Large-scale EXecution for Industry & Society

Deliverable D6.4

Pilot results: improved scenarios measurements and evaluation

Co-funded by the Horizon 2020 Framework Programme of the European Union
Grant Agreement Number 825532
ICT-11-2018-2019 (IA - Innovation Action)

DELIVERABLE ID | TITLE	D6.4 | Pilot results: improved scenarios measurements and evaluation
RESPONSIBLE AUTHOR	Lorenza Bovio (ITHACA)
WORKPACKAGE ID | TITLE	WP6 | Earthquake and Tsunami large scale pilot
WORKPACKAGE LEADER	CEA
DATE OF DELIVERY (CONTRACTUAL)	31/12/2021 (M36)
DATE OF DELIVERY (SUBMITTED)	22/01/2022
VERSION | STATUS	V1.0 | Final
TYPE OF DELIVERABLE	R (Report)
DISSEMINATION LEVEL	PU (Public)
AUTHORS (PARTNER)	Lorenza Bovio (ITHACA), Andrea Ajmar (ITHACA); Sven Harig (AWI), Natalja Rakowsky (AWI); Tomáš Martinovič (IT4I); Stéphane Louise (CEA); Danijel Schorlemmer (GFZ), Marius Kriegerowski (GFZ)
INTERNAL REVIEW	Martin Golasowski (IT4I); Domokos Sarmany (ECMWF)

Project Coordinator: Dr. Jan Martinovič – IT4Innovations, VSB – Technical University of Ostrava
E-mail: jan.martinovic@vsb.cz, **Phone:** +420 597 329 598, **Web:** https://lexis-project.eu

DOCUMENT VERSION

VERSION	MODIFICATION(S)	DATE	AUTHOR(S)
0.1	Table of content and main contributions by all partners	17/11/2021	Lorenza Bovio, Andrea Ajmar (ITHACA) ; Tomas Martinovic (IT4I); Sven Harig, Natalja Rakowsky (AWI); Stéphane Louise (CEA); Danijel Schorlemmer, Marius Kriegerowski (GFZ)
0.2	First version for review	15/12/2021	Lorenza Bovio , Andrea Ajmar (ITHACA)
0.3	Reviewed	20/12/2021	Martin Golasowski (IT4I); Domokos Sarmany (ECMWF)
0.4	Addressing of the comment from reviewers	29/12/2021	Lorenza Bovio (ITHACA); Sven Harig, Natalja Rakowsky (AWI); Tomáš Martinovič (IT4I); Stéphane Louise (CEA); Danijel Schorlemmer, Marius Kriegerowski (GFZ)
0.5	Final check of the deliverable I	04/01/2022	Tomáš Martinovič (IT4I)
1.0	Final check of the deliverable II Integration updates, added ZENODO link from partners	21/01/2022	Kateřina Slaninová, Jan Martinovič, Tomáš Martinovič (IT4I)

GLOSSARY

ACRONYM	DESCRIPTION
AOI	Area Of Interest
BPMN	Business Process Model Notation: a graphical notation to describe and model business processes.
CEMS	Copernicus Emergency Management Service
DDI	Distributed Data Infrastructure
EQ	Earthquake
FEP	First Estimate Product
GIS	Geographical Information System: a database and additional functions for data with a geographical meaning and coordinates.
MOC	Model of Computation
ODbL	Open Database License
QML	The name of an actor in the MoC producing a QuakeML file

QUAKEML	Quake Markup Language: a flexible, extensible and modular XML representation of seismological data, e.g. epicentre, hypocentre, magnitude.
RISE	Project in Horizon 2020: "Real-time earthquake rIsk reduction for a reSilient Europe"
RM	Rapid Mapping
SDF	Synchronous Data Flow
SEM	Satellite-based Emergency Mapping: the process of using remote sensing data to produce maps for emergencies
TSUNAWI	Tsunami simulation code developed at AWI

TABLE OF PARTNERS

ACRONYM	PARTNER
Avio Aero	GE AVIO SRL
Atos	BULL SAS
AWI	ALFRED WEGENER INSTITUT HELMHOLTZ ZENTRUM FUR POLAR UND MEERESFORSCHUNG
BLABS	BAYNCORE LABS LIMITED
CEA	COMMISSARIAT A L ENERGIE ATOMIQUE ET AUX ENERGIES ALTERNATIVES
CIMA	CENTRO INTERNAZIONALE IN MONITORAGGIO AMBIENTALE - FONDAZIONE CIMA
CYC	CYCLOPS LABS GMBH
ECMWF	EUROPEAN CENTRE FOR MEDIUM-RANGE WEATHER FORECASTS
EURAXENT	MARC DERQUENNES
GFZ	HELMHOLTZ ZENTRUM POTSDAM DEUTSCHESGEOFORSCHUNGSZENTRUM GFZ
ICHEC	NATIONAL UNIVERSITY OF IRELAND GALWAY / Irish Centre for High-End Computing
IT4I	VYSOKA SKOLA BANSKA - TECHNICKA UNIVERZITA OSTRAVA / IT4Innovations National Supercomputing Centre
ITHACA	ASSOCIAZIONE ITHACA
LINKS	FONDAZIONE LINKS / ISTITUTO SUPERIORE MARIO BOELLA ISMB
LRZ	BAYERISCHE AKADEMIE DER WISSENSCHAFTEN / Leibniz Rechenzentrum der BAdW
NUM	NUMTECH
O24	OUTPOST 24 FRANCE
TESEO	TESEO SPA TECNOLOGIE E SISTEMI ELETTRONICI ED OTTICI

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

EXECUTIVE SUMMARY

This deliverable describes the analysis, evolution, and the possible steps forward of the main software components of the earthquake and tsunami large scale pilot of the LEXIS project. This pilot is a use case and exercise for the federation of resources and technologies at the convergence of cloud, big data and high-performance computing that the LEXIS project is building, in order to manage complex workflows with real-time constraints.

Position of the deliverable in the whole project context

The LEXIS project relies on three pilots to validate and deploy its technology and infrastructure advances, giving to each pilot its own work package. Work package 6 (WP6) is dedicated to the Earthquake and tsunami large scale pilot. Following on Task 6.1 of scenarios and baseline requirements definition, this work package is running Task 6.2, a development task where individual components of the pilot are improved to benefit from the technology of the project, and to prepare them for integration on the LEXIS platform. Integration, effectively undertaken in Task 6.3 and Task 6.4, is dedicated to further developments of the pilot's individual components and to the integration of the entire workflow.

Description of the deliverable

This deliverable states the advances from the developments carried out in Deliverable D6.3 [1], and its completion leads to the first fully working workflow. As such, this deliverable also shows the integration of both the entire workflow and the individual components within the working workflow, using the Coquimbo case (Chile). Section 0 describes the evolution of the individual components of the workflow since Deliverable D6.3, while Section 2 describes how the components integrate in the first working workflow.

1 EVOLUTION AND INTEGRATION OF THE WORKFLOW

1.1 OPEN BUILDING MAP AND LOSS ASSESSMENT DATA

The detailed description of the whole workflow is provided in Deliverable D6.3 [1].

1.1.1 Loss Calculator

We have expanded the functionality of the loss calculator to improve its flexibility with three new features:

- New framework for fragility functions,
- Database import format for the loss calculator,
- Result aggregation per tile or building.

We have refactored the use of fragility functions. Previously, these functions were provided as discrete functions in CSV files and they were applied asset-wise by reading the fragility-function data from the file and running the damage-probability computation. With new types of continuous fragility functions in XML, we have redesigned the processing such that there is an internal class structure for both types of functions and the loss calculator is reading any set of functions in first and applying them using the internal representation.

To comply with the input data licenses, we have changed the data input formats. The exposure data used in the pilot are licensed under CC BY-SA 4.0[1], while the building footprints and other building-related information is coming from OpenStreetMap and licensed under ODbL[2]. These two datasets cannot be combined in a way that each building dataset carries the exposure information because distributing these data would violate the contributing-data licences. Therefore, we are storing the input data in a Spatialite[3] database in different tables for the different input data and connecting them via their IDs. This way, the input data can be distributed with different licenses applying to different tables of the database.

As much as building-wise damage calculations are useful for emergency management, risk assessment, and mitigation efforts, they may violate privacy concerns. In order to be able to disseminate our damage assessments, we added a tile-based aggregation of damage data to the output of the loss calculator. For this, we use zoom-level 18 tiles of the Quadtree, roughly 100m x 100m size. Instead of providing probabilities for all given damage states (from no damage to complete destruction) per building, the loss calculator can provide number of buildings per damage state for each tile, see Figure 1.

