



eFlows4HPC

Enabling dynamic and Intelligent workflows
in the future EuroHPC ecosystem

D6.2 Description of the use cases for Pillar III

Version 1.0

Documentation Information

Contract Number	9555558
Project Website	www.eFlows4HPC.eu
Contractual Deadline	31.08.2021
Dissemination Level	PU
Nature	R
Author	Jacopo Selva (INGV)
Contributors	Fabrizio Bernardi, Alberto Michelini, Fabrizio Romano, Roberto Tonini (INGV), Jorge Macías, Carlos Sanchez Linares (UMA), Marisol Monterrubio (BSC), Marta Pienkowska (ETHZ)
Reviewer	Steven Gibbons (NGI)
Keywords	urgent computing, earthquakes, tsunamis, natural hazards



This project has received funding from the European High-Performance Computing Joint Undertaking (JU) under grant agreement No 955558. The JU receives support from the European Union's Horizon 2020 research and innovation programme and Spain, Germany, France, Italy, Poland, Switzerland, Norway.

Change Log

Version	Description Change
V0.1	Proposed ToC and First Draft
V0.2	Updated ToC
V0.3	Cleaned, after significant reviews
V0.4	Finalized and submitted for revision
V0.5	Revised based on reviewer's comments and suggestions
V1.0	Final version

Table of Contents

1. Executive Summary	3
2. Introduction	3
3. Data required and computational resources for implementing and testing seismic and tsunami workflows	5
3.1 The Seismic Workflow (UCIS4EQ)	5
3.2 The Tsunami Workflow (PTF/FTRT)	8
4. Use cases: specific data and computational resources	10
4.1 Target regions and selected events	11
4.1.1 Mediterranean Sea	11
4.1.2 Mexico	12
4.1.3 Iceland	13
4.1.4 Chile	15
4.2 The Seismic Workflow (UCIS4EQ)	16
4.2.1 Global datasets/repositories	16
4.2.2 Data for the Mediterranean Sea	17
4.2.3 Data for Mexico	19
4.2.4 Data for Iceland	21
4.2.5 Estimation of Computational Resources	22
4.3 The Tsunami Workflow (PTF/FTRT)	23
4.3.1 Global datasets/repositories	23
4.3.2 Data for the Mediterranean Sea	23
4.3.3 Data for Chile	25
4.3.4 Estimation of Computational Resources	27
5. Conclusions	27
6. Acronyms and Abbreviations	28
7. References	28

1. Executive Summary

This deliverable is a compendium of datasets and models employed for development and validation of Pillar III (Urgent Computing) workflows. In **Section 2**, we introduce the main data to be collected for these workflows, in order to introduce the rationale adopted in selecting the use cases. In particular, we defined a few regions for which urgent computing will be made possible through the retrieved datasets. This completely defines the urgent computing models, building a running implementation of the workflows described in Deliverable D6.1. Among the selected regions, one is common to both seismic and tsunami workflows, enabling potential multi-hazard use cases. In each region, we selected a few earthquakes to run and test the workflows, defining specific seismic and tsunami use cases. In **Section 3**, we describe in details the data collected for the seismic and the tsunami workflows, highlighting their role within the computation framework. In **Section 4**, we describe the specific regions selected and the collected databases that will enable the implementation of the use cases. The datasets have been organized in a single repository for Pillar III, in b2drop (<https://b2drop.bsc.es/>).

2. Introduction

Both seismic and tsunami urgent computing procedures quantify the potential hazard due to an earthquake just after its occurrence, that is within a few hours from its occurrence. This of course depends on the end-user to which such estimations should be delivered. For example, the ARISTOTLE-eENHSP Project (<http://aristotle.ingv.it/tiki-index.php>) typically delivers the hazard reports in approximately 3 hours, leaving a couple of hours for computation time. The hazard is quantified by simulating the propagation of the elastic waves in the solid earth and, potentially, of the tsunami generated by the seismic source, while accounting for the uncertainty due to the scarce knowledge of the model parameters (such as the seismic source characteristics or the velocity/structure model) and the wave modelling procedure.

In general terms, the input data to both the seismic and the tsunami workflows fall in three categories:

1. **Earthquake input:** urgent computing modelling evaluates the potential impact of an earthquake that just occurred. Therefore, all simulations are initialized by the data coming from the monitoring systems (e.g., NEIC/USGS, GFZ, EMSC, INGV) that provide a first estimation of the source parameters. In general, such parameters should include values estimating the energy and the location of the seismic events, as well as a characterization of the dimensions, the geometry, and the kinematics of the originating fault.
2. **Simulation settings:** wave propagation modelling requires the definition of the medium through which waves will propagate. For seismic waves, the medium is described by topobathymetric data describing the free surface, as well as 3D velocity models for the crust. For tsunamis the computational domain is described by the topo-bathymetric data. The tsunami models are usually depth averaged, and the horizontal wave speed is controlled by the water depth given by the topo-bathymetric data.
3. **Waveform data:** during an event, the physical observables of the earthquake (seismic and tsunami waves) start becoming available from seismic stations and other sensors recording

tsunamis (e.g. tide-gauges) surrounding the seismic source. Such data should be used to constrain and/or to check urgent computing forecasts.

Notably, earthquake and waveform data (points 1 and 3 above) are related to a specific seismic event, while simulation settings depend on the target area only, and can be treated as time-independent. Therefore, data concerning simulation settings can be set in advance, and only earthquake and waveform data have to be retrieved in real-time. Once simulation settings are set for a given source area of interest, the urgent computing simulations can be run for any seismic event occurring in the source area.

To guarantee a range of potential use cases, we first define a list of target regions for which all the required simulation settings are to be collected. In this way, the urgent computing environment will be set for such regions and can be applied to any source event occurring in the target source area. Then, we pre-select one or two specific historical events of interest in each region for which we will collect seismic source and seismic/tsunami wave data to test our procedures.

To select such regions, we have to consider the peculiarity of the seismic and tsunami cases. Indeed, even if they share a common source (earthquake), they typically have different needs. For example, seismic waves are relevant mainly in the surroundings, closer to the seismic source, e.g. within a distance of a few hundreds of kilometres from the originating fault. Significant waves may be originated also by moderate earthquakes ($M > 5$), and good recordings require a dense network of strong motion instruments, ideally all around the source. The earthquakes that generate tsunamis occur offshore or near to the coast, and consequently they may be of scarce interest for seismic urgent computing, as few or no lands may be within the near field (e.g., < 200 km) where seismic shaking is significant. In addition, as seismic stations are mainly deployed on land, seismic waveform records may be also scarce or hill distributed (e.g., only on one side) for tsunamigenic earthquakes. Consequently, an earthquake generating a tsunami is not necessarily of interest for seismic urgent computing. On the other hand, many earthquakes of potential interest for seismic urgent computing do not generate significant tsunamis, such as moderate ($M < 6.5$) or inland earthquakes. Consequently, an earthquake generating significant seismic shaking is not necessarily of interest for tsunami urgent computing.

However, there are several examples of earthquakes that are relevant for both workflows, that is, near-coast large earthquakes for which seismic waves may be relevant in the nearby coastal area and tsunami waves may be generated. For example, the 1908 Mw 7.1 Messina Straits earthquake originated under the sea and caused at least 80,000 deaths due to both seismic and tsunami waves, and a few hundred were added by the tsunami (e.g., Guidoboni et al. 2019; Rovida et al, 2020). More in general, all near-coast large magnitude ($M > 6.5$) events may generate both seismic and tsunami hazards, and thus can be considered of common interest.

Therefore, we decided to define a few regions for each workflow, considering the peculiarity of each application. In addition, we included one common region, for which a multi-hazard earthquake-tsunami use case may be defined. In particular, we defined the Mediterranean basin as the common region for which several recent events of interest occurred and for which significant data can be retrieved. Then, for testing the *seismic workflow UCIS4EQ*, two additional regions will be considered, located in Mexico and in Iceland. For testing the *tsunami workflow PTF/FTRT*, we will consider the Chilean source area, extending the target area to the entire Pacific Ocean.

3. Data required and computational resources for implementing and testing seismic and tsunami workflows

In this section we describe the type of data and their role that are used in the use cases relative to the seismic and tsunami workflows. This includes input data for the workflows and data necessary for testing/evaluating the results.

3.1 The Seismic Workflow (UCIS4EQ)

The data required for UCIS4EQ to execute an end-to-end run is divided into three types following a creation-consumption pattern: sources or inputs, intermediate or temporal, and outputs. The data that flows through UCIS4EQ is read or created by the different building blocks in an external host or in a HPC environment, that is:

- transferred from a remote data repository,
- queried in streaming,
- pre-computed and stored in a target machine, or
- generated from a determined action produced in a building block.

The data structure in UCIS4EQ includes different formats such as key-value structures, multi-dimensional arrays of numerical data, time-series, raw data, serialized objects, images, and documents.

The computational resources needed to solve seismic simulations depend on two factors, the targeted resolved frequency and the size of the computational domain. Higher frequency implies both increasing the spatial mesh resolution and decreasing the time step to solve the elastic wave equation. On the other hand, the size of the domain depends on how far from the hypocentral location the seismic effects should be computed.

The functional data required in UCIS4EQ is divided into the region-specific information and the earthquake-specific data, as follows:

For the regions:

- **Velocity models:** Seismic waves travel from the rupture fault through the heterogeneous Earth crust to the surface where we experience the destructive ground shaking. The wave speeds change for different medium properties, and representing those properties is crucial to successful numerical modelling of seismic waves. A tomographic velocity model is a spatial representation that characterizes those properties for a given region to the best of our knowledge.
- **Topography and bathymetry:** Seismic waves interact with 3-D structures and interfaces. In particular, the topography and bathymetry on the Earth's surface are known to have significant focussing/defocussing effects that impact the ground shaking. Discretising and

representing topography and bathymetry in the simulation domain allows us to account for and study such effects. Topography and bathymetry are available from a range of sources, in particular Earth2014 global topography (relief) model (Hirt and Rexer, 2015), NASA's Shuttle Radar Topography Mission (SRTM, Farr et al, 2007), or the Global Multi-Resolution Topography (GMRT; Ryan et al. (2009)).

- **Faulting mechanism forecast and CMT catalogue:** Faulting mechanism forecasts may be used to initialize fault geometry uncertainty distribution when only location and magnitude data are available (Selva et al., 2021b), as well as to treat uncertainty when the first real-time solutions (see below, under data for specific earthquakes) start being available. In UCIS4EQ, the faulting mechanism initialization is presently based on the local Centroid Moment Tensor (CMT) catalogue (Monterrubio-Velasco et al., *in review Frontiers in Earth Science*). These catalogs are queried from the Global Centroid Moment Tensor (GCMT) catalog which includes the hypocenter location, earthquake size, faulting geometry, the CMT, and the FM solution of moderate to large events with magnitude $M \geq 4.5$. At present, the GCMT catalog contains more than 40,000 earthquakes (Dziewonski et al., 1981; Ekström et al., 2012). Other global or local fault mechanism forecasting models may be available (e.g., Kagan and Jackson, 2014; Selva et al. 2016; Roselli et al., 2018; Taroni and Selva 2021) and could be retrieved for input or for testing purposes.
- **Receivers and infrastructure locations:** The location of seismic stations, critical infrastructures, cities and towns that are relevant to seismic hazard assessment. We compute the simulation results in these specified sites to evaluate and calibrate the outputs.

