

Article

Digital Twins-Assisted Design of Next-Generation Advanced Controllers for Power Systems and Electronics: Wind Turbine as a Case Study

Meisam Jahanshahi Zeitouni ¹, Ahmad Parvaresh ² , Saber Abrazeh ³, Saeid-Reza Mohseni ¹, Meysam Gheisarnejad ⁴ and Mohammad-Hassan Khooban ^{5,*}

¹ Shiraz University of Technology, Tehran 1458889694, Iran; Jahanshahi.Meisam@gmail.com (M.J.Z.); sr.mohseni@alum.sharif.edu (S.-R.M.)

² Department of Electrical Engineering, Shahid Bahonar University of Kerman, Kerman 76169-14111, Iran; parvaresh.ahmad@yahoo.com

³ Shiraz University, Shiraz 71454, Iran; saber.abrazeh@yahoo.com

⁴ Department of Electrical Engineering, Najafabad Branch, Islamic Azad University, Isfahan 8514143131, Iran; me.gheisarnejad@gmail.com

⁵ DIGIT, Department of Engineering, Aarhus University, 8200 Aarhus, Denmark

* Correspondence: khooban@eng.au.dk

Received: 11 April 2020; Accepted: 29 April 2020; Published: 8 May 2020



Abstract: This paper proposes a novel adaptive controller based on digital twin (DT) by integrating software-in-loop (SIL) and hardware-in-loop (HIL). This work aims to reduce the difference between the SIL controller and its physical controller counterpart using the DT concept. To highlight the applicability of the suggested methodology, the regulation control of a horizontal variable speed wind turbine (WT) is considered for the design and assessment purposes. In the presented digital twin framework, the active disturbance rejection controller (ADRC) is implemented for the pitch angle control of the WT plant in both SIL and HIL environments. The design of the ADRC controllers in the DT framework is accomplished by adopting deep deterministic policy gradient (DDPG) in two stages: (i) by employing a fitness evaluation of wind speed error, the internal coefficients of HIL controller are adjusted based on DDPG for the regulation of WT plant, and (ii) the difference between the rotor speed waveforms in HIL and SIL are reduced by DDPG to obtain a similar output behavior of the system in these environments. Some examinations based on DT are conducted to validate the effectiveness, high dynamic performance, robustness and adaptability of the suggested method in comparison to the prevalent state-of-the-art techniques. The suggested controller is seen to be significantly more efficient especially in the compensation of high aerodynamic variations, unknown uncertainties and also mechanical stresses on the plant drive train.

Keywords: pitch angle control; active disturbance rejection controller (ADRC); digital twin (DT); deep deterministic policy gradient (DDPG)

1. Introduction and Preliminaries

Undeniably, renewable energy can promote countries to meet their valuable development goals through the provision of access to clean, secure, reliable and affordable energy [1,2]. Wind energy is one of the fastest growing and most promising energy sources, and its development has progressed tremendously worldwide. Therefore, the growth of wind turbine (WT) power generation has been increasing during the past decades [3,4]. Nowadays, multi mega-watt WTs are common in both off-shore and on-shore physical products of wind farms [5]. On the other hand, with the application of new-generation information technologies such as digital twins (DT) [6,7] in the wind turbine industry

and manufacturing, the convergence between physical products and virtual space has been expedited. Digital twin technology in WT systems are composed of physical products, virtual products and connection data that ties physical and virtual products together [8].

Any kind of WTs has been categorized into one of the two significant classifications: fixed-speed wind turbine (FS-WT) and variable-speed wind turbine (VS-WT) [9]. Many benefits such as improved energy capture, reduction in transient load, and better power conditioning can be achieved by comparing FS-WT and VS-WT [10]. Totally, for any type of WTs, it has been scientifically proven that control strategies play a significant role in WT characteristics and performances [11,12]. Depending on the rated wind speed, two main operating regions can be classified in VS-WT types, below and above rated wind speed. The main purpose of the controller at below-rated wind speed is the optimization of captured wind energy by rejecting the uncertainties in the turbine components. At the above-rated wind speed, the most important aim is to maintain the power of WT [13].

The global architecture of the pitch-controlled WT plant is illustrated in Figure 1. The efficient and reliable operation of a wind power plant heavily depends on the control systems applied to the WT at different operating regions. To restrict the aerodynamic power captured by the wind turbine at above-rated wind speed region, plentiful classic and modern control strategies to design efficient pitch angle controllers have been suggested, namely, proportional integral derivative (PID) and its variants [14,15], fuzzy PID control [16,17], linear parameter-varying (LPV) [18], nonlinear controller [19], optimal control [20], robust control [21], sliding mode control (SMC) [22] and model predictive control (MPC) [23]. On one hand, the classical PID controllers and their variants can hardly achieve excellent control performance under turbulent areas because newly WTs are high-order, multi-variable and highly coupled nonlinear systems. On the other hand, Fuzzy PID control, which usually uses a fuzzy controller to adjust the PID gains, is suitable for the high-order nonlinear system [24]. More importantly, variable gain PID control and MPC control both need to set up the precise model of wind turbine, which is very difficult to implement practically [25]. Another method using the SMC controller with a linear matrix inequality approach was proposed [26], which gives a good performance of the turbine output power as well as the robustness to the variations of the wind speed and the turbine parameters. However, the undesirable phenomenon, which is known as “chattering” is the main obstacle for its implementation. It is extremely harmful due to the fact that it leads to low control accuracy, high wear of mechanical parts movements and high heat losses in power circuit systems of WTs [27].

According to the nonlinear characteristics of VS-WTs and also to the existence of uncertainty, modeling these systems is time-consuming and complex. Due to the use of model-based controllers they need high order controllers of high complexity. An attempt has been made by the investigators to compensate for the uncertainties and non-linearity in the system by active disturbance rejection control (ADRC) [28,29]. Due to the inherent feature of disturbance rejection, ADRC becomes more popular and suitable in industries mainly for control purposes [30]. This type of controller has been inherited originally from classical PID controller, which is exactly contained in tracking differentials (TD), extended state observer (ESO), and non-linear state error feedback (NLSEF) controller [31]. The most important challenge in adopting ADRC as a control methodology to solve various control problems is tuning its numerous parameters, while the acceptable performance of ADRC has an irrefutable relation with tuning these coefficients. Although in past researches these parameters have been tuned by fuzzy method or optimization algorithms, these methods suffer from lack of adaptability and inability to learn [32–34]. Furthermore, these algorithms can only optimize and tune the parameters in certain cycles, and their performance decreases with changes in WT situations and operating points.

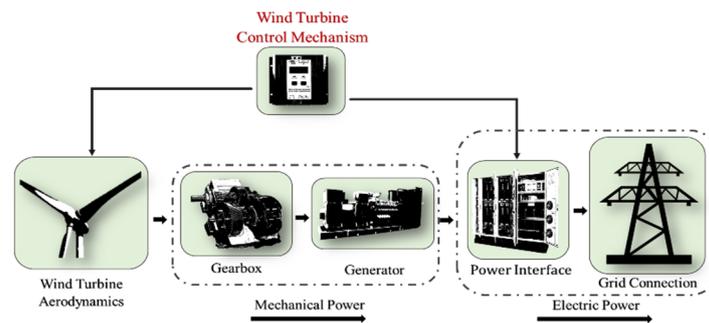


Figure 1. Global scheme of the WT plant.

Nowadays, reinforcement learning (RL) algorithms have been used increasingly in a wide range of systems, namely, robotic, energy management, etc. [35]. They are a type of machine learning technique that enables an agent to learn in an interactive environment by trial and error using feedback from its actions and experiences, and it is based on artificial neural networks with representation learning. The most popular algorithms of RL are the SARSA [36] and deep Q network (DQN) [37,38]. The key advances of DQN were the inclusion of an experience replay buffer (to overcome data correlation) and a different approach for the target Q-Network, whose weights change with the update of the main Q-Network to break the correlation between both networks. However, it was not designed for continuous states, which are deeply related to VS-WT control systems. Recently, to solve the DQN problems, a new deep RL algorithm, called deep deterministic policy gradients (DDPG) [39–42], has achieved good performance in many simulated continuous control problems. The main contributing advantage of DDPG algorithm can generate continuous actions, which is very valuable in a practical process.

