LABORATÓRIO NACIONAL
DE ENGENHARIA CIVIL

DIGITAL COAST: A SCIENTIFIC PROPOSAL FOR IT-BASED RESEARCH IN COASTAL REGIONS FOR THE NEXT DECADE

Programa apresentado a concurso para a obtenção
do título de Habilitado para o Exercício de Funções
de Coordenação Científica,
na área científica de Hidráulica Marítima

LISBOA ● setembro de 2022

I & D DHA

ABSTRACT

This document corresponds to the Research Program and associated Post-Graduation Program elaborated by the author in the scope of the process of certification for the functions of coordination of scientific research, according to Decree-Law no. 124/99, of April 20th.

This Research Program identifies several research themes for the next decade related to the application of information technologies in coastal science and innovation of research. The selection of these themes is framed in the scientific national and international context, focused in particular in the activity of LNEC in this area through the research of the Information Technology in Water and Environment research group, led by the applicant, in collaboration with other divisions of the Hydraulics and Environment Department.

After a brief overview of the theme and presentation of the rationale for the development of this work, the national and global context for the Program is presented, wrapping up with the presentation of the research strategy for the Information Technology in Water and Environment research group. From this strategy, the two research areas of this Habilitation Program are identified and briefly described. The first area is the creation and development of reliable, cross-scale, multi-process, on-demand coastal forecast framework for oceans to hydrographic basin application, from hydrodynamics to biogeochemistry. The second area is the creation and development of intelligent, high-resolution, user-centered and inclusive coastal digital twins.

The two following chapters present the state-of-the-art in these two areas, the challenges to be overcome and the general roadmaps for the tools to be developed in the next decade to address the societal challenges in the coastal regions. The two Research Studies are presented next, organized along 19 projects. For each project, the applicant presents the rationale behind it, along with its goals, describes the methodologies for its implementation and the results to be generated. The resources necessary for its implementation along with the expected partnerships and adequate funding sources are also described.

Finally, the Post-Graduation Program is presented, providing multiple M.Sc. and Ph.D. education opportunities framed in the previous Research Program. A total of 9 Ph.D. and 6 M.Sc. proposals are presented.

RESUMO

Este documento corresponde ao Programa de Investigação e ao Programa de Pós-graduação elaborados pela autora nos termos estabelecidos no artigo 30º do Decreto-lei nº 124/99, de 20 de abril, com vista à prestação de provas de habilitação para o exercício de funções de coordenação de investigação científica.

O presente Programa de Investigação identifica várias temáticas de investigação na aplicação das tecnologias da informação à engenharia costeira para a próxima década. A seleção destas áreas foi enquadrada nos atuais desafios nacionais e internacionais nesta área científica e na atividade do LNEC no Departamento de Hidráulica e Ambiente através da investigação do Grupo de Tecnologias da Informação em Água e Ambiente em parceria com outros setores deste departamento.

Após uma breve introdução da temática e apresentação da sua importância, este Programa começa por apresentar o contexto nacional e internacional, terminando com a apresentação da estratégia para a próxima década do Grupo de Tecnologias da Informação em Água e Ambiente do LNEC, a qual identifica as duas áreas principais de investigação a seguir, as quais são objeto deste programa de habilitação. A primeira área é a previsão em tempo real de elevada precisão, confiável, multi-modelo, desde a bacia hidrográfica até ao mar, com a dimensão urbana, incluindo desde a hidrodinâmica até à biogeoquímica. A segunda área são os Gémeos Digitais Costeiros (*Coastal Digital Twins*), integrando inteligência e serviços digitais de alta resolução, centrados na interação com os utilizadores, desde os gestores costeiros ao público em geral.

Os dois capítulos seguintes identificam o conhecimento atual em cada área, os desafios a ultrapassar e as ferramentas a desenvolver para aplicação prática em resposta aos desafios societais na zona costeira da próxima década. Apresenta-se de seguida os dois Programas de Estudos, organizados em torno de 19 projetos. Para cada um destes projetos, estabelecem-se o respetivo enquadramento, os objetivos a atingir, a metodologia para a sua implementação, detalhada ao nível de tarefa, os resultados esperados e os recursos necessários para a sua implementação, terminando com as oportunidades de financiamento e de parceria.

Finalmente, é apresentado o Programa de Pós-graduação, que visa promover múltiplas oportunidades de desenvolvimento de teses de doutoramento e de mestrado com base na investigação proposta anteriormente. São assim propostas 9 dissertações de doutoramento e 6 teses de mestrado.

Keywords: Coastal Digital Twins, Hybrid forecast systems, Artificial intelligence, Multi-process, Cross-scale modeling, Data-based methodologies, Information quality and reliability, Early-warning systems, Collaboratories

Palavras-chave: Gémeos digitais costeiros, Sistema de previsão híbridos, Inteligência artificial, Modelação integrada multi-processo, Metodologias baseadas em dados, Qualidade e confiabilidade de dados, Sistemas de alerta, Colaboratórios

Table of Contents

List of Figures

List of Tables

Acronyms

AI - artificial intelligence

APA - Portuguese Agency for the Environment

AIR Centre - Atlantic International Research Centre

ANN - artificial neural networks

ATTRACT-DIH - ArTificial InTelligence and High-PeRformAnce CompuTing @ Portugal

CBEPS - Chesapeake Bay Ecological Prediction System

CDT - Coastal Digital Twins

CFS - Cyber-Physical Systems

CMOP - Center for Coastal Margin Observation & Prediction

DANUBIUS-RI - International Centre for Advanced Studies on River-Sea Systems

DGRM - Directorate-General for Natural Resources, Safety and Maritime Services

DHA – LNEC's Hydraulics and Environment Department

E2I – LNEC's Innovation and Research Strategy

ENGIZC - Integrated Coastal Zone Management National Strategy

ENPCP - National Strategy for a Preventive Civil Protection 2030

EOSC - European Open Science Cloud

GIS - Geographical Information Systems

GTI: Information Technology for Water and Environment Research Group of LNEC

HPC - High Performance Computing

IH – Instituto Hidrográfico

INCD - Infraestrutura para a Computação Distribuída

IT – Information Technology

NEC - Estuaries and Coastal Zones division of LNEC

NPE - Ports and Maritime structures division of LNEC

OGC - Open Geospatial Consortium

OSSEs - Observing System Simulation Experiments

POC - Coastal Edge Programs

POE - Estuaries plans

RNCA - Rede Nacional de Computação Avançada

SIARL - Information system for the administration of the resource “coast”

SIIN2A - Innovative Information Systems for Smart Water and Environmental applications

SVM - support vector machine models

UG - unstructured grids

VRE - Virtual Research Environments

WebGIS – internet-based geographical information systems

WIFF - Water Information Forecast Framework

WWTP – Wastewater treatment plants

Acknowledgements

This Research Program is dedicated to my wonderful sons David and Artur. Their love and companionship have always filled my life with happiness and joy. My sons' support along with that of my brother and my sister-in-law in the last months has been fundamental to have time and peace of mind for this work as my mother's health required more help than ever. The ever-present righteous and hard-working example of my mother and the memory of my long-deceased loving father gave me strength and motivation to overcome the difficulties and to be able to develop this work while maintaining my coordination role at the Information and Technology research group of LNEC and the coordination and participation in multiple projects for the vast majority of the last year.

Happiness is not achieved with human interaction only and I am the proud owner of four beautiful and mischievous pets, who fill my days with joy (and sometimes prevent me from sleeping at night). To Tofu, Penny, Simba and Fiona, I thank you for making my life brighter and happier.

The work proposed herein is built on top of many strongly collaborative projects that started over 10 years ago at the Estuaries and Coastal Zones division (NEC). The many fruitful discussions, creation of project proposals and innovative execution of projects with my colleagues André, Marta, Alberto and Paula are at the core of most of the ideas proposed herein. OPENCoastS' concept was born from a think-tank coffee break with Alberto and many of its current capacities were born from NEC's requirements and ideas. The many collaborations of colleagues from other institutions in Portugal and abroad has also enriched my knowledge and helped to raise many questions to be addressed at the core of coastal digital twins. I would like to highlight the provocative discussions more than a decade ago with António Baptista, my former Ph.D. advisor and longtime mentor, and the ground breaking work at his STC CMOP center of excellence in the USA.

However, this work would have never been possible without the collaboration of my colleagues from the Information Technology Division in Water and Environment (GTI), in particular Gonçalo, João, Ana, Joana, Pedro, Ricardo and Américo. They have transformed the coastal engineer that I was in 2010, into a "digital" coastal engineer, by questioning and aligning my every idea to an IT-context, allowing me to grow in creativity and collaborative work in applied IT and providing the grounds for melanging coastal engineering and IT concepts and tools into state-of-the-art research. For their support and dedication, my whole-hearted thanks.

I would like to acknowledge my many other colleagues and friends at LNEC and abroad that have kept me motivated and supported me with very kind words during these past months, in particular Elsa, Armando, Teresa Viseu, Luís David and Lina Nunes. To my former and present secretaries, Marisa and Alda, thank you for your help and patience as the bureaucracy became even more annoying in the past months and my patience for it grew thinner as the Program progressed.

I would also like to thank the present and the previous Boards of Directors of LNEC and the past and current Heads of the Hydraulic and Environment Department of LNEC for the support and conditions to develop this work.

Finally, I would like to acknowledge two projects that were at the heart of the OPENCoastS service development: H2020 EOSC-hub and H2020 EGI-ACE. The opportunities of research proposed herein were strongly inspired by the vision of open science, shared resources and collaborative research promoted in these projects.

Part A - Research Program

1. Introduction

1.1. Initial considerations

The digital dimension is nowadays present in all Water-related arenas, whether the concern has economy, safety (risk) or ecology contexts. Its role has been growing in importance in parallel with public awareness and public participation in water affairs, with a strong connection with economic growth and adaptation to climate change impacts.

In the blue economy context, Information technology (IT) concepts and tools have been gaining increasing importance. Major international efforts have been developed in the last decade to address ocean problems through collaborative IT-based research. However, much remains to be done at the estuarine and coastal dimension, given the complexity of coastal processes, their interactions with land and ocean inputs and local management context. This Research Program aims at addressing some of the current key research challenges for IT-based developments applied to coastal management and protection.

The present Research Program, together with the Post-Graduation and Education Program of Part B, are destined to be defended in public examination for access to the degree of “Habilitação para o exercício de funções de coordenação de investigação científica" in the scientific area of “Hidráulica Marítima” (Coastal and maritime hydraulics), following Decree Law nº 124/99, of April, 20th, which regulates the Statute of the Scientific Research Careers in Portugal.

It aims to contribute to the Hydraulics and Environment Department’s (DHA) strategy for IT-based research in the estuarine and coastal domains for the next decade. It is rooted in the collaborative activity of the Information Technology for Water and Environment Research Group (GTI), led by the applicant, with the Estuaries and Coastal Zones Division (NEC) for the last 10 years, and on her previous nearly 20 years of coastal modeling research developed at NEC, between 1990 and 2009.

The research proposed in this Program is both framed by LNEC’s 2014-2020 research strategy, currently in operation, through the applicant’s P2I project “SIIN2A - Innovative Information Systems for Smart Water and Environmental applications" and LNEC’s main research strategy topics for 2021-2027, namely Industry 4.0, Climate Change and Circular Economy. Through the concepts, methodologies and tools proposed herein, the applicant aims at promoting LNEC’s as a center of excellence on IT-based coastal research and consolidating and improving LNEC’s competences and capacities to address IT-related requests from the National water authority, coastal managers and other stakeholders in the estuarine and coastal zones. The Program also aims to address the several challenges and opportunities at national and international in the Digital-centered domains, such as the Digital Twins, Hybrid modeling, Artificial Intelligence applications and Data quality and reliability, and adapt them to the estuarine and coastal zones requirements and characteristics. Finally, the applicant also aims to

propose new IT-based methodologies and tools that can contribute to current challenges in coastal management, such as climate change impacts at short to long term scales.

The present Program aims at exploring Coastal and Estuarine Information in multiple perspectives, to build knowledge, provide confidence, explore interactions and predict events, sharing all through user-friendly, agile tools that seek user uptaking and adoption for both formal purposes and informal uses, supported by e-infrastructures that guarantee timely delivery of data and information products. Herein, “Information” is defined in a broad sense, including both data collected in the real world, from in-situ or remote sensing networks, and virtual “data” generated by models, either numerical or data-based ones.

The proposed work is organized and framed along GTI’s 4 research activity domains:

- Real time forecast frameworks for coastal phenomena, including numerical model and AI-based predictors
- Near real coastal observatories and WebGIS platforms, integrating data and model results in a seamless way through indicators targeting knowledge creation
- Intelligent systems for coastal risk and emergency management, including early-warning systems, providing accurate, timely warning tailored to the clients´ needs
- Multi-source coastal data management, including data reliability and open data sharing and providing cascading data integration procedures towards support of everyday, emergency or long term management actions.

The research strategy is presented below to address multiple challenges and problems in coastal regions, materialized through **2 Research Studies** and **19 research projects**, all detailed in goals, methodologies, resources and delivery products description. This research agenda is then organized as post-graduation opportunities, along **6 M.Sc. thesis** and **9 Ph.D. dissertation** opportunities.

1.2. Rationale

Scientifically-based coastal management is becoming increasingly important given the impact of climate change, the need to comply with European directives and the pressure of a growing population living near the coast, both from an economic and environmental perspectives, exploring its resources in a sustainable way, and from an emergency and risk context.

The fast evolution of information technology has prompted the development of multiple tools such as forecast systems, WebGIS, low cost in-situ monitoring networks and remote sensing data. This wealth of information has in turn prompted the development of multiple platforms to convey knowledge and support coastal management. However, an integrated approach that mimics the complexity of coastal dynamics and has the ability to accurately forecast it at the time scales required for everyday and emergency management is still to be done. The atmosphere and ocean community are now taking the first steps in digital representations of Earth denoted **Digital Twins.** The integration of these compartments with the inland water dimension (river to coast) is the next step towards achieving a fully-scientifically based coastal management supported by IT tools. In this Program, the applicant proposes an ambitious strategy to take full advantage of existing IT-based methodologies and concepts for

prediction and monitoring the coastal zones. It explores IT assets to propose a suite of tools for multi-purpose coastal management support that minimizes future maintenance and has a broad scope, accounting for uncertainty at all levels of the cascading modeling procedure and integrating data through artificial intelligence (AI) models with process-based predictions. Both short term (emergency, early-warning) and long term (risk, climate change impacts) processes at coasts and connected water domains (ocean, river, city) can be dealt with the concept of coastal Digital Twins and used to address city to water management issues, climate change impacts, multiple sources of risk (inundation, wave overtopping, contamination, etc..) and support to pre- and post- emergency actions.

Several European Directives with impact at the estuarine and coastal dimensions, such as Water Framework Directive (WFD), the Bathing Water Directive, the Floods Directive and the Marine Strategy Framework Directive, have promoted detailed monitoring programs of these areas, targeting the good status of the waters and the preservation of ecosystems and biodiversity. In Portugal, the application of these Directives implied a considerable monitoring effort (in particular the WFD) and produced a considerable number of reports and scientific publications. However, most data are scattered and often inaccessible, real time continuous monitoring is often out of operation/unavailable, automatic data quality procedures are not implemented and no usage of artificial intelligence methodologies are applied to extract knowledge in a continuous way. Open data repositories and web portals are limited and the combination of efforts to support multiple goals and directives through a single portal is not implemented, duplicating time and human resources to maintain in operation a multitude of portals which become outdated in content and technology every day. Furthermore, integration of other open data sources of relevance, such as Copernicus sentinel remote sensing data, or the exploitation of high resolution local cameras is not part of the existing coastal surveillance assets.

The concept of multi-purpose coastal observatories, developed recently at LNEC and FCUL (projects UBEST; COASTNET), based on the concepts and vision of the Center for Coastal Margin Observation & Prediction (CMOP, U.S.A.) and other international initiatives, are a possible avenue to address these challenges at national level. However, much work remains to be done regarding automatic data reliability and quality assessment and the creation of knowledge using AI-based methodologies. The present Program addresses these challenges by merging existing knowledge and tools with several tailored AI technologies and products to build trustworthy observatories, aiming at addressing multiple end user needs through a single one-stop-shop portal. This portal will manage the whole cycle of data and predictions automatically with quality assessment. These new observatories can be used to track continuously seasonal variability and climate changes impact, support long term goals and short term operations and address risk and emergency in a comprehensive and timely way.

The development of the Research Program is also relevant and timely for LNEC and Portugal as many initiatives are being promoted at national and international level towards taking advantage of IT to better know and manage our coasts. The development of the concept of web-based coastal services, targeting user creation of knowledge and products, such as OPENCoastS (Oliveira et al., 2020, 2021) or WORSICA (https://indico.egi.eu/event/5464/contributions/15666/attachments/14141/17811/WORSICA_EGI2021_

v1.pdf) along with the development of global initiatives for virtual representation of the Earth (DestinEarth) and oceans in particular (Digital Twins of the Earth) have set the stage for a user-centered digital vision for support of all coastal interventions and knowledge creation. Additionally, the continuous improvement in shared computational resources such as EOSC, Supercomputing and national infrastructures availability, such as the *Rede Nacional de Computação Avançada* (RNCA) in Portugal, has allowed for the creation of Virtual Research Environments (VRE) to support modeling and data analysis for cross-scale research, from the physics to the biogeochemistry, integrating uncertainty, data assimilation and inclusive web platforms.

At international level, the United Nations Sustainable Development Goals (UN SDGs) provide a framework for objectives to be reached by 2030 that are important in the estuarine and coastal zones. Goal 14 is the main target, but activities in this program are also framed in Goals 6, 11 and 13. Horizon Europe also addresses estuarine and coastal water in a hydrographic basin-to-ocean perspective in its "Healthy oceans, seas, coastal and inland waters" Mission. The EU Mission "Restore our Ocean and Waters" goals by 2030 are to achieve "cleaning marine and fresh waters, restoring degraded ecosystems and habitats, decarbonising the blue economy in order to sustainably harness the essential goods and services they provide". The approach is based on the vision to "address the ocean and waters as one and play a key role in achieving climate neutrality and restoring nature". Finally, the UN Decade of Ocean Science promotes several objectives related to coastal regions and is already promoting research through two programmes: CoastPredict, dedicated to facilitating the setup of forecast systems in coasts worldwide, and DITTO, centered on ocean digital twins. Addressing these challenges at local and global level requires the availability of 1) trustworthy information and IT tools for monitoring its implementation and indicator evolution, 2) digital twins to predict short to long term dynamics at the coasts, that are integrated in river to ocean vision and modeling capacity, and finally 3) the capacity to convey the relevant messages to the right end users, being them the national water authority, the civil protection agencies or the population at risk, through dedicated, tailored, easy-to-use IT portals rooted on the best information available .

Several research agendas in the Information Technology areas are also a motivation for the choice of the chosen scientific themes. The Digital Europe Programme covers several areas of interest, including the creation of Data Spaces for the Green Deal, the Digital Earth Initiative and the AI-on-demand platform. Behind the scenes, the EU data strategy provides the background for all aspects of data, from data lakes to FAIR data. All of these challenges and opportunities are addressed in this Research Program in an integrated, seamless way that are centered at coastal problems, but with the potential to evolve to the whole water cycle. Details on a selected list of the most relevant initiatives and their relevance for the present work are provided in the next chapters.

Finally, the choice of themes on this Research Program was also influenced by the applicant's professional career, as the creation of this proposal required a strong background on estuarine and coastal sciences and the knowledge of IT concepts, tools and capacities to connect and integrate coastal sciences achievements into tools to address management needs. After a nearly 20-year period of research at the Estuaries and Coastal Zones Division, with extensive work in multiple areas such as

hydrodynamics, non-cohesive sediment and inlet dynamics, and water quality research, the 10+ years at the leadership of the Information Technology Division and Research Group (GTI) has led to over 20 national and international research projects applied to coastal areas. These projects promoted the competences of DHA in coastal research abroad and led to the creation and implementation of many tools that are at the core of the present Research Program, such as the Water Information Forecast Framework (WIFF), the OPENCoastS e-service, the coastal observatories and AI-based data reliability and quality assessment framework. This work was achieved through the strong and very competent team of computer science researchers and IT experts at GTI and strong connections to other Divisions at DHA, which paved the way for the applicant proposal of the present strategy for Coastal hydroinformatics at LNEC. Finally, the knowledge gained and networking developed by the applicant at the European Open Science Cloud (EOSC) and CoastPredict initiatives has motivated the development of the present Research Program as a contribution towards coastal science.

1.3. Structure/Organization of the Program

This Research Program is organized as follows, after the present Introduction. Chapter 2 presents the National and International context and framing of the Research themes, including LNEC's dimension. Chapter 3 and 4 are dedicated to the State-of-the-art in the two main research topics: Forecast systems and Digital Twins. Chapter 5 presents the proposed Research Studies, along with its Research projects, followed by Chapter 6, which presents the Post-Graduation and Education Program. The document closes with the Final Considerations (Chapter 7) and the Reference list.

2. Research Program's national and international context

2.1. IT-related coastal research at LNEC

The Present Research Program and related Post Graduation Program are expected to be promoted and developed at DHA, through the collaboration of GTI and other divisions working in the estuarine and coastal zones area (NEC and NPE), as well as other partners from academia and industry.

Over the past 20 years, the need for Information Technology tools have been identified by all divisions of DHA to address multiple requests from end-users and to support and disseminate research in several scientific areas. The development of risk and emergency plans through Geographical Information Systems (GIS), the development of databases to contain water related information and some preliminary interfaces for sharing relevant information with several users are among the first developments, along with the IT support on the usage of multi-processor resources for numerical modeling. In the last decade, the experience of the applicant in promoting multidisciplinary projects at NEC has changed the focus of the IT division to embrace more complex tools, such as generic forecast platforms, artificial intelligence methodologies, the integration of multiple sources of monitoring sources with emphasis to remote sensing, and the development of modular, generic platforms for risk, emergency and general information sharing that are tailored in each project to address its requirements. The vast majority of these projects were applied to the estuarine and coastal zones, or to broad water compartments that also included these areas and its interfaces to freshwater domains or the bordering cities.

This new vision and strategy promoted several research lines, framed in the past E2I research plan, that include forecast, risk and early-warning systems, WebGIS platforms, geographical data specification and data management plans, and High Performance Computing (HPC) applications.

The present Research Plan aims at establishing the new strategy for GTI and IT-related research at DHA for the next decade, with a focus on maritime hydraulics themes. It is framed in the new, proposed LNEC research strategy for 2021-2027, with developments mainly integrated in line "Industry 4.0" which aims at producing, exploring and managing information and developing information systems for decision support. In particular, all the proposed projects in Chapter 5 are aligned with several challenges of this LNEC's Research Line:

- "developing and putting in operation information systems for decision support": all proposed projects address this concern. Research study 1 aims at creating an integrated, basin-to-ocean forecast framework for on-demand applications, which can be applied to multiple coastal management problems such as inundation, contamination, erosion, evaluating engineering interventions at the coast as well as all aspects related to adaptation and evolution due to climate change concerns. Research study 2 will propose Coastal digital twins as a tool for future short to long term management, through a combination of modeling and data tools for

integrated city and coastal management portal to linked management concerns: coastal city inundation and urban beach contamination. Finally, Research study 2 also illustrates the integration of all methodologies proposed into the building of a one-stop-shop information portal for coastal management through the proposal of os a coastal digital twin that address Portuguese management concerns that have been conveyed to the applicant in the last decade

- “Big Data and Data Analytics' ': this Research Program deals with information from multiple sources, its processing, validation, quality assessment and management at all levels. Therefore, data analytics (considered here in the broad sense of information analytics) are present in all proposed Research studies. In particular, they integrate directly data operations, and the development/adaptation of artificial intelligence methodologies for data creation and processing. Both Research study 1 and 2 have to address data concerns in all steps of the data management cycle, from data creation, data validation and data sharing. Big data in the field of coastal science is frequently associated with remote sensing, herein present from both satellite and local camera sources. In the proposed research studies, several methodologies are used to handle information at the scale of Big data, either to address remote sensing data or to process and integrate it with other information sources such as other monitoring sources (e.g. low cost sensors) and forecast system outputs. Processing of information to generate products to address directly society needs (e.g through early-warning issuing) or to make data available for information systems feeding is present on all projects.

The present work also contributes with methodologies and tools that serve societal needs along the research Lines “Circular Economy”, which aims at promoting sustainability in exploring natural and asset resources, closing cycles and avoiding waste, and “Climate change”, which aims at exploring and reducing uncertainty, managing risks and mitigation and adapting to climate change. Research study 1 promotes the knowledge in advance of our coastal system dynamics, permitting decisions to take into account the preservation of coastal resources and anticipate risks and emergencies, supporting adequate early warning. Research study 2 explores multiple sources of data and proposes tools based on artificial intelligence that provide data-based knowledge for data quality and confidence. Achieving a circular economy cannot be met without reliable information. Likewise, quality data and data forecasts are mandatory for climate change adaptation and mitigation. The Coastal Digital Twins from Research Study 2 will provide the tools to addresses the societal concerns in both research lines. It also illustrates how these tools can be merged and integrated into a multiple stakeholder concern one-stop-shop tool to address these and other concerns tailored to the Portuguese coast and the needs of coastal managers, both in terms of protecting coastal assets and dealing with climate change concerns.

Finally, the participation of LNEC in two digital networks through the applicant and her research group, the Infraestrutura para a Computação Distribuída (INCD) and the European Digital Innovation Hub for ArTificial InTelligence and High-PeRformAnce CompuTing @ Portugal (ATTRACT-DIH) and the opportunities for collaborative research for society therein, contributes to the financing of the present Research Program and to promote the post graduate Program among students of the several participating universities. LNEC’s participation in INCD over the last three years has also funded the initial developments of several tools, such as OPENCoastS and Worsica services and UBEST coastal

observatories, that are at the heart of the present Research Project. This research has also promoted the participation in several e-infrastructures H2020 projects, such as EOSC-hub, EOSC-Synergy and EGI-ACE, which have contributed to the setting of the vision towards HPC and GPU computing in coastal modeling, forecasting and remote sensing data processing proposed herein. The combination of knowledge, tools and products in this Research Program are expected to be used in ATTRACT-DIH to address both societal needs at our public entities and promote the economy through handing over of research products to SMEs.

2.2. National estuarine and coastal management programs and plans and other relevant framing initiatives

2.2.1 Overview

Estuaries and coastal zones' management have to take into account the implementation of several national plans and programs besides the applications of a number of European Directives, with many entities being involved in the management decisions or having an impact on the coast.

The Portuguese Agency for the Environment (APA), as the National Water Authority, holds the major responsibilities in managing these regions and is responsible for several plans and programs besides being the entity in charge of the application of several European directives in Portugal. Through its several Programs and Plans, APA defines directions, guidance and criteria aiming at the good quality of the water bodies and towards harmonizing water and bordering territories management, targeting the preservation of natural resources and assets. The combination of management initiatives provides the adequate national strategy for management and usage of the water resources and its linkage with land management and the preservation of the environment. The planning is centered on the hydrographic regions, supported by the Water law, and promoting harmonized special programs for the coast and the estuaries.

The Directorate-General for Natural Resources, Safety and Maritime Services is responsible for the remaining Directives more focused on the sea, fisheries, maritime protected areas and port management. Finally, there are other initiatives that have an impact on the coast and also contribute towards its management. The Civil Protection Authority is one of the entities that contributes to coastal management through initiatives dedicated to risk and emergency.

Several other institutions and initiatives were identified below as contributing towards estuarine and coastal research management and the general themes of this Research Program. Below, a summary of the most relevant plans, programs and initiatives is presented, along with the implications and relationships with the present Research Program.

2.2.2 Integrated Coastal Zone Management National Strategy - ENGIZC (*Estratégia Nacional para a Gestão Integrada das Zonas Costeiras*)

The ENGIZC aims at achieving an economically developed, sustainable coastal region based on a systemic strategy that values its resources and cultural assets, supported by knowledge and managed through cooperative actions bringing together institutions, politics and tools for a broad participation of all involved actors.

Its thematic goals are:

i) Preserve and value coastal resources and the natural, cultural and landscape assets;

ii) Anticipate, prevent and manage risk events and situations that have an impact on the environmental, economic or social nature;

iii) Promote sustainable activities that have a considerable economic impact and contribute towards valuation of coastal resources;

iv) Improve scientific knowledge on coastal systems and ecosystems.

The strategy was approved in 2009 and its vision aims at a 20-year period.

The present Research project is fully aligned with the thematic objectives of this strategy and can contribute in a relevant way with methodologies and tools that can enhance the knowledge and the management of the coast and its evolution in a climate change context, promoting solutions for protection and adaptation in a scientifically, process-based way.

Focusing on specific topics of this strategy in the context of the present research Program, the following items can benefit from this work's implementation:

- identification of the lack of a system that integrates all information on the coast and makes it available for decision-makers, planners and general managers, economic agents and society in general - while the COSMOS and the SIARL platforms from APA were developed to address this concern among other uses, neither have a complete repository of all coastal data (remote sensing for instance is missing), neither the data is provided with a quality classification. Therefore, both the data reliability and quality control and the digital twin approach proposed in this Research Program could enhance the existing information channels and materialize the full vision of the ENGIZC through the observatories proposed in project 5.
- Initiative 18 - to develop a national monitoring system of the coastal system, biotic communities and environmental quality - this target is today scattered through many portals at APA and other institutions (IH, IPMA, Coastnet,...), making the goal of analyzing and cross-validating data a huge effort. The integration of multiple sources of information from Research Study 2 along with its observatories can contribute towards a quality control system, that integrates data from multiple providers, quality assesses them automatically and make it available through indicators target at distinct management goals.

- Initiative 20 - develop an information program targeting the general public knowledge on the coast, including the development of an interactive Web plataform - this goal is partly addressed by the COSMOS program although the navegation on the portal can require technical knowledge. The coastal collaboratories proposed in Research Study 2 can include a dimension for public information, taking advantage of the flow of validated data behind it.

Link: https://files.dre.pt/1s/2009/09/17400/0605606088.pdf

2.2.3 Coastal Edge Programs - POC (*Planos da Orla Costeira*)

The POC are implementations of the National Soil, Urban and Land Management Law (Law 104/2014, Série I of 2014-05-30) for the coast and aim at the safekeeping and valorization of the coastal edge, through actions that establish which interventions are allowed in these regions. POC is one of Special Plans of Law 104/2014 and its management actions rule over local and regional planning programs. Mainland Portugal is divided into six regions and a POC is issued for each region. These plans have a period of operation and are updated after that period.

While these programs are static in time unlike most of the work proposed herein, they rely on the good quality of the coastal information for the setup of the coastal planning measures and the monitoring of the implementation of the measures proposed to address for instance erosion risk. Therefore, future editions of the POC may be able to take advantage of the data processing workflows proposed herein and of the digital twins as a scientific backup for informed decision making.

Link: https://apambiente.pt/agua/planos-e-programas-da-orla-costeira

2.2.4 Estuaries plans (POE)

The Estuaries Plans (POE) were defined in Decret-Law n.º 129/2008 (21 of July) and also in the Special Plan of Law 104/2014. They are applicable to Estuaries, incluíndo transitional waters, their margins and beds, and also over the estuarine border, which is the 500 m wide, measured from the margin, land protection area. These plans aim to protect estuarine waters, beds and margins and their ecosystems, through an integrated management that values these regions from an environmental, economic, social and cultural dimensions.

Up to now, only the Plan for the Tagus estuary was developed (2009). The requirements for the Douro estuary have already been published but have not been executed. The original goals included also the Mondego and the Vouga estuaries.

The Tagus estuary POE was presented in a dedicated publication identified below, encompassing the several themes relevant for the systems management and planning, from the physical processes to the ecological dimensions. In the context of the present Research Program, it is important to highlight the recognition, more than 10 years ago, of the importance of the IT-based tools as means to convey information between decision makers and to the general public. The dynamic and interactive tools to support decision making and public information identified in the POE can now be implemented through the applicant's proposals, merging data from multiple sources and high performance modeling from the

physical to the biogeochemistry processes. In particular, the new collaboratories proposed in Research Study 2, which are built from the concept developed in project UBEST and applied to the Tagus estuary, provide a unique decision support tool that address most areas outlined in this POE.

Links: https://apambiente.pt/agua/programas-de-estuarios

(Tagus estuary): https://apambiente.pt/sites/default/files/_SNIAMB_A_APA/Publicacoes/Tagides/tagidespoe.pdf

2.2.5 National Strategy for the Sea 2021-2030

The Portuguese National Strategy for the Sea for the period 2021-2030 aims at promoting a healthy ocean and coasts zones, the well being of the Portuguese people and place Portugal as a lider in ocean governance, supported by the best scientific knowledge. It organized around 10 strategic goals, dealing with the multiple challenges from governance to education, including also the digital dimension and the response to climate change and associated risks.

The present Research Program is well aligned with this strategy and can contribute to several of these goals:

- OE10, in the areas of rescue at sea prediction and the development of platforms for data sharing and tools for remote and in-situ data integration;
- OE9, in the areas of information systems and AI tools for ocean and coastal engineering;
- OE1, in the areas of coastal risk and emergency tools.

All of the tools proposed herein can also contribute to the ocean literacy of the public (OE8) and to the development of new research and innovation in particular in the areas of remote sensing platforms (OE7).

Link:

https://www.portugal.gov.pt/download-ficheiros/ficheiro.aspx?v=%3d%3dBQAAAB%2bLCAAAAAAABAAzNLQwsQQAODaj3AUAAAA%3d

2.2.6 National Strategy for a Preventive Civil Protection 2030 (*ENPCP - Estratégia Nacional para uma Proteção Civil Preventiva 2030*)

The National Strategy for a Preventive Civil Protection is framed by three risk management global strategies, adopted in 2015: The Sendai framework for disaster risk, the Paris agreement on climate change, and the UN sustainable development goals. It is based on 5 strategic goals that promote over 100 measures for its implementation at local and national level. This Research Program is aligned with many of these strategic goals:

- Goal 2, to improve risk knowledge, has produced a web portal that shares information from ANPC projects on risks (http://www.pnrrc.pt/index.php/geovisualizador/). In the cope of coastal zones risk, the maps of coastal inundation and overtopping risk are of particular importance and can be added as a webservice in the forecast framework platform of Research Study 1 or

in the Digital Twins of Research Study 2, as part of a customization for risk prevention and response;

- Goal 3, on risk reduction, will generate updated cartography on inundation risk including riverine and coastal areas. This cartography, like the layers on the previous goal, can also be added to some of the products of this Research Program and integrate long term risk analysis with short term hazard evaluation for emergency purposes.
- Goal 4, on improvement of preparedness for an effective response under risk events, has produced the National Risk monitoring and communication, Alert and Population Early-warning system (Sistema Nacional de Monitorização e Comunicação de Risco, de Alerta Especial e de Aviso à População - https://files.dre.pt/1s/2019/01/00800/0010500108.pdf). While LNEC has no formal responsibility in alerting ANPC on coastal risks, channels for informal communication of information generated by LNEC's monitoring networks and model prediction has already occured in the past in the scope of several projects on inundation risk (MOSAIC and MOLINES projects are examples). The implementation of the several tools proposed herein and their customization for risk management can improve in the future the quality of the information shared with ANPC either by adding new sources of information to the traditionally used (e.g. local cameras and their images automatic processing for inundation detection) or by providing a quality assessment for data supporting an alert emission. Furthermore, the forecast framework of Research Study 1 and their integrated approach on neighbor water compartments can improve current capacity to predict hazardous events.

Link: https://files.dre.pt/1s/2021/08/15500/0013300156.pdf

2.2.7 Research and Innovation Climate change thematic agenda

This thematic agenda is developed in cooperation between FCT and the Portuguese Environmental Agency (APA), framed by the recommendations of the National Strategy for Climate Change Adaptation https://files.dre.pt/1s/2015/07/14700/0511405168.pdf), which identifies the sea level rise, the changes in the waves climate and in the meteorological surge, and the modifications in temperature and precipitation as the main consequences of climate change in the coast. These changes also affect sediment balance, generating coastal erosion, change the frequency and magnitude of coastal inundations, and also alter the water quality in estuaries, coastal lagoons and coastal aquifers.

This agenda aims at promoting the necessary developments to achieve a carbon neutral and resilient to climate change society, through proposals of adaptation and mitigation measures and initiatives to reduce and manage risks. It is supported by the best scientific knowledge and proposes a strategic vision for research and innovation to be implemented until 2030.

This agenda has a strong link with coastal and estuarine initiatives given the extent of the Portuguese coasts and its vulnerability to multiple risks. Here the focus will be on the aspects of this agenda that influence or are related with coastal environments.

The Agenda identifies 4 themes on research and 2 on innovation. Among the first ones, the "Impacts, vulnerabilities and risks in systems and sectors" subtheme deals with the understanding on how climate

change will affect natural and anthropogenic systems as well as their interactions. The natural systems in this subtheme include the estuarine and coastal regions. One of the two innovation sub themes - Technological innovation of products and services - promotes in particular the development and application of assets towards risk and disaster reduction. It proposes the development of innovative early warning systems, new sensing methods, use of remote sensing data and real time information and artificial intelligence methods. Both subjects are well aligned with the goals of the present Research Project, in particular in the development of tools for risk reduction and anticipation of hazardous events, such as the forecast framework for integrated prediction (Research study 1) and the digital twins that can be used not only for short term emergency but also to assess climate change impacts and measure implementation evaluation (Research study 2).

Link: https://www.fct.pt/agendastematicas/docs/agenda_alteracoes_climaticas_pre_finalizada.pdf

2.2.8 Research and Innovation Thematic Agenda for Research and Innovation: Sea

FCT's thematic agenda for research and innovation for the Sea is expected to become the main source of information to inspire and support decision processes in Sea's research and innovation. It identified the following four grand areas of intervention:

(1) Integrated knowledge of the sea;

(2) Sustainable and informed exploitation of the sea resources

(3) Evaluation of natural and anthropogenic risks, of climate change impacts and the development and implementation of measure for their mitigation; and

(4) Knowledge of the ocean's past and promotion of a participative relationship of society.

These areas are then mapped in the several themes and challenges, organized along the labels of "Research and Technology" and "Innovation and Technology". To each label, the authors identified several key issues for the next 10 years as well as the major limiting factors to the development of these key issues. Finally, three priority programs were suggested, whose implementation is expected to occur in the scope of the AirCentre (described in the next section).

In the Agenda´s themes for "Research and Technology", several issues are identified that have a strong connection with the present Research Program:

- Integrated knowledge of the sea; Topic "Observation and Modeling"
    - integration of monitoring networks data into forecast models and creation of products for the general public - This goal is fully aligned with the proposal of Digital Twins of Researc Study 2, and can further improved if data reliability and data quality procedures proposed herein are integrated in the workflow.
    - develop cascading modeling from the regional dimension to the coast - this modeling effort has already been implemented in many national and European institutions. In the present Research Program, we take this effort one step further and propose a

framework for prediction that not only integrates global to local scales, but also includes uncertainty in the predictions and integrates the freshwater and the coastal cities water compartments

- Global changes, natural and anthropogenic risks; Topic “Extreme events and natural risks”: this topic aims at developing local and regional early-warning multi-risk systems (storms, coastal erosion and inundation, algae blooms). Several of the projects under both Research studies aim at developing all the infrastructure and information generation and processing necessary to build early-warning systems. The Digital twins integrate all processes to generate high accuracy predictions integrated with data-based models and the quality assessment of data for reliability. Finally, the coastal observatories provide the capacity to access all information tailored to the specific users (water authorities, civil protection agencies or the general public. Therefore, the present Research Program can contribute decisively to this topic and is fully aligned with its goals. The same tools can be applied for the modeling and prediction proposed in Topic “Anthropogenic impacts”.

Other topics, related with the improvement of the ocean dynamics are also linked to the research proposed herein in the sense that the proposed outcomes and tools can contribute to a better monitoring, prediction and understanding of the coastal zones dynamics.

In the agenda’s themes on “Innovation and Technology”, other issues are identified that also have a strong connection with the present Research Program:

- Monitoring technologies and integrated study of the sea; Topics “Operation Modeling” & “IT-based platforms for data management”- this topic aims at developing an operational forecast infrastructure for the sea to coastal regions, together with the necessary data infrastructure and viewer, that can be used by coastal managers and all players at the coast. While some of this effort was already achieved for the physical processes by several institutions in Portugal in the last years (e.g. IH’s regional operational forecast of circulation and waves, LNEC’s water information forecast framework for coastal and estuarine circulation and its on-demand service OPENCoastS, IPMA’s regional wave and circulation forecasts,…), a lot of effort is still necessary to expand to water quality and ecology and to make the results available to the end-users in a tailored way. Herein, we propose to expand the WIFF framework and the OPENCoastS platform to address this challenge, integrating data-based models and quality-certified data streams into digital twins and coastal observatories tailored for the users.

Link: https://www.fct.pt/agendastematicas/docs/Agenda_Mar_Versao_Finalizacao.pdf

2.2.9 Atlantic International Research Centre (AIR Centre)

Portugal created an international collaboration organization devoted to the Atlantic area that deals with ocean challenges integrated with climate and space. Technological innovations and advances in data science are at the core of the Air Centre initiatives and thus are closely linked with the present Research Proposal.

The challenges outlined in the present work are well aligned with four out of the five thematic missions of this initiative:

- Clean and Productive Bays and Estuaries - identifying the need for comprehensive monitoring an evaluation of impacts on ecosystem health
- Resilience to Coastal Natural Hazards - identifying the need to build reliable early warning systems
- Improved Management of Marine and Coastal Resources - identifying the need to maintain the health of sea and coastal ecosystems
- Improved Environmental and Maritime Monitoring - identifying the need to better observe and understand the ocean

The present Research Program aims at developing several tools that can contribute towards AIR Centre challenges, either through data reliability tools that provide data quality assessment on data through AI-based methodologies, integration of multiple sources of data into virtual representation of the coastal Earth system or through coastal observatories that can import and export information to the ocean dimension. By focusing on nearshore and estuarine domains, the tools in this Research Program can complement the work developed in AIR Centre, in particular through its projects.

Indeed, many of the current projects of this organization are linked with the research projects proposed herein. A few examples are outlined below:

- AIR_DATANET - The AIR Centre Data Intelligence Network- The resources to be produced in this project both at infrastructure and data repository level can contribute to the implementation of the present Research Program. As FAIR principles will apply to this project, it will be able to share and upload information from both the applications of the Digital Twin project as well as the Observatories project. In turn, the work to be developed in Research Study 2 for data reliability and quality assessment can be a valuable asset to increase the value of the information in AIR_DATANET's data repository (https://www.aircentre.org/projects/air_datanet-the-air-centre-data-intelligence-network/)
- NEXTOCEAN – Next Generation of Fishing and Aquaculture Services - This project takes advantage of remote sensing data and the GEOSS catalogue to build services directed to society needs in the area of fisheries and aquaculture, similarly to the main concept behind part of Research Study 2 and the past developments in the OPENCoastS and WORSICA services.
- ATLANTIC CITIES AND PORTS (ARIA3) - this project aims at developing customised Earth Observation-based information services to support decision making. The concept is aligned with the proposed Digital Twins project proposed herein in the sense of integrating remote and in situ data towards the development of services for the society. ARIA3 is devoted to nearshore and coastal cities challenges. Unlike ARIA3 the Digital Twin project also integrates prediction tools for better support anticipation of hazardous events at several time and spatial scales and promote mitigation actions.

Link: https://www.aircentre.org/

2.3. National data and forecast portals and repositories

2.3.1 Overview

Over the past decade many information hubs were created to address the implementation of European directives and other formal responsibilities, to support public institutions mission's implementation or to explore the open or consortium-restricted wealth of remote and in-situ data. Furthermore, the growth and availability of computational resources has increased the capacity to predict estuarine, coastal and ocean dynamics, promoting the outburst of publicly available short term forecasts at local, regional and ocean-wide scales. In Portugal, several platforms were created that are related to the present Research Program goals. Some of them outlived the end of their funding projects, while others are simply out of operation or without any data income for a long time. Herein, the applicant selected the most relevant ones, including those that could benefit from the Program's outcomes and those that can be sources of information to be integrated in the products proposed in the projects.

2.3.2 Coastal Monitoring Programme of Continental Portugal - COSMO portal

The COSMO monitoring program collects and processes information on the evolution of beaches, dunes, nearshore seabed and sea cliffs along the Continental Portuguese coastline. It aims at supporting informed coastal management and planning, in particular to understand the evolution of both sandy an cliffed coasts and the impacts of climate change an engineering works on them; and to support short, medium and long term planning, by anticipating the need for new/rehabilitated coastal defenses/protection mechanisms and to optimize risk mitigation strategies.

The portal information is open for visualization and download through a very well documented workflow and is properly meta-informed to allow other users to take adequate advantage of it (Figures 2.1 to 2.2). The concept behind COSMO is fully aligned with the present Research Program goals. First, COSMO's data can be a valuable resource for the proposed portals in both Research Studies and the data fusion and data quality methodologies can be added to the portal with clear advantage for coastal management activities at APA. The concept of Digital Twins and its management portal can be an avenue to enhance COSMO's portal by including the virtual dimension (through AI data-based or numerical model predictions).

The current portal is however fully dedicated to APA's monitoring efforts and does not allow for on-the-fly upload of data from other providers, using for instance web services, so an effort will be required for optimal integration of data.

Figure 2.1 - Sample data information and download page

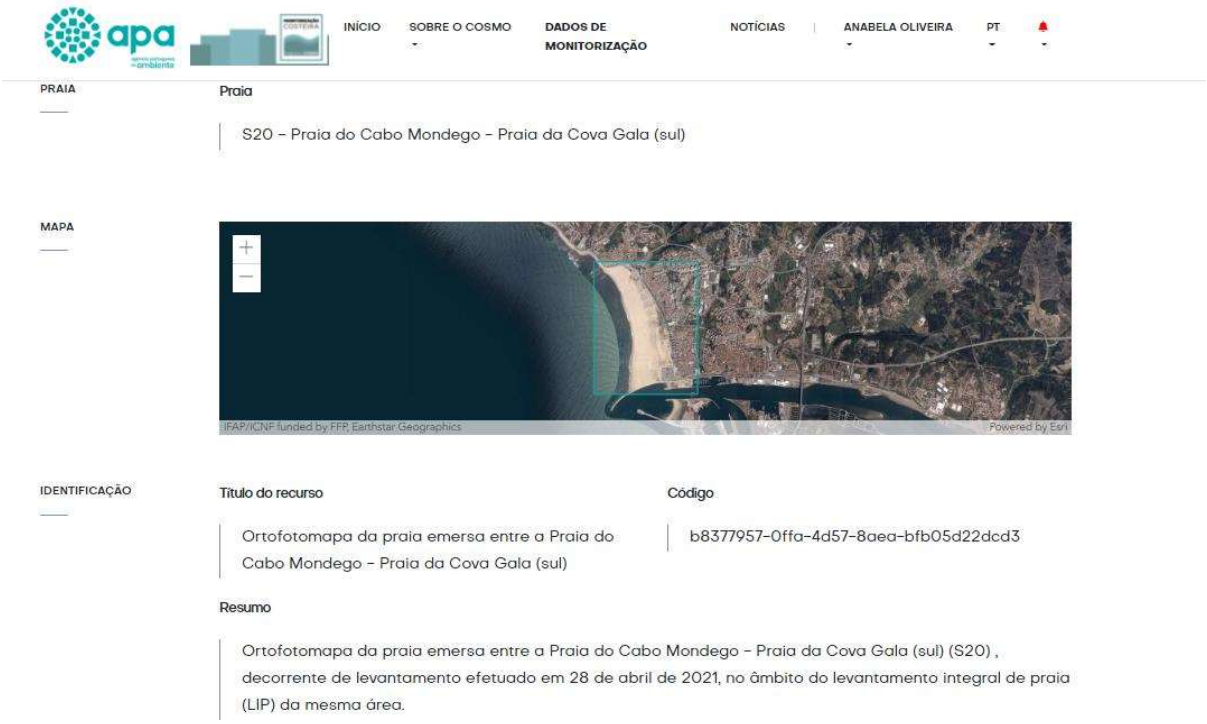


Figure 2.2 – Metadata for a specific dataset

Link: https://cosmo.apambiente.pt/

2.3.3 Information system for the administration of the resource "coast" (*Sistema de Administração do Recurso Litoral - SIARL*)

SIARL is a geoportal that aims at hosting the most relevant information on the portuguese coast to support political and technical decision at local, regional and national level. It takes advantage of information produced at other institutions and provides a mean for the data to be visualized and

combined in a geoportal. Figure 2.3 illustrates the example of the information that can be searched for and visualized in SIARL. Concerns on metainformation and compliance with the Inspire directive are at the core of this portal's conception and implementation.

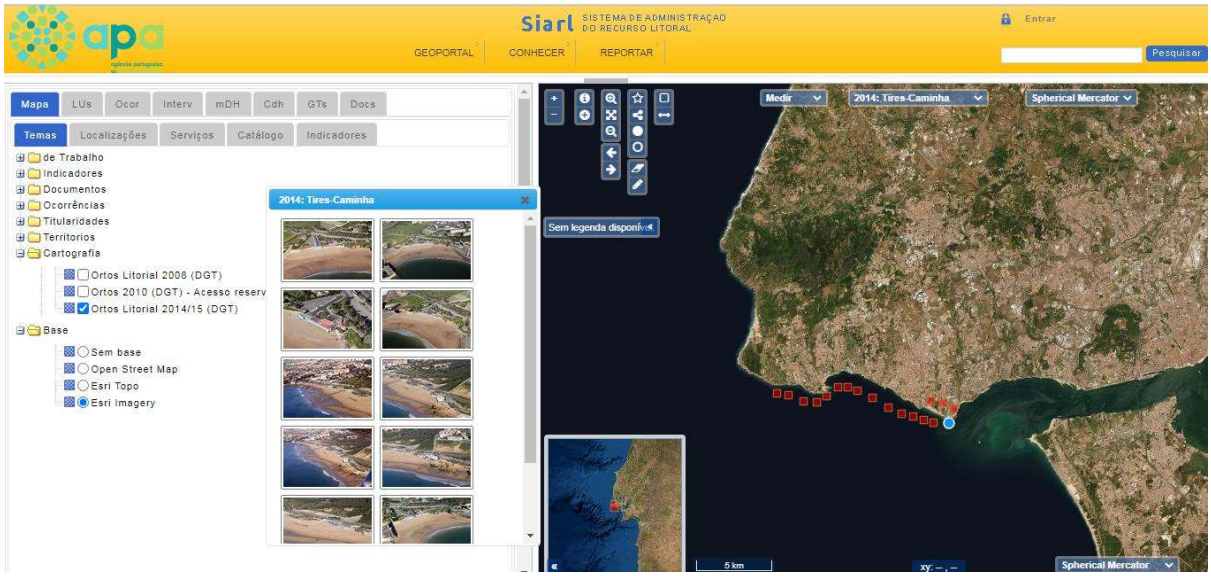


Figure 2.3 – Sample information provided by SIARL

LNEC, through the applicant and several colleagues from NEC, participated in several meetings with the SIARL main promoter at APA, towards the integration of LNEC's products (forecast and early-warning products and near-real time monitoring datasets) in SIARL. While SIARL's implementation is still far from the initial ambitious goals of its promoter, its geographical information design (where information is linked through web services) allows for easy integration with the outcomes of the present Research program. Moreover, the work developed by the applicant and her team on the OPENCoastS service and the proposed enhancements proposed herein, can contribute towards the scientifically-based coastal management vision of SIARL.

Link: https://siarldev.apambiente.pt/destaques.aspx

2.3.4 Geoportal for the reference status of the portuguese seas (*Geoportal 'Situação de Referência do Mar Português*)

The Directorate-General for Natural Resources, Safety and Maritime Services has the responsibility for implementing the Common Fisheries Policy through the management of natural resources, the fleet and the coordination of activity monitoring. The DGRM also has responsibilities in the maritime-port sector, and is responsible for most maritime safety and security state responsibilities and the certification of vessels and seafarers. It also assumes responsibilities for monitoring of coastal navigation as well as for managing maritime spatial planning, including carrying out port security works in fishing ports. DGRM is also responsible for proposing the creation of protected marine areas, in conjunction with the national authority for the conservation of nature and biodiversity. DGRM is also

responsible for ensuring the management of MPAs of national interest and collaborating in the management of those of a regional or local scope.

In order to support the actions regarding many of these responsibilities and to contribute towards sharing data among institutions with responsibilities at the sea and coast, DGRM created a Web geoportal and associated GIS infrastructure that provides the current status of the Portuguese sea. This Web portal integrates and makes available data from several providers and web map services that allows for search and visualization of spatial data. The applicant coordinated the LNEC's projet to generate maximum wave characteristics in the Mainland Portuguese coast and to provide them to DGRM for visualization in the portal (Figure 2.4). The vision of this portal is aligned with the goals of the present Research Program and the outputs of the proposed projects can contribute to the enrichment of this portal and to the sharing of relevant information for DGRM for their management responsabilities. The capacity of the portal to upload data through web services (Figure 2.4) greatly facilitates this exchange.

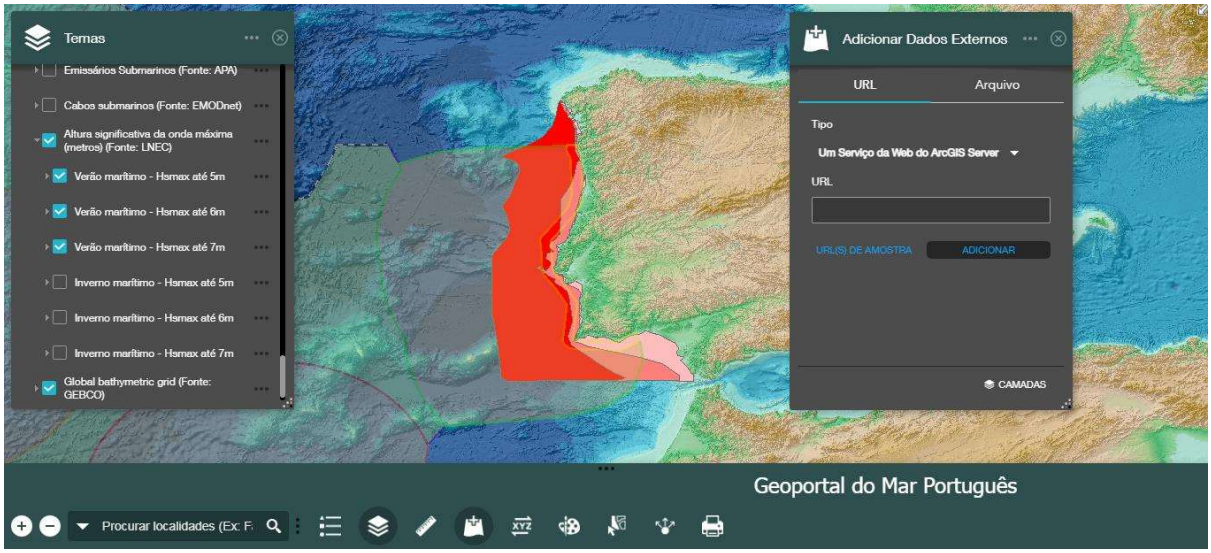
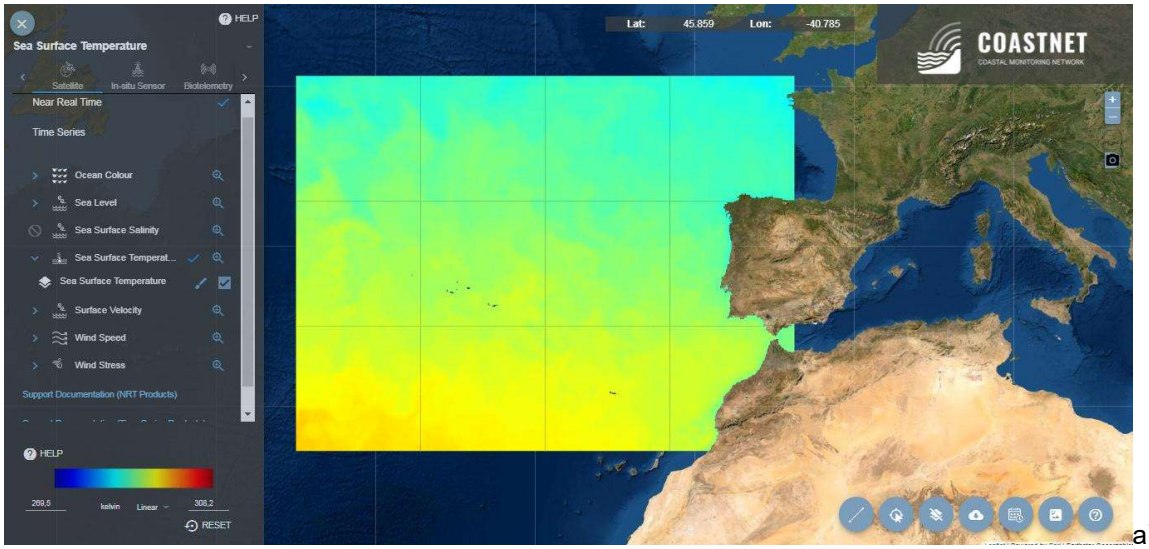


Figure 2.4 – Sample results from LNEC's wave analysis in DGRM's portal

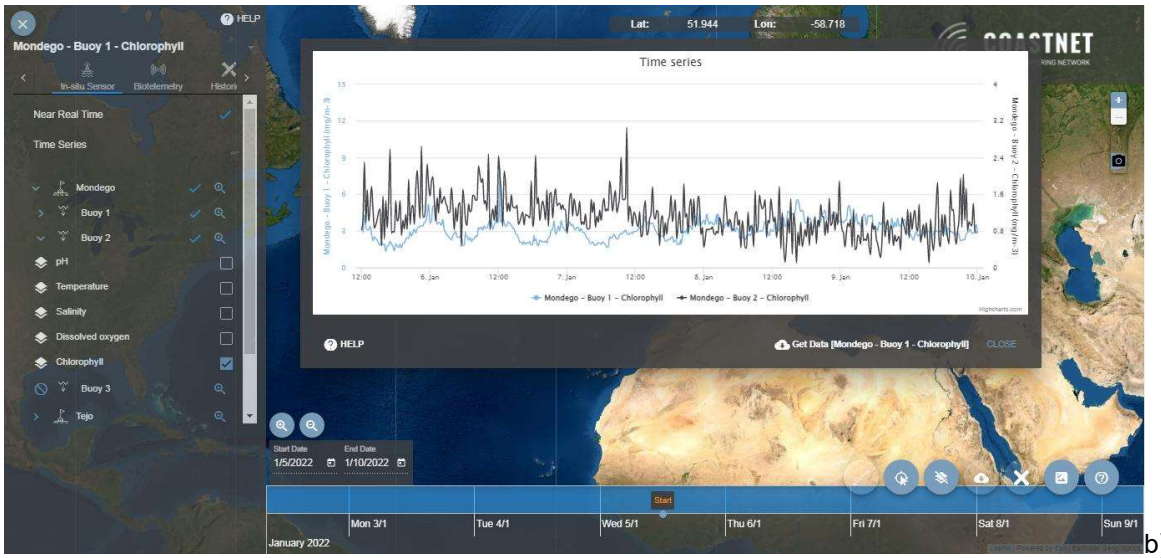
Link:
https://webgis.dgrm.mm.gov.pt/portal/apps/webappviewer/index.html?id=df8accb510bc4f33963d9b03bf3674b8

2.3.5 Portuguese Coastal Monitoring Network (CoastNet)

CoastNet is a research infrastructure that aims at providing near-real time remote and in-situ monitoring of important ecosystems, as well as historical data. It is devoted to physical, chemical and biological monitoring and provides free access to raw and processed data through a geoportal. A few reanalysis products are also available in time series repositories.



a)



b)

Figure 2.5 – Sample data requested at the CoastNet portal: a) remote sensing, b) in-situ data

While mostly focused on monitoring data (Figure 2.5), CoastNet shares the vision of Collaborative coastal observatories with LNEC's UBEST Coastal Observatories, and therefore can easily be integrated with the research proposed in Chapter 5 towards a national collaborative coastal observatories portal that address physical, chemical and biological processes in an integrated way towards policy support, European directives compliance and adaptation and mitigation to climate change impacts, besides being a one-stop-shop open repository to scientific research development. The major differences between CoastNet and LNEC's UBEST observatories lay on the strong presence of modeling and model-data processed products in the latter, following the groundbreaking concept of the Columbia River collaborative observatory, and the ownership of most in-situ data spots in the former. Furthermore, the generic concept of coastal observatories proposed by both portals is the key for their application at national level, leveraging the knowledge created by the pilots of Mondego, Tagus and Mira in CoastNet and the pilots of Tagus and Ria Formosa in UBEST. The present Research Program evolves this concept by bringing in the "on-demand" concept of LNEC's OPENCoastS service

to envision a portal where any coastal user can build their own observatory by on-the-fly integration of both monitoring sources (such as Copernicus, local cameras or other remote devices and open data, standard-based in-situ sensors) and tailored predictions, taking advantage of the WorSiCa service for remote sensing and OPENCoastS for on-demand forecasting.

Link: https://coastnet.pt/

2.3.6 Hidrográfico+

Hidrographic+ is a Geoportal created by the Portuguese Hydrographic Institute (IH) to organize, systematize and make available data collected by this institution, promoting data transfer to other institutions and the general public. The portal allows for search and selection of several types of data from bathymetry to the navigation channels or the location of the tidal data network, and a few modeling products (Figure 2.6). Downloading of data is restricted to authorized users, but visualization and comparison with external data is possible through WMS. The shared data is fully metainformed, facilitating the data interoperability and reuse.

The vision behind this portal is aligned with the proposed Research Program in the sense of data sharing and contribution of data for the development of Digital Twins. The use of WMS can easily contribute towards the work proposed in Research Study 2, while the bathymetric data is useful across the whole program as well as regional model predictions. Likewise, the research proposed herein can contribute towards enrichment of IH's data repository and portal, either on a case by case with manual upload or through a strong connection between tools in the future. Finally, like many data portals, IH's Hidrographic+ has no data quality assessment and therefore the outputs of this analysis could benefit the service provided by this portal by adding a reliability and data quality layer through its integration in the portal.

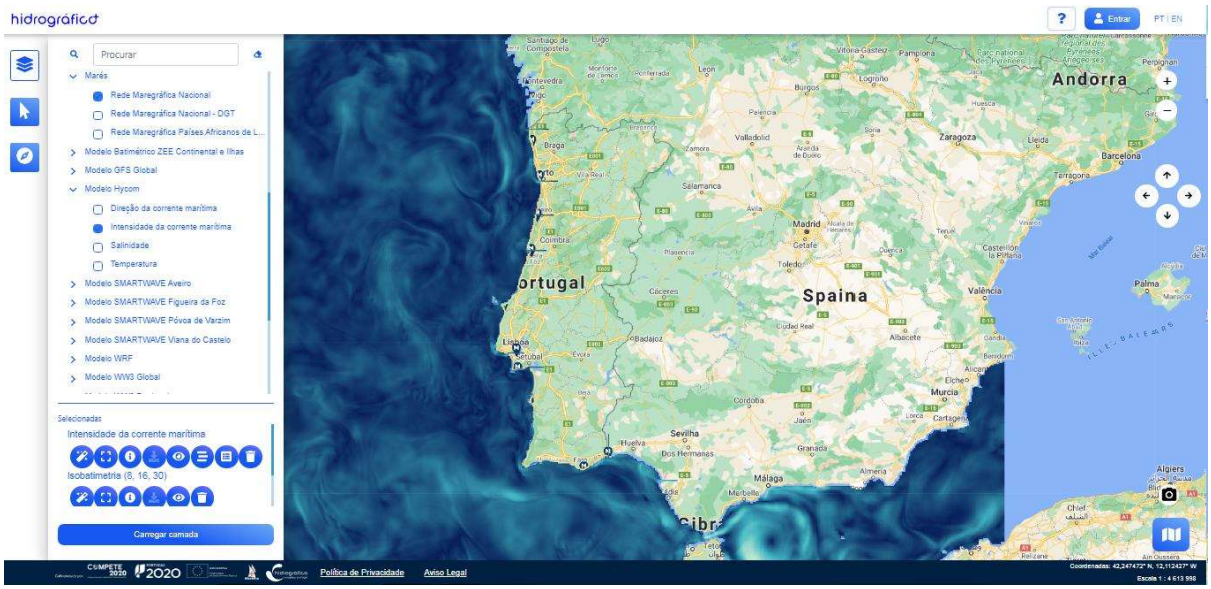


Figure 2.6 – The Hidrographic+ portal: sample visualization with HYCOM model outputs, nearshore isobaths and the tidal sensor network

Link: https://geomar.hidrografico.pt/

2.3.7 SNIMAR

The geoportal SNIMAR was developed by IPMA through a EEA Grants. It aims at centralizing multiple types of oceanographic data in Portugal, made available by several contributing institutions.

The geoportal is the web interface of the spatial data infrastructure SNIMAR (Figure 2.7). It allows for many functionalities including downloading of the SNIMAR data and the uploading of external geographical layers available as web services.

In the scope of the present Research Project, this portal can provide important data to feed some of the foreseen outcomes and projects. At the same time, it can take advantage of the reliable data provided by the web portals and reliability framework of Research Study 2, the implementation of the forecasting infrastructure (in particular for nearshore and estuarine predictions) in Research study 1 and the digital twin outputs.

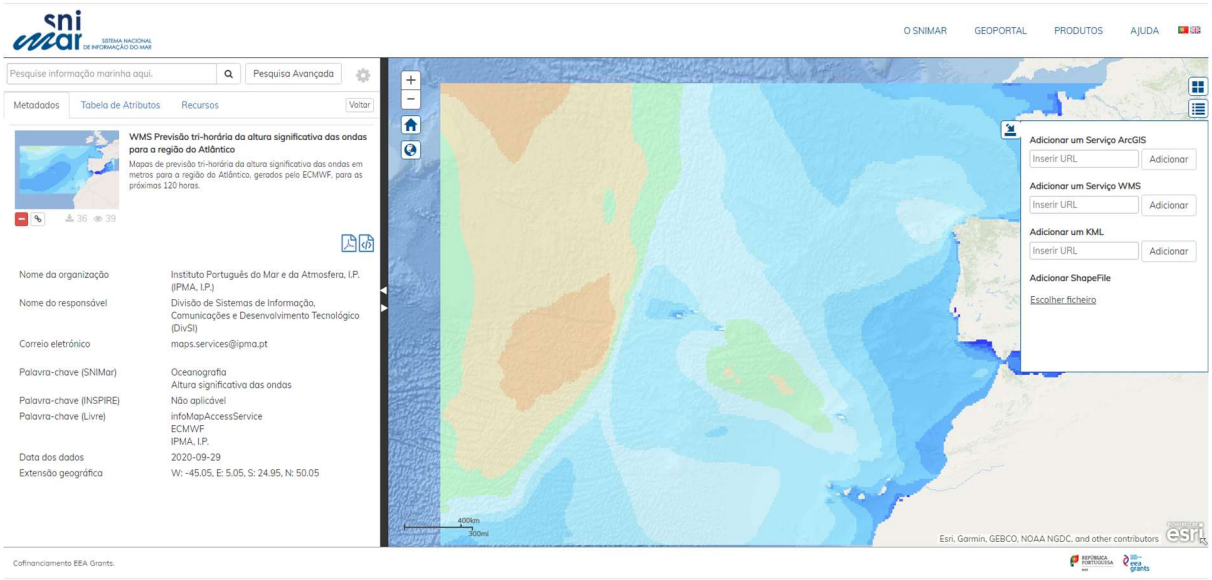


Figure 2.7 – The SNIMAR portal: sample visualization with ECMWF Wave model outputs

Link: http://snimar.pt/

2.3.8 MF2 - Meteorological danger of forest fires portal

The MF2 web portal aims at providing access to meteorological models (ECMWF and AROME) forecasts and satellite data outputs for forest fire prevention, to monitor the occurrence of extreme events such as heavy precipitation, strong winds, thunderstorms, extreme temperatures and rural fire risk (Figure 2.8).

While the goals behind the portal are not directly for coastal zones, the information is quite important for many coastal uses. The Web portal has two main goals: to provide an agile way to look at data with the possibility to overlap layers and to provide a system for data sharing that can serve people and

computers alike, by using standard map services and file formats (netcdf). Metainformation is provide for all layers, allowing for facilitated data interoperability and reuse. This portal follows an open data vision, aiming at the implementation of the FAIR principles. The availability of this source of data is very important for coastal management concerns as it provides forcing conditions for coastal models and a way to anticipate the occurrence of extreme coastal events.

The MF2 portal is a relevant data provider for all Research program projects proposed herein. This is however a one-way sharing as this portal does not allow for the uploading of other data sources.

