

# Seismic Damage Estimation of an Actual Reinforced Concrete Structure Using Deep Learning

Shigeru KUSHIYAMA

**Abstract**— When estimating low damage probability of a building to the  $10^{-6}$  level, even with subset MCMC, response analysis of over 150,000 times is necessary. It requires much CPU time and therefore unrealistic. In this paper, once after reducing the number of response analysis without using subset method, we create a n dimensional regression hypersurface (a deep learning regression model) that predicts low damage probability of a building. This enables quantitative estimation with dramatically less effort.

**Keywords**— seismic low damage probability, deep learning, importance analysis, gradient boosting decision tree, Keras, TensorFlow

## I. Introduction

In our past research [1], we estimated a low damage probability considering an earthquake risk problem by using subset MCMC (Markov Chain Monte Carlo) [2] instead of standard MCS (Monte Carlo Simulation). We have pointed out that the relative magnitude relation of the influence on the maximum ductility factor:  $\mu_{max}$  of the weakest story is the story yield shear strength:  $V_{jy}$  > the second shear stiffness:  $SK2$  > the first shear stiffness:  $SK1$ . However, the calculation required too much CPU time. For example, if one analysis took 1 (sec), it would require 41.7 hours for 150,000 analyses. Furthermore, since subset MCMC takes only a certain number of samples  $n_s$  as seeds when proceeding to the next chain level for generating samples in subset space, there is a limit to eliminating the periodicity of random variables in every  $n_s$  steps, and it is also difficult to judge the convergence of a simulation [3].

In order to solve this issue requiring too much CPU time, we first used GBDT (Gradient Boosting Decision Tree), and examined the influence on the fracture index ( $\mu_{max}$  in this example) of the input variable parameters (33 pieces:  $V_{jy}$ ,  $SK1$ ,  $SK2$  in all stories), and selected important parameters. Next, we reduced the number of response analysis by 1/100 of standard MCS (or 1/15 of subset MCMC) by introducing a new idea, followed by creating a n dimensional regression hypersurface (a regression learning model) that predicts  $\mu_{max}$ , finally estimating the low damage probability by comparing the  $\mu_{max}$  obtained from the test data set to the definition of limit state.

## II. Ground Motion, Analysis Model and Assumption

The input ground motion is JMA Kobe 1995 NS shown in Figure-1. Note this strong motion is a level 2 motion that standardized PGV (Peak Ground Velocity) to 50 kine. The analysis model is an actual RC building with story height and weight shown in Table-1. This is the same model as the analysis model described in Reference [1], and is replaced with a shear type frame with 2 spans in the X direction and 2 spans in the Y direction.

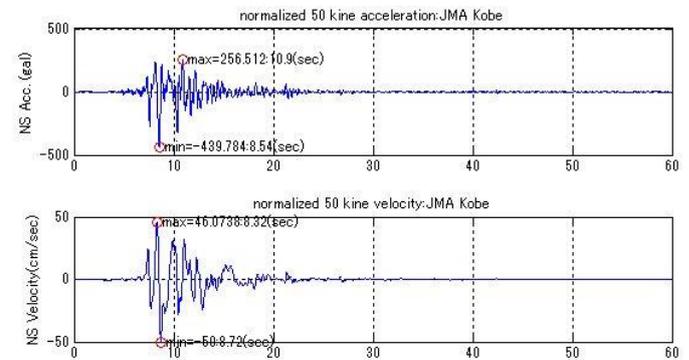


Figure-1 JMA Kobe 1995 NS normalized Earthquake

Table-1 Structural height and story weight of analysis model

floor	11F	10F	9F	8F	7F	6F	5F	4F	3F	2F	1F
structural height (cm)	295	295	295	295	295	295	295	295	295	295	490
story weight (kN)	4022	4176	4282	4282	4311	4330	4426	4446	4486	4550	4724

In nonlinear dynamic response analysis, it is assumed that a skeleton curve is tri-linear type, a hysteresis rule is Masing type, a viscous damping is Rayleigh damping that has 5% attenuation of the natural vibration mode up to the first 2<sup>nd</sup> order. For the numerical integration of the equation of motion, Wilson's  $\theta$  method ( $\theta = 1.42$ ) was used. Since this analysis model is a 11 story building, it is assumed that there is no large sway after the ground motion stopped, and we also adopted the following calculation cut-off time to minimize calculation time.

$$\text{cut-off time} = \text{maximum amplitude occurrence time} + 5 \text{ seconds}$$

Here, the maximum amplitude occurrence time means the time that occurs most late among PGA, PGV, and PGD generation times.

