



interTwin

D7.3 First version of the thematic modules for the environment domain

Status: APPROVED BY THE EC
Dissemination Level: public



Funded by the
European Union

Disclaimer: Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union. Neither the European Union nor the granting authority can be held responsible for them


Abstract

Key Words

Digital Twins, Thematic modules, environment domain, development, integration

This report describes the status of development of thematic modules that have been identified at this stage of the project to support the development of Digital Twins in the environmental domain.



Document Description			
D7.3 First version of the thematic module for the environment domain			
Work Package number 6			
Document type	Deliverable		
Document status	APPROVED BY THE EC	Version	1
Dissemination Level	Public		
Copyright Status	 <p>This material by Parties of the interTwin Consortium is licensed under a Creative Commons Attribution 4.0 International License.</p>		
Lead Partner	UNITN		
Document link	https://documents.egi.eu/document/3957		
DOI	https://zenodo.org/records/14974970		
Author(s)	<ul style="list-style-type: none"> • Sandro Fiore (UNITN), • Massimiliano Fronza (UNITN) • Donatello Elia (CMCC), • Davide Donno (CMCC), • Emanuele Donno (CMCC) • Bjorn Backeberg (DELTARES) • Alexander Jacob (EURAC), • Michele Claus (EURAC) • Christian Pagè (CERFACS) 		
Reviewers	<ul style="list-style-type: none"> • Estíbaliz Parceros Iglesias (UPV) • Kostas Koumantaros (GRNET) 		
Moderated by:	<ul style="list-style-type: none"> • Sandro Fiore (UNITN), • Massimiliano Fronza (UNITN) • Sjomara Specht (EGI) 		
Approved by	AMB		

Revision History			
Version	Date	Description	Contributors
V0.1	05/09/2023	ToC	Sandro Fiore (UNITN)
V0.2	01/11/2023		Sandro Fiore (UNITN), Massimiliano Fronza (UNITN), Donatello Elia (CMCC), Davide Donno (CMCC), Emanuele Donno (CMCC), Bjorn Backeberg (DELTARES), Alexander Jacoc (EURAC), Michele Claus (EURAC), Christian Pagè (CERFACS)
V0.3	08/11/2023	Introduction, executive summary, and conclusions included. Report sent to reviewers	Sandro Fiore (UNITN)
V0.4	22/11/2023	Review deliverable	Estíbaliz Parcero Iglesias (UPV) Kostas Koumantaros (GRNET)
V0.5	29/11/2023	Version Ready for QA	Charis Chatzikyriakou (EODC)
V1.0	29/11/2023	Final	

Terminology / Acronyms	
Term/Acronym	Definition
AI	Artificial Intelligence
API	Application Programming Interface
CMIP6	Coupled Model Intercomparison Project, 6th phase
CNN	Convolutional Neural Network
DestinE	Destination Earth
DT	Digital Twin
DTE	Digital Twin Engine
ECMWF	European Centre for Medium-Range Weather Forecasts
EO	Earth Observation
ERA5	Fifth generation ECMWF reanalysis for the global climate and weather
FESOM	Finite-Element/volumE Sea ice-Ocean Model
FIAT	Fast Impact Assessment Tool
GNN	Graph Neural Network
GPU	Graphics Processing Unit
IBTrACS	International Best Track Archive for Climate Stewardship
icclim	Index Calculation for CLIMate
JSON	JavaScript Object Notation



D7.3 First version of the thematic modules for the environment domain

ML	Machine Learning
SFINCS	Super-Fast INundation of CoastS
STAC	SpatioTemporal Asset Catalog
TC	Tropical Cyclone
VGG	Visual Geometry Group

Terminology / Acronyms: <https://confluence.egi.eu/display/EGIG>



Table of Contents

1	9
1.1	9
1.2	9
2	10
2.1	10
2.2	11
2.3	12
3	14
3.1	14
3.1.1	15
3.1.2	15
3.2	15
3.2.1	16
3.2.2	16
3.3	16
3.3.1	17
3.3.2	17
3.4	17
3.4.1	18
3.4.2	18
3.5	18
3.5.1	19
3.5.2	19
3.6	20
3.6.1	20
3.6.2	20
3.7	21
3.7.1	21
3.7.2	21
3.8	22
3.8.1	22
3.8.2	22
3.9	22
3.9.1	23
3.9.2	23
3.10	23
3.10.1	24
3.10.2	24



D7.3 First version of the thematic modules for the environment domain

3.11	24
3.11.1	25
3.11.2	25
3.12	26
3.12.1	27
3.12.2	27
3.13	27
3.13.1	28
3.13.2	28
3.14	29
3.14.1	30
3.14.2	30
3.15	31
3.15.1	32
3.15.2	32
4	33
5	34
6	35



Executive summary

This report describes the status of development of thematic modules that have been identified at this stage of the project to support multiple Digital Twins (DTs) in the environmental domain.

It addresses the Digital Twins thematic modules activities foreseen in T7.4 Climate analytics and data processing, T7.5 Earth Observation Modelling and Processing, and T7.6 Hydrological model data processing.

The deliverable provides a total of 15 thematic modules which are described according to a common template highlighting, among others, description, value proposition, licence, user & technical documentation, responsible, and code repository; release notes and future work are also provided.

A summary of the integration status regarding the Digital Twins applications foreseen in WP4 is also provided at the end of this document.

1 Introduction

1.1 Scope

This deliverable summarises the status of development of the 15 thematic modules relevant to implement DTs across multiple domains.