For the specific earthquakes:

- **Regional waveforms & source parameter estimations:** regional waveforms are usually analysed in real-time to extract estimations of the seismic source parameters. Such estimations evolve through time, as data from new stations becomes available. Certain estimations, like for example a more accurate source geometry, may become available at later times (several hours if not days). All parameters may either be directly calculated from waveforms or retrieved from web services.
- **Real-time faulting mechanism:** moment tensor inversion provides the first estimation of the faulting mechanism, which defines the geometry of the fault and the direction of the dislocation. Moment tensor solutions start being available after a few minutes after the event, and typically include an ambiguity for the faulting mechanisms between two potential focal mechanisms, which may be solved using local data (Selva et al. 2016; Taroni and Selva, 2021).
- **Finite fault:** Finite fault slip models require information to simulate the fault rupture. This data is usually available from seismic agencies a few hours (or even the following day in some cases) after the earthquake. In UCIS4EQ this information is used to estimate the kinematic fault rupture that will be input to the wave-propagation HPC code. To bridge the lapsed time until a finite-fault slip inversion has been computed we use the Graves-Pitarka (GP) rupture generator that estimates the full kinematic rupture history based on empirical laws and stochastic distributions (Graves and Pitarka, 2010), taking in as input externally estimated parameters like fault location and geometry (length, width, strike and dip), seismic moment or magnitude, rupture initiation point (hypocenter), and slip direction

(rake). For uncertainty quantification and testing purposes, we will gather the finite fault slip inversions computed in a post-processing stage.

- **Local strong motion data:** local strong motion records measure the actual ground motion caused by the earthquake in the target area. They may represent both a term of comparison for testing UCIS4EQ results, as well as potential input to the workflow allowing a forecast update through assimilation techniques
- **Shakemaps:** shakemaps provide an estimate of the spatial distribution of ground motions and shaking intensities following a significant, potentially destructive event. Such maps adopt recorded data whenever available, and otherwise use ground motion models corrected for local site effects. They are used both for rapid response and for preparedness planning. Simulation results for past events need to be compared with existing shakemaps for testing purposes.

Moreover, **computational meshes** are required for the simulations. Regional meshes need to be built in advance and stored on the HPC facilities, and they are generated for a target frequency resolution using the region-specific velocity model, topography and bathymetry (see above). Therefore, the size of this data depends on the maximum resolved frequency of the simulation, which can range from 1 Hz (small runs) to 10 Hz (large scale runs), as well as on the spatial extent of the target area. In the current setting, the extent of the target area can change depending on the earthquake size. Ultimately we aim to provide very large meshes for entire regions, with an ability to “mask” elements and run simulations on target areas only. For the purpose of this deliverable, we provide a few sample meshes for the Mediterranean region (see Section 4.2).

As a summary, **Table 3.1** includes the required data and the building block that uses each database.

Table 3.1 Definition of the data sets used in the different Building Blocks (BB) of the UCIS4EQ workflow (BB#2 Source parameters acquisition, BB#3 Source Building, BB#4 EQ. HPC simulations, BB#8 Results post-processing; see deliverable D6.1 for further BB details).

Data sets	Used in	Comments
Historical CMT catalogs	BB #2	The results generated after use the CMT catalogs are used in BB #7 and BB #3
Velocity models	BB #3 and mesh generation	The 1D-velocity models are used in BB #3. The 3D velocity models are used to generate the computational mesh.
Regional bathymetry and topography	mesh generation	Topography and bathymetry are used to generate the computational mesh.
Computational Meshes	BB#4	Precomputed. Depend on required frequency resolution and target area.
Waveforms at seismic stations	BB#8	Time series from regional seismic networks
Infrastructure locations and cities	BB#8	This information is used to compute synthetic waveforms at specific locations
Finite Fault models	BB #3	This dataset is obtained in streaming by BB#5. Once this information is available is used in BB#3.

Shake maps of ground motion proxies (PGA, PGV, PSA, shaking duration)	BB#8	This information is used to validate and compare the simulation results and that computed using real data in the post-processing stage
GMPE's	BB#8	This information is used to the uncertainty quantification and the calibration in the post-processing stage

3.2 The Tsunami Workflow (PTF/FTRT)

The PTF/FTRT workflow (**Fig. 3.1**) is composed of three stages. Notably, considering that the workflow is still in development, with many blocks that are only partially developed, or even have not been realized yet (see Deliverable D6.1), the PTF/FTRT workflow that we describe is still subject to potential changes in the course of this project.

The PTF/FTRT workflow is initialized by a seismic trigger indicating a large earthquake with a potential to generate a tsunami. Both automatic and human user triggers may be implemented. In this phase, taking as reference for urgent computing the ARISTOTLE-eENHSP Project (<http://aristotle.ingv.it/tiki-index.php>), we target human trigger (e.g., the person on duty for tsunami hazard evaluation). Block #1 retrieves information about the (potentially) tsunamigenic seismic event and Block #2 initializes an ensemble of seismic sources covering the uncertainty on source characteristics. Blocks #3 simulate tsunami wave propagation for the seismic scenarios in the ensemble, dividing the ensemble in groups of scenarios. The size of groups is defined based on the available computational resources. Likely, many (at least tens) of groups are required to complete the ensemble, defining many cycles in the process. After each cycle, the results are aggregated (Blocks #5 & #6) to produce a forecast of the potential tsunamis at target points, and the asymptotic convergence of forecasts is checked (Block #7), to evaluate if the already simulated scenarios sufficiently cover the ensemble statistics. At each cycle, the ensemble manager (Block #2) checks whether new data for the event are available, and potentially updates the ensemble based on this new information, potentially reweighting or eliminating ensemble members. When a sufficiently large number of simulations is available, tsunami propagation may be emulated based on machine learning procedures (Block #4), instead of explicitly simulating the tsunamis (Block #3), significantly reducing the computational effort. Overall, after a certain number of cycles, this workflow will start producing stable forecasts of the tsunami size at target points based on all the available information at the time of the estimate. Workflow execution is ended by Block #7 based on a predefined rule, for example a predefined delivery time based on when urgent computing forecasts have to be delivered to end-users, triggering the production of forecast visualization (Block #8) and the long-term storage of all produced simulation results (Block #9). The timing of this delivery depends on specific agreement with the end-user: taking again as reference the ARISTOTLE-eENHSP Project, reports should be typically finalized within 3 hours from the request, leaving approximately 2 hours for computation time.

The data required for PTF/FTRT to execute an end-to-end run can be divided according to their role in the workflow, as defined in Section 2. Simulation settings data (type 2 in Section 2) should be fully available at the time of the execution, and may probably be preinstalled in the HPC server. Workflow execution may start when earthquake input data (type 1 in Section 2) are retrieved by Block #1, allowing to initialize the ensemble.

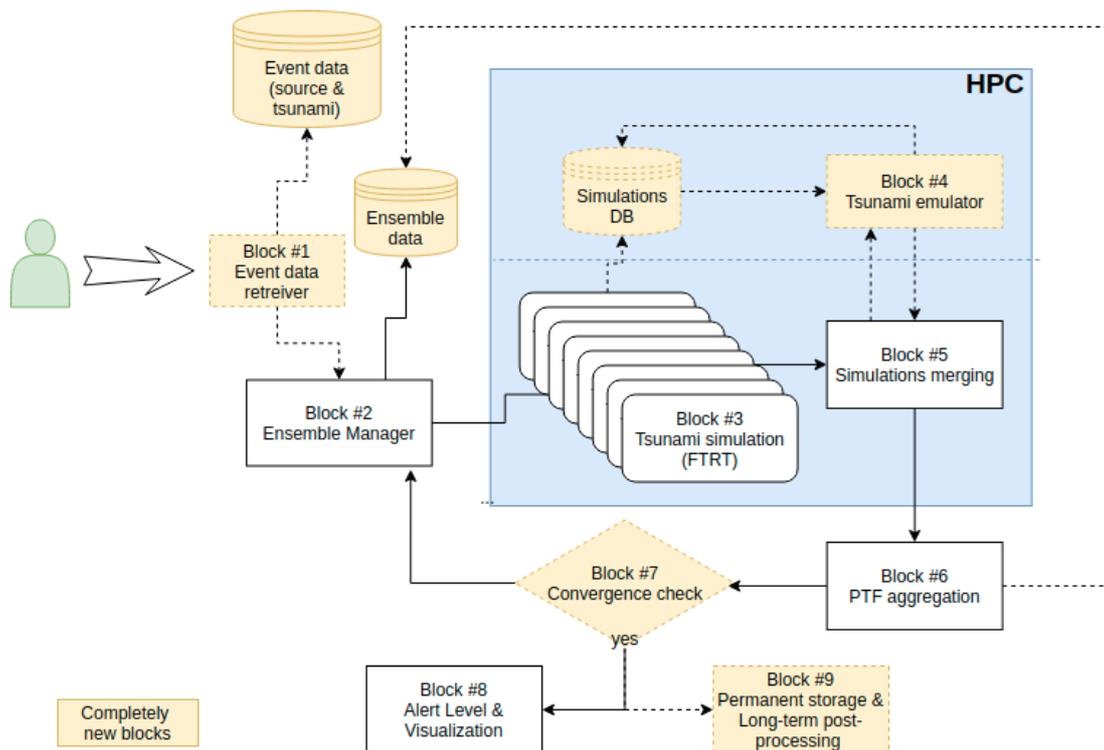


Figure 3.1 Schematic representation of PTF/FTRT workflow, as described in deliverable D6.1. Blocks highlighted in yellow have to be completely designed and developed within this project, either in Phase 1 or 2. The other blocks will be instead updated from existing methods / scripts.

During the execution, incoming earthquake input and tsunami waveform data types 1 and 3 in Section 2, respectively) are acquired continuously to update PTF/FTRT forecasts. Such new data are again retrieved in real-time by Block #1, and Block #2 manages the updating of the ensemble, taking into account the newly acquired data. Notably, the same data may be used for further testing purposes also after the urgent computing is finished. At the end of the execution, when Block #7 indicates convergence, updated forecasts are produced and simulation results are stored for more in-depth analysis purposes.

While a quite large range of data formats are necessary (e.g. mseed, quakeML, generic ASCII files, see Section 4.3) for intermediate stages of the workflows, following international standards at different levels, the PTF/FTRT workflow is oriented to producing files in NetCDF format, tentatively following the CF (Climate and Forecast) convention (<https://cfconventions.org/>).

PTF/FTRT data can be divided into regional data, i.e. regionally specific simulation settings defined before an event occurs in the area, and event related data, i.e. data describing one specific seismic event occurring in the source area.

Regional data include:

- **Topo-bathymetric data:** they represent the basic data to allow modelling tsunami propagation from any source. They should cover the entire simulation domain and include the source area. If inundation is modelled, nested grids at increasing resolution may be used, scaling up to the metric scale in the innermost grid.

- **Source discretization:** to enable fast solutions, a predefined list of potential sources in the source areas can be considered. This list may include different types of seismicity (e.g., prevalent and background, after Selva et al. 2016; Basili et al., 2021). For early-warning purposes, this list can be connected to a list of pre-computed tsunami scenarios to be used in real time (Selva et al., 2021b), otherwise each scenario simulation is run on the appropriate HPC environment.
- **List of forecast points:** this is a list of the target points for which PTF will be evaluated. It must be inside the target area, and it must include the position of all the available instruments measuring the tsunami (tide-gauges, DART, etc.).

