wcrp-climate.org/>), the **Intergovernmental Panel on Climate Change (IPCC)**, (<https://www.ipcc.ch/>), etc..

The core idea is to develop a hands-on and practical set of technical features based on best practice from other European and international projects to develop the SIP.

The SIP will need to support:

- as part of **WP2**, the activities with stakeholders aimed at enabling access, for each CCLL, to both historical data (e.g., time-series of coastal meteo-oceanographic parameters, pre-existing databases and GIS mappings) and data from sensor networks owned/operated by local authorities;
- as part of **WP3**, the procurement of the raw data to be fed into the models and the sharing with other WPs of the processed data about projections and evolution trends;
- as part of **WP4**, the acquisition and storage of data produced by either SCORE sensors deployed in the CCLLs and citizen science activities using DIY kits, low-cost commercial sensors/acquisition devices, smartphones, etc.;



- as part of **WP6** and **WP7**, the acquisition of WP3's projections to be processed with suitable models for assessing and disseminating the results about the risk analysis and the socio-economic impact, respectively, of climate changes;
- as part of **WP8**, complement the digital twin solutions developed therein.

As such, to guarantee the relevance of the previously used best practice examples, the technical features of the existing databases of projects in the areas of “Coastal Resilience, “Ecosystem-Based Approaches” and “Living Labs” will be reported in the following sections to develop relevant synthesis matrices for the SIP. The synthesis matrix is a valuable and practical tool to organise/tabulate the findings from the review of relevant reports on existing databases and thus provide a visual representation of the technical features of the examined existing databases and to identify overlaps in terms of technical features of best practice existing databases.

1.2.1. Synthesis matrix parameters

A preliminary assessment of deliverables from other European projects (analysed in more detail in the following sections) has allowed drafting the following list of technical features to be included in the synthesis matrix:

In relation to the database management aspects.

- **Data models.** Collection of concepts, graphics, data definitions, dictionaries, and common vocabularies to describe and communicate the data elements, data structure, relationships, and data processing operations in a database. Thus, help to understand the phenomena and real-world problems that a database is designed to address.
- **Multi-model data platform.** Manages multiple data types and database models such as relational, non-relational (document, graph, key-value, etc.) or object-oriented. Integrates structured data, semi-structured data and unstructured data and content such as text, social media postings, documents, photos, satellite images, graphics, animations, audio, or video. Also includes spatial and non-spatial data.
- **Data quality management.** Use of policies, procedures, metrics, technology, and resources for documenting, ensuring and improving the data quality. Monitoring and use of data quality measures, best practices during data creation, and data quality techniques such as data profiling, data cleansing, data validation, and data quality checks and reviews.
- **Data governance.** Definition of data standards, policies, and rules to control data usage, availability, usability, integrity, and security throughout the information lifecycle of a database system. Practices to ensure data is secure, reliable, trustworthy, and properly used. Including actions to manage sensitive data, data privacy and data protection. Elements for documenting data assets, their classification and flows through a system, business glossaries, common terms and definitions, data catalogues and standard data formats applied in all system units to monitor the usage of data and enforce policies.
- **Data integration.** Integration environment and methods available to help users manage and simplify the process of preparing, transforming, exchanging, and combining multiple data sources for operational and analytical uses. Offering tools such as ETL (extract, transform and load), ELT (extract, load, transform), big data integration, real-time integration, streaming integration, APIs, etc.
- **Data interoperability.** Capabilities to communicate, execute processes, exchange and transfer data among various internal system environments and external information systems, e.g., by enabling standard APIs.
- **Meta-data management.** Creation and use of meta-data: structured reference data that describe the characteristics and lineage of other data. It provides context and better organises a data object by using terms associated with that particular object. It helps to sort, search and identify the attributes of the information described by the meta-data. The goal of meta-data management is to enable access to high



quality, integrated meta-data to make it easier to locate a specific dataset through the design of a meta-data repository.

