Proceedings Paper:
Sanusi, I., Mills, A., Trodden, P. et al. (2019) Reinforcement learning for condition-based control of gas turbine engines. In: 2019 18th European Control Conference (ECC). 2019 18th European Control Conference (ECC), 25-28 Jun 2019, Naples, Italy. IEEE, pp. 3928-3933. ISBN: 9783907144008.

https://doi.org/10.23919/ECC.2019.8795878

Reinforcement learning for condition-based control of gas turbine engines

Ibrahim Sanusi[1], Andrew Mills[2], Paul Trodden[3], Visakan Kadirkamanathan[4], Tony Dodd[5]

***Abstract*— A *condition-based control* framework is proposed for gas turbine engines using reinforcement learning and adaptive dynamic programming (RL-ADP). The system behaviour, specifically the fuel efficiency function and constraints, exhibit unknown degradation patterns which vary from engine to engine. Due to these variations, accurate system models to describe the true system states over the life of the engines are difficult to obtain. Consequently, model-based control techniques are unable to fully compensate for the effects of the variations on the system performance. The proposed RL-ADP control framework is based on Q-learning and uses measurements of desired performance quantities as reward signals to learn and adapt the system efficiency maps. This is achieved without knowledge of the system variation or degradation dynamics, thus providing a through life adaptation strategy that delivers improved system performance. In order to overcome the long standing difficulties associated with the application of adaptive techniques in a safety critical setting, a dual-control loop structure is proposed in the implementation of the RL-ADP scheme. The overall control framework maintains guarantees on the main thrust control loop whilst extracting improved performance as the engine degrades by tuning sets of variable geometry components in the RL-ADP control loop. Simulation results on representative engine data sets demonstrate the effectiveness of this approach as compared to an industry standard gain scheduling.**

I. INTRODUCTION

Most engineering systems are subject to degradation, yet their control systems are not designed to explicitly account for it. While the dynamics that govern the operation of the systems are usually modelled or identified for the control design, the degradation dynamics are not; typically, these evolve over long timescales and in non-deterministic ways. This affects the states of the component health of the systems resulting in reduced performance and increased fuel consumption over time [1]. Opportunities to mitigate the effects of degradation therefore cannot be over emphasised as in the case of the civil gas turbine engine (GTE) where the cost of fuel accounts for about 15% to 25% of the total aircraft operating cost [2]. In addition to the gradual degradation, performance of the GTEs are also affected by fleet variations from engine build differences and changing operating conditions. Optimising the system performance as a result of these varying factors pose a major challenge to the GTE control.

The authors are with the Rolls-Royce Control and Monitoring Systems Technology Centre, Department of Automatic Control and Systems Engineering, University of Sheffield, Sheffield, United Kingdom.

[1]iesanusi1@sheffield.ac.uk, [2]a.r.mills@sheffield.ac.uk, [3]p.trodden@sheffield.ac.uk, [4]visakan@sheffield.ac.uk, [5]t.j.dodd@sheffield.ac.uk

The unknown degradation dynamics and variations affecting the GTE states reflect as changes in the measured/estimated system performance characteristics such as the system efficiency index and life [2]. Whilst monitoring of these performance characteristics can help to reduce the cost of operation from economic and performance perspectives, the opportunities to complement the GTE control design have received little attention e.g. monitoring of the fuel consumption and component temperatures have been used to predict the system life necessary for maintenance scheduling but not for feedback control [2], [3]. It is therefore increasingly important to use the information about the system performance characteristics in optimising the GTE control design whilst considering the reliability of its implementation.

In this paper, techniques that enable such capabilities are termed *condition-based* and are aimed at maintaining the GTE safety and reliability whilst optimising the system performance. Condition-based control (CBC) techniques can therefore be classed as types of adaptive control schemes that focus on optimising to slow and varying changes in the system performance. This combined with an appropriate adaptation strategy and architecture increases the feasibility of the scheme to a fully intelligent control and health management technology for industrial applications.

Existing techniques that have explored the possibility of CBC are mostly model-based including life extending control, performance seeking control (PSC) and model-predictive control (MPC) schemes [1]. Of note are the PSC schemes [4], [5] that match the actual degrading engine conditions by assuming that the degradation behaviour is well understood and modelled through high fidelity on-board models. These high fidelity models require a lot of effort and are expensive to build as reported in [6], [7] where the models took 15+ years of experience with the F100 class of engines and accurate nonlinear simulations. Similarly, recent work by General Electric (GE) aviation [8] which makes use of a tracking filter to estimate some engine deviation parameters (EDP) to account for degradation and engine-to-engine variations relied on the use of high fidelity dynamic engine models. Obtaining these dynamic degradation models in practice for the model-based performance control strategies is usually expensive and can be impractical in many situations.