In this work, a new application of DT based control strategy is introduced and implemented on a WT plant. The DT controller in this application is a virtual replica of the physical controller and can update itself concerning the measured information from its pre-designed physical counterpart. The ADRC controller has been adopted in HIL and SIL environments, and the parameters of the established controllers have been adjusted by the DDPG algorithm in the DT manner. For the adaptive realization of DT concept, the DDPG algorithm is applied to the WT system in two stages : (i) for the regulation of the wind speed in the HIL environment and (ii) for minimization of the system output behavior of the HIL and SIL environments. Several scenarios in the context of WT have been carried out to validate the correctness and applicability of the suggested DT controller method.

This paper is organized as follows. Section 2 establishes the nonlinear model of variable speed wind turbines and calculates the state-space equations for it. Then, Section 3 introduces the optimized control system with an integrated control algorithm combining ADRC with reinforcement learning. In Section 4, we describe and discuss the digital twin concept for implementing the proposed controller in the HIL and SIL environments. The results of simulation in the MATLAB platform and of implementing this controller on DSP hardware as a digital twin concept are presented in Section 5. Finally, the concluding remarks are summarized in Section 6.

2. Variable Speed Wind Turbine Model

The two-mass model structure of WT depicted in Figure 2, which is commonly used in the literature, is considered in the current work to illustrate the WT dynamics [43].

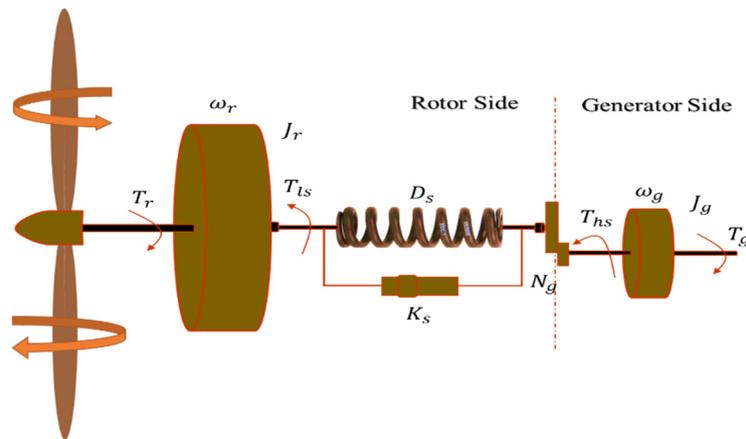


Figure 2. Two mass-model structure of considered wind turbine (WT).

The total power of wind has a direct relation with wind speed as in the following equation [23]:

$$P_W = \frac{1}{2} \rho A V^3 \quad (1)$$

where ρ is the air density (kg/m^3), A is the swept area of the turbine (m^2), and V is the wind speed (m/s). It has been long proven that if the wind speed is zero after passing the turbine, the total wind energy will be absorbed by the turbine [23]. However, due to the wind losses, it is practically impossible to transfer all the energy. Due to this reason, the power coefficient (C_p) is presented, which represents the aerodynamic efficiency of the wind turbine. Using C_p , the aerodynamic power of the turbine (P_a) can be expressed as follows:

$$P_a = \frac{1}{2} \cdot \rho \cdot C_p \cdot A \cdot V^3 \quad (2)$$

The power coefficient is a nonlinear function that depends on two paramount factors: the tip speed ratio (λ) and the blade pitch angle (β) as in the following numerical equation:

$$C_p(\lambda, \beta) = 0.5176 \left(\frac{116}{\lambda_i} - 0.4\beta - 5 \right) e^{-21/\lambda_i} + 0.0068 \quad (3)$$

The parameter λ_i can be calculated as follows:

$$\frac{1}{\lambda_i} = \frac{1}{\lambda + 0.08\beta} - \frac{0.035}{\beta^3 + 1} \quad (4)$$

The parameter λ is calculated by the blade tip speed and wind speed upstream of the rotor as:

$$\lambda = \frac{R\omega_r}{V} \quad (5)$$

with ω_r being the rotor angular speed. The power coefficient curves of the wind turbine have been shown as a function of λ and β in Figure 3.

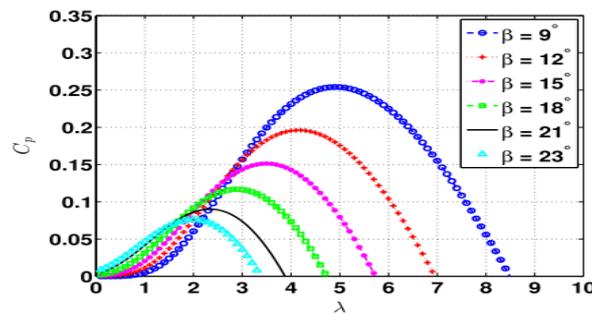


Figure 3. Curves of power coefficient $C_p(\lambda, \beta)$.

Nonlinear wind turbine model is shown in a generalized nonlinear form as follows [43]:

$$\begin{aligned} \dot{X} &= G(X) + Bu \\ &= \begin{bmatrix} \frac{P_r(x_1, x_4, V)}{x_1 J_r} - \frac{x_1 D_s}{J_r} + \frac{x_2 D_s}{N_g J_r} - \frac{x_3 K_s}{J_r} \\ \frac{x_1 D_s}{N_g J_g} - \frac{x_2 D_s}{N_g^2 J_g} + \frac{x_3 K_s}{N_g J_g} - \frac{T_g}{J_g} \\ x - \frac{x_2}{N_g} \\ -\frac{1}{\tau_\beta} x_4 \end{bmatrix} \\ &+ \begin{bmatrix} 0 \\ 0 \\ 0 \\ \frac{1}{\tau_\beta} \end{bmatrix} u \end{aligned} \tag{6}$$

The state vector x , control input u , and nonlinear vector $G(X)$ are defined as:

$$X = \left[\omega_r \quad \omega_g \quad \delta \quad \beta \right]^T \tag{7}$$

$$Y = \omega_r \tag{8}$$

where ω_r is the rotor speed, ω_g is the generator speed and δ is the twist angle. τ_β is the time constant of pitch actuator and β_r is the pitch angle control. T_g is the generator torque, J_r and J_g are the rotor and generator inertia, N_g is the gear ratio, D_s and K_s are the drive-train damping and spring constant, respectively.

The objective of this paper is to develop a novel digital twin-based pitch angle controller for rotor speed regulation at Region III of the wind turbine operation, by restricting the power derived from the wind turbine. The parameters of the wind turbine system are borrowed from [43].

3. Design of Proposed Controller

3.1. ADRC Technique

The ADRC is generally regarded as a model-free technique since it does not need complete knowledge of the system. This method is introduced to deal with general nonlinear uncertain plants as it can eliminate the impact of the internal/external disturbances in real-time. Since the ADRC originates from the traditional PID controller, it has the same benefit of fast response and strong robustness as the PID controller. The block diagram of the ADRC controller is illustrated in Figure 4, which consists of TD, NLSEF and ESO. In Figure 4, w denotes the external disturbances, y is the output signal, and u is the control signal.

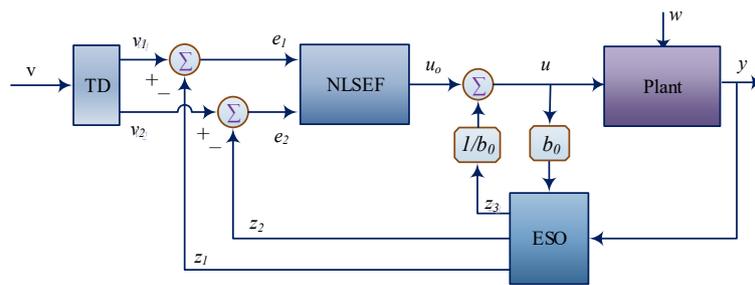


Figure 4. The block diagram of the active disturbance rejection controller (ADRC) controller.