Random variables are  $SK1$ ,  $SK2$  and  $V_{jy}$  of each story, and each parameter is a set of i.i.d. (independent identically distributed) according to Gaussian normal distribution. Table-2 shows the expected values prescribed by replacing the curvilinear relationship between inter story drift and

story shear force, which are obtained from the pushover analysis, into a tri-linear relationship. In this table, the values of X direction frame which is slightly inferior in earthquake resistance is shown.

Table-2 Story yield strength and 1<sup>st</sup> and 2<sup>nd</sup> story shear stiffness (expected value)

floor	X direction frame		
	story yield strength (kN)	1st story stiffness (kN/cm)	2nd story stiffness (kN/cm)
11F	4089	7193	3205
10F	6237	9609	4119
9F	8296	10604	4368
8F	9891	11296	5010
7F	11267	12083	6010
6F	12862	13175	8755
5F	14122	14441	9629
4F	15025	15372	10634
3F	15824	16617	11393
2F	16389	18916	10287
1F	16869	27463	13560

### III. GBDT Importance Analysis of Input Parameters

$V_{ij}$ , SK1, and SK2 of each story are set up as input parameters (this is called 'feature' in machine learning field) and the significance of parameters was analyzed by GBDT (Gradient Boosting Decision Tree). GBDT is one of the methods of branching features in stages, drawing tree diagrams and classifying them, enabling easier interpretation of output results. Furthermore, preprocessing accompanying scaling such as standardization and normalization of training data is unnecessary, and continuous data and discrete data can be handled at the same time.

Figure-2 shows the importance analysis result in the case of COV (Coefficient of Variation) 5% and the response analysis  $nt=1,000$  times. The features are 33 pieces:  $V_{ij}$ , SK1 and SK2 for all stories shown earlier. On the other hand, the target value is class level 1 to 10, which are obtained by dividing the  $\mu_{max}$  range [minimum value, maximum value] for the weakest story obtained from each response analysis into 10 classes in the order of small value to large value. Here, the weakest story means, the floor having the largest frequency where  $\mu_{max}$  is the largest among all stories for each response analysis.

From Figure-2, the importance of features giving a large influence to  $\mu_{max}$  of the weakest story (3<sup>rd</sup> FL in this example) are as follows:

- The importance of  $V_{ij}$  in the weakest story is very large (23%).
- Next, the importance of  $V_{ij}$  of the lower and the upper story adjacent to the weakest story is large (5 to 6%).
- The importance of the other features are smaller than above values, and there is not much difference between SK1, SK2 and  $V_{ij}$ .

Considering the above, we selected 11 features (only  $V_{ij}$  of each floor) as important features. Figure-3 shows the importance analysis result on 11 features. In this example, the weakest floor is 1<sup>st</sup> FL. However, this result is induced from the result among only 1,835 times where the weakest story of each time coincides with the weakest story (1<sup>st</sup> FL) defined above, within the response analysis  $nt=5,000$  times

shown in Figure-4. In the figure, although the frequency of dividing  $\mu_{max}$  of the weakest story (1<sup>st</sup> FL) into 10 classes is biased, the validation accuracy of GBDT was 84.5%.

```
n_features=33
[0.0086233 0.01543476 0.02079628 0.01312414 0.00958999 0.02922866
0.01568326 0.02126588 0.018102 0.02282655 0.02738191 0.00912781
0.02547183 0.01817363 0.01847061 0.0218824 0.01994607 0.02037109
0.02578091 0.03612174 0.03092708 0.02284066 0.01956999 0.0523466
0.0110117 0.01450942 0.2329489 0.02020643 0.01815058 0.05878598
0.02539321 0.01808 0.04816797]
```

