Flight Testing Reinforcement-Learning-Based Online Adaptive Flight Control Laws on CS-25-Class Aircraft

Konatala, Ramesh; Milz, Daniel; Weiser, Christian; Looye, Gertjan; van Kampen, E.

DOI
10.2514/1.G008321

Publication date
2024

Document Version
Final published version

Published in
Journal of Guidance, Control, and Dynamics

JOURNAL OF GUIDANCE, CONTROL, AND DYNAMICS
Vol. 47, No. 11, November 2024

Technical Notes

Flight Testing Reinforcement-Learning-Based Online Adaptive Flight Control Laws on CS-25-Class Aircraft

Ramesh Konatala,* Daniel Milz,†
Christian Weiser,‡ and Gertjan Looye§
German Aerospace Center (DLR), 82234 Weßling, Germany
and
E. van Kampen¶
Delft University of Technology, 2626HS Delft, The Netherlands

https://doi.org/10.2514/1.G008321

I. Introduction

LOSS of control in-flight (LOC-I) is an off-nominal flying condition where the aircraft deviates from the normal flight envelope and is a leading cause of accidents in commercial aviation [1]. With the increasing trend toward autonomous and complex systems, one can expect an increasing trend of such LOC-I incidents unless proactive measures are taken. Developing an integrated fault-tolerant resilient flight control law (FCL) is imperative to enhance safety under off-nominal conditions, addressing parametric failures and abnormal flight scenarios. Main challenges in designing such a controller include low confidence in the aircraft model post failure, which degrades the model-dependent controller performance, nonlinearities in the model post failure, and the need for rapid adaptation of the controller to restore the aircraft within the safe flight envelope.

Self-learning adaptive flight control system (FCS) algorithms were initially tested in the 1960s on the X-15 research aircraft [2]. Some of the open challenges in realizing adaptive FCS include sample efficiency and convergence, controller robustness, and interpretability of the controller's adaptive mechanism [3]. Reinforcement learning (RL), a bio-inspired machine learning approach, has been used for adaptive flight control since the early 2000s, e.g., the work from Enns and Si on helicopter control using neuro-dynamic programming [4] or the work from Ferrari and Stengel on applying RL for control of a business jet type of aircraft [5]. An advantage of RL is that it can be used as a model-free controller, meaning that no information about the plant that is to be controlled has to be known before the start of training. Another advantage is that it is by definition an adaptive controller, and hence can be used when online adaptation is required, for example after a fault or failure of a part of the aircraft. More recent applications of RL to flight control can be seen in [6,7], where RL controllers are designed that make use of an incremental model of the plant, which is identified online. Nevertheless, RL's application in flight control has thus far been limited to validating control laws on models of the target system rather than real aircraft. A key challenge that limits the adoption of RL-based methods for flight control is validation through flight testing on a CS-25-class aircraft, which could aid in the certification of RL-based FCL for fault tolerance [8].

Although several variants of RL-based FCL's were developed [7,9], practical verification and validation (V&V) constraints guided the choice of this RL-based FCL design. The following desirable features for the control method are defined as follows: a simpler control design strategy with fewer and more interpretable learning parameters, the ability to adapt in real-time to fast-changing nonlinear dynamics of aircraft in case of a failure, an algorithm that is sample efficient [10,11] and fast to converge on the conventional flight control computer (FCC). In this regard, Incremental Approximate Dynamic Programming (iADP)-based FCS [12] is a good choice among the RL-based FCS with practical interests. This algorithm identifies a local linearized incremental model online to estimate and minimize an infinite horizon quadratic cost-to-go, exclusively using the collected aircraft state data [13–17]. The iADP algorithm demonstrated effective online adaptation for a F-16 aircraft model, with good tracking performance both in normal and failure conditions [18].

The main contributions of this paper are an RL-based iADP FCS design for a CS-25-class aircraft and a validation of this RL-based FCS design through flight tests on the PH-LAB Research Aircraft shown in Fig. 1. The outcome from the flight test campaigns is detailed, viz., the ability of the controller to capture pitch and roll rate tasks without a priori knowledge of aircraft model or any pretraining of the controller, stable continuous learning of the controller, adaptability of the controller to aircraft configurations assessed by comparing the adapting parameters across different configurations, and finally a discussion on observed challenges of some flight test trials. The technical scope of this study is limited to V&V, excluding the interaction of the pilot with the adaptive system.

The structure of the paper is organized as follows: Section II contains the iADP FCL Architecture Design for the Cessna Citation-II Aircraft, FCL evaluation strategy, software and hardware Integration of the controller, and a summary of V&V procedures for the FCL clearance for flight tests. Section III delves into the validation of controller functionality, focusing extensively on the results obtained from the flight test campaign. Lastly, Sec. IV includes concluding remarks and a note on how this research could aid aviation safety. A detailed explanation of the methodology behind the iADP algorithm, its extensions for flight control, and a discussion on flight testing experiences, challenges, features, and limitations of the iADP FCL is provided in [19]. The process of clearing this online adaptive FCL for flight testing on CS-25-class Citation-II aircraft, including V&V, using various tools, methods, and a framework is detailed in [20], which underpins this work.

II. The iADP-Based Flight Control System Design

A. Control Law Architecture

The iADP algorithm is used to design the FCL for the inner loop, tracking desired pitch and roll rate commands using three control surfaces: aileron, elevator, and rudder, as shown in Fig. 2. As this FCL is solely sensor-based, signal processing of sensor measurements is performed to reduce the impact of noisy sensor signals. Smooth

Received 4 March 2024; revision received 29 June 2024; accepted for publication 8 July 2024; published online 27 August 2024.

*Research Associate, Department of Aircraft System Dynamics; also Ph.D. Researcher, Faculty of Aerospace Engineering, Delft University of Technology, P.O. Box 5058, 2626 HS Delft, The Netherlands; ramesh.konatala@dlr.de.

†Research Associate, Department of Aircraft System Dynamics; daniel.milz@dlr.de.

‡Research Associate, Department of Aircraft System Dynamics; christian.weiser@dlr.de.

§Head of Department, Department of Aircraft System Dynamics; gertjan.looye@dlr.de.

¶Associate Professor, Faculty of Aerospace Engineering, Control and Simulation Division, P.O. Box 5058; E.vanKampen@tudelft.nl.

Fig. 1 Cessna Citation II (PH-LAB) research aircraft captured by Alan Wilson. Image licensed under CC BY-SA 2.0.