The TD is established to improve the ideal transient process and smooth signals for the controller. One feasible second-order TD can be formulated as:

$$\dot{v}_1 = v_2 \tag{9}$$

$$\dot{v}_2 = fhan(v_1 - v(t).v_2.r.h_0) \tag{10}$$

where the term v denotes the control objective, v_1 and v_2 are the desired trajectory and its derivative of v , respectively. Likewise, r and h_0 are the speed and filtering factors, respectively. The formulation of $fhan(v_1 - v(t).v_2.r.h_0)$ is given as:

$$\begin{aligned} \dot{v}_1 &= v_2 \\ \dot{v}_2 &= fhan(v_1 - v(t).v_2.r.h_0) \\ d &= rh_0^2.a_0 = v_2.h_0 \\ y &= v_1 - v(t) + a_0 \\ a_1 &= \sqrt{d^2 + 8.d|y|} \\ a_2 &= a_0 + sign(y)(a_1 - d)/2 \\ s_y &= (sign(y + d) - sign(y - d))/2 \\ a &= (a_0 + y - a_2)s_y + a_2 \\ s_a &= (sign(a + d) - sign(a - d))/2 \\ fhan &= -r\left(\frac{a}{d} - sign(a)\right)s_a - rsign(a) \end{aligned} \tag{11}$$

ESO is adopted as a type of robust control that could assess all the external disturbances and internal perturbations and then compensate all of them to get the correct response. The relationship between input and output in ESO is illustrated below:

$$\begin{aligned} e(t) &= z_1 - y \\ \dot{z}_1 &= z_2 - \beta_1 e \\ \dot{z}_2 &= z_3 - \beta_2 fal(e.a_1.\delta) + b_0 u \\ \dot{z}_3 &= -\beta_3 fal(e.a_2.\delta) \end{aligned} \tag{12}$$

where z_1, z_2 and z_3 are the observer output signals and the terms of β_1, β_2 and β_3 denote the design variables of ESO.

The NLSEF is established in the ADRC structure to achieve the control input of the system by combination estimated states (z_1, z_2) of ESO and output signals (e_1, e_2) of TD. The low control of NLSEF is described by:

$$\begin{aligned} e_1 &= v_1 - z_1 \\ e_2 &= v_2 - z_2 \\ u_0 &= k_1 fal(e_1.a_1.\delta) + k_2 fal(e_2.a_2.\delta) \end{aligned} \tag{13}$$

where k_1 and k_2 are the proportional and differential parameters, respectively, and $fal(\cdot)$ is a nonlinear function, which is expressed as:

$$fhan(x,a,\delta) = \begin{cases} \frac{x}{\delta^{1-a}} \cdot |x| \leq a \\ sign(x)|x|^a \cdot |x| > a \end{cases} \quad (14)$$

Finally, the controller is achieved by

$$u = u_0 - \frac{z_3}{b_0} \quad (15)$$

3.2. Deep Reinforcement Learning

Reinforcement Learning (RL) is a widely used methodology in machine learning due to its potential to learn highly in complex environments. It is applicable in many research and practical areas such as game theory, control theory, simulation-based optimization, multi-agent systems and statistics. Practically, there are numerous challenges such as the temporal correlation of data, divergence of learning or continuous nature of inputs and outputs, to implement RL in practical control problems. Recently, the Deep Q-Network (DQN) has revealed a new set of possibilities to solve most of the mentioned problems; however, control actions of DQN are restricted to a small action space that decrease its applicability. According to the advances of DQN and the actor-critic paradigm, deep deterministic policy gradients (DDPG) had been proposed by Lillicrap et al. [44] as an algorithm that solves continuous control problems by integrating neural networks in the RL paradigm.

Reinforcement Learning is concerned with how agents ought to take actions in an environment to maximize the reward. More specifically, all RL applications involve interaction between an active decision-making agent and its environment, within which the agent seeks to achieve a goal despite uncertainty about its environment. This interaction process is formulated as a Markov Decision Process (MDP) which is described by the concepts below:

- Environment: Space through which the agent moves and responds to the agent. The environment takes the agent's current state and action as input and returns as output the agent's reward and its next state.
- Agent: An agent tries to find an optimal policy to map the state of the environment to an action that will maximize the rewards of accumulated future in turn. State ($s \in \mathbb{S}$): \mathbb{S} is state-space or all possible states of the agent in the environment.
- Policy (π): The policy is the strategy that the agent employs to determine the next action based on the current state. It maps states to actions, the actions that promise the highest reward.
- Action ($a \in \mathbb{A}$): It is the set of all possible moves that the agent can make.
- Reward ($r \in \mathbb{R}$): A reward is feedback by which the success or failure of an agent's actions in a given state is evaluated.
- Value function (V^π): It is the expected long-term return with a discount, as opposed to the short-term reward.
- Q-value or action-value (Q): Q-value is similar to V^π , except that it takes an extra parameter, the current action a .

The final goal of an RL subject is to learn a policy $\pi : S \rightarrow A$, which maximizes the expectation of a long-term discounted reward as below:

$$J = \mathbb{E}_{r_i, s_i \sim E, a_i \sim \pi} [G_1] \quad (16)$$

where $G_t = \sum_{k=0}^{\infty} \gamma^k r_{t+k}$ is the total long term discounted reward at each step, and $\gamma \in (0, 1]$ refers to the discount factor which dampens the rewards' effect on the agent's choice of action to make future rewards worth less than immediate rewards.

Considering the target policy $\pi : S \rightarrow A$, which maps each state-space to deterministic actions, a value function V^π is formulated as a depiction of total discounted reward G_t for each $s \in \mathbb{S}$.

$$V^\pi(s) = \mathbb{E}_\pi[G_t | s_t = s] \tag{17}$$

Using the Bellman equation, V^π can be recursively described as below:

$$V^\pi(s) = \mathbb{E}_\pi[r_t + \gamma V^\pi(s_{t+1}) | s_t = s] \tag{18}$$

Action-value function Q^π is represented in the equation below as the value function V^π defined based on the Bellman equation:

$$Q^\pi(s, a) = \mathbb{E}_\pi[r_t + \gamma Q^\pi(s_{t+1}, a_{t+1}) | s_t = s, a_t = a] \tag{19}$$

The policy that maximizes the action-value function or the value function is the optimal policy ($\pi^* = \operatorname{argmax}_a Q^*(s, a)$). Regarding DDGP algorithm, it contains two neural networks, $Q(s_t, a_t | \theta^Q)$ and $\mu(s_t | \theta^\mu)$, which have proven to perform well to solve continuous problems. In the algorithm, both functions, $Q(s_t, a_t)$ and $\mu(s_t)$, are approximated by the aforementioned neural networks respectively, where θ^Q and θ^μ are the weights of the critic and actor networks. The critic network is updated by minimizing the loss function based on the stochastic gradient descent:

$$\mathcal{L}(\theta^Q) = E_{(s,a)} \left[\left(y_t - Q(s_t, a_t | \theta^Q) \right)^2 \right] \tag{20}$$

where

$$y_t = r_t(s_t, a_t) + \gamma Q(s_{t+1}, \mu(s_t | \theta^\mu) | \theta^Q) \tag{21}$$

The actor network's coefficient θ^μ is updated based on the following policy gradient:

$$\begin{aligned} \nabla_{\theta^\mu} J^{\theta^\mu} &\approx \mathbb{E}_{s_t \sim \rho^\beta} \left[\nabla_{\theta^\mu} Q(s, a | \theta^Q) \Big|_{a=\mu(s | \theta^\mu)} \nabla_{\theta^\mu} \mu(s | \theta^\mu) \right] \\ &= \mathbb{E}_{s_t \sim \rho^\beta} \left[\nabla_a Q(s, a | \theta^Q) \Big|_{a=\mu^\theta(s)} \nabla_{\theta^\mu} \mu(s | \theta^\mu) \right] \end{aligned} \tag{22}$$

In Equation (22), ρ is the discounted distribution, and β is a specific policy to the current policy π .

In the DDGP algorithm, a replay buffer is used to weaken the correlations existing in the input experiences, and target network approaches are exploited to stabilize the training procedure. According to the replay buffer mechanism, which uses a finite-size memory, each experience tuple $e = (s_t, a_t, r_t, s_{t+1})$ of each time step saved in an R -sized experience memory $\mathcal{D} = \{e_1, e_2, \dots, e_R\}$. In each step of the training process, a mini-batch of previous saved experiences is uniformly sampled from the memory \mathcal{D} . In each step of the training process, a mini-batch of previous saved experiences is uniformly sampled from the memory R to update the neural network at a time step. In terms of the stability of the DDPG learning method, two additional neural networks, $Q'(s, a | \theta^{Q'})$ and $\mu'(s | \theta^{\mu'})$, named target networks, are also adopted for the actor and critic neural networks to avoid the instability of DDPG learning. Both weights of $\theta^{Q'}$ and $\theta^{\mu'}$ are updated from the current networks at each time step. Moreover, in the training phase, a Laplacian exploration noise (\mathcal{N}), which is represented in Equation (23), is added to the actions provided by the agent (i.e., $a_t = \mu(s_t | \theta^\mu) + \mathcal{N}$) for exploration purposes.

$$\mathcal{N}(x|b) \sim \frac{1}{2b_t} \exp\left(-\frac{x}{b_t}\right) \tag{23}$$

The pseudo-code for the standard DDPG algorithm is presented in Algorithm 1 [45], and the DDPG based online learning framework is illustrated in Figure 5 [46].

