Seismology Toward Research Innovation with
Data of Earthquake (STAR-E Project) promoted by MEXT

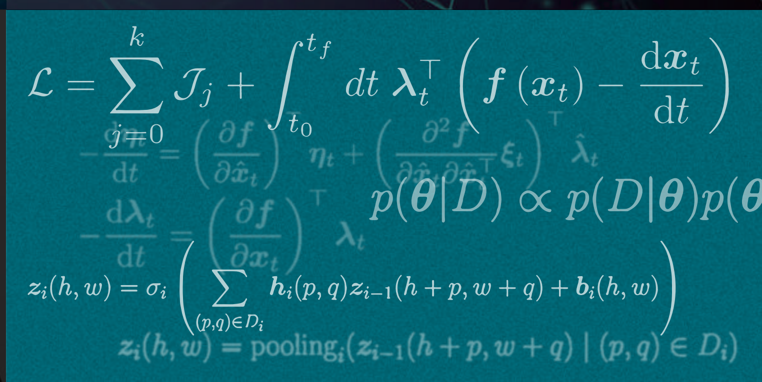
SYNTHA-Seis

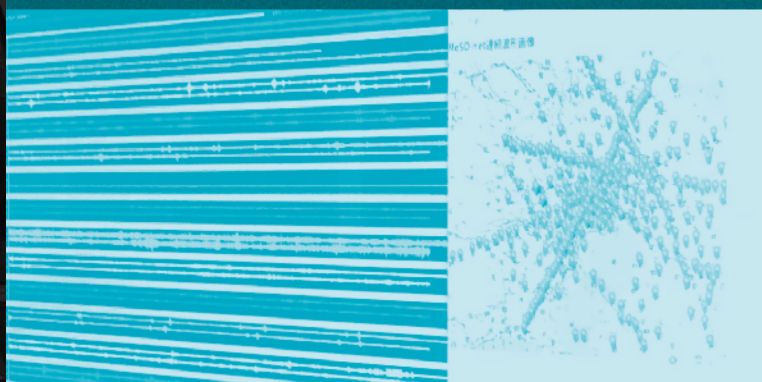
NEWSLETTER

Vol. 02
August 2024

Synergy effect Through Human and Artificial Intelligence Towards New Era in Seismology

A New Wind in Interdisciplinary Research through International Exchange





We're just entering an era of data science. It's time to fuse AI and seismology.

Currently, thousands of seismographs are deployed throughout Japan to monitor seismic activity.
Earthquake research begins with the collection of past records.
Research into the phenomena and tendency of seismic activity is progressing with "Earthquake catalogs", which collect information on detected earthquakes, such as arrival time, epicenter, and so on. However, These catalogs do not include small earthquakes that are not felt by the body, for example recently discovered slow earthquakes and low-frequency microtremors, despite their great importance for research.
AI is capable of comprehensively processing vast amounts of data and is by far the most cost-effective and fastest way to collect more data. Furthermore, machine learning, a type of AI, is also good at recognizing patterns in data, such as voice and image recognition, and is expected to have a high affinity with earthquake data recorded in waveforms. In the future, the relationship between earthquakes and microtremors may be newly elucidated through the re-detection of microtremors, and the introduction of AI is expected to generate more accurate earthquake prediction models.
Even with seemingly all-powerful AI, the ability of skilled professionals to instantly detect anomalies from a big-picture perspective, or natural intelligence, is still high. In addition, the deep learning process is a black box, and human judgment is always required as to whether the conclusions are correct and how the results should be handled. Recently, new methods have been developed to incorporate scientific knowledge and physical laws based on human experience and knowledge into the machine learning process in order to draw more useful conclusions. Thus, the key is to carefully consider and examine the respective strengths of artificial and natural intelligence, utilize them, provide appropriate feedback, and mutually develop them.
We will accelerate the practical application of "AI x Seismology". This new initiative will make earthquake research more fruitful and contribute to the prevention of excessive damage from natural disaster.