The possibility of developing a model-free CBC can however be achieved by exploiting frameworks based on *reinforcement learning* (RL). RL schemes do not require knowledge of the system dynamics (including gradual degradation), but learn by interacting with the system and gradually modifying their actions based on some received reward signals [9]. The reward signals are direct measurements from the system that determine the utility of the current control strategy against a desired control goal or cost. RL schemes are therefore termed *goal-directed* control schemes [10], as only the reward signals are used to steer the system to optimal operating points without any knowledge of the system dynamics. In control, mathematical implementation of RL has been enabled through approximate/adaptive dynamic programming (ADP) and is theoretically linked to both adaptive and optimal control methods [11].

This paper proposes a RL-ADP scheme that provides a natural mathematical framework for the CBC problem by using as its reward signals the measurements of the system performance characteristics and without knowledge of the system degradation dynamics. The RL-ADP scheme uses measurements of the reward signals to learn and adapt the system efficiency maps and to extract improved system performance. Furthermore, in order to overcome the long standing difficulties associated with adaptive techniques in a safety critical setting such as the problem of bursting and potential instability [12], a dual-control loop structure is proposed in the implementation of the RL-ADP scheme. The proposed framework maintains guarantees on the main thrust control loop whilst extracting improved performance as the engine degrades by tuning sets of variable geometry components (VGC) in the RL-ADP control loop. This approach is essential to providing a potential route to certification of the overall control framework.

The rest of this paper is organised as follows. Section II provides the problem formulation for the CBC in the GTE control architecture. In Section III, an RL-ADP solution that addresses the GTE CBC problem is proposed along with the corresponding algorithm and control architecture. Section IV discusses the simulation of the proposed scheme and results from the algorithm implementation on representative engine data sets while Section V gives the concluding remarks and future works.

II. PROBLEM FORMULATION

Gas turbine engines consist of a compressor to draw in and compress air, a combustor to mix and burn fuel with the compressed air, and a turbine to extract power from the hot stream of air to generate thrust [13]. The desired level of thrust is mainly regulated via the control of fuel flow, with modern engines having other variable geometry components (VGC) such as the variable stator vanes (VSV), variable inlet guide vanes (VIGV) and the exhaust nozzle area [2]. A conceptual mathematical model for the GTE dynamics is given as:

$$\boldsymbol{x}_{k+1} = F(\boldsymbol{x}_k, \boldsymbol{u}_k, \boldsymbol{d}_k) \tag{1}$$

where $\boldsymbol{x} \in \mathbb{R}^n$ are the system states such as the shaft speeds (NH), engine pressure ratio (EPR), pressure and temperature. The control inputs $\boldsymbol{u} = \begin{bmatrix} u^{main} \\ \boldsymbol{u}^{aux} \end{bmatrix} \in \mathbb{R}^m$ consist of the main fuel flow input denoted by $u^{main} \in \mathbb{R}^{m_1} \subset \mathbb{R}^m$ and the additional control parameters such as the VGCs employed in many GTE designs [14] denoted by $\boldsymbol{u}^{aux} \in \mathbb{R}^{m_2} \subset \mathbb{R}^m$. The component health states $\boldsymbol{d} \in \mathbb{R}^d$ denotes the system performance characteristics such as the compressor and turbine efficiencies that change slowly over time due to degradation [1]. Typically, $\boldsymbol{d}$ is difficult to estimate as it is governed by non-deterministic processes which vary across fleets and from engine to engine. Conventionally, thrust regulation is achieved by designing the control system at some identified nominal models of the system (i.e at predetermined configurations of $\boldsymbol{d}$) [1], [13], [15].

Assumption 1. The control design for $u^{main} = h(\boldsymbol{y})$ guarantees the thrust response by regulating the system measurements $\boldsymbol{y} = c(\boldsymbol{x})$ to the desired reference i.e. $\boldsymbol{y} \rightarrow \boldsymbol{y}_{ref}$. This represents the conventional main control loop with $\boldsymbol{y}$ as the primary system measurements. The regulated states are kept within their prescribed limits using limit management controllers such as the min-max limiter logic [16].

