

Article

Optimal Reinforcement Learning-Based Control Algorithm for a Class of Nonlinear Macroeconomic Systems

Qing Ding ^{1,2}, Hadi Jahanshahi ³, Ye Wang ^{4,5,*} , Stelios Bekiros ^{6,7,*}  and Madini O. Alassafi ⁸ 

- ¹ College of Liberal Arts and Sciences, National University of Defense Technology, Changsha 410073, China; dingqing@hufe.edu.cn
- ² College of Mathematics and Statistics, Hunan University of Finance and Economics, Changsha 410205, China
- ³ Department of Mechanical Engineering, University of Manitoba, Winnipeg, MB R3T 5V6, Canada; jahanshahi.hadi90@gmail.com
- ⁴ Department of Mathematics, Huzhou University, Huzhou 313000, China
- ⁵ Institute for Advanced Study Honoring Chen Jian Gong, Hangzhou Normal University, Hangzhou 311121, China
- ⁶ Department of Banking and Finance, FEMA, University of Malta, MSD 2080 Msida, Malta
- ⁷ Department of Economics, European University Institute, I-50014 Florence, Italy
- ⁸ Department of Information Technology, Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah 21589, Saudi Arabia; malasafi@kau.edu.sa
- * Correspondence: 03019@zjhu.edu.cn (Y.W.); stelios.bekiros@eui.eu (S.B.)

Abstract: Due to the vital role of financial systems in today's sophisticated world, applying intelligent controllers through management strategies is of crucial importance. We propose to formulate the control problem of the macroeconomic system as an optimization problem and find optimal actions using a reinforcement learning algorithm. Using the Q-learning algorithm, the best optimal action for the system is obtained, and the behavior of the system is controlled. We illustrate that it is possible to control the nonlinear dynamics of the macroeconomic systems using restricted actuation. The highly effective performance of the proposed controller for uncertain systems is demonstrated. The simulation results evidently confirm that the proposed controller satisfies the expected performance. In addition, the numerical simulations clearly confirm that even when we confined the control actions, the proposed controller effectively finds optimal actions for the nonlinear macroeconomic system.

Keywords: macroeconomic system; reinforcement learning; intelligent control; optimal controller



Citation: Ding, Q.; Jahanshahi, H.; Wang, Y.; Bekiros, S.; Alassafi, M.O. Optimal Reinforcement Learning-Based Control Algorithm for a Class of Nonlinear Macroeconomic Systems. *Mathematics* **2022**, *10*, 499. <https://doi.org/10.3390/math10030499>

Academic Editors: Ioan Batrancea, Ramona-Anca Nichita, Lucian Gaban and Mircea-Iosif Rus

Received: 27 December 2021

Accepted: 1 February 2022

Published: 3 February 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Nowadays, the control and synchronization of nonlinear systems have attracted attention due to their appreciable usage in a wide variety of fields [1–7]. Up to now, intelligent control systems have effectively solved some problems in the control of modern tools and methods, such as self-organizing maps, neural networks, fuzzy logic, expert systems and various nature-inspired algorithms [8,9]. Several research studies have applied neural networks as disturbance estimators to observe the disturbances and uncertainties in nonlinear systems [10–12]. On the other hand, the robustness of the control strategies plays an inevitable role in the performance of the systems. Hence, researchers also propound robust controllers for complex systems [13–15].

A macroeconomic model describes the dynamics of an economic system over long- and short-run periods. The macroeconomic model can consist of difference or differential equations. Usually, the variables in the macroeconomic model are categorized into the following three main classes: (1) variables that represent the state of the system, (2) variables that indicate an uncontrollable environment and (3) control variables that denote leverage, which can be employed as economic policies to push the system towards a target [16]. It is noteworthy that the existence of chaos in the macroeconomic time series has not been

conclusively confirmed so far. This issue belongs to the unsolved scientific problems. In [17], Barnett and He have thoroughly explained this matter in their study.

Though studies have made impressive progress in economic systems, the investigation of macroeconomic systems and control methods for these systems have room for improvement [18,19]. For instance, in the literature, a few studies exist that have investigated the control of macroeconomic systems [20]. In this study, a fuzzy controller has been designed for the macroeconomic system. Nevertheless, this control scheme has some disadvantages. For instance, the fuzzy control may not give assurance that the system reaches the desired value in the presence of uncertainties. Actually, controllers that have been proposed for macroeconomic systems are neither intelligent nor robust enough for many uncertain situations. Due to the system uncertainty, complexity, and nonlinearity in economic systems, mathematical models may not be able to consider all the variations in the system [21]. Therefore, further studies on robust and intelligent control techniques are required to achieve appropriate performance for the stabilization and control of macroeconomic systems.

These issues motivated this study. In the current paper, a macroeconomic system will be considered in the presence of dynamic uncertainties, and we aim to enhance the performance of the system by taking advantage of artificial intelligence techniques. Since applying optimal policies for economic systems is crucial, we propose to formulate the control of the macroeconomic system as an optimization problem and find optimal actions based on a reinforcement learning algorithm. The significant advantage of the proposed reinforcement learning-based method is its optimal performance. In addition, using the proposed method, we can control the system's behavior with only restricted and specific actions. In the current study, we use Goodwin's non-linear model and generate data from it. Actually, for real-world application, we need a model to train the Q-learning algorithm and extract the control laws. After that, the proposed controller can be applied to the real-world economic system. This is a theoretical study that aims to show the application of the Q-learning algorithm in economic systems. Hence, we use a well-known mathematical model in this study.

2. Literature Review

Up to now, various control techniques, such as predictive control [22–24], sliding mode control [25–32], fuzzy control [33–35], adaptive control [36–39], and so on, have been proposed for nonlinear systems. Among the stated control methods for various systems, the methods that can act intelligently in unpredictable conditions have gained significant attention because of their unique properties [40,41]. Artificial intelligence does combine a wide variety of state-of-the-art technologies to give systems an ability to adaptively make decisions in new and unknown situations [42,43]. In some applications, due to the high value of their tasks, and their considerable risks, implementing a robust and reliable control scheme is a significant concern. In the presence of high disturbances and uncertainties, classical control methods may fail [44]. Moreover, today, technologies have reached a point where machine intelligence can be utilized for many systems. Intelligent controllers can make systems smarter and able to find more effective ways, hopefully for better actions in real-world applications [45,46].