OPEN SCIENCE

Comprehensive Innovation
Open Access
Society 5.0
Open Data & Code

Empirical science → Theoretical science → Computational science → Data science

GREETINGS FROM THE PRINCIPAL INVESTIGATOR

SYNTHA-Seis Deepens "Information Science x Seismology"

SYNTHA-Seis, which started in July 2021, is already in its fourth year. Toward the realization of the research project's goal of "dialogue and collaboration between artificial intelligence and natural intelligence," SYNTHA-Seis is making remarkable progress with three project researchers and individual advice from professors of information science. As a new development in FY2023, five members, including myself, visited the Seismology Laboratory of the California Institute of Technology (Caltech), a pioneer in the development of deep learning models in seismology. We were greatly inspired by the excellent research environment and free atmosphere of the institute. We, the two organizations, have concluded an international exchange agreement to begin regular personnel exchanges. In April, 2024, Dr. Zachary E. Ross and two of his students from Caltech visited our institute and discussed future collaborations in the interdisciplinary field between information science and seismology.
We will continue to develop this field with such international and domestic collaborations. Please stay tuned for our activities in SYNTHA-Seis.

Principal Investigator: Hiromichi Nagao
Associate Professor, Earthquake Research Institute, The University of Tokyo

Earned Doctor of Science at Graduate School of Science, Kyoto University in 2002. After studying at the Japan Atomic Energy Agency, the Japan Agency for Marine-Earth Science and Technology, and the Institute of Statistical Mathematics, he has been working at the current position since 2013. He specializes in the interdisciplinary research between applied mathematics including statistics and solid earth science.

1960 | 1980 | 2000 | 2020

Fullness of libraries & Popularization of GPU

1st AI Boom → Solve a puzzle → 2nd AI Boom → Expert System → 3rd AI Boom → Deep Learning → Generative AI

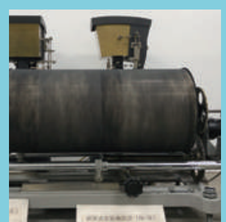
Continuous observation

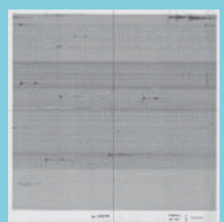
1995 Southern Hyogo Prefecture Earthquake
2011 the Pacific coast of Tohoku Earthquake
2017 AI+Seismology

Improved Seismographs | Networking | Discovery of Slow Earthquakes | Development of Earthquake Alerts | Introducing Deep Learning

Analog data on paper

Independent government/university/institute observatories

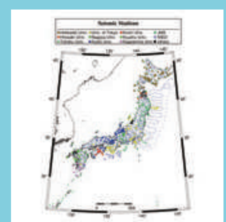


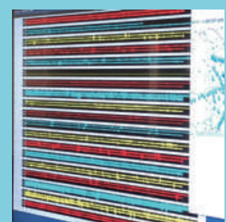


Buried Microtremor Detection

Digital data on the network

Nation-wide unification





High-precision Re-detection

Big data from diverse observatories

SYNTHA-Seis

AI Performance Evaluation
AI Big Data Analysis
AI Auto Detection

The penciled record of Wakayama Observatory is provided by ERI paleo earthquake and tsunami record committee.

Earthquake Detection Based on Deep Learning Method Inspired by Observation of Seismic Waveforms

In recent years, AI technology has come to be used for earthquake detection. By using a large amount of waveform data and having AI learn the characteristics of seismic waves, it is now possible to detect earthquakes more accurately than ever before. In this study, we developed a new AI method based on the idea from observing seismic waves with the "human eye", improving an existing AI method. The key idea was to incorporate not only the entire seismic wave but also local waveform patterns of seismic waves into AI learning, so that earthquakes can be detected based on both local and global characteristics. As a result, the weakness of the existing method, which tends to misclassify earthquakes, has been overcome and earthquakes can be detected more accurately.

Tomoki Tokuda
Project Assistant Professor, ERI, UTokyo

His expertise is in the development and application of machine learning methods, especially clustering methods. In this project, he is developing earthquake detection methods based on deep learning, while utilizing his expertise.

