17th World Conference on Earthquake Engineering, 17WCEE
Sendai, Japan - September 13th to 18th 2020

FUNDAMENTAL STUDY ON VERTICAL RESPONSE CONTROL OF ISOLATED BUILDING USING DEEP REINFORCEMENT LEARNING

M. Yukawa(1), T. Furuhashi(2), T. Iwasaki(3), Y. Takeuchi(4), Y. Akama(5), K. Ichikawa(6)

(1) *Engineer, Obayashi Corporation, yukawa.masaki@obayashi.co.jp*
(2) *Professor, Nihon University, furuhashi.takeshi@nihon-u.ac.jp*
(3) *Engineer, Obayashi Corporation, iwasaki.tomohiro@obayashi.co.jp*
(4) *Engineer, Obayashi Corporation, takeuchi.yoshitaka@obayashi.co.jp*
(5) *Engineer, Ark Information Systems INC., akama.yuki@ark-info-sys.co.jp*
(6) *Engineer, GROOVE X INC., kyosuke.ichikawa@groove-x.com*
…

Abstract

Horizontal response acceleration can be reduced by the adoption of laminated rubber bearings in base-isolated building at the time of earthquake, but the vertical response acceleration may be amplified depending on the type of base isolation devices adopted. Especially in nuclear facilities where equipment and piping are installed, amplification of vertical response acceleration causes a problem in the design of equipment and piping. In order to control the vertical response acceleration of base-isolated building, application of inertial mass dampers (DM dampers) is assumed, and the possibility of semi-active control methods using deep reinforcement learning in switching controller of inertial mass ratio is investigated. As a result, it is confirmed that the learning system is effective for controlling vertical response acceleration by using the controller obtained from the non-linear models. Seismic performance of the base isolation devices and upper building, and equipment design can be improved by the application of the proposed method in this study. The findings obtained in this study are shown below:

1) In the vertical response control for a base-isolated building using DM dampers, a controller can be generated reasonably by performing a parameter study on the following: the final value of ε in ε- greedy, the number of experiences to be saved, the update interval of the target network, and the action interval setting of 0.1 second. In addition, it is shown that the system is effective for reducing responses in non-linear models, however the applicable condition is limited.

2) Even in the case where the response displacement is small, it is confirmed that the DM dampers are effective for reducing responses. Applying a controller based on passive control, the building response tends to be increased due to the phase difference between the maximum response displacement and the maximum response acceleration, whereas the use of controller based on semi-active control obtained from the deep reinforcement learning enables the response increase caused by the phase difference to be avoided.

3) Semi-active control with a reward function, which corresponds to the design target value to use DM damper as effectively as possible and to reduce the displacement deviating from the design range, could improve the seismic feasibility than passive control.

Keywords: base isolation; Deep reinforcement learning; inertial mass; semi-active control; nuclear power plant

17th World Conference on Earthquake Engineering, 17WCEE
Sendai, Japan - September 13th to 18th 2020

1. Introduction

Horizontal response acceleration can be reduced by the adoption of laminated rubber bearings in base-isolated buildings at the time of earthquake, but the vertical response acceleration may be amplified depending on the type of base isolation device adopted. For a base-isolated reactor building subjected to a large input motion, the exceedance of design criteria due to tensile deformation of laminated rubber and amplification of the vertical response acceleration of equipment and piping installed inside the building are specifically considered as study objects. The laminated rubber bearing, e.g. [1], which has been developed to be applied to base-isolated reactor building, has higher compression stiffness than the horizontal one, hence quite small vertical response displacement and velocity, and which makes it difficult for a hysteretic or viscous damper being installed to exert its power of control sufficiently. In addition, an asymmetric bilinear hysteresis with a stiffness on the tension side different from the compression side is used for the non-linear characteristics of the base isolation device, which may cause a complicated non-linear vibration in which the sequential input level and the initial condition would vary [2] in the transient response. The variation makes it difficult to perform optimal design by means of conventional passive damping devices.

In this study, considering a base-isolated PWR (Pressurized Water Reactor) reactor building as a target, the inertial mass damper (DM damper) [3, 4], which is effective even for a small deformation, is used so as to carry out numerical analyses to improve seismic safety margin of base isolation devices and upper buildings and to reduce responses of the equipment installed in the building by constructing a control system that reduces the non-linear vibration by switching the inertial mass ratio. Considering the necessity of building a complicated control system due to the non-linearity of subjected vibration system and the damper performance, deep reinforcement learning methods are employed to implement semi-active control which switch the inertial mass ratio of DM damper sequentially.

