EARTHQUAKE GROUND MOTION MATCHING ON SHAKING TABLE USING REINFORCEMENT LEARNING CONTROLLER

S. Larbi[(1)], N. Bourahla[(2)], Y. Bourahla[(3)], K. Choutri[(4)]

[(1)] *PhD student, Civil Engineering Department, university of Saad Dahlab Blida1, Algeria, selma.h.larbi@univ-blida.dz*
[(2)] *Professor, Civil Engineering Department, Ecole Nationale Polytechnique, Algeria, nouredine.bourahla@g.enp.edu.dz*
[(3)] *PhD student, Informatics department, University of Sciences and Technology Houari Boumedien, Algeria, bourahla.yasser@gmail.com*
[(4)] *Research assistant, Aeronautic department, University of Saad Dahlab Blida, Algeria, k.choutri@univ-blida.dz*

Abstract

Shaking tables are an essential tool in dynamic testing to analyse the nonlinear behaviour of structures, structural and non-structural elements, evaluate and predict their seismic responses to dynamic excitations similar to real earthquakes. A high fidelity in signal reproduction on the shaking table platforms, achieved by a robust control system, is therefore required. Many control techniques are developed and implemented in shaking tables to enhance their performance over a wide range of frequencies. Several controllers including linear, nonlinear, predictive, iterative and adaptive controllers are among others popular shaking tables controllers. However, the distortion between the original earthquake records and the reproduced signals is still problematic in shaking table tests. Recently, in contrast to conventional control methodologies, several Artificial Intelligence algorithms have proven to be more robust to learn nonlinear control laws. In this paper, the potential of a Reinforcement Learning (RL) approach to reduce signal distortion for an electric bi-axial shaking table is investigated through the application of a Q-learning algorithm. A numerical model of the shaking table system is developed in MATLAB/Simulink to simulate the environment in which the RL controller will interact and receives information about the system in form of a state signal. Then, the controller selects the action (command signal) from a discretized action space, to be applied to the shaking table with the main goal of maximizing the cumulative discounted reward. The action is defined as good if the error between the desired and the reproduced acceleration is remaining within the limits of the acceleration error tolerance and then the agent receives a positive reward. Otherwise, it receives a negative reward. The purpose of this control scheme is to develop an intelligent controller that learns, in long term, to select an appropriate action directly from the environment which is a shaking table system. The superior performance achieved by the RL controller is assessed in terms of the root mean square (RMSE) error and mean square error (MSE) between the measured response and the desired signal after applying the optimal command signal for several simulations using real earthquake records as acceleration input signals. The main benefit in practice of the RL controller over the classical controller is its continuing improvement (learning) with ongoing use of the testing on the shaking table.

Keywords: shaking table, control, machine learning, Reinforcement Learning, Q-learning.

