Performance Enhancement of the Artificial Neural Network based Reinforcement Learning for Wind Turbine Yaw Control

Aitor Saenz-Aguirre [1,*], Ekaitz Zulueta [1], Unai Fernandez-Gamiz [2], Alain Ulazia [3] and Daniel Teso-Fz-Betono [1]

[1] Automatic Control and System Engineering Department, University of the Basque Country (UPV/EHU), Nieves Cano, 12, 01006 Vitoria-Gasteiz, Spain; asaenz012@ehu.eus (A.S.); ekaitz.zulueta@ehu.eus (E.Z.); daniel.teso@ehu.eus (D.T.)

[2] Nuclear Engineering and Fluid Mechanics Department, University of the Basque Country (UPV/EHU), Nieves Cano, 12, 01006 Vitoria-Gasteiz, Spain; unai.fernandez@ehu.eus

[3] Nuclear Engineering and Fluid Mechanics Department, University of the Basque Country (UPV/EHU), Avenida Otaola, 29, 20600 Eibar, Spain; alain.ulazia@ehu.eus

* Correspondence: asaenz012@ehu.eus; Tel.: +34-945-01-4066

Received: date; Accepted: date; Published: date

Abstract: The yaw angle control of a wind turbine allows maximization of the power absorbed from the wind and, thus, the increment of the system efficiency. Conventionally, classical control algorithms have been used for the yaw angle control of wind turbines. Nevertheless, in recent years advanced control strategies have been designed and implemented for this purpose. These advanced control strategies are considered to offer improved features in comparison to classical algorithms. In this paper, an advanced yaw control strategy based on Reinforcement Learning (RL) is designed and verified in simulation environment. The proposed RL algorithm considers multivariable states and actions, as well as the mechanical loads due to the yaw rotation of the wind turbine nacelle and rotor. Furthermore, a Particle Swarm Optimization (PSO) and Pareto optimal Front (PoF) based algorithm has been developed in order to find the optimal actions that satisfy the compromise between the power gain and the mechanical loads due to the yaw rotation. Maximizing the power generation and minimizing the mechanical loads in the yaw bearings in an automatic way are the objectives of the proposed RL algorithm. The data of the matrices Q(s,a) of the RL algorithm are stored as continuous functions in an Artificial Neural Network (ANN) avoiding any quantification problem. The NREL 5MW reference wind turbine has been considered for the analysis and real wind data from Salt Lake, Utah, USA have been used for the validation of the designed yaw control strategy via simulations with the aeroelastic code FAST.

Keywords: wind turbine control; yaw control; reinforcement learning; artificial neural network; optimization; Pareto front.

Acronyms and Symbols

The following acronyms and symbols are used in this manuscript:

RL	Reinforcement Learning	ANN	Artificial Neural Network
PSO	Particle Swarm Optimization	PoF	Pareto optimal Front
PID	Proportional Integral Derivative	PI	Proportional Integral
MLP-BP	MultiLayer Perceptron with BackPropagation		
FAST	Fatigue, Aerodynamics, Structure and Turbulence		
FPGA	Field Programmable Gate Array		

1. Introduction

The promotion of the renewable energies has nowadays emerged as a major necessity in order to overcome the problems associated with the combustion of conventional fossil fuels. In this context, as presented in the work Nehrir et al. [1], extensive research has been conducted with the objective of discovering alternative sustainable energy resources. Additionally, many efforts are directed to the technological development and efficiency enhancement of the existing renewable energy generation systems [2-4].

One of the fields on the focus is the improvement of the control system of the sustainable energy generation systems. The design of an adequate control strategy enables maximization of the power generated by the system, and thus, its efficiency. As introduced by Njiri et al. [5], the principal objectives of the control system implemented in a wind turbine are to guarantee the safety of the workers and the turbine and to maximize its power output.

From a control design perspective, conventionally algorithms based on classical PIDs (Proportional, Integral, Derivative) or PIDs with slight variations have been implemented in industrial wind turbines. In the work of Habibi et al. [6] an adaptive PID strategy is designed for the output power regulation of a wind turbine. A Fuzzy logic based PI (Proportional Integral) controller to optimize the power generation of a wind turbine is presented by Aissaoui et al. [7].

Nowadays, with the objective of providing control systems of wind turbines with additional features, advanced control strategies are being introduced. The application of different advanced control strategies to the operation and grid connection of wind turbines has been found in the literature. Kim [8] presents a data driven robust $H\infty$ controller which is aimed to improve the operation of a wind turbine. A non-linear control strategy for variable speed wind turbines based on Fuzzy Logic is proposed in the work of Liu et al. [9]. Jafarnejadsani et al. [10] present in their work a gain scheduled optimal control of a wind turbine. An advanced control strategy for the generator of a wind turbine based on Sliding Mode Control is introduced by Merabet et al. [11] and Evangelista et al. [12]. A novel multifrequency power oscillations mitigation algorithms to improve the grid connection of the wind turbine is proposed in the work of Moriano et al. [13].

Regarding the yaw operation of a wind turbine, different studies such as the one presented by Gebraad et al. [14], have been introduced in the literature to optimize the power production of a wind farm by calculation of the optimal yaw angle for individual wind turbines. Munters et al. [15] present a gradient-descent based algorithm for the calculation of this optimal yaw angle. Dar et al. [16] present an optimization technique for the yaw angle of individual wind turbines in a wind farm through a dynamic programming formulation.

Additionally, some advanced control strategies applied to enhance the operation of the yaw system of individual wind turbines have been found in the literature. Song et al. [17] present two variations of a predictive control strategy applied to the yaw operation of the wind turbine. The use of estimators to anticipate the wind direction is shown to improve the performance of the classical yaw control methods. Saenz-Aguirre et al. [18] present an Artificial Neural Network (ANN) based Reinforcement Learning (RL) control strategy for the yaw control of a wind turbine. According to the work of Saenz-Aguirre et al. [18], a RL based yaw control algorithm shows important advantages in comparison to the conventional PID based yaw control methods [19,20], especially in form of lack of control parameters tuning necessity and a fully automatic performance, due to the self-learning process.