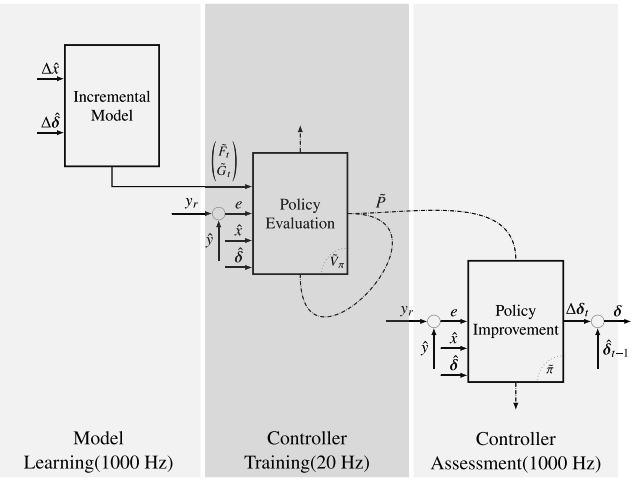


Fig. 3 Structure of the reinforcement learning agent of iADP flight control law (FCL).

sensor measurements are obtained by low-pass filtering relevant aircraft states and actuator position measurements. Aerodynamic angles, namely, angle of attack and sideslip angle, are acquired through a boom with attached vanes on the aircraft. Complementary filtering of these angles is executed by combining them with inertial-reference sensors. The signal processing block does not contain any knowledge of the aircraft model, ensuring a model-free and aircraft-independent inner loop control structure. Pilot-in-the-loop studies are excluded from this flight control design, and thus, the desired pitch and roll rate commands are generated automatically. However, the iADP controller is not aware of these reference commands a priori.

Direct control over the aircraft's attitude angles is absent, as the cost function does not consider attitude angle errors, as only the one-step rate errors are fed back as the reward signal. Additionally, airspeed and altitude information is excluded from the state vector due to slower local variations, which could impact incremental model identification and subsequent value function approximation. Decoupled longitudinal and lateral controllers are designed: the longitudinal control loop tracks a pitch rate command using the elevator, while the lateral control loop tracks a roll rate command using the aileron and/or rudder. Only one axis is controlled during a maneuver, which means during the evaluation of the longitudinal task, aileron and rudder maintain the trimmed control input and vice versa.

B. The iADP Control Law Evaluation Strategy

The iADP controller operates in real-time through three phases: model learning, controller training, and controller assessment, as illustrated in Fig. 3. Model learning provides the latest model estimates using the RLS algorithm. Controller training evaluates (V_π) the control policy using incremental model estimates and one-step cost. Controller assessment takes actions and improves control policy (π)-based on policy evaluation. The frequency at which each subsystem on the flight control computer runs is indicated at the bottom.

During the model learning phase, the controller estimates incremental model parameters using the RLS approach. These estimates are fed to the controller training phase, where, using the latest available model estimates and control policy, the value function (cost-to-go) estimate is improved. This value function provides a measure of goodness of the underlying control policy. To enable smoothness in parameter update, the value function update is done batchwise considering data over a window of "x" number of samples. The optimal window length is determined from offline simulation analysis and based on how much computational load the FCC could handle using HIL ground tests. The controller assessment phase evaluates the controller against a commanded reference signal using the converged control parameters obtained from the controller training phase.

Two experimental approaches are considered for controller evaluation, as shown in Figs. 4a and 4b:

1) *Sequential learning approach (SLA)*: This approach runs each phase sequentially, starting with the model learning phase, followed by the controller training and controller assessment phases. The model learning phase is open loop and lasts for 20–25 s.

2) *Continuous learning approach (CLA):* In this approach, all three phases run concurrently. The controller updates its policy at every time step using the latest model parameters from the model learning loop and controller parameters from the controller training loop. The controller assessment loop runs concurrently, allowing the controller to update the model and control along the commanded trajectory defined for the controller assessment. To facilitate the learning process, the model learning phase operates in an open loop for the first 20–25 s of the trial. Subsequently, it adjusts the model parameters, taking into account the influence of the controller in the loop.

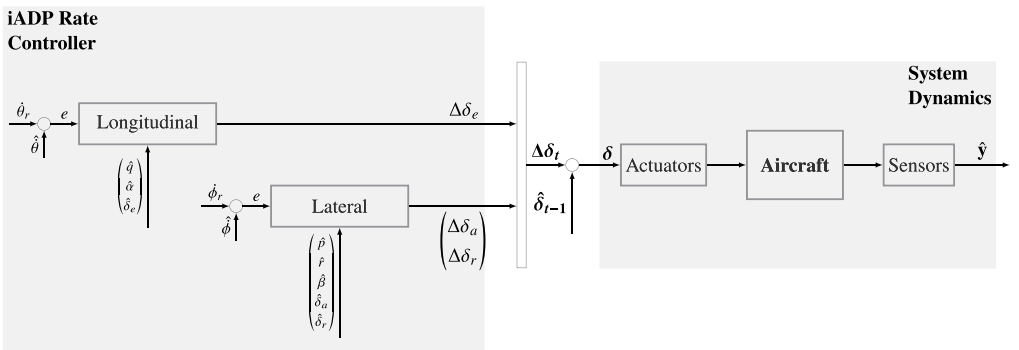
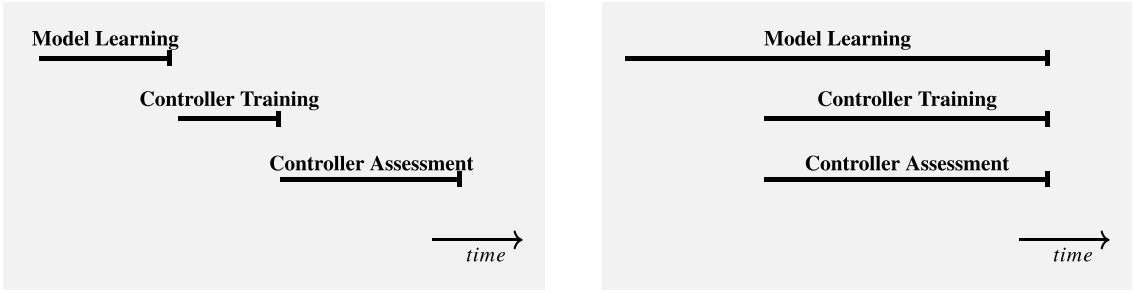


Fig. 2 Incremental approximate dynamic-programming-based flight control law (FCL) architecture for inner-loop rate tracking.



a) Sequential Learning **b) Continuous Learning**

Fig. 4 Comparison of controller learning approaches.

III. Flight Tests

Controller validation involves assessing the performance of the FCL's against predefined criteria. iADP FCL validation occurred through flight tests on the PH-LAB research aircraft, conducted in November 2022 and August 2023, departing from Rotterdam The Hague Airport.

