# Deep Reinforcement Learning as a Rectification Agent in Process Control for Alarm Reduction

Ammar N. Abbas

*Data Science, Software Competence Center Hagenberg, Austria. E-mail: ammar.abbas@scch.at*

Georgios C. Chasparis

*Data Science, Software Competence Center Hagenberg, Austria. E-mail: georgios.chasparis@scch.at*

John D. Kelleher.

*ADAPT Research Centre, Technological University of Dublin, Ireland. E-mail: john.d.kelleher@tudublin.ie*

Conventional process controllers (such as proportional integral derivative controllers and model predictive controllers) are simple and effective once they have been calibrated for a given system. However, it is difficult and costly to re-tune these controllers if the system deviates from its normal conditions and starts to deteriorate. Recently, reinforcement learning has shown significant improvement in learning process control policies through direct interaction with a system, without the need of a process model or the system characteristics as it learns the optimal control by interacting with the environment directly. However, developing such a black-box system is a challenge when the system is complex and it may not be possible to capture the complete dynamics of the system just with a single reinforcement learning agent. Therefore, in this paper, we propose a simple architecture that does not replace the conventional proportional integral derivative controllers but instead augments the control input to the system with a reinforcement learning agent that adds a correction factor to the output provided by such controllers so as to maintain optimal process control even when the system is not operating under its normal condition.

*Keywords*: Deep Reinforcement Learning (DRL), process control, optimal control, alarm management.

## 1. Introduction

Industrial processes control has become autonomous with the advent of sophisticated control strategies such as Proportional Integral Derivative (PID) or *Model Predictive Control (MPC)* Efheij et al. (2019), based on the look-ahead optimization. However, one of the major disadvantages of such control laws is that their implementation requires an explicit understanding of the system dynamics and sometimes also knowledge of the environment. Furthermore, once the controller is tuned to the specific model or set-points of the system it only provides the optimal control under set system specificities. If the system deteriorates or the environmental conditions and set-points drift from the normal conditions, the controller starts deviating and provides sub-optimal control strategies and sometimes fails to control the process at all. In these cases, it becomes necessary to optimize the controller performance by re-tuning the controller parameters and system re-identification, tasks which lead to process shutdowns and massive time consumption Spielberg et al. (2019).

Recent developments in model-free *Deep Reinforcement Learning (DRL)* have demonstrated the feasibility of replacing such controllers with fully autonomous controllers that interact with the environment in an online setting and create their own understanding of the model of the environment, therefore eliminating the need for system re-identification Spielberg et al. (2019). Reinforcement Learning (RL) is a branch of machine learning that learns through interaction with the environment without having any prior knowledge of the dataset Sutton and Barto (2018). The majority of work on DRL for process control replaces the conventional controllers entirely with the DRL controller as suggested in Spielberg et al. (2019); Nian et al. (2020); McClement et al. (2021); Conradie and Aldrich (2001); Mageli (2019). Such an approach is well suited for simpler control prob-

Proceedings of the 32nd European Safety and Reliability Conference.
*Edited by* Maria Chiara Leva, Edoardo Patelli, Luca Podofillini and Simon Wilson
 *Published by* Research Publishing, Singapore
ISBN: 981-973-0000-00-0 :: doi: 10.3850/981-973-0000-00-0_output

lems. However, developing a controller for sophisticated control scenarios generally requires either proper domain knowledge or a very complex DRL algorithm structure that is not easily generalizable.

Process control is a critical optimization problem that needs to consider optimizing every time-step to be able to run the process smoothly because if it fails at any instant then the process trips (shutdown) and this may lead to catastrophic failures. DRL was developed to solve an optimization problem without considering the path that optimal policy takes to achieve the maximum cumulative reward. Therefore, it is not necessarily appropriate to replace conventional control with DRL as the trajectory a process follows can have a major impact on the process control. Hence, we argue that it is best to use DRL in a hybrid setting with the conventional controllers, as also recommended by Shin et al. (2019).

