IOSR Journal of Electrical and Electronics Engineering (IOSR-JEEE)
e-ISSN: 2278-1676,p-ISSN: 2320-3331, Volume 13, Issue 1 Ver. III (Jan. – Feb. 2018), PP 23-29
www.iosrjournals.org

UAV Heading Control in Windy and Turbulent Conditions Using Reinforcement Learning

Kimathi S*., Kang'ethe S.,** and Kihato P.**
**Kimathi S., Department of Electrical and Electronic Engineering, DeKUT*
*** Kang'ethe S., and Kihato P., Department of Electrical, Electronic and Information Engineering, JKUAT*
Corresponding Author: Kimathi S.

Abstract: *Due to the high non-linearity and coupling of a system model in an Unmanned Aerial Vehicle, UAV, the control of the heading has been a challenging task especially under windy and turbulent conditions. In this paper an online adaptive method using reinforcement learning is proposed to counter the effects of wind disturbances. The heading controller is designed in Matlab/Simulink for controlling a UAV in an X-Plane test platform. Through the X-Plane test platform, the performance of the designed controller is shown using real time simulations under different cross wind conditions. The performance of the proposed method is compared to that of a well tuned PID controller. The results show that the proposed method performs better in tracking a given heading angle under windy conditions.*

Keywords: *UAV, Heading, Reinforcement Learning, Matlab, X-Plane*

Date of Submission: 25-01-2018 Date of acceptance: 15-02-2018

I. Introduction

Unmanned Aerial Vehicles, UAVs have become an interesting resource for applications where manned operations are considered inefficient and/or dangerous. Due to their design, various platforms exist such as the fixed wing and the quadcopter with each having different features. These devices are perking up interest in different applications including surveillance, search and rescue [1], target tracking, naval operations, weather observations and agricultural services. For the UAV to accomplish these tasks, it is paramount for it to have a good heading controller.

Due to system dynamics and the nonlinearity of an UAV system, various control techniques have been used to control and stabilize it. Those include PID control [1] [3], H_∞ control [4], fuzzy systems in [5] [6], active disturbance rejection control, ADRC in [7] [8] and [9] an adaptive backstepping approach where the dynamics of the cross track error was derived using the lateral system equations of motion.

In this paper an adaptive control strategy based on reinforcement learning technique for heading control of a fixed wing UAV in windy and turbulent conditions is presented. Due to the high nonlinearity of the system dynamics associated with small aerial vehicles and lack of complete knowledge of vehicle dynamics for parameter estimation, and adaptive method based on reinforcement learning is proposed. Reinforcement learning explores actions from available courses of action and chooses the best course of action based on the reward it gets, hence suitable for this kind of application. A nonlinear model of a small fixed wing UAV is taken, which is linearized about a stable trim point and decoupled into longitudinal and lateral designs. The lateral design is used to design the controller. The proposed controller will act on the deflection angles of the two lateral control surfaces i.e. the aileron and rudder in the presence of simulated windy conditions in X-Plane test platform.

This rest of this paper is organized as follows: Section II presents the basics of UAV control, section III introduces reinforcement learning principles, section IV gives a brief preview of X-Plane test platform, section V gives the model of the UAV and the design of the controller, section VI provides the results and discussion and the conclusion is given in the last section.

UAV Control Basics

Figure 1 shows that a UAV can move about the three axes of motion (*x,y,z*) from its centre of gravity.

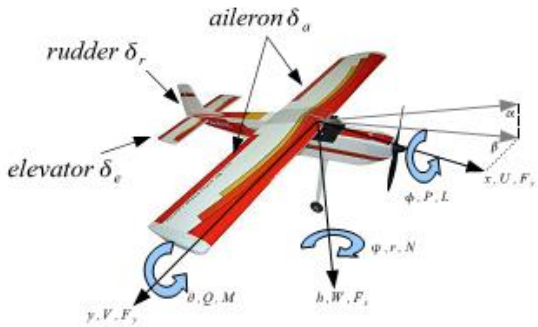

Figure 1: UAV Control surfaces

The relative position of a UAV is converted to angular control in the three principle control axes; roll (φ), pitch (θ) and yaw (Ψ) [2]. The control surfaces for a fixed wing UAV are as illustrated in Figure 1; ailerons to control the rolling elevator to control the pitching and rudder to control the yawing movement. In addition to these three control surfaces, the engine throttle controls the engines power.