A. Experimental Objectives and Setup

Flight control design specifications act as a guide for control engineers, directing them in the design process to verify that the controller meets specific criteria [21]. For iADP FCL validation, the following FCS design requirements are formulated to ensure the controller commands the aircraft to follow pitch and roll rates, rate tracking in different operating conditions (altitude and velocity changes), consistent controller behavior under similar conditions (reproducibility), stable continuous learning over longer maneuvers, and rate tracking through parameter adaptation in different aircraft configurations.

The dimensions of the available airspace for the tests were decided preflight, in consultation with pilots. Operating conditions for which the aircraft equipment is certified, in consultation with technicians, provided estimates on the boundaries of the flight envelope. These factors determined the higher limit on the trial's duration. The lower limit on trial duration was determined by the minimum time/samples required for the iADP algorithm to converge, determined through offline desktop simulations and HIL tests. To test adaptability, a set of feasible aircraft configurations was decided preflight. These configurations should be observable by the iADP controller and sufficiently alter the aircraft's dynamics to assess FCL's adaptability. For these tests, the higher limit on the airspeed is imposed by the fact that landing gear and flap maximum extension configurations can be deployed only below a certain airspeed.

B. The iADP-Based Control Law Flight Performance

Table 1 provides a summary of all the flight test trials, offering a concise overview of the outcomes along with the operating condition at which the aircraft is trimmed before the beginning of the flight trial. The trials are listed chronologically, and the trial ID follows the notation: N22 for November 2022, A23 for August 2023, F# indicates the flight test day, and T# indicates the trial number. The *Axis* indicates the actively controlled channel of the aircraft via the FBW during the trial. V_{TAS} represents true airspeed, and h represents altitude. The *Outcome* column reflects whether the controller response aligned with the design specifications. *Config.* denotes aircraft configuration.

SLA is adopted from trial N22-F2-T1 to A23-F1-T2, and all subsequent trials have adopted CLA. SLA takes precedence to identify and assess converged parameters, addressing issues such as the ideal open-loop aircraft identification input signal, determining acceptable phase durations for parameter convergence, and refining the reference signals to ensure the aircraft stays within airspace limits. The controller parameters, including model parameters, are reset after each trial. Consequently, no information about the aircraft model or controller parameters from previous flight tests is retained or transferred to subsequent attempts. For stable parameter convergence in the RLS algorithm, we combine frequency-rich persistently exciting signals with the calculated control input [22]. The selected persistent excitation signal for the iADP FCL is a sinusoidal signal with a small amplitude.

Table 1 Overview of the iADP controller flight testing campaigns

Trial ID	Axis	V_{TAS}, m/s	h, m	Brief description	Outcome
N22-F2-T1	Pitch	101	3600	Oscillatory response; convergence in model prediction	✓
N22-F2-T2	Pitch	106	3650	Off-nominal flight; inverted incremental model parameters	×
N22-F2-T3	Pitch	104	3650	First success; decent tracking; slight elevator oscillations	✓
N22-F2-T4	Pitch	105	3550	Inverted controller commands; inverted model parameters	×
N22-F2-T5	Pitch	94	3500	Better tracking; increased elevator oscillations	✓
N22-F3-T1	Pitch	102	2100	Decent tracking; high model prediction error	✓
N22-F3-T2	Roll	91	2150	Oscillatory response; high model prediction error	×
N22-F3-T3	Roll	96	2000	Aircraft deviated from level flight post model learning phase	×
A23-F1-T1	Pitch	99	2750	Oscillatory response; model learning duration too short	✓
A23-F1-T2	Roll	101	2800	Deviated from level flight; model learning duration too short	×
A23-F1-T3	Pitch	102	2750	First success with continuous learning; decent tracking	✓
A23-F1-T4	Pitch	101	2800	Reproducible continuous learning; better tracking	✓
A23-F1-T5	Roll	100	2800	First success in lateral with continuous learning	✓
A23-F1-T6	Roll	101	2800	Reproducible continuous learning; good tracking response	✓
A23-F2-T1	Pitch	97	3050	Stable continuous learning; decent tracking response	✓
A23-F2-T2	Roll	101	3050	Nominal config.; reproducible continuous learning	✓
A23-F2-T3	Roll	95	3100	Nominal config.; reproducible continuous learning	✓
A23-F2-T4	Roll	97	3100	Nominal config.; reproducible continuous learning	✓
A23-F2-T5	Roll	97	3050	Landing gear down config.; stable continuous learning	✓
A23-F2-T6	Roll	97	3050	Flaps 15° config.; stable continuous learning	✓
A23-F2-T7	Roll	98	3100	Flaps 40° config.; slightly oscillatory tracking	✓

1. Longitudinal Controller Assessment: Sequential Learning Approach

The first successful trial of the iADP algorithm for longitudinal rate control is presented in Fig. 5. The left-aligned plots illustrate the model learning phase, an open-loop period where a 3211 maneuver is commanded by the elevator. The online RLS algorithm updates model parameters at each time step during this 20 s phase, with fixed parameters subsequently passed to the controller training phase. The choice of the 3211 signal is based on its proven effectiveness in previous system identification flight tests on the Citation aircraft. Control over the system's excited frequency can be achieved by adjusting the duration of individual step commands in the 3211 signal. The 3211 signal also serves the functionality of a persistently exciting signal. The forgetting factor (γ_{RLS}) is tuned using the Multi-Objective Parameter Synthesis (MOPS) tool [23]. Results comparing measured longitudinal states against predictions from the RLS algorithm show a good fit. Although model parameters seem to converge quickly, some oscillations are observed.

The right-aligned plots depict results from the controller training and assessment phase. Parameters for this phase, including the discount factor (γ) and weighting matrices (Q and R), are tuned using the MOPS tool. The controller training phase lasted for 40 s (from 20 to 60 s), during which the controller loop is closed and internally a pitch rate reference command to evaluate the policy is generated. The controller, along with model parameters estimated from the model learning phase and observed one-step error in pitch rate tracking, has to improve its estimates of the cost-to-go function. Controller parameters are updated during a brief 5 s phase (55–60 s). This computationally intensive phase updates kernel matrix parameters using data collected over a 20 s window (data from 35–55 s is used for the update from 55th second onward). This loop is running at a much lower 20 Hz due to real-time constraints. After the controller training phase concludes, the parameters are fixed and passed to the subsequent controller assessment phase.

During the controller assessment phase, an internal pitch rate command is generated, and the controller's objective is to track this reference command. The results, from 60 to 100 s, show the aircraft effectively tracking a pitch rate command, which can also be interpreted as a reduction in the cost-to-go plot. The tracking performance improved with higher values of Q, but this made the controller response more oscillatory (from flight trial N22-F2-T5 in Table 1).