Today, earthquake detection is based on seismometers that measure ground motion. Since the ground surface is constantly shaking due to various factors including human activities, it is necessary to distinguish seismic waves from the measured shaking in order to detect earthquakes. Conventional detection methods evaluate rapid temporal changes in shaking amplitude and determine seismic waves when they exceed a threshold value. In recent years, with the development of AI technology, earthquake detection AI based on deep learning models has been actively developed. By having AI learn from the vast amount of seismic wave data, it is now possible to detect earthquakes more accurately by capturing features of seismic waves that have been overlooked by conventional methods. The key to the success of these AI techniques is that they are not restricted to the characteristics of rapid temporal changes in amplitude, but rather they learn the unique characteristics of seismic waves. On the other hand, most of the AI techniques used to date have been developed and used in fields other than seismology, such as image recognition, and are not specialized for earthquake detection. There is a wide variety of deep learning models, and it is not easy to give a clear answer to the question of which model should be used. In an attempt to address these issues, this study developed a new AI technique based on the characteristics of seismic waves.

The inspiration for the AI technique devised in this study came from qualitative and quantitative observations of waveform data. Figure 1 shows the waveforms of three phases of an earthquake (P-wave, S-wave, and noise; several waveforms are overlapped), with the maximum amplitude set to 1. It can be seen from this figure that the waveform patterns of the three phases are different, and it is also important to note that the three phases have different characteristics not only in the entire waveform but also in the first half or the second half of the waveform. The basis of this idea is something we experience in our daily life. For example, a butterfly/moth or a plum tree/peach tree can be misidentified by looking at only the whole from a distance, but can be accurately identified by paying attention to the wings and flowers (local information). Similarly, if the differences among the three phases in the local area have *different information* from the overall waveform, we can expect to be able to distinguish the three phases more accurately by utilizing this information. To confirm this, a statistical method called multiple clustering was applied to the waveform data. While conventional clustering methods group objects based on all features, multiple clustering automatically selects multiple features and identifies multiple grouping patterns. In this study, the multiple clustering method was applied to waveform data including P-wave, S-wave, and noise, using the amplitude of the waveform at each time point as the feature value. The results show that the data points are divided into four regions (black, green, blue, and red), as indicated by the color coding of the waveforms in Figure 1, and that each region has a different pattern of waveform groups. This result suggests that the first half and the second half of the four regions have *different information* about the waveform classification of the three phases, which is consistent with the waveform observation results described above.

Inspiration from data

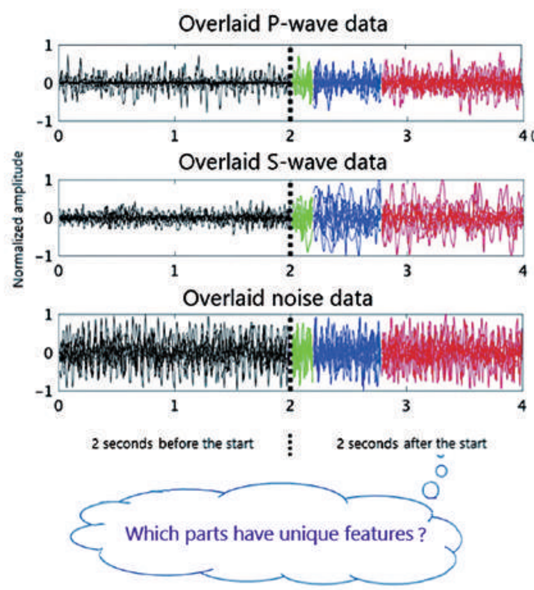


Figure1: The 4-second waveforms of P wave, S wave, and noise were overlaid with 10 instances. The vertical axis denotes normalized amplitude, whereas the horizontal axis elapsed time with initial motion aligned in the center. The four color-coded sections are likely to be distinctive.