Therefore, we propose a novel yet simple methodology that merges the conventional controller with a DRL-based correction factor applied to each controller output. The corrected signal is then fed as an input to the plant. This correction factor aids the adjustment of the control in the case of system disturbances or when the controller requires re-tuning. DRL interacts with the process in real-time and generates an additional control signal that rectifies the output provided by the conventional controller (PID/ MPC) and results in the optimal control with reduced alarming scenarios and reduced operator burden.

## 2. Related Literature

An adaptive and self-learning model-free DRL controller is proposed by Spielberg et al. (2019). The proposed controller learns while interacting with the process in real-time, hence it is a data-based approach. The proposed system uses an actor-critic Konda and Tsitsiklis (1999) architecture for the DRL agent based on the Deep Policy Gradient (DPG) Lillicrap et al. (2015). In order to make the DRL agent completely aware of the system dynamics, the state is defined as the current state as well as the previous states, and the current control action taken by the RL agent as well as the previous control actions, up to a predefined number of the previous time steps. In addition, the state also incorporates the current deviation from the system-defined set-points. The approach is validated on the set-point tracking problem in control theory where the controller has to reach the predefined set-point with minimal oscillations and time while reducing the error caused by the deviation of the system state from the defined set-points. The performance of the DRL controller is evaluated through simulation experiments with a number of use cases, including (i) a paper machine, (ii) a distillation column, (iii) and a heating, ventilation, and air conditioning (HVAC) system.

A multi-criteria decision-making control process using DRL has been implemented by He et al. (2021) and has been evaluated using the case study of a textile manufacturing process. Process optimization for the textile industry includes various parameters to be tuned simultaneously, and DRL is well suited for such multi-objective optimization.

Panzer and Bender (2021) provide a literature review on the use of DRL in production systems. All of the work cited in this review paper replaces the conventional control methods and utilizes DRL for process optimization. The research reviewed was applied across a number of case studies, such as the liquid level control of multiple connected tanks, single- and multi-input and -output processes, and chemical-mechanical polishing Noel and Pandian (2014); Spielberg et al. (2017); Yu and Guo (2020). In all cases, the DRL-based controller achieves optimal performance from the conventional control strategies with reduced maintenance and cost along with increased process stability.

Mageli (2019) used a DRL agent to replace regular controllers in a case study of tank level regulation. The DRL controller was compared with a *Proportional* controller, a type of PID controller where only the first component Proportional (P) is used. The results showed that the P-controller performed better with stable controller output changes, whereas DRL with larger output changes resulted in system oscillation. The research indicates the complexity of the DRL controller does not outperform a simple well-tuned controller. However, with the increased complexity of the system nonlinearities and the ability to incorporate system deviation from the standard operating conditions for which the controller was tuned, DRL has great potential.

A generalizable approach to process control using DRL is used by McClement et al. (2021). The approach can be integrated within the existing control structures and be used to tune the PID or MPC controllers or can be used as an independent controller without the aid of any other existing control. For example, DRL is used as a set-point decision-maker Hernández-del Olmo et al. (2018) in a wastewater treatment plant, where the suggested set-points are then controlled using a PID controller.

Finally, Shin et al. (2019) present a brief introduction to RL and its use in process control followed by its limitations and comparison with conventional controllers. They argue that model-based/mathematical programming-based controllers such as MPC are limited in their ability to incorporate stochasticity of the environment and that RL can overcome these issues. Furthemore, they identify three strategies for implementing RL in process control: (i) replacing the conventional control with RL, (ii) hybrid RL and conventional controller, and (iii) RL to manage the control systems (PID tuning or MPC gain adjustments). In this paper, the second method of using a hybrid model is proposed.

### 2.1. *Literature Gap*

Several hybrid structures of MPC with RL were proposed by Lee and Wong (2010). The first method is a hierarchical structure where MPC determines the state regions to focus on for RL. The second includes learning value function for states in order to capture the uncertainties within the system model and incorporate them within MPC formulation. The third approach uses switching between MPC and RL where the MPC is used instead of the RL when a new state has been observed. Another example is the Dual MPC methodology introduced by Morinelly and Ydstie (2016), where RL is used to incorporate the predicted information within the model.