[1] CC BY-SA 4.0: https://creativecommons.org/licenses/by-sa/4.0/
[2] ODbL: https://opendatacommons.org/licenses/odbl/
[3] Spatialite database: https://www.gaia-gis.it/fossil/libspatialite/index

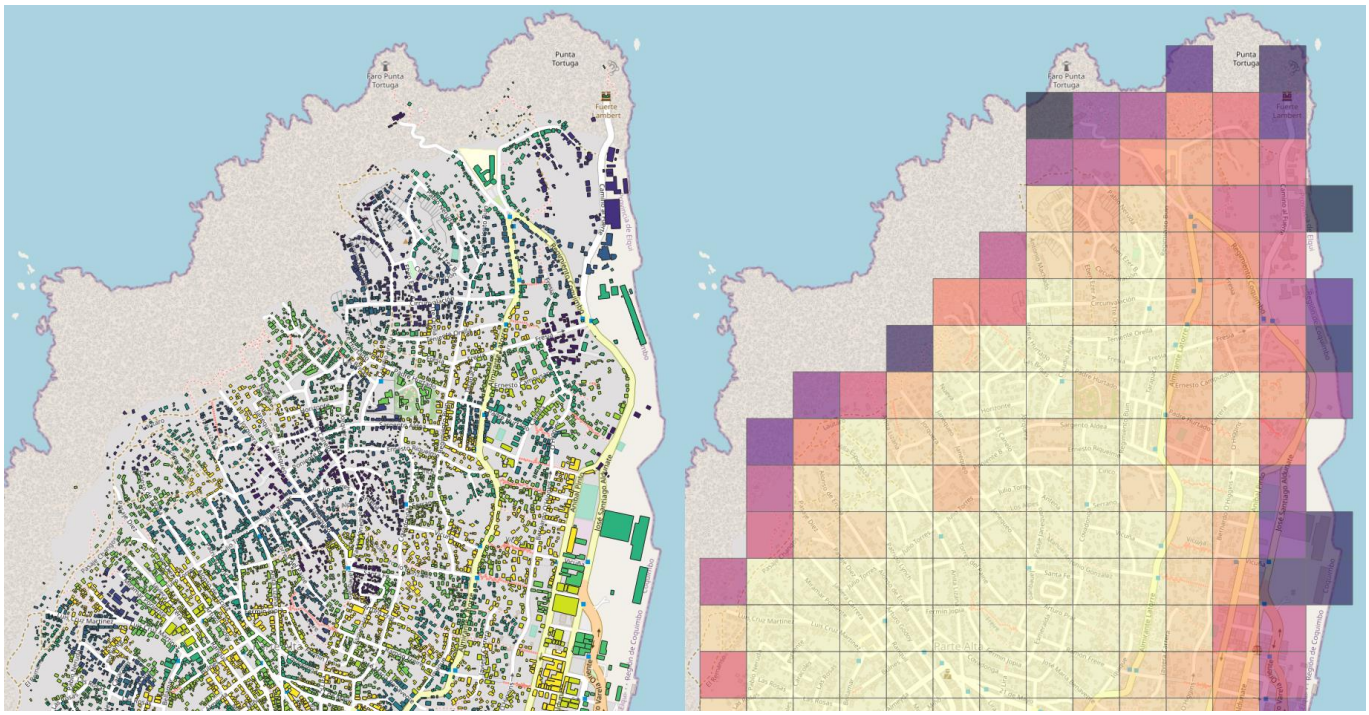

Figure 1 Comparison of building-wise slight-damage probabilities (left) with tile-wise number of slightly damaged buildings (right)

To make visualizations of damage assessment results, we have added a Geopackage[4] output to the loss calculator. Geopackage is an open platform-independent and standard-based format for geospatial data. It is internally based on a Spatialite database and can be considered the successor of Shapefiles. Previously, only classical CSV files were output which makes it more difficult to visualize results on maps in GIS applications. The new Geopackage output, either for building-wise or tile-wise damage computations, allows for quick data inspections in tools like QGIS[5] by simply dragging and dropping the results file.

1.1.2 Shakemap micro-service

The loss calculator harnesses an asynchronous shakemap micro-service that was developed throughout the LEXIS project (https://git.gfz-potsdam.de/marius/shakemapi). It is built around the OpenQuake engine (https://github.com/gem/oq-engine) for seismic hazard and risk analysis. The service was extended to implement additional intensity measure types (IMT) such as peak ground acceleration (PGA), peak ground velocity (PGV) and spectral acceleration (SA). These three IMTs are used by the loss calculator. The shakemap service uses the shear-wave velocity time-averaged down to a depth of 30 m depth (global V_S30) as a proxy for site amplification and allows forward modelling based on double-couple (DC) seismic source parameters defined by strike, dip, rake and earthquake magnitude, given as so-called moment magnitude.

1.2 TSUNAWI

The TsunAWI simulation provides an estimate of the inundation by the tsunami caused by the earthquake. In Deliverable D6.3 [1], we showed the improvements in time to solution by single precision arithmetic and MPI parallelization. In the meantime, we could go into more detail of TsunAWI's runtime behaviour. When regarding the strong scaling of the MPI parallel code on Karolina, the new HPC system at IT4I with AMD Rome CPUs, we observe superlinear scaling for the larger setups. The effect is stronger than on other systems with older CPUs

[4] Geopackage: https://www.ogc.org/standards/geopackage
[5] QGIS: https://qgis.org

(ollie.awi.de[6], Intel Xeon Broadwell) or more power-hungry CPUs (emmy.hlrn.de[7], Intel Cascade Lake) compared to AMD Rome, see Figure 2.

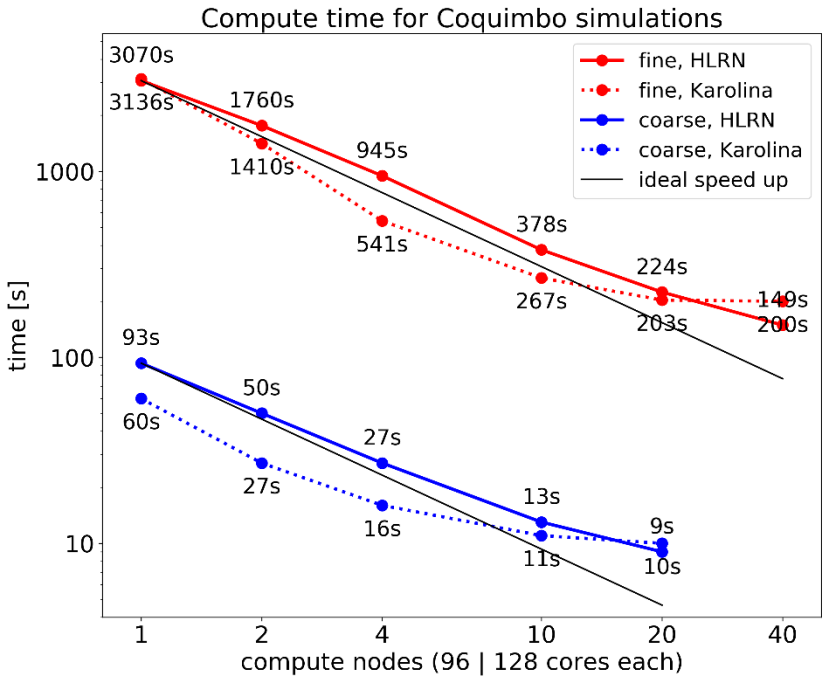

Figure 2 Scaling of the TsunAWI simulations for the Coquimbo testcases with 1.7M vertices (coarse) and 4.9M vertices (fine) on emmy.hlrn.de (Intel Cascade Lake) and karolina.it4i.cz (AMD Rome)

In particular, the coarse Coquimbo setup with 1.7 Million mesh vertices is considerably faster on one node of Karolina than on one node at HLRN, with 60s on 128 AMD Rome cores compared to 93s on 96 Intel Cascade Lake Cores. However, the fine setup with 4.9 Million mesh vertices is even a bit slower on the AMD system. On the other hand, as soon as the MPI domain decomposition distributes the mesh on two compute nodes with 2.4M vertices on each node, the AMD Rome based compute nodes overtake the Intel based ones, see Table 1. Summing up, the reason seems to be that the ratio of compute power and memory bandwidth for the AMD Rome 7H12 is larger than on the older Intel systems. To underline the assumption, we tested with a simple stream benchmark, written in Fortran and compiled with the same options as TsunAWI on the specific systems.

SYSTEM	ARCHITECTURE	PERFORMANCE/CORE	AGGREGATED: PERFORMANCE/NODE
ollie.awi.de	2x 18 Core Intel Boadwell	285 MFlop/s	10 GFlop/s
emmy.hlrn.de	2x 48 Core Intel Cascade Lake	360 MFlop/s	34.5 GFlop/s
karolina.it4i.cz	2x 64 Core AMD Rome	187 MFlop/s	24 GFlop/s

Table 1 Simple stream-like benchmark to assess the effect of the memory bandwidth. Performance of a(:) = a(:)+alpha*b(:) for large vectors, in parallel on one MPI-task per compute core.

[6] High-Performance Computers at AWI: https://www.awi.de/en/science/special-groups/scientific-computing/high-performance-computing/computers/hardware.html
[7] HLRN-IV-System: https://www.hlrn.de/supercomputer/hlrn-iv-system/

For large submeshes, on each compute node the memory bandwidth determines the run time, and soon as the subdomains provided by the domain decomposition for the MPI parallelization fit into the caches, the computation is no longer slowed down by the access to the main memory, finally resulting in superlinear scaling.