The two principal RL algorithms, SARSA [21] and Q-Learning [22], are introduced in the work of Liu et al. [23]. The RL algorithms are based on the knowledge acquired by a system via its interaction with the environment. For that purposed, a $Q : S\,x\,A \rightarrow R$ function is defined by the RL algorithm. In this function, S refers to the range of considered states of the system, A is the available set of different actions that can be taken in a given state and R refers to the reward obtained by the system if the action a is taken in a state s. The definition of this matrix Q is the main difference between the RL algorithms SARSA and Q-Learning. While Q-Learning considers quantified states and actions in a quantified matrix $Q(s,a)$, SARSA considers the matrix Q as a continuous function $Q(s(t),a(t))$ calculated from an initial time to an horizon time value. The main characteristic of both RL algorithms is the fully automatic performance that is achieved after a training process covering the whole range of states and actions considered for the system.

The use of ANNs have also been introduced in the field of the renewable energies, especially with the objective of obtaining data driven models. Lopez-Guede et al. [24] present an ANN based modelling technique of photovoltaic modules. A modelling of the wind power output, the vibration of the drive train and the vibration of the tower of a wind turbine using ANNs is introduced in the work of Kusiak et al. [25]. Although the use of ANNs as controllers is not generalized, some examples of ANNs in the control system of a wind turbine have been found in the literature. Shi et al. [26] present in their work a neural network based power coefficient compensation to optimize the power production of a wind turbine. Li et al. [27] introduce the process of the digital implementation of an ANN in a Field Programmable Gate Array (FPGA) to be implemented in a wind turbine and optimize its operation.

In this paper, an improved version of the ANN based RL yaw control strategy introduced by Saenz-Aguirre et al. [18] is developed. The performance enhancement of the ANN based RL yaw control strategy is aimed to improve the operation of the yaw control system for wind speed values over the rated value and to reduce the mechanical moments in the yaw system bearings. To that end, the error and action variables of the RL algorithm have been converted into multivariate variables and the mechanical loads due to the correction of the yaw angle during operation of the wind turbine have been considered as a reward value and incorporated to the calculation of the matrices Q(s,a). Another innovative element introduced in the ANN based RL algorithms presented in this paper is the implementation of a Particle Swarm Optimization (PSO) algorithm [28,29] and a Pareto optimal Front (PoF) [30,31] based optimization algorithm in order to calculate the optimal actions that maximize the power gain as a result of the yaw correction and minimize the mechanical loads induced in the yaw system bearings due to it.

A MultiLayer Perceptron with BackPropagation (MLP-BP) neural network is designed in this paper to store the matrices $Q(s,a)$ correspondent to the RL algorithm as continuous functions and avoid quantification problems. Furthermore, the use of a MLP-BP is expected to erase the needs for management of large amounts of data during operation of the wind turbine. The NREL 5MW reference wind turbine, introduced by Jonkman et al. [32], has been considered for the analysis presented in this paper. The simulations for the validation of the designed ANN based RL yaw control strategy have been carried out with the aeroelastic code FAST [33], widely-used for the analysis of the performance and mechanical loads during operation of wind turbines. As shown in the works of Rahimi et al. [34-36], the skewed wake model implemented in the aeroelastic code is observed to affect the calculation of the mechanical loads in the wind turbine in cases of yaw misalignment. All the simulation results presented in this paper are based on the Pitt and Peters[37] skewed wake model. Real wind data from Salt Lake, Utah, USA have been introduced as input of the aeroelastic code FAST [33] and used for the validation process.

In comparison to conventional control algorithms, the ANN based RL yaw control strategy presented in this paper is considered to offer the same advantages as the strategy presented in the work of Saenz-Aguirre et al. [18], i.e., online learning capability (during operation of the wind turbine), fully autonomous performance and lack of design of a controller. However, the additional features introduced to the strategy are supposed to improve the performance of the system, especially in form of reduction of the mechanical loads in the yaw system components, which in absence of an adequate control strategy could become too high and endanger the safe operation of the wind turbine.

The paper is structured as follows: The main characteristics of the NREL 5MW wind turbine and the method for the calculation of the power generated by the wind turbine and the mechanical loads in the yaw bearing are presented in Section 2. In Section 3 the structure of the ANN based RL yaw control algorithm proposed in this paper is given. The synthesis and design process of the ANN based RL yaw control strategy is shown in Section 4. Finally, Sections 5 and 6 correspond to the validation results and the conclusions, respectively.

2. Wind Turbine Characterization

The NREL 5MW wind turbine, introduced by Jonkman et al. [32] and considered as the reference wind turbine for many offshore applications, has been adopted for the simulations presented in this document. The main features of the NREL 5MW wind turbine are presented in Table 1.

The yaw control strategy presented in this document is based on the selection of the optimal control action that maximizes the power generated by the wind turbine while minimizing the mechanical moments induced in the bearings of the yaw system. Consequently, for the development of the proposed yaw control strategy the characterization of the generated power and the z axis mechanical moment in the yaw bearings for different operating points of the wind turbine is of great importance. The mechanical moment with respect to the z axis has been found to be the most critical load in cases of yaw rotations.

As it can be observed in Figure 1 (a), the power generated by the wind turbine is fully defined and can be easily calculated with the power curve of the wind turbine and setting the wind speed and the yaw angle (misalignment angle between the incident wind and the orientation of the wind turbine rotor) as inputs of the power curve. Hence, the power curve of the NREL 5MW wind turbine has been stored in a 2-D Look-up Table for further access during the training process of the RL algorithm.

Regarding the mechanical moment with respect to the z-axis induced in the yaw bearings, more than two variables are necessary to estimate its value. Figure 1 (b) shows the variables involved in an accurate estimation of this magnitude.

As it can be observed in Figure 1 (b), the z axis mechanical moment in the yaw bearings has been approximated to be a function of some external factors, such as the wind speed and the yaw angle, and the control action taken by the yaw control system of the wind turbine (YawRateK [-] refers to the yaw rotation speed factor and YawToMove [deg] refers to the duration of the yaw rotation). The larger the values of the YawRateK [-] and the YawToMove [deg] parameters are, the higher the z axis mechanical moment induced in the yaw bearings is. The mechanical moment in the yaw bearings is also known to be affected by other external factors such as the skewed wake model employed in the aeroelastic simulation [34-36].