Reinforcement learning is one of the earliest learning techniques that emerged in the field of machine learning, and soon became a popular method [47,48]. In [49,50], the authors have used reinforcement learning methods to explore the system's response to possible actions, and in this way, they found optimal actions by calculating how the previous actions move the system to a favorable state. After the learning process, the controller uses the optimal learned policies. It is shown in [51,52] that reinforcement learning can be utilized for the control of complex and unknown systems, as it does not require a precise mathematical model of systems. Hence, as is stated in [53–55], reinforcement learning-based control methods have been used in various fields of study, including the speed control of turbines, drug dosing, image evaluation and robotics.

3. Methodology

In the current study, we propose a reinforcement learning-based method for optimal control of the macroeconomic system. Traditionally, designing optimal control for nonlinear systems is challenging [56]. When the model of a system is known, using the algebraic Riccati equation, we can design the optimal control law for linear systems. However, when designing an optimal control law for nonlinear systems, the Hamilton–Jacobi–Bellman equation must be solved [57].

Watkin’s Q-learning is a reinforcement learning-based method that has attracted significant attention in recent years. This method does not need a priori information about systems and can effectively be employed online when the model of the system changes during the learning process [58]. Figure 1 illustrates the structure of reinforcement learning. In this process, agents or control signals apply an action on the system and observe the corresponding reward to find effective control policies or action plans.

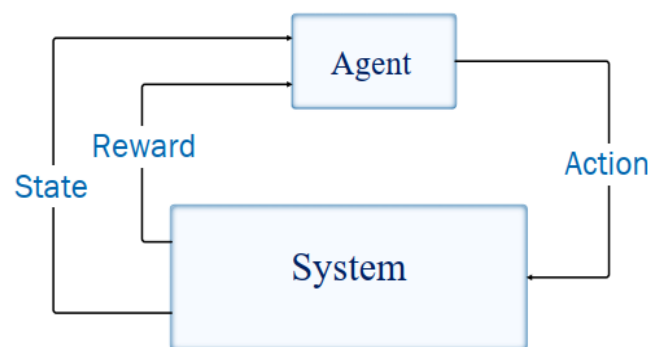


Figure 1. Structure of reinforcement learning.

Reinforcement learning starts by choosing an initial arbitrary policy and then, by interacting with the system, it learns the optimal policy. In general, a policy can be a rule base, such as “if in this state, then do this” [59]. Reinforcement learning progresses iteratively by interacting with the system. Hence, the agent’s decision sets an optimal control policy based on the rewards it obtains [60,61]. In the Q-learning algorithm, each batch of information, including action, reward and state, is utilized to update the Q table. In the Q table, the entry $Q_k(s_k, a_k)$ denotes the desirability of actions in the finite sequence $(A_j)_{j \in J^+}$ with respect to states of the finite sequence $(S_i)_{i \in I^+}$. As is demonstrated in Figure 1, the central part of the reinforcement learning includes a system and an agent [62].

At time step k , firstly, the agent observes the current state s_k ; after that it selects action a_k from the sequence of actions (A) . The results of the selected action a_k are scored based on an appropriate reward $(r_{k+1} \in R)$. Based on the value of the reward, the agent realizes whether the last action that had been selected was “bad” or “good.” The agent uses the Q-learning algorithm to obtain an optimal policy that maximizes the expected value $\mathbb{E}[\cdot]$ of the discounted reward, which is given by:

$$J(r_k) = \mathbb{E} \left[\sum_{k=1}^{\infty} \theta^{k-1} r_k \right] \quad (1)$$

In Equation (1), the importance of future and immediate rewards is represented by $\theta \in [0, 1]$, where for $\theta = 0$, the agent only considers the current reward, and for θ approaching 1, the agent takes into account current and future rewards. In this regard, when the agent calculates action a_k and reward r_{k+1} , with respect to the state transition $s_k \rightarrow s_{k+1}$, the Q table will be updated based on the Q-learning algorithm, which is given by the following equation:

$$Q_k(s_k, a_k) = Q_{k-1}(s_k, a_k) + \eta_k(s_k, a_k) \times \left[r_{k+1} + \theta \max_{a_{k+1}} Q_{k-1}(s_{k+1}, a_{k+1}) - Q_{k-1}(s_k, a_k) \right] \quad (2)$$

where the learning rate that impacts the size of the correction after each iteration is represented by $\eta_k(s_k, a_k) \in [0, 1)$, $k = 1, 2, \dots$. It is noteworthy that the Q-learning algorithm begins with an initial $Q_1(s_1, a_1)$. Then, the Q table will be updated based on the observations.

To assign the minimum threshold for convergence, it is common to use a tolerance parameter δ with condition $Q_k |Q_k - Q_{k-1}| \leq \delta$. Several sources, such as references, can be consulted for more details on the conditions required and the proof of convergence of the Q-learning algorithm [63,64]. The algorithm used in this paper has been delineated in Algorithm 1, in which episode is defined as the process of reaching the final state from the initial state.

Algorithm 1. Q-learning algorithm.

- 1: Initialize Q-table.
 - 2: **Loop** {for all of episodes}.
 - 3: Initialize state s .
 - 4: **Repeat** {for each step-in episode}.
 - 5: Calculate firing strength of A at state s_i .
 - 6: Choose action (a_i) for each of the rules at state s_i .
 - 7: Take action a_{i+1} , observe reward r_i .
 - 8: Calculate state value of state s_{i+1} .
 - 9: Update Q-table.
 - 10: **Until** s is terminal state.
 - 11: **End loop**.
-

Figure 2 demonstrates the structure of the training sequence that has been used to obtain the optimal Q table for the macroeconomic system. In the current study, we consider the state s_k of the macroeconomic model in terms of the available output of tracking error. In the macroeconomic system, the target is to obtain the optimal sequence of actions, which results in a minimum value of error.

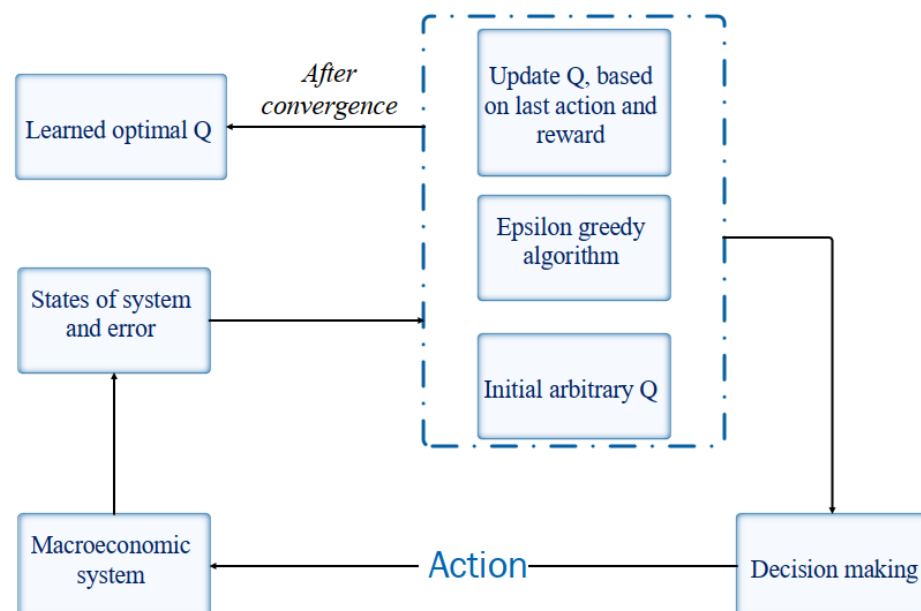


Figure 2. Structure of training sequence to reach optimal Q table.