For the specific earthquakes:

- **Real-time Earthquake source parameter estimations:** as for the seismic workflow, regional seismic waveforms are usually analysed in real-time to extract estimations of the seismic source parameters. Such estimations evolve through time, as data from additional stations are acquired. Specific estimations, like for example source geometry, may become available at later times. All parameters may either be calculated from waveforms, or retrieved from web services. For PTF, the starting solution is represented by estimations from the Early-Est system (Lomax et al. 2009; Lomax and Michelini, 2009; Bernardi et al. 2015), which include magnitude and location estimation and it is updated every minute after the first event identification). Regarding magnitude estimation, Early-Est provides mb, Mwp and Mw_{pd} magnitude estimations. The rules to select the appropriate magnitude type are described in (Bernardi et al. 2015) and updated in (Amato et al. 2021).
- **Real-time Faulting mechanism:** At later stages, other solutions become available, better characterizing also the faulting mechanisms of the earthquake. This information is fundamental to initialize and update the ensemble, to manage the uncertainty of the tsunami. Faulting mechanism is typically estimated starting from the focal mechanisms derived from moment tensor solutions, which are provided by global and local web services. This information is then combined with other local information (e.g. local faults), to provide a first estimation of the faulting mechanism (Taroni and Selva, 2021; Selva et al. 2016).
- **Tsunami records:** tsunami records measure the tsunami waves at discrete points within the target area. They include coastal (e.g., tide-gauges) and deep-sea observations (e.g., DART, smart cables, etc.). They may represent both a term of comparison for PTF/FTRT results, as well as potential input to the workflow allowing to update the forecast through assimilation techniques.

4. Use cases: specific data and computational resources

Four regions have been defined as use cases for eFlows4HPC:

- For both seismic and tsunami use cases, we selected the Mediterranean region. This area includes the entire Mediterranean basin and all relative coastal areas (up to 200 km inland). The Kos-Bodrum 2017 Mw6.6 (USGS) earthquake and tsunami has been selected as a

specific event for both seismic and tsunami applications. To also cover the western Mediterranean area, a further event will be considered as a use case for the tsunami workflow that is the 2003 M6.8 Zemmouri-Boumedes earthquake and tsunami.

- For seismic use cases, two further regions have been selected: the Mexican intraplate seismic source zone and the Iceland southern seismic area. As target events, we selected the 19 September 2017 M7.1 Puebla earthquake in Mexico, and the M6.5 21st June 2000 South Icelandic Seismic Zone (SISZ) earthquake.
- For tsunami use cases, also the Chilean subduction zone has been selected. Several large tsunamis originated in this area and propagated all over the Pacific Ocean. Consequently, the source area is focused in Chile, but the target area extends to the entire Pacific. As an initial target event in this area, we selected the 2015 M8.3 Illapel earthquake and tsunami.

In Section 4.1 we briefly describe the selected regions. Then, we summarize the data retrieved and the main computational constraints in Sections 4.2 and 4.3 for the seismic and tsunami workflows respectively. The retrieved data guarantee the possibility to implement the defined use cases, as presently defined. Such data may be updated or integrated during the implementation of the use case. Computational constraints are reported as presently known, and future developments of the workflow may significantly change this information.

All the collected data have been organized in a common repository for Pillar III, exploiting b2drop services (<https://b2drop.bsc.es/>). The repository has been organized by type of hazard (seismic or tsunami) and by region, with one folder containing all global data (applicable to multiple regions). Each regional folder contains all data collected for a single region, and includes both the regional-specific data and the event-specific data. The repository is available, under restricted access, at the following link:

https://b2drop.bsc.es/index.php/apps/files/?dir=/eFlows4HPC_Pillar3/D6.2%20Use%20Cases&fileid=408947

4.1 Target regions and selected events

4.1.1 Mediterranean Sea

The Mediterranean region is characterized by intense tectonic activity mainly driven by the still active convergence between Africa and Europe, with a complex deformation pattern that includes smaller plates and blocks forming promontories and tectonic subregions (**Fig. 4.1**). The boundary between the African and Eurasian plates has different characteristics. This complex geological setting is associated with frequent and intense seismicity, which includes subduction-related earthquakes in the Calabrian, Hellenic, and Cyprian arcs, as well as crustal events, on thrust (e.g., Maghrebides, Dinarides, Eastern Alps), normal (e.g., the Messina Straits, the Apennines, the Corinth Gulf), and strike-slip faults (e.g., the Kefalonia-Lefkada, the North Anatolia faults). This complex tectonic setting is also associated with significant volcanism, including syn- and post-orogenic volcanoes (see Selva et al. 2021a and references therein).

Two specific events have been selected in this region: the 2003 M6.8 Zemmouri-Boumerdes (Algeria) earthquake in the western Mediterranean, and the 2017 Mw 6.6 Kos-Bodrum earthquake in the eastern Mediterranean. Both these events are near-coast earthquakes that caused

significant damage from both seismic and tsunami waves. The 2003 Mw 6.8 Zemmouri-Boumerdes earthquake occurred on the Tell-Atlas fold-and-thrust belt and triggered a tsunami that hit several harbours in the western Mediterranean (e.g., Alasset et al., 2006; Sahal et al., 2009; Heidarzadeh & Satake, 2013). The 2017 Mw 6.6 Kos-Bodrum earthquake occurred between the Bodrum Peninsula in Turkey and the Greek Island of Kos, in the eastern Aegean Sea, with E–W trending normal fault consistent with the general N–S extension in the SE Aegean region, and caused two deaths and hundreds of injuries, significant structural damage in Bodrum and Kos Island, and a local tsunami with wave heights reaching 1.4 m (Heidarzadeh et al. 2017; Konca et al., 2019; Dogan et al., 2019).

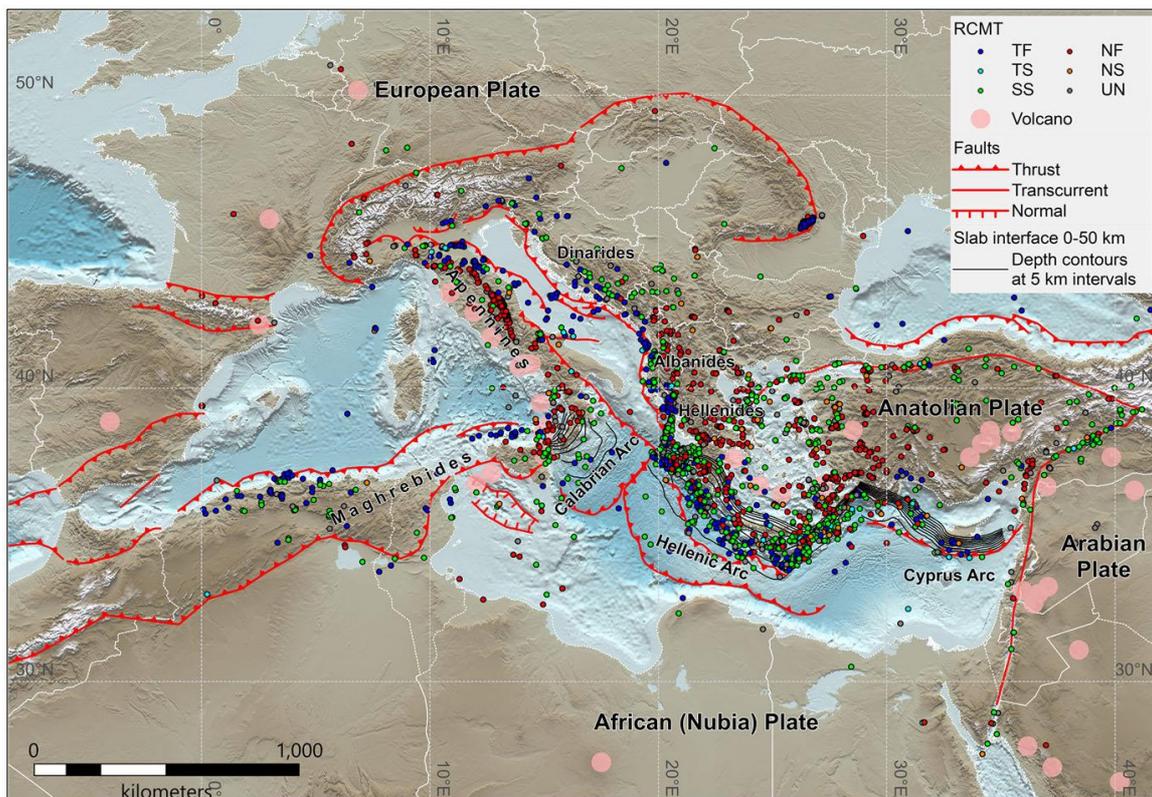


Fig. 4.1 Tectonic setting of Mediterranean region (modified from Selva et al., 2021a)

4.1.2 Mexico

México is a natural seismic laboratory located in a tectonically active setting (**Fig. 4.2**). Most of the recorded seismicity is due to the interaction between five major tectonic plates (**Fig. 4.2a**), namely: North American, Pacific, Cocos, Rivera and Caribbean. The Mexican National Seismological Service (SSN, Spanish acronym, www.ssn.unam.mx) which is the agency responsible for providing information about the Mexican earthquakes.

The selected event for specific case study of the seismic workflow is the M7.1 19 September Puebla 2017 earthquake. The 2017 Mw 7.1 Puebla event occurred on the same day but 32 years after the destructive 1985 Mw 8.1 Michoacan earthquake that killed more than 9,000 people and left more than 100,000 homeless. The hypocentral location of the Puebla event was 18.40°N, 98.72°W and 57 km depth (**Fig. 4.2b**).