Since the estimation of the z axis mechanical moment in the yaw bearings has been observed to depend on the control action of the yaw control system, several simulations covering all possible scenarios considered in this analysis have been performed with the aeroelastic code FAST. As it is shown in Figure 2, wind direction variations of -90:10:90 deg with a constant wind speed value have been considered. These scenarios has been repeated for constant winds from 3:2:25 m/s. Finally, each wind scenario has been simulated for the whole range of the control action YawRateK=0:0.1:1 considered in the analysis and the values of the sum of the z axis mechanical moment induced in the yaw bearings for each action YawRateK [-] and YawToMove [deg] has been calculated.

The results of the simulations have been stored in a 4-D Look-up Table for further access during the training process of the RL algorithm.

3. Structure of the proposed Yaw Control Strategy

The yaw control strategy introduced in this paper is based on the ANN based RL yaw control strategy presented in the work of Saenz-Aguirre et al. [18] and it is intended to improve its performance by the introduction of additional features.

The yaw control strategy introduced in this paper considers new state and actions that define the operation of the wind turbine more accurately. Furthermore, the mechanical moment with respect to the z axis induced in the yaw bearings as a result of the yaw rotation has been considered as a reward variable of the RL algorithm. Finally, a PSO and PoF based optimization algorithm has been designed to respond to the necessity of compromise between the power increment with a severe yaw control action and the mechanical costs associated to it.

This section is divided in 3 subsections: An extended explanation of the ANN based RL algorithm is presented in Subsections 3.1. The introduction of the PSO and PoF based optimization algorithm in the yaw control system of a wind turbine is given in Subsection 3.2. Finally, Subsection 3.3 presents the Decision Making algorithm associated to the selection of one of the possible optimal actions presented by the PoF.

3.1 Artificial Neural Network based Reinforcement Learning algorithm

A multivariate RL algorithm (two states, two actions and two immediate reward variables are considered) is proposed in this document. The objective of considering an extended RL algorithm, in comparison to the simple RL algorithm considered in the work Saenz-Aguirre et al. [18], is to provide an improved characterization of the states, actions and rewards of the RL algorithm associated to the yaw control system of the wind turbine. The following state, action and reward variables have been considered in the RL algorithm proposed in this paper.

- Two states *s* are considered:
    - YawAngle [deg]: Represents the misalignment angle between the orientation of the rotor of the wind turbine and the direction of the incoming wind, as shown in Eq. (1).

$$\theta_{yaw} = \theta_{wind} - \theta_{nacelle} \tag{1}$$

    - WindSpeed [m/s]: Determines the operating point of the wind turbine. As a result of the control system implemented in wind turbines, the power loss due to yaw misalignments is not equal for every wind speed value. The consideration of the wind speed as a state enables the particularization of the effect of each yaw angle to a determined operating point of the wind turbine.

- Two actions *a* are considered:
    - YawRateK [-]: Represents the gain associated to the yaw rotational speed of the wind turbine, as shown in Eq. (2).

$$\Omega_{yaw} = YawRateK \cdot \theta_{yaw} \tag{2}$$

- YawToMove [deg]: Limits the rotation of the rotor of the wind turbine to a certain value, as described in Eq. (3).

$$\Delta\theta_{yaw} \in [-YawToMove, YawToMove] \quad (3)$$

- Two immediate reward variables *r* are considered:
  - PowerGain [%]: Indicates the power gain the wind turbine could achieve by performing a concrete action (YawRateK [-] and YawToMove [deg]) in a defined state (YawAngle [deg] and WindSpeed [m/s]). The expression to calculate PowerGain [%] is given in Eq. (4).

$$PowerGain = \frac{P\_control - P\_no\_control}{P\_no\_deviation} \cdot 100 \quad (4)$$

where, as it is described in the work of Saenz-Aguirre et al. [18], *P_control* [W] refers to the power generation of the wind turbine when the yaw control system is activated, *P_no_control* [W] refers to the power generation of the wind turbine when the yaw control is not activated and *P_no_deviation* [W] refers to the power generation of the wind turbine if the yaw angle was zero.

  - YawMoment [N·m]: Indicates the value of the sum of the mechanical moment with respect to the z axis induced in the yaw system bearing by performing a concrete action (YawRateK [-] and YawToMove [deg]) in a defined state (Yaw angle [deg] and Wind speed [m/s]).

The mathematical procedure to calculate the function *Q(s,a)* corresponding to both reward variables is given by an exponential moving average from the instant in which the action is taken to a predefined time horizon, in this case set to T=60 s., as described in Eq. (5). Since no difference of importance between the responses of the system until the end of the time horizon is considered, the discount factor γ is set to 1.

$$Q(s,a) = \sum_{i=0}^{i=T} r_{t+i} \cdot \gamma^{i} \quad (5)$$

The expressions of the matrix *Q(s,a)* for each one of the considered reward variables are presented in Eq. (6) and Eq. (7) respectively.

$$Q\_P(s,a) = \frac{\frac{1}{T}\int_{t}^{t+T}(P\_control(t) - P\_no\_control(t))\cdot dt}{\frac{1}{T}\int_{t}^{t+T} P\_no\_deviation\cdot dt} \cdot 100 \qquad [\%] \quad (6)$$

$$Q\_M(s,a) = \int_{t}^{t+T} YawMoment\ (t) \cdot dt \qquad [N \cdot m] \quad (7)$$

Once the matrices *Q_P(s,a)* and *Q_M(s,a)* have been calculated, they are stored in an ANN with the objective of avoiding quantification problems and eliminating the need of management of big amounts of data during operation of the wind turbine. In addition, the use of an ANN to store the matrices *Q_P(s,a)* and *Q_M(s,a)* as continuous functions *Q_P(s(t),a(t))* and *Q_M(s(t),a(t))* allows simple estimation of the non-simulated scenarios. Finally, due to the use of an ANN, the refreshment policy of the RL algorithm is incorporated in the training process of the ANN. The inputs and outputs of the designed ANN are presented in Figure 3.