2. Lateral Controller Assessment: Continuous Learning Approach

Figures 6 and 7 depict the flight test results from trial A23-F2-T2, focusing on achieving a roll rate tracking task with a CLA.

The controller's objective is to command the aircraft to follow a roll rate reference and demonstrate stable continuous learning capability. Figure 6 illustrates the performance of the incremental model identification stage. Doublets are initially commanded at the aileron and rudder to aid the identification process. A sinusoidal signal is superimposed on commanded aileron and rudder to ensure

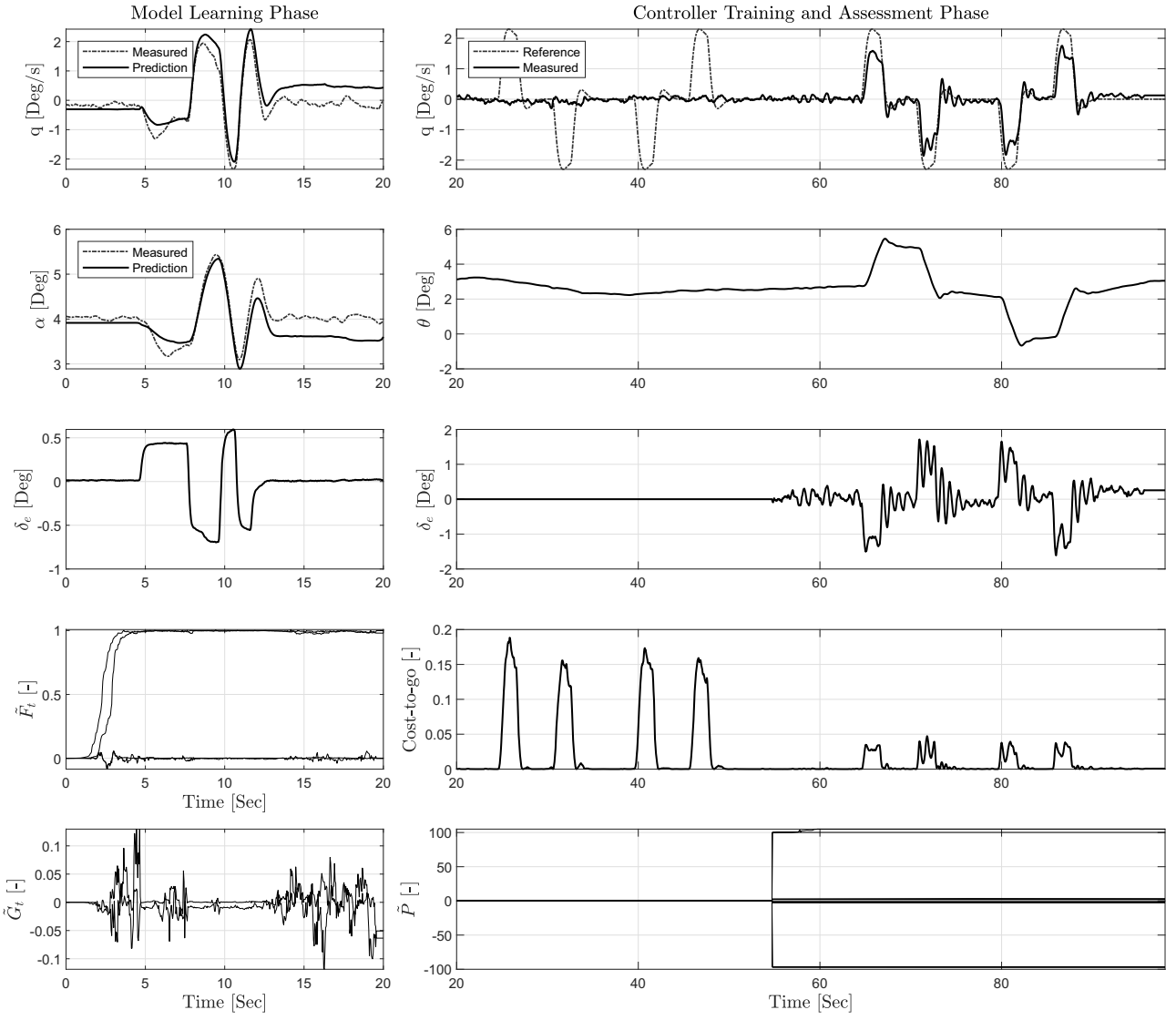


Fig. 5 Flight test data (trial ID : N22-F2-T3), PH-LAB performing a longitudinal maneuver: iADP flight control law (FCL) designed for pitch rate capture. Sequential learning approach (SLA) with fixed parameters post model learning and controller training.