As was previously mentioned, the reward function is defined to guide the agent in recognizing whether the last action was appropriate or not. Actually, this information is

utilized to reinforce the decision making of the agent. At every time step, the controller will choose the action a_k as follows:

$$a_k = (A_j)_{j \in \mathbb{J}^+} \quad j = \operatorname{argmax}(Q_k(s_k, \cdot)) \quad (3)$$

The function that is utilized for calculating the reward of the agent for the transition from state s_k to state s_{k+1} is as follows:

$$r_{k+1} = \begin{cases} \left| \frac{e(kT) - e((k+1)T)}{e(kT)} \right|, & |e((k+1)T)| < |e(kT)| \\ -\xi & |e((k+1)T)| > |e(kT)| \end{cases} \quad (4)$$

where $\xi \geq 0$ and $e(t)$ represent the error of the system. The more the agent explores the system, the more it learns. When $k \rightarrow \infty$, the algorithm is able to converge to the optimal Q table. In addition, in most cases, for a finite value of k , systems converge to their optimal solution with an acceptable tolerance δ [51].

4. Macroeconomic System

Puu [65] proposed a modification of the Goodwin model by allowing the accelerator function to be non-monotonic. He has supported the idea that when the income sharply drops, the investment might rise. Moreover, when the income is rising, especially sharply, the investment might fall. This model is described as follows [65]:

$$Z(t) = C(t) + I(t) + G(t), \quad (5)$$

$$C(t) = cY(t) - u(t), \quad (6)$$

$$\frac{dI(t)}{dt} = -\beta(I(t) - B(t)), \quad (7)$$

$$\frac{dY(t)}{dt} = -\alpha(Y(t) - Z(t)). \quad (8)$$

where the aggregate demand (Z) consists of consumption (C), autonomous expenditures of the government (G) and investment (I). Consumption (C) depends on income (Y) and is disturbed by a spontaneous change in the u coefficient, in which u is defined by a step function. In addition, parameter c denotes the marginal propensity to consume. Equation (7) represents the rate of investment, where β is the speed of the response of the investment to changes in production, and the speed of the response of production to changes in demand is represented by α . In addition, $B(t) = \Phi\left(v \frac{dY(t)}{dt}\right)$. In this nonlinear macroeconomic model, the rate of investment ($\frac{dI(t)}{dt}$) depends on investment I and the amount of the investment decision B . Moreover, as is described by this model, the decision to invest B depends nonlinearly, by Φ , on the rate of change in production ($\frac{dY(t)}{dt}$). The nonlinear accelerator Φ is given by:

$$\Phi\left(v \frac{dY(t)}{dt}\right) = M \left(\frac{L + M}{Le^{v \frac{dY(t)}{dt}} + M} - 1 \right) \quad (9)$$

where M indicates the scrapping rate of capital equipment and L is the net capacity of the capital goods trades. Substituting Equation (9) into Equation (7) results in the final shape of the governing equation of the Goodwin's model, which is given by:

$$\frac{dI(t)}{dt} = -\beta \left(I(t) - M \left(\frac{L + M}{Le^{v \frac{dY(t)}{dt}} + M} - 1 \right) \right), \quad (10)$$

5. Results

5.1. Dynamical investigation

The time history of the macroeconomic system is shown in Figure 3. It illustrates the time history of the system with different initial conditions. It is noteworthy that since we did not examine the absolute values of the system's variables, and only their rates of change have been examined, they can be negative and positive. For this simulation, $c = 0.25$, $\beta = 3$, $v = 3$, $\alpha = 4$, $M = -1$ and $L = 2$. Additionally, for case 1 and case 2, the initial conditions are considered to be $[-2, 0, 1, -1]$ and $[-8, 1, 1\frac{1}{4}, 6\frac{3}{4}]$, respectively. As is depicted in Figure 3, the behavior of the system depends on the initial condition [66].

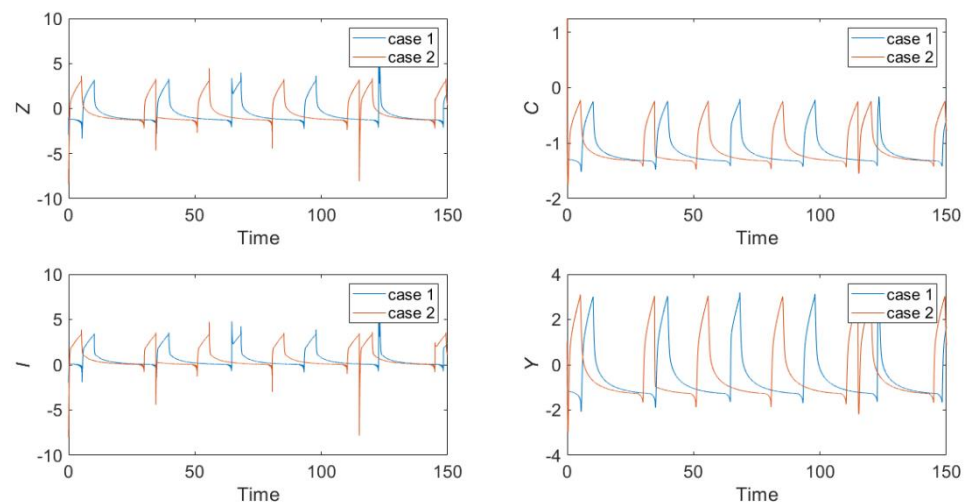


Figure 3. The time history of the system without controller (in case 1 and case 2, the initial conditions are considered to be $[-2, 0, 1, -1]$ and $[-8, 1, 11/4, 63/4]$, respectively).