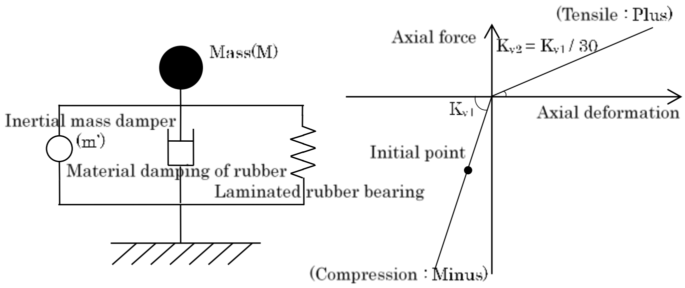
2. Analysis Model

Vibration characteristics used for the seismic response analysis and the analysis model are shown in Table 1 and Fig. 1, respectively. The direct integration method is used for the seismic response analysis. The subject building of this analysis is PWR reactor building [5], and the isolation device employed in this analysis is thick LRB (Lead Rubber Bearing). In order to focus on the responses from base isolation devices, the building is modelled as a single-lumped-mass model where the weight of the building is concentrated at the upper basemat level. The inertial mass ratio (γ=m'/M) of DM damper is chosen among γ=0.00, 0.25, 0.50 and 1.00 based on the constitutive law obtained by the machine learning, and the vibration control is performed by sequential switching of the inertial mass ratio. The reason of the selection of inertial mass ratio values being equal to or lower than 1.0 was because it is known that the acceleration response decreases as the inertial mass ratio increases in the resonance region but increases in the large frequency ratio region. The ratio was set at equal to or lower than 1.0 because normally equipment with wide range of frequencies is installed in a reactor building. The initial ratio of γ=0.5 was set for the machine learning, and the difference occurred in internal force in DM damper by switching was processed by the convergent calculation. If the difference could not converge by the convergent calculation within the convergence criteria, the internal force is carried over to the next integral step as a residual force.

Table 1 – Vibration characteristics of analysis model

Mass	(ton)	236190
Stiffness	(kN/m)	9.32×10^{8}
Damping coefficient*	(kNs/m)	5.94×10^{5}
Inertial mass ratio	—	0, 0.25, 0.5, 1.0
Initial deformation	(m)	-1.86×10^{-3}
Total thickness of rubber	(m)	0.52

*: Material damping ratio of rubber (2.0%)



(a) Isolation layer (b) Hysteresis of laminated rubber bearing

Fig. 1 – Analysis model of isolation device

3. Learning method

Deep reinforcement learning which enables automatic extraction of feature quantity is employed in this study to construct the control system which can correspond to non-linear vibrations of base isolation devices and/or a mode change due to the switching of inertial mass ratio. To deal with the expected complicative response, Double Deep Q Network (DDQN), with which bias (overvaluation) of the action value Q can be reduced, is used as an algorithm [6] in this study. ChainerRL, a library of Python, a general programming language, is used in the DDQN.

Input values to Neural Network (NN) are composed of 201 items in total, including the response displacement and response acceleration for the last 100 steps and the inertia mass ratio of the upper building used in the current step. The middle layer of NN is composed of 5 layers. ReLU (Rectified Linear Unit), as an activating function; Adam (Adaptive Moment Estimation), to minimize the error function; and stochastic gradient descent, to update the Online Network are used in this study. The ε-greedy is applied to search for the optimal solution. In the deep reinforcement learning, the setting of the hyperparameter is known to affect the result of learning significantly, however the selection method of its optimal values for the response control is unestablished. Therefore, in this study several parameter studies are conducted on the hyperparameters, such as update interval at which the inertial mass ratio is switched, the final value of ε in the ε-greedy, the number of experiences saved in the learning process, and the update interval of the Target Network, to investigate the impact on the result of learning. The hyperparameters used in this study are shown in Table 2.

Table 2 – Hyper parameters

	case0-0-0	case0-0-2	case0-1-0	case0-1-2	case2-0-0	case2-0-2	case2-1-0	case2-1-2
Last value of ε	0.01	0.01	0.01	0.01	0.20	0.20	0.20	0.20
Replay buffer size (episode)	1	1	100	100	1	1	100	100
target update interval (episode)	1	50	1	50	1	50	1	50
Number of episode (episode)	5000							
Learning rate	1.0×10^{-5}							
Last step of ε damping (episode)	2500							
Initial value of ε	1.00							
Batch size	32							

The reward functions used in this study are shown in Fig. 2. Reward functions are set for the deformation in axial direction and absolute acceleration, respectively. For the deformation in axial direction, large penalty is given to the response exceeding the design criteria (i.e. tensile deformation corresponding to the tensile surface pressure of 1MPa: approximately 12mm), whereas the penalty for the absolute acceleration is given using the quadratic function which takes half the value of the penalty given to the deformation in axial direction at 1.0G. These settings of reward functions are determined expecting to secure the deformability without deviating from the design criteria and the utmost exertion of the response reduction effect of DM damper because the vertical response displacements of base isolation devices are quite small. The learning is performed for the time histories for the whole period. For this study, the optimal control is to obtain the response reduction effect in the responses from multiple equipment installed inside the building, and therefore the reduction of floor response spectra for the wide range of frequencies is expected to be achieved by the reduction of time histories for the whole period rather than the reduction of the maximum acceleration.