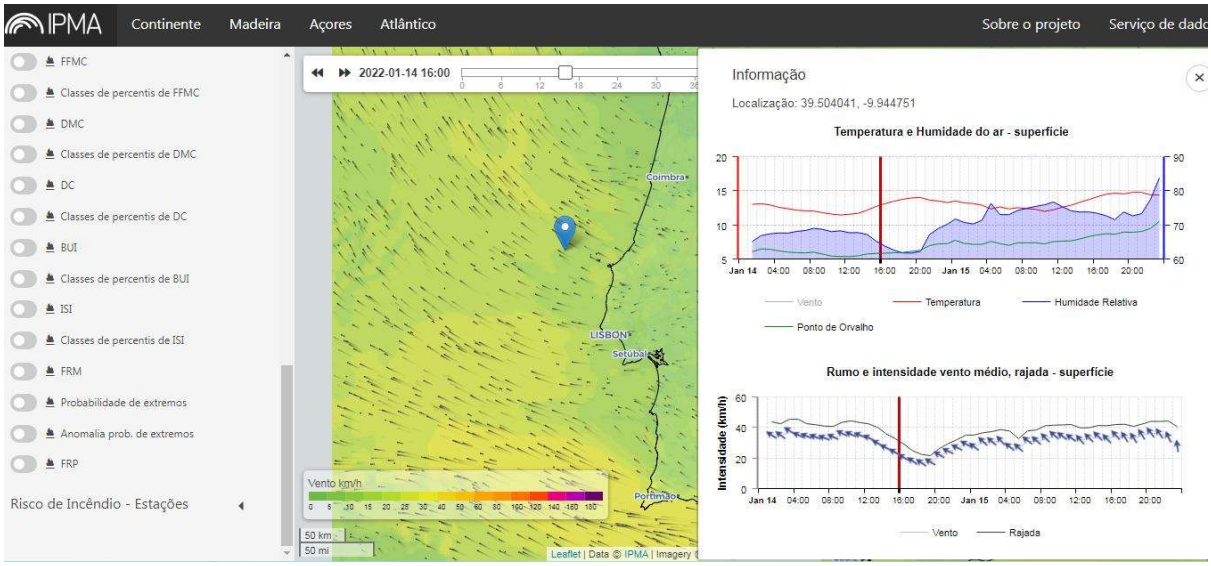


Figure 2.8 – Sample output of wind predictions from the AROME model

Link: http://mf2.ipma.pt/

2.3.9 OPENCOastS - coastal circulation on-demand forecast platform

The OPENCoastS service assembles on-demand circulation forecast systems for selected coastal areas and keeps them running operationally for a period defined by the user. It was jointly developed by LNEC, LIP, CNRS-La Rochelle and the University of Cantábria, being coordinated by the applicant, and has been in operation since 2018. This service generates daily forecasts of water levels, wave characteristics and vertically averaged or 3D velocities, salinities and temperature over the region of interest for 48 hours, based on numerical simulations of the relevant physical processes and the usage of the Water Information Forecast Framework (WIFF), developed by LNEC. OPENCoastS also provides an automatic connection with EMODnet physical data, offering the possibility for comparison between predictions and data for the stations inside the computational domain. Users can setup, manage and visualize through the dedicated Web portal and download both input and output files (Figure 2.9). This service is based on an open research vision, and both the usage of the service and the source code are open, after a simple registration. New developments continue to be done by LNEC to integrate more

complex physics and water quality predictions. Over the past 3 years, this service has attracted over 500 users from more than 65 countries.

OPENCoastS and its underlying forecast engine WIFF are part of the main assets that supports the implementation of the present Research Program, and a fundamental stepping stone for the integrated water forecast framework from Research Study 1, the Digital Twins and the integrated city-to-coast contamination and inundation platform of Research study 2. The availability of this service and all related software suite provides the core for the development of the proposed services targeting multiple society needs.

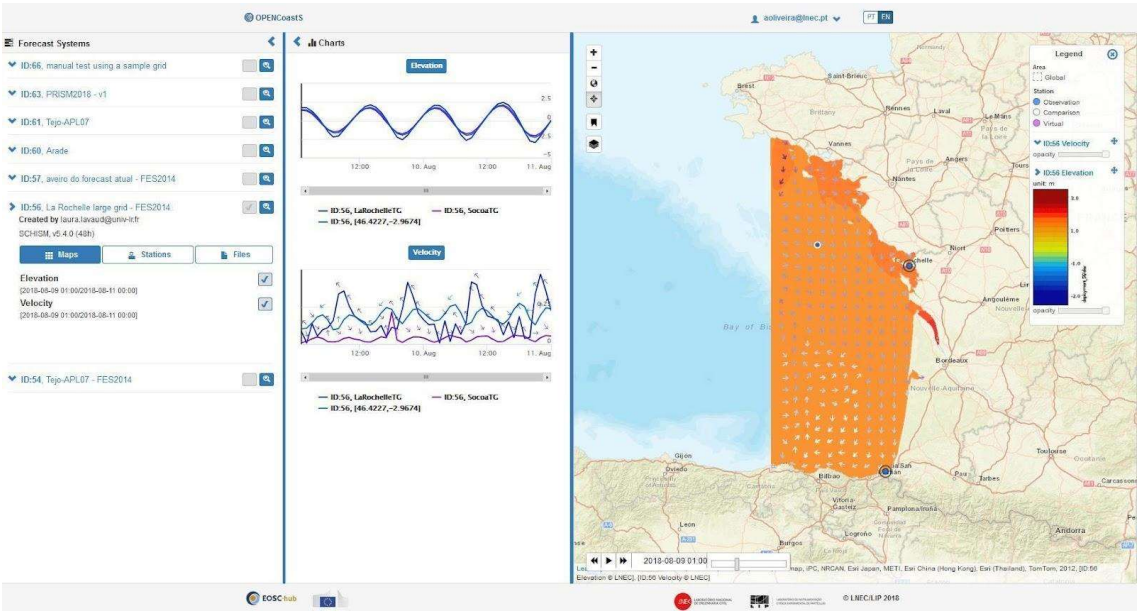


Figure 2.9 – An example of OPENCoastS' viewer products

Links:

- site: http://opencoasts.lnec.pt/index_en.php
- service: https://opencoasts.ncg.ingrid.pt/

2.3.10 WORSICA - Water Monitoring Sentinel Cloud Platform

The Water Monitoring Sentinel Cloud Platform (WORSICA) is an open service that integrates remote sensing and in-situ data for the determination of water presence in coastal and inland areas, applicable to a range of purposes from the determination of flooded areas (from rainfall, storms, hurricanes or tsunamis) to the detection of large water leaks in major water distribution networks. It is a one-stop-shop service to provide access to customized remote sensing services based on Copernicus data, and is currently applied to the detection of the coastal and inland water land interface and the water mains infrastructure leak detection (Figure 2.10). Development is led by LNEC in collaboration with LIP, the applicant's being the co-coordinator of LNEC's participation.

The vision and strategy for future developments of this service is to integrate other remote sensing related services, either provided by LNEC or other users. In this sense, the research proposed herein in the fields of remote sensing data exploration, quality assessment and management fits in this vision and it is expected to contribute for the enrichment of WORSICA's portal, either through new satellite base services or by exploring new possibilities using near field, local cameras to build society-requested products. Integration with WORSICA is expected to occur in a more significant way in Research study 2, but its outputs will also contribute to Research study 1.

Figure 2.10 – An example of WORSICA's coastaline detection products

Links:

- site: http://worsica.lnec.pt/
- service: https://worsica.incd.pt/index/

2.4. International estuarine and coastal management programs and plans and other relevant framing initiatives

2.4.1 Overview

A multitude of directive, laws, initiatives and open-access tools are available nowadays to rule, influence or promote estuarine and coastal zones knowledge and management. Herein the applicant reviews the most relevant ones in the context of this Research Program, without the pretension to be exhaustive. The alignment and the ways that the Research Program contributes or takes advantage of these resources to better accomplish the Program's goals is also discussed.

2.4.2 Marine Strategy Framework Directive

The Marine Strategy Framework Directive was adopted in 2008 with the aim to more effectively protect the marine environment across Europe and to achieve or maintain good environmental status by the year 2020. Coastal waters are an integral part of the marine environment, and, as such, they should

also be covered by this Directive. One of the most relevant measures implemented was the establishment of marine protected areas, where monitoring and assessment of measures to guarantee conservation has a special emphasis.

The 2020's report on the first implementation cycle of the Directive identified that 46% of the European coastal waters still have poor eutrophication status and that the coastal bed is still disturbed by specific fishing methods, and issued several recommendations towards achievement of good environmental status. In the scope of this Research Program, the recommendation for marine data to be comparable across regions emphasizes the importance of the integrated approaches of the forecast framework and the usage of standards among providers to assure that AI data models can be built across regions. Overall, this and the following directives are motivations for the conceptualization of tools that can contribute to their goals.

Link: https://ec.europa.eu/environment/marine/eu-coast-and-marine-policy/marine-strategy-framework-directive/index_en.htm

2.4.3 Maritime Spatial Planning Directive

The Maritime Spatial Planning Directive aims at managing the use of the maritime space to better integrate multiple activities competing for the same area. It promotes an integrative and holistic vision to cope with the increasing demand for maritime space from traditional and emerging sectors while preserving the proper functioning of the marine ecosystems. It aims at reducing conflicts, encouraging economic growth and promoting cross-border collaboration. In order to promote the sustainable use of maritime space, land-sea interactions should be considered as often marine and coastal activities are interrelated. Thus, maritime spatial planning should aim to integrate the maritime dimension of some coastal uses or activities and their impacts to allow an integrated and strategic vision.

Effective marine spatial planning needs tools that enable different end users to access up-to-date information in each marine area. Therefore, this directive motivates the conceptualization and development of most of the work developed herein. While data is at the core of the implementation of this Directive, the possibility to investigate distinct scenarios of occupation can be a potential usage for both forecast framework and digital twins. The quality of data takes a leading role as decisions need to be done over solid and reliable data. Finally, the monitoring of the implementation of the planning requires a dedicated coastal observatory tailored to these goals (Research project 2). A simple implementation for the mediterranean, based mostly on CMEMS was developed, denoted Seawetra. It is is a good pilot on how IT tools can support this directive. Seawetra is a near real-time integrated system for marine ecosystem monitoring and conservation that aims 1) to collect, archive, visualize and share geographical information, 2) to provide stakeholders and different end users with dedicated tools.

Link: https://eur-lex.europa.eu/eli/dir/2014/89/oj

2.4.4 Water Framework Directive

The EU Water Framework Directive, adopted in 2000, requires all Member States to protect and improve water quality in all waters so that we achieve good ecological status by 2027 at latest. It is implemented through river basin management plans and it is applicable to rivers, lakes, groundwater, transitional and coastal waters. The Directive specifies a structured method for developing these plans and classifies each area regarding surface water ecological and chemical status, groundwater chemical and quantitative status among other elements.

One of the challenges for the joint implementation of the Water Framework Directive (WFD) was the sharing of information between countries and making it available for the interested public in an appealing way. To this end the Commission created a platform denoted CIRCABC, a web service for the creation of collaborative workspaces for users to work together and share information, documents and other resources.

The Directive information for each River basin management plan is available at the Water Information System for Europe (WISE) that now provides access to freshwater and marine data, integrating the relevant information not only for this directive but also for all major directives (Floods, Bathing Waters, Maritime Spatial Planning and Marine Strategy). Some countries also developed specific interfaces, such as the Irish platform Catchments.ie, a web portal that provides multiple facilities and data visualization, to facilitate the access to all relevant information. This platform bears many similarities with the proposed estuarine and coastal collaborative observatories, but differing in the availability of forecasts.

The quantification of the several elements in this Directive is supported by a huge monitoring effort in each country, further confirming the importance of field data to support management at all levels. The present Research Program aims at going one step further and looks at integration of data from multiple sources (e.g. remote sensing from camaras and satellites) and AI-based data quality assessment as possible enhancements to the applications of this and the next Directive.

Link: https://ec.europa.eu/environment/water/water-framework/info/intro_en.htm

2.4.5 Floods Directive

This directive, adopted in 2007, aims at establishing a framework for the assessment and management of flood risks to reduce the negative consequences of flooding on human health, economic activities, the environment and cultural heritage in the European Union. Each country needs to assess what are the water courses and coastlines at risk of flooding, determine flood extent and the humans and assets at risk, and then take measures to reduce this risk. For the areas at risk, flood hazard and risk maps must be produced. All this information should be part of the Flood Risk Management Plans, that should indicate the measures to reduce flooding and its impact.

The information generated in this Directive, as in the previous one, is a relevant part of the knowledge to be included in coastal observatories, contributing towards an integrated management that accounts for the several challenges in estuarine and coastal regions, and supporting the UN SDG vision for

sustainability. The present Research program addresses the several challenges in this directive by handling and integrating information from several sources into an information-rich environment in the proposed collaborative observatories.

Link: https://ec.europa.eu/environment/water/flood_risk/

2.4.6 United Nations Agenda 2030 for Sustainable Development

The United Nations Agenda 2030 for Sustainable Development, adopted in 2015, sets a plan of action to mobilize all countries, organizations and citizens to address sustainable development and overcome its challenges. It comprises 17 Sustainable Development Goals (SDG), to erradicate hunger, poverty and inequality and developing human societies, while respecting the planet's sustainability. The importance of coastal ecosystems and the vulnerability of coastal areas in a climate change context are among the challenges identified, and addressed in SDG 14 (Conserve and sustainably use the oceans, seas and marine resources for sustainable development) and SDG 13 (Take urgent action to combat climate change and its impacts). They are most important for the present Research Project.

SDG 14 is mostly focused on the conservation of ocean and coastal ecosystems and in the health of the seas in general. Their preservation or recovery depend however on management decisions that require both innovative methodologies that integrate the wealth of ocean and coastal information (such as the one provided by EMODNET, GOOS, SeaDataNet or GEOSS among others) with the social and economic data to create innovative assessments. In this context, the present Research Program can contribute not only through science-based digital representations of the coastal regions, which allow for evaluation of interventions, but also through collaborative coastal observatories that can accurately predict coastal evolution in a climate change environment.

In SDG13, the increase of resilience and adaptation to climate-related hazards and natural disasters is one of the promoted targets. The availability of tools that can simulate multiple scenarios, thus accurately informing decision makers on the consequences of the hazards and effect of adaptation measures, is one of the usages for the Coastal Digital Twins proposed herein.

The extension of the IT tools developed here to other water compartments can also contribute to other targets in this Agenda.

Link: https://www.un.org/ga/search/view_doc.asp?symbol=A/RES/70/1&Lang=E

2.4.7 European Green Deal

The European Green Deal is one of the top priorities for 2019-2024 for the European Commision. It aims at turning the EU into the first climate neutral continent by 2050. One of the 8 actions in this deal is devoted to Environment and oceans, as they are a source of natural and economic wealth for Europe and thus need to be preserved and protected. This Deal is the umbrella framework for all EU Missions and the top engine for providing funding and opportunities to address the blue economy, protect the coastal and estuarine ecosystems from pollution and protect coastal populations from climate change impacts. Therefore, its vision and missions along with the Digital age priority framed and motivated the

research proposed herein. The outcomes expected to be delivered from the implementation of the Research project aim at contributing towards the goals of this initiative.

Link: https://ec.europa.eu/info/strategy/priorities-2019-2024/european-green-deal_en

2.4.8 EU Mission: Restore our Ocean and Waters

This Mission is one of the 5 initiatives to put research and innovation in practice in the implementation of the European Green Deal. The Restore our Ocean and Waters mission aims at supporting the marine and freshwater goals of the European Green Deal, such as restoring ecosystems, preventing and eliminating pollution and protecting 30% of the European sea area, and protecting the coast from extreme weather events such as floods. It aims at reducing human pressures on marine and freshwater environments, restoring degraded ecosystems and harnessing the essential goods and services they provide in a sustainable way. Protection and conservation efforts must address the entire ocean and water system in a holistic fashion. One of the planned actions that is relevant in the scope of this Research Program is to create a network of Lighthouses at sea and river basin scale to implement the mission and expand the networks of marine protected areas. These lighthouses are organized per basin: Atlantic and Arctic sea basin; Baltic and North Sea basin; Mediterranean Sea basin and Danube River Basin were the choices in the first funding schemes. This mission also includes several cross-cutting enabling actions, such as the development of the Digital Twin of the Ocean and a broad public mobilization and engagement to achieve the mission's goals.

This Mission's implementation plan is well aligned to the present Research program's goals and strengthens the importance of the proposed methodologies and tools. It aims to provide a Global Digital Twin by 2030 and an European one in 2025, and promote deployment of selected coastal-scale forecasting and climate services by 2025. These products can be of great value to provide forcing conditions for the development of the on-demand products such forecast or digital twin's implementations. Likewise, the outcomes of the Research Program can contribute to the goals of the mission.

Link: https://ec.europa.eu/info/research-and-innovation/funding/funding-opportunities/funding-programmes-and-open-calls/horizon-europe/missions-horizon-europe/healthy-oceans-seas-coastal-and-inland-waters_en

https://ec.europa.eu/info/publications/mission-starfish-2030-restore-our-ocean-and-waters_en

2.4.9 EU Mission: Adaptation to Climate Change

The goal of this mission is to support at least 150 European regions and communities to become climate resilience by 2030. To achieve this goal, it will promote the development of innovative solutions to adapt to climate change and encourage regions, cities and communities to share experiences and solutions and promote societal transformation. Like the previous one, this mission also aims to contribute to the European Green Deal objectives (e.g. climate change adaptation, disaster risk, biodiversity). It has three specific objectives:

- Preparing and planning for climate resilience, by supporting European regions and communities to better understand, prepare for and manage climate risks and opportunities;
- Accelerating transformations to climate resilience, by working with at least 150 regions and communities for climate resilient future;
- Demonstrating systemic transformations to climate resilience in 75 large-scale demonstrations.

The Research Program is inspired by these goals and contributes through IT-based tools (such as the on-demand Digital Twins or the collaborative coastal observatories) to the capacity of the regions to understand climate change and evaluating solutions and measures for climate resilient future. This approach has already started to be explored by the applicant and colleagues in the Resilience Assessment Tool App, developed in the scope of the H2020 RESCCUE project, and demonstrated in 3 European cities.

2.4.10 United Nations Decade of Ocean Science for Sustainable Development

The Ocean Decade is promoted by the United Nations and runs from 2021 to 2030. It aims to ensure that ocean science can fully support countries to achieve the 2030 Agenda for Sustainable Development by promoting scientific research and innovative technologies that can connect ocean science with the needs of society. According to the Chair of the Intergovernmental Oceanographic Commission of UNESCO the Decade of Ocean Science "will help to build a shared information system, based on trustworthy, scientific data, from all parts of the world's ocean". The decade has very strict objectives to be achieved (https://unesdoc.unesco.org/ark:/48223/pf0000265198):

- a clean ocean, by identifying and removing the sources of pollution
- a healthy and resilient ocean where ecosystems are mapped and protected
- a predicted ocean, where society has the capacity to predict current and future ocean conditions
- a safe ocean where people are protected from ocean hazards

Several specific aspects to achieve these goals are related to the Research program thematics, namely support to coastal zone management and adaptation, open access to data, information and technologies, and the development of early warning systems.

The Decade has nine challenges to be addressed, out of which the following three were inspirational for the present Research Program:

- Create a digital representation of the Ocean,
- Expand the Global Ocean Observing System,
- Increase community resilience to ocean hazards.

The Research Program implementation may also benefit and/or contribute from several Programmes and Projects in the Decade that have already started:

- ForeSea (Programme) - The Ocean Prediction Capacity of the Future: goals are to (1) improve the science, capacity, efficacy, use, and impact of ocean prediction systems and (2) build a

seamless ocean information value chain, from observations to end users, for economic and societal benefit.

- Digital Twins of the Ocean (DITTO, Programme) - establish and advance a digital framework on which all marine data, modeling and simulation along with AI algorithms and specialized tools including best practice will enable shared capacity to access, manipulate, analyze, and visualize marine information. https://ditto-oceandecade.org/
- CoastPredict (Programme) - CoastPredict aims at transforming the science of observing and predicting the Global Coastal Ocean, from river catchments, including urban scales, to oceanic slope waters. The applicant is part of the Steering Committee of this programme. https://www.coastpredict.org/

The development of tools and methodologies proposed herein aim at contributing to address the Decade challenges for coastal regions and are well aligned with several programmes that have just started. In particular, the integrated forecast framework, the AI data-base models, the digital twins) and the comprehensive coastal observatories aim at developing and sharing knowledge and generic tools that can be applied anywhere. At the same time, the concern on the quality of data and related software from as well as the integration of uncertainty in Research study 1 will contribute towards a portfolio of trustworthy information.

Link: https://forum.oceandecade.org/

https://unesdoc.unesco.org/ark:/48223/pf0000261962

2.4.11 The Sendai Framework for Disaster Risk Reduction 2015-2030

The Sendai Framework aims at “the substantial reduction of disaster risk and losses in lives, livelihoods and health and in the economic, physical, social, cultural and environmental assets of persons, businesses, communities and countries”.

It recognizes that climate change has been exacerbating the impacts of disasters and these have been growing in number and intensity. Exposure is therefore growing and at a faster pace than the efforts to reduce vulnerability. A distinct perspective needs to be used, by looking at disaster risk in a multi-hazard and multi-sectoral approach. At the same time, decision making needs to be supported by the sharing of data as well as easy to access and to understand, up-to-date and science based risk information and communication to all.

The Sendai framework has four priorities for action to prevent new and reduce existing disaster risks: (i) Understanding disaster risk; (ii) Strengthening disaster risk governance to manage disaster risk; (iii) Investing in disaster reduction for resilience and; (iv) Enhancing disaster preparedness for effective response, and to "Build Back Better" in recovery, rehabilitation and reconstruction.

In the scope of the present Research Program, IT-based tools can contribute towards strengthening technical and scientific capacity to capitalize on and consolidate existing knowledge and to operationalize methodologies and models to assess disaster risks, vulnerabilities and exposure to all

hazards, including GIS based platforms. They can also be used to develop and maintain multihazard, multisectoral forecasting and early warning systems, disaster risk and emergency communications mechanisms and hazard-monitoring.

Thus this framework influenced the conceptualization of the forecast framework and the Digital Twins, making risk management a necessary application field for the proposed tools. The products of Research study 1 also explore the tailored forecast through AI-based data models, much simpler to implement and in particular to operate in low income countries. Finally, project 4 illustrates an application to a multi-hazard situation where flood and contamination risk are present due to the exchanges between water compartments. The applicant hopes that the proposed work can contribute toward the implementation of this framework.

Link: https://www.undrr.org/publication/sendai-framework-disaster-risk-reduction-2015-2030

2.4.12 World Health Organization Guidelines on recreational water quality: volume 1 – coastal and fresh waters

The 2021 WHO guidelines aims to protect public health by ensuring that the quality of recreational waters is safely managed, with an emphasis on preventive risk management. They contain the current state of art on the possible adverse health effects of recreational use of coastal, estuarine and freshwater environments; and a set of recommendations for setting national health-based targets; conducting risk assessments; and putting in place management approaches to identify, monitor and control these hazards, and associated public health surveillance and communication. The recommendations are the following:

1. Set national health-based targets for recreational water bodies
2. Develop and implement recreational water safety plans for priority bathing sites
3. Conduct ongoing public health surveillance and risk communication of recreational water - related illness

In the scope of the present Research Program, recommendation 2 identifies a number of challenges to be addressed that are aligned with the Program's goals:

- Develop predictive models for real-time operational monitoring and public communications
- Communicating to the public on the safety of recreational water bodies, based on predictions for on-the-day warnings
- Develop operational monitoring to support rapid warning

The availability of online, near real time data is identified as a relevant asset for rapid identification of hazardous events. Given the complexity of water quality online monitoring and the strict requirements for maintenance of these networks, the possibility to assess the quality of this source of data and the quality classification of the data have a special importance. Therefore, the work proposed herein on data evaluation and quality assessment can contribute to the accuracy of the implementation of these recommendations in estuarine and coastal waters.

The rapid monitoring is presented in this guide in articulation with predictive models, both directed at same-day response on hazardous events. Predictive models provide water users and other beach users with near-real-time information on likely water quality conditions. The proper integration of these forecasts into smart tools (available as apps and websites) using adequate communication procedures can help users to make informed choices on whether to use the recreational water site or not. Thus, the several AI-based tools proposed herein can contribute to these goals in several ways:

- the forecast framework can improve the quality of the local water quality predictions by guaranteeing the full integration of all relevant processes and their timely delivery, accounting for uncertainty and other sources of errors
- the AI-based models can use past water quality data to build site-specific predictive models that can have environmental and other conditions as inputs and thus integrate the previous model into digital twins of the recreational waters
- the enhanced coastal observatories can be used not only for coastal managers to assess site evolution in a climate change context, but also be an information-rich environment to convey timely, accurate information to the recreational site users.

Link: https://www.who.int/publications/i/item/9789240031302

2.4.13 DANUBIUS-RI- International Centre for Advanced Studies on River-Sea Systems

The International Centre for Advanced Studies on River-Sea Systems (DANUBIUS-RI) is a distributed Research Infrastructure (RI) for observation, experimentation and modeling in a range of European River-Sea Systems. Among the RI roadmap ocean and coastal assets (https://www.euro-argo.eu/content/download/157674/file/2021-Report%20on%20the%20EuroGOOS%20event%20for%20Marine%20RIs_FinalVersion.pdf), DANUBIUS is the one that has the closest focus to this Research Program. It aims at providing solutions to societal risks and challenges arising from global and climate change, and extreme events, offering a river to sea perspective. River-Sea Systems comprise rivers and their catchments, estuaries, deltas and lagoons, as well as their adjacent coastal seas, covering freshwater, transitional and coastal waters.

DANUBIUS-RI goals relevant to the present REsearch Program are:

- to enable interdisciplinary research along River-Sea continuum,
- to integrate existing knowledge and providing new interdisciplinary knowledge,
- to use standardized methods and providing access to comparable data.

While DANUBIUS-RI is still an ESFRI and has not reached the ERIC level in the last evaluation cycle, the present Research Program is aligned with its work and vision, namely with the need to combine data and modeling, and the integrated vision from the hydrographic basin to the sea, to produce relevant and innovative knowledge and tools to manage and investigate estuaries and coastal zones. The DANUBIUS data portal is still in its first stages, with only a few datasets available (https://gis.geoecomar.ro/danubius/dataportal/menu.php), but the vision of integrating multiple sources

of information in a single portal is aligned with the concept of coastal observatories addressed herein. The implementation of the proposed enhanced coastal observatories, where AI-models are also part of the modeling suite, can benefit from the availability of data portals such as DANUBIUS's.

Link: https://www.danubius-ri.eu/

2.4.14 Blue-Cloud: European marine thematic platform

Blue Cloud is the "Future of Seas and Oceans Flagship Initiative" European Open Science Cloud (EOSC) marine thematic platform. It delivers a collaborative virtual research environment that integrates multidisciplinary data repositories, analytical tools and computing facilities. Several marine thematic services are available as demonstrators of the core services providers. Users can test and explore two main services: a data discovery and access service and a Blue-Cloud lab.

Several other initiatives have explored this concept of cloud and generic e-services as providers of resources and capacities to build innovative thematic services. EOSC-hub, EGI-ACE, EOSC-Synergy and DICE are examples of these projects and they have motivated the vision behind the integration of IT resources in a seamless way with coastal thematic services proposed in the present Research Program. The vision, facilities and e-services proposed in Blue Cloud may be a valuable asset for the implementation of the projects proposed here as they are also targeted to blue economy purposes. Furthermore, the proposed projects may in the future be integrated in the Blue Cloud portal if it continues beyond the funding project timeline and becomes a repository and main access gate for ocean- and coastal-related services.

Link: https://www.blue-cloud.org/

2.5. International coastal data and forecast platforms and repositories, and coastal IT service providers

2.5.1 Overview

Several data and forecast platforms exist today at European and Worldwide level, to promote data preservation and sharing and support coastal and oceanic research as well as the management of these areas, in particular in a context of climate change. Some of these platforms also integrate predictions at several spatial and temporal scales, with the general aim of supporting emergency and other real time concerns. Finally, integrators and indicators are also computed and shared, to support climate change impact evolution and overall dynamics and at the same time provide necessary information for compliance with European Directives. Herein, the applicant selected the most relevant assets, including those that could benefit from the Program's outcomes and those that can be sources of information to be integrated in the products proposed in the projects.

2.5.2 European Marine Observation and Data Network (EMODnet)

EMODNET is a network of organizations that share data through a common data repository and dedicated portals. Data is collected, processed following international standards and made freely available through interoperable data layers and data products. There are seven portals, each dedicated to one area: bathymetry, physics (Figure 2.11), chemistry, biology, geology, seabed habitats and human activities. Besides providing data, EMODNET also allows institutions to upload their own data, thus providing data storage, preservation and a channel for data sharing.

EMODNET currently offers data from almost all large ocean and coastal data providers in Europe and a few outside Europe. Several aggregators of data such as PANGAEA and SeaDataNet (described next) also offer their data through EMODNET.

Data is easily accessed by people (Figure 2.12) or machines, the last having several options available that follow the Open Geospatial Consortium (OGC) specifications (www.opengeospatial.org) and other standards. WMS and WFS are common to all portals, while some offer OPeNDAP and Rest.

In the context of this Research Project, EMODNET offers the opportunity for multiple exploitations of data, either as part of coastal observatories or AI-models feed data providers. Given the opportunity to upload 3rd party information, EMODNET can also be a channel to share new aggregator data or to integrate new algorithms for data processing. This Research Program shares several principles with EMODNET such as the interoperability of data and open access to data.

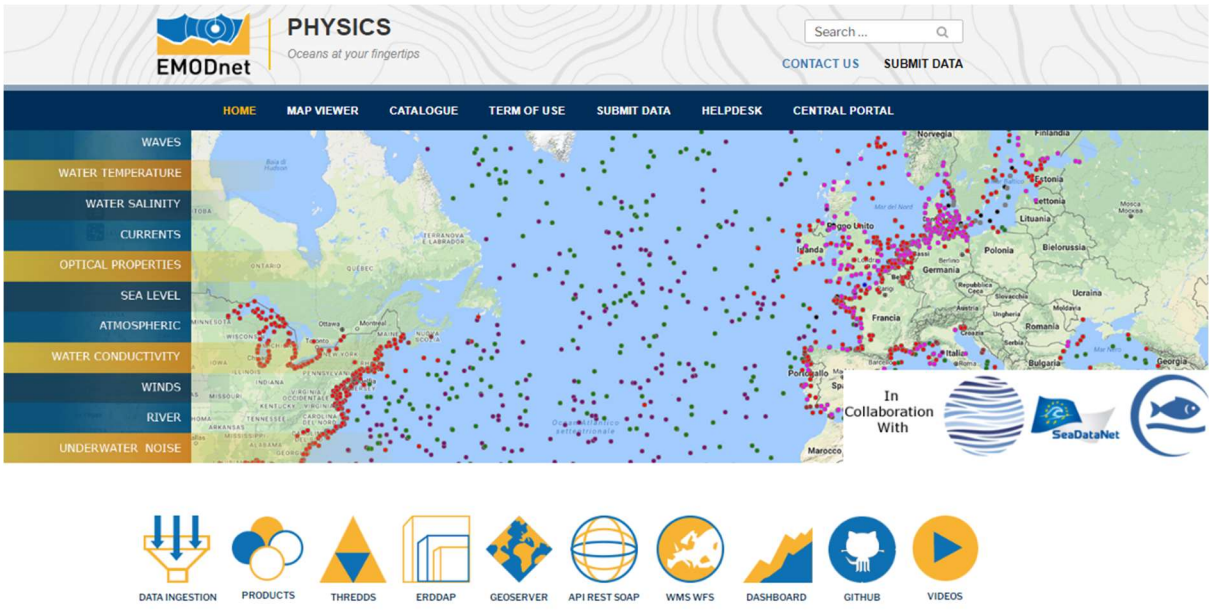


Figure 2.11 – Main portal of EMODNET-Physics

Figure 2.12 – Sample query and output on EMODNET Physics

Link: https://emodnet.ec.europa.eu/

2.5.3 Copernicus-Marine Environment Monitoring Service

The Copernicus Marine Service (or Copernicus Marine Environment Monitoring Service) is the marine component of the Copernicus Programme of the European Union. It provides free-of-charge ocean data and information based on satellite Earth Observation, in situ (non-space) data and numerical models. The Programme is coordinated and managed by the European Commission and is implemented in partnership with the Member States, the European Space Agency (ESA), the European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT), the European Centre for Medium-Range Weather Forecasts (ECMWF), EU Agencies and Mercator Ocean International. CMEMs aims at contributing to several ocean and coastal concerns such as pollution, marine protection, maritime safety and routing, sustainable use of ocean resources, developing renewable marine energy resources, supporting blue growth, climate monitoring, forecasting, among others. The information services provided are freely and openly accessible to its users.

CMEMS is organized along 4 services: data catalogue (Figure 2.13) and portal, indicators, status of the ocean report and visualization.

Access to data is facilitated as each product in the online catalogue of products has its own individual URL that allows direct access to a product sheet. WMS services are also available for uploading of CMEMS products to other interfaces.

CMEMS is a very valuable resource for the development of the Research Program as it is a provider of forcing conditions for the forecast framework of Research study 1, of data for both studies, as well as a provider of boundary conditions for the application of SCHISM and other models in Research studies 1 and 2.

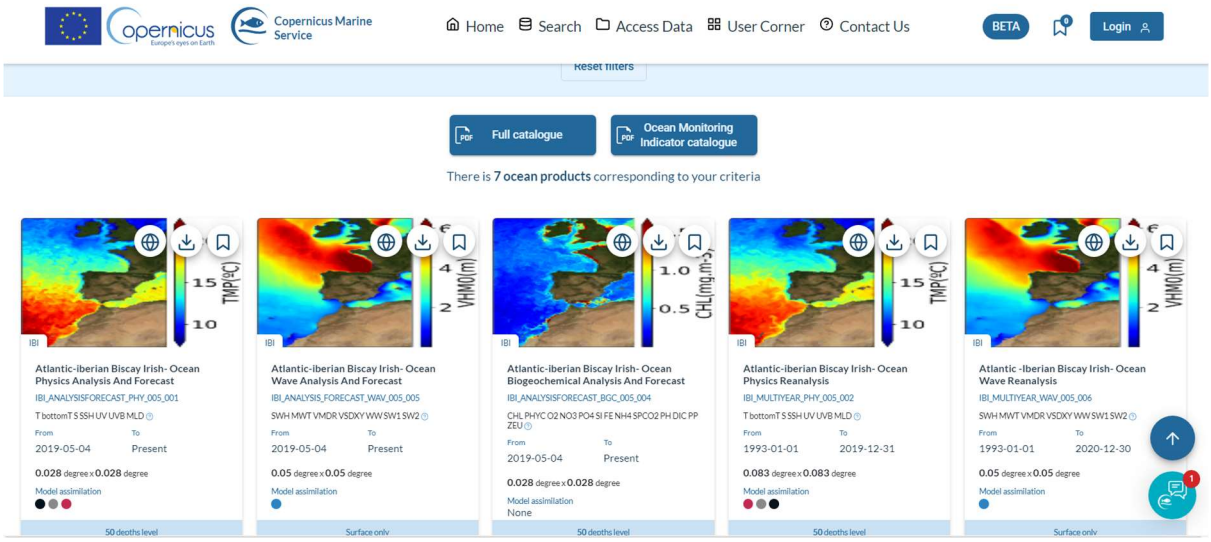


Figure 2.13 – Sample catalogue list for a query on "Iberian"

Link: https://marine.copernicus.eu/

2.5.4 SeaDataNet

Together with CMEMS and EMODNET, SeaDataNet is one of the main Marine data management infrastructures. It is a Pan-European network providing on-line access to in-situ data, metadata and data products through a unique portal that gives access to the interoperable node platforms of the several contributing data centers. Data quality, compatibility and coherence is one of the main concerns, with methods in place to detect data outliers, duplicates and errors generated in processing or transmitting, aiming towards setting a quality flag to each value in a dataset. Values are not corrected as no data recovery algorithm is in operation, which makes the inclusion of the proposed AI-based data processing methodology a relevant addition to this infrastructure. Another interesting feature of SeaDatNet is its virtual research environment (VRE) where registered users can take advantage of the underlying software for its own datasets. Finally, SeaDataNet uses the SEANOE (SEA scieNtific Open data Edition) service to facilitate the publication of marine research data as citable resources.

Like the predecessors, SeaDataNet can be both a provider of quality data for the application of the methodologies and tools proposed herein (if authentication as a service can be used), but it could also benefit from the integration of these tools in its services.

Data access is possible after registration, and its visualization can be done through a web viewer (Figure 2.14). Download of data may require authorization (Figure 2.15) and data cannot be shared with others Web services do not appear to be available for automatic integration with other portals.

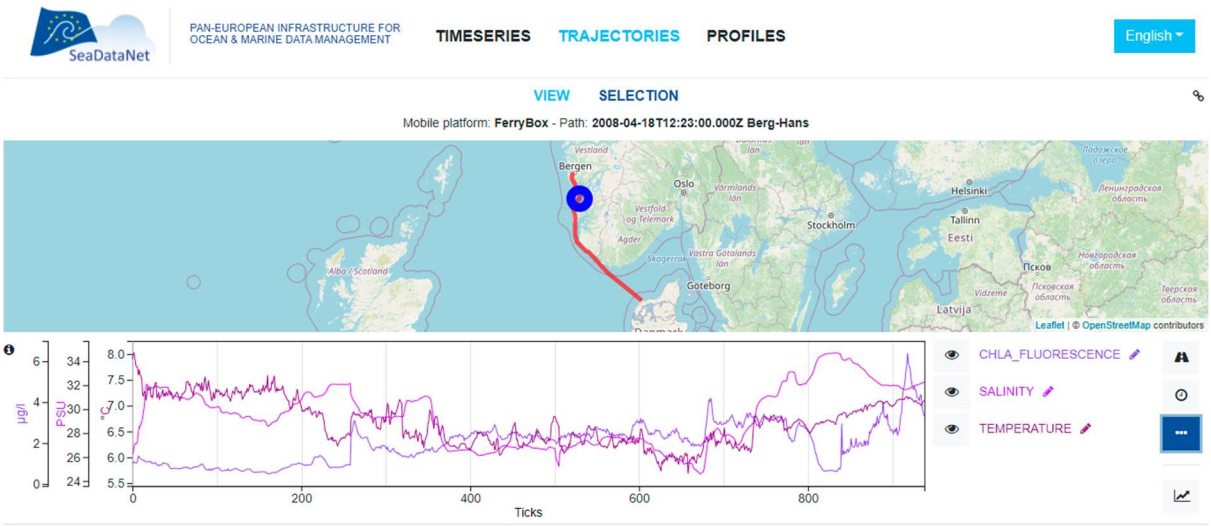


Figure 2.14 – Sample visualization of data in SeaDataNet viewer

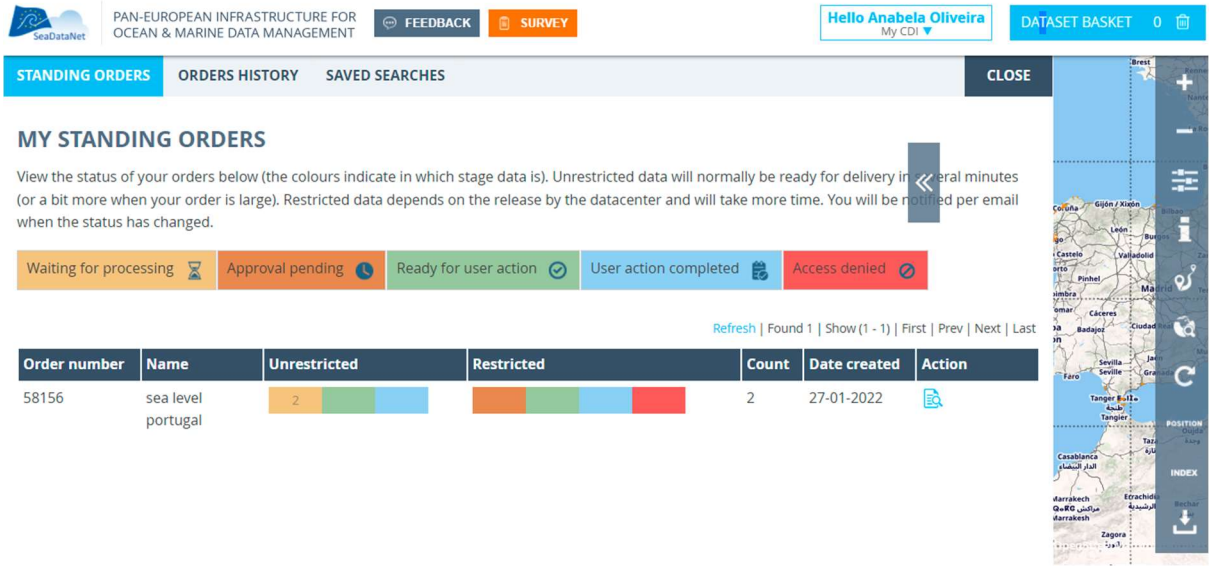


Figure 2.15 – Data search and request example

Link: https://www.seadatanet.org/

2.5.5 GOOS (Global Ocean Observing System)

Besides other missions, GOOS is a worldwide collaborative effort to build a platform of access, download and visualization of observations from a network of independently managed institutions. Data from GOOS is accessible through the OceanOPS which offers network visualization, monitoring and support to implementers and users of the system. Depending on the datasets, quality control and access is offered as well as metadata on the datasets (Figure 2.16). For general users, no automatic

web services are available to map these datasets in an external platform. The possibility of uploading data is allowed.

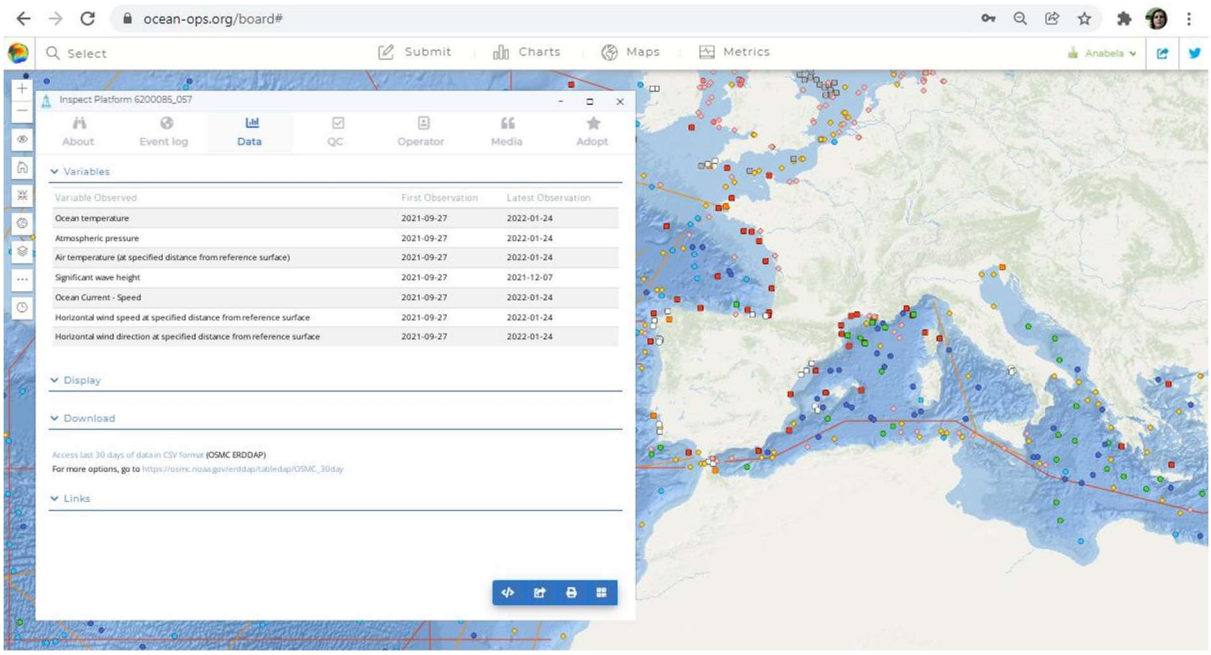


Figure 2.16 – Data access example

At European level, the EOOS (European Ocean Observing System) is the European fit-for-purpose framework for contributing to GOOS and also GEOSS (described below) with the European ocean observing community and networks information.

Link: https://www.ocean-ops.org/

2.5.6 PANGAEA (Data Publisher for Earth & Environmental Science)

PANGAEA is an Open Access data library aimed at archiving, publishing and distributing georeferenced data from geosciences research. Most of the data are freely available and the license is included in the data set description as well as all metadata fields necessary for its reuse (Figure 2.17). Publishing data in PANGAEA is open to all to use or to archive and publish data

Each dataset can be identified, shared, published and cited by using a Digital Object Identifier (DOI Name). Access to data by machines is assured through the data archives with some WMS services and other tools for efficient and standards-complaint queries. PANGAEA can also be used as data deposit with an associated DOI for data publication in the scope of scientific papers publications as supplements to science articles or as citable data collections. Archived data are machine readable and mirrored into our data warehouse which allows efficient compilations of data. PANGAEA is also a data warehouse for several data portals such as GEOSS and OBIS.

Currently only a few datasets from European estuaries and coasts are available. Portuguese data was only found in one dataset.

The present Research Program is aligned with the concepts of open data and FAIR principles that are behind PANGAEA vision and implementation. Interaction with PANGAEA can occur at two levels: PANGAEA's data can contribute to the implementation and operation of the several tools proposed herein, with emphasis to the AI-based data processing and the Digital Twins of of Research study 2; at the same time, data generated in the implementation of these tools in real sites can be deposited in PANGAEA and thus shared openly with all.

Figure 2.17 – An example of a data set in PANGAEA

Link: https://www.pangaea.de/

2.5.7 GEO Blue Planet Initiative

The GEO Blue Planet Initiative aims at promoting collection of continuous ocean observations, processing of data into information and linking this information with societal needs. The 10 thematic areas in this initiative cover many concerns on coastal regions such as sustainable coasts, disaster warning and mitigation, water quality, sustainable fisheries and aquaculture and climate.

The importance of coastal data and forecasts are identified as key to address these areas and to support adequate management, the issuing of early-warnings and the identification of pollution pathways and sources. These tools are also key to protect people and assets at the coast, subject to storm surges, inundation and erosion in particular in a context of changing climate and promote coastal development in a sustainable way. The importance of forecast frameworks that can be deployed anywhere were also identified as a relevant asset for coastal inundation response and management in particular on coral-reef islands. Finally, the gathering of information through dedicated information systems to address water quality and pollution issues, merging remote sensing and in-situ data are also part of the proposals of this initiative.

These concerns and others are promoted through dedicated projects that bring together providers (typically research institutions) and recipients (stakeholders at local, regional and national scales), linked with national and international initiatives. A portal (Oceanscape) was developed to promote the identification of the many organizations that work in the ocean and coastal areas and to promote opportunities to make connections between them.

While this initiative is not an information portal (but rather a portal to find information gateways), the promoted projects identify several of the challenges and tools that are promoted in the Present Research project, thereby supporting the importance and the need of the proposed research.

Link: https://geoblueplanet.org/

2.5.8 Global Earth Observation System of Systems (GEOSS)

As the name states, GEOSS provides access to and links a set of coordinated, independent Earth observation, information and processing systems. GEOSS also ensures that all data are accessible, of identified quality and provenance, and is interoperable to facilitate the development of IT tools and the delivery of information services, thus supporting the development of new systems where gaps currently exist. It also promotes common standards so that data from the multiple providers can be combined into coherent data sets for better Earth knowledge and prediction. GEOSS's priority engagement areas include the United Nations 2030 Agenda for Sustainable Development, the Paris Agreement, and the Sendai Framework for Disaster Risk Reduction. Data is accessible through the GEOSS portal.

The concept behind GEOSS to integrate data providers seamlessly and to guarantee free and interoperable data creates the opportunity for new IT-tools such as the ones proposed herein to address the many societal needs. The priority areas of GEOSS are at the core of the present Research Program's goals such as Climate change adaptation, Disaster risk reduction, Sustainable ecosystems and biodiversity. Furthermore, the work proposed herein in terms of mechanisms for data integration and quality, and forecast and early warning systems can contribute to GEOSS analytics as open data and shared research tools.

Link: https://earthobservations.org/geoss.php;

2.6. National and international digital-related relevant initiatives

2.6.1. Overview

In this section, several initiatives that have contributed towards applied information technology research and innovation in the last years are summarized. These initiatives are effectively contributing towards integration of a digital dimension in coastal zone research, innovation and management or will contribute in the near future. Given the information technology focus on the coastal research proposed herein, the alignment of the proposed projects with these initiatives is particularly important both for funding opportunities in IT-related calls in the future and for taking advantage of the existing knowledge, methodologies and tools to be integrated in coastal works.

The review of initiatives starts with several international and European Commission initiatives, followed by two important works promoted in Portugal.

2.6.2. Destination Earth (DestinE)

Destination Earth is a joint initiative of the European Commission's Green Deal and Digital Strategy. It aims at developing a high precision digital model of the Earth, to support multiple actions:

- climate change impact assessments and prediction;
- predict major environmental degradation and disasters, helping to anticipate events and plan adequate measure;
- assess the impact of environmental policy and legislative measure;
- perform high precision simulations of the oceans and coasts;
- take advantage of developments in the areas of artificial intelligence and high-performance computing to improve modeling, prediction and predictive data analytics.

Several replicas of specific aspects of the Earth system, denoted Digital Twins, will be developed to address specific aspects such as extreme weather events or oceans. Public data sharing will also be promoted, thereby contributing to the European Strategy for Data. This initiative started in autumn 2021 and is expected to last 10 years.

The vision behind DestinE is at the core of motivation for the present Research Plan. Herein, the applicant aims at going a step further by giving users the capacity to build their own AI-based models, forecast workflows or Digital Twins in an on-demand, user selected perspective. While DestinE is focused at a global scale, the products proposed herein have a more local scale, being however generic to be applied anywhere. Integration of these services in DestinE is envisioned by the applicant.

Link: https://digital-strategy.ec.europa.eu/en/policies/destination-earth

2.6.3. Earth System Modeling Framework (ESMF)

The ESMF is a joint American initiative of NOAA, NASA, the USA department of Defence and NSF, to build a framework to support the construction of Earth system models from interoperable components. The vision behind ESMF is that complex applications should be able to be broken into components with interoperable interfaces. It is a high-performance software infrastructure for building and coupling weather, climate, and related Earth science applications, such as coastal or estuarine models or applications to a specific site. ESMF defines an architecture for creating the connections for coupling models and also includes data structures and utilities for developing individual models.

This initiative motivated the applicant´s vision for building a forecast framework that handles physical and biogeochemical processes for the river to the sea, accounting for the urban dimension, uncertainty and error cascading. While the forecast framework proposed herein has smaller spatial dimension of actuation, it is also rooted, like ESMF, on the concept of connecting building blocks to reach the full complexity of predicting estuarine and coastal dynamics.

Link: https://earthsystemmodeling.org/

2.6.4. NOAA's Digital Coast

Digital Coast is a NOAA's initiative to address coastal management in its many dimensions, including Climate Adaptation, Coastal Conservation, Coastal Hazards, Ocean Planning and Water Quality. This site contains data and predictions approved by NOAA, providing services based on visualization tools, predictive tools, and tools for data discovery and facilitated use (Figure 2.18). It also includes tools for citizen science data upload and products dedicated to the general public. In the context of the present Research Program, this successful initiative, dedicated to US territory, motivates the development of integrated resources where data and predictive tools are used to generate activity-targeted services. The concept of Digital Coast bears some similarities with the Coastal Observatories tools and is adapted herein in project 5's collaborative coastal observatories.

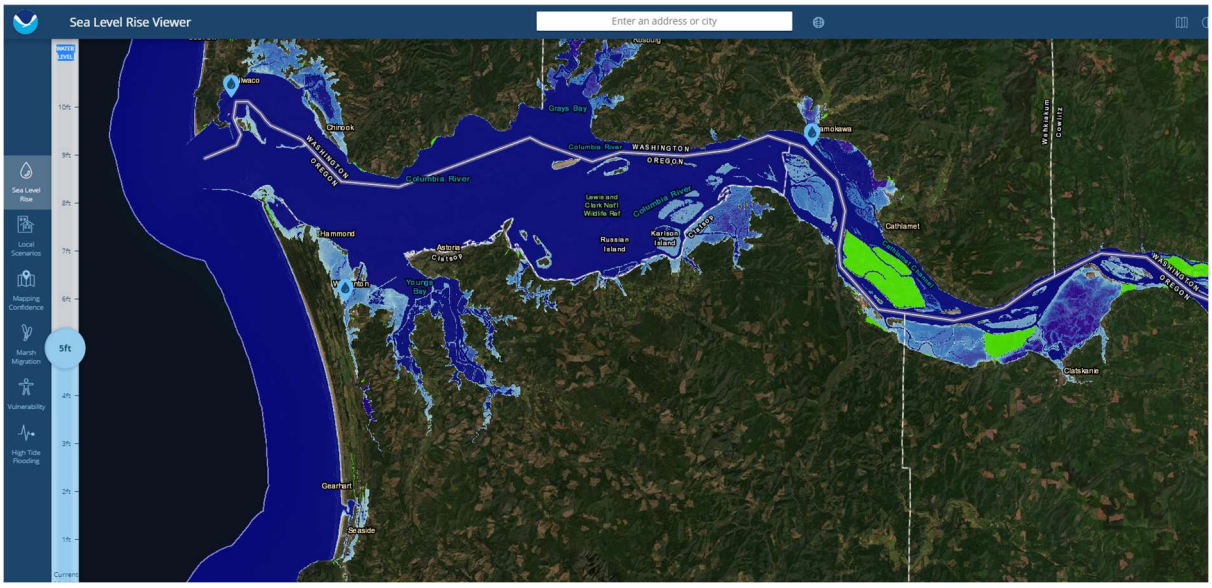


Figure 2.18 – An interactive simulator of sea level rise inundation

Digital Coast is just one of the data and forecast portals operated by NOAA. Other examples include nowCOAST, a web platform to share atmospheric, river, coastal and oceanic forecast for the US territory, and OneStop Data portal, an advanced search portal for multiple types of data in US territory.

Link: https://coast.noaa.gov/digitalcoast/

2.6.5. European Commission priority: A Europe fit for the digital age and the several underlying actions on AI, HPC and data

This initiative of the European Commission aims at harnessing the way Digital technology is changing people's lives. It includes several actions, part of which are relevant in the context of this Research Program: Artificial Intelligence, European data strategy and High Performance Computing.

1) Action: Excellence and trust in artificial intelligence

AI can have a major role in the European economy as it helps to find solutions to many of society's problemas. In particular, the use of AI systems can have a significant role in achieving the Sustainable

Development Goals, in particular in the fields of Climate Change and ocean and coastal zones protection.

The Commission has several initiatives that are being built. The following are of particular importance to this Research Programme:

- Digital Innovation Hubs, one-stop shops to provide access to technical expertise and experimentation, so that companies can "test before invest" and the public administration can uptake their products in a trusted way. The applicant is part of the team of the Attract-DIH project in Portugal and as a European digital innovation hub, devoted to AI and HPC services, which provided motivation for some of the thematics proposed.
- An AI-on-demand Platform as a central European toolbox of AI resources (e.g. expertise, algorithms, software frameworks, development tools) needed for industry and public sector uses. The platform is currently under construction (https://www.ai4europe.eu/) but its services may be of usage to complement and strengthen the research proposed herein in this area. Furthermore, as the work developed in the applicant's division is open source in the vast majority of the projects, the outcomes of the work may be integrated in the future in this platform.

2) European data strategy

This strategy aims at creating a single market for data, which will allow it to be explored freely within the EU and across sectors for the benefit of businesses, researchers and public administrations. Free access to data in the coastal zone can benefit its management, promote the development of tailored products and promote the development of new businesses. In the scope of this Research Programme, data-driven innovation can bring major benefits and the development of data-based artificial intelligence models, as proposed herein, can be an avenue to explore both data and AI methods.

3) High Performance Computing

High performance computing refers to computing systems with extremely high computational power that are able to solve complex and demanding problems. The expansion of data availability and the onset of free remote sensing data and low cost sensing has triggered the need for vast computational resources. At the same time the increased complexity in climate change impact analysis, complex, multi-scale forecast systems and the development of digital twins for the Earth require the availability of HPC to provide timely answers. Therefore, HPC is central to many of the goals from the initiatives outlined above and is also a requirement for the implementation of the several projects in this Research Program.

Link: https://ec.europa.eu/info/strategy/priorities-2019-2024/europe-fit-digital-age_en

2.6.6. Artificial Intelligence national strategy (AI Portugal 2030) - An innovation and growth strategy to foster Artificial Intelligence in Portugal in the European context

The AI national strategy is one of the outcomes developed in the scope of the public policy initiative INCoDe.2030 Program, aimed at enhancing digital competences in Portugal. In this strategy, AI is defined as the ”scientific area and the suite of technologies that use programs and physical devices to mimic advanced facets of human intelligence”. In the scope of this Research Program, the most relevant abilities that AI tools can display are: problem solving, reasoning, inference, decision making, diagnosis, prediction, learning from experience and object recognition.

The strategy has 3 main objectives:

- on economic growth by the integration of AI technologies in companies and the growth of knowledge intensive AI companies;
- on the scientific excellence by improving fundamental and applied AI research of the Portuguese research institutions, promoting an increased collaboration between academia and companies/public sector;
- on the human development, by improving AI competences in the labor force.

While most of this strategy is geared towards improving AI usage for economic growth in industry, AI applications in several fields emerge as part of the program’s vision. In particular, application of new AI developments in environment related, biodiversity areas, including blue economy is proposed as one of the living labs targets. In this context, the present Research program is fully aligned with this target and to the development and application of new AI-based tools for prediction of coastal dynamics based on remote and in-situ sensor networks. Research study 2 explores this concept integrating it with data quality assessment, data reliability and data fusion. It also explores distinct information sources (model scenarios) to build prediction AI models for optimized infrastructure operation to minimize fecal contamination at urban beaches and coastal city inundation. Finally, Research study 2 explores data based models using AI methods as part of the construction of coastal digital twins, fostering real time decision making on coastal management and coastal risk emergency actions.

Link: https://www.incode2030.gov.pt/sites/default/files/julho_incode_brochura.pdf

2.6.7. Portuguese Thematic Agenda for Research and Innovation: Cyber-physical Systems and Advanced Forms of Computation and Communication

FCT’s Research and Innovation Thematic Agendas aim to identify medium- and long-term challenges and opportunities in the Portuguese scientific and technological system. The thematic agenda on Cyber-Physical Systems (CFS) and Advanced Forms of Computation and Communication presents a strategic vision for research and innovation in the development of CFS. It is organized along 4 main areas:

- CFS infrastructure and advanced computational systems;

- Cross-cutting issues: System and data reliability, data and privacy, fault resilience and tolerance, UI usability and man-machine interfaces;
- Methodologies and tools of CFS and advanced computing;
- Cutting-edge technologies and applications (machine learning, big data and intelligent systems research).

The present Research Program includes several topics that are aligned with the areas above, applied to coastal areas challenges. Research study 2 addresses data quality and the associated data reliability evaluation and the usage of data fusion. It also uses machine learning methods and addresses big data issues in particular to process and integrate remote sensing data and models. The proposed Digital twins are indeed Cyberphysical systems, being digital representations of the real world processes, and therefore entail several common concerns, such as CFS modeling, data reliability and real time operation requirements. Reliable early warning and monitoring systems, a strong topic of this work, is one of the examples given in this agenda for the importance of CFS to society.

Link:
https://www.fct.pt/agendastematicas/docs/sistemas_ciberfisicos_e_formas_avancadas_de_computacao_e_comunicacao.pdf

2.7. Strategy for the next decade for GTI and IT-related coastal research at LNEC

The proposed strategy for IT-related coastal research at LNEC for the next decade, to address the compliance/integration/addressing the needs/exploring of the previous initiatives and targeting the placement of LNEC as a center of excellence in this area, is therefore the following:

- to build a coastal forecast infrastructure that addresses all concerns of predicting coastal conditions for short (emergency) to long term (risk, sustainable planning). It should account for the dynamics of all bordering water compartments (ocean, rivers, city) in an integrated way, quantifying the cascading error propagation, the forcing and parametrization uncertainty and solves all relevant processes and their interactions from hydrodynamics to ecosystem modeling. Finally, it should be available as a service, following the work developed in OPENCoastS, and aiming at its use by all interested in the coastal zone, either for knowledge building, management, exploring and protecting its resources.
- to create and develop the concept of Coastal Digital Twins (CDT) as the tool of excellence to integrate data and prediction tools in a seamless way, bringing in process-based or AI-based tools towards hybrid prediction and making the capacity to build replica of real systems available to all in a user-friendly way. CDT will be tailored and customized to user needs either to address site-specific multi-dimensional problems or thematic concerns across water bodies, aiming at building collaboratory platforms targeting co-creation of solutions by all involved and the capacity to anticipate interventions impact, climate change actions or simply to monitor coastal domains under multiple stress factors. Usability strategies will make complex tools and

knowledge from the scientific and technical communities available for non-experts to ask “what if” questions and to share the work with other users in an open science, open information vision. Finally, this effort will be fully integrated with national and international initiatives at the ocean scales, in order to address global challenges such as the UN decade of the ocean.

3. State-of-the art: Forecasting the dynamics of the estuaries and coasts using IT tools

3.1 Introduction

In the late 20th and early 21st centuries, the oceanographic community started developing forecast systems to provide short-term predictions of ocean hydrodynamics (e.g., Clancy and Sadler, 1992; Brassington et al., 2007; Baptista et al., 2008; Mehra and Rivin, 2010). These forecast systems used hydrodynamic models forced by atmospheric model predictions to provide, among other variables, sea surface elevations, currents and wave spectra a few days in advance. Ocean forecast systems are nowadays very complex and detailed tools, including data assimilation from both in-situ, moving and remote sensing sources and biogeochemistry predictions. The Harvard Ocean Prediction System is one of the pioneering ocean forecast systems, which has evolved to the MSEAS system (http://mseas.mit.edu/) that includes uncertainty and data assimilation analysis as well as biogeochemistry forecasts.

Coastal forecast systems started then to be developed, fueled by the success in the adjacent ocean forecasting tools. Coastal forecast systems are used for many purposes, including harbor management, search and rescue operations, and response to extreme events. Like its ocean counterparts, coastal forecast systems provide predictions of environmental variables at time scales of a few days (Oliveira et al., 2020). Environmental variables include water levels, velocities, wave parameters, pollutant concentrations and sediment fluxes. In the last decade, these systems have been applied to multiple uses: coastal (Viegas et al., 2009; Bedri et al., 2014; Oliveira et al., 2021, Zou et al., 2021) and harbor management (Poseiro 2019), civil protection (Breivik and Allen, 2008; Fortunato et al., 2017; Ferrarin et al., 2019; Stokes et al., 2020), navigation (Orseau et al., 2021), military operations (Blain and Preller, 2007) and recreation (e.g. windguru.cz, magicseaweed.com/Spain-Portugal-Surf-Forecast/8/).

Much work remains yet to be done in coastal forecast systems to achieve the sophistication of ocean forecasts and the stringent requests for predictions in estuarine and coastal zones. The advances in ocean forecasts need to be adapted for the complexity and non-linearity of the coastal and estuarine physical processes, and to the requirements of coastals users (e.g. small spatial scale representation, products tailored to coastal uses, proper integration of urban inputs and freshwater sources and its

associated uncertainty, accurate representation of complex dynamics of transitional waters from the hydrodynamics to the ecological modeling, …).

In this chapter, the applicant reviews the most relevant issues in the estuarine and coastal forecasting, based on several reviews on this theme and a collection of carefully chosen publications. In the end of the chapter, the open issues and directions for future research are identified, as a stepping stone for the proposed Research Studies.

3.2 Numerical model-based forecasts systems

The vast majority of forecast systems are based on applications of process-based numerical models, given its generic nature and applicability in multiple systems. The quality of the forecast outputs is thus strongly dependent on the underlying model, the processes that it solves and the quality of the model application (Fortunato et al., 2017, Oliveira et al., 2020, Umgiesser et al., 2020). The efforts of the scientific community in process knowledge and numerical modeling improvements and the developments in computational techniques allowed to increase the simulation capacity, improve the accuracy and resolution of the numerical models and reduce the computational processing time (Umgiesser et al., 2020). As a consequence, forecast systems have also been evolving to very complete prediction of the relevant processes in each application. For instance, hydrodynamics modeling in the near shore and at tidal inlets have been improved considerably in the last years (Bertin et al., 2014, Fortunato et al., 2017).

The initial forecast systems for estuarine and coastal systems followed the similar path as numerical modeling. Finite difference models, faster and computationally less demanding, were the first computational engines and several well known prediction systems were developed for site specific applications (Shaffer et al., 1989). The increase in computational efficiency in unstructured grid models in the last decades, in particular by resorting to parallel computing paradigms, and the availability of extensive computational resources, has triggered a major shift towards these models. By facilitating spatial representation at the necessary scale both in the horizontal and vertical dimensions, unstructured grid models can represent complex processes in a more efficient and accurate way and are now at the heart of the most recent and complete forecast systems (Akbar et al., 2013, Karna and Baptista, 2016, Rodrigues et al., 2021, Moghimi et al., 2021) although finite differences systems also coexist (Mateus et al., 2012, Solano et al., 2021). Even for oceanic scales, unstructured grid modeling is starting to play a relevant role (Akbar et al., 2013, Moghimi et al., 2021), allowing for the construction of large forecast systems that encompass several orders of magnitude in spatial resolution (from the river to the deep ocean), allowing for the accurate simulation of hurricanes and other phenomena that travel considerably across scales (Akbar et al., 2013, Moghimi et al., 2021).

Other types of numerical models are also used for forecasting coastal and estuarine dynamics. Particle models are particularly suited for some water quality predictions on top of grid-based models (Daniel 1996) as well as for small scale forecasting of wave overtopping in structures (Neves et al., 2021). These predictions are generally implemented in small localized areas, given the huge computational

effort necessary for the massive number of particles, but their results are far better than the traditional use of empirical formulae commonly used in port early-warning systems (Fortes et al., 2020).

Recently, model ensembling is also being explored for estuarine and coastal forecasting (Iglesias et al, 2022, Umgiesser et al., 2020), aiming at reducing the uncertainty and errors associated with model physics and numerical characteristics, and model parametrization. These methodologies are described in more detail below.

3.3 Data-based forecast systems

Data-based models are typically site dependent tools, based on the analysis of the data characterizing the system under study. They comprise statistical approaches (Solomatine and Ostfeld, 2008, Kim et al., 2016, Ganguli and Merz, 2019) and artificial intelligence (AI) approaches (French et al., 2017, Mosavi et al., 2018). Examples of statistical models for flood prediction are autoregressive moving average, multiple linear regression, and autoregressive integrated moving average models (Mosavi et al., 2018). However, their accuracy is strongly dependent on data stream size and have been reported to be inadequate for short term predictions (Mosavi et al., 2018). AI-based models are becoming increasingly more popular given the availability of computational power necessary for the model training phase, a direct measure of the final quality of the predictions, and low cost data sources for domain and process representation. Model response under emergency conditions is much faster for AI data-based models, once the training is concluded. Recent developments have started to explore the best of both worlds by combining artificial intelligence and process based models (French et al., 2017).

Artificial intelligence models depend on significant volumes of data to mimic the complex mathematical expressions of physical processes of water dynamics. Many ML algorithms, e.g., artificial neural networks (ANNs), neuro-fuzzy, support vector machine (SVM) and support vector regression (SVR) were reported as effective for both short-term and long-term flood forecast (Mosavi et al., 2018). They do not require hydro-geomorphological data (only the guarantee that the data points are properly selected to cover all relevant areas of the domain at the necessary time pace to represent correctly the physical processes), can handle well flow nonlinearity, are quicker to develop with minimal inputs and can be used by non-experts on the physical processes at stake, although a minimal base of knowledge is required to specify the relevant spatial and temporal scales for the setup of the AI-models. On the contrary, process-based models require hydrogeomorphological information to set up the computational grid, boundary conditions and in-situ or remote sensing data for its calibration and validation. AI methods are becoming very popular in several water domains research such as river floods because they provide better performance and cost-effective solutions (Mosavi et al., 2018). AI models have been used successfully in rivers (Hagen et al., 2020), with models of varying complexity, such as ANNs, neuro-fuzzy, adaptive neuro-fuzzy inference systems, support vector machines, wavelet neural networks, and multilayer perceptron. These types of models can be explored in coastal environments too. The usage of AI-based models for estuarine and coastal operational purposes is still limited although on-going experiences are providing promising results (Umgiesser et al., 2020). The

importance of proper training in AI-based models for accurate predictors has been identified in several coastal applications, given its complexity (Ferreira et al., 2018, Jesus et al., 2015).

In some simple contexts, where probabilistic relations between offshore forcing conditions (e.g., wave height), local hazard intensity (e.g., erosion and inundation) and impact at the receptors can be established, Bayesian networks can be used successfully to replace process-based models (Ferreira et al., 2018). Forecasts using ML or other AI methods are however generally concerned with the lead time to hazardous occurrence, and do not address the spatial distribution of the events (Mosavi et al., 2018). In coastal regions the spatial dimension of the hazardous events, either being floods or contamination events is very important. Therefore, the joint application of both AI and physical process models or its combination into an assimilated result appears to be an avenue to be pursued in the future (French et al., 2017).

3.4 Forecasts systems for estuarine and coastal circulation prediction

3.4.1 Overview

Circulation forecasts are nowadays a powerful tool for coastal managers and companies operating on the coast and are also a helpful resource for recreational uses, given the free access policy and intuitive forecast-sharing interfaces for the vast majority of the forecast providers. Requirements for coastal circulation forecasts are dependent on the specific event to be predicted (e.g. hazardous event or daily management of coastal areas) and on the nature of application of the forecast outputs (e.g. civil protection agency early warning time concerns, inundation area accuracy for safe landing in military operations). There are many circulation forecast systems in operation throughout the world and in Europe in particular, at local, national, regional or global scales (Kourafalou et al., 2015).

The Atlantic coast of Europe has been subject to severe storm surge events in the last decade, mostly due a combination of hazards and to large continental shelves where wind-driven surges are much larger than in deep waters (Umgiesser et al., 2021). The impact of these events is expected to increase due to two effects: 1) predicted climate change increases the hazards of sea level rise and storms; and 2) continued growth of coastal occupation will raise the exposure (Dongeren et al., 2018). The consequences of floods for Atlantic coastal regions has fuelled the development of several forecast systems that simulate tides and surges, accounting also in some cases for riverine inputs and for wave and current interactions.