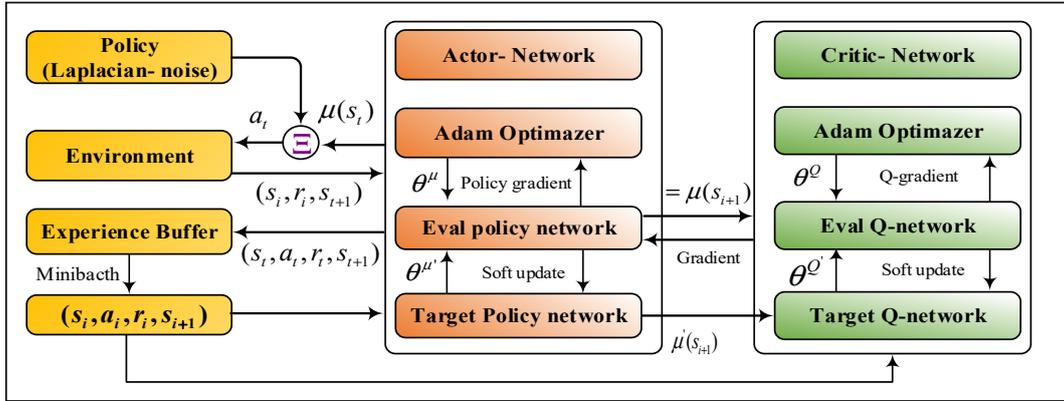


Figure 5. Illustration of an actor-network critic-network.

The pseudo-code for the standard DDPG algorithm is presented in Algorithm 1 [45], and the DDPG based online learning framework is illustrated in Figure 5 [46].

Algorithm 1: Framework of the DDPG for the WMR system.

- 1: Randomly initialize critic $Q(s, a | \theta^Q)$ and actor $\mu(s | \theta^\mu)$ networks with weights θ^Q and θ^μ
- 2: Initialize target networks Q' and μ' with weights $\theta^{Q'} \leftarrow \theta^Q, \theta^{\mu'} \leftarrow \theta^\mu$
- 3: Set up empty replay buffer R
- 4: **for** episode = 1 to M **do**
- 5: Begin with a Laplacian noise \mathcal{N} for exploration
- 6: Receive initial observation state
- 7: **for** $t = 1$ to T **do**
- 8: Apply action $a_t = \mu(s_t | \theta^\mu) + \mathcal{N}$ to environment
- 9: Observe next state s_{t+1} and reward r_t
- 10: Store following transitions (s_t, a_t, r_t, s_{t+1}) into replay buffer R
- 11: Sample random minibatch of K transitions from R
- 12: Set $y_i = r_i + \gamma Q'(s_{i+1}, \mu'(s_{i+1} | \theta^{\mu'}) | \theta^{Q'})$
- 13: Update critic by the loss: $L = \frac{1}{N} \sum_i (y_i - Q(s_i, a_i | \theta^Q))^2$
- 14: Update the actor policy using the sampled policy gradient:

$$\nabla_{\theta^\mu} J^{\theta^\mu} \approx \frac{1}{N} \sum_i \nabla_a Q(s, a | \theta^Q) \Big|_{a=\mu^\theta(s)} \nabla_{\theta^\mu} \mu(s | \theta^\mu)$$

- 15: Update the target networks:

$$\theta^{Q'} \leftarrow \tau \theta^Q + (1 - \tau) \theta^{Q'}, \theta^{\mu'} \leftarrow \tau \theta^\mu + (1 - \tau) \theta^{\mu'}$$

- 16: **end for**
 - 17: **end for**
-

4. Digital Twin Controller of WT System

For the aim of the control problem of a typical WT system, the digital twin (DT) method has been suggested as a significant strategy for performing, testing, and improving. Digital twin technology in WT systems has been composed of physical products, virtual products and connection data that ties physical and virtual products together. The DT is a digital replica or representation of a physical object

or an intangible system that can be examined, altered and tested without interacting with it in the real world and avoiding negative consequences. It is a bridge between the digital world and the physical world. Digital twins have achieved remarkable popularity in recent years mainly in the industrial field.

Under this approach, a digital twin of the wind turbine pitch angle controller, which is proposed in this paper, has been defined and implemented on a Texas Instruments (TI) digital signal processor (DSP) computing device. As shown in Figure 6, the hardware-in-loop (HIL) [47,48] concept enables the test of controller algorithms on the actual controller hardware deployed on the wind turbine. In contrast, the software-in-the-loop (SIL) concept allows the test of the algorithms but neglects the test of the controller hardware.

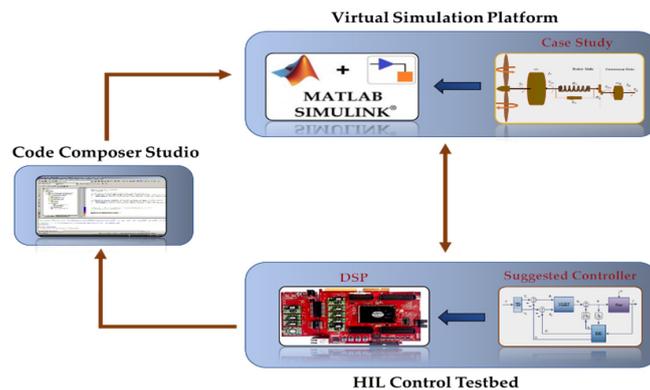


Figure 6. The concept of hardware-in-loop (HIL) testing.

The purpose of using digital twins (DT) here is to design the controllers in such a way that the system in the SIL environment behaves similarly to the HIL one. In the suggested methodology, the ADRC controller is adopted in both the SIL and HIL environments for the pitch angle control of the WT plant. In this work, the design of the DT control-based strategy has been realized in two stages by the DDPG algorithm as depicted in Figure 7, which are discussed in the following sub-sections.

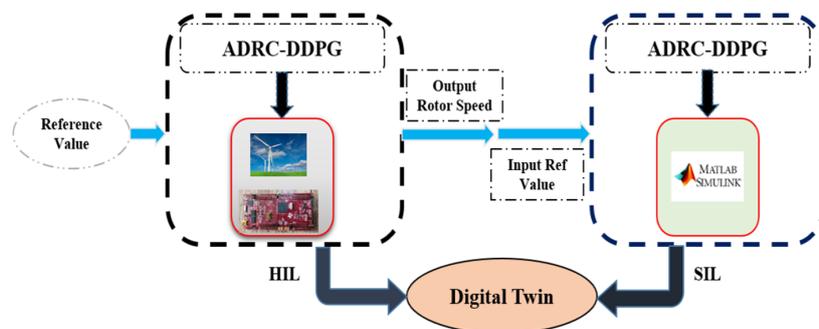


Figure 7. A proposed strategy for the combination of HIL and SIL testing.

4.1. Design of the HIL Controller

Firstly, the DDPG scheme with the actor-critic architecture is applied as a parameter tuner to provide the regulative signals to set the NLSEF gains of the HIL setup adaptively. The following equation shows the coefficients of the self-adapting ADRC method:

$$\begin{cases} k_1^{HIL} = k_{1,0}^{HIL} + \Delta k_{1,0}^{HIL} \\ k_2^{HIL} = k_{2,0}^{HIL} + \Delta k_{2,0}^{HIL} \end{cases} \quad (24)$$

where $k_{1,0}^{HIL}$ and $k_{2,0}^{HIL}$ are the NLSEF initial coefficients in the HIL setup and $\Delta k_{1,0}^{HIL}$ and $\Delta k_{2,0}^{HIL}$ are the regulative signals, which are tuned by the DDPG algorithm. The state variables including the rotor

speed of the HIL setup ω_r^{HIL} , its error ($e^{HIL} = \omega_r^{HIL} - \omega_{ref}^{HIL}$) and its error integral are expressed as $s_t^{HIL} = \{\omega_r^{HIL} \cdot e^{HIL} \cdot \int e^{HIL} dt\}$.