- **Data standards.** Use of rules and guidelines by which data are described, documented, recorded, and encoded to facilitate data exchange, understanding and interoperability. For example, dataset-level standards and parameter-level standards (developed by CEN, ISO, Open Geospatial Consortium, etc.).
- **Querying and reporting tools.** End-user tools and applications to run ad hoc queries against a database. They facilitate data search and discovery, filtering and simple formatting for further analysis and for designing report layouts and creating organised listings.
- **Analytics.** Tools, functionalities, and methods available for computational analysis of data and statistics to discover, interpret, and communicate significant data patterns to generate new insights and knowledge. Offering tools and interfaces such as data visualization (charts, graphs, dashboards, maps), storytelling, data and text mining, predictive analytics, data science, simulations and decision-making applications, metrics, etc.
- **Data sharing and Open Data.** End-user tools (download functionalities, APIs, web services, etc.) to provide data access, use and sharing by anyone. Making data available online in a common understandable format and under a license that should allow for free re-use, transformation, and combination of data, including for commercial purposes.
- **Crowdsourced data latency.** Use of methods and tools to control latency of data provided by a volunteer or citizen scientist to improve the system performance (important if it is a time-critical situation such as crisis response or crisis management for instance). For example, by controlling the time it takes to recruit a volunteer, the speed of response to initiate data collection and transfer, managing the size and location of an available pool of volunteers, the batch size of the crowdsourced task, the design of the task or the assignment of tasks, etc.

Other aspects:

- **Case studies catalogues.** Listings and tools to explore the case studies available in the system, filter them through different criteria and their geographic location, providing access to the extended case study description and linked reports and outcomes.



2. RELEVANT DATABASES

Several initiatives and projects were identified to review and compare their data management aspects. Their databases are relevant to the SCORE activities as they manage information and knowledge on climate change adaptation and local and regional resilience. In particular, coastal resilience and protection, coastal vulnerability, coastal change and erosion, flood risks and impacts, etc. As well as climate data management, models, simulations and predictions, and knowledge for understanding climate variability and extreme weather events, for decision-making, and planning adaptation and response measures.

2.1. Review cases

A presentation of the reviewed cases will be provided below in order to understand more about the databases synthesized in the matrix 2.1.1.

The **Adriadapt** project (<https://adriadapt.eu/>) aimed at supporting the building of local and regional resilience by developing the knowledge base required to identify and plan appropriate climate change adaptation options. In order to achieve this goal, from the beginning of 2019 to June 2021, a set of activities was carried out. High-resolution climate projections with detailed information on climate parameters for the Adriatic regions have been produced, and this information is available on the platform. This knowledge platform for the Adriatic region contains an overview of different adaptation options, case studies, guidance documents, legal frameworks, and other useful material on climate change adaptation. The knowledge platform was tested during the implementation of Adriadapt pilot projects.

The **Operandum** project (<https://www.operandum-project.eu/>) is an important project funded by the European Union's Horizon 2020 research and innovation programme and is based on Nature-Based Solutions (NBS). Specifically, all the site-specific and innovative NBS are co-designed, co-developed, deployed, tested and demonstrated with partners and local stakeholders in 10 open-air laboratories (OALs). These provide concrete, flexible and transportable frameworks in order to expand the adoption of green/blue/hybrid infrastructures across Europe and in developing countries.

The **CoastalRes (CRM)**, (<https://coastalmonitoring.org/ccoresources/coastalres/>) was a project funded by the UK Strategic Fund Climate Resilience Programme with the purpose of investigating the changing nature of coastal and estuarine erosion and flood-risk in order to develop new policies and strategies. In this project, to get a deeper understanding of the problem, they involved a large number of stakeholders and several workshops were held in London, Havant and York. The report shows a summary of the data and geospatial data development process undertaken by the GeoData Institute in order to demonstrate the results of a new Coastal Resilience Model (CRM).

The **Climate Data Management System Specifications** (https://library.wmo.int/doc_num.php?explnum_id=7867) is a detailed publication developed by the World Meteorological Organization (WMO) in order to establish an overview of climate data management system (CDMS).