Among these domains, there are:

- Climate analytics and data processing (Task 7.4)
- Earth Observation Modelling and Processing (Task 7.5)
- Hydrological model data processing (Task 7.6)

1.2 Document Structure

In [Section 2](#) we introduce the list of thematic modules under development. In [Section 3](#) we provide details on their functionalities, value, licence, documentation; together with the description and other details, each module also has its own Release Notes subsection with the information about the current status of development and the Future Plans subsection. All the software is released as open-source code and available on git/Github. In [Section 4](#) we report about the integration status with respect to the Digital Twins activity performed in Work Package 4.

2 Thematic Modules Architecture for the environmental domain

2.1 T7.4 Climate analytics and data processing

Task 7.4 aims to develop thematic software modules to support climate data processing, in particular for the DT application for extreme weather events on climate projection data (e.g., CMIP6 and potentially DestinE data). Such thematic modules will support capabilities regarding data preparation and augmentation, data-driven approaches for extreme events prediction, post-processing, multi-model analysis, extreme events characterization, compound event analysis, etc. The modules can be used to support DTs in the context of WP4, in particular for use cases developed in tasks T4.5, T4.6 and T4.7:

- Digital Twins from T4.5:
 - a. **Wildfire danger prediction on climate projections:** this DT will focus on the generation of wildfire danger maps (e.g., burned areas maps) on future climate projections;
 - b. **Tropical cyclones detection and tracking on climate projections:** this DT will support the analysis of tropical cyclones on future projection data in order to evaluate how climate change affects their frequency and duration. The flexibility of this DT will allow to choose between a TensorFlow/Keras implementation and a PyTorch one, respectively working with CNN and GNN models;
 - c. **Eddies prediction on unstructured meshes:** this DT will allow the analysis and detection of oceanic mesoscale eddies. The neural model will work directly with unstructured FESOM2 grids, and once trained it will be able to work with graph-like data and distinguish between cyclonic and anticyclonic eddies in the ocean. It will be a useful tool for the stakeholder that wants to detect these phenomena faster and without the need for prior interpolations to regular grids. It will also allow to study the possible interactions between oceanic currents and atmospheric events like tropical cyclones;
- Digital Twin from T4.6:
 1. **Drought early warning system for the Alps:** this DT will focus on the generation of drought warnings at the river basin scale. The prototype is developed and integrated into openEO as a user interface for researchers and decision makers. Through openEO the user should be able to:
 - 1.1 Run a trained model for a specific area of interest and temporal extent using OpenEO.
 - 1.2 Validate results using historical observations.



D7.3 First version of the thematic modules for the environment domain

1.3 Run the model driven by seasonal forecasts to identify areas affected by hydrological drought.

- Digital Twin from T4.7
 - a. **Climate Change Impacts of Extreme Events:** this DT will support the analysis of future climate simulations in order to evaluate how the characteristics of climate extremes will change in the future, in order to assess their impact by users of the DT through what-if scenarios. The DT will implement a generic method to quantify the changes of the occurrence, the intensity, the spatial coverage, and the duration of climate extremes in specific regions and time periods of the future climate.

The set of thematic modules defined in deliverable D7.1 [R1] has been expanded with additional modules including data-drive models for detection and prediction, as well as compound events analysis. Moreover, some of the thematic modules have been reorganised to better fit new DTs requirements. In particular, the set of capabilities identified in the first deliverable (described in the following list) are actually provided by three different modules addressing different tasks from an application perspective (storms, eddies, and fires). Such modules, from a logical point of view, provide libraries of capabilities to support:

- Data gathering from various data sources.
- Pre-processing functionalities.
- Augmentation and feature scaling (when required).
- Configurable data-driven models (DNN).
- Results post-processing (when required).

Overall, the following thematic modules are being developed in T7.4 and will be reported on in more details in [subsection 3.1](#):

- Thematic modules for tropical cyclones-related data (*ML TC detection*).
- Thematic modules for wildfires-related data (*ML4Fires*).
- Thematic modules for eddies-related data (*eddiesGNN*).
- Generic detection and characterization of climate extreme changes and impacts in the future climate projections (*xtclim*).
- Downscaling Climate Data (*downscaleML*).
- Detection of time or periods of emergence for compound events (*CompEvPoEToE*).

2.2 T7.5 Earth Observation Modelling and Processing

Task 7.5 aims to develop thematic software modules to support the Earth Observation data Modelling and Processing blocks to run Digital Twins based on EO data, with openEO as the driving technology. These modules can be used to support DTs in the context of WP4, in particular for use cases developed in tasks T4.6, the Early Warning for Extreme Events (floods & droughts). In deliverable D7.1, section 2.2.1, the C4 model explains the interaction between the modules for both flood and drought scenarios. Moreover, in



D7.3 First version of the thematic modules for the environment domain

deliverable D7.1, section 2.2.2, the foreseen thematic modules to be developed within T7.5 have been described in detail.

The thematic modules are reported on in more detail in [subsections 3.7, 3.8](#) and [3.9](#):

- Thematic module to run openEO process graphs. The key components for parsing and running an openEO workflow consist in the two following Python projects:
 - *openeo-processes-dask*
 - *openeo-pg-parser-networkx*.
- Thematic module to Index raster data in openEO:
 - *raster-to-stac*.

2.3 T7.6 Hydrological model data processing

Task 7.6 aims to develop thematic software modules to support data processing for hydrological models. The focus is on developing the necessary modules to facilitate near-automatic setting up of local flood hazard and impact models anywhere on Earth, including Earth Observation data processing pipelines. The software modules will support two DT applications:

1. DT for flood early warning in coastal and inland regions, which will focus on the generation of flood risk maps that can be used to trigger early warning alerts when a flood is predicted.
2. DT for flood climate impact in coastal and inland regions, which will focus on the generation of flood maps and quantifying impacts on buildings, utilities, roads, and accessibility under future climate conditions.