1. Introduction

Optimal control techniques are efficient alternatives to adapt controller's parameters for dynamic systems, to deal with the system nonlinearities and to enhance the tuning process of the controller without causing system divergence or instability. A considerable research in adaptive control has been developed for shaking tables and has shown the important capabilities of the adaptive tuning of such complex systems [1-5]. However, these techniques generally depends on an accurate knowledge of the plant, which is hard to achieve in most cases. Moreover, the performance of the adaptive control approaches turn out to be inefficient when facing fast variations of the system characteristics [6]. In the last decades, the artificial intelligence (AI) algorithms have been combined to classic controllers to obtain different schemes of the adaptive self-tuning controllers. Particularly the fuzzy-PID [7, 8] and the NN-PID [9-12] are the most used hybrid control combinations. In fact, using learning paradigms to capture the system behaviour, in order to estimate the optimal controller's gains, is widely used. Traditional AI techniques learn from labelled or unlabelled data examples. While, supervised learning maps between collected inputs and outputs with the objective of producing the correct output for data that does not exist in the training examples, the unsupervised learning deals with unlabelled data examples to find an inherent structure or hidden patterns from inputs only. The Reinforcement Learning (RL), however, which is the third class of machine learning, learns from an environment how to generate actions to achieve a task in an optimal way, relying on the main objective of maximizing rewards. Unlike conventional controllers, RL is model-free algorithm that does not require an accurate model of the system dynamics, which is a crucial advantage when the system is complex [13]. Also, the agent is learning by interacting directly with the environment, not from data as it is for the classical AI controllers. This main advantage makes the RL suitable in robotic [14-16] and autonomous system [17-19]. Despite of some feature similarities in terms of nonlinear complexity, application of modern computational AI techniques to control shaking tables is still lacking. Yet, few recent works [20-22] demonstrate the high performance achieved by the proposed intelligent controllers which appears promising in that regard. The continuing learning process as a key feature of RL algorithm to capture the nonlinear behaviour of complex systems with various disturbances and changing conditions is motivating at first stage to explore, numerically, the potential of a Q-learning algorithm in controlling a shaking table. Since the main challenge in shaking table control is to replicate faithfully the prescribed earthquake records, the task of the RL agent in this case is to provide an optimal discretized control signal (optimal actions), which minimizes the acceleration errors between the reproduced accelerations and the desired ones. An action get more reward when the error remains within the limits of the acceleration error tolerance. In the present paper, a numerical framework for control shaking table using a RL approach is presented in which the RL agent will learn through a simulated environment, represented by a modelled shaking table system. The RL agent receives the state signal through measured acceleration responses. A Q-learning algorithm is designed as a model-free RL control algorithm applied to the shaking table, performing offline and providing states-actions pairs in a form of a Q-matrix.

The overall performance of the Q-learning controller is evaluated using the standard Root-Mean-Square Errors (RMSE) as a performance index to assess the quality of the matching between the measured accelerations and targets. A comparative study between the PDFF controller and the proposed controller is carried out to demonstrate the advantages of the Q-learning algorithm over the traditional controller.

2. Shaking table numerical model

A numerical model of a shaking table is developed to simulate the behaviour of the environment in which the RL agent will interact. The advantage of applying a RL methodology is that the agent learns to perform a sequence of actions in an uncertain and complex environment, without the need of an accurate system model [23]. In this study, the environment is represented by a model of an electric linear motor bi-axial shaking table, composed of two moving stages: a bottom stage that moves along the horizontal x-axis, and a top stage that moves along the transversal y-axis, as depicted in Fig.1. A simplified diagram of the shaking table system is illustrated in Fig.2.

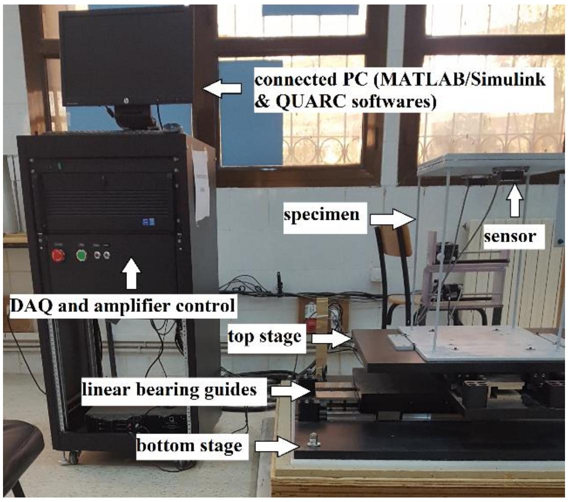

Fig. 1 – Experimental setup of the electric linear motor bi-axial shaking table

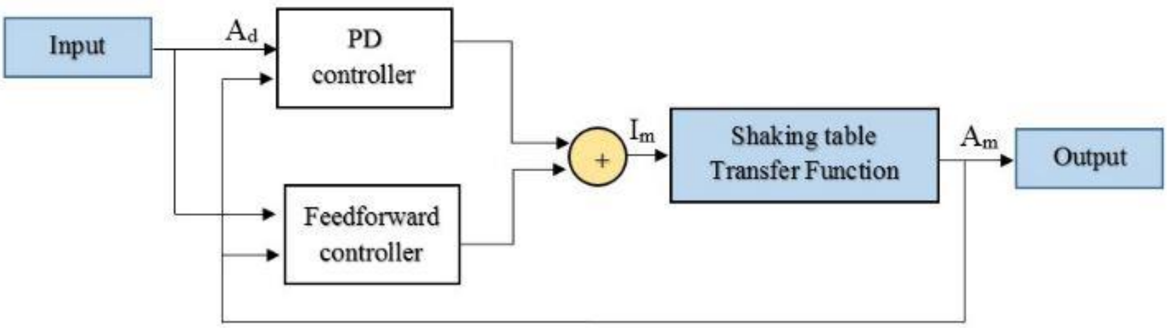

Fig. 2 – Simplified diagram of the shaking table and the original PDFF controller