Equation (25) shows the reward function for the designing the HIL controller, which is a basis for the evaluation of the DDPG control actions (k_1 and k_2).

$$r_t^{HIL} = \frac{1}{(e^{HIL})^2} \tag{25}$$

With an increase in rotor speed error that results from a system perturbation, r_t^{HIL} comes down, and the weights of the actor and critic networks need to be updated accordingly. For more specification, to mitigate the effect of the perturbation, the actor-network senses the state variables s_t^{HIL} and then generates two continuous regulative signals. Then, the critic network receives s_t^{HIL} , $k_{1,0}^{HIL}$ and $k_{2,0}^{HIL}$, and the weights of the critic network are trained. Then the function $Q(s_t, a_t)$ is derived in the output layer, which leads to an updated DDPG network with adapted regulative signals to feed the controller.

4.2. Design of the Digital Twin Controller Based on the System Output Specification of the HIL Setup

In the second step, the rotor speed output of the HIL environment is introduced as the reference input for the rotor speed regulation of the SIL environment. To do this, similarly to the HIL controller design, the NLSEF coefficients are adaptively regulated by employing the DDPG method, given as:

$$\begin{cases} k_1^{SIL} = k_{1,0}^{SIL} + \Delta k_{1,0}^{SIL} \\ k_2^{SIL} = k_{2,0}^{SIL} + \Delta k_{2,0}^{SIL} \end{cases} \tag{26}$$

where $k_{1,0}^{SIL}$ and $k_{2,0}^{SIL}$ are the NLSEF initial coefficients of SIL setup, which are tuned by the DDPG algorithm according to the pre-designed HIL controller.

For the design of the SIL controller, the state variables are chosen as $s_t^{HIL} = \{\omega_r^{HIL} \cdot e^{HIL} \cdot \int e^{HIL} dt\}$, where ω_r^{SIL} and e^{SIL} are the rotor speed and rotor speed error in the SIL.

A reward function is also constructed for the optimal setting of the DT controller in the SIL setup, described as:

$$r_t = \frac{1}{|\omega_r^{SIL} - \omega_r^{HIL}|} \tag{27}$$

Based on the reward function of (27), the actor and critic network of the DDPG scheme are trained in a way that minimizes the difference between the output responses of the WT system in the SIL and HIL environments. To do this, the actor takes the state variable s_t^{SIL} and generates continuous regulatory signals. Likewise, s_t^{SIL} , $k_{1,0}^{SIL}$ and $k_{2,0}^{SIL}$ are considered as the input of the critic network, and a continuous Q-Value is produced in the output of the network.

5. Experimental Results

For the optimal design of the ADRC controller based digital twin concept, the actor and critic networks are trained over 200 episodes. The DDPG learning agent interacts with the environment at a frequency of 10 KHz, which corresponds to one training step. The weights of both actor and critic network are being optimized with a base learning rate of 10^{-4} and 10^{-3} , respectively, employing Adam optimizer.

In the following section, the effectiveness and efficiency of the proposed control system is tested by Real-Time SIL (RT-SIL) MATLAB simulation experiments, as well as Real-Time HIL (RT-HIL) board. The output results are evaluated and verified under the following three typical scenarios of WT process: (i) the step changes of wind speed, (ii) the random changes of wind speed and (iii) the parametric uncertainty in the turbine model. In the comparative analysis of the real-time setup, the output results of the proposed method are compared with the ADRC and PI controllers in HIL and SIL environments.

5.1. Scenario I: The Step Changes in Wind Speed

In the first scenario, a multi-step variation of wind speed (which is varied within [12 m/s 21 m/s]) is applied to the nonlinear WT plant, as depicted in Figure 8. The average accumulated reward for the full-simulated training phase of 200 episodes in HIL is depicted in Figure 9. As shown in this figure, the reward chart follows an upward trend since episode 5 and has been almost constant since episode 20 onwards. This indicates that the rotor speed error at the HIL output is significantly reduced, and the DDPG algorithm calculates the coefficients controller k_1^{HIL} and k_2^{HIL} accurately.

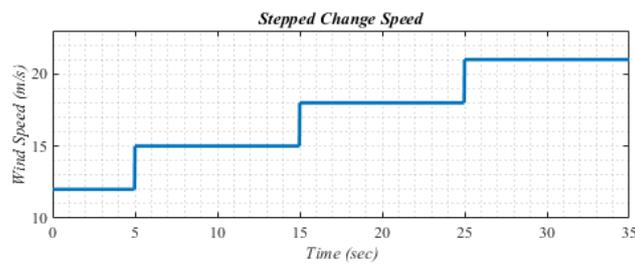


Figure 8. Profile of the multi-step change of wind speed.

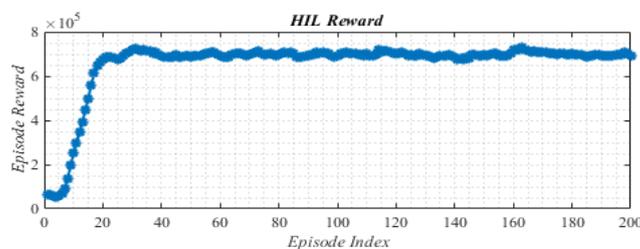


Figure 9. SIL and HIL rewards.

The HIL output comparative results of ADRC-DDPG, ADRC and PI controllers under the multi-step disturbance are presented in Figure 10. From Figure 10, it is clear that the ADRC-DDPG and ADRC controllers obtain satisfactory performance to control the WT system, but the rotor speed outcomes of the PI controller experience large deviations. It is also observed that the transient specifications of the rotor speed in terms of settling time and overshoot have been remarkably ameliorated in the suggested controller compared to the other two types of pitch angle control strategies.

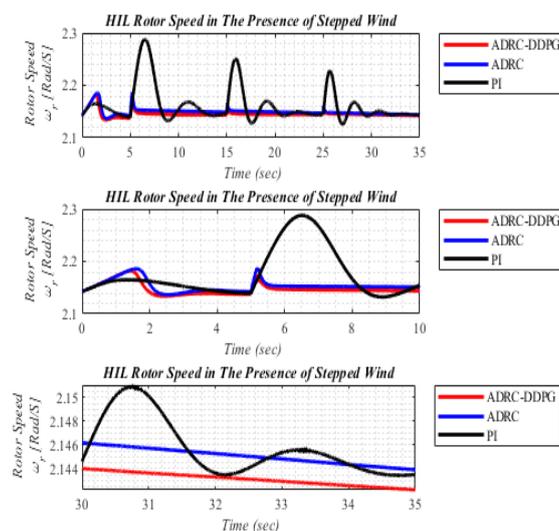


Figure 10. HIL output comparative results of active disturbance rejection controller-deep deterministic policy gradient (ADRC-DDPG), active disturbance rejection controller (ADRC) and PI controllers according to the scenario I.

Besides, the average reward of the DDPG agent when minimizing the difference between the output response specification of HIL and SIL environments is presented as illustrated in Figure 11. From Figure 11, it is noted that the average reward is eventually increased and stabilized during the 200 episodes, which proves the correctness and usefulness of the DDPG agent for the studied WT system based digital twin concept.

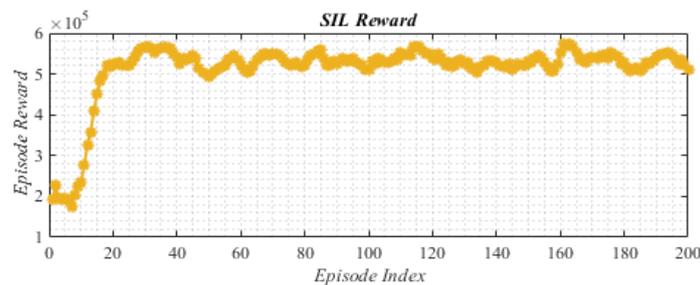


Figure 11. SIL and HIL rewards.

The SIL responses for the ADRC-DDPG, ADRC and PI controllers under the concerning wind speed disturbance are shown in Figure 12. The outcomes of Figure 12 reveal that with the application of the suggested controller, a superior performance of rotor speed response is achieved to that of the other two pitch angle control strategies. By comparing the curves of Figures 10 and 12, it can be inferred that under the actions of the ADRC-DDPG, the difference of rotor speed waveforms of the HIL and SIL is further reduced.