For the control system learning, 22 seismic waves conforming to the notification spectrum with maximum acceleration of 6.0 m/s^2 provided based on the past reports [7], 40 learning waves generated by positive/negative inversion, and 4 verification waves are used (Fig. 3). Furthermore, random phase characteristics and the envelope function of the aging characteristics using the method of Jennings et al. are used. Input values of assumed magnitude M and equivalent source distance Xeq to the envelope function are

M8.3 and Xeq119.466km, respectively, which are equivalent to those of the assumed Sanriku-oki earthquake [8]. One seismic wave is set to be one episode under learning process, and the learning order of each seismic wave is determined by a uniform random number.

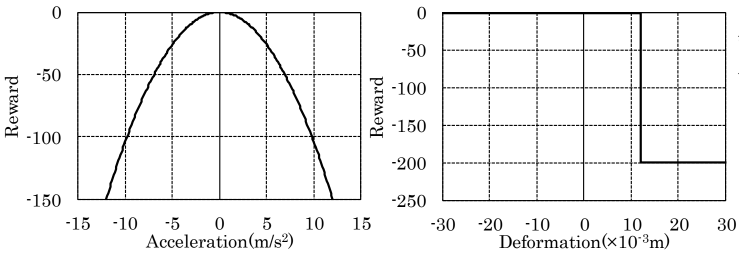


Fig. 2 – Reward function

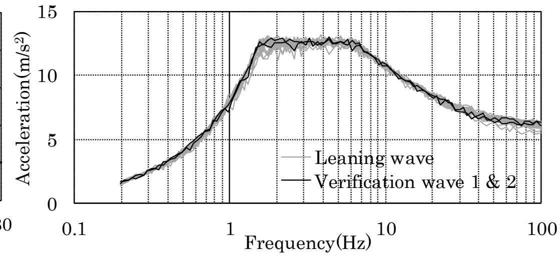


Fig. 3 – Analysis model of isolation device

4. Result of Learning

The semi-active control for the non-linear vibration system is constructed using the base-isolated building model shown in Chapter 2. As the frequency sequentially changes in the non-linear vibration system, it is difficult to switch the inertial mass once every period like in the linear vibration system. Therefore, in this study the results of learning using the action intervals of 0.01 second, 0.1 second, and 1.0 second respectively are compared. Considering the compression frequency without DM damper is 10Hz, it is possible to select actions once or more per period when the action intervals of 0.1 second is used. As stated previously, the evaluation of the performance of semi-active control using the control system of the analysis case producing the highest rewards in this report.

4.1 Active interval of 0.1 second

The results of learning for each analysis case are shown in Table 3. The control system in the episode with the maximum rewards is used to implement the response analysis for 4 verification waves to calculate the average value of the respective response to give the response values in the table. The analysis results by the passive control with a constant inertial mass ratio is also described. γ=0.00 indicates the analysis result without DM damper. It was confirmed in Table 3 that the case0-1-0 provides the maximum rewards. It was also confirmed that the use of semi-active control tends to provide larger rewards, thereby the adequate learning is carried out. The relationship between the number of episodes and the rewards is shown in Fig. 4. It was confirmed in Fig. 4 that the rewards mostly converged after 2200 episodes, although more learning experiences than in linear vibration system are necessary, adequate learning can be achieved even for the non-linear vibration system.

The time history response waveform under semi-active and passive controls using verification wave 1, the relationship between the response displacement and the response acceleration, and the plot of the selected inertial mass ratio are shown in Fig. 5, Fig.6, and Fig.7, respectively. According to Fig. 5, the response displacement and the response acceleration tend to increase under the passive control with γ=1.00 and γ=0.00 when the tensile response occurs, whereas the increase in the response is not confirmed even on the tensile side under semi-active control. According to Fig. 6, the non-linear vibration significantly develops toward the tensile side under the passive control, whereas the response does not increase even on the tensile side under the semi-active control and the significant non-linear behavior is not confirmed. Hence, through this learning, the control law corresponding to the non-linear vibration appears to be constructed. Under the semi-active control, the response acceleration decreases compared to the case under the passive control with γ=0.0, but the reduction effect for the response displacement which becomes equal to or less than the design target value for the reward function setting is restricted. According to Fig. 5, it is more remarkable that the response displacement in non-linear model tends to be evaluated larger under the passive control with γ=1.00 than the case with γ=0.00, due to the phase difference, than in linear model. Under the semi-active control on the other hand, the maximum displacement is suppressed almost the same as the case under the passive control with γ=0.00, and it is confirmed that the degradation of the DM damper performance against the non-linear vibration system can be prevented thereby.

