Benchmark on External Events Hazard Frequency and Magnitude Statistical Modelling in KAERI

Beom-Jin Kim[a], Minkyu Kim[b], Daegi Hahm[c]

[a] Korea Atomic Energy Research Institute, Daejeon, Republic of Korea, beomjin88@kaeri.re.kr
[b] Korea Atomic Energy Research Institute, Daejeon, Republic of Korea, minkyu@kaeri.re.kr
[c] Korea Atomic Energy Research Institute, Daejeon, Republic of Korea, dhahm@kaeri.re.kr

Abstract: This benchmark study aims to apply statistical modeling for frequency and magnitude estimation based on external event hazard assessment data. Based on the results of this study, it is believed that an approach to the quantification of external event initiating events (IEs) can be formulated and evaluated through the application of an effective statistical model. This study's analysis was based on two cases that considered benchmarks provided by the Nuclear Energy Agency of Organization for Economic Co-operation and Development (OECD/NEA). Each case was given a magnitude according to the return period. An appropriate statistical model was applied through regression analysis for each case based on this data. Based on the results, the magnitudes of 500, 5,000, 50,000, and 500,000 years were predicted and presented.

The result of this study, statistical analysis was applied to the estimation of two cases presented by the OECD/NEA. In any statistical analysis, it is important to understand the characteristics of the data set. For the given problems here, the range of the return period was 10–10,000 years, while that of the magnitude range was 0.4–5.0 meters. Therefore, the coefficient of the synthetic model had a significant influence on the analysis results. This study demonstrates that employing the full extent of the significant figures is important to handle the different ranges of data values. In the future, it is expected that data-based statistical values can be better estimated through various verified statistical models.

1. INTRODUCTION

In recent years, the intensity of external hazard and the frequency of typhoons have increased due to an abnormal climate [1]. As an input into risk analysis modelling and simulation, a hazard initiating event (IE) is typically considered as the starting point for risk models. For example, IE is an important factor when using event tree or fault tree for probabilistic safety assessment (tsunami, earthquake, flood and other natural hazard). Since the IE both contributes to the risk quantification results and provides the boundary conditions for the rest of the hazard scenarios, effective modelling of the frequency and magnitude for different external events using data-driven methods can be applied. However, current practice has shown a variety of technical approaches, models, and limitations in validation of these approaches. Consequently, this benchmark study is intended to demonstrate and capture commendable practices in formulating and assessing the quantification of external event IEs when using statistical models [2].

2. SYNTHETIC DATA ANALYSIS

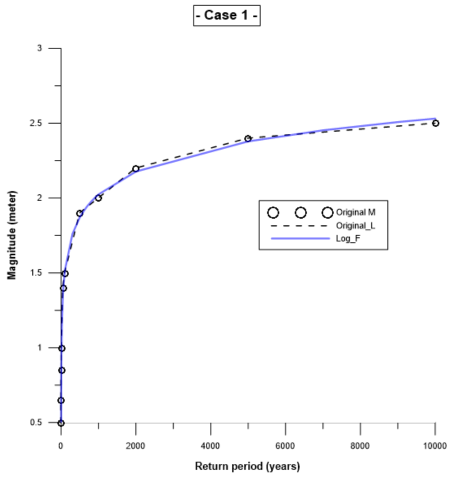
2.1. Study of case 1

Table 1 shows the synthetic data for Case 1 as provided by the OECD. Using regression analysis, log regression showed appropriate fitting results for the relationship between magnitude and return period, as shown in Figure 1.

Table 1. Synthetic data for Case 1

Return period (years)	**1**	**2**	**5**	**10**	**50**	**100**	**500**	**1,000**	**2,000**	**10,000**
Original M (meters)	0.50	0.65	0.85	1.00	1.40	1.50	1.90	2.00	2.20	2.50

Figure 1. Log regression fitting



As a result of the regression analysis on the magnitude of Case 1, the log regression equation including variables A and B is shown in Eq. (1). The magnitude of the return period from 1 to 10,000 years was estimated by Eq. (1); Table 2 compares the values to the original magnitude values proposed by the OECD.