Based on this idea, we developed a new deep learning model. Specifically, we improved an existing detection method called the GPD (Generalized Phase Detection) method, which is a typical deep learning model consisting of convolutional neural network (CNN) layers and fully connected layers. The CNN layer automatically extracts waveform features by multiplying a large number of weight matrices to the data, whereas the fully connected layer integrates these features to distinguish the three phases by calculating the detection probability of P-wave, S-wave, and noise. For continuous waveform data, P-wave, S-wave, and noise are distinguished sequentially by shifting the 4-second time window little by little (e.g., in 0.1-second increments). In this study, the GPD method was improved to incorporate not only the entire seismic waveform but also local information based on the aforementioned idea (Figure 2). In the improved model, seismic detection models were constructed for both the whole waveform and the local waveform, and the final results were obtained by integrating the results from each model. By incorporating local waveform information into the model in an explicit form, the model can more accurately distinguish waveforms that tend to be misclassified. When applied to continuous waveform data, it was confirmed that false positives were reduced (Figure 3).

In this study, we developed a new AI technology inspired by what we observed with "human eyes". One of the drawbacks of AI technology is that it is black-boxed, i.e., the process of reaching a conclusion is not clear. In practical use, there may be no particular problem if the correct conclusion is obtained using AI, but it is also true that the black-boxing makes it difficult to see the direction of AI technology development. Our approach in this research provides one solution to such a problem, which we find very useful for the development of AI specialized for the field of seismology. We would like to continue our efforts to further develop AI technology needed in the field of seismology by making the best use of human "natural intelligence".

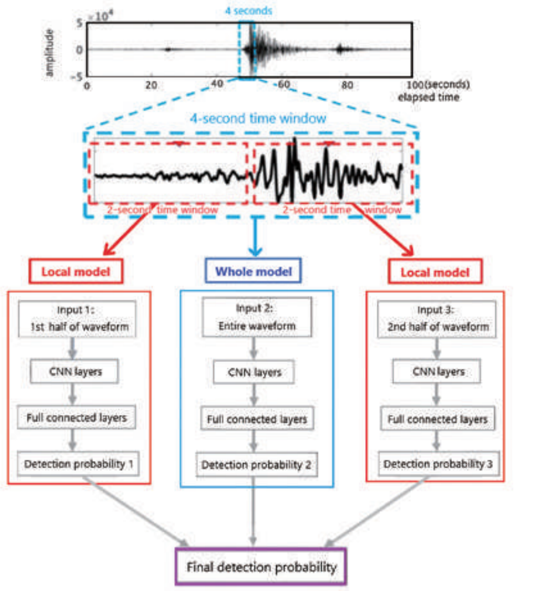


Figure2: Schematic diagram of the proposed method. The 4-second waveform is divided into first half, and second half. On top of the whole waveform, these local waveforms are also subjected to a deep learning model of P-wave, S-wave, and noise. The resultant detection probabilities are integrated to obtain the final probability.

Seismic detection results before and after the 2016 Bombay Beach earthquake swarm