The CDMS is an integrated computer system that facilitates the effective storage, management, analysis, delivery and use of a wide range of climatic data. The publication framework illustrates a series of interconnected building blocks - but not exclusively to be used together - called components. Each component describes a specific functional requirement of a CDMS and contains references where appropriate for more information.

The “**Implementing resilience: the barriers and opportunities in inventorying natural resiliency project proposals**” report (published 2019), starts from the assumption that Natural and nature-based features (NNBFs) can mitigate the impacts of flooding and storm surge on coastal communities, providing critical habitat and improving water quality. The report’s goal was to create a database in order to connect projects with possible funding streams, proposing an inventory of grants to embed within the database.

https://static1.squarespace.com/static/56af7134be7b96f50a2c83e4/t/5d715eba77a52f0001781832/1567710915543/Final_CZM_Database_Report_webupload.pdf

The **Resilience Evaluation and Needs Assessment (RENA)**, (<https://deq.nc.gov/about/divisions/coastal-management/coastal-adaptation-and-resiliency/dcm-coastal-resilience-pilot-program>) pilot project is the first part of a five-year project that aims to provide a comprehensive guide or online portal for building resilience in USA North Carolina's coastal communities. As reported by the final Report, they will serve local governments in four points:

- illustrating successful planning of frameworks to engage communities in resilience planning;
- identifying common needs from local governments and compiling state and federal resources that can address those needs;
- providing case studies across the North Carolina coast that can help guide other local governments with planning;
- discussing adaptation and mitigation strategies that can be applied in a variety of situations experienced by coastal communities in the state of North Carolina.

The **Climate Adaptation Platform (Climate-ADAPT)**, (<https://climate-adapt.eea.europa.eu/>) has been developed in partnership between the European Commission and the European Environmental Agency (EEA). This platform aims to be the European reference information system for climate change adaptation, supporting better informed decision-making to contribute to a more climate-resilient Europe. Their main objectives are: (a) to facilitate the collection, sharing and use of information on the impacts of climate change; (b) to help decision-makers effectively assimilate the relevant knowledge; and (c) to contribute to better coordination among sectors and institutional levels. Climate-ADAPT helps users to access and share data and information on Europe’s projected climate change, regional vulnerability, adaptation strategies and actions, adaptation case studies and possible adaptation options, as well as tools to support adaptation planning. The information is accessible and organised under the following main categories: (a) EU Policy, including EU Adaptation Policy, Adaptation in EU Policy Sectors (e.g., Coastal areas, Forestry, Water management, Marine and fisheries, etc.) and EU Regional Policy; (b) Countries, Transnational regions, Cities; (c) Knowledge: Topics, Data and indicators, Research and innovation projects, Tools, Practice; (d) European Climate and Health Observatory; and (e) Networks.



The **Natural Water Retention Measures (NWRM)** (<http://nwrn.eu/>) EU platform provides a comprehensive database of NBAs that aim to protect water resources and address water-related challenges, with technical specifications and several case studies applications throughout the EU, including coastal protection. The platform presents information and a collaborative knowledge base on the potential benefits and impacts of a wide range of NWRM and their practical applications in different European contexts.

DINAS-COAST (<https://cordis.europa.eu/project/id/EVK2-CT-2000-00084/it>) was a project funded by the European Union, which brought together the best available science and data to help policy to interpret and evaluate coastal vulnerability to climate change. It provides a database that contains regional and international coastal information. Databases also inform administration, research, and business experts, as well as students about coastal-related conferences, projects, documents and courses, and offer photographs of the coast.

The **European Marine Observation and Data Network (EMODnet)** (<https://emodnet.ec.europa.eu/en>) is a network of organisations supported by the European Union's integrated maritime policy. Rapid access to reliable and accurate information is vital in the development of policies and legislation to protect vulnerable areas of coasts and oceans, in understanding trends and in forecasting future changes. Towards these goals, EMODnet's organisations work together to observe the sea, process the data according to international standards and make that information freely available as interoperable data layers and data products. EMODnet provides a database with European marine data, including data concerning the coast.