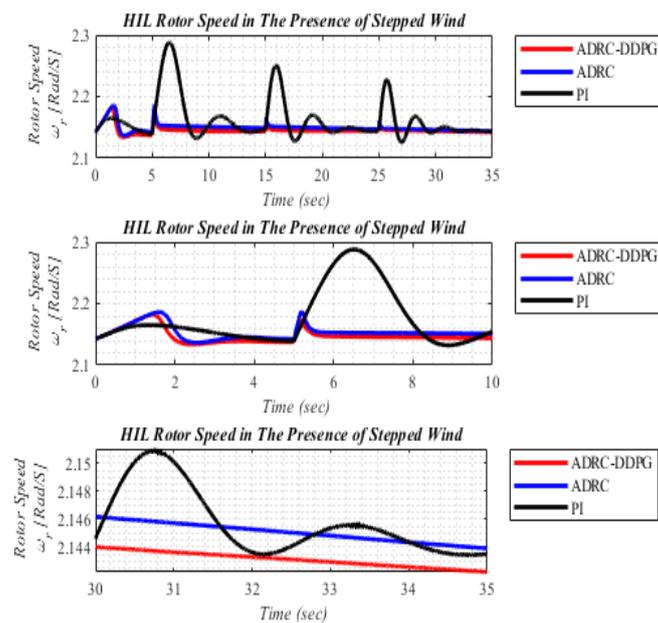


Figure 12. SIL output comparative results of ADRC-DDPG, ADRC and PI controllers according to the scenario I.

The performance index corresponds to the dynamic specifications of the WT system under the multi-step wind speed, such as settling time, overshoot and output error, which are furnished in Table 1. From the quantitative analysis of Table 1, it is noticed that with the application of the ADRC-DDPG, the considered dynamic specifications are greatly improved and outperform the ADRC and PI controllers for the same investigated plant.

Table 1. Settling time, overshoot and output error comparison outcomes according to the scenario I.

Performance Measurements	ADRC-DDPG		ADRC		PI	
	HIL	SIL	HIL	SIL	HIL	SIL
Settling time	2.1	2.3	5.2	5.8	27	29
Overshoot	1.78%	3.13%	2.10%	5.00%	5.93%	6.86%
Error	0.6577	0.7143	0.9315	1.0875	16.5921	17.3392

5.2. The Random Changes in wind Speed

To study the feasibility of the adaptive ADRC-DDPG controller in a more realistic condition of the WT plant, a random variation of wind speed (which is fluctuated within [12 m/s 18 m/s]) is applied to the system as depicted in Figure 13.

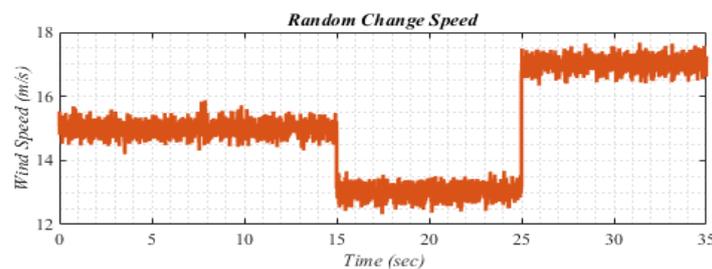


Figure 13. Profile of the random change of wind speed.

Figures 14 and 15 illustrate the rotor speed responses of the HIL and SIL environments in the scenario for the pitch angle controllers, respectively. From Figures 14 and 15, it is found that the suggested adaptive ADRC-DDPG controller ameliorates the performance of the WT plant in both environments compared to other pitch angle controllers, especially from settling time as well as the amplitude of fluctuations point of view.

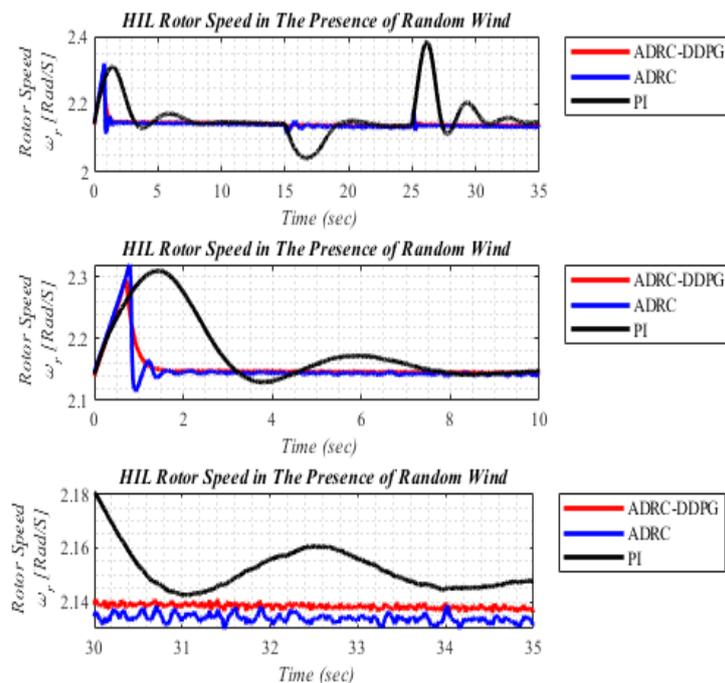


Figure 14. HIL output comparative results of ADRC-DDPG, ADRC and PI controllers according to scenario II.

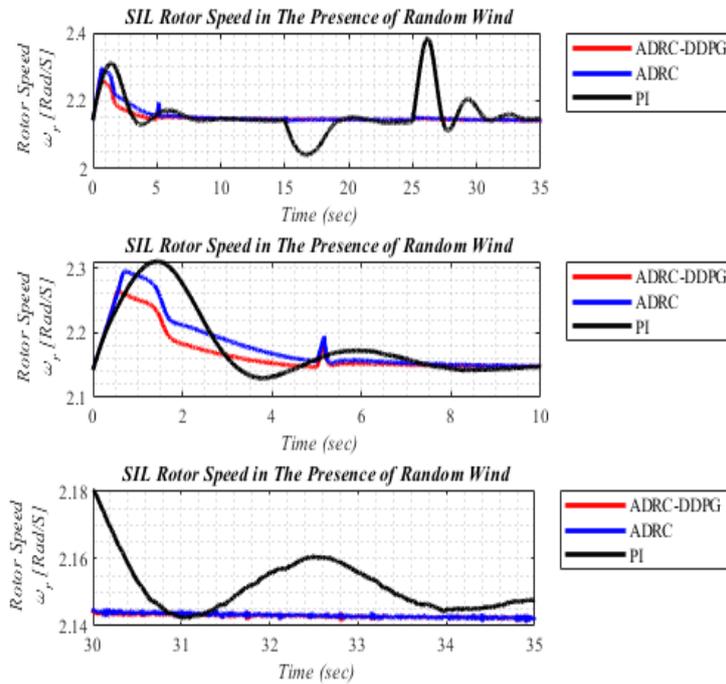
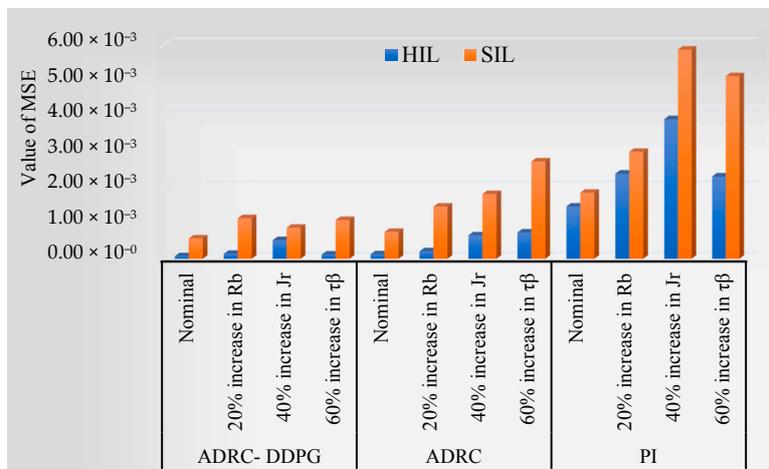


Figure 15. SIL output comparative results of ADRC-DDPG, ADRC and PI controllers according to scenario II.