As an additional improvement of TsunAWI, we added asynchronuous I/O. For the workflow, the benefit is small, as the main product is the estimate of the inundation which is written after the simulation has finished. However, one is often interested in snapshots of the wave propagation on the fly. The implementation is based on one-sided MPI. The communicator is split and one MPI task is dedicated to collecting and writing the data, while the others continue to compute. With modern many-core architectures, the loss of one compute task is negligible compared to the time spend on writing. The effect, of course, also depends on the speed of the file system.

1.3 SUPPORT TO SATELLITE EMERGENCY MAPPING MECHANISMS

This final task has included the development of a code related to the generation of raster outputs for the prioritization of impacted areas for the eventual satellite tasking operations after an earthquake/tsunami event.

Main data inputs of the code are loss assessment data, tsunami simulations and exposure data mainly related to occupancy status of buildings during night.

From an operational point of view, the proposed code can manage three different scenarios, corresponding to the availability of both loss assessment data and outputs of tsunami simulations. The code execution is triggered by the presence of relevant input files in specific folders.

The presence of a loss assessment file, including complete information of damage level of buildings, in the proper folder, guarantees the triggering of the workflow for an earthquake event. For this scenario, the code generates raster outputs using a grid-based approach: exposure data about population and buildings and probabilities of expected damage information provided by loss assessment at a single building level, are integrated and subsequently aggregate at a map scale (200 m GSD) suitable for regional level analyses. The main products expected by the code are: Relative Potential Damage map, Damage Absolute Probability map, and Night Population map. Relative Potential Damage map provides the indication of expected most affected areas by the combination of different classes of damage probabilities and the exposure information about building occupancy. Damage Absolute Probability map provides the spatial distribution of the absolute values of probability of damage derived considering the most severe damage classes. Night Population map provide an estimation of a population exposed to the event in the examined area, based on the population occupancy during night hours. The additional presence of the outputs of TsunAWI simulation in a different folder fit into the second scenario, related to a tsunami-triggered earthquake. This means that in addition to the previously mentioned products, the code generates two additional raster outputs, which are obtained integrating the same exposure data on population used in the case of the earthquake event and the maximum wave height expected over inundated areas generated by TsunAWI. Also in this case, the procedure is based on a grid approach to aggregate all information at a map scale (200 m GSD) suitable for regional level analyses. The derived additional products are: Potential Damage map and Building Exposure map.

The last scenario is related to the analysis of damages due only to a tsunami event, e.g. a tsunami triggered by a seismic event that did not cause any damage, and for which earthquake loss assessment data are lacking. This scenario is triggered by the presence of outputs of TsunAWI simulation in the appropriate folder. As the exposure information provided at single building level is crucial data to derive the proposed map products the code runs a check on the structural fields of loss assessment database. This scenario is activated once the code detects missing field values for buildings damage.

1.4 INTEGRATION PROCESS

The workflow evolved slightly during the LEXIS project time. In some cases, the initial version was adjusted to the limitation of the frameworks, such as removing explicit cycles from the workflow, and in other the tools in LEXIS

were updated to support needs of the workflow set in the urgent computing, for example implementation of deadlines in the orchestration service.

The current workflow that was integrated in the LEXIS is composed of two parallel branches: one for the TsunAWI computation and one for the damage assessment and shakemap computation, which are then combined into one for the final image product creation. The overview of the workflow is shown in Figure 3. The whole workflow starts with the QuakeML [2] file describing the event[8]. This file is then copied to the TsunAWI job, which is run at an HPC cluster, and to the ShakeMap and Loss assessment dockers for the damage assessment. Results are then copied and mounted to the docker container containing postprocessing scripts which creates several geotiff[9] images containing information about different kinds of potential damage and risks. These then can be ingested by the external geoinformation systems like FLOREON+[10].

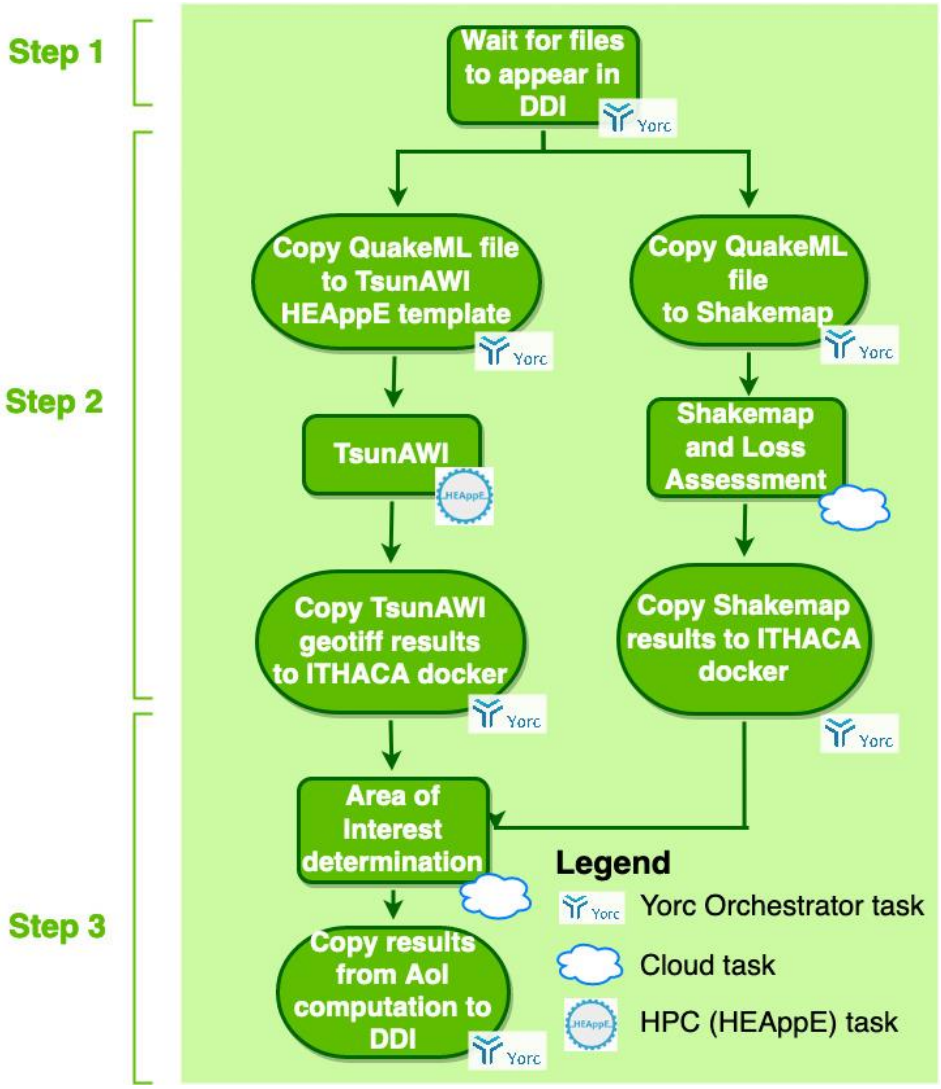

Figure 3 WP6 Workflow diagram

The Figure 4 shows the workflow execution progress in the LEXIS Portal. This view contains more tasks than the high level workflow shown in the Figure 3, because it shows all the tasks done by the YSTIA orchestrator.

To summarize now - this workflow contains one HPC job - TsunAWI tsunami inundation computation which must be deployed on the cluster alongside meshes for the areas which are to be used by the system. This script is now

[8] QuakeML: https://quake.ethz.ch/quakeml
[9] Geotiff: https://github.com/OSGeo/libgeotiff
[10] FLOREON+: https://floreon.eu/mapa/

deployed at IT4I clusters and LRZ cluster as a Generic HEAppE command template and can be run more simultaneously taking into account the results which are computed first. This can be easily extended to more clusters, to increase the resilience to outages and maintenance. On the other side, there are two docker images which are used for the damage assessment shakemap docker and loss assessment docker provided by GFZ. The set of the docker images also needs to have the data about the buildings which are in the affected area. Such data may be uploaded to the LEXIS DDI and updated regularly, or they can be downloaded from the GFZ OBM server. The last piece is the docker image used to post-process results of the previously mentioned components. Since these components are docker images, as long as the given centre has a cloud service working these are easily deployed and managed by the YSTIA orchestrator.