The floor response spectra under the semi-active and the passive controls (γ=0.00, 1.00) are shown in Fig. 8. According to response spectra, the response reduction effect is confirmed in the wide range of frequencies under the semi-active control. At the peak frequency, the response reduction of 30 to 50 percent is achieved. In comparison with the case under the passive control with γ=1.00, the dominant frequency occurs at the shorter period. The shift in dominant frequency appears to be due to the fact that the time history other than γ=1.00 is selected exists under the semi active control, and the control law effectively works against the large response acceleration that occurs consistent with the resonance period. Under the passive control, the response increases in the wide range of frequencies due to possible impact of the mollification of the resonance curve, whereas under the semi-active control, the range of dominant frequency is slightly narrower, so that it was confirmed that the semi-active control is effective to reduce the responses of upper building and/or equipment installed in it.

Table 3 – Result of learning model with action interval of 0.1s

	Reward ($\times 10^4$)	Max tensile deformation ($\times 10^{-3}$m)	Max respose acceleration (m/sec^2)	Variance response displacement ($\times 10^{-3}$m^2)	Variance response acceleration ($\times 10^{-2}$m^2/sec^4)
γ=0.00	-7.25	3.28	16.24	4.36	6.33
γ=1.00	-5.17	7.89	15.22	14.58	4.52
case0-0-0	-4.23	5.36	12.50	6.60	3.70
case0-0-2	-4.15	4.41	13.83	6.13	3.63
case0-1-0	-3.91	3.19	14.06	6.12	3.42
case0-1-2	-4.08	2.18	10.17	4.76	3.57
case2-0-0	-4.34	3.54	10.35	5.24	3.80
case2-0-2	-4.58	2.48	11.06	5.46	4.01
case2-1-0	-4.15	6.27	13.08	5.45	3.63
case2-1-2	-4.07	2.60	10.52	5.31	3.56

*Underline show the best value of each case

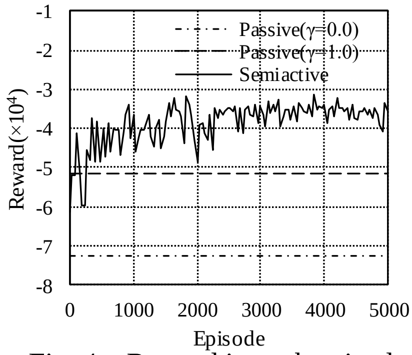


Fig. 4 – Reward in each episode (case0-1-0)

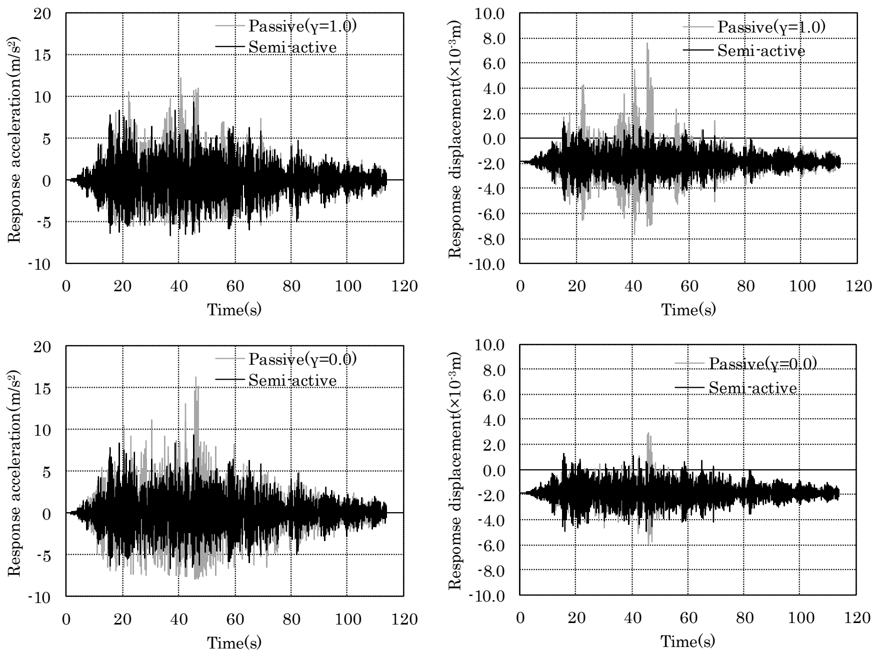


Fig. 5 – Time history of response waveform of analysis (case0-1-0)

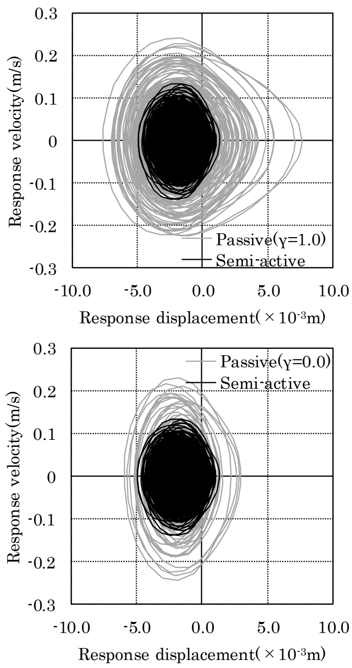


Fig. 6 – Relationship between response displacement and response velocity (case0-1-0)