5.3. The Parametric Uncertainty in the Turbine Model

In this scenario, the robustness and supremacy of the suggested controller are evaluated by imposing some uncertainties in the WT model for both the HIL and SIL environments as follows: $R_b = +20\%$, $J_r = +40\%$ and $t_B = +60\%$. The effects of these variations on the output rotor speed are calculated using two standard error measurement criteria including Mean Square Error (MSE) and Root MSE (RMSE). Figure 16a,b depicts the bar chart curves of MSE and RMSE corresponding to the designed pitch angle controllers for both the HIL and SIL. Looking at the information presented in this figure, not only does the suggested ADRC-DDPG controller have the best performance against uncertainties but also, and more importantly, based on the digital twin concept, SIL output variations follow almost completely the HIL output variations.



(a)

Figure 16. Cont.

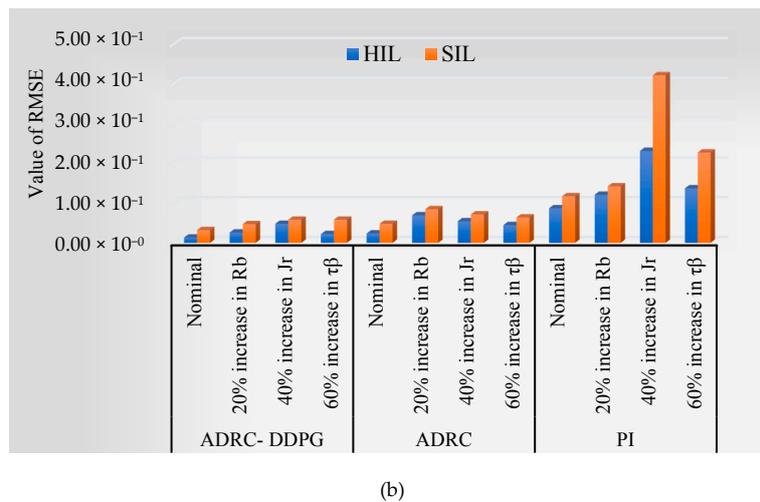


Figure 16. Bar chart comparison of standards under the parametric uncertainties: (a) values of MSE; (b) values of RMSE.

Remark 1. The outcomes of the three scenarios reveal that in spite the ADRC controller can compensate the disturbances and uncertainties in the WT system, its obtained outcomes are not optimal. On the other hand, the PI controller performs poorly that it could not satisfy the standards from a power engineering point of view. However, the suggested DDPG based ADRC scheme provides a better level of rotor speed regulation, reaching the pitch angle control goals in a lesser time responding and smoother outcome.

6. Conclusions

This paper concentrates on developing a novel adaptive ADRC controller-based digital twin for pitch control of a nonlinear speed WT plant. In this application, to regulate the speed rotor of a WT in HIL, the ADRC controller is firstly designed by the DDPG algorithm for this environment. Then, the output response of HIL is considered as the reference for the design of SIL controller. To do this, the ADRC of the SIL controller is designed by the DDPG algorithm minimizing the difference between the speed rotor waveforms of the HIL and SIL. To verify the efficiency of the suggested digital twin controller, critical examinations are carried out for pitch angle control of the WT plant in both the SIL and HIL environments. Comprehensive examinations demonstrate the dynamic behavior improvement of the digital twin-based system compared to state-of-the-art schemes.

Author Contributions: Investigation, S.A.; methodology, M.J.Z., A.P. and S.-R.M.; software, M.G.; supervision, M.-H.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

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

References

- Dehghani, M.; Khooban, M.H.; Niknam, T.; Rafiei, S.M.R. Time-varying sliding mode control strategy for multibus low-voltage microgrids with parallel connected renewable power sources in islanding mode. *J. Energy Eng.* **2016**, *142*, 05016002. [[CrossRef](#)]
- Gheisarnejad, M.; Mohammadi-Moghadam, H.; Boudjadar, J.; Khooban, M.H. Active Power Sharing and Frequency Recovery Control in an Islanded Microgrid With Nonlinear Load and Nondispatchable DG. *IEEE Syst. J.* **2019**, *14*, 1058–1068. [[CrossRef](#)]
- Gao, Z.; Tang, C.; Zhou, X.; Ma, Y.; Wu, Y.; Yin, J.; Xu, X. An overview on development of wind power generation. In Proceedings of the 2016 Chinese Control and Decision Conference (CCDC), Yinchuan, China, 28–30 May 2016; pp. 435–439.

4. Heydari-Doostabad, H.; Khalghani, M.R.; Khooban, M.H. A novel control system design to improve LVRT capability of fixed speed wind turbines using STATCOM in presence of voltage fault. *Int. J. Electr. Power Energy Syst.* **2016**, *77*, 280–286. [[CrossRef](#)]
5. Mahdizadeh, A.; Schmid, R.; Oetomo, D. Fatigue load mitigation in multi-megawatt wind turbines using output regulation control. In Proceedings of the 2017 21st International Conference on System Theory, Control and Computing (ICSTCC), Sinaia, Romania, 19–21 October 2017; pp. 163–168.
6. Erikstad, S.O. *Merging Physics, Big Data Analytics and Simulation for the Next-Generation Digital Twins*. *Hiper*, No. September; High-Performance Marine Vehicles: Zevenwacht, South-Africa, 2017; pp. 139–149.
7. Wagg, D.J.; Gardner, P.; Barthorpe, R.J.; Worden, K. On Key Technologies for Realising Digital Twins for Structural Dynamics Applications. In *Model Validation and Uncertainty Quantification*, 2019 ed.; Springer: Berlin, Germany, 2020; Volume 3, pp. 267–272.
8. Sivalingam, K.; Sepulveda, M.; Spring, M.; Davies, P. A review and methodology development for remaining useful life prediction of offshore fixed and floating wind turbine power converter with digital twin technology perspective. In Proceedings of the 2018 2nd International Conference on Green Energy and Applications (ICGEA), Singapore, 24–26 March 2018; pp. 197–204.
9. Rezaei, V. Advanced control of wind turbines: Brief survey, categorization, and challenges. In Proceedings of the 2015 American Control Conference (ACC), Chicago, IL, USA, 1–3 July 2015; pp. 3044–3051.
10. Gharibeh, H.F.; Khiavi, L.M.; Farrokhifar, M.; Alahyari, A.; Pozo, D. Capacity Value of Variable-Speed Wind Turbines. In Proceedings of the 2019 IEEE Milan PowerTech, Milan, Italy, 23–27 June 2019; pp. 1–5.
11. Chen, J.; Chen, J.; Gong, C. New overall power control strategy for variable-speed fixed-pitch wind turbines within the whole wind velocity range. *IEEE Trans. Ind. Electron.* **2012**, *60*, 2652–2660. [[CrossRef](#)]
12. Miao, L.; Wen, J.; Xie, H.; Yue, C.; Lee, W.-J. Coordinated control strategy of wind turbine generator and energy storage equipment for frequency support. *IEEE Trans. Ind. Appl.* **2015**, *51*, 2732–2742. [[CrossRef](#)]
13. Xie, K.; Jiang, Z.; Li, W. Effect of wind speed on wind turbine power converter reliability. *IEEE Trans. Energy Convers.* **2012**, *27*, 96–104. [[CrossRef](#)]
14. Anjun, X.; Hao, X.; Shuju, H.; Honghua, X. Pitch control of large scale wind turbine based on expert PID control. In Proceedings of the 2011 International Conference on Electronics, Communications and Control (ICECC), Ningbo, China, 9–11 September 2011; pp. 3836–3839.
15. Kim, J.-S.; Jeon, J.; Heo, H. Design of adaptive PID for pitch control of large wind turbine generator. In Proceedings of the 2011 10th International Conference on Environment and Electrical Engineering, Rome, Italy, 8–11 May 2011; pp. 1–4.
16. Cheng, X.; Lei, Z.; Junqiu, Y. Fuzzy PID controller for wind turbines. In Proceedings of the 2009 Second International Conference on Intelligent Networks and Intelligent Systems, Tianjin, China, 1–3 November 2009; pp. 74–77.
17. Baburajan, S. Improving the efficiency of a wind turbine system using a fuzzy-pid controller. In Proceedings of the 2018 Advances in Science and Engineering Technology International Conferences (ASET), Dubai, Sharjah, 6 February–5 April 2018; pp. 1–5.
18. Bakka, T.; Karimi, H.-R.; Christiansen, S. Linear parameter-varying modelling and control of an offshore wind turbine with constrained information. *IET Control Theory Appl.* **2014**, *8*, 22–29. [[CrossRef](#)]
19. Boukhezzar, B.; Siguerdidjane, H. Nonlinear control of a variable-speed wind turbine using a two-mass model. *IEEE Trans. Energy Convers.* **2010**, *26*, 149–162. [[CrossRef](#)]
20. Ma, Z.; Yan, Z.; Shaltout, M.L.; Chen, D. Optimal real-time control of wind turbine during partial load operation. *IEEE Trans. Control Syst. Technol.* **2015**, *23*, 2216–2226. [[CrossRef](#)]
21. da Costa, J.P.; Pinheiro, H.; Degner, T.; Arnold, G. Robust controller for DFIGs of grid-connected wind turbines. *IEEE Trans. Ind. Electron.* **2010**, *58*, 4023–4038. [[CrossRef](#)]
22. Beltran, B.; Ahmed-Ali, T.; Benbouzid, M.E. Sliding mode power control of variable-speed wind energy conversion systems. *IEEE Trans. Energy Convers.* **2008**, *23*, 551–558. [[CrossRef](#)]
23. Dang, D.; Wang, Y.; Cai, W. Offset-free predictive control for variable speed wind turbines. *IEEE Trans. Sustain. Energy* **2012**, *4*, 2–10. [[CrossRef](#)]
24. Civelek, Z.; Lüy, M.; Çam, E.; Barışçı, N. Control of pitch angle of wind turbine by fuzzy PID controller. *Intell. Autom. Soft Comput.* **2016**, *22*, 463–471. [[CrossRef](#)]
25. Henriksen, L.C. Model Predictive Control of Wind Turbines. Ph.D. Thesis, Technical University of Denmark, Kgs. Lyngby, Denmark, 2011.