The VGCs ($\boldsymbol{u}^{aux}$) on the other hand are set via fixed (open-loop) gain schedules against the system outputs or flight parameters σ (such as altitude, mach number (Mn) and temperature) [13], [14]. These gain schedules are designed for the worst case degradation condition resulting in large efficiency margins and increased system life cycle costs during actual system operation.

Assumption 2. Secondary system measurements $\boldsymbol{y}^p$, that reflect changes in the system performance characteristics mainly due to degradation are assumed to be available. These measurements are normally used for engine health monitoring to schedule maintenance actions and are hitherto not used for control [1]. Additional measurements that provide limitations for GTE safety and stability $\boldsymbol{g}^p$, are equally assumed to be available. These limits are calculated through a standard design practice to 'stack' uncertainties (actuation and sensing errors, operational uncertainties e.t.c.) into safety margins for the main control loop [17].

The CBC challenge within the current GTE architecture is to use the secondary system measurements $\boldsymbol{y}^p$ in addition to the primary measurements $\boldsymbol{y}$ for control decisions such that:

- The system maintains the desired thrust response control i.e $\boldsymbol{y}_k \rightarrow \boldsymbol{y}_{ref}$ as $k \rightarrow \infty$.
- The system performance measurements are optimised subject to the gradual engine degradation i.e $min \sum_{n=k}^{\infty} \boldsymbol{y}^p_n$.

- The system safety/stability is guaranteed i.e the measurements $\boldsymbol{g}_k^p \leq$ specified limits, $\boldsymbol{G}_{limits} \quad \forall k$.

The VGCs are known to have a large effect on the system performance such as fuel consumption [18], [19] and are envisaged to increase and provide extra degrees of freedom. A candidate solution approach to the CBC problem is therefore to devise a feedback tuning strategy for $\boldsymbol{u}^{aux}$ in place of their conventional fixed gain scheduling by solving a performance optimisation problem as:

$$\begin{aligned} \boldsymbol{u}_k^{*aux} &= \arg\min \sum_{n=k}^{\infty} \boldsymbol{y}_n^p \\ \textit{subject to: } \boldsymbol{x}_{k+1} &= F(\boldsymbol{x}_k, \boldsymbol{u}_k, \boldsymbol{d}_k), \quad \boldsymbol{y}_k = c(\boldsymbol{x}_k) \\ u_k^{main} &= h(\boldsymbol{y}_k), \quad \boldsymbol{g}_k^p \leq \boldsymbol{G}_{limits} \end{aligned} \tag{2}$$

Solving (2) is difficult due to the unknown $F(\cdot)$ in (1). A standard system identification approach will result in the nonlinear Hamilton-Jacobi Bellman (HJB) equations which are often impossible to solve analytically [20]. RL solves the problem by not requiring models of the system but incrementally improves the desired control performance using the secondary performance measurements. The proposed solution approach is given in the next section.

III. PROPOSED RL-ADP SOLUTION

RL problem is concerned with optimising the expected value of some desired cost through a sequence of observations, actions and rewards over time [10]. Practical methods for solving the RL problems have been based on approximate dynamic programming (ADP) which are able to solve the sequence of operations using dynamic programming and function approximations [21], [22]. For the formulated GTE condition-based control problem, let the desired cost to be optimised at discrete time steps j be given as:

$$Q(\boldsymbol{x}_j, \boldsymbol{u}^{aux}(x_j)) = \sum_{n=j}^{N} \lambda^{n-j} R(\boldsymbol{x}_n, \boldsymbol{u}_n) \tag{3}$$

where N is the discrete time interval considered for optimisation, $\lambda \in [0, 1]$ is a discount factor and $R(\boldsymbol{x}, \boldsymbol{u})$ is the observed scalar reward measurement assumed to be the system performance measurements $\boldsymbol{y}^p$. Function approximation for the cost is given as:

$$Q(\boldsymbol{x}_j, \boldsymbol{u}^{aux}(\boldsymbol{x}_j); \Theta) \approx \sum_{n=j}^{N} \lambda^{n-j} \boldsymbol{y}_n^p \tag{4}$$

In RL literature, this is known as the approximated state-action value function or *Q-function*. Learning is achieved by minimising the temporal-difference (TD) error and using a recursive relationship known as the Bellman equation [11]:

$$\begin{aligned} e_j &= \sum_{n=j}^{N} \lambda^{n-j} \boldsymbol{y}_n^p - Q(\boldsymbol{x}_j, \boldsymbol{u}^{aux}(\boldsymbol{x}_j); \Theta) \\ &= \boldsymbol{y}_j^p + \lambda^j Q(\boldsymbol{x}_{j+1}, \boldsymbol{u}^{aux}(\boldsymbol{x}_{j+1}); \Theta) - Q(\boldsymbol{x}_j, \boldsymbol{u}^{aux}(\boldsymbol{x}_j); \Theta) \end{aligned} \tag{5}$$