ANYWHERE (<http://anywhere-h2020.eu>) is a project funded by the European Union's H2020 research and innovation programme. The principal objective of ANYWHERE is to enable society as a whole and the main civil protection agencies to respond rapidly to extreme climate and weather events. The Project provides a platform on extreme climate risks that will enable us to identify critical situations that could lead to loss of life and economic damages. Such early-warnings should enable us to improve protection measures. The platform will therefore serve as a decision-making tool for various authorities when faced with a situation of crisis.

Med-CORDEX (<https://www.medcordex.eu/>) is an initiative proposed by the Mediterranean climate research community as a follow-up of previous and existing initiatives. More specifically, it is a coordinated contribution to CORDEX that is supported by HyMeX and MedCLIVAR international programs. Med-CORDEX is a unique framework where the research community will make use of both regional atmospheric, land surface, river and oceanic climate models, coupled to regional climate system models for increasing the reliability of past and future regional climate information and understanding the processes that are responsible for the Mediterranean climate variability and trends. Med-CORDEX takes advantage of new fully coupled Regional Climate System Models (RCSMs) and provides a database of numerical simulations, which is a repository of numerical files produced by research groups within their project activities.



The goal of the **COordinated Regional climate Downscaling EXperiment (CORDEX)**, (<https://cordex.org/>) is to produce a regional downscaling, which is both an important research topic and an opportunity to engage a broader community of climate scientists in its activities.

The **HYdrological cycle in the Mediterranean EXperiment (HyMeX)**, (<https://www.hymex.org>) project aims at a better understanding, quantification and modelling of the hydrological cycle in the Mediterranean, with emphasis on the predictability and evolution of extreme weather events, inter-annual to decadal variability of the Mediterranean coupled system, and associated trends in the context of global change. The multidisciplinary research and associated database developed within HyMeX aim to improve the observational and modelling systems, better predict extreme events, simulate the long-term water-cycle, and provide guidelines for adaptation measures.

Mediterranean CLimate VARIability and predictability (MedCLIVAR), (<http://www.medclivar.eu/>) is a scientific network to promote better communication among different scientific disciplines and to develop a multidisciplinary vision of the evolution of the Mediterranean climate through studies that integrate atmospheric, marine, and terrestrial climate components at time scales ranging from paleo reconstructions to future climate scenarios. The program deals with scientific issues including past climate variability; connections between the Mediterranean and global climate; Mediterranean Sea circulation and sea level; feedback on the global climate system; regional responses to greenhouse gases, air pollution, and aerosols, as well as regional impacts of climate change.

MAREGOT (<http://interreg-maritime.eu/fr/web/maregot/projet> and <http://interreg-maritime.eu/it/web/maregot/progetto>) is a project aimed at prevention and joint management of risks deriving from coastal erosion in the maritime area that affects Corsica, Liguria, Sardinia, the PACA Region and Tuscany. The project intends to launch a shared project action which, thanks to a better knowledge of erosion phenomena and the dynamics of the coasts, identifies the best practices for the management of the transboundary coastal territory and allows to reduce the risks associated with climate change.

CONSCIENCE (<http://www.conscience-eu.net/>) was a project funded by the European Union with a view to enhancing the implementation of a scientifically based sustainable coastal erosion management in Europe. It has been testing scientific concepts and tools in six pilot sites around Europe. The field sites have been selected to represent the most common European coastal types: sand, gravel/shingle and cliffed coasts; high-energy open coasts (Atlantic coasts), medium-energy coasts (North Sea coasts, English Channel coasts) and low-energy coasts (Mediterranean coasts, Baltic Sea coast, Black Sea coasts). This project provides many reports and documents about management of coastal erosion.



2.1.1. Synthesis matrix

The following table presents all the cases synthesized in a matrix using the variables defined in section 1.2.1.