26. Kachroo, P. Existence of solutions to a class of nonlinear convergent chattering-free sliding mode control systems. *IEEE Trans. Autom. Control* **1999**, *44*, 1620–1624. [[CrossRef](#)]
27. Kachroo, P.; Tomizuka, M. Chattering reduction and error convergence in the sliding-mode control of a class of nonlinear systems. *IEEE Trans. Autom. Control* **1996**, *41*, 1063–1068. [[CrossRef](#)]
28. Li, S.; Cao, M.; Li, J.; Cao, J.; Lin, Z. Sensorless-Based Active Disturbance Rejection Control for a Wind Energy Conversion System With Permanent Magnet Synchronous Generator. *IEEE Access* **2019**, *7*, 122663–122674. [[CrossRef](#)]
29. Li, S.; Li, J. Output predictor-based active disturbance rejection control for a wind energy conversion system with PMSG. *IEEE Access* **2017**, *5*, 5205–5214. [[CrossRef](#)]
30. Kourchi, M.; Rachdy, A. Nonlinear ADRC Applied on Wind Turbine Based on DFIG Operating at its Partial Load. In Proceedings of the 2019 International Conference of Computer Science and Renewable Energies (ICCSRE), Agadir, Morocco, 22–24 July 2019; pp. 1–8.
31. Anjun, X.; Xu, L.; Shuju, H.; Nianhong, L.; Honghua, X. A new pitch control method for large scale wind turbine based on ADRC. In Proceedings of the 2013 International Conference on Materials for Renewable Energy and Environment, Chengdoun, China, 19–21 August 2014; pp. 373–376.
32. Huang, C.; Yin, Y. Wind Turbine Pitch Control Based on Error-based ADRC Approach Optimized by Brain Storm Optimization Algorithm. In Proceedings of the 2019 1st International Conference on Industrial Artificial Intelligence (IAI), Shenyang, China, 23–27 July 2019; pp. 1–6.
33. Khooban, M.H.; Soltanpour, R.M. Swarm optimization tuned fuzzy sliding mode control design for a class of nonlinear systems in presence of uncertainties. *J. Intell. Fuzzy Syst.* **2013**, *24*, 383–394. [[CrossRef](#)]
34. Rahimi, A.; Bavafa, F.; Aghababaei, S.; Khooban, M.H.; Naghavi, S.V. The online parameter identification of chaotic behaviour in permanent magnet synchronous motor by self-adaptive learning bat-inspired algorithm. *Int. J. Electr. Power Energy Syst.* **2016**, *78*, 285–291. [[CrossRef](#)]
35. Wiering, M.A.; van Hasselt, H. Ensemble algorithms in reinforcement learning. *IEEE Trans. Syst. Man Cybern. Part B* **2008**, *38*, 930–936. [[CrossRef](#)]
36. Arvind, C.; Senthilnath, J. Autonomous RL: Autonomous Vehicle Obstacle Avoidance in a Dynamic Environment using MLP-SARSA Reinforcement Learning. In Proceedings of the 2019 IEEE 5th International Conference on Mechatronics System and Robots (ICMSR), Singapore, 3–5 May 2019; pp. 120–124.
37. Hu, Y.; Li, W.; Xu, K.; Zahid, T.; Qin, F.; Li, C. Energy management strategy for a hybrid electric vehicle based on deep reinforcement learning. *Appl. Sci.* **2018**, *8*, 187. [[CrossRef](#)]
38. Hasanvand, S.; Rafiei, M.; Gheisarnejad, M.; Khooban, M.-H. Reliable Power Scheduling of an Emission-Free Ship: Multi-Objective Deep Reinforcement Learning. *IEEE Trans. Transp. Electrif.* **2020**. [[CrossRef](#)]
39. Qiu, C.; Hu, Y.; Chen, Y.; Zeng, B. Deep Deterministic Policy Gradient (DDPG)-Based Energy Harvesting Wireless Communications. *IEEE Internet Things J.* **2019**, *6*, 8577–8588. [[CrossRef](#)]
40. Khooban, M.H.; Gheisarnejad, M. A Novel Deep Reinforcement Learning Controller Based Type-II Fuzzy System: Frequency Regulation in Microgrids. *IEEE Trans. Emerg. Top. Comput. Intell.* **2020**. [[CrossRef](#)]
41. Gheisarnejad, M.; Khooban, M.H. An Intelligent Non-integer PID Controller-based Deep Reinforcement Learning: Implementation and Experimental Results. *IEEE Trans. Ind. Electron.* **2020**. [[CrossRef](#)]
42. Rodriguez-Ramos, A.; Sampedro, C.; Bavle, H.; De La Puente, P.; Campoy, P. A deep reinforcement learning strategy for UAV autonomous landing on a moving platform. *J. Intell. Robot. Syst.* **2019**, *93*, 351–366. [[CrossRef](#)]
43. Ren, Y.; Li, L.; Brindley, J.; Jiang, L. Nonlinear PI control for variable pitch wind turbine. *Control Eng. Pract.* **2016**, *50*, 84–94. [[CrossRef](#)]
44. Gu, S.; Lillicrap, T.; Sutskever, I.; Levine, S. Continuous deep q-learning with model-based acceleration. In Proceedings of the International Conference on Machine Learning, New York, NY, USA, 19–24 June 2016; pp. 2829–2838.
45. Zhu, M.; Wang, X.; Wang, Y. Human-like autonomous car-following model with deep reinforcement learning. *Transp. Res. Part C Emerg. Technol.* **2018**, *97*, 348–368. [[CrossRef](#)]
46. Li, F.; Jiang, Q.; Quan, W.; Cai, S.; Song, R.; Li, Y. Manipulation Skill Acquisition for Robotic Assembly Based on Multi-Modal Information Description. *IEEE Access* **2019**, *8*, 6282–6294. [[CrossRef](#)]

47. Khooban, M.H. Hardware-in-the-loop simulation for the analyzing of smart speed control in highly nonlinear hybrid electric vehicle. *Trans. Inst. Meas. Control* **2019**, *41*, 458–467. [[CrossRef](#)]
48. Khooban, M.-H.; Dehghani, M.; Dragičević, T. Hardware-in-the-loop simulation for the testing of smart control in grid-connected solar power generation systems. *Int. J. Comput. Appl. Technol.* **2018**, *58*, 116–128. [[CrossRef](#)]



© 2020 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 (<http://creativecommons.org/licenses/by/4.0/>).