Th equations of motion for a fixed wing UAV are derived in [10] [11]. The nonlinear equations are linearized about a stable trim level. The linear models obtained describe the aircraft natural longitudinal and lateral modes.

A lateral state space model is decoupled from within the linear model and used with inputs of aileron and rudder to control the heading of an aircraft. The decoupled lateral model state space equation is given as [12]

$$\dot{x} = A_{lat}x_{lat} + B_{lat}u_{lat}$$

Where A_{lat} is the state matrix, x_{lat} is the decoupled lateral state space model with $[\rho \;\; \beta \;\; r \;\; \varphi\,]^T$ as the state variables. ρ is the roll rate, β is the sideslip angle, r is the yawing rate and φ is the roll angle . B_{lat} the input matrix, u_{lat} is the control input and comprises of δ_a the aileron deflection and δ_r the rudder deflection.

II. Reinforcement Learning

Reinforcement learning, RL is learning what to do – how to map situations to actions, so as to maximize some numerical reward. The learning agent is not told the correct actions; instead it explores the possible actions and remembers the reward it receives. It is inspired by natural learning mechanisms where animals adjust their actions based on the reward or punishment stimuli received from interacting with the environment. The RL model consists of a set of environment states $s_t \in S$; a set of actions $a_t \in A$ that an agent can perform at each state, and as a consequence of its action, the agent receives a numerical reward r_t. At each time step, an agent implements a mapping from states to probabilities of selecting each possible action. This mapping is called the agent's policy and denoted as π_t which maximizes the cumulative reward of an agent over time as [13]

$$R = \sum_{t=0}^{\infty}\gamma^t r_t$$

where $0 < \gamma < 1$ is a discount factor, which reduces the value of future rewards. Reinforcement learning encompasses dynamic programming, DP Monte Carlo methods and temporal difference, TD methods. TD methods comprise of Q-learning and *sarsa* algorithm where the latter is an online learning method [14].

TD methods can learn directly from raw experience without a model of the environment dynamics and like DP, TD methods updates estimates based in part on other learned estimates without having to wait for a final outcome. The simplest TD method is given as

$$V(s_t) \leftarrow V(s_t) + \alpha[r_{t+1} + \gamma V(s_{t+1}) - V(s_t)]$$

SARSA algorithm being an online TD method, estimates $Q^\pi(s,a)$ for the current behavior policy π for all states *s* and actions *a.* This is done using the same TD method described above but the transitions from state-action pair to state-action pair are considered rather than from a state to a state transition, hence the value of the state-action pairs

$$Q(s_t,a_t) \leftarrow Q(s_t,a_t) + \alpha[r_{t+1} + \gamma Q(s_{t+1},a_{t+1}) - Q(s_t,a_t)]$$

This update is done after every transition from a non terminal state s_t. Therefore with *sarsa* there is a continual estimate of Q^π for the behavior policy π and at the same time a change of π towards greediness with respect to Q^π.

X-Plane

X-Plane is powerful flight simulator; it is not a game but rather an engineering tool that can be used to predict the flying qualities of fixed and rotary wing aircraft with considerable accuracy [15].

It has the capacity to send and receive data to and from other devices using the User Datagram Protocol, UDP. Each data packet is configured to carry specific aircraft parameters that are selected in the X-Plane Input and Output data interface. Once these are selected on X-Plane, through loopback addresses it is possible to receive them in Matlab/Simulink in the same computer using X-Plane Communication Library in Simulink. The received packets are repackaged for use in Matlab/Simulink environment and the sent data is repacked in a format that can be received and processed by X-Plane.

X-Plane also has the functionality of altering the weather conditions i.e. wind speed, shear speed and direction and turbulence of an altitude layer as shown in Figure 2. This allows for close to real life flying conditions hence a robust simulation environment.