A batch or recursive least squares solution is determined for the parameters of the Q-function at each time step for Θ using the TD error. This can be cast into a Kalman Filtering problem with the additional advantage of compensating for time varying parameters and measurement noise assumed to be zero-mean. Online approximation of the Q-function using the system performance measurements corresponds to determining the desired operating points for the GTE subject to the gradual engine degradation and variations. This RL-ADP framework therefore belongs to the class of critic-only policy iteration algorithms where the Q-function parameters are adapted to recursively solve the Bellman equation and thereafter used to prescribe a near-optimal policy [11].

On convergence of the Q-function parameters, an optimisation sub-problem is solved for the VGC set-points update and constitutes a policy update step [10]. In contrast to the conventional Q-learning policy update, a constrained optimisation problem that guarantees the GTE safety/stability limits is solved as:

$$\begin{aligned} \boldsymbol{u}^{*aux}(\boldsymbol{x}_j) &= \arg\min Q(\boldsymbol{x}_j, \boldsymbol{u}^{aux}(\boldsymbol{x}_j); \Theta) \\ \text{subject to: } & \boldsymbol{g}_j^p \leq \boldsymbol{G}_{limits} \end{aligned} \tag{6}$$

In order for the RL-ADP update framework to fit into the overall GTE control architecture, a dual-loop control structure shown in Fig. 1 is considered, where the conventional main loop regulates the fuel flow while the RL-ADP loop continually updates the VGC set-points in the optimisation sub-problem. Transient interaction between the two control

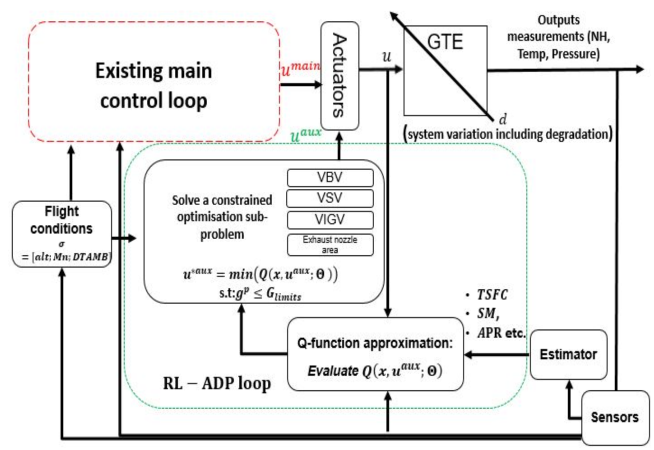


Fig. 1: Block diagram of the RL-ADP dual-control loop for GTE condition-based control. The existing main control loop guarantees the thrust response control while the RL-ADP control loop continually updates the VGC set-points.

loops is minimised by triggering the RL-ADP adaptation only at steady-state operating conditions where the most benefits in fuel savings is achievable [2]. Algorithm 1 gives the overall template for the RL-ADP CBC framework.

Algorithm 1 RL-ADP for GTE condition-based control

1: Initialise the Q-function model parameters Θ_0

Main control loop: at discrete flight time steps k:

2: Existing controller computes $u_k^{main} = h(\boldsymbol{y}_k)$ while the VGC set-points i.e $\boldsymbol{u}^{aux}(\boldsymbol{x}_k)$ are kept fixed till the next update.

RL-ADP loop: triggered at steady-state intervals

Q-function update step for $j = k$ till convergence

3: Apply $\boldsymbol{u}^{aux}(\boldsymbol{x}_j)$ and obtain measurements for $\boldsymbol{y}_j^p$, $\boldsymbol{g}_j^p$, $\boldsymbol{x}_j$ and $\boldsymbol{x}_{j+1}$.

4: Compute the TD error from (5), and solve the least squares solution for Θ_{j+1}.

VGC set-points update

5: Solve a constrained optimisation sub-problem in (6) using the updated steady-state Q-functions.

6: Repeat steps 2 to 5 till end of flight.