	PROJECTS			
Technical Features	Adriadapt	Operandum	CRM	WMO
Data models	Available (Summaries and climate projections)	Available (maps, graphs and summary matrices)	Available (maps, graphs and summary matrices)	Available (Summary matrix)
Multi-model data platform	Multiple data available	Multiple data available	Multiple data available	Multiple data available
Data quality management	Available. Integrated Adaptation Planning Tool	Data profiling, data cleaning, data validation, and data quality checks and reviews	Data cleaning and profiling	Data profiling, data cleaning, data validation, and data quality checks and reviews
Data governance	Not specified	Available	Not specified	Not specified
Data integration	Not specified	Available integration	Not specified	Available integration
Data interoperability	Not specified	API information available	Not specified	API information available
Meta-data management	Not specified	Available	Available	Available
Data standards	Available. Guideline and handbook	Available	Not specified	Available
Querying and reporting tools	Not specified	Available	Not specified	Available
Analytics	Not specified	Available	Available	Available
Data sharing and Open Data	Not specified	Available	Available	Available
Crowdsourced data latency	Not specified	Not specified	Not specified	Not specified
Case studies catalogues	Not specified	Provided	Provide several case studies	Provide several case studies

	PROJECTS				
Technical Features	CZM	RENA	Climate-ADAPT	NWRM	DINAS-COAST



Data models	Available (summary matrix)	Available (maps, graphs and summary matrices)			
Multi-model data platform	Multiple data available	Multiple data available	Multiple data available	Multiple data available	Multiple data available
Data quality management	Data cleaning and profiling	Data cleaning and profiling	Data quality assurance, review procedures	Review procedures	Data quality assurance, review procedures
Data governance	Not specified	Not specified	Implemented	Not specified	Available
Data integration	Not specified	Not specified	Not implemented	Not implemented	Not specified
Data interoperability	Not specified	Not specified	Not implemented	Not implemented	API information available
Meta-data management	Not specified	Not specified	Available	Available	Available
Data standards	Available	Not specified	Not specified	Not specified	Available
Querying and reporting tools	Not specified	Not specified	Available	Available	Not specified
Analytics	Available	Available	Available	Not implemented	Not specified
Data sharing and Open Data	Not specified	Not specified	Available	Available	Available
Crowdsourced data latency	Not specified	Not specified	Not implemented	Not implemented	Not specified
Case studies catalogues	Provided	Provided	Available	Available	Available

	PROJECTS			
Technical Features	EMODnet	ANYWHERE	Med-CORDEX	CORDEX
Data models	Available (maps graphs and summary matrix)	Available (collection of concept, maps and summary matrix)	Available (collection of concept and summary matrix)	Available (collection of concept and summary matrix)
Multi-model data platform	Multiple data available	Multiple data available	Multiple data available	Multiple data available



Data quality management	Data profiling and data cleaning	Data profiling	Data profiling	Data profiling
Data governance	Available	Available	Available	Available
Data integration	Not specified	Not specified	Available	Available
Data interoperability	API information available	Not specified	API information available	API information available
Meta-data management	Available	Not specified	Available	Available
Data standards	Available	Not specified	Available	Available
Querying and reporting tools	Available	Not specified	Available	Not specified
Analytics	Available	Not specified	Available	Available
Data sharing and Open Data	Available	Not specified	Available	Available
Crowdsourced data latency	Not specified	Not specified	Not specified	Not specified
Case studies catalogues	Available	Available	Available	Available

	PROJECTS			
Technical Features	HyMeX	MedCLIVAR	MAREGOT	CONSCIENCE
Data models	Available (collection of concept and summary matrix)	Available (collection of concept, maps and summary matrix)	Available (maps graphs and summary matrix)	Available (collection of concept, maps and summary matrix)
Multi-model data platform	Multiple data available	Multiple data available	Multiple data available	Multiple data available
Data quality management	Data profiling	Data profiling and data cleaning	Best practice and data profiling	Best practice and data profiling
Data governance	Available	Available	Not specified	Available
Data integration	Not specified	Available	Not specified	Available
Data interoperability	Not specified	Not specified	Not specified	Not specified
Meta-data management	Available	Available	Not specified	Available



Data standards	Available	Available	Not specified	Available
Querying and reporting tools	Available	Available	Not specified	Available
Analytics	Available	Available	Not specified	Not specified
Data sharing and Open Data	Available	Not specified	Not specified	Available
Crowdsourced data latency	Not specified	Not specified	Not specified	Not specified
Case studies catalogues	Available	Available	Available	Available

2.1.2. Accessibility of data

The following table shows whether a project is still active today, and whether its outcomes and data are accessible and available for use.