The following thematic modules will be reported on in more detail in [Section 3](#):

- **FloodAdapt**: A software package Support System which can be used to assess the benefits and costs of Flood Resilience measures in a community. It uses SFINCS (Super-Fast INundation of CoastS) and Delft-FIAT as calculation cores.
- **HydroMT**: An open-source Python package that facilitates the process of building and analysing spatial geoscientific models with a focus on water system models.
 - **HydroMT-SFINCS**: A plugin that provides an implementation of the model API for the SFINCS model.
 - **HydroMT-FIAT**: A plugin which can be used to build a Delft-FIAT model anywhere in the world.
- **SFINCS**: A reduced-complexity model designed for super-fast modelling of compound flooding events in a dynamic way.
- **Delft-FIAT**: A fast and flexible flood impact assessment tool which combines hazard, vulnerability, and exposure to calculate damages and risk.
- **Hython Wflow_SBM Hydrological Model**: A spatially distributed hydrologic model to estimate hydrological fluxes, including snow accumulation and melt, interception, evapotranspiration, soil moisture, streamflow and actual evapotranspiration.



D7.3 First version of the thematic modules for the environment domain

The application of the above thematic modules will be demonstrated in Jupyter Notebooks that describe setting up the necessary models, configuring them and running scenarios and events and analysing their output. The Jupyter Notebooks are published here: <https://github.com/interTwin-eu/DT-flood/tree/DemonstrationNotebooks>

Note that these Notebooks depend on Deltares' background IPR, namely FloodAdapt, HydroMT and the SFINCS and FIAT plugins, as well as SFINCS and Delft-FIAT themselves. These tools and models will not be modified in the context of interTwin, hence none of these will be published to the interTwin repository, rather these will be referred to / pulled from their existing repositories.



3 Thematic Modules

3.1 ML TC detection

Component name	Thematic modules for tropical cyclones (TCs)
Description	Provides a set of Python modules for supporting processing and analysis of TC-related data and data-driven models
Value proposition	Address tropical cyclones analysis by providing the tools for gathering and pre-processing data, training ML models, post-process results. A version that uses GNNs instead of CNNs is available and written in PyTorch.
Users of the Component	<ul style="list-style-type: none"> • Developers of DTs • Expert scientists
User Documentation	CMCC: https://github.com/CMCC-Foundation/ml-tropical-cyclones-detection/blob/main/README.md UNITN: https://github.com/HPCI-Lab/TC-GNN/blob/main/README.md
Technical Documentation	CMCC: https://github.com/CMCC-Foundation/ml-tropical-cyclones-detection/blob/main/README.md UNITN: https://github.com/HPCI-Lab/TC-GNN/blob/main/README.md
Responsible	CMCC (Davide Donno) and UNITN (Massimiliano Fronza)
Licence	GPLv3
Source code	CMCC: https://github.com/CMCC-Foundation/ml-tropical-cyclones-detection UNITN: https://github.com/HPCI-Lab/TC-GNN



3.1.1 Release notes

The development of the Python packages is currently underway. In this version a set of data-driven models based on VGG [R2] have been developed and tested. An ML ensemble of VGG neural networks has also been implemented as well as an alternative version of the pipeline that makes use of GNNs rather than CNNs.

The library includes a set of features for preparing the data for training the model including: capabilities for splitting the input gridded data into non-overlapping patches, patches selection and storage as TFRecords. Moreover, the code includes the trainable model for TC detection.

3.1.2 Future plans

Next versions will also include features for gathering and managing climate change future projection data like from CMIP6 archive. However, since such data will be needed also by the wildfires and eddies-related ML models, the gathering capabilities could be developed as a separate thematic module. A multi-model thematic module will be also developed to support ensemble detection over multiple CMIP6 data. Finally, we will explore the possibility of implementing a data-driven approach for TC tracking.

3.2 ML4Fires

Component name	ML4Fires
Description	Provides a set of Python modules for supporting processing and analysis of wildfires-related data
Value proposition	Address wildfires analysis and prediction providing tools allowing users to pre-process data, choose model architecture and train the model, post-process results (basic visualisation).
Users of the Component	<ul style="list-style-type: none">• Developers of DTs• Expert scientists
User Documentation	CMCC: https://github.com/CMCC-Foundation/ML4Fires/blob/main/README.md
Technical Documentation	CMCC: https://github.com/CMCC-Foundation/ML4Fires/blob/main/README.md



D7.3 First version of the thematic modules for the environment domain

Responsible	CMCC (Emanuele Donno)
Licence	Apache v2.0
Source code	https://github.com/CMCC-Foundation/ML4Fires

3.2.1 Release notes

The development of the Python package is currently underway. In this first version, a Machine Learning pipeline has been developed for running the different thematic modules. The capabilities include subsetting SeasfireCube data [R3] in Training, Validation and Testing, after a preliminary selection of drivers and target variables. The ML model used for this version is a standard UNet++ network [R4]; users can customise the network depth. The module includes model training and data visualisation utilities.

3.2.2 Future plans

Train, validate and test a ML model for the fire ignition prediction by selecting the best hyperparameter configuration, including potentially different ML architectures. Validate the resulting model on CMIP6 reanalysis data and build inference pipelines with different CMIP6 scenarios. Support for multi-model (on CMIP6 data) ensembles will also be evaluated.

