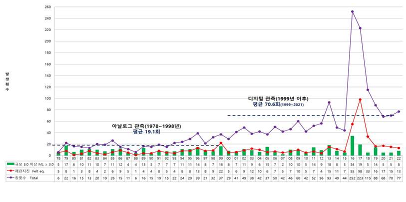
Preliminary study of estimation of earthquake characteristics in unmeasured site through machine learning of earthquake observation records

Yongmoon Hwang[a], Junghoon Lee[b*]
[a]Structural and Seismic Safety Research Division, Korea Atomic Energy Research Institute, Daedeok-Daero 989-111, Yuseong-Gu, Daejeon
[b*]Applied Science Research Institute, Korea Advanced Institute of Science and Technology, Daehak-Ro 291, Yuseong-Gu, Daejeon

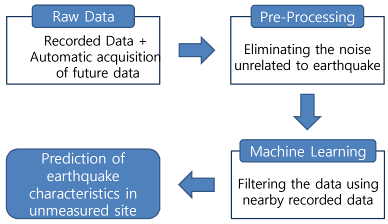
Korean Nuclear Society Spring Meeting, Korea, May 17-19, 2023

Introduction

- Frequency of Earthquake in Korea is gradually increased recently.
  - Gyeongju earthquake: magnitude 5.8
  - Pohang earthquake: magnitude 5.4
- Recently, studies using machine learning to predict the responses of unmeasured sites using monitoring system are being conducted.
- The objective of this study is to develop a technology that can estimate the earthquake characteristics of a unmeasured site through the surrounding seismic measurement record data.



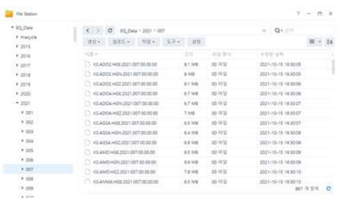
▲ Current status of Earthquakes in Korea (1978~2022)



▲ Flowchart of prediction of earthquake characteristics in unmeasured site

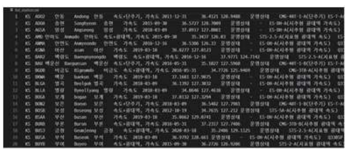
Acquisition of recorded data

- Acquisition and construction of seismological observatory continuous waveform data
  - Recorded seismic data are collected from seismological observatory in NECIS operated by the Korea Meteorological Administration
  - Earthquake data were obtained from January 1, 2015 to December 14, 2021 based on UTC.
  - Measurement data was sampled at 100 Hz with 3-axis accelerometer, with directions of X(EW), Y(NS, and Z(Vertical) axes.
- Classification of acquired earthquake data list
  - The acquired earthquake data was organized into a list of magnitude 3.0 or higher.
  - The data was organized by time of occurrence (UTC), magnitude, latitude, and longitude.
  - Data file name: KS.{point code}.HG(direction}.{year}.{No.}.00.00.00
- Arrangement of seismological observatory in Korea
  - As of January 1, 2022, the operating seismological observatory has been compiled.
  - The seismological observatory is organized with point code, latitude, longitude, X-, Y-, Z-axis, and gain recorded data.



▲ List of classified earthquake list

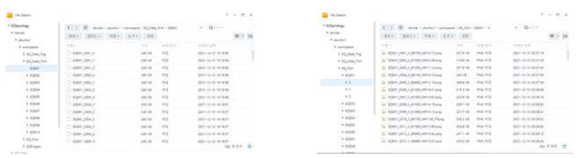
▲ Arrangement of earthquake list



▲ Arrangement of observation site list

Data Processing For Learning

- Data Processing for each seismological observatory
  - Data load of earthquake and seismological observatory → filtering by classifying earthquake occurrence time and observatory
  - Calculate the distance from the epicenter of the earthquake to the seismological observatory → sort it by closest to the epicenter of earthquake
  - Load the data of the filtered seismological observatory → cut the data from the time of the occurrence of the earthquake to 5 min. → simply shift the average value to 0 → resave the data as a new file
  - File name: EQ{Earthquake No.}_S(point code}_{direction}

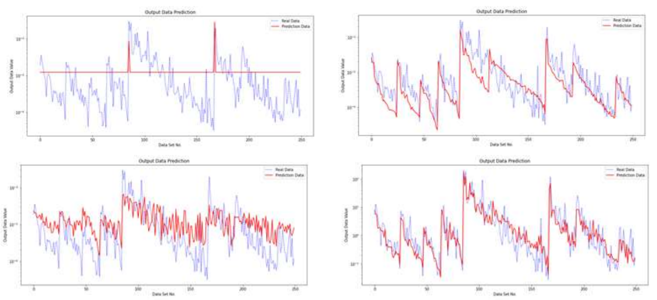


▲ Observation data Processing

Learning cases & Results

- Develop a machine learning model to predict the PGA in unmeasured site
  - Classify into input and output data to perform machine learning on measurement data
    - Input variables – 1) latitude and longitude (deg.) of the occurrence of the earthquake, 2) earthquake magnitude, 3) latitude and longitude (deg.) of the seismological observatory, 4) distance from earthquake to seismological observatory (km), and 5) orientation of the observatory from the earthquake (west direction (-), east direction (+) relative to the north)
    - Output variable – Peak Ground Acceleration
  - Utilize Multi-Layer Perceptron-based Deep Neural Network Model[1]
    - MLP model structure: one input layer, two hidden layers, and one output layer
    - Input layer – 7 nodes, each hidden layer – 100 nodes, output layer – 1 node
    - Node activation function: Rectified Linear Unit (ReLU)[2]
    - 80% of training and 20% validation set randomly for each epoch
    - Evaluation index – Mean Squared Error (MSE) & Mean Absolute Error (MAE)
- Learning Case for Predicting unmeasured site

Case	Features	Limitations
1	Using recorded data without Preprocessing	Inaccurate
2	Training data with log-scale	Reversed tendencies
3	Relative distance from epicenter to observatory ($1/r^2$)	Inaccurate
4	Normalization of recorded data	Better



▲ Comparison of results according to the learning cases

Summary

- Acquisition and classification of the recorded earthquake data through NECIS
- Consideration of the input variables for machine learning
- Comparison of results according to the learning cases (Case 1 to 4)
- Furth study: Prediction of the response spectrum of the unmeasured site and application to NPP

Acknowledgement

- This work was supported by the National Research Foundation of Korea (NRF) grand founded by the Korea government (MSIT) (RS-2022-00144425 and 2021R1I1A1A01050935).

References

[1] Shiblee, Md, Perm Kumar Kalra, and B. Chander, "Time series prediction with multiplayer perceptron (MLP): a new generalized error based approach," Advances in Neuro-Information Processing: 15th International Conference, ICONIP 2008.
[2] Brownlee J., "A gentle introduction to the rectified linear unit (ReLU)," Machine Learning Mastery, 6, 2019.