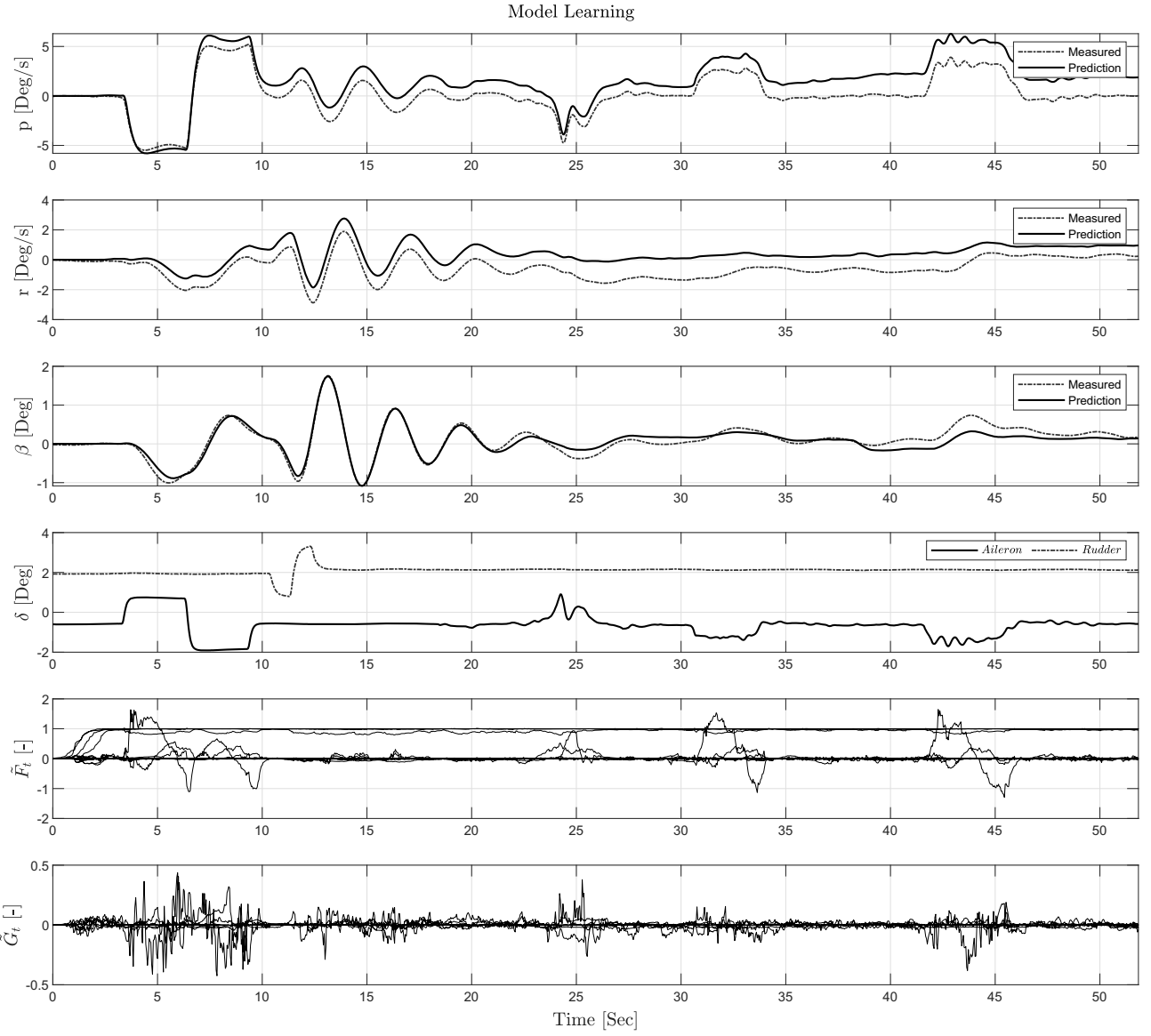


Fig. 6 Flight test data (trial ID : A23-F2-T2), PH-LAB performing a lateral maneuver: plots compare the predictions from the incremental model to the measured states. The iADP flight control law (FCL) designed for roll rate capture. Continuous learning approach (CLA) with real-time parameter adaptation.

RLS parameter convergence. When comparing the measured state variable with the prediction, the RLS algorithm appears to offer a good fit in state prediction throughout the maneuver. However, after 30 s, error in predicting certain states was observed to have an increasing trend, likely attributed to sharp control inputs occurring around the moment the control loop was closed at 20 s.

The performance of the controller training and assessment stage is presented in Fig. 7. Comparing the reference to the measured roll rate output in the first row, a good tracking response is observed. The bottom two plots show the reduction in the cost-to-go estimate and the evolution of the controller parameters, respectively.

3. *Adaptability to Different Aircraft Configurations*

To assess the adaptability of the iADP controller, five flight test trials were conducted with varying aircraft configurations. The controller's objective is to track roll rate commands using a CLA. The operating conditions (trimmed aircraft velocity and altitude) are consistent across all trials. The controller objective, approach, operating conditions, and hyper-parameters of the iADP controller are kept constant, enabling an evaluation of controller performance solely against changes in aircraft configuration. Four aircraft configurations are considered:

1) *Nominal (N):* Landing gear up, flaps completely retracted (0° extension), similar to a normal cruise flight.

2) *Landing gear down (G-D):* Landing gear down, flaps completely retracted (0° extension).

3) *Flaps 15° extension (F-15):* Landing gear up, flaps extended to 15°.

4) *Flaps 40° extension (F-40):* Landing gear up, flaps extended to 40°.

Each configuration change is introduced one at a time, allowing an assessment of adaptability to individual changes. Pilots change the aircraft configuration by deploying landing gear or extending flaps first, and the aircraft is trimmed in each configuration before engaging the iADP controller. Following each configuration change, the iADP FCL undergoes three phases: model learning, controller training, and controller assessment phase.

Two trials, designated as N-1 and N-2, were conducted in the nominal configuration, followed by the deployment of landing gear (G-D) and flap extensions (F-15 and F-40). The adaptability test aims to evaluate differences in model/controller parameters with respect to configuration changes. Figure 8 illustrates the controller's performance across these configurations. Plots show performance from controller training and assessment phase across different aircraft