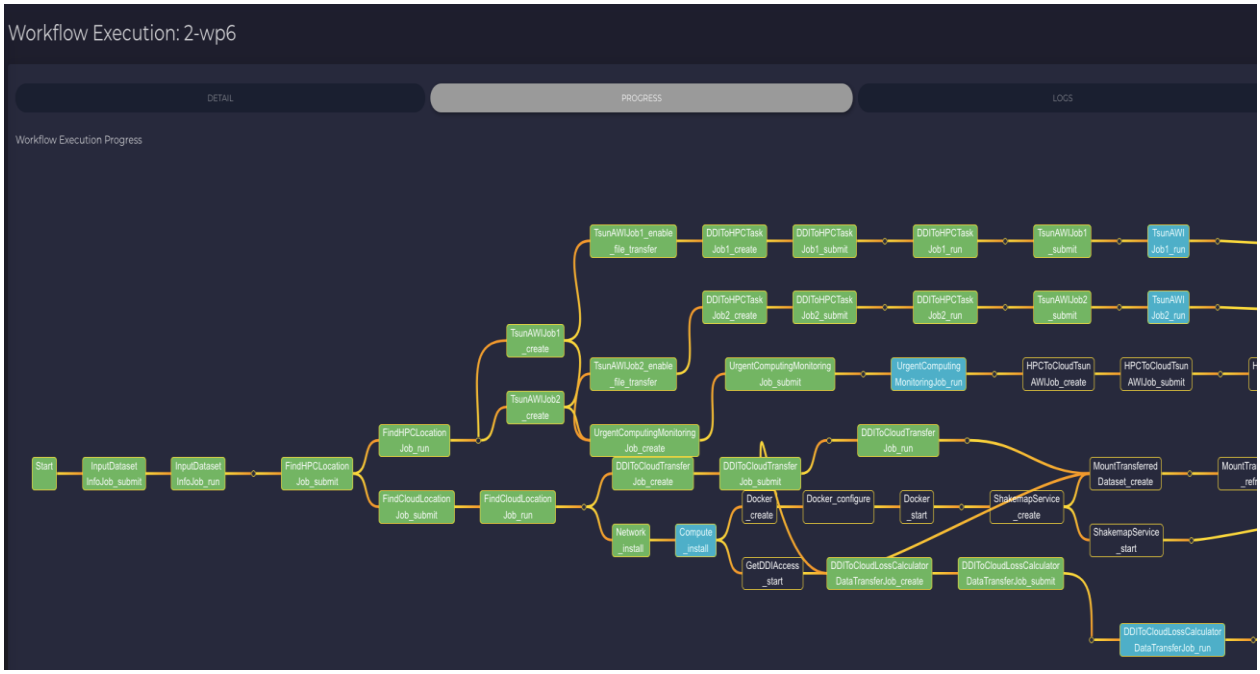

Figure 4 WP6 workflow execution progress in LEXIS Portal

2 FIRST WORKING WORKFLOW ON THE CHILE CASE

2.1 OPEN BUILDING MAP AND LOSS ASSESSMENT DATA

For production-ready and easy setup both the shakemap and the loss-calculator service come with a Dockerfile. This ensures quick, reproduceable and scalable setup of the micro-service, as well as stable integration into the LEXIS workflow. To further simplify the setup, both containerized applications were pre-built and hosted on Docker-Hub[11], an exchange platform for production-ready applications wrapped in docker-containers. The currently hosted containers are only compatible with arm64 architecture systems. As an alternative, the full micro-service system running the demos can be spawned with docker-compose which will build the containers from scratch, locally. For further reference see the lexis-scenario-examples at ZENODO[12].

[11] Docker-Hub: https://hub.docker.com/u/emrius11

[12] LEXIS scenario examples: https://doi.org/10.5281/zenodo.5887012

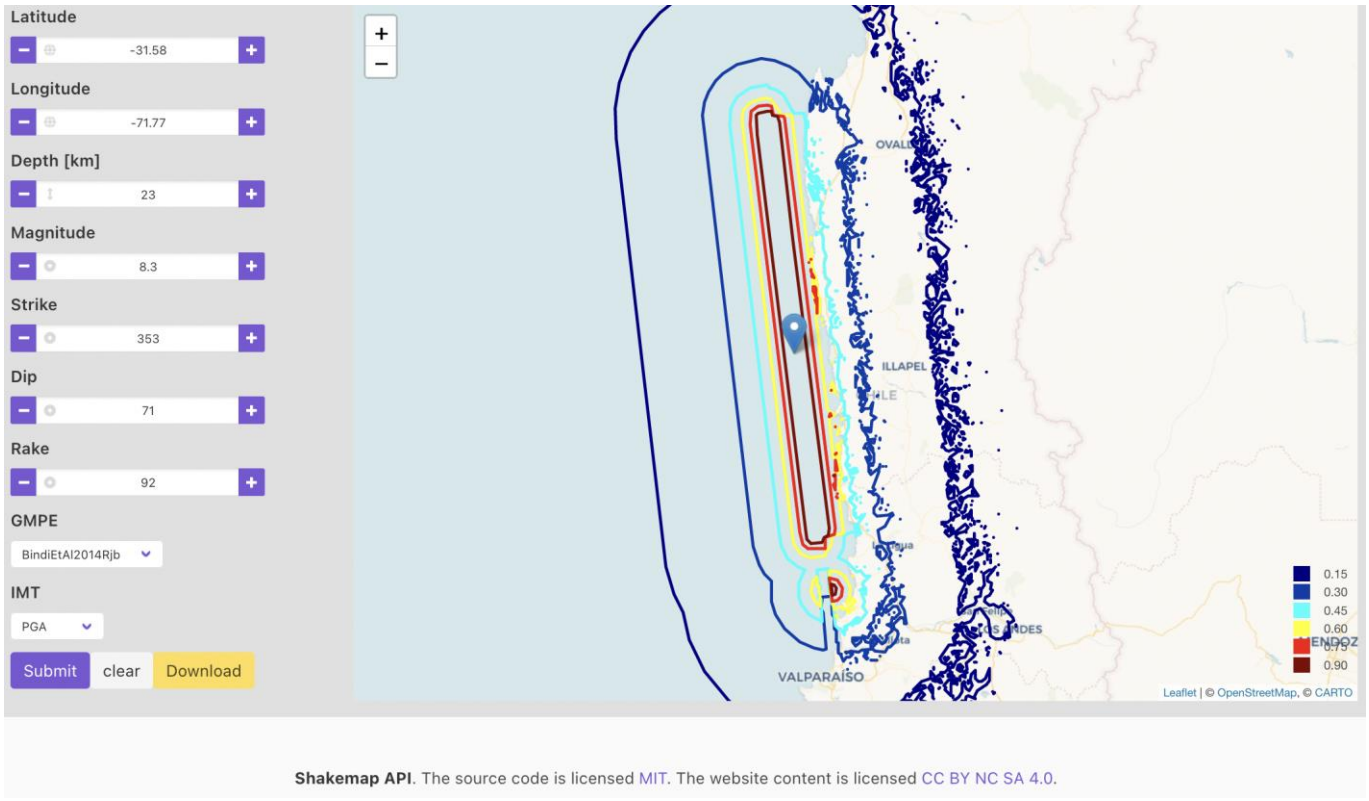

Figure 5 Shakemap of the 2015 M8.3 Coquimbo (Chile) event generated by the developed shakemap micro-service and displayed in the interactive web-frontend.

The shakemap service exposes shakemap generation through an asynchronous REST-full API and includes a web-front-end that allows to interactively query and inspect the generated shakemap (see example in Figure 5).

The loss calculator and the shakemap services can be triggered by posting a seismic source in QuakeML format against each service to trigger computations. Internally, the loss calculator, which is the entry point of this part of the workflow, forwards the source information to the shakemap requesting the ground shaking in different intensity measurement types (IMT). To achieve minimum latencies the loss calculator implements an asynchronous API run by ‚uvloop', a minimum latency asynchronous event loop.

Depending on the complexity, size of the ground shaking area and number of buildings in the targeted region a loss estimate is computed within a few seconds to approximately one minute on a standard desktop computer after triggering the loss calculator by posting an event.

2.2 TSUNAWI

The long compute time of more than 50 minutes of the TsunAWI simulation for the fine Coquimbo on one Karolina node made us reconsider the setup for the LEXIS workflow. The time can be reduced to just 4 minutes 30 seconds on 10 compute nodes. However, to account for warning centres and other potential use cases where hardware is more limited than in the LEXIS framework, a shorter runtime on just one node would be an advantage. We therefore consider two aspects: the simulated timeframe of the tsunami propagation and the mesh size.

Regarding the first issue, the initial setup was to simulate 10 hours of tsunami propagation, which is a valid choice regarding tsunami early warning in general. In real events at the Chilean Coast, rather high crests are observed even hours after the event. It is crucial to get the full picture when deciding to send an all-clear message. However, a model time of just 4 hours is sufficient to cover the highest crests and thus the largest extent of the inundation, which is the only required data product for the LEXIS workflow. This simple reasoning reduces the compute time by 60%. On the other hand, for the coarse 150 m resolution, we had to decrease the time step from 2 s to 1.8 s, as

experiments with stronger magnitudes could crash. To choose an optimal time step is always crucial, as very specific local mesh geometry can require a smaller timestep than usual to ensure stability for a given resolution.

Second, the initial meshes have several foci with high resolution in addition to Coquimbo. This is also a typical choice for general early warning, but for the workflow, it is sufficient to restrict the high resolution to Coquimbo. The tsunami simulations are carried out in different meshes, an initial fast simulation in a rather coarse triangular mesh (resolution 10 km - 100 m) is refined with a second run of the same model in a refined triangulation (resolution 10 km - 20 m). Both simulations are carried out with the model TsunAWI, due to numerical constraints the time step in the fine mesh is considerably smaller than in the coarse run. Therefore, the refined simulations are time consuming, and optimizations of the mesh were carried out to improve the run time. Originally, all meshes for Chile were designed to cover a large part of the Chilean coast with high resolution in several locations of interest (coastal cities, tide gauge locations). Since the pilot area in the present workflow in Chile is Coquimbo bay, we generated an optimized mesh with focus in that specific area only. Additionally, we adjusted the model domain in the coastal part of the pilot area. The mesh resolution in the inundation domain is highest with triangle edge length going down to 20 m and we reduced the mesh size significantly by limiting the mesh extent to the area with higher risk of inundation.