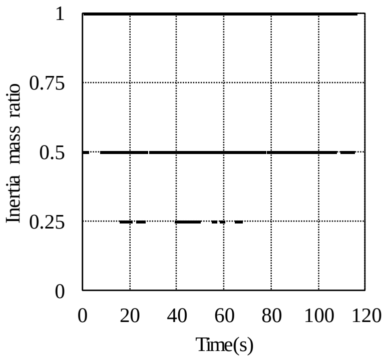


Fig. 7 – Time history of selected inertia mass ratio (case0-1-0)

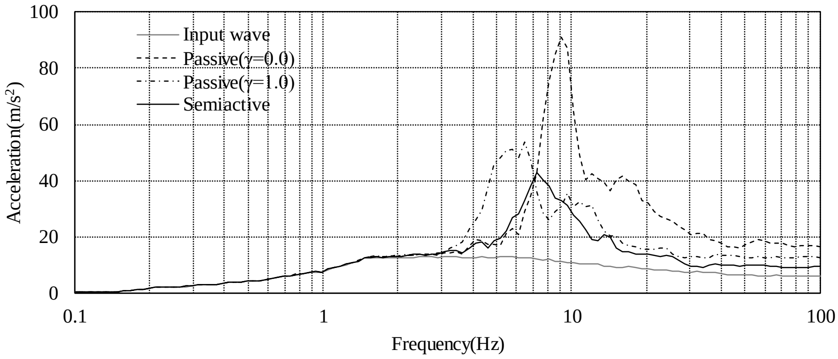


Fig. 8 – Response acceleration spectra (h=5.0%)

4.2 Active Interval of 1.0 Second

In this section, the result of learning with the active interval of 1.0 second is described to study whether the number of inertial mass switching can be reduced by the learning method. The list of the results of learning is shown in Table 4. The control system obtained from case0-0-2, whose rewards are the highest in the list, is used to perform the evaluation using verification waves.

The time history response waveforms by semi-active and passive control, the plots of the selected inertial mass ratio, and the floor response spectra comparison are shown in Fig. 10, Fig. 11, and Fig. 9, respectively. From the time history response waveforms in Fig. 10 it can be said that some response reduction effect is confirmed in the vicinity of 37 seconds, etc., but the response for most of the time range is similar to the one obtained under the passive control with γ=1.00.

As shown in the response spectra in Fig. 9, some reduction in spectrum peak at dominant frequency under semi-active control compared to the one under passive control with γ=0.00 is confirmed, but the peak is nearly equal to the one under passive control with γ=1.00. Hence, the response reduction effect here is mostly due to the presence of DM damper, and by this learning method the selection of wider active interval reduces the performance of the semi-active control. It is confirmed in Fig. 11 that γ=1.00 is selected in the range where relatively large response including the principal motion of 17 to 48 seconds are shown, whereas γ=0.50 is selected in other range where the response is smaller. Hence, by setting the active interval longer than the natural period of the building, the control system close to the sequence control, which increases the inertial mass ratio when the responses including the principal motion are large whereas reduces when the response is small, is constructed. This may provide the evaluation results under semi-active control nearly equal to those of under passive control with γ=1.00 in the rewards and/or the response spectra in which impacts from the responses including principal motion are dominant.

Table 4 – Result of learning model by Action interval 1.0s

	Reward ($\times 10^4$)	Max tensile deformation ($\times 10^{-3}$m)	Max respose acceleration (m/sec^2)	Variance response displacement ($\times 10^{-3}$m^2)	Variance response acceleration ($\times 10^{-2}$m^2/sec^4)
γ=0.00	-4.78	3.28	16.24	4.36	6.33
γ=1.00	-3.45	7.89	15.22	14.58	4.52
case0-0-0	-3.40	8.08	15.29	13.30	4.46
case0-0-2	-3.32	6.35	13.21	12.39	4.37
case0-1-0	-3.39	7.89	15.22	13.48	4.46
case0-1-2	-3.40	7.89	15.22	13.95	4.48
case2-0-0	-3.42	7.12	13.92	13.38	4.49
case2-0-2	-3.39	9.66	15.68	13.82	4.46
case2-1-0	-3.41	7.89	15.22	14.08	4.48
case2-1-2	-3.44	8.04	15.27	14.23	4.53

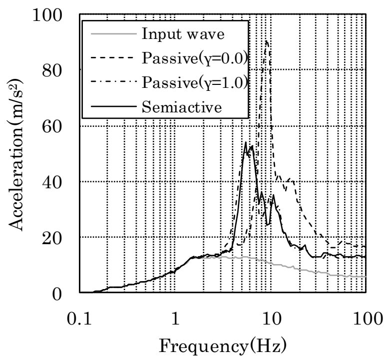


Fig. 9 – Response acceleration spectra (h=5.0%)