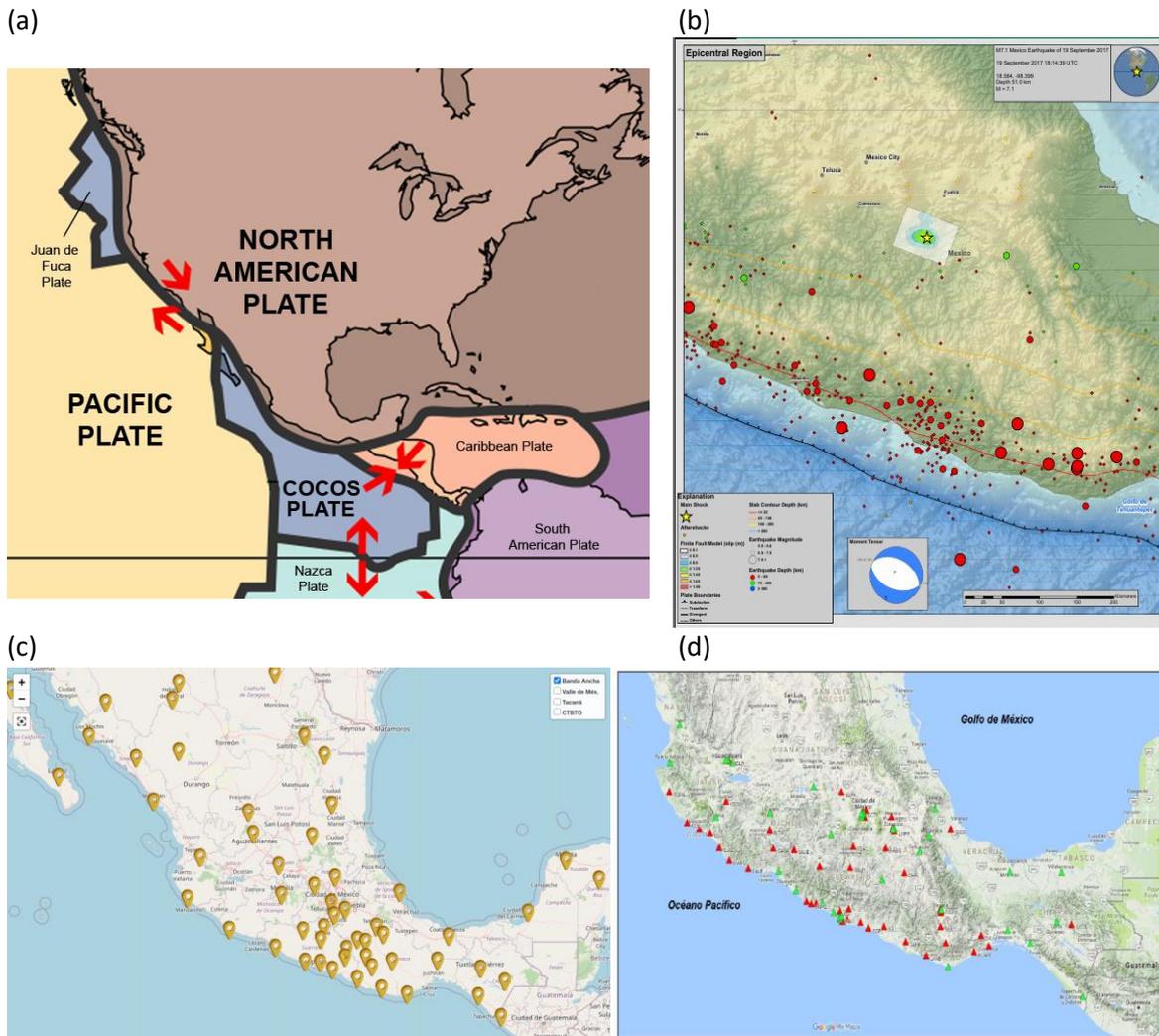


Fig. 4.2 (a) Schematic representation of the major tectonic plates in Mexico. (b) seismicity map updated by U.S. Geological Survey National Earthquake Information Center, the yellow star indicates the epicentral location of the M7.1 19 September Puebla 2017 earthquake. (c) SSN Broadband seismic network (<http://www.ssn.unam.mx/acerca-de/estaciones/>). (d) Stations of the Accelerograph Network of the Engineering Institute (red triangles). Green triangles are the seismic stations in real time transmission (Ramírez Guzmán et al., 2017).

Focal mechanism solutions indicate that the earthquake occurred on a moderately dipping normal fault, striking either to the southeast, or to the northwest. It was the best inland recorded event ever by different regional networks, the SSN (**Fig. 4.2c**), and the accelerograph stations from the Engineering Institute and the Institute of Geophysics at Autonomous University of Mexico (UNAM) (**Fig. 4.2d**). The earthquake was a severe event (VIII) in the Modified Mercalli Intensity scale with a total of 370 killed people and around 6000 injured and elevated structural damage. The accelerographic register is public and can be downloaded from the Engineering Institute database (<http://aplicaciones.iingen.unam.mx/AcelerogramasRSM>: registration required).

4.1.3 Iceland

Iceland is a superstructural part of the Mid-Atlantic Ridge situated in the North Atlantic Ocean where the Icelandic Hot Spot, a broad, localized upwelling of magma from deep within the mantle, elevates the seafloor so that it is partly exposed and forms land. The Ridge marks the boundary

between the North American and the Euro-Asian Plate and creates a belt of seismic activity from the Azores in the south towards Jan Mayen in the north. Across Iceland from southwest to the north, the plate boundary is displaced to the east through two major fracture zones, the South Iceland Seismic Zone (SISZ) in the south and the Tjörnes Fracture Zone in the north (TFZ, **Fig. 4.3**, Solnes et al., 2004). Iceland is the most seismically active region in northern Europe. The largest historic earthquakes in Iceland have occurred within these zones and have exceeded magnitude 7.

South Iceland is a relatively densely populated farming area with many small towns and several critical infrastructures, such as hydropower plants and associated reservoirs, geothermal power plants, industrial plants and transportation infrastructures (Panzera et al., 2016). Therefore the importance of studying the affectation of earthquakes occurring in this region is highly relevant. In Southern Iceland, two tectonics regions are defined: the South Iceland Seismic Zone (SISZ) and the Reykjanes Peninsula Oblique Rift (RPOR) take up the transform motion of the eastward jump of the Mid-Atlantic Ridge in southwest Iceland. Instead of a sinistral transform fault system, linear and parallel to the direction of plate spreading, a “bookshelf” faulting system of near vertical dextral transform faults perpendicular to the direction of plate spreading takes up the transform motion across the SISZ and RPOR (Einarsson, P., & Björnsson, S., 1979; Kowsari et al., 2020).

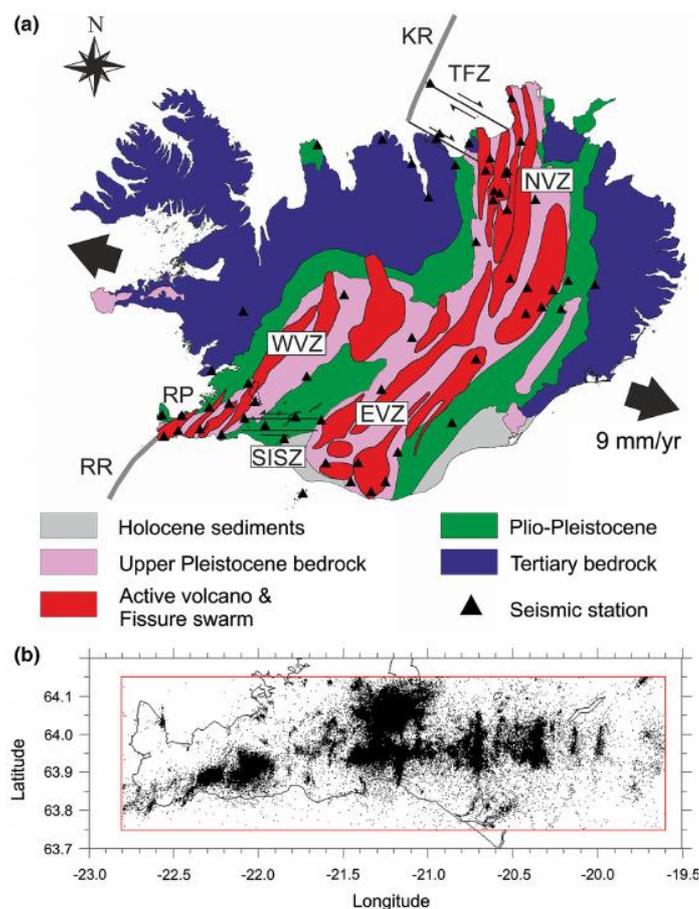


Fig. 4.3 a) Schematic representation of the plate boundary in Iceland (modified from ANGELIER et al. 2004). b) Map of the study area. Black dots events in the selected catalogue. EVZ Eastern Volcanic Zone, NVZ Northern Volcanic Zone, KR Kolbeinsey Ridge, RP Reykjanes Peninsula, RR Reykjanes Ridge, TFZ Tjörnes Fracture Zone, SISZ South Iceland Seismic Zone, WVZ Western Volcanic Zone (Panzera et al., 2016).

The SISZ has been the location of large destructive earthquakes in the past, relieving the tectonic stress that accumulates on the transverse shift in the Mid-Atlantic Ridge during the drift of the North American and Eurasian plates (Halldorsson et al., 2007). The City of Reykjavik and surrounding towns are located close to the seismic delineation of the Reykjanes Peninsula where the Mid Atlantic Ridge extends into the peninsula from the southwest. Moreover, the South Icelandic Seismic Zone is within 70 km distance of the Capital (Solnes et al., 2000). The Mw 6.5 earthquake of 21 June 2000, selected as a Use Case, is located in the SISZ with a hypocentral location 63.98°N, 20.71°W and 5.1km depth. The Harvard CMT solution has a strike of 20°, dip 85° and rake 173° (Pedersen et al., 2003).

4.1.4 Chile

Chile is located along the plate boundary where the oceanic Nazca plate subducts beneath the South America plate for more than 6,000 km, from the northern border with Peru to the southern Chilean margin, the latter characterized by the triple junction with the Antarctic plate (Fig. 4.4a). The relative converging movement between the Nazca and the South America plates is responsible for several geologic processes such as the formation of the Andes Mountains together with copious volcanic and seismic activity. Many great earthquakes occurred off-shore Chile (Hayes et al, 2017). Among the others, we recall the 1960 Valdivia earthquake (Mw 9.5), the 2010 Maule earthquake (Mw8.8), the 2014 Iquique earthquake (Mw 8.1), and the 2015 Illapel earthquake (Mw 8.3).

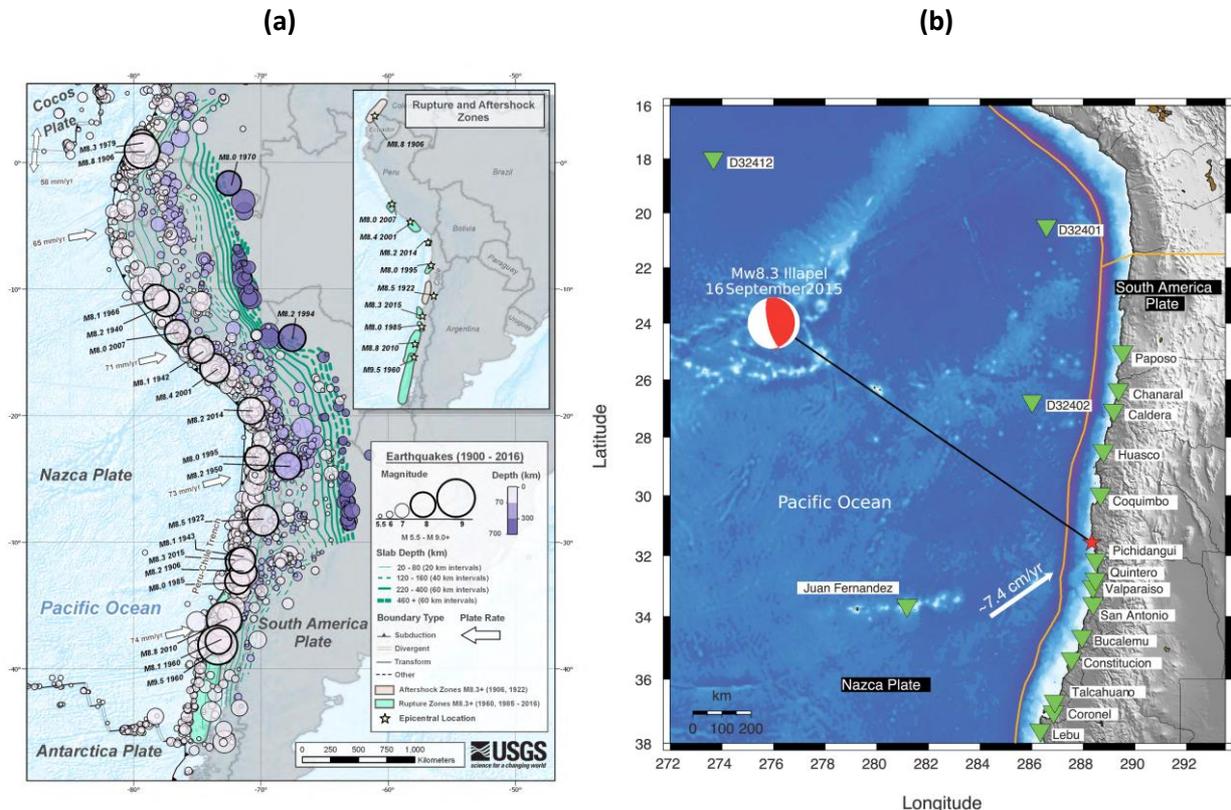


Fig 4.4 (a) Tectonic settings and seismicity of Nazca Plate Region (from USGS event page). (b) Location of the Illapel epicenter together with tide gauge and DART stations (from Romano et al., 2016).