3.3 eddiesGNN

Component name	eddiesGNN
Description	Provides a set of Python modules for supporting processing and analysis of eddy-related data
Value proposition	Address oceanic mesoscale eddies analysis by providing the tools for pre-processing of FESOM2 data and training GNN models
Users of the Component	<ul style="list-style-type: none">• Developers of DTs• Expert scientists
User Documentation	https://github.com/HPCI-Lab/eddiesGNN/blob/main/README.md
Technical Documentation	https://github.com/HPCI-Lab/eddiesGNN/blob/main/README.md
Responsible	UNITN Contact Point: Massimiliano Fronza



D7.3 First version of the thematic modules for the environment domain

Licence	GPLv3
Source code	https://github.com/HPCI-Lab/eddiesGNN

3.3.1 Release notes

The development of the Python packages is currently underway. In this version a pre-processing pipeline has been developed, with both a complete notebook and a demonstration one. They allow interpolations from unstructured to regular grids and vice versa to prepare the dataset for the training. The actual training has both a notebook for demonstration with plots and a Python script to be run with Slurm on the chosen cluster. Both use the PyTorch Geometric framework to run the deep learning process, and both are natively built to make use of GPUs, if the underlying system offers them. Custom loss functions and a Graph U-Net model were customised for this task.

3.3.2 Future plans

Next versions will improve the prediction capabilities of the trained network. Improvements will include a data augmentation step for the generation of sub-regions to be used to improve the generalisation capabilities of the GNN, as well as a final post-processing step to test the network on a bigger region. Other improvements aim at improving the prediction results. Unfortunately, it's not possible to use future projection data like CMIP6 due to its resolution not being high enough for the detection of oceanic mesoscale eddies. However, it will be possible to have a tool that can produce segmentation masks directly on unstructured climate grids and do it faster than the classic physics-based algorithms that only work on regular matrix data. A multi-model (both for data and ML models) thematic module could be also explored to support ensemble detection approaches.

3.4 xtclim

Component name	xtclim
Description	xtclim is a python package implementing an unsupervised Deep Learning method, a Convolutional Variational Auto-Encoder (CVAE) that can characterise generic climate extreme events
Value proposition	Base methods and functions to provide the extraction of generic characteristics of climate extremes. It will enable users to explore the impacts on extreme events on specific users' applications in the context of selected climate simulations.
Users of the Component	<ul style="list-style-type: none">• Developers of DTs



D7.3 First version of the thematic modules for the environment domain

	<ul style="list-style-type: none">• Expert scientists
User Documentation	https://github.com/cerfacs-globc/xtclim/blob/master/README.md
Technical Documentation	https://github.com/cerfacs-globc/xtclim/blob/master/README.md
Responsible	CERFACS (Anne Durif, Christian Pagé)
Licence	Apache 2
Source code	https://github.com/cerfacs-globc/xtclim

3.4.1 Release notes

The development of the xtclim Python package is currently under heavy development. The level of the software can be considered alpha, as many aspects are still under development.

Currently the xtclim package can identify generic combined climate extremes of temperature, wind, and precipitation for four seasons using a Convolutional Variational Auto-Encoder (CVAE) using pyTorch, on global climate simulations over a selected geographical region.

The main workflow consists of data pre-processing to tailor data appropriately for the CVAE method, training, then applying the method. End-user products are then generated at the last step of the workflow such as plots and graphs.

3.4.2 Future plans

First Release will happen no later than the end of the first quarter of 2024.

- Assess required training sample size.
- Validate against iclim results (analytical method).
- Fix issues related to the land-sea mask.
- Implement spatial extent characteristic retrieval.
- Implement time duration characteristic retrieval.
- Assess the method performance (computing time, memory needed).
- Assess different climate variables.

3.5 downscaleML: Downscaling Climate Data

Component name	downscaleML: Downscaling Climate Data
Description	downScaleML is an open-source Python package, designed to streamline the process of climate data downscaling using machine learning techniques. It offers an



D7.3 First version of the thematic modules for the environment domain

	automated workflow tailored for downscaling seasonal forecast climate variables, specifically temperature and precipitation, with a particular emphasis on addressing climate extremes.
Value proposition	It eases forecast data preprocessing, statistical downscaling through a selection of machine learning techniques and result validation. It provides a flexible module for any modelling scheme requiring tailored climate inputs, enables scalability and applicability to other domains, resolutions, and datasets.
Users of the Component	<ul style="list-style-type: none">• Expert users
User Documentation	https://github.com/interTwin-eu/downScaleML#readme
Technical Documentation	https://github.com/interTwin-eu/downScaleML#readme
Responsible	EURAC (Alice Crespi, Suriyah Dhinakaran)
Licence	GNU GPL v3
Source code	https://github.com/interTwin-eu/downScaleML

3.5.1 Release notes

The development of the downScaleML Python package is currently underway, and it is being piloted on a virtual machine hosted at Eurac Research.

The current version of the downScaleML package comprises two distinct workflows: one for preprocessing the input data and another for performing downscaling. In the current developmental stage, the downscaling component identifies the appropriate model and parameters through a grid-search process.

3.5.2 Future plans

For upcoming package updates, our objectives include:

- Implementing the most effective downscaling machine learning models to process real-time seasonal forecast data from ECMWF, with a focus on downscaling climate variables, particularly temperature and precipitation, while emphasising climate extremes.



D7.3 First version of the thematic modules for the environment domain

- Containerizing the entire package environment using Docker and establishing a streamlined pipeline for managing input and output data through openEO processes.