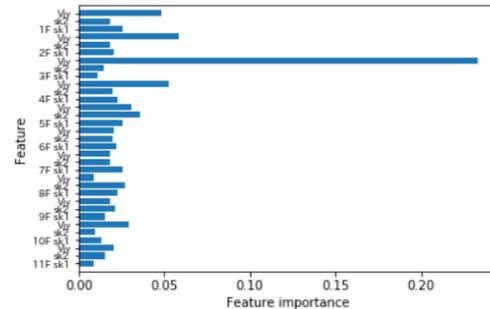


Figure-2 Importance analysis of features (input features = 33,  $nt=1000$ ,  $COV = 5\%$ )

```
n_features=11
[0.02937364 0.02775854 0.03869304 0.0273882 0.02820438 0.02914368
0.02623706 0.04557286 0.10735621 0.16703154 0.44004084]
```

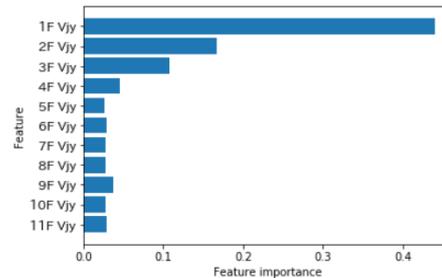


Figure-3 Importance analysis of features (features=11,  $nt=5000$ ,  $COV=2\%$ )

```
>> DeepLearning_Ap2_mdof_RC11
analysis case name=X-case08-d
MCS executing...
Earthquake name=JMA Kobe-NS
simulation no=1
new_nt=5000
the weakest story : weak_FL=1(F)
相対頻度で規定した最弱層と各回判定最弱層が一致した総数=1835
target-value: Distributed number after 10 class classification
elements in class(1)=11
elements in class(2)=121
elements in class(3)=535
elements in class(4)=1195
elements in class(5)=1463
elements in class(6)=1036
elements in class(7)=472
elements in class(8)=133
elements in class(9)=27
elements in class(10)=7
Number judged as the weakest story(RF→1F)
0 0 0 0 0 10 78 619 1516 942 1835
analysis_time=4078.85(sec)
>>
```

Figure-4 Frequency in class of  $\mu_{max}$  and the number judged as the weakest story

## IV. Derivation Policy of Damage Probability

### A. Setting Up of COV and Response Analysis Times for Deep Learning

We use Latin hypercube sampling to generate sample parameters. However, low damage probability is quite influenced by the COV value. Therefore, in Figure-5, we confirmed the following relationship between COV and the ratio of minimum probability variate to the mean value. The minimum probability variate means the probability variate corresponding to the equal division minimum point of the cumulative distribution function when samples are generated using the Latin hypercube sampling.

From the figure, the ratio of minimum probability variate to the mean value is over 75% for risk= $5 \times 10^{-5}$ % of COV = 5%, and is about 90% for COV = 2%. With a large COV, we can determine the low damage probability with relatively little response analysis times: nt. However, a large nt is required in the case of small COV, such as COV = 2% in estimation of probability exceeding life safe limit state. For example, standard MCS (Monte Carlo Simulation) would require one million analyses in such a case. Considering this point, we assume that nt are 1,000,000, 100,000, 10,000, 5,000, respectively for COV = 2, 3, 4, 5%. In these cases, the ratio of minimum probability variate to the mean value remains as high as approximately 82% for even the smallest case, and the values seem roughly acceptable.