The event selected as a use case in eFlow4HPC is the September 16, 2015 M 8.3 earthquake west of Illapel, Chile, which occurred as the result of thrust faulting on the interface between the Nazca and South America plates in Central Chile (**Fig. 4.4b**). At the location of this event, the Nazca plate is moving towards the east-northeast at a velocity of 74 mm/yr with respect to South America, and begins its subduction beneath the continent at the Peru-Chile Trench, 85 km to the west of the Illapel's epicenter. Several finite fault solutions have been proposed for this earthquake (Satake & Heidarzadeh, 2017).

4.2 The Seismic Workflow (UCIS4EQ)

In this section we discuss the specific datasets collected for each of the seismic use cases. Some data repositories are use-case specific, while other repositories or catalogs are held in common, that is, a subset of a global dataset is retrieved for the given use case. In Section 4.2.1 we first describe the shared datasets/repositories and their formats, while in Sections 4.2.2-4.2.4 we list the details for each use case.

4.2.1 Global datasets/repositories

Topography and bathymetry

Information about the topography and bathymetry is extracted from three different sources for each of the use cases (where available). All three are in netCDF file format.

- **Earth2014 and EGM2008:** The prepared lower resolution global topography and bathymetry files are common to all seismic use cases, with the relevant sections of the globe extracted for a given region at the stage of mesh generation. We use the Earth2014 global topography (relief) model (Hirt and Rexer, 2015; <https://www.bgu.tum.de/iapg/forschung/topographie/earth2014/>) and the Earth Gravitational Model 2008 (EGM2008, Pavlis et al., 2012; https://en.wikipedia.org/wiki/Earth_Gravitational_Model). Specifically, spherical harmonic coefficients for the Earth surface (Earth2014_SUR) and the for the Topo bedrock, bathymetry, and ice (Earth2014_TBI) are used from the Earth2014, while the Geoid undulations with respect to WGS84 are used from EGM2008. The preprocessed Earth2014 and EGM2008 datasets are provided by Mondaic for download in netCDF file format, and we use the finest provided spatial resolution at 1.8 km. See the following website for a detailed description and download links: https://mondaic.com/docs/0.11.35/data/seismology/global_topography_data.
- **SRTM:** NASA's Shuttle Radar Topography Mission provides surface topography elevation measurements at 90m and 30m resolution for most of the globe (Farr et al., 2007). The data is provided with reference to the WGS84 ellipsoid. It is available for download for a specified region in netCDF format via the AppEEARS web service run by the USGS (<https://lpdaacsvc.cr.usgs.gov/appears/>).
- **GMRT:** The Global Multi-Resolution Topography (GMRT; Ryan et al. (2009)) is "a multi-resolutional compilation of edited multibeam sonar data collected by scientists and institutions worldwide, that is (...) merged into a single continuously updated compilation

of global elevation data” (<https://www.gmrt.org/>). The data (with reference to the WGS84 ellipsoid) can be downloaded for a specified region via the GMRT MapTool in netCDF file format (<https://www.gmrt.org/GMRTMapTool/>).

Given the low maturity of the workflow we include three datasets: each offers different advantages and can be employed with relative ease for different mesh types. We have not taken the final decision on the target production dataset and we continue exploring the three.

Faulting mechanism forecast and CMT Catalog

Presently, UCIS4EQ faulting mechanism forecast is based on the CMT catalogue. IRIS’ Searchable Product Depository (SPUD; Trabant et al., 2012) provides Global Centroid-Moment-Tensors from the GCMT project (Ekström et al., 2012) at Lamont-Doherty Earth Observatory. Initial quick CMT solutions are available within minutes at SPUD, and are later updated to GCMT solutions when updates arrive. The SPUD Moment Tensor Product Query is available at <http://ds.iris.edu/spud/momenttensor> and data subsets can be downloaded in XML format.

Another global forecasting model, based on the GCMT catalogue, has been recently discussed and tested in Taroni and Selva (2021) and could be incorporated in the workflow in the future. The forecast models are available in ASCII tables and binary format (Matlab file), from Taroni and Selva (2021; <https://github.com/MatteoTaroniINGV/Testable-Worldwide-Earthquake-Faulting-Mechanism-Model>) and Selva et al. (2021c). At regional scale, other models may be available, eventually derived from regional or local hazard quantifications.

Infrastructure locations

The infrastructure locations including cities, hospitals, and critical facilities are taken from the GeoNames that is the geographical database that covers all countries and contains over eleven million place names that are available for download free of charge (<https://www.geonames.org/>).

Computational meshes

The meshes are generated for a given region with a regional velocity model (1D or 3D) and can, but do not have to, include the topography and bathymetry. They are generated with the Salvus software suite provided by Mondaic (<https://mondaic.com>) that supports the HPC wave propagation simulations in the workflow. The meshes are in HDF5 file format.

The mesh determines the maximum possible resolved frequency of a given simulation. Therefore, mesh sizes vary not only with the domain size, but also with the desired frequency content.

4.2.2 Data for the Mediterranean Sea

Region-specific Data

- **Regional velocity model:** 3D S-wave velocity (SV and SH) model for Eastern Mediterranean extracted from the Collaborative Seismic Earth Model (CSEM; Fichtner et al., 2019; <https://polybox.ethz.ch/index.php/s/NZCFUImMtR1LSWJ>). The gridded Earth model is provided as a netCDF file conforming with the format of the IRIS Earth Model Collection (EMC; <http://ds.iris.edu/ds/products/emc/>). Moreover, we also collect the 1D velocity model for the region obtained from earthquake relocations in Konka et al., (2019).

- **Topography and bathymetry:** Earth2014 and EGM2008, SRTM and GMRT datasets in netCDF.
- **CMT catalog:** The catalogue has a total of **622** events recorded in the region of interest from 2011 to the present. The catalogue includes earthquakes of magnitude in the range of $4.6 < M_w < 7.0$, located between 31.3° - 40.0° latitude, and 14.87° - 29.68° longitude.
- **Receivers and infrastructure locations:**
 - The receiver locations corresponding to existing seismometers and accelerographs in the Mediterranean area are retrieved from the AFAD in Turkey (<https://deprem.afad.gov.tr/istasyonlar?lang=en#>), from the Hellenic Unified Seismic Network (<http://www.gein.noa.gr/en/networks/husn>) and from the European Integrated Data Archive (<https://www.orfeus-eu.org/data/eida/>). Format: the locations, along with station and network codes, are collated in the yaml file (key-value format).
 - The location of cities and critical infrastructures are collected from the GeoNames database (<https://www.geonames.org/>) and Geographical Information System of the European Commission (GISCO).

Event-specific Data: the 2017 M6.8 Kos-Bodrum earthquake

- **Regional seismic waveforms, time location and magnitude estimation:** Seismic waveforms are downloaded via the FSDN service. From the earthquake epicenter, we mined all broadband data available in a circular area of 80 degree for all networks. The length of the data is about 1800 seconds after event origin time. We filtered the dataset for only the stations generally used by Early-Est and available in real time (See table `rt_ee_miniseed_station_coordinates.csv`). The waveform data are in mseed file format. The Real Time location estimate for the 2017 Kos-Bodrum earthquake was produced by Early-Est at the INGV, using the version 1.1.9. The location's parameters are listed into an ASCII table. A location map is also provided in pdf file format, generated in real time. For these events, the location covariance matrix is not available since it is not stored after computation. Off-line relocation has been produced also in hindcasting mode, using the data collected via FSDN applying version 1.1.9 of Early-Est. The tables of the relocated events include the location covariance matrix. Regarding magnitude estimation, Early-Est provides mb, Mwp and Mwpc magnitude estimations. The rules to select the appropriate magnitude type are described in Bernardi et al. (2015) and updated in Amato et al. (2021). For the relocated event, we provided only the appropriate magnitude type values.
- **Real-time moment tensor:** Data from global and local institutions have been retrieved, and release time is added when available. As global sources, we considered USGS (which includes wPHASE, Global CMT, and USGS-NEIC body waves solutions, <https://earthquake.usgs.gov/earthquakes/search/>; quakeML format) and GFZ-Geophon (<https://geofon.gfz-potsdam.de/>, txt and quakeML formats), while local data (coming from other seismic networks, including the local ones) are retrieved through EMSC-CSEM services (https://www.emsc-csem.org/Earthquake/index_tensors.php, txt mixed format). For this event, EMSC-CSEM services include GFZ, GCMT, INGV-RCMT, IPGP, KOERI, NOA, CPPT, AUTH, UOA, USGS. Apart from data, also several informative PDF files provided by the same agencies have been collected. For this event, we also retrieved INGV (quakeML

format, <http://terremoti.ingv.it/en>) and Regional CMT solutions (quakeML and json formats; txt for the quick solution; <http://rcmt2.bo.ingv.it/>).

- **Finite fault:** The information for the finite fault slip model is taken from [Konca et al., \(2019\)](#).
- **Local strong motion data:** The strong motion data can be retrieved from i) AFAD through the provided event service (<https://tadas.afad.gov.tr/list-event>) in the near source epicentral distance (range 0-200 km) and ii) from the Engineering Strong Motion DB (<https://esm-db.eu/>). ESM gathers the strong motion data available through EIDA (and other networks not available in EIDA) and proceeds with manual processing. The broadband data can be retrieved using the EIDA web services provided through ORFEUS (<https://www.orfeus-eu.org/data/eida/webservices/>). For this event, data is gathered from the Engineering Strong Motion database (through the webpage https://esm-db.eu/#/event/EMSC-20170720_0000091) and from AFAD (<https://tadas.afad.gov.tr/list-event>).
 - AFAD provides three-component strong motion data from 39 stations within 200 km in three different formats - ascii, asdf and miniSEED.
 - ESM provides three-component strong motion data from 44 stations belonging to the FDSN networks with codes GE, HI, HL and KO. The data are in three different formats - ascii, asdf and miniSEED.
- **Shakemaps:** The USGS provides the information related to their Shakemap for different intensity measures including uncertainty information for this event. (<https://earthquake.usgs.gov/earthquakes/eventpage/us20009ynd/shakemap/intensity>)
- **Reference** **GMPEs:** (see file <https://earthquake.usgs.gov/product/shakemap/us20009ynd/atlas/1594399583291/download/info.json>)

4.2.3 Data for Mexico

Region-specific Data

- **Regional Velocity model:** 1D velocity and 2-D S-wave velocity model obtained from the inversion of local dispersion curves that were reconstructed from the tomographic solutions ([Iglesias et al., 2010](#)).
- **Topography and bathymetry:** Earth2014 and EGM2008, SRTM and GMRT datasets in netCDF.
- **Receivers and infrastructure locations:** The location of cities and critical infrastructures are collected from the GeoNames database (<https://www.geonames.org/>). Moreover, we include complementary data from governmental agencies such as National Institute of Statistic and Geography (<http://en.www.inegi.org.mx/default.html>).
- **CMT catalog:** A total of 1824 events from 1950 to date, in XML format.