In the Iberian Peninsula several hydrodynamic forecast systems are operating daily and sharing forecasts on the web. In Portugal several institutions operate daily circulation forecast systems that predict tides and storm surges and other phenomena covering the adequate spatial scales, and publishing the forecasts openly on the Web. LNEC has been operating the Water Information Forecast System for the North Atlantic and Iberian coast water levels and wave parameters for over a decade (Fortunato et al., 2017) and offers in the last years an open-source, open access, on-demand service -

OPENCoastS - for users to setup, manage, download and visualize their own 2D and 3D wave and current integrated forecast system (Oliveira et al., 2021, https://opencoasts.ncg.ingrid.pt/). Tailored predictions for port areas are also provided by the HidraAlerta system (Poseiro, 2019). Navigation support forecast is also available for the Ria de Aveiro, through the Routinav app, supported by a carefully calibrated and validated hydrodynamic forecast system (Picado et al., 2022). IPMA is the official hydrodynamic prediction provider and offers images of predictions for waves and sea level at the official page (https://www.ipma.pt/en/maritima/ssh/) while IH presents wave predictions and tide predictions also in map format (https://www.hidrografico.pt/prev.mare). IPMA's products are geared towards several applications such as supporting the Civil Protection actions, while IH's products are more devoted to maritime activities such as navigation or surf. LNEC's forecast systems are tailored to specific end users needs (e.g. local Civil Protection agents) and the on-demand nature of OPENCoastS allows for the users to tailor their forecast for their specific needs (from research to recreational uses, including also coastal management actions). Several universities have similar systems that operate typically in the scope of research projects (e.g. Mateus et al., 2012) In Spain, Puertos del Estado are the official providers of regional and local map and time series predictions (https://www.puertos.es/en-us/oceanografia/Pages/portus.aspx), while several other providers such as MeteoGalicia provide not only time series but also web services to access predictions. In France, MeteoFrance is the official provider and wave and current circulation forecasts are simulated daily and provided as webservices (https://portail-api.meteofrance.fr/devportal/apis/80afa827-d24c-4c8e-8c53-a556b3c80f3a/overview). In the UK, the British Met Office provides forecasts from the regional to the local scale (https://www.metoffice.gov.uk/weather/specialist-forecasts/coast-and-sea), supporting warnings and emergency operations. Similar systems are available worldwide.

These forecast systems are the backbone of many local forecast systems, tailored to the needs of water and port authorities, companies supported by coastal and estuarine resources, environmental protection and civil protections agencies and the society at large. Indeed, forecast products could address many critical and relevant coastal issues, either at emergency (floods, structure overtopping/destruction, search and rescue, pollution events,...) or as a support for daily and long term management such as marine spatial planning, integrated coastal zone management, climate change evaluation and mitigation.

3.4.2 Estuarine and coastal flood and overtopping prediction

Coastal flooding is by far the most common usage for circulation forecasts, aiming towards issuing accurate early warnings. Indeed, early warning systems are nowadays fundamental instruments in flood risk emergencies and their accuracy is deeply rooted in the use of operational flood forecast systems (Hagen et al., 2020). Accurate storm surge forecasts are crucial to predict in a timely way extreme sea levels that flood lowland coastal areas (Ferrarin et al., 2019) and generate saline plumes that infiltrate shallow coastal aquifers (Giambastiani et al.2017). Coastal flooding and defense structures overtopping is one of the application areas that has triggered considerable research in the last decade.

In order to obtain accurate time of arrival of coastal and estuarine flood and flooding areas it is essential to predict the hydrodynamics (and often the morphodynamics) from the ocean to the land scales. Many forecast systems have been developed, using process-based models, empirical formulae, data-driven models or a combination of all (e.g. LNEC's WIFF, UK Met Office MOGREPS), The vast majority of these systems account for hydrodynamics only (Fortunato et al., 2017), neglecting morphodynamic effects and some of them do not account for the current and wave interaction (Hidralerta, Poseiro 2019). Recent developments have included interaction between waves and currents (Zou et al., 2021, Oliveira et al., 2020, Oliveira et al., 2021) and are starting to include morphodynamic evolution also (Nahon et al., 2020, Valchev et al., 2017).

User requirements for flood forecast systems can be quite distinct depending on the application to be done to its results. Inundation areas are the most important output of inundation forecasts, along with the timing of the flood wave. For assets near the coastline (such as buildings or roads), the concern may also include the possibility and timing of the overtopping of coastal defense or for ports and marinas, the overtopping of port structures (Poseiro, 2019). Military operations, such as landings, can be affected by inundation by changing the water/land interface, changing soil conditions in the inundation area and altering the currents at the landing spots (Blain and Preller, 2007). Forecast systems are also starting to be used for other purposes than early warning systems in the risk management workflow, by supporting humanitarian aid prior to anticipated flood events. These initiatives are still at pilot stage in Red Cross National Societies across Africa, Asia and Latin-America (see https://www.forecast-basedfinancing.org/our-projects/".

Typically, estuarine and coastal inundation and overtopping forecast systems are forced at their ocean boundaries by global tidal models (such as FES2014 - Lyard et al., 2020) and by inverse barometer effects to account for pressure (e.g. Fortunato et al., 2017). In the last decade, several regional coupled storm surge and tidal models based on unstructured grid were developed, aiming at a Pan-European Storm Surge Forecasting System with better accuracy than a decoupled forcing (e.g. EU-SSF, Fernández-Montblanc et al., 2019; Regional WIFF, Fortunato et al., 2017). However, the experience in some paradigmatic cases such as the Xynthia event showed that this coupling is not sufficient and that wave action and wave and current interactions, namely infragravity waves and wave run-up, are needed for an accurate flood prediction (Bertin et al., 2014). Therefore, one of the avenues to be pursued in the future is the development of fully coupled regional forecast providers for local forecasts, all integrated in a seamless way. The tools to address this challenge already exist, such as the new version of OPENCoastS (OPENCoastS+, opencoasts.lnec.pt).

Integration of wave overtopping modeling in detailed flood prediction has started to be considered in the last decade (Poseiro, 2019). Traditionally wave overtopping with empirical formulae was typically used with surge and tidal levels obtained by external sources without considering the interaction between waves and currents and their effects on water levels and velocities (Fortes et al., 2020). Recently, the need to accurately simulate overtopping events in ports, marinas and in coastlines with fixed defense structures has promoted the improvement of forecast systems to include artificial intelligence models for overtopping simulation, using artificial neural networks (ANN, Oliver et al., 2021,

Fortes et al., 2020), and the use of circulation models for water level and velocity predictions. Indeed, flooding and overtopping in urban waterfronts is mainly controlled by high energy wave action and high-water level, being a complex problem to simulate due to plunging wave breaking and strong water/air mixture turbulent flow. ANN are powerful analytical, parallel processing tools that can successfully approximate any complex non-linear processes (Oliver et al., 2021). Fine detail overtopping forecasting, using for instance particle and volume of fluid models, is still limited by computational constraints for operational purposes (Neves et al., 2021). Indeed, a 600 s real time 3D simulation of a typical beach, employing 10 million particles (23 per wavelength), took about 170 hours (3,4 hours per wave period). Exploring highly efficient particle models (for instance from the water quality coastal modeling community) and high performance computational resources may be a future avenue of research in this area.

As numerical modeling provides a better representation of physical processes, accounting for all relevant processes, grid resolution needs to increase drastically to be able to represent fine bathymetric and topographic features as well as man-made structures such as dykes and barriers, built to prevent flood progression inland. Moreover, during severe events, these man-made structures may fail, breach or collapse and the capacity to simulate these collapsing processes requires very fine resolution. While the consideration of these structures is usual in river flooding studies, they are frequently disregarded in coastal studies (Umgiesser et al., 2020). As an example, Bertin et al. (2014) found that a resolution of a few meters was required to represent dykes and barriers and thus achieve acceptable accuracy in Xynthia's flooding modeling, together with the need to simulate the whole north Atlantic to represent properly all processes influencing flooding, which led to a computational grid of over 1,5 million nodes.

The computational resources to address these stringent conditions are nowadays a major limitation to the generalized application of forecast systems, as in-house resources may not be sufficient for timely daily runs of fine-grid forecast systems to achieve the desired accuracy. Resorting to shared e-infrastructures, providing cloud or grid resources, is starting to be an attractive solution. Indeed, cloud resources are becoming increasingly popular in environmental modeling, due to their flexibility to provide resources on request and to the simplicity of their use (Chandrasekar et al., 2012). Cloud resources are today the backbone of several complex modeling and data-based applications, providing computational power to many application fields (e.g. Glenis et al., 2013). Multi-tenancy and node-to-node communication in parallel computations remains an issue for environmental applications (O'Donncha et al., 2016), which can be further aggravated by the need to guarantee the delivery time of the forecasts (Rogeiro et al., 2018). A comparison of model performance indicators for an operational hydrodynamic and water quality forecast simulation executed in local workstations, a HPC cluster and a pilot cloud revealed the good performance of cloud computing resources (Rogeiro et al., 2018).

3.4.3 Baroclinic forecast predictions

Baroclinic processes are a fundamental part of ocean and estuarine dynamics. In estuaries, the balance of forcings (tides, freshwater flow and atmospheric conditions) lead to very distinct circulation characteristics, from barotropic to baroclinic dominated flows (Karna and Baptista, 2016). This variability

of conditions, in addition with complex geometry and bathymetry, large velocities and sharp gradients between water masses make estuarine baroclinic circulation modeling a complex task. As a consequence, the quality of baroclinic flows' prediction is strongly dependent on the quality of the prior calibration and validation of these models and the quality of the forcing conditions. Unlike typical 2D circulation models, the calibration and validation of baroclinic models should cover the most complete sampling of system conditions to occur, requiring therefore long term monitoring (Karna and Baptista, 2016). In these systems, the representation of the river plume dynamics is also crucial to achieve reliable coastal forecasts (Kourafalou et al. 2015). These difficulties are extended to baroclinic flow forecast systems, associated water quality and ecology predictions and other types of phenomena. For instance, Li and Nie (2017) point out that the prediction of a Typhoon cyclone, requires ocean waves, heavy precipitation, inundation, estuarine block and salty invasion prediction for an accurate overall risk estimation. Climate change impacts on temperature and river flow will have a major impact on estuarine baroclinic circulation, with consequences on biogeochemistry, thus making the accuracy of river flow estimation a fundamental one (Pereira et al., 2022).

Addressing freshwater predictions is thus a crucial element in estuarine forecasting. Traditional approaches are based on the adoption of climatological values for the river boundaries (Ferrarin et al., 2019, Oliveira et al., 2014), but they fail to represent typical high frequency variations in river inputs in particular during river floods. Other approaches include the integration of river flow forecast systems as part of the modeling chain (Campuzano et al., 2016, HypeWeb - https://hypeweb.smhi.se/explore-water/forecasts/). River flood forecasting systems have been developed in the last decades based on numerical models (Campuzano et al., 2016, Jain et al., 2018; Barthélémy et al. 2018), data-based models (Okuno et al., 2021) and hybrid approaches using machine learning tools (Young et al., 2017). For freshwater systems controlled by dams, a prediction of river flow at the estuary head under rainy/flood conditions is difficult to achieve through data extrapolation based on previous days or through river flow forecast systems that do not account for dam management procedures. Machine learning tools applied over long datasets provide good predictions (Valipour et al., 2013) but these datasets are seldom available in most basins and changes in regulations for dam operation may hinder their accuracy.

The integration of a hybrid, dynamic, generic approach that includes both basin forecast models and real time upstream and downstream data inputs, all merged into a ML flow model, dully integrated into the coastal forecast chain through standard connectors that allow the system to include new modules, new structures and new operation rules, appear to be the avenue for future basin-to-sea full fledge forecast systems. This approach can easily be extended to address other sources of freshwater (for instance urban discharges), by applying the same standard, FAIR-complaint connectors, and thus account for the full set of processes influencing baroclinic flows in estuaries and coastal regions.

3.4.4 The importance of morphodynamics and the challenges to address bathymetry evolution in forecast systems

The quality and resolution of the bathymetry to set up a coastal forecast system is critical but maintaining its accuracy over time is an even bigger challenge (Kourafalou et al., 2015). While there are open source regional and global providers (e.g. GEBCO 2020) and local coastal sources (such as Instituto Hidrográfico in Portugal) for the model setup, coastal topography and bathymetry change on both intra- and inter-storm scales (in particular during severe winter storms) and many coastal and port engineering interventions change the geomorphology of coasts, such as altering the coastal boundaries, increase port entrance channel depths or implement beach nourishments, to name a few. Reliable automatic ways to update the topography and bathymetry in coastal and estuarine forecasts are not available yet and the long term accuracy of forecasts could benefit from its integration in forecast frameworks. Recent works on extracting topography and bathymetry from satellite images are promising (Fassoni et al., 2021) but they are not intended at short term (storm) events geometry update.

Additionally, the vast majority of process-based forecast systems, either for hydrodynamics or water quality predictions, assume that the bathymetry is fixed and that sediments do not interact with other processes. In some forecast systems water quality modeling of decay processes consider sediments as a sink for contaminants (Rodrigues et al., 2021), but without simulating the actual sediment dynamics. In other forecast systems, morphodynamics is simulated as the last step in the modeling chain (Ferreira et al., 2017), but bathymetry is not updated in the whole forecast model chain. Part of this simplification is due to the complexity of modeling all relevant morphodynamic processes in estuarine and coastal zones, the computational time required for a proper representation of morphodynamic processes and the lack of stability of many of these models (Fortunato et al., 2022CSM bottoms (Nahon et al., 2020, Oliveira (J) et al., 2020, Zhou and Li, 2005).

Morphodynamic processes can play a very significant role in coastal forecast systems continuous quality by providing the means to update the bathymetry after a major hazardous events. According to Zhou and Li (2005), a strong storm may remove sediments from the beach or near-shore environment and depositing overwash fans across back-barrier marshes, lakes, and lagoons. Tropical cyclones may account for up to 90% of coastline retreat and a single typhoon may render a waterway useless in a few hours period, endangering navigation and blocking economic activities. These authors investigated the importance of bathymetry variation in a harbor during a typhoon, and found that erosion and deposition patterns are dependent on the typhoon track.

Coastal erosion early warning systems are disaster risk reduction tools that have been receiving considerable attention recently (Matheen et al., 2021) providing hydrodynamic and bathymetric evolution predictions several days in advance of a predicted storm. These systems typically incorporate coastal numerical models within their workflow to simulate nearshore hydrodynamic and morphodynamic processes, with the goal of predicting the location and magnitude of coastal flooding and erosion along coastlines. Recent examples of these tools are iCoast (Gracia et al., 2014) and Risk-kit (Dongeren et al., 2018) that propose early-warning procedures based on a chain of models that

integrate morphodynamic forecasts. iCoast (Integrated Coastal Alert System) forecast workflow uses GFS, WRF-AWS (in a two-level nesting) and the CALMET model (down to 400 m resolution) for meteorological prediction and the Local Analysis and Prediction System (LAPS) to perform data assimilation of meteorological observations. These atmospheric fields force both circulation model ROMS and wave model SWAN, using CMEMs as boundary conditions. The nearshore hydrodynamics and morphodynamics module is modeled with the wave-averaged model XBEACH in open sandy beaches and model SWASH is used for harbors and pocket beaches. Risk-kit Early Warning System uses a similar approach based on the DELFT-FEWS forecast infrastructure (Dongeren et al., 2018), tailoring the sequence of models to the specific problem at hand (Ferreira et al., 2018, Valchev et al., 2018). One important advantage of Risk-kit EWS is the integration of a Bayesian network tool at the end of the process-based modeling chain, using the results from the last modeling component, to provide probabilistic relations between offshore forcing conditions, local hazard intensity (e.g., erosion and inundation) and impact at the coast (Ferreira et al., 2018).

However, these systems are only triggered for specific events and are not operating in a continuous way, updating the bathymetry in the modeling chain, as part of operational forecasting systems to address this and other hazards. In order to operate in such a way, they require an initial bathymetry of the coastal region before the event, either assimilated through remote sensing data or by typical profiles depending on the season. This discontinuous procedure can lead to major errors and be very ineffective as a management tool.

Integrating sediment and morphodynamics mode in the operational workflows of coastal and estuarine forecast systems to address all emergency and planning challenges is thus the adequate way to address the importance of bathymetric evolution. In order to avoid the accumulation of errors in the forecast chain, assimilation mechanisms based on the available data should be implemented along with tools for morphodynamic model stability such as filters or similar approaches (Fortunato and Oliveira, 2004). Besides satellite data, in-situ cameras have proved to be a valuable tool for topographic and bathymetric features detection (Martins et al., 2022), resorting to AI approaches for better efficiency (Buscombe et al., 2022, Jesus et al., 2022) to be integrated in the assimilation procedures. More details on assimilation in coastal modeling and forecasting are provided below.

3.5 Water quality and biogeochemistry predictions

3.5.1 Overview

Coastal eutrophication and hypoxia, human exposure to waterborne pathogens or radiation, harmful algal blooms, habitat loss and modification, maintenance of the ecological health of coastal systems are some of the issues that require the development of water quality and biogeochemistry prediction systems to support daily to long term management actions and facilitate emergency operations (Kourofalou et al., 2015). These modeling-based systems should include realistic representations of circulation coupled with biogeochemical or ecological models (Rodrigues et al., 2021). The biogeochemical/ecological model components can range in complexity from simple parameterizations

to explicit representations of multiple nutrients and functional groups. These models can be used to three major purposes: (1) assessment of past and current states and trends of the modeled system using hindcasts or reanalyses, (2) multiple scales forecasts ranging from short-term (days to weeks) to seasonal (months) predictions, and (3) scenario simulations including climate change or anthropogenic action projections. In this context, operational systems include both short-term to long-term forecasts (Fennel et al., 2019).

A multitude of open-source numerical models are nowadays available to address water quality and biogeochemistry processes in estuarine and coastal environments, based on numerical representations of the governing equations and/or process parametrization, combining physical, chemical and biological processes modeling to address a specific problem (Rodrigues et al., 2021). Typically, these models use unstructured grids for better representation of complex bathymetry and geometry and also complex vertical exchange processes. SCHISM (Zhang et al., 2016), DELFT-3D (https://www.deltares.nl/en/software/module/d-water-quality/) and FV-COM (Zang et al., 2021) are examples of such modeling systems with forecast implementations that have been applied to algae blooms, oil spills, biogeochemical dynamics and fecal contamination among other uses. Computational costs of these complex models is often the limitation for its application in real problems that require a large number of simulations or a very fine scale in large domains. Model parallelization along with the availability of shared e-infrastructures facilitates this task, but limitations still persist for some applications (Rogeiro et al., 2018).

Given the existing lack of full knowledge of coastal biogeochemistry processes and to facilitate computational usage, particle models are also often used in the prediction workflow to simulate oils spills, larvae movement, water discharges and other processes (e.g. Mateus et al., 2012, Chiu et al., 2018), accounting for advection, diffusion and problem-specific transformation processes. In the last years, particle models have become very popular to simulate microplastics transport and fate, allowing for distinction between different size and characteristics of these contaminants (Domercq et al., 2022). Particle models are also very useful in simulating marine debris pathways (Mcknight et al., 2017). A particle forecast engine, built on top of a circulation forecast, that could handle multiple options and allow for on-the-fly customization by the users, does not exist yet and would be a valuable addition for coastal management actions.

As computational resources availability becomes facilitated and with the onset of satellite data at smaller scales and advances in robotic and low cost sensors, the development of water quality data-based models has been growing. Among data-based models, the major past limitation of AI models in coastal biogeochemistry areas was the scarcity of field data to train and constrain the model under a forecasting framework. Recently, in particular in the field of algae blooms, significant advances have been made and several data-based forecasting systems have been proposed either as stand-alone or combined with numerical models. The Chesapeake Bay Ecological Prediction System (CBEPS) provides short-term predictions of various species and hydrodynamic and biogeochemical variables using the ROMS model with fully mechanistic physical and biogeochemical components. It also includes multivariate empirical habitat suitability models that predict several noxious organisms, three species of HABs and

several water-borne pathogens (Brown et al., 2013). Empirical habitat suitability models for each of the target species were developed by relating the species' presence or abundance to available environmental variables using a suite of modeling approaches, including relatively simple approaches, such as logistic regression, and advanced computing techniques, such as artificial neural networks (ANN). Machine-Learning methods, including ANN, genetic programming, support vector machine (SVM) and Random Forest, have been used with success for algae prediction, ANN standing out as the most accurate (Deng et al., 2021).

For coastal regions, there are fewer applications than for ocean basins. Deng et al. (2021) uses ANN and SVM to predict algae growth trend and magnitude using 30 years of data in Tolo harbor (Hong Kong) with success, using Chl-a concentrations to predict blooms 1-2 weeks ahead. Generalization of these approaches towards the concept of a service for water quality/biogeochemistry/ecology modeling, where the user can set up the dependencies and the type of methodology to apply, is an important avenue for the future in this area.

The accuracy of water quality and biogeochemistry forecasts in coastal zones is also very dependent on the quality of underlying circulation forecasts as well as the accuracy of the contaminant sources specification in space and time. In general, these sources are calculated with water quality drainage system models (either for contaminants or for sources of nitrogen and phosphorus) or discharge outflow models (Rodrigues et al., 2013). For riverine inputs, the best option would be to simulate the whole water cycle from the hydrographic basin to the sea. While some modeling systems already do this for physical processes (Moghimi et al., 2021), no such forecast system exists for water quality/biogeochemistry. Therefore, the simplest existing approaches extrapolate both freshwater water quantity and quality from data (Rodrigues et al., 2013) and the most sophisticated ones use a river basin hydrodynamic forecast system coupled with correlation rules for water quality (Brown et al., 2013).

3.5.2 Data/model integration and the contributions of in-situ and remote sensing

Water quality data plays a major role in forecasting efforts as process knowledge is still far from complete in this area, either in general biogeochemistry, fecal contamination or algae bloom modeling. Many models supporting prediction frameworks are built using correlations only applicable for a specific site, while others are more generic (Rodrigues et a., 2013, Rodrigues et al., 2021). However, all models still require a massive amount of data and detailed calibration and validation to provide reliable real time predictions. Access to real-time or near-real time biogeochemical data at adequate spatial and temporal scales is still limited, the main biogeochemical data stream used in assimilation is satellite ocean color, but this measurement is limited to the surface ocean and provides an imperfect proxy of phytoplankton biomass (Fennel et al., 2019). It is seldomly used at estuarine and coastal scales.

Additionally, established assimilation methods cannot be directly applied to biogeochemical and ecological variables due to the non-Gaussian nature of biogeochemical observations and the frequent lack of direct correspondence between observed data and model variables. Moreover, traditional assimilation methods applied to the ocean circulation part of a coupled hydrodynamic-biogeochemical model tend to lower the accuracy of the biogeochemical variables (Fennel et al., 2019). Some

assimilation methods are being developed to address this issue (Song et al., 2016) based on a 4DVar method, but their usefulness for estuarine environments have not been tested yet. Moreover, the associated computational cost increase needs to be assessed in order to avoid delayed predictions. As the degradation of biogeochemical variables during hydrodynamic assimilation appears to arise when physical and biogeochemical variables are updated independently, accounting for their correlation may be an alternative to reduce this problem.

The onset of low cost sensors and their installation in surface and underwater drones are starting to provide space and time-resolution data that is adequate for high accuracy modeling with assimilation at smaller time scales (Lyhne Christensen et al., 2015). The combination of the multiple sources of data, representing well the small spatial and temporal processes scales of estuarine and coastal environments, and tailored assimilation methods may be an avenue for investigation targeting high resolution couple biogeochemical forecasts.

Some water quality simulations are complex to operate in real time given their dependency on accurate initial conditions, toxic algae bloom modeling being the most critical example. As different species of algae have different triggers for bloom occurrence, although some general conditions are known, forecasting bloom onset is very difficult (Davidson et al., 2016). On-site samples are typically necessary to define whether the species in a bloom is a harmless or a toxic one. Recent developments in remote sensing and AI combination have provided a methodology for harmful algae bloom early detection, for a particular species (https://www.habtrail.com/).

Oil spill forecasting also requires an accurate determination of the initial conditions, for the characteristics of the spilled material (Azevedo, 2010), but that can sometimes be obtained with some accuracy from ship information. Knowing the spilled material characteristics determine the ratio between surface/mixed over depth material and its capacity to evaporate. A way to establish the possible pathways for distinct materials is to perform on-demand simple particle simulations for multiple types of spill characteristics and determine the forecasted affected area (Oliveira et al., 2014), to help deployment of containment devices. Remote sensing plays an important role in conforming forecast simulations, not only for assimilation mechanisms but also to help estimate the often unknown spill characteristics. An innovative methodology was proposed by Chiu et al. (2018) combining a locally-installed nautical X-band radar and a high resolution modeling system (Zhang et at., 2016). The X-band radar processed outputs detect the location of the surface slick which is then used in the spill forecast engine. This analysis used the simple oil particle model of SCHISM but its results could be improved using the module VELA-OIL (Azevedo et al., 2014), which accounts for 3D oil simulation and high-resolution oil beaching and removal. Site sampling is also often necessary to establish local remediation measures or to establish responsibility over a non-reported spill accident.

Quality control is also more important for biogeochemical data than for its physics counterparts, as many more sources of errors are observed due to the complexity of the sensors and their low resilience to biological interferences for instance (Gomes et al., 2015). Human verification is common in monitoring networks for these variables with several levels of quality (Maier et al., 2012). Automatic quality evaluation methodologies such as the one proposed by Jesus et al. (2021b) become trickier to

apply as the volume of data is smaller. The usage of model outputs, also analyzed in this reference, may be a good starting point for a semi-automatic procedure.

3.5.3 Coastal and estuarine observatories and data management

Given the importance of taking the best advantage of data and model information in water quality and biogeochemistry analysis due to the high cost of collecting that data, a huge effort is being made in estuarine and coastal communities to develop tools and platforms for data sharing, following the vision implemented for oceans for a long time (e.g. IOOS, in the USA, and Jerico and EMODnet in Europe are examples of these assets, Kourafalou et al., 2015). Similarly to other types of data, usage of standards and availability of metadata is also fundamental for these areas. These combination of actions allow for the setting up of coherent and rich observatories of water quality and biogeochemistry information, aggregating data and models results (either from scenario analysis or forecast systems - Rodrigues et a., 2021) into a single place and allowing for the calculation of indicators and other products suitable for decision- and policy-making in the socioeconomically vital and often environmentally stressed coastal regions. It also allows for the recovery of historical data of utmost importance for climate change impact studies and to handle access to open and confidential data through authentication mechanisms.

Estuarine and coastal observatories are in general dedicated to one specific site and region, as a one-stop-shop portal for the majority of the relevant information for that site, either from local monitoring networks and processed remote sensing products for the site or from dedicated application of integrated modeling systems that address both physical water quality/biogeochemistry processes. The modeling can be done through a coupled system (Rodrigues et al., 2021, Sanchez-Arcilla et al., 2021) or through a chain of models that address both forcing conditions (e.g. ocean, atmosphere, city outflows, freshwater inputs) and the site processes (Rodrigues et al., 2015). Early applications of this concept are the Land-Margin Ecosystem Research (LMER) sites in the USA (e.g. Boynton et al., 2001) and the LTER worldwide coastal sites (example for the Aveiro lagoon site data: https://deims.org/dfc24538-730e-4e4b-9f04-8e84608b9999). Both programs are targeted at advancing ecology and biogeochemistry knowledge and its relation to physical processes. Physically-based definitions of the concept of residence time, a widely used indicator to assess physical impact on scalars, was developed in the scope of the Columbia River estuary LMER program (Oliveira and Baptista, 1997). Over the last two decades, highly sophisticated observatories dedicated to multiple water quality and biogeochemistry issues and their physical counterparts have been developed and are generally openly available to all users. Examples of these assets are the Saturn observatory, provided by the National Science Foundation CMOP center (http://www.stccmop.org/saturn), the International Centre for Advanced Studies on River-Sea Systems (DANUBIUS-RI), a pan-European distributed research infrastructure supporting interdisciplinary research on River-Sea Systems (https://gis.geoecomar.ro/danubius/) with on-going development at 12 supersites across Europe, the Swan-Canning Estuary Virtual Observatory, in Australia (https://swan.science.uwa.edu.au/index.html) or the S. Francisco bay data center (https://www.sfei.org/sfeidata.htm). At national level, COASTNET includes continuous measurements of environmental and biological variables along the estuarine

gradients of the Mondego, Tagus and Mira estuaries and processed products of the remote sensing data from Copernicus (http://geoportal.coastnet.pt/), UBEST (http://portal-ubest.lnec.pt/ubest/) aims at improving the global understanding of the biogeochemical buffering capacity of estuaries and its susceptibility to future scenarios of anthropogenic inputs and climate change, integrates products based on process-based numerical models and in-situ observations, and is applied to the Tagus estuary and the Ria Formosa, and CONPRAR (http://conprar.lnec.pt/), dedicated to shellfish protection and evaluation of the impact of wastewater discharges on the water quality of coastal lagoon, also applied to the Ria Formosa, that also integrates real time and campaign data and high resolution numerical models. Each of these examples provided access to site information and several processed results targeted at the goal of the observatory. Water quality indicators, as single numbers, or as spatial and temporal variation products are provided to facilitate the analysis of contamination levels, understanding of trends or as support to early warning alert issuing, all for the dedicated sites. Basic information, such as the status of the nodes of the monitoring networks as well as the modeling output production are also provided in dedicated dashboards to facilitate their maintenance and to keep users informed of the update status of the provided products.

A distinct type of biogeochemistry and water quality assets, focused on a service vision and aiming at user-driven, generic application to any estuarine and coastal system, has been developed in the last years. These services started with physical processes provisioning and are now being extended to water quality and biogeochemistry. OPENCoastS^{+} is an example of a generic platform for on-demand forecasting of faecal contamination and hydrodynamics for any estuarine and coastal site in the world (https://opencoasts.ncg.ingrid.pt/). Aquaculture monitor, developed in the scope of the Blue-Cloud project (https://blue-cloud.d4science.org/web/aquacultureatlasgeneration) is another example, aiming at producing national aquaculture sector overviews allowing a country to use OGC compliant data services to monitor its aquaculture sector and take advantage of interoperable services where teams can compute and publish reproducible experiments.

This new vision to address water quality and biogeochemistry problems has still major challenges to be addressed to include both process-based and AI models, uncertainty and data assimilation, and to make the outcomes available to the users in a usability mode that allows for user-customized products. The capacity to build complex forecast workflows, picking the tools most appropriate for a specific problem, and on-demand processing methodologies is not available yet. Moreover, their availability through a user-friendly, guided platform, such as OPENCoastS, still requires the development of core components only achievable through multi-disciplinary teams that include coastal scientists and engineers and computer science experts.

3.6 Forecast quality and uncertainty evaluation

3.6.1 Role of (near)real time data and forecast error quantification

Over the last decade, the online sharing of data and the building of large data repositories have become common and are essential for independent evaluation of forecast system results and to setup the initial

conditions at the initial establishment of the forecast deployment. EMODnet (https://map.emodnet-physics.eu/), the IOC Sea Level Station Monitoring Facility (http://www.ioc-sealevelmonitoring.org/map.php), the GOOS observation and data portal - OceanOPS (https://www.ocean-ops.org/board), among others, are examples of such resources, and provide automatic ways for data to be fetched and products to be build within forecast interfaces to provide users with results quality procedures. Based on those global or even regional repositories, most coastal and ocean forecast systems provide some near-real time, on-the-fly model data/model evaluation, from simple time series comparison with online stations (e.g. Mateus et a., 2012) to sophisticated interfaces that allow for interactions between the data and the user and the access to both in-situ and remote sensing data (e.g. GLOFFIS and GLOSSIS from Deltares, or OPENCoastS from LNEC).

For estuarine and coastal forecasting, data is also necessary from rivers and other freshwater data sources, both for model forcing and to establish the adequate modeling approach depending on environmental conditions. If forecasts go beyond physical processes, freshwater quality data is also fundamental to establish for instance outflow discharges.

The multiple sources of real or near-real time in-situ and remote sensing data are also a fundamental component of forecast evaluation and reevaluation as new processes or seasonal behavior emerge from data analysis that were not solved in previous deployments (Kourafalou et al., 2015). Several error measures have been proposed throughout the last decades, targeting at evaluating distinct characteristics of model forecasts. The average of simulated and observed values (BIAS), the root mean square error (RMSE) or Nash Sutcliffe Efficiency (NSE) are popular choices (Ferrarin et al., 2019).

However, a detailed evaluation of a forecast system goes beyond simple, averaged quantities, as accuracy is not the only factor for a reliable, robust and spatially and temporally accurate prediction system. With a focus on forecast evaluation from an end-user perspective and targeting the development of adequate tools for forecast-based financing of populations at risk, Hagen et al (2020) proposed a model suitability matrix. It considers needs at end-user level through quantitative score assignment on several criteria using a decision tree: data, software, computational efficiency, flexibility, requirements of technical expertise, forecast skill and uncertainty. A holistic and flexible framework for model forecast evaluation is presented in this work, targeted at the concession of forecast-based funding.

The metrics suite for a detailed evaluation encompasses average indicators such as Nash Sutcliffe Efficiency (NSE), root-mean squared error (RMSE), normalized goodness-of-fit statistics, along with some detailed calculations such as the Kling Gupta Efficiency (KGE), a decomposition of the NSE to correlation, bias and variability, absolute value error statistics and the Index of Agreement (AINDEX), which states the ratio of mean squared error to potential error. It also includes non average metrics such as the hit rate (HR) and the false alarm rate (FAR), of huge importance in early-warning systems. This combination of average-based metrics and HR/FAR can be used for evaluation of forecast skill; firstly, to constrain predicted flows to the range of observed flows, and secondly, to ensure that the model differentiates flows above and below the forecast threshold.

For probabilistic forecasts, discussed next, the Ensemble Verification System (EVS) using a suite of metrics to assess resolution, reliability and discrimination of forecast probabilities is proposed, includingthe Brier skill score (BSS), the mean continuous ranked probability skill score (CRPSS) and the relative operating characteristic score (ROCS). These evaluators evaluate the ensemble skill (and not just the ensemble spread) with reference to the sample climatology. To assess the ensemble spread, rank histograms were proposed.

After assessing forecast skill, other issues are evaluated such as complexity, flexibility and data – and code availability. The model suitability matrix for forecast-based financing encompasses the following generic steps: 1. Define criteria of interest; 2. Select metrics and suitability thresholds for quantitative score assignment; 3. Embed suitability thresholds in decision tree for transparency; 4. Select forecast lead time and use a decision tree to consistently assign scores; and 5. Normalize scores in the suitability matrix and display in radar charts. These items are defined as:

- ID Criteria of interest description

- Data: The degree to which data used in model setup is available and free of charge regardless of geographical location. Data availability is fundamental not only for forecast evaluation but also to its setup or building of data-based models.

- Software: The degree to which open-source code comprises the model structure. While closed-source models are becoming the exception, applications in some low income countries with weak technical staff still rely on models whose characteristics cannot be fully analyzed and assessed;

- Computational efficiency: The time required to generate forecasts relative to the forecast lead-time. As complexity increases, the time for predictions availability also increase. A balance is necessary between quality and usefulness and new approaches based on AI-based or hybrid models are becoming attractive for problems requiring very fast answers or without access to proper computational resources.

- Flexibility: The degree to which the model can adapt to catchment changes and incorporate observations through data assimilation. Most estuarine and coastal forecast systems are still limited to address river-to-ocean scales and do not include assimilation.

- Forecast skill: the application of the rather detailed suite of evaluators is fundamental for a proper skill evaluation.

- Uncertainty: the main sources for uncertainty are listed and they should be taken into account in forecast operations and clearly displayed in model outputs.

Sanchez-Arcilla et al (2021) proposes a different set of forecast error evaluation indicators, targeting both ensemble and stand-alone forecasts and the assimilation with in-situ and remote sensing. Robust error metrics, such as Mean Absolute Error (MAE), Random Error Reduction (RER) and Continuous Rank Probability Score (CRPS) are proposed to determine forecast skill and long-term model evolution. RER and MAE are simple metrics as requested by CMEMS product verification, while CRPS score is suggested for ensemble forecasts for comparison with RER or MAE for deterministic forecasts.

Citizen science data is also starting to become an asset in validating and introducing on-the-fly corrections to forecast predictions, allowing for an improved prediction in the future (Loftis et al., 2019). While still limited to flood validation, the emergence of several mobile applications and citizen science flood mapping projects such as King Tides (kingtides.net) for flood reporting are promoting an involvement of local communities to help improve forecasting models. The concept can be enlarged to other phenomena where web-shared photos can improve predictions such as contamination events of algae blooms.

3.6.2 Uncertainty

Uncertainty and inaccuracies in forecast systems are strongly related with the uncertainty and errors of the underlying models in the cascading forecast procedure (or in the single model for simpler forecasts). The sources of uncertainty in operational numerical models are associated with process representation and its physical and numerical parametrization, scale representation (horizontal and vertical spatial resolutions) and the boundary conditions prediction (Feng et al., 2011, Iglesias et al., 2022).

Given the importance of accurate predictions for most forecast system applications, care is necessary to address uncertainty and other error sources. For instance, for search and rescue operations, forecast errors can lead to fatalities or a huge resource allocation to avoid them. Given the complexity of coastal dynamics, errors will vary with forecast durations: for short times, the unresolved processes are the major limitations, while as time goes by, the variability of coastal currents leads to a growing forecast error (Révelard et al., 2021). Uncertainties can propagate and be amplified through the model execution, potentially leading to a large forecast error (Mohan Das et al., 2017). The need for accurate forecasts, where errors of this type are minimized, highlights the importance of finding and implementing new solutions to deliver better numerical results. Traditionally several models of the same system are used when available to determine a range of variability (Révelard et al., 2021), often complemented with assimilation of in-situ data from drifters or HFR radars. More recently and by inspiration from atmospheric modeling, ensemble modeling is starting to be considered as one of the best solutions because it can minimize the combined uncertainty in input data, model parameters and model structure, improving the final performance of the predictions (Iglesias et al., 2022).

Next to the natural randomness associated with the temporal and spatial variation of natural processes, there are other sources of numerical model uncertainties. They can result from model initialization, due to, for example, incomplete required data, measurement errors or errors in data processing, or reside in the model itself, due to imprecise parametrization of physical processes, inaccuracies in the definition of parameters, the existence of unresolved scales, or errors in the boundary conditions (Buizza et al., 2005; Weigel et al., 2008; Tebaldi and Knutti, 2007). Furthermore, it must not be forgotten that any attempt to represent nature in a set of equations will always be a simplification of reality that inevitably introduces uncertainties. Although the partial differential equations that govern estuarine hydrodynamic phenomena are well known, they must be truncated to a finite-dimensional set of linear equations to be computationally integrated and solved. The uncertainties associated with this process can propagate and produce effects across the entire spectrum of scales predicted by the model (Palmer et al., 2004;

c). To produce a perfect forecast, a sound knowledge of each process and state of the system would be needed, and this should be perfectly implemented in the model. However, this can only be done in an approximate way, as even the actual quantification of the system is accompanied by inherent uncertainties. All uncertainties, even minor ones, propagate through the model processes, adding up and eventually leading to a large forecast error (Mohan Das et al., 2017). In forecast systems where a chain of models is operated, uncertainties are propagated and amplified from each level to the next one (Rodrigues et al., 2021). The need for accurate forecasts, where errors of this type are minimized, highlights the importance of finding and implementing new solutions to deliver better numerical results. In this context, ensemble modeling is considered as one of the best solutions because it can minimize the combined uncertainty in input data, model parameters and model structure, improving the performance of the models (Viney et al., 2005; Suh et al., 2012; Mohan Das et al., 2017).

Multiple models and the construction of a weighted ensemble, giving more weight for better performing models (Umgiesser et al., 2020), is being used at regional sea level at the Baltic operational oceanographic system (http://www.boos.org/multi-model-ensemble-of-forecast-products/) and could be adapted for coastal and estuarine scales. This application is being used for national warning systems for storm surges, generally with an accuracy of 10-20 cm for 1-20 year events (Umgiesser et al., 2020). An alternative to multi-simulation approaches is dress ensembling (Mel and Lionello, 2016), shown to be computationally cheap and providing reasonably realistic results. It replaces the explicit computation of uncertainty by ensemble forecast with an empirical estimate of the intended variable, by assigning an error distribution form, which includes a dependence of the uncertainty on surge level and lead time and of the uncertainty of the meteorological forcing, on the other hand.

Uncertainty consideration and forecast error assessment are still at a research level for estuarine and coastal forecast frameworks. Indeed, integration of uncertainty in these frameworks through simultaneous consideration of ensemble predictions in external forcings (either river or atmospheric), multiple model configurations and multiple models usage and the corresponding ensemble are the next steps. The availability of large computational resources at e-infrastructures and the free availability of the vast majority of models has paved the way to overcome this limitation.

3.7 River-to-ocean integrated approaches: integrating the whole hydrographic basin in the forecast framework

The recent developments in modeling systems accuracy, data availability and its global coverage, and strong and reliable computational resources have created the conditions to build river-to-ocean integrated forecasts. This challenge entails solving the relevant processes at each water compartment and to handle interactions between compartments, and a harmonized handling of spatial and temporal scales. The importance of this integrated approach has been broadly recognized (e.g. Dekleermaeker et al, 2017, Stanev et al., 2016, Ferrarin et al., 2019). Indeed, integrated coastal forecasting systems should appropriately account for the dominant coastal processes at wide characteristic scales, accounting for the exchanges along and across the shelf break, storm surges, tides, internal waves, surface waves, fronts, slope currents, estuarine processes, river plumes and suspended sediment

dynamics (Stanev et al., 2016). Close to the coast, biogeochemical processes show large diversity and strong gradients, as a consequence with the interactions with land and inland waters, requiring high confidence and timeliness predictions. The challenge is to use multidisciplinary and multiscale observations and modeling suites that represent the individual subsystems: atmosphere, waves, circulation and biogeochemistry. These coupled systems differ considerably from uncoupled modeling approaches (e.g. Rodrigues et al., 2013), where information is shared in the two directions but on different time scales. Errors arise from this approach as interactions may determine the dynamics of each compartment.

This goal can be achieved through the implementation of numerical models based on a unique unstructured grid able to describe processes at different spatial scales (Zhang et al. 2016; Stanev et al. 2016), or through nesting of models (structured and unstructured) at different resolution (Kourafalou et al. 2015; Trotta et al. 2016, 2017; Fortunato et al. 2017). Indeed, unstructured grid (UG) models, such as SCHISM (Zhang et al., 2016), SHYFEM (Umgiesser et al., 2004) and ADCIRC (Luettich et al., 1992), are fundamental tools to perform cross-scale modeling and to develop operational implementations that downscale from regional processes down to estuaries, rivers and coastlines. UG-based forecast systems have been developed in the last decade with success, reaching spatial resolutions on the order of a few meters, thus being able to accurately tackle complex coastlines and wave and current interactions (Lavaud et al., 2020). Structured grid models have also been used in an operational context, resorting to multiple level of grid nesting to better address complex coastlines (Mateus et al., 2012).

Some works are starting to include rivers and creeks in the estuarine and coastal modeling procedure, through the use of compound modeling strategies, either using statistical (Ganguli and Merz, 2019) or process-based modeling (Ye et al., 2020, Moghimi et al., 2021). Other approaches (e.g. Dekleermaeker et al, 2017) use the same forecast infrastructure to handle river forecasts (such as GLOFFIS) and coastal forecasts (such as GLOSSIS). The Delft-FEWS infrastructure is an example of such an approach through the GLOFFIS and GLOSSIS forecast systems.

Taking advantage of the availability of meteorological forecasts, global scale precipitation products and global scale hydrologic and hydrodynamic models, forecast systems such as GLOSSIS, GLOFFIS or WIFF (Fortunato et al., 2017) and others can be applicable at global scale. However, an intertwined, integrated operation with spatial and time scales compatibility from river-to-ocean processes, accounting for the urban dimension, has not been reported in the literature. The integration of multi-platform observatories with models that resolve river, estuarine and coastal dynamics is clearly a key feature of successful operational systems. Such integrated systems must be linked to larger scale systems toward the achievement of seamless data sets, nowcasts and forecasts across the river-to-ocean scales.

3.8 Open, on-demand, automatic forecast and pre-operacional modeling frameworks and platforms for management, research and recreation uses

3.8.1 Overview

Open data is a global trend nowadays from single sensors to large monitoring networks. In Europe, open data is part of the Digital strategy and the Green deal, strengthening the recognition that the economy and mankind as a whole benefit greatly from sharing information openly. The concept of open forecast data has been gaining importance in forecast systems in operation throughout the world, whether they address global, regional or local dynamics. While there are many systems in operation (a large number of examples are listed here, although some of them are no longer active, https://www.godae-oceanview.org/science/task-teams/coastal-ocean-and-shelf-seas-tt/coss-tt-system-information-table/), most systems just publish images without providing access to quantified predictions that could in turn be used to produce new services and thus have a greater societal impact. The initiatives under the UN Ocean Decade such as the DITTO (https://ditto-oceandecade.org/) or CoastPredict (https://www.coastpredict.org/) projects aim at opening forecast information to all and address dynamics from the global to the coastal dimension. Estuaries and the interaction between land and the ocean are external to these initiatives. Herein, several examples of open data forecast systems are reviewed, covering from global to regional and local geographical settings.

Setting up model or forecast systems are complex tasks that require considerable expertise of coastal dynamics, numerical modeling and computer science. In the last few years, several initiatives have emerged to provide simplified ways to address this challenge and provide user-friendly tools to set up models and their forecast systems, with linkage to global or regional forcings and access to data comparison in near real time. These platforms aim at expanding the application of forecast systems worldwide, by overcoming current barriers. Below the main platforms are also summarized, emphasizing their characteristics and range of applicability.

3.8.2 DELFT-FEWS and its platforms for flood forecasting GLOSSIS and GLOFFIS – river to ocean forecast and generic data handling platform

Delft-FEWS (Delft-Flood Early Warning System, Werner et al., 2013) is an open data handling platform initially developed as a flood forecasting and warning system. Nowadays it is a vast collection of modules designed for building a water forecast system customized to the specific requirements of a user, and range from basin to ocean applications. Delft-FEWS is open usage software, but it is not open source. It can be used as a stand-alone application at the user's computer or in centralized IT facilities.

One of the key features of the system is its flexibility in integrating (third-party) models and data, producing forecasts on a daily basis. Delft-FEWS can integrate third-party models through available connectors, or by building new ones through collaborations with Deltares. Rather than integrating models at the algorithm level (such as the OpenMI interface standard), models are integrated in Delft-

FEWS as an external process. Delft-FEWS provides the required input data and parameters, executes the model, and reads the results. For coastal and estuarine applications, the Delft3D-FM is one of the available choices. The users can also select available forcings and real time data for setting a new forecast system, through interaction with the Delft-FEWS application interface. On-the-fly qualitative comparison with in-situ and satellite data is available (Dekleermaeker et al, 2017).

Delft-FEWS is also a data-centric platform, with an extensive library of data processing functions, to address data quality control, serial and spatial interpolation, aggregation and disaggregation, and merging data (Werner et al., 2013).

Inundation forecasts are one of the major applications of forecast systems in water domains, Deltares being one of the most active institutions to address this challenge. Inundation occurs in the coastal and estuary regions as well as in the rivers and in cities, with distinct sources of flood events. Deltares has adopted an uncoupled approach to handle flooding, through two global flood forecasting systems: Global Flood Forecasting Information System (GLOFFIS) and Global Storm Surge Information System (GLOSSIS), based on Delft-FEWS, an open software platform that integrates data and models (Dekleermaeker et al, 2017). FEWS has many applications in riverine systems across the globe and a few storm surge ones. Access to forecast time series for several variables from GLOSSIS and GLOFFIS is freely available at https://blueearthdata.org/ for the whole globe. Spatial outputs are not available yet.

3.8.3 Science and Technology University Research Network Collaboratory - SATURN

Saturn is a river-to-shelf Columbia river collaboratory that integrates an Observation Network, with physical and biogeochemical data from multiple stations, and the Virtual Columbia River, a skill-assessed modeling environment with daily forecasts, simulation databases and scenario simulations. SATURN's virtual environment and observation network are integrated through a computational infrastructure with a web interface that provides convenient real-time access to both model and observation data products (Baptista et al., 2015). The Virtual Columbia River aims to provide an ongoing prediction of the estuary dynamics, including the freshwater and ocean contributions, accounting for multiple processes, disciplines and space and time scales. It includes forecasts, decadal historical simulations and climate change or anthropogenic future scenarios databases, a Climatological Atlas and targeted simulations of processes or scenarios, using multiple models.

Saturn was originally developed by the NSF Center for Coastal Margin Observation & Prediction (CMOP) and is now maintained by The Columbia River Inter-Tribal Fish Commission. This is a one-of-a-kind transition of the operation of a scientific product to the end-users, following the vision of CMOP of Saturn as a 'forever resource' for the Columbia river region. The tribal community is one of the communities of practice, individuals and groups who developed their activity using the information and insight generated by Saturn's outputs.

Besides sharing openly all outputs of models and data at its several levels of quality assessment, CMOP also opened several of the Saturn software components to all. Furthermore, several interactive tools

are available to explore data and forecast information such as the Data Explorer, for Web analysis of observational river and ocean data, and Watches, for tracking specific features including estuarine hypoxia and estuarine turbidity maxima. The integration of these services in an on-demand platform such as OPENCoastS (described below) would foster the capacity of coastal managers to support their decisions based on these data-rich environments.

3.8.4 SOCIB - the new Balearic Islands Coastal Ocean Observing and Forecasting System

SOCIB is a multi-platform integrated prediction system that provides free access to oceanographic data and modeling services to support operational oceanography in the Mediterranean Sea, focused on the Balearic Islands (Tintoré et al., 2019). It is composed of three major subsystems: an observing sub-system, a forecasting sub-system and a data management sub-system. Predictions are provided for oceanographic parameters based on ROMS and SWAN high resolution simulations and model outputs are provided for both human users and M2M protocols. Data is integrated through assimilation of satellite, Argo floats and Ibiza Channel HF radar. These predictions are then used in the development of techniques for predicting larval dynamics in the Balearic Sea (Tintoré et al., 2019).

SOCIB stands out in the operational oceanography arena for its comprehensive compliance with FAIR principles, promoting its reuse for other users. The existence of a Digital Object Identifier in most data sets is for instance a very important asset that could be replicated in other systems.

3.8.5 Generic infrastructures for setting up and operating models before operational implementation –WebMARVL

The Web-based MARine Virtual Laboratory (WebMARVL, Oke et al., 2015, http://www.marvl.org.au/) is an interface to establish models, allowing the user to quickly and easily configure all steps, from grid generation, setup of boundary conditions and model selection. WebMARVL can be applied to ocean circulation or wave modeling through a simple interface, after which a series of processing scripts are run to build a bundle for simulation usage. Currently, MARVL can be used to:

- Configure several open, community coastal/ocean and wave models for any region, for any time period, with model specifications of their choice, through a user-friendly web application,
- Facilitated access to data sets for model initial and boundary conditions;
- Access and assemble ocean observations from for model evaluation or data assimilation.

Through this interface any user can prepare a simulation environment for future use in a forecast system. Usage of WebMARVL portal is open to anyone in the Australian community, but the source code is open and available for installation at the user selected infrastructure to all (https://code.google.com/archive/p/marine-virtual-laboratory-information-system/). Several features in WebMARVL can complement the open platforms described in the next section and the framework proposed in the studies section, in particular in the grid generation task.

3.8.6 RELOcatable ocean nowcast/forecast system - RELO

RELO is the USA's Naval Research Laboratory (NRL) forecast platform, based on the Navy Coastal Ocean Model (NCOM) and the Navy Coupled Ocean Data Assimilation (NCODA) system (Rowley and Mask, 2014). It provides a capability for relocatable ocean forecast modeling and data assimilation that is used for operational forecast support for USA's military operations. The concept of relocation is a modeling or forecast system that can be moved to different geographic regions or scales easily, with minimal changes in domain configuration.

RELO is dedicated to large scale modeling, without the capacity to solve the spatial scales required by nearshore processes. To overcome this limitation, RELO was coupled with the Advanced Circulation Model for Shelves, Coastal Seas, and Estuaries (ADCIRC) for simulations in coastal, estuarine and shallow regions (Chu et al., 2009).

Both structured and unstructured versions of RELO are closed software code, and most of its results are available through a closed access NAVOCEAN. The RELO system consists of a forecast model component, a data analysis component and supporting codes and scripts. As all platforms in this section, they are component based IT infrastructures that automatize the routine procedures of a forecast system.

3.8.7 Structured and unstructured infrastructure for ocean to coast modeling forecast – SURF

The Structured and Unstructured Relocatable Ocean Model for Forecasting (SURF) provides a mechanism to deploy a nested high-resolution numerical model forced by larger-scale ocean forecasts (Trotta et al., 2021). SURF is a fully baroclinic hydrodynamic model downscaling tool and integrates two circulation models, SHYFEM for coastal and estuarine application using unstructured grids (Umgiesser et al., 2004) and NEMO for open ocean and shelf applications based on finite difference and structured grids (Madec, 2008).

The SURF infrastructure follows a decade-long research on the concept of relocatable forecasts, based on the application of nested regular grid applications. Unlike the previous efforts, SURF allows forcing from both low resolution ocean models analysis and reanalysis and thus is particularly suited to address risk and other long term studies. This chain of nested models can reach spatial scales on the order of a few hundred meters (Trotta et al., 2021), being thus appropriate for coastal scales but not to near shore or estuarine forecasting such as coastal beach inundation and overtopping or marsh and estuarine beach flooding is not accounted for.

This software package is open source and can be deployed at a user selected infrastructure, through a Virtual Machine Environment (https://www.surf-platform.org/). Currently it allows for generic application of the NEMO model forced by CMEMS forecasts, handling the creation of inputs, model operation and outputs processing. The integration of SHYFEM is underway as well as the Web interface for visualization of the results.

3.8.8 Open forecasts to all – the OPENCoastS initiative

Coastal forecast systems now have a solid role in coastal management procedures for all entities with responsibilities or economic activities in these areas. Their setup, usage, and continuous maintenance are, however, limited by technological evolution constraints and personnel cost, thus, preventing their broad use for many systems.

The OPENCoastS framework, freely available to all at https://opencoasts.ncg.ingrid.pt/, was designed to address these challenges in a comprehensive way (Oliveira et al., 2020). OPENCoastS is a computational service designed to build circulation forecast systems for any coastal system in the world. To facilitate its use by users that are not modeling experts, it requires the minimum information from the user and is supported by a simple workflow of 7-8 steps in a WebGIS interface. All the data required at each step is clearly identified and often available as a dropdown menu. For the setup of each simulation, sets of pre-defined modeling parameters are provided, retaining the ability for most experienced users to define their own values.

OPENCoastS aims at making operational systems open to every user and promoting that every coastal system should have a forecast implementation, anchored on the availability of relevant computational resources at the European Open Science Cloud (EOSC). This user-centered approach started with simple 2D circulation physics in 2018 and has attracted over 500 users from over 60 countries in the last 3 years. From the simple 2D tidal dynamics in 2018, OPENCoastS has grown to become a sophisticated tool that simulates both 2D and 3D physics, including circulation and water quality, and accounting for wave and current interaction. The simple way to build, manage, and visualize results in a WebGIS platform is considered a major advantage by the users and has promoted the application of OPENCoastS in estuarine and coastal systems worldwide (Oliveira et al., 2021). In parallel with the many options for physical and water quality processes, OPENCoastS is linked to EMODNETPhysics and processed Sentinel data from on-the-fly data comparison and allows for several options for atmospheric and oceanographic forcings. Land and freshwater inputs can be introduced in the forecasts through climatology or by linking with existing drainage or river forecast systems.

The concept of OPENCoastS is deeply rooted in an open science vision. Besides the open access to the platform (only a registration is needed), an open grid repository was also built, as the availability of a computational grid could be a limiting factor for its use (https://github.com/LNEC-GTI/OPENCoastS-Grids/tree/master/docs/Portugal). While the complexity of OPENCoastS simulations have benefited from the availability of EOSC resources, the users can select a local implementation of the framework through installation at their computational infrastructure of choice of the frontend and backend software openly available at https://gitlab.com/opencoasts/eosc-hub.

While the openness and flexibility of OPENCoastS have contributed decisively to the creation of many forecast systems worldwide, the availability of a computational grid is still a considerable limiting factor to many users. Many applications are available to generate these grids (e.g. SMS, Gmsh, ADMESH, OceanMesh2D), but they require the knowledge to build it and do not operate in an automatic way, using only geometric information of the area of interest. Recently, new tools have been developed for

the automatic generation of unstructured grids (e.g OCSMesh - https://github.com/noaa-ocs-modeling/OCSMesh, Mani, 2021, Soroosh et al., 2021) which can be integrated within OPENCoastS workflow.

Data assimilation and uncertainty are also missing in OPENCoastS's current version. Data sources are integrated in OPENCoastS only for fast data/prediction comparisons, using EMODNet Physics water elevation data and Sentinel images after processing (Oliveira et al., 2021). Improvement of circulation predictions could be achieved by assimilation of the multiple inundation lines obtained using for instance the Sentinel-based WORSICA service (worsica.lnec.pt) after integration with the topo-bathymetric extractor of Fassoni et al. (2021). Likewise, assimilation with in-situ cameras information could also improve local deployments. In addition to spatial information, the frequent time sampling of cameras is now being explored for coastal information extraction (Martins et al., 2022, Jesus et al., 2022).

Data assimilation for water quality simulations is also of relevance in particular for phenomena that are difficult to predict such as the conditions for harmful algae bloom appearance. Recent developments on satellite imagery processing (such as HABTrail) were able to identify the triggering factors that appear on HAB generation. Linking these algorithms to OPENCoastS are an avenue to explore and build an on-demand application that can help decision makers to react quickly at the beginning of the bloom thus avoiding danger to the population using the affected area or for water-borne industries such as aquaculture or fisheries.

The integration of uncertainty, in particular associated with atmospheric forcings, could also be included taking advantage of the multiple sources available within OPENCoastS. Automatic error calculation, for the several options or weighted ensembles, could be offered along with the identification of the most reliable choice.

3.9 Open problems and challenges for the future

In the review above many issues were raised that are worth investigating in the future. Herein the applicant selected those tasks that contribute towards a richer contribution of modeling and forecasting tools towards the development of coastal digital twins, addressed in the next chapter and the main goal of this research plan.

The selected avenues for future research are the following:

- The integration of multi-platform observatories with models to forecast river, estuarine and coastal dynamics, including city water cycles, is clearly a key feature of successful, comprehensive operational systems. However, intertwined, integrated operation of forecast systems with spatial and time scales compatibility from river-to-ocean processes, accounting for the urban dimension, has not been reported in the literature. Such integrated systems must be linked to larger scale ocean and atmospheric systems toward the achievement of seamless data sets, nowcasts and forecasts across the river-to-ocean scales. Building an infrastructure that predicts watershed to ocean dynamics and integrating it with data through the concept of collaboratories is a challenging theme to address in future research.

- Huge computational resources requirements for river to ocean forecast frameworks, in particular for integrated hydrodynamics and water quality forecast, can be solved through integration in e-infrastructures. Unfortunately, it is not a turn-key procedure and integration of platforms and frameworks require adaptation to these resources (e.g. integration in a virtual research environment - VRE) and exploitation with multiple core services such as authentication, orchestration (to take advantage of multiple resources provides) or simply to use compute services.
- Current forecast systems offer little or no flexibility in forecast setting up. Developing a forecast framework as a service where users can choose which dimensions/processes and their interactions to include, is an important contribution towards user-empowered forecast systems. The concept of OPENCoastS for broader process coverage and further flexibility in choice of process and their interactions, as well as modeling choices, is also an avenue of research for the future.
- Merging/fusing process-based numerical models and AI data-based forecasts can be a way to improve forecast results. Therefore, the joint application of both AI and physical process models or its combination into hybrid forecasting considering also data assimilation appears to be an attractive methodology to explore.
- integration of uncertainty in forecast frameworks either through consideration of ensemble predictions in external forcings (either river or atmospheric), multiple model configurations or multiple models usage are one of the most relevant issues to be addressed. Ensemble modeling, based on multiple models or several configurations of models through weighted contributions, should be supported by on-going forecast evaluation of accuracy and timeliness. Automatic identification of best performance based on error calculation against real time data networks would be one of the products with particular interest for managers.
- Acute changes in bathymetry after hazardous events compromise the quality of the predictions afterwards. Integrating morphodynamics in the forecast modeling chain can provide the necessary update of bathymetry in all chain components. To avoid instability, data assimilation in particular based on remote sensing (satellites and cameras) should be part of this integration.
- The combination of the multiple sources of data, representing well the small spatial and temporal processes scales of estuarine and coastal environments, and tailored assimilation methods may be an avenue for investigation targeting high resolution couple biogeochemical forecasts. For instance, integrating remote sensing data assimilation in water quality predictions could improve the usefulness of on-demand forecast systems for early detection of emergency events such as algae blooms. Recent advances in both areas have set the stage for this integration and creation of an accurate early-warning, on-demand tool for algae blooms and other hazardous events.
- Similarly, the development of a particle forecast engine, built on top of a circulation forecast, that could handle multiple options and allow for customization on-the-fly by the users, would be a valuable addition for coastal management actions.

- Data quality assessment at biogeochemical and water quality predictions is still at its infancy, with manual assessment that is not compatible with real time requirements. Automatic quality evaluation methodologies such as the one proposed by Jesus et al. (2021a,b) may be an avenue for research and the usage of model outputs, also analyzed in this reference, may be a good starting point for a semi-automatic procedure.

4 State of the art: Coastal Digital Twins

4.1 Introduction

In this chapter the concept and capabilities of digital twins and several methodologies for its implementation are reviewed, along with some examples in selected areas. Existing applications to the water domains are also presented, with a focus on estuarine and coastal uses.

The terminology “Digital Twin” was first used in Hernandez and Hernandez (1997), but the first works on the full scope of the concept (“digital equivalent to a physical asset”) were indeed proposed later as product avatar (Hribernik et al., 2006) and in the scope of the emergence of coastal observatories (Baptista et al., 2008). Most of the earliest works using this concept came from the manufacturing area, where Digital Twin was defined as a methodology to simulate the physical components of a production system (Liu et al., 2021). In the last two decades, the concept of Digital Twins has evolved considerably and has been used in many contexts and applications with some differences in its definition and scope (see Table 5 in Liu et al., 2021). Nowadays, Digital Twin integrates continuously, in an interactive, two-way data connection, the physical and the virtual assets. They provide a virtual representation of a physical asset enabled through data and models and can be used for multiple applications such as real-time forecast, system optimization, monitoring and controlling, and support enhanced decision making (Rasheed et al., 2020). Recent tools take advantage of the huge online volume of data streams provided by satellites, IoT sensing and many real time surveillance platforms, the availability of powerful computational resources that made process-solving high resolution models or AI-based models possible, to build high accuracy replicas of the real world (Barbie et al., 2021).

4.2 Digital twin applications and their implementation in the water domains

Existing applications of Digital Twin tools span many themes, from manufacturing (Liu et al., 2021) to space exploration (Shafto et al., 2010). Each area has its adapted definition of digital twins, depending on the characteristics of the assets being virtualized. Focus may address in particular issues like dynamic, bidireccional integration and simulation, real time control and optimization, and virtual and/or dynamic replicas. For instance, NASA’s digital twin is defined as “an integrated multi-physics, multi-scale, probabilistic simulation of a vehicle or system that uses the best available physical models, sensor updates, fleet history, etc., to mirror the life of its flying twin. The digital twin integrates sensor data from the vehicle’s on-board integrated vehicle health management system, maintenance history, and all available historical/fleet data obtained using data mining and text mining”. Application of DT in this context includes recommendations of changes in the vehicule mission aiming at mitigating damage or degradation to increase both the duration and success of the mission (Shafto et al., 2010). An application of Digital Twins to disaster management (Fan et al., 2021) aims at improving disaster

response by integrating information and communication technology tools and artificial intelligence approaches with social science knowledge. AI methods are fundamental to address speech recognition, event extraction and image processing, along with data reliability and processing while ICT tools address information sharing and communication. The multimodal data hub comprises remote sensing (from Satellites and UAVs, to address building and infrastructure destruction, storm surge water level and land occupation during and after events), social sensing (for power outage, contamination of water and other infrastructure failures as well as monitoring people's reactions during an event) and crowdsourcing (for data collection and analysis, bringing in the agents reporting in the field during events).

Digital Twins and forecast systems are distinct, the latter being an integral part of the former. Digital Twins are user-centered, having user interaction embedded in their design, including visualization, user-driven data transformation and processing, and user selection of tools to produce a specific output. On the contrary, forecast systems are typically product centered, aiming at providing pre-defined services to users. OPENCoastS, a framework for on-demand coastal circulation and water quality forecasts (Oliveira et al., 2020, Oliveira et al., 2021), is a step closer to the concept of Digital Twins when compared to traditional forecast systems, but it still lacks the level of user interaction and data integration that is at the core of DT.

Digital Twins are composed of several components, defined according to the problem being addressed. For instance, Zhang et al (2019) proposes a framework with 4 major layers, for the manufacturing area: Physical layer, the Model layer, the data layer and the Service layer that promotes interactions between the previous two. Interaction with users are embedded in several of these steps, unlike the next examples from other areas. A framework for disaster management is proposed in Fan et al. (2021), containing 4 components: 1) multi-data sensing for data collection, (2) data integration and analytics, (3) multi-actor game-theoretic decision making, and (4) dynamic network analysis. Given the emergency nature of this DT, interaction with actors/users is much stronger than in the previous example, being present at data collection, decision making and network analysis.

A final example of a distinct structure is the Digital Twin of the Ocean proposed by UN decade project DITTO (https://ditto-oceandecade.org/). In the scope of the UN Ocean Decade, two projects were proposed that deal with digital twins and its forecast engines: the DITTO project and the CoastPredict project. DITTO's DTO is organized around 5 components: 1) Observing system; 2) Data Space; 3) Data Analytics and Prediction Engine; 4) Interactive and Provisioning layer; and 5) Outreach and Training. This framework is by far the most sophisticated, providing several ways for users to interact and tailor the usage of the DT to their needs, either by tailoring the data analysis, selecting models for prediction engines (both process-based or AI-based) or configuring visualization interfaces for optimal outputs.

In the UK, the use of Digital twins goes one dimension higher, by establishing a National Digital Twin – an ecosystem of connected digital twins, securely sharing infrastructure information to support better outcomes (https://www.waterdigitaltwin.com/). This system of systems aims at supporting government, industry and academics by creating a central hub for all infrastructure data in the UK, facilitating

operations that have interdependencies between sectors such as road, rail, waste management and telecoms, and their infrastructures.

In the water domains, the early implementations reported in the literature, labeled as Digital Twins, are applied to the urban water systems (Pedersen et al., 2021), given their similarity with manufacturing workflows. Here, Digital Twins are virtual models that replicate the real behavior of the system, integrating both elements and the dynamics between them. Using DT, simulations of the system's response under any conditions are provided, allowing water organizations to avoid problems (as a full understanding of the processes taking place is necessary to build the DT), preventing downtimes, planning future scenarios and interventions using simulations and optimizing system performance. The interactive nature of Digital Twins makes them a powerful tool for water managers, and greatly facilitates the handling and the analytics of data from the huge number of sensors typically installed in large infrastructures. Still, Digital Twins have started to be implemented for water utilities mostly as data aggregators with limited capacity to forecast scenarios based solely on simple data manipulation. Further evolution is needed, to achieve the full potential of DT in this area. In particular, the integration with hydraulic models and Artificial Intelligence algorithms is one of the avenues for the future.

In the ocean and coastal domains, the concept of digital twins was first applied in the Columbia River river-to-coast observatory (Baptista et al., 2008). Denoted Saturn, the virtual Columbia River, it provided a digital replica of the real system and a suite of services that ranged from multiple hindcast simulations and their ensemble, to address uncertainty, and forecast simulations using several modeling configurations, to support system's management and research advances. Saturn was one of the 3 pillars of the Columbia River observatory, along with an extensive monitoring network and a computational infrastructure that bounded all together. The infrastructure integrates continuously, in an interactive, two-way connection, the physical and the virtual assets (Baptista et al., 2015). This colaboratory shared both a product-centered approach, typical of forecast frameworks, and a user-centered approach through a data explorer tool where users could build knowledge and indicators by processing on-the-fly the information from both monitoring and prediction repositories. The user does not have, however, the freedom to build its own forecasts of the Columbia river or to apply the Saturn infrastructure on other systems.

The difficulty to monitor ocean processes in an accurate and efficient way, given the very expensive nature of ocean instrumentation, has motivated the development of digital twins of a robotic sensing network, to autonomously respond and adapt to changes in the environment (Barbie et al., 2021). Digital Twin prototypes connect the Physical Twin and the Virtual one. Sensors provide environmental information to the digital twin to mirror and model the physical twin, and provide actuators if an event is detected and monitoring strategies need to be adapted to the new data. Each monitoring node has its digital twins and synchronization of actions is provided in time. A digital twin for oceanic simulation of sea surface height and its application to the North Sea area was presented in Jiang et al (2021), exploring neural network methods to build a physics-informed surrogate model based on numerical model NEMO simulations.