Most of the hybrid DRL-based conventional controller either uses RL to predict uncertainties within the environment and then incorporate such information within the mathematical modeling or uses RL independently with the conventional controllers with a switching probability. However, we propose to use DRL alongside the process controllers and to act as a correcting agent that feeds in the information of the current state and action proposed by the conventional control and outputting a correction factor added to the output of the conventional control. Such a method can help correct the control signal during the process disturbances and abnormalities where the probability of occurrence of multiple alarms is high and it can help minimize or mitigate such alarm scenarios.

## 3. Proposed Methodology

Our proposed methodology is to add the DRL agent integrated within the industrial process. The agent continuously observes the state of the system and at every time step provides corrected signals to be added to the output of the PID/MPC controller, which is then fed as the control signal to the plant. We propose two different architectures in terms of the state representative for the RL as shown in fig. 1 and fig. 2. The first architecture shown in fig. 1 represents the state as a function of the industrial process (plant) output concatenated with the output signal from the PID controller.

The modified architecture as shown in fig. 2 represents the state as the function of output from the plant, output signal from the PID controller, deviation of process variables from the setpoint, and the action proposed by the DRL agent.

Three different reward functions can be used (i) L1 norm, (ii) L2 norm, and (iii) polar reward as shown in Spielberg et al. (2019). The first two reward functions will enable an agent to learn faster but will likely result in more oscillation in the control signals than the third. The third reward function stops penalizing the agent once it sees that the agent is starting to improve. The objective function of the agent is to minimize the deviation from the setpoints during the disturbance phase where the PID/MPC controller fails to mitigate the

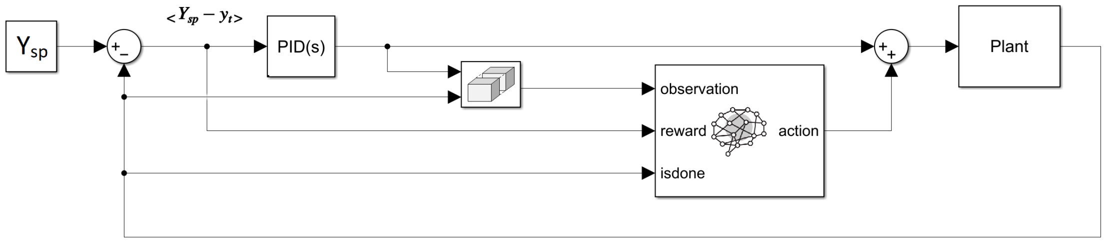


Fig. 1. Simplified DRL-RA methodology.

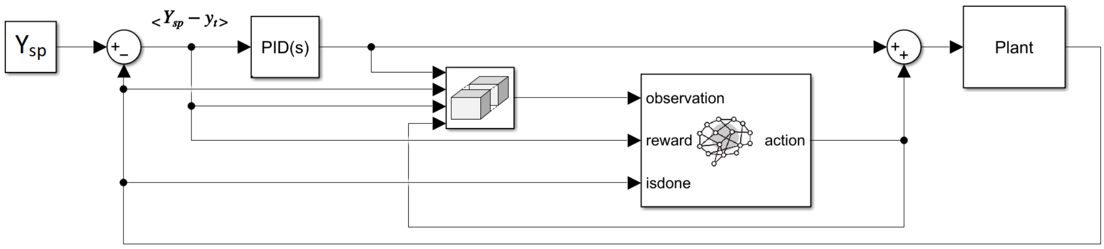


Fig. 2. Enriched state modified DRL-RA methodology (ES-DRL-RA).

error and reduce the number of alarms.

## 4. Future Work

In this paper, we proposed a novel hybrid architecture of DRL-based conventional process control and its potential applications in the case of alarm reduction and mitigation to help the operators in an abnormal situation where handling multiple alarms simultaneously becomes difficult, which causes occlude the root cause failure of the system. In the future, we will try to use this methodology in several real-world case studies with historic data or with the help of a simulator and compare the performance of such hybrid architecture over the conventional control. The presented state and reward architectures of the simplified and modified methodology will be compared and evaluated against the benchmark of an average of the total number of alarms generated compared to the conventional control.