IV. SIMULATION OF THE PROPOSED RL-ADP CBC FRAMEWORK

The proposed RL-ADP scheme is demonstrated on representative GTE data sets in MATLAB/SIMULINK environment. The data sets are cruise data from Roll-Royce RB3039-06B model for different synthesised degradation conditions between cycle 0 as nominal and cycle 3000 as fully degraded. Inputs to the system are given as the fuel flow (WFE) for the main control variable u^{main} and two sets of variable inlet guide vanes (VIGV) for the auxiliary control variables $\boldsymbol{u}^{aux}$: the high pressure (HP VIGV) and intermediate pressure (IP VIGV). WFE is allowed to vary between $\pm 2.5\%$ of its nominal value at cruise and in steps of 0.5% while the IP and HP VIGV vary in steps between -6.67 to 14 and -7.5 to 25 degrees respectively. System performance measurements $\boldsymbol{y}^p$ and $\boldsymbol{g}^p$ that reflect changes in the system health due to degradation are available in the data sets. These are given as the thrust specific fuel consumption (TSFC), surge margin (SM) and various air pressure ratio (APR) measurements.

Based on Assumption 1, the main control loop computes the required WFE settings and guarantees the thrust response control (i.e. pre-stabilised) with $u^{main} = h(\boldsymbol{y})$. Similarly, fixed gain schedules for the VGCs are designed for the worst case degradation condition. Fig. 2 shows the offline static variations of the system performance measurements with the control inputs (WFE, IP and HP VIGV) for different degradation cycles. Representative of the conventional design approach, fixed VIGV set-points are then scheduled against the steady-state WFE settings ($WFE_{min} : WFE_{max}$) at the worst degradation cycle (i.e. cycle 3000) that satisfy the system constraints and are given in Fig. 4. Clearly, fixing the VIGV angles for the worst degradation condition will lead to increased fuel consumption at the other conditions. The formulated RL-ADP scheme is then applied to continually adapt the VIGV gains as the engine degrades using the system performance measurements as the reward signals.

Algorithm implementation

In order to initialise the Q-function model parameters for the system performance measurements, second-order quadratic polynomials were fitted to the offline test engine data as:

$$Q(\boldsymbol{x}, \boldsymbol{u}^{aux}) \approx \Theta^\top \Phi(\boldsymbol{z}); \quad \Theta \in \mathbb{R}^p \tag{7}$$

with $\Phi(\boldsymbol{z}) = \begin{bmatrix} WFE^2 & IP^2 & HP^2 & WFE & IP & HP & 1 \end{bmatrix}$. These were found to give a cross-validated R^2 test statistic of 0.94 negating the need to investigate more complex basis function. The least squares estimates for the Q-function parameters in Algorithm 1 is cast as a Kalman Filter (KF) parameter estimation problem modelled as:

$$\Theta_{j+1} = \Theta_j + w_j; \quad w_j \sim \mathcal{N}(0, \mathcal{Q}) \tag{8}$$

where (8) assumes a random walk model for the parameters. The TD error from (5) becomes:

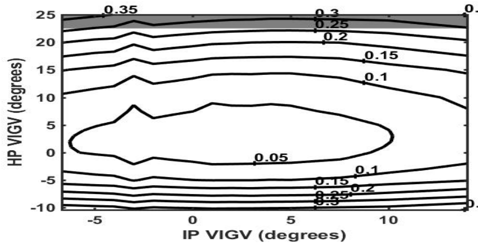
$$e_j = \boldsymbol{y}_j^p + \Theta_{j+1}^\top \big(\lambda \Phi(\boldsymbol{z}_{j+1}) - \Phi(\boldsymbol{z}_j)\big) \quad e_j \sim \mathcal{N}(0, \mathcal{R}) \tag{9}$$

$\mathcal{Q}$ and $\mathcal{R}$ are respectively the process and the measurement noise co-variance matrices. The KF parameter estimation operates in a cycle of predict-correct stages as follows:

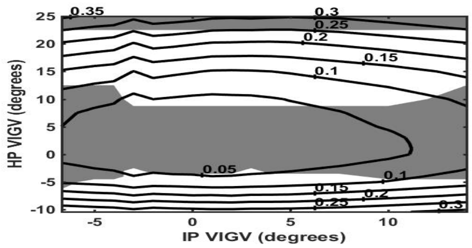
$$\begin{gathered} \Theta_{j+1}^- = \Theta_j; \quad P_{j+1}^- = P_j + \mathcal{Q} \\ K_{gain} = P_{j+1}^- \Phi(\boldsymbol{z_j})^\top \big(\Phi(\boldsymbol{z}_j) P_{j+1}^- \Phi(\boldsymbol{z}_j)^\top + \mathcal{R}\big)^{-1} \\ \Theta_{j+1} = \Theta_{j+1}^- + K_{gain} e_j; \quad P_{j+1} = \big(I - K_{gain}\Phi(\boldsymbol{z}_j)\big) P_{j+1}^- \end{gathered} \tag{10}$$