This vision is well aligned with the application of the Digital Twins proposed by the Digital Twin of the Ocean expert group, where the applicant participates in the WP2 - Ocean modeling efforts. Figure 4.1 illustrates this complex interaction. Herein, this approach is further exploited by integrating the user dimension and their capacity to interact with both real and virtual dimensions. This vision stems from Mercator Ocean Digital Twins initiative (https://digitaltwinocean.mercator-ocean.eu/), which aims to develop a full-fledged Digital Twin of the Ocean, defined as a coherent, high-resolution, multi-dimensional, multi-variable and near real-time representation of the ocean. Based on the existing operational oceanography framework, it aims to integrate multiple existing and new remote and in-situ data sources, advanced ocean modeling, artificial intelligence and high performance computing with new ways to promote user interactivity to build new knowledge and support informed decisions.

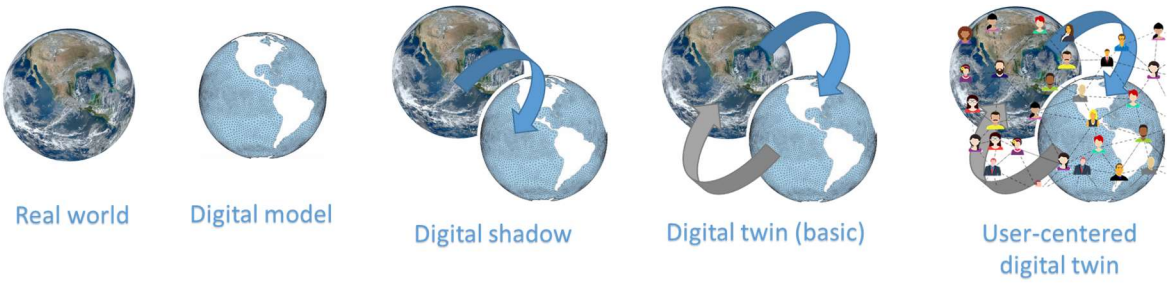


Figure 4.1 – The evolving transformation from monitoring networks to full-fledge Digital Twins

Given the scope of application, goals and similarity to physical and biogeochemistry coastal processes, a review is presented below on the DT components and the assets available for its construction, following loosely the definition and basic components of Digital Twins proposed by the DITTO project and in line with the methodology proposed in the Digital Ocean Forum, which identified the 4 pillars of the ocean value chain as “Real time ocean observing systems”, “Routine ocean monitoring and predicting (large scale)”, “Integrated infrastructure for on-demand modeling and data analytics” and “Customized applications” (https://www.youtube.com/watch?v=a6aSWEk9Fbk).

4.3 Coastal data lakes, reliability and quality assessment

4.3.1 Coastal data lakes

In the last decades a huge increase in data sensing, collection and availability in coastal and estuarine regions has occurred, thanks to low cost sensors, multiple data sources, open data policies in many data providers and the availability of data repositories based on shared, collaborative e-infrastructures. Herein, the most relevant data sources for estuarine and coastal regions are reviewed, organized along data types. Some discussion is also provided on their usefulness for service provisioning or support in the context of digital twins, being the fundamental components of their data lakes. Then, the challenges to integrate data into digital twins either in an automatic way or through dedicated user services is

reviewed with an emphasis on FAIR (Findable/accessible/interoperable/reusable) principles and data standards.

Remote sensing solutions have become major data sources for many coastal and estuarine data acquisition systems, given the technological advances of remote sensing devices, such as HFR radar, cameras or sensors mounted in drones, and the open availability of medium resolution (~10 m) satellite data from the Sentinel constellation. Different solutions provide data for different variables and devices can be applied independently or in an integrated fashion, to promote knowledge, validate forecast simulations or be part of assimilation procedures.

While open satellite data has been available for a long time (from USA providers), their resolution was only fitted for ocean applications given the small spatial scale of resolution. The Sentinel remote sensing data has brought satellite data to coastal and estuarine applications (e.g Pottier et al., 2021), and has been integrated in coastal forecast systems as a near real time data source for model/data comparison (e.g. Mateus et al., 2012, Oliveira et al., 2021). Some difficulties arise in altimetry data exploitation mostly to tides but recent developments were able to overcome these limitations (Cipollini et al., 2010). Typical products from the Sentinel constellation are wind fields from Sentinel 1 and Sentinel 3, significant wave heights from Sentinel 3, or sea surface color from Sentinel 2, amongst others examples (Sanchez-Arcilla et al., 2021). Exploitation of these products for both data assimilation and model validation has led to considerable improvements in coastal wave predictions, some improvements in velocities, but requiring correction with HF radar, and good prediction of storm conditions, but at distances of more than 10 km from the coastline. These results outline the importance of combining high-resolution coastal models with in-situ and satellite data to produce reliable ensemble forecasts (Sanchez-Arcilla et al., 2021).

Recent exploitation of Sentinel data for estuarine and coastal applications has promoted the development of several services that facilitate the access to data, process it in automatic ways and make it available through user friendly platforms. As the computation effort for these services is high, integration of these services in federated computational infrastructures such as European Open Science Cloud (EOSC) is fundamental to guarantee the capacity to serve multiple users simultaneously. Data Analytics in Earth observation workflows require a federated infrastructure with a core cloud computing and storage architecture optimized for very large data handling and fast user query response (Schumacher et al., 2022). WORSICA (https://indico.egi.eu/event/5464/contributions/15666/attachments/14141/17811/WORSICA_EGI2021_v1.pdf), for water bodies and coastal line extraction and water leak detection, estuarine bathymetry and intertidal topography extraction (Fassoni-Andrade et al., 2021), and coastal marine ecosystems mapping are examples of these services that are readily integrated in DT as part of the core technologies. They can also be combined with numerical or AI models to provide added-value results. In the scope of a user-driven interactive context, these methodologies can be components in a user specified framework that for instance allows for the selection of one remote sensing tool in a remote sensing tech pool, to be combined with a specific numerical model output, selected among several modeling services, to produce a new indicator or a new emergency system. Some core services such

as EO data archive, EO compute service and Analytics platform are already provided by the European C-Scale project (https://c-scale.eu/).

Space (10m) and time (~every 5 days) resolution of open satellite data are not, however, fitted for many coastal applications such as run-up and wave breaking location (Martins et al., 2022) and inundation extent evaluation (Gomes et al., 2015). Therefore, the use of fixed, locally-installed high resolution cameras has gained popularity and are already being used in many locations to support the determination of relevant quantities for coastal hydrodynamics (Harley et al., 2014). These cameras are relatively cheap to acquire, with high resolution in space and time (sampling for short waves). Some efforts are also being made to take advantage of non-stationary cameras (used for security or leisure uses) to obtain relevant coastal data (Conlin et al., 2022). Typically, the current research efforts explore the spatial dimension with some work to determine wave breaking through the multiple time frames.

The usage of cameras to assess critical hazardous conditions by combining time and space dimensions have not however been fully explored, in particular in articulation with forecast systems.

As the exploitation of fixed cameras required a considerable processing procedure (line detection, georeference, vertical projection) to extract the data that can be readily used in DT, their integration in drones, unmanned and manned small aircraft has been explored both as part of geodesics (to obtain topography or extract water delimitation/coast lines, Krapez et al., 2022), but also to detect hydrodynamics and water quality features. This concept was also explored in inland studies to detect water leaks (Krapez et al., 2022).

The usage of autonomous surface and underwater water drones with embedded sensors has also experienced considerable developments in the last decades, in particular in the scope of the Internet of Things (IoT) context (Petritoli et al., 2020). In order to achieve an efficient and low cost solution, the usage of distinct nodes within the swarm was proposed with some drones with long term communication capacity that interact/exchange information with external sources, while other drones only communicate with close partners. This will lead to a decentralized control that allows nodes to be able to take and articulate decisions even when external communication is missed (Lyhne Christensen et al., 2015). Collected data is typically of physical processes, but some water drones are already able to collect water samples (Kunze et al., 2020). This paper also proposes the use of fog cloud architecture that decentralizes cloud computing through distributed nodes. One avenue of research in this area is to use water drone swarms, that allow individuals to act as a single coordinated entity, in combination with numerical models and fixed monitoring networks, following earlier efforts that used boats as the mobile devices (Dang et al., 2010). Numerical or AI models are used as controlling devices of swarm paths through forecasts and as validators of the collected data.

The onset of small, low cost sensors, based for instance on sonar, radar or GPS technologies have also considerably increased the capacity to monitor coastal dynamics, by increasing the number of nodes in coastal water level and wave fixed monitoring networks (Kang et al., 2021, Raghukumar et al., 2019, Marimon et al., 2011). Besides encompassing a broader area covered by the surveillance system, these low cost sensors also allow for the installation of redundant sensors at the same location as part of the data quality guarantee of data provisioning (Casimiro et al., 2019).

The combined usage of the multiple sources of data acquisition technologies described herein are a fundamental part of the construction of resilient and redundant monitoring networks and coastal data lakes. Its usefulness was recently demonstrated in the Inner-Shelf Dynamics Experiment (ISDE), an intensive, coordinated, multi-institution field experiment conducted from September–October 2017 in the USA to better understand cross-shelf exchange through the inner shelf which is important for coastal water quality and ecology as well as transport of sediments.

4.3.2 Data reliability and quality assessment

Multi-source estuarine and coastal monitoring networks that combine in-situ and remote sensing, fixed and mobile devices, covering multiple space and time scales, are becoming the strategy to address coastal management supported by a data-rich environment, with data available at fine spatial and temporal scales (Fascista 2021). As data lakes become vaster and the number and importance of data-based tools grow in operational settings such as CDT or even only in data-based forecast systems (e.g. data-driven and hybrid models, coastal data assimilation, data-driven early-warnings) the quality (Sharma et al., 2022), availability and, in some situations, the security of the underlying information becomes fundamental. Understanding which faults can affect sensors in coastal regions, and how they can affect the quality of the measured data, and identifying methodologies to quantify and improve data quality is a pressing research theme that has been the focus of a recent Ph.D thesis developed at LNEC and FCUL (Jesus, 2019).

There are many sources of data failures in coastal environmental systems given the typically harsh conditions in these regions (Jesus et al., 2021), such as communication failures or sensor faults. These harsh conditions require frequent sensor maintenance operations and considerable human effort to validate the quality of data (Jesus et al., 2017). Several methodologies are being investigated for application to coastal environments to provide multiple, reliable and low cost communication mechanisms (Gaitan et al., 2022, Casimiro et al., 2021). Coastal environments are particularly complex as data communication over a body of water entails challenges (Gaitan et al., 2022). Furthermore, the existence of complex, spatial and temporal variable phenomena may bring difficulties in distinguishing a faulty behavior from real phenomenon signals.

Faults in sensor data can be due to sensor malfunctioning, incorrect sensor calibration, errors due to environmental conditions and other reasons. Faults can compromise sensor data quality through several forms either as outliers, offset and drifting failures, in particular when they are integrated in multiple node networks and used in automatic ways (e.g. to support the issuing of an alert in an early-warning system). Data quality is often linked to other terms such as validity (used when comparable reference is available), confidence (used for instance for probabilistic analysis), reliability (measuring the correctness of values) and trustworthiness (associated with security). For multi-sensor networks, the detection of sensor faults has improved considerably in recent years as the size of networks grows and provides additional sensors or environmental model outputs to be used in correlation analysis and afterwards in sensor fusion techniques (Jesus et al., 2021).

Detecting outliers in wireless sensor networks requires tailored techniques given their specific requirements. Zhang et al (2010) presents a review of these techniques that was later updated in Jesus et al. (2017). The former also presents a guideline for the choice of the most adequate technique, depending on the type of data and type and degree of outlier. A new methodology was proposed in Jesus et al. (2021a), based on dependability concepts, that detects outliers and distinguishes them from real events, using Artificial Neural Networks (ANN), quantifies the quality of each data set and finally corrects the low quality measurements. While the application of the methodology is site-specific as ANN needs to be trained with local information, the generic approach can be applied anywhere and use multiple sources of information for training.

Systematic failures such as drifts and offsets are much more complex to detect and their presence in a data repository can have many serious consequences as it is difficult to detect by hand when the failure starts to occur. Furthermore, this type of failure is very complex to distinguish from deviations arising from natural phenomena (Jesus et al., 2021b). A few studies are available to address each of these failures based on AI and statistical techniques. The framework proposed in Jesus et al. (2021a) was also implemented to detect outliers and drifts (Jesus et al., 2021b). In this application, several models are built for each sensor, based on past correct behavior. Detection is performed by performing temporal correlation of past values of the target sensor, space and value correlations of the sensor's past measures and other sensors. Both instantiations of the methodology were developed for runtime application and are therefore particularly suited for integration in DT.

Extension of these concepts to multiple sources of data including remote sensing are one of the avenues for future research, along with the implementation of a quality layer within a DT workflow. Furthermore, these concepts can also be applied to evaluate numerical or hybrid forecast results, to assess systematic deviations due to numerical errors.

4.3.3 FAIR data and mechanisms for compliance

Digital twins will generate, through both surveillance systems and prediction models, and their information processing to generate indicators and products, a considerable volume of information on a regular basis (typically on a daily basis but some simulations such as flash floods may create information even more frequently). Following international efforts for open data, data sharing and optimization of resources (European Commission.Directorate-General for Research & Innovation, 2016: https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf), the concept of FAIR (Findable, Accessible, Interoperable and Reusable) data has been gaining importance and is at the core of most shared ocean data repositories worldwide. Providing long-term preservation and continued access to open research data is fundamental to maximize its reuse. In Portugal, FCT is promoting the development of an Open Data strategy to be applied to national datasets both public and private.

Given that data source integration in the CDT data lake is one of the services for a continuously enhanced digital twin of the coast, to be done in a decentralized way by each data steward, the capacity

for the automatic evaluation of data FAIRness is a fundamental feature in DT. This step should follow the evaluation of data reliability and quality classification as discussed in the previous section.

The applicability of FAIR definitions is closely linked to the usage of standards in metadata and data formats. The evaluation of FAIRness of data requires the evaluation of several components, being it for data or for any digital object. Some of the aspects to address are a) data needs to be accompanied by proper metadata, following standards both data and metadata; b) the existence of a Persistent Identifier that should be accessible via protocols like HTTP, and supported by services such as repositories. Devaraju and Huber (2021) propose a set of metrics to assess the application of the FAIR principles and a set of practical tests to evaluate its compliance, with automatic assessment in mind. This work complements the RDA FAIR Data Maturity Model and was produced in the H2020 FAIRsFAIR project. This project aimed at developing practical tools for the application of FAIR principles throughout the research data life cycle in the European Open Science Cloud (EOSC) data repositories.

The proposed 17 indicators that address each of the 4 attributes are summarized in Table 4.1. Similar, but simpler, indicators are provided in FAIR_EVA (Orviz et al., 2020).

The development of this set of indicators is necessary for the automatic validation procedures. In the scope of a Digital Twin data lake, mechanisms for automatic certification of data fairness need to be part of the data processing workflow, as dataset fairness may need frequent compliance checks and the volume of data at stake prevents any human verification. Several tools are already available to fulfill this task and they can benefit and improve their quality by being integrated in Software Quality Assurance as a Service (SQAaaS) tools. Currently, two softwares stand out in this regard: F-UJI (Devaraju and Huber (2021), a web service to programmatically assess FAIRness of research data objects based on metrics developed by the FAIRsFAIR project (Table 4.1), and the FAIR evaluator for DIGITAL.CSIC, developed in the scope of the EOSC-Synergy H2020 project (Orviz et al., 2020). The last one is being integrated in the SQAaaS tool developed in that project.

Table 4.1 - FAIR indicators (adapted from Devaraju and Huber 2021)

Indicator number	Attribute	Description	Validation tests
1	Findable	data are assigned a globally unique identifier	a) The object should be given a unique identifier (URI) that follows a proper syntax b)The identifier is Web accessible (not broken)
2	Findable	data are assigned a persistent identifier	A data identifier is specified based on a commonly accepted persistent identifier scheme suitable for digital objects The identifier resolves to a landing page with metadata of the data object
3	Findable	metadata include descriptive core	Limited core metadata properties are specified

		elements to support data findability	Citation metadata properties are specified (creator, title, publication date, publisher, identifier, resource type) through appropriate metadata fields All minimum descriptive metadata properties (creator, title, publisher, publication date, summary, keywords, identifier, resource type) are specified through appropriate metadata fields
4	Findable	metadata include the identifier of the data they describe	Metadata contain a PID or URL that represents the Web location of the downloadable data content A data identifier is included in the metadata, and it matches the identifier provided as part of the assessment request
5	Findable	metadata are offered in such a way that they can be retrieved by machines	Metadata of the object are retrievable programmatically through at least one of the following methods: structured data embedded in the landing page of the data object typed links of metadata document or signposting header links content negotiation with a PID provider service
6	Accessible	metadata contain access level and access conditions of the data	Metadata include the level of data access (e.g., public, embargoed, restricted) and their access conditions using appropriate metadata fields Access level metadata are machine readable, and this is verified against controlled vocabularies (COAR, Eprints, EU Vocabulary, and OpenAIRE)
7	Accessible	metadata are accessible through a standardized communication protocol	The metadata URI's scheme is based on a common application protocol The metadata are accessible through the identifier provided
8	Accessible	data are accessible through a standardized communication protocol	The data URI's scheme is based on a shared application protocol The data are accessible through the identifier provided
9	Accessible	metadata remain available, even if the data are no longer available	Preservation of data and metadata is an explicit role of the repository. Therefore, it should be assessed at the level of a repository, not at the level of individual objects.

10	Interoperable	metadata are represented using a formal knowledge representation language	The metadata of the object are available in a formal knowledge representation language; e.g., through at least one of the following mechanisms: ○ parsable, structured data are embedded in the landing page ○ parsable, formal metadata (e.g., RDF, JSON-LD) are accessible through content negotiation, typed links, or SPARQL endpoint
11	Interoperable	metadata use semantic resources	Namespaces of known semantic resources (excluding common namespaces; e.g., RDF, RDFS, XSD, OWL) are present in the metadata of an object
12	Interoperable	metadata include links between the data and its related entities	Metadata capture the relation between a data object and its related entity. The relation should be expressed using a relation type and, if a URI is used to represent a related entity, it should be accessible
13	Reusable	metadata specify the content of the data	Metadata include the type of the object and the technical properties of its data file, such as format, size, and observed variables Metadata values of the properties comply with the actual data file
14	Reusable	metadata include license information under which data can be reused	Metadata contain license information represented using an appropriate metadata element A standard, machine-readable license is specified
15	Reusable	metadata include provenance information about data creation or generation	Metadata include properties representing data creation, such as creator, contributors, creation and modification dates and version, source, and relations that indicate data creation activities Provenance metadata are available in a machine-readable version of PROV-O or PAV
16	Reusable	metadata follow a standard recommended by the target research community of the data	Metadata are available through at least one of the domain metadata standards listed in the RDA Metadata Standards Catalog
17	Reusable	data are available in a file format recommended by	Data are available in a long-term file format as defined in ISO/TR 22299 Data are available in an open format

		the target research community	Data are available in a scientific file format (e.g., Library of Congress dataset formats, Wolfram Alpha supported file formats)

4.4 Prediction engines and analytics component

4.4.1 Combined prediction models

Prediction of coastal and estuarine conditions have been done for several decades by traditional physics-based numerical models (e.g. Baptista et al., 2008). The accuracy of these simulators have been increasing steadily as coastal processes knowledge has evolved, supported by a fast increase in computational methods and resources, and new data sources (described in the previous section). Most of the prediction engines behind the first digital twins in the ocean and coastal regions are also based on this type of models (https://digitaltwinocean.mercator-ocean.eu/). The numerically-based forecast engines have been described in detail in Chapter 3, along with a detailed review on model accuracy and uncertainty assessment. Data-based forecast alternatives were also detailed in that chapter.

Herein, alternative modeling choices for the estuarine and coastal prediction engines of DT are reviewed, denoted as surrogate or hybrid models, that are based on the integration of computer science AI tools for data-based prediction, data sources and the traditional numerical models (Frolov et al., 2009). AI tools are becoming popular to address issues such as climate change in complex environment such as the atmosphere and the ocean (an example of such efforts is Climate Change AI, a global initiative to catalyze impactful work at the intersection of climate change and machine learning - https://www.climatechange.ai/).

Computational costs associated with numerical coastal models have been pointed out as a limitation for high-accuracy numerical predictions in particular through ensemble simulations. The motivation behind the development of ensemble models is to take into account for instance the uncertainty in oceanographic and freshwater inputs, but computational costs can become prohibitive for many applications. At the same time, AI based model surrogates (such as Machine Learning and in particular neural network models) provide a fast execution time that can permit an ensemble analysis that could take into account in a detailed way atmospheric uncertainty cascading to ocean and coastal simulations and also permit iterative parameter optimizations (van der Merwe et al., 2007, Frolov et al., 2009). Hybrid or surrogate models can be several orders of magnitude faster than equivalent numerical models (van der Merwe et al., 2007, Wolff et al., 2021) and are less demanding regarding the volume of field data than pure ML models (Wolff et al., 2021). Physics-informed neural networks are an example of such a hybrid approach that combines a data-based neural network model and a physics-informed mechanistic model, by minimizing the error with the data as well as the error with the physical equations. Recent strategies for these models include methodologies for optimal weighting of the error functions and have been applied to the incompressible Navier Stokes equations. While the data-model applications are still restricted to the studied coastal site, their portability to other systems through the

combined methodology is high. Training is done using numerical models duly calibrated and applied to the coastal site of interest and the site data. After training is completed, the emulators are able to provide accurate predictions over new forcing conditions.

Most of these hybrid models have been applied to ocean domains, without accounting for the strong nonlinear estuarine and coastal processes. Moreover, these applications do not account also for condition changing environments such as intertidal flats or go beyond hydrodynamics. Given the flexibility of these hybrid approaches, their application should be explored in complex intertidal environments, accounting for morphodynamic simulations and / or water quality/ecology processes. Ecology understanding in climate change context in ocean settings through hybrid modeling is started to be investigated in the OcéanIA project (https://oceania.inria.cl/).

4.4.2 Data assimilation in prediction engines

Data assimilation is a fundamental component in the construction of digital shadows and digital twins, linking data and models in a seamless and quality-driven way (Frolov et al., 2009). Its application has been undergoing for several decades in the ocean community, supported by vast cooperative monitoring programs such as the ARGO initiative (Roemmich et al., 2009) and satellite remote sensing data (LeTrahon, 2011) in particular from the Copernicus Sentinel program (https://www.copernicus.eu/en), and the development of the Global Ocean Data Assimilation Experiment (Smith, 2000) with considerable increases in output quality (Jacobs et al., 2021). Ocean forecasting systems use data assimilation to estimate initial and boundary data at each forecast time, to interpolate and smooth sparse or noisy observations, and to quality control observing network and forecast models (https://www.climate-policy-watcher.org/oceanography/introduction-to-ocean-data-assimilation.html). The main concept of data assimilation modeling is to produce estimates of the relevant variables that are consistent with data and model dynamics, allowing for errors in both. Data are assimilated to produce a revised prediction through optimal combination of information from the model and observations. The analysis is then used as the initial condition for the next prediction cycle, and the process repeats. Nowadays, the vast majority of ocean prediction systems integrate data through assimilation, CMEMS being one of the top examples of numerical model prediction quality assurance through data assimilation (Le Traon et al., 2019). Data assimilation and models do not only promote model improvement, but they can also be used to design and assess ocean observing systems through the creation and establishment of Observing System Simulation Experiments (OSSEs, Jacobs et al., 2021).

Several concerns have limited the application of data assimilation schemes in estuarine and coastal zones until recent years (Stanev et al., 2016), as the small temporal scales and horizontal scales are both computationally and scientifically challenging for data assimilation. Several issues are identified in Stanev et al. (2016), herein the most relevant ones are described. The difficulty to maintain operational, real time monitoring networks in these regions due to the strong dynamics, variable salinity conditions and exposure to vandalism has limited the necessary data acquisition for continuous long periods. The recent development of low cost sensors, deployed over fixed structures or in swarms of autonomous

vehicles (Wynn et al., 2014) /multiple drifters (Novelli et al., 2017), and locally-based remote sensors (Martins et al., 2022) has changed our capacity to build dense, reliable monitoring networks in coastal zones and has alleviated this constraint. Additionally, the complexity of the non-linear processes in coastal regions is a limiting factor to the application of simple assimilation techniques. The loss of predictability is associated with the nonlinear transfer and growth of errors (Stanev et al., 2016). One traditional approach to assimilation in estuaries was based on nudging factors, applied directly to the equations of motion, and used for instance with HF radar data. This approach is however limited to the surface currents and inconsistencies in the vertical may arise. Finally, many data assimilation methodologies require the specification of the data error, which can be quite challenging in the fast varying coastal environments. Extending the spatial and temporal coverage of remote sensing data whether from HF radar (Stanev et al., 2016) or Satellite data (Etala et al., 2016) appear to be a promising approach to overcome assimilation inconsistencies.

Data assimilation methods are generally classified into optimal control theory-based or sequential methods (Bertino et al., 2002). The use of several variations of the Kalman filter approach have also been applied with relative success in 3D hydrodynamic modeling, accounting for the uncertainty in forcings, in the data and in the system characteristics (Bertino et al., 2002). Application to an ecological model strengthens the importance of these techniques for dynamics that are very dependent on initial conditions. This method has also proved to handle the nonlinearity of estuarine and near coast dynamics well. Recent developments (Asher et al., 2019, Zijl et al., 2015) have promoted new assimilation schemes that handle this complexity at least for hydrodynamic modeling. Optimal interpolation is one specific type of sequential methods, which uses the variability in the observations and an error covariance matrix, assumed a priori, to create a spatial field that integrates model and observations, and can emulate physical behavior by selecting appropriate covariance function(s) (Asher et al., 2019).

Storm surge and flood predictions are one of the most active areas in the use of assimilation techniques in estuarine and coastal zones. Errors in the predictions of the atmospheric sources are one of the most limiting factors, along with model shortcomings and parameterization in operational context. The 4D-Var technique (Lionello et al., 2006) is an example of a popular approach which defines a cost function describing the discrepancy between model results and observations. The adjoint and the conjugate gradient methods are used for the computation of the cost function gradient and the search of its minimum, applied to each forecast run. During each analysis period, the observations are assimilated in the model to identify the optimal initial condition for the next surge prediction. The procedure is capable of compensating for the forcing and model shortcoming errors and improves the accuracy of the storm surge forecast. Zijl et al. (2015) applies a sequential method based on a computationally efficient steady-state Kalman filter to address errors in the meteorological forcings during severe storm surges, achieving considerable enhancements in forecast quality at shorter time scales.

The possibility of using multiple sources of data in the assimilation procedure raises the potential for better predictions but requires the capacity to ensemble data from multiple sources, with potentially distinct time and space scales, heterogeneous formats and semantics (Fan et al., 2021). Examples of

such approaches are the combination of satellite altimetry and water level data that was used to improve storm surge forecasts (Etala et al., 2015) and the integration of remote sensing, social sensing and crowdsource data (Schnebele and Cervone, 2013). Recent studies have attempted to develop new models and algorithms to integrate multi-source data in particular using authoritative and social-originated data.

However, the integration of multiple sources of data may hinder even more the computational cost associated with data assimilation. Depending on the methodology chosen and the level of refinement in the model simulations, data assimilation may increase CPU time by an order of magnitude (Asher et al., 2019). The use of the proposed optimal interpolation can be a trade-off, with smaller increase in accuracy but with a smaller computation effort. Alternative approaches for computational speedup have also been proposed in Frolov et al. (2009), applying a reduced-dimension Kalman filter, without requiring an adjoint calculation and maintaining accuracy. To facilitate data assimilation procedures, several common general frameworks for data assimilation in ocean models have been proposed, which allows for the choice of the methodology being implemented, such as COSTA (van Velzen and Verlaan 2007) and OpenDA-NEMO (van Velzen et al., 2016), part of the SANGOMA Tools for Data Assimilation (https://www.researchgate.net/publication/266139953_The_SANGOMA_Tools_for_Data_Assimilation?msclkid=9433c2cecfa911ecbb0838b5814f6c46).

4.5 User interaction mechanisms for knowledge creation and coastal management support

4.5.1 User-oriented platforms

The process of sharing estuarine and coastal information to end-users, being them coastal managers, companies operating on the coast or even the general public in support of recreational activities, is challenging given the complexity of coastal processes and their varying time and spatial scales and seasonal signal. Existing web platforms and observatories are typically dedicated to specific sites (e.g the Columbia river collaboratory - Baptista et al., 2008, the Ria Formosa and Tagus estuary observatory - Rodrigues et al., 2021), to specific goals (e.g. inundation and erosion - Dongeren et al., 2018, water quality/biogeochemistry - Rodrigues et al., 2021, oil spills - Oliveira et al., 2014) or to targeted user communities (researchers, managers, public, ...). WebGIS platforms are by far the most advanced solutions, providing interactive capabilities (e.g. probing over prediction results) along with the capacity to visualize results down to very fine spatial scales (maps), integrating also the time domain through sliders and other technological solutions. The most relevant current limitation of WebGIS is the capacity to renderize results on the screen in a very fast way without requiring sophisticated resources from the user side (Gomes et al., 2015). Recent works in the scope of the Digital Twin of the Earth initiatives have developed a web interface to explore and analyze CMEMs data, based on Open Geospatial Consortium (OGC) standards and enabling the possibility for product sharing through the platform.

The potential usages of the Coastal Digital Twins are very broad by definition and a user exploratory and on-the-fly product creation nature is the expected level of service delivery. The user interfaces of

CDT should be flexible on the nature of application and the degree of complexity of the products. They should provide on-demand services tailored to the knowledge of the users, guiding them but providing freedom of choice of models, either data- or numerical equations based, type and nature of outputs and capacity to create new products based on existing assets. CDT should also promote a culture of knowledge sharing, providing the capacity for knowledge and product sharing amongst users in a public way or within communities (from Digital Ocean Forum outputs: https://www.youtube.com/watch?v=a6aSWEk9Fbk).

Several approaches have been proposed in the literature to organize products and services in an efficient way, targeting making data, forecasts and their analysis available for users in a facilitated way. Thematic access portals (Rosebrock et al., 2015, Rodrigues et al., 2021, Rocha et al., 2021), dedicated to themes like public health, storm surge emergency, search and rescue or ecosystem evolution, are some of the observatories/collaboratories developed, providing distinct services for distinct uses, being it emergency, daily management, strategic planning or simply recreation. Examples include on-the-fly oil spill simulations customized by users in a simplistic way, based on very complex circulation outputs, for real time emergency usage, shared along pre-processed simulations for the most probable spill locations (Oliveira et al., 2014) or public health public reporting, creating confidence and promoting transparency in knowledge sharing (http://marvlis.blogspot.com/).

Typically, products are available or can be set up with limited flexibility to input user knowledge or own tools. On-demand services are however distinct from pre-customized tools. They are a key component of CDT, as user interaction is a fundamental and innovative capability that makes them distinct from other IT tools. Early approaches to user-enabled capacities were proposed in the Saturn observatory of the Columbia river, through a number of information processing tools that users can build based on model and data information. Examples include 2D and 3D correlation of variables, setting of early warning conditions or efficient capacity to browse quickly through the gigantic archives for “near-me” data (Baptista et al 2008). These early initiatives have been followed by increasing complex services that combine models, fixed and mobile data and analytics to build intelligent outputs. For instance, CoTrack is a Collaborative Tracking framework, that allows mobile sensors to cooperate with fixed sensors and numerical models to accurately track dynamic features in a coastal and estuarine environment. In these early efforts mobile sensors were boat-hosted, but the advances in surface and underwater water drones can extend these tools to intelligent, two-way interaction where users or automatic forecasts can correct drone paths to a specific goal accomplishment.

Most challenges associated with user-driven products are related to the difficulty to set up the information aggregation level for a specific user, as different aggregation levels for each user may be required to achieve a specific goal. The user should be able to select the granularity of the data processing. For quick, emergency issues, a very simple and fast approach to interpret information should be used (e.g the overtopping risk spatial representation of Hidralerta system - Fortes et al., 2020, or the inundation tables at predefined risk spots of the MOLINES system - Gomes et al., 2015). However, simplistic views should always be complemented with additional, optional layers of in-depth information, to support the understanding of the sources of risk and the uncertainty associated with the

predictions. Most existing platforms follow only one of the approaches, either failing to support users because of too complex, hard-to-understand information or by lacking the information that would support a simple interpretation.

Building simplistic, compact services to digest complexity into usable information require considerable knowledge on the dynamics being measured and simulated. For instance, complex numerical models have been used in the past to setup the optimal locations for sensor network node location (Baptista et al., 2015), in order to sample conditions at the relevant hot spots. This ingenious idea can also be used to convey compact information on the relevant distinct locations within a sea of Terabytes of model outputs. Time series and statistics/indicators computed at these locations, meta-informated on the dynamic “cells” that they represent, can facilitate the understanding of key messages to non-expert managers and to the general public (to explain for instance the best locations for fishing or the safest locations for recreation). While some examples have already been developed, many more can be created in close articulation with end-users, both in terms of pre-cooked data or tools for flexible building. Iterative tools could be built with some basic capacities for indicator creation and new capacities be added as new information sources with more complex data are available.

While compact approaches can be the best approach for some societal applications, the exploitation of the innovative tools in visualization technologies using extended reality (mixed or augmented reality) can expand the usefulness of CDT products, as they provide ways to visualize a combination of real and digital worlds in an enhanced way. In particular, for scenarios analysis (e.g. for climate change impacts) or any what-if simulations, extended reality can facilitate understanding of impacts, allowing the user to configure on-the-fly what is being seen, adding or removing elements as he/she see fit and combine model and data outputs in a seamless way. Some examples of these technologies (mostly coming from the videogame area) as stand-alone works have already been applied to estuarine and coastal environments (e.g. Katsiokalis et al., 2020; Templin et al., 2022) to address coastal erosion and safe navigation. Their integration in CDT can have benefits at several levels, from technical to educacional contexts.

However, there are still challenges in the application of these technologies in water domains, as limitations arise in the two types of approaches to achieve both good accuracy and high efficiency. Sensor-based methods are efficient, but sensors offer only low accuracy in mobile devices. Vision methods are very precise but typically require significant computational and memory space for image processing. Hybrid approach that integrates combining vision-based and sensor-based methods can be a potential avenue for improvement but much work remains to be done (Templin et al., 2022). The application of extended reality in CDT may have smaller requirements for technical uses as near real time answers may not be necessary, as decisions in estuaries and coasts involve in general very expensive interventions and need to be carefully analyzed.

4.5.2 Analytics for user-driven products creation

Both numerical models, AI tools and their integrated solutions generate complex outputs that are difficult to analyze by non-experts and to process into meaningful indicators/aggregators for knowledge creation

and system management. Additionally, the onset of computational resources has triggered the development of increasingly higher resolution forecasts that generate vast amounts of information daily that are not possible to be analyzed by humans. Big data is a well known challenge in many fields and analytics have been developed to process coastal big data information for multiple uses (Duque Ordoñez, 2022).

The availability of integrated solutions comprising data and models in digital twins brings in new challenges where both information sources can be combined to produce new services and boost knowledge creation. Many tools in the field of oceanography and atmospheric sciences have been proposed to process the large volume of information and make processed results available to users. Most of these tools are dedicated to the research community and provide very flexible ways to experiment, change processing workflow, integrate new algorithm, etc… without the need to download data and taking advantage of federated data and computer resources (e.g. ECAS Climate analytical service, Bendouka et al., 2019) Other tools perform complex computations and produce indicators/aggregated information that are made available to users typically through user-targeted applications such as the ones presented in the previous section (Jusa and Tintoré, 2021, Rodrigues et al., 2021). However, none of these approaches fulfill the basic requirements for user-promoted usage of CDT: each user should be able to access information at the aggregation level necessary for his/hers needs and be able to customize products to the final goals. On-demand coastal products that do not entail knowledge of complex tools or processing environments such as Jupyther notebooks but are yet flexible to comply with user requests are quite rare and are often very limited in scope.

Future directions in the scope of CDT point out services built on top of Jupyther notebooks and similar tools as a possible avenue for simple uptake of user processing tools and sharing across the communities of practice. The possibility for direct edition of the workflows and methodologies through interactive interfaces is a major future component of CDT to facilitate the uptake of very complex, high accuracy products by the economy end-users. The availability, edition and sharing of multiple notebooks can be achieved with notebook computing environments such as BINDER (Project Jupyter et al., 2018). Binder is an open source web service that lets users create cloud-based sharable, interactive, reproducible environments where pre-existing analytics workflows become interactive versions of GitHub and other repositories.

Analytics are also available as a prediction tool builder for optimized prediction and numerical modeling research. Firedrake is an example of such a tool that can be used through Thetis to build custom numerical formulation of finite element models. Karna et al. (2018) illustrates the construction and evaluation of such a tailored model applied to 3D estuarine circulation and its demonstration in a complex river-estuary system. Implementation of these modeling builders is an avenue for exploration in the context of CDT.

Finally, GIS-related analyses are part of almost all coastal information platforms addressing spatially relevant issues (Oliveira et al., 2014, Knight et al., 2015, Rocha et al., 2021) and spatial analytics are a very important asset for knowledge creation and management support. Recently multivariate spatial analysis has been gaining importance, in particular the most recent trends of explicit formulations that

allow for spatially varying outputs. Its application to data lakes containing crowdsourced data has proved to be very effective (Malczewski and Jankowski, 2020).

In order to provide a suite of analytical tools for integration in a DT, harmonization of processing platforms is necessary as well as the usage of metadata and file formats following oceanographic and coastal area standards. The construction of a tool lake, where pre-cooked indicators and services coexist with on-demand construction of new methodologies requires the usage of platforms like Jupyter and its processing environment BINDER. These and other tools can provide a capacity for knowledge creation, training and management support with precedence in coastal research and coastal engineering and management.

4.6 Open questions

The efforts towards building digital twins for coastal regions are still at their infancy, most of the efforts still concentrated in ocean scales, and many challenges lay ahead. The reviews above highlight the past work that can contribute to the construction of these societal-driven tools, taking advantage of the achievements and tools already available. Herein several avenues of research are identified along with some core steps that still need to be developed.

Central questions that need to be addressed at this stage are:

- Should we have all coastal DT as part of the large scale ocean DTs, in a seamless continuum or should two-ways connections be made between bordering DTs and this way optimize resource distribution and tailor DT to each usage?
- Should all possible services be available at all coastal DTs or should we customize each site tool for the relevant issues, allowing for on-demand improvements and additions as needed?
- Should specific prediction and data processing tools deemed as the best be the only ones in DTs or should we build a tool lake, where new tools can be integrated on-the-fly, and look for optimal performances in each application (dependent for instance on the current forcing conditions?
- What are the best ways for conveying information to distinct end-users in particular for mediation of conflicting uses? Should user-developed editable products and methods sharing be a mechanism to facilitate mediation?

The analysis of the reviews above, looking for answers to the above questions and others, allow for the identification of the following open research issues in coastal and estuarine DT:

1) Coastal and estuarine digital twins are not available yet as generic platforms with capacity for integration of multiple prediction engines (that solve the relevant processes at the correct space and time scales), multiple data sources and advanced processing tools. The development of these sophisticated tools as a generic framework, applicable to coastal regions worldwide, is one of the most relevant open questions in estuarine and coastal research and a most needed tool for conflict-mediated, science-based coastal management.

2) The effort to create multiple coastal DT either in a decentralized/soft-coupled or centralized way can be huge for a widespread application of this concept, with considerable duplication of efforts and low capacity for technology updates. In a vision similar to OPENCoastS, building a digital twin builder, to deploy customized frameworks where tools´ and thematic-related services lakes are the ones relevant for the specific site, while allowing for on-the-fly addition of processing workflows, new prediction tools or tailored services, is an avenue for research, entailing challenges on the technology and concept sides. The outcome of the previous open issue would constitute the backbone of the proposed DT builder.
3) The requirements for prediction/modeling engines in coastal DT far outweigh the existing prediction tools, mostly without ensembling features or without including multiple-source data assimilation, and without automatic mechanisms for optimal prediction choice. At the same time, new approaches have been proposed for new data-based and hybrid modeling engines that need now to be integrated in prediction workflows, along with efficient data assimilation schemes.

In the scope of this open research question, the following themes are relevant issues to be addressed and developed in the future:

- Creating a multi-option forecast engine that provides ensemble of distinct prediction engines, built with the prediction components chosen by the user and including a classificator for selection of the best prediction depending on the specific environment conditions;
- Given the flexibility of hybrid modeling approaches, their application should be explored in complex intertidal environments, accounting for morphodynamic simulations and / or water quality/ecology processes;
- Development of methodologies to assimilate multi-source data in a seamless and quality controlled way, keeping computational cost under acceptable conditions. These methodologies should have the capacity to setup an assimilation strategy that can be applied for a multi-scale (from the basin to the sea), multi-process digital twin (including hydrodynamics, coastal erosion and water biogeochemistry) and be applied to data-based or numerical models, in a user-driven, on-demand way.

4) Data and general information are core components of DT and their quality must be known as well as its reliability (Sharma et al., 2022). Building space-wide quality classificators, for early-detection of outliers, drifts and other faulty behavior are necessary to support data integration in DT´s data lakes, but are not available yet. Extension of the single monitoring network node methodology concepts to multiple sources of data including remote sensing are one of the avenues for future research, along with the implementation of a quality layer within a DT workflow. These concepts can also be applied in the future to evaluate numerical or hybrid forecast results, to assess systematic deviations due to numerical errors. These developments become even more relevant for real time information, as a continuous, automatic evaluation of the reliability of real-time monitoring data and the implementation of operational procedures for

data quality assessment are currently not used in operational coastal systems. They are a fundamental component of the DT by providing confidence levels to the continuous flow of real time data and correct faulty data and fill in data gaps.

5) Implementing the previous challenge requires a novel integrated reliability infrastructure that could address information in general, evaluating data and model results in a seamless way where both approaches could contribute towards data fusion and failure compensation. This infrastructure should also accomodate ensemble simulations to tackle uncertainty and optimal choice of simulation for a specific environmental context and data assimilation mechanisms, bringing in several approaches for quality improvement, targeting an overall framework where optimal, quality-labeled information would be produced to support user-driven services.
6) Building new services tailored to user needs at specific coastal systems requires the best exploitation of the available sources of data. Open issues to be explored include workflows for data exploitation available at each DT. Several examples worth investigating include: 1) selecting one or several remote sensing tools from a remote sensing tech pool, to be combined with specific numerical models output, selected among several modeling services, to produce a new indicator or a new emergency system; 2) using water drones swarms with mobile sensors to cooperate between them and with fixed sensors and numerical models to accurately track dynamic features in the coastal environment, embedded in a user-dedicated platform for on-the-fly control; 3) Using local cameras to assess critical hazardous conditions by combining time and space dimensions of images that have not however being fully explored, in particular in articulation with forecast systems
7) The implementation of automatic data FAIR assessment tools in data lake integration in the DT, as part of the quality control workflow, has not been implemented yet in coastal research. This task involves several challenges in terms of standardization of data and metadata formats as well as unification of ontologies.
8) User experience and interactive portals are a vital component to bring DT to all its potential users and for all economic benefits. Although many works have been developed in multiple components of this target, their efforts have not been integrated yet and many issues require further research. For instance:
    - multi-level, multi choice web portal that allows users to select the path to follow according to their needs is necessary, providing “highways” for selection of thematic, nature of use, high level of detail/aggregate simple outputs, open-page configuration experience or pre-cooked indicators; and helping users to navigate to the desired outcome, learning from each user experience to propose intelligently the next steps. Many challenges exist today on the conceptualization of this tool, such as the technological choices for optimal user experience and the harmonization of the many building blocks.
    - CDT requires portals that grow in service offering as knowledge is created and be flexible by providing different aggregation levels for each user as required to achieve a specific goal. The user should be able to select the granularity of the data processing, the information analytics processing methodologies and new products/methods built on the fly

should be integrated in the list of offered services to other users. Portal usability conceptualization for these goals is challenging along with the technology and the resources to support it.

- To allow for the possibility to view/create thematic services, to serve specific communities and promote interaction between peers, or the possibility to build services addressing one specific type of information per concern that span several disciplines (e.g. sea water level rise impact on inundation, ecology, ...).
- While focusing on specific themes, CDT portals should also provide cross-discipline capacities, building new tools for knowledge creation by mimicking concepts from one area to the next, thus exploring cross-fertilization to create out-of-the-box products. For instance, time series and statistics/indicators computed at hotspot locations, meta-informated on the dynamic “cells” that they represent, can facilitate the understanding of key messages to non-expert managers or to the general public, as opposed to complex 2D or 3D results visualization. This field data-inspired for model processing is an example of the capacity that a user-exploration, on-demand service building tool should provide and the conceptualization and implementation of such a tool remains to be addressed. Building cross-fertilization algorithms is challenging and their automatic implementation requires artificial intelligence paradigms.
- Coastal portals typically provide access to information on pre-defined grounds. The capacity to build new knowledge even if on-demand, exploratory tools are provided may be complex for many users. Taking advantage of extended reality may facilitate product exploitation and creation but the integration of these methodologies encompass challenges on the conceptual and technological sides, along with the adequate choice for the resources available. Hybrid approach that integrates combining vision-based and sensor-based methods can be a potential avenue for improvement but much work remains to be done. For experienced, expert users, the possibility to build tailored automatic services, using Jupyther notebooks and services built on top of DT is also a possible avenue for simple uptake of user processing tools and sharing across the communities of practice.

5 Research Studies presentation

5.1 Introduction

The Present Research Program and related Post Graduation Program are framed in the scientific area of “Maritime Hydraulics” and its implementation is expected to address the challenges and opportunities provided by Information technology tools towards supporting the coastal engineering research developed at LNEC and the societal needs in a variety of areas such as emergency response, sustainable coastal management and adaptation to climate change and anthropogenic actions.

The previous chapters introduced the Program’s theme, framed it in the national and international coastal research context as well as in the research strategy at LNEC, in particular in the scope of the Hydraulics and Environment department competence areas. Then a review of the state-of-the-art was presented organized along the two main themes of this Research program: coastal forecast systems and coastal digital twins. Each of the chapters finished with the main open research topics for each theme.

In this chapter the proposed Research Studies and its research projects are presented, aiming at addressing most of the identified research gaps, framed in LNEC’s research program. Research at LNEC’s Information Technology in Water and Environment Research Group (GTI) for the next decade will incorporate in its general actuation lines the following themes:

- Digital twins for the estuaries, coasts and all bordering water systems, available for generic or thematic creation, to address all societal needs for information access and sharing and knowledge creation;
- User centered services encompassing on-demand forecast platforms, integrating process-based, data-based and hybrid modeling, applicable from the ocean to the hydrographic basin, and integrating physical to biogeochemistry processes, and multi-source information-centered observatories.

In the next subchapter the conceptualization of the Research Program is presented, along with the description of each Research Study and its research projects. The next two chapters detail each Research Studies along with its research projects, identifying goals, methodologies and activities, as well as the human and financial resources for its accomplishment. Outputs and societal benefits are also described in these subchapters, framed in national and international initiatives.

5.2 Vision, strategy and implementation of the Research Program

The proposed Research Program aims to advance the state of the art in digital coastal research and promote the research strategy for the next decade at GTI. It also aims to place LNEC at the top of the research institutions in Portugal that integrate IT and natural science knowledge to produce actionable

research to address societal needs and in particular the products to support science-based coastal management. The proposed work is thus expected to be implemented through joint collaboration of GTI and the coastal-related divisions at LNEC, as well as external partners at national and international level, continuing GTI's collaborative work developed at DHA through multiple projects and networks, some of them already approved and others to be submitted later. At international level, these Research Studies are framed in the activities of LNEC at ATTRACT-DIH - Digital Innovation Hub for Artificial Intelligence and High-Performance Computing, a European Digital Innovation Hub, and EGI-ACE, an EOSC-powered thematic services project, both led at LNEC by the applicant. Nationally, this research is mainly framed in the ANI-funded SINERGEA project also led at LNEC by the applicant.

This strategy's main pillar is the development of Digital Twins (Research Program 2), focused herein in coasts and estuaries but conceptualized in a generic way so that they can be ported to other areas in the Hydraulics and Environment Department of LNEC. As digital twins aim to address societal needs in a user-centered, operational way, their development and evolution require the development of several lines of research, including on-demand, hybrid forecasting (Research program 1); multi-source, quality-assessed, FAIR data lakes; and exploitation of virtual research environments (Research program 1 and 2). These research lines are combined to develop high-resolution, thematic services made available through Digital Twins platforms and user interfaces, tailored to the needs of each user, from the research communities, to managers, industry and the public.

The Research Studies are expected to be developed over a 10-year timeline and are integrated in the current and the next LNEC's Research and Innovation programs, the last one about to be released. The relationships between the Programs' and LNEC's research are illustrated in Figure 5.1, framed in IT-related LNEC/DHA's program. They will also contribute to several other Research and Innovation projects at DHA, providing tools and information to address issues such as coastal dynamics, impacts of climate change and anthropogenic actions, and emergency and risk management.

The first Research Study, ***OPENCoastS4ALL - Reliable, quality-controlled, multi-process, on-demand coastal forecast framework for oceans to hydrographic basin application***, addresses the core engine for forecasting the coast in all its dimensions, based on the pioneering work of the applicant in the user-centered, on-demand forecast service OPENCoastS. It is anchored on LNEC's SIIN2A project (Innovative Information Systems for Smart Water and Environmental applications) with important contributions towards development of the tools for early-warning of LNEC's "Risk and Safety management in Hydraulics and Environment" project and the development of tailored multi-process forecast systems and data lakes for LNEC's "Environmental Management of water and coastal systems".

CODIT - Intelligent, high-resolution, user-centered and inclusive coastal digital twins, is the second Research Study, and it aims at creating for the first time high-accuracy representations of any real coastal systems through two-way interactions between digital representations and the real world, centered and operated by any users. OPENCoastS4ALL is one of the core components of CODIT, individualized for the complexity of the research to be undertaken in Research Study 1. CODIT is also loosely framed on LNEC's SIIN2A project (Innovative Information Systems for Smart Water and

Environmental applications), on its forecasting, intelligent data and web platforms current research, as well as the other two LNEC's project described above. However, its scope goes way beyond their scope in the breadth of mathematical, physical and computer sciences involved. A dedicated LNEC research strategy study will be proposed during the current year by the applicant to frame the proposed activities and promote Digital Twins as a leading IT concept in the Hydraulics and Environment research at LNEC.

In addition to their framing in the Maritime Hydraulics research area, these Research Studies are also integrated in the goals of LNEC's current research line "Tools for prediction and analysis" and in the general goals of the draft programme "Industry 4.0" from the 2021-2027 Research Program. The new project to be submitted will be aligned with these research lines.

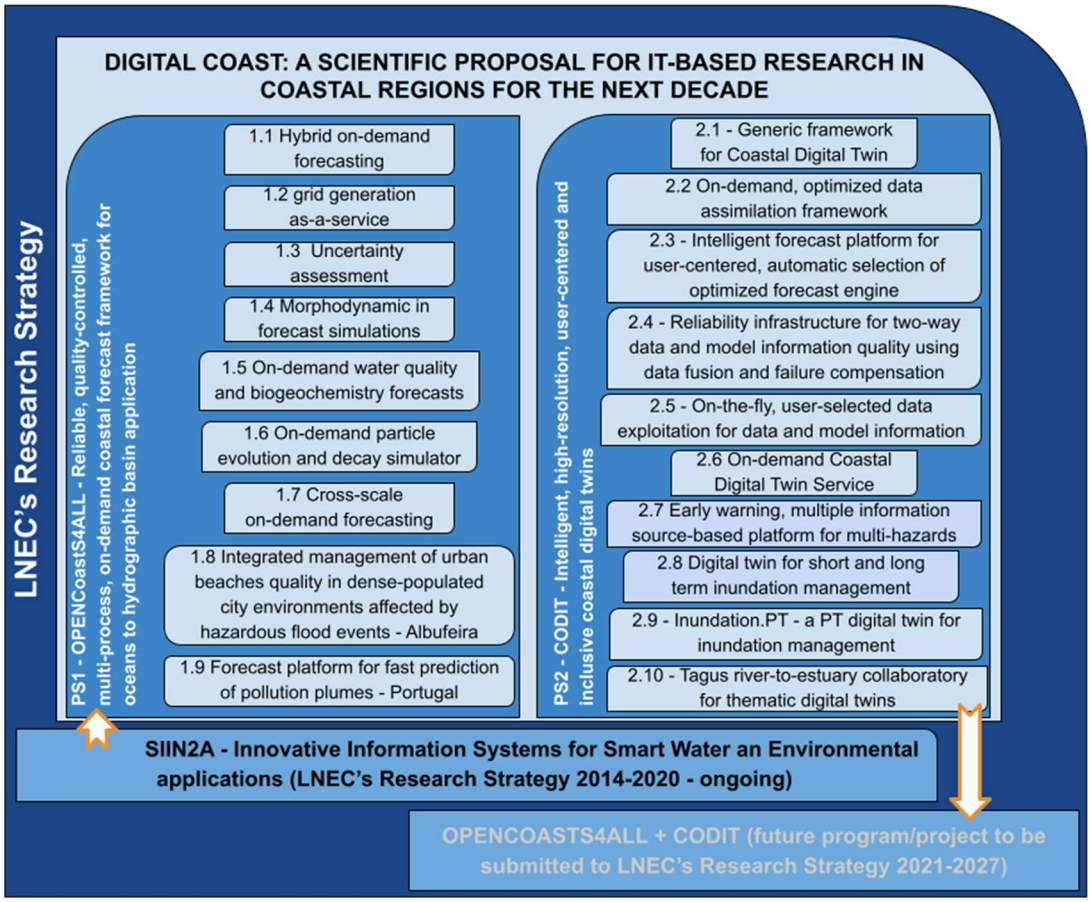


Figure 5.1 – Framing of proposed Research Program in LNEC's research strategy

The full list of Research Studies and associated projects proposed herein are:

- **PS1 - OPENCoastS4ALL - Reliable, quality-controlled, multi-process, on-demand coastal forecast framework for oceans to hydrographic basin application**
    - Project 1.1 - Hybrid on-demand forecasting in coastal regions: a melange of process-based and data-based models
    - Project 1.2 - Computational grid generation as-a-service
    - Project 1.3 - Uncertainty assessment methodology and tool

    - Project 1.4 - Morphodynamic processes integration in forecast simulations. Methodology and tools
    - Project 1.5 - On-demand water quality and biogeochemistry forecasts for multiple applications in coastal regions
    - Project 1.6 - On-demand particle evolution and decay simulator for multiple applications
    - Project 1.7 - Cross-scale on-demand forecasting from the ocean to the hydrographic basin including the urban dimension
    - Project 1.8 - Integrated management of urban beaches quality in dense-populated city environments affected by hazardous flood events - the coastal city of Albufeira case
    - Project 1.9 - Multi-purpose forecast platform for fast prediction of pollution plumes and its application to portuguese estuaries and coasts
- **PS2 - CODIT - Intelligent, high-resolution, user-centered and inclusive coastal digital twins**
    - Project 2.1 - Generic framework for Coastal Digital Twin
    - Project 2.2 - On-demand, optimized data assimilation framework for user-selected multi-source data
    - Project 2.3 - Intelligent forecast platform for user-centered, automatic selection of optimized forecast engine
    - Project 2.4 - Reliability infrastructure for two-way data and model information quality using data fusion and failure compensation
    - Project 2.5 - On-the-fly, user-selected data exploitation for data and model information
    - Project 2.6 - On-demand Coastal Digital Twin Service
    - Project 2.7 - Early warning, multiple information source-based platform for multi-hazards
    - Project 2.8 - Digital twin for short and long term inundation management
    - Project 2.9 - Inundation.PT - a digital twin for inundation management at the Portuguese coast and estuaries
    - Project 2.10 - Tagus river-to-estuary collaboratory for thematic digital twins and collaborative management

The list of the Research Studies and its underlying projects is also presented in Table 5.1. This table also characterizes:

- scope: the breath of the research context: national, european and international;
- potential partnerships: partners for project implementation, organized as academia, industry, managers;
- potential end-users; entities that will benefit from the Research Program outcomes; and
- existing and future funding sources: funding entities that are already or may fund in the future the proposed research, organized by name in Portugal and by typology abroad.

The details of each Study and its associated projects are detailed in the next sections, including goals, methodology, outcomes, temporal planning and associated costs.

The financial plans will be presented in full. The costs can be co-funded by FCT on the grantee monthly stipend and one attendance to a conference, but other funding sources are also available. As the co-funded value will depend on the supporting project's funding and its rules (e.g. FCT, CE, ANI, INTERREG all have different funding schemes), the full budget without any co-funding is presented, organized by typology.

The associated costs consider several typologies and are organized in Table 5.2:

- human resources: defined by LNEC's table of human resources costs, with 100% overheads, and by FCT's cost table for human resources;
- equipment and service provisioning: presented without depreciation but considering the corresponding overheads. Overheads will be included at LNEC rates (20% for travel costs and equipment).
- Missions and field work: presented with a 20% overhead rate as explained above.

Table 5.1 - Research studies and related projects characterization

Research Study	Project	Scope	Partnerships	End-users	Funding sources
PS 1 - OPENCoastS4All - Reliable, cross-scale, multi-process, on-demand coastal forecast framework for oceans to hydrographic basin application	1.1 Hybrid on-demand forecasting	International	Portugal: University of Lisbon, Aveiro, Oporto, Minho and Algarve, IPMA, IH Abroad: NOAA, VIMS, CNRS, UN Decade partners	Port and coastal authorities, APA, researchers, consultancy companies	EC, FCT, INTERREG LNEC, APA, Port & coastal authorities
	1.2 Grid generation as-a-service	International	Portugal: University of Aveiro, Abroad: NOAA, VIMS, CNRS, UN Decade partners	Port and coastal authorities, APA, researchers, consultancy companies	EC, FCT, INTERREG LNEC
	1.3 Uncertainty assessment	International	Portugal: University of Lisbon, Aveiro and Algarve, IPMA, IH Abroad: VIMS, CNRS, UN Decade partners	researchers, consultancy companies	EC, FCT, INTERREG LNEC

Research Study	Project	Scope	Partnerships	End-users	Funding sources
	1.4 Morphodynamic in forecast simulations	International	Portugal: University of Aveiro and Algarve Abroad: VIMS, CNRS	Port and coastal authorities, APA, researchers, consultancy companies	EC, FCT, INTERREG LNEC, APA, Port & coastal authorities
	1.5 On-demand water quality and biogeochemistry forecasts	International	Portugal: University of Lisbon, Aveiro and Algarve, IH Abroad: VIMS, UN Decade partners	Port and coastal authorities, APA, researchers, consultancy companies	EC, FCT, INTERREG LNEC, APA, Port & coastal authorities
	1.6 - On-demand particle evolution and decay simulator for multiple applications	International	Portugal: University of Lisbon, Aveiro and Algarve, IH Abroad: VIMS, UN Decade partners	Port and coastal authorities, APA, researchers, consultancy companies	EC, FCT, INTERREG LNEC, APA, Port & coastal authorities
	1.7 Cross-scale on-demand forecasting	International	Portugal: University of Lisbon, Aveiro, Oporto, Minho and Algarve, IPMA, IH Abroad: NOAA, VIMS, CNRS, UN Decade partners	River, coastal & city authorities, APA, researchers, consultancy companies	EC, FCT, INTERREG LNEC, river, coastal & city authorities
	1.8 Integrated management of urban beaches quality in dense-populated city environments affected by hazardous flood events - Albufeira	International	Portugal: University of Lisbon, Aveiro, Oporto, Minho and Algarve, IPMA Abroad: NOAA, VIMS, CNRS, UN Decade partners	River, coastal & city authorities, APA, researchers, consultancy companies	EC, FCT, INTERREG LNEC, river, coastal & city authorities

Research Study	Project	Scope	Partnerships	End-users	Funding sources
PS2 - CODIT - Intelligent, high-resolution, user-centered and inclusive coastal digital twins	2.1 Generic framework for Coastal Digital Twin	International	Portugal: University of Oporto and Minho Abroad: NOAA, VIMS, UN Decade partners	Researchers	EC, FCT, INTERREG, LNEC
	2.2 On-demand, optimized data assimilation framework	International	Portugal: University of Lisbon, Aveiro and Algarve, IPMA, IH Abroad: NOAA, VIMS, CNRS, UN Decade partners	Researchers, consultancy companies	EC, FCT, INTERREG, LNEC
	2.3 Optimized forecast engine	International	Portugal: University of Lisbon,Oporto and Minho, IPMA, IH Abroad: NOAA, VIMS, CNRS, UN Decade partners	River, coastal & city authorities, APA, researchers, consultancy companies	EC, FCT, INTERREG, LNEC, river, coastal & city authorities
	2.4 Reliability infrastructure	International	Portugal: University of Lisbon, Oporto and Minho, IPMA, IH Abroad: NOAA, UN Decade partners	River, coastal & city authorities, APA, researchers, consultancy companies	EC, FCT, INTERREG, LNEC, river, coastal & city authorities
	2.5 User-selected data exploitation	International	Portugal: University of Lisbon, Oporto, Minho and Algarve, IPMA, IH	River, coastal & city authorities, APA, researchers, consultancy companies	EC, FCT, INTERREG, LNEC, river, coastal & city authorities
	2.6 On-demand Coastal Digital Twin Service	International	Portugal: University of Lisbon, Aveiro, Oporto, Minho and Algarve, IPMA, IH Abroad: NOAA, VIMS, CNRS, UN	River, coastal & city authorities, APA, researchers, water-related and consultancy companies	EC, FCT, INTERREG, LNEC, river, port, coastal & city authorities

Research Study	Project	Scope	Partnerships	End-users	Funding sources
			Decade partners		
	2.7 Early warning, multiple information source-based platform for multi-hazards	International	Portugal: University of Lisbon, CES and Algarve, IPMA Abroad: UN Decade partners	River, coastal & city authorities, APA, researchers, consultancy companies	EC, FCT, INTERREG, LNEC, river, port, coastal & city authorities
	2.8 Digital twin for short and long term inundation management	International	Portugal: University of Lisbon, CES and Algarve, IPMA, IH Abroad: UN Decade partners	River, coastal & city authorities, APA, researchers, consultancy companies	EC, FCT, INTERREG, LNEC, river, port, coastal & city authorities
	2.9 Inundation.PT	National	Portugal: University of Lisbon, CES and Algarve, IPMA, IH	Portuguese water authorities, municipalities, coast-related companies, general public	FCT, INTERREG, LNEC, river, port, coastal & city authorities
	2.10 Tagus river-to-estuary collaboratory	National	Portugal: University of Lisbon, IPMA, IH	Portuguese water authorities, Tagus regional authorities, consulting and Tagus-related companies, general public	FCT, INTERREG, LNEC, river, port, coastal & city authorities

Table 5.2 - Costs for Research studies and related projects implementation

Position at LNEC	Unit monthly cost (euros)
Principal researcher and researcher with habilitation	6500
Senior and Assistant researchers	5700
Computer Science expert	3500
Ph.D grantee (FCT)	1 144,64
M.Sc. grantee (FCT)	875,98

5.3 Research Study 1: OPENCoastS4All - Reliable, cross-scale, multi-process, on-demand coastal forecast framework for oceans to hydrographic basin application

5.3.1 Introduction

This research study aims to address the open questions identified in Chapter 3. It addresses current gaps in coastal forecast frameworks building on top of the OPENCoastS platform for on-demand forecasts, conceptualized by the applicant, and extending it to a broader scope of applications and facilitating its access to all users regardless of their modeling background. The program addresses at the same time forecast quality, through uncertainty consideration, intelligent multi-model usage and data assimilation, taking advantage of current data richness in coastal regions. Automatic grid generation, the major difficulty in the current OPENCoastS platform, will be integrated to make on-demand forecasting accessible to all. It also adapts research from other areas to create a robust, accurate and quality- assessed forecast-engine for application in the scope of Coast Digital Twins, the new IT-based paradigm to support research, management and knowledge creation in coastal regions, addressed in the Research Study 2. The proposed projects within this Study are organized in two dimensions: core developments (CD) and demonstrations to the Portuguese coast (DEMO-PT). The classification is indicated at the end of the project title. Figure 5.2 illustrates the several projects, their workflow and their dependencies.

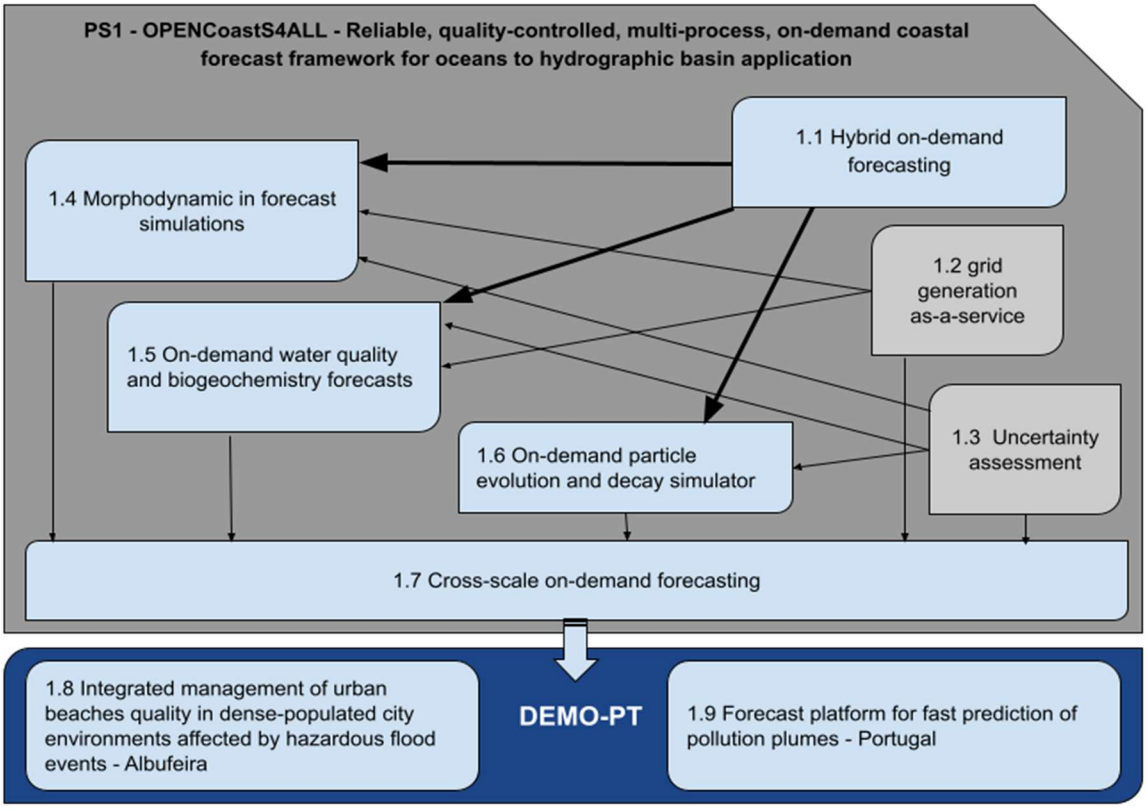


Figure 5.2 - Research study 1 and its projects: workflow and dependencies

5.3.2 Project 1.1 - Hybrid on-demand forecasting in coastal regions: a melange of process-based and data-based models (CD)

5.3.2.1 Rationale

The large computational cost of process-based models can be a limiting factor for very complex simulations, high-resolution large domain forecasts or applications in low income countries where computational resources may be scarce. Hybrid or surrogate models are an attractive alternative several orders of magnitude faster than equivalent numerical models. At the same time, they are also less demanding regarding the volume of field data than pure AI models, and they can be developed in a generic way to be applicable in multiple systems. Training is done using numerical models duly calibrated and applied to the coastal site of interest and the site data. After training is completed, the emulators are able to provide accurate predictions over new forcing conditions. These characteristics make hybrid forecast systems a very attractive solution in particular in emergency context as issuers of early warnings. Having additional time available for population warning or to protect material and ecological assets may have very important economic, social and ecological benefits. Coastal hybrid forecasting has not been developed yet and there is a clear societal need for these tools.

LNEC has developed the world' first relocatable and on-demand coastal forecast web platform, OPENCoastS, through the joint work of the Information Technology Research Group and the Estuaries and Coastal Zones division. OPENCoastS is built on top of the Water Information Forecast Framework (WIFF), also developed at LNEC over the past 12 years. WIFF is a highly sophisticated and flexible forecast framework that can integrate any news models and process solving in a simple and straightforward way. This project aims at merging data, data-based models and their AI tools into coastal forecasting and building a hybrid on-demand forecast framework, based on the current WIFF framework and OPENCoastS platform.

The resulting framework will extend the robustness, accuracy and applicability of coastal forecast systems for multiple purposes and provide the grounds for timely early-warning of coastal phenomena at any time- and space scales and accounting for multiple complex processes. This framework will also be the core engine for the Coastal Digital Twins proposed in the following projects and provide LNEC with an advanced on-demand forecast engine that can be applied to multiple engineering applications.

5.3.2.2 Goals

This project aims at building a robust, quality-controlled, hybrid on-demand forecast framework, starting from OPENCoastS, WIFF and their data lake. The specific goals of the project are:

- to select the adequate AI methodologies for coastal dynamics modeling starting from the analysis developed in the Ph.D. dissertation of Jesus (2019) and the state-of-the-art review on hybrid models of chapter 3;
- to define the suite of open data sources to complement/support the data-based models to be used in the hybrid forecasting and to develop those models;

- to develop the methodology for the scenario building to be used in the AI training procedure, with possible usage of data sources;
- to build an on-demand workflow for user-selected building of a hybrid forecast system at a selected site;
- To compare the several approaches of hybrid modeling in a data-rich system.

To implement these goals, access to high performance computing resources will be requested through both INCD and RNCA. The proposed work will be framed in the European Digital Hub ATTRACT-DIH and the several European and national Projects at LNEC in the area of estuarine and coastal forecasting.