5.2. Control Results

The proposed controller learns the optimal strategies based on the response of the system to different actions. For control of the system, the consumption of Equation (6) in the macroeconomic system, including the disturbance and control input, is given by:

$$C(t) = cY(t) - u(t) + U_c + D \quad (11)$$

where U_c is the control action and D represents disturbances that are imposed on the system. The time response of the macroeconomic system in the presence of uncertainties is simulated to test the effectiveness of the proposed control scheme. In the simulations, the number of iterations for the learning algorithm is set on 500,000 scenarios. In each scenario, the macroeconomic system is simulated (see Figure 2), and the scenarios indicate the series of transitions from arbitrary initial states to the terminal state s_k . In addition, $(A_j)_{j \in \mathbb{J}^+}$ denotes the action a_k , where $\mathbb{J}^+ = \{1, 2, \dots, 24\}$. Actually, we consider 24 possible actions for the control of the system, and the learning algorithm learns how and when to apply each control action. Table 1 lists the norms of regulation errors and control inputs.

Table 1. The norms of the regulation errors and control input.

State (s_k)	$e(t)$	State (s_k)	$e(t)$
1	$[-\infty, -15]$	13	$[-0.15, -0.1]$
2	$[-15, -9]$	14	$[0.1, 0.15]$
3	$[-9, -8]$	15	$[0.15, 0.2]$
4	$[-8, -6.5]$	16	$[0.2, 0.5]$
5	$[-6.5, -5]$	17	$[0.5, 1]$
6	$[-5, -4]$	18	$[1, 1.5]$
7	$[-4, -3]$	19	$[1.5, 3]$
8	$[-3, -1.5]$	20	$[3, 4]$
9	$[-1.5, -1]$	21	$[4, 5]$
10	$[-1, -0.5]$	22	$[5, 8]$
11	$[-0.5, -0.2]$	23	$[8, 15]$
12	$[-0.2, -0.15]$	24	$[15, \infty]$

These intervals are user-defined and could be changed. In the current study, we have classified these intervals into 24 levels, and the controller will try to find an appropriate control action for each of them. After convergence of the Q table, the agent will select an action $a_k = (A_j)_{j \in \mathbb{J}^+}$, where $j = \operatorname{argmax} Q_k(s_k, \cdot)$.

Furthermore, the learning rates that influence the value of the correction after each iteration are considered as follows:

$$\eta_k(s_k, a_k) = \zeta \frac{Ep_i}{Ep_f} \quad (12)$$

where $\zeta = 0.5$ and Ep_i denote the number of the current episode, and Ep_f is the maximum value of the episode that we want the simulation to run. To control the system, two cases have been considered. Namely, in the first simulation, we consider the maximum action (maximum value for the controller) to be 40, and the minimum is equal to -40 . In the second simulation, the maximum and minimum values of actions are considered to be 30 and -30 , respectively. Note that two different reinforcement learning agents have been trained for each of the aforementioned cases.

For both cases, the desired value of income is considered to be two ($Y_d = 2$), and the aim is to reduce the absolute value of error ($|e(t)| = |Y - Y_d|$) until it becomes equal to zero. The reward is calculated based on Equation (4), in which $\xi = -0.1$. The discount factor $\theta = 0.7$. In order to reduce the convergence time, we set the control policy as: when $|e(t)| < 0.1$ $U_c = K e(t)$ and parameter K is equal to ten, and other actions are considered as:

$$A = U_{max}[0, 0.001, 0.002, -0.002, 0.005, -0.005, 0.1, -0.1, 0.2, -0.2, 0.3, -0.3, 0.4, -0.4, 0.6, -0.6, 0.7, -0.7, -0.8, 0.8, -0.9, 0.9, -1, 1]. \quad (13)$$

At first, we consider $U_{max} = 40$. The stabilized states of the macrosystem are depicted in Figure 4. As is shown in this figure, the states of the system are stabilized in finite time using the proposed controller.

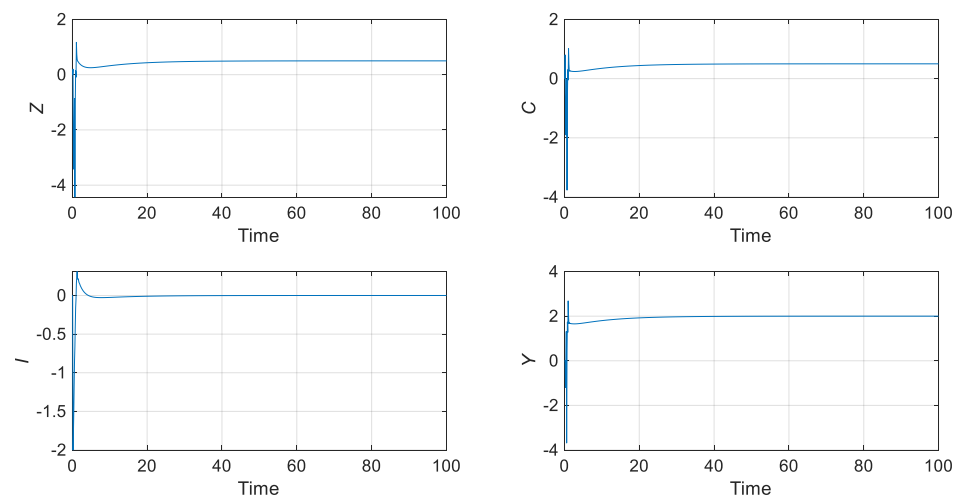


Figure 4. The time history of the macroeconomic system based on the proposed reinforcement learning controller when $U_{max} = 40$.

Additionally, Figure 5 depicts the control input based on the applied optimal reinforcement learning-based controller. Based on the numerical simulation, using the proposed controller, after a short period of time, the system is completely stabilized, and the states of the system reach their desired value.

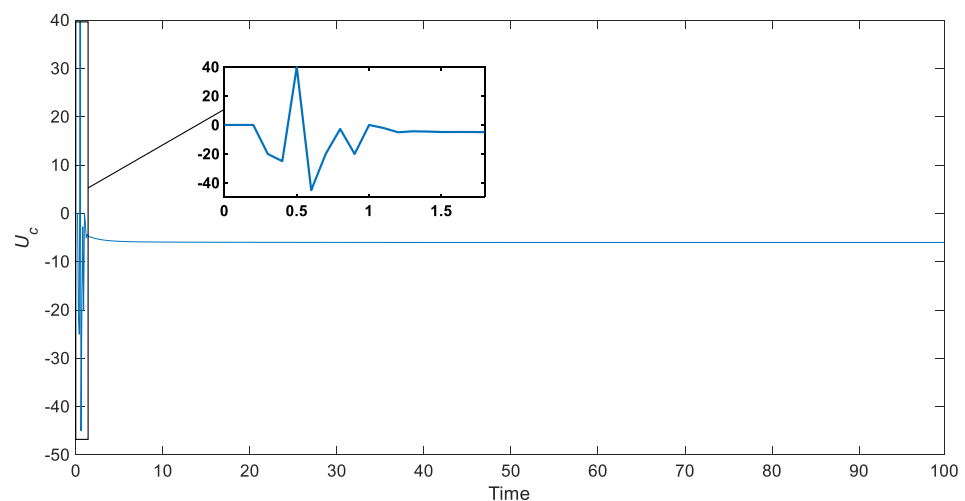


Figure 5. The control input of the closed-loop system based on the proposed reinforcement learning controller when $U_{max} = 40$.