3.6 CompEvPoEToE

Component name	CompEVPoEToE
Description	Provides a set of R functions for determining if periods of emergence (PoE) and/or time of emergence (ToE) of compound events probabilities have emerged in data
Value proposition	This module allows to statistically model if and how compound events have significantly evolved through time, based on reanalysis or observational data. The definition of the compound (i.e., involved variables) is made by the user.
Users of the Component	<ul style="list-style-type: none">• Expert scientists
User Documentation	Not available yet
Technical Documentation	Not available yet
Responsible	CNRS (J. Schmutz, M. Vrac, G. Levavasseur)
Licence	CeCill-C
Source code	https://gitlab.in2p3.fr/ipsi/espri/espri-mod/intertwin

3.6.1 Release notes

The development of the R package is currently underway. In this version, the statistical protocol has been well-defined, developed and tested. A restricted set of marginal distribution functions and copula families are provided.

3.6.2 Future plans

Next versions should include a larger set of marginal distribution functions and copula families. It will also include features for specifically managing climate change future



D7.3 First version of the thematic modules for the environment domain

projection data like from CMIP6 archive, e.g., for direct comparison to observations or reanalysis data. Documentation will be provided as first priority action.

3.7 openeo-processes-dask

Component name	openeo-processes-dask
Description	Python implementation of openEO processes.
Value proposition	Base component necessary to run openEO process graphs. All the processes are implemented using Dask, making them easily scalable and parallelizable.
Users of the Component	<ul style="list-style-type: none">• Expert users• Flood and drought modellers
User Documentation	https://open-eo.github.io/openeo-python-client/cookbook/localprocessing.html
Technical Documentation	https://github.com/Open-EO/openeo-processes-dask
Responsible	EODC (Lukas Weidenholzer, Valentina Hutter) and EURAC (Michele Claus)
Licence	Apache 2.0
Source code	https://github.com/Open-EO/openeo-processes-dask

3.7.1 Release notes

The xArray/Dask implementation of the openEO processes is already deployed in the EODC openEO back-end (part of the DTE), available via openEO Platform (openeo.cloud) or directly (openeo.eodc.eu/openeo/1.1.0). The same code is also used for the Client-Side Processing functionality, allowing to prototype an openEO workflow in a local machine with the same Python syntax (<https://open-eo.github.io/openeo-python-client/cookbook/localprocessing.html>).

3.7.2 Future plans

More processes are currently under development and will be released soon:

- mask
- mask_polygon
- filter_spatial



D7.3 First version of the thematic modules for the environment domain

- aggregate_spatial.

3.8 openeo-pg-parser-networkx

Component name	openeo-pg-parser-networkx
Description	Parse OpenEO process graphs from JSON to traversable Python objects.
Value proposition	Base component necessary to parse openEO process graphs, before calling openeo-processes-dask.
Users of the Component	<ul style="list-style-type: none">• Expert users
User Documentation	https://github.com/Open-EO/openeo-pg-parser-networkx/blob/main/README.md
Technical Documentation	https://github.com/Open-EO/openeo-pg-parser-networkx/blob/main/README.md
Responsible	EODC (Lukas Weidenholzer, Valentina Hutter) and EURAC (Michele Claus)
Licence	Apache 2.0
Source code	https://github.com/Open-EO/openeo-pg-parser-networkx

3.8.1 Release notes

Similarly, with the openeo-processes-dask component, the openeo-pg-parser-networkx is already in use in the EODC back-end and in the Client-Side Processing.

3.8.2 Future plans

Maintain the functionality and adapt to new API changes. If new openEO processes are created and released, we will make sure to support them, parsing the I/O and parameters correctly.

3.9 raster-to-stac

Component name	raster-to-stac
Description	Create STAC metadata for raster datasets.



D7.3 First version of the thematic modules for the environment domain

Value proposition	Makes a resulting dataset easily accessible, interoperable, and shareable.
Users of the Component	<ul style="list-style-type: none"> Expert users
User Documentation	https://gitlab.inf.unibz.it/earth_observation_public/raster-to-stac/-/blob/main/README.md?ref_type=heads
Technical Documentation	https://gitlab.inf.unibz.it/earth_observation_public/raster-to-stac/-/blob/main/README.md?ref_type=heads
Responsible	Michele Claus
Licence	MIT
Source code	https://gitlab.inf.unibz.it/earth_observation_public/raster-to-stac/

3.9.1 Release notes

The component is still work in progress. The first release will allow the creation of a STAC Collection starting from an xArray object.

3.9.2 Future plans

Multiple features are planned to be included:

- Writing to S3 bucket.
- openEO integration in the save_result process.
- More file formats allowed.

3.10 FloodAdapt

Component name	FloodAdapt
Description	A software package Support System which can be used to assess the benefits and costs of Flood Resilience measures in a community. It uses SFINCS and FIAT-Objects in the background.
Value proposition	FloodAdapt is a decision-support tool that seeks to advance and accelerate flooding-related adaptation planning. It brings rapid, physics-based compound flood modelling and detailed impact modelling into an easy-to-use system, allowing non-



D7.3 First version of the thematic modules for the environment domain

	expert end-users to evaluate a wide variety of compound events, future conditions, and adaptation options in minutes. FloodAdapt serves as a connector between scientific advances and practitioner needs, improving and increasing the uptake and impact of adaptation research and development.
Users of the Component	<ul style="list-style-type: none">• Non-expert end-users• Decision makers• Planners
User Documentation	https://www.deltares.nl/en/software-and-data/products/floodadapt
Technical Documentation	https://github.com/Deltares/FloodAdapt#readme
Responsible	Kathryn Roscoe
Licence	MIT
Source code	https://github.com/Deltares/FloodAdapt

3.10.1 Release notes

The deployment of the FloodAdapt backend is achieved via Jupyter Notebooks developed in interTwin (<https://github.com/interTwin-eu/DT-flood/tree/DemonstrationNotebooks>).