5.3.2.3 Methodology and task description

The methodology comprises the development of the several building blocks that are required for multi-choice hybrid forecasting in coastal regions, and their integration in the WIFF framework and OPENCoastS platform. Each activity aims at achieving each of the specific goals identified above.

Activity 1 - AI framework development

Task 1: Review and identification of relevant AI methodologies and their comparison

Task 2: Implementation of AI methodologies or adaptation of existing tool for coastal dynamics prediction

Task 3: Creation of the on-demand layer on top of the AI tool, by identification of the relevant variables for user selection

Task 4: Setting up of virtual research environment for on-demand AI tool operation in shared e-infrastructures

Activity 2 - Development of data-based models and their integration in the hybrid framework

Task 1: Selection of generic data sources and their possibilities of integration for data-modeling building

Task 2: Comparison of data-model methodology building workflows, selection of the most appropriate ones and their application in the selected AI-tools

Task 3: Validation at selected data-rich site

Activity 3 - Scenario building tool for AI tool training

Task 1: Definition of the methodology for scenario building under multiple forcing conditions and contexts

Task 2: Tool creation for automatic scenario building and implementation of the virtual research environment for simulation operation

Task 3: AI tool training application for hybrid model building

Task 4: Validation in a selected schematic domain

Activity 4 - Integration in OPENCoastS framework

Task 1: Requirements analysis for OPENCoastS integration

Task 2: Development of multi-model component

Task 3: Integration of new data sources in OPENCoastS data

Task 4: Integration of new data- based and hybrid models in OPENCoastS modeling workflow

Task 5: Development of user-selection engine for configuring the modeling type choice for forecast application

Task 6: Setting up of virtual research environment for hybrid forecasting in shared e-infrastructures

Activity 5 - Comparative analysis on accuracy, robustness and performance of process-based, data-model and hybrid modeling on a data rich environment

Task 1: Selection of research question and site for its evaluation

Task 2: Setting up of multiple choice forecasting at selected site

Task 3: Selection of evaluation indicators

Task 4: Application for a selected period and evaluation report

Activity 6 - Dissemination and education

Task 1: Presentation at national and international congresses and conferences

Task 2: Publication in conference proceedings and WoS journals

Task 3: Elaboration of M.Sc. thesis and Ph.D. dissertation

Task 4: User training events

These atividades are expected to be executed over the timeline presented in Table 5.3.

Table 5.3 - Project 1.1: Timeline

Year	Year 1				Year 2				Year 3				Year 4			
Activity	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**
1 - AI framework development																
2 - Data-based models integration in the hybrid framework																
3 - Scenario building tool for AI tool training																

Year	Year 1				Year 2				Year 3				Year 4			
Activity	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**
4 - Comparative analysis on accuracy, robustness and performance																
5 - Integration in OPENCoastS																
6 - Dissemination and education																

5.3.2.4 Expected outcomes

The outcomes of the project will be divided in two categories: contributions to the advance of the state of the art and benefits for LNEC.

In the first category, this research project will produce the following new results:

- on-demand AI tool tailored for coastal environments;
- comparative analysis of AI-, process- and hybrid-based forecasts in data rich real environments;
- tool for scenario building aimed at AI training;
- enhancement of the OPENCoastS service to accommodate the alternative of hybrid modeling.

LNEC will benefit of the project by the availability of the following products for application in coastal engineering research and consultancy projects:

- a new, unique AI tool tailored for coastal studies that can be used to provide faster answers than just the traditional process-based models;
- a new tool for scenario building that can be used for multiple applications in coastal projects such as future engineering works or climate change impacts;
- a more accurate, faster and less computationally intensive OPENCoastS, further contributing to expand LNEC's world's positioning in the prediction area.

5.3.2.5 Resources

The list below summarizes the expected costs associated with this work.

a) Personnel costs
   - 24 months of Ph.D. grant
   - 12 months of senior researchers
   - 2 months of computer expert
   - Total: 102 871,36 euros

b) Travel: 4 participation in conferences (2 national, 2 Europe) - 8500 euros
c) Equipment:
    - laptop - 1500 euros
    - e-infrastructure resources - no costs (assuming that an application to free resources is done in the Rede Nacional de Computação Avançada calls)
d) Open access publication fees: 3000 euros (fee waivers will be promoted, one publication is expected to be in a journal without waivers)
e) Grand total (with overheads): 192 071,36 euros

5.3.2.6 Funding and partnerships

For the funding of the project, applications should be submitted to FCT project and Ph.D. grant calls, to fund the PhD grant and the travel and equipment values. As FCT does not fund personnel with a permanent contract, a research proposal should be submitted to the Horizon Europe, INTERREG, ANI or similar calls to fund or co-fund the senior researchers. The EU MIssions calls are a good example of adequate calls.

Regarding partnerships, many can be established for this project with the partners/funders of OPENCOastS development and follow-up initiatives (CNRS, VIMS, NOAA, UN Decade CoastPredict team) and the entities that are involved in coastal applications (University of Lisbon, Aveiro and Algarve). New partnerships should also be built with end-users and consultancy companies, framed in the scope of the Digital Innovation Hub ATTRACT-DIH (Led by the universities of Minho and Oporto). The H2020 EGI-ACE project will also contribute to some minor features. Finally, forecasting is an important activity in IPMA and IH, the other two state-labs that deal with modeling the coastal zone. Therefore, a long-term collaboration protocol should be built to frame this and most of the projects in this Research Program.

5.3.3 Project 1.2 - Computational grid generation as-a-service (CD)

5.3.3.1 Rationale

Unstructured grid modeling has gained a prominent place in estuarine and coastal dynamics studies for a better representation of complex bathymetry and geometry and of complex vertical exchange processes. The availability of horizontal and vertical computational grids for the systems under study is often a limitation for usage of unstructured grid-based forecast systems by users without a strong knowledge of the processes at stake and of the numerical methods of the modeling tool. The experience of LNEC in the several training events of the OPENCoastS on-demand platform has identified grid generation as the major limitation for its uptake by both coastal managers and researchers from non-modeling backgrounds.

The availability of an automatic grid generator that is provided as a service through an easy-to-use interface is thus an important tool to facilitate access to forecast services as well as to support unstructured grid models usage in general, as it can encapsulate expert knowledge for different type of applications (e.g. 3D baroclinic, wave and current interaction, …). Even for experienced modelers,

building an unstructured grid that provides good results can be a time consuming task, the availability of a high accuracy grid generator being an attractive asset to explore.

As different models may also have different constraints on grid characteristics (e.g. Delaunay triangulations) and allow for different combinations of element type (triangular, quadrangle, ...) and its combinations, the grid-generator-as-a-service should address most common aspects for the creation of a robust representation of the system of interest.

Additionally, the capacity to integrate geometry and bathymetry into the grid and to take them into account in grid generation is also an important aspect. As high resolution bathymetry or configuration for future engineering works in the margins is often not public, the capacity for the user to upload bathymetry/geometry data is also quite relevant.

5.3.3.2 Goals

This project aims at developing a grid generation tool that can be used for circulation, water quality and biogeochemical modeling grid creation. It should be based on the following characteristics, available through user interaction with a simple-to-use platform:

- simulation model and its grid constraints;
- type of simulation (2D/3D, circulation/wave and current/water quality/biogeochemical, ...);
- region of interest to be simulated (the area where results are looked for, not necessarily the grid domain);
- identification of hotspots that will require higher grid resolution (e.g. discharge points);
- sources of geometry/bathymetry or usage of open information data.

Each of the following tasks address the several steps of grid generation outlined above and are a combination of hydraulic modeling knowledge steps and IT tasks. This work is framed in the development of OPENCoastS core capacities in the H2020 EGI-ACE project and will be implemented for application of the modeling suite SCHISM (Zhang et al., 2016). The core grid generator can be based on the work developed by NOAA in the OCSMesh code or similar works.

5.3.3.3 Methodology and task description

The methodology comprises the development of the several building blocks that are required for a grid generator for forecasting in coastal regions, and their integration in the OPENCoastS platform.

Activity 1 - Requirements analysis for unstructured grid generation

Task 1: requirements analysis for multiple type of simulations

Task 2: requirements analysis for selected models horizontal and vertical grid generation

Task 3: specification of grid quality indicators as a function of application type

Activity 2 - Selection of automatic grid generator and adaptation for present purpose

Task 1: review of grid generators and selection of the most appropriate for the problem at stake

Task 2: extension/adaptation of the selected tool for Activity 1's requirements analysis

Task 3: validation in selected applications against existing expert grid modeling

Activity 3 - Integration in OPENCoastS framework

Task 1: Requirements analysis for OPENCoastS integration

Task 2: Integration of the open source bathymetry and coastal geometry databases and development of the bathymetry dashboard, including upload of bathymetry/geometry layers

Task 3: Integration of the selected grid generation engine in OPENCoastS and development of the user-selection grid-generation dashboard

Task 6: Setting up of virtual research environment for grid generation in shared e-infrastructures

Task 7: Testing for simple cases of hydrodynamic and water quality simulations

Activity 4 - Dissemination and education

Task 1: Presentation at national and international congresses and conferences

Task 2: Publication in conference proceedings and WoS journals

Task 3: Elaboration of a M.Sc. thesis

These atividades are expected to be executed over the timeline presented in Table 5.4.

Table 5.4 - Project 1.2: Timeline

Year	Year 1				Year 2				Year 3			
Activity	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**
1 - Requirements analysis for unstructured grid generation												
2 - Selection of automatic grid generator and adaptation for present purpose												
3 - Integration in OPENCoastS framework												
6 - Dissemination and education												

5.3.3.4 Expected outcomes

The outcomes of the project will be divided in two categories: contributions to the advance of the state of the art and benefits for LNEC.

In the first category, this research project will create a novel automatic grid generator applied to any type of simulations in coastal regions

LNEC will benefit of the project by:

- Obtaining an automatic grid generator that can be applied to any consultancy projects, facilitating and speeding-up the development of any modeling studies;
- implement automatic generation in OPENCoastS thus increasing its visibility and usefulness for end-users;
- increase its outreach by brand-developing a very useful tool for researchers, consultancy companies and end-users.

5.3.3.5 Resources

The list below summarizes the expected costs associated with this work.

a) Personnel costs
    - 12 months of M.Sc. grant
    - 1,5 months of a senior researcher
    - 1 month of computer expert
    - Total: 22 561,76 euros
b) Travel: 1 participation in european conference - 3000 euros
c) Equipment:
    - laptop - 1500 euros
    - e-infrastructure resources - no costs (assuming that an application to free resources is done in the Rede Nacional de Computação Avançada calls)
d) Grand total: 38 211,76 euros

5.3.3.6 Funding and partnerships

For the funding of the project, an application should be submitted to FCT project calls, to fund the M.Sc. grant and the travel and equipment values. As FCT does not fund personnel with a permanent contract, a research proposal can be submitted to modeling related calls such as CMEMs ones, in partnership with Portuguese or international institutions involved in unstructured grid modeling. Given the importance of the tool for LNEC's consultancy studies, a contribution from LNEC's own funds may be adequate.

5.3.4 Project 1.3 - Uncertainty assessment methodology and tool (CD)

5.3.4.1 Rationale

Uncertainty evaluation is seldom a part of forecast systems workflow, potentially jeopardizing their application for many problems. Forecast uncertainty may arise from physical processes representation and numerical formulations, as well as in the construction and application of the numerical models including uncertainties in initial conditions, model boundary conditions and model parameters. As models become more complex, for instance by integrating waves, morphodynamic or biogeochemical

simulations, uncertainty importance grows as errors propagate and are amplified in the cascade of model simulations.

Recently, ensemble modeling is emerging as one of the best solutions for uncertainty as it can minimize the combined uncertainty in input data, model parameters and model structure, shown to improve forecast performance. In addition to the use of multiple models and the construction of a weighted ensemble, giving more weight for better performing models, the identification of the dependency of these weights with environmental and other system conditions, creating on-the-fly, optimized weighted ensembles. These ensembles consider multiple external forcings (either river or atmospheric), multiple model configurations or multiple models usage. Ensemble modeling should be supported by on-going forecast evaluation of accuracy and timeliness, based on a detailed suite of error measures, calculated against real time data networks.

5.3.4.2 Goals

This project aims at developing an uncertainty tool that can be used for multiple types of simulations (hydrodynamic, water quality, biogeochemistry,...) comprising all uncertainty sources. It will be integrated in a user interface for configuration of the uncertainty methodology to be implemented in each forecast application, as part of the OPENCoastS platform. In order to obtain an optimized uncertainty workflow, this tool should also provide an opportunity for the user to define which uncertainty sources should be considered and allow for the definition of the error measure indicators and the sources for the real time data networks to be used in the forecast error performance analysis.

The specific goals of the project are:

- to select the main sources of uncertainty in coastal forecasts
- to select a suite of error measures and an ensemble methodology for multi-model, multi-parameter and multi-input uncertainty assessment
- to develop an optimized algorithm for uncertainty quantification under distinct environmental conditions
- to implement the above products in OPENCoastS as part of the quality assurance and robustness of the prediction framework

5.3.4.3 Methodology and task description

The methodology comprises the development of the several building blocks of an uncertainty engine for application in forecasting coastal regions, and their integration in the OPENCoastS platform.

Activity 1 - Detailed requirements analysis for uncertainty evaluation

Task 1: Definition of the uncertainty sources for each type of simulation, from processes to model configuration

Task 2: Definition of the sources of data and error measure suite for forecast error evaluation according to simulation type

Task 3: Selection of a range of models for ensemble modeling according to simulation type

Task 4: Definition of environmental and system conditions to be used in the optimization procedure according to simulation type

Task 5: Definition of algorithm for optimization implementation

Activity 2 - Optimized ensemble engine development

Task 1: Implementation of ensemble engine, based on requirement analysis and available open software

Task 2: Test of ensemble engine for each type of simulation using a data-rich environment

Task 3: Development of optimized ensemble engine

Task 4: Test of optimized ensemble engine in several complex use cases with multiple data sources

Activity 3 - Integration in OPENCoastS framework

Task 1: Requirements analysis for OPENCoastS integration

Task 2: Development of user interface for uncertainty configuration

Task 3: Integration of uncertainty component (multi-simulation) engine

Task 4: Integration of user-selection engine for configuring the choice of uncertainty components for forecast application

Task 5: Setting up of virtual research environment for uncertainty simulations in shared e-infrastructures

Activity 4 - Dissemination and education

Task 1: Presentation at national and international congresses and conferences

Task 2: Publication in conference proceedings and WoS journals

Task 3: Elaboration of M.Sc. thesis and Ph.D. dissertation

These atividades are expected to be executed over the timeline presented in Table 5.5.

Table 5.5 - Project 1.3: Timeline

Year	Year 1				Year 2				Year 3				Year 4			
Activity	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**
1 - Detailed requirements analysis for uncertainty evaluation																
2 - Optimized ensemble engine development																

Year	Year 1				Year 2				Year 3				Year 4			
Activity	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**
3 - Integration in OPENCoastS framework																
4 - Dissemination and education																

5.3.4.4 Expected outcomes

The outcomes of the project will be divided in two categories: contributions to the advance of the state of the art and benefits for LNEC.

In the first category, this research project will:

- novel methodology for uncertainty assessment covering all error sources
- optimized engine for uncertainty evaluation and proposal of optimal ensembling building depending on the current environmental conditions

LNEC will benefit of the project by:

- increasing the accuracy and robustness of the OPENCoastS predictions, by evaluating uncertainty and expanding the prediction suite to ensembles and optimal ensemble creation
- by using the optimal ensembling engine on non-forecast simulations also, LNEC will be able to provide a high level of reliability on its modeling studies

5.3.4.5 Resources

The list below summarizes the expected costs associated with this work.

f) Personnel costs
   - 48 months of Ph.D. grant
   - 12 months of senior researchers
   - 3 months of computer expert
   - Total: 133 842,72 euros

g) Travel: 4 participation in conferences (2 national, 2 Europe) - 8500 euros

h) Equipment:
   - laptop - 1500 euros
   - e-infrastructure resources - no costs (assuming that an application to free resources is done in the Rede Nacional de Computação Avançada calls)

i) Open access publication fees: 3000 euros (fee waivers will be promoted, one publication is expected to be in a journal without waivers)

j) Grand total (with overheads) : 226 542,72 euros

5.3.4.6 Funding and partnerships

For the funding of the project, applications should be submitted to FCT project and grant calls, to fund the PhD grant and the travel and equipment values. As FCT does not fund personnel with a permanent contract, a research proposal should be submitted to the Horizon Europe, INTERREG, ANI or similar calls to fund or co-fund the senior researchers.

Partnerships can be established for this project with the IPMA and IH as both have the same time of concerns for their forecasting activities. The more scientific tasks will benefit partnerships from coastal modeling related universities (University of Lisbon, Aveiro and Algarve). Given that SCHISM will be one of the primary models VIMS and NOAA will be definitely important partners, framed in the context of the Ocean Modeling Collaboration Forum. Finally, the UN Decade CoastPredict team are also interesting partners for a joint European project given their experience in high resolution coastal modeling.

5.3.5 Project 1.4 - Morphodynamic processes integration in forecast simulations. Methodology and tools (CD)

5.3.5.1 Rationale

Maintaining the accuracy of the bathymetric conditions is a fundamental part of the long term operation of coastal forecast systems, in particular after highly energetic events, such as storms, and anthropogenic interventions, such as channel deepening. Furthermore, the bathymetric evolution can alter radically the nature of coastal systems, limiting for instance the capacity to exchange water between the coast and coastal lagoons, due to the migration and closure of lagoon inlets. Automatic ways to update the topography and bathymetry in coastal and estuarine forecasts are the way to proceed and several efforts are starting to be made through integration of remote sensing data from satellites and camaras for bathymetric update in hydrodynamic models, through the inclusion of morphodynamic models in the forecast workflows and by combining both approaches through data assimilations. As current forecast systems do not currently possess the capacity to accurately assess bathymetry evolution and to include it in the modeling framework, a new approach is proposed here to accomplish that goal for non-cohesive sediments, taking advantage of the several data sources openly available and high resolution numerical models. As several process-based morphodynamic models are currently available, the possibility for selecting and comparing bathymetric predictions will be included through both ensembles evaluated against usage of single models. Remote sensing data, from satellites and camaras, will also be explored for data assimilation in a combined way into morphodynamic forecasts. The work will start with the identification of the controlling bathymetric features that need to be accurately represented in the forecast procedure to prevent modeling divergence from reality. Assimilation from multiple data sources will be used in multiple morphodynamic forecasts including several models and their ensemble.

5.3.5.2 Goals

This project aims at integrating morphodynamic prediction and bathymetry update in a generic forecast framework, taking advantage of the several models and data sources available, and exploring technologies such as model ensembling and data assimilation to create an innovative tool that can accurately handle coastal prediction for long term periods without robustness issues that are quite common in morphodynamic modeling.

It will be integrated in a user interface for configuration of the morphodynamic prediction and data sources to be used in the assimilation procedure, to be implemented in each forecast application, as part of the OPENCoastS platform. In order to obtain a user-centered workflow, this tool should also provide an opportunity for the user to define how morphodynamic modeling should be considered and allow for the definition of the error measure indicators and the sources for the real time data networks to be used in the assimilation methodology.

The specific goals of the project are:

- to select a suite of morphodynamic models and to build model ensembles;
- to detect the most important features controlling morphodynamic evolution and to define a methodology for its proper representation using a combination of process-based and AI models, eventually complemented with existing robustness techniques;
- to integrate the morphodynamic engine in OPENCoastS, as one of the workflow components if selected by the user.

This work is framed in the continuation of the activities of the FCT project MOSAIC.pt, which is addressing the prediction of erosion in coastal regions through a suite of numerical models for region to local hydrodynamic prediction and estimation of coastal morphodynamics. It aims at extending this analysis to include several data sources for assimilation and to complement the SCHISM+XBeach workflow with additional models and ensemble building.

5.3.5.3 Methodology and task description

The methodology comprises the development of an innovative morphodynamic engine and its validation in data-rich use cases. It also includes their integration in the OPENCoastS platform. The following activities and tasks are organized sequentially to achieve the above goals.

Activity 1 - Core components analysis for morphodynamic forecasting

Task 1: Identification of the main bathymetric features to be detected, monitored and accurately predicted in morphodynamic predictions

Task 2: Definition of the methodology for bathymetry features data integration from multiple sources of remote (cameras, satellite, drones) and in-situ data

Task 3: Definition of the methodology for generic morphodynamic forecast based on numerical models, AI methods and robustness tools

Task 3: State-of-the-art-review of open access morphodynamic models, their range of applicability and selection of the components for the forecast engine

Task 4: Definition of the methodology for morphodynamic ensemble building and error measure selection for accuracy evaluation

Activity 2 - Development of morphodynamic engine for single model and ensemble simulations

Task 1: Implementation of a common framework for morphodynamic model operation under common formats and inputs

Task 2: Integration of the selected morphodynamic models in framework

Task 3: Development of the ensemble predictor based on any combination of models

Task 4: Testing of morphodynamic engine in data-rich use cases

Activity 3 - Development of assimilation strategy for bathymetry update in the morphodynamic engine

Task 1: Review of bathymetric assimilation procedures and selection of the most appropriate ones for the several combinations of remote and in-situ data sources

Task 2: Implementation of the assimilation procedures

Task 3: Testing of assimilation strategy in data-rich use cases

Activity 4 - Integration of morphodynamic forecasting in OPENCoastS framework

Task 1: Requirements analysis for OPENCoastS integration

Task 2: Integration of new data sources/processing tools for morphodynamic module usage

Task 3: Integration of assimilation tool for bathymetry update

Task 4: Integration of morphodynamic engine in OPENCoastS

Task 5: Development of user-selection engine for specifying morphodynamic update choice

Task 6: Setting up of virtual research environment for morphodynamics simulation and assimilation in shared e-infrastructures

Activity 5 - Dissemination and education

Task 1: Presentation at national and international congresses and conferences

Task 2: Publication in conference proceedings and WoS journals

Task 3: Elaboration of M.Sc. thesis and Ph.D. dissertations

Task 4: User training in OPENCoastS morphodynamic platform

These atividades are expected to be executed over the timeline presented in Table 5.6.

Table 5.6 - Project 1.4: Timeline

Year	Year 1				Year 2				Year 3			
Activity	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**
1 - Core components analysis for morphodynamic forecasting												
2 - Development of morphodynamic engine for single model and ensemble simulations												
3 - Development of assimilation strategy for bathymetry update in the morphodynamic engine												
4 - Integration in OPENCoastS framework												
5 - Dissemination and education												

5.3.5.4 Expected outcomes

The outcomes of the project will be divided in two categories: contributions to the advance of the state of the art and benefits for LNEC.

In the first category, this research project will:

- create and develop a unique morphodynamic forecast engine for single model and ensemble coastal simulations
- expand the on-demand forecast service OPENCoastS to become the first on-demand forecast framework that can be applied to problems with highly sediment dynamics environments

LNEC will benefit of the project by:

- having the capacity and tools to predict highly sediment dynamics environments in real time
- expand OPENCoastS capacities to a unique position in operational oceanography

5.3.5.5 Resources

The list below summarizes the expected costs associated with this work.

a) Personnel costs
    - 12 months of M.Sc. grant
    - 6 months of a senior researcher
    - 2 month of computer expert
    - Total: 51 711,76 euros

b) Travel: 1 participation in european conference - 3000 euros
c) Equipment:
    - laptop - 1500 euros
    - e-infrastructure resources - no costs (assuming that an application to free resources is done in the Rede Nacional de Computação Avançada calls)
d) Grand total: 96 511,76 euros

5.3.5.6 Funding and partnerships

For the funding of the project, applications should be submitted to FCT projects to fund the M.Sc. grant and the travel and equipment values. As FCT does not fund personnel with a permanent contract, a research proposal should be submitted to the Horizon Europe, INTERREG, ANI or similar calls to fund or co-fund the senior researcher and the computer expert. Given the importance of morphodynamics on hazardous events forecasts and the possibility of having much more accurate predictions of storm impacts, a co-funding should also be requested to the coastal authorities (APA) and coastal municipalities. The concept of collaborative funding, successfully applied at the urban water division of LNEC, can be an avenue to be explored for this project.

Partnerships can be established for this project with the several teams working in morphodynamics in Portugal: University of Lisbon, Aveiro and Algarve. Given that SCHISM will be one of the primary models in this project, VIMS and NOAA will be definitely important partners, framed in the context of the Ocean Modeling Collaboration Forum. CNRS-La Rochelle and the Prof. Xavier Bertin team, long term collaborators of LNEC in this area, are a core partner for any proposals in morphodynamics. They were also a partner in the development of OPENCoastS and SCHISM's morphodynamic models.

5.3.6 Project 1.5 - On-demand water quality and biogeochemistry forecasts for multiple applications in coastal regions (CD)

5.3.6.1 Rationale

Accurately forecasting water quality and biogeochemistry in estuarine and coastal regions is necessary to address both long term trends that may have serious impact in ecological assets and to support short term emergency actions under hazardous events such as coastal eutrophication and hypoxia, exposure to waterborne pathogens or harmful algal blooms.

Today's availability of multiple sources of data, including complex and low cost in-situ sensors, remote sensing or water drones, are providing the channels to obtain the information necessary to build hybrid modeling tools and to perform accurate assessments of predictions in comparison with real time data representing well the small spatial and temporal processes scales of estuarine and coastal environments or to increase the accuracy of predictions through data assimilation. Two more complex and important challenges arise in this area. The first is related to data management and data lakes creation, as a backbone for the necessary data-sharing in this highly expensive monitoring area. The second, and even more important one, is the automatic assessment of data quality in order for data to be used in the above procedures. Exporting the concepts and tools developed in Jesus (2019) for error

detection and correction and data classification is a necessary measure for the achievement of reliable water quality and biogeochemistry predictions.

At the same time, the growth of computational resources and the growth of shared e-infrastructures facilitating access to large HPC resources, has paved the way to address uncertainty and to perform long term and scenarios analysis to investigate climate changes or anthropogenic interventions impacts.

A new framework for water quality and biogeochemistry prediction is proposed herein to address all these issues, targeting a flexible, accurate and user-configurable forecast framework that is rooted on the work already developed in the scope of the OPENCoastS platform for on-demand forecasting.

5.3.6.2 Goals

This project aims at developing a generic water quality and biogeochemistry forecast framework and customizing it for specific problems such as algae blooms or oil spills, by integrating initial conditions generators based on remote sensing. The forecast framework will also support the capacity to perform uncertainty analysis by integrating multiple models (data-based, process-based and hybrid models) and ways to assess uncertainty in initial and boundary conditions and model parameters. Exploring technologies such as data assimilation will also be done to create an innovative tool that can accurately handle predictions for long term periods.

This forecast engine and its services will be integrated in a dedicated dashboard in the OPENCoastS platform, to facilitate access and configuration for the new features.

The specific goals of the project are:

- to select a suite of water quality and biogeochemistry models, including data-based, process-based and hybrid, for usage in the forecast engine, including model ensembles;
- to define tailored uncertainty strategy (based on project 1.3) and assimilation procedure and implement them as tools;
- to define and implement generators of initial conditions for selected problems (algae blooms, oil spills, …);
- to integrate all these tools in OPENCoastS, building a dedicated dashboard for these simulations.

This work aims at extending the work developed in the FCT UBEST project, recently finished. Its applicability falls in the scope of the newly approved ATTRACT DIH national and european projects.

5.3.6.3 Methodology and task description

The methodology comprises the development of an innovative water quality and biogeochemistry forecast engine, including modules for data assimilation and uncertainty analysis, and its validation in data-rich use cases. It also includes their integration in the OPENCoastS platform. The following activities and tasks are organized sequentially to achieve the above goals.

Activity 1 - Requirement analysis for on-demand water quality and biogeochemistry forecasts

Task 1: Definition of the parameters for pre-selected forecast typologies and selection of the models formulations for these typologies

Task 2: Definition of the parameters for selection in the on-demand forecast engine for pre-defined typologies

Task 2: Definition of the generic parameters for selection in the on-demand forecast engine for user-defined tracer

Activity 2 - Development of initial conditions remote sensing data-based processing tool for selected typologies

Task 1: Definition of the data sources for each tracer typology

Task 2: Development of data acquisition and input file preparation scripts

Task 3: Creation of data repository for each typology of tracer

Task 4: Testing of tool in selected use cases

Activity 3 - Development of assimilation module for water quality and biogeochemistry

Task 1: Selection of assimilation workflow and data sources for each typology

Task 2: Implementation of tailored module for each typology

Task 3: Implementation of user-selected workflow module

Task 4: Testing of final data layers for each typology in selected uses cases

Activity 4 - Development of water quality and biogeochemistry forecast engine

Task 1: Implementation of a common framework for water quality and biogeochemistry model operation under common formats and inputs

Task 2: Integration of the selected models in framework

Task 3: Development of the ensemble predictor based on any combination of models

Task 4: Customization of predictor to address specific problems including integration of initial conditions tailored generator

Task 4: Testing of forecast engine in data-rich use cases

Activity 5 - Integration of the uncertainty and assimilation modules in the on-demand water quality and biogeochemistry forecast engine

Task 1: Integration of uncertainty module from project 1.3 to water quality and biogeochemistry simulations

Task 2: Integration of data sources (if not yet available through other presently proposed projects) and assimilation procedure in predictios

Task 3: Evaluation of the performance of the several options using multiple error measures

Activity 6 - Integration in OPENCoastS framework

Task 1: Requirements analysis for OPENCoastS integration

Task 2: Integration of water quality and biogeochemistry engine in OPENCoastS

Task 3: Uncertainty assessment tool integration in water quality workflow and dashboard

Task 4: Integration in OPENCoastS of new water quality/biogeochemistry data sources/processing tools for model evaluation and assimilation module usage

Task 5: Development of dashboard for user-selection engine for specifying water quality and biogeochemistry problem choice

Task 6: Setting up of virtual research environment for water quality and biogeochemistry in shared e-infrastructures

Activity 7 - Dissemination and education

Task 1: Presentation at national and international congresses and conferences

Task 2: Publication in conference proceedings and WoS journals

Task 3: Elaboration of M.Sc. thesis and Ph.D. dissertations

Task 4: User training

These atividades are expected to be executed over the timeline presented in Table 5.7.

Table 5.7 - Project 1.5: Timeline

Year	Year 1				Year 2				Year 3				Year 4				Year 5	
Activity	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**	**1**	**2**
1 - Requirements analysis for on-demand water quality and biogeochemistry forecasts																		
2 - Development of initial conditions remote sensing data-based processing tool for selected typologies																		
3 - Development of assimilation module																		

Year	Year 1				Year 2				Year 3				Year 4				Year 5	
Activity	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**	**1**	**2**
for water quality and biogeochemistry																		
4 - Development of innovative water quality and biogeochemistry forecast engine																		
5 - Integration of the uncertainty and assimilation modules in the on-demand forecast engine																		
6 - Integration in OPENCoastS framework																		
7 - Dissemination and education																		

5.3.6.4 Expected outcomes

The outcomes of the project will be divided in two categories: contributions to the advance of the state of the art and benefits for LNEC.

In the first category, this research project will develop:

- a new framework for water quality and biogeochemistry prediction, targeting a flexible, accurate and user-configurable forecast framework, and implement it as complex, comprehensive water quality and biogeochemistry forecast tool integrated in an on-demand platform
- new modules for initial conditions for specific water quality problems such as algae blooms or oil spills based on remote sensing, to be used in the previous outcome

LNEC will benefit of the project by getting:

- a unique forecast platform for water quality and biogeochemistry to be used in its consultancy and research projects
- an enhancement of the current usage of OPENCoastS for water quality problems, expanding its users base

5.3.6.5 Resources

The list below summarizes the expected costs associated with this work.

a) Personnel costs
    - 48 months of Ph.D. grant
    - 12 months of a senior researcher
    - 3 months of computer expert
    - Total: 133 842,72 euros
b) Travel: 4 participation in conferences (2 national, 2 Europe) - 8500 euros
c) Equipment:
    - laptop - 1500 euros
    - e-infrastructure resources - no costs (assuming that an application to free resources is done in the Rede Nacional de Computação Avançada calls)
d) Grand total (with overheads): 226 542,72 euros

5.3.6.6 Funding and partnerships

For the funding of the project, applications should be submitted to FCT projects and grant calls, to fund the PhD grant and the travel and equipment values. As FCT does not fund personnel with a permanent contract, a research proposal should be submitted to the Horizon Europe, INTERREG, ANI or similar calls to fund or co-fund the senior researchers. The EU MIssions calls are a good example of adequate calls as well as INTERREG. Given the importance of water quality forecasts and the possibility of having much more accurate predictions of contamination a co-funding should also be requested at the coastal authorities (APA), municipalities and wastewater companies. The concept of collaborative funding, successfully applied at the urban water division of LNEC, can be an avenue to be explored for this project.

Regarding partnerships, the Algarve, Aveiro and Lisbon university stand out as the most adequate taking into account recent collaborations in the scope of the UBEST, CONPRAR and the Ph.D. and pos-doc of Dr. Marta Rodrigues. Collaborations with the computer science team of the AQUAMON project are also adequate for the technological components. Finally, collaborations can also be established for this project with the partners/funders of OPENCOastS development and follow-up initiatives (VIMS, NOAA). New partnerships should also be built with end-users and consultancy companies, framed in the scope of the Digital Innovation Hub ATTRACT-DIH (led by the universities of Minho and Oporto). The H2020 EGI-ACE project will also contribute to some minor features. Finally, water quality forecasting is an important activity in IPMA and IH, the other two state-labs that deal with modeling the coastal zone. Therefore, a long-term collaboration protocol should be built to frame this and most of the projects in this Research Program. In the oil spill area, collaborations with the Port of Aveiro and the Aveiro university, in Portugal, and the IH Cantabria in Spain, will be promoted as follow up on the SPRES and pac:man projects.

5.3.7 Project 1.6 - On-demand particle evolution and decay simulator for multiple applications (CD)

5.3.7.1 Rationale

Given the limited capacity for water quality and biogeochemical process representation, the usage of complex, time- and resource-consuming models may not always be the best approach for some water quality applications. Likewise, some uncertainty or comprehensive scenario database building requiring thousands of simulations may be undoable in a reasonable time frame with domain discretization methods. The development of a particle forecast engine, built on top of a flexible circulation forecast, that could handle multiple types of simulations and allow for customization on-the-fly by the users, can be a valuable asset for many coastal management actions and for scenario-based analysis. Integrating uncertainty in forcing conditions in the simulations as well as adequate decay formulations may also increase the quality of the simple particle models.

While particle models are available in forecast systems, a configurable engine that could simulate oil spills, larvae movement, marine debris pathways, water discharges and other processes selected or configured by the user is not yet available. Moreover, linking these models with remote sensing data sources that could provide adequate initial conditions and accounting for uncertainty are not available either in particle forecasting. Such a flexible and user-centered approach can be a valuable asset for coastal managers and researchers.

5.3.7.2 Goals

This project aims at developing a user configurable particle forecast engine, applicable for multiple water quality and biogeochemical problems, forced by the relevant circulation conditions, and its integration in the OPENCoastS framework, thus taking advantage of the available flexible, high accuracy circulation prediction forcing.

The specific goals of the project are:

- to select a suite of pre-defined relevant transport and transformation problems, such as oil spills, marine debris pathways, algae blooms, larvae movement, and to select the most accurate formulations for their pathways and evolution;
- to define adequate sources for initial conditions for these problems, based on remote sensing data and development of the initial condition forecast engine;
- to define the relevant characteristics in generic transport and transformation problems and to develop a user configurable particle forecast engine;
- to integrate the above tools in OPENCoastS, taking advantage of the multiple circulation options, and create a particle transport and evolution forecast dashboard.

5.3.7.3 Methodology and task description

The methodology comprises the development of a flexible and user configurable particle forecast engine, including modules for specific problems such as oil spills or marine debris tailing, and its validation in data-rich use cases. It also includes their integration in the OPENCoastS platform. The following activities and tasks are organized sequentially to achieve the above goals:

Activity 1 - Requirement analysis for configurable particle forecast engine

Task 1: Definition of the range of particle conditions and process type formulations for configurable particle runs characterization

Task 2: Definition of the parameters for user configuration for integration in the dedicated dashboard in OPENCoastS

Task 3: Definition of the predefined application types and selection of the corresponding most accurate formulations

Task 4: Definition of data sources for initial conditions establishment

Activity 2 - 2D and 3D particle model engine development with multi-mode flexible problem options and pre-defined application types

Task 1: Development of generic particle forecast engine with user-specified configuration or predefined typology

Task 2: Development of pre-defined typology forecast engine

Task 3: Integration of uncertainty option through multi-scenario or multi-model approaches

Task 3: Testing for multiple typology cases with available field data

Activity 3 - Development of initial conditions provider engine

Task 1: Development of initial conditions engine for selected pre-defined transport typologies (oil spills, algae blooms, …)

Task 2: Integration in particle forecast engine

Task 3: Testing in selected use cases with available field data

Activity 4 - Integration in OPENCoastS framework

Task 1: Requirements analysis for OPENCoastS integration

Task 2: 2D and 3D particle model engine integration with multi-mode flexible problem configuration and uncertainty tool integration (linked to project 1.3)

Task 3: Development of pre-defined engine for oil spill, marine litter, microplastic and other pre-defined typology dynamics in operational or scenario mode

Task 4: Development of user-selection dashboard for specifying scalar problem choice and configuration of decay and evolution properties

Task 7: Setting up of virtual research environment for multi-mode particle simulations in shared e-infrastructures

Activity 5 - Dissemination and education

Task 1: Presentation at national and international congresses and conferences

Task 2: Publication in conference proceedings and WoS journals

Task 3: Elaboration of M.Sc. thesis

These atividades are expected to be executed over the timeline presented in Table 5.8.

Table 5.8 - Project 1.6: Timeline

Year	Year 1				Year 2				Year 3			
Activity	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**
1 - Requirement analysis for configurable particle forecast engine												
2 - 2D and 3D particle model engine development with multi-mode flexible problem options and predefined application types												
3 - Development of initial conditions provider engine												
4 - Integration in OPENCoastS framework												
5 - Dissemination and education												

5.3.7.4 Expected outcomes

The outcomes of the project will be divided in two categories: contributions to the advance of the state of the art and benefits for LNEC.

In the first category, this research project will provide:

- a novel, fast user configurable particle forecast engine, forced by user-specified, on-demand hydrodynamic conditions, with several predefined options
- a new hybrid initial conditions engine for several predefined transport and decay problems

LNEC will benefit of the project by:

- adding 2D and 3D fast but detailed transport and decay particle models to water quality module of OPENCoastS, of particular relevance for consultancy projects.

5.3.7.5 Resources

The list below summarizes the expected costs associated with this work.

a) Personnel costs
    - 12 months of M.Sc. grant

- 4 months of a senior researcher
- 2 months of computer expert
- Total: 40 311,76 euros

b) Travel: 2 participation in european conference - 6000 euros
c) Equipment:
    - laptop - 1500 euros
    - e-infrastructure resources - no costs (assuming that an application to free resources is done in the Rede Nacional de Computação Avançada calls)
d) Grand total: 77 311,76 euros

5.3.7.6 Funding and partnerships

For the funding of the project, applications should be submitted to FCT projects to fund the M.Sc. grant and the travel and equipment values. As FCT does not fund personnel with a permanent contract, a research proposal should be submitted to the INTERREG, ANI or similar calls to fund or co-fund the senior researchers and computer experts.

Partnerships can be established for this project with the several teams working in water quality in Portugal: University of Lisbon, Aveiro and Algarve. This new module also constitutes an attractive add-on to the Digital Twin efforts, so collaborations with the CoastPredict team are envisioned.

5.3.8 Project 1.7 - Cross-scale on-demand forecasting from the ocean to the hydrographic basin including the urban dimension (CD)

5.3.8.1 Rationale

Coastal and estuarine dynamics are strongly coupled with the bordering water systems, in terms of circulation, sediment dynamics and biogeochemistry. In order to obtain reliable and accurate predictions in these regions, the consideration of the exchanges with the atmosphere, ocean, riverine and land inputs at the adequate spatial and temporal scales at each compartment is fundamental. For instance, close to the coast, biogeochemical processes show large variability and strong gradients, due to the interactions with land and inland waters, requiring high confidence and timeliness predictions of these nearby domains.

The recent availability of multiple sources of data across domains and high resolution modeling suites that are able to address cross-scale domains at the necessary time and spatial has paved the way to build integrated prediction systems from the ocean to the watershed including the urban dimension.

While computational costs concern in these simulations are strongly alleviated with the availability of shared computational infrastructures, the uncertainty requirements at its many dimensions becomes critical as errors propagate across the modeling chain. To minimize error generation and propagation, assimilation strategies with data from the several compartments becomes a major asset to achieve high accuracy predictions. Therefore, this project should be built on top of the outcomes of the previously proposed projects, closing the loop of development and implementation of reliable, robust, accurate and problem-driven on-demand coastal forecasting tools proposed in this first Research Study of Research Program.

5.3.8.2 Goals

This project aims at developing a cross-scale forecast framework that integrates processes from the ocean to the watershed, including the urban dimension. It will be built integrating the outcomes of the previous projects, accounting for uncertainty and integrating data assimilation procedures, applicable from circulation to biogeochemical problems, forced by the relevant circulation conditions, and its integration in the OPENCoastS framework, thus taking advantage of the available flexible, high accuracy circulation prediction forcing.

The specific goals of the project are:

- to define a framework for integrated forecasting across river-to-ocean scales to predict jointly the dynamics of the adjoint water bodies using a combination of models at the correct time and space scales.
- to select the modeling suite, combining data-models, process-based and hybrid ones to build a model cascade, linking in a bidirectional way with compartments at larger scales (e.g. linking the forecast framework with multiple providers of oceanic and atmospheric predictions to handle uncertainty and provide land and coastal inputs to these compartments)
- To the continuum of river to ocean modeling, an urban/land input component will be added, allowing also for bidirectional interactions, fundamental for including the coastal impact in the city (for instance for sea level rise or salinization of the drainage networks with potential severe impacts on WWTPs). Both hybrid and process based models will be considered and integrated.
- to explore the several tools developed in the previous projects to build a reliable, high-accuracy framework that can serve a number of purposes and have the capacity to be customized / tailored to specific themes

The work proposed herein is framed in LNEC's new research strategy. Along with the previous projects and contributing to the next Research Study, it will constitute the forecast core of the new Coastal Digital Twins, proposed in the next study.

5.3.8.3 Methodology and task description

The project will build the cross-scale platform, following a typical methodology for information technology development, adapted for maritime engineering purposes. The building blocks from the previous project will be integrated in a complex framework that address a wide range of scales and bidirectional interactions. Uncertainty concerns as well as quality control through assimilation will be addressed. The system will be validated in data-rich use cases, with varying processes combinations. It also includes their integration in the OPENCoastS platform.

The following activities and tasks are organized sequentially to achieve the above goals:

Activity 1 - Requirements analysis for cross-scale prediction

Task 1: Definition of the interdependencies among modeling compartments and data exchange protocols and standards

Task 2: Selection of river and urban modeling tools including hybrid approaches

Task 3: Definition of the integrated forecast workflow, data sources for validation and assimilation and sources of uncertainty

Activity 2 - Development of new core forecast engines

Task 1: Development/adaptation of riverine forecast engine

Task 2: Development/adaptation of urban forecast engine

Task 3: Testing in data-rich sites

Activity 3 - Integration of quality-controlled data assets in cross-scale forecast framework

Task 1: Integration of new data sources for the new domains

Task 2: Integration of optimized uncertainty module with adaptation to new uncertainty sources

Task 3: Adaptation and integration of assimilation module for freshwater and urban domains

Task 4: Testing in data-rich sites

Activity 4 - Integration in OPENCoastS framework

Task 1: Requirements analysis for OPENCoastS integration

Task 2: Integration of basin and river forecasting engine components

Task 3: Integration of urban discharge engine component for both inundation and water quality

Task 4: Development of user-selection engine for specifying water pathway components choice

Task 5: Setting up of virtual research environment for cross-scale forecasting in shared e-infrastructures

Activity 5 - Dissemination and education

Task 1: Presentation at national and international congresses and conferences

Task 2: Publication in conference proceedings and WoS journals

Task 3: Elaboration of M.Sc. thesis and Ph.D. dissertation

These atividades are expected to be executed over the timeline presented in Table 5.9.

Table 5.9 - Project 1.7: Timeline

Year	Year 1				Year 2				Year 3				Year 4			
Activity	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**
1 - Requirements analysis for cross-scale prediction																

Year	Year 1				Year 2				Year 3				Year 4			
Activity	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**
2 - Development of new core forecast engines																
3 - Integration of quality-controlled assets in cross-scale forecast framework																
4 - Integration in OPENCoastS framework																
5 - Dissemination and education																

5.3.8.4 Expected outcomes

The outcomes of the project will be divided in two categories: contributions to the advance of the state of the art and benefits for LNEC.

In the first category, this research project will:

- a novel framework for integrated forecasting across river-to-ocean scales to predict jointly the dynamics of the adjoint water bodies using a combination of models at the correct time and space scales, including the urban dimension and with bidirectional interaction between neighboring water compartments;
- an implementation of this framework based on high accuracy multiple models, accounting for uncertainty.

LNEC will benefit of the project by:

- the availability of a high-accuracy single forecast framework and platform that address all areas of actuation of the Hydraulics and Environment Department at the adequate space and time scales, that would promote an integrated vision to prediction in the water domain;
- the extension of the OPENCoastS platform usability to the full water cycle (except the atmosphere component).

The outcome of this project will be integrated as the core forecast engine for the coastal digital twins' framework and platform, to be proposed in the Research Study 2 of this Research Program.

As a standalone product it will also be used in the following two projects under this Research Study 1, to illustrate its usefulness for two relevant forecast problems in support of the management of the Portuguese coast.

5.3.8.5 Resources

The list below summarizes the expected costs associated with this work.

a) Personnel costs
   - 48 months of Ph.D. grant
   - 12 months of a senior researcher
   - 3 months of computer expert
   - Total: 133 842,72 euros
b) Travel: 4 participation in conferences (2 national, 2 Europe) - 8500 euros
c) Equipment:
   - laptop - 1500 euros
   - e-infrastructure resources - no costs (assuming that an application to free resources is done in the Rede Nacional de Computação Avançada calls)
d) **Grand total (with overheads): 226 542,72 euros**

5.3.8.6 Funding and partnerships

For the funding of the project, applications should be submitted to FCT projects and Ph.D. calls to fund the Ph.D. grant and the travel and equipment values. As FCT does not fund personnel with a permanent contract, a research proposal should be submitted to the Horizon Europe, INTERREG, ANI or similar calls to fund or co-fund the senior researchers. Given the importance of integrated forecasting to support water management and the possibility of having much more accurate predictions of urban impacts on estuarine and coastal systems, a co-funding should also be requested at the coastal authorities (APA), municipalities and wastewaters companies. The concept of collaborative funding, successfully applied at the urban water division of LNEC, can be an avenue to be explored for this project.

Partnerships can be established for this project with all teams working in estuarine and coastal modeling and its interactions with land inputs in Portugal, with emphasis on the University of Lisbon, Aveiro and Algarve. Given that SCHISM will be one of the primary models, VIMS and NOAA will be definitely important partners, framed in the context of the Ocean Modeling Collaboration Forum.

5.3.9 Project 1.8 - Integrated management of urban beaches quality in dense-populated city environments affected by hazardous flood events - the coastal city of Albufeira case (DEMO-PT)

5.3.9.1 Rationale

Drainage infrastructures have a fundamental role in the quality of life of cities, avoiding inundations, by regulating precipitation effects, and avoiding water and atmospheric contamination through adequate water treatment. In a context of climate changes, their impact is even greater, due to the expected

increasing of precipitation negative effects through short term inundation events (flash floods) and greater contamination discharges to nearby receiving water bodies.

However, these infrastructures require a continuous rigorous maintenance and operation, which can lead to very large energy consumptions, which can be further aggravated by climate change. The continuous increase in surface impermeabilization, the climate change related aggravated conditions in precipitation and sea level rise and the joint drainage for pluvial and wastewater flows in many cities leads to very complex conditions that need to be tackled jointly at all water compartments and address quantity and quality. One of the greatest challenges of drainage network infrastructures is thus to find an optimized way for their operation that minimizes energy consumption while protecting receiving water bodies and minimizing urban flooding. A management support system to tackle this challenge is not available yet.

The cross-scale forecast framework developed in the previous project is herein proposed to be applied to this challenge using the Albufeira coastal city as a demonstrator, taking at the same time into account the several tools developed for uncertainty, scenario building, hybrid modeling and data assimilation, and simulation in real time. A scenario database accounting for all relevant processes at stake in the several water compartments will also be produced. An optimization engine is added to the tool suite, in order to address the multi-objective goals, supported by a suite of indicators for both inundation, beach water quality and energy.

The Albufeira coastal city is frequently subject to large inundation events downtown that have considerable economic and social impacts and also often lead to the contamination and occasional closing of the urban beaches. Several options exist for the management of pluvial waters, through careful management of the time and space distribution of precipitation flow and its retention to allow for adequate water treatment before discharge in the nearby coastal zone.

5.3.9.2 Goals

This project aims at customizing the cross-scale framework of project 1.7 to the optimization of drainage infrastructure operation in coastal cities targeting the minimization of city-originated contamination and city inundation. The project extends the analysis and modeling efforts developed in the ANI-funded SINERGEA project, scientifically led by the applicant and currently under development at LNEC in partnership with SIEMENS, IST and Algarve university.

The specific goals of this project are:

- To establish a hybrid model forecasting application to the urban and coastal water domains, based on the numerical model applications of sinergea and several sources of in-situ and remote data
- To establish a suite of scenarios for environmental conditions and anthropogenic operation of the infrastructures and its expected changes due to climate change, accounting for uncertainty, and build a fast-response model for daily management support
- To establish an early warning system based on the two previous tools

- To build an optimization engine that uses both the database from scenarios and the daily predictions to propose the optimal infrastructure operation under everyday environmental and infrastructure regular and failure conditions

To avoid handling a huge volume of simulations and data, the concept of hotspots defined in collaboration with the local authorities and infrastructure managers will be used to provide fast answers in time for real time optimization.

5.3.9.3 Methodology and task description

The methodology is based on the development and application of the several components for the early-warning and operation optimizer in the Albufeira case study. It will build on top of the outcomes of the previous projects and the preliminary analysis developed in the SINERGEA project.

Activity 1 - Requirements analysis for the multi-objective coastal management system

Task 1: Review of the numerical models implemented and definition of the data sources to be used for the hybrid modeling

Task 2: Review of the environmental and CC scenarios to be simulated updated to current knowledge of the system and status of implementation of the flood defense project, and evaluation of the computational resources required

Task 3: Establishment of the suite of indicators to be used in the system performance assessment and optimizing engine

Task 4: Definition of the optimization algorithm

Task 5: Definition of the hotspot locations for inundation and coastal water quality to serve as decision guides for optimal performance

Task 6: Definition of the early-warning conditions

Activity 2 - Development of the updated scenario database and hotspot analysis

Task 1: Simulation of scenarios and creation of hotspots timeseries

Task 2: Development of AI models based on hotspots results

Task 3: Development of fast search algorithm to support fast predictions using scenario database

Task 4: Application and validation of both approaches with past events

Activity 3 - Development and testing of the optimization operation engine

Task 1: Development of optimization engine

Task 2: Testing under several environmental conditions

Activity 4 - Implementation of the hybrid forecast system with uncertainty and data assimilation

Task 1: Implementation of the hybrid models

Task 2: Implementation of the cross-scale forecast systems using those hybrid models

Task 3: Integration of the uncertainty engine fully adapted for this cross-scale modeling

Task 4: Evaluation and possible implementation of data assimilation using remote sensing at the cost and real time data at the city infrastructure

Task 5: Testing under past events conditions

Activity 5 - Implementation of the early-warning system and optimization conditions proposal

Task 1: Implementation of the early warning system for inundation and beach contamination based on the tools from activities 2 and 4

Task 2: Implementation of the optimized proposal of operation for real time operation

Task 3: Development of dedicated dashboard in OPENCoastS for the early warning and proposal of operation results

Task 4: Test and validation of the whole system under past events conditions, with the support of the management entities

Activity 6 - Dissemination and education

Task 1: Training and technology transfer to the management entities to support a harmonization of management decisions

Task 2: Presentation at national and international congresses and conferences

Task 3: Publication in conference proceedings and WoS journals

Task 4: Elaboration of M.Sc. thesis and Ph.D. dissertations

These atividades are expected to be executed over the timeline presented in Table 5.10.

Table 5.10 -- Project 1.8: Timeline

Year	Year 1				Year 2				Year 3				Year 4			
Activity	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**
1 - Requirements analysis for multi-objective coastal management system																
2 - Development of the updated scenario database and hotspot analysis																

Year	Year 1				Year 2				Year 3				Year 4			
Activity	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**
3 - Development and testing of the optimization operation engine																
4 - Implementation of the hybrid forecast system with uncertainty and data assimilation																
5 - Implementation of the early-warning system and optimization conditions proposal																
6 - Dissemination and education																

5.3.9.4 Expected outcomes

The outcomes of the project will be divided in two categories: contributions to the advance of the state of the art and benefits for LNEC.

In the first category, this research project will provide:

- a methodology for optimization of drainage infrastructure operation in coastal cities, combining multiple models and data sources
- an optimization engine that uses both the database from scenarios and the daily predictions to propose the optimal infrastructure operation under everyday environmental and infrastructure regular and failure conditions
- a demonstration in a very complex situation where bathing water quality and inundation need to be handled simultaneously, along with concerns of energy savings.

LNEC will benefit of the project by:

- the availability of a optimization procedure applicable to coastal cities management, combining scenarios and real time predictions with the potential to be explored in other optimization areas at the coast or in other water compartments
- a unique application to the multi-objective case of the city of Albufeira, which can be used to provide consultancy services to the several end-users: APA, Águas do Algarve and Albufeira municipality.

5.3.9.5 Resources

The list below summarizes the expected costs associated with this work.

a) Personnel costs
    - 48 months of Ph.D. grant
    - 12 months of a senior researcher
    - 1 month of computer expert
    - Total: 126 842,72 euros
b) Travel: 4 participation in conferences (2 national, 2 Europe) - 8500 euros
c) Equipment:
    - laptop - 1500 euros
    - e-infrastructure resources - no costs (assuming that an application to free resources is done in the Rede Nacional de Computação Avançada calls)
d) Grand total (with overheads): 212 542,72 euros

5.3.9.6 Funding and partnerships

For the funding of the project, applications should be submitted to FCT projects and Ph.D. calls to fund the Ph.D. grant and the travel and equipment values. As FCT does not fund personnel with a permanent contract, a research proposal should be submitted to the INTERREG, ANI or similar calls to fund or co-fund the senior researchers. Given the importance of integrated forecasting to support water management and the possibility of having much more accurate predictions of urban impacts on estuarine and coastal systems, a co-funding should also be requested at the coastal authorities (APA), municipalities and wastewaters companies. The concept of collaborative funding, successfully applied at the urban water division of LNEC, can be an avenue to be explored for this project.

5.3.10 Project 1.9 - Multi-purpose forecast platform for fast prediction of pollution plumes and its application to portuguese estuaries and coasts (DEMO-PT)

5.3.10.1 Rationale

The impact of pollution accidents in the Portuguese estuaries and coasts can have severe consequences both in economic and social contexts. The availability of an open access, easy-to-use tool can contribute towards human and environmental assets protection, participative management and harmonization of the sometimes conflicting uses of coastal regions. It can be used for instance to forecast particular events by coastal managers (e.g. spill from a ship) or to estimate the exposure of a particular recreational spot by a company operating in that coastal region, without requiring any technical background. Likewise, this tool can be used by researchers without expert modeling skills to assess the impact of physical processes on transport, water quality or ecology and thus contribute towards knowledge creation in biogeochemical and ecological processes.

The platform will be based on the outcomes of the previous development projects and will be built as an independent web tool in Portuguese, with usability conceptualization for multiple types of users, based on GTI's experience in the creation of multiple portals such as the UBEST, MOSAIC or the SPRES web platforms.

5.3.10.2 Goals

This project aims at offering a fast, simple and user-friendly pollution prediction tool applicable to any generic tracer to be customized by the user or to pre-identified relevant contamination types for the Portuguese coast. The tool is expected to be tailored for usage by multiple types of users with distinct backgrounds without requiring numerical modeling skills but retaining high accuracy to also support research studies.

5.3.10.3 Methodology and task description

Unlike most applications of models at the coasts, this project starts by an in-depth collaborative work with end-users to guarantee that the prediction tool is comprehensive but simple enough to address the main contamination problems at the Portuguese coast. Then it sets up detailed, uncertainty-aware circulation forecasts for the areas identified as critical, to serve as the hydrodynamic engine for any pollution predictions. OPENCoastS is then customized to this problem by creating an early warning and an information sharing dashboards to promote knowledge and harmonized management of conflicting areas.

Activity 1 - Requirements analysis

Task 1: Participatory identification of the most relevant water quality problems in the Portuguese coast, with coastal managers and representatives of associations

Task 2: Identification of the most relevant hotspots for the identified pollutants (including oil spills, microplastic and wastewater discharges) in the Portuguese coast

Task 3: Selection of the relevant physical processes for each hotspot location

Task 3: Selection of the relevant decay processes for each type of pollutants

Activity 2 - Implementation of the hotspots' circulation forecasts

Task 1: Grid generation for the selected areas using OPENCoastS grid generator

Task 2: Setting up of uncertainty-oriented, multiple choice circulation forecasts in OPENCoastS and operation for a defined period

Task 3: Calibration and validation through OPENCoastS' data lakes

Activity 3 - Application of OPENCoastS for on-demand pollution events and pre-defined continuous sources

Task 1: Requirements for integration in OPENCoastS

Task 2: Implementation of early-warning dashboard for continuously operating pollution sources dashboard comprising pre-defined and user-tailored tracers

Task 3: Online information sharing dashboard creation for multiple end-users

Activity 4 - Dissemination and education

Task 1: Presentation at national congresses and conferences

Task 2: Publication in conference proceedings and WoS journals

Task 3: Elaboration of M.Sc. thesis

Task 4: Onsite training for distinct end-users on pre-defined products and on-demand events simulation

These atividades are expected to be executed over the timeline presented in Table 5.11.

Table 5.11 - Project 1.9: Timeline

Year	Year 1				Year 2				Year 3	
Activity	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**	**1**	**2**
1 - Requirement analysis										
2 - Implementation of the hotspots' circulation forecasts										
3 - Application of OPENCoastS for on-demand pollution events										
4 - Dissemination and education										

5.3.10.4 Expected outcomes

The outcomes of the project will be divided in two categories: contributions to the advance of the state of the art and benefits for LNEC.

In the first category, this research project will:

- participative, fast, simple and user-friendly pollution prediction tool applicable to any generic tracer to be customized by the user or to pre-identified relevant contamination types for the Portuguese coast

LNEC will benefit of the project by:

- the availability of fast and simple tool for exploratory works
- the provisioning of a collaborative tool for end-users aiming at supporting coastal management

5.3.10.5 Resources

The list below summarizes the expected costs associated with this work.

a) Personnel costs
    - 12 months of M.Sc. grant
    - 6 months of a senior researcher
    - 2 months of computer expert
    - Total: 51 711,76 euros
b) Travel: 1 participation in european conference - 3000 euros
c) Equipment:
    - laptop - 1500 euros
    - e-infrastructure resources - no costs (assuming that an application to free resources is done in the Rede Nacional de Computação Avançada calls)
d) Grand total: 100 111,76 euros

5.3.10.6 Funding and partnerships

For the funding of the project, applications should be submitted to FCT projects calls to fund the M.Sc. grant and the travel and equipment values. As FCT does not fund personnel with a permanent contract, a research proposal should be submitted to the INTERREG, ANI or similar calls to fund or co-fund the senior researchers and computer experts. Given the flexibility of the proposed tools, it can be of utmost value for the management of contamination events. Therefore, a co-funding should also be requested at the coastal authorities (APA), municipalities and wastewaters companies. The concept of collaborative funding, successfully applied at the urban water division of LNEC, can be an avenue to be explored for this project.

5.4 Research Study 2: CODIT - Intelligent, high-resolution, user-centered and inclusive Coastal Digital Twins

5.4.1 Introduction

This research study aims to address the open questions identified in Chapter 4. It proposes the roadmap for the development of a new paradigm in coastal research, Coastal Digital Twins (CDT), created and developed in close articulation with Destination Earth goals and the current initiatives to develop the Digital twins of the Ocean. The concept presented herein is strongly inspired by the water industry vision of Digital Twins.

CDT provides a virtual representation of a physical asset enabled through data and models, which can be used for multiple applications such as real-time forecast, system optimization, monitoring and controlling, and support enhanced decision making. They can integrate continuously, in an interactive, two-way data connection, the physical and the virtual assets. They can take advantage of the huge online volume of data streams provided by satellites, IoT sensing and many real time surveillance platforms, the availability of powerful computational resources that made process-solving high resolution models or AI-based models possible, to build high accuracy replicas of the real world.

CDT are still in the conceptual phase and have not been detailed or implemented yet in operational mode. CDT and coastal forecast systems are distinct, the latter being an integral part of the former.

Digital Twins are user-centered, having user interaction embedded in their design, including visualization, user-driven data transformation and processing, and user selection of tools to produce a specific output. On the contrary, forecast systems are typically product centered, aiming at providing pre-defined services to users. The research projects proposed in the previous study aim at advancing the current definition of coastal forecast systems to create new enhanced products but also to develop the on-demand forecast core to be used in the CDT proposed herein.

The research projects within this Study are organized along three main tracks: Research dimension (RD), Data dimension (DD) and Service dimension (SD). A fourth dimension - Applications to Portuguese coast - is intertwined with all of the others and will not be presented separately but rather integrated as the demonstration layer in most of the projects. The classification is indicated at the end of each project's title. This suite of projects aims at establishing the concept and building the core capabilities of Coastal Digital Twins, anchored on the most advanced knowledge and tools available in coastal science and engineering and applied Information technology, taking advantage also of the advances and outcomes proposed in the Research Study 1.

Figure 5.3 illustrates the several projects, their relationships and context and in some cases their dependencies.

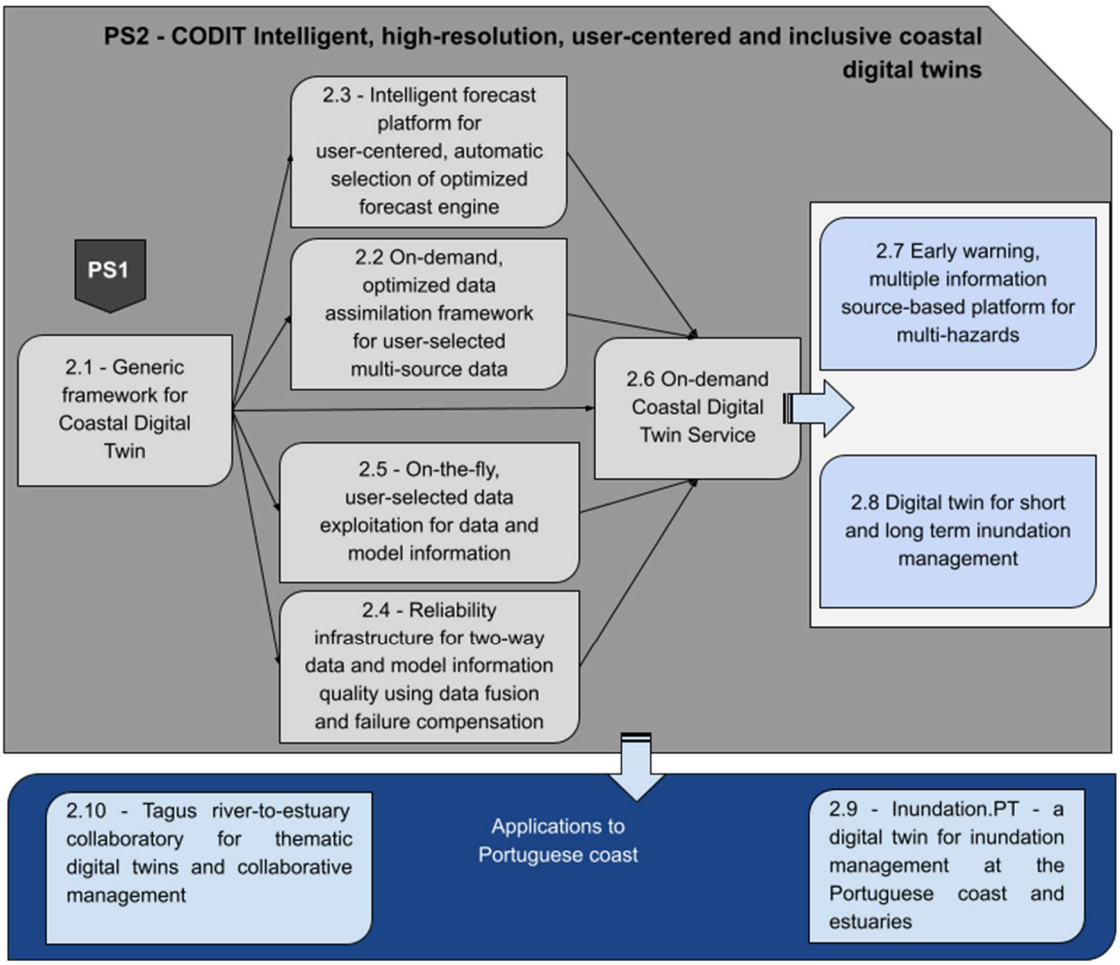


Figure 5.3 - Research study 2 and its projects: characterization, workflow and dependencies

5.4.2 Project 2.1 - Generic framework for Coastal Digital Twin (RD)

5.4.2.1 Rationale

The current expectations of CDT's usefulness, completeness and flexibility are very high as the conceptual vision targets a world "controlled at the end of our fingertips", with information harnessed to produce knowledge and optimized decisions automatically. These expectations are framed by the ever-growing need from researchers to create knowledge from the huge volume of data freely available, that can no longer be analyzed by humans without automatic computational procedures, and the fast pace required to management actions that need to respond on frequently unreasonable time frames to the evolution of a constantly changing coast due to climate change or anthropogenic actions, dealing with growing legislative restrictions and very high needs to preserve ecosystems and food chains for the services that provide to mankind.

In order to move towards this highly ambitious vision, the construction of a solid backbone is necessary, in order to accommodate any new developments and requirements in the future. Herein, the applicant proposes the development of a generic framework that will integrate the current vision for the requirements and is prepared to integrate a high level of interaction between tools, data and models and provide scientifically-based support to collaborative management of coastal zones.

Building from the experience of LNEC's on-demand forecast framework OPENCoastS, CODIT's generic framework target robustness, accuracy, flexibility and wide, multi-purpose applicability in any coastal system in the world. This framework will also provide LNEC with the capacity to build an advanced, tailored, on-demand tool that can be customized to multiple engineering applications and be handed out at the end of the project to the clients for their current and future use.

5.4.2.2 Goals

This project aims at conceptualizing and building a robust, flexible framework to serve as a computational basis for the development of complex, flexible and interactive CDT. The OPENCoastS and WIFF frameworks will be used as basis, given their on-demand capacity and simple user-interaction properties.

The specific goals of the project are:

- to identify the core requirements for a CDT, accounting for multiple service tools, real time operations and usability concerns for the user interaction;
- to structure a comprehensive data model to support the CDT data/information lake and account for information dependencies, multiple sources of data/models and compliance with FAIR data principles;
- to decouple the CDT structure with generic building blocks that can then be used to mount the tailored workflows for multiple applications through a CDT builder tool;
- to select and integrate several predefined tools (models, data sources, ...) for immediate setup of simple CDT using the outcomes of Research Study 1.

The proposed work will be framed in the European Digital Hub ATTRACT-DIH and the several other European and national projects at LNEC in the area of estuarine and coastal forecasting.

5.4.2.3 Methodology and task description

The implementation will start with the requirement analysis both from an IT and coastal engineering perspectives. The necessary basic components for a generic application will be identified and implemented in a flexible way to grow to further options in the future (e.g. the simulation block will be built for a set of simulation types with some options for the modeling tools based on the enhanced OPENCoastS forecast engine of Research Study 1). The data model to support the implementation will be conceptualized in a generic way accounting for data standards from the coastal scientific community (e.g. netcdf files for model outputs).

Activity 1 - CDT requirements analysis

Task 1: Identification of relevant functions for a CDT and identification of their properties

Task 2: Identification of the relevant types of tools for implementation and identification of its characteristics and dependencies

Task 3: Identification of the relevant data formats and data dependencies for creation of a data /information model and its implementation as a data lake

Task 4: Identification of the user interaction modes and development of a usability strategy to accommodate them in a platform dashboard

Activity 2 - Data/information model and lake construction

Task 1: Implementation of a data/information repository following the requirements analysis

Task 2: Evaluation of standard data and FAIR compliance through FAIR compliance software

Task 3: Validation in multiple types of data structures for several selected use cases

Activity 3 - CDT workflow builder

Task 1: Definition of the generic building blocks in the CDT and its tailoring for predefined typologies of problems

Task 2: Adaptation/Extension of WIFF´s and OPENCoastS´ building blocks for CDT

Task 3: Development of a workflow builder according to specifications for specific applications and for a generic, customizable problem

Task 4: Testing of the procedure in selected use cases

Activity 4 - Development of simple CDTs for predefined cases using Research Study 1 outcomes

Task 1: Identification of the thematic focus of the areas to be addressed by these simple CDTs

Task 2: Identification of the services and capacities to be available in each simple CDT

Task 3: Creation of the thematic CDTs using the builder

Task 4: Testing of the coherence and services of the CDT for selected demo applications

Activity 5 - Dissemination and education

Task 1: Presentation at selected IT-oriented water congresses and conferences

Task 2: Publication in conference proceedings and 1 WoS journal

Task 3: Elaboration of 1-2 M.Sc. thesis and 1 Ph.D. dissertation.