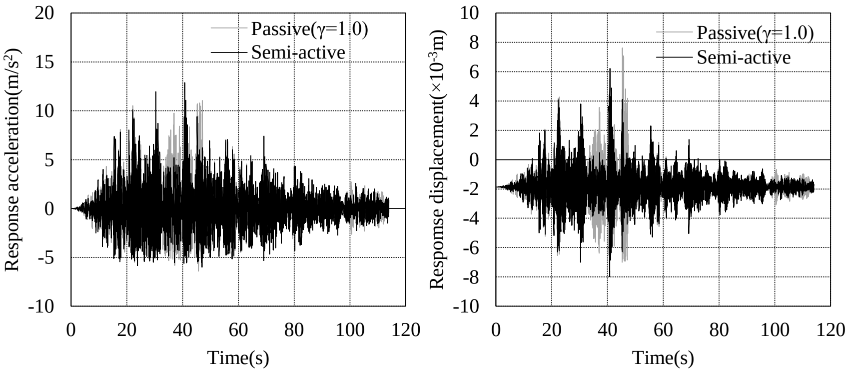


Fig. 10 – Time history of response waveform of analysis (case0-0-2)

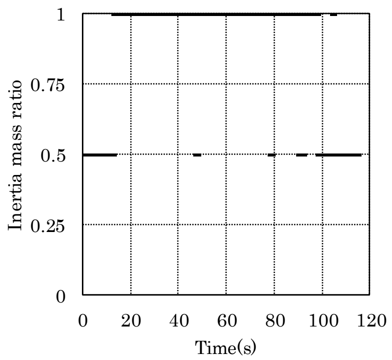


Fig. 11 – Time history of selected inertia mass ratio (case0-0-2)

5. Consideration on Excessive Input

In Chapter 4, the study considering spectrum wave with the maximum acceleration of 6.0m/s^2 as the design level input was described. It is reported that the margin against the horizontal shear breaking strain of the base-isolated building is about 2.0 times [9], therefore the impact study on the similar level of input for response analysis in vertical direction may also be conducted. In that case, because the periodic solution depends on the input level in the non-linear vibration system, the control system constructed in Section 4.2 (hereinafter referred to as system A) may not operate effectively. Hence, the control system was constructed by learning using input acceleration multiplied by the coefficient generated by uniform random number which takes the values between 1.0 and 2.0 (hereinafter referred to as system B) for the comparison with system A. The active interval of 0.1 second was employed for this comparison.

5.1 Construction of System B

The constitutive law is constructed for semi-active control using learning wave factored by uniform random numbers. The same learning method as described in Chapter 3 is used, and 12 verification waves (4 seismic waves times 3 factors, namely, 1.0, 1.5, and 2.0) are used in the performance evaluation.

The list of the results of learning is shown in Table 5. The analysis case with the highest rewards was case2-1-0 in which the hyperparameter different from those in Section 4.2 is used. The smaller the distributions of the response displacement and the response acceleration, the rewards tend to become larger, which indicates that the reward function of the deformation in the axial direction and that of the response acceleration mutually affect each other. The relationship between the number of episodes and the reward function is shown in Fig. 12. As shown in Fig. 12, the rewards tend to be converged at equal to or greater than 1000 episodes, it appears that the adequate learning is carried out.

Table 5 – Result of learning model by Action interval 1.0s (Exceed design wave)

	Reward ($\times 10^4$)	Max tensile deformation ($\times 10^{-3}$m)	Max respose acceleration (m/sec^2)	Variance response displacement ($\times 10^{-3}$m^2)	Variance response acceleration ($\times 10^{-2}$m^2/sec^4)
γ=0.00	-20.93	31.17	49.73	30.54	17.93
γ=1.00	-15.55	47.50	29.35	126.45	10.82
case0-0-0	-11.48	45.51	39.60	39.14	9.51
case0-0-2	-11.58	39.50	30.16	41.48	9.71
case0-1-0	-11.08	42.19	39.64	35.09	9.30
case0-1-2	-11.09	21.28	37.64	25.91	9.46
case2-0-0	-11.68	36.48	36.04	48.67	9.43
case2-0-2	-12.01	32.16	37.01	45.62	10.04
case2-1-0	-10.93	32.37	38.98	30.17	9.32
case2-1-2	-11.26	29.77	44.29	35.66	9.41

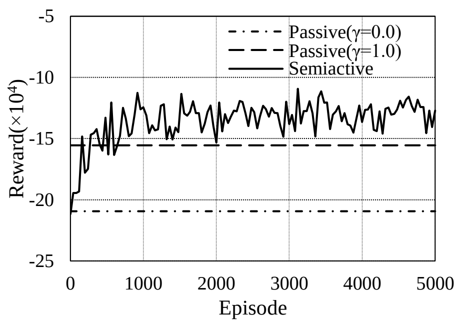


Fig. 12 – Time history of selected inertia mass ratio (case2-1-0)