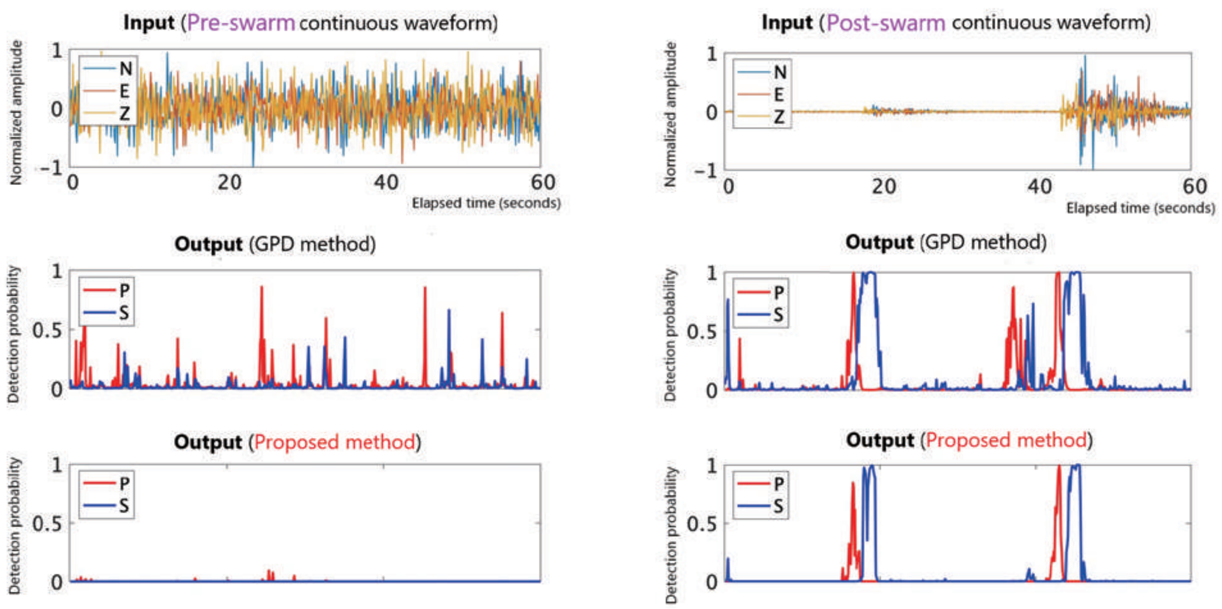


Figure3: The panels in the first row show waveforms of the north-south (N), east-west (E), and up-down (Z) components. The panels in the second and third rows show the detection probability of earthquakes (P-wave in red; S-wave in blue) evaluated every 0.1 second by the GPD method and the proposed method.

Reference: Tomoki Tokuda, Hiromichi Nagao (2023) Geophysical Journal International, doi:10.1093/gji/ggad270

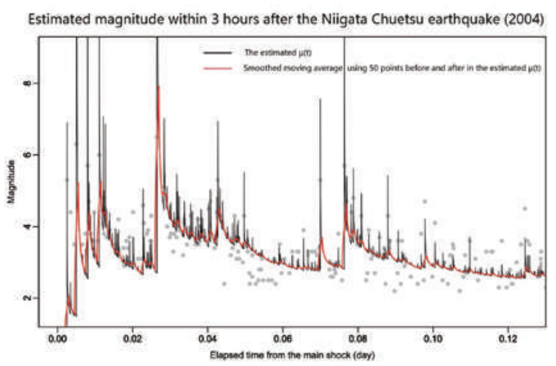
Kosuke Morikawa
Associate Professor, Graduate School of Engineering Science, Osaka University

Specialty: Mathematical statistics, missing-value data analysis, and semiparametric inference. His recent research focuses mainly on point process data analysis, including statistical seismology and survival time analysis.

Estimation for aftershocks distribution immediately after a large main shock using the detection probability of self-exciting aftershocks

Since seismic activity becomes active immediately after the main shock, numerous aftershocks occur and some large aftershocks of the relatively same magnitude as the main shock are observed. For disaster damage prevention, it is necessary to develop a method to estimate the "characteristics of aftershock activity" quickly from the data. Previous studies have been able to estimate the probability distribution of aftershocks based on the occurrence frequency and size of aftershocks, but the results are biased because the number of aftershocks is underestimated due to active aftershock activity. Therefore, it becomes possible to construct an estimation method without the bias caused by underestimation of the number of aftershocks by introducing aftershock detection probabilities that depend on the arrival time of aftershocks and their magnitudes and incorporating them into statistical models.

Meanwhile, in seismology, the Omori-Utsu law is known as a probability law for aftershock frequency distribution. This probability law is valid globally, but a parametric model that is not suitable for expressing detailed local probability laws. Then, the Epidemic Type Aftershock Sequence (ETAS) model, which uses a self-exciting model of historical event information, can represent the probability law of the frequency distribution well, even locally. This study proposes a detection probability of aftershocks that is expressive enough even for parametric model, by using a self-exciting model that utilizes historical aftershock data as well as the ETAS model (see the figure below).