where Θ_{j+1}^- and P_{j+1}^- are respectively the predicted parameter and error co-variance estimates, K_{gain} is the Kalman Filter gain, while Θ_{j+1} and P_{j+1} are respectively the parameter and error co-variance updates. The matrix $\mathcal{Q}$ is selected as $8e^{-8}$ for the slowly varying efficiency measurements due to degradation while $\mathcal{R}$ was selected as $4e^5$ for noisy measurements. This is done till convergence of the parameters and constitutes the Q-function update step.

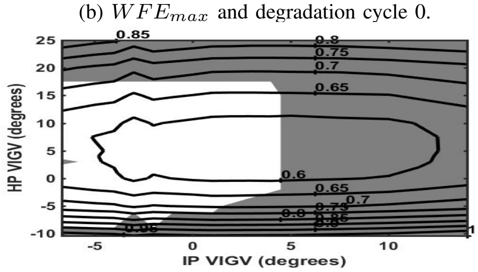
A nonlinear constrained optimisation problem is solved for the VGC set-points update step described in Algorithm 1. Due to the computational complexity of gradient based optimisation methods, an adapted direct search method from [23] called 'constrained scan and zoom' was used. This is a derivative free method which executes disciplined search for points around the current iterate using the adapted Q-functions, and systematically proceeds to points where the objective function value is reduced and satisfies constraints. The set-points for the VGCs are then updated to the identified optimal points and the process is continued till end of flight. Fig. 3 shows snapshots of the adapted online Q-functions and the identified set-points, representative of the actual (but assumed unknown) TSFC and constraint variations at the steady-state. Fig. 5(a) and Fig. 5(b) show the identified VIGV angles from the algorithm as the engine undergoes step changes in degradation from cycle 0 to 3000 while Fig. 5(c) shows the achieved fuel consumption as compared with their conventional fixed gains. This resulted in fuel savings



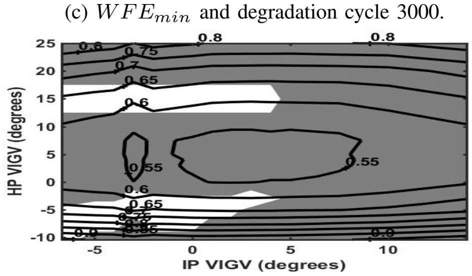
(a) WFE_{min} and degradation cycle 0.



(b) WFE_{max} and degradation cycle 0.

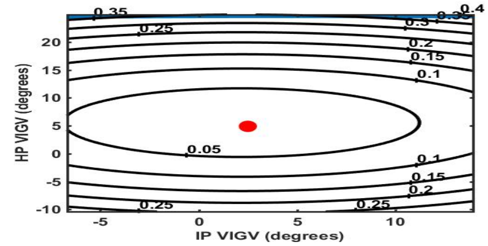


(c) WFE_{min} and degradation cycle 3000.

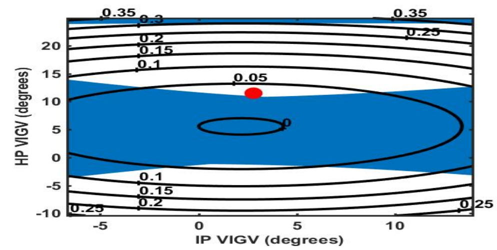


(d) WFE_{max} and degradation cycle 3000.

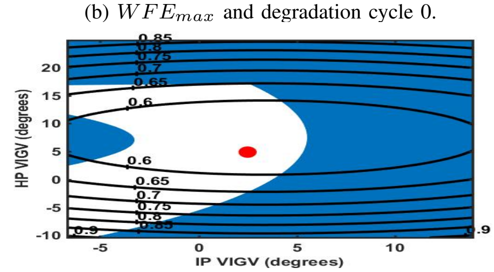
Fig. 2: Contour plots showing the variation of TSFC as functions of IP and HP VIGV, at two sample steady-state WFE settings (WFE_{min} and WFE_{max}) and degradation cycles 0 and 3000. The shaded regions indicate infeasible regions of operation due to safety/stability constraints.