A pilot implementation is being run on a Virtual Machine at DESY where the Notebooks are being adapted to leverage the pilot interTwin datalake.

This release offers examples of how to set up and run SFINCS and Delft-FIAT.

3.10.2 Future plans

In the next release examples of how to include a hydrological model (Wflow) and the Resilience Assessment and Adaptation for Critical infrastructureE – model (RA2CE) will be included.

3.11 HydroMT-SFINCS

Component name	HydroMT-SFINCS
Description	HydroMT (Hydro Model Tools) is an open-source Python package that facilitates the process of building and analysing spatial



D7.3 First version of the thematic modules for the environment domain

	geoscientific models with a focus on water system models. It does so by automating the workflow to go from raw data to a complete model instance which is ready to run and to analyse model results once the simulation has finished. This plugin provides an implementation of the model API for the SFINCS model.
Value proposition	Easily build and update the SFINCS model with a single line of code.
Users of the Component	<ul style="list-style-type: none">• Expert users• Flood modellers
User Documentation	https://deltares.github.io/hydromt_sfincs/latest/index.html
Technical Documentation	https://deltares.github.io/hydromt_sfincs/latest/getting_started/intro
Responsible	Roel de Goede
Licence	GNU GPL v3
Source code	https://github.com/Deltares/hydromt_sfincs

3.11.1 Release notes

The application of HydroMT-SFINCS is demonstrated in Jupyter Notebooks developed in interTwin (<https://github.com/interTwin-eu/DT-flood/tree/DemonstrationNotebooks>).

A pilot implementation is being run on a Virtual Machine at DESY where the Notebooks are being adapted to leverage the pilot interTwin datalake.

This release offers examples of how to set up and run SFINCS and Delft-FIAT. The HydroMT-SFINCS plugin offers capabilities to easily set up and run SFINCS using globally available data.

Users can select flood walls, levees, pumps and culverts as flood mitigation and adaptation measures.

3.11.2 Future plans

In the next release capabilities will be created so that a user can

- Easily run long hindcast and ensemble simulations and produce probabilistic flood maps using the flood inundation model (SFINCS).
- Easily run the flood inundation model (SFINCS) using locally available data.
- Select urban green infrastructure as a flood mitigation and adaptation measure.



3.12 HydroMT-FIAT

Component name	HydroMT-FIAT
Description	HydroMT is an open-source python package, developed by Deltares, to build and analyze hydro models. It provides a generic model api with attributes to access the model schematization, (dynamic) forcing data, results, and states. This plugin provides an implementation for the Delft-FIAT model.
Value proposition	<p>With the HydroMT-FIAT plugin, users can easily benefit from the rich set of tools of the HydroMT package to build and update Delft-FIAT models from available global and local data.</p> <p>This plugin assists the FIAT modeller in:</p> <ul style="list-style-type: none"> • Quickly setting up a Delft-FIAT model based on existing hazard maps, global and user-input exposure layers, and a global database of vulnerability curves. • Adjusting and updating components of a FIAT model and their associated parameters in a consistent way, e.g., to test measures that affect the exposure or vulnerability of a FIAT model or to improve an existing FIAT model with better quality data. • Building FIAT models in a reproducible and consistent way.
Users of the Component	<ul style="list-style-type: none"> • Expert users • Flood risk specialists
User Documentation	https://deltares.github.io/hydromt_fiat/latest/index.html
Technical Documentation	https://deltares.github.io/hydromt_fiat/latest/index.html
Responsible	Frederique de Groen



D7.3 First version of the thematic modules for the environment domain

Licence	GNU GPL v3
Source code	https://github.com/Deltares/hydrontfiat/tree/main

3.12.1 Release notes

The application of HydroMT-FIAT is demonstrated in Jupyter Notebooks developed in interTwin (<https://github.com/interTwin-eu/DT-flood/tree/DemonstrationNotebooks>).

A pilot implementation is being run on a Virtual Machine at DESY where the Notebooks are being adapted to leverage the pilot interTwin datalake.

This release offers examples of how to set up and run SFINCS and Delft-FIAT. The HydroMT-FIAT plugin offers capabilities to easily set up and run Delft-FIAT using globally available data.

Users can select raising properties, floodproofing properties and buying out neighbourhoods as flood mitigation and adaptation measures.

3.12.2 Future plans

In the next release capabilities will be created so that a user can

- Estimate damages to buildings and utilities using locally available data
- use Docker / Singularity containers to execute Delft-FIAT on heterogeneous computing infrastructures.

3.13 SFINCS

Component name	SFINCS
Description	SFINCS is a new fast numerical model to simulate 2D compound flooding dynamically for large scale coastal systems, within a fraction of the time required by the Delft3D-1D2D models.
Value proposition	Compound flooding during extreme events can result in tremendous amounts of property damage and loss of life. Early warning systems and multi-hazard risk analysis can reduce these impacts. However, traditional approaches either do not involve relevant physics or are too computationally expensive to do so for large stretches of coastline. The SFINCS model (Super-Fast INundation of CoastS) is



D7.3 First version of the thematic modules for the environment domain

	a new reduced-complexity engine recently developed by Deltares, that is capable of simulating compound flooding including a high computational efficiency balanced with good accuracy.
Users of the Component	<ul style="list-style-type: none">• Expert users• Flood modellers
User Documentation	https://www.deltares.nl/en/software-and-data/products/SFINCS
Technical Documentation	https://sfincs.readthedocs.io/en/latest/
Responsible	Tim Leijnse Roel de Goede
Licence	GNU GPL v3
Source code	https://github.com/Deltares/SFINCS

3.13.1 Release notes

The application of SFINCS is demonstrated in Jupyter Notebooks developed in interTwin (<https://github.com/interTwin-eu/DT-flood/tree/DemonstrationNotebooks>).