Event-specific Data: the 2017 M7.1 Puebla earthquake

- **Regional seismic waveforms, time location and magnitude estimation:** Seismic waveforms are downloaded via FSDN service. From the earthquake epicenter we mined all broadband data available in a circle area of 80 degree for all networks. The length of the data is about 1800 seconds after event origin time. We filtered the dataset for only the stations generally used by Early-Est and available in real time (See table [rt_ee_miniseed_station_coordinates.csv](#)). The waveform data are in mseed file format. Real Time location for the 2017 Puebla earthquake was produced by Early-Est at the INGV, using the version 1.1.9. The location's parameters are listed into an ASCII table. A location map is also provided in pdf file format, generated in real time. For these events the location covariant matrix is not available since it is not stored after computation. Off line relocation has been produced also in hindcasting mode, using the data collected via FSDN applying version 1.1.9 of Early-Est. The tables of the relocated events include the location covariant matrix.
- **Real-time moment tensor:** Data from global and local institutions have been retrieved, and release time is added when available. As global sources, we considered USGS (which includes wPHASE, Global CMT, and USGS-NEIC body waves solutions, <https://earthquake.usgs.gov/earthquakes/search/>; quakeML format) and GFZ-Geophon (<https://geofon.gfz-potsdam.de/>, txt and quakeML formats), while local data (coming from other seismic networks, including the local ones) are retrieved through EMSC-CSEM services (https://www.emsc-csem.org/Earthquake/index_tensors.php, txt mixed format). Apart from data, also several informative PDF files provided by the same agencies have been collected. For this event, EMSC-CSEM services include GFZ, GCMT, USGS, and CPPT.
- **Finite fault:** The Finite Fault solution is collected from the United States Geological Survey (USGS) agency (<https://earthquake.usgs.gov/earthquakes/eventpage/us2000ar20/finite-fault>).
- **Local Strong motion data:** The Strong motion data is downloaded from RAII-UNAM Strong Motion open source Database from Engineering Institute (<http://aplicaciones.iingen.unam.mx/AcelerogramasRSM/Inicio.aspx>). In the case of broadband stations the data must be retrieved through a petition to SSN (ongoing work). We have downloaded the data and we are in contact with RAII-UNAM to seek a formal agreement about the use of these data in the project.
- **Shakemaps:** shakemaps are obtained from two sources, USGS (<https://earthquake.usgs.gov/earthquakes/eventpage/us2000ar20/shakemap/intensity>) and [Ramírez Guzmán et al., \(2017\)](#) report.
- **Reference GMPEs:** Four Ground Motion Prediction Equations (GMPE's) are retrieved from [Çelebi et al., \(2018\)](#) and references therein. See also those used by USGS for the shakemap generation in the file (<https://earthquake.usgs.gov/product/shakemap/us2000ar20/atlas/1594399976203/download/info.json>)

4.2.4 Data for Iceland

Region-specific Data

- **Regional Velocity model:** For the SISZ-RPOR region we have 1D and 3D velocity models taken from [Tryggvason et al., \(2002\)](#). The data is in plain-text format.
- **Topography and bathymetry:** Earth2014 and EGM2008 and GMRT datasets. SRTM data is not available for this region.
- **Receivers and infrastructure locations:** File in plain-text format that contains the location of seismic receivers and cities. The location of cities and critical infrastructures are collected from the GeoNames database (<https://www.geonames.org/>). The seismic receivers locations are provided by the data published in [Sonnemann \(2019\)](#).
- **CMT catalog:** The catalog has a total of **164** recorded events from 1976 to the present in the magnitude range $4.6 < M_w < 6.5$. The events in the dataset are in a domain delimiting by 62.48° and 67.66° latitude, and -25.14° and -16.29 longitude.

Event-specific Data: the 2000 June 21, M6.5 SISZ earthquake

- **Regional seismic waveforms, time location and magnitude estimation:** Seismic waveforms are downloaded via the FSDN service. From the earthquake epicenter, we mined all broadband data available in a circle area of 80 degree for all networks. The length of the data is about 1800 seconds after event origin time. We filtered the dataset for only the stations generally used by Early-Est and available in real time (See table `rt_ee_miniseed_station_coordinates.csv`). The waveform data are in mseed file format. Real Time location for the 2000 SISZ earthquake is not available, while an off line relocation has been produced also in hindcasting mode, using the data collected via FSDN applying version 1.1.9 of Early-Est. The tables of the relocated events include the location covariant matrix.
- **Real-time moment tensor:** Data from global and local institutions have been retrieved, and release time is added when available. As global sources, we considered USGS (which includes wPHASE, Global CMT, and USGS-NEIC body waves solutions, <https://earthquake.usgs.gov/earthquakes/search/>; quakeML format) and GFZ-Geophon (<https://geofon.gfz-potsdam.de/>, txt and quakeML formats), while local data (coming from other seismic networks, including the local ones) are retrieved through EMSC-CSEM services (https://www.emsc-csem.org/Earthquake/index_tensors.php, txt mixed format). Apart from data, also several informative PDF files provided by the same agencies have been collected. For this event, only USGS solutions (GCMT and body-waves) are available, as the other services were not available at the time of the event.
- **Finite fault:** In [Pedersen et al., \(2003\)](#) a finite fault slip distribution is estimated from joint inversion of InSAR and GPS measurements. The format is in plain text.
- **Local Strong motion data:** The University of Iceland, Engineering Research Institute, Applied Mechanics Laboratory, Reykjavik, provide the ground acceleration records from 24 stations (http://isesd.hi.is/ESD_Local/frameset.htm). We have downloaded the data and we are in contact with the Engineering Research Institute to seek a formal agreement about the use of these data in the project.

- **Shakemaps:** USGS shakemaps for this event are available at <https://earthquake.usgs.gov/earthquakes/eventpage/usp0009ux8/shakemap/intensity>
- **Reference GMPEs (?):** Six different GMPEs for the SISZ region are available in [Kowsari et al., \(2020\)](#) . The USGS shakemaps have used the following the GMPEs included in the json file available at <https://earthquake.usgs.gov/product/shakemap/usp0009ux8/atlas/1594169845271/download/info.json>

4.2.5 Estimation of Computational Resources

The computational requirements of UCIS4EQ vary not only with the targeted frequency, but also with the domain size and with the number of instances launched. As the workflow is not mature, the estimates provided below are subject to change with further technical developments.

- **Size of input (simulation settings) for the region:** The size of the velocity models is below 10 MB and are stored as netCDF files. The size of the regional bathymetry and topography maps is below 1 MB and are stored in netCDF file format as well. It should be noted that these files are necessary only to generate computational meshes that are then the direct input for the HPC simulations. The computational meshes are stored in HDF5 file format and the size of these files depends on the frequency used in the simulation and on the target area within the given region. Typically, the size will be below 5 GB for a maximum resolved frequency of 5 Hz, below 50 GB for a maximum resolved frequency of 10 Hz, and below 250 GBs for a maximum resolved frequency of 20 Hz. Sample meshes for maximum resolved frequency of 1 and 5 Hz have been generated for the Mediterranean use case and can be found in b2drop repository.
- **Size of input (seismic source) for the specific event:** Fault kinematic source description. This data is stored as a key-value plain text file. The size of this file is typically below 1 GB. UCIS4EQ launches in parallel HPC simulations considering different fault kinematic sources coming from statistical results and assimilated data (around 50-100 different instances, details TBD).
- **Job size and approximate total number of jobs:** The job size depends on the domain size, the target frequency and the seismogram length. For example, on the MareNostrum4 supercomputer, a simulation with a maximum resolved frequency of 5 Hz required 2160 ranks (48 ranks per node) and completed in 3.5 hours for 50 seconds of elastic wave propagation.
- **Required local memory:** TBD. This is subject to changes in the software.
- **Size of output for decision making use (extraction for maps):** TBD.
- **Size of output for second use (all output files):** For a 5Hz simulation the current minimum size of output files per each job is ~2GB.

4.3 The Tsunami Workflow (PTF/FTRT)

4.3.1 Global datasets/repositories

Faulting mechanism forecast

As already described in Section 4.2.1 for the seismic workflow, a faulting mechanism forecast is fundamental to initializing a fault geometry uncertainty distribution when only location and magnitude data are available. This is at the base of the PTF development for early-warning (Selva et al., 2021b). A global forecasting model has been recently discussed and tested in Taroni and Selva (2021), based on the GCMT catalogue. The forecast models are available in binary format (Matlab file), from Selva et al. (2021c). At regional scale, other models may be available, eventually derived from regional or local hazard quantifications.

4.3.2 Data for the Mediterranean Sea

Region-specific Data

- **Topo-bathymetric data:** Different data sources have been used to construct the topo-bathymetric data for the Mediterranean Sea.
 - GEBCO's current gridded bathymetric data set, the GEBCO_2020 Grid, a global terrain model for ocean and land, providing elevation data, in meters, on a 15 arc-second interval grid. (More information can be found here: https://www.gebco.net/data_and_products/gridded_bathymetry_data/).
 - EMODnet Digital Terrain Model (DTM), generated for European sea regions (36W,15N; 43E,90N) from selected bathymetric survey data sets, composite DTMs and Satellite Derive Bathymetry (SDB) data products with grid resolution of 1/16 * 1/16 arc minutes (circa 115 * 115 meters) (More information can be found here: <https://portal.emodnet-bathymetry.eu/>).

The continuous topo-bathymetry model is provided with an interpolated resolution of 30 arcsec, with reference to the WGS84 ellipsoid, and in .grd format - netCDF3 CLASSIC (32-bit offset) storage format.

- **Forecast points:** a total of 1107 forecast points distributed all over the Mediterranean have been derived from the regional hazard model NEAMTHM18 (Basili et al. 2018, 2021). These data represent the Mediterranean subset of the Point of Interest adopted by the regional hazard model NEAMTHM18. They are stored in a text file (CSV format) and in a binary file (.mat Matlab format), and can be retrieved from the NEAMTHM18 documentation (<http://www.tsumaps-neam.eu/documentation/>) and from Selva et al. (2021b).
- **Source discretization:** the source discretization in the Mediterranean is derived from NEAMTHM18 (Basili et al. 2018, 2021). The data are in a binary file in Matlab format (Discretization.mat), and can be downloaded from NEAMTHM18 documentation (<http://www.tsumaps-neam.eu/documentation/>).

Event-specific Data: the 2017 M6.8 Kos-Bordum earthquake and tsunami

- **Regional seismic waveforms, time location and magnitude estimation:** as already described in Section 4.2.2, seismic waveforms are downloaded via FSDN service. From the earthquake epicenter we mined all broadband data available in a circle area of 80 degree for all networks. The length of the data is about 1800 seconds after event origin time. We filtered the dataset for only the stations generally used by Early-Est and available in real time (See table `rt_ee_miniseed_station_coordinates.csv`). The waveform data are in mseed file format. Real Time location for the 2017 Kos-Bodrum earthquake was produced by Early-Est at the INGV, using the version 1.1.9. The location's parameters are listed into an ASCII table. A location map is also provided in pdf file format, generated in real time. For these events the location covariant matrix is not available since it is not stored after computation. Off line relocation has been produced also in hindcasting mode, using the data collected via FSDN applying version 1.1.9 of Early-Est. The tables of the relocated events include the location covariant matrix.
- **Real-time moment tensor:** as already described in Section 4.2.2, data from global and local institutions have been retrieved, and release time is added when available. As global sources, we considered USGS (which includes wPHASE, Global CMT, and USGS-NEIC body waves solutions, <https://earthquake.usgs.gov/earthquakes/search/>; quakeML format) and GFZ-Geophon (<https://geofon.gfz-potsdam.de/>, txt and quakeML formats), while local data (coming from other seismic networks, including the local ones) are retrieved through EMSC-CSEM services (https://www.emsc-csem.org/Earthquake/index_tensors.php, txt mixed format). For this event, EMSC-CSEM services include GFZ, GCMT, INGV-RCMT, IPGP, KOERI, NOA, CPPT, AUTH, UOA, USGS. Apart from data, also several informative PDF files provided by the same agencies have been collected. For this event, we also retrieved INGV (quakeML format, <http://terremoti.ingv.it/en>) and Regional CMT solutions (quakeML and json formats; txt for the quick solution; <http://rcmt2.bo.ingv.it/>).
- **Tsunami records:**
 - waveforms from tide-gauges: raw data downloaded from the sea level station monitoring facility (<http://www.ioc-sealevelmonitoring.org>) and then filtered out from astronomical tide
 - Tsunami run-up data: the used dataset was retrieved from the datasets published after the post event field survey in Dogan et al. 2019

Event-specific Data: the 2003 M6.8 Zemmouri-Boumerdes earthquake and tsunami

- **Regional seismic waveforms, time location and magnitude estimation:** Seismic waveforms are downloaded via FSDN service. From the earthquake epicenter we mined all broadband data available in a circle area of 80 degree for all networks. The length of the data is about 1800 seconds after event origin time. We filtered the dataset for only the stations generally used by Early-Est and available in real time (See table `rt_ee_miniseed_station_coordinates.csv`). The waveform data are in mseed file format. Real Time location for the 2003 Zemmouri-Boumerdes earthquake is not available, while an off line relocation has been produced in hindcasting mode, using the data collected via FSDN applying version 1.1.9 of Early-Est. The tables of the relocated events include the

location covariant matrix. Regarding magnitude estimation, Early-Est provides mb, Mwp and Mw_{pd} magnitude estimations. The rules to select the appropriate magnitude type are described in (Bernardi et al. 2015) and updated in (Amato et al. 2021). For the relocated event, we provided only the appropriate magnitude type values.