The result of estimating the magnitude μ(t) for detecting an aftershock with a 50% probability at time t (day)
(Data of the 2004 Niigata Chuetsu earthquake in the JMA catalog)

Yoshikazu Terada
Associate Professor, Graduate School of Engineering Science, Osaka University

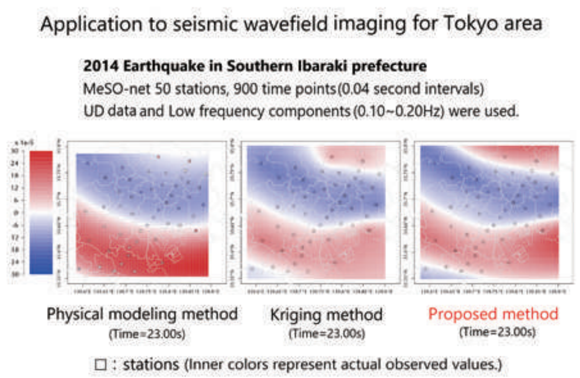
Specialty: Statistical science and machine learning with a focus on unsupervised learning. His recent research focuses mainly on the theory and applications of clustering methods and functional data analysis for large-scale data.

A fast smoothing method for spatio-temporal data and application to seismic wavefields imaging

When an earthquake occurs, early assessment of shaking in urban areas is useful for estimating the damage and for prompt recovery efforts. However, it is not possible to observe seismic motions at all locations. This study deals with a smoothing problem for spatio-temporal data, which aims to predict and complement long-period seismic motion at unobserved times and locations using time-series data of limited stations.

In general, smoothing methods for spatio-temporal data are computationally expensive. For example, the kriging method requires parameter estimation for each unobserved location, which is computationally expensive for imaging applications and also has the problem of poor accuracy when the stationarity assumption is not appropriate. Therefore, we focused on smoothing methods based on basis function expansions, which do not require parameter estimation for each unobserved location. On the other hand, existing methods based on basis function expansions are computationally expensive when the number of bases increases, making it difficult to use them for immediate evaluation. Therefore, we proposed a fast smoothing method using orthogonalized basis functions and a computationally efficient regularization method. Through numerical experiments, we confirmed that the proposed method can significantly speed up the estimation process without sacrificing the accuracy.

We also applied the proposed method to the imaging the long-period seismic wavefield of the earthquake that occurred in the southern part of Ibaraki Prefecture in 2014, and the proposed method can reconstruct it in about 3 seconds on a general laptop computer, including the selection of tuning parameters (selected from 1000 candidates) and complementary calculations using the cross-validation method.



Results of applying the proposed method to seismic wavefields imaging

"Caltech Seismo Lab" Visit Report

DATE 2023.12.4-8

Reported by Jeffrey Church
Doctoral student 2nd year, UTokyo

People around the world have various experiences and points of view, and scientists are no exception. Naturally, scientific progress benefits greatly from incorporating such diversity. This belief was one of the driving factors of my decision to pursue a doctoral degree at the University of Tokyo. I wanted to learn from the diverse field of researchers at the University of Tokyo' s Earthquake Research Institute (ERI) and incorporate their various perspectives into my own research. In December 2023, I had the opportunity to visit the Caltech Seismological Laboratory with my colleagues from the Nagao & Ito laboratory, ERI, with the goal of fostering international collaboration on AI seismology research. This visit to Caltech was an invaluable opportunity to learn from brilliant researchers on the other side of the world and expand my desire for collaboration to a global scale.

During our visit we had the pleasure of interacting with Dr. Zachary Ross and his group. Dr. Ross was the first scientist to apply deep learning to the problem of seismic phase picking with his Generalized Phase Detection (GPD) architecture. Dr. Ross' AI phase picking technique, along with others developed in recent years, have greatly increased the size of earthquake catalogs. As my research interests at the University of Tokyo are also in the area of deep learning phase picking architectures, I was excited for the opportunity to converse with Dr. Ross and to learn about his group's current research. I hoped that what I learned might inspire the future direction of my research at the University of Tokyo.