A pilot implementation is being run on a Virtual Machine at DESY where the Notebooks are being adapted to leverage the pilot interTwin datalake.

This release offers examples of how to set up and run SFINCS and Delft-FIAT. SFINCS offers capabilities to simulate floodings using globally available data.

Users can specify a geographic area of interest and a minimal number of model-specific parameters, and SFINCS to simulate flood inundation is automatically set up using globally available data.

Globally available static and dynamic input data can be used to create deterministic flood maps based on SFINCS.

A user can specify a simulation period and the flood inundation model (SFINCS) can be run on heterogeneous computing infrastructures using Docker or Singularity containers.

A user can select scenarios to simulate, including mitigation and adaptation measures and the flood inundation model (SFINCS) and the necessary static and dynamic input data are automatically prepared to run the scenario.

3.13.2 Future plans

In the next release capabilities will be created so that:

- Users can specify a geographic area of interest and a minimal number of model-specific parameters, and SFINCS to simulate flood inundation is automatically set up using locally available data.



D7.3 First version of the thematic modules for the environment domain

- Users can easily run the flood inundation model (SFINCS) using locally available data.

3.14 Delft-FIAT

Component name	Delft-FIAT
Description	Delft-FIAT is a fast, free, Python-based tool developed and continuously improved by Deltares to rapidly assess direct economic impacts to buildings, utilities, and roads for user-input flood maps.
Value proposition	<p>Fast impact modelling removes traditional bottlenecks in climate adaptation planning, making it possible to (1) understand the effectiveness of adaptation options and (2) quantify changes in damage and risk as climate and socio-economic conditions change.</p> <p>Fast and automated Delft-FIAT is fast and can be automated. This makes it possible to evaluate future risk caused by changing drivers like growing populations and economies. It also makes it possible to evaluate the effectiveness of interventions by assessing flood damages - now and under changing conditions (a combinations of) interventions, like home elevations, buy-outs, or floodproofing.</p> <p>Flexible Delft-FIAT has a flexible architecture and is data-agnostic. Exposure data can easily be modified, and hazard data - the flood maps - can come from any source. For example, a user may want to try out different depth-damage functions or include a different class of damage than the traditional structure and content damages.</p> <p>Furthermore, any damage type that can be</p>



D7.3 First version of the thematic modules for the environment domain

	<p>described with a depth-damage function can be analysed in Delft-FIAT.</p> <p>Customisable Delft-FIAT is also customisable. It can be connected to a tailored user-interface to make a custom damage modelling tool for less-technical users.</p>
Users of the Component	<ul style="list-style-type: none">• Expert users• Flood risk specialists
User Documentation	https://www.deltares.nl/en/software-and-data/products/delft-fiat-flood-impact-assessment-tool
Technical Documentation	https://github.com/Deltares/Delft-FIAT#readme
Responsible	Frederique de Groen
Licence	GNU GPL v3
Source code	https://github.com/Deltares/Delft-FIAT

3.14.1 Release notes

The application of Delft-FIAT is demonstrated in Jupyter Notebooks developed in interTwin (<https://github.com/interTwin-eu/DT-flood/tree/DemonstrationNotebooks>).

A pilot implementation is being run on a Virtual Machine at DESY where the Notebooks are being adapted to leverage the pilot interTwin datalake.

This release offers examples of how to set up and run SFINCS and Delft-FIAT. Delft-FIAT offers capabilities to quantify the damages resulting from a flood.

Users can specify a geographic area of interest and a minimal number of model-specific parameters and using the output flood maps from SFINCS to estimate damages to buildings and utilities (Delft-FIAT) are automatically set up using globally available data.

Users can configure the input, i.e., the flood map, needed for the flood impact assessment tool (Delft-FIAT) and execute it from a command line interface.

Users can select scenarios to simulate, including mitigation and adaptation measures and the flood inundation model (SFINCS), the flood impact assessment tool (Delft-FIAT) and the necessary static and dynamic input data are automatically prepared to run the scenario.

3.14.2 Future plans

In the next release capabilities will be created so that a user can



D7.3 First version of the thematic modules for the environment domain

- Estimate damages to buildings and utilities using locally available data
- Use Docker / Singularity containers to execute Delf-FIAT on heterogeneous computing infrastructures.

3.15 Hython Wflow_SBM Hydrological Model

Component name	Surrogate Model of WFLOW_SBM:
Description	The trained surrogate model reproduces the physical representation of hydrological processes of the model Wflow, by learning complex physical aspects using spatially distributed data.
Value proposition	The developed model integrated into openEO as a user interface for researchers who can test the performance of a drought early warning system using a trained model and EO data. They can also contribute to further enhancing the trained model. The model is also helpful for public authorities in the field of agriculture and river basin management to identify areas potentially affected by hydrological drought.
Users of the Component	<ol style="list-style-type: none">1. Researchers2. Local/Regional public authorities in the field of agriculture3. hydrology and river basin management authorities4. Journalist for environmental topics with little expertise about technical data
User Documentation	https://github.com/interTwin-eu/hython/blob/main/README.md
Technical Documentation	https://github.com/interTwin-eu/hython/blob/main/workflow_time_series_predictions2.ipynb
Responsible	EURAC (Muhammad Usman Liaqat, Alasawedah Mohammad Hussein)
Licence	CC-BY-4.0 Licence



Source code

<https://github.com/interTwin-eu/hython>

3.15.1 Release Notes

The current version of Hython Python package is currently underway, and it is being under development to examine seasonal forecast of hydrological drought for the alps at river basin scale. The current Hython version is only supporting LSTM. In the current version of the Hython package, a surrogate model of the hydrological model is developed by linking dynamic and static data parameters for two target variables.