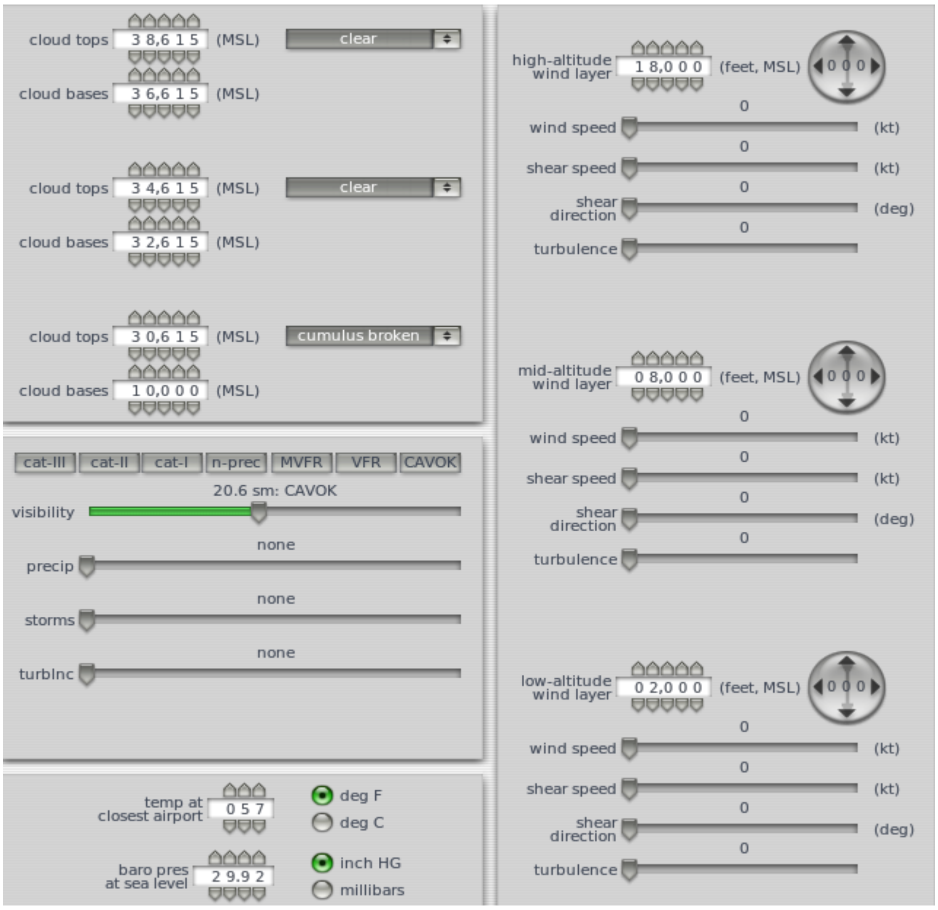

Figure 2: X-Plane atmospheric layers

UAV Model

An Ultra Stick 25E UAV is modeled using the aerodynamic coefficients in the mathematical model as provided for in [11] and taken from [4] and [16] with physical characteristics as $m = 1.9kg, b = 1.2m, g = 9.8\frac{m}{s^2}$, $S = 0.32\ m^2, \bar{c} = 0.3\ m, \rho = 1.225kg/m^3, \frac{1}{\pi eAR} = 0.0815$.

With the above parameters, the state space matrix was calculated under trim conditions as given in [12] as

$$A = \begin{bmatrix} -1.4000 & 0 & 0 & 9.4953 \\ -30.9000 & -12.8000 & 14.4000 & 0 \\ 1.4781 & -0.4480 & -6.080 & 0 \\ 0 & 1.0000 & 0 & 0 \end{bmatrix}$$

$$B = \begin{bmatrix} 0 & 0.7412 \\ 61.4000 & 12.4000 \\ -3.6700 & 15.0000 \\ 0 & 0 \end{bmatrix}$$

With states as $[\rho\ \ \beta\ \ \ r\ \ \varphi\ ]$ as was defined in literature above.

Controller Design and Implementation

From the state space model above, the associated Ricatti coefficient, P is calculated which is formulated as the Algebraic Ricatti Equation. The Ricatti coefficient is used to calculate the Cost function for each state-action pair using a simple Lyapunov function $V = X^T \boldsymbol{P} X$, which is reformulated to include references as

$$Q(s,a) = (X - Z)^T \boldsymbol{P} (X - Z)$$

where X are states and Z, the references.