If the effect of the state variables *s* and the actions *a* on both *Q_P(s(t),a(t))* and *Q_M(s(t),a(t))* matrices is studied in detail, a necessity for compromise in the selection of the optimal action can be observed. On the one hand, the bigger the value of the YawRateK [-] and YawToMove [deg] is, the larger the power gain of the wind turbine will be. Nevertheless, the higher the value of the YawRateK [-] and YawToMove [deg] is, the larger the z axis mechanical moment in the yaw bearings will be.

Due to the existence of two *Q(s(t),a(t))* functions, the output values of which are subjected to a compromise, a PSO and PoF based optimization algorithm is proposed to calculate the optimal set of solutions for the yaw control system of the wind turbine.

3.2 Particle Swarm Optimization and Pareto optimal Front

The objective of the PSO and PoF based optimization algorithm is to give response to the compromise situation explained in Subsection 3.1 and to calculate the optimal set of combinations of the RL actions YawRateK [-] and YawToMove [deg]. A maximum power generated by the wind turbine and a minimum z axis mechanical moment in the yaw bearing are desired.

According to Ho et al. [28], one characteristic aspect of the PSO algorithm is that it works with a potential group of solutions instead of a unique solution. Moreover, instead of evolutionary aspects to generate new generations of populations, in PSO there is a parameter space in which the particles move according to their own experience and the experience of the other particles. As a result, each particle moves toward a weighted average of its own maximum and the maximum of the rest of the particles. Nevertheless, as introduced in the work of Ehrgott et al. [38], in a multiobjective optimization problem there are solutions in which one of the optimization objectives is not fulfilled. This set of solutions is called non-dominated and form the PoF.

In the analysis presented in this document, the states of the system are defined as YawAngle [deg] and WindSpeed [m/s]. Hence, for a given known state of the system, the PSO and PoF based algorithm should find the set of optimal actions (YawRateK [-] and YawToMove [deg]) that maximize the power generated by the wind turbine and minimize the z axis mechanical load in the yaw bearing. As it was previously explained, an ANN has been trained to store the data of both matrices *Q_P(s,a)* and *Q_M(s,a)* as continuous functions *Q_P(s(t),a(t))* and *Q_M(s(t),a(t))*. Hence, the PSO and PoF based optimization algorithm will access the ANN to optimize its output values. At the end of the optimization process, a PoF with a set of 20 optimal solutions for both *Q_P(s(t),a(t))* and *Q_M(s(t),a(t))* functions is obtained. Each solution correspond to the optimal solution of a 5% wide window of the whole range considered for the output of the *Q_P(s(t),a(t))* function, i.e. [0, 100] %.

A pseudocode with the principal aspects of the PSO and PoF based optimization algorithm is presented in Figure 4.

3.3 Decision Making Algorithm

After calculation of the PoF, a Decision Making algorithm is designed to select one action (YawRateK [-] and YawToMove [deg]) from the set of optimal actions proposed by the PSO-PoF algorithm for a given known state of the system (YawAngle [deg] and WindSpeed [m/s]).

The Decision Making algorithm proposed in this paper is based on two concepts:

- All the solutions, in which the output of the function *Q_M(s(t),a(t))* is bigger than a predefined upper threshold value, are discarded.
- From the rest of the solutions, the one with the biggest output value of the function *Q_P(s(t),a(t))* is selected.

4 Design of the proposed Yaw Control Strategy

The design process of the ANN based RL yaw control strategy of the wind turbine has been carried out with a simplified model of the yaw control system, see Figure 5. This simplified model is very similar to the model introduced in the work of Saenz-Aguirre et al. [18]. However, a limiter to include the actuation of the RL action YawToMove [deg] has been introduced. The successful completion of the design process with the simplified model would prove its validity and enables its use for the online training of the system during operation of the wind turbine.

Another important aspect of the training process of the RL algorithm is that the operation of the yaw control system of the wind turbine is prevented to actuate when the possible power gain as a result of the yaw rotation is not significant. To that purpose, a parameter named DeserveMove [%] has been created for the training process. If the output of the function *Q_P(s(t),a(t))* is smaller than the parameter DeserveMove [%], the value of the corresponding RL actions YawRateK [-] and YawToMove [deg] are directly set to 0.

The states YawAngle [deg] and WindSpeed [m/s] of the RL algorithm are based on 60 s filtered measurements of the wind direction and wind speed, respectively. The time constant of the filter is related to the time horizon selected for the RL algorithm, which is an adaptable parameter subject to any kind of restriction associated to the yaw system of the wind turbine or its control execution management. The objective of the filtering step is to reduce the possible affection of sudden and short-term wind gusts or failed measurements on the RL algorithm.

As it was explained in Section 2 and shown in Figure 2, simulations with constant wind values and the whole range of considered yaw angle values have been performed for the training process of the RL algorithm. Furthermore, these simulations are repeated for the whole range of values of the control actions considered in the analysis. The objective is to train the system with cases correspondent to the whole operating range of the wind turbine.

In this case, the training process of the RL algorithm has been performed offline, i.e., not during operation of the wind turbine, and considering all possible winds and yaw control actions, so an adequate response of the system for the whole range of possible scenarios is achieved. Furthermore, an online training process linked to the actual operational conditions of the wind turbine could be implemented to keep the system learning during its operation.

Similarly to how it is done in the work of Saenz-Aguirre et al. [18], once the training process is finished and the matrices *Q_P(s,a)* and *Q_M(s,a)* have been obtained, a MultiLayer Perceptron with BackPropagation neural network is trained to store the data correspondent to the matrices as a continuous functions *Q_P(s(t),a(t))* and *Q_M(s(t),a(t)*.

The list of the parameters considered for the design process of the ANN based RL yaw control is presented in Table 2.