3.15.2 Future Work

In updated versions our plan to include following features:

- Train surrogate model other catchments of Alps.
- Calibrate parameters of the hydrological (surrogate) model.
- More preprocessing options: normalization, handling NaNs. Hyperparameter Optimization, Data augmentation, Visualization, Parallel and Distributed ML tasks, Explainable AI.
- Add more timeseries models like GRU, ARIMA.



4 Summary of integration status

The first release of the thematic modules from T7.4 includes a preliminary configuration of the DT applications from WP4 based on the thematic modules [3.1](#), [3.2](#). Initial evaluation of the data-driven models capabilities has been performed (thematic module [3.3](#)). Moreover, a simplified version of the tropical cyclones DT, based on the thematic modules [3.1](#), has been tested on the itwinai AI workflow library¹ developed in T6.5. The integration will be expanded in the next months with the finalisation of the thematic modules ([3.1](#), [3.2](#), [3.3](#)). Furthermore, a sample of the dataset required for the training of the models has been copied to the WP5 data lake for testing data access.

Moreover, an early version of a novel AI-based method implemented in the generic extreme events impact detection DT, xtclim, based on the thematic module [3.4](#), has been successfully evaluated. It is implemented in a jupyter notebook environment. xtclim robustness is currently being assessed and the method is under heavy development in order to improve its detection performance. Additionally, CompEVPoEToE, a module to determine if periods of emergence (PoE) and/or time of emergence (ToE) of compound events probabilities have emerged in data (thematic module [3.6](#)) is under development. An initial version has been implemented and tested.

A first draft of the flood monitoring workflow has been developed with openEO ([3.7](#), [3.8](#)) and tested with the modules from T7.5, available for local prototyping and also on the EODC openEO back-end. Additionally, openEO ties with a drought early warning system where the first version of Hython Wflow_SBM Hydrological Model ([3.15](#)) and Downscaling Climate Data ([3.5](#)) have been developed and available for testing.

Finally, the thematic modules under T7.6 are implemented in jupyter notebooks built to guide a user in setting up the FloodAdapt ([3.10](#)) backend, which currently includes setting up a SFINCS ([3.13](#)) inundation model and a Delft-FIAT ([3.14](#)) impact assessment model using the HydroMT plugins for SFINCS ([3.11](#)) and Delft-FIAT ([3.12](#)). The notebooks are deployed on a VM hosted at DESY in order to adapt the tools (HydroMT-plugins) and models (SFINCS and Delft-FIAT) to leverage available data on the prototype interTwin datalake. Example datasets have been uploaded to the prototype interTwin datalake based on Rucio. Once the datalake is available the notebooks will be adjusted accordingly to read from the data lake. Then we will also be able progress on integration activities and test running the models in docker/singularity containers.

¹ Itwinai: <https://github.com/interTwin-eu/itwinai>



5 Conclusions

This report provides an overview about the first version of the thematic modules for the environmental domain.

It addresses the thematic modules activities foreseen in T7.4 Climate analytics and data processing, T7.5 Earth Observation Modelling and Processing, and T7.6 Hydrological model data processing.

About 15 thematic modules are presented through a common template jointly with some release notes and future work. New modules as well as updated versions of the ones presented in this document will be reported in the deliverable D7.7 Second Release of the thematic modules for the environment domain.

A summary of the preliminary integration status regarding the Digital Twins applications foreseen in WP4 and the thematic modules developed so far is provided at the end of this document.

As a future work, the activity will proceed toward two integration objectives:

- A horizontal integration of the thematic modules into the Digital Twins application will be further addressed in D4.3; yet such integration will be strengthened over the second and third period of the project, by considering the updated requirements (D7.5) and leveraging the outcome of the second release of the thematic modules (D7.7).
- A vertical integration of the thematic modules with the core ones foreseen in WP6 to fully address the Digital Twins requirements.

6 References

Reference	
No	Description / Link
R1	M. Claus et al., (2023). interTwin D7.1 Report on requirements and thematic modules definition for the environment domain (V1 Under EC review). DOI: https://doi.org/10.5281/zenodo.8036991
R2	Karen Simonyan and Andrew Zisserman. 2015. Very Deep Convolutional Networks for Large-Scale Image Recognition. arXiv:1409.1556
R3	Alonso, L., Gans, F., Karasante, I., Ahuja, A., Prapas, I., Kondylatos, S., Papoutsis, I., Panagiotou, E., Mihail, D., Cremer, F., Weber, U., & Carvalhais, N. (2023). SeasFire Cube: A Global Dataset for Seasonal Fire Modeling in the Earth System (0.3) [Data set]. DOI: https://doi.org/10.5281/zenodo.8055879
R4	Zhou, Z., Rahman Siddiquee, M.M., Tajbakhsh, N., Liang, J. (2018). UNet++: A Nested U-Net Architecture for Medical Image Segmentation. In: Stoyanov, D., et al. Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support. DLMIA ML-CDS 2018 2018. Lecture Notes in Computer Science, vol 11045. Springer, Cham. DOI: https://doi.org/10.1007/978-3-030-00889-5_1