However, if nt exceeds 10,000, it requires extended CPU time. Therefore, we considered one strategy to reduced nt. First, as shown in Figure-6, focusing on the normalized  $V_{jy}$  of the weakest story (1<sup>st</sup> FL in this example), we can see that samples with a large  $\mu_{max}$  (except blue points in the figure) correspond to a small value on the left half, and furthermore, the level of  $\mu_{max}$  becomes larger as it approaches the left end. It suggests that the smaller the value of  $V_{jy}$  of the weakest story, the larger the value of  $\mu_{max}$ . However, there is no such trend like this in the other floors. The trend was same even if the weakest story is different. Based on this finding and the fact that preliminary prediction of the weakest story is difficult, we assumed that  $V_{jy}$  takes a restricted value for all floors, not a whole fluctuation range but a lower fluctuation range 1/R in this paper. According to this, giving restriction to the fluctuation width of  $V_{jy}$  and the additional condition reducing nt to nt/100, the following equation holds.

$$(1/R)^{11}=1/100 \quad (1)$$

By taking the logarithm of both sides, we obtain a R value such as  $1.519911=1/0.657933$ . Therefore, in the case of nt = 1,000,000, that is, the interval [0, 1] of CDF shown in Figure-7 equally divided into 1,000,000 by the Latin hypercube sampling, the above expression is also satisfied when using generate probability variate of all floors from the smallest one until the 657,933th. When nt is equal to 100,000 and the right side of equation (1) is set to 1/10, reduced nt is set to 10,000, the generate probability variate is picked up until the 81,113th.

If we restrict the fluctuation range of  $V_{jy}$  of other floors in addition to the weakest story, and reduce the response

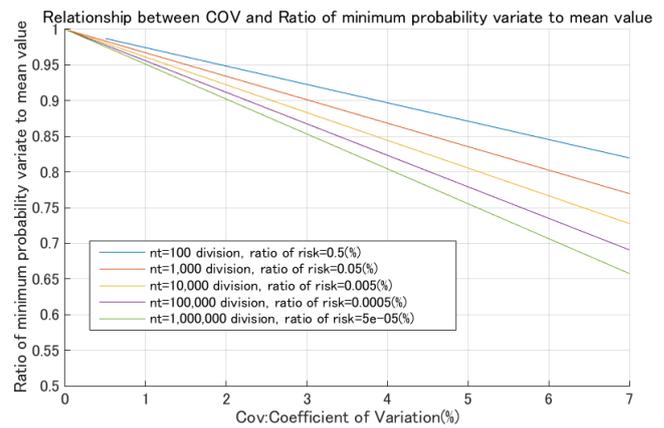


Figure-5 Relationship between COV and ratio of minimum probability variate to mean value

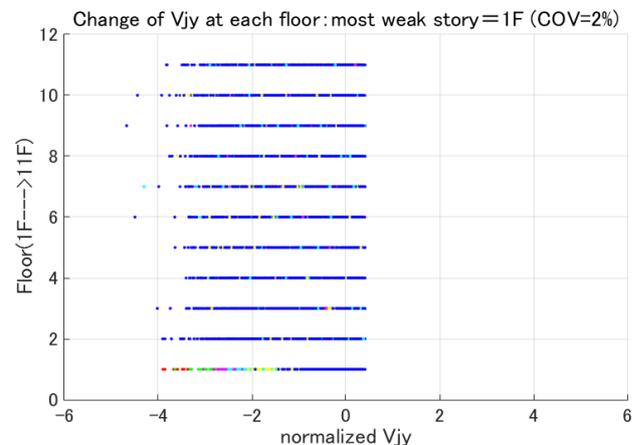


Figure-6 Relationship between the level of  $\mu_{max}$  (large red  $\rightarrow$  small blue) and  $V_{jy}$

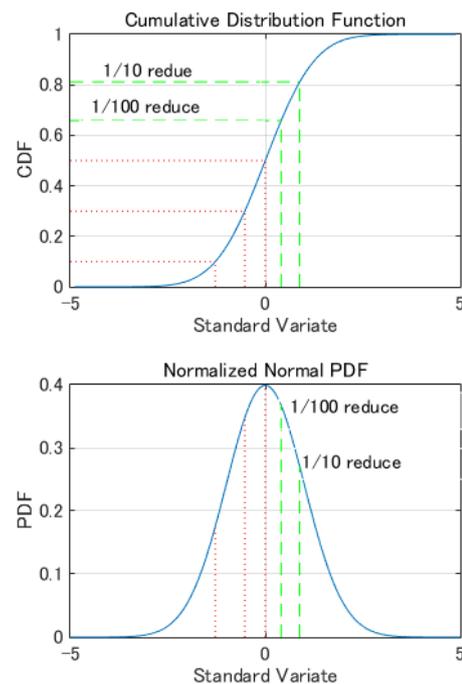


Figure-7 Latin Hypercube Sampling

analysis times to 1/100 and 1/10, there seems to be a considerable number about overlooking of sample cases where  $\mu_{\max}$  is large. However, assuming that the region of large  $\mu_{\max}$  is mostly covered, we consider that it is possible to create a 12 dimensional regression hypersurface using deep learning.