The original control system is a classic position PDFF controller represented in Fig.3, where $A_d(s)$ is the desired acceleration, $X_d(s)$ is the desired position, $X(s)$ is the measured position, b_{sd} is the set-point velocity weight, $I_{ff}(s)$ and $I_{pd}(s)$ are the command signals provided by the FF and the PD controllers, respectively.

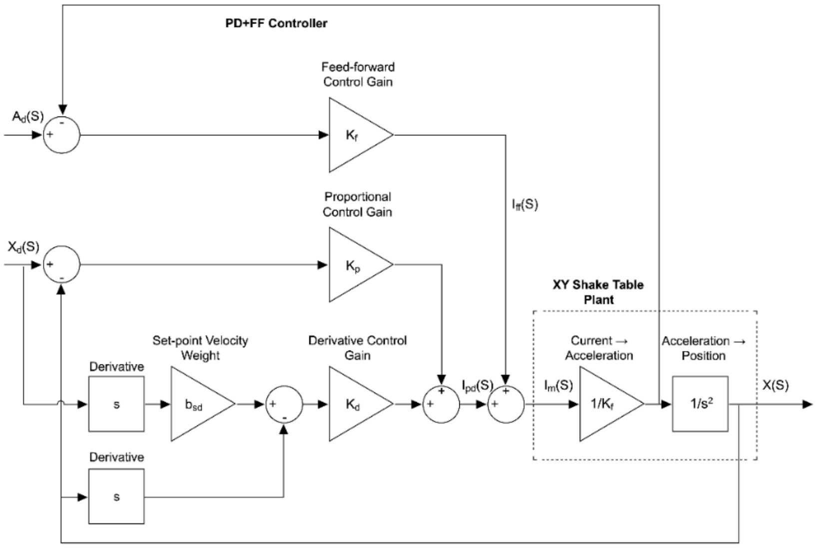

Fig. 3 – Schematic of the shaking table PDFF position controller

The x-axis model of the shaking table can be represented by the transfer function expressed by the following equation:

$$X(s) = (1/K_{f,x} * s^2)\ I_{m,x}(s) \tag{1}$$

Where $X(s) = \mathcal{L}[x(t)]$ is the Laplace transform of the stage position, $I_{m,x}(s)$ is the Laplace transform of the applied motor current provided by the PDFF shaking table controller and $K_{f,x}$ is the feedforward control gain along the x-axis.

The original motor current $I_m(s)$ is expressed as follows:

$$I_m(t) = K_p\,(x_d(t) - x(t)) + K_d\,(\dot{x}_d(t) - \dot{x}(t)) + K_f\,(\ddot{x}_d(t) - \ddot{x}(t)) \tag{2}$$

The RL agent is aimed to replace the PDFF controller by providing the command signal directly to the shaking table system. The later will be implemented in the numerical model for comparison purposes only.

3. RL control methodology

The basic principle of RL is illustrated in Fig.4. Real earthquake records are simulated through the numerical model of the shaking table, to provide a quasi-continuous interaction loop between the RL agent (controller) and the environment. In the figure below, $A_{d,t}$ represents the desired acceleration, $A_{m,t}$ represents the measured signal that constitute the state signal, R_t is the immediate reward.