SETUP		COQUIMBO REGION COARSE	COQUIMBO REGION FINE	
		ORIGINAL	ORIGINAL	FOCUS ON COQUIMBO
NUMBER OF VERTICES		1,709,500	4,888,000	1,474,000
RESOLUTION		150 m - 15 km	20 m – 15 km	20 m – 12 km
TIME STEP DT		1.8 s	0.15 s	0.2 s
SIMULATE TSUNAMI TIMEFRAME		4 h	4 h	4 h
COMPUTE TIME FOR TIME STEPPING WITH SINGLE PRECISION ARITHMETIC ON KAROLINA.IT4I.CZ				
1 NODE	128 tasks	26 s	1,200 s	182 s

Table 2 Compute time for the time stepping of the improved model setups with the model time reduced from 10h to 4h and an alternative mesh for the fine resolution case. The original compute times are listed in Deliverable D6.3 [1].

The following Figure 6 shows the resolutions and extents of the meshes in use. Due to the flexible triangular structure the deep sea region is discretized very similarly in all grids since the wave in that part of the domain is well captured even by low resolution of about 10 km. Only the coastal regions differ strongly in all the mesh realizations. The final Figure 7 of the section highlights the difference between coarse and fine mesh structure and corresponding model results for identical events.

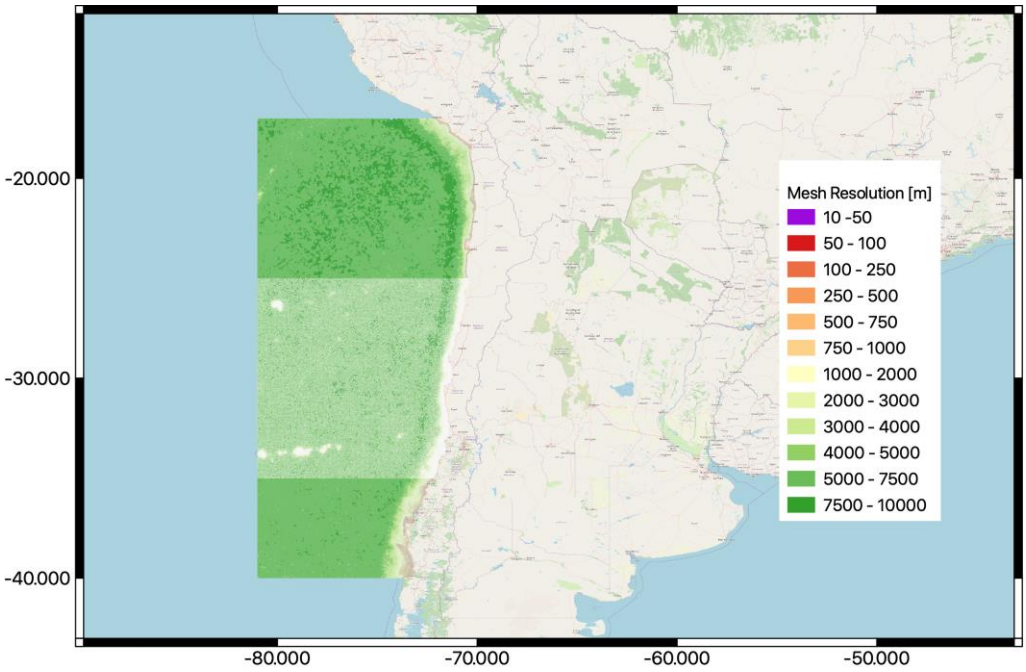

Figure 6 Mesh domain and resolution of the Chilean high resolution meshes available in LEXIS.
The optimized mesh for Coquimbo is marked by white shading. Both triangulations use the same resolution in the deep ocean. However, the original large domain contains several highly resolved areas.

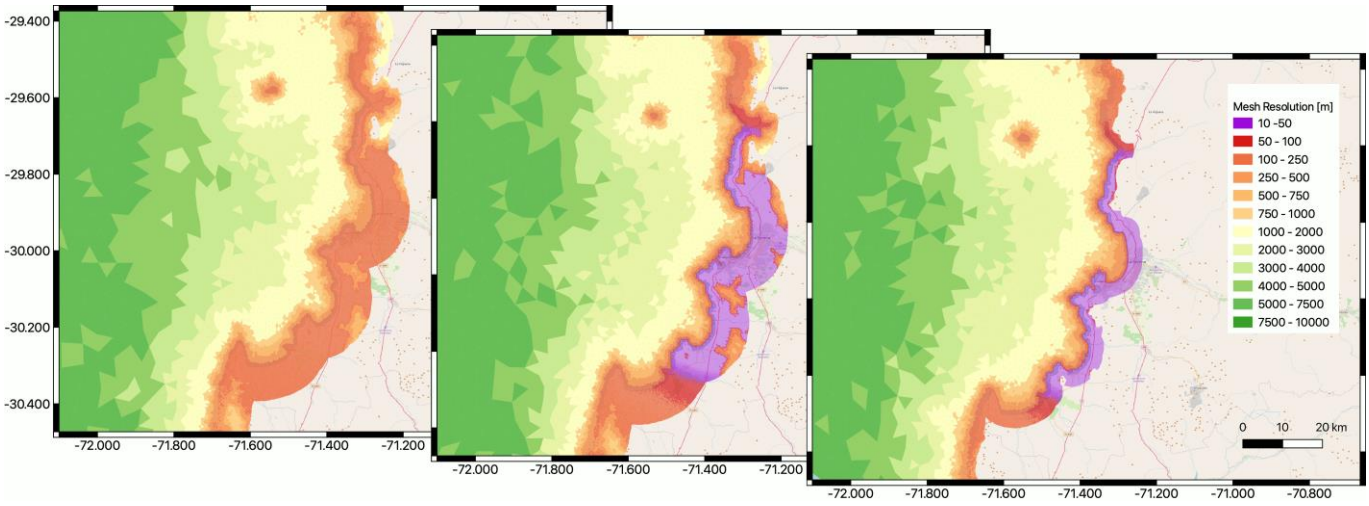

Figure 7 Triangular mesh resolution in the original (left panel) and optimized version in the pilot area (Coquimbo bay).
The finest resolution is the same in both triangulations, however the optimized variant is restricted to the area prone to inundation.

Despite the mesh optimisation, all versions are available at ZENODO[13] and can be used in the workflow. For example, an initial coarse estimate may be refined in the optimised fine mesh also with updated sources, whereas the large-scale fine mesh may still be used for estimates of the impact in remote regions. Inundation estimates in the coarse and in the optimized fine mesh are shown in the following Figure 8.

[13] Tsunami Simulation: TsunAWI Testcase "Coquimbo 2015", Setups and Example Results: DOI:10.5281/zenodo.5799056, https://zenodo.org/record/5799056