These atividades are expected to be executed over the timeline presented in Table 5.12.

Table 5.12 - Project 2.1: Timeline

Year	Year 1				Year 2				Year 3			
Activity	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**
1 - CDT requirement analysis												
2 - Data/information model and lake construction												
3 - CDT workflow builder												
4 - Development of simple CDTs for predefined cases using Research Study 1 outcomes												
5 - Dissemination and education												

5.4.2.4 Expected outcomes

The outcomes of the project will be divided in two categories: contributions to the advance of the state of the art and benefits for LNEC.

In the first category, this research project will:

- develop the first coast digital twin infrastructure through the development of a generic framework that is prepared to integrate a high level of interaction between tools, data and models and provide scientifically-based support to collaborative management of coastal zones

LNEC will benefit of the project by:

- being the first world research institute that promote CDT as effective and operational collaborative management tools
- developing the backbone for the future CDT and being able to apply it to address a set of coastal management problems in a collaborative, multi-user way.

5.4.2.5 Resources

The list below summarizes the expected costs associated with this work.

a) Personnel costs
    - 24 months of Ph.D. grant
    - 12 months of senior researchers
    - 4 months of computer expert
    - Total: 109 871,36 euros
b) Travel: 4 participation in conferences (2 national, 2 Europe) - 8500 euros
c) Equipment:
    - laptop - 1500 euros
    - e-infrastructure resources - no costs (assuming that an application to free resources is done in the Rede Nacional de Computação Avançada calls)
d) Open access publication fees: 3000 euros (fee waivers will be promoted, one publication is expected to be in a journal without waivers)
e) Grand total (with overheads): 206 071,36 euros

5.4.2.6 Funding and partnerships

For the funding of the project, applications should be submitted to FCT project and grant calls, to fund the PhD grant and the travel and equipment values. As FCT does not fund personnel with a permanent contract, a research proposal should be submitted to the Horizon Europe, INTERREG, ANI or similar calls to fund or co-fund the senior researchers and computer experts. The EU Missions calls are a good example of adequate calls as well as the calls dedicated to Digital Twins.

Regarding partnerships, many can be established for this project with the partners/funders of OPENCOastS development and follow-up initiatives (CNRS, VIMS, NOAA, UN Decade CoastPredict team) and usage for coastal applications (University of Lisbon and Algarve). New partnerships should also be built with end-users and consultancy companies, framed in the scope of the Digital Innovation Hub ATTRACT-DIH (led by the universities of Minho and Oporto). The conceptualization of the CDT should be done in close cooperation with those that should be using it in the future. The partnerships with the two UN-DEcade Programs linked with this area (DITTO and CoastPredict) are already being built and will be an important part of the success of this project.

Finally, building and maintaining CDT in Portugal should be a joint effort of all state labs operating in this area. Therefore, a long-term collaboration protocol with IPMA and IH will be promoted to frame this effort.

5.4.3 Project 2.2 - On-demand, optimized data assimilation framework for user-selected multi-source data (RD)

5.4.3.1 Rationale

The recent availability of multiple data sources at adequate spatial and temporal scales has promoted the development of several strategies for assimilation of data in numerical modeling simulations. In particular, remote sensing solutions have become major data sources for many coastal and estuarine data acquisition systems, given the technological advances of remote sensing devices, such as HFR radar, cameras or sensors mounted in drones and satellite remote sensing. These different solutions provide data for different variables and devices can be applied independently or in an integrated fashion, to promote knowledge, validate forecast simulations or to be part of assimilation procedures. At the same time the expansion of IoT networks and the development of several types of low cost sensors has also provided additional data sources in estuarine and coastal regions. Finally, many of the data providers are now following a policy of open access to most of the data sources (e.g. open availability of medium resolution (~10 m) satellite data from the Sentinel constellation), thus facilitating the creation of added value service to improve the quality of digital representations of real systems such as forecast systems or CDT. Combining data assimilation and models do not only promote model improvement, but they can also be used to design and assess ocean observing systems through the creation and establishment of Observing System Simulation Experiments (OSSEs).

The complexity of the non-linear processes in coastal regions is in general a limiting factor to the application of simple assimilation techniques, in particular for those requiring an estimator of the errors. Using a combination of many data sources, to build dense, reliable monitoring networks covering the small time and spatial scales, may be an avenue to overcome this problem and promote useful developments. The possibility of using multiple sources of data in the assimilation procedure requires the capacity to ensemble data from multiple sources, with potentially distinct time and space scales, heterogeneous formats and semantics. Furthermore, the choice of the assimilation technique may not be a single one for the same problem as environmental conditions may shift and the optimal procedure may be space and time varying. Exploring this variability into an optimized tool has not been explored yet.

5.4.3.2 Goals

In this project the applicant aims at developing an optimized methodology to assimilate multi-source data in a seamless and quality controlled way, keeping computational cost under acceptable conditions. These methodologies should have the capacity to setup an assimilation strategy that can be applied for a multi-scale (from the basin to the sea), multi-process digital twin (including hydrodynamics, coastal erosion and water biogeochemistry) and to data-based, numerical or hybrid models, in a user-driven, on-demand way, and switching from one methodology to the next depending on their performance over specific environmental conditions.

The specific goals are:

- to define data sources, error measures and the goals for each assimilation procedure targeting optimization
- select the best assimilation algorithms for estuarine and coastal dynamics and implement them for generic application to multiple model types
- integrate the methodologies in OPENCoastS to serve in automatic forecasting, user configuration and construction of new tools.

5.4.3.3 Methodology and task description

The work will start with the data sources identification, processing and integration, followed by the algorithm comparative choice. Based on these results, an automatic algorithm selector will be built and integrated in the on-demand tools strategy to allow users to configure their choice of assimilation or use the optimized procedure. These tasks will be available by integration of the tools in OPENCoastS.

Activity 1 - Requirements analysis and data characterization for assimilation purposes

Task 1: Definition of the data sources and their characterization

Task 2: Definition of the errors measures to be used in data assimilation assessment

Task 3: Definition of the goals for data assimilation

Activity 2 - Selection and implementation of data assimilation methodology

Task 1: Update of data assimilation state-of-the-art

Task 2: Selection of most relevant algorithms and their implementation

Task 3: Comparison of algorithms in simple use cases and classification of their adequacy depending on dominant processes

Task 4: Development of automatic selector for proposal to users

Activity 3 - Development of assimilation on-demand capacity and integration in OPENCoastS

Task 1: Identification of the user-selected parameters

Task 2: Adaptation of OPENCoastS frontend to accommodate data assimilation

Task 3: Linkage between data-sources interface and data assimilation through user selection of sources and algorithm choice

Task 4: Integration of data assimilation in the hybrid forecast workflow

Task 4: Setting up of virtual research environment for data assimilated simulations in shared e-infrastructures

Activity 4 - Dissemination and education

Task 1: Presentation at national and international congresses and conferences

Task 2: Publication in conference proceedings and WoS journals

Task 3: Elaboration of M.Sc. thesis and Ph.D. dissertation

These atividades are expected to be executed over the timeline presented in Table 5.13.

Table 5.13 - Project 2.2: Timeline

Year	Year 1				Year 2				Year 3				Year 4			
Activity	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**
1 - State-of-the-art review and data characterization																
2 - Selection and implementation of data assimilation methodology																
3 - Development of assimilation on-demand capacity and integration in OPENCoastS																
4 - Dissemination and education																

5.4.3.4 Expected outcomes

The outcomes of the project will be divided in two categories: contributions to the advance of the state of the art and benefits for LNEC.

In the first category, this research project will:

- provide a new methodology for adaptive data assimilation in estuaries and coasts that selects the best assimilation algorithms for current estuarine and coastal dynamics and implement them for generic application to multiple model type
- integrate the data assimilation strategy in OPENCoastS to serve in automatic forecasting, user configuration choice and support the construction of new tools in the CDT

LNEC will benefit of the project by:

- the availability of an innovative strategy for multiple model type, adaptive data assimilation that can be applied in a CDT forecast context or as a stand-alone tool, taking the best advantage of the multiple sources of open data

5.4.3.5 Resources

The list below summarizes the expected costs associated with this work.

e) Personnel costs
    - 48 months of Ph.D. grant
    - 12 months of a senior researcher
    - 3 months of computer expert
    - Total: 133 842,72 euros
f) Travel: 4 participation in conferences (2 national, 2 Europe) - 8500 euros
g) Equipment:
    - laptop - 1500 euros
    - e-infrastructure resources - no costs (assuming that an application to free resources is done in the Rede Nacional de Computação Avançada calls)
h) Grand total (with overheads): 226 542,72 euros

5.4.3.6 Funding and partnerships

For the funding of the project, applications should be submitted to FCT project and grant calls, to fund the PhD grant and the travel and equipment values. As FCT does not fund personnel with a permanent contract, a research proposal should be submitted to the Horizon Europe, INTERREG, ANI or similar calls to fund or co-fund the senior researchers. The Digital Twins calls as well as some ICT ones are adequate funding calls.

Regarding partnerships, the teams at Portugal working on data lake construction and reliability are adequate partners for funding search. Previous projects with the University of Lisbon and the partnerships established in the scope of the Digital Innovation Hub ATTRACT-DIH (led by the universities of Minho and Oporto) are among the preferred ones as work developed previously can be used to enrich the proposed research (e.g. the reliability framework developed in the scope of the Ph.D. of Dr. Gonçalo de Jesus).

5.4.4 Project 2.3 - Intelligent forecast platform for user-centered, automatic selection of optimized forecast engine (RD)

5.4.4.1 Rationale

Hybrid forecast models, combining numerical and AI-based models, are becoming a popular tool in ocean operational services, given the fastest response of these models and their smaller need for computational resources in ensemble or scenario settings. The motivation behind the development of ensemble models is to take into account for instance the uncertainty in oceanographic and freshwater inputs, but computational costs can become prohibitive for many traditional model applications, making hybrid models an attractive alternative. While the traditional data-model applications are still restricted to the studied coastal site, their portability to other systems through the combined methodology is high. Training is done using numerical models duly calibrated and applied to the coastal site of interest and the site data. After training is completed, the emulators are able to provide accurate predictions over new forcing conditions. Physics-informed neural networks, for instance, are an example of a hybrid

approach that combines a data-based neural network model and a physics-informed mechanistic model, by minimizing the error with the data as well as the error with the physical equations.

The application of hybrid models to coastal domains is still very limited due to the highly non-linear nature of coastal dynamics. To overcome that limitation, the relative weight of data-based or process-based models in a hybrid forecast simulation could be optimized, taking advantage of assimilation and indicator calculations, developed in the previous project. Herein we explore the hybrid forecast framework and the data assimilation module developed in the previous research projects by including methodologies for optimal weighting of the error functions. These optimized settings will be made available to the users in the on-demand dashboards, allowing the user to configure, constrain and select the procedure to his/her goals. Integration in OPENCoastS allows for a centralized multi-service strategy behind CDT and the exploitation of the forecast system advances of Research Study 1 will provide the backbone for this work.

Application of the optimized forecasting will cover from hydrodynamics to biogeochemistry applications, including when needed morphodynamic updates.

5.4.4.2 Goals

This project aims at setting up an optimized forecast infrastructure from circulation to biogeochemistry, based on hybrid modeling and assimilation and uncertainty engines, and targeting the minimization of errors against simulations and data. The outcomes will be integrated in the generic framework for CDT, based on OPENCoastS and developed in the previous projects.

The specific goals are:

- defining the error measures and control parameters for the core optimization procedure;
- Selection of the optimization algorithm and definition of the smooth transition procedure between distinct weighted runs;
- development of optimized selector forecast infrastructure;
- integration in CDT OPENCoastS platform.

5.4.4.3 Methodology and task description

This work will first evaluate the necessary core elements for forecast evaluation, targeting automatic model selection and operation. Next, the selector engine will be build, based on the combination of outcomes from previous projects. Finally, all tools will be integrated in the CDTs OPENCoastS platform.

Activity 1 - Requirements analysis for optimized forecasts

Task 1: Definition of control parameters and error measures for hydrodynamic and biogeochemistry simulations

Task 2: Selection of the optimization algorithm and transition mechanisms between forecast engines

Task 3: Specification of user-centered dashboard

Activity 2 - Development of optimed forecast selector infrastructure

Task 1: Development of optimized selector

Task 2: Integration of uncertainty and assimilation modules

Task 2: Testing and validation under controlled conditions

Activity 3 - Implementation in CDT OPENCoastS framework

Task 1: Requirements analysis for OPENCoastS integration

Task 2: Integration of forecast selector and development of dedicated dashboard

Task 3: Development of user-selection engine for specifying optimization conditions and choice of proposed alternatives

Task 4: Setting up of virtual research environment for selector simulations in shared e-infrastructures

Activity 4 - Dissemination and education

Task 1: Presentation at national and international congresses and conferences

Task 2: Publication in conference proceedings and WoS journals

Task 3: Elaboration of M.Sc. thesis

These atividades are expected to be executed over the timeline presented in Table 5.14.

Table 5.14 - Project 2.3: Timeline

Year	Year 1				Year 2				Year 3			
Activity	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**
1 - Requirements analysis for optimized forecasts												
2 - Development of optimed forecast selector infrastructure												
3 - Implementation in CDT OPENCoastS framework												
4 - Dissemination and education												

5.4.4.4 Expected outcomes

The outcomes of the project will be divided in two categories: contributions to the advance of the state of the art and benefits for LNEC.

In the first category, this research project will provide:

- an optimized forecast engine applicable from circulation to biogeochemistry, based on hybrid modeling and assimilation and uncertainty engines, to provide the best predictions within the CDT

LNEC will benefit of the project by:

- having a highly efficient and accurate CDT that provides predictions for all model accounting for uncertainty and exploring data through assimilation. This tool can be used in research projects to advance knowledge or to build end user-tailored CDT

5.4.4.5 Resources

The list below summarizes the expected costs associated with this work.

a) Personnel costs
    - 24 months of Ph.D. grant
    - 8 months of senior researchers
    - Total: 95 871,36 euros
b) Travel: 4 participation in conferences (2 national, 2 Europe) - 8500 euros
c) Equipment:
    - laptop - 1500 euros
    - e-infrastructure resources - no costs (assuming that an application to free resources is done in the Rede Nacional de Computação Avançada calls)
d) Open access publication fees: 3000 euros (fee waivers will be promoted, one publication is expected to be in a journal without waivers)
e) Grand total (with overheads): 132 471,36 euros

5.4.4.6 Funding and partnerships

For the funding of the project, applications should be submitted to FCT project and grant calls, to fund the PhD grant and the travel and equipment values. As FCT does not fund personnel with a permanent contract, a research proposal should be submitted to the Horizon Europe, INTERREG, ANI or similar calls to fund or co-fund the senior researchers. The EU Missions calls are a good example of adequate calls as well as the calls dedicated to Digital Twins. Given the capacity and the characteristics of the final tool to support country wide prediction, co-funding from APA should be sought in partnership with local end users (municipalities, businesses at the coast, ...) through collaborative funding schemes.

Regarding partnerships, many can be established for this project with the partners/funders of OPENCOastS development and follow-up initiatives (CNRS, VIMS, NOAA, UN Decade CoastPredict team) and usage for coastal applications (University of Lisbon and Algarve). New partnerships should also be built with end-users and consultancy companies, framed in the scope of the Digital Innovation Hub ATTRACT-DIH (led by the universities of Minho and Oporto). The enhancement of the CDT proposed in project 2.1 proposed herein should also be done in close cooperation with the two UN-Decade Programs linked with this area (DITTO and CoastPredict), reinforcing the current collaboration initiatives.

5.4.5 Project 2.4 - Reliability infrastructure for two-way data and model information quality using data fusion and failure compensation (DD)

5.4.5.1 Rationale

A data-rich environment, with multi-source data available at fine spatial and temporal scales, is a fundamental component of a robust and accurate CDT. Nowadays, multi-source estuarine and coastal monitoring networks that combine in-situ and remote sensing, fixed and mobile devices covering multiple space and time scales are becoming the strategy to address properly achieve data-informed coastal management. However, these vast amounts of data that no longer can be scanned and evaluated by humans, pose a new challenge to data reliability and quality which become fundamental to support in a robust way the usage of data-rich CDTs or even simple management decisions based on this data. The work of Jesus (2019), based on in-situ sensors and a single type of simulation, is proposed herein to be extended to the complex web of multi-source data and hybrid modeling of CDTs and to multiple process analysis, from circulation to water quality.

The implementation of this reliability infrastructure at the core of the CDT and the possibility to make it configurable by the users allows for simultaneous robustness and flexibility.

5.4.5.2 Goals

The goal of this project is to create the reliability layer in the CDT, using a composed analysis as proposed by Jesus (2019), applied to the multitude of data and models sources currently available. This layer constitutes a fundamental robustness and trustworthiness tool for formal application of CDTs for decision making.

The specific goals are:

- to create a reliability workflow for two-way data/model quality evaluation;
- to label information in CDT's data lake to support quality indexing of data and of the services built on top of it;
- to develop a dedicated dashboard in CDT's OPENCoastS platform to convey optimized results and allow users to create their own reliability workflow.

5.4.5.3 Methodology and task description

This project builds a reliability infrastructure from existing methods for automatic fault detection and correction, data quality assessment and classification, applied through the CDT's OPENCoastS platform. A user-centered dashboard allows users to make their own choices of data sources for the evaluation procedure or to apply the optimal procedure automatically proposed.

Activity 1 - Requirements analysis for reliable CDT data lakes

Task 1: Review of existing reliability procedures and identification of requirements for CDT

Task 2: Definition of the reliability and data quality assessment procedures using data and model in a composed way

Task 3: Requirements for integration in CDT's framework and OPENCoastS platform

Activity 2 - Reliability engine for optimal multi-source information ingestion in CDT data lakes

Task 1: Comparison of fault evaluation mechanisms for data and model information, for multiple types of simulations and data sources, using model/data composed analysis

Task 2: Selection of optimal procedure for fault detection and compensation

Task 3: Development of automatic engine for reliability evaluation of composed information from data sources and models

Activity 3 - Information classifier engine for optimal multi-source information ingestion in CDT data lakes

Task 1: Definition/selection of information classifier according to typology of conditions

Task 2: Fairness automatic evaluation engine integration

Task 3: Integration with reliability engine for optimal information ingestion in CDT data lakes

Task 4: Tests for multiple process conditions in data-rich use cases

Activity 4 - Implementation in CDT OPENCoastS framework

Task 1: Requirements analysis for OPENCoastS integration

Task 2: Integration of reliability and information classifier and development of dedicated dashboard

Task 3: Development of user-selection engine for specifying reliability formulation and choice of information sources

Task 4: Setting up of virtual research environment for reliability analysis in shared e-infrastructures

Activity 5 - Dissemination and education

Task 1: Presentation at national and international congresses and conferences

Task 2: Publication in conference proceedings and WoS journals

Task 3: Elaboration of M.Sc. thesis

These atividades are expected to be executed over the timeline presented in Table 5.15.

Table 5.15 - Project 2.4: Timeline

Year	Year 1				Year 2				Year 3			
Activity	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**
1 - Requirements analysis for reliable CDT data lakes												
2 - Reliability engine for optimal multi-source information ingestion in CDT data lakes												
3 - Information classifier engine for optimal multi-source information ingestion												
4 - Implementation in CDT OPENCoastS framework												
5 - Dissemination and education												

5.4.5.4 Expected outcomes

The outcomes of the project will be divided in two categories: contributions to the advance of the state of the art and benefits for LNEC.

In the first category, this research project will:

- automatic optimal information ingestion engine in CDT data lakes, including reliability and information classification engines and fairness evaluation

LNEC will benefit of the project by:

- automatic mechanism to evaluate and compensate for faults and guarantee the quality of information for application in coastal systems, which can increase the quality of the outcomes of LNEC's projects in the coastal area
- reliability layer as part of the CDT and the core OPENCoastS, further promoting the usefulness of this on-demand forecast tool

5.4.5.5 Resources

The list below summarizes the expected costs associated with this work.

a) Personnel costs
   - 12 months of M.Sc. grant
   - 6 months of a senior researcher
   - 2 months of computer expert

- Total: 51 711,76 euros

b) Travel: 1 participation in european conference - 3000 euros
c) Equipment:
    - laptop - 1500 euros
    - e-infrastructure resources - no costs (assuming that an application to free resources is done in the Rede Nacional de Computação Avançada calls)
d) Grand total: 100 111,76 euros

5.4.5.6 Funding and partnerships

For the funding of the project, applications should be submitted to FCT projects calls, to fund the M.Sc. grant and the travel and equipment values. As FCT does not fund personnel with a permanent contract, a research proposal should be submitted to the Horizon Europe, INTERREG, ANI or similar calls to fund or co-fund the senior researchers and IT personnel. Given the importance of reliable data lakes for decision making and the generic nature of the tool proposed for all water bodies, co-funding from APA should be sought in partnership with local end users (municipalities, businesses at the coast, ...) through collaborative funding schemes.

Regarding partnerships, two past collaborations will be continued: The University of Lisbon /FCUL collaboration in the area of reliability, partners in the AQUAMON project and Ph.D. dissertation of Dr. Gonçalo de Jesus, and the ATTRACT-DIH projects in the AI area. The creation of a reliable data lake inside the CDT proposed herein should also be done in close cooperation with the two UN-Decade Programs linked with this area (DITTO and CoastPredict), reinforcing the current collaboration initiatives.

5.4.6 Project 2.5 - On-the-fly, user-selected data exploitation for data and model information (SD)

5.4.6.1 Rationale

The volume of information to be created in the scope of CDT, along with its forecast systems and data sources outputs, is no longer possible to be processed in a non-automatic way. This project addresses this need at basic output processing and high level indicator production, to support both knowledge creation and management/legislation goals, using popular strategies such as Jupyter notebooks for maximum user uptake and customization. Sharing of algorithms is also a fundamental component of CDT, to maximize information exploitation and promote transparency. Therefore, the applicant proposes the integration of these algorithms in the CDT framework developed in the previous projects, as part of its available features. Besides flexibility and on-the-fly customization, the capacity to grow by integrating future notebooks is also a challenge to be addressed in this project, promoting a continuous improvement of CDT as knowledge advances and new tools are available.

The generation of time series of indicators and of other information at selected hotspots by the users through the application of the notebook algorithms creates new opportunities of semi-automatic

knowledge creation. The user can select outputs to be applied to AI-based algorithms for pattern detection and maximize the exploitation of the CDT information.

5.4.6.2 Goals

The aim of this project is to build a suite of data exploitation algorithms and make them available for automatic operation, sharing and customization to specific purposes, all within a CDT workflow. These algorithms will take advantage of popular technologies such as notebooks to facilitate user uptaking.

5.4.6.3 Methodology and task description

The work will start with a user requirements analysis, followed by identification of existing tools. These tools along with new ones for unresolved requirements will all be integrated in CDT framework through notebooks strategy.

Activity 1 - Requirements analysis and resources workflow

Task 1: Identification of user requirements

Task 2: Review of available information processing tools and selection of relevant ones for CDT

Task 3: Identification of resources for processing tools operation and establishment of workflow for operation in shared infrastructure

Activity 2 - Creation of dedicated notebooks for algorithm implementation and sharing

Task 1: Development/adaptation of dedicated notebooks for selected algorithms

Task 2: Development of notebooks for addressing the remaining requirements

Task 3: Integration in CDT framework for operation, sharing and on-the-fly customization

Task 4: Test in data- and result-rich CDT deployments

Activity 3 - Use of AI for automatic knowledge creation

Task 1: Implementation/adoption of AI engine in CDT

Task 2: Creation of dashboard for selection of information to be applied in the AI engine and selection of AI algorithm

Task 3: Test in data-rich CDT deployment

Activity 4 - Implementation in CDT OPENCoastS framework

Task 1: Requirements analysis for OPENCoastS integration

Task 2: Integration of data exploitation and AI engine and development of dedicated dashboard

Task 3: Setting up of virtual research environment for reliability analysis in shared e-infrastructures

Activity 5 - Dissemination and education

Task 1: Presentation at national and international congresses and conferences

Task 2: Publication in conference proceedings and WoS journals

Task 3: Elaboration of M.Sc. thesis

Task 4: Training for user uptake

These atividades are expected to be executed over the timeline presented in Table 5.16.

Table 5.16 - Project 2.5: Timeline

Year	Year 1				Year 2				Year 3			
Activity	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**
1 - Requirements analysis and resources workflow												
2 - Creation of dedicated notebooks for algorithm implementation and sharing												
3 - Use of AI for automatic knowledge creation												
4 - Implementation in CDT OPENCoastS framework												
5 - Dissemination and education												

5.4.6.4 Expected outcomes

The outcomes of the project will be divided in two categories: contributions to the advance of the state of the art and benefits for LNEC.

In the first category, this research project will:

- an infrastructure for information processing and sharing, to be used in automatic and user-customizable way, available as a service inside CDT

LNEC will benefit of the project by having:

- an automatic and customizable to specific purposes data exploitation algorithms engine to be applied in research and consultancy projects
- an enhanced OPENCoastS CDT, whose user-friendly information processing tools can be useful for many LNEC's end users.

5.4.6.5 Resources

The list below summarizes the expected costs associated with this work.

a) Personnel costs
    - 12 months of M.Sc. grant
    - 6 months of a senior researcher
    - 2 months of computer expert
    - Total: 51 711,76 euros
b) Travel: 1 participation in european conference - 3000 euros
c) Equipment:
    - laptop - 1500 euros
    - e-infrastructure resources - no costs (assuming that an application to free resources is done in the Rede Nacional de Computação Avançada calls)
d) Grand total: 100 111,76 euros

5.4.6.6 Funding and partnerships

For the funding of the project, applications should be submitted to FCT projects calls, to fund the M.Sc. grant and the travel and equipment values. As FCT does not fund personnel with a permanent contract, a research proposal should be submitted to the Horizon Europe, INTERREG, ANI or similar calls to fund or co-fund the senior researchers and IT personnel.

Regarding partnerships, the creation of knowledge will benefit from collaboration in the teams involved in data gathering and processing at the universities. Examples include the Algarve, Aveiro and Lisbon universities, in continuation of projects such as UBEST, AQUAMON, CONPRAR and SPRES/pac:man.

5.4.7 Project 2.6 - On-demand Coastal Digital Twin Service (RD)

5.4.7.1 Rationale

The experience of operating the OPENCoastS service for almost 5 years showed that many users wanted to build their own forecasts systems for their coastal site of interest if a user-friendly tool was available to guide them, hiding the inherent complexity of operating complex numerical models in e-infrastructures and providing the necessary computational resources. This success motivates the adaptation of the web-based, on-demand concept to CDT. Since CDT provides a virtual representation of a physical asset that can be used for multiple applications such as real-time forecast, system optimization, monitoring and controlling, and support enhanced decision making, facilitating the access to these tools and allowing users to configure them for their specific goals can be a valuable contribution towards achieving UN's SDG and the Ocean decade goals.

Developing useful on-demand CDT requires a careful balance between the flexibility and features offered at the web interface and the usability and friendliness of the service. Taking advantage of the features and tools developed in the previous projects, a huge number of features can be built and integrated as an option in the interface. However, this complexity may deter non-expert users and limit the usefulness of the tool for management purposes. Therefore, usability and portal conceptualization is a key challenge in this project.

Following the same open science vision of OPENCoastS, the applicant proposes an open access service for on-demand CDT. Balancing computational resources for multiple users in a very demanding context constitutes a new challenge that requires the best usage of the several e-core services for shared computational resources developed in the EOSC community. Dedicated implementations of the on-demand CDT may be necessary for specific applications. The platform should be thus developed as a stand-alone bundle that can be implemented easily in distinct infrastructures, taking advantage of technology such as docker or udocker (Gomes et al., 2018).

5.4.7.2 Goals

This project aims at developing an on-demand CDT portal to support CDT creation for specific user needs. The following specific goals apply:

- develop a set of scientific and technical requirements that address usability, adequate content and friendliness in a innovative way
- integrate the multiple features developed in the previous projects
- develop an optimized workflow for shared e-infrastructures usage that promote the maximum exploitation of computational resources
- validate the portal for multiple thematic and community (site) CDT

5.4.7.3 Methodology and task description

The project will start with a detailed requirements analysis both from the scientific and computational perspective. Integration of the selected features resulting from previous projects along with linkage with the CDT core framework is performed next, and the outcome is validated for both thematic and community portals. User training both on the on-demand portal and on the specific application portals is included as a key element for user uptake.

Activity 1 - Conceptualization and requirements analysis

Task 1: Identification of the core features for implementation

Task 2: Usability requirements for on-demand CDT

Activity 2 - Development of CDT-as-a-service platform and backend

Task 1: Development of the frontend application according to requirements

Task 2: Integration of CDT on-demand hybrid forecast in the backend

Task 3: Integration of CDT on-demand data assimilation engine

Task 4: Integration of reliability engine

Task 5: Integration of on-demand processing tools and indicators engine

Activity 3 - Development of optimized workflow for computational resources usage

Task 1: Identification of the relevant EOSC core e-services for CDT usage

Task 2: Implementation of these services in the on-demand CDT processing workflow

Task 3: Test at INCD infrastructure

Task 4: Users training on the on-demand platform application

Activity 4 - Application of on-demand CDT to thematic digital twins portals creation

Task 1: Identification of a few themes for implementation in cooperation with end-users

Task 2: Creation of simple thematic digital twins

Task 3: Validation in selected coastal sites

Task 4: Training for end-users and general public on the platform use

Activity 5 - Application to community portals creation dedicated at specific coastal sites to support governance

Task 1: Identification of a few communities for implementation in cooperation with end-users

Task 2: Creation of simple community digital twins

Task 3: Validation in specific “what-if” scenarios at the sites

Task 4: Training for end-users and general public on the platform use

Activity 4 - Dissemination and education

Task 1: Presentation at national and international congresses and conferences

Task 2: Publication in conference proceedings and WoS journals

Task 3: Elaboration of Ph.D. dissertation

These atividades are expected to be executed over the timeline presented in Table 5.17.

Table 5.17 - Project 2.6: Timeline

Year	Year 1				Year 2				Year 3				Year 4			
Activity	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**
1 - Conceptualization and requirements analysis																
2 - Development of CDT-as-a-service platform and backend																
3 -Development of optimized workflow for computational resources usage																

Year	Year 1				Year 2				Year 3				Year 4			
Activity	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**
4 - Application of on-demand CDT to thematic digital twins portals creation																
5 - Application to community portals creation dedicated at specific coastal sites to support governance																
6 - Dissemination and education																

5.4.7.4 Expected outcomes

The outcomes of the project will be divided in two categories: contributions to the advance of the state of the art and benefits for LNEC.

In the first category, this research project will:

- develop CDT-as-a-service platform for generation of user selected thematic or collaborative CDT

LNEC will benefit of the project by:

- having the capacity to build thematic and collaborative CDTs as part of consultancy projects
- exploring CDTs build for specific scientific questions to perform research and advance the state of the art without the need to setup models and evaluate data quality, having trustworthiness in the processing analysis and being able to create knowledge from AI applications
- get a world recognition for being a lead institute in developing tailored CDT

5.4.7.5 Resources

The list below summarizes the expected costs associated with this work.

a) Personnel costs
    - 24 months of Ph.D. grant
    - 12 months of senior researchers
    - 4 months of computer expert
    - Total: 109 871,36 euros
b) Travel: 4 participation in conferences (2 national, 2 Europe) - 8500 euros
c) Equipment:
    - laptop - 1500 euros

- e-infrastructure resources - no costs (assuming that an application to free resources is done in the Rede Nacional de Computação Avançada calls)

d) Open access publication fees: 3000 euros (fee waivers will be promoted, one publication is expected to be in a journal without waivers)

e) Grand total (with overheads): 206 071,36 euros

5.4.7.6 Funding and partnerships

For the funding of the project, applications should be submitted to FCT project and grant calls, to fund the PhD grant and the travel and equipment values. As FCT does not fund personnel with a permanent contract, a research proposal should be submitted to the Horizon Europe, INTERREG, ANI or similar calls to fund or co-fund the senior researchers. The EU Missions calls are a good example of adequate calls as well as the calls dedicated to Digital Twins.

Regarding partnerships, many can be established for this project with the partners/funders of OPENCOastS development and follow-up initiatives (CNRS, VIMS, NOAA, UN Decade CoastPredict team) and usage for coastal applications (University of Lisbon, Aveiro and Algarve). New partnerships should also be built with end-users and consultancy companies, framed in the scope of the Digital Innovation Hub ATTRACT-DIH (led by the universities of Minho and Oporto). The implementation of the on-demand CDT should be done in close cooperation with those that should be using it in the future. The partnerships with the two UN-Decade Programs linked with this area (DITTO and CoastPredict) are already being built and will be an important part of the success of this project.

Finally, building and maintaining an on-demand CDT in Portugal should be a joint effort of all state labs operating in this area. Therefore, a long-term collaboration protocol with IPMA and IH will be promoted to frame this effort. Maintaining thematic CDT should also be done with the entities (research and management) that most profit from the availability of the service. For instance, a thematic CDT for water quality should be supported both financially and in collaborative development/upgrade with port authorities, APA or the relevant municipalities.

5.4.8 Project 2.7 - Early warning, multiple information source-based platform for multi-hazards (SD)

5.4.8.1 Rationale

Emergency situations in the coastal zone are often a combination of the action of multiple hazards. Unfortunately, most of the existing early-warning systems do not account for the several sources of danger or concentrate the most accurate tools for one particular hazard, addressing the others in simplistic or correlated ways. For instance, the combined action of waves, surges and tides can cause flooding, erosion and dune and structure overtopping in many coastal regions (Rocha et al., 2021). Another example are contamination events in coastal lagoons with high mobile tidal inlets, with water quality and erosion/inlet closure concerns that need to be included in the virtual asset. Any early-warning system that addresses these situations needs to include a combination of targeted campaigns and real-time data that measure all phenomena at stake and can be used to develop comprehensive

monitoring platforms. These monitoring platforms can support the development of prediction tools that address all hazards in an integrated way. The continuous usage of the prediction tools provides information to revisit the location of the monitoring stations, adapting it to the ever-changing environment at the coast, providing a two-way interaction between data and models, besides the reliability service described above.

Herein, a generic methodology focused on multi-hazard coastal alert is proposed, along with its implementation in the on-demand CDT framework and platform, motivated by the work developed in the MOSAIC project for inundation and erosion for the Portuguese coast.

5.4.8.2 Goals

The goal of this project is to build a multi-hazard layer within the CDT prepared to build emergency early warning systems for a set of user-selected hazards, combining data and prediction tools within the CDT to support timely and high precision warnings. The specific goals of the project are:

- to build a dashboard for user interaction and feature selection, after defining the requirements for the early warning procedure
- to build a forecast workflow based on the typology of the processes selected
- to build a monitoring infrastructure combining data and models and their assimilation
- to develop a generic workflow for early-warning issuing based on the information from the prediction and data tools

5.4.8.3 Methodology and task description

The work will start with the definition of early-warning, model and data configurations for the multi-hazard CDT, along with the platform for its setup. Afterwards, an intelligent monitoring tool will be built according to user choices, including an information reliability and quality assessment. Finally, the early-warning will be set up using the on-demand CDT and validated in distinct types of multi-hazard use cases.

Activity 1 - Requirements analysis

Task 1: Definition of the early-warning criteria for multiple hazards

Task 2: Definition of the features for user-interaction dashboard

Task 3: Definition of the data sources and modeling tools for multiple hazards

Activity 2 - Combining information assets for intelligent monitoring

Task 1: Implement a multi-hazard CDT using the on-demand portal

Task 2: Process feature user selection at the dashboard

Task 3: Establishing on-demand forecast systems for multi-hazard typology

Task 4: Establishment of data sources acquisition to optimize operation

Task 5: Define workflow of information integration and reliability assessment: from models to data assets or bidirecional

Activity 3 - Implementation and validation of early-warning in generic, on-demand CDT

Task 1: Implement and operate early-warning system using the on-demand CDT for thematic applications

Task 2: Validate in distinct multi-hazard configurations and test sites

Activity 4 - Dissemination and education

Task 1: Presentation at national and international congresses and conferences

Task 2: Publication in conference proceedings and WoS journals.

Task 3: Elaboration of Ph. D thesis

Task 4: Training to users for several types of multi-hazard warning

These atividades are expected to be executed over the timeline presented in Table 5.18.

Table 5.18 - Project 2.7: Timeline

Year	Year 1				Year 2				Year 3			
Activity	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**
1 - Requirements analysis												
2 - Combining information assets for intelligent monitoring												
3 - Implementation and validation of early-warning in generic, on-demand CDT												
4 - Dissemination and education												

5.4.8.4 Expected outcomes

The outcomes of the project will be divided in two categories: contributions to the advance of the state of the art and benefits for LNEC.

In the first category, this research project will build:

- multi-hazard CDT for site-independent emergency support
- Site independent early-warning system built with the on-demand CDT for thematic applications

LNEC will benefit of the project by the availability:

- on-demand CDT to build multi-hazard emergency platforms tailored to user needs to be used in consultancy projects
- exploratory platform to study impacts of climate change and anthropogenic actions on hazard analysis on the coast

5.4.8.5 Resources

The list below summarizes the expected costs associated with this work.

a) Personnel costs
    - 24 months of Ph.D. grant
    - 12 months of senior researchers
    - 2 months of computer expert
    - Total: 102 871,36 euros
b) Travel: 4 participation in conferences (2 national, 2 Europe) - 8500 euros
c) Equipment:
    - laptop - 1500 euros
    - e-infrastructure resources - no costs (assuming that an application to free resources is done in the Rede Nacional de Computação Avançada calls)
d) Open access publication fees: 3000 euros (fee waivers will be promoted, one publication is expected to be in a journal without waivers)
e) Grand total (with overheads): 192 071,36 euros

5.4.8.6 Funding and partnerships

For the funding of the project, applications should be submitted to FCT project and grant calls, to fund the PhD grant and the travel and equipment values. As FCT does not fund personnel with a permanent contract, a research proposal should be submitted to the Horizon Europe, INTERREG, ANI or similar calls to fund or co-fund the senior researchers. Given the capacity and the characteristics of the final tool to support country wide hazard prediction and support for emergency, co-funding from APA and the Civil Protection should be sought in partnership with local end users (municipalities, businesses at the coast,,..) through collaborative funding schemes. The on-going project MOSAIC will also contribute to preliminary definition of the requirements for the multi-hazard CDT, while ATTRACT DIH activities may also fund some of the IT developments.

Partnerships can be established for this project with the collaborations from previous projects in the risk area (such as SPRES, MADyCOS, pac:man, MOSAIC, MOLINES), such as the University of Lisbon, Aveiro and Algarve in Portugal and CNRS/La Rochelle and IH Cantabria abroad. New partnerships should also be built with end-users and consultancy companies, framed in the scope of the Digital Innovation Hub ATTRACT-DIH (led by the universities of Minho and Oporto).

5.4.9 Project 2.8 - Digital twin for short and long term inundation management (SD)

5.4.9.1 Rationale

Accurate inundation forecasts are crucial to predict in a timely way extreme sea levels that flood lowland coastal and estuarine areas and generate saline plumes that infiltrate shallow coastal aquifers and endanger water uptakes. Often these high water levels occur simultaneously with large precipitation events that further enhance the risk of flooding in coastal cities as drainage networks are often insufficient for fast water drainage in presence of high water levels at receiving water bodies. In order to obtain accurate time of arrival of coastal and estuarine flood and flooding areas it is essential to predict the hydrodynamics (and often the morphodynamics) from the ocean to the land scales. Many forecast systems exist but the vast majority of these systems account for hydrodynamics only, neglecting morphodynamic effects, and do not integrate real time data at the adequate space and time scales. Moreover, short term and long term inundation management is performed in a decoupled way, often with tools of distinct precision.

The proposed CDT addresses these challenges and promotes an innovative surveillance system that allows users to configure their digital twin for inundation integrating a new monitoring system, hybrid forecasting and scenario analysis to seamlessly address inundation for multiple users in any coastal system accounting for the relevant processes. This CDT will be built on top of the research proposed in Research Study 1 and the previous projects in Research Study 2.

5.4.9.2 Goals

The goal of this project is to build a CDT for inundation management accounting simultaneously for short term events and long term analysis (considering climate change or anthropogenic changes). This innovative CDT will allow for automatic integration of the daily simulations and data acquisition in the long term scenario analysis, allowing for a continuous update of hazard quantification. Short term management will benefit by data and models usage in an interactive to set up an innovative, generic local remote sensing monitoring that provides continuous assessment of hazard conditions for both circulation and morphodynamics.

5.4.9.3 Methodology and task description

The work will start with user requirements, followed by the short term tools development (monitoring and forecasting), followed by the long term database and management intelligent tool for the continuous surveillance and hazard preparedness.

Activity 1 - Configuration setup for a core inundation CDT

Task 1: Identify the relevant generic processes for inundation in coastal regions

Task 2: Identify the relevant components from the core CDT for circulation and morphodynamics and establishment of options for user-specified process and tool selection using the multiple hazards tool

Task 3: Identification of the scenarios for the database creation for long term analysis

Activity 2 - Short term inundation detection based on intelligent local camera deployment and processing

Task 1: Definition of the methodology for assessment of the best location for camera placement and camera installation

Task 2: Definition of hotspots for alert virtual stations through numerical modeling processing

Taks 3: Definition of alert levels and associated thresholds

Task 4: Testing and validation at data rich use cases

Activity 3 - Short term inundation platform based on intelligent local data processing and hybrid inundation forecast services

Task 1: Application of the multiple hazard dashboard (previous project) for user-specified process and tool selection (numerical model, data sources, inundation thresholds for multiple types of coastal observatories)

Task 2: Integration of camera detection tool providing updates and initial conditions for predictions

Task 3: Operation of the hybrid circulation and morphodynamic forecasting platform from Research Study 1

Task 4: Application of the early warning service in coastal digital twin inundation infrastructure

Task 5: Testing and validation at data rich use cases

Activity 4 - Long term, adaptive inundation platform based on scenarios and short term outputs

Task 1: Establishment of long-term scenario database using a dedicated dashboard built on top of the on-demand CDT and a suite of bathymetric typologies

Task 2: Creation of inundation hazard CDT for long term management purposes, accounting for adaptive update using short term results and applying dedicated indicators at user-selected hotspots

Task 3: Demonstration at selected use cases under dynamic conditions

Activity 5 - Dissemination and education

Task 1: Presentation at national and international congresses and conferences

Task 2: Publication in conference proceedings and WoS journals

Task 3: Elaboration of M.Sc. thesis

Task 4: User training on short a long term inundation analysis CDT

These atividades are expected to be executed over the timeline presented in Table 5.19.

Table 5.19 - Project 2.8: Timeline

Year	Year 1				Year 2				Year 3			
Activity	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**
1 - Configuration setup for a core inundation CDT												
2 - Short term inundation detection based on intelligent local camera deployment and processing												
3 - Short term inundation platform based on intelligent local data processing and hybrid inundation forecast services												
4 - Long term, adaptive inundation platform based on scenarios and short term outputs												
5 - Dissemination and education												

5.4.9.4 Expected outcomes

The outcomes of the project will be divided in two categories: contributions to the advance of the state of the art and benefits for LNEC.

In the first category, this research project will develop:

- a CDT for inundation management accounting simultaneously for short term events and long term analysis (considering climate change or anthropogenic changes)
- a methodology for real time detection of inundation using local remote sensing using camaras

LNEC will benefit of the project by:

- a generic CDT infrastructure to handle inundation at short and long term to use in consultancy projects for the emergency authorities
- a tool for easy replication for real time detection of inundation using local remote sensing using camaras

5.4.9.5 Resources

The list below summarizes the expected costs associated with this work.

a) Personnel costs
   - 24 months of Ph.D. grant
   - 12 months of senior researchers
   - 2 months of computer expert

- Total: 102 871,36 euros

b) Travel: 4 participation in conferences (2 national, 2 Europe) - 8500 euros
c) Equipment:
    - laptop - 1500 euros
    - e-infrastructure resources - no costs (assuming that an application to free resources is done in the Rede Nacional de Computação Avançada calls)
d) Open access publication fees: 3000 euros (fee waivers will be promoted, one publication is expected to be in a journal without waivers)
e) Grand total (with overheads) 192 071,36 euros

5.4.9.6 Funding and partnerships

For the funding of the project, applications should be submitted to FCT project and grant calls, to fund the PhD grant and the travel and equipment values. As FCT does not fund personnel with a permanent contract, a research proposal should be submitted to the Horizon Europe, INTERREG, ANI or similar calls to fund or co-fund the senior researchers. Given the capacity and the characteristics of the final tool to support country wide inundation hazard prediction and support for emergency, co-funding from APA and the Civil Protection should be sought in partnership with local end users through collaborative funding schemes. The on-going project MOSAIC will also contribute to the preliminary definition of the requirements for the inundation CDT and some work on the local cameras data processing, while ATTRACT DIH activities may also fund some of the IT developments.

Partnerships can be established for this project with the collaborations from previous projects in the risk area (such as MOSAIC and MOLINES), such as the University of Lisbon and CES in Portugal and CNRS/La Rochelle abroad. New partnerships should also be built with end-users and consultancy companies, framed in the scope of the Digital Innovation Hub ATTRACT-DIH (led by the universities of Minho and Oporto).

5.4.10 Project 2.9 - Inundation.PT - a digital twin for inundation management at the Portuguese coast and estuaries (DEMO-PT)

5.4.10.1 Rationale

In the last decade, the Portuguese coast has been subject to numerous inundation events (Tavares et al., 2021), as a result of the highly energetic wave regime and frequent storm surges. The events had severe human and material consequences as most economic assets are located at the coast, along with a high population density.

In spite of this importance and vulnerability, real time, high accuracy inundation management systems do not exist, except for several time-limited research tools, promoted by universities and research institutes. Wave and storm surge forecasts are operated by the official entities and research institutes such as LNEC, based on high resolution numerical models, but their impact on the coast on flood or structure overtopping is not produced on a regular basis to support inundation management. Coastal monitoring is supported by the long-term COSMO initiative on the Atlantic coast, but the transitional waters only have a few water level monitoring and even less frequent camera surveillance. HR radars

at strategic points, operated by IH, contribute to the monitoring effort. Long-term management is guided in general to the compliance to the Flood Directive although some more detailed efforts exist at a local scale. Finally, the availability of services such as OPENCoastS, to produce on-demand forecasting of inundation, or WORSICA, to produce inundation lines from remote sensing, provide the opportunity along with the other monitoring sources to build reliable, high accuracy CDT that allow end-users to promote collaborative short and long term inundation management.

5.4.10.2 Goals

This project aims at developing a digital twin for inundation in the Portuguese coast and apply it for collaborative inundation management at selected hotspots.

5.4.10.3 Methodology and task description

This work will be built using the tool from the previous project accounting for the characteristics of the Portuguese coast and the needs of Portuguese end-users, from managers, authorities, companies and other entities with an interest on the coast. It will start with requirements from the users to the scientists, targeting collaborative decision making, including the identification of the most critical hotspots. The multiple data sources currently deployed at the Portuguese coast will be then integrated into CDT's data lake along with dedicated implementations of hybrid forecasting.

Activity 1 - Requirements analysis for addressing inundation management in Portugal

Task 1: Review of existing legislation and identification of end-users from local to national breath of intervention

Task 2: Collaborative initiatives for requirements identification and harmonization and identification of the most critical hotspots

Task 3: Evaluation of integration of erosion processes

Task 4: IT-based requirements analysis execution

Activity 2 - Integration of national data-sources into a national data hub for inundation

Task 1: Identification of relevant data sources and data repositories and their interoperability properties

Task 2: Creation of mechanisms for data harvesting and harmonization towards integrated data and data processing layers

Task 3: Integration of data products into web interface

Activity 3 - Implementation of observatories at the most critical hotspots

Task 1: Implementation of complementary real time in-situ and remote monitoring networks

Task 2: Implementation of assimilation data layers combining available data sources at each hotspot

Task 3: Development of on-demand forecast application to the hotspots based on the OPENCoastS framework and the assimilation layers

Task 4: Identification and implementation of the most relevant digital services, including hybrid early warning system

Task 5: Validation of the CDT services for collaborative decision making by end-users

Task 6: Dissemination and training for multiple tipologies of end-users

Activity 4 - Dissemination and education

Task 1: Training courses for end-user engagement

Task 2: Presentation at national conferences

Task 3: Publication in conference proceedings and WoS journals

Task 4: Elaboration of M.Sc. thesis

These atividades are expected to be executed over the timeline presented in Table 5.20.

Table 5.20 - Project 2.9: Timeline

Year	Year 1				Year 2				Year 3			
Activity	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**
1 - Requirements analysis for addressing inundation management in Portugal												
2 - Integration of national data-sources into a national data hub for inundation												
3 - Implementation of observatories at the most critical hotspots												
4 - Dissemination and education												

5.4.10.4 Expected outcomes

The outcomes of the project will be divided in two categories: contributions to the advance of the state of the art and benefits for LNEC.

In the first category, this research project will:

- establish an inundation CDT in Portugal, innovative in the methodology for its implementation, in particular in the setting of a national data hub and the use of hotspots as a selection strategy for the implementation of inundation observatories

LNEC will benefit of the project by:

- establishing an inundation CDT for the Portuguese coast that can support coastal management authorities as well as private companies that want to implement infrastructures or modify the coast in any way
- establish LNEC as the center of inundation research and innovation in Portugal regarding our coast

5.4.10.5 Resources

The list below summarizes the expected costs associated with this work.

e) Personnel costs
    - 12 months of M.Sc. grant
    - 6 months of a senior researcher
    - 2 months of computer expert
    - Total: 51 711,76 euros
f) Travel: 1 participation in european conference - 3000 euros
g) Equipment:
    - laptop - 1500 euros
    - e-infrastructure resources - no costs (assuming that an application to free resources is done in the Rede Nacional de Computação Avançada calls)
h) Grand total: 100 111,76 euros

5.4.10.6 Funding and partnerships

For the funding of the project, applications should be submitted to FCT project calls, to fund the M.Sc. grant and the travel and equipment values. As FCT does not fund personnel with a permanent contract, a research proposal should be submitted to ANI to fund or co-fund the senior researchers. Given the capacity and the characteristics of the final tool to support country wide inundation hazard prediction and support for emergency, co-funding from APA and the Civil Protection should be sought in partnership with local end users through collaborative funding schemes. The on-going project MOSAIC will also contribute to the preliminary definition of the requirements for the Portuguese west coast observatories.

Partnerships can be established for this project with the collaborations from previous projects in the risk area (such as MOSAIC and MOLINES), such as the University of Lisbon and CES in Portugal and CNRS/La Rochelle abroad.

5.4.11 Project 2.10 - Tagus river-to-estuary collaboratory for thematic digital twins and collaborative management (DEMO-PT)

5.4.11.1 Rationale

The Tagus estuary is the largest estuarine region in the Iberian Peninsula and holds a multitude of services of huge economic, environmental and social value. The management of this large system is

quite complex and there are often conflicting uses that require high resolution, complex tools to understand and predict its dynamics and support any interventions. At the same time, the Tagus basin is subject to a number of concerns related to inundation and erosion (Fortunato et al., 2021) and water quality (Rodrigues et al., 2020). A variety of models have been applied here (a recent review is presented at Vaz et al., 2020) to address multiple concerns from physical to water quality and ecology. At the same time, the Tagus holds several observatories supported by data (e.g. COASTNet, http://geoportal.coastnet.pt/) and integrated model and data (UBEST, http://ubest.lnec.pt/). In spite of all these efforts, no integrated infrastructure, from river to ocean, accounting for the city of Lisbon and other important cities´ drainage, is available to support management and research alike, allowing for users to interact with data and models to build customized knowledge. The work developed in the previous projects creates the opportunity for the construction of a river-to-estuary with urban dimension collaboratory that can be used to build several thematic digital twins and provide a coherent channel for information sharing.

This project aims at demonstrating how all of the products presented previously can be used jointly for a collaborative management of a conflicting-uses system, accounting for all stresses and addressing changes due to climate. Always starting from user needs and organized along thematic issues for in-depth evaluation of processes, the collaborative management platform proposed herein can be an exploration tool for specific issues and what-if questions answering, and at the same time, a platform for human intervention harmonization in a climate-change aware setting. The methodology can be applied to any estuarine and coastal system in Portugal and aims to address end-users needs at local, regional and national level.

5.4.11.2 Goals

The goal of this project is to demonstrate the integration of the multiple tools developed in the Research Program under a collaboratory-concept context that provides the backbone for the creation of multiple thematic digital twins that address the several areas of concern of the largest estuary in the Iberian Peninsula. A user-centered, participatory view is behind all work, providing for the first time an infrastructure for knowledge creation and sharing, mixing research and management towards collaborative actions. The specific goals are:

- Mixing data-rich repositories with numerical models and AI to build hybrid forecast systems for all relevant processes
- Evaluation the impact of the city of Lisbon and climate change on the estuary dynamics from circulation to biogeochemistry through a high-resolution scenario approach
- Addressing joint floods and droughts in the Tagus through an integrated river-to-ocean approach that addresses risk at the highest spatial and temporal resolution
- Provide a collaborative forum for integrated knowledge facilitation and sharing to bring together all users and their interests, thus providing a solid, science-based, harmonization-targeted tool for sustainable use of this valuable asset.

5.4.11.3 Methodology and task description

Each of the following tasks aim at addressing the above goals.

Activity 1 - Hybrid forecast system for physical to biochemistry processes using on-demand digital twin with OPENCoastS

Task 1: Development of a data-based forecast system based on in-situ and satellite and camera remote sensing data integration for multiple processes (water samples, sensors, drones, ship-of-opportunity, cameras, satellite, …)

Task 2: Implementation of urban discharge models in UBEST numerical forecasts

Task 3: Construction of hybrid forecast systems through integration of data-based forecasts and enhanced UBEST numerical model forecasts

Task 4: Implementation simple on-demand water quality simulator for multiple sources of contamination

Task 5: Customization of twin platform and outputs in CDT for the several thematic twins support

Activity 2 - Urban fingerprint digital twin

Task 1: End-user centered requirements for collaborative use of the estuary

Task 2: Anthropogenic and climate change scenario database building using hybrid forecast system

Task 3: Development of tools and indicators for hands-on management simulator

Activity 3 - River-to-estuary inundation and salinity management digital twin

Task 1: End-user centered requirements for collaborative use of the river to estuary dimensions in water exceedance and shortage conditions

Task 2: Integration of agricultural water pumping models in hybrid forecast

Task 3: Integration of data-based AI dams model

Task 4: Implementation of early warning systems for exceedance of salinity at water intake locations

Task 5: Implementation of early warning systems for local inundation events

Activity 4 - Collaborative management digital platform

Task 1: Integration of multiple thematic twins and creation of hot spot products and alerts

Task 2: Integration of simulation for on-demand “What-if” scenarios simulation and automatic assessment over the thematic twins

Task 3: Training for end-users and general public on the platform use

Activity 5 - Dissemination and education

Task 1: Presentation at national and international congresses and conferences

Task 2: Publication in conference proceedings and WoS journals

Task 3: Elaboration of several M.Sc. theses and a Ph.D. thesis

These atividades are expected to be executed over the timeline presented in Table 5.21.

Table 5.21 - Project 2.10: Timeline

Year	Year 1				Year 2				Year 3				Year 4			
Activity	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**
1 - Hybrid forecast system for physical to biochemistry processes using on-demand digital twin with OPENCoastS																
2 - Urban fingerprint digital twin																
3 - River-to-estuary inundation and salinity management digital twin																
4 - Collaborative management digital platform																
5 - Dissemination and education																

5.4.11.4 Expected outcomes

The outcomes of the project will be divided in two categories: contributions to the advance of the state of the art and benefits for LNEC.

In the first category, this research project will:

- build a “system of systems” using on-demand CDT to build multiple thematic CDTs and an overall CDT for collaboratory purposes;
- demonstrate in a complex estuarine and coastal system how CDT can be a comprehensive management tool, addressing multiple concerns at the same time and providing a collaboratory environment where users can experiment and assess impact measures aiming at a compromise that best serves the interests of all.

LNEC will benefit from the project by:

- building a comprehensive CDT for the Tagus estuary and nearby coastal area that addresses multiple concerns through high accuracy hybrid models that can be used to answers multiple end user needs through consultancy projects

- building the backbone for the continuation of the estuarine and coastal process research, started more than 3 decades ago, based on the state-of-the-art tools that seamlessly integrate data and models.

5.4.11.5 Resources

The list below summarizes the expected costs associated with this work.

a) Personnel costs
    - 48 months of Ph.D. grant
    - 18 months of senior researchers
    - Total: 157 542,72 euros
b) Travel: 4 participation in conferences (2 national, 2 Europe) - 8500 euros
c) Equipment:
    - laptop - 1500 euros
    - e-infrastructure resources - no costs (assuming that an application to free resources is done in the Rede Nacional de Computação Avançada calls)
d) Open access publication fees: 3000 euros (fee waivers will be promoted, one publication is expected to be in a journal without waivers)
e) Grand total: 273942,72 euros

5.4.11.6 Funding and partnerships

LNEC has a long tradition of research and consultancy projects involving the Tagus estuary and its bordering areas, for multiple end-users (municipalities around the estuary, Water authority, Port of Lisbon, etc..) with several partners from academia and industry. The work proposed herein aims at addressing several concerns voiced by multiple end-users in a concerted way that looks at each problem in a multi-scale, multi-process dimension, facilitating end-user daily interventions and long term planning.

Therefore, it is expected that funding will be obtained in the short term or already at place from:

- AML project for inundation (in preparation in partnership with multiple players developing activity in the Tagus river and estuary)
- H2020 EGI project (providing the necessary developments in OPENCoastS)
- ATTRACT DIH, through its national (IAPMEI) and european funding (HEuropa) as an opportunity to joint estuarine and coastal experts with computer scientists for the development of the intelligent monitoring framework and customizable on-demand forecast

Given the importance of the Tagus estuary, other projects will be proposed to both local stakeholders and national and international funding agencies.

Part B - Post-graduation Program

6 Post-graduation Program

6.1 Introduction

The present Post Graduation Program, part of this Habilitation program, is expected to be developed at national and international universities in close collaboration with LNEC, in particular with GTI and other divisions at LNEC working in estuarine and coastal zones research. It aims at addressing the scientific gaps identified in the state-of-the-art (chapters 3 and 4) and contributing towards the implementation of the Research studies (chapter 5).

The Post Graduation Program aims also at educating and giving research and technical skills to young researchers in the Maritime Hydraulics area framed in the Information Technology subarea. It targets a new generation of researchers for the Hydraulics an Environment Department with a strong background and competence in the Digital areas that can advance the existing research in the estuarine and coastal areas, taking advantage of IT to build new tools and methodologies that can address the challenges in the coastal areas for the next ten years.

Given the multidisciplinary nature of the research proposed in the previous chapters, the proposed plans for M.Sc. and Ph.D. degrees can be framed in several existing university programs in multiple areas such as Civil Engineering, Oceanography, Mechanical engineering, Advanced Computing/Computational Sciences and Applications, Geomatics, Information technology, Mathematics and Computer science, taking also advantage of the joint initiatives with North American universities such as the University of Texas at Austin, Carnegie Mellon University and the Massachusetts Institute of Technology.

The program proposes 6 M.Sc. and 9 Ph.D. proposals to be developed in the ten years of expected implementation of the Habilitation Program (Figure 6.1, Tables 6.1-6.2). The outcomes of these theses include Coastal Digital Twins builders, through advanced frameworks and user friendly platforms for their developments, new instruments to manage knowledge creation and user-centered management of coastal resources, tailored to specific problems such as inundation and contamination. Visionary, high accuracy collaborative tools that promote co-creation of the sustainable solutions for the coast by multiple players, from stakeholders, industry and the public, will also result from these graduation efforts. Competences in numerical modeling, data management and exploitation, forecast systems in a context of information quality and sharing will be obtained by the trainees, as a result of the proposed proposals, contributing towards DHA's future research in the scientific area of Maritime Hydraulics.

The full list of Ph.D. dissertations and M.Sc. theses are:

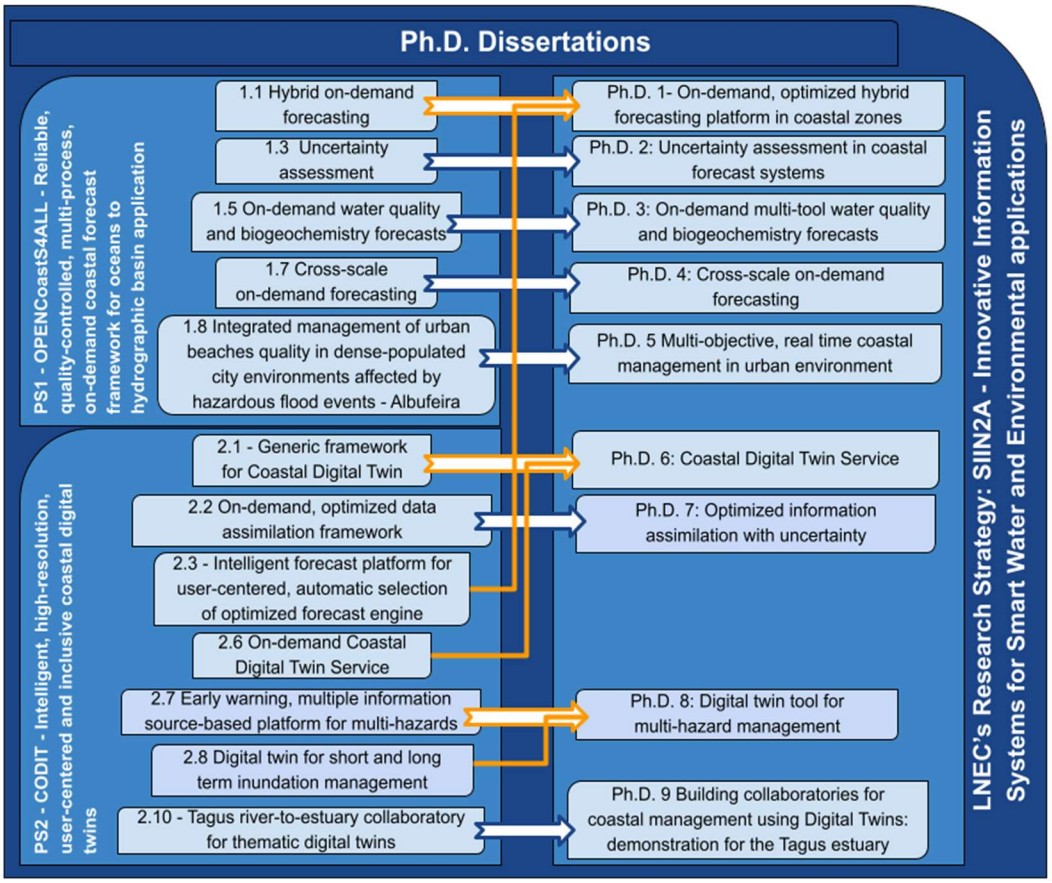
1) Ph.D. dissertations
   - Ph.D. 1: On-demand, optimized hybrid forecasting platform in coastal zones
   - Ph.D. 2: Uncertainty assessment in coastal forecast systems
   - Ph.D. 3: On-demand multi-tool water quality and biogeochemistry forecasts

- Ph.D. 4: Cross-scale on-demand forecasting for generic management of coastal zones
- Ph.D. 5: Multi-objective, real time coastal management in a Portuguese urban coastal environment
- Ph.D. 6: Coastal Digital Twin Service
- Ph.D. 7: Optimized information assimilation with uncertainty consideration
- Ph.D. 8: Digital twin tool for multi-hazard management and application to short and long term inundation
- Ph.D. 9: Building collaboratories for coastal management using Digital Twins and demonstration for the Tagus estuary

2) M.Sc. theses
   - M.Sc. 1: Automatic grid generator for forecast applications
   - M.Sc. 2: Morphodynamic updater in forecast platforms
   - M.Sc. 3: Operational fast assessment pollution tool
   - M.Sc. 4: Information quality assessment for coastal applications
   - M.Sc. 5: Multi-dimensional, automatic and on-the-fly data analysis in a Coastal Digital Twin
   - M.Sc. 6: Inundation management in Portugal using Digital Twins

The post-graduation proposals involve the availability of several grants framed in the scope of national or european research projects or national calls for Ph.D. grants. Several on-going and proposed projects at GTI are already planned to fund several of these grants. The costs associated with traveling for presentation of the post-graduation work results will also be considered in the projects/FCT grants or obtained at competitive grant programs funded for instance by FLAD (https://www.flad.pt/papers-usa/).

a)



b)

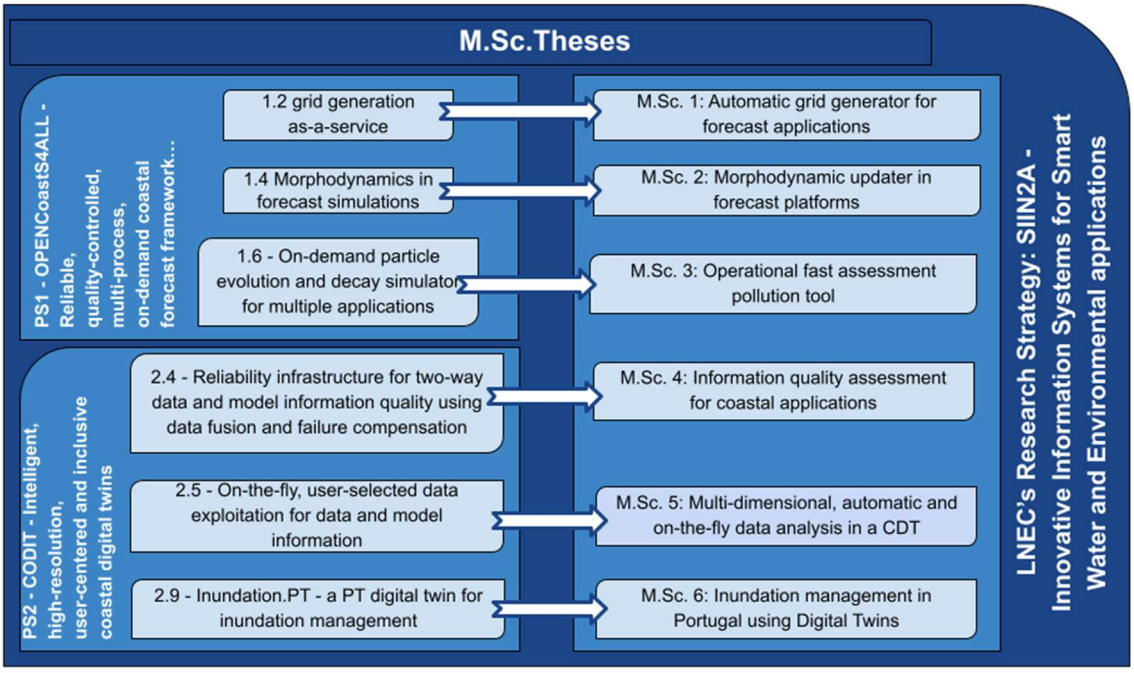


Figure 6.1 – Framing of the post-graduation research program in the Research Studies and LNEC's strategic research: a) Ph.D. dissertations b) M.Sc. theses.

Table 6.1 - Summary of the Ph.D. proposals and their relationship with the Research projects

Projeto	Ph.D Proposals								
	Ph.D. 1 On-demand, optimized hybrid forecasting platform in coastal zones	Ph.D. 2 Uncertainty assessment in coastal forecast systems	Ph.D. 3 On-demand multi-tool water quality and biogeochemistry forecasts	Ph.D. 4 Cross-cale on-demand forecasting for generic management of coastal zones	Ph.D. 5 Multi-objective, real time coastal management in urban environment	Ph.D. 6 Coastal Digital Twin Service	Ph.D. 7 Optimized information assimilation with uncertainty consideration	Ph.D. 8 Digital twin tool for multi-hazard management and application to short and long term inundation	Ph.D. 9 Building collaboratories for coastal management using Digital Twins: demonstration for the Tagus estuary
1.1 Hybrid on-demand forecasting	●								
1.3 Uncertainty assessment		●							
1.5 On-demand water quality and biogeochemistry forecasts			●						
1.7 Cross-scale on-demand forecasting				●					

Projeto	Ph.D Proposals								
1.8 Integrated management of urban beaches quality in dense-populated city environments affected by hazardous flood events - Albufeira					●				
2.1 Generic framework for Coastal Digital Twin						●			
2.2 On-demand, optimized data assimilation framework							●		
2.3 Optimized forecast engine	●								
2.6 On-demand Coastal Digital Twin Service						●			
2.7 Early warning, multiple information source-based platform for multi-hazards								●	

Projeto	Ph.D Proposals								
2.8 Digital twin for short and long term inundation management								●	
2.10 Tagus river-to-estuary collaboratory									●

Table 6.2 - Summary of the M.Sc. proposals and their relationship with the Research projects

Projeto	M.Sc. Proposals					
	M.Sc. 1 Automatic grid generator for forecast applications	M.Sc. 2 Morphodynamic updater in forecast platforms	M.Sc. 3 Operational fast assessment pollution tool	M.Sc. 4 Information quality assessment for coastal applications	M.Sc. 5 Multi-dimensional, automatic and on-the-fly data analysis in a Coastal Digital Twin	M.Sc. 6 Inundation management in Portugal using Digital Twins
1.2 Grid generation as-a-service	●					
1.4 Morphodynamic in forecast simulations		●				
1.6 - On-demand particle evolution and decay simulator for multiple applications			●			
2.4 Reliability infrastructure				●		
2.5 User-selected data exploitation					●	
2.9 Inundation.PT						●

The financial plans will be presented in full, including an expected time for supervision of the work, assuming a university supervisor and a LNEC co-supervisor. The costs can be co-funded by FCT on the grantee monthly stipend and one attendance to a conference, but other funding sources are also available. As the co-funded value will depend on the supporting project's funding and its rules (e.g. FCT, CE, ANI, INTERREG all have different funding schemes), the full budget without any co-funding is

presented, organized by typology. Overheads will be included at LNEC rates (100% for LNEC's human resources, 20% for travel costs and equipment).