### Acknowledgement

This research is supported by the project Collaborative Intelligence for Safety-Critical systems (CISC); the CISC project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under the Marie Skłodowska-Curie grant agreement no. 955901. The work of Kelleher is also partly funded by the ADAPT Centre which is funded under the Science Foundation Ireland (SFI) Research Centres Programme (Grant No. 13/RC/2106_P2).


## References


Conradie, A. and C. Aldrich (2001). Plant-wide neurocontrol of the tennessee eastman challenge process using evolutionary reinforcement learning. In *In 3rd international conference in intelligent processing and manufacturing of materials*.

Efheij, H., A. Albagul, and N. Ammar Albraiki (2019). Comparison of model predictive control and pid controller in real time process control system. In *2019 19th International Conference on Sciences and Techniques of Automatic Control and Computer Engineering (STA)*, pp. 64–69.

He, Z., K.-P. Tran, S. Thomassey, X. Zeng, J. Xu, and C. Yi (2021). A deep reinforcement learning based multi-criteria decision support system for optimizing textile chemical process. *Computers in Industry 125*, 103373.

Hernández-del Olmo, F., E. Gaudioso, R. Dormido, and N. Duro (2018). Tackling the start-up of a reinforcement learning agent for the control of wastewater treatment plants. *Knowledge-Based Systems 144*, 9–15.

Konda, V. and J. Tsitsiklis (1999). Actor-critic algorithms. *Advances in neural information processing systems 12*.

Lee, J. H. and W. Wong (2010). Approximate dynamic programming approach for process control. *Journal of Process Control 20*(9), 1038–1048.

Lillicrap, T. P., J. J. Hunt, A. Pritzel, N. Heess, T. Erez, Y. Tassa, D. Silver, and D. Wierstra (2015). Continuous control with deep reinforcement learning. *arXiv preprint arXiv:1509.02971*.

Mageli, E. R. (2019). Reinforcement learning in process control. Master's thesis, NTNU.

McClement, D. G., N. P. Lawrence, P. D. Loewen, M. G. Forbes, J. U. Backström, and R. B. Gopaluni (2021). A meta-reinforcement learning approach to process control. *IFAC-PapersOnLine 54*(3), 685–692.

Morinelly, J. E. and B. E. Ydstie (2016). Dual mpc with reinforcement learning. *IFAC-PapersOnLine 49*(7), 266–271.

Nian, R., J. Liu, and B. Huang (2020). A review on reinforcement learning: Introduction and applications in industrial process control. *Computers & Chemical Engineering 139*, 106886.

Noel, M. M. and B. J. Pandian (2014). Control of a nonlinear liquid level system using a new artificial neural network based reinforcement learning approach. *Applied Soft Computing 23*, 444–451.

Panzer, M. and B. Bender (2021). Deep reinforcement learning in production systems: a systematic literature review. *International Journal of Production Research*, 1–26.

Shin, J., T. A. Badgwell, K.-H. Liu, and J. H. Lee (2019). Reinforcement learning–overview of recent progress and implications for process control. *Computers & Chemical Engineering 127*, 282–294.

Spielberg, S., R. Gopaluni, and P. Loewen (2017). Deep reinforcement learning approaches for process control. In *2017 6th international symposium on advanced control of industrial processes (AdCONIP)*, pp. 201–206. IEEE.

Spielberg, S., A. Tulsyan, N. P. Lawrence, P. D. Loewen, and R. Bhushan Gopaluni (2019). Toward self-driving processes: A deep reinforcement learning approach to control. *AIChE journal 65*(10), e16689.

Sutton, R. S. and A. G. Barto (2018). *Reinforcement learning: An introduction*. MIT press.

Yu, J. and P. Guo (2020). Run-to-run control of chemical mechanical polishing process based on deep reinforcement learning. *IEEE Transactions on Semiconductor Manufacturing 33*(3), 454–465.