The training process of the MLP-BP has been completed with a correlation coefficient of 0.9999 and a Mean Squared Error (MSE) value of $1.62 \cdot 10^{-6}$. Both values indicate that the training process of the MLP-BP has been successful and the neural network is accurate enough to adequately store the data of the matrices *Q_P(s,a)* and *Q_M(s,a)*.

In order to prove the correct training process of the RL algorithm a comparison between the output values of the matrices *Q_P(s,a)* and *Q_M(s,a)* and the functions *Q_P(s(t),a(t))* and *Q_M(s(t),a(t)* for three different set of RL actions (YawRateK = 0.5 and YawToMove =70 deg, YawRateK = 0.5 and YawToMove = 30 deg and YawRateK = 0.1 and YawToMove =70 deg) is presented in Figure 6 and Figure 7. For each one of the three RL action cases two different WindSpeed [m/s] states are defined: WindSpeed = 11 m/s and WindSpeed = 21 m/s.

As it can be observed in Figure 6, the value of the power gain that can be achieved with the yaw control depends on 4 different factors. First, as a result of the control system implemented in the wind turbines, the value of the power gain depends on the wind speed value. As it can be seen in Figure 6 (a), the amount of power that can be gained in a state YawAngle = 50 deg and WindSpeed = 11 m/s is around 60%. However, in Figure 6 (b), in case of YawAngle = 50 deg and WindSpeed = 21 m/s the possible power gain is 0%. This is due to the fact that despite the misalignment of 50 deg, the wind turbine operates in the rated power zone and there is no loss of power.

The YawAngle [deg] also affects the power gain that can be achieved with the yaw control, since the bigger the YawAngle [deg] is, the bigger the power loss is, unless the YawAngle [deg] value is not large enough to make the wind turbine operate outside the rated power zone. Finally, the two other factors that influence the power gain are the RL actions YawRateK [-] and YawToMove [deg]. As it can be seen in Figure 6 (a), if the YawToMove [deg] action limits the rotation capability of the wind turbine, the power gain that can be extracted in high YawAngle [deg] values is severely decreased The effect of the YawRateK [-] is as well clearly observable in Figure 6 (a), where the power gain achieved by the wind turbine has been reduced with the reduction of the YawRateK [-].

As it can be observed in Figure 7, the value of the z axis mechanical moment in the yaw bearing does not vary significantly with the WindSpeed [m/s] but it does with the YawAngle [deg] state and the YawRateK [-] and the YawToMove [deg] actions. The larger the value of the YawAngle [deg], the YawRateK [-] and the YawToMove [deg] are, the higher the value of the z axis mechanical moment in the yaw bearing is.

Both Figure 6 and Figure 7 shown that the training process of the MLP-BP neural network has been successful since there is a complete correspondence between the values of the matrices *Q_P(s,a)* and *Q_M(s,a)* and the values of the functions *Q_P(s(t),a(t))* and *Q_M(s(t),a(t))* modelled by the MLP-BP.

Once the training process of the RL algorithm is finished, the PSO and PoF based optimization algorithm must be designed. The objective of the PSO and PoF based algorithm is to find the set of optimal actions (YawRateK [-] and YawToMove [deg]) that maximize the power generated by the wind turbine and minimize the z axis mechanical moment in the yaw bearing for a given known state of the system (YawAngle [deg] and WindSpeed [m/s]).

A list of the parameters considered for the design process of the PSO and PoF based optimization algorithm is presented in Table 3.

To ensure the correct performance of the designed PSO and PoF optimization algorithm, the intermediate solutions of the algorithm for two different states of the RL algorithm are presented in Figure 8. The RL states are defined as YawAngle = 90 deg and WindSpeed =11 m/s in Figure 8 (a), and YawAngle = 30 deg and WindSpeed =11 m/s in Figure 8 (b).

Furthermore, the output result of the PSO and PoF based optimization algorithm for both cases is presented in Figure 9.

As it can be observed in Figure 8 and Figure 9, the PoFs represented in Figure 9 correspond to the highlighted zone of the Figure 8, which indicates that the performance of the optimization algorithm is correct. If the PoF is analyzed in detail it is to be observed the compromise between the power gain of the system and the z axis mechanical moment in the yaw bearings. In the case of YawAngle=30 deg in Figure 8 (b) the maximum output of the function *Q_P(s(t),a(t))* is seen to be smaller than 20%, which is concordance with the results obtained in Figure 6 (a).

Finally, the parameters defined for the Decision Making process are presented in Table 4.

The selection of the optimal solution with respect to the Decision Making process parameter defined in Table 4 is shown in Figure 10.

The ANN based RL yaw control strategy introduced in this document has been verified with the aeroelastic code FAST using with real wind speed data [39] from a meteorological station located in Salt Lake, Utah, USA. The location of the meteorological station is defined with the following geographical coordinates [-112.0621°, 40.5938°] and it is formed by seven measuring stations containing ultrasonic anemometers capable of recording data at sampling rates higher than 1 Hz. The measurements have been afterwards averaged to 1 second rates. The collection of the data was carried out from November 10, 2010 to February 2, 2011.

The meteorological station is situated at a height of 10 m and the data have been transformed to the hub height of the NREL 5MW wind turbine, i.e., 90 m. Due to the adequate location of the meteorological station in a flat terrain, without obstacles for the wind, the logarithmic law have been used for this transformation.

The use of real wind speed data in the FAST simulation environment is important since it allows a detailed analysis of the performance of the designed control strategy in a realistic scenario. In fact, the data collected from this meteorological station has been used in several publications [40,41].

A detailed analysis of the available wind data has been conducted and 6 different cases with a variety of stable wind conditions have been identified and isolated to be used in the study of the performance of the proposed yaw control system. One example case where the wind conditions remain rather stable during a time span of 10000 s is represented in Figure 11.

Once the wind cases have been identified, the operation of the proposed ANN based RL yaw control system is verified with the aeroelastic code FAST. First, the 60 s filtered wind direction and wind speed values correspond to the states YawAngle [deg] and WindSpeed [m/s] of the RL algorithm. When the states are known, the PSO-PoF optimization algorithm is executed and the optimal front of Pareto is obtained. Once this is obtained, one of the solutions is selected with the Decision Making algorithm, and the optimal actions YawRateK [-] and YawToMove [deg] are calculated. Finally, at the end of the simulation the power gain and the z axis mechanical moment in the yaw bearing are analyzed. The values of the states and actions of the RL algorithm and the power gain and the z axis mechanical moment for the considered 6 scenarios are listed in Table 5.