A reward function is calculated as the deviation of the target state from the desired state, which in this work is taken as the heading error.

$$r(s) = -C_1(\emptyset - \emptyset_{ref})$$

The total value function is calculated, which is a sum of previous state-action value function, the current reward and the current state-action value function according to *sarsa algorithm* as

$$Q(s_1, a_1) \leftarrow Q(s_1, a_1) + \alpha[r_2 + \gamma Q(s_2, a_2) - Q(s_1, a_1)]$$

$$Q(s_2, a_2) \leftarrow Q(s_2, a_2) + \alpha[r_3 + \gamma Q(s_3, a_3) - Q(s_2, a_2)]$$

…and so forth.

After each cycle, this is updated as the total value function for the next cycle of learning.

According to [17] it is allowed to have a *one step gradient search* of the value function i.e. exploitation for ease of real time implementation and offers less computational burden. This was formulated as the temporal difference as;

$$\delta_t = r_{t+1} + \gamma Q(s_{t+1}, a_{t+1}) - Q(s_t, a_t)$$

An optimal control effort is given as $u^* = -\mathrm{K}\mathrm{X}$ but according to differential games algorithm, this is expressed as

$$u^* = -\frac{1}{2} R^{-1} g^T \nabla V^*$$

where ∇V^* is taken as the change in the optimal cost function. In this work the temporal difference between successive cost functions is used as the optimal cost function which will be reinforcement signal.

The control signals from a feed forward neural network are compared with this control signal. Then back propagation algorithm updates the feedforward neural network weights using back propagation to compensate for the difference. This implies that we are correcting the error in the control deflection in the next control deflection through update of weights thus slowly taking our deflections to the best available control effort in each consecutive cycle. This process is shown in Figure 3.

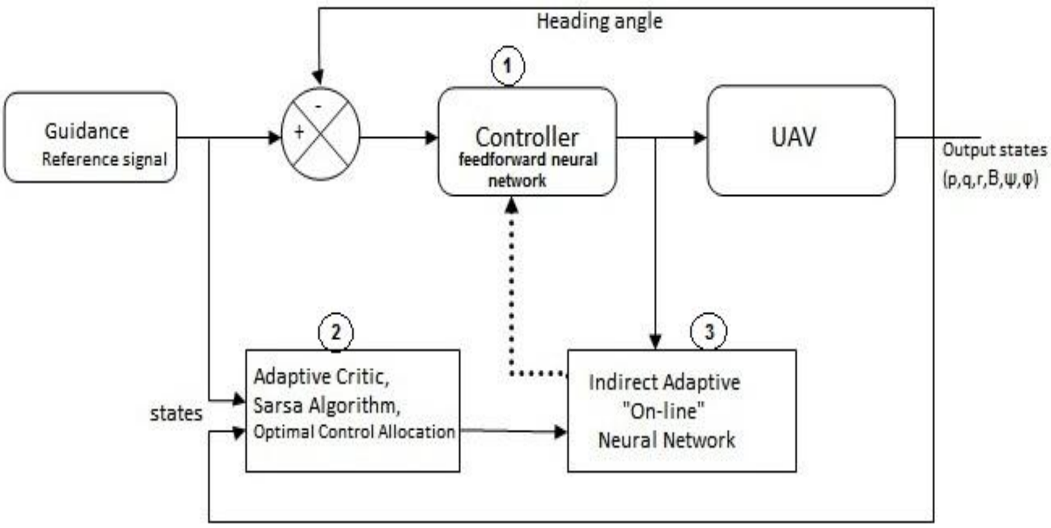

Figure 3: Experiment Block Diagram

III. Results And Discussion

The controller designed as above was used to control a UAV in X-Plane as in [12] where the disturbances were introduced using the weather setting in X-Plane as explained above.

A crosswind of 5 knots, 345^0 wind direction and 2^0 wind shear was set. The tracking ability of the controller under this condition was as shown in Figure 4 below.