According to this idea, we set the reduced nt at 10,000 times for COV = 2 ~ 4% and 5,000 times for COV = 5% as shown in Table-3. Here, vt in the table is the data set number to test after creation of the learning model. After executing nonlinear dynamic response analysis of MDOF (Multi Degree of Freedom),  $V_{jy}$  and  $\mu_{\max}$  at each story for all analyses are saved in a file and handed over to the program to predict  $\mu_{\max}$ .

Table-3 response analysis times: nt, test data set: vt

COV(%)	nt	vt	reduced nt
2	1,000,000	1,000,000	around 10,000
3	100,000	100,000	around 10,000
4	10,000	10,000	—
5	5,000	5,000	—

note) nt: calculation number of MDOF  
 vt: number of test data set

## B. Creation of a Learning Model to Predict $\mu_{\max}$ of Each Story

$V_{jy}$  and  $\mu_{\max}$  of each story are input data of python program creating the learning model. The learning model was created using Keras and TensorFlow. Keras is a high-level API (Application Programming Interface) that simplifies algorithm implementation. It is easy to handle without extensive knowledge of machine learning [4]. Here, we assume that set of  $V_{jy}$  at each story for each response analysis constitutes features and  $\mu_{\max}$  of the specified floor constitutes target values, and the regression learning model is individually created for each specified floor. In the following calculation examples, cross-validation with  $k = 4$  was performed under the setting condition shown in Table-4. Data set to create the learning model is equal to reduced nt, that is, 10,000 or 5,000 for COV = 2, 3, 4 and 5%, and the test data sets number is as shown in the vt column of Table-3. Although vt is larger than the reduced nt for COV = 2, 3%, this is based on the idea mentioned in section 4.A, that is, reduced nt corresponds to original nt, covering large  $\mu_{\max}$  area. As a result, we consider that the learning model is almost same to the learning model derived from response calculations of 1,000,000 or 100,000 times.

Table-4 Execution conditions for Keras regression

- features 11, data number nt=10,000 or 5,000, (training: validation = 0.75: 0.25), preparation of test data separately
- hidden layer=3, node number=16  
total parameter = 481, batch size=16, epoch number=100
- activation function: 'softplus' in hidden layer,  
not applicable in output layer
- loss function: 'mean\_squared\_error',  
optimization function: 'adam'

## v. Quantitative Estimation of Damage Probability

Analyses in this section were executed on four cases shown in Table-5. The first letter of the case name is the initial character of the observation site of the seismic ground motion, and the second letter is the value of COV. The right end column of the table is the reduced nt. The value corresponding to COV = 2, 3% becomes around 10,000 since the set of  $V_{jy}$  samples changes each time these samples are generated. Table-6 shows the cross validation ( $k=4$ ) average of MAE (Mean Absolute Error) at the final epoch of the analysis case K-2. Since the number of features is equal to 11, the dimension size becomes 12 when  $\mu_{\max}$  is added. Therefore, the regression status cannot be recognized visually. Since average MAE of 4<sup>th</sup> FL is 0.004809 for the  $\mu_{\max}$  range of 4<sup>th</sup> FL is [ 0.5098, 4.7924] shown in Figure-8, its ratio is very small at less than 1%.

Figure-9 shows the transition of the average MAE of cross-validation ( $k=4$ ), and Figure-10 shows the loss transition of the final 4<sup>th</sup> fold. Both figures are examples of K-2, the weakest floor is the 1<sup>st</sup> floor. From Figure-9, we can see the average MAE is decreasing roughly monotonously. And from Figure-10, the loss value rapidly decreases at epoch = 2. There is no sign of over fitting. The transition tendency was similar to the other floors.