6.2 Ph.D. 1: On-demand, optimized hybrid forecasting platform in coastal zones

6.2.1 Rationale

The lag of time between alert issuing and the occurrence of a hazardous event can have huge impacts in emergency operations to assure the protection of people and assets, having very important economic, social and ecological benefits. The large computational cost of process-based models can be sometimes a limiting factor for very complex emergency simulations, in particular when applied to high-resolution large domain forecasts or applications in low income countries where computational resources may be scarce. Hybrid or surrogate models are an attractive alternative to process/based models, being several orders of magnitude faster than equivalent numerical models. They are therefore a natural candidate for forecast workflows as a stand/alone tool or combined with process/based models used in scenario mode. Coastal hybrid forecasting has not been developed yet and there is a clear societal need for these tools.

This proposal aims at extending the work developed by LNEC in the area of relocatable and on-demand coastal forecast systems. The OPENCoastS framework and its forecast engine WIFF can thus be improved to build an optimized forecast platform that advances the state-of-the-art by using hybrid models in forecast frameworks, with advantages on accuracy, speed of response and robustness.

The use of hybrid modeling requires the specification of weights to be attributed to each modeling component, to account for the specific environmental conditions and model uncertainty. The relative weight of data-based or process-based models in a hybrid forecast simulation could be optimized, taking advantage of assimilation and indicator calculations. These optimized settings can be made available to the users in the on-demand dashboards, allowing the user to configure, constrain and select the procedure to his/hers goals. Integration in OPENCoastS of the optimized feature provides the user with the most efficient and accurate solution, thus contributing to a novel strategy of computational resource efficiency at the maximum accuracy and speed of results availability.

This innovative solution will extend the robustness, accuracy and applicability of coastal forecast systems and provide the grounds for better and faster timely early-warning of coastal phenomena at any time- and space scales and accounting for multiple complex processes.

This proposal aims at contributing to projects 1.1 "Hybrid on-demand forecasting in coastal regions: a melange of process-based and data-based models", of Research Study 1 "OPENCoastS4ALL - Reliable, quality-controlled, multi-process, on-demand coastal forecast framework for oceans to hydrographic basin application" and project 2.3 "Intelligent forecast platform for user-centered, automatic selection of optimized forecast engine" of Research Study 2 "CODIT - Intelligent, high-resolution, user-centered and inclusive coastal digital twins".

6.2.2 Goals

The main goal of this Ph.D. proposal is to develop an innovative, optimized, hybrid forecast engine and implement it in the OPENCoastS platform for on-demand, hybrid forecasting. The specific goals are:

- to conceptualize, develop and implement a hybrid modeling methodology appropriate for coastal environments based on the WIFF framework;
- to identify and selection the error measures and control parameters for optimization algorithm and the definition of the new, smooth transition procedure between distinct weighted runs;
- to create, develop and implement an optimization algorithm based on the previous hybrid engine to maximize accuracy of the hybrid procedure under distinct environmental conditions and accounting for uncertainty in the whole modeling/monitoring procedure;
- to implement the above results in the OPENCoastS platform.

6.2.3 Methodology and activity planning

The activities to be developed in this Ph.D. dissertation are:

Activity 1 - State-of-the-art review and AI framework development

Task 1: Review and identification of relevant AI methodologies and their comparison

Task 2: Implementation of AI methodology or adaptation of existing tools for coastal dynamics prediction

Task 3: Creation of the on-demand layer on top of the AI tool, by identification of the relevant variables for user selection

Task 4: Setting up of virtual research environment for on-demand AI tool operation in shared national e-infrastructures

Activity 2 - State of the art review and development of data-based models and their integration in the hybrid framework

Task 1: Selection of generic data sources and their possibilities of integration for data-modeling building

Task 2: State-of-the-art review on data-based models

Task 3: Comparison of data-model methodology building workflows, selection of the most appropriate ones and their application in the selected AI-tools

Task 4: Validation at selected data-rich site

Activity 3 - Requirements and development of optimed forecasts

Task 1: Definition of control parameters and error measures for hydrodynamic and biogeochemistry simulations

Task 2: Selection of the optimization algorithm and transition mechanisms between forecast engines

Task 3: Development of optimized selector

Task 4: Testing and validation under controlled conditions

Activity 4 - Comparative analysis on accuracy, robustness and performance of process-based, data-model, hybrid modeling and optimized hybrid modeling on a data rich environment

Task 1: Selection of research question and site for its evaluation

Task 2: Setting up of multiple choice forecasting at selected site

Task 3: Selection of evaluation indicators

Task 4: Comparative application for a selected period and evaluation report

Activity 5 - Integration in OPENCoastS framework

Task 1: Requirements analysis for OPENCoastS integration

Task 2: Development of multi-model component

Task 3: Integration of new data sources in OPENCoastS data

Task 4: Integration of new data- based and hybrid models in OPENCoastS modeling workflow

Task 5: Integration of user-selection forecast selector engine (modeling choice, optimization parameters) and development of dedicated dashboard

Task 6: Setting up of virtual research environment for optimized hybrid forecasting in shared e-infrastructures

Activity 6 - Scientific publications and dissemination

Task 1: Conference articles preparation and public presentation (4, at the end of each year)

Task 2: Journal papers publication (3, at the end of years 2,3,4)

Task 3: Dissertation preparation

Table 6.3 summarizes the timeline associated with this work.

Table 6.3 - Timeline for Ph.D. proposal 1

Year	Year 1				Year 2				Year 3				Year 4			
Activity	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**
1 - SoA review and AI framework development																
2 - SoA review data-based models integration in the hybrid framework																

Year	Year 1				Year 2				Year 3				Year 4			
Activity	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**
3 - Requirements and development of optimed forecasts																
4 - Comparative analysis on accuracy, robustness and performance																
5 - Integration in OPENCoastS																
6 - Scientific publications and dissemination																

6.2.4 Financial plan

The list below summarizes the expected costs associated with this work.

a) Personnel costs
   - 48 months of Ph.D. grant
   - 4 months of a senior researcher
   - Total: 77,742.72 euros
b) Travel: 4 participation in conferences (2 national, 2 Europe) - 8500 euros
c) Equipment:
   - laptop - 1500 euros
   - e-infrastructure resources - no costs (assuming that an application to free resources is done in the Rede Nacional de Computação Avançada calls)
d) Open access publication fees: 3000 euros (fee waivers will be promoted, one publication is expected to be in a journal without waivers)
e) Grand total (with overheads): 114 342,72 euros

6.3 Ph.D. 2 Uncertainty assessment in coastal forecast systems

6.3.1 Rationale

As computational resources and scientific knowledge of processes grows, the complexity and breadth of application of coastal and estuarine forecasts become larger. Many forecasting systems currently in operation range from circulation to biogeochemistry and some also integrate the urban dimension, taking advantage of chains of linked models and multiple sources of initial and boundary conditions and requiring a large number of parametrizations for adequate results. In oceanography and meteorology, inclusion of uncertainty has been frequent for some time, through the use of ensemble modeling, but most coastal forecast systems do not account for its effect.

Forecast uncertainty may arise from physical processes representation and numerical formulations, as well as in the construction and application of the numerical models including uncertainties in initial conditions, model boundary conditions and model parameters. As models become more complex, for instance by integrating waves, morphodynamic or biogeochemical simulations, uncertainty importance grows as errors propagate and are amplified in the cascade of model simulations.

The present proposal aims at addressing this limitation and proposes to advance the state-of-the-art in coastal simulations by creating, implementing and testing a novel uncertainty tool that can be used for multiple types of simulations comprising all uncertainty sources.

Ensemble modeling will be the core of the new strategy, as it is emerging in other fields as one of the best solutions by minimizing the combined uncertainty in input data, model parameters and model structure, and was shown to improve forecast performance. To optimize the use of multiple models, the construction of a weighted ensemble is proposed, giving more weight for better performing models. Furthermore, the identification of the dependency of these weights with environmental and other system conditions will be included, creating on-the-fly, optimized weighted ensembles. These ensembles consider multiple external forcings (either river or atmospheric), multiple model configurations or multiple models usage. An on-going forecast evaluation of accuracy and timeliness, based on a detailed suite of error measures, calculated against real time data networks, will also be proposed.

This proposal aims at contributing to Project 1.3 - Uncertainty assessment methodology and tool, of Research Study 1 “OPENCoastS4ALL - Reliable, quality-controlled, multi-process, on-demand coastal forecast framework for oceans to hydrographic basin application”.

6.3.2 Goals

This dissertation aims at developing an uncertainty tool for multiple types of simulations comprising all uncertainty sources. In order to obtain an optimized uncertainty workflow, this tool should also provide an opportunity for the user to define which uncertainty sources should be considered and allow for the definition of the error measure indicators and the sources for the real time data networks to be used in the forecast error performance analysis. It will be integrated in a user interface for configuration of the uncertainty methodology to be implemented in each forecast application, as part of the OPENCoastS platform.

The specific goals of the project are:

- to identify and select the main sources of uncertainty in coastal forecasts;
- to identify and select a suite of error measures and an ensemble methodology for multi-model, multi-parameter and multi-input uncertainty assessment;
- to conceptualize, develop and implement an optimized algorithm for uncertainty quantification under distinct environmental conditions and automatic selection of the modeling source;
- to implement the above products in OPENCoastS as part of the quality assurance and robustness of the prediction framework.

6.3.3 Methodology and activity planning

The activities to be developed in this Ph.D. dissertation are:

Activity 1 - Requirements analysis for uncertainty evaluation

Task 1: State-of-the-art review of the uncertainty sources for each type of simulation, from processes to model configuration

Task 2: State-of-the-art review of the sources of data and error measure suite for forecast error evaluation according to simulation type

Task 3: Selection of a range of models for ensemble modeling according to simulation type

Task 4: Definition of environmental and system conditions to be used in the optimization procedure according to simulation type

Task 5: Definition of algorithm for optimization implementation

Activity 2 - Optimized ensemble engine development

Task 1: Implementation of ensemble engine, based on requirement analysis and available open software

Task 2: Test of ensemble engine for each type of simulation using a data-rich test case

Task 3: Development of optimized ensemble engine

Task 4: Test of optimized ensemble engine in several use cases with multiple data sources

Activity 3 - Integration in OPENCoastS framework

Task 1: Requirements analysis for OPENCoastS integration

Task 2: Development of user interface for uncertainty configuration

Task 3: Integration of uncertainty component (multi-simulation) engine

Task 4: Integration of user-selection engine for configuring the choice of uncertainty components for forecast application

Task 5: Setting up of virtual research environment for uncertainty simulations in shared e-infrastructures

Activity 4 - Scientific publications and dissemination

Task 1: Conference articles preparation and public presentation (4, at the end of each year)

Task 2: Journal papers publication (3, at the end of years 2,3,4)

Task 3: Dissertation preparation

Table 6.4 summarizes the timeline associated with this work.

Table 6.4 - Timeline for Ph.D. proposal 2

Year	Year 1				Year 2				Year 3				Year 4			
Activity	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**
1 - Requirements analysis for uncertainty evaluation																
2 - Optimized ensemble engine development																
3 - Integration in OPENCoastS framework																
4 - Scientific publications and dissemination																

6.3.4 Financial plan

The list below summarizes the expected costs associated with this work.

a) Personnel costs
   - 48 months of Ph.D. grant
   - 4 months of a senior researcher
   - Total: 77,742.72 euros
b) Travel: 4 participation in conferences (2 national, 2 Europe) - 8500 euros
c) Equipment:
   - laptop - 1500 euros
   - e-infrastructure resources - no costs (assuming that an application to free resources is done in the Rede Nacional de Computação Avançada calls)
d) Grand total (with overheads): 114 342,72 euros

6.4 Ph.D. 3: On-demand multi-tool water quality and biogeochemistry forecasts

6.4.1 Rationale

Accurately simulating water quality and biogeochemistry in estuarine and coastal regions is necessary to address both climate change long term trends that may have serious impact in ecological assets and to support short term emergency actions under hazardous events such as coastal eutrophication and hypoxia, exposure to waterborne pathogens or harmful algal blooms. Integrating uncertainty that

explores the multiple sources of data, including complex and low cost in-situ sensors, remote sensing or water drones, are key components of that effort, along with a modeling vision that accounts for hybrid modeling tools and to perform accurate assessments of predictions in comparison with real time data. The recent growth of computational resources and the availability of considerable shared e-infrastructures facilitating access to large HPC resources, has paved the way to address uncertainty and to perform long term and scenarios analysis to investigate climate changes or anthropogenic interventions impacts.

Two more complex and important challenges arise in this area. The first is related to data management and data lakes creation, as a backbone for the necessary data-sharing in this highly expensive monitoring area. The second, and even more important one, is the automatic assessment of data quality in order to be used in the above procedures, taking advantage of the recent work developed at LNEC (Jesus (2019)) for error detection and correction and data classification, necessary for the achievement of reliable water quality and biogeochemistry predictions.

This work aims at conceptualizing, implementing and testing a new framework for water quality and biogeochemistry prediction, targeting a flexible, accurate and user-configurable on-demand forecast framework that advances the state of the art in this area. Furthermore, the outcomes of this work will be integrated in the OPENCoastS platform for on-demand forecasting, facilitating its sharing and application with any users.

This proposal aims at contributing to Project 1.5 - On-demand water quality and biogeochemistry forecasts for multiple applications in coastal regions, of Research Study 1 “OPENCoastS4ALL - Reliable, quality-controlled, multi-process, on-demand coastal forecast framework for oceans to hydrographic basin application”. Its outcomes will be integrated in Project 2.10 - Tagus river-to-estuary collaboratory for thematic digital twins and collaborative management of Research Study 2 “CODIT - Intelligent, high-resolution, user-centered and inclusive coastal digital twins”.

6.4.2 Goals

This dissertation aims at conceptualizing and developing a new, generic water quality and biogeochemistry forecast framework and customizing it for specific problems such as algae blooms or oil spills. An innovative approach will be taken for initial conditions establishment by integrating automatic several sources of remote sensing. The forecast framework will also support the capacity to perform uncertainty analysis by integrating multiple models (data-based, process-based and hybrid models) and ways to assess uncertainty in initial and boundary conditions. Exploring technologies such as data assimilation will also be done to create an innovative tool that can accurately handle predictions for long term periods.

This forecast engine and its services will be integrated in a dedicated dashboard in the OPENCoastS platform.

The specific goals of the project are:

- to select a suite of water quality and biogeochemistry models, including data-based, process-based and hybrid, for usage in the forecast engine, including model ensembles;
- to define tailored uncertainty strategy (based on project 1.3 and the previous Ph.D. thesis) and implement it as a tool;
- to define and implement generators of initial conditions for selected problems (algae blooms, oil spills, …);
- to integrate all these tools in OPENCoastS, building a dedicated dash board for these simulations.

6.4.3 Methodology and activity planning

The activities to be developed in this Ph.D. dissertation are:

Activity 1 - Requirement analysis for on-demand water quality and biogeochemistry forecasts

Task 1: Definition of the parameters for pre-selected forecast typologies and selection of the models formulations for these typologies

Task 2: Definition of the parameters for selection in the on-demand forecast engine for pre-defined typologies

Task 2: Definition of the generic parameters for selection in the on-demand forecast engine for user-defined tracer

Activity 2 - Development of initial conditions remote sensing data-based processing tool for selected typologies

Task 1: Definition of the data sources for each tracer typology

Task 2: Development of data acquisition and input file preparation scripts

Task 3: Creation of data repository for each typology of tracer

Task 4: Testing of tool in selected use cases

Activity 3 - Development of assimilation module for water quality and biogeochemistry

Task 1: Selection of assimilation workflow and data sources for each typology

Task 2: Implementation of tailored module for each typology

Task 3: Implementation of user-selected workflow module

Task 4: Testing of final data layers for each typology in selected uses cases

Activity 4 - Development of innovative water quality and biogeochemistry forecast engine

Task 1: Implementation of a common framework for water quality and biogeochemistry

model operation under common formats and inputs

Task 2: Integration of the selected models in framework

Task 3: Development of the ensemble predictor based on any combination of models and the work of project 1.1

Task 4: Customization of predictor to address specific problems including integration of initial conditions tailored generator

Task 4: Testing of forecast engine in data-rich use cases

Activity 5 - Integration in OPENCoastS framework

Task 1: Integration of uncertainty module from project 1.3 to water quality and biogeochemistry simulations

Task 2: Integration of data sources and assimilation procedure in predictions

Task 3: Uncertainty assessment tool integration in water quality workflow and dashboard

Task 4: Integration in OPENCoastS of new water quality/biogeochemistry data sources/processing tools for model evaluation

Task 5: Integration of water quality and biogeochemistry engine in OPENCoastS through a dedicated dashboard

Task 6: Evaluation of the performance of the several options using multiple error measures

Activity 6 - Scientific publications and dissemination

Task 1: Conference articles preparation and public presentation (4, at the end of each year)

Task 2: Journal papers publication (3, at the end of years 2,3,4)

Task 3: Dissertation preparation

Table 6.5 summarizes the timeline associated with this work.

Table 6.5 - Timeline for Ph.D. proposal 3

Year	Year 1				Year 2				Year 3				Year 4			
Activity	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**
1 - Requirements analysis for on-demand water quality and biogeochemistry forecasts																
2 - Development of initial conditions remote sensing data-based processing tool for selected typologies																

Year	Year 1				Year 2				Year 3				Year 4			
Activity	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**
3 - Development of assimilation module for water quality and biogeochemistry																
4 - Development of innovative water quality and biogeochemistry forecast engine																
5 - Integration in OPENCoastS framework																
6 - Scientific publications and dissemination																

6.4.4 Financial plan

The list below summarizes the expected costs associated with this work.

a) Personnel costs
   - 48 months of Ph.D. grant
   - 6 months of a senior researcher

   Total: 89 142,72 euros

b) Travel: 4 participation in conferences (2 national, 2 Europe) - 8500 euros
c) Equipment:
   - laptop - 1500 euros
   - e-infrastructure resources - no costs (assuming that an application to free resources is done in the Rede Nacional de Computação Avançada calls)

d) Grand total (with overheads): 137 142,72 euros

6.5 Ph.D. 4: Cross-scale on-demand forecasting for generic management of coastal zones

6.5.1 Rationale

Coastal and estuarine dynamics are strongly coupled with the bordering water systems, in terms of circulation, sediment dynamics and biogeochemistry. In order to obtain reliable and accurate predictions in these regions, the consideration of the exchanges with the atmosphere, ocean, riverine and land

inputs at the adequate spatial and temporal scales at each compartment is fundamental. Forecast systems that address all relevant compartments at these scales that can be relocatable to any coastal location do not exist yet. This work aims at developing such a system, taking advantage of the work from the projects proposed in the previous chapter and the vast availability of data and computational resources. For instance, the complexity of the modeling web for a river to ocean forecasting that also accounts for the urban dimension makes uncertainty and error cascading a fundamental concern that can be now addressed by project 1.3. To minimize error generation and propagation, assimilation strategies with data from the several compartments becomes a major asset to achieve high accuracy predictions.

This proposal aims at contributing to projects 1.7 “Cross-scale on-demand forecasting from the ocean to the hydrographic basin including the urban dimension”, of Research Study 1 “OPENCoastS4ALL - Reliable, quality-controlled, multi-process, on-demand coastal forecast framework for oceans to hydrographic basin application”.

6.5.2 Goals

This work aims at developing a cross-scale forecast framework that integrates processes from the ocean to the watershed, including the urban dimension. It will account for uncertainty and integrating data assimilation procedures, applicable from circulation to biogeochemical problems, forced by the relevant circulation conditions, and include its integration in the OPENCoastS framework.

The specific goals of the project are:

- to define a framework for integrated forecasting across river-to-ocean scales to predict jointly the dynamics of the adjoint water bodies using a combination of models at the correct time and space scales;
- to select the modeling suite, combining data-models, process-based and hybrid ones to build a model cascade, linking in a bidirectional way with compartments at larger scales (e.g. linking the forecast framework with multiple providers of oceanic and atmospheric predictions to handle uncertainty and provide land and coastal inputs to these compartments);
- To the continuum of river to ocean modeling, an urban/land input component will be added, allowing also for bidirectional interactions, fundamental for including the coastal impact in the city (for instance for sea level rise or salinization of the drainage networks with potential severe impacts on WWTPs).

6.5.3 Methodology and activity planning

The activities to be developed in this Ph.D. dissertation are:

Activity 1 - Requirements analysis for cross-scale prediction

Task 1: Definition of the interdependencies among modeling compartments and data exchange protocols and standards

Task 2: Selection of river and urban modeling tools including hybrid approaches, to complement OPENCoastS current assets for the coasts

Task 3: Definition of the integrated forecast workflow, data sources for validation and assimilation and sources of uncertainty

Activity 2 - Development/adaptation of new core forecast engines

Task 1: Development/adaptation of riverine forecast engine

Task 2: Development/adaptation of urban forecast engine

Task 3: Testing of new modules in data-rich use case

Activity 3 - Integration of quality-controlled assets in cross-scale forecast framework

Task 1: Integration of new data sources for the new modeling domains

Task 2: Integration of optimized uncertainty module with adaptation to new uncertainty sources

Task 3: Adaptation and integration of assimilation module for freshwater and urban domains

Activity 4 - Integration in OPENCoastS framework

Task 1: Requirements analysis for OPENCoastS integration

Task 2: Integration of basin and river forecasting engine components in the on-demand workflow

Task 3: Integration of urban discharge engine component for both inundation and water quality in the on-demand workflow

Task 4: Development of user-selection engine for specifying water pathway components choice

Task 5: Setting up of virtual research environment for cross-scale forecasting in shared e-infrastructures

Activity 5 - Scientific publications and dissemination

Task 1: Conference articles preparation and public presentation (4, at the end of each year)

Task 2: Journal papers publication (3, at the end of years 2,3,4)

Task 3: Dissertation preparation

Table 6.6 summarizes the timeline associated with this work.

Table 6.6 - Timeline for Ph.D. proposal 4

Year	Year 1				Year 2				Year 3				Year 4			
Activity	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**
1 - Requirements analysis for cross-scale prediction																

Year	Year 1				Year 2				Year 3				Year 4			
Activity	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**
2 - Development/adaptation of new core forecast engines																
3 - Integration of quality-controlled assets in cross-scale forecast framework																
4 - Integration in OPENCoastS framework																
5 - Scientific publications and dissemination																

6.5.4 Financial plan

The list below summarizes the expected costs associated with this work.

a) Personnel costs
   - 48 months of Ph.D. grant
   - 4 months of a senior researcher
   - Total: 77,742.72 euros
b) Travel: 4 participation in conferences (2 national, 2 Europe) - 8500 euros
c) Equipment:
   - laptop - 1500 euros
   - e-infrastructure resources - no costs (assuming that an application to free resources is done in the Rede Nacional de Computação Avançada calls)
d) Grand total (with overheads): 114 342,72 euros

6.6 Ph.D. 5: Multi-objective, real time coastal management in a Portuguese urban coastal environment

6.6.1 Rationale

Coastal cities are subject to multiple risks in particular due to an increasing concentration of population in those areas and the impact of climate changes. Drainage infrastructures have a fundamental role in the quality of life of cities, avoiding inundations, by regulating precipitation effects, and avoiding water and atmospheric contamination through adequate water treatment. In a context of climate changes, their impact is even greater, due to the expected increasing of precipitation negative effects through short term inundation events (flash floods) and greater contamination discharges to nearby receiving water

bodies, often used as recreational bathing areas. The continuous increase in surface impermeabilization, the climate change related aggravated conditions in precipitation and sea level rise and the joint drainage for pluvial and wastewater flows in many cities leads to very complex conditions that need to be tackled jointly at all water compartments and address quantity and quality.

One of the current greatest challenges of drainage network infrastructure management is thus to find an optimized, multi-objective way for their real time operation that minimizes energy consumption while protecting receiving bathing waters and minimizing urban flooding. A management support system to tackle this challenge and adequately support drainage networks operation is not available yet.

This Ph.D. proposal aims at contributing to Project 1.8 “Integrated management of urban beaches quality in dense-populated city environments affected by hazardous flood events - the coastal city of Albufeira case”, of Research Study 1 “OPENCoastS4ALL - Reliable, quality-controlled, multi-process, on-demand coastal forecast framework for oceans to hydrographic basin application”. It also takes advantage of an on-going project (ANI-funded SINERGEA) to establish the demo model application and to explore the data network.

6.6.2 Goals

This Ph.D. proposal aims at developing a multi-objective, real time, optimized coastal management system for coastal cities subject to inundations and contamination events associated with their drainage system. It will take advantage of the cross-scale framework of project 1.7, for both real time forecast and modeling scenarios analysis, and the uncertainty framework of project 1.3. Demonstration of the outcome will be performed in the very complex case of the Albufeira city in the south of Portugal, where combined effects have led to considerable economic losses in recent events. The specific goals of the work are:

- To create a new hybrid model forecasting application to the urban and coastal water domains, based on the numerical model applications of ANI’s sinergea and several sources of in-situ and remote data;
- To establish a suite of scenarios for environmental conditions and anthropogenic operation of the infrastructures and its expected changes due to climate change, accounting for uncertainty, and build a fast-response model for daily management support;
- To build an optimization engine that uses both the database from scenarios and the daily predictions to propose the optimal infrastructure operation under everyday environmental and infrastructure conditions;
- To implement an early warning system based on the previous tools.

6.6.3 Methodology and activity planning

The activities to be developed in this Ph.D. dissertation are:

Activity 1 - Requirements analysis for multi-objective coastal city management system

Task 1: State-of-the-art review of the numerical models suites openly available and definition of the possible data sources to be used for the hybrid modeling

Task 2: Review of the environmental and CC scenarios to be simulated for a coastal city, using Albufeira as a example of an Atlantic ocean bordering system

Task 3: Review and establishment of the suite of indicators to be used in the system performance assessment and optimizing engine

Task 4: Creation of the optimization algorithm

Task 5: Creation of an early-warning system for the multi-objective goal

Activity 2 - Development of a fast-response scenario database

Task 1: Setting up of virtual research environment for simulation of scenarios and its application

Task 2: Development of AI models based on process-based model results

Task 3: Development of fast search algorithm to support fast predictions using scenario database

Task 4: Testing in past events of the case study

Activity 3 - Development and testing of the optimization operation engine

Task 1: Development of optimization engine

Task 2: Testing under several environmental conditions for the case study

Activity 4 - Implementation of the hybrid forecast system with uncertainty

Task 1: Implementation of the hybrid models for the case study

Task 2: Implementation of the cross-scale forecast systems using those hybrid models

Task 3: Integration of the uncertainty engine fully adapted for this cross-scale modeling

Task 4: Testing under past events conditions

Activity 5 - Implementation of the early-warning system and optimization conditions proposal

Task 1: Implementation of the early warning system for inundation and beach contamination based on the tools from activities 2 and 4

Task 2: Implementation of the optimized proposal of operation for real time operation

Task 3: Development of dedicated dashboard in OPENCoastS for the early warning and proposal of operation results

Task 4: Test and validation of the whole system under past events conditions

Activity 6 - Scientific publications and dissemination

Task 1: Conference articles preparation and public presentation (4, at the end of each year)

Task 2: Journal papers publication (3, at the end of years 2,3,4)

Task 3: Dissertation preparation

Table 6.7 summarizes the timeline associated with this work.

Table 6.7 - Timeline for Ph.D. proposal 5

Year	Year 1				Year 2				Year 3				Year 4			
Activity	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**
1 - Requirements analysis for multi-objective system																
2 - Development of a fast-response scenario database																
3 - Development and testing of the optimization operation engine																
4 - Implementation of the hybrid forecast system with uncertainty																
5 - Implementation of the early-warning system and optimization conditions proposal																
6 - Scientific publications and dissemination																

6.6.4 Financial plan

The list below summarizes the expected costs associated with this work.

a) Personnel costs
   - 48 months of Ph.D. grant
   - 4 months of a senior researcher
   - Total: 77,742.72 euros
b) Travel: 4 participation in conferences (2 national, 2 Europe) - 8500 euros
c) Equipment:
   - laptop - 1500 euros

- e-infrastructure resources - no costs (assuming that an application to free resources is done in the Rede Nacional de Computação Avançada calls)

d) Grand total (with overheads): 114 342,72 euros

6.7 Ph.D. 6: Coastal Digital Twin Service

6.7.1 Rationale

The current expectations of CDT's usefulness, completeness and flexibility are very high as the conceptual vision targets a world "controlled at the end of our fingertips", with information harnessed to produce knowledge and optimized decisions automatically. These expectations are framed by the ever-growing need from researchers to create knowledge from the huge volume of data freely available, that can no longer be analyzed by humans without automatic computational procedures, and the fast pace required to management actions that need to respond on frequently unreasonable time frames to the evolution of a constantly changing coast due to climate change or anthropogenic actions, dealing with growing legislative restrictions and very high needs to preserve ecosystems and food chains for the services that provide to mankind.

This proposal, framed in the context of computer science applied to maritime hydraulics, aims at building a coastal digital twin service from the ground, taking advantage of Research Study 1 outcomes, in particular of the hybrid forecast service. Research will entail building the foundations for the service, through a generic framework that is prepared to integrate a high level of interaction between tools, data and models, and the construction of the CDT's service for user interaction, built to accommodate multi-users and an on-demand vision.

This proposal aims at contributing to projects 2.1 "Generic framework for Coastal Digital Twin" and 2.6 "On-demand Coastal Digital Twin Service" of Research Study 2 "CODIT - Intelligent, high-resolution, user-centered and inclusive coastal digital twins".

6.7.2 Goals

This Ph.D. proposal will advance the state-of-the-art by addressing two main challenges:

- construction of a solid backbone to accommodate any new developments and requirements in the future: the development of a generic framework that will integrate the current vision for the requirements and is prepared to integrate a high level of interaction between tools, data and models and provide scientifically-based support to collaborative management of coastal zones
- construction of an on-demand CDT portal that addresses usability, adequate content and friendliness in a innovative way, for multiple types of users

6.7.3 Methodology and activity planning

The activities to be developed in this Ph.D. dissertation are:

Activity 1 - CDT requirements analysis

Task 1: Identification of relevant functions for a CDT and of the core features for implementation of an on-demand CDT portal

Task 2: Identification of the relevant types of tools for implementation and identification of its characteristics and dependencies

Task 3: Identification of the relevant data formats and data dependencies for creation of a data /information model and its implementation as a data lake

Task 4: Identification of the user interaction modes and development of an usability strategy to accommodate them in an platform dashboard

Activity 2 - Data/information model and lake construction

Task 1: Implementation of a data/information repository following the requirements analysis

Task 2: Validation in multiple types of data structures for several selected use cases

Activity 3 - CDT workflow builder

Task 1: Definition of the generic building blocks in the CDT and its tailoring for predefined typologies of problems

Task 2: Adaptation/Extension of WIFF´s and OPENCoastS´ building blocks for CDT

Task 3: Development of a workflow builder according to specifications for specific applications and for a generic, customizable problem

Task 4: Testing of the procedure in selected use cases

Activity 4 - Development of CDT-as-a-service platform and backend

Task 1: Development of the frontend application according to requirements

Task 2: Integration of CDT on-demand hybrid forecast in the backend

Task 3: Integration of CDT on-demand data assimilation engine

Task 4: Integration of reliability engine

Task 5: Integration of on-demand processing tools and indicators engine

Activity 5 - Application of on-demand CDT to a thematic digital twins portal creation

Task 1: Identification of a theme for implementation in cooperation with end-users/coastal researchers

Task 2: Creation of simple thematic digital twin

Task 3: Validation in selected coastal sites

Activity 6 - Scientific publications and dissemination

Task 1: Conference articles preparation and public presentation (4, at the end of each year)

Task 2: Journal papers publication (3, at the end of years 2,3,4)

Task 3: Dissertation preparation

Table 6.8 summarizes the timeline associated with this work.

Table 6.8 - Timeline for Ph.D. proposal 6

Year	Year 1				Year 2				Year 3				Year 4			
Activity	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**
1 - CDT requirements analysis																
2 - Data/information model and lake construction																
3 - CDT workflow builder																
4 - Development of CDT-as-a-service platform and backend																
5 - Implementation of the early-warning system and optimization conditions proposal																
6 - Scientific publications and dissemination																

6.7.4 Financial plan

The list below summarizes the expected costs associated with this work.

a) Personnel costs
   - 48 months of Ph.D. grant
   - 6 months of a senior researcher
   - Total: 89 142,72 euros
b) Travel: 4 participation in conferences (2 national, 2 Europe) - 8500 euros
c) Equipment:
   - laptop - 1500 euros
   - e-infrastructure resources - no costs (assuming that an application to free resources is done in the Rede Nacional de Computação Avançada calls)
d) Grand total (with overheads): 137 142,72 euros

6.8 Ph.D. 7: Optimized information assimilation with uncertainty consideration

6.8.1 Rationale

Data assimilation is a fundamental component in the construction of digital replica of water systems, linking data and models in a seamless and quality-driven way. The recent availability of multiple data sources at adequate spatial and temporal scales has promoted the development of several strategies for assimilation of each of these sources data in numerical modeling simulations. In particular, remote sensing solutions have become major data sources for many coastal and estuarine data acquisition systems, given the technological advances of remote sensing devices, such as HFR radar, cameras or sensors mounted in drones and satellite remote sensing. These different solutions provide data for different variables and devices can be applied independently or in an integrated fashion, to promote knowledge, validate forecast simulations or to be part of assimilation procedures. The main concept of data assimilation modeling is to produce estimates of the relevant variables that are consistent with data and model dynamics, allowing for errors in both. Data are assimilated to produce a revised prediction through optimal combination of information from the model and observations. The analysis is then used as the initial condition for the next prediction cycle, and the process repeats.

Many of the data providers are now following a policy of open access to most of the data sources (e.g. open availability of medium resolution (~10 m) satellite data from the Sentinel constellation), thus facilitating the creation of added value service to improve the quality of digital representations of real systems such as forecast systems or CDT.

The complexity of the non-linear processes in coastal regions is in general a limiting factor to the application of simple assimilation techniques, in particular for those requiring an estimator of the errors. Using a combination of many data sources, to build dense, reliable monitoring networks covering the small time and spatial scales, may be an avenue to overcome this problem and promote useful developments. The possibility of using multiple sources of data in the assimilation procedure requires the capacity to ensemble data from multiple sources, with potentially distinct time and space scales, heterogeneous formats and semantics. Furthermore, the choice of the assimilation technique may not be a single one for the same problem as environmental conditions may shift and the optimal procedure may be space and time varying. Exploring this variability into an optimized assimilation tool has not been explored yet and will be developed in this dissertation.

This proposal aims at contributing to Project 2.2 “On-demand, optimized data assimilation framework for user-selected multi-source data” of Research Study 2 “CODIT - Intelligent, high-resolution, user-centered and inclusive coastal digital twins”.

6.8.2 Goals

This proposal aims at developing an optimized, adaptive methodology to assimilate multi-source data in a seamless and quality controlled way, and to integrate it in the OPENCoastS service. The innovative

assimilation strategy should be flexible enough be applied for a multi-scale (from the basin to the sea), multi-process digital twin and to data-based, numerical or hybrid models, in a user-driven, on-demand way. It should also be able to adapt to specific environmental conditions.

The specific goals are:

- to define data sources, error measures and the goals for each assimilation procedure targeting optimization
- select the best assimilation algorithms for estuarine and coastal dynamics and implement them for generic application to multiple model types, accounting for uncertainty in data and models outputs
- integrate the methodologies and its automatic selection according to environmental conditions in OPENCoastS to serve in automatic forecasting, subject to user configuration.

6.8.3 Methodology and activity planning

The activities to be developed in this Ph.D. dissertation are:

Activity 1 - State-of-the-art review and data characterization for assimilation purposes

Task 1: Definition of the data sources and their characterization

Task 2: Definition of the errors measures to be used in data assimilation assessment

Task 3: Definition of the goals for data assimilation

Task 4: State/of/the/art review data assimilation strategies

Activity 2 - Selection and implementation of data assimilation methodology

Task 1: Selection of most relevant algorithms and their implementation

Task 2: Comparison of algorithms in simple use cases and classification of their adequacy depending on dominant processes

Task 2: Development of automatic selector for proposal to users

Activity 3 - Development of assimilation on-demand capacity and integration in OPENCoastS

Task 1: Identification of the user-selected parameters for dashboard creation

Task 2: Adaptation of OPENCoastS frontend to accommodate data assimilation

Task 3: Linkage between data-sources interface and data assimilation through user selection of sources and algorithm choice

Task 4: Integration of data assimilation in the hybrid forecast workflow of OPENCoastS

Task 5: Setting up of virtual research environment for data assimilated simulations in shared e-infrastructures

Activity 4 - Scientific publications and dissemination

Task 1: Conference articles preparation and public presentation (4, at the end of each year)

Task 2: Journal papers publication (3, at the end of years 2,3,4)

Task 3: Dissertation preparation

Table 6.9 summarizes the timeline associated with this work.

Table 6.9 - Timeline for Ph.D. proposal 7

Year	Year 1				Year 2				Year 3				Year 4			
Activity	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**
1 - State-of-the-art review and data characterization																
2 - Selection and implementation of data assimilation methodology																
3 - Development of assimilation on-demand capacity and integration in OPENCoastS																
4 - Scientific publications and dissemination																

6.8.4 Financial plan

The list below summarizes the expected costs associated with this work.

a) Personnel costs
   - 48 months of Ph.D. grant
   - 4 months of a senior researcher
   - Total: 77,742.72 euros
b) Travel: 4 participation in conferences (2 national, 2 Europe) - 8500 euros
c) Equipment:
   - laptop - 1500 euros
   - e-infrastructure resources - no costs (assuming that an application to free resources is done in the Rede Nacional de Computação Avançada calls)
d) Open access publication fees: 3000 euros (fee waivers will be promoted, one publication is expected to be in a journal without waivers)
e) Grand total (with overheads): 114 342,72 euros

6.9 Ph.D. 8: Digital twin tool for multi-hazard management and application to short and long term inundation

6.9.1 Rationale

Emergency situations in the coastal zone are often a combination of the action of multiple hazards. Unfortunately, most of the existing early-warning systems do not account for the several sources of danger or concentrate the most accurate tools for one particular hazard, addressing the others in simplistic or correlated ways. For instance, the combined action of waves, surges and tides can cause flooding, erosion and dune and structure overtopping in many coastal regions (Rocha et al., 2021). Accurate inundation forecasts are crucial to predict in a timely way extreme sea levels that flood lowland coastal and estuarine areas, with risk for people and goods. Often these high water levels occur simultaneously with large precipitation events that further extend the simultaneous risk of flooding to coastal cities as drainage networks are often insufficient for fast water drainage in presence of high water levels at receiving water bodies. Thus, to obtain accurate time of arrival of coastal and estuarine flood and flooding areas, it is essential to predict the hydrodynamics from the ocean to the land scales. Many forecast systems exist but the vast majority of these systems account for hydrodynamics only and do not integrate real time data at the adequate space and time scales. Moreover, short term and long term inundation management is performed in a decoupled way, often with tools of distinct precision.

A multi-hazard coastal management system that addresses all relevant water compartments at the adequate spatial and temporal scales, applicable to short term (emergency) and long term (planning) time frames has not been developed yet at the accuracy necessary to support alerts and collaborative decision making. Besides models, data is a key component of this infrastructure through a combination of targeted campaigns and real-time data that measure all phenomena at stake and can be used to develop comprehensive monitoring platforms. These monitoring platforms can support the development of prediction tools that address all hazards in an integrated way. The continuous usage of the prediction tools provides information to revisit the location of the monitoring stations, adapting it to the ever-changing environment at the coast, providing a two-way interaction between data and models.

This proposal aims at contributing to Projects 2.7 “Early warning, multiple information source-based platform for multi-hazards” and 2.8 “Digital twin for short and long term inundation management” of Research Study 2 “CODIT - Intelligent, high-resolution, user-centered and inclusive coastal digital twins”.

6.9.2 Goals

This Ph.D. proposal will build a new, generic methodology focused on multi-hazard coastal management and its implementation in the on-demand CDT framework and platform devoted to short and long term inundation management. It will also create and implement an innovative surveillance system that allows users to configure their digital twin for inundation integrating a new monitoring system, hybrid forecasting and scenario analysis to seamlessly address inundation for multiple users in any coastal system accounting for the relevant processes.

6.9.3 Methodology and activity planning

The activities to be developed in this Ph.D. dissertation are:

Activity 1 - Requirements for a multi-hazards CDT and its configuration to inundation

Task 1: State-of-the-art review the relevant processes for multi-hazard, including inundation, in coastal regions

Task 1: Definition of the early-warning criteria for multiple hazards

Task 2: Definition of the data sources and modeling tools for multiple hazards

Task 3: Definition of the features for user-interaction multiple hazards dashboard

Task 4: Identification of the scenarios for the database creation for long term analysis

Activity 2 - Combining information assets for intelligent multi-hazard CDT

Task 1: Implement a multi-hazard CDT using the on-demand portal and configure it for short and long term inundation

Task 2: Establishing on-demand forecast systems for multi-hazard typology

Task 3: Establishment of data sources acquisition to optimize operation

Task 4: Define a two-way workflow of information integration and reliability assessment

Task 5: Validate in distinct multi-hazard configurations and test sites

Activity 3 - Short term inundation platform based on intelligent local data processing and hybrid inundation forecast services

Task 1: Application of the multiple hazard dashboard for user-specified process and tool selection (numerical model, data sources, inundation thresholds for multiple types of coastal observatories)

Task 2: Operation of the hybrid circulation and morphodynamic forecasting platform from Research Study 1

Task 3: Application of the early warning service in coastal digital twin inundation infrastructure

Task 5: Testing and validation at data rich use cases

Activity 4 - Long term, adaptive inundation platform based on scenarios and short term outputs

Task 1: Establishment of long-term scenario database using the dedicated dashboard

Task 2: Creation of inundation hazard CDT for long term management purposes, accounting for adaptive update using short term results and applying dedicated indicators at user-selected hotspots

Task 3: Demonstration at selected use cases under dynamic conditions

Activity 5 - Scientific publications and dissemination

Task 1: Conference articles preparation and public presentation (4, at the end of each year)

Task 2: Journal papers publication (3, at the end of years 2,3,4)

Task 3: Dissertation preparation

Table 6.10 summarizes the timeline associated with this work.

Table 6.10 - Timeline for Ph.D. proposal 8

Year	Year 1				Year 2				Year 3				Year 4			
Activity	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**
1 - Requirements for a multi-hazards CDT and its configuration to inundation																
2 - Combining information assets for intelligent multi-hazard CDT																
3 - Short term inundation platform based on intelligent local data processing and hybrid inundation forecast services																
4 - Long term, adaptive inundation platform based on scenarios and short term outputs																
5 - Scientific publications and dissemination																

6.9.4 Financial plan

The list below summarizes the expected costs associated with this work.

a) Personnel costs
   - 48 months of Ph.D. grant
   - 4 months of a senior researcher
   - Total: 77,742.72 euros
b) Travel: 4 participation in conferences (2 national, 2 Europe) - 8500 euros
c) Equipment:
   - laptop - 1500 euros
   - e-infrastructure resources - no costs (assuming that an application to free resources is done in the Rede Nacional de Computação Avançada calls)

d) Grand total (with overheads): 114 342,72 euros

6.10 Ph.D. 9: Building collaboratories for coastal management using Digital Twins and demonstration for the Tagus estuary

6.10.1 Rationale

Managing coastal regions subject to multiple and often conflicting uses is an unresolved challenge for many regions throughout the world. Collaborative initiatives have been promoted for several years, developed with particular systems in mind, but a generic infrastructure for collaboration that provides access to high accuracy assets from data and models, in real time and for long term planning, and that allows users to experiment in a virtual way several solutions or anticipate the impact of climate change is not available yet.

This Ph.D. dissertation aims at taking advantage of on-demand CDTs and all the tools proposed in the previous plans and projects to create a collaboratory-concept tool that provides the backbone for the creation of multiple thematic, collaborative digital twins that can address the several areas of concern in coastal regions and implement it as a collaborative CDT. A demonstration in the Tagus estuary, the most complex and probably the most conflicting estuarine and coastal region in Portugal, will at the same time validate the tool and advance the state-of-the-art in CDT use for coastal regions.

The work developed in this dissertation aims to contribute and provide tools to Project 2.10 - Tagus river-to-estuary collaboratory for thematic digital twins and collaborative management of Research Study 2 “CODIT - Intelligent, high-resolution, user-centered and inclusive coastal digital twins”.

6.10.2 Goals

This dissertation will create a collaborative CDT concept and implement it using the on-demand CDT, followed by a knowledge-creating demo for the Tagus estuary. A user-centered, participatory view is behind all work, providing for the first time an infrastructure for knowledge creation and sharing, mixing research and management towards collaborative actions.

6.10.3 Methodology and activity planning

The activities to be developed in this Ph.D. dissertation are:

Activity 1 - Requirements and core application for collaborative CDTs

Task 1: End-user centered requirements for collaborative use of the estuary, for multiple uses including inundation and water quality protection, water shortage and exceedance avoidance,

Task 2: Establishment of the thematic CDTs for each concern and development of corresponding core CDTS

Task 3: Establishment of the hot spot products and alerts

Activity 2 - Collaborative management digital platform

Task 1: Integration of multiple thematic twins and implementation of hot spot products and alerts

Task 2: Integration of simulation for on-demand “What-if” scenarios simulation and automatic assessment over the thematic twins

Task 3: Usability analysis over the final platform

Activity 3 - Tagus estuary demo: Hybrid forecast system for physical to biochemistry processes using on-demand digital twin

Task 1: Development of a data-based forecast system based on in-situ and satellite and camera remote sensing data integration for multiple processes (water samples, sensors, drones, ship-of-opportunity, cameras, satellite, ...)

Task 2: Implementation of urban discharge models in UBEST numerical forecasts

Task 3: Construction of hybrid forecast systems through integration of data-based forecasts and enhanced UBEST numerical model forecasts

Task 4: Implementation simple on-demand water quality simulator for multiple sources of contamination

Task 5: Customization of twin platform and outputs in CDT for the several thematic twins support

Activity 4 - Tagus estuary demo: Urban fingerprint digital twin

Task 2: Anthropogenic and climate change scenario database building using hybrid forecast system

Task 3: Development of tools and indicators for hands-on management simulator for the Metropolitan Lisbon area

Activity 5 - Tagus estuary demo: River-to-estuary inundation and salinity management digital twin

Task 1: End-user centered requirements for collaborative use of the river to estuary dimensions in water exceedance and shortage conditions

Task 2: Integration of agricultural water pumping models in hybrid forecast

Task 3: Integration of data-based AI dams model

Task 4: Implementation of early warning systems for exceedance of salinity at water intake locations

Task 5: Implementation of early warning systems for local inundation events

Activity 6 - Tagus estuary demo: Collaborative management digital platform

Task 1: Integration of multiple thematic twins and creation of targeted hot spot products and alerts for the Tagus

Task 2: Exploitation of the collaborative CDT under different user roles and stress conditions towards achieving harmonized decisions that minimize impacts.

Activity 7 - Scientific publications and dissemination

Task 1: Conference articles preparation and public presentation (4, at the end of each year)

Task 2: Journal papers publication (3, at the end of years 2,3,4)

Task 3: Dissertation preparation

Table 6.11 summarizes the timeline associated with this work.

Table 6.11 - Timeline for Ph.D. proposal 9

Year	Year 1				Year 2				Year 3				Year 4			
Activity	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**
1 - Requirements and core application for collaborative CDTs																
2 - Collaborative management digital platform																
3 - Tagus estuary demo: Hybrid forecast system for physical to biochemistry processes using on-demand CDT																
4 - Tagus estuary demo: Urban fingerprint digital twin																
5 - Tagus estuary demo: River-to-estuary inundation and salinity management digital twin																
6 - Tagus estuary demo: Collaborative management digital platform																
7 - Scientific publications and dissemination																

6.10.4 Financial plan

The list below summarizes the expected costs associated with this work.

e) Personnel costs
   - 48 months of Ph.D. grant
   - 6 months of a senior researcher
   - Total: 89 142,72 euros
f) Travel: 4 participation in conferences (2 national, 2 Europe) - 8500 euros
g) Equipment:
   - laptop - 1500 euros
   - e-infrastructure resources - no costs (assuming that an application to free resources is done in the Rede Nacional de Computação Avançada calls)
h) Grand total: 137 142,72 euros

6.11 M.Sc. 1: Automatic grid generator for forecast applications

6.11.1 Rationale

The capacity of unstructured grid modeling to handle complex bathymetries and geometries is making it the primary choice for high resolution simulations in coastal zones, supported by the exponential growth in computational resources necessary for its application. Many open source models under this discretization paradigm are available (ADCIRC, SELFE, SCHISM; DELF3D, …) and have been used for many purposes from inundation to biogeochemical simulations.

Creating an adequate horizontal computational grid remains however as the major limitation to the application of these models. Non-experienced users have pointed out, in OPENCoastS training events, that the difficulty to build a grid prevents them from using services that require the availability of this asset. Even for experienced modelers, the construction of a good quality grid that can provide high accuracy results is a very time consuming task.

To address this issue, many grid generation packages have been developed in recent years, facilitating the grid generation problem but still requiring considerable knowledge on unstructured numerical modeling and grid generation procedure. This approach is not however a solution for non-expert users.

This proposal aims at developing an automatic grid generator where users do not need numerical modeling expertise to build an adequate grid. This automatic grid generator will be provided as a service, allowing users to interact with a friendly web interface built to facilitate the grid construction procedure, and will encapsulate expert knowledge for different type of applications (e.g. 3D baroclinic, wave and current interaction, ...). This work will advance the state-of-the-art in mechanisms for automatic grid generation for unstructured modeling in coastal regions and will be provided as part of the OPENCoastS service, thus facilitating the execution of the whole workflow for on-demand forecasting.

This proposal aims at contributing to Project 1.2 - Computational grid generation as-a-service, of Research Study 1 “OPENCoastS4ALL - Reliable, quality-controlled, multi-process, on-demand coastal forecast framework for oceans to hydrographic basin application”.

6.11.2 Goals

This project aims at developing a grid generation tool that can be used for 2D coastal modeling grid creation. It should be based on the following characteristics, available through user interaction with a simple-to-use platform:

- simulation model
- type of simulation (circulation/wave and current/water quality)
- region of interest to be simulated (the area where results are looked for, not necessarily the grid domain)
- identification of hotspots that will require higher grid resolution (e.g discharge points)

As this work is proposed in the scope of a M.Sc. thesis, global bathymetric databases will be used. The work will take advantage of the core grid generator developed by NOAA in the OCSMesh code.

This work is framed in the development of OPENCoastS core capacities in the H2020 EGI-ACE project. It contributes towards the development of project 1.2 “Computational grid generation as-a-service” of Research Study 1 “OPENCoastS4All - Reliable, cross-scale, multi-process, on-demand coastal forecast framework for oceans to hydrographic basin application”.

6.11.3 Methodology and activity planning

The activities to be developed in this M.Sc. thesis are:

Activity 1 - Requirements analysis for unstructured grid generation

Task 1: Requirements analysis for multiple type of simulations

Task 2: Specification of grid quality indicators as a function of application type

Activity 2 - Adaptation of automatic grid generator and validation

Task 1: Review of OCSMesh grid generator

Task 2: Extension/adaptation of the selected tool for Activity 1’s requirements analysis

Task 3: Validation in selected applications against existing expert modelers grids

Activity 3 - Integration in OPENCoastS framework

Task 1: Requirements analysis for OPENCoastS integration

Task 2: Integration of the open source bathymetry and coastal geometry databases

Task 3: Integration of the selected grid generation engine in OPENCoastS and development of the user-selection grid-generation dashboard

Task 6: Setting up of virtual research environment for grid generation in shared e-infrastructures

Task 7: Testing for simple cases of hydrodynamic simulations

Activity 4 - Scientific publications and dissemination

Task 1: Conference article preparation and public presentation (at the end of the year)

Task 2: Thesis preparation

Table 6.12 summarizes the timeline associated with this work.

Table 6.12 - Timeline for M.Sc. proposal 1

Year/Month	1 Year											
Activity	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
1 - Requirements analysis												
2 - Adaptation of automatic grid generator and validation												
3 - Integration in OPENCoastS												
4 - Scientific publications and dissemination												

6.11.4 Financial plan

The list below summarizes the expected costs associated with this work.

i) Personnel costs
   - 12 months of M.Sc. grant
   - 1,5 months of a senior researcher
   - Total: 19 061.76 euros

j) Travel: 1 participation in european conference - 3000 euros

k) Equipment:
   - laptop - 1500 euros
   - e-infrastructure resources - no costs (assuming that an application to free resources is done in the Rede Nacional de Computação Avançada calls)

l) Grand total: 31 211,76 euros

6.12 M.Sc. 2: Morphodynamic updater in forecast platforms

6.12.1 Rationale

Estuarine and coastal systems are frequently affected by bathymetric changes that can alter their hydrodynamic and water quality characteristics. Furthermore, the bathymetric evolution can alter

radically the nature of coastal systems, limiting for instance the capacity to exchange water between the coast and coastal lagoons, due to the migration and closure of lagoon inlets. The accuracy of the long term operation of forecast systems is thus dependent on the continuous maintenance of the bathymetric status.

Automatic ways to update the topography and bathymetry in coastal and estuarine forecasts are the way to proceed and several efforts are starting to be made through integration of remote sensing data from satellites and camaras for bathymetric update in hydrodynamic models, through the inclusion of morphodynamic models in the forecast workflows and by combining both approaches through data assimilations.

This proposal aims at exploring these recent tools and combining them, validating the procedure in at least one use case. As morphodynamic simulations often suffer from stability or divergence from reality issues, the new methodology will start by identifying the core bathymetric features that need to be controlled in the automatic bathymetry update procedure.

This proposal aims at contributing to Project 1.4 ”Morphodynamic processes integration in forecast simulations. Methodology and tools”, of Research Study 1 “OPENCoastS4ALL - Reliable, quality-controlled, multi-process, on-demand coastal forecast framework for oceans to hydrographic basin application”.

6.12.2 Goals

This proposal aims at integrating morphodynamic prediction and bathymetry update in the OPENCoastS forecast framework, taking advantage of the morphodynamic module of the SCHISM model and data sources available.

The specific goals of the project are:

- to select a suite of SCHISM morphodynamic model configurations and to build model ensembles;
- to detect the most important features controlling morphodynamic evolution and to define a methodology for its proper representation;
- to integrate the ensemble morphodynamic engine in OPENCoastS, as one of the workflow components available for selection by the user.

6.12.3 Methodology and activity planning

The activities to be developed in this M.Sc. thesis are:

Activity 1 - Core components analysis for morphodynamic forecasting

Task 1: State-of-the-art-review of open access morphodynamic models, their range of applicability and selection of the components from SCHISM for the forecast engine

Task 2: Identification of the main bathymetric features to be detected, monitored and accurately predicted in morphodynamic predictions

Task 3: Definition of the methodology for bathymetry features data integration from selected sources of remote (cameras, satellite)

Task 4: Definition of the methodology for morphodynamic ensemble building and error measure selection for accuracy evaluation

Activity 2 - Development of morphodynamic engine

Task 1: Development of the ensemble predictor based on any combination of models

Task 2: Testing of morphodynamic engine in one use case

Activity 3 - Development of assimilation strategy for bathymetry update

Task 1: Review of bathymetric assimilation procedures and selection of a simple one for the selected remote data sources

Task 2: Implementation of the assimilation procedure

Task 3: Testing of assimilation strategy in one use case

Activity 4 - Integration of morphodynamic forecasting in OPENCoastS

Task 1: Requirements analysis for OPENCoastS integration

Task 2: Integration of assimilation tool for bathymetry update

Task 3: Integration of morphodynamic engine in OPENCoastS

Task 4: Setting up of virtual research environment for morphodynamics simulation and assimilation in shared e-infrastructures

Activity 5 - Scientific publications and dissemination

Task 1: Conference article preparation and public presentation (at the end of the year)

Task 2: Thesis preparation

Table 6.13 summarizes the timeline associated with this work.

Table 6.13 - Timeline for M.Sc. proposal 2

Year/Month	**1 Year**											
Activity	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
1 - Core components analysis for morphodynamic forecasting												
2 - Development of morphodynamic engine												

Year/Month	1 Year											
Activity	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
3 - Development of assimilation strategy												
4 - Integration in OPENCoastS												
5 - Scientific publications and dissemination												

6.12.4 Financial plan

The list below summarizes the expected costs associated with this work.

a) Personnel costs
   - 12 months of M.Sc. grant
   - 1,5 months of a senior researcher
   - Total: 19 061.76 euros
b) Travel: 1 participation in european conference - 3000 euros
c) Equipment:
   - laptop - 1500 euros
   - e-infrastructure resources - no costs (assuming that an application to free resources is done in the Rede Nacional de Computação Avançada calls)
d) Grand total: 31 211,76 euros

6.13 M.Sc. 3: Operational fast assessment pollution tool

6.13.1 Rationale

Particle models remain a valuable asset in estuarine and coastal studies given their flexibility, low computational costs and ease of operation, making them an attractive choice for large number scenario and uncertainty calculations or a multiple contamination source setting. The development of a particle forecast engine, built on top of a flexible circulation forecast, that could handle multiple types of simulations and allow for customization on-the-fly by the users, does not exist yet and can be a valuable asset for many coastal management actions and for scenario-based analysis. Integrating uncertainty in forcing conditions in the simulations as well as adequate decay formulations may also increase the usability and accuracy of the simple particle models.

While particle models are available in forecast systems, a configurable engine that could simulate multiple types of contamination such as oil spills, larvae movement, marine debris pathways, water discharges and other processes selected or configured by the user is not yet available. Moreover, linking

these models with remote sensing data sources that could provide adequate initial conditions and accounting for uncertainty are not available either in particle forecasting. Such a flexible and user-centered approach can be a valuable asset for coastal managers and researchers and extends the current state-of-the-art.

This thesis aims at developing such a tool, framed in project “On-demand particle evolution and decay simulator for multiple applications”, of Research Study 1 “OPENCoastS4ALL - Reliable, quality-controlled, multi-process, on-demand coastal forecast framework for oceans to hydrographic basin application”.

6.13.2 Goals

This work aims at developing a user configurable particle forecast engine, applicable for multiple water quality and biogeochemical problems, forced by the relevant circulation conditions, and its integration in the OPENCoastS framework. Specific goals are:

- to select a suite of pre-defined relevant transport and transformation problems, such as oil spills, marine debris pathways, algae blooms, larvae movement, and to select the most accurate formulations for their pathways and evolution;
- to define adequate sources for initial conditions for these problems, based on remote sensing data and development of the initial condition forecast engine;
- to define the relevant characteristics in generic transport and transformation problems and to develop a user configurable particle forecast engine;
- to integrate the above tools in OPENCoastS, taking advantage of the multiple circulation options, and create a particle transport and evolution forecast dashboard.

6.13.3 Methodology and activity planning

The activities to be developed in this M.Sc. thesis are:

Activity 1 - Requirement analysis for multi-type particle forecast engine

Task 1: Definition of the predefined application types and selection of the corresponding most accurate formulations

Task 2: Definition of the parameters for user configuration for integration in the dedicated dashboard in OPENCoastS

Task 3: Definition of data sources for initial conditions establishment

Activity 2 - 3D particle model engine with multi-mode flexible problem options and predefined application types

Task 1: Development/adaptation of generic particle forecast engine with user-specified configuration or predefined typology

Task 2: Development of pre-defined typology forecast engine

Task 3: Integration of uncertainty option through multi-scenario or multi-model approaches

Task 3: Testing for multiple typology cases with available field data

Activity 3 - Initial conditions provider engine development

Task 1: Development of initial conditions engine for selected pre-defined transport typologies (oil spills, algae blooms, …)

Task 2: Integration in particle forecast engine

Task 3: Testing in selected use cases with available field data

Activity 4 - Integration in OPENCoastS framework

Task 1: Requirements analysis for OPENCoastS integration

Task 2: 3D particle model engine integration with multi-mode flexible problem configuration and uncertainty tool integration (linked to project 1.3)

Task 3: Development of pre-defined engine for oil spill, marine litter and other to be defined pre-defined typology dynamics in operational or scenario mode

Task 4: Development of user-selection dashboard for specifying scalar problem choice and configuration of decay and evolution properties

Task 7: Setting up of virtual research environment for multi-mode particle simulations in shared e-infrastructures

Activity 5 - Scientific publications and dissemination

Task 1: Conference article preparation and public presentation (at the end of the year)

Task 2: Thesis preparation

Table 6.14 summarizes the timeline associated with this work.

Table 6.14 - Timeline for M.Sc. proposal 3

Year/Month	1 Year											
Activity	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
1 - Requirements analysis for multi-type particle forecast												
2 - 2D and 3D particle model engine												
3 - Initial conditions provider engine												

Year/Month	1 Year											
Activity	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
3 - Integration in OPENCoastS												
4 - Scientific publications and dissemination												

6.13.4 Financial plan

The list below summarizes the expected costs associated with this work.

a) Personnel costs
   - 12 months of M.Sc. grant
   - 1,5 months of a senior researcher
   - Total: 19 061.76 euros
b) Travel: 1 participation in european conference - 3000 euros
c) Equipment:
   - laptop - 1500 euros
   - e-infrastructure resources - no costs (assuming that an application to free resources is done in the Rede Nacional de Computação Avançada calls)
d) Grand total: 31 211,76 euros

6.14 M.Sc. 4: Information quality assessment for coastal applications

6.14.1 Rationale

The current vision of CDT entails the intelligent and automated exploitation of the multiple sources of data and forecast results currently available at fine spatial and temporal scales. The volume of information at hand far outweighs the capacity of humans to process it and produce intelligent knowledge and decision taking. The usage of automatic procedures becomes thus very dependentent of the quality of the information being used and pose a new challenge to data reliability and quality which become fundamental to support in a robust way the usage of data-rich CDTs or even simple management decisions based on this data.

The work of Jesus (2019), based on in-situ sensors and a single type of simulation, is proposed in this M.Sc. work to be extended to the information generated by multi-source data and hybrid models, for future integration in CDTs.

This proposal aims at contributing to Project 2.4 “Reliability infrastructure for two-way data and model information quality using data fusion and failure compensation” of Research Study 2 “CODIT - Intelligent, high-resolution, user-centered and inclusive coastal digital twins”.

6.14.2 Goals

The goal of this M.Sc. proposal is to create a first reliability infrastructure for future application in the CDT, starting from the outcomes of the work of Jesus (2019), applied to the several data and models sources currently available in OPENCoastS (in-situ and satellite remote sensing data and several modeling options).

The specific goals are:

- to create a reliability workflow for two-way data/model quality evaluation
- to label information to support quality indexing of data and of the services built on top of it
- to develop a dedicated dashboard in OPENCoastS platform for information quality assessment

6.14.3 Methodology and activity planning

The activities to be developed in this M.Sc. thesis are:

Activity 1 - Requirements analysis for reliable data lakes

Task 1: Review of existing reliability procedures

Task 2: Definition of the reliability and data quality assessment procedures using data and model in a composed way

Activity 2 - Reliability and information classifier engine for multi-source information ingestion in data lakes

Task 1: Comparison of fault evaluation mechanisms for data and model information, for multiple types of simulations and data sources, using model/data composed analysis

Task 2: Selection of procedure for fault detection and compensation and for quality assessment

Task 3: Development of engine for reliability and quality classification evaluation

Task 4: Tests in data-rich use cases

Activity 3 - Implementation in OPENCoastS framework

Task 1: Requirements analysis for OPENCoastS integration

Task 2: Integration of reliability and information classifier and development of dedicated dashboard

Task 3: Development of user-selection dashboard for specifying reliability formulation and choice of information sources

Task 4: Setting up a virtual research environment for reliability analysis in shared e-infrastructures…

Activity 4 - Scientific publications and dissemination

Task 1: Conference article preparation and public presentation (at the end of the year)

Task 2: Thesis preparation

Table 6.15 summarizes the timeline associated with this work.

Table 6.15 - Timeline for M.Sc. proposal 4

Year/Month	1 Year											
Activity	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
1 - Requirements analysis for reliable data lakes												
2 - Reliability and information classifier engine												
3 - Integration in OPENCoastS												
4 - Scientific publications and dissemination												

6.14.4 Financial plan

The list below summarizes the expected costs associated with this work.

e) Personnel costs
   - 12 months of M.Sc. grant
   - 1,5 months of a senior researcher
   - Total: 19 061.76 euros

f) Travel: 1 participation in european conference - 3000 euros

g) Equipment:
   - laptop - 1500 euros
   - e-infrastructure resources - no costs (assuming that an application to free resources is done in the Rede Nacional de Computação Avançada calls)

h) Grand total: 31 211,76 euros

6.15 M.Sc. 5: Multi-dimensional, automatic and on-the-fly data analysis in a Coastal Digital Twin

6.15.1 Rationale

The volume of information currently being generated and expected to be created in the scope of CDT, including forecast systems and data sources outputs, requires automatic processing towards knowledge creation and indicator calculation. This thesis addresses this need at basic output processing and high level indicator production, to support both knowledge creation and management/legislation goals. To facilitate user uptake and take advantage of algorithms already shared in the ocean/coastal communities, popular strategies such as Jupyter notebooks will be used. This thesis also addresses the

integration of these algorithms, promoting their reuse and adaptation to new products and indicators, and their operationalization inside OPENCoastS.

Finally, this thesis also explores the concept of hotspots, to provide easy to understand products for end user exploration. The generation of time series of indicators and of other information at selected hotspots by the users through the application of the notebook algorithms creates new opportunities of semi-automatic knowledge creation.

This proposal aims at contributing to Project 2.5 “On-the-fly, user-selected data exploitation for data and model information” of Research Study 2 “CODIT - Intelligent, high-resolution, user-centered and inclusive coastal digital twins”.

6.15.2 Goals

The aim of this thesis is to build a suite of data exploitation tools and make them available for automatic operation, sharing and customization to specific purposes inside OPENCoastS. These algorithms will take advantage of popular technologies such as notebooks to facilitate user uptaking.

6.15.3 Methodology and activity planning

The activities to be developed in this M.Sc. thesis are:

Activity 1 - Requirements analysis and resources workflow

Task 1: Identification of user requirements for information processing

Task 2: Review of available information processing tools and selection of relevant ones for this work

Task 3: Identification of resources for processing tools operation and establishment of workflow for operation in shared infrastructure

Activity 2 - Creation of dedicated notebooks for algorithm implementation and sharing

Task 1: Development/adaptation of dedicated notebooks for selected algorithms

Task 2: Development of notebooks for addressing the remaining requirements

Task 4: Test in information-rich use cases

Activity 3 - Integration in OPENCoastS

Task 1: development of dedicated dashboard for processing tools/algorithms selection

Task 2: Integration in OPENCoastS framework for operation, sharing and on-the-fly customization

Task 4: Test in information-rich deployments

Activity 4 - Scientific publications and dissemination

Task 1: Conference article preparation and public presentation (at the end of the year)

Task 2: Thesis preparation

Table 6.16 summarizes the timeline associated with this work.

Table 6.16 - Timeline for M.Sc. proposal 5

Year/Month	1 Year											
Activity	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
1 - Requirements analysis and resources workflow												
2 - Creation of dedicated notebooks												
3 - Integration in OPENCoastS												
4 - Scientific publications and dissemination												

6.15.4 Financial plan

The list below summarizes the expected costs associated with this work.

i) Personnel costs
   - 12 months of M.Sc. grant
   - 1,5 months of a senior researcher
   - Total: 19 061.76 euros

j) Travel: 1 participation in european conference - 3000 euros

k) Equipment:
   - laptop - 1500 euros
   - e-infrastructure resources - no costs (assuming that an application to free resources is done in the Rede Nacional de Computação Avançada calls)

l) Grand total: 31 211,76 euros

6.16 M.Sc. 6: Inundation management in Portugal using Digital Twins

6.16.1 Rationale

Climate change and its consequences together with a highly energetic wave regime and frequent storm surges are driving the onset of multiple inundation events in the Portuguese coast whose consequences are further aggravated by the large concentration of population in that area.

The thematic CDT developed in projects 2.7 and 2.8 is explored in this thesis to build an inundation management tool that brings together and explores data and predictions at fine resolution, processed and explored to generate timely and high accuracy information to support both short term emergency and long term planning of this hazard in our coasts. The innovation also lies in the collaborative

conceptualization of the tool that aims at creating knowledge and information to support harmonization of procedures and interventions towards a safer and sustainable coast. The outcome of the thesis is expected to be the prototype of a formal inundation portal to be used and operated for the whole coast.

This proposal aims at contributing to Project 2.9 “Inundation.PT - a digital twin for inundation management at the Portuguese coast and estuaries” of Research Study 2 “CODIT - Intelligent, high-resolution, user-centered and inclusive coastal digital twins”, with strong uptake of the outcomes of projects 2.7 and 2.8.

6.16.2 Goals

This proposal aims at developing and demonstrating a digital twin for inundation in the Portuguese coast and apply it for collaborative inundation management at selected hotspots.