As it can be observed in Table 5, the power increment that can be achieved with the yaw control is dependent on both, the YawAngle [deg] and the WindSpeed [m/s]. For instance, the power increment in case 6 is larger than in case 3, being the value of the WindSpeed [m/s] smaller. However, a similar YawAngle [deg] in case 4 and case 5 does not cause a larger power increment. This is due to the fact that the WindSpeed [m/s] is high enough to keep the wind turbine operating at the rated power zone, so the effect of the yaw misalignment is not significant. Regarding the z axis mechanical moment in the yaw bearings, it is to be noted that its value is above all dependent on the duration of the yaw movement, which is related to the YawAngle [deg] state and the YawToMove [deg] action. In this way, it is to be observed that the biggest values of the z axis mechanical moment correspond to the longest yaw rotations.

The results obtained with the ANN based RL yaw control algorithm presented is this document have also been compared to the results obtained with yaw control algorithm in the work of Saenz-Aguirre et al. [18] for the same input wind conditions. The objective is to characterize the improvements achieved with the enhancement of the control strategy. To that end, the values of the action YawRateK [-] of the RL algorithm that would be obtained with the yaw control algorithm in the work of Saenz-Aguirre et al. [18] are listed in Table 6.

In Table 6 it is to be seen that the values of the action YawRateK [-] are considerably higher than in Table 5, which should translate in higher z axis mechanical moments in the yaw bearings. The RL action YawToMove [deg] is not included in Table 6 because this action was not considered in the work of Saenz-Aguirre et al. [18], so it is considered to have a value of 90 deg.

The values of the power gain and the z axis mechanical moment in the yaw bearings for each one of the analyzed yaw control strategies and a comparison between these values are presented in Table 7. The columns 2 and 3 correspond to the results obtained for the ANN based RL yaw control strategy presented in this paper. The columns 4 and 5 correspond to the results obtained for the yaw control strategy presented by Saenz-Aguirre et al. [18].

The results presented in Table 7, show a considerable performance improvement of the ANN based RL yaw control strategy presented in this document with respect to the yaw control algorithm presented in the work of Saenz-Aguirre et al. [18]. As it can be seen in the power increment and z axis mechanical moment comparison, the values of the mechanical moments are drastically reduced, while the value of the power gain has been kept similar. The reason for that is that the designed PSO-PoF algorithm calculates the best actions to limit the z axis mechanical moment in the yaw bearings and maximize the power gain. The absence of such an algorithm in the strategy in the work of Saenz-Aguirre et al. [18] causes the system to operate with the only objective of maximizing the power and not caring about the mechanical loads.

Especially remarkable is the performance of the proposed RL based yaw control algorithm in the cases 4 and 5 presented in Table 7. In these cases, in which the wind turbine is operating in its rated power zone, no power gain could be achieved with the yaw correction, but high mechanical loads will occur as a result of it. With the proposed yaw control algorithm significant moment reductions with no power loss are achieved in these both cases, which is translated in a longer lifetime of the wind turbine.

5 Conclusions

An enhanced performance of the ANN based RL yaw control strategy is presented and verified in this document. The proposed yaw control algorithm has been observed to drastically reduce the mechanical moments in the components of the yaw system while keeping similar values of the power gain in comparison to similar strategies previously found in the literature.

The extension of the RL algorithm by considering new states and actions and the execution of the PSO-PoF optimization algorithm allow the calculation of a set of optimal solutions from which a desired one can be selected in every case.

In comparison to conventional yaw control strategies, the ANN based RL yaw control strategy introduced in this document is designed to achieve a completely automatic operation of the yaw system after the training process of the RL algorithm. The off-line training proposed in this paper, based on data obtained from simulations in FAST, tries to cover all the possible scenarios in the operation of the wind turbine. However, one important aspect of the ANN based RL yaw control strategy presented in this document is that the training process of the RL algorithm could be performed on-line during the operation of the wind turbine and feed the system with real-time data.

Moreover, the yaw control strategy introduced in this document eliminates the need for tuning the controller of the yaw control system of the wind turbine and, hence, erases the risk of an inadequate tuning and possible damages to the wind turbine components. Furthermore, the possibility to select one optimal solution from a set of optimal solutions enables the wind turbine operator to adequate the operation of the wind turbine to its condition or the need of energy production.

Finally, the validation of the proposed ANN based RL yaw control strategy with the aeroelastic code FAST and using real wind speed data gives certainty about its correct operation and its applicability in real wind generation systems.

Funding: The Regional Development Agency of the Basque Country (SPRI) is gratefully acknowledged for economic support through the research project AIRJET, grant number: KK-2018/00109. This research was partially funded by Fundation VITAL Fundazioa (FP18/36).

Acknowledgments: The authors are grateful to the Government of the Basque Country and the University of the Basque Country UPV/EHU through the SAIOTEK (S-PE11UN112) and EHU12/26 research programs, respectively.

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

References

1. Nehrir MH, Wang C, Strunz K, et al. A review of hybrid renewable/alternative energy systems for electric power generation: Configurations, control, and applications. *IEEE Transactions on Sustainable Energy*. 2011;2(4):392-403. doi: 10.1109/TSTE.2011.2157540.

2. Astolfi D, Castellani F, Terzi L. Wind turbine power curve upgrades. *Energies*. May 2018;11(5):1300.

3. Terzi L, Lombardi A, Castellani F, Astolfi D. Innovative methods for wind turbine power curve upgrade assessment. *J Phys , Conf Ser (UK)*. 2018;1102:012036 (11 pp.); 012036(11.)-012036 (11 pp.). doi: 10.1088/1742-6596/1102/1/012036.