Since the analysis building is the same as the reference [1], we defined the same design criteria shown in Table-7. Although  $\mu_{\max}$  in the table seems to be inconsistent to the definition of limit state column, it reflects that the response analysis result of the middle and low-rise buildings tends to give a strict evaluation rather than an actual damage.

Figure-11 shows the prediction result of  $\mu_{\max}$  and damage probability in case K-2. The vertical red dotted lines in the figure represent each limit state listed in Table-7. Here, the total frequency is larger than 1,000,000 for  $\mu_{\max} \geq 2.2$ . This is because the case of  $\mu_{\max} \geq 2.2$  may occur simultaneously on multiple floors in one response analysis. The damage probability:  $P_f$  is calculated by dividing the total frequency exceeding each limit state by the total number of test data set, so  $P_f = 1.70 \times 10^{-5}$  for  $\mu_{\max} \geq 4.4$ . However, since the set of  $V_{jy}$  samples changes each time these samples are generated, the value changed somewhat such as  $P_f = (1.70 \sim 3.30) \times 10^{-5}$ .

Table-8 shows the probability exceeding incipient collapse damage and life safe for all cases. Also, Figure-12 shows this by a diagram. From these figures, when COV is increased for  $\mu_{\max} \geq 4.4$ ,  $P_f$  increases. This indicates the expected trend. However,  $P_f$  is equal to 1 for  $\mu_{\max} \geq 2.2$ , and then incipient collapse cannot be avoided.

However, this probability is the probability under the occurrence of the level 2 JMA Kobe 1995 NS. Here, assuming that the level 2 strong motion has exceeding probability of 10% in 50 years, the annual exceeding probability is as follows.

$$\text{We put } PE_{50} = 0.1 \text{ in } PE_{50} = 1 - (1 - PE_{ann})^{50},$$

$$\text{it will be } PE_{ann} = 1 - 0.9^{1/50} = 0.0021$$

The occurrence probability of the above strong motion is  $2.1 \times 10^{-3}$  (= 1/475: return period 475 years). Therefore, it is

important to note that the annual risk of building damage becomes the value obtained by multiplying  $P_f$  in Table-8 by the occurrence probability of earthquake  $2.1 \times 10^{-3}$ .

In the reference [1], it was an analysis condition in the case where only the  $V_{jy}$  and SK2 of the weakest story fluctuate with  $COV = 2\%$ , which is different from the analysis condition that each story  $V_{jy}$  is fluctuating in this paper. Therefore, we cannot compare them simply. However, if we read the value corresponding to  $\mu_{max} = 4.4$  as a reference value in Figure-13 quoted from reference [1], the approximate values in 50 simulations are as follows:

- lower limit value:  $P_f = 2.0 \times 10^{-6}$ ,
- ensemble average:  $P_f = 5.0 \times 10^{-5}$ ,
- upper limit value:  $P_f = 4.5 \times 10^{-4}$ .

$P_f = 1.70 \times 10^{-5}$  in Table-8 is slightly smaller than the above ensemble average. As described in section 4.A, although we adopt reduced nt to obtain low damage probability for  $COV=2\%$ , it seems that approximate acceptable results were obtained.

Table-5 Analysis case name

case name	COV(%)	reduced/original nt
K-2	2	reduced nt=9,975
K-3	3	reduced nt=9,935
k-4	4	original nt=10,000
k-5	5	original nt=5,000

Table-6 Average MAE of validation data (K-2)