Here, we consider $U_{max} = 30$. Actually, to investigate the proposed control in another condition, we confine the controller with actions that have fewer values. Figure 6 demonstrates the time history of the macroeconomic system by applying the designed control when $U_{max} = 30$. In addition, the control input is displayed in Figure 7. The simulation results evidently corroborate that the proposed controller satisfies the expected performances. Moreover, the numerical simulations clearly corroborate that the proposed controller effectively finds optimal actions for the control of the nonlinear macroeconomic system, even when we confined the control actions. It is noteworthy that the results of the examined model show that it is possible to control and stabilize the variables at constant values, e.g., Y (national income) = 2 and I (investments) = 0. This means simple reproduction, which is normal when investments are zero.

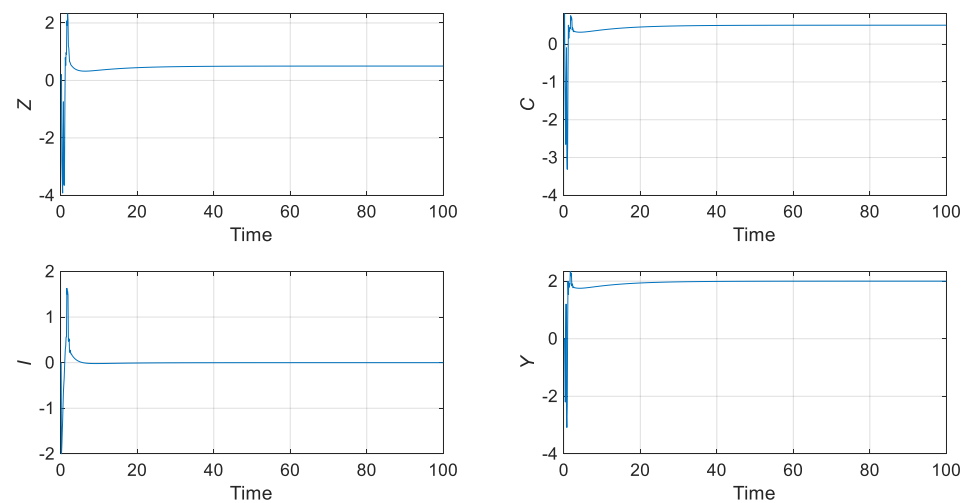


Figure 6. The time history of the macroeconomic system based on the proposed reinforcement learning controller when $U_{max} = 30$.

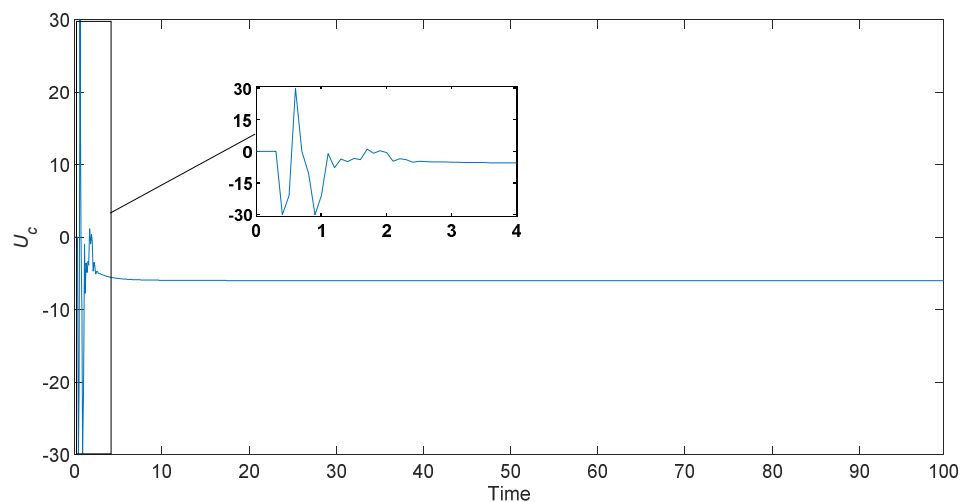


Figure 7. The control input of the closed-loop system based on the proposed reinforcement learning controller when $U_{max} = 30$.

In [20], a fuzzy controller has been presented for macroeconomic models. Although a fuzzy controller is a good candidate for economic systems, in that study, only some simple fuzzy rules are considered, and the controller is not robust against uncertainties and disturbances. In addition, that controller is only suitable for continuous systems, which is not the case in most studies on macroeconomic models.

Remark 1. The numerical results show that the proposed controller can be applied to nonlinear economic systems. Since the controller can be used for tracking control and stabilization, from an economic point of view, this method is useful to remove business cycles, even when these systems are not chaotic and have deterministic results.

Remark 2. The nonlinear Goodwin model and other models that are similar to it are the most plausible in science, but they have serious limitations that the researcher must be aware of. There are differences between the data generated by Goodwin's simple nonlinear model and actual macroeconomic data. The former is purely deterministic, while the latter is burdened with significant noise and may be complex. In the current study, we only use this model to investigate the developed control approach. Consequently, applying the developed method to a real-world economic system

will provide better insight into the control of the macroeconomic model, and it is the next step for this study. Hence, we will consider it as a future direction for our study.

6. Conclusions

The control and stabilization of macroeconomic systems using the reinforcement learning-based controller were studied. The proposed control scheme uses a reinforcement learning algorithm to find optimal control actions. Based on the Q-learning algorithm, the best optimal action for the system was obtained. The proposed method can control the system using restricted actuation, and, this way, it provides highly effective performance and optimal strategies using possible actions. Through numerical simulation, the adequate performance of the proposed reinforcement learning-based controllers was demonstrated by using different numerical simulations. The possibility of controlling the behavior of the macroeconomic system using the proposed reinforcement learning-based controller suggests the extension of the application of the reinforcement learning algorithms to the control of other nonlinear economic systems. Future studies will also be devoted to a complete comparison with other learning-based strategies, as well as alternative algorithms from the Q-learning method.