- Case 1: $M = 0.2199 * ln\ (x) + 0.5034\ (parameters\ A\ and\ B)$ (1)

Table 2. Regression results for Case 1

Return period (years)	**1**	**2**	**5**	**10**	**50**	**100**	**500**	**1,000**	**2,000**	**10,000**
Original M (meters)	0.50	0.65	0.85	1.00	1.40	1.50	1.90	2.00	2.20	2.50
Log Magnitude (meters)	0.503	0.656	0.857	1.010	1.364	1.516	1.870	2.022	2.175	2.529

An error analysis was then performed using the SUMXMY2 function to verify the statistical justification of the estimated magnitudes. The SUMXMY2 function squares and sums the difference between two corresponding values, and therefore, the closer to 0, the smaller the error between the two variables, and the more statistically valid the estimation can be considered. The SUMXMY2 function is expressed as Eq. (2).

- $SUMXMY2 = \sum(x - y)^2$ (2)

Here, x is the value of the original magnitude, and y is the value of the estimated magnitude. As a result of error analysis using the SUMXMY2 function, the sum square error (SSE) value was calculated as 0.005, which is very close to zero. It was therefore judged that the estimated magnitude values were very similar to the original values and valid. However, to minimize SSE and more precisely estimate the magnitude values, a solver function was used. The target of the SSE value was set to 0, and an optimization analysis was performed on parameters A and B of Eq. (1). The results are shown in Table 3.

Table 3. Optimization for parameters

Parameter	Original	SSE_Solver
A	0.2199	0.219861834
B	0.5034	0.503363549

These optimized parameters are then used in the regression equation for Case 1, as shown in Eq. (3) below. Table 4 compares the magnitude values estimated by Eq. (3) with those estimated by Eq. (1) and the original magnitude values.

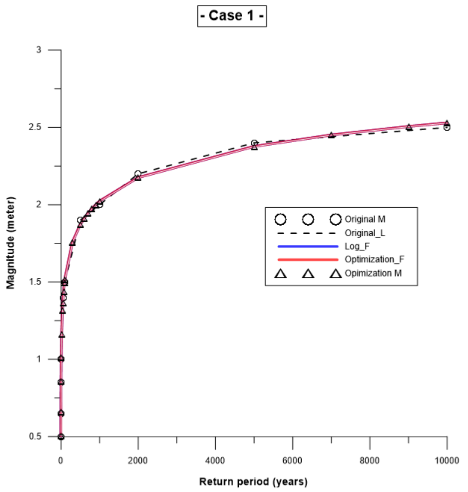
- Case 1: $M = 0.219861834 * ln\ (x) + 0.503363549\ (parameters\ A\ and\ B)$ (3)

Table 4. Comparison of magnitudes

Return period (years)	1	2	5	10	50	100	500	1,000	2,000	10,000
Original M (meters)	0.500	0.650	0.850	1.000	1.400	1.500	1.900	2.000	2.200	2.500
Log Magnitude (meters)	0.503	0.656	0.857	1.010	1.364	1.516	1.870	2.022	2.175	2.529
Optimized magnitude (meters)	0.503	0.656	0.857	1.010	1.363	1.516	1.870	2.022	2.175	2.528

For Case 1, the values of the optimized parameters A and B were similar to the existing values. Likewise, the SSE value of 0.0049 was also similar to the existing value of 0.005. Fitting was then performed based on the optimized magnitude values; results are shown in Figure 2.

Figure 2. Fitting result for Case 1



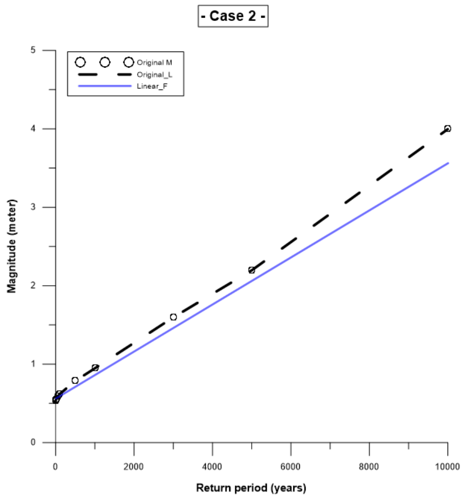
2.2. Study of case 2

Table 5 shows the synthetic data for Case 2 provided by the OECD. In this case, regression analysis found linear regression to give the best fit between magnitude and return period, as shown in Figure 3.