4. Massardo AF, McDonald CF, Korakianitis T. Microturbine/fuel-cell coupling for high-efficiency electrical-power generation. *Journal of Engineering for Gas Turbines and Power*. 2000;124(1):110-116. http://dx.doi.org/10.1115/1.1398552.

5. Njiri JG, Söffker D. State-of-the-art in wind turbine control: Trends and challenges. *Renewable and Sustainable Energy Reviews*. 2016;60:377-393.

6. Habibi H, Nohooji HR, Howard I. Adaptive PID control of wind turbines for power regulation with unknown control direction and actuator faults. *IEEE Access*. 2018;6:37464-37479. doi: 10.1109/ACCESS.2018.2853090.

7. Aissaoui AG, Tahour A, Essounbouli N, Nollet F, Abid M, Chergui MI. A fuzzy-PI control to extract an optimal power from wind turbine. *Energy Conversion and Management*. 2013;65:688-696.

8. Kim Y-. Robust data driven H-infinity control for wind turbine. *Journal of the Franklin Institute*. 2016;353(13):3104-3117.

9. Liu J, Gao Y, Geng S, Wu L. Nonlinear control of variable speed wind turbines via fuzzy techniques. *IEEE Access*. 2017;5:27-34. doi: 10.1109/ACCESS.2016.2599542.

10. Jafarnejadsani H, Pieper J, Ehlers J. Adaptive control of a variable-speed variable-pitch wind turbine using radial-basis function neural network. *IEEE Trans Control Syst Technol*. 2013;21(6):2264-2272. doi: 10.1109/TCST.2012.2237518.

11. Merabet A, Ahmed KT, Ibrahim H, Beguenane R. Implementation of sliding mode control system for generator and grid sides control of wind energy conversion system. *IEEE Transactions on Sustainable Energy*. 2016;7(3):1327-1335. doi: 10.1109/TSTE.2016.2537646.

12. Evangelista CA, Pisano A, Puleston P, Usai E. Receding horizon adaptive second-order sliding mode control for doubly-fed induction generator based wind turbine. *IEEE Trans Control Syst Technol*. 2017;25(1):73-84. doi: 10.1109/TCST.2016.2540539.

13. Moriano J, Rizo M, Bueno EJ, Martin R, Rodriguez FJ. A novel multifrequency current reference calculation to mitigate active power fluctuations. *IEEE Trans Ind Electron*. 2018;65(1):810-818. doi: 10.1109/TIE.2017.2686319.

14. Gebraad P, Teeuwisse F, Wingerden J, et al. Wind plant power optimization through yaw control using a parametric model for wake effects—a CFD simulation study. *Wind Energy*. 22 December 2014.

15. Munters W, Meyers J. Dynamic strategies for yaw and induction control of wind farms based on large-eddy simulation and optimization. *Energies*. 2018;11(1):177. doi: 10.3390/en11010177.

16. Dar Z, Kar K, Sahni O, Chow JH. Windfarm power optimization using yaw angle control. *IEEE Transactions on Sustainable Energy*. 2017;8(1):104-116. doi: 10.1109/TSTE.2016.2585883.

17. Song D, Yang J, Fan X, et al. Maximum power extraction for wind turbines through a novel yaw control solution using predicted wind directions. *Energy Conversion and Management*. 2018;157:587-599. doi: 10.1016/j.enconman.2017.12.019.

18. Saenz-Aguirre A, Zulueta E, Fernandez-Gamiz U, Lozano J, Lopez-Guede JM. Artificial neural network based reinforcement learning for wind turbine yaw control. *Energies*. Jan 2019.

19. Karakasis N, Mesemanolis A, Nalmpantis T, Mademlis C. Active yaw control in a horizontal axis wind system without requiring wind direction measurement. *IET Renewable Power Generation*. 2016;10(9):1441-1449. doi: 10.1049/iet-rpg.2016.0005.

20. Shariatpanah H, Fadaeinedjad R, Rashidinejad M. A new model for PMSG-based wind turbine with yaw control. *IEEE Trans Energy Convers*. 2013;28(4):929-937. doi: 10.1109/TEC.2013.2281814.

21. Adam S, Busoniu L, Babuska R. Experience replay for real-time reinforcement learning control. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*. 2012;42(2):201-212. doi: 10.1109/TSMCC.2011.2106494.

22. Watkins CJCH, Dayan P. Q-learning. *Mach Learning*. 1992;8(3):279-292. https://doi.org/10.1007/BF00992698. doi: 10.1007/BF00992698.

23. Liu C, Xu X, Hu D. Multiobjective reinforcement learning: A comprehensive overview. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*. 2015;45(3):385-398. doi: 10.1109/TSMC.2014.2358639.

24. Lopez-Guede JM, Ramos-Hernanz JA, Zulueta E, Fernandez-Gamiz U, Oterino F. Systematic modeling of photovoltaic modules based on artificial neural networks. *Int J Hydrogen Energy*. 2016;41(29):12672-12687. doi: 10.1016/j.ijhydene.2016.04.175.

25. Kusiak A, Zhang Z, Li M. Optimization of wind turbine performance with data-driven models. *IEEE Transactions on Sustainable Energy*. 2010;1(2):66-76. doi: 10.1109/TSTE.2010.2046919.

26. Shi aKL, McLaren PG. Neural-network-based sensorless maximum wind energy capture with compensated power coefficient. *IEEE Trans Ind Appl*. 2005;41(6):1548-1556. doi: 10.1109/TIA.2005.858282.

27. Li H, Zhang D, Foo SY. A stochastic digital implementation of a neural network controller for small wind turbine systems. *IEEE Transactions on Power Electronics*. 2006;21(5):1502-1507. doi: 10.1109/TPEL.2006.882420.

28. Ho SL, Lo EWC, Wong HC. A particle swarm optimization-based method for multiobjective design optimizations. *IEEE Trans Magn*. 2005;41(5):1756-1759. doi: 10.1109/TMAG.2005.846033.

29. Coelho LdS, Ayala HVH, Alotto P. A multiobjective gaussian particle swarm approach applied to electromagnetic optimization. *IEEE Trans Magn*. 2010;46(8):3289-3292. doi: 10.1109/TMAG.2010.2047250.

