SEISMOSTORM

Making Analog Seismograms FAIR to Enable Research

DURATION
15/03/2021 – 15/03/2023

BUDGET
409.850 €

PROJECT DESCRIPTION

Context

Seismographs have recorded ground motion since the end of the XIXth century until digital recordings became available in the 1970s. Pre-digital seismic records were typically created by using ink on white paper, scratching black-smoked paper, or by using light on photographic paper.
With most of these legacy seismic data now stacked in archives, several projects in the last 20 years (e.i., EUROSISMOS project, Ferrari & Pino, 2004) started to dedicate resources to digitize them to preserve the scientific wealth they retain and, in turn, introduce them into the age of modern seismology.
This effort has become increasingly important due to the risk of permanently losing those aging paper seismograms combined with their recently found exclusive potential in recovering the global oceanic climate for the last century (Lecocq et al., 2020).

General objectives and underlying research questions

The ROB has an extensive archive of legacy seismic data that have recently been scanned. Our project aims to digitize the wiggles on the scanned images using computer vision and machine learning methodologies to convert them into calibrated and time-coded seismic times series. These will, in turn, be made publicly available on international web services following community-defined standards. This will give access to the scientific community to continuous seismic data covering the last century, with a particular interest in the study of past oceanic climates.

Methodology

The methodology has been separated into 4 steps spreading across 6 work packages:
The first step leverages the metadata on the former seismic instruments used at the Observatory to construct the corresponding frequency/amplitude instrument response functions (WP1). This will rely on the published seismic bulletins of the observatory to convert the digitized waveforms into ground motion.

The second step is dedicated to the development of a new tool to digitize scanned analog seismograms (WP2, WP3). this advanced tool was developed to digitize analog records using an algorithm based on the principles of image processing, data/noise discovery, and time-keeping while minimizing human interactions to facilitate the processing of large collections. The main challenge behind these principles is to extract a maximum of continuous seismic data from scanned analog seismograms while avoiding the unintentional rejection of potentially useful data. This motivates the combination of traditional computer vision used for the vectorization of the scans with machine learning techniques. This is meant to help maximize the digitization of challenging waveforms, such as earthquakes with large amplitudes that might cause lines to cross each other on the paper.

The vectorized traces then can be validated by specifically comparing the microseism amplitude and frequency content to the corresponding microseism generation model (WP4) based on the WAVEWATCH III (WWIII). The comparison first requires transforming the modeled wave climate into microseismic amplitudes at the station location on land (e.g. Stutzmann et al., 2012). Transforming the wave spectra into theoretical microseisms provides a first-order comparison with the digitized waveform to validate the resulting time series.

Once digitized and validated, the data will be made openly available to the community through a web service (WP5, WP6). Seismology has always been at the forefront of open and free data and metadata distribution and it is critical that this extends to legacy seismic data for their preservation and valorization. This requires making old seismograms compliant with modern standards prior to their dissemination.

SEISMOSTORM

Potential impact of the research on Science, Economy, Society, Culture, Environment and/or quality of life, Public policy or services, and/or management, conservation and valorisation of Federal Scientific Institution's collections.

Archives of legacy seismic data are the results of a century of scientific and financial investments that are not used in their paper form and are at risk of being lost. Digitizing all this seismic data will allow it to be preserved and analyzed using modern techniques. This does not only valorize the information they contain but also the initial investment that made them possible in the first place and the preservation effort that kept them usable to this day, all while giving them a new scientific potential.
The tools developed within this project are specifically thought out with reusability and open science in mind. This implies a broader impact on the valorization of legacy seismograms at observatories worldwide which will be facilitated by the connection of the project to the ESC working group 02-12 “Preservation, valorisation and analysis of seismological legacy data” (https://www.legacy-seismograms.eu/). This will be critical for projects that require global contributions such as extracting only quantitative observations of the oceanic climate at the global scale that is retained in legacy seismograms covering the XXth century.

Description of the expected final research results (analysis tools, studies, recommendations, conferences, models, scenarios, reports, publications, etc…) and valorisation perspectives at short and medium terms

The project will produce a few key tools with a global interest:
- The waveform digitization algorithm
- The tool transforming the wave spectra of global ocean wave models into theoretical microseisms at any stations on land

Both tools will be the subject of technical publications to describe them followed by scientific publication on the outcome of the SEISMOSTORM project.
The ongoing progress of the project have already been presented in an international conference (3ECEES, September 2022) and should be presented at more upon completing more work packages.
In the short term, the digitized legacy data from the ROB will be made available to the community through a web server and following modern standards. That will allow modern seismology technique to be implemented of on these old data.
The global ocean wave model generated from 1900 will also be made available for other observatory needing to validate their own digitized legacy data. These models are processed at the global scale and therefore good to be used for any seismic station around the world. In the medium and long term this will help reconstructing observations of the oceanic climate for the entire XXth century.

CONTACT INFORMATION

Coordinator

Dr. Thomas Lecocq
Royal Observatory of Belgium - Seismology & Gravimetry
thomas.lecocq@seismology.be
https://www.geophysique.be/thomas-lecocq/

Partners

Dr. Olivier Debeir
Université Libre de Bruxelles (ULB) – Laboratory of Image Synthesis and Analysis (LISA)
Olivier.Debeir@ulb.be
https://lisa.polytech.ulb.be/en/team/academics/pr-olivier-debeir

Dr Céline Hadziioannou
University of Hamburg, Germany
celine.hadziioannou@uni-hamburg.de

Fabrice Ardhuin
Université de Brest – Institut Français de Recherche pour l'Exploitation de la Mer (IFREMER)
Fabrice.Ardhuin@ifremer.fr

LINKS

https://seismologie.be/seismostorm/index.html