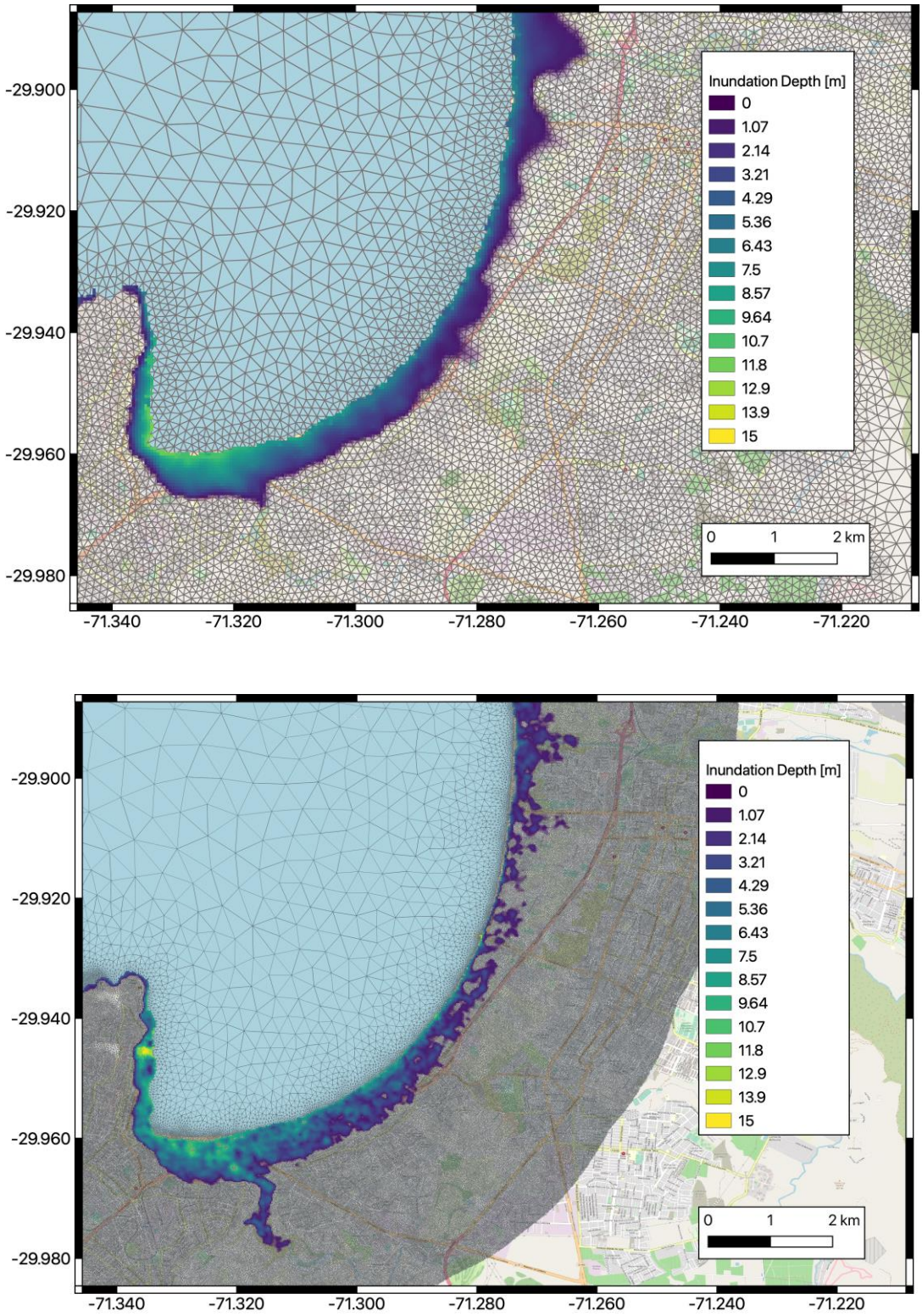

Figure 8 Small section of the triangulation in the coarse (upper panel) and fine mesh realisation in Coquimbo bay. The figures also contain the inundation result for identical sources.

2.3 SCENARIO EVALUATION

Copernicus Emergency Mapping Service (CEMS) provides on-demand detailed information for selected emergency situations that arise from natural or man-made disasters anywhere in the world. CEMS Rapid Mapping (RM) provides geospatial information within hours or days of a service request to support emergency management

activities in the immediate aftermath of a disaster: the delivery of CEMS-RM products is strongly connected to the acquisition of satellite data.

Wania et al. [3] show that the step devoted to the satellite image acquisition takes most of time in the general SEM workflow and that the delay between a user request and the delivery of proper imagery could limit the efficiency of the service. Experience over the past 3 years of CEMS-RM activations has shown that the observed average time lapse between an earthquake event and the triggering of the related activation request is 20 hours in average (in a range variable between 3 to 107 hours, considering 10 activations). This time lapse could lead to a delay in image acquisition of at least one day if the image request comes too late to allow data acquisition on the same day, considering different request cut-off times for each image providers. As an example, Table 3 shows the values of service hours for the handling of rush orders, the deadlines for the data providers to receive rush tasking requests (cut-off times) and Table 4 expected delivery timeliness for some frequently exploited Copernicus Contributing Missions (CCMEs).

TIME DELTA [hh:mm]	ACTIVATION (FROM EVENT)	1st USABLE IMAGERY (FROM ACTIVATION)	1st MAPPING PRODUCT FROM EVENT	1ST MAPPING PRODUCT FROM IMAGE AVAILABILITY
MEAN	20:29	25:25	57:55	14:40
MIN	02:53	05:55	26:58	05:32
MAX	107:00	83:01	157:12	46:42

Table 3 Time delta between time event and (1) activation request, (2) activation time from image availability, (3) delivery of the first product from time event and (4) delivery of the first product from image availability

IMAGE PROVIDER	TASKING CUT- OFF TIME (UTC)	RUSH DATA DELIVERY TIMELINESS
e-GEOS	7:30 for data collection from 18:21 same day to 6:21 the next day; 17:30 for data collection from 6:21 same day to 18:21 the next day;	Within 3 hours from order confirmation (archive) and 5 hours from new acquisition downlink (new)
AIRBUS Defence and Space	08:30 of day 0 for tasking opportunities from 18:02 of day 0 to 07:17 of day 1; 21:30 of day 0 from tasking opportunities from 07:17 of day 1 to 18:01 to day 1	Within 3 hours from order confirmation (archive) and for new acquisition: Europe, within 3 hrs, Rest of the globe within 7 hrs
Planet	09:00 for tasking opportunities from 13:00 to 23:59 of the same day; 18:00 for tasking opportunities from 00:00 of day 1 to 18:01 to day 1	Within 3 hours from order confirmation (archive) or sensing (new)

Table 4 Some examples of emergency service hours, emergency orders cut-off times and expected rush delivery timeliness (source: CCMEs Rush Service Support - CSCDA (Copernicus Space Component Data Access))

As explained by Ajmar et al. [4], despite technical limitations of tasking satellites for a new acquisition (e.g. orbits and cut-off-times for scheduling a new tasking), in case of tsunami or earthquake events, time can be gained by submitted a tasking as soon as an event has occurred.

Based on these considerations, the first expected benefit of the LEXIS workflow proposed in the frame of this project, one that exploits loss assessment data and simulation provided by TsunAWI, compared with the standard

one, and activated on-demand by authorized users, is to allow the reduction of time required to receive the needed post-event satellite imagery, in some cases up to one day in advance.

The evaluation of the benefits due to a possible exploitation and integration of LEXIS system in the operational CEMS workflow has been carried out considering, as a case study, the Chile earthquake and tsunami events occurred in 2015.

In the Figure 9, the Chile earthquake and tsunami activation timelines related to current CEMS activities are summarized and compared with the one based on workflow including LEXIS products.

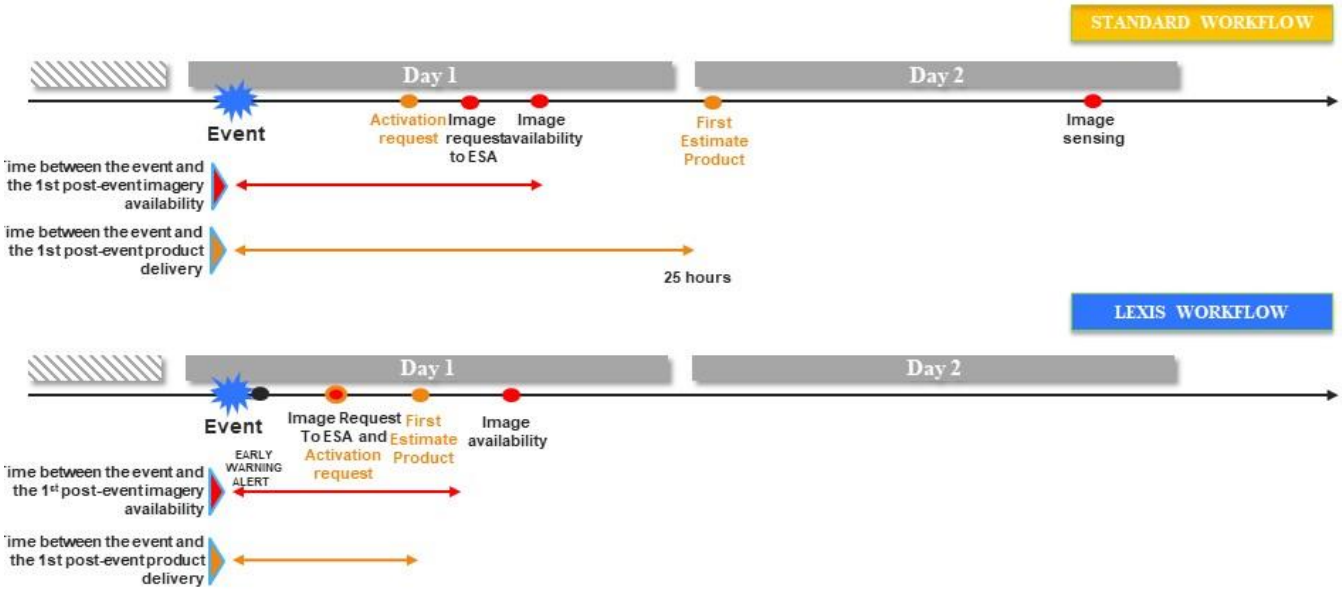

Figure 9 Comparison, in terms of delivery timeliness, of a CEMS-RM standard and one based on Lexis workflow

The event occurred on 16 September at 22.54 UTC. CEMS-RM[14] and the International Charter Space and Major Disasters[15] were triggered about 9 hours after the event. In its first request, the Emergency Response Coordination Centre (ERCC), acting as authorized user of CEMS-RM, required maps for the area of Coquimbo city and surrounding, based on media reports and knowledge received from the ground. The results of first satellite-based damage assessment maps were published 25 hours after the earthquake and 16 hours after the user request. Considering this delivery time, it can be stated that the general timeline for CEMS-RM activation was extremely good and effective. Very important for this purpose was the scheduling of the satellite tasking over the affected areas at such short notice (request in the early morning and acquisition in the afternoon on the same day). The availability of the first usable imagery from activation generally takes one day, depending on request time with respect to the satellite tasking cut-off times.