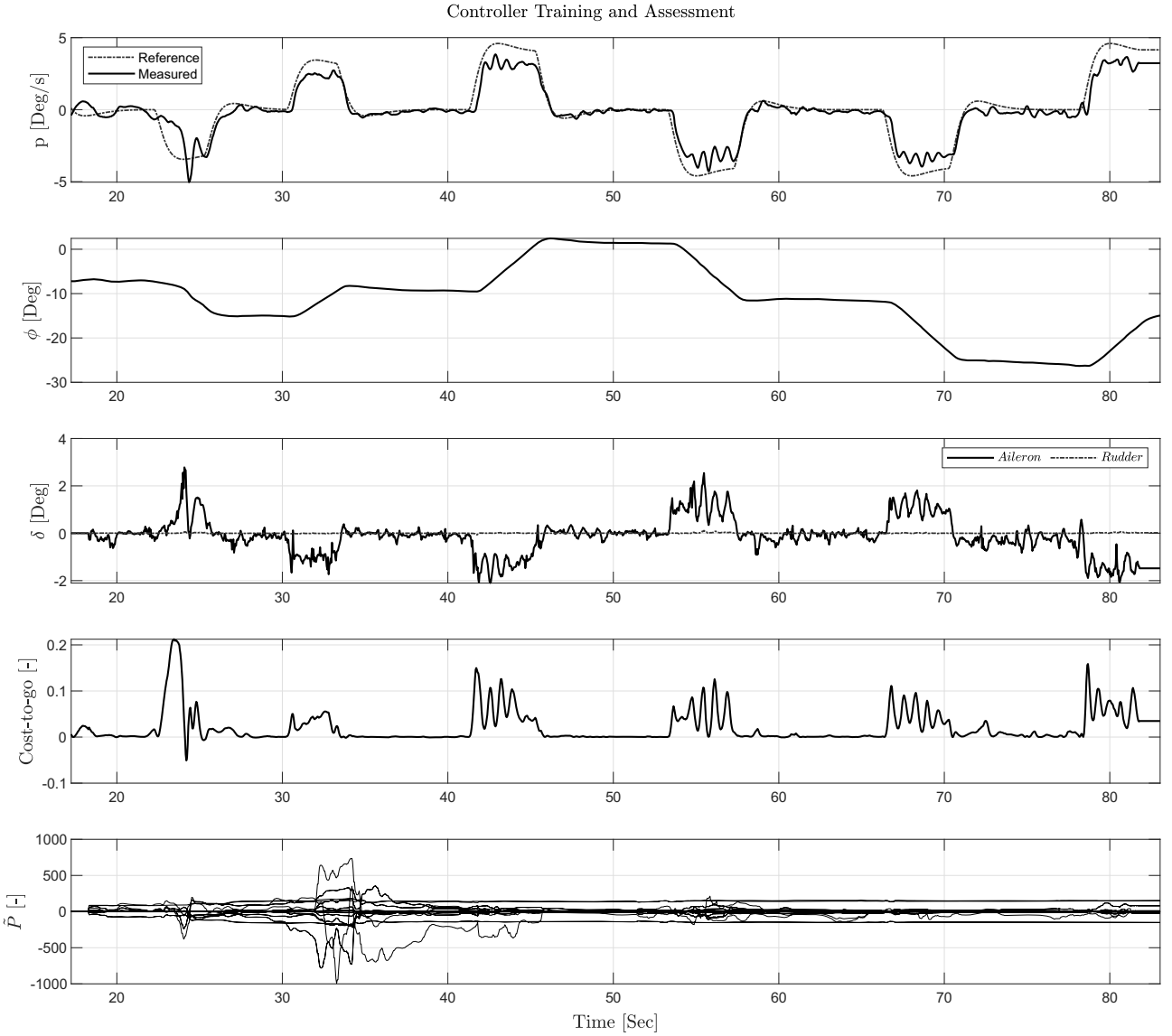


Fig. 7 Flight test data (trial ID : A23-F2-T2), PH-LAB performing a lateral maneuver: plots show performance from controller training and assessment phase. The iADP flight control law (FCL) designed for roll rate capture. Continuous learning approach (CLA) with real-time parameter adaptation.

configurations. Model and controller parameters undergo continuous updates at each time step, i.e., CLA. When comparing Nominal-1 configuration with others, the tracking performance exhibits the most significant difference with the landing gear and 40° flaps extension, while showing similar performance with Nominal-2 and 15° flaps extension configurations. The controller response appears oscillatory when flaps are at their maximum (40°) extension, likely due to the substantial alteration in the aircraft's aerodynamic properties. Flap deployment modifies the aerodynamic forces on the wings, influencing the rolling moment of the aircraft. The landing gear-down configuration also results in a different controller response, possibly due to the changes in the moment of inertia about the roll axis due to mass redistribution, consequently affecting the lateral stability characteristics of the aircraft. Aileron effectiveness is expected to be impacted by both flap extension and landing gear down configurations.

Despite the variations in aircraft configurations, it is interesting to observe that the controller, despite being unaware of these model changes, effectively guides the aircraft in the lateral axis throughout the entire maneuver. Additionally, the controller parameter updates exhibit stability throughout the entire maneuver across all four configurations.

To quantify the adaptability of the controller further, time-evolving parameters are compared against different configurations. Four different metrics are considered for comparison:

1) *Tracking error:* Evaluates controller tracking performance, assessing the control objective.

2) *Incremental model state matrix* ($\tilde{F}_t$): Measures identified incremental model parameters related to state transitions, containing state derivatives.

3) *Incremental model control effectiveness matrix* ($\tilde{G}_t$)*:* Measures identified incremental model parameters related to control effectiveness, containing control derivatives.

4) *Kernel matrix* ($\tilde{P}$): Measures learned control policy parameters.

For an accurate comparison, flight test data is aligned, and values are smoothed using a Gaussian-weighted moving average filter to remove noise artifacts. Additionally, only data from 35 to 100 s is utilized in this analysis to mitigate the impact of transients. The Frobenius norm of the difference in matrices is selected to assess the similarity of these values. Although nuclear and spectral norms are considered as alternatives for comparison, the results appear insensitive to the choice of the norm, and only Frobenius norm-based evaluation is presented. For example, the norm for comparing N-1 configuration data with flap 15 configuration is defined as follows:

$$\|\Delta P\|_F = \|P_{N1} - P_{F15}\|_F$$

The evolving matrix $\|\Delta P\|_F$, sized according to each trial's duration, represents a time-dependent parameter. A cumulative sum of this