The red line shows PID controller the designed and the Blue line shows the RL controller

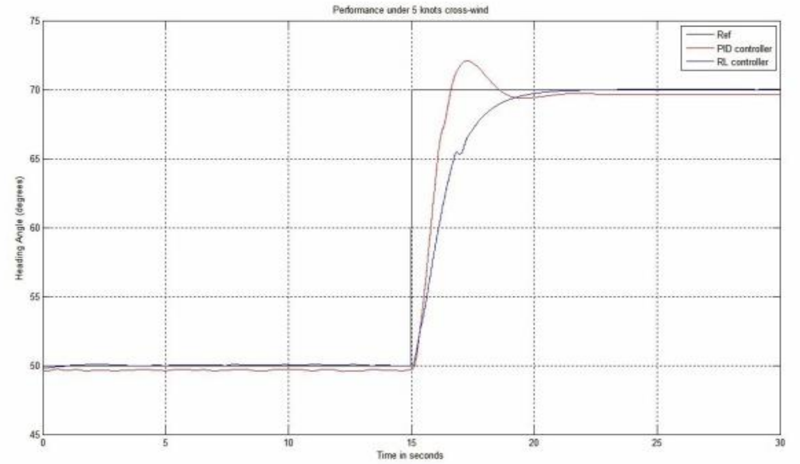

Figure 4: 5 knots wind disturbance

The response with these disturbance show that the UAV is able to track well the reference in the presence of these disturbances for both the controllers. The overshoot in the step heading change is explained in [12] [2].
A 10 knots, 345^0 wind direction and 2^0 wind shear was set in X-Plane weather widget, and the tracking performance of the controllers are shown in Figure 5.

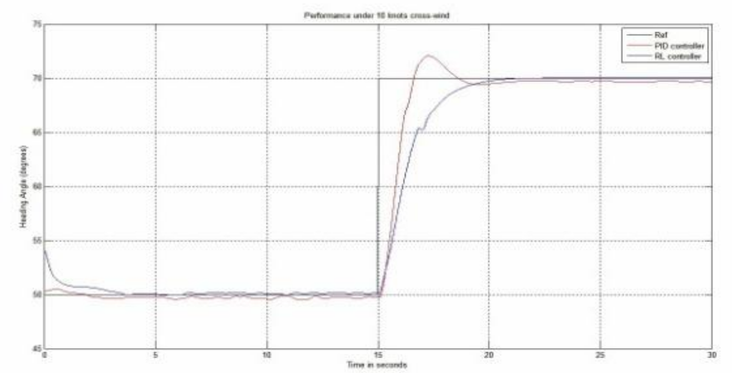

Figure 5: 10knots wind disturbance

Here, it can be seen that the disturbance is substantial to cause some slight bumpy deviations from the reference. But the designed controllers are able to keep the UAV heading close to reference heading.
The crosswind was adjusted to 20 knots, 345^0 wind direction and 5^0 wind shear. A similar flight regime as above was used. The response was as shown in Figure 6 below.

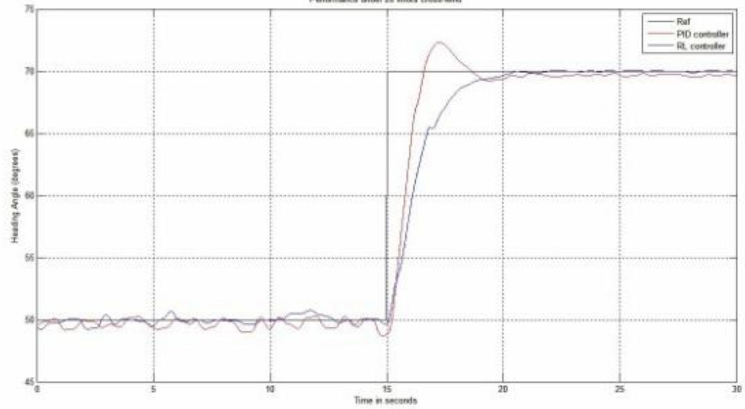

Figure 6: 20 knots wind disturbance

The disturbances are big enough to cause some bumpy deviations from the reference. The response from the RL controller has lesser bumpy deviations from the reference. Therefore, the RL controller performs slightly better.
A 40 knots, 345^0 wind direction and 10^0 wind shear was introduced as Figure 7 shows the responses achieved by the two controllers.