[14] CEMS-RM: https://emergency.copernicus.eu/mapping/list-of-components/EMSR137
[15] Charter Space and Major Disasters: https://disasterscharter.org/ru/web/guest/activations/-/article/ocean_wave-in-chile

The LEXIS workflow is triggered directly when receiving an earthquake alert, allowing the production of a TsunAWI fast simulation, a shakemap and a fast loss assessment. The outputs of these operations, normally produced within few minutes, activate the computation of the LEXIS SEM workflow aiming to generate raster products in support to the prioritization of impacted areas supporting satellite early tasking operations. In this scenario, the first post-event image available for damage assessment is expected to be delivered up to one day in advance. In the specific Chile case, the discrepancy between the two workflows, in terms of timeliness for satellite tasking, is not relevant because, as said, the timeline for CEMS-RM activation was particularly favourable.

Nevertheless, a significant difference in the timelines obtained using the original CEMS-RM workflow and the modified one is observed with regard to the timeliness of delivery of the First Estimate Product (FEP). In the workflow obtained exploiting LEXIS platform, after receiving the earthquake moment tensor, approximately 10 minutes after the initial earthquake event, a precise wave height and an inundation map are produced. Those datasets, together with the precise damage assessment, generated within an hour, can then be used to generate the CEMS FEP. The novel approach is the delivery of a FEP without waiting for the acquisition of post- disaster images which allow a huge advance on products delivery times.

2.3.1 FEP product for earthquake and tsunami event

Vector building features affected by the tsunami inundation, including information about people potentially affected and the expected wave height to which they could be subjected, constitute the base for a first damage assessment product in case of tsunami event. This vector layer is generated by the following input data:

- Building geometry and number of people present in the building during night-time hours,
- Simulated inundation.

Tsunami inundation area, wave height and spatial distribution of buildings are very useful bits of information to indicate which buildings are potentially affected by the event.

Figure 10 Buildings potentially affected by inundation on the left and results from damage surveys completed in the field after the event. Buildings affected can be detected for higher wave height simulation.

Considering the Chile test case, Figure 10 shows an example of the potentiality of these data in support to rapid mapping activities. On the left, buildings potentially affected by inundation and the expected wave height at their location, derived by a TsunAWI simulation is displayed. On the right, results from damage surveys completed in the field after the event (results reported in Paulik et al., [5]) are shown. Information related to expected wave heights and the area affected by inundation allow for a first damage assessment based only on simulation modelling. The model results seem to agree with field surveys carried out in Coquimbo.

As far as earthquake events are concerned, loss assessment dataset containing different classes of damage probabilities and exposure information about building occupancy, is suitable to derive a FEP based only on simulation modelling.

Considering the example of Chile test case, loss assessment data derived from an earthquake event shows reduced level of damage. This has been confirmed by satellite-based damage assessment operations carried out in these areas and by information concerning damage to buildings, reported in various field survey results (around 60 houses were destroyed and nearly 200 damaged, mostly in the Region of Coquimbo[16]).

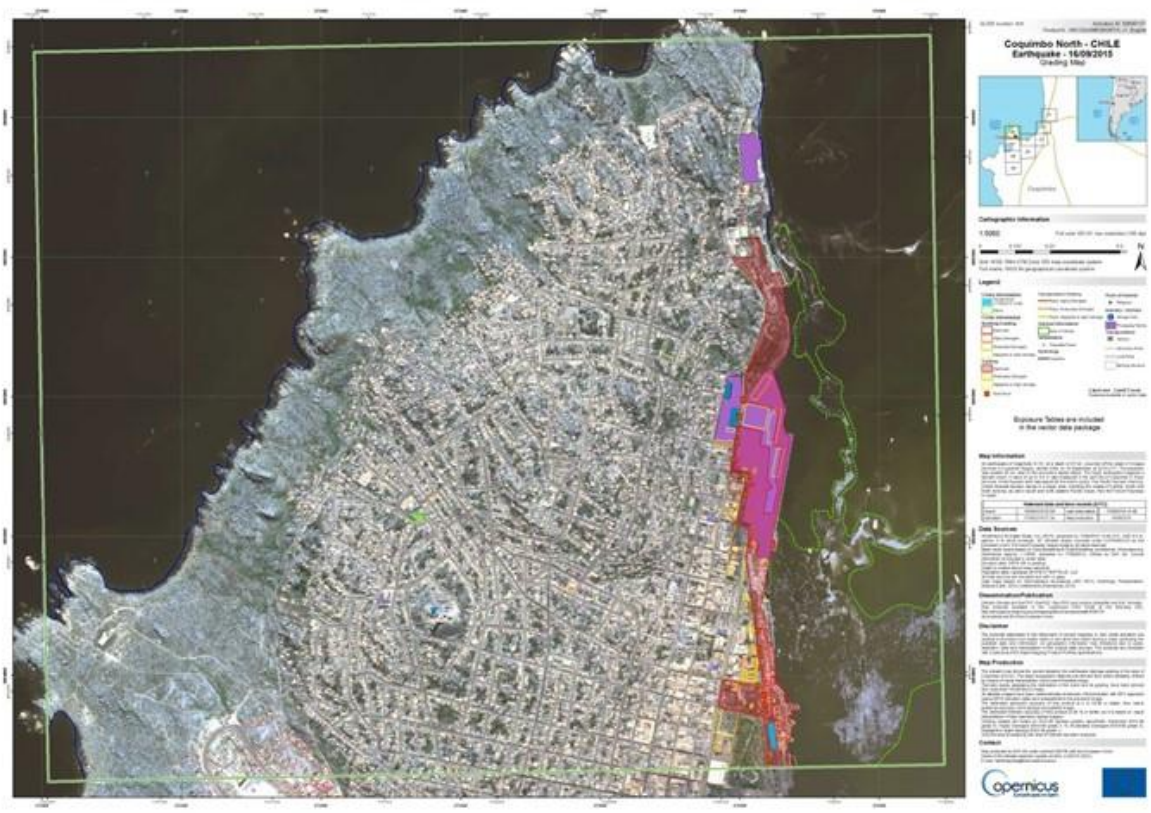

Figure 11 Example of one of the grading products issued in the CEMS-RM activation.
No actual damage has been reported in the product, excluding the coastal areas affected by tsunami inundation (source: Copernicus Emergency Management Service © European Union 2015, [EMSR137] Coquimbo North: Grading Map).

2.3.2 Satellite tasking and AOI prioritization

As already mentioned, a critical aspect to allow effective CEMS-RM activities, is to grant the early-tasking of satellite acquisition to improve the time delivery of the products and to support the prioritization of proper Areas of Interest (AOIs) to be submitted to the satellite data providers.

[16] Field survey result: https://reliefweb.int/node/1177936/

In the frame of the CEMS-RM activation EMSR137, 15 areas of interest (AOIs) have been defined along Chile's coast, as depicted in the Activation Extent Map[17] . In particular, 9 AOIs have been defined, covering the areas surrounding Coquimbo (see Figure 12).

According to the observed outcomes of the grading information extracted from satellite imagery acquired in the frame of this activation and supplied to the users, only 4 AOIs were actually affected, as damage to buildings was observed in them. These AOIs are visible in Figure 13. They are represented using different colour to indicate the qualitative whole damage level observed in the AOI: red areas contain the higher number of affected buildings, characterized by different severity grades of damage (buildings with negligible, moderate, high damage levels and/or completely destroyed), while the yellow ones contain mainly a very reduced number of affected buildings (from 10 to 20 buildings).