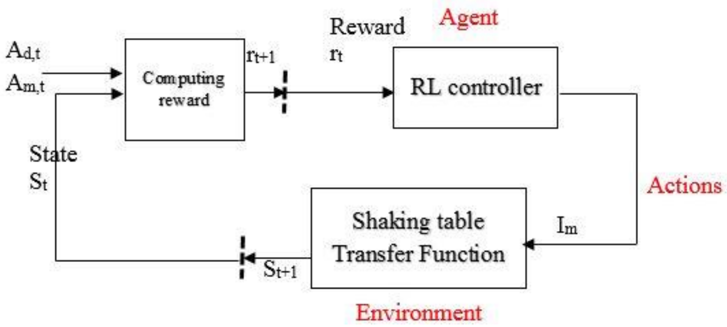

Fig. 4 – RL agent-environment interaction framework

At each time step, the RL agent receives a feedback from the shaking table (environment) in form of an acceleration time history (state signal) and generates an action (command signal) chosen from a discrete actions space, constituted by all the possible taken actions. Immediately, the agent receives a single reward R_t. The goal of the agent is to maximize the total reward it receives for a long-term. The long-term quality of the chosen action is evaluated through an action-value function, which is the function to maximize. Value estimation has been the most important activity in most RL methods [24].

Basically, the Q-function is obtained by experience, using an iterative process called Bellman equation, expressed in Eq. (3) and Eq. (4):

$$Q\ (s, a) = E\ [(R_t \mid S_t = S,\ a_t = a,\ \pi\ )] \tag{3}$$

$$Q\ (s_t, a_t) \leftarrow R_t + \alpha\ (r + \gamma \max Q\ (s_{t+1}, a_t) \tag{4}$$

The optimal Q-function Q*(s, a) is expressed as follows:

$$Q^*\ (s, a) = \max_\pi Q(s, a) \tag{5}$$

Where γ is the discount factor, $0 \leq \gamma \leq 1$. The discount factor takes into account the future rewards.

4. Q-learning control algorithm

The designed RL agent is aimed to replace the classic PDFF controller by providing the optimal command signal $I_m(s)$ sequentially to the shaking table based on the measured accelerations. The main objective of the intelligent RL controller is to minimize the error between the desired acceleration and the measured ones. An absolute error between the reproduced and the desired accelerations, less than 0.05g is judged to be reasonable [25], then the action is evaluated as good and the reward will be positive. Otherwise, the agent receives a negative reward.

The expression of the immediate reward in the following equation:

$$r = \begin{cases} +1, & if\ e_a \leq 0.05\ g \\ -1, & otherwise \end{cases} \tag{6}$$

Thus, the discretized state-signal of the environment at each time step, can be expressed as:

$$S_t = \{a_{m,t}\} \tag{7}$$

The action-signal can be defined easily since the only parameter that can actively be set by the RL agent is the command current applied to the shaking table. For a given earthquake record, the current signal values are bounded between -10 and +10A. The discretization step was set to 0.01 based on a compromise between the accuracy in making decision and the computation time.

$$a_t = [-10A:0.01:10A] \tag{8}$$

The RL agent learns the best behaviour, based on an immediate reward (short term reward) that he receives, in order to maximize the long term reward. Given the nature of the task and the definition of the reward-signal, the highest possible reward equals +1, corresponds to a quasi-perfect matching between desired and output accelerations.

Through this work, the implementation of a Q-learning algorithm has been chosen to be an interesting solution of the RL problem proposed in this study. The information about the interactions between the agent and the environment is stored in a so-called Q-Table. Having provided the definition of the state and action-signals, the structure of the Q-Table is that of a 2-dimensional matrix given in Eq. (9).

$$\text{Q-table} = [\text{states} \ ; \text{actions}] \tag{9}$$

The objective of the agent is to select the optimum action "a" having the highest value of Q(s, a). Initially, the Q(s, a) function values are initialized to zeros. After each iteration, the values are updated using the Bellman equation given in Eq. (4). The iterative process can be summarized as follows:

- Initialize (s, a) arbitrarily
- Repeat (for each episode) Initialize s
- Repeat (for each step of episode):
    - Take action a randomly, observe r and s'
    - Choose a' from s' with policy derived from Q (e.g. ε-greedy)

    $(s, a) \leftarrow Q(s, a) + \alpha \, [r + \gamma \, \max s' \, Q(s', a') - Q(s, a)]$