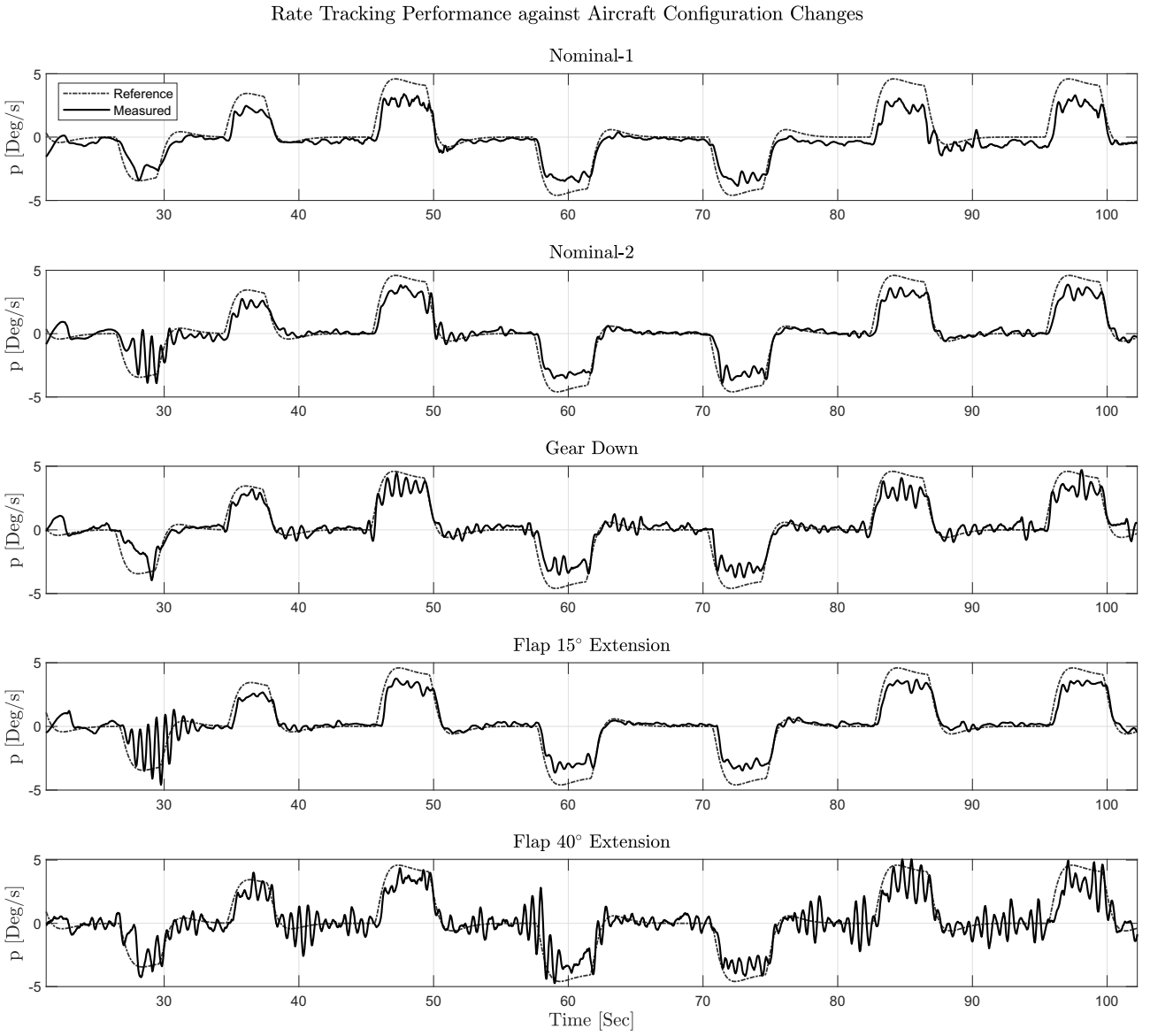


Fig. 8 Flight test data (trial ID : A23-F2-T3 to A23-F2-T7), PH-LAB performing lateral maneuvers: experimental setup to test the controller adaptability to change in aircraft configurations.

value serves as a metric indicating the similarity between P_{N1} and P_{F15} parameters. This similarity measure is depicted for various values and compared across different aircraft configurations in Fig. 9. Configuration labels denote N-1 (first trial in nominal configuration), N-2 (second trial in nominal configuration), G-D (landing gear down), F-15 (flaps 15° extension), and F-40 (flaps 40° extension)

Comparing the first column of all four plots, i.e., comparing N-1 to {N-2, G-D, F-15, and F-40}, the difference in the tracking error seems to be minimum between N-1 and N-2. However, examining the $\|\Delta\tilde{P}\|_F$ plot indicates that the similarity is least between N-1 and F-40, contrasting with the N-1 and N-2 comparison. This observation suggests that controller parameters undergo updates to accommodate

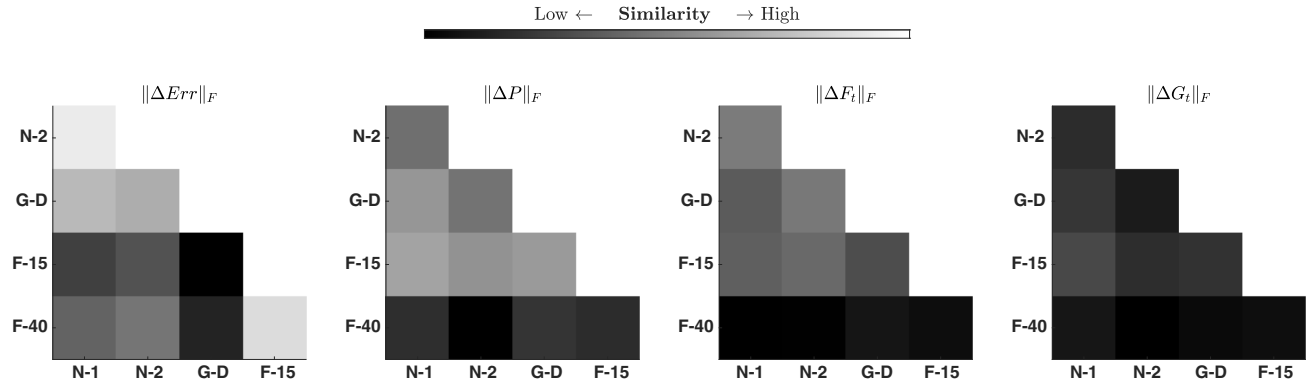


Fig. 9 Quantifying similarity index of adaptive control parameters using Frobenius norm across aircraft configurations.

aircraft configuration alterations. The $\|\tilde{F}_t\|_F$ plot indicates the greatest difference in the state derivative matrix during the F-40 configuration, as expected due to significant changes in the aircraft's aerodynamic properties during maximum flap extension. The variability of the control effectiveness matrix $\tilde{G}_t$ appears high, making comparisons challenging. But $\|\tilde{G}_t\|_F$ plot indicates the lowest similarity when comparing {N-1, N-2} with F-40, and G-D. This suggests the model learning phase adapting to identify the control derivative parameters specific to F-40 and G-D configurations. To summarize, kernel matrix parameters $\tilde{P}$ are correlated to aircraft configurations, indicating adaptation of control policy by the RL agent. State transition matrix parameters ($\tilde{F}_t$) from the Incremental Model are correlated to configurations, from high sensitivity to Flaps 40° Extension.

IV. Conclusions

This paper presents the RL-based iADP FCL design for a CS-25-class aircraft and reports the findings from the maiden flight test campaign of this controller. The flight tests validate the FCL for the stable longitudinal axis through pitch rate captures and the unstable lateral axis through roll rate captures. Postflight comparison of adaptive parameters indicates the capability of the controller to adapt to different configurations while retaining parameter interpretability. The outcome of flight tests was reviewed, encompassing discussions on challenges faced, potential improvements to the FCL, and scenarios for future fault tolerance validation through flight tests. With the aim of tackling the LOC-I problem in aircraft, this research initially focuses on designing a model-free FCL and subsequently developing the necessary toolchains to V&V these FCLs. Furthermore, the toolchains and routines developed for V&V processes could be extended to other online learning-based control laws, accelerating the development of fault-tolerant FCS. This model-free and adaptive FCS could work as a potential lifeline in flight emergencies, complementing traditional FCS and aiding aircraft in fault recovery when facing controllable failures, with online adaptive fault-tolerant control capabilities.

Acknowledgments

The authors express their gratitude to the personnel at the Delft University of Technology, including Alexander in't Veld, Hans Mulder, Menno Klaassen, Ferdinand Postema, Olaf Stroosma, René van Paassen, Fred den Toom, and Isabelle El-Hajj, for providing the framework and expertise that facilitated the successful execution of the flight tests. Special acknowledgment also goes to the DLR (German Aerospace Center) team members involved in the flight test campaign, namely, Richard Kuchar, Marc May, Reiko Müller, Stefan Langen, and Christina Schreppel, for their valuable support and expertise. A special acknowledgment goes to Carsten Oldemeyer for carefully reviewing the manuscript.