Table 5. Synthetic data for Case 2

Return Period (years)	1	2	5	10	50	100	500	1,000	3,000	10,000
Original M (meters)	0.53	0.53	0.54	0.55	0.59	0.62	0.79	0.95	1.60	4.00

Figure 3. Linear regression fitting



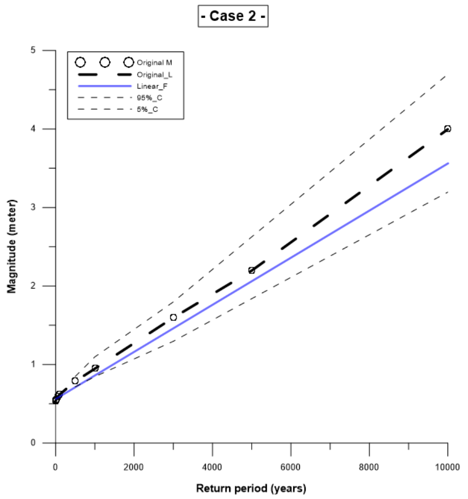
Equation (4) below is the linear regression equation including variables A and B as a result of the regression analysis on the magnitude of Case 2. Magnitudes of return periods from 1 to 10,000 years were estimated by Eq. (4) and compared to the values of the original magnitude proposed by the OECD (Table 7). In addition, 95% and 5% confidence interval values provided by the OECD are also given in Table 6, with fitting results plotted in Figure 4.

- Case 2: $M = 0.0003 * x + 0.5651\ (parameters\ A\ and\ B)$ (4)

Table 6. Regression results for Case 2

Return period (years)	1	2	5	10	50	100	500	1,000	3,000	10,000
Original M (meters)	0.53	0.53	0.54	0.55	0.59	0.62	0.79	0.95	1.60	4.00
Linear Magnitude (meters)	0.565	0.566	0.567	0.568	0.580	0.595	0.715	0.865	1.465	3.565
Original M 95% (meters)	—	—	—	—	—	—	0.85	1.1	1.8	4.7
Original M 5% (meters)	—	—	—	—	—	—	0.72	0.85	1.3	3.2

Figure 4. Linear regression fitting according to confidence interval



As a result of the regression analysis, the estimated magnitude values fell within the confidence interval and cannot be judged as inappropriate. The estimated magnitude values for the initial return period were similar to those proposed by the OECD. However, as the return period increases, errors in the magnitude values were found to occur. Error analysis in this case using the SUMXMY2 function gave an SSE value of 0.22443. In order to minimize this error in Case 2, the solver function was used, where again the target of the SSE value was set to 0 and an optimization analysis was performed on parameters A and B from Eq. (4). The results are shown in Table 7.

Table 7. Optimization for parameters

Parameter	Original	SSE_Solver
A	0.0003	0.000344252
B	0.5651	0.565051348

Based on the optimized parameters, linear regression analysis for Case 2 was re-estimated via Eq. (5). Table 8 compares the magnitude values from the three sources [original, Eq. (4), and Eq. (5)] along with the confidence intervals.

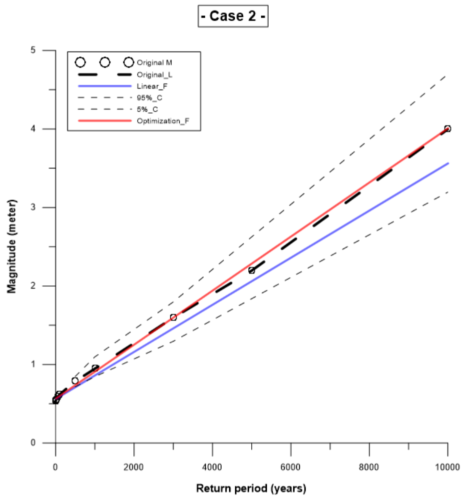
- Case 1: $M = 0.000344252 * x + 0.565051348\ (Parameters\ A\ and\ B)$ (5)

Table 8. Compare for magnitude

Return period (years)	1	2	5	10	50	100	500	1,000	2,000	10,000
Original M (meters)	**0.53**	**0.53**	**0.54**	**0.55**	**0.59**	**0.62**	**0.79**	**0.95**	**1.6**	**4**
Linear Magnitude (meters)	**0.565**	**0.566**	**0.567**	**0.568**	**0.58**	**0.595**	**0.715**	**0.865**	**1.465**	**3.565**
Optimized magnitude (meters)	**0.565**	**0.566**	**0.567**	**0.568**	**0.582**	**0.599**	**0.737**	**0.909**	**1.598**	**4.008**
Original M 95% (meters)	—	—	—	—	—	—	**0.85**	**1.1**	**1.8**	**4.7**
Original M 5% (meters)	—	—	—	—	—	—	**0.72**	**0.85**	**1.3**	**3.2**

In Case 2, the SSE of the optimized model was calculated to be 0.01. Accordingly, it was judged that the values of parameters A and B were significantly improved compared to the existing values. Fitting was then performed based on the optimized magnitude values. The results are shown in Figure 5.