Author Contributions: Conceptualization, Q.D., H.J., Y.W., S.B. and M.O.A.; methodology, Q.D., H.J., Y.W., S.B. and M.O.A.; software, Q.D., H.J., Y.W., S.B. and M.O.A.; validation, Q.D., H.J., Y.W., S.B. and M.O.A.; formal analysis, Q.D., H.J., Y.W., S.B. and M.O.A.; investigation, Q.D., H.J., Y.W., S.B. and M.O.A.; writing—original draft preparation, Q.D., H.J., Y.W., S.B. and M.O.A.; writing—review and editing, Q.D., H.J., Y.W., S.B. and M.O.A.; supervision, Q.D., H.J., Y.W., S.B. and M.O.A. All authors have read and agreed to the published version of the manuscript.

Funding: The Deanship of Scientific Research (DSR) at King Abdulaziz University, Jeddah, Saudi Arabia has funded this project, under grant no. (FP-153-43). This work was supported by the Hunan Provincial Department of Education Scientific Research Outstanding Youth Project (Grant No.: 20B093), the Hunan Philosophy and Social Science Foundation Project (Grant No.: 20JD008) and the Natural Science Foundation of China (Grant No.: 71873045).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Lai, Q.; Wan, Z.; Kuate, P.D.K.; Fotsin, H. Coexisting attractors, circuit implementation and synchronization control of a new chaotic system evolved from the simplest memristor chaotic circuit. *Commun. Nonlinear Sci. Numer. Simul.* **2020**, *89*, 105341. [\[CrossRef\]](#)
2. Zhu, Z.-Y.; Zhao, Z.-S.; Zhang, J.; Wang, R.-K.; Li, Z. Adaptive fuzzy control design for synchronization of chaotic time-delay system. *Inf. Sci.* **2020**, *535*, 225–241. [\[CrossRef\]](#)
3. Jahanshahi, H.; Yousefpour, A.; Munoz-Pacheco, J.M.; Kacar, S.; Pham, V.-T.; Alsaadi, F.E. A new fractional-order hyperchaotic memristor oscillator: Dynamic analysis, robust adaptive synchronization, and its application to voice encryption. *Appl. Math. Comput.* **2020**, *383*, 125310. [\[CrossRef\]](#)
4. Han, Z.; Li, S.; Liu, H. Composite learning sliding mode synchronization of chaotic fractional-order neural networks. *J. Adv. Res.* **2020**, *25*, 87–96. [\[CrossRef\]](#) [\[PubMed\]](#)
5. Kosari, A.; Jahanshahi, H.; Razavi, S. An optimal fuzzy PID control approach for docking maneuver of two spacecraft: Orientational motion. *Eng. Sci. Technol. Int. J.* **2017**, *20*, 293–309. [\[CrossRef\]](#)
6. Wang, B.; Jahanshahi, H.; Volos, C.; Bekiros, S.; Khan, M.; Agarwal, P.; Aly, A. A New RBF Neural Network-Based Fault-Tolerant Active Control for Fractional Time-Delayed Systems. *Electronics* **2021**, *10*, 1501. [\[CrossRef\]](#)
7. Wang, H.; Jahanshahi, H.; Wang, M.-K.; Bekiros, S.; Liu, J.; Aly, A. A Caputo–Fabrizio Fractional-Order Model of HIV/AIDS with a Treatment Compartment: Sensitivity Analysis and Optimal Control Strategies. *Entropy* **2021**, *23*, 610. [\[CrossRef\]](#)
8. Jahanshahi, H.; Yousefpour, A.; Munoz-Pacheco, J.M.; Moroz, I.; Wei, Z.; Castillo, O. A new multi-stable fractional-order four-dimensional system with self-excited and hidden chaotic attractors: Dynamic analysis and adaptive synchronization using a novel fuzzy adaptive sliding mode control method. *Appl. Soft Comput.* **2020**, *87*, 105943. [\[CrossRef\]](#)