Project	Connectivity to database (Y/N)	Ongoing (Y/N)	Accessible data	Link
Adriadapt	N	Y	Not available, ongoing project	N/A
Operandum	Y	Y	Data models, API information, case studies	https://geoikp.operandum-project.eu/
CRM	Y	N	Data models, several case studies	https://coastalmonitoring.org/realtimedata/
WMO	Y	N	Data models, API information, guidelines	https://climatedata-catalogue.wmo.int/explore
CZM	Y	N	Data models, case study	N/A
RENA	Y	N	Data models, GIS datasets	https://deq.nc.gov/about/divisions/coastal-management/coastal-adaptation-and-resiliency/data-support-coastal-adaptation
Climate-ADAPT	Y	Y	Data models, indicators, several case	https://climate-adapt.eea.europa.eu/knowledge/adaptation-information/climate-services/



			studies	
NWRM	Y	N	Data models, data access per sector, case studies	http://nwrn.eu/list-of-all-case-studies
DINAS-COAST	Y	N	Data models, API information, case study	N/A
EMODnet	Y	Y	Data models, API information, data portal and data portfolio,	https://emodnet.ec.europa.eu/en/portals
ANYWHERE	Y	Y	Data models, access to deliverables documents,	http://aqua.upc.es/anywhere-catalogue-v2/
Med-CORDEX	Y	Y	Data models, API information, case studies	https://www.medcordex.eu/medcordex.php
CORDEX	Y	N	Data models, API information, impact and regional data portals	https://cordex.org/data-access/
HyMeX	Y	N	Data models, case studies	https://mistrals.sedoo.fr/HyMeX/
MedCLIVAR	Y	Y	Data models, case studies	http://www.medclivar.eu/index.php/how-to-join
MAREGOT	Y	Y	Data models, access to deliverables documents	http://interreg-maritime.eu/fr/web/maregot/projet
CONSCIENCE	Y	N	Data models, access to deliverables documents, case studies	N/A



3. DISCUSSION & CONCLUSIONS

This section focuses on the analysis of the similarities and differences between the analysed projects' databases in order to derive a set of practical recommendations for the database to be developed for the SCORE project.

3.1. Similarities and differences between databases

From a first overview, the databases shown in the tables 2.1.1 present different characteristics in most of the variables. Most of the databased projects illustrate information such as data summaries, graphs, spatial maps and synthesis matrices in order to describe and communicate better the data management, the construction of the databases and as described previously the real-world problems. From what is described in the available databases, multiple data are available i.e., photos, satellite images and also spatial and non-spatial data are provided. For example, many projects integrate structured data, semi-structured data and unstructured data in their reports, the Climate Data Management System Specifications provided interesting guidelines defining the functionality of the CDMS.

The management of the data information is available and illustrates different tools, such as adaption and planning tools, regarding the **Operandum** project different information is available, it is possible to find the profiling of the data, the data cleaning, data validation, and data quality checks. Many projects include different reviews, for example in the **NWRM** project there is a detailed review of the procedures.

Most of the databases do not specify information about data governance, however some of them including **Operandum** project, **ANYWHERE** project, **CORDEX**, provided different details such as data standards and control of the used data. We found that the integration of the software environment and methods available to help users manage and simplify the data preparation are not always available, however the projects that do, **Operandum**, **WMO**, **DINAS-COAST**, **EMODnet**, **Med-CORDEX** and **CORDEX** provide detailed APIs information.

In order to help to search and identify the attributes of the information, different meta-data management information is available with different data standards, for example, **Adriadapt** and **CWO** published guidelines and handbooks to improve readability and understanding of the data. Querying and reporting tools and analytics are most of the time available with end-user tools and facilities. Use of methods and tools to control latency of data are not available, instead many study cases are reported in the projects.