References

[1] "Loss of Control In-Flight Accident Analysis Report," Tech. Rept., International Air Transport Association (IATA), Geneva, 2019, https://www.iata.org/contentassets/b6eb2adc248c484192101edd1ed36015/loc-i_2019.pdf.

[2] Dydek, Z. T., Annaswamy, A. M., and Lavretsky, E., "Adaptive Control and the NASA X-15-3 Flight Revisited," *IEEE Control Systems Magazine*, Vol. 30, No. 3, 2010, pp. 32–48. https://doi.org/10.1109/MCS.2010.936292

[3] Annaswamy, A. M., and Fradkov, A. L., "A Historical Perspective of Adaptive Control and Learning," *Annual Reviews in Control*, Vol. 52, Jan. 2021, pp. 18–41. https://doi.org/10.1016/j.arcontrol.2021.10.014

[4] Enns, R., and Si, J., "Neuro-Dynamic Programming Applied to Helicopter Flight Control," *AIAA Guidance, Navigation, and Control Conference and Exhibit*, AIAA Paper 2000-4280, 2000. https://doi.org/10.2514/6.2000-4280

[5] Ferrari, S., and Stengel, R. F., "Online Adaptive Critic Flight Control," *Journal of Guidance, Control, and Dynamics*, Vol. 27, No. 5, 2004, pp. 777–786. https://doi.org/10.2514/1.12597

[6] Helder, B., van Kampen, E.-J., and Pavel, M., "Online Adaptive Helicopter Control Using Incremental Dual Heuristic Programming," *AIAA Scitech 2021 Forum*, AIAA Paper 2021-1118, 2021. https://doi.org/10.2514/6.2021-1118

[7] Heyer, S., Kroezen, D., and Van Kampen, E., "Online Adaptive Incremental Reinforcement Learning Flight Control for a CS-25 Class Aircraft," *AIAA Scitech 2020 Forum*, AIAA Paper 2020-1844, 2020. https://doi.org/10.2514/6.2020-1844

[8] Jacklin, S. A., "Closing the Certification Gaps in Adaptive Flight Control Software," *AIAA Guidance, Navigation and Control Conference and Exhibit*, AIAA Paper 2008-6988, 2008. https://doi.org/10.2514/6.2008-6988

[9] Zhou, Y., Van Kampen, E.-J., and Chu, Q., "Incremental Model-Based Online Dual Heuristic Programming for Nonlinear Adaptive Control," *Control Engineering Practice*, Vol. 73, April 2018, pp. 13–25. https://doi.org/10.1016/j.conengprac.2017.12.011

[10] Recht, B., "A Tour of Reinforcement Learning: The View from Continuous Control," *Annual Review of Control, Robotics, and Autonomous Systems*, Vol. 2, No. 1, 2019, pp. 253–279. https://doi.org/10.1146/annurev-control-053018-023825

[11] Dean, S., Mania, H., Matni, N., Recht, B., and Tu, S., "On the Sample Complexity of the Linear Quadratic Regulator," *Foundations of Computational Mathematics*, Vol. 20, No. 4, 2019, pp. 633–679. https://doi.org/10.1007/s10208-019-09426-y

[12] Konatala, R., Kampen, E.-J. V., and Looye, G., "Reinforcement Learning Based Online Adaptive Flight Control for the Cessna Citation II (PH-LAB) Aircraft," *AIAA Scitech 2021 Forum*, AIAA Paper 2021-0883, 2021. https://doi.org/10.2514/6.2021-0883

[13] Zhou, Y., Van Kampen, E., and Chu, Q., "Nonlinear Adaptive Flight Control Using Incremental Approximate Dynamic Programming and Output Feedback," *Journal of Guidance Control and Dynamics*, Vol. 40, No. 2, 2017, pp. 493–496. https://doi.org/10.2514/1.G001762

[14] Zhou, Y., van Kampen, E.-J., and Chu, Q., "Incremental Approximate Dynamic Programming for Nonlinear Adaptive Tracking Control with Partial Observability," *Journal of Guidance, Control, and Dynamics*, Vol. 41, No. 12, 2018, pp. 2554–2567. https://doi.org/10.2514/1.G003472

[15] Lewis, F. L., and Vrabie, D., "Reinforcement Learning and Adaptive Dynamic Programming for Feedback Control," *IEEE Circuits and Systems Magazine*, Vol. 9, No. 3, 2009, pp. 32–50. https://doi.org/10.1109/MCAS.2009.933854

[16] Bradtke, S. J., and Ydstie, B. E., "Adaptive Linear Quadratic Control Using Policy Iteration," *Proceedings of 1994 American Control Conference-ACC'94*, Vol. 3, Inst. of Electrical and Electronics Engineers, New York, 1994, pp. 3475–3479. https://doi.org/10.1109/ACC.1994.735224

[17] Bellman, R., "Dynamic Programming," *Science*, Vol. 153, No. 3731, 1957, pp. 34–37. https://doi.org/10.1126/science.153.3731.34

[18] Dias, P. M., Zhou, Y., and Kampen, E.-J. V., "Intelligent Nonlinear Adaptive Flight Control Using Incremental Approximate Dynamic Programming," *AIAA Scitech 2019 Forum*, AIAA Paper 2019-2339, 2019. https://doi.org/10.2514/6.2019-2339

[19] Konatala, R., Kampen, E.-J. V., Looye, G., Milz, D., and Weiser, C., "Flight Testing Reinforcement Learning Based Online Adaptive Flight Control Laws on CS-25 Class Aircraft," *AIAA Scitech 2024 Forum*, AIAA Paper 2024-2402, 2024. https://doi.org/10.2514/6.2024-2402

[20] Konatala, R., Müller, R., May, M., Looye, G., and van Kampen, E.-J., "Verification & Validation (V&V) of Reinforcement Learning Based Online Adaptive Flight Control Laws on CS-25 Class Aircraft," *Proceedings of the 2024 CEAS EuroGNC Conference*, Council of European Aerospace Societies, Bristol, U.K., Paper CEAS-GNC-2024-062, 2024.

[21] "Flight Control Design—Best Practices," Defense Technical Information Center, Research, and Technology Organization, NATO, 2000, https://apps.dtic.mil/sti/tr/pdf/ADA387777.pdf.

[22] Narendra, K. S., and Annaswamy, A. M., "Persistent Excitation in Adaptive Systems," *International Journal of Control*, Vol. 45, No. 1, 1987, pp. 127–160. https://doi.org/10.1080/00207178708933715

[23] Joos, H.-D., "A Methodology for Multi-Objective Design Assessment and Flight Control Synthesis Tuning," *Aerospace Science and Technology*, Vol. 3, No. 3, 1999, pp. 161–176. https://doi.org/10.1016/S1270-9638(99)80040-6

K. A. Wise
Associate Editor