- **Real-time moment tensor:** Data from global and local institutions have been retrieved, and release time is added when available. As global sources, we considered USGS (which includes wPHASE, Global CMT, and USGS-NEIC body waves solutions, <https://earthquake.usgs.gov/earthquakes/search/>; quakeML format) and GFZ-Geophon (<https://geofon.gfz-potsdam.de/>, txt and quakeML formats), while local data (coming from other seismic networks, including the local ones) are retrieved through EMSC-CSEM services (https://www.emsc-csem.org/Earthquake/index_tensors.php, txt mixed format). Apart from data, also several informative PDF files provided by the same agencies have been collected. For this event, only USGS solutions are available, as the other services were not available at the time of the event.
- **Tsunami records:**
 - Tide-gauges data: the used dataset was retrieved from the table published in Heidarzadeh & Satake (2013) containing the maximum wave heights (peak-to-trough) recorded at several available tide gauges for this event.
 - Synthetic tide-gauges: waveforms are synthetic sea level records obtained with T-HySea and using a set of 12 finite fault solutions described in Selva et al. (2021b)

4.3.3 Data for Chile

Region-specific Data

All the retrieved files are reported in the folder in B2Drop, in the Pillar III folder, in folder Tsunami/Chile.

- **Topo-bathymetric data:** Earth2014 global topography (relief) model with 1 arcmin of resolution is used to cover the whole Pacific with reference to the WGS84 ellipsoid, and in .grd format - netCDF3 CLASSIC (32-bit offset) storage format.
- **Forecast points:** a possible list of potential target points is defined in Selva et al. (2021b), and include all DART stations and all locations of local tide-gauges. The file is binary in Matlab format, and can be downloaded from Selva et al. (2021c).
- **Source discretization:** a possible source discretization strategy is defined in Selva et al. (2021b). The subduction zone mesh is a ASCII file, and can be downloaded from Selva et al. (2021c).
- **Available seismic and tsunami instrumentation in the region:** The coordinates of both DART stations and Tide-gauges can be found within the files topobath/DART_station_table_Final.dat and topobath/TG_station_table_Final.dat, respectively. These files have 3 columns (ID/Name, Lon, Lat).

Event-specific Data: the 2015 M8.2 Illapel (Chile) earthquake and tsunami

- **Regional seismic waveforms, time location and magnitude estimation:** Seismic waveforms are downloaded via FSDN service. From the earthquake epicenter we mined all broadband data available in a circle area of 80 degree for all networks. The length of the data is about 1800 seconds after event origin time. We filtered the dataset for only the stations generally used by Early-Est and available in real time (See table rt_ee_miniseed_station_coordinates.csv). The waveform data are in mseed file format. Real Time location for the 2015 Illapel earthquake was produced by Early-Est at the INGV, using the version 1.1.8. The location's parameters are listed into an ASCII table. A location map is also provided in pdf file format, generated in real time. For these events the location covariant matrix is not available since it is not stored after computation. Off line relocation has been produced also in hindcasting mode, using the data collected via FSDN applying version 1.1.9 of Early-Est. The tables of the relocated events include the location covariant matrix.
- **Real-time faulting mechanism:** Data from global and local institutions have been retrieved, and release time is added when available. As global sources, we considered USGS (which includes GCMT, <https://earthquake.usgs.gov/earthquakes/search/>; quakeML format) and GFZ-Geophon (<https://geofon.gfz-potsdam.de/>, txt and quakeML formats), while local data (coming from other seismic networks, including the local ones) are retrieved through EMSC-CSEM services (https://www.emsc-csem.org/Earthquake/index_tensors.php, txt mixed format). For this event, EMSC-CSEM services include EOT, GFZ, GCMT, IPGP, and USGS.
- **Tsunami records:**
 - DART signals are reported in the tsunami subfolder, in a folder named `sea_level_data/2015-09-16-Illapel/DART_ILLAPEL_signals_detide`; each file has 2 columns (time in minutes from the earthquake origin time, wave amplitude in meters). Each waveform has been obtained by removing with a LOWESS (Romano et al., 2016) procedure the tidal component from the original recorded signal (<https://nctr.pmel.noaa.gov/Dart/>).
 - Tide-gauge signals are reported in the tsunami subfolder, in a folder named `sea_level_data/2015-09-16-Illapel/TG_ILLAPEL_signals_detide`; each file has 2 columns (time in minutes from the earthquake origin time, wave amplitude in meters). Each waveform has been obtained by removing with a LOWESS (Romano et al., 2016) procedure the tidal component from the original recorded signal (<http://www.ioc-sealevelmonitoring.org>).
 - RUNUP data are reported in the tsunami subfolder, in a folder named `runup_data/2015-09-16-Illapel/RUNUP_ILLAPEL_NGDC`; each file has 3 columns (longitude, latitude, runup in meters). The runup observations have been obtained by the NGDC/WDS Global Historical Tsunami Database (https://www.ngdc.noaa.gov/hazard/tsu_db.shtml).

4.3.4 Estimation of Computational Resources

The computational requirements of PTF/FTRT vary with the domain size and the number of target points for which to store the output information. It may also strongly vary if nested grids are used to increase the resolution of the simulation in the target area. Here, we consider the simplest configuration, and the numbers are subject to change with further technical developments.

- **Size of input (simulation settings) for the region:** For the Mediterranean region, ~50Mb (30arcsec Mediterranean grid), covering the entire Mediterranean region. For Chile, ~360 Mb (1arcmin Pacific grid), covering the entire Pacific Ocean.
- **Size of input (seismic source) for the specific events:** < 1MB
- **Approximate total number of jobs:** 10,000 to 100,000 events (event dependent)
- **Single Job size:** < 1Mb / TBD
- **GPU hours per job:** for a single THySEA job/simulation, the GPU hours depend (among others) on the simulation time. For example, in the case of the Mediterranean-8h-30arcsec simulation, it takes around 5 minutes. In contrast, in the Pacific, where we typically simulate 40 hours of simulation time, it takes around 2 hours of runtime.
- **Required local memory:** For the Mediterranean region, 4GB per GPU. For Chile, 6GB per GPU.
- **Size of output for decision making use (extraction for maps):** < 100 MB
- **Size of output for second use (all output files):** 50GB-4TB (event dependent: output variables, simulation time, number of pois, etc.

5. Conclusions

In this deliverable we collected a series of datasets that enable, in principle, the application of the seismic and the tsunami workflows in the selected areas, for whatever seismic event occurring in these areas. We also collected the data that will be used to implement specific events in these areas.

The collected datasets represent a starting point, to guarantee the applications in the use cases. It will be possible to extend or update such data in the future, if found appropriate during the actual implementation of the use cases.

6. Acronyms and Abbreviations

Term or abbreviation	Description
CF	Climate and Forecast convention (https://cfconventions.org/)
CMT	Centroid Moment Tensor
ESM	Engineering Strong Motion DB (https://esm-db.eu/)
GCMT	Global CMT
GMRT	Global Multi-Resolution Topography
GMPE	Ground Motion Prediction Equation
GP	Graves-Pitarka algorithm
HPC	High Performance Computing
MLEsmap	Machine-Learning based Estimator for ground motion Shaking maps
NEAM	North-eastern Atlantic, the Mediterranean and connected seas
NEAMTHM18	NEAM Tsunami Hazard Model 2018
PTF/FTRT	Probabilistic Tsunami Forecasting / Faster-than Real Time tsunami simulations
RCMT	Regional CMT
RPOR	Reykjanes Peninsula Oblique Rift
SISZ	South Iceland Seismic Zone
SRTM	NASA's Shuttle Radar Topography Mission
SSN	Seismological Mexican Service
TFZ	Tjörnes Fracture Zone
UCIS4EQ	Urgent Computing Integrated Services for EarthQuakes
UNAM	Autonomous University of Mexico
USGS	U.S. Geological Survey

7. References

Alasset, P.-J., Hébert, H., Maouche, S., Calbini, V. & Meghraoui, M. The tsunami induced by the 2003 Zemmouri earthquake (MW= 6.9, Algeria): modelling and results. *Geophys J Int* 166, 213–226 (2006).

Amato A , Avallone A, Basili R, Bernardi F, Brizuela B, Graziani L, Herrero A, Lorenzino M.C., Lorito S, Mele F.M, Michelini A, Piatanesi A, Pintore S, Romano F, Selva J, Stramondo S, Tonini R, Volpe M; From Seismic Monitoring to Tsunami Warning in the Mediterranean Sea. *Seismological Research Letters* 2021;; 92 (3): 1796–1816. doi: <https://doi.org/10.1785/0220200437>

Basili R, Brizuela B, Herrero A, Iqbal S, Lorito S, Maesano FE, Murphy S, Perfetti P, Romano F, Scala A, Selva J , Taroni M, Tiberti MM, Thio HK, Tonini R, Volpe M, Glimsdal S, Harbitz CB, Løvholt F, Baptista MA, Carrilho F, Matias LM, Omira R, Babeyko A, Hoechner A, Gurbuz M, Pekcan O, YalcÄ±ner A, Canals M, Lastras G, Agalos A, Papadopoulos G, Triantafyllou I, Benchekroun S, Jaouadi HA, Ben Abdallah S, Bouallegue A, Hamdi H, Oueslati F, Amato A, Armigliato A, Behrens J, Davies G, Di Bucci D, Dolce M, Geist E, Gonzalez Vida JM, Gonzalez M, MacÃ±as Sanchez J, Meletti C, Ozer Sozdinler C, Pagani M, Parsons T, Polet J, Power W, Sørensen W, Zaytsev A (2020), The

making of the NEAM Tsunami Hazard Model 2018 (NEAMTHM18). , *Frontiers in Earth Science* 9:82, DOI : <https://doi.org/10.3389/feart.2020.616594>

Basili, R., Brizuela, B., Herrero, A., Iqbal, S., Lorito, S., Maesano, F. E., et al. (2018). NEAM tsunami hazard model 2018 (NEAMTHM18): online data of the probabilistic tsunami hazard model for the NEAM region from the TSUMAPS-NEAM project. Roma: Istituto Nazionale di Geofisica e Vulcanologia (INGV). doi:10.13127/tsunami/neamthm18

Bernardi, F., Lomax, A., Michelini, A., Lauciani, V., Piatanesi, A., and Lorito, S.: Appraising the Earliest earthquake monitoring system for tsunami alerting at the Italian Candidate Tsunami Service Provider, *Nat. Hazards Earth Syst. Sci.*, 15, 2019–2036, <https://doi.org/10.5194/nhess-15-2019-2015>, 2015.