5.2 Comparison of System A with System B

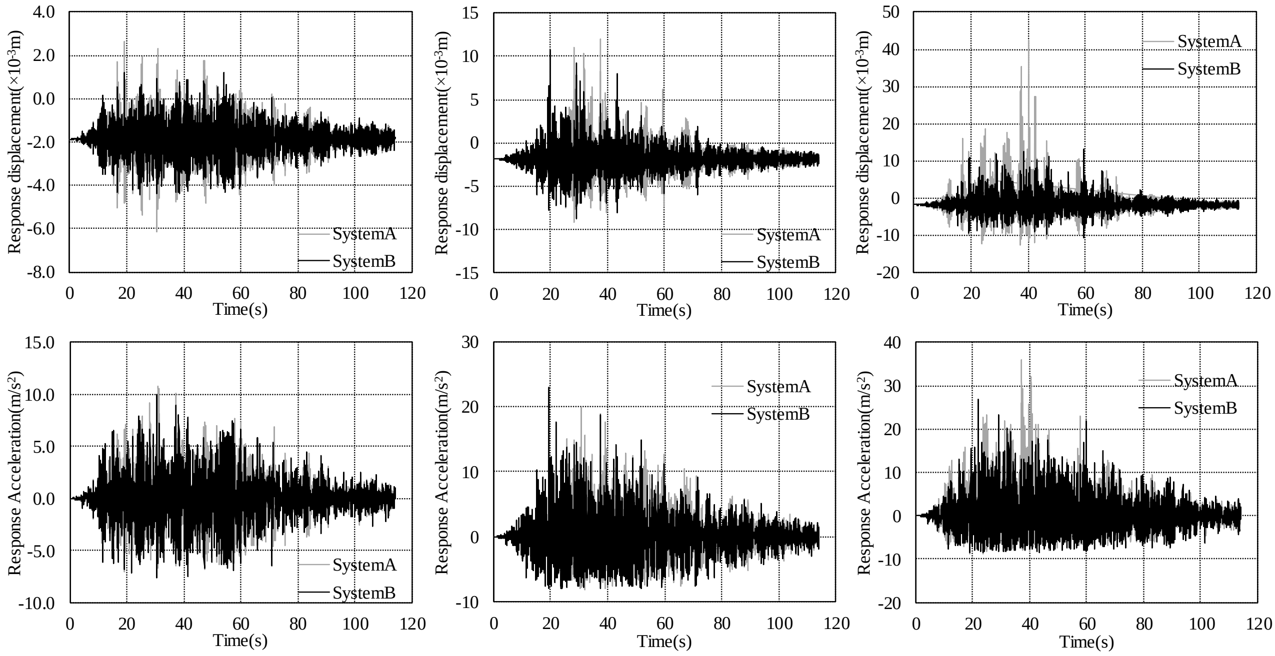
The response analyses are implemented using verification wave 2 multiplied by 1.0 times, 1.5 times, and 2.0 times, responsively to evaluate performances of system A and system B. The list of the rewards and the maximum response values for each analysis case is shown in Table 6. It is found that the rewards for every factored input in both systems A and B are evaluated larger than the ones under the passive control. When the passive control is used, significant response amplification is occurred using excessive input with coefficients 1.5 and 2.0 times in comparison with the coefficient 1.0 times. The response amplification from 1.0 times input to 1.5 times input in system A and to 2.0 times input in system B is reduced compared to the ones under passive control. Under passive control, the tensile response exceeds the design target value (12mm) at 1.5 times input, whereas under both systems A and B it satisfies the design target under semi-active control. Therefore, it is confirmed that the improvement of the earthquake resistance can be expected by the semi-active control.

Next, the comparison of time history response waveforms is shown in Fig. 13. As shown in the comparison of maximum responses, the response reduction effect is larger in System A than in System B for 1.0 times input and 1.5 times input. Regarding 2.0 times input, for the system A the response value is as large as under passive control, whereas for the system B it is as low as the one with 1.5 times input under semi-active control. This is because the control law is constructed to be most effective for approximately 2.0 times input with large response in system B and therefore exhibits insufficient exertion for smaller input such as approximately 1.0 to 1.5 times input. Hence, it was confirmed from the result of this study that the control system that is effective to reduce the response up to approximately 1.5 times input level can be constructed even when the design level input only is used as learning wave. It was also confirmed that a control system, which shows a response reduction effect equivalent to or greater than that under passive control even for design level input and is highly effective for response reduction when subjected to excessive input, can be constructed by using the excessive wave as a learning wave.