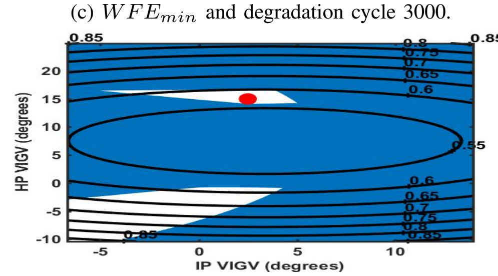
(a) WFE_{min} and degradation cycle 0.



(b) WFE_{max} and degradation cycle 0.



(c) WFE_{min} and degradation cycle 3000.



(d) WFE_{max} and degradation cycle 3000.

Fig. 3: Adapted Q-function of the system performance measurements as functions of IP and HP VIGV, at two sample steady-state WFE settings and degradation cycles 0 and 3000. The shaded regions indicate infeasible regions of operation due to the constraints, while the red dots represent sample identified optimal points from the RL-ADP scheme.

of about 0.6% at the early degradation stages. The later degradation stages correspond to the worst case for the initial set-points design where the algorithm slowly converges to. As 1% of cruise TSFC can be worth about $\$150,000$ per year on a four-engined civil aircraft [2], the proposed RL-ADP framework therefore leads to a simple, yet effective and practical means of improving the performance of GTEs across fleets subject to unknown degradation patterns and using only measurements of desired reward signals.

V. CONCLUSIONS

Conventional control approaches are unable to compensate for gradual degradation affecting system performance. Consequently, this paper has proposed and demonstrated the suitability of a RL framework for the condition-based control of GTEs to extract improved performance due to the unknown variations and degradation. A proposed dual-loop control structure which is essential to providing a potential route to certification for the overall framework integrates the RL adaptations into the existing controller structure. Simulation results on representative data sets delivered improved fuel consumption to the GTE as compared to the conventional static scheduling by adapting to through life degradation and variations.

REFERENCES

[1] J. S. Litt, D. L. Simon, S. Garg, T.-H. Guo, C. Mercer, R. Millar, A. Behbahani, A. Bajwa, D. T. Jensen, A survey of intelligent control and health management technologies for aircraft propulsion systems,

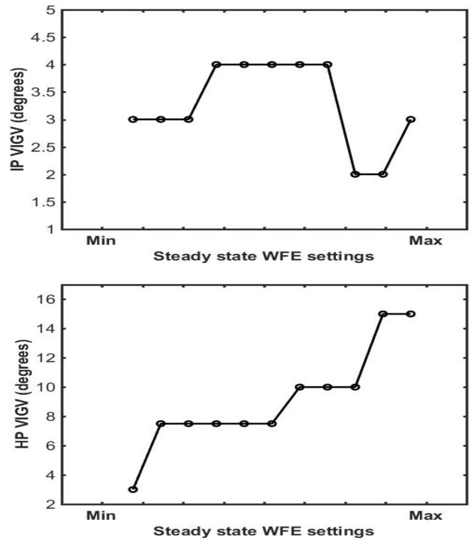
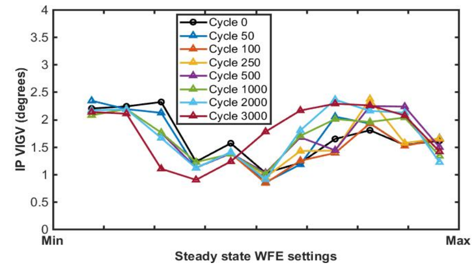
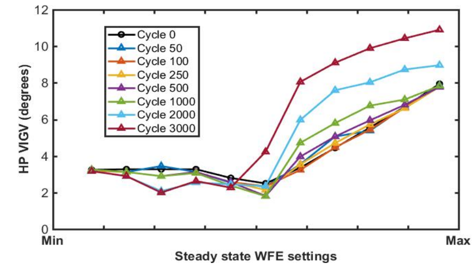


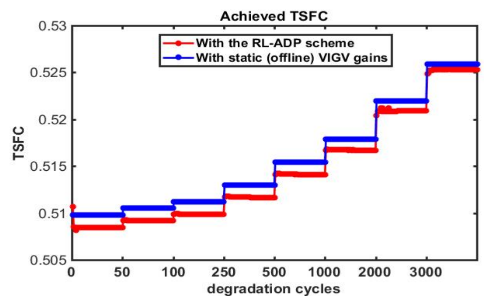
Fig. 4: Fixed schedules for the IP and HP VIGV angles designed at the worst case system condition to satisfy system safety/stability constraints.