Dziewonski, A. M., T.-A. Chou and J. H. Woodhouse, Determination of earthquake source parameters from waveform data for studies of global and regional seismicity, *J. Geophys. Res.*, 86, 2825-2852, 1981. doi:10.1029/JB086iB04p02825

Dogan, G.G., Annunziato, A., Papadopoulos, G.A. et al. The 20th July 2017 Bodrum–Kos Tsunami Field Survey. *Pure Appl. Geophys.* 176, 2925–2949 (2019). <https://doi.org/10.1007/s00024-019-02151-1>

Einarsson, P., & Björnsson, S. (1979). Earthquakes in Iceland. *Jökull*, 29, 37-43.

Ekmström, G., M. Nettles, and A. M. Dziewonski, The global CMT project 2004-2010: Centroid-moment tensors for 13,017 earthquakes, *Phys. Earth Planet. Inter.*, 200-201, 1-9, 2012. doi:10.1016/j.pepi.2012.04.002

Farr, T. G., et al. (2007), The Shuttle Radar Topography Mission, *Reviews of Geophysics*, 45, RG2004, doi:10.1029/2005RG000183.

Fichtner, A. et al. (2018), The Collaborative Seismic Earth Model: Generation 1, *Geophysical Research Letters*, 45 (9), 4007-4016, doi:10.1029/2018GL077338.

Franco, S. I., Iglesias, A., & Fukuyama, E. (2020). Moment tensor catalog for Mexican earthquakes: almost two decades of seismicity. *Geofísica internacional*, 59(2), 54-80.

Guidoboni, E., Ferrari, G., Tarabusi, G., Sgattoni, G., Comastri, A., Mariotti, D., ... & Valensise, G. (2019). CFTI5Med, the new release of the catalogue of strong earthquakes in Italy and in the Mediterranean area. *Scientific data*, 6(1), 1-15.

Halldorsson, B., Olafsson, S., & Sigbjörnsson, R. (2007). A fast and efficient simulation of the far-fault and near-fault earthquake ground motions associated with the June 17 and 21, 2000, earthquakes in South Iceland. *Journal of Earthquake Engineering*, 11(3), 343-370.

Hayes, G.P., Myers, E.K., Dewey, J.W., Briggs, R.W., Earle, P.S., Benz, H.M., Smoczyk, G.M., Flamme, H.E., Barnhart, W.D., Gold, R.D., and Furlong, K.P., 2017, Tectonic summaries of magnitude 7 and greater earthquakes from 2000 to 2015: U.S. Geological Survey Open-File Report 2016–1192, 148 p., <https://doi.org/10.3133/ofr20161192>

Heidarzadeh, M. & Satake, K. (2013) The 21 May 2003 Tsunami in the Western Mediterranean Sea: Statistical and Wavelet Analyses. *Pure Appl. Geophys.* 170,1449–1462

Heidarzadeh M., Necmioglu O., Ishibe T., Yalçiner A. (2017). Bodrum–Kos (Turkey–Greece) Mw 6.6 earthquake and tsunami of 20 July 2017: a test for the Mediterranean tsunami warning system, *Geosci. Lett.*, 4, 31.10.1186/s40562-017-0097-0

Hirt, C. and Rexer, M. (2015), Earth2014: 1 arc-min shape, topography, bedrock and ice-sheet models - available as gridded data and degree-10,800 spherical harmonics. *International Journal of Applied Earth Observation and Geoinformation*, 39, 103-112, <https://doi.org/10.1016/j.jag.2015.03.001>

Iglesias, A., R. W. Clayton, X. Pérez-Campos, S. K. Singh, J. F. Pacheco, D. García, and C. Valdés-González (2010), S wave velocity structure below central Mexico using high-resolution surface wave tomography, *J. Geophys. Res.*, 115, B06307, doi:10.1029/2009JB006332.

Kagan, Y. Y., and D. D. Jackson (2014). Statistical earthquake focal mechanism forecasts, *Geophys. J. Int.* 197, no. 1, 620–629

Konca, A. O., Guvercin, S. E., Ozarpaci, S., Ozdemir, A., Funning, G. J., Dogan, U., & Reilinger, R. (2019). Slip distribution of the 2017 M_w6.6 Bodrum–Kos earthquake: resolving the ambiguity of fault geometry. *Geophysical Journal International*, 219(2), 911-923.

Kowsari, M., Sonnemann, T., Halldorsson, B., Hrafinkelsson, B., Snæbjörnsson, J. Þ., & Jonsson, S. (2020). Bayesian inference of empirical ground motion models to pseudo-spectral accelerations of south Iceland seismic zone earthquakes based on informative priors. *Soil Dynamics and Earthquake Engineering*, 132, 106075.

Lomax, A., Michelini, A. & Curtis, A. Earthquake Location, Direct, Global-Search Methods. in *Encyclopedia of Complexity and Systems Science* (ed. Meyers, R. A.) 2449–2473 (Springer, 2009). doi:10.1007/978-0-387-30440-3_150.

Lomax, A. & Michelini, A. Mw_{pd}: A duration–amplitude procedure for rapid determination of earthquake magnitude and tsunamigenic potential from P waveforms. *Geophys J Int* 176, 200–214 (2009).

Monterrubio-Velasco, M., Carrasco-Jimenez, J.C., Rojas, O., Rodríguez, J. Fichtner, A., de la Puente, J. (in review *Frontiers in Earth Science*) A statistical approach towards fast estimates of moderate-to-large earthquake focal mechanisms.

Panzer, F., Zechar, J. D., Vogfjörð, K. S., & Eberhard, D. A. (2016). A revised earthquake catalogue for South Iceland. *Pure and Applied Geophysics*, 173(1), 97-116.

Pavlis, N. K., Simon A. Holmes, Steve C. Kenyon, and John K. Factor. (2012). The Development and Evaluation of the Earth Gravitational Model 2008 (EGM2008) *Journal of Geophysical Research: Solid Earth*, 117 (4): 1–38, <https://doi.org/10.1029/2011JB008916>

Romano, F., Piatanesi, A., Lorito, S., Tolomei, C., Atzori, S., & Murphy, S. (2016). Optimal time alignment of tide-gauge tsunami waveforms in nonlinear inversions: Application to the 2015 Illapel (Chile) earthquake. *Geophysical Research Letters*, 43, 11226–11235. doi:10.1002/2016GL071310.

Roselli, P., W. Marzocchi, M. T. Mariucci, and P. Montone (2018). Earthquake focal mechanism forecasting in Italy for PSHA purposes, *Geophys. J. Int.* 212, no. 1, 491–508.

Rovida, A., Mario, L., Romano, C., Lolli, B., & Gasperini, P. (2020). The Italian earthquake catalogue CPT15. *Bulletin of Earthquake Engineering*, 18(7), 2953-2984.

Ryan, W.B.F., S.M. Carbotte, J.O. Coplan, S. O'Hara, A. Melkonian, R. Arko, R.A. Weissel, V. Ferrini, A. Goodwillie, F. Nitsche, J. Bonczkowski, and R. Zensky (2009), Global Multi-Resolution Topography synthesis, *Geochemistry, Geophysics, Geosystems*, 10, Q03014, doi: [10.1029/2008GC002332](https://doi.org/10.1029/2008GC002332)

Sahal, A. et al. The tsunami triggered by the 21 May 2003 Boumerdès-Zemmouri (Algeria) earthquake: field investigations on the French Mediterranean coast and tsunami modelling. *Natural Hazards and Earth System Sciences* 9, 1823–1834 (2009).

Satake, K., Heidarzadeh, M. A Review of Source Models of the 2015 Illapel, Chile Earthquake and Insights from Tsunami Data. *Pure Appl. Geophys.* 174(1), 1–9 (2017). <https://doi.org/10.1007/s00024-016-1450-5>

Selva J, Tonini R, Molinari I, Tiberti MM, Romano F, Grezio A, Melini D, Piatanesi A, Basili R, Lorito S (2016), Quantification of source uncertainties in Seismic Probabilistic Tsunami Hazard Analysis (SPTHA), *Geophys. J. Int.*, 205: 1780-1803, DOI: 10.1093/gji/ggw107

Selva J, Amato A, Armigliato A, Basili R, Bernardi F, Brizuela B, Cerminara M, de' Micheli Vitturi M, Di Bucci D, Di Manna P, Esposti Ongaro T, Lacanna G, Lorito S, Løvholt F, Mangione D, Panunzi E, Piatanesi A, Ricciardi A, Ripepe M, Romano F, Santini M, Scalzo A, Tonini R, Volpe M, Zaniboni F (2021a), Tsunami risk management for crustal earthquakes and non-seismic sources in Italy. , *La Rivista del Nuovo Cimento* 44: 69-144, DOI : <https://doi.org/10.1007/s40766-021-00016-9>

Selva J, Lorito S, Volpe M, Romano F, Tonini R, Perfetti P, Bernardi F, Taroni M, Scala A, Babeyko A, Løvholt F, Gibbons S, Macias J, Castro MJ, Gonzalez-Vida JM, Sanchez-Linares C, Bayraktar HB, Basili R, Maesano FE, Tiberti MM, Mele F, Piatanesi A, Amato A (2021b): Probabilistic Tsunami Forecasting for Early Warning. *Nature Communications*, in press

Selva J, Lorito S, Volpe M, Romano F, Tonini R, Perfetti P, Bernardi F, Taroni M, Scala A, Babeyko A, Løvholt F, Gibbons S, Macias J, Castro MJ, Gonzalez-Vida JM, Sanchez-Linares C, Bayraktar HB, Basili R, Maesano FE, Tiberti MM, Mele F, Piatanesi A, Amato A (2021c): Data for: Probabilistic Tsunami Forecasting for Early Warning. figshare. Dataset. <https://doi.org/10.6084/m9.figshare.15015132>

Solnes, J., Sigbjörnsson, R., & Eliasson, J. (2004, August). Probabilistic seismic hazard mapping of Iceland. In *Proceedings of the 13th World conference on earthquake engineering, Vancouver, BC, Canada*.

Solnes, J., Sigbjörnsson, R., & Eliasson, J. (2000). Earthquake hazard mapping and zoning of Reykjavik.

Sonnemann, 2019. Earthquake Source Modelling and Broadband Ground Motion Simulation in South Iceland for Earthquake Engineering Applications, PhD thesis. Faculty of Civil and Environmental Engineering School of Engineering and Natural Sciences, University of Iceland, Reykjavik.

Trabant, C., A. R. Hutko, M. Bahavar, R. Karstens, T. Ahern, and R. Aster (2012), Data Products at the IRIS DMC: Stepping Stones for Research and Other Applications, *Seismological Research Letters*, 83(5), 846–854, <https://doi.org/10.1785/0220120032>.

Tryggvason, A., Rögnvaldsson, S. T., & Flóvenz, O. G. (2002). Three-dimensional imaging of the P- and S-wave velocity structure and earthquake locations beneath Southwest Iceland. *Geophysical Journal International*, 151(3), 848-866.