The table in paragraph 2.1.2 shows the accessibility of the data, specifying if the project is ongoing, the possibility of connection to the database and the type of data available. Most of the cases provide access to the database, including, as mentioned before, intuitive tools for the readability of the data and results, some projects results ongoing (e.g., **Adriadapt**, **Climate-ADAPT**, **EMODnet** etc.) and some projects are concluded (e.g., **CRM**, **WMO**, **RENA**, **CORDEX** etc.). Regarding the accessibility of the data, API information is provided in the above mentioned projects with different case studies. The **RENA** project makes available a Geographic Information System (GIS) dataset divided into three categories: environmental, infrastructure and social vulnerability.

Some of the studied databases including **ANYWHERE**, **MAREGOT**, **CONSCIENCE** report that all the derivables of the projects provide interesting data accessible to the public without any need for authorization, while other projects including **CORDEX** and **EMODnet**, propose a different division, the first with an impact and regional data portals, the second by making available a data portal and a data portfolio.



3.1.1. Recommendations

A set of recommendations were extracted from the comparison and analysis of the technical characteristics of the assessed databases which should be used as a starting point for the development of the SCORE project database:

- future proof and guarantee of the interoperability of datasets;
- details such as data standards and control of the used data should be provided;
- software environment and methods to help users manage and simplify the data preparation, should be integrated in the database and made available to users;
- information such as data summaries, graphs, spatial maps and synthesis matrices should be illustrated in order to describe and communicate better the data management;
- multiple data formats should be available, such as photos, satellite images and also spatial and non-spatial data by integrating structured data, semi-structured data and unstructured data;
- information about the management of the data should be available and should illustrate different tools, such as adaption and planning tools, profiling of the data, data cleaning and validation procedures, and data quality checks;
- a meta-data management structure needs to be in place to help to search and identify the attributes of the information, following best practice data standards;
- end users tools for querying, reporting and analytics should be available and should be tailored to different end-users;
- intuitive tools to provide access to the data and results stored in the database;
- API information should be provided.

3.2. Conclusions

The analysis of the databases developed for the projects listed in Section 2 highlighted the high variability between the adopted approaches and technical solutions, but also allowed us to identify a series of best practices which are listed in the previous section. The high heterogeneity of the data types and formats also identified potential issues with interoperability and a need to future proof the datasets produced as part of the activities in the SCORE project. Databases of projects funded by the European Commission play an essential role in the promotion of Open Data practices across and within each country and region, as they represent central points to publish and access research datasets. In order to promote the sustainability of these open (or partially open) data infrastructures, there are some principles that should be followed, as recommended by the “The Future of Open Data Portals” report published by the European Commission in 2020¹:

- Organising for use of the datasets (rather than simply for publication);
- Learning from the techniques utilised by recently emerged commercial data marketplaces; promoting use via the sharing of knowledge, co-opting methods common in the open source software community;
- Investing in discoverability best practices, borrowing from e-commerce;
- Publishing good quality meta-data, to enhance reuse;
- Adopting standards to ensure interoperability;
- Co-locating tools, so that a wider range of users and re-users can be engaged with;
- Linking datasets to enhance value;

¹ “Analytical Report 8: The Future of Open Data Portals”, European Commission, 2020. Accessed on 14/12/21: https://data.europa.eu/sites/default/files/edp_analyticalreport_n8.pdf



- Being accessible by offering both options for big data, such as APIs, and options for more manual processing, such as comma separated value files, thus ensuring that a wide range of user needs are met;
- Co-locating documentation, so that users do not need to be domain experts in order to understand the data.

The ten recommendations are not to be understood as an abstract list, but more like a tool that can be used in critically assessing the As-Is state of portals and taking an honest inventory. With the envisaged To-Be state in mind, portal owners should consider addressing front-end issues to meet user needs, improving the engagement mechanisms with data providers not only to deliver the content in appropriate formats, but also to share their domain knowledge and engaging with other portals to solve joint challenges.