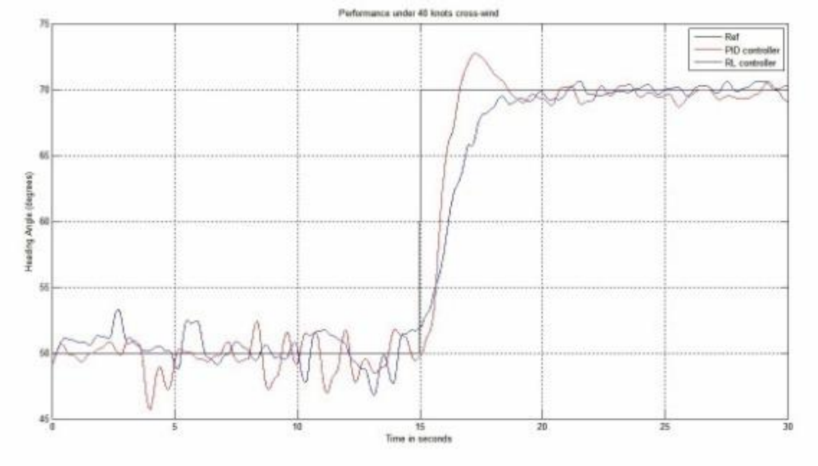

Figure 7: 40 Knots wind disturbance

This represented severe wind conditions for a small UAV. The tracking ability of the controllers under this conditions showed big deviations from the reference heading. This similar response was reported in [18].
Finally, turbulence was introduced which included 20 knots, a wind shear of 5 knots and a 10^0 shear speed in X-Plane. This represents extreme weather turbulent conditions for a small UAV. The tracking performance under this conditions was as shown in Figure 8.

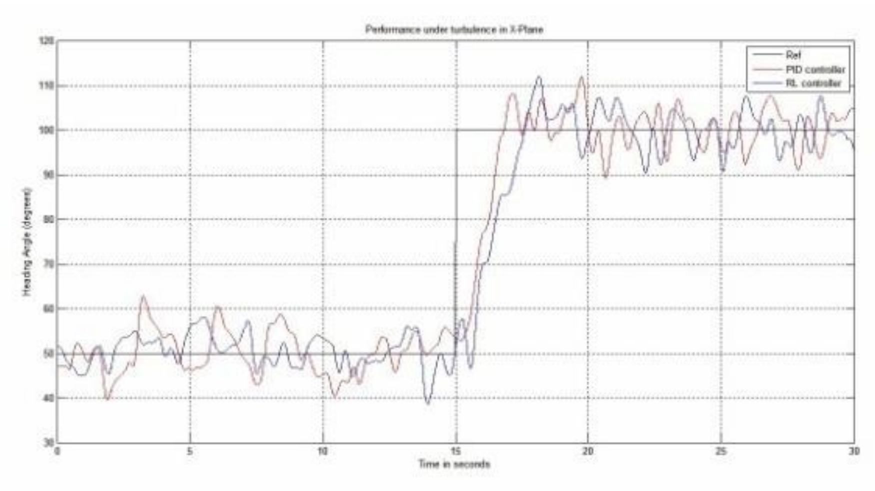

Figure 8: Turbulence in X-Plane

The response under this condition is poor; the UAV heads in the right direction but with big deviations and oscillations from the desired reference. It can also be noted that the PID controller response had bigger deviations than the RL controller. The reason for this behavior is highlighted in [19] [20]which states that the sensor and actuator dynamics do not scale with aircraft dynamics hence a limitation in lateral response for UAVs in turbulent conditions.

IV. Conclusion

In this paper, a heading controller for a small fixed wing UAV in windy and turbulent conditions was proposed. A controller based on reinforcement learning was used to achieve that, where the lateral dynamics of the UAV with respect to the desired heading was derived. The simulation results show the usability of the proposed method in comparison to the already existing PID control. Future work should be done to validate the proposed method using a prototype.