Floor	Average MAE at last epoch
1	0.002745
2	0.002643
3	0.003803
4	0.004809
5	0.004229
6	0.003941
7	0.002079
8	0.002137
9	0.001882
10	0.001217
11	0.001425

note) MAE: Mean Absolute Error for validation data

```
>> final_result_display
Range of maximum ductility factor=[1.57311,4.25826] at 1(FL)
Range of maximum ductility factor=[1.01289,3.98744] at 2(FL)
Range of maximum ductility factor=[0.924474,4.72254] at 3(FL)
Range of maximum ductility factor=[0.509819,4.79242] at 4(FL)
Range of maximum ductility factor=[0.76673,4.25301] at 5(FL)
Range of maximum ductility factor=[0.703271,3.90949] at 6(FL)
Range of maximum ductility factor=[0.880415,2.74277] at 7(FL)
Range of maximum ductility factor=[0.878403,2.06693] at 8(FL)
Range of maximum ductility factor=[0.956953,1.59029] at 9(FL)
Range of maximum ductility factor=[0.86815,1.34723] at 10(FL)
Range of maximum ductility factor=[0.687024,0.917854] at 11(FL)
```

Figure-8 Maximum ductility factor of each story

Average MAE at last epoch  
0.00274549348771692

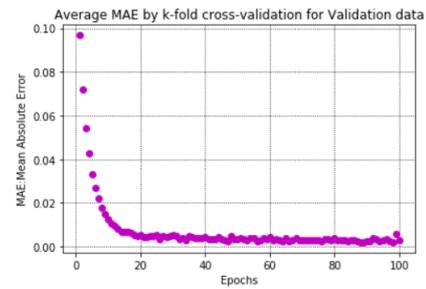


Figure-9 Transition of average MAE (K-2, 1<sup>st</sup> floor)

Loss of last epoch at fold=4 cross-validation  
1.2980899365507817e-05

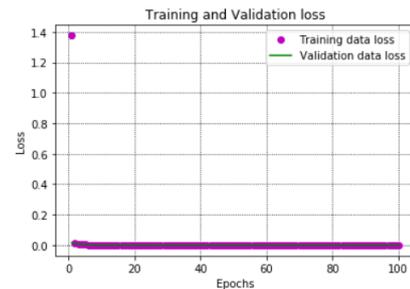


Figure -10 Loss transition (K - 2, 1<sup>st</sup> FL, 4<sup>th</sup> fold)

Table-7 Design Criteria (Definition of Limit State)

MCDR	$\mu_{max}$	Limit State
1/800(0.125%)	0.27	Fully operational(cracking occurrence of nonstructural elements)
1/200(0.5%)	1.1	Operational(cracking occurrence of structural elements)
1/100(1%)	2.2	Incipient collapse(yielding occurrence of structural elements)
1/50(2%)	4.4	Life safe

MCDR:Maximum Column Drift Ratio,  $\mu_{max}$ :Maximum Ductility Factor

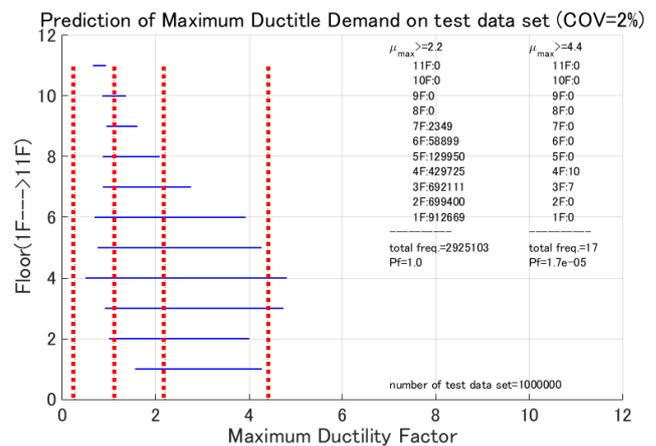


Figure-11 Prediction result and damage probability (K-2)

Table-8 Damage Probability of Incipient Collapse, Life Safe

case name	COV(%)	$P_f(\mu_{max}=4.4)$	$P_f(\mu_{max}=2.2)$	repeated times	CPU time(sec)	weakest floor
K-2	2	$1.70 \times 10^{-9}$	1.00	reduced nt=9,975	8111.09	1 FL
K-3	3	$3.70 \times 10^{-9}$	1.00	reduced nt=9,935	7927.72	1 FL
K-4	4	$2.62 \times 10^{-2}$	1.00	original nt=10,000	8100.1	3 FL
K-5	5	$8.16 \times 10^{-2}$	1.00	original nt=5,000	4143.28	3 FL