Figure 5. Fitting result for Case 2



3. RESULT OF CASE STUDY

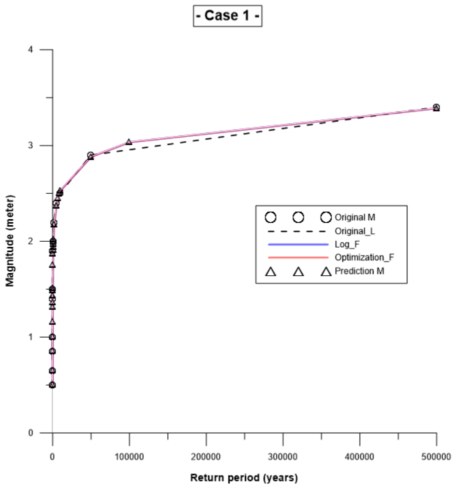
3.1. Case 1 result

The results of the optimized fit for Case 1 were estimated to be similar to the size values presented by the OECD. The trend lines calculated from the estimated magnitude values (Figures 1 and 2) were also similarly estimated. Therefore, it can be judged that the fitting result for Case 1 in this study is valid. Additionally, based on the regression analysis estimates in Case 1, magnitude values were predicted for 500, 5,000, 50,000, and 500,000 years return period. The results are presented in Table 9 and Figure 6.

Table 9. Magnitude prediction by return period (Case 1)

Return period (years)	500	5,000	50,000	500,000
Magnitude (meters) Case 1 Exact	1.9	2.4	2.9	3.4
Log KAERI mean (meters)	1.870	2.376	2.883	3.389
Optimized KAERI mean (meters)	1.870	2.376	2.882	3.388

Figure 6. Magnitude prediction fitting for Case 1



3.2. Case 2 result

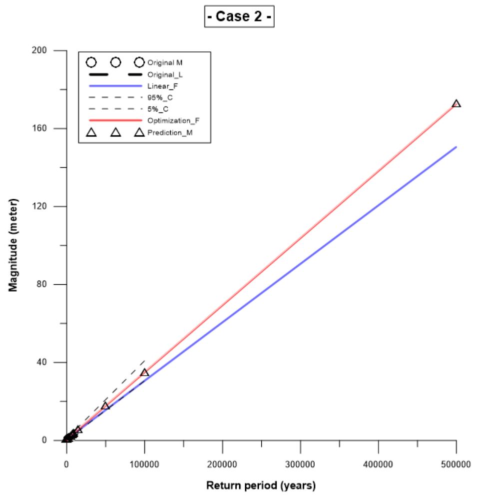
It was found that when the optimization technique was applied to Case 2, the model performance further improved, as seen in Figure 6. In other words, the optimized fitting was able to estimate values similar to the magnitude values proposed by the OECD. Comparing Figures 5 and 6, the trend lines calculated from the estimated magnitude values were also similarly estimated. Therefore, like Case 1, it can be judged that the fitting result for Case 2 in this study is valid.

Then, based on the regression analysis estimates in Case 2, magnitude values for return periods of 500, 5,000, 50,000, and 500,000 years were predicted. The results are shown in Table 10 and Figure 7.

Table 10. Magnitude prediction by return period (Case 2)

Return period (years)	500	5,000	50,000	500,000
Magnitude (meters) Case 2 Exact	0.78	2.2	28	2,000
Linear KAERI mean (meters)	0.715	2.065	15.565	150.565
Optimized KAERI mean (meters)	0.737	2.286	17.778	172.691

Figure 7. Magnitude prediction fitting for case 2



4. CONCLUSION

In this study, statistical analysis was applied to the estimation of two cases presented by the OECD/NEA. In any statistical analysis, it is important to understand the characteristics of the data set. For the given problems here, the range of the return period was 10–10,000 years, while that of the magnitude was only 0.4–5.0 meters. Therefore, the coefficient of the synthetic model had a great influence on the analysis results. This study demonstrates that employing the full extent of the significant figures is important to handle the different ranges of data values. In the future, it is expected that data-based statistical values can be better estimated through various verified statistical models.

Acknowledgements

This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIP) (RS-2022-00144493)

References

[1] National Institute of Meteorological Sciences. “Global Climate Change Forecast Report; National Institute of Meteorological Sciences: Jeju-do”, Korea, (2019).

[2] Nuclear Energy Agency, “*Benchmark on External Event Hazard Frequency and Magnitude Statistical Modelling*,” Organisation for Economic Co-operation and Development, 2020.