    $s \leftarrow s' \, ; a \leftarrow a'$

    Until s is terminal

5. Performance of the Q-learning control algorithm

This work focuses on improving shaking table control performance in term of reduction in the signal distortion affecting the quality of the reproduced signal at the base of the specimen tested. In order to verify the efficiency of the proposed control algorithm, numerical simulations of the shaking table was performed using several earthquake records such as El-Centro, Northridge and Cape-Mendocino. The discount factor, the learning rate and ε are equal to 0.9, 0.7 and 0.9, respectively [14]. Fig.5 shows the cumulative rewards vs. episodes. It is visible that the controller reached a maximum cumulative rewards within almost 60 episodes. After that, the value of the cumulative reward is increasing which means that the controller has achieved a stage of learning to choose the best possible action that produce the less error between the desired and the measured accelerations.

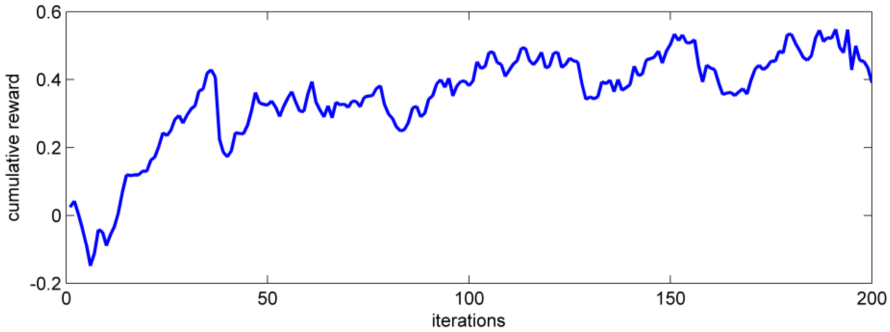

Fig. 5 – Reward graph showing the performance of the RL agent

After applying the optimal command signal to the shaking table transfer function, the acceleration response is illustrated in Fig.6 revealing the enhanced tracking performance of the system. A comparative analysis between the two feedback signals obtained by the Q-learning controller and the traditional PDFF controllers is carried out. To evaluate quantitatively the improvement in the signal reproduction provided by the Q-learning, the Root-Mean-Square Error (RMSE) and the Mean Square Error (MSE) are used as indices assessment and their values are presented in Table 1.

Table 1 – RMS values for the PDFF and the Q-learning controllers

Earthquake record	**RMS (g)**		**MSE (g)**	
	PDFF controller	**Q-learning controller**	**PDFF controller**	**Q-learning controller**
El Centro	0.0237	0.0081	$5.6\ 10^{-4}$	$6.6\ 10^{-5}$
Cape-Mendocino	0.0374	0.0402	0.0014	0.0016
Northridge	0.0297	0.0257	$8.82\ 10^{-4}$	$6.73\ 10^{-5}$

Despite the fact that the response accelerations are closer to the target with the Q-learning, this approach has the advantage of selecting appropriate actions through the direct interaction with the environment, based on the long term maximized reward, which makes the RL agent able to obtain similar results with any predefined signal to be reproduced.