CPU time: using note book computer: DELL Precision M3800

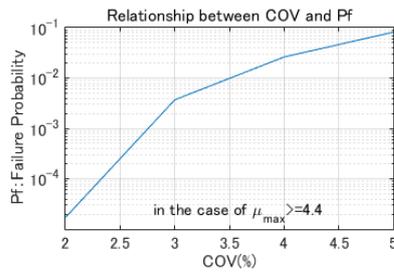


Figure-12 Relationship between COV and damage probability

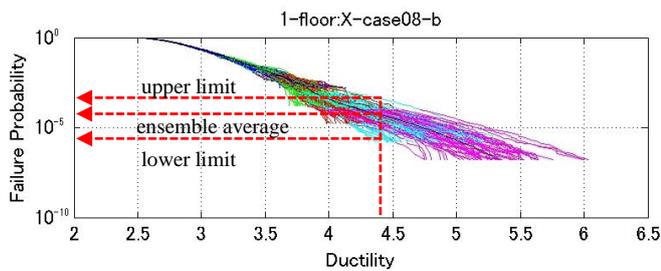


Figure-13  $\mu_{max}$  and damage probability using subset MCMC

## VI. Conclusion

We attempted quantitative damage estimations of an actual RC building assuming level 2 earthquake ground motion of JMA Kobe 1995 NS using deep learning. The results obtained are as follows:

- 1) Importance analysis of input data was carried out using GBDT ( Gradient Boosting Decision Tree). In this paper, input data are story yield shear strength  $V_{jy}$ , first shear stiffness SK1 and second shear stiffness SK2 of each story, and target value is 1 to 10 class of the fracture index (maximum ductility factor:  $\mu_{max}$ ) for the weakest story. It is clarified that  $V_{jy}$  is more influential than SK2, SK1.
- 2) In order to obtain low damage probability of  $10^{-6}$  level, a large number of response analyses is required, but when sample parameters are generated by equally division of the cumulative distribution function using the Latin hypercube sampling, focusing on the lower side in the interval  $[0, 1]$ , we proposed an idea to reduce the number of response calculation to 1/100 or 1/10. This makes it possible to reduce the number of calculation from 1,000,000 times in the standard Monte Carlo Simulation (or 150,000 times in subset MCMC) to 10,000 times.
- 3) Based on the reduced response analysis results, the 12 dimensional regression hypersurface using deep learning was created and it was clarified that the average MAE (Mean Absolute Error) is as small as less than 1% of the maximum ductility factor, in case of  $COV = 2\%$ .
- 4) When  $COV$  of  $V_{jy}$  was set to 2%,  $Pr = (1.70 \sim 3.30) \times 10^{-5}$  was obtained as damage probability exceeding the life safety limit state. Since the analysis condition is different with subset MCMC, simple comparison cannot be done. However, it was confirmed that  $Pr$  is slightly smaller than the ensemble average of  $Pr$  obtained by subset MCMC. Future works are as follows:

- 1) To develop a more precise reduction method covering the region where maximum ductility factor becomes large.
- 2) To examine behaviors for various strong motions other than JMA Kobe 1995 NS.

## References

- [1] Shigeru Kushiyama, "Seismic Damage Estimation of an Actual Reinforced Concrete Structure Using Subset MCMC", Scientific Research, Open Journal of Civil Engineering, 2013, 3, pp 136-142.
- [2] S. K. Au and J. L. Beck, "Subset Simulation and its Application to Seismic Risk Based on Dynamics Analysis", Journal of Engineering Mechanics, 2003, pp.901-917.
- [3] Shigeru Kushiyama, "Basic Study on Convergence using Generic Metro-polis Algorithm (in Japanese)", Bulletin of the Faculty of Engineering, Hokkai-Gakuen University (No. 32), 2005 February, pp. 27 - 40.
- [4] François Chollet, "Deep Learning with Python", Manning Publication, 2017.

About Author (s):



We introduced that we can calculate low damage probability of an actual RC building using deep learning instead of subset MCMC. The regression hypersurface is created with number of response calculation within allowable range.