9. Chen, Y.-J.; Chou, H.-G.; Wang, W.-J.; Tsai, S.-H.; Tanaka, K.; Wang, H.O.; Wang, K.-C. A polynomial-fuzzy-model-based synchronization methodology for the multi-scroll Chen chaotic secure communication system. *Eng. Appl. Artif. Intell.* **2020**, *87*, 103251. [\[CrossRef\]](#)
10. Wang, B.; Liu, J.; Alassafi, M.O.; Alsaadi, F.E.; Jahanshahi, H.; Bekiros, S. Intelligent parameter identification and prediction of variable time fractional derivative and application in a symmetric chaotic financial system. *Chaos Solitons Fractals* **2021**, *154*, 111590. [\[CrossRef\]](#)
11. Wang, R.; Zhang, Y.; Chen, Y.; Chen, X.; Xi, L. Fuzzy neural network-based chaos synchronization for a class of fractional-order chaotic systems: An adaptive sliding mode control approach. *Nonlinear Dyn.* **2020**, *100*, 1275–1287. [\[CrossRef\]](#)
12. Jahanshahi, H.; Shahriari-Kahkeshe, M.; Alcaraz, R.; Wang, X.; Singh, V.P.; Pham, V.-T. Entropy Analysis and Neural Network-Based Adaptive Control of a Non-Equilibrium Four-Dimensional Chaotic System with Hidden Attractors. *Entropy* **2019**, *21*, 156. [\[CrossRef\]](#) [\[PubMed\]](#)
13. Rostam, M.; Nagamune, R.; Grebenyuk, V. A hybrid Gaussian process approach to robust economic model predictive control. *J. Process Control* **2020**, *92*, 149–160. [\[CrossRef\]](#)
14. Rajaei, A.; Vahidi-Moghaddam, A.; Ayati, M.; Baghani, M. Integral sliding mode control for nonlinear damped model of arch microbeams. *Microsyst. Technol.* **2018**, *25*, 57–68. [\[CrossRef\]](#)
15. Yousefpour, A.; Vahidi-Moghaddam, A.; Rajaei, A.; Ayati, M. Stabilization of nonlinear vibrations of carbon nanotubes using observer-based terminal sliding mode control. *Trans. Inst. Meas. Control* **2019**, *42*, 1047–1058. [\[CrossRef\]](#)
16. Rao, M. *Filtering and Control of Macroeconomic Systems: A Control System Incorporating the Kalman Filter for the Indian Economy*; Elsevier: Amsterdam, The Netherlands, 2013.
17. Barnett, W.A.; He, S. *Unsolved Econometric Problems in Nonlinearity, Chaos, and Bifurcation*; University of Kansas, Department of Economics: Lawrence, KS, USA, 2012.
18. Jahanshahi, H.; Yousefpour, A.; Wei, Z.; Alcaraz, R.; Bekiros, S. A financial hyperchaotic system with coexisting attractors: Dynamic investigation, entropy analysis, control and synchronization. *Chaos Solitons Fractals* **2019**, *126*, 66–77. [\[CrossRef\]](#)
19. Wang, S.; He, S.; Yousefpour, A.; Jahanshahi, H.; Repnik, R.; Perc, M. Chaos and complexity in a fractional-order financial system with time delays. *Chaos Solitons Fractals* **2020**, *131*, 109521. [\[CrossRef\]](#)
20. Keller, A. Fuzzy control of macroeconomic models. *Int. J. Appl. Math. Comput. Sci.* **2009**, *5*, 115.
21. Wang, S.; Bekiros, S.; Yousefpour, A.; He, S.; Castillo, O.; Jahanshahi, H. Synchronization of fractional time-delayed financial system using a novel type-2 fuzzy active control method. *Chaos Solitons Fractals* **2020**, *136*, 109768. [\[CrossRef\]](#)
22. Jahanshahi, H.; Sajjadi, S.S.; Bekiros, S.; Aly, A.A. On the development of variable-order fractional hyperchaotic economic system with a nonlinear model predictive controller. *Chaos Solitons Fractals* **2021**, *144*, 110698. [\[CrossRef\]](#)
23. Allgöwer, F.; Zheng, A. *Nonlinear Model Predictive Control*; Birkhäuser: Basel, Switzerland, 2012; Volume 26.
24. Camacho, E.; Alba, C. *Model Predictive Control*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2013.
25. Jahanshahi, H.; Sari, N.N.; Pham, V.-T.; E Alsaadi, F.; Hayat, T. Optimal adaptive higher order controllers subject to sliding modes for a carrier system. *Int. J. Adv. Robot. Syst.* **2018**, *15*, 1729881418782097. [\[CrossRef\]](#)
26. Jahanshahi, H. Smooth control of HIV/AIDS infection using a robust adaptive scheme with decoupled sliding mode supervision. *Eur. Phys. J. Spec. Top.* **2018**, *227*, 707–718. [\[CrossRef\]](#)
27. Yousefpour, A.; Jahanshahi, H. Fast disturbance-observer-based robust integral terminal sliding mode control of a hyperchaotic memristor oscillator. *Eur. Phys. J. Spec. Top.* **2019**, *228*, 2247–2268. [\[CrossRef\]](#)
28. Xiong, P.-Y.; Jahanshahi, H.; Alcaraz, R.; Chu, Y.-M.; Gómez-Aguilar, J.; Alsaadi, F.E. Spectral Entropy Analysis and Synchronization of a Multi-Stable Fractional-Order Chaotic System using a Novel Neural Network-Based Chattering-Free Sliding Mode Technique. *Chaos Solitons Fractals* **2021**, *144*, 110576. [\[CrossRef\]](#)
29. Wang, Y.-L.; Jahanshahi, H.; Bekiros, S.; Bezzina, F.; Chu, Y.-M.; Aly, A.A. Deep recurrent neural networks with finite-time terminal sliding mode control for a chaotic fractional-order financial system with market confidence. *Chaos Solitons Fractals* **2021**, *146*, 110881. [\[CrossRef\]](#)
30. Wang, B.; Jahanshahi, H.; Volos, C.; Bekiros, S.; Yusuf, A.; Agarwal, P.; Aly, A. Control of a Symmetric Chaotic Supply Chain System Using a New Fixed-Time Super-Twisting Sliding Mode Technique Subject to Control Input Limitations. *Symmetry* **2021**, *13*, 1257. [\[CrossRef\]](#)
31. Wang, B.; Jahanshahi, H.; Dutta, H.; Zambrano-Serrano, E.; Grebenyuk, V.; Bekiros, S.; Aly, A.A. Incorporating fast and intelligent control technique into ecology: A Chebyshev neural network-based terminal sliding mode approach for fractional chaotic ecological systems. *Ecol. Complex.* **2021**, *47*, 100943. [\[CrossRef\]](#)
32. Wang, B.; Derbeli, M.; Barambones, O.; Yousefpour, A.; Jahanshahi, H.; Bekiros, S.; Aly, A.A.; Alharthi, M.M. Experimental validation of disturbance observer-based adaptive terminal sliding mode control subject to control input limitations for SISO and MIMO systems. *Eur. J. Control* **2021**. [\[CrossRef\]](#)
33. Jahanshahi, H.; Rajagopal, K.; Akgul, A.; Sari, N.N.; Namazi, H.; Jafari, S. Complete analysis and engineering applications of a megastable nonlinear oscillator. *Int. J. Non-linear Mech.* **2018**, *107*, 126–136. [\[CrossRef\]](#)
34. Li, J.-F.; Jahanshahi, H.; Kacar, S.; Chu, Y.-M.; Gómez-Aguilar, J.; Alotaibi, N.D.; Alharbi, K.H. On the variable-order fractional memristor oscillator: Data security applications and synchronization using a type-2 fuzzy disturbance observer-based robust control. *Chaos Solitons Fractals* **2021**, *145*, 110681. [\[CrossRef\]](#)