30. Goudos SK, Zaharis ZD, Kampitaki DG, Rekanos IT, Hilas CS. Pareto optimal design of dual-band base station antenna arrays using multi-objective particle swarm optimization with fitness sharing. *IEEE Trans Magn*. 2009;45(3):1522-1525. doi: 10.1109/TMAG.2009.2012695.

31. Agrawal S, Dashora Y, Tiwari MK, Son Y. Interactive particle swarm: A pareto-adaptive metaheuristic to multiobjective optimization. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*. 2008;38(2):258-277. doi: 10.1109/TSMCA.2007.914767.

32. Jonkman JM, Butterfield S, Musial W, Scott G. Definition of a 5MW reference wind turbine for offshore system development. *National Renewable Energy Laboratory (NREL)*. 2009. doi: 10.2172/947422.

33. NREL NWTC FAST version 7. available online: Https://Nwtc.nrel.gov/FAST7/ (accessed on 21 oct 2018).

34. Rahimi H, Martinez Garcia A, Stoevesandt B, Peinke J, Schepers G. An engineering model for wind turbines under yawed conditions derived from high fidelity models. *Wind Energy*. 2018;8:618-633.

35. Rahimi H, Hartvelt M, Peinke J, Schepers J. Investigation of the current yaw EngineeringModels for simulation of wind turbines in bem and comparison with cfd and experiment. *IOP Publishing*. 2016;753.

36. Rahimi H, Dose B, Stoevesandt B, Peinke J. Investigation of the validity of bem for simulation of wind turbines in complex load cases and comparison with experiment and CFD. *IOP Publishing*. 2016;749.

37. Pitt DM, Peters DA. Theoretical prediction of dynamic-inflow derivatives. *Vertica*. 1981;5(1):21-34.

38. Ehrgott M, Gandibleux X. A survey and annotated bibliography of multiobjective combinatorial optimization. *OR-Spektrum*. 2000;22(4):425-460. https://doi.org/10.1007/s002910000046. doi: 10.1007/s002910000046".

39. UCAR/NCAR - earth observing laboratory. 2013. PCAPS ISFS 1 second data. version 1.0. UCAR/NCAR - earth observing laboratory. https://Doi.org/10.5065/D6QV3JRP. accessed 15 jul 2018.

40. Saenz-Aguirre A, Fernandez-Gamiz U, Zulueta E, Ulazia A, Martinez-Rico J. Optimal wind turbine operation by artificial neural network-based active gurney flap flow control. *Sustainability*. 2019;11:2809.

41. Foster CS. Simulation of a cold-air pool in utah's salt lake valley. *The University of Utah*. 2015.

Table 1. Principal characteristics of the NREL 5MW reference wind turbine

Turbine model	NREL 5MW
Skewed wake correction	Pitt and Peters
Rated power	5 MW
Rotor diameter	126 m
Hub height	90 m
Cut-in wind speed	3 m/s
Rated wind speed	11.4 m/s
Cut-out wind speed	25 m/s

Table 2. Parameter definition for the ANN based RL yaw control training process

Parameter	**Value**
Wind Speed [m/s]	3:2:25
DeserveMove [%]	5
YawRateK [-]	0:0.1:1
YawToMove [deg]	0:10:90
ANN Input neurons [-]	4
ANN Hidden Neurons [-]	[75 25]
ANN Output neurons [-]	2
ANN Learning Rate [-]	$1\cdot10^{-50}$
Training Ratio [%]	90
Validation Ratio [%]	5
Test Ratio [%]	5

Table 3. Parameters definition for the ANN based RL yaw control PSO and PoF optimization algorithm

Parameter	**Symbol**	**Value**
Population [-]	P	1000
Iterations [-]	n	30
phi_1_max [-]	phi_1_max	1.5
phi_2_max [-]	phi_2_max	0.1
Inertia_max [-]	I_max	0.5

Table 4. Parameters definition for the ANN based RL yaw control. Decision Making process.

Parameter	**Wind Speed**
Maximum Mechanical Moment [N·m]	$2.5\cdot10^{5}$

Table 5. Validation results of the proposed ANN based RL yaw control strategy

Case	YawAngle [deg]	WindSpeed [m/s]	YawRateK [-]	YawToMove [deg]	Generated power gain [%]	z axis yaw mechanical moment [N·m]
Case 1	72.5	9.1	0.1368	68.87	42.64	$3.657 \cdot 10^5$
Case 2	38.8	10.73	0.0521	65.8992	21.86	$8.317 \cdot 10^4$
Case 3	-49.1	15.3	0.2821	7.6969	12.92	$1.533 \cdot 10^5$
Case 4	43.4	15.7	0	0	0	0
Case 5	42.1[9]	25	0	0	0	0
Case 6	-46.3	6.3	0.0272	37.72	19	$1.227 \cdot 10^5$

Table 6. State and action variables of the ANN based RL yaw control strategy proposed in [18]

Case	YawAngle [deg]	YawRateK [-]
Case 1	72.5	0.5
Case 2	38.8	0.7
Case 3	-49.1	0.9
Case 4	43.4	0.6
Case 5	42.1[9]	0.6
Case 6	-46.3	0.9

Table 7. Power gain and mechanical moment comparison

Case	Power gain [%]	Mechanical moment [N·m]	Power gain Old [%]	Mechanical moment Old [N·m]	Δ Generated power gain [%]	Δ z axis yaw mechanical moment [%]
Case 1	42.64	$3.657 \cdot 10^5$	47.83	$5.185 \cdot 10^5$	-5.19	-29.45
Case 2	21.86	$8.317 \cdot 10^4$	27.43	$3.325 \cdot 10^5$	-5.57	-74.99
Case 3	12.92	$1.533 \cdot 10^5$	13.92	$5.114 \cdot 10^5$	-1	-70.08
Case 4	0	0	0	$5.609 \cdot 10^5$	0	-100
Case 5	0	0	0	$8.308 \cdot 10^5$	0	-100
Case 6	19	$1.227 \cdot 10^5$	32.5	$4.487 \cdot 10^5$	-13.5	-72.65