6.16.3 Methodology and activity planning

The activities to be developed in this M.Sc. thesis are:

Activity 1 - Requirements analysis for addressing inundation management in Portugal

Task 1: Review of existing legislation and identification of end-users from local to national breath of intervention

Task 2: Promote trial collaborative initiatives for requirements identification and harmonization

Activity 2 - Integration of national data-sources into a national data hub for inundation

Task 1: Identification of relevant data sources and data repositories and their interoperability properties

Task 2: Creation of tools for data harvesting and harmonization towards integrated data and data processing layers

Task 3: Integration of data products into CDT’s web interface

Activity 3 - Implementation of demo observatories at some hotspots

Task 1: Implementation of complementary real time in-situ and remote monitoring networks

Task 2: Development of on-demand forecast application to the hotspots based on the OPENCoastS framework

Task 3: Identification and implementation of the most relevant digital services, including hybrid early warning system

Task 4: Testing at a few hotspots

Activity 4 - Scientific publications and dissemination

Task 1: Conference article preparation and public presentation (at the end of the year)

Task 2: Thesis preparation

Table 6.17 summarizes the timeline associated with this work.

Table 6.17 - Timeline for M.Sc. proposal 6

Year/Month	1 Year											
Activity	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
1 - Requirements analysis for addressing inundation management in Portugal												
2 - Integration of national data-sources into a national data hub												
3 - Implementation of demo observatories at some hotspots												
4 - Scientific publications and dissemination												

6.16.4 Financial plan

The list below summarizes the expected costs associated with this work.

m) Personnel costs
   - 12 months of M.Sc. grant
   - 1,5 months of a senior researcher
   - Total: 19 061.76 euros

n) Travel: 1 participation in european conference - 3000 euros

o) Equipment:
   - laptop - 1500 euros
   - e-infrastructure resources - no costs (assuming that an application to free resources is done in the Rede Nacional de Computação Avançada calls)

p) Grand total: 31 211,76 euros

7 FINAL CONSIDERATIONS

Information technology advances in the last decades have changed the future for all other sciences, comprising achievements such as on-demand forecast systems, artificial intelligence tools, data reliability methodologies, hybrid data/process-based modeling, IoT, low cost sensors and remote sensing monitoring. In parallel, the access to vast, shared e-infrastructures and the change of paradigm from proprietary models and closed access data into an open access vision for data, models and science, creates the capacity for everyone to explore these assets.

The present Research Program, framed in the area of “Maritime Hydraulics”, aims at advancing the state-of-the-art in the application of information technology towards science, short and long-term management and overall knowledge in estuarine and coastal regions by proposing a suite of methodologies, strategies, computational engines and user platforms geared at all users operating at the coast: researchers, managers, companies, associations and the general public. By integrating IT-based and conventional coastal tools, exploring the man-power in the Hydraulics and Environment Department at the coastal-related divisions and the Information Technology Research Group (GTI), this Research Program aims at placing LNEC at the top of applied coastal research institutes and reinforcing its capacity to develop the consultancy studies requested by the national entities operating at the coast using the best IT-based methodologies and tools currently available.

The theme selected for the this work, entitled “Digital coast: a scientific proposal for it-based research in coastal regions for the next decade” is framed in the research activity of GTI and in particular in LNEC’s Strategy SIIN2A project (Innovative Information Systems for Smart Water and Environmental applications), where the main foundations for this proposal, scientifically led by the applicant, were built: the Water Information Forecast Framework (WIFF) and the on-demand coastal forecast platform (OPENCoastS). The conceptualization of the program is also deeply rooted in the coastal and estuarine background of the applicant and the many collaborative projects over the past 30 years with the Estuaries and Coastal Zones division of LNEC.

Part A of this Research Program starts with a National and International framing of the research theme, followed by a state-of-the-art of the two selected areas of research and innovation, where the issues lacking investigation and their importance for coastal regions are presented in detail. These areas are:

- Reliable, quality-controlled, multi-process, on-demand coastal forecasts for oceans to hydrographic basin applications: that addresses all concerns of predicting coastal conditions for short (emergency) to long term (risk, sustainable planning), account for the dynamics of all bordering water compartments (ocean, rivers, city) in an integrated way, quantifying the cascading error propagation, the forcing and parametrization uncertainty and solves all relevant processes and their interactions from hydrodynamics to ecosystem modeling, available as a service, following the work developed in OPENCoastS.

- Coastal Digital Twins (CDT): as tools of excellence to integrate data and prediction tools in a seamless way, bringing in process-based or AI-based tools towards hybrid prediction and making the capacity to build replica of real systems available to all in a user-friendly way. CDT must be tailored and customized to user needs either to address site-specific multi-dimensional problems or thematic concerns across water bodies, aiming at building collaboratory platforms targeting co-creation of solutions by all involved and the capacity to anticipate interventions impact, climate change actions or simply to monitor coastal domains under multiple stress factors. Usability strategies will make complex tools and knowledge from the scientific and technical communities available for non-experts to ask “what if” questions and to share the work with other users in an open science, open information vision.

Two Research Studies were then proposed to address these gaps, proposing the conceptualization, development, implementation and validation of a suite of tools towards addressing these needs, using data, process/based models, artificial intelligence and related assets, in a combination of coastal science and innovation and information technology tools. The Studies are expected to be implemented over a 10-year period.

The first Research Study, OPENCoastS4ALL - Reliable, quality-controlled, multi-process, on-demand coastal forecast framework for oceans to hydrographic basin application, addresses the core engine for forecasting the coast in all its dimensions, based on the pioneering work of the applicant in the user-centered, on-demand forecast service OPENCoastS and its core framework WIFF. It integrates multiple approaches to forecast tools, from process/based to AI/models and their combination, and addresses uncertainty in all its dimensions, applicable from hydrodynamics to biogeochemistry and from the hydrographic basin to the sea, including the urban dimension.

The second Research Study, CODIT - Intelligent, high-resolution, user-centered and inclusive coastal digital twins, aims at creating for the first time high-accuracy representations of any real coastal systems through two-way interactions between digital representations and the real world, centered and operated by the users. It adresses optimal approaches to forecasting, data assimilation, reliability, quality and exploitation and aims at building flexible and customizable services, all available under the CDT concept. Like forecasting, an on/demand concept is also proposed, targeting the creation of thematic and collaborative platforms to address specific areas (pollution, inundation, ...) or conflicting uses.

The research proposed in Part A provides fertile ground for several post graduation proposals. Therefore, Part B proposes in detail 6 M.Sc. and 9 Ph.D. proposals to be developed in the ten years of expected implementation of the Habilitation Program.

These proposals aim at contributing towards DHA’s future research in the scientific area of Maritime Hydraulics and will produce Coastal Digital Twins builders, through advanced frameworks and user friendly platforms for their developments, new instruments to manage knowledge creation and user-centered management of coastal resources, tailored to specific problems such as inundation and contamination, and high accuracy collaborative tools that promote co-creation of the sustainable solutions for the coast by multiple players. The outcomes of this work are schematically represented in Figure 7.1.

The Program's implementation starts with funding from several on-going research projects, funded by national and international funding agencies. It is the applicant's goal to take advantage of the work proposed herein to promote several other projects, through collaborations inside LNEC and with national and international partners.

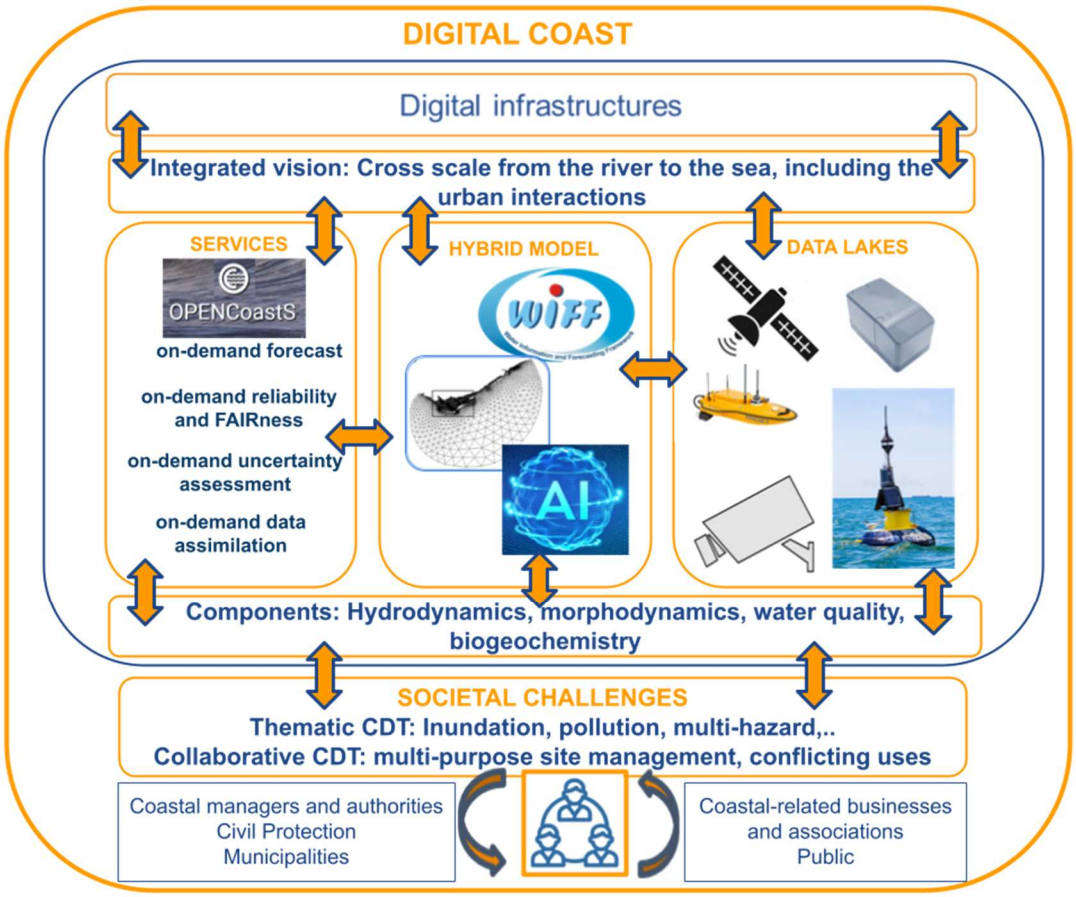


Figure 7.1 – Research program vision and outcomes.

LIST OF REFERENCES

Adams, P., Palmsten, M., 2022. On the Potential for Remote Observations of Coastal Morphodynamics from Surf-Cameras. Remote Sensing 14, 1706. https://doi.org/10.3390/rs14071706

Akbar, M., Aliabadi, S., Patel, R., Watts, M., 2013. A fully automated and integrated multi-scale forecasting scheme for emergency preparedness. Environmental Modelling & Software 39, 24–38. https://doi.org/10.1016/j.envsoft.2011.12.006

Asher, T.G., Luettich Jr., R.A., Fleming, J.G., Blanton, B.O., 2019. Low frequency water level correction in storm surge models using data assimilation. Ocean Modelling 144, 101483. https://doi.org/10.1016/j.ocemod.2019.101483

Azevedo, A., 2010. Sistema integrado de modelação para apoio à prevenção e mitigação de acidentes de hidrocarbonetos em estuários e orla costeira. Lisbon University.

Azevedo, A., Oliveira, A., Fortunato, A.B., Zhang, J., Baptista, A.M., 2014. A cross-scale numerical modeling system for management support of oil spill accidents. Marine Pollution Bulletin 80, 132–147. https://doi.org/10.1016/j.marpolbul.2014.01.028

Baptista, A., Howe, B., Freire, J., Maier, D., Silva, C., 2008. Scientific Exploration in the Era of Ocean Observatories. Comput. Sci. Eng. 10, 53–58. https://doi.org/10.1109/MCSE.2008.83

Baptista, A.M., Seaton, C., Wilkin, M.P., Riseman, S.F., Needoba, J.A., Maier, D., Turner, P.J., Kärnä, T., Lopez, J.E., Herfort, L., Megler, V.M., McNeil, C., Crump, B.C., Peterson, T.D., Spitz, Y.H., Simon, H.M., 2015. Infrastructure for collaborative science and societal applications in the Columbia River estuary. Front. Earth Sci. 9, 659–682. https://doi.org/10.1007/s11707-015-0540-5

Barbie, A., Pech, N., Hasselbring, W., Flogel, S., Wenzhofer, F., Walter, M., Shchekinova, E., Busse, M., Turk, M., Hofbauer, M., Sommer, S., 2022. Developing an Underwater Network of Ocean Observation Systems with Digital Twin Prototypes - A Field Report from the Baltic Sea. IEEE Internet Comput. 1–1. https://doi.org/10.1109/MIC.2021.3065245

Barthélémy, S., Ricci, S., Morel, T., Goutal, N., Le Pape, E., Zaoui, F., 2018. On operational flood forecasting system involving 1D/2D coupled hydraulic model and data assimilation. Journal of Hydrology 562, 623–634. https://doi.org/10.1016/j.jhydrol.2018.05.007

Bedri, Z., Corkery, A., O'Sullivan, J.J., Alvarez, M.X., Erichsen, A.Chr., Deering, L.A., Demeter, K., O'Hare, G.M.P., Meijer, W.G., Masterson, B., 2014. An integrated catchment-coastal modelling system for real-time water quality forecasts. Environmental Modelling & Software 61, 458–476. https://doi.org/10.1016/j.envsoft.2014.02.006

Bendoukha, S., Weigel, T., Fiore, S., Elia, D., 2019. Enabling Server-Based Computing and FAIR Data Sharing with the ENES Climate Analytics Service, in: 2019 15th International Conference on EScience (EScience). IEEE, San Diego, CA, USA, pp. 651–653. https://doi.org/10.1109/eScience.2019.00103

Bertin, X., Li, K., Roland, A., Zhang, Y.J., Breilh, J.F., Chaumillon, E., 2014. A modeling-based analysis of the flooding associated with Xynthia, central Bay of Biscay. Coastal Engineering 94, 80–89. https://doi.org/10.1016/j.coastaleng.2014.08.013

Bertino, L., Evensen, Geir, Wackernagel, Hans, 2002. Combining geostatistics and Kalman filtering for data assimilation in an estuarine system. Inverse Problems 18.

Blain, C.A., Preller, R.H., 2007. High Resolution Modeling of Coastal Inundation: User Requirements and Current Practice, a Navy Perspective. mar technol soc j 41, 76–83. https://doi.org/10.4031/002533207787442312

Boynton, W.R., Hagy, J.D., Breitburg, D.L., 2001. CHAPTER 11. Issues of Scale in Land-Margin Ecosystems, in: Gardner, R.H., Kemp, W.Michael., Kennedy, V.S., Petersen, J.E. (Eds.), Scaling Relations in Experimental Ecology. Columbia University Press, pp. 299–330. https://doi.org/10.7312/gard11498-014

Brassington, G.B., Oke, P.R., Pugh, T., 2007. BLUElink: OPERATIONAL OCEAN PREDICTION IN AUSTRALIA. pp. 87–95. https://doi.org/10.1142/9789812707215_0011

Breivik, Ø., Allen, A.A., 2008. An operational search and rescue model for the Norwegian Sea and the North Sea. Journal of Marine Systems 69, 99–113. https://doi.org/10.1016/j.jmarsys.2007.02.010

Brown, C.W., Hood, R.R., Long, W., Jacobs, J., Ramers, D.L., Wazniak, C., Wiggert, J.D., Wood, R., Xu, J., 2013b. Ecological forecasting in Chesapeake Bay: Using a mechanistic–empirical

modeling approach. Journal of Marine Systems 125, 113–125. https://doi.org/10.1016/j.jmarsys.2012.12.007
Buizza, R., Houtekamer, P.L., Pellerin, G., Toth, Z., Zhu, Y., Wei, M., 2005. A Comparison of the ECMWF, MSC, and NCEP Global Ensemble Prediction Systems. Monthly Weather Review 133, 1076–1097. https://doi.org/10.1175/MWR2905.1
Buscombe, D., Goldstein, E.B., Sherwood, C.R., Bodine, C., Brown, J.A., Favela, J., Fitzpatrick, S., Kranenburg, C.J., Over, J.R., Ritchie, A.C., Warrick, J.A., Wernette, P., 2022. Human-in-the-Loop Segmentation of Earth Surface Imagery. Earth and Space Science 9. https://doi.org/10.1029/2021EA002085
Campuzano, F., Brito, D., Juliano, M., Fernandes, R., de Pablo, H., Neves, R., 2016. Coupling watersheds, estuaries and regional ocean through numerical modelling for Western Iberia: a novel methodology. Ocean Dynamics 66, 1745–1756. https://doi.org/10.1007/s10236-016-1005-4
Casimiro, A., Cecílio, J., Ferreira, P., Oliveira, A., Freire, P., Rodrigues, M., Almeida, L., 2019. AQUAMON – A dependable Monitoring Platform based on Wireless Sensor Networks for Water Environments, in: 38th International Conference on on Computer Safety, Reliability and Security (SAFECOMP 2019). p. 3 pp.
Chandrasekar, K., Pathirage, M., Wijeratne, S., Mattocks, C., Plale, B., 2012. Middleware alternatives for storm surge predictions in Windows Azure, in: Proceedings of the 3rd Workshop on Scientific Cloud Computing Date - ScienceCloud '12. ACM Press, Delft, The Netherlands, p. 3. https://doi.org/10.1145/2287036.2287040
Chiu, C.-M., Huang, C.-J., Wu, L.-C., Zhang, Y.J., Chuang, L.Z.-H., Fan, Y., Yu, H.-C., 2018. Forecasting of oil-spill trajectories by using SCHISM and X-band radar. Marine Pollution Bulletin 137, 566–581. https://doi.org/10.1016/j.marpolbul.2018.10.060Chu, P., Blain, C.A., Linzell, R.S., 2009. Development of a relocatable coastal forecast system for the US Navy, in: OCEANS 2009. IEEE, Biloxi, MS, pp. 1–8. https://doi.org/10.23919/OCEANS.2009.5422137
Cipollini, P., 2010. The Role of Altimetry in Coastal Observing Systems, in: Proceedings of OceanObs'09: Sustained Ocean Observations and Information for Society. European Space Agency, pp. 166–176. https://doi.org/10.5270/OceanObs09.cwp.16
Clancy, R.M., Sadler, W.D., 1992. The Fleet Numerical Oceanography Center Suite of Oceanographic Models and Products. Wea. Forecasting 7, 307–327.
Conlin, M., Adams, P., Palmsten, M., 2022. On the Potential for Remote Observations of Coastal Morphodynamics from Surf-Cameras. Remote Sensing 14, 1706. https://doi.org/10.3390/rs14071706
Dang, T., Bulusu, N., Feng, W., Frolov, S., Baptista, A., 2010. CoTrack: A Framework for Tracking Dynamic Features with Static and Mobile Sensors, in: 2010 Proceedings IEEE INFOCOM. IEEE, San Diego, CA, USA, pp. 1–5. https://doi.org/10.1109/INFCOM.2010.5462234
Daniel, P., 1996. Operational forecasting of oil spill drift at Météo-France. Spill Science & Technology Bulletin 3, 53–64. https://doi.org/10.1016/S1353-2561(96)00030-8
Das, D.M., Singh, R., Kumar, A., Mailapalli, D.R., Mishra, A., Chatterjee, C., 2016. A Multi-Model Ensemble Approach for Stream Flow Simulation, in: Modeling Methods and Practices in Soil and Water Engineering. p. 32.
Davidson, K., Anderson, D.M., Mateus, M., Reguera, B., Silke, J., Sourisseau, M., Maguire, J., 2016. Forecasting the risk of harmful algal blooms. Harmful Algae 53, 1–7. https://doi.org/10.1016/j.hal.2015.11.005
Dekleermaeker et al, 2017. Global operational data services for storm surge and fluvial flood forecasting. https://doi.org/10.36334/MODSIM.2017.H6.dekleermaeker
Deng, T., Chau, K.-W., Duan, H.-F., 2021. Machine learning based marine water quality prediction for coastal hydro-environment management. Journal of Environmental Management 284, 112051. https://doi.org/10.1016/j.jenvman.2021.112051
Devaraju, A., Huber, R., 2021. An automated solution for measuring the progress toward FAIR research data. Patterns 2, 100370. https://doi.org/10.1016/j.patter.2021.100370
Domercq, P., Praetorius, A., MacLeod, M., 2022. The Full Multi: An open-source framework for modelling the transport and fate of nano- and microplastics in aquatic systems. Environmental Modelling & Software 148, 105291. https://doi.org/10.1016/j.envsoft.2021.105291
Duque Ordoñez, J.P. 2022. Towards a Digital Twin of the Italian Coast, M.Sc. thesis, Politecnico di Milano, 101 pp.
Etala, P., Saraceno, M., Echevarría, P., 2015. An investigation of ensemble-based assimilation of satellite altimetry and tide gauge data in storm surge prediction. Ocean Dynamics 65, 435–447. https://doi.org/10.1007/s10236-015-0808-z

Fan, C., Zhang, C., Yahja, A., Mostafavi, A., 2021. Disaster City Digital Twin: A vision for integrating artificial and human intelligence for disaster management. International Journal of Information Management 56, 102049. https://doi.org/10.1016/j.ijinfomgt.2019.102049
Fascista, A., 2022. Toward Integrated Large-Scale Environmental Monitoring Using WSN/UAV/Crowdsensing: A Review of Applications, Signal Processing, and Future Perspectives. Sensors 22, 1824. https://doi.org/10.3390/s22051824
Fassoni-Andrade, A.C., Durand, F., Moreira, D., Azevedo, A., dos Santos, V.F., Funi, C., Laraque, A., 2021. Comprehensive bathymetry and intertidal topography of the Amazon estuary. Earth Syst. Sci. Data 13, 2275–2291. https://doi.org/10.5194/essd-13-2275-2021
Feng, J., Lee, D.-K., Fu, C., Tang, J., Sato, Y., Kato, H., Mcgregor, J.L., Mabuchi, K., 2011. Comparison of four ensemble methods combining regional climate simulations over Asia. Meteorol Atmos Phys 111, 41–53. https://doi.org/10.1007/s00703-010-0115-7
Fennel, K., Gehlen, M., Brasseur, P., Brown, C.W., Ciavatta, S., Cossarini, G., Crise, A., Edwards, C.A., Ford, D., Friedrichs, M.A.M., Gregoire, M., Jones, E., Kim, H.-C., Lamouroux, J., Murtugudde, R., Perruche, C., the GODAE OceanView Marine Ecosystem Analysis and Prediction Task Team, 2019. Advancing Marine Biogeochemical and Ecosystem Reanalyses and Forecasts as Tools for Monitoring and Managing Ecosystem Health. Front. Mar. Sci. 6, 89. https://doi.org/10.3389/fmars.2019.00089
Fernández-Montblanc, T., Vousdoukas, M.I., Ciavola, P., Voukouvalas, E., Mentaschi, L., Breyiannis, G., Feyen, L., Salamon, P., 2019. Towards robust pan-European storm surge forecasting. Ocean Modelling 133, 129–144. https://doi.org/10.1016/j.ocemod.2018.12.001
Ferrarin, C., Davolio, S., Bellafiore, D., Ghezzo, M., Maicu, F., Mc Kiver, W., Drofa, O., Umgiesser, G., Bajo, M., De Pascalis, F., Malguzzi, P., Zaggia, L., Lorenzetti, G., Manfè, G., 2019. Cross-scale operational oceanography in the Adriatic Sea. Journal of Operational Oceanography 12, 86–103. https://doi.org/10.1080/1755876X.2019.1576275
Ferreira, O., Viavattene, C., Jiménez, J.A., Bolle, A., das Neves, L., Plomaritis, T.A., McCall, R., van Dongeren, A.R., 2018. Storm-induced risk assessment: Evaluation of two tools at the regional and hotspot scale. Coastal Engineering 134, 241–253. https://doi.org/10.1016/j.coastaleng.2017.10.005
Fortes, C.J.E.M., Reis, M.T., Pinheiro, L., Poseiro, P., Serrazina, V., Mendonça, A., Smithers, N., Santos, M.I., Barateiro, J., Azevedo, E.B., Salvador, M., Reis, F.V., 2020. The HIDRALERTA system: Application to the ports of Madalena do Pico and S. Roque do Pico, Azores. Aquatic Ecosystem Health & Management 23, 398–406. https://doi.org/10.1080/14634988.2020.1807295
Fortunato, A.B., Azevedo, A., Martins, R., 2022. ASSIMILATION OF SHORELINE DATA INTO MORPHODYNAMIC MODEL PREDICTIONS, in: MEC Proceedings. Lisbon, Portugal, pp. 33–34.
Fortunato, A.B., Freire, P., Mengual, B., Bertin, X., Pinto, C., Martins, K., Guérin, T., Azevedo, A., 2021. Sediment dynamics and morphological evolution in the Tagus Estuary inlet. Marine Geology 440, 106590. https://doi.org/10.1016/j.margeo.2021.106590
Fortunato, A.B., Oliveira, A., 2004. A modeling system for tidally driven long-term morphodynamics. Journal of Hydraulic Research 42, 426–434. https://doi.org/10.1080/00221686.2004.9728408
Fortunato, A.B., Oliveira, A., Rogeiro, J., Tavares da Costa, R., Gomes, J.L., Li, K., de Jesus, G., Freire, P., Rilo, A., Mendes, A., Rodrigues, M., Azevedo, A., 2017. Operational forecast framework applied to extreme sea levels at regional and local scales. Journal of Operational Oceanography 10, 1–15. https://doi.org/10.1080/1755876X.2016.1255471
French, J., Mawdsley, R., Fujiyama, T., Achuthan, K., 2017. Combining machine learning with computational hydrodynamics for prediction of tidal surge inundation at estuarine ports. Procedia IUTAM 25, 28–35. https://doi.org/10.1016/j.piutam.2017.09.005
Frolov, S., Baptista, A.M., Leen, T.K., Lu, Z., van der Merwe, R., 2009. Fast data assimilation using a nonlinear Kalman filter and a model surrogate: An application to the Columbia River estuary. Dynamics of Atmospheres and Oceans 48, 16–45. https://doi.org/10.1016/j.dynatmoce.2008.10.004
Gaitan, M.G., P. M. d'Orey, J. Cecilio, M. Rodrigues, P. M. Santos, L. Pinto, A. Oliveira, A. Casimiro, L. Almeida, (accepted for publication, 2022). Modeling LoRa Communications in Estuaries for IoT Environmental Monitoring Systems, IEEE Sensors.
Ganguli, P., Merz, B., 2019. Extreme Coastal Water Levels Exacerbate Fluvial Flood Hazards in Northwestern Europe. Sci Rep 9, 13165. https://doi.org/10.1038/s41598-019-49822-6
Giambastiani, B.M.S., Colombani, N., Greggio, N., Antonellini, M., Mastrocicco, M., 2017. Coastal aquifer response to extreme storm events in Emilia-Romagna, Italy. Hydrol. Process. 31, 1613–

1621. https://doi.org/10.1002/hyp.11130
Glenis, V., McGough, A.S., Kutija, V., Kilsby, C., Woodman, S., 2013. Flood modelling for cities using Cloud computing. J Cloud Comp 2, 7. https://doi.org/10.1186/2192-113X-2-7
Gomes, J., Bagnaschi, E., Campos, I., David, M., Alves, L., Martins, J., Pina, J., López-García, A., Orviz, P., 2018. Enabling rootless Linux Containers in multi-user environments: The udocker tool. Computer Physics Communications 232, 84–97. https://doi.org/10.1016/j.cpc.2018.05.021
Gomes, J.L., Jesus, G., Rogeiro, J., Oliveira, A., Costa, R. da, Fortunato, A.B., 2015. Molines – towards a responsive web platform for flood forecasting and risk mitigation. pp. 1171–1176. https://doi.org/10.15439/2015F265
Gracia, V., García-León, M., Sánchez-Arcilla, A., Gault, J., Oller, P., Fernández, J., Sairouní, A., Cristofori, E., Toldrà, R., 2014. A NEW GENERATION OF EARLY WARNING SYSTEMS FOR COASTAL RISK. THE ICOAST PROJECT. Int. Conf. Coastal. Eng. 1, 18. https://doi.org/10.9753/icce.v34.management.18
Hagen, J.S., Cutler, A., Trambauer, P., Weerts, A., Suarez, P., Solomatine, D., 2020. Development and evaluation of flood forecasting models for forecast-based financing using a novel model suitability matrix. Progress in Disaster Science 6, 100076. https://doi.org/10.1016/j.pdisas.2020.100076
Harley, M.D., Andriolo, U., Armaroli, C., Ciavola, P., 2014. Shoreline rotation and response to nourishment of a gravel embayed beach using a low-cost video monitoring technique: San Michele-Sassi Neri, Central Italy. J Coast Conserv 18, 551–565. https://doi.org/10.1007/s11852-013-0292-x
Hernandez, L.A. Hernandez, S., 1997. Application Of Digital 3D Models On Urban Planning And Highway Design, in: Urban and Maritime Transport 2022. p. 12.
Hribernik, K.A., Rabe, L., Thoben, K.D., Schumacher, J., 2006. The product avatar as a product-instance-centric information management concept. IJPLM 1, 367. https://doi.org/10.1504/IJPLM.2006.011055
Iglesias, I., Pinho, J.L., Avilez-Valente, P., Melo, W., Bio, A., Gomes, A., Vieira, J., Bastos, L., Veloso-Gomes, F., 2022. Improving Estuarine Hydrodynamic Forecasts Through Numerical Model Ensembles. Front. Mar. Sci. 9, 812255. https://doi.org/10.3389/fmars.2022.812255
Jacobs, G., D'Addezio, J.M., Ngodock, H., Souopgui, I., 2021. Observation and model resolution implications to ocean prediction. Ocean Modelling 159, 101760. https://doi.org/10.1016/j.ocemod.2021.101760
Jain, Sharad Kumar, Mani, P., Jain, Sanjay K., Prakash, P., Singh, V.P., Tullos, D., Kumar, S., Agarwal, S.P., Dimri, A.P., 2018. A Brief review of flood forecasting techniques and their applications. International Journal of River Basin Management 16, 329–344. https://doi.org/10.1080/15715124.2017.1411920
Jesus, G., 2019. A dependability framework for WSN-based aquatic monitoring systems. University of Lisbon, Lisbon, Portugal.
Jesus, Gonçalo, Casimiro, A., Oliveira, A., 2021a. Using Machine Learning for Dependable Outlier Detection in Environmental Monitoring Systems. ACM Trans. Cyber-Phys. Syst. 5, 1–30. https://doi.org/10.1145/3445812
Jesus, G., Casimiro, A., Oliveira, A., 2017. A Survey on Data Quality for Dependable Monitoring in Wireless Sensor Networks. Sensors 17, 2010. https://doi.org/10.3390/s17092010
Jesus, G., Casimiro, A., Oliveira, A., 2022. Systematic Failure Detection and Correction in Environmental Monitoring Systems. Sensors & Transducers 25, 28–34.
Jesus, G., Martins, R., Rogeiro, J., Oliveira, A., Azevedo, A., Fortunato, A.B., Oliveira, F., Nahon, A., 2022. Using Microsoft Azure Cognitive Services for Flood Risk Classification, in: 3º Congresso de Ensaios e Experimentação Em Engenharia Civil. Caparica, Portugal, p. 2 pp.
Jesus, G., Oliveira, A., Azevedo, A., Casimiro, A., 2015. Improving sensor-fusion with environmental models, in: 2015 IEEE SENSORS. IEEE, Busan, pp. 1–4. https://doi.org/10.1109/ICSENS.2015.7370654
Jesus, G., Oliveira, A., Casimiro, A., 2021b. Detecting Drifts and Offsets in Environmental Monitoring Networks, in: Proceedings of the 7th International Conference on Sensors and Electronic Instrumentation Advances. pp. 96–100.
Jiang, P., Meinert, N., Jordão, H., Weisser, C., Holgate, S., Lavin, A., Lütjens, B., Newman, D., Wainwright, H., Walker, C., Barnard, P., 2021. Digital Twin Earth -- Coasts: Developing a fast and physics-informed surrogate model for coastal floods via neural operators. https://doi.org/10.48550/ARXIV.2110.07100
Jupyter, P., Bussonnier, M., Forde, J., Freeman, J., Granger, B., Head, T., Holdgraf, C., Kelley, K., Nalvarte, G., Osheroff, A., Pacer, M., Panda, Y., Perez, F., Ragan-Kelley, B., Willing, C., 2018.

Binder 2.0 - Reproducible, interactive, sharable environments for science at scale. Austin, Texas, pp. 113–120. https://doi.org/10.25080/Majora-4af1f417-011
Juza, M., Tintoré, J., 2021. Multivariate Sub-Regional Ocean Indicators in the Mediterranean Sea: From Event Detection to Climate Change Estimations. Front. Mar. Sci. 8, 610589. https://doi.org/10.3389/fmars.2021.610589
Kang, S., David, D.S.K., Yang, M., Yu, Y.C., Ham, S., 2021. Energy-Efficient Ultrasonic Water Level Detection System with Dual-Target Monitoring. Sensors 21, 2241. https://doi.org/10.3390/s21062241
Kärnä, T., Baptista, A.M., 2016. Evaluation of a long-term hindcast simulation for the Columbia River estuary. Ocean Modelling 99, 1–14. https://doi.org/10.1016/j.ocemod.2015.12.007
Kärnä, T., Kramer, S.C., Mitchell, L., Ham, D.A., Piggott, M.D., Baptista, A.M., 2018. Thetis coastal ocean model: discontinuous Galerkin discretization for the three-dimensional hydrostatic equations. Geosci. Model Dev. 11, 4359–4382. https://doi.org/10.5194/gmd-11-4359-2018
Katsiokalis, M., Ragia, L., Mania, K., 2020. Outdoors Mobile Augmented Reality for Coastal Erosion Visualization Based on Geographical Data, in: Nternational Workshop on Cross-Reality (XR) Interaction 2020.
Kim, S., Matsumi, Y., Pan, S., Mase, H., 2016. A real-time forecast model using artificial neural network for after-runner storm surges on the Tottori coast, Japan. Ocean Engineering 122, 44–53. https://doi.org/10.1016/j.oceaneng.2016.06.017
Kourafalou, V.H., De Mey, P., Le Hénaff, M., Charria, G., Edwards, C.A., He, R., Herzfeld, M., Pascual, A., Stanev, E.V., Tintoré, J., Usui, N., van der Westhuysen, A.J., Wilkin, J., Zhu, X., 2015. Coastal Ocean Forecasting: system integration and evaluation. Journal of Operational Oceanography 8, s127–s146. https://doi.org/10.1080/1755876X.2015.1022336
Krapez, J.-C., Sanchis Muñoz, J., Mazel, C., Chatelard, C., Déliot, P., Frédéric, Y.-M., Barillot, P., Hélias, F., Barba Polo, J., Olichon, V., Serra, G., Brignolles, C., Carvalho, A., Carreira, D., Oliveira, A., Alves, E., Fortunato, A.B., Azevedo, A., Benetazzo, P., Bertoni, A., Le Goff, I., 2022. Multispectral Optical Remote Sensing for Water-Leak Detection. Sensors 22, 1057. https://doi.org/10.3390/s22031057
Kunze, J., Mayer, V., Thiergart, L., Javed, S., Scheppe, P., Tran, T., Haug, M., Avezum, M., Bruegge, B., Ezin, E.C., 2020. Towards SWARM: A Smart Water Monitoring System, in: 2020 IEEE Conference on Industrial Cyberphysical Systems (ICPS). Presented at the 2020 IEEE Conference on Industrial Cyberphysical Systems (ICPS), IEEE, Tampere, Finland, pp. 332–337. https://doi.org/10.1109/ICPS48405.2020.9274796
Lavaud, L., Bertin, X., Martins, K., Arnaud, G., Bouin, M.-N., 2020. The contribution of short-wave breaking to storm surges: The case Klaus in the Southern Bay of Biscay. Ocean Modelling 156, 101710. https://doi.org/10.1016/j.ocemod.2020.101710
Le Traon, P.-Y., 2011. Satellites and Operational Oceanography, in: Schiller, A., Brassington, G.B. (Eds.), Operational Oceanography in the 21st Century. Springer Netherlands, Dordrecht, pp. 29–54. https://doi.org/10.1007/978-94-007-0332-2_2
Le Traon, P.Y., Reppucci, A., Alvarez Fanjul, E., Aouf, L., Behrens, A., Belmonte, M., Bentamy, A., Bertino, L., Brando, V.E., Kreiner, M.B., Benkiran, M., Carval, T., Ciliberti, S.A., Claustre, H., Clementi, E., Coppini, G., Cossarini, G., De Alfonso Alonso-Muñoyerro, M., Delamarche, A., Dibarboure, G., Dinessen, F., Drevillon, M., Drillet, Y., Faugere, Y., Fernández, V., Fleming, A., Garcia-Hermosa, M.I., Sotillo, M.G., Garric, G., Gasparin, F., Giordan, C., Gehlen, M., Gregoire, M.L., Guinehut, S., Hamon, M., Harris, C., Hernandez, F., Hinkler, J.B., Hoyer, J., Karvonen, J., Kay, S., King, R., Lavergne, T., Lemieux-Dudon, B., Lima, L., Mao, C., Martin, M.J., Masina, S., Melet, A., Buongiorno Nardelli, B., Nolan, G., Pascual, A., Pistoia, J., Palazov, A., Piolle, J.F., Pujol, M.I., Pequignet, A.C., Peneva, E., Pérez Gómez, B., Petit de la Villeon, L., Pinardi, N., Pisano, A., Pouliquen, S., Reid, R., Remy, E., Santoleri, R., Siddorn, J., She, J., Staneva, J., Stoffelen, A., Tonani, M., Vandenbulcke, L., von Schuckmann, K., Volpe, G., Wettre, C., Zacharioudaki, A., 2019. From Observation to Information and Users: The Copernicus Marine Service Perspective. Front. Mar. Sci. 6, 234. https://doi.org/10.3389/fmars.2019.00234
Li, J., Nie, B., 2017. Storm surge prediction: present status and future challenges. Procedia IUTAM 25, 3–9. https://doi.org/10.1016/j.piutam.2017.09.002
Lionello, P., Sanna, A., Elvini, E., Mufato, R., 2006. A data assimilation procedure for operational prediction of storm surge in the northern Adriatic Sea. Continental Shelf Research 26, 539–553. https://doi.org/10.1016/j.csr.2006.01.003
Liu, M., Fang, S., Dong, H., Xu, C., 2021. Review of digital twin about concepts, technologies, and industrial applications. Journal of Manufacturing Systems 58, 346–361. https://doi.org/10.1016/j.jmsy.2020.06.017

Loftis, J.D., Mitchell, M., Schatt, D., Forrest, D.R., Wang, H.V., Mayfield, D., Stiles, W.A., 2019. Validating an Operational Flood Forecast Model Using Citizen Science in Hampton Roads, VA, USA. JMSE 7, 242. https://doi.org/10.3390/jmse7080242

Luettich, R.A., Luettich, J.J., Scheffner, N.W., 1992. ADCIRC: an advanced three-dimensional circulation model for shelves coasts and estuaries, report 1: theory and methodology of ADCIRC-2DDI and ADCIRC-3DL (No. Dredging Research Program Technical Report DRP-92-6). U.S. Army Engineers Waterways Experiment Station, Vicksburg.

Lyard, F.H., Allain, D.J., Cancet, M., Carrère, L., Picot, N., 2021. FES2014 global ocean tide atlas: design and performance. Ocean Sci. 17, 615–649. https://doi.org/10.5194/os-17-615-2021

Lyhne Christensen, A., Oliveira, S., Postolache, O., João de Oliveira, M., Sargento, S., Santana, P., Nunes, L., Velez, F., Sebastião, P., Costa, V., Duarte, M., Gomes, J., Rodrigues, T., Silva, F., 2015. Design of Communication and Control for Swarms of Aquatic Surface Drones:, in: Proceedings of the International Conference on Agents and Artificial Intelligence. SCITEPRESS - Science and and Technology Publications, Lisbon, Portugal, pp. 548–555. https://doi.org/10.5220/0005281705480555

Madec, G. 2008. NEMO ocean engine (Note du Pole de Modélisation, 27) France. Institut Pierre-Simon Laplace (IPSL) 300pp.

Maier, D., Megler, V.M., Baptista, A.M., Jaramillo, A., Seaton, C., Turner, P.J., 2012. Navigating Oceans of Data, in: Ailamaki, A., Bowers, S. (Eds.), Scientific and Statistical Database Management. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 1–19.

Malczewski, J., Jankowski, P., 2020. Emerging trends and research frontiers in spatial multicriteria analysis. International Journal of Geographical Information Science 34, 1257–1282. https://doi.org/10.1080/13658816.2020.1712403

Mani, S., 2021. OCSMesh: a data-driven automated unstructured mesh generation software for coastal ocean modeling. https://doi.org/10.25923/CSBA-M072

Marimon, M.C., Villegas, E.I.T., Borja, M.J.H., Tangonan, G.L., Libatique, N.J.C., 2011. Low cost sensor system for wave monitoring, in: 2011 Seventh International Conference on Intelligent Sensors, Sensor Networks and Information Processing. IEEE, Adelaide, Australia, pp. 312–316. https://doi.org/10.1109/ISSNIP.2011.6146584

Martins, R., Jesus, G., Oliveira, A., Azevedo, A., Fortunato, A.B., Oliveira, F., Nahon, A., 2022. A METHODOLOGY FOR WATER BOUNDARY DETECTION USING HIGH - RESOLUTION REMOTE CAMERA IMAGES, in: Proceedings Do 3º Congresso de Ensaios e Experimentação Em Engenharia Civil. p. 2 pp.

Mateus, M., Riflet, G., Chambel, P., Fernandes, L., Fernandes, R., Juliano, M., Campuzano, F., de Pablo, H., Neves, R., 2012. An operational model for the West Iberian coast: products and services. Ocean Sci. 8, 713–732. https://doi.org/10.5194/os-8-713-2012

Matheen, N., Harley, M.D., Turner, I.L., Splinter, K.D., Simmons, J.A., Thran, M.C., 2021. Bathymetric Data Requirements for Operational Coastal Erosion Forecasting Using XBeach. JMSE 9, 1053. https://doi.org/10.3390/jmse9101053

Mcknight, L., Rodrigues, M., Fortunato, A.B., Clarindo, G., Rodríguez, G., 2017. Nearshore Circulation in the Confital Bay: Implications on Marine Debris Transport and Deposition at Las Canteras Beach, in: Fate and Impact of Microplastics in Marine Ecosystems. Elsevier, p. 87. https://doi.org/10.1016/B978-0-12-812271-6.00088-0

Mehra, A., Rivin, I., 2010. A Real Time Ocean Forecast System for the North Atlantic Ocean. Terr. Atmos. Ocean. Sci. 21, 211. https://doi.org/10.3319/TAO.2009.04.16.01(IWNOP)

Mel, R., Lionello, P., 2016. Probabilistic Dressing of a Storm Surge Prediction in the Adriatic Sea. Advances in Meteorology 2016, 1–8. https://doi.org/10.1155/2016/3764519

Mohan Das, D., Singh, R., Kumar, A., Mailapalli, D.R., Mishira, A., Chatterjee, C., 2016. A Multi-Model Ensemble Approach For Stream Flow Simulation, MODELING METHODS AND PRACTICES IN SOIL AND WATER ENGINEERING, chapter 4, 72-100.

Moghimi, S., Myers, E., Pe'eri, S., Zhang, Y., Yi, F., 2021. Forecasting Compound Floods in Complex Coastal Regions. Eos 102. https://doi.org/10.1029/2021EO210604

Mosavi, A., Ozturk, P., Chau, K., 2018. Flood Prediction Using Machine Learning Models: Literature Review. Water 10, 1536. https://doi.org/10.3390/w10111536

Nahon, A., Fortunato, A.B., Azevedo, A., Oliveira, F.S.B.F., Oliveira, J.N.C., Rogeiro, J., Jesus, G., Oliveira, A., Silva, P.A., Freire,P. 2020. Implementation and validation of an operational forecasting system for nearshore hydrodynamics with OPENCoastS. Proceedings as 6as Jornadas de Engenharia Hidrográfica / 1as Jornadas Luso-Espanholas de Hidrografia. 6.as Jornadas de Engenharia Hidrográfica, 203-206, Instituto Hidrográfico, Lisboa.

Neves, M.G., Didier, E., Brito, M., Clavero, M., 2021. Numerical and Physical Modelling of Wave

Overtopping on a Smooth Impermeable Dike with Promenade under Strong Incident Waves. JMSE 9, 865. https://doi.org/10.3390/jmse9080865
Novelli, G., Guigand, C.M., Cousin, C., Ryan, E.H., Laxague, N.J.M., Dai, H., Haus, B.K., Özgökmen, T.M., 2017. A Biodegradable Surface Drifter for Ocean Sampling on a Massive Scale. Journal of Atmospheric and Oceanic Technology 34, 2509–2532. https://doi.org/10.1175/JTECH-D-17-0055.1
O'Donncha, F., Ragnoli, E., Venugopal, S., James, S.C., Katrinis, K., 2016. On the Efficiency of Executing Hydro-environmental Models on Cloud. Procedia Engineering 154, 199–206. https://doi.org/10.1016/j.proeng.2016.07.447
Oke, P.R., Proctor, R., Rosebrock, U., Brinkman, R., Cahill, M.L., Coghlan, I., Divakaran, P., Freeman, J., Pattiaratchi, C., Roughan, M., Sandery, P.A., Schaeffer, A., Wijeratne, S., 2016. The Marine Virtual Laboratory (version 2.1): enabling efficient ocean model configuration. Geosci. Model Dev. 9, 3297–3307. https://doi.org/10.5194/gmd-9-3297-2016
Okuno, S., Ikeuchi, K., Aihara, K., 2021. Practical Data-Driven Flood Forecasting Based on Dynamical Systems Theory. Water Res 57. https://doi.org/10.1029/2020WR028427
Oliveira, A., Baptista, A.M., 1997. Diagnostic modeling of residence times in estuaries. Water Resour. Res. 33, 1935–1946. https://doi.org/10.1029/97WR00653
Oliveira, A., Fortunato, A.B., Rodrigues, M., Azevedo, A., Rogeiro, J., Bernardo, S., Lavaud, L., Bertin, X., Nahon, A., de Jesus, G., Rocha, M., Lopes, P., 2021. Forecasting contrasting coastal and estuarine hydrodynamics with OPENCoastS. Environmental Modelling & Software 143, 105132. https://doi.org/10.1016/j.envsoft.2021.105132
Oliveira, A., Fortunato, A.B., Rogeiro, J., Teixeira, J., Azevedo, A., Lavaud, L., Bertin, X., Gomes, J., David, M., Pina, J., Rodrigues, M., Lopes, P., 2020. OPENCoastS: An open-access service for the automatic generation of coastal forecast systems. Environmental Modelling & Software 124, 104585. https://doi.org/10.1016/j.envsoft.2019.104585
Oliveira, A., Jesus, G., Gomes, J.L., Rogeiro, J., Azevedo, A., Rodrigues, M., Fortunato, A.B., Dias, J.M., Tomas, L.M., Vaz, L., Oliveira, E.R., Alves, F.L., Boer, S. den, 2014. An interactive WebGIS observatory platform for enhanced support of integrated coastal management. Journal of Coastal Research 70, 507–512. https://doi.org/10.2112/SI70-086.1
Oliver, J.M., Esteban, M.D., López-Gutiérrez, J.-S., Negro, V., Neves, M.G., 2021. Optimizing Wave Overtopping Energy Converters by ANN Modelling: Evaluating the Overtopping Rate Forecasting as the First Step. Sustainability 13, 1483. https://doi.org/10.3390/su13031483
Orseau, S., Huybrechts, N., Tassi, P., Kaidi, S., Klein, F., 2021. NavTEL: Open-Source Decision Support Tool for Ship Routing and Underkeel Clearance Management in Estuarine Channels. J. Waterway, Port, Coastal, Ocean Eng. 147, 04020053. https://doi.org/10.1061/(ASCE)WW.1943-5460.0000610
Orviz, P., Campos, I., EOSC-SYNERGY, 2020. EOSC-SYNERGY EU DELIVERABLE: D3.2 First prototype of the Service Integration platform. https://doi.org/10.20350/DIGITALCSIC/12721
Palmer, T.N., Alessandri, A., Andersen, U., Cantelaube, P., Davey, M., Délécluse, P., Déqué, M., Díez, E., Doblas-Reyes, F.J., Feddersen, H., Graham, R., Gualdi, S., Guérémy, J.-F., Hagedorn, R., Hoshen, M., Keenlyside, N., Latif, M., Lazar, A., Maisonnave, E., Marletto, V., Morse, A.P., Orfila, B., Rogel, P., Terres, J.-M., Thomson, M.C., 2004. Development of A European Multimodel Ensemble System For Seasonal-To-Interannual Prediction (Demeter). Bulletin of the American Meteorological Society 85, 853–872. https://doi.org/10.1175/BAMS-85-6-853
Panigrahi, B., Goyal, M.R., 2016. Modeling methods and practices in soil and water engineering. ISBN: 978-1-315-36643-2.
Petritoli, E., Cagnetti, M., Leccese, F., 2020. Simulation of Autonomous Underwater Vehicles (AUVs) Swarm Diffusion. Sensors 20, 4950. https://doi.org/10.3390/s20174950
Pedersen, A.N., Borup, M., Brink-Kjær, A., Christiansen, L.E., Mikkelsen, P.S., 2021. Living and Prototyping Digital Twins for Urban Water Systems: Towards Multi-Purpose Value Creation Using Models and Sensors. Water 13, 592. https://doi.org/10.3390/w13050592
Pereira, H., Sousa, M.C., Vieira, L.R., Morgado, F., Dias, J.M., 2022. Modelling Salt Intrusion and Estuarine Plumes under Climate Change Scenarios in Two Transitional Ecosystems from the NW Atlantic Coast. JMSE 10, 262. https://doi.org/10.3390/jmse10020262
Picado, A., Abreu, J., Carvalho, D.M., Santos, H., Leitão, P., Rodrigues, J., Chambel Leitão, J., Dias, J.M., 2022. Optimization of safe navigability conditions in shallow inland waters: The Routinav app. Ocean Engineering 259, 111946. https://doi.org/10.1016/j.oceaneng.2022.111946
Poseiro, P.G.G., 2019. Forecast and Early Warning System for Wave Overtopping and Flooding in Coastal and Port Areas: Development of a Model and Risk Assessment., Tese de Doutoramento em Engenharia Civil, Instituto Superior Técnico, Universidade de Lisboa. ed.

Pottier, A., Catry, T., Trégarot, E., Maréchal, J.-P., Fayad, V., David, G., Sidi Cheikh, M., Failler, P., 2021. Mapping coastal marine ecosystems of the National Park of Banc d'Arguin (PNBA) in Mauritania using Sentinel-2 imagery. International Journal of Applied Earth Observation and Geoinformation 102, 102419. https://doi.org/10.1016/j.jag.2021.102419

Raghukumar, K., Chang, G., Spada, F., Jones, C., Janssen, T., Gans, A., 2019. Performance Characteristics of "Spotter," a Newly Developed Real-Time Wave Measurement Buoy. Journal of Atmospheric and Oceanic Technology 36, 1127–1141. https://doi.org/10.1175/JTECH-D-18-0151.1

Révelard, A., Reyes, E., Mourre, B., Hernández-Carrasco, I., Rubio, A., Lorente, P., Fernández, C.D.L., Mader, J., Álvarez-Fanjul, E., Tintoré, J., 2021. Sensitivity of Skill Score Metric to Validate Lagrangian Simulations in Coastal Areas: Recommendations for Search and Rescue Applications. Front. Mar. Sci. 8, 630388. https://doi.org/10.3389/fmars.2021.630388

Rocha, M., Oliveira, A., Freire, P., Fortunato, A.B., Nahon, A., Barros, J.L., Azevedo, A., Oliveira, F.S.B.F., Rogeiro, J., Jesus, G., Martins, R.J., Santos, P.P., Tavares, A.O., Oliveira, J., 2021. Multi-Hazard WebGIS Platform for Coastal Regions. Applied Sciences 11, 5253. https://doi.org/10.3390/app11115253

Rodrigues, M., Costa, J., Jesus, G., Fortunato, A.B., Rogeiro, J., Gomes, J.L., Oliveira, A., David, L., 2013. Application of an estuarine and coastal nowcast-forecast information system to the Tagus estuary, in: 6th SCACR – International Short Course/Conference on Applied Coastal Research. p. 10 pp.

Rodrigues, M., Cravo, A., Freire, P., Rosa, A., Santos, D., 2020. Temporal assessment of the water quality along an urban estuary (Tagus estuary, Portugal). Marine Chemistry 223, 103824. https://doi.org/10.1016/j.marchem.2020.103824

Rodrigues, M., Jesus, G., Rogeiro, J., Fortunato, A.B., Gomes, J.L., Oliveira, A., 2015. RealQual: Towards real-time high-resolution monitoring and prediction of water quality in estuaries and coastal areas - Final Report (No. Relatório 68/2015). LNEC.

Rodrigues, M., Martins, R., Rogeiro, J., Fortunato, A.B., Oliveira, A., Cravo, A., Jacob, J., Rosa, A., Azevedo, A., Freire, P., 2021. A Web-Based Observatory for Biogeochemical Assessment in Coastal Regions. J ENVIRON INFORM. https://doi.org/10.3808/jei.202100450

Roemmich, D., Johnson, G., Riser, S., Davis, R., Gilson, J., Owens, W.B., Garzoli, S., Schmid, C., Ignaszewski, M., 2009. The Argo Program: Observing the Global Oceans with Profiling Floats. Oceanog. 22, 34–43. https://doi.org/10.5670/oceanog.2009.36

Rogeiro, J., Rodrigues, M., Azevedo, A., Oliveira, A., Martins, J.P., David, M., Pina, J., Dias, N., Gomes, J., 2018. Running high resolution coastal models in forecast systems: Moving from workstations and HPC cluster to cloud resources. Advances in Engineering Software 117, 70–79. https://doi.org/10.1016/j.advengsoft.2017.04.002

Rosebrock, U., Oke, P.R., Carroll, G., 2015. An Application Framework for the Rapid Deployment of Ocean Models in Support of Emergency Services: Application to the MH370 Search, in: Denzer, R., Argent, R.M., Schimak, G., Hřebíček, J. (Eds.), Environmental Software Systems. Infrastructures, Services and Applications, IFIP Advances in Information and Communication Technology. Springer International Publishing, Cham, pp. 235–241. https://doi.org/10.1007/978-3-319-15994-2_23

Rowley, C., Mask, A., 2014. Regional and Coastal Prediction with the Relocatable Ocean Nowcast/Forecast System. oceanog 27, 44–55. https://doi.org/10.5670/oceanog.2014.67

Sanchez-Arcilla, A., Staneva, J., Cavaleri, L., Badger, M., Bidlot, J., Sorensen, J.T., Hansen, L.B., Martin, A., Saulter, A., Espino, M., Miglietta, M.M., Mestres, M., Bonaldo, D., Pezzutto, P., Schulz-Stellenfleth, J., Wiese, A., Larsen, X., Carniel, S., Bolaños, R., Abdalla, S., Tiesi, A., 2021. CMEMS-Based Coastal Analyses: Conditioning, Coupling and Limits for Applications. Front. Mar. Sci. 8, 604741. https://doi.org/10.3389/fmars.2021.604741

Schumacher, B., Griffiths, P., Pebesma, E., Dries, J., Jacob, A., Thiex, D., Mohr, M., Briese, C., 2022. openEO Platform: Enabling analysis of large-scale Earth Observation data repositories with federated computational infrastructure (other). display. https://doi.org/10.5194/egusphere-egu22-9101

Shaffer, W. A., C. P. Jelesnianski and J. Chen, 1989. Hurricane Storm Surge Forecasting. 11th Conference on Probability and Statistics in Atmospheric Sciences, Amer. Meteor. Soc., Monterey, CA, pp. J53–J58.

Shafto, M., Conroy, M., Doyle, R, Glaessgen, E., Kemp, C., LeMoigne, J., Wang, L. 2010. Modeling, Simulation, InformationTechnology & Processing Roadmap Technology Area 11, NASA, 38 pp.

Sharma, A., Kosasih, E., Zhang, J., Brintrup, A., Calinescu, A., 2022. Digital Twins: State of the art theory and practice, challenges, and open research questions. Journal of Industrial Information

Integration 30, 100383. https://doi.org/10.1016/j.jii.2022.100383
Smith, N.R., 2000. The Global Ocean Data Assimilation Experiment. Advances in Space Research 25, 1089–1098. https://doi.org/10.1016/S0273-1177(99)00868-6
Solano, M., Canals, M., Leonardi, S., 2018. Development and validation of a coastal ocean forecasting system for Puerto Rico and the U.S. virgin islands. Journal of Ocean Engineering and Science 3, 223–236. https://doi.org/10.1016/j.joes.2018.08.004
Solomatine, D.P., Ostfeld, A., 2008. Data-driven modelling: some past experiences and new approaches. Journal of Hydroinformatics 10, 3–22. https://doi.org/10.2166/hydro.2008.015
Song, H., Edwards, C.A., Moore, A.M., Fiechter, J., 2016. Data assimilation in a coupled physical-biogeochemical model of the California current system using an incremental lognormal 4-dimensional variational approach: Part 3—Assimilation in a realistic context using satellite and in situ observations. Ocean Modelling 106, 159–172. https://doi.org/10.1016/j.ocemod.2016.06.005
Soroosh, M., Calzada, J., Moghimi, S., Zhang, J., Myers, E., Pe'eri, S., 2021. OCSMesh: a data-driven automated unstructured mesh generation software for coastal ocean modeling (NOAA Technical Memorandum NOS CS 47).
Stanev, E.V., Schulz-Stellenfleth, J., Staneva, J., Grayek, S., Grashorn, S., Behrens, A., Koch, W., Pein, J., 2016. Ocean Forecasting: From Regional to Coastal Scales (preprint). All Depths/Numerical Models/Shelf Seas/Tides. https://doi.org/10.5194/os-2016-25
Stokes, K., Poate, T., Masselink, G., King, E., Saulter, A., Ely, N., 2021. Forecasting coastal overtopping at engineered and naturally defended coastlines. Coastal Engineering 164, 103827. https://doi.org/10.1016/j.coastaleng.2020.103827
Suh, M.-S., Oh, S.-G., Lee, D.-K., Cha, D.-H., Choi, S.-J., Jin, C.-S., Hong, S.-Y., 2012. Development of New Ensemble Methods Based on the Performance Skills of Regional Climate Models over South Korea. Journal of Climate 25, 7067–7082. https://doi.org/10.1175/JCLI-D-11-00457.1
Tavares, A.O., Barros, J.L., Freire, P., Santos, P.P., Perdiz, L., Fortunato, A.B., 2021. A coastal flooding database from 1980 to 2018 for the continental Portuguese coastal zone. Applied Geography 135, 102534. https://doi.org/10.1016/j.apgeog.2021.102534
Tebaldi, C., Knutti, R., 2007. The use of the multi-model ensemble in probabilistic climate projections. Phil. Trans. R. Soc. A. 365, 2053–2075. https://doi.org/10.1098/rsta.2007.2076
Templin, T., Popielarczyk, D., Gryszko, M., 2022. Using Augmented and Virtual Reality (AR/VR) to Support Safe Navigation on Inland and Coastal Water Zones. Remote Sensing 14, 1520. https://doi.org/10.3390/rs14061520
Tintoré, J., Pinardi, N., Álvarez-Fanjul, E., Aguiar, E., Álvarez-Berastegui, D., Bajo, M., Balbin, R., Bozzano, R., Nardelli, B.B., Cardin, V., Casas, B., Charcos-Llorens, M., Chiggiato, J., Clementi, E., Coppini, G., Coppola, L., Cossarini, G., Deidun, A., Deudero, S., D'Ortenzio, F., Drago, A., Drudi, M., El Serafy, G., Escudier, R., Farcy, P., Federico, I., Fernández, J.G., Ferrarin, C., Fossi, C., Frangoulis, C., Galgani, F., Gana, S., García Lafuente, J., Sotillo, M.G., Garreau, P., Gertman, I., Gómez-Pujol, L., Grandi, A., Hayes, D., Hernández-Lasheras, J., Herut, B., Heslop, E., Hilmi, K., Juza, M., Kallos, G., Korres, G., Lecci, R., Lazzari, P., Lorente, P., Liubartseva, S., Louanchi, F., Malacic, V., Mannarini, G., March, D., Marullo, S., Mauri, E., Meszaros, L., Mourre, B., Mortier, L., Muñoz-Mas, C., Novellino, A., Obaton, D., Orfila, A., Pascual, A., Pensieri, S., Pérez Gómez, B., Pérez Rubio, S., Perivoliotis, L., Petihakis, G., de la Villéon, L.P., Pistoia, J., Poulain, P.-M., Pouliquen, S., Prieto, L., Raimbault, P., Reglero, P., Reyes, E., Rotllan, P., Ruiz, S., Ruiz, J., Ruiz, I., Ruiz-Orejón, L.F., Salihoglu, B., Salon, S., Sammartino, S., Sánchez Arcilla, A., Sánchez-Román, A., Sannino, G., Santoleri, R., Sardá, R., Schroeder, K., Simoncelli, S., Sofianos, S., Sylaios, G., Tanhua, T., Teruzzi, A., Testor, P., Tezcan, D., Torner, M., Trotta, F., Umgiesser, G., von Schuckmann, K., Verri, G., Vilibic, I., Yucel, M., Zavatarelli, M., Zodiatis, G., 2019. Challenges for Sustained Observing and Forecasting Systems in the Mediterranean Sea. Front. Mar. Sci. 6, 568. https://doi.org/10.3389/fmars.2019.00568
Trotta, F., Fenu, E., Pinardi, N., Bruciaferri, D., Giacomelli, L., Federico, I., Coppini, G., 2016. A Structured and Unstructured grid Relocatable ocean platform for Forecasting (SURF). Deep Sea Research Part II: Topical Studies in Oceanography 133, 54–75. https://doi.org/10.1016/j.dsr2.2016.05.004
Trotta, F., Pinardi, N., Fenu, E., Grandi, A., Lyubartsev, V., 2017. Multi-nest high-resolution model of submesoscale circulation features in the Gulf of Taranto. Ocean Dynamics 67, 1609–1625. https://doi.org/10.1007/s10236-017-1110-z
Trotta, F., Federico, I., Pinardi, N., Coppini, G., Causio, S., Jansen, E., Iovino, D., Masina, S., 2021. A Relocatable Ocean Modeling Platform for Downscaling to Shelf-Coastal Areas to Support Disaster Risk Reduction. Front. Mar. Sci. 8, 642815. https://doi.org/10.3389/fmars.2021.642815

Umgiesser, G., Bajo, M., Ferrarin, C., Cucco, A., Lionello, P., Zanchettin, D., Papa, A., Tosoni, A., Ferla, M., Coraci, E., Morucci, S., Crosato, F., Bonometto, A., Valentini, A., Orlic, M., Haigh, I.D., Nielsen, J.W., Bertin, X., Fortunato, A.B., Pérez Gómez, B., Alvarez Fanjul, E., Paradis, D., Jourdan, D., Pasquet, A., Mourre, B., Tintoré, J., Nicholls, R.J., 2020. The prediction of floods in Venice: methods, models and uncertainty (preprint). Sea, Ocean and Coastal Hazards. https://doi.org/10.5194/nhess-2020-361

Umgiesser, G., Canu, D.M., Cucco, A., Solidoro, C., 2004. A finite element model for the Venice Lagoon. Development, set up, calibration and validation. Journal of Marine Systems 51, 123–145. https://doi.org/10.1016/j.jmarsys.2004.05.009

UNESCO/IOC, 2016. H2020 Programme: Guidelines on FAIR Data Management in Horizon 2020 Version 3.0. https://doi.org/10.25607/OBP-774

Valchev, N., Eftimova, P., Andreeva, N., 2018. Implementation and validation of a multi-domain coastal hazard forecasting system in an open bay. Coastal Engineering 134, 212–228. https://doi.org/10.1016/j.coastaleng.2017.08.008

Valipour, M., Banihabib, M.E., Behbahani, S.M.R., 2013. Comparison of the ARMA, ARIMA, and the autoregressive artificial neural network models in forecasting the monthly inflow of Dez dam reservoir. Journal of Hydrology 476, 433–441. https://doi.org/10.1016/j.jhydrol.2012.11.017

van der Merwe, R., Leen, T.K., Lu, Z., Frolov, S., Baptista, A.M., 2007. Fast neural network surrogates for very high dimensional physics-based models in computational oceanography. Neural Networks 20, 462–478. https://doi.org/10.1016/j.neunet.2007.04.023

van Dongeren, A., Ciavola, P., Martinez, G., Viavattene, C., Bogaard, T., Ferreira, O., Higgins, R., McCall, R., 2018. Introduction to RISC-KIT: Resilience-increasing strategies for coasts. Coastal Engineering 134, 2–9. https://doi.org/10.1016/j.coastaleng.2017.10.007

van Velzen, N., Altaf, M.U., Verlaan, M., 2016. OpenDA-NEMO framework for ocean data assimilation. Ocean Dynamics 66, 691–702. https://doi.org/10.1007/s10236-016-0945-z

van Velzen, N., Verlaan, M., 2007. COSTA a problem solving environment for data assimilation applied for hydrodynamical modelling. metz 16, 777–793. https://doi.org/10.1127/0941-2948/2007/0241

Vaz, N., Mateus, M., Pinto, L., Neves, R., Dias, J.M., 2019. The Tagus Estuary as a Numerical Modeling Test Bed: A Review. Geosciences 10, 4. https://doi.org/10.3390/geosciences10010004

Viegas et al., 2009. Streams contribution on bathing water quality after rainfall events in Costa do Estoril-a tool to implement an alert system for bathing water quality.

Viney, N.R., Croke, B.F.W., Breuer, L., Bormann, H., Bronstert, A., Frede, H., Gräff, T., Hubrechts, L., et al, 2005. Ensemble modelling of the hydrological impacts of land use change, in: MODSIM05 - International Congress on Modelling and Simulation: Advances and Applications for Management and Decision Making, Proceedings.

Weigel, A.P., Liniger, M.A., Appenzeller, C., 2008. Can multi-model combination really enhance the prediction skill of probabilistic ensemble forecasts? Q.J.R. Meteorol. Soc. 134, 241–260. https://doi.org/10.1002/qj.210

Werner, M., Schellekens, J., Gijsbers, P., van Dijk, M., van den Akker, O., Heynert, K., 2013. The Delft-FEWS flow forecasting system. Environmental Modelling & Software 40, 65–77. https://doi.org/10.1016/j.envsoft.2012.07.010

Wolff, T., Carrillo, H., Marti, L., Sanchez-Pi, N., 2021. Towards Optimally Weighted Physics-Informed Neural Networks in Ocean Modelling.

Wynn, R.B., Huvenne, V.A.I., Le Bas, T.P., Murton, B.J., Connelly, D.P., Bett, B.J., Ruhl, H.A., Morris, K.J., Peakall, J., Parsons, D.R., Sumner, E.J., Darby, S.E., Dorrell, R.M., Hunt, J.E., 2014. Autonomous Underwater Vehicles (AUVs): Their past, present and future contributions to the advancement of marine geoscience. Marine Geology 352, 451–468. https://doi.org/10.1016/j.margeo.2014.03.012

Ye, F., Zhang, Y.J., Yu, H., Sun, W., Moghimi, S., Myers, E., Nunez, K., Zhang, R., Wang, H.V., Roland, A., Martins, K., Bertin, X., Du, J., Liu, Z., 2020. Simulating storm surge and compound flooding events with a creek-to-ocean model: Importance of baroclinic effects. Ocean Modelling 145, 101526. https://doi.org/10.1016/j.ocemod.2019.101526

Young, C.-C., Liu, W.-C., Wu, M.-C., 2017. A physically based and machine learning hybrid approach for accurate rainfall-runoff modeling during extreme typhoon events. Applied Soft Computing 53, 205–216. https://doi.org/10.1016/j.asoc.2016.12.052

Zang, Z., Ji, R., Feng, Z., Chen, C., Li, S., Davis, C.S., 2021. Spatially varying phytoplankton seasonality on the Northwest Atlantic Shelf: a model-based assessment of patterns, drivers, and implications. ICES Journal of Marine Science 78, 1920–1934. https://doi.org/10.1093/icesjms/fsab102

Zhang, Y., Meratnia, N., Havinga, P., 2010. Outlier Detection Techniques for Wireless Sensor Networks:

A Survey. IEEE Commun. Surv. Tutorials 12, 159–170. https://doi.org/10.1109/SURV.2010.021510.00088
Zhang, Y.J., Ye, F., Stanev, E.V., Grashorn, S., 2016. Seamless cross-scale modeling with schism. Ocean Model. 102, 64–81.
Zhang, C., Xu, W., Liu, J., Liu, Z., Zhou, Z., Pham, D.T., 2021. Digital twin-enabled reconfigurable modeling for smart manufacturing systems. International Journal of Computer Integrated Manufacturing 34, 709–733. https://doi.org/10.1080/0951192X.2019.1699256
Zhou, J.-F., Li, J.-C., 2005. Modeling Storm-Induced Current Circulation and Sediment Transport in a Schematic Harbor. J. Waterway, Port, Coastal, Ocean Eng. 131, 25–32. https://doi.org/10.1061/(ASCE)0733-950X(2005)131:1(25)
Zijl, F., Sumihar, J., Verlaan, M., 2015. Application of data assimilation for improved operational water level forecasting on the northwest European shelf and North Sea. Ocean Dynamics 65, 1699–1716. https://doi.org/10.1007/s10236-015-0898-7
Zou, J., Zhan, C., Song, H., Hu, T., Qiu, Z., Wang, B., Li, Z., 2021. Development and Evaluation of a Hydrometeorological Forecasting System Using the Coupled Ocean-Atmosphere-Wave-Sediment Transport (COAWST) Model. Advances in Meteorology 2021, 1–17. https://doi.org/10.1155/2021/6658722

Lisboa, LNEC, setembro de 2022

AUTORIA

_____

Anabela Oliveira