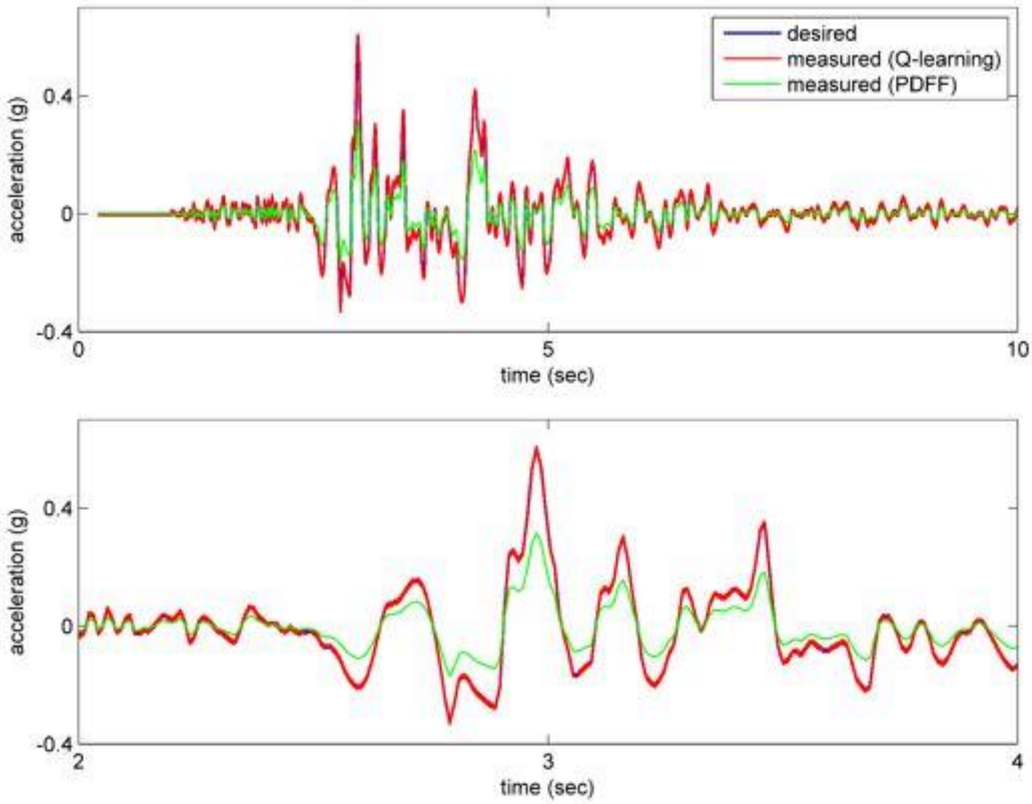

Fig. 6 – Numerical replication of Northridge earthquake record with the Q-learning and PDFF controller

6. Conclusion

In this study, a numerical framework was proposed to investigate the potential of a novel control technique based on RL applied to shaking tables. It is shown that, from direct interaction with the simulated environment, the RL controller can learn to choose the optimal command values, based on the reward signal that it receives for each chosen action. The Q-learning algorithm was performed offline and the optimal actions obtained have been applied to the system producing a quasi-perfect replication of the desired signal. The superior performance achieved by the proposed RL control method over the traditional PDFF controller was evaluated through a direct comparison of the time-history replication of several earthquake records and two performance assessment indices: the Root-Mean-Square-Error (RMSE) and Mean-Square-Error (MSE). However, further issues are being addressed to demonstrate the robustness of the controller by introducing different nonlinear specimen and developing a natural extension from the obtained Q-matrix to a Deep Q-Network (DQN) controller. The key feature of the RL controller is its continuing ability to (learn) improve with continuing use of the shaking table (environment). This will ultimately eliminate the signal matching phase at the beginning of shaking table tests.

7. References

[1] Shen G, Lv GM, Ye ZM, Cong DC, Han JW (2011): Implementation of electrohydraulic shaking table controllers with a combined adaptive inverse control and minimal control synthesis algorithm. *Earthquake IET Control theory and applications*, **5** (13), 1471-1483.

[2] Chen PC, Chang CM, Spencer Jr BF, Tsai K-C (2014): Adaptive model-based tracking control for real-time hybrid simulation. *Bulletin of Earthquake Engineering*, **13** (6), 1633–1653.

[3] Yachun T, Peng P, Dongbin Z, Yi Z (2018): A two-loop control method for shaking table tests combining model reference adaptive control and TVC. *Frontiers in Built Environment*, **4** (54), 1–12.

[4] Stoten DP, Gomez EG (2001): Adaptive control of shaking tables using the minimal control synthesis algorithm. *Philosophical Transactions: Mathematical, Physical and Engineering Science*, **359** (1786), 1697–1723.

[5] Gizatullin AO, Edge KA (2007): Adaptive control for a multi-axis hydraulic test rig. *Proceedings of the Institution of Mechanical Engineers, Part I: Journal of Systems and Control Engineering*, **221** (2), 183-198.

[6] Enokida R, Kajiwara K (2019): Nonlinear signal-based control for single axis shaking table supporting nonlinear structure systems. *Structural control and Health Monitoring*, **26** (9), 1-20.

[7] Jigang H (2019): An anti-windup self-tuning fuzzy PID controller for speed control of brushless DC motor. *Journal of Control, measurements, electronics, computing and communication*, **60** (2), 135-148.