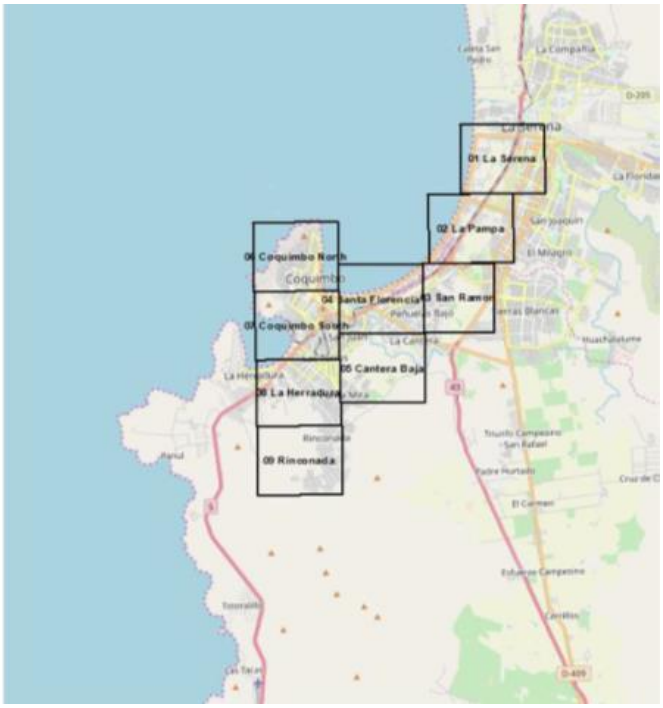

Figure 12 Location of 9 AOIs near Coquimbo mapped by the CEMS-RM (EMSR137)

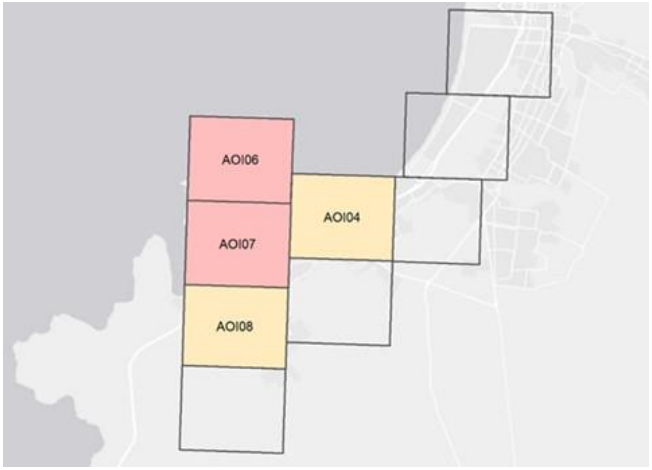

Figure 13 Affected AOIs.
Red areas contain the higher number of affected buildings, while the yellow ones contain mainly a very reduced number of affected buildings.

[17] Activation Extent Map: https://emergency.copernicus.eu/mapping/sites/default/files/thumbnails/EMSR137-AEM-1442912620-r04-v1.jpg

Figure 14 shows the map products obtained for this case study integrating the earthquake loss assessment dataset and tsunami model outputs with exposed assets. In particular, Relative Potential Damage map and Absolute Probability map can help to identify the expected most affected areas and, with caution, the possible level of damage in those areas, in case of earthquake event, while the Tsunami Potential Damage map helps to delineate the expected most affected areas, in case of tsunami event.

According to the earthquake Relative Potential Damage map, main affected areas could have been identified in the CEMS-RM AOI06, AOI07 and in the areas near the AOI01 and AOI08. Nevertheless, the absolute values of probability of damage derived considering the most severe damage classes, reported on the Damage Absolute Probability map, clearly indicated very unlikely significant damages in these areas, as indeed was indicated by the results of the analysis of the satellite images carried out during the CEMS EMSR137 mapping activities. In addition, general low levels of damaged buildings due to the occurred earthquake event have been confirmed in this case study by several sources, such as local authorities and post event field survey activities. Moreover, the analysis of the tsunami Potential Damage map gives an indication of areas where higher levels of damage could have been expected, exploiting exposure information about possibly affected buildings, their occupancy and the expected maximum wave height provided by TsunAWI simulations. In this case, the main affected areas, obviously located along the coastline, belong to the CEMS-RM AOI07 and AOI02, and partially to the AOI04.

Finally, it can be observed that the content of all proposed products derived from the LEXIS models, properly integrated and interpreted, seems to be sufficient, in the frame of this specific case study, to effectively support the early identification and prioritization of areas where to trigger satellite acquisitions and perform subsequent satellite-based damage assessment operations.

In particular, the operations of prioritization of areas of interest based on these data would have led to the selection of a reduced area on the overall, if compared to the one adopted for the purposes of the CEMS-RM activation, including only some regions that are in the CEMS-RM AOI02, AOI03, AOI04, AOI06 and AOI07. This appears to be in agreement with the results of field surveys carried out in the areas near Coquimbo (see, for instance, results reported in Paulik et al. [5].

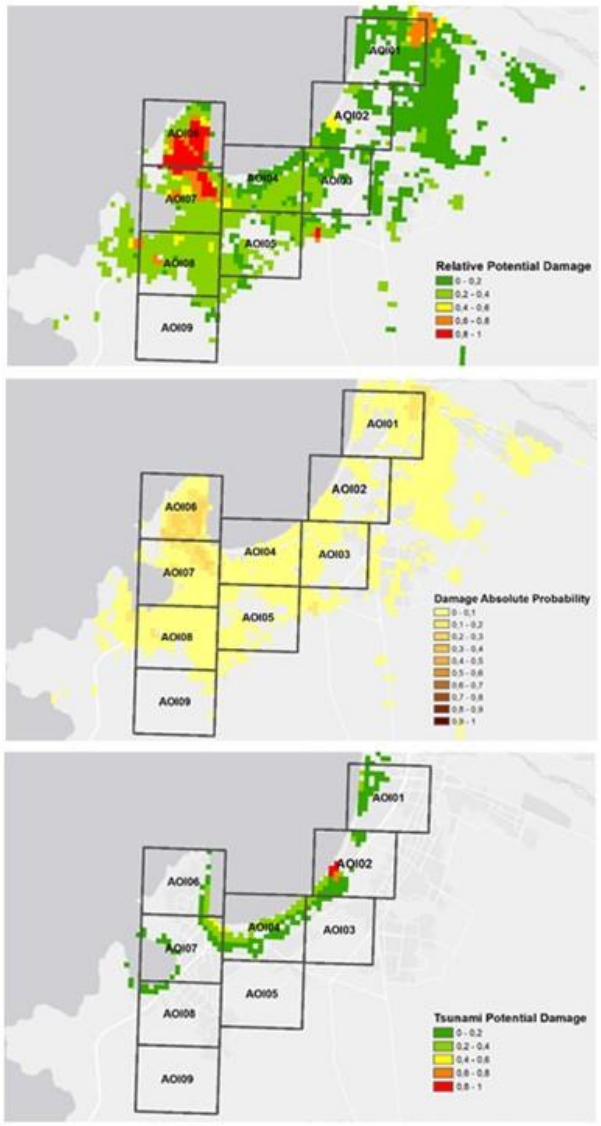

Figure 14 Map products obtained for the case study.
From top to bottom prioritization of areas of interest based on Relative Potential Damage map, Damage Absolute Probability map and Tsunami Potential Damage Map in the area of Coquimbo and La Serena.

2.4 FUTURE EVOLUTIONS

The presented workflow provides the first step on a roadmap that was set to be more ambitious.

One future step would be combination of outputs of both computing services (either in HPC centres or cloud infrastructures) to have redundancy and more accurate results in the case both are available. This would require to set up a virtual input file for both locations for the earthquake event, a virtually shared (non-blocking) directory for the inundation and damage assessment from the simulation, and a redundant data-fusion process from the virtually shared directory, which would give a redundant output (if both sites are available). The virtually shared directory (simply utilizing the DDI already implemented in LEXIS) would act as the data-fusion directory for the several tsunami simulations, as long as we can ensure that different sets of tsunami simulations are processed with the results in this virtually shared directory. This way each site would have an extended number of simulated scenarios and therefore can achieve a better accuracy. Even though the LEXIS platform already implements this through the DDI.

The second possible evolution would be to integrate the rapid inundation and loss assessments with more accurate but longer simulation at least for the inundation with fine grained grid resolution with a result a few minutes after the initial estimations.

3 SUMMARY

This document, based on what was proposed in Deliverable D6.3 [1], describes further developments carried out and the final working workflow proposed for earthquake and tsunami large scale pilot of the LEXIS project.

As far as the loss assessment calculation is concerned, the main evolution is the refactoring of the use of fragility functions and the change of the input data formats and output data aggregation: those changes make the existing tools more flexible, more compliant with input data licensing conditions and more aware of possible privacy concerns.

The TsunAWI calculation process has been further improved based on the results of performance tests. Adding an asynchronous I/O allows the user to have access to a snapshot of the wave propagation on the fly while the process is running.

Furthermore, outcomes and impacts obtainable from the adoption of this workflow to support SEM activities have been presented and discussed in the frame of a real tsunami event occurred in Chile in 2015.

REFERENCES

[1] LEXIS Deliverable, *D6.3 Pilot improvements: Evaluation of Software Development.*

[2] D. Schorlemmer, F. Euchner, P. Kaestli, J. Saul and QuakeML Working Group, "QuakeML: Status of the XML-based Seismological Data Exchange Format," *Annals of Geophysics,* vol. 54, no. 1, pp. 59-65, 2011.

[3] A. Wania, I. Joubert-Boitat, F. Dottori, M. Kalas and P. Salamon, "Increasing Timeliness of Satellite-Based Flood Mapping Using Early Warning Systems in the Copernicus Emergency Management Service," *Remote Sens.,* vol. 13, no. 11, 2021.

[4] A. Ajmar, A. Annunziato, P. P. Boccardo, F. G. Tonolo and A. Wania, "Tsunami modelling and satellite-based emergency mapping: Workflow integration opportunities. no. 314," *Geosciences,* vol. 13, no. 11, 2019 Switzerland.

[5] R. Paulik, Williams P.A, N. Horspool and P. Catalan, "The 16 September 2015 Illapel Earthquake and Tsunami: Post-Event Tsunami Inundation, Building and Infrastructure Damage Survey in Coquimbo, Chile," *Pure and Applied Geophysics,* pp. 1-15, 2021.