(a) Adapted IP VIGV set-points through degradation cycles 0 to 3000



(b) Adapted HP VIGV set-points through degradation cycles 0 to 3000



(c) Comparison of the achieved $TSFC$ using the RL-ADP condition-based control framework with the conventional fixed VIGV gains.

Fig. 5: Identified optimal VIGV operating points in the simulation of the RL-ADP for GTE condition-based control.

Journal of Aerospace Computing, Information, and Communication 1 (12) (2004) 543–563.

[2] Rolls-Royce, The Jet Engine, 5th Edition, Rolls-Royce Technical Publications, London, 2005.

[3] M. Bohlin, K. Doganay, P. Kreuger, R. Steinert, M. Warja, Searching for gas turbine maintenance schedules, AI Magazine 31 (1) (2010) 21–36.

[4] R. Smith, J. Chisholm, J. Stewart, Optimizing aircraft performance with adaptive, integrated flight propulsion control, Journal of engineering for gas turbines and power 113 (1) (1991) 87–94.

[5] J. S. Orme, G. B. Gilyard, Subsonic flight test evaluation of a propulsion system parameter estimation process for the f100 engine, NASA Technical Memorandum 4426.

[6] G. B. Gilyard, J. S. Orme, Performance seeking control: program overview and future directions, NASA Technical Memorandum 4531.

[7] J. S. Orme, T. R. Conners, Supersonic flight test results of a performance seeking control algorithm on a NASA f-15 aircraft, AIAA Paper (1994) 94–3210.

[8] D. E. Viassolo, S. Adibhatla, B. J. Brunell, J. H. Down, N. S. Gibson, A. Kumar, H. K. Mathews, L. D. Holcomb, Advanced estimation for aircraft engines, in: 2007 American Control Conference, IEEE, 2007, pp. 2807–2821.

[9] R. Hafner, M. Riedmiller, Reinforcement learning in feedback control, Machine learning 84 (1-2 (2011) 137–169.

[10] R. S. Sutton, A. G. Barto, Reinforcement learning: An introduction, Vol. 1, MIT press Cambridge 1998.

[11] F. L. Lewis, D. Vrabie, K. G. Vamvoudakis, Reinforcement learning and feedback control: Using natural decision methods to design optimal adaptive controllers, IEEE Control Systems 32 (6) (2012) 76–105.

[12] B. D. Anderson, A. Dehghani, Challenges of adaptive control–past, permanent and future, Annual reviews in control 32 (2) (2008) 123–135.

[13] A. Linke-Diesinger, Systems of commercial turbofan engines: an introduction to systems functions, Springer Science & Business Media, 2008.

[14] V. V. Silva, W. Khatib, P. J. Fleming, Performance optimization of gas turbine engine, Engineering Applications of Artificial Intelligence 18 (5) (2005) 575–583.

[15] H. A. Spang III, H. Brown, Control of jet engines, Control Engineering Practice 7 (9) (1999) 1043–1059.

[16] H. Richter, Advanced control of turbofan engines, Springer Science & Business Media, 2011.

[17] Eckert, C. & Isaksson, O., Safety Margins and Design Margins: A Differentiation between Interconnected Concepts, Procedia CIRP, Elsevier, 2017, 60, 267-272

[18] R. Kurz, K. Brun, Degradation in gas turbine systems, Journal of Engineering for Gas Turbines and Power 123 (1) (2001) 70–77.

[19] C. Bringhenti, J. Barbosa, Methodology for gas turbine performance improvement using variable-geometry compressors and turbines, Proceedings of the Institution of Mechanical Engineers, Part A: Journal of Power and Energy 218 (7 (2004) 541–549.

[20] F. L. Lewis, K. G. Vamvoudakis, Optimal adaptive control for unknown systems using output feedback by reinforcement learning methods, in: Control and Automation (ICCA), 2010 8th IEEE International Conference on, IEEE, 2010, pp. 2138–2145.

[21] F. L. Lewis, D. Vrabie, Reinforcement learning and adaptive dynamic programming for feedback control, IEEE circuits and systems magazine 9 (3) (2009) 32–50.

[22] L. Busoniu, R. Babuska, B. De Schutter, D. Ernst, Reinforcement learning and dynamic programming using function approximators, Vol. 39, CRC press, 2010.

[23] P. Venkataraman, Applied optimization with MATLAB programming, John Wiley & Sons, 2009.