Table 6 – Comparison of System A and System B (Verification wave2)

	Amplification of input wave	Reward ($\times 10^4$)	Max response displacement ($\times 10^{-3}$m)		Max response acceleration (m/s^2)	
Passive (γ=0.0)	1.0	-8.4	3.3	(1.00)	13.6	(1.00)
	1.5	-26.1	18.8	(5.73)	36.5	(2.68)
	2.0	-36.5	31.2	(9.51)	45.3	(3.32)
Passive (γ=1.0)	1.0	-5.4	7.1	(1.00)	12.8	(1.00)
	1.5	-11.9	22.9	(3.22)	21.3	(1.66)
	2.0	-29.6	47.5	(6.68)	29.3	(2.29)
Semi-active	1.0	-4.0	2.6	(1.00)	10.7	(1.00)
	1.5	-9.3	9.2	(3.52)	14.6	(1.36)
	2.0	-17.2	35.3	(13.51)	36.0	(3.36)

*The ratio with amplification 1.0 in parenthese



(a) 1.0 times (b) 1.5 times (c) 2.0 times

Fig. 13 – Time history of response waveform of analysis

6. Conclusion

In this study, DM damper is applied in the aim of controlling vertical response in base-isolated reactor building to study the feasibility of semi-active control using deep reinforcement learning for the inertial mass ratio switching. On the construction of control system, learning is performed using non-linear models as a subject in order to confirm the effectiveness of this system as well as the possibility of construction of a system which can control the significant non-linear vibration for excessive input. The improvement of the seismic performance of base isolation devices, the upper building, and equipment design can be expected by use of this system. The findings obtained in this study are as follows:

1) An appropriate control system can be constructed for the vertical response control by DM damper in the base-isolated reactor building, with active interval of 0.1s, by performing a parameter study for final value of ε in ε- greedy, the number of experiences to be saved in the learning process and the Target Network update interval. This method also exerted the response reduction effect where the vibration system is either linear or non-linear, and the method is also effective for the containment of the non-linear vibration in the non-linear vibration system. Note that this is applicable to the restricted number of analysis case, the further consideration of combination of parameters is necessary.

2) DM damper effectively exerts the response reduction effect nevertheless the response displacement is small. The phenomenon found in the study using DM damper under passive control that the building response increases due to the phase difference between the response displacement and the response acceleration is confirmed to be avoidable by the semi-active control with control system obtained by this learning.

3) Within the design range, the performance of DM damper can be used as effectively as possible by setting reward functions corresponding to the design target value so that the reduction of responses is avoided. Furthermore, it was confirmed that seismic feasibility can be improved by reducing the displacement deviating from the design range compared to that of passive control. On constructing control system corresponding to excessive input, it was found necessary to set reward function which increases the amount of penalty as displacement increases over design range.

The subject of this study was base-isolated reactor building. Nevertheless, damage on base-isolated buildings due to the vertical response is not reported at this time, the vertical response acceleration over 1.0G is reported in the observation records of the 2016 Kumamoto Earthquake and the 2018 Hokkaido Eastern Iburi Earthquake. In the future, other than the base-isolated reactor building, the large vertical response may have the impact on the design of the base-isolated building adopting base isolation devices with weaker vertical stiffness and/or the countermeasures against the vertical vibration for floor/beam of the upper building. Further study on the improvement of precision of machine learning and the applicability of DM dampers should be carried out.

7. References

[1] S. Harada, S. Nakagawa, Y. Okouchi, S. Inaba, Y. Takeuchi, M. Yukawa (2016): A development of base-isolation system applied in large input region. (Part1) The outline, Summaries of Technical Papers of Annual Meeting, Architectural Institute of Japan, Structure Ⅱ, 1337-1338

[2] T. Aihara, H. Kumano (2008): Forced Vibration in an Unsymmetric Piecewise-Linear System Excited by General Periodic Force Functions (Response Vibration in Main and Superharmonic Response and Analysis of Stability Criterion), Bulletin of the Japan Society of Mechanical Engineers, Vol.74, No.743, 8-15

[3] T. Furuhashi, S. Ishimaru (2006): RESPONSE CONTROL OF MULTI-DEGREE SYSTEM BY INERTIAL MASS (Study on response control by inertial mass No.2), Journal of Structural and Construction Engineering (Transactions of AIJ), No.601, 83-90

[4] T. Furuhashi, S. Ishimaru (2008): MODE CONTROL SEISMIC DESIGN WITH DYNAMIC MASS, 11-0028, The 14th World Conference on Earthquake Engineering, Beijing, China, 12-17

[5] The Japan Electric Association (1987): Technical guidelines for aseismic design of nuclear power plants JEAG4601-1987

[6] Hado van Hasselt, Arthur Guez, David Silver (2016): Deep Reinforcement Learning with Double Q-Learning, Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence (AAAI-16), pp.2094-2100

[7] H. Shimizu, et al. (2010): Development of seismic isolated light-water reactor plant in the next generation (Part2: Design Conditions of seismic isolation system), Summaries of Technical Papers of Annual Meeting, Architectural Institute of Japan, Structure Ⅱ, 1133-1134

[8] Nuclear Regulation Authority: http://www2.nsr.go.jp/data/000227027.pdf (accessed 2018.4.12)

[9] S. Yoshida, et al. (2015): Development of seismic isolation technology for nuclear plant. (Part23: Study on partial failure of seismic isolators for fragility evaluation of isolated reactor building.), Summaries of Technical Papers of Annual Meeting, Architectural Institute of Japan, Structure Ⅱ, 1163-1164