[8] Chao CT, Sutarna N, Chiou JS, Wang CJ (2019): An optimal Fuzzy PID controller design based on conventional PID control and nonlinear factors. *Applied Sciences*, **9** (6), 1224, 1-18.

[9] Muderrisoglu K, Arisoy DO, Ahan AO, Akdogan E (2016): PID parameters prediction using NN for a linear quarter car suspension. *International Journal of intelligent systems and applications in engineering*, **4** (1), 20-24.

[10] Fu T, Wang X (2016): Design of brushless DC motor controller based on adaptive RBF NN. *International Journal of Control and automation*, **9** (1), 459-470.

[11] Hernández-Alvarado R, García-Valdovinos LG, Salgado-Jiménez T, Gómez-Espinosa A, Fonseca-Navarro F (2016): NN-based self-tuning PID control for underwater vehicles. *Sensors (Basel)*, **16** (9), 1429, 1-18.

[12] Rosales C, Soria CM, Rossomando FG (2018): Identification and adaptive PID control for hexacopter 4AV based on NN. *International journal of Adaptive control and signal processing*, **33** (1), 74-91.

[13] Han M, Zhang X, Xu L, May R, Pan S, Wu J (2018): A review of reinforcement learning methodologies on control systems for building energy. *Working papers in transport, tourism, information technology and microdata analysis*, **2**, 1-26.

[14] Rahman MM, Rashid SMH, Hossain MM (2018): Implementation of Q learning and deep Q network for controlling a self-balancing robot model. *Journal of Robotics and Biomimics*, **5** (8), 1-6.

[15] Tai L, Liu M (2016): Mobile robots exploration through CNN-based reinforcement learning. *Journal of Robotics and Biomimics*, **3** (24), 1-8.

[16] Zamora I, Lopez NG, Vilches VM, Cordero AH (2016): Extending the OpenAI Gym for robotics: a toolkit for reinforcement learning using ROS and Gazebo. *ArXiv*, abs/1608.05742.

[17] Tran L, Cross C, Montague G, Motter M, Neilan J, Quallsk G, Rothhaar P, Trujillo A, Allen B.D (2015): Reinforcement Learning with Autonomous Small Unmanned Aerial Vehicles in Cluttered Environments. *In: 15th AIAA aviation technology, integration, and operations conference*.

[18] Kim HJ, Jordan MI, Sastry S, Ng AY (2004): Autonomous helicopter flight via reinforcement learning. *Advances in neural information processing systems*, 16 (NIPS2003).

[19] Jamone L, Natale L, Nori F, Metta G, Sandini G (2012): Autonomous online learning of reaching behavior in a humanoid robot. *International Journal of Humanoid Robotics*, **9** (3), 1250017, 1-26.

[20] Tang Y, Zhu Z, Shen G, Zhang W (2017): Real Time Acceleration Tracking of Electro-Hydraulic Shake Tables Combining Inverse Compensation Technique and Neural-Based Adaptive Controller. *IEEE Acces*, **5**, 23681 - 23694.

[21] Larbi SH, Bourahla N, Benchoubane H, Choutri K, Badaoui M (2016): Developing a Neural Network algorithm as an additional online controller to the PID controller. *Proceedings of the 16WCEE Santiago, Chile.*

[22] Soleymani M, Soltani AK, Ghanbari SB (2019): Fuzzy-sliding-mode supervisory control of a seismic shake table with variable payload for robust and precise acceleration tracking. *Journal of earthquake Engineering*, **23** (4), 539-559.

[23] Bae H, Kim G, Kim J, Qian D, Lee S (2019): Multi-Robot Path Planning Method Using Reinforcement Learning. *Applied Sciences*, **9** (15), 3057, 1-16.

[24] Marinescu A (2016): Prediction-based multi-agent Reinforcement Learning for inherently non-stationary environments. *Doctor of Philosophy dissertation in computer science, University of Dublin.*

[25] Guo W, Shao P, Li HY, Long Y, Mao JF (2019): Accuracy Assessment of Shake Table Device on Strong Earthquake Output. *Advances in Civil Engineering*, **2019**, Article ID 9372505, 1-23.