35. Bekiros, S.; Jahanshahi, H.; Bezzina, F.; Aly, A. A novel fuzzy mixed H_2/H_∞ optimal controller for hyperchaotic financial systems. *Chaos Solitons Fractals* **2021**, *146*, 110878. [\[CrossRef\]](#)
36. Tutueva, A.V.; Moysis, L.; Rybin, V.G.; Kopets, E.E.; Volos, C.; Butusov, D.N. Fast synchronization of symmetric Hénon maps using adaptive symmetry control. *Chaos Solitons Fractals* **2021**, *155*, 111732. [\[CrossRef\]](#)
37. Liu, Z.; Jahanshahi, H.; Volos, C.; Bekiros, S.; He, S.; Alassafi, M.O.; Ahmad, A.M. Distributed Consensus Tracking Control of Chaotic Multi-Agent Supply Chain Network: A New Fault-Tolerant, Finite-Time, and Chatter-Free Approach. *Entropy* **2021**, *24*, 33. [\[CrossRef\]](#) [\[PubMed\]](#)
38. Al-Hussein, A.-B.; Tahir, F.; Ouannas, A.; Sun, T.-C.; Jahanshahi, H.; Aly, A. Chaos Suppressing in a Three-Buses Power System Using an Adaptive Synergetic Control Method. *Electron.* **2021**, *10*, 1532. [\[CrossRef\]](#)
39. Yousefpour, A.; Jahanshahi, H.; Bekiros, S.; Muñoz-Pacheco, J.M. Robust adaptive control of fractional-order memristive neural networks. In *Mem-Elements for Neuromorphic Circuits with Artificial Intelligence Applications*; Elsevier BV: Amsterdam, The Netherlands, 2021; pp. 501–515.
40. Bhuvaneswari, N.; Uma, G.; Rangaswamy, T. Adaptive and optimal control of a non-linear process using intelligent controllers. *Appl. Soft Comput.* **2009**, *9*, 182–190. [\[CrossRef\]](#)
41. Zak, M. Expectation-based intelligent control. *Chaos Solitons Fractals* **2006**, *28*, 616–626. [\[CrossRef\]](#)
42. Chen, X. Research on application of artificial intelligence model in automobile machinery control system. *Int. J. Heavy Veh. Syst.* **2020**, *27*, 83. [\[CrossRef\]](#)
43. Das, P.; Chanda, S.; De, A. Artificial Intelligence-Based Economic Control of Micro-grids: A Review of Application of IoT. In *Lecture Notes in Electrical Engineering*; Springer Science and Business Media LLC: Berlin/Heidelberg, Germany, 2019; pp. 145–155.
44. Yousefpour, A.; Jahanshahi, H.; Munoz-Pacheco, J.M.; Bekiros, S.; Wei, Z. A fractional-order hyper-chaotic economic system with transient chaos. *Chaos Solitons Fractals* **2020**, *130*, 109400. [\[CrossRef\]](#)
45. Ho, Y.-C. Neuro-fuzzy And Soft Computing - A Computational Approach To Learning And Machine Intelligence [Book Reviews]. *Proc. IEEE* **1998**, *86*, 600–603.
46. Woelfel, J. Convergences in cognitive science, social network analysis, pattern recognition and machine intelligence as dynamic processes in non-Euclidean space. *Qual. Quant.* **2020**, *54*, 263–278. [\[CrossRef\]](#)
47. Sutton, R.; Barto, A. *Introduction to Reinforcement Learning*; MIT press Cambridge: Cambridge, MA, USA, 1998; Volume 135.
48. Kaelbling, L.P.; Littman, M.L.; Moore, A.W. Reinforcement Learning: A Survey. *J. Artif. Intell. Res.* **1996**, *4*, 237–285. [\[CrossRef\]](#)
49. Szepesvári, C. Algorithms for reinforcement learning. *Synth. Lect. Artif. Intell. Mach. Learn.* **2010**, *4*, 1–103. [\[CrossRef\]](#)
50. Mnih, V.; Kavukcuoglu, K.; Silver, D.; Rusu, A.A.; Veness, J.; Bellemare, M.G.; Graves, A.; Riedmiller, M.; Fidjeland, A.K.; Ostrovski, G.; et al. Human-level control through deep reinforcement learning. *Nature* **2015**, *518*, 529–533. [\[CrossRef\]](#) [\[PubMed\]](#)
51. Padmanabhan, R.; Meskin, N.; Haddad, W.M. Reinforcement learning-based control of drug dosing for cancer chemotherapy treatment. *Math. Biosci.* **2017**, *293*, 11–20. [\[CrossRef\]](#) [\[PubMed\]](#)
52. Bucci, M.A.; Semeraro, O.; Allauzen, A.; Wisniewski, G.; Cordier, L.; Mathelin, L. Control of chaotic systems by deep reinforcement learning. In *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*; The Royal Society: London, UK, 2019; Volume 475, p. 20190351.
53. Mao, Y.; Wang, J.; Jia, P.; Li, S.; Qiu, Z.; Zhang, L.; Han, Z. A Reinforcement Learning Based Dynamic Walking Control. In *Proceedings of the 2007 IEEE International Conference on Robotics and Automation*, Roma, Italy, 10–14 April 2007; pp. 3609–3614.
54. Qiao, J.; Hou, Z.; Ruan, X. Application of reinforcement learning based on neural network to dynamic obstacle avoidance. In *Proceedings of the 2008 International Conference on Information and Automation*, Changsha, China, 20–23 June 2008; pp. 784–788.
55. Wei, C.; Zhang, Z.; Qiao, W.; Qu, L. Reinforcement-Learning-Based Intelligent Maximum Power Point Tracking Control for Wind Energy Conversion Systems. *IEEE Trans. Ind. Electron.* **2015**, *62*, 6360–6370. [\[CrossRef\]](#)
56. Balashevich, N.; Gabasov, R.; Kalinin, A.; Kirillova, F. Optimal control of nonlinear systems. *Comput. Math. Math. Phys.* **2002**, *42*, 931–956.
57. Aliyu, M.D.S. An improved iterative computational approach to the solution of the Hamilton–Jacobi equation in optimal control problems of affine nonlinear systems with application. *Int. J. Syst. Sci.* **2020**, *51*, 2625–2634. [\[CrossRef\]](#)
58. Watkins, C.; Dayan, P. Q-learning. *Mach. Learn.* **1992**, *8*, 279–292. [\[CrossRef\]](#)
59. Sutton, R.S.; Barto, A.G. *Reinforcement Learning: An Introduction*; MIT Press: Cambridge, MA, USA, 2018.
60. Sutton, R.S.; Barto, A.G.; Williams, R.J. Reinforcement learning is direct adaptive optimal control. *IEEE Control Syst.* **1992**, *12*, 19–22. [\[CrossRef\]](#)
61. Kearns, M.; Singh, S. Near-Optimal Reinforcement Learning in Polynomial Time. *Mach. Learn.* **2002**, *49*, 209–232. [\[CrossRef\]](#)
62. Rummery, G.; Niranjan, M. *On-Line Q-Learning Using Connectionist Systems*; University of Cambridge, Department of Engineering: Cambridge, UK, 1994; Volume 37.
63. Wei, Q.; Lewis, F.L.; Sun, Q.; Yan, P.; Song, R. Discrete-Time Deterministic Q- Learning: A Novel Convergence Analysis. *IEEE Trans. Cybern.* **2016**, *47*, 1224–1237. [\[CrossRef\]](#)
64. Melo, F.S.; Ribeiro, M.I. Convergence of Q-learning with linear function approximation. In *Proceedings of the 2007 European Control Conference (ECC)*, Kos, Greece, 2–5 July 2007; pp. 2671–2678.

-
65. Puu, T. Multiplier-Accelerator Models Revisited. In *Economics of Space and Time*; Springer Science and Business Media LLC: Berlin/Heidelberg, Germany, 1997; pp. 145–159.
 66. White, M.V.; Rosser, J.B. From Catastrophe to Chaos: A General Theory of Economic Discontinuities. *South. Econ. J.* **1992**, *59*, 350. [[CrossRef](#)]