After a tour of the South Mudd building, home to the Caltech Seismological Laboratory, the visit kicked off with a pair of seminars with the Ross Laboratory, in which members of both groups introduced their research topics. Dr. Ross' group presented the innovative PhaseNO phase picking architecture, which produces cutting edge phase picking results by simultaneously analyzing data recorded at multiple seismic stations, an approach which better emulates the manual process by which human experts perform the task of phase picking. I found the seminars to be very inspiring, not only because of the impressive results of the Ross group's current research, but also because of the history. Dr. Ross began the pursuit of AI phase picking with his 2018 GPD paper, and his group continues to push the limits of what is possible today. For our goals, I believe there may be no better group in the world to collaborate with, and I look forward to a future exchange of ideas.

The visit concluded with a presentation by Dr. Nagao at the Caltech Seismological Laboratory Seminar. In his presentation, Dr. Nagao introduced current approaches to AI seismology in Japan. Highlighted was the increased volume of seismic data available to researchers in Japan, and the corresponding need for improved AI tools to process this glut of information. In pursuit of such AI advancements, Dr. Nagao introduced a new five-year collaboration between the ERI and the Caltech Seismological Laboratory, including the exchange of visiting researchers. During the presentation, I felt that these two groups were a perfect match for such a collaborative effort, as both groups are concerned with generating and processing large quantities of seismic data. I am very optimistic that our collaboration will be productive, and I am excited to make my own contributions to the project. During our visit to Caltech, the Ross group demonstrated what is possible to achieve in the field of AI seismology, and I am filled with renewed determination to become a researcher of the same caliber. After our visit to Caltech, my commitment is even stronger to producing useful research in the future that will make the world a safer place from seismic hazards.

NEW RESEARCHERS

Gerardo Manuel Mendo Pérez

Project Researcher ERI, UTokyo

Development of automatic seismic wave processing package

My motivation to research is to understand the "hidden" processes behind generating earthquakes and volcanoes. My research in this project aims to develop new methods to identify and process seismic slow earthquakes, LFE and tectonic tremor, and retrieve critical parameters that contribute to the understanding of this phenomenon. Concretely, I am developing an alternative method that combines the calculation of Root Mean Square (RMS) amplitudes to identify events with the advantage that the duration of these events depends on the RMS amplitude decay in time, retrieving events with variable durations. Afterwards, these events are classified using clustering algorithms. I hope to contribute to creating knowledge that helps understand complex earthquake-related phenomena.

Katoh Shinya

Project Researcher ERI, UTokyo

Development of an Automatic Processing Package for Seismic Waves

Seismic cataloging is essential for seismological analysis, but it requires human manual labor. This task is very hard and is not suitable for human processing of huge amounts of data. Therefore, I am developing an automatic seismic wave processing package using deep learning. This package will automatically perform travel-time determination, earthquake detection, P-wave initial motion polarity determination, and epicenter determination from observed continuous seismic waveform data, and create an earthquake catalog. The package under development has improved the ability to detect small earthquakes and significantly exceeds the processing speed of human manual work. We expect that this package will not only reduce the human workload but also contribute to a detailed understanding of seismic activity and crustal structure by being able to detect a large number of earthquakes.

AWARDS

Prof. Masaaki Imaizumi received the Young Scientist Award of Commendation for Science and Technology 2023 [2023.4.7]

He was awarded the Young Scientist Award of Commendation for Science and Technology 2023 from the Minister of Education, Culture, Sports, Science and Technology. This is for his work in "Research on Statistical Theory Describing Data Analysis Principles using Deep Learning".

Prof. Kazushige Obara received the Imperial Prize and Japan Academy Prize [2024.3.12]

He was awarded the Imperial Prize and Japan Academy Prize from the Japan Academy. This is for his work in "Discovery of Deep Low-Frequency Tremor at Subducting Plate Interface and Development of Slow-Earthquake Seismology".

UTokyo 大阪大学 OSAKA UNIVERSITY MEXT Earthquake Research Institute, The University of Tokyo 基礎工学部 大学院基礎工学研究科 STAR-E

SYNTHA-Seis

SYNTHA-Seis Secretariat Earthquake Research Institute, The University of Tokyo 1-1-1 Yayoi, Bunkyo-ku, Tokyo 113-0032, Japan
Email: syntha-seis-secretariat-group@g.ecc.u-tokyo.ac.jp