References

[1]. A. Mansoor et al., "Heading control of a Fixed wing UAV using Alternate control surfaces," IEEE, vol. 2, December 2012.
[2]. S. Kimathi, S. Kang'ethe, and P. Kihato, "UAV heading controller using Reinforcement learning," in Pan African Conference on Science, Computing and Telecommunications, PACT, Nairobi, 2017.
[3]. Z. Qing and G. Zhiqiang, "On Practical application of Active Disturbance Rejection Control," in Proceedings of the 29th Chinese Control Conference, Beijing, 2010.
[4]. H. Ferreira et al., "Disturbance Rejection in a Fixed Wing UAV using nonlinear H-Infinty state feedback," in 9th International Conference on Control and Automation, Santiago, USA, 2011.
[5]. V. Nagababu, and M. Prabhudas, "A Fuzzy logic strategy on attitude controlling of Longitudinal autopilot for better disturbance rejection," International Journal of Engineering Research and Technology, vol. Vol. 2, no. Issue 12, December 2013.
[6]. D. Stojcsics, "Fuzzy controller for small size Unmanned Aerial Vehicles," in 10th IEEE International Symposium on Applied machine Intelligence and Informatics, Herl'any, Slovakia, 2012.
[7]. X. Hua et al., "Disturbance rejection in UAV's velocity and Altitude Control: Problems and Solutions," in Proceedings of the 30th Chinese Control Conference, Yantai, China, 2011.
[8]. L. Jing-Mei and Z. Ke, "Design of Active Disturbance Rejection Controller for Autonomous Aerial Refuelling UAV," in IEEE, 2013.
[9]. A. Brezoescu et al., "Adaptive Trajectory following for a fixed wing UAV in Presence of Crosswind," Journal of Intelligent Robotic Systems, vol. Vol. 69, 2013.
[10]. R. Nelson, *Flight Stability and Automatic Control*, 2nd ed. Singapore: McGraw Hill, 1998.
[11]. R. Beard and T. McLain, *Small Unmanned Aircraft; Theory and Practice*, 1st ed. Princeton, New Jersey: Princeton University Press, 2012.
[12]. S. Kimathi, S. Kang'ethe, and P. Kihato, "Application of Reinforcement Learning in heading control of a fixed wing UAV using X-Plane test platform," International Journal of Scientific & Technology Research, vol. 6, no. 02, February 2017.
[13]. R. Sutton and A. Barto, *Reinforcement Learning; An Introduction*, 1st ed. Massachusetts: MIT press, 2005.
[14]. S. Rusell and P.Norvig, *Artificial Intelligence; A Modern Approach*, 3rd ed. New Jersey: Pearson Education, 2010.
[15]. A. Meyer, *X-Plane 10: Operation Manual*, 2012.
[16]. Y. Paw, "Synthesis and Validation of Flight Control for UAV," University of Minnesota, Minnesota, PhD Thesis, Dec. 2009.
[17]. R. Sutton et al., Experiments with Reinforcement Learning in Problems with Continuous State and Action Spaces, Unpublished.
[18]. A. Bittar, O. Nuesa, and H. Figueiredo, "Hardware in the loop simulation with X-Plane of Attitude Control of a SUAV Exploring Atmospheric Conditions," Journal of Intelligent Robotic Systems, 2014.
[19]. K. Zhukov et al., "Effects of atmospheric turbulence on UAV," in IFFK, Budapest, 2014.
[20]. W. Pisano and D. Lawrence, "Control Limitations of Small Unmanned Aerial Vehicles in Turbulent Environments," in AIAA Guidance, Navigation and Control Conference, Chicago, 2009.
[21]. D. Lundstrom and P. Krus, "Testing of Atmospheric Turbulence effects on the Performance of Micro Air Vehicles," International Journal of Micro Air Vehicles, 2012.

IOSR Journal of Electrical and Electronics Engineering (IOSR-JEEE) is UGC approved Journal with Sl. No. 4198, Journal no. 45125.

Kimathi, S "UAV Heading Control in Windy and Turbulent Conditions Using Reinforcement Learning." IOSR Journal of Electrical and Electronics Engineering (IOSR-JEEE